IBM

IBM System Storage N series Reference Architecture for Virtualized Environments

- Learn how to integrate VMware vSphere with IBM System x and N series
- Understand the importance of storage as a foundation for a cloud solution
- Design VMware vSphere solutions with N series

Roland Tretau
Jacky Ben-Bassat
Michel Chalogany

Redpaper

ibm.com/redbooks

International Technical Support Organization

IBM System Storage N series Reference Architecture for Virtualized Environments

December 2012

REDP-4865-00

Note: Before using this information and the product it supports, read the information in “Notices” on page v.

First Edition (December 2012)

This edition applies to the IBM System Storage N series portfolio and Data ONTAP 8.1.1 7-mode as of November 2012.

This document was created or updated on January 25, 2013.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

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, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: 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 states 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 supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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.

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 the names and addresses used by an actual business enterprise 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.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

BladeCenter®
IBM®
RackSwitch™
Redbooks®
Redpaper™
Redbooks (logo) ®
System Storage®
System x®
Tivoli®
VMready®
X-Architecture®

The following terms are trademarks of other companies:

Intel Xeon, 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.

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

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

Preface

This IBM® Redpaper™ publication provides deployment guidelines, workload estimates, and preferred practices for clients who want a proven IBM technology stack for VMware environments. The result is a Reference Architecture for Virtualized Environments (RAVE) that uses VMware vSphere, IBM System x® or IBM BladeCenter® server, IBM System Networking, and IBM System Storage® N series as a storage foundation. The reference architecture can be used as a foundation to create dynamic cloud solutions and make full use of underlying storage features and functions.

This paper provides a blueprint that illustrates how clients can create a virtualized infrastructure and storage cloud to help address current and future data storage business requirements. It explores the solutions that IBM offers to create a storage cloud solution addressing client needs. This paper also shows how the Reference Architecture for Virtualized Environments and the extensive experience of IBM in cloud computing, services, proven technologies, and products support a Smart Storage Cloud solution that is designed for your storage optimization efforts.

Clients face many common storage challenges, and some clients have variations that make them unique. With RAVE, a proven and scalable solution has been created that consists of a rack of servers, storage, and networking components. Thus, we have carefully sized three scenarios, *Entry*, *Mainstream*, and *Advanced*, each based on preferred practices for real world workloads.

When used as the storage foundation, the IBM System Storage N series offers unified storage solutions. These solutions provide industry-leading technologies for storage efficiencies, instantaneous virtual machine and data store cloning for virtual servers and virtual desktops, and virtual data center backup and business continuance solutions.

This paper is for anyone who wants to learn how to successfully deploy a virtualized environment. It is also written for anyone who wants to understand how IBM addresses data storage and compute challenges with IBM System Storage N series solutions with IBM servers and networking solutions. This paper is suitable for IT architects, business partners, IBM clients, storage solution integrators, and IBM sales representatives.

The team who wrote this paper

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center.

Roland Tretau is an Information Systems professional with more than 15 years of experience in the IT industry. He holds Engineering and Business Masters degrees, and is the author of many storage-related IBM Redbooks® publications. Roland's areas of expertise range from project management, market enablement, managing business relationships, product management, and consulting to technical areas including operating systems, storage solutions, and cloud architectures.

Jacky Ben-Bassat is an IT architect at NetApp who currently focuses on worldwide market and business development as part of the relationship between IBM and NetApp. For the past six years, Jacky has participated in developing multiple cloud and other solutions around the globe with IBM STG, GTS, and IBM Software Group and GBS. The solutions developed

included N series storage, virtualization components from VMware, Microsoft and Citrix, and IBM Tivoli® management and orchestration tools. Prior to joining NetApp, Jacky held senior SI, IT, and IM positions for 15 years, which included responsibilities for pre-sales engineering, professional services, IT operations, infrastructure and strategy. He worked closely with senior business leaders to align IT with various transformation initiatives across the globe.

Michel Chalogany is a Technical Partner Manager at NetApp. He is responsible for managing the technical relationship between IBM and NetApp in EMEA. In his current role, he provides a high level of technical support including sales, pre-sales and technical workshops to IBM and its Business Partner community. For the past five years, Michel has spent significant time on various enablement activities around NetApp technology and solutions. He has also contributed to increasing IBM N series awareness in the marketplace through roadshows, technical refresh campaigns, and solutions designed around virtualized environments. Prior to working at NetApp, Michel held several pre-sales and systems engineering roles in enterprise systems management, application service management, network performance solutions, data protection, and data management solutions.

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

Bertrand Dufrasne
IBM International Technical Support Organization, San Jose Center

Tony Huynh, Oren Yehudai
NetApp

Stephan Fleck
IBM

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 ITSO 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 client 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 papers to be as helpful as possible. Send us your comments about this paper 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, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

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

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

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

1

Architecture

This chapter describes the Reference Architecture for Virtualized Environments (RAVE) using VMware vSphere, IBM System x Server/BladeCenter, IBM System Networking, and IBM System Storage N series as a storage foundation. This architecture can be also be used as a foundation to create dynamic cloud solutions, and make full use of underlying storage features and functions. Sample workloads and an approach and details of storage sizing are also provided.

This chapter includes the following sections:

- Architecture overview
- Architectural approach
- Configuration and components
- Solution classification

1.1 Architecture overview

The IBM System x and IBM System Storage N series with VMware vSphere as virtualization software offer an architecture that can be easily sized to fit the needs of the SMB to enterprise clients. Each selected architecture is a predefined, robust solution with immediate usability.

System x and N series with VWware vSphere 5 is a tested architecture and end-to-end solution that comprises all parts of a future-ready data center in a modular, extensible, flexible, and scalable rack server format. This format optimizes value in terms of providing dynamically changeable capability based on changing user needs, ease of deployment, administration, support, expansion, and most importantly, affordability.

To support business agility, IT teams must have the ability to add or reduce resources such as memory and disk storage capacity when required, in order to accommodate business demands by users. IBM System x and IBM System Storage N series support this dynamic scaling of resources. Furthermore, the combination of System x, N series, and VMware vSphere creates mutual benefits and enables the creation of a dynamic cloud computing architecture.

In addition to VMware vSphere, the following key infrastructure components provide the foundations for this cloud computing architecture:

- IBM System x3650M4, HS23, and HX5 servers are optimized to meet the virtualization management and advanced workload demands of private cloud data centers.
- The IBM N series N3xxx and N62xx Storage Systems were selected to bring high performance SAN and NAS features in a unified system with efficiency features including deduplication and thin provisioning. In addition, N series products offer data protection features for vSphere such as no-performance-impact snapshots and thin replication for enabling and supporting disaster recovery and business continuity.
- IBM Storage Networking Gbps Ethernet components form the backbone for data and network connectivity. The combination of high performance adapters and low latency, cut-through switches enables the high-speed infrastructure that is critical for resource utilization and load balancing within the cloud. To provide additional flexibility, clients can choose either IBM Fibre Channel (FC) or Ethernet switches for their deployment.

Performance is a key consideration to support the virtualization and elasticity requirements. Reliability and high availability are of paramount importance when designing architectures that offer cloud services. This chapter provides information about three reference configurations, all based on a common architecture, that incorporate these critical requirements for the successful delivery and ongoing operation of various solutions.

1.2 Architectural approach

Business agility in today's world relies on IT infrastructure more than ever before due to the fast pace of changing demand. The higher demand for changes dictates the need for flexible infrastructures that are scalable, reliable, and economical. The challenge for IT is to build and maintain infrastructures that can cater to these business necessities.

Modern IT infrastructures have moved from isolated designs, through virtualized environments, to internal and external cloud solutions. The architectural approach for IT infrastructures needs to address hardware and software scaling to address today’s business demands.

The cloud-based architectural approach illustrated in Figure 1-1 provides a fast and predictable way to scale IT infrastructures. It describes the transformation from application-based silos with dedicated hardware to a shared infrastructure that enables cloud solutions through the phase of adopting virtualization.

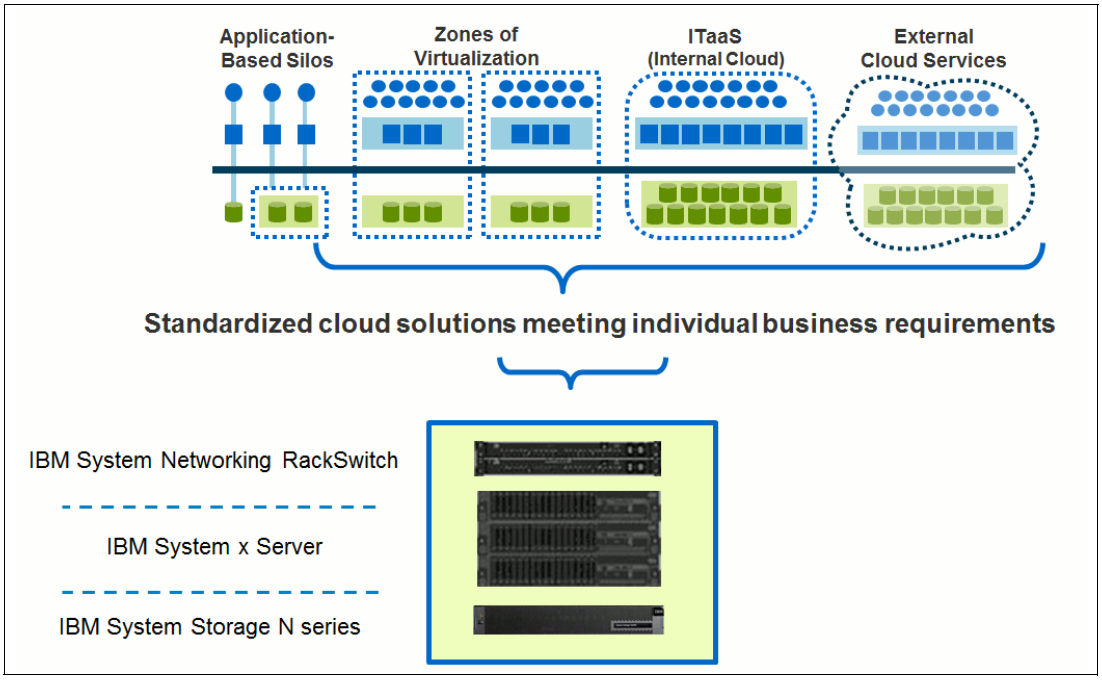

Figure 1-1 Cloud-based architectural approach

This paper provides guidelines and details about the architecture that will benefit most businesses. As mentioned, three different configurations are provided and they are all based on VMware vSphere, x Series servers, IBM System Storage N series and IBM networking components. But although these configurations share the same architecture, each one is designed to meet a different size of workload and different performance requirements, which correlate to organizations of different sizes. The three configurations are:

- Entry configuration
- Mainstream configuration
- Advanced configuration

This approach offers reasonably simple scaling capabilities. To dynamically scale with this architecture, IT can simply add additional VMware ESXi servers. Or if your data growth requires additional storage capacity or performance, you can add more storage shelves. Upgrading the storage controllers is also a simple task.

Thanks to its unified architecture, the IBM System Storage N series offers unique scalability in multiple dimensions, both out and up. In some cases, it might be more economical to scale by adding more storage controllers, versus the scale up approach of upgrading to a stronger storage model that can support higher capacity and better performance.

The decision of scaling up or out will depend on a comprehensive understanding of the full operational complexity of the environment. This will include the performance characteristics, the management aspects, the integration between the hardware component of the infrastructure and the virtualization layer, and the specific ability of the application. In some cases, infrastructure cost is less expensive when scaling out by using multiple smaller servers and storage units, but the management cost and operational limitation might offset those cost savings.

1.3 Configurations and components

The set of components consists primarily of software from VMware and hardware from IBM. The Entry and Mainstream configurations use IBM System x3650 M4 compute nodes. The Entry configuration is approximately half the compute capacity of the Mainstream configuration. Businesses starting with the Entry configuration can scale up to the Mainstream configuration in response to an increase in demand. The Advanced configuration uses IBM System HX5 compute nodes for higher performance and support for a larger number of virtual machines.

As shown in Figure 1-2 (the Mainstream configuration), the hardware building blocks consist of the following items:

- Networking components
- Compute nodes
- Storage subsystem

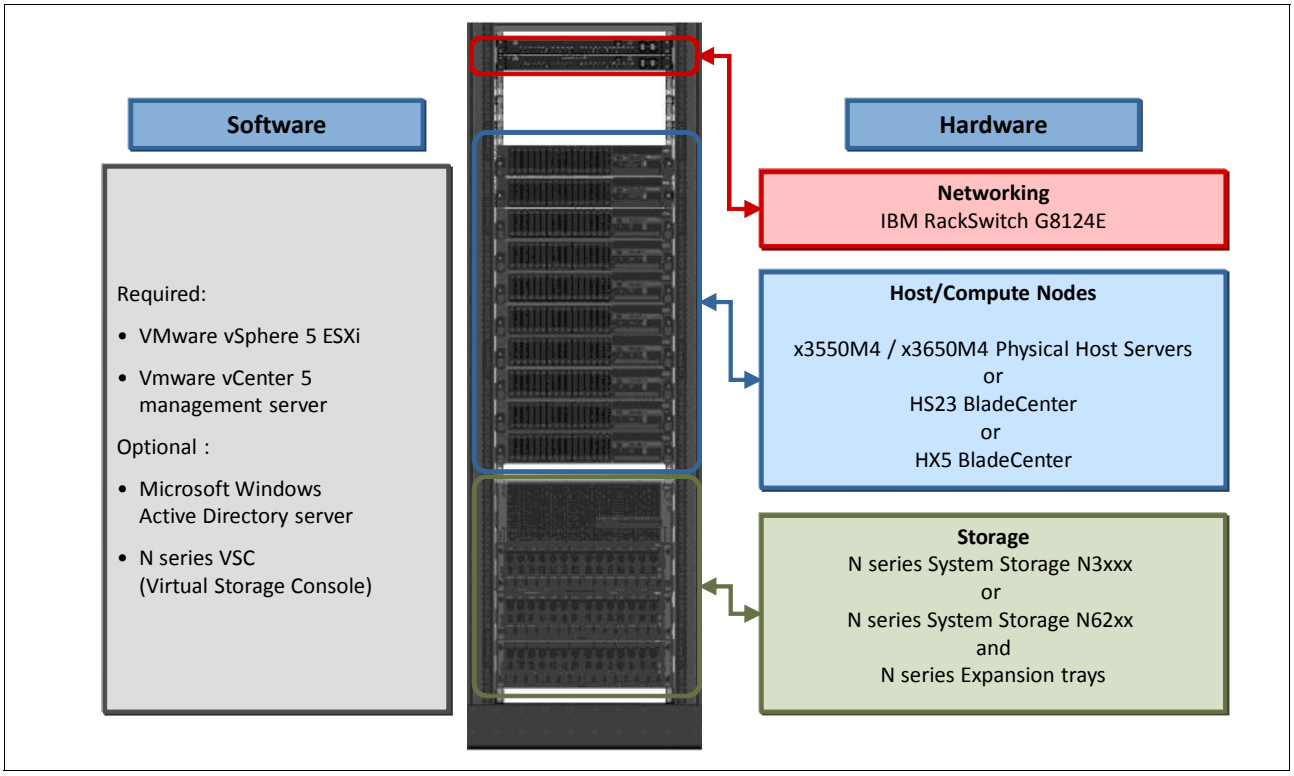

Figure 1-2 Architectural components

Next, we examine the three main configurations in the reference architecture in more detail:

- Entry configuration

  The Entry configuration is suitable for small organizations with dynamic IT infrastructure needs. In our reference, the Entry configuration supports about 300 users and uses a System x3550M4 or x3650M4 server. This provides seamless scale up capabilities such as VM support or storage.

  This configuration can consist of the same components as the Mainstream configuration, but with half the number of computer servers and with entry level storage, the N3150 or the N 3220/N3240. These are more economical choices for small organizations because they offer the same unified architecture with the full range of features at a low price point.

- Mainstream configuration

  The Mainstream configuration is a superset of the Entry configuration, with a seamless scale up of the Entry configuration for environments that require additional resources.

  The Mainstream configuration is suitable for larger organizations than the Entry configuration, and in our reference we show an example of a workload that supports 1500 users. The component building blocks of the Mainstream and Entry configurations are the same.

  However, the Mainstream configuration can also be composed using the HS23 server. The N6240 is the suggested storage model for this configuration.

- Advanced configuration

  The Advanced configuration is suitable for larger organizations with increased demands regarding capacity, speed, and concurrent users. The Advanced configuration uses HX5 servers as the host node servers. It supports about four times the number of VMs as the Mainstream configuration.

  As a stand-alone implementation, the Advanced configuration uses the latest N6270 storage system. The N6270 offers the second highest capacity and performance in the N series line (second only to the N7950T). In addition to scalability and performance, the Advanced configuration offers greater flexibility in terms of consolidating multiple different workloads into a single, highly available storage platform.

1.4 Solution classification

In this section we discuss solution classification and the characteristics of different workloads.

1.4.1 Solution classification overview

As infrastructures become increasingly complex and heterogeneous, the need for more cost-effective end-to-end solutions that are easier to manage is also growing. IT organizations are seeking solutions that can manage and scale across the entire infrastructure. So it is important to develop an architecture that easily adapts to clients'

businesses needs and supports dynamic growth. Figure 1-3 gives an overview of the reference architecture that enables clients to handle today's IT demands.

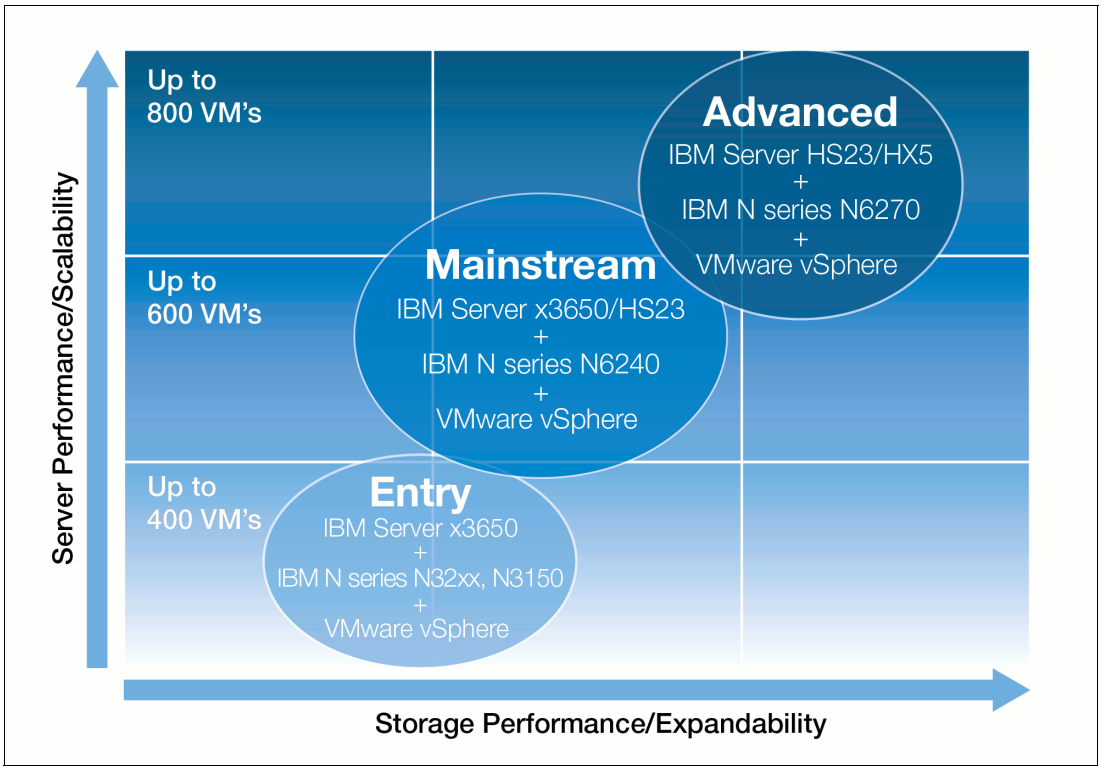

Figure 1-3 Reference architecture overview

Table 1-1 lists the various load characteristics based on the solution components introduced earlier.

Table 1-1 Reference architecture solution classification

	Entry	Mainstream	Advanced
Networking	IBM G8124E	G8124E or BladeCenter	G8124E or BladeCenter
IBM server platform[a]	x3550M4 or x3650M4	x3650M4 or HS23	HS23 or HX5
Storage[b]	N3150/N3220/N3240	N6240	N6270
Range of VMs (estimated)[c]	Up to 400 Up to 250 for N3150[d]	Up to 600	Up to 800

a. To provide additional flexibility, clients can choose either IBM Fibre Channel (FC) or Ethernet switches for their deployment.
b. All N series attachments are through FCP or NFS.
c. Range of VMs is based on mixed workload deployments.
d. N 3150 iSCSI or NFS.

Estimating hardware requirements: The guidance provided in this paper regarding estimation of hardware requirements of this cloud solution are assumptions based on empirical data gathered from existing installations. Individual performance requirements might vary based on individual client environments and applications.

1.4.2 Sample workloads

There are many performance counters that can become important when sizing VMware environments. Note that this section is not a sizing guide, and that the sizing information described is simply a summary of the elaborate sizing calculations incorporating the various parameters that are required to size each of the workloads described here.

Obtaining sizing support and guidance: Sizing practices and calculations are complex. As in our examples, clients should engage with IBM pre-sales engineers, who are familiar with various sizing tools and preferred practices, to identify the various sizing parameters for each workload, application, and task performed on the servers. Sizing calculations should also incorporate additional room for growth and for background tasks such as storage deduplication and data replication.

To match the Entry, Mainstream, and Advanced architecture classifications, we used sample workload from popular client applications:

- VMware View 4.5
- Microsoft Exchange 2010
- Microsoft SharePoint 2010
- Microsoft SQL Server 2008R2

These applications are meant to be reference configurations only, to help explain the sizing approach and methodology. Clients might end up with a different mix of applications and a different number of users for their specific workloads, but the sizing approach methodology will be the same.

Additional workloads: In addition to these sample application workloads, all solutions will be able to provide file-based CIFS or NFS storage for the environment. When designing a solution, be sure to also consider the impact of such additional workloads.

Following are the three sample workloads and their simplified storage considerations.

Table 1-2 lists a sample Entry level workload.

Table 1-2 Entry level sample workload

Application	Total users	Active users	IOPS per Active user	Total IOPS	kb/s per Active User	Total MB/s
VMware View	300	210	12	2520	70	14.36
MS SQL Server	300	150	3	450	52	7.62
MS Exchange	300	210	2	420	25	5.13
MS SharePoint	300	60	1.5	90	2	0.12
SUM				3480		27.22

Table 1-3 lists a sample Mainstream level workload.

Table 1-3 Mainstream level sample workload

Application	Total users	Active users	IOPS per Active user	Total IOPS	kb/s per Active User	Total MB/s
VMware View	1500	1050	12	12600	70	71.78
MS SQL Server	1500	750	3	2250	52	38.09
MS Exchange	1500	1050	2	2100	25	25.63
MS SharePoint	1500	300	1.5	450	2	0.59
SUM				17400		136.08

Table 1-4 lists a sample Advanced level workload.

Table 1-4 Advanced level sample workload

Application	Total users	Active users	IOPS per Active user	Total IOPS	kb/s per Active User	Total MB/s
VMware View	2500	1750	12	21000	70	119.63
MS SQL Server	2500	1250	3	3750	52	64.48
MS Exchange	2500	1750	2	3500	25	42.72
MS SharePoint	2500	500	1.5	750	2	0.98
SUM				29000		226.81

Several assumptions were made in these samples regarding the number of active users and other sizing parameters, and they are the same across all three workloads. For details about the storage sizing assumptions and the approach used, see 1.4.3, “Storage sizing approach and details” on page 8.

Remember: Your client-specific workloads might vary from these samples, and different performance metrics might apply. For guidance, and prior to any implementation, work with a pre-sales engineer to determine which configuration and sizing is correct for your enterprise.

1.4.3 Storage sizing approach and details

Each of the sample workloads and configurations included dual N series controllers (HA pair) for high availability in active-active mode. This meant that the entire workload was shared by the two storage controllers. If one of the controllers failed, then the second, active controller would assume ownership of all disks and connections. Thus, it would assume ownership over the total combined workload, which is known as *controller failover mode*.

In the sample workloads, it is assumed that the storage system is able to sustain the full workload even in controller failover mode when only one storage controller is active. This assumption translates to limiting the total workload per individual storage controller to less than 50 percent utilization, with some spare to be added for unforeseen spikes in workloads and limited higher utilization during initial loads.

And although it is also important to account for storage tasks such as deduplication and data replication, it is preferable to schedule such tasks for off-business hours. As a result, it is

unlikely that the utilization consumption of such tasks will compete with regular workload during business hours.

Clients, with the help of an IBM or IBM Business Partner pre-sales engineer, might conclude that higher utilization per controller is acceptable. For example, the storage systems can be sized for 60 - 70 percent utilization during steady state (per each of the two active-active controllers). This will provide for “headroom” for some internal storage processes, higher utilization during initial loads, and some room for growth.

In case of a controller failure, the remaining active controller will have to support up to 140 percent utilization. As a result, some performance impact is to be expected. A client’s IT staff, together with the lines of business managers, will have to determine whether the performance impact is tolerable for the brief duration of the failure period. If so, then much higher utilization can be acceptable and the figure given in 1.4.2, “Sample workloads” on page 7 can be much higher, or even doubled in some cases.

Note that the total capacity served and the number of connections will not be impacted; only users will likely experience slower response time, due to higher latency. The IBM System Storage N series provides a Quality of Service (QoS) feature known as FlexShare. This feature allows storage administrators to assign different priorities to different workloads and thereby minimize the impact to users during a controller failover. Critical applications will get higher priority.

Use individual client requirements to determine the correct storage utilization. The guidelines are common and include Total Cost of Ownership (TCO), acceptable and expected performance, service level agreements, business risks such as impact on productivity, project delays, and maintenance and support that will minimize the impact to only few hours in some cases.

When sizing the sample workloads, we made the following additional assumptions. We also list here the considerations that clients should keep in mind when their environments are being sized.

- For the performance of users and VMs, we used sizing best practices and empirical data gathered from existing installations, which are commonly used in the industry.

  However, in reality clients might find that their specific requirements are higher or lower. Therefore, a proper sizing calculation should be conducted for the specific workloads and environment of each client.

- When sizing storage, evaluate the following parameters and account for them across all workloads and for each user and VM type:
  - Required (usable) capacity on day one and over time (when considering Thin Provisioning)
  - IOPS and/or KB/s / MB/s
  - Latency
  - Random, sequential, or mixed workloads
  - Read/Write ratio
  - Maximum users
  - Concurrent users - in our samples, certain assumptions were made and the results were reflected in the active users
  - Storage tasks like data replication (SnapMirror), Deduplication and other housekeeping tasks

- Protocols - In the sample workloads, we used multiple protocols to offer the most efficient operation. However, different protocols have different performance characteristics, so accurate sizing is more complex and requires thorough calculations.

- For the sample workloads on N 6240 and N 6270, a Flash Cache module of 512 GB was used as part of the configuration.

To determine the specific value of all the required storage parameters, the pre-sales engineer, as part of the sizing calculation, will consider key parameters for each application that is part of a client's environment. The value for these parameters will be determined based on input from the client, best practices, and output from special sizing tools. In essence, sizing is an accurate exercise that accounts for all known variables but also includes reasonable assumptions that are validated by the users for some unknown parameters such as concurrent users, expected growth, or increased performance requirements.

In our case, we used common enterprise applications as described in 1.4.2, "Sample workloads" on page 7. Both Microsoft and VMware provide detailed information about those required parameters. Links to their resources are listed in "Related publications" on page 57.

2

vSphere and N series integration

This chapter highlights the benefits of using IBM System Storage N series storage with VMware vSphere. For detailed information about the introduced N series software features and for information about VMware vSphere itself, see the resources referenced in “Related publications” on page 57.

This chapter includes the following sections:

- Introduction to vSphere and N series integration
- Multiprotocol capability for datastores
- Flash Cache
- Virtual Storage Tier
- Virtual Storage Console
- FlexClone
- FlexShare
- Snapshots
- Storage configuration
- Using deduplication or compression with VMware

2.1 Introduction to vSphere and N series integration

With unmatched storage efficiency, performance and data protection capabilities, N series storage solutions complement the manageability, utilization, and cost-saving benefits of VMware virtualization software. N series products integrate with VMware, including the vStorage API for Array Integration and the vCloud suite, to deliver cost-effective, flexible, and scalable data center solutions that can respond immediately to changing business needs.

IBM System Storage Systems N series storage, from entry-level arrays to large enterprise systems, offers unmatched business agility, superior application uptime, simplicity of management, and breakthrough value for your enterprise, as explained here:

- Flash Cache optimizes the performance of random read-intensive workloads without the complexity of another storage tier.
- Virtual Storage Console (VSC) is a vCenter plug-in that provides access to advanced storage capabilities such as cloning, backup, recovery, and disaster recovery. It also simplifies the configuration and management of N series arrays.
- The Rapid Cloning Utility (RCU) is part of the Virtual Storage Console. RCU allows you to quickly and efficiently create, deploy, and manage the lifecycle of virtual machines (VMs) from an easy-to-use interface integrated into VMware vCenter 4.0 and later.
- FlexVol technology enables IT organizations to create multiple flexible volumes on a large pool of disks, and aggregate I/O across all physical disks on NFS, iSCSI or Fibre Channel. Users can dynamically allocate storage and add storage capacity as needed without downtime.
- FlexClone provides instant clones of virtual machines without impacting server resources and requiring additional storage capacity.
- FlexShare enables administrators to assign five priority levels to share. The levels can be set on a volume-by-volume basis. FlexShare allows you to prioritize specific data sets and change priorities as often as needed.
- Snapshot allows VMware users to make fast and space-efficient backup copies of active data, with negligible performance and storage overhead.
- SnapMirror Software protects data and accelerates recovery by providing a simple, efficient way to replicate data between systems in multiple locations using multi-transport (IP and FC) and frequency choice (sync, semi-sync, and a-sync) options, while only transferring changed blocks.
- *Deduplication* (wherein multiple instances of the same information are stored into a single point, and a pointer is used to refer to it on the next occurrence, so that files which potentially might be stored in an environment many times are stored only once) is ideal for VMware environments because the VMs deduplicate extremely well. The N series provides block-level deduplication within the entire flexible volume.
- Data compression is a new feature that compresses data as it is written to N Series storage systems or the N series gateway. Like deduplication, N series data compression works in both SAN and NAS environments, and is application- and storage tier-agnostic. Data Compression can be set to work inline, or as a post process.
- Virtual Storage Tier is a self-managing data-driven service layer for storage infrastructure. It provides real-time assessment of workload-based priorities and enables I/O data requests to be optimized for cost and performance without requiring complex data classification. Because only hot data blocks are promoted (not large chunks of data that might include both hot and cold blocks), Flash technology can be used sparingly without the need for excessive data movement. This reduces the processing burden on system resources and reduces traffic in the disk subsystem. Data promotion is performed on

demand based on actual usage patterns, enabling an immediate response to changing workload demands.

2.2 Multiprotocol capability for datastores

The N series storage system provides flexibility in the method and protocol used to connect to storage. Each method and protocol has benefits and disadvantages, depending on the existing solution and VMware environment requirements.

Traditionally, most VMware scenarios use standard Fibre Channel SAN connectivity. With N series, you can keep using this method if it is already in the environment. However, fiber connectivity can be expensive if new purchases are required. For this reason, more environments are now implementing network connectivity methods to storage. Such methods include iSCSI, Network File System (NFS), and Common Internet File System (CIFS), as illustrated in Figure 2-1.

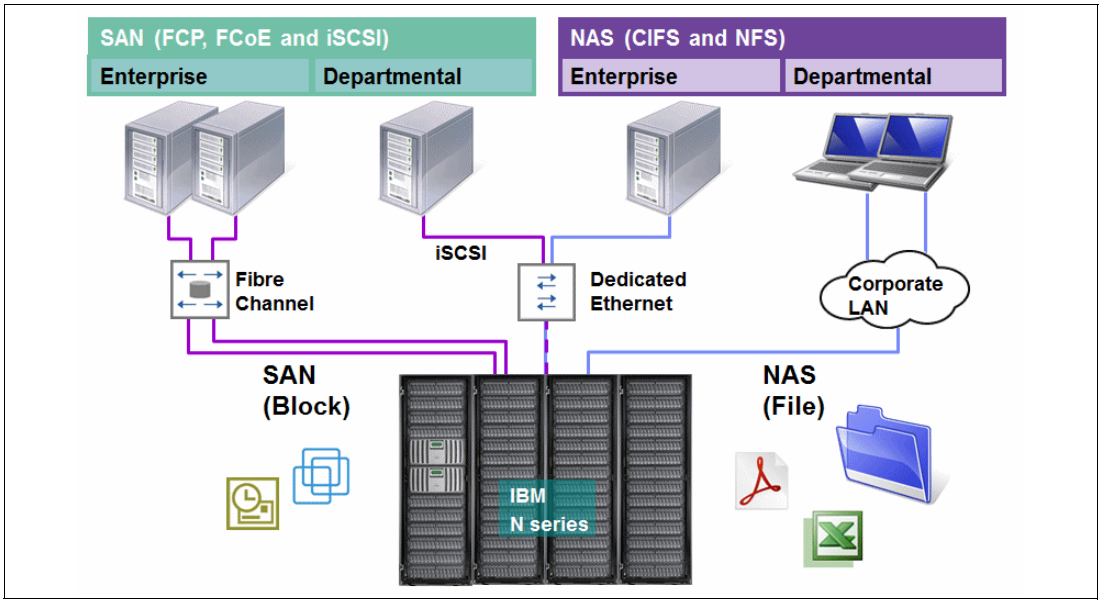

Figure 2-1 Storage protocols used by VMware and available on the N series family

Currently, VMware vSphere supports FC, iSCSI, and NFS connections for datastores. All three methods are fully supported by N series systems. In addition, N series systems offer native Fibre Channel over Ethernet (FCoE) SAN solutions, which is an end-to-end Ethernet storage solution that delivers increased data center efficiency and reduced overall cost.

Figure 2-2 further exemplifies how N series key technologies reduce storage costs and capacity requirements. As a combined hardware and software strategy, these are the technologies that help improve clients' storage efficiency.

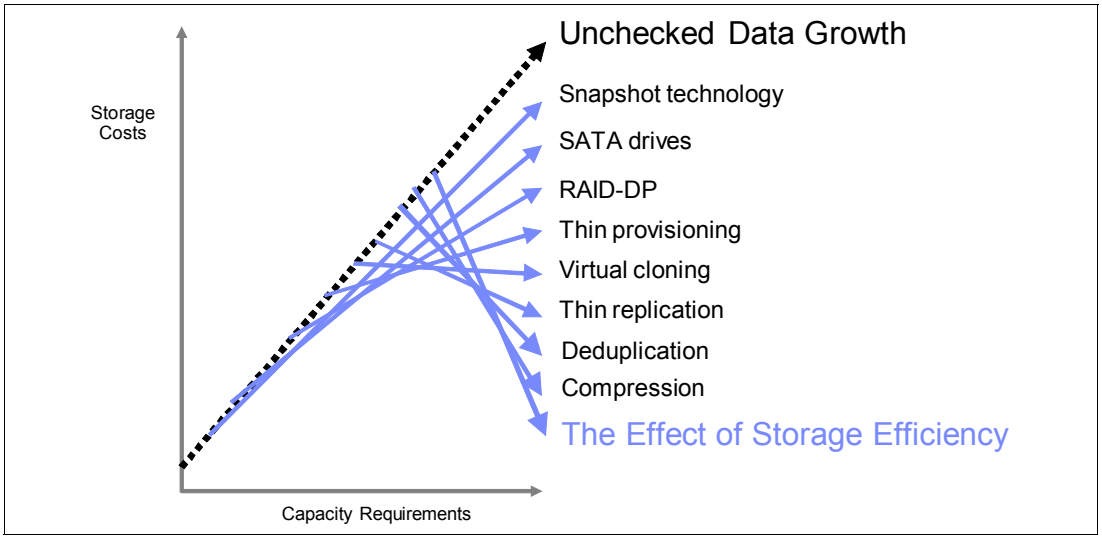

Figure 2-2 Storage efficiencies technologies

The diagonal dotted line in the graph represents the storage growth experienced in many data centers. As clients' data continues to grow at a rapid pace, they will not have to grow their storage footprint at anywhere near that pace. All these features are available for virtualized environments, and they enable clients to grow at a much slower pace because of the cumulative effect of all these technologies. N series solves many data center challenges though a unified and flexible storage management approach.

2.3 Flash Cache

Flash Cache (previously called PAM II) is a set of solutions that combine software and hardware within IBM N series storage controllers to increase system performance without increasing the disk drive count. Flash Cache is implemented as software features in Data ONTAP and PCIe-based modules with either 256 GB, 512 GB, or 1 TB of Flash memory per module. The modules are controlled by custom-coded Field Programmable Gate Array processors. Multiple modules may be combined in a single system and are presented as a single unit. This technology allows sub-millisecond access to data that previously was served from disk at averages of 10 milliseconds or more.

Tip: This solution is suitable for all types of workloads but provides the greatest benefit from IBM System Storage N series storage subsystems serving intensive random read transactions.

2.3.1 Flash Cache module

The Flash Cache option offers a way to optimize the performance of an N series storage system by improving throughput and latency while reducing the number of disk spindles/shelves required and the power, cooling, and rack space requirements.

A Flash Cache module provides an additional 256 GB, 512 GB or 1 TB (PAM II) of extended cache for your IBM System Storage N series storage subsystem, depending on the model. Up to eight modules can be installed. Each module must be installed on a PCI express slot, and it only consumes an additional 18 watts of power per module. Extra rack space and ventilation is not required, making it an environmentally friendly option. Figure 2-3 shows the Flash Cache module.

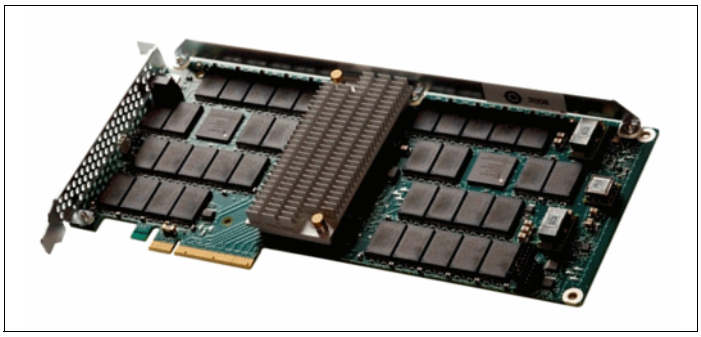

Figure 2-3 Flash Cache module

2.3.2 How Flash Cache works

Flash Cache replaces disk reads with access to an extended cache contained in one or more hardware modules. Your workload is accelerated in direct proportion to the disk reads replaced. The remainder of this chapter describes different workloads and how they are accelerated. It also explains how to choose and configure the best mode of operation, and how to observe Flash Cache at work.

2.4 Virtual Storage Tier

The N series Virtual Storage Tier offers a unique approach to automated storage tiering. Enabled by our foundational strengths in storage efficiency and intelligent caching, the Virtual Storage Tier (Figure 2-4 on page 16) provides the following benefits:

- A real-time, data-driven response to your most demanding application workloads
- The ability to consolidate your data onto fewer storage tiers
- Industry-leading efficiency through integration of data deduplication and thin cloning
- Ready for immediate use
- Automated support for PCI-e Flash and SSD technologies

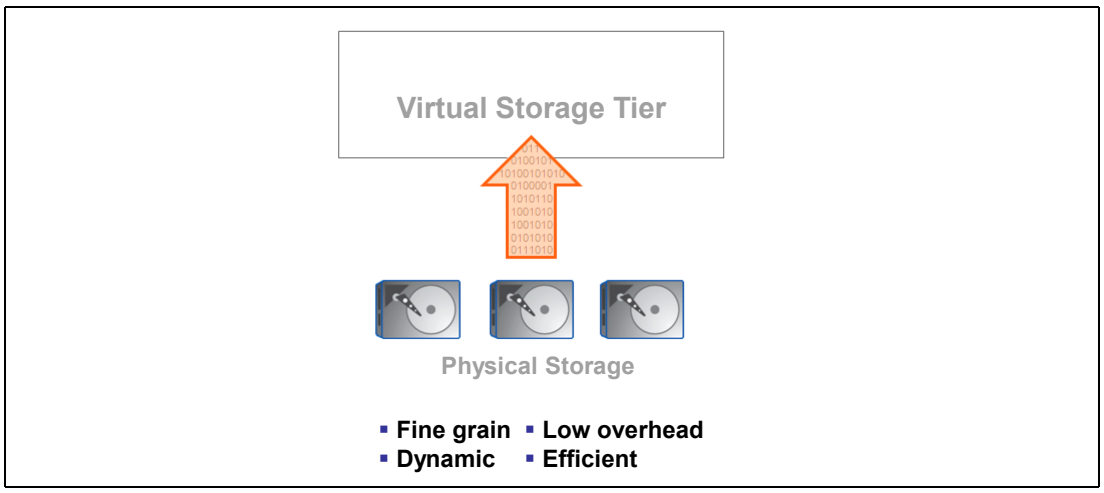

Figure 2-4 N series Virtual Storage Tier

Virtual Storage Tier provides the following benefits:

- The N Series Virtual Storage Tier provides fully automated use and optimization of Flash technology, both controller-based PCI-e-based Flash and solid-state disk (SSD).
- IBM N series Flash Cache PCI-e modules improve performance for workloads that are random read-intensive, reducing latency by a factor of 10 or more compared to hard disk drives.
- Flash Cache modules are available in capacities up to 1 terabyte and provide controller-based caching.
- IBM N series Flash Pool provides caching of both random read and write operations through the automated use of SSD drives, thereby enabling the use of capacity-optimized hard disk drive technology across the majority of application workloads.
- Flash Pool enables the creation of a Data ONTAP software RAID-protected aggregate that is composed of a combination of hard disk drives (HDDs) and solid-state disk drives.
- With Flash Cache and Flash Pool you can significantly decrease the cost of your disk purchases and make your storage environment more efficient. Specific workload testing showed the following results:
  - File Services Workload: Combining Flash Cache with SATA disks can significantly improve I/O throughput and response time (compared to high-performance HDD configurations) while lowering the cost per terabyte of storage and saving on power.
  - OLTP Workload: Combining Flash Pool with SATA disks can match the performance of high performance HDD configurations (Fibre Channel or SAS), while providing more capacity, lowering the cost per terabyte of storage, and saving significantly on power.
  - When placing a pool of VMs on an aggregate that is utilizing the Virtual Storage Tier technology, changes in the required performance on individual VMs will automatically rebalance the workload across the VMs existing in that aggregate.

2.5 Virtual Storage Console

The N series Virtual Storage Console (VSC) delivers comprehensive management of storage operations in both SAN-based and NAS-based VMware infrastructures. These operations include discovery, health monitoring, capacity management, provisioning, cloning, backup,

and recovery. These capabilities are integrated into the VMware vCenter framework, providing VMware administrators with the tools to easily improve both server and storage visibility and efficiencies, while still enabling storage administrators to own and control storage policies. VSC provides VMware administrators with a window into the storage domain and the tools to effectively and efficiently manage the lifecycle of virtual server and desktop environments (see Figure 2-5).

The Virtual Storage Console (VSC) software is a single vCenter Server plug-in. It provides end-to-end virtual machine lifecycle management for VMware environments running N series storage. The plug-in provides these features:

- Storage configuration and monitoring, using the Monitoring and Host Configuration capability (previously called the Virtual Storage Console capability)
- Datastore provisioning and virtual machine cloning, using the Provisioning and Cloning capability
- Backup and recovery of virtual machines and datastores, using the Backup and Recovery capability

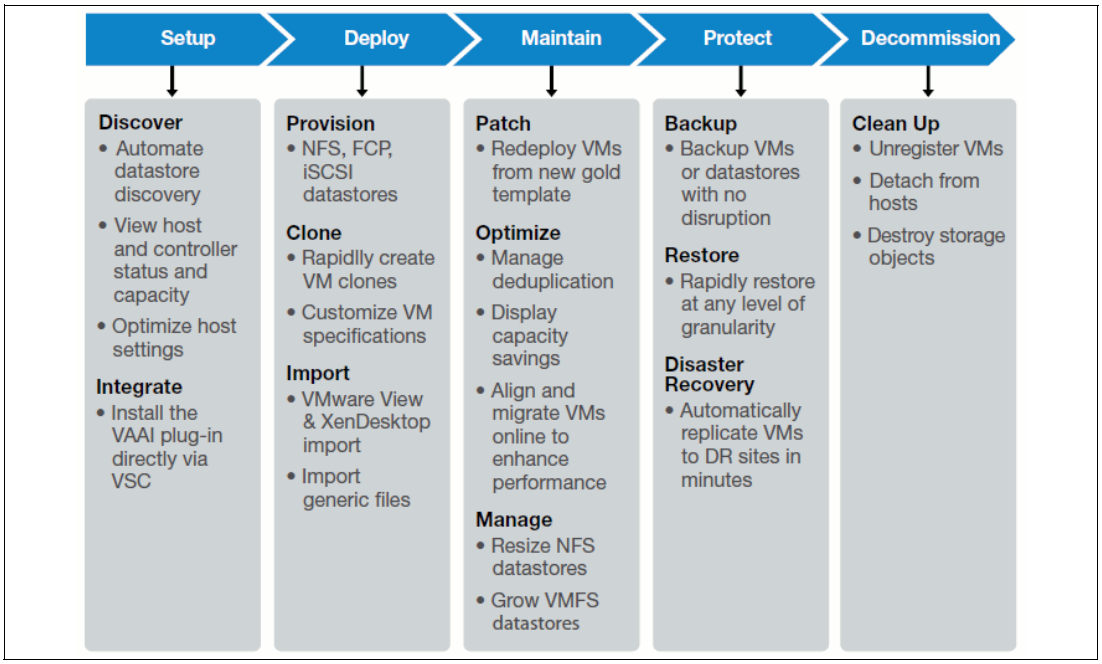

Figure 2-5 Virtual Storage Console end-to-end lifecycle management

As a vCenter Server plug-in, shown in Figure 2-6 on page 18, the VSC is available to all vSphere clients that connect to the vCenter Server. This availability is different from a client-side plug-in that must be installed on every vSphere client. You can install the VSC

software on a Windows server in your data center, but you must *not* install it on a client computer.

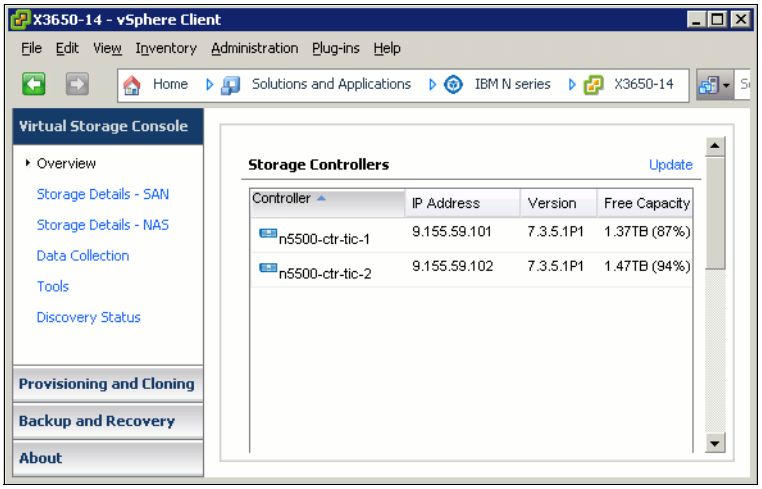

Figure 2-6 Virtual Storage Console 2

Virtual Storage Console integrates VSC storage discovery, health monitoring, capacity management, and preferred practice-based storage setting. It offers additional management capabilities with two capability options in a single vSphere client plug-in. Thus it enables centralized, end-to-end management of virtual server and desktop environments running on N series storage.

VSC is composed of three main components:

- Virtual Storage Console capability (base product)

  This provides a storage view of the VMware environment with a VM administrator perspective. It automatically optimizes the client's host and storage configurations, including HBA time-outs, NFS parameters, and multipath configurations. Using the Virtual Storage Console, a VM administrator can quickly and easily view controller status and capacity information. Also, the administrator can accurately report back utilization information, to make more informed decisions about VM object placement.

- Provisioning and Cloning capability

  This provides end-to-end datastore management (provisioning, resizing, and deletion). It also offers rapid, space-efficient VM server and desktop cloning, patching, and updating by using FlexClone technology.

- Backup and Recovery capability (formerly SnapManager for Virtual Infrastructure)

  This automates data protection processes by enabling VMware administrators to centrally manage backup and recovery of datastores and VMs. This management can be performed without impacting guest performance. The administrator can also rapidly recover from these backup copies at any level of granularity: datastore, VM, VMDK, or guest file.

Additionally, other important capabilities are offered:

- VAAI support for SAN and for NFS through the VIB plug-in

  This integration enables a variety of resource-consuming activities to be offloaded from physical VM hosts to N series storage, thereby improving host resource utilization and performance.

- VSC lets you easily determine when VAAI support is enabled on a specific N series controller. It simplifies operations by enabling you to centrally install the NFS VAAI plug-in through the VSC interface.
- Nondisruptive VM optimization and migration

  VSC optimization and migration provides a critical feature that is exclusive to N series systems. VSC lets you scan datastores stored on N series and non N series storage, individually or many at a time, to determine which VMs are misaligned. You can then improve I/O performance by migrating the misaligned VMs to optimized new or existing datastores nondisruptively while the VMs are online.
- Enabling Cloud Computing

  By supporting N series MultiStore technology, VSC enables you to provision and manage VMs in secure multi-tenant cloud environments. Whether you are a client, service provider, or cloud provider, you can securely administer partitions of shared application, compute, and storage resources (vFiler units) from within the vCenter framework, maintaining desired service levels and security for each tenant.

VSC is designed to simplify storage management operations, improve efficiencies, enhance availability, and reduce storage costs in both SAN-based and NAS-based VMware infrastructures. It provides VMware administrators with a window into the storage domain. It also provides the tools to effectively and efficiently manage the lifecycle of virtual server and desktop environments running on N series storage.

2.5.1 Provisioning and Cloning usage

The Provisioning and Cloning capability of Virtual Storage Console helps you to provision datastores and quickly create multiple clones of virtual machines in the VMware environment. Using FlexClone technology, the Provisioning and Cloning capability allows you to efficiently create, deploy, and manage the lifecycle of virtual machines. These tasks can be performed from an easy-to-use interface integrated into the VMware environment. It is ideal for virtual server, desktop, and cloud environments. You can use this capability for the following purposes:

- Clone individual virtual machines and place in new or existing datastores
- Create, resize, or delete datastores
- Apply guest customization specifications and power up new virtual machines
- Run deduplication operations
- Monitor storage savings
- Redeploy virtual machines from a baseline image
- Replicate NFS datastores across sites
- Import virtual machines into virtual desktop infrastructure connection brokers and management tools

2.5.2 Cloning virtual machines

The Provisioning and Cloning capability can theoretically create thousands of virtual machine clones and hundreds of datastores at one time. In practice, however, multiple executions of fewer requests are preferred. The exact size of these requests depends on the size of the vSphere deployment and the hardware configuration of the vSphere Client managing the ESX hosts. To clone virtual machines simply select in the vSphere Client Inventory, right-click a powered-down virtual machine (Figure 2-7) or template, and select **N series** → **Provisioning and Cloning** → **Create rapid clones**.

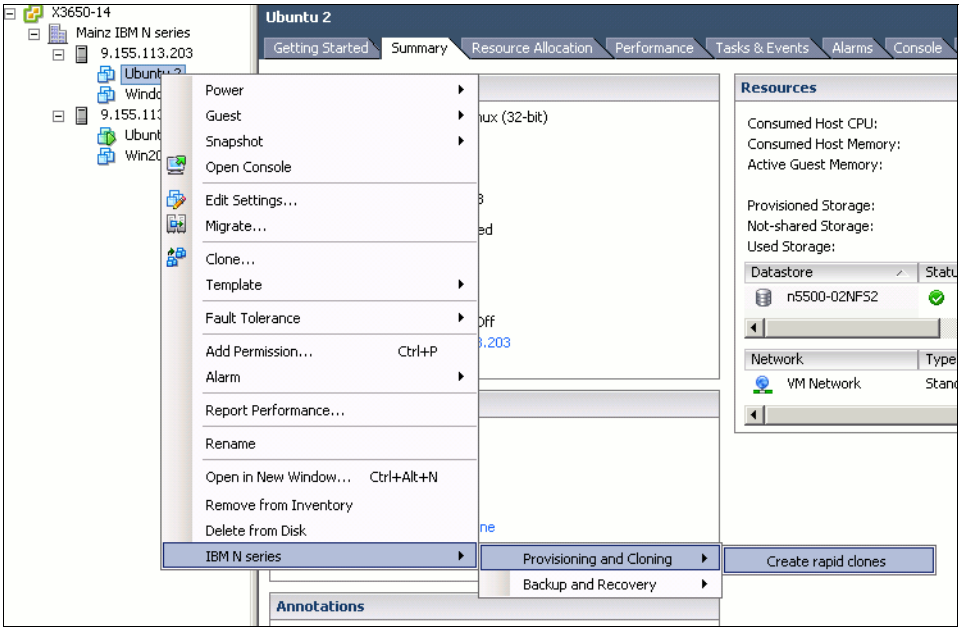

Figure 2-7 Cloning a selected VM

2.5.3 How VSC enables infrastructure automation

Although the three basic Virtual Storage Console (VSC) functional areas are monitoring, provisioning, and backup, it is also useful to focus on the larger topic of infrastructure automation.

When clients need seamless integration across your infrastructure, the storage administrator and the VMware administrator at a basic level must execute in a coordinated way so that systems are configured optimally to support the services being provided. Some form of “cross-domain” or “cross-management-boundary” execution is needed to provide this coordination.

The Virtual Storage Console provides this capability for storage and VMware administrators. It changes the role of the storage administrator from someone who provisions storage for the VMware administrator, which the VMware administrator then consumes, to a model where the

storage administrator can assign pools of storage resources to the VMware administrator (Figure 2-8).

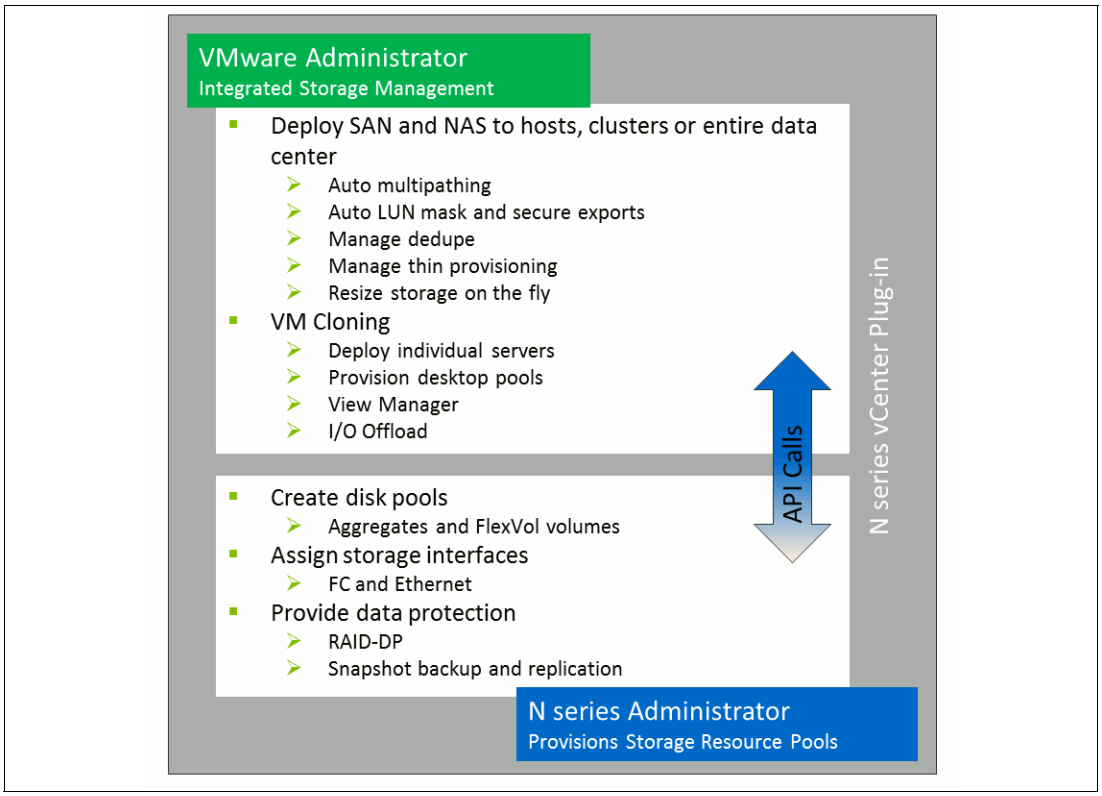

Figure 2-8 vSphere and storage administrator tasks using the VSC vCenter plug-in

These resources can be either physical resources such as sets of disk drives (an aggregate on N series platforms) or physical interfaces such as Ethernet ports. Or, they can be virtual or logical resources such as flexible volumes (FlexVol volumes) and virtual interfaces. A virtual resource can also be a “virtual storage profile.” This feature is known as a vFiler unit enabled by the MultiStore feature. Instead of a physical storage controller, the VMware administrator is given control over a virtual storage system that is a subset of a physical storage system.

Whether physical or virtual, these resources are not provisioned to the VMware administrator; in effect storage administrators grant control of specific resources to the VMware administrator. Further, you can lock that down and apply enterprise-level policies, including backup, compliance, and security. Through the VSC plug-in for VMware vCenter, the VMware administrator can now consume these resources on demand; can never exceed the resource limits assigned by the storage administrator; and must comply with the policies that the storage administrator has put in place.

VSC makes sure that your storage preferred practices are applied to resources consumed by the VMware administrator. This gives you a top-down, bottom-up intersection between storage technology and VMware technology. Thus, the workload of the storage administrator to support the VMware administrator is reduced. Services can be brought online more quickly, and are more dynamic. Service levels and performance are assured, and checks and balances are automatically implemented.

As a result, clients will experience improved IT infrastructure operations, a reduction in the time used for operational tasks, and a reduction in resources consumed. For complete end-to-end automation, N series can also be integrated with VMware vCloud.

2.6 FlexClone

FlexClone technology enables multiple, instant data set clones with no storage impact. It provides dramatic improvements for application test and development environments. It is also tightly integrated with file system technology and a microkernel design in a way that renders competitive methods archaic.

Within VMware vSphere environments, the FlexClone feature can be used for cloning VMs, datastore mounts, VMDK recovery to alternate datastore, and single file restore (SFR). Thus FlexClone technologies allow administrators to easily provision virtual machines.

With FlexClone, you can clone a volume, a file, or LUN and make it available to other servers. This method can be used to deploy multiple ESXi hosts. For example, you can install the ESXi operating system on a single server, and then use FlexClone to make a copy of that LUN to multiple servers. This N series feature is also helpful when you want to reproduce your production environment on a test area. FlexClone functionality is shown in Figure 2-9.

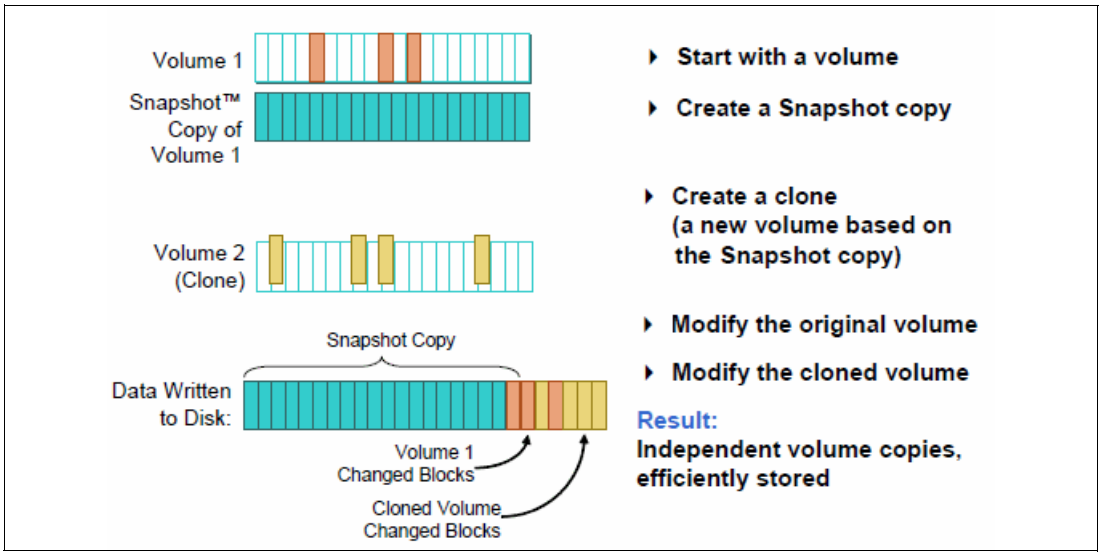

Figure 2-9 FlexClone cloning and space savings

Customizing the ESXi operating system: After using FlexClone, the ESXi operating system must to be customized to avoid IP and name conflicts with the original server from which the FlexClone was taken. VSC supports the vCenter administrator by simplifying these tasks.

2.7 FlexShare

IBM System Storage N series FlexShare is a control-of-service tool designed to give administrators the control needed to prioritize applications based on how critical they are to the business. It also provides a priority mechanism to give preferential treatment to higher priority tasks using the methods discussed in this section.

Priorities are assigned to volumes in order to assign relative priorities between these possibilities:

- Using different volumes

  You can specify, for example, that operations on volume cifs_vol3 are more important than operations on volume cifs_vol2 and other volumes.

- Client data accesses and system operations

  You can specify that client accesses are more important than SnapMirror operations.

- Cache utilization options

  You can configure the cache to retain data in cache or reuse the cache depending on workload characteristics. Optimizing cache usage can significantly increase performance for data that is frequently read or written.

2.8 Snapshots

VMware is capable of taking a snapshot of guests. This enables you to make point-in-time copies that provide the fastest means to recover a guest to a previous point in time. N series storage systems have been providing clients with the ability to create snapshot copies of their data since its introduction.

The basic concept of a snapshot is similar between VMware and N series systems. However, it is important to understand both the major differences between the two technologies, and when to use one rather than the other.

2.8.1 VMware snapshots

VMware snapshots provide simple point-in-time versions of guests, allowing quick recovery. The benefit of VMware snapshots is that they are easy to create and use because they can be executed and scheduled from within vCenter. For more information about native VMware snapshots, including usage guidelines, see the *ESXi and vCenter Server 5 Documentation* section at the following website:

http://pubs.vmware.com/vsphere-50/index.jsp

2.8.2 N series Snapshot technology

The patented N series Snapshot technology can easily be integrated into VMware environments. This technology provides crash-consistent versions of guests for full guest recovery, full guest cloning, or site replication and disaster recovery.

The benefit of N series Snapshot is that it is the storage industry's only Snapshot technology that does not have a negative impact on system performance. VMware states that, for optimum performance and scalability, hardware-based Snapshot technology is preferred over software-based solutions.

The Virtual Storage Console completes the vCenter management by integrating N series Snapshot capabilities with single VMware management.

2.9 Storage configuration

In this section we discuss elemental N series storage technologies.

2.9.1 RAID data protection

When focusing on storage availability, many levels of redundancy are available for deployment. Examples include purchasing physical servers with multiple storage host bus adapters (HBAs), and deploying redundant storage networking and network paths to use storage arrays with redundant controllers. If you have deployed a storage design that meets all of the criteria, you might think that you have eliminated all single points of failure. But actually, data protection requirements in a virtual infrastructure are even greater than in a traditional physical server infrastructure. Data protection has become a paramount feature of shared storage devices.

RAID-DP in Data ONTAP is an advanced RAID technology that is provided as the default RAID level on all IBM System Storage N series storage systems. RAID-DP provides protection from the simultaneous loss of two drives in a single RAID group. RAID-DP is economical to deploy, because the impact with the default RAID group size is a mere 12.5 percent. This level of resiliency and storage efficiency makes data residing on RAID-DP safer than data stored on RAID 5 and more cost effective than RAID 10. Use RAID-DP on all RAID groups that store VMware data.

2.9.2 Aggregates

An *aggregate* is the virtualization layer of Data ONTAP that abstracts physical disks from logical data sets, which are referred to as *flexible volumes*. Aggregates provide a means whereby the total IOPS available to all of the physical disks is pooled as a resource. This design is better suited to meet the needs of an unpredictable and mixed workload.

Whenever possible, use a small aggregate as the *root aggregate*, which stores the files that are required for running and providing GUI management tools for the N series storage system. Place the remaining storage in a small number of large aggregates.

Because the overall disk I/O from the VMware Virtual Infrastructure 3 environment is traditionally random by nature, this storage design ensures optimal performance because a large number of physical spindles are available to service I/O requests. On smaller N series storage systems, it might be impractical to have more than a single aggregate because of a restricted number of disk drives on the system. In these cases, it is acceptable to have only a single aggregate.

2.9.3 Flexible volumes

Flexible volumes (Figure 2-10 on page 25) contain either LUNs or virtual disk files that are accessed by hosts. Use a one-to-one (1:1) alignment of VMware Virtual Infrastructure datastores to flexible volumes. This design makes it easy to understand the VMware ESX Server data layout when viewing the storage configuration from the N series storage system.

This mapping model also provides an easy means to implement Snapshot backups or SnapMirror replication policies at the datastore level. This is because Data ONTAP implements these storage-side features at the flexible volume level.

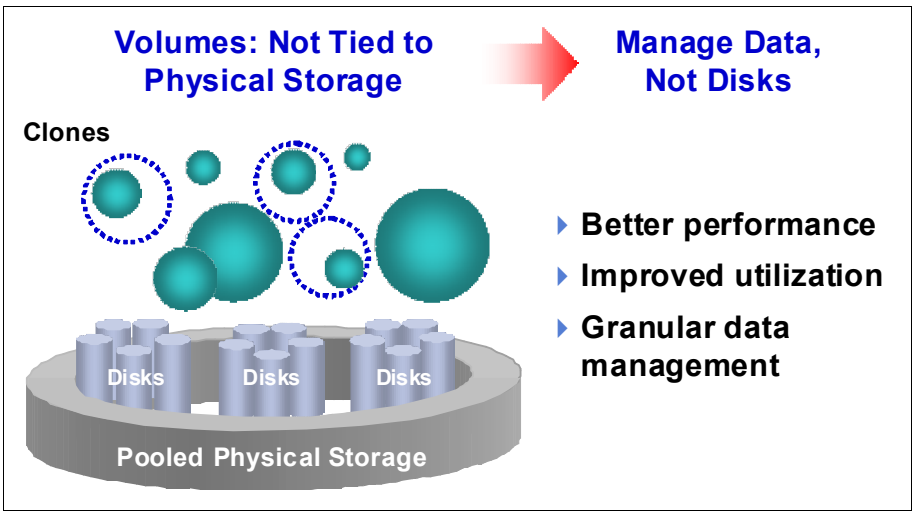

Figure 2-10 Flexible volumes

2.10 Using deduplication or compression with VMware

As previously explained, *deduplication* refers to the concept of storing multiple instances of the same information into a single point. Then a pointer is used to refer to it on the next occurrence, so files that potentially might be stored in an environment many times are stored only once. Microsoft Exchange and Symantec Vault are commercial products known for the usage of deduplication.

VMware environments deduplicate extremely well. N series deduplication provides Advanced Single Instance Storage (A-SIS) at the storage level, rather than the application level. This

significantly reduces the amount of storage that is used when the same files are stored multiple times. The deduplication process is shown in Figure 2-11.

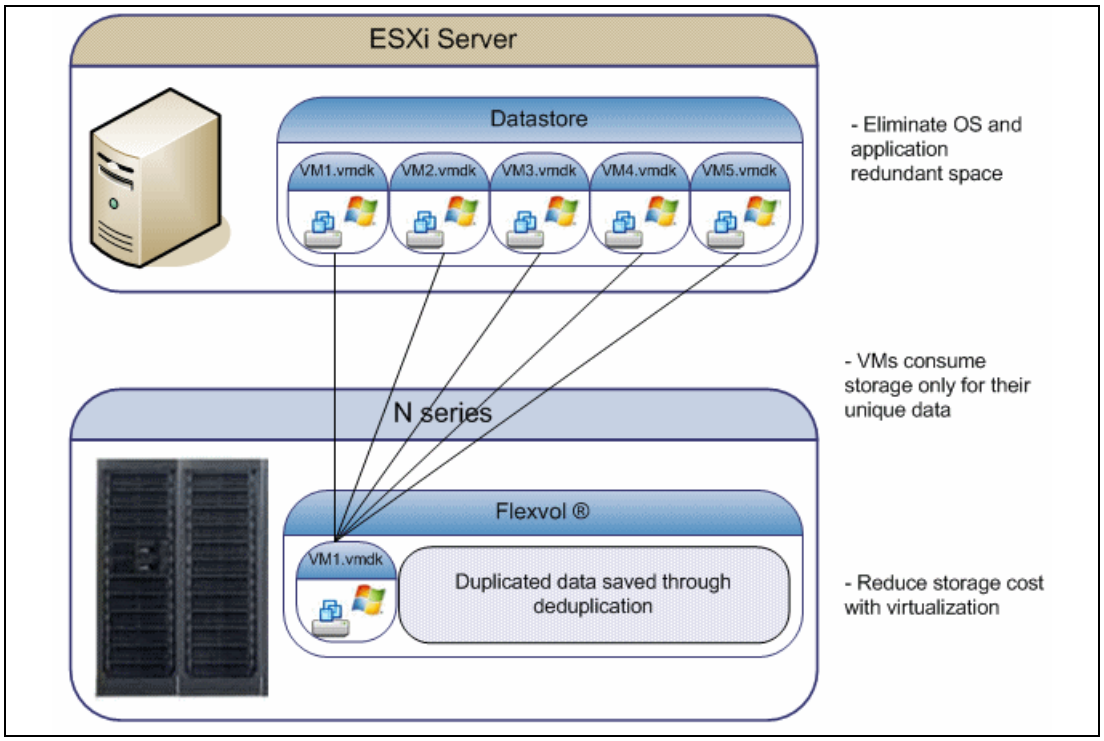

Figure 2-11 N series deduplication benefits

Although VMware environments deduplicate well, when considering the VMDK and datastore layouts, keep the following points in mind:

- Operating system VMDKs deduplicate well because the binary files, patches, and drivers are highly redundant between virtual machines (VMs). Maximum savings can be achieved by keeping these in the same volume. These VMDKs typically do not benefit from compression over what deduplication can already achieve.

  Further, because compressed blocks bypass the Flash Cache card, compressing the operating system VMDK can negatively impact performance during a boot storm. For these reasons, we advise that you do not add compression to an operating system VMDK.
- Application binary VMDKs compress or deduplicate to varying degrees. Duplicate applications deduplicate well. Applications from the same vendor commonly have similar libraries installed and deduplicate somewhat successfully. Applications written by different vendors do not deduplicate at all.
- When compressed or deduplicated, application datasets have varying levels of space savings and performance impact based on application and intended use. Careful consideration is needed, just as with non-virtualized environments, before deciding to keep the application data in a compressed or deduplicated volume.
- Transient and temporary data such as VM swap files, page files, and user and system temp directories do not compress or deduplicate well and potentially add significant performance pressure when compressed or deduplicated. Therefore, it is advisable to keep this data on a separate VMDK and volume that are not compressed or deduplicated.
- The IBM System Storage N series includes a performance enhancement referred to as *intelligent cache*. Although it is applicable to many different environments, intelligent caching is particularly applicable to VM environments, where multiple blocks are set to

zero as a result of system initialization. These zero blocks are all recognized as duplicates and are deduplicated quite efficiently.

The warm cache extension enhancement provides increased sequential read performance for such environments, where there are significantly large amounts of deduplicated blocks. Examples of sequential read applications that benefit from this performance enhancement include NDMP, SnapVault, and some NFS-based applications. This performance enhancement is also beneficial to the boot-up processes in VDI environments.

- The expectation is that about 30 percent space savings will be achieved overall. This is a conservative figure, and in some cases users have achieved savings of up to 80 percent. The major factor that affects this percentage is the amount of application data. New installations typically deduplicate extremely well, because they do not contain a significant amount of application data.

VMware considerations:

- In VMware environments, the need for proper partitioning and alignment of the VMDKs is extremely important (not just for deduplication). VMware must be configured so that the VMDKs are aligned on N series WAFL (Write Anywhere File Layout - File System) 4K block boundaries as part of a standard VMware implementation.
- The applications in which performance is heavily affected by deduplication (when these applications are run without VMware) are likely to suffer the same performance impact from deduplication when they are run with VMware.

2.11 Further information

For additional information about how to integrate VMware vSphere with the IBM System Storage N series system, including preferred practices and detailed descriptions, see “Related publications” on page 57.

For details about IBM VMware offerings, refer to the following website:

http://www.ibm.com/systems/xbc/cog/vmwareesx.html

For further details about how to integrate N series with VMware vSphere, see the IBM Redbooks publication *IBM System Storage N series with VMware vSphere 4.1*, SG24-7636, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/sg247636.html?Open

3

Server

This chapter describes the IBM System x components that are used as building blocks for the VMware vSphere environment.

This chapter includes the following sections:

- Rack and power infrastructure
- Host/compute solution classification
- Entry x3650 M4 host/compute nodes
- Mainstream HS23 host/compute nodes
- Advanced HX5 host/compute nodes
- Management mode vCenter server
- Active Directory Server

3.1 Rack and power infrastructure

Optimized infrastructure equipment is critical to drive improved IT efficiency and availability for the data centers of today and tomorrow. The IBM rack and power infrastructure offerings are custom designed for IBM System x servers, and they provide the following benefits:

- Improved data center efficiency
  - Increased power efficiency
  - Increased space efficiency (avoid over-design)
  - Lower cost through better data center utilization
- Improved IT availability
  - Improved uptime
  - Act before downtime impacts business
  - Match utilization, power resources, and capacity planning

In addition, IT availability and efficiency are primary drivers to data center spending:

- Servers per rack are up 50 percent since the year 2000
- Energy consumption is up 20 percent due to more memory, and improved utilization due to virtualization
- Higher power densities at the server and rack levels

In today's online environment, even minutes of downtime can have a significant impact on an organization's operations, client satisfaction, and financial results, thus making high availability an essential feature. The technology fundamentals for today's data center require a solid foundation of rack and power infrastructure that delivers the ability to securely manage and control power resources, servers, and appliances in the data center and across the network. This is imperative to maintain the highest levels of IT availability, and drive operational efficiencies.

IBM has announced over 40 new products, refreshing the offerings across the entire rack and power options portfolio, including the following items:

- Three new racks that are 1200 mm deep. This new lineup includes a new 47U tall rack and new 42U versions including a "dynamic" rack that is ship-loadable.
- An IBM lineup of optional universal power supply (UPS) units that includes new rack-mounted and tower units supporting voltages and configurations not previously available, with new 1500, 2200, 3000, and 6000 volt-ampere (VA) units.
- A new line of 0U Strip Power Distribution Units (PDUs), designed for tool-less installation in the new racks. These PDUs have 24 outlets for today's server-dense rack installations.
- IBM is also offering new Local and Global Console Managers that support unique cabling options ("conversion options") to enable chaining up to 1,024 managed devices that can be managed from a single console.

These offerings are shown in Figure 3-1.

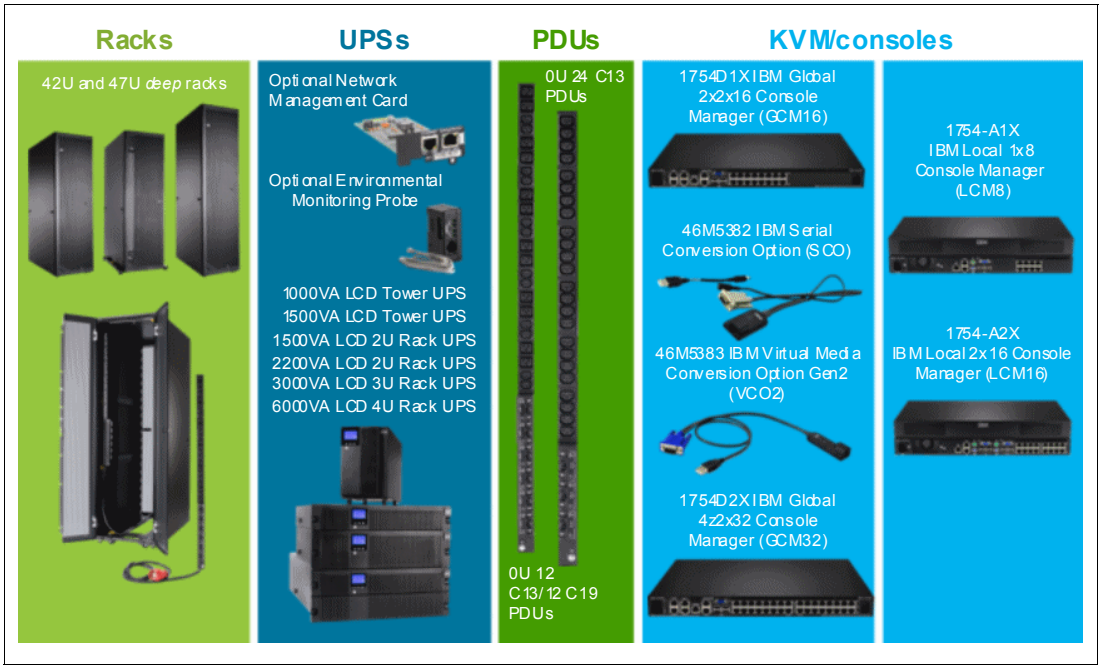

Figure 3-1 System x rack and power choices

Table 3-1 lists the rack and power items used in all of these offerings.

Table 3-1 Rack and power parts list

Part number	Description	Quantity
1754D1X	IBM Global 2x2x16 Console Manager	1
46M5383	IBM Virtual Media Conversion Option VCO2	11
172319X	1U 19-inch Flat Panel Monitor Console Kit with DVD	1
40K5372	IBM Keyboard with Integrated Pointing Device	1
53956KX	IBM 6000VA 4U Rack UPS (230V)	2
46M4110	IBM LCD UPS Network Management Card	2
46M4004	IBM 1U 12 C13 Switched and Monitored DPI PDU	4
40K9614	IBM DPI 30A Cord (NEMA L6-30P)	4
93604PX	IBM 42U 1200 mm Deep Dynamic Rack	1

3.2 Host/compute solution classification

The compute nodes are the processing elements for the offering. Virtualization technology in VMware vSphere allows each user to see the compute node as a dedicated resource, even though it is shared among other users. Table 3-2 on page 32 lists the reference architecture solution classification.

Table 3-2 Reference architecture solution classification

	Entry	Mainstream	Advanced
Networking	IBM G8124E	G8124E or BladeCenter	BladeCenter
IBM server platform	x3550M4 or x3650M4	x3650M4 or HS23	HS23 or HX5
Storage[a]	N3150/N3220/N3240	N6240	N6270
Range of VMs (estimated)[b]	Up to 400 Up to 250 for N3150[c]	Up to 600	Up to 800

a. All N series attachments are through FCP or NFS.
b. Range of VMs are based on mixed workload deployments.
c. N 3150 iSCSI or NFS.

The Entry configuration uses four compute nodes, and the Mainstream configuration has eight compute nodes, as explained here:

- Entry has four x3550M4 or x3650 M4 servers.
- Mainstream has eight x3650 M4 or HS23 servers.
- Advanced has eight HS23 or HX5 servers.

IBM is a leader in technology and innovation, and has a deep understanding of virtual environments. With substantial investments in green initiatives and energy-smart designs, IBM not only provides high performing, easy-to-manage servers, but can also help minimize costs for power and cooling.

3.3 Entry x3650 M4 host/compute nodes

The x3650 M4 is an outstanding 2U two-socket business-critical server, offering improved performance and pay-as-you grow flexibility along with new features that improve server management capability. This powerful system is designed for your most important business applications and cloud deployments.

Combining balanced performance and flexibility, the x3650 M4 is a great choice for a range of small and medium businesses, and even up to large enterprises. It can provide outstanding uptime to keep business-critical applications and cloud deployments running safely. Ease of use and comprehensive systems management tools make it easy to deploy. Outstanding RAS and high-efficiency design improve your business environment and help save operational costs.

The x3650 M4 offers numerous features to boost performance, improve scalability, and reduce costs:

- The Intel Xeon processor E5-2600 product family improves productivity by offering superior system performance with 8-core processors and up to 2.9 GHz core speeds, up to 20 MB of L3 cache, and up to two 8 GT/s QPI interconnect links.
- Up to two processors, 16 cores, and 32 threads maximize the concurrent execution of multithreaded applications.
- Intelligent and adaptive system performance with Intel Turbo Boost Technology 2.0 allows CPU cores to run at maximum speeds during peak workloads by temporarily going beyond processor TDP.

- Intel Hyper-Threading Technology boosts performance for multithreaded applications by enabling simultaneous multithreading within each processor core, up to two threads per core.
- Intel Virtualization Technology integrates hardware-level virtualization hooks that allow operating system vendors to better utilize the hardware for virtualization workloads.
- Intel Advanced Vector Extensions (AVT) significantly improve floating-point performance for compute-intensive technical and scientific applications compared to Intel Xeon 5600 series processors.
- Twenty-four Load Reduced DIMMs (LRDIMMs) of 1333 MHz DDR3 ECC memory provide speed, high availability, and a memory capacity of up to 768 GB (running at 1066 MHz).
- Theoretical maximum memory bandwidth of the Intel Xeon processor E5 family is 51.6 GBps, which is 60 percent more than in the previous generation of Intel Xeon processors.
- The use of solid-state drives (SSDs) instead of, or along with, traditional spinning drives (HDDs) can significantly improve I/O performance. An SSD can support up to 100 times more I/O operations per second (IOPS) than a typical HDD.
- Up to 16 drive bays, together with internal backup and an optical drive at the same time, provide a flexible and scalable all-in-one platform to meet your increasing demands.
- The server has four integrated Gigabit Ethernet ports and two optional 10 Gb Ethernet ports with mezzanine cards that do not consume PICe slots.
- The server offers PCI Express 3.0 I/O expansion capabilities that improve the theoretical maximum bandwidth by 60 percent (8 GT/s per link) compared to the previous generation of PCI Express 2.0.
- With Intel Integrated I/O Technology, the PCI Express 3.0 controller is integrated into the Intel Xeon processor E5 family. This integration helps to dramatically reduce I/O latency and increase overall system performance.

Figure 3-2 shows the front view of the x3650 M4.

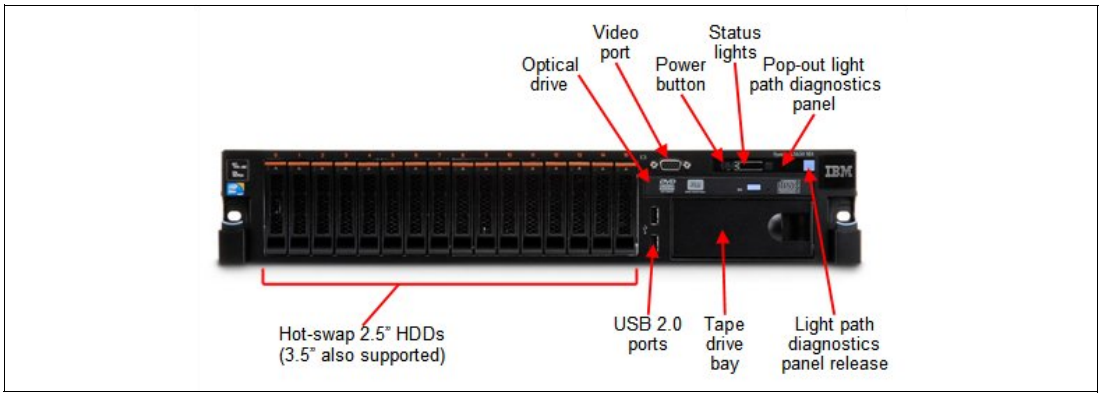

Figure 3-2 x3650 M4 front view

Figure 3-3 shows the rear view of the x3650 M4.

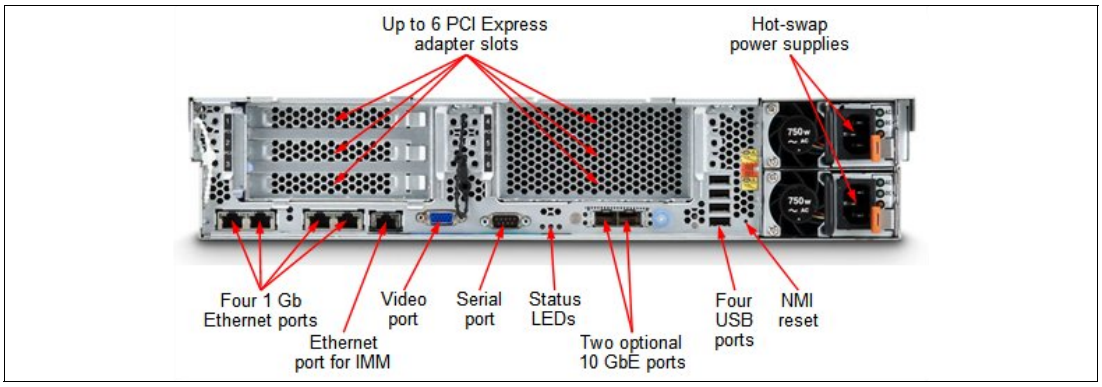

Figure 3-3 x3650 M4 rear view

For additional details about the x3650 M4, refer to the *IBM System x3650 M4 Product Guide*, available at the following URLs:

- http://www.ibm.com/systems/x/hardware/rack/x3650m4/
- http://www.redbooks.ibm.com/technotes/tips0850.pdf

3.4 Mainstream HS23 host/compute nodes

The Advanced configuration has servers with four sockets. In essence, this doubles the number of CPUs on the host node from 16 to 32. Because of this increase in the number of sockets, the Advanced configuration offers an almost doubling of the number of VMs, making for a rich mix of scaleup possibilities in an actively growing data center.

The IBM BladeCenter HS23 is a next-generation two-socket blade server running the Intel Xeon processor E5-2600 product family. With its industry-leading RAS features, energy efficiency, outstanding performance, flexible and scalable I/O, and complete systems management, HS23 offers a robust platform optimized for your mission-critical applications. Standard 30 mm single-wide form-factor protects your investments by providing compatibility with the IBM BladeCenter H, E, S, and HT chassis. The suggested use is a versatile platform to run a broad range of workloads, including infrastructure, virtualization, and enterprise applications.

The IBM BladeCenter HS23 gives you the networking capacity that you need to manage your data center. The new Virtual Fabric-capable integrated 10 GbE offers extreme speed, and the HS23 is designed with highly scalable I/O to give you a total of up to four 10 Gb physical ports that can be divided into up to 14 virtual ports, and the ability to run multiple I/O protocols (FCoE/iSCSI). Sixteen DIMM slots supporting up to 256 GB of DDR3 memory allow you to fit more and larger virtual machines per blade. In addition, the HS23 is backward-compatible with all BladeCenter chassis, including the original BladeCenter E.

The BladeCenter HS23 offers numerous features to boost performance, improve scalability, and reduce costs:

- The Intel Xeon processor E5-2600 product family improves productivity by offering superior system performance with up to 8-core processors and up to 3.0 GHz core speeds depending on the CPU's number of cores, up to 20 MB of L3 cache, and QPI interconnect links of up to 8 GT/s.

- Up to two processors, 16 cores, and 32 threads maximize the concurrent execution of multithreaded applications.
- Intelligent and adaptive system performance with Intel Turbo Boost Technology 2.0 allows CPU cores to run at maximum speeds during peak workloads by temporarily going beyond processor TDP.
- Intel Hyper-Threading Technology boosts performance for multithreaded applications by enabling simultaneous multithreading within each processor core, up to two threads per core.
- Intel Virtualization Technology integrates hardware-level virtualization hooks that allow operating system vendors to better utilize the hardware for virtualization workloads.
- Intel Advanced Vector Extensions (AVT) can significantly improve floating point performance for compute-intensive technical and scientific applications.
- Up to 16 DDR3 ECC memory RDIMMs provide speeds up to 1600 MHz and a memory capacity of up to 256 GB.
- The theoretical maximum memory bandwidth of the Intel Xeon processor E5 family is 51.6 GBps, which is 60 percent more than in the previous generation of Intel Xeon processors.
- The use of solid-state drives (SSDs) instead of, or along with, traditional spinning drives (HDDs) can significantly improve I/O performance. An SSD can support up to 100 times more I/O operations per second (IOPS) than a typical HDD.
- The HS23 scales to 18 I/O ports on a single-wide blade with integrated Gigabit Ethernet and 10 Gb Ethernet ports and optional expansion cards, offering the choice of Ethernet, Fibre Channel, SAS, iSCSI, and FCoE connectivity.
- The HS23 offers PCI Express 3.0 I/O expansion capabilities that improve the theoretical maximum bandwidth by almost 100 percent (8 GTps per link using 128b/130b encoding) compared to the previous generation of PCI Express 2.0 (5 GTps per link using 8b/10b encoding).
- With Intel Integrated I/O Technology, the PCI Express 3.0 controller is integrated into the Intel Xeon processor E5 family. This helps to dramatically reduce I/O latency and increase overall system performance.

Figure 3-4 shows a view of the HS23 BladeCenter module.

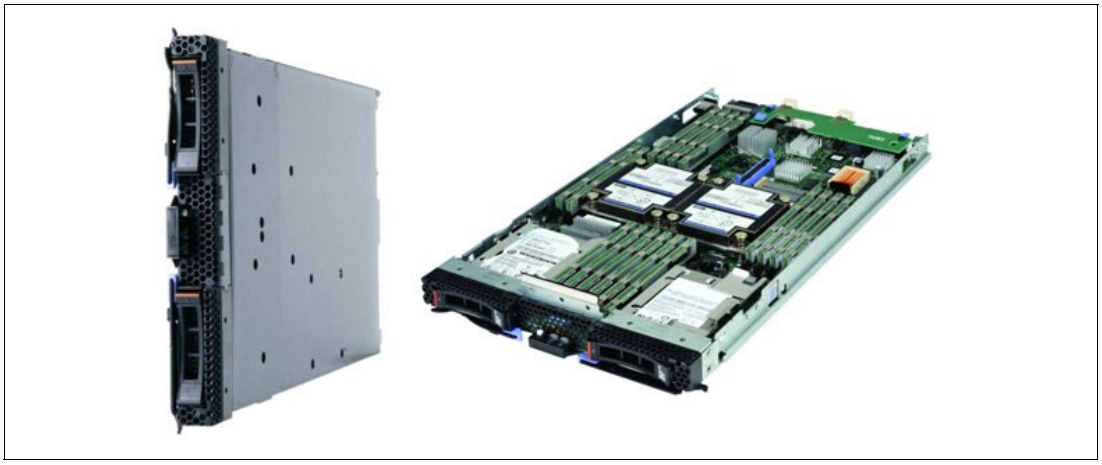

Figure 3-4 View of HS23 BladeCenter module

For additional details about the HS23, refer to the *IBM System HS23 Product Guide*, available at the following URLs:

- http://www.ibm.com/systems/bladecenter/hardware/servers/hs23/
- http://www.redbooks.ibm.com/technotes/tips0843.pdf

3.5 Advanced HX5 host/compute nodes

IBM has been designing and implementing chipsets under the IBM X-Architecture® name since 2001. The eX5 technology represents the fifth generation of products based on the same design principle that IBM began in 1997. That principle is to offer Intel Xeon processor-based systems that are expandable, offer “big iron” reliability, availability, and serviceability (RAS) features, with extremely competitive price/performance.

The eX5 technology is primarily designed around three major workloads: database servers, server consolidation using virtualization services, and Enterprise Resource Planning (application and database) servers.

The BladeCenter HX5 offers numerous features to boost performance, improve scalability, and reduce costs:

- The HX5 offers a choice of high-performance 4-, 6-, 8-, and 10-core Xeon processors with dual integrated memory controllers, clock rates of 1.86 GHz to 2.67 GHz, 12 MB to 30 MB of integrated Level 3 cache, and Intel Virtualization Technology (VT), Hyper-Threading (HT) technology, and Turbo Boost technology.
- By scaling to two blade servers, the HX5 can support up to 512 GB of memory, which is generally enough for even the most demanding virtualization, database, or ERP needs. If even that is not enough memory, however, then by using one blade server and one memory expansion blade, the HX5 can scale to 640 GB.
- Alternatively, with 40 DIMM slots the server and MAX5 can be populated with inexpensive 2 GB DIMMs (for 80 GB), while gaining the performance boost from using 8 memory controllers (4 in the server and 4 more in the MAX5 unit).
- Embedded virtualization (optional on all models) offers extremely high performance, enhanced security, and a zero-gigabyte HDD footprint. (In other words, there is no mechanical HDD to fail.)
- Solid-state drives (SDDs) use only 2 watts of energy per drive, versus 9 to 10 watts for 2.5-inch HDDs. This is as much as 80 percent less energy than a HDD uses (with a corresponding reduction in heat output).
- The extremely high degree of integration in the various BladeCenter chassis reduces the need for server components, thus replacing numerous fans, KVM and Ethernet cables, power supplies, external switches and other components with fewer shared hot-swap/redundant components in the BladeCenter chassis itself. This integration also can greatly reduce the amount of power consumed and heat produced, relative to an equivalent number of 1U servers. This can significantly reduce a data center power bill. The reduced data center footprint can also save on infrastructure cost.
- The midplanes used in all chassis provide high-speed blade-to-blade, blade-to-switch-module, and module-to-module communications internally and externally. The midplanes used in the BladeCenter H and BladeCenter HT chassis provide four 10 Gb data channels to each blade, supporting high-speed switch modules including 4X InfiniBand and 10 Gb Ethernet.
- The various BladeCenter chassis use ultra-high efficiency power supplies. Most industry-standard servers use power supplies that are between 70-75 percent efficient at

converting power from AC wall current to the DC power used inside servers. BladeCenter power modules are up to 92 percent efficient. This helps save even more money, because more of the power input you are paying for is used for processing, rather than released into the data center as waste heat that requires even more energy to cool.

- BladeCenter design also reduces the number of parts required to run the system. Sharing fans, systems management, and optical media means fewer parts to buy and maintain, and fewer items that can fail and bring the overall solution down.

The HX5 can grow with your application requirements, due to the following abilities:

- The ability to grow from a single-wide 2-socket server blade to a double-wide 4-socket server configuration, optimized for compute-intensive workloads (with up to 4 processors/40 cores, 512 GB of memory, 4 PCIe cards, 16 I/O ports, and 4 SSDs in two server blades). Or, for memory-intensive workloads, with one server blade and one MAX5 memory expansion blade (2 processors/20 cores, 640 GB of memory, 4 PCIe cards, 16 I/O ports, and 4 SSDs).
- A choice of processor speeds (1.73 to 2.0 GHz), and shared L3 cache sizes (12 MB, 18 MB, 24 MB, or 30 MB).
- Up to two internal hot-swap 1.8-inch solid-state drives, and access to terabytes of external storage through the BladeCenter S chassis or IBM System Storage SAN and NAS storage devices. SSDs consume only 10-20 percent of the energy required by 2.5-inch HDDs.
- Two Gigabit Ethernet ports standard, plus optional 2-port or 4-port expansion cards or a BladeCenter PCI Express I/O Expansion Unit 3. One HX5 model includes a 10 Gb expansion card (optional in the other models).

Figure 3-5 shows a view of the HS23 BladeCenter module.

Figure 3-5 View of HX5 BladeCenter module

For additional details about the HX5, refer to the *IBM System HSX5 Product Guide*, available at the following URLs:

- http://www.ibm.com/systems/bladecenter/hardware/servers/hx5/
- http://public.dhe.ibm.com/common/ssi/ecm/en/blo03039usen/BLO03039USEN.PDF

3.6 Management node vCenter server

The management node for all three configurations needs to host the VMware vCenter server. Compared to other virtualization solutions, you can run the vCenter server in a virtual machine. As a result, you do not need dedicated hardware to manage your cloud environments.

To implement a vCenter server, simply install the vCenter server in the Microsoft Windows virtual machine that runs on an ESXi host. By doing so, you realize additional benefits compared to using dedicated hardware:

- You have support for VMware High Availability (HA).
- You have support of VMware virtual machine migration functions in case of maintenance or other necessary downtime of the physical equipment.
- You are able to utilize VM-level utilities (for example, VMware Data Recovery) and VMware snapshot mechanisms for possible restore points. Thus, you create an additional level of security for your vCenter management server.

For more information and preferred practices about how to install the vCenter server, refer to:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2003790

3.7 Active Directory server

This section addresses the possible integration of a Microsoft Windows Active Directory server in the solution. This optional server provides an Active Directory and DNS environment for your servers. These services are required for the servers and computers to participate in a Microsoft Active Directory environment. If an Active Directory server is already available in your environment, this specific kind of server is not needed.

The Active Directory server can be virtualized with minimal impact on performance. If additional capacity is required (depending on individual workloads), a physical Active Directory server can be added, as is optional with the Mainstream and Advanced configurations. For more detailed information about Active Directory virtualization, refer to the VMware white paper *Virtualizing a Windows Active Directory Domain Infrastructure*, which is available at the following address:

http://www.vmware.com/files/pdf/Virtualizing_Windows_Active_Directory.pdf

3.8 Further information

The following list includes links for additional information about IBM System x server:

- IBM System x Server Rack servers

  http://www.ibm.com/systems/x/hardware/rack/index.html
- IBM System x Configuration and Options Guide

  http://www.ibm.com/systems/xbc/cog/
- IBM x86 Server Reference

  http://www.redbooks.ibm.com/redpapers/pdfs/redpxref.pdf

4

Networking

This chapter explains how using a solution built around iSCSI SANs with high bandwidth 10 Gbps Ethernet makes advanced networking and storage architectures accessible to the value-conscious installation.

This chapter includes the following sections:

- Ethernet switches
- Architecture with multiswitch link aggregation
- Storage load balancing

4.1 Ethernet switches

The IBM System Networking Virtual Fabric 10 Gb G8124E RackSwitch™ for IBM System x provides exceptional performance that is both lossless and low latency. In addition, the G8124E delivers excellent cost savings regarding acquisition costs and energy costs, along with feature-rich design regarding virtualization, CEE/FCoE, high availability, and its enterprise class Layer 2 and Layer 3 functionality.

Figure 4-1 shows the IBM RackSwitch G8124E.

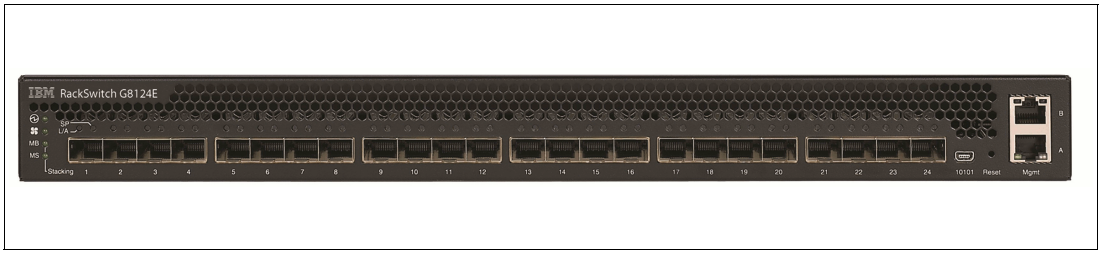

Figure 4-1 IBM RackSwitch G8124E

With support for 1 G or 10 G, this switch is designed for clients that are leveraging 10 G Ethernet today or have plans to do so in the future. This is the first Top of Rack 10 Gb switch for IBM System x designed to support IBM Virtual Fabric, which helps clients significantly reduce cost and complexity when it comes to the I/O requirements of most virtualization deployments today.

Virtual Fabric can help clients reduce the number of multiple I/O adapters down to a single dual-port 10 G adapter, in addition to reducing the number of cables and upstream switch ports required. Virtual Fabric allows clients to carve up a dual-port 10 G adapter into eight virtual NICs (vNICs) and create dedicated virtual pipes between the adapter and the switch for optimal performance, higher availability, and better security. This functionality provides the ability to dynamically allocate bandwidth per vNIC in increments of 100 Mb, while being able to adjust over time without downtime.

The IBM System Networking RackSwitch G8124E offers the following benefits:

- High performance

  The 10 G Low Latency (as low as 570 nanoseconds) switch provides the best combination of extremely low latency, non-blocking line-rate switching and ease of management.

- Lower power and better cooling

  The G8124E uses as little power as two 60-watt light bulbs, which is a fraction of the power consumption of most competitive offerings. Unlike side-cooled switches, which can cause heat recirculation and reliability concerns, the G8124E rear-to-front cooling design reduces data center air conditioning costs by having airflow match the servers in the rack. In addition, variable speed fans assist in automatically reducing power consumption.

- Virtual Fabric

  Virtual Fabric can help clients address I/O requirements for multiple NICs, while also helping reduce cost and complexity. Virtual Fabric for IBM allows for the carving up of a physical NIC into multiple virtual NICs (2 - 8 vNICs), and it creates a virtual pipe between the adapter and the switch for improved performance, availability, and security while reducing cost and complexity.

- VM-aware networking

  IBM VMready® software on the switch helps reduce configuration complexity while significantly improving security levels in virtualized environments. VMready automatically detects virtual machine movement from one physical server to another, and instantly reconfigures the network policy of each VM across VLANs to keep the network up and running without interrupting traffic or impacting performance. VMready works with all leading VM providers such as VMware, Citrix, Xen, and Microsoft.

- Layer 3 functionality

  The switch includes Layer 3 functionality, which provides security and performance benefits because inter-VLAN traffic stays within the chassis. This switch also provides the full range of Layer 3 protocols from static routes for technologies such as Open Shortest Path First (OSPF) and Border Gateway Protocol (BGP) for enterprise clients.

- Seamless interoperability

  IBM switches interoperate seamlessly with other vendors' upstream switches. For more information, see Tolly Reports: Tolly Functionality and Certification: RackSwitch G8000 and G8124 and Cisco Catalyst Interoperability Evaluation, located at: http://www.bladenetwork.net/userfiles/file/PDFs/Tolly209116BladeRackSwitchInteroperability.pdf

- Fault tolerance

  These switches learn alternate routes automatically and perform faster convergence in the unlikely case of a link, switch, or power failure. The switch uses proven technologies such as L2 trunk failover, advanced VLAN-based failover, VRRP, HotLink, Uplink Failure Detection (UFD), IGMP V3 snooping, and OSPF.

- Converged fabric

  The switch is designed to support CEE/DCB and connectivity to FCoE gateways. CEE helps enable clients to combine storage, messaging traffic, VoIP, video, and other data on a common data center Ethernet infrastructure. FCoE helps enable highly efficient block storage over Ethernet for consolidating server network connectivity. As a result, clients can deploy a single server interface for multiple data types. This can simplify both deployment and management of server network connectivity, while maintaining the high availability and robustness required for storage transactions.

For more information and part numbers, refer to the following website:

http://www.redbooks.ibm.com/Redbooks.nsf/RedbookAbstracts/tips0787.html?Open

4.2 Architecture with multiswitch link aggregation

In this configuration, the storage network architecture uses multiswitch link aggregation (MSLA). The IP switches used for the Ethernet storage network support multiswitch link aggregation. Therefore, each storage controller requires one physical connection to each switch. The two ports connected to each storage controller are then combined into one multimode LACP VIF with IP load balancing enabled.

This design provides multiple active connections to each storage controller and provides a means to scale throughput by simply adding more connections. It requires multiple IP addresses per controller. Additionally, each connection uses two physical links for each active network connection to achieve path high availability (HA).

MSLA provides the following benefits:

- It provides multiple active connections to each storage controller.
- It easily scales to more connections by adding NICs and aliases.
- It provides two active connections to each storage controller.
- Storage controller connection load balancing is automatically managed by the EtherChannel IP load-balancing policy.
- It requires only one VMkernel port for IP storage to make use of multiple physical paths.

Figure 4-2 provides a sample design for multiswitch link aggregation.

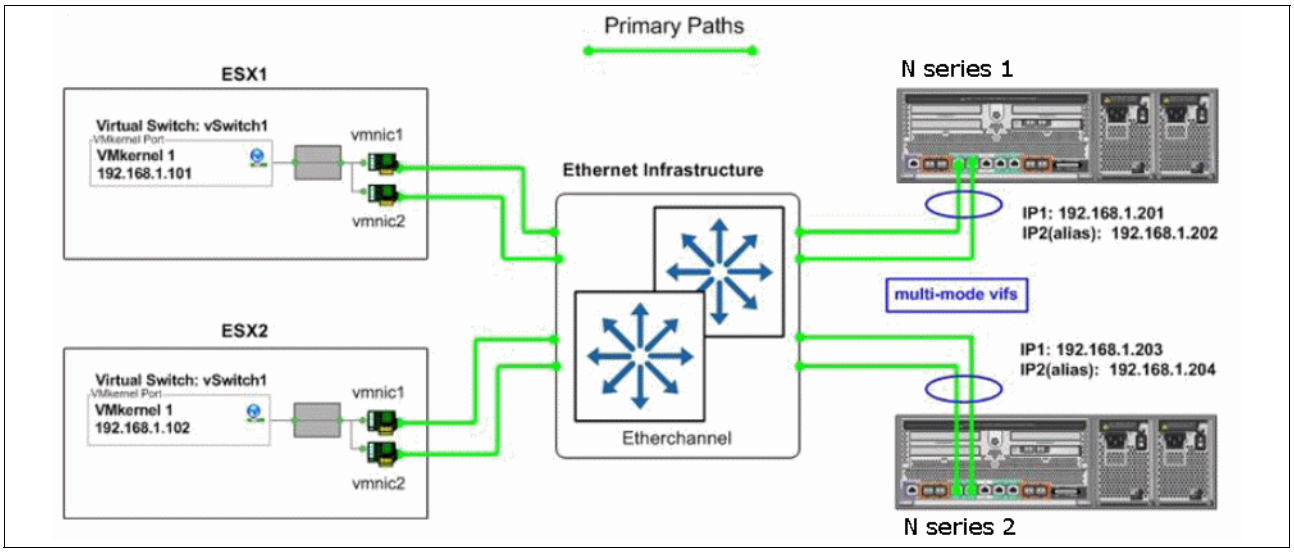

Figure 4-2 Multiswitch link aggregation design

4.3 Storage load balancing

The previous design can be improved with storage load balancing. Using multiple physical paths simultaneously on an IP storage network requires EtherChannel ports and multiple IP addresses on the storage controller, and multiple VMkernel ports defined for storage I/O in the ESX/ESXi hosts. This model results in a design that balances datastore connectivity across all interfaces. This balancing is handled by the RCU at the time the datastore is provisioned.

The layered multimode design requires each storage controller to have at least four physical network connections, as depicted in Figure 4-3 on page 43.

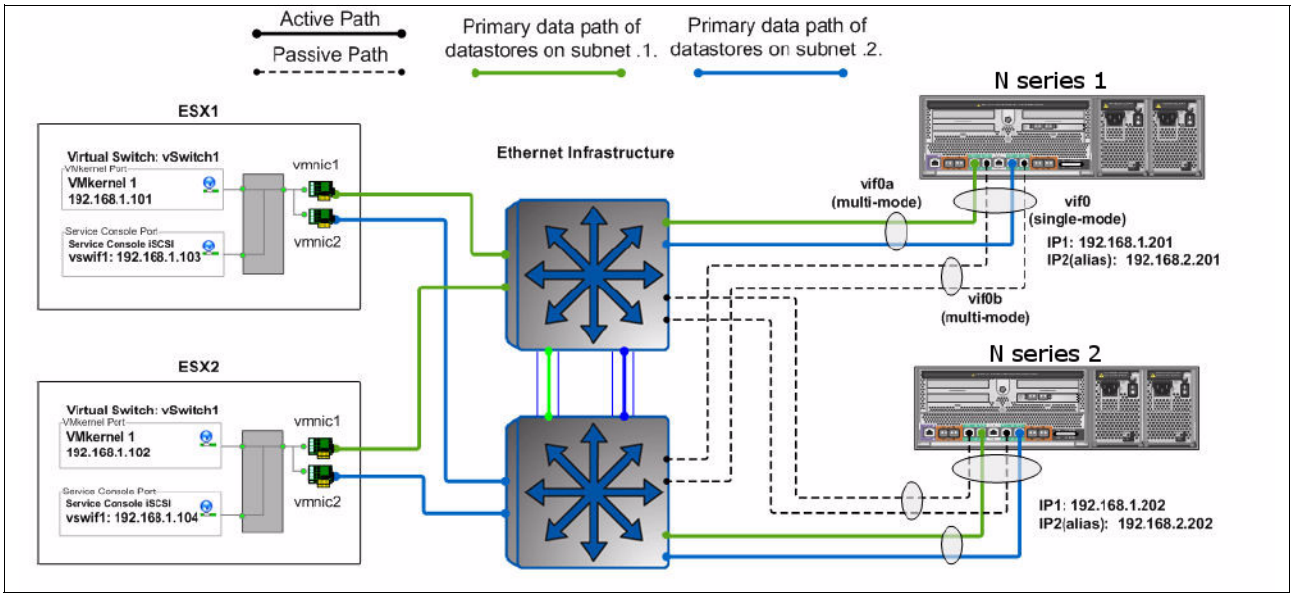

Figure 4-3 Layered multimode network design

The connections are divided into two multimode (active-active) EtherChannels, or VIFs, with IP load balancing enabled. One virtual interface (VIF) is connected to each of the two switches. These two VIFs are then combined into one single mode (active-passive) VIF. This configuration is also referred to as a second-level VIF.

This option also requires multiple IP addresses on the storage appliance. You can assign multiple IP addresses to the single-mode VIF by using IP address aliases or by using virtual local area network (VLAN) tagging.

Layered multimode EtherChannel provides the following benefits:

- The EtherChannel IP load balancing policy automatically manages storage controller connection load balancing.
- Data I/O to a single IP is aggregated over multiple links.

4.4 Further information

This section includes links for additional information for networking components as part of the solution design.

IBM Systems Networking Switches:

http://www.ibm.com/systems/networking/switches/rack.html

IBM System x Top of Rack switches Machine Type Model information

http://www.ibm.com/systems/xbc/cog/network_switches/network_switches.html

5

Storage

This chapter describes the IBM System Storage N series 3000 system that will be used as the unified storage foundation for the solution architecture that was introduced previously.

This chapter includes the following sections:

- Entry portfolio
- N3150
- N3220
- N3240
- N32x0 common information
- Mainstream and Advanced portfolio
- Common functions and features of all mid-range models
- N6240 and N6270 hardware overview

5.1 Entry portfolio

Figure 5-1 shows the N3000 modular disk storage system Entry portfolio. These systems are designed to provide primary and secondary storage for entry to midsize enterprises. N3000 systems offer integrated data access, intelligent management software, data protection capabilities, and expendability to 432 TB of raw capacity, all in a cost-effective package. N3000 series innovations also include internal controller support for Serial-Attached SCSI (SAS) or SATA drives, expandable I/O connectivity, and onboard remote management.

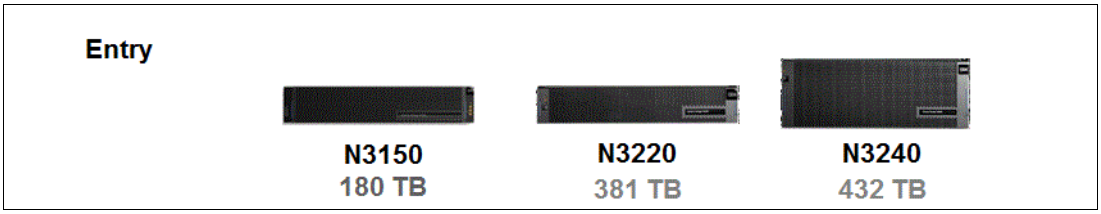

Figure 5-1 Entry systems

The following N3000s are available:

- IBM System Storage N3150 is available as a single-node (Model A15) and as a dual-node (Model A25) (active-active) base unit.
- IBM System Storage N3220 is available as a single-node (Model A12) and as a dual-node (Model A22) (active-active) base unit.
- The IBM System Storage N3240 consists of single-node (Model A14) and dual-node (Model A24) (active-active) base units.

To summarize the differences, Table 5-1 provides a comparison of the N3000 series.

Table 5-1 N3000 series comparison

N3000 series overview[a]	N3150	N3220	N3240
Form factor	2U/12 Drive	2U/24 Drive	4U/24 Drive
Dual controllers	Yes	Yes	Yes
Maximum raw capacity [TB]	180	381	432
Maximum disk drives	60	144	144
Maximum Ethernet ports	8	8	8
Onboard SAS port	4	4	4
Flash Pool support	Yes	Yes	Yes
8 Gb FC support	No	Yes[b]	Yes[b]
10 Gbe support	No	Yes[b]	Yes[b]
Remote management	Yes	Yes	Yes
Storage protocols	iSCSI, NFS, CIFS	FCP, iSCSI, NFS, CIFS	FCP, iSCSI, NFS, CIFS

a. All specifications are for dual-controller, active-active configurations.
b. Based on optional dual-port 10 GbE or 8 Gb FC mezzanine card and single slot per controller.

5.1.1 N3150 models

In this section we discuss the N series 3150 models.

N3150 notes: Be aware of the following points regarding N3150 models.

- N3150 models do not support the FibreChannel protocol.
- Compared to N32xx systems, the N3150 models have newer firmware, and no mezzanine card option is available.

N3150 Model 2857-A15

The N3150 Model A15 is a single-node storage controller that is designed to provide HTTP, Internet Small Computer System Interface (iSCSI), NFS, and CIFS support through optional features. Model A15 is a 2U storage controller that must be mounted in a standard 19-inch rack. Model A15 can be upgraded to a Model A25. However, this is a disruptive upgrade.

N3150 Model 2857-A25

The N3150 Model A25 is designed to provide identical functions as the single-node Model A15, but with the addition of a second Processor Control Module (PCM) and the Clustered Failover (CFO) licensed function. Model A25 consists of two PCMs that are designed to provide failover and failback function, thus helping to improve overall availability. Model A25 is a 2U rack-mountable storage controller.

N3150 hardware

The N3150 hardware includes the following highlights.

- It has a 2U footprint with 12 horizontal disks.
- It leverages the SAS shelf architecture.
- It has the same capacity HDD disks as EXN3000.
- It has the same SSD disks as EXN3000 shelves.
- The N3150 Processor Control Module (PCM) has newer firmware but there is no mezzanine card option available.
- It has 4x GbE ports and 2x 6 Gb SAS ports per PCM.
- There are 6- disk and 12-disk orderable configurations.
- Supported shelves and modules are EXN3000 and EXN3500 using 3 GB and 6 GB SAS IO modules.

Figure 5-2 shows the front and rear views of the N3150.

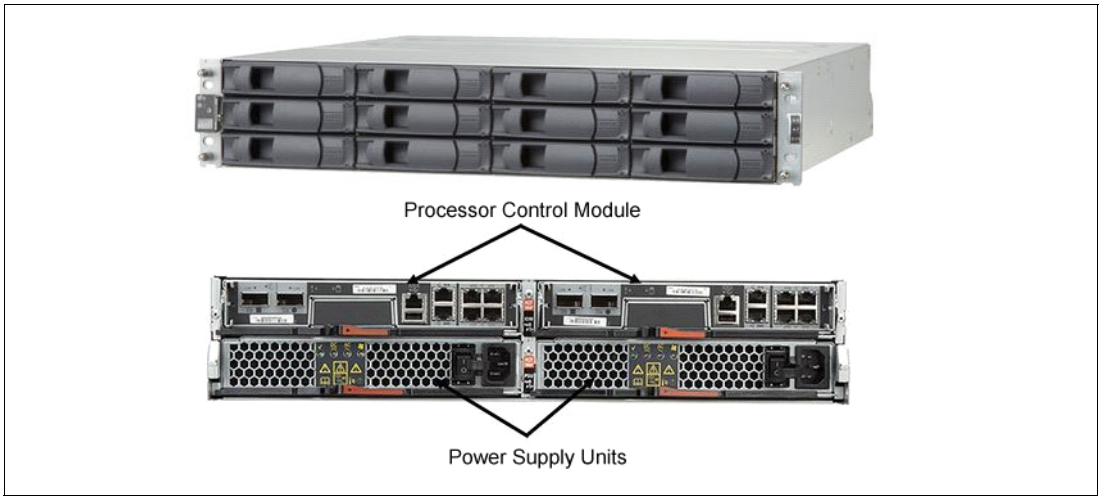

Figure 5-2 N3150 front and rear views

5.1.2 N3220

In this section we discuss N series 3220 Models.

N3220 Model 2857-A12

The N3220 Model A12 is a single-node storage controller designed to provide HTTP, Internet Small Computer System Interface (iSCSI), NFS, CIFS, and Fibre Channel Protocol (FCP) support through optional features. Model A12 is a 2U storage controller that must be mounted in a standard 19-inch rack. Model A12 can be upgraded to a Model A22. However, this is a disruptive upgrade.

N3220 Model 2857-A22

N3320 Model A22 is designed to provide identical functions as the single-node Model A12, but with the addition of a second processor control module (PCM) and the Clustered Failover (CFO) licensed function. Model A22 consists of two PCMs that are designed to provide failover and failback function, thus helping to improve overall availability. Model A22 is a 2U rack-mountable storage controller.

N3220 hardware

The N3220 hardware includes the following highlights.

- It is based on the EXN3500 expansion shelf.
- It has 24 2.5” SFF SAS disk drives:
  - Minimum initial order of 12 disk drives
- It has the following specifications (single node, 2x for dual node):
  - 2U, standard 19-inch rackmount enclosure (single or dual node)
  - One 1.73 GHz Intel dual-core processor
  - 6 GB random access ECC memory (NVRAM 768 MB)
  - Four integrated Gigabit Ethernet RJ45 ports
  - Two SAS ports
  - One serial console port & one integrated RLM port
  - One optional expansion I/O adapter slot on mezzanine card
  - 10 GbE or 8 Gb FC card provides two ports
  - Redundant hot-swappable, auto-ranging power supplies and cooling fans

Figure 5-3 shows the front and rear views of the N3220.

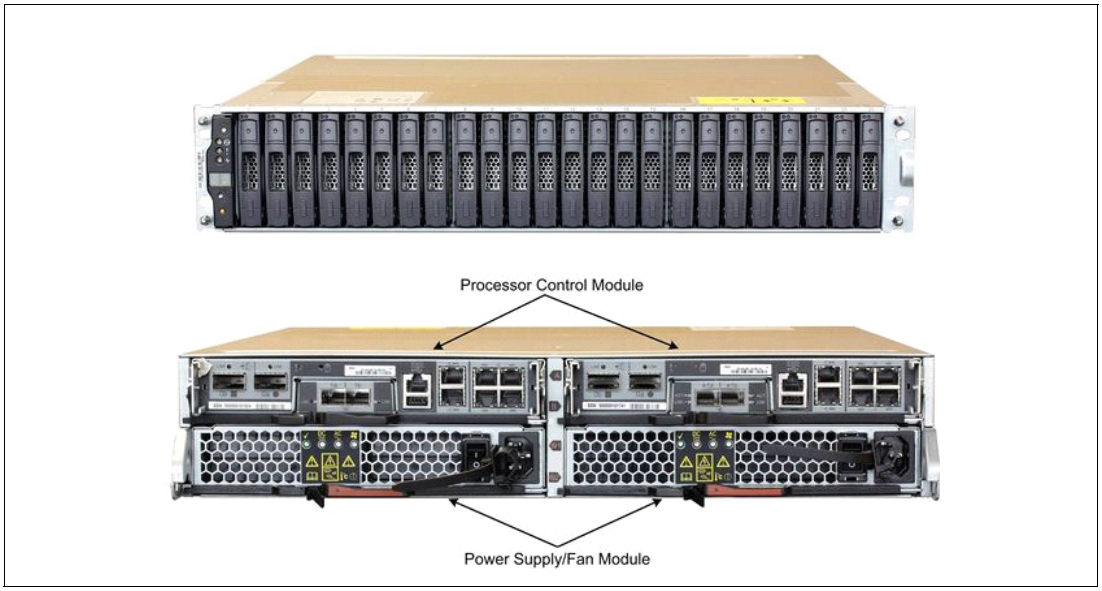

Figure 5-3 N3220 front and rear views

5.1.3 N3240

In this section we discuss the N series 3240 models.

N3240 Model 2857-A14

The N3240 Model A14 is designed to provide a single-node storage controller with HTTP, iSCSI, NFS, CIFS, and Fibre Channel Protocol (FCP) support through optional features. The N3240 Model A14 is a 4U storage controller that must be mounted in a standard 19-inch rack. Model A14 can be upgraded to a Model A24. However, this is a disruptive upgrade.

N3240 Model 2857-A24

The N3240 Model A24 is designed to provide identical functions as the single-node Model A14, but with the addition of a second processor control module (PCM) and the Clustered Failover (CFO) licensed function. Model A24 consists of two PCMs that are designed to provide failover and failback function, thus helping to improve overall availability. Model A24 is a 4U rack-mountable storage controller.

N3240 hardware

The N3240 hardware includes the following highlights.

- It is based on the EXN3000 expansion shelf.
- It has 24 SATA disk drives:
  - Minimum initial order of 12 disk drives.
- It has the following specifications (single node, 2x for dual node):
  - 4U, standard 19-inch rackmount enclosure (single or dual node).
  - One 1.73 GHz Intel dual-core processor.
  - 6 GB random access ECC memory (NVRAM 768 MB).
  - Four integrated Gigabit Ethernet RJ45 ports .
  - Two SAS ports.
  - One serial console port and one integrated RLM port.
  - One optional expansion I/O adapter slot on mezzanine card.

– 10 GbE or 8 Gb FC card provides two ports.
– Redundant hot-swappable, auto-ranging power supplies and cooling fans.

Figure 5-4 shows the front and rear views of the N3240.

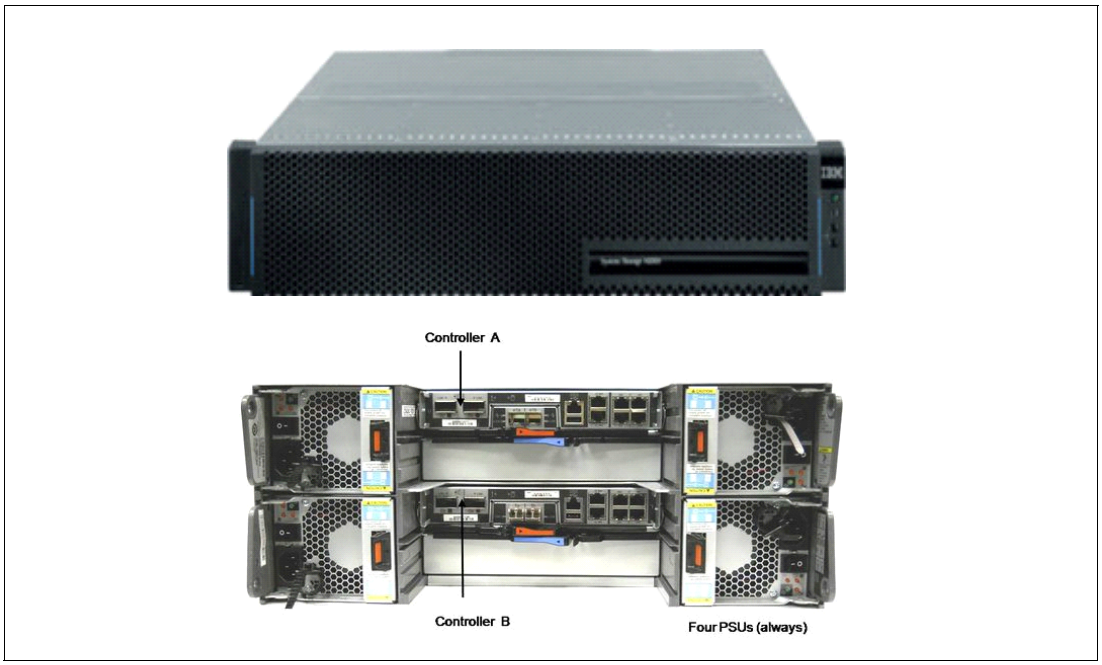

Figure 5-4 N3240 front and rear views

Figure 5-5 shows the Controller with the 8 Gb FC mezzanine card option.

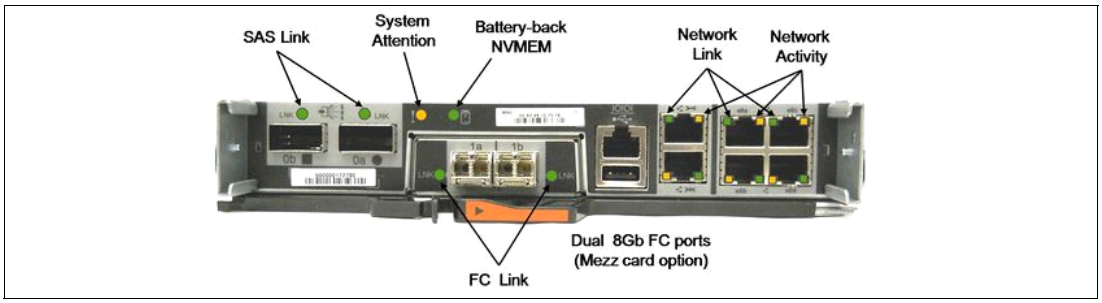

Figure 5-5 Controller with 8 Gb FC mezzanine card option

Figure 5-6 shows the Controller with the 10 GbE mezzanine card option.

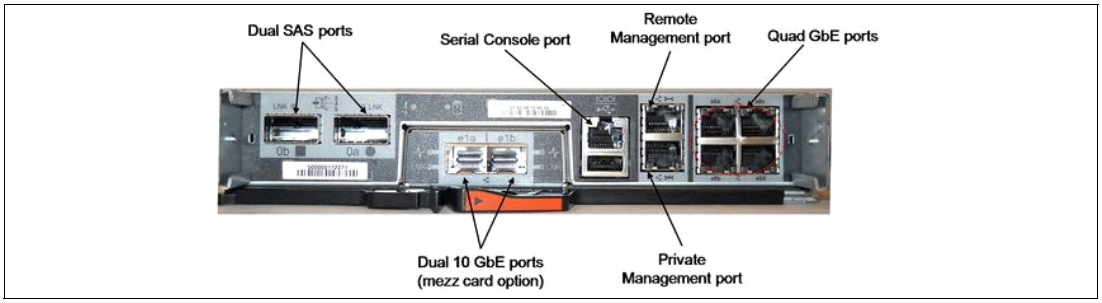

Figure 5-6 Controller with 10 GbE mezzanine card option

5.1.4 N32x0 common information

Table 5-2 lists ordering information for N32x0 systems.

Table 5-2 N32x0 configuration

Model	Form factor	HDD	PSU	Select PCM
N3220-A12, A22	2U chassis	24 SFF SAS 2.5”	2	One or two controllers, each with: ► No mezzanine card, or ► Dual FC mezzanine card, or ► Dual 10 GbE mezzanine card
N3240-A14, A24	4U chassis	24 SATA 3.5”	4	

Table 5-3 lists controller information for N32x0 systems with mezzanine cards.

Table 5-3 N32x0 controller configuration

Feature code	Configuration
	Controller with no mezzanine card (blank cover)
2030	Controller with dual-port FC mezzanine card (include SFP+)
2031	Controller with dual-port 10 GbE mezzanine card (no SFP+)

Table 5-4 lists information about the maximum number of supported shelves by expansion type.

Table 5-4 N3000 number of supported shelves

Expansion shelf (total 114 spindles)	Number of shelves supported
EXN 1000	Up to 6 Shelves (500 GB, 750 GB and 1 TB SATA disk drives)
ESN 3000	Up to 5 Shelves (300 GB, 450 GB, 600 GB, 900 GB SAS) or (500 GB, 1 TB, 2 TB, and 3 TB SATA disk drives)
EXN 3500	Up to 5 Shelves (450 GB, 600 GB, 900 GB SAS SFF disk drives)
EXN 4000	Up to 6 Shelves (144 GB, 300 GB, 450 GB and 600 GB F/C disk drives)

5.2 Mainstream and Advanced portfolio

Figure 5-7 on page 52 shows the N62x0 modular disk storage systems, which are designed to provide the following benefits:

- Increase NAS storage flexibility and expansion capabilities by consolidating block and file data sets onto a single multiprotocol storage platform
- Achieve performance when your applications need it most with high bandwidth, 64-bit architecture and the latest I/O technologies
- Maximize storage efficiency and growth and preserve investments in staff expertise and capital equipment with data-in-place upgrades to more powerful IBM System Storage N series
- Improve your business efficiency by using N6000 series capabilities, also available with a Gateway feature, to reduce data management complexity in heterogeneous storage environments for data protection and retention

Figure 5-7 Mainstream and Advanced systems

With IBM System Storage N62x0 series systems, you can meet your Network Attached Storage (NAS) needs and provide high levels of application availability for everything from critical business operations to technical applications. You can also address NAS and Storage Area Network (SAN) as primary and secondary storage requirements. In addition, you get outstanding value because our flexible systems offer excellent performance and impressive expendability at a low total cost of ownership.

Common features

The following list is an overview of common features:

- Simultaneous multiprotocol support for FCoE, FCP, iSCSI, CIFS, NFS, HTTP, and FTP
- File-level and block-level service in a single system
- Support for Fibre Channel, SAS and SATA disk drives
- Data ONTAP software
- Broad range of built-in features
- Multiple supported backup methods including disk-based and host-based backup and tape backup to direct, SAN, and GbE attached tape devices

Hardware summary

The following list is a hardware summary:

- Up to 2880 TB raw storage capacity
- 4 GB to 32 GB random access memory
- 512 MB to 4 GB nonvolatile memory
- Integrated Fibre Channel, Ethernet and SAS ports
- Quad-port 4 Gbps adapters (optional)
- Up to four Performance Acceleration Modules (Flash Cache)
- Diagnostic LED/LCD
- Dual redundant hot-plug integrated cooling fans and autoranging power supplies
- 19 inch, rack-mountable

N6240

The IBM System Storage N6240 storage controllers include the following models:

- Model C21, which is an active/active dual-node base unit
- Model E11, which is a single-node base unit
- Model E21, which is the coupling of two Model E11s

Exx models contain an I/O expansion module that provides additional PCIe slots. Note that I/O expansion is not available on Cxx models.

N6270

The IBM System Storage N6270 storage controllers include the following models:

- Model C22, which is an active/active dual-node base unit consisting of a single chassis with two controllers and no I/O expansion modules
- Model E12, which is a single-node base unit consisting of a single chassis with one controller and one I/O expansion module
- Model E22, which is the coupling of two E12 models

Exx models contain an I/O expansion module that provides additional PCIe slots. Note that I/O expansion is not available on Cxx models.

5.2.1 Common functions and features of mid-range models

This section describes the functions and features that are common to all mid-range models.

FC, SAS, and SATA attachment

FC, SAS, and SATA attachment options for disk expansion units are designed to allow deployment in multiple environments, including data retention, NearStore, disk-to-disk backup scenarios, and high performance, mission-critical I/O-intensive operations.

The IBM System Storage N series supports the EXN4000 FC storage expansion units, the EXN3000 SAS/SATA expansion unit, EXN3200 SATA expansion unit, and the EXN3500 SAS expansion unit. At least one storage expansion unit must be attached to the N series system.

All eight models must be mounted in a standard 19-inch rack. None of the eight models include storage in the base chassis.

Dynamic removal and insertion of the controller

The N6000 controllers are hot-pluggable. You do not have to turn off PSUs to remove a controller in a dual-controller configuration.

PSUs are independent components. One PSU can run an entire system indefinitely. There is no “two-minute rule” if you remove one PSU. PSUs have internal fans for self-cooling only.

RLM design and internal Ethernet switch on the controller

The Data ONTAP management interface, known as e0M, provides a robust and cost-effective way to segregate management subnets from data subnets without incurring a port penalty. On the N6000 series, the traditional RLM port on the rear of the chassis (now identified by a wrench symbol) connects first to an internal Ethernet switch that provides connectivity to the RLM and e0M interfaces.

Because the RLM and e0M each have unique TCP/IP addresses, the switch can discretely route traffic to either interface. You do not need to use a data port to connect to an external Ethernet switch. Setup of VLANs and VIFs is not required and not supported because e0M allows clients to have dedicated management networks without VLANs.

The e0M interface can be thought of as another way to remotely access and manage the storage controller, much like the serial console, RLM, or standard network interface. Use the e0M interface for network-based storage controller administration, monitoring activities, and ASUP reporting. The RLM is used when you require its higher level of support features. Connect host-side application data to the appliance on a separate subnet from the management interfaces.

RLM-assisted cluster failover

To decrease the time required for cluster failover (CFO) to occur when there is an event that the RLM is aware of, the RLM can communicate with the partner node instance of Data ONTAP. This capability was available in other N series models prior to the N6000 series, but the internal Ethernet switch makes the configuration much easier and facilitates quicker cluster failover, with some failovers occurring within 15 seconds.

5.2.2 N6240 and N6270 hardware overview

Figure 5-8 shows the IBM N6240/N6270 configuration flexibility.

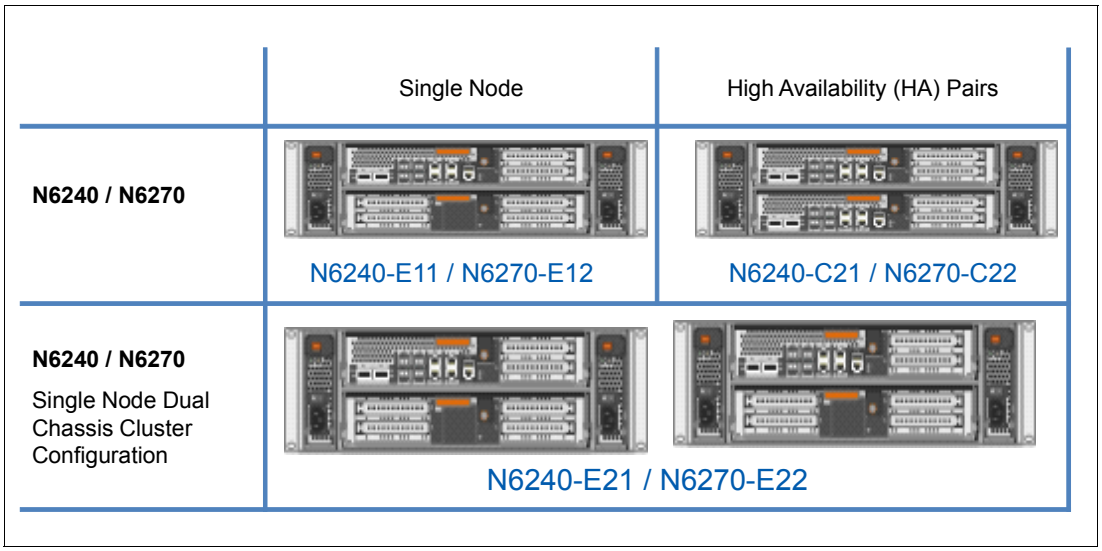

Figure 5-8 IBM N6240/N6270 configuration flexibility

The IBM N62x0 slots and interfaces for the controller module provide the following connectivity features:

- 2 PCIe v2.0 (Gen 2) x 8 slots
  - Top full height, full length
  - Bottom full height, ¾ length
- 2 x 6 Gb SAS (0a, 0b)
- 2 x HA interconnect (c0a, c0b)
- 2 x 4 Gb FCP (0c, 0d)
- 2 x GbE (e0a, e0b)
- USB port (not currently used)
- Management (wrench)
  - SP and e0M
- Private management
  - ACP (wrench w/lock)
- Serial console port
- I/O expansion module
  - 4 x PCIe 8x

– Full length, full height slots

Figure 5-9 shows the IBM N62x0 Controller I/O module.

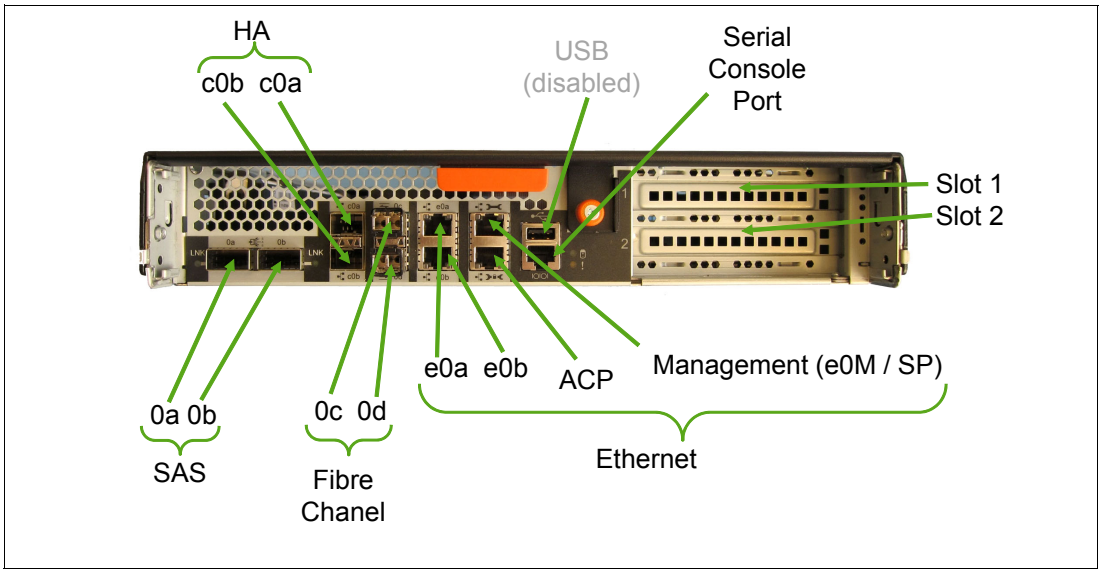

Figure 5-9 IBM N62x0 Controller I/O module

Figure 5-10 shows the N62x0 I/O configuration flexibility.

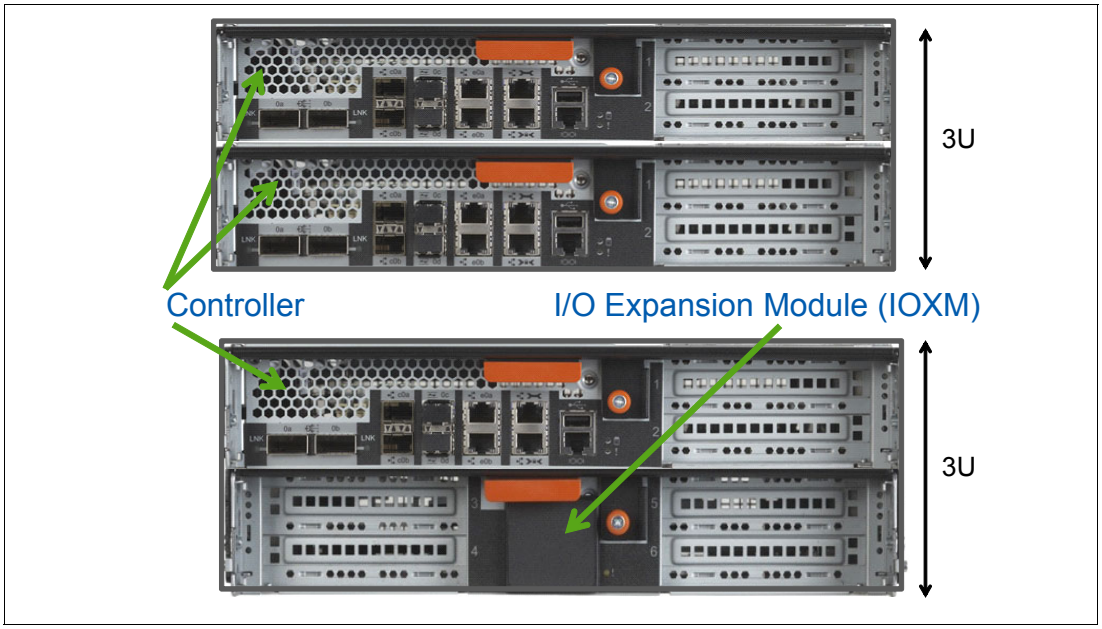

Figure 5-10 IBM N62x0 I/O configuration flexibility

The IBM N62x0 I/O Expansion Module (IOXM), which is displayed in Figure 5-10, has non-hot-swappable components. Note the following points:

- The controller will panic if removed.
- If inserted into a running IBM N6200, the IOXM will not be recognized until the controller is rebooted.
- It has 4 full-length PCIe v1.0 (Gen 1) x 8 slots.

5.3 Further information

Links for additional information about N series unified NAS storage solutions are listed here.

- For further N series 3000 systems information and specifications, refer to the following websites:

  http://www.ibm.com/systems/storage/network/n3000/appliance/index.html
  http://www.ibm.com/systems/storage/network/n3000/appliance/specifications.html

- For further N series 6000 systems information, refer to the following websites:

  http://www.ibm.com/systems/storage/network/n6000/appliance/index.html
  http://www.ibm.com/systems/storage/network/n6000/appliance/specifications.html

- For more detailed information about N series hardware features, see the IBM Redbooks publication *IBM System Storage N series Hardware Guide*, SG24-7840

  http://www.redbooks.ibm.com/abstracts/sg247840.html?Open

- For more detailed information about N series software features, see the IBM Redbooks publication *IBM System Storage N series Software Guide*, SG24-7129

  http://www.redbooks.ibm.com/abstracts/sg247129.html?Open

- IBM System Storage N series Machine Types and Models (MTM) Cross Reference

  http://www.ibm.com/support/docview.wss?uid=ssg1S7001844

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

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.

- *IBM System Storage N series Hardware Guide*, SG24-7840
- *IBM System Storage N series Software Guide*, SG24-7129
- *Managing Unified Storage with IBM System Storage N series Operation Manager*, SG24-7734
- *Using the IBM System Storage N series with IBM Tivoli Storage Manager*, SG24-7243
- *IBM System Storage N series and VMware vSphere Storage Best Practices,* SG24-7871
- *IBM System Storage N series with VMware vSphere 4.1*, SG24-7636
- *Designing an IBM Storage Area Network*, SG24-5758
- *Introduction to Storage Area Networks and System Networking*, SG24-5470
- *IP Storage Networking: IBM NAS and iSCSI Solutions*, SG24-6240
- *Storage and Network Convergence Using FCoE and iSCSI,* SG24-7986
- *IBM Data Center Networking: Planning for Virtualization and Cloud Computing*, SG24-7928
- *IBM N Series Storage Systems in a Microsoft Windows Environment*, REDP-4083
- *Using an IBM System Storage N series with VMware to Facilitate Storage and Server Consolidation*, REDP-4211
- *IBM System Storage N series MetroCluster,* REDP-4259
- *IBM System Storage N series with FlexShare*, REDP-4291
- *IBM System Storage N series A-SIS Deduplication Deployment and Implementation Guide*, REDP-4320
- *IBM System Storage N series with VMware vSphere 4.1 using Virtual Storage Console 2*, REDP-4863

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

Other publications

These publications are also relevant as further information sources:

- Network-attached storage

  http://www.ibm.com/systems/storage/network/

- IBM support documentation

  http://www.ibm.com/support/entry/portal/Documentation

- IBM Storage – Network Attached Storage: Resources

  http://www.ibm.com/systems/storage/network/resources.html

- IBM System Storage N series Machine Types and Models (MTM) Cross Reference

  http://www-304.ibm.com/support/docview.wss?uid=ssg1S7001844

- IBM N Series to NetApp Machine type comparison table

  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD105042

- Interoperability matrix

  http://www-304.ibm.com/support/docview.wss?uid=ssg1S7003897

- VMware documentation

  http://www.vmware.com/support/pubs/

- VMware vSphere 5 documentation

  http://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html
  http://pubs.vmware.com/vsphere-50/index.jsp

- VMware Capacity Planner

  http://www.vmware.com/products/capacity-planner/

- VMware vSphere 4.1 configurations maximum

  http://www.vmware.com/pdf/vsphere4/r41/vsp_41_config_max.pdf

- VMware vCloud suite

  http://www.vmware.com/products/datacenter-virtualization/vcloud-suite/overview.html

- Microsoft Mailbox Server Storage Design

  http://technet.microsoft.com/en-us/library/dd346703.aspx

- Microsoft Mailbox Server Processor Capacity Planning

  http://technet.microsoft.com/en-us/library/ee712771.aspx

- Microsoft Planning and architecture for SharePoint Server 2010

  http://technet.microsoft.com/en-us/library/cc261834.aspx

- Microsoft Hardware and Software Requirements for Installing SQL Server 2012

  http://technet.microsoft.com/en-us/library/ms143506.aspx

Online resources

These websites are also relevant as further information sources:

- IBM NAS support website

  http://www.ibm.com/storage/support/nas/

- NAS product information

  http://www.ibm.com/storage/nas/

- IBM Integrated Technology Services

  http://www.ibm.com/planetwide/

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM System Storage N series Reference Architecture for Virtualized Environments

Learn how to integrate VMware vSphere with IBM System x and N series

Understand the importance of storage as a foundation for a cloud solution

Design VMware vSphere solutions with N series

This IBM Redpaper publication provides deployment guidelines, workload estimates, and preferred practices for clients who want a proven IBM technology stack for VMware environments. The result is a Reference Architecture for Virtualized Environments (RAVE) that uses VMware vSphere, IBM System x or IBM BladeCenter server, IBM System Networking, and IBM System Storage N series as a storage foundation. The reference architecture can be used as a foundation to create dynamic cloud solutions and make full use of underlying storage features and functions.

This paper provides a blueprint that illustrates how clients can create a virtualized infrastructure and storage cloud to help address current and future data storage business requirements. It explores the solutions that IBM offers to create a storage cloud solution addressing client needs. This paper also shows how the Reference Architecture for Virtualized Environments and the extensive experience of IBM in cloud computing, services, proven technologies, and products support a Smart Storage Cloud solution that is designed for your storage optimization efforts.

When used as the storage foundation, System Storage N series offers unified storage solutions. These solutions provide industry-leading technologies for storage efficiencies, instantaneous virtual machine and data store cloning for virtual servers and virtual desktops, and virtual data center backup and business continuance solutions.

This paper is for anyone who wants to learn how to successfully deploy a virtualized environment. It is also written for anyone who wants to understand how IBM addresses data storage and compute challenges with IBM System Storage N series solutions with IBM servers and networking solutions. This paper is suitable for IT architects, business partners, IBM clients, storage solution integrators, and IBM sales representatives.

REDP-4865-00

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks