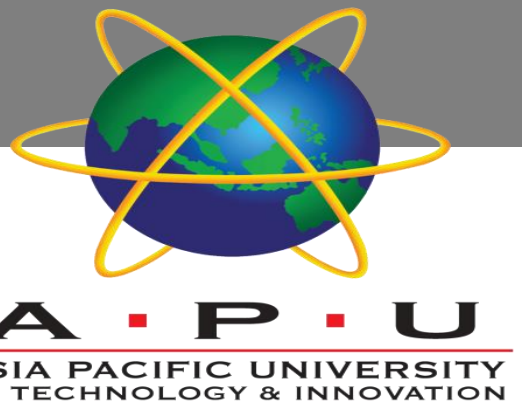
Data Protection and Management



Data Protection Architecture

Learning Objectives

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

- Describe the building blocks of a data protection architecture
- Describe data source components – application, hypervisor, and primary storage
- Describe data protection applications and storage

Data Protection Architecture

A data protection architecture is a blueprint that specifies the protection components and their interrelationships and guides an organization to provide centralized data protection services.

It enables cost-optimized and consolidated data protection, simplifies data protection management, and helps organizations to meet service level requirements.

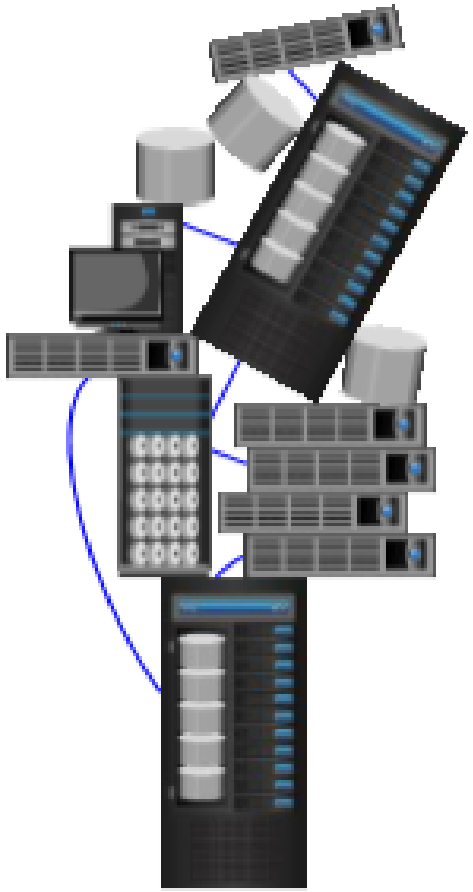
An intentional data protection architecture is explicitly identified and then implemented.

Accidental Architecture

An accidental architecture consists of a fragmented set of data protection processes, multiple unconnected data protection tools, and infrastructure silos. Multiple entities within an organization perform their own data protection operations without a clear picture of the ownership of protection processes and resources.

Why Accidental Architecture Happen?

Data protection without an intentional architecture results in an accidental architecture

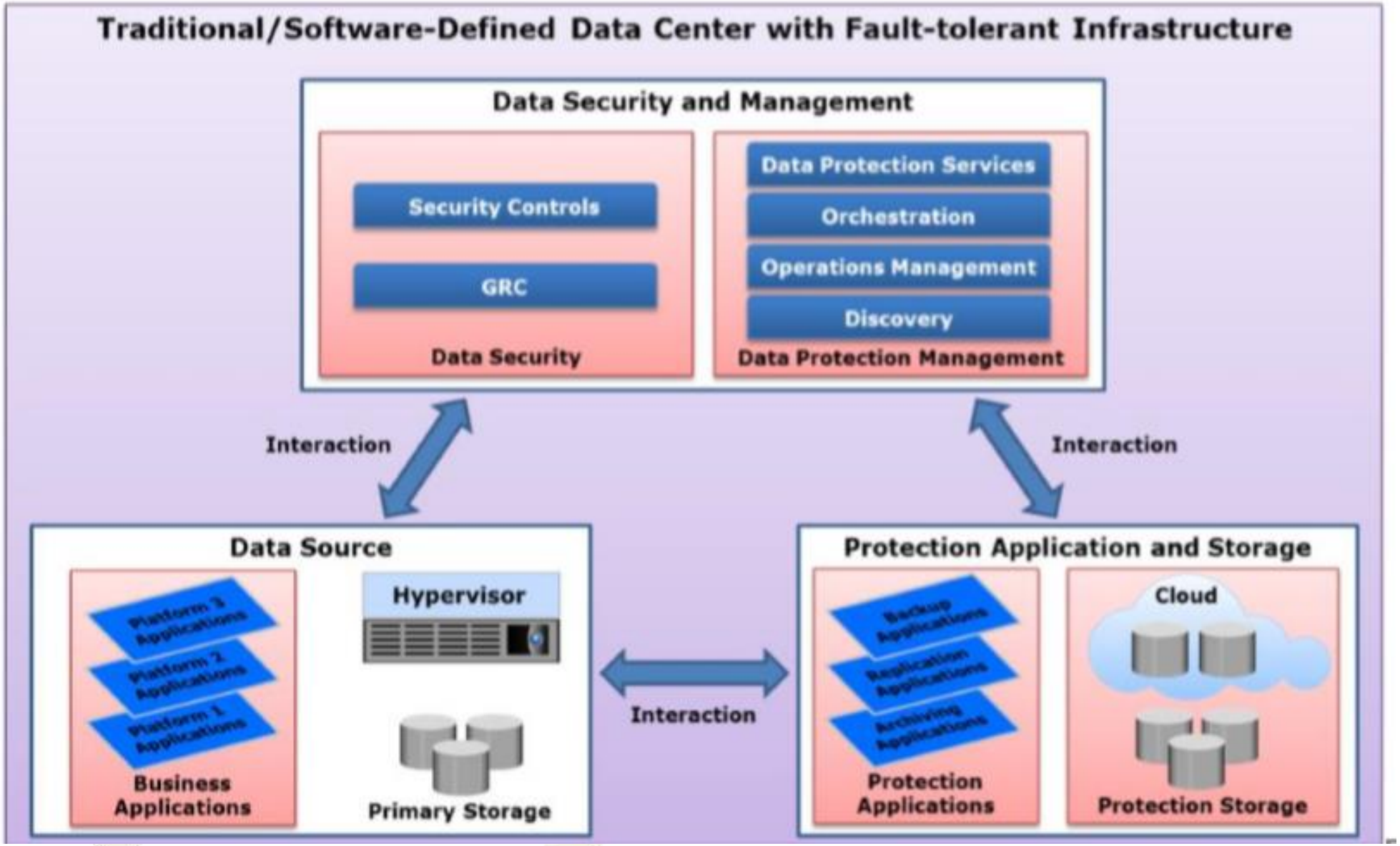


- Ad hoc and arbitrarily implemented solutions
- Unclear ownership of processes and resources
- Multiple unconnected tools and no central visibility
- Complexity in scaling resources
- Difficulty in meeting SLAs
- Expenditure increases manifold with data growth

Quick Review

- Describe using your own word the term “Accidental Architecture” in data protection strategy
- How to avoid Accidental Architecture

Data Protection Architecture



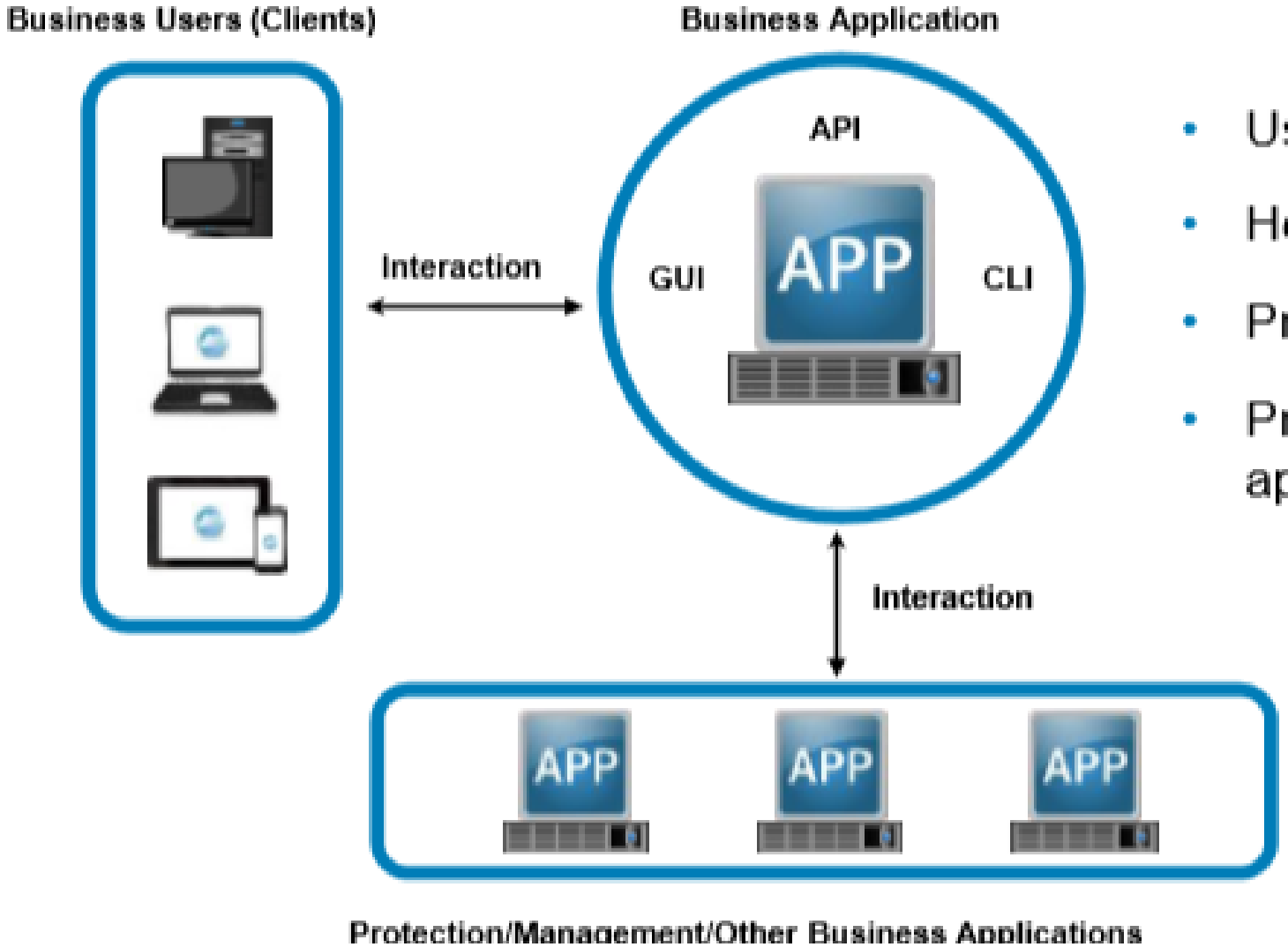
Data Source

It is the source of the data that must be protected.

The data source can be:

- a business application
- a hypervisor
- a primary storage.

Data Source – Business Application



- Used to perform business operations
- Helps in increasing productivity
- Provides user interfaces – CLI, GUI
- Provides API for application-to-application interaction

Third Platform Applications

Third platform applications are based on cloud, Big Data analytics, mobile, and social technologies

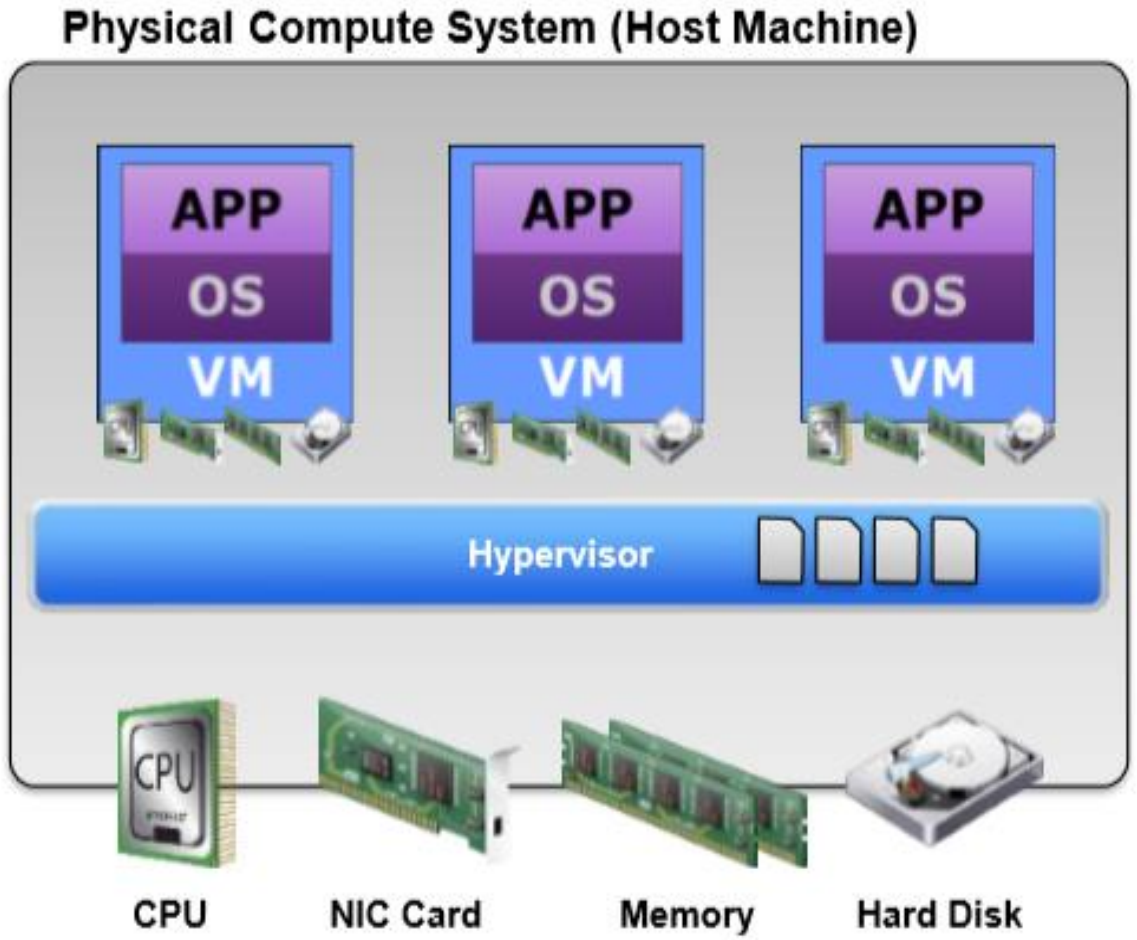
Applications are massively scalable

Applications support anytime access from worldwide locations

Data Source – Hypervisor

A VM appears as a physical compute system with its own CPU, memory, network controller, and disks

Hypervisor allows multiple OSs to run concurrently on a single compute system



OS and application run inside a VM

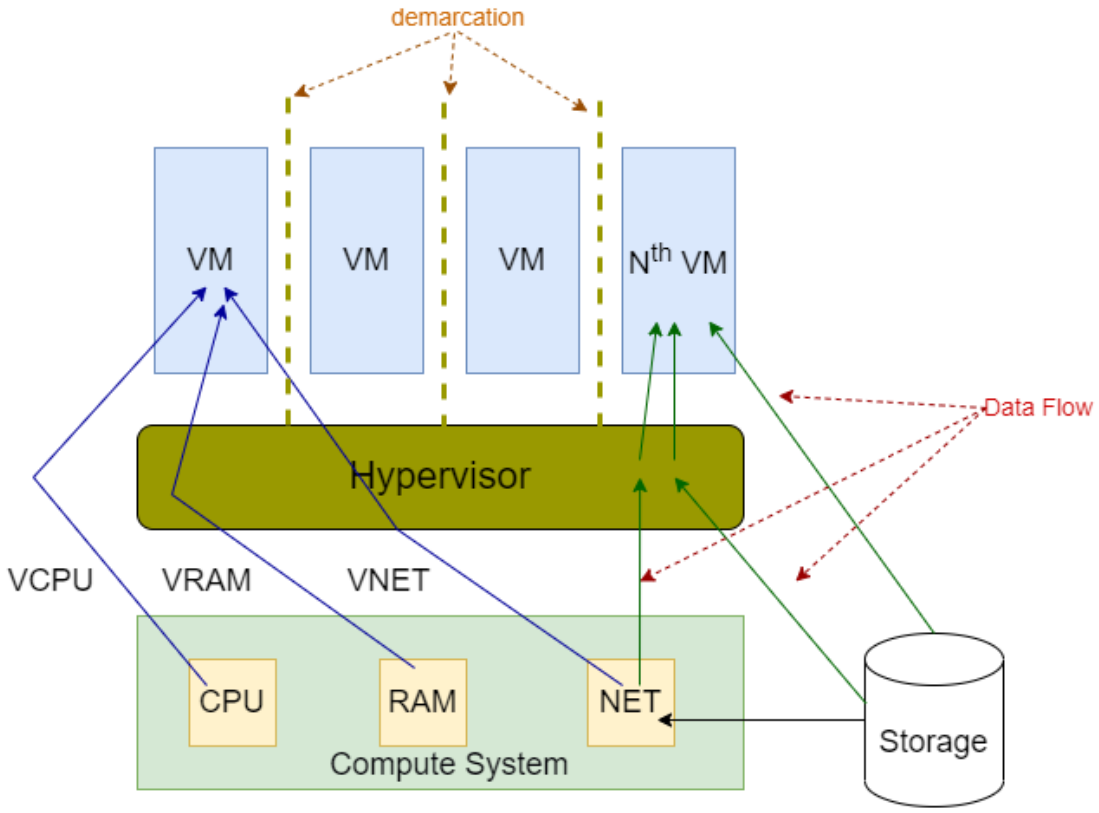
Each VM is isolated from the other VMs on the same physical compute system

From a hypervisor's perspective, a VM is a set of files

Hypervisor is the source of VM files for data protection

Hypervisor

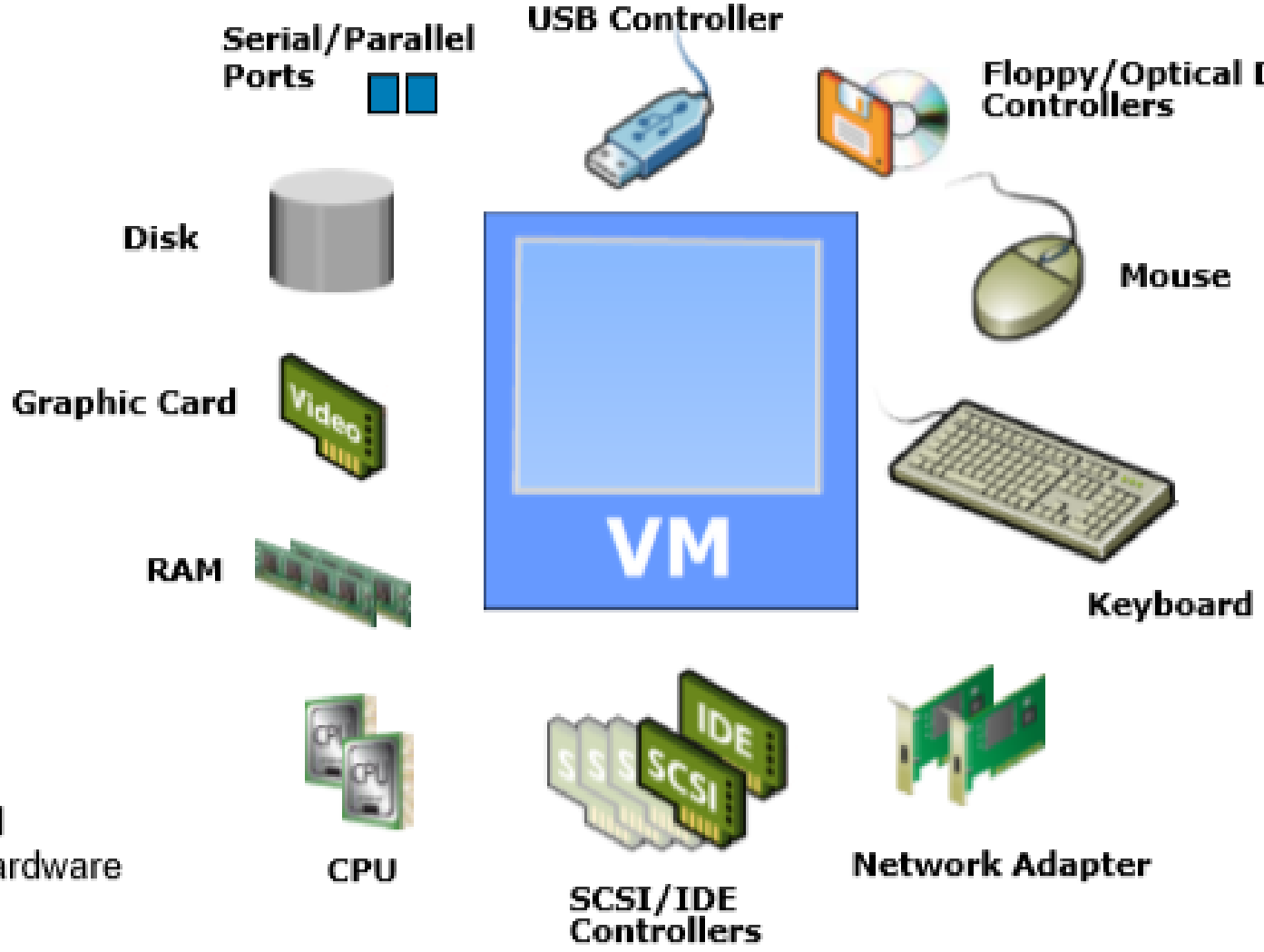
- The hypervisor translates the VM's resource requests and maps the virtual hardware of the VM to the hardware of the physical compute system.
- From a hypervisor's perspective, a VM is a discrete set of files on a storage device.
- https://www.youtube.com/watch?v=LMAEbB2a50M



VM - VIRTUAL MACHINE
VCPU - VIRTUAL CENTRAL PROCESSING UNIT
VRAM - VIRTUAL RAM
VNET - VIRTUAL NETWORK

VM

VM is a logical compute system with virtual hardware



Hypervisor maps the virtual hardware to the physical hardware

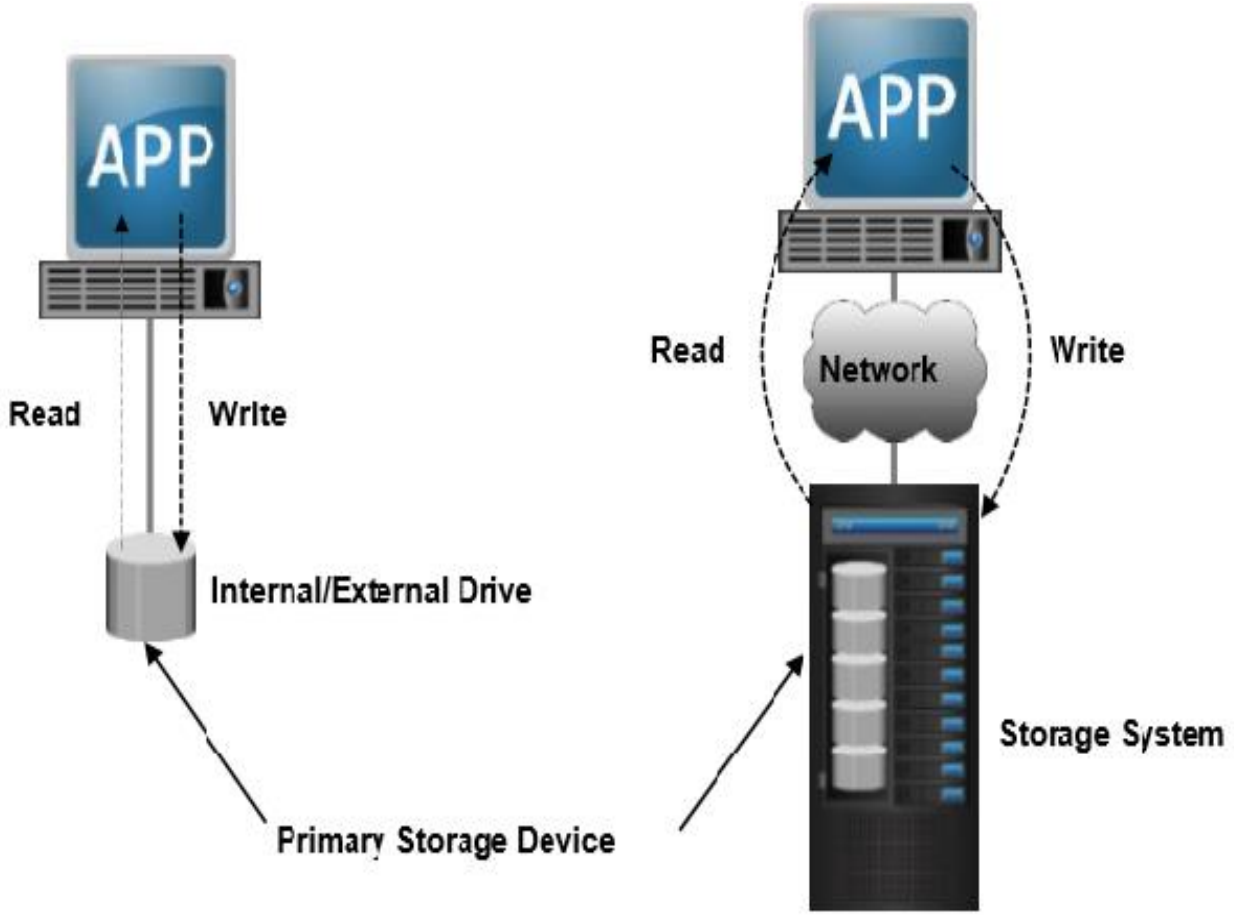
Quick Review

- Explain the purpose of Hypervisor in data protection architecture

Data Source – Primary Storage

It is the persistent storage for data used by business applications to perform transactions.

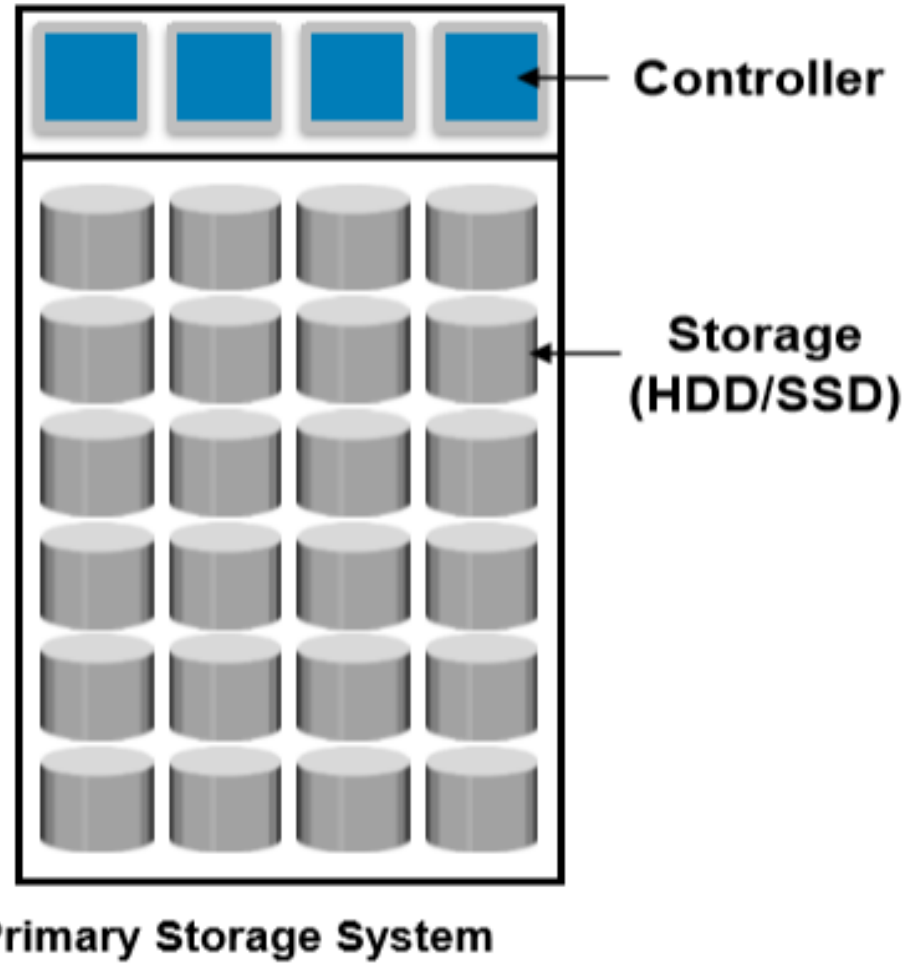
Data from primary storage can be directly copied or moved



Data Source – Primary Storage

- A primary storage device can be leveraged as a data source during protection operations. Data from a primary storage device can be copied or moved directly to protection storage without using the CPU cycles of the compute systems that run business applications and hypervisors. Therefore, application performance is not impacted during data protection. This may also improve the performance of data protection operations.

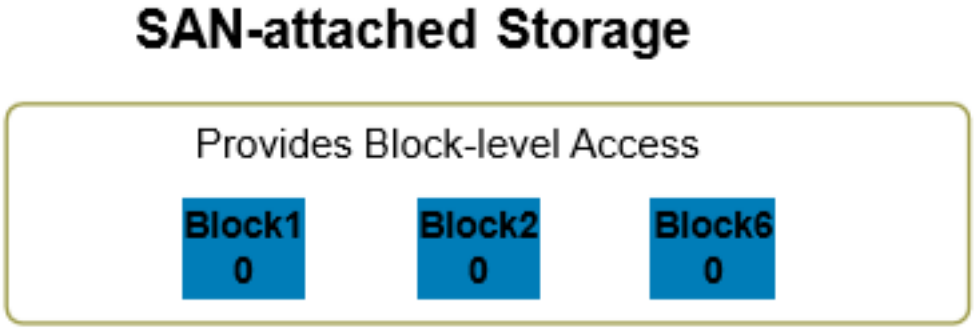
Architecture of a Primary Storage System

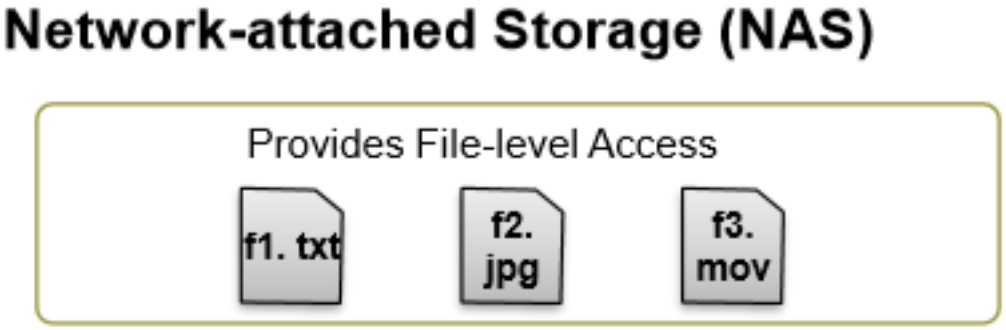


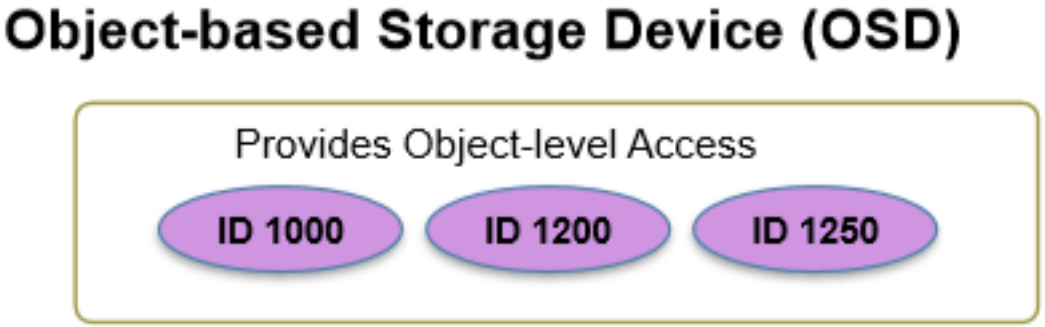
Key components are:

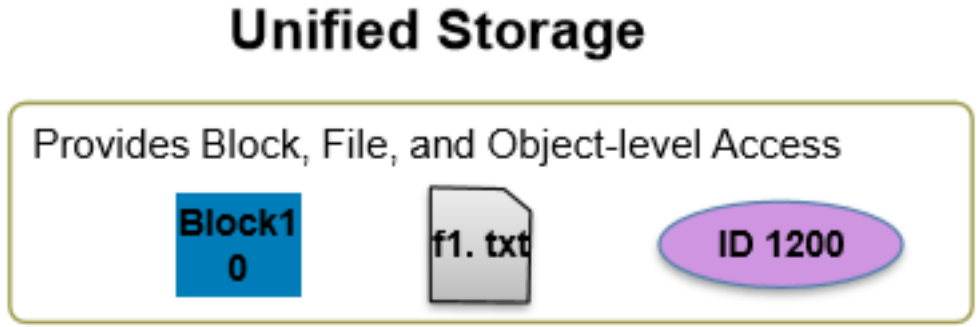
- Controller
  - Manages all storage activities
  - Runs a purpose-built OS
  - Consists of processors and cache
- Storage
  - HDDs/SSDs/combination

Common Types of Primary Storage Systems

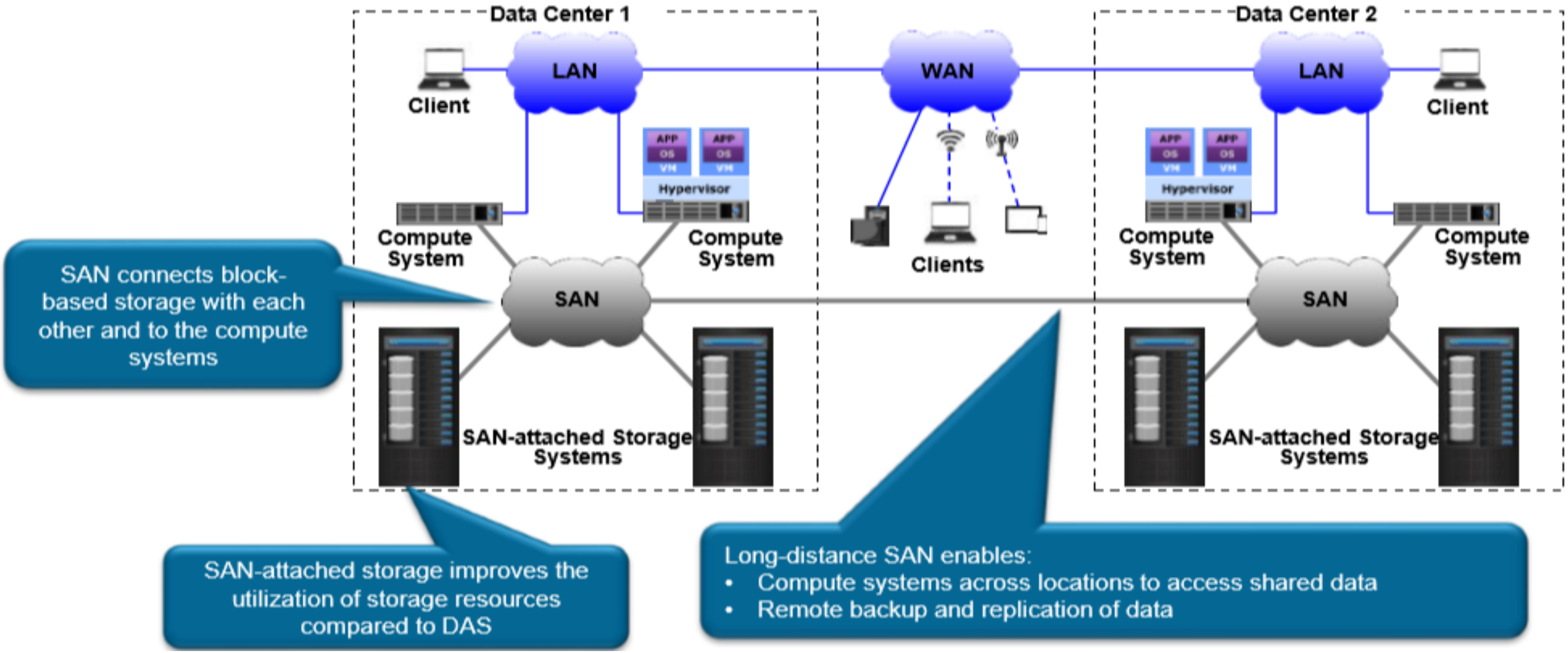




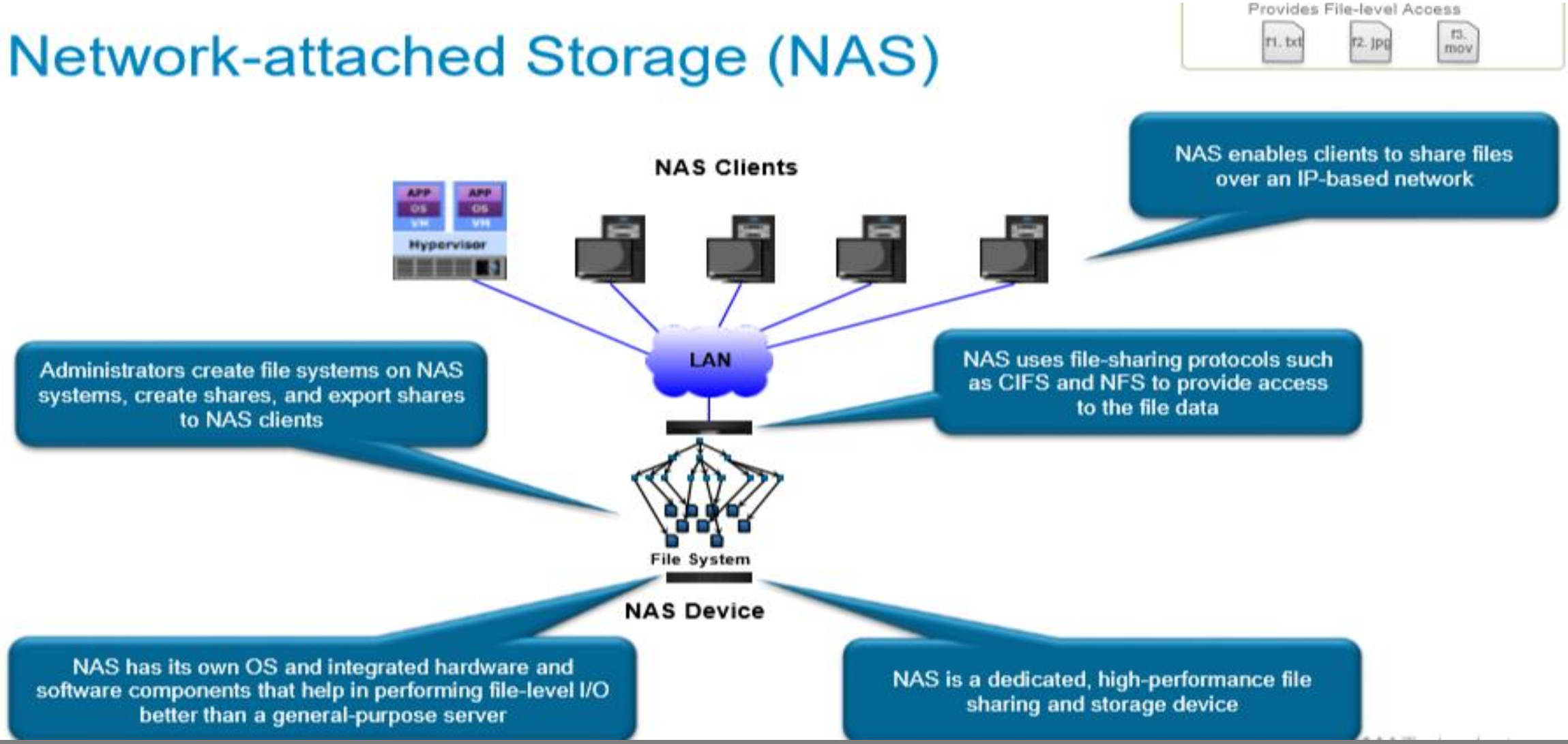




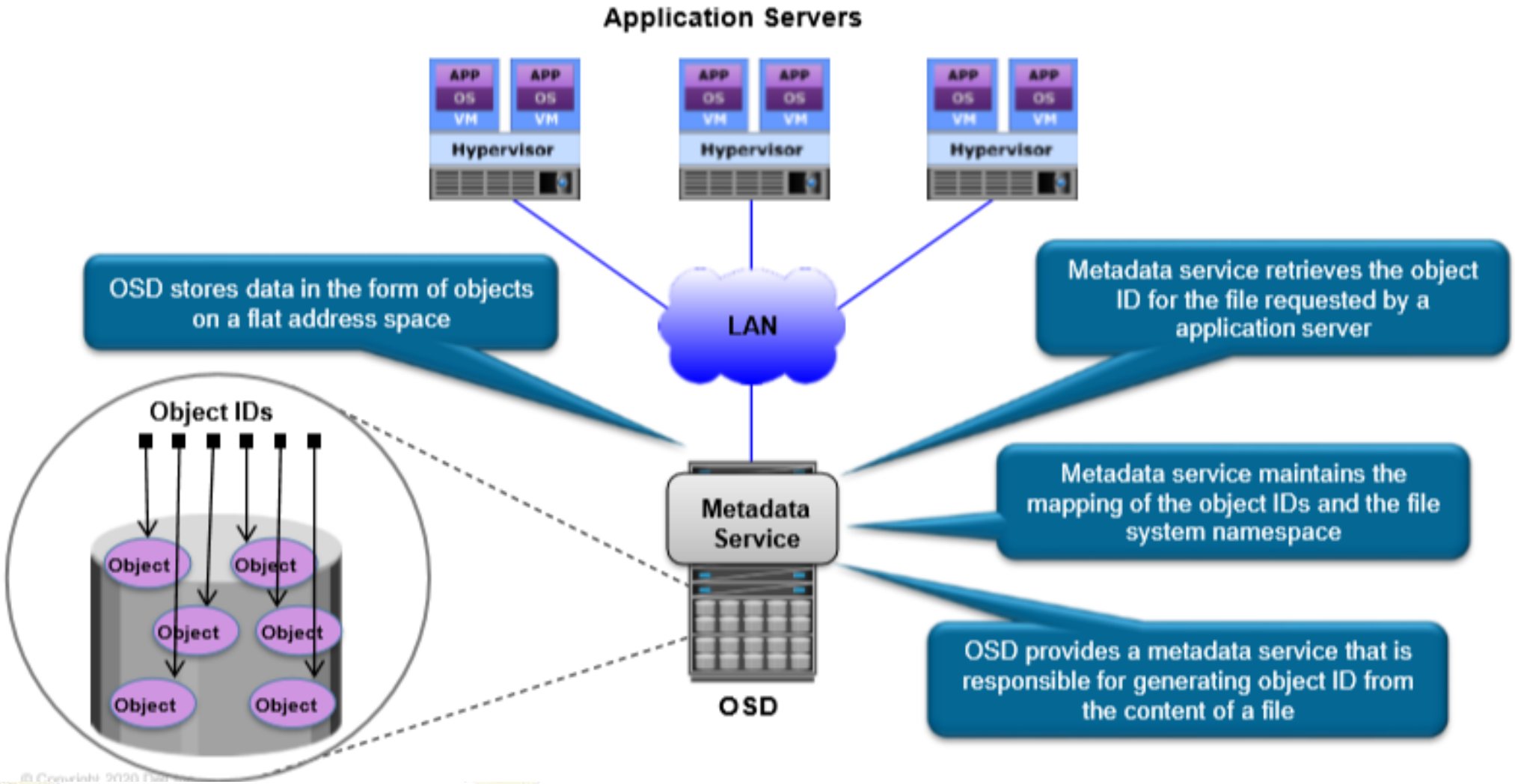
SAN-attached Storage



Network Attached Storage



Object-based Storage Device (OSD)

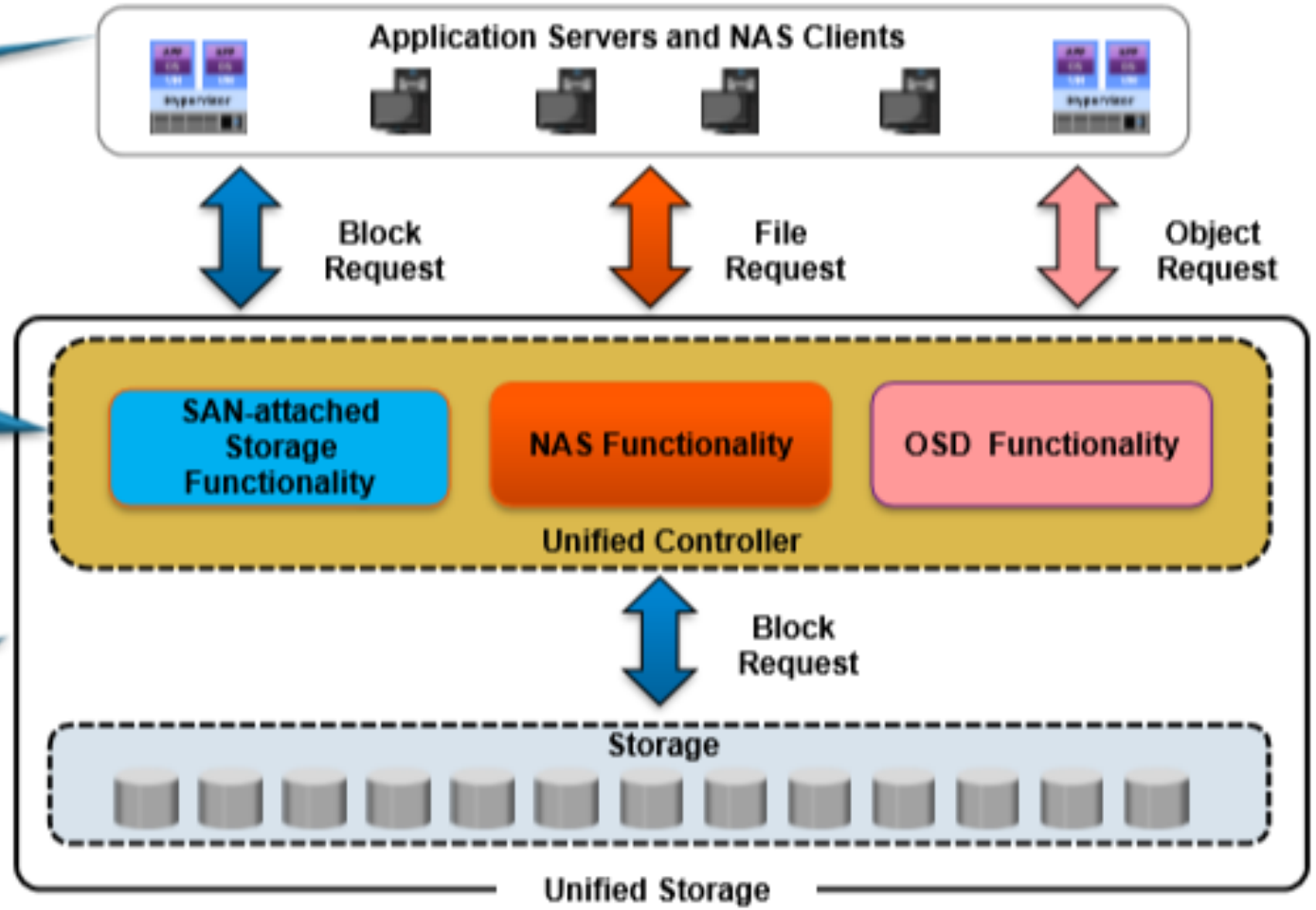


Unified Storage

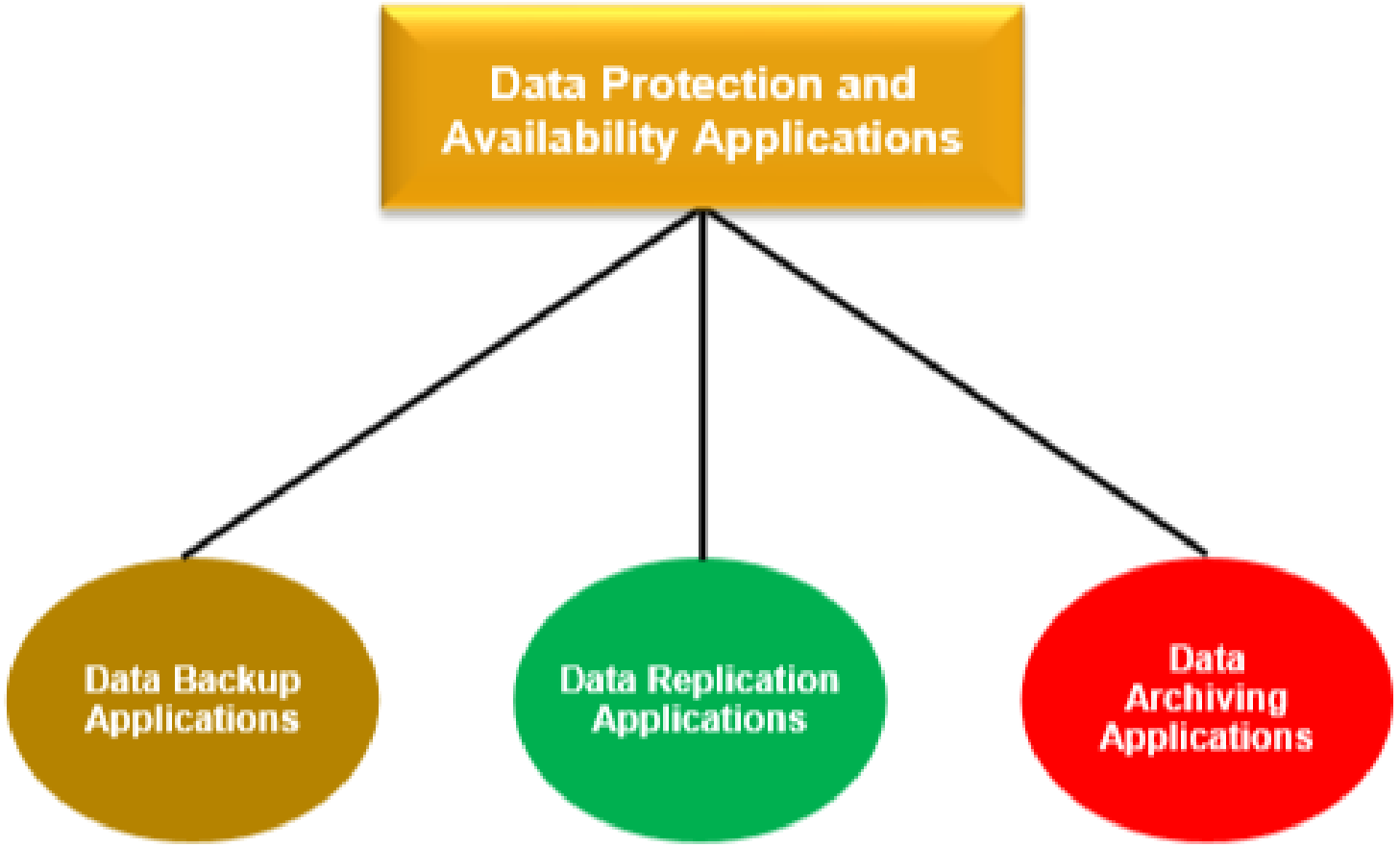
Unified storage resources can be shared by servers and NAS clients with varied levels of data access requirements

Unified controller is capable of processing block-level, file-level, and object-level I/O requests concurrently

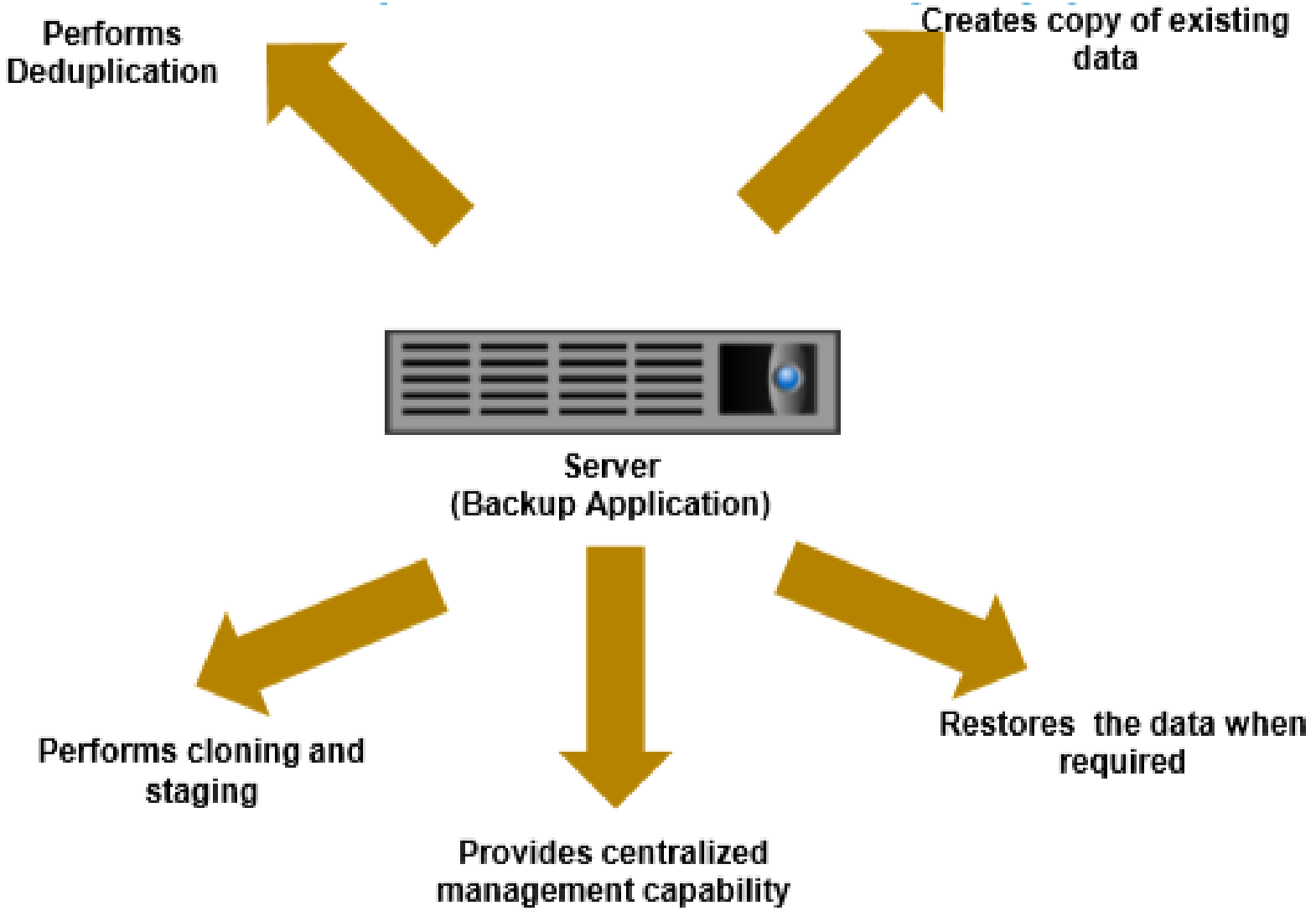
Unified storage consolidates block-level, file-level, and object-level access



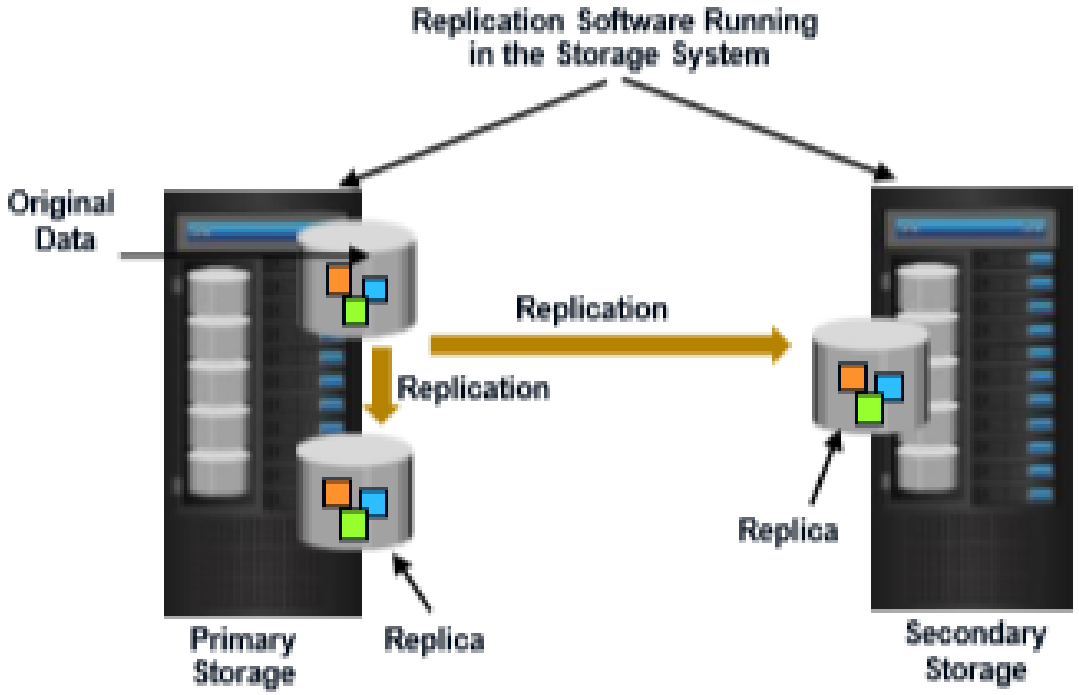
Data Protection and Availability Application Overview



Functions of Backup and Recovery Applications



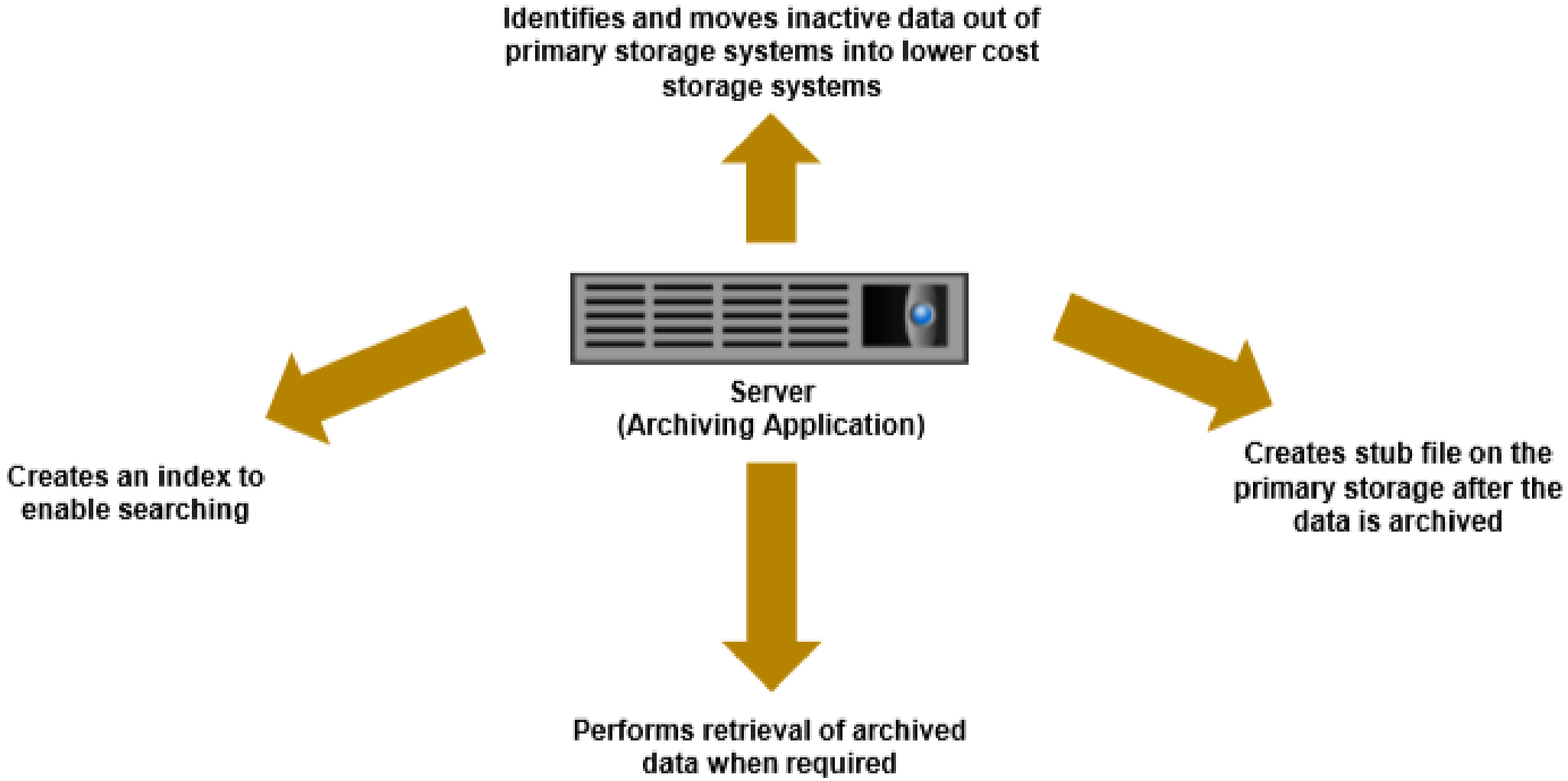
Functions of Replication Applications



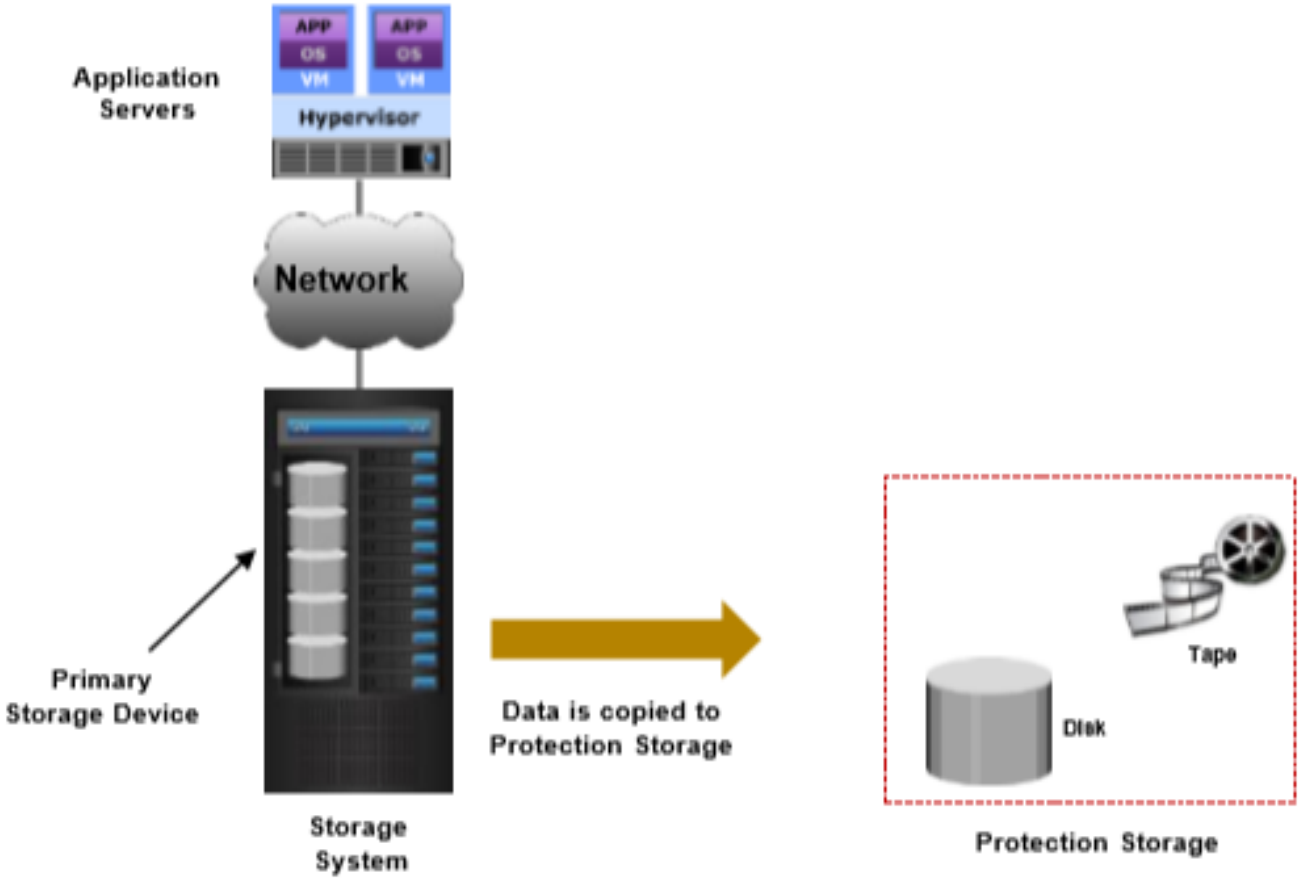
Key Functions of replication applications

- Creates a copy of data
- Creates both local and remote copies of data
- Performs data migration
- Performs compression and encryption
  - Reduces network bandwidth and improves data security

Functions of Archiving Applications



Protection Storage Overview

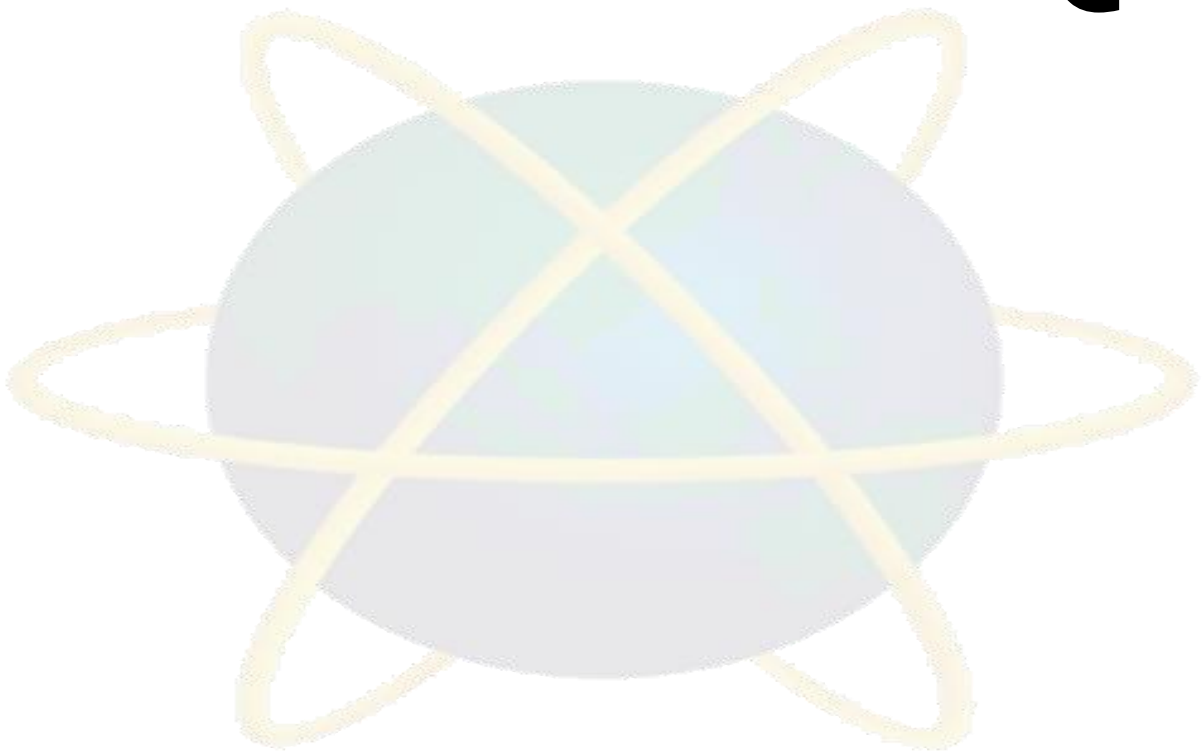


- Protection storage is used to store the data to be protected
- Types of protection storage:
  - Disk-based
  - Tape-based
- Protection storage can reside within a data center or in the cloud

Summary

The data protection architecture is based on the concept of a fault-tolerant data center infrastructure that assures continuous availability of data and services.

Q & A



Next Topic

Data Protection Architecture- Security