



Section 3: Backup, Archive, and Replication

This section contains four modules:

- Module 9: Introduction to Business Continuity
- Module 10: Backup and Archive
- Module 11: Local Replication
- Module 12: Remote Replication

Welcome to Section 3: Backup, Archive, and Replication. This section covers four modules, namely Introduction to Business Continuity, Backup and Archive, Local Replication, and Remote Replication.



Module 9: Introduction to Business Continuity

Upon completion of this module, you should be able to:

- Define business continuity (BC) and information availability (IA)
- Explain the impact of information unavailability
- Describe BC planning process
- Explain business impact analysis (BIA)
- Explain BC technology solutions

This module focuses on the importance of business continuity, the factors that can affect information availability, and the consequences of information unavailability. This module also details on BC planning process and BC technology solutions, specifically on eliminating single points of failure.

Module 9: Introduction to Business Continuity

Lesson 1: Business Continuity Overview

During this lesson the following topics are covered:

- Business continuity
- Information availability metrics

This lesson covers the importance of business continuity to an organization, factors that can affect information availability and the consequences of information unavailability. Also this lesson focuses on information availability metrics namely mean time between failure (MTBF) and mean time to repair (MTTR).

Why Business Continuity?

- Information is an organization's most important asset
- Continuous access to information ensures smooth functioning of business operations
- Cost of unavailability of information to an organization is greater than ever

In today's world, continuous access to information is a must for the smooth functioning of business operations. The cost of unavailability of information is greater than ever, and outages in key industries cost millions of dollars per hour. There are many threats to information availability, such as natural disasters, unplanned occurrences, and planned occurrences, that could result in the inaccessibility of information. Therefore it is critical for businesses to define appropriate strategies that can help them to overcome these crises. Business continuity is an important process to define and implement these strategies.

What is Business Continuity?

Business Continuity

It is a process that prepares for, responds to, and recovers from a system outage that can adversely affect business operations.

- An integrated and enterprise-wide process that includes set of activities to ensure “information availability”
- BC involves proactive measures and reactive countermeasures
- In a virtualized environment, BC solutions need to protect both physical and virtualized resources

Business continuity (BC) is an integrated and enterprise-wide process that includes all activities (internal and external to IT) that a business must perform to mitigate the impact of planned and unplanned downtime. BC entails preparing for, responding to, and recovering from a system outage that adversely affects business operations. It involves proactive measures, such as business impact analysis, risk assessments, BC technology solutions deployment (backup and replication), and reactive measures, such as disaster recovery and restart, to be invoked in the event of a failure. The goal of a BC solution is to ensure the “information availability” required to conduct vital business operations.

In a virtualized environment, BC technology solutions need to protect both physical and virtualized resources. Virtualization considerably simplifies the implementation of BC strategy and solutions.

Information Availability

Information Availability

It is the ability of an IT infrastructure to function according to business expectations, during its specified time of operation.

- Information availability can be defined with the help of:
 - ▶ Accessibility
 - ▶▶ Information should be accessible to the right user when required
 - ▶ Reliability
 - ▶▶ Information should be reliable and correct in all aspects
 - ▶ Timeliness
 - ▶▶ Defines the time window during which information must be accessible

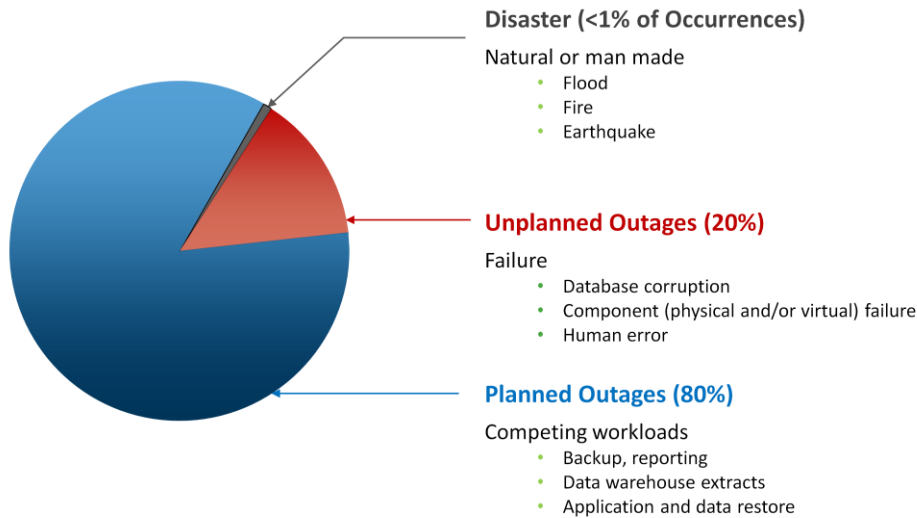
Information availability (IA) refers to the ability of an IT infrastructure to function according to business expectations during its specified time of operation. IA ensures that people (employees, customers, suppliers, and partners) can access information whenever they need it. IA can be defined in terms of the accessibility, reliability, and timeliness of the information.

Accessibility: Information should be accessible to the right user when required.

Reliability: Information should be reliable and correct in all aspects. It is “the same” as what was stored and there is no alternation or corruption to the information.

Timeliness: Defines the time window (a particular time of the day, week, month, and year as specified) during which information must be accessible. For example, if online access to an application is required between 8:00 am and 10:00 pm each day, any disruptions to data availability outside of this time slot are not considered to affect timeliness.

Causes of Information Unavailability



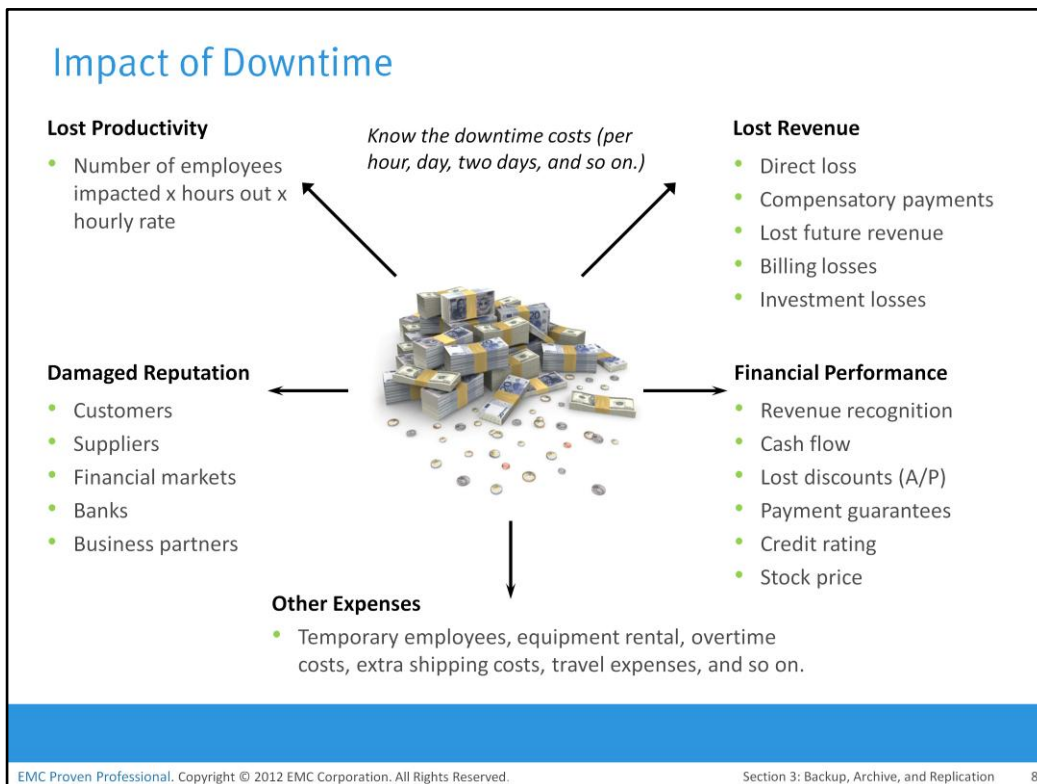
EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication

7

Various planned and unplanned incidents result in information unavailability. *Planned outages* include installation/integration/maintenance of new hardware, software upgrades or patches, taking backups, application and data restores, facility operations (renovation and construction), and refresh/migration of the testing to the production environment. *Unplanned outages* include failure caused by human errors, database corruption, and failure of physical and virtual components.

Another type of incident that may cause data unavailability is natural or man-made disasters, such as flood, fire, earthquake, and so on. As illustrated in figure on the slide, the majority of outages are planned. Planned outages are expected and scheduled but still cause data to be unavailable. Statistically, the cause of information unavailability due to unforeseen disasters is less than 1 percent.



Information unavailability or downtime results in loss of productivity, loss of revenue, poor financial performance, and damages to reputation. Loss of productivity include reduced output per unit of labor, equipment, and capital. Loss of revenue includes direct loss, compensatory payments, future revenue loss, billing loss, and investment loss. Poor financial performance affects revenue recognition, cash flow, discounts, payment guarantees, credit rating, and stock price. Damages to reputations may result in a loss of confidence or credibility with customers, suppliers, financial markets, banks, and business partners. Other possible consequences of downtime include the cost of additional equipment rental, overtime, and extra shipping.

The business impact of downtime is the sum of all losses sustained as a result of a given disruption. An important metric, *average cost of downtime per hour*, provides a key estimate in determining the appropriate BC solutions. It is calculated as follows:

Average cost of downtime per hour = average productivity loss per hour + average revenue loss per hour

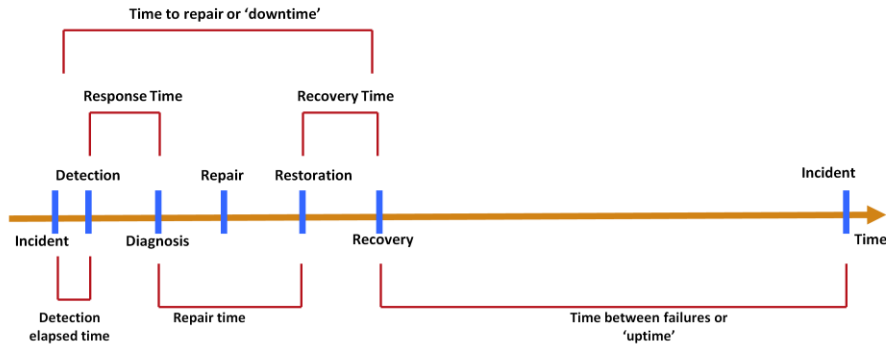
Where:

Productivity loss per hour = (total salaries and benefits of all employees per week) / (average number of working hours per week)

Average revenue loss per hour = (total revenue of an organization per week) / (average number of hours per week that an organization is open for business)

The average downtime cost per hour may also include estimates of projected revenue loss due to other consequences, such as damaged reputations, and the additional cost of repairing the system.

Measuring Information Availability



- MTBF: Average time available for a system or component to perform its normal operations between failures

$$MTBF = \text{Total uptime} / \text{Number of failures}$$

- MTTR: Average time required to repair a failed component

$$MTTR = \text{Total downtime} / \text{Number of failures}$$

$$IA = MTBF / (MTBF + MTTR) \text{ or } IA = \text{uptime} / (\text{uptime} + \text{downtime})$$

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication

9

Information availability relies on the availability of both physical and virtual components of a data center. Failure of these components might disrupt information availability. A failure is the termination of a component's ability to perform a required function. The component's ability can be restored by performing an external corrective actions, such as a manual reboot, a repair, or replacement of the failed component(s). Proactive risk analysis, performed as part of the BC planning process, considers the component failure rate and average repair time, which are measured by MTBF and MTTR:

Mean Time Between Failure (MTBF): It is the average time available for a system or component to perform its normal operations between failures. It is the measure of system or component reliability and is usually expressed in hours.

Mean Time To Repair (MTTR): It is the average time required to repair a failed component. MTTR includes the total time required to do the following activities: detect the fault, mobilize the maintenance team, diagnose the fault, obtain the spare parts, repair, test, and restore the data. MTTR is calculated as: $\text{Total downtime} / \text{Number of failures}$

IA can be expressed in terms of system uptime and downtime and measured as the amount or percentage of system uptime:

$$IA = \text{system uptime} / (\text{system uptime} + \text{system downtime})$$

Where *system uptime* is the period of time during which the system is in an accessible state; when it is not accessible, it is termed as *system downtime*.

In terms of MTBF and MTTR, IA could also be expressed as: $IA = MTBF / (MTBF + MTTR)$

Availability Measurement – Levels of ‘9s’ Availability

Uptime (%)	Downtime (%)	Downtime per Year	Downtime per Week
98	2	7.3 days	3hrs, 22 minutes
99	1	3.65 days	1 hr, 41 minutes
99.8	0.2	17 hrs, 31 minutes	20 minutes, 10 secs
99.9	0.1	8 hrs, 45 minutes	10 minutes, 5 secs
99.99	0.01	52.5 minutes	1 minute
99.999	0.001	5.25 minutes	6 secs
99.9999	0.0001	31.5 secs	0.6 secs

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 10

Uptime per year is based on the exact timeliness requirements of the service. This calculation leads to the number of “9s” representation for availability metrics. Table on the slide lists the approximate amount of downtime allowed for a service to achieve certain levels of 9s availability.

For example, a service that is said to be “five 9s available” is available for 99.999 percent of the scheduled time in a year (24×365).

Module 9: Introduction to Business Continuity

Lesson 2: BC Planning and Technology Solutions

During this lesson the following topics are covered:

- BC terminologies
- BC planning
- Business impact analysis
- Single points of failure
- Multipathing software

This lesson covers various BC terminologies and BC planning. This lesson also focuses on eliminating single points of failure and multipathing software.

BC Terminologies – 1

- Disaster recovery
 - ▶ Coordinated process of restoring systems, data, and infrastructure required to support business operations after a disaster occurs
 - ▶ Restoring previous copy of data and applying logs to that copy to bring it to a known point of consistency
 - ▶ Generally implies use of backup technology
- Disaster restart
 - ▶ Process of restarting business operations with mirrored consistent copies of data and applications
 - ▶ Generally implies use of replication technologies

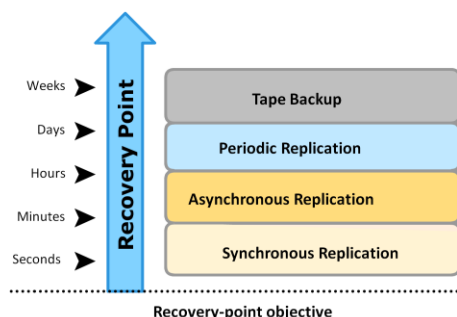
Disaster recovery: This is the coordinated process of restoring systems, data, and the infrastructure required to support ongoing business operations after a disaster occurs. It is the process of restoring a previous copy of the data and applying logs or other necessary processes to that copy to bring it to a known point of consistency. After all recovery efforts are completed, the data is validated to ensure that it is correct.

Disaster restart: This is the process of restarting business operations with mirrored consistent copies of data and applications.

BC Terminologies – 2

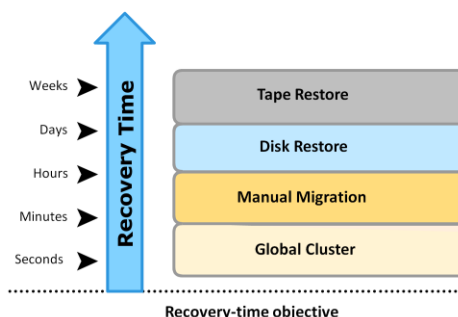
Recovery-Point Objective (RPO)

- Point-in-time to which systems and data must be recovered after an outage
- Amount of data loss that a business can endure



Recovery-Time Objective (RTO)

- Time within which systems and applications must be recovered after an outage
- Amount of downtime that a business can endure and survive



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 13

Recovery-Point Objective (RPO): This is the point-in-time to which systems and data must be recovered after an outage. It defines the amount of data loss that a business can endure. Based on the RPO, organizations plan for the frequency with which a backup or replica must be made. An organization can plan for an appropriate BC technology solution on the basis of the RPO it sets. For example:

RPO of 24 hours: Backups are created at an offsite tape library every midnight. The corresponding recovery strategy is to restore data from the set of last backup tapes.

RPO of 1 hour: Shipping database logs to the remote site every hour. The corresponding recovery strategy is to recover the database to the point of the last log shipment.

RPO in the order of minutes: Mirroring data asynchronously to a remote site.

RPO of zero: Mirroring data synchronously to a remote site.

Recovery-Time Objective (RTO): The time within which systems and applications must be recovered after an outage. It defines the amount of downtime that a business can endure and survive. Some examples of RTOs and the recovery strategies to ensure data availability are listed below:

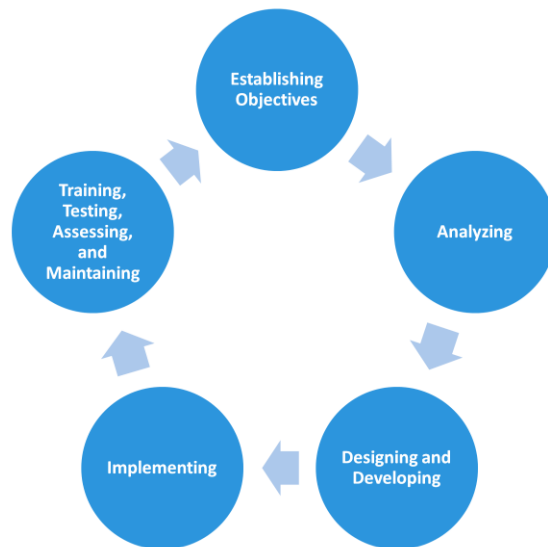
RTO of 72 hours: Restore from tapes available at a cold site.

RTO of 12 hours: Restore from tapes available at a hot site.

RTO of few hours: Use disk-based backup technology, which gives faster restore than a tape backup.

RTO of a few seconds: Cluster production servers with bidirectional mirroring, enabling the applications to run at both sites simultaneously.

BC Planning Lifecycle



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 14

BC planning must follow a disciplined approach like any other planning process. Organizations today dedicate specialized resources to develop and maintain BC plans. From the conceptualization to the realization of the BC plan, a lifecycle of activities can be defined for the BC process. The BC planning lifecycle includes five stages:

1. Establishing objectives

- Determine BC requirements.
- Estimate the scope and budget to achieve requirements.
- Select a BC team that includes subject matter experts from all areas of the business, whether internal or external.
- Create BC policies.

2. Analyzing

- Collect information on data profiles, business processes, infrastructure support, dependencies, and frequency of using business infrastructure.
- Conduct a business impact analysis.
- Identify critical business processes and assign recovery priorities.
- Perform risk analysis for critical functions and create mitigation strategies.
- Perform cost benefit analysis for available solutions based on the mitigation strategy.
- Evaluate options.

Cont..

3. Designing and developing

- Define the team structure and assign individual roles and responsibilities. For example, different teams are formed for activities, such as emergency response, and infrastructure and application recovery.
- Design data protection strategies and develop infrastructure.
- Develop contingency solution and emergency response procedures.
- Detail recovery and restart procedures.

4. Implementing

- Implement risk management and mitigation procedures that include backup, replication, and management of resources.
- Prepare the DR sites that can be utilized if a disaster affects the primary data center.
- Implement redundancy for every resource in a data center to avoid single points of failure.

5. Training, testing, assessing, and maintaining

- Train the employees who are responsible for backup and replication of business-critical data on a regular basis or whenever there is a modification in the BC plan.
- Train employees on emergency response procedures when disasters are declared.
- Train the recovery team on recovery procedures based on contingency scenarios.
- Perform damage-assessment processes and review recovery plans.
- Test the BC plan regularly to evaluate its performance and identify its limitations.
- Assess the performance reports and identify limitations.
- Update the BC plans and recovery/restart procedures to reflect regular changes within the data center.

Business Impact Analysis

- Identifies which business units and processes are essential to the survival of the business
- Estimates the cost of failure for each business process
- Calculates the maximum tolerable outage and defines RTO for each business process
- Businesses can prioritize and implement countermeasures to mitigate the likelihood of such disruptions

A *business impact analysis* (BIA) identifies which business units, operations, and processes are essential to the survival of the business. It evaluates the financial, operational, and service impacts of a disruption to essential business processes. Selected functional areas are evaluated to determine resilience of the infrastructure to support information availability. The BIA process leads to a report detailing the incidents and their impact over business functions. The impact may be specified in terms of money or in terms of time. Based on the potential impacts associated with downtime, businesses can prioritize and implement countermeasures to mitigate the likelihood of such disruptions. These are detailed in the BC plan. A BIA includes the following set of tasks:

- Determine the business areas.
- For each business area, identify the key business processes critical to its operation.
- Determine the attributes of the business process in terms of applications, databases, and hardware and software requirements.
- Estimate the costs of failure for each business process.
- Calculate the maximum tolerable outage and define RTO for each business process.
- Establish the minimum resources required for the operation of business processes.
- Determine recovery strategies and the cost for implementing them.
- Optimize the backup and business recovery strategy based on business priorities.
- Analyze the current state of BC readiness and optimize future BC planning.

BC Technology Solutions

- Solutions that enable BC are:
 - ▶ Resolving single points of failure
 - ▶ Multipathing software
 - ▶ Backup and replication
 - ▶▶ Backup
 - ▶▶ Local replication
 - ▶▶ Remote replication

After analyzing the business impact of an outage, designing the appropriate solutions to recover from a failure is the next important activity. Following are the solutions and supporting technologies that enable business continuity and uninterrupted data availability:

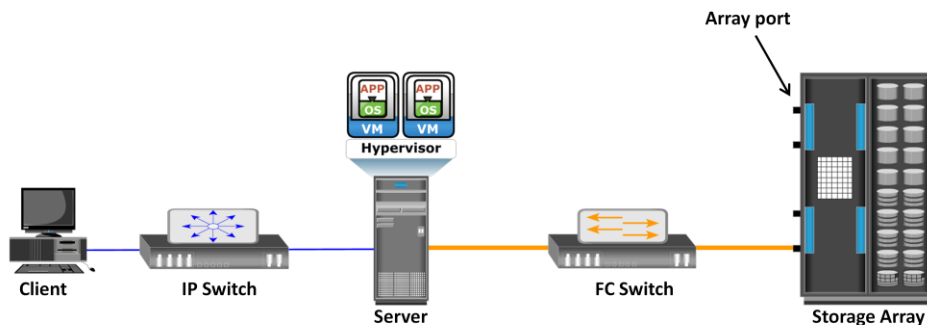
- Resolving single points of failure
- Multipathing software
- Backup and replication

Note: Backup and Replication will be discussed in forthcoming modules.

Single Points of Failure

Single Points of Failure

It refers to the failure of a component of a system that can terminate the availability of the entire system or IT service.

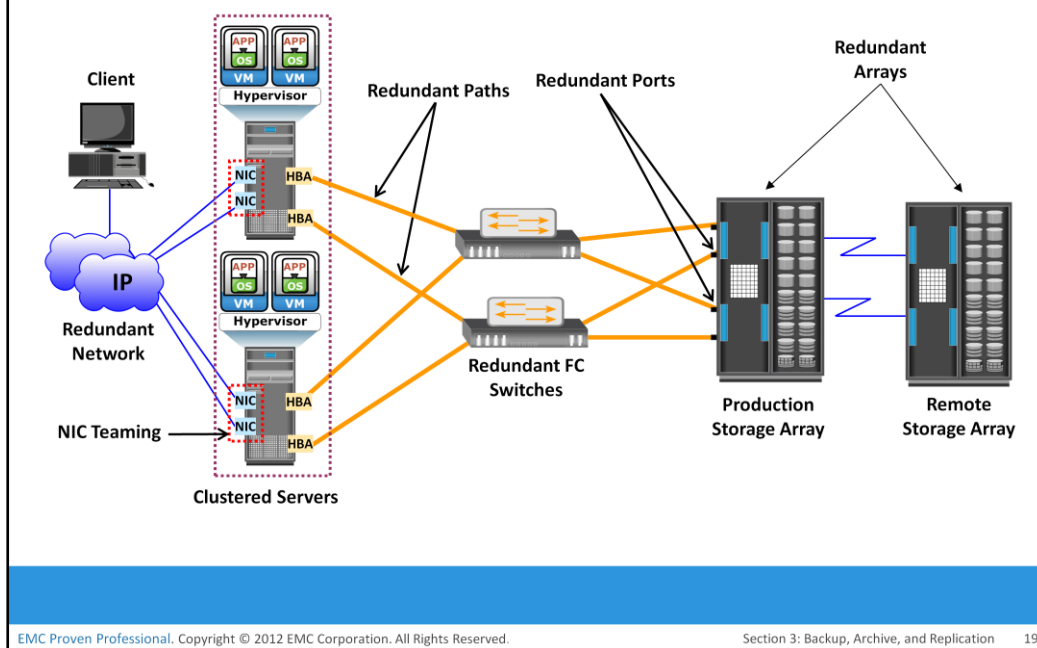


EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 18

A *single point of failure* refers to the failure of a component that can terminate the availability of the entire system or IT service. The figure depicts a system setup in which an application, running on a VM, provides an interface to the client and performs I/O operations. The client is connected to the server through an IP network, and the server is connected to the storage array through an FC connection. In this setup, each component must function as required to ensure data availability, the failure of a single physical or virtual component causes the unavailability of an application. This failure results in disruption of business operations. For example, failure of a hypervisor can affect all the running VMs and virtual network, which are hosted on it. In the figure on the slide, several single points of failure can be identified. A VM, a hypervisor, or an HBA/NIC on the server, the physical server itself, the IP network, the FC switch, the storage array port, or even the storage array could be a potential single point of failure.

Resolving Single Points of Failure



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 19

To mitigate single points of failure, systems are designed with redundancy, such that the system fails only if all the components in the redundancy group fail. This ensures that the failure of a single component does not affect data availability. Data centers follow stringent guidelines to implement fault tolerance for uninterrupted information availability. Careful analysis is performed to eliminate every single point of failure. The example shown in figure on the slide represents all enhancements in the infrastructure to mitigate single points of failure:

- Configuration of redundant HBAs at a server to mitigate single HBA failure.
- Configuration of NIC teaming at a server allows protection against single physical NIC failure. It allows grouping of two or more physical NICs and treating them as a single logical device. With NIC teaming, if one of the underlying physical NICs fails or its cable is unplugged, the traffic is redirected to another physical NIC in the team. Thus, NIC teaming eliminates the single point of failure associated with a single physical NIC.
- Configuration of redundant switches to account for a switch failure.
- Configuration of multiple storage array ports to mitigate a port failure.
- RAID and hot spare configuration to ensure continuous operation in the event of disk failure.

Cont..

- Implementation of a redundant storage array at a remote site to mitigate local site failure.
- Implementing server (or compute) clustering, a fault-tolerance mechanism whereby two or more servers in a cluster access the same set of data volumes. Clustered servers exchange *heartbeat* to inform each other about their health. If one of the servers or hypervisors fails, the other server or hypervisor can take up the workload.
- Implementing a VM Fault Tolerance mechanism ensures BC in the event of a server failure. This technique creates duplicate copies of each VM on another server so that when a VM failure is detected, the duplicate VM can be used for failover. The two VMs are kept in synchronization with each other in order to perform successful failover.

Multipathing Software

- Recognizes and utilizes alternate I/O path to data
- Provides load balancing by distributing I/Os to all available, active paths:
 - ▶ Improves I/O performance and data path utilization
- Intelligently manages the paths to a device by sending I/O down the optimal path:
 - ▶ Based on the load balancing and failover policy setting for the device

Configuration of multiple paths increases the data availability through path failover. If servers are configured with one I/O path to the data, there will be no access to the data if that path fails. Redundant paths to the data eliminate the possibility of the path becoming a single point of failure. Multiple paths to data also improve I/O performance through load balancing among the paths and maximize server, storage, and data path utilization.

In practice, merely configuring multiple paths does not serve the purpose. Even with multiple paths, if one path fails, I/O does not reroute unless the system recognizes that it has an alternative path. Multipathing software provides the functionality to recognize and utilize alternative I/O paths to data. Multipathing software also manages the load balancing by distributing I/Os to all available, active paths.

Multipathing software intelligently manages the paths to a device by sending I/O down the optimal path based on the load balancing and failover policy setting for the device. It also takes into account path usage and availability before deciding the path through which to send the I/O. If a path to the device fails, it automatically reroutes the I/O to an alternative path.

In a virtual environment, multipathing is enabled either by using the hypervisor's built-in capability or by running a third-party software module, added to the hypervisor.

Module 9: Introduction to Business Continuity

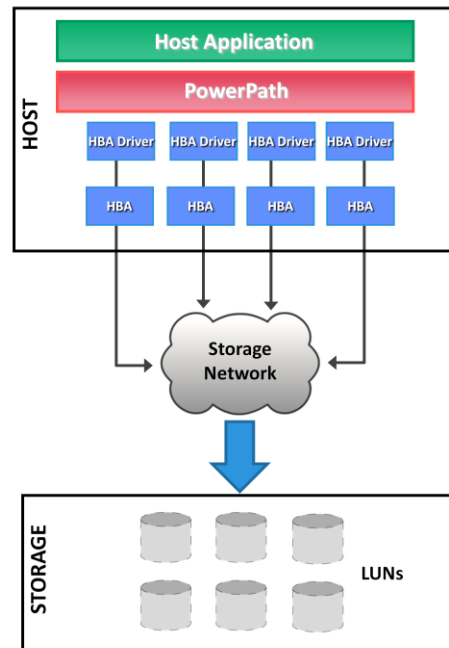
Concept in Practice

- EMC PowerPath

The concept in practice section covers EMC PowerPath.

EMC PowerPath

- Host-based multipathing software
- Provides path failover and load-balancing functionality
- Automatic detection and recovery from host-to-array path failures
- PowerPath/VE software allows optimizing virtual environments with PowerPath multipathing features



EMC PowerPath is host-based multipathing software. Every I/O from the host to the array must pass through the PowerPath software, which allows PowerPath to provide intelligent I/O path management. PowerPath provides path failover and dynamic load balancing. PowerPath/VE software allows optimizing virtual environments with PowerPath multipathing features.

Module 9: Summary

Key points covered in this module:

- Importance of business continuity
- Impact of information unavailability
- Information availability metrics
- Business impact analysis
- Single points of failure
- Multipathing software

This module covered the importance of business continuity, impact of information unavailability, and information availability metrics. This module also focused on business continuity planning and business impact analysis. Further, this module detailed on single points of failure and multipathing software.

BC entails preparing for, responding to, and recovering from a system outage that adversely affects business operations. Information unavailability or downtime results in loss of productivity, loss of revenue, poor financial performance, and damages to reputation.

Information availability metrics are MTBF and MTTR. MTBF defines average time available for a system or component to perform its normal operations between failures. MTTR defines the average time required to repair a failed component. A business impact analysis identifies which business units, operations, and processes are essential to the survival of the business.

A single point of failure refers to the failure of a component that can terminate the availability of the entire system or IT service. Multipathing software provides the functionality to recognize and utilize alternate I/O paths to data. Multipathing software also manages the load balancing by distributing I/Os to all available, active paths.

Check Your Knowledge – 1

- Which statement is true if the recovery-point objective (RPO) of an application is 2 hours?
 - A. Time to resume application operations must be less than 2 hours
 - B. Time to resume application operations must equal to 2 hours
 - C. No more than 2 hours of production data can be lost
 - D. Mean time between application failure is 2 hours
- Which allows grouping of two or more physical NICs and treating them as a single logical device?
 - A. NIC streaming
 - B. NIC porting
 - C. NIC teaming
 - D. NIC zoning

Check Your Knowledge – 2

- Which best describes recovery-time objective (RTO)?
 - A. Point-in-time to which data must be recovered after an outage
 - B. Time available for a system or component to perform its normal operations between failures
 - C. Time within which systems and applications must be recovered after an outage
 - D. Amount of data loss that a business can endure
- Which expression represents availability of a system in terms of MTBF and MTTR?
 - A. $MTTR / (MTBF \times MTTR)$
 - B. $MTBF / (MTBF \times MTTR)$
 - C. $MTTR / (MTBF + MTTR)$
 - D. $MTBF / (MTTR + MTBF)$

Check Your Knowledge – 3

- A department requires access to the database application from Monday to Friday, 9 AM to 5 PM. Last Thursday at 1 PM the application crashed and it took six hours to fix the problem. What was the availability of the application during last week?
 - A. 85%
 - B. 90%
 - C. 95%
 - D. 100%

Exercise 1: MTBF and MTTR

- A system has three components and requires all three to be operational for 24 hours from Monday to Friday. Failure of component 1 occurs as follows:

- ▶ Monday = No failure
- ▶ Tuesday = 5 am to 7 am
- ▶ Wednesday = No failure
- ▶ Thursday = 4 pm to 8 pm
- ▶ Friday = 8 am to 11 am

Calculate the MTBF and MTTR of component 1.

Exercise 2: Information Availability

- A system has three components and requires all three to be operational from 8 am to 5 pm business hours, Monday to Friday. Failure of component 2 occurs as follows:

- ▶ Monday = 8 am to 11 am
- ▶ Tuesday = No failure
- ▶ Wednesday = 4 pm to 7 pm
- ▶ Thursday = 5 pm to 8 pm
- ▶ Friday = 1 pm to 2 pm

Calculate the availability of component 2.

Module – 10

Backup and Archive



Module 10: Backup and Archive

Upon completion of this module, you should be able to:

- Describe backup granularities
- Explain backup and recovery operations
- Describe various backup targets
- Explain data deduplication
- Describe backup in virtualized environment
- Explain data archive

This module focuses on backup granularities and backup operations. This module also focuses on various backup targets and data deduplication. Further, this module details backup in virtualized environment. Additionally, this module focuses on data archive.

Module 10: Backup and Archive

Lesson 1: Backup Overview

During this lesson the following topics are covered:

- Backup granularity
- Backup method
- Backup architecture
- Backup and recovery operations

This lesson covers various backup granularities and backup method. This lesson also covers backup architecture and operations.

What is Backup?

Backup

It is an additional copy of production data that is created and retained for the sole purpose of recovering lost or corrupted data.

- Organization also takes backup to comply with regulatory requirements
- Backups are performed to serve three purposes:
 - ▶ Disaster recovery
 - ▶ Operational recovery
 - ▶ Archive

A *backup* is an additional copy of production data, created and retained for the sole purpose of recovering lost or corrupted data. With growing business and regulatory demands for data storage, retention, and availability, organizations are faced with the task of backing up an ever-increasing amount of data. This task becomes more challenging with the growth of information, stagnant IT budgets, and less time for taking backups. Moreover, organizations need a quick restore of backed up data to meet business service-level agreements (SLAs).

Backups are performed to serve three purposes: disaster recovery, operational recovery, and archival.

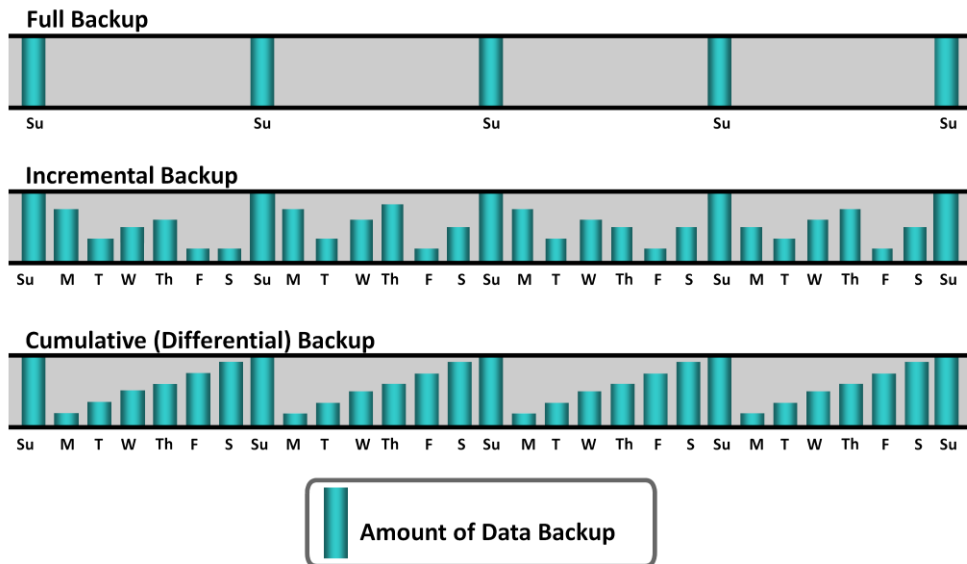
Backups can be performed to address disaster recovery needs. The backup copies are used for restoring data at an alternate site when the primary site is incapacitated due to a disaster. Based on Recovery-point objective (RPO) and Recovery-time objective (RTO) requirements, organizations use different data protection strategies for disaster recovery.

Data in the production environment changes with every business transaction and operation. Backups are used to restore data if data loss or logical corruption occurs during routine processing. The majority of restore requests in most organizations fall in this category. For example, it is common for a user to accidentally delete an important e-mail or for a file to become corrupted, which can be restored using backup data.

Backups are also performed to address archival requirements. Although content addressed storage (CAS) has emerged as the primary solution for archives (CAS is discussed in module 8), traditional backups are still used by small and medium enterprises for long-term preservation of transaction records, e-mail messages, and other business records required for regulatory compliance.

Note: Backup window is the period during which a source is available for performing a data backup.

Backup Granularity



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

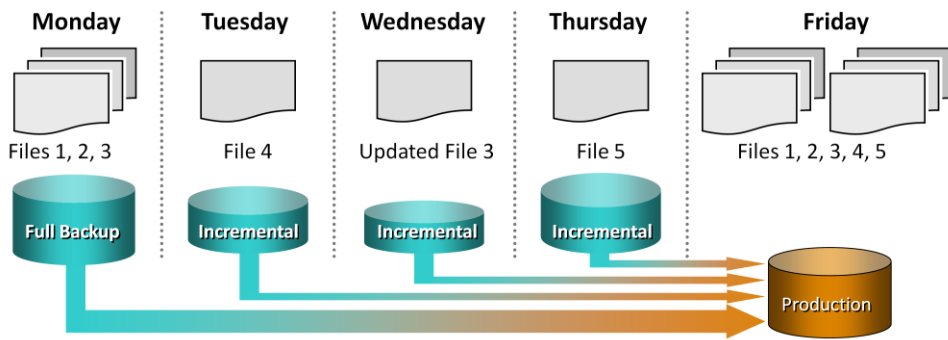
Section 3: Backup, Archive, and Replication 34

Backup granularity depends on business needs and the required RTO/RPO. Based on the granularity, backups can be categorized as full, incremental, and cumulative (or differential). Most organizations use a combination of these three backup types to meet their backup and recovery requirements. Figure on the slide depicts the different backup granularity levels.

Full backup is a backup of the complete data on the production volumes. A full backup copy is created by copying the data in the production volumes to a backup storage device. It provides a faster recovery but requires more storage space and also takes more time to back up. *Incremental backup* copies the data that has changed since the last full or incremental backup, whichever has occurred more recently. This is much faster than a full backup (because the volume of data backed up is restricted to the changed data only) but takes longer to restore. *Cumulative backup* copies the data that has changed since the last full backup. This method takes longer than an incremental backup but is faster to restore.

Another way to implement full backup is *synthetic (or constructed) backup*. This method is used when the production volume resources cannot be exclusively reserved for a backup process for extended periods to perform a full backup. It is usually created from the most recent full backup and all the incremental backups performed after that full backup. This backup is called *synthetic* because the backup is not created directly from production data. A synthetic full backup enables a full backup copy to be created offline without disrupting the I/O operation on the production volume. This also frees up network resources from the backup process, making them available for other production uses.

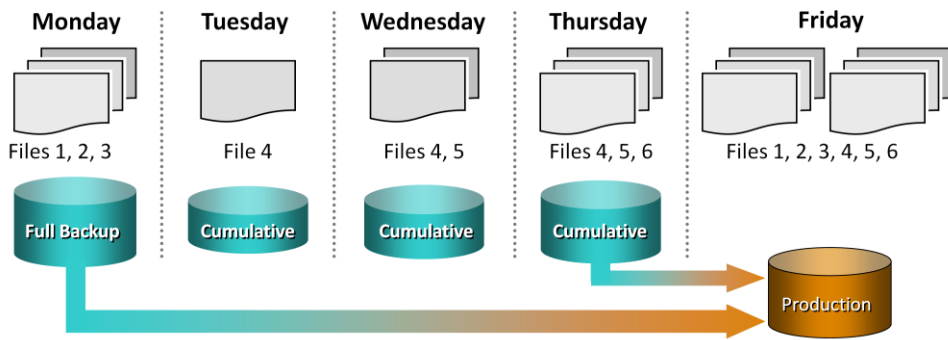
Restoring from Incremental Backup



- Less number of files to be backed up, therefore, it takes less time to backup and requires less storage space
- Longer restore because last full and all subsequent incremental backups must be applied

The process of restoration from an incremental backup requires the last full backup and all the incremental backups available until the point of restoration. Consider an example, a full backup is performed on Monday evening. Each day after that, an incremental backup is performed. On Tuesday, a new file (File 4 as shown in the figure) is added, and no other files have changed. Consequently, only File 4 is copied during the incremental backup performed on Tuesday evening. On Wednesday, no new files are added, but File 3 has been modified. Therefore, only the modified File 3 is copied during the incremental backup on Wednesday evening. Similarly, the incremental backup on Thursday copies only File 5. On Friday morning, there is data corruption, which requires data restoration from the backup. The first step toward data restoration is restoring all data from the full backup of Monday evening. The next step is applying the incremental backups of Tuesday, Wednesday, and Thursday. In this manner, data can be successfully recovered to its previous state, as it existed on Thursday evening.

Restoring from Cumulative Backup

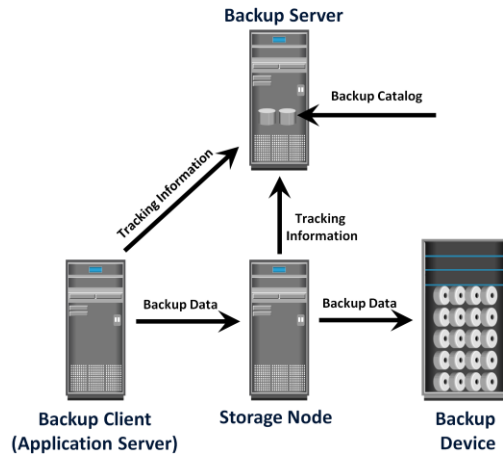


- More files to be backed up, therefore, it takes more time to backup and requires more storage space
- Faster restore because only the last full and the last cumulative backup must be applied

Consider an example, a full backup of the business data is taken on Monday evening. Each day after that, a cumulative backup is taken. On Tuesday, File 4 is added and no other data is modified since the previous full backup of Monday evening. Consequently, the cumulative backup on Tuesday evening copies only File 4. On Wednesday, File 5 is added. The cumulative backup taking place on Wednesday evening copies both File 4 and File 5 because these files have been added or modified since the last full backup. Similarly, on Thursday, File 6 is added. Therefore, the cumulative backup on Thursday evening copies all three files: File 4, File 5, and File 6. On Friday morning, data corruption occurs that requires data restoration using backup copies. The first step in restoring data is to restore all the data from the full backup of Monday evening. The next step is to apply only the latest cumulative backup, which is taken on Thursday evening. In this way, the production data can be recovered faster because it needs only two copies of data—the last full backup and the latest cumulative backup.

Backup Architecture

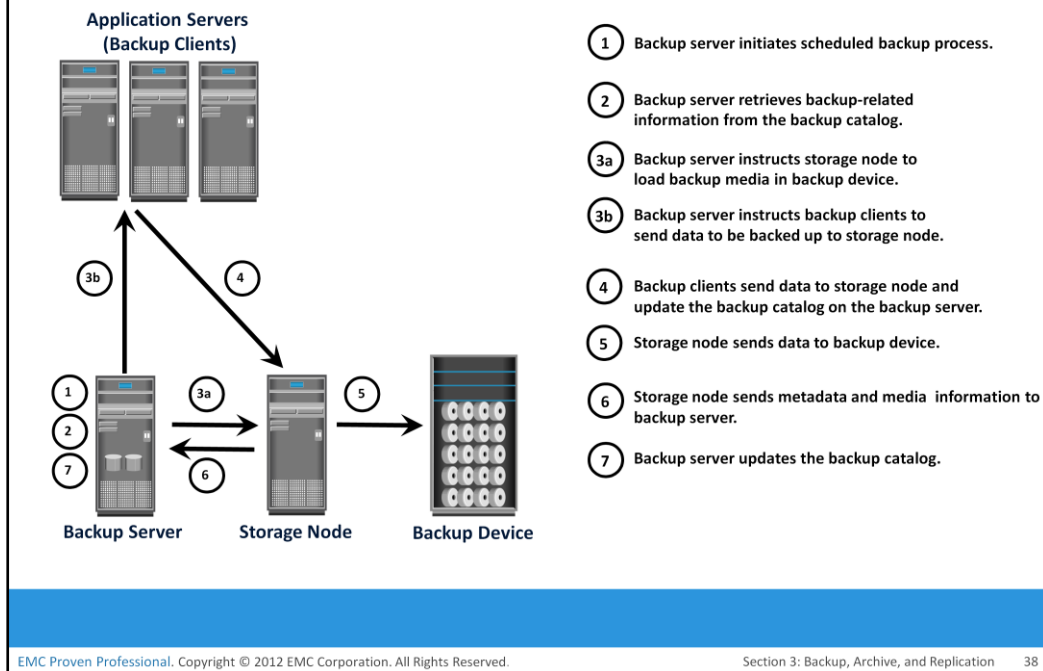
- Backup client
 - ▶ Gathers the data that is to be backed up and send it to storage node
- Backup server
 - ▶ Manages backup operations and maintains backup catalog
- Storage node
 - ▶ Responsible for writing data to backup device
 - ▶ Manages the backup device



A backup system commonly uses the client-server architecture with a backup server and multiple backup clients. Figure on the slide illustrates the backup architecture. The backup server manages the backup operations and maintains the backup catalog, which contains information about the backup configuration and backup metadata. Backup configuration contains information about when to run backups, which client data to be backed up, and so on, and the backup metadata contains information about the backed up data. The role of a backup client is to gather the data that is to be backed up and send it to the storage node. It also sends the tracking information to the backup server.

The storage node is responsible for writing the data, to the backup device. In a backup environment, a *storage node* is a host that controls backup devices. The storage node also sends tracking information to the backup server. In many cases, the storage node is integrated with the backup server, and both are hosted on the same physical platform. A backup device is attached directly or through a network to the storage node's host platform. Some backup architecture refers the storage node as the *media server* because it manages the storage device.

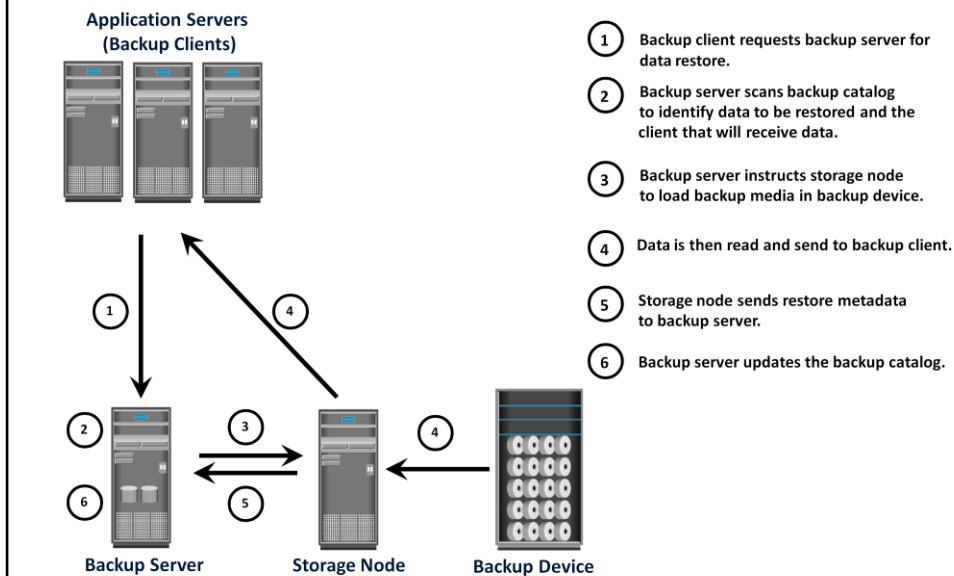
Backup Operation



When a backup operation is initiated, significant network communication takes place between the different components of a backup infrastructure. The backup operation is typically initiated by a server, but it can also be initiated by a client. The backup server initiates the backup process for different clients based on the backup schedule configured for them. For example, the backup for a group of clients may be scheduled to start at 3:00 a.m. every day.

The backup server coordinates the backup process with all the components in a backup environment. The backup server maintains the information about backup clients to be backed up and storage nodes to be used in a backup operation. The backup server retrieves the backup-related information from the backup catalog and, based on this information, instructs the storage node to load the appropriate backup media into the backup devices. Simultaneously, it instructs the backup clients to gather the data to be backed up and send it over the network to the assigned storage node. After the backup data is sent to the storage node, the client sends some backup metadata (the number of files, name of the files, storage node details, and so on) to the backup server. The storage node receives the client data, organizes it, and sends it to the backup device. The storage node then sends additional backup metadata (location of the data on the backup device, time of backup, and so on) to the backup server. The backup server updates the backup catalog with this information.

Recovery Operation



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 39

After the data is backed up, it can be restored when required. A restore process must be manually initiated from the client. Some backup software has a separate application for restore operations. These restore applications are usually accessible only to the administrators or backup operators. Figure on the slide depicts a restore operation.

Upon receiving a restore request, an administrator opens the restore application to view the list of clients that have been backed up. While selecting the client for which a restore request has been made, the administrator also needs to identify the client that will receive the restored data. Data can be restored on the same client for whom the restore request has been made or on any other client. The administrator then selects the data to be restored and the specified point in time to which the data has to be restored based on the RPO. Because all this information comes from the backup catalog, the restore application needs to communicate with the backup server.

The backup server instructs the appropriate storage node to mount the specific backup media onto the backup device. Data is then read and sent to the client that has been identified to receive the restored data.

Some restorations are successfully accomplished by recovering only the requested production data. For example, the recovery process of a spreadsheet is completed when the specific file is restored. In database restorations, additional data, such as log files, must be restored along with the production data. This ensures consistency for the restored data. In these cases, the RTO is extended due to the additional steps in the restore operation.

Backup Methods

- Two methods of backup, based on the state of the application when the backup is performed
 - ▶ Hot or Online
 - ▶▶ Application is up and running, with users accessing their data during backup
 - ▶▶ Open file agent can be used to backup open files
 - ▶ Cold or Offline
 - ▶▶ Requires application to be shutdown during the backup process

Hot backup and cold backup are the two methods deployed for backup. They are based on the state of the application when the backup is performed. In a *hot backup*, the application is up-and-running, with users accessing their data during the backup process. This method of backup is also referred to as online backup. A *cold backup* requires the application to be shutdown during the backup process. Hence, this method is also referred to as offline backup.

The hot backup of online production data is challenging because data is actively being used and changed. If a file is open, it is normally not backed up during the backup process. In such situations, an *open file agent* is required to back up the open file. These agents interact directly with the operating system or application and enable the creation of consistent copies of open files. The disadvantage associated with a hot backup is that the agents usually affect the overall application performance. Consistent backups of databases can also be done by using a cold backup. This requires the database to remain inactive during the backup. Of course, the disadvantage of a cold backup is that the database is inaccessible to users during the backup process. All the files must be backed up in the same state for consistent backup of a database that comprises many files.

Bare-Metal Recovery

- OS, hardware, and application configurations are appropriately backed up for a full system recovery
- Server configuration backup (SCB) can also recover a server onto dissimilar hardware

In a disaster recovery environment, *bare-metal recovery* (BMR) refers to a backup in which OS, hardware, and application configurations are appropriately backed up for a full system recovery. BMR builds the base system, which includes partitioning, the file system layout, the operating system, the applications, and all the relevant configurations. BMR recovers the base system first before starting the recovery of data files. Some BMR technologies—for example server configuration backup (SCB)—can recover a server even onto dissimilar hardware.

Server Configuration Backup

- Creates and backs up server configuration profiles, based on user-defined schedules
 - ▶ Profiles are used to configure the recovery server in case of production server failure
 - ▶ Profiles include OS configurations, network configurations, security configurations, registry settings, application configurations
- Two types of profiles used
 - ▶ Base profile
 - ▶▶ Contains the key elements of the OS required to recover the server
 - ▶ Extended profile
 - ▶▶ Typically larger than base profile and contains all necessary information to rebuild application environment

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 42

Most organizations spend a considerable amount of time and money protecting their application data but give less attention to protecting their server configurations. During disaster recovery, server configurations must be re-created before the application and data are accessible to the user. The process of system recovery involves reinstalling the operating system, applications, and server settings and then recovering the data. During a normal data backup operation, server configurations required for the system restore are not backed up. *Server configuration backup* (SCB) creates and backs up server configuration profiles based on user-defined schedules. The backed up profiles are used to configure the recovery server in case of production-server failure. SCB has the capability to recover a server onto dissimilar hardware.

In a server configuration backup, the process of taking a snapshot of the application server's configuration (both system and application configurations) is known as *profiling*. The profile data includes operating system configurations, network configurations, security configurations, registry settings, application configurations, and so on. Thus, profiling allows recovering the configuration of the failed system to a new server regardless of the underlying hardware.

There are two types of profiles generated in the server configuration backup environment: base profile and extended profile. The base profile contains the key elements of the operating system required to recover the server. The extended profile is typically larger than the base profile and contains all the necessary information to rebuild the application environment.

Key Backup/Restore Considerations

- Customer business needs determine:
 - ▶ What are the restore requirements – RPO & RTO?
 - ▶ Which data needs to be backed up?
 - ▶ How frequently should data be backed up?
 - ▶ How long will it take to backup?
 - ▶ How many copies to create?
 - ▶ How long to retain backup copies?
 - ▶ Location, size, and number of files?

The amount of data loss and downtime that a business can endure in terms of RPO and RTO are the primary considerations in selecting and implementing a specific backup strategy. The RPO determines backup frequency. For example, if an application requires an RPO of 1 day, it would need the data to be backed up at least once every day. Another consideration is the retention period, which defines the duration for which a business needs to retain the backup copies.

The backup media type or backup target is another consideration that is driven by RTO and impacts the data recovery time. Organizations must also consider the granularity of backups. The development of a backup strategy must include a decision about the most appropriate time for performing a backup to minimize any disruption to production operations. The size, number of files and data compression should also be considered because they might affect the backup process. Backing up large-size files (for example, ten 1 MB files) takes less time, compared to backing up an equal amount of data composed of small-size files (for example, ten thousand 1 KB files). Data compression and data deduplication (discussed later in the module) are widely used in the backup environment because these technologies save space on the media.

Location is an important consideration for the data to be backed up. Many organizations have dozens of heterogeneous platforms locally and remotely supporting their business. The backup process must address these sources for transactional and content integrity.

Module 10: Backup and Archive

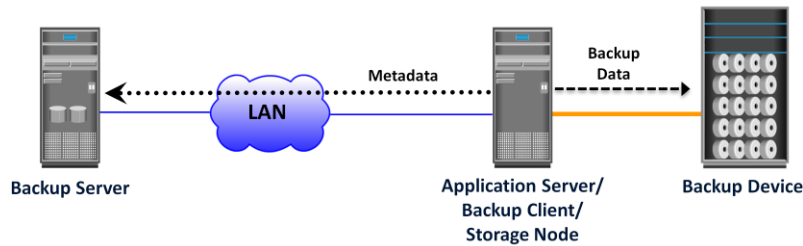
Lesson 2: Backup Topologies and Backup in NAS Environment

During this lesson the following topics are covered:

- Common backup topologies
- Backup in NAS environment

This lesson covers various backup topologies such as Direct-attached, LAN-based, SAN-based and mixed backup. This lesson also covers backup in NAS environment.

Direct-Attached Backup

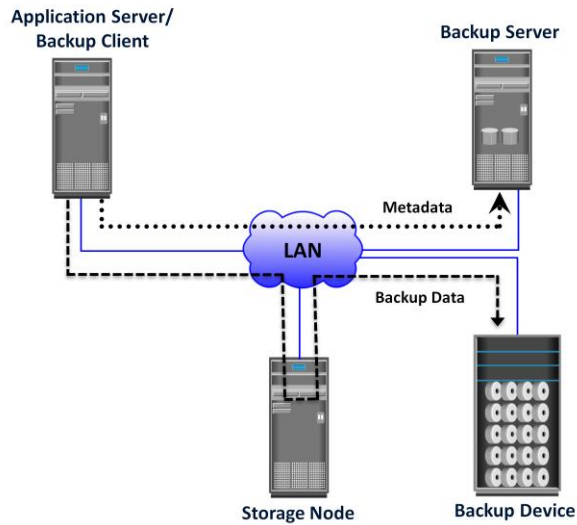


EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 45

In a *direct-attached backup*, the storage node is configured on a backup client, and the backup device is attached directly to the client. Only the metadata is sent to the backup server through the LAN. This configuration frees the LAN from backup traffic. As the environment grows, there will be a need for centralized management and sharing of backup devices to optimize costs. An appropriate solution is required to share the backup devices among multiple servers. Network-based topologies (LAN-based and SAN-based) provide the solution to optimize the utilization of backup devices.

LAN-based Backup



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 46

In a *LAN-based backup*, the clients, backup server, storage node, and backup device are connected to the LAN. The data to be backed up is transferred from the backup client (source) to the backup device (destination) over the LAN, which might affect network performance. This impact can be minimized by adopting a number of measures, such as configuring separate networks for backup and installing dedicated storage nodes for some application servers.

SAN-based Backup

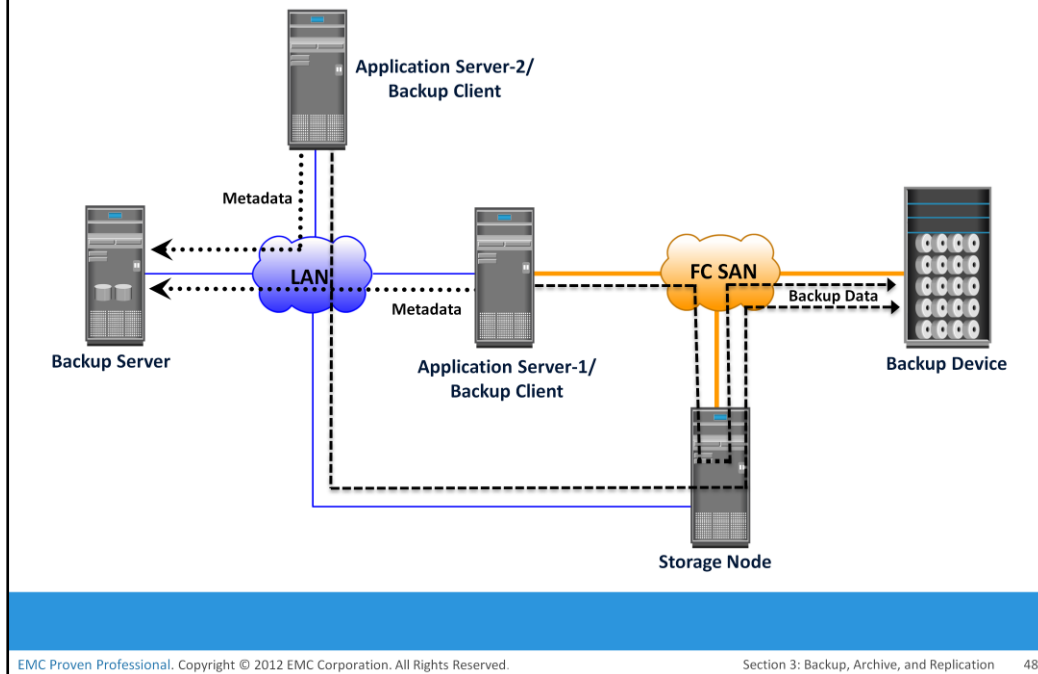
The diagram illustrates a SAN-based backup architecture. It features four main components: a Backup Server, an Application Server/Backup Client, a Storage Node, and a Backup Device. The Backup Server and Application Server/Backup Client are connected via a LAN. The Application Server/Backup Client and Storage Node are connected via an FC SAN. The Storage Node and Backup Device are also connected via an FC SAN. Data flow is indicated by arrows: Metadata is sent from the Application Server/Backup Client to the Backup Server via the LAN; Backup Data is sent from the Application Server/Backup Client to the Backup Device via the FC SAN. The Storage Node is connected to the FC SAN, which in turn connects to the Backup Device.

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 47

The emergence of low-cost disks as a backup medium has enabled disk arrays to be attached to the SAN and used as backup devices. A tape backup of these data backups on the disks can be created and shipped offsite for disaster recovery and long-term retention.

Mixed Backup Topology

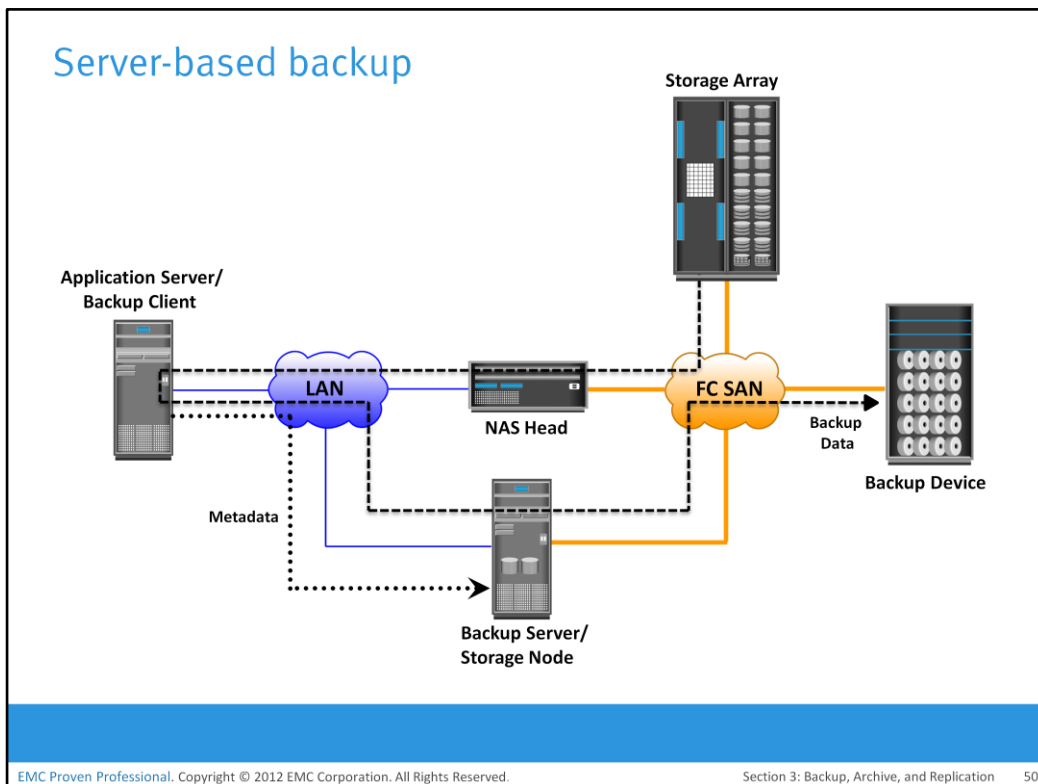


The *mixed topology* uses both the LAN-based and SAN-based topologies. This topology might be implemented for several reasons, including cost, server location, reduction in administrative overhead, and performance considerations.

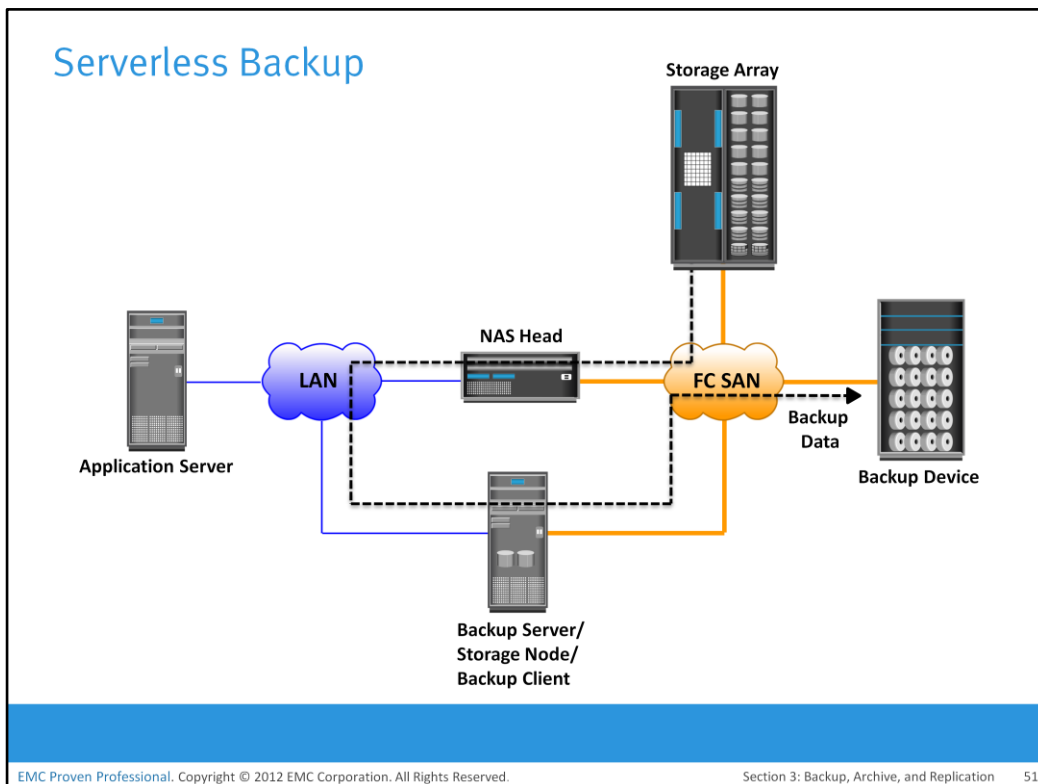
Backup in NAS Environment

- Common backup implementations in a NAS environment are:
 - ▶ Server-based backup
 - ▶ Serverless backup
 - ▶ NDMP 2-way backup
 - ▶ NDMP 3-way backup

The use of a NAS head imposes a new set of considerations on the backup and recovery strategy in NAS environments. NAS heads use a proprietary operating system and file system structure that supports multiple file-sharing protocols. In the NAS environment, backups can be implemented in different ways: server-based, serverless, or using Network Data Management Protocol (NDMP). Common implementations are NDMP 2-way and NDMP 3-way.

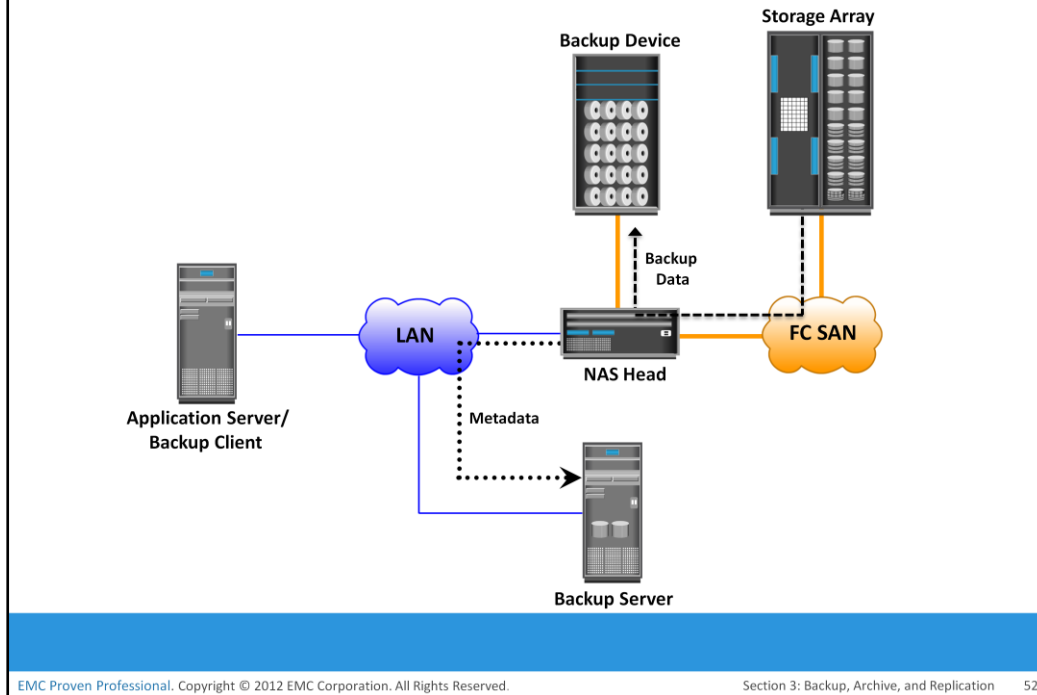


In an *application server-based backup*, the NAS head retrieves data from a storage array over the network and transfers it to the backup client running on the application server. The backup client sends this data to the storage node, which in turn writes the data to the backup device. This results in overloading the network with the backup data and using application server resources to move the backup data.



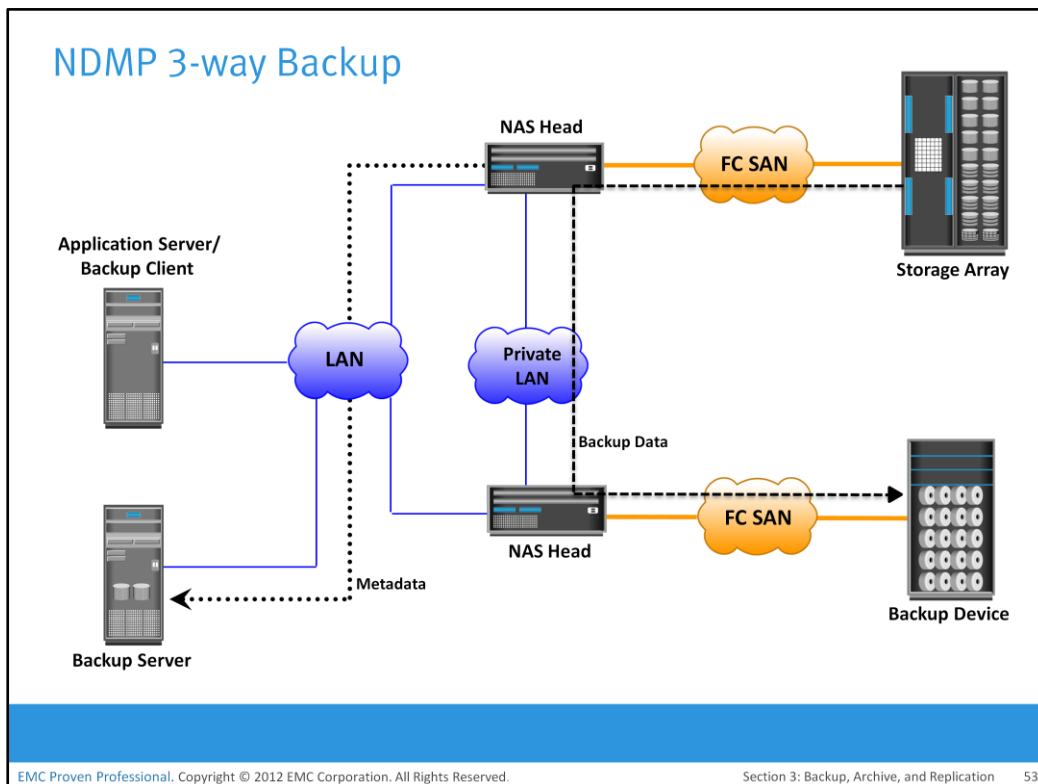
In a *serverless backup*, the network share is mounted directly on the storage node. This avoids overloading the network during the backup process and eliminates the need to use resources on the application server. In this scenario, the storage node, which is also a backup client, reads the data from the NAS head and writes it to the backup device without involving the application server. Compared to the previous solution, this eliminates one network hop.

NDMP 2-way Backup



NDMP is an industry-standard TCP/IP-based protocol specifically designed for a backup in a NAS environment. It communicates with several elements in the backup environment (NAS head, backup devices, backup server, and so on) for data transfer and enables vendors to use a common protocol for the backup architecture. Data can be backed up using *NDMP* regardless of the operating system or platform. Due to its flexibility, it is no longer necessary to transport data through the application server, which reduces the load on the application server and improves the backup speed. *NDMP* optimizes backup and restore by leveraging the high-speed connection between the backup devices and the NAS head. In *NDMP*, backup data is sent directly from the NAS head to the backup device, whereas metadata is sent to the backup server.

Figure on the slide illustrates backup in the NAS environment using *NDMP* 2-way. In this model, network traffic is minimized by isolating data movement from the NAS head to the locally attached backup device. Only metadata is transported on the network. The backup device is dedicated to the NAS device, and hence, this method does not support centralized management of all backup devices.



In the *NDMP 3-way* method, to avoid the backup data traveling on the production LAN, a separate private backup network must be established between all NAS heads and the NAS head connected to the backup device. Metadata and NDMP control data are still transferred across the public network. Figure on the slide depicts NDMP 3-way backup. An NDMP 3-way is useful when backup devices need to be shared among NAS heads. It enables the NAS head to control the backup device and share it with other NAS heads by receiving the backup data through the NDMP.

Module 10: Backup and Archive

Lesson 3: Backup Targets

During this lesson the following topics are covered:

- Backup to Tape
- Backup to Disk
- Backup to Virtual Tape

This lesson covers various backup targets such as physical tape, disk and virtual tape.

Backup to Tape

- Traditionally low cost solution
- Tape drives are used to read/write data from/to a tape
- Sequential/linear access
- Multiple streaming to improve media performance
 - ▶ Writes data from multiple streams on a single tape

Tapes, a low-cost solution, are used extensively for backup. Tape drives are used to read/write data from/to a tape cartridge (or cassette). Tape drives are referred to as sequential, or linear, access devices because the data is written or read sequentially. A tape cartridge is composed of magnetic tapes in a plastic enclosure. Tape Mounting is the process of inserting a tape cartridge into a tape drive. The tape drive has motorized controls to move the magnetic tape around, enabling the head to read or write data.

Tape drive *streaming* or *multiple streaming* writes data from multiple streams on a single tape to keep the drive busy. Multiple streaming improves media performance, but it has an associated disadvantage. The backup data is interleaved because data from multiple streams is written on it. Consequently, the data recovery time is increased because all the extra data from the other streams must be read and discarded while recovering a single stream.

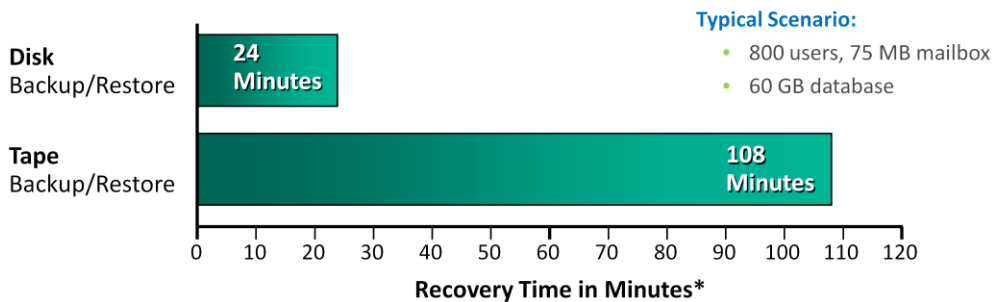
Limitation of Tape

- Backup and recovery operations are slow due to sequential access
- Controlled environment is required for tape storage
- Wear and tear of tape
- Shipping/handling challenges
- Causes “shoe shining effect” or “backhitching”

Data access in a tape is sequential, which can slow backup and recovery operations. Tapes are primarily used for long-term offsite storage because of their low cost. Tapes must be stored in locations with a controlled environment to ensure preservation of the media and to prevent data corruption. Tapes are highly susceptible to wear and tear and usually have shorter shelf life. Physical transportation of the tapes to offsite locations also adds to management overhead and increases the possibility of loss of tapes during offsite shipment. Many times, even the buffering and speed adjustment features of a tape drive fail to prevent the gaps, causing the “*shoe shining effect*” or “*backhitching*.” *Shoe shining* is the repeated back and forth motion a tape drive makes when there is an interruption in the backup data stream. This repeated back-and-forth motion not only causes a degradation of service, but also excessive wear and tear to tapes.

Backup to Disk

- Enhanced overall backup and recovery performance
 - ▶ Random access
- More reliable
- Can be accessed by multiple hosts simultaneously



Source: EMC Engineering and EMC IT

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 57

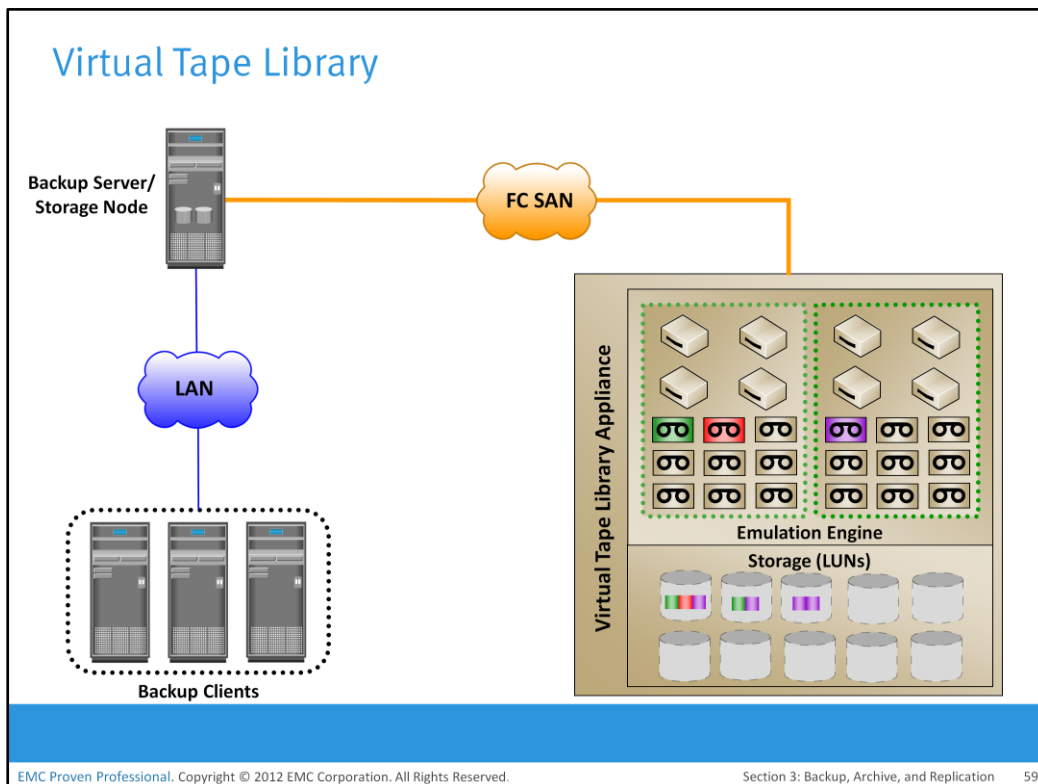
Availability of low cost disks have now replaced tapes as the primary device for storing backup data because of their performance advantages. Backup-to-disk systems offer ease of implementation, reduced TCO, and improved quality of service. Apart from performance benefits in terms of data transfer rates, disks also offer faster recovery when compared to tapes.

Backing up to disk storage systems offers clear advantages due to their inherent random access and RAID-protection capabilities. In most backup environments, backup to disk is used as a staging area where the data is copied temporarily before transferring or staging it to tapes. This enhances backup performance. Some backup products allow for backup images to remain on the disk for a period of time even after they have been staged. This enables a much faster restore. Figure on the slide illustrates a recovery scenario comparing tape versus disk in a Microsoft Exchange environment that supports 800 users with a 75 MB mailbox size and a 60 GB database. As shown in the figure, a restore from the disk took 24 minutes compared to the restore from a tape, which took 108 minutes for the same environment.

Backup to Virtual Tape

- Disks are emulated and presented as tapes to backup software
- Does not require any additional modules or changes in the legacy backup software
- Provides better single stream performance and reliability over physical tape
- Online and random disk access
 - ▶ Provides faster backup and recovery

Virtual tapes are disk drives emulated and presented as tapes to the backup software. The key benefit of using a virtual tape is that it does not require any additional modules, configuration, or changes in the legacy backup software. This preserves the investment made in the backup software. Compared to physical tapes, virtual tapes offer better single stream performance, better reliability, and random disk access characteristics. Backup and restore operations are benefited from the disk's random access characteristics because they are online and provide faster backup and recovery. A virtual tape drive does not require the usual maintenance tasks associated with a physical tape drive, such as periodic cleaning and drive calibration. Compared to backup-to-disk devices, a virtual tape library offers easy installation and administration because it is preconfigured by the manufacturer.



A *virtual tape library* (VTL) has the same components as that of a physical tape library, except that the majority of the components are presented as virtual resources. For the backup software, there is no difference between a physical tape library and a virtual tape library. Figure on the slide shows a virtual tape library. Virtual tape libraries use disks as backup media. Emulation software has a database with a list of virtual tapes, and each virtual tape is assigned space on a LUN. A virtual tape can span multiple LUNs if required. File system awareness is not required while backing up because the virtual tape solution typically uses raw devices.

Similar to a physical tape library, a robot mount is virtually performed when a backup process starts in a virtual tape library. However, unlike a physical tape library, where this process involves some mechanical delays, in a virtual tape library it is almost instantaneous. Even the *load to ready* time is much less than a physical tape library. After the virtual tape is mounted and the virtual tape drive is positioned, the virtual tape is ready to be used, and backup data can be written to it. In most cases, data is written to the virtual tape immediately. Unlike a physical tape library, the virtual tape library is not constrained by the sequential access and shoe shining effect. When the operation is complete, the backup software issues a rewind command. This rewind is also instantaneous. The virtual tape is then unmounted, and the virtual robotic arm is instructed to move it back to a virtual slot.

The steps to restore data are similar to those in a physical tape library, but the restore operation is nearly instantaneous. Even though virtual tapes are based on disks, which provide random access, they still emulate the tape behavior.

Backup Target Comparison

	Tape	Disk	Virtual Tape
Offsite Replication Capabilities	No	Yes	Yes
Reliability	No inherent protection methods	RAID, spare	RAID, spare
Performance	Low	High	High
Use	Backup only	Multiple (backup and production)	Backup only

The table on the slide provides the comparison among various backup targets.

Module 10: Backup and Archive

Lesson 4: Data Deduplication

During this lesson the following topics are covered:

- Deduplication overview
- Deduplication methods
- Deduplication implementations
- Key benefits of deduplication

This lesson covers different deduplication method. This lesson also covers deduplication implementation such as source-based deduplication and target-based deduplication. Further this lesson details various key benefits of data deduplication.

What is Data Deduplication?

Data Deduplication

It is a process of identifying and eliminating redundant data.

- Deduplication methods
 - ▶ File level
 - ▶ Subfile level
- Deduplication implementations
 - ▶ Source-based
 - ▶ Target-based

Traditional backup solutions do not provide any inherent capability to prevent duplicate data from being backed up. With the growth of information and 24x7 application availability requirements, backup windows are shrinking. Traditional backup processes back up a lot of duplicate data. Backing up duplicate data significantly increases the backup window size requirements and results in unnecessary consumption of resources, such as storage space and network bandwidth.

Data deduplication is the process of identifying and eliminating redundant data. When duplicate data is detected during backup, the data is discarded and only the pointer is created to refer the copy of the data that is already backed up. Data deduplication helps to reduce the storage requirement for backup, shorten the backup window, and remove the network burden. It also helps to store more backups on the disk and retain the data on the disk for a longer time.

There are two methods of data deduplication, file level and subfile level. Determining the uniqueness by implementing either method offers benefits; however, results can vary. The differences exist in the amount of data reduction each method produces and the time each approach takes to determine the unique content.

Deduplication can occur close to where the data is created, which is often referred to as “Source-based Deduplication”. It can also occur close to where the data is stored, which is commonly called “Target-based Deduplication”.

Data Deduplication Methods

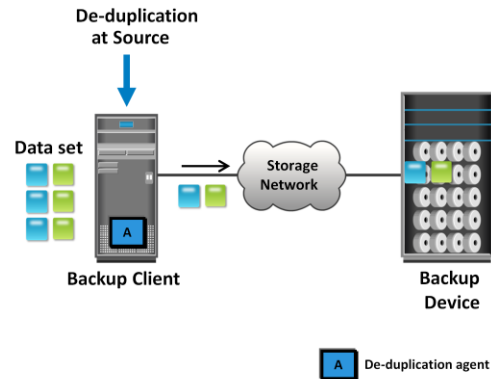
- File-level deduplication (single-instance storage)
 - Detects and removes redundant copies of identical files
 - After a file is stored, all other references to the same file refer to the original copy
- Subfile deduplication
 - Detects redundant data within and across files
 - Two methods
 - Fixed-length block
 - Variable-length segment

File-level deduplication (also called *single-instance storage*) detects and removes redundant copies of identical files. It enables storing only one copy of the file; the subsequent copies are replaced with a pointer that points to the original file. File-level deduplication is simple and fast but does not address the problem of duplicate content inside the files. For example, two 10-MB PowerPoint presentations with a difference in just the title page are not considered as duplicate files, and each file will be stored separately.

Subfile deduplication breaks the file into smaller chunks and then uses specialized algorithm to detect redundant data within and across the file. As a result, subfile deduplication eliminates duplicate data across files. There are two forms of subfile deduplication: fixed-length block and variable-length segment. The *fixed-length block deduplication* divides the files into fixed-length blocks and uses a hash algorithm to find the duplicate data. Although simple in design, fixed-length blocks might miss many opportunities to discover redundant data because the block boundary of similar data might be different. Consider the addition of a person's name to a document's title page. This shifts the whole document, and all the blocks appear to have changed, causing the failure of the deduplication method to detect equivalencies. In variable-length segment deduplication, if there is a change in the segment, the boundary for only that segment is adjusted, leaving the remaining segments unchanged. This method vastly improves the ability to find duplicate data segments compared to fixed-block.

Data Deduplication Implementation – Source-based

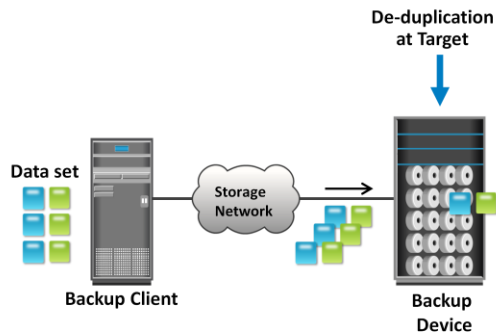
- Data is deduplicated at the source (backup client)
- Backup client sends only new, unique segments across the network
- Reduced storage capacity and network bandwidth requirements
- Increased overhead on the backup client



Source-based data deduplication eliminates redundant data at the source before it transmits to the backup device. Source-based data deduplication can dramatically reduce the amount of backup data sent over the network during backup processes. It provides the benefits of a shorter backup window and requires less network bandwidth. There is also a substantial reduction in the capacity required to store the backup images. Source-based deduplication increases the overhead on the backup client, which impacts the performance of the backup and application running on the client. Source-based deduplication might also require a change of backup software if it is not supported by backup software.

Data Deduplication Implementation – Target-based

- Data is deduplicated at the target
 - ▶ Inline
 - ▶ Post-process
- Offloads the backup client from deduplication process
- All the backup data traverse the network



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 65

Target-based data deduplication is an alternative to source-based data deduplication. Target-based data deduplication occurs at the backup device, which offloads the backup client from the deduplication process. Figure on the slide illustrates target-based data deduplication. In this case, the backup client sends the data to the backup device and the data is deduplicated at the backup device, either immediately (Inline) or at a scheduled time (Post-process).

Inline deduplication performs deduplication on the backup data before it is stored on the backup device. Hence, this method reduces the storage capacity needed for the backup. Inline deduplication introduces overhead in the form of the time required to identify and remove duplication in the data. So, this method is best suited for an environment with a large backup window.

Post-process deduplication enables the backup data to be stored on the backup device first and then deduplicated later. This method is suitable for situations with tighter backup windows. However, post-process deduplication requires more storage capacity to store the backup images before they are deduplicated.

Because deduplication occurs at the target, all the backup data need to be transferred over the network, which increases network bandwidth requirements. Target-based data deduplication does not require any changes in the existing backup software.

Data Deduplication – Key Benefits

- Reduces infrastructure costs
 - ▶ By eliminating redundant data, less storage is required to hold the backup images
- Enables longer retention periods
 - ▶ Reduces the amount of redundant content in the daily backup, and hence, users can extend their retention policies
- Reduces backup window
 - ▶ Less data to be backed up, which reduces backup window
- Reduces backup bandwidth requirement
 - ▶ Source based de-duplication eliminates redundant data before data is sent over the network

Reduces infrastructure costs: By eliminating redundant data from the backup, far less infrastructure is required to hold the backup images. Data de-duplication directly results in reduced storage capacities to hold backup images. Smaller capacity requirements means lower acquisition costs as well as reduced power and cooling costs.

Enables longer retention periods: As data de-duplication reduces the amount of content in the daily backup, users can extend their retention policies. This can have a significant benefit to users who currently require longer retention.

Reduces backup window: Data de-duplication eliminates redundant content of backup data, which makes less data to be backed up and reduces backup window.

Reduces backup bandwidth requirement: By utilizing data de-duplication at the client (source-based), redundant data is removed before the data is transferred over the network. This considerably reduces the network bandwidth required for backup.

Use Case: Remote Office/Branch Office Backup

- Protecting data at an organization's branch and remote offices, across multiple locations, is critical for business
- Backing up data from remote offices to a centralized data center was restricted due to
 - ▶ Time and cost involved in sending huge volumes of data over the network
- Disk-based backup solution, along with source-based deduplication, eliminates the challenges in centrally backing up remote-office data
 - ▶ Reduces the network bandwidth requirement
 - ▶ Reduces the backup window

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 67

Today, businesses have their remote or branch offices spread over multiple locations. Typically, these remote offices have their local IT infrastructure. This infrastructure includes file, print, Web, or email servers, workstations, and desktops, and might also house some applications and databases. Too often, business-critical data at remote offices are inadequately protected, exposing the business to the risk of lost data and productivity. As a result, protecting the data of an organization's branch and remote offices across multiple locations is critical for business. Traditionally, remote-office data backup was done manually using tapes, which were transported to offsite locations for DR support. Some of the challenges with this approach were lack of skilled onsite technical resources to manage backups and risk of sending tapes to offsite locations, which could result in loss or theft of sensitive data. Backing up data from remote offices to a centralized data center was restricted due to the time and cost involved in sending huge volumes of data over the WAN. Therefore, organizations needed an effective solution to address the data backup and recovery challenges of remote and branch offices.

Disk-based backup solutions along with source-based deduplication eliminate the challenges associated with centrally backing up remote-office data. Deduplication considerably reduces the required network bandwidth and enables remote-office data backup using the existing network. Organizations can now centrally manage and automate remote-office backups while reducing the required backup window.

Module 10: Backup and Archive

Lesson 5: Backup in Virtualized Environment

During this lesson the following topics are covered:

- Traditional backup approach
- Image-based backup

This lesson covers traditional backup approach and image-based backup in virtualized environment.

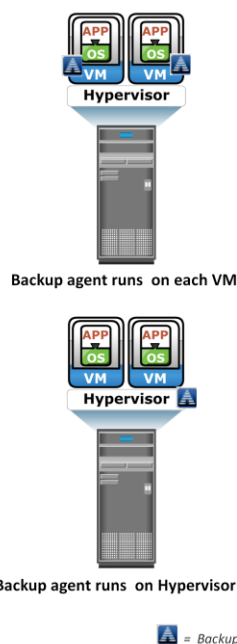
Backup in Virtualized Environment Overview

- Backup options
 - ▶ Traditional backup approach
 - ▶ Image-based backup approach
- Backup optimization
 - ▶ Deduplication

In a virtualized environment, it is imperative to back up the virtual machine data (OS, application data, and configuration) to prevent its loss or corruption due to human or technical errors. There are two approaches for performing a backup in a virtualized environment: the traditional backup approach and the image-based backup approach. Owing to the increased capacity requirements in a virtualized environment, backup optimization methods are necessary. The use of deduplication techniques significantly reduces the amount of data to be backed up in a virtualized environment. The effectiveness of deduplication is identified when VMs with similar configurations are deployed in a data center. The deduplication types and methods used in a virtualized environment are the same as in the physical environment.

Traditional Backup Approaches

- Backup agent on VM
 - ▶ Requires installing a backup agent on each VM running on a hypervisor
 - ▶ Can only backup virtual disk data
 - ▶ Does not capture VM files such as VM swap file, configuration file
 - ▶ Challenge in VM restore
- Backup agent on Hypervisor
 - ▶ Requires installing backup agent only on hypervisor
 - ▶ Backs up all the VM files



In the *traditional backup approach*, a backup agent is installed either on the virtual machine (VM) or on the hypervisor. If the backup agent is installed on a VM, the VM appears as a physical server to the agent. The backup agent installed on the VM backs up the VM data to the backup device. The agent does not capture VM files, such as the virtual BIOS file, VM swap file, logs, and configuration files. Therefore, for a VM restore, a user needs to manually re-create the VM and then restore data on to it.

If the backup agent is installed on the hypervisor, the VMs appear as a set of files to the agent. So, VM files can be backed up by performing a file system backup from a hypervisor. This approach is relatively simple because it requires having the agent just on the hypervisor instead of having on all the VMs. The traditional backup method can cause high CPU utilization on the server being backed up. In the traditional approach, the backup should be performed when the server resources are idle or during a low activity period on the network. Also consider allocating enough resources to manage the backup on each server when a large number of VMs are in the environment.

Image-based Backup

- Creates a copy of the guest OS, its data, VM state, and configurations
 - ▶ The backup is saved as a single file – “image”
 - ▶ Mounts image on a proxy server
 - ▶ Offloads backup processing from the hypervisor
- Enables quick restoration of VM

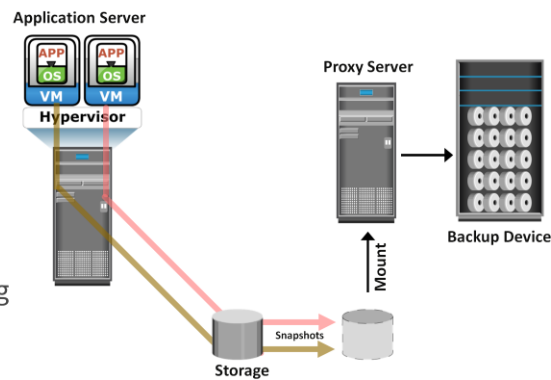


Image-based backup operates at the hypervisor level and essentially takes a snapshot of the VM. It creates a copy of the guest OS and all the data associated with it (snapshot of VM disk files), including the VM state and application configurations. The backup is saved as a single file called an “image” and this image is mounted on the proxy server (acts as a backup client). The backup software then backs up these image files normally. This effectively offloads the backup processing from the hypervisor and transfers the load on the proxy server, thereby reducing the impact to VMs running on the hypervisor. Image-based backup enables quick restoration of a VM.

Module 10: Backup and Archive

Lesson 6: Data Archive

During this lesson the following topics are covered:

- Fixed content
- Data archive
- Archive solution architecture

This lesson covers fixed content and challenges in storing fixed content. This lesson also focuses on data archive solution architecture.

Fixed Content

- Fixed content is growing at more than 90% annually
 - ▶ Significant amount of newly created information falls into this category
 - ▶ New regulations require retention and data protection

Examples of Fixed Content		
<i>Electronic Documents</i> • Contracts and claims • Email attachments • Financial spread sheets • CAD/CAM designs • Presentations	<i>Digital Records</i> • Documents • Checks, securities trades • Historical preservation • Photographs • Personal/professional • Surveys • Seismic, astronomic, geographic	<i>Rich Media</i> • Medical • X-rays, MRIs, CT Scan • Video • News/media, movies • Security surveillance • Audio • Voicemail • Radio

In the life cycle of information, data is actively created, accessed, and changed. As data ages, it is less likely to be changed and eventually becomes “fixed” but continues to be accessed by applications and users. This data is called *fixed content*. All organizations may require to retain their data for an extended period of time due to government regulations and legal/contractual obligations. Organizations also make use of this fixed content to generate new revenue strategies and improve service levels.

Currently, fixed content data is the fastest growing sector of the data storage market. Assets such as X-rays, MRIs, CAD/CAM designs, surveillance video, MP3s and financial documents are just a few examples of an important class of data that is growing at over 90% annually.

Data Archive

- A repository where fixed content is stored
- Enables organizations retaining their data for an extended period of time in order to
 - ▶ Meet regulatory compliance
 - ▶ Plan new revenue strategies
- Archive can be implemented as
 - ▶ Online
 - ▶ Nearline
 - ▶ Offline

Data archive is a repository where fixed content is stored. It enables organizations retaining their data for an extended period of time in order to meet regulatory compliance and generate new revenue strategies. An archive can be implemented as an online, nearline, or offline solution:

Online archive: A storage device directly connected to a host that makes the data immediately accessible.

Nearline archive: A storage device connected to a host, but the device where the data is stored must be mounted or loaded to access the data.

Offline archive: A storage device that is not ready to use. Manual intervention is required to connect, mount, or load the storage device before data can be accessed.

Challenges of Traditional Archiving Solutions

- Both tape and optical are susceptible to wear and tear
 - ▶ Involve operational, management, and maintenance overhead
- Have no intelligence to identify duplicate data
 - ▶ Same content could be archived many times
- Inadequate for long-term preservation (years-decades)
- Unable to provide online and fast access to fixed content

Typically, long-term preservation is required (years-decades) for fixed content and it is also important to have simultaneous, fast online access. The increase in fixed content is driven by regulatory requirements. Because of this explosive growth and changes to user requirements, traditional storage options are inadequate.

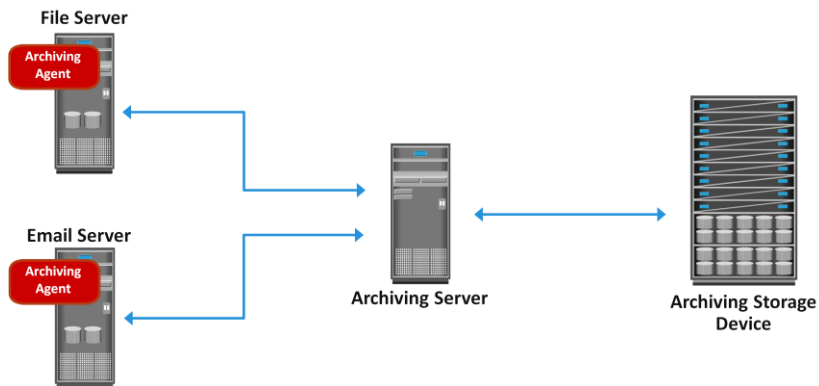
Optical media are typically *write once read many* (WORM) devices that protect the original file from being overwritten. Some tape devices also provide this functionality by implementing file-locking capabilities. Although these devices are inexpensive, they involve operational, management, and maintenance overhead. The traditional solutions using optical discs and tapes is not optimized to recognize the content, so that the same content could be stored several times. Additional costs are involved in offsite storage of media and media management. Tapes and optical media are also susceptible to wear and tear. Frequent changes in these device technologies lead to the overhead of converting the media into new formats to enable access and retrieval. Government agencies and industry regulators are establishing new laws and regulations to enforce the protection of archives from unauthorized destruction and modification. These regulations and standards have established new requirements for preserving the integrity of information in the archives. These requirements have exposed the shortcomings of the traditional tape and optical media archive solutions.

Content Addressed Storage – An Archival Solution

- Disk-based storage that has emerged as an alternative to traditional archiving solutions
- Provides online accessibility to archive data
- Enables organization to meet the required SLAs
- Provides features that are required for storing archive data
 - ▶ Content authenticity and content integrity
 - ▶ Location independence
 - ▶ Single-instance storage
 - ▶ Retention enforcement
 - ▶ Data protection

Content addressed storage (CAS) a disk based storage that has emerged as an alternative to tape and optical solutions. CAS meets the demand to improve data accessibility and to protect, dispose off, and ensure service level agreements (SLAs) for archive data. CAS is detailed in module 8.

Archiving Solution Architecture



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 77

Archiving solution architecture consists of three key components; archiving agent, archiving server, and archiving storage device.

An archiving agent is software installed on the application server. The agent is responsible for scanning the data that can be archived based on the policy defined on the archiving server. After the data is identified for archiving, the agent sends the data to the archiving server. Then the original data on the application server is replaced with a stub file. The stub file contains the address of the archived data. The size of this file is small and significantly saves space on primary storage. This stub file is used to retrieve the file from the archive storage device.

An archiving server is software installed on a host that enables administrators to configure the policies for archiving data. Policies can be defined based on file size, file type, or creation/modification/access time. The archiving server receives the data to be archived from the agent and sends it to the archive storage device.

An archive storage device stores fixed content.

Use Case: Email Archiving

- Moves the emails from primary to archive storage, based on policy
- Saves space on primary storage
- Enables to retain emails in the archive for longer period to meet regulatory requirements
- Gives end users virtually unlimited mailbox space
- File archiving is another use case that benefits from an archival solution

E-mail is an example of an application that benefits most by an archival solution. Typically, a system administrator configures small mailboxes that store a limited number of e-mails. This is because large mailboxes with a large number of e-mails can make management difficult, increase primary storage cost, and degrade system performance. When an e-mail server is configured with a large number of mailboxes, the system administrator typically configures a quota on each mailbox to limit its size. Configuring fixed quota on mailboxes impacts end users. A fixed quota for a mailbox forces users to delete e-mails as they approach the quota size. End users often need to access e-mails that are weeks, months, or even years old.

E-mail archiving provides an excellent solution that overcomes the preceding challenges. Archiving solutions move e-mails that have been identified as candidates for archive from primary storage to the archive storage device based on a policy—for example, “e-mails that are 90 days old should be archived.” After the e-mail is archived, it is retained for years based on the retention policy. This considerably saves space on primary storage and enables organizations to meet regulatory requirements. Implementation of an archiving solution gives end users virtually unlimited mailbox space.

A file sharing environment is another environment that benefits from an archival solution. Typically, users store a large number of files in the shared location. Most of these files are old and rarely accessed. Administrators configure quotas on the file share that forces the users to delete these files. This impacts users because they may require access to files that may be months or even years old. In some cases the user may request an increase in the size of the file share. This in turn increases the cost of primary storage. A file archiving solution archives the files based on the policy such as age of files, size of files, and so on. This considerably reduces the primary storage requirement and also enables users to retain the files in the archive for longer periods.

Module 10: Backup and Archive

Concepts in Practice

- EMC NetWorker
- EMC Avamar
- EMC Data Domain

The concept in practice section covers various EMC backup and archive products.

EMC NetWorker

- Centralizes, automates, and accelerates data backup and recovery operations across the enterprise
- Key features
 - ▶ Supports heterogeneous platforms such as Windows, UNIX, Linux, and also supports virtual environments
 - ▶ Supports different backup targets – tapes, disks, and virtual tapes
 - ▶ Supports Multiplexing (or multi-streaming) of data
 - ▶ Provides both source-based and target-based deduplication capabilities by integrating with EMC Avamar and EMC Data Domain respectively
 - ▶ Cloud-backup option enables backing up data to cloud

The EMC NetWorker backup and recovery software centralizes, automates, and accelerates data backup and recovery operations across the enterprise. The features of EMC NetWorker are listed on the slide.

EMC Avamar

- Disk-based backup and recovery solution that provides source-based data deduplication
- Three major components include Avamar server, Avamar backup clients, and Avamar administrator
- Avamar server includes
 - ▶ Software only, Avamar Data Store, Avamar Virtual Edition

EMC Avamar is a disk-based backup and recovery solution that provides inherent source-based data deduplication. With its unique global data deduplication feature, Avamar differs from traditional backup and recovery solutions, by identifying and storing only unique sub-file data objects. Redundant data is identified at the source, the amount of data that travels across the network is drastically reduced, and the backup storage requirement is also considerably reduced. The three major components of an Avamar system includes Avamar server, Avamar backup clients, and Avamar administrator. Avamar server provides the essential processes and services required for client access and remote system administration. The Avamar client software runs on each computer or network server that is being backed up. Avamar administrator is a user management console application that is used to remotely administer an Avamar system. The three Avamar server editions include software only, Avamar Data Store, and Avamar Virtual Edition. The features of EMC Avamar are as follows:

- **Fault tolerance:** Uses RAID, RAIN, checkpoints, and replication to provide data integrity and protection.
- **Standard IP network leveraging:** Optimizes the use of network for backup; dedicated backup networks are not required. Daily full backups are possible using the existing networks and infrastructure.
- **Scalable server architecture:** Additional storage nodes can be added non-disruptively to accommodate increased backup storage requirements.
- **Centralized management:** Enables remote management of Avamar servers from a centralized location and through the use of the Avamar Enterprise Manager and Avamar Administrator interfaces.

EMC Data Domain

- Target-based deduplication solution
- Provides technological advantages
 - ▶ Data invulnerability architecture
 - ▶ Data Domain Stream-Informed Segment Layout (SISL) scaling architecture
 - ▶ Support native replication technology
 - ▶ Global compression
- EMC Data Domain Archiver
 - ▶ Solution for long term retention of backup and archive data
 - ▶ Designed with internal tiering approach
 - ▶ Supports deduplication technology

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 82

The EMC Data Domain deduplication storage system is a target-based data deduplication solution. Using high-speed, inline deduplication technology, the Data Domain system provides a storage footprint that is significantly smaller on an average, than that of the original data set. Data Domain systems can scale from smaller remote office appliances to large data-center systems. These systems are available as integrated appliances or as gateways that use external storage.

Data Domain deduplication storage systems provide the following unique advantages:

Data invulnerability architecture: Provides unprecedented levels of data integrity, data verification, and self-healing capabilities, such as RAID 6 protection. Continuous fault detection, healing, and write verification ensure that the backup is accurately stored, available, and recoverable.

Data Domain SISL (Stream-Informed Segment Layout) scaling architecture: Enables scaling of CPUs to add a direct benefit to system throughput scalability.

Support native replication technology: Enables automatic, secure transfer of compressed data over the wide area network (WAN) with minimum bandwidth requirement.

Global compression: Highly efficient deduplication and compression technology, which radically changes storage economics.

EMC Data Domain Archiver is a solution for long term retention of backup and archive data. It is designed with internal tiering approach to enable cost effective, long term retention of data on disk by implementing deduplication technology.

Module 10: Summary

Key points covered in this module:

- Backup granularity
- Backup and recovery operations
- Backup topologies
- Backup targets
- Data deduplication
- Backup in virtualized environment
- Data archive

This module covered various backup granularities and backup operations. This module also discussed various backup topologies and backup targets. Further this module covered data deduplication and backup in virtualized environment. Additionally, this module also covered data archive in detail.

A *backup* is an additional copy of the production data, created and retained for the sole purpose of recovering lost or corrupted data. Based on the granularity, backups can be categorized as full, cumulative, and incremental.

The three basic topologies used in a backup environment are direct-attached backup, LAN-based backup, and SAN-based backup.

A wide range of technology solutions are currently available for backup targets. Tape and disk libraries are the two most commonly used backup targets. Virtual tape library (VTL) is one of the options that uses disks as backup medium. VTL emulates tapes and provides enhanced backup and recovery capabilities.

Data deduplication is the process of identifying and eliminating redundant data. When duplicate data is detected during backup, the data is discarded and only the pointer is created to refer the copy of the data which is already backed up.

In a virtualized environment, it is imperative to back up the virtual machine data (OS, application data, and configuration) to prevent its loss or corruption due to human or technical errors.

Data archive is a repository where fixed content is stored. It enables organizations to retain their data for an extended period of time, in order to meet regulatory compliance and generate new revenue strategies.

Check Your Knowledge – 1

- Which is true about incremental backup?
 - A. Restore requires only last full and last incremental backup
 - B. Restore requires only last incremental backup
 - C. Copies the data that has changed since last full or incremental backup
 - D. Copies the data that has changed since last full backup
- What is an advantage of image-based backup over traditional backup approach in a virtualized environment?
 - A. Offloads backup processing from the hypervisor
 - B. Faster because it copies only virtual machine disk data
 - C. Faster because it copies only virtual machine configuration data
 - D. Space required is a fraction of total backup data

Check Your Knowledge – 2

- Which accurately describes the role of a backup server?
 - A. Gathers the data that is to be backed up and send it to storage node
 - B. Responsible for writing the data, which client sends, to backup device
 - C. Manages the backup operation and maintains backup catalog
 - D. Controls the robotic arm in the tape library

- In backup to tape environment, what does 'shoe shining' mean?
 - A. Writing data from multiple streams on a single tape
 - B. Process of emulating disk drives and presenting as tapes to backup software
 - C. Repeated back and forth motion that a tape drive makes when there is an interruption in the backup data stream
 - D. Process of deleting redundant content in the backup data

Check Your Knowledge – 3

- What is an advantage of source-based deduplication?
 - A. Improves the performance of backup client
 - B. Improves the performance of backup server
 - C. Reduces backup window to zero
 - D. Reduces the network bandwidth requirement for backup

Exercise: Backup/Recovery

- Current situation
 - ▶ Full backup is performed on every Sunday and incremental on remaining days
 - ▶ Database have to be shut down during backup
 - ▶ Multiple redundant copies of backup data
 - ▶ Network bandwidth constraint
- Business requirement
 - ▶ Eliminate the need to shutdown the database for backup
 - ▶ Need faster backup and restore
 - ▶ Eliminate redundant copies of backup data
- Task
 - ▶ Suggest a solution and justify

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 87

Business profile:

An organization uses tape as their primary backup storage media for their applications.

Current situation:

- Full backup is performed on every Sunday and incremental on remaining days
- Database have to be shut down during backup
- Multiple redundant copies of backup data
- Network bandwidth constraint

Requirements:

- Eliminate the need to shutdown the database for backup
- Need faster backup and restore
- Eliminate redundant copies of backup data

Task:

Propose a solution to address the organization's concern and justify your solution.

Module – 11

Local Replication



Module 11: Local Replication

Upon completion of this module, you should be able to:

- Describe various uses of local replica
- Describe how consistency is ensured in file system and database replication
- Describe host-based, array-based, and network-based local replication technologies
- Explain restore and restart considerations
- Describe local replication in virtualized environment

This module focuses on various uses of local replica along with file system and database consistency in replication. This module also focuses on various host-based, storage-based, and network-based local replication technologies. Further, this module focuses on restore and restart considerations. Additionally, this module details local replication in a virtualized environment.

Module 11: Local Replication

Lesson 1: Local Replication Overview

During this lesson the following topics are covered:

- Uses of local replica
- File system and database consistency

This lesson covers various uses of local replica and also covers how consistency is ensured in file system and database replication.

What is Replication?

Replication

It is a process of creating an exact copy (replica) of data.

- Replication can be classified as
 - ▶ Local replication
 - ▶▶ Replicating data within the same array or data center
 - ▶ Remote replication
 - ▶▶ Replicating data at remote site



In today's business environment, it is imperative for an organization to protect mission-critical data and minimize the risk of business disruption. If a local outage or disaster occurs, fast data restore and restart is essential to ensure business continuity (BC). Replication is one of the ways to ensure BC. It is the process to create an exact copy (replica) of data. These replica copies are used for restore and restart operations if data loss occurs. These replicas can also be assigned to other hosts to perform various business operations, such as backup, reporting, and testing.

Replication can be classified into two major categories: local and remote. Local replication refers to replicating data within the same array or the same data center. Remote replication refers to replicating data at a remote site. Remote replication is discussed in the next module.

Uses of Local Replica

- Alternate source for backup
- Fast recovery
- Decision support activities
- Testing platform
- Data Migration

One or more local replicas of the source data may be created for various purposes, including the following:

Alternative source for backup: Under normal backup operations, data is read from the production volumes (LUNs) and written to the backup device. This places an additional burden on the production infrastructure because production LUNs are simultaneously involved in production operations and servicing data for backup operations. The local replica contains an exact point-in-time (PIT) copy of the source data, and therefore can be used as a source to perform backup operations. This alleviates the backup I/O workload on the production volumes. Another benefit of using local replicas for backup is that it reduces the *backup window* to zero.

Fast recovery: If data loss or data corruption occurs on the source, a local replica might be used to recover the lost or corrupted data. If a complete failure of the source occurs, some replication solutions enable replica to be used to restore data on to a different set of source devices or production can be restarted on the replica. In either case, this method provides faster recovery and minimal RTO compared to traditional recovery from tape backups.

Cont..

Decision-support activities, such as reporting or data warehousing: Running the reports using the data on the replicas greatly reduces the I/O burden placed on the production device. Local replicas are also used for data-warehousing applications. The data-warehouse application may be populated by the data on the replica and thus avoid the impact on the production environment.

Testing platform: Local replicas are also used for testing new applications or upgrades. For example, an organization may use the replica to test the production application upgrade; if the test is successful, the upgrade may be implemented on the production environment.

Data migration: Another use for a local replica is data migration. Data migrations are performed for various reasons, such as migrating from a smaller capacity LUN to one of a larger capacity for newer versions of the application.

Replica Characteristics

- Recoverability/Restartability
 - Replica should be able to restore data on the source device
 - Restart business operation from replica
- Consistency
 - Replica must be consistent with the source
- Choice of replica tie back into RPO
 - Point-in-Time (PIT)
 - Non-zero RPO
 - Continuous
 - Near-zero RPO

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 94

A replica should have following characteristics:

Recoverability: Enables restoration of data from the replicas to the source if data loss or corruption occurs.

Restartability: Enables restarting business operations using the replicas.

Consistency: Replica must be consistent with the source so that it is usable for both recovery and restart operations. Ensuring consistency is the primary requirement for all the replication technologies.

Replicas can either be point-in-time (PIT) or continuous:

Point-in-Time: The data on the replica is an identical image of the production at some specific timestamp. For example, a replica of a file system is created at 4:00 PM on Monday. This replica would then be referred to as the Monday 4:00 PM PIT copy. The RPO will map to the time when the PIT was created to the time when any kind of failure on the production occurred. If there is a failure on the production at 8:00 PM and there is a 4:00 PM PIT available, the RPO would be 4 hours ($8 - 4 = 4$). To minimize RPO, take periodic PITs.

Continuous replica: The data on the replica is in-sync with the production data at all times. The objective with any continuous replication is to reduce the RPO to zero or near-zero.

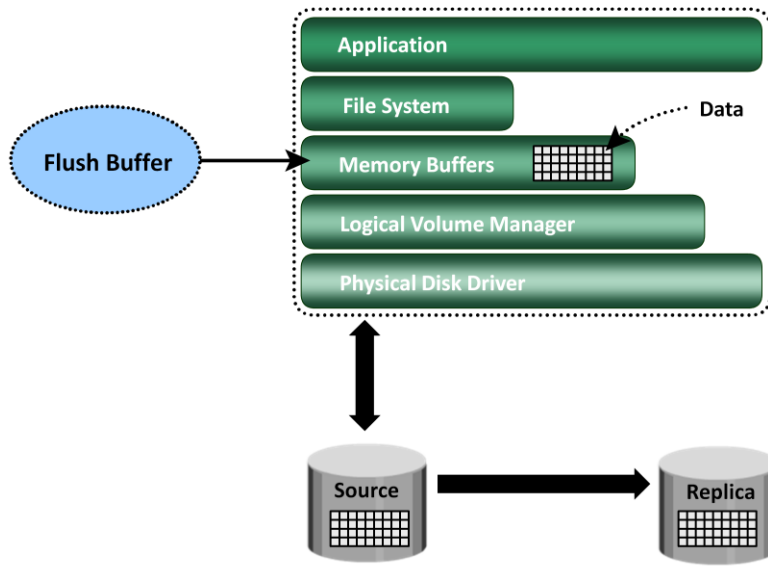
Understanding Consistency

- Consistency ensures the usability of replica
- Consistency can be achieved in various ways for file system and database

	Offline	Online
File System	Unmount file system	Flushing host buffers
Database	Shutdown database	a) Using dependent write I/O principle b) Holding I/Os to source before creating replica

Consistency is a primary requirement to ensure the usability of replica device. In case of file systems, consistency can be achieved either by taking FS offline i.e. by un-mounting FS or by keeping FS online by flushing host buffers before creating replica. Similarly in case of databases, consistency can be achieved either by taking database offline for creating consistent replica or by keeping online. Consistent replica of online database can be created by using dependent write I/O principle or by holding I/Os before creating replica.

File System Consistency: Flushing Host Buffer



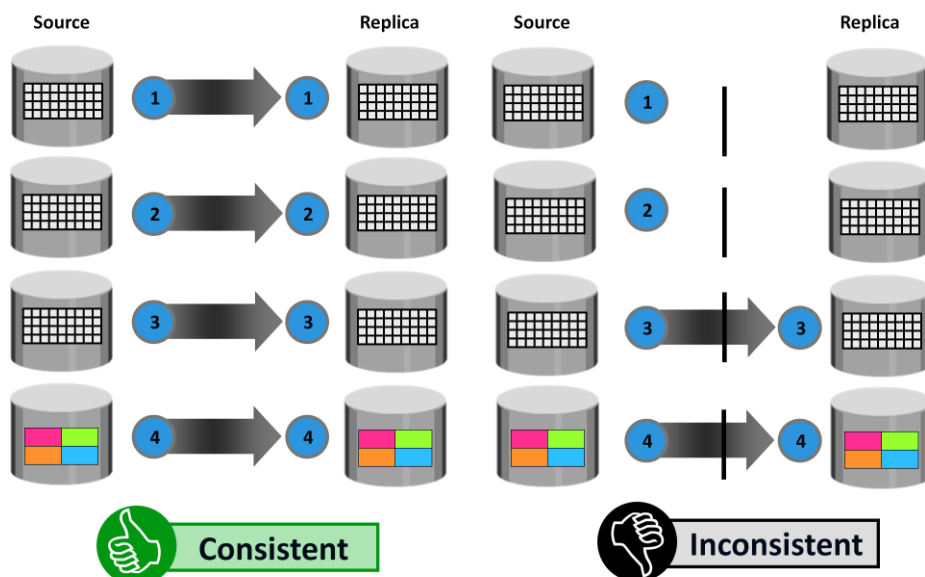
EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 96

File systems buffer the data in the host memory to improve the application response time. The buffered data is periodically written to the disk. In UNIX operating systems, *sync daemon* is the process that flushes the buffers to the disk at set intervals. In some cases, the replica is created between the set intervals, which might result in the creation of an inconsistent replica. Therefore, host memory buffers must be flushed to ensure data consistency on the replica, prior to its creation. If the host memory buffers are not flushed, the data on the replica will not contain the information that was buffered in the host. If the file system is unmounted before creating the replica, the buffers will be automatically flushed and the data will be consistent on the replica.

If a mounted file system is replicated, some level of recovery, such as *fsck* or *log replay*, is required on the replicated file system. When the file system replication and check process are completed, the replica file system can be mounted for operational use.

Database Consistency: Dependent Write I/O Principle



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 97

If the database is online, it is available for I/O operations, and transactions to the database update the data continuously. When a database is replicated while it is online, changes made to the database at this time must be applied to the replica to make it consistent. A consistent replica of an online database is created by using the dependent write I/O principle or by holding I/Os momentarily to the source before creating the replica.

A *dependent write I/O* principle is inherent in many applications and database management systems (DBMS) to ensure consistency. According to this principle, a write I/O is not issued by an application until a prior related write I/O has completed. For example, a data write is dependent on the successful completion of the prior log write.

For a transaction to be deemed complete, databases require a series of writes to have occurred in a particular order. These writes will be recorded on the various devices/file systems.

When the replica is created, all the writes to the source devices must be captured on the replica devices to ensure data consistency. Figure on the slide illustrates the process of replication from the source to the replica. The I/O transactions 1 to 4 must be carried out for the data to be consistent on the replica. It is possible that I/O transactions 3 and 4 were copied to the replica devices, but I/O transactions 1 and 2 were not copied. Figure on the slide also shows this situation. In this case, the data on the replica is inconsistent with the data on the source. If a restart were to be performed on the replica devices, I/O 4, which is available on the replica, might indicate that a particular transaction is complete, but all the data associated with the transaction will be unavailable on the replica, making the replica inconsistent.

Another way to ensure consistency is to make sure that the write I/O to all source devices is held for the duration of creating the replica. This creates a consistent image on the replica. However, databases and applications might time out if the I/O is held for too long.

Module 11: Local Replication

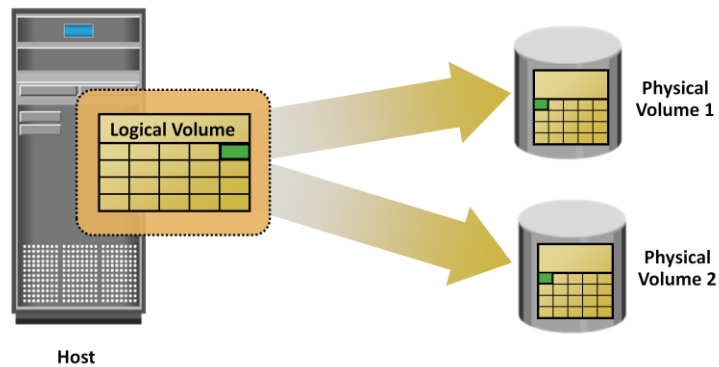
Lesson 2: Local Replication Technologies

During this lesson the following topics are covered:

- Local replication technologies
- Restore and restart considerations

This lesson covers various host-based, network-based and storage array-based local replication technologies and also covers restore and restart considerations.

Host-based Replication: LVM-based Mirroring



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 99

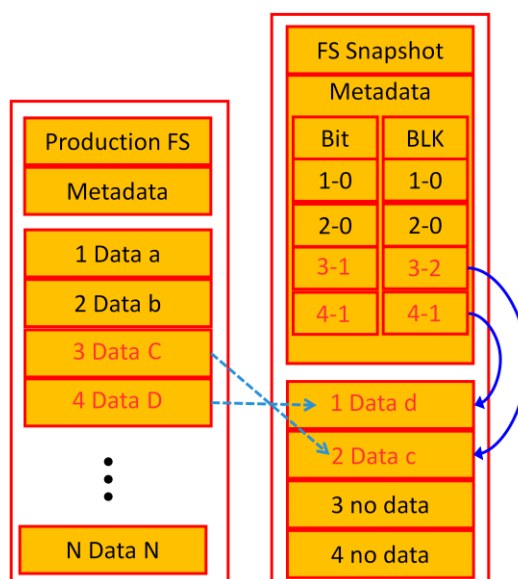
In *LVM-based replication*, the logical volume manager is responsible for creating and controlling the host-level logical volumes. An LVM has three components: physical volumes (physical disk), volume groups, and logical volumes. A *volume group* is created by grouping one or more physical volumes. *Logical volumes* are created within a given volume group. A volume group can have multiple logical volumes. In LVM-based replication, each *logical block* in a logical volume is mapped to two physical blocks on two different physical volumes, as shown in figure on the slide. An application write to a logical volume is written to the two physical volumes by the LVM device driver. This is also known as *LVM mirroring*. Mirrors can be split, and the data contained therein can be independently accessed.

Advantages: The LVM-based replication technology is not dependent on a vendor-specific storage system. Typically, LVM is part of the operating system, and no additional license is required to deploy LVM mirroring.

Limitations: Every write generated by an application translates into two writes on the disk, and thus, an additional burden is placed on the host CPU. This can degrade application performance. Presenting an LVM-based local replica to another host is usually not possible because the replica will still be part of the volume group, which is usually accessed by one host at any given time. If the devices are already protected by some level of RAID on the array, then the additional protection that the LVM mirroring provides is unnecessary. This solution does not scale to provide replicas of federated databases and applications. Both the replica and source are stored within the same volume group. Therefore, the replica might become unavailable if there is an error in the volume group. If the server fails, both the source and replica are unavailable until the server is brought back online.

Host-based Replication: File System Snapshot

- Pointer-based replication
- Uses Copy on First Write (CoFW) principle
- Uses bitmap and block map
- Requires a fraction of the space used by the production FS



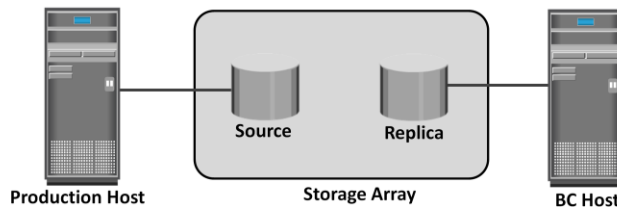
EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 100

File system (FS) snapshot is a pointer-based replica that requires a fraction of the space used by the production FS. It uses the Copy on First Write (CoFW) principle to create snapshots. When a snapshot is created, a bitmap and blockmap are created in the metadata of the Snap FS. The bitmap is used to keep track of blocks that are changed on the production FS after the snap creation. The blockmap is used to indicate the exact address from which the data is to be read when the data is accessed from the Snap FS. Immediately after the creation of the FS Snapshot, all reads from the snapshot are actually served by reading the production FS. In a CoFW mechanism, if a write I/O is issued to the production FS for the first time after the creation of a snapshot, the I/O is held and the original data of production FS corresponding to that location is moved to the Snap FS. Then, the write is allowed to the production FS. The bitmap and blockmap are updated accordingly. Subsequent writes to the same location will not initiate the CoFW activity. To read from the Snap FS, the bitmap is consulted. If the bit is 0, then the read is directed to the production FS. If the bit is 1, then the block address is obtained from the blockmap and the data is read from that address on the snap FS. Read requests from the production FS work as normal.

Storage Array-based Local Replication

- Replication performed by the array operating environment
- Source and replica are on the same array
- Types of array-based replication
 - ▶ Full-volume mirroring
 - ▶ Pointer-based full-volume replication
 - ▶ Pointer-based virtual replication



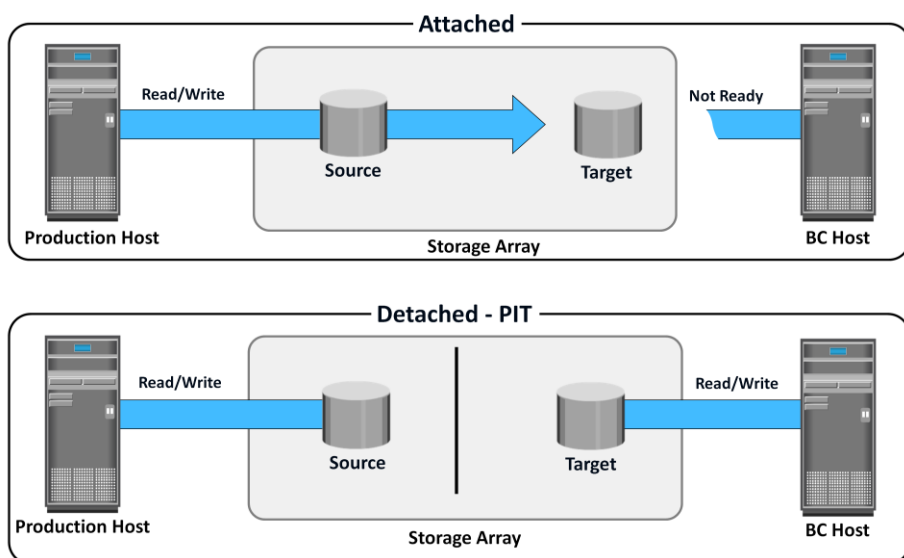
EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 101

In *storage array-based local replication*, the array-operating environment performs the local replication process. The host resources, such as the CPU and memory are not used in the replication process. Consequently, the host is not burdened by the replication operations. The replica can be accessed by an alternate host for other business operations.

In this replication, the required number of replica devices should be selected on the same array and then data should be replicated between the source-replica pairs. Figure on the slide shows a storage array-based local replication, where the source and target (replica) are in the same array and accessed by different hosts. Storage array-based local replication is commonly implemented in three ways full-volume mirroring, pointer-based full-volume replication, and pointer-based virtual replication.

Full-Volume Mirroring



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 102

In *full-volume mirroring*, the target is attached to the source and established as a mirror of the source. The data on the source is copied to the target. New updates to the source are also updated on the target. After all the data is copied and both the source and the target contain identical data, the target can be considered as a mirror of the source. While the target is attached to the source it remains unavailable to any other host. However, the production host continues to access the source.

After the synchronization is complete, the target can be detached from the source and made available for other business operations. Both the source and the target can be accessed for read and write operations by the production and business continuity hosts respectively. After detaching from the source, the target becomes a point-in-time (PIT) copy of the source. The PIT of a replica is determined by the time when the target is detached from the source. For example, if the time of detachment is 4:00 pm., the PIT for the target is 4:00 pm. After detachment, changes made to both the source and replica can be tracked at some predefined granularity. This enables incremental resynchronization (source to target) or incremental restore (target to source). The granularity of the data change can range from 512 byte blocks to 64 KB blocks or higher.

Pointer-based Full-Volume Replication

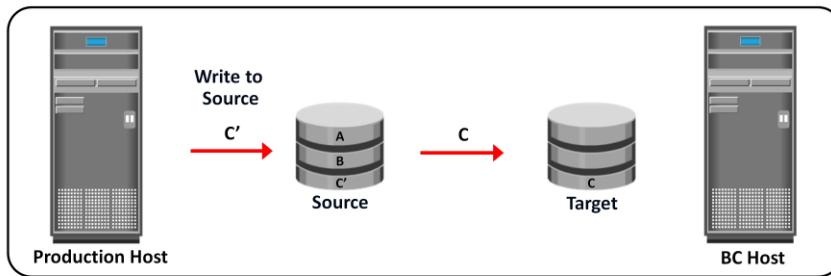
- Provides full copy of source data on the target
- Target device is immediately accessible by the BC host after the replication session is activated
- PIT is determined by time of session activation
- Target device is at least as large as the source device
- Two modes
 - ▶ Full copy mode
 - ▶▶ After session starts, all the data from source is copied to the target in the background
 - ▶ Copy on First Access (deferred)

Another method of array-based local replication is *pointer-based full-volume replication*. Similar to full-volume mirroring, this technology can provide full copies of the source data on the targets. Unlike full-volume mirroring, the target is immediately accessible by the BC host after the replication session is activated. Therefore, data synchronization and detachment of the target is not required to access it. Here, the time of replication session activation defines the PIT copy of the source.

Pointer-based, full-volume replication can be activated in either Copy on First Access (CoFA) mode or Full Copy mode. In either case, at the time of activation, a protection bitmap is created for all data on the source devices. The protection bitmap keeps track of the changes at the source device. The pointers on the target are initialized to map the corresponding data blocks on the source. The data is then copied from the source to the target based on the mode of activation.

In a Full Copy mode, all data from the source is copied to the target in the background. Data is copied regardless of access. If access to a block that has not yet been copied to the target is required, this block is preferentially copied to the target. In a complete cycle of the Full Copy mode, all data from the source is copied to the target. If the replication session is terminated now, the target contains all the original data from the source at the point-in-time of activation. This makes the target a viable copy for restore or other business continuity operations.

Copy on First Access: Write to the Source



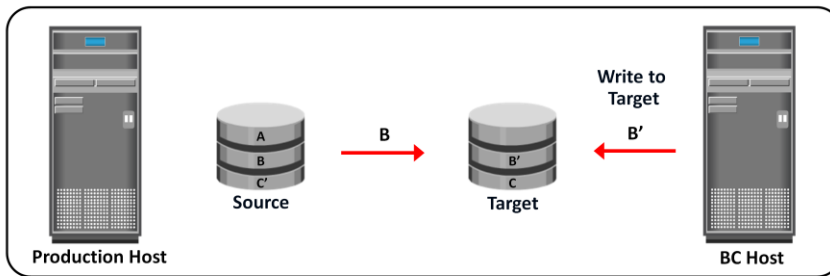
- When a write is issued to the source for the first time after replication session activation:
 - ▶ Original data at that address is copied to the target
 - ▶ Then the new data is updated on the source
 - ▶ This ensures that original data at the point-in-time of activation is preserved on the target

In CoFA, after the replication session is initiated, the data is copied from the source to the target only when the following condition occurs:

- A write I/O is issued to a specific address on the source for the first time.
- A read or write I/O is issued to a specific address on the target for the first time.

When a write is issued to the source for the first time after replication session activation, the original data at that address is copied to the target. After this operation, the new data is updated on the source. This ensures that the original data at the point-in-time of activation is preserved on the target.

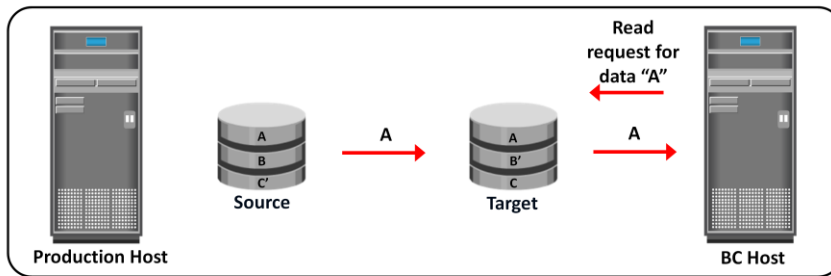
Copy on First Access: Write to the Target



- When a write is issued to the target for the first time after replication session activation:
 - ▶ The original data is copied from the source to the target
 - ▶ Then the new data is updated on the target

When a write is issued to the target for the first time after the replication session activation, the original data is copied from the source to the target. After this, the new data is updated on the target (see figure on the slide).

Copy on First Access: Read from Target



- When a read is issued to the target for the first time after replication session activation:
 - ▶ The original data is copied from the source to the target and is made available to the BC host

When a read is issued to the target for the first time after replication session activation, the original data is copied from the source to the target and is made available to the BC host.

In all cases, the protection bit for the data block on the source is reset to indicate that the original data has been copied over to the target. The pointer to the source data can now be discarded. Subsequent writes to the same data block on the source, and the reads or writes to the same data blocks on the target, do not trigger a copy operation, therefore this method is termed “Copy on First Access.”

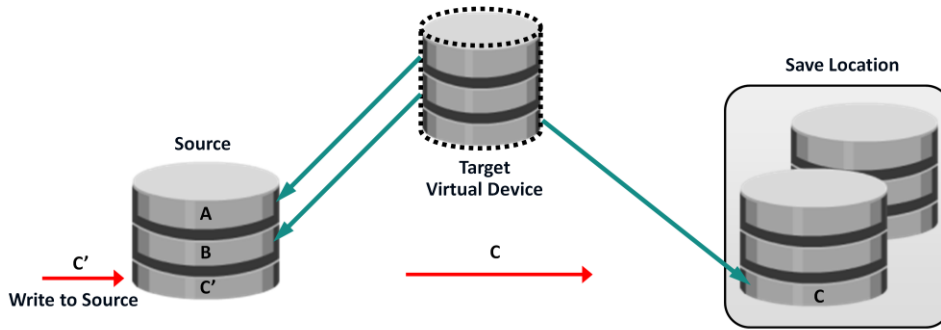
If the replication session is terminated, then the target device has only the data that was accessed until the termination, not the entire contents of the source at the point-in-time. In this case, the data on the target cannot be used for restore because it is not a full replica of the source.

Pointer-based Virtual Replication

- Targets do not hold data, but hold pointers to where the data is located
 - ▶ At the start of the session the target device holds pointers to data on source device
 - ▶ Target requires a small fraction of the size of the source volumes
- Target devices are accessible immediately when the session is started
- Uses CoFW principle
- This method is recommended, if the changes to the source are typically less than 30%

In *pointer-based virtual replication*, at the time of the replication session activation, the target contains pointers to the location of the data on the source. The target does not contain data at any time. Therefore, the target is known as a *virtual replica*. Similar to pointer-based full-volume replication, the target is immediately accessible after the replication session activation. This replication method uses CoFW technology and typically recommended when the changes to the source are less than 30%.

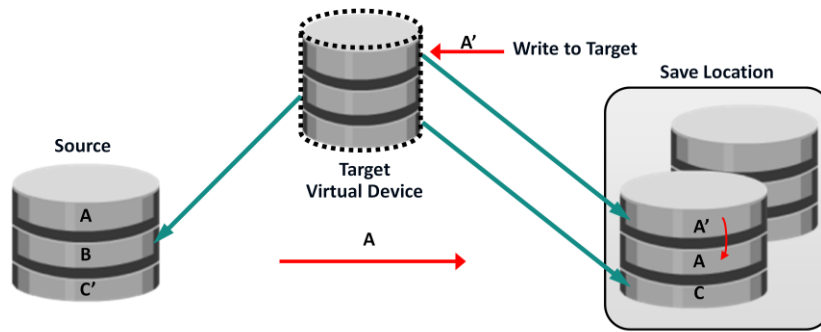
Pointer-based Virtual Replication (CoFW): Write to Source



- When a write is issued to the source for the first time after replication session activation:
 - ▶ Original data at that address is copied to save location
 - ▶ The pointer in the target is updated to point to this data in the save location
 - ▶ Finally, the new write is updated on the source

When a write is issued to the source for the first time after the replication session activation, the original data at that address is copied to a predefined area in the array. This area is generally known as the *save location*. The pointer in the target is updated to point to this data in the save location. After this, the new write is updated on the source.

Pointer-based Virtual Replication (CoFW): Write to Target



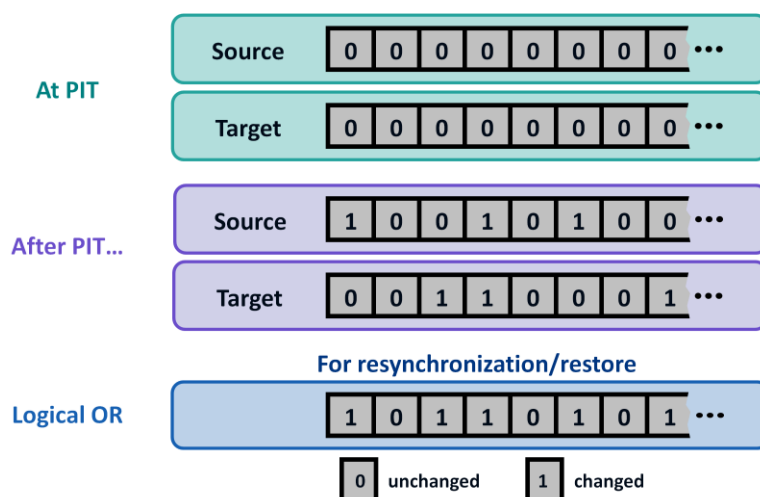
- When a write is issued to the target for the first time after replication session activation:
 - ▶ Original data from the source device is copied to the save location
 - ▶ The pointer is updated to the data in save location
 - ▶ Another copy of the original data is created in the save location before the new write is updated on the save location

When a write is issued to the target for the first time after replication session activation, the data is copied from the source to the save location, and the pointer is updated to the data in the save location. Another copy of the original data is created in the save location before the new write is updated on the save location. Subsequent writes to the same data block on the source or target do not trigger a copy operation.

When reads are issued to the target, unchanged data blocks since the session activation are read from the source, whereas data blocks that have changed are read from the save location.

Data on the target is a combined view of unchanged data on the source and data on the save location. Unavailability of the source device invalidates the data on the target. The target contains only pointers to the data, and therefore, the physical capacity required for the target is a fraction of the source device. The capacity required for the save location depends on the amount of the expected data change.

Tracking Changes to Source and Target



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 110

Updates can occur on the source device after the creation of point-in-time local replicas. If the primary purpose of local replication is to have a viable point-in-time copy for data recovery or restore operations, then the replica devices should not be modified. Changes can occur on the replica device if it is used for other business operations. To enable incremental resynchronization or restore operations, changes to both the source and replica devices after the point-in-time should be tracked. This is typically done using bitmaps, where each bit represents a block of data. The data block sizes can range from 512 bytes to 64 KB or greater. For example, if the block size is 32 KB, then a 1 GB device would require 32,768 bits (1GB divided by 32KB). The size of the bitmap would be 4 KB. If the data in any 32 KB block is changed, the corresponding bit in the bitmap is flagged. If the block size is reduced for tracking purposes, then the bitmap size increases correspondingly.

The bits in the source and target bitmaps are all set to 0 (zero) when the replica is created. Any changes to the source or replica are then flagged by setting the appropriate bits to 1 in the bitmap. When resynchronization or restore is required, a *logical OR* operation between the source bitmap and the target bitmap is performed. The bitmap resulting from this operation references all blocks that have been modified in either the source or replica. This enables an optimized resynchronization or a restore operation, because it eliminates the need to copy all the blocks between the source and the replica. The direction of data movement depends on whether a resynchronization or a restore operation is performed. If resynchronization is required, changes to the replica are overwritten with the corresponding blocks from the source. If a restore is required, changes to the source are overwritten with the corresponding blocks from the replica. In either case, changes to both the source and the target cannot be simultaneously preserved.

Restore and Restart Considerations

- Source has a failure
 - Logical corruption or physical failure of source devices
 - Solution
 - Restore data from target to source
 - Restore would typically be done incrementally
 - Applications can be restarted even before synchronization is complete
- OR-----
- Start production on target
 - Create a “Gold” copy of target device before restarting on target
 - Resolve issues with source while continuing operations on target
 - After resolving the issue, restore latest data on target to source

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 111

Local replicas are used to restore data to production devices. Alternatively, applications can be restarted using the consistent point-in-time replicas.

Replicas are used to restore data to the production devices if logical corruption of data on production devices occurs—that is, the devices are available but the data on them is invalid. Examples of logical corruption include accidental deletion of data (tables or entries in a database), incorrect data entry, and incorrect data updates. Restore operations from a replica are incremental and provide a small RTO. In some instances, the applications can be resumed on the production devices prior to the completion of the data copy. Prior to the restore operation, access to production and replica devices should be stopped.

Production devices might also become unavailable due to physical failures, such as production server or physical drive failure. In this case, applications can be restarted using the data on the latest replica. As a protection against further failures, a “Gold Copy” (another copy of replica device) of the replica device should be created to preserve a copy of data in the event of failure or corruption of the replica devices. After the issue has been resolved, the data from the replica devices can be restored back to the production devices.

Full-volume replicas (both full-volume mirrors and pointer-based in Full Copy mode) can be restored to the original source devices or to a new set of source devices. Restores to the original source devices can be incremental, but restores to a new set of devices are full-volume copy operations.

In pointer-based virtual and pointer-based full-volume replication in CoFA mode, access to data on the replica is dependent on the health and accessibility of the source volumes. If the source volume is inaccessible for any reason, these replicas cannot be used for a restore or a restart operation.

Comparison of Local Replication Technologies

Factor	Full-Volume Mirroring	Pointer-based Full-Volume Replication	Pointer-based Virtual Replication
Performance impact on source due to replica	No impact	Full copy mode – no impact CoFA mode – some impact	High impact
Size of target	At least the same as the source	At least the same as the source	Small fraction of the source
Availability of source for restoration	Not required	Full copy mode – not required CoFA mode – required	Required
Accessibility to target	Only after synchronization and detachment from the source	Immediately accessible	Immediately accessible

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 112

This table summarizes the comparison between full-volume mirroring, pointer-based full-volume, and pointer-based virtual replication technologies.

Note:

Most array based replication technologies allow the source devices to maintain replication relationships with multiple targets.

- More frequent replicas will reduce the RPO.
- Each PIT could be used for a different BC activity and also as restore points.

Network-based Local Replication: Continuous Data Protection

- Replication occurs at the network layer between the hosts and storage arrays
 - ▶ Ideal for highly heterogeneous environment
- Typically provides the ability to restore data to any previous point-in-time
 - ▶ RPOs are random and do not need to be defined in advance
- Data changes are continuously captured and stored in a separate location from the production data

In network-based replication, the replication occurs at the network layer between the hosts and storage arrays. Network-based replication combines the benefits of array-based and host-based replications. By offloading replication from servers and arrays, network-based replication can work across a large number of server platforms and storage arrays, making it ideal for highly heterogeneous environments. *Continuous data protection* (CDP) is a technology used for network-based local and remote replications. CDP for remote replication is detailed in module 12.

In a data center environment, mission-critical applications often require instant and unlimited data recovery points. Traditional data protection technologies offer limited recovery points. If data loss occurs, the system can be rolled back only to the last available recovery point. Mirroring offers continuous replication; however, if logical corruption occurs to the production data, the error might propagate to the mirror, which makes the replica unusable. In normal operation, CDP provides the ability to restore data to any previous PIT. It enables this capability by tracking all the changes to the production devices and maintaining consistent point-in-time images.

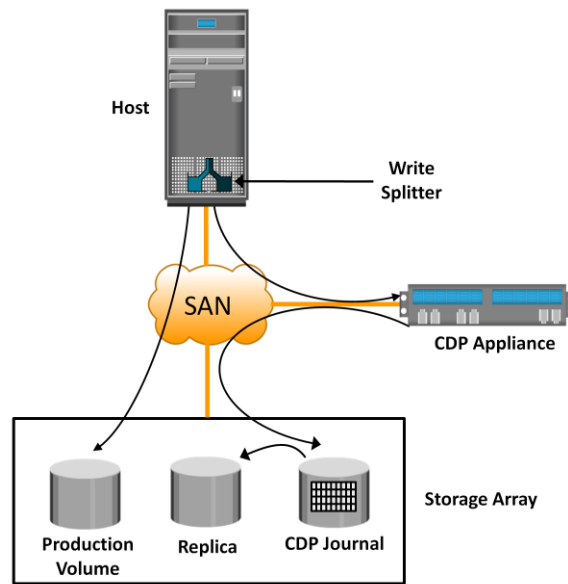
In CDP, data changes are continuously captured and stored in a separate location from the primary storage. Moreover, RPOs are random and do not need to be defined in advance. With CDP, recovery from data corruption poses no problem because it allows going back to a PIT image prior to the data corruption incident.

Continuous Data Protection: Components

- Journal volume
 - ▶ Stores all data changes on the primary storage
- CDP appliance
 - ▶ Intelligent hardware platform that runs the CDP software and manages local and remote data replications
- Write splitter
 - ▶ Intercept writes to the production volume from the host and split each write into two copies

CDP uses a *journal volume* to store all data changes on the primary storage. The journal volume contains all the data that has changed from the time the replication session started. The amount of space that is configured for the journal determines how far back the recovery points can go. CDP also uses *CDP appliance* and *write splitter*. CDP implementation may also be host-based in which CDP software is installed on a separate host machine. CDP appliance is an intelligent hardware platform that runs the CDP software and manages local and remote data replications. Write splitters intercept writes to the production volume from the host and split each write into two copies. Write splitting can be performed at the host, fabric, or storage array.

CDP Local Replication Operation



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 115

Figure on the slide describes CDP local replication. In this method, before the start of replication, the replica is synchronized with the source and then the replication process starts. After the replication starts, all the writes to the source are split into two copies. One of the copies is sent to the CDP appliance and the other to the production volume. When the CDP appliance receives a copy of a write, it is written to the journal volume along with its timestamp. As a next step, data from the journal volume is sent to the replica at predefined intervals.

While recovering data to the source, the CDP appliance restores the data from the replica and applies journal entries up to the point in time chosen for recovery.

Module 11: Local Replication

Lesson 3: Local Replication in Virtualized Environment

During this lesson the following topics are covered:

- Mirroring of a virtual volume
- Replication of virtual machines

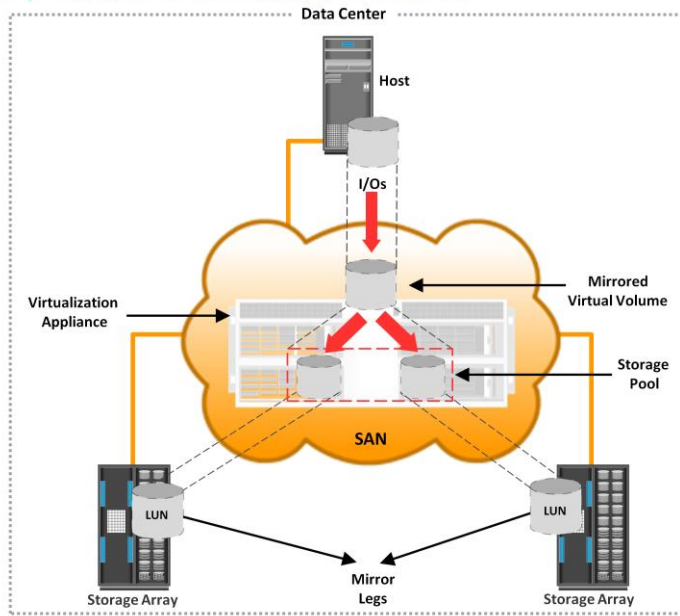
This lesson covers mirroring of a virtual volume and also replication of virtual machines.

Local Replication in Virtualized Environment

- Local replication (mirroring) of a virtual volume assigned to a host
 - ▶ Mirroring is performed by a virtualization appliance
- Replication of virtual machines
 - ▶ VM snapshot
 - ▶ VM clone

The discussion so far has focused on local replication in a physical infrastructure environment. In a virtualized environment, along with replicating storage volumes, replication of virtual machine (VM) is performed by the hypervisor. For hypervisor-based local replication, two options are available - VM Snapshot, and VM Clone. However, in a virtualized environment, the data in the virtual volumes can also be replicated with the help of a virtualization layer in a SAN.

Local Replication of Virtual Volume



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 118

In a virtualized environment, the virtualization appliance at the SAN abstracts the LUNs and creates virtual volumes. These virtual volumes are presented to the host. The virtualization appliance has the ability to mirror the data of a virtual volume between the LUNs. Figure on the slide provides an illustration of a virtual volume that is mirrored between arrays within a data center. Each I/O to the virtual volume is mirrored to the underlying LUNs on the arrays. If one of the arrays incurs an outage, the virtualization appliance will be able to continue processing I/O on the surviving mirror leg. Upon restoration of the failed storage array, the data from the surviving LUN is resynchronized to the recovered leg.

VM Snapshot

- Captures the state and data of a running VM at a specific PIT
- Uses a separate delta file to record all the changes to the virtual disk since the snapshot session is activated
- Restores all settings configured in a guest OS to the PIT

VM Snapshot captures the state and data of a running virtual machine at a specific point in time. The VM state includes VM files, such as BIOS, network configuration, and its power state (powered-on, powered-off, or suspended). The VM data includes all the files that make up the VM, including virtual disks and memory. A VM Snapshot uses a separate delta file to record all the changes to the virtual disk since the snapshot session is activated. Snapshots are useful when a VM needs to be reverted to the previous state in the event of logical corruptions. Reverting a VM to a previous state causes all settings configured in the guest OS to be reverted to that PIT when that snapshot was created. There are some challenges associated with the VM Snapshot technology. It does not support data replication if a virtual machine accesses the data by using raw disks. Also, using the hypervisor to perform snapshots increases the load on the compute and impacts the compute performance.

VM Clone

- An identical copy of an existing VM
 - ▶ Clones are created for different use such as testing
 - ▶ Changes made to a clone VM do not affect the parent VM and vice versa
- Clone VM is assigned a separate network identity
 - ▶ Clone has its own separate MAC address
- Useful when multiple identical VMs need to deploy

VM Clone is another method that creates an identical copy of a virtual machine. When the cloning operation is complete, the clone becomes a separate VM from its parent VM. The clone has its own MAC address, and changes made to a clone do not affect the parent VM. Similarly, changes made to the parent VM do not appear in the clone. VM Clone is a useful method when there is a need to deploy many identical VMs. Installing guest OS and applications on multiple VMs is a time-consuming task; VM Clone helps to simplify this process.

Module 11: Local Replication

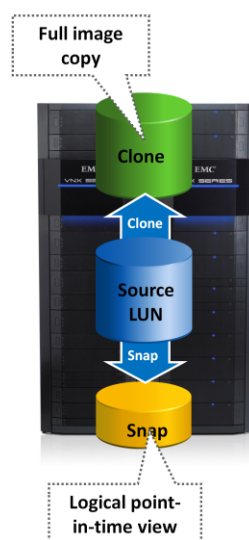
Concept in Practice

- EMC SnapView
- EMC TimeFinder
- EMC RecoverPoint

The concepts in practice section covers various EMC local replication products such as EMC SnapView, EMC TimeFinder, and EMC RecoverPoint.

EMC SnapView

- SnapView Snapshot
 - ▶ Logical view of the production volume
 - ▶ Uses CoFW principle
- SnapView Clone
 - ▶ Full volume copies that require same disk space as the source
 - ▶ Becomes a PIT copy once the clone is fractured from the source



SnapView is an EMC VNX array-based local replication software that creates pointer-based virtual copy and full-volume mirror of the source using SnapView snapshot and SnapView clone respectively.

SnapView Snapshot

A SnapView snapshot is not a full copy of the production volume; it is a logical view of the production volume based on the time at which the snapshot was created. Snapshots are created in seconds, and can be retired when no longer needed. A snapshot “roll back” feature provides instant restore to the source volume. The key terminologies of SnapView snapshot are as follows:

SnapView session: The SnapView snapshot mechanism is activated when a session starts and deactivated when a session stops. A snapshot appears “offline” until there is an active session. Multiple snapshots can be included in a session.

Reserved LUN pool: This is a private area, also called a save area, used to contain CoFW data. The “Reserved” part of the name refers to the fact that the LUNs are reserved and therefore cannot be assigned to a host.

SnapView Clone

SnapView Clones are full-volume copies that require the same disk space as the source. These PIT copies can be used for other business operations, such as backup and testing. SnapView Clone enables incremental resynchronization between the source and replica. Clone fracture is the process of breaking off a clone from its source. After the clone is fractured, it becomes a PIT copy and available for other business operations.

EMC TimeFinder

- TimeFinder/Snap
 - ▶ Creates space-saving, logical PIT (snapshots)
 - ▶ Allows creating multiple snapshots from a single source
- TimeFinder/Clone
 - ▶ Creates PIT copy of the source volume
 - ▶ Uses pointer-based full-volume replication technology
 - ▶ Allows creating multiple clones from a single source device

The TimeFinder family of products consists of two base solutions and four add-on solutions. The base solutions are TimeFinder/Clone and TimeFinder/Snap. The add-on solutions are TimeFinder/Clone Emulation, TimeFinder/Consistency Groups, TimeFinder/Exchange Integration Module, and TimeFinder/SQL Integration Module. TimeFinder is available for both open systems and mainframes. The base solutions support the different storage array-based local replication technologies discussed in this module. The add-on solutions are customizations of the replicas for specific application or database environments.

TimeFinder/Snap

TimeFinder/Snap creates space-saving, logical PIT images called snapshots. The snapshots are not full copies, but contain pointers to the source data. The target device used by TimeFinder/Snap is called a virtual device (VDEV). It keeps pointers to the source device or SAVE devices. The SAVE devices keep the point-in-time data that has changed on the source after the start of the replication session. TimeFinder/Snap allows creating multiple snapshots from a single source device.

TimeFinder/Clone

TimeFinder/Clone creates a PIT copy of the source volume that can be used for backups, decision support, or any other process that requires parallel access to production data. TimeFinder/Clone uses pointer-based full-volume replication technology. TimeFinder/Clone allows creating multiple clones from a single production device, and all the clones are available immediately for read and write access.

EMC RecoverPoint

- Provides continuous data protection and recovery to any PIT
- Uses splitting technology at server, fabric, or array to mirror a write to a RecoverPoint appliance
- Provides automatic RecoverPoint appliance failover
- Family of product includes
 - ▶ RecoverPoint/CL
 - ▶ RecoverPoint/EX
 - ▶ RecoverPoint/SE

RecoverPoint is a high-performance, cost-effective, single product that provides local and remote data protection for both physical and virtual environments. It provides faster recovery and unlimited recovery points. RecoverPoint provides continuous data protection and performs replication between the LUNs. RecoverPoint uses lightweight splitting technology either at the application server, fabric, or arrays to mirror a write to a RecoverPoint appliance. The RecoverPoint family of products include RecoverPoint/CL, RecoverPoint/EX, and RecoverPoint/SE.

RecoverPoint/CL is a replication product for a heterogeneous server and storage environment. It supports both EMC and non-EMC storage arrays. This product supports host-based, fabric-based, and array-based write splitters. *RecoverPoint/EX* supports replication between EMC storage arrays and allows only array-based write splitting.

RecoverPoint/SE is a version of RecoverPoint that is targeted for VNX series arrays and enables only Windows-based host and array-based write splitting.

Module 11: Summary

Key points covered in this module:

- Uses of local replicas
- Consistency in file system and database replication
- Host-based, storage array-based, and network-based replication
- Restore and restart considerations
- Local replication of a virtual volume
- VM snapshot and VM clone

This module covered various uses of local replica along with file system and database consistency in replication. This module also covered various host-based, storage-based, and network-based local replication technologies. Further, this module focused restore and restart considerations for local replica. Finally, this module detailed local replication in virtualized environment.

Local replicas are used for several purposes such as alternate source for backup, fast recovery, decision support, testing platform, and data migration.

Replica must be consistent with the source so that it is usable for both recovery and restart operations. Ensuring consistency is the primary requirement for all the replication technologies.

Host-based, storage array-based, and network-based replications are the major technologies used for local replication. File system replication and LVM-based replication are examples of host-based local replication. Storage array-based replication can be implemented with distinct solutions, namely, full-volume mirroring, pointer-based full-volume replication, and pointer-based virtual replication. Continuous Data Protection (CDP) is an example of network-based replication.

In a virtualized environment, virtual machine (VM) replication is a critical requirement for successful BC process. Typically, local replication of VMs (VM snapshot and VM clone) is performed by the hypervisor at the compute level. However, it can also be performed at the storage level using array-based local replication, similar to the physical environment.

Check Your Knowledge – 1

- What is an advantage of pointer-based virtual replication?
 - A. Source device need not be healthy for restore
 - B. Save location is not required to activate session
 - C. Less storage space is required for creating replica
 - D. No performance impact on source due to replication
- Which is true about pointer-based full-volume replication?
 - A. Size of the target is a small fraction of the source
 - B. Size of the target is at least the same as the source
 - C. Target can be accessed only after synchronization and detachment from source
 - D. Target contains only pointers to save location at all time

Check Your Knowledge – 2

- In continuous data protection technology, which factor determines how far back the recovery points can go?
 - A. Amount of space that is configured for the replica
 - B. Type of write splitter used in replication
 - C. Amount of space that is configured for the journal
 - D. Rate of changes happening to the replica
- When is the data copied from source to target in CoFA mode of replication?
 - A. A read occurs for the first time from a location on the source
 - B. A read or write occurs for the first time to a location on the source
 - C. A read or write occurs for the first time to a location on the target
 - D. All writes issued to a location on the target

Check Your Knowledge – 3

- What occurs to the guest OS configuration when a VM is reverted from its snapshot?
 - A. Guest OS configurations are reverted to the PIT of snapshot creation
 - B. Current guest OS configurations are preserved
 - C. Guest OS configurations are duplicated to a new VM
 - D. Guest OS settings are lost and need manual configuration

Exercise: Local Replication

- Scenario
 - Organization's mission critical data is stored on RAID 1 volumes
 - Database application uses 1TB storage
 - Average data that changes in 24 hours is 60 GB
- Requirements
 - Need solution to address logical corruption of database
 - Maximum RPO of 1 hour
 - Solution should support restore request for up to 8 hours old data
 - Minimize the amount of storage used for data protection
- Task
 - Suggest an appropriate local replication solution to meet RPO requirement with minimum amount of storage
 - Estimate the physical storage required by this solution

Scenario:

A manufacturing organization stores data of their mission critical applications on a high end storage array with RAID 1 configuration. The database application has 1 TB of storage and needs 24x7 availability. Average data that changes in 24 hours is 60 GB.

Requirements:

Need solution to address logical corruption of database

Maximum RPO of 1 hour

Solution should support restore request for up to 8 hours old data

Minimize the amount of storage used for data protection

Task:

Suggest an appropriate local replication solution to meet RPO requirement with minimum amount of storage. Estimate the physical storage required by this solution.

Module – 12

Remote Replication



Module 12: Remote Replication

Upon completion of this module, you should be able to:

- Explain synchronous and asynchronous replication mode
- Describe host-based, array-based, and network-based remote replication technologies
- Describe three-site remote replication
- Explain data migration solution
- Describe remote replication and migration in virtualized environment

This module focuses on synchronous and asynchronous remote replication mode. This module also focuses on host-based, array-based, and network-based remote replication technologies. Further, this module details three-site remote replication and data migration solution. Additionally, this module details remote replication and migration in virtualized environment.

Module 12: Remote Replication

Lesson 1: Remote Replication Overview

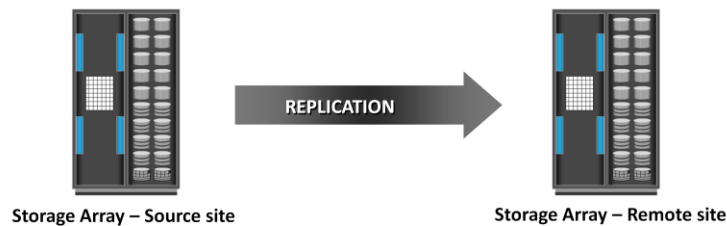
During this lesson the following topics are covered:

- Synchronous and asynchronous remote replication
- Bandwidth requirement for synchronous and asynchronous remote replication

This lesson covers synchronous and asynchronous remote replication mode. This lesson also covers on bandwidth requirement for both synchronous and asynchronous remote replication.

What is Remote Replication?

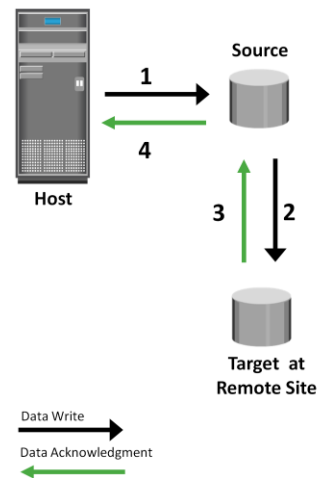
- Process of creating replicas at remote sites
 - ▶ Addresses risk associated with regionally driven outages
- Modes of remote replication
 - ▶ Synchronous
 - ▶ Asynchronous



Remote replication is the process to create replicas of information assets at remote sites (locations). Remote replication helps organizations mitigate the risks associated with regionally driven outages resulting from natural or human-made disasters. During disasters, the workload can be moved to a remote site to ensure continuous business operation. Similar to local replicas, remote replicas can also be used for other business operations. Two modes of remote replications are synchronous and asynchronous replication.

Synchronous Replication – 1

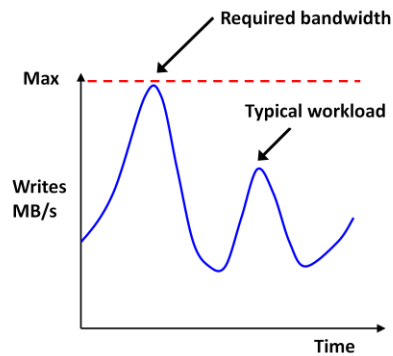
- A write is committed to both source and remote replica before it is acknowledged to the host
- Ensures source and replica have identical data at all times
 - ▶ Maintains write ordering
- Provides near-zero RPO



In *synchronous remote replication*, writes must be committed to the source and remote replica (or target), prior to acknowledging “write complete” to the host. Additional writes on the source cannot occur until each preceding write has been completed and acknowledged. This ensures that data is identical on the source and replica at all times. Further, writes are transmitted to the remote site exactly in the order in which they are received at the source. Therefore, write ordering is maintained. If a source-site failure occurs, synchronous remote replication provides zero or near-zero recovery-point objective (RPO).

Synchronous Replication – 2

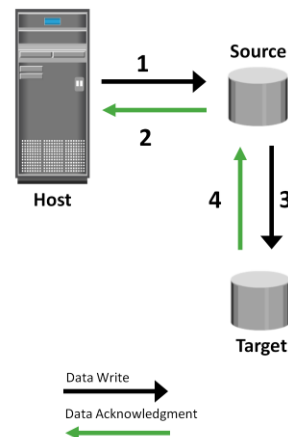
- Response time depends on bandwidth and distance
- Requires bandwidth more than the maximum write workload
- Typically deployed for distance less than 200 km (125 miles) between two sites



Application response time is increased with synchronous remote replication because writes must be committed on both the source and target before sending the “write complete” acknowledgment to the host. The degree of impact on response time depends primarily on the distance between sites, bandwidth and quality of service (QOS) of the network connectivity infrastructure. If the bandwidth provided for synchronous remote replication is less than the maximum write workload, there will be times during the day when the response time might be excessively elongated, causing applications to time out. The distances over which synchronous replication can be deployed depend on the application’s capability to tolerate extensions in response time. Typically, it is deployed for distances less than 200 KM (125 miles) between the two sites.

Asynchronous Replication – 1

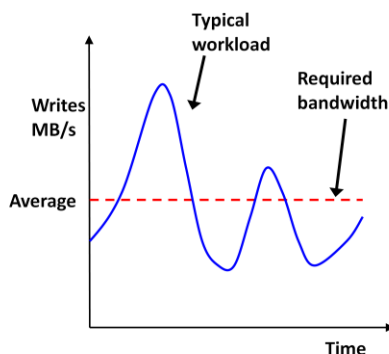
- A write is committed to the source and immediately acknowledged to the host
- Data is buffered at the source and transmitted to the remote site later
- Finite RPO
 - ▶ Replica will be behind the source by a finite amount



In *asynchronous remote replication*, a write is committed to the source and immediately acknowledged to the host. In this mode, data is buffered at the source and transmitted to the remote site later. Data at the remote site will be behind the source by at least the size of the buffer. Hence, asynchronous remote replication provides a finite (nonzero) RPO disaster recovery solution.

Asynchronous Replication – 2

- RPO depends on size of buffer and available network bandwidth
- Requires bandwidth equal to or greater than average write workload
- Sufficient buffer capacity should be provisioned
- Can be deployed over long distances



Asynchronous replication eliminates the impact to the application's response time because the writes are acknowledged immediately to the source host. This enables deployment of asynchronous replication over distances ranging from several hundred to several thousand kilometers between the primary and remote sites. In this case, the required bandwidth can be provisioned equal to or greater than the average write workload. Data can be buffered during times when the bandwidth is not enough and moved later to the remote site.

Therefore, sufficient buffer capacity should be provisioned. RPO depends on the size of the buffer, the available network bandwidth, and the write workload to the source.

Asynchronous replication implementation can take advantage of *locality of reference* (repeated writes to the same location). If the same location is written multiple times in the buffer prior to transmission to the remote site, only the final version of the data is transmitted. This feature conserves link bandwidth.

Note: In both synchronous and asynchronous modes of replication, only writes to the source are replicated; reads are still served from the source.

Module 12: Remote Replication

Lesson 2: Remote Replication Technologies

During this lesson the following topics are covered:

- Host-based, storage array-based, and network-based remote replication technologies
- Three-site remote replication
- Data migration solution
- Remote replication in virtualized environment

This lesson focuses on host-based, array-based, and network-based remote replication technologies. Further this lesson focuses on three-site remote replication and data migration solution. Additionally this lesson also details remote replication and migration in virtualized environment.

Host-based Remote Replication: LVM-based Replication

- Replication is performed by host-based software
- All writes to the source volume group are replicated to the target volume group by the LVM
- Can be synchronous or asynchronous

Host-based remote replication uses the host resources to perform and manage the replication operation. There are two basic approaches to host-based remote replication: Logical volume manager (LVM) based replication and database replication via log shipping.

LVM-based remote replication is performed and managed at the volume group level. Writes to the source volumes are transmitted to the remote host by the LVM. The LVM on the remote host receives the writes and commits them to the remote volume group. Prior to the start of replication, identical volume groups, logical volumes, and file systems are created at the source and target sites. Initial synchronization of data between the source and replica is performed. One method to perform initial synchronization is to backup the source data and restore the data to the remote replica. Alternatively, it can be performed by replicating over the IP network. Until the completion of the initial synchronization, production work on the source volumes is typically halted. After the initial synchronization, production work can be started on the source volumes and replication of data can be performed over an existing standard IP network. LVM-based remote replication supports both synchronous and asynchronous modes of replication. If a failure occurs at the source site, applications can be restarted on the remote host, using the data on the remote replicas.

LVM-based remote replication is independent of the storage arrays and therefore supports replication between heterogeneous storage arrays. Most operating systems are shipped with LVMs, so additional licenses and specialized hardware are not typically required. The replication process adds overhead on the host CPUs. CPU resources on the source host are shared between replication tasks and applications. This might cause performance degradation to the applications running on the host.

Host-based Remote Replication: Log Shipping

- Commonly used in a database environment
- All relevant components of source and target databases are synchronized prior to the start of replication
- Transactions to source database are captured in logs and periodically transferred to remote host

Database replication via log shipping is a host-based replication technology supported by most databases. Transactions to the source database are captured in logs, which are periodically transmitted by the source host to the remote host. The remote host receives the logs and applies them to the remote database. Prior to starting production work and replication of log files, all relevant components of the source database are replicated to the remote site. This is done while the source database is shut down.

After this step, production work is started on the source database. The remote database is started in a standby mode. Typically, in standby mode, the database is not available for transactions. All DBMSs switch log files at preconfigured time intervals or when a log file is full. The current log file is closed at the time of log switching, and a new log file is opened. When a log switch occurs, the closed log file is transmitted by the source host to the remote host. The remote host receives the log and updates the standby database. This process ensures that the standby database is consistent up to the last committed log. RPO at the remote site is finite and depends on the size of the log and the frequency of log switching. Available network bandwidth, latency, rate of updates to the source database, and the frequency of log switching should be considered when determining the optimal size of the log file.

Similar to LVM-based remote replication, the existing standard IP network can be used for replicating log files. Host-based log shipping requires low network bandwidth because it transmits only the log files at regular intervals.

Storage Array-based Remote Replication

- Replication is performed by array-operating environment
- Three replication methods: synchronous, asynchronous, and disk buffered
- Synchronous
 - ▶ Writes are committed to both source and replica before it is acknowledged to host
- Asynchronous
 - ▶ Writes are committed to source and immediately acknowledged to host
 - ▶ Data is buffered at source and transmitted to remote site later

In *storage array-based remote replication*, the array-operating environment and resources perform and manage data replication. This relieves the burden on the host CPUs, which can be better used for applications running on the host. A source and its replica device reside on different storage arrays. Data can be transmitted from the source storage array to the target storage array over a shared or a dedicated network. Replication between arrays may be performed in synchronous, asynchronous, or disk-buffered modes.

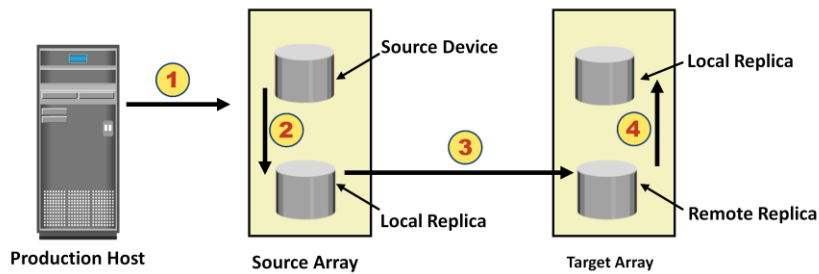
In array-based synchronous remote replication, writes must be committed to the source and the target prior to acknowledging “write complete” to the production host. Additional writes on that source cannot occur until each preceding write has been completed and acknowledged. To optimize the replication process and to minimize the impact on application response time, the write is placed on cache of the two arrays. The storage arrays destage these writes to the appropriate disks later. If the network links fail, replication is suspended; however, production work can continue uninterrupted on the source storage array. The array operating environment keeps track of the writes that are not transmitted to the remote storage array. When the network links are restored, the accumulated data is transmitted to the remote storage array. During the time of network link outage, if there is a failure at the source site, some data will be lost, and the RPO at the target will not be zero.

Cont...

In *asynchronous remote replication mode*, a write is committed to the source and immediately acknowledged to the host. Data is buffered at the source and transmitted to the remote site later. The source and the target devices do not contain identical data at all times. The data on the target device is behind that of the source, so the RPO in this case is not zero. Similar to synchronous replication, asynchronous replication writes are placed in cache on the two arrays and are later destaged to the appropriate disks. Some implementations of asynchronous remote replication maintain write ordering. A timestamp and sequence number are attached to each write when it is received by the source. Writes are then transmitted to the remote array, where they are committed to the remote replica in the exact order in which they were buffered at the source. This implicitly guarantees consistency of data on the remote replicas.

In asynchronous remote replication, the writes are buffered for a predefined period of time. At the end of this duration, the buffer is closed, and a new buffer is opened for subsequent writes. All writes in the closed buffer are transmitted together and committed to the remote replica. Asynchronous remote replication provides network bandwidth cost-savings because the required bandwidth is lower than the peak write workload. During times when the write workload exceeds the average bandwidth, sufficient buffer space must be configured on the source storage array to hold these writes.

Storage Array-based Remote Replication – Disk Buffered Mode



- 1 Production host writes data to source device.
- 2 A consistent PIT local replica of the source device is created.
- 3 Data from local replica is transmitted to the remote replica at target.
- 4 Optionally a PIT local replica of the remote replica on the target is created.

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 143

Disk-buffered replication is a combination of local and remote replication technologies. A consistent PIT local replica of the source device is first created. This is then replicated to a remote replica on the target array. Figure on the slide shows the sequence of operations in a disk-buffered remote replication. At the beginning of the cycle, the network links between the two arrays are suspended, and there is no transmission of data. While production application runs on the source device, a consistent PIT local replica of the source device is created. The network links are enabled, and data on the local replica in the source array transmits to its remote replica in the target array. After synchronization of this pair, the network link is suspended, and the next local replica of the source is created. Optionally, a local PIT replica of the remote device on the target array can be created. The frequency of this cycle of operations depends on the available link bandwidth and the data change rate on the source device. Because disk-buffered technology uses local replication, changes made to the source and its replica are possible to track. Therefore, all the resynchronization operations between the source and target can be done incrementally. When compared to synchronous and asynchronous replications, disk-buffered remote replication requires less bandwidth.

Network-based Replication – Continuous Data Protection

- Provides any-point-in-time recovery capability during its normal operation
- Components of CDP
 - ▶ CDP appliance
 - ▶ Write splitter
 - ▶ Journal volume
- CDP appliances are present at both source and remote sites
- Supports both synchronous and asynchronous replication modes

In network-based remote replication, the replication occurs at the network layer between the host and storage array. Continuous data protection technology, discussed in the previous module, also provides solutions for network-based remote replication. In normal operation CDP remote replication, provides any-point-in-time recovery capability, which enables the target LUNs to be rolled back to any previous point in time. Similar to CDP local replication, CDP remote replication typically uses *journal volume*, *CDP appliance* or CDP software installed on a separate host (*host based CDP*), and *write splitter* to perform replication between sites. The CDP appliance is maintained at both source and remote sites.

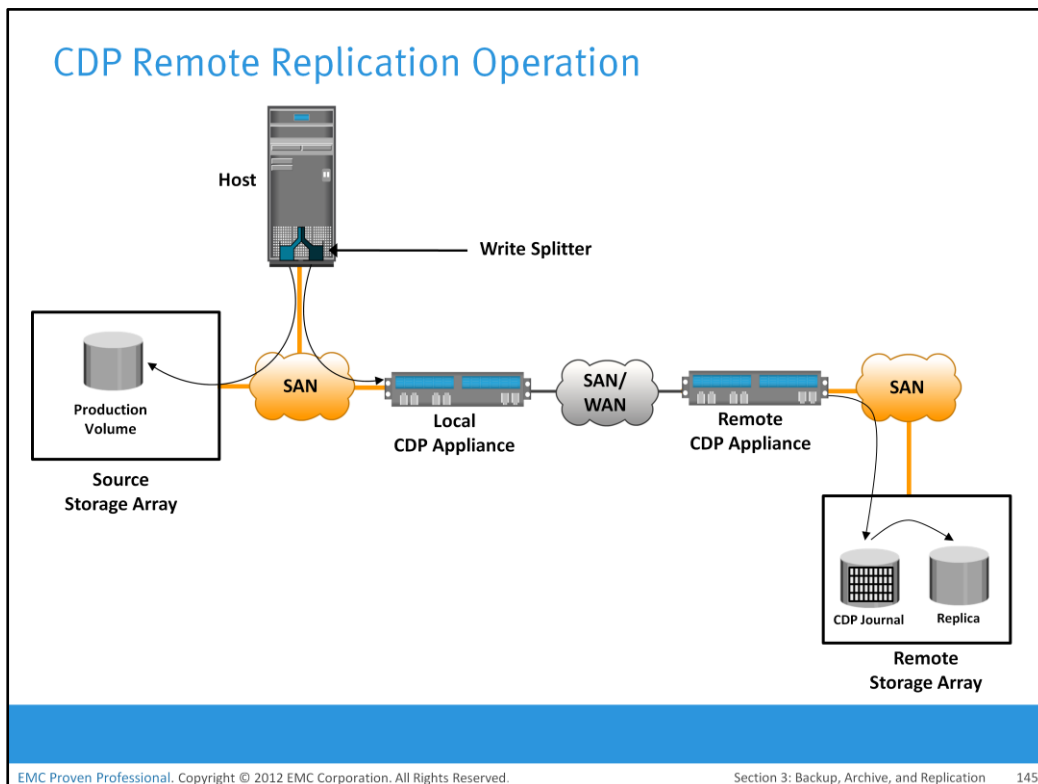


Figure on the slide describes CDP remote replication. In this method, the replica is synchronized with the source, and then the replication process starts. After the replication starts, all the writes from the host to the source are split into two copies. One of the copies is sent to the local CDP appliance at the source site, and the other copy is sent to the production volume. After receiving the write, the appliance at the source site sends it to the appliance at the remote site. Then, the write is applied to the journal volume at the remote site. For an asynchronous operation, writes at the source CDP appliance are accumulated, and redundant blocks are eliminated. Then, the writes are sequenced and stored with their corresponding timestamp. The data is then compressed, and a checksum is generated. It is then scheduled for delivery across the IP or FC network to the remote CDP appliance. After the data is received, the remote appliance verifies the checksum to ensure the integrity of the data. The data is then uncompressed and written to the remote journal volume. As a next step, data from the journal volume is sent to the replica at predefined intervals.

In the asynchronous mode, the local CDP appliance instantly acknowledges a write as soon as it is received. In the synchronous replication mode, the host application waits for an acknowledgment from the CDP appliance at the remote site before initiating the next write. The synchronous replication mode impacts the application's performance under heavy write loads.

Three-site Replication

- Data from source site is replicated to two remote sites
 - ▶ Replication is synchronous to one of the remote sites and asynchronous or disk buffered to the other remote site
- Mitigates the risk in two site replication
 - ▶ No DR protection after source or remote site failure
- Implemented in two ways:
 - ▶ Cascade/multihop
 - ▶ Triangle/multitarget

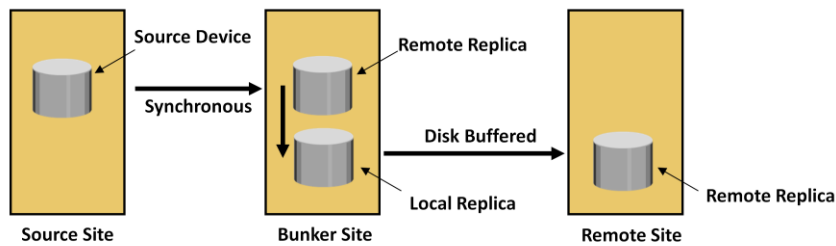
In synchronous replication, the source and target sites are usually within a short distance. Therefore, if a regional disaster occurs, both the source and the target sites might become unavailable. This can lead to extended RPO and RTO because the last known good copy of data would need to come from another source, such as an offsite tape library.

A regional disaster will not affect the target site in asynchronous replication because the sites are typically several hundred or several thousand kilometers apart. If the source site fails, production can be shifted to the target site, but there is no further remote protection of data until the failure is resolved.

Three-site replication mitigates the risks identified in two-site replication. In a three-site replication, data from the source site is replicated to two remote sites. Replication can be synchronous to one of the two sites, providing a near zero-RPO solution, and it can be asynchronous or disk buffered to the other remote site, providing a finite RPO. Three-site remote replication can be implemented as a cascade/multihop or a triangle/multitarget solution.

Three-site Replication: Cascade/Multihop – 1

- Synchronous + Disk Buffered



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

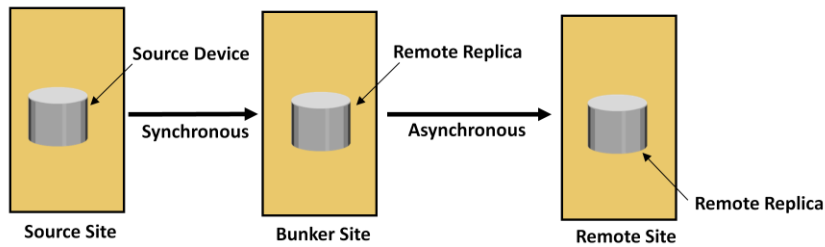
Section 3: Backup, Archive, and Replication 147

In the *cascade/multihop* three-site replication, data flows from the source to the intermediate storage array, known as a *bunker*, in the first hop, and then, from a bunker to a storage array at a remote site in the second hop. Replication can be performed in two ways, synchronous+ disk-buffered or synchronous+asynchronous.

Synchronous + Disk Buffered: This method employs a combination of local and remote replication technologies. Synchronous replication occurs between the source and the bunker; a consistent PIT local replica is created at the bunker. Data is transmitted from the local replica at the bunker to the remote replica at the remote site. Optionally, a local replica can be created at the remote site after data is received from the bunker. In this method, a minimum of 4 storage volumes are required (including the source) to replicate one storage device. RPO at the remote site is usually in the order of hours for this implementation.

Three-site Replication: Cascade/Multihop – 2

- Synchronous + Asynchronous



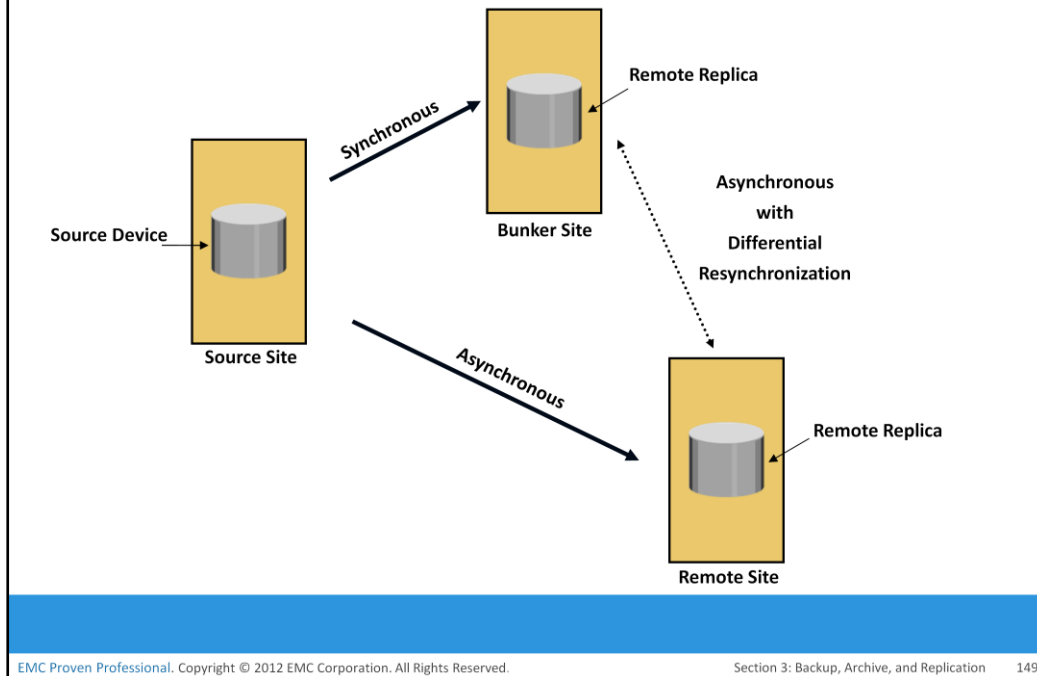
EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 148

Synchronous + Asynchronous: Synchronous replication occurs between the source and the bunker. Asynchronous replication occurs between the bunker and the remote site. The replica in the bunker acts as the source for asynchronous replication to create a remote replica at the remote site. RPO at the remote site is usually in the order of minutes. In this method, a minimum of 3 storage volumes are required (including the source). If a disaster occurs at the source, production operations are failed over to the bunker site with zero or near-zero data loss. If there is a disaster at the bunker site or if there is a network link failure between the source and bunker sites, the source site will continue to operate as normal but without any remote replication. This situation is very similar to remote site failure in a two-site replication solution. The updates to the remote site cannot occur due to the failure in the bunker site. Therefore, the data at the remote site keeps falling behind, but the advantage here is that if the source fails during this time, operations can be resumed at the remote site. RPO at the remote site depends on the time difference between the bunker site failure and source site failure.

A *regional disaster* in cascade/multihop replication is similar to a source site failure in two-site asynchronous replication. Operations are failover to the remote site with an RPO in the order of minutes. There is no remote protection until the regional disaster is resolved. Local replication technologies could be used at the remote site during this time. If a disaster occurs at the remote site, or if the network links between the bunker and the remote site fail, the source site continues to work as normal with disaster recovery protection provided at the bunker site.

Three-site Replication: Triangle/Multitarget



In *three-site triangle/multitarget replication*, data at the source storage array is concurrently replicated to two different arrays at two different sites, as shown in the figure on the slide. The source-to-bunker site (target 1) replication is synchronous with a near-zero RPO. The source-to-remote site (target 2) replication is asynchronous with an RPO in the order of minutes. The distance between the source and the remote sites could be thousands of miles. This implementation does not depend on the bunker site for updating data on the remote site because data is asynchronously copied to the remote site directly from the source. The triangle/multitarget configuration provides consistent RPO unlike cascade/multihop solutions in which the failure of the bunker site results in the remote site falling behind and the RPO increasing. The key benefit of three-site triangle/multitarget replication is the ability to failover to either of the two remote sites in the case of source-site failure, with disaster recovery (asynchronous) protection between the bunker and remote sites. Disaster recovery protection is always available if any one-site failure occurs.

During normal operations, all three sites are available and the production workload is at the source site. At any given instant, the data at the bunker and the source is identical. The data at the remote site is behind the data at the source and the bunker. The replication network links between the bunker and remote sites will be in place but not in use. The difference in the data between the bunker and remote sites is tracked, so that if a source site disaster occurs, operations can be resumed at the bunker or the remote sites with incremental resynchronization between these two sites.

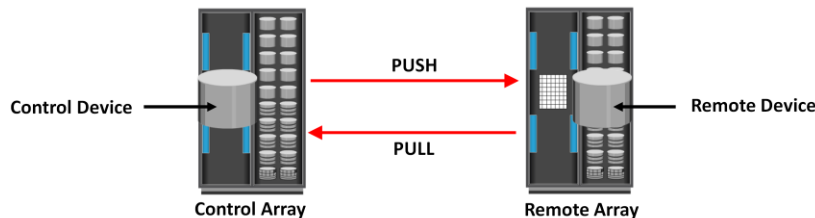
Cont..

A *regional disaster* in three-site triangle/multitarget replication is very similar to a source site failure in two-site asynchronous replication. If failure occurs, operations failover to the remote site with an RPO in the order of minutes. There is no remote protection until the regional disaster is resolved.

A failure of the bunker or the remote site is not really considered as a disaster because the operation can continue uninterrupted at the source site while remote DR protection is still available. A network link failure to either the source-to-bunker or the source-to-remote site does not impact production at the source site while remote DR protection is still available with the site that can be reached.

Data Migration Solution

- Specialized replication technique that enables creating remote point-in-time copies
 - Used for data mobility, migration, and disaster recovery
- Moves data between heterogeneous storage arrays
 - Array performing replication operations is called control array
 - Push: Data is pushed from control array to remote array
 - Pull: Data is pulled to the control array from remote array



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 151

A *data migration and mobility solution* is a specialized replication technique that enables creating remote point-in-time copies. These copies can be used for data mobility, migration, content distribution, and disaster recovery. This solution moves data between heterogeneous storage arrays. This technology is application- and server-operating-system independent because the replication operations are performed by one of the storage arrays.

The array performing the replication operations is called the *control array*. Data can be moved from/to devices in the control array to/from a remote array. The terms control or remote do not indicate the direction of data flow; they indicate only the array that is performing the replication operation. Data migration solutions perform *push* and *pull operations* for data movement. These terms are defined from the perspective of the control array. In the *push* operation, data is moved from the control array to the remote array. In the *pull* operation, data is moved from the remote array to the control array.

Cont..

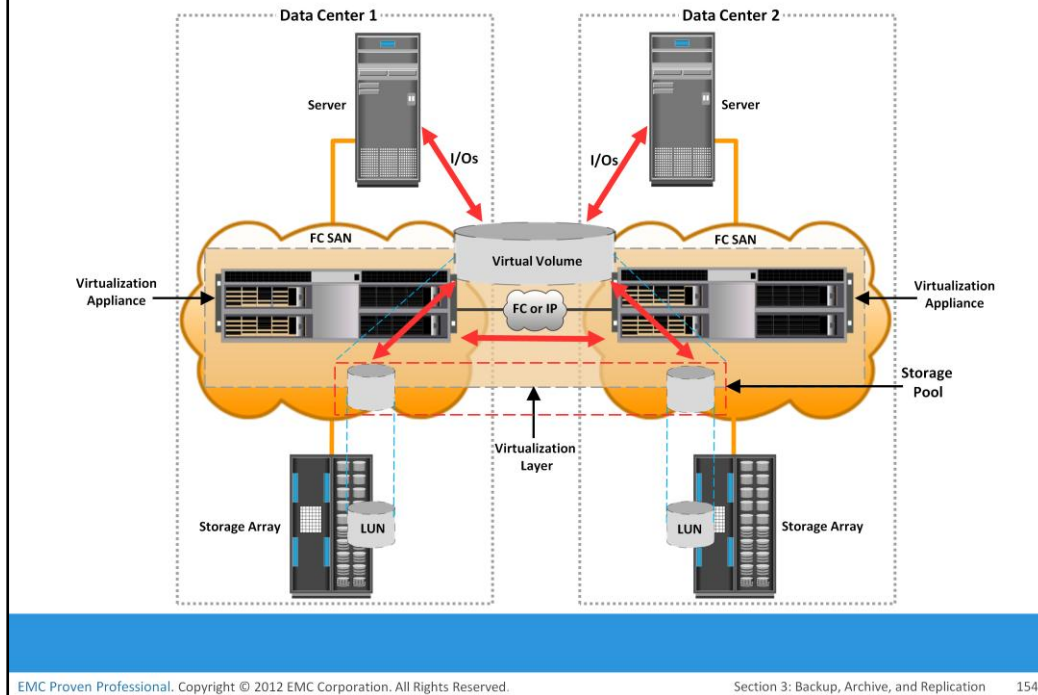
When a push or pull operation is initiated, the control array creates a protection bitmap to track the replication process. Each bit in the protection bitmap represents a data chunk on the control device. When the replication operation is initiated, all the bits are set to one, indicating that all the contents of the source device need to be copied to the target device. As the replication process copies data, the bits are changed to zero, indicating that a particular chunk has been copied. During the push and pull operations, host access to the remote device is not allowed because the control array has no control over the remote array and cannot track any change on the remote device. Data integrity cannot be guaranteed if changes are made to the remote device during the push and pull operations. The push/pull operations can be either *hot* or *cold*. These terms apply to the control devices only. In a cold operation, the control device is inaccessible to the host during replication. Cold operations guarantee data consistency because both the control and the remote devices are offline. In a hot operation, the control device is online for host operations. During hot push/pull operations, changes can be made to the control device because the control array can keep track of all changes, and thus ensure data integrity.

Remote Replication/Migration in Virtualized Environment

- Remote mirroring of virtual volume
 - ▶ Virtual volumes assigned to hosts are mirrored to two different sites
- VM migration
 - ▶ Moving VMs from one location to another without powering off VMs
 - ▶ Commonly used techniques for VM migration are:
 - ▶▶ Hypervisor-to-hypervisor
 - ▶▶ Array-to-array

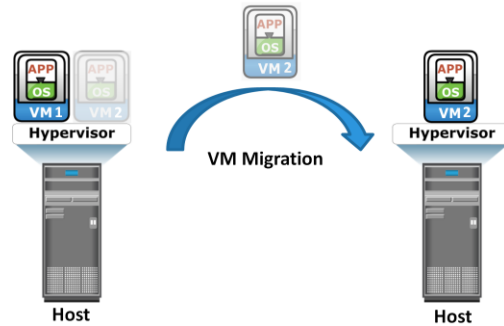
In a virtualized environment, the data residing in a virtual volume is replicated (mirrored) to two storage arrays located at two different sites. Virtual machine migration is another technique used to ensure business continuity in case of hypervisor failure or scheduled maintenance. VM migration is the process of moving VMs from one location to another without powering off the virtual machines. VM migration also helps in load balancing when multiple virtual machines running on the same hypervisor contend for resources. Two commonly used techniques for VM migration are hypervisor-to-hypervisor and array-to-array migration.

Remote Mirroring of Virtual Volume



Virtualization enables mirroring of a virtual volume data between data centers. It provides the capability to connect the virtualization layers (appliance) at multiple data centers. The connected virtualization layers are managed centrally and work as a single virtualization layer that is stretched across data centers. The virtual volumes are created from the federated storage resources across data centers. The virtualization appliance has the ability to mirror the data of a virtual volume between the LUNs located in two different storage arrays at different locations. Figure on the slide provides an illustration of a virtual volume that is mirrored between arrays across data centers. Each I/O from a host to the virtual volume is mirrored to the underlying LUNs on the arrays. If an outage occurs at one of the data centers, the virtualization appliance will be able to continue processing I/O on the surviving mirror leg. Data on the virtual volume can be mirrored synchronously or asynchronously.

VM Migration: Hypervisor-to-Hypervisor

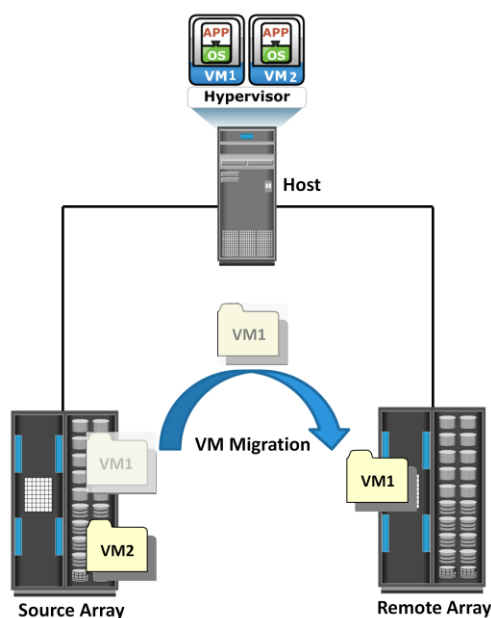


- Active state of a VM is moved from one hypervisor to another
 - ▶ Copies the contents of virtual machine memory from the source hypervisor to the target
- This technique requires source and target hypervisor access to the same storage

In hypervisor-to-hypervisor VM migration, the entire active state of a VM is moved from one hypervisor to another. This method involves copying the contents of virtual machine memory from the source hypervisor to the target and then transferring the control of the VM's disk files to the target hypervisor. Because the virtual disks of the VMs are not migrated, this technique requires both source and target hypervisor access to the same storage.

VM Migration: Array-to-Array

- VM files are moved from source array to remote array
- Can move VMs across dissimilar storage arrays
- Balances storage utilization by redistributing VMs to different storage arrays



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 156

In array-to-array VM migration, VM files are moved from the source array to the remote array. This approach enables the administrator to move VMs across dissimilar storage arrays. Array-to-array migration starts by copying the metadata about the VM from the source array to the target. The metadata essentially consists of configuration, swap, and log files. After the metadata is copied, the VM disk file is replicated to the new location. During replication, there might be a chance that the source is updated; therefore, it is necessary to track the changes on the source to maintain data integrity. After the replication is complete, the blocks that have changed since the replication started are replicated to the new location. Array-to-array VM migration improves performance and balances the storage capacity by redistributing virtual disks to different storage devices.

Module 12: Remote Replication

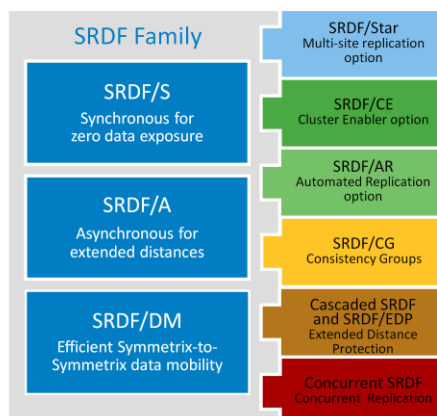
Concept in Practice

- EMC Symmetrix Remote Data Facility (SRDF)
- EMC MirrorView
- EMC RecoverPoint

The concepts in practice section covers various EMC remote replication products such as EMC SRDF, EMC MirrorView, and EMC RecoverPoint.

EMC SRDF

- Offers a family of solutions to implement array-based remote replication
- Minimizes performance impact on applications and hosts



EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Section 3: Backup, Archive, and Replication 158

SRDF offers a family of technology solutions to implement storage array-based remote replication. The SRDF family of software includes the following:

SRDF/Synchronous (SRDF/S): A remote replication solution that creates a synchronous replica at one or more Symmetrix targets located within campus, metropolitan, or regional distances. SRDF/S provides a no-data-loss solution (near zero RPO) if a local disaster occurs.

SRDF/Asynchronous (SRDF/A): A remote replication solution that enables the source to asynchronously replicate data. It incorporates delta set technology, which enables write ordering by employing a buffering mechanism. SRDF/A provides minimal data loss if a regional disaster occurs.

SRDF/DM: A data migration solution that enables data migration from the source to target volume over extended distances.

SRDF/Automated Replication (SRDF/AR): A remote replication solution that uses both SRDF and TimeFinder/Mirror to implement disk-buffered replication technology. It is offered as SRDF/AR Single-hop for two-site replication and SRDF/AR Multi-hop for three-site cascade replication. SRDF/AR provides long distance solution with RPO in the order of hours.

SRDF/Star: Three-site multi-target remote replication solution that consists of primary (production), secondary (bunker), and tertiary (remote) sites. The replication between the primary and secondary sites is synchronous, whereas the replication between the primary and tertiary sites is asynchronous. In the event of primary site outage, EMC's SRDF/Star solution allows to quickly move operations and re-establish remote replication between the remaining two sites.

EMC MirrorView

- Replicates data from a primary volume to a secondary volume that reside on different VNX storage systems
- Uses a bitmap to track host writes while the link to the secondary array is down
 - ▶ When secondary is available, sends only changed data
- MirrorView family consists of:
 - ▶ MirrorView/Synchronous (MirrorView/S)
 - ▶ MirrorView/Asynchronous (MirrorView/A)

The MirrorView software enables EMC VNX storage array–based remote replication. It replicates the contents of a primary volume to a secondary volume that resides on a different VNX storage system. The MirrorView family consists of MirrorView/Synchronous (MirrorView/S) and MirrorView/Asynchronous (MirrorView/A) solutions.

MirrorView/S is a synchronous product that mirrors data between local and remote storage systems. MirrorView/A is an asynchronous product that offers extended distance replication based on periodic incremental update model. It periodically updates the remote copy of the data with all the changes that occurred on the primary copy since the last update.

EMC RecoverPoint

- RecoverPoint Continuous Remote Replication (CRR) provides both synchronous and asynchronous remote replication
- Dynamically switches between synchronous and asynchronous replication
 - ▶ Based on the policy for performance and latency
- Capable to recover data remotely to any PIT

EMC RecoverPoint Continuous Remote Replication (CRR) provides bi-directional synchronous and asynchronous replication. In normal operations, RecoverPoint CRR enables users to recover data remotely to any point in time. RecoverPoint dynamically switches between synchronous and asynchronous replication based on the policy for performance and latency.

Module 12: Summary

Key points covered in this module:

- Synchronous and asynchronous replication mode
- Host-based, array-based, and network-based remote replication
- Three-site remote replication
- Data migration solution
- Remote replication and migration in virtualized environment

This module covered various host-based, network-based, and array-based remote replication technologies. This module also covered three-site remote replication and data migration solution. Finally, this module covered remote replication in a virtualized environment.

Remote replication is the process of creating replicas of information assets at remote sites (locations). Two basic modes of remote replications are synchronous and asynchronous replication. Remote replication of data can be handled by the hosts or storage arrays. Other options include specialized network-based appliances to replicate data over the LAN or SAN.

In host-based remote replication, there are two basic approaches for performing replication: Logical volume manager (LVM) based replication and database replication via log shipping. In storage array-based remote replication, replication between arrays can be performed in synchronous, asynchronous, or disk-buffered modes. In network-based remote replication, continuous data protection technology is used for performing remote replication.

Three-site replication is used to mitigate the risks identified in two-site replication. In a three-site replication, data from the source site is replicated to two remote sites.

Data Migration and mobility solution is a specialized replication technique that enables creating remote point-in-time copies. These copies can be used for data mobility, migration, content distribution, and disaster recovery.

In a virtualized environment, the data residing in virtual volume is replicated (mirrored) to two storage arrays located in two different sites. Virtual machine migration is another technique used to ensure business continuity in case of hypervisor failure or scheduled maintenance. VM migration is the process of moving VMs from one location to another without powering off the virtual machines.

Check Your Knowledge – 1

- Which factor influences the write response time in a synchronous remote replication solution?
 - A. Number of LUNs on source array
 - B. Number of LUNs on remote array
 - C. Distance between source and remote site
 - D. RAID level configured on the remote array
- What accurately describes asynchronous remote replication?
 - A. Provides near zero RPO in the event of disaster
 - B. Not suitable for distance beyond 200 km
 - C. Writes are committed to both source and replica before acknowledging host
 - D. Writes are committed to source and immediately acknowledged to host

Check Your Knowledge – 2

- Which determines the RPO in an asynchronous remote replication?
 - A. Application response time
 - B. Average read workload
 - C. Size of the buffer
 - D. Number of LUNs on remote array
- What is the minimum number of storage volumes required in the cascade/multihop (synchronous+disk buffered) three-site replication?
 - A. 2
 - B. 3
 - C. 4
 - D. 5

Check Your Knowledge – 3

- Which is true about hypervisor-to-hypervisor VM migration?
 - A. Moves VM files from the source array to the remote array
 - B. Enables to move VMs across dissimilar storage arrays
 - C. Requires both source and target hypervisor access to the same storage
 - D. Requires VMs to go offline during the migration

Section 3: Summary

Key points covered in this section:

- Importance of business continuity
- Business continuity planning and single points of failure
- Backup operations, topologies, and targets
- Data deduplication and backup in a virtualized environment
- Data archive
- Host-based, storage array-based, and network-based local and remote replication
- Local and remote replication in a virtualized environment

This section covered the importance of business continuity, business continuity planning, and single points of failure. It also covered various backup operations, topologies, targets, data deduplication options, and backup in a virtualized environment. Further, this section covered data archiving. This section also covered various host-based, storage-based, and network-based local and remote replication technologies. Finally, this section detailed local and remote replication in a virtualized environment.