Student Name: Tevin Achong
Student ID: 816000026
Course Code: INFO3606
Course Title: Cloud Computing
Assignment: 2

March 10, 2020

Question 1

Limitations of leaf spine architecture:

- A large amount of cables and network equipment is required to scale the bandwith since each leaf must be connected to every spine device.
  - This can lead to more expensive spine switches with high port counts.
- The number of hosts that can be supported can be limited due to spine port counts restricting the number of leaf switch connections.

Why we need Software-Defined Networking:

- *Low-Cost Network Devices*:
  - Since intelligence is removed from devices in SDN it is not necessary to use costly network devices in the infrastructure layer.
  - SDN promotes using low-cost network switches.
  - An SDN controller manages the network using higher level policies as defined by applications.
  - It performs all the control logic and gives instructions to network devices, which forward the packets or perform actions in accordance with instructions from the controller.
- *Separation of Control Plane from Data Plane*:
  - The actions of the forwarding plane depend on different tables and logic such as whether to forward the incoming packet, drop the packet, or replicate the packet.
  - The logic and algorithm underlying the actions of the forwarding plane need global knowledge of the network.
  - Actions in a conventional network model are performed in the individual control planes of devices and the state of each and every device remains synchronized so that the entire network can continue working.
  - But with the introduction of network virtualization and multi-tenancy, several virtual devices need to be configured according to the requirements of the application.
  - Hence SDN totally deactivates the control planes of individual devices and replaces them with a centralized controller that manages all forwarding planes.
  - The first task of SDN is to separate the control plane of individual devices.
- *Network Abstraction and Automation*:
  - SDN aims to provide network operators with a global view by hiding the internal workings of the network. SDN provides complete abstraction, so much so that the network programmer can easily obtain a global view without being aware of hardware-specific details.
  - He can program the network requirements of applications simply by using NBI interfaces.

    - Network programmers no longer need concern themselves with individual device configuration, physical vendor-specific hardware details, or proprietary interfaces.

- *Openness*:
    - SDN emphasizes open and standard protocols for communication between the controller and network devices (SBI) and communication between applications and the controller (NBI).
    - Open interfaces allow interoperability among devices manufactured by different vendors. Open interfaces will also reduce the cost of networking devices.

Question 2

Network Function Virtualization (NFV), like SDN, is a network-related technique, the primary objective of which is to virtualize network functions such as load balancing, implementing firewall policies and routing, WAN optimization, and deep packet inspection. Back in the day, these functions were implemented using specialized expensive hardware. With the advent of NFV these functions are now implemented in software; this reduces the costs involved in setting up the network infrastructure of large enterprises. Because these functions are implemented using virtual machines and commodity hardware the network provisioning is simplified. NFV plays a crucial role in simplifying network management and provision in the transport, session, presentation, and application layers. The European Telecommunications Standards Institute (ETSI) defined the architecture of NFV and broke it down into three components: NFV infrastructure, virtualized network functions (VNFs), NFV management and orchestration (NFV MANO). Further, NFV helps operations support systems (OSS) to provide internal network requirements and business support systems (BSS) to deal with end users.

The NFV infrastructure comprises physical hardware made up of computing servers, storage devices, and network devices that are distributed across different geographical locations. This physical infrastructure is virtualized into a single logical virtual resource with the help of a virtualization layer on which different virtualized network functions are deployed.

Different networking functions that back in the day were implemented using specialized hardware - such as routers, firewalls, load balancers, switches, access control mechanisms, network address translation (NAT), content delivery network (CDN), and radio access network (RAN) - are now implemented using software, virtual machines (VMs), and commodity hardware. The virtualized network function elements are deployed on top of the NFV infrastructure.

NFV management and orchestration consists of three components: NFV orchestrator, virtualized infrastructure manager, and VNF manager; MANO provides orchestration and; life cycle management for virtualized software resources and other virtualization-related management tasks.

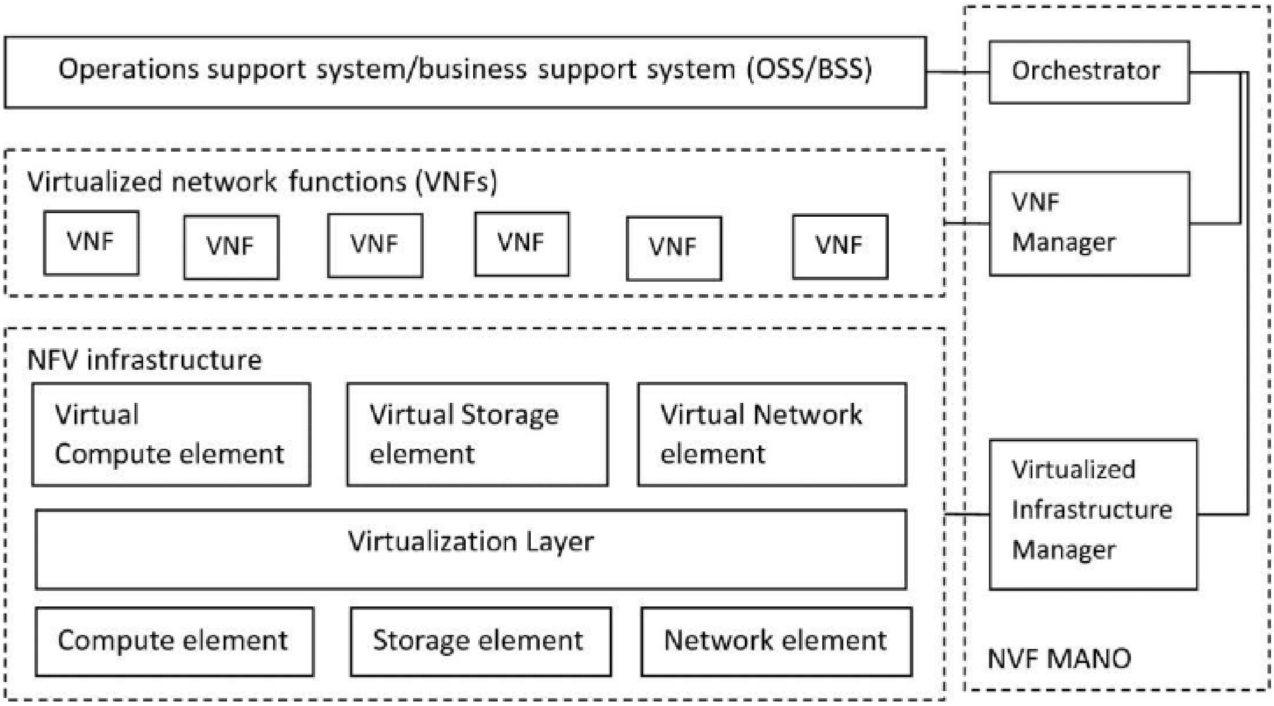

Figure 1: Architecture of NFV

Question 3

- Modern enterprise networks and data centers are large and span a number of buildings.
- They employ virtualization to hide heterogeneity in hardware.
- This led inevitably to huge east-west traffic within enterprise networks.
- Traffic within large enterprises needs to be effectively handled with deterministic latency.
- The number of servers and devices keep on increasing as do the number of virtual servers.
- For example, around 20 virtual servers can be configured on a single physical server.
- Data centers typically contain around 100,000 physical servers.
- Hence a huge number of virtual servers are involved in data centers.
- To support multiple tenants data centers adopt network virtualization.
- Each tenant has his/her own virtual network, which must be logically separated from other virtual networks.
- Everyday digitization of data drives tremendous changes in customer demands and hence enterprises have to evolve to satisfy a wide variety of use cases:
  - videoconferencing applications
  - video sharing and video games
  - video surveillance applications
  - real-time processing of streaming data
  - situational awareness applications with robotics
  - big data applications
  - IoT applications
  - natural language applications
- Typically, most use cases either require
  - huge storage of data
  - huge computing power to process data
  - huge networking capabilities to send and receive data
- The need for business agility drives the automation of networking.
- For example, after testing a code in the development environment let us consider the code is to be moved to the integration and production environments.
- When software is moved to different environments, configuring the required security policies, access policies, and routing policies in each environment becomes essential.

- In addition, the movement of software among different environments will happen on a daily or weekly basis in an Agile enterprise.
- In such a scenario manual configuration of the policies required in each environment will become impractical.
- This drives the automation of networking configuration.
- Network performance includes such features as high bandwith, low latency, low jitter, and no packet loss.
- On the one hand, external customers expect very low latency when they interact with enterprise websites.
- On the other hand, such business processes as manufacturing and monitoring require very frequent data exchange with very low latency.
- Hence it is essential to include various enterprise networking solutions such as
  - WAN optimization
  - static routes
  - migration of servers to nearest points, and
  - caching contents
- As the enterprise network spans across different locations, visibility and control of network devices has to be met with a global view of devices.
- It is up to enterprises to ensure that services are offered in line with agreed services levels (as per the SLA) and have sufficient path redundancy.
- Suitable load balancing has to be ensured for peak loads.
- Security is brought about by making use of a number of security mechanisms such as
  - authentication and identity management
  - access control
  - firewalling
  - single sign-on
  - data loss prevention
  - fraud detection
  - malware detection
  - intrusion detection and prevention
  - audit records
- Security is an important concern as enterprise networks span multiple domains and different service provider sites.
- These requirements make it clear that enterprise networking has to offer quality of service not only in bandwith or performance but in many other areas as well.

- Enterprise networks cannot stand alone from computing or storage technologies.
- Networking techniques should grow in line with the speed of computing and storage because the use cases of business processes are becoming increasingly agile and dynamic.
- The digitization of data drives three major tasks:
  - data storage
  - data transformation
  - data transfer
- All three tasks essentially need networking capabilities.
- Ensuring networking capabilities are optimized is essential when data and processes are not in same location.
- Since data gets distributed the processing of data is done in a distributed manner along with parallel processing techniques.
- This is essentially responsible for bringing a very large amount of internal traffic to the enterprise network.
- Hence large enterprises and data centers prefer another network architecture called leaf spine architecture.
- In leaf spine architecture servers are connected to leaf switches
- Each leaf is connected to every spine
- There is no direct spine-to-spine or leaf-to-leaf connection in this architecture
- The leaf layer consists of access switches that connect to devices such as servers.
- The spine layer is the backbone of the network and is responsible for interconnecting all leaf switches.
- The path is randomly chosen so that traffic load is evenly distributed among the switches of the spine layer.
- The ease of expansion optimizes the IT department's process of scaling the network.
- The important aspect of this architecture is that every server can reach every other server with the same number of hops.
- This makes network latency predictable and consistent.
- In this architecture the spine layer is the backbone of the network and is responsible for interconnecting all leaf switches.
- As enterprises move from on-premises enterprise infrastructures to the cloud infrastructure there is a need not only to consider the architecture of applications but also network architecture models.

- Traditionally, enterprises provided the typical tiers of an application - client tier, business logic tier, and database tier - in different servers and interconnected those servers by means of the network infrastructure.
- In the 2010s enterprises adopted the service-oriented architecture (SOA) model to develop business applications and achieve business agility.
- One of the tenets of this development was to promote reusability and interoperability.
- SOA provides application-to-application interactions with the help of an enterprise service bus (ESB).
- An ESB serves as the central location to integrate applications using a loosely coupled, asynchronous, reliable, and secure messaging model.
- Although an ESB provides centralized integration of applications, it is inflexible.
- An ESB is basically used to integrate different applications, whereas recent microservices architecture (MSA) provides the architecture to develop applications in the form of microservices.
- Microservices are autonomous, self-contained, and self-deployable. This makes deployment much easier.
- Whenever there is a change in a service, only that particular service is redeployed.
- In contrast to a traditional monolithic or SOA-based application, which treats all functions as a single process, MSA treats each service as a process.
- These services can communicate with one another using HTTP.
- Moreover, a microservice can be manipulated through its application programming interface (API).
- In much the same way as they have modernized their application architecture, enterprises primarily focus on the recent software development process model - namely, Agile - in contrast to the old waterfall or other models.
- They do so because the Agile model accepts changing and dynamic customer demands and emphasizes continuous integration and continuous delivery of software development.
- When an application is developed in a traditional manner, continuous delivery becomes difficult.

Question 4

NAS devices make use of the TCP/IP network for file sharing. During this file sharing, at a high level:

- Clients mount the remote file system into their machines (a Windows client has to mount CIFS and a Unix client mounts NFS).
- Clients can view files, create files, retrieve files, change files, etc.

Internally:

- When clients make a request to retrieve a file they use an NFS or CIFS request.
- Clients submit a high-level file I/O request to the NAS gateway.
- The NAS gateway converts the high-level file I/O request into a block-level request with details of the actual blocks where the file is stored, can be written to, or retrieved from the disk.
- The disk produces a block-level I/O response.
- The NAS gateway again converts the block-level I/O response into a file-level I/O response and returns it to the client. It is the NAS gateway that performs all the internal operations and allows very simple file-sharing access to the clients in a network.

Since NAS devices make use of the TCP/IP network for file sharing, a considerable amount of the bandwith available will be utilized by NAS devices.
NAS is less efficient since data transfer is on top of the TCP/IP.

Question 5

From the description of the case study, we know that each of the company's branches has its own dedicated FC SAN; so we know that they have access to the high performance and high scalability provided by FC SAN. However, these advantages of FC SAN bring with them the burden of the additional cost of buying FC components such as HBA, FC switches, and establishing the FC fabric. **One solution Gamma can use to interconnect their storage networks at low cost with no distance limitation is to use Internet Protocol Storage Area Network (IP SAN), using the internet Small Computer System Interface (iSCSI).**

IP SAN is a type of Storage Area Network that has the aim of transporting storage traffic over the existing TCP/IP network using protocols like internet SCSI (iSCSI). Since, from the specification of the case study, we know that the different locations are connected through TCP/IP, using IP SAN is a definite possibility.

The architecture of IP SAN is as follows:

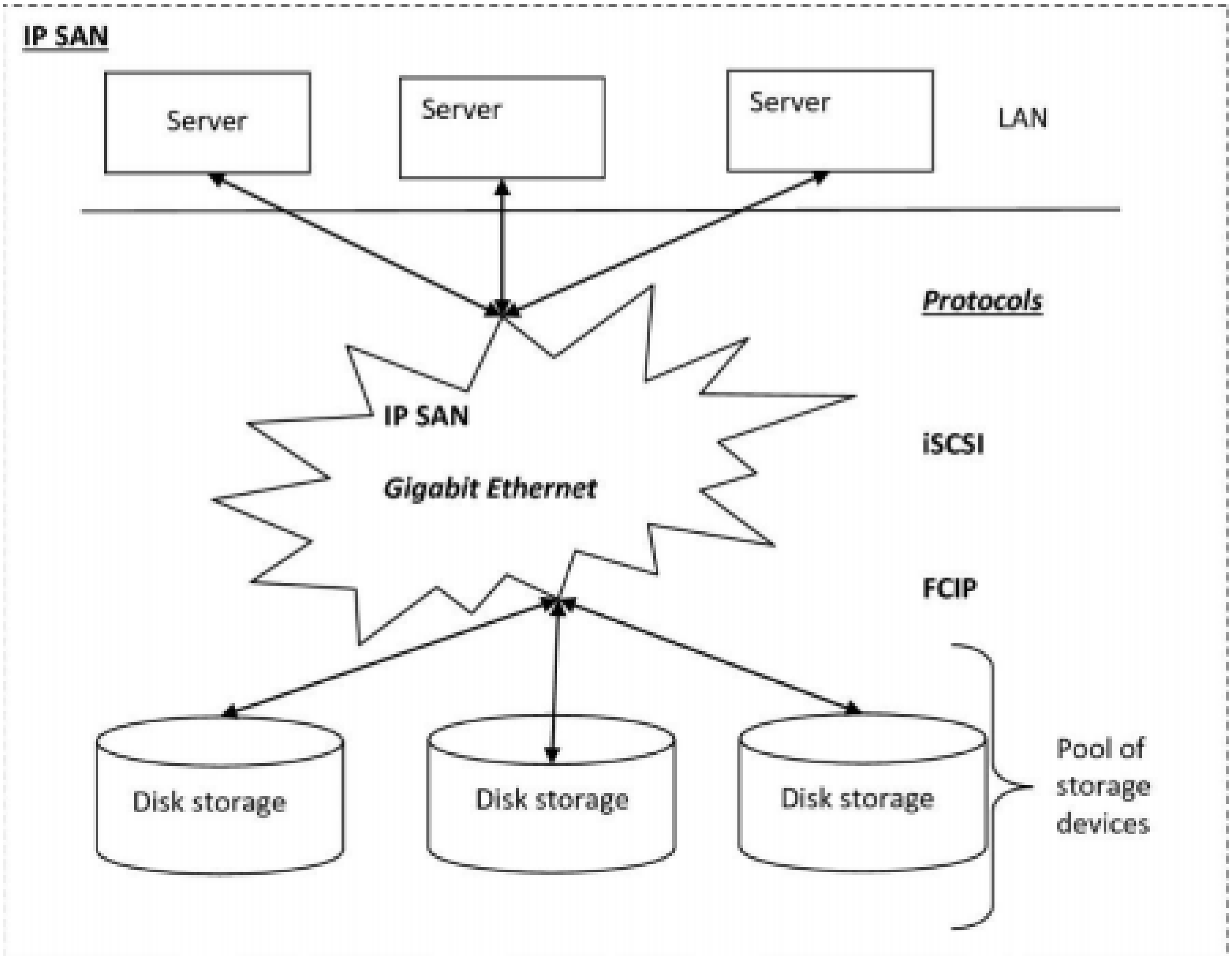

Figure 2: The architecture of IP SAN

- The internet is the primary medium and backbone used to share massive amounts of all kinds of digital information, such as text, audio, and video, in a highly distributed way.

- The growing trend of digitization requires a huge amount of network storage space.
- Each of the protocols in the TCP/IP network are used for communication.
- The primary advantage of IP SAN is that networked storage can be made available at any place wherever there is a TCP/IP network.
- Another advantage is that there is no need to procure special hardware since the network is already existing.
- Hence the cost involved in IP SAN is very low.

Internet Small Computer System Interface (ISCSI)

- **iSCSI is the most common protocol used in IP SAN.**
- Small Computer System Interface (SCSI) is one of the standard interfaces and command sets used to transfer data between computers and disk storage, whereas internet SCSI (iSCSI) uses the same SCSI command set to communicate between computing devices and storage devices via a TCP/IP network.
- iSCSI is a storage networking protocol stack that allows storage resources to be shared over an IP network. iSCSI is a mapping of the SCSI protocol over TCP/IP.

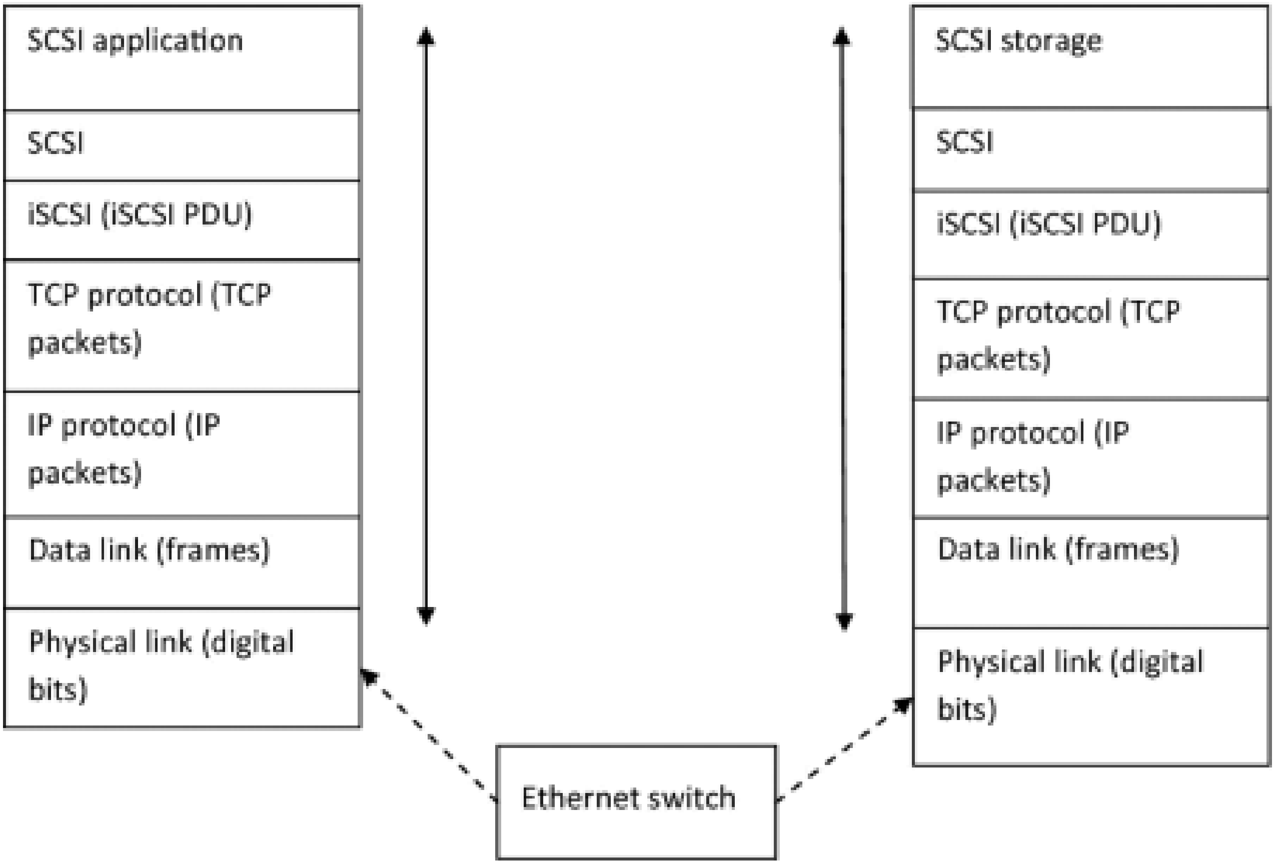

Figure 3: iSCSI over the TCP/IP storage network

- Using SCSI initiates and SCSI request that is mapped onto the internet by the iSCSI layer.

- At they physical layer, iSCSI supports a Gigabit Ethernet interface that enables systems supporting iSCSI interfaces to be directly connected to standard Gigabit switches and IP routers.

- The iSCSI protocol sits above the physical and data link layers and interfaces to the operating system's standard SCSI access method command set. IP SAN uses TCP as a transport mechanism for storage over Ethernet and iSCSI encapsulates SCSI commands into TCP packets, thus enabling the transport of I/O block data over IP networks.

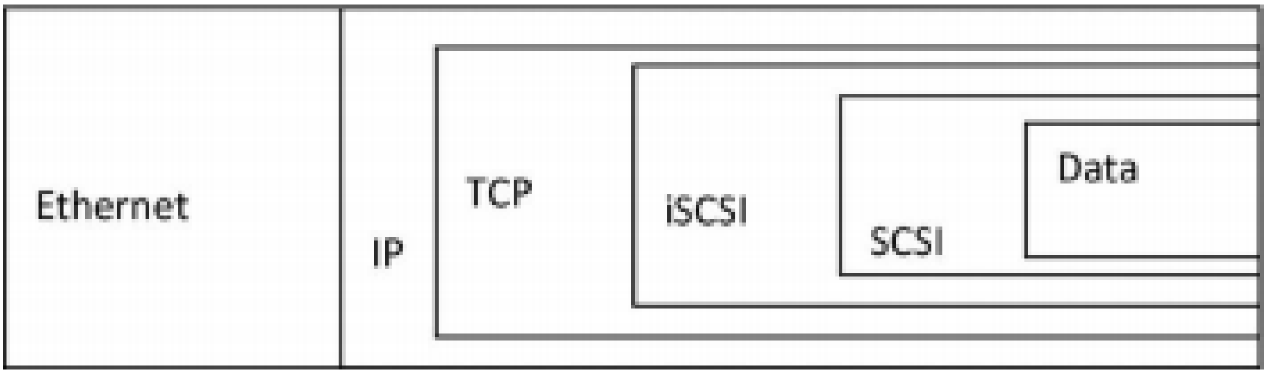

Figure 4: Encapsulation of SCSI commands into TCP/IP

- iSCSI can be supported over all physical media that support TCP/IP as a transport, but today's iSCSI implementations are on Gigabit Ethernet.

- The iSCSI protocol runs on host initiator and on the receiving target device. iSCSI can run in software over a standard Gigabit Ethernet network interface card (NIC) or can be optimized in hardware on an iSCSI host bus adapter (HBA) for better performance.

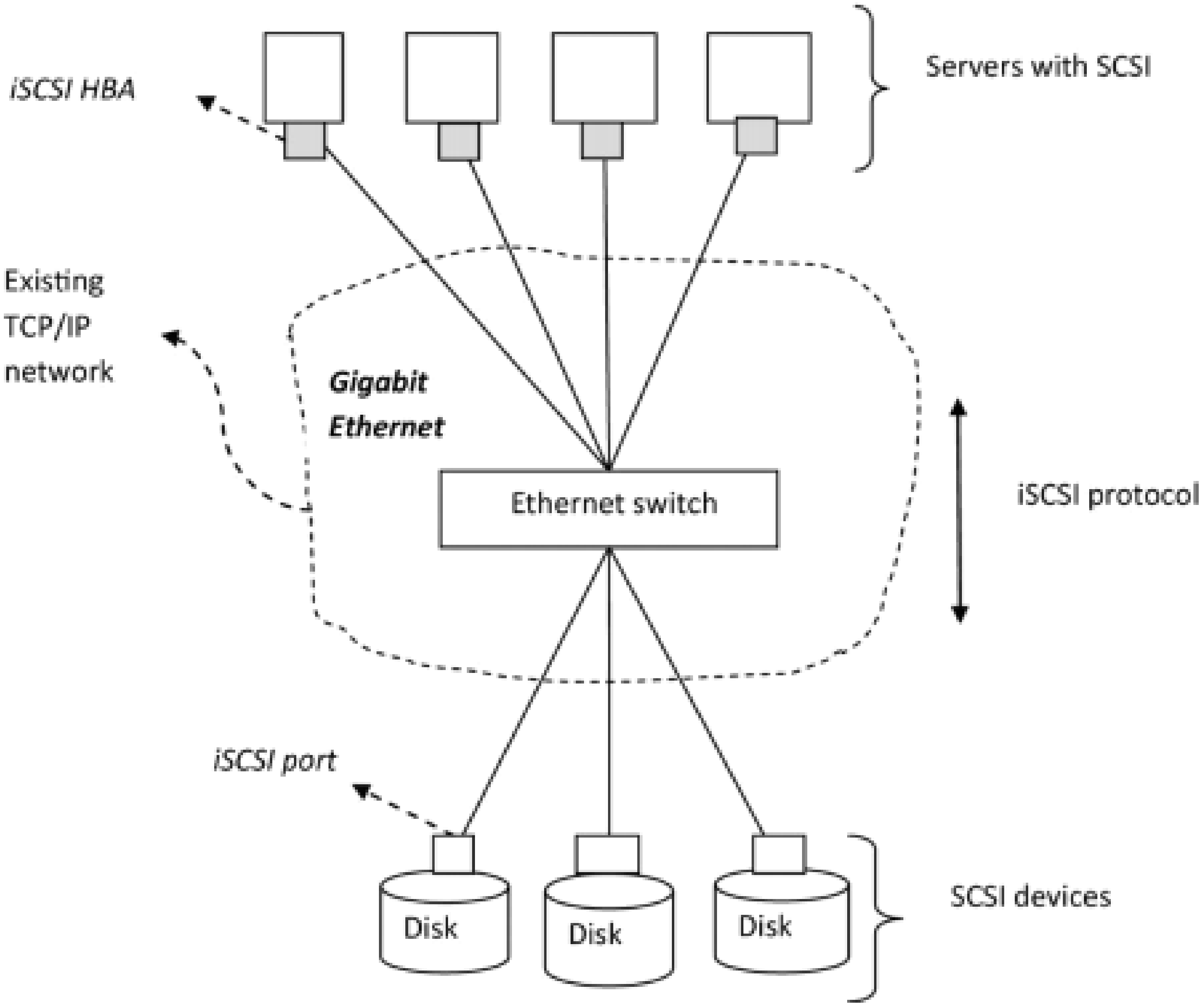

Figure 5: Implementation of IP SAN and iSCSI

- There are three key components for iSCSI communication in IP SAN:
  - iSCSI initiators
  - iSCSI targets
  - An IP-based network
- An iSCSI initiator sends commands and associated data to a target and the target returns data and responses to the initiator.
- IP SAN using iSCSI is composed of initiators and targets connected by an IP network in which the server acts as the initiator or iSCSI host and the iSCSI storage device acts as the target. iSCSI storage devices and iSCSI servers are connected using one of the following four types of iSCSI interface:
  - a standard NIC with a software iSCSI adapter
  - a TCP offload engine (TOE) NIC with a software iSCSI adapter

  - an iSCSI HBA
  - a converged network adapter (CNA)

Justification of the use IP SAN:

- IP SAN does not require a separate network to manage connections to storage systems. It uses the existing TCP/IP network to create the storage area network. Hence the cost of creating an IP SAN is much less than is the case with other networks.
- **IP SAN does not have any distance limitation.** Hence an enterprise can establish its data center at any location where its network is over a LAN, WAN, or the internet. This allows IP SAN to be used flexibly and conveniently at any place to establish disaster recovery.
- IP SAN can co-exist with FC SAN. Hence an enterprise has the option of combining IP SAN with an already existing FC SAN to keep investment costs for storage to a minimum. Since each branch of Gamma has its own dedicated FC SAN, this is ideal for them.
- IP SAN can use either a specialized HBA to connect servers to the SAN or just use standard NIC cards/Ethernet ports to do the same. This enables server I/O consolidation and reduces complexity/cost.
- Gigabit Server Adapters (NIC cards) can be used to connect to the network and thereby provide the Gigabit speeds so necessary today.
- Implementation and maintenance of IP SAN is easier than is the case with other networks.
- IP SAN is highly suitable when implementing a SAN in virtual server environments.

Question 6

There are three standard data access methods:

1. File-level access
2. Block-level access
3. Object-level access

- *File-Level Access*
  - In file-level access the file system is created on a separate file server at the storage device end and the file-level request is sent over a network.
  - Because data are accessed at the file level this method has a higher overhead than data accessed at the block level.
  - File-based storage is usually accessed via a standard protocol like NFS or SMB/CIFS.
  - Fixed file attributes, such as type, size, date created, and date modified, are stored in the file system.
  - File-based storage is good for sharing files and sharing directories over a LAN or WAN.
- *Block-Level Access*
  - In block-level access, the file system is created on a client compute system, and data are accessed on a network at the block level.
  - In this case raw disks or logical volumes are assigned to the compute system, which the client compute system can format to create its own file system.
  - Access to block storage is usually through a client on the operating system over the FC.
  - Block storage is ideal for databases or VMs.
  - The block-level access mechanism is the typical data access mechanism used in SAN.
  - Data access in this mechanism is done in terms of blocks of fixed size.
  - The tpyical block size in most scenarios is 512 bytes.
- *Object-Level Access*
  - In object-level access data are accessed in terms of variable-sized chunks called objects.
  - Each object gets a unique identifier called an object identifier (OID), which is calculated from the file content and the metadata.
  - Applications access the object using this OID. The OID is generated with the help of a 128-bit random number generator, which helps to ensure that the OID is unique. Other details about the object, such as location and size, are stored in the form of metadata. Data that are stored in object-based storage devices can be accessed using web service APIs such as Representational State Transfer (REST) and Simple Object Access Protocol (SOAP). Some types of object-based storage devices also offer support for protocols such as Hyper Text Transfer Protocol (HTTP) and XML.

Advantages of Object-Level Data Access:

- Object-based storage devices incur much less overhead when performing concurrent read/writes, file locks, and permissions.
- This significantly improves performance and gives massive scaling capabilities to object-based storage devices.
- In addition, the amount of rich metadata associated with each object helps in carrying out analytical operations very efficiently.
- Hence object-based storage devices are ideal candidates for storing data that are generated/used by high-performance big data applications.

Question 7

The point of software-defined storage:

- The goal of software-defined storage is to provide administrators with flexible management capabilities through programming.
- A single software interface can be used to manage a shared storage pool that runs on commodity hardware.
- Whether storage is virtualized or not, SDS provides a simple managing interface that automates the tasks of managing storage.
- SDS is a management API that provides a standard way of configuring software devices.
- This API hides the heterogeneity in devices, vendors, hardware, etc.
- This makes it easier for database administrators to configure devices automatically through software.
- There is no need to set the required parameters in hardware, and by so doing human error during configuration is prevented.

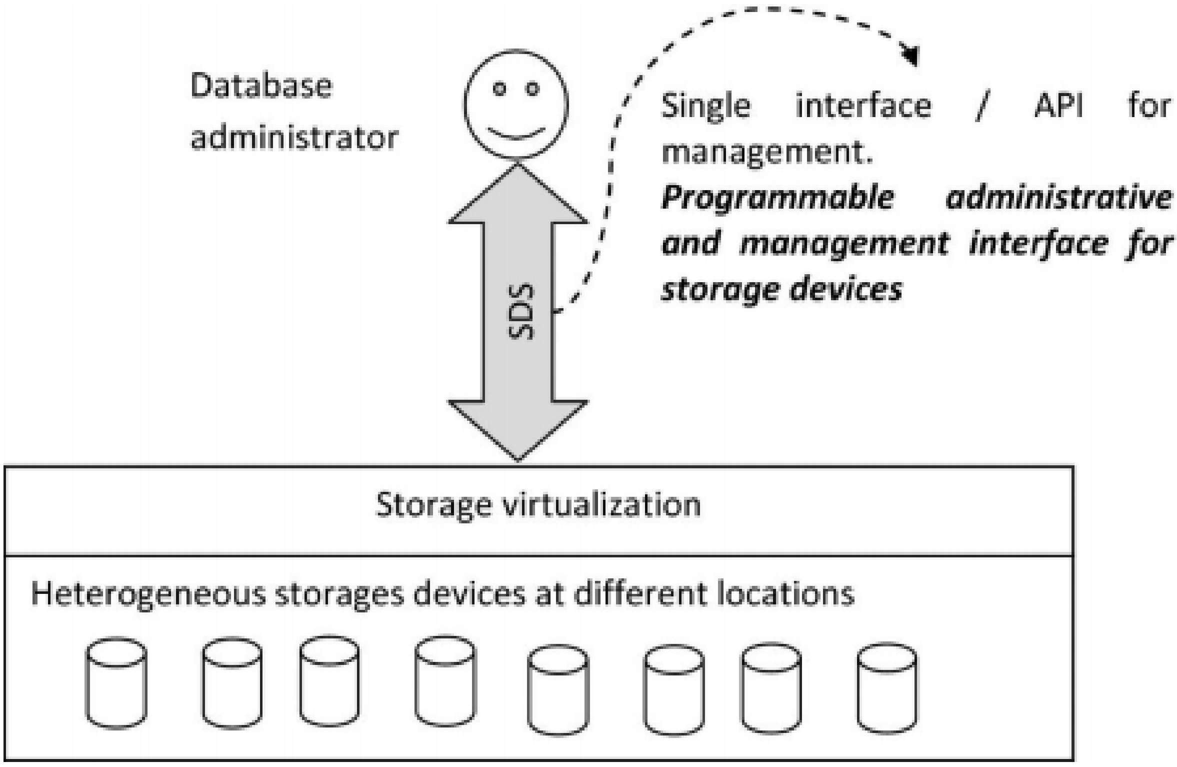
Database administrator
SDS
Single interface / API for management.
Programmable administrative and management interface for storage devices
Storage virtualization
Heterogeneous storages devices at different locations

Figure 6: Concept of software-defined storage

Key characteristics of Software-Defined Storage:

- Abstraction: In SDS architecture the software that manages storage decouples and abstracts the heterogeneity in hardware, vendors, etc.

- Automation: In SDS the configuration and management tasks of storage devices are automated using software configurations, scripts, etc.

- Industry standards: SDS solutions rely on industry standards and hence prevent vendor lock-in.

- Scalability: SDS supports the concepts of virtualization and hence SDS makes it easier to add or remove storage devices.

- Flexibility: SDS allows users to choose hardware according to their needs and removes the constraint of having to deal with a specific vendor.

- Virtualization: SDS architecture pools together storage resources and manages them as a cohesive unit. It virtualizes storage devices in much the same way as server or network virtualization.

Running in a Docker container is an example of a Container-based software-defined storage that exists in the market today.