# MODULE 5

***Syllabus: Securing and Managing Storage Infrastructure Securing and Storage Infrastructure: Information Security Framework, Risk Triad, Storage Security Domains, Security Implementations in Storage Networking, Securing Storage Infrastructure in Virtualized and Cloud Environments. Managing the Storage Infrastructure Monitoring the Storage Infrastructure, Storage Infrastructure Management activities, Storage Infrastructure Management Challenges, Information Lifecycle management, Storage Tiering***

## Chapter 14: Securing and managing storage infrastructure

14.1 Information Security Framework

The basic information security framework is built to achieve four security goals: confidentiality, integrity, and availability (CIA), along with accountability.

**Confidentiality:**

- Provides the required secrecy of information and ensures that only authorized users have access to data. This requires authentication 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.

**Integrity:**

- Ensures that the information is unaltered.
- Ensuring integrity requires detection of and protection against unauthorized alteration or deletion of information.
- Ensuring integrity stipulates measures such as error detection and correction for both data and systems.

**Availability:**

- This ensures that authorized users have reliable and timely access to systems, data, and applications residing on these systems.
- Availability requires protection against unauthorized deletion of data and denial of service .
- Availability also implies that sufficient resources are available to provide a service.

**Accountability service:**

- Refers to accounting for all the events and operations that take place in the data center infrastructure.
- The accountability service maintains a log of events that can be audited or traced later for the purpose of security.

**Risk triad**

- Raid 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.

**14. 2.1 Assets**

- Information is one of the most important *assets* for any organization.
- 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.

**14.2.2 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.

***Passive Attacks***

**Eavesdropping:** When someone overhears a conversation, the unauthorized access to this information is called eavesdropping.

**Snooping:** This refers to accessing another user's data in an unauthorized way. In general, snooping and eavesdropping are synonymous.

**14. 2. 3Vulnerability**

- 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 adequate security controls at all the access points on an access path.
- 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.
- *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.
- An ***attack vector*** is a step or a series of steps necessary to complete an attack vector 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.
- ***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.
- 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.

### 14.3 Storage Security Domains

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*. Figure 14-1 depicts the three security domains of a storage system environment.

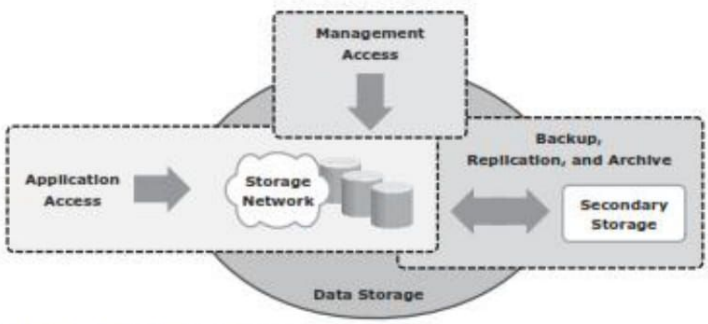



**Figure 14-1:** Storage security domains

- The first security domain involves application access to the stored data through 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.

### 14.3.1 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.
- Figure 14-2 shows application access in a storage networking environment.

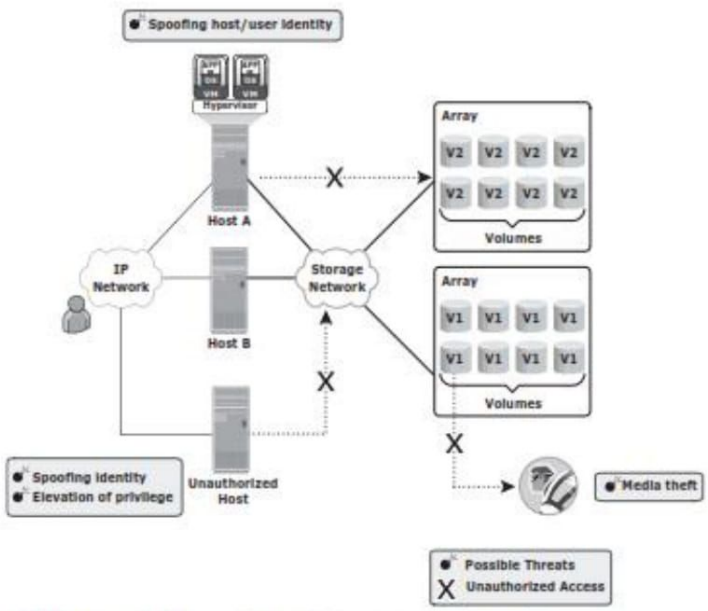



**Figure 14-2:** Security threats in an application access domain

- 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's resources.

- Another threat could be that an unauthorized host gains access the network; the attacker on this host may try to spoof the identity of another host and tamper 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.

***Controlling User Access to Data***

- 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 network 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 spoofing.
- 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.
- 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 controls 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.

***Protecting the Storage Infrastructure***

- The security controls for protecting the network fall into two general categories: *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
- It is also advisable to consider administrative controls, such as "separation of duties," 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.

***Data Encryption***

- 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.

### 14.3.2 Securing the Management Access Domain

- Implementing appropriate controls for securing storage management applications is important because the damage that can be caused by using these applications can be far more extensive.
- Figure 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 to remote storage array B for replication purposes.

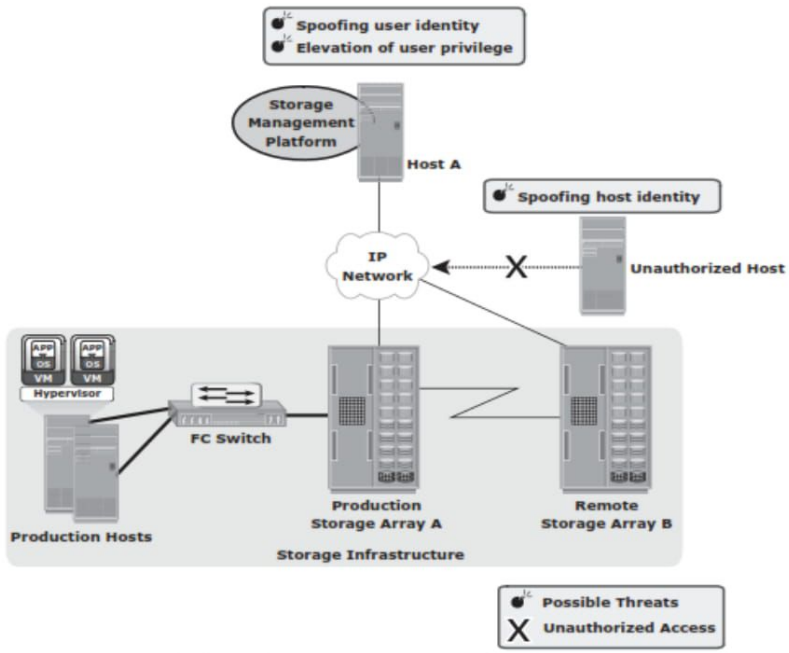



**Figure 14-3:** Security threats in a management access domain

- 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.
- Providing management access through an external network increases the potential for an unauthorized host or switch to connect to that network.
- 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 log monitoring helps to identify unauthorized access and unauthorized changes to the infrastructure.

***Controlling Administrative Access***

- Controlling administrative access to storage aims to safeguard against the threats of an attacker spoofing an administrator's identity 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 with third-party authentication directories, such as Lightweight Directory Access Protocol (LDAP) or Active Directory.

***Protecting the Management Infrastructure***

- Mechanisms to protect the management network infrastructure include encrypting 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.
- Access controls need to be enforced at the storage-array level to specify which host has management access to which array.
- A separate private management network is highly recommended for management 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.

### 14.3.3 Securing Backup, Replication, and Archive

BURA is the third domain that needs to be secured against attack. BURA is complex and is based on the BURA software accessing the storage arrays.

Organizations must ensure that the DR site maintains the same level of security for the backed up data. Protecting the BURA infrastructure requires addressing several threats, including spoofing the legitimate identity of a DR site, tampering with data, network snooping, DoS attacks, and media theft.

Figure 15-4 illustrates a generic remote backup design whereby data on a storage array is replicated over a disaster recovery (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 claims to be the backup server at the DR site, which may lead to a remote backup being performed to an unauthorized and unknown site. In addition, attackers can use the connection to the DR network to tamper with data, snoop the network for authentication data, and create a DoS attack against the storage devices.

**Security Implementations in Storage Networking**

### 14.4.1 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.
- Many FC SAN security mechanisms have evolved from their counterpart in IP networking, thereby bringing in matured security solutions.

***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.

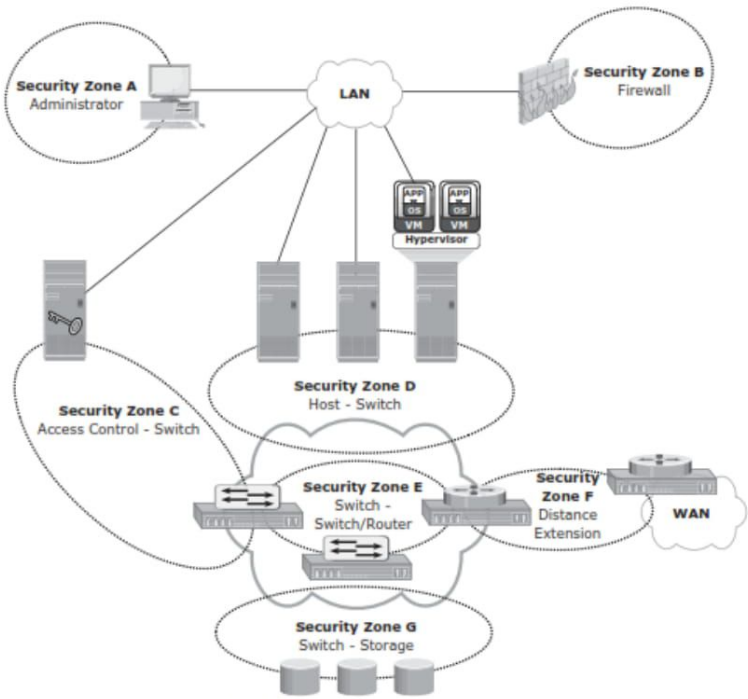



**Figure 14-5:** FC SAN security architecture

- Figure 14-5 illustrates various levels (zones) of a storage networking environment that must be secured and the security measures that can be deployed.

**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. |
| 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. |

### *Basic SAN Security Mechanisms*

Most commonly used SAN security methods.

1. LUN masking and zoning
2. switch-wide and fabric-wide access control
3. RBAC
4. logical partitioning of a fabric (Virtual SAN)

### *LUN Masking and Zoning*

- LUN masking and zoning are the basic SAN security mechanisms used to protect against unauthorized access to storage.
- The standard implementations of LUN masking on storage arrays mask the LUNs presented to a frontend 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.
- *Hard zoning* or *port zoning* is the mechanism of choice in security-conscious environments. Unlike soft zoning or WWPN zoning, it actually filters frames to ensure that only authorized zone members can communicate.

- However, it lacks one significant advantage of WWPN zoning: The zoning configuration must change if the source or the target is relocated across ports in the fabric. There is a trade-off between ease of management and the security provided by WWPN zoning and port zoning.

***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.
- *Persistent port disable* prevents a switch port from being enabled even after a switch reboot.

***Switch-Wide and Fabric-Wide Access Control***

- Network security can be configured on the FC switch by using *access control lists* (ACLs) and on the fabric by using fabric binding.
- **ACLs** incorporate the device connection control and switch connection 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 allowed to be part of the fabric, preventing unauthorized 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.

***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.
- Figure 14-6 depicts logical partitioning in a VSAN.

- The SAN administrator can create distinct VSANs by populating each of them with switch port
- In the example, the switch ports are distributed over two VSANs: 10 and 20 — for the Engineering and HR divisions, respectively.
- Although they share 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.

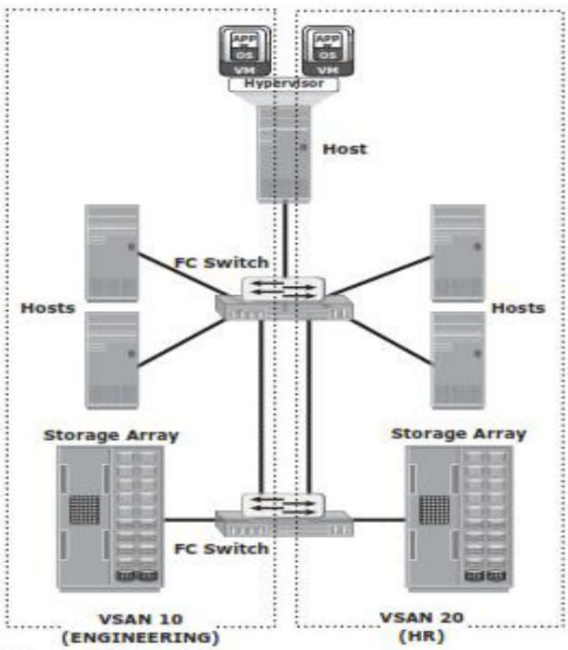



**Figure 14-6:** Securing SAN with VSAN

## NAS

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

***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, is used to determine 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 ownership.

- The owner of an object has hard-coded rights to that object, and these rights do not have to be explicitly granted in the SACL.

***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. UNIX permissions specify the operations that can be performed by any ownership relation with respect to a file.
- In simpler terms, these permissions specify what the owner can do, what the owner group can do, and what everyone else can do with the file.

***Authentication and Authorization***

- In a file-sharing environment, NAS devices use standard file-sharing protocols, NFS and CIFS. 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 directory, and Kerberos for network security.
- NAS devices use the same authentication techniques to validate network user credentials. Figure 15-7 depicts the authentication process 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.

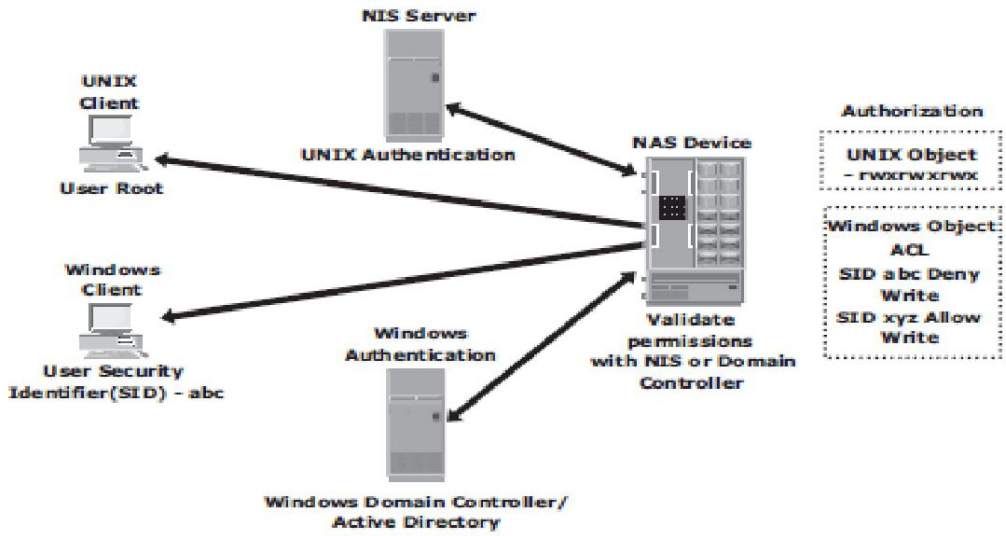



**Figure 14-7:** Securing user access in a NAS environment

## *Kerberos*

- Kerberos is a network authentication protocol. It is designed to provide strong authentication for client/server applications by using secret-key cryptography.
- The term *Kerberos server* generally refers to the Key Distribution Center (KDC).
- The KDC implements the 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.
- The Kerberos authorization process shown in Figure 15-8 includes the following steps:

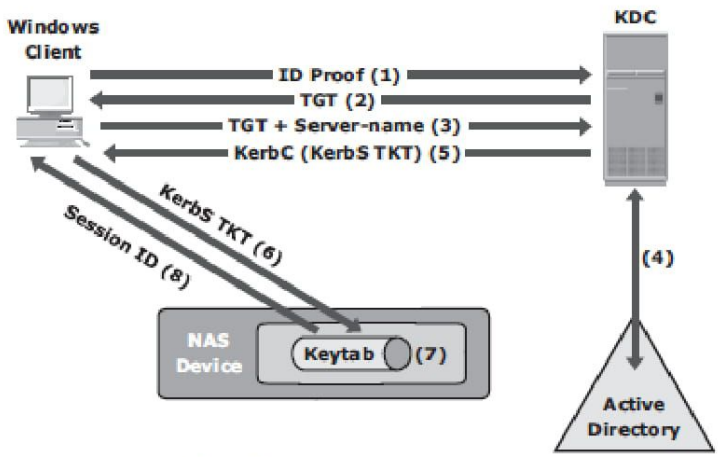



**Figure 14-8:** Kerberos authorization

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. (Note that this step is not explicitly shown in Figure 15-8.)

2. The KDC responds with a TGT (TKT is a key used for identification and has limited validity period). It contains two parts, one decryptable by the client and the other by the KDC.

3. When the client requests a service from a server, it sends a request, consist of the previously generated TGT 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 that is hosting the service.

6. The client then sends the service ticket to the server that houses the desired resources.

7. The server, in this case the NAS device, decrypts the server portion of the ticket and stores the information in a key tab 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 appropriate resources.

8. A client/server session is now established. The server returns a session ID to the client, which is used to track client activity, such as file locking, as long as the session is active.

**Network-Layer Fire walls**

- These network- layer firewalls are capable of examining network packets and comparing 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 requested 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.
- Figure 15-9 depicts a typical firewall implementation. Demilitarized zone (DMZ) is commonly used in networking environments.
- A DMZ provides a means of securing internal assets while allowing Internet-based access to various resources. In a DMZ environment, servers that need to be accessed through the

Internet are placed between 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. 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 that is accessing data stored on a NAS device, which may be located on the internal private network.

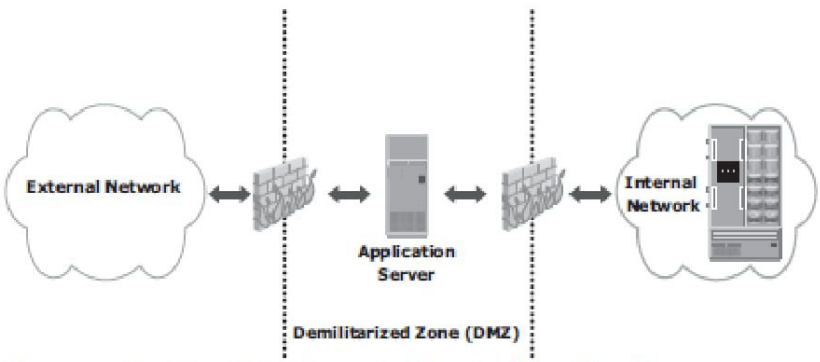



**Figure 14-9:** Securing a NAS environment with a network-layer firewall

## IP SAN

- The Challenge-Handshake Authentication Protocol (CHAP) is a basic authentication mechanism 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 wire; 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. Figure 15-10 depicts the CHAP authentication process.
- 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, comprising the name of a node and the secret associated with the node, is maintained by the target and the initiator.

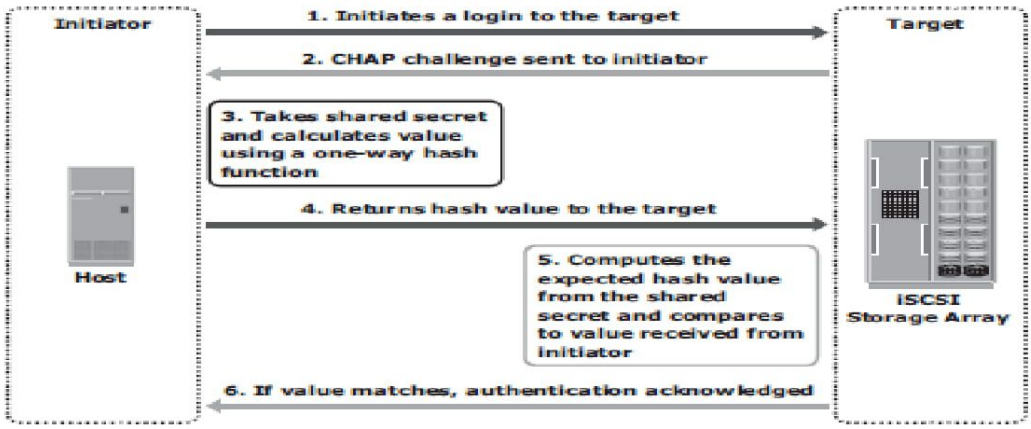



**Figure 14-10:** Securing IPSAN with CHAP authentication

- 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. In order for devices to communicate with one another, they must be configured in the same discovery domain.
- State change notifications (SCNs) tell the iSNS server when devices are added or removed from a discovery domain. Figure 15-11 depicts the discovery domains in iSNS.

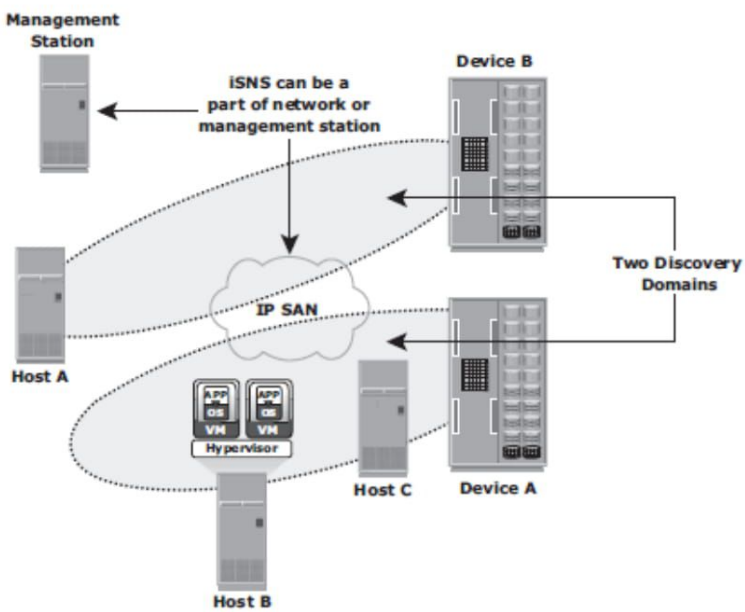



**Figure 14-11:** Securing IPSAN with iSNS discovery domains

## Securing Storage Infrastructure in Virtualized and Cloud Environments

- These environments have additional threats due to multitenancy and lack of control over the cloud resources
- Virtualization-specific security concerns are common for all cloud models
- In public clouds, there are additional security concerns, which demand specific countermeasures
- Clients have less control to enforce security measures in public clouds
- Difficult for cloud service provider(CSP) to meet the security needs of all the clients

## Security Concerns

- Multitenancy: Enables multiple independent tenants to be serviced using the same set of storage resources.
- Business critical data of one tenant is accessed by other competing tenants who run applications using the same resources.
- Velocity-of-attack: 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
- Data privacy is also a major concern in a virtualized and cloud environment.

## Security Measures

- **Security at the computer Level:**
  - Enforce the security of physical server, VMs, and hypervisor
  - Physical server: user authentication and authorization mechanism
  - Hypervisor: security-critical hypervisor updates should be installed regularly
- **Security at the network level:**
  - To minimize vulnerabilities at the network layer: firewall, intrusion detection, DMZ, encryption.
  - Virtual firewall
  - Provides packet filtering and monitoring of the VM-to-VM traffic
  - DMZ and data encryption
- **Security at the storage Level:**
  - Adequate security measures at computer & network levels helps to ensure storage security
  - Access control: to regulate which users and processes access data on storage system
  - Data encryption: encrypt backup, store encryption keys separately from data
  - Use separate LUNs for VM configuration files and VM data
  - Segregate VM traffic from management traffic

## Chapter 2

# Managing the Storage Infrastructure

## Monitoring the Storage Infrastructure

- 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.
- The major storage infrastructure components that should be monitored
  - Servers, databases and applications
  - Network (SAN and IP)
  - Storage arrays
- Each of these components should be monitored for accessibility, capacity, performance and security.

### Monitoring Parameters

- **Accessibility** refers to the availability of a component to perform a desired operation. A component is said to be accessible when it is functioning without any fault at any given point in time.
- **Monitoring hardware components** (e.g., a SAN interconnect device, a port, an HBA, or a disk drive) or **software components** (e.g.,a database instance) for accessibility involves checking their availability status by listening to pre-determined alerts from devices. For example, a port may go down resulting in a chain of availability alerts.
- A storage infrastructure uses redundant components to avoid a single point of failure. Failure of a component may cause an outage that affects application availability, or it may cause serious performance degradation even though accessibility is not compromised. For example, an HBA failure can restrict the server to a few paths for access to data devices in a multipath environment, potentially resulting in degraded performance.
- Continuously monitoring for expected accessibility of each component and reporting any deviations 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
  - The numbers of ports available on a switch.
- Inadequate capacity may lead to degraded performance or affect accessibility or even application/service availability.
- Capacity monitoring ensures uninterrupted data availability and scalability by averting outages before they occur. For example, if a report indicates that 90 percent of the ports are utilized in a particular SAN fabric, a new switch should be added if more arrays and servers need to be installed on the same fabric.
- Capacity monitoring is preventive and predictive, usually leveraged with advanced analytical tools for trend analysis. These trends help to understand emerging challenges, and can provide an estimation of time needed to meet them.
- **Performance monitoring** evaluates how efficiently different storage infrastructure components are performing and helps to identify bottlenecks.
- Performance monitoring usually measures and analyzes behavior in terms of response time or the ability to perform at a certain predefined level.
- It also deals with 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 to disks, application response time, network utilization, and server CPU utilization are examples of performance monitoring.
- **Security monitoring** helps to tracks unauthorized configuration changes of storage infrastructure elements. For example, security monitoring tracks and reports the initial zoning configuration performed and all subsequent changes.
- Physical security of a storage infrastructure is also continuously monitored using badge readers, biometric scans, or video cameras.

**Components Monitored**

Hosts, networks, and storage are the components within the storage environment that should be monitored for accessibility, capacity, performance, and security.

**Hosts**

- **Accessibility**
  - Hardware components: HBA, NIC, graphic card, internal disk
  - For example, an application crash due to host hardware failure can cause instant unavailability of the data to the user. Servers are used in a cluster to ensure high availability
  - Status of various processes/applications
- **Capacity**
  - File system utilization
  - For example capacity monitoring helps in estimating the file system's growth rate and predicting when it will reach 100 percent. Accordingly, the administrator can extend (manually or automatically) the file system's space proactively to prevent a failure resulting from a file system being full.
  - Database: Table space/log space utilization
  - User quota
- **Performance**
  - CPU and memory utilization
  - For example, if a server running an application is experiencing 80 percent of CPU utilization server may be running out of processing power, which can lead to degraded performance and slower response time. Administrators can take upgrading or adding more processors, shifting the workload to different servers, and restricting the number of simultaneous client access.
  - Memory utilization is measured by the amount of free memory available. Insufficient memory leads to excessive swapping and paging on the disk, which in turn affects response time to the applications.
  - Transaction response times
- **Security**
  - Login and authorization
  - For example, an administrator can block access to an unauthorized user if multiple login failures are logged.
  - Physical security (Data center access)

**Storage Network**

Uninterrupted access to data over the storage network depends on accessibility of the physical and logical components in storage network. Physical components of a storage network include elements such as switches, ports, cables, GBICs, and power supplies. The logical components include constructs, such as zones and fabrics. Any failure in the physical or logical components may cause data unavailability.

- **Accessibility**
  - Fabric errors, zoning errors, GBIC failure

- Device status/attribute Change
- Processor cards, fans, power supplies

- **Capacity**
  - ISL and port utilization
  - Availability of ports on a switch, number of available ports in the entire fabric, utilization of ISLs, individual ports, and each interconnect device in the fabric
- **Performance**
  - Connectivity ports
    - Link failures, Loss of signal, Link utilization
  - Connectivity devices
    - Port statistics
- Measuring receive or transmit link utilization metrics, which indicate how busy the switch port is, based on expected maximum throughput. Heavily used ports can cause queuing delays on the server.
- For IP networks, monitoring performance includes monitoring network latency, packet loss, bandwidth utilization for I/O, network errors, and collisions.
- **Security**
- Storage network security monitoring provides information for any unauthorized change to the configuration of the fabric—for example, changes to the zone policies that can affect data security. Login failures and unauthorized access to switches for performing administrative changes should be logged and monitored continuously
  - Zoning and LUN Masking
  - Administrative Tasks and physical security
  - Authorized Access, Strict Passwords

**Storage**

- **Accessibility**
  - All Hardware components
  - For example, the failure of a replication task affects disaster recovery capabilities. Some storage arrays also provide the capability to send a message to the vendor's support center in the event of hardware or process failures, referred to as a *call home*.
  - Array Operating Environment
    - RAID processes
    - Environmental Sensors
    - Replication processes
- **Capacity**
- Capacity monitoring of a storage array enables the administrator to respond to storage needs as they occur. Information about fan-in or fan-out ratios and the availability of front-end ports is useful when a new server is given access to the storage array.
  - Configured/un-configured capacity
  - Allocated/unallocated storage
  - Fan-in/fan-out ratios
- **Performance**
  - Performance metrics, such as utilization rates of the various storage array components, I/O response time, and cache utilization. A high utilization rate of

storage array components may lead to performance degradation.
    - FE and BE utilization/throughput
    - I/O profile, response time, cache metrics
- Security
    - Monitoring security helps to track unauthorized configuration of the storage array or corruption of data and ensures that only authorized users are allowed to access it.
    - Physical and administrative security

**Monitoring Examples**

A storage infrastructure requires implementation of an end-to-end solution to actively monitor all the parameters of its components.

**Accessibility Monitoring**

Failure of any component may affect the accessibility of one or more components due to their interconnections and dependencies, or it may lead to overall performance degradation.

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 storage array through two switches, SW1 and SW2, as shown in below figure. The three servers share two storage ports on the storage array. Path failover software has been installed on all three servers.

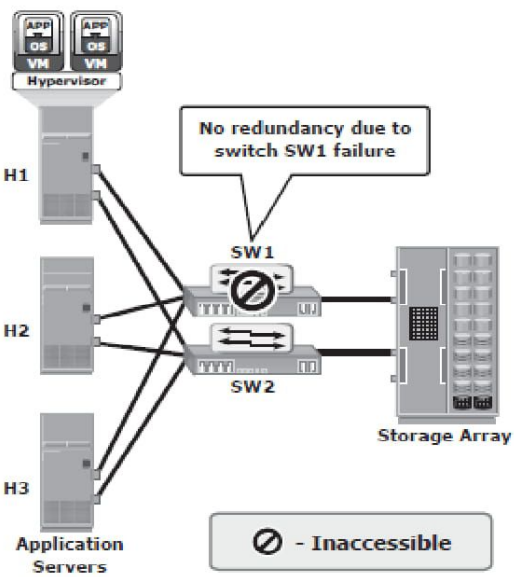



**Figure 15-1:** Switch failure in a storage infrastructure

If one of the *storage array ports* fails, all the storage volumes that were accessed through the switch connected to that port may become unavailable, depending on the type of storage array. If the storage volume becomes unavailable, path failover software initiates a path failover. However, due to redundant ports, the servers continue to access data through another switch, SW2. The servers H1, H2, and H3 may experience degraded performance due to an increased load on the path through SW2.

In the same example, if a single HBA fails on server H1, the server experiences path failure as shown in figure. However, due to redundant HBAs, H1 can still access the storage device but it may experience degraded application response time (depends on I/O load).

**Capacity Monitoring**

In the figure shown below the servers is allocated storage on the storage array. When a new server is deployed in this configuration, the applications on the new servers have to be given access to the storage devices from the array through switches SW1 and SW2.

Monitoring the available capacity on the array helps to proactively decide whether the array can provide the required storage to the new server. Other considerations include the availability of ports on SW1 and SW2 to connect to the new server as well as the availability of storage ports to connect to the switches.

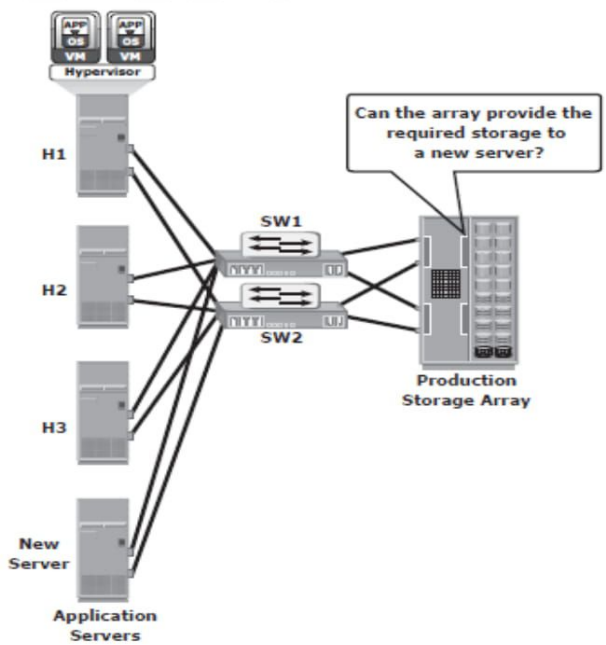



**Figure 15-2:** Monitoring storage array capacity

The following example illustrates the importance of monitoring file system capacity on servers. If file system capacity monitoring is not implemented, as shown in figure, and the file system is full, the application most likely will not function properly.

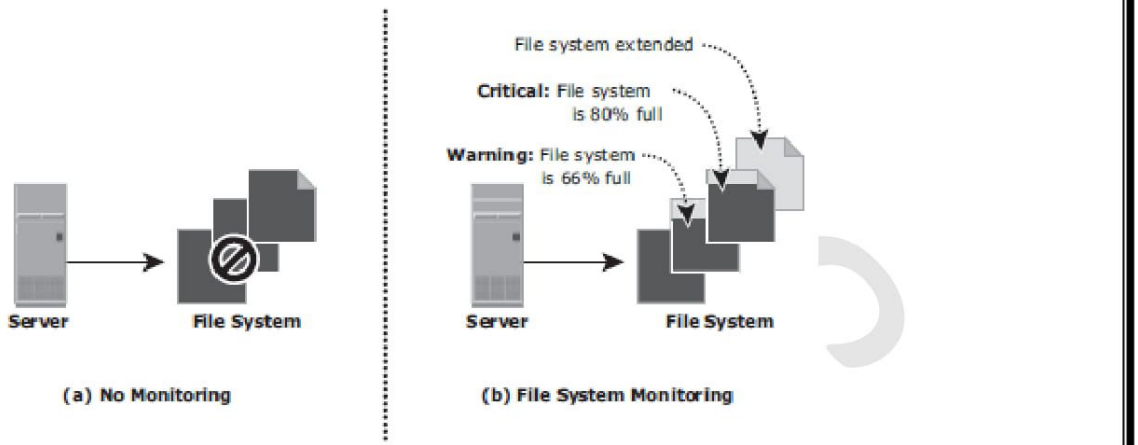



**Figure 15-3:** Monitoring server file system space

Monitoring can be configured to issue a message when thresholds are reached on file system capacity. For example, when the file system reaches 66 percent of its capacity a warning message is issued, and a critical message when the file system reaches 80 percent of its capacity.

**Performance Monitoring**

In the example shown below, servers H1, H2, and H3 (with two HBAs each) are connected to the storage array through switches SW1 and SW2. The three servers share the same storage ports on the storage array.

A new server H4 running an application with high work load, has to be deployed to share the same storage ports as H1, H2, and H3.

Monitoring array port utilization ensures that the new server does not adversely affect the performance of the other servers. In this example, utilization for the shared ports is shown by the solid and dotted lines in the line graph for the storage ports.

Notice that the port represented by a solid line is close to 100 percent utilization. If the actual utilization of both ports prior to deploying the new server is closer to the dotted line, there is room to add the new server. Otherwise, deploying the new server will affect the performance of all servers.

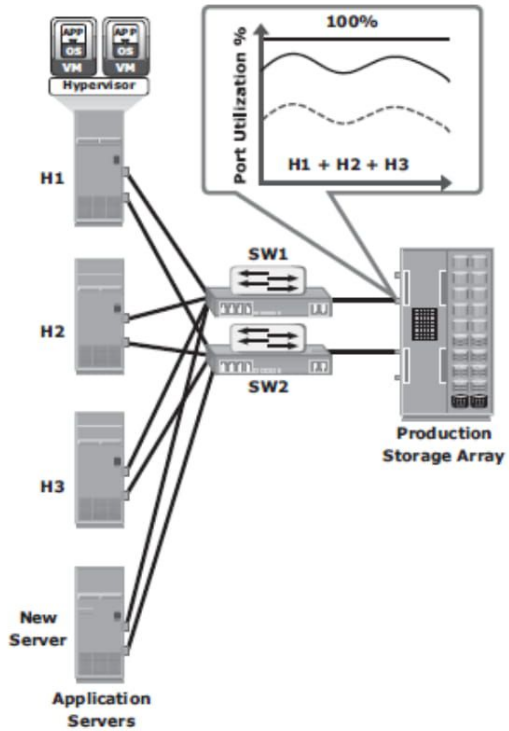



**Figure 15-4:** Monitoring array port utilization

Most servers offer tools that enable interactive monitoring of server CPU usage. For example, Windows Task Manager displays CPU and memory usage, as shown in slide. These interactive tools are useful only when a few servers need to be managed.

A storage infrastructure requires performance monitoring tools that are capable of monitoring many servers simultaneously.

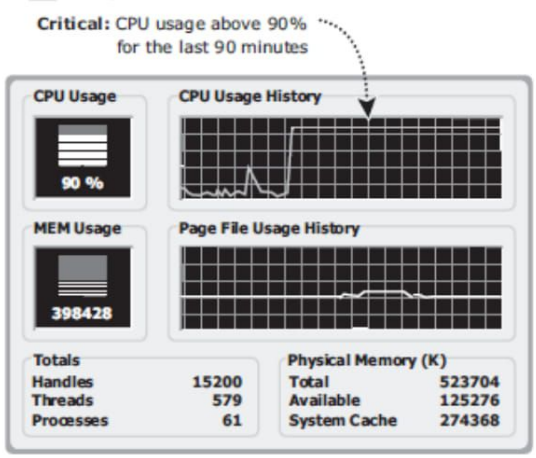



**Figure 15-5:** Monitoring the CPU and memory usage of a server

**Security Monitoring**

In this example shown below, the storage array is shared between two workgroups, WG1 and WG2. The data of WG1 should not be accessible by WG2. Likewise, WG2 should not be accessible by WG1. A user from WG1 may try to make a local replica of the data that belongs to WG2.

However, if this action is not monitored or recorded, it is difficult to track such a violation of security protocols.

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.

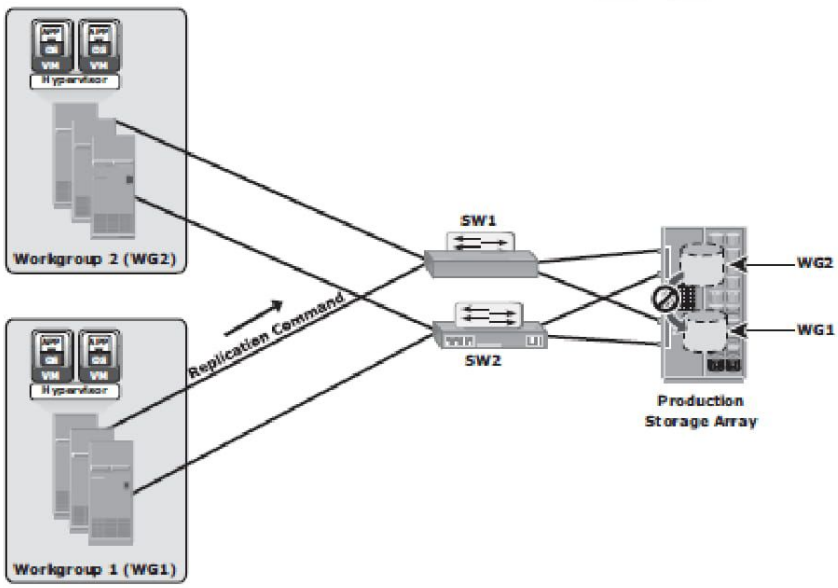



**Figure 15-6:** Monitoring security in a storage array

***Alerts***

- Alerting is an integral part of monitoring
- Monitoring tools enables administrators to assign different severity levels for different events
- Level of alerts based on severity
    - **Information alert**: Provide useful information and may not require administrator intervention
        - Creation of zone or LUN
    - **Warning alerts**: Require administrative attention
        - File systems becoming full/Soft media errors
    - **Fatal alert**: Require immediate administrative attention
        - Power failures/Disk failures/Memory failures/Switch failures

## Storage Infrastructure Management Activities

The key storage infrastructure management activities performed in a data center can be broadly categorized into availability management, capacity management, performance management, security management, and reporting.

**Availability Management**

- Establishing guidelines for all configurations based on service levels
- To ensure high availability by:
    - Eliminating single points of failure deploy/configure
        - Deploying Two or more HBAs
        - Multipathing software with path failover capability, and server clustering
        - RAID protection
        - Redundant Fabrics
    - Configuring data backup and replication
    - Deploying virtualized environment

**Capacity Management**

- Ensures adequate availability of resources based on their service level requirements
- Capacity management provides capacity analysis, comparing allocated storage to forecasted storage on a regular basis.
- Manages resource allocation
- Key activities
    - Trend and Capacity analysis: actual utilization of allocated storage and rate of consumption
    - Storage provisioning
    - Examples
        - Host: Host configuration and file system/DB management
        - SAN: Unused Ports and Zoning
        - Storage: Device configuration and LUN Masking

**Performance Management**

- Configure/design for optimal operational efficiency
- Helps to identify the performance of storage infrastructure components. This analysis provides the information — whether a component is meeting expected performance

levels. Several performance management activities are initiated for the deployment of an application or server in the existing storage infrastructure.

- Performance analysis
    - Identify bottlenecks
    - Fine tuning for performance enhancement
- Key activities
    - Host: Volume management, database/application layout
    - SAN: Designing sufficient ISLs with adequate bandwidth
    - Storage Array: Choice of RAID type and layout of devices (LUNs) and choice of front-end ports

**Security Management**

- Prevent unauthorized activities or access
- For example, while deploying an application or a server, the security management tasks include managing user accounts and access policies, that authorizes users to perform role-based activities.
- Key activities
    - Server:
        - Creation of user logins, user privileges
    - SAN:
        - Configuration of zoning to restrict unauthorized HBA's
    - Storage Array:
        - LUN masking prevents data corruption on the storage array by restricting host access to a defined set of logical devices

**Reporting**

- Reporting on a storage infrastructure involves keeping track and gathering information from various components/processes
- This information is compiled to generate reports for trend analysis, capacity planning, chargeback, performance, and to illustrate basic configuration of storage infrastructure components
- Also used to provide information for Capacity, Availability, Security and Performance Management

## Storage Management Examples

### *Example 1: Storage Allocation to a New Server/Host*

Consider a deployment of the new RDBMS server to the existing non-virtualized SAN environment. Following are the management activities carried out to allocate storage to new host.

**Storage array management activities:** The administrator needs to configure new volumes on the array then assign those volumes to the array front-end ports. Administrator also need to configure LUN masking on the storage array by assigning new servers and volumes to the storage group.

**Server management activities:** The installation and configuration of the HBA hardware (at least two to ensure redundancy) and driver has to be performed on the server before it can be physically connected to the SAN. Server reconfiguration may be required, depending on the operating system installed on the server, so it can recognize the new devices. Optional multipathing software can be installed on the server, which might require additional configuration. The volume management tasks on the host involve the creation of volume groups, logical volumes, and file systems. On the application side, whether it is a database or any other type of application, administrator tasks include installation of the database or the application on the logical volumes or file systems that were created. Figure 15-7 illustrates the activities performed on a server, a SAN, and a storage array for the allocation of storage to a new server.

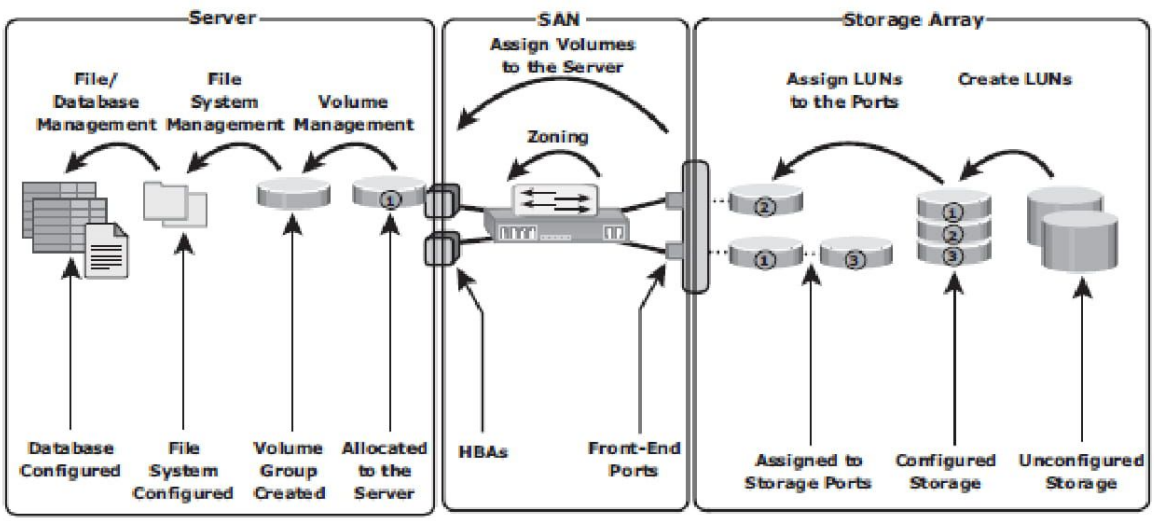



**Figure 15-7:** Storage allocation tasks

**SAN management activities:** The administrator configures the fabric's zoning policies for the new server's HBA, allowing the host to access the storage array port via the specific HBA port. This operation should probably be done at two or more fabrics to ensure redundant paths between the hosts and the storage array. The switches should have free ports available for the new server, and the array port utilization is validated against the required I/O performance of the server if the port is shared between many servers.

### *Example 2: File System Space Management*

To prevent a fi le system from running out of space, administrators need to perform tasks to offload data from the existing fi le 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. Figure 15-8 shows the steps and considerations for the extension of file systems in the flow chart.

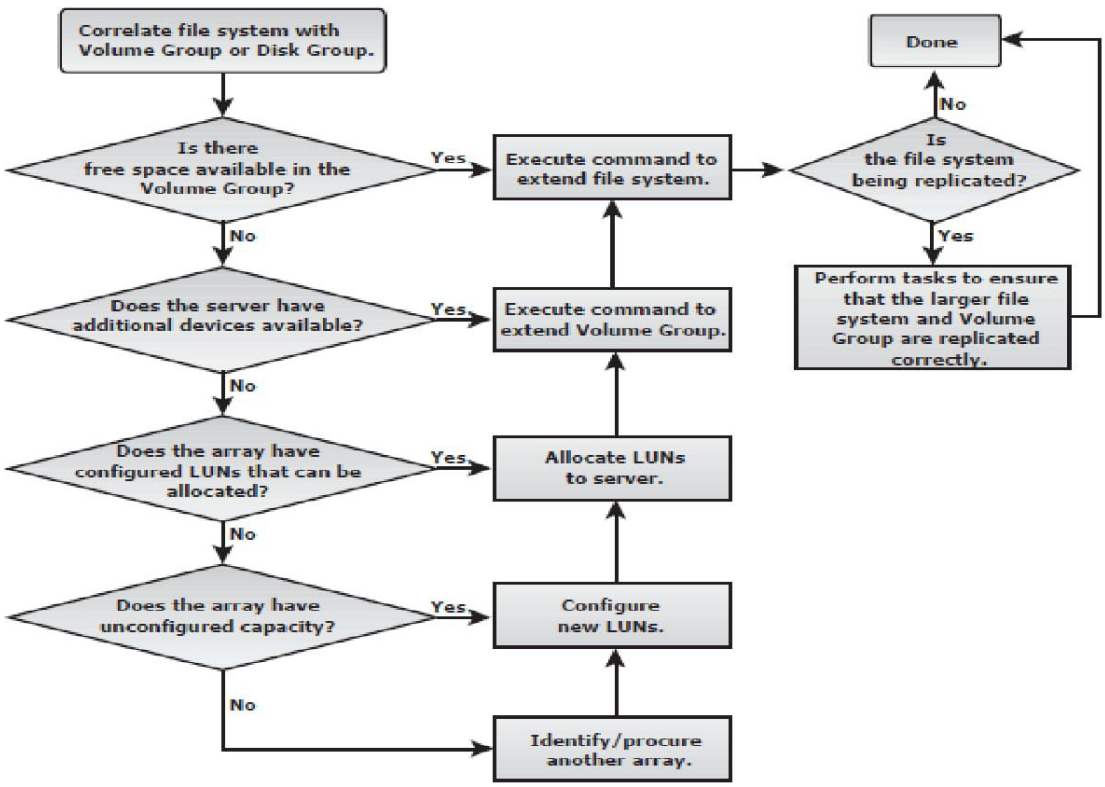



**Figure 15-8:** Extending a file system

**Storage Infrastructure Management Challenges**

Monitoring and managing today's complex storage infrastructure environment has become very challenging due to the number and variety of storage arrays, networks, servers, databases, and applications.

- Variety of storage devices varying in capacity, performance, and protection methodologies
- Storage infrastructures deploy both SAN and IP networks
- Servers with different operating systems such as UNIX, LINUX, Windows, or mainframe

These products and services from multiple vendors may have interoperability issues which add complexity in managing storage infrastructure. All of these components are provided with vendor-specific tools to manage and monitor them.

## Information Lifecycle Management

The 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 repurposing has contributed to the multifold increase of information growth.
- **Increasing dependency on information**: The strategic use of information plays an important role in determining the success of a business and provides competitive advantages in the marketplace.
- **Changing value of information:** Information that is valuable today might become less important tomorrow. The value of information often changes over time.

When information is first 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.

- 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 Figure 15-11).
- High value → new sales order and process to deliver product
- Medium or low value → after order fulfillment → data can transfer to less expensive secondary storage
- No value → warranty becomes void → can dispose the data

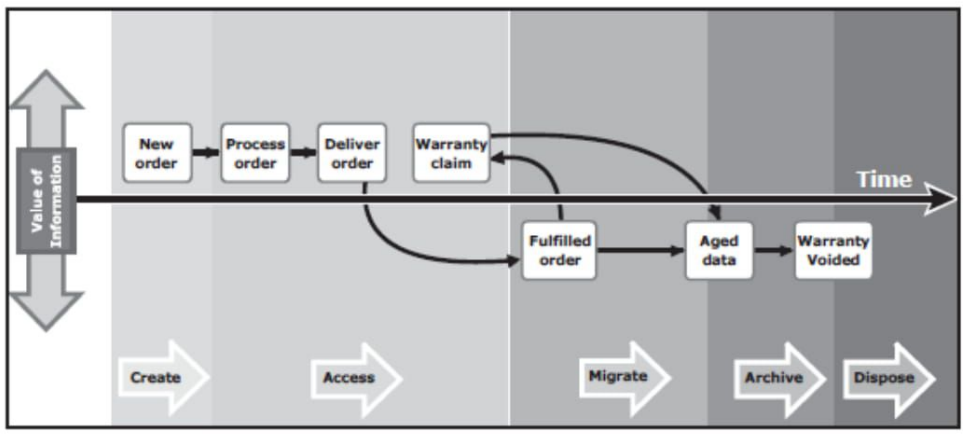



**Figure 15-11:** Changing value of sales order information

**Information Lifecycle Management (ILM)** is a proactive strategy that enables an IT organization to effectively manage information throughout its life cycle based 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 ILM strategy 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-value data.
- **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 of time
- **Optimized utilization:** By deploying storage tiering

## Storage Tiering

It is a technique of establishing a hierarchy of storage types and identifying the candidate data to relocate to the appropriate storage type to meet service level requirements at a minimal cost.

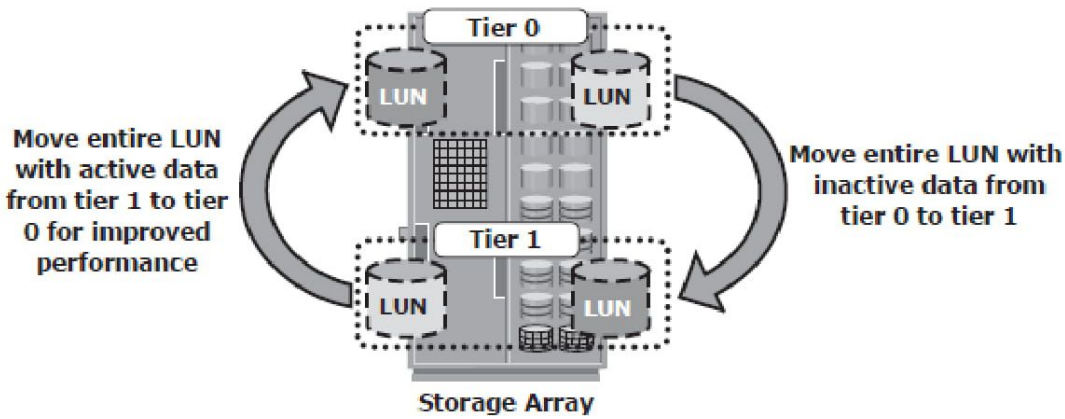
- Each tier has different levels of protection, performance, and cost
- Efficient storage tiering requires defining tiering policies

- Storage tiering implementations are:
  - Manual storage tiering
  - Automated storage tiering
- Data movement occurs between tiers
  - Within a storage array (Intra-array)
  - Between storage arrays (Inter-array)

**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.

- LUN tiering
  - Moves entire LUN from one tier to another
  - Does not give effective cost and performance benefits




(a) LUN Tiering

- Sub-LUN tiering
  - A LUN is broken down into smaller segments and tiered at that level
  - Provides effective cost and performance benefits

**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. Figure 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.

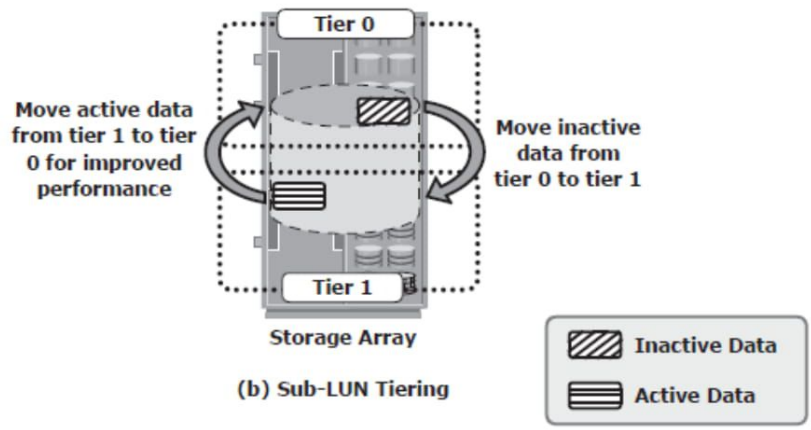



**Figure 15-12:** Implementation of intra-array storage tiering

**Cache Tiering**

- Enables creation of a large capacity secondary cache using SSDs
- Enables tiering between DRAM cache and SSDs (secondary cache)
- Most reads are served directly from high performance tiered cache

**Benefits**

- Enhances performance during peak workload
- Non-disruptive and transparent to applications

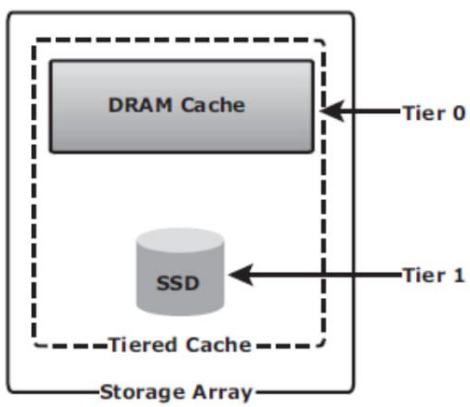



**Figure 15-13:** Cache tiering

********** END**********