FUTURE VISION BIE

One Stop for All Study Materials

& Lab Programs

Future Vision

By K B Hemanth Raj

Scan the QR Code to Visit the Web Page

Or

Visit : https://hemanthrajhemu.github.io

Gain Access to All Study Materials according to VTU,
CSE – Computer Science Engineering,
ISE – Information Science Engineering,
ECE - Electronics and Communication Engineering
& MORE...

Join Telegram to get Instant Updates: https://bit.ly/VTU_TELEGRAM

Contact: MAIL: futurevisionbie@gmail.com

INSTAGRAM: www.instagram.com/hemanthraj_hemu/

INSTAGRAM: www.instagram.com/futurevisionbie/

WHATSAPP SHARE: https://bit.ly/FVBIESHARE

BMS
Institute of Technology and Management

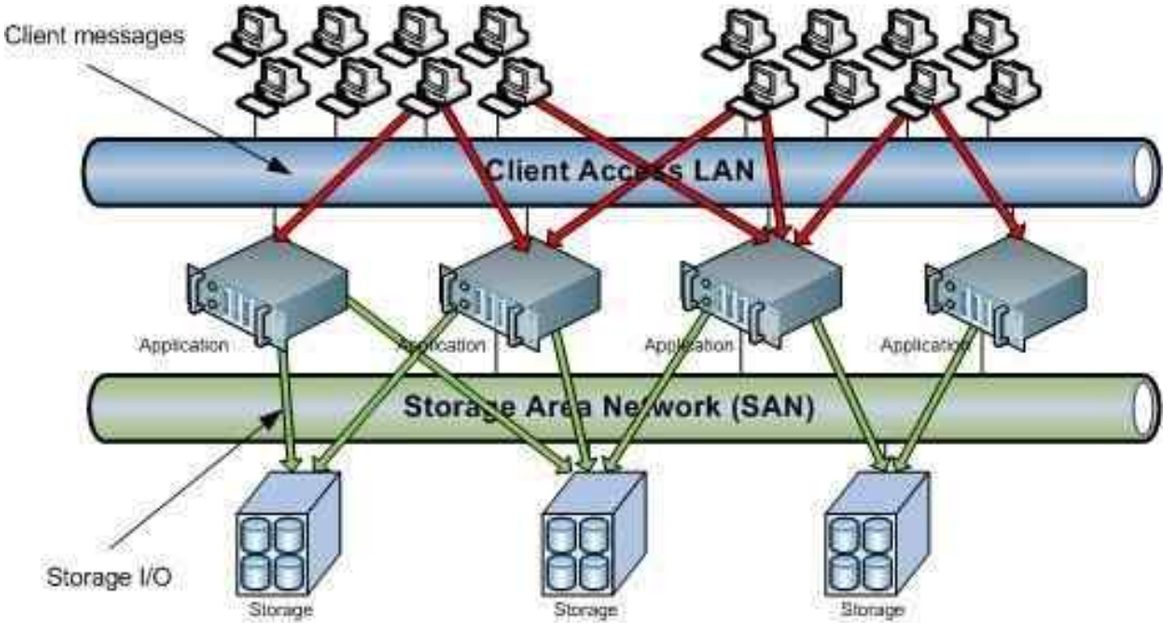
Avalahalli, Doddaballapur Main Road, Bengaluru – 560064

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Storage Area Networks

(17CS754)

SANs are primarily used to access storage devices, such as disk arrays and tape libraries from servers so that the devices appear to the operating system as direct- attached storage.



STORAGE AREA NETWORKS
[As per Choice Based Credit System (CBCS) scheme]
(Effective from the academic year 2017 - 2018) SEMESTER – VII

Subject Code	17CS754	IA Marks	40
Number of Lecture Hours/Week	3	Exam Marks	60
Total Number of Lecture Hours	40	Exam Hours	03

CREDITS – 03

Module – 1	**Teaching Hours**
Storage System Introduction to evolution of storage architecture, key data centre Elements, virtualization, and cloud computing. Key data centre elements – Host (or compute), connectivity, storage, and application in both classic and virtual Environments. RAID implementations, techniques, and levels along with the Impact of RAID on application performance. Components of intelligent storage systems and virtual storage provisioning and intelligent storage system Implementations.	**8 Hours**
Module – 2	
Storage Networking Technologies and Virtualization Fibre Channel SAN components, connectivity options, and topologies including access protection mechanism „zoning", FC protocol stack, addressing and operations, SAN-based virtualization and VSAN technology, iSCSI and FCIP(Fibre Channel over IP) protocols for storage access over IP network, Converged protocol FCoE and its components, Network Attached Storage (NAS) - components, protocol and operations, File level storage virtualization, Object based storage and unified storage platform.	**8 Hours**
Module – 3	
Backup, Archive, and Replication This unit focuses on information availability and business continuity solutions in both virtualized and non-virtualized environments. Business continuity terminologies, planning and solutions, Clustering and multipathing architecture to avoid single points of failure, Backup and recovery - methods, targets and topologies, Data deduplication and backup in virtualized environment, Fixed content and data archive, Local replication in classic and virtual environments, Remote replication in classic and virtual environments, Three-site remote replication and continuous data protection	**8 Hours**
Module – 4	
Cloud Computing Characteristics and benefits This unit focuses on the business drivers, definition, essential characteristics, and phases of journey to the Cloud. ‚Business drivers for Cloud computing, Definition of Cloud computing, Characteristics of Cloud computing, Steps involved in transitioning from Classic data center to Cloud computing environment Services and deployment models, Cloud infrastructure components, Cloud migration considerations	**8 Hours**
Module – 5	
Securing and Managing Storage Infrastructure This chapter focuses on framework and domains of storage security along with covering security. implementation at storage networking. Security threats, and countermeasures in various domains Security solutions for (Fiber channel)FC-SAN, IP-SAN, and NAS	**8 Hours**

managing various information infrastructure components in classic and virtual environments, Information lifecycle management (ILM) and storage tiering, Cloud service management activities	

Course outcomes: The students should be able to:

- Identify key challenges in managing information and analyze different storage networking technologies and virtualization
- Explain components and the implementation of NAS
- Describe CAS architecture and types of archives and forms of virtualization
- Illustrate the storage infrastructure and management activities

Question paper pattern:

The question paper will have ten questions.
There will be 2 questions from each module.
Each question will have questions covering all the topics under a module.
The students will have to answer 5 full questions, selecting one full question from each module.

Text Books:

1. Information Storage and Management, Author :EMC Education Services, Publisher: Wiley ISBN: 9781118094839
2. Storage Virtualization, Author: Clark Tom, Publisher: Addison Wesley Publishing Company ISBN: 9780321262516

Table of Content

Module – 5

Securing and Managing Storage Infrastructure

Information Security Framework

The basic information security framework is built to achieve four security goals: confidentiality, integrity, and availability (CIA), along with accountability. This frameworkincorporatesall security standards, procedures, and controls, required to mitigate threats in the storage infrastructure environment.

- **Confidentiality:** Provides the required secrecy of information and ensures thatonlyauthorizedusershaveaccesstodata. Thisrequiresauthentica- tion of users who need to access information.

  Data in transit (data transmitted over cables) and data at rest (data residing on a primary storage, backup media, or in the archives) can be encrypted to maintain its confidentiality. In addition to restricting unauthorized users from accessing information, confidentiality also requires implementing traffic flow protection measures as part of the security protocol. These protection measures generally include hiding source and destinationaddresses, frequencyof databeingsent, andamount of datasent.

- **Integrity:** Ensures that the information is unaltered. Ensuring integrity requires detection of andprotection against unauthorized alteration or deletion of information. Ensuring integrity stipulatesmeasures suchas error detection and correction for both data and systems.

- **Availability:** Thisensuresthatauthorizedusershavereliableand timely access to systems, data, and applications residing on these systems. Availability requires protection against unauthorized deletion of data and denial of service (discussed in section —14.2.2 Threatsǁ). Availability alsoimplies thatsufficient resources are available to provide aservice.

- **Accountabilityservice:**Referstoaccounting forallthe eventsand operationsthattakeplaceinthedatacenterinfrastructure. Theaccountability service maintains a log of events that can be audited or traced later for the purpose of security.

Risk Triad

Risk triad defines risk in terms of threats, assets, and vulnerabilities. Risk arises when a threat agent (an attacker) uses an existing vulnerability to compromise the security services of an asset, for example, if a sensitive document is transmitted without any protection over an insecure channel, an attacker might get unauthorized access to the document and may violate its confidentiality and integrity. This may, in turn, result in business loss for the organization. In this scenario potential business loss is the risk, which arises because an attacker

uses the vulnerability of the unprotected communication to access the document and tamper with it.

To manage risks, organizations primarily focus on vulnerabilities because they cannot eliminate threat agents that appear in various forms and sources to its assets. Organizations can enforce countermeasures to reduce the possibility of occurrence of attacks and the severity of their impact.

Risk assessment is the first step to determine the extent of potential threats and risks in an IT infrastructure. The process assesses risk and helps to identify appropriate controls to mitigate or eliminate risks. Based on the value of assets, risk assessment helps to prioritize investment in and provisioning of security measures. To determine the probability of an adverse event occurring, threats to an IT system must be analyzed with the potential vulnerabilities and the existing security controls.

The severity of an adverse event is estimated by the impact that it may have on critical business activities. Based on this analysis, a relative value of criti- cality and sensitivity can be assigned to IT assets and resources. For example, a particular IT system component may be assigned a high-criticality value if an attack on this particular component can cause a complete termination of mission-critical services.

The following sections examine the three key elements of the risk triad. Assets, threats, and vulnerabilities are considered from the perspective of risk identification and control analysis.

Assets

Information is one of the mostimportant *assets* foranyorganization. Other assets include hardware, software, and other infrastructure components required to access the information. To protect these assets, organizations must develop a set of parameters to ensure the availability of the resources to authorized users and trusted networks. These parameters apply to storage resources, network infrastructure, and organizational policies.

Security methods have two objectives. The first objective is to ensure that the network is easily accessible to authorized users. It should also be reliable and stable under disparate environmental conditions and volumes of usage. The second objective is to make it difficult for potential attackers to access and compromise the system.

The security methods should provide adequate protection against unauthor- ized access, viruses, worms, trojans, and other malicious software programs. Security measures should also include options to encrypt critical data and dis- able unused services to minimize the number of potential security gaps. The security method must ensure that updates to the operating system and other software are installed regularly. At the same time, it must provide adequate redundancy in the form of replication and mirroring of the production data

to prevent catastrophic data loss if there is an unexpected data compromise. For the security system to function smoothly, all users are informed about the policies governing the use of the network.

The effectiveness of a storage security methodology can be measured by two key criteria. One, the cost of implementing the system should be a fraction of the value of the protected data. Two, it should cost heavily to a potential attacker, in terms of money, effort, and time.

Threats

Threats are the potential attacks that can be carried out on an IT infrastructure. These attacks can be classified as active or passive. *Passive attacks* are attempts to gain unauthorized access into the system. They pose threats to confidentiality of information. *Active attacks* include data modification, denial of service (DoS), and repudiation attacks. They pose threats to data integrity, availability, and accountability. In a data modification attack, the unauthorized user attempts to modify information for malicious purposes. A modification attack can target the data at rest or the data in transit. These attacks pose a threat to data integrity.

Denial of service (DoS) attacks prevent legitimate users from accessing resources and services. These attacks generally do not involve access to or modification of information. Instead, they pose a threat to data availability. The intentional flooding of a network or website to prevent legitimate access to authorized users is one example of a DoS attack.

Repudiation is an attack against the accountability of information. It attempts to provide false information by either impersonating someone or denying that an event or a transaction has taken place. For example, a repudiation attack may involve performing an action and eliminating any evidence that could prove the identity of the user (attacker) who performed that action. Repudiation attacks include circumventing the logging of security events or tampering with the security log to conceal the identity of the attacker.

Vulnerability

The paths that provide access to information are often vulnerable to potential attacks. Each of the paths may contain various access points, which provide different levels of access to the storage resources. It is important to implement adequatesecurity controls at all the accesspoints on an accesspath. Implementing security controls at each access point of every access path is known as *defense in depth*.

Defense in depth recommends using multiple security measures to reduce the risk of security threats if one component of the protection is compromised. It is also known as a —layered approach to security.‖ Because there are multiple measures for security at different levels, defense in depth gives additional time to detect and respond to an attack. This can reduce the scope or impact of a security breach.

Attack surface, *attack vector*, and *work factor* are the three factors to consider when assessing the extent to which an environment is vulnerable to security threats. *Attack surface* refers to the various entry points that an attacker can use to launch an attack. Each component of a storage network is a source of potential vulnerability. An attacker can use all the external interfaces supported by that component, such as the hardware and the management interfaces, to execute various attacks. These interfaces form the attack surface for the attacker. Even unused network services, if enabled, can become a part of the attack surface.

An *attack vector* is a step or a series of steps necessary to complete an attack. For example, an attacker might exploit a bug in the management interface to execute a snoop attack whereby the attacker can modify the configuration of the storage device to allow the traffic to be accessed from one more host. This

redirected traffic can be used to snoop the data in transit.

Work factor refers to the amount of time and effort required to exploit an attack vector. For example, if attackers attempt to retrieve sensitive information, they consider the time and effort that would be required for executing an attack on a database. This may include determining privileged accounts, determining the database schema, and writing SQL queries. Instead, based on the work factor, they may consider a less effort-intensive way to exploit the storage array by attaching to it directly and reading from the raw disk blocks.

Having assessed the vulnerability of the environment, organizations can deploy specific control measures. Any control measures should involve all the three aspects of infrastructure: people, process, and technology, and the relationships among them. To secure people, the first step is to establish and assure

their identity. Based on their identity, selective controls can be implemented for

their access to data and resources. The effectiveness of any security measure is primarily governed by processes and policies. The processes should be based on a thorough understanding of risks in the environment and should recog- nize the relative sensitivity of different types of data and the needs of various stakeholders to access the data. Without an effective process, the deployment

of technology is neither cost-effective nor aligned to organizations' priorities. Finally, the technologies or controls that are deployed should ensure compliance with the processes, policies, and people for its effectiveness. These secu- rity technologies are directed at reducing vulnerability by minimizing attack surfaces and maximizing the work factors. These controls can be technical or nontechnical. Technical controls are usually implemented through computer systems, whereas nontechnical controlsare implementedthrough administrative and physical controls. Administrative controls include security and personnel policies or standard procedures to direct the safe execution of various opera- tions. Physical controls include setting up physical barriers, such as security guards, fences, or locks.

Based on the roles they play, controls are categorized as preventive, detective, and corrective. The preventive control attempts to prevent an attack; the detective control detects whether an attack is in progress; and after an attack is discovered, the corrective controls are implemented. *Preventive controls* avert the vulnerabilities from being exploited and prevent an attack or reduce its

impact. *Corrective controls* reduce the effect of an attack, whereas *detective controls*

discover attacks and trigger preventive or corrective controls. For example, an Intrusion Detection/Intrusion Prevention System(IDS/IPS) isadetective control that determines whether an attack is underway and then attempts to stop it by terminating a network connection or invoking a firewall rule to block traffic.

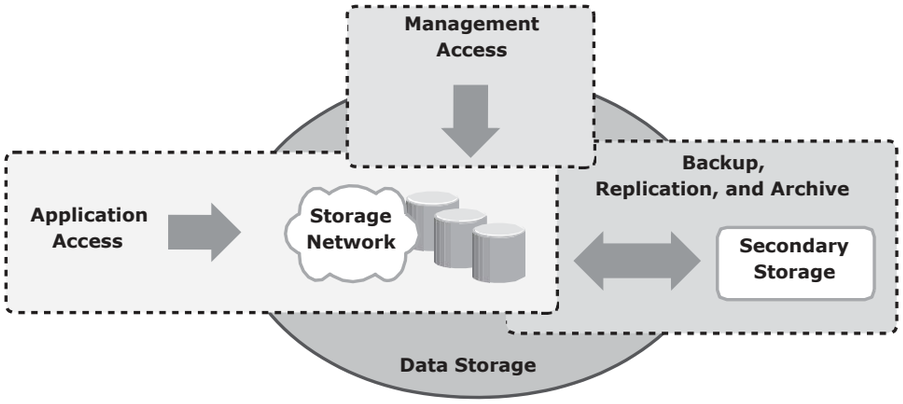
Storage Security Domains

Storage devices connected to a network raise the risk level and are more exposed to security threats via networks. However, with increasing use of networking in storage environments, storage devices are becoming highly exposed to secu- rity threats from a variety of sources. Specific controls must be implemented to secure a storage networking environment. This requires a closer look at storage networking security and a clear understanding of the access paths leading to storage resources. If a particular pathis unauthorized and needs to be prohibited by technical controls, ensure that these controls are not compromised. If each component within the storage network is considered a potential access point, the attack surface of all these access points must be analyzed to identify the associated vulnerabilities.

To identify the threats that apply to a storage network, access paths to data storage can be categorized into three security domains: *application access*, *management access*, and *backup, replication, and archive*. - 14-1 depicts the three security domains of a storage system environment.

Thefirst securitydomaininvolvesapplicationaccess to the storeddatathrough the storage network. The second security domain includes management access to storage and interconnect devices and to the data residing on those devices.

This domain is primarily accessed by storage administrators who configure and manage the environment. The third domain consists of backup, replication, and archive access. Along with the access points in this domain, the backup media also needs to be secured.



- 14-1: Storage security domains

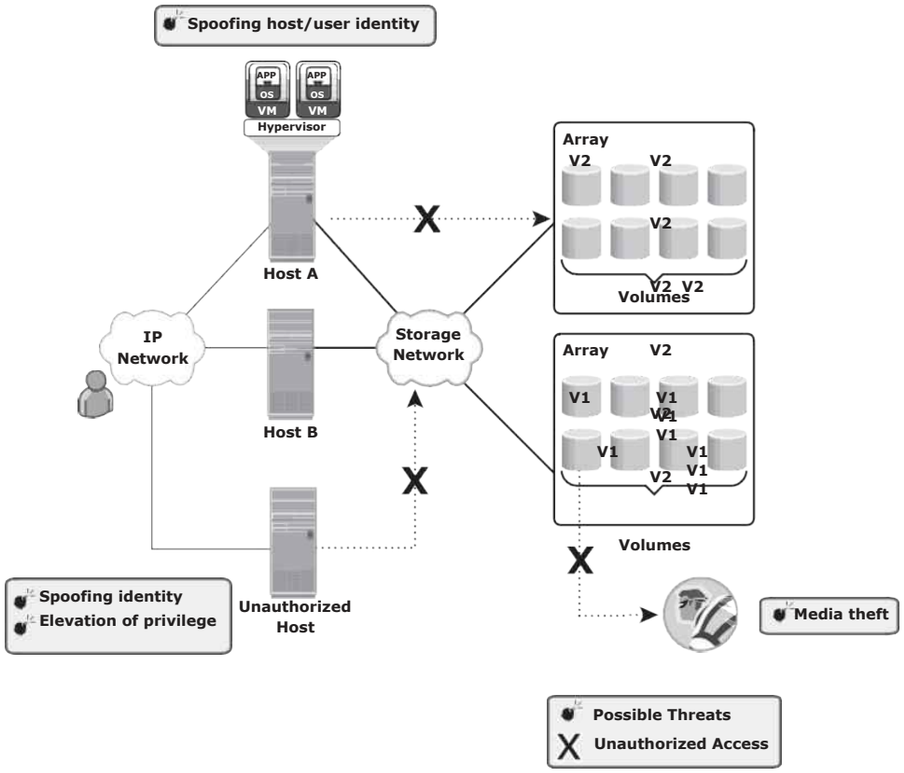
To secure the storage networking environment, identify the existing threats within each of the security domains and classify the threats based on the type of security services — availability, confidentiality, integrity, and accountability. The next step is to select and implement various controls as countermeasures to the threats.

Securing the Application Access Domain

The *application access domain* may include only those applications that access the
data through the file system or a database interface.

An important step to secure the application access domain is to identify the threats in the environment and appropriate controls that should be applied. Implementing physical security is also an important consideration to prevent media theft.

- 14-2 showsapplicationaccessin a storagenetworkingenvironment. Host A can access all V1 volumes; host B can access all V2 volumes. These volumes are classified according to the access level, such as confidential, restricted, and public. Some of the possible threats in this scenario could be host A spoofing the identity or elevating to the privileges of host B to gain access to host B'sresources. Another threat could be that an unauthorized host gains access to the network; the attacker on this host may try to spoof the identity of another host and tam- per with the data, snoop the network, or execute a DoS attack. Also any form of media theft could also compromise security. These threats can pose several serious challenges to the network security; therefore, they need to be addressed.



- 14-2: Security threats in an application access domain

Controlling User Access to Data

Access control services regulate user access to data. These services mitigate the threats of spoofing host identity and elevating host privileges. Both these threats affect data integrity and confidentiality.

Access control mechanisms used in the application access domain are user and host authentication (technical control) and authorization (administrative control). These mechanisms may lie outside the boundaries of the storage net- work and require various systems to interconnect with other enterprise identity management and authentication systems, for example, systems that provide strong authentication and authorization to secure user identities against spoof- ing. NAS devices support the creation of *access control lists* that regulate user access to specific files. The Enterprise Content Management application enforces access to data by using Information Rights Management (IRM) that specifies which users have what rights to a document. Restricting access at the host level starts with authenticating a node when it tries to connect to a network.

Different storage networking technologies, such as iSCSI, FC, and IP-based storage, use various authentication mechanisms, such as Challenge-Handshake Authentication Protocol (CHAP), Fibre Channel Security Protocol (FC-SP), and IPSec, respectively, to authenticate host access.

After a host has been authenticated, the next step is to specify security con- trols for the storage resources, such as ports, volumes, or storage pools, that the host is authorized to access. *Zoning* is a control mechanism on the switches that segments the network into specific paths to be used for data traffic; *LUN masking* determines which hosts can access which storage devices. Some devices support mapping of a host's WWN to a particular FC port and from there to a particular LUN. This binding of the WWN to a physical port is the most secure. Finally, it is important to ensure that administrative controls, such as defined security policies and standards, are implemented. Regular auditing is required to ensure proper functioning of administrative controls. This is enabled by log- ging significant events on all participating devices. Event logs should also be protected from unauthorized access because they may fail to achieve their goals if the logged content is exposed to unauthorized modifications by an attacker.

Protecting the Storage Infrastructure

Securing the storage infrastructure from unauthorized access involves pro- tecting all the elements of the infrastructure. Security controls for protecting the storage infrastructure address the threats of unauthorized tampering of data in transit that leads to a loss of data integrity, denial of service that compromises availability, and network snooping that may result in loss of confidentiality.

The security controls for protecting the network fall into two general cat- egories: *network infrastructure integrity* and *storage network encryption*. Controls for ensuring the infrastructure integrity include a fabric switch function that ensures fabric integrity. This is achieved by preventing a host from being added to the SAN fabric without proper authorization. Storage network encryption methods include the use of IPSec for protecting IP-based storage networks, and FC-SP for protecting FC networks.

In secure storage environments, root or administrator privileges for a specific device are not granted to every user. Instead, *role-based access control* (RBAC) is deployed to assign necessary privileges to users, enabling them to perform their roles. A role may represent a job function, for example, an administrator. Privileges are associated with the roles and users acquire these privileges based upon their roles.

It is also advisable to consider administrative controls, such as —separation of duties,ll when defining data center procedures. Clear separation of duties ensures that no single individual can both specify an action and carry it out. For example, the person who authorizes the creation of administrative accounts

should not be the person who uses those accounts. Securing management access is covered in detail in the next section.

Management networks for storage systems should be logically separate from other enterprise networks. This segmentation is critical to facilitate ease of management and increase security by allowing access only to the components existing within the same segment. For example, IP network segmentation is enforced with the deployment of filters at Layer 3 by using routers and firewalls, and at Layer 2 by using VLANs and port-level security on Ethernet switches.

Finally, physical access to the device console and the cabling of FC switches must be controlled to ensure protection of the storage infrastructure. All other
established security measures fail if a device is physically accessed by an unau- thorized user; this access may render the device unreliable.

Data Encryption

The most important aspect of securing data is protecting data held inside the storage arrays. Threats at this level include tampering with data, which violates data integrity, and media theft, which compromises data availability and confi- dentiality. To protect against these threats, encrypt the data held on the storage media or encrypt the data prior to being transferred to the disk. It is also criti- cal to decide upon a method for ensuring that data deleted at the end of its life cycle has been completely erased from the disks and cannot be reconstructed for malicious purposes.

Data should be encrypted as close to its origin as possible. If it is not possible to perform encryption on the host device, an encryption appliance can be used for encrypting data at the point of entry into the storage network. Encryption devices can be implemented on the fabric that encrypts data between the host and the storage media. These mechanisms can protect both the data at rest on the destination device and data in transit.

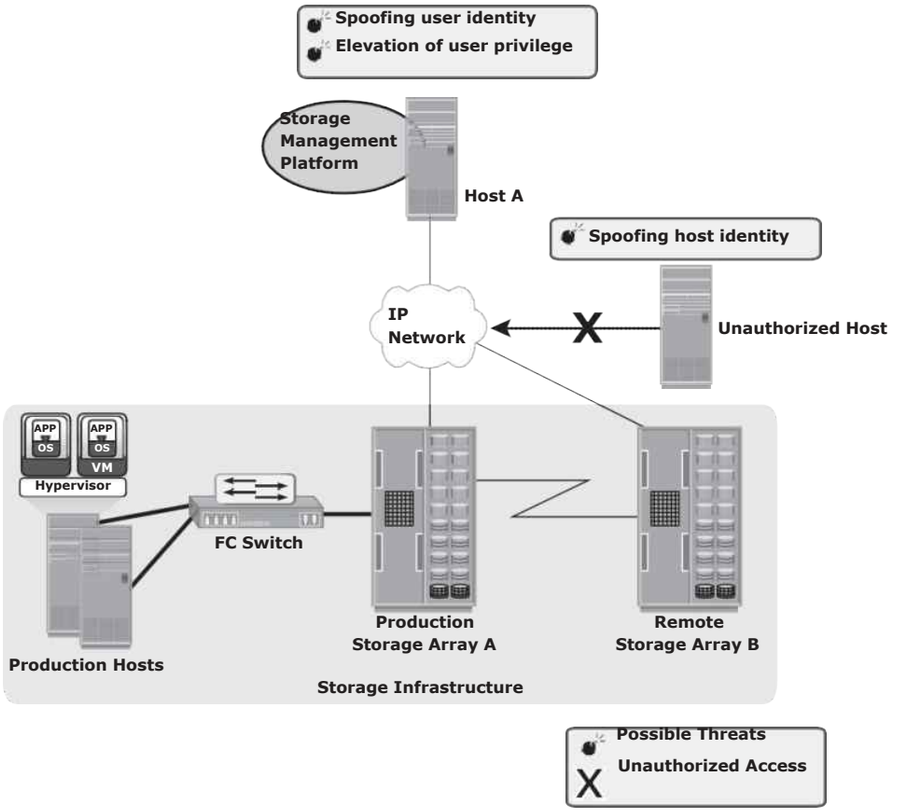
On NAS devices, adding antivirus checks and file extension controls can further enhance data integrity. In the case of CAS, use of MD5 or SHA-256 cryptographic algorithms guarantees data integrity by detecting any change in content bit patterns. In addition, the data erasure service ensures that the data
has been completely overwritten by bit sequence before the disk is discarded. An organization's data classification policy determines whether the disk should
actually be scrubbed prior to discarding it and the level of erasure needed based on regulatory requirements.

Securing the Management Access Domain

Management access, whether monitoring, provisioning, or managing storage resources, is associated with every device within the storage network. Most management software supports someform of CLI, system management console,

or a web-based interface. Implementing appropriatecontrols forsecuring storage management applications is important because the damage that can be caused by using these applications can be far more extensive.

- 14-3 depicts a storage networking environment in which production hosts are connected to a SAN fabric and are accessing production storage array A, which is connected toremotestorage array B forreplication purposes. Further, this configuration has a storage management platform on Host A. A possible threat in this environment is an unauthorized host spoofing the user or host identity to manage the storage arrays or network. For example, an unauthorized host may gain management access to remote array B.



- 14-3: Security threats in a management access domain

Providing management access through an external network increases the potential for an unauthorized host or switch to connect to that network. In such circumstances, implementing appropriate security measures prevents certain types of remote communication from occurring. Using secure communication

channels, such as Secure Shell (SSH) or Secure Sockets Layer (SSL)/Transport Layer Security (TLS), provides effective protection against these threats. Event logmonitoring helps to identify unauthorized access and unauthorized changes to the infrastructure. Event logs should be placed outside the shared storage systems where they can be reviewed if the storage is compromised.

The storage management platform must be validated for available security controls and ensures that these controls are adequateto secure the overall storage environment. The administrator's identity and role should be secured against any spoofing attempts so that an attacker cannot manipulate the entire storage array and cause intolerable data loss by reformatting storage media or making data resources unavailable.

Controlling Administrative Access

Controlling administrative access to storage aims to safeguard against the threats of an attacker spoofing an administrator'sidentity or elevating privileges to gain administrative access. Both of these threats affect the integrity of data and devices. To protect against these threats, administrative access regulation and various auditing techniques are used to enforce accountability of users and processes. Access control should be enforced for each storage component. In some storage environments, it may be necessary to integrate storage devices withthird-party authenticationdirectories, such as Lightweight Directory Access Protocol (LDAP) or Active Directory.

Security best practices stipulate that no single user should have ultimate con- trol over all aspects of the system. If an administrative user is a necessity, the number of activities requiring administrative privileges should be minimized. Instead, it is better to assign various administrative functions by using RBAC. Auditing logged events is a critical control measure to track the activities of an administrator. However, access to administrative log files and their content must be protected. Deploying a reliable Network Time Protocol on each system that can be synchronized to a common time is another important requirement to ensure that activities across systems can be consistently tracked. In addition, having a Security Information Management (SIM) solution supports effective analysis of the event log files.

Protecting the Management Infrastructure

Mechanisms to protect the management networkinfrastructure include encrypt- ing management traffic, enforcing management access controls, and applying IP network security best practices. These best practices include the use of IP routers and Ethernet switches to restrict the traffic to certain devices. Restricting network activity and access to a limited set of hosts minimizes the threat of an unauthorized device attaching to the network and gaining access to the

management interfaces. Access controls need tobe enforced at the storage-array level to specify which host has management access to which array. Some stor- age devices and switches can restrict management access to particular hosts and limit the commands that can be issued from each host.

A separate privatemanagement network is highly recommended formanage- ment traffic. If possible, management traffic should not be mixed with either production data traffic or other LAN traffic used in the enterprise. Unused network services must be disabled on every device within the storage network. This decreases the attack surface for that device by minimizing the number of interfaces through which the device can be accessed.

To summarize, security enforcement must focus on the management communication between devices, confidentiality and integrity of management data, and availability of management networks and devices.

Securing Backup, Replication, and Archive

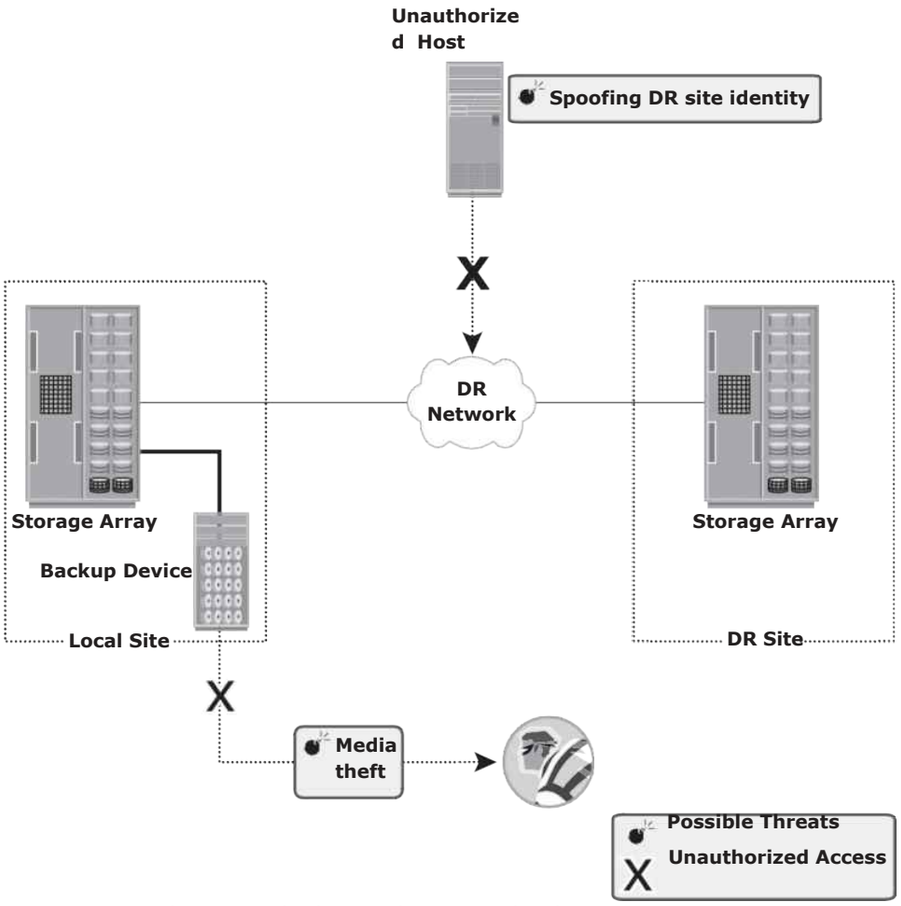
Backup, replication, and archive is the third domain that needs to be secured against an attack. As explained in Chapter 10, a backup involves copying the data from a storage array to backup media, such as tapes or disks. Securing backup is complex and is based on the backup software that accesses the storage arrays. It also depends on the configuration of the storage environments at the primary and secondary sites, especially with remote backup solutions performed directly on a remote tape device or using array-based remote replication.

Organizations must ensure that the disaster recovery (DR) site maintains the same level of security for the backed up data. Protecting the backup, replication, and archive infrastructure requires addressing several threats, including spoof- ing the legitimate identity of a DR site, tampering with data, network snooping, DoS attacks, and media theft. Such threats represent potential violations of integrity, confidentiality, and availability. - 14-4 illustrates a generic remote backup design whereby data on a storage array is replicated over a DR network to a secondary storage at the DR site. In a remote backup solution where the storage components are separated by a network, the threats at the transmission layer need to be countered. Otherwise, an attacker can spoof the identity of the backup server and request the host to send its data. The unauthorized host claiming to be the backup server may lead to a remote backup being performed to an unauthorized and unknown site. In addition, attackers can use the DR network connection to tamper with data, snoop the network, and create a DoS attack against the storage devices.

The physical threat of a backup tape being lost, stolen, or misplaced, especially

if the tapes contain highly confidential information, is another type of threat.

Backup-to-tape applications are vulnerable to severe security implications if they do not encrypt data while backing it up.



- 14-4: Security threats in a backup, replication, and archive environment

Security Implementations in Storage Networking

The following discussion details some of the basic security implementations in FC SAN, NAS, and IP-SAN environments.

FC SAN

Traditional FC SANs enjoy an inherent security advantage over IP-based networks. An FC SAN is configured as an isolated private environment with fewer nodes than an IP network. Consequently, FC SANs impose fewer security

threats. However, this scenario has changed with converged networks and storage consolidation, driving rapid growth and necessitating designs for large, complex SANs that span multiple sites across the enterprise. Today, no single comprehensive security solution is available for FC SANs. Many FC SAN secu- rity mechanisms have evolved from their counterpart in IP networking, thereby bringing in matured security solutions.

Fibre Channel Security Protocol (FC-SP) standards (T11 standards), published in 2006, align security mechanisms and algorithms between IP and FC inter- connects. These standards describe protocols to implement security measures in a FC fabric, among fabric elements and N_Ports within the fabric. They also include guidelines for authenticating FC entities, setting up session keys, negotiating the parameters required to ensure frame-by-frame integrity and confidentiality, and establishing and distributing policies across an FC fabric.

FC SAN Security Architecture

Storage networking environments are a potential target for unauthorized access, theft, and misuse because of the vastness and complexity of these environments. Therefore, security strategies are based on the *defense in depth* concept, which recommends multiple integrated layers of security. This ensures that the failure of one security control will not compromise the assets under protection. - 14-5 illustrates various levels (zones) of a storage network- ing environment that must be secured and the security measures that can be deployed.

FC SANsnotonly suffer from certain risks and vulnerabilities that are unique, but also share common security problems associated with physical security and remote administrative access. In addition to implementing SAN-specific

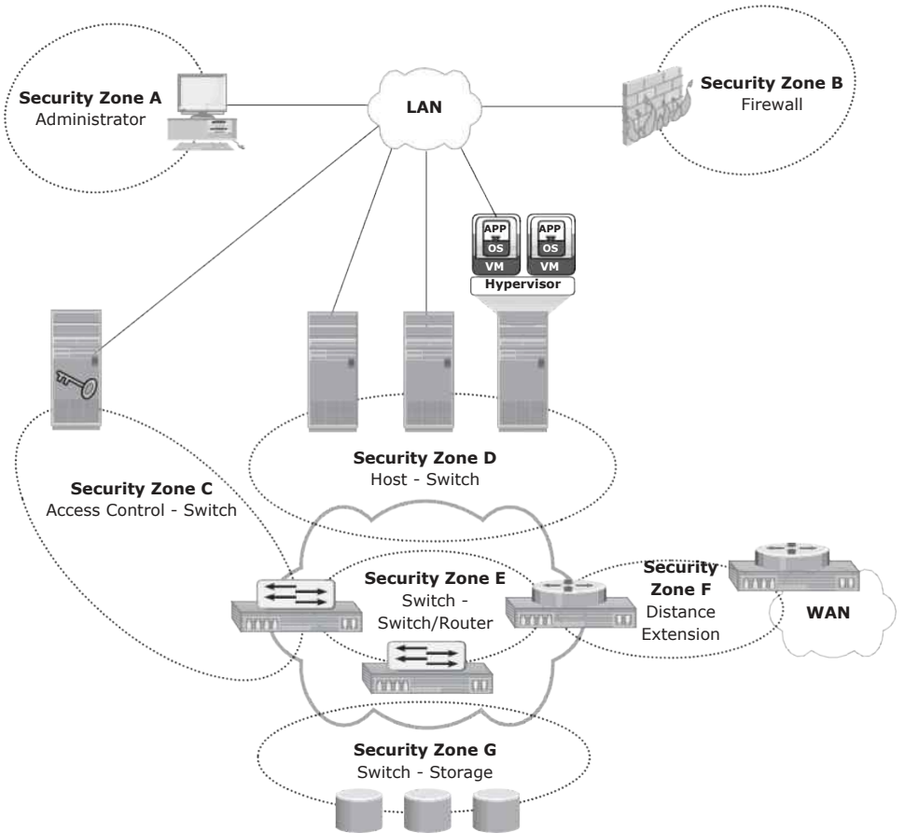
security measures, organizations must simultaneously leverage other security implementations in the enterprise. Table 14-1 provides a comprehensive list of

protection strategies that must be implemented in various security zones. Some

of the security mechanisms listed in Table 14-1 are not specific to SAN but are commonly used data center techniques. For example, two-factor authentica- tion is implemented widely; in a simple implementation it requires the use of a username/password and an additional security component such as a smart card for authentication.

Basic SAN Security Mechanisms

LUN masking and zoning, switch-wide and fabric-wide access control, RBAC, and logical partitioning of a fabric (Virtual SAN) are the most commonly used SAN security methods.



- 14-5: FC SAN security architecture

Table 14-1: Security Zones and Protection Strategies

SECURITY ZONES	PROTECTION STRATEGIES
Zone A (Authentication at the Management Console)	(a) Restrict management LAN access to authorized users (lock down MAC addresses); (b) implement VPN tunneling for secure remote access to the management LAN; and (c) use two-factor authentication for network access.
Zone B (Firewall)	Block inappropriate traffic by (a) filtering out addresses that should not be allowed on your LAN; and (b) screening for allowable protocols, block ports that are not in use.
Zone C (Access Control-Switch)	Authenticate users/administrators of FC switches using Remote Authentication Dial In User Service (RADIUS), DH-CHAP (Diffie-Hellman Challenge Handshake Authentication Protocol), and so on.

SECURITY ZONES	PROTECTION STRATEGIES
Zone D (Host to switch)	Restrict Fabric access to legitimate hosts by (a) implementing ACLs: Known HBAs can connect on specific switch ports only; and (b) implementing a secure zoning method, such as port zoning (also known as hard zoning).
Zone E (Switch to Switch/Switch to Router)	Protect traffic on fabric by (a) using E_Port authentication; (b) encrypting the traffic in transit; and (c) implementing FC switch controls and port controls.
Zone F (Distance Extension)	Implement encryption for in-flight data (a) FC-SP for long-distance FC extension; and (b) IPSec for SAN extension via FCIP.
Zone G (Switch to Storage)	Protect the storage arrays on your SAN via (a) WWPN-based LUN masking; and (b) S_ID locking: masking based on source FC address.

LUN Masking and Zoning

LUN masking and zoning are the basic SAN security mechanisms used to protect against unauthorized access to storage. LUN masking and zoning are detailed in Chapter 4 and Chapter 5, respectively. The standard implementations of LUN masking on storage arrays mask the LUNs presented to a front-end storage port based on the WWPNs of the source HBAs. A stronger variant of LUN masking may sometimes be offered whereby masking can be done on the basis of source FC addresses. It offers a mechanism to lock down the FC address of a given node port to its WWN. *WWPN zoning* is the preferred choice in security-conscious environments.

Securing Switch Ports

Apart from zoning and LUN masking, additional security mechanisms, such as port binding, port lockdown, port lockout, and persistent port disable, can be implemented on switch ports. *Port binding* limits the number of devices that can attach to a particular switch port and allows only the corresponding switch port to connect to a node for fabric access. Port binding mitigates but does not eliminate WWPN spoofing. *Port lockdown* and *port lockout* restrict a switch port's type of initialization. Typical variants of port lockout ensure that the switch port cannot function as an E_Port and cannot be used to create an ISL, such as a rogue switch. Some variants ensure that the port role is restricted to only FL_Port, F_Port, E_Port, or a combination of these. *Persistent port disable* prevents a switch port from being enabled even after a switch reboot.

Switch-Wide and Fabric-Wide Access Control

As organizations grow their SANs locally or over longer distances, there is a greater need to effectively manage SAN security. Network security can be configured on the FC switch by using *access control lists* (ACLs) and on the fabric by using fabric binding.

ACLsincorporate the device connection control and switchconnection control policies. The device connection control policy specifies which HBAs and storage ports can be a part of the fabric, preventing unauthorized devices from accessing it. Similarly, the switch connection control policy specifies which switches are allowedto be part of the fabric, preventingunauthorized switches from joining it. *Fabric binding* prevents an unauthorized switch from joining any existing switch in the fabric. It ensures that authorized membership data exists on every switch and any attempt to connect any switch in the fabric by using an ISL causes the fabric to segment.

Role-based access control provides additional security to a SAN by preventing

unauthorized activity on the fabric for management operations. It enables the security administrator to assign roles to users that explicitly specify privileges or access rights after logging into the fabric. For example, the *zone admin* role can modify the zones on the fabric, whereas a basic user may view only fabric- related information, such as port types and logged-in nodes.

Logical Partitioning of a Fabric: Virtual SAN

VSANs enable the creation of multiple logical SANs over a common physical SAN. They provide the capability to build larger consolidated fabrics and still maintain the required security and isolation between them. - 14-6 depicts logical partitioning in a VSAN.

The SAN administrator can create distinct VSANs by populating each of them withswitchports. In the example, the switchportsare distributedovertwo VSANs: 10and 20 — forthe Engineeringand HR divisions, respectively. Althoughtheyshare

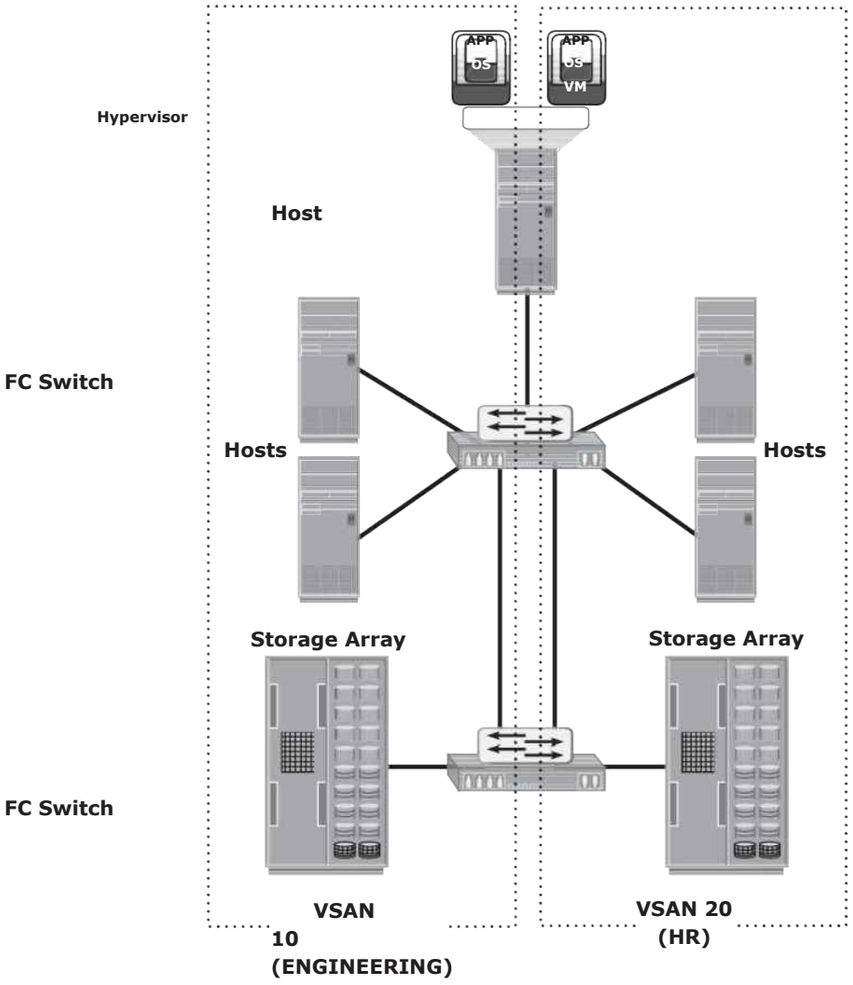
physical switching gear with other divisions, they can be managed individually as standalone fabrics. Zoning should be done for each VSAN to secure the entire

physical SAN. Each managed VSAN can have only one active zone set at a time.

VSANs minimize the impact of fabricwide disruptive events because management and control traffic on the SAN — which may include RSCNs, zone set activation events, and more — does not traverse VSAN boundaries. Therefore, VSANs are a cost-effective alternativeforbuilding isolated physical fabrics. They contribute toinformation availability and security by isolating fabric events and providing authorization control within a single fabric.

NAS

NASis open tomultipleexploits, including viruses, worms, unauthorized access, snooping, and data tampering. Various security mechanisms are implemented in NAS to secure data and the storage networking infrastructure.



- 14-6: Securing SAN with VSAN

Permissions and ACLs form the first level of protection to NAS resources by restricting accessibility and sharing. These permissions are deployed over and above the default behaviors and attributes associated with files and folders. In addition, various other authentication and authorization mechanisms, such as Kerberos and directory services, are implemented to verify the identity of net- work users and define their privileges. Similarly, firewalls protect the storage infrastructure from unauthorized access and malicious attacks.

NAS File Sharing: Windows ACLs

Windows supports two types of ACLs: *discretionary access control lists* (DACLs) and
system access control lists (SACLs). The DACL, commonly referred to as the

ACL, that determines access control. The SACL determines what accesses need to be audited if auditing is enabled.

In addition to these ACLs, Windows also supports the concept of object own- ership. The owner of an object has hard-coded rights to that object, and these rights do not need to be explicitly granted in the SACL. The owner, SACL, and DACL are all statically held as attributes of each object. Windows also offers the functionality to inherit permissions, which allows the child objects existing within a parent object to automatically inherit the ACLs of the parent object.

ACLs are also applied to directory objects known as security identifiers (SIDs). These are automatically generated by a Windows server or domain when a user or group is created, and they are abstracted from the user. In this way, though a user may identify his login ID as —User1,‖ it is simply a textual representation of the true SID, which is used by the underlying operating sys- tem. Internal processes in Windows refer to an account's SID rather than the account's username or group name while granting access to an object. ACLs are set by using the standard Windows Explorer GUI but can also be configured with CLI commands or other third-party tools.

NAS File Sharing: UNIX Permissions

For the UNIX operating system, a *user* is an abstraction that denotes a logical entity for assignment of ownership and operation privileges for the system.A user can be either a person or a system operation. A UNIX system is only aware of the privileges of the user to perform specific operations on the system and identifies each user by a user ID (UID) and a username, regardless of whether it is a person, a system operation, or a device.

In UNIX, users can be organized into one or more groups. The concept of group serves the purpose to assign sets of privileges for a given resource and sharing them among many users that need them. For example, a group of people working on one project may need the same permissions for a set of files.

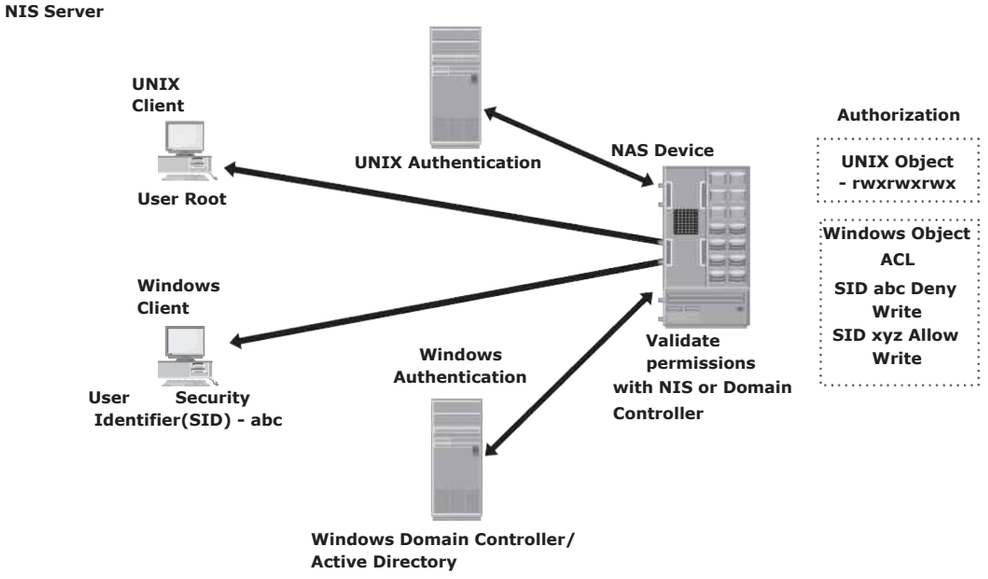
UNIX permissions specify the operations that can be performed by any own- ership relation with respect to a file. In simpler terms, these permissions specify whatthe owner can do, whatthe ownergroup can do, and whateveryone else can do with the file. For any given ownership relation, three bits are used to specify access permissions. The first bit denotes read (r) access, the second bit denotes write (w) access, and the third bit denotes execute (x) access. Because UNIX defines three ownership relations (Owner, Group, and All), a triplet (defining the access permission) is required for each ownership relationship, resulting in nine bits. Each bit can be either set or clear. When displayed, a set bit is marked by its corresponding operation letter (r, w, or x), a clear bit is denoted by a dash (-), and all are put in a row, such as rwxr-xr-x. In this example, the owner can do anything with the file, but group owners and the rest of the world can read or execute only. When displayed, a character denoting the mode of the file may

precede this nine-bit pattern. For example, if the file is a directory, it is denoted as —dll; and if it is a link, it is denoted as —l.ll

NAS File Sharing: Authentication and Authorization

In a file-sharing environment, NAS devices use standard file-sharing protocols, NFS and CIFS. Therefore, authentication and authorization are implemented and supported on NAS devices in the same way as in a UNIX or Windows file- sharing environment.

Authentication requires verifying the identity of a network user and therefore involves a login credential lookup on a Network Information System (NIS) server in a UNIX environment. Similarly, a Windows client is authenticated by a Windows domain controller that houses the Active Directory. The Active Directory uses LDAP to access information about network objects in the direc- tory and Kerberos for network security. NAS devices use the same authentica- tion techniques to validate network user credentials. - 14-7 depicts the authentication process in a NAS environment.



- 14-7: Securing user access in a NAS environment

Authorization defines user privileges in a network. The authorization techniques for UNIX users and Windows users are quite different. UNIX files use mode bits to define access rights granted to owners, groups, and other users, whereas Windows uses an ACL to allow or deny specific rights to a particular user for a particular file.

Although NAS devices support both of these methodologies for UNIX and Windows users, complexities arise when UNIX and Windows users access and share the same data. If the NAS device supports multiple protocols, the integrity of both permission methodologies must be maintained. NAS device vendors provide a method of mapping UNIX permissions to Windows and vice versa, so a multiprotocol environment can be supported. However, consider these com- plexities of multiprotocol support when designing a NAS solution. At the same time, validate the domain controller and NIS server connectivity and bandwidth. If multiprotocol access is required, specific vendor access policy implementations need to be considered.

Kerberos

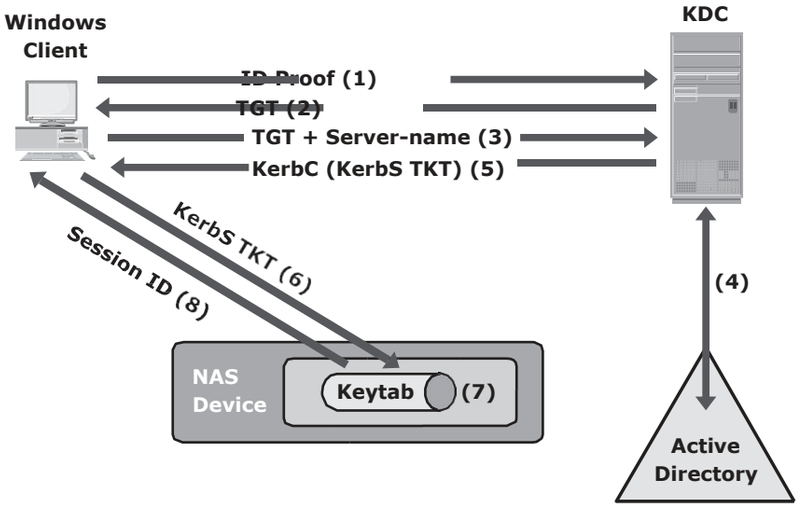
Kerberosisanetworkauthenticationprotocol, which isdesignedtoprovidestrong authentication for client/server applications by using secret-key cryptography. It uses cryptography so that a client and server can prove their identity to each other across an insecure network connection. After the client and server have proven their identities, they can choose to encrypt all their communications to ensure privacy and data integrity.

In Kerberos, authentications occur between clients and servers. The client gets a ticket for a service and the server decrypts this ticket by using its secret key. Any entity, user, or host that gets a service ticket for a Kerberos service is called a *Kerberos client*. The term *Kerberos server* generally refers to the Key Distribution Center(KDC). The KDC implementsthe Authentication Service (AS) and the Ticket Granting Service (TGS). The KDC has a copy of every password associated with every principal, so it is absolutely vital that the KDC remain secure. In Kerberos, users and servers for which a secret key is stored in the KDC database are known as *principals*.

In a NAS environment, Kerberos is primarily used when authenticating against a Microsoft Active Directory domain, although it can be used to execute security functions in UNIX environments. The Kerberos authentication process shown in - 14-8 includes the following steps:

1. The user logs on to the workstation in the Active Directory domain (or forest) using an ID and a password. The client computer sends a request to the AS running on the KDC for a Kerberos ticket. The KDC verifies the user's login information from Active Directory. (This step is not explicitly shown in - 14-8.)
2. The KDC responds with an encrypted Ticket Granting Ticket (TGT) and an encryptedsessionkey. TGThasalimitedvalidityperiod. TGT can bedecrypted only by the KDC, and the client can decrypt only the session key.
3. When the client requests a service from a server, it sends a request, consisting of the previously generated TGT, encrypted with the session key and the resource information to the KDC.

4. The KDC checks the permissions in Active Directory and ensures that the user is authorized to use that service.
5. The KDC returns a service ticket to the client. This service ticket contains
   fields addressed to the client and to the server hosting the service.
6. Theclient thensendstheservice ticket tothe serverthat houses therequired resources.
7. The server, in this case the NAS device, decrypts the server portion of the ticket and stores the information in a keytab file. As long as the client's Kerberos ticket is valid, this authorization process does not need to be repeated. The server automatically allows the client to access the appro- priate resources.
8. A client-server session is now established. The server returns a session ID to the client, which tracks the client activity, such as file locking, as long as the session is active.



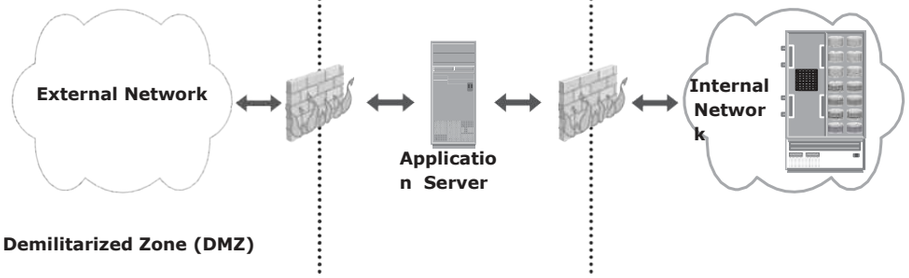
- 14-8: Kerberos authorization

Network-Layer Firewalls

Because NAS devices utilize the IP protocol stack, they are vulnerable to vari- ous attacks initiated through the public IP network. Network layer firewalls are implemented in NAS environments to protect the NAS devices from these security threats. These network-layer firewalls can examine network packets and compare them to a set of configured security rules. Packets that are not authorized by a security rule are dropped and not allowed to continue to the destination. Rules can be established based on a source address (network or host), a destination address (network or host), a port, or a combination of those

factors(source IP, destination IP, and port number). The effectiveness of a firewall depends on how robust and extensive the security rules are. A loosely defined rule set can increase the probability of a security breach.

- 14-9 depicts a typical firewall implementation. A demilitarized zone (DMZ) is commonly used in networking environments. A DMZ provides a means to secure internal assets while allowing Internet-based access to various resources. In a DMZ environment, servers that need to be accessed through the Internet are placedbetween two sets of firewalls. Application-specific ports, such as HTTP or FTP, are allowed through the firewall to the DMZ servers. However, no Internet-based traffic is allowed to penetrate the second set of firewalls and gain access to the internal network.

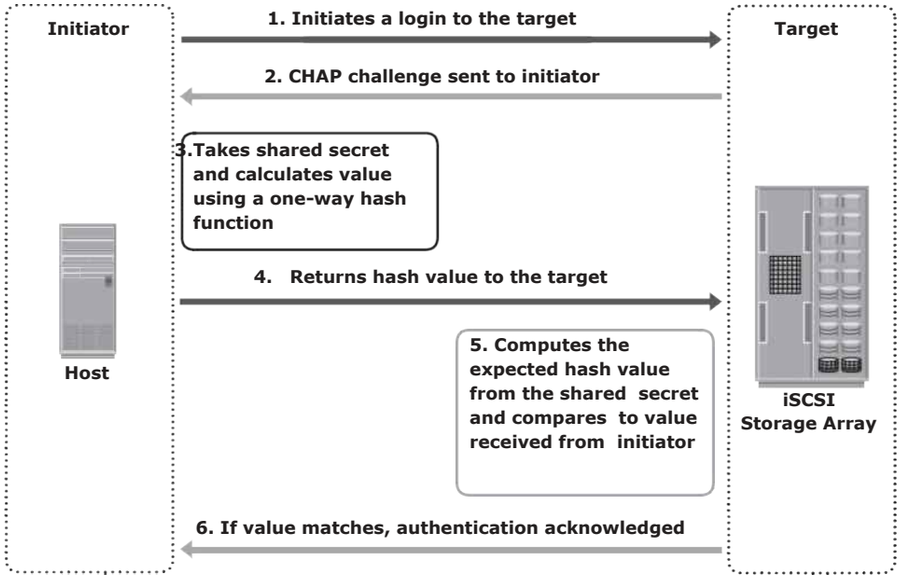


- 14-9: Securing a NAS environment with a network-layer firewall

The servers in the DMZ may or may not be allowed to communicate with internal resources. In such a setup, the server in the DMZ is an Internet-facing web application accessing data stored on a NAS device, which may be located on the internal private network. A secure design would serve only data to internal and external applications through the DMZ.

IP SAN

This section describes some of the basic security mechanisms used in IP SAN environments. The *Challenge-Handshake Authentication Protocol* (CHAP) is a basic authenticationmechanism that has been widely adopted by network devices and hosts. CHAP provides a method for initiators and targets to authenticate each other by utilizing a secret code or password. CHAP secrets are usually random secrets of 12 to 128 characters. The secret is never exchanged directly over the communication channel; rather, a one-way hash function converts it into a hash value, which is then exchanged. A hash function, using the MD5 algorithm, transforms data in such a way that the result is unique and cannot be changed back to its original form. - 14-10 depicts the CHAP authentication process.



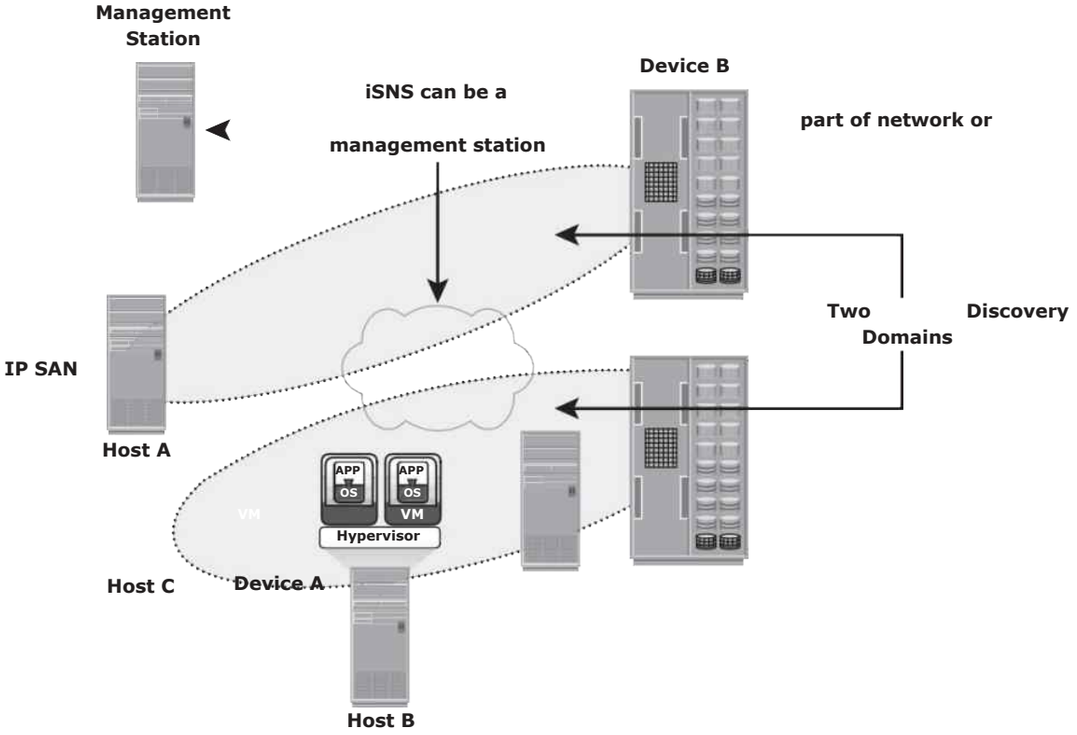
- 14-10: Securing IPSAN with CHAP authentication

If the initiator requires reverse CHAP authentication, the initiator authenticates the target by using the same procedure. The CHAP secret must be configured on the initiator and the target. A CHAP entry, composed of the name of a node and the secret associated with the node, is maintained by the target and the initiator. The same steps are executed in a two-way CHAP authentication scenario.

After these steps are completed, the initiator authenticates the target. If both authentication steps succeed, then data access is allowed. CHAP is often used because it is a fairly simple protocol to implement and can be implemented across a number of disparate systems.

iSNS discovery domains function in the same way as FC zones. Discovery domains provide functional groupings of devices in an IP-SAN. For devices to

communicate with one another, they must be configured in the same discovery domain. State change notifications (SCNs) inform the iSNS server when devices are added to or removed from a discovery domain. - 14-11 depicts the discovery domains in iSNS.



- 14-11: Securing IPSAN with iSNS discovery domains

Securing Storage Infrastructure in Virtualized and Cloud Environments

This chapter, so far, focused only on the security threats and measures in a traditional data center. These threats and measures are also applicable to infor- mation storage in virtualized and cloud environments. However, virtualized and cloud computing environments pose additional threats to an organiza- tion's data due to multitenancy and lack of control over the cloud resources. A public cloud has more security concerns compared to a private cloud and demands additional counter measures. This is because in a public cloud, cloud users (consumers) usually have limited control over resources, and therefore, enforcement of security mechanisms by consumers is comparatively difficult.

From a security perspective, both consumers and cloud service providers (CSP) have several security concerns and face multiple threats. Security concerns and security measures are detailed next.

Security Concerns

Organizations are rapidlyadoptingvirtualization and cloudcomputing, however they have some security concerns. These key security concerns are multitenancy, velocity of attack, information assurance, and data privacy.

Multitenancy, by virtue of virtualization, enables multiple independent tenants to be serviced using the same set of storage resources. In spite of the benefits offered by multitenancy, it is still a key security concern for users and service providers. Colocation of multiple VMs in a single server and sharing the same resources increase the attack surface. It may happen that business critical data of one tenant is accessed by other competing tenants who run applications using the same resources.

Velocity-of-attack refers to a situation in which any existing security threat in the cloud spreads more rapidly and has a larger impact than that in the traditional data center environments. *Information assurance* for users ensures confidentiality, integrity, and availability of data in the cloud. Also the cloud user needs assurance that all the users operating on the cloud are genuine and access the data only with legitimate rights and scope.

Data privacy is also a major concern in a virtualized and cloud environment. A CSP needs to ensure that Personally Identifiable Information (PII) about its clients is legally protected from any unauthorized disclosure.

Security Measures

Security measures can be implemented at the compute, network, and storage levels. These security measures implemented at three layers mitigate the risks in virtualized and cloud environments.

Security at the Compute Level

Securing a compute infrastructure includes enforcing the security of the physical server, hypervisor, VM, and guest OS (OS running within a virtual machine). *Physical server security* involves implementing user authentication and authoriza- tionmechanisms. These mechanisms identify users and provideaccessprivileges on the server. To minimize the attack surface on the server, unused hardware components, such as NICs, USB ports, or drives, should be removed or disabled. A *hypervisor* is a single point of security failure for all the VMs running on it. Rootkits and malware installed on a hypervisor make detection difficult for the antivirus software installed on the guest OS. To protect against attacks,

security-critical hypervisor updates should be installed regularly. Further, the hypervisor management system must also be protected. Malicious attacks and infiltration to the management system can impact all the existing VMs and allow attackers to create new VMs. Access to the management system should be restricted to authorized administrators. Furthermore, there must be a separate firewall installed between the management system and the rest of the network. *VM isolation* and *hardening* are some of the common security mechanisms to effectively safeguard a VM from an attack. VM isolation helps to prevent a com- promised guest OS from impacting other guest OSs. VMisolationis implemented at the hypervisor level. Apart from isolation, VMs should be hardened against security threats. Hardening is a process to change the default configuration to achieve greater security.

Apart from the measures to secure a hypervisor and VMs, virtualized and cloud environments also require further measures on the guest OS and application levels.

Security at the Network Level

The key security measures that minimize vulnerabilities at the network layer are firewall, intrusion detection, demilitarized zone (DMZ), and encryption of data-in-flight.

A *firewall* protects networks from unauthorized access while permitting only legitimate communications. In a virtualized and cloud environment, a firewall can also protect hypervisors and VMs. For example, if remote administration is enabled on a hypervisor, access to all the remote administration interfaces should be restricted by a firewall. A firewall also secures VM-to-VM traffic. This firewall service can be provided using a *Virtual Firewall* (VF). A VF is a firewall service running entirely on the hypervisor. A VF provides packet filtering and monitoring of the VM-to-VM traffic. A VF gives visibility and control over the VM traffic and enforces policies at the VM level.

Intrusion Detection (ID) is the process to detect events that can compromise the confidentiality, integrity, or availability of a resource. An ID System (IDS)

automatically analyzes events to check whether an event or a sequence of events match a known pattern for anomalous activity, or whether it is (statistically) different from most of the other events in the system. It generates an alert if an irregularity is detected. DMZ and data encryption are also deployed as security measures in the virtualized and cloud environments. However, these deployments work in the same way as in the traditional data center.

Security at the Storage Level

Major threats to storage systems in virtualized and cloud environments arise due to compromises at compute, network, and physical security levels. This is because access to storage systems is through compute and network infrastruc- ture. Therefore, adequate security measures should be in place at the compute and network levels to ensure storage security. Common security mechanisms that protect storage include the following:

- Access control methods to regulate which users and processes access the data on the storage systems
- Zoning and LUN masking
- Encryption of data-at-rest (on the storage system) and data-in-transit. Data encryption should also include encrypting backups and storing encryption keys separately from the data.
- Data shredding that removes the traces of the deleted data

Apart from these mechanisms, isolation of different types of traffic using VSANs further enhances the security of storage systems. In the case of storage utilized by hypervisors, additional security steps are required to protect the storage. Storage for hypervisors using clustered file systems supporting multiple VMs may require separate LUNs for VM components and VM data.

Concepts in Practice: RSA and VMware Security Products

RSA, the security division of EMC, is the premier provider of security, risk, and compliance solutions, helping organizations to solve their most complex and sensitive security challenges.

VMware offers secure and robust virtualization solutions for virtualized and cloud environments. This section provides a brief introduction to RSA SecureID, RSA Identity and Access Management, RSA Data Protection Manager, and VMware vShield.

RSA SecureID

RSA SecurID two-factor authentication provides an added layer of security to ensure that only valid users have access to systems and data. RSA SecurID is based on something a user knows (a password or PIN) and something a user has (an authenticator device). It provides a much more reliable level of user authentication than reusable passwords. It generates a new one-time password code every 60 seconds, making it difficult for anyone other than the genuine user to input the correct token code at any given time. To access their resources, users combine their secret Personal Identification Number (PIN) with the token code that appears on their SecurID authenticator display at that given time. The result is a unique, one-time password to assure a user's identity.

RSA Identity and Access Management

The RSA Identity and Access Management product provides identity, security, and access-controls management for physical, virtual, and cloud-based environments through access management. It enables trusted identities to freely and securely interact with systems and access. The RSA Identity and Access Management family has two products: *RSA Access Manager* and *RSA Federated Identity Manager*. RSA Access Manager enables organizations to centrally manage authentication and authorization policies for a large number of users, online web portals, and application resources. Access Manager provides seamless user access with single sign-on (SSO) and preserves identity context for greater security. RSA Federated Identity Manager enables end users to collaborate with business partners, outsourced service providers, and supply-chain partners or across multiple offices or agencies all with a single identity and logon.

RSA Data Protection Manager

RSA Data Protection Manager enables deployment of encryption, tokenization, and enterprise key management simply and affordably. The RSA Data Protection Manager family is composed of two products: *Application Encryption and Tokenization* and *Enterprise Key Management*.

- Application Encryption and Tokenization with RSA Data Protection Manager helps to achieve compliance with regulations related to PII by quickly embedding the encryption and tokenization of sensitive data and helping to prevent data loss. It works at the point of creation, ensuring that the data stays encrypted as it is transmitted and stored.
- Enterprise keymanagementis an easy-to-use managementtoolfor encrypting keys at the database, file server, and storage layers. It is designed to simplify the deployment of encryption throughout the enterprise. It also helps to ensure that information is properly secured and fully accessible when needed at any point in its life cycle.

VMware vShield

The VMware vShield family includes three products: *vShield App*, *vShield Edge*, and
vShield Endpoint.

VMware vShield App is a hypervisor-based application-aware firewall solution. It protects applications in a virtualized environment from network-based threats by providing visibility into network communications and enforcing granular policies with security groups. VMware vShield App observes network activity between virtual machines to define and refine firewall policies and secure business processes through detailed reporting of application traffic.

VMware vShield Edge provides comprehensive perimeter network security for a virtualized environment. It is deployed as a virtual appliance and serves as a network security gateway for all the hosts within the virtualized environment. It provides many services including firewall, VPN, and Dynamic Host Configuration Protocol (DHCP) services.

VMware vShield Endpoint consists of a hardened special security VM with a third party antivirus software. VMware vShield Endpoint streamlines and accelerates antivirus and antimalware deployment because antivirus engine and signature files are updated only within the special security VM. VMware vShield Endpoint improves VM performance by offloading file scanning and other tasks from VMs to the security VM. It prevents antivirus storms and bottlenecks associated with multiple simultaneous antivirus and antimalware scans and updates. It also satisfies audit requirements with detailed logging of antivirus and antimalware activities.

Monitoring the Storage Infrastructure

Monitoringisone of the mostimportant aspects that formsthe basis formanaging storage infrastructure resources. Monitoring provides the performance and accessibility status of various components. It also enables administrators to perform essential management activities. Monitoring also helps to analyze the utilization and consumption of various storage infrastructure resources. This analysis facilitates capacity planning, forecasting, and optimal use of these resources. Storage infrastructure environment parameters such as heating and power supplies are also monitored.

Monitoring Parameters

Storageinfrastructure componentsshould be monitoredforaccessibility, capacity, performance, and security. *Accessibility* refers to the availability of a component to perform its desired operation during a specified time period. Monitoring the accessibility of hardware components (for example, a port, an HBA, or a disk drive) or software component (for example, a database) involves checking their availability status by reviewing the alerts generated from the system. For example, a port failure might result in a chain of availability alerts.

A storage infrastructure uses redundant components to avoid a single point

of failure. Failure of a component might cause an outage that affects application availability, or it might cause performance degradation even though accessibility is not compromised. Continuously monitoring for expected accessibility of each component and reporting any deviation helps the administrator to identify failing components and plan corrective action to maintain SLA requirements.

Capacity refers to the amount of storage infrastructure resources available. Examples of capacity monitoring include examining the free space available on a file system or a RAID group, the mailbox quota allocated to users, or the numbers of ports available on a switch. Inadequate capacity leads to degraded performance or even application/service unavailability. *Capacity monitoring* ensures uninterrupted dataavailability and scalability by avertingoutagesbefore they occur. For example, if 90 percent of the ports are utilized in a particular

SAN fabric, this could indicate that a new switch might be required if more arrays and servers need to be installed on the same fabric. Capacity monitoring usually leverages analytical tools to perform trend analysis. These trends help to understand future resource requirements and provide an estimation on the time line to deploy them.

Performance monitoring evaluates how efficiently different storage infrastructure components are performing and helps to identify bottlenecks. Performance monitoring measures and analyzes behavior in terms of response time or the ability to perform at a certain predefined level. It also deals with the utilization of resources, which affects the way resources behave and respond. Performance measurement is a complex task that involves assessing various components on several interrelated parameters. The number of I/Os performed by a disk, application response time, network utilization, and server-CPU utilization are examples of performance parameters that are monitored.

Monitoring a storage infrastructure for security helps to track and prevent unauthorized access, whether accidental or malicious. *Security monitoring* helps to track unauthorized configuration changes tostorageinfrastructure resources. For example, security monitoring tracks and reports the initial zoning configuration performed and all the subsequent changes. Security monitoring also detects unavailability of information to authorized users due to a security breach. Physical security of a storage infrastructure can also be continuously monitored using badge readers, biometric scans, or video cameras.

Components Monitored

Hosts, networks, and storage are the componentswithin the storage environment that should be monitored for accessibility, capacity, performance, and security. These components can be physical or virtualized.

Hosts

The accessibility of a host depends on the availability status of the hardware components and the software processes running on it. For example, a host's NIC failure might cause inaccessibility of the host to its user. Server clustering is a mechanism that provides high availability if a server failure occurs.

Monitoring the file system capacity utilization of a host is important to ensure that sufficient storage capacity is available to the applications. Running out of file system space disrupts applicationavailability. Monitoring helps estimate the file system's growth rate and predict when it will reach 100 percent. Accordingly, the administrator can extend (manually or automatically) the file system's space proactivelytopreventapplicationoutage. Use of virtual provisioning technology

enables efficient management of storage capacity requirements but is highly dependent on capacity monitoring.

Host performance monitoring mainly involvesastatus check on the utilization of various server resources, such as CPU and memory. For example, if a server running an application is experiencing 80 percent of CPU utilization continu- ously, it indicates that the server may be running out of processing power, which can lead to degraded performance and slower response time. Administrators can take several actions to correct the problem, such as upgrading or adding more processors and shifting the workload to different servers. In a virtualized environment, additional CPU and memory may be allocated to VMs dynami- cally from the pool, if available, to meet performance requirements.

Security monitoring on servers involves tracking of login failures and execution of unauthorized applications or software processes. Proactive measures against unauthorized access to the servers are based on the threat identified. For example, an administrator can block user access if multiple login failures are logged.

Storage Network

Storage networks need to be monitored to ensure uninterrupted communication between the server and the storage array. Uninterrupted access to data over the storage network depends on the accessibility of the physical and logical com- ponents of the storage network. The physical components of a storage network include switches, ports, and cables. The logical components include constructs, such as zones. Any failure in the physical or logical components causes data unavailability. For example, errors in zoning, such as specifying the wrong WWN of a port, result in failure to access that port, which potentially prevents access from a host to its storage.

Capacity monitoring in a storage network involves monitoring the number of available ports in the fabric, the utilization of the interswitch links, or individual ports, and each interconnect device in the fabric. Capacity monitoring provides all the required inputs for future planning and optimization of fabric resources.

Monitoring the performance of the storage network enables assessing individual component performance and helps to identify network bottlenecks. For example, monitoring port performance involves measuring the receive or transmit link utilization metrics, which indicates how busy the switch port is. Heavily used ports can cause queuing of I/Os on the server, which results in poor performance.

For IP networks, monitoring the performance includes monitoring network latency, packet loss, bandwidth utilization for I/O, network errors, packet retransmission rates, and collisions.

Storage network security monitoring provides information about any unauthorized change to the configuration of the fabric — for example, changes to the zone policies that can affect datasecurity. Login failures and unauthorized access to switches for performing administrative changes should be logged and monitored continuously.

Storage

The accessibility of the storage array should be monitored for its hardware components and various processes. Storage arrays are typically configured with redundant components, and therefore individual component failure does not usually affect their accessibility. However, failure of any process in the stor- age array might disrupt or compromise business operations. For example, the failure of a replication task affects disaster recovery capabilities. Some storage arrays provide the capability to send messages to the vendor's support center if hardware or process failures occur, referred to as a *call home*.

Capacity monitoring of a storage array enables the administrator to respond to storage needs preemptively based on capacity utilization and consumption trends. Information about unconfigured and unallocated storage space enables the administrator to decide whether a new server can be allocated storage capac- ity from the storage array.

Astoragearray can bemonitored using a number of performance metrics, such as utilization rates of the various storage array components, I/O response time, and cache utilization. For example, an over utilized storage array component might lead to performance degradation.

A storage array is usually ashared resource, which maybe exposed to security

threats. Monitoring security helps to track unauthorized configuration of the storage array and ensures that only authorized users are allowed to access it.

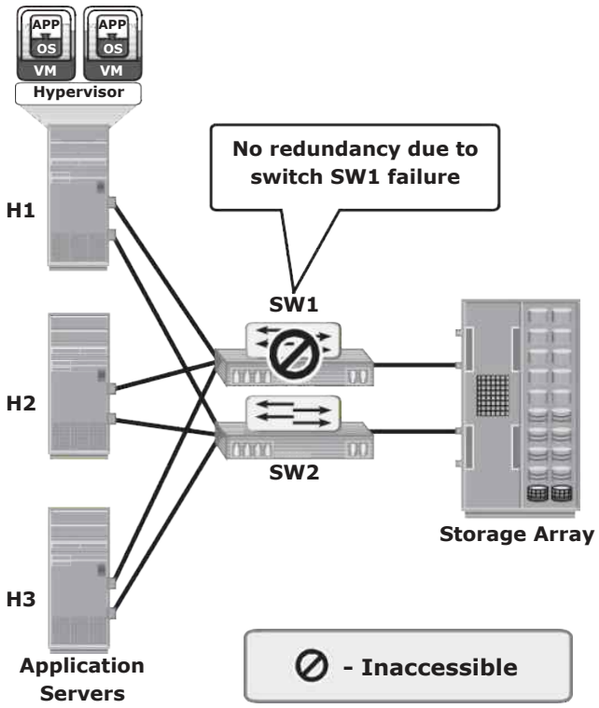
Monitoring Examples

A storage infrastructure requires implementation of an end-to-end solution to actively monitor all the parameters of its components. Early detection and preemptive alerting ensure uninterrupted services from critical assets. In addition, the monitoring tool should analyze the impact of a failure and deduce the root cause of symptoms.

Accessibility Monitoring

Failure of anycomponentmight affect the accessibility of one or morecomponents due to their interconnections and dependencies. Consider an implementation in a storage infrastructure with three servers: H1, H2, and H3. All the servers

are configured with two HBAs, each connected to the production storage array through two switches, SW1 and SW2, as shown in - 15-1. All the servers share two storage ports on the storage array and multipathing software is installed on all the servers.



- 15-1: Switch failure in a storage infrastructure

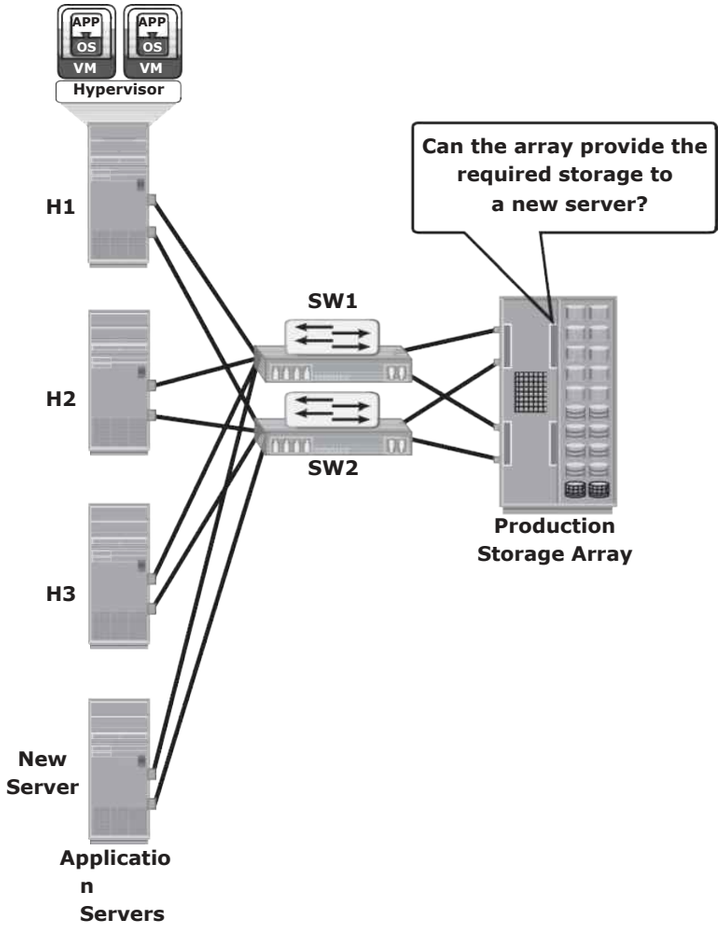
If one of the switches (SW1) fails, the multipathing software initiates a path failover, and all the servers continue to access data through the other switch, SW2. However, due to the absence of a redundant switch, a second switch failure could result in inaccessibility of the array. Monitoring for accessibility enables detecting the switch failure and helps an administrator to take corrective action before another failure occurs.

In most cases, the administrator receives symptom alerts for a failing component and can initiate actions before the component fails.

Capacity Monitoring

In the scenario shown in - 15-2, servers H1, H2, and H3 are connected to the production array through two switches, SW1 and SW2. Each of the servers

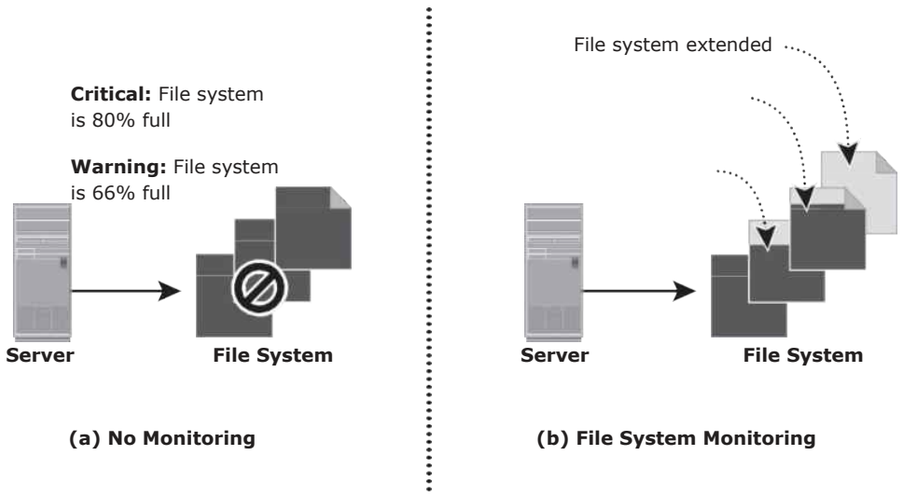
is allocated storage on the storage array. When a new server is deployed in this configuration, the applications on the new server need to be given storage capacity from the production storage array. Monitoring the available capacity (configurable and unallocated) on the array helps to proactively decide whether the array can provide the required storage to the new server. Also, monitoring the available number of ports on SW1 and SW2 helps to decide whether the new server can be connected to the switches.



- 15-2: Monitoring storage array capacity

The following example illustrates the importance of monitoring the file system capacity on file servers. - 15-3 (a) illustrates the environment of a file system when full and that results in application outage when no capacity

monitoring is implemented. Monitoring can be configured to issue a message when thresholds are reached on the file system capacity. For example, when the file system reaches 66 percent of its capacity, a warning message is issued, and a critical message is issued when the file system reaches 80 percent of its capacity (see - 15-3 [b]). This enables the administrator to take action to extend the file system before it runs out of capacity. Proactively monitor- ing the file system can prevent application outages caused due to lack of file system space.



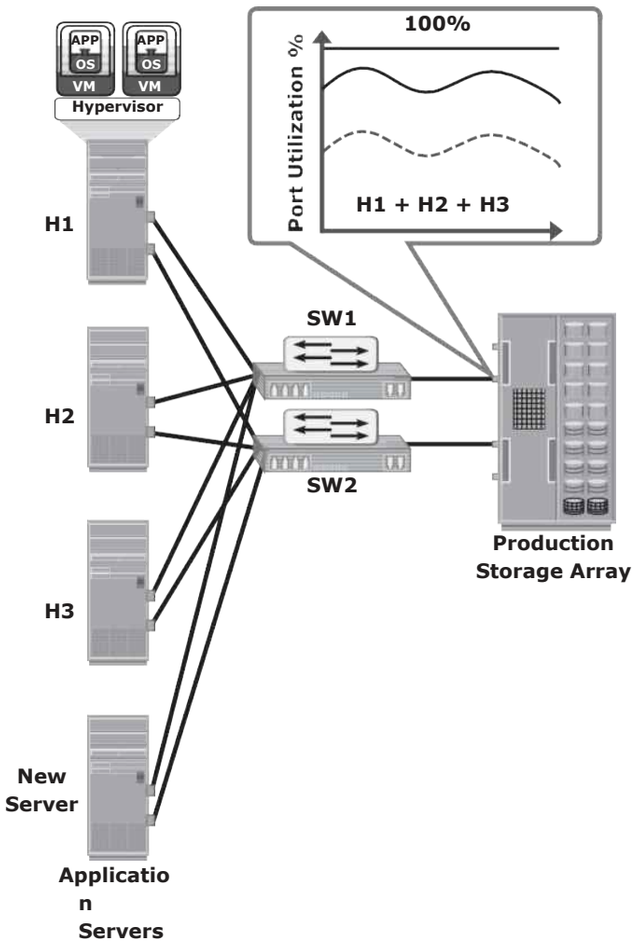
- 15-3: Monitoring server file system space

Performance Monitoring

The example shown in - 15-4 illustrates the importance of monitoring performance on storage arrays. In this example, servers H1, H2, and H3 (with two HBAs each) are connected to the storage array through switch SW1 and SW2. The three servers share the same storage ports on the storage array to access LUNs. A new server running an application with a high work load must be deployed to share the same storage port as H1, H2, and H3.

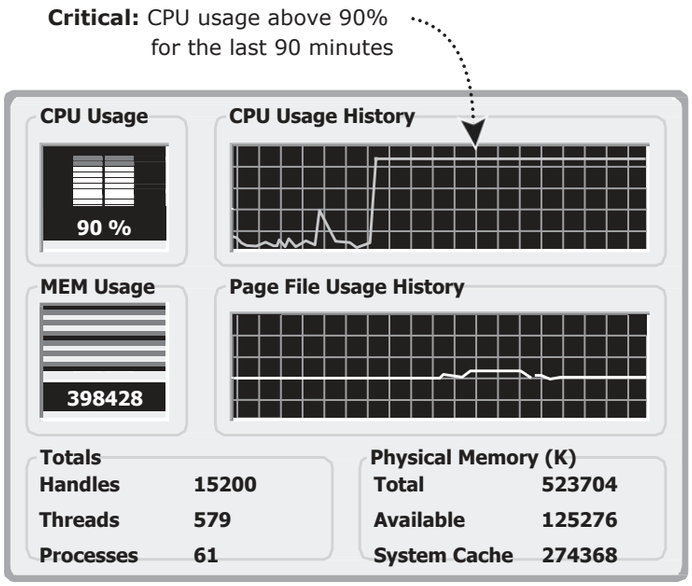
Monitoringarrayport utilizationensuresthat thenewserverdoesnotadversely affect the performance of the other servers. In this example, utilization of the shared storage port is shown by the solid and dotted lines in the graph. If the port utilization prior to deploying the new server is close to 100 percent, then deploying the new server is not recommended because it might impact the

performance of the other servers. However, if the utilization of the port prior to deploying the new server is closer to the dotted line, then there is room to add a new server.



- 15-4: Monitoring array port utilization

Most servers offer tools that enable monitoring of server CPU usage. For example, Windows Task Manager displays CPU and memory usage, as shown in - 15-5. However, these tools are inefficient at monitoring hundreds of servers running in a data-center environment. A data-center environment requires intelligent performance monitoring tools that are capable of monitoring many servers simultaneously.



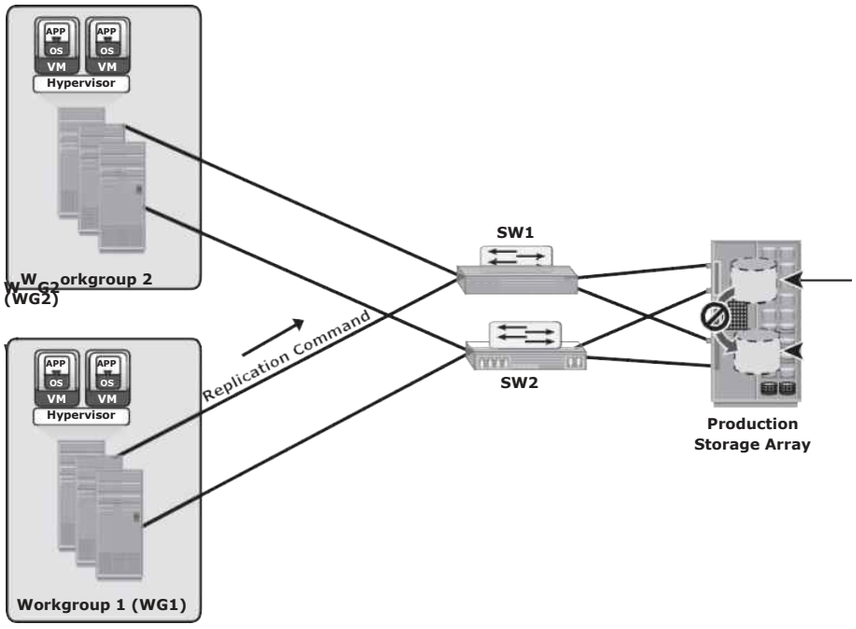
- 15-5: Monitoring the CPU and memory usage of a server

Security Monitoring

The example shown in - 15-6 illustrates the importance of monitoring security in a storage array.

In this example, the storage array is shared between two workgroups, WG1 and WG2. The data of WG1 should not be accessible to WG2 and vice versa. A user from WG1 might try to make a local replica of the data that belongs to WG2. If this action is not monitored or recorded, it is difficult to track such a violation of information security. Conversely, if this action is monitored, a warning message can be sent to prompt a corrective action or at least enable discovery as part of regular auditing operations.

An example of host security monitoring is tracking of login attempts at the host. The login is authorized if the login ID and password entered are correct; or the login attempt fails. These login failures might be accidental (mistyping) or a deliberate attempt to access a server. Many servers usually allow a fixed number of successive login failures, prohibiting any additional attempts after these login failures. In a monitored environment, the login information is recorded in a system log file, and three successive login failures trigger a message, warning of a possible security threat.



- 15-6: Monitoring security in a storage array

Alerts

Alerting of events is an integral part of monitoring. Alerting keeps administrators informed about the status of various components and processes — for example, conditions such as failure of power, disks, memory, or switches, which can impact the availability of services and require immediate administrative attention. Other conditions, such as a file system reaching a capacity threshold or a soft media error on disks, are considered warning signs and may also require administrative attention.

Monitoringtoolsenableadministrators to assign different severitylevelsbased on the impact of the alerted condition. Whenever a condition with a particular severity level occurs, an alert is sent to the administrator, a script is triggered, or an incident ticket is opened to initiate a corrective action. Alert classifications can range from information alerts to fatal alerts. *Information alerts* provide use- ful information but do not require any intervention by the administrator. The creation of a zone or LUN is an example of an information alert. *Warning alerts* require administrative attention so that the alerted condition is contained and

does not affect accessibility. For example, if an alert indicates that the number of soft media errors on a disk is approaching a predefined threshold value, the administrator can decide whether the disk needs to be replaced. *Fatal alerts* require immediate attention because the condition might affect overall perfor- mance, security, or availability. For example, if a disk fails, the administrator must ensure that it is replaced quickly.

Continuous monitoring, with automated alerting, enables administrators to respond to failures quickly and proactively. Alerting provides information that helps administrators prioritize their response to events.

Storage Infrastructure Management Activities

The pace of information growth, proliferation of applications, heterogeneous infrastructure,andstringentservice-levelrequirementshaveresultedinincreased complexity of managing storage infrastructures. However, the emergence of storage virtualization and other technologies, such as data deduplication and compression, virtual provisioning, federated storage access, and storage tiering, have enabled administrators to efficiently manage storage resources.

Thekeystorageinfrastructure managementactivitiesperformedinadatacenter can be broadly categorized into availability management, capacity management, performance management, security management, and reporting.

Availability Management

A critical task in availability management is establishing a proper guideline based on defined service levels to ensure availability. *Availability management* involves all availability-related issues for components or services to ensure that service levels are met. A key activity in availability management is to provision redundancy at all levels, including components, data, or even sites. For example, when a server is deployed to support a critical business function, it requires high availability. This is generally accomplished by deploying two or more HBAs, multipathing software, and server clustering. The server must be connected to the storage array using at least two independent fabrics and switches that have built-in redundancy. In addition, the storage arrays should have built-in redundancy for various components and should support local and remote replication.

Capacity Management

The goal of *capacity management* is to ensure adequate availability of resources based on their service level requirements. Capacity management also involves optimizationofcapacitybasedonthecostandfutureneeds.Capacitymanagement

provides capacity analysis that compares allocated storage to forecasted storage on a regular basis. It also provides trend analysis based on the rate of consump- tion, which must be rationalized against storage acquisition and deployment timetables. Storage provisioning is an example of capacity management. It involves activities, such as creating RAID sets and LUNs, and allocating them to the host. Enforcing capacity quotas for users is another example ofcapacity management. Provisioning a fixed amount of user quotas restricts users from exceeding the allocated capacity.

Technologies, such as data deduplication and compression, have reduced the amount of data to be backed up and thereby reduced the amount of storage
capacity to be managed.

Performance Management

Performance management ensures the optimal operational efficiency of all components. Performance analysis is an important activity that helpsto identify the performance of storage infrastructure components. This analysis provides information on whether a component meets expected performance levels.

Several performance management activities need to be performed when deploying a new application or server in the existing storage infrastructure. Every component must be validated for adequate performance capabilities as defined by the service levels. Forexample, to optimize the expected performance levels, activities on the server, such as the volume configuration, database design
or application layout, configuration of multiple HBAs, and intelligent multipa- thing software, must be fine-tuned. The performance management tasks on a SAN include designing and implementing sufficient ISLs in a multiswitch fabric with adequate bandwidth to support the required performance levels. The storage array configuration tasks include selecting the appropriate RAID type, LUN layout, front-end ports, back-end ports, and cache configuration, when considering the end-to-end performance.

Security Management

The key objective of the *security management* activity is to ensure confidentiality, integrity, and availability of information in both virtualized and nonvirtualized environments. Security management prevents unauthorized access and con- figuration of storage infrastructure components. For example, while deploying an application or a server, the security management tasks include managing the user accounts and access policies that authorize users to perform role-based activities. The security management tasks in a SAN environment include con- figuration of zoning to restrict an unauthorized HBA from accessing specific storage array ports. Similarly, the security management task on a storage array includes LUN masking that restricts a host's access to intended LUNs only.

Reporting

Reporting on a storage infrastructure involves keeping track and gathering information from various components and processes. This information is compiled to generate reports for trend analysis, capacity planning, chargeback, and performance. Capacity planning reports contain current and historic information about the utilization of storage, file systems, database tablespace, ports, and so on. Configuration and asset management reports include details about device allocation, local or remote replicas, and fabric configuration. This report also lists all the equipment, with details, such as their purchase date, lease status, and maintenance records. Chargeback reports contain information about the allocation or utilization of storage infrastructure components by various departments or user groups. Performance reports provide details about the performance of various storage infrastructure components.

Storage Infrastructure Management in a Virtualized Environment

Virtualization technology has dramatically changed the complexity of storage infrastructure management. In fact, flexibility and ease of management are key drivers for wide adoption of virtualization at all layers of the IT infrastructure.

Storage virtualization has enabled dynamic migration of data and extension of storagevolumes. Due todynamic extension, storagevolumes can be expanded nondisruptively to meet both capacity and performance requirements. Because

virtualization breaks the bond between the storage volumes presented to the host and its physical storage, data can be migrated both within and across data centers without any downtime. This has made the administrator's tasks easier while reconfiguring the physical environment.

Virtual storage provisioning is another tool that has changed the infrastructure management cost and complexity scenario. In conventional provisioning, storage capacity is provisioned upfront in anticipation of future growth. Because growth is uneven, some users or applications find themselves running out of capacity, whereas others have excess capacity that remains underutilized. Use of virtual provisioning can address this challenge and make capacity management less challenging. In virtual provisioning, storage is allocated from the shared pool to hosts on-demand. This improves the storage capacity utilization, and thereby reduces capacity management complexities.

Virtualization has also contributed to network management efficiency. VSANs and VLANs made the administrator's job easier by isolating different

networks logically using management tools rather than physically separating them. Disparate virtual networks can be created on a single physical network, and reconfiguration of nodes can be done quickly without any physical changes. It has also addressed some of the security issues that might exist in a conven- tional environment.

On the host side, compute virtualization has made host deployment, reconfiguration, and migration easier than physical environment. Compute, application, and memory virtualization have not only improved provisioning, but also contributed to the high availability of resources.

Storage Management Examples

The following section provides examples of various storage management activities.

Example 1: Storage Allocation to a New Server/Host

Consider the deployment of a new RDBMS server to the existing nonvirtualized storage infrastructure. As a part of storage management activities, first, the administrator needs to install and configure the HBAs and device drivers on the server before it is physically connected to the SAN. Optionally, multipathing software can be installed on the server, which might require additional configuration. Further, storage array ports should be connected to the SAN.

As the next step, the administrator needs to perform zoning on the SAN switches to allow the new server access to the storage array ports via its HBAs. To ensure redundant paths between the server and the storage array, the HBAs of the new server should be connected to different switches and zoned with different array ports.

Further, the administrator needs to configure LUNs on the array and assign these LUNs to the storage array front-end ports. In addition, LUN masking configuration is performed on the storage array, which restricts access to LUNs by a specific server.

Theserver then discovers the LUNs assigned to it byeither a *bus rescan* process or sometimes through a server reboot, depending upon the operating system installed. A volume manager may be used to configure the logical volumes and file systems on the host. The number of logical volumes or file systems to be created depends on how a database or an application is expected to use the storage.The administrator's task also includes installation of a database or an application on the logical volumes or file systems that were created.

The last step is to make the database or application capable of using the new file system space. - 15-7 illustrates the activities performed on a server, a SAN, and a storage array for the allocation of storage to a new server.

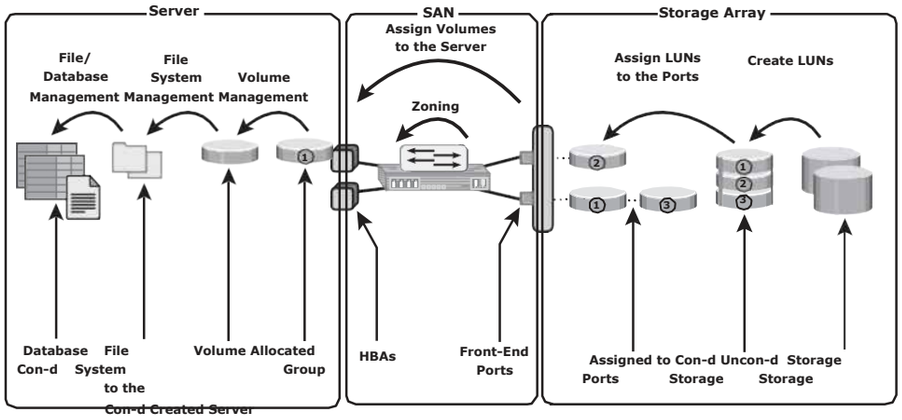
In a virtualized environment, provisioning storage to a VM that runs an RDBMS requires different administrative tasks.

Similar to a nonvirtualized environment, a physical connection must be established between the physical server, which hosts the VMs, and the storage array through the SAN. At the SAN level, a VSAN can be configured to transfer

data between the physical server and the storage array. The VSAN isolates this storage traffic from any other traffic in the SAN. Further, the administrator can

configure zoning within the VSAN.

At the storage side, administrators need to create thin LUNs from the shared storage pool and assign these thin LUNs to the storage array front-end ports. Similar to a physical environment, LUN masking needs to be carried out on the storage array.



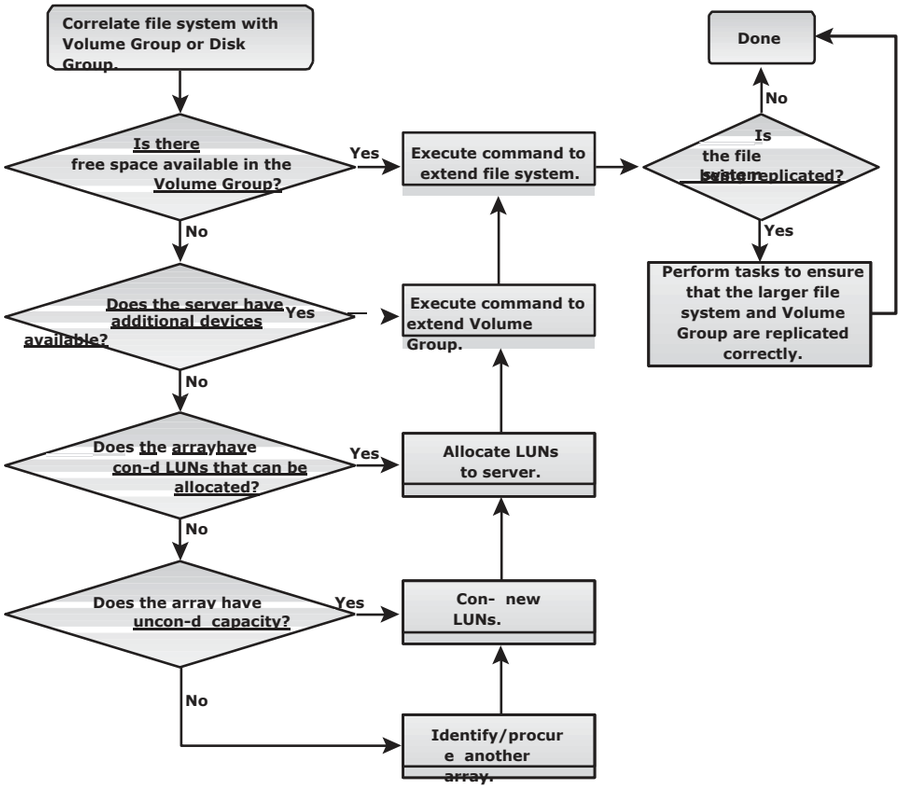
- 15-7: Storage allocation tasks

At the physical server side, the hypervisor discovers the assigned LUNs. The hypervisor creates alogical volume and filesystem to store and manage VM files. Then, the administrator creates a VM and installs the OS and RDBMS on the VM. While creating the VM, the hypervisor creates a virtual disk file and other VM files in the hypervisor file system. The virtual disk file appears to the VM as a SCSI disk and is used to store the RDBMS data. Alternatively, the hypervisor enables virtual provisioning to create a thin virtual disk and assigns it to the VM. Hypervisors usually have native multipathing capabilities. Optionally, a third-party multipathing software may be installed on the hypervisor.

Example 2: File System Space Management

To prevent a file system from running out of space, administrators need to per- form tasks to offload data from the existing file system. This includes deleting unwanted files or archiving data that is not accessed for a long time.

Alternatively, an administrator can extend the file system to increase its size and avoid an application outage. The dynamic extension of file systems or a logical volume depends on the operating system or the logical volume manager (LVM) in use. - 15-8 shows the steps and considerations for the extension of file systems in the flow chart.



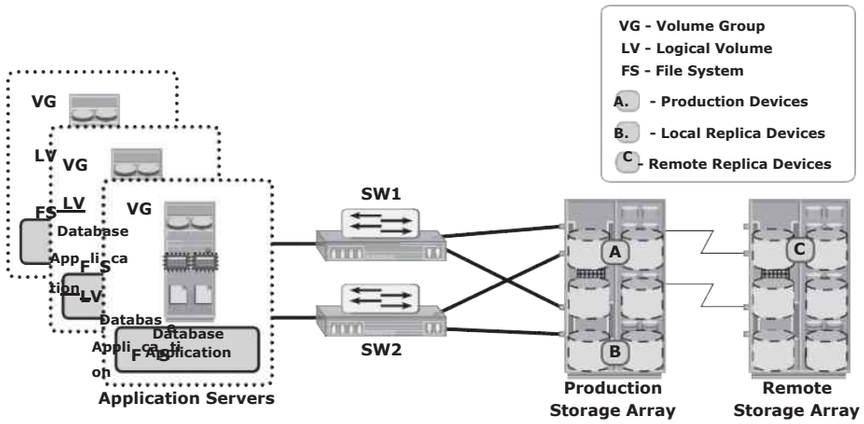
- 15-8: Extending a file system

Example 3: Chargeback Report

This example explores the storage infrastructure management tasks necessary to create a chargeback report.

- 15-9 shows a configuration deployed in a storage infrastructure. Three servers with two HBAs each connect to a storage array via two switches, SW1 and SW2. Individual departmental applications run on each of the servers. Array replication technology is used to create local and remote replicas. The production device is represented as A, the local replica device as B, and the remote replica device as C.

A report documenting the exact amount of storage resources used by each application is created using a chargeback analysis for each department. If the unit for billing is based on the amount of raw storage (usable capacity plus protection provided) configured for an application used by a department, the exact amount of raw space configured must be reported for each application. - 15-9 shows a sample report. The report shows the information for two applications, Payroll_1 and Engineering_1.

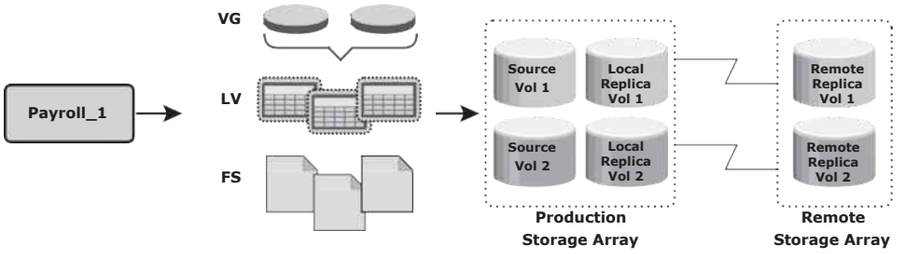


Application	Storage (GB)	Production Storage Raw (GB)	Local Replica Storage Raw (GB)	Remote Replica Storage Raw (GB)	Total Storage Raw (GB)	Chargeback Cost $ 5/Raw (GB)
Payroll_1	100	200	100	125	425	$ 2125
Engineering_1	200	250	200	250	700	$ 3500

- 15-9: Chargeback report

The first step to determine chargeback costs is to correlate the application with the exact amount of raw storage configured for that application.

As indicated in - 15-10, the Payroll_1 application storage space is traced from file systems to logical volumes to volume groups and to the LUNs on the array. When the applications are replicated, the storage space used for local replication and remote replication is also identified. In the example shown, the application is using Source Vol 1 and Vol 2 (in the production array). The replication volumes are Local Replica Vol 1 and Vol 2 (in the production array) and Remote Replica Vol 1 and Vol 2 (in the remote array).



- 15-10: Correlation of capacity con-d for an application

The amount of storage allocated to the application can be easily computed after the array devices are identified. In this example, consider that Source

Vol 1 and Vol 2 are each 50 GB in size, the storage allocated to the application is 100 GB (50 + 50). The allocated storage for replication is 100 GB for local replication and 100 GB for remote replication. From the allocated storage, the raw storage configured for the application is determined based on the RAID protection that is used for various array devices.

If the Payroll_1 application's production volumes are RAID 1-protected, the raw space used by the production volumes is 200 GB. Assume that the local replicas are on unprotected volumes, and the remote replicas are protected with a RAID 5 configuration, then 100 GB of raw space is used by the local replica and 125 GB by the remote replica. Therefore, the total raw capacity used by the Payroll_1 application is 425 GB. The total cost of storage provisioned for Payroll_1 application will be $2,125 (assume cost per GB of storage is $5). This exercise must be repeated for each application in the enterprise to generate the chargeback report.

Chargeback reports can be extended to include a pre-established cost of other resources, such as the number of switch ports, HBAs, and array ports in the configuration. Chargeback reports are used by data center administrators to ensure that storage consumers are well aware of the costs of the services that they have requested.

Storage Infrastructure Management Challenges

Monitoring and managing today'scomplexstorageinfrastructure is challenging. This is due to the heterogeneity of storage arrays, networks, servers, databases, and applications in the environment. For example, heterogeneous storage arrays vary in their capacity, performance, protection, and architectures.

Each of the components in a data center typically comes with vendor-specific tools for management. An environment with multiple tools makes understanding the overall status of the environment challenging because the tools may not be interoperable. Ideally, management tools should correlate information from all components in one place. Such tools provide an end-to-end view of the environment, and a quicker root cause analysis for faster resolution to alerts.

Developing an Ideal Solution

An ideal solution should offer meaningful insight into the status of the overall infrastructure and provide root cause analysis for each failure. This solution should also provide central monitoring and management in a multivendor storage environment and create an end-to-end view of the storage infrastructure.

The benefit of end-to-end monitoringis the ability to correlate one component's behavior with the other. In many cases, looking at each component individu- ally may not be sufficient to reveal the actual cause of the problem. The central monitoring and management system should gather information from all the components and manage them through a single-user interface. In addition, it must provide a mechanism to notify administrators about various events using methods, such as e-mail and Simple Network Management Protocol (SNMP) traps. It should also have the capability to generate monitoring reports and run automated scripts for task automation.

Theidealsolutionmustbebased on industry standards, by leveragingcommon APIs, data model terminology, and taxonomy. This enables the implementation of policy-based management across heterogeneous devices, services, applica- tions, and deployed topologies.

Traditionally, SNMP protocol was the standard used to manage multivendor SAN environments. However, SNMP was inadequate for providing the detailed information required to manage the SAN environment. The unavailability of automatic discovery functions and weak modeling constructs are some inad- equacies of SNMP in a SAN environment. Even with these limitations, SNMP still holds a predominant role in SAN management, although newer open stor- age SAN management standards have emerged to monitor and manage storage environments more effectively.

Storage Management Initiative

The Storage Networking Industry Association (SNIA) has been engaged in an initiative to develop a common storage management interface. SNIA has developed a specification called Storage Management Initiative-Specification (SMI-S). This specification is based on the Web-Based Enterprise Management (WBEM) technology, and Distributed Management TaskForce's(DMTF) Common Information Model (CIM). The initiative was formally created to enable broad interoperability and management among heterogeneous storage and SAN components. For more information, see www.snia.org.

SMI-S offers substantial benefits to users and vendors. It forms a normalized, abstracted model to which a storage infrastructure's physical and logical com- ponents can be mapped. This model is used by management applications, such as storage resource management, device management, and data management, for standardized, end-to-end control of storage resources.

UsingSMI-S,devicesoftwaredevelopershaveaunifiedobjectmodelwithdetails about managing the breadth of storage and SAN components. SMI-S-compliant products lead to easier, faster deployment and accelerated adoption of policy-

based storage management frameworks. Moreover, SMI-S eliminates the need for the development of vendor-proprietary management interfaces and enables

vendors to focus on value-added features.

Enterprise Management Platform

An enterprise managementplatform(EMP) is a suite of applicationsthat provides an integrated solution for managing and monitoring an enterprise storage infra- structure. These applications have powerful, flexible, unified frameworks that provide end-to-end management of both physical and virtual resources. EMP provides a centrally managed, single point of control for resources throughout the storage environment.

These applications can proactively monitor storage infrastructure components and alert users about events. These alerts are either shown on a console depicting
the faultycomponentin a differentcolor, or they can be configured to sendthe alert by e-mail. In addition to monitoring, an EMP provides the necessary management functionality, which can be natively implemented into the EMP or can launch the proprietary management utility supplied by the componentmanufacturer.

An EMP also enables easy scheduling of operations that must be performed regularly, such as the provisioning of resources, configuration management, and fault investigation. These platforms also provide extensive analytical, remedial, and reporting capabilities to ease storage infrastructure management. EMC ControlCenter and EMC Prosphere, described in section 15.7 —Concepts in Practice,‖ are examples of an EMP.

Information Lifecycle Management

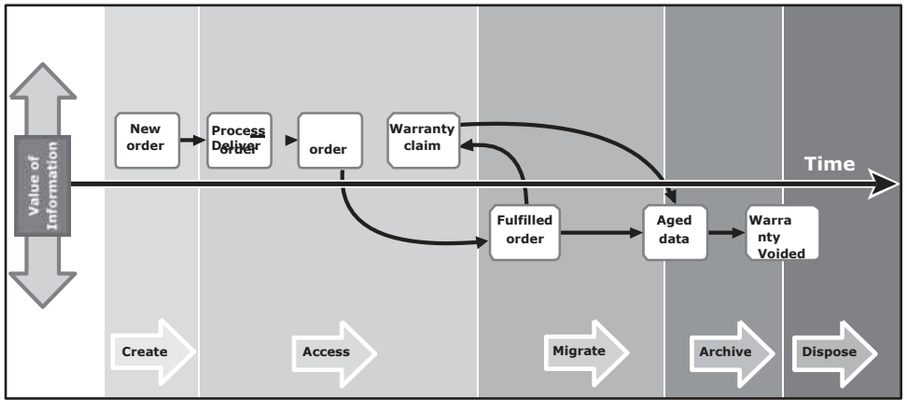
In both traditional data center and virtualized environments, managing information can be expensive if not managed appropriately. Along with the tools, an effective management strategy is also required to manage information efficiently. This strategy should address the following key challenges that exist in today's data centers:

- **Exploding digital universe:** The rate of information growth is increasing exponentially. Creating copies of data to ensure high availability and repurposinghascontributedtothemultifoldincreaseofinformationgrowth.
- **Increasing dependency on information:** The strategic use of information plays animportant role in determining the success ofabusiness and provides competitive advantages in the marketplace.
- **Changingvalueofinformation**:Informationthatisvaluabletodaymight become less important tomorrow. The value of information often changes over time.

Framing a strategy to meet these challenges involves understanding the value of information overitslife cycle. When information isfirst created, it often has the highest value and is accessed frequently. As the information ages, it is accessed less frequently and is of less value to the organization. Understanding the value

of information helps to deploy the appropriate infrastructure according to the changing value of information.

For example, in a sales order application, the value of the information (customer data) changes from the time the order is placed until the time that the warranty becomes void (see - 15-11). The value of the information is highest when a company receives a new sales order and processes it to deliver the product. After the order fulfillment, the customer data does not need to be available for real-time access. The company can transfer this data to less expensive secondary storage with lower performance until a warranty claim or another event triggers its need. After the warranty becomes void, the company can dispose of the information.



- 15-11: Changing value of sales order information

Information Lifecycle Management (ILM) is a proactive strategy that enables an ITorganization to effectively manage information throughout its lifecyclebased on predefined business policies. From data creation to data deletion, ILM aligns the business requirements and processes with service levels in an automated fashion. This allows an IT organization to optimize the storage infrastructure for maximum return on investment. Implementing an ILMstrategy has the following key benefits that directly address the challenges of information management:

- **Lower Total Cost of Ownership (TCO):** By aligning the infrastructure and management costs with information value. As a result, resources are not wasted, and complexity is not introduced by managing low-value data at the expense of high-valuedata.
- **Simplified management:** By integrating process steps and interfaces with individual tools and by increasing automation
- **Maintaining compliance:** By knowing what data needs to be protected for what length oftime
- **Optimized utilization:** By deploying storage tiering

Storage Tiering

Storage tiering is a technique of establishing a hierarchy of different storage types (tiers). This enables storing the right data to the right tier, based on service level requirements, at a minimal cost. Each tier has different levels of protection, performance, and cost. For example, high performance solid-state drives (SSDs) or FC drives can be configured as tier 1 storage to keep frequently accessed data, and low cost SATA drives as tier 2 storage to keep the less frequently accessed data. Keeping frequently used data in SSD or FC improves application performance. Moving less-frequently accessed data to SATA can free up storage capacity in high performance drives and reduce the cost of storage. This movement of data happens based on defined tiering policies. The tiering policy might be based on parameters, such as file type, size, frequency of access, and so on. For example, if a policy states —Move the files that are not accessed for the last 30 days to the lower tier,‖ then all the files matching this condition are moved to the lower tier.

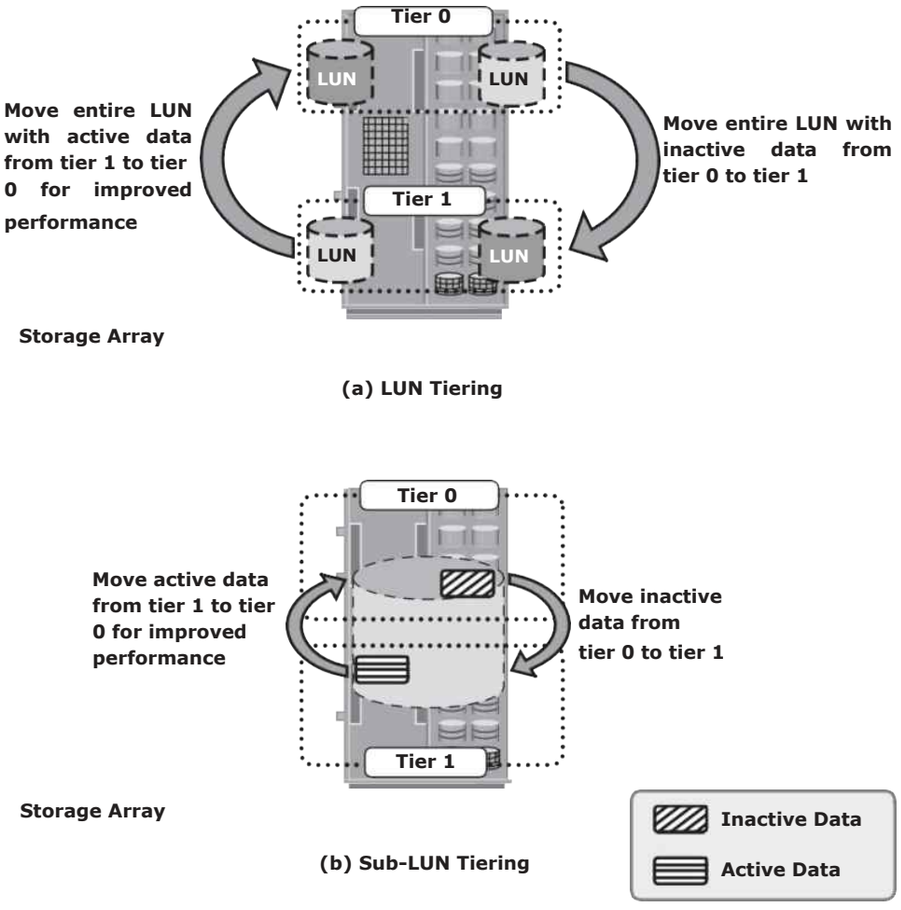
Storage tiering can be implemented as a manual or an automated process. *Manual storage tiering* is the traditional method where the storage administrator monitors the storage workloads periodically and moves the data between the tiers. Manual storage tiering is complex and time-consuming. *Automated storage tiering* automates the storage tiering process, in which data movement between the tiers is performed nondisruptively. In automated storage tiering, the applica- tion workload is proactively monitored; the active data is automatically moved to a higher performance tier and the inactive data to a higher capacity, lower performance tier. Data movements between various tiers can happen within (intra-array) or between (inter-array) storage arrays.

Intra-Array Storage Tiering

The process of storage tiering within a storage array is called *intra-array storage tiering*. It enables the efficient use of SSD, FC, and SATA drives within an array and provides performance and cost optimization. The goal is to keep the SSDs busy by storing the most frequently accessed data on them, while moving out the less frequently accessed data to the SATA drives. Data movements executed between tiers can be performed at the LUN level or at the sub-LUN level. The performance can be further improved by implementing tiered cache. LUN tiering, sub-LUN tiering, and cache tiering are detailed next.

Traditionally, storage tiering is operated at the LUN level that moves an entire LUN from one tier of storage to another (see - 15-12 [a]). This movement includes both active and inactive data in that LUN. This method does not give effective cost and performance benefits. Today, storage tiering

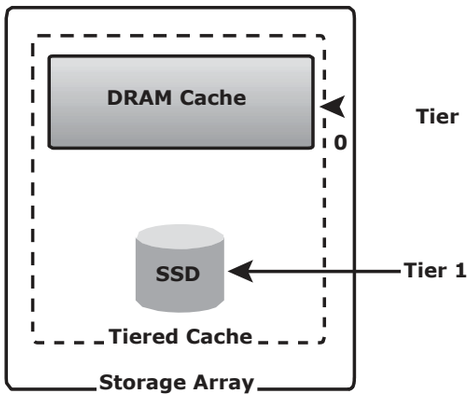
can be implemented at the sub-LUN level (see - 15-12 [b]). In sub-LUN level tiering, a LUN is broken down into smaller segments and tiered at that level. Movement of data with much finer granularity, for example 8 MB, greatly enhances the value proposition of automated storage tiering. Tiering at the sub-LUN level effectively moves active data to faster drives and less active data to slower drives.



- 15-12: Implementation of intra-array storage tiering

Tiering is also be implemented at the cache level, as shown in - 15-13. A large cache in a storage array improves performance by retaining a large amount of frequently accessed data in a cache, so most reads are served directly from the

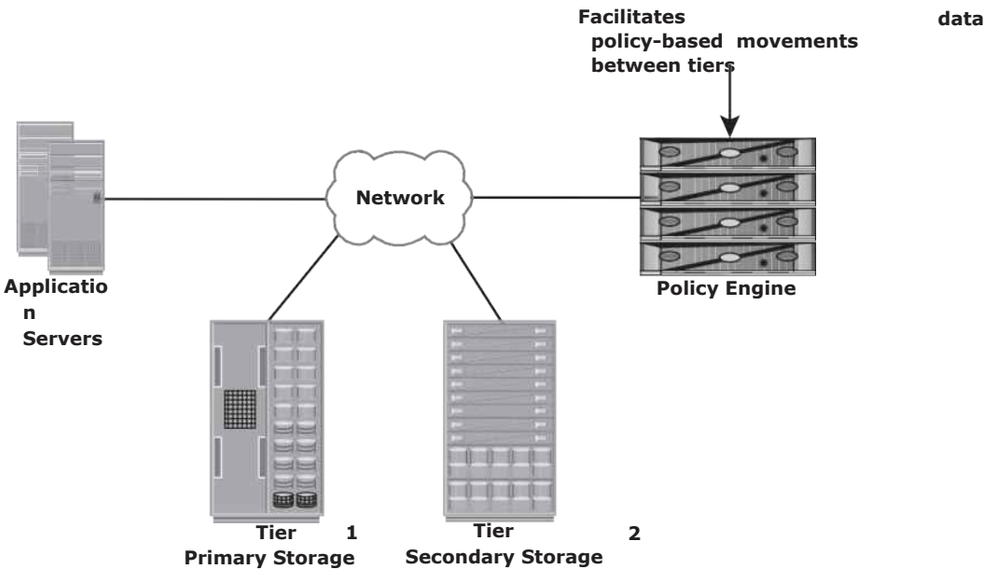
cache. However, configuring a large cache in the storage array involvesmore cost. An alternative way to increase the size of the cache is by utilizing the SSDs on the storage array. In cache tiering, SSDs are used as a large capacity secondary cache to enable tiering between DRAM (primary cache) and SSDs (secondary cache). Server flash-caching is another tier of cache in which a flash-cache card is installed in the server to further enhance the application's performance.



- 15-13: Cache tiering

Inter-Array Storage Tiering

The process of storage tiering between storage arrays is called *inter-array storage tiering*. Inter-array storage tiering automates the identification of active or inactive data to relocate them to different performance or capacity tiers between the arrays. - 15-14 illustrates an example of a two-tiered storage environment. This environment optimizes the primary storage for performance and the secondary storage for capacity and cost. The policy engine, which can be software or hardware where policies are configured, facilitates moving inactive or infrequently accessed data from the primary to the secondary storage. Some prevalent reasons to tier data across arrays is archival or to meet compliance requirements. As an example, the policy engine might be configured to relocate all the files in the primary storage that have not been accessed in one month and archive those files to the secondary storage. For each archived file, the policy engine creates a small space-saving stub file in the primary storage that points to the data on the secondary storage. When a user tries to access the file at its original location on the primary storage, the user is transparently provided with the actual file from the secondary storage.



- 15-14: Implementation of inter-array storage tiering

Concepts in Practice: EMC Infrastructure Management Tools

Businesses today face challenges in managing their IT infrastructure due to the large number of heterogeneous resources in their environment. These resources may be physical resources, virtualized resources, or cloud resources. EMC offers different tools that satisfy different requirements of the business. EMC ControlCenter and ProSphere are suites of software that can perform end-to-end management of storage infrastructure, while EMC Unisphere is software that managesEMCstoragearrays,suchasVNX andVNXe.EMCUnifiedInfrastructure Manager (UIM) is software that manages the Vblock infrastructure (cloud resources). For more information, visit www.emc.com/.

EMC ControlCenter and Prosphere

EMCControlCenterisafamilyofstorageresourcemanagement(SRM)applications that provide a unified solution to manage a multivendor storage infrastructure. It helps address the challenges to manage a large, complex storage environment that includes hosts, storage networks, storage, and virtualization across all the layers. ControlCenter provides capabilities, such as storage planning,

provisioning, monitoring, and reporting. Itenablesimplementingan ILMstrategy by providing comprehensive management of tiered storage infrastructure. It also provides an end-to-end view of the entire networked storage infrastructure that includes SAN, NAS, and host storage resources, including a virtualized environment. It provides a central administrative console, discovery of new components, quota management, event management, root cause analysis, and chargeback. ControlCenter comes with built-in security features that provide access control, data confidentiality, data integrity, logging, and auditing. It offers an intuitive, easy-to-use interface that provides insight into the complex relationships of the environment. ControlCenter uses an agent to discover the components in the environment.

EMC ProSphere is also storage resource management software built to meet the demands of the new cloud computing era. EMC ProSphere improves produc- tivity and service levels in the virtualized and cloud environment. ProSphere includes the following key capabilities:

- **End-to-end visibility:** It offers an intuitive, easy-to-use interface that provides insight into the complex relationships between objects in large, virtualized environments.
- **Multi-site management:** From a single console, ProSphere's federated architecture aggregates information from across sites and simplifies informationmanagementbetweendata centers. ProSphere ismanaged from a web browser to allow easy access over the Internet for remote management.
- **Improved productivity in growing virtualized environments:** ProSphere introduces an innovative technology called Smart Groups, which groups objects with similar characteristics into a user-defined group for perform- ing management tasks. This enables IT to take apolicy-based approach to manage objects or to set data collection policies.
- **Fast, easy, and efficient deployment:** Agent-less discovery eliminates the burdenofdeploying andmanaginghostagents. ProSphereispackaged as a virtual appliance that can be installed in a short time.
- **Delivery of IT as a service:** With ProSphere, service levels can now be monitored from host-to-storage layers. This allows organizations to main- tain consistent service levelsat an optimal price-performance ratio to meet business objectives to delivering IT-as-a-service.

EMC Unisphere

EMCUnisphereis a unified storagemanagementplatformthat providesintuitive user interfacesformanaging EMCVNX and EMCVNXestoragearrays.Unisphere

is web-enabled and supports remote management of storage arrays. Some of the key capabilities offered by Unisphere follow:

- Provides unified management for file, block, and object storage
- Provides single sign-on for all devices in a management domain
- Supports automated storage tiering and ensures that data is stored in the correct tier to meet performance and cost
- Provides management of both physical and virtual components

EMC Unified Infrastructure Manager (UIM)

EMC Unified Infrastructure Manageris a unifiedmanagementsolutionfor Vblocks. (Vblock is covered in Chapter 13.) It enables configuring the Vblock infrastructure resourcesand activating cloudservices. It providesa single user interface to manage multiple Vblocks and eliminates the need for configuring compute, network, and storage separately using different virtual infrastructure management tools. UIM provides a dashboard that shows how the Vblock infrastructure is config- ured and how the resources are used. This enables an administrator to monitor the configuration and utilization of the Vblock infrastructure resources and to plan for capacity requirements. UIM also provides a topology or a map view of the Vblock infrastructure, which enables an administrator to quickly locate and understand the interconnections of the Vblock infrastructure components and services. It provides an alerts console, which allows an administrator to see the alerts against the Vblock infrastructure resources and the associated ser- vices affected by problems. UIM performs a compliance check during resource configuration. It validates compliance with configuration best practices. It also prevents conflicting resource identity assignments, for example, accidentally assigning a MAC address to more than one virtual NIC.