Business continuity

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

- Explain cloud service availability
- List various fault tolerance mechanisms
- Explain data protection solutions
- Explain data protection as a service

Business continuity overview

This lesson covers the following topics:

- Importance of business continuity
- Cloud service availability
- Causes and impact of cloud service unavailability
- Methods to achieve the required service availability

What is business continuity?

BC entails preparing for, responding to, and recovering from service outage that adversely affects business operations.

- It enables continuous availability of information and services to meet the required SLA
- It involves various proactive and reactive countermeasures
- Prevents interruption of mission-critical services
  - Re-establishes the impacted services as swiftly and smoothly as possible by using an automated process
- Goal of BC solution is to ensure “cloud service availability”

Microsoft 365 and Power BI Services Interrupted Worldwide

Some Microsoft customers experienced service delays or timeouts on Tuesday, which occurred worldwide.

The service interruption lasted about two hours and 40 minutes, according to a Jan. 25 "Azure Status History" page entry.

Microsoft's public Azure customers using services such as Microsoft 365 and Power BI were affected. The service problems occurred "between 07:05 UTC and 09:45 UTC on 25 January 2023," Microsoft stated.

That time period translates to a near three-hour period starting at 11:05 p.m. Pacific Time on Tuesday Jan. 24.

Microsoft is attributing the service delays to an undescribed change made by its personnel to the Microsoft Wide Area Network.

The glitch affected client connections to Azure services. It also affected service connectivity between regions, "as well as ExpressRoute connections." Azure ExpressRoute is Microsoft's network with hosting partners to provide high-bandwidth private Internet connections.

Microsoft fixed the connection problem, and most services came back on line thereafter.

"Most impacted Microsoft services automatically recovered once network connectivity was restored, and we worked to recover the remaining impacted services," Microsoft explained.

Microsoft is planning to release a "preliminary Post Incident Review" report in three days that will describe the "initial root cause" of the incident. It plans to issue a final Post Incident Review report "14 days later" that will offer a "deep dive" description.

Microsoft's services go down all of the time. This particular incident was widely reported. The only area in the world that was not affected was China, according to a *Reuters'* report.

Microsoft does offer service-level agreements (SLAs) for its various services. The SLAs typically assure 99.95 percent or 99.99 percent uptimes. A service credit might be permitted if that uptime percentage drops. However, the SLAs come with fine print exceptions. For instance, the SLA won't apply for failures in a "single Microsoft Datacenter location" when organizations are specifically dependent on that location.

Microsoft's SLA terms are rather involved and they can get changed. The latest SLA documents can be downloaded at this page.

https://rcpmag.com/articles/2023/01/25/worldwide-microsoft-service-outage.aspx

Cloud service availability

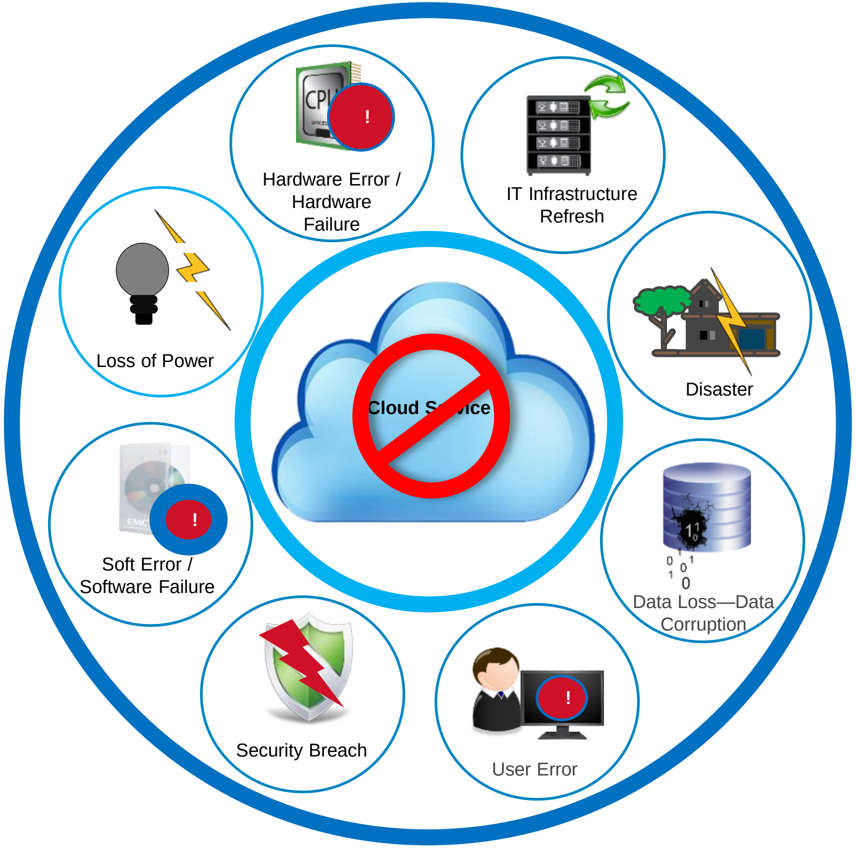
Refers to the ability of a cloud service to perform its agreed function according to business requirements and customer expectations during its specified time of operation.

- Service availability depends primarily on the reliability of the cloud infrastructure components, applications, and the availability of data
- Cloud service availability is measured as percentage of uptime in a given year

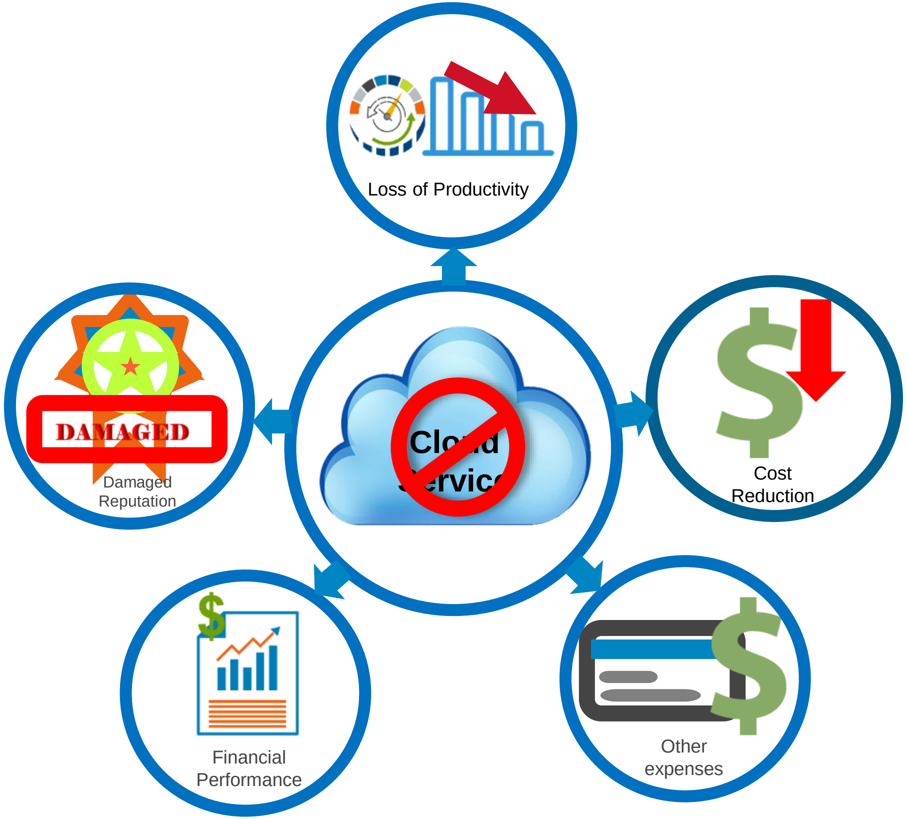
$$\text{Cloud Service Availability} = \frac{\text{Agreed Service Time} - \text{Downtime}}{\text{Agreed Service Time}} \text{ x } 100$$

Service availability—99.99%

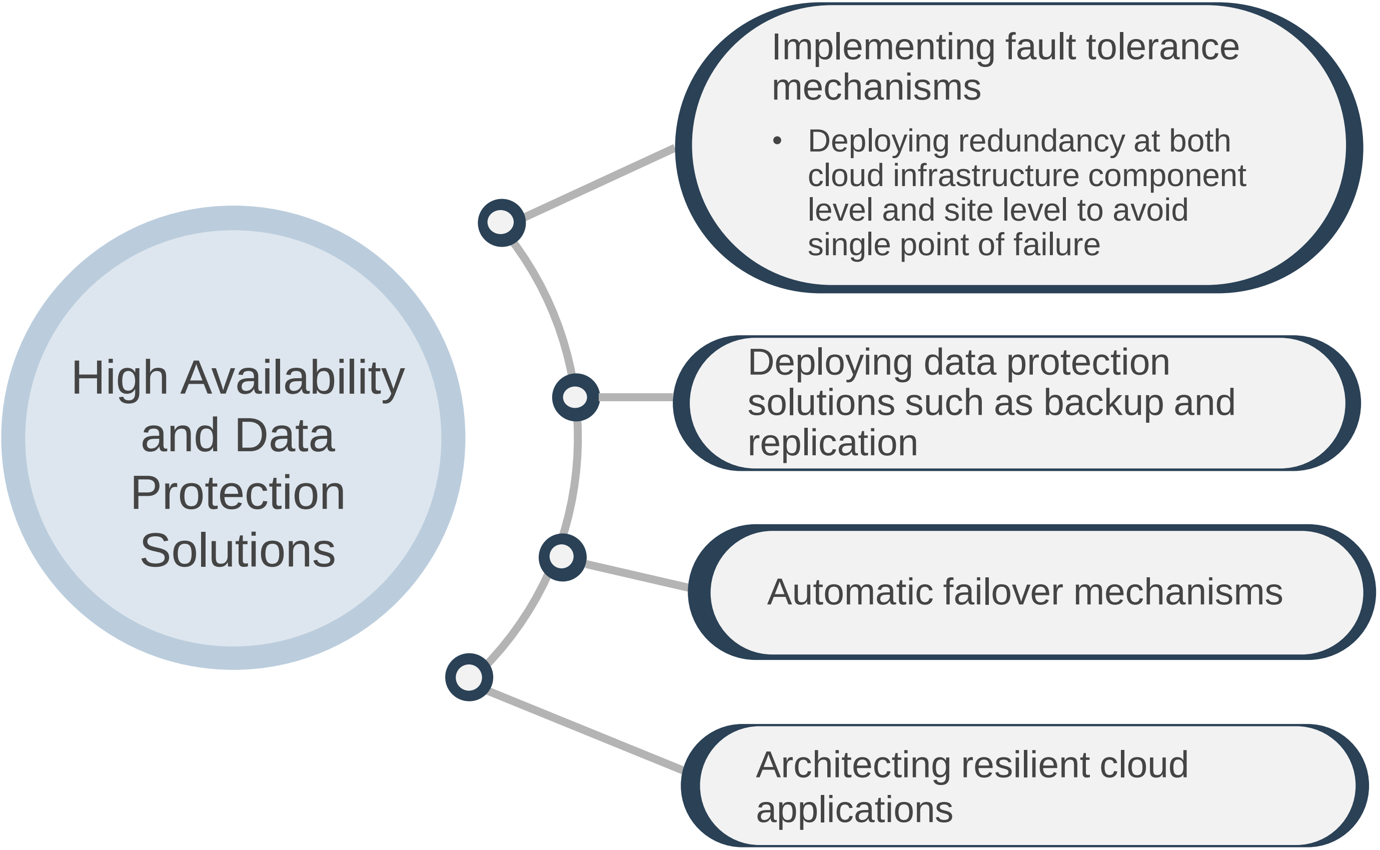
Service unavailability: Causes


Hardware Error / Hardware Failure
IT Infrastructure Refresh
Loss of Power
Disaster
Cloud Service
Soft Error / Software Failure
Data Loss—Data Corruption
Security Breach
User Error

Impact of service unavailability



Methods to achieve required service availability



High Availability and Data Protection Solutions

Implementing fault tolerance mechanisms

- Deploying redundancy at both cloud infrastructure component level and site level to avoid single point of failure

Deploying data protection solutions such as backup and replication

Automatic failover mechanisms

Architecting resilient cloud applications

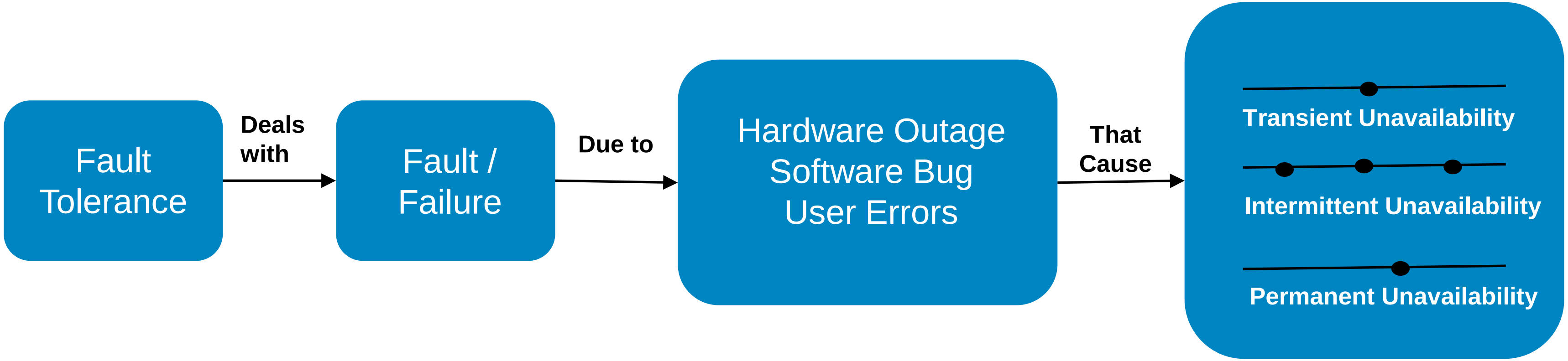
Fault tolerance IT infrastructure

This lesson covers the following topics:

- Need for fault tolerance
- Key fault tolerance techniques
- Application resiliency techniques

What is fault tolerance?

Ability of a system to continue functioning in the event of a fault within or failure of some of its components



Fault Tolerance

Deals with

Fault / Failure

Due to

Hardware Outage
Software Bug
User Errors

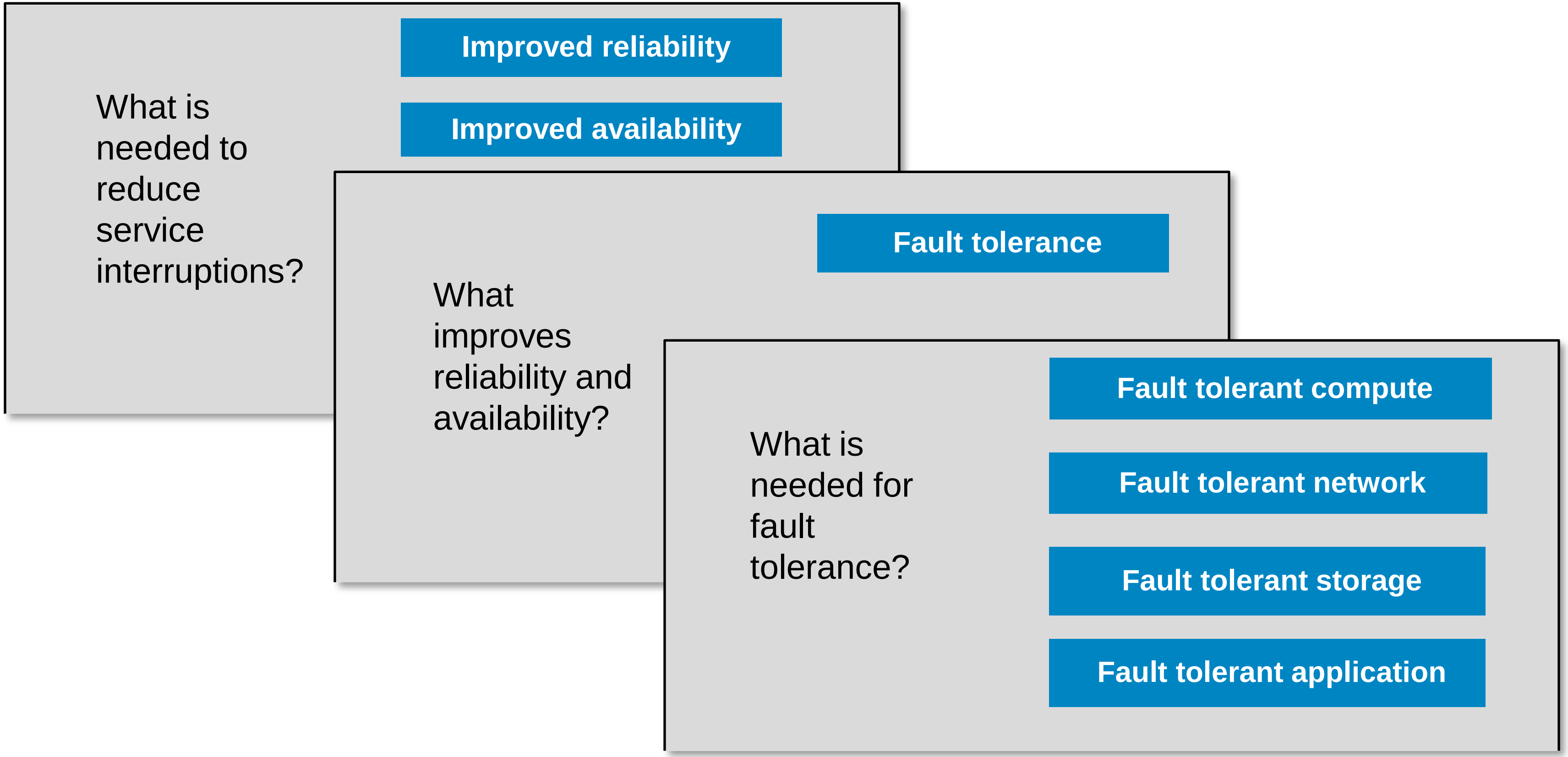
That Cause

Transient Unavailability

Intermittent Unavailability

Permanent Unavailability

Need for fault tolerance



What is needed to reduce service interruptions?

- Improved reliability
- Improved availability

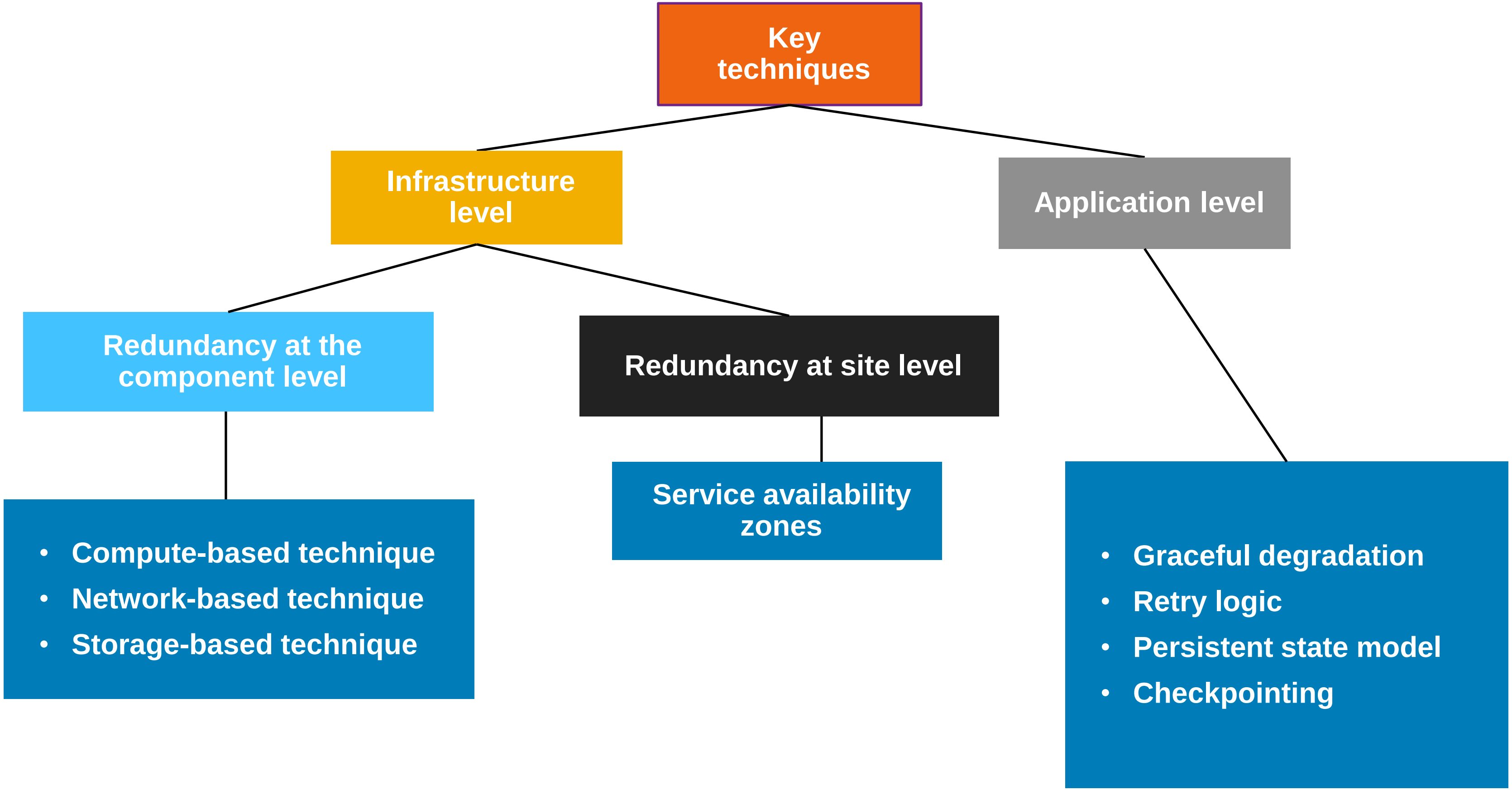
What improves reliability and availability?

- Fault tolerance

What is needed for fault tolerance?

- Fault tolerant compute
- Fault tolerant network
- Fault tolerant storage
- Fault tolerant application

Key fault tolerance techniques



- Compute-based technique
- Network-based technique
- Storage-based technique

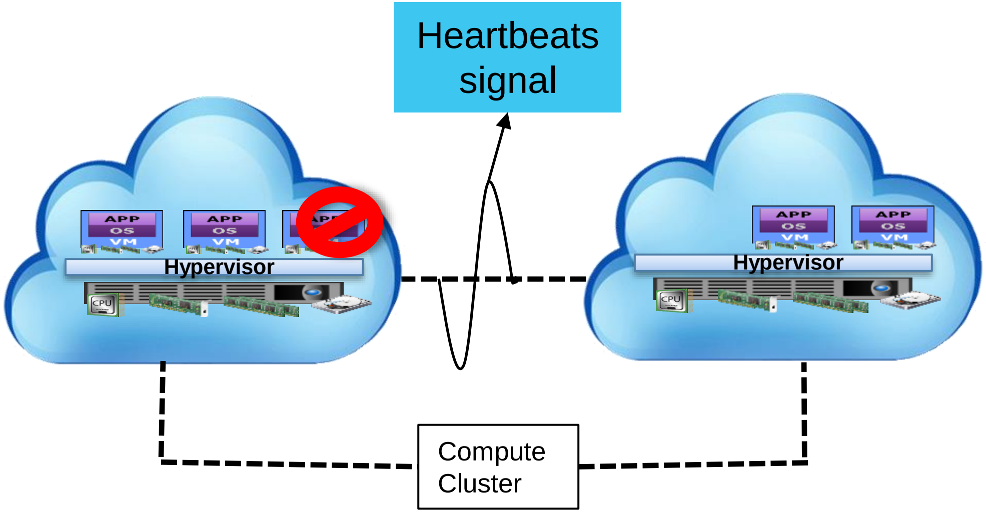
- Graceful degradation
- Retry logic
- Persistent state model
- Checkpointing

Fault tolerance technique

Compute-based technique

Compute Clustering

- Two or more compute instances / hypervisors are clustered to provide high availability and load balancing
- Service running on a failed compute system moves to another compute system
- Heartbeat mechanism determines the health of compute systems in a cluster
- The two common cluster implementations are: Active-Active and Active-Passive

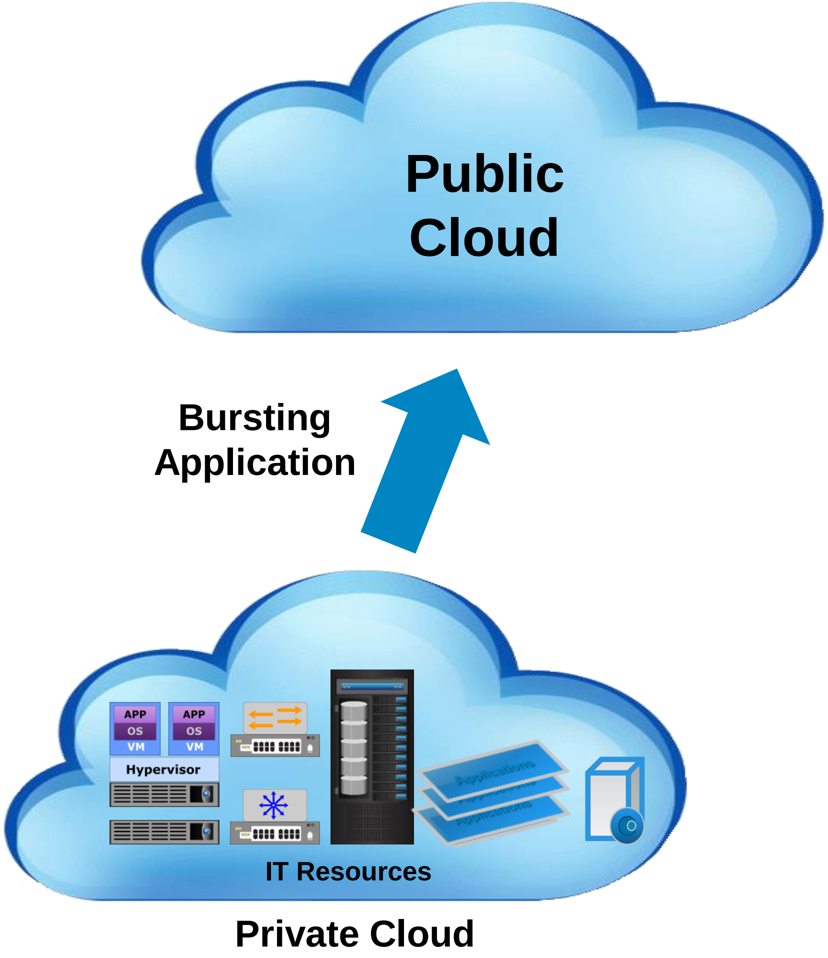


Example: Compute clustering can be implemented among multiple physical compute systems, or multiple VMs, or VM and physical compute system, or multiple hypervisors.

DELLEMC

Cloud Bursting: Avoids Downtime

- Cloud bursting represents one of the key advantages of hybrid cloud technology
- Business will be able to ensure service availability as well as realize cost savings by not having to invest in excess infrastructure to meet peak demands
- Consider factors such as compliance, load balancing, application portability and compatibility and performance implication due to latency



Fault tolerance technique

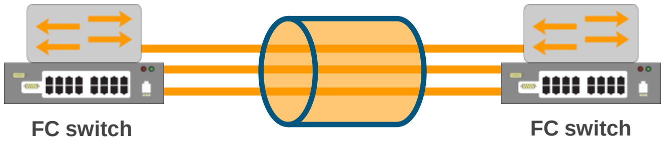
Network-based technique

Link aggregation

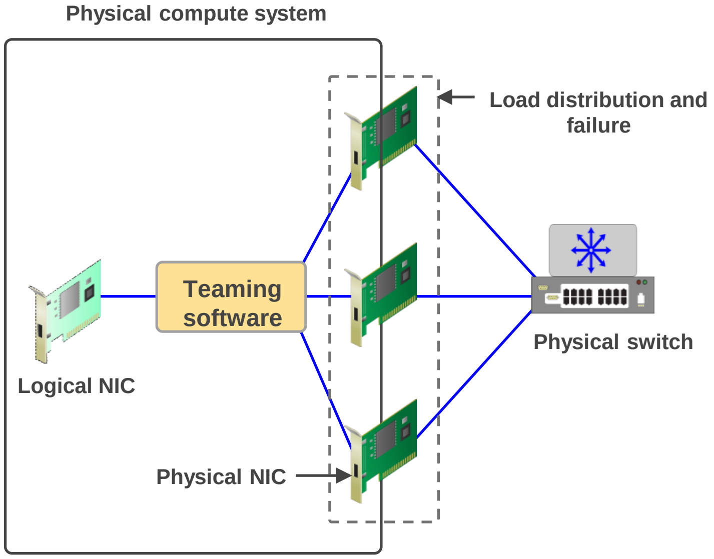
- Combines links between two switches and also between a switch and a node
- Enables network traffic failover in the event of a link failure in the aggregation

NIC teaming

- Groups NICs so that they appear as a single, logical NIC to the OS or hypervisor
- Provides network traffic failover in the event of a NIC/link failure
- Distributes network traffic across NICs



Link aggregation



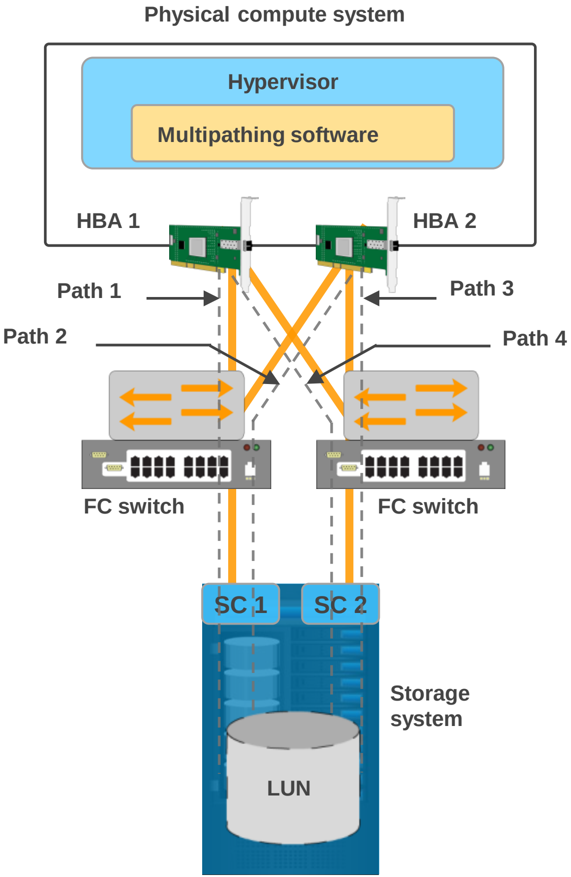
NIC teaming

Fault tolerance technique

Network-based technique

Multipathing

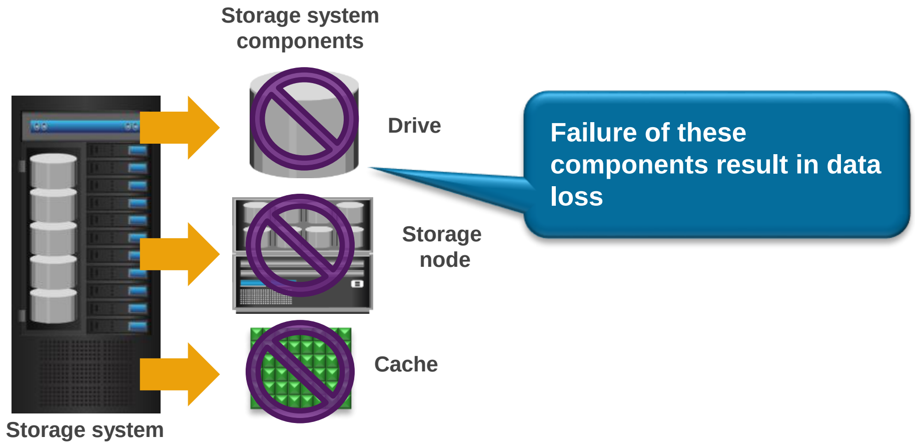
- Enables a compute system to use multiple paths for transferring data to a LUN
- Enables failover by redirecting I/O from a failed path to another active path
- Performs load balancing by distributing I/O across active paths
  - Standby paths become active if one or more active paths fail



Fault tolerance technique

Storage-based technique

- Cloud comprises of large number of storage media such as disk drive and solid state drives
- Failure of these drives can result in data loss
- Greater the number of drives, greater the probability of the drive failure
- Some of the key fault tolerance techniques are RAID, RAIN architecture, erasure coding, hot drive sparing, and cache protection

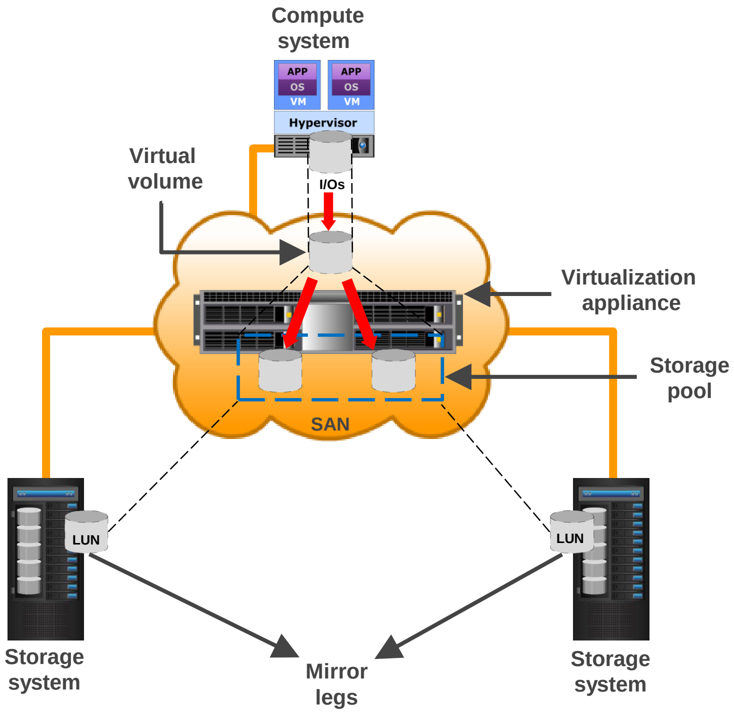

Storage system components
Drive
Storage node
Cache
Storage system
Failure of these components result in data loss

Fault tolerance storage techniques allow the storage system to continue working if drive, node, and cache fail

Fault tolerance technique

Storage resiliency using virtualization

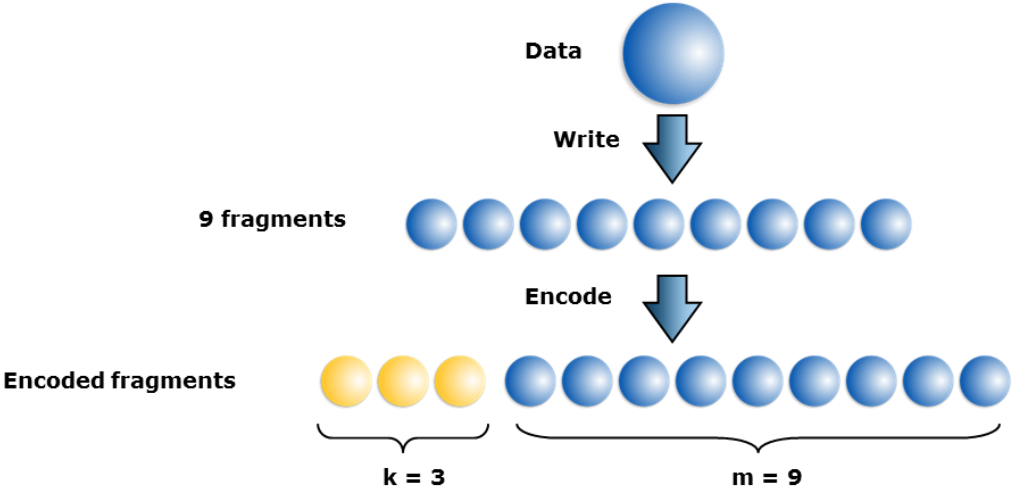
- Virtual volume is created using virtualization appliance
  - Each I/O to the volume is mirrored to the LUNs on the storage systems
- Virtual volume is continuously available to compute system
  - Even if one of the storage systems is unavailable due to failure



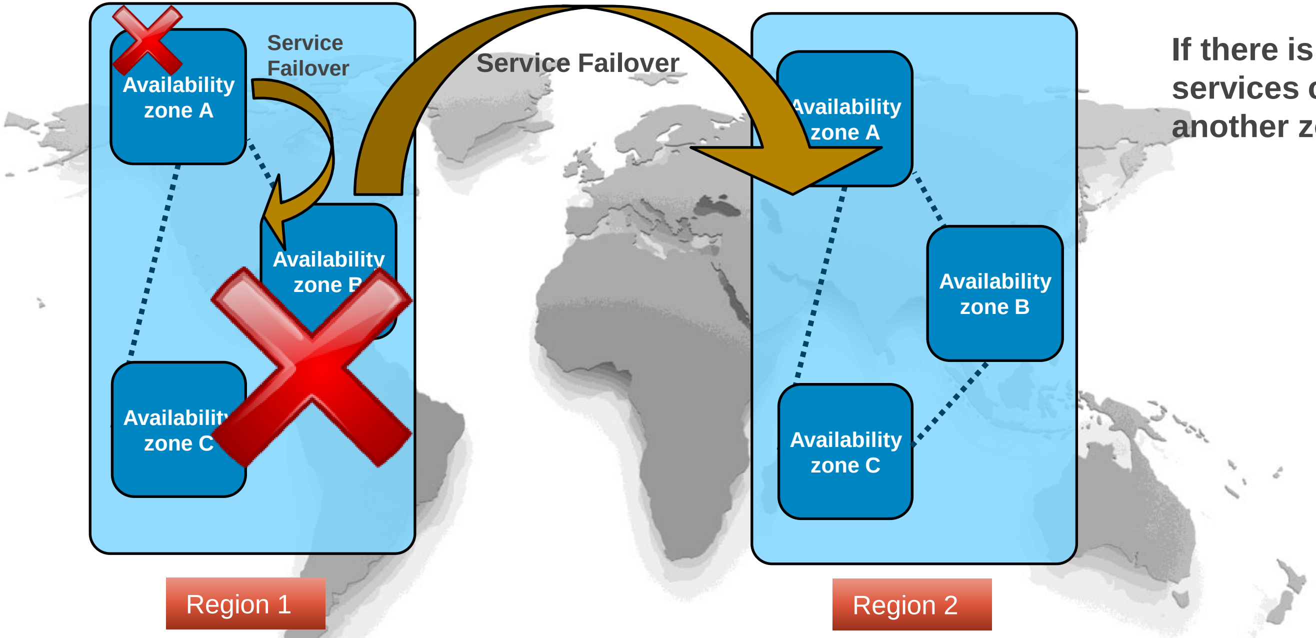
Fault tolerance technique

Erasure coding technique

- Provides space-optimal data redundancy to protect data loss against multiple drive/node failures
  - A set of n disks is divided into m disks to hold data and k disks to hold coding information
  - Coding information is calculated from data



Redundancy at site level

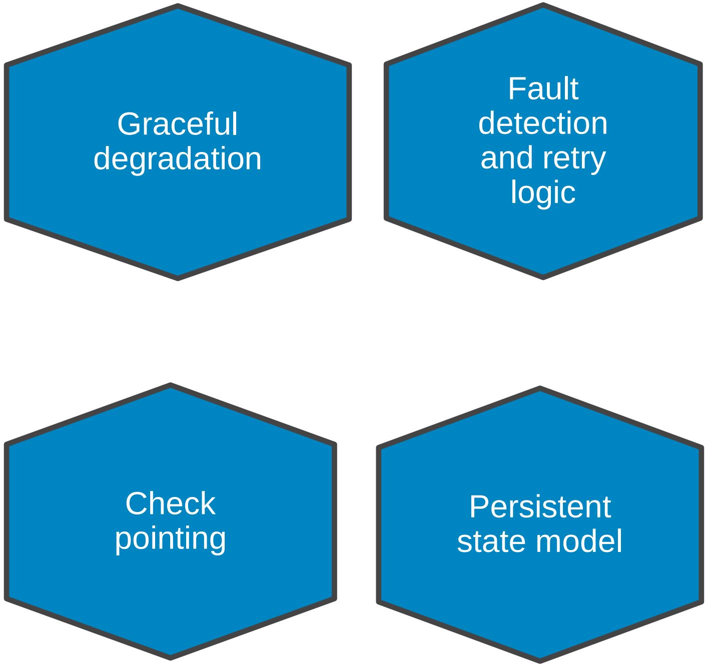


If there is a zone outage, services can fail over to another zone.

Availability zone is a location with its own set of resources and isolated from other zones.

Availability zones, although isolated from each other, are connected through low-latency network links.

Application resiliency techniques



Graceful degradation

Fault detection and retry logic

Check pointing

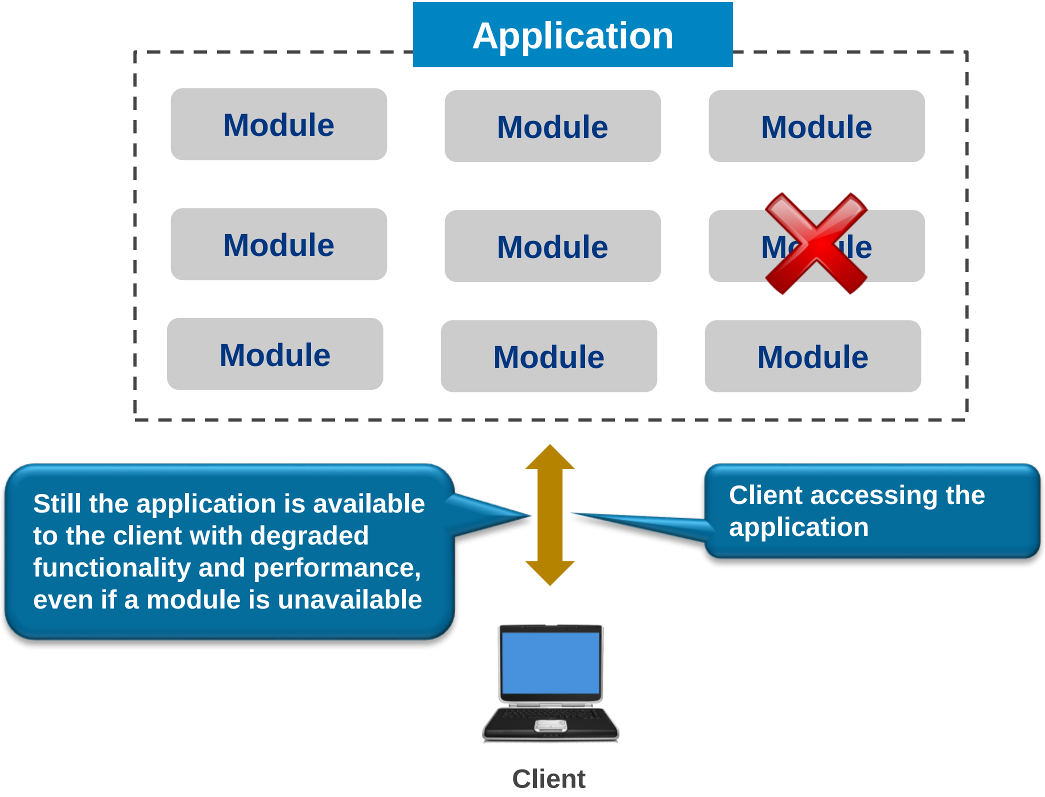
Persistent state model

Key application design strategies

In the design of an application the failure of individual resources often has to be foreseen to ensure availability.

Fault-tolerant applications have logic to detect and handle fault conditions to avoid application downtime

Graceful degradation



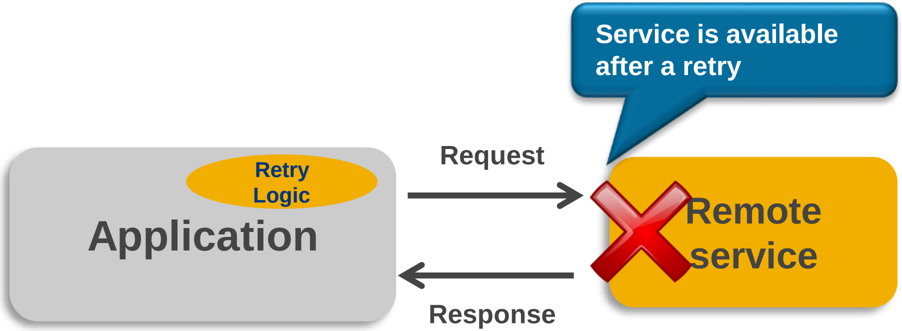
Still the application is available to the client with degraded functionality and performance, even if a module is unavailable

Client accessing the application

- Application maintains limited functionality even when some of the modules or supporting services are not available
- Unavailability of certain application component or modules should not make the entire application down

Example: e-commerce application consists of modules such as product catalog, shopping cart, order status, order submission, and order processing. Unavailability of order status module does not impact the entire application

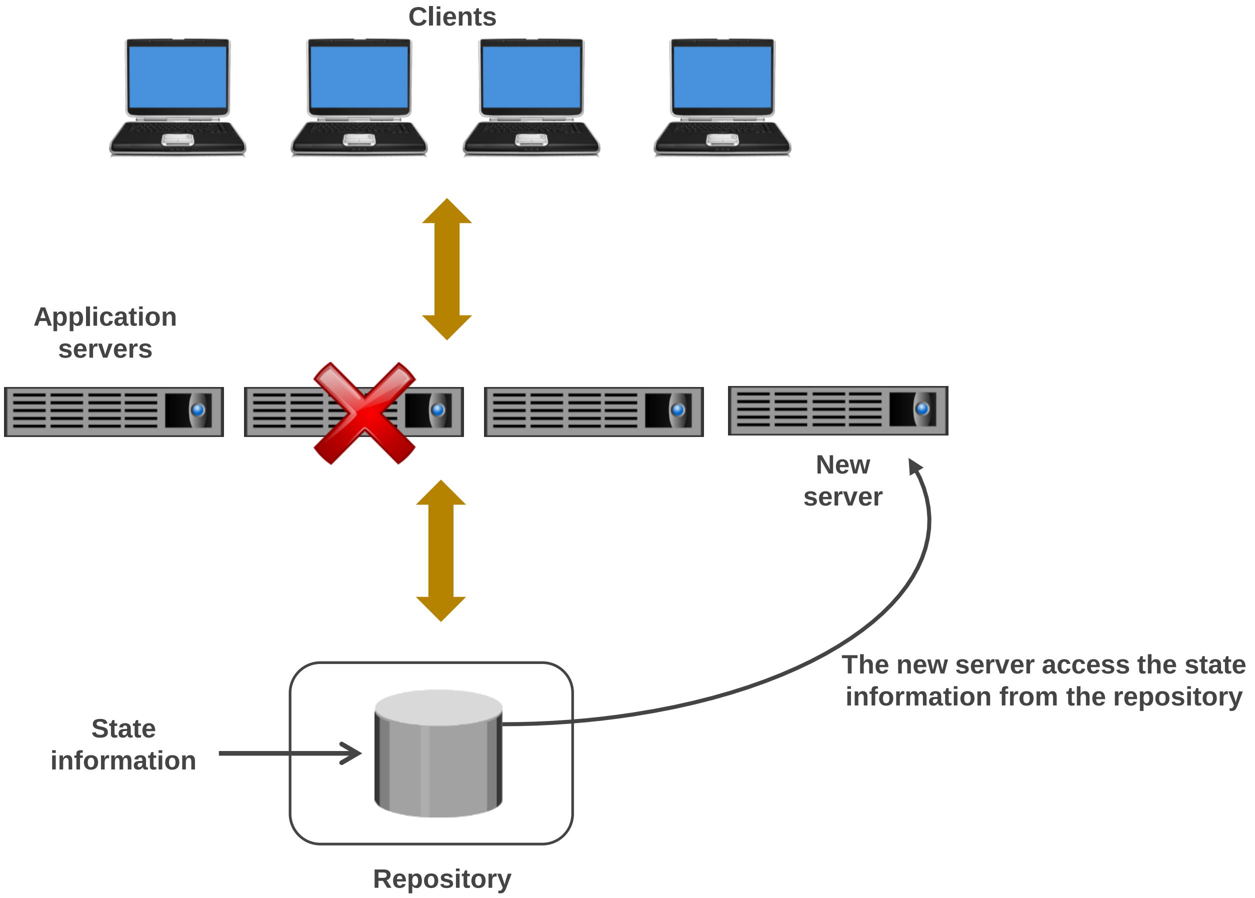
Fault detection and retry logic



Refer to a mechanism that implements a logic in the code of an application to improve the availability

Detect and retry the service that is temporarily down which may result in successful restore of the service

Application state model


Clients
Application servers
New server
The new server access the state information from the repository
State information
Repository

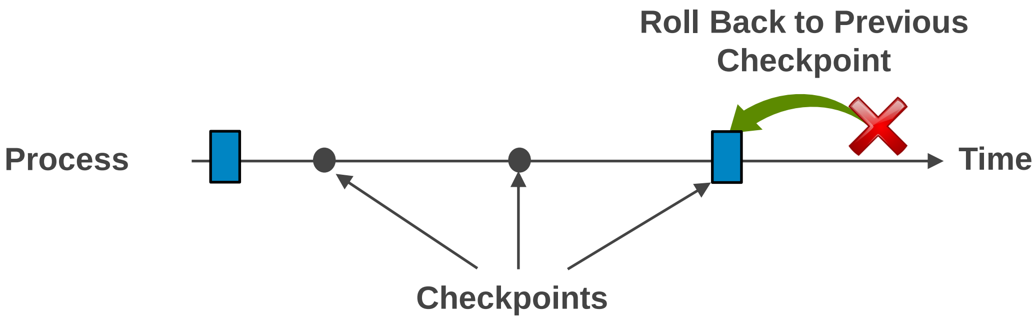
State information is stored out of the memory and stored in a data repository.

If an instance fails, the state information is still available in the repository.

In a stateless application model, the server does not store any session state information about the client session.

Checkpointing

- Saves a copy of the state-checkpoint of a process or application periodically
- Enables rolling back to a previous state and continuing tasks
- Provides protection against transient unavailability



Data protection: Backup

This lesson covers the following topics:

- Data protection overview
- Evolution of data protection solutions
- Backup and recovery

Data protection

- Protecting critical data ensures availability of services
  - Seamless service failover requires the availability of data
- Businesses also implement data protection solutions to comply with regulatory requirements
- Individual services and associated data sets have different business values, require different data protection strategies
- Two common data protection solutions:
  - Backup
  - Replication

Evolution of data protection solutions

Server-centric Backup

- Data source – server
- Separate backup infrastructure
- One-size-fits-all

Infrastructure-centric Data Protection

- Data source – Infrastructure components
- Use embedded backup/replication/archiving features

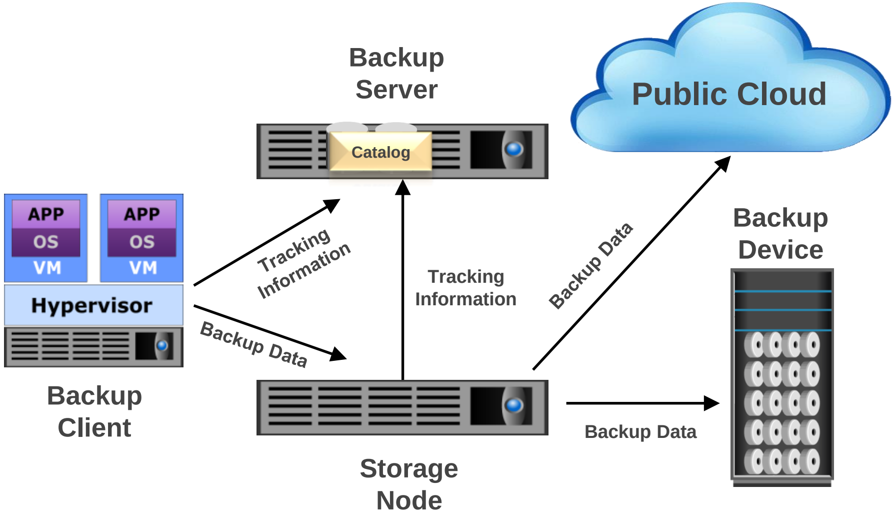
Cloud Service-centric Data Management

- SLA-driven protection
- Optimized for data management

Backup and recovery

An extra copy of production data, created and retained for the sole purpose of recovering lost or corrupted data.

- RPO and RTO are the primary considerations in selecting and implementing a specific backup strategy
  - RPO specifies the time interval between two backups
  - RTO relates to the time taken to recover data from backup
- To implement a successful backup and recovery solution
  - Service providers need to evaluate the backup methods along with their recovery considerations and retention requirements



Backup and recovery forecast

By 2021, 50% of organizations will augment or replace their current backup application with another solution, compared to what they deployed at the beginning of 2017.

- Gartner

By 2020, the number of enterprises using the cloud as a backup target will double, up from 10% at the beginning of 2017.

- Gartner

Backup methods

- Agent-based backup approach
- Image-based backup

Agent-based Backup approach

- Backup agent is installed on each VM
  - Performs file-level backup and recovery
  - Does not backup VM configuration files
- Performing backup on multiple VMs on a compute system may consume more resources and lead to resource contention
  - Impacts performance of applications running on VMs

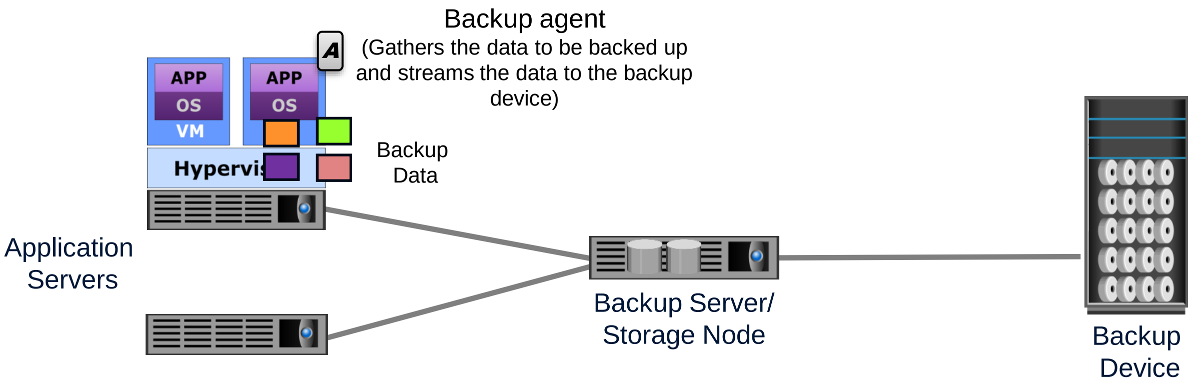
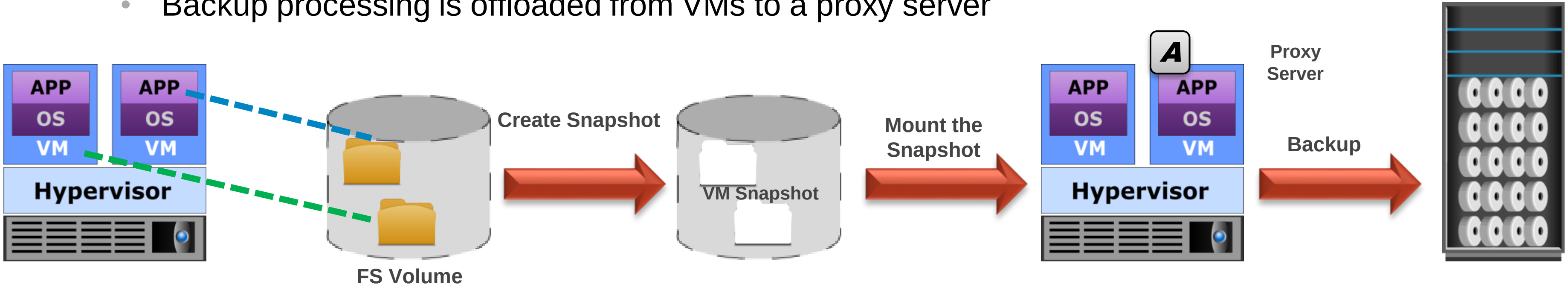


Image-level Backup

Creates a copy of the entire virtual disk and configuration data associated with a particular VM

- Backup is saved as a single entity called a VM image
  - Provides VM image-level and file-level recovery
- No backup agent is required inside the VM to backup
- Backup processing is offloaded from VMs to a proxy server



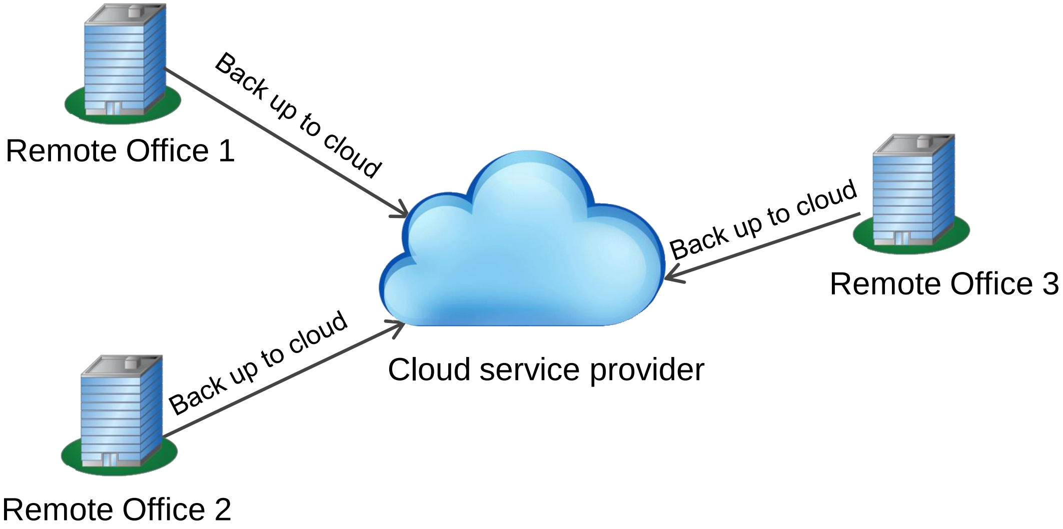
Use case of ROBO backup in cloud

Challenges associated with ROBO backup

- ✓ Lack of qualified IT staff with backup skills
- ✓ Less IT infrastructure to manage the backup copies
- ✓ Huge volume of redundant content
- ✓ Silos of data repository leads to security threat
- ✓ High cost to manage backup across remote offices

Backing up ROBO data to the Cloud addresses these challenges

Use case of ROBO backup in cloud



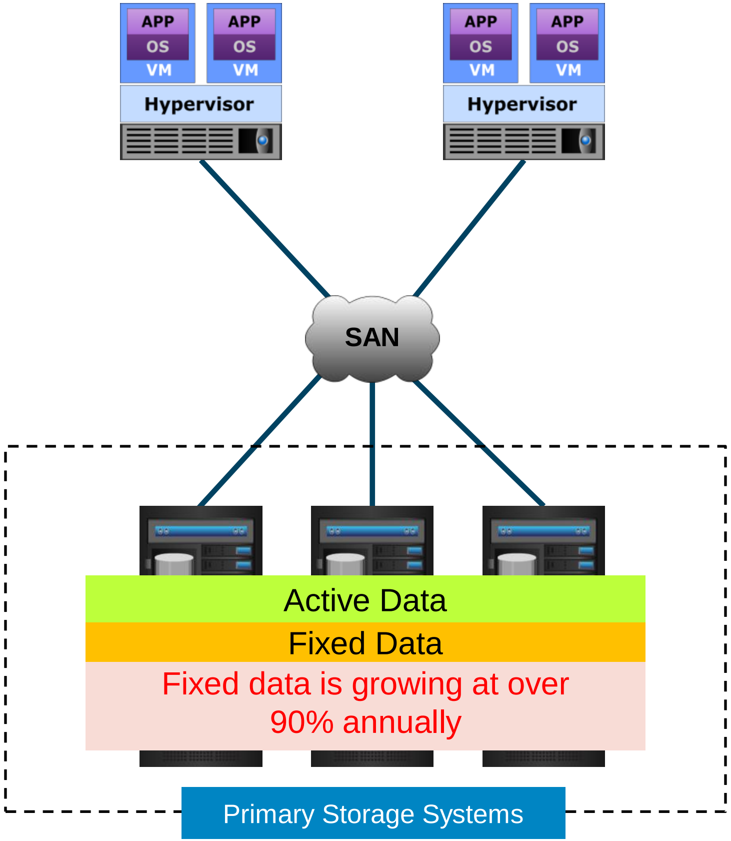
- Cloud backup service typically deploys disk-based backup solutions along with source-based deduplication to eliminate the challenges associated with centrally backing up remote-office data
- Performing backup to the cloud, reduces the cost of managing the organization’s ROBO backup environment

Data protection: Archiving and replication

This lesson covers the following topics:

- Archiving
- Replication
- Role of deduplication in a data protection environment

Why do we need data archiving?



What are the challenges of keeping fixed data in primary storage?

- Increasing consumption of expensive primary storage
- High-performance storage for less frequently accessed data
- Risk of compliance breach
- Increased data backup window and cost

Data archiving addresses these challenges

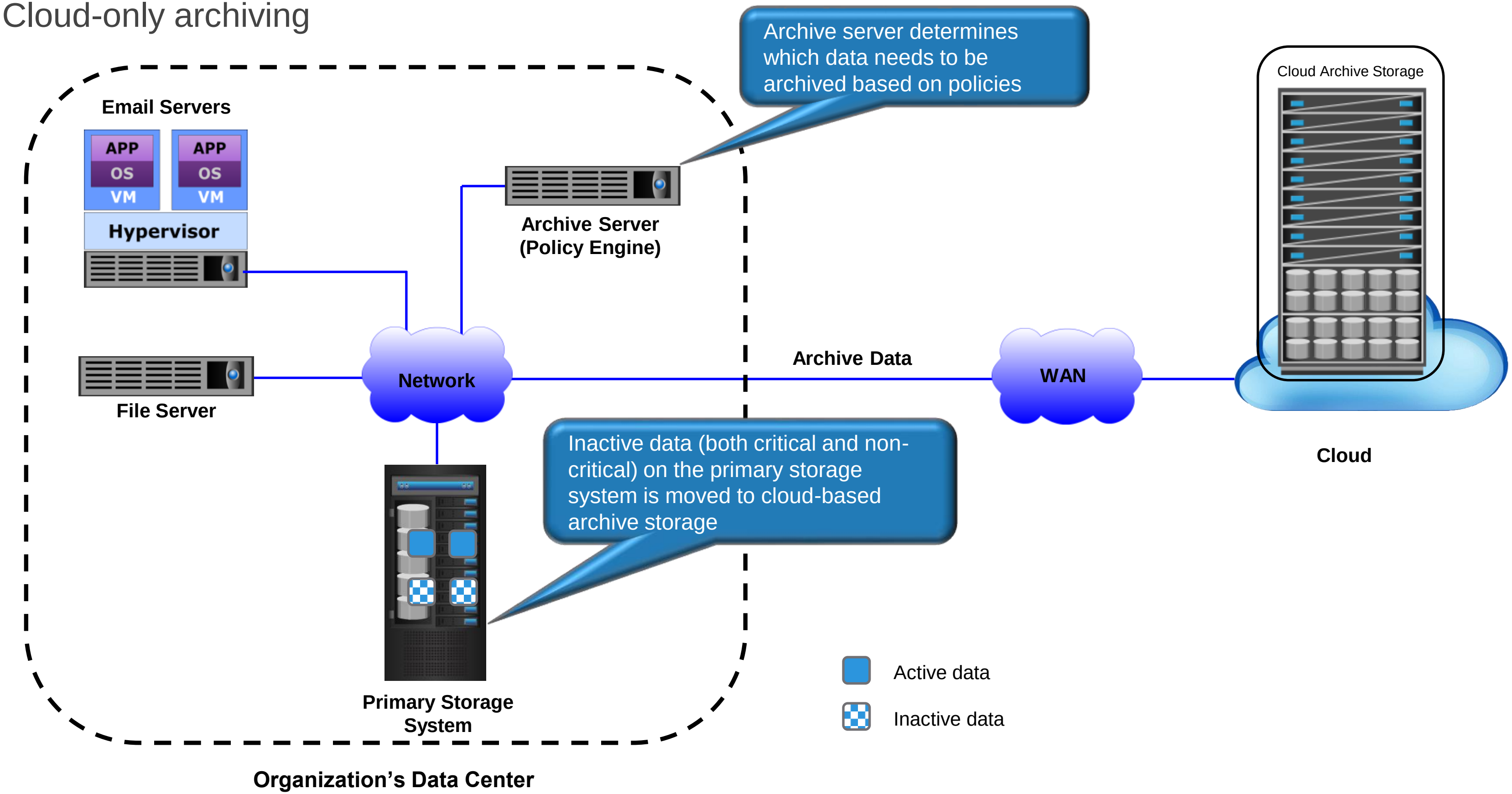
Data archiving and its benefits

Data archiving moves fixed data that is no longer actively accessed to a separate low-cost archive storage system for long-term retention and future reference.

- Saves primary storage capacity
- Reduces backup window and backup
- Moves less frequently accessed data to lower-cost archive storage
- Preserves data for future reference and adherence to regulatory compliance

Cloud-based archiving options

- Cloud-only archiving



Archive server determines which data needs to be archived based on policies

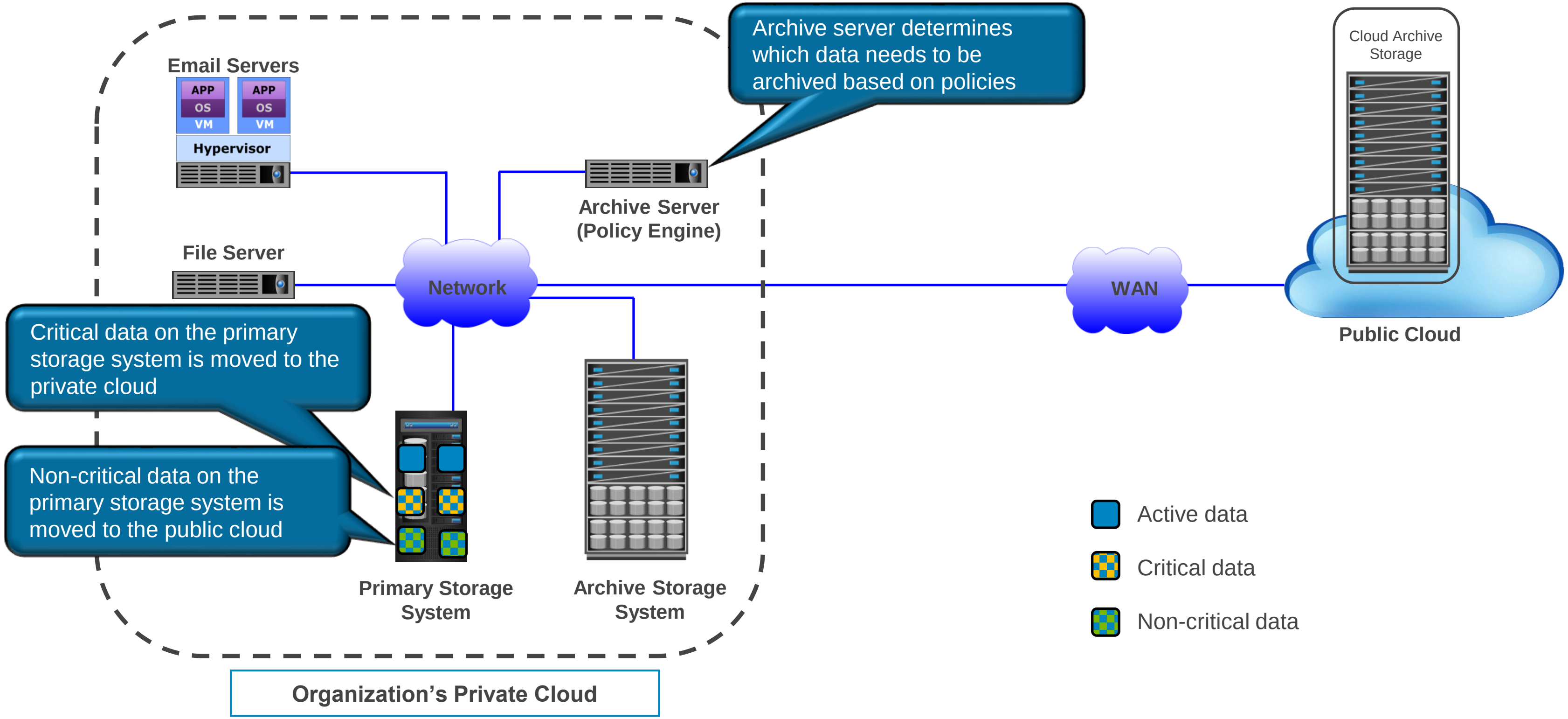
Inactive data (both critical and non-critical) on the primary storage system is moved to cloud-based archive storage

🟦 Active data

▦ Inactive data

Cloud-based archiving options

- Hybrid archiving



Archive server determines which data needs to be archived based on policies

Critical data on the primary storage system is moved to the private cloud

Non-critical data on the primary storage system is moved to the public cloud

🟦 Active data

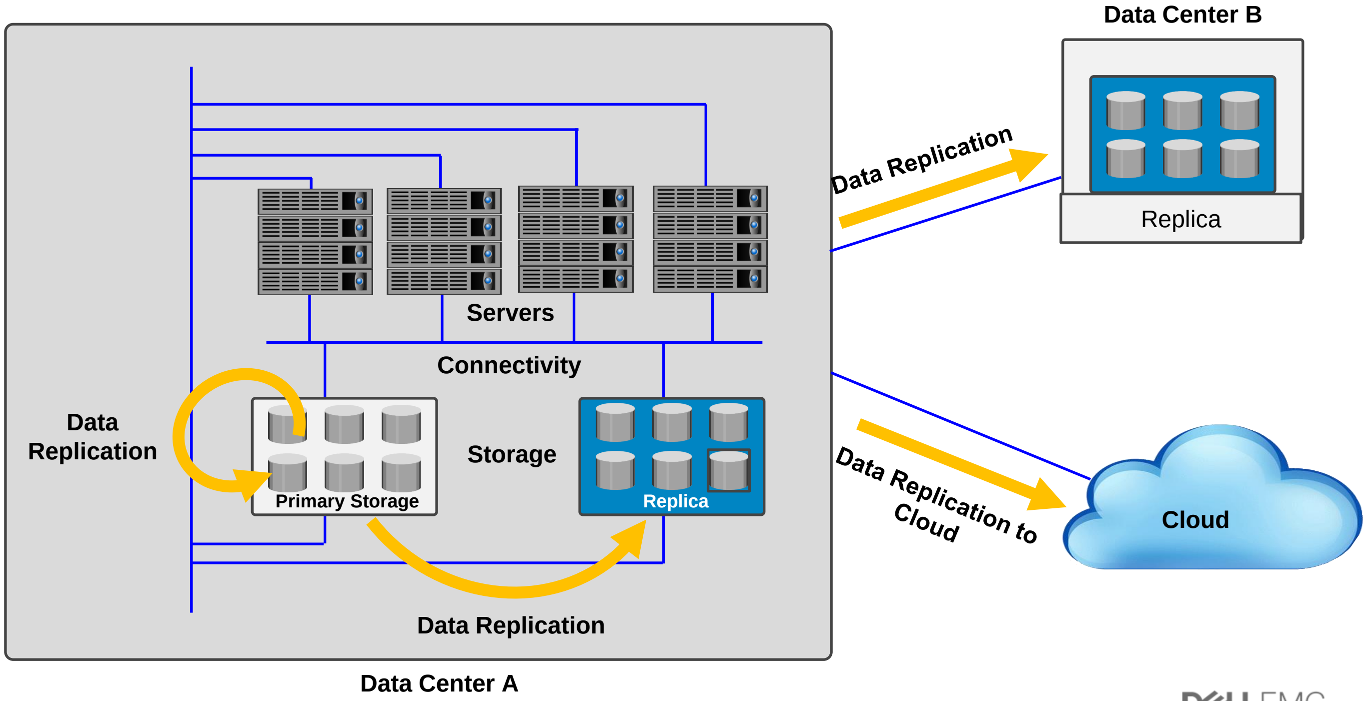
▨ Critical data

▨ Non-critical data

Introduction to Replication

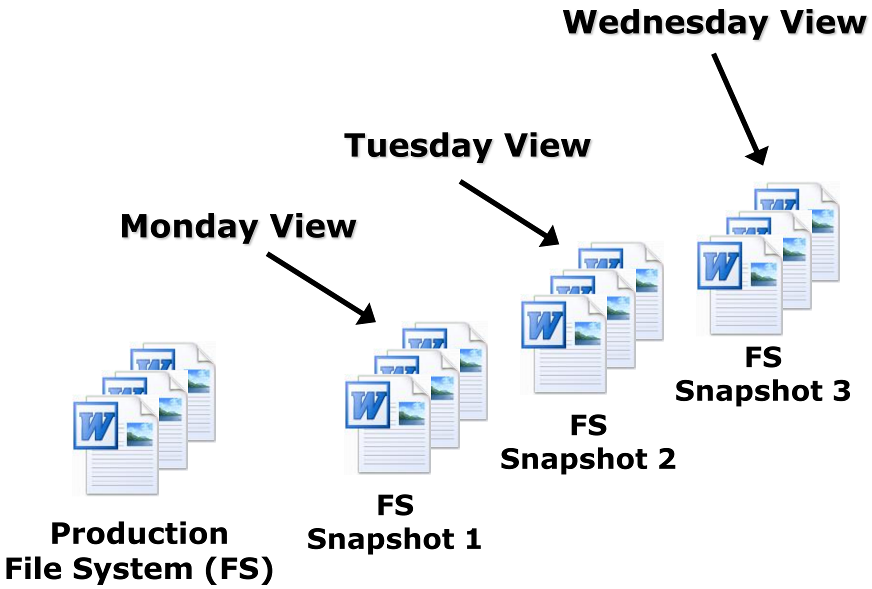
Process of creating an exact copy or replica of the data to ensure business continuity if there is a local outage or disaster

- Key replication technique
  - Snapshot
  - Synchronous and asynchronous replication
  - Continuous Data Protection (CDP)



Snapshot

File system (FS) snapshot creates a copy of a file system at a specific point-in-time, even when the original file system continues to be updated and used normally



Wednesday View

Tuesday View

Monday View

FS Snapshot 3

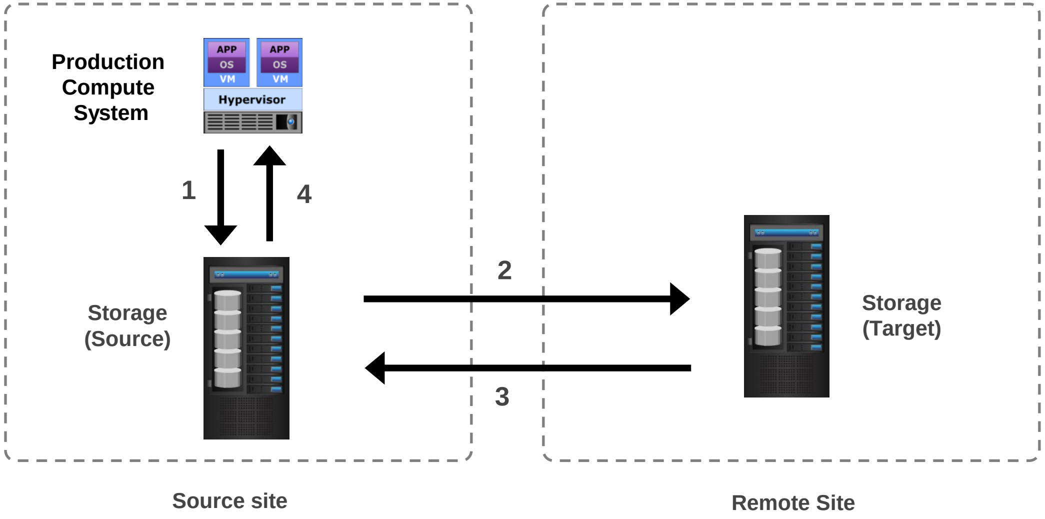
FS Snapshot 2

FS Snapshot 1

Production File System (FS)

Remote Replication: Synchronous

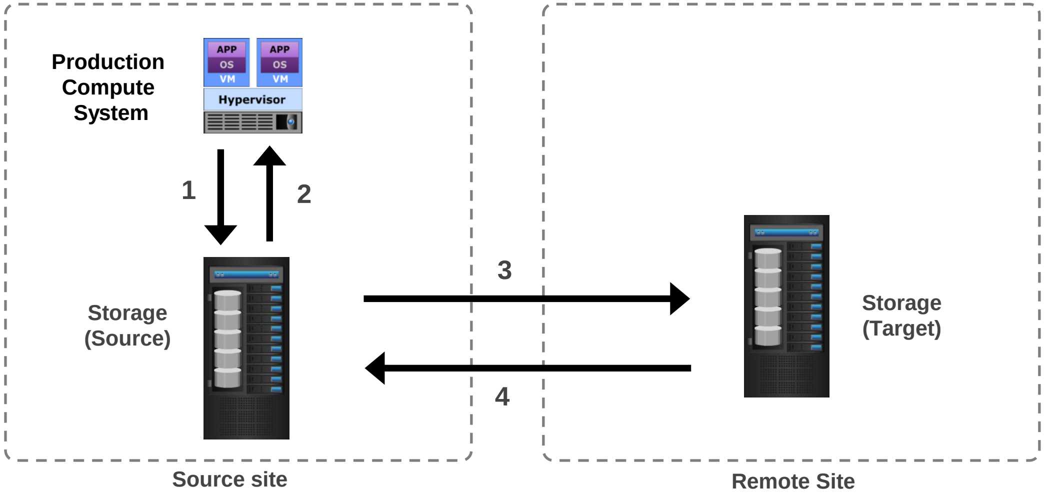
- Write is committed to both the source and the remote replica before it is acknowledged to the compute system
- Allows to restart business operations at a remote site with zero data loss
  - Provides near zero RPO



1. The write I/O is received from production compute system into cache of source and placed in queue
2. The write I/O is transmitted to the cache of the target storage
3. Receipt acknowledgment is provided by target storage back to cache of the source
4. Source storage system sends an acknowledgment back to the production compute system

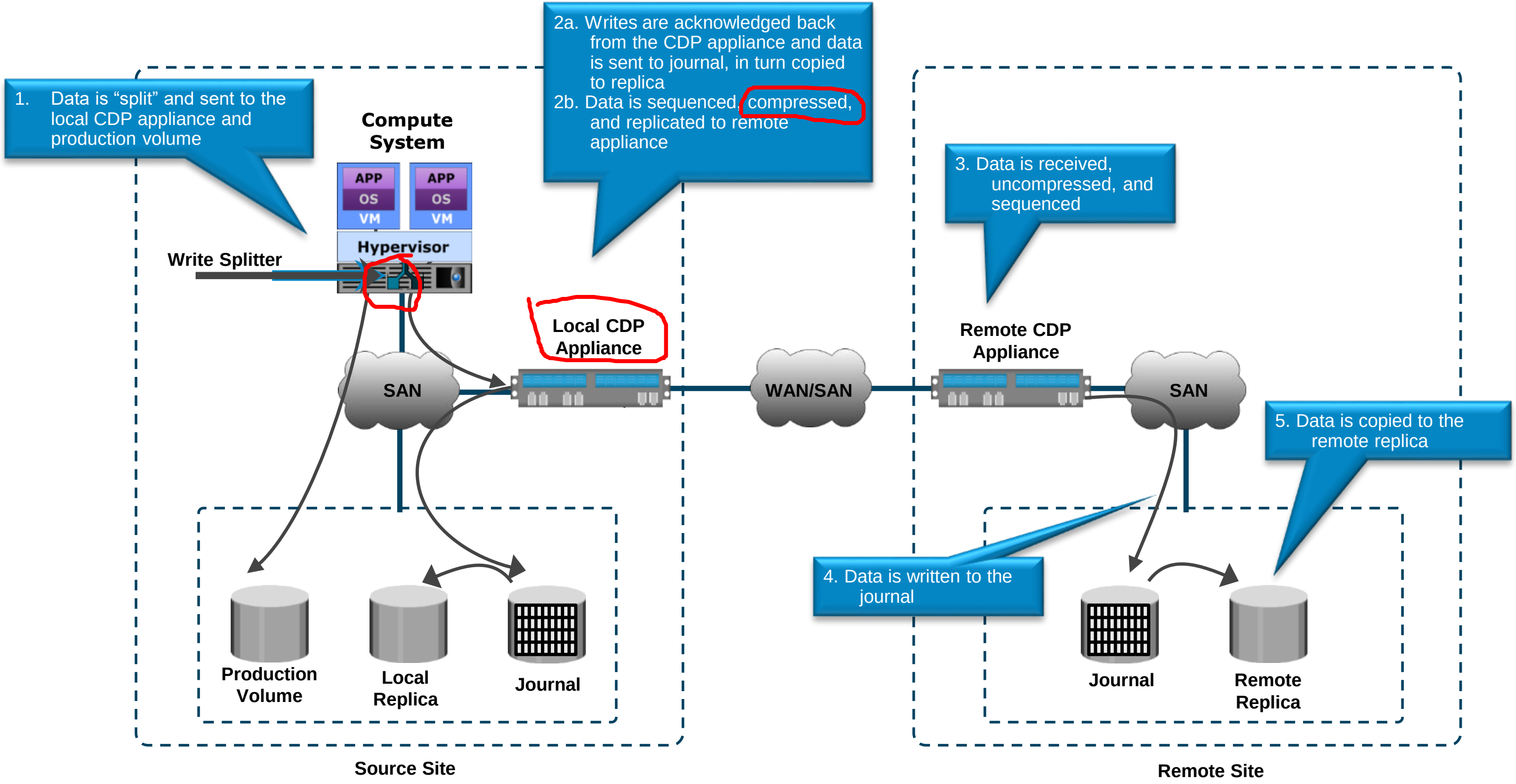
Remote Replication: Asynchronous

- Write is committed to the source and immediately acknowledged to the compute system
  - Data is buffered at the source and sent to the remote site periodically
  - Application write response time is not dependent on the latency of the link
  - Replica will be behind the source by a finite amount (finite RPO)



1. The write I/O is received from production compute system into cache of source and placed in queue
2. Receipt acknowledgment is provided by source storage back to production compute system
3. The write I/O is transmitted to the cache of the target storage
4. Target acknowledges back to source

CDP Operations: Local and remote replication



1. Data is “split” and sent to the local CDP appliance and production volume

2a. Writes are acknowledged back from the CDP appliance and data is sent to journal, in turn copied to replica

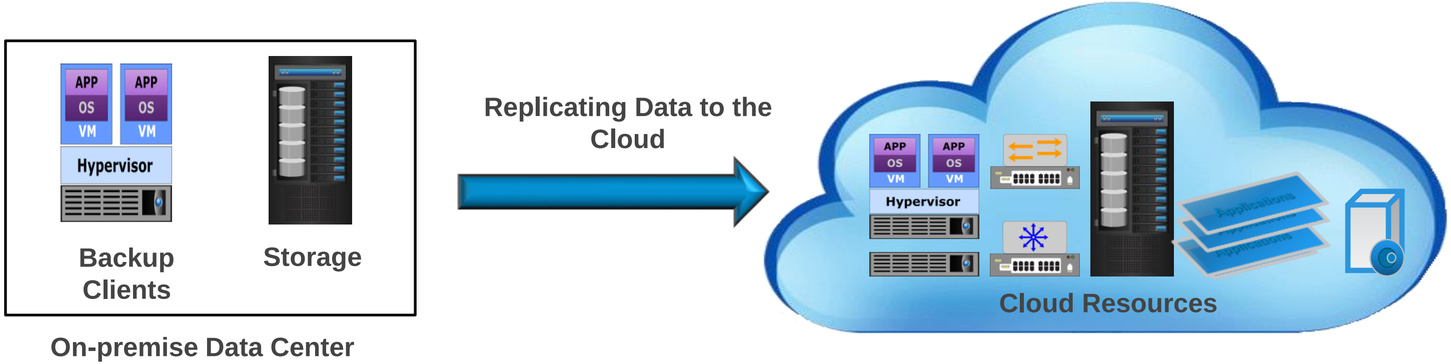
2b. Data is sequenced, compressed, and replicated to remote appliance

3. Data is received, uncompressed, and sequenced

4. Data is written to the journal

5. Data is copied to the remote replica

Replication to cloud



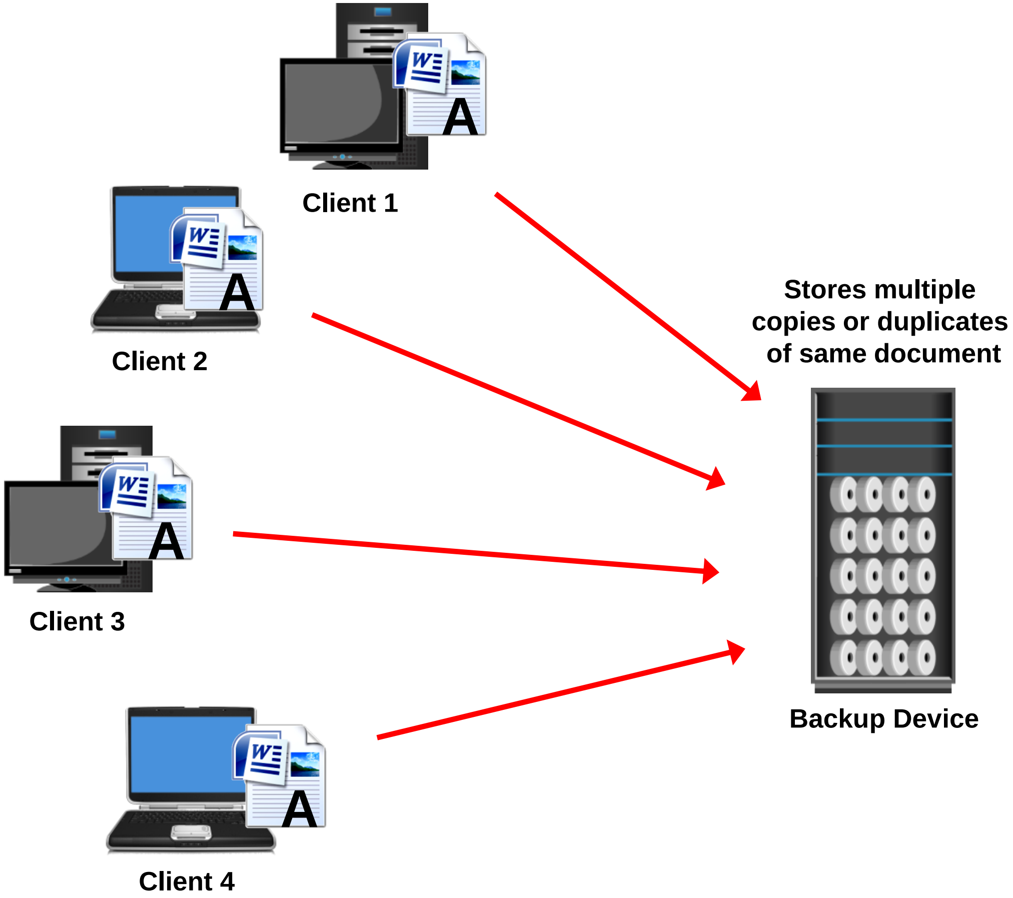
- Replicating application data and VM to the cloud enable organization to restart the application from the cloud
- Replication to the cloud can be performed using compute-based, network-based, and storage-based replication techniques

Examples:

DELL EMC Isilon cloudpools

DELL EMC CloudBoost

Why do we need data deduplication?



Client 1

Client 2

Client 3

Client 4

Stores multiple copies or duplicates of same document

Backup Device

What are the challenges of duplicate data in a data center?

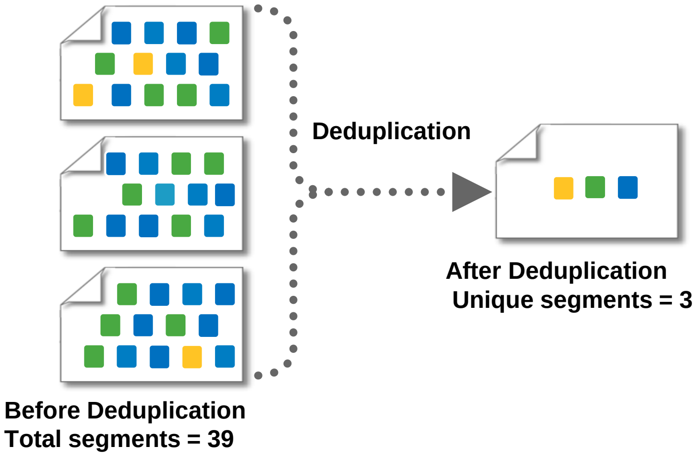
- Difficult to protect the data within the budget
- Impacts the backup window
- Requirement of more bandwidth

Data Deduplication addresses these challenges

Data deduplication

The process of detecting and identifying the unique data segments within a given set of data to eliminate redundancy.

- Deduplication process
  - Chunk the data set
  - Identify duplicate chunk
  - Eliminate the redundant chunk



Deduplication methods

Source based Deduplication

- Eliminates redundant data at the source i.e. backup client
- Client sends only new, unique segments across the network
- Reduces storage and network bandwidth requirements
- Increases overhead on the backup client

Target based Deduplication

- Offloads deduplication process from the backup client
- Data is deduplicated at the target either inline or post-process

Deduplication benefits

Benefits	Description
Reduces infrastructure costs	Elimination of redundant data leads to less storage space consumption during data backup
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 back up, which reduces backup window
Reduces network bandwidth requirement	Eliminating the redundant data reduces the amount of data sent over the network

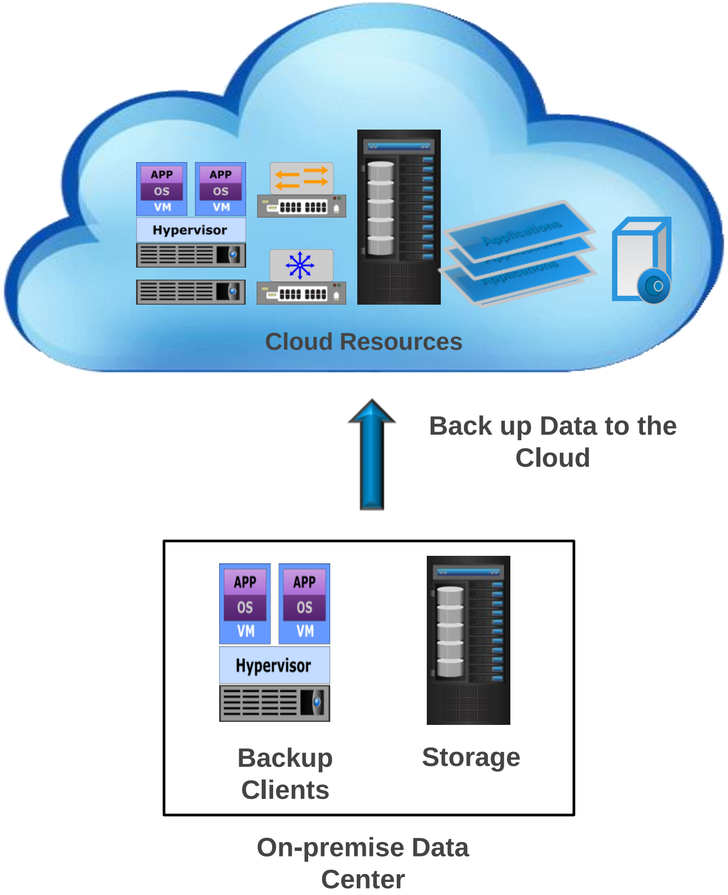
Data Protection as a Service

This lesson covers the following topics:

- Backup as a Service
- Disaster Recovery as a Service

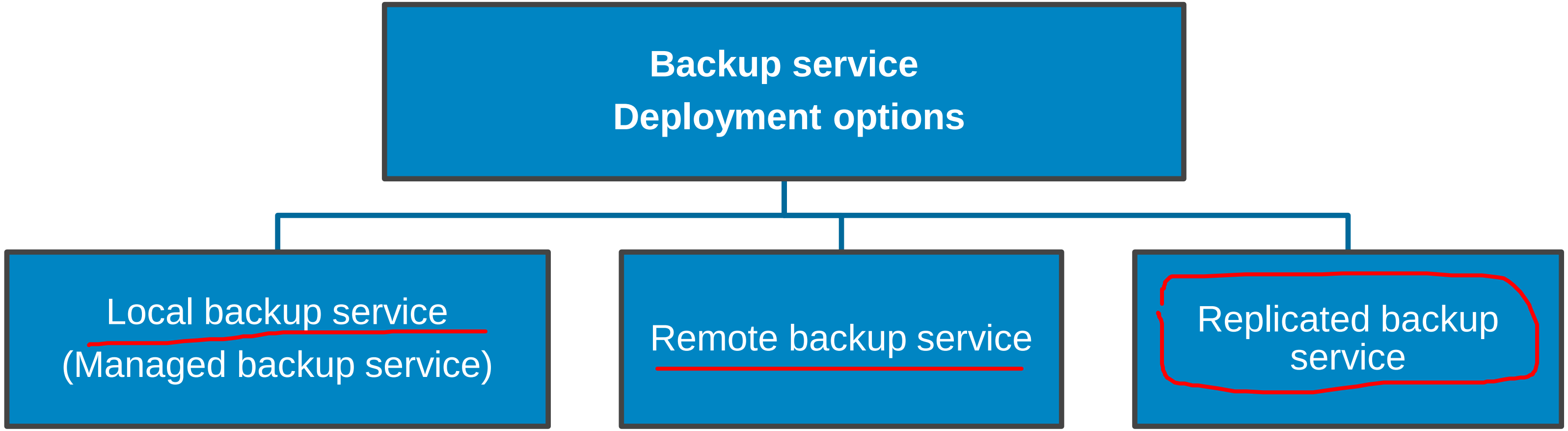
Backup as a service

- ✓ Enables consumers to procure backup services on-demand
- ✓ Reduces the backup management overhead
- ✓ Backing up to cloud ensures regular and automated backup of data
- ✓ Gives the consumers the flexibility to select a backup technology based on their current requirements



DELLEMC

Types of backup services



Backup service
Deployment options

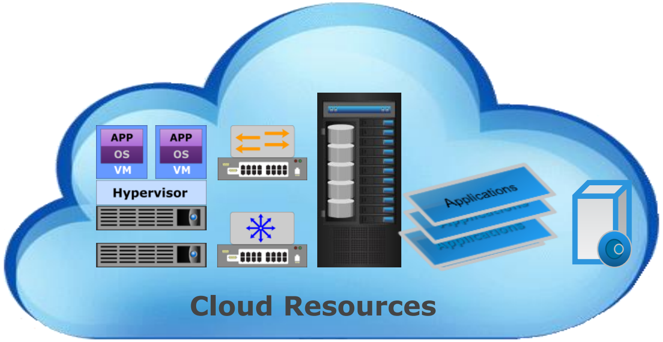
Local backup service
(Managed backup service)

Remote backup service

Replicated backup service

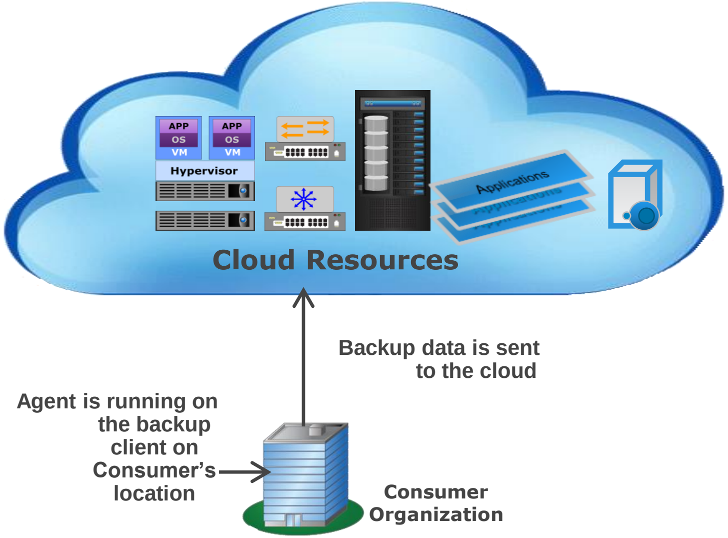
Managed and remote backup services

Managed backup service



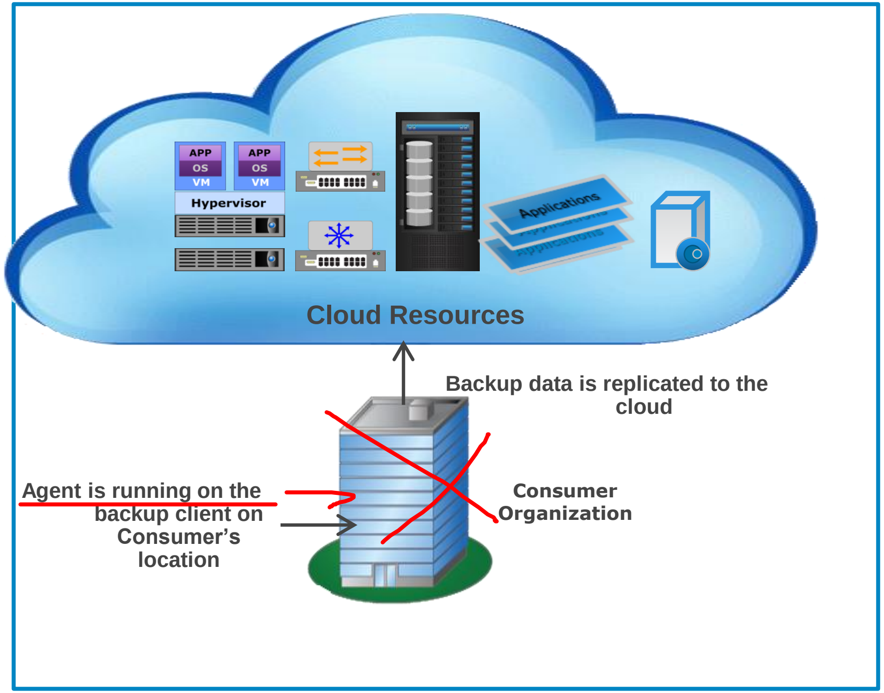
- Suitable when a cloud service provider already hosts consumer applications and data
- Provider provides backup service to protect consumer's data
- The service provider manages backup service

Remote backup service


APP
OS
VM
APP
OS
VM
Hypervisor
Applications
Cloud Resources
Backup data is sent to the cloud
Agent is running on the backup client on Consumer's location
Consumer Organization

- Service provider receives data from consumers
- The service provider manages backup service

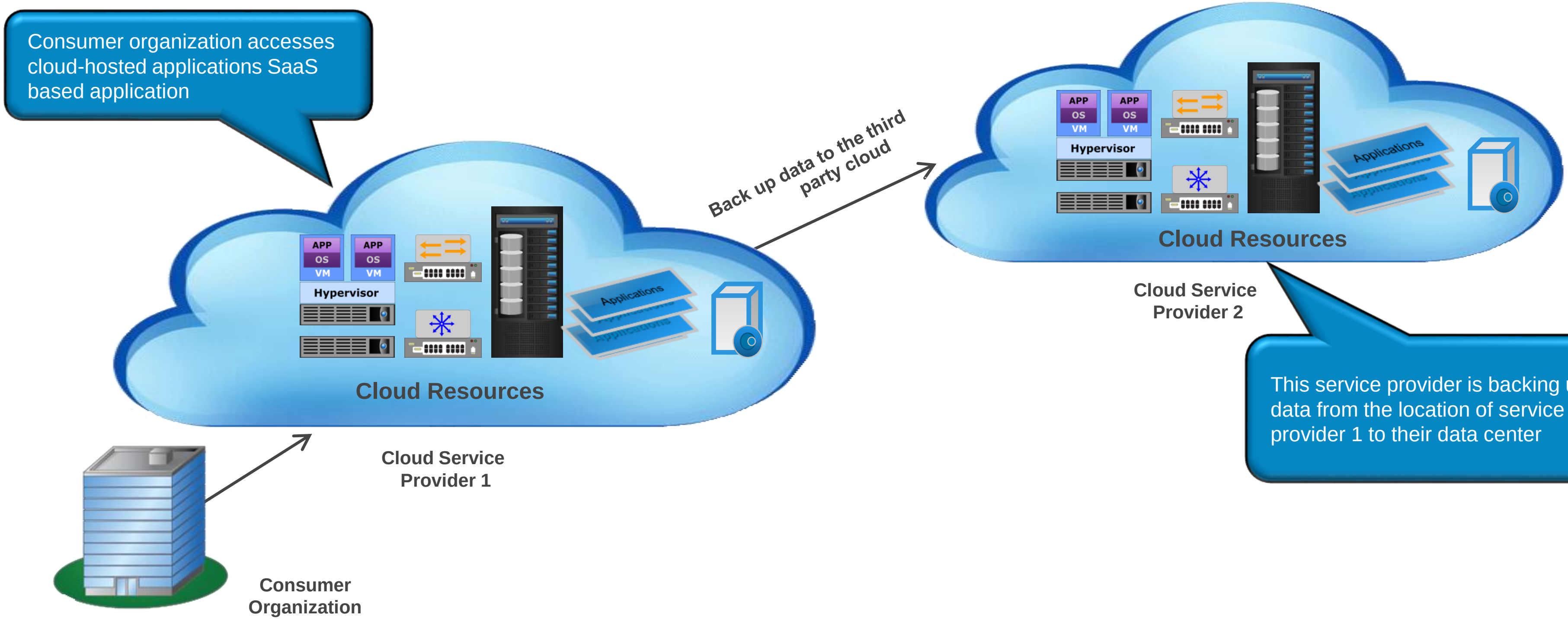
Replicated backup service


APP
OS
VM
APP
OS
VM
Hypervisor
Applications
Cloud Resources
Backup data is replicated to the cloud
Consumer Organization
Agent is running on the backup client on Consumer's location

- Service provider only manages data replication and IT infrastructure, at disaster recovery site
- Consumers manage local backups

Cloud-to-cloud backup

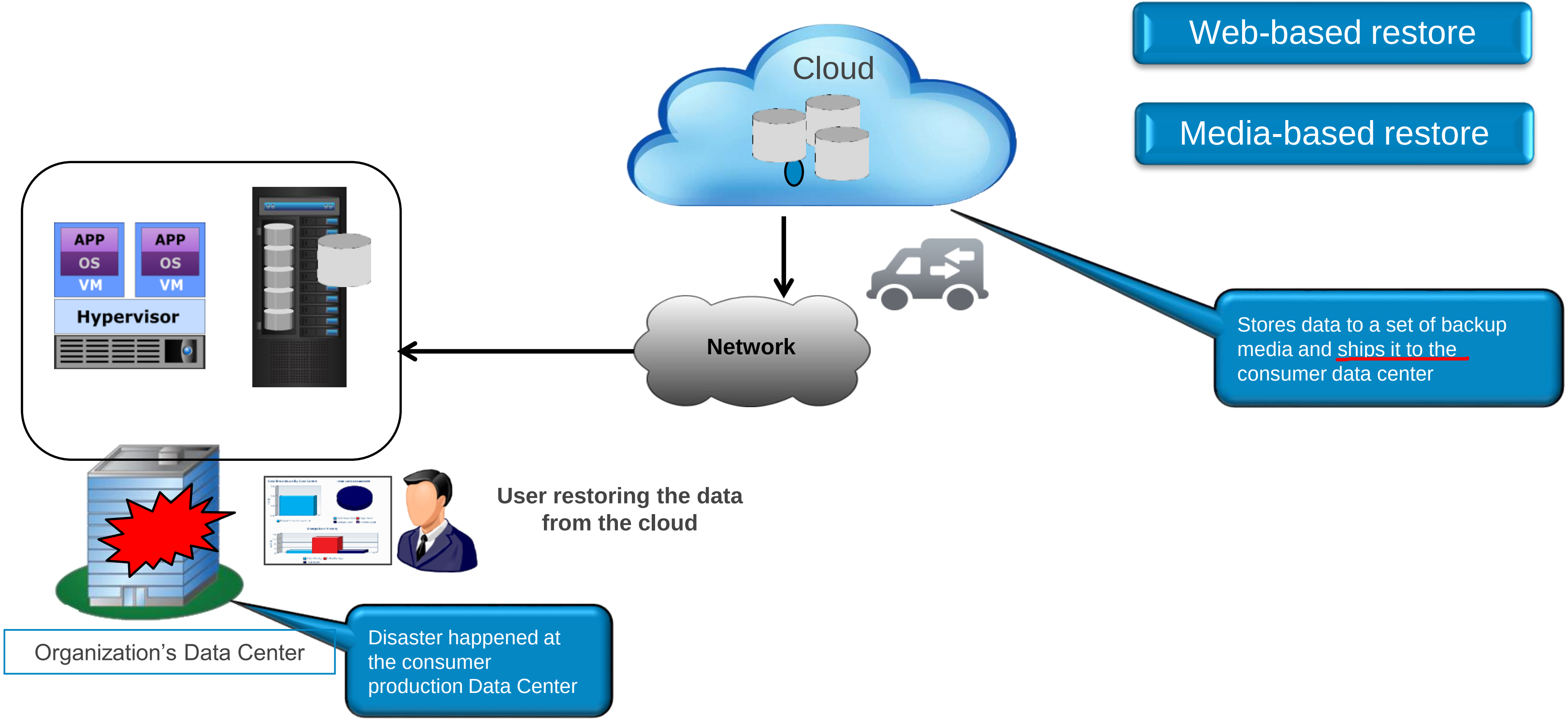
Allows consumers to back up cloud-hosted applications - SaaS data to other cloud



Consumer organization accesses cloud-hosted applications SaaS based application

This service provider is backing up data from the location of service provider 1 to their data center

Restoring data from cloud



Web-based restore

Media-based restore

Stores data to a set of backup media and ships it to the consumer data center

User restoring the data from the cloud

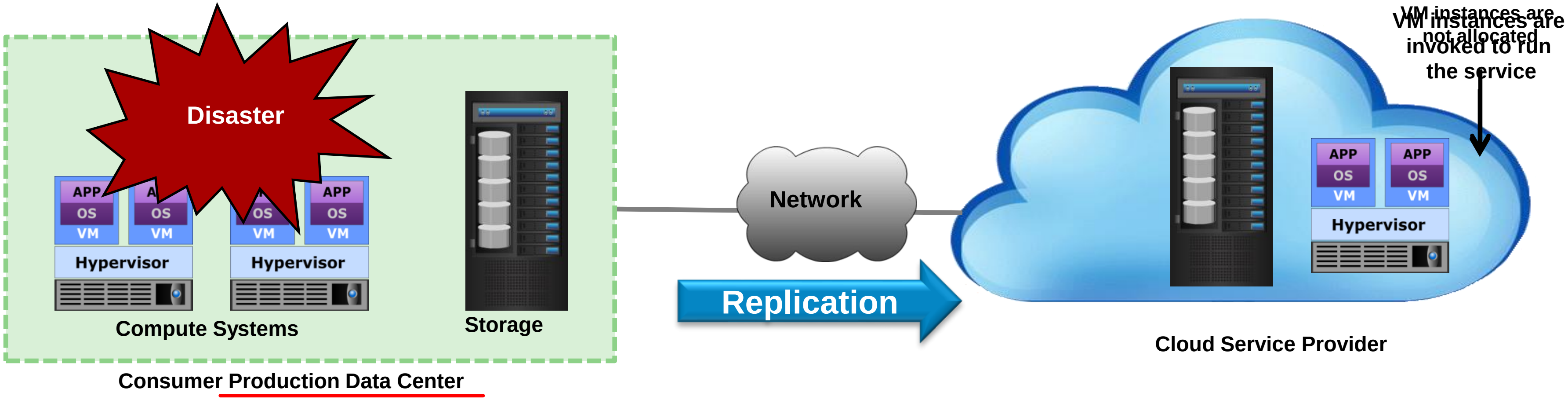
Organization's Data Center

Disaster happened at the consumer production Data Center

Disaster recovery as a service

Service provider offers resources to enable consumers to run their IT services if there is a disaster

Business Disruption



Exercise on service availability

Scenario:

- A storage system is used to provide a data backup service
- Scheduled service time of the service = 24×365 hours
- Last year the storage system failed twice
- Storage system failures resulted in a service downtime of three days

Deliverables:

- What is the achieved availability of the data backup service in the last year?

Exercise debrief

- Scheduled service time of the service = 24 × 365 hours

  = 8760 hours

- Achieved availability of the service in the last year =

$$\frac{\text{Agreed Service Time} - \text{Downtime}}{\text{Agreed ServiceTime}} \times 100 = \frac{8760 - (24 \times 3)}{8760} \times 100$$

$$= 0.9918 \times 100 = 99.18\%$$

Summary

Key points covered in this module:

- Cloud service availability
- Fault Tolerance mechanism
- Data protection solutions
- Data protection as a service