Data storage and data availability

# Storage technology

* Storage devices
* Data access - block, files, object
* Data center End to End View – overview of complete stack including storage, network, host, cluster, applications, virtual machines, cloud storage
* Storage virtualization technologies - RAID level, storage pooling, storage provisioning
* Advance topics in storage virtualization – storage provisioning, thin-provisioning
* Cloud storage – S3, glacier, storage tiering

# High Availability

* Introduction to high availability
* clustering, failover, parallel access

# Disaster Recovery

* Need of disaster recovery (DR)
* Building blocks-global cluster, wide-area-connector (WAC), heartbeat
* Split-brain–problem and solutions
* Preparing for DR–firedrill

# Storage technology

## Storage Devices

We will look at the various persistent storage devices in following section.

### Hard Disk Device (HDD)

The hard disk drives have been the main form of persistent data storage in computer systems for decades and much of the development of file system technology is predicated on their behavior.

#### Geometry

A platter is a circular hard surface on which data is stored persistently by inducing magnetic changes to it. Each platter had 2 sides, which are called as surfaces. These platters are usually made of some hard material (such as aluminum), and then coated with a thin magnetic layer that enables the drive to persistently store bits even when the drive is powered off.

The platters are all bound together around the spindle, which is connected to a motor that spins the platters around at a constant rate when disk is turned on. The rate of rotation is often measured in rotations per minute (RPM), and typical modern values are in the 7,200RPM to 15,000 RPM range.

Data is encoded on each surface in concentric circles of sectors. One such concentric circle is called as a track. A single surface contains many thousands of tracks, tightly packed together.

Process of reading and writing is accomplished by the disk head; there is one such head per surface of the drive. The disk head is attached to a single disk arm, which moves across the surface to position the head over the desired track. Reading is accomplished by sensing magnetic patterns. Writing is accomplished by inducing a change in magnetic pattern.

#### Disk scheduling

There are different I/O scheduling algorithms available.

* Shortest Seek Time First (SSTF)
* Elevator – Also called as SCAN or C-SCAN

### Solid State Devices (SSD)

Solid-state drives (SSDs) are devices that do not have spinning disks. Today's solid-state technologies, such as DRAM and NAND flash, provide faster data access, are more efficient, and have a smaller footprint than traditional spinning disks. The data center uses solid-state technologies in many form factors: in-server, all flash arrays, all flash appliances, and mixed with traditional HDD arrays. Each form factor offers a different value proposition. SSDs also have many connectivity types: PCIe, FC, SATA, and SAS.

Due to the current cost per gigabyte of SSD devices, the best value of SSDs is not as high capacity storage devices. The benefit of adopting SSDs is to improve performance and reduce the cost per I/O per second (IOPS). Data efficiency and placement is critical to maximizing the returns on any data center’s investment in solid state.

File systems that create and remove files often reuse storage blocks by overwriting a storage block with new contents. A Solid-State Drive (SSD) device cannot overwrite a block of storage without erasing it first. This behavior causes a performance cost for writes to the previously used blocks, when compared to writes to unused or erased blocks. To avoid this cost, the TRIM operation informs the SSD which blocks of data are no longer in use and can be erased. The SSDs erase the unused blocks before the blocks are required for reuse, which improves the performance of the future write I/Os to the SSD. The TRIM operation also reduces wear leveling and fragmentation, because unused blocks are erased. The unused data does not get moved during a garbage collection or a cleaning cycle.

### Memory devices

Except for SSDs and HDDs, few other persistent memory devices are being developed.

* MRAM – Magneto resistive RAM
* Phase-change memory
* CRAM - Chalcogenide RAM

## Data Access

Data on storage devices can be accessed using various units.

### Block

A block is minimum unit of size can be addressed by file system individually on the disk. A block consists of sequence of bytes. Normally value of blocks are power of 2.

Disk space is allocated in 512-byte sectors to form logical blocks. The smallest available block size for VxFS is 1 KB. VxFS supports logical block sizes of 1024, 2048, 4096 and 8192 bytes.

A block size is chosen based on the type of application being run. For example, if there are many small files, a 1 KB block size may save space. For large file systems, with relatively few files, a larger block size is more appropriate. Larger block sizes use less disk space in file system overhead (like metadata), but it consumes more space for files that are not a multiple of the block size. The easiest way to judge which block sizes provide the greatest system efficiency is to try representative system loads against various sizes and pick the fastest.

### Files

A file is a named collection of related information which is recorded on storage devices. From user’s perspective, a file is the small allocation unit on storage device.

#### File types

* Regular files
* Directories
* Symbolic links
* Hard Links
* Named pipes
* Special files

#### Basic file properties

* File type
* Owner and group
* Access permissions
* Link counts
* Size
* Name
* Time stamps – creation, modification, access

### Objects

Object are minimum units of data that you can store in cloud storage. Normally objects are arranged in buckets in the cloud storage. Objects typically represent a file in cloud storage. Objects normally consist of 2 components.

* Object Data
* Object metadata

#### Properties of objects

* Name
* Version or generation number

An object is variable-length and can be used to store any type of data, such as files, database records, medical images or multimedia. A single object could even be used to store an entire filesystem or database.

Data center

In this topic we will look at End to End View of a data center. It will include overview of complete stack including storage, network, host, cluster, applications, virtual machines, cloud storage. A data center is dedicated space with building or set of buildings used to place computer systems for storage of data.

### Storage

Many data centers still rely on three traditional configurations:

1. direct-attached storage (DAS)
2. network-attached storage (NAS)
3. storage area network (SAN)

#### Direct-attached storage (DAS)

In case of Direct-attached storage (DAS), HDDs or SDDs are directly attached to the computer rather than over network like ethernet, fiber optics etc.

##### Typical DAS interface

* Serial-Attached SCSI (SAS)
* Serial Advanced Technology Attachment (SATA)
* Small Computer System Interface (SCSI)
* Peripheral Component Internet Express (PCIe)

##### Advantages of DAS

* It typically provides better performance than networked storage solutions such as NAS and SAN. Data is in proximity of computer system. It avoids network bottleneck such as network speed, congestion.
* It is cheaper and easier to implement and maintain.

##### Disadvantages of DAS

* Limited Scalability

#### Network-attached storage (NAS)

NAS is a file-level storage device that enables multiple users and applications to access data from a centralized system via the network. A NAS device is an independent node on the local area network (LAN) with its own IP address. It is essentially a server that contains multiple HDDs or SSDs, along with processor and memory resources. The device typically runs a lightweight operating system (OS) that manages data storage and file sharing.

##### NAS protocols

* Network File System (NFS)
* Common Internet File System (CIFS)
* Server Message Block (SMB).
* Internetwork Packet Exchange (IPX)
* NetBIOS Extended User Interface (NetBEUI)
* Apple Filing Protocol (AFP)
* Gigabit Ethernet (GigE)

##### Advantages of NAS

* easy to deploy and operate and relatively inexpensive when compared to SANs
* can be scaled out or integrated with cloud services
* built-in redundancy while offering a great deal of flexibility

##### Disadvantages of NAS

* compete with other traffic on the network, can be mitigated by use of private network
* scalability in case of large number of users

#### Storage Area Network (SAN)

A SAN is a dedicated, high-speed network that interconnects one or more storage systems and presents them as a pool of block-level storage resources.

##### Components of SAN

* application servers
* storage management software
* host bus adapters (HBAs)
* high-speed cabling
* special switches for routing traffic

##### Networks used in SAN

* Fibre Channel
* Internet SCSI (iSCSI)
* Fibre Channel over Ethernet (FCoE)
* NVMe over Fabrics (NVMe-oF)

##### Advantages of SAN

* high availability and scalability
* centralized management, failover protection, and disaster recovery
* improved storage resource utilization

##### Disadvantages of SAN

* Complex environment that can be difficult to deploy and maintain
* Costly to build and maintain

### Network

Types of network topologies used at the data centers.

* Clos Network
* Fat-tree
* Jellyfish
* BCube

### Hosts

In many ways, servers are the engines of data center networking architecture. They store valuable data, provide the processing power for computing workloads, and host various applications and services. While they may appear to take up a relatively small space on a typical data center topology map, it’s important to remember that the entire network infrastructure is set up to facilitate server performance.

High-density server deployments tend to have higher requirements in terms of cabling, cooling, and power supplies. Many colocation customers often want to place their equipment in racks with easy access to direct connections and single cross-connects that offer them improved performance and speed with minimal risk of downtime.

### Cluster

Cluster Server connects multiple, independent systems into a management framework for increased availability. Each system, or node, runs its own operating system and cooperates at the software level to form a cluster. VCS links commodity hardware with intelligent software to provide application failover and control. When a node or a monitored application fails, other nodes can take predefined actions to take over and bring up services elsewhere in the cluster.

#### Types of cluster

* Campus cluster – Servers in the system are co-located at same data center.
* Global cluster – Servers are in multiple different data centers.

### Applications

* Database
* Media server
* Web hosts
* Big Data
* Cloud computing
* Video streaming services

### Virtual machines

A virtual machine (VM) is a virtual environment that functions as a virtual computer system with its own CPU, memory, network interface, and storage, created on a physical hardware system (located off- or on-premises). Software called a hypervisor separates the machine’s resources from the hardware and provisions them appropriately so they can be used by the VM.

The physical machines equipped with a hypervisor is called the host machine. The many VMs that use its resources are guest machines. The hypervisor treats compute resource like CPU, memory, and storage from host machine as a pool of resources that can easily be relocated between existing guest machines or to new virtual machines.

VMs are isolated from the rest of the system, and multiple VMs can exist on a single piece of hardware, like a server. They can be moved between host servers depending on demand or to use resources more efficiently.

VMs allow multiple different operating systems to run simultaneously on a single computer—like a Linux® distro on a MacOS laptop. Each operating system runs in the same way an operating system or application normally would on the host hardware, so the end user experience emulated within the VM is nearly identical to a real-time operating system experience running on a physical machine.

#### Types of hypervisors

* Type 1 – A type 1 hypervisor is on bare metal. VM resources are scheduled directly to the hardware by the hypervisor. KVM is an example of a type 1 hypervisor. KVM was merged into the Linux® kernel in 2007, so if you’re using a modern version of Linux, you already have access to KVM.
* Type 2 – A type 2 hypervisor is hosted. VM resources are scheduled against a host operating system, which is then executed against the hardware. VMware Workstation and Oracle VirtualBox are examples of type 2 hypervisors.

Server consolidation is a top reason to use VMs. Most operating system and application deployments only use a small amount of the physical resources available when deployed to bare metal. By virtualizing your servers, you can place many virtual servers onto each physical server to improve hardware utilization.

This keeps you from needing to purchase additional physical resources, like hard drives or hard disks, as well as reducing the need for power, space, and cooling in the datacenter. VMs provide additional disaster recovery options by enabling failover and redundancy that could previously only be achieved through additional hardware.

A VM provides an environment that is isolated from the rest of a system, so whatever is running inside a VM won’t interfere with anything else running on the host hardware.

Because VMs are isolated, they are a good option for testing new applications or setting up a production environment. You can also run a single purpose VM to support a specific process.

### Cloud storage

The cloud storage is a storage space available to store data on remote servers which can be accessed from the cloud (or the internet). Cloud storage uses data centers with massive computer servers that physically store the data and make it available online to users via web.

Storage virtualization technologies

Storage virtualization is the pooling of multiple physical storage arrays from SANs and making them appear as a single virtual storage device. It is also known as Software defined storage (SDS). The pool can integrate storage hardware from different networks, vendors, or data centers into one logical view and manage them from a single point.

The Storage Network Industry Association (SNIA) has a unique storage virtualization definition: ‘The application of virtualization to storage services or devices for the purpose of aggregating functions or devices, hiding complexity, or adding new capabilities to lower level storage resources’

Challenges overcame by storage virtualization

* Vendor lock-in
* Data migration across arrays
* Scalability
* Redundancy
* Performance
* High costs
* Management

### RAID level

Redundant Array of Independent Disks (RAID) is one of the oldest and still active technologies to achieve always-on status. Developers designed RAID to improve redundancy and performance in storage systems.

#### Types of RAID

* Hardware RAID – A dedicated hardware controller provides hardware-based RAID services. IT can deploy hardware RAID two ways: an external RAID Controller Card or internal RAID-on-Chip.
  + **RAID Controller Card:** This plug-in expansion card connects to a PCIe or PCI-X motherboard slot. The card contains a RAID processor and I/O processors with drive interfaces. The cards are expensive, but since they are independent of the host, all RAID operations are offloaded from the CPU to the dedicated card.
  + **RAID-on-Chip:** A single chip on the motherboard integrates the host interface, I/O interfaces for HDDs, the RAID processor, and a memory controller.
* Software RAID – Software-based RAID delivers RAID services from the host. Software RAID comes in two flavors: pure software defined running from the OS, and hybrid software that contains a hardware component to relieve the load on the CPU.
  + **Software-only:** Software RAID is the least expensive of the RAID types and is often included as a native function on the OS. It is a host-based software application that manages RAID calculations for attached hard disk drives. It’s attached via a HBA or native I/O interface and activates when the OS loads the RAID driver.
  + **Hybrid:** This software-based RAID uses a hardware component to deliver RAID BIOS functions from RAID BIOs on the motherboard or on an HBA. This technology offers a layer of redundant protection from a faulty boot process. Software-only RAID boots from the operating system, and boot errors could affect the entire RAID subsystem. The addition of a RAID BIOS hardware component protects the subsystem from operating system boot errors.

#### **RAID Levels**

* RAID 0: Striping

Striping (RAID-0) is useful if you need large amounts of data written to or read from physical disks, and performance is important. Striping is also helpful in balancing the I/O load from multi-user applications across multiple disks. By using parallel data transfer to and from multiple disks, striping significantly improves data-access performance. Striping maps data so that the data is interleaved among two or more physical disks. A striped plex contains two or more subdisks, spread out over two or more physical disks. Data is allocated alternately and evenly to the subdisks of a striped plex. The subdisks are grouped into “columns,” with each physical disk limited to one column. Each column contains one or more subdisks and can be derived from one or more physical disks. The number and sizes of subdisks per column can vary. Additional subdisks can be added to columns, as necessary.

A picture containing chart

Description automatically generated

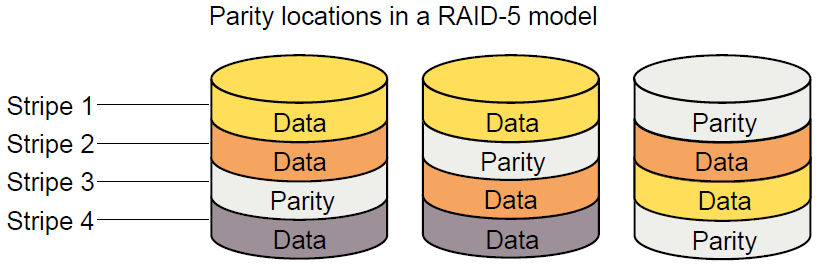
* RAID 1: Mirroring

Mirroring uses multiple mirrors (plexes) to duplicate the information contained in a volume. In the event of a physical disk failure, the plex on the failed disk becomes unavailable, but the system continues to operate using the unaffected mirrors. Similarly, mirroring two LUNs from two separate controllers lets the system operate if there is a controller failure. Although a volume can have a single plex, at least two plexes are required to provide redundancy of data. Each of these plexes must contain disk space from different disks to achieve redundancy. When striping or spanning across many disks, failure of any one of those disks can make the entire plex unusable. Because the likelihood of one out of several disks failing is reasonably high, you should consider mirroring to improve the reliability (and availability) of a striped or spanned volume.

* RAID 5: Striping with Parity

RAID-5 provides data redundancy by using parity. Parity is a calculated value used to reconstruct data after a failure. While data is being written to a RAID-5 volume, parity is calculated by doing an exclusive OR (XOR) procedure on the data. The resulting parity is then written to the volume. The data and calculated parity are contained in a plex that is “striped” across multiple disks. If a portion of a RAID-5 volume fails, the data that was on that portion of the failed volume can be recreated from the remaining data and parity information. It is also possible to mix concatenation and striping in the layout.

Every stripe has a column containing a parity stripe unit and columns containing data. The parity is spread over all of the disks in the array, reducing the write time for large independent writes because the writes do not have to wait until a single parity disk can accept the data.



* RAID 6: Striping with double parity

This RAID level operates like RAID 5 with distributed parity and striping. The main operational difference in RAID 6 is that there is a minimum of four disks in a RAID 6 array, and the system stores an additional parity block on each desk. This enables a configuration where two disks may fail before the array is unavailable. Its primary usage case or application servers and large storage arrays.

RAID 6 offers higher redundancy than 5 and increased read performance. It can suffer from the same server performance overhead with intensive write operations. This performance hit depends on the RAID system architecture: hardware or software, if it’s located in firmware, and if the system includes processing software for high-performance parity calculations.

* RAID 10: Striping and Mirroring

Striping above mirroring (RAID10) layout is also called a striped-mirror layout. Putting mirroring below striping mirrors each column of the stripe. If there are multiple subdisks per column, each subdisk can be mirrored individually instead of each column. A striped-mirror volume offers the dual benefits of striping to spread data across multiple disks, while mirroring provides redundancy of data. In addition, it enhances redundancy, and reduces recovery time after disk failure.

Diagram

Description automatically generated

### Storage pooling

Storage pools are capacity aggregated from disparate physical storage resources in a shared storage environment. Storage pools can be configured in varying sizes and provide several benefits, including performance, management and data protection improvements. Pools can be provisioned to include any amount of capacity and use any combination of physical storage space in a storage area network (SAN). In virtual server environments, virtual machines (VMs) can be stored on dedicated pools, ensuring critical VMs have access to the proper amount of storage.

vendors often provide the capability for administrators to provision storage pools through the management interface. For example, the Storage Spaces feature in Windows Server 2012 provides pooling capabilities, and array vendors such as IBM and EMC also allow this type of provisioning as well.

A storage volume is the basic unit of storage, such as allocated space on a disk or a single tape cartridge. A storage pool is a collection of storage volumes. The server uses the storage volumes to store backed-up, archived, or space-managed files. The group of storage pools that you set up for the TSM server to use is called server storage. Storage pools can be arranged in a storage hierarchy.

The server has two types of storage pools that serve different purposes: primary storage pools and copy storage pools.

#### Primary Storage Pool

When a user tries to restore, retrieve, recall, or export file data, the requested file is obtained from a primary storage pool if possible. Primary storage pool volumes are always located onsite.

A primary storage pool can use random access storage (DISK device class) or sequential access storage (for example, tape or FILE device classes).

The server has a default DISKPOOL storage pool that uses random access disk storage. You can easily create other disk storage pools and storage pools that use tape and other sequential access media by using Device Configuration Wizard in the TSM Console.

The server does not require separate storage pools for archived, backed-up, or space-managed files. However, you may want to have a separate storage pool for space-managed files. Clients are likely to require fast access to their space-managed files. Therefore, you may want to have those files stored in a separate storage pool that uses your fastest disk storage.

### Copy Storage Pool

When an administrator backs up a primary storage pool, the data is stored in a copy storage pool. A copy storage pool can use only sequential access storage (for example, a tape device class or FILE device class).

The copy storage pool provides a means of recovering from disasters or media failures. For example, when a client attempts to retrieve a file and the server detects an error in the file copy in the primary storage pool, the server marks the file as damaged. At the next attempt to access the file, the server obtains the file from a copy storage pool. For details, see Restoring Storage Pools, Using Copy Storage Pools to Improve Data Availability, recovering a Lost or Damaged Storage Pool Volume, and Maintaining the Integrity of Files.

You can move copy storage pool volumes offsite and still have the server track the volumes. Moving copy storage pool volumes offsite provides a means of recovering from an onsite disaster.

### Storage provisioning

Storage provisioning is the process of assigning storage, usually in the form of server disk drive space to optimize the performance of a storage area network (SAN). The storage template is an efficient way to quickly create storage based on the configurations that are mentioned in the template. This mechanism lets you specify storage attributes and conditions, and the template can be used by the storage operators to allocate storage. It ensures that the allocated storage contains attributes mentioned in the template, and it adheres to the conditions that are specified in the template.

Storage provisioning lets you create storage templates from existing virtual storage devices such as volumes and file systems. It extracts all the properties of the virtual storage device and recreates the entire stack.

With storage provisioning, you determine the virtual device characteristics of the storage object using the file system, volume manager, and operating system utilities and can create a storage template based on this information. A storage template lets you migrate existing virtual devices to new ones with different properties. For example, you can migrate volumes from thick LUNs to thin LUNs.

### Thin provisioning

Thin provisioning is a storage array feature that optimizes storage use by allocating and reclaiming the storage on demand. With thin provisioning, the array allocates storage to applications only when the storage is needed, from a pool of free storage.

Thin provisioning solves the problem of under-utilization of available array capacity. Administrators do not have to estimate how much storage an application requires. Instead, thin provisioning lets administrators provision large thin or thin reclaim capable LUNs to a host. When the application writes data, the physical storage is allocated from the free pool on the array to the thin-provisioned LUNs.

The two types of thin provisioned LUNs are thin-capable or thin-reclaim capable. Both types of LUNs provide the capability to allocate storage as needed from the free pool. For example, storage is allocated when a file system creates or changes a file. However, this storage is not released to the free pool when files get deleted. Therefore, thin-provisioned LUNs can become 'thick' over time, as the file system starts to include unused free space where the data was deleted. Thin-reclaim capable LUNs address this problem with the ability to release the once-used storage to the pool of free storage. This operation is called thin storage reclamation.

The thin-reclaim capable LUNs do not perform the reclamation automatically. The server using the LUNs must initiate the reclamation. The administrator can initiate a reclamation manually, or with a scheduled reclamation operation.

Array-based options like Thin Storage and Thin Provisioning help storage administrators to meet the challenges in managing their storage. These challenges include provisioning the storage, migrating data to maximize storage utilization, and maintaining the optimum storage utilization.

#### Reclaiming space on Solid State Devices (SSDs) with the TRIM operation

File systems that create and remove files often reuse storage blocks by overwriting a storage block with new contents. A Solid-State Drive (SSD) device cannot overwrite a block of storage without erasing it first. This behavior causes a performance cost for writes to the previously used blocks, when compared to writes to unused or erased blocks. To avoid this cost, the TRIM operation informs the SSD which blocks of data are no longer in use and can be erased. The SSDs erase the unused blocks before the blocks are required for reuse, which improves the performance of the future write I/Os to the SSD. The TRIM operation also reduces wear leveling and fragmentation, because unused blocks are erased. The unused data does not get moved during a garbage collection or a cleaning cycle.

## Cloud storage

Cloud storage is becoming quite popular because it’s a powerful tool that allows businesses to meet their computing and storage needs while keeping business and customer data safe. For companies looking to get in on the online storage bandwagon, this guide will help them decide on the best storage option for them. It also discusses online storage, how it works, how safe it is, the different types of storage options available, and the advantages and disadvantages.

#### What is Cloud Storage?

It refers to file storage online. Instead of using local storage options (think flash drive, hard disk, or external hard disk) to store your files, you save them online. This cloud computing model allows the storage of data on remote servers and accesses them over the Internet (cloud). It is managed, maintained, and operated by a service provider on storage servers built on virtualization techniques.

The primary premise of the online storage service is user convenience. In today’s dynamic business world, your customers and employees need instant access to information and data. Ergo, online storage bridges this gap allowing you to access a file or app on your device (PC, tablet, phone, or desktop). It also offers you enhanced performance, extra storage space, and a secure platform to store a vital file, app, or data log if you do not wish to use your on-premises storage.

#### How Does Cloud Storage Work?

There are numerous different online storage systems in existence. Some have a specific focus such as storing digital pictures or web email messages while others store various forms of digital data. Some systems are small operations that use only one data server while others (enterprise storage solutions) are so large that their physical equipment fills up a warehouse.

An online storage system only needs one data server that is connected to the Internet to work. Here, you (the client) send a copy of your file to the data center over the Internet, which then records the information. When you wish to retrieve the file, you can access the data server via a web-based interface. The server then allows you access to the file or sends it back to you.

Cloud-based storage systems usually rely on many data servers because, since computers occasionally require repair or maintenance, it is essential to back-up the data on multiple machines running on different power supplies. Without this redundancy, the online storage systems could not ensure clients instant access to their data at any given time.

Cloud backup is essential to your business because it allows you to recover your information if something happens to your on-premises storage.

Reputable online storage services protect your data behind encryption and require a password to gain access. Other times, your file can be stored behind multiple-factor authentication that requires a password plus another code generated independently upon the login request to gain access.

There are different methods of uploading a file or app to your online account. Some online storage services support in-browser uploads only, which means that you must log in to the service’s website to upload data. Others provide you with a desktop app that makes it easier to drag-and-drop a file into the service’s dedicated folder.

Once your files and data are stored online, the features you receive include the ability to: access a file or app from your mobile device, stream music and videos, share files with others via a special link, encrypt your files to prevent others from accessing them, download files back to your computer, delete files to free up space, and more.

#### What is the Best Cloud Storage Service?

With the multitude of CSPs in the market, you need to be wise about choosing a provider who will offer you the maximum amount of bandwidth and low-cost storage while keeping your data safe. While there are several reliable public CSPs in the market, you need to select the best option that meets all your needs.

Many companies are turning to a hybrid type of cloud computing and storage or integrated information-management platforms. The hybrid model combines the advantages of both public and private clouds infrastructure to come up with better flexibility and security at a lower cost. This model bridges the gap between the low-security public cloud and the high-security private cloud. You should note that an arbitrary server that is connected to a public CSP is not a hybrid network. To be considered a hybrid, the private cloud infrastructure must run on a cloud service.

The best online storage system must do the following:

* Improve resource utilization
* Lower IT-related capital and operational expenses
* Introduce quick and flexible functionality
* Strengthen disaster recovery capabilities
* Increase the response and delivery speed of IT resources to end users
* Enable balancing of cost, scaling and isolation requirements
* Help enterprises build an exit strategy
* Drive innovation and help build a future-proof business
* Offers better network optimization since it shifts the heavy lifting to in-house servers, therefore, improving connection reliability by leveraging the CSP’s network investment.

#### How Safe is Enterprise Cloud Storage?

The safety of your data stored on the cloud is relative to the online storage solution you use. Cloud providers store their backups in multiple locations across different power supplies to prevent data loss. These data centers use sophisticated systems that provide emergency power, detect smoke, and suppress fires. Furthermore, they are heavily guarded, reinforced, and internally protected to prevent intrusion, stealing the storage software, or physical destruction from disgruntled employees.

Cloud systems use authentication processes such as passwords and usernames to limit access hence safeguarding your data and preventing unauthorized access. Data encryption and multiple-factor authentication are also used to protect data that is intercepted in route or stolen from the server. However, passwords can be hacked, and a rogue employee can gain access to your data’s encryption keys. Also, your data is not immune to government search and seizure.

So is your data safe on the cloud? Well, the truth is that no cloud service provider can assure you of 100 percent data security. In any case, since most online storage companies operate in a cut-throat market, they rely heavily on their reputation, and you can rest assured that they will take great pains to ensure that they have the most advanced security measures available to protect clients’ data.

This is an age where governments have been exposed for unwarranted tapping into otherwise private cloud data. You would be wise not to store sensitive personal or company information on the cloud. CSPs, however, use more complex security methods to protect data than most local storage options.

While all online storage systems keep data secure from loss caused by fires, hurricanes, floods, and computer meltdowns, your information is still vulnerable because it is with a third party. Therefore, before selecting a provider, do your research.

What’s more, online storage is more reliable when it is used together with another storage system; for example, Google Drive. Since the biggest concern with cloud services is lost data and not hacked data, the issue is eliminated if you use the cloud as more of a sharing platform rather than a storage one.

Although online storage services may not be 100 percent secure, their benefits far outweigh the potential risk. This is evidenced by many companies that use these services extensively.

#### What are the Different Types of Cloud Storage?

1. Public Cloud

This cloud service allows you to have your cloud in a shared infrastructure. Here, the provider owns and manages all the cloud infrastructure (hardware and software) and delivers it over the Internet as a cloud service. Imagine the public cloud as a large apartment complex. The CSP is the landlord while individual and enterprise clients using the service are the tenants. The principle of shared resources allows the CSP to offer lower prices.

You benefit a lot from sharing of these resources since industry regulation compliance is high and infrastructure investments are top quality to attract customers. The other benefits of the public cloud are:

* Highly scalable
* Cost-effective
* Highly reliable
* No maintenance costs incurred on businesses
* Expert monitoring

1. Private Cloud

Unlike the public cloud that relies on sharing of resources, the private cloud is used by one organization alone. It is a preferred solution for larger enterprises with mission-critical data that requires additional security. Financial institutions, government agencies, and healthcare organizations are likely to prefer a private cloud solution.

The infrastructure can either reside with a third-party vendor or on-premises. Regardless, the hardware and services remain dedicated. Its advantages include:

* Allows you greater control over your data
* Exclusive access
* Highly scalable
* Highly efficient
* Offer greater security
* It's customizable

The private cloud model is costly compared to a public one because it requires you to own and manage your data center and IT team. Outsourcing from a third-party vendor is also expensive.

1. Hybrid Cloud

The hybrid system employs both private and public clouds meaning that while your IT team manages a section of the cloud in-house, the rest is managed off-site.

1. Community Cloud

This cloud infrastructure allows data access from different organizations or departments. This collaborative effort allows the sharing of infrastructure between several entities from a specific community that has common concerns (compliance, security, jurisdiction, etc.) whether managed by a third-party or internally. For instance, different branches of the same company that are in different towns. Different government functionalities can also access government data via the community cloud infrastructure.

1. Mobile Cloud

Mobile online storage allows you to store and manage your files, photos, videos, and music from your mobile devices (tablets, laptops, and smartphones). Today, many new mobile devices are outfitted with preloaded and configured online storage that can be used to back-up your files. Android OS devices use Google Drive, Apple iOS devices use iCloud, while Samsung Galaxy has partnered with Dropbox.

#### What are the Advantages of Enterprise Cloud Storage Service?

* Cost

Your company is more likely to cut back on operating costs by using cloud-based services compared to using external hard drives or in-house hosting solutions. It also eliminated the cost of time taken to complete routing backups manually. You do not need to spend more money on additional servers, which are expensive to maintain. Online storage gives you plenty of space to store your data, often for low, affordable prices.

* Accessibility

Local storage only allows you access to your files or data from a location. However, with online storage, the devices that you use daily become access points. Thus, you can access all your files, photos, apps, videos, and folders in the cloud from any place across the world. Of course, you must have Internet access and the necessary credentials. This eliminates the stress of transferring files between devices every time you want to access them, which is, frankly, annoying and complicated.

* Recovery

Perhaps the most significant advantage of using online storage is that you are assured of a back-up solution should something go wrong and your data is lost or corrupted. If something happens to the files stored on your computer, you can always access and retrieve them from the cloud with minimal downtime.

* Syncing

You can sync all your files in the cloud; therefore, if you make any changes to one file, the cloud will automatically sync your changes across all affiliated devices. This means that your files are similar across all devices and you will have the latest version of your file irrespective of how you chose to retrieve it.

* Increased Security

Most CSPs add extra security layers to their servers and protocols to protect all data stored therein from outside hacking or physical attacks. To gain a competitive edge, CSPs employ the latest enterprise-grade security measures that businesses would not usually apply. Data is encrypted at rest and in flight making unauthorized users unable to access the data.

* Collaboration Tools

Sending dozens of emails to share files is cumbersome. The cloud makes it simple to sync and share one or more files with several recipients. Additionally, storage services enable you to collaborate with peers and availing editing capabilities to authorized users. Furthermore, such changes are saved automatically and shared easily with all collaborators.

* Space Efficiency

Once you move your data into the cloud, you free up space on your local devices and hardware. Therefore, it no longer takes up valuable space in your premises since it is stored virtually on your CSP’s servers.

#### What are the Disadvantages of Cloud Storage?

* Dependent on Internet Connection

You must have an Internet connection to gain access to the cloud. Since online storage solutions are dependent on your Internet upload and download speed, low latency can seriously impede you from accessing your app or data in real time.

* Backups May Be Slower

Online storage and internet bandwidth max ingest speeds may be more limited than your local network or disk. This may not be an issue for file servers and workstations backups since they can run in the background as work continues. Once the initial full backup is complete, incremental backups reduce the backup size and times while deduplication and compression help, too.

However, specialized applications such as databases (for instance Microsoft SQL Server) can be affected by long backup times. In this case, you want these applications backed up within your maintenance windows or during low activity periods.

* Restores May Be Slower

Here, it is all about rated online storage speed and Internet bandwidth. Restoring the entire server after a crash or damage may take longer, but file-level restore is equally as fast. The takeaway is to ensure that you meet your recovery time objectives. If this is not possible, you may consider performing hybrid backups on your critical servers and send the backups to both cloud and local storage. Two backup copies allow fast local restores (with no reliance on cloud vendor and Internet availability) as well as the benefit of off-site storage protection for disaster recovery.

* Higher Internet Utilization

If your backups run during peak business hours, the heavy Internet use affects all Internet-related activity performance. You need to control your Internet bandwidth by setting up utilization rules to make sure that your Internet connection is not saturated during times when your Internet access is required for other critical business activity.

* High Costs

For enterprises and corporations, online storage devices are sound solutions; however, private solutions can be quite costly for small businesses.

* Improper Handling of Login Details

Beware of improper handling of your login credentials since if you misplace them and someone gains access to your user-ID and password they can gain access to your account. You must use complex passwords that no one else can guess and avoid keeping them in your devices such as HDD or pen drive. Finally, if you forget your credentials, recovering them can be a cumbersome process.

* Customer Support

Customer support has not been one of the strongest suits of many CSPs, which is why you should take a closer look at the company’s terms and agreements as well as its online forums and FAQ section.

* Privacy Issues

Even the best CSP relationship suffers from one issue; a third party manages your data. This means that you have limited control over the data if you opt for a public CSP. There is also the issue of ownership since after migrating your data to your CSP’s servers, it may be difficult to determine who, between you and the company, owns the information.

#### What is Hybrid Cloud Storage?

Just like its name suggests, the hybrid cloud converges the private and public cloud solutions into a unique and customizable type of online storage, which offers you the advantages of both. Here, applications and data can move freely between private and public clouds when needed, providing increased flexibility.

Most enterprises can use the public cloud for some aspects of their work that are not sensitive such as webmail. However, these businesses still need a private cloud service to store their data logs hence the hybrid cloud solution. Its advantages include:

* It allows for greater control
* It is cost-effective
* It is highly customizable

For instance, banks could adopt the hybrid cloud system and have their client communication accounts stored in a public cloud, but their confidential account details stored in a private cloud service.

Enterprise cloud storage has become the leading and preferred method of digital data storage, especially for small and medium-sized enterprises. It is safer than local storage because it utilizes industry-grade safety measures to secure data. It is also space and cost-efficient, accessible across the globe, allows you to sync and share files quickly, and prevents loss of data.

## Amazon S3

Amazon Simple Storage Service (Amazon S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance. This means customers of all sizes and industries can use it to store and protect any amount of data for a range of use cases, such as data lakes, websites, mobile applications, backup and restore, archive, enterprise applications, IoT devices, and big data analytics. Amazon S3 provides easy-to-use management features so you can organize your data and configure finely tuned access controls to meet your specific business, organizational, and compliance requirements. Amazon S3 is designed for 99.999999999% (11 9's) of durability, and stores data for millions of applications for companies all around the world.

#### AWS S3 Benefits

* Industry-leading performance, scalability, availability, and durability
* Wide range of cost-effective storage classes
* Unmatched security, compliance, and audit capabilities
* Easily manage data and access controls
* Query-in-place services for analytics
* Most supported cloud storage service

#### Use Cases

* Backup and restore
* Disaster recovery (DR)
* Archive
* Data lakes and big data analytics
* Hybrid cloud storage
* Cloud-native applications

## Amazon Glacier

Amazon S3 Glacier and S3 Glacier Deep Archive are a secure, durable, and extremely low-cost Amazon S3 cloud storage classes for data archiving and long-term backup. They are designed to deliver 99.999999999% durability and provide comprehensive security and compliance capabilities that can help meet even the most stringent regulatory requirements. Customers can store data for as little as $1 per terabyte per month, a significant savings compared to on-premises solutions. To keep costs low yet suitable for varying retrieval needs, Amazon S3 Glacier provides three options for access to archives, from a few minutes to several hours, and S3 Glacier Deep Archive provides two access options ranging from 12 to 48 hours.

#### Benefits

* Retrievals as quick as 1-5 minutes
* Unmatched durability and scalability
* Comprehensive security and compliance capabilities
* Supported by partners, venders and AWS services
* Low cost

#### Use cases

* Media asset workflows
* Healthcare information archiving
* Regulatory and compliance archiving
* Scientific data storage
* Digital preservation
* Magnetic tape replacement

## Storage tiering

Storage tiering is an underlying principle of ILM (information lifecycle management). It is a storage networking method where data is stored on various types of media based on performance, availability and recovery requirements.

Storage tiering enables to allocate file storage space from different storage tiers according to created rules. Static storage tiering involves a manual one-time assignment of application files to a storage class, which is inflexible over a long term. Hierarchical Storage Management solutions typically require files to be migrated back into a file system name space before an application access request can be fulfilled, leading to latency and run-time overhead.

Storage tier allows organizations to

* Optimize storage assets by dynamically moving a file to its optimal storage tier as the value of the file changes over time
* Automate the movement of data between storage tiers without changing the way users or applications access the files
* Migrate data automatically based on policies set up by administrators, eliminating operational requirements for tiered storage and downtime commonly associated with data movement

#### Devices used for storage tiering

* SSD – Solid State Devices
* Hard Disks
* Tapes
* Cloud storage

# High Availability

## Introduction to high availability

High Availability (HA) describes systems that are dependable enough to operate continuously without failing. They are well-tested and sometimes equipped with redundant components.

“Availability” includes two periods of time: how much time a service is accessible, and how much time the system needs to respond to user requests. High availability refers to those systems that offer a high level of operational performance and quality over a relevant time period.

### Clustering

Clustering connects multiple, independent systems into a management framework for increased availability. Each system, or node, runs its own operating system and cooperates at the software level to form a cluster. Clustering solutions link commodity hardware with intelligent software to provide application failover and control. When a node or a monitored application fails, other nodes can take predefined actions to take over and bring up services elsewhere in the cluster.

Each node adds its processing power to the cluster as a whole and can increase overall throughput or performance. Each node runs its own operating system. A cluster interconnect enables cluster communications. A public network connects each node to a LAN for client access. Shared storage is accessible by each node that needs to run the application.

### Failover

In failover, failure of an application is detected by issuing specific commands, tests, or scripts to monitor the overall health of an application. Also, the health of underlying resources can be determined by supporting the applications such as file systems and network interfaces.

A redundant network heartbeat can be used to differentiate between the loss of a system and the loss of communication between systems. SCSI3-based membership coordination and data protection can be used for detecting failure on a node and on fencing.

Switchover and failover are the processes of bringing up application services on a different node in a cluster. The difference between the two processes is as follows:

* A switchover is an orderly shutdown of an application and its supporting resources on one server and a controlled startup on another server.
* A failover is like a switchover, except the ordered shutdown of applications on the original node may not be possible due to failure of hardware or services, so the services are started on another node.

#### Application availability

When an application or node failure is detected, application services are brought up on a different node in a cluster.

Diagram

Description automatically generated

IP addresses and system names are virtualized to ensure application availability. For example, in a two-node cluster consisting of db-server1 and db-server2, a virtual address may be called db-server. Clients access db-server and are unaware of which physical server hosts the db-server.

##### Procedures for application availability

* Start procedure

The application must have a command to start it and all resources it may require. Availability solutions bring up the required resources in a specific order, then brings up the application by using the defined start procedure.

For example, to start an Oracle database, Availability solution must know which Oracle utility to call, such as sqlplus. Availability solutions must also know the Oracle user, instance ID, Oracle home directory, and the pfile.

* Stop procedure

An individual instance of the application must be capable of being stopped without affecting other instances. For example, you cannot kill all httpd processes on a Web server because it also stops other Web servers. If availability solution cannot stop an application cleanly, it may call for a more forceful method, like a kill signal. After a forced stop, a clean-up procedure may be required for various process-specific and application-specific items that may be left behind. These items include shared memory segments or semaphores.

* Monitor procedure

The application must have a monitor procedure that determines if the specified application instance is healthy. The application must allow individual monitoring of unique instances. For example, the monitor procedure for a Web server connects to the specified server and verifies that it serves Web pages. In a database environment, the monitoring application can connect to the database server and perform SQL commands to verify read and write access to the database. If a test closely matches what a user does, it is more successful in discovering problems. Balance the level of monitoring by ensuring that the application is up and by minimizing monitor overhead.

##### Application restart

When you take an application offline, the application must close out all tasks, store data properly on shared disk, and exit. Stateful servers must not keep that state of clients in memory. States should be written to shared storage to ensure proper failover.

Commercial databases such as Oracle, Sybase, or SQL Server are good examples of well-written, crash-tolerant applications. On any client SQL request, the client is responsible for holding the request until it receives acknowledgement from the server. When the server receives a request, it is placed in a special redo log file. The database confirms that the data is saved before it sends an acknowledgement to the client. After a server crashes, the database recovers to the last-known committed state by mounting the data tables and by applying the redo logs. This returns the database to the time of the crash. The client resubmits any outstanding client requests that are unacknowledged by the server, and all others are contained in the redo logs.

If an application cannot recover gracefully after a server crashes, it cannot run in a cluster environment. The takeover server cannot start up because of data corruption and other problems.

### Parallel access

In parallel access, applications run simultaneously on more than one system in the cluster. Parallel access is more complex than failover. Parallel access is appropriate for applications that manage multiple application instances that run simultaneously without data corruption. In parallel access, standby node need not remain idle. It could be used to serve other applications in parallel. This utilizes more capacity of the available resources.

Multiple parallel applications that require flexible sharing of data in a data warehouse are currently deployed on separate clusters. Access across clusters is provided by NFS or other distributed file system technologies. Multiple parallel applications are deployed that require flexible sharing of data within a single cluster.

In a data center, multiple clusters exist with their dedicated fail over nodes.

There is a need to optimize the deployment of these disjoint clusters as a single large cluster.

Business critical applications require dedicated hardware to avoid the impact of configuration changes of one application on other applications. For example, when a node leaves or joins the cluster, it affects the cluster and the applications running on it. If multiple applications are configured on a large cluster, configuration changes have the potential to cause application downtime.

The application isolation feature provides logical isolation between applications at the disk group boundary. This is very helpful when applications require occasional sharing of data. Data can be copied efficiently between applications by using Veritas Volume Manager snapshots and disk group split, join, or move operations. Updates to data can be optimally shared by copying only the changed data. Thus, existing configurations that have multiple applications on a large cluster can be made more resilient and scalable with the application isolation feature.

# Disaster Recovery

Disaster recovery (DR) refers to the security planning area that aims to protect your organization from the negative effects of significant adverse events. It allows an organization to either maintain or quickly resume its mission-critical functions following a data disaster without incurring significant loses in business operations or revenues.

Disasters come in different shapes and sizes. They do not only refer to catastrophic events such as earthquakes, tornadoes, or hurricanes, but also security incidents such as equipment failures, cyber-attacks, or even terrorism classified as disasters.

In preparation, organizations and companies create DR plans detailing processes to follow and actions to take to resume their mission-critical functions.

### What is Disaster Recovery?

Disaster recovery focuses on IT systems that help support an organization’s critical business functions. It is often associated with the term business continuity, but the two are not entirely interchangeable. DR is part of business continuity. It focuses more on keeping all business aspects running despite disasters. Since IT systems have become critical to business success, disaster recovery is now a primary pillar within the business continuity process.

Most business owners do not usually consider that they may be victims of a natural disaster until an unforeseen crisis happens, which ends up costing their company a lot of money in operational and economic losses. These events can be unpredictable, and as a business owner, you cannot risk not having a disaster preparedness plan in place.

### What Kind of Disasters Do Businesses Face?

Business disasters can either be technological, natural or human-made. Examples of natural disasters include floods, tornadoes, hurricanes, landslides, earthquakes, and tsunamis. Whereas, human-made and technological disasters involve things like hazardous material spills, power or infrastructural failure, chemical and biological weapon threats, nuclear power plant blasts or meltdowns, cyberattacks, acts of terrorism, explosions, and civil unrest.

Potential disasters to plan for include:

* Application failure
* VM failure
* Host failure
* Rack failure
* Communication failure
* Data center disaster
* Building or campus disaster
* Citywide, regional, national, and multinational disasters

## Need of disaster recovery (DR)

Regardless of size or industry, when unforeseen events take place, causing daily operations to come to a halt, your company needs to recover quickly to ensure that you continue providing your services to customers and clients.

Downtime is perhaps among the biggest IT expenses that a business faces. Based on 2014-2015 disaster recovery statistics from Infrascale, one hour of downtime can cost small businesses as much as $8,000, mid-size companies $74,000, and large organizations $700,000.

For small and mid-sized businesses (SMBs), extended loss of productivity can lead to the reduction of cash flow through lost orders, late invoicing, missed delivery dates, and increased labor costs due to extra hours resulting from downtime recovery efforts.

If businesses do not anticipate major disruptions and address them appropriately, they risk incurring long-term negative consequences and implications as a result of the occurrence of unexpected disasters.

Having a DR plan in place can save companies from multiple risks, including:

* Reputation loss
* Out of budget expenses
* Data loss
* Negative impact on your clients and customers

As businesses become more reliant on high availability, their tolerance for downtime has decreased. Therefore, many have a DR in place to prevent adverse disaster effects from affecting their daily operations.

The two critical measurements in DR and downtime are:

* Recovery Point Objective (RPO):

It refers to the maximum age of files that your organization must recover from its backup storage to ensure its normal operations resume after a disaster. It determines the minimum backup frequency. For instance, if organization has a four-hour RPO, its system must back up every four hours.

* Recovery Time Objective (RTO):

It refers to the maximum amount of time your organization requires to recover its files from backup and resume normal operations after a disaster. Therefore, RTO is the maximum downtime amount that organization can handle. If the RTO is two hours, then operations can’t be down for a period longer than that.

Once RPO and RTO is identified, administrators can use the two measures to choose optimal disaster recovery strategies, procedures, and technologies.

To recover operations during tighter RTO windows, organization needs to position its secondary data optimally to make it easily and quickly accessible. One method used to restore data quickly is recovery-in-place because it moves all backup data files to a live state, which eliminates the need to move it across a network. It can protect against server and storage system failure.

Before using recovery-in-place, organizations need to consider three things:

* Its disk backup appliance performance
* The time required to move all data from its backup state to a live one
* Failback

Also, since recovery-in-place can sometimes take up to 15 minutes, replication may be necessary if a quicker recovery time is needed. Replication refers to the periodic electronic refreshing or copying of a database from computer server A to server B, which ensures that all users in the network always share the same information level.

### Disaster Recovery Plan (DRP)

A disaster recovery plan refers to a structured, documented approach with instructions put in place to respond to unplanned incidents. It’s a step-by-step plan that consists of the precautions put in place to minimize a disaster’s effects so that organization can quickly resume its mission-critical functions or continue to operate as usual.

Typically, DRP involves an in-depth analysis of all business processes and continuity needs. What’s more, before generating a detailed plan, organization should perform a risk analysis (RA) and a business impact analysis (BIA). It should also establish its RTO and RPO.

1. Recovery Strategies

A recovery strategy should begin at the business level, which allows to determine the most critical applications to run organization. Recovery strategies define organization’s plans for responding to incidents, while DRPs describe in detail how it should respond.

When determining a recovery strategy, it should consider issues such as:

* Budget
* Resources available such as people and physical facilities
* Management’s position on risk
* Technology
* Data
* Suppliers
* Third-party vendors

Management must approve all recovery strategies, which should align with organizational objectives and goals. Once the recovery strategies are developed and approved, one can then translate them into DRPs.

1. Disaster Recovery Planning Steps

The DRP process involves a lot more than only writing the document. A business impact analysis (BIA) and risk analysis (RA) help determine areas to focus resources in the DRP process.

The BIA is useful in identifying the impacts of disruptive events, which makes it the starting point for risk identification within the DR context. It also helps generate the RTO and RPO.

The risk analysis identifies vulnerabilities and threats that could disrupt the normal operations of processes and systems highlighted in the BIA. The RA also assesses the likelihood of the occurrence of a disruptive event and helps outline its potential severity.

A DR plan checklist has the following steps:

* Establishing the activity scope
* Gathering the relevant network infrastructure documents
* Identifying severe threats and vulnerabilities as well as the organization’s critical assets
* Reviewing the organization’s history of unplanned incidents and their handling
* Identifying the current DR strategies
* Identifying the emergency response team
* Having the management review and approve the DRP
* Testing the plan
* Updating the plan
* Implementing a DR plan audit

1. Creating a DRP

An organization can start its DRP with a summary of all the vital action steps required and a list of essential contacts, which ensures that crucial information is easily and quickly accessible.

The plan should also define the roles and responsibilities of team members while also outlining the criteria to launch the action plan. It must then specify, in detail, the response and recovery activities. The other essential elements of a DRP template include:

* Statement of intent
* The DR policy statement
* Plan goals
* Authentication tools such as passwords
* Geographical risks and factors
* Tips for dealing with the media
* Legal and financial information
* Plan history

1. DRP Scope and Objectives

A DRP can range in scope (i.e., from basic to comprehensive). Some can be upward of 100 pages.

DR budgets can vary significantly and fluctuate over time. Therefore, organization can take advantage of any free resources available such as online DR plan templates from the Federal Emergency Management Agency. There is also a lot of free information and how-to articles online.

A DRP checklist of goals includes:

* Identifying critical IT networks and systems
* Prioritizing the RTO
* Outlining the steps required to restart, reconfigure, or recover systems and networks

The plan should, at the very least, minimize any adverse effects on daily business operations. Employees should also know the necessary emergency steps to follow in the event of unforeseen incidents.

Distance, though important, is often overlooked during the DRP process. A DR site located close to the primary data center is ideal in terms of convenience, cost, testing, and bandwidth. However, since outages differ in scope, a severe regional event may destroy both the primary data center and its DR site when the two are located close together.

1. Types of Disaster Recovery Plans

DRP can be tailored for a given environment.

* Virtualized DRP:

Virtualization allows to implement DR using an efficient and straightforward way. Using a virtualized environment, you can create new virtual machines (VMs) instances immediately and provide high availability application recovery. What’s more, it makes testing easier to achieve. Your plan must include validation ability to ensure that applications can run faster in DR mode and return to normal operations within the RTO and RPO.

* Network DRP:

Coming up with a plan to recover a network gets complicated with the increase in network complexity. Ergo, it is essential to detail the recovery procedure step-by-step, test it correctly, and keep it updated. Under a network DRP, data is specific to the network; for instance, in its performance and networking staff.

* Cloud DRP:

A cloud-based DR can range from file backup to a complete replication process. Cloud DRP is time-, space-, and cost-efficient; however, maintaining it requires skill and proper management. IT manager must know the location of both the physical and virtual servers. Also, the plan must address security issues related to the cloud.

* Data Center DRP:

This plan focuses on data center facility and its infrastructure. One key element in this DRP is an operational risk assessment since it analyzes the key components required, such as building location, security, office space, and power systems and protection. It must also address a broader range of possible scenarios.

## Building blocks of disaster recovery

### Global cluster

Local clustering provides local failover for each site or building. But these configurations do not provide protection against large-scale disasters such as major floods, hurricanes, and earthquakes that cause outages for an entire city or region. The entire cluster could be affected by an outage. In such situations, global clusters can ensure availability by migrating applications to remote clusters located considerable distances apart.

Diagram

Description automatically generated

Let us take the example of an Oracle database configured in a global cluster. Oracle is installed and configured in both clusters, local and global. Oracle data is located on shared disks within each cluster and is replicated across clusters to ensure data concurrency. The Oracle application is online on a system in cluster A and is configured to fail over globally, on clusters A and B.

Clustering solution continuously monitors and communicates events between clusters. Inter-cluster communication ensures that the global cluster is always aware of the state of global application service.

In the event of a system or application failure, VCS fails over the Oracle service group to another system in the same cluster. If the entire cluster fails, VCS fails over the service group to the remote cluster, which is part of the global cluster. VCS also redirects clients once the application is online on the new location.

### Wide-area-connector (WAC)

The wide-area connector (WAC) is a failover Application resource that ensures communication between clusters.

Diagram

Description automatically generated

The WAC process runs on one system in each cluster and connects with peers in remote clusters. It receives and transmits information about the status of the cluster, service groups, and systems. This communication enables clustering solution to create a consolidated view of the status of all the clusters configured as part of the global cluster. The process also manages wide-area heart beating to determine the health of remote clusters. The process also transmits commands between clusters and returns the result to the originating cluster.

Clustering solutions provide the option of securing the communication between the wide-area connectors.

### Heartbeat

Heartbeat communication is used to transmit the cluster status and configuration information between nodes. This allows the managing software of the cluster to send and receive these updates among all cluster nodes. This communication is vital to the well-being of the cluster. If an active cluster node is unable to communicate with the other passive nodes, this may cause minor to severe issues with cluster behavior and performance. This may impede access to clustered applications.

Private network

Links are used for heartbeat communication, as well as sharing status and update information between clustered nodes. These private links are used for cluster communication only. No other application or resource should be configured to share these links. This may interfere with cluster communication and cause issues with cluster functionality.

#### Configuration

The following recommended configuration are used to set up heartbeat communication. This ensures that there is no single point of failure, and that the private communication will function optimally.

* Do not assign IP addresses to private heartbeats network adapters.
* Ensure that all adapters, and switch ports, are set to auto-negotiate, or are set to use matching speed settings.
* All private heartbeats should be connected via a network hub or a switch. The best practice is to use only one hub, or switch, per heartbeat connection. This avoids a single point of failure.
* Although cross-linked cables can be used for heartbeat purposes, we recommend that hubs, or switches, be used to ensure reliable connection.
* Ensure that the same driver revision is used across all NICs (Network Interface Cards).
* Ensure that the same make and model of NICs are used throughout the network, to promote consistency.
* The public network connection should also have its own NIC and hub, or switch. We recommend using the same NIC make and model, and the same NIC driver be used for the public network, that are used for the private heartbeat connections.
* If using three, or more, network adapters, set at least two of them as private heartbeat adapters in any manual configuration, or when using the Veritas Cluster Configuration Wizard. If required, set the third as a low priority (lowpri) heartbeat adapter, which uses the public network if both private heartbeats fail. This maintains cluster communication over the public network, as a low priority, until the private heartbeat links can be restored.

#### Disaster Recovery-as-a-Service (DRaaS)

Disaster recovery-as-a-service is a cloud-based DR method that has gained popularity over the years. This is because DRaaS lowers cost, it is easier to deploy, and allows regular testing.

Cloud testing saves company money because they run on shared infrastructure. They are also quite flexible, allowing to sign up for only the needed services, and DR tests can be completed by only spinning up temporary instances.

DRaaS expectations and requirements are documented and contained in a service-level agreement (SLA). The third-party vendor then provides failover to their cloud computing environment, either on a pay-per-use basis or through a contract.

However, cloud-based DR may not be available after large-scale disasters since the DR site may not have enough room to run every user’s applications. Also, since cloud DR increases bandwidth needs, the addition of complex systems could degrade the entire network’s performance.

Perhaps the biggest disadvantage of the cloud DR is that organization have little control over the process; thus, service provider needs to be trusted to implement the DRP in the event of an incident while meeting the defined recovery point and recovery time objectives.

Costs vary widely among vendors and can add up quickly if the vendor charges based on storage consumption or network bandwidth. Therefore, before selecting a provider, you need to conduct a thorough internal assessment to determine DR needs.

#### Disaster Recovery Sites

A DR site allows to recover and restore technology infrastructure and operations when primary data center is unavailable. These sites can be internal or external. These sites are necessary for companies with aggressive RTOs and large information requirements. Some considerations to make when building your internal recovery site are hardware configuration, power maintenance, support equipment, layout design, heating and cooling, location, and staff. Though much more expensive compared to an external site, an internal DR site allows to control all aspects of the DR process.

External sites are owned and operated by third-party vendors. They can either be:

* Hot: It’s a fully functional data center complete with hardware and software, round the clock staff, as well as personnel and customer data.
* Warm: It’s an equipped data center with no customer data. Clients can install additional equipment or introduce customer data.
* Cold: It has the infrastructure in place to support data and IT systems. It, however, has no technology until client organizations activate DR plans and install equipment. It, sometimes, supplements warm and hot sites during long-term disasters.

## Split-brain problem and solutions

Under normal conditions, when a cluster system ceases heartbeat communication with its peers due to an event such as power loss or a system crash, the peers assume the system has failed and issue a new, "regular" membership excluding the departed system. A designated system in the cluster then takes over the service groups running on the departed system, ensuring the application remains highly available. However, heartbeats can also fail due to network failures. If all network connections between any two groups of systems fail simultaneously, a network partition occurs. When this happens, systems on both sides of the partition can restart applications from the other side resulting in duplicate services, or "split-brain". A split brain occurs when two independent systems configured in a cluster assume, they have exclusive access to a given resource (usually a file system or volume). The most serious problem caused by a network partition is that it affects the data on shared disks. All failover management software uses a predefined method to determine if its peer is "alive". If the peer is alive, the system recognizes it cannot safely take over resources. Split brain occurs when the method of determining peer failure is compromised. In virtually all failover management software (FMS) systems, split-brain situations are rare. A true split brain means multiple systems are online and have accessed an exclusive resource simultaneously.

A split-brain means cluster membership was affected in such a way that multiple systems use the same exclusive resources, usually resulting in data corruption. The goal is to minimize the chance of a system taking over an exclusive resource while another has it active, yet accommodate a system powering off. In other words, a way to discriminate between a system that has failed and one that is simply not communicating.

### Split-Brain Prevention

Cluster software provides several functions aimed at the prevention of split-brain situations. The following list contains a brief explanation of each prevention method.

* Private Heartbeat

It is recommended to have a minimum of two dedicated 100-megabit private links between cluster nodes. These must be completely isolated from each other so the failure of one heartbeat link cannot possibly affect the other.

Configuring private heartbeats to share any infrastructure is not recommended. Configurations such as running two shared heartbeats to the same hub or switch or using a single virtual local area network (VLAN) to trunk between two switches induce a single point of failure in the heartbeat architecture. The simplest guideline is "No single failure, such as power, network equipment or cabling can disable both heartbeat connections."

* Low-Priority Heartbeat

Heartbeat over public network does minimum traffic over the network until you get down to one normal heartbeat remaining. Then it becomes a full functional heartbeat.

Use of a low priority link is also recommended to provide further redundancy. The low priority link prevents a jeopardy condition on loss of any single private link and provides additional redundancy (consider low-priority heartbeat along with two private network heartbeats).

* Disk Heartbeat

With disk heartbeat configured, each system in the cluster periodically writes to and reads from specific regions on a dedicated shared disk. This exchange consists of heartbeat only and does not include communication about cluster status.

With disk heartbeat configured in addition to the private network connections, clustering software has multiple heartbeat paths available. For example, if one of two private network connections fail, clustering software has the remaining network connection and the disk heartbeat region that allow heartbeats to continue normally.

* SCSI II Disk Reservations

Reserves and monitors SCSI disks for a system, enabling a resource to go online on that system, when using the Disk Reservation agent. The agent supports all SCSI II disks. Use this agent to specify a list of raw disk devices, and reserve all or a percentage of accessible disks for an application. The reservation prevents disk data corruption by restricting other systems from accessing and writing to the disks. An automatic probing feature allows systems to maintain reservations even when the disks or bus are reset.

* IP Checking

This method is used in either the pre online-ipc event trigger, or simply make an IP resource the first resource to online in the service group. Both methods check to make sure the IP addresses for this service group are not being used by another system before onlining the service group.

* I/O Fencing SCSI III Reservations

Clustering software can have parallel or failover service groups with disk group resources in them. If the cluster has a split-brain, I/O fencing should force one of the sub clusters to commit suicide in order to prevent data corruption. The sub cluster which commits suicide should never gain access to the disk groups without joining the cluster again. In parallel access clusters, it is necessary to prevent any active processes from writing to the disks. In failover cluster, however, access to the disk only needs to be prevented when clustering software fails over the service group to another node. Some multipathing products will be supported with I/O Fencing.

### Preparing for DR–firedrill

A DR fire drill is a zero-downtime test that mimics the configuration, application data, and fail over behavior of critical service groups. A successful DR fire drill indicates that it is highly likely for a critical service group to fail over as intended or as configured on to a remote cluster, when it is needed.

The disaster recovery fire drill procedure tests the fault-readiness of a configuration by mimicking a failover from the primary site to the secondary site. This procedure is done without stopping the application at the primary site and disrupting user access, interrupting the flow of replicated data, or causing the secondary site to need resynchronization.

The initial steps to create a fire drill service group on the secondary site that closely follows the configuration of the original application service group and contains a point-in-time copy of the production data in the Replicated Volume Group (RVG). Bringing the fire drill service group online on the secondary site demonstrates the ability of the application service group to fail over and come online at the secondary site, should the need arise. Fire drill service groups do not interact with outside clients or with other instances of resources, so they can safely come online even when the application service group is online.

The DR fire drill feature lets organization do the following:

* Verify that replication for an application works correctly.
* Verify that a DR service group can be brought online successfully.

Testing substantiates all DRPs. It identifies deficiencies in the plan and provides opportunities to fix any problems before a disaster occurs. Testing can also offer proof of the plan’s effectiveness and hits RPOs.

IT technologies and systems are continually changing. Therefore, testing ensures that DRP is up to date.

Some reasons for not testing DRPs include budget restrictions, lack of management approval, or resource constraints. DR testing also takes time, planning, and resources. It can also be an incident risk if it involves the use of live data. However, testing is an essential part of DR planning that it should never be ignored.

DR testing ranges from simple to complex:

* A plan review involves a detailed discussion of the DRP and looks for any missing elements and inconsistencies.
* A tabletop test sees participants walk through the plan’s activities step by step. It demonstrates whether DR team members know their duties during an emergency.
* A simulation test is a full-scale test that uses resources such as backup systems and recovery sites without an actual failover.
* Running in disaster mode for a period is another method of testing your systems. For instance, you could failover to your recovery site and let your systems run from there for a week before falling back.

Organization should schedule testing in its DR policy; however, be wary of its intrusiveness. This is because testing too frequently is counter-productive and draining on your personnel. On the other hand, testing less regularly is also risky. Additionally, always test DR plan after making any significant system changes.

To get the most out of testing:

* Secure management approval and funding
* Provide detailed test information to all parties concerned
* Ensure that the test team is available during the test date
* Schedule your test correctly to ensure that it doesn’t conflict with other activities or tests
* Confirm that test scripts are correct
* Verify that your test environment is ready
* Schedule a dry run first
* Be prepared to stop the test if needed
* Have a scribe take notes
* Complete an after-action report detailing what worked and what failed
* Use the results gathered to update your DR plan

### References

* Arpaci-Dusseau, Remzi H.; Arpaci-Dusseau, Andrea C. (2014). "Operating Systems: Three Easy Pieces, Chapter: Hard Disk Drives" (PDF). Arpaci-Dusseau Books. Archived (PDF) from the original on February 16, 2015. Retrieved March 7, 2014
  + <http://pages.cs.wisc.edu/~remzi/OSTEP/file-disks.pdf>
* <https://cloud.google.com/storage/docs/key-terms>
* <https://www.red-gate.com/simple-talk/sql/database-administration/storage-101-data-center-storage-configurations/>
* <https://www.vxchnge.com/blog/data-center-networking-101-everything-to-know>
* <https://www.redhat.com/en/topics/virtualization/what-is-a-virtual-machine>
* <https://www.datacore.com/storage-virtualization/>
* <https://www.enterprisestorageforum.com/storage-management/raid-levels.html>
* <https://searchservervirtualization.techtarget.com/definition/storage-pools>
* <https://publib.boulder.ibm.com/tividd/td/TSMCW/GC32-0782-00/en_US/HTML/anrwgd91.htm>
* <https://www.veritas.com/information-center/enterprise-cloud-storage-ultimate-guide>
* <https://aws.amazon.com/s3/>
* <https://aws.amazon.com/glacier/>
* <https://avinetworks.com/glossary/high-availability/>
* <https://www.veritas.com/information-center/disaster-recovery-guide>