IT8001
Information Storage and Management

Professional Elective

Topic

Intelligent Storage System

Varun C.M.

Assistant Professor, Department of Information Technology,
St. Xavier's Catholic College of Engineering

Expected Outcome

CO1

Understand the logical and physical components of a Storage infrastructure.

Intelligent Storage System

These intelligent storage systems are feature-rich RAID arrays that provide highly optimized I/O processing capabilities.

Components of Intelligent Storage System

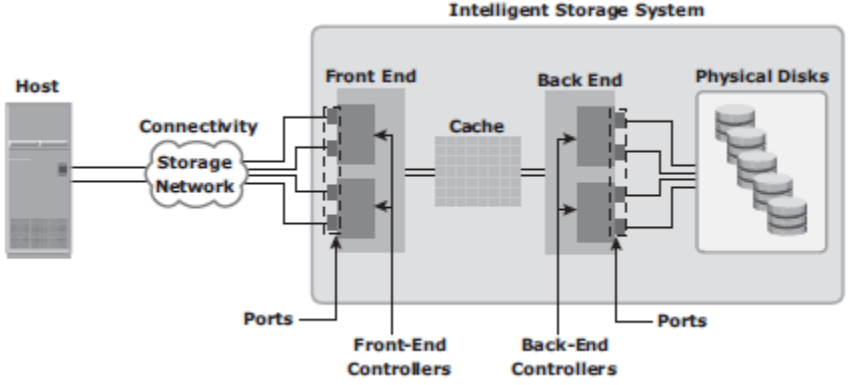


Image courtesy EMC Corporation, "Information Storage and Management"

1. Front End

The front end provides the interface between the storage system and the host. It consists of two components: front-end ports and front-end controllers.

Front end has redundant controllers for high availability, and each controller contains multiple ports that enable large numbers of hosts to connect to the intelligent storage system.

2. Cache

Cache is semiconductor memory where data is placed temporarily to reduce the time required to service I/O requests from the host.

Cache improves storage system performance by isolating hosts from the mechanical delays associated with rotating disks or hard disk drives (HDD).

Accessing data from cache is fast and typically takes less than a millisecond. On intelligent arrays, write data is fi rst placed in cache and then written to disk.

3. Back End

The back end provides an interface between cache and the physical disks. It consists of two components: back-end ports and back-end controllers.
The back-end controls data transfers between cache and the physical disks.
The back-end controller communicates with the disks when performing reads and writes and also provides additional, but limited, temporary data storage.
For high data protection and high availability, storage systems are configured with dual controllers with multiple ports. Such configurations provide an alternative path to physical disks if a controller or port failure occurs.

4. Physical Disk

Physical disks are connected to the back-end storage controller and provide persistent data storage.

Modern intelligent storage systems provide support to a variety of disk drives with different speeds and types, such as FC, SATA, SAS, and flash drives.

They also support the use of a mix of flash, FC, or SATA within the same array.

Storage Provisioning

Storage provisioning is the process of assigning storage resources to hosts based on capacity, availability, and performance requirements of applications running on the hosts.

Storage Provisioning

In ***traditional storage provisioning***, physical disks are logically grouped together and a required RAID level is applied to form a set, called a RAID set.

Virtual provisioning enables creating and presenting a LUN with more capacity than is physically allocated to it on the storage array.

Traditional Storage Provisioning

RAID sets usually have a large capacity because they combine the total capacity of individual drives in the set.

Logical units are created from the RAID sets by partitioning the available capacity into smaller units. These units are then assigned to the host based on their storage requirements.

Logical units are spread across all the physical disks that belong to that set. Each logical unit created from the RAID set is assigned a unique ID, called a ***logical unit number (LUN)***.

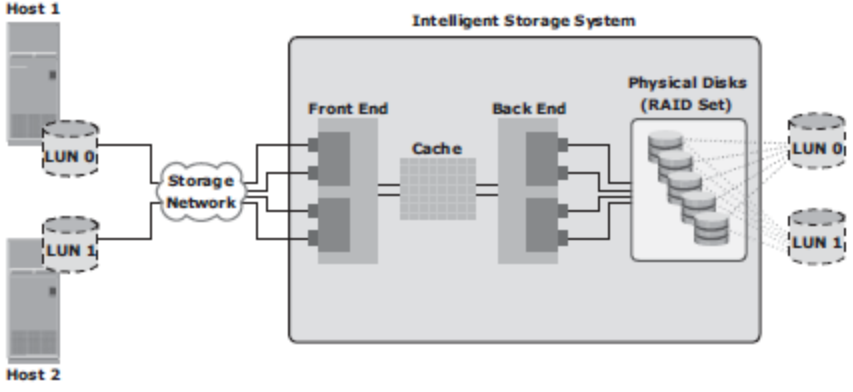


Image courtesy EMC Corporation, "Information Storage and Management"

Virtual Storage Provisioning

The LUN created using virtual provisioning is called a ***thin LUN*** to distinguish it from the traditional LUN.

Thin LUNs do not require physical storage to be completely allocated to them at the time they are created and presented to a host.

Physical storage is allocated to the host "on-demand" from a shared pool of physical capacity.

A shared pool consists of physical disks. A shared pool in virtual provisioning is analogous to a RAID group, which is a collection of drives on which LUNs are created.

Virtual provisioning enables more efficient allocation of storage to hosts. Virtual provisioning also enables oversubscription, where more capacity is presented to the hosts than is actually available on the storage array.

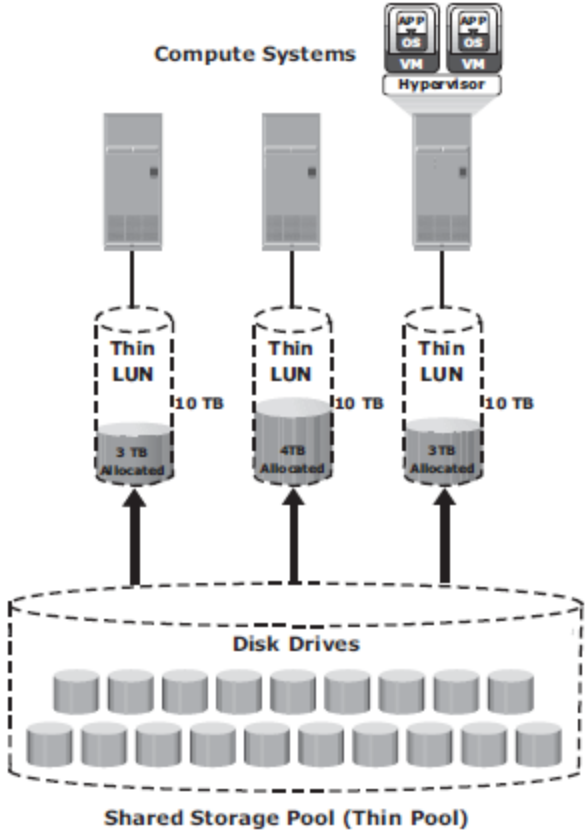


Image courtesy EMC Corporation, "Information Storage and Management"

Types of Intelligent Storage System

High-end storage systems, referred to as active-active arrays, are generally aimed at large enterprise applications. These systems are designed with a large number of controllers and cache memory.

Midrange storage systems are also referred to as active-passive arrays and are best suited for small and medium sized enterprise applications. They also provide optimal storage solutions at a lower cost.

High End Storage Systems

High End Storage Systems provide

- Large storage capacity
- Large amounts of cache to service host I/Os optimally
- Fault tolerance architecture to improve data availability
- Connectivity to mainframe computers and open systems hosts
- Availability of multiple front-end ports and interface protocols to serve a large number of hosts
- Availability of multiple back-end controllers to manage disk processing
- Scalability to support increased connectivity, performance, and storage capacity requirements
- Ability to handle large amounts of concurrent I/Os from a number of hosts and applications
- Support for array-based local and remote data replication

Mid Range Storage Systems

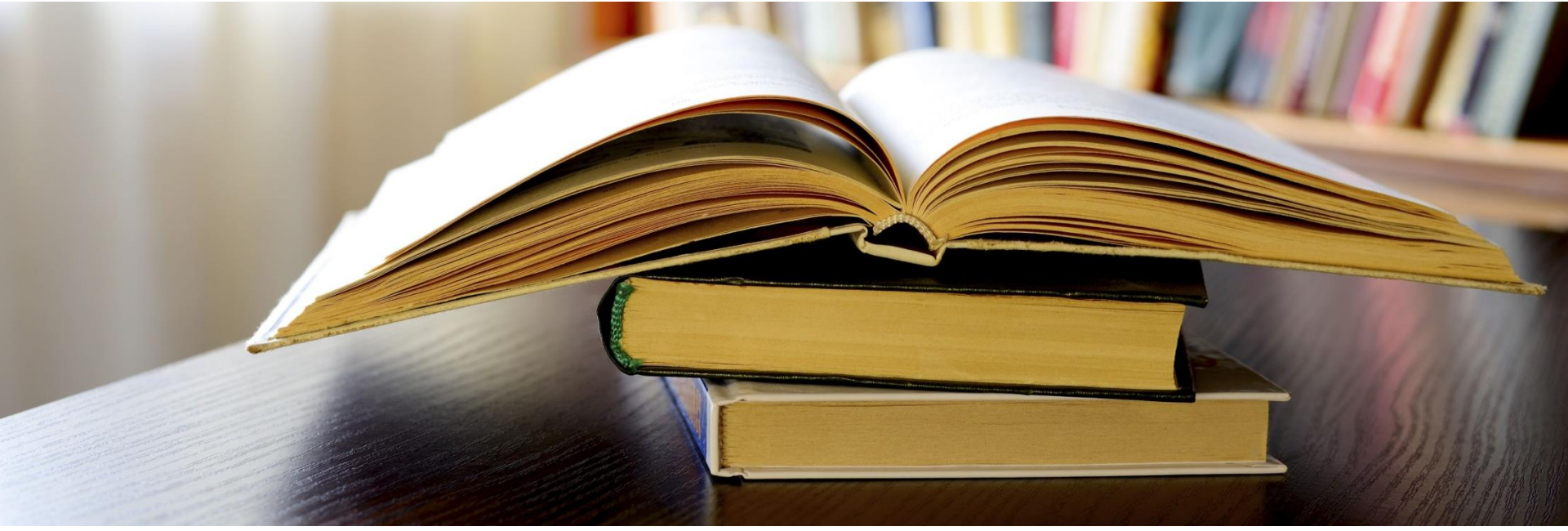
In an active-passive array, a host can perform I/Os to a LUN only through the controller that owns the LUN.

The host can perform reads or writes to the LUN only through the path to controller A because controller A is the owner of that LUN. The path to controller B remains passive and no I/O activity is performed through this path.

Midrange storage systems are typically designed with two controllers, each of which contains host interfaces, cache, RAID controllers, and interface to disk drives.

Midrange arrays are designed to meet the requirements of small and medium enterprise applications; therefore, they host less storage capacity and cache than high-end storage arrays. There are also fewer front-end ports for connection to hosts.

However, they ensure high redundancy and high performance for applications with predictable workloads. They also support array-based local and remote replication.



References

EMC Corporation, "Information Storage and Management" , Wiley, India.
Robert Spalding, "Storage Networks: The Complete Reference", Tata McGraw Hill, Osborne, 2003.

Varun C.M.
Assistant Professor, Department of Information Technology, St. Xavier's Catholic College of Engineering