

1. Key characteristics of Data center elements.

- * **Availability** : Data center should ensure the availability of information whenever required. Unavailability of information may cost millions of dollars per hour, eg: financial services e-commerce
- * **Security** : Data center should establish policies, procedures and core element integration to prevent unauthorized access to information.
- * **Performance** : All elements of data center should provide optimal performance based on required service levels.
- * **Scalability** : Data center resource should scale based on the requirement, ~~also~~ without interrupting business operation.
- * **Capacity** : Data center operations require adequate resources to store & process large amount of data efficiently. When ~~data~~ capacity requirement increases the data center should provide additional capacity without interrupting availability.
- * **Manageability** : Data center should provide easy and integrated management of its elements. This can be achieved with minimum intervention of human in common tasks.
- * **Data Integrity** : This refers to mechanisms like error correction codes or parity bits which ensure that the data is stored and retrieved exactly as it was received.

2. The evolution of ~~dynamic~~ logical volume managers enabled dynamic extension of the file system capacity and efficient storage management. The LVM software which runs on the compute system & manages logical and physical storage. LVM is an integrated layer between the file system and the physical disk. It can partition a large-~~a~~ capacity disk into virtual, smaller capacity volumes or aggregate several smaller disk to form a larger virtual volume.

1 of 4
Vlll

Disk partition is used to improve flexibility and utilization of disk drives. The disk drive is divided into logical containers during partitioning which is called as logical volumes. The partitions are created from groups of contiguous cylinders when the hard disk is initially set up on the host. The host's file system access the logical volumes without any knowledge of partitioning and physical structure of the disk. The LVM converts the physical storage provided by the physical volumes to a logical view of storage which is used by the OS & applications. A volume group is created by grouping together one or more physical volumes. A unique physical volume is assigned to each physical volumes when it is initialized for use by LVM.

Each physical volumes is partitioned into equal size data blocks which is called as physical extents when volume group is created. Logical volumes is created within a given volume group which can be thought of as a disk partition whereas the volume group can be thought of as a disk.

3. Raid implementation methods are software and hardware raid.

→ Software RAID: uses host based software to provide RAID functions. It is implemented at the OS levels and don't use a dedicated hardware controller to manage the RAID array. Software RAID provides cost & simplicity benefits over hardware RAID.

Limitations:

- Performance: Software raid affects overall performance due to additional CPU cycles required to perform RAID calculation.
- Supported features: Software raid doesn't support all RAID levels.

Operating system compatibility: Software raid is tied to the host OS hence upgrades the software RAID. This leads to inflexibility in the data processing environment.

Hardware RAID: This has a specialized hardware controller which is implemented on host or the array.

Host based hardware RAID - controller card RAID which a specialized RAID controller is installed in the host & disk drives are connected to it. This is not an efficient solution in data center environment which has a large no of hosts.

Array based hardware RAID is internal RAID controller which acts as an interface between the host and the disk. It presents storage volumes to the host & the host manages these volumes as physical drives.

Function of RAID controller.

- Management & control of Disk aggregations.
- Data regeneration in event of disk failure
- Translation of I/O request between physical & logical disk..

4. Flushing is the process which commits data from cache to the disk.

→ Idle flushing: which occurs continuously at a modest rate when the cache utilization level is between the high & low watermark.

→ High watermark flushing: Activated when cache utilization hits the high watermark. The storage system dedicates ~~some~~ ~~resources~~ on I/O processing some additional resources for flushing & have impact on I/O processing.

→ Forced flushing: This occurs in the event of a large I/O burst when cache reaches 100 percent of its capacity which significantly affects the I/O response time. This flushes the cache on priority by allocating more resources.

cache management algorithms.

- Least Recently used : This continuously monitors the data access in cache & identifies the cache page that have not been accessed for long time. LRU either forces up these pages or mark them for reuse. It is based on the assumption that data that has not been accessed for a while will not be requested by the host. If the page contains write data that has not yet been committed to disk. The data is first written to disk before page is reused.
- Most recently used : It is opposite of LRU, where the pages that have been accessed most recently are forced up or marked for reuse. It is based on the assumption that recently accessed data may not be required for a while.