Homework 5

1. Rack Server: A rack server, also known as a rack-mounted server, is a computer that is dedicated to serving as a server and is built to be positioned in a rack. The rack is divided into bays, each of which is designed to carry a hardware item that is secured with screws. In contrast to a tower server, which is housed into an upright, independent cabinet, a rack server has a low-profile enclosure. Multiple servers can be stacked one on top of the other in a single rack, combining network resources and reducing the amount of floor space required. The rack server architecture also makes network component cabling easier. A unique cooling system is required in an equipment rack filled with servers to prevent excessive heat accumulation that would otherwise occur when several power-dissipating devices are used.

   Blade Server: A blade server is a modular server that allows for the consolidation of numerous servers into a smaller space. Physically, these servers are very thin, with only CPUs, memory, integrated network controllers, and occasionally storage disks built in. The server chassis will facilitate any video cards or other components that are required. This is where the blades will be inserted. Large data centers frequently use blade servers. Because of their ability to accommodate many servers onto a single rack and provide great processing power. In most cases, a single big chassis, such as HPE's BladeSystem, is put into a server rack, and several blade servers are subsequently slotted into the chassis. The chassis can then provide power, network management, and other functions. This increases the efficiency of each blade server and reduces the number of internal components required. Blade servers are typically used in conjunction with some form of Enterprise Storage System, such as Network Attached Storage (NAS) or a Storage Area Network (SAN) (SAN). They make the most of available space by offering the most processors per RU. Blade Servers also enable quick serviceability by allowing components to be swapped out without shutting down the machine.

   The motivation behind the two forms of server placement is to reduce the complexity in connection, space occupancy, and to provide flexibility.

   - Space: A blade server (also known as a blade server) is even smaller than a rack-mounted server. They've been designed to have a small physical footprint and a simple connecting system. In the face of an ever-increasing number of servers that must be crammed into a data center's limited space, such optimization is required.

   - Computing Power: The reduced form factor is achieved by removing non-computer-related components, such as cooling. As a result, a blade might simply be a computer circuit board with a processor, memory, I/O, and an auxiliary interface. A blade like this can't possibly function on its own. Only when it's installed into a chassis with the missing modules does it work. Multiple blades can be accommodated in the chassis.

- Flexibility: A blade server acts as a switch, connecting the servers within to the outside network. It's worth mentioning that the chassis, like a rack-mounted server, fits into a rack.

(Reference: http://whatis.techtarget.com/definition/rack-server-rack-mounted-server, http://searchdatacenter.techtarget.com/definition/blade-server, https://en.wikipedia.org/wiki/Blade_server)

2. A data center's servers must be interconnected, as well as connected to the rest of the globe. More wires must fit into a given location as the number of services increases. The two ways to connecting, Top-of-Rack (TOR) and End-of-Row (EOR), result in distinct cable alternatives. Ethernet Technology is used to construct both TOR and EOR switches. Ethernet technology is particularly relevant to data centers because it has the ability to eliminate the need for separate storage and interprocessor traffic transmission systems (e.g. FC). In data centers, Ethernet is the most common network protocol for computer-to-computer communication. When the network or devices are busy, however, Ethernet is meant to be a best-effort network that may incur packet loss.

   - The transport protocols, such as the Transmission Control Protocol, are responsible for transport reliability in IP networks under the end-to-end principle (TCP). One area in which Ethernet is evolving is the addition of extensions to the existing protocol suite in order to ensure reliability without the complexity of TCP.

   - With the transition to 10 Gbit/s and greater transmission rates, there is a requirement for finer granularity in bandwidth allocation control and to ensure that it is used more effectively. These improvements are especially significant for making Ethernet a more feasible storage and server cluster transport.

   - The susceptibility of Fiber Channel over Ethernet to frame loss is a primary motive. The higher-level goal is to utilize a single set of Ethernet physical devices, or to use a single set of Ethernet physical devices. for computers to talk to a Storage Area Network, Local Area network and InfiniBand fabric.

   Reference: [https://en.wikipedia.org/wiki/Data_center_bridging]

3. Network Attached Storage (NAS): It's a one-of-a-kind instrument. It consists of hard disks and management software. The sole purpose of a NAS is to serve files over a network. NAS stands for Network Attached Storage, and it is a type of storage that is shared through a network. You'll see special files called 'Shares' that can be accessed over the network once you've connected. To give different levels of access, several user logins can be formed. NAS is perfect for SMBs because it provides a cost-effective means for numerous customers to instantly access data at the file level. SMBs can

benefit from its performance and boost their output. Other benefits include simple setup and deployment against SAN, maximum storage resource usage, and the ability to give RAID redundancy to a large number of users.

Storage Area Networks (SAN): A SAN is a dedicated storage network with excellent performance. It is used to transport data between servers and storage devices. It works independently of the LAN. Fiber channel is used in the SAN system to connect devices like RAID arrays, DAS, and tape libraries to servers. The capacity to transport huge data blocks is the main benefit of SAN. This is particularly useful for bandwidth-hungry applications like photography, database (cloud computing, virtual environments), and transaction processing. Furthermore, SAN provides comprehensive data dependability and availability 24 hours a day, seven days a week.

Limitations of DAS:

The storage device is directly connected to the computer in DAS. A USB-connected external hard drive, for example. DAS units can be connected using a variety of cables, including fiber optics, SAS, SATA, and so on. Unlike NAS and SAN, which are designed to be shared resources, a DAS is designed to be utilized by only one computer. When compared to SAN storage, the key advantages of DAS are its high performance, ease of setup and configuration, and often lower cost. It has the disadvantage of not being able to be administered over a network and possibly not having the same amount of redundancy as a NAS or SAN.

Through a point-to-point link, it connects directly to the processor. NAS (Network Attached Storage) and SAN (Storage Area Network) are two types of storage that are shared across a network. In the case of SAN, this network is designed specifically for and dedicated to storage traffic. The semantics of the interface are the distinction between NAS and SAN. Files or objects are NAS units, while disk blocks are SAN units. Another distinction is that SAN requires specialist transport, FC, which is optimized for storage traffic, whereas NAS only requires an IP network. Although NAS and SAN are easily adaptable to Cloud Computing, DAS has a drawback. When a virtual machine is moved to a new physical host, the storage associated with it must also be moved, results in consuming bandwidth and time.

DAS is ideal for storing local data such as boot images and swap space because it is not affected by network delays. DAS can be internal or external, depending on where the storage device is in relation to the host.

(Reference: http://www.computerweekly.com/tip/DAS-vs-NAS-vs-SAN-Which-is-best-for-virtual-storage)

4. A phy layer handles with out-of-band signal line coding and other serial transmission preparations. It has an 8-bit identifier that is device-specific. A management function assigns the identifier.

   A phy is a combination of the physical layer, phy layer, and link layer functions, as specified by SAS. A SAS physical connection pathway must have at least two phys (one at the initiator and the other at the target).

   SAS physical connections (phys) are two differential signal pairs made up of four wires. One differential signal sends data in one way, while the other sends data in the other direction. Data can be sent in both directions at the same time. SAS ports, which contain one or more phys, contain phys. If there are more than one phy in a port, it is called a broad port. It is a narrow port if there is only one phy in it. A SAS global name is used to identify a port (also called SAS address).

   Physical and electrical features of cables, connections, and transceivers are dealt with in the physical layer.

   (Reference:https://www.snia.org/sites/default/education/tutorials/2007/spring/networking/SAS-Overview.pdf)

5. Process control, file manipulation, device manipulation, information maintenance, and communication are the five types of system calls. open(), write(), read(), poll(), and close() are all functions that can be used.

   RPC (Remote Procedure Call) is a strong technology for building client-server systems that are distributed. It works by extending standard local procedure calling so that the called procedure does not have to be in the same address space as the calling procedure. The two operations could be on the same system, or they could be on distinct systems connected via a network.

   Because there is no alteration of the file in this situation, and the original stateless design of servers does not retain note of past recovery, the NFS shut system call does not invoke RPC.
   When a file operation is performed remotely, RPC may not be invoked. The open, shut, read, and write system calls are generic file-related system calls.

   Even though a remote file operation has an RPC counterpart, it does not always result in an RPC invocation. When the information is saved in the client cache, no such invocation is required, reducing the number of remote procedure calls and increasing efficiency. Nonetheless, caching makes maintaining file consistency challenging.

There is no RPC invocation for the close file> system calls for the following reasons:

I. The NFS protocol lacks a close procedure due to the stateless architecture of servers, which do not maintain account of previous requests to aid in crash detection.

II. There is no file alteration in this situation.

(Reference: http://www2.cs.uregina.ca/~hamilton/courses/330/notes/unix/filesyscalls.html, https://en.wikipedia.org/wiki/Remote_procedure_call)

6. The point-to-point architecture is the simplest, with two ports connected directly. It has the same effect as DAS, but it can travel longer distances and work faster.

   The topology of the fabric is the most adaptable. It consists of a set of ports connected by distinct physical links to a network of interconnected FC switches. A 24-bit address space is organised hierarchically in the switching network (or fabric) according to domains and areas. During the fabric login procedure, a unique address is assigned to each attached port. The exact address is usually determined by the fabric's physical port of connection (or switch, to be precise).

   The fabric routes frames individually based on the destination port address in each frame header.

   The arbitrated loop topology allows three or more ports to interconnect without a fabric. In FC-2M, there are three types of connection topologies are supported:

   I. Point to point

   II. Fabric

   III. Arbitrated Loop

   The structure of fabric connections is the most adaptable. It consists of a set of ports connected by distinct physical links to a network of interconnected FC switches. The switching network uses a 24-bit address space that is organized into domains and areas. During the fabric login procedure, a unique address is assigned to each attached port. The exact address is usually determined by the fabric's physical port of connection. Based on the destination port address in each frame header, the fabric routes each frame separately.

(Reference: https://en.wikipedia.org/wiki/Switched_fabric,
https://en.wikipedia.org/wiki/Arbitrated_loop,
https://en.wikipedia.org/wiki/Point-to-point_(telecommunications))

7. Based on the advertisement, an ENode chooses a compatible FCF and sends a discovery solicitation, at which point the capability negotiation begins.
-The FCF responds to the ENode with a requested discovery advertisement after receiving the solicitation, confirming the negotiated capabilities.
-The ENode can build up a virtual link to the FCF after receiving the solicited discovery advertisement. The technique is identical to that of FC's fabric login procedure.
-When the login operation is completed successfully, a virtual port on the ENode, a virtual port on the FCF, and a virtual connection between them are created.
- Based on the advertisement, an ENode chooses a compatible FCF and sends a discovery solicitation, at which point the capability negotiation begins. The FCF responds to the ENode with a requested discovery advertisement after receiving the solicitation, confirming the negotiated capabilities. The ENode can build up a virtual link to the FCF after receiving the solicited discovery advertisement. The technique is identical to that of FC's fabric login procedure. After successfully completing the login operation, a virtual port on the ENode, a virtual port on the FCF, and a virtual link between them are created.

   (Reference: Cloud Computing: Business Trends and Technologies)

8. A. TCP is used in iSCSI to provide characteristics such as dependable in-order delivery, automated retransmission of unacknowledged packets, and congestion control that are critical to SCSI operations. Various iSCSI nodes can be contacted at the same address, and the same iSCSI node can be reached at multiple addresses, hence TCP features are used in iSCSI. As a result, numerous TCP connections can be used to increase the performance of a communication session between two iSCSI nodes.

   B. At the same address, several iSCSI nodes can be contacted, and the same iSCSI node can be reached at many addresses. As a result, numerous TCP connections can be used to increase the performance of a communication session between two iSCSI nodes. Because of dependable in-order delivery, automatic re-transmission of unacknowledged packets, and congestion control, these capabilities are critical to SCSI operations.

   C. The Stream Control Transmission Protocol (SCTP) is comparable to TCP in that it supports SCSI-related capabilities. However, at the time of iSCSI standardization, the SCTP was seen to be too new to be trusted. The SCTP, or Stream Control Transmission Protocol, is similar to TCP in that it supports the capabilities required for SCSI operations. However, at the time of iSCSI standardization, the SCTP was seen to be too new to be trusted.

   D. iSCSI has no built-in means for securing a connection or a session. All native iSCSI communication is unencrypted, making it vulnerable to eavesdropping and active attacks. iSCSI should be used in conjunction with IPsec in an untrusted environment. iSCSI has no built-in means for securing a connection or a session. All native iSCSI communication is unencrypted, making it vulnerable to eavesdropping and active attacks. When working in an untrustworthy environment, iSCSI should be utilized in conjunction with IPsec.

(Reference: Cloud Computing: Business Trends and Technologies)

9. An iSCSI session is a collection of TCP connections that connects an initiator to a destination. This collection may expand and contract over time, allowing us to aggregate many TCP connections for increased throughput. With the availability of many connections comes the challenge of correctly exploiting them in the context of I/O. Separate connections for control and data transfer are undoubtedly sensible to ensure that a connection is always available for task management. However, such a strategy necessitates monitoring and coordination across numerous connections and may even necessitate the use of distinct adaptors on the initiator and target.

   - Connection allegiance is a method used by iSCSI to avoid this complexity. The initiator can issue a command over any connection, but all subsequent communications must be sent over the same connection.

   - It's necessary to keep track of the iSCSI sessions. The iSCSI login mechanism handles a significant portion of session management. The login procedure results in the creation of a new session or the addition of a connection to an existing session.

   - The initiator must know the name and address of the storage device (i.e., the target) that will be used in the procedure. One approach is to pre-configure such information in the initiator. Any subsequent change will necessitate reconfiguration.

   Reference: Cloud Computing: Business Trends and Technologies

10. The ANSI INCITS 458-201140 access control method is based on the concept of capability and credential.

- A capacity is a term that represents a client's access permissions to an object, such as read, write, create, and delete.
- A credential is essentially a cryptographically secured tamper-proof capability that involves a shared key and a keyed-Hash Message Authentication Code (HMAC)41. A credential is a structure that consists of the following elements:

<capability, object storage identifier, capability key>,
where
capability key = HMAC (secret key, capability ‖ object storage identifier).

Example: Based on the capability key, the standardized scheme generates a proof. According to the negotiated security approach, the proof is a quantity computed with the capability key over selective request components.
It should be verifiable, tamper-proof, difficult to forge, and secure against unauthorized usage at the very least.
All but the last condition are met by a credential; there is no built-in way to link it to the acquiring client or the communication channel between the client and the storage device. (A driver's license, on the other hand, features a photograph of the driver to connect the license to the driver, though this isn't necessary for the matter at hand). This is obviously not a good thing, especially if the credential is vulnerable to eavesdropping via an insecure storage transfer. As a result, a new proof methodology is required.

Reference: [Cloud Computing: Business Trends and Technologies]

11. There are three approaches to block-level virtualization depending on where virtualization is done: the host, the network, or the storage device.

i. Virtualization is handled by a volume manager in the host-based method, which could be part of the operating system. The volume manager is in charge of mapping native blocks into logical volumes and monitoring overall storage utilization. The mapping should, ideally, be able to be dynamically modified to allow virtual storage capacity to expand or shrink in response to the most recent needs of a given application.
ii. Virtualization is managed by the storage system's controller in the storage device-based method. This technique tends to produce good performance due to the controller's near proximity to physical storage.
iii. Virtualization is handled by a particular function in a storage network, which may be part of a switch, in the network-based approach. As long as hosts and storage systems support the proper storage network protocols, the technique is transparent (such as FC, FCoE, or iSCSI). It can be classed as in-band (symmetric) or out-of-band (asymmetric) depending on how control and application traffic are handled (asymmetric).

There are three approaches to block-level virtualization depending on where virtualization is done: the host, the network, or the storage device.

- Host-based: In this approach, virtualization is handled by the volume manager which could be part of the operating system. The volume manager is responsible for mapping native blocks into logical volumes while keeping track of the overall storage utilization. A major drawback of the approach is that per-host control is not favorable to optimal storage utilization in a multi host environment, not to mention the operational overhead of the volume manager is multiplied.

- Storage Device-based: In this approach, virtualization is handled by the controller of a storage system. Because of the close proximity of the controller to physical storage, this approach tends to result in good performance.

- Network-based: In this approach, virtualization is handled by a special function in a storage network, which may be part of a switch. The approach is transparent to hosts and storage systems as long as they support the appropriate storage network protocols. Depending on how control traffic and application traffic are handled, it can be further classified as in-band (symmetric) or out-of-band (asymmetric).

In-band approach, where the virtualization function for mapping and I/O redirection is always in the path of both the control and application traffic.

Advantages	Disadvantages
I. On the positive side, the central point of control afforded by the in-band approach simplifies administration and support for advanced storage features such as snapshots, replication and migration.	I. Naturally the virtualization function could become a bottleneck and a single point of failure.
II. The snapshot feature is of particular relevance to Cloud Computing. It can be applied to capture the state of a virtual machine at a certain point in time, reflecting the run-time conditions of its components (e.g., memory, disks, and network interface cards).	II. There is a trade-off as in this case the performance of other virtual machines on the same host may suffer when the snapshot of a virtual machine is being taken

Out-of-band approach, where the virtualization function is in the path of the control traffic but not the application traffic. The virtualization function directs the application traffic.

Advantages	Disadvantages
I. In comparison with the in-band approach, the approach results in better performance since the application traffic can go straight	I. This approach does not lend itself to supporting advanced storage features. More important, it imposes an additional

to the destination without incurring any processing delay in the virtualization function.	requirement on the host to distinguish the control and application traffic and route the traffic appropriately. As a result, the host needs to add a virtualization adaptor, which, incidentally, may also support caching of both metadata and application data to improve performance..
	II. Per-host caching, however, faces the challenging problem of keeping the distributed cache consistent

Given its greater openness and flexibility in storage pooling, I believe the network-based method is most suited for cloud computing. Storage can be assigned to VM hosts, who can then allocate the assigned virtual storage to VMs using their own virtualization capabilities.

(Reference: Cloud Computing: Business Trends and Technologies)

12. The basic build features of NOR flash are similar to those of a NOR gate. NOR flash is quick (at least faster than hard drive) and can address a particular byte at random. Its storage density, on the other hand, is restricted.

    The basic build features of NAND flash are comparable to those of a NAND gate. NAND flash, on the other hand, only enables random access in quantities larger than a byte. NAND flash has made a big impression in consumer electronics, and it's used in a lot more places than NOR flash: digital cameras, portable music players, and smart phones, to name a few examples.

NOR flash	NAND flash
I. Its basic construct has properties resembling those of a NOR gate.	I. Its basic construct has properties similar to those of a NAND gate.
II. It is fast (at least faster than hard disk), and it can be randomly addressed to a given byte.	II. It allows random access only in units that are larger than a byte.
III. Its storage density is limited	

III. It has made a splash in consumer electronics

IV. It is more widely than NOR flash – in digital cameras, portable music players, and smart phones.

Reference:https://www.techtarget.com/searchstorage/definition/NOR-flashmemory#:~:text=NOR%20flash%20vs.,NAND%20flash,control%2C%20address%20and%20data%20information.

13. Inorder to be deployed, in the cloud, the solidstate drives must overcome three limitations inherent to NAND Flash:
    I. A write operation over the existing content requires that this content be erased first. (This makes Write operations much slower than Read operations).

    II. Erase operations are done on a block basis, while write operations on a page basis.

    III. Memory cells erase out after a limited number or write-erase cycles.

    Because of the limits, immediately altering the contents of a page in place will result in a large delay because the full block must be read, erased, and rewritten. Obviously, this is undesirable, leading to the relocation-on-write technique (or out-of- place write).

    (Reference: Cloud Computing: Business Trends and Technologies)

14. Caching the work load data may require more than one server, depending on the size of DRAM available. The hash table is shared across numerous servers in this example, forming a cluster with aggregated DRAM. Memcached servers are not aware of one another and are not coordinated centrally by design. A client's responsibility is to choose which server to utilize, and the client does so based on the key of the data item to be cached (armed with information of the servers in use).

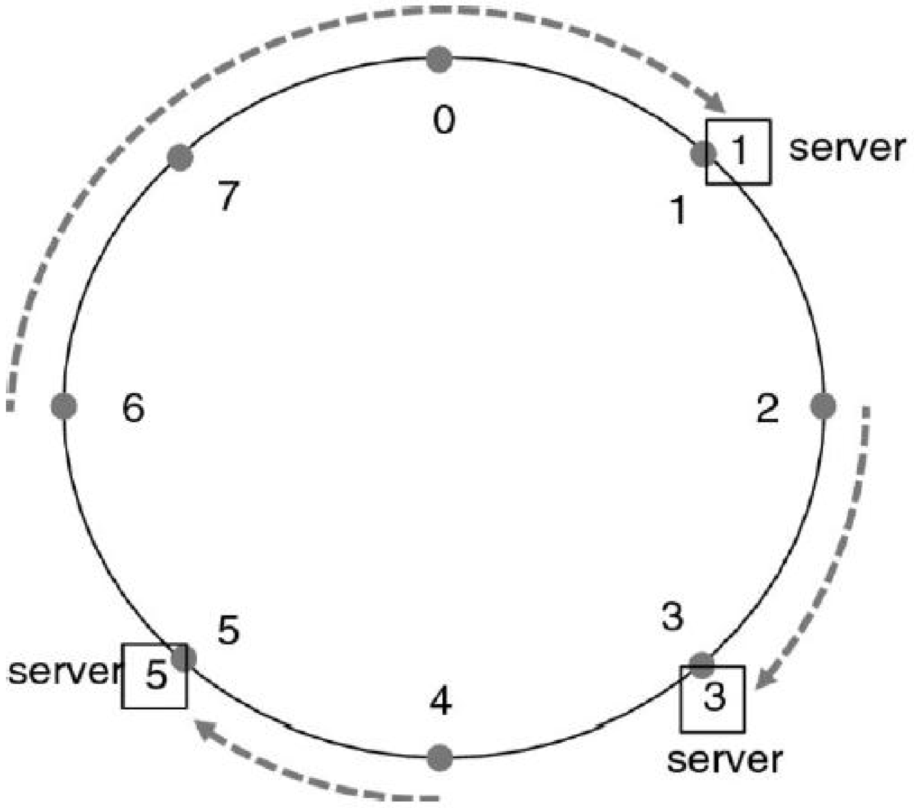
    s=H(k)mod n; where H(k) is a hashing function, k is the key, n is the number of servers, and s is the server label, which is assigned the remainder of H(k) divided by n. . The approach works as long as n is constant, but when the number of servers grows or shrinks dynamically, as it often does in Cloud Computing, it will most likely yield a different server. Cache misses abound, application performance suffers, and all servers in the most recent cluster must be updated as a result.

Since this is undesirable, and so another scheme is in order.

- Memcached implementations usually employ variants of consistent hashing to minimize the updates required as the server pool changes and maximize the chance of having the same server for a given key.
- The basic algorithm of consistent hashing can be outlined as follows:

  -Map the range of a hash function to a circle, with the largest value wrapping around to the smallest value in a clockwise fashion
  -Assign a value (i.e., a point on the circle) to each server in the pool as its identifier49
  -To cache a data item of key k, select the server whose identifier is equal to or larger than H(k).

- An immediate result of consistent hashing is that a departure or an arrival of a server only affects its immediate neighbors. In other words, when a new server p joins the pool, certain keys that were previously assigned to the original p's successor will now be reassigned to server p, while other servers are not affected.
- Similarly, when an old server p leaves the pool, the keys previously assigned to it will now be reassigned to p's successor while other servers are not affected.
- The basic algorithm allows the server pool to scale effectively and provides a sound foundation for further enhancements.



An immediate result of consistent hashing is that a departure or an arrival of a server only affects its immediate neighbors. In other words, when a new server p joins the pool, certain keys that were previously assigned to the original p's successor will now be reassigned to server p, while other servers are not affected. Similarly, when an old server p leaves the pool, the keys previously assigned to it will now be reassigned to p's successor while other servers are not affected. Adding a new server 7 would result in reassigning keys 6 and 7 to the new server; removing server 3 would result in reassigning keys 2 and 3 to server 5.

(Reference: Cloud Computing: Business Trends and Technologies)