Redbooks
ibm.com/redbooks

IBM Power Security Catalog

Tim Simon

Felipe Bessa

Hugo Blanco

Carlo Castillo

Rohit Chauhan

Kevin Gee

Gayathri Gopalakrishnan

Samvedna Jha

Andrey Klyachkin

Andrea Longo

Ahmed Mashhour

Amela Peku

Prashant Sharma

Vivek Shukla

Dhanu Vasandani

Henry Vo

Power Systems

Security

IBM

Redbooks

IBM Redbooks

IBM Power Security Catalog

August 2024

SG24-8568-00

Note: Before using this information and the product it supports, read the information in "Notices" on page xi.

First Edition (August 2024)

This edition applies to:
Hardware Management Console V10R1 or later
PowerVM Virtual I/O Server 3.1.3 and Virtual I/O Server 3.1.4
PowerVM Virtual I/O Server 4.1.0.10 or later
AIX 7.2 TL4 or later, AIX 7.3 TL0 or later
IBM i 7.3 TR13, IBM i 7.4 TR7, IBM i 7.5 or later
Red Hat Enterprise Linux Server 8.4 (ppc64le)
Red Hat Enterprise Linux Server 9.2 (ppc64le)
SLES 15 SP 5 (ppc64le)
Ubuntu 22.04.3 (ppc64le) or later
Red Hat OpenShift Version 4.14 or later
PowerSC 2.0 or later

Contents

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at https://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

AIX®
DB2®
DS8000®
FlashCopy®
GDPS®
IBM®
IBM Cloud®
IBM FlashSystem®
IBM Instana™
IBM Security®
IBM Z®
IBM z Systems®
Instana®
POWER®
Power9®
PowerHA®
PowerPC®
PowerVM®
QRadar®
Redbooks®
Redbooks (logo) ®
System z®
SystemMirror®
Tivoli®
WebSphere®
X-Force®
z Systems®
z/OS®

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

ITIL is a Registered Trade Mark of AXELOS Limited.

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Ansible, Ceph, Fedora, JBoss, OpenShift, Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

VMware, and the VMware logo are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

Other company, product, or service names may be trademarks or service marks of others.

Preface

IT security is paramount in today's digital age. As businesses increasingly rely on technology to operate, protecting sensitive data and preventing cyberattacks becomes a top priority. Cloud adoption introduces additional security risks, including data breaches and loss of access. A strong IT security infrastructure safeguards customer information, financial data, intellectual property, and overall business operations. By investing in robust security measures, organizations can mitigate risks, maintain trust with customers, and ensure business continuity.

A multi-layered security architecture is essential for protection. Key areas to focus on include:

- Hardware-Level Security: Prevent physical tampering and ensure data integrity.
- Virtualization Security: Isolate environments and control resource access.
- Management Tool Security: Secure hardware and cloud resources.
- Operating System Security: Continuously update for robust security.
- Storage Security: Protect data at rest and in transit.
- Networking Security: Prevent unauthorized access and data breaches.

This Redbook describes how the IBM Power ecosystem provides advanced security capabilities at each of these layers. IBM Power systems are designed with security as a core consideration.

At the hardware level, advanced technology includes tamper-resistant features built into the processor to prevent unauthorized access and modifications, secure cryptographic engines to provide strong encryption of data, and Trusted Boot to ensure that only authorized software components are loaded during system startup.

At the virtualization level, the hypervisor – which manages virtual machines – is designed to be secure and resistant to attacks. The hypervisor isolates workloads within a single physical server, allowing for secure resource sharing within your infrastructure. The Hardware Management Console (HMC) provides centralized management and control of Power systems in a secure manner.

The operating systems that run on IBM Power servers – AIX, IBM i, and Linux on Power – offer robust security features, including user authentication, access controls, and encryption support. In addition, tools such as IBM PowerSC provide a comprehensive security and compliance solution that helps manage security policies, monitor threats, and enforce compliance.

Security also requires solid management and control. This book describes best practices such as conducting regular security audits, keeping operating systems and applications up-to-date with the latest security patches, and implementing strong user authentication and authorization policies. Other critical elements include the implementation of data encryption for both data at rest and in flight, and strong network security processes utilizing firewalls, intrusion detection systems, and other security measures.

By combining these hardware, software, and management practices, IBM Power systems provide a robust foundation for security in your IT environment.

Authors

This book was produced by a team of specialists from around the world working at IBM Redbooks, Center.

Tim Simon is an IBM® Redbooks® Project Leader in Tulsa, Oklahoma, USA. He has over 40 years of experience with IBM, primarily in a technical sales role working with customers to help them create IBM solutions to solve their business problems. He holds a BS degree in Math from Towson University in Maryland. He has extensive experience creating customer solutions using IBM Power, IBM Storage, and IBM System z® throughout his career.

Felipe Bessa is an IBM Brand Technical Specialist and Partner Technical Advocate on IBM Power. He works for IBM Technology in Brazil and has over 25 years of experience in the areas of research, planning, implementation, and administration of IT infrastructure solutions. Before joining IBM, he was recognized as a Reference Client for IBM Power Technologies for SAP and SAP HANA, IBM PowerVC, IBM PowerSC, Monitoring and Security, IBM Storage, and the Run SAP Like a Factory (SAP Solution Manager) Methodology. He was chosen as an IBM Champion for IBM Power for 2018 - 2021.

Hugo Blanco is an IBM Champion based in Madrid, has been working with Power systems since 2008. He began his career as an instructor and has since taken on a variety of roles at SIXE, IBM BP, gaining extensive experience across different roles and functions. Hugo is deeply passionate about AIX, Linux on Power, and various cybersecurity solutions. He has contributed to the development of several IBM certification exams and actively participates in Common Iberia, Common Europe, and TechXchange. He enjoys delivering technical talks on emerging technologies and real-world use cases. Beyond his technical pursuits, he is also a dancer, DJ, and event producer.

Carlo Castillo is a Client Services Manager for Right Computer Systems (RCS), an IBM Business Partner and Red Hat partner in the Philippines. He has over thirty years of experience in pre-sales and post-sales support, designing full IBM infrastructure solutions, creating pre-sales configurations, performing IBM Power installation, implementation and integration services, and providing post-sales services and technical support for customers, as well as conducting presentations at customer engagements and corporate events. He was the very first IBM-certified IBM AIX Technical Support engineer in the Philippines in 1999. As training coordinator during RCS' tenure as an IBM Authorized Training Provider from 2007 to 2014, he also administered the IBM Power Systems curriculum, and conducted IBM training classes covering AIX, PureSystems, PowerVM, and IBM i. He holds a degree in Computer Data Processing Management from the Polytechnic University of the Philippines.

Rohit Chauhan is a Senior Technical Specialist with expertise in IBM i architecture, working at Tietoevry Tech Services, Stavanger, Norway, an IBM Business Partner and also one of the biggest IT service provider in Nordics. He has over 12 years of experience working on the IBM Power platform with design, planning and implementation of IBM i infrastructure including High availability/Disaster recovery solutions for many customers during this tenure. Before his current role, Rohit worked for clients in Singapore and U.A.E in the technical leadership and security role on IBM Power domain. He possesses rich corporate experience in architecting solution design, implementations and system administration. He is also a member of Common Europe Norway with strong focus on IBM i platform and security. He is also recognized as IBM Advocate, Influencer and Contributor for 2024 through IBM Rising Champions Advocacy Badge program. He holds a bachelor degree in Information Technology. He is a IBM certified technical expert and also holds ITIL CDS certificate. His areas of expertise includes overall IBM i, IBM Hardware Management Console (HMC), Security enhancements, IBM PowerHA®, Systems Performance analysis and tuning, BRMS, External storage, Power VM and providing solutions to the customers on IBM i platform.

Kevin Gee has over 30 years of IT experience, mostly prominently with IBM technology solutions in support, systems engineering, consulting, and technical sales roles. He has broad experience in server and storage infrastructure and performance, high availability and disaster recovery, enterprise backup and recovery, and product development and go-to-market strategy. Kevin has authored hundreds of white papers, training guides, and presentations for clients, co-authored IBM Redbooks, and helped develop certification exams for the AIX, PowerHA for AIX, and PowerVC products. He frequently presents at IBM technical conferences and contributes content to others' presentations. He has also published research into machine learning, graph processing, and natural language processing. Kevin holds a master's degree in Computer Science from The University of Texas at Arlington and bachelor's degrees in Computer Science and Spanish from Brigham Young University. He has been an IBM Champion since 2019.

Gayathri Gopalakrishnan works for IBM India and has over 22 years of experience as a technical solution and IT architect, working primarily in consulting. She is a results-driven IT Architect with extensive working experience in spearheading the management, design, development, implementation, and testing of solutions. A recognized leader, applying high-impact technical solutions to major business objectives with capabilities transcending boundaries. She is adept at working with management to prioritize activities and achieve defined project objectives with an ability to translate business requirements into technical solutions.

Samvedna Jha is a Senior Technical Staff Member in the IBM Power Systems organization, Bengaluru, India. She holds a masters degree in Computer Application and has more than twenty years of work experience. In her current role as Security Architect, IBM Power, she has worldwide technical responsibility to handle security and compliance requirements of Power products. Samvedna is a recognized speaker in conferences, has authored blogs and published disclosures. She is also the security focal point for Power products secure release process.

Andrey Klyachkin is a solution architect at eNFence, an IBM business partner in Germany. He has more than 25 years experience in UNIX systems, designing and supporting AIX® and Linux systems for different customers all over the world. He is co-author of many IBM AIX and IBM Power certifications, IBM Champion and IBM AIX Community Advocate. He is also a Red Hat Certified Engineer and Red Hat Certified Instructor. Andrey blogs a lot about IBM AIX automation using Ansible on LinkedIn where you can always find and ask him questions. Another way to ask him a question is to meet at international and local events about IBM Power like IBM TechXchange, GSE and Common Europe Congress.

Andrea Longo is a Partner Technical Specialist for IBM Power in Amsterdam, the Netherlands. He has a background in computational biology research and holds a degree in Science and Business Management from Utrecht University. He is also serving IBM in the role of Quantum Ambassador to prepare academia and industry leaders to be quantum-safe and to experiment the immense possibilities of the technology.

Ahmed Mashhour is a Power Technology Services Consultant Lead at IBM Saudi Arabia. He is an IBM L2 certified Expert. He holds IBM AIX, Linux, and IBM Tivoli® certifications. He has 19 years of professional experience in IBM AIX and Linux systems. He is an IBM AIX back-end SME who supports several customers in the US, Europe, and the Middle East. His core expertise is in IBM AIX, Linux systems, clustering management, IBM AIX security, virtualization tools, and various IBM Tivoli and database products. He authored several publications inside and outside IBM, including co-authoring other IBM Redbooks publications. He also hosted IBM AIX, Security, PowerVM®, IBM PowerHA, PowerVC, PowerVS and IBM Storage Scale classes worldwide.

Amela Peku is a Partner Technical Specialist with broad experience in leading technology companies. She holds an MS in Telecommunication Engineering and is part of the IBM Power team, working with Business Partners and customers to showcase the value of IBM Power solutions. Previously, she provided technical support for Next Generation Firewalls, Webex, Webex Teams focusing on performance and networking, and handled escalations, working closely with engineering teams. She is certified in Networking, Security, and IT Management.

Prashant Sharma is the IBM Power Technical Product Leader for the Asia Pacific region, based in Singapore. He holds a degree in Information Technology from the University of Teesside, England, and an MBA from the University of Western Australia. With extensive experience in IT infrastructure enterprise solutions, he specializes in pre-sales activities, client and partner consultations, technical enablement, and the implementation of IBM Power servers, IBM i, and IBM Storage. He drives technical strategy and product leadership for IBM Power Systems, ensuring the delivery of innovative solutions to diverse markets.

Vivek Shukla is a Technical Sales Specialist for IBM Power, Hybrid Cloud, AI, and Cognitive Offerings in Qatar working for GBM. He has rich experience in Sales, Application Modernization, Digital Transformation, Infrastructure Sizing, Cyber Security & consulting, SAP HANA/Oracle/Core Banking. He is an IBM Certified L2 (Expert) Brand Technical Specialist. He has over 22 years of IT experience in Technical Sales, Infrastructure Consulting, IBM Power servers, AIX, IBM i and IBM Storage implementations. He has hands-on experience on IBM Power servers, AIX, PowerVM, PowerHA, PowerSC, Requests for Proposals, Statements of Work, sizing, performance tuning, root cause analysis, DR, and mitigation planning. In addition to writing multiple IBM Power FAQs, he is also a Redbook Author. He is a presenter, mentor, and profession champion accredited by IBM. He graduated with a bachelor's degree (BTech) in electronics and telecommunication engineering from IETE, New Delhi, and a master's degree (MBA) in information technology from IASE University. Red Hat OpenShift, IBM Cloud® Paks, Power Enterprise Pools, and Hybrid Cloud are among his areas of expertise.

Dhanu Vasandani is a Staff Software Test Engineer with over 13 years of experience, specializing in AIX Operating System Security Testing at IBM Power Systems in Bangalore, India. She holds a Bachelor of Technology degree in Computer Science and has been instrumental in testing multiple AIX releases across various Power Server Models. In her current role, Dhanu serves as the Component Lead for the AIX Operating System Security Guild, overseeing various sub-components. She is responsible for conducting comprehensive system testing for Pre-GA and Post-GA phases of multiple AIX releases across different Power Server Models. Dhanu is known for her expertise in areas such as Encryption, Trustchk, Audit, RBAC, and other security aspects, contributing significantly to IBM's Lighthouse Community and Knowledge Centre. She is recognized for her proficiency in identifying and addressing high-impact AIX defects within the ISST System organization, ensuring the delivery of top-quality products to customers.

Henry Vo is an IBM Redbooks Project Leader with 10 years experience in IBM. He has technical expertise in business problem solving, risk/root-cause analysis, and writing technical plans for business. He has held multiple roles at IBM including Project management, ST/FT/ETE Test, Back End Developer, DOL agent for NY. He is a certified IBM zOS Mainframe Practitioner including IBM Z® System programming, Agile, and Telecommunication Development Jumpstart. Henry holds a Master of MIS (Management Information System) from the University of Texas in Dallas.

Thanks to the following people for their contributions to this project:

Stephen Dominguez, WW Lead Consultant for AIX and Linux Security
IBM Technology Expert Labs, Austin, TX

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an IBM Redbooks residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, IBM Redbooks
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on LinkedIn:

  https://www.linkedin.com/groups/2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/subscribe

- Stay current on recent Redbooks publications with RSS Feeds:

  https://www.redbooks.ibm.com/rss.html

1

Security and cyber security challenges

This chapter delivers a thorough examination of the security challenges confronting organizations in today's intricate digital environment. It begins by outlining the diverse range of security issues, highlighting the growing sophistication of threats and the imperative need for strong security measures to protect sensitive data and infrastructure. As IT operations extend beyond traditional data centers into hybrid and multi-cloud settings, new vulnerabilities and challenges arise, necessitating adjustments to security strategies.

The chapter delves into the concept of cyber resilience, focusing on the zero trust security model, which mandates continuous verification of users, devices, and network components in an environment with both internal and external threats. It also provides an in-depth look at IBM's security approach, showcasing how its advanced technologies and methodologies are designed to defend against various threats.

In summary, this chapter offers a comprehensive analysis of the security and cybersecurity challenges faced by organizations, presenting detailed insights into strategies and technologies to mitigate these threats and enhance overall security resilience, with a particular emphasis on IBM's response to these challenges.

This chapter focuses in the following topics:

1.1 Protection at every layer

Cloud adoption, edge computing, and hybrid infrastructures, while offering significant benefits in terms of flexibility and scalability introduce security concerns. These include data breaches, loss of control over data, and the complexities of managing security across diverse cloud platforms. Cyber attacks and ransomware attacks are becoming more prevalent and more sophisticated which can have devastating impacts on organizations. A multi-layered security architecture is crucial for robust protection, encompassing several implementation layers.

At the hardware level, built-in protections that prevent physical tampering and ensure data integrity are needed. Virtualization technologies need to enhance security by isolating environments and controlling resource access. The security of the hypervisor, a critical component in virtualized environments, is paramount in preventing attacks that could compromise multiple virtual machines. In the IBM Power environment, logical partitioning (LPAR) provides strong isolation between different workloads on the same physical hardware, enhancing security. Figure 1-1 shows how IBM Power10 works to provide protection at every layer.

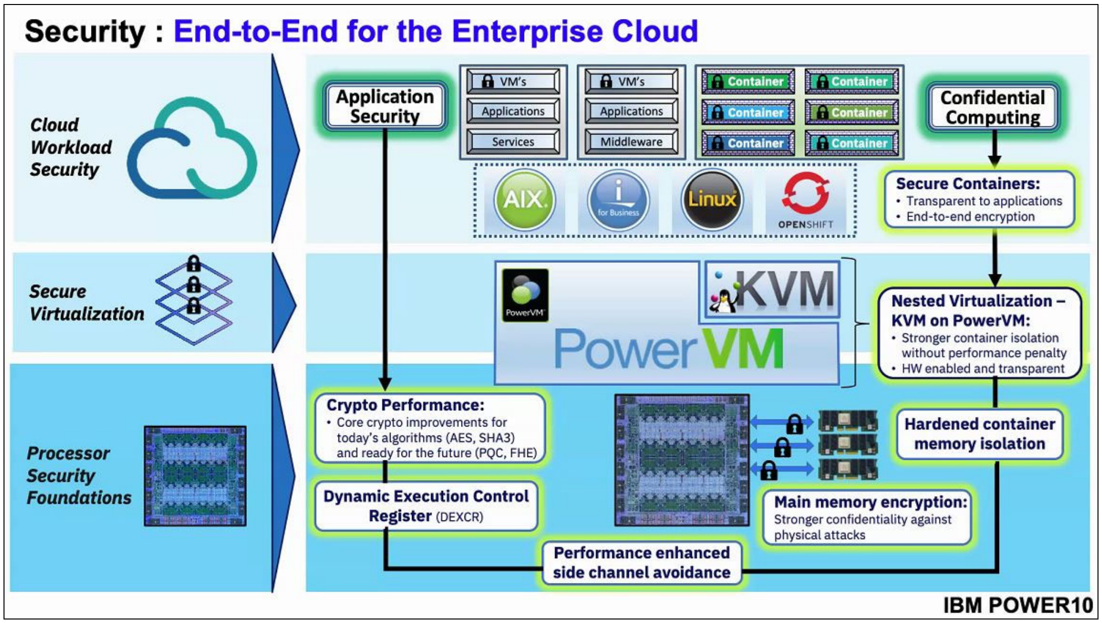

Security : End-to-End for the Enterprise Cloud
Cloud Workload Security
Secure Virtualization
Processor Security Foundations
Application Security
VM's
Applications
Services
VM's
Applications
Middleware
Container
Container
Container
Container
Container
Container
Confidential Computing
AIX
for Business
Linux
OPENSHIFT
Secure Containers:
• Transparent to applications
• End-to-end encryption
PowerVM
KVM
PowerVM
Nested Virtualization – KVM on PowerVM:
• Stronger container isolation without performance penalty
• HW enabled and transparent
Crypto Performance:
• Core crypto improvements for today's algorithms (AES, SHA3) and ready for the future (PQC, FHE)
Dynamic Execution Control Register (DEXCR)
Hardened container memory isolation
Main memory encryption: Stronger confidentiality against physical attacks
Performance enhanced side channel avoidance
IBM POWER10

Figure 1-1 Protection at every layer[1]

Management tools like the Hardware Management Console (HMC) and Cloud Management Console (CMC) play a vital role in securing hardware and cloud resources. Operating systems need to continuously provide better security features as they are often a vector of attack and their contribution is critical to the overall security posture of a system.

Storage security involves protecting data at rest and in transit through techniques such as encryption and access controls. Methods for creating secure, resilient copies of data – known as safeguarded copy – and data resiliency are needed to protect against data corruption or loss. Finally, networking security is integral to overall security with a focus on secure network design, monitoring, and protection mechanisms to prevent unauthorized access and data breaches.

[1] https://hc32.hotchips.org/assets/program/conference/day1/HotChips2020_Server_Processors_IBM_Starke_POWER10_v33.pdf

1.2 IBM Systems: Built to Protect

In today's digital landscape, IBM Infrastructure serves as a formidable shield against increasingly sophisticated cyber threats through its robust and integrated security solutions. IBM weaves security into the fabric of its systems and platforms, allowing businesses to operate confidently amid evolving risks.

At the heart of IBM's approach is the integration of security throughout its systems, building trust and resilience from the ground up. This includes safeguarding firmware integrity with secure boot processes and bolstering data protection through hardware-based encryption acceleration.

IBM goes beyond basic protection with a proactive cybersecurity strategy. It offers secure storage solutions and advanced threat prevention and detection mechanisms. In the event of an incident, IBM provides rapid response and recovery options to minimize downtime and effectively manage operational risks.

Privacy and confidentiality are paramount, supported by IBM's advanced encryption technologies. These include pervasive encryption throughout the data lifecycle and quantum-safe cryptography, designed to guard against emerging threats such as quantum computing.

IBM simplifies regulatory compliance with continuous compliance and audit capabilities. Automated monitoring and enforcement tools ensure adherence to industry standards, while unified security management tools facilitate consistent governance across diverse IT environments.

Collaborating closely with ecosystem partners, IBM integrates security across hybrid cloud environments, networks, software systems, architectures, and chip designs. This comprehensive approach ensures holistic protection and resilience across all facets of IT infrastructure.

By consolidating security insights across various domains, IBM enables informed decision-making and proactive threat management. This integrated approach dissolves traditional security silos, turning security into a catalyst for innovation and business growth.

In summary, IBM Infrastructure sets a high standard for security excellence by embedding advanced features into its solutions and equipping businesses to address both current and future cybersecurity challenges with confidence. Through collaborative efforts with ecosystem partners and a focus on regulatory compliance, IBM delivers secure, resilient, and compliant infrastructure solutions, empowering businesses to thrive in the digital age amidst evolving cyber threats.

1.3 Overview of Security Challenges

Digital transformation is reshaping the landscape of modern business, driving the creation and modification of products, services, and operations through digital technology. This integration touches every area of business, fundamentally altering operations and customer value delivery.

The necessity for digital transformation spans businesses of all sizes, from small enterprises to large corporations. This message is conveyed clearly through virtually every keynote, panel discussion, article, or study related to how businesses can remain competitive and relevant as the world becomes increasingly digital. However, there are many considerations, with security being one of the most important. Ensuring that the outcome of digital transformation is more secure than before, and that the transition process is handled securely, is crucial.

In the era of digital transformation, many organizations have reported experiencing at least one data breach due to the digital transformation process. Besides data breaches, there are other concerns organizations need to address, such as ensuring a secure expansion beyond their data centers, secure cloud adoption, and mitigating cyberattacks and ransomware.

1.3.1 Expansion Beyond the Data Center

Expanding operations beyond the traditional data center into cloud environments, edge computing, and hybrid infrastructures introduces several security challenges. These challenges arise from increased complexity, diverse environments, and an evolving threat landscape. To navigate these complexities, organizations must address key areas such as data protection, access control, visibility, compliance, threat management, configuration, and network security.

Data protection and privacy are paramount as data flows between data centers, cloud services, and edge devices. Ensuring data is encrypted in transit and at rest across various platforms is essential. Proper encryption and access controls safeguard stored data, while compliance with data sovereignty regulations ensures data is processed and stored according to regional laws. Businesses need to implement end-to-end encryption, enforce access controls, and stay updated on data protection regulations.

Access control and identity management become more complex in hybrid environments. Consistent identity and access management (IAM) across on-premises and cloud environments is crucial. Managing and monitoring privileged access helps prevent unauthorized access and insider threats. Strong authentication methods, such as multi-factor authentication (MFA), enhance security by adding additional layers of protection. Implementing robust IAM solutions, continuously monitoring access, and ensuring the use of MFA are key steps.

Visibility and monitoring across hybrid and multi-cloud environments are critical for detecting anomalies and threats. Achieving comprehensive visibility involves implementing unified monitoring solutions that provide a holistic view of the entire infrastructure. Consistent logging and auditing mechanisms are necessary to track activities and support incident response. Network monitoring helps detect and respond to threats in real time. Organizations should invest in integrated monitoring tools, establish thorough logging practices, and deploy real-time network monitoring systems.

Compliance and regulatory requirements present another set of challenges. Navigating diverse regulations across different regions and sectors requires a deep understanding of relevant laws. Ensuring systems and processes are audit-ready demonstrates compliance, while maintaining data classification schemes ensures appropriate protection measures are in place. Businesses need to stay informed about regulatory changes, conduct regular audits, and implement robust data classification protocols.

Managing advanced threats is an ongoing battle. Defending against sophisticated threats such as advanced persistent threats, zero-day vulnerabilities, and ransomware requires a proactive approach. Securing endpoints, including edge devices with varying security capabilities, is crucial. Keeping systems up to date with security patches across different environments helps mitigate vulnerabilities. Organizations should adopt advanced threat detection solutions, enforce stringent endpoint security measures, and establish effective patch management processes.

Configuration management ensures consistent security configurations across diverse environments. Detecting and correcting misconfiguration that can introduce vulnerabilities is vital. Organizations should use automated tools to manage configurations and regularly audit systems to identify and rectify misconfiguration.

Network security involves implementing network segmentation to limit the spread of threats. Ensuring secure connections between data centers, cloud environments, and edge locations is essential. Deploying and managing firewalls and intrusion detection/prevention systems in a coordinated manner strengthens network security. Businesses should design segmented network architectures, secure connectivity channels, and maintain robust firewalls as well as Intrusion Detection Systems (IDS) and Intrusion Protection Systems (IPS).

Expanding operations beyond the traditional data center offers numerous benefits, but it also introduces a range of security challenges that organizations must proactively address. By prioritizing comprehensive data protection, robust access control, enhanced visibility, regulatory compliance, advanced threat management, consistent configuration, and strong network security, businesses can mitigate these risks and fully leverage the advantages of hybrid and multi-cloud environments. Security in these complex infrastructures is an ongoing process that requires vigilance, adaptability, and a commitment to staying ahead of emerging threats.

1.3.2 Cloud adoption

The migration to cloud computing has revolutionized how businesses operate, offering unparalleled scalability, cost-efficiency, and flexibility. However, alongside these advantages come significant security challenges that organizations must address to ensure their data and operations remain secure.

One of the most critical security challenges in cloud adoption is the potential for data breaches and data loss. Sensitive information stored in the cloud can be an attractive target for cyber criminals. Unauthorized access can lead to the exposure of confidential data, resulting in financial losses, reputational damage, and legal repercussions. To mitigate these risks, businesses should implement end-to-end encryption for data at rest and in transit, enforce strict access control policies, and conduct regular security audits and vulnerability assessments.

Compliance and regulatory issues add another layer of complexity to cloud security. Cloud environments often span multiple jurisdictions, each with its own set of regulations and compliance requirements. Ensuring that cloud operations comply with laws such as GDPR, HIPAA, and CCPA can be complex and resource-intensive. Organizations must stay informed about relevant regulations, utilize compliance management tools, and engage third-party auditors to verify compliance.

Insider threats, whether from malicious intent or inadvertent actions, pose a significant risk. Employees, contractors, or third-party vendors with access to cloud systems can potentially misuse their access, leading to data leaks or disruptions. To counter these threats, businesses should implement regular security training programs, utilize monitoring and anomaly detection systems, and apply the principle of least privilege to limit access based on necessity.

The shared responsibility model in cloud security, where both the cloud provider and the customer share security responsibilities, can lead to confusion and security gaps. Clear definitions of security responsibilities in contracts, regular reviews of cloud provider security documentation, and ongoing collaboration between IT teams and cloud providers are essential to avoid misunderstandings and ensure comprehensive security coverage.

Application Programming Interfaces (APIs) are essential for cloud integration and operations but can also introduce vulnerabilities. Poorly secured APIs can become entry points for attackers. To secure APIs, organizations should adopt secure coding practices, use API gateways to manage and secure API traffic, and implement rate limiting to prevent abuse.

Cloud accounts are vulnerable to hijacking through phishing, credential stuffing, or other attack methods. Once compromised, attackers can gain control over cloud resources and data. Enforcing multi-factor authentication (MFA), implementing strong password policies, and continuously monitoring account activities for suspicious behavior are crucial steps to protect cloud accounts from hijacking.

Interfaces and APIs, as gateways to cloud services, need robust security measures. If not properly secured, they can be exploited to gain unauthorized access or disrupt services. Following best practices in API design and security, conducting regular penetration testing and vulnerability assessments, and implementing strong authentication and authorization measures are necessary to secure these critical components.

Adopting cloud technology offers numerous benefits but also introduces a range of security challenges that organizations must proactively address. By implementing robust security measures, maintaining regulatory compliance, and fostering a culture of security awareness, businesses can mitigate these risks and fully leverage the advantages of the cloud. Security in the cloud is an ongoing process that requires vigilance, adaptability, and a commitment to staying ahead of emerging threats.

1.3.3 Cyber Attacks and Ransomware

The digital transformation era has enhanced business efficiency and connectivity, but it has also introduced sophisticated cyber attacks and ransomware threats. Addressing these challenges requires a comprehensive and proactive approach to protect sensitive data and ensure business continuity.

Cyber attacks, such as phishing, malware, and advanced persistent threats, disrupt operations and compromise data. To combat these threats, organizations must implement robust firewalls, intrusion detection and prevention systems, and leverage threat intelligence.

Ransomware, which encrypts data and demands a ransom for its release, has become particularly destructive. Preventing ransomware involves regular software updates, strong email filtering, and anti-phishing solutions. Regular data backups stored securely and offline, along with tested recovery processes, are critical for mitigating ransomware impact. IBM Storage Defender is a storage software solution that can help protect your data and accelerate recovery in the event of a cyber attack or other catastrophic events. It includes immutable backups, early threat detection, data copy management, and automated recovery capabilities.

Employee training and awareness are essential, as many attacks exploit human vulnerabilities. Regular training can help employees recognize phishing attempts and follow best practices for data security, acting as a front line defense against cyber threats.

Endpoint security is crucial as employees access corporate resources from various devices. Advanced endpoint protection solutions, including anti-virus software, endpoint detection and response (EDR) tools, and mobile device management (MDM) systems safeguard endpoints from malicious activity.

Network segmentation limits the spread of ransomware and other threats. Dividing the network into smaller segments and implementing strong access controls and monitoring can contain damage and prevent lateral movement by attackers.

Incident response planning is vital for minimizing the impact of attacks. An up-to-date incident response plan with clear communication protocols, roles, and procedures for isolating affected systems and restoring operations is essential. Regular drills ensure the readiness of the response team.

Cyber insurance provides additional protection, covering the costs of recovery, legal fees, data restoration, and customer notification in case of an attack.

In conclusion, addressing cyber attacks and ransomware requires a comprehensive strategy, including multi-layered defenses, regular backups, employee training, endpoint security, network segmentation, incident response planning, and cyber insurance. By staying vigilant and continually enhancing security measures, organizations can mitigate risks and protect against these threats.

Cyber Resilience - Zero Trust

In today's digital landscape, organizations face increasingly sophisticated cyber threats that require robust defense strategies. Cyber resilience and the zero trust model have emerged as critical frameworks to strengthen these defenses and ensure business continuity in the face of evolving risks.

Cyber resilience encompasses strategies to prepare for, respond to, and recover from cyber incidents effectively. This includes comprehensive risk management to prioritize critical assets and threats, incident response planning with clear protocols, regular data backups, and continuous improvement through assessments and updates.

The zero trust model challenges traditional security approaches by assuming no implicit trust based on network location. Instead, it verifies and validates all devices, users, and applications attempting to connect, regardless of their location. Key principles include explicit verification, least privilege access, micro-segmentation to limit lateral movement, and continuous monitoring of network traffic and user behavior.

By integrating cyber resilience with zero trust principles, organizations enhance their ability to detect, respond to, and mitigate cyber threats. Continuous monitoring and analysis of network activity and user behavior enable prompt threat detection and response. Dynamic, risk-based access controls based on real-time assessments improve security without hindering productivity. Robust backup and recovery measures combined with strict access controls ensure data integrity and availability, even in the event of a breach.

In conclusion, cyber resilience and the zero trust model are essential for organizations striving to fortify their security posture amidst a complex threat landscape. By adopting proactive strategies and integrating zero trust principles into their security framework, businesses can safeguard critical assets, maintain operational continuity, and mitigate the impact of cyber attacks. These frameworks not only strengthen defenses but also foster a culture of security awareness and readiness across the organization, ensuring ongoing protection against evolving cyber threats.

1.3.4 Government regulations

Government regulations play a crucial role in shaping and enforcing security standards across various sectors. These regulations aim to protect sensitive information, ensure data privacy, and maintain overall cybersecurity. In this section we discuss some key aspects of government regulations in relation to security and provide some examples.

Data Protection and Privacy Laws

Different jurisdictions have specific data protection and privacy laws that regulate how data is secured and how consumers can protect their privacy. A couple of examples are:

- General Data Protection Regulation (GDPR)

  Enforced by the European Union, GDPR mandates stringent data protection and privacy measures for organizations handling EU residents' data. It includes requirements for data breach notifications, consent management, and data subject rights.

- California Consumer Privacy Act (CCPA)

  This law, from the United States state of California, provides residents with rights over their personal data, including access, deletion, and opt-out options for data sales.

Cybersecurity Frameworks

There are groups of regulations that address cybersecurity requirements, such as:

- NIST Cybersecurity Framework

  Developed by the National Institute of Standards and Technology (NIST), this framework provides guidelines for improving cybersecurity practices. While not mandatory, many organizations adopt it to align with best practices and regulatory expectations.

- Federal Information Security Management Act (FISMA)

  In the United States, FISMA requires federal agencies and contractors to implement information security programs and comply with NIST standards.

Industry-Specific Regulations

There are specific industry requirements that address data management and security, such as:

- Health Insurance Portability and Accountability Act (HIPAA)

  For the healthcare industry in the U.S., HIPAA sets standards for protecting patient health information and requires secure handling and storage of sensitive data.

- Payment Card Industry Data Security Standard (PCI DSS)

  This set of standards applies to organizations handling credit card information and mandates secure processing, storage, and transmission of payment data.

Data Breach Notification Laws

Many jurisdictions have laws requiring organizations to notify affected individuals and relevant authorities in the event of a data breach. These laws often specify time lines for notification and the types of information that must be disclosed.

Critical Infrastructure Protection

In the United States, the Cybersecurity and Infrastructure Security Agency (CISA) works to protect critical infrastructure from cyber threats and provides guidance, support, and coordination for incident response.

Export Control Regulations

These regulations control the export of certain technologies and information, including cybersecurity tools and data, to prevent unauthorized access or use by foreign entities. Some examples are the International Traffic in Arms Regulations (ITAR) and Export Administration Regulations (EAR)

Surveillance and Data Retention Laws

Governments may impose regulations on data retention and surveillance, requiring organizations to store certain data for specified periods or provide access to law enforcement under certain conditions.

Compliance and Enforcement

Regulatory bodies have the authority to enforce compliance with security regulations through audits, fines, and other penalties. Organizations must stay informed about relevant regulations and ensure they adhere to legal requirements to avoid sanctions.

Overall, government regulations help establish a baseline for security practices, protect sensitive information, and promote trust in digital systems. Organizations must understand and comply with these regulations to safeguard their operations and avoid legal repercussions.

1.4 Architecture and implementation layers

This section delves into how security and resiliency are inherently woven into the IBM Power stack across its various layers. It also underscores key considerations for designing and implementing your IBM Power infrastructure to ensure optimal security and performance.

Figure 1-2 illustrates how the IBM Power ecosystem with IBM Power10 processors provides protection at every layer.

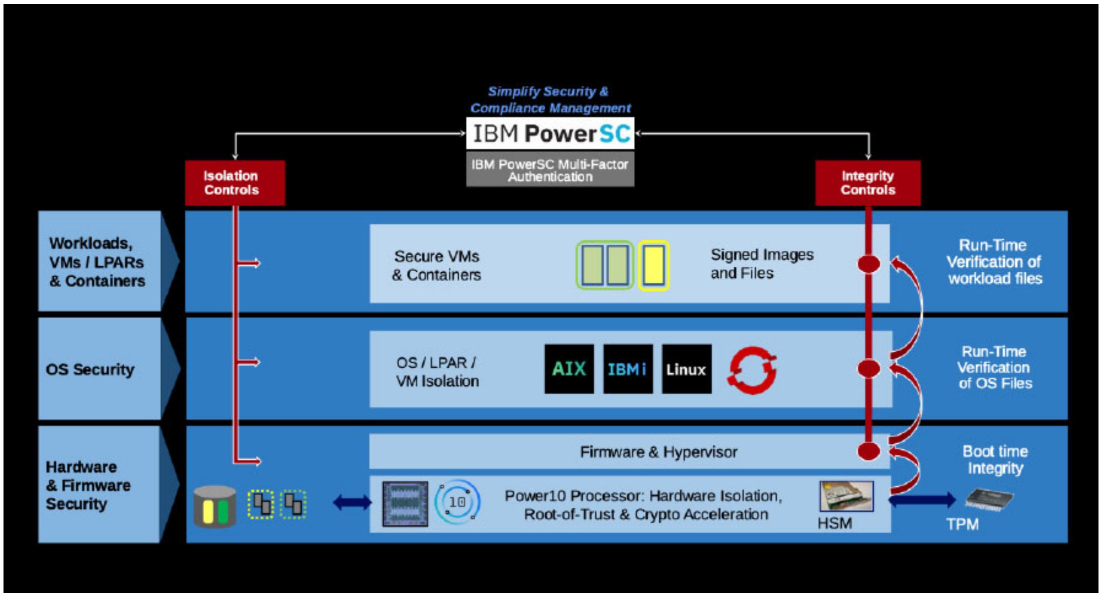

Simplify Security & Compliance Management
IBM PowerSC
IBM PowerSC Multi-Factor Authentication
Isolation Controls
Integrity Controls
Workloads, VMs / LPARs & Containers
Secure VMs & Containers
Signed Images and Files
Run-Time Verification of workload files
OS Security
OS / LPAR / VM Isolation
AIX
IBM i
Linux
Run-Time Verification of OS Files
Hardware & Firmware Security
Firmware & Hypervisor
Boot time Integrity
Power10 Processor: Hardware Isolation, Root-of-Trust & Crypto Acceleration
10
HSM
TPM

Figure 1-2 Layers of protection with IBM Power10[2]

[2] https://events.ibs.bg/events/itcompass2021.nsf/IT-Compass-2021-S06-Power10.pdf

1.4.1 Principle of Least Privilege

This principle involves limiting access to the minimum necessary to perform authorized activities, thereby significantly reducing the risk of an attacker gaining access to critical systems or sensitive information. To be most effective, this principle should be applied at every level of your infrastructure. Implementing this principle involves:

- Access Control Policies

  Establishing comprehensive policies that dictate who can access specific resources based on their job requirements.

- Role-Based Access Control (RBAC)

  Implementing RBAC systems that assign permissions to roles rather than individuals, making it easier to manage and audit access rights across a large organization.

- Regular Audits and Reviews

  Periodically reviewing access rights and usage logs to ensure compliance with the principle of least privilege and to detect any unauthorized access attempts or policy violations.

By adopting these practices, users of IBM Power systems can not only bolster their defenses against current threats but also foster a more resilient posture to adapt to future security challenges.

1.4.2 Hardware

The security of physical hardware is fundamental to protecting the overall integrity of IBM Power systems. This section explores the multiple facets of hardware security, from the physical measures used to protect equipment to the embedded technologies designed to safeguard data and systems from cyber threats.

Physical Security Controls

Physical security controls are critical for protecting systems from unauthorized access, physical damage, or theft. This section delves into various measures and technologies that enhance the physical security of your critical infrastructure locations and components.

- Physical Barriers

  Physical barriers serve as the first line of defense in protecting sensitive hardware. Barriers such as walls, gates, and secure enclosures prevent unauthorized physical access.

- Fencing and Gates

  Perimeter fencing and security gates serve as the first line of defense, controlling access to your facilities. To enhance security, consider reinforced materials and integrating advanced features such as biometric locks. This comprehensive approach provides a robust barrier against unauthorized entry.

- Secure Enclosures

  For shared locations, the use of secure racks and cages may be necessary to protect critical hardware. These enclosures guard against unauthorized access and can be equipped with additional sensors and alarms for enhanced security.

It is important to integrate physical barriers with electronic security measures, such as surveillance and access control systems, to create a comprehensive security envelope around sensitive hardware. Providing access logging capabilities also helps identify personnel who have accessed the environment.

Access Controls

Access controls are designed to ensure that only authorized individuals can enter specific physical areas where sensitive hardware is located. There are multiple types of access control systems that work with different authentication methodologies. These systems can be broadly categorized into:

- Biometric Systems

  Biometric systems use fingerprint scanners, retina scans, and facial recognition technologies to provide a high level of security by verifying the unique physical characteristics of individuals.

- Electronic Access Cards

  Technologies such as RFID cards, magnetic stripe cards, and smart cards grant access based on credentials stored on the card. Many of these can be managed centrally to update permissions as needed.

- PIN Codes and Keypads

  Requiring the entry of personal identification numbers (PINs) into keypads provides a method of access control that can be easily updated and managed remotely.

Surveillance Systems

Surveillance systems are essential for monitoring physical environments to detect, deter, and document unauthorized activities. This section introduces the purpose and strategic placement of surveillance systems within power system facilities. Depending on your requirements, you may need multiple complementary surveillance systems. Here are some types of surveillance technologies you can consider:

- CCTV Cameras

  Closed-circuit television cameras provide identification and monitoring capabilities. Different types, such as dome, bullet, and PTZ (pan-tilt-zoom) cameras, should be strategically placed and monitored.

- Motion Detectors

  Motion detectors that trigger alerts or camera recordings can enhance the efficiency of surveillance by focusing resources on areas where activity is detected.

- Advanced Surveillance Technologies

  Newer technologies like thermal imaging and night vision cameras, which capture video in low light or through obstructions, can enhance your around-the-clock surveillance capabilities.

> **Important:** Data management and privacy considerations are involved with the collection and storage of surveillance information. It is important to manage surveillance footage securely, including storage, access controls, and compliance with privacy laws and regulations to protect the rights of individuals.

1.4.3 Embedded Security Features

Embedded security features are integral components of modern hardware. They are designed to provide built-in protection against various threats. This section explores various embedded security technologies, their functions, and how they contribute to the overall security architecture of IBM Power systems.

Trusted Platform Module (TPM)

The Trusted Platform Module (TPM) is a dedicated micro-controller designed to secure hardware through integrated cryptographic keys. TPM plays an important role in enhancing hardware security by providing hardware-based, security-related functions.

TPM supports Secure Boot to ensure that only trusted software is loaded during the system's startup process. TPM checks the digital signatures of operating system and application software before allowing them to execute, preventing unauthorized software from running.

TPM also provides key storage and management. It safeguards cryptographic keys at a hardware level, preventing them from being exposed to outside threats. These keys can be used for encrypting data and securing communications.

Secure Boot

IBM Power10 and Power9® servers incorporate Secure Boot, a critical security feature that ensures the integrity of the system's firmware and operating system at startup. This safeguards against unauthorized modifications and potential attacks, providing a robust foundation for secure operations.

Secure Boot verifies the integrity of the firmware, boot loader, and operating system to prevent unauthorized code from running during the boot process. It ensures that only trusted software signed with a valid certificate is executed, protecting against rootkits and boot-level malware that could compromise the system's security before the operating system starts.

Secure Boot utilizes digital signatures and certificates to validate the authenticity and integrity of firmware and software components. Each component in the boot process is signed with a cryptographic key, and the system verifies these signatures before allowing the component to execute.

A trusted key management capability, such as TPM, is required to store and manage cryptographic keys and to support the Secure Boot process. Organizations can manage keys and certificates used in Secure Boot through configuration settings, enabling them to control which software and firmware are trusted.

Secure Boot helps prevent unauthorized code execution during the boot process, protecting the system from early-stage attacks and aiding in meeting compliance requirements for security standards and regulations that mandate secure boot processes.

Hardware Encryption

Hardware encryption involves using dedicated processors that perform cryptographic operations directly within the hardware itself, enhancing security by isolating the encryption process from software vulnerabilities.

Encryption can be implemented at several layers, providing protection of your data as it moves through the system. Power10 provides encryption acceleration built into the chip. This allows the system to encrypt memory by default within the system with no performance impact. As data leaves the processor, encryption can be utilized at the disk level, file system level, and network level providing complete protection for your data.

Hardware Security Modules

Hardware Security Modules (HSMs) are physical devices that manage digital keys for strong authentication and provide secure crypto-processing. HSMs generate, store, and manage encryption keys in a tamper-resistant environment, ensuring that keys are never exposed to potentially compromised operating systems. HSMs are integral to digital signing processes, ensuring the integrity and authenticity of software updates and communications. Hardware security modules are discussed in 2.1.2, “Encryption enablement in hardware” on page 32.

1.4.4 Risk Management in Hardware

A significant part of security involves identifying and managing risk. Understanding your environment and its vulnerabilities is critical to creating a plan to protect your enterprise from attacks that can compromise your data or interrupt business operations.

Regular vulnerability assessments are vital for identifying weaknesses in hardware that could be exploited by attackers or fail under operational stress. These assessments should include physical inspections, cybersecurity evaluations, and testing against environmental and operational conditions. Techniques such as penetration testing and red team exercises can simulate real-world attack scenarios to test the resilience of hardware components.

Protecting your environment should include the use of continuous monitoring technologies, including hardware sensors and network monitoring tools, which play a critical role in the early detection of potential failures or security breaches.

Regular reviews ensure that risk management strategies and practices stay relevant as new threats emerge and business needs change. This involves re-evaluating and updating risk assessments, mitigation strategies, and response plans at defined intervals or after significant system changes.

Having detailed incident response and recovery plans is essential for minimizing downtime and restoring functionality in the event of hardware failure or a security incident. These plans need to include roles and responsibilities, communication strategies, and recovery steps.

Training programs for IT staff, operators, and other stakeholders involved in hardware management are crucial for maintaining system security. Effective documentation and reporting are also fundamental to the risk management process. It is important to be transparent in reporting to stakeholders and regulatory bodies.

1.4.5 Virtualization

Virtualization has become a cornerstone of modern IBM Power systems, enabling enhanced flexibility and efficiency. However, the shift to virtual environments also introduces specific security challenges that must be addressed to protect these dynamic and often complex systems.

Security in a Virtualized Environment

A virtualized environment provides better utilization of compute resources and allows more flexibility in setting up and running your environment. Virtualization also simplifies creating highly available and resilient systems, allowing workloads to be moved to support hardware maintenance and outages.

The function that allows virtualization in a system is called a hypervisor, also known as a virtual machine monitor (VMM). The hypervisor is a type of computer software that creates and runs virtual machines, which are also called logical partitions (LPARs). The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of the guest operating systems. Hypervisors are generally classified into two types. A Type 1 hypervisor is a native hypervisor that runs on bare metal. In contrast, a Type 2 hypervisor is hosted on an underlying operating system. The Type 1 hypervisor is considered more secure as it can provide better isolation between the VMs and generally offers better performance to those VMs.

IBM's solution for virtualization on a Power server, PowerVM, is a combination of hardware, firmware, and software components that provide a foundation for virtualizing CPU, storage, and network resources. At the heart of PowerVM is the Power hypervisor, which is built into the Power Systems' firmware. As a Type 1 hypervisor, PowerVM allows the creation of multiple logical partitions (LPARs) – also referred to as Virtual Machines (VMs) – to run on a single Power server and provides the necessary LPAR isolation. Each LPAR is assigned a set of resources such as memory, CPU, disk space, and adapters to connect to other resources. The operating system in each LPAR is only aware of the resources it has been assigned.

Figure 1-3 illustrates how PowerVM works.

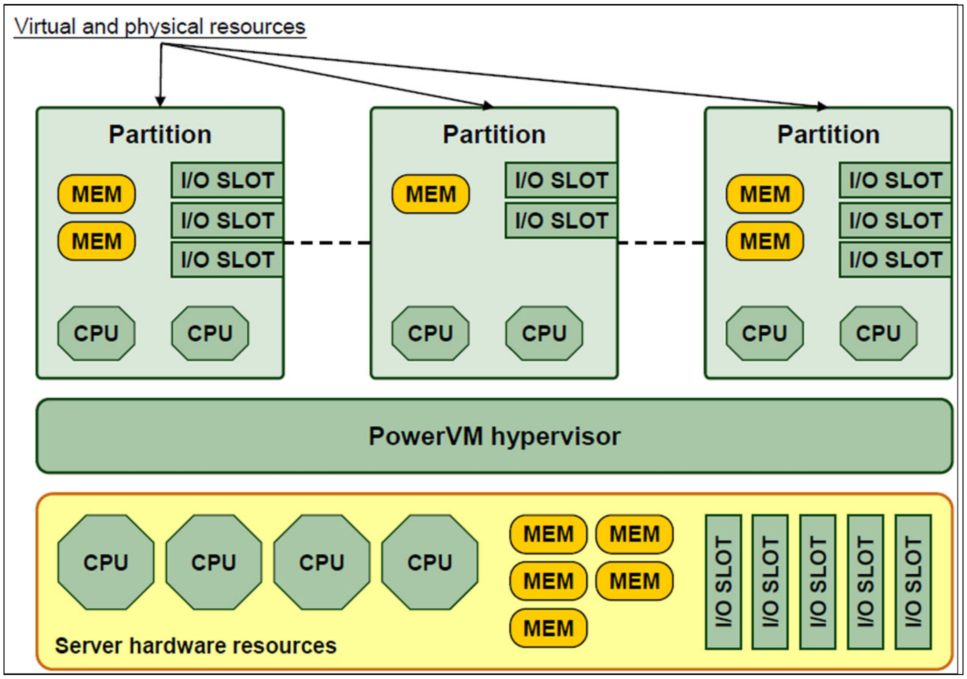


Figure 1-3 PowerVM illustration

The following list provides some security implications that need to be addressed by the virtualization layer:

- Isolation Failures

  As there are multiple VMs running at any one time on the same physical hardware, it is imperative that the hypervisor maintain strict isolation between virtual machines to prevent a breach in one VM from compromising others.

- Hypervisor Security

  The hypervisor, being the hardware and software layer that enables virtualization, presents a critical security focal point. Ensuring that the hypervisor is secure and kept up to date is key.

- VM Sprawl

  VM sprawl refers to the rapid increase in the number of VMs as your environment grows. This rapid increase can lead to management challenges and potential security oversights. Strategies for controlling VM sprawl, such as inventory management, LPAR consolidation, and lifecycle control policies, are important.

- Secure VM Configuration

  Standard processes for maintaining the security of your VMs should be followed. This includes practices such as using hardened base images, applying least privilege principles for VM access, and encrypting VM data both at rest and in transit.

Monitoring and Management

Effective management and continuous monitoring are essential for maintaining the security of virtual environments. These include:

- Real-time Monitoring Tools

  Integrate tools that provide real-time monitoring of virtual environments, highlighting anomalies and potential security threats. This includes solutions that offer visibility into VM operations and network traffic patterns.

- Configuration Management

  Maintain a consistent and secure configuration across all virtual assets. This includes using configuration management tools that can automate the application of security settings and patches.

- Log Management and Analysis

  It is important to collect and analyze logs from the virtual systems. Log management solutions can help detect, investigate, and respond to security incidents within virtual environments.

- Compliance Auditing

  Conduct regular audits to ensure that virtual environments comply with relevant security standards and regulations. This includes using automated tools to streamline the auditing process and ensure continuous compliance.

1.4.6 HMC and CMC

The Hardware Management Console (HMC) and Cloud Management Console (CMC) are critical components in managing and monitoring the physical and virtual environments running on IBM Power systems. This section explores their roles, security challenges, and best practices to ensure their security.

HMC

The Hardware Management Console (HMC) is used to configure and manage IBM Power systems. Its capabilities encompass logical partitioning, centralized hardware management, Capacity on Demand (CoD) management, advanced server features, redundant and remote system supervision, and security.

The HMC provides a reliable and secure console for IBM Power systems. It is built as an appliance on a highly secured system, tied to specific hardware, and not compatible with other systems. This stringent build process includes incorporating advanced hardware and software security technologies from IBM. Furthermore, HMCs are closed and dedicated, meaning users cannot add their own software. These features work together to create a highly secure environment.

CMC

The Cloud Management Console (CMC) allows you to securely view information and gain insights about your Power Systems infrastructure across multiple locations. Dynamic views of performance, inventory, and logging for your complete Power enterprise—whether on-premises or off-premises—simplify and unify information in a single location. CMC provides consolidated information and analytics, which can be key enablers for the smooth operation of infrastructure. Hosted on IBM Cloud, the CMC is a highly secure cloud-based service accessible from mobile devices, tablets, and PCs.

1.4.7 Operating Systems

IBM Power offers unparalleled flexibility by enabling you to consolidate diverse operating environments onto a single system. From industry-leading options like AIX and IBM i to the widely adopted Linux and Red Hat OpenShift platforms, you can harness the power of IBM Power to consolidate mission-critical applications across any number of systems. This consolidation provides enhanced reliability, availability, and security.

Supported Operating Systems

As of the time of this publication, Power10 processor-based systems support the following platforms/operating system versions shown in Table 1-1.

Table 1-1 Power10 platform / operating system support matrix

Operating System	Supported versions
Red Hat OpenShift Container Platform	4.9 or later
PowerVM Virtual I/O Server	4.1.0.0 or later 3.1.4.10 or later 3.1.3.10 or later 3.1.2.30 or later 3.1.1.50 or later
AIX	7.3 TL0 or later 7.2 TL4 or later (with any I/O configuration) 7.1 TL5 or later (through VIOS only)
IBM i	7.5 or later 7.4 TR5 or later 7.3 TR11 or later
Red Hat Enterprise Linux	8.4 or later 9.0 or later
SUSE Linux Enterprise Server	15.3 or later 12.5 or later
Ubuntu	22.04 or later

A full list of operating systems that run on IBM Power is also available at `https://www.ibm.com/power#Operating+systems`.

Note: Table 1-1 on page 16 shows the supported operating systems of the Power E1080. Software maps detailing which versions are supported on which specific IBM Power server models (including previous generations of IBM Power) can be found in the IBM Support page at `https://www.ibm.com/support/pages/node/6020068`.

1.4.8 Storage

An enterprise's data is a critical resource and must be available play a crucial role in data management, ensuring that vast amounts of operational and historical data are securely stored, readily accessible, and efficiently managed. This section delves into the various aspects of storage technology, highlighting types of storage architectures, security measures, and best practices for managing storage in power systems.

Storage topologies

There are multiple methods of connecting storage to your servers. The different options have evolved over time to meet different requirements and each type has benefits and disadvantages. They also vary in performance, availability, and price.

Direct Attached Storage

Direct Attached Storage (DAS) is a type of storage that is directly connected to a computer or server without a network in between. This contrasts with Network Attached Storage (NAS) or Storage Area Networks (SAN), which involve network connections. DAS is characterized by these characteristics:

- Usage
  - Small Businesses: Small businesses often use DAS for straightforward, cost-effective storage solutions, including external hard drives or RAID systems directly connected to a server or workstation.
  - Enterprise: In enterprise environments, DAS can be utilized for high-performance tasks where large amounts of data need to be quickly accessible, such as for databases or virtual machines.
- Benefits
  - High Performance: Since DAS is directly connected, it often provides faster access speeds and lower latency compared to network-based storage solutions. This is crucial for applications requiring rapid data retrieval.
  - Simplicity: Setting up DAS is straightforward because it doesn't involve network configurations. This makes it easy to install and manage, especially for non-technical users.
  - Cost-Effective: Generally, DAS solutions are less expensive than NAS or SAN systems because they don't require additional network infrastructure or management tools.
  - Control and Security: With DAS, data is stored directly on the device connected to the computer or server, which can enhance control and security as it is not accessible over a network.
- Disadvantages
  - Scalability: Expanding storage with DAS can be cumbersome. Adding more storage often requires physically connecting additional drives or upgrading existing hardware, which can be less flexible compared to network-based solutions.
  - Limited Sharing: DAS typically does not support easy sharing across multiple computers or users. Each device connected to DAS usually has exclusive access, making it less suitable for environments requiring collaborative access.

  - Management Complexity: In environments with multiple DAS devices, managing and backing up data can become complex. Unlike NAS or SAN, which often include centralized management tools, DAS requires individual management of each device.
  - Redundancy and Backup: Implementing redundancy and backup solutions with DAS can be more challenging. While RAID configurations can provide redundancy, managing and monitoring these setups can be more labor-intensive compared to solutions with built-in redundancy features in networked storage systems.

In summary, DAS offers high performance and simplicity but can be limited in scalability and sharing capabilities. Its benefits make it suitable for scenarios where high speed and control are priorities, while its disadvantages suggest it may not be ideal for environments needing extensive collaboration or large-scale storage expansion.

Network Attached Storage

Network Attached Storage (NAS) is a dedicated file storage system connected to a network, allowing multiple users and devices to access and share data over the network. NAS devices typically contain one or more hard drives and have their own operating system and management interface. NAS generally has the following characteristics:

- Usage
  - Small and Medium Businesses (SMBs): SMBs use NAS for file sharing, backup solutions, and as a centralized repository for documents and other business-critical data. NAS devices in this context can offer features like user authentication, access control, and remote access.
  - Enterprise Environments: In enterprises, NAS systems are used for departmental file sharing, backup, and collaboration. Advanced NAS devices can support high-capacity storage, multiple RAID configurations for redundancy, and integration with enterprise applications.
- Benefits
  - Ease of Access: NAS provides a centralized location for data, making it accessible from any device on the network. This facilitates file sharing and collaboration among multiple users.
  - Scalability: NAS systems can often be expanded by adding additional drives or connecting multiple NAS units. This allows for scalable storage solutions as data needs grow.
  - Cost-Effective: NAS is generally more affordable compared to SAN solutions, offering a good balance between performance and cost, especially for SMBs and home users.
  - Centralized Management: NAS devices come with management interfaces that allow for easy setup, monitoring, and maintenance. They often include features like data encryption, access controls, and user management.
  - Data Redundancy: Many NAS devices support RAID configurations, which provide redundancy and protection against data loss due to drive failure.
- Disadvantages
  - Network Dependency: NAS performance is dependent on the network’s speed and reliability. High network traffic or network issues can impact access speeds and performance.
  - Limited Performance: While NAS provides adequate performance for many applications, it may not be suitable for high-performance tasks that require extremely fast data access, such as high-frequency trading or large-scale data processing.
  - Complexity in Large Deployments: In larger environments with numerous NAS devices, managing multiple units and ensuring consistent backup and security policies can become complex.

  - Security Risks: Since NAS devices are network-connected, they are vulnerable to network-based threats. Proper security measures, such as encryption, firewalls, and access controls, are necessary to protect data.
  - Cost of Advanced Features: While basic NAS units are affordable, those with advanced features like high capacity, high availability, or enterprise-level functionality can become expensive.

In summary, NAS offers centralized, accessible, and scalable storage solutions suitable for a wide range of environments, from home use to enterprise settings. It excels in providing file sharing and backup capabilities but may face limitations in performance and complexity as needs grow. Proper network infrastructure and security measures are crucial for optimizing NAS performance and protecting data

Storage Area Network:

A Storage Area Network (SAN) is a specialized network that provides high-speed, dedicated access to consolidated storage resources. Unlike Network Attached Storage (NAS), which operates over a standard network, SAN is designed specifically for storage and often uses high-performance connections like Fibre Channel or iSCSI. SAN storage generally has these characteristics:

- Usage
  - Enterprise Environments: SAN is commonly used in large-scale enterprise environments where high performance, scalability, and reliability are critical. It is often employed for mission-critical applications, large databases, and virtualized environments.
  - Data Centers: SANs are widely used in data centers to provide centralized storage for multiple servers. They support high-capacity and high-performance storage needs, facilitating efficient data management and backup.
  - High-Performance Computing: Applications that require fast data access and large volumes of data, such as scientific simulations or financial transactions, benefit from SAN's high throughput and low latency.
- Benefits
  - High Performance: SANs provide high-speed data access, which is essential for performance-intensive applications. Technologies like Fibre Channel offer very low latency and high throughput.
  - Scalability: SANs can be easily scaled by adding more storage devices or connecting additional servers, making them suitable for growing data needs and expanding workloads.
  - Centralized Storage Management: SANs consolidate storage into a single, centralized system, simplifying storage management, backup, and recovery processes.
  - Data Redundancy and Reliability: SANs often support advanced redundancy features, such as multiple paths to storage devices and RAID configurations, enhancing data protection and availability.
  - Virtualization Support: SANs are well-suited for virtualized environments, providing flexible and efficient storage allocation and management for virtual machines.
- Disadvantages
  - Cost: SANs can be expensive to implement and maintain, especially with high-performance components like Fibre Channel switches and storage arrays. The initial investment and ongoing operational costs can be significant.
  - Complexity: SANs are complex to set up and manage. They require specialized knowledge and expertise for configuration, management, and troubleshooting. This complexity can lead to increased administrative overhead.

  - Infrastructure Requirements: SANs require dedicated hardware and infrastructure, such as Fibre Channel switches or iSCSI adapters, which can add to the cost and complexity of the overall system.
  - Network Congestion: While SANs are designed to avoid network congestion, in some cases, the network infrastructure supporting the SAN can become a bottleneck if not properly managed or if it is shared with other traffic.
  - Security Risks: SANs are often accessed over dedicated networks, but they still require robust security measures to protect data from unauthorized access or breaches. Proper access controls and encryption are essential to safeguard data.

In summary, SAN provides high-performance, scalable, and centralized storage solutions ideal for enterprise and data center environments. It excels in performance and reliability but can be costly and complex to implement and manage. Organizations using SANs need to balance their need for high-speed data access with the associated infrastructure and operational costs.

Cloud Storage

Cloud storage refers to the practice of storing data on remote servers that can be accessed over the internet. Providers manage these servers and offer various services for storing, managing, and retrieving data. This model contrasts with traditional on-premises storage solutions, where data is stored locally on physical devices. Cloud storage can be characterized as:

- Usage
  - Small and Medium Businesses (SMBs): SMBs leverage cloud storage for file sharing, collaboration, and remote work. It provides a cost-effective way to scale storage needs without investing in physical infrastructure.
  - Large Enterprises: Enterprises use cloud storage for scalable data storage solutions, disaster recovery, and global access. It supports extensive data needs, facilitates collaboration, and integrates with various enterprise applications.
  - Developers and IT Professionals: Cloud storage is used for hosting applications, managing databases, and providing scalable storage solutions for big data and analytics.
- Benefits
  - Scalability: Cloud storage offers virtually unlimited storage capacity. Users can scale their storage up or down based on their needs without needing to invest in physical hardware.
  - Accessibility: Data stored in the cloud can be accessed from anywhere with an internet connection. This supports remote work, collaboration, and access from multiple devices.
  - Cost-Effectiveness: Typically, cloud storage operates on a pay-as-you-go model, allowing users to pay only for the storage they use. This reduces upfront costs associated with purchasing and maintaining physical storage hardware.
  - Automatic Updates and Maintenance: Cloud providers handle software updates, security patches, and hardware maintenance, freeing users from these tasks and ensuring that the storage environment is up-to-date.
  - Disaster Recovery: Many cloud storage services include built-in redundancy and backup solutions, providing enhanced data protection and recovery options in case of data loss or system failures.
- Disadvantages
  - Security and Privacy: Storing data off site introduces concerns about data security and privacy. Users must trust cloud providers to protect their data from breaches and unauthorized access. Encryption and other security measures are essential but may not be foolproof.

- Internet Dependence: Access to cloud storage is dependent on a stable internet connection. Poor connectivity can hinder access to data and affect performance.
- Ongoing Costs: While cloud storage can be cost-effective, ongoing subscription fees can add up over time, especially for large amounts of storage or high-performance requirements.
- Data Transfer Speeds: Uploading and downloading large amounts of data can be slow, especially with limited internet bandwidth. This can impact the speed at which data is accessible or transferred.
- Vendor Lock-In: Different cloud providers use proprietary technologies and formats, which can make it challenging to migrate data between services or integrate with other systems. This can lead to vendor lock-in and potential difficulties if you decide to switch providers.

Cloud storage offers flexible, scalable, and accessible solutions suitable for personal, business, and enterprise needs. Its benefits include scalability, cost-effectiveness, and automatic maintenance, making it an attractive option for modern data management. However, concerns about security, reliance on internet connectivity, ongoing costs, and potential vendor lock-in are important considerations that users must address when opting for cloud storage solutions.

Security Considerations for Storage

When it comes to securing storage systems, whether on-premises or in the cloud, several key considerations are crucial to protecting data from unauthorized access, breaches, and other security threats. Here’s an overview of important security considerations for storage:

1. Data Encryption

   At Rest: Encrypt data stored on physical media to protect it from unauthorized access if the storage device is stolen or compromised.

   In Transit: Use encryption protocols like TLS/SSL for data transmitted over networks to prevent interception and unauthorized access during transmission.

2. Access Controls

   Authentication: Implement strong authentication mechanisms, such as multi-factor authentication (MFA), to ensure that only authorized users can access storage systems.

   Authorization: Define and enforce access controls to limit what users can see and do based on their roles and permissions. Implement principles of least privilege to minimize access rights.

3. Data Integrity

   Checksums and Hashes: Use checksums or cryptographic hashes to ensure data integrity and detect any unauthorized alterations or corruption of data.

   Version Control: Maintain version histories of important data to recover from accidental deletions or modifications.

4. Backup and Recovery

   Regular Backups: Perform regular backups of critical data and store backups securely, preferably in a different location from the primary storage to protect against site-specific disasters.

   Test Recovery: Periodically test backup and recovery processes to ensure data can be restored quickly and accurately in case of data loss or corruption.

5. Physical Security

   Data Center Security: For on-premises or co-located storage, ensure that data centers have robust physical security measures, such as restricted access, surveillance, and environmental controls.

   Device Security: Secure physical storage devices to prevent unauthorized access and theft. Consider measures like locked server rooms or safes for sensitive equipment.

6. Network Security

   Firewalls and Intrusion Detection: Use firewalls and intrusion detection/prevention systems to protect storage networks from unauthorized access and cyber threats.

   Segmentation: Segment storage networks from other networks to limit exposure and potential attack vectors.

7. Monitoring and Auditing

   Activity Monitoring: Implement logging and monitoring to track access and changes to storage systems. This helps in identifying suspicious activities and responding to potential security incidents.

   Regular Audits: Conduct regular security audits and assessments to ensure compliance with security policies and identify vulnerabilities.

8. Compliance

   Regulatory Requirements: Adhere to relevant regulations and standards, such as GDPR, HIPAA, and PCI-DSS, which mandate specific security practices for protecting data.

   Data Sovereignty: Understand and comply with data residency requirements and local laws related to data storage and protection.

9. Vendor Management

   Cloud Providers: When using cloud storage, carefully evaluate the security measures and practices of the cloud service provider. Ensure they meet your security requirements and comply with relevant regulations.

   Third-Party Risk: Assess and manage risks associated with third-party vendors who have access to your storage systems or handle your data.

10. Incidentnt Response

    Plan and Preparation: Develop and maintain an incident response plan to address security breaches or data loss. This plan should include procedures for containment, investigation, and communication.

    Training: Regularly train staff on security best practices and how to respond to security incidents to ensure preparedness.

Securing storage systems involves a comprehensive approach that includes data encryption, robust access controls, regular backups, physical security, network protections, monitoring, compliance, and vendor management. By addressing these considerations, organizations can significantly reduce the risk of data breaches, loss, and other security incidents.

Best Practices for Storage Management

Beyond the security aspects of your data, there are some other considerations that need to be addressed in your storage environment including compliance requirements and planning for recovery in case of cyber attack or ransomware. Effective storage management ensures that data is stored securely, efficiently, and cost-effectively.

Here are some best practices for managing storage across various environments, including on-premises and cloud-based solutions:

1. Capacity Planning

   Assess Needs: Regularly evaluate current and future storage needs based on data growth projections, application requirements, and business goals.

   Optimize Usage: Use tools to monitor storage utilization and optimize space. Consider implementing data deduplication and compression to reduce the amount of storage required.

2. Data Classification and Organization

   Classify Data: Categorize data based on its importance, sensitivity, and usage patterns. This helps in applying appropriate storage and security policies.

   Organize Efficiently: Structure storage systems to facilitate easy access and retrieval. Use logical grouping and hierarchical storage management to keep data organized.

3. Data Backup and Recovery

   Regular Backups: Implement a consistent backup schedule to ensure that data is regularly backed up. Consider full, incremental, and differential backups based on the needs.

   Test Recovery: Periodically test backup and recovery procedures to ensure data can be restored quickly and accurately in case of loss or corruption.

4. Data Retention Policies

   Define Policies: Establish clear data retention policies based on legal requirements, business needs, and data usage patterns. Determine how long different types of data should be kept before deletion or archiving.

   Automate Management: Use automated tools to enforce retention policies, manage data lifecycle, and handle the archiving or deletion of obsolete data.

5. Performance Optimization

   Monitor Performance: Regularly monitor storage performance metrics, such as I/O operations and response times, to identify and address bottlenecks.

   Optimize Storage: Use techniques such as tiered storage to allocate high-performance storage to critical applications while using lower-cost storage for less critical data.

6. Cost Management

   Budgeting: Develop and adhere to a storage budget that aligns with business needs and growth projections.

   Cost Optimization: Regularly review storage costs and explore cost-saving options, such as moving infrequently accessed data to lower-cost storage tiers or cloud storage solutions.

7. Disaster Recovery Planning

   Plan Development: Create a disaster recovery plan that outlines procedures for data backup, restoration, and continuity of operations in case of catastrophic events.

   Regular Updates: Review and update the disaster recovery plan regularly to adapt to changes in the business environment or technology.

8. Compliance and Auditing

   Ensure Compliance: Stay compliant with relevant regulations and standards (e.g., GDPR, HIPAA, PCI-DSS) regarding data storage and protection.

   Conduct Audits: Perform regular security and compliance audits to identify and address any gaps or issues in storage management practices.

9. Vendor Management

   Evaluate Providers: When using third-party storage solutions, thoroughly evaluate vendors based on their security, reliability, and performance.

   Manage Contracts: Clearly define service level agreements (SLAs) and terms in contracts with storage vendors to ensure they meet your performance and security requirements.

10. Documentation and Training

    Document Procedures: Maintain comprehensive documentation of storage management policies, procedures, and configurations.

    Train Staff: Provide training for staff involved in storage management to ensure they are knowledgeable about best practices, tools, and security measures.

11. Automation and Tools

    Implement Automation: Use storage management tools and automation to streamline tasks such as provisioning, monitoring, and maintenance.

    Utilize Management Software: Leverage software solutions for centralized management of

Effective storage management involves strategic planning, organization, and implementation of best practices in data classification, backup, security, and performance optimization. By adhering to these practices, organizations can ensure that their storage systems are reliable, secure, and aligned with business objectives while managing costs and compliance requirements efficiently.

Safeguarded Copy and data resiliency

Safeguarded Copy is a term used in data management and backup solutions to refer to backup copies of data that are specifically protected to ensure their reliability and integrity. The concept focuses on creating and maintaining backup copies that are not only reliable but also secure from various threats, including corruption, tampering, and unauthorized access. It’s an essential part of a robust data protection strategy, especially in environments where data integrity and availability are critical.

The key aspects of SafeGuarded Copy are:

- Data Integrity:
  - Consistency: Safeguarded copies are created to reflect a stable state of the data at a specific point in time. This can involve techniques like snapshots or consistent backups to ensure that all parts of the data are accurately captured.
  - Validation: Integrity checks, such as checksums or cryptographic hashes, are used to verify that backup data has not been altered or corrupted. This ensures that the data in the backup matches the original data.
- Security:
  - Encryption: Backup copies are encrypted both at rest and during transmission. Encryption protects data from unauthorized access and ensures that even if a backup is compromised, the data remains secure.
  - Access Controls: Strict access controls are enforced to limit who can access or manage the backup copies. This helps prevent unauthorized access or tampering with the backup data.
- Protection from Ransomware:
  - Immutable Backups: Implement features such as write-once, read-many (WORM) to make backup copies immutable. This means that once a backup is created, it cannot be altered or deleted by ransomware or malicious actors.

  - Isolated Storage: Store backup copies in a separate location or isolated environment that is not directly accessible from the primary network. This reduces the risk of backups being affected by ransomware or other attacks.
- Automated Management:
  - Scheduling: Automated scheduling ensures that backups are performed regularly and consistently without relying on manual intervention. This helps maintain up-to-date backup copies.
  - Verification: Automated verification processes check the integrity of backup copies to ensure they are usable and not corrupted.
- Disaster Recovery Readiness:
  - Testing: Regularly test backup and recovery processes to ensure that safeguarded copies can be restored effectively. This includes performing periodic restore tests to verify the accuracy and completeness of backups.
  - Documentation: Maintain detailed documentation of backup procedures, configurations, and recovery plans. This helps ensure that backup and restoration processes are clear and can be executed quickly in case of an emergency.
- Compliance:
  - Regulatory Requirements: Ensure that safeguarded copies meet regulatory and industry standards for data protection and privacy. This includes adhering to requirements for data retention, security, and access control.

In summary, Safeguarded Copy refers to backup copies of data that are specifically protected to ensure their integrity, security, and reliability. This involves creating consistent and reliable backups, encrypting and securing backup data, protecting against threats like ransomware, and automating management processes. By implementing SafeGuarded copies, organizations can ensure that their data backups are robust, secure, and capable of supporting effective disaster recovery and data protection strategies.

> **Important:** Safeguarded Copy is not just a physical copy of the data. It involves automation and management to take regular copies, validate that they are valid and are stored so that they cannot be modified. Of equal importance is the ability to quickly recognize when your data has been compromised and recover to a last good state. It also involves business processes to be able to recover applications and databases to minimize data loss.

IBM Storage provides a Safeguarded Copy capability in both the IBM DS8000® and the IBM FlashSystem® systems. For more information on the IBM solutions see:

  - *IBM Storage DS8000 Safeguarded Copy: Updated for DS8000 Release 9.3.2*, REDP-5506
  - *Data Resiliency Designs: A Deep Dive into IBM Storage Safeguarded Snapshots*, REDP-5737
  - *Cyber Resiliency with IBM Storage Sentinel and IBM Storage Safeguarded Copy*, SG24-8541

1.4.9 Networking

Security considerations for networking involve several key aspects to protect data integrity, confidentiality, and availability across networked systems. Whether using physical networking connections or virtualize the network functions, the considerations are generally the same. Here are some essential considerations:

- Network Segmentation

  Dividing a network into segments can limit the spread of attacks and contain potential breaches. Segmentation helps isolate sensitive data and systems from less critical areas.

- Firewalls

  Firewalls act as barriers between internal networks and external threats. They filter incoming and outgoing traffic based on predefined security rules.

- Intrusion Detection and Prevention Systems (IDPS)

  These systems monitor network traffic for suspicious activity and can either alert administrators or take action to block potential threats.

- Encryption

  Encrypting data transmitted over the network ensures that even if data is intercepted, it remains unreadable without the proper decryption keys.

- Access Controls

  Implementing strong access controls, including multi-factor authentication and least privilege principles, ensures that only authorized users and devices can access network resources.

- Regular Updates and Patching

  Keeping network devices and software up to date with the latest security patches helps protect against known vulnerabilities and exploits.

- Network Monitoring

  Continuous monitoring of network traffic and device behavior helps detect and respond to anomalies and potential security incidents in real-time.

- Secure Configuration

  Ensuring that network devices (e.g., routers, switches) are securely configured according to best practices reduces the risk of exploitation through bad actor connections and potential threats helps reduce the risk of human error and social engineering attacks.

- Incident Response Planning

  Having a plan in place for responding to network security incidents helps minimize damage and recover quickly from breaches.

- Security Policies and Training

  Establishing clear security policies and training employees on best practices and potential threats helps reduce the risk of human error and social engineering attacks.

Addressing these considerations helps build a robust network security posture and protect against various cyber threats.

2

Protection across every layer

An infrastructure built using IBM Power servers benefits from the robust security technologies integrated into both the hardware and software stacks. IBM Power servers offer advanced security features at every level of the system, ensuring comprehensive protection for sensitive data and applications. These features include advanced encryption technologies, secure boot capabilities, and integrated firmware updates. Additionally, IBM Power servers leverage IBM's extensive expertise in securing mission-critical workloads, making them a popular choice for organizations seeking a secure environment for their digital assets.

Workloads on the Power10 server see significant benefits from improved cryptographic accelerator performance compared to previous generations. Specifically, the Power10 chip supports accelerated cryptographic algorithms such as AES, SHA2, and SHA3, resulting in considerably higher per-core performance for these algorithms. This enhancement allows features like AIX Logical Volume Encryption to operate with minimal impact on system performance.

The processor-core technology of the Power10 incorporates integrated security protections aimed at:

- Improved Cryptographic Performance: Support for fully homomorphic encryption (FHE) and quantum-safe cryptography helps you stay ahead of both current and future data threats.
- Increased Application Security: Hardened defenses against return-oriented programming (ROP) attacks.
- Simplified Hybrid Cloud Security: Easy-to-use, setup-free hybrid cloud security administration with a single interface.
- Enhanced Virtual Machine Isolation: Providing the industry's most secure virtual machine isolation technology, which boasts significantly fewer Common Vulnerability Exposures (CVEs) compared to hypervisors associated with x86 processor-based servers[1]. This technology effectively shields your applications and data from potential exploits.

[1] https://itic-corp.com/ibm-z-ibm-power-systems-lenovo-thinksystem-servers-most-secure-toughest-to-crack/

In this chapter we discuss the following topics:

- 2.1, “Encryption technologies and their applications” on page 29
- 2.2, “Compliance automation and real-time monitoring” on page 37
- 2.3, “Endpoint detection and response” on page 38
- 2.4, “Malware prevention technologies” on page 38
- 2.5, “Trusted boot” on page 39
- 2.6, “Hypervisor security” on page 40

2.1 Encryption technologies and their applications

Power10 emphasizes comprehensive security throughout its design, offering multiple encryption options. Key among these are Transparent Memory Encryption (TME), Fully Homomorphic Encryption (FHE), and Quantum-Safe Encryption (QSE).

Transparent Memory Encryption (TME) encrypts data in memory to protect it from unauthorized access and tampering during runtime. Operating at the hardware level, TME utilizes the Power10 processor's cryptographic engines to perform encryption and decryption tasks efficiently. Each Power10 core includes four cryptographic engines to handle these operations without significantly impacting performance. TME ensures pervasive protection of data in memory with minimal impact on system performance, as encryption and decryption are managed at the chip level. Its integration into normal operations is seamless and automatic.

Fully Homomorphic Encryption (FHE) allows computations to be performed directly on encrypted data without decrypting it first. This ensures that sensitive data remains confidential even during processing. FHE operates at the software level and involves sophisticated mathematical algorithms to enable computations on ciphertexts. Implementing FHE requires specialized libraries and frameworks. However, FHE is computationally intensive and can introduce performance overhead compared to conventional hardware-only encryption methods due to the complexity of the algorithms.

Quantum-Safe Encryption (QSE) is designed to be resistant to quantum attacks, securing data against the computational capabilities of future quantum computers that could potentially break current cryptographic algorithms. QSE employs cryptographic algorithms believed to be resistant to quantum attacks, such as lattice-based, hash-based, and multivariate-quadratic-equations-based cryptography. Many quantum-safe algorithms are still undergoing testing and standardization to ensure they provide robust security in the face of future quantum advancements. QSE is typically used for securing long-term data, sensitive communications, and critical infrastructure.

The relevant features and differences in these technologies is shown in Table 2-1.

Table 2-1 Key Differences

Feature	TME	FHE	QSE
Encryption Scope	Secures data in memory	Allows computations on encrypted data	Prevents against future quantum computing threats
Implementation Level	Implemented in hardware	Implemented using a combination of hardware and software	Implemented using a combination of hardware and software
Performance Impact	Hardware accelerated through the Power10 cryptographic engines and designed to have minimal performance impact	Involves substantial computational overhead	The impact of QSE varies; some quantum-safe algorithms may introduce performance overhead; this is a subject of ongoing research
Use Cases	Used to protect data in memory	Used for performing secure computations on sensitive data without decrypting it	Provide long-term data protection and secure communications in the future.

2.1.1 Quantum-Safe Encryption

Quantum-Safe Encryption, also known as post-quantum cryptography, refers to encryption methods that are secure against both classical and quantum computers. As quantum computers advance, they may pose a threat to existing cryptographic systems, potentially compromising their security. Quantum-Safe Encryption (QSE) is essential for protecting sensitive data, communication channels, and user identities in the age of quantum computing.

The urgency of adopting QSE stems from two primary concerns.

- Advanced quantum computers could allow adversaries to intercept and decrypt protected digital communications through Harvest Now, Decrypt Later (HNDL) strategies, even before reaching Q-Day. Q-Day is the anticipated point in time when quantum supremacy becomes widespread and many of the current encryption algorithms are no longer effective.
- Transitioning to QSE might require over a decade due to the complexities of organizational structures and IT infrastructure.

Consequently, organizations should start evaluating and implementing QSE solutions immediately to ensure continued protection and maintain trust among their stakeholders.

Delaying QSE adoption could have severe consequences. Legacy cryptographic systems left unaltered could be compromised in the event of a successful quantum attack, exposing sensitive data and risking confidential business transactions and individual privacy. Financial institutions, critical infrastructure providers, and government agencies would face significant challenges in maintaining operational integrity and confidentiality. Therefore, prioritizing QSE implementation is crucial for long-term cybersecurity resilience.

Power10 is designed to supports these quantum-safe algorithms, ensuring robust security even as quantum computing advances. Power10's support of Quantum-Safe features provide the following.

- Data Encryption Breakage Protection
  - **Risk**: Quantum computers could break widely used cryptographic algorithms such as RSA, ECC (Elliptic Curve Cryptography), and traditional Diffie-Hellman key exchange protocols. Shor's algorithm, for instance, could efficiently factor large integers and solve discrete logarithms, compromising the security of these algorithms.
  - **Protection**: Power10 supports quantum-safe algorithms like lattice-based, hash-based, code-based, and multivariate quadratic equations-based cryptography, which are believed to be resistant to quantum attacks. The crypto engines in Power10 enhance the performance of these algorithms, ensuring secure encryption and key exchange processes with minimal performance degradation.
- Secure Communications
  - **Risk**: Quantum computers could intercept and decrypt secure communications, undermining protocols that currently rely on classical encryption methods.
  - **Protection**: Power10 secures communication channels with quantum-resistant protocols, ensuring data confidentiality even in the presence of quantum adversaries. End-to-end encryption is maintained throughout the data lifecycle, from storage to transmission, using algorithms resistant to quantum attacks.
- Data Integrity and Authenticity
  - **Risk**: Quantum computers could forge digital signatures or tamper with data.
  - **Protection**: Power10 utilizes quantum-safe digital signature algorithms such as XMSS (eXtended Merkle Signature Scheme) and UOV (Unbalanced Oil and Vinegar), providing strong security against quantum attacks. Secure boot processes ensure firmware and software integrity using quantum-resistant cryptographic techniques.

- Long-Term Data Protection
  - **Risk**: Sensitive data stored today could be harvested and decrypted in the future as quantum computers become more powerful, threatening long-term confidentiality.
  - **Protection**: Implementing quantum-safe encryption methods ensures that data remains secure over time, even as quantum computing capabilities evolve. Power10's architecture supports updates to cryptographic libraries and protocols, enabling the adoption of new quantum-safe algorithms as they are developed and standardized.
- Physical and Memory Attack Protection
  - **Risk**: Physical attacks on memory, such as cold boot attacks, could expose sensitive data if not adequately protected.
  - **Protection**: Power10's Transparent Memory Encryption (TME) ensures that data in memory is encrypted, protecting it from physical attacks during runtime and mitigating risks from quantum-assisted physical attacks.

Quantum-Safe Algorithms Supported by Power10:

The following quantum-safe algorithms are supported by Power10:

- Lattice-Based Cryptography
  - **Algorithms**: NTRUEncrypt, Learning With Errors (LWE), and Ring-Learning With Errors (Ring-LWE).
  - **Characteristics**: Secure against quantum attacks based on lattice problems, relatively efficient for hardware and software implementations.
- Hash-Based Cryptography
  - **Algorithms**: Merkle Signature Scheme (MSS), XMSS (eXtended Merkle Signature Scheme), and SPHINCS+.
  - **Characteristics**: Secure based on hash functions, though generally produces larger signatures and keys.
- Code-Based Cryptography
  - **Algorithms**: McEliece Cryptosystem, BIKE (Bit Flipping Key Encapsulation), and HQC (Hamming Quasi-Cyclic).
  - **Characteristics**: Quantum-resistant based on decoding random linear codes, though public keys can be large.
- Multivariate Quadratic Equations
  - **Algorithms**: Unbalanced Oil and Vinegar (UOV), Rainbow.
  - **Characteristics**: Secure against solving systems of multivariate quadratic equations, generally efficient in signature generation but may involve larger key sizes.
- Supersingular Elliptic Curve Isogeny (SIKE)
  - **Characteristics**: Based on finding isogenies between supersingular elliptic curves. Thus approach generally has smaller key sizes but may involve complex computations.

Power10 Implementation

Power10 processors support these quantum-safe algorithms through:

- Crypto Engines: Multiple engines per core enable efficient execution of cryptographic operations.
- Software Updates: The architecture allows updates to cryptographic libraries, ensuring the integration of new quantum-safe algorithms as they become standardized.

Power10's design and capabilities ensure robust security against future quantum threats by leveraging hardware acceleration and flexible software updates, maintaining high-security standards as the cryptographic landscape evolves.

2.1.2 Encryption enablement in hardware

There are two options for accelerating encryption in an IBM Power10 server. The first option is to use the built in encryption acceleration built into the Power10 chip. In addition, IBM Power supports a PCIe based encryption accelerator.

On-chip encryption support in Power10

IBM Power10 chip is designed to effectively support future encryption – including Fully Homomorphic Encryption (FHE) and Quantum-Safe Cryptography – in order to be ready for the Quantum age. The Power10 processor-chip instruction set architecture (ISA) is tailored for these solutions' software libraries, which are currently available or will soon be made available in the corresponding open source communities.

Workloads on the Power10 benefit from cryptographic algorithm acceleration, enabling much higher per-core performance than POWER9 processor-based servers for algorithms like Advanced Encryption Standard (AES), SHA2, and SHA3. Features like AIX Logical Volume Encryption can be activated with minimal performance overhead thanks to this performance enhancement.

With four times as many AES encryption engines, Power10 processor technology is designed to offer noticeably quicker encryption performance. Power10 is more advanced than IBM POWER9 processor-based servers, with updates for the most stringent standards of today as well as future cryptographic standards including post-quantum and fully homomorphic encryption. It also introduces additional improvements to container security. Through the use of hardware features for a seamless user experience, transparent memory encryption aims to simplify encryption and support end-to-end security without compromising performance.

IBM Cryptographic Coprocessor cards

IBM PCIe Cryptographic Coprocessors are a family of high-performance hardware security modules (HSM). These programmable PCIe cards work with certain IBM Z, x64 and IBM Power servers to offload computationally intensive cryptographic processes such as secure payments or transactions from the host server.

These coprocessors enable you to accelerate cryptographic processes that safeguard and secure your data, while protecting against a wide variety of attacks. The IBM 4770, 4769, 4768 and 4767 HSMs deliver security-rich, high-speed cryptographic operations for sensitive business and customer information with the highest level of certification for commercial cryptographic devices.

Cryptographic Coprocessor cards relieve the main processor from cryptographic tasks. The IBM HSMs have a PCIe local-bus-compatible interface and have tamper responding, programmable, cryptographic coprocessors. Each coprocessor contains a CPU, encryption hardware, RAM, persistent memory, hardware random number generator, time-of-day clock, infrastructure firmware, and software. Their specialized hardware performs AES, DES, DES, RSA, ECC, AESKW, HMAC, DES/3DES/AES MAC, SHA-1, SHA-224 to SHA-512, SHA-3, and other cryptographic processes. this relieves the main processor from these tasks. The coprocessor design protects your cryptographic keys and any sensitive customer applications.

Customizable to Meet Special Requirements

The firmware running in the coprocessor together with the software running on your host can be customized to meet any special requirements that your enterprise has. For the IBM 4769 and IBM 4767, the Cryptographic Coprocessor Toolkit (CCTK) is available for purchase from IBM, subject to the export regulations of the United States Government. The CCTK can enable developers to build applications for the HSM, authenticate programs, and load programs into the HSM. The custom programming toolkit includes a custom software interface reference which describes the function calls that applications running in the HSM use to obtain services from the HSM operating system, and from the HSM host system device driver. Another included reference provides the method for extending the CCA host API and the API reference for the user-defined extensions programming environment. Finally, an Interactive Code Analysis Tool (ICAT) is provided that developers can use to debug applications running on the HSM. Frequently a custom contract provides consultation to hasten application development, and sometimes provides for initial development by IBM. Whenever needed, IBM is also able to bid on developing your custom solution or extension.

Secure Administration of HSMs

For the IBM 4769 and IBM 4767, IBM offers GUI-based utilities to administer the HSM cards, including loading of initial keys and setup of the access control system. Each of these can use smart cards as part of the administrative process to carry key parts securely and to identify administrators and allow them to perform sensitive functions. On IBM Power servers running AIX (as well as on Intel x64 systems), the Smart Card Utility Program (SCUP), Cryptographic Hardware Initialization and Maintenance (CHIM), and/or CNM (Cryptographic Node Management - 4767 only) utilities are provided with the HSM software.

The CHIM workstation connects via secure sessions to the cryptographic coprocessors to let authorized personnel perform the following tasks:

– View coprocessor status
– View and manage coprocessor configuration
– Manage coprocessor access control (user roles and profiles)
– Generate and load coprocessor master keys
– Create and load operational key parts

Figure 2-1illustrates the secure management of the IBM HSM cards.

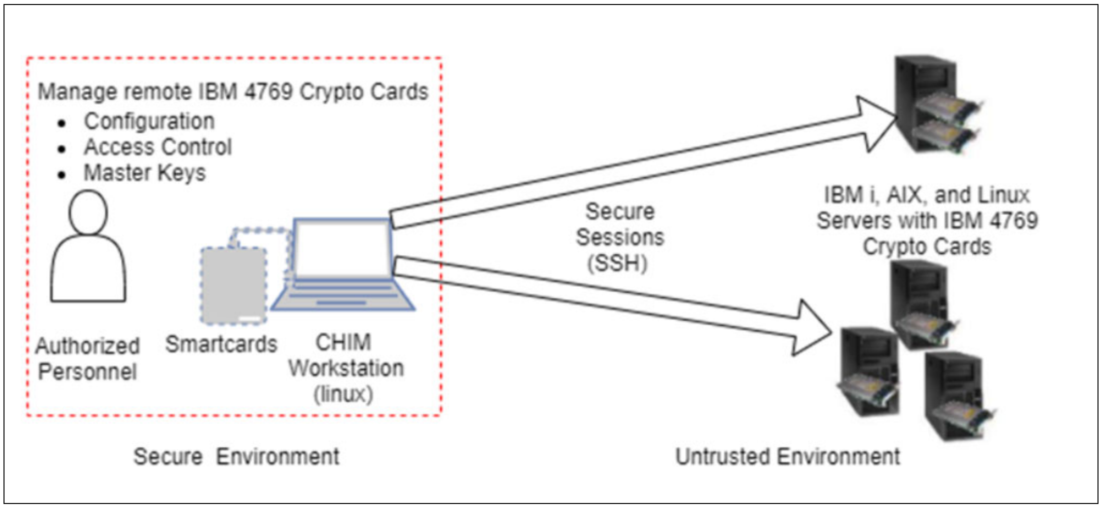


Figure 2-1 Secure management of IBM 4769 Crypto cards

Cryptographic Hardware Initialization and Maintenance on IBM i

CHIM is a PCI-compliant interface to configure the IBM 4769 Cryptographic Coprocessor. CHIM allows you to securely manage remote IBM PCIe Cryptographic Coprocessors with a secure connections even in hostile environment. Management tasks are done using a specialized workstation, the CHIM workstation. CHIM uses smart cards for profile authentication and storage of coprocessor master key parts.

IBM i Common Cryptographic Architecture Cryptographic Service Provider (CCA CSP), delivered as IBM i Option 35, includes the support for IBM Cryptographic Hardware Initialization and Maintenance (CHIM) Catcher. This support is provided with IBM i PTFs

The following are the requirements to use CHIM to manage IBM 4769 Cryptographic Coprocessor(s) located in IBM i systems:

- Installation of the following products (with appropriate PTF levels):
  - 5770SS1 option 35 - CCA Cryptographic Service Provider
  - 5733CY3 - Cryptographic Device Manager
  - 5733SC1 option 1 - OpenSSH, OpenSSL, zlib
- The Secure Shell (SSH) server daemon must be active (use STRTCPSVR *SSHD), must be configured to allow local port forwarding from the CHIM workstation to the CHIM catcher port (which defaults to 50003) on localhost, and must have logging configured for at least the INFO level (the default).
- The CHIM catcher must be active (use STRTCPSVR *CHIM). The CHIM catcher will not start successfully if the previous requirements are not met.
- Cryptographic device descriptions must be created for each IBM 4769 Cryptographic Coprocessor being managed (use CRTDEVCRP) and must be in *ACTIVE status (use VRYCFG or WRKCFGSTS).
- The IBM i user profile used when authenticating from the CHIM workstation must have *IOSYSCFG special authority and have *USE authority for the cryptographic device descriptions for each IBM 4769 Cryptographic Coprocessor being managed.

The CHIM catcher is controlled like all other TCP servers on IBM i. The `STRTCPSVR`, `ENDTCPSVR`, and `CHGTCPSVR` commands can be used to manage the CHIM catcher. The server application value for CHIM is *CHIM. The CHIM catcher port is configured with service name "chim" which is set to port 50003. The CHIM catcher will only listen for incoming connections on localhost. The CHIM catcher will end itself if no server activity occurs for 1 hour.

Smart Cards on Linux

For the 4769 and 4767, IBM provides the SCUP and CHIM applications to manage smart cards with an IBM HSM. SCUP and CHIM run on x64 systems with Linux and can target IBM HSMs installed in x64 and Power systems running AIX. Customers can use SCUP to initialize smart cards that can then be used with CHIM to generate and store CCA master key parts on supported smart cards, load CCA master key parts stored on supported smart cards, and log on to CCA using smart card CCA profiles tied to an RSA key pair associated with a particular smart card and user profile. Smart cards are available for purchase from IBM.

Benefits

IBM PCIe Cryptographic Coprocessors provide you:

- **Improved performance**

  Gain significant performance and architectural advantages and enable future growth by offloading cryptographic processing from the host server.
- **Keep data safe and secure**

  Safeguard data with a tamper-responding design and sensors that protect against module penetration and power or temperature manipulation attacks.
- **Choose your platform**

  Available on select IBM z Systems® servers, on z/OS® or Linux; IBM LinuxONE Emperor, Rockhopper; x64 servers with certain RHEL releases and IBM Power servers.

Note: At the time of this publication, IBM Power supports both the **4769** and **4767** HSMs. The **4769** is currently available, while the **4767** has been withdrawn from marketing.

Features

IBM Cryptographic Coprocessor cards provide the following features:

- High-end secure coprocessors

  Delivers high-speed cryptographic functions for data encryption and digital signing, secure storage of signing keys or custom cryptographic applications.

- Highest level of certification: FIPS PUB 140-2, Level 4

  Validated to FIPS PUB 140-2, Security Requirements for Cryptographic Modules, Overall Security Level 4, the highest level of certification achievable.

- Performance and architectural improvements

  IBM 4769 can exceed 23,000 PIN conversion operations per second, contains custom symmetric key and hashing engines and supports asymmetric algorithms.

- Tamper-responding design

  Sensors protect against a wide variety of attacks on the system and immediately destroy all keys and sensitive data if tampering is detected.

- Common Cryptographic Architecture, Enterprise PKCS #11 APIs

  Performs cryptographic functionality common in the finance industry and business applications, with custom functions available through a programming toolkit.

- Embedded certificate for external verification

  Generates a unique public or private key pair with a certificate that is stored in the device, with safeguards to ensure that the HSM is genuine and untampered.

The remainder of this section covers the 4769 Cryptographic Coprocessor, which is the currently available HSM option for IBM Power.

IBM 4769 Cryptographic HSM Highlights

Each of IBM's HSM devices offer the highest cryptographic security available commercially. Federal Information Processing Standards (FIPS) publication 140-2 defines security requirements for cryptographic modules. It is issued by the U.S. National Institute of Standards and Technology (NIST) and is widely used as a measure of the security of HSMs. The cryptographic processes of each of the IBM HSMs are performed within an enclosure on the HSM that is designed to provide complete physical security.

The 4769 Cryptographic Coprocessor is a PCI Express (PCIe) generation 3 (Gen3) x4 adapter. The secure-key adapter provides both cryptographic coprocessor and cryptographic accelerator functions in a single PCIe card. The 4769 Cryptographic Coprocessor is suited to applications that require high-speed, security-sensitive, RSA acceleration, cryptographic operations for data encryption and digital signing. Additionally, the adapter is useful in secure management, use of cryptographic keys, or custom cryptographic applications. It provides secure storage of cryptographic keys in a tamper-resistant hardware security module that is designed to meet FIPS 140-2 level 4 security requirements. The adapter runs in dedicated mode only.

The IBM 4769 is available as FC EJ35, Customer Card Identification Number (CCIN) C0AF (without blind-swap cassette custom carrier) and as FC EJ37, CCIN C0AF (with blind-swap cassette custom carrier) on IBM Power10 servers, either on IBM AIX, IBM i, or Power Linux (Red Hat Enterprise Linux (RHEL) or SUSE Linux Enterprise Server (SLES)) operating systems. It is also available as FC EJ35 and EJ37 on IBM Power9® servers, either on IBM AIX or IBM i.

An image of the 4769 Cryptographic Coprocessor can be seen in Figure 2-2.

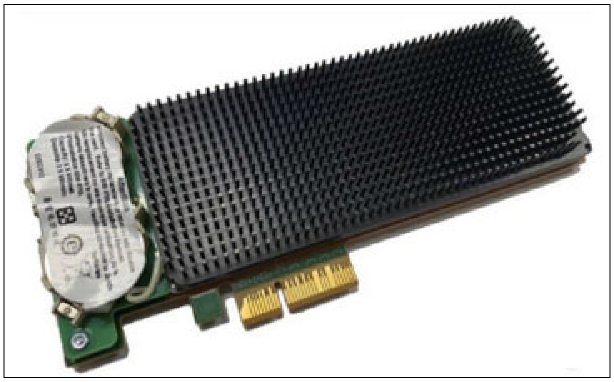


Figure 2-2 4769 Cryptographic Coprocessor

The IBM 4769 hardware provides significant performance improvements over its predecessors while enabling future growth. The secure module contains redundant IBM PowerPC® 476 processors, custom symmetric key and hashing engines to perform AES, DES, TDES, SHA-1 and SHA- 2, MD5 and HMAC as well as public key cryptographic algorithm support for RSA and Elliptic Curve Cryptography. Other hardware support includes a secure real-time clock, hardware random number generator and a prime number generator. The secure module is protected by a tamper responding design that protects against a wide variety of attacks against the system.

IBM CEX7S / 4769

The IBM 4769 is validated by NIST to FIPS 140-2 Level 4, the highest level of certification achievable for commercial cryptographic devices.[2] FIPS 140 defines security requirements for cryptographic modules. It is issued by the U.S. National Institute of Standards and Technology (NIST) and is widely used as a measure of the security of HSMs.

The “Payment Card Industry Hardware Security Module” standard, PCI HSM, is issued by the PCI Security Standards Council. It defines physical and logical security requirements for HSMs that are used in the finance industry. The IBM CEX7S with CCA 7.x has PCI HSM certification.[3]

The following attributes are provided by the 4769:

- Supported cryptographic mode: Common Cryptographic Architecture (CCA)
- PPC 476 Processors run in lockstep and the outputs of each core are compared cycle by cycle
- Error Checking and Correction (ECC) protection on DDR3 memory
- Cryptographic key generation and random number generation
- Over 300 cryptographic algorithms and modes
- Byte wide parity protection on all internal registers and data paths wider than two bits
- RSA/ECC engines are protected by a duplicate engine which predicts the CRC of the result
- SHA, MD5, AES and DES engines are protected by running the same operation on two independent engines and the outputs are compared cycle by cycle.

[2] https://csrc.nist.gov/projects/cryptographic-module-validation-program/certificate/4079

[3] https://listings.pcisecuritystandards.org/popups/pts_device.php?appnum=4-20358

The IBM 4769 is designed for improved performance and security rich services for sensitive workloads, and to deliver high throughput for cryptographic functions. For a detailed summary of the capabilities and specifications of the IBM 4769, refer to the IBM 4769 Data Sheet.

- Reliability, Availability, and Serviceability (RAS)

  Hardware has also been designed to support the highest level of RAS requirements that enables the secure module to self-check at all times. This is achieved by running a pair of PowerPC processors in lock step and comparing the result from each cycle by cycle. Also all interfaces, registers, memory, cryptographic engines, and buses are protected at all times using parity, ECC (Error Correcting Codes), or CRC. Power on self-tests that are securely stored inside the secure module verify the hardware and firmware loaded on the module is secure and reliable at every power on. Then, the built-in RAS features check it continuously in real time.

- Embedded Certificate

  During the final manufacturing step, the coprocessor generates a unique public/private key pair which is stored in the device. The tamper detection circuitry is activated at this time and remains active throughout the useful life of the coprocessor, protecting this private key as well as other keys and sensitive data. The public key of the coprocessor is certified at the factory by an IBM private key and the certificate is retained in the coprocessor. Subsequently, the private key of the coprocessor is used to sign the coprocessor status responses which, in conjunction with a series of public key certificates, demonstrate that the coprocessor remains intact and is genuine.

- Tamper Responding Design

  The IBM 4769 HSM is validated by NIST to meet the FIPS 140-2 Level 4 requirements by protecting against attacks that include probe penetration or other intrusion into the secure module, side-channel attacks, power manipulation, and temperature manipulation. From the time of manufacture, the hardware is self-protecting by using tamper sensors to detect probing or drilling attempts. If the tamper sensors are triggered, the HSM destroys critical keys and certificates, and is rendered permanently inoperable. Note therefore that the HSM must be maintained at all times within the temperature, humidity, and barometric pressure ranges specified.

2.2 Compliance automation and real-time monitoring

IBM Power10 systems offer tools to automate compliance tasks and monitor your IT environment in real-time. This helps you meet regulations and stay secure. You can use IBM PowerSC for automated compliance management. For more details see 9.1, “Compliance automation” on page 220.

PowerVM, the virtualization management tool for IBM Power, provides real-time monitoring of virtualized environments. It helps track performance, resource utilization, and security status across virtual machines and physical servers. In addition the Hardware Management Console (HMC) offers real-time monitoring and management of IBM Power systems. It provides insights into system health, performance metrics, and potential security issues.

Power10 systems can be configured to generate real-time alerts for various events, including security incidents, system performance issues, and hardware faults. These alerts can be integrated with enterprise monitoring solutions for centralized management. Integration with Security Information and Event Management (SIEM) systems allows for real-time analysis of security events and incidents. This helps in detecting and responding to potential threats as they occur.

IBM Power10 systems provide a robust framework for compliance automation and real-time monitoring through integrated tools and features. By leveraging solutions like IBM PowerSC, advanced monitoring tools, and real-time alert systems, organizations can ensure continuous compliance with security standards, automate policy enforcement, and monitor system performance and security in real time. This combination of capabilities helps maintain a secure and compliant IT environment, reducing risks and enhancing operational efficiency.

2.3 Endpoint detection and response

Endpoint detection and response (EDR) is a system that monitors and analyzes security threats from endpoints, such as computers and mobile devices. It uses machine learning and analytics to identify patterns that indicate suspicious activity or known threats in real time. The goal of EDR is to find security breaches as they happen and respond quickly to potential or discovered threats. An endpoint detection and response (EDR) solution proactively and automatically blocks and isolates malware while equipping security teams with the right tools to confidently deal with these challenges. A modern EDR can ensure business continuity by effectively mitigating fast-growing, automated and advanced threats, such as ransomware or other attacks, without increasing analyst workloads or requiring highly skilled security specialists.

Here are some options within the IBM Power ecosystem to support EDR:

- IBM PowerSC is a security and compliance solution optimized for virtualized environments on IBM Power servers running AIX, IBM i or Linux. PowerSC sits on top of the IBM Power server stack, integrating security features built at different layers. You can now centrally manage security and compliance on Power for all IBM AIX and Linux on Power endpoints.

  For more information see 9.3, “Endpoint Detection and Response” on page 221.

- IBM Security® QRadar® EDR remediates known and unknown endpoint threats in near real time with easy-to-use intelligent automation that requires little-to-no human interaction.

  For more information see “IBM QRadar Suite (Palo Alto Networks)” on page 250.

2.4 Malware prevention technologies

One of the security exploits that has been identified is an exploit of Return Oriented Programming (ROP). ROP is a type of software attack technique that exploits vulnerabilities in programs that do not properly validate user-supplied input. By carefully crafting input data, attackers can manipulate the program's control flow to execute arbitrary code, often leading to severe consequences like remote code execution.

The Power10 processor architecture incorporates several features designed to enhance control flow security and mitigate the risk of ROP attacks. These features include improved hardware-based encryption and advanced protection against side-channel attacks. While these features can mitigate some attacks, they do not make ROP attacks impossible, but rather more challenging.

Modern compilers and operating systems for Power10 can include additional security features and mitigations. Developers should ensure that their software is built with the latest security practices and that the operating system is up-to-date with relevant patches.

2.5 Trusted boot

Trusted Boot (TB), also known as Secure Boot, is a security process that ensures a system boots using only software that is trusted by the manufacturer. It prevents unauthorized and potentially malicious software from loading during the boot process, which protects against various types of attacks and ensures the integrity of the system from the moment it starts.

TB helps organizations meet security compliance requirements by ensuring that their systems are protected from unauthorized software and firmware changes.

Here are the key components of any Trusted Boot process:

- Boot Process Integrity

  TB verifies the integrity and authenticity of each component in the boot process, from the firmware and boot loader to the operating system kernel. It ensures that only trusted, signed software components are executed.

- Cryptographic Verification

  During TB, each stage of the boot process is cryptographically verified before it is executed. This involves checking digital signatures against known, trusted keys stored in secure hardware or firmware.

- Root of Trust

  TB relies on a Root of Trust, which is a secure, hardware-based mechanism that serves as the foundation for all other trust decisions. This is typically implemented in a Trusted Platform Module (TPM) or similar secure hardware component.

- Chain of Trust

  The verification process creates a Chain of Trust, where each successive component in the boot sequence verifies the integrity of the next component before passing control to it. This ensures that the entire boot process is secure and unmodified.

- Detection and Prevention of Unauthorized Changes

  TB helps detect and prevent unauthorized changes to the system's firmware, boot loader, and operating system. If an untrusted component is detected, the boot process is halted, and the system may alert the user or administrator of the security issue.

Power10 and Trusted Boot

Power10's hypervisor is called PowerVM; here is a flow of how PowerVM ensures trusted boot during the boot process.

1. Root of Trust Initialization

   When the system powers on, the Trusted Platform Module (TPM) initializes, and it measures and records the initial state of the firmware.

2. Firmware Verification

   System firmware is stored in a secure location and contains digital signatures that are verified against trusted keys stored in the TPM, to ensure that the firmware has not been tampered with.

3. Hypervisor Verification

   After the firmware is verified, the PowerVM hypervisor (PHYP) is loaded. The PHYP is signed, and its integrity is checked against the trusted keys. The TPM logs the measurement of the hypervisor, to detect any modifications.

4. Boot loader and Operating System Verification

   Once the hypervisor is loaded and running, it loads the boot loader and the operating system for each logical partition (LPAR), each of which is verified against their respective digital signatures. The TPM continues to log these measurements, maintaining a chain of trust from the firmware to the hypervisor and then to the OS. This ensures that each LPAR starts in a secure state.

5. Continuous Monitoring and Attestation

   Continuous integrity checks may be performed during runtime to detect any deviations from the trusted state.

2.5.1 Quantum safe boot

Quantum-Safe Boot (QSB) is a related concept to Trusted Boot (TB), but it addresses a different aspect of system security. While both concepts aim to enhance system security, TB focuses on ensuring the initial integrity of the system boot process and components using hardware-based mechanisms, whereas QSB prepares systems to withstand potential future cryptographic threats posed by quantum computing. They complement each other by addressing different layers of security concerns in computing environments, rather than being logical equivalents.

TB focuses on establishing a chain of trust from hardware to software. It ensures the integrity and authenticity of the boot process and critical system components from the initialization of the hardware through the completion of the operating system startup process. It relies on hardware-based security like a Trusted Platform Module (TPM) to measure and verify the integrity of firmware, boot-loaders, and operating system components.

QSB has the same goal as TB, but rather than focusing on hardware encryption mechanisms, QSB implements post-quantum cryptographic (PQC) algorithms resistant to attacks from quantum computers. Such algorithms include lattice-based cryptography, code-based cryptography, and hash-based cryptography.

2.6 Hypervisor security

Hypervisors, as the backbone of virtualized environments, require robust security measures to protect against various threats and vulnerabilities that can compromise the entire infrastructure. This section delves into the security aspects of hypervisors, detailing the potential attack vectors, security best practices, and the latest technologies for safeguarding these critical components.

Hypervisor vulnerabilities

The first step in securing hypervisors is understanding the unique vulnerabilities they face. The following are some vulnerabilities and a discussion of how to avoid them in your IBM Power environment.

Hyperjacking

Hyperjacking is a type of advanced cyber attack where threat actors take control of a crucial element called the hypervisor, which handles the virtualized environment within a main computer system. Their ultimate aim is to deceive the hypervisor into executing unauthorized tasks without leaving traces elsewhere in the computer.

Hyperjacking involves exploiting a vulnerability in a user's browser extension or add-on to gain control over their computer without their knowledge. By hijacking the user session, cybercriminals can conduct surveillance, manipulate connected devices remotely, and potentially steal sensitive information.

VM Escape

VM escape is a security vulnerability that lets attackers breach virtual machines and obtain unauthorized access to the underlying physical hardware, such as the hypervisor or host system. It circumvents the virtualization layer's isolation barriers, enabling potential exploits.

IBM Power and PowerVM provide many options to help preventing a VM Escape type attack. PowerVM has excellent LPAR isolation that prevents an LPAR from seeing resources out of its defined VM. Additionally, Power10 servers have memory encryption enabled that will further protect from exposure of data to another LPAR.

Resource Exhaustion

Attackers can target the resource allocation features of a hypervisor, leading to denial of service (DoS) by exhausting resources such as CPU and memory, which affects all VMs hosted on the hypervisor.

Within PowerVM, when a resource is defined to an LPAR there are limits that are enforced by the hypervisor to protect from overallocation. Defining the minimum, maximum and requested values for memory and CPU correctly will ensure that you will avoid resource exhaustion attacks.

Ensure that the administrator credentials used for access to PowerVM are protected and use RBAC to limit the scope of changes that can be made.

Protecting from hypervisor vulnerabilities

To mitigate the risks associated with hypervisor vulnerabilities, use the same strategies that are used to protect other components in your environment:

- Turning on Trusted boot.

  IBM’s PowerVM hypervisor is built into the firmware on the server. This, along with capabilities like Trusted Boot (see section 2.5, “Trusted boot” on page 39) protect you from having a bad actor insert a false hypervisor into the system.

- Protecting access to PowerVM

  Most hypervisor intrusions are due to gaining access to the hypervisor using exposed credentials. Implement good authentication control and define users in PowerVM using RBAC to limit their access to resources.

- Protecting access to the VMs

  Ensure that you follow best practices for protecting each VM. Manage credentials, use multi-factor authentication, and define user credentials with the minimum capabilities that they need to do their job.

- Plan for isolation of networks and storage

  Protect your VMs by isolating network traffic between VMs to prevent eavesdropping and network attacks. Techniques such as VLANs, firewall rules, and virtual network appliances can be used.

  Isolate storage access among VMs to prevent data leakage or corruption. This includes using separate storage accounts for sensitive data and implementing robust access controls. Use encryption to avoid improper access to storage.

– Maintain currency in your firmware

  Be aware the availability of new firmware versions and schedule any updates on a regular basis.

– Maintain currency in your Virtual Machine operating systems

  Plan to monitor updates available for the operating systems in your VMs and schedule any updates on a regular basis.

For more information on securing your PowerVM environment see section and 3.1, “Hardware Management Console Security” on page 44 and section 3.3, “VIOS Security” on page 63

2.6.1 LPAR Isolation

Logical Partitioning (LPAR) is a technology used primarily in enterprise computing environments to divide a computer's resources, such as CPU, memory, and storage, into multiple, separate virtual machines. Each LPAR operates as a standalone environment with its own operating system and applications, making it an invaluable tool for optimizing resource use, improving system security, and increasing availability in power systems.

LPAR isolation is a basic tenet of IBM PowerVM, the IBM Power hypervisor. PowerVM is designed to share resources from a single machine across all of the LPARs defined on that machine. Additionally, PowerVM has additional capabilities that allow LPARs to be non-disruptively moved from one host machine to another to provide load balancing and high availability configurations. LPAR restart technologies also support disaster recovery options to restart workloads in another site in case of site failure.

One of the strengths of PowerVM is its flexibility in sharing processing resources. CPUs can be defined to an LPAR as dedicated or shared, capped or uncapped and donating or not donating. This allows you to effectively allocate resource among LPARs, ensuring that each partition receives the necessary resources to perform optimally without affecting the performance of others. This includes dynamic resource allocation techniques that allow resources to be reallocated based on workload demands.

3

Security in the Virtualization and Management Layer

Virtualization is a keystone of the IBM Power ecosystem, allowing clients to take full advantage of the performance, reliability and security built into IBM Power servers while reducing the cost and complexity of running isolated workloads on dedicated machines. Virtualization by definition involves sharing hardware infrastructure across multiple workloads. This provides significant benefits but it also provides additional challenges in keeping different workloads isolated and secure.

The task of managing the virtualization layer in the IBM Power ecosystem is divided into two distinct areas: hardware management, and I/O virtualization. The hardware management aspect is handled by the Hardware Management Console (HMC) which is an appliance designed to define the logical partitions in each server, dividing and sharing the installed resources across the various virtual machines that are supported. An HMC can manage multiple servers, but as your infrastructure grows across multiple locations and large number of servers the Cloud Management Console provides a solution to having a single tool for consolidating information across a number of HMCs.

The Virtual I/O Server (VIOS) is a special partition running in an IBM Power server that allows sharing of physical devices across multiple LPARs. The purpose of the VIOS is to virtualize the physical adapters in the system, reducing the number of adapters to be reduced. Systems with virtualized i/O can also be more easily moved to other servers as needed for load balancing and high availability during planned or unplanned outages—providing a more available and resilient environment.

Each of these functions: HMC, CMC and VIOS are discussed in this chapter:

- 3.1, “Hardware Management Console Security” on page 44
- 3.2, “Cloud Management Console security” on page 55
- 3.3, “VIOS Security” on page 63

3.1 Hardware Management Console Security

The Hardware Management Console (HMC) is a specialized device designed for configuring and managing IBM Power systems. It facilitates basic virtualization management by supporting the setup of logical partitions (LPARs) and dynamic resource allocation, including adjustments to processor and memory settings for IBM Power servers. Additionally, the HMC offers advanced service functions such as guided repair and verification, concurrent firmware updates for managed systems, and continuous error reporting through the Electronic Service Agent for expedited support. The latest model, the 7063-CR2, operates on an IBM Power9 server. This dedicated device is exclusively used for controlling and servicing IBM Power servers and cannot be utilized as a general-purpose computing resource.

HMC Packaging

Initially, the HMC was delivered solely as a traditional hardware appliance, with the software and hardware bundled together and installed on-site. As client environments grew, there was a demand to virtualize the HMC function to minimize infrastructure needs. In response, IBM introduced the virtual HMC (vHMC), which enables you to use your own hardware and server virtualization to host the IBM-provided HMC virtual appliance. The vHMC image is available for both x86 and IBM Power servers and supports the following hypervisors:

- For x86 virtualization:
  - Kernel-based Virtual Machine (KVM) on Ubuntu 18.04 LTS or Red Hat Enterprise Linux 8.0 or 9.0
  - Xen on SUSE Linux Enterprise Server 12
  - VMware ESXi 6.5, 7.0 or 7.0.2
- For Power virtualization:
  - PowerVM

The distribution of HMC service packs and fixes is consistent for both hardware and virtual HMCs. However, for vHMC on PowerVM, Power Firmware updates are managed by IBM. For vHMC on x86 systems, if security vulnerabilities arise, you should consult with the hypervisor and x86 system vendors for any necessary updates to the hypervisor and firmware. The steps for enabling secure boot differ between hardware and virtual HMCs due to architectural differences. For detailed instructions on enabling the secure boot function, refer to section 3.1.9, “Secure boot” on page 52.

For additional information on the vHMC see:
https://www.ibm.com/support/pages/virtual-hmc-appliance-vhmc-overview

HMC Functions

The HMC enables you to create and manage logical partitions, including the ability to dynamically add or remove resources from active partitions. It also handles advanced virtualization functions such as Capacity Upgrade on Demand and Power Enterprise Pools.

In addition, the HMC provides terminal emulation for the logical partitions on your managed systems. You can connect directly to these partitions from the HMC or configure it for remote access. This terminal emulation feature ensures a reliable connection, useful if other terminal devices are unavailable or not operational. It is particularly valuable during the initial system setup, before configuring your preferred terminal.

Using its service applications, the HMC communicates with managed systems to detect, consolidate, and relay information to service and support teams for analysis. For a visual representation of how the HMC integrates into the management and serviceability of IBM Power systems, refer to Figure 3-1.

3

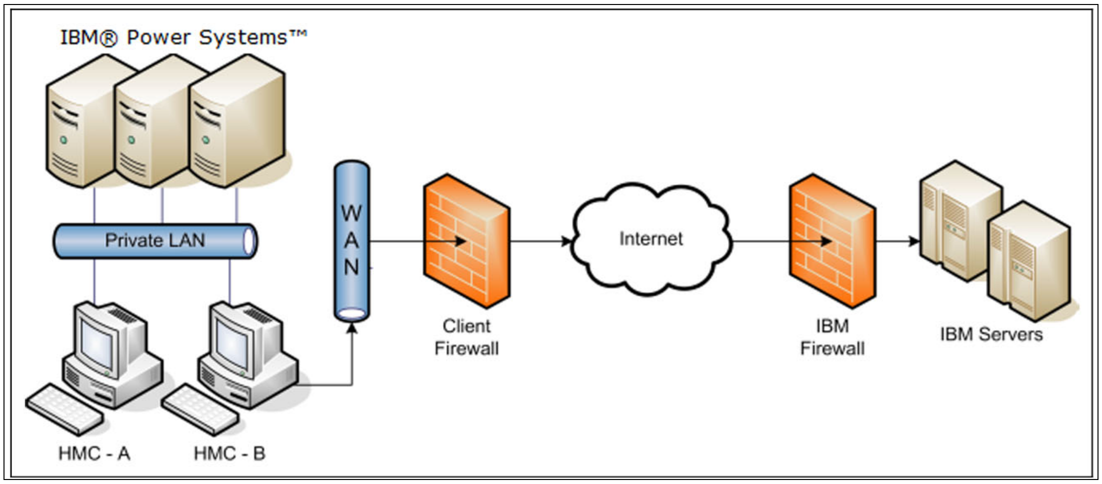


Figure 3-1 HMC used for configuration and serviceability functions

One HMC can oversee multiple servers and multiple HMCs can connect to a single server. If a single HMC fails or loses connection to the server firmware, the server will continue to operate normally, but changes to the logical partition configuration won't be possible. To mitigate this, you can connect an additional HMC as a backup to ensure a redundant path between the server and service and support.

Each HMC comes preinstalled with the HMC Licensed Machine Code to ensure consistent functionality. You have several options for configuring HMCs to provide flexibility and availability:

- Local HMC

  A local HMC is situated physically close to the system it manages and connects via a private or public network. In a private network setup, the HMC acts as a DHCP server for the system’s service processors. Alternatively, it can manage the system over an open network, where the service processor’s IP address is manually assigned through the Advanced System Management Interface (ASMI).

- Remote HMC

  A remote HMC is located away from its managed systems, which could be in a different room, building, or even a separate site. Typically, a remote HMC connects to its managed servers over a public network, although it can also be configured to connect via a private network.

IBM has created a document that provides a starting point on understanding the connectivity used by the HMC and how to make it secure. The HMC 1060 Connectivity Security White Paper is a good starting point for enabling a secure HMC environment in your enterprise.

3.1.1 Security levels for the HMC

The HMC offers REST, GUI, and CLI interfaces for user interaction. Security requirements vary depending on the model and the interface used. The HMC provides recommendations for different levels of security, which are outlined in this section.

Level 1

Level 1 defines the security actions that you must have. These are the minimum measures recommended to secure your HMCs:

- Change the default password of *hscroot* user.
- Enable grub password if your HMC is not in a physically secure environment.

  ```
  chhmc -c grubpasswd -s enable --passwd <new grub password>
  ```
- If you have configured Integrated Managemet Module (IMM) on HMC, set strong IMM password.
- Set strong password for admin and general users on all servers.
- Keep HMC updated with all released security fixes. Fixes are available in Fix Central.

Level 2

Level 2 defines some actions that you should consider when you have multiple HMC users defined in the environment. If you have multiple users defined to use the HMC then consider:

- HMC supports fine-grained control on resources and roles. Create account for each user on HMC.
- Assign only necessary roles to users.
- Assign only necessary resources (Systems, Partitions, etc.) to users.
- Both resources and roles assigned to the users must be minimum that's required for doing the job. Create custom roles if necessary.
- Enable user data replication between HMCs with different modes.
- Import a certificate signed by Certificate Authority.
- Enable secure boot.
- Enable Multi Factor Authentication.
- Enable PowerSC profile.

Level 3

Level 3 defines additional considerations when you have multiple HMCs in the environment. If you have many HMCs and Sysadmins:

- Use centralized authentication—LDAP or Kerberos (HMC does not support SSO feature).
- Enable user data replication between HMCs.
- Put HMC in NIST SP 800-131A mode so that it uses strong ciphers only.
- Block unnecessary ports in firewall.

For more information see this IBM Document on HMC Security.

3.1.2 Port security

The HMC is primarily a Java-based application running on Linux, utilizing various open-source components. It communicates with different services through various ports, all of which are encrypted using TLS 1.2 or higher. For remote access, it is recommended to only expose the following ports:

- SSH (port 22),
- HTTPS (port 443)
- VTerm (port 9960)

All other ports should be kept within a private or isolated network for security purposes.

3.1.3 Securing connections to Power systems

The HMC connects to managed systems via an integrated service adapter built into the system. Depending on the model of the Power system, this adapter may be an FSP (Field Service Processor) or an eBMC (Enterprise Baseboard Management Controller). Connectivity varies slightly based on the endpoint:

- For FSP and Hypervisor: Management uses a proprietary binary protocol known as NETC, with communication encrypted using TLS 1.2 or higher.
- For eBMC: The connection is established through the Redfish REST API, with encrypted communication supported by TLS 1.2 or higher and appropriate certificates.

3.1.4 NIST SP 800-131A compliance mode

HMC can operate in two modes: legacy and nist_sp800_131a. Once you set HMC in nist_sp800_131a mode, only strong ciphers listed by NIST SP 800-131A will be used.

Use this command on the HMC to set the NIST SP 800-131A mode:

```
chhmc -c security -s modify --mode nist_sp800_131a
```

If you wish to return the HMC to legacy mode, use this command:

```
chhmc -c security -s modify --mode legacy
```

3.1.5 Encryption

All communication channels utilized by the HMC are encrypted. By default, the HMC employs Transport Layer Security (TLS) and HTTPS with secure cipher sets bundled with the HMC. The default ciphers provide strong encryption and are used for secure communication on ports 443, 17443, 2301, and 5250 proxy, as well as for internal HMC communication.

> **Note:** For details on the encryption ciphers used by the HMC, you can execute the `lshmcencr` command in the HMC command-line interface (CLI). If your organization's corporate standards require different ciphers, you must use the **`chhmcencr`** command to modify them. For more details see:
> https://www.ibm.com/docs/en/power10/000V-HMC?topic=sh-managing-https-ciphers-hmc-web-interface-by-using-hmc

The HMC supports both self-signed and CA-signed certificates for encryption. Starting with HMC Version 10.2.1040.0 and later, you can select the key size for certificates when generating a certificate signing request (CSR), with options of 2048, 3072, or 4096 bits. When using CA-signed certificates, ensure a minimum of 2048-bit RSA encryption is employed. By default, the HMC uses a self-signed certificate with the SHA256 algorithm and 2048-bit RSA encryption.

3.1.6 Certificate management

Security certificates are crucial for ensuring that the HMC operates securely in client/server mode, where the managed machines act as servers and the managed users are clients. Communication between the server and client occurs over the Secure Sockets Layer (SSL) protocol, which provides server authentication, data encryption, and data integrity.

When a user seeks remote access to the HMC user interface via a web browser, they initiate a request for the secure page using *https://hmc_hostname*. The HMC then presents its certificate to the remote client (web browser) during the connection process. The browser verifies the certificate by checking that it was issued by a trusted authority, that it is still valid, and that it was specifically issued for that HMC.

3.1.7 User management

On an HMC, a user can be a member of various task roles. Each task role allows the user to access different parts of the HMC and to perform different tasks on the managed system.

HMC task roles are either predefined or customized. When you create an HMC user, you must assign a task role to that user. Each task role allows the user varying levels of access to tasks that are available on the HMC interface. You can assign managed systems and logical partitions to individual HMC users, allowing you to create a user that has access to managed system A but not to managed system B. Each grouping of managed resource access is called a managed resource role. Table 3-1 lists the predefined HMC task roles, which are the default on the HMC.

Table 3-1   HMC predefined roles

Task Role	Description
hmcservicerep	A service representative is an employee who is at your location to install, configure, or repair the system.
hmcviewer	A viewer can view HMC information, but cannot change any configuration Information.
hmcoperator	The operator is responsible for daily system operation.
hmcpe	A product engineer assists in support situations but cannot access HMC user management functions. To provide support access for your system, you must create and administer user IDs with the product engineer role.
hmcsuperadmin	The super administrator acts as the root user or manager of the HMC system. The super administrator has unrestricted authority to access and modify most of the HMC system.

You can create customized HMC task roles by modifying predefined HMC task roles. Creating customized HMC Task Roles is useful for restricting or granting specific task privileges to a certain user.

Authentication

User authentication is the first step to protecting your HMC and ensuring that only authorized users are access the management console. The HMC supports various authentication methods to validate users:

- Local Authentication

  If you select local authentication then password and number of days the password is valid are required to be set.

- Kerberos Authentication

  If you select Kerberos Authentication, specify a Kerberos remote user ID and configure HMC to use Kerberos. When a user logs in to the HMC, authentication is first verifies against a local password file. If a local password file is not found, the HMC can contact a remote Kerberos server for authentication. You must configure your HMC so that it uses Kerberos remote authentication.

  Information on setting up kerberos can be found in this IBM Document Link.

- LDAP Authentication

  If you select LDAP Authentication, configure HMC to use LDAP server. Information on configuring HMC for LDAP can be found in this IBM Documentation link.

  - Automanage Authentication

    Use this option to indicate whether the HMC should automatically manage remotely authenticated LDAP users. Valid values are 0 to disable automatic management, or 1 to enable automatic management.

    When automatic management is enabled, an LDAP user can log in to the HMC. An HMC user will be automatically created for the LDAP user if the HMC user does not exist when the LDAP user logs in. If the HMC user already exists, it will be updated with the current user definition retrieved from the LDAP server when the LDAP user logs in.

User properties

The User Properties has the following properties that you can set:

- Timeout Values

  These values specify values for various timeout situations.

  - Session timeout minutes

    Specifies the number of minutes, during a logon session, that a user is prompted for identity verification. If a password is not re-entered within the amount of time that was specified in the Verify timeout minutes field, then the session is disconnected. A zero (0) is the default and indicates no expiration. You can specify up to a maximum value of 525600 minutes (equivalent to one year).

  - Verify timeout minutes

    Specifies the amount of time that is required for the user to re-enter a password when prompted, if a value was specified in the Session timeout minutes field. If the password is not re-entered within the specified time, the session will be disconnected. A zero (0) indicates there is no expiration. The default is 15 minutes. You can specify up to a maximum value of 525600 minutes (equivalent to one year).

  - Idle timeout minutes

    Specifies the number of minutes the user’s session can be idle. If the user does not interact with the session in the specified amount of time, the session becomes disconnected. A zero (0) is the default and indicates no expiration. You can specify up to a maximum value of 525600 minutes (equivalent to one year).

  - Minimum time in days between password changes

    Specifies the minimum amount of time in days that must elapse between changes for the user’s password. A zero (0) indicates that a user’s password can be changed at any time.

- Inactivity Values

  These define what actions to take due to various periods of inactivity:

  - Disable for inactivity in days

    This value defines the number of days of inactivity after which a user is temporarily disabled. A value of zero (0) means that the user will not be disabled regardless of the duration of inactivity.

  - Never disable for inactivity

    If you do not want to not disable user access based on inactivity, select “**Never disable for inactivity”**.

  - Allow remote access using the web

    To enable remote web server access for the user you are managing, select "**Allow remote access via web.**" If this option is not selected, the user will only have local access to the HMC or access via the command line using an SSH session.

- User Lockout

  The user will be locked out of the HMC after a specified number of invalid login attempts through any interface, including CLI, GUI, or REST. By default, the system is set to lock out the user after 3 failed attempts, with a lockout period of 5 minutes. You can configure these lockout parameters using this command:

  ```
  chhmcusr -t default -i "max_login_attempts=3,login_suspend_time=5"
  ```

Password policy

The HMC ships with default password policies which can be used to meet general corporate requirements. To meet specific requirements, users can create a custom password policy and apply it on the HMC. Password policies are enforced for locally authenticated HMC users only.

To see what password policies are defined on the HMC use the **lspwdpolicy**[1] command as follows:

- List all of the HMC password policies:

  ```
  lspwdpolicy -t p
  ```

- List just the names of all of the HMC password policies:

  ```
  lspwdpolicy -t p -F name
  ```

- List HMC password policy status information:

  ```
  lspwdpolicy -t s
  ```

The "HMC Medium Security Password Policy" is defined by default but not activated. It has the following settings:

- min_pwage=1
- pwage=180
- min_length=8
- hist_size=10
- warn_pwage=7
- min_digits=0
- min_uppercase_chars=1
- min_lowercase_chars=6
- min_special_chars=0
- inactivity_expiration=180

The policy can be activated with the **chpwpolicy**[2] command: **chpwdpolicy -o -n "HMC Medium Security Password Policy".** It can be deactivated by: **chpwdpolicy -o d**. If you deactivate a password policy, be sure to activate another policy to protect your system.

An additional defined policy, "HMC Standard Security Password Policy", is also available and might be acceptable for use depending on your corporate requirements. Its setting are defined as:

- min_lowercase_chars=1
- min_uppercase_chars=1
- min_digits=1

[1] https://www.ibm.com/docs/en/power10/7063-CR1?topic=commands-lspwdpolicy
[2] https://www.ibm.com/docs/en/power10/7063-CR1?topic=commands-chpwdpolicy

- min_special_chars=1
- pwage=90
- min_length=15

If you wish to create your own policy, use the **mkpwpolicy**[3] command. One example of creating a new password policy is shown in Example 3-1.

Example 3-1 Make password policy example

```
mkpwdpolicy -i "name=xyzPolicy,description=Company xyz policy,
pwage=90,min_digits=2,min_uppercase_chars=0,min_lowercase_chars=0"
```

The “-i” flag shown uses the command line input to define the parameters of the policy. Using the “-f” flag allows the use of a file with the parameters defined to simplify the entry of the command and to provide consistency across your HMCs. Reminder, once the policy is defined it still needs to be activated before it is effective.

Deleting password policies is done using the **rmpwdpolicy** command. The single parameter is “-n” to specify the name to be deleted. For example **rmpwdpolicy -n xyzPolicy** would delete the policy “xyzPolicy”.

WebUI Session limit

The HMC provides a feature to limit the number of concurrent web user interface (webui) logins for each user. As there is a maximum number of concurrent webui sessions supported by the HMC, limiting the number of webui logins per user will help avoid user lockouts due to exceeding the maximum webui session limit and potential denial of service attacks. Below are description of the attributes that can be used to configure these session limits.

- Maximum webui sessions per user: Specify the maximum number of web user interface sessions that are allowed for a logged in user. By default, 100 web user interface sessions are allowed for a user. The value for maximum webui sessions ranges between 50-200.

  The command below can be used to set the session limit per user:

  ```
  chhmcusr -t default -i “max_webui_sessions_per_user =50”
  ```

- Console maximum webui session: Specifies the maximum number of web user interface sessions that are allowed in the HMC. By default, 1000 web user interface sessions are allowed in the HMC. This is a currently a read-only parameter that cannot be modified.

Enabling MFA

Multi-Factor Authentication is disabled on the HMC by default. For HMC GUI login, when MFA is enabled and the user is configured on the PowerSC MFA server, enter the Cache Token Credential (CTC) code in the password field. For Secure Shell (SSH) login, when MFA is enabled, all users that login through SSH are prompted for a CTC code. If the user is configured on the PowerSC MFA server, then you can enter the CTC code at the prompt. If the user is not configured on the PowerSC MFA server, press **Enter** when prompted for CTC code, and then enter the password of the user at the prompt. When HMC is enabled with Power SC MFA, all the users including local, LDAP, Kerberos will be prompted with PowerSC MFA authentication process. If for any reason you don’t want MFA for certain user, HMC is enabled with an PowerSC MFA allow list for users exempt from MFA. The users which are added for allow list will be exempted from PowerSC MFA authentication on HMC.

[3] https://www.ibm.com/docs/en/power10/7063-CR1?topic=commands-mkpwdpolicy

3.1.8 Auditing capabilities of the HMC

A secure system also requires strong auditing capabilities. This section describes some of the logging and auditing functions on the HMC.

Most tasks performed on the HMC (either locally or remotely) are logged by the HMC. These entries can be viewed by using the Console Events Log task, under **Serviceability** → **Console Events Log** or by using the **`lssvcevents`** command from the restricted shell.

A log entry contains the time stamp, the user name, and the task being performed. When a user logs in to the HMC locally or from a remote client, entries are also recorded. For remote login, the client host name or IP address is also captured, as in Example 3-2.

Example 3-2 User login entry

```
lssvcevents -t console
time=11/11/2015 09:52:55,"text=User hscroot has logged on from location <ip
address> to session id 32. The user's maximum role is ""hmcsuperadmin""."
```

Standard log entries from *syslogd* can also be seen on the HMC by viewing the */var/hsc/log* file. This file can be read by users with the *hmcsuperadmin* role. It is under *logrotate* control. A valid user can simply use the **`cat`** or **`tail`** command to view the file.

A user with the *hmcsuperadmin* role can also use the **`scp`** command to securely copy the file to another system. If you want to copy *syslogd* entries to a remote system, you may use the **`chhmc`** command to change the **`/etc/syslog.conf`** file on the HMC to specify a system to which to copy. For example, the following command causes the syslog entries to be sent to the *myremotesys.company.com* host name:

```
chhmc -c syslog -s add -h myremotesys.company.com
```

The systems administrator must be sure that the *syslogd* daemon running on the target system is set up to receive messages from the network. On most Linux systems, this can be done by adding the -r option to the SYSLOGD_OPTIONS in /etc/sysconfig/syslog file.

3.1.9 Secure boot

Secure boot feature is enabled on the HMC hardware appliance (7063-CR2). Secure boot is also supported when using the virtual Hardware Management Console (vHMC) when running on ESXi or KVM—on Ubuntu and RHEL. This provide integrity of kernel of the system. The kernel images are signed by IBM private keys to ensure the HMC boots up only with IBM HMC supplied kernel images. The keys are validated as the first step in the boot process.

Due to the difference in the architecture, different steps are required to enable secure boot on physical and virtual HMCs:

- Full documentations for secure boot feature enablement steps for hardware HMC can be found at:
  https://www.ibm.com/docs/en/power10/7063-CR2?topic=rack-enabling-secure-boot-7063-cr2-hmc
- The dedicated steps to enable secure boot for virtual HMC based on VMWare ESXi can be found at:
  https://www.ibm.com/docs/en/power10?topic=ihvax-installing-hmc-virtual-appliance-enabled-secure-boot-by-using-vmware-esxi
- The document with steps to enable secure boot for virtual HMC for KVM Hypervisor on RHEL is available at:

https://www.ibm.com/docs/en/power10?topic=ihvax-installing-hmc-virtual-appliance-enabled-secure-boot-by-using-kvm-hypervisor-rhel

- The document with steps to enable secure boot for virtual HMC for KVM Hypervisor on Ubuntu is available at:
  https://www.ibm.com/docs/en/power10/000V-HMC?topic=ihvax-installing-hmc-virtual-appliance-enabled-secure-boot-by-using-kvm-hypervisor-ubuntu

3.1.10 Enabling PowerSC profile for the HMC

A PowerSC agent on the Hardware Management Console (HMC) can be enabled for PowerSC server. You can enable PowerSC communication with the HMC using firewall settings and by installing PowerSC server certificates on the HMC. The PowerSC server detects the HMC and starts managing it as an endpoint. The HMC profile available from PowerSC can then be applied on HMC and monitored for compliance.

For additional information on the enabling the PowerSC profile in the HMC see:

https://www.ibm.com/docs/en/power10/000V-HMC?topic=hmc-enabling-powersc-profile
https://www.ibm.com/docs/en/powersc-standard/2.1?topic=concepts-hmc-hardening-profile

3.1.11 Managing and understanding security vulnerabilities on the HMC

HMC users can subscribe to email notification of corrective service at the Fix Central website. Whenever a vulnerability is discovered on the HMC, a bulletin describing how to obtain the fix will be sent to users. In most cases, because of the closed nature of the HMC and the presence of the restricted shell, some vulnerabilities found on non-HMC systems will not apply. Each time a new release of the HMC code is made available on the support website, a list of security fixes included in the release is also published. To find fixes for your HMC environment see IBM Support Fix Central.

3.1.12 Electronic Service Agent Setup Wizard

The HMC includes a Call Home feature to notify IBM of any issues that occur. It can be configured to send call-home data via direct LAN-based connections or indirect SSL connections to the internet. Additionally, internet support can be provided through a proxy server if needed.

Comprehensive documentation for configuring the HMC to send call-home data, test problem reporting, manage user authorization, and handle information transmission can be found at:

:https://www.ibm.com/docs/en/power10?topic=menus-configuring-local-console-report-errors-service-support

Full documentation to enable call home in 7063 CR2 is available at:
https://www.ibm.com/support/pages/node/6485099

3.1.13 Inband Communication setup on 7063-CR2

In order for the HMC to self-monitor for problem reporting it must be able to communicate with the management controller built in to the HMC. The management controller interface is used to poll for platform events and check the status of hardware components. For general information on how to configure the BMC on the 7063-CR2 HMC, see “How to configure the BMC on HMC 7063-CR2”.

To enable the communication from the HMC to the BMC of the inband connection, the user needs to configure credentials to allow the HMC to connect to the BMC for periodic monitoring of hardware problem events and other management console functions.

To communicate with OpenBMC, two things are needed:

- An inband or “pass-through” interface for the OS to connect to the BMC.

  The 7063-CR2 HMC uses an usb0-to-usb0 model of communication. There is a usb0 interface on the HMC OS, and an usb0 interface on the BMC. The two use a predefined set of IPs. The two interfaces are preconfigured, no user intervention is needed for this step.

  The interfaces are defined as:

  - BMC usb0 IP: 169.254.95.120
  - HMC usb0 IP: 169.254.95.121

- Administrator privilege credentials used to access the OpenBMC to allow discovery and retrieval of events.

  The administrator privilege credentials are necessary for the HMC to communicate with the API of the BMC.The default administrator privilege credentials for the BMC are:

  - Username: root
  - Password: 0penBmc (zero for the O)

  > **Note:** The default password auto-expire3s on the first access by the user, and must be changed.

  It is recommended that a local user, other than root, is configured with administrator privilege to be used for console inband communications. See article “How to add a user to the BMC on the 7063-CR2 HMC” for detailed steps.

Managing HMC to BMC credentials

Once the HMC has booted, a periodic pop-up event is displayed (after 20 minutes), to remind the user to set the inband credentials, unless the user has previously set them. If any users are accessing a shell (ssh or rshterm), they will also receive a “wall” message. The notifications repeat every 24 hrs while the credentials are not set.

How to configure the Console Inband Communication Credentials

This process if fully documented in this IBM Document. To start the process:

1. Select a local user on the BMC with administrator privilege.

   It is recommended that you not use root on the BMC, but instead create a new user with administrator privilege. See “How to add a user to the BMC on the 7063-CR2 HMC”.

2. If running HMC below HMC V10r2.1030:

   – On the HMC, select **Console Settings** → **Console Inband Communication Credentials**

   For HMC v10r2.1030 or later:

   – On the HMC select **HMC Management** → **Inband BMC credentials**

3. When the task loads, it checks if current credentials exist, and validates them. The user is informed if the currently set credentials are valid, failed, not set, or expired.

   There are two types of credentials-related tasks available, Set Credentials or Change Expired Password.

The default task is Set Credentials, unless the previously provided password is expired, in which case it loads in the Change Expired Password task.

> **Note:** The Change Expired Password task, cannot be selected by the user, it is only available when the previously provided password has expired. This scenario can be common on first time setups when the user has yet to configure the BMC and the default credentials of root/0penBMC are still in place.

4. If the current credentials are valid, there is nothing else to do, click **Close** on the message panel, and then click **Close** to end the task.

   If the credentials are failed or not set, then update the credentials by providing a valid username and password and clicking on the **Set Credentials** button. If the credentials are accepted click **Close** to exit.
5. If the credentials are expired, clicking on **Close** on the message panel switches to the Change Expired Password task. Provide a new password (twice to confirm), to update the new password for the user, on the BMC. Click **Change Expired Password**.

3.1.14 Summary

The HMC is an integral part of the management of the IBM Power ecosystem. To provide the range of functions, it needs to connect to: the servers it is managing, to IBM for call home, and to the users that are administering the environment. It is designed to make these connections in a secure manner. We have discussed many areas where you can ensure that you are taking full advantage of the security options available.

3.2 Cloud Management Console security

As private and hybrid cloud deployments expand, enterprises require new management insights into these environments. Tools that offer consolidated information and analytics are crucial for smooth infrastructure operations. The IBM Cloud Management Console (CMC) for Power Systems delivers a unified view of your Power Systems cloud landscape, regardless of the number of systems or data centers involved.

3.2.1 Overview of Cloud Management Console

The Cloud Management Console (CMC) is an IBM Cloud-based solution designed to manage your IBM Power environment across all data centers. It offers system inventories, including virtual components, and consolidated performance data to optimize utilization throughout your data centers. Additionally, CMC provides aggregated logging information for deeper insights and features Patch Planning to help identify and apply necessary updates.

Hosted in the IBM Cloud, the CMC ensures secure, anytime access, enabling system administrators to easily generate reports and gain insights into their Power cloud deployments. The CMC is required for IBM Power Enterprise Pools 2, a cloud solution that facilitates dynamic resource allocation and pay-as-you-go capacity management within the IBM Power environment.

CMC is not a single product but a platform through which IBM delivers applications and microservices in a DevOps model. The solution is designed for mobile devices, tablets, and desktop browsers, ensuring convenient access for cloud operators.

Cloud Connector is the service that runs on the IBM Power Hardware Management Console (HMC) and sends system resource usage data to the IBM Cloud Management Control (CMC) service. The Cloud Connector and the Cloud Management Console provide the applications shown in Table 3-2 for the IBM Power ecosystem.

Table 3-2 Cloud Management Console applications for IBM Power

Feature	Benefits
Inventory Aggregation	► Enterprise views of Power Servers, HMCs, LPARs and resources associated with these components which provides insight into health and status of the enterprise ► Centralized Hardware Inventory ► Grouping of resources using customer supplied names which customize views
Performance Monitoring	► Enterprise views providing resource consumption and performance for Power servers, LPARs and I/O components ► Guest Operating Systems performance metrics that point out potential areas for improvement ► Enterprise views of energy consumed by Power Systems
Log Trends	► System log aggregation across the Power Systems enterprise which provides a central point to view log trends from Live Partition Mobility, remote restart, and the lifecycle of LPARs
Patch Planning	► Know all your patch planning needs at glance! (incl. Firmware, VIOS, OS and HMC) ► Identify all patch dependencies ► Integrated, collaborative planning with stakeholders
Power Enterprise Pools 2	► Provides support for managing resource sharing across systems defined in Power Enterprise Pools 2 (PEP2). PEP2 provides an advanced resource sharing option across multiple servers and multiple sites to allow an enterprise to run workloads on any server in the pool.

3.2.2 User and resource roles management on CMC

Users access the CMC portal using a valid IBMid. IBM supports federated authentication with the IBMid, allowing organizations to use their own login pages and security controls to securely access the CMC.

User management is handled by the administrator of the organization registered for IBM Cloud Management Console for Power Systems services. Administrators can manage users from the **Settings** page. To access this page, click the navigation menu icon in the CMC portal header, then select the **Settings** icon. On the **Settings** page, click the **Manage Users** tab to view all users configured for your organization. Users without administrator privileges will have limited access to specific applications.

To be added to the CMC, users must have a valid IBMid. In addition to the IBMid, users must be added to the CMC application by the administrator within your company. The resource role assignment feature allows administrators to assign appropriate tasks to users.

Resource roles can be managed from the CMC Portal **Settings** page. On this page, click the **Manage Resource Roles** tab to view and manage resource roles. Administrators can add, modify, or delete resource roles for other users from this page.

3.2.3 Protection for sensitive data

The Cloud Management Console provides techniques for protecting sensitive data. Although actual user data is not uploaded from the Cloud Collector, some of the information in the configuration data could be considered to be sensitive and data from some systems might be considered sensitive. To provide flexibility in managing what data is uploaded, CMC provides a method for identifying systems that should not upload data to the CMC. CMC also provides a methodology for masking sensitive metadata within your configuration.

Blocklist, Allowlist, and No List

The three options for filtering systems from uploading data are:

- Blocklist: To prevent the Cloud Connector from uploading data for specific managed systems, add these systems to the blocklist. The **Blocklist** tab in the **Managed System Filter** area displays the managed systems that are on the blocklist, including their model, type, and machine serial number (MTMS). Data from these systems will not be uploaded to the cloud.
- Allowlist: To specifically permit the Cloud Connector to upload data from certain managed systems, add these systems to the allowlist. The **Allowlist** tab in the **Managed System Filter** area shows the managed systems that are on the allowlist, including their model, type, and MTMS. Only data from managed systems on the allowlist will be uploaded to the cloud.

> **Attention:** Data filtering for the allowlist is supported only with HMC version 1020 or later. If your version does not meet this requirement, data from systems not on the allowlist will still be uploaded to the CMC.

- No List: Selecting **No List** disables both filtering types and shows data from all managed systems.

To view the current managed systems on the blocklist or allowlist, click the **Blocklist** and **Allowlist** tabs in the Managed System Filter area.

> **Important:** Only one filter type, Blocklist, Allowlist, or No List, can be active at a time

Managing the Blocklist:

To add a managed system to the blocklist, click the Blocklist tab and confirm your selection. In the **Managed System Filter** area, click **Edit Blocklis**t, enter the model, type, and serial number of the system you want to block, and click **Add**. To remove a system from the blocklist, click the minus sign (-) next to the system's name and confirm the removal.

Note that adding a system to the blocklist does not automatically remove its existing data from the cloud. To purge this data, ensure the Cloud Connector is running, and use the command run `chsvc -s cloudconn -o stop --purge` from the management console command line.

> **Important:** Systems in Power Enterprise Pool 2.0 cannot be blocklisted.

Managing the Allowlist:

To add a managed system to the allowlist, click the Allowlist tab and confirm your selection. In the Managed System Filter area, click Edit Allowlist, specify the model, type, and serial number of the system you want to allow, and click Add. To remove a system from the allowlist, click the minus sign (-) next to the system's name and confirm the removal.

No List Option:

Click No List to disable the blocklist and allowlist filters and display data from all managed systems.

Data Filter

To filter the data in Cloud Connector from getting pushed to cloud storage, you add the systems to this table. Selections are available to filter the System IP Address and Logical Partition/Virtual IO Server IP Address. These systems can be re-enabled if you want to.

After the selection is made, it will take about 5 to 10 minutes to reflect the data in the CMC. The patch planning data gets updated once a day, so you might see delay in reflecting the changes in the Patch Planning app. The purge operation is supported if the HMC is at Version 8 Release 8.6.0 Service Pack 2 or higher.

CMC Attribute Masking

By enabling the attribute masking feature, you can ensure that sensitive data does not leave your data center. After enabling this feature, Cloud Connector masks sensitive data and sends the masked data to the CMC server, and the CMC UI displays these masked values of the resource attributes on all the CMC pages and apps. When Attribute Masking is enabled, the CMC APIs also contain masked data in their response.

To enable Attribute Masking, from the Cloud Console interface click **Settings** → **Cloud Connector** → **Cloud Connector Management**, scroll down to the end of the page, and then set **Attribute Masking** to On.

The attribute masking feature is available with HMC 1040 and later only. Data from earlier HMC versions is not masked and will continue to be displayed unmasked even when Attribute Masking is enabled. For details on what fields are masked see Attribute Masking.

3.2.4 Cloud Connector connections

Cloud Connector needs to connect to multiple end points to send and receive the data needed used by the Cloud Management Console. These are all managed through outbound connection to the IPs and the ports for the different components. There are multiple sources requiring outbound connections:

- HMC Cloud Connector to CMC Cloud Portal Server
- HMC Cloud Connector to CMC Cloud Database (Configuration/Certificates)
- HMC Cloud Connector to CMC Cloud Data Ingestion Node

Cloud Connector connections from HMC to CMC Cloud Portal

Cloud Connector provides connection between the HMC in your data center and the CMC instance. It is a component that uploads data to the CMC cloud. Cloud Connector is preinstalled on the HMC, but is not started by default. The Cloud Connector can be started using a key which is available in:

CMC Portal → **Settings** → **Cloud Connector Settings**.

Cloud Connector supports either a direct outbound connection or a connection through a proxy server to the IBM CMC portal. The Cloud Connector supports a basic authentication protocol as well as other authentication protocols, such as Kerberos, LDAP, and Digest-MD5 to connect to the proxy server. This section provides details on various security aspects of CMC.

IP connections and ports required

The IPs and Ports required for connecting to the CMC Cloud Portal can be found on the CMC Portal (C**MC Portal** → **Settings** → **Cloud Connector Settings)**. The HMCs need either direct outbound connectivity to those endpoints or then need a proxy connection set up. For the connection to the cloud portal an HTTP proxy can be used.

> **Important:** Starting with HMC version 9.1.941.0, the Cloud Connector supports an HTTP proxy. If you are using versions of the HMC older than that, the Cloud Connector requires a SOCKS5 proxy.

Cloud Connector supports Kerberos, LDAP, and Digest-MD5 based proxy server authentication in addition to Basic authentication. While starting Cloud Connector, an attribute can be specified which designates the authentication type to be used for the proxy connection. The default authentication used is Basic.

Starting the Cloud Connector

To start the Cloud Connector, complete the following steps on the HMC:

1. In the navigation area, click *HMC management.*
2. In the content panel, click *Cloud connector*. The *State, Authentication Type, HTTP Proxy,* and *Sock Proxy* of the Cloud Connector is displayed.
3. To start the Cloud Connector, click *Start cloud connector* and follow the steps in the wizard.

Cloud Connector utilizes a one-way push model where It initiates all outbound communication. For automatic network based configuration, where Cloud Connector pulls the configuration file from the cloud database, HTTPS is used; and for application data flow (push) between Cloud Connector and the CMC data ingestion node, TCP with SSL is used. All communication from the Cloud Connector to the CMC are secured using the Transport Layer Security Version 1.2 protocol (TLSv1.2).

The startup key for the HMC based cloud connector is used to establish a valid connection between the connector and the CMC Cloud Portal Server (cloud portal). This key is also used for connection between the Connector and the configuration database. Once a valid connection is established to the cloud portal, credentials are returned to the Cloud Connector allowing for dynamic configuration and reconfiguration.

Figure 3-2 shows the Cloud Connector establishing trust with the cloud portal via pushing the user provided key to a cloud portal key verification endpoint.

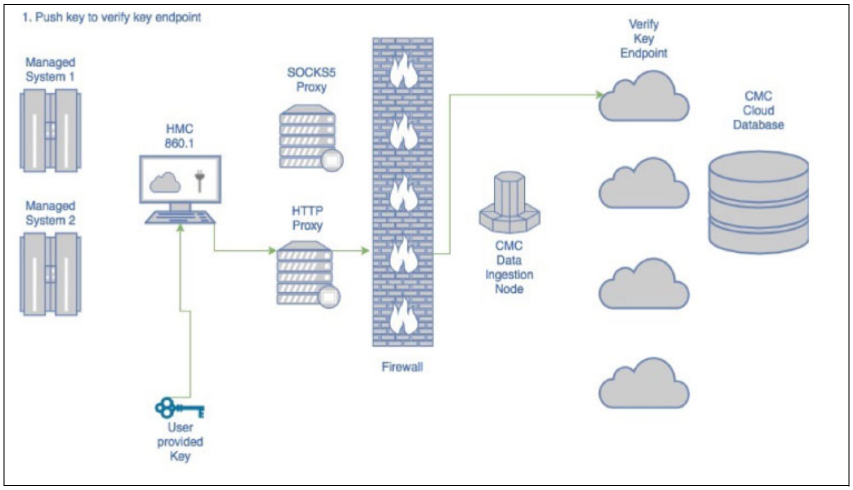


Figure 3-2 Cloud Connector push key to cloud portal[4]

Once the key verification is successful, the CMC cloud portal Server returns credentials for pulling the Cloud Connector configuration file and SSL certificates as shown in Figure 3-3.

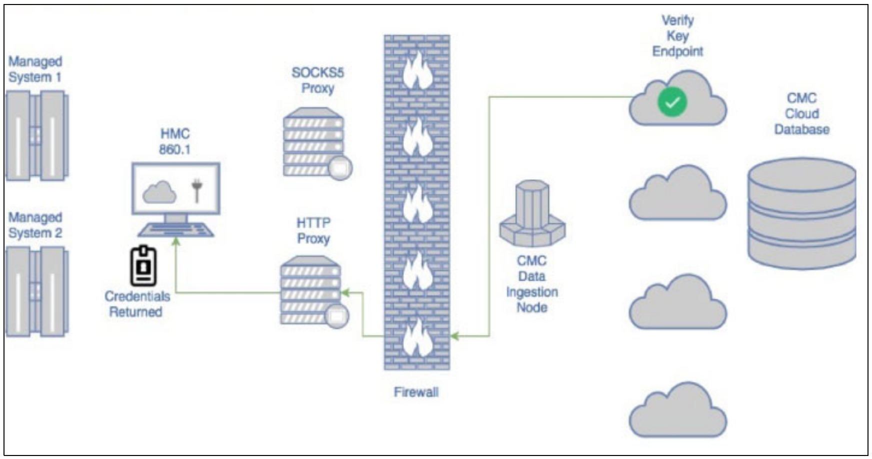


Figure 3-3 Cloud portal returns credentials to Cloud Connector[4]

A security test is executed to assert that the startup key provided is valid. The test begins with a GET request from the connector to the cloud portal which will return a cross-site request forgery (XSRF) header. This XSRF header, along with a portion of the decoded key, the POST operation is performed to the same cloud portal endpoint. If the key is considered valid, the cloud portal will respond with a set of encoded credentials giving cloud connector access to a database containing the customers Cloud Connector configuration file.

Cloud Connector connections from HMC to CMC Cloud Database

Once Cloud Connector established successful connection with the cloud portal, a secure SSL connection is then established between Cloud Connector and the cloud database to fetch the configuration file. This configuration file contains:

- Cloud applications enabled by the customer
- Data to push for those applications
- Data to filter (block listed managed systems, selected the system and partition IP addresses)
- IP address of the cloud data ingestion node.

In addition, it provides credentials for fetching an SSL certificate and key pair used in communication between the Connector and the cloud data ingestion node. The credentials are used to access a separate database from the one used to fetch the configuration file. However, the underlying network location and mechanism used to fetch the certificates is the same.

An SSL connection is established and the data is returned to the connector. Every minute the cloud connector fetches a new configuration to ensure changes are handled. All communication from the connector to the cloud database are secured using the Transport Layer Security Version 1.2 protocol (TLSv1.2). Using the received credentials as shown in Figure 3-3, the cloud connector pulls the customer specific cloud connector configuration file from the CMC Cloud Configuration Database as shown in Figure 3-4 on page 61.

[4] https://ibmcmc.zendesk.com/hc/en-us/article_attachments/360083545614/CloudConnectorSecurityWhitePaper.pdf

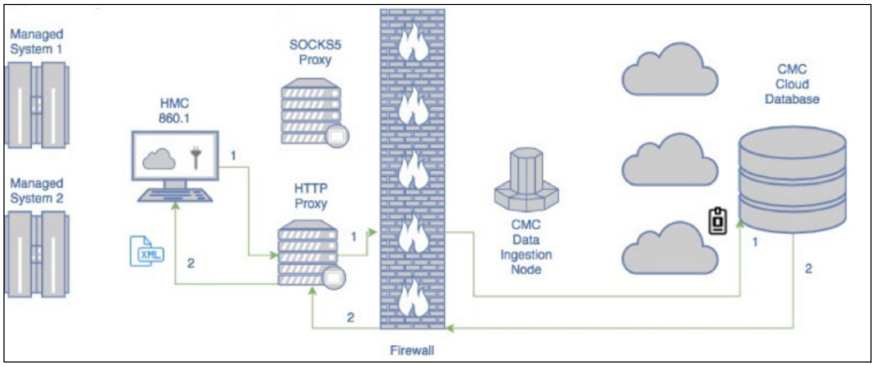


Figure 3-4 Cloud Connector pulls configuration file from CMC database[4]

Once credentials from the configuration file is collected as shown in Figure 3-3 on page 60, the cloud connector pulls the SSL certificates and key from the CMC Cloud Certificate Database as shown in Figure 3-5.

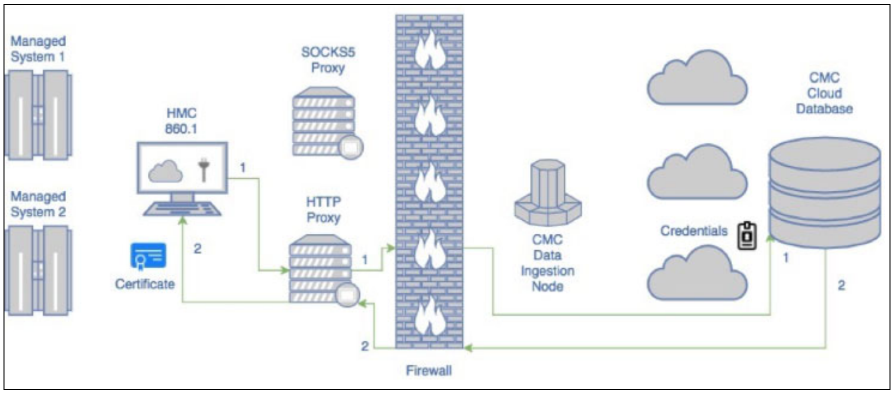


Figure 3-5 Cloud Connector pulls SSL keys from cloud database[4]

Cloud Connector connections from HMC to CMC Cloud Data Ingestion Node

Once the Cloud Connector is configured via the automated configuration process, it will begin collecting data and pushing that data to the data ingestion node. This channel is secured using SSL with mutual authentication using the certificate and key mentioned in Figure 3-5. Using mutual authentication ensures that the connector only sends data to trusted data ingestion nodes. The certificate and key are stored on the HMC filesystem but are only accessible by the root user.

Figure 3-6 on page 62 shows the connection between the HMC and the ingestion node through SOCKS5 proxy using the certificate obtained in Figure 3-5. With 9.1.941.0 version of HMC, cloud connector can be started only with HTTP proxy option.

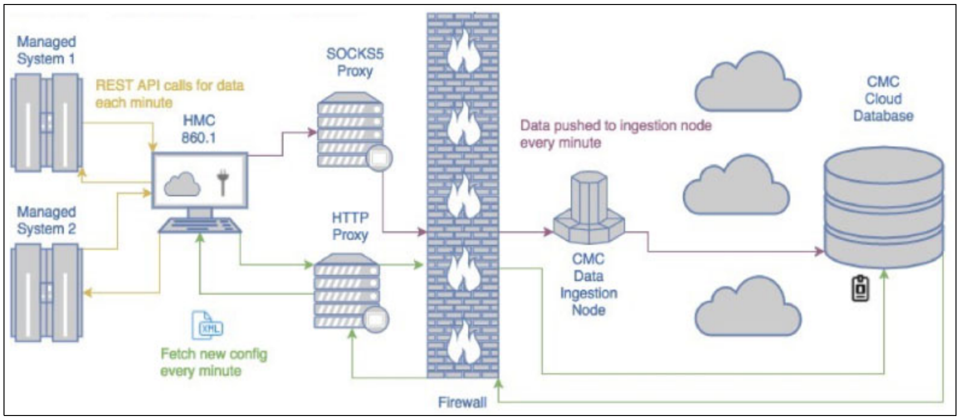


Figure 3-6 Cloud Connector authentication to CMC data ingestion node SOCK5 version[4]

If the Cloud Connector is started with only HTTP proxy, then it uses HTTP proxy to establish connection between HMC and ingestion node as shown in Figure 3-7. The proxy options shown in Figure 3-6 are still supported in current versions of the HMCs, when the Cloud Connector is started with both HTTP and SOCKS5 proxies.

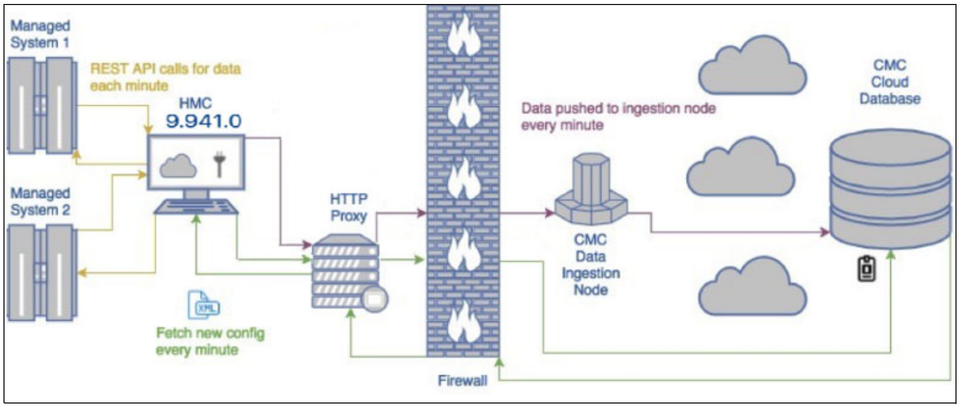

Managed System 1
REST API calls for data each minute
HMC 9.941.0
Data pushed to ingestion node every minute
CMC Cloud Database
Managed System 2
HTTP Proxy
CMC Data Ingestion Node
Fetch new config every minute
Firewall

Figure 3-7 Cloud Connector authentication to CMC data ingestion node new HTTP proxy mode[4]

3.2.5 Hints for using the CMC

This section presents some suggestions on how to personalize the CMC to meet specific needs.

Enabling and disabling CMC applications

An administrator can enable or disable applications that are hosted by the CMC servers. In the navigation area, select **Settings**. In the contents area, click **Apps**. Set the switch next to each application to enable or disable that application.

IBM Power Enterprise Pools 2.0 provides enhanced multi-system resource sharing and by-the-minute consumption of on-premises compute resources to clients who deploy and manage a private cloud infrastructure. A Power Enterprise Pool 2.0 is monitored and managed by the IBM Cloud Management Console. The CMC Enterprise Pools 2.0 application helps you to monitor base and metered capacity across a Power Enterprise Pool 2.0 environment, with both summary views and sophisticated drill-down views of real-time and historical resource consumption by logical partitions.

Important: To use the **Power Enterprise Pools 2.0** app, the **Capacity Monitoring** app must be enabled first for the collection of performance data.

Viewing logging dashboards

CMC provides insights into various virtualization operations in the *Dashboard*. The *Dashboard* information includes Live Partition Mobility, remote restart, and other partition activities. The logging dashboard is available in the navigation area. Select *Logging* and then In the contents area, select the PowerVM Virtualization actions that you want to view.

View Patch Planning Information

You can get a comprehensive view of your inventory with information about the current patch state of resources in your environment. You can also view a list of resources that need to be updated, and the recommended service level for each resource. The Patch Planning > Inventory > All view lists all the resources in your environment including the operating systems, firmware, Virtual I/O Servers, Adapters and Hardware Management Consoles.

3.3 VIOS Security

The Virtual I/O Server (VIOS) is a specialized AIX-based appliance designed to virtualize I/O adapters for Virtual I/O client LPARs running AIX, IBM i, and Linux. Typically, two VIOS instances are installed on the same managed system to ensure high availability and to minimize risks from human errors, maintenance windows, and potential hardware failures.

Given that all virtualized I/O traffic routes through VIOS, securing it is crucial. If an attacker compromises VIOS, they could gain access to all virtualized network and storage traffic on the system, potentially infiltrating client LPARs.

After deploying VIOS, the first priority for an administrator is to configure it securely. Many of the security settings applicable to AIX can also be applied to VIOS. Given VIOS's appliance nature, if you're unsure about applying specific security configurations, contact IBM support for assistance.

Although VIOS does not have its own published security benchmarks, you can refer to the Center for Internet Security (CIS) AIX benchmark to guide your VIOS security configuration. Note that VIOS version 3.1 is based on AIX 7.2, while VIOS version 4.1 is based on AIX 7.3.

VIOS currency and security fixes

As with all other operating systems, you need to keep your VIOS server updated and current. You must regularly check for updates and be sure that you are using a supported version of VIOS. Unsupported versions do not get any security fixes and as such should not be considered to be secure.

The easiest way to get information about new VIOS releases, service packs and security fixes is to subscribe to Virtual I/O Server notifications on the IBM support portal.

Occasionally, VIOS receives security fixes to address newly identified vulnerabilities. You should apply these fixes in accordance with your company's security and compliance policies. Many VIOS administrators delay installing security updates, opting to wait for the next service pack instead. This approach can be acceptable if you've evaluated the risks associated with a compromised virtualization infrastructure and your organization is prepared to accept those risks.

If your infrastructure is designed for high availability with dual VIOS instances per managed system and all client LPARs are correctly connected, you can install security fixes without disrupting operations, potentially even during normal business hours. To do this, first install the fixes on one VIOS instance, then reboot that instance to activate the updates. After the VIOS restarts, verify that all client LPARs have their storage paths and network connections restored. Next, repeat the process on the second VIOS instance.

Decide which VIOS to update first based on the needs of your client LPARs. To prevent log messages during the update, you can disable the storage paths to the VIOS that is being updated.

3.3.1 Virtual I/O Server system security hardening

Beginning with Version 1.3 of the Virtual I/O Server, you can set security options that provide tighter security controls over your Virtual I/O Server environment. These options allow you to select a level of system security hardening and specify the settings allowable within that level. The Virtual I/O Server security feature also allows you to control network traffic by enabling the Virtual I/O Server firewall. You can configure these options by using the **viosecure** command.

To help you set up system security when you initially install the Virtual I/O Server, the Virtual I/O Server provides the configuration assistance menu. You can access the configuration assistance menu by running the **cfgassist** command. Using the **viosecure** command, you can set, change, and view current security settings. By default, no Virtual I/O Server security levels are set. You must run the **viosecure** command to change the settings.

The system security hardening feature protects all elements of a system by tightening security or implementing a higher level of security. Although hundreds of security configurations are possible with the Virtual I/O Server security settings, you can easily implement security controls by specifying a high, medium, or low security level.

Using the system security hardening features provided by Virtual I/O Server, you can specify values such as the following:

- Password policy settings
- Actions such as usrck, pwdck, grpck, and sysck
- Default file-creation settings
- Settings included in the **crontab** command

Configuring a system at too high a security level might deny services that are needed. For example, **telnet** and **rlogin** are disabled for high-level security because the login password is sent over the network unencrypted. If a system is configured at too low a security level, the system might be vulnerable to security threats. Since each enterprise has its own unique set of security requirements, the predefined High, Medium, and Low security configuration settings are best suited as a starting point for security configuration rather than an exact match for the security requirements of a particular enterprise. As you become more familiar with the security settings, you can make adjustments by choosing the hardening rules that you want to apply. You can get information about the hardening rules by running the **man** command.

Configuring Virtual I/O Server system security hardening

The following illustrates how you can set the security level to specify security hardening rules for your Virtual I/O Server system.

To implement system security hardening rules, you can use the **viosecure** command to specify a security level of high, medium, or low. A default set of rules is defined for each level. You can also set a level of default, which returns the system to the system standard settings and removes any level settings that have been applied.

The low-level security settings are a subset of the medium level security settings, which are a subset of the high-level security settings. Therefore, the high level is the most restrictive and provides the greatest level of control. You can apply all of the rules for a specified level or select which rules to activate for your environment. By default, no Virtual I/O Server security levels are set; you must run the **viosecure** command to modify the settings.

Use the following tasks to configure the system security settings.

To set a Virtual I/O Server security level of high, medium, or low, use the command **viosecure -level**.

```
viosecure -level low -apply
```

Changing the settings in a security level

To set a Virtual I/O Server security level in which you specify which hardening rules to apply for the setting, run the **viosecure** command interactively.

For example:

1. At the Virtual I/O Server command line, type **viosecure -level high**. All the security level options (hardening rules) at that level are displayed ten at a time (pressing **Enter** displays the next set in the sequence).
2. Review the options that are displayed and make your selection by entering the numbers, which are separated by a comma, that you want to apply, or type **ALL** to apply all the options or **NONE** to apply none of the options.
3. Press **Enter** to display the next set of options, and continue entering your selections.

Note: To exit the command without making any changes, type "“q”".

Viewing the current security setting

To display the current Virtual I/O Server security level setting, use the **viosecure** command with the **-view** flag.

```
viosecure -view
```

Removing security level settings

To unset any previously set system security levels and return the system to the standard system settings, run the command **viosecure -level default**

To remove the security settings that have been applied, run the command **viosecure -undo**.

3.3.2 Virtual I/O Server firewall

Using the Virtual I/O Server firewall, you can enforce limitations on IP activity in your virtual environment. With this feature, you can specify which ports and network services are allowed access to the Virtual I/O Server system. For example, if you need to restrict login activity from an unauthorized port, you can specify the port name or number and specify deny removing it from the allow list. You can also restrict a specific IP address.

Configuring Virtual I/O Server firewall settings

The following illustrates how you can enable the Virtual I/O Server firewall to control IP activity.

The Virtual I/O Server firewall is not enabled by default. To enable the Virtual I/O Server firewall, you must turn it on by using the **`viosecure`** command with the **`-firewall`** option. When you enable it, the default setting is activated, which allows access for the following IP services:

- ftp
- ftp-data
- ssh
- web
- https
- rmc
- cimom

> **Note:** The firewall settings are contained in the *viosecure.ctl* file in the */home/ios/security* directory. If for some reason the *viosecure.ctl* file does not exist when you run the command to enable the firewall, you receive an error. You can use the **`-force`** option to enable the standard firewall default ports.

You can use the default setting or configure the firewall settings to meet the needs of your environment by specifying which ports or port services to allow. You can also turn off the firewall to deactivate the settings.

Configuring the VIOS firewall

Use the following tasks at the Virtual I/O Server command line to configure the Virtual I/O Server firewall settings:

1. Enable the Virtual I/O Server firewall by running the command:

   ```
   viosecure -firewall on
   ```

2. Specify the ports to allow or deny, by using the command:

   ```
   viosecure -firewall allow | deny -port number
   ```

3. View the current firewall settings by running the command:

   ```
   viosecure -firewall view
   ```

4. If you want to disable the firewall configuration, run the command:

   ```
   viosecure -firewall off
   ```

> **Tip:** Additional information on securing your Virtual I/O Server can be found in this IBM Document.

3.3.3 User management and authentication

Similar to AIX you can have multiple user accounts on VIOS. The reason to have multiple accounts is to limit access to padmin account - the main VIOS administrative account. You may want to create a view-only account for your monitoring software or a separate account for automating your VIOS operations with Ansible.

Creating an additional administrative account on VIOS

Use the command **mkuser** to create a new user's account on VIOS. An example is shown in Example 3-3.

Example 3-3 Create administrative user account on VIOS

```
$ mkuser admin
Changing password for “admin”
admin’s New password:
Enter the new password again:
```

The account “admin” will get all privileges to change VIOS configuration but will not be able to switch into oem_setup_env mode and execute root commands.

Creating view-only user

Use the command mkuser with additional attributes to create a new read-only account on VIOS as shown in Example 3-4.

Example 3-4 Create view-only user account on VIOS

```
$ mkuser -attr pgrp=view monitor
Changing password for "monitor"
monitor's New password:
Enter the new password again:
```

The account “monitor” will be able to login into VIOS and see the current configuration of the VIOS, but will not be able to make changes to the configuration.

Listing VIOS users

Use the **lsuser** command to list users on VIOS. This is shown in Example 3-5.

Example 3-5 List existing users on VIOS

```
$ lsuser
padmin roles=PAdmin,CacheAdm default_roles=PAdmin,CacheAdm account_locked=false
expires=0 histexpire=52 histsize=4 loginretries=0 maxage=13 maxexpired=4
maxrepeats=8 minage=4 minalpha=2 mindiff=0 minlen=10 minother=0 pwdwarntime=5
registry=files SYSTEM=compat
admin roles=Admin default_roles=Admin account_locked=false expires=0 histexpire=52
histsize=4 loginretries=0 maxage=13 maxexpired=4 maxrepeats=8 minage=4 minalpha=2
mindiff=0 minlen=10 minother=0 pwdwarntime=330 registry=files SYSTEM=compat
```

The super-administrators with access to *oem_setup_env* mode will have role *PAdmin*. All other administrators have role *Admin*. View-only users have role *ViewOnly*.

Changing user privileges

To change user privileges use the **chuser** command. For example, if you wish to degrade some user to read-only user, set its roles to ViewOnly as in this command:

```
chuser -attr roles=ViewOnly default_roles=ViewOnly pgrp=view admin
```

If you want to assign admin privileges to a user use this command:

```
chuser -attr roles=Admin default_roles=Admin pgrp=system admin
```

After you change the privileges, the user must re-login to the system in order the changes be effective.

Changing password policy for user

You can set password policy for users by changing the attributes shown in Table 3-3 with the command **chuser**

Table 3-3 Attributes for user passwords

Attribute	Meaning
expires	Expiration date of the account in format MMDDhhmmyy. If the value is set to 0, the account doesn't expire
histexpire	Period in weeks when the user can't reuse old password
histsize	Number of previous passwords which user can't reuse
loginretries	Number of unsuccessful logins attempts before the account is locked
maxage	Maximum number of weeks when the password is valid.
maxexpired	Maximum number of weeks when the user can change its expired password.
maxrepeats	Maximum number of times a character can be repeated in the password.
minage	Minimum number of weeks when the user can't change the password after the new password is set.
minalpha	Minimum number of alphabetic characters in the password.
mindiff	Minimum number of characters in the new password which should differ from the old password
minlen	Minimum length of the password
minother	Minimum number of non-alphabetic characters in the password
pwdwarntime	Number of days before a warning about password expiration is shown.

Locking and unlocking accounts

You can temporary lock user's account by setting the following attribute:

```
chuser -attr account_locked=true admin
```

To unlock the account, you must set the attribute to false.

Removing unneeded accounts

If you don't need a user account, you can delete it using **rmuser** command:

```
rmuser -rmdir admin
```

The option *-rmidr* will remove user's home directory. Please note: the files on VIOS which are owned by the removed user, will not change their ownership automatically,

3.3.4 Auditing of VIOS commands

All commands which are issued by administrative or view-only accounts are logged. The logs are saved in file */home/ios/logs/ioscli_global.log*. You can access the file only in *oem_setup_env* mode. That’s why it is important to have separate accounts for administrators and not use *padmin* user for daily tasks!

Local date, time, user account and the issued command are saved in the file. An example is shown in Example 3-6.

Example 3-6 ioscli_global.log

```
Jun 24 2024, 18:55:20 monitor  vfcmap  -vadapter vfchost28 -fcp
Jun 24 2024, 18:56:10 admin    lsuser
Jun 24 2024, 18:58:01 padmin   chuser  -attr roles=ViewOnly default_roles=ViewOnly
pgrp=view admin
```

4

AIX Security

Security is a very hot topic for a good reason. So much of our personally identifying data is now being stored that the security break-ins that have been happening have most likely affected everyone reading this chapter. Additionally, the penalties now for breaches of the various standards (HIPAA, PCI, etc) are significant. Good security requires a multi-layered approach that starts with people, then physical security, and then the various layers. It's critical to look at the whole environment and see how security can be applied at each level.

In this chapter, we will cover some of the basics of locking down your AIX LPARs. Some of security hardening is implemented in AIX 7.3 by default. Usually you must check and implement some additional security settings according to your environmental requirements.This includes setting default permissions and umasks, using good usernames and passwords, hardening the security with AIX Expert, protecting the data at rest directly on the disk or at the logical volume layer via encryption, removing insecure daemons and integrating with LDAP directory services or Microsoft Active Directory (AD)).

This chapter describes the following:

4.1 AIX Security Checklist

If you are setting up a new AIX system and need guidance on basic security measures and initial steps, IBM has published a security checklist[1] to assist you. We have included it here for your convenience.

While this checklist is not exhaustive, it provides a solid foundation for developing a comprehensive security plan tailored to your environment. We will cover these recommendations and introduce additional considerations in the following sections.

Here is a checklist of security actions to perform on a newly installed or existing system.

When installing a new system:

- Use secure base media to install AIX.
- Avoid installing desktop software, such as CDE, GNOME, or KDE, on servers.
- Apply necessary security fixes and any recommended maintenance and technology level updates. For the latest service bulletins, security advisories, and fix information, visit the IBM Support Fix Central website (`https://www.ibm.com/support/fixcentral`).
- Back up the system after the initial installation and store the backup in a secure location.
- Set up access control lists for restricted files and directories.
- Disable unnecessary user and system accounts, such as daemon, bin, sys, adm, lp, and uucp. Deleting accounts is not advised as it removes account information, such as user IDs and user names, which may still be linked to data on system backups. If a user is recreated with a previously deleted ID and the backup is restored, the new user might gain unintended access.
- Regularly review and remove unnecessary daemons and services from the /etc/inetd.conf, /etc/inittab, /etc/rc.nfs, and /etc/rc.tcpip files.
- Verify that the permissions for the files listed in Example 4-1 are correctly set.

Example 4-1 Validate permissions of files

```
-rw-rw-r-- root     system  /etc/filesystems
-rw-rw-r-- root     system  /etc/hosts
-rw------- root     system  /etc/inittab
-rw-r--r-- root     system  /etc/vfs
-rw-r--r-- root     system  /etc/security/failedlogin
-rw-rw---- root     audit   /etc/security/audit/host
```

- Disable the root account from being able to remotely log in. The root account should be able to log in only from the system console.
- Enable system auditing. For more information, see Auditing overview in AIX documentation.
- Enable a login control policy. For more information, see Login control in AIX documentation.
- Disable user permissions to run the xhost command. For more information, see Managing X11 and CDE concerns in AIX documentation.
- Prevent unauthorized changes to the PATH environment variable. For more information, see PATH environment variable in AIX documentation.

[1] `https://www.ibm.com/docs/en/aix/7.3?topic=security-checklist`

- Disable telnet, rlogin, and rsh. For more information, see TCP/IP security in AIX documentation.
- Establish user account controls. For more information, see User account control in AIX documentation.
- Enforce a strict password policy. For more information, see Passwords in AIX documentation.
- Establish disk quotas for user accounts. For more information, see Recovering from over-quota conditions in AIX documentation.
- Allow only administrative accounts to use the **su** command. Monitor the **su** command's logs in the /var/adm/sulog file.
- Enable screen locking when using X-Windows.
- Restrict access to the **cron** and **at** commands to only the accounts that need access to them.
- Use an alias for the **ls** command to show hidden files and characters in a file name.
- Use an alias for the **rm** command to avoid accidentally deleting files from the system.
- Disable unnecessary network services. For more information, see Network services in AIX documentation.
- Perform frequent system backups and verify the integrity of backups.
- Subscribe to security-related e-mail distribution lists.

4.2 Encrypted File Systems

The Encrypted File System enables individual users on the system to encrypt their data on the JFS2 file system using their personal key stores. Each user is associated with a unique key. These keys are stored in a cryptographically protected key store, and upon successful login, the user's keys are loaded into the kernel and associated with the process's credentials.

Later, when the process needs to open an EFS-protected file, these credentials are checked. If a key matching the file protection is found, the process can decrypt the file key and, consequently, the file content. Group-based key management is also supported

> Note: EFS is part of an overall security strategy. It is designed to work in conjunction with sound computer security practices and controls.

4.2.1 Encrypted File System usability

Encrypted File System (EFS) key management, file encryption, and file decryption are transparent to users in normal operations.

EFS is part of the base AIX operating system. To enable EFS, *root* – or any user with the RBAC authority **aix.security.efs** authorization – must use the **efsenable** command to activate EFS and create the EFS environment. See section 4.2.3, “Root Access to User Keys” on page 74 for more information on who can manage EFS. This is a one time system enablement.

After EFS is enabled, when a user logs in, their key and keystore are silently created and secured or encrypted with the user’s login password. The user's keys are then used automatically by the JFS2 file system for encrypting or decrypting EFS files. Each EFS file is protected with a unique file key, which is in turn secured or encrypted with the file owner’s or group’s key, depending on the file permissions. By default, a JFS2 File System is not EFS-enabled.

When a file system is EFS-enabled, the JFS2 File System transparently handles encryption and decryption in the kernel for read and write requests. User and group administration commands (such as **mkgroup**, **chuser**, and **chgroup**) manage the keystores for the users and groups seamlessly.

The following EFS commands are provided to allow users to manage their keys and file encryption:

efskeymgr	Manages and administers the keys.
efsmgr	Manages the encryption of files/directories/file system.

4.2.2 User Keystores

The user keystore is managed automatically for basic operations. Users can perform advanced tasks and create encrypted files/directories using the **efskeymgr** and **efsmgr** commands respectively. Access to group keystores is automatically granted when a user is added to a group. Similar to UNIX ACLs, file owners can control access to encrypted files.

Users can change their login password without affecting open keystores and the keystore password can be different from the login password. When the user password differs from the keystore password, requiring manual loading with efskeymgr.

Keystore Details

The keystore has the following characteristics:

- Protected with passwords and stored in PKCS #12 format.
- Location:
  - User: /var/efs/users/<username>/keystore
  - Group: /var/efs/groups/<groupname>/keystore
  - efs_admin: /var/efs/efs_admin/keystore
- Users can choose encryption algorithms and key lengths.
- Access is inherited by child processes.

Group Key Management

Keys are kept ab both the user level and the group level. For group keys:

- Only group members can add/remove group keys in Root Guard mode.
- User keystores contain user private keys and passwords to access group keystores.
- Group keystores contain the group's private keys.

4.2.3 Root Access to User Keys

Root privileges can be configured to provide either unrestricted or limited access to user keys. Regardless of configuration, root cannot directly assume a user's identity (using su) to access encrypted files or keystores.

Unrestricted Root Access

In this mode, root can reset a user's keystore password, potentially gaining access to the keys within. This configuration offers greater flexibility for system administration.

Restricted Root Access

In this mode, root can only reset a user's login password, not their keystore password. Root cannot impersonate a user (via su) to access an open keystore. Although root can create, modify, and delete users and their associated keystores, accessing the keys within these keystores remains prohibited. This configuration provides enhanced protection against malicious root activity.

4.2.4 EFS Keystore Management Modes

There are two main modes for managing EFS keystores: Root Admin and Root Guard. Additionally, a special “efs_admin” key provides root-level access to all keystores.

Root Admin Mode

Root Admin mode is the default setting. It offers full access to all keystores, including user and group keystores. In this mode Root can reset user keystore passwords, potentially compromising data if a user forgets their password.

Root Guard Mode

This mode provides strong security against unauthorized access, even for root users. However, losing the user’s keystore password will result in data loss as there is no recovery methodology. Some keystore operations (adding/removing group access keys, regenerating private keys) might require user intervention.

efs_admin Key

The efs_admin key is a special key stored in the root user's keystore that grants root-level access to all keystores in Root Admin mode. Permissions to access this key can be granted/revoked to specific users or groups using the **`efskeymgr`** command. This requires the aix.security.efs RBAC authorization for users to manage EFS.

> **Note:** The EFS keystore is opened automatically as part of the standard AIX login only when the user’s keystore password matches their login password. This is set up by default during the initial creation of the user’s keystore. Login methods other than the standard AIX login, such as loadable authentication modules and pluggable authentication modules may not automatically open the keystore.

4.2.5 Encryption and inheritance

EFS is a feature of JFS2. The filesystem's *efs* option must be set to **`yes`** (see the **`mkfs`** and **`chfs`** commands). EFS automatically encrypts and decrypts user data. However, if a user has read access to an EFS activated file but does not have the right key, then the user cannot read the file in the normal manner. If the user does not have a valid key, it is impossible to decrypt the data.

All cryptographic functions come from the *CLiC* kernel services and *CLiC* user libraries.

By default, a JFS2 File System is not EFS-enabled. A JFS2 File System must be EFS-enabled before EFS inheritance can be activated or any EFS encryption of user data can take place. A file is created as an encrypted file either explicitly with the **efsmgr** command or implicitly via EFS inheritance. EFS inheritance can be activated either at the File System level, at a Directory level, or both.

The **ls** command lists entries of an encrypted file with a preceding **e**.

The **cp** and **mv** commands can handle metadata and encrypted data seamlessly across EFS-to-EFS and EFS-to-non-EFS scenarios.

The **backup**, **restore**, and **tar** commands and other related commands can back up and restore encrypted data, including the EFS meta-data used for encryption and decryption.

4.2.6 Backup and restore

It is important to properly manage the archiving or backup of the keystores associated with the archived EFS files. You must also manage and maintain the keystore passwords associated with the archived or backup keystores. Failure to do either of these tasks may result in data loss.

When backing up EFS encrypted files, you can use the –Z option with the **backup** command to back up the encrypted form of the file, along with the file's cryptographic meta-data. Both the file data and meta-data are protected with strong encryption. This has the security advantage of protecting the backed-up file through strong encryption. It is necessary to back up the keystore of the file owner and group associated with the file that is being backed up. These key stores are located in the following files:

users keystores	/var/efs/users/*user_login*/*
group keystore	/var/efs/groups//keystore
efsadmin keystor	/var/efs/efs_admin/keystore

Use the **restore** command to restore an EFS backup that was made with the **backup** command and –Z option. The **restore** command ensures that the crypto-meta data is also restored. During the restore process, it is not necessary to restore the backed-up keystores if the user has not changed the keys in their individual keystore. When a user changes their password to open their keystore, their keystore internal key is not changed. Use the **efskeymgr** command to change the keystore internal keys.

If the user's internal, keystore key remains the same, the user can immediately open and decrypt the restored file using their current keystore. However, if the key internal to the user's keystore has changed, the user must open the keystore that was backed up in association with the backed-up file. This keystore can be opened with the **efskeymgr –o** command. The **efskeymgr** command prompts the user for a password to open the keystore. This password is the one used in association with the keystore at time of the backup.

For example, assume that a user Bob's keystore was protected with the password *foo* (the password '*foo*' is not a secure password and only used in this example for simplicity sake) and a backup of Bob's encrypted files was performed in January along with Bob's keystore. In this example, Bob also uses *foo* for his AIX login password. In February, Bob changed his password to *bar*, which also had the effect of changing his keystore access password to *bar*. If, in March, Bob's EFS files were restored, then Bob would be able to open and view these files with his current key store and password, because he did not change the internal key of the keystore.

If however, it was necessary to change the internal key of Bob's keystore (with the **efskeymgr** command), then by default the old keystore internal key is deprecated and left in Bob's keystore. When the user accesses the file, EFS will automatically recognize that the restored file used the old internal key, and EFS will then use the deprecated key to decrypt it.

During this same access instance, EFS will convert the file over to using the new internal key. There is not a significant performance impact in the process, because it is all handled via the key store and file's crypto meta-data, and does not require that the file data is re-encrypted.

If the deprecated internal key is removed through **efskeymgr**, then the old keystore containing the old internal key must be restored and used in conjunction with the files encrypted with this internal key.

This raises the question of how to securely maintain and archive old passwords. There are methods and tools to archive passwords. Generally, these methods involve having a file which contains a list of all old passwords, and then encrypting this file and protecting it with the current keystore, which in turn is protected by the current password. However, IT environments and security policies vary from organization to organization, and consideration and thought should be given to the specific security needs of your organization to develop security policy and practices that are best suited to your environment.

4.2.7 JFS2 EFS internal mechanism –

Each JFS2 EFS encrypted file is associated with a special extended attribute which contains EFS metadata used to validate crypto authority and information used to encrypt and decrypt files such as keys, and crypto algorithm

The content of the extended attribute (EA) content is opaque for JFS2. Both user credentials and EFS meta-data are required to determine a crypto authority (access control) for any given EFS-activated file.

Note: Special attention should be given to situations where a file or data may be lost—for example, removal of the file's EA.

EFS Protection Inheritance

After a directory is EFS-activated, any newly created immediate children are automatically EFS-activated if not manually overridden. The EFS attributes of the parent directory take precedence over the EFS attributes of the file system.

The scope of the inheritance of a directory is exactly one level. Any newly created child also inherits the EFS attributes of its parent if its parent directory is EFS-activated. Existing children maintain their current encrypted or non-encrypted state. The logical inheritance chain is broken if the parent changes its EFS attributes. These changes do not propagate down to the existing children of the directory and must be applied to those directories separately,

Workload Partition considerations

Before enabling or using Encrypted File System within a Workload Partition, EFS must first be enabled on the global system with the **efsenable** command. This enablement only needs to be performed once. Additionally, all filesystems, including EFS-enabled filesystems, must be created from the global system.

4.2.8 Setting up the Encrypted File System

This section details how to set up an encrypted file system. Follow these steps:

1. Install the *clic.rte* fileset. This fileset contains the cryptographic libraries and kernel extension required by EFS. The *clic.rte* fileset is part of the base AIX image from ESS.

   Enable EFS on the system with the command **efsenable –a**. When prompted for a password, it is reasonable to use the root password. Users keystores are created automatically when the user logs in, or re-logs in, after the **efsenable** command has been run.

2. Once **efsenable –a** has been run on a system, then the system is EFS-enabled and the command does not need to be run again.
3. Create an EFS-enabled filesystem using the **crfs** command **–a efs=yes** option. For example,

   ```
   crfs -v jfs2 -m /foo –A yes -a efs=yes -g rootvg -a size=20000
   ```

4. After mounting the filesystem, turn on the cryptographic inheritance on the EFS-enabled filesystem. This can be done with the **efsmgr** command. To continue the previous example where the filesystem */foo* was created, run this command:

   ```
   efsmgr –s –E /foo
   ```

   This allows every file created and used in this filesystem to be an encrypted file.

If a filesystem already exists, it can be enabled for encryption by using **chfs** command, for example:

```
chfs -a efs=yes /foo
```

It is impossible to disable EFS on a filesystem once it is enabled.

The following filesystems can't be converted to EFS-enabled filesystems:

– /
– /opt
– /usr
– /var

From this point forward, when a user or process with an open keystore creates a file on this filesystem, the file will be encrypted. When the user or file reads the file, the file is automatically decrypted for users who are authorized to access the file.

4.2.9 Centralizing access to Encrypted File System keystores

In an enterprise environment, you can centralize your Encrypted File System (EFS) keystores. When you store the databases that control the keystores on each system independently, it can be difficult to manage the keystores. AIX Centralized EFS Keystore allows you to store the user and group keystore databases in Lightweight Directory Access Protocol (LDAP) so that you can centrally manage the EFS keystore.

The Lightweight Directory Access Protocol (LDAP) defines a standard method for accessing and updating information in a directory (a database) either locally or remotely in a client-server model.

You can store all of the AIX EFS keystore databases in LDAP, which includes the following EFS databases:

- User Keystore
- Group Keystore
- Admin Keystore
- Cookies

The AIX operating system provides utilities to help you perform the following management tasks:

- Export local keystore data to an LDAP server
- Configure the client to use EFS keystore data in LDAP
- Control access to EFS keystore data
- Manage LDAP data from a client system

All of the EFS keystore database management commands are enabled to use the LDAP keystore database. If the system-wide search order is not specified in the */etc/nscontrol.conf* file, keystore operations are dependent on the user and group *efs_keystore_access* attribute. If you set the *efs_keystore_access* to **ldap**, the EFS commands perform keystore operations on the LDAP keystore. Table 4-1 describes changes to EFS commands for LDAP.

Table 4-1 EFS command enablement for LDAP

Command	**LDAP information**
Any EFS command	When you set the *efs_keystore_access* attribute to ldap, you do not need to use the special option **-L domain** with any command in order to perform keystore operations on LDAP.
efskeymgr	Includes the **-L load_module** option so that you can perform explicit keystore operations on LDAP.
efsenable	Includes the **-d Basedn** option so that you can perform the initial setup on LDAP for accommodating the EFS keystore. The initial setup includes adding base distinguished names (DNs) for the EFS keystore and creating the local directory structure (/var/efs/).
efskstoldif	Generates the EFS keystore data for LDAP from the following databases on the local system: • /var/efs/users/*username*/keystore • /var/efs/groups/*groupname*/keystore • /var/efs/efs_admin/keystore • Cookies, if they exist, for all the keystores

All of the keystore entries must be unique. Each keystore entry directly corresponds to the DN of the entry that contains the user and group name. The system queries the user IDs (uidNumber), group IDs (gidNumber), and the DNs. The query succeeds when the user and group names match the corresponding DNs. Before you create or migrate EFS keystore entries on LDAP, ensure that the user and group names and IDs on the system are unique.

Exporting Encrypted File System keystore data to LDAP

You must populate the LDAP server with the keystore data to use LDAP as a centralized repository for the Encrypted File System (EFS) keystore.

Before you create or migrate EFS keystore entries on LDAP, ensure that the user and group names and IDs on the system are unique.

To populate the LDAP server with the EFS keystore data, complete the following steps:

1. Install the EFS keystore schema for LDAP on to the LDAP server:
   a. Retrieve the EFS keystore schema for LDAP from the `/etc/security/ldap/sec.ldif` file on the AIX system.
   b. Run the **ldapmodify** command to update the schema of the LDAP server with the EFS keystore schema for LDAP.
2. Run the **efskstoldif** command to read the data in the local EFS keystore files and output the data in a format that is suitable for LDAP.

   To maintain unique keystore access, consider placing the EFS keystore data that resides in LDAP under the same parent distinguished name (DN) as the user and group data.
3. Save the data to a file.
4. Run the **ldapadd -b** command to populate the LDAP server with the keystore data.

Configuring an LDAP client for Encrypted File System keystore

To use Encrypted File System (EFS) keystore data that is stored in LDAP, you must configure a system as an LDAP client.

To configure an LDAP client for EFS keystore, complete the following steps:

1. Run the **`/usr/sbin/mksecldap`** command to configure a system as an LDAP client.

   The **`mksecldap`** command dynamically searches the specified LDAP server to determine the location of the EFS keystore data. Then, it saves the results to the *`/etc/security/ldap/ldap.cfg`* file. The **`mksecldap`** command determines the location for user, group, admin, and EFS cookies keystore data.
2. Complete one of the following steps to enable LDAP as a lookup domain for EFS keystore data:
   - Set the user and group **efs_keystore_access** attribute to **file** or **ldap**.
   - Define the search order for the keystore at the system level by using the `/etc/nscontrol.conf` file.

Table 4-2 shows an example.

Table 4-2 Example configuration for the /etc/nscontrol.conf file

Attribute	Description	Search order
efsusrkeystore	This search order is common for all users.	LDAP, files
efsgrpkeystore	This search order is common for all groups.	files, LDAP
efsadmkeystore	This search order locates the admin keystore for any target keystore.	LDAP, files

Attention: The configuration defined in the */etc/nscontrol.conf* file overrides any values set for the user and group *efs_keystore_access* attribute. The same is true for the user *efs_adminks_access* attribute.

After you configure a system as an LDAP client and enable LDAP as a lookup domain for EFS keystore data, the */usr/sbin/secldapclntd* client daemon retrieves the EFS keystore data from the LDAP server whenever you perform LDAP keystore operations.

4.3 Logical Volume Encryption

Encrypted file system (EFS) provides data encryption at a file system level. The EFS manages the data encryption key at a file level and protects the data encryption key for each user. If you want to avoid the complexity of fine granular control of file system encryption and selective file encryption, you can choose logical volume encryption. Starting with AIX 7.2 TL5, AIX added LV encryption (encryption at logical volume level). This is one choice for data at rest encryption within AIX. Using this feature, you can encrypt the data to prevent data exposure due to lost or stolen hard disk drives or inappropriately decommissioned computers. The term data at rest refers to data that is stored physically in any digital form.

Some organizations are required to show that data at rest is encrypted. A common example is the payment card industry PCI DSS requirement to encrypt sensitive data such as a direct link between card holder name and card number.

Using LV encryption is similar to physical disk encryption. Once operational, the application environment does not even know the data is encrypted. The encryption is only noticeable when the (disk) storage is mounted somewhere else and the data is unreadable. Outside of the configured environment information in the logical volume cannot be accessed.

Using logical volume encryption has the following advantages:

- The data owner controls the encryption keys.
- The data that is transmitted over the network (Fibre Channel or Ethernet) are encrypted and protected. These characteristics are important for virtual servers that are hosted in the cloud environment.

For more information about the LV encryption architecture, see the blog: AIX 72 TL5: Logical Volume Encryption[2].

Logical volume encryption (LV encryption) is simple to use and is transparent to the applications. Once the system has been booted and an authorized process or user is active on the system the data is accessible to authorized users based on classic access controls such as ACLs.

When enabled—by default data encryption is not enabled in logical volumes—each LV is encrypted with a unique key. Data encryption must be enabled at the volume group level before you can enable the data encryption option at the logical volume level. The logical volume data is encrypted as the data is written to the physical volume. and decrypted when it is read from the physical volume.

Enabling LV encryption creates one data encryption key for each logical volume. The data encryption key is protected by storing the keys separately in other data storage devices. The following types of key protection methods are supported:

- Paraphrase
- Key file
- Cryptographic key server
- Platform keystore (PKS) which is available in IBM PowerVM firmware starting at firmware level FW950

[2] https://community.ibm.com/community/user/power/blogs/xiaohan-qin1/2020/11/23/aix-lv-encryption

LV encryption enhancements in AIX 7.3

Logical volume encryption was further enhanced in AIX 7.3 and starting from AIX 7.3, the following enhancements are added to the LV encryption function:

- You can encrypt LVs in the root volume group (rootvg) that are used in the boot process. The LV encryption option must be selected during the installation of the base operating system. For more information, see BOS installation options.
- After you install the base operating system, you can use the **hdcryptmgr** conversion commands to change the encryption setting of an LV. However, the conversion of an LV in the rootvg is different from the conversion of an LV in a user volume group.
  - When you run the **hdcryptmgr** conversion command to change the encryption status of an LV in a rootvg, the **hdcryptmgr** command creates an LV to store the conversion recovery data.
  - When you run the **hdcryptmgr** conversion command to change the encryption status of an LV in a user volume group, the **hdcryptmgr** command stores the conversion recovery data in a file that is in the /var/hdcrypt directory.

    Therefore, the rootvg must have at least one free logical partition for successful conversion. When the conversion status of the encryption is successful, the LV that contains the conversion recovery data is deleted.
- When the rootvg is varied on, the network is not available. Hence, the platform keystore (PKS) authentication method must be available for LVs that are used in the boot process.
  - If the PKS authentication method is not available for an encrypted LV in the rootvg, the LV remains locked and thus not accessible until it is explicitly unlocked later.
  - You cannot delete a valid PKS authentication method from an LV in the rootvg that is used in the boot process.
  - If you convert an unencrypted LV, which is used in the boot process, to an encrypted LV, the PKS authentication method is automatically added to the LV.
  - If the PKS authentication method is not available or is corrupted for an encrypted LV that is used in the boot process, you must boot the operating system in maintenance mode and repair the PKS authentication method before you can resume the normal boot operation.
- The command in s are enhanced to support LV encryption: **cplv**, **splitvg**, **splitlvcopy**, **chlvcopy**, **snapshot**, **savevg**, and **restvg**.
- You can encrypt an LV in concurrent mode. If you change the encryption status of an LV in a node that is in concurrent mode, you cannot access the other nodes until the encryption conversion is complete.
- AIX 7.3 TL1 supports Hyper Protect Crypto Services (HPCS) for AIX logical volume encryption. To use HPCS with AIX, you must provision Power Systems Virtual Server. The **keysvrmgr** command provides options to manage the integration.

4.3.1 LV encryption commands

You can use the command in s to manage encryption keys and key server information.

The hdcryptmgr command

The **hdcryptmgr** utility manages the encrypted LVs that includes the tasks such as displaying logical volume and volume encryption information, controlling authentication, and many other functions. The utility and its help messages are built in a hierarchical and self-explanatory manner. Example 4-2 on page 83 shows a summary of the command usage. For a detailed manual page, see `hdcryptmgr` command.

Example 4-2   Help page from hdcryptmgr command

```
# hdcryptmgr -h
Usage: hdcryptmgr <action> <..options..>

Display :
showlv        : Displays LV encryption status
showvg        : Displays VG encryption capability
showpv        : Displays PV encryption capability
showmd        : Displays encryption metadata related to device
showconv      : Displays status of all active and stopped conversions

Authentication control :
authinit      : Initializes master key for data encryption
authunlock    : Authenticates to unlock master key of the device
authadd       : Adds additional authentication methods
authcheck     : Checks validity of an authentication method
authdelete    : Removes an authentication method
authsetrvgpwd : Adds "initpwd" passphrase method to all rootvg's LVs

PKS management :
pksimport     : Import the PKS keys
pksexport     : Export the PKS keys
pksclean      : Removes a PKS key
pksshow       : Displays PKS keys status

Conversion :
plain2crypt   : Converts a LV to encrypted
crypt2plain   : Converts a LV to not encrypted

PV encryption management :
pvenable      : Enables the Physical Volume Encryption
pvdisable     : Disables the Physical Volume Encryption
pvsavemd      : Save encrypted physical volume metadata to a file
pvrecovmd     : Recover encrypted physical volume metadata from a file
```

The keysvrmgr command

For the key server method, you can use the **keysvrmgr** utility to manage Object Data Manager (ODM) entries that are associated with the key server information such as the key server hostname or IP address, the connection port, and certification location. Example 4-3 shows a summary of the command usage. For a detailed manual page see **keysvrmgr** command.

Example 4-3   Help page from keysvrmgr command

```
# keysvrmgr -h
Usage: keysvrmgr <action> [-h] -t <server_type> <options> server_name
Manage ODM data for key server and HPCS.

<action> is one of the following:
add     : Add a new key server or HPCS to ODM.
modify  : Modify a key server or HPCS ODM record.
remove  : Remove a keyserver or HPCS ODM record.
show    : Display key server or HPCS ODM records.
verify  : Verify a HPCS ODM record (HPCS only).
rekey   : Generate a new API key for a HPCS ODM record (HPCS only).

<server_type> is one of the following:
keyserv : For (KMIP compliant) key management server.
hpcs    : For IBM Cloud Hyper Protect Crypto Services.

For more details on <options> run : keysvrmgr <action> -h
```

4.3.2 Prerequisites for using LV encryption

Before you implement logical volume encryption, make sure that you meet the following prerequisites.

- Use AIX 7.2.5 or later to encrypt a logical volume.
- These filesets must be installed to encrypt the LV data. These filesets are included in the base operating system.
  - bos.hdcrypt
  - bos.kmip_client
  - bos.rte.lvm
  - security.acf
  - openssl.base
  - oss.lib.libcurl
  - oss.lib.libjson-c

> **Note:** The bos.hdcrypt and bos.kmip_client filesets are not installed automatically when you run the smit update_all command or during an operating system migration operation. You must install it separately from your software source such as a DVD or an ISO image.

4.3.3 Creating and authenticating an encrypted logical volume

To create an encrypted logical volume, complete the following procedures:

1. Create an encryption-enabled volume group.
2. Create an encryption-enabled logical volume.
3. Authenticate the primary encryption key of the logical volume.

Create an encryption-enabled volume group

To create an encryption-enabled volume group, complete the following steps:

1. Create a volume group in which the data encryption option is enabled by running this command:

   ```
   mkvg -f -y testvg -k y hdisk2
   ```

   where *testvg* is the name of the new volume group and *hdisk2* is the physical volume that is used for the volume group.

2. Check the details of the new volume group by running the command shown in Example 4-4. Note that *ENCRYPTION* is set to **yes.**

Example 4-4 Showing the volume group settings

```
# lsvg testvg
VOLUME GROUP:       testvg                   VG IDENTIFIER: 00fb294400004c0000000176437c6663
VG STATE:           active                   PP SIZE:        8 megabyte(s)
VG PERMISSION:      read/write               TOTAL PPs:      637 (5096 megabytes)
MAX LVs:            256                      FREE PPs:       637 (5096 megabytes)
LVs:                0                        USED PPs:       0 (0 megabytes)
OPEN LVs:           0                        QUORUM:         2 (Enabled)
TOTAL PVs:          1                        VG DESCRIPTORS: 2
STALE PVs:          0                        STALE PPs:      0
ACTIVE PVs:         1                        AUTO ON:        yes
MAX PPs per VG:     32512
MAX PPs per PV:     1016                     MAX PVs:        32
LTG size (Dynamic): 512 kilobyte(s)          AUTO SYNC:      no
HOT SPARE:          no                       BB POLICY:      relocatable
PV RESTRICTION:     none                     INFINITE RETRY: no
DISK BLOCK SIZE:    512                      CRITICAL VG:    no
FS SYNC OPTION:     no                       CRITICAL PVs:   no
ENCRYPTION:         yes
```

3. Check the encryption state of varied on volume groups by running the command shown in Example 4-5.

Example 4-5 Check encryption state of volume groups

```
# hdcryptmgr showvg
VG NAME / ID         ENCRYPTION ENABLED
testvg                        yes
rootvg                        no
```

4. Check the volume group encryption metadata by running the command shown in Example 4-6.

Example 4-6 Validating volume group metadata

```
# hdcryptmgr showmd testvg
.....    Mon Dec  7 21:19:00 2020
.....    Device type : VG
.....    Device name : testvg
.....
============== B: VG HEADER ===============
Version                      : 0
Timestamp                    : Mon Dec  7 21:16:04 2020
Default data crypto algorithm: AES_XTS
Default MasterKey size       : 16 bytes
Auto-auth (during varyonvg)  : Enabled
============== E: VG HEADER ===============
============== B: VG TRAILER ==============
Timestamp        : Mon Dec  7 21:16:04 2020
============== E: VG TRAILER ==============
```

Create an encryption-enabled logical volume

To create an encryption-enabled logical volume, complete the following steps:

1. Create a logical volume in which the data encryption option is enabled by running the command shown in Example 4-7.

Example 4-7 Creating encrypted logical volume

```
# mklv -k y -y testlv testvg 10
testlv
mklv: Please run :
hdcryptmgr authinit lvname [..] to define LV encryption options.
```

2. Check the details of the new volume group by running the command shown in Example 4-8.

Example 4-8 Validate logical volume details

```
# lslv testlv
LOGICAL VOLUME:     testlv                             VOLUME GROUP:   testvg
LV IDENTIFIER:      00fb294400004c0000000176437c6663.1 PERMISSION:     read/write
VG STATE:           active/complete                    LV STATE:       closed/syncd
TYPE:               jfs                                WRITE VERIFY:   off
MAX LPs:            512                                PP SIZE:        8 megabyte(s)
COPIES:             1                                  SCHED POLICY:   parallel
LPs:                10                                 PPs:            10
STALE PPs:          0                                  BB POLICY:      relocatable
INTER-POLICY:       minimum                            RELOCATABLE:    yes
INTRA-POLICY:       middle                             UPPER BOUND:    32
MOUNT POINT:        N/A                                LABEL:          None
MIRROR WRITE CONSISTENCY: on/ACTIVE
EACH LP COPY ON A SEPARATE PV ?: yes
Serialize IO ?:     NO
INFINITE RETRY:     no                                 PREFERRED READ: 0
ENCRYPTION:         yes
```

3. Check the authentication state of the logical volume by running the command shown in Example 4-9.

Example 4-9 Validate authentication state of logical volume

```
# hdcryptmgr showlv testlv
LV NAME   CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)  CONVERSION
testlv        yes              no            100             done
```

Authenticate the primary encryption key of the logical volume

To authenticate the primary encryption key of the logical volume, complete the following steps:

1. Initialize the primary key for an encrypted logical volume by running the command shown in Example 4-10. The logical volume is not accessible until the first passphrase method is initialized.

Example 4-10 Setting authentication for logical volume

```
# hdcryptmgr authinit testlv
Enter Passphrase:
Confirm Passphrase:
Passphrase authentication method with name "initpwd" added successfully.
```

2. Check the authentication status and authentication methods for the logical volume by running the command shown in Example 4-11.

Example 4-11 Check authentication for the LV

```
# hdcryptmgr showlv testlv -v
LV NAME    CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv         yes              yes             100              done

-- Authentication methods ------------
INDEX         TYPE              NAME
#0            Passphrase        initpwd
```

3. Vary off and vary on the volume group by running the command in s:

```
# varyoffvg testvg
# varyonvg testvg
```

4. Check the authentication status of the logical volume by running the command shown in Example 4-12.

Example 4-12 Check authentication of LV after vary on

```
# hdcryptmgr showlv testlv
LV NAME      CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv           yes              no              100              done
```

The output shows that the logical volume **testlv** is not authenticated.

5. Unlock the authentication of the logical volume by running the command: shown in Example 4-13.

Example 4-13 Unlock logical volume

```
# hdcryptmgr authunlock testlv
Enter Passphrase:
Passphrase authentication succeeded.
```

6. Check the authentication state of the logical volume again as shown in Example 4-14.

Example 4-14   Validate authentication status

```
# hdcryptmgr showlv testlv
LV NAME      CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv           yes              yes             100              done
```

4.3.4 Adding the platform keystore (PKS) authentication method

To add the Platform keystore (PKS) authentication method, complete the following steps:

1. Check the LPAR PKS status by running the command shown in Example 4-15.

Example 4-15   Check status of PKS

```
# hdcryptmgr pksshow
3020-0349 PKS is not supported or PKS is not activated.
3020-0218 hdcrypt driver service error. QUERY_PKS service failed with error 124: An attempt
was made to set an attribute to an unsupported value.
```

The output in this example shows that the PKS is not activated. The keystore size of a logical partition is set to 0 by default.

2. Shut down the LPAR and increase the keystore size in the associated HMC. The keystore size is in the range 4 KB to 64 KB. You cannot change the value of the keystore size when the LPAR is active.
3. Check the LPAR PKS status again by running the command shown in Example 4-16.

Example 4-16   Recheck PKS status

```
# hdcryptmgr pksshow
PKS uses 32 bytes on a maximum of 4096 bytes.
PKS_Label (LVid) Status
PKS_Label (objects)
```

4. Add the PKS authentication method to the logical volume by running the command shown in Example 4-17.

Example 4-17   Add PKS authentication method

```
# hdcryptmgr authadd -t pks -n pks1 testlv
PKS authentication method with name "pks1" added successfully.
```

5. Check the encryption status of the logical volume by running the command shown in Example 4-18.

Example 4-18   Check status of LV encryption

```
# hdcryptmgr showlv testlv -v
LV NAME        CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv         yes              yes              100              done
-- Authentication methods ------------
INDEX         TYPE                 NAME
#0            Passphrase           initpwd
#1            PKS                  pks1
```

6. Check the PKS status by running the command shown in Example 4-20.

Example 4-19 Check PKS status

```
# hdcryptmgr pksshow
PKS uses 116 bytes on a maximum of 4096 bytes.
PKS_Label (LVid) Status
00fb294400004c0000000176437c6663.1 VALID KEY
PKS_Label (objects)
```

PKS is an automatic authentication method that means the **varyonvg** command automatically unlocks the authentication of the logical volume.

7. Vary off the volume group by running this command:

   ```
   # varyoffvg testvg
   ```

8. Check the PKS status by running the command shown in Example 4-20.

Example 4-20 Check PKS status again

```
# hdcryptmgr pksshow
PKS uses 116 bytes on a maximum of 4096 bytes.
PKS_Label (LVid) Status
00fb294400004c0000000176437c6663.1 UNKNOWN
PKS_Label (objects)
```

9. Vary on the volume group by running this command:

   ```
   # varyonvg testvg
   ```

10. Check the encryption status of the logical volume by running the command shown in Example 4-21

Example 4-21 Validate LV encryption status

```
# hdcryptmgr showlv testlv
LV NAME         CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv          yes              yes              100              done
```

4.3.5 Adding the key server authentication method

You can use any Key Management Interoperability Protocol (KMIP) compliant key management server to use this type of authentication method. In this example, the AIX logical partition is installed and configured with the IBM Security Key Lifecycle Manager (SKLM) V4.0 for AIX. The Security Key Lifecycle Manager key is used as an encryption key server.

To add the key server authentication method, complete the following steps:

1. Check the key servers in the LPAR by running the command shown in Example 4-22.

Example 4-22 Show no key server defined.

```
# keysvrmgr show
3020-0279 No key server in database
```

2. Add encryption key server with the name **keyserver1** by running the command shown in Example 4-23.

Example 4-23 Add key server

```
# keysvrmgr add -i 9.X.X.X -s /tmp/sklm_cert.cer -c /tmp/ssl_client_cer.p12 keyserver1
Key server keyserver1 successfully added
```

3. Check the key servers in the LPAR again by running the command shown in Example 4-24.

Example 4-24 Validate key server is defined

```
# keysvrmgr show
List of key servers:
ID                    PWD            IP:PORT
keyserver1            N              9.X.X.X:5696
```

4. Check the encryption key server information that is saved in the ODM KeySvr object class by running the command shown in Example 4-25.

Example 4-25 Check key server definition

```
# odmget KeySvr
KeySvr:
        keysvr_id = "keyserver1"
        ip_addr = "9.X.X.X"
        port = 5696
        svr_cert_path = "/tmp/sklm_cert.cer"
        cli_cert_path = /tmp/ssl_client_cer.p12 "
        flags = 0
```

5. Add the key server authentication method to the logical volume by running the command shown in Example 4-26.

Example 4-26 Add key server authentication method to LV

```
# hdcryptmgr authadd -t keyserv -n key1_testlv -m keyserver1 testlv
Keyserver authentication method with name "key1_testlv" added successfully.
```

6. Check the encryption status of the logical volume by running the command shown in Example 4-28 on page 90.

Example 4-27 Check encryption status of LV

```
#hdcryptmgr showlv -v testlv
LV NAME          CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv           yes              yes              100              done
-- Authentication methods ------------
INDEX                 TYPE              NAME
#0                    Passphrase        initpwd
#1                    PKS               pks1
#2                    Keyserver         key1_testlv
```

4.3.6 Adding key file authentication method

To add key file authentication method, complete the following steps:

1. Create a file named *testfile* that contains the passphrase text by running the command shown in Example 4-28.

Example 4-28 Create file

```
# cat /testfile
Add1ng Key f1le authent1cation meth0d
```

2. Add the key file authentication method to the logical volume by running the command in Example 4-29.

Example 4-29 Add key authentication

```
# hdcryptmgr authadd -t keyfile -n key1_file -m /testfile testlv
Keyfile authentication method with name "key1_file" added successfully.
```

3. Check the contents of the testfile file by running the command in Example 4-30.

Example 4-30 Check contents of file

```
# cat /testfile
Add1ng Key f1le authent1cation meth0d
00fb294400004c0000000176437c6663.1 xdxKjlJvZU+f9lFTgSM63kGoIoKW6Yxc+bKrk5GgCzc=
```

4. Check the encryption status of the logical volume by running the command in Example 4-31.

Example 4-31 Validate encryption status

```
# hdcryptmgr showlv testlv -v
LV NAME          CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv           yes              yes              100              done
-- Authentication methods ------------
INDEX              TYPE               NAME
#0                 Passphrase         initpwd
#1                 PKS                pks1
#2                 Keyserver          key1_testlv
#3                 Keyfile            key1_file
```

4.3.7 Adding passphrase authentication method

To add the passphrase authentication method, complete the following steps:

1. Add the passphrase authentication method to the logical volume by running the command shown in Example 4-32.

Example 4-32 Add passphrase

```
# hdcryptmgr authadd -t pwd -n test_pwd testlv
Enter Passphrase:
Confirm Passphrase:
Passphrase authentication method with name "test_pwd" added successfully.
```

2. Check the encryption status of the logical volume by running the command shown in Example 4-33.

Example 4-33 Check encryption status of LV

```
# # hdcryptmgr showlv testlv -v
LV NAME              CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv               yes              yes              100              done
-- Authentication methods ------------
INDEX         TYPE            NAME
#0            Passphrase      initpwd
#1            PKS             pks1
#2            Keyserver       key1_testlv
#3            Keyfile         key1_file
#4            Passphrase      test_pwd
```

4.3.8 Migrating the PKS to another LPAR before the volume group is migrated

To migrate the platform keystore (PKS) to another LPAR, complete the following steps:

1. Export the PKS keys into another file by running the command shown in Example 4-34.

Example 4-34 Export PKS keys

```
# hdcryptmgr pksexport -p /tmp/pksexp testvg
Enter Passphrase:
Confirm Passphrase:
1 PKS keys exported.
```

2. Import the volume group to another LPAR by running this command:

   `# importvg -y testvg hdisk2`

3. Check the encryption status of the logical volume by running the command shown in Example 4-35.

Example 4-35 Check encryption status

```
# hdcryptmgr showlv testlv -v
LV NAME          CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv           yes              yes              100              done
-- Authentication methods ------------
INDEX         TYPE             NAME
#0            Passphrase       initpwd
#1            PKS              pks1
#2            Keyserver        key1_testlv
#3            Keyfile          key1_file
#4            Passphrase       test_pwd
```

4. Check whether the authentication method is valid and accessible by running the command shown in Example 4-36.

Example 4-36 Test authentication

```
# hdcryptmgr authcheck -n pks1 testlv
3020-0199 Key does not exist in PKS storage.
3020-0127 hdcryptmgr authcheck failed for LV testlv.
```

5. Move the PKS key file to a new LPAR and run the command shown in Example 4-37.

Example 4-37 Import key

```
#  hdcryptmgr pksimport -p /tmp/pksexp testvg
Enter Passphrase:
3020-0341 Key having LVid 00fb294400004c0000000176437c6663.1 is successfully imported in LV
testlv.
1 PKS keys imported.
```

6. Check whether the authentication method is valid and accessible by running the command shown in Example 4-38.

Example 4-38 Validate authentication method

```
# hdcryptmgr authcheck -n pks1 testlv
PKS authentication check succeeded.
```

4.3.9 Changing the encryption policy of the volume group

Encryption metadata is saved at the end of each disk in the volume group. Enabling the volume group encryption requires free physical partitions on each disk in the volume group.

1. Change the data encryption option of the volume group by running the command shown in Example 4-39.

Example 4-39 Change encryption option

```
# chvg -k y testvg
0516-1216 chvg: Physical partitions are being migrated for volume group
                descriptor area expansion.  Please wait.
```

2. Check the details of the volume group by running the command shown in Example 4-40.

Example 4-40 List VG details

```
# lsvg testvg
VOLUME GROUP:       testvg              VG IDENTIFIER:  00fb294400004c000000017648ff8d32
VG STATE:           active              PP SIZE:        8 megabyte(s)
VG PERMISSION:      read/write          TOTAL PPs:      636 (5088 megabytes)
MAX LVs:            256                 FREE PPs:       506 (4048 megabytes)
LVs:                1                   USED PPs:       130 (1040 megabytes)
OPEN LVs:           0                   QUORUM:         2 (Enabled)
TOTAL PVs:          1                   VG DESCRIPTORS: 2
STALE PVs:          0                   STALE PPs:      0
ACTIVE PVs:         1                   AUTO ON:        yes
MAX PPs per VG:     32512
MAX PPs per PV:     1016                MAX PVs:        32
LTG size (Dynamic): 512 kilobyte(s)     AUTO SYNC:      no
HOT SPARE:          no                  BB POLICY:      relocatable
PV RESTRICTION:     none                INFINITE RETRY: no
DISK BLOCK SIZE:    512                 CRITICAL VG:    no
FS SYNC OPTION:     no                  CRITICAL PVs:   no
ENCRYPTION:         yes
```

4.3.10 Changing the encryption policy of the logical volume

To change the encryption policy, complete the following steps:

Note: This capability is for experimental use only.

1. Enable the logical volume encryption by running the command shown in Example 4-41.

Example 4-41 Enable LV encryption

```
# hdcryptmgr plain2crypt testlv
Enter Passphrase:
Confirm Passphrase:
Passphrase authentication method with name "initpwd" added successfully.
Created recovery file : /var/hdcrypt/conv.004200021607542921
In case of error or if the conversion is canceled, this file may be
necessary to be able to recover the LV. If the conversion is fully
successful, then the file will be removed automatically
Successfully converted LV testlv to an encrypted LV.
```

This command performs the following operations:

- Enables the encryption policy of the logical volume
- Initializes the master-key and encryption metadata for an encrypted logical volume
- Encrypts the data in the logical volume

2. Check the details of the logical volume by running the command shown in Example 4-42.

Example 4-42 List the LV details

```
# lsvg testvg
VOLUME GROUP:       testvg                   VG IDENTIFIER:  00fb294400004c000000017648ff8d32
VG STATE:           active                   PP SIZE:        8 megabyte(s)
VG PERMISSION:      read/write               TOTAL PPs:      636 (5088 megabytes)
MAX LVs:            256                      FREE PPs:       506 (4048 megabytes)
LVs:                1                        USED PPs:       130 (1040 megabytes)
OPEN LVs:           0                        QUORUM:         2 (Enabled)
TOTAL PVs:          1                        VG DESCRIPTORS: 2
STALE PVs:          0                        STALE PPs:      0
ACTIVE PVs:         1                        AUTO ON:        yes
MAX PPs per VG:     32512
MAX PPs per PV:     1016                     MAX PVs:        32
LTG size (Dynamic): 512 kilobyte(s)          AUTO SYNC:      no
HOT SPARE:          no                       BB POLICY:      relocatable
PV RESTRICTION:     none                     INFINITE RETRY: no
DISK BLOCK SIZE:    512                      CRITICAL VG:    no
FS SYNC OPTION:     no                       CRITICAL PVs:   no
ENCRYPTION:         yes
```

3. Check the encryption status of the logical volume by running the command shown in Example 4-43.

Example 4-43 Check encryption status

```
# hdcryptmgr showlv testlv -v
LV NAME              CRYPTO ENABLED   AUTHENTICATED    ENCRYPTION (%)   CONVERSION
testlv               yes              yes              100              done
-- Authentication methods ------------
INDEX         TYPE                 NAME
#0                 Passphrase    initpwd
```

4.3.11 Best practices

When setting up and using LV encryption we suggest that you follow these suggested practices.

- Use an inline log device for any file system that is created from an encrypted logical volume.
- If the file system is created with an external log device and the log device is shared across multiple file systems, unlock the authentication (`hdcryptmgr authunlock)` for all encrypted logical volumes before you mount the file system.
- Use one of the non-PKS authentication methods to unlock the authentication of the snapshot volume group.
- To copy an encrypted logical volume by using the `cplv` command:
  - Create a logical volume in which encryption is enabled and
  - Use this newly created logical volume as the destination logical volume to copy the source logical volume.

4.3.12 Limitations of LV encryption

If an LV is encrypted, the following LV commands or functions are not supported:

AIX Live Update The Live Update operation is not supported if LV encryption is enabled.

I/O serialization The I/O serialization is not guaranteed while LV encryption conversion is in progress.

4.3.13 File system consideration for LV encryption

Consider the following items when you create or modify file systems that are associated with an encrypted LV:

- When you create or mount a file system on to an encrypted LV, ensure that the encrypted LV is unlocked and activated.
- If an encrypted LV, which is hosting a file system by using the Network File System (NFS) /etc/exports file, is not unlocked during system boot, the mount operation of the file system fails and the table of physical file systems in the /etc/exports file is not updated. After the encrypted LV is unlocked and the file system is mounted, you can run the **exportfs -a** command to update the /etc/exports file.
- In Enhanced Journaled File System (JFS2), you can use a single log device across multiple file systems. If the log device is shared across multiple file systems and if the LV that is used by file systems is encrypted, the LV must be unlocked before mounting the file systems.

4.4 Physical Volume Encryption

Physical volume (PV) encryption protects user data by encrypting data that is written to the physical volume. The base operating system performs physical volume data encryption and decryption during I/O operations. The data is encrypted before it is sent to an external storage area network (SAN) device to protect data on the SAN. Physical volume encryption also protects data exposure because of lost or stolen hard disk drives or because of inappropriately decommissioned computers or storage devices. Applications that perform I/O operations can use the protected data without any modifications. The encrypted physical volumes can be used in the same way as unencrypted physical volumes. However, the *rootvg* volume group cannot contain any encrypted physical volumes.

With AIX 7.3 TL1, IBM continues to address clients' need to protect data by introducing encrypted physical volumes. This capability encrypts data at rest on disks, and since the data is encrypted in the OS, the disk data in flight is encrypted as well.

You must install the following filesets to encrypt the physical volume data. These filesets are included in the base operating system.

- bos.hdcrypt
- bos.kmip_client
- security.acf
- openssl.base

AIX has historically supported encrypted files using the Encrypted File System (EFS). More recently, AIX 7.2 TL 5 introduced support for logical volume encryption, as detailed in 4.3, "Logical Volume Encryption" on page 81.

Now, AIX offers a new level of security with physical volume encryption. This feature allows for the encryption of entire physical volume, providing enhanced protection for applications that don't rely on volume groups or logical volumes, such as certain database applications. However, it's also possible to create volume groups and logical volumes on encrypted disks.

Physical volume encryption leverages the infrastructure developed for logical volume encryption. Therefore, many of the concepts and features described in previous section on logical volume encryption also apply to encrypted physical volumes. For instance, both types support the same key management functions.

The **hdcryptmgr** command is used to manage encrypted physical volumes, and the *hdcrypt* driver handles the encryption process. While the core functionality remains similar, some new options and actions have been added to the **hdcryptmgr** command specifically for physical volume encryption.

4.4.1 Configuring PV encryption

To use physical volume encryption, the disk must first be formatted for encryption. This operation erases any data on the disk. There is no support for directly encrypting existing data on a disk. To encrypt existing data, you can allocate a new physical volume, enable encryption on the new physical volume, and then copy the existing data to the new physical volume.

The size of the encrypted physical volume is smaller than the size of the physical volume before encryption because the encryption feature reserves some space on the physical volume for the encryption process.

The command to enable encryption on disk hdisk10 is **hdcryptmgr pvenable hdisk10**. This command prompts the user for a passphrase to use to unlock the disk and then reserves some space at the beginning of the disk for metadata. As with logical volume encryption, a data encryption key is created automatically when the disk is initialized for encryption. The **pvenable** action also prompts the user to add a passphrase wrapping key to encrypt the data encryption key. Additional wrapping keys may be added using the **authadd** action of the **hdcryptmgr** command. Note that since space is reserved for metadata, the space available for user data on an encrypted physical volume is slightly smaller than the total size of the physical volume.

Once the disk is initialized for encryption and unlocked, it may be used just as any other disk in AIX—the exception is that encrypted disks cannot be used as part of the *rootvg* volume group. As the OS writes data to the disk, the data is encrypted; when data is read from the disk it is decrypted before being passed to the user.

4.4.2 Using encrypted physical volumes

Encrypted physical volume supports the same methods of key storage and retrieval as encrypted logical volumes (LV). The key can be a typed passphrase, can be obtained from platform keystore (PKS), or can be obtained from a network key manager.

When the key is stored in a PKS or in a network key manager, the physical volume is unlocked automatically during the boot process. The **authunlock** action parameter of the **hdcryptmgr** command can be used to manually unlock an encrypted physical volume. Any attempts to perform I/O operation on a locked encrypted physical volume fails with a permission denied error until that physical volume is unlocked.

If the AIX LPAR is rebooted, encrypted disks that use only the passphrase wrapping key protection method must be manually unlocked using the **hdcryptmgr authunlock** action. If one of the other methods, such as using a key server or PKS, has been added to the disk using the **authadd** action, AIX attempts to automatically unlock the disk during boot. Any attempt to do I/O to an encrypted disk that is still locked fails. Figure 4-1 illustrates the encryption process.

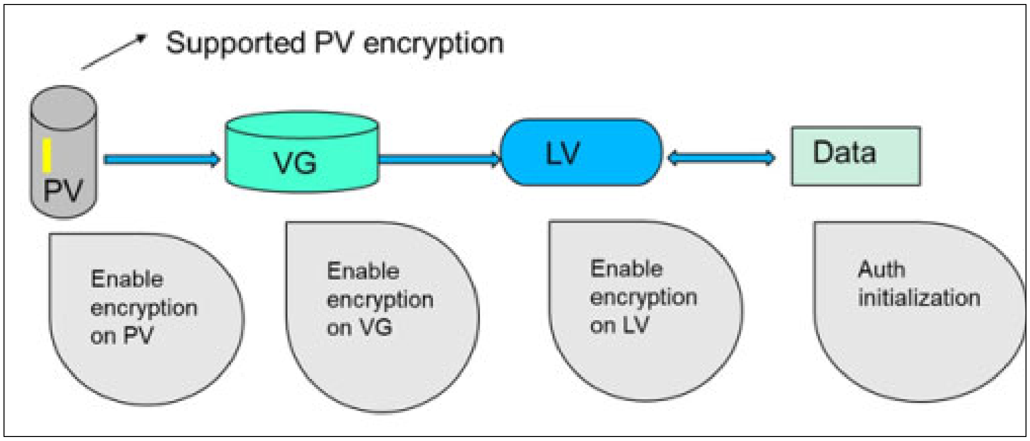
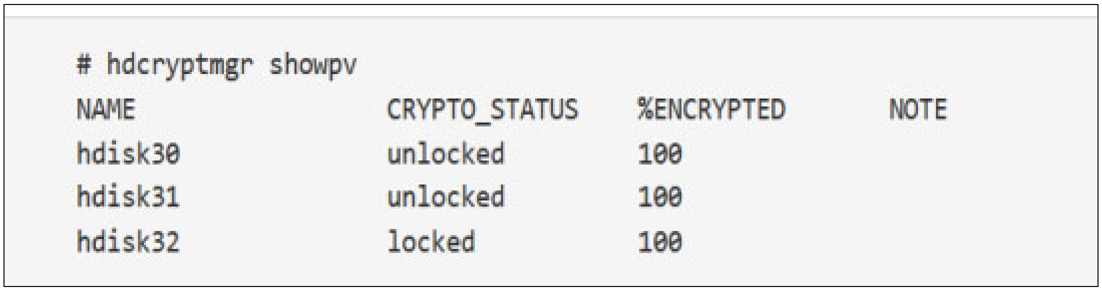


Figure 4-1 Physical volume Encryption

Figure 4-2 shows the output of the **showpv** and **showmd** actions of the **hdcryptmgr** command. The **showpv** output displays three encrypted disks, two that are unlocked (able to be read from or written to) and one that is locked. The locked disk requires **hdcrpytmgr authunlock hdisk32** before it is usable.



```
# hdcryptmgr showpv
NAME                CRYPTO_STATUS   %ENCRYPTED      NOTE
hdisk30             unlocked        100
hdisk31             unlocked        100
hdisk32             locked          100
```

Figure 4-2 Output from hdcryptmgr showpv

4.4.3 Limitations of encrypted PV

The encrypted physical volumes have the following restrictions:

- The *rootvg* volume group must not contain any encrypted physical volume. If *rootvg* contains one or more encrypted physical volumes, the AIX boot process fails. The **mkvg** and **extendvg** commands prevent using encrypted physical volumes with *rootvg*.
- The existing physical volumes cannot be converted from unencrypted physical volumes to encrypted physical volumes, or vice versa. Enabling encryption on a physical volume deletes all the existing data on that disk.
- Physical volume encryption requires additional disk attributes that are provided by the AIX operating system. If a disk is defined by using object data manager (ODM) definitions from another storage vendor, new ODM definitions from that vendor must be acquired to support physical volume encryption.
- Encrypted physical volumes can be shared with other AIX logical partitions that are running AIX 7.3 Technology Level 1, or later. Sharing an encrypted physical volume with an older level of AIX corrupts data because the older level of AIX does not recognize that the physical volume is encrypted.
- Physical volumes that are encrypted with PKS authentication can be used as a traditional dump device if it does not belong to the *rootvg* volume group.
- Encrypted physical volumes cannot be used as the destination disk when you use the **alt_disk_copy** and **alt_disk_mksysb** commands because the *rootvg* volume group does not support the encrypted physical volumes.
- Only SCSI physical volumes can be encrypted. You cannot encrypt NVMe or vPMEM disks.
- The same AIX operating system image cannot use geographical logical volume manager (GLVM) or AIX storage data caching (**cache_mgt** command) with other SCSI disks while using encrypted physical volumes. GLVM or storage data caching can be used with NVMe disks or with vPMEM disks.

4.4.4 Disk backup considerations for encrypted PV

The various methods of backing up data on the physical disk have different characteristics when encrypted physical volumes are used.

If the data backup operation is running in the operating system instance, the operating system reads data and decrypts that data before sending it to the backup software. The backup media contains the decrypted user data. The metadata related to encryption is not stored in the backup media. If this backup data is restored to another physical volume, data is encrypted only if encryption is enabled for that physical volume. If encryption is not enabled for the destination physical volume, the restored data is not encrypted and can be used directly even by older levels of AIX.

If data is backed up by using a storage device such as snapshot or IBM FlashCopy®, the data that is backed up is encrypted. The backup data in the storage device includes both the encryption metadata and the encrypted user data. The storage-based backup is a block-for-block copy of the encrypted data and the storage cannot determine that the data is encrypted by the operating system.

For additional information on physical volume encryption in AIX see:

- https://community.ibm.com/community/user/power/blogs/gary-domrow/2023/02/08/understanding-aix-physical-volume-encryption?CommunityKey=daa942cb-b783-4fd3-ba27-a2d7462f9530
- https://www.ibm.com/docs/en/aix/7.3?topic=system-encrypted-physical-volumes
- https://www.ibm.com/docs/en/aix/7.3?topic=io-encrypting-physical-volumes

4.5 AIX Access Control Lists

A Standard UNIX environment uses only three attributes to determine which user has access to a file or a directory:

- owner of the file or the directory
- group of the file or the directory
- permissions mode to define what the owner, the group and all other can do with the file (read, write, execute).

In addition to the standard UNIX discretionary access control (DAC) AIX has Access Control Lists (ACL). ACLs enable you to define access to files and directories more granularly. Typically an ACL consists of series of entries called an Access Control Entry (ACE). Each ACE defines the access rights for a user in relationship to the object.

When an access is attempted, the operating system will use the ACL associated with the object to see whether the user has the rights to do so. These ACLs and the related access checks form the core of the Discretionary Access Control (DAC) mechanism supported by AIX.

The operating system supports several types of system objects that allow user processes to store or communicate information. The most important types of access controlled objects are as follows:

- Files and directories
- Named pipes
- IPC objects such as message queues, shared memory segments, and semaphores

All access permission checks for these objects are made at the system call level when the object is first accessed. Because System V Interprocess Communication (SVIPC) objects are accessed statelessly, checks are made for every access. For objects with file system names, it is necessary to be able to resolve the name of the actual object. Names are resolved either relatively (to the process' working directory) or absolutely (to the process' root directory). All name resolution begins by searching one of these directories.

The discretionary access control mechanism allows for effective access control of information resources and provides for separate protection of the confidentiality and integrity of the information. Owner-controlled access control mechanisms are only as effective as users make them. All users must understand how access permissions are granted and denied, and how these are set.

For example, an ACL associated with a file system object (file or directory) could enforce the access rights for various users in regards to access of the object. It is possible that such an ACL could enforce different levels of access rights, such as read or write, for different users.

Typically, each object will have a defined owner and, in some cases, be associated to a primary group. The owner of a specific object controls its discretionary access attributes. The owner's attributes are set to the creating process's effective user ID.

The following list contains direct access control attributes for the different types of objects.

Owner

For System V Interprocess Communication (SVIPC) objects, the creator or owner can change the object's ownership. SVIPC objects have an associated creator that has all the rights of the owner (including access authorization). The creator cannot be changed, even with root authority.

SVIPC objects are initialized to the effective group ID of the creating process. For file system objects, the direct access control attributes are initialized to either the effective group ID of the creating process or the group ID of the parent directory (this is determined by the group inheritance flag of the parent directory).

Group

The owner of an object can change the group. The new group must be either, the effective group ID of the creating process, or the group ID of the parent directory. (As above, SVIPC objects have an associated creating group that cannot be changed, and share the access authorization of the object group.)

Mode

The **chmod** command (in numeric mode with octal notations) can set base permissions and attributes. The **chmod** subroutine that is called by the command, disables extended permissions. The extended permissions are disabled if you use the numeric mode of the **chmod** command on a file that has an ACL. The symbolic mode of the **chmod** command disables extended ACLs for NSF4 ACL type but does not disable extended permissions for AIXC type ACLs. For information about numeric and symbolic mode, see **chmod**.

Many objects in the operating system, such as sockets and file system objects, have ACLs associated for different subjects. Details of ACLs for these object types could vary from one to another.

Traditionally, AIX has supported mode bits for controlling access to the file system objects. It has also supported a unique form of ACL around mode bits. This ACL consisted of base mode bits and also allowed for the definition of multiple ACE entries; each ACE entry defining access rights for a user or group around the mode bits. This classic type of ACL behavior will continue to be supported, and is named AIXC ACL type.

Note that support of an ACL on file system objects depends on the underlying physical file system (PFS). The PFS must understand the ACL data and be able to store, retrieve, and enforce the accesses for various users. It is possible that some of the physical file systems do not support any ACLs at all (may just support the base mode bits) as compared to a physical file system that supported multiple types of ACLs. Few of the file systems under AIX have been enhanced to support multiple ACL types. JFS2 and GPFS will have the capability to support NFS version 4 protocol based ACL type too. This ACL has been named NFS4 ACL type on AIX. This ACL type adheres to most of the ACL definition in the NFS version 4 protocol specifications. It also supports more granular access controls as compared to the AIXC ACL type and provides for capabilities such as inheritance.

4.6 Role-based access control

System administration is an important aspect of daily operations, and security is an inherent part of most system administration functions. Also, in addition to securing the operating environment, it is necessary to closely monitor daily system activities.

Most environments require that different users manage different system administration duties. It is necessary to maintain separation of these duties so that no single system management user can accidentally or maliciously bypass system security. While traditional UNIX system administration cannot achieve these goals, role-based access control (RBAC) can.

4.6.1 AIX Role-based access control

AIX provided a limited role-based access control (RBAC) implementation prior to AIX 6.1.

Beginning with AIX 6.1, a new implementation of RBAC provides for a very fine granular mechanism to segment system administration tasks. Since these two RBAC implementations differ greatly in functionality, the following terms are used:

- **Legacy RBAC Mode**: The historic behavior of AIX roles that apply to versions before AIX 6.1
- **Enhanced RBAC Mode:** The new implementation introduced with AIX 6.1

Both modes of operation are supported. However, Enhanced RBAC Mode is the default on a newly installed AIX systems after AIX 6.1. The following sections provide a brief discussion of the two modes and their differences. We also include information on configuring the system to operate in the desired RBAC mode.

Legacy RBAC Mode

Prior to AIX 6.1, AIX provided limited RBAC functionality that allowed non-root users to perform certain system administration tasks.

In this RBAC implementation, when a given administrative command is invoked by a non-root user, the code in the command determines if the user is assigned a role with the required authorization. If a match is found, the command execution continues. If not, the command fails with an error. It is often required that the command being controlled by an authorization be **setuid** to the root user for an authorized invoker to have the necessary privilege to accomplish the operation.

This RBAC implementation also introduced a predefined but user-expandable set of authorizations that can be used to determine access to administrative commands. Additionally, a framework of administrative commands and interfaces to create roles, assign authorizations to roles, and assign roles to users is also provided.

While this implementation provides the ability to partially segment system administration responsibilities, it functions with the following constraints:

1. The framework requires changes to commands and applications to be RBAC-enabled.
2. Predefined authorizations are not granular and the mechanisms to create authorizations are not robust.
3. Membership in a certain group is often required as well as having a role with a given authorization in order to run a command.
4. Separation of duties is difficult to implement. If a user is assigned multiple roles, there is no way to act under a single role. The user always has all of the authorizations for all of their roles.

5. The least privilege principle is not adopted in the operating system. Commands must typically be SUID to the root user.

Legacy RBAC Mode is supported for compatibility, but Enhanced RBAC Mode is the default RBAC mode. Enhanced RBAC Mode is preferred on AIX.

Enhanced RBAC Mode

A more powerful implementation of RBAC is provided with starting with AIX 6.1. Applications that require administrative privileges for certain operations have new integration options with the enhanced AIX RBAC infrastructure.

These integration options center on the use of granular privileges and authorizations and the ability to configure any command on the system as a privileged command. Features of the enhanced RBAC mode will be installed and enabled by default on all installations of AIX beginning with AIX 6.1.

The enhanced RBAC mode provides a configurable set of authorizations, roles, privileged commands, devices and files through the following RBAC databases listed below. With enhanced RBAC, the databases can reside either in the local filesystem or can be managed remotely through LDAP.

- Authorization database
- Role database
- Privileged command database
- Privileged device database
- Privileged file database

Enhanced RBAC mode introduces a new naming convention for authorizations that allows a hierarchy of authorizations to be created. AIX provides a granular set of system-defined authorizations and an administrator is free to create additional user-defined authorizations as necessary.

The behavior of roles has been enhanced to provide separation of duty functionality. Enhanced RBAC introduces the concept of role sessions. A role session is a process with one or more associated roles. A user can create a role session for any roles that they have been assigned, thus activating a single role or several selected roles at a time. By default, a new system process does not have any associated roles. Roles have further been enhanced to support the requirement that the user must authenticate before activating the role to protect against an attacker taking over a user session since the attacker would then need to authenticate to activate the user’s roles.

The introduction of the privileged command database implements the least privilege principle. The granularity of system privileges has been increased, and explicit privileges can be granted to a command and the execution of the command can be governed by an authorization. This provides the functionality to enforce authorization checks for command execution without requiring a code change to the command itself. Use of the privileged command database eliminates the requirement of SUID and SGID applications since the capability of only assigning required privileges is possible.

The privileged device database allows access to devices to be governed by privileges, while the privileged file database allows unprivileged users access to restricted files based on authorizations. These databases increase the granularity of system administrative tasks that can be assigned to users who are otherwise unprivileged.

The information in the RBAC databases is gathered and verified and then sent to an area of the kernel designated as the Kernel Security Tables (KST). It is important to note that the state of the data in the KST determines the security policy for the system. Entries that are modified in the user-level RBAC databases are not used for security decisions until this information has been sent to the KST with the **setkst** command.

Note: A full discussion of Role-based access control on AIX can be found in IBM Documentation at https://www.ibm.com/docs/en/aix/7.3?topic=system-role-based-access-control.

4.6.2 AIX Toolbox for Open Source Software

It is important to note that many of the most prominent tools used in AIX and Linux are open source and should be considered when designing and maintaining virtual machines on IBM Power platforms. This list does not serve as an endorsement of these tools, just a note that they may be of value in your specific environment.

Note: For AIX users, these commands are available in the IBM AIX Toolbox for Open Source Software at https://www.ibm.com/support/pages/aix-toolbox-open-source-software-overview.

In addition to the GNU Public License (GPL), each of these packages includes its own licensing information, so remember to consult the individual tools for their licensing information.

Important: The freeware packages provided in the AIX Toolbox for Open Source Software are made available as a convenience to IBM customers. IBM does not own these tools, did not develop or exhaustively test them, nor do they provide support for these tools. IBM has compiled the these tools so that they will run with AIX.

The following tools are available:

- Role-Based Access Control
  - sudo: Grants specific users or groups the ability to execute commands as root or other users, enhancing security by limiting the need for users to have full root access.
- System Monitoring and Management:
  - Nagios: Monitor system metrics, services, and network protocols. Nagios Core is the open-source version.
  - Zabbix: An enterprise-level monitoring solution that provides real-time monitoring and alerting of various metrics.
  - Prometheus: A monitoring and alerting toolkit with a flexible query language and powerful visualization capabilities when paired with Grafana.
- Intrusion Detection:
  - OSSEC: A host-based intrusion detection system (HIDS) that performs log analysis, file integrity checking, and more.
  - Snort: An open-source network intrusion detection system (NIDS) that can also be configured as an intrusion prevention system (IPS).

- Network Analysis:
  - Wireshark: A widely used network protocol analyzer that helps in capturing and analyzing network traffic.
  - tcpdump: A command-line packet analyzer used for network diagnostics and monitoring.
- Log Management and Analysis:
  - ELK Stack (Elasticsearch, Logstash, Kibana): A powerful suite for log management and analysis, which helps in searching, analyzing, and visualizing log data.
  - Graylog: A log management tool that provides real-time analysis, alerting, and visualization capabilities.
- File Synchronization
  - rsync: A file synchronization and backup tool that is used for incremental backups and mirroring data.
- Package Management:
  - YUM/DNF: Package managers for RPM-based distributions that handle package installations, updates, and dependencies.

4.7 AIXpert Security compliance

AIX Security Expert is a comprehensive tool for managing system security settings, including TCP, NET, IPSEC, system configurations, and auditing. Part of the bos.aixpert fileset, it facilitates system security hardening through a user-friendly interface. The tool offers predefined security levels—High, Medium, Low, and AIX Standard Settings—that cover over 300 security configurations. It also allows advanced administrators to customize each security setting as needed.

With AIX Security Expert, you can easily apply a chosen security level without the need for extensive research and manual implementation of individual security elements. Additionally, the tool enables you to create a security configuration snapshot, which can be used to replicate the same settings across multiple systems, streamlining security management and ensuring consistency across an enterprise environment.

AIX Security Expert can be accessed either through SMIT or by using the **`aixpert`** command.

AIX Security Expert settings

The following coarse-grain security settings are available:

- High Level Security: Applies high level security settings definition
- Medium Level Security: Applies Medium level security settings definition
- Low Level Security: Applies the low level security settings definition
- Advanced Security: Applies custom user-specified security settings
- AIX Standard Settings: Uses original system default security settings
- Undo Security: Allows some AIX Security Expert configuration settings to be undone
- Check Security: Provides a detailed report of current security settings

4.7.1 AIX Security Expert security hardening

Security hardening protects all elements of a system by tightening security or implementing a higher level of security. It helps ensure that all security configuration decisions and settings are adequate and appropriate. Hundreds of security configuration settings might need to be changed to harden the security of an AIX system.

AIX Security Expert provides a menu to centralize effective and common security configuration settings. These settings are based on extensive research on properly securing UNIX systems. Default security settings are provided for broad security environment needs (High Level Security, Medium Level Security, and Low Level Security), and advanced administrators can set each security configuration setting independently.

Configuring a system at too high a security level might deny necessary services. For example, telnet and rlogin are disabled for High Level Security because the login password is sent over the network unencrypted. Conversely, if a system is configured at too low a security level, it can be vulnerable to security threats. Since each enterprise has its own unique set of security requirements, the predefined High Level Security, Medium Level Security, and Low Level Security configuration settings are best used as a starting point rather than an exact match for the security requirements of a particular enterprise.

The practical approach to using AIX Security Expert is to establish a test system (in a realistic test environment) similar to the production environment in which it will be deployed. Install the necessary business applications and run AIX Security Expert via the GUI. The tool will analyze this running system in its trusted state. Depending on the security options you choose, AIX Security Expert will enable port scan protection, turn on auditing, block network ports not used by business applications or other services, and apply many other security settings. After re-testing with these security configurations in place, the system is ready to be deployed in a production environment. Additionally, the AIX Security Expert XML file defining the security policy or configuration of this system can be used to implement the exact same configuration on similar systems in your enterprise.

Note: For more information on security hardening, see NIST Special Publication 800-70, NIST Security Configurations Checklist Program for IT Products. The fourth revision of the document is at: https://csrc.nist.gov/pubs/sp/800/70/r4/final.

A full discussion of AIX Security Expert on AIX v7.3 is available on IBM Documentation at https://www.ibm.com/docs/en/aix/7.3?topic=security-aix-expert.

4.8 File Permission Manager

The AIX File Permission Manager manages the permissions on commands and daemons that are owned by privileged users with setuid or setgid permissions.

The **`fpm`** command allows administrators to harden their system by setting permissions for important binaries and dropping the setuid and setgid bits on many commands in the operating system. This command is intended to remove the setuid permissions from commands and daemons that are owned by privileged users, but you can also customize it to address the specific needs of unique computer environments.

The setuid programs on the base AIX operating system have been grouped to allow for levels of hardening. This grouping allows administrators to choose the level of hardening according to their system environment. Also, you can use the **`fpm`** command to customize the list of programs that need to be disabled in your environment. Review the levels of disablement and choose the right level for your environment.

Changing execution permissions of commands and daemons with the **fpm** command affects non-privileged users, denying their access to these commands and daemons or functions of the commands and daemons. Also, other commands that call or depend on these commands and daemons can be affected. Any user-created scripts that depend on commands and daemons with permissions that were altered by the **fpm** command cannot operate as expected when run by non-privileged users. Give full consideration to the effect and potential impact of modifying default permissions of commands and daemons.

Perform appropriate testing before using this command to change the execution permissions of commands and daemons in any critical computer environment. If you encounter problems in an environment where execution permissions have been modified, restore the default permissions and recreate the problem in this default environment to ensure that the issue is not due to lack of appropriate execution permissions.

The **fpm** command provides the capability to restore the original AIX installation default permissions by using the **-l default** flag.

Also, the **fpm** command logs the permission state of the files before changing them. The **fpm** log files are created in the **/var/security/fpm/log/*date_time*** file. If necessary, you can use these log files to restore the system's file permissions that are recorded in a previously saved log file.

When the **fpm** command is used on files that have extended permissions, it disables the extended permissions, though any extended permission data that existed before the **fpm** invocation is retained in the extended ACL.

Customized configuration files can be created and enacted as part of the high, medium, low, and default settings. File lists can be specified in the **/usr/lib/security/fpm/custom/high/*** directory, the **/usr/lib/security/fpm/custom/medium/*** directory, and the **/usr/lib/security/fpm/custom/default/*** directory. To take advantage of this feature, create a file containing a list of files that you want to be automatically processed in addition to the **fpm** commands internal list. When the *fpm* command is run, it also processes the lists in the corresponding customized directories. To see an example of the format for a customized file, view the **/usr/lib/security/fpm/data/high_fpm_list** file. The default format can be viewed in the **/usr/lib/security/fpm/data/default_fpm_list.example** file. For the customization of the **-l low** flag, the fpm command reads the same files in the **/usr/lib/security/fpm/custom/medium** directory, but removes the setgid permissions, whereas the **-l medium** flag removes both the setuid and setgid permissions.

The **fpm** command cannot run on TCB-enabled hosts.

AIX File Permission Manager examples

1. To apply the fpm command's low-level security settings, enter:

   ```
   fpm –l low
   ```

   This command also processes any file list in the **/usr/lib/security/fpm/custom/med/** directory.

2. To check if the system commands are presently set to fpm low-level permissions, enter:

   ```
   fpm –c –l low
   ```

   This command reports any file with permissions out of conformance.

3. To restore the traditional out-of-the-box default permissions, enter:

   ```
   fpm –l default
   ```

   This command also processes any file list in the **/usr/lib/security/fpm/custom/default/** directory.

4. To list, or give a preview of what permission changes are to be done to make the system compliant with the fpm command's high-level security without changing any file permissions, enter:

   ```
   fpm -l high –p
   ```

   This command also previews any file list in the **/usr/lib/security/fpm/custom/high/** directory.

5. To apply the fpm command's high level security settings, enter:

   ```
   fpm –l high
   ```

   This command also processes any file list in the **/usr/lib/security/fpm/custom/high/** directory.

6. To list the current status of the system as changed through the fpm command, enter:

   ```
   fpm –s
   ```

7. If the fpm -l level command was run on 7 January 2024 at 8:00 AM then the permission state of the affected files was captured by the fpm command before it made any changes. To restore the file permissions to their state of 7 January 2007 at 8:00 AM, enter:

   ```
   fpm –l default –f /var/security/fpm/log/01072024_08:00:00
   ```

4.9 Trusted Execution

Trusted Execution (TE) refers to a collection of features that are used to verify the integrity of the system and implement advanced security policies, which together can be used to enhance the trust level of the complete system.

The usual way for a malicious user to negatively impact the system typically involves gaining unauthorized access and subsequently installing harmful programs like Trojans, rootkits, or modifying sensitive security files, thereby rendering the system vulnerable and prone to exploitation. Trusted Execution aims to prevent such activities or, in cases where incidents do occur, quickly identify them.

Using the functionality provided by Trusted Execution, the system administrator can define the exact set of executables that are permitted to run or specify the kernel extensions that are allowed to load. Additionally, it can be utilized to examine the security status of the system and identify files that have been updated, thereby raising the trustworthiness of the system and making it harder for an attacker to cause damage.

The set of features under TE can be grouped into the following:

- Managing Trusted Signature Database
- Auditing integrity of the Trusted Signature Database
- Configuring Security Policies
- Trusted Execution Path and Trusted Library Path

Important: A Trusted Computing Base (TCB) functionality already exists in the AIX operating system. *However, TCB features can only be enabled during the Base Operating System installation process by expressly switching on its corresponding menu option, or by performing a Preservation install on an already installed AIX system.*

Trusted Execution is a more powerful and enhanced mechanism that overlaps some of the TCB functionality and provides advance security policies to better control the integrity of the system. While the TCB is still available, TE introduces a new and more advanced concept of verifying and guarding the system integrity.

AIX Trusted Execution uses whitelisting to prevent or detect malware that is executed on your AIX system. It provides the following features:

- Provides cryptographic checking that will allow you to determine if a hacker has replaced an IBM published file with his own Trojan horse
- Provides the ability to scan for root kits
- Provides the ability to detect if various attributes of a file have been altered
- Provides the ability to correct certain file attribute errors
- Provides “white listing” functionality
- Provides a numerous configuration options
- Provides the ability to detect and/or prevent malicious scripts, executables, kernel extensions and libraries
- Provides functionality for protecting files from alteration by a hacker that has gained root access
- Provides functionality for protecting the Trusted Execution’s configuration from a hacker that has gained root access
- Provides functionality for utilizing digital signatures to verify IBM and non-IBM published files haven’t been altered by an attacker

Trusted Execution is available in AIX version 6 and all higher releases

Trusted Signature Database Management

Similar to that of Trusted Computing Base (TCB) there exists a database which is used to store critical security parameters of trusted files present on the system. This database, called Trusted Signature Database (TSD), resides in /etc/security/tsd/tsd.dat.

Auditing the integrity of Trusted Signature Database

The trustchk command can be used to audit the integrity state of the file definitions in the Trusted Signature Database (TSD) against the actual files.

Security policies configuration

The Trusted Execution (TE) feature provides you with a run-time file integrity verification mechanism. Using this mechanism, the system can be configured to check the integrity of the trusted files before every request to access those file, effectively allowing only the trusted files that pass the integrity check to be accessed on the system.

Quick reference for security checks

The `trustchk` command can be used to enable or disable the Trusted Library Path or Trusted Execution Library and to set the colon-separated path list for both, using Trusted Execution Path and Trusted Library Path command-line attributes of the trustchk command.

Table 4-3 compares the Trusted Computing Base (TCB) function with the Trusted Execution (TE) functionality.

Table 4-3 Comparison of TE and TCB functionality

Function	TE	TCB
Integrity Checking reference	System and runtime checking	System checking only
System Enablement	Enabled at any time	Installation time option
Security Database Files	/etc/security/tsd/tsd.dat	/etc/security/sysck.cfg
Management Commands	trustchk	tcbck

Trusted execution Policy Management

To enable or disable different security policies that are used with the Trusted Execution mechanism, use the **`trustchk`** command. You can specify policies shown in Table 4-4.

Table 4-4 Policies

Policy	Policy Description
CHKEXEC:	Checks integrity of commands before running the commands.
CHKSHLIBS:	Checks integrity of shared libraries before loading the libraries.
CHKSCRIPTS:	Checks integrity of shell scripts before running the scripts.
CHKKERNEXT:	Checks integrity of kernel extensions before loading the extensions.
LOCK_TSD:	Disables modification of TSD.
LOCK_TSD_	FILES: Disables modification of TSD files.
STOP_UNTRUSTD:	Does not load files unless in TSD.
STOP_ON_CHKFAIL:	If integrity check fails, do not load file
TEP:	Allows execution of commands from a defined list of directories.
TLP:	Allows library loads from a defined list of directories.

In order for TE to work, the CryptoLight for C library (CLiC) and kernel extension must be installed. To see if it is installed and loaded into the kernel, run the commands shown in Example 4-44.

Example 4-44 Check for CLIC install

```
# lslpp -l "clic*"
File set Level State Description
----------------------------------------------------------------------------
Path: /usr/lib/objrepos
clic.rte.kernext 4.3.0.0 COMMITTED CryptoLite for C Kernel
clic.rte.lib 4.3.0.0 COMMITTED CryptoLite for C Library
Path: /etc/objrepos
clic.rte.kernext 4.3.0.0 COMMITTED CryptoLite for C Kernel
# genkex|grep clic
4562000 37748 /usr/lib/drivers/crypto/clickext
```

If the file set is not installed, install it on your system and load it into the Kernel when installation completes successfully, by running:

```
# /usr/lib/methods/loadkclic
```

We have a database (Trusted Signature Database) that is used to store critical Security parameters of trusted files present on the system. This database is located at /etc/security/tsd/tsd.dat and comes with any AIX media. In TE’s context, Trusted files are files that are critical from the security perspective of a system and if compromised can jeopardize the security of the entire system. Typically the files that match this definition are:

- Kernel (operating system)
- All SUID root programs
- All SGID root programs
- Any program that is exclusively run by root or by a member of the system group
- Any program that must be run by the administrator while on the trusted
- Communication path (for example, the ls command)
- The configuration files that control system operation
- Any program that is run with the privilege or access rights to alter the kernel
- The system configuration files

Every trusted file must ideally have an associated stanza or a file definition stored in the TSD. A file can be marked as trusted by adding its definition in the TSD using the **trustchk** command. This command can be used to add/delete/list entries from the TSD. The TSD can be locked so even root cannot write to it any longer. Locking the TSD becomes immediately effective. Example 4-45 shows how the **ksh** command appears in the TSD db file.

Example 4-45 ksh entry in TSD database file

```
/usr/bin/ksh:
Owner = bin
Group = bin
Mode = TCB,555
Type = FILE
Hardlinks = /usr/bin/sh,/usr/bin/psh,/usr/bin/tsh,/usr/bin/rksh
Symlinks =
Size = 294254
Cert_tag = 00af4b62b878aa47f7
Signature =
8e8118ec793fd4899ccc38c0f4ab88571b0488024aff80f83d0bde2380f3ae44137a26607cd5d4c5e58e02ad1f872ca1c398f8
702ad38f3a0f0a584c2061bb09de3e5218405f1b07d80efe0be192d3333b8cd49a4ff980ce5e1f15f6b64d3b38f75d0cc6fb5e
f9e7d8b410547c40181847c5ae980979abf3279f25c6b512178a
hash_value = f3a2e9b92e2cfc10ffb2274680c97f29742ff2dd12dda04de85544fd8c039fd8
t_accessauths = aix.mls.system.access.dir
t_innateprivs = PV_DAC_R,PV_DAC_X,PV_MAC_R


/usr/lib/drivers/igcts:
Owner = root
Group = system
Mode = 555
Type = HLINK
Size = 7714
Cert_tag = 00af4b62b878aa47f7
Signature =
b47d75587bbd4005c3fe98015d9c0776fd8d40f976fb0f529796ffe1b2f9028500ffd2383ca31cd2f39712f70e36c522dc1ba5
2c44334781a389ea06cdabd82c72d705fd94
bffe59817b5a4d45651e2d5457cb83ebdb3b705a3b5c981c51eae79facfe271fbde0e396b7ea64d4dbd6ab753a3fa7a9578b7f
5e6458b83d8f08df
Hash_value = 6d13bbd588ecfdd06cbb2dc3a17eabad6b51a42bd1fd62e7ae5402a75116e8bd
```

To enable TSD protection, run the commands shown in Example 4-46.

Example 4-46 Enabling TSD protection

```
# trustchk -p tsd_lock=on
# trustchk -p te=on
```

The TSD is immediately protected against any kind of modification by either the **trustchk** command or by manually editing the file. This is shown in Example 4-47.

Example 4-47 TSD is protected

```
# trustchk -d /usr/bin/ps
Error writing to database file
# echo >> /etc/security/tsd/tsd.dat
Operation not permitted.
```

To enable the TSD for write access again, you either need to turn off TE completely or set tsd_lock to off using the **trustchk** command.

When the system is blocking any untrusted shell scripts by using the CHKSCRIPT policy as shown in Example 4-48 make sure all scripts needed by your services are included in the TSD.

Example 4-48 Commands to set chkscript on and stop untrusted executions

```
# trustchk -p stop_untrustd=on
# trustchk -p chkscript=on
```

For example, if you are using OpenSSH make sure the sshd and ksshd start and stop scripts in **/etc/rc.d/rc2.d** are in the TSD. Otherwise, **sshd** does not start when the system is restarted and it will not be shut down on a system shutdown.

When you try to start a script with chkscript=on and that script is not included in the TSD, its execution is denied, regardless of its permissions, even when root is starting it. This is shown in Example 4-49.

Example 4-49 Show permission denied for access

```
# ./foo
ksh: ./foo: 0403-006 permission denied.

# ls -l foo
-rwx------- root system 17 May 10 11:51 foo
```

The Trusted Execution Path defines a list of directories that contain the trusted commands. When Trusted Execution Path verification is enabled, the system loader allows commands in the specified paths to run.

The Trusted Library Path has the same function as Trusted Execution Path with the only Difference that it is used to define the directories that contain trusted libraries of the system. When TLP is enabled, the system loader allows the libraries from this path to be linked to the commands.

The trustchk command can be used to enable or disable the Trusted Execution Path or Trusted Execution Library as well as to set the Colon-separated path list for both, using Trusted Execution Path and Trusted Library Path command-line attributes of trustchk:

Important: Be careful when you are changing either Trusted Execution Path or Trusted Library Path. We do not recommend removing paths from their default settings, which are currently set to:

```
TEP=/usr/bin:/usr/sbin:/etc:/bin:/sbin:/sbin/helpers/jfs2:/usr/lib/instl:/usr/ccs/bin
TLP=/usr/lib:/usr/ccs/lib:/lib:/var/lib
```

Doing so most probably results in a system that will not restart and function properly since it cannot access necessary files and data any longer.

Here are some common command usages.

- To perform a system check comparison with the TSD and report errors run the following command:

  `# trustchk -n ALL`

- To delete the entry for /usr/bin/ls in the TSD run this command:

  `# trustchk -d /usr/bin/ls`

- Enable policy for checking commands listed in TSD on every load run this command:

  `# trustchk -p CHKEXEC=ON`

- To turn on runtime TSD checking, run this command:

  `# trustchk -p TE=ON`

- To check the current runtime policy in effect, run this command:

  `# trustchk -p`

Here are some additional examples:

- Adding STOP_ON_CHKFAIL option so it stops commands failing the test is shown in Example 4-50.

  The trustchk flags used are:

  - TE=[ON|OFF]: Turns runtime checks
  - CHKEXEC=[ON|OFF]: Turns commands checking
  - STOP_ON_CHKFAIL= [ON|OFF]: Stop commands failing the test
  - STOP_UNTRUSTD= [ON|OFF]: Stop commands not listed in /etc/security/tsd/tsd.dat

Example 4-50 Demonstrating STOP_ON_CHKFAIL

```
# openssl dgst -sha256 /usr/bin/ls | awk '{print $2}'
8f3505509771df3915b6f8c7e45fc6a56ec68d4c082bfb640f89c2251bf9550c

# trustchk -q /usr/bin/ls | grep hash_value | awk '{print $3}'
8f3505509771df3915b6f8c7e45fc6a56ec68d4c082bfb640f89c2251bf9550c

# cp /usr/bin/ls /usr/bin/.goodls
- Hash value of "/usr/bin/ls" command changed
# trustchk -p TE=ON CHKEXEC=ON STOP_ON_CHKFAIL=ON

# ls
ksh: ls: 0403-006 permission denied.

# cp /usr/bin/ls /usr/bin/.badls
# cp /usr/bin/.goodls /usr/bin/ls
# chown bin:bin /usr/bin/ls

# ls
file1 file2 dir1
```

2) Using STOP_UNTRUSTD=ON option so it stops executables not listed in /etc/security/tsd/tsd.dat. This is shown in Example 4-51.

Example 4-51 Demonstrating STOP_UNTRUSTD=ON

```
# trustchk -p TE=ON CHKEXEC=ON STOP_UNTRUSTD=ON
# ls
file1 file2 dir1
# /usr/bin/.goodls
ksh: /usr/bin/.goodls: 0403-006 permission denied.
# ls -l /usr/bin/.goodls
-r-xr-xr-x 1 bin bin 26732 May 28 17:39 /usr/bin/.ls
```

4.10 IPSec Network Filtering

This sections talks about IBM AIX IPSEC network filters and tunneling.

With the constant threat of security breaches, companies are under pressure to lock down every aspect of their applications, infrastructure, and data.

One method of securing IBM AIX network transactions is to establish networks based on the IPSEC protocol. Internal Protocol Security (IPSEC) is an IBM AIX network based protocol that defines how to secure a computer network at the IP layer.When determining how to secure your IPSEC connections, you may need to consider these items:

- Connectivity architecture—whether it is an internal or external connection.
- Encryption mechanisms or a use of authentication services.

IBM AIX IPSEC native tool uses the **mkfilt** and **genfilt** binaries to activate and add the filter rules. It can also be used to control the filter logging functions.This works on IP version 4 and IP version 6. With the IPSEC feature enabled, you can also create IP filtering rules to block the IP address from accessing hosts or exact ports.

One of the interesting features of IPSEC is IP Security dynamic tunnels. These tunnels use the Internet Key Exchange (IKE) Protocol to protect IP traffic by authenticating and encrypting IP data. The **ike** command performs several functions such as activate, remove, or list IKE and IP Security tunnels.

For more information, please see these links:

https://www.ibm.com/support/pages/ibm-aix-using-ipsec-rules-filter-network-traffic
https://www.ibm.com/docs/en/aix/7.3?topic=m-mkfilt-command
https://www.ibm.com/docs/en/aix/7.3?topic=g-genfilt-command
https://www.ibm.com/docs/en/aix/7.3?topic=i-ike-command
https://www.ibm.com/docs/en/aix/7.3?topic=tunnels-command-line-interface-ike-tunnel-configuration

4.11 AIX Event Monitoring

AIX has built in functionality for monitoring and auditing events that occur within the environment. These are discussed in this section.

4.11.1 Auditing

Auditing is the process of examining systems, accounts, or activities to verify accuracy, compliance, and efficiency. An auditing subsystem records system events to monitor, analyze transactions, and ensure transparency and traceability.

By default, auditing is disabled in AIX. When activated, the auditing subsystem begins collecting information based on your configuration settings. The frequency of auditing depends on your environment and usage patterns. While recommended for enhanced security and troubleshooting, the decision to enable auditing and its frequency is ultimately yours.

Any security-relevant occurrence on the system is considered an auditable event. Auditing involves detecting, recording, and analyzing these events to maintain system security and integrity. The set of auditable events determines which occurrences can be audited and the level of detail provided. By examining audit trails, organizations can identify patterns, trends, and anomalies that may indicate potential security threats or issues.

Understanding the AIX Auditing Subsystem

The auditing subsystem provides a mechanism to record security-related information and alert system administrators of potential or actual security policy violations. Collected information includes the auditable event name, status (success or failure), and event-specific security details.

The auditing subsystem comprises detection, collection, and processing functions:

- Event Detection: Distributed throughout the Trusted Computing Base (TCB), both in the kernel and trusted programs. Detects security-relevant occurrences and reports them to the system audit logger.
- Event Information Collection: The kernel audit logger records selected auditable events, constructing a complete audit record consisting of a header and trail. The audit trail can be written in BIN or STREAM mode.
- Audit Trail Information Processing: The operating system offers various options for processing the audit trail, including compression, filtering, and formatting.

The system administrator can configure each of these functions.

Auditing event detection

Event detection is integrated throughout the Trusted Computing Base (TCB), encompassing both the kernel (supervisor state code) and trusted programs (user state code). An auditable event is any occurrence that is relevant to system security. This includes changes to the system's security state, attempted or actual violations of access control or accountability policies, or both. Programs and kernel modules responsible for detecting these events must report them to the system audit logger, which operates as part of the kernel. The logger can be accessed via a subroutine (for trusted program auditing) or through a kernel procedure call (for supervisor state auditing). The reported information includes the event name, its success or failure status, and any additional details pertinent to security auditing.

Event Detection Configuration:

Configuring event detection involves enabling or disabling the audit subsystem and specifying which events to audit for particular users. To manage event detection, use the audit command to activate or deactivate the subsystem. The /etc/security/audit/config file contains the configuration details for the events and users processed by the audit subsystem.

Event information collection

Information collection encompasses logging the selected auditable events. This function is performed by the kernel audit logger, which provides both a system call and an intra-kernel procedure call interface that records auditable events.

The audit logger is responsible for constructing the complete audit record, consisting of the audit header, that contains information common to all events (such as the name of the event, the user responsible, the time and return status of the event), and the audit trail, which contains event-specific information. The audit logger appends each successive record to the kernel audit trail, which can be written in either (or both) of two modes:

- BIN mode

  The trail is written into alternating files, providing for safety and long-term storage.

- STREAM mode

  The trail is written to a circular buffer that is read synchronously through an audit pseudo-device. STREAM mode offers immediate response.

Information collection can be configured at both the front end (event recording) and at the back end (trail processing). Event recording is selectable on a per-user basis. Each user has a defined set of audit events that are logged in the audit trail when they occur. At the back end, the modes are individually configurable, so that the administrator can employ the back-end processing best suited for a particular environment. In addition, BIN mode auditing can be configured to generate an alert in case the file system space available for the trail is getting too low.

Audit trail information processing

The operating system offers various options for processing the kernel audit trail. In BIN mode, the audit trail can be compressed, filtered, or formatted for output, or a combination of these methods can be used before archiving the audit trail.

- Compression: This is performed using Huffman encoding.
- Filtering: Achieved through SQL-like audit record selection (via the auditselect command), which allows for selective viewing and retention of records.
- Formatting: This enables examination of the audit trail, generation of periodic security reports, and printing of a paper audit trail.

These processing options help manage and analyze audit data effectively.

The STREAM mode audit trail can be monitored in real time, to provide immediate threat-monitoring capability. Configuration of these options is handled by separate programs that can be invoked as daemon processes to filter either BIN or STREAM mode trails, although some of the filter programs are more naturally suited to one mode or the other.

To ensure that the AIX audit subsystem can retrieve information from the AIX security audit, you must set the following files to the AIX server to be monitored.

- streamcmds
- config
- events
- objects

For more information on how to configure the AIX audit subsystem for collecting, recording and auditing the events, please check below links:

https://www.ibm.com/support/pages/aix-audit-audit-subsystem-aix
https://www.ibm.com/docs/en/aix/7.3?topic=files-config-file

https://www.ibm.com/docs/en/aix/7.3?topic=files-events-file
https://www.ibm.com/docs/en/powersc-standard/2.0?topic=logging-configuring-aix-audit-subsystem
https://www.ibm.com/docs/en/aix/7.3?topic=system-auditing-overview
https://www.ibm.com/docs/en/aix/7.2?topic=management-enhanced-auditing

4.11.2 Accounting

The accounting subsystem provides features for monitoring system resource utilization and billing users for the use of resources. Accounting data can be collected on a variety of system resources: processors, memory, disks, and such.

Another kind of data collected by the accounting system is connect-time usage accounting, which lets us know how many users are connected to a system and for how long. The connect time data enables us to detect unused accounts, which have to be invalidated (for security reasons) or even erased to save resources. Also, the connect-time usage may enable the discovery of suspect activities (such as too many unsuccessful logon attempts) that signal that security measures should be adopted.

The data collected by the accounting subsystem is used to automatically generate reports, such as daily and weekly reports. The reports can be generated at any time, using accounting-specific commands. The accounting subsystem provides tools that enable us to observe how the system reacts at a particular moment in time (for instance, when executing a specific command or task).

Accounting data provides valuable information to:

- Develop effective charge-back policies.
- Assess the adequacy of the current resources.
- Effectively balance and control resource allocation.
- Forecast future needs.

For more details, on how to set up accounting subsystem and accounting internals, please click on the below links:

https://www.ibm.com/docs/en/aix/7.3?topic=accounting-administering-system

4.11.3 AIX Event Infrastructure for AIX and AIX clusters

The AIX Event Infrastructure is an event monitoring framework for monitoring predefined and user-defined events.

An event in the AIX Event Infrastructure refers to any detectable change in a system's state or values by the kernel or its extensions at the moment the modification takes place. These events are stored as files within a specialized file system known as the pseudo file system. The AIX Event Infrastructure offers several benefits, including:

- There is no need for constant polling. Users monitoring the events are notified when those events occur.
- Detailed information about an event (such as stack trace and user and process information) is provided to the user monitoring the event.
- Existing file system interfaces are used so that there is no need for a new application programming interface (API).
- Control is handed to the AIX Event Infrastructure at the exact time the event occurs.

AHAFS architecture

The AIX Event Infrastructure is made up of the following four components:

1. The kernel extension implementing the pseudo file system.
2. The event consumers that consume the events.
3. The event producers that produce events.
4. The kernel component that serve as an interface between the kernel extension and the event producers.

Figure 4-3 illustrates the architecture.

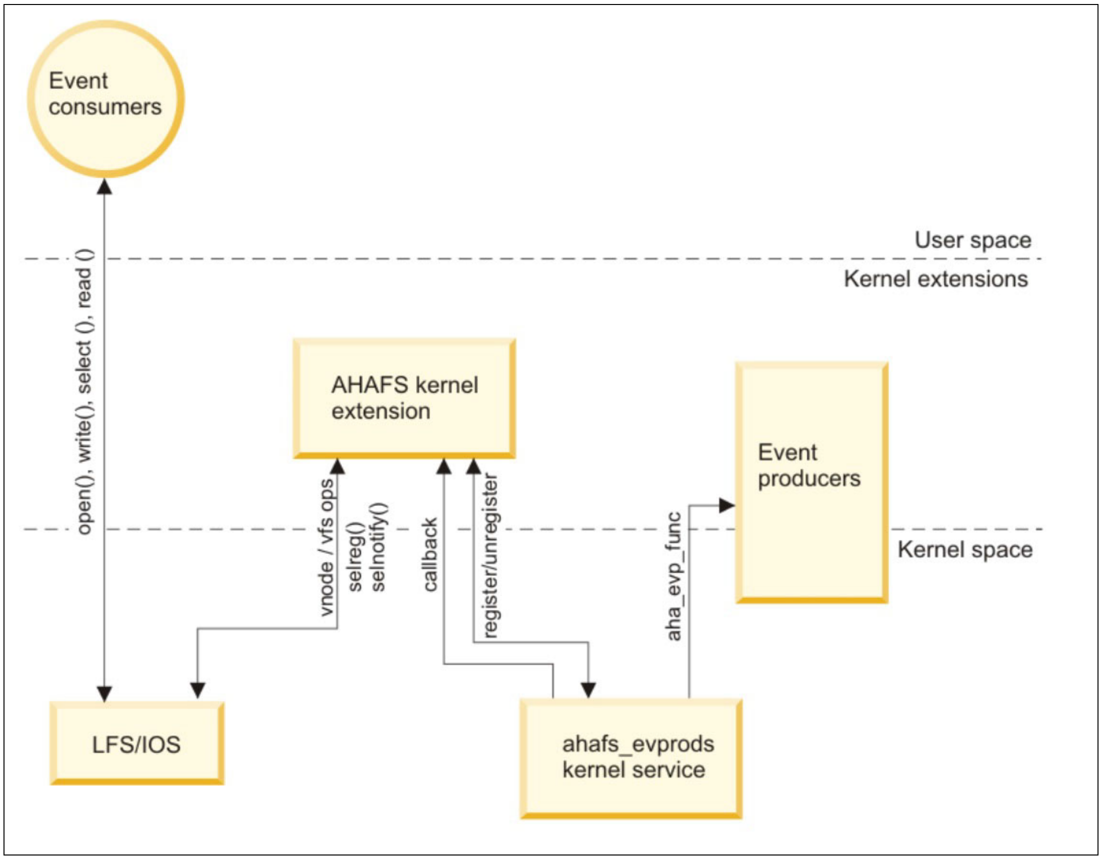


Figure 4-3 AHAFS architecture[3]

For more information see the AIX Event Infrastructure documentation.

4.12 In-core Cryptography Support

The use of encryption to protect data requires a significant amount of processing to encrypt and decrypt data. To reduce the performance impact of these activities, IBM Power processors provide on-chip capabilities to off-load these processor intensive cryptographic operations. These benefits are provided by on-chip accelerators and by using new crypto processor instructions (in-core) to accelerate these functions. These hardware accelerators can greatly reduce the CPU utilization and improve the performance of AIX applications which use crypto features.

[3] https://www.ibm.com/docs/en/aix/7.3?topic=ahafs-aix-event-infrastructure-components

IBM POWER7+ was the first IBM Power processor to include Nest Accelerator (NX) for symmetric (shared key) cryptography. The accelerators are shared among the logical partitions (LPARs) under the control of the PowerVM Hypervisor, and accessed via Hypervisor call. The internal NX crypto API calls require extra pages of memory to perform the relevant Hypervisor calls. The overhead of NX calls makes them suitable for large size of data only. A tuning parameter (min_sz) for data size is implemented, to gate what minimal data size for NX accelerator operations.

The POWER8 processor provided a new set of VMX/VSX in-core symmetric cryptographic instructions that are aimed at improving performance of various crypto operations. In most circumstances, the in-core crypto instructions provide better performance with lower latency and no extra page requirements. To be able to use the in-core crypto instructions in kernel, there is a small amount of overhead to save and restore the vector register content.

Additional improvements were made in the IBM Power9 and IBM Power10 chips to increase the encryption capabilities and greatly improve system performance.

Advance Crypto Facility (ACF) in AIX

Advance Crypto Facility (ACF) is the AIX cryptographic framework that provides crypto services (APIs) for kernel and user space applications. It implemented all the supported crypto algorithms in software that can be replaced by other crypto providers, like crypto cards and hardware accelerations when the respective hardware acceleration is enabled. The leverage of hardware acceleration is done in a manner transparent to the callers.

The ACF kernel services are implemented in pkcs11 device driver (kernel extension), providing services for other kernel subsystems like EFS, IPSec and LV-Encryption. User space applications can also use ACF kernel services by calling the AIX PKCS #11 subsystem library (/usr/lib/pkcs11/ibm_pkcs11.so).

- The purpose of this feature is to improve the performance in pkcs11(Public Key Cryptography Standards) kernel extension.
- Use of the AES instruction set can greatly reduce the CPU utilization and improve the performance of AIX applications which use AES crypto features—EFS, IPSEC, Trusted Execution.
- This feature enables the in-core vector AES crypto instructions under CLiC interfaces of pkcs11 kernel extension.
- Customers can enable/disable the In-Core support in ACF kernel extension though a CLI interface.
- ODM Support is provided for enabling or disabling the feature after reboots.
- Display the Status of the in-core crypto enablement.
- Supported IBM POWER8 and above.
- Prerequisites:
  - OS Level: 7.2 TL5 and later
  - VIOS:v3.1
  - Hardware: POWER8 or Higher
  - Firmware: Any
- Enablement, how to turn it on:
  - Two flags introduced:
    - in_core_capable
    - in_core_enabled (acfo -t in_core_enabled=1)

- Recovery, how to turn it off or disable it if something goes wrong:
  - acfo -t in_core_enabled=0
- Default settings:
  - For POWER8 and above, default values are:
    - in_core_Capable=1 and in_core_enabled=0

For more information, please see this link.

4.13 LDAP (Lightweight Directory Access Protocol)

Lightweight Directory Access Protocol (LDAP) is a distributed hierarchical directory-service access protocol that is used to access repositories of users and other network-related entities. Lightweight Directory Access Protocol (LDAP) is an open standard for managing directory data.

LDAP defines a message protocol used by directory clients and directory servers. LDAP originated from X.500 Directory Access Protocol, which is considered as heavyweight. X.500 needs the entire OSI protocol stack, where LDAP is built on the TCP/IP stack. LDAP is also considered lightweight because it omits many X.500 operations that are rarely used.

An application-specific directory stores the information which do not have general search capabilities. Keeping multiple copies of information up-to-date and synchronized is difficult. What is needed is a common, application-independent directory. A dream of single common Directory can be achieved using LDAP. Clients can interact independent of the platform. Also clients can be setup without any dependency.

LDAP works with most vendor directory services, such as Active Directory (AD). With LDAP, sharing information about users, services, systems, networks, and applications from a directory service to other applications and services becomes easier to implement. When using LDAP, the client access is independent of the platform. Since LDAP is a standard protocol, clients can be setup without any dependency on the specific LDAP server being utilized.

For example if you have a Microsoft Active Directory (LDAP Server), you can configure an LDAP client with IBM TDS file-sets and access the data from the server. Example 4-52 shows a sample of an LDAP entry for multiple applications.

Example 4-52 Sample LDAP entries

```
Application 1: FirstName, LastName
Application 2: name
Application 3: firstname, middlename, lastname
```

When setup to use LDAP, multiple applications such as IBM Verse, intranet page, BestQuest, RQM, ClearQuestcan be connected to a user entry in the same directory. If a user changes their password once, it is reflected in all the applications.

Figure 4-4 illustrates this concept.

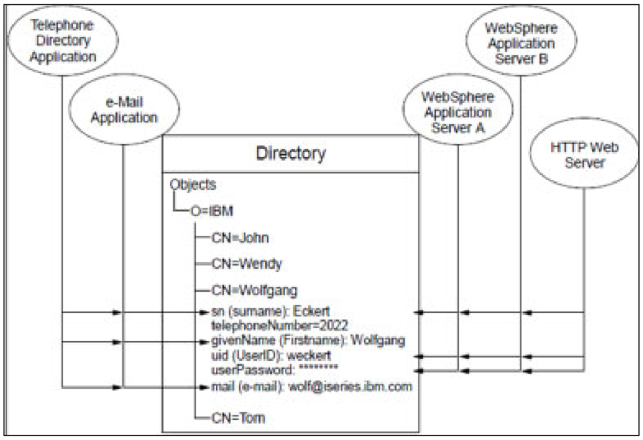


Figure 4-4 Several applications using the attributes of a single entry

4.13.1 How LDAP works

An application that wants to read or write information in a directory does not access the directory directly. It calls a function or API that causes a message to be sent to another process. This second process accesses the information in the directory on behalf of the requesting application via TCP/IP.

- The default TCP/IP ports are 636 for secure communications and 389 for unencrypted communications.
- You use the **mksecdlap** command in AIX to set up LDAP.
- DB2 is the database used.

4.13.2 Where LDAP is used

Most of the applications need the similar functionality of authenticating the user and storing some user/device information. In normal method, we need to create database schema, create tables etc and maintain the database. Also, later we need to create application program which access the database. Using LDAP, all the above work will be done by LDAP server. The user can just use the services using APIs available.

Here is a list of possible LDAP Configurations:

- Server configurations :
  - Master/Replica
  - Peer-peer
- Client configurations:
  - LDAP Anonymous bind – Automated as per ISST Auto bucket
  - LDAP Automount – Automated as per ISST auto bucket
  - LDAP Netgroups – Automated as per ISST auto bucket
  - LDAP with Auth only mode – Automated as per ISST auto bucket
  - IPSEC with LDAP – Automation Code is ready, need to verify
  - LDAP with SSL communication

4.13.3 LDAP Exploitation in AIX

IBM Security Verify Directory is a highly scalable and robust LDAP directory server. The AIX security subsystem's use of IBM Security Verify Directory allows for centralized security authentication, as well as access to user and group information. This functionality can be used in a host clustering environment to keep authentication, user, and group information common. The **mksecldap** command sets up an AIX cluster that consists of one or more servers, and one or more clients that use the IBM Security Verify Directory (LDAP) for security authentication, user and group management.

The LDAP exploitation in AIX provides:

- Seamless Integration
  - Automated client and server configuration.
  - Centralized AIX user/group/network management.
  - User/group/network export tools.
- Features
  - RFC 2307 based implementation
    - Heterogeneous LDAP environment compatibility.
  - Client side daemon manages requests and connections.
  - Fault tolerant and priority based server failover and reconnection.

For more information on how to set up an LDAP server and to configure clients in AIX, see the following:

- *Integrating AIX into Heterogeneous LDAP Environments*, SG24-7165
- http://theaix.blogspot.com/2009/10/ldap-in-aix.html
- https://community.spiceworks.com/t/how-to-install-ldap-on-aix-7-1-and-configure-as-ldap-server/836720/2
- https://www.ibm.com/docs/bg/aix/7.2?topic=module-setting-up-ldap-client

4.13.4 LDAP-Based User and Group Management on AIX

The LDAP exploitation of the security subsystem is implemented as the LDAP authentication load module. It is conceptually similar to the other load modules such as NIS, DCE, and KRB5. Load modules are defined in the /usr/lib/security/methods.cfg file. The LDAP load-module provides user authentication and centralized user and group management functionality through the LDAP protocol. A user defined on a LDAP server can be configured to log in to an LDAP client even if that user is not defined locally.

The AIX LDAP load module is fully integrated within the AIX operating system. After the LDAP authentication load module is enabled to serve user and group information, high-level APIs, commands, and system-management tools work in their usual manner. An -R flag is introduced for most high-level commands to work through different load modules.

AIX supports LDAP-based user and group management, integrating with IBM Security Verify Directory servers, non-IBM RFC 2307-compliant servers, and Microsoft Active Directory. The recommended option for use in defining AIX users and groups is IBM Security Verify Directory. Refer to Setting up an IBM Security Verify Directory Server for more information on setting up the server.

AIX supports non-IBM directory servers as well. A directory server that is RFC 2307 compliant is supported and AIX treats these servers similarly to IBM Security Verify Directory Servers. Directory servers that are not RFC 2307 compliant can be used but they require additional manual configuration to map the data schema. There may be some limitations due to the subset of user and group attributes in RFC 2307 compared to the AIX implementation. LDAP Version 3 protocol support is required.

AIX also supports Microsoft Active Directory (AD) as an LDAP server for user and group management. This requires the UNIX supporting schema be installed (included in Microsoft Service For UNIX). AIX supports AD running on Windows 2000, 2003, and 2003 R2 with specific SFU schema versions.

Some AIX commands may not function with LDAP users if the server is AD due to differences in user and group management between UNIX and Windows systems. Most user and group management commands (e.g., lsuser, chuser, rmuser, lsgroup, chgroup, rmgroup, id, groups, passwd, chpasswd) should work, depending on access rights.

4.13.5 Setting up the IBM LDAP server

IBM Security Verify Directory is a highly scalable and robust LDAP directory server. The AIX security subsystem's use of IBM Security Verify Directory allows for centralized security authentication, as well as access to user and group information. This functionality can be used in a host clustering environment to keep authentication, user, and group information common. The mksecldap command sets up an AIX cluster that consists of one or more servers, and one or more clients that use the IBM Security Verify Directory (LDAP) for security authentication, user and group management.

The procedure to set up the AIX Security Subsystem to use IBM Security Verify Directory (LDAP) involves two steps. The first step sets up a IBM Security Verify Directory Server that serves as a centralized repository for user and group information when authenticating. The second step in the procedure sets up the host systems (clients) to use the IBM Security Verify Director server for authentication and to retrieve user/group information.

Instructions for installing the IBM AIX LDAP can be found here:

https://www.ibm.com/support/pages/ldap-aix-step-step-instructions-installing-ldap-client-filesets-aix

5

IBM i Security

The IBM Power family covers a wide range of users. Security on the IBM i platform is flexible enough to meet the requirements of this wide range of users and situations. This chapter provides information so that you can understand the features and options available and can adapt them to your own security requirements.

This chapter discusses the following:

5.1 Introduction to IBM i security

To effectively create a security policy and plan security measures for your system, you need to understand the following security concepts, some of which are general concepts and some of which are specific to the hardware type.

A small system might have three to five users and a large system might have several thousand users. Some installations have all their workstations in a single, relatively secure area. Others have widely distributed users, including users who connect by dialing in and indirect users connected through personal computers or system networks. Security on IBM i is flexible enough to meet the requirements of this wide range of users and situations.

System security has some important objectives. Each security control or mechanism should satisfy one or more of the following security goals:

Confidentiality

Confidentiality concerns include:

- Protecting against disclosing information to unauthorized people
- Restricting access to confidential information
- Protecting against curious system users and outsiders

Integrity

Integrity is an important aspect when applied to data within your enterprise. Integrity goals include:

- Protecting against unauthorized changes to data
- Restricting manipulation of data to authorized programs
- Providing assurance that data is trustworthy

Availability

Systems are often critical to keep an enterprise running. Availability includes:

- Preventing accidental changes or destruction of data
- Protecting against attempts by outsiders to abuse or destroy system resources

Authentication

Ensuring that your data is only accessible by entities that are authorized is one of the basic tenets of data security. Proper authentication methodologies are important to:

- Determines whether users are who they claim to be. The most common technique to authenticate is by user profile name and password.
- Provide additional methods of authentication such as using Kerberos as an authentication protocol in a single sign-on (SSO) environment.

Authorization

Once a user is authenticated, it is also important to ensure that they only access the data and tasks that are relevant to their job. Proper authorization is important to:

- Permit a user to access resources and perform actions on them.
- Define access permissions (public or private rights) to objects to ensure that they are not accessed except by those that have authorization.

Auditing or logging

Auditing and logging are important in discovering and stopping access threats before the system becomes compromised. It is important to:

- As soon as your security plan is implemented, you must monitor the system for any out-of-policy security activity and resolve any discrepancies created by the activity.
- Depending on your organization and security policy, you may also need to issue a security warning to the person who performed the out-of-policy security activity so that they know not to perform this action in the future.

System security is often associated with external threats, such as hackers or business rivals. However, protection against system accidents by authorized system users is often the greatest benefit of a well-designed security system. In a system without good security features, pressing the wrong key might result in deleting important information. System security can prevent this type of accident.

The best security system functions cannot produce good results without good planning. Security that is set up in small pieces, without planning, can be confusing and is difficult to maintain and to audit. Planning does not imply designing the security for every file, program, and device in advance. It does imply establishing an overall approach to security on the system and communicating that approach to application designers, programmers, and system users.

As you plan security on your system and decide how much security you need, consider these questions:

- Is there a company policy or standard that requires a certain level of security?
- Do the company auditors require some level of security?
- How important is your system, and the data on it, to your business?
- How important is the error protection provided by the security features?
- What are your company security requirements for the future?

To facilitate installation, many of the security capabilities on your system are not activated when your system is shipped. Recommendations are provided in this chapter to bring your system to a reasonable level of security. Always consider the security requirements of your own installation as you evaluate any recommendations.

5.2 Staying current on IBM i

As we discussed in earlier sections, one of the most important task in maintaining a system’s security is to ensure that all available operating system updates are installed promptly. Fixes are essential for system maintenance, ensuring optimal availability, reduced downtime, and added functionality.

IBM periodically releases fixes to address issues discovered in IBM i programs. These fixes are bundled into cumulative PTF packages, which contain recommended fixes for specific time periods. Consider installing cumulative PTF packages twice a year in dynamic environments and less frequently in stable ones. Additionally, apply them when making major hardware or software changes.

By prioritizing fixes, fix groups, cumulative packages, and high-impact pervasive (HIPER) fixes, you can prevent security issues caused by failing to implement operating system fixes to correct known issues.

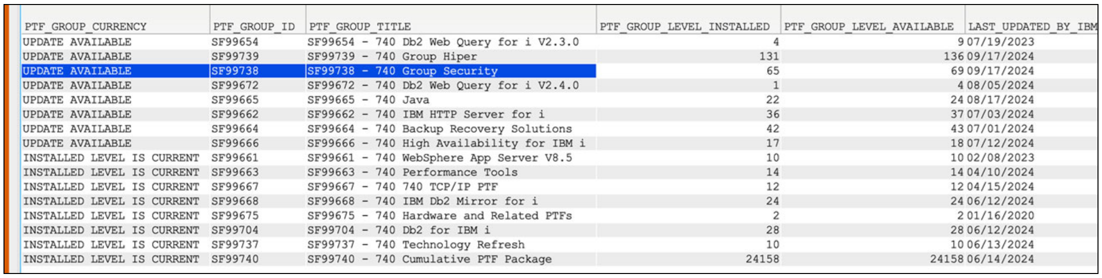
The IBM Navigator for i web based tool contains technology for doing system management tasks across one or more systems at the same time. IBM Navigator for i provides wizards that simplify fix management. The wizards allow you to easily send, install or uninstall fixes.You can also use the compare and update wizard to compare a model system to multiple target systems to find missing or extra fixes. Additionally tools like IBM Administration Runtime Expert for i and Ansible can be used to not only compare, but automate this process.

Another option for managing fixes is to use a SQL query to identify any issues. This is documented in this IBM document. The query is shown in Example 5-1.

Example 5-1 Group PTF currency query

```
--
-- Derive the IBM i operating system level and then
-- determine the level of currency of PTF Groups
--
With iLevel(iVersion, iRelease) AS
(
select OS_VERSION, OS_RELEASE from sysibmadm.env_sys_info
)
  SELECT P.*
     FROM iLevel, systools.group_ptf_currency P
     WHERE ptf_group_release =
           'R' CONCAT iVersion CONCAT iRelease concat '0'
     ORDER BY ptf_group_level_available -
        ptf_group_level_installed DESC;
```

An example output is shown in Figure 5-1.



PTF_GROUP_CURRENCY	PTF_GROUP_ID	PTF_GROUP_TITLE	PTF_GROUP_LEVEL_INSTALLED	PTF_GROUP_LEVEL_AVAILABLE	LAST_UPDATED_BY_IBM
UPDATE AVAILABLE	SF99654	SF99654 - 740 Db2 Web Query for i V2.3.0	4	9	07/19/2023
UPDATE AVAILABLE	SF99739	SF99739 - 740 Group Hiper	131	136	09/17/2024
UPDATE AVAILABLE	SF99738	SF99738 - 740 Group Security	65	69	09/17/2024
UPDATE AVAILABLE	SF99672	SF99672 - 740 Db2 Web Query for i V2.4.0	1	4	08/05/2024
UPDATE AVAILABLE	SF99665	SF99665 - 740 Java	22	24	08/17/2024
UPDATE AVAILABLE	SF99662	SF99662 - 740 IBM HTTP Server for i	36	37	07/03/2024
UPDATE AVAILABLE	SF99664	SF99664 - 740 Backup Recovery Solutions	42	43	07/01/2024
UPDATE AVAILABLE	SF99666	SF99666 - 740 High Availability for IBM i	17	18	07/12/2024
INSTALLED LEVEL IS CURRENT	SF99661	SF99661 - 740 WebSphere App Server V8.5	10	10	02/08/2023
INSTALLED LEVEL IS CURRENT	SF99663	SF99663 - 740 Performance Tools	14	14	04/10/2024
INSTALLED LEVEL IS CURRENT	SF99667	SF99667 - 740 740 TCP/IP PTF	12	12	04/15/2024
INSTALLED LEVEL IS CURRENT	SF99668	SF99668 - 740 IBM Db2 Mirror for i	24	24	06/12/2024
INSTALLED LEVEL IS CURRENT	SF99675	SF99675 - 740 Hardware and Related PTFs	2	2	01/16/2020
INSTALLED LEVEL IS CURRENT	SF99704	SF99704 - 740 Db2 for IBM i	28	28	06/12/2024
INSTALLED LEVEL IS CURRENT	SF99737	SF99737 - 740 Technology Refresh	10	10	06/13/2024
INSTALLED LEVEL IS CURRENT	SF99740	SF99740 - 740 Cumulative PTF Package	24158	24158	06/14/2024

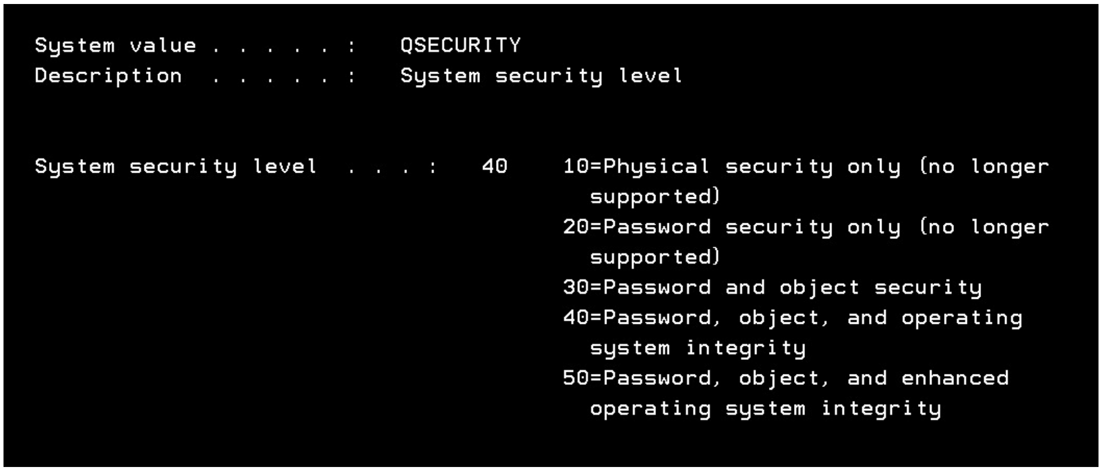
Figure 5-1 PTF currency query results.

For more information on staying current on IBM i see this document on using fixes.

5.3 Security levels

Security on IBM I systems is designed in a series of levels with each level offering a greater degree of security and protection of your data than the previous level. You can choose how much security you want the system to enforce by setting the security level (QSECURITY) system value. IBM i supports these fully-integrated system security levels: IBM i platform offers five levels of security.

Figure 5-2 shows the QSECURITY panel where the level can be set.



```
System value . . . . . :   QSECURITY
Description  . . . . . :   System security level


System security level  . . . :   40    10=Physical security only (no longer
                                         supported)
                                       20=Password security only (no longer
                                         supported)
                                       30=Password and object security
                                       40=Password, object, and operating
                                         system integrity
                                       50=Password, object, and enhanced
                                         operating system integrity
```

Figure 5-2 The QSECURITY system value and the various security levels on IBM i

Level 10: Password security (changing to level 10 is no longer supported)

At security level 10, you have no security protection and is not recommended or supported.Running at this security level is both a security and integrity risk.

Level 20: Password security (changing to level 20 is no longer supported)

At security level 20, users require both a unique user ID and password to access the system which is created by the system administrator. A significant issue with this level is that it grants users ALLOBJ special authority, allowing them unrestricted access to all system data, files, and objects. This means there's no way to restrict individual user privileges, posing a severe security risk. Given the potential for unauthorized access and data breaches, running at security level 20 is strongly discouraged and no longer supported. Higher security levels (40 and 50) offer enhanced protection mechanisms that can significantly reduce risks.

Level 30: Password and resource security

Level 30 provides more security functions in addition to what is provided at level 20. Users must have specific authority to use resources on the system. Users do not have automatic access to everything on the system and the system administrator must define a valid user ID and password for them. User access is limited by the security policies of the business. Level 30 is not considered a secure level as the integrity protection features available on security level 40 and 50 are not activated at security level 30. Running at this security level is both a security and integrity risk as you do not have the protection of the higher security levels, 40 and 50, activated and being enforced.

Level 40: Integrity protection

At this security level, resource security and integrity protection are enforced, and the system itself is protected against users. Integrity protection functions, such as the validation of parameters for interfaces to the operating system, help protect your system and the objects on it from tampering by experienced system users. For example, user-written programs cannot directly access the internal control blocks through pointer manipulation. Level 40 is the default security level for every new installation and is the recommended security level for most installations.

Level 50: Advanced integrity protection

At this security level, advanced integrity protection is added to the resource security and level 40 integrity protection enforcement. Advanced integrity protection includes further restrictions, such as the restriction of message-handling between system state programs and user state programs. Not only is the system protected against user-written programs, but it ensures that users only have access to data on the system, rather than information about the system itself. This offers greater security against anyone attempting to learn about your system. Level 50 is the recommended level of security for most businesses, because it offers the highest level of security currently possible.Security level 50 is intended for IBM i platforms with high security requirements, and it is designed to meet Common Criteria (CC) security requirements.

5.4 System values

System values are part of the global settings of your system. They allow you to customize many characteristics of your system. The security system values are used to control the security settings on your system and are broken into four groups:

- System values that control passwords
- System values that control auditing
- General security system values
- Other system values related to security

System values also provide customization on many characteristics of your IBM i platform. You can use system values to define system-wide security settings. To access the jobs category of system values from IBM Navigator for i, select Configuration and Services and then select System Values. This is shown in Figure 5-3.

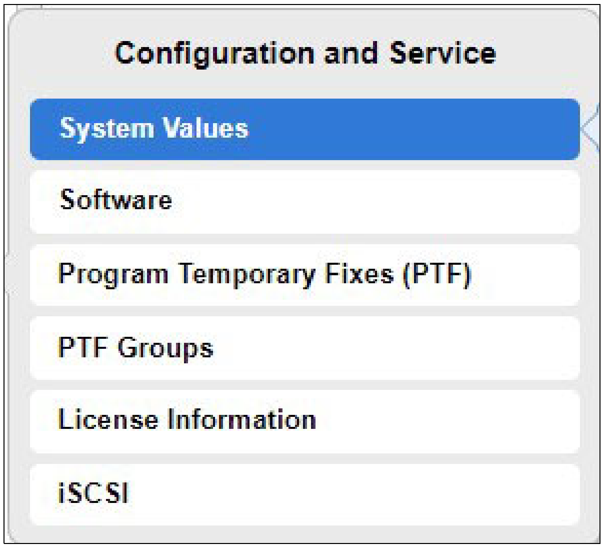


Figure 5-3 “System values” option under “Configuration and Service” menu within IBM Navigator for i

For example, you can specify the following settings:

- How many sign-on attempts you allow at a device.
- Whether the system automatically signs off an inactive workstation.
- How often passwords need to be changed.
- The length and composition of passwords.

You can restrict users from changing the security-related system values. The Change SST Security Attributes (CHGSSTSECA) command, system service tools (SST) and dedicated service tools (DST) provide an option to lock these system values. By locking the system values, you can prevent even a user with *SECADM and *ALLOBJ authority from changing these system values with the CHGSYSVAL command. In addition to restricting changes to these system values, you can also restrict adding digital certificates to digital certificate store with the Add Verifier API and restrict password resetting on the digital certificate store.

To see a list of all the security-related system values, from the IBM Navigator for i go to **Security** → **Security Config. info**. This is typically related to your security environment requirement and may differ slightly for every organization.

5.5 Authentication

Authentication is the set of methods used by organizations to ensure that only the authorized personnel, services, and applications with the correct permissions can get access to company resources. There are those who wish to gain access to your systems with ill intentions, thus making authentication a critical part of cybersecurity. These bad actors will try to steal credentials from users who already have access to your environment. Therefore, your authentication process should primarily include these three steps:

1. Identification - ensure that the one requesting access is who they are, usually through a user name or other type of login ID.
2. Authentication - users will usually provide a password (a random word or phrase or sequence of characters that a user is the only one who is supposed to know) to prove that they are who they claim to be, but if you want to strengthen security, organizations may also require the user to provide something they have (a phone or token device) to further prove their identity, or a unique characteristic that is part of their person (a face or fingerprint scan).
3. Authorization - The system then verifies that the user is indeed who he/she claims to be and allows them access to the system or application that they are trying to gain access to.

Authentication is important because it helps your organization to protect your applications, systems, data, websites, and networks from internal and external attacks. It also aids in keeping an individual’s personal data confidential, allowing them to conduct their everyday business online with less risk. When authentication systems are weak, it allows attackers to compromise a user account either by guessing ones password or tricking a person into handing over their credentials. This may lead to any of the risks below:

- Exfiltration or data breach
- Installation of various types of malware (in enterprise environments, the most prevalent of these is ransomware)
- Non-compliance with different data privacy regulations

Hence, because compromising a user’s access to a system is a common method for attackers to obtain unauthorized access to an organization’s resources, it becomes of utmost importance that strong authentication security be enforced.

5.5.1 Single sign-on enablement (SSO)

Single sign-on is an authentication process in which a user can access more than one system by entering a single user ID and password. In today's heterogeneous networks with partitioned systems and multiple platforms, administrators must cope with the complexities of managing identification and authentication for network users.

To enable a single sign-on environment, IBM provides two technologies that work together to enable users to sign in with their Windows user name and password and be authenticated to

IBM i platforms in the network. Network Authentication Service (NAS) and Enterprise Identity Mapping (EIM) are the two technologies that an administrator must configure to enable a single sign-on environment. Windows operating systems, AIX, and z/OS use Kerberos protocol to authenticate users to the network. A secure, centralized system, called a key distribution center, authenticates principals (Kerberos users) to the network.

While Network Authentication Service (NAS) allows a IBM i platform to participate in the Kerberos realm, EIM provides a mechanism for associating these Kerberos principals to a single EIM identifier that represents that user within the entire enterprise. Other user identities, such as an IBM i user name, can also be associated with this EIM identifier. When a user signs on to the network and accesses an IBM i platform, that user is not prompted for a user ID and password. If the Kerberos authentication is successful, applications can look up the association to the EIM identifier to find the IBM i user name. The user no longer needs a password to sign on to IBM i platform because the user is already authenticated through the Kerberos protocol. Administrators can centrally manage user identities with EIM while network users need only to manage one password. You can enable single sign-on by configuring Network Authentication Service (NAS) and Enterprise Identity Mapping (EIM) on your system.

Note: Full documentation of Single sign-on for IBM i 7.5 can be found at `https://www.ibm.com/docs/en/ssw_ibm_i_75/pdf/rzamzpdf.pdf.`

5.5.2 User profiles

On the IBM i operating system, every system user has a user profile, you must create a user profile before a user can sign on.

The user profile is a powerful and flexible tool. It controls what the user can do and customizes the way the system appears to the user. The following list describes some of the important security features of the user profile:

Special authority

Special authorities determine whether the user is allowed to perform system functions, such as creating user profiles or changing the jobs of other users. The special authority available are enumerated in the Table 5-1.

Table 5-1 Security-specific special authorities

Special Authority	Description
***ALLOBJ**	All-object (*ALLOBJ) special authority allows the user to access any resource on the system whether private authority exists for the user.
***SECADM**	Security administrator (*SECADM) special authority allows a user to create, change, and delete user profiles.
***JOBCTL**	The Job control (*JOBCTL) special authority allows a user to change the priority of jobs and of printing, end a job before it has finished, or delete output before it has printed. *JOBCTL special authority can also give a user access to confidential spooled output, if output queues are specified OPRCTL(*YES).
***SPLCTL**	Spool control (*SPLCTL) special authority allows the user to perform all spool control functions, such as changing, deleting, displaying, holding and releasing spooled files.

Special Authority	Description
*SAVSYS	Save system (*SAVSYS) special authority gives the user the authority to save, restore, and free storage for all objects on the system, regardless of whether the user has object existence authority to the objects.
*SERVICE	Service (*SERVICE) special authority allows the user to start system service tools using the STRSST command. This special authority allows the user to debug a program with only *USE authority to the program and perform the display and alter service functions. It also allows the user to perform trace functions.
*AUDIT	Audit (*AUDIT) special authority gives the user the ability to view and change auditing characteristics.
*IOSYSCFG	System configuration (*IOSYSCFG) special authority gives the user the ability to change how the system is configured. Users with this special authority can add or remove communications configuration information, work with TCP/IP servers, and configure the internet connection server (ICS). Most commands for configuring communications require *IOSYSCFG special authority.

Initial menu and initial program

The initial menu and program determine what the user sees after signing on the system. You can limit a user to a specific set of tasks by restricting the user to an initial menu.

Limit capabilities

The limit capabilities field in the user profile determines whether the user can enter commands and change the initial menu or initial program when signing on. The Limit capabilities field in the user profile and the ALWLMTUSR parameter on commands apply only to commands that are run from the command line, the Command Entry display, FTP, REXEC, using the QCAPCMD API, or an option from a command grouping menu. Users are not restricted to perform the following actions:

- Run commands in CL programs that are running a command as a result of taking an option from a menu
- Run remote commands through applications

5.5.3 Signing for objects

You can reinforce integrity by signing software objects that you use.

A key component of security is integrity: being able to trust that objects on the system have not been tampered with or altered. Your IBM i operating system software is protected by digital signatures.

Signing your software object is particularly important if the object has been transmitted across the Internet or stored on media which you feel might have been modified. The digital signature can be used to detect if the object has been altered.

Digital signatures, and their use for verification of software integrity, can be managed according to your security policies using the Verify Object Restore (QVFYOBJRST) system value, the Check Object Integrity (CHKOBJITG) command, and the Digital Certificate Manager tool. Additionally, you can choose to sign your own programs (all licensed programs shipped with the system are signed).

You can restrict adding digital certificates to a digital certificate store using the Add Verifier API and restrict resetting passwords on the digital certificate store. System Service Tools (SST) provides a new menu option, entitled “Work with system security” where you can restrict adding digital certificates.

5.5.4 Group profiles

A group profile is a special type of user profile. Rather than giving authority to each user individually, you can use a group profile to define authority for a group of users.

A group profile can own objects on the system. You can also use a group profile as a pattern when creating individual user profiles by using the copy profile function.

5.6 Transport Layer Security

Transport Layer Security (TLS), originally know as Secure Socket Layer or SSL, uses certificates to establish an encrypted link between a server and a client. This allows sensitive information like credit card details to be transmitted securely over the internet. The certificate contains a public key that authenticates the identity of a website and allows for encrypted data transfer through asymmetric, or public-key cryptography. The matching private key is kept secret on the server.

There are two types of encryption keys used in SSL/TLS:

- Asymmetric keys – The public and private key pair are used to identify the server and initiate the encrypted session. The private key is known only to the server, while the public key is shared via a certificate.
- Symmetric session keys – Disposable keys are generated for each connection and used to encrypt/decrypt transmitted data. The symmetric keys are securely exchanged using asymmetric encryption.

SSL/TLS supports multiple symmetric ciphers and asymmetric public key algorithms. For example, AES with 128-bit keys is a common symmetric cipher, while RSA and ECC commonly use asymmetric algorithms.

5.6.1 Secure Sockets Layer/Transport Layer Security (SSL/TLS) on IBM i

This section explores the Secure Sockets Layer (SSL) and Transport Layer Security (TLS) implementations available on IBM i. These protocols provide secure communication channels for data transmission between applications.

Overview

The IBM i system offers multiple SSL/TLS implementations, each adhering to industry-defined protocols and specifications set by the Internet Engineering Task Force (IETF). These implementations cater to different application needs and offer varying functionalities. The specific implementation used by an application depends on the chosen API set.

For Java applications, the configured JSSE provider determines the implementation, as Java interfaces are standardized. Alternatively, an application can embed its own implementation for exclusive use.

Available Implementations

Here's a breakdown of the available SSL/TLS implementations on IBM i:

- System SSL/TLS:
  - Primarily used by ILE applications.
  - Certificate management handled by the Digital Certificate Manager (DCM) with certificates stored in CMS (Certificate Management Services) format (.KDB files).
  - While Java applications can utilize System SSL/TLS, it's not the typical approach.
  - Even rarer is a Java application concurrently using both System SSL/TLS and a Java Keystore.
- IBMJSSE2 (IBMJSSEProvider2):
  - A pure Java implementation of SSL/TLS protocols available on various platforms.
  - Known as com.ibm.jsse2.IBMJSSEProvider2 in the java.security provider list.
  - Default JSSE provider for all JDK versions on IBM i, making it the most commonly used option for Java applications.
  - Certificates typically reside in Java Keystore files (.JKS) and are managed via Java keytool or IBM Key Management (iKeyman).
  - For general JSSE information, refer to the Java Secure Socket Extension (JSSE) documentation.
  - Specific details for IBMJSSE2 can be found in the platform-independent documentation for your corresponding JDK version. For JDK 8, refer to the “Security Reference for IBM SDK, Java Technology Edition, Version 8."
- OpenSSL:
  - Open-source toolkit offering SSL/TLS protocol implementation and a comprehensive cryptography library.
  - Limited availability; only accessible within the IBM Portable Application Solutions Environment for i (PASE for i).
  - Certificates typically stored in PEM files and managed using OpenSSL commands.
  - Primarily used by applications like the Common Information Model Object Manager (CIMOM).
  - Refer to the “Common Information Model” documentation for further details.

System SSL/TLS

System SSL/TLS is a set of generic services that are provided in the IBM i Licensed Internal Code to protect TCP/IP communications by using the SSL/TLS protocol. System SSL/TLS is tightly coupled with the operating system and the LIC sockets code specifically providing extra performance and security.

System TLS has the infrastructure to support multiple protocols. The following protocols can be supported by System TLS:

- Transport Layer Security version 1.3 (TLSv1.3)
- Transport Layer Security version 1.2 (TLSv1.2)
- Transport Layer Security version 1.1 (TLSv1.1)
- Transport Layer Security version 1.0 (TLSv1.0)
- Secure Sockets Layer version 3.0 (SSLv3)

System TLS also supports the following cipher suites:

- AES_128_GCM_SHA256
- AES_256_GCM_SHA384
- CHACHA20_POLY1305_SHA256
- ECDHE_ECDSA_AES_128_GCM_SHA256
- ECDHE_ECDSA_AES_256_GCM_SHA384
- ECDHE_RSA_AES_128_GCM_SHA256

- ECDHE_RSA_AES_256_GCM_SHA384
- ECDHE_ECDSA_CHACHA20_POLY1305_SHA256
- ECDHE_RSA_CHACHA20_POLY1305_SHA256

System values for setting protocols and cipher suites

The QSSLPCL system value setting identifies the specific protocols that are enabled on the system. Applications can negotiate secure sessions with only protocols that are listed in QSSLPCL. For example, to restrict the System TLS implementation to use only TLSv1.3 and not allow any older protocol versions, set QSSLPCL to contain only *TLSV1.3.

The QSSLPCL special value *OPSYS allows the operating system to change the protocols that are enabled on the system. The value of QSSLPCL remains the same when the system upgrades to a newer operating system release. If the value of QSSLPCL is not *OPSYS, then the administrator must manually add newer protocol versions to QSSLPCL after the system moves to a new release.

For the most current information on System SSL/TLS support for protocols and cipher suites see this IBM document on System SSL/TLS.

Important: IBM strongly recommends that you always run your IBM i server with the following network protocols **disabled**. Using configuration options that are provided by IBM to enable the weak protocols results in your IBM i server being configured to allow use of the weak protocols. This configuration results in your IBM i server potentially being at risk of a network security breach.

- Transport Layer Security version 1.1 (TLSv1.1)
- Transport Layer Security version 1.0 (TLSv1.0)
- Secure Sockets Layer version 3.0 (SSLv3)
- Secure Sockets Layer version 2.0 (SSLv2)

The QSSLCSL system value setting identifies the specific cipher suites that are enabled on the system. Applications can negotiate secure sessions with only a cipher suite that is listed in QSSLCSL. No matter what an application does with code or configuration, it cannot negotiate secure sessions with a cipher suite if it is not listed in QSSLCSL. Individual application configuration determines which of the enabled cipher suites are used for that application.

To restrict the System TLS implementation from using a particular cipher suite, follow these steps:

- Change QSSLCSLCTL system value to special value *USRDFN to allow the QSSLCSL system value to be edited.
- Remove all cipher suites to be restricted from the list in QSSLCSL.

The QSSLCSLCTL system value special value *OPSYS allows the operating system to change the cipher suites that are enabled on the system. The value of QSSLCSLCTL remains the same when the system upgrades to a newer operating system release. If the value of QSSLCSLCTL is *USRDFN, then the administrator must manually add in newer cipher suites to QSSLCSL after the system moves to a new release. Setting QSSLCSLCTL back to *OPSYS also adds the new values to QSSLCSL.

A cipher suite cannot be added to QSSLCSL if the TLS protocol that is required by the cipher suite is not set in QSSLPCL.

5.7 Service tools

Service tools are used to configure, manage, and service all models of the IBM i.

Service tools can be accessed from dedicated service tools (DST) or system service tools (SST). Service tools user IDs are required if you want to access DST, SST, and to use the IBM Navigator for i functions for disk unit management.

Service tools user IDs have been referred to as DST user profiles, DST user IDs, service tools user profiles, or a variation of these names. Within this topic collection, the term “service tools user IDs” is used.

> **Note:** Full documentation of Service Tools for IBM i 7.5 can be found at
> `https://www.ibm.com/docs/en/ssw_ibm_i_75/pdf/rzamhpdf.pdf`

5.7.1 System Service Tool (SST)

If your user profile has the required authorizations, you can use system service tools (SST) to access service tools.

The service tools user ID you use to access SST needs to have the functional privilege to use SST. The IBM i user profile needs to have the following authorizations:

- Authorization to the Start SST (STRSST) CL command.
- Service special authority (*SERVICE).

To access service tools using SST, complete the following steps:

1. Enter **`STRSST`** (Start SST) on an IBM i command line. The Start SST Sign On display is shown.
2. Enter the following information:
   – Service Tools User ID: The service tools user ID you sign on with.
   – Password: The password associated with this user ID.
3. Press **`Enter`**.
4. The System Service Tools (SST) display is shown.

To exit from SST after performing the desired action, press **`F3 (Exit)`** until you get to the Exit System Service Tools display then press **`Enter`** to end SST.

5.7.2 Dedicated Service Tool (DST)

To access service tools, you can use dedicated service tools (DST) from the system console.

The service tools user ID that you use to access service tools with DST needs to have the functional privilege to use DST. You can start the DST by using function 21 from the system control panel or by using a manual initial program load (IPL).

Accessing service tools using DST from the system control panel

To access service tools using DST from the control panel, complete the following steps:

1. Put the control panel in manual mode.
2. Use the control panel to select function 21 and press Enter. The DST Sign On display appears on the console.
3. Sign on to DST using your service tools user ID and password. The Use dedicated service tools (DST) display appears.

4. Select the appropriate option from the list and press Enter.
   – Select option 5 (Work with DST environment) to get to additional options for working with service tools user IDs.
   – Select option 7 (Start a service tool) to start any of the service tools available from DST.
   – Select any of the other options, as appropriate.

Accessing service tools using DST from a manual IPL

To access service tools using DST from a manual initial program load (IPL), complete the following steps:

1. Put the control panel in manual mode.
2. Take either of the following actions:
   – If the system is powered off, turn the system on.
   – If the system is powered on, enter the Power Down System (PWRDWNSYS) command, `PWRDWNSYS *IMMED RESTART(*YES)`, on a command line to turn off the system and restart it.
3. Sign on to DST using your service tools user ID and password. The IPL or Install the System menu is shown.
4. At the IPL or Install the System menu:
   – Select option 3 (Use Dedicated Service Tools (DST)).
   – Select the appropriate option from the list and press `Enter`.
   – Select option 5 (Work with DST environment) to get additional options for working with service tools user IDs.
   – Select option 7, (Start a service tool) to start any of the service tools available from DST.
   – Select any of the other options, as appropriate.

Exiting from DST

To exit from DST after performing the desired action, press `F3 (Exit)` until you return to the Exit Dedicated Service Tools (DST) display then select option 1 (Exit Dedicated Service Tools (DST)).

5.8 Digital Certificate Manager

Digital Certificate Manager (DCM) allows you to manage digital certificates for your network and use Transport Layer Security (TLS) to enable secure communications for many applications.

A digital certificate is an electronic credential that you can use to establish proof of identity in an electronic transaction. There are an increasing number of uses for digital certificates to provide enhanced network security measures. For example, digital certificates are essential to configuring and using the TLS. Using TLS allows you to create secure connections between users and server applications across an untrusted network, such as the Internet. TLS provides one of the best solutions for protecting the privacy of sensitive data, such as user names and passwords, over the Internet. Many IBM i platforms and applications, such as FTP, Telnet, HTTP Server provide TLS support to ensure data privacy.

IBM i provides extensive digital certificate support that allows you to use digital certificates as credentials in a number of security applications. In addition to using certificates to configure TLS, you can use them as credentials for client authentication in both TLS and virtual private network (VPN) transactions. Also, you can use digital certificates and their associated security keys to sign objects. Signing objects allows you to detect changes or possible tampering to object contents by verifying signatures on the objects to ensure their integrity.

Capitalizing on the IBM i support for certificates is easy when you use Digital Certificate Manager, a free feature, to centrally manage certificates for your applications. DCM allows you to manage certificates that you obtain from any Certificate Authority (CA). Also, you can use DCM to create and operate your own local CA to issue private certificates to applications and users in your organization.

Proper planning and evaluation are the keys to using certificates effectively for their added security benefits.

5.9 Encryption and Cryptography

Cryptography is the study and implementation of processes which manipulate data for the purpose of hiding and authenticating information. A comprehensive cryptography solution is an important part of a successful security strategy. The use of cryptography for encryption of data as it is processed within the IBM i partition provides enhanced security for memory and stored data as part of the IBM Power10 infrastructure.

As discussed in 2.1, “Encryption technologies and their applications” on page 29, Power10 provides Transparent Memory Encryption which transparently encrypts and protects memory within the system utilizing the encryption acceleration processors built into the Power10 processing chip, providing protection without performance penalties.

IBM i offers various levels of encryption for databases and attached storage devices. Using Field Procedures within IBM DB2®, IBM i provides field-level encryption to directly protect sensitive data fields within the database. Additionally, IBM i supports encryption for directly attached storage devices to safeguard data at rest within the system.

IBM i includes both software cryptography and a range of cryptographic hardware options for data protection and secure transaction processing. Users can leverage the built-in encryption acceleration processors on the Power10 chip or integrate specialized cryptographic coprocessors—both options provide robust security without compromising performance.

IBM i cryptographic services help ensure data privacy, maintain data integrity, authenticate communicating parties, and prevent repudiation when a party denies having sent a message.

Cryptographic Services Key Management

Cryptographic services key management for the IBM i operating system allows you to store and manage master keys and keystores. Since you are exchanging sensitive data to manage master keys and keystores, it is recommended that you use a secure session.

Cryptographic Services supports a hierarchical key system. At the top of the hierarchy is a set of master keys. These keys are the only key values stored in the clear (unencrypted). Cryptographic services securely stores the master keys within the IBM i Licensed Internal Code (LIC).

Eight general-purpose master keys are used to encrypt other keys which can be stored in keystore files. Keystore files are database files. Any type of key supported by cryptographic services can be stored in a keystore file, for example AES, RC2, RSA, SHA1-HMAC.

In addition to the eight general-purpose master keys, cryptographic services supports two special-purpose master keys. The ASP master key is used for protecting data in the Independent Auxiliary Storage Pool (in the Disk Management GUI this is known as an Independent Disk Pool). The save/restore master key is used to encrypt the other master keys when they are saved to media using a Save System (SAVSYS) operation.

You can work with Cryptographic services key management using the IBM Navigator for i interface as shown in Figure 5-4.

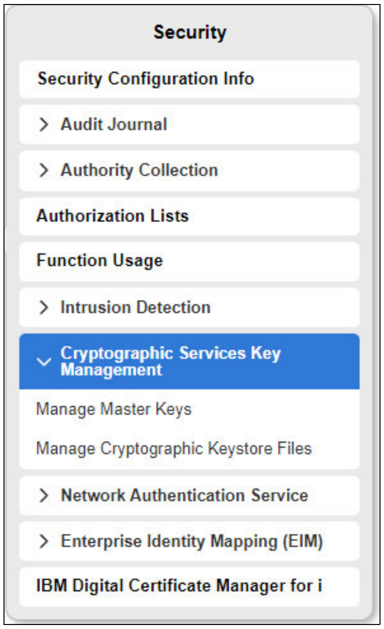

Security
Security Configuration Info
Audit Journal
Authority Collection
Authorization Lists
Function Usage
Intrusion Detection
Cryptographic Services Key Management
Manage Master Keys
Manage Cryptographic Keystore Files
Network Authentication Service
Enterprise Identity Mapping (EIM)
IBM Digital Certificate Manager for i

Figure 5-4 “Security menu within IBM Navigator for i.

After you connect to IBM Navigator for i, click **Security** → **Cryptographic Services Key Management**. You can, thereafter, work with managing master keys and cryptographic keystore files.

You can also use the cryptographic services APIs or the control language (CL) commands to work with the master keys and keystore files.

Note: You should use Transport Layer Security (TLS) to reduce the risk of exposing key values while performing key management functions.

4769 Cryptographic Coprocessor

IBM offers Cryptographic Coprocessors, which are available on a variety of system models. Cryptographic Coprocessors contain hardware engines, which perform cryptographic operations used by IBM i application programs and IBM i TLS transactions. The 4769 (which requires IBM i 7.3 or later) appears as hardware feature #EJ35 or #EJ37 on Power10-based systems. More information on the IBM Cryptographic cards is provided in section “IBM Cryptographic Coprocessor cards” on page 32.

Note: The IBM 4767 Cryptographic Coprocessor is no longer available but it is still supported.

Full documentation for cryptography can be found at `https://www.ibm.com/docs/en/ssw_ibm_i_75/pdf/rzajcpdf.pdf.`

5.10 Resource security

The ability to access an object is called authority. Resource security on the IBM i operating system enables you to control object authorities by defining who can use which objects and how those objects can be used.

You can specify detailed authorities, such as adding records or changing records. Or you can use the system-defined subsets of authorities: *ALL, *CHANGE, *USE, and *EXCLUDE.

Files, programs, and libraries are the most common objects requiring security protection, but you can specify authority for any object on the system. The following list describes the features of resource security:

Group profiles

A group of similar users can share the same authority to use objects. See 5.5.4, “Group profiles” on page 132 for more information.

Authorization lists

Objects with similar security needs can be grouped in one list. Authority can be granted to the list rather than to the individual objects.

Object ownership

Every object on the system has an owner. Objects can be owned by an individual user profile or by a group profile. Correct assignment of object ownership helps you manage applications and delegate responsibility for the security of your information.

Primary group

You can specify a primary group for an object. The primary group’s authority is stored with the object. Using primary groups may simplify your authority management and improve authority checking performance.

Library authority

You can put files and programs that have similar protection requirements into a library and restrict access to that library. This is often easier than restricting access to each individual object.

Directory authority

You can use directory authority in the same way that you use library authority. You can group objects in a directory and secure the directory rather than the individual objects.

Object authority

In cases where restricting access to a library or directory is not specific enough, you can restrict authority to access individual objects.

Public authority

For each object, you can define what kind of access is available for any system user who does not have any other authority to the object. Public authority is an effective means for securing information and provides good performance.

Adopted authority

Adopted authority adds the authority of a program owner to the authority of the user running the program. Adopted authority is a useful tool when a user needs different authority for an object, depending on the situation.

Authority holder

An authority holder stores the authority information for a program-described database file. The authority information remains, even when the file is deleted. Authority holders are commonly used when converting from the System/36, because System/36 applications often delete files and create them again.

Field level authority

Field level authorities are given to individual fields in a database file. You can use SQL statements to manage this authority.

5.11 Security audit journal

You can use security audit journals to audit the effectiveness of security on your system.

The IBM i operating system provides the ability to log selected security-related events in a security audit journal. Several system values, user profile values, and object values control which events are logged.

The security audit journal is the primary source of auditing information about the system. This section describes how to plan, set up, and manage security auditing, what information is recorded, and how to view that information.

A security auditor inside or outside your organization can use the auditing function that is provided by the system to gather information about security-related events that occur on the system.

You can define auditing on your system at three different levels:

- System-wide auditing that occurs for all users.
- Auditing that occurs for specific objects.
- Auditing that occurs for specific users.

When a security-related event that might be audited occurs, the system checks whether you have selected that event for audit. If you have, the system writes a journal entry in the current receiver for the security auditing journal (*QAUDJRN* in library *QSYS*).

When you want to analyze the audit information you have collected in the journal you can use IBM Navigator for i to display the output. You can also use SQL commands as documented in this link: `https://www.ibm.com/docs/en/i/7.5?topic=services-journal.`

5.12 Independent disk pool

An Independent disk pool, also known as an Independent Auxiliary Storage Pool, provides the ability to group together storage that can be taken offline or brought online independent of system data or other unrelated data. The terms independent auxiliary storage pool (iASP) and independent disk pool are synonymous.

Auxiliary storage is the permanent disk space that we assign to the system in the form of disk units which of course can be physical or virtual. The disk pool can contain objects, libraries, directories and many other objects such as object attributes. The concept of IASP also forms a base for high availability or disaster recovery solutions like PowerHA.

The concept of iASP is very straightforward and there are many solutions built around it. iASPs provide an attractive solutions for clients who are looking at server consolidation and continuous availability with a minimum amount of downtime. Using the iASP provides both technical and business advantages on IBM i.

The key difference between the system auxiliary storage pool (ASP) and an IASP is that the system ASP is always accessible when system is up and running, while IASP can be brought online or offline independent of the system activity on any other pools.

An IASP must be brought online or "varied on" to make it visible to the system before making any attempt to access data on it. If you want to make the IASP inaccessible by the system you "Vary off" the IASP. The vary on process is not instantaneous and can take several minutes š the amount of time required depends on several factors. Figure 5-5 shows a system with its SYSBAS or ASP and an iASP defined.

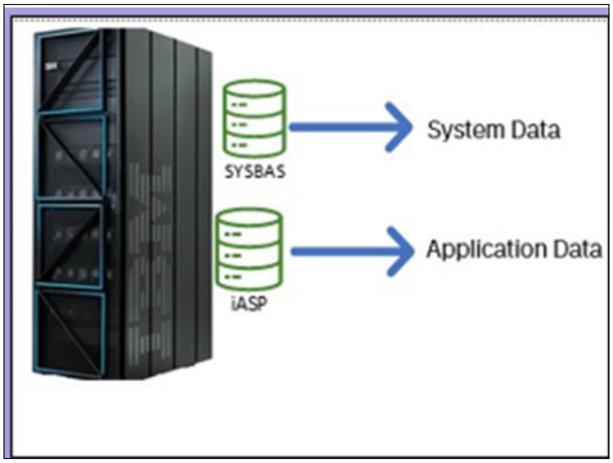
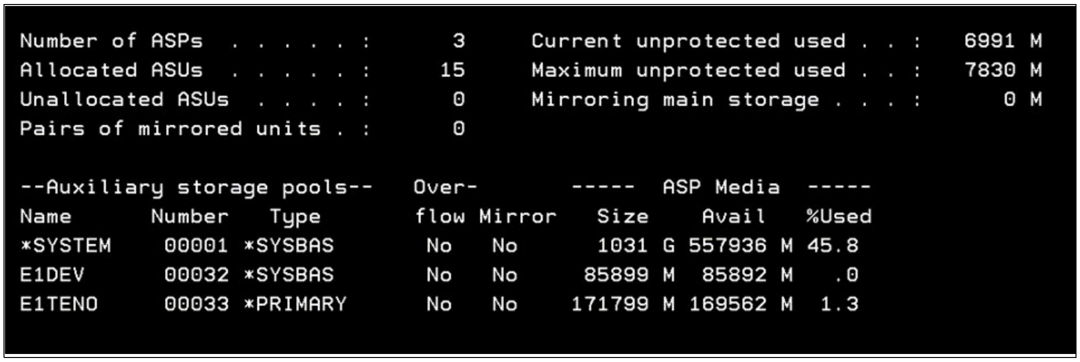


Figure 5-5 Single system having an iASP where application data resides

Independent ASPs are always numbered starting from 33 up through 255 while the basic ASPs are always numbered 2 through 32. All basic ASPs are automatically available when the system is online and cannot be independently varied on or off. Figure 5-6 shows a system with a system pool, a user ASP, and an iASP defined.



```
Number of ASPs  . . . . . :      3     Current unprotected used . . :   6991 M
Allocated ASUs  . . . . . :     15     Maximum unprotected used . . :   7830 M
Unallocated ASUs  . . . . :      0     Mirroring main storage . . . :      0 M
Pairs of mirrored units . :      0

--Auxiliary storage pools--   Over-       -----  ASP Media  -----
Name       Number    Type       flow Mirror    Size     Avail   %Used
*SYSTEM     00001 *SYSBAS        No    No      1031 G 557936 M 45.8
E1DEV       00032 *SYSBAS        No    No     85899 M  85892 M   .0
E1TENO      00033 *PRIMARY       No    No    171799 M 169562 M  1.3
```

Figure 5-6 System having user ASP and iASP

An iASP can be deployed using external storage—often referred to as SAN storage—and also can be defined on internal disks. However, there are many benefits of using SAN instead of internal disk as SAN storage—combined with an IASP—allows cluster configurations where the use of replication technologies available on the SAN storage can provide high availability and disaster recovery options. Many clients across many different sectors—Manufacturing, Insurance, Aviation, Banking and Retail for example—currently running with this configuration.

An independent disk pool can either be:

Switchable An IASP which is used across two or more IBM i partitions in a clustered environment. When the switch occurs and the independent disk pool is “varied on”, the entire contents on the IASP is available.

Non-Switchable An IASP which is created locally to the system and is not shared with any other system.

When considering IASP implementation, business needs should be considered first and accordingly a plan should be made to implement this in the client environment. At the application level you should have a good understanding about where objects reside, who are the users and how the program and data is accessed. There are certain type of objects that while they are supported in IASP, they should remain in the system ASP only in order to maintain the expected or normal behavior of the system. Some work management related changes will need to be made with the introduction of an IASP. In general there are two environments in which IASP can be used.

- Single system environment

  In this case you have an Independent disk pool on the local system. It can be brought online or offline without impacting other storage pools on the system or the need to do an Initial Program Load (IPL). This is often used in a local system which contains multiple databases located on IASPs. The iASP can be made available while system is active without the need to perform an IPL and the independent disk pool can remain offline until it is needed. This kind of setup is very common if you want to segregate application data, keep historical and archived data on same system, maintain multiple application versions or to meet data compliance rules where you need to have data in different pools and keep it offline unless needed by the business.

- Multi-system environment

  In this case you have one or more IASPs which are shared between multiple IBM i partitions—on the same system or on different systems, possibly even in another locations—that are members of the cluster. In this kind of setup this IASP can be switched between these systems without the need of any IPL for any of the partitions. This is quite a significant advantage because it allows continuous availability of the data. There can be various reasons to implement IASPs in multi-system environments. For example, if you are implementing a new disaster recovery or high availability solution then you would normally choose a switchable IASP setup for the most flexible implementation.

Some of the practical implementation of Independent Disk pool are as follows.

- Disaster recovery and High availability Scenarios:

  When planning for a high availability (HA) or disaster recovery (DR) solution, one of the key prerequisites for PowerHA—which provides geographical mirroring, Hyperswap, metro distance replication or global replication—is the use of an iASP. This allows you to have an active partition at the target site for minimal interruption while moving workloads between systems. PowerHA is an integrated extension of the IBM i operating system and offers environmental, application, and data resiliency solutions for managing data access and replicated storage in the event of planned or unplanned outages. This is a solution which provides near continuous availability of data. The time to recover user access to their applications depends on many factors in your environment which include the distance between your data centers and the bandwidth available between them. Using a tool like PowerHA allows you to avoiding the use of a software replication tool.

Use of an iASP can also be combined with IBM DB2 mirror for i where continuous availability is implemented by use of IASP for Integrated File System (IFS). There is very minimal maintenance required when this solution is implemented. The data can remain available to an application even in the case of a system outage. Based on implementation choices, switchover is automated.

- Data Isolation or Protection:

  An iASP can also be used when you want to divide data between multiple databases. It can also be used to store data that is only accessed occasionally based on business requirements.Usin the IASP in this manner means that the data is only brought online when needed which provides additional protection in terms of data access/ The data remains offline until needed and hence no one can access this portion of storage until it is varied on. When it is varied on you must have proper rights to access it. Customers can use this to store and retain any historical data which they might need to use from time to time and as a part of security compliance where it can remain offline and inaccessible. Most common use is also to maintain different application versions on the same system and make efficient use of it.

- Flash copy

  Using an iASP also provides the ability to do a full system flash copy where a flash is taken for the IBM i environment. This delivers an almost instantaneous copy of the data in the shortest possible time. This copy can be varied on to a separate partitions using the IASP. Once the iASP is varied on, the data is now available to access in order to save the data to physical or virtual tapes. This process can be automated using Backup Recovery and Media Services (BRMS) is fully integrated with IASP to automate the full process and also provide reporting.

5.13 Exit points

An *exit point* signifies the point in a system function or program where control is turned over to one or more exit programs to perform a function.

The *registration facility* provides a central point to store and retrieve information about IBM i and non-IBM i exit points and their associated exit programs. This information is stored in the registration facility repository and can be retrieved to determine which exit points and exit programs already exist.

You can use the registration facility APIs to register and unregister exit points, to add and remove exit programs, and to retrieve information about exit points and exit programs. You can also perform some of these functions by using the Work with Registration Information (`WRKREGINF`) command.

The *exit point provider* is responsible for defining the exit point information, defining the format in which the exit program receives data, and calling the exit program. There are particularly four areas in which Exit points provide a another layer of security.

Securing network access

When it comes to different network protocols on IBM i like FTP, Netserver, JDBC, ODBC, DDM, DRDA and others which are used extensively by users to connect to back end database on IBM i. IBM provides dozens of exit points that cover most network access protocols, which means that exit programs can be created and assigned to these exit points, not only to monitor and log activity but, most importantly, to control access by a variety of criteria

Securing communication port access

There are a handful of network protocols that don't have their own exit points and thus can't be protected in the same way as protocols that have specific exit points. These network protocols include SSH, SFTP, SMTP, and others. In addition, organizations may need to control communication access in a way network or other types of exit points cannot because it is not possible to specify a port number in these other types of exit points. For example, it may be important for a specific type of network connection to be able to use only one or more secured ports.

IBM provides socket exit points that make it possible to develop exit programs for securing connections to your IBM i by specific ports and/or IP addresses

Securing database access

One particularly powerful exit point is called Open Database File, and it allows development of exit programs that protect sensitive data from any kind of access. The added layer of security this exit point provides is significant because of its ability to invoke an exit program whenever a specified file on the system is opened, whether it's a physical file, logical file, SQL table, or SQL view. As with other exit points, your exit program can be defined to audit all activity, such as the user, the method of access, the date/time, and the operation (read, update, add, or delete). Plus, the exit program can contain a granular set of rules that control under what conditions the file can be accessed and by whom.

Securing command access

IBM provides exit points that cover the use of commands, thus making it possible to develop exit programs that allow or disallow access to any command within very specific circumstances, regardless of whether the access attempt comes from a user performing a command-line function directly within the IBM i, through network access, or otherwise. Because command exit programs supersede normal object-level security, they add an additional, very useful layer of security that can control the use of commands even for users with powerful authorities such as *ALLOBJ or *SECADM. As with other types of exit points, command exit programs can be defined in such a way that each command can have its own specific rules of usage, while providing logging of any activity.

5.14 Function usage

Function usage allows you to define who can use an application, the parts of an application, or the functions within a program.

This support is not a replacement for resource security. Function usage does not prevent a user from accessing a resource (such as a file or program) from another interface. Function usage support provides APIs to perform the following tasks:

- Register a function
- Retrieve information about the function
- Define who can or cannot use the function
- Check to see if the user is allowed to use the function

To use a function within an application, the application provider must register the functions when the application is installed. The registered function corresponds to a code block for specific functions in the application. When the user runs the application, before the application invokes the code block, it calls the check usage API to verify that the user has the authority to use the function that is associated with the code block. If the user is allowed to use the registered function, the code block runs. If the user is not allowed to use the function, the user is prevented from running the code block.

The system administrator specifies who is allowed or denied access to a function. The administrator can either use the *Work with Function Usage Information* (**WRKFCNUSG**) command to manage the access to program function or use **Security** → **Function Usage** in the IBM Navigator for i.

Separation of duties

Separation of duties helps businesses comply with government regulations and simplifies the management of authorities. It provides the ability for administrative functions to be divided across individuals without overlapping responsibilities, so that one user does not possess unlimited authority—such as with *ALLOBJ authority. The function, QIBM_DB_SECADM, provides a user with the ability to grant authority, revoke authority, change ownership, or change primary group, but without giving access to the object or, in the case of a database table, to the data that is in the table or allowing other operations on the table.

QIBM_DB_SECADM function usage can be given only by a user with *SECADM special authority and can be given to a user or a group.

QIBM_DB_SECADM is also responsible for administering Row and Column Access Control. Row and

Column Access Control provides the ability to restrict which rows a user is allowed to access in a table and whether a user is allowed to see information in certain columns of a table. For more information, see Row and column access control (RCAC).

> **Note:** You can find extensive documentation on IBM i security by viewing the IBM i *7.5 Security Reference* at https://www.ibm.com/docs/en/ssw_ibm_i_75/pdf/sc415302.pdf and IBM i *7.5 Security - Planning and setting up system security* at https://www.ibm.com/docs/en/ssw_ibm_i_75/pdf/rzamvpdf.pdf for a more in-depth discussion.
>
> There is also a dedicated topic on IBM i security found in IBM Documentation at https://www.ibm.com/docs/en/i/7.5?topic=security.

5.15 IFS Security on IBM i

IFS refers to “Integrated File System.” This integrated file system is a component of the IBM i operating system, facilitating stream input/output and storage management similar to personal computer and UNIX systems, while offering a cohesive structure for all information stored within the system. Ensuring IFS security is a primary concern when developing security strategies for IBM i. One of the significant challenges faced by security administrators is managing security related to the root directory ('/'). This root directory houses the majority of IBM i products, third-party applications, configuration files, code, and data.

A major concern is that the root directory “/” is publicly accessible, with the default setting allowing full access for public users. Upon installation of a new IBM i operating system, the default permission for root is set to *RWX, which poses a considerable risk and should be modified. The IFS enables users to store and manage various types of data, such as documents, images, program source code, and more. It offers a unified interface for accessing and managing files across different platforms, simplifying the integration of IBM i systems with other systems in a diverse IT environment. Often, the data stored in IFS is sensitive and requires robust security measures.

Figure 5-7 is a conceptual view of the Integrated File System in IBM i.

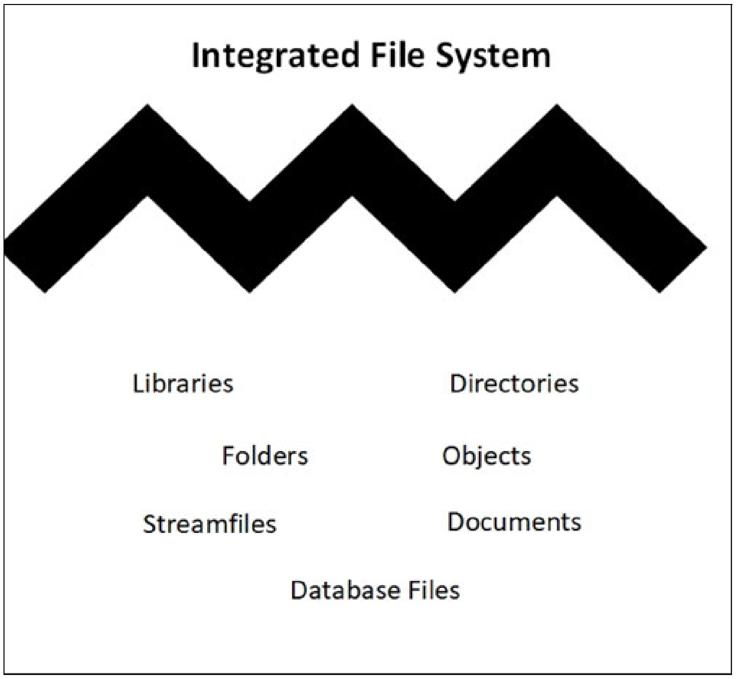


Figure 5-7 A structure for all information stored in the IBM i *operating system*

Virus Scanning for the IFS

Viruses are designed to target specific computer architectures. The unique architecture of the IBM i system significantly reduces the likelihood of a virus being developed to exploit it. Viruses intended for PC environments cannot operate on the IBM i operating system, and IBM does not offer any dedicated Anti-Virus, Anti-Spyware, Anti-Malware, or Anti-Ransomware solutions for this platform. However, if a file infected on a PC is transferred to the Integrated File System (IFS) and subsequently shared with another PC, it can potentially spread the virus to that new machine. Similarly, if a network drive is connected to the IFS, a virus from a PC that can affect files on a network drive may also compromise files stored on the IFS.

IBM i does support scanning for malicious activities through third-party software. Users can scan objects within the integrated file system, providing them with the flexibility to determine the timing of scans and the actions to take based on the outcomes. There are two exit-points related to this support are:

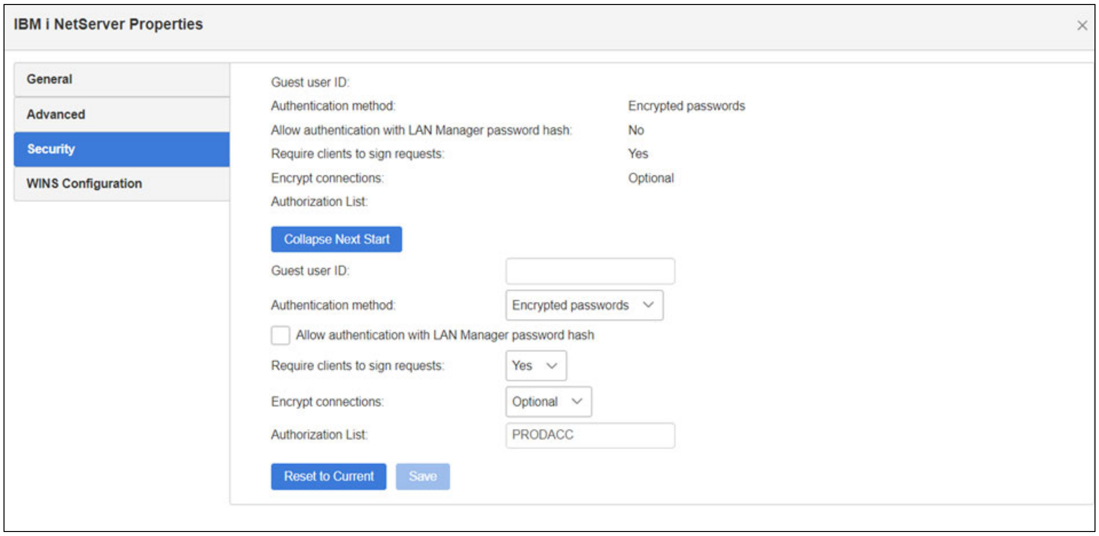
- "QIBM_QP0L_SCAN_OPEN - Integrated File System Scan on Open Exit Program. For this exit point, the integrated file system scan on open exit program is called to do scan processing when an integrated file system object is opened under certain conditions.
- "QIBM_QP0L_SCAN_CLOSE - Integrated File System Scan on Close Exit Program. For this exit point, the integrated file system scan on close exit program is called to do scan processing when an integrated file system object is closed under certain conditions.

Note: Only objects in file systems that have been fully converted to *TYPE2 directories will be scanned.

Example for securing a file share directory

Standard IBM i security processes are used to set the security of files in the integrated file system.

Figure 5-8 shows setting a file share directory (/PTF) which is secured by limiting access to members of the authorization list PRODACC.



IBM i NetServer Properties

Guest user ID:

Authentication method: Encrypted passwords

Allow authentication with LAN Manager password hash: No

Require clients to sign requests: Yes

Encrypt connections: Optional

Authorization List:

Guest user ID:

Authentication method: Encrypted passwords

□ Allow authentication with LAN Manager password hash

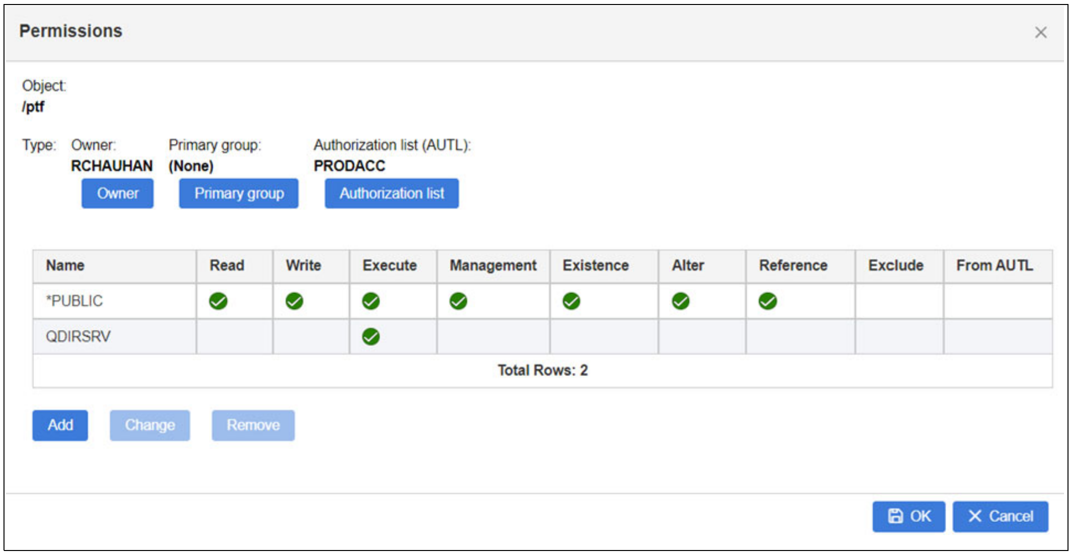
Require clients to sign requests: Yes

Encrypt connections: Optional

Authorization List: PRODACC

Figure 5-8 Figure 5-9 Limiting access of /ptf to an access list

Figure 5-9 shows the interface to display the current access permissions for directory /ptf. Additional access can be set from this screen.



Permissions

Object:
/ptf

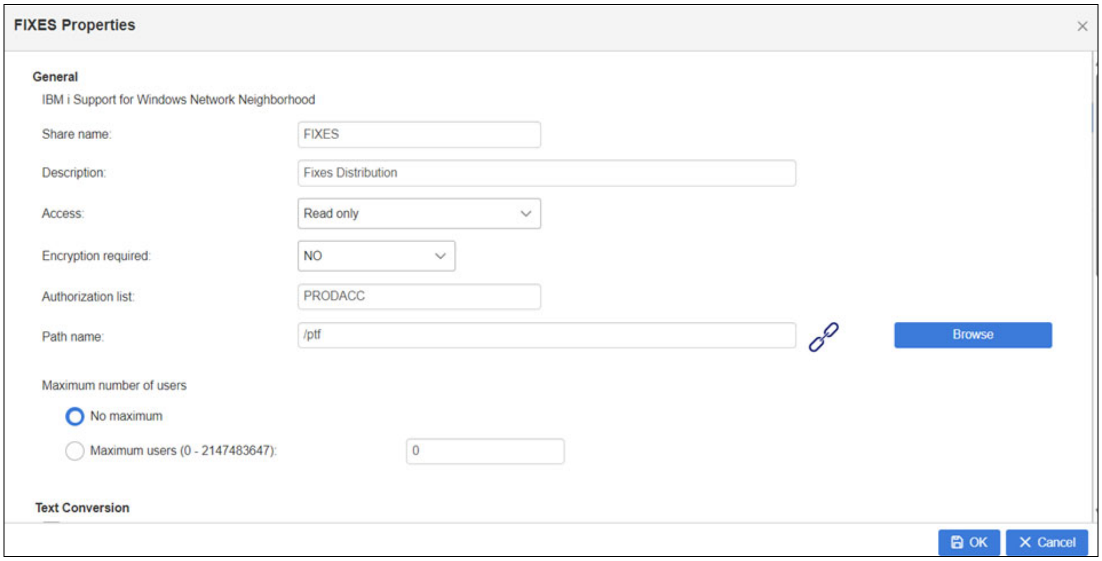
Type: Owner: **RCHAUHAN** Primary group: **(None)** Authorization list (AUTL): **PRODACC**

Name	Read	Write	Execute	Management	Existence	Alter	Reference	Exclude	From AUTL
*PUBLIC	✓	✓	✓	✓	✓	✓	✓		
QDIRSRV			✓						

Total Rows: 2

Figure 5-9 Permissions display for /ptf

Finally, Figure 5-10 displays the definition of a shared directory (FIXES) which points to the /ptf directory and defines the access as limited to members in the PRODACC group.



FIXES Properties

General

IBM i Support for Windows Network Neighborhood

Share name:	FIXES
Description:	Fixes Distribution
Access:	Read only
Encryption required:	NO
Authorization list:	PRODACC
Path name:	/ptf

Maximum number of users

● No maximum

○ Maximum users (0 - 2147483647): 0

Text Conversion

Figure 5-10 Share definition for FIXES which connects to /ptf

Tips to enhance security measures

Consider the following to enhance security measures for the integrated file system:

1. Review File Shares

   A systematic review of the file shares should take place on a regular basis. This evaluation is important, as it can greatly lower risk by ensuring that any unnecessary shares in the system are removed.

2. Set shares to Read-only wherever possible

   Acknowledging the potential risks associated with write access, particularly in the context of ransomware or malware, it is vital to implement access review processes. The primary goal of these exercises is to lower the risk of security breaches by routinely examining access rights. This involves making educated decisions about who has access to sensitive data and critical resources, and determining whether such access is warranted for each user. This includes root and should be changed to *PUBLIC use only.

3. Do not share the root or /QSYS.lib

   Sharing the root directory ("/") or /QSYS.lib poses a considerable security threat, especially in the absence of effective access controls. If someone maps to root they can see the entire structure of then system. This practice reveals all files and directories under the root, which is not advisable; therefore, it is better to create specific shares as needed, no higher.

4. Use Authority Collection if have unclear on the usage of shares by users

   The authority collection feature is included as part of the core operating system. It operates by collecting data linked to the run-time authority verification processes embedded within the IBM i system. The primary goal of this feature is to support security administrators and application providers in safeguarding application objects with the least amount of authority required for effective operation. Utilizing authority collection to reduce or prevent excessive authority contributes to strengthening the overall security of the objects employed by an application.

5. Restrict server and share access with Authorization list

   Starting in IBM i 7.5, user access to IBM i NetServer and specific shares can be restricted by assigning an authorization list. IBM i NetServer now allows assigning an authorization list object to the server and individual shares. The authorization list is used as an extra layer of protection for shared resources. Updating the configuration can be performed through Navigator by changing IBM i NetServer Properties or Share properties or by using the IBM i NetServer APIs. A user must have at least *USE authority to the authorization list to access the server when an authorization list is assigned.

> **Important:** Authorization lists do not restrict access to users with *ALLOBJ special authority. Any user profile with *ALLOBJ special authority will be able to access IBM i NetServer as if there is no authorization list restriction in place. This can be used to create administrative shares that can only be accessed by IBM i administrative profiles by specifying an authorization list that only lists public *EXCLUDE.

You can find extensive documentation on IBM i Netserver security at https://www.ibm.com/docs/en/i/7.5?topic=netserver-i-security

Additional information on the IBM i 7.5 Integrated File system can be found at https://www.ibm.com/docs/en/ssw_ibm_i_75/pdf/rzaaxpdf.pdf

5.16 Security Compliance Tools for IBM i

IBM Technology Expert Labs for Security is your accelerator to maximum value with our unmatched IBM product expertise. Our mission is to successfully deploy, optimize and expand our clients’ IBM Security platforms, utilizing the most knowledgeable IBM Security Software experts.

The IBM Technology Expert Labs team for IBM i Security is an IBM team that helps specializing in IBM i security services such as security assessments, system hardening and developing IBM i utilities. This family of utilities goes under the name “Security Compliance Tools for IBM i.”

For more details on these offerings see “Security assessment for IBM Power from IBM Technology Expert Labs” on page 244

6

Linux security and compliance on IBM Power

IBM Power servers allow for the operation of Linux applications that leverage the superior capabilities of IBM Power hardware, including high performance, dependability, and resilience. With the increasing adoption of Linux workloads on IBM Power systems, ensuring stringent and accurate security measures across those Linux landscapes assumes paramount importance.

This chapter provides an overview of Linux on Power, highlighting its unique features and challenges. It discusses various supported Linux environments and offers guidance on implementing robust security measures to establish a secure and high-performing Linux system. By combining the strengths of Linux and IBM Power technology, organizations can benefit from a powerful and flexible infrastructure.

We discuss the following:

- “Overview” on page 152
- “Linux on IBM Power” on page 153
- “Hardening Linux Systems” on page 156
- “Best practices” on page 182
- “Develop Incident Response Plan” on page 183

6.1 Overview

We are moving towards a multi-architecture world, and that has implications for security on IBM Power systems. As more mission-critical environments are being implemented on Linux running on IBM Power, it is even more important to guarantee the security and compliance of those systems.

Linux is an open source based system in nature. In contrast to AIX or IBM i, which experienced significantly fewer than ten reported vulnerability reports in 2023, the Linux kernel suffered more than a hundred documented flaws during the same period. Given its open nature and extensive user base, this outcome was predictable and it makes the task of protecting Linux workloads even more critical.

This chapter provides a comprehensive guide to understanding the various threats, implementing effective security measures, and adopting best practices to ensure data integrity and confidentiality on Linux systems running on IBM Power architecture, regardless of the distribution chosen. In the chapter we will address pertinent security and Linux compliance concerns.

In regard to security, the discussion will encompass practices, processes, and tools that are specifically designed to safeguard Linux systems on Power from cyber threats, thereby ensuring the confidentiality, integrity, and availability (CIA triad) of these systems.

Ensuring robust security in Linux environments necessitates a comprehensive strategy that integrates multiple layers of defense. This strategy should encompass meticulous configuration management, proactive vulnerability assessments, and strict adherence to regulatory compliance frameworks.

The intricate nature of Linux systems demands a diverse set of tools and methodologies to effectively reduce the attack surface and bolster defenses against both established and emerging threats.

Note: In our laboratory setting, we utilize a variety of distributions, including Red Hat, SUSE, and Ubuntu, as well as Debian, CentOS, Fedora, Alma, Rocky, and OpenSUSE, which offers robust support for the ppc64le architecture.

6.1.1 Implementation and Best Practices

The method of implementing each security measure and maximizing the utility of the available tools will vary depending on the distribution and version chosen. Accordingly, this guide will delineate general principles without delving into the minutiae of specific configurations, which may vary and be subject to constant change. However, we will provide straightforward examples and guidelines on how to apply this knowledge using open source software that has been tested on ppc64le and the solutions that we believe are most effective in achieving our objective of making our Linux systems on Power as secure as possible.

6.2 Threats

Linux systems serve numerous use cases effectively, yet they remain susceptible to various threats that could jeopardize data integrity and security, thereby increasing the attack surface in Power Systems due to their extensive and uniform deployment across a vast user base. To address these challenges, adopting a holistic and cross-functional strategy is crucial to comprehend and counteract the diverse array of threats, including malware, unauthorized access, and data breaches, amongst others.

On the other hand, Linux serves as a valuable asset in enhancing system security, offering a wide range of potent tools, many of which are freely available and open-source and compatible with ppc64le architecture. By implementing these tools and adhering to best practices, it's possible to bolster the overall security posture of Power Infrastructure, starting from a basic LPAR equipped with Linux and suitable software configurations.

6.2.1 Malware

Malware, including viruses, worms, Trojans, and ransomware, poses significant risks to Linux systems on Power. These malicious programs can disrupt operations, steal sensitive information, and cause substantial financial and reputational damage.

Implementing a comprehensive security strategy that includes anti-virus solutions, regular system scans, security patches, strong authentication measures, and user awareness training is crucial for safeguarding Linux systems from malware threats.

6.2.2 Unauthorized Access

Unauthorized access can materialize via multiple channels, including taking advantage of compromised credentials, unaddressed weaknesses, or manipulative tactics known as social engineering. Counteracting unauthorized entry necessitates establishing robust authentication protocols such as multi-factor authentication (MFA), frequently updating systems, and training end-users in cybersecurity fundamentals.

6.2.3 Data Leaks

Data leaks, whether accidental or malicious, can lead to severe consequences, including regulatory penalties and loss of customer trust. Effective measures to prevent data leaks include data encryption, stringent access controls, and regular security audits.

6.2.4 Misconfiguration and human errors

Incorrectly configured systems can leave them open for exploitation. This includes weak passwords, open ports, and incorrect permissions among other issues. Applying security profiles such as CIS Benchmarks and custom security policies is essential to mitigate its impact.

6.3 Linux on IBM Power

IBM’s long history and strong commitment to open source is the best kept secret in open source. While open source communities have long appreciated IBM’s role in the movement, until the recent acquisition of Red Hat, not many people outside of those communities would have associated IBM with open source.

IBM was one of the earliest champions of open source, backing influential communities like Linux, Apache, and Eclipse, pushing for open licenses, open governance, and open standards. Beginning in the late 1990s, IBM supported Linux with patent pledges, a $1 billion investment of technical and other resources, and helped to establish the Linux Foundation in

2000. Since then, IBM has been consistently behind open source initiatives in general, and Linux and accompanying technologies in particular. Proof of this is IBM's support of the Linux operating system on its own hardware, including IBM Power.

IBM Power supports a variety of Linux distributions, each offering unique features and capabilities. This section provides an overview of the supported distributions and their main security features and utilities.

6.3.1 Linux Distributions on IBM Power

When we talk about Linux on Power we often stick to Red Hat Enterprise Linux and SUSE Linux Enterprise Server, when in fact there are a few more alternatives. IBM has been a long supporter of Linux across all of its hardware platforms. Linux is supported in both IBM Z—either natively or running under z.VM—and IBM Power. IBM has been supporting Linux on IBM Power systems from as early as the POWER4-based processors.

Red Hat-Based Distributions

Red Hat Enterprise Linux (RHEL) and its derivatives, CentOS, Fedora, Alma and Rocky Linux, are widely used distributions known for their stability, security, and enterprise-grade support for different use cases from classic workloads such as application servers and big databases to containers, machine learning and others.

For more information about Red Hat Enterprise Linux see this Red Hat website.

Debian-based distributions

Debian is a popular and widely-used operating system, primarily known for its stability, reliability, security and extensive software repositories. It is a Linux distribution consisting entirely of free software. Debian is the foundation for many other distributions, most notably Ubuntu also supported on Power.

Ubuntu is optimized for workloads in the mobile, social, cloud, Big Data, analytics and machine learning spaces. With its unique deployment tools (including Juju and MAAS), Ubuntu makes the management of those workloads trivial. Starting with Ubuntu 22.04 LTS, POWER9 and POWER10 processors are supported. For more information about Ubuntu Server see this website: https://ubuntu.com/server

SUSE-Based Distributions

SUSE Linux Enterprise Server (SLES) traditionally used for SAP HANA on Power environments is an alternative also for classic workloads to RHEL. In addition OpenSUSE Leap is a community-driven, open-source Linux distribution, developed by the OpenSUSE Project. It shares its core with SUSE Linux Enterprise (SLE), providing a highly stable and well-tested base. receives the same security fixes as soon as they are released to SLE customers.

SUSE Linux Enterprise Server for IBM POWER® is an enterprise-grade Linux distribution optimized for IBM POWER-based systems. It is designed to deliver increased reliability and provide a high-performance platform to meet increasing business demands and accelerate innovation while improving deployment times.

SUSE Linux Enterprise Server for POWER at a Glance:

SUSE Linux Enterprise Server for POWER was the first Linux distribution optimized for IBM Power Systems, and the first to run in POWER9 base mode.

- Increase reliability and reduce costs for mission critical applications with advanced RAS capabilities optimized to support IBM Power systems features
- Deliver a high-performance platform to meet increasing business demands with improved application performance and instant access to data
- Accelerate innovation and improve deployment times with support for POWER9 Little Endian Mode (ppc64le) for a broad choice of open source and partner solutions.

For more information about SUSE Linux Enterprise Server for IBM Power see this website: (https://www.suse.com/products/power/)

Supported Distributions

In the previous section we discussed a number of Linux distributions that are available including versions for IBM Power. Table 6-1 is a table of Linux distributions that are supported by IBM on IBM Power10 based systems. Also listed are the Ubuntu distributions where the support comes directly from Canonical.

Table 6-1 A list of supported Linux distributions on IBM Power10-based systems.

IBM Power10 processor-based systems	PowerVM LPARs
► 9043-MRX (IBM Power E1050) ► 9105-22A (IBM Power S1022) ► 9105-22B (IBM Power S1022s) ► 9105-41B (IBM Power S1014) ► 9105-42A (IBM Power S1024) ► 9786-22H (IBM Power L1022) ► 9786-42H (IBM Power L1024)	► Red Hat Enterprise Linux 9.0, any subsequent RHEL 9.x releases ► Red Hat Enterprise Linux 8.4, any subsequent RHEL 8.x releases ► SUSE Linux Enterprise Server 15 SP3, any subsequent SLES 15 updates ► Red Hat OpenShift Container Platform 4.9, or later ► Ubuntu 22.04, or later[a]
► 9080-HEX (IBM Power E1080)	► Red Hat Enterprise Linux 9.0, any subsequent RHEL 9.x releases ► Red Hat Enterprise Linux 8.4, any subsequent RHEL 8.x releases ► Red Hat Enterprise Linux 8.2 (POWER9 Compatibility mode only)[b] ► SUSE Linux Enterprise Server 15 SP3, any subsequent SLES 15 updates ► SUSE Linux Enterprise Server 12 SP5 (POWER9 Compatibility mode only) ► Red Hat OpenShift Container Platform 4.9, or later ► Ubuntu 22.04, or later[a]
► 9028-21B (IBM Power S1012)	► Red Hat Enterprise Linux 9.2, for PowerLE, or later ► Red Hat OpenShift Container Platform 4.15, or later ► Ubuntu 22.04, or later[a]

a. Ubuntu on Power support is available directly from Canonical.

b. Red Hat Business Unit approval is required for using RHEL 8.2 on IBM Power10 processor based systems.

IBM Power10 processor-based systems support the following configurations per logical partition (LPAR):

- SUSE Linux Enterprise Server 15 SP4: up to 64 TB of memory and 240 processor cores.
- SUSE Linux Enterprise Server 15 SP3: up to 32 TB of memory and 240 processor cores.
- Red Hat Enterprise Linux 8.6, or later: up to 64 TB of memory and 240 processor cores.
- Red Hat Enterprise Linux 8.4 and 9.0: up to 32 TB of memory and 240 processor cores.
- SUSE Linux Enterprise Server 12 SP5 and RHEL 8.2: up to 8 TB of memory and 120 processor cores.

The recommended Linux distribution for a particular server is always the latest level distribution that is optimized for the server. The listed distributions are the operating system versions that are supported for the specific hardware. For information about product lifecycle for Linux distributions, see the support site for each distribution.

- SUSE Linux Enterprise Server: SUSE Product Support Lifecycle (`https://www.suse.com/lifecycle/`)
- Red Hat Enterprise Linux: Red Hat Enterprise Linux Life Cycle (`https://access.redhat.com/support/policy/updates/errata/`)
- Ubuntu: Ubuntu Release Life Cycle (`https://ubuntu.com/about/release-cycle`)

For libraries and tools that can aid in leveraging the capabilities of Linux on Power10 servers, see *IBM Software Development Kit for Linux on Power* tools (`https://developer.ibm.com/linuxonpower/sdk/`). Other information about packages and migration assistance can be found in the *Find packages built for POWER* (`https://developer.ibm.com/linuxonpower/open-source-pkgs/`) in the IBM Linux on Power developer portal.

Cores is supported as a part of OpenShift Container Platform (OCP). For more information about OCP, see *Getting started with Red Hat OpenShift on IBM Cloud* (`https://cloud.ibm.com/docs/openshift?topic=openshift-getting-started`) and *Architecture and dependencies of the service* (`https://cloud.ibm.com/docs/openshift?topic=openshift-service-arch`).

6.4 Hardening Linux Systems

System hardening in Linux is an ongoing, dynamic process that involves the careful application of security principles and practices to safeguard the system against threats.

Given the complexity of Linux systems, a variety of tools and methodologies are necessary to effectively minimize the attack surface and strengthen defenses against both established and emerging threats.

Implementing security measures and utilizing available tools will vary depending on the chosen distribution and version. This guide outlines general principles without focusing on specific configurations, which may change over time.

This section covers essential aspects of hardening GNU/Linux OS on IBM Power from a distribution-neutral perspective. We will provide practical examples and guidelines using open-source software tested on ppc64le, specifically in Debian and Fedora, to ensure our Linux systems on Power are as secure as possible using an open-source first approach.

Finally, it is important to note that, despite having the source code for all applications and security tools in Linux, the methods for building the binaries are often unavailable, and dependencies—growing in number and variety—are frequently missing for them to function on ppc64le. However, this gap is narrowing as we move towards a multi-architecture world of data centers, with automated build processes and increased awareness among Linux application developers and maintainers about the potential beyond x86. When considering what tool to choose where multiple alternatives exist, choose one based on its ppc64le support.

6.4.1 Compliance

Compliance ensures that Linux deployments meet the minimum required standards in terms of configuration, patching, security, and regulatory compliance. CIS Benchmarks and DISA STIGs provide detailed guidelines for securing IT systems including Linux OS.

CIS Benchmarks are developed by the Center for Internet Security and offer best practices for securing a wide range of systems and applications, including various Linux distributions. They are community-driven and cover a broad spectrum of security configurations.

DISA STIGs, on the other hand, are developed by the Defense Information Systems Agency and are tailored to the stringent security requirements of the U.S. Department of Defense. These guides provide highly detailed security configurations and are mandatory for DoD-related systems. DISA STIGs offer comprehensive security measures that address potential threats specific to defense environments. Implementing these guidelines ensures that systems meet federal security standards and are protected against sophisticated threats.

For our purpose of providing a good basis for Linux security in Power, we will use CIS as a reference, but other standards such as PCI-DSS may be more appropriate depending on the environment.

SCAP Security Guide

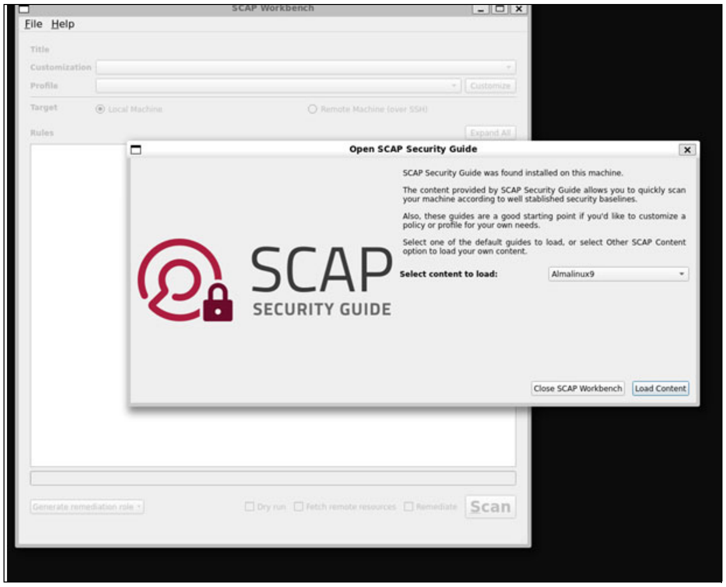
The **SCAP Security Guide (SSG)** is a comprehensive collection of security policies, baselines, and configuration guides developed by the open-source community and maintained by the National Security Agency (NSA). It provides a standardized approach to configuring and managing security settings for various operating systems and applications. The SSG has 3 key components:

1. **Security Baselines**: Predefined sets of security policies that align with various regulatory standards such as CIS, DISA STIG, PCI-DSS, and others. These baselines help organizations comply with industry-specific security requirements.
2. **SCAP Content**: Machine-readable files written in XML that describe the security policies and configurations. These files include benchmarks, rules, checks, and remediation scripts.
3. **Automation Tools**: SSG provides tools and scripts for automating the process of scanning, evaluating, and remediating security configurations. These tools can generate reports, fix scripts, and apply configurations across multiple systems.

OpenSCAP

OpenSCAP, often referred to by its command-line tool `oscap`, is an open-source framework that implements SCAP standards. It provides tools to audit and verify system configurations, vulnerabilities, and compliance against the content provided by SCAP, including the content from the SCAP Security Guide (SSG).

It can verify that the PPC64LE system adheres to various security benchmarks and standards such as CIS (Center for Internet Security) benchmarks, NIST (National Institute of Standards and Technology) guidelines, custom security policies or vulnerability lists. It also has a GUI, **`scap-workbench`**, available at least on RHEL-based distributions on Power such as Alma Linux 9 which is shown in Figure 6-1 on page 158.



Open SCAP Security Guide

SCAP Security Guide was found installed on this machine.

The content provided by SCAP Security Guide allows you to quickly scan your machine according to well stablished security baselines.

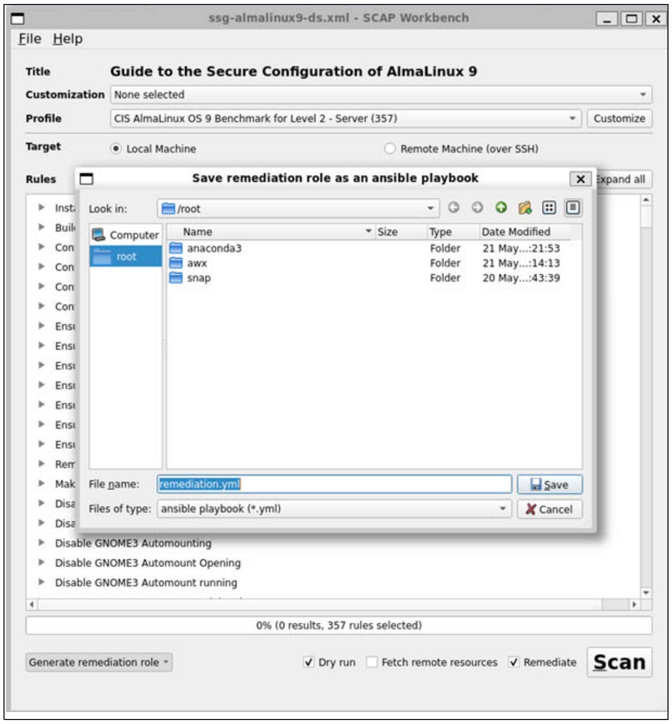
Also, these guides are a good starting point if you'd like to customize a policy or profile for your own needs.

Select one of the default guides to load, or select Other SCAP Content option to load your own content.

Select content to load: Almalinux9

Figure 6-1 SCAP workbench GUI

Figure 6-2 shows one of the management screens in the GUI.



Title Guide to the Secure Configuration of AlmaLinux 9

Customization None selected

Profile CIS AlmaLinux OS 9 Benchmark for Level 2 - Server (357)

Target ● Local Machine ○ Remote Machine (over SSH)

Save remediation role as an ansible playbook

Look in: /root

Name	Size	Type	Date Modified
anaconda3		Folder	21 May...:21:53
awx		Folder	21 May...:14:13
snap		Folder	20 May...:43:39

File name: remediation.yml

Files of type: ansible playbook (*.yml)

Computer
root
Save
Cancel
Inst
Buil
Con
Con
Con
Con
Ensu
Ensu
Ensu
Ensu
Ensu
Ensu
Ensu
Rem
Mak
Disa
Disa
Disable GNOME3 Automounting
Disable GNOME3 Automount Opening
Disable GNOME3 Automount running

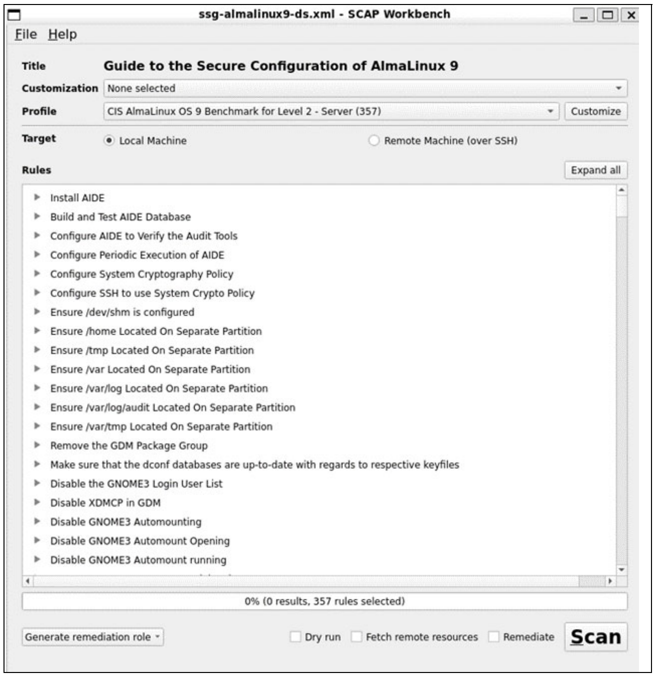
0% (0 results, 357 rules selected)

Generate remediation role ☑ Dry run □ Fetch remote resources ☑ Remediate

Figure 6-2 Management screen in GUI

To comply with security regulations and policies, we take the following approach:

1. Install the Linux ISO of your choice
2. Decide which set of rules should use (always start with a dry run). Figure 6-3 on page 159 shows using CIS Level 2 benchmark using scap-workbench (GUI)


ssg-almalinux9-ds.xml - SCAP Workbench
File
Help
Customize
Expand all
Generate remediation role
Scan

Title Guide to the Secure Configuration of AlmaLinux 9

Customization None selected

Profile CIS AlmaLinux OS 9 Benchmark for Level 2 - Server (357)

Target ● Local Machine ○ Remote Machine (over SSH)

Rules

- Install AIDE
- Build and Test AIDE Database
- Configure AIDE to Verify the Audit Tools
- Configure Periodic Execution of AIDE
- Configure System Cryptography Policy
- Configure SSH to use System Crypto Policy
- Ensure /dev/shm is configured
- Ensure /home Located On Separate Partition
- Ensure /tmp Located On Separate Partition
- Ensure /var Located On Separate Partition
- Ensure /var/log Located On Separate Partition
- Ensure /var/log/audit Located On Separate Partition
- Ensure /var/tmp Located On Separate Partition
- Remove the GDM Package Group
- Make sure that the dconf databases are up-to-date with regards to respective keyfiles
- Disable the GNOME3 Login User List
- Disable XDMCP in GDM
- Disable GNOME3 Automounting
- Disable GNOME3 Automount Opening
- Disable GNOME3 Automount running

0% (0 results, 357 rules selected)

☐ Dry run ☐ Fetch remote resources ☐ Remediate

Figure 6-3 Choosing benchmark via GUI

3. Automatically address these compliance gaps when technically feasible, with Bash scripts and Ansible playbooks, as shown in the screen shot or via the command line as shown below:

   ```
   oscap xccdf generate fix --profile [PROFILE_ID] --output remediation_script.sh
   \  usr/share/xml/scap/ssg/content/ssg-[OS].xml
   ```

It is crucial to be aware that automated remediation may yield unexpected results on systems that have already been modified. Therefore, administrators are strongly advised to thoroughly evaluate the potential impact of remediation actions on their specific systems. You might want to make a snapshot / backup before moving on.

> **Tip:** Under normal conditions the remediation of compliance issues will be the result of several iterations and some backtracking by recovering snapshots or backups until we reach a level of security adequate for our purposes. Always in balance with the usability of the system.

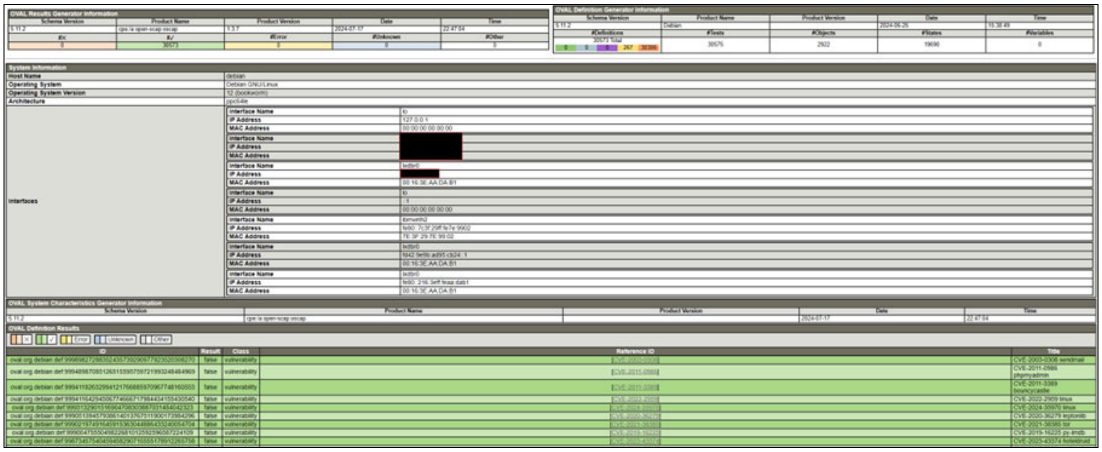
OpenSCAP is not only about compliance. It can also help you to check if there is any vulnerability in our current OS version using **OVAL** (Open Vulnerability and Assessment Language) and generating a report.

Example 6-1 shows how we can do it using Debian.

Example 6-1 Generating vulnerability report

```
wget https://www.debian.org/security/oval/oval-definitions-$(lsb_release -cs).xml.bz2
bzip2recover oval-definitions-$(lsb_release -cs).xml.bz2
oscap oval eval --report report.html oval-definitions-$(lsb_release -cs).xml
```

Figure 6-4 shows the generated HTML report with no vulnerabilities found.



OVAL Results Generator Information				
Schema Version	Product Name	Product Version	Date	Time
5.11.2	cpe:/a:open-scap:oscap	1.3.7	2024-07-17	22:47:04
#✕	#✓	#Error	#Unknown	#Other
0	30573	0	0	0

OVAL Definition Generator Information				
Schema Version	Product Name	Product Version	Date	Time
5.11.2	Debian		2024-06-25	15:38:49
#Definitions	#Tests	#Objects	#States	#Variables
30573 Total 0 0 0 267 30306	30675	2922	19690	0

System Information		
Host Name	debian	
Operating System	Debian GNU/Linux	
Operating System Version	12 (bookworm)	
Architecture	ppc64le	
Interfaces	Interface Name	lo
	IP Address	127.0.0.1
	MAC Address	00:00:00:00:00:00
	Interface Name	[redacted]
	IP Address	[redacted]
	MAC Address	[redacted]
	Interface Name	lxdbr0
	IP Address	[redacted]
	MAC Address	00:16:3E:AA:DA:B1
	Interface Name	lo
	IP Address	::1
	MAC Address	00:00:00:00:00:00
	Interface Name	ibmveth2
	IP Address	fe80::7c3f:29ff:fe7e:9902
	MAC Address	7E:3F:29:7E:99:02
	Interface Name	lxdbr0
	IP Address	fd42:9e9b:ad95:cb24::1
	MAC Address	00:16:3E:AA:DA:B1
	Interface Name	lxdbr0
	IP Address	fe80::216:3eff:feaa:dab1
	MAC Address	00:16:3E:AA:DA:B1

OVAL System Characteristics Generator Information				
Schema Version	Product Name	Product Version	Date	Time
5.11.2	cpe:/a:open-scap:oscap		2024-07-17	22:47:04

OVAL Definition Results

✕ ✓ Error Unknown Other

ID	Result	Class	Reference ID	Title
oval:org.debian:def:99969827286352435739290977923520308270	false	vulnerability	[CVE-2003-0308]	CVE-2003-0308 sendmail
oval:org.debian:def:99948987085126815595759721993248484969	false	vulnerability	[CVE-2011-0986]	CVE-2011-0986 phpmyadmin
oval:org.debian:def:99941182632994121766885970967748160553	false	vulnerability	[CVE-2011-3389]	CVE-2011-3389 bouncycastle
oval:org.debian:def:99941164294506774666717984434155430540	false	vulnerability	[CVE-2022-2909]	CVE-2022-2909 linux
oval:org.debian:def:99901329015169047083038870314540423 23	false	vulnerability	[CVE-2024-35970]	CVE-2024-35970 linux
oval:org.debian:def:99905139457938614013767511900173984296	false	vulnerability	[CVE-2020-36279]	CVE-2020-36279 leptonlib
oval:org.debian:def:99902197491640915363044886433240054704	false	vulnerability	[CVE-2021-38385]	CVE-2021-38385 tor
oval:org.debian:def:99900475550498226810125925960587224109	false	vulnerability	[CVE-2019-16225]	CVE-2019-16225 py-imdb
oval:org.debian:def:99673457540459456290715555178912263758	false	vulnerability	[CVE-2023-43374]	CVE-2023-43374 hoteldruid

Figure 6-4   Vulnerability report

Tip: If you are satisfied with the image you have just evaluated and you use PowerVC (IBM ‘s Power virtualization solution based in OpenStack) it is a good time to capture this system as a template or even create an OVA. Be careful when using “default” installations as they may be missing important security protection settings. Always utilize appropriate compliance policies to ensure that your Linux systems running on IBM Power are all well configured and protected.

To summarize, while compliance signifies that organizations adhere to minimal required standards, it doesn't automatically imply full security. A firm might fulfill every criterion for a PCI DSS assessment without ensuring effective employee training, resulting in inadequate execution and a higher likelihood of security breaches.

6.4.2 Network Security

Network security measures protect Linux systems from external and internal threats. This includes implementing intrusion detection and prevention systems and encrypting data in transit.

Firewall Technologies

Firewalls are a critical component of network security, essential for controlling the flow of incoming and outgoing traffic based on predefined security rules. Effective firewall management on Linux systems involves various tools, each offering different levels of control, efficiency, and ease of use. This section explores the primary tools used in Linux firewall implementations, their relationships, and practical guidance on their use.

Linux firewalls have evolved significantly over time, starting from simple packet filtering mechanisms to more sophisticated and user-friendly management tools. The primary tools used in Linux firewall implementations include iptables, nftables, firewalld, and UFW (Uncomplicated Firewall). Understanding the background and functionality of these tools helps in choosing the right one for your specific needs (including the distribution you chose)

Netfilter is a framework within the Linux kernel that provides various networking-related operations such as packet filtering, network address translation (NAT), and packet mangling. It is the core infrastructure that enables these operations, with hooks in the kernel where modules can register callback functions to handle network packets. Both *iptables* and *nftables* are user-space utilities that interact with the *netfilter* framework,

iptables is a command-line utility that allows administrators to configure the Linux kernel firewall. It enables granular control over packet filtering and manipulation, offering precise management of data flow and security policies. Despite its powerful features, *iptables* can be complex to configure due to its detailed syntax and structure.

This command allows the use of SSH:

```
sudo iptables -A INPUT -p tcp --dport 22 -j ACCEPT
```

nftables is the successor to *iptables*, designed to provide a more efficient and streamlined framework for packet filtering and Network Address Translation (NAT). Introduced in the Linux kernel since version 3.13, *nftables* offers a simplified syntax and enhanced performance. It has been gradually adopted by many distributions as the default backend for firewall configurations, aiming to overcome part of the complexity and performance limitations of iptables. The command to permit SSH in nftables is:

```
sudo nft add rule inet filter input tcp dport 22 accept
```

bpfilter is a newer and primary kernel-space framework designed to improve firewall management in Linux by leveraging eBPF (Extended Berkeley Packet Filter). *bpfilter* aims to replace the older *iptables* and *nftables* with a more flexible and efficient packet filtering mechanism, but it is still under heavy development and has not yet seen widespread adoption.

Firewall tools

Within the front end tools for creating and maintaining firewall rules in Linux (using iptables or nftables) we have two main options:

firewalld is a dynamic firewall management tool included in Red Hat Enterprise Linux (RHEL) since version 7. It simplifies firewall management by using the concept of network zones, which define the trust level of network connections and interfaces. Firewalld allows for real-time changes without needing to restart the firewall, providing a flexible and dynamic approach compared to traditional static tools like iptables.Firewalld uses nftables as its backend by default on modern systems and firewall-cmd as the command line tool.

```
sudo firewall-cmd --zone=public --add-port=22/tcp --permanent
```

UFW (Uncomplicated Firewall) is designed to provide an easy-to-use interface for managing firewall settings. With its straightforward command-line interface, UFW allows users to implement basic firewall rules with minimal effort. It is especially useful for users who may not have extensive networking or firewall management experience but still need to ensure system security. UFW is included by default in Ubuntu and can be installed on Debian systems. A command for allowing SSH using *UFW* would be:

```
sudo ufw allow 22/tcp
```

Recommendations for Linux Firewalls

The Center for Internet Security (CIS) advises setting a default deny policy for both incoming and outgoing traffic, ensuring that only explicitly allowed traffic is permitted. This involves allowing essential services such as SSH from trusted networks, and loop-back traffic, while restricting other services to mitigate unauthorized access.

Regular reviews and updates of firewall rules are also recommended to maintain compliance with security policies and adapt to emerging threats. These measures collectively aim to fortify Linux systems against a variety of network-based threats.

A best practice would be to forward Linux firewall logs to a solution that automates their analysis, alerts, and, when possible, the responses as well. For an example using IBM QRadar see this document.

In Example 6-2 we show a simple firewall configuration on Linux on Power using **firewall-cmd**.

Example 6-2 Configure firewall with firewall-cmd

```
# Install firewalld (if not already installed)
sudo dnf install firewalld -y

# Start and enable firewalld
sudo systemctl start firewalld
sudo systemctl enable firewalld

# Allow SSH traffic
sudo firewall-cmd --permanent --add-service=ssh

# Set default deny policies
# Firewalld uses zones to manage rules. The default zone is "public".
# By default, firewalld allows all outgoing traffic. To mimic iptables behavior, you can configure it 
to deny by default.
sudo firewall-cmd --permanent --set-target=DROP

# Allow loopback traffic
sudo firewall-cmd --permanent --add-interface=lo --zone=trusted
sudo firewall-cmd --permanent --zone=trusted --add-source=127.0.0.1

# Enable logging (optional)
sudo firewall-cmd --set-log-denied=all

# Allow specific outgoing traffic (optional)
sudo firewall-cmd --permanent --add-port=80/tcp
sudo firewall-cmd --permanent --add-port=443/tcp

# Reload firewall to apply changes
sudo firewall-cmd --reload

# Review firewalld rules
sudo firewall-cmd --list-all
```

Example 6-3 shows this same configuration using UFW.

Example 6-3 Configuring firewall with UFW

```
# Install UFW
sudo apt-get install ufw

# Allow SSH traffic
sudo ufw allow ssh

# Set default deny policies
sudo ufw default deny incoming
sudo ufw default deny outgoing

# Enable UFW
sudo ufw enable

# Allow specific outgoing traffic (optional)
sudo ufw allow out to any port 80
sudo ufw allow out to any port 443

# Review UFW status
sudo ufw status verbose
```

Tip: Take care not to lock yourself out. Although new rules will not apply to existing connections, make sure you either have a script to disable the firewall automatically after a few seconds or direct console access in case of emergency.

Additionally, CIS emphasizes the importance of logging and auditing firewall activity to detect and respond to suspicious behavior, and suggests using stateful inspection and rate limiting to prevent attacks like Denial of Service.

Intrusion Detection and Prevention

Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) monitor network traffic for suspicious activities. Tools such as *Snort, OSSEC, Zeek* and *Suricata* can be deployed to detect and prevent potential threats.

We will be using *Suricata* because it has strong support for ppc64le architecture. *Suricata* is a versatile and high-performance Network Security Monitoring (NSM) tool capable of detecting and blocking network attacks. By default, *Suricata* operates as a passive Intrusion Detection System (IDS), scanning for suspicious traffic on a server or network and generating logs and alerts for further analysis. Additionally, it can be configured as an active Intrusion Prevention System (IPS) to log, alert, and completely block network traffic that matches specific rules. *Suricata* is open source and managed by the community-run non-profit organization, the Open Information Security Foundation.

Example 6-4 shows setting up *Suricata*

Example 6-4 Intrusion detection using suricata

```
root@debian:~/snort3# systemctl status suricata

suricata.service - Suricata IDS/IDP daemon
Loaded: loaded (/lib/systemd/system/suricata.service; enabled; preset: enabled)
Active: active (running) since Wed 2024-07-17 19:07:00 BST; 2s ago
Docs: man:suricata(8)
      man:suricatasc(8)
      https://suricata-ids.org/docs/
    Process: 2965553 ExecStart=/usr/bin/suricata -D --af-packet -c /etc/suricata/suricata.yaml
--pidfile /run/suricata.pid (code=exited, status=0/SUCCESS)
   Main PID: 2965554 (Suricata-Main)
      Tasks: 38 (limit: 9635)
     Memory: 255.8M
        CPU: 335ms
     CGroup: /system.slice/suricata.service
             ··2965554 /usr/bin/suricata -D --af-packet -c /etc/suricata/suricata.yaml --pidfile
/run/suricata.pid

##
## Step 3: Configure common capture settings
##
## See "Advanced Capture Options" below for more options, including Netmap
## and PF_RING.
##

# Linux high speed capture support
af-packet:
  - interface: ibmveth2 <- modify this
    # Number of receive threads. "auto" uses the number of cores
    #threads: auto
```

Suricata includes a tool called **`suricata-update`** that can fetch rule sets from external providers. Example 6-5 shows how to download the latest rule set for your *Suricata* server.

Example 6-5 Installing suricata-update

```
root@debian:~# sudo suricata-update -o /etc/suricata/rules
20/7/2024 -- 21:38:39 - <Info> -- Using data-directory /var/lib/suricata.
20/7/2024 -- 21:38:39 - <Info> -- Using Suricata configuration /etc/suricata/suricata.yaml
20/7/2024 -- 21:38:39 - <Info> -- Using /etc/suricata/rules for Suricata provided rules.
20/7/2024 -- 21:38:39 - <Info> -- Found Suricata version 6.0.10 at /usr/bin/suricata.
20/7/2024 -- 21:38:39 - <Info> -- Loading /etc/suricata/suricata.yaml
20/7/2024 -- 21:38:39 - <Info> -- Disabling rules for protocol http2
20/7/2024 -- 21:38:39 - <Info> -- Disabling rules for protocol modbus
20/7/2024 -- 21:38:39 - <Info> -- Disabling rules for protocol dnp3
20/7/2024 -- 21:38:39 - <Info> -- Disabling rules for protocol enip
20/7/2024 -- 21:38:39 - <Info> -- No sources configured, will use Emerging Threats Open
20/7/2024 -- 21:38:39 - <Info> -- Fetching
https://rules.emergingthreats.net/open/suricata-6.0.10/emerging.rules.tar.gz.
20/7/2024 -- 21:40:43 - <Info> -- Loaded 367 rules.
20/7/2024 -- 21:40:43 - <Info> -- Disabled 14 rules.
20/7/2024 -- 21:40:43 - <Info> -- Enabled 0 rules.
20/7/2024 -- 21:40:43 - <Info> -- Modified 0 rules.
20/7/2024 -- 21:40:43 - <Info> -- Dropped 0 rules.
20/7/2024 -- 21:40:43 - <Info> -- Enabled 0 rules for flowbit dependencies.
20/7/2024 -- 21:40:43 - <Info> -- Backing up current rules.
20/7/2024 -- 21:40:43 - <Info> -- Writing rules to /etc/suricata/rules/suricata.rules: total: 367;
enabled: 311; added: 367; removed 0; modified: 0
20/7/2024 -- 21:40:43 - <Info> -- Writing /etc/suricata/rules/classification.config
20/7/2024 -- 21:40:43 - <Info> -- Testing with suricata -T.
20/7/2024 -- 21:40:43 - <Info> -- Done.
```

For additional information on Suricata, including installation instructions, see the Suricata documentation.

Encryption in Flight

Encrypting data in transit protects it from being intercepted and read by unauthorized parties. Protocols such as SSL/TLS are used to secure communications over networks.

- **SSL/TLS** are Secure protocols for encrypting web traffic, email, and other communications.

  To secure your web server with SSL/TLS, you first need to obtain a digital certificate. Certbot is an automated tool designed to streamline the process of acquiring and installing SSL/TLS certificates. It is one of many technology projects developed by the Electronic Frontier Foundation (EFF) to promote online freedom.

  Certbot is available in different Linux repositories including ppc64le versions, making installation straightforward, It has plug-ins for both apache and nginx among other typical deployments and includes a tool to automatically renew these certificates.

  Example 6-6 shows how to install Certbot in a Debian Linux system. Other Linux versions might differ slightly.

Example 6-6 Installing Certbot

```
sudo apt-get install certbot python3-certbot-apache
sudo apt-get install certbot python3-certbot-nginx
```

To obtain and automatically install the certificate for your web server run these commands.

`sudo certbot –apache`

or

`sudo certbot --nginx`

Certbot will prompt you to enter your email address and agree to the terms of service. It will then interact with your web server to perform the domain verification process and install the SSL/TLS certificate.

Check their website for more information and how-to guides.

- Virtual Private Networks (VPNs) create encrypted tunnels to provide secure remote access. Setting up a VPN is essential for protecting communications and ensuring the privacy of transmitted data. VPNs are highly recommended by the Center for Internet Security (CIS) as a best practice for securing remote connections.

  To install and configure OpenVPN on Linux on Power, you can follow the specific guides for different Linux on Power distributions found at these links:

  - `https://wiki.debian.org/OpenVPN`—Debian based
  - `https://fedoraproject.org/wiki/OpenVPN`—RHEL based
  - `https://documentation.suse.com/sles/15-SP5/html/SLES-all/cha-security-vpnserver.html`–SUSE based

Encryption at rest

Encryption at Rest is a form of encryption that is designed to prevent an attacker from accessing data by ensuring it is encrypted when stored on a persistent device. This can be done at different layers, from physical storage systems to the OS. If you choose to encrypt at the OS level, it is best to employ full disk encryption using LUKS with LVM (Debian / RHEL-Based) or BTRFS (SUSE).

Linux Unified Key Setup-on-disk-format (LUKS) offers a suite of tools designed to simplify the management of encrypted devices. LUKS allows you to encrypt block devices and supports multiple user keys that can decrypt a master key. This master key is used for the bulk encryption of the partition.

You can configure disk encryption at the installation time or later using **`cryptsetup`**, a command-line tool used to conveniently set up disk encryption based on the *dm-crypt* kernel module. It offers a range of functionalities, including creating, opening, and managing encrypted volumes.

Prerequisites for installing LUKS are:

- – A Linux system with LVM installed.
- – cryptsetup and LUKS packages installed.
- – Root or sudo privileges

Example 6-7 provides an example of activating encryption at rest.

Example 6-7 Setting up encryption at rest

```
sudo apt-get install lvm2 cryptsetup

Initialize the physical volume
sudo pvcreate /dev/sdX

Create a volume grouo
sudo vgcreate acme_vg /dev/sdX
```

```
Create a logical volume
sudo lvcreate -n acme_lv01 -L 10G acme_vg

Encrypt the new LV
sudo cryptsetup luksFormat /dev/acme_vg/acme_lv01

WARNING!
========
This will overwrite data on /dev/sdX irrevocably.

Are you sure? (Type uppercase yes): YES
Enter LUKS passphrase:
Verify passphrase:
Command successful.
Create a mapping
sudo cryptsetup open /dev/acme_vg/acme_lv01 encrypted_lv01

Format the encrypted logical volume
sudo mkfs.ext4 /dev/mapper/encrypted_lv01

Mount it (and use it)
sudo mount /dev/mapper/encrypted_lv01 /media
```

Automating Encryption at Boot

To ensure the encrypted LVM is available at boot, you need to configure /etc/crypttab and /etc/fstab.

To configure /etc/crypttab add the following line:

```
encrypted_lv01 /dev/acme_vg/acme_lv01 none luks
```

To configure /etc/fstab add the following line:

```
/dev/mapper/encrypted_lv01 /mnt ext4 defaults 0 2
```

For more information see this link: https://gitlab.com/cryptsetup/cryptsetup.

6.4.3 User policies and access controls

Proper user policies and administration are vital for maintaining a secure environment. This includes defining and enforcing password policies and managing user access. The Center for Internet Security provides some recommendations that we can apply on Linux on Power.

Linux utilizes Pluggable Authentication Modules (PAM) in the authentication process, serving as an intermediary layer between users and applications. PAM modules are accessible on a system-wide basis, allowing any application to request their services. The PAM modules implement most of the user security measures that are defined in various files within the /etc directory, including LDAP, Kerberos and Active Directory connections or MFA options.

Access control mechanisms ensure that only authorized users can access specific resources. This includes configuring SUDO, managing user groups, and maintaining access logs.

Password Policies

Enforcing strong password policies is crucial to prevent unauthorized access. Policies should mandate complex passwords, regular password changes, and account lockout mechanisms after multiple failed login attempts.

A password policy can specify the minimum length a password must have and the maximum duration it can be used before needing to be changed. All users under this policy must create passwords that are long enough and update them regularly. Implementing password policies helps mitigate the risk of passwords being discovered and misused.

A minimum password policy for Linux on Power should contain at least:

- **Complex passwords:** Require a mix of uppercase, lowercase, numbers, and special characters. This can be enforced with the **pam_pwquality** PAM module. CIS recommends that passwords should be at least 14 characters long with no limit on the enforced maximum number of characters among other requirements.

Example 6-8 is an excerpt of a sample configuration of **/etc/security/pwquality.conf** (Ubuntu.

Example 6-8 Example security configuration

```
# Minimum acceptable size for the new password (plus one if
# credits are not disabled which is the default). (See pam_cracklib manual. Cannot be set
to lower value than 6.
minlen = 15

# The maximum credit for having digits in the new password. If less than 0 it is the
minimum number of digits in the new password.
dcredit = -1

# The maximum credit for having uppercase characters in the new password. # If less than 0
it is the minimum number of uppercase characters in the new
# password.
ucredit = -1

..
```

- **Regular changes**: CIS recommends specific password change policies for Linux systems to enhance security. These include setting a maximum password age of 90 days or less to ensure regular password updates, a minimum password age of 7 days to prevent rapid password changes that could cycle back to previous passwords, and a password expiration warning of 7 days to notify users in advance of impending password expiry. These guidelines help maintain robust security by ensuring that passwords are regularly updated and users are adequately informed.

  Enforce password expiration policies are defined in **login.defs**. Example 6-9 shows an excerpt of a sample configuration of **/etc/login.defs** (Ubuntu) to log both successful logins and su activity.

Example 6-9 Sample login definition

```
#
# Enable logging of successful logins
#
LOG_OK_LOGINSyes


#
# If defined, all su activity is logged to this file.
#
SULOG_FILE/var/log/sulog
```

- **Account lockout**: Typically, it is advised by CIS to lock the account after five unsuccessful attempts and to unlock it automatically after a period, such as 15 minutes. These settings help mitigate the risk of unauthorized access by deterring repeated login attempts and ensuring that legitimate users can regain access after a brief lockout period.

  This is configured in the /etc/pam.d/common-auth file (Debian/Ubuntu/SUSE) or /etc/pam.d/system-auth file (RHEL-based). The pam_faillock module performs a function similar to the legacy pam_tally and pam_tally2 but with more options and flexibility. Check which is the recommended method in your chosen distribution and current version.

  This is an excerpt of a sample legacy configuration (SUSE):

  ```
  auth   required  pam_tally2.so  onerr=fail deny=3 unlock_time=1800
  ```

Groups

Grouping users based on their roles and responsibilities helps in managing permissions efficiently. Assigning users to appropriate groups ensures they have access only to the necessary resources.

For example, a file with permissions *rw-rw---- (660)* allows the owner and the group to read and write the file, but others cannot access it. This reduces the risk of accidental or malicious modifications to sensitive files.

In this way developers can be part of a *dev* group with access to development files, while the production team is part of a *prod* group with access to production files.

CIS advises regular audits of group memberships to ensure that users have appropriate permissions and to remove any unnecessary or outdated group assignments. Additionally, the creation of custom groups for specific tasks or roles is recommended to further refine access control and minimize potential security risks.

Access Control Lists

Access Control Lists (ACLs) provide more fine-grained control over permissions for files and directories than user and group permissions. Example 6-10 shows displaying the ACL for a file.

Example 6-10 Viewing ACL

```
touch testfile
ls -l testfile
-rw-r--r-- 1 user user 0 Jul 17 14:05 testfile
```

Example 6-11 shows the process to grant read and write permissions to another user (e.g., john). After setting the ACL for the user the **getfacl** command can be used to display the ACL as shown in the example.

Example 6-11 Setting ACL for a specific user

```
sudo setfacl -m u:john:rw testfile
getfacl testfile
# file: testfile
# owner: user
# group: user
user::rw-
user:john:rw-
group::r--
mask::rw-
other::r--
```

Multi-Factor Authentication

Multi-Factor Authentication (MFA) is a security process that requires users to verify their identity through multiple methods before gaining access to a system or application. Unlike single-factor authentication (e.g., a password), MFA enhances security by combining two or more independent credentials from the following categories:

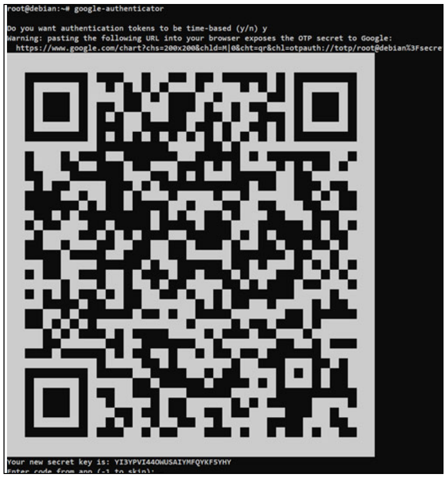
- Something you know: Typically a password or PIN.
- Something you have: A physical device such as a smart phone
- Something you are: Biometric verification like fingerprints, facial recognition, or iris scans.

When attempting to login to a system secured by multi-factor authentication (MFA), users are required to supply extra credentials beyond their standard username and password. In the context of Linux systems, Secure Shell (SSH) serves as a common method for remotely accessing the system. To enhance security further, it's advisable to incorporate MFA when SSH is used.

One method of implementing MFA is the use of IBM PowerSC. However, MFA can also be implemented using native tools like Google authenticator. This can be done using:

- **`google-authenthicator libpam-google-authenticator`** for Debian based systems
- **`google-authenticator-libpam`** in SUSE based systems
- **`google-authenticator`** in Extra Packages for Enterprise Linux (EPEL) for Red Hat Enterprise Linux based systems

Google authenticator has a setup script for configuration that works out of the box. This uses the Google Authenticator app available for Android and iOS to generate authentication codes. The authentication code is shown in Figure 6-5.



```
root@debian:~# google-authenticator

Do you want authentication tokens to be time-based (y/n) y
Warning: pasting the following URL into your browser exposes the OTP secret to Google:
  https://www.google.com/chart?chs=200x200&chld=M|0&cht=qr&chl=otpauth://totp/root@debian%3Fsecre

Your new secret key is: YI3YPVI44OWUSAIYMFQYKF5YHY
Enter code from app (-1 to skip):
```

Figure 6-5 Google authenticator screen

For more information on adding MFA to other distribution, see the following links:

https://ubuntu.com/tutorials/configure-ssh-2fa#1-overview
https://fedoramagazine.org/two-factor-authentication-ssh-fedora/

Role Based Access Control

While organizations have various user provisioning methodologies to choose from, Role-Based Access Control (RBAC) is among the most prevalent. RBAC is a method of restricting system access to authorized users based on their roles within an organization. In RBAC, permissions are not assigned directly to users but to roles, and users are assigned to these roles. This abstraction simplifies the management of permissions.

RBAC offers a more detailed approach to identity and access management (IAM) compared to access control lists (ACLs), yet it is simpler and more straightforward to implement than attribute-based access control (ABAC). Other IAM methods, such as mandatory access control (MAC) or discretionary access control (DAC), can be effective for particular scenarios.

The following list provides some methods to help assist with setting appropriate access controls within your system:

- **SELinux** (RHEL / SUSE based) allows for the definition of roles and the assignment of domains (or types) to these roles. Users are then assigned roles, and the roles define the allowable operations on objects within the system which makes it a RBAC-like solution. SELinux utilizes security policies that are label-based, identifying applications through their file system labels. SELinux might be complex to configure and manage.For additional information see this link: https://github.com/SELinuxProject.
- **AppArmor** (Debian-based) employs security profiles that are path-based, identifying applications by their executable paths. This means it does not have a traditional RBAC approach but allows defining profiles for applications, which can be seen as a form of access control. For more information see: https://apparmor.net
- **FreeIPA** is an integrated security information management solution combining Fedora Linux, 389 Directory Server, Kerberos, NTP, DNS, and Dogtag (Certificate System). It provides centralized identity management and includes support for RBAC, allowing administrators to define roles and associate permissions and policies with these roles across a network of Linux systems. For more information see: https://www.freeipa.org/
- **RHEL System Roles** is a collection of Ansible roles and modules that provide a stable and consistent configuration interface to automate and manage multiple releases of Red Hat Enterprise Linux. The RHEL System Roles are supported as provided from the following methods:
  - As an RPM package in the RHEL 9 or RHEL 8 Application Streams repositories
  - As a supported collection in the Red Hat Automation Hub

  More information see: https://access.redhat.com/articles/3050101

Each solution offers different features and complexities, allowing administrators to choose the most appropriate tool based on their specific security requirements and environment. Red Hat based distributions come preconfigured with a lot of SELinux policies but the configuration might be more complex than FreeIPA or AppArmour. Using RHEL roles will typically be part any of automation policies.

SUDO

The widespread reliance on sudo in most Linux distributions over other choices is attributed to its ease of use and the granular control it provides over user permissions. Sudo simplifies the delegation of limited root access, specifies allowed commands through the sudoers file, and maintains an audit trail, which makes it highly practical for routine administrative tasks. Other tools, while powerful, involve complex management and a level of detail that is typically unnecessary for everyday operations, making them more suitable for specialized use cases.

The fundamental approach to configuring sudo is consistent across Debian-based, Red Hat-based, and SUSE-based distributions, the specifics of default configurations and group usage vary.

These variations are:

- **Debian-based**: Focus on sudo group, root user disabled by default (Ubuntu).
- **Red Hat-based**: Use of wheel group, root user enabled by default.
- **SUSE-based**: Flexible with sudo or users groups, root user enabled by default.

That said, to implement RBAC using sudo, we can follow these steps.

1. Determine the different roles in the organization and the specific permissions or commands each role needs.
2. Create Unix groups corresponding to each role. For example, admin, developer, auditor, etc. Example 6-12 shows adding groups.

Example 6-12 Create groups for SUDO

```
    sudo groupadd admin
    sudo groupadd developer
    sudo groupadd auditor
```

Add users to the appropriate groups based on their roles as seen in Example 6-13.

Example 6-13 Add users to sudo

```
sudo usermod -aG admin alice
    sudo usermod -aG developer bob
    sudo usermod -aG auditor charlie
```

3. Edit the *sudoers* file to grant permissions to groups. This is done using the **`visudo`** command to ensure proper syntax and prevent mistakes.

   `sudo visudo`

   In the sudoers file, define the commands that each group can execute. Example 6-14 shows group permissions.

Example 6-14 Group permissions

```
     %admin ALL=(ALL) ALL
     # Admins can execute any command
     %developer ALL=(ALL) /usr/bin/git, /usr/bin/make, /usr/bin/gcc
     # Developers can run git, make, gcc
    %auditor ALL=(ALL) /usr/bin/less, /bin/cat, /usr/bin/tail            # Auditors can run less, cat, tail
```

4. Users can now use the sudo command to execute commands based on their roles as shown in Example 6-15.

Example 6-15 Role definitions in sudo

```
    sudo git commit -m "example commit"   # For a developer
     sudo less /var/log/syslog          # For an auditor
     sudo systemctl restart apache2       # For an admin
```

In this example, admin role has full control over the system, the developer role grants access to development tools like git, make, gcc and the auditor role has read-only access to logs and configuration files. You can learn more about sudo at `https://www.sudo.ws/.`

6.4.4 Logging, audits and file integrity monitoring

Access logging provides a record of user activities, which is crucial for auditing and identifying suspicious behavior. They offer insights into system activities and ensuring compliance with security policies. We can leverage several tools. These commands and typical use cases are:

rsyslog: Centralize logs using syslog to monitor and analyze security events.

auditd: A powerful auditing tool that logs system calls and user activities.

ausearch: A tool that can query the audit daemon logs based for events based on different search criteria.

We will show how to deploy and combine all these tools in a practical example.

Syslog is a standard for message logging that allows separation of the software that generates messages from the system that stores them and the software that reports and analyzes them. `Rsyslog` is an enhanced version of `syslog`, It builds upon the foundation of `syslog`, providing advanced features and greater flexibility.

Install rsyslog (Debian-based) using this command:

```
sudo apt-get install rsyslog
```

Edit `/etc/rsyslog.conf` to configure log levels and destinations as shown in Example 6-16.

Example 6-16 Configuring log levels

```
# Log all user messages to /var/log/user.log user.
* /var/log/user.log
# Log all auth messages to /var/log/auth.log auth.
* /var/log/auth.log
```

Restart syslog service using this command:

```
sudo systemctl restart rsyslog
```

Auditd is the userspace component of the Linux Auditing System, which is used to collect, filter, and store audit records generated by the kernel. These records can include information about system calls, file accesses, user logins, and other significant security events. The audit daemon (auditd) is responsible for writing these records to disk and managing the log files. Install `auditd` using this command:

```
sudo apt-get install auditd audispd-plugin
```

Next, edit the audit.rules file using this command:

```
vi /etc/audit/rules.d/audit.rules
```

Append the lines shown in Example 6-17.

Example 6-17 Lines to append to audit.rules

```
# Monitor changes to /etc/passwd
-w /etc/passwd -p wa -k passwd_changes

# Monitor changes to /etc/shadow
-w /etc/shadow -p wa -k shadow_changes

# Monitor use of privileged commands
-a always,exit -F arch=b64 -S execve -C uid!=euid -F euid=0 -k privileged
```

Restart auditd to make the changes take effect.

```
sudo systemctl restart auditd
```

To validate that the changes took effect change one password.

```
#passwd hugo
```

Now search the audit log for password changes.

```
sudo ausearch -k passwd_changes
```

The results are shown in Example 6-18.

Example 6-18 Audit display of password changes

```
----
time->Wed Jul 17 18:35:16 2024
type=PROCTITLE msg=audit(1721237716.900:98064):
proctitle=2F7362696E2F617564697463746C002D52002F6574632F61756469742F61756469742E72756C6573
type=SYSCALL msg=audit(1721237716.900:98064): arch=c0000015 syscall=335 success=yes
exit=1084 a0=3 a1=7fffd8d50114 a2=43c a3=0 items=0 ppid=2962757 pid=2962773 auid=4294967295
uid=0 gid=0 euid=0 suid=0 fsuid=0 egid=0 sgid=0 fsgid=0 tty=(none) ses=4294967295
comm="auditctl" exe="/usr/sbin/auditctl" subj=unconfined key=(null)
type=CONFIG_CHANGE msg=audit(1721237716.900:98064): auid=4294967295 ses=4294967295
subj=unconfined op=add_rule key="passwd_changes" list=4 res=1
```

You can also use **ausearch** in conjunction with **aureport** for detailed reports as shown in Example 6-19. The command is:

```
sudo aureport -k
```

Example 6-19 Audit report

```
Key Report
===============================================
# date time key success exe auid event
===============================================
17/07/24 18:35:16 passwd_changes yes /usr/sbin/auditctl -1 98064
17/07/24 18:35:16 shadow_changes yes /usr/sbin/auditctl -1 98065
17/07/24 18:35:16 privileged yes /usr/sbin/auditctl -1 98066
17/07/24 18:35:33 shadow_changes yes /usr/bin/passwd 1000 98076
```

6.4.5 File system security

With detailing and auditing we can know who does what, when and how. But, how can we know if among all the millions of events that can happen in a system, one of them, for example, modifies a critical file or even worse, thousands of them at the same time (ransomware attack)?

AIDE helps to monitor and verify the integrity of files and directories on a system. It helps detect unauthorized changes, such as modifications, deletions, or additions, by creating a database of file attributes and comparing the current state to the baseline. It is a File Integrity Monitor initially developed as a free and open source replacement for Tripwire licensed under the terms of the GNU General Public License

To install AIDE (Debian), use this command:

```
sudo apt-get install aide
```

To begin using AIDE, you must make sure the database is present:

```
ls /var/lib/aide
```

If you see the file *aide.db* in the output of the **ls** command, then proceed to the initialization step. If, instead, you see the file *aide.db.new*, then you need to rename the *aide.db.new* file to *aide.db* using this command:

```
sudo mv /var/lib/aide/aide.db.new /var/lib/aide/aide.db
```

Once the AIDE database is in place, you can initialize the database with this command from a terminal prompt: (this can take a while, go for a coffee)

```
aide --config /etc/aide/aide.conf --init
```

To perform an initial check of the directories and files specified in */etc/aide/aide.conf*, enter this command in a terminal prompt:

```
sudo aide -check
```

If everything in the monitored directories and files is correct, you will see the following message when the check completes:

```
All files match AIDE database. Looks okay!
```

AIDE will also run daily via the */etc/cron.daily/aide* crontab, and the output will be emailed to the user specified in the **MAILTO=** directive of the */etc/default/aide* configuration file as mentioned above.

Set up a cron job for regular checks

```
sudo crobtab -e 0 0 * * * /usr/bin/aide --check
```

AIDE is able to determine what changes were made to a system, but is not able to determine who made the change, when the change occurred, and what command was used to make the change. For that, you use **auditd** and **ausearch**

By combining these tools, you establish a robust system for logging, integrity checking, and auditing. This multi-layered approach enhances the security and integrity of your Linux installation on ppc64le architecture, providing early detection of potential security incidents and unauthorized changes.

> **Tip:** Forwarding these events to a SIEM or remote log solution (including PowerSC Trusted Logging on VIOS) would be the best practice to ensure these logs are tamper-proof stored and therefore cannot be modified or deleted.This applies to any other log or audit file of security interest.

6.4.6 SIEM & EDR Integration

CIS recommends configuring Linux systems to send logs to a centralized log server. This enhances security by protecting logs from being tampered with on the local machine and simplifies log management.

Integrating Linux on IBM Power systems with SIEM tools such as IBM QRadar involves several steps to ensure that logs from the Linux systems are properly collected, transmitted, and ingested by the SIEM platform. The same steps apply if instead of a classic SIEM tool it is a remote log collector or other observability tool that allows us to centralize the logs from different environments for their secure storage and correct subsequent analysis.

Figure 6-6 shows the concept.

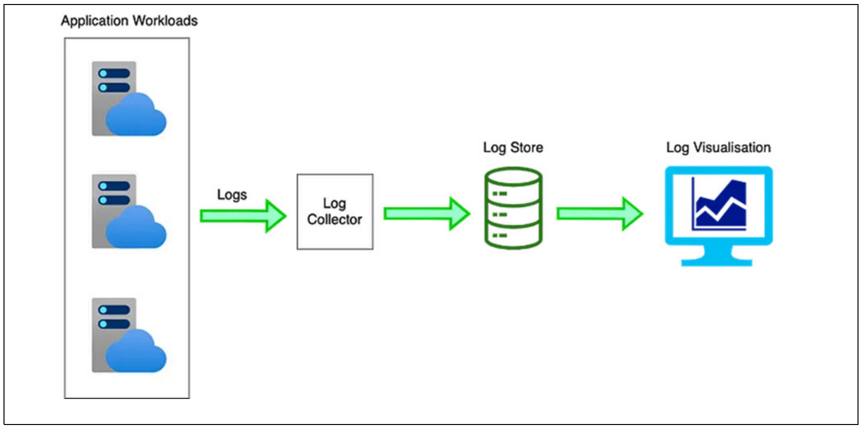


Figure 6-6 Log management

There are two approaches. The classical one will involve using syslog based demon. Syslog is a protocol created in the 1980s to handle the syslog protocol. It remains the default on OpenBSD. We typically have two options:

syslog-ng
: Developed in the late 1990s as a robust replacement for syslog. It introduced support for TCP, encryption, and numerous other features. Syslog-ng became the standard and was included in distributions such as SUSE, Debian, and Fedora for many years.

Example 6-20 shows a configuration example to forward auth logs to a remote SIEM on IP 1.2.3.4 (Debian based)

Example 6-20 Configuration to forward logs to remote SIEM

```
# Source configuration
source s_src {
    system();
    internal();
};

# Destination configuration for QRadar
destination d_remotesiem {
    udp("1.2.3.4" port(514));
};

# Filter for authentication logs
filter f_auth { facility(auth, authpriv); };

# Log path for authentication logs to QRadar
log {
    source(s_src);
    filter(f_auth);
    destination(d_remotesiem);
};
```

rsyslog

Launched in 2004 as a competitor to syslog-ng, it has become the default syslog daemon on Ubuntu, Red Hat Enterprise Linux, and many other distributions. If you are using a common, up-to-date Linux distribution, you are likely using rsyslog by default.

Example 6-21 shows forwarding the auth log using rsyslog (Red Hat Enterprise Linux based)

Example 6-21 Forwarding auth log

```
# Load necessary modules
module(load="imfile")  # File reader module
module(load="omfwd")   # Forwarding module

# Input configuration for auth logs
input(type="imfile" File="/var/log/secure" Tag="auth-log"
StateFile="state-auth-log")

# Forward auth logs to remote SIEM
if $syslogtag == 'auth-log' then {
    action(type="omfwd" target="1.2.3.4" port="514" protocol="udp"
template="RSYSLOG_ForwardFormat")
```

The problem with syslog messages is they need to be parsed to extract every field using regular expressions which is very time consuming.

A second approach would be to send JSON or field-based logs to IBM QRadar without using traditional syslog daemons or after storing this messages in a database. This can be done with tools like Fluentd or even your own scripts in python.

Fluentd, a extensively deployed open-source log collector written in Ruby, stands out for its versatile pluggable architecture. This design enables it to seamlessly connect to a broad array of log sources and storage solutions, including Elasticsearch, Loki, Rsyslog, MongoDB, AWS S3 object storage, and Apache Kafka, among others. Figure 6-7 shows how Fluentd can help with log management.

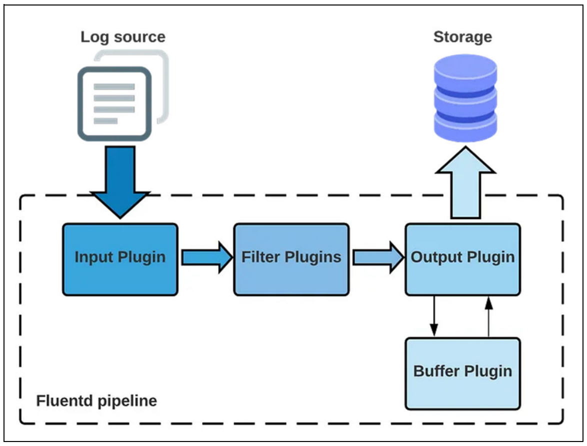


Figure 6-7 Fluentd used for log management

IBM leverages Fluentd to streamline its log management processes across diverse environments including sending logs from kubernetes based deployments on IBM Cloud or OpenShift `https://cloud.ibm.com/docs/containers?topic=containers-health` and can be deployed in Power using a docker image `https://hub.docker.com/r/ppc64le/fluentd`.

6.4.7 Malware Protection

Malware, which encompasses any malicious software designed to harm or exploit programmable devices, services, or networks, includes ransomware-a specific type of malware that encrypts victims' files and demands a ransom for their release-and viruses, which attach to programs or files, spreading from one computer to another.

To prevent these threats on Linux systems, ClamAV can be utilized to detect and remove various forms of malware through regular scans and real-time protection, while chkrootkit can identify and report signs of rootkits, both tools enhancing security by ensuring the system remains free from unauthorized access and malicious activity.

Both tools are open source and available for ppc64le.

Virus detection

ClamAV (Comprehensive Malware Detection and Removal) is a versatile and powerful open-source anti-virus engine designed for detecting Trojans, viruses, malware, and other malicious threats. It offers several features that make it a valuable tool for enhancing Linux system security:

- Regular Scans: ClamAV can be configured to perform regular scans of the system, ensuring that any new or existing malware is promptly detected and addressed.
- Real-Time Protection: With the ClamAV daemon, real-time scanning can be enabled to monitor file activity continuously, providing immediate detection and response to potential threats.
- Automatic Updates: ClamAV includes an automatic update mechanism for its virus definitions, ensuring that the system is protected against the latest threats.
- Cross-Platform Support: ClamAV supports multiple platforms, making it a flexible solution for various environments working on Linux on Power but also in AIX and IBM i (PASE)

To install and configure ClamAV on a Linux system, follow these steps (Debian based):

- Install ClamAV using this command:

```
sudo apt-get install clamav clamav-daemon
```

- Update ClamAV database using this command:

```
sudo freshclam
```

- Start ClamAV daemon using this command:

```
sudo systemctl start clamav-daemon
```

- Schedule a daily scan (add this line to your crontab) and send a report by email using this command:

```
MAILTO=admin@acme.com 0 1 * * * /usr/bin/clamscan -ri --no- summary /
```

Tip: ClamAV is also available for AIX and IBM i.

Rootkit detection

The tool `chkrootkit` is a rootkit detector that checks for signs of rootkits on Unix-based systems. It scans for common signatures of known rootkits and helps ensure the system remains uncompromised.

You can scan for many types of rootkits and detect certain log deletions using chkrootkit. While it doesn't remove any infected files, it does specifically tell you which ones are infected, so that you can remove/reinstall/repair the file or package.

To install **chkrootkit** use this commands:

```
sudo apt-get install chkrootkit.
```

To then execute **chkrootkit** use this command:

```
sudo chkrootkit
```

You can also schedule a daily scan via **cron**.

6.4.8 Backup strategy

Regular backups are essential for data recovery in case of an attack or system failure. Implementing automated backup solutions is a must. There are open source projects with support for ppc64le like Bacula and enterprise-grade solutions such as IBM Storage Protect and others.

Bacula is an enterprise-grade open-source backup solution designed to automate the process of backing up, recovering, and verifying data across a network of computers. It is highly flexible and scalable, making it suitable for various environments, from small businesses to large enterprises. Bacula is particularly known for its robustness, extensive feature set, and support for multiple operating systems, including Linux / ppc64le (fedora)

IBM Storage Protect is a comprehensive, enterprise-grade data protection solution engineered to safeguard critical data across diverse environments. Designed to automate the processes of data backup, recovery, and archiving, IBM Storage Protect ensures that business-critical information remains secure, available, and verifiable. Its robust architecture and extensive feature set make it an ideal choice for organizations of all sizes, from small businesses to large enterprises. Both the clients and the server itself are well supported on Linux / ppc64le. For more information see the IBM Storage Protect documentation.

6.4.9 Consistent update strategy

Updating and maintaining Linux environments is crucial for ensuring security, performance, and compliance. For enterprises using IBM Power (ppc64le) systems, integrating Foreman/Katello and their supported products such as Red Hat Satellite and Orcahino offers a robust solution for consistent and automated updates. To date, server packages are not available on ppc64le but client packages are.

Using Formeman and Katello together with Red Hat Satellite provides one option for update management.

Foreman is an open-source lifecycle management tool that allows system administrators to manage servers throughout their lifecycle, from provisioning and configuration to monitoring and management. It provides an integrated, comprehensive solution for managing large-scale infrastructures.

Katello is a plug-in for Foreman that adds content management and subscription management capabilities. It allows administrators to manage software repositories, handle updates, and ensure compliance with subscription policies.

Figure 6-8 shows how Red Hat Satellite works.

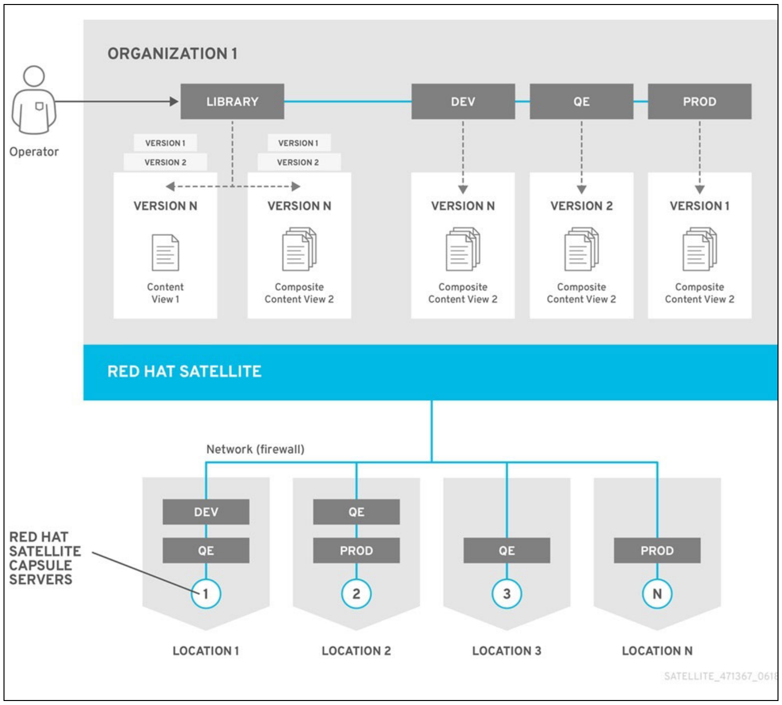
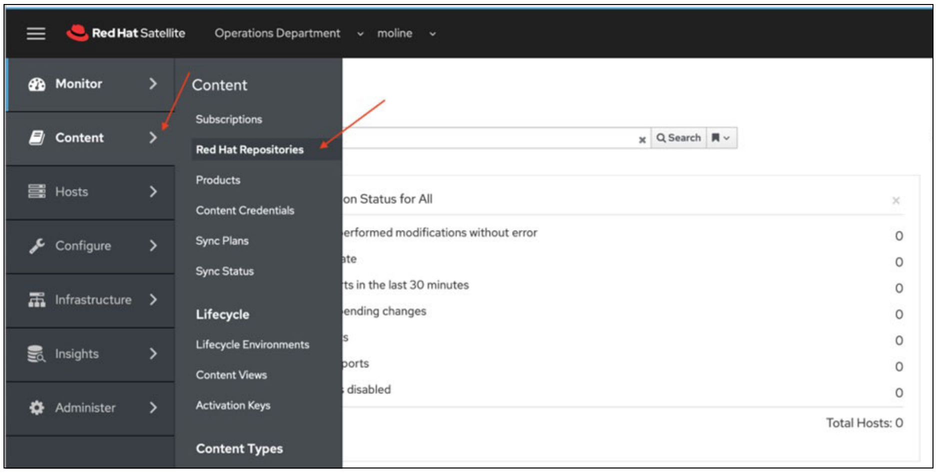


Figure 6-8 Red Hat Satellite

Figure 6-9 shows Red Hat Satellite's GUI.



on Status for All

erformed modifications without error	0
ate	0
ts in the last 30 minutes	0
ending changes	0
s	0
ports	0
disabled	0

Total Hosts: 0

Figure 6-9 Red Hat Satellite GUI

Using Red Hat Satellite along with Foreman/Katello manages package and patch lifecycles, including update distribution. Using Red Hat Satellite in this environment can initiate updates. However, Ansible offers a more comprehensive automation solution for keeping systems up to date. Ansible can:

- Perform prechecks, backups, and snapshots
- Initiate patch updates,
- Reboot systems
- Conduct post-checks for complete patch automation

Thus, combining Satellite and Ansible is optimal. Satellite handles lifecycle management and package provision, while Ansible automates the entire patching process. This integration ensures efficient and consistent updates across your environment.

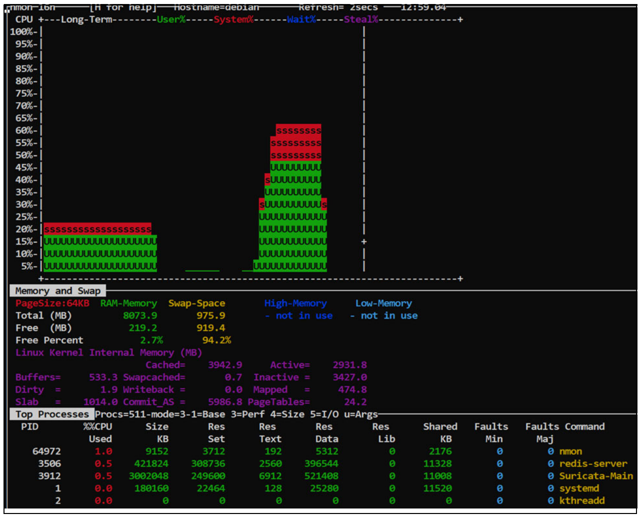
6.4.10 Monitoring

Monitoring Linux on Power systems plays a vital role in ensuring their security, supplementing the specialized tools mentioned in this chapter and providing additional insights.

Specific examples include detecting abnormal CPU or network consumption by unfamiliar applications, which may indicate underlying issues. Proper sizing of workloads and appropriate distribution of system resources contribute significantly to their accessibility.

There are many options for monitoring Linux systems in Power. Most commercial and community solutions have ppc64le agents. As an example, consider Pandora FMS. There are also solutions that support monitoring Linux on Power and also fit into a complete monitoring infrastructure across all of your IBM Power workloads running on Linux, AIX, and IBM i where you can visualize the status any partition, and generate alerts which can be redirected to a centralized monitoring environment.

One of the simplest options for this multiple architecture monitoring tool would be nmon. Nmon was originally written for AIX and is now an integrated tool within AIX. A version for Linux was written by IBM and later released as open source for Linux across multiple platforms including x86, IBM Power, IBM Z and even ARM. There are multiple integrations for using and analyzing nmon data, including charts and spreadsheet integrations. There is even a newer version (nmonj) that saves the performance data in JSON format for better integration into databases and for web browser graphing. Figure 6-10 show nmon reporting on utilization.

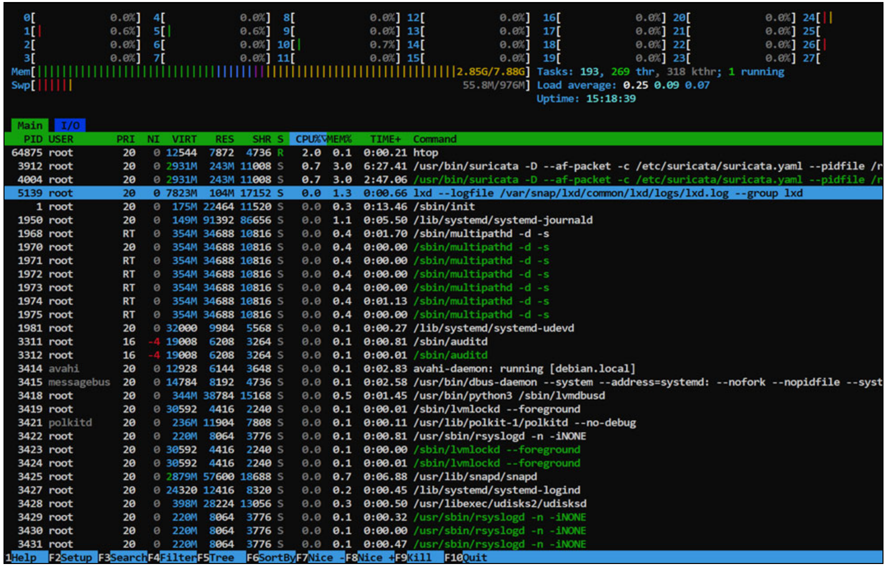


```
nmon-16n      [H for help]   Hostname=debian       Refresh= 2secs   12:59.04
 CPU +---Long-Term--------User%-----System%------Wait%-----Steal%-------------+
100%-|                                                 |
 95%-|                                                 |
 90%-|                                                 |
 85%-|                                                 |
 80%-|                                                 |
 75%-|                                                 |
 70%-|                                                 |
 65%-|                                                 |
 60%-|                                     ssssssss    |
 55%-|                                    sssssssss    |
 50%-|                                    sssssssss    |
 45%-|                                    UUUUUUUUU    |
 40%-|                                   sUUUUUUUUU    |
 35%-|                                   UUUUUUUUUU    |
 30%-|                                  sUUUUUUUUUUs   |
 25%-|                                  UUUUUUUUUUUU   |
 20%-|sssssssssssssssssss               UUUUUUUUUUUU   |
 15%-|UUUUUUUUUUUUUUUUUUUU              UUUUUUUUUUUU   +
 10%-|UUUUUUUUUUUUUUUUUUUU              UUUUUUUUUUUU   |
  5%-|UUUUUUUUUUUUUUUUUUUU    ______  _UUUUUUUUUUUUU   |
     +--------------------------------------------------------------------------+
 Memory and Swap
 PageSize:64KB  RAM-Memory  Swap-Space       High-Memory     Low-Memory
 Total (MB)         8073.9       975.9       - not in use    - not in use
 Free  (MB)          219.2       919.4
 Free Percent          2.7%       94.2%
 Linux Kernel Internal Memory (MB)
                       Cached=    3942.9     Active=    2931.8
 Buffers=     533.3 Swapcached=      0.7  Inactive =    3427.0
 Dirty  =       1.9 Writeback =      0.0  Mapped   =     474.8
 Slab   =    1014.0 Commit_AS =   5986.8 PageTables=      24.2
 Top Processes Procs=511-mode=3-1=Base 3=Perf 4=Size 5=I/O u=Args
  PID       %%CPU     Size      Res      Res      Res      Res   Shared  Faults  Faults Command
             Used       KB      Set     Text     Data      Lib       KB     Min     Maj
   64972      1.0     9152     3712      192     5312        0     2176       0       0 nmon
    3506      0.5   421824   308736     2560   396544        0    11328       0       0 redis-server
    3912      0.5  3002048   249600     6912   521408        0    11008       0       0 Suricata-Main
       1      0.0   180160    22464      128    25280        0    11520       0       0 systemd
       2      0.0        0        0        0        0        0        0       0       0 kthreadd
```

Figure 6-10 A nmon display

Another tool is htop, an interactive process viewer that offers several enhanced functionalities that make it particularly user-friendly and versatile. For example, allowing users to scroll through and select processes for detailed information, and to make changes in priority and terminate processes directly from the interface. Figure 6-11shows an example screen from htop.

Figure 6-11 Screen shot of htop



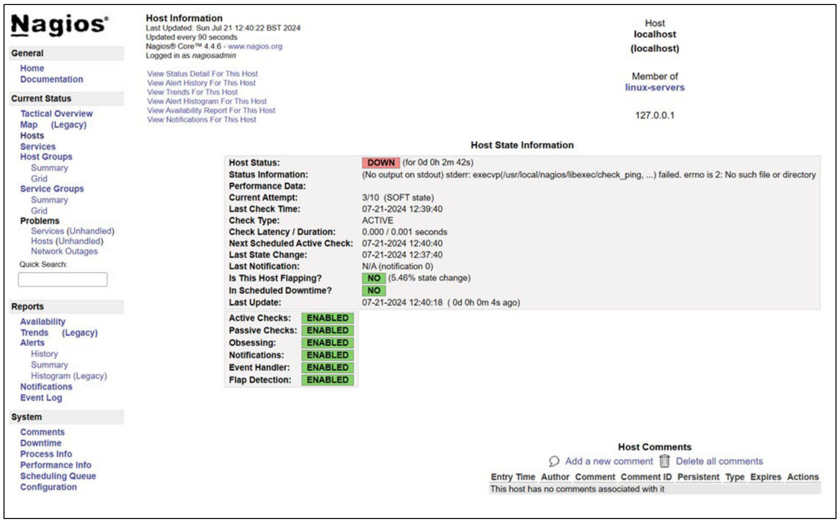
```
    0[           0.0%]  4[           0.0%]  8[           0.0%] 12[           0.0%] 16[           0.0%] 20[           0.0%] 24[||
    1[|          0.6%]  5[|          0.6%]  9[           0.0%] 13[           0.0%] 17[           0.0%] 21[           0.0%] 25[
    2[           0.0%]  6[           0.0%] 10[|          0.7%] 14[           0.0%] 18[           0.0%] 22[           0.0%] 26[|
    3[           0.0%]  7[           0.0%] 11[           0.0%] 15[           0.0%] 19[           0.0%] 23[           0.0%] 27[
  Mem[||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||2.85G/7.88G] Tasks: 193, 269 thr, 318 kthr; 1 running
  Swp[|||||                                                                 55.8M/976M] Load average: 0.25 0.09 0.07
                                                                                         Uptime: 15:18:39

 Main  I/O
   PID USER       PRI  NI  VIRT   RES   SHR S  CPU%▽MEM%   TIME+  Command
 64875 root        20   0 12544  7872  4736 R   2.0  0.1  0:00.21 htop
  3912 root        20   0 2931M  243M 11008 S   0.7  3.0  6:27.41 /usr/bin/suricata -D --af-packet -c /etc/suricata/suricata.yaml --pidfile /r
  4004 root        20   0 2931M  243M 11008 S   0.7  3.0  2:47.06 /usr/bin/suricata -D --af-packet -c /etc/suricata/suricata.yaml --pidfile /r
  5139 root        20   0 7823M  104M 17152 S   0.0  1.3  0:00.66 lxd --logfile /var/snap/lxd/common/lxd/logs/lxd.log --group lxd
     1 root        20   0  175M 22464 11520 S   0.0  0.3  0:13.46 /sbin/init
  1950 root        20   0  149M 91392 86656 S   0.0  1.1  0:05.50 /lib/systemd/systemd-journald
  1968 root        RT   0  354M 34688 10816 S   0.0  0.4  0:01.70 /sbin/multipathd -d -s
  1970 root        20   0  354M 34688 10816 S   0.0  0.4  0:00.00 /sbin/multipathd -d -s
  1971 root        RT   0  354M 34688 10816 S   0.0  0.4  0:00.00 /sbin/multipathd -d -s
  1972 root        RT   0  354M 34688 10816 S   0.0  0.4  0:00.00 /sbin/multipathd -d -s
  1973 root        RT   0  354M 34688 10816 S   0.0  0.4  0:00.00 /sbin/multipathd -d -s
  1974 root        RT   0  354M 34688 10816 S   0.0  0.4  0:01.13 /sbin/multipathd -d -s
  1975 root        RT   0  354M 34688 10816 S   0.0  0.4  0:00.00 /sbin/multipathd -d -s
  1981 root        20   0 32000  9984  5568 S   0.0  0.1  0:00.27 /lib/systemd/systemd-udevd
  3311 root        16  -4 19008  6208  3264 S   0.0  0.1  0:00.81 /sbin/auditd
  3312 root        16  -4 19008  6208  3264 S   0.0  0.1  0:00.01 /sbin/auditd
  3414 avahi       20   0 12928  6144  3648 S   0.0  0.1  0:02.83 avahi-daemon: running [debian.local]
  3415 messagebus  20   0 14784  8192  4736 S   0.0  0.1  0:02.58 /usr/bin/dbus-daemon --system --address=systemd: --nofork --nopidfile --syst
  3418 root        20   0  344M 38784 15168 S   0.0  0.5  0:01.45 /usr/bin/python3 /sbin/lvmdbusd
  3419 root        20   0 30592  4416  2240 S   0.0  0.1  0:00.01 /sbin/lvmlockd --foreground
  3421 polkitd     20   0  236M 11904  7808 S   0.0  0.1  0:00.11 /usr/lib/polkit-1/polkitd --no-debug
  3422 root        20   0  220M  8064  3776 S   0.0  0.1  0:00.81 /usr/sbin/rsyslogd -n -iNONE
  3423 root        20   0 30592  4416  2240 S   0.0  0.1  0:00.00 /sbin/lvmlockd --foreground
  3424 root        20   0 30592  4416  2240 S   0.0  0.1  0:00.01 /sbin/lvmlockd --foreground
  3425 root        20   0 2879M 57600 18688 S   0.0  0.7  0:06.88 /usr/lib/snapd/snapd
  3427 root        20   0 24320 12416  8320 S   0.0  0.2  0:00.45 /lib/systemd/systemd-logind
  3428 root        20   0  398M 28224 13056 S   0.0  0.3  0:00.50 /usr/libexec/udisks2/udisksd
  3429 root        20   0  220M  8064  3776 S   0.0  0.1  0:00.32 /usr/sbin/rsyslogd -n -iNONE
  3430 root        20   0  220M  8064  3776 S   0.0  0.1  0:00.00 /usr/sbin/rsyslogd -n -iNONE
  3431 root        20   0  220M  8064  3776 S   0.0  0.1  0:00.47 /usr/sbin/rsyslogd -n -iNONE
1Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

There are more and more projects being developed using python and other languages that are easily portable between architectures. Some of them have good export capabilities to InfluxDB, Cassandra, OpenTSDB, StatsD, ElasticSearch or RabbitMQ.

Nagios

At a next level, we would have the deployment of complete monitoring environments such as Nagios or Zabbix. These frameworks support extensive customization and scalability. Their source code can be easily downloaded and compiled on IBM Power, with agents / integrations for both AIX and IBM i. Some of them have enterprise support options or require licenses from a number of instances.

Figure 6-12 is an example of Nagios (core version) running on Debian on Power.



Host Information
Last Updated: Sun Jul 21 12:40:22 BST 2024
Updated every 90 seconds
Nagios® Core™ 4.4.6 - www.nagios.org
Logged in as *nagiosadmin*

Host
localhost
(localhost)

Member of
linux-servers

127.0.0.1

Host State Information

Host Status:	DOWN (for 0d 0h 2m 42s)
Status Information:	(No output on stdout) stderr: execvp(/usr/local/nagios/libexec/check_ping, ...) failed. errno is 2: No such file or directory
Performance Data:	
Current Attempt:	3/10 (SOFT state)
Last Check Time:	07-21-2024 12:39:40
Check Type:	ACTIVE
Check Latency / Duration:	0.000 / 0.001 seconds
Next Scheduled Active Check:	07-21-2024 12:40:40
Last State Change:	07-21-2024 12:37:40
Last Notification:	N/A (notification 0)
Is This Host Flapping?	NO (5.46% state change)
In Scheduled Downtime?	NO
Last Update:	07-21-2024 12:40:18 (0d 0h 0m 4s ago)

Active Checks:	ENABLED
Passive Checks:	ENABLED
Obsessing:	ENABLED
Notifications:	ENABLED
Event Handler:	ENABLED
Flap Detection:	ENABLED

Host Comments

Entry Time	Author	Comment	Comment ID	Persistent	Type	Expires	Actions
This host has no comments associated with it							

Figure 6-12 Nagios running on Linux on IBM Power

IBM Instana

In the field of commercial monitoring solutions, we highlight IBM Instana™. It leverages various open-source projects to provide advanced monitoring and observability capabilities, making it an excellent enterprise-supported solution for monitoring Linux on Power (ppc64le) systems but also AIX and IBM i.

IBM Instana® integrates with technologies such as Apache Kafka for real-time data processing, Prometheus for metrics collection, Grafana for data visualization, OpenTelemetry for tracing and metrics, Elastic Stack (ELK) for log management, Kubernetes for container orchestration, and Jenkins for continuous integration and delivery.

With support for Debian, Red Hat, and SUSE on ppc64le, Instana ensures comprehensive, real-time visibility into the performance and health of applications and systems, backed by IBM's robust enterprise support

More information can be found at `https://www.ibm.com/products/instana/`

6.5 Best practices

To ensure the security and reliability of Linux on IBM Power Systems (ppc64le), several best practices should be implemented. These practices include system hardening, regular updates, access control, and data protection strategies.

System Hardening

- Minimal Installation: Begin with a minimal base installation to reduce the attack surface. Only install necessary software and services.This will reduce the surface attack by limiting installed software and services. It is always easier to add software than to remove it.
- Compliance: use tools such as OpenSCAP or PowerSC to help ensure minimum levels of compliance in all systems. This can be done by generating a base image and then being rigorous in change control using tools such as Ansible for configuration management to enforce consistent security policies across all systems.
- Patch Management: Regularly apply security patches and updates. Use automated tools like Red Hat Satellite to keep the system not only current but also consistent. Automation tools will help you to do the hard work.
- File System Security: Implement secure file system permissions and use encryption for sensitive data. Regularly audit file permissions and access controls. Use tools like AIDE to monitor and verify the integrity of files and directories in order to protect sensitive data.

Access Control

- User Authentication: Implement strong authentication mechanisms, including multi-factor authentication (MFA). Use SSH key pairs instead of passwords for remote access, if possible in combination with second method of authentication.
- Role-Based Access Control (RBAC): Assign permissions based on roles rather than individual users. Sudo is a great, powerful and probably the easiest tool to do it locally.
- Password Policies: Enforce strong password policies, including complexity requirements, expiration, and account lockout mechanisms.

Data Protection

- Encryption: Use encryption for data at rest and in transit. Implement SSL/TLS for network communications and encrypt sensitive files on disk.
- Backup Strategies: Regularly back up critical data and test restore procedures. Use tools like Bacula or IBM Storage Protect for automated backups.

Regular monitoring and logging

- Effective monitoring and logging are essential for detecting and responding to security c analysis and compliance requirements. Search and analytics engines like Elasticsearch are a great choice for different use cases.
- SIEM Integration: Integrate logs with Security Information and Event Management (SIEM) systems for real-time analysis and alerting.
- Intrusion Detection and Prevention: Deploy IDS/IPS tools like Suricata to monitor network traffic and detect suspicious activities.
- Regular Audits: Perform regular audits of log files and security configurations to identify and address potential issues. Use tools like auditd and ausearch for detailed audit logs. They can be also forwarded and analyzed by external tools from traditional SIEM's to ELK (Elasticsearch, Logstash, Kibana) stacks.

Summary

In summary, layered security provides enhanced safety. However, it's important to note that even the most robust defenses have weaknesses, which makes their effectiveness dependent on the least secure component. Achieving the right balance between security and usability is essential; while technologies advance and operating systems change, core problems remain and new ones emerge.

6.6 Develop Incident Response Plan

An Incident Response Plan (IRP) is a comprehensive, systematic approach to handling and managing security breaches or cyberattacks. It outlines the procedures and actions an organization must take to detect, respond to, and recover from security incidents, aiming to minimize the impact on business operations, data integrity, and overall security posture.

A well-defined incident response plan is crucial for minimizing the impact of security incidents. Here are the key components of an effective incident response plan:

Preparation: Define roles and responsibilities for incident response specific to Linux on Power environments. Ensure all team members are trained and familiar with the response procedures for this architecture. Make sure you have a clearly defined architecture where the people who specialize in each technology: PowerVM, SUSE, Red Hat, Ubuntu, databases, applications, storage and communications are located and understand the Linux on Power environment.

Identification: Implement monitoring and alerting mechanisms to quickly identify potential security incidents in the Power environment. Utilize log analysis and Security Information and Event Management (SIEM) tools tailored for ppc64le to detect anomalies. Ensure compatibility and optimization of these tools for the Power architecture.

Containment: Develop strategies for containing incidents to prevent further damage. This may involve isolating affected Power Systems or networks. Consider the specific containment techniques suitable for Power hardware, such as leveraging virtualization features to isolate affected Logical Partitions (LPARs), VLANs or shared storage.

Eradication: Identify and remove the root cause of the incident in the Power environment. This may involve applying patches, removing malware, or addressing configuration issues specific to ppc64le systems. Ensure the incident response team is familiar with patch management and malware removal tools compatible with Linux on Power.

Recovery: Restore affected Power Systems to normal operation. This may involve restoring data from backups, rebuilding compromised LPARs, or reconfiguring network settings specific to the Power architecture. Ensure that recovery procedures are tested and validated for ppc64le environments.

Lessons Learned: After resolving an incident, conduct a postmortem analysis to identify lessons learned and improve future response efforts. Update incident response plans and security policies based on findings, considering any unique aspects of the Power environment. Document any architecture-specific issues and resolutions to enhance future readiness.

Tip: Regularly test and update your incident response plan to ensure it remains effective and relevant to the current threat landscape.

7

Red Hat OpenShift Security

This chapter describes the Red Hat OpenShift platform by giving a basic overview of its history, its key features, the role it plays in the modern IT landscape, and how it compares to similar platforms. This chapter also discusses Kubernetes as the foundation of OpenShift, and how OpenShift improves on the fundamental features of Kubernetes. We also see the advantages of using OpenShift for orchestrating containers in your environment. The chapter concludes by providing a brief description of Red Hat OpenShift when implemented on IBM Power infrastructure.

7.1 Red Hat OpenShift Fundamentals

This section provides a quick introduction to Red Hat OpenShift which is a leading enterprise cloud platform that can help you design and build applications to run in a hybrid cloud environment. Red Hat OpenShift can be run on your on-premise infrastructure, across a wide variety of cloud vendors, and even in your Edge environments, allowing seamless migration of services across the hybrid cloud environment.

7.1.1 What is Red Hat OpenShift?

Red Hat OpenShift is a leading enterprise Kubernetes platform that provides a robust foundation for developing, deploying, and scaling cloud-native applications. It extends Kubernetes with additional features and tools to enhance productivity and security, making it an ideal choice for businesses looking to leverage container technology at scale.

Red Hat OpenShift is a unified platform to build, modernize, and deploy applications at scale. Work smarter and faster with a complete set of services for bringing apps to market on your choice of infrastructure. OpenShift delivers a consistent experience across public cloud, on-premise, hybrid cloud, or edge architecture.

Red Hat OpenShift offers you a unified, flexible platform to address a variety of business needs spanning from an enterprise-ready Kubernetes orchestrator to a comprehensive cloud-native application development platform that can be self-managed or used as a fully managed cloud service.

Figure 7-1 shows how Kubernetes is only one component (albeit a critical one) in Red Hat OpenShift.

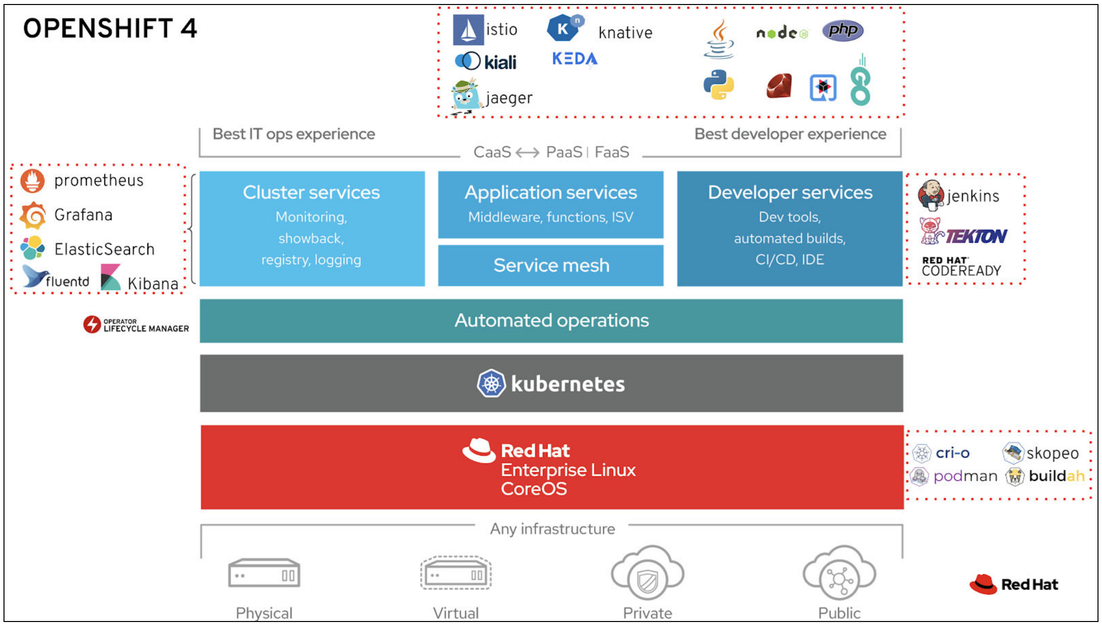


Figure 7-1 OpenShift components

Red Hat OpenShift provides:

- The ability to deploy and run in any environment, the flexibility to build new applications, modernize existing applications, run third-party ISV applications, or use public cloud services under a single platform.
- The tools necessary to help customers integrate data analytics, artificial intelligence and machine learning (AI/ML) capabilities into cloud-native applications to deliver more insight and value.
- Consistency and portability to deploy and manage containerized workloads, make infrastructure and investments future-ready, and deliver speed and flexibility on-premise, across cloud environments, and to the edge of the network.
- Advanced security and compliance capabilities, allowing end-to-end management and observability across the entire architecture.

Built by open source leaders, Red Hat OpenShift includes an enterprise-ready Kubernetes solution with a choice of deployment and usage options to meet the needs of your organization. From self-managed to fully managed cloud services, you can deploy the platform in the data center, in cloud environments, and at the edge of the network. With Red Hat OpenShift, you have the option to get advanced security and compliance capability, end-to-end management and observability, and cluster data management and cloud-native data services. Red Hat Advanced Cluster Security for Kubernetes modernizes container and Kubernetes security, letting developers add security controls early in the software life cycle. Red Hat Advanced Cluster Management for Kubernetes lets you manage your entire application life cycle and deploy applications on specific clusters based on labels, and Red Hat OpenShift Data Foundation supports performance at scale for data-intensive workloads.

7.1.2 Distinguishing features

There are many cloud and container management options available. OpenShift has integrated these features to enhance the cloud and container management experience:

- With a focus on Developer-Centric Tools, OpenShift enhances Kubernetes with developer-friendly tools, including OpenShift Console – a developer-centric view for application management – and Source-to-Image (S2I) technology. This simplifies the process of building reproducible container images from source code.
- Designed with Advanced Security in mind, built-in security at every layer of the application stack – from the operating system (Red Hat Enterprise Linux CoreOS) to the application services – ensures that compliance features and security best practices are built in from the start.
- Focused on Hybrid Cloud capabilities, Red Hat OpenShift is designed to operate across on-premise, public cloud, and hybrid cloud environments, providing consistent application portability and flexibility in deployment options.

Red Hat OpenShift is an enterprise level production product that entitles enterprise level support based on Kubernetes and Kubernetes management. Red Hat OpenShift provides the following benefits:

- Red Hat OpenShift offers automated installation, upgrades, and lifecycle management throughout the container stack – the operating system, Kubernetes, cluster services, and applications – on any cloud.
- Red Hat OpenShift helps teams build with speed, agility, confidence, and choice. Get back to doing work that matters.

- Red Hat OpenShift is focused on security at every level of the container stack and throughout the application lifecycle. It includes long-term, enterprise support from one of the leading Kubernetes contributors and open source software companies.

7.1.3 The role of OpenShift in modern IT

OpenShift plays a crucial role in modern IT by facilitating the DevOps approach, improving software delivery speed, and enabling a more agile development environment. It offers a scalable platform that supports both microservices and traditional application models, accommodating a wide range of programming languages and frameworks. Through its comprehensive tool set, OpenShift addresses the needs of developers, system administrators, and IT managers, making it a pivotal tool in enterprise digital transformation strategies.

7.1.4 Key takeaways

When considering which cloud management platform to use, consider this:

- OpenShift provides extensive enterprise features out-of-the-box, including advanced security features, integrated developer tools, and extensive automation capabilities that may not be as comprehensive in other Kubernetes services.
- OpenShift's ability to deploy across multiple environments (cloud, on-premise, and hybrid) with consistency makes it a strong choice for organizations with complex infrastructure needs.
- OpenShift distinctly benefits developers with features like Source-to-Image (S2I), a comprehensive web console, and application templates, which facilitate a smoother and more productive development experience compared to basic Kubernetes services.

Red Hat OpenShift is a strong leader in the cloud landscape of Kubernetes platforms, and is chosen for its strengths in enterprise environments, multi-environment consistency, and developer-centric features.

7.1.5 Kubernetes fundamentals

Kubernetes serves as the backbone of OpenShift, providing the essential framework for orchestrating containerized applications. Understanding these core concepts is crucial for leveraging OpenShift effectively. This section provides a detailed exploration of the fundamental components and mechanisms of Kubernetes as implemented in OpenShift:

Basic components

The basic components of Kubernetes can be described as:

Pods The smallest deployable units created and managed by Kubernetes. A pod is a group of one or more containers that share storage, network, and specifications on how to run the containers. Pods are ephemeral by nature; they are created and destroyed to match the state specified by users.

Nodes The physical or virtual machines where Kubernetes runs the pods. A node can be a worker node or a master node, although with the latest Kubernetes (and by extension OpenShift) practices, the distinction is often abstracted away, especially in managed environments.

Clusters A cluster consists of at least one worker node and at least one master node. The master node manages the state of the cluster, including scheduling workloads and handling scaling and health monitoring.

Figure 7-2 shows a basic cluster architecture. While a cluster can technically be created with one master node and two worker nodes, best practices generally recommend at least three master nodes, which can share functions and provide failover, and three or more worker nodes to provide failover and scalability.

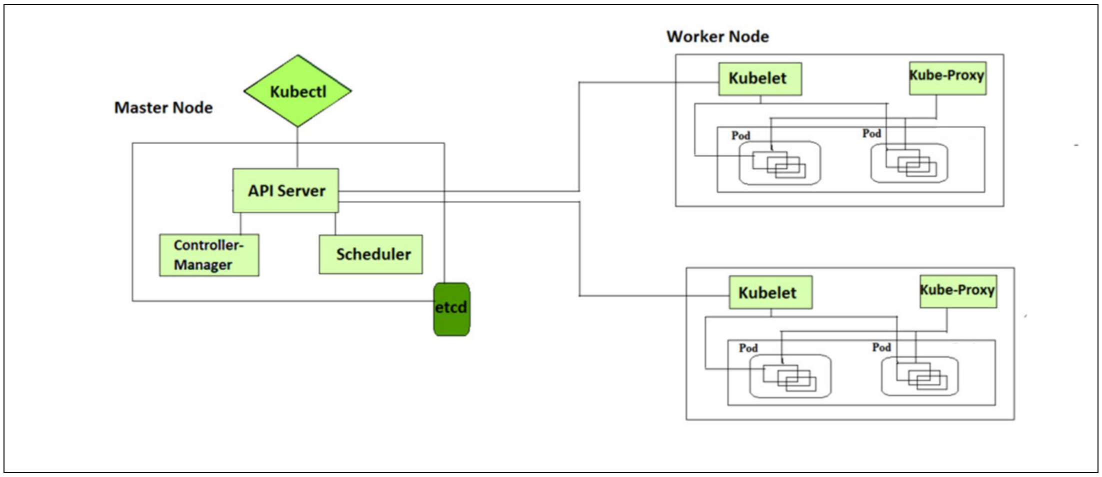


Figure 7-2 Base Kubernetes cluster

Control Plane Components

As described in the previous section, there are two distinct types of nodes in a Kubernetes cluster, a master node and a worker node. The worker nodes are designated to run containers with the applications that run your business. The master nodes run the services that control the cluster and manage the pods which are running the application code, collectively the master nodes create what is called the control plane.

The major services that are running in the control plane are:

API Server Acts as the front end for Kubernetes. The API server is the component that clients and external tools interact with.

etcd A highly-available key-value store used as Kubernetes' backing store for all cluster data. It maintains the state of the cluster.

Scheduler Watches for newly created pods with no assigned node, and selects a node for them to run on based on resource availability, policies, and specifications.

Controller Manager Runs controller processes, which are background tasks in Kubernetes that handle routine tasks such as ensuring the correct number of pods for replicated applications.

Workload Resources

The control plane is in charge of setting up and managing the worker nodes which are running the application code. Workload components can be described as:

Deployments A deployment specifies a desired state for a group of pods. You describe a desired state in a deployment, and the Deployment Controller changes the actual state to the desired state at a controlled rate. You can define deployments to create new ReplicaSets, or to remove existing deployments and adopt all their resources into new deployments.

ReplicaSet A ReplicaSet's purpose is to maintain a stable set of replica pods running at any given time. As such, it is often used to guarantee the availability of a specified number of identical pods. This maintains a stable set of replica pods running at any given time.

StatefulSets Used for applications that require persistent storage and a unique identity for each pod, making them ideal for databases and other stateful applications.

DaemonSets Ensures that each node in the cluster runs a copy of a pod. Useful for deploying system services that need to run on all or certain nodes.

Networking

Networking connectivity between pods and between pods and outside services is managed within a Kubernetes cluster. The following functions are maintained by the cluster:

- **Service**

  An abstraction that defines a logical set of pods and a policy by which to access them. Services enable communication between different pods and external traffic routing into the cluster.

- **Ingress**

  Manages external access to the services in a cluster, typically HTTP. Ingress can provide load balancing, SSL termination, and name-based virtual hosting.

Storage

Containers are by definition ethereal as is any data stored in the container. To enable persistent storage, Kubernetes uses the following concepts:

- **Persistent Volumes** (PV)

  PVs are resources in the cluster which can be connected to containers to provide persistent storage.

- **Persistent Volume Claims (PVC)**

  PVCs are requests for storage by users. These requests are satisfied by allocating PVs.

Configurations and Secrets

- **ConfigMaps**: Allows you to decouple configuration artifacts from image content to keep containerized applications portable.
- **Secrets**: Used to store and manage sensitive information such as passwords, OAuth tokens, and ssh keys.

Security

- **Role-Based Access Control (RBAC)**: Controls authorization – determining what operations a user can perform on cluster resources. It's crucial for maintaining the security of the cluster.

7.1.6 OpenShift Enhancements to Kubernetes

While Kubernetes offers a robust platform for container orchestration, OpenShift enhances this foundation with additional features and tools designed to meet the needs of enterprise environments and developer workflows. These enhancements improve usability, security, and operational efficiency and include:

  – A user-friendly web console for easier management and monitoring

  - Enhanced security features tailored to strict compliance requirements
  - Built-in tools for CI/CD to streamline the development process

Here's a detailed look at how OpenShift builds on the core Kubernetes architecture:

- Enhanced Developer Productivity
  - OpenShift includes a sophisticated web-based console that provides a more user-friendly interface than the standard Kubernetes dashboard. This console allows developers to manage their projects, visualize the state of their applications, and access a broad range of development tools directly.
  - Code-Ready Containers simplifies the setup of local OpenShift clusters for development purposes, providing a minimal, preconfigured environment that can run on a developer s workstation. It s particularly useful for simplifying the “getting started” experience.
  - The Source-to-Image (S2I) tool is a powerful feature for building reproducible container images from source code. This tool automates the process of downloading code, injecting it into a container image, and assembling a new image. The new image incorporates runtime artifacts necessary to execute the code, thus streamlining the workflow from source code to deployed application.
- Advanced Security Features
  - OpenShift enhances Kubernetes security by implementing Security Context Constraints. SCCs are akin to Pod Security Policies but provide more granular security controls over the deployment of pods. They allow administrators to define a set of conditions that a pod must run with to be accepted into the system, such as forbidding running containers as root.
  - OpenShift integrates an OAuth server that can connect to external identity providers, allowing for a streamlined authentication and authorization process. This integration enables users to log into OpenShift using their corporate credentials, simplifying access management and enhancing security.
  - OpenShift provides extensive support for Kubernetes network policies, which dictate how pods communicate with each other and other network endpoints. OpenShift takes this further with the introduction of egress firewall capabilities, allowing administrators to control outbound traffic from pods to external networks.
- Operational Efficiency
  - OpenShift fully embraces the Kubernetes Operator pattern, which extends Kubernetes capabilities by automating the deployment, scaling, and management of complex applications. OpenShift includes the Operator Hub, a marketplace where users can find and deploy Operators for popular software stacks.
  - OpenShift offers a streamlined and highly automated installation process that simplifies the setup of production-grade Kubernetes clusters. This extends to updates, which can be applied automatically across the cluster, reducing downtime and manual intervention.
  - OpenShift includes built-in monitoring and telemetry capabilities that are preconfigured to collect metrics from all parts of the cluster. This feature provides insights into the performance and health of applications and infrastructure, enabling proactive management and troubleshooting.
- Enterprise Integration and Support
  - OpenShift integrates Istio-based service mesh capabilities directly into the platform, facilitating microservices architecture by providing service discovery, load balancing, failure recovery, metrics, and monitoring, along with complex operational requirements like A/B testing, canary releases, and more.

  - OpenShift integrates with various Continuous Integration and Continuous Deployment tools, offering built-in Jenkins support and integrations with other major CI/CD platforms. This integration supports automation of the build, test, and deployment lifecycle within the same platform.

7.1.7 Key Features of OpenShift

OpenShift is strongly focused on the developer's experience and has integrated many features that are designed to make development of applications more efficient and productive. This section provides an overview of some of those enhancements.

Developer Productivity

OpenShift is designed to enhance developer productivity by streamlining processes and reducing the complexities typically associated with deploying and managing applications. Here is a detailed look at how OpenShift achieves this through its key features:

- Developer-Focused User Interface
  - The OpenShift Console is a powerful, user-friendly interface that provides developers with an overview of all projects and resources within the cluster. It offers a perspective tailored to developers' needs, allowing them to create, configure, and manage applications directly from the browser. Features like the Topology view let developers visualize their applications and services in a graphical interface, making it easier to understand and manage the relationships between components.
  - OpenShift includes a Developer Catalog that offers a wide array of build and deploy solutions, such as databases, middleware, and frameworks, which can be deployed on the cluster with just a few clicks. This self-service portal accelerates the setup process for developers, allowing them to focus more on coding and less on configuration.
- Code-Ready Workspaces
  - OpenShift integrates with Code-Ready Workspaces, a Kubernetes-native IDE that developers can use within their browser. This IDE provides a fully featured development environment, complete with source code management, runtimes, and dependencies that are all managed and kept consistent across the development team. This ensures that the entire team works within a controlled and replicable environment, reducing “works on my machine” problems.
- Application Templates and S2I
  - OpenShift application templates are predefined configurations for creating applications based on specific languages, frameworks, or technologies. These templates include everything needed to build and deploy an application quickly, such as build configurations, deployment strategies, and required services.
  - S2I is a tool for building reproducible Docker images from source code. S2I lets developers build containerized applications without needing to write Dockerfiles or become experts in Docker. It combines source code with a base Docker image that contains the appropriate runtime environment for the application. The result is a ready-to-run Docker image built according to best practices.
- Automated Build and Deployment Pipelines
  - OpenShift has robust support for CI/CD processes, integrating tools like Jenkins, GitLab CI, and others directly into the platform. It automates the build, test, and deployment pipeline, enabling developers to commit code changes frequently without the overhead of manual steps.

  - OpenShift can automatically trigger builds and deployments when code changes are pushed to a source code repository or when other specified events occur. This feature ensures that applications are always up-to-date with the latest code changes.
- Live Application Development
  - OpenShift supports hot deployments, where changes to application code can be made active without restarting the entire application. This capability is crucial for environments where uptime is critical, and it allows developers to see changes instantly.
  - Developers can access real-time logs and debugging tools directly through the OpenShift console, making it easier to diagnose and resolve issues in development and production environments.

By focusing on these aspects of developer productivity, OpenShift significantly lowers the barrier to entry for deploying applications in a Kubernetes environment, simplifies the management of these applications, and accelerates the development cycle. This enables developers to spend more time coding and less time dealing with deployment complexities, leading to faster innovation and deployment cycles in a cloud-native landscape.

7.1.8 Enterprise-Grade Security

OpenShift is designed with security as a foundational aspect, integrating robust security features that support the demanding requirements of enterprise environments. This includes everything from strict access controls to ensuring container and platform integrity. Here is a deeper look into how OpenShift delivers enterprise-grade security:

- Security Context Constraints (SCC)
  - OpenShift enhances the security of container environments by using Security Context Constraints (SCC) to define a set of conditions that a container must comply with to run on the platform. These role-based constraints can limit the actions that a pod can perform and the resources it can access, significantly reducing the risk of unauthorized actions.
  - Fine-grained Permissions: Administrators can use SCCs to manage permissions at a granular level, controlling whether pods can run as privileged containers, access sensitive volumes, or use host networking and ports, among other security settings.
- Integrated Authentication and Authorization
  - OpenShift integrates with existing enterprise authentication systems, such as LDAP, Active Directory, and public OAuth providers, to provide a robust user authentication process seamlessly across the organization.
  - RBAC (role based access control) in OpenShift allows administrators to regulate access to resources based on the roles of individual users within the enterprise. This ensures that only authorized users have access to control critical operations, thereby securing the environment against internal and external threats.
- Network Policies and Encryption
  - OpenShift allows administrators to define network policies that govern how pods communicate with each other and with other network endpoints. This ensures that applications are isolated and protected from network-based attacks.
  - Data in transit and at rest can be encrypted, providing an additional layer of security. OpenShift supports TLS for all data in transit and can integrate with enterprise key management solutions to manage encryption keys for data at rest.

- Security Enhancements and Compliance
  - OpenShift provides automated mechanisms to apply security patches and updates to the container host, runtime, and the application containers themselves. This helps in maintaining security compliance and reducing the vulnerability window.
  - OpenShift includes features to support compliance with various regulatory requirements such as PCI DSS, HIPAA, and GDPR. It provides extensive logging and auditing capabilities that help in tracking all user actions and system changes, crucial for forensic analysis and compliance reporting.
- Container Security and Image Assurance
  - OpenShift integrates with tools like Quay.io to provide automated container image scanning. This scans images for vulnerabilities before they are deployed, and image signing ensures that only approved and verified images are used in the environment.
  - Running on Red Hat Enterprise Linux, OpenShift leverages SELinux to enforce mandatory access control policies that isolate containers from each other and from the host system. This prevents a compromised container from affecting others or gaining undue access to host resources.
- Secure Default Settings and Practices
  - OpenShift encourages the use of minimal base images that contain only the essential packages needed to run applications, reducing the potential attack surface.
  - OpenShift is preconfigured with security best practices and regularly updated security benchmarks that guide users in setting up and maintaining a secure environment.

By providing these comprehensive security features, OpenShift addresses the complex security challenges faced by enterprises today, ensuring that their deployments are secure by design, compliant with industry standards, and capable of withstanding modern cybersecurity threats. This security-first approach is integral to maintaining trust and integrity in enterprise applications and data.Add text here (Body0).

7.2 Designing for security

Red Hat OpenShift provides a methodology to build applications as a set of independent containerized microservices, containing their dependencies and running on a shared host in a multi-tenant fashion. In this regard, the paradigm of application development is experiencing a shift from monolithic application to modern containerized applications, with the latter allowing a faster release cycle and the possibility to upscale and downscale microservices according to business needs – for example a sudden increase in customer demand. The operational advantages that OpenShift adoption brings to enterprises are scalability, flexibility and maintainability.

As the application modernization effort is increasingly embraced by organizations, the decision of the preferred container platform inevitably needs to take into considerations security as a main requirement. In fact, a microservices architecture also introduces security challenges that need to be considered by security teams in the enterprise workforce. This chapter aims to highlight the unique security proposition of Red Hat which positions the container platform as the leading in the enterprise dimension.

The major security challenges linked with distributed environments are the following:

Complexity and visibility Microservices challenge security due to the distributed nature of its building blocks. As containers are independent and built on various frameworks (e.g., different language, different libraries), the security challenges require a preventive strategy to monitor containers.

Communication Microservices communicate with each other via APIs, increasing the attack surface. Encryption of data and authentication are measures that need to address the issue effectively.

Access control Access needs to be monitored granularly and specific policies are required to guarantee a balance between smooth development workflow and highly secure environment.

Beginning with the Operating System layer, this section will then explore the Compute layer, specifically focusing on the IBM Power server, to emphasize the security features integrated into its hardware design. Before delving into more detailed discussions, we will also introduce the Network and Storage layers, highlighting how the OpenShift platform provides strategies to address the challenges mentioned earlier.

7.2.1 Operating system layer

Red Hat OpenShift Container Platform is secure by design, and a significant contribution is due to its operating system. The host OS addresses the security challenges (complexity and visibility, communication and access control) by securing the host from container vulnerabilities as well as isolating containers from each other.

Red Hat OpenShift Container Platform leverages Red Hat CoreOS, a container-oriented operating system which implements the Security Enhanced Linux (SELinux) kernel to achieve container isolation and supports access control policies. CoreOS includes:

- Ignition: first boot system configuration responsible for starting and configuring machines
- CRI-O: container runtime integrating with the OS, responsible for running, stopping and restarting containers (it replaces the Docker Container Engine)
- Kubelet: node agent responsible for monitoring containers

An additional security measure is implemented by namespaces, which enable to abstract the resources consumed, including the OS, so that the running container appear as running its own OS in the effort to limit the attack surface and prevent that vulnerabilities contaminate other containers. Compromised containers are a vector for the host OS and for other containers not running SELinux, therefore with the control groups, the administrator can set a limitation to the resources that a collection of containers can consume from the host. Secure computing profiles can be defined to limit the system calls available to a collection of containers.

Ultimately, SELinux isolates namespaces, control groups and secure computing nodes.

As a first good practice to secure a multi-tenant environment is to start designing a container with the least privileges (described in 7.4.1, “Privileges” on page 204) possible. As it will be illustrated in section 7.3, “Securing your container environment” on page 198, the administrator can (and should) apply mandatory access control for every user and application while making sure that control groups limit the resources that containers consume from the host.

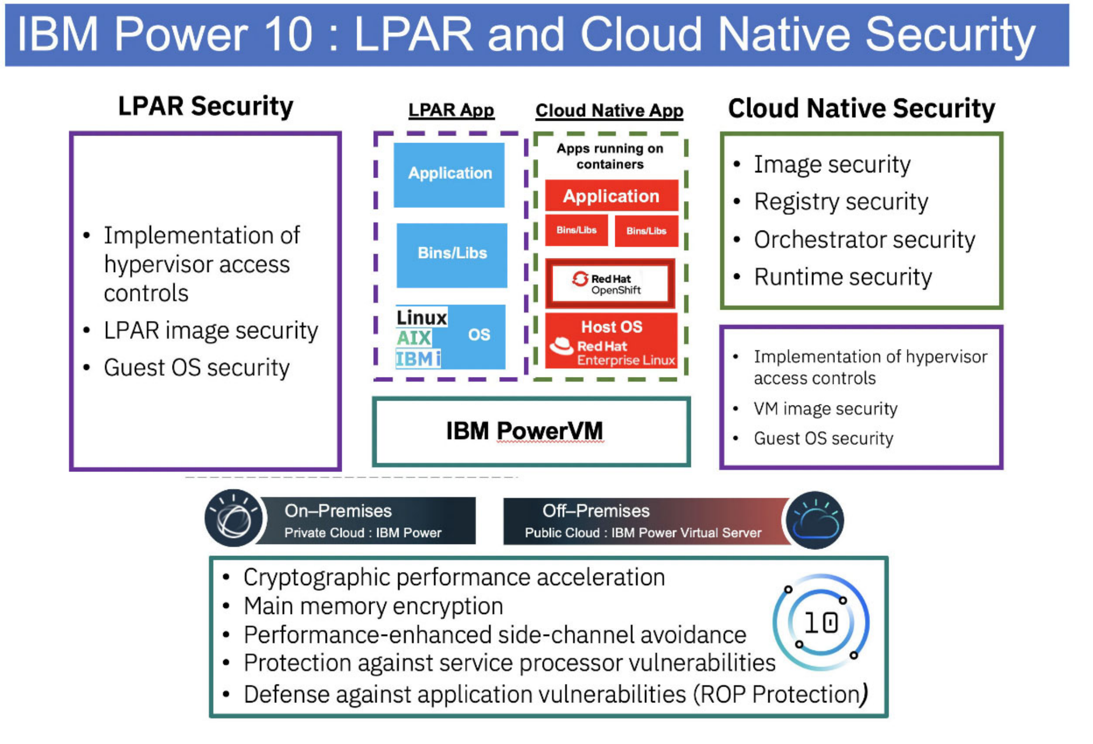
7.2.2 Compute layer

The combination of Red Hat OpenShift and IBM Power servers is synergistic from a scalability point of view, as it can achieve more than 3x container density compared to x86 based servers. Even more importantly is the advantage that IBM Power provides from a security standpoint. IBM Power10 encryption of data at rest and in motion are clear assets to securing cloud native applications in a hybrid-cloud environment.

This is illustrated in Figure 7-3.

Importantly, IBM Power10 has in-core hardware that protects against the Return-Oriented Programming (ROP) cyberattacks with incredibly limited performance overhead (1-2%). ROP attacks are difficult to identify and contain, as they are based on collecting and reusing existing code from memory (also known as “gadgets”), rather than injecting new code in the system. In fact, hackers chain the commands already existing in the memory to perform malicious actions.

IBM Power10 isolates the Baseboard Management Controller (BMC), the micro-controller embedded on the motherboard responsible to control remote management capabilities, and implements allowlist and blocklist approaches to limit the CPU resources that the BMC can access.



IBM Power 10 : LPAR and Cloud Native Security

LPAR Security

- Implementation of hypervisor access controls
- LPAR image security
- Guest OS security

LPAR App

Application

Bins/Libs

Linux
AIX
IBM i
OS

Cloud Native App

Apps running on containers

Application

Bins/Libs

Bins/Libs

Red Hat OpenShift

Host OS
Red Hat Enterprise Linux

IBM PowerVM

Cloud Native Security

- Image security
- Registry security
- Orchestrator security
- Runtime security

- Implementation of hypervisor access controls
- VM image security
- Guest OS security

On–Premises
Private Cloud : IBM Power

Off–Premises
Public Cloud : IBM Power Virtual Server

- Cryptographic performance acceleration
- Main memory encryption
- Performance-enhanced side-channel avoidance
- Protection against service processor vulnerabilities
- Defense against application vulnerabilities (ROP Protection)

Figure 7-3 BM Power10 security for LPARs and Cloud Native applications

For additional information please refer to section 1.4, “Architecture and implementation layers” on page 9.

7.2.3 Network layer

When working with containerized applications distributed across multiple hosts and nodes, the network becomes crucial in securing communication, which has been referred to as one of the three main challenges above. This section aims to introduce the role of networking in OpenShift Container Platform so that in the later section “Network isolation and API endpoint security” the components will be put into context.

Red Hat OpenShift comes with Red Hat Single Sign-On (SSO) which acts as a API authentication and authorization measure to secure platform endpoints.

As previously mentioned, Kubernetes clusters are composed of at least one master node (preferably more for redundancy purposes) and multiple worker nodes, which are virtual or physical machines on top of which containers run. Each node has an IP address and containerized applications are deployed on these nodes as pods. Each pod is also identified with a unique IP address and this results in network management ease, as the pod can be treated as a physical host or VM in terms of port allocation, naming and load balancing.

The apparent complexity of communication in this distributed infrastructure architecture is solved by the implementation of virtual networking. In fact, each container in a pod shares network ports, facilitating the port allocation task. While containers in a pod communicate with each other via localhost, the communication with containers belonging to different pods requires coordination of networking resources. The connectivity is implemented by the Red Hat Software-Defined Networking (SDN).

The Red Hat SDN utilizes Open vSwitch to manage network traffic and resources as software, allowing policy-based management. SDN controllers satisfy applications requests by managing networking devices and routing the data packages to their destination.

The network components in a cluster are managed by a Cluster Network Operator (CNO), which runs in turn on an OpenShift cluster.

Leveraging Single Root I/O Virtualization (SR-IOV) on IBM Power servers, the network design becomes more flexible.

Before moving to storage, another functional aspect of OpenShift is Network Files System (NFS), which is the method used to share files across clusters over the network. While NFS is an excellent solution for many environments, understanding the workload requirements of an application is important when selecting NFS based storage solutions.

7.2.4 Storage layer

Storage considerations in respect of Red Hat OpenShift workload are derived from a preliminary intrinsic condition of containers: whereas monolithic applications reserve storage resources in a static fashion, containers require more flexibility underneath as they need to be agile, easy to manage and quickly movable between environments.

The storage layer section aims to address the first of the challenges mentioned above: complexity and visibility.

When a container is created, a transient layer handling all read/write data is present within it. However, when the container stops running, this ephemeral layer is lost. Certainly, according to the nature of the container, administrators decide to assign either volumes (bound to the lifetime of the pod), or persistent volumes (persisting longer than the lifetime of the pod).

The dynamic provisioning requirements which benefit from microservices architecture is facilitated by the Container Storage Interface (CSI) which allows a vendor neutral management of file and block storage. Using a CSI API allows to:

- provision or deprovision a determined volume
- attach or detach volume from a node
- mount or unmount volume from a node
- consume block and mountable volumes
- create or delete a snapshot
- provision a volume from a snapshots

With the Red Hat OpenShift Platform Plus plan, the enterprise can leverage Red Hat Data Foundation, a software-defined storage orchestration platform for container environments. The data fabric capabilities of the OpenShift Data Platform are derived from the combination of Red Hat Ceph (software-defined storage platform), Rook.io (storage operator) and NooBaa (storage gateway). OpenShift Data Platform can be deployed as internal storage cluster or external storage cluster and it utilizes CSI to serve storage to the OpenShift Container Platform pods. The capabilities provided allow to manage block, file and object storage to serve databases, CI/CD tools and S3 API endpoints to the nodes.

Once clarified the contextual framework of the storage layer in OpenShift, here are reported the security measures that Red Hat Caph enforces to address threat and vulnerability management, encryption and identity and access management.

- Maintaining upstream relationships and community involvement to help focus on security from the start.
- Selecting and configuring packages based on their security and performance track records.
- Building binaries from associated source code (instead of simply accepting upstream builds).
- Applying a suite of inspection and quality assurance tools to prevent an extensive array of potential security issues and regressions.
- Digitally signing all released packages and distributing them through cryptographically authenticated distribution channels.
- Providing a single, unified mechanism for distributing patches and updates.

For additional information, refer to this Red Hat documentation article.

7.3 Securing your container environment

Red Hat OpenShift provides the container environment with capabilities that are missing in K8s distribution. The following sections of this chapter aim to describe the security features that position OpenShift as the enterprise choice for containerized workloads.

Source trusting

When pulling code from a Github repository, the first consideration should be whether on not you can trust the third-party developer. Inevitably developers might overlook at vulnerabilities of libraries or other dependencies used in the code, therefore it is recommended to conduct a proper due diligence before deploying a container in your enterprise environment.

To mitigate the risk, Red Hat provides Quay, a security focused container image registry which is included in the Red Hat OpenShift Platform Plus.

On the other hand, if it is preferred to scan for vulnerabilities with different tools, administrators should be aware of the possibility to integrate with OpenShift scanners such as OpenSCAP, BlackDuck Hub, JFrog Xray and Twislock.

Protecting software build process

Source to Image (S2I) provides a framework to integrate the code with the runtime libraries and other dependencies. The best practice is to integrate automated security scanning tools in the CI/CD pipeline of enterprises. RESTful APIs, in this context, allow the integration of the workflow with Static Application Security Testing (SAST) or Dynamic Application Security Testing (DAST) tools such as IBM AppScan or HCL AppScan.

Deployments on cluster

It is recommended to leverage automated policy-based tools to deploy containers in production environments. In this regard, Security Context Constraints (SCCs), packaged in Red Hat OpenShift Container Platform (extensively discussed in 8.3.1), support administrators in securing sensitive information by allowing/denying access to volumes, accept/deny privileges and extending/limiting capabilities that a container requires.

Orchestrating securely

Red Hat OpenShift extends K8s capabilities in terms of secure containers orchestration by:

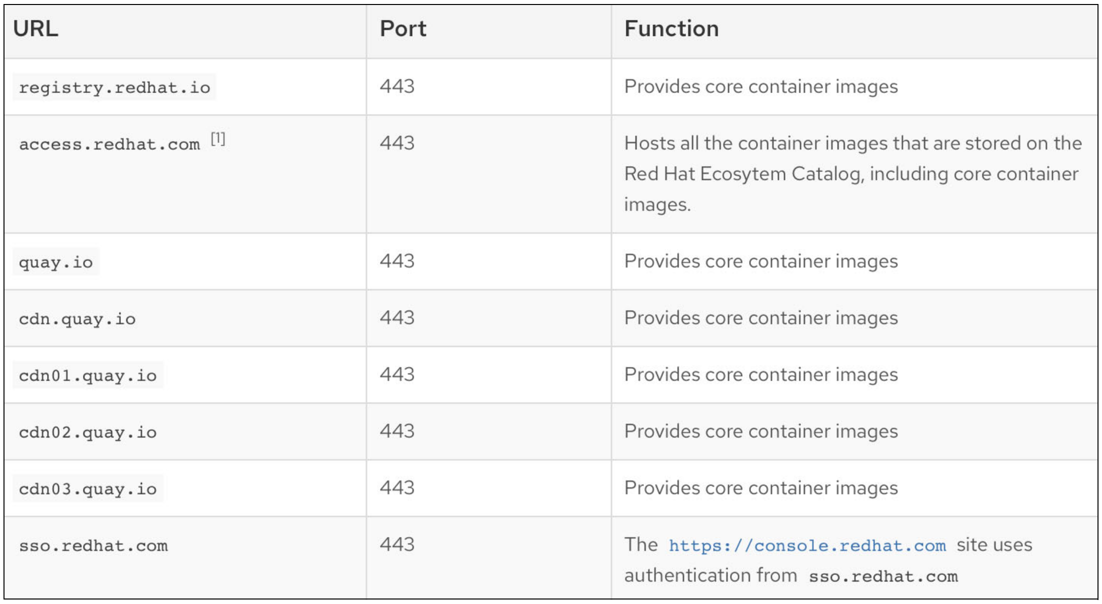
- Handling access to the master node via Transport Layer Security (TLS), which ensures that the data over the internet are encrypted
- Ensuring that the API server access is based on X.509 certificates or OAuth access tokens
- Avoiding the exposure of etcd (open source key-value store database for critical data) to the cluster
- SELinux

Network isolation and API endpoint security

The SDN previously introduced facilitates the management and visibility over the complex distributed workload. In fact, it is possible to control the outbound traffic of data out of the cluster and further network control is implemented via router or firewall to have a view on which IP addresses are allow/deny access to all others.

Moreover, Red Hat Single Sign-On (SSO), API authentication and authorization service, features client adapters for Red Hat JBoss, a Node.js and Lightweight Directory Access Protocol (LDAP)-based directory services. An API management tool advised in this context is Red Hat 3scale API management.

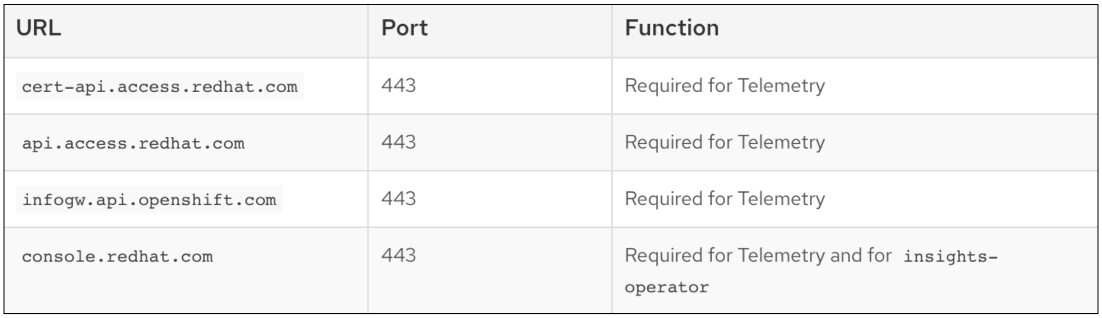
To configure a firewall for OpenShift Container Platform 4.12, it is required to define the sites that OCP requires so that the firewall grants access to those. As a first step it is recommended to create an allowlist containing the URLs in Figure 7-4 on page 200. Obviously, if a specific framework requires additional resources, this is the step at which it is recommended to include them.



URL	Port	Function
registry.redhat.io	443	Provides core container images
access.redhat.com [1]	443	Hosts all the container images that are stored on the Red Hat Ecosytem Catalog, including core container images.
quay.io	443	Provides core container images
cdn.quay.io	443	Provides core container images
cdn01.quay.io	443	Provides core container images
cdn02.quay.io	443	Provides core container images
cdn03.quay.io	443	Provides core container images
sso.redhat.com	443	The https://console.redhat.com site uses authentication from sso.redhat.com

Figure 7-4 Allowlist for OCP firewall

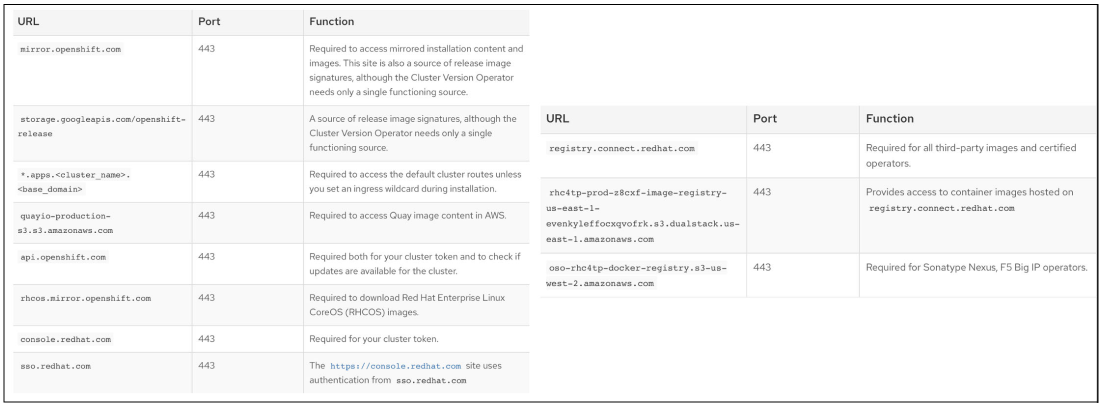
If you wish to use Telemetry to monitor the health, security and performance of application components, the URLs shown in Figure 7-5 in order to access Red Hat Insights.



URL	Port	Function
cert-api.access.redhat.com	443	Required for Telemetry
api.access.redhat.com	443	Required for Telemetry
infogw.api.openshift.com	443	Required for Telemetry
console.redhat.com	443	Required for Telemetry and for insights-operator

Figure 7-5 Telemetry URLs to allow

If the environment extends to Alibaba, AWS, GCP or Azure to host the cluster, it will be necessary to grant access to the provider API and DNS for the specific cloud. An example of this is shown in Figure 7-6.



URL	Port	Function
mirror.openshift.com	443	Required to access mirrored installation content and images. This site is also a source of release image signatures, although the Cluster Version Operator needs only a single functioning source.
storage.googleapis.com/openshift-release	443	A source of release image signatures, although the Cluster Version Operator needs only a single functioning source.
*.apps.<cluster_name>.<base_domain>	443	Required to access the default cluster routes unless you set an ingress wildcard during installation.
quayio-production-s3.s3.amazonaws.com	443	Required to access Quay image content in AWS.
api.openshift.com	443	Required both for your cluster token and to check if updates are available for the cluster.
rhcos.mirror.openshift.com	443	Required to download Red Hat Enterprise Linux CoreOS (RHCOS) images.
console.redhat.com	443	Required for your cluster token.
sso.redhat.com	443	The https://console.redhat.com site uses authentication from sso.redhat.com

URL	Port	Function
registry.connect.redhat.com	443	Required for all third-party images and certified operators.
rhc4tp-prod-z8cxf-image-registry-us-east-1-evenkyleffocxqvofrk.s3.dualstack.us-east-1.amazonaws.com	443	Provides access to container images hosted on registry.connect.redhat.com
oso-rhc4tp-docker-registry.s3-us-west-2.amazonaws.com	443	Required for Sonatype Nexus, F5 Big IP operators.

Figure 7-6 Cloud connections to allow

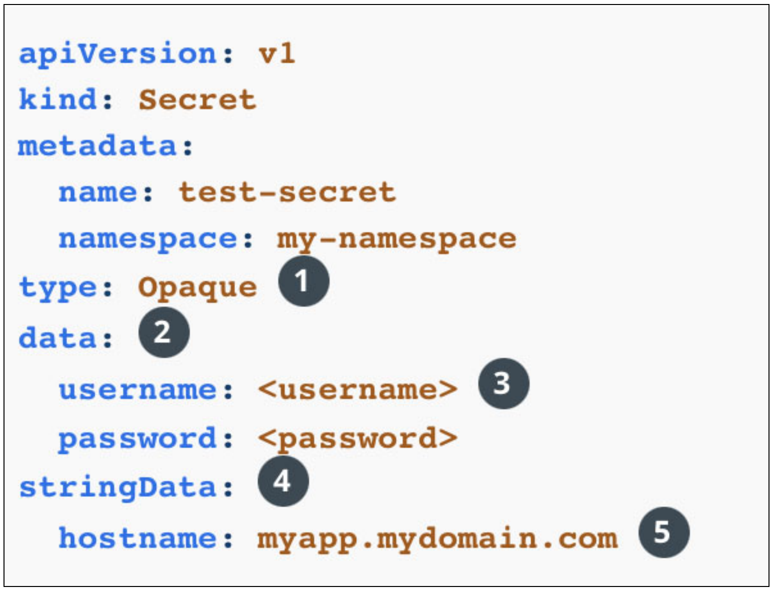
If the preferred option is the default Red Hat Network Time Protocol (NTP) server, *rhel.pool.ntp.org* is also required.

For additional information please consult:
https://docs.openshift.com/container-platform/4.12/installing/install_config/configuring-firewall.html.

Federation of containerized applications

Federations are deployment models composed of different meshes that are managed by different administrative domains. Federated namespaces create namespaces in the federation control plane so that the pods have consistent port ranges and IP addresses assigned. Federations can share services and workloads while ensuring extensive security via “secrets”. A secret is an object that holds sensitive information (e.g., passwords) decoupling them from the pod. Secret data volumes can be shared within a namespace and are never at rest on a node.

Figure 7-7, shows an example of a YAML definition of a secret object type and describes some of the contents.



```
apiVersion: v1
kind: Secret
metadata:
  name: test-secret
  namespace: my-namespace
type: Opaque
data:
  username: <username>
  password: <password>
stringData:
  hostname: myapp.mydomain.com
```

Figure 7-7 YAML file to describe secret

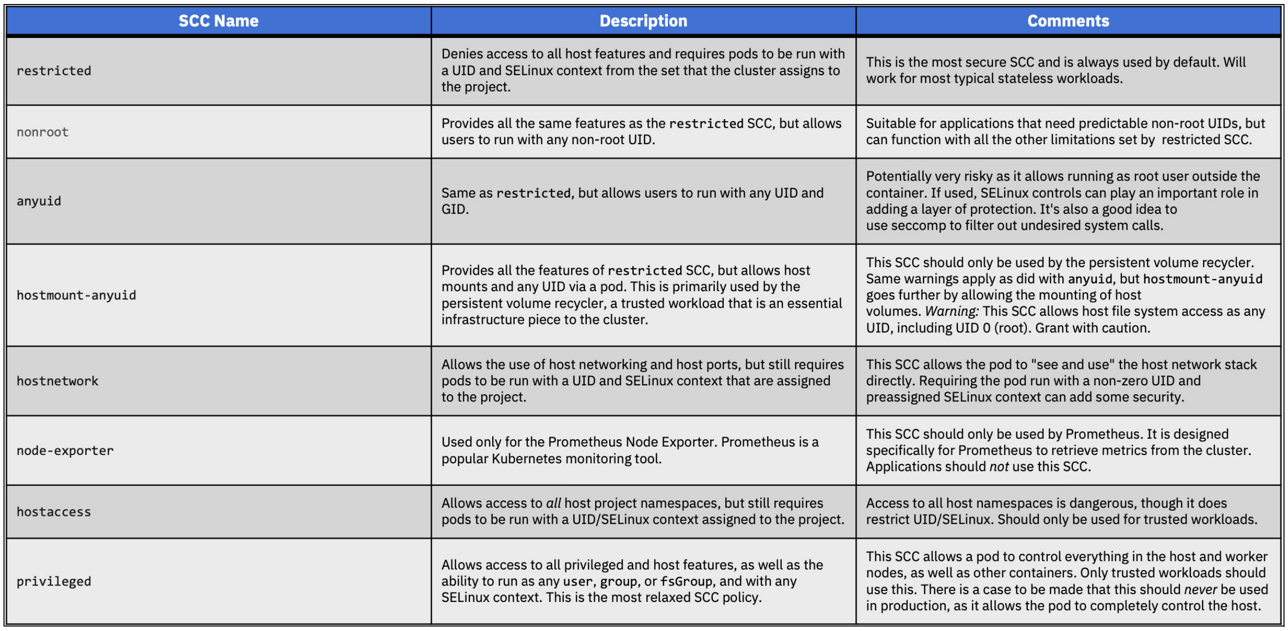
1. Indicates the structure of the secret (in this case, opaque identifies a key-value pair)
2. The format for the keys in “data” must meet the guidelines for DNS_SUBDOMAIN of the K8s glossary. More information found at this link: https://github.com/kubernetes/kubernetes/blob/v1.0.0/docs/design/identifiers.md
3. Values associated with the keys in “data” must be base64 converted
4. Entries in “stringdata” are converted to base64 and will be moved to “data” automatically
5. Plain text strings associated with the “stringdata” key

7.4 Security contexts and security context constraints

Access control to any shared computing environment, VMs and containers, is an essential task for the Chief Security Officer’s (CSO) team. Red Hat OpenShift provides security contexts (SCs) and security context constraints (SCCs) to help manage security in the container environment.

Security contexts and security context constraints are required for a container to configure access to protected Linux operating system functions on an OpenShift Container Platform cluster. While SCs are defined by the development team, SCCs are determined by cluster administrators. An application's security context specifies the permissions that the application needs, whereas the cluster's security context constraints specify the permissions that the cluster allows. An SC with an SCC enables an application to request access while limiting the access that the cluster will grant.

An example of SCC and SC implementation is shown in Figure 7-8.



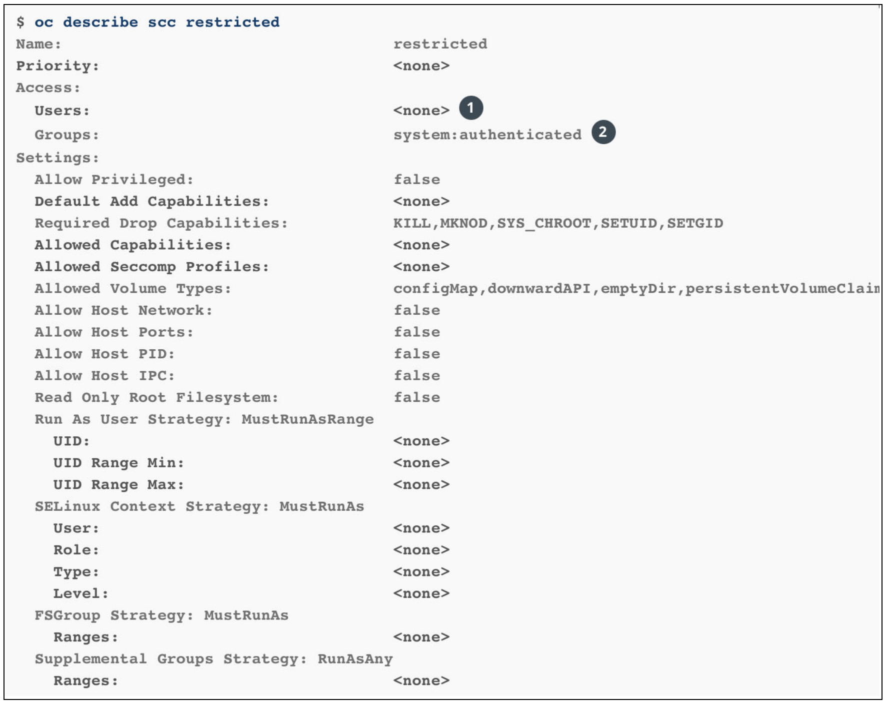
SCC Name	Description	Comments
`restricted`	Denies access to all host features and requires pods to be run with a UID and SELinux context from the set that the cluster assigns to the project.	This is the most secure SCC and is always used by default. Will work for most typical stateless workloads.
`nonroot`	Provides all the same features as the `restricted` SCC, but allows users to run with any non-root UID.	Suitable for applications that need predictable non-root UIDs, but can function with all the other limitations set by restricted SCC.
`anyuid`	Same as `restricted`, but allows users to run with any UID and GID.	Potentially very risky as it allows running as root user outside the container. If used, SELinux controls can play an important role in adding a layer of protection. It's also a good idea to use seccomp to filter out undesired system calls.
`hostmount-anyuid`	Provides all the features of `restricted` SCC, but allows host mounts and any UID via a pod. This is primarily used by the persistent volume recycler, a trusted workload that is an essential infrastructure piece to the cluster.	This SCC should only be used by the persistent volume recycler. Same warnings apply as did with `anyuid`, but `hostmount-anyuid` goes further by allowing the mounting of host volumes. *Warning:* This SCC allows host file system access as any UID, including UID 0 (root). Grant with caution.
`hostnetwork`	Allows the use of host networking and host ports, but still requires pods to be run with a UID and SELinux context that are assigned to the project.	This SCC allows the pod to "see and use" the host network stack directly. Requiring the pod run with a non-zero UID and preassigned SELinux context can add some security.
`node-exporter`	Used only for the Prometheus Node Exporter. Prometheus is a popular Kubernetes monitoring tool.	This SCC should only be used by Prometheus. It is designed specifically for Prometheus to retrieve metrics from the cluster. Applications should *not* use this SCC.
`hostaccess`	Allows access to *all* host project namespaces, but still requires pods to be run with a UID/SELinux context assigned to the project.	Access to all host namespaces is dangerous, though it does restrict UID/SELinux. Should only be used for trusted workloads.
`privileged`	Allows access to all privileged and host features, as well as the ability to run as any `user`, `group`, or `fsGroup`, and with any SELinux context. This is the most relaxed SCC policy.	This SCC allows a pod to control everything in the host and worker nodes, as well as other containers. Only trusted workloads should use this. There is a case to be made that this should *never* be used in production, as it allows the pod to completely control the host.

Figure 7-8 SCC and SC implementation

By default, OpenShift prevents the containers running in a cluster from accessing protected functions. These functions – Linux features such as shared file systems, root access, and some core capabilities such as the KILL command – can affect other containers running in the same Linux kernel, so the cluster limits access to them. Most cloud-native applications work fine with these limitations, but some (especially stateful workloads) need greater access. Applications that need these functions can still use them, but they need the cluster's permission.

SCs are defined as a YAML file within the pod that attempts to deploy the application into production. SCCs determine which Linux functions a pod can request for its application. The pod requesting access to specific functions via SCs will fail to launch unless SCCs give permission to proceed.

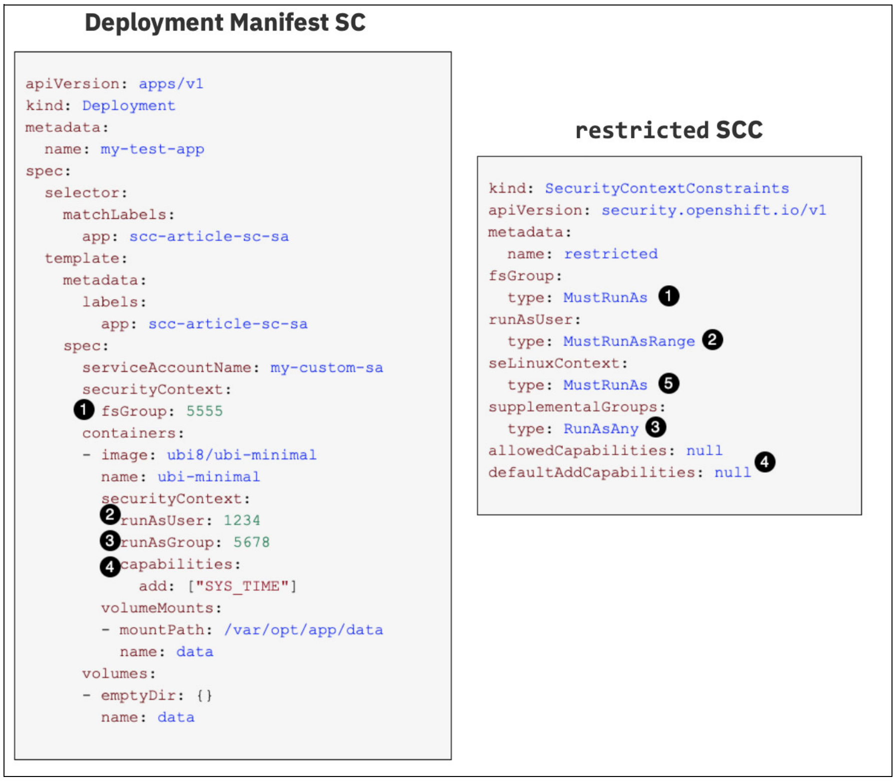
A list of predefined (default) SCCs is shown in Figure 7-9.



```
$ oc describe scc restricted
Name:                                   restricted
Priority:                               <none>
Access:
  Users:                                <none> ❶
  Groups:                               system:authenticated ❷
Settings:
  Allow Privileged:                     false
  Default Add Capabilities:             <none>
  Required Drop Capabilities:           KILL,MKNOD,SYS_CHROOT,SETUID,SETGID
  Allowed Capabilities:                 <none>
  Allowed Seccomp Profiles:             <none>
  Allowed Volume Types:                 configMap,downwardAPI,emptyDir,persistentVolumeClai
  Allow Host Network:                   false
  Allow Host Ports:                     false
  Allow Host PID:                       false
  Allow Host IPC:                       false
  Read Only Root Filesystem:            false
  Run As User Strategy: MustRunAsRange
    UID:                                <none>
    UID Range Min:                      <none>
    UID Range Max:                      <none>
  SELinux Context Strategy: MustRunAs
    User:                               <none>
    Role:                               <none>
    Type:                               <none>
    Level:                              <none>
  FSGroup Strategy: MustRunAs
    Ranges:                             <none>
  Supplemental Groups Strategy: RunAsAny
    Ranges:                             <none>
```

Figure 7-9 Predefined SCCs

Taking a closer look to one of the SCCs illustrated above, the object would look like the one represented in Figure 7-10.



Deployment Manifest SC

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: my-test-app
spec:
  selector:
    matchLabels:
      app: scc-article-sc-sa
  template:
    metadata:
      labels:
        app: scc-article-sc-sa
    spec:
      serviceAccountName: my-custom-sa
      securityContext:
     ❶ fsGroup: 5555
      containers:
      - image: ubi8/ubi-minimal
        name: ubi-minimal
        securityContext:
       ❷runAsUser: 1234
       ❸runAsGroup: 5678
       ❹capabilities:
            add: ["SYS_TIME"]
        volumeMounts:
        - mountPath: /var/opt/app/data
          name: data
      volumes:
      - emptyDir: {}
        name: data
```

restricted SCC

```
kind: SecurityContextConstraints
apiVersion: security.openshift.io/v1
metadata:
  name: restricted
fsGroup:
  type: MustRunAs ❶
runAsUser:
  type: MustRunAsRange ❷
seLinuxContext:
  type: MustRunAs ❺
supplementalGroups:
  type: RunAsAny ❸
allowedCapabilities: null
defaultAddCapabilities: null ❹
```

Figure 7-10 Description of the “restricted” SCC

Note that field designated by the number 1 in Figure 7-10 on page 203 represents the users to which the SCC “restricted” applies to, whereas field 2 identifies the group to which the SCC “restricted” applies.

It is highly recommended to avoid modifying default SCCs, however, administrators can create new and customized SCCs that better fit specific requirements and policies in the organizational processes. An example of how a new SCC can be created is shown in section 7.4.2, “Access controls” on page 204 and in Example 7-4 on page 205.

The following sections discuss protected Linux functions such as privileges, access controls and capabilities.

7.4.1 Privileges

Privileges describe the authority of a determined pod and the containerized applications running within it. There are two places that privileges can be assigned – either in the SC when privilege is set equal to true in the SC request, or set in the SCC where privilege is set to true. This is shown in Example 7-1.

Example 7-1 Privileged container settings in an SCC

```
allowPrivilegedContainer
allowPrivilegedEscalation
```

In Example 7-1, the first line indicates that the container will run with specified privileges, whereas the second line grants the possibility for a pod derived by the parent pod to be allowed to run with additional privileges than the parent pod.

The request for privileges in an SC is shown in Example 7-2. It is worth noting that when privileges are requested from an SCs perspective, the developer needs to only request for privileges, whereas from the SCCs perspective, the administrator is expected to be specific about the set of privileges that are allowed.

Example 7-2 Request for privileges in an SC

```
securityContext.privileged: true
```

It is good practice to keep in mind that privileged pods might endanger the host and other containers, therefore only well-trusted processes should be allowed privileges.

7.4.2 Access controls

Administrators define “access control” within SCCs to manage user IDs and group IDs accessing pods. Example 7-3 highlights some examples.

Example 7-3 Examples of access controls

```
1. runAsUser: the command specifies the range of user IDs that are allowed to run
within a pod, accessing all its respective containers
2. supplementalGroups: determines the group IDs allowed to run all containers
within a pod
3. fsGroup: entitles a range of group IDs to control the pod storage volumes
4. seLinuxContext: the command specifies the values for setting SELinux user,
role, type and level
5. mustRunAs: assigned to fields 1:4 and enforces the range of user IDs that a
container can request
```

```
6. mustRunAsRange: assigned to fields 1:4 and enforces the range of user IDs that
a container can request
7. RunaAsAny: this command allows a user ID to be requested by a pod even if its
ID is not within a specified range in SCCs (also referred to as root ID). Careful
using it because it inevitably creates a privileged access which needs to be
cautiously assigned
8. MustRunAsNonRoot: ensures that any non-root IDs need to be specified
```

As previously illustrated, the correct syntax for the development team to include these requests is:

`securityContext.field`

Where **`field`** is any of the fields in Example 7-3 on page 204.

Once the request is made, it will be processed and validated against the cluster SCCs.

Example 7-4 shows how a new SCC would look, integrating the fields listed in Example 7-3 on page 204.

Example 7-4 Example SCC

```
kind: SecurityContextConstraints
apiVersion: v1
metadata:
  name: scc-admin
allowPrivilegedContainer: true
runAsUser:
  type: RunAsAny
seLinuxContext:
  type: RunAsAny
fsGroup:
  type: RunAsAny
supplementalGroups:
  type: RunAsAny
users:
- my-admin-user
groups:
- my-admin-group
```

7.4.3 Capabilities

Some capabilities, specifically Linux OS capabilities, take precedence over the pod's settings. A list of these capabilities can be found in this document. For completeness, Example 7-5 shows some of the most popular ones.

Example 7-5 Capabilities that are override a pod's settings

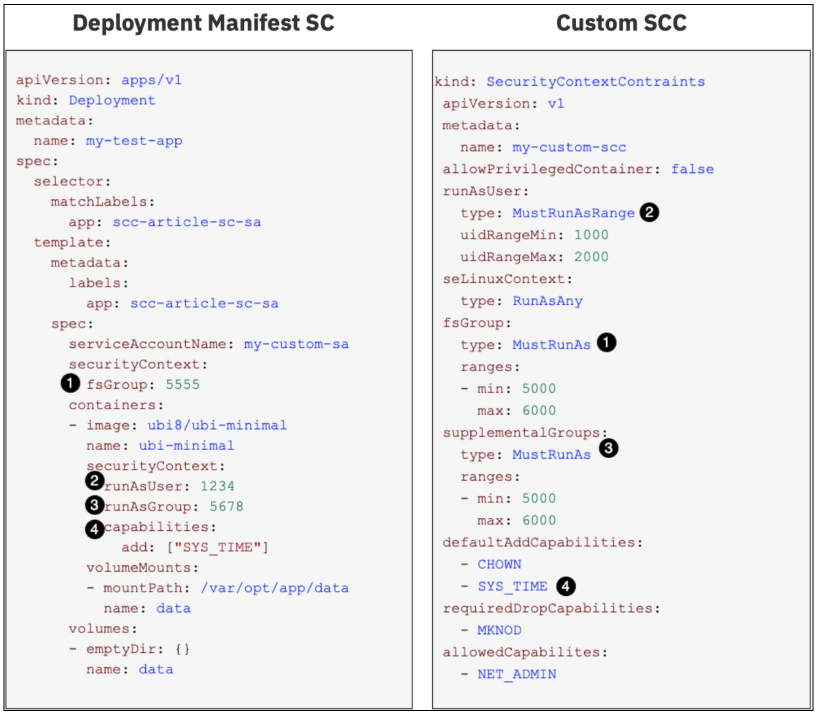
```
chown
kill
setcap
```

7.4.4 Deployment examples

The purpose of this final section on SCs and SCCs is to combine the pieces described so far, illustrating a deployment scenario in which an SC manifest is validated, first against a default “restricted” SCC (Figure 7-11) and then against a customized SCC (Figure 7-12 on page 207).

In Figure 7-11, the SC fails to pass due to three critical issues shown as points 1, 2 and 4.

- First, in the attempt to control the pod storage volumes, the SC requests *fsGroup* 5555. The reason this fails is that SCC “restricted” does not specify a range for *fsGroup*, therefore the default range is used (*1000000000-1000009999*) excluding the requested *fsGroup* 5555. (1)
- Secondly, the SC is asking permission to *runAsUser* 1234. However the SCC “*restricted*” once again takes into consideration the default range (*1000000000-1000009999*)so the request is failed as not within the range. (2)
- Finally, the deployment manifest requests the capability “*SYS_TIME*” (gives ability to manipulate the system clock). This request fails as the SCC does not specify “*SYS_TIME*” either in “*allowedCapabilities*” nor is it included in “*defaultAddCapabilities*” (4). The only requests that passes is (3). The SC requests to *runAsGroup* 5678 and this is allowed by the *runAsAny* field of the “restricted” SCC.
- As a final remark, (5) is a note to highlight that the container is assigned to the project default context value as the *seLinuxContext* is set as *MustRunAs* but lacks the specific context.



Deployment Manifest SC

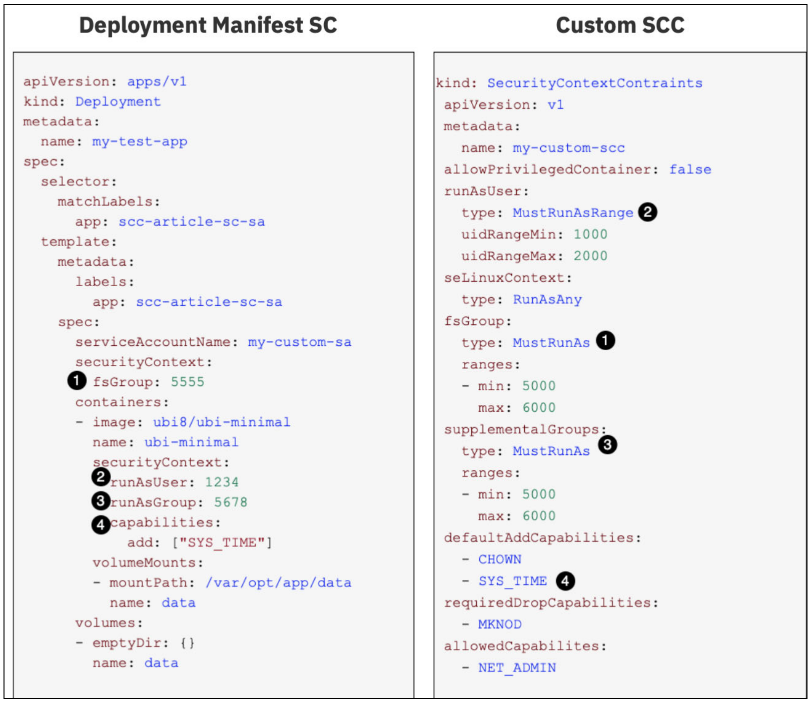
```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: my-test-app
spec:
  selector:
    matchLabels:
      app: scc-article-sc-sa
  template:
    metadata:
      labels:
        app: scc-article-sc-sa
    spec:
      serviceAccountName: my-custom-sa
      securityContext:
    ❶ fsGroup: 5555
      containers:
      - image: ubi8/ubi-minimal
        name: ubi-minimal
        securityContext:
       ❷runAsUser: 1234
       ❸runAsGroup: 5678
       ❹capabilities:
            add: ["SYS_TIME"]
        volumeMounts:
        - mountPath: /var/opt/app/data
          name: data
      volumes:
      - emptyDir: {}
        name: data
```

Custom SCC

```
kind: SecurityContextContraints
 apiVersion: v1
 metadata:
   name: my-custom-scc
 allowPrivilegedContainer: false
 runAsUser:
   type: MustRunAsRange ❷
   uidRangeMin: 1000
   uidRangeMax: 2000
 seLinuxContext:
   type: RunAsAny
 fsGroup:
   type: MustRunAs ❶
   ranges:
   - min: 5000
     max: 6000
 supplementalGroups:
   type: MustRunAs ❸
   ranges:
   - min: 5000
     max: 6000
 defaultAddCapabilities:
   - CHOWN
   - SYS_TIME ❹
 requiredDropCapabilities:
   - MKNOD
 allowedCapabilites:
   - NET_ADMIN
```

Figure 7-11 Validating an SC against the “restricted” SCC

Figure 7-12 shows how the SC request can be satisfied when against a customized SCC "*my-custom-scc*".



Deployment Manifest SC

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: my-test-app
spec:
  selector:
    matchLabels:
      app: scc-article-sc-sa
  template:
    metadata:
      labels:
        app: scc-article-sc-sa
    spec:
      serviceAccountName: my-custom-sa
      securityContext:
    ❶ fsGroup: 5555
      containers:
      - image: ubi8/ubi-minimal
        name: ubi-minimal
        securityContext:
       ❷runAsUser: 1234
       ❸runAsGroup: 5678
       ❹capabilities:
            add: ["SYS_TIME"]
        volumeMounts:
        - mountPath: /var/opt/app/data
          name: data
      volumes:
      - emptyDir: {}
        name: data
```

Custom SCC

```
kind: SecurityContextContraints
 apiVersion: v1
 metadata:
   name: my-custom-scc
 allowPrivilegedContainer: false
 runAsUser:
   type: MustRunAsRange ❷
   uidRangeMin: 1000
   uidRangeMax: 2000
 seLinuxContext:
   type: RunAsAny
 fsGroup:
   type: MustRunAs ❶
   ranges:
   - min: 5000
     max: 6000
 supplementalGroups:
   type: MustRunAs ❸
   ranges:
   - min: 5000
     max: 6000
 defaultAddCapabilities:
   - CHOWN
   - SYS_TIME ❹
 requiredDropCapabilities:
   - MKNOD
 allowedCapabilites:
   - NET_ADMIN
```

Figure 7-12 Validating an SC against a custom SCC

7.5 Monitoring and logging

Logging and monitoring enable earlier detection of vulnerabilities in Red Hat OpenShift Container Platform by providing essential context during active security incident investigations and postmortem analyses. They facilitate proactive monitoring of security-related activities and help confirm the effectiveness and integrity of the existing security configuration. This chapter delves into three primary areas: monitoring containers and Red Hat OpenShift Container Storage security, audit logs, and Red Hat OpenShift File Integrity Operator monitoring.

Monitoring a containerized environment involves tracking and measuring various key performance indicators (KPIs) to ensure optimal performance of decoupled applications, often within a microservices architecture. Effective monitoring helps maintain application health, performance, and security, and involves several critical aspects and challenges.

Monitoring containers in a dynamic and rapidly changing environment presents several challenges:

- **Rapid Provisioning and Termination**: Containers are designed to be provisioned and terminated quickly, making it difficult to track changes in environments with continuously fluctuating numbers of containers and instances. This rapid churn requires monitoring tools to be highly adaptive and capable of real-time tracking.
- **Ephemeral Nature of Containers**: Since containers are temporary, their metrics, logs, and other data disappear immediately upon termination. It is crucial to collect and store this data in a central location before the containers shut down. This necessitates a robust logging and monitoring infrastructure that can handle high data throughput and ensure data persistence.

- **Shared Resources**: Containers share resources such as memory, CPU, and operating systems, complicating the measurement of individual container performance. Resource contention and performance bottlenecks can arise, making it essential to have tools that can isolate and identify issues at a granular level.
- **Inadequacy of Traditional Monitoring Tools**: Many conventional monitoring tools are often insufficient for effectively monitoring containerized environments due to their inability to handle the dynamic and scalable nature of containers. Traditional tools may not provide the necessary visibility into container orchestration layers or the ephemeral nature of container lifecycles.

To address these challenges, several strategies can be employed:

- **Monitoring the Entire Stack**: Achieving full application visibility requires monitoring the entire stack, including containers, clusters, networking, and inter-container communications. This holistic approach ensures that all aspects of the application and infrastructure are observed, providing a complete picture of system health and performance.
- **Granular Visibility**: Multiple levels of granularity are required to get a comprehensive picture. Drilling down by degrees of granularity helps pinpoint the exact locations of issues. This involves monitoring at the node, pod, and container levels, as well as observing network traffic, storage I/O, and other critical metrics.
- **Contextualized Alerting**: In a containerized environment, alerts should include relevant context, as an issue in one container might be related to its interaction with another container. Contextualized alerts help in understanding the root cause of issues and in taking appropriate corrective actions.

The benefits of monitoring containers are substantial:

- **Problem Resolution**: Monitoring helps determine the cause of issues, facilitates their resolution, and allows for cataloging the “lessons learned” for future reference. This continuous improvement cycle helps in building more resilient and reliable applications.
- **Resource Usage Analysis**: Monitoring enables analysis of how containerized applications use cloud resources and assists in cost apportionment. By understanding resource utilization patterns, organizations can optimize their infrastructure and reduce costs.
- **Future Resource Planning**: Historical monitoring data helps organizations plan future computing resource requirements. This data-driven approach ensures that adequate resources are allocated to meet future demand, avoiding both under-provisioning and over-provisioning.

The Red Hat OpenShift Container Monitoring Platform addresses many of these monitoring challenges through a preconfigured, automatically updating stack based on Prometheus, Grafana, and Alertmanager. Key components of this platform include:

- **Prometheus**: Used as a backend to store time-series data, Prometheus is an open-source solution for cloud-native architecture monitoring. It offers powerful querying capabilities and a flexible data model, making it suitable for a wide range of monitoring scenarios.
- **Alertmanager**: Handles alarms and sends notifications. It integrates seamlessly with Prometheus, allowing for sophisticated alerting rules and notification mechanisms. Alertmanager supports multiple notification channels, including email, Slack, and PagerDuty, ensuring that alerts reach the right people at the right time.
- **Grafana**: Provides visual data representation through graphs. Grafana's rich visualization capabilities allow users to create dynamic and interactive dashboards, making it easier to interpret monitoring data and identify trends and anomalies.

The platform includes default alerts that notify administrators immediately about cluster issues. Default dashboards in the Red Hat OpenShift Container Platform web console offer visual representations of cluster metrics, aiding in quick understanding of cluster states. The “Observe” section of the web console allows access to metrics, alerts, monitoring dashboards, and metrics targets. Cluster administrators can optionally enable monitoring for user-defined projects, allowing for customized monitoring of services and pods. This flexibility ensures that different teams and projects can tailor monitoring to their specific needs.

As cloud-native applications continue to grow in scale and complexity, Application Performance Monitoring (APM) observability provides constant visibility into the health of the app and its infrastructure. APM observability is crucial for highly distributed and scalable cloud-native and hybrid apps, ensuring optimal performance and resiliency.

IBM Instana enhances the observability and APM functions provided by the default Red Hat OpenShift container monitoring tools. Instana is an automated system and APM service that visualizes performance through machine learning-generated graphs. It increases application performance and reliability through deep observability and applied intelligence. Instana excels in cloud-based microservices architectures, enabling development teams to iterate quickly and address issues before they impact customers. Instana provides several key capabilities:

- **Automatic Discovery and Instrumentation**: Instana automatically discovers applications and their dependencies, and instruments them without requiring manual intervention. This reduces the overhead associated with setting up monitoring and ensures that all components are monitored from the outset.
- **Real-Time Data Collection**: Instana collects data in real-time, providing immediate insights into application performance and health. This real-time visibility is critical for identifying and resolving issues before they affect users.
- **Machine Learning-Based Analytics**: Instana uses machine learning algorithms to analyze performance data and detect anomalies. This predictive capability helps in identifying potential issues early and taking preemptive action.
- **Comprehensive Dashboards**: Instana offers comprehensive dashboards that provide a unified view of application performance, infrastructure health, and user experience. These dashboards can be customized to meet the specific needs of different stakeholders, from developers to operations teams.

By integrating IBM Instana with Red Hat OpenShift, organizations can elevate their monitoring and observability capabilities, ensuring that their cloud-native applications remain perform-ant, resilient, and reliable.

In addition to monitoring containers and application performance, maintaining security and integrity within the Red Hat OpenShift environment involves leveraging audit logs and the Red Hat OpenShift File Integrity Operator.

Audit logs provide a detailed record of all activities and changes within the system. They are crucial for tracking user actions, detecting unauthorized access, and investigating security incidents. Effective audit logging helps in maintaining compliance with regulatory requirements and provides an audit trail that can be used for forensic analysis.

The Red Hat OpenShift File Integrity Operator enhances security by monitoring file integrity within the cluster. It detects unauthorized changes to critical system files, ensuring that the integrity of the operating environment is maintained. The File Integrity Operator works by periodically checking the hashes of monitored files and comparing them to known good values. Any discrepancies trigger alerts, allowing administrators to investigate and remediate potential security breaches.

7.6 Authorization and authentication

Users accessing the Red Hat OpenShift Container Platform must authenticate initially to the cluster. Authentication verifies the identity of the user making requests to the platform's API. Subsequently, the authorization layer evaluates the user's permissions to determine if the requested actions are permitted. Configuration of authentication settings within the platform is managed by the cluster administrator.

The authentication process in Red Hat OpenShift Container Platform involves multiple layers to ensure secure access to its resources. Users authenticate primarily through OAuth access tokens or X.509 client certificates. OAuth tokens are obtained via the platform's built-in OAuth server, which supports authentication flows such as Authorization Code Flow and Implicit Flow. The server integrates seamlessly with various identity providers, including LDAP, Keystone, GitHub, and Google, enabling organizations to leverage existing user management systems securely.

X.509 client certificates are utilized for HTTPS-based authentication, providing a robust mechanism for verifying the identity of clients interacting with the OpenShift API server. These certificates are verified against a trusted Certificate Authority (CA) bundle, ensuring the integrity and authenticity of client connections.

In OpenShift, users are classified into different categories based on their roles and responsibilities within the platform. Regular users are typically individuals who interact directly with applications and services deployed on OpenShift. System users, on the other hand, are automatically generated during the platform's setup and are associated with specific system-level tasks, such as managing cluster nodes or executing infrastructure-related operations.

Service accounts represent a specialized type of system user tailored for project-specific roles and permissions. These accounts enable automated processes within projects, ensuring that applications and services can securely access resources without compromising system integrity.

Groups play a pivotal role in managing authorization policies across OpenShift environments. Users can be organized into groups, facilitating streamlined assignment of permissions and simplifying the enforcement of access control policies. Alongside user-defined groups, OpenShift automatically provisions virtual groups, which include system-defined roles and default access configurations. This hierarchical group structure ensures efficient management of user permissions while adhering to organizational security policies and compliance requirements.

The internal OAuth server in OpenShift acts as a central authority for managing authentication and authorization workflows. It issues and validates OAuth tokens used by clients to authenticate API requests, ensuring that only authorized users and applications can access protected resources. Administrators can configure the OAuth server to integrate seamlessly with various identity providers, including htpasswd, Keystone, LDAP, and external OAuth providers like GitHub or Google. Each identity provider offers distinct authentication mechanisms, such as simple bind authentication for LDAP or OAuth 2.0 flows for external identity providers, enhancing flexibility and compatibility with diverse organizational environments.

Role-Based Access Control (RBAC) is fundamental to enforcing granular access control policies within OpenShift, allowing administrators to define fine-grained permissions through roles and role bindings. Roles specify a set of permissions (verbs) that dictate actions users can perform on specific API resources (objects). Role bindings associate these roles with individual users, groups, or service accounts, enabling administrators to implement the principle of least privilege effectively.

ClusterRoles extend RBAC capabilities by providing cluster-wide permissions that apply to all users within the platform. ClusterRoleBindings establish associations between ClusterRoles and subjects (users or groups), allowing administrators to manage permissions consistently across large-scale deployments.

Prometheus system metrics capture comprehensive data on authentication attempts within OpenShift, providing administrators with actionable insights into access patterns and security incidents. Metrics include counts of successful and failed login attempts across different authentication methods, such as password-based authentication via CLI or web console logins. Monitoring these metrics enables proactive detection of anomalous login activities, supporting timely mitigation of security threats and optimization of authentication workflows.

Administrators can configure and manage RBAC roles and role bindings using command-line interfaces (CLI) or graphical user interfaces (GUI) provided by OpenShift. Practical examples illustrate the steps for creating, modifying, and deleting roles and bindings, ensuring precise control over access permissions across diverse user populations and project environments. Role-based access control strategies empower organizations to align access policies with business requirements, enforcing security best practices while facilitating seamless collaboration and application deployment within OpenShift Container Platform.

7.7 Tools

There are multiple tools available to assist you in setting up and monitoring security in your OpenShift environment. This section describes some of them.

7.7.1 Aqua

This section delves into Aqua, a robust security tool designed explicitly for safeguarding workloads hosted on Red Hat OpenShift running on IBM Power servers. Developed by an IBM Business Partner, Aqua addresses the intricate security challenges inherent in cloud-native environments, spanning the entire lifecycle of containerized applications.

Aqua is positioned as a pivotal component in the Cloud-Native Application Protection Platform (CNAPP). This platform is designed to secure applications from development through to deployment and runtime on cloud-native architectures. It supports the shift from traditional software security models reliant on vendor-provided patches to a proactive, integrated security approach suited for DevOps environments.

Below is a list of Aqua features and capabilities

- **Image Scanning**: Aqua performs comprehensive vulnerability and malware scans on container images. This scanning process is integral to the CI/CD pipeline, ensuring that vulnerabilities are identified and mitigated early in the development cycle. Aqua utilizes both proprietary scanning engines and open-source tools to detect security issues, ensuring that only secure images are deployed into production.

- **Runtime Protection**: Once containers are deployed, Aqua provides robust runtime protection. This includes implementing network security policies, access controls, and process-level isolation to prevent unauthorized access, privilege escalation, and network-based attacks. This proactive approach minimizes security risks during application execution.
- **Compliance and Governance**: Aqua enables organizations to enforce compliance with regulatory standards and internal security policies. It offers detailed auditing and reporting capabilities, essential for demonstrating compliance and maintaining security posture across diverse environments and regulatory frameworks.
- **Centralized Management**: The Aqua platform provides a unified management interface via a web-based console and APIs. This centralized management facilitates seamless oversight and control across multiple Kubernetes clusters and namespaces, enhancing operational efficiency and security management at scale.
- **Secrets Management**: Aqua ensures the secure management of secrets, credentials, and sensitive data within container environments. It offers features such as secure storage, encryption, and fine-grained access controls to protect critical information from unauthorized access and breaches.

Aqua integrates seamlessly with Red Hat OpenShift on IBM Power by deploying an Aqua Enforcer container on each node within the cluster. These enforcers communicate with the Aqua Security Control Plane, enabling the enforcement of security policies and providing real-time visibility into the security status of the cluster. This integration augments native OpenShift security controls, enhancing overall security posture without compromising platform compatibility or performance.

Recognizing the trend towards hybrid and multi-cloud deployments, Aqua supports security management across diverse infrastructure environments. It enables organizations to maintain consistent security policies and compliance measures across on-premises data centers and public cloud platforms, thereby reducing the attack surface and mitigating risks associated with complex deployment landscapes.

7.7.2 Red Hat ACS

Red Hat Advanced Cluster Security for Kubernetes is a Kubernetes-native security platform that equips you to build, deploy, and run cloud-native applications with more security.

The solution helps protect containerized Kubernetes workloads in all major clouds and hybrid platforms, including Red Hat OpenShift, Amazon Elastic Kubernetes Service (EKS), Microsoft Azure Kubernetes Service (AKS), and Google Kubernetes Engine (GKE).

Red Hat Advanced Cluster Security for Kubernetes is included with Red Hat OpenShift Platform Plus, a complete set of powerful, optimized tools to secure, protect, and manage the applications. See the detailed informaion at the following link:

https://www.redhat.com/en/technologies/cloud-computing/openshift/advanced-cluster-security-kubernetes

A good feature in Red Hat ACS is that it works to prevent risky workloads from being deployed or running. Red Hat Advanced Cluster Security monitors, collects, and evaluates system-level events such as process execution, network connections and flows, and privilege escalation within each container in your Kubernetes environments. Combined with behavioral baselines and “allowlisting”, it detects anomalous activity indicative of malicious intent such as active malware, cryptomining, unauthorized credential access, intrusions, and lateral movement.

See the full features at the Red Hat ACS Data Sheet.

Certifications

This chapter discusses certifications security standard and its importance, as the security certifications provide an extra layer of validation on security best practices to help ensure data protection from being compromised and also to ensure compliance with evolving data privacy and business resiliency laws and regulations. Use this chapter as reference in order to minimize financial risk and avoid incurring steep fines for non-compliance.

This chapter describes the following:

- 8.1, “Why do we need certifications?” on page 214
- 8.2, “FIPS” on page 214
- 8.3, “ISO Standards” on page 215
- 8.4, “Security Technical Implementation Guides (STIGs)” on page 216
- 8.5, “Center for Internet Security” on page 216

8.1 Why do we need certifications?

The entire purpose of this publication is to provide a comprehensive idea of how to properly implement security on IBM Power and accompanying technologies (e.g, IBM Storage, networking, etc.). In order to do this, IBM Power and complementary products need to comply with as many industry certifications as possible. But this isn't as simple as IBM claiming that we provide compliant products. A more credible (and certainly more reliable) method of signifying compliance to industry-wide security standards is for third-parties themselves to certify that IBM indeed provides solutions that can adhere to these demanding security guidelines.

Certifications provide customers third-party validations that vendor products are developed using security best practices. This helps ensure that your data is protected from being compromised, that your environment complies with evolving data privacy and business resiliency laws and regulations, and that financial risk is minimized to avoid steep fines for non-compliance.

8.2 FIPS

The Federal Information Processing Standards (FIPS) are a set of publicly announced standards that the National Institute of Standards and Technology (NIST) has developed for use in computer of agencies and contractors. FIPS standards establish requirements for ensuring computer security and interoperability, and are intended for cases in which suitable industry standards do not already exist.

If we are focusing on IBM Power Servers security standard, we shall mention IBM PCIe Cryptographic Coprocessors of which they are a family of high-performance hardware security modules (HSM). These programmable PCIe cards work with IBM Power servers to offload computationally intensive cryptographic processes such as secure payments or transactions from the host server. Using these HSMs allow you to gain significant performance and architectural advantages and enable future growth by offloading cryptographic processing from the host server in addition to delivering a high-speed cryptographic functions for data encryption and digital signing, secure storage of signing keys or custom cryptographic applications. That has been validated to FIPS PUB 140-2, Security Requirements for Cryptographic Modules, Overall Security Level 4, the highest level of certification achievable.

See the link below for more information:

https://www.ibm.com/products/pcie-cryptographic-coprocessor

Since each of IBM's HSM devices offer the highest cryptographic security available commercially. Federal Information Processing Standards (FIPS) publication 140-2 defines security requirements for cryptographic modules. It is issued by the National Institute of Standards and Technology (NIST) and is widely used as a measure of the security of HSMs. The cryptographic processes of each of the IBM HSMs are performed within an enclosure on the HSM that is designed to provide complete physical security.

See the high security modules page:

https://www.ibm.com/docs/en/cryptocards?topic=hsm-highlights

8.2.1 Software Layer Support

The evolutions for security standard are not only specific to IBM Power operating systems, but also for IBM Java SDK that works on top of IBM Power. See the pages below for more information:

– https://www.ibm.com/docs/en/sdk-java-technology/8?topic=guide-fips-140-3-evaluation-technology
– https://csrc.nist.gov/projects/cryptographic-module-validation-program/certificate/4079

IBM AIX operating system supports FIPS, you can read more information at the below links:

– https://www.stigviewer.com/stig/ibm_aix_7.x/2023-08-23/
– https://csrc.nist.gov/CSRC/media/projects/cryptographic-module-validation-program/documents/security-policies/140sp3269.pdf

If you are dealing with the Red Hat Enterprise Linux CoreOS (RHCOS) machines in your OpenShift cluster, this can be applied when the machines are deployed based on the status of certain installation options that governs the cluster options of which a user can change during cluster deployment. With Red Hat Enterprise Linux (RHEL) machines, you must enable FIPS mode when you install the operating system on the machines that you plan to use as worker machines. These configuration methods ensure that your cluster meets the requirements of a FIPS compliance audit: only FIPS validated or Modules In Process cryptography packages are enabled before the initial system boot.

Read more information at the following link:

– https://community.ibm.com/community/user/powerdeveloper/blogs/paul-bastide/2023/11/21/enabling-fips-compliance-in-openshift-cluster-plat

8.3 ISO Standards

Why the need for ISO standards? From a client's perspective, they choose an ISO-certified supplier for general trust and the reliability of the services provided. Clients need to work with reliable suppliers for different reasons, such as audited processes, documented processes, and other behind-the-scenes requirements that can be extra proof for professionalism and long-term stability. An ISO-certified supplier is thus more reliable than one who doesn't have any ISO certifications. It is the same for other reputable certifications, like the SOC 2 Type II or PCI DSS.

From a supplier's perspective, having ISO certifications in place leads to continuous improvement of the services offered. Another benefit is the presence of auditable processes; there is a significant difference between just having various internal processes and the processes that can be audited. Even more so, there are clear guidelines in place on how to audit these processes.

Common Criteria (ISO 15408) is the only global mutually recognized product security standard. The goal of the Common Criteria is to develop confidence and trust in the security characteristics of a system and in the processes used to develop and support it.

The ISO 15408 international standard is specifically for computer security certification. The full description of the ISO 15408 standard can be found at:

– `https://www.iso.org/standard/72891.html`.

8.4 Security Technical Implementation Guides (STIGs)

The United States Department of Defense (DoD) systems have yet another layer of requirements promulgated by the Defense Information Systems Agency (DISA). Though more of a set of guidelines than a certification, the Security Technical Implementation Guides (STIGs) and Security Requirements Guides for the Department of Defense (DOD) information technology systems describe security hardening guidelines as mandated by DODI 8500.01.

Federal IT security pros within the DoD must comply with the STIG technical testing and hardening frameworks. According to DISA (https://disa.mil/), STIGs “are the configuration standards for DOD [information assurance, or IA] and IA-enabled devices and system. The STIGs contain technical guidance to ‘lock down’ information systems/software that might otherwise be vulnerable to a malicious computer attack.”

You can search through a publicly available document library of STIGs at https://public.cyber.mil/stigs/downloads/. Also, Table 8-1 below lists some specific operating systems and components and their corresponding STIGs:

Table 8-1 STIGs for various IBM components

Product	Link to STIG
Hardware Management Console	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_IBM_HMC_Y24M07_STIG.zip
AIX v7.1 and v7.2	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_IBM_AIX_7-x_V2R9_STIG.zip
Red Hat Enterprise Linux 8	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_RHEL_8_V1R14_STIG.zip
Red Hat Enterprise Linux 9	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_RHEL_9_V2R1_STIG.zip
SUSE Linux Enterprise Server 12	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_SLES_12_V2R13_STIG.zip
SUSE Linux Enterprise Server 15	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_SLES_15_V2R1_STIG.zip
IBM DB2 V10.5	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_IBM_DB2_V10-5_LUW_V2R1_STIG.zip
IBM WebSphere® Liberty Server	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_IBM_WebSphere_Liberty_Server_V1R2_STIG.zip
IBM WebSphere Traditional V9.x	https://dl.dod.cyber.mil/wp-content/uploads/stigs/zip/U_IBM_WebSphere_Traditional_V9-x_V1R1_STIG.zip

8.5 Center for Internet Security

The Center for Internet Security (CIS), a nonprofit founded in October 2000, unites the global IT community to develop, validate, and promote best practices in cybersecurity. Over the years, CIS has created and distributed numerous free tools and solutions aimed at enhancing cybersecurity readiness for organizations of all sizes.

CIS is best known for its CIS Controls, a comprehensive framework consisting of 20 essential safeguards and countermeasures designed to improve cyber defense. These controls offer a prioritized checklist that organizations can use to significantly reduce their vulnerability to cyberattacks. Additionally, CIS produces CIS Benchmarks, which provide best practice recommendations for secure system configurations, referencing these controls to guide organizations in building stronger security measures.

8.5.1 CIS benchmarks

Created by a global network of cybersecurity experts, CIS Benchmarks provide best practices for securely configuring IT systems, software, networks, and cloud infrastructure. Published by the Center for Internet Security (CIS), there are currently over 140 CIS Benchmarks across seven core technology categories.

These benchmarks are developed through a consensus-based process involving cybersecurity professionals and subject matter experts worldwide. This collaborative approach ensures that security best practices are continuously updated and validated.

CIS benchmarks align closely with-security and data privacy regulatory frameworks including the NIST (National Institute of Standards and Technology) Cybersecurity Framework, the PCI DSS (Payment Card Industry Data Security Standard) (PCI DSS), HIPAA (Health Insurance Portability and Accountability Act), and ISO/EIC 2700. As a result, any organization operating in an industry governed by these types of regulations can make significant progress toward compliance by adhering to CIS benchmarks. In addition, CIS Controls and CIS Hardened Images can help support an organization's compliance with GDPR (the European Union's General Data Protection Regulation).

Each CIS Benchmark offers configuration recommendations organized into two profile levels: Level 1 and Level 2. Level 1 profiles provide base-level configurations that are easier to implement with minimal impact on business operations. Level 2 profiles are designed for high-security environments, requiring more detailed planning and coordination to implement while minimizing business disruption.

There are seven (7) core categories of CIS benchmarks:

1. Operating systems benchmarks cover security configurations of core operating systems, such as Microsoft Windows, Linux and Apple OSX. These include best-practice guidelines for local and remote access restrictions, user profiles, driver installation protocols and internet browser configurations.
2. Server software benchmarks cover security configurations of widely used server software, including Microsoft Windows Server, SQL Server, VMware, Docker and Kubernetes. These benchmarks include recommendations for configuring Kubernetes PKI certificates, API server settings, server admin controls, vNetwork policies and storage restrictions.
3. Cloud provider benchmarks address security configurations for Amazon Web Services (AWS), Microsoft Azure, Google, IBM and other popular public clouds. They include guidelines for configuring identity and access management (IAM), system logging protocols, network configurations and regulatory compliance safeguards.
4. Mobile device benchmarks address mobile operating systems, including iOS and Android, and focus on areas such as developer options and settings, OS privacy configurations, browser settings and app permissions.
5. Network device benchmarks offer general and vendor-specific security configuration guidelines for network devices and applicable hardware from Cisco, Palo Alto Networks, Juniper and others.

6. Desktop software benchmarks cover security configurations for some of the most commonly used desktop software applications, including Microsoft Office and Exchange Server, Google Chrome, Mozilla Firefox and Safari Browser. These benchmarks focus on email privacy and server settings, mobile device management, default browser settings and third-party software blocking.
7. Multi-function print device benchmarks outline security best practices for configuring multi-function printers in office settings and cover such topics as firmware updating, TCP/IP configurations, wireless access configuration, user management and file sharing.

Currently, there are more than 100 CIS Benchmarks that are available through free PDF download for non-commercial use.

CIS benchmarks for IBM Power

CIS has developed specific benchmarks for AIX, IBM i and various Linux distributions which run on IBM Power.

Here are some CIS Benchmarks for that are relevant to IBM Power systems:

- CIS Benchmark for IBM AIX:

  This benchmark provides security configuration guidelines for IBM's AIX operating system, which is commonly used on IBM Power Systems. It includes best practices for system configuration to enhance security and reduce vulnerabilities.

- CIS Benchmark for IBM i

  This benchmark offers recommendations for securely configuring the IBM i operating system, It focuses on system settings, security policies, and configurations to improve overall security posture.

- CIS Benchmarks for Linux

  For IBM Power Systems running Linux there is a generic Linux benchmark as well as benchmarks for Red Hat Enterprise Linux, SUSE Enterprise Linux and Ubuntu Linux.

These benchmarks are regularly updated to reflect the latest security practices and vulnerabilities. You can find the most recent versions and additional details on the CIS website or through their publications and resources. Table 8-2 provides a more comprehensive list.

Table 8-2 CIS compliance of IBM Power supported operating systems

Operating system	Recent versions available for CIS Benchmark
AIX	IBM AIX 7.2 (1.1.0) IBM AIX 7.1 (2.1.0)
IBM i	IBM i V7R5M0 (2.0.0) IBM i V7R4M0 (2.0.0) IBM i V7R3M0 (1.0.0)
Red Hat Enterprise Linux	Red Hat Enterprise Linux 9 (2.0.0) Red Hat Enterprise Linux 8 (3.0.0) Red Hat Enterprise Linux 8 STIG (1.0.0)
SUSE Linux Enterprise Server	SUSE Linux Enterprise 15 (1.1.1) SUSE Linux Enterprise 12 (3.1.0)
Ubuntu Linux	Ubuntu Linux 22.04 LTS (2.0.0)

For the most accurate and current information, always refer to the CIS official website.

9

PowerSC

IBM PowerSC is a security and compliance solution optimized for virtualized environments on IBM Power servers running AIX, IBM i or Linux.

PowerSC sits on top of the IBM Power server stack, integrating security features built at different layers. You can now centrally manage security and compliance on Power for all IBM AIX and Linux on Power endpoints. In this way you can get better support for compliance audits, including GDPR.

This chapter discusses PowerSC security features and most usable components such as:

- “Compliance automation” on page 220
- “Real-time file integrity monitoring” on page 220
- “Endpoint Detection and Response” on page 221
- “Anti-malware integration” on page 221
- “Multi factor authentication” on page 222

9.1 Compliance automation

The PowerSC Security and Compliance Automation feature is an automated method to configure and audit systems in accordance with Department of Defense (DoD) Security Technical Implementation Guide (STIG), the Payment Card Industry (PCI) data security standard (DSS), the Sarbanes-Oxley act, COBIT compliance (SOX/COBIT), the Health Insurance Portability and Accountability Act (HIPAA), Center for Internet Security (CIS) benchmarks compliance for AIX, and IBM i best practices.

PowerSC helps to automate the configuration and monitoring of systems that must be compliant with the Payment Card Industry (PCI) data security standard (DSS). Therefore, the PowerSC Security and Compliance Automation feature is an accurate and complete method of security configuration automation that is used to meet the IT compliance requirements of the DoD UNIX STIG, the PCI DSS, the Sarbanes-Oxley act, SOX/COBIT, and HIPAA.

The PowerSC Security and Compliance Automation feature creates and updates ready XML profiles that are used by IBM Compliance Expert express (ICEE) edition. You can use the PowerSC XML profiles with the **pscxpert** command.

The preconfigured compliance profiles delivered with PowerSC reduce the administrative workload of interpreting compliance documentation and implementing the standards as specific system configuration parameters. This technology reduces the cost of compliance configuration and auditing by automating the processes. IBM PowerSC is designed to help effectively manage the system requirement associated with external standard compliance that can potentially reduce costs and improve compliance.

See more information in this IBM PowerSC document.

9.2 Real-time file integrity monitoring

The PowerSC GUI includes File Integrity Monitoring (FIM). FIM includes Real Time Compliance (RTC) for AIX, IBM i audit, and Linux Auditd events. By using the PowerSC GUI, you can configure and view real-time events. This allows you to manage extensive profiles by editing and customizing the reporting capabilities.

The interesting part is that it takes care of critical files that exist on a system that contain sensitive data, such as configuration details and user information. From a security perspective, it is important to monitor changes that are made to these sensitive files.

File Integrity Monitoring (FIM) is a method that can detect all of that not only for critical files, but also for binaries and libraries.

PowerSC has the capability to generate real-time alerts whenever the contents of a monitored file is changed and even when a file’s characteristics are modified. By using the AHAFS event monitoring technology, PowerSC RTC monitors all of these changes and will generate alerts using the following methods:

- Email alerts
- Log message to a file
- SNMP message to your monitoring server
- Alert to PowerSC GUI server

See more information in this document and in this PowerSC GUI description.

9.3 Endpoint Detection and Response

With the recent increase in ransomware and other cybersecurity attacks, PowerSC has added endpoint detection and response (EDR) capabilities. Aspects of EDR include the following items:

- Intrusion detection and prevention
- log inspection and analysis
- Incident response and event triggering and filtering

More information can be found in this document and in this IBM Support page.

One of the EDR forms in PowerSC is that you can configure IDP for a specific endpoint. For AIX, the PowerSC GUI allows you to use the IP Security (IPSec) facility of AIX to define parameters for intrusion detection. The IP Security (IPSec) facility of AIX must already be installed on the AIX endpoint. For Red Hat Enterprise Linux Server and SUSE Linux Enterprise Server, you must install the **psad** package on each endpoint on which you want to run **psad**, as described in Installing PowerSC on Linux systems, before you can use it with PowerSC GUI.

The PowerSC GUI uiAgent monitors the endpoint for port scan attacks on the ports listed in IPSec filter rules. By default, PowerSC creates an IPv4 rule in */etc/idp/filter.rules* to monitor operating system network ports. PowerSC also creates the */var/adm/ipsec.log* log file. The IP Security (IPSec) facility of AIX also parses IPv6 rules in */etc/idp/filter.rules* and the IPv6 addresses appear in the event list.

See both links below for more setup and information:

- – https://www.ibm.com/docs/en/powersc-standard/2.2?topic=ids-configuring-intrusion-detection-prevention-idp-aix-endpoints
- – https://www.ibm.com/docs/en/powersc-standard/2.2?topic=security-configuring-intrusion-detection-system-ids

9.4 Anti-malware integration

This section shows how IBM PowerSC can act as an anti malware defense.

IBM PowerSC has the capability to integrate with ClamAV global anti-virus software toolkit to help prevent malware attacks and detect Trojans, viruses, malware and other malicious threats. This happens by scanning all incoming data to prevent malware from being installed and infecting the server.

Through the PowerSC server UI, you can configure anti-malware settings for specific endpoints. ClamAV will then move or copy any detected malware to the quarantine directory on the PowerSC uiAgent, assigning a time-stamped prefix and nullifying file permissions to prevent access. Note that ClamAV is not included in the initial PowerSC package, so you'll need to install it on the uiAgent before it can be utilized with the PowerSC GUI.

See the following links for installing the ClamAV toolkit on the operating systems:

- – https://www.ibm.com/docs/en/powersc-standard/2.2?topic=malware-installing-anti-aix
- – https://www.ibm.com/docs/en/powersc-standard/2.2?topic=cam-installing-anti-malware-red-hat-enterprise-linux-server-suse-linux-enterprise-server
- – https://www.ibm.com/docs/en/powersc-standard/2.2?topic=malware-installing-configuring-anti-i

For generic PowerSC ClamAV detailed configuration and features, please follow the below link:

- https://www.ibm.com/docs/en/powersc-standard/2.2?topic=security-configuring-anti-malware

9.5 Multi factor authentication

This section shows how IBM PowerSC can be an authentication server with multi-layered levels of authentication.

IBM PowerSC has a capability of deploying a Multi-Factor Authentication (MFA) for mitigating the risk of data breach caused by compromised credentials. PowerSC Multi-Factor Authentication (PMFA) provides numerous flexible options for implementing MFA on Power. PMFA is implemented with a Pluggable Authentication Module (PAM), and can be used on AIX, VIOS, RHEL, SLES, IBM i, HMC, and PowerSC Graphical User Interface server.

The National Institute of Standards and Technology (NIST) defines MFA as authentication that uses two or more factors to achieve authentication.

Factors include “something that you know”, such as password or personal identification number. Factors include “something that you have”, such as a cryptographic identification device or a token. Factors include “something that you are”, such as a biometric.

IBM PowerSC authentication factors improves the security of user accounts. It allows the user to either provide the credentials directly in the application (in-band) or out-of band.

For in-band authentication, users can generate a token to satisfy a policy and use that token to directly log in, however out-of-band authentication allows users to authenticate on a user-specific web page with one or more authentication methods to retrieve a cache token credential (CTC) that they then use to log in. For more information, see Out-of-band authentication type.

IBM PowerSC MFA server can be installed on AIX, IBM i or Linux operating systems. See the links for installation procedures:

- https://www.ibm.com/docs/en/powersc-mfa/2.2?topic=installing-powersc-mfa-server-aix
- https://www.ibm.com/docs/en/powersc-mfa/2.2?topic=installing-powersc-mfa-server-pase-i
- https://www.ibm.com/docs/en/powersc-mfa/2.2?topic=installing-powersc-mfa-server-linux

See also the full use guide and installation for IBM PowerSC MFA:

- https://www.ibm.com/docs/en/SS7FK2_2.2/pdf/powersc_mfa_users_pdf.pdf
- https://www.ibm.com/docs/en/SS7FK2_2.2/pdf/powersc_mfa_install_pdf.pdf

9.5.1 IBM PowerSC MFA high availability

Since IBM PowerSC depends on Postgres database, there is a capability to configure the IBM PowerSC MFA Postgres database for streaming replication between the primary IBM PowerSC MFA server and a secondary IBM PowerSC MFA server. The feature exists to improves availability and is not a load balancing solution.

In the replication model, the Postgres database on the secondary IBM PowerSC MFA server is a read-only copy of the database on the primary IBM PowerSC MFA server. In the event that the database on the primary IBM PowerSC MFA server becomes unavailable, you can promote the database on the secondary IBM PowerSC MFA server to be the primary. Only one database can be the primary database at any time.

Before you configure IBM PowerSC MFA for high availability, satisfy the following prerequisites:

- The primary and secondary server must use the same operating system.
- Updates to any files in /opt/IBM/powersc/MFA/mfadb are not preserved if you reinstall the IBM PowerSC MFA server.
- If the secondary server uses Red Hat Enterprise Linux Server or SUSE Linux Enterprise Server, install Postgres, openCryptoki and opencryptoki-swtok on the secondary server.

See more information in the following link:

  – https://www.ibm.com/docs/en/powersc-mfa/2.2?topic=configuring-powersc-mfa-high-availability

10

IBM Power Virtual Server Security

As the migration of IT services and workloads to the cloud accelerates, cloud security becomes paramount. Given the hybrid cloud's essential role in deploying critical business applications, prioritizing security is imperative.

With the introduction of PowerVS and its ability to run AIX, IBM i, and Linux on Power in the cloud, understanding PowerVS security is crucial for establishing a reliable and secure environment.

This chapter is designed to give a high level overview of security in PowerVS. For additional information, reference the links in section 10.7, “Additional References” on page 229

In this chapter we introduce and discuss:

- “Introduction to PowerVS” on page 226
- “Authentication and Authorization” on page 226
- “IBM Cloud Key Management Services” on page 227
- “Network interfaces” on page 227
- “Network security” on page 228
- “Security and Compliance” on page 228
- “Additional References” on page 229

10.1 Introduction to PowerVS

The IBM Power Virtual Server (PowerVS) offering allows users to deploy Power Servers running AIX, IBM i, or Linux in the cloud. This chapter provides details on various security features and compliance met by Power Virtual Server.

10.2 Authentication and Authorization

PowerVS user access is controlled using IBM Cloud Identity and access management (IAM) service. Table 1 displays the IAM platform access roles and the corresponding type of control that is allowed by the Power Virtual Server

Table 10-1   IAM platform access roles:

Platform access role	Type of access allowed
Viewer	View instances and list instances.
Operator	View instances and manage aliases, bindings (IBM Power Virtual Server Private Cloud only), and credentials.
Editor	View instances, list instances, create instances, and delete instances.
Administrator	View instances, list instances, create instances, delete instances, and assign policies to other users.

You can use the service access roles to define the actions that the users can perform on Power Virtual Server resources. Table 10-2 displays the IAM service access roles and the corresponding actions that a user can complete by using the Power Virtual Server:

Table 10-2   IAM service access roles

Service access role	Description of actions
Reader	View all resources (such as SSH keys, storage volumes, and network settings). You cannot make changes to the resources
Manager	Configure all resources. You can perform the following actions: ▸ Create instances ▸ Increase storage volume sizes ▸ Create SSH keys ▸ Modify network settings ▸ Create boot images ▸ Delete storage volumes

When you assign access to the Power Virtual Server service, you can set the access scope to:

- All resources
- Specific resources, which support the following selections:
  - Resource group
  - Service instance

Power Virtual Server requires extra access for features such as Direct Link, Transit Gateway service, and Virtual Private Cloud. You might require these extra access capabilities based on your resource requirements. Table 10-3 displays the additional access roles that are required for the corresponding type of services that are allowed by Power Virtual Server:

Table 10-3 Additional access roles

Additional access role	Resources Attribute
Editor, Manager, Operator, Reader, Viewer	Power Virtual Server service
Editor, Manager, Operator, Reader, Viewer, VPN Client	VPC Infrastructure Services service
Editor, Operator, Viewer	Transit Gateway service
Reader, Viewer	All resources in account (Including future IAM enabled services)
Editor, Operator, Viewer	Direct Link service
Viewer	All resource group
Viewer	Satellite service

10.3 IBM Cloud Key Management Services

IBM Cloud provides two Cloud key management services that integrate with Power Virtual Server workloads:

- IBM Cloud Hyper Protect Crypto Services (HPCS) is a dedicated key management service and hardware security module (HSM) based on IBM Cloud. You can integrate HPCS with Power Virtual Server to securely store and protect encryption key information for AIX and Linux.
- IBM Key Protect is a full-service multi-tenant encryption solution that allows data to be secured and stored in IBM Cloud with the envelope encryption techniques. You can integrate Key Protect with Power Virtual Server to securely store and protect encryption key information for AIX and Linux.

10.4 Network interfaces

PowerVS is logically isolated from all other public cloud tenants and infrastructure components, creating a private, secure place on the public cloud. This isolation includes all Logical partitions, network and storage. When you create a Power Virtual Server you can choose a connectivity type from various available options. When you create a logical partition on PowerVS, you can bind a public interface or multiple private network interfaces to it. A public network interface is implemented by using an IBM Cloud Virtual Router Appliance (VRA) and a Direct Link Connect connection. The public network is protected by a firewall and supports SSH (port 22), HTTPS (port 443), Ping, IBM i 5250 terminal emulation with SSL (port 992) network protocols and ports. Private network interfaces use a Direct Link Connect connection to connect to your IBM Cloud account network and resources.

10.5 Network security

As discussed in previous sections, Power Virtual Server internal networks are isolated. In order to meet different network requirements, IBM Cloud offers many connectivity options to multiple environments on the internal private network, public networks or on premise networks. Cloud Connections (also called Direct Link Connect) facilitates connectivity between PowerVS and other IBM Cloud environments (Classic, VPC, etc). Cloud Connections can connect directly to specific VPCs, the classic environment, or to an IBM Cloud Transit Gateway. The IBM Cloud Transit Gateway is a network service that interconnects IBM PowerVS, IBM Cloud VPCs and classic infrastructure, allowing users to build a global network. Transit Gateway is deployed in a hub and spoke model, where the Transit Gateway is the hub and IBM Cloud VPC, PowerVS, and Classic Infrastructure are the spokes. Transit Gateways can be scoped locally (Local Transit Gateway) or across regions (Global Transit Gateways). Using Transit Gateway, environments can be configured across geographies for High Availability/Disaster Recovery. A Generic Routing Encapsulation (GRE) tunnel connects two endpoints (a firewall or a router and another network appliance) in a point-to-point logical link. Finally, Power Edge Router (PER) is a high-performance networking component which provides a direct access to the IBM Cloud services from the Power Virtual Server workspace. It also provides a direct access to the Power Virtual Server from a client-managed environment by using a Direct Link connect or Direct Link dedicated.

10.6 Security and Compliance

Power Virtual Server meets below compliance requirements of industry:

- General Data Protection Regulation (GDPR) as data protection law framework across the European Union (EU).
- Financial Services Validated as per the IBM Cloud framework for financial services control requirements.
- System and Organization Controls (SOC) audit of the internal controls at a service organization that is implemented to protect client-owned data involved in client financial reporting. Also includes audit based on the Statement on Standards for Attestation Engagements (SSAE 18) and the International Standards for Assurance Engagements No. 3402 (ISAE 3402). Below SOC reports are available for Power Virtual Server:
  - System and Organization Controls (SOC) 1 Type II report
  - System and Organization Controls (SOC) 2 Type II report
- International Organization for Standardization (ISO): The Power Virtual Server provides services that are delivered from global data centers that are a component of the IBM Cloud IaaS ISO certification. The ISO certification covers a family of 4 standards as follows:
  - ISO/IEC 27001:2013
  - ISO/IEC 27017:2015
  - ISO/IEC 27018:2019
  - ISO/IEC 27701:2019
- PCI-DSS to meet Payment Card Industry (PCI) data security standards. A Service Responsibility Matrix (SRM) guide for Power Virtual Server will be available on request.
- US Health Insurance Portability and Accountability Act (HIPAA) to build HIPAA-ready environments and applications by using Power Virtual Server.

10.7 Additional References

The following links provide additional information on security options in PowerVS.

- *IBM Power Systems Cloud Security Guide: Protect IT Infrastructure In All Layers*, REDP-5659
- https://cloud.ibm.com/docs/power-iaas?topic=power-iaas-getting-started
- https://cloud.ibm.com/docs/power-iaas?topic=power-iaas-network-security
- https://cloud.ibm.com/docs/power-iaas?topic=power-iaas-about-power-iaas#public-private-networks
- https://cloud.ibm.com/catalog/architecture/deploy-arch-ibm-pvs-inf-2dd486c7-b317-4aaa-907b-42671485ad96-global
- https://medium.com/@jose.guerra.m/ibm-cloud-powervs-networking-concepts-17bd04644419
- https://cloud.ibm.com/docs/power-iaas?topic=power-iaas-managing-resources-and-users

11

Lessons Learned and Future Directions in Power System Security

This chapter provides a summary of what our residents have found in real life. In addition it presents the findings of a recent IBM study on current threats in the Cyber Security environment. Topics included in this chapter are:

- “Lessons Learned from Real-World Breaches” on page 232
- “Basic AIX Security Strategies and Best Practices” on page 233
- “Fix Level Recommendation Tool for IBM Power” on page 238
- “Physical security” on page 239

11.1 Lessons Learned from Real-World Breaches

A strong security culture is the backbone of any effective security program. People are the final line of defense, and their awareness and actions directly impact an organization's vulnerability. While security policies and procedures are essential, their success hinges on employees understanding the risks and consistently practicing safe behaviors.

While learning from real-world incidents is valuable, proactive measures are crucial to prevent costly breaches. Security experts emphasize the importance of cultivating a security-conscious workforce through targeted training and awareness campaigns. By fostering a culture where security is a shared responsibility, organizations can significantly reduce their risk exposure.

11.1.1 Recommendations to Reduce Data Breach Costs

IBM published Cost of a Data Breach Report 2024 listing findings of research from IBM and Ponemon Institute and providing insights from the experiences of 604 organizations and 3,556 cybersecurity and business leaders hit by a breach. Out of the research came the following recommendations on how to mitigate the risks of a breach:

- Comprehensive Data Visibility: Gain a complete understanding of data locations (on-premises, cloud, etc.) and implement robust data security measures across all environments.
- AI and Automation: Leverage AI and automation for enhanced threat detection, response, and vulnerability management.
- Gen AI Security: Prioritize security in AI initiatives by protecting data, models, and infrastructure.
- Incident Response Preparedness: Conduct regular cyber range simulations and train employees on incident response procedures.

By following these recommendations, organizations can significantly reduce the financial and reputational impact of a data breach.

11.1.2 Summary of IBM X-Force Threat Intelligence Index 2024[1]

X-Force® is a team of elite cybersecurity professionals, including hackers, incident responders, researchers, and analysts. With a deep understanding of the threat landscape, they offer a comprehensive approach to defending against cyberattacks. The Red Team thinks like attackers to uncover vulnerabilities, while the Incident Response team focuses on preventing, detecting, and responding to threats. The researchers stay ahead of emerging threats, and the analysts transform complex data into actionable insights.

IBM X-Force published the IBM X-Force Threat Intelligence Index 2024. The following is a summary of the findings:

- Identity-Centric Attacks: Cybercriminals increasingly target identities as the easiest point of entry, with a significant rise in credential theft and abuse.
- Ransomware Decline, Data Theft Surge: While ransomware attacks decreased, data theft and leaks became the primary motivation for cyberattacks.
- Infostealer Malware Growth: The use of infostealer malware to steal credentials has skyrocketed, fueling the dark web's stolen credential market.

[1] IBM X-Force Threat Intelligence Index 2024

- Misconfiguration and Legitimate Tool Abuse: Security misconfiguration and the misuse of legitimate tools contributed significantly to breaches.
- Emergence of AI as a Target: The rapid adoption of AI is creating a new attack surface, and cybercriminals are likely to focus on AI platforms once they achieve market dominance.
- Manufacturing Remains Top Target: The manufacturing industry continues to be the most targeted sector, with malware and ransomware being the primary threats.

Overall, the report highlights a shift in cybercrime tactics towards identity-based attacks and data theft, while also warning of the growing threat posed by AI. Organizations must prioritize identity protection, implement strong security measures, and stay vigilant against evolving threats.

11.1.3 Best practices for data breach prevention

This IBM blog on data breach prevention presents the following best practices. It states that “To enhance your cyber resilience, it is vital to build security in every stage of software and hardware development”. You can strengthen your data breach prevention strategy by:

- Proactive Risk Management: Employ a zero-trust security framework and conduct rigorous testing to identify and eliminate vulnerabilities before breaches occur.
- Data Protection: Safeguard sensitive information with multi-factor authentication, strong passwords, and employee training to prevent data loss and identity theft.
- Business Continuity: Implement robust data backup strategies and well-rehearsed incident response plans to minimize downtime and financial losses in case of emergencies.

11.1.4 Summary

The importance of fixing the basics is key. In other words, security is built from steps such as asset inventory, patching and training. Some important points to take in consideration:

- Develop an automated methodology for secure assessments and detection.
- Establish a risk management framework that includes cyber insurance.
- Maintain a dedicated environment for testing security patches.
- Ensure rollback options are available in all scenarios.

11.2 Basic AIX Security Strategies and Best Practices

Security is a very hot topic for a good reason. So much of our personally identifying data is now being stored that the security break-ins that have been happening have most likely affected everyone reading this article. Additionally, the penalties now for breaches of the various standards (HIPAA, PCI, etc) are significant. Good security requires a multi-layered approach that starts with people, then physical security, and then the various layers. It s critical to look at the whole environment and see how security can be applied at each level. In this section we will cover some of the basics of locking down your LPARs. This is not done by default, but it is fairly easy to do. It ranges from default permissions and umasks, good usernames and passwords, logging, patching, and removing insecure daemons to integration with LDAP or active directory (AD).

11.2.1 Usernames and Passwords

This is one of the most basic protections. In order to have longer usernames and passwords, you need to make a system change. Changing the username length is almost always required if you want to integrate with LDAP or AD (active directory) and it requires a reboot. Below is the command to increase the maximum username length to 36:

```
chdev -l sys0 -a max_logname=36
```

Then you check it as follows:

```
# lsattr -El sys0 | grep max_log
max_logname     36          Maximum login name length at boot time           True
```

The above change requires a reboot of the LPAR. In order to have longer passwords, you need to use the **chsec** command. The version below causes the system to use ssha256 (up to 255 characters) for passwords. The next time local users change their password they will get a much longer, more secure password.

```
chsec -f /etc/security/login.cfg -s usw "pwd_algorithm=ssha256"
```

Then you can use getconf to check the setting

```
getconf PASS_MAX
255¬†
```

Finally, I normally set the system up to automatically create home directories, this is important in an LDAP or AD environment. An illustration of this is shown in Example 11-1.

Example 11-1 Automatically creating home directories

```
chsec -f /etc/security/login.cfg -s usw -a mkhomeatlogin=true
tail /etc/security/login.cfg
    pwd_algorithm = ssha256
        mkhomeatlogin = true
```

11.2.2 Logging

Logging is a critical part of any system-protection strategy. Without logs, it is impossible to know what has been happening on the system. The syslog daemon (syslogd) starts by default on AIX, but the log configuration file is not set up to actually log everything. The first step is to correctly set up /etc/syslog.conf. It is best to set up a separate file system for logs (e.g., /usr/local/logs) rather than use the default of /var/spool. If /var fills up, the system will crash; if your separate file system fills up, it will just stop logging. Although file systems should be monitored, it is still wise to store logs in their own file system to protect against large logs bringing down the system. Logs can be written to a file, sent to the console, logged to a central host across the network (be wary of this as the traffic can be substantial), e-mailed to an administrator or sent to all logged-in users or any combination thereof. The most commonly used method is writing to a file in a file system. Once the file system is set up, code a /etc/syslog.conf file. Example 11-2 on page 235 shows an example file that writes to a local filesystem. It keeps the logs to no more than 2MB, then rotates and compresses them, keeping the last 10 logs. I do this on all LPARs and VIO servers.

Example 11-2 Sample log configuration

```
mail.debug      /usr/local/logs/mailog rotate size 2m files 10 compress
*.emerg         /usr/local/logs/syslog rotate size 2m files 10 compress
*.alert         /usr/local/logs/syslog rotate size 2m files 10 compress
*.crit          /usr/local/logs/syslog rotate size 2m files 10 compress
*.err           /usr/local/logs/syslog rotate size 2m files 10 compress
auth.notice     /usr/local/logs/infolog rotate size 2m files 10 compress
*.info          /usr/local/logs/messages rotate size 2m files 10 compress
```

Go into /usr/local/logs and create each of the files above using **touch**. Now you can stop (**stopsrc -s syslogd**) and start (**startsrc -s syslogd**) the logging daemon.

11.2.3 Insecure Daemons

The /etc/inetd.conf is over 120 lines long Some items are commented out but most are not. Many of the protocols listed in there have known security holes. One of the first things I do after setting up an lpar is to secure those protocols. I take a copy of /etc/inetd.conf to /etc/inetd.conf-orig and then delete everything in /etc/inetd.conf except the items I want to keep. I then do a refresh -s inetd. I do this on all my LPARs and my VIO servers. Typically, my inetd.conf looks like Example 11-3.

Example 11-3 inetd.conf file

```
#ftp stream      tcp6  nowait   root /usr/sbin/ftpd ftpd
#telnet stream   tcp6  nowait   root /usr/sbin/telnetd telnetd -a
#xmquery dgram   udp6  wait     root /usr/bin/xmtopas
#dtspcd stream   tcp   nowait   root /usr/dt/bin/dtspcd /usr/dt/bin/dtspcd
```

You will notice it is only four lines and everything is commented out. On a NIM server you will see tftp and bootp uncommented. Occasionally when you do maintenance it uncomments or adds back in services. When the file is only 4 lines you can see immediately that it did that. I do not use ftp and telnet because they are insecure and I have ssh and sftp instead. If you have to use telnet or ftp then you can uncomment them, but remember they send passwords, etc in clear text. I would also recommend looking at /etc/rc.tcpip to see if snmp, sendmail and other daemons are starting. If you need snmp or sendmail to run then they should be properly configured to keep hackers from taking advantage of default exploits.

11.2.4 Time Synchronization

In a previous life I used to do a lot of security investigations. When issues involve multiple systems it is a nightmare trying to follow the trails when the timestamps do not match. A simple fix for this is to implement NTP (network time protocol). Most companies have an NTP server set up on their AD servers, but you can always point to the atomic clocks if necessary. You should set up ntp on your LPARs and your VIO servers.

11.2.5 Patching

At a minimum, make sure you are running a fully supported version of the OS (VIOS, AIX, IBM i, Linux). You can check this using the FLRT (fix level recommendation tool). It is important to keep your patching up to date to proactively solve problems.

In the AIX/VIOS world there are two different kinds of patching. The first is fix packs (technology levels and service packs) and the second is efixes or ifixes (emergency or interim fixes). Fix packs are installed using **install** and efixes/ifixes are installed using **emgr**.

Technology levels and service packs are found at Fix Central. You should check here regularly for updates to your LPARs, VIO servers, server and I/O firmware and HMCs. Additionally, there are products installed – even at the latest service pack – that need updating. Typically this includes Java, OpenSSH and OpenSSL. Java patches are downloaded at Fix Central. OpenSSH and OpenSSL are downloaded at the Web Applications site. I try to get a full patching window every six months unless it is an emergency. You can use the FLRT and FLRTVC tools to determine what patching needs to occur.

Typically, I update the HMC first, then the server firmware, then the I/O firmware and VIOS servers and finally the LPARs. However, you should look at the readme/description files for every update to make sure IBM does not have prerequisites that must be followed. This is particularly important with the HMC and server firmware interaction. There are also some requirements with POWER9 and adapter firmware because of the new trusted boot settings.

FLRTVC (efixes and ifixes)

Fix Level Recommendation Tool Vulnerability Checker (FLRTVC) comes in two flavors:

- The first is a script that you can run on the system that uses data from a file (apar.csv) to compare installed filesets and interim fixes against known security problems.
- The second option is to use the web based online tool (FLRTVC Online). This option allows you to upload the output from two commands to the web page and it produces the output that you need to identify security holes that need to be closed. I normally run it directly on the LPAR or VIO server. This is discussed further in section 11.3, “Fix Level Recommendation Tool for IBM Power” on page 238.

To run **flrtvc** you first need to download the zip file and then unzip it. You may also need to download the apar.csv file. If your LPAR/VIO does not have access to IBM, then you will need to get the file from IBM and upload it to the LPAR yourself. You then edit the script and change SKIPDOWNLOAD from 0 to 1. It will now look for the apar.csv file in the same directory the script is in. Once that is done you can run it in compact mode and produce an output file as follows:

- cd /directory where flrtvc is
- ksh93 ./flrtvc.ksh >systemname-flrtvc-output.csv

Then **sftp** or **scp** (as ASCII) the systemname-flrtvc-output.csv file to your computer and open it with Excel as a csv file. The delimiter is |.

There are a number of flags that you can use but for the most part I do not use any of them as I want to get everything. I tend to have the output go to an NSF mounted filesystem so that all of my security reports are in one place. That way you can concatenate them together or at least just download them all from one place. You can also write scripts that **grep** on certain things in the output and email those to yourself.

The compact output from the FLRTVC script is best viewed in a spreadsheet and is broken down into the following columns:

- Fileset: shows the name of the fileset that is of issue, for example bos.net.tcp.client
- Current Version: shows the currently installed level for example7.1.3.45
- Type: will be either sec (for security) or hiper
- EFix: you only see a value here if the actual efix for the problem is installed.
- Abstract: a description of the problem i.e. Vulnerability in BIND
- Unsafe Versions: a list of the fileset levels that are impacted, for example 7.1.3.0-7.1.3.45
- APARs: provide the actual APAR number i.e. CVE-2015-5477 or IV75031
- Bulletin URL: provides the URL where you can go to read about the vulnerability to get more information
- Download URL: provides the links to get the fix.

You can run **flrtvc** ahead of time and then download and prestage the updates. **Flrtvc** typically identifies efixes and ifixes that need to go on as well as Java, OpenSSH, OpenSSL and other updates that need to go on to the system.

11.2.6 Server Firmware and I/O Firmware

A large number of security and other fixes are done in firmware of some kind. Server firmware and I/O firmware updates can resolve many issues and should be done at least once a year, preferably every six months. Those updates can be downloaded ahead of time and prestaged so that you are not dependent on the IBM sites during the maintenance window.

Typically I will wait until firmware, technology levels or service packs have been out for at least one month (preferably two) before I update to them. At that time, I will update my NIM server and then start to migrate the updates through test, dev, QA and finally, production. Having a good update strategy will save you a lot of downtime and will help with securing your systems.

11.2.7 Active Directory and LDAP Integration

When users are created on AIX systems (or any UNIX*) they are assigned a default user (UID) and group (GID) number. The system permissions for files use those numbers. For that reason, you want to be sure that a user has the same UID and GID on every system. Doing this and setting permissions correctly should ensure they can not access files they are not entitled to. You can either do this manually (think huge spreadsheets) or you can integrate your normal users into an AD or LDAP environment. Implementing AD or LDAP integration on your AIX systems is much easier than it used to be. You will need to work closely with your AD/LDAP admin as well as the security team. A word to the wise, do not put root or other system accounts under the control of AD or LDAP – those have to be local. Instead, you should restrict them to console access only.

11.2.8 Enhanced Access

If you have admins or others who need enhanced access, then this should be provided using **sudo** or some other tool. If multiple users are logging in as root then there is no accountability. Using **sudo** causes everything to be logged. We used to have to install it using rpm and figure out all the prerequisites. Today, we go to the IBM Linux toolbox and download yum.sh and run it. This installs **rpm** and **yum** (it does require **ftp** access to IBM. Once **yum** is installed you can use **yum** to install **sudo** or other tools. You then use **visudo** to put together the rules. You can allow a user to become root with or without a password and you can also restrict them to only being able to use certain commands as root. This is very useful for level 1 support and DBAs who need privileges to perform certain tasks.

11.2.9 Backups

Everyone thinks about taking backups for data, but data is of no use if you have no OS. It s critical to take regular **mksysb** (OS bootable) backups. I normally take them to my NIM server as that is where I would restore the system from. When discussing backups, you need to make sure these bare metal **mksysb** backups are part of any backup and disaster recovery plan. A **mksysb** should be taken at least monthly, and before and after any system maintenance. Additionally, I always have two disks (even on the SAN) on the system reserved for rootvg. One is active and the other is one I use to take an alt_disk_copy backup of rootvg before I make changes. You can never have enough backups!

11.2.10 A Multi-Silo Approach to Security

Security needs to be a team, multi-silo approach where everyone works together to provide the multiple layers that ensure the systems are as secure as possible. The security team should be able to provide you with the policies regarding usernames and settings. If you re integrating with AD or LDAP then most of those policies are implemented there. While it takes a little time to implement the basic security measures above, it s worth it to make your systems harder to break into. I build security in by default whenever I set up an LPAR or VIO server. It makes life much simpler to do it at that point. However, it s still relatively easy to implement later although it will most likely require a reboot.

11.2.11 References

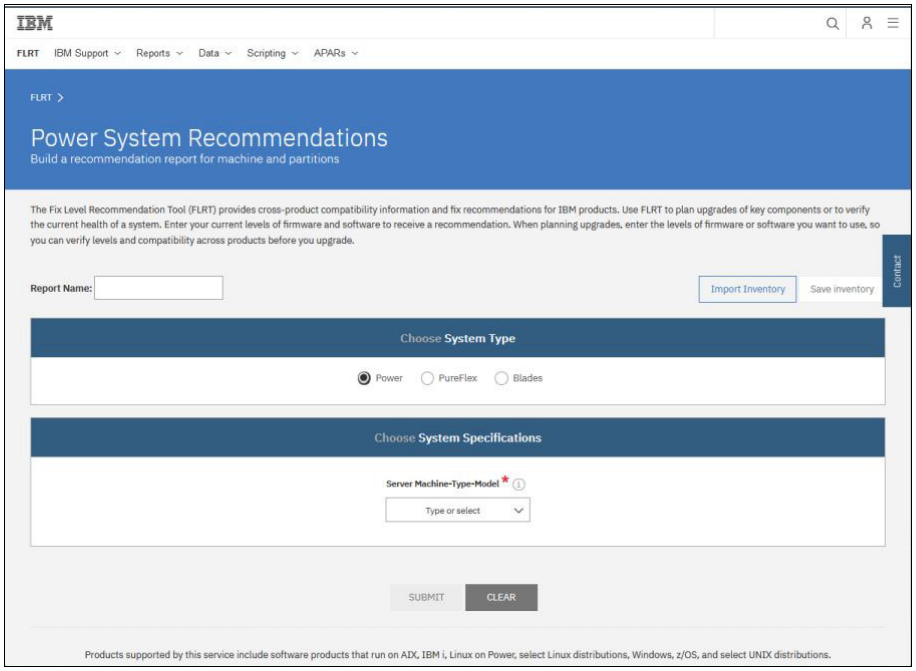
The following links will be helpful as you set up your AIX security.

- FLRT Home Page
- FLRTVC Home Page
- Apar.csv file
- FLRTVC Online Tool
- Fix Central (patches and updates)
- FLRT LITE (Check firmware and software supported levels)
- Web Applications (OpenSSH, ldap, OpenSSL, Kerberos)
- AIX Linux Toolbox

11.3 Fix Level Recommendation Tool for IBM Power

The Fix Level Recommendation Tool (FLRT) provides cross-product compatibility information and fix recommendations for IBM products. Use FLRT to plan upgrades of key components or to verify the current health of a system. Enter your current levels of firmware and software to receive a recommendation. When planning upgrades, enter the levels of firmware or software you want to use, so you can verify levels and compatibility across products before you upgrade.

Figure 11-1 shows an image of the Fix Level Recommendation Tool for IBM Power.



Power System Recommendations
Build a recommendation report for machine and partitions

The Fix Level Recommendation Tool (FLRT) provides cross-product compatibility information and fix recommendations for IBM products. Use FLRT to plan upgrades of key components or to verify the current health of a system. Enter your current levels of firmware and software to receive a recommendation. When planning upgrades, enter the levels of firmware or software you want to use, so you can verify levels and compatibility across products before you upgrade.

Report Name:

Choose System Type

● Power ○ PureFlex ○ Blades

Choose System Specifications

Server Machine-Type-Model *

Type or select

Products supported by this service include software products that run on AIX, IBM i, Linux on Power, select Linux distributions, Windows, z/OS, and select UNIX distributions.

Figure 11-1 The IBM Fix Level Recommendation Tool for IBM Power

Note: You can find the Fix Level Recommendation Tool for IBM Power at `https://esupport.ibm.com/customercare/flrt/power`.

11.4 Physical security

Often overshadowed by its digital counterpart, physical security is a critical component of safeguarding people, property, and assets. While locks and alarms are foundational, they represent just the tip of the iceberg.

Protecting hardware, data, and backup systems from damage or theft is paramount. A robust physical security framework is essential for any organization, serving as the bedrock upon which other security measures are built. Without it, securing information, software, user access, and networks becomes significantly more challenging.

Beyond internal systems, physical security encompasses protecting facilities and equipment from external threats. Building structures, such as fences, gates, and doors, form the initial defense against unauthorized access. A comprehensive approach considers both internal and external factors to create a secure environment.

11.4.1 Key Physical Security Measures: A Layered Approach

Access control encompasses the measures taken to limit exposure of certain assets to authorized personnel only. Examples of these corporate barriers often include ID badges, keypads and security guards. However, these obstacles can vary greatly in terms of method, approach and cost.

Effective physical security is essential for protecting facilities, assets, and personnel. A comprehensive strategy involves a layered approach that combines various security measures to deter, detect, delay, and respond to potential threats.

Deterrence

Discourage unauthorized access through visible security measures such as:

- Clear signage indicating surveillance
- Robust physical barriers like fences and gates
- High-quality security cameras
- Controlled access systems (card readers, keypads)

Detection

Identify potential threats early with:

- Motion sensors and alarms
- Advanced video analytics
- Environmental sensors (temperature, humidity)
- Real-time monitoring systems

Delay

Hinder intruders and buy time for response through:

- Multiple points of entry and exit
- Sturdy doors, locks, and window reinforcements
- Access control measures (biometrics, mobile credentials)
- Security personnel or guards

Response

Swiftly address security incidents with:

- Emergency response plans and procedures
- Integration of security systems with communication tools
- Trained personnel for incident management
- Collaboration with law enforcement

By strategically combining these elements, organizations can create a robust physical security framework that mitigates risks and protects critical assets.

11.4.2 Perimeter Security and Beyond

Having physical protection throughsolid building construction and perimiter protectio and controll is only part of the equiation. Maximizing the physical security measures to limit and control who has access to sites, facilities and materials is paramount. Additionally, a good physical security process include monitoring, emergency preparedness, reliable power supplies, adequate climate control, and effective system documentations.

Perimeter security forms the initial line of defense for any facility. Physical barriers like fences, gates, and surveillance systems create a deterrent against unauthorized access. Strategic landscaping and lighting can further enhance perimeter protection by improving visibility and restricting movement.

Access Control

Granting authorized access while preventing unauthorized entry is crucial. Modern access control systems, such as key cards, biometric readers, and mobile credentials, offer convenience and security. Restricted areas demand stricter controls, often incorporating multi-factor authentication and surveillance.

Monitoring and Detection

Surveillance systems, including cameras and sensors, play a vital role in detecting and deterring threats. Video analytics and sensor technology provide real-time monitoring and can trigger alerts for suspicious activity. Detailed logs of access attempts and system events are essential for incident investigation and security audits.

The Human Factor

While advanced technology is a cornerstone of modern security, human involvement is equally critical. Trained security personnel and informed employees form a powerful defense against threats. Regular security drills and comprehensive training empower staff to recognize and respond to potential dangers.

It's essential to remember that even the most sophisticated security systems are only as effective as the people who use them. Employees who understand their role in security can significantly enhance a facility's protection. By equipping your staff with the knowledge and skills to handle emergencies, you create a safer environment for everyone.

A comprehensive physical security strategy integrates these elements to create a layered defense. By combining physical barriers, access control, monitoring, and human involvement, organizations can effectively protect their assets and personnel.

A

IBM Technology Expert Labs Offerings

IBM Expert Labs is a professional services organization powered by a highly experienced team of product specialists. This knowledgeable team offers deep technical expertise across various areas, including IBM Data and AI, Automation, Sustainability, Security, Software Defined Networking, IBM Power, IBM Storage, IBM Z and LinuxONE, IBM GDPS®, and IBM Cloud.

They utilize proven methodologies, practices, and patterns to help partners develop complex solutions, achieve better business outcomes, and drive client adoption of IBM software, servers, and storage.

This chapter will focus on the Security offerings from IBM Technology Expert Labs. For more information on Technology Expert Labs broader offerings see the Technology Expert Labs website.

Security assessment for IBM Power from IBM Technology Expert Labs

If you want to avoid the pressure and confusion of ensuring that your IBM Power environment is safe and secure, there may be no better way than to have IBM secure your environment for you by employing the services of IBM Technology Expert Labs.

Engaging IBM Technology Expert Labs allows you to properly secure your IBM Power environment by having them make a proper assessment of your setup. The purpose of this services activity is to help you to assess system security on IBM Power and it provides a comprehensive security analysis of either a single AIX, IBM i, or Linux instance or a single Red Hat OpenShift cluster.

This service is designed to help you address issues that affect IT compliance and governance standards.

Assess IBM Power Security for AIX, Linux, or Red Hat OpenShift

The goal of this service is to assist the client in assessing system security on IBM Power, providing a thorough security analysis of the AIX instance, Linux instance, or Red Hat OpenShift cluster. This service is aimed at helping the client address issues related to IT compliance and governance standards.

IBM will:

1. Conduct a comprehensive security analysis of the following environments:
   - AIX
   - Red Hat Enterprise Linux
   - SUSE Linux Enterprise Server (SLES)
   - Ubuntu
   - Red Hat OpenShift v4 cluster
2. Evaluate the security configuration details of the AIX, Linux or Red Hat OpenShift system.
3. For Red Hat OpenShift, analyze security recommendations for master node configuration files, API server, controller manager, scheduler, etcd, control plane configuration, worker nodes, and kubelet configuration.
4. For AIX or Linux, review administrative privileges, logging, monitoring, vulnerability management, malware defenses, and the limitation and control of network ports, protocols, and services.
5. Provide guidance on security best practices based on the Center for Internet Security (CIS) Critical Controls and CIS Benchmarks.
6. Offer detailed recommendations for potential adjustments and remediation to enhance overall security.

Assess IBM Power Security for IBM i

The purpose of this services activity is to help a client to assess system security on IBM Power and it provides a comprehensive security analysis an IBM i environment. The service is designed to help the Client address issues that affect IT compliance and governance standards.

IBM will:

1. Perform a comprehensive analysis of the security to the IBM i environment.
2. Identify a comprehensive analysis of user profiles, special authorities, system service tool (SST) Profiles.
3. Assess authentication of password policy, default and dictionary passwords, user groups, multi-factor authentication (MFA), single-sign on (SSO).
4. Perform a comprehensive analysis of confidentiality and data access library authorities, public authority, integrated file system (IFS) root.
5. Assess system integrity and config for system values, product temporary fix (PTFs), work management.
6. Assess logging and auditing with audit journaling, syslog.
7. Assess IP network settings for example transport layer security (TLS), NetServer shares, distributed data management/distributed relational database architecture (DDM/DRDA), guest access, exit points, ransomware.
8. Provide in-depth recommendations for potential adjustments and remediation, if necessary, to improve overall security.

Other offerings

These security assessment offerings, as well as a wide range of offerings covering areas such as performance and availability among others, are provided by IBM Technology Expert Labs and are generally available worldwide. The list of IBM Power offerings available when this document was published is shown here:

- Assess IBM Power System Health
- Assess IBM Power Availability
- Assess IBM Power Performance
- Assess IBM Power Database Performance
- Assess IBM Power Security for AIX, Linux or Red Hat OpenShift
- Assess IBM Power Security for IBM i
- Assess IBM PowerVM Health
- Assess IBM Power Capacity
- Assess Oracle Licensing
- Assess IBM i Performance
- Assess DB2 Mirror for IBM i
- Plan Migration to IBM Power10
- Plan Oracle Exadata Migration to IBM Power
- Install and Configure Linux on IBM Power
- Install and Configure IBM License Management Tool (ILMT) for Software Asset Management
- Install and Configure Security and Compliance Tools for IBM i
- Build IBM Power VM Recovery Manager
- Build IBM PowerHA
- Build IBM PowerSC
- Build IBM PowerHA SystemMirror® for AIX
- Build IBM PowerHA SystemMirror for IBM i
- Build HA/DR Solution with PowerHA Tools for IBM i IASP Manager
- Build Safeguarded Copy with IBM i
- Build Cyber Vault with IBM i
- Build Full System FlashCopy and Replication for IBM i
- Migrate to IBM i Infrastructure
- Perform IBM i Security Services

- Perform AIX and Linux Security Optimization
- Perform AIX Upgrade
- IBM Expertise Connect for AIX on IBM Power
- IBM Expertise Connect for IBM i on IBM Power

The complete list of standard services offered by IBM Technology Expert Labs for IBM Power can be found at:
https://www.ibm.com/support/pages/ibm-technology-expert-labs-power-offerings

The offerings may differ according to your geographical region. For details specific to your region, it would be best to get in touch with an IBM Technology Expert Labs seller to determine details.

Security and Compliance Tools for IBM i

The IBM Technology Expert Labs team for IBM i Security is an IBM team that helps clients make the most of the IBM i purchase. They do this by offering services such as security assessments and system hardening as well as developing IBM i utilities. This family of utilities goes under the name “Security Compliance Tools for IBM i.”

The utilities range from simple to complex and complement the tools provided natively in IBM i. Each tool has its own purchase price and is available directly from IBM Technology Expert Labs. A quick summary of the tools is as follows:

- Compliance Automation Reporting Tool (CART)

  After a Security Assessment and subsequent remediation, systems must be monitored to maintain compliance. Without monitoring, the state of the system is unknown. And so, while your system might have been secure at one point in time, without ongoing monitoring you cannot be sure of your current status. While there are many security tools available, most of them do not focus on IBM i. In fact, several do not even run on IBM i nor analyze IBM i security attributes. For this reason, the IBM Technology Expert Labs security and database teams collaborated to create a tool specifically for IBM i, taking advantage of the unique features of our system. This tool provides built-in reports and dashboards for monitoring security attributes that highlight where vulnerabilities or configuration mistakes may exist.

- Advanced Authentication for IBM i

  The primary purpose of this tool is to provide a second factor that users must enter when attempting to gain access to a system. In addition to the standard user password (which should expire on a regular basis), users need to provide a unique six (6) digit code that changes every 30 seconds on a hardware token or software pap. This is known as a time-based one-time password (TOTP) and is based on RFC 6238. This forces users to not only provide something they know (their standard password) but also something they have (the hardware token or software pap). Without both items, access to the system is denied

- Syslog Reporting Manager for IBM i

  The primary purpose of this tool is to provide a simple way to extract native IBM i logs and send them to a centralized security information and event monitoring (SIEM) solution. We do this by extracting entries from various native IBM i logging facilities, transforming them into properly formatted syslog messages (per RFC 3164 and 5424), and sending them over to a central collection system. In addition to the native logs, our tool can also monitor and report on changes to IFS files.

- Network Interface Firewall for IBM i

  The primary purpose of this tool is to restrict access to various remote services on IBM i to only an approved list of users. If the user is not authorized to use that particular service, then they will be blocked. Blocking is done per service, per user and even by IP address, if desired. For example, you could allow user JSMITH access over FTP only if coming from another server.

- Privileged Elevation Tool for IBM i

  Without careful control, privileged users can pose a risk to your system security. This tool enables the security administrator to reduce privileged accounts, with a mechanism to temporarily elevate privileges to users when needed. Privileged Elevation Tool for IBM i is fully auditable and provides notifications when invoked. This allows compliance to industry guidelines on privileged users.

- Single Sign On Suite for IBM i

  IBM Technology Expert Labs Power Delivery Practice is proud to provide a suite of tools uniquely designed to help your company get started with Single Sign On (SSO) with IBM i. SSO involves setting up Network Authentication Services (NAS) and then mapping Windows user profiles to IBM i user profiles using Enterprise Identity Mapping (EIM).

- Password Validation Tool for IBM i

  The primary purpose of this tool is to provide a more complex method of password validation than the operating system alone provides. Despite warnings, one in five users choose a non-compliant password to protect their identity. We've developed a program to validate and ensure passwords meet company and industry recommended rules and guidelines.

- Certificate Expiration Manager

  The primary purpose of this tool is to provide a simple way to get notified about upcoming certificate expirations. In a modern network, TLS encryption is crucial to provide encrypted communications. But this encryption only works when the certificates are in their valid date period. Expired certificates can lead to outages in an otherwise healthy network. Because of this, staying on top of certificates is a key item. Using the Certificate Expiration Manager will notify you well in advance of your certificate expiration to allow you the time needed to ensure uninterrupted service.

For additional information on these offerings check out this link:

https://ibm.biz/IBMiSecurity

B

Ecosystem and products

One path you can choose to properly secure your environment is to employ solutions and services that can do the job for you. Though not intended to be a comprehensive list, here we enumerate some of these security solutions and services that are available on IBM Power through IBM or third-party providers.

This appendix contains the following:

BigFix (HCL Technologies)

If you choose to automate the assessment and application of security patches and fixes in your IBM Power environment, one option is to utilize an automation management solution such as BigFix. Formerly from IBM, BigFix (now owned by HCL) is a solution for endpoint and security management. If you wish to stay ahead of cyber attacks, improve workstation and server patching, proactively resolve tickets, or enhance your digital employee experience, BigFix enables you to do it from a central platform.

BigFix allows you to:

- Revolutionize Workspace and Enterprise Management

  Leverage Artificial Intelligence (AI) technologies to elevate your digital experience and automate infrastructures with seamless, secure and AI-enabled intelligent management.

- Automate IT Operations and Lower Costs

  By leveraging the world's largest library of automations and a comprehensive unified endpoint management solution, you can automate OS and software patch management and streamline management processes. BigFix can reduce IT complexity and cost by consolidating patch management, software asset management and endpoint security with a single, comprehensive, BigFix offering that supports multiple operating systems.

- Achieve and Maintain Continuous Compliance

  Automatically bring non-compliant endpoints back to a compliant state using out-of-the-box industry checklists containing over 44,000 compliance checks. Ensure continuous compliance with constant low impact monitoring and automatic remediation that protects your endpoints against cybersecurity threats and provides near real-time compliance reporting.

- Discover, Prioritize and Remediate Vulnerabilities FAST

  Leverage the world's fastest vulnerability remediation solution to discover, prioritize and mitigate critical security vulnerabilities using threat sources from MITRE and CISA and vulnerability scan data from Tenable, Rapid7, Qualys and others.

BigFix provides an AI Digital+ endpoint management platform that leverages AI to improve employee experience and intelligently automate infrastructure management. It aims to secure and manage endpoints across nearly 100 different operating systems, ensure continuous compliance with industry benchmarks, and revolutionize vulnerability management with cybersecurity analytics.

Note: You can find more information on BigFix at https://www.hcl-software.com/bigfix,

IBM QRadar Suite (Palo Alto Networks)

IBM QRadar Suite is a modernized threat detection and response solution designed to unify the security analyst experience and accelerate their speed across the full incident lifecycle. The portfolio is embedded with enterprise-grade AI and automation to dramatically increase analyst productivity, helping resource-strained security teams work more effectively across core technologies.

Acquisition by Palo Alto of QRadar Suite SaaS offerings was closed as of Sept 4, 2024. QRadar Suite SaaS offerings are to be integrated into Cortex XSIAM, see https://www.paloaltonetworks.com/cortex/cortex-xsiam.

Note: QRadar on-premises remains with IBM.

With a common user interface, shared insights and connected workflows, it offers integrated products for:

- Endpoint security (EDR, MDR)

  Endpoint detection and response (EDR) solutions are more important than ever, as endpoints remain the most exposed and exploited part of any network. The rise of malicious and automated cyber activity targeting endpoints leaves organizations struggling against attackers who easily exploit zero-day vulnerabilities with a barrage of ransomware attacks.

  IBM QRadar EDR provides a more holistic EDR approach that:

  - Remediates known and unknown endpoint threats in near real time with intelligent automation
  - Enables informed decision-making with attack visualization storyboards
  - Automates alert management to reduce analyst fatigue and focus on threats that matter
  - Empowers staff and helps safeguard business continuity with advanced continuous learning AI capabilities and a user-friendly interface

- SIEM

  As the cost of a data breach rises and cyberattacks become increasingly sophisticated, the role of security operations center (SOC) analysts is more critical than ever. IBM QRadar SIEM is more than a tool; it is a teammate for SOC analysts—with advanced AI, powerful threat intelligence and access to the latest detection content.

  IBM QRadar SIEM uses multiple layers of AI and automation to enhance alert enrichment, threat prioritization and incident correlation—presenting related alerts cohesively in a unified dashboard, reducing noise and saving time. QRadar SIEM helps maximize your security team's productivity by providing a unified experience across all SOC tools, with integrated, advanced AI and automation capabilities.

- SOAR

  The IBM QRadar SOAR platform is built to optimize your security team's decision-making processes, improve your security operations center (SOC) efficiency, and ensure your incident response processes are met with an intelligent automation and orchestration solution.

  Winner of a Red Dot User Interface Design Award, QRadar SOAR helps your organization:

  - Cut response time with dynamic playbooks, customizable and automated workflows and recommended responses
  - Streamline incident response processes by time-stamping key actions and aiding in threat intelligence and response
  - Manage incident response to over 200 international privacy and data breach regulations with Breach Response

For more information on th QRadar Suite see the QRadar Web Page.

Trend Vision One (Trend Micro)

Trend Vision One Security is a solution for IBM Power customers looking to protect across clouds, networks, devices, and endpoints with an AI-powered cybersecurity platform. With full support to run all components of the Trend Vision One platform on IBM Power, Trend Vision One aims to provide administration and DevOps teams greater control over their environment with central visibility and management. Utilizing Trend Vision One on IBM Power can help your organization modernize, simplify, and converge your security operations, enabling better protection against cyber threats across diverse hybrid IT environments.

In today's complex threat environment, the ability to stay ahead of adversaries, design for resilience, and create secure work environments is paramount. Trend Micro's XDR services are engineered to provide advanced threat defense through technologies and human intelligence that proactively monitor, detect, investigate, and respond to attacks. The IBM Power partnership ensures data is protected with comprehensive end-to-end security at every layer of the stack. These integrated security features are designed to ensure compliance with security regulatory requirements.

Trend Vision One delivers real-time insights neatly displayed on your executive dashboard. No more manual tasks—just efficient, informed decision-making. While IBM Power frees up client resources, allowing them to focus on strategic business outcomes, Trend Vision One automates cyber security reporting and playbooks for more efficient and productive security operations. Security teams can stay ahead of compliance regulations, with real-time updates ensuring their enterprise security posture remains robust.

Server and workload protection features

- Intrusion and vulnerability prevention
  - Protect your environment from attacks on known and zero-day vulnerabilities, SQL injections, cross-site scripting, and other web application vulnerabilities
  - Utilize intrusion prevention rules when patches are unavailable for known vulnerabilities in applications or operating systems
  - Intercept traffic trying to exploit unpatched vulnerabilities, keeping your assets protected until patches are released, tested, and deployed
- File integrity monitoring
  - Scan for unexpected changes to registry values and keys, services, processes, installed software, ports, and files
  - Using a baseline secure state as a reference, the integrity monitoring module scans the above and logs an event if unexpected changes are detected
- Log inspection
  - Identify significant events that might be buried in your operating system and application logs
  - Send these events to a security information and event management (SIEM) system or centralized logging server for correlation, reporting, and archiving
- SAP Scanner
  - Scan files on demand to protect critical information within SAP environments
  - Experience seamless certified integrations with both SAP NetWeaver and the SAP HANA platform
  - Analyze uploaded data and identify possible malicious script content that might be embedded or disguised within documents

  - Auto-tag and report malicious content to SAP systems through the NetWeaver Virus Scan Interface (VSI), where administrators can set or enforce policies and actions

Other features

- Antimalware
- Web reputation service
- Activity monitoring
- Activity firewall
- Application control
- Behavioral analysis
- Machine learning
- EDR and XDR
- Device control
- Virtualization protection

> **Note:** Information on Trend Vision One can be found at `https://www.trendmicro.com/en_us/business/products/one-platform.html` while the Trend Vision One on IBM Power solution brief can be found at `https://www.trendmicro.com/en_us/business/products/endpoint-security.html?modal=s7d-card-btn-ibm-power-sb-859b9e`.

Anypoint Flex Gateway (Salesforce/Mulesoft)

To fully realize the value of enterprise data, businesses have often utilized application programming interfaces (APIs) to gain the benefits. APIs improve existing products, operations, and systems, open new streams of revenue, and provide richer insights that result in enhanced business strategies and provide richer customer experiences.

To transport data through APIs, however, requires a protection layer to ensure security of data and accessibility only to known actors. Mulesoft has partnered with IBM to provide Anypoint Flex Gateway on IBM Power.

In today's digital landscape, seamless connectivity and rapid data exchange are crucial for business success. Organizations constantly seek innovative solutions to streamline operations, and MuleSoft Anypoint Flex Gateway provides that capability.

Many companies leverage Salesforce's MuleSoft to manage and secure application programming interface (APIs) across cloud-native, containerized environments. Now, IBM Power users can tap into the power of Anypoint Flex Gateway's advanced API protection layer to modernize applications and accelerate API-driven initiatives.

Empowering integration on IBM Power

IBM Power is renowned for its robust performance, reliability, and scalability. With the native integration of Anypoint Flex Gateway, businesses using IBM Power servers can leverage one of the industry-leading API Management Platform from MuleSoft to seamlessly connect diverse systems, applications, and data sources.

MuleSoft Anypoint Flex Gateway is an Envoy-based, ultra-fast, lightweight API gateway built on Envoy technology. Designed for seamless integration with DevOps and CI/CD workflows, Anypoint Flex Gateway delivers the performance needed for demanding applications and microservices, while ensuring enterprise-grade security and manageability across any environment.

This synergy unlocks new levels of agility, innovation, and efficiency for your digital transformation journey. The native integration enables smooth installation and operation of the API gateway, effectively safeguarding your IBM Power applications.

Key use cases and benefits

This approach empowers you to accelerate modernization with API-led integration. Easily enable a hybrid retail model with a container-based solution and an API integration layer or simplify SAP S/4HANA integration with other systems.

Deploying Anypoint Flex Gateway close to your IBM Power-hosted applications, APIs, and data significantly enhances the customer experience, enforces security policies, reduces data latency, and boosts application performance. You can deploy the gateway on Red Hat OpenShift, Red Hat Enterprise Linux (RHEL), and SUSE Linux Enterprise Server (SLES).

Here are the key benefits:

- Seamless connectivity: Connect seamlessly across on-premises, cloud, and hybrid environments, facilitating real-time data exchange and decision-making.
- Unified integration platform: Access a unified platform for integration and API management, streamlining development, deployment, and management of integration solutions, reducing complexity and accelerating time-to-market.
- Scalability and flexibility: Handle a few transactions or millions of events with unmatched scalability and flexibility, adapting to evolving business needs and ensuring future-proof integration solutions.
- Effortless integration: Connect applications, data sources, and devices across your IBM Power server environment with ease.
- Dynamic scaling: Scale your integration infrastructure dynamically to meet evolving business demands without compromising performance.
- Unwavering reliability: Ensure continuous operation and data integrity with resilient integration solutions designed for IBM Power servers.
- Enhanced security: Safeguard your critical assets and data with enterprise-grade security features embedded within Anypoint Flex Gateway.

By combining the strengths of MuleSoft's Anypoint Platform with the performance and reliability of IBM Power servers, businesses can confidently embark on their digital transformation journeys, equipped with the tools and capabilities to drive innovation, agility, and growth.

> **Note:** IBM and Mulesoft’s partnership announcement for Anypoint Flex Gateway on IBM Power can be found at `https://www.ibm.com/blog/announcement/ibm-and-mulesoft-expand-global-relationship/` while the solution brief can be found at `https://www.mulesoft.com/sites/default/files/cmm_files/MuleSoft_AnypointFlexGateway_IBM%20Power_0.pdf`.

Active IBM i Security Ecosystem Companies

The IBM i ecosystem includes several companies dedicated to enhancing security for IBM i environments. These companies provide a range of solutions to address various security needs, ensuring robust protection for IBM i systems. Notable companies active in IBM i security include:

Fortra (formerly HelpSystems)

Fortra offers a comprehensive suite of security solutions designed to protect IBM i environments. Their products cover areas such as data encryption, compliance management, and threat detection. Fortra's IBM i security solutions are renowned for their ease of use, robust features, and comprehensive reporting capabilities. They cater to a wide range of industries, including finance, healthcare, and retail, ensuring that their clients meet stringent regulatory requirements and safeguard critical data.

Raz-Lee Security

Raz-Lee Security specializes in providing advanced security solutions for IBM i. Their offerings include tools for real-time threat detection, audit and compliance management, and vulnerability assessment. Raz-Lee's iSecurity suite is highly regarded for its powerful and customizable security modules, which help organizations proactively manage and mitigate security risks. Their customer base spans various sectors such as banking, insurance, manufacturing, and government, reflecting their ability to address diverse security challenges across different industries.

Precisely

Precisely provides a range of IBM i solutions aimed at ensuring data integrity, availability, security, and compliance. Their IBM i security solutions include tools for access control, monitoring, privacy, and malware defense. Precisely is known for its robust, scalable solutions that can integrate seamlessly into existing IT infrastructures. These solutions deliver market-leading IBM i security capabilities that help organizations successfully comply with cybersecurity regulations and reduce security vulnerabilities. In addition, the security offerings seamlessly integrate with Precisely's IBM i HA solutions to deliver an even greater level of business resilience. Precisely's customers ranges from large enterprises to SMB in sectors like telecommunications, financial services, and logistics.

Precisely also offers a free assessment tool for IBM i. Assure Security Risk Assessment checks over a dozen categories of security values, compares them to recommended best practices, reports on findings, and makes recommendations. You can find this security risk assessment at: https://www.precisely.com/product/precisely-assure/assure-security.

Fresche Solutions

Fresche Solutions offers a comprehensive IBM i Security Suite designed to protect IBM i systems from modern security threats. Their solutions include tools for real-time monitoring, vulnerability assessment, and compliance management. Fresche's security suite is noted for its innovative approach to security management, combining ease of deployment with powerful analytical capabilities. Their customer base includes businesses of all sizes, from SMBs to large enterprises, in industries such as retail, manufacturing, and services, demonstrating their versatile and scalable security offerings.

These companies, among others, play a vital role in the IBM i ecosystem by continuously innovating and providing security solutions tailored to the unique needs of IBM i users. Their diverse customer bases and strong industry reputations underscore their effectiveness in delivering reliable, high-quality security solutions.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *Security Implementation with Red Hat OpenShift on IBM Power Systems, REDP-5690*
- *Implementing, Tuning, and Optimizing Workloads with Red Hat OpenShift on IBM Power*, SG24-8537
- *IBM Storage DS8000 Safeguarded Copy: Updated for DS8000 Release 9.3.2*, REDP-5506
- *Data Resiliency Designs: A Deep Dive into IBM Storage Safeguarded Snapshots*, REDP-5737.
- *IBM Power Systems Cloud Security Guide: Protect IT Infrastructure In All Layers, REDP-5659*
- *Implementing, Tuning, and Optimizing Workloads with Red Hat OpenShift on IBM Power, SG24-8537*
- *Introduction to IBM PowerVM*, SG24-8535

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Online resources

These websites are also relevant as further information sources:

- Cloud Management Console Cloud Connector Security White Paper

  https://www.ibm.com/downloads/cas/OGGYD90Y
- IBM AIX Documentation on Security

  https://www.ibm.com/docs/en/aix/7.3?topic=security
- Red Hat OpenShift Documentation on Configuring your Firewall

  https://docs.openshift.com/container-platform/4.12/installing/install_config/configuring-firewall.html
- Modernizing Business for Hybrid Cloud on OpenShift Video Series

  https://community.ibm.com/community/user/power/blogs/jenna-murillo/2024/01/29/modernizing-business-for-hybrid-cloud-on-openshift

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

IBM Power Security Catalog

SG24-8568-00

ISBN

(1.5” spine)
1.5”<-> 1.998”
789 <->1051 pages

Redbooks

IBM Power Security Catalog

SG24-8568-00

ISBN

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

Redbooks

IBM Power Security Catalog

SG24-8568-00

ISBN

(0.5” spine)
0.475”<->0.873”
250 <-> 459 pages

Redbooks

IBM Power Security Catalog

(0.2”spine)
0.17”<->0.473”
90<->249 pages

(0.1”spine)
0.1”<->0.169”
53<->89 pages

Redbooks

IBM Power Security Catalog

SG24-8568-00

ISBN

IBM

(2.5” spine)
2.5”<->nnn.n”
1315<-> nnnn pages

Redbooks

IBM Power Security Catalog

SG24-8568-00

ISBN

IBM

(2.0” spine)
2.0” <-> 2.498”
1052 <-> 1314 pages

SG24-8568-00

ISBN

Printed in U.S.A.