



CLOUD COMPUTING

Cloud Storage and Enablers for Storage Virtualization

Dr. Prafullata Kiran Auradkar

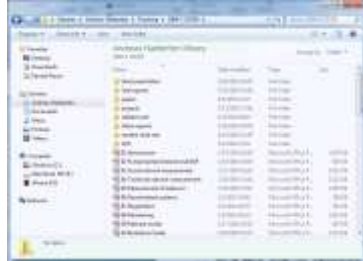
Department of Computer Science and Engineering

Acknowledgements:

Significant information in the slide deck presented through the Unit 3 of the course have been created by **Dr. H.L. Phalachandra** and would like to acknowledge and thank him for the same. There have been some information which I might have leveraged from the content of **Dr. K.V. Subramaniam's** lecture contents too. I may have supplemented the same with contents from books and other sources from Internet and would like to sincerely thank, acknowledge and reiterate that the credit/rights for the same remain with the original authors/publishers only. These are intended for classroom presentation only.

What is Cloud Storage : Complex service, simple storage

Consider the following files which you see regularly



Variable-size files

- read, write, append
- move, rename
- lock, unlock
- ...

Operating system



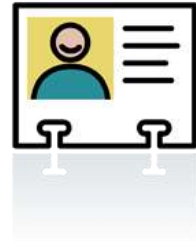
Fixed-size blocks

- read
- write

- PC users see a rich, powerful interface
 - Hierarchical namespace (directories); can move, rename, append to, truncate, (de)compress, view, delete files, ...
- But the actual storage device is very simple
 - HDD only knows how to read and write fixed-size data blocks
- Translation done by the operating system

CLOUD COMPUTING

What is Cloud Storage : Analogy to cloud storage



Shopping carts
Friend lists
User accounts
Profiles
...

Web service



Key/value store
- read, write
- delete

- Many cloud services have a similar structure
 - Users see a rich interface (shopping carts, product categories, searchable index, recommendations, ...)
- But the actual storage service is very simple
 - Read/write 'blocks', similar to a giant hard disk
- Translation done by the web service

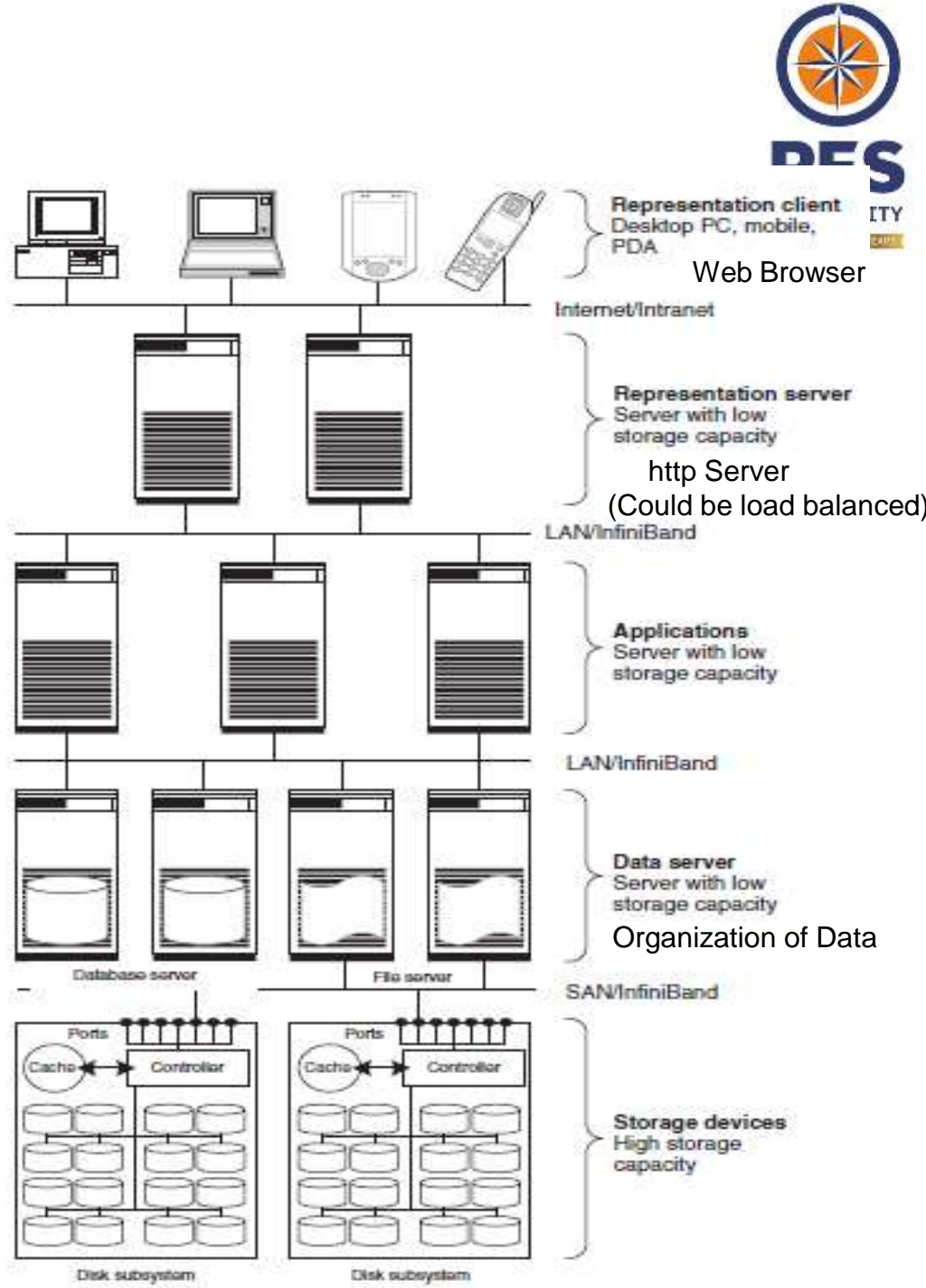
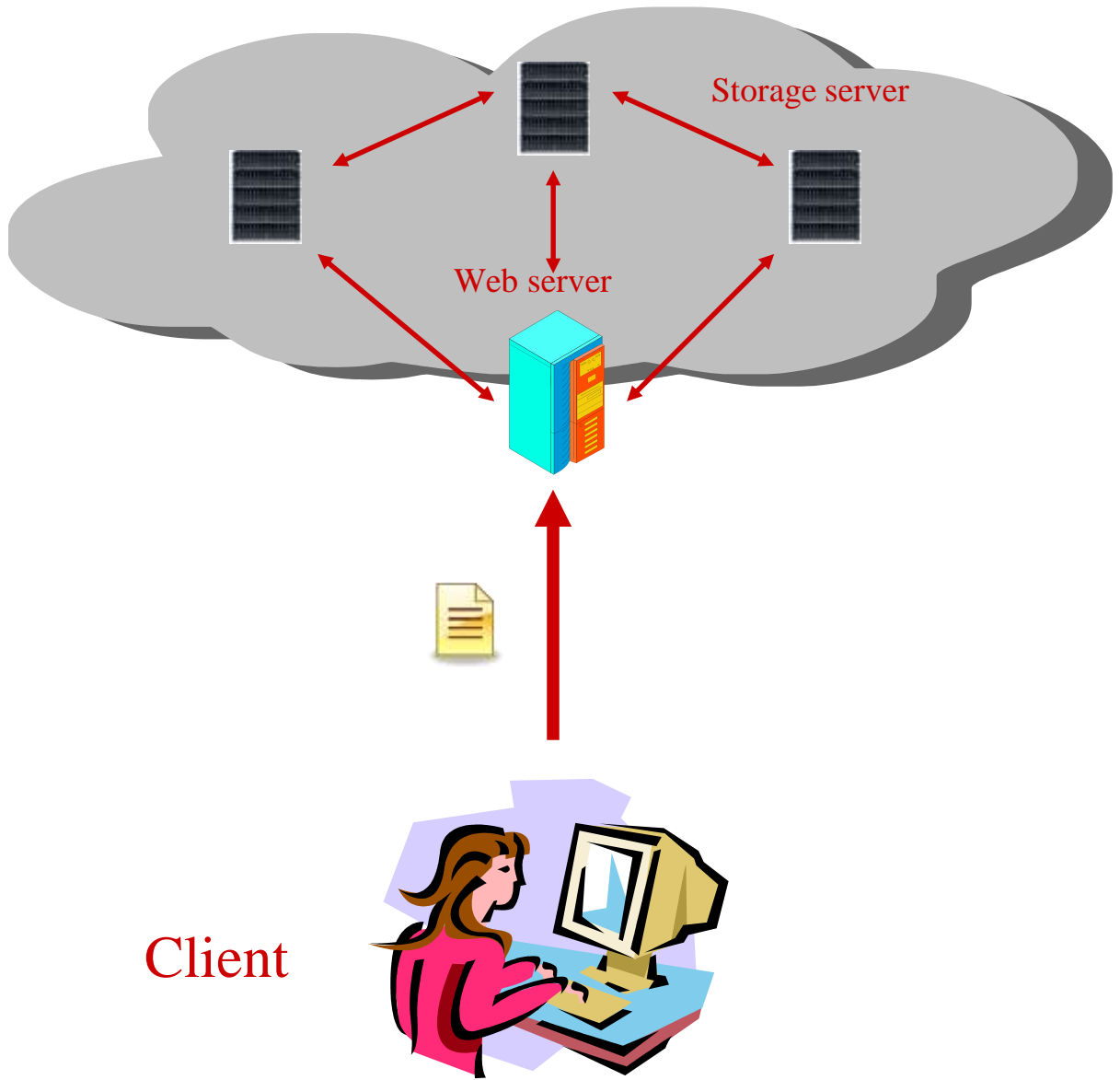
- We have seen the Applications can be thrown into the cloud, and we don't care what server the application lands on. This application needs to store/write or read data, there needs to be an ability for the application to access this seamlessly. This data should also be accessible in case the VM migrates too.
- Cloud Storage provides an ability to an application running some-where, to save data and files in an off-site location, and access the same either through the public internet or a dedicated private network connection from where-ever.
- Storage Virtualization helps supporting access, utilization, availability and other features which are needed by the applications.
- Terminologies like Private cloud, Public cloud, Hybrid Cloud, Internal cloud, external cloud can be applied to the storage based on the where the storage is located, but finally its just the storage systems having the capability of virtual storage pool and multi-tenancy.
- Cloud Storage infrastructure includes the hardware and software cloud components. Object based storage is the prominent approach and access to the infrastructure is via web services API

- **Cloud storage** is a data storage service model in which data is maintained, managed, and backed up remotely and made available to users over a network (typically the internet)
- These cloud storage providers are responsible for keeping the data available and accessible, and the physical environment, protected and running.
- People and organizations buy or lease storage capacity from the providers to store user, organization, or application data.
- Cloud Storage or cloud enabled storage can also be visualized as virtual storage pool
- Object storage services like Amazon S3 and Microsoft Azure Storage, object storage software like Openstack Swift, object storage systems like EMC Atmos, EMC ECS and Hitachi Content Platform, and distributed storage research projects like OceanStore and VISION Cloud are all examples of storage that can be hosted and deployed with cloud storage characteristics.

CLOUD COMPUTING

Storage in the Cloud – Cloud Storage Service

Cloud
Storage
Provider



CLOUD COMPUTING

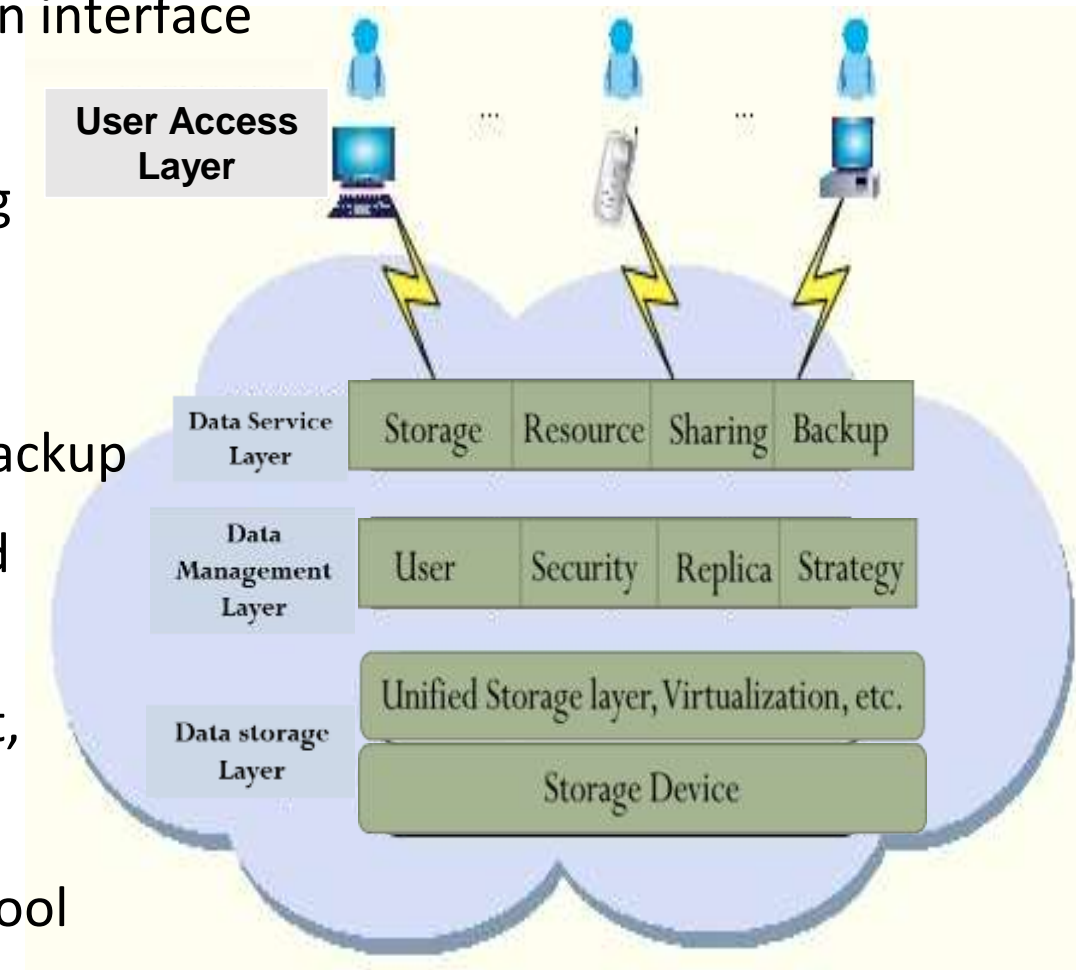
E.g. Architecture of a cloud storage platform

User access layer: an authorized user can log into the cloud storage platform from any location via a standard public application interface and access cloud storage

Data service layer: deals directly with users and depending on user demands, different application interfaces can be developed to provide services such as data storage, space leasing, public resource, multi-user data sharing, or data backup

Data management: provides the upper layer with a unified public management interface for different services. With functions such as user management, security management, replica management, and strategy management

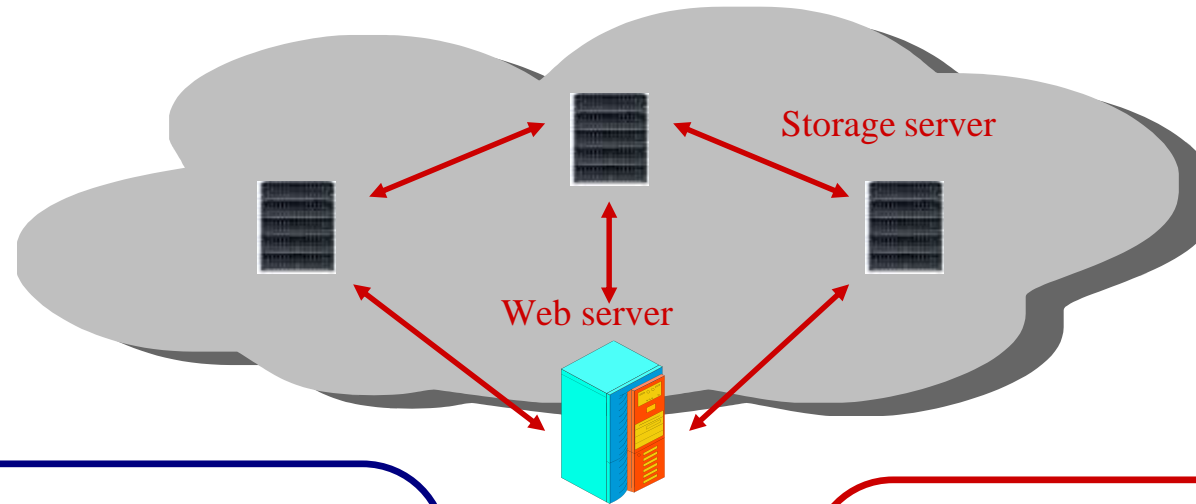
Data storage: data stored in the system forms a massive pool and needs to be organized



CLOUD COMPUTING

Cloud storage : Cloud Storage Service

Cloud
Storage
Provider



Pros:

- Lower cost
- Easier management
- Enables sharing and access from anywhere

Cons:

- Loss of control
- No guarantees of data availability
- Provider failures

Client



1. Dramatic reduction in TCO

- Cuts storage cost by more than 10 times compared to block or file storage

2. Unlimited scalability

- Since built using distributed technologies, has unlimited scalability
- Seamlessly add or remove storage systems from the pool

3. Elasticity

- Storage virtualization decouples and abstracts the storage pool from its physical implementation. So we can get an virtual elastic (grow and shrink as required) and unified storage pool

4. On-Demand

- Uses a pay-as-you-go model, where you pay only for the data stored and the data accessed. For a private cloud, there is a minimal cluster to start with, beyond which it is on-demand.
- This can result in huge cost savings for the storage user.

5. Universal Access

- Traditional storage has limitations like for block storage, the server needing to be on the same SAN , but Cloud storage offers flexibility on the number of users and from where to access the same.

6. Multitenancy

- Cloud Storage is typically multi-tenant and supports centralized management, higher storage utilization and lower costs

7. Data durability and availability

- Runs on commodity hardware but still highly available even with partial failures of the storage system supported by software layer providing the availability

8. Usability

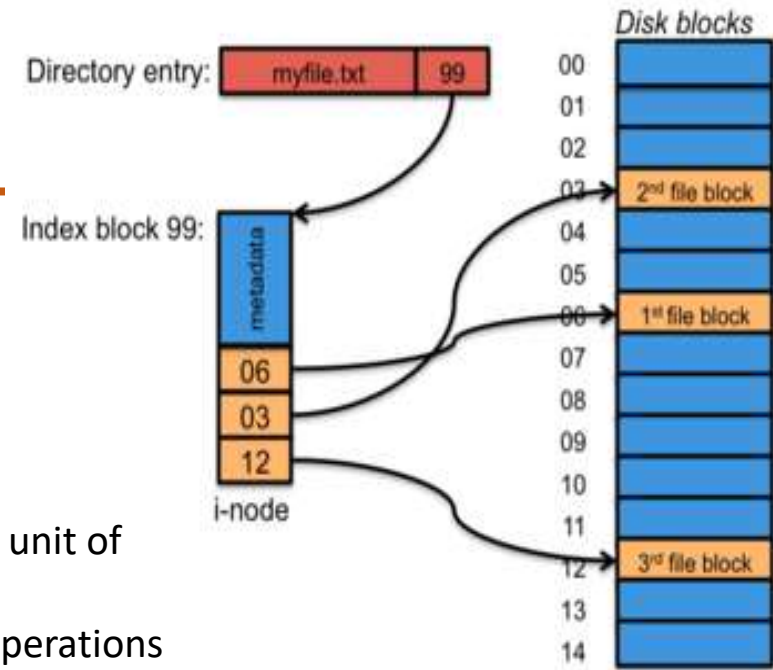
9. Disaster Recovery

- Storage like Compute resources is enabled for the cloud by Virtualization
- Storage virtualization could be implemented in hardware or software.
- Storage virtualization could also be implemented in the Server, in the storage device and in the network carrying the data
- Techniques and components like File Systems, Volume Manager, Logical Volume Managers enable Storage virtualization in the Server
- Techniques like RAID and Logical Volume Management is also used with Storage virtualization in the storage device
- We will discuss a few of these enablers now and discuss the Storage Virtualization itself in the next session.

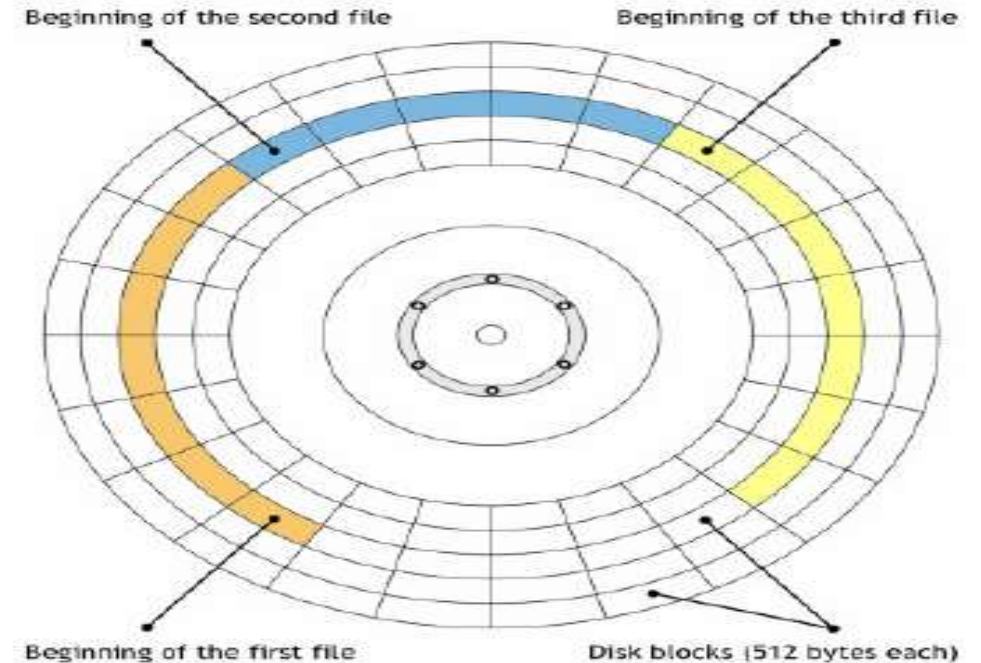
CLOUD COMPUTING

Enablers for Storage Virtualization – File Systems

- Data is separated and grouped into pieces and given a name called a **file**
- The *structure and logic rules* used *to manage the groups of information (files) and their names* is what forms a file system
- **File system** is used to control how data is stored and retrieved.
Without a file system, information placed in a storage medium would be one large body of data with no way to tell where one piece of information stops and the next begins.



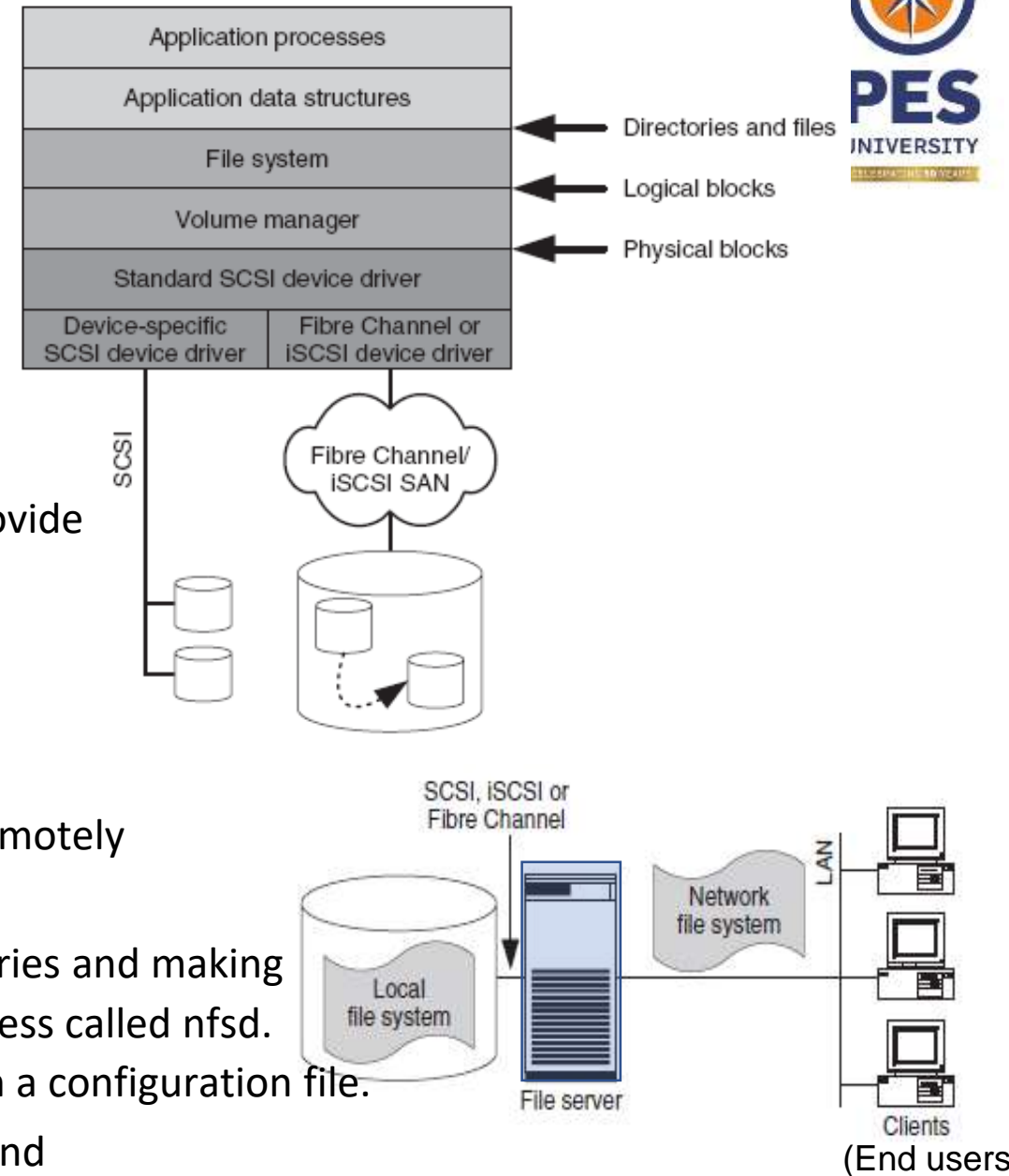
Block is a basic unit of storage for IO (Read/Write) operations



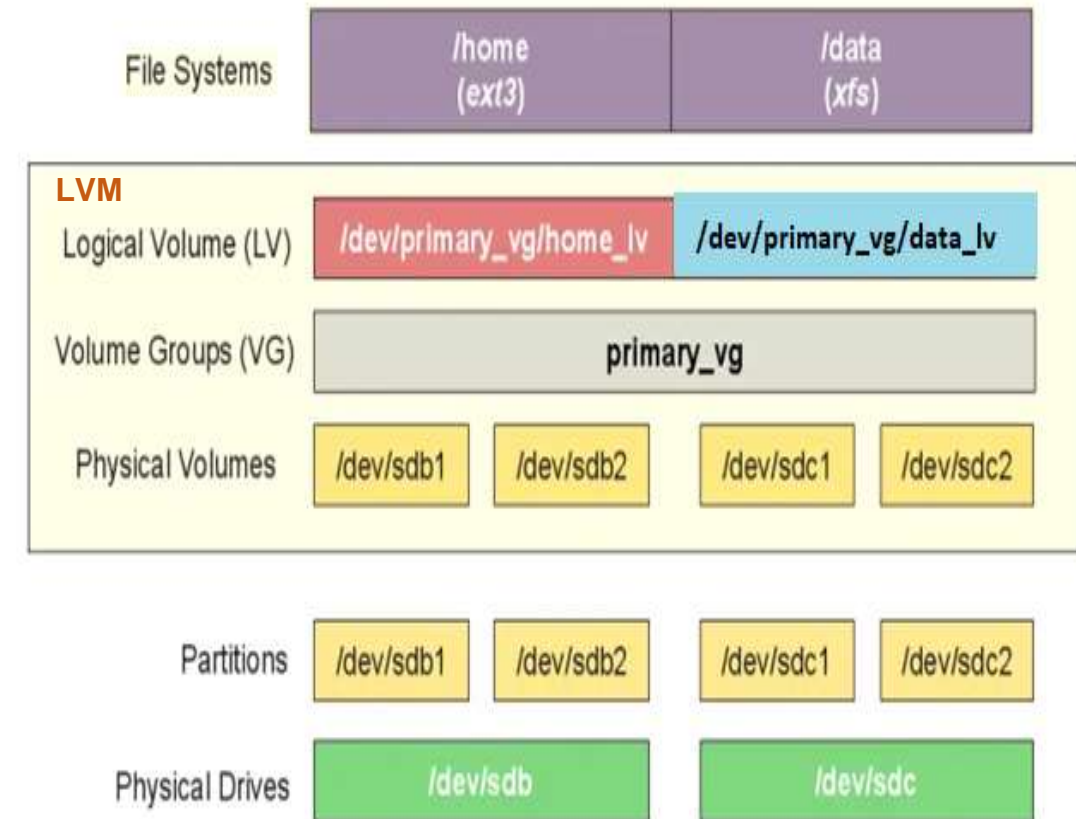
CLOUD COMPUTING

Enablers for Storage Virtualization – File Systems

- Local file systems could be seen in this fashion
- File systems
 - Can be used on numerous different types of storage devices that uses different kinds of media
 - There are also many different kinds of file systems.
- Some filesystems are used on local data storage devices and others provide file access via a network interface and is responsible for arranging storage space;
- Network file systems** make local files and directories available over the LAN. Several end users can thus work on common files
- Network File Systems supports applications to share and access files remotely from various computers
- NFS functionality (in a typical server dedicated to host files and directories and making them available to be accessed across the network), has a daemon process called nfsd. Server administrator exports the directories and advertises the same in a configuration file.
- NFS client requests to the exported directories with the mount command
- Once mounted, its transparent on where the data is being accessed and can be controlled by permissions

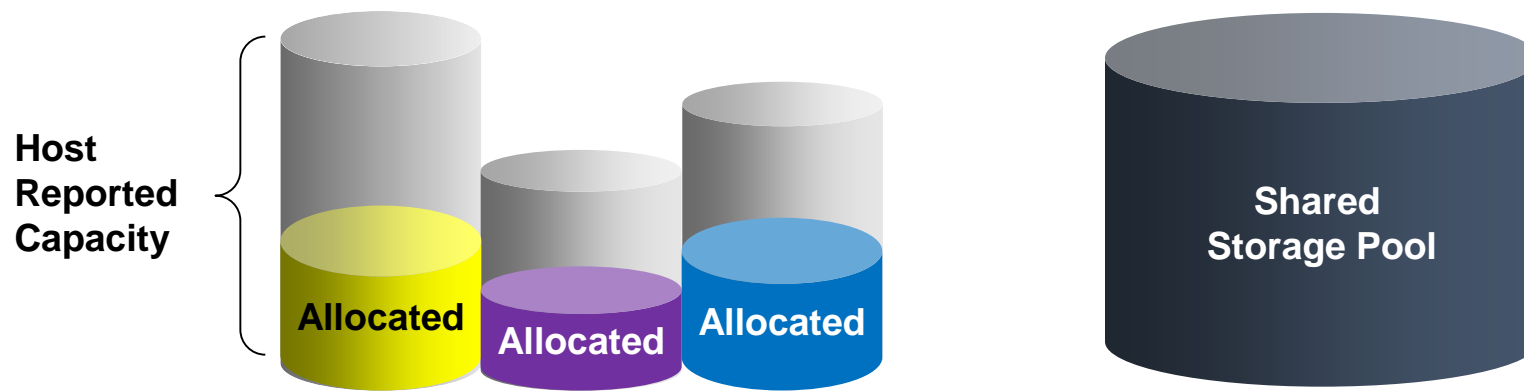


- Independent layer between the File System and the disk drives
- You could create partitions on the physical disk and create physical volumes.
- These physical volumes could be grouped into a volume group
- This volume group could be broken up into logical volumes
- File system can be created on these Logical volumes and mounted



Thin Provisioning or Virtual Provisioning

- Capacity-on-demand from a shared storage pool
 - Logical units presented to hosts have more capacity than physically allocated
(The physical resources are thinly or virtually provisioned)
 - Physical storage is allocated only when the host requires it
 - Provisioning decisions not bound by currently available storage



Storage perceived by the application is larger than physically allocated storage



THANK YOU

Prafullata Kiran Auradkar

Department of Computer Science and Engineering

prafullatak@pes.edu