

Storage Methodologies In Cloud Computing And Its Role In Data Centers

By

ODION MARLENE ARHELO

A capstone project submitted in partial fulfillment of the requirements for the
degree of

Master of Science

In

Internetworking

Department of Computing Science /Department of Electrical and Computer Engineering

University of Alberta

Supervisor
[Juned Noonari]

© Odion Marlene Arhelo, 2016

ACKNOWLEDGEMENT

I will like to thank my supervisor for his feedbacks and also really appreciate my husband for his support and encouragement.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	i
TABLE OF CONTENTS	ii
LIST OF FIGURES	iv
LIST OF TABLES	v
CHAPTER ONE: INTRODUCTION	1
1.1 Definition of cloud storage and models	1
1.1.1 The private cloud storage	2
1.1.2 Public cloud storage.....	2
1.1.3 Hybrid cloud storage.....	2
1.1.4 Advantages of cloud storage.....	3
1.1.5 Disadvantages of cloud storage.....	4
1.2 Role of storage in data centers	4
1.3 History of storage and fundamentals.....	5
1.4 Storage hierarchy	5
1.5 Storage SLA and RAID architecture	7
1.5.1 Storage SLA	7
1.5.2 RAID Architecture	8
1.6 Storage Logical Unit Number (Storage LUN).....	10
CHAPTER TWO: STORAGE TOPOLOGIES	12
2.1 Direct Attached Storage (DAS).....	12
2.1.1 Advantages of Direct Attached Storage (DAS)	13
2.1.2 Disadvantages of Direct Attached Storage (DAS)	13
2.2 Storage Area Network (SAN).....	14
2.2.1 Advantages of Storage Area Network.....	15
2.2.2 Disadvantages of Storage Area Network	16
2.2.3 Storage Area Network (SAN) Components	16
2.2.4 Storage Area Network Connectivity	18
2.3 Network Attached Storage (NAS).....	18

2.3.1	NAS Protocol	19
2.3.2	Advantages of Network Attached Storage.....	20
2.3.3	Disadvantages of Network Attached Storage (NAS)	20
CHAPTER THREE:	STORAGE PROTOCOLS.....	22
3.1	File Level Storage	22
3.1.1	Features of file level storage	22
3.2	File-Oriented Protocol	22
3.2.1	Network File System (NFS).....	22
3.2.2	Common Internet File System / Server Message Block (CIFS/SMB).....	26
3.3	Block level storage.....	29
3.3.1	Features of block level storage.....	29
3.4	Block-oriented protocol.....	29
3.4.1	Internet Small Computer System Interface (iSCSI)	29
3.4.2	Fiber Channel	30
CHAPTER FOUR:	COMPARISON OF STORAGE TOPOLOGIES AND TRENDS	33
4.1	Pros and Cons of various storage Topologies	33
4.2	Traditional storage vs. cloud storage	35
4.2.1	Comparing traditional storage with cloud-based storage	36
4.3	New market trends and direction of storage technologies	37
CHAPTER 5:	CONCLUSION.....	39
REFERENCES.....		40

LIST OF FIGURES

Figure 1: Storage in cloud	1
Figure 2: Cloud storage models.....	3
Figure 3: Storage Hierarchy	6
Figure 4: Direct Attached Storage	12
Figure 5: Storage Area Network.....	15
Figure 6: Network Attached Storage.....	19
Figure 7: iSCSI Initiators and an iSCSI Target Connected via an IP-based Network	30
Figure 8: Fiber Channel layers vs OSI layers.....	31
Figure 9: Traditional vs cloud storage	35

LIST OF TABLES

Table 2: SLA parameters for storage as a service.....	8
Table 3: NFS protocol stack.....	25
Table 4: Summary of NFS procedures	26
Table 5: CIFS packet header.....	27

CHAPTER ONE: INTRODUCTION

1.1 Definition of cloud storage and models

Cloud storage is a method of data storage in which digital data is hosted, managed and backed up remotely and this stored data can be accessed by authorized users over a network, thereby providing storage as a service to users (Coyne, Shivramakrishnan, & Sing, 2014). This stored digital data is managed by a hosting company who is in charge of ensuring the data is available when needed, it is easily accessible and also protected from unauthorized persons.

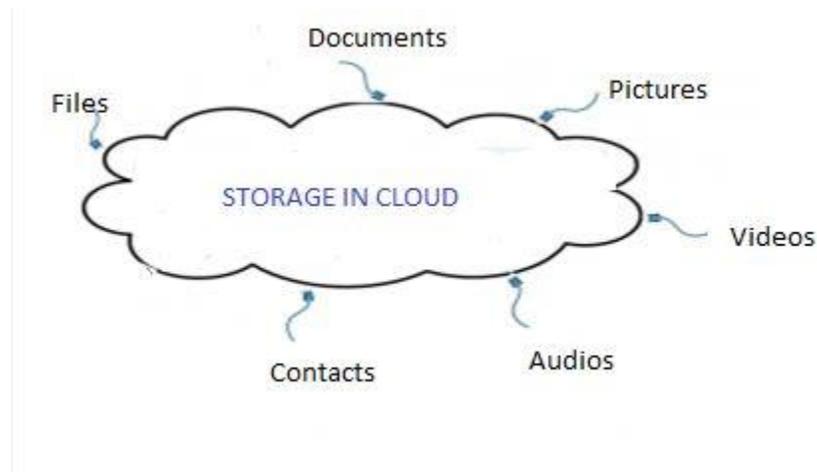


Figure 1: Storage in cloud

Cloud storage evolved at the same time with cloud computing. Cloud storage refers to storage involved with cloud computing. Users accessing data stored in the cloud can have access through different means; access is made through pay by the use of time (pay-per-use), space (pay-per-storage-space) or a combination of both (Jiehui, Jiyi, Jianqing, & Zhijie, 2011). As more companies are sourcing their data storage to cloud providers to save cost and promote ease of access and use of their data, individuals also utilize cloud storage to store their data and easily share such data with others. Data stored in the cloud can range from company's critical documents to individual's pictures or videos. Cloud storage can be classified into three models; these are private cloud storage model, public cloud storage model and the hybrid cloud storage model (Mell & Grance, 2010).

1.1.1 The private cloud storage

The private cloud storage is also referred to as internal cloud storage; it is commonly deployed in a large enterprise, and it is usually designed to meet the specific needs of such enterprise (Mell & Grance, 2010). The enterprise often serves as a cloud service provider to its internal business units which obtain various cloud storage benefits without the burden of providing their own infrastructure (Coyne, Shivaramakrishnan, & Sing, 2014). This model provides the benefits available with public cloud storage, but at the same time addressing issues of security and performance concerns.

1.1.2 Public cloud storage

Public cloud storage model is designed based on a typical cloud computing model. Cloud storage services are accessible to the public via the internet (Mell & Grance, 2010). Accessing the storage services is usually made possible based on authorization, and it may be free or offered on a pay-per-usage model (Buyya, Broberg, & Goscinski, 2011). The public cloud storage is usually cheaper as service provider takes care of the bigger cost involved. Other benefits include its scalability to meet needs, and no resources are wasted as you only pay for what you use (Mell & Grance, 2010). Public cloud storage offers organizations the opportunity to realize many of the advantages provided by private cloud storage, but without the massive costs associated with private cloud storage. However, the choice of public cloud storage can present many drawbacks that may subsequently outweigh the potential costs' savings for the company. Security, reliability and regulatory compliance remain as major concerns with public cloud storage (Hoover, 2013). There is a persistent concern about potential vulnerabilities to exploit other users' sessions with this storage model just like there is always a security concern with online-based services.

1.1.3 Hybrid cloud storage

Hybrid cloud storage combines private and public cloud storage model (Buyya et al., 2011). With the hybrid cloud storage model, the organizations' data which are very critical and sensitive get hosted in the company's private cloud while the other data which are less susceptible are stored and accessed from a public cloud provider (Mell & Grance, 2010). This storage model is

of a huge advantage because it combines the scalability, fast deployment and low cost of public cloud storage with full control and security control associated with private cloud storage. Hybrid clouds have aspects of the other cloud model, and this is often to optimize these services. For hybrid cloud storage to work, it has to meet specific requirements. It must act like homogenous storage; it has to be transparent and also be able to keep active data which needs to be regularly on-site while less accessed data gets moved to the public cloud (Mell & Grance, 2010).

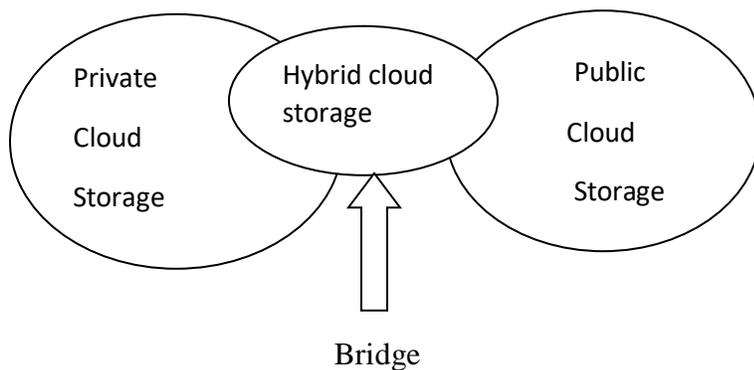


Figure 2: Cloud storage models

Having considered the various storage models available, the option to choose from by an enterprise will usually depend on its specific needs, the size of the enterprise and its budget.

1.1.4 Advantages of cloud storage

- Cloud storage helps in cost savings. When implemented correctly to meet business needs, organizations can significantly reduce annual operating expenses.
- Data is readily available. As long as there is an internet connection, stored files can be easily accessed from anywhere via Internet connection.
- Efficient disaster recovery. In the event of a disaster, businesses have to recover lost data. Cloud storage is very useful in this aspect as it can serve as a backup plan by businesses because a second copy of important files which got damaged or lost can be easily retrieved. Since with cloud storage, the data is stored at a remote location, and they are easily accessed through an internet connection.
- Sharing of data with other users is straightforward and convenient as only a few mouse clicks will get the job done.

- Files across different devices are easily synced, and this helps users to obtain updated data quickly regardless of the device used (Coyne, Shivaramakrishnan, & Sing, 2014).

1.1.5 Disadvantages of cloud storage

- Bandwidth problems. Bandwidth issues can be a drawback as some cloud storage services have a predetermined bandwidth allowance which an organization can access and once surpassed, it incurs heavy charges. However, some service providers allow unlimited bandwidth.
- Without an internet connection, stored data cannot be accessed.
- With remote storage, there are security concerns for sensitive data stored (Dhamdhere, 2014).

1.2 Role of storage in data centers

Wu & Buyya (2015) describes the data center as a facility housing computing systems which are used in the collection, sharing, and management of large volume of data. The data center acts as a central location for information gathering and distribution to other data centers or end-users. The traditional data center has undergone tremendous changes as cloud computing continues to spread.

Storage systems serve as the core of any data center, whether small, medium or large (Wu & Buyya, 2015). When storage systems are implemented correctly in data centers, they help the business to grow and thrive. However, if executed poorly, such companies are at a high risk of been destabilized. With the widespread of cloud computing, companies not yet using the cloud have been virtualizing and considering migration to the cloud. These goals are accomplished through storage systems.

The common storage technologies in use by most data centers are the Network Attached Storage (NAS) and the Storage Attached Network (SAN). Recently, other storage technologies such as Object Storage (OS) are utilized by data centers (Lee, 2014). Several trends such as increased number of users, a lot more data, and IT consumerization have had a tremendous impact on storage model. As earlier defined that the data center serves as the central point for information gathering and distribution to either other data centers or end-users, it is at the epicenter of cloud storage and as such it has evolved to support more agile, and scalable storage platforms. As

today's data center infrastructure is tasked with supporting lots of applications, users and workload, storage has evolved tremendously to support these business demands efficiently.

1.3 History of storage and fundamentals

From the very first developed computers to our current day handheld devices, all through the history of computers, storage has always been a major factor to consider. The oldest form of data storage, which is the punch card, was developed by Basile Bouchon in 1725 (Reilly, 2003). In the early days of storage, punched cards were the traditional media for information storage (Goda & Kitsuregawa, 2012). With the invention of early computers, the punch cards got replaced with magnetic tapes although punch cards remained very much in use (Leggett, 2014).

The magnetic tape was invented by Fritz Pfleumer in 1928, but the quality was generally poor (Leggett, 2014). This was eventually improved. The magnetic tape was used for sound recordings until the arrival of CD. The magnetic tape was originally used for audio recording which is accessed sequentially (Goda & Kitsuregawa, 2012). To access any part of the tape will need rewinding and this often cause major delays, and this is a major disadvantage. In contrast, magnetic disks were developed to overcome this drawback; it permitted random accesses with little delay (Goda & Kitsuregawa, 2012). The magnetic tape got replaced with magnetic. Magnetic disk development began with the IBM 350 Disk File (Noyes & Dickinson, 2000). As magnetic disks continue to evolve, there is a more drive to increase capacity and its reliability.

Hard drives were introduced in the 1950s, with the first hard drive unveiled by IBM in 1956 (Leggett, 2014). These days, hard drives are much smaller, have better performance, can store more data and are way cheaper. Continuous development of storage technologies led to the invention of the compact disk, DVD in the 1990s, flash drive, blue ray and now the IT world is on to storage in the cloud (Goda & Kitsuregawa, 2012). As Internet bandwidth continue to increase, storage in the cloud is rapidly evolving. Just as technology has continued to evolve, methods of data storage will continue to evolve as well, this rapid development is particularly in the aspect of performance, its speed and size of the storage medium.

1.4 Storage hierarchy

Storage is one of the essential components of any cloud infrastructure. Computer storage is made up of primary and secondary storage (Wu & Buyya, 2015). The secondary storage is at the lower level, and it can also be referred to as remote storage; it is different from the primary storage as it

cannot be directly accessed by the CPU (Wu & Buyya, 2015). Secondary storage provides high capacity, but the data may not be immediately available. Computer storage hardware can be classified into five categories according to their ranking with off-line storage occupying the bottom of the hierarchy structure and static random access memory at the top (Hoskins, 2005).

- Static Random Access Memory (SRAM)
- RAM (dynamic RAM or DRAM)
- Solid State Disk (SSD)
- Hard Disk Drive (HDD)
- Tape drive/optical storage (Wu & Buyya, 2015)

These five categories are different types of storage components with different technologies, installed in various locations on a server. They are ranked with respect to their size, cost and speed (Hoskins, 2005). Tape drive also known as off-line storage are the least expensive but they are also the slowest and are huge in size. The Static Random Access Memory (SRAM) are at the top of the hierarchy. They are the most expensive and fastest but their size is considerably small.

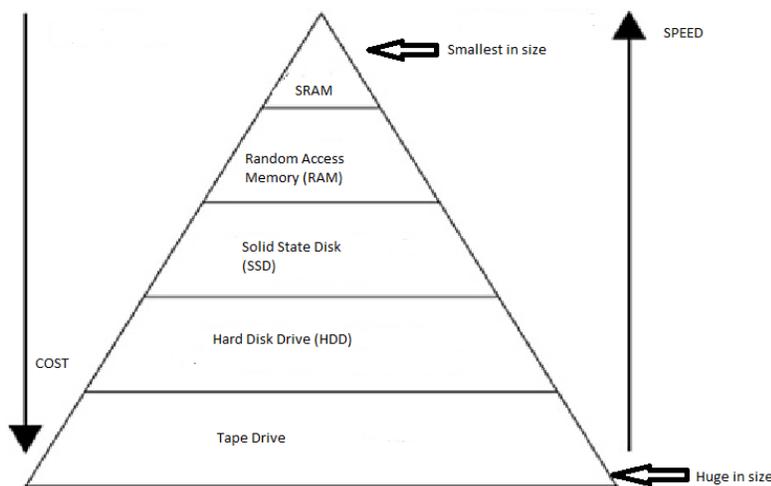


Figure 3: Storage Hierarchy

1.5 Storage SLA and RAID architecture

1.5.1 Storage SLA

A Service Level Agreement refers to a formal contract between a service provider and the customer which specifies the services been delivered and all attached conditions (Wieder, 2011). The SLA can also be referred to as a formal contract binding both the service providers and the customer. Elements that are defined in the SLAs includes the parties involved, the cost involved, the contract date and the terms of the agreement between both parties (Hoskins, 2005). It is crucial to understand the SLA, so as to ensure it meets the customer's needs and in some cases, storage providers may be willing to make adjustments to meet the customer's business needs.

Wieder (2011) described different SLA parameters for services that can be offered in a cloud computing environment, one of them been storage as a service. These parameters are utilized by the customer to negotiate with storage providers when seeking storage services. A brief description of these parameters is given in the table below.

Parameter	Description
Storage space	Units of data storage
Security	Description of security issues, especially authorization and authentication.
Storage billing	Details of storage cost is described
Scalability	This parameter specifies the ability to increase storage space as needed or decrease it if less capacity is needed.
Backup	Describes the location and how images of data are stored.
Privacy	This parameter defines how the data is transferred for maximum security.
Recovery	Measure in place in the event of any disaster.
Geographic location	Location where the data is stored.
System throughput	This refers to the amount of data that can be retrieved in a unit of time
Transferring bandwidth	This the communication channel's capacity.

Table 1: SLA parameters for storage as a service

(Wieder, 2011)

Since SLA is the performance metric that both service providers and consumers have agreed upon for service availability within the specified period, it will help both parties to visualize the service cost and consequences if the service provider fails to meet the agreed SLA.

In (Wu & Buyya, 2015), the authors gave a formula utilized in measuring service availability;

$$Availability = \frac{Uptime}{Uptime+Downtime} \times 100\%$$

A well-defined SLA will provide precise levels of service, for example, 99.5% uptime and in the event if the service provider cannot adequately meet up, compensation for the discrepancy is usually specified.

1.5.2 RAID Architecture

RAID is an acronym for Redundant Array of Independent/Inexpensive Disks (Poelker & Nikitin, 2009). RAID architecture is a way of combining groups of HDDs to prevent HDD failure (May, 2001). HDD failure is prevented because of the presence of redundancy which is implemented by the RAID architecture. If one disk fails, information is retrieved from the other disk. With RAID architecture, storage systems perform much better when compared to the performance from individual disk drives. This grouping of physical drives together to form a bigger drive is called a RAID set (Paulsen, 2011). The RAID set groups smaller physical drives together and presents it to the server as one logical disk (Poelker & Nikitin, 2009). The logical disk is a logical unit number (LUN). The concept of RAID architecture is that data is spread consistently across each drive in the array. The data is first broken into a consistent size and written to the drive based on the RAID level used (May, 2001). At the time the data is to be read, the process is turned around, and hence it will appear as if multiple drives are one large drive (Poelker & Nikitin, 2009). From this redundancy implementation, there are two significant advantages of implementing the RAID architecture; higher performance and availability. The high performance is as a result of the server's ability to read and write from multiple drives. The increase in availability is attributed to the RAID controller usage of parity information to regenerate any lost data, and this parity information is the checksum calculated from the original data (Poelker &

Nikitin, 2009). This technology is particularly beneficial in the storage of a large amount of data. Other benefits of RAID architecture include enhanced speed, increased storage capacity and lower risk of disk failure.

Wu & Buyya, (2015) described two types of RAID implementation. Software RAID implementation and hardware RAID implementation. In comparison to hardware RAID implementation, software RAID implementation is cheaper to implement as expensive disk controller cards are not required. In contrast, hardware RAID has a dedicated hardware controller to manage RAID configurations. Wu & Buyya (2015) further stated that data loss is prevented in Hardware based RAID by using an intelligent drive controller and this, in turn, improves the performance of read/write operations. Hardware RAID provides excellent performance for cloud storage. Hardware RAID consists of four building blocks: storage host server, RAID controller card, logical array (logical volumes) and hard disk array (Wu & Buyya, 2015). The various RAID configurations are described below:

RAID 0: Also known as disk striping. RAID 0 is the particular case of RAID, and it is often not considered a RAID because it doesn't have any redundancy (Wu & Buyya, 2015). In this RAID configuration, HDD failure will lead to data loss or the destruction of the entire disk array (Jang, 2003). RAID 0 has high performance because the capacity utilization is 100%. For better performance, RAID 0 is highly recommended, but it is not so good for high availability because basically, there is no parity generated for this RAID type. The RAID 0 configuration poses the problem of single point of failure (Jang, 2003).

RAID 1: also called disk mirroring as it adopts the mirroring technique to create HDD redundancy, giving 100% redundancy (Jang, 2003). All the data gets recorded twice, thereby giving a usable capacity of only 50%. The disks are exact images of each other. This means when one disk fails; then data is retrieved from the other disk. Write operations are slow with this RAID type since data is written twice. RAID 1 will need a minimum of two physical disks (Wu & Buyya, 2015).

RAID 3: RAID 3 stores parity information usually computed by the controller in the RAID (Wu & Buyya, 2015). This generated parity is not stored on the actual disk but a separate disk. This RAID type is suited for a large amount of data transfer but performs poorly when used for several small requests and RAID 3 needs at least three physical disks (May, 2001).

RAID 4: This RAID configuration uses a parity disk dedicated for that purpose, and it performs block-level striping across multiple disks (Wu & Buyya, 2015). It is well suited when accessing sequential data; it is not currently used very often.

RAID 5: It utilizes disk striping with parity, this RAID type stripes both the data and the parity information across all the disk (May, 2001). If the disk fails, the parity information is used to reconstruct the lost data (Poelker & Nikitin, 2009). This RAID type presents an excellent balance between performance and availability, and this has made it a standard RAID type in use today. RAID 5 eliminates the write issues encountered in RAID 4. RAID five needs at least three physical disks (Wu & Buyya, 2015).

RAID 6: The RAID 6 type uses two parity stripes to increase reliability (May, 2001). The idea of two parity stripes is to allow for two disk failures before losing data (Wu & Buyya, 2015). So with this RAID configuration, if there are "N" number of HDDs in the disk array, the usable capacity of the HDDs will be $N-2$ (Wu & Buyya, 2015). For example, if "N" is equal to 10, the usable space is 80% of raw capacity. RAID 6 is beneficial for data archiving or backup.

RAID 10: also called RAID 1+0, is a combination of disk mirroring and disk striping (May, 2001). RAID 0 and RAID 1 combined gives this RAID type. If the striping HDDs are combined first, it is nested as RAID 10, but if the mirroring HDDs are combined first, it will become RAID 01 (Wu & Buyya, 2015). Lots of RAID 10 configurations have the data mirrored first and then striped. RAID 10 requires a minimum of four physical disks (Wu & Buyya, 2015).

1.6 Storage Logical Unit Number (Storage LUN)

Disk drive serves as the basis for data storage. However, operating systems are unable to get access directly to the physical disk storage, and as a result of this, physical disk storage must be converted into a logical space which the operating system can access. This logical space is also known as a logical unit, and it is identified by the Logical Unit Number (LUN) (Poelker & Nikitin, 2009). According to Long (2006), a logical unit number is a number used to identify logical unit so they can be easily accessed. LUNs refer to specific storage in a storage attached network (SAN). LUNs are responsible for deciding which hosts will access which storage, and this makes them very crucial to the management and efficiency of storage systems. What a LUN does is that it partitions a group of HDDs that form a redundant array of an independent disk (RAID) set (May, 2001). LUNs will represent the entire RAID set capacity by dividing the large

volume of a RAID set into some smaller slices where each LUN is mapped to an individual host server (Wu & Buyya, 2015) Sometimes, the problem of insufficient LUN storage capacity for a particular business application is encountered, and this issue can be resolved by joining other LUNs to increase the storage capacity, a process referred to as concatenation (Wu & Buyya, 2015).

Wu & Buyya (2015) noted that concatenation serves as a quick solution to increase LUN storage capacity, but it will not provide any benefits for LUN performance. Apart from capacity increment, performance must also be increased and as such; there is the need to adopt Meta-LUN striping (Wu & Buyya, 2015). The more HDDs there is, the higher the performance will be. With storage LUN, the issue of security is critical. It is important to ensure that users and servers have authorization. This security measure is achieved through LUN masking (Wu & Buyya, 2015). The function of LUN masking is to prevent unauthorized hosts that are trying to access an LUN, which has already been mapped to another host. LUN masking defines which host can access which LUN in a shared storage environment. An example of this is seen in a case where there are three LUNs mapped to three lines of business, namely; sales, production and finance. If LUN masking is not implemented, each line of business can access all LUNs (Wu & Buyya, 2015). LUN masking ensures each line of business can only access its own LUN.

CHAPTER TWO: STORAGE TOPOLOGIES

In today's digitalized world, storage solutions pose as a primary concern in the IT world as the reliability, availability and security of data are essential to the business success. Different storage topologies are available today that support the requirements of a small, medium and enterprise level businesses. The three main storage methods are:

- Direct Attached Storage (DAS)
- Storage Area Network (SAN)
- Network Attached Storage (NAS)

2.1 Direct Attached Storage (DAS)

Direct Attached Storage (DAS) is the direct connection of the servers to the storage (Sosinsky, 2009). This direct link of the servers to the storage is done without a storage network or any network device (Wu & Buyya, 2015). However, a lack of storage network does not mean there is no interface connection. With DAS, different types of interfaces are connected to the server, such as SATA, IDE/ATA, fiber channel, etc.

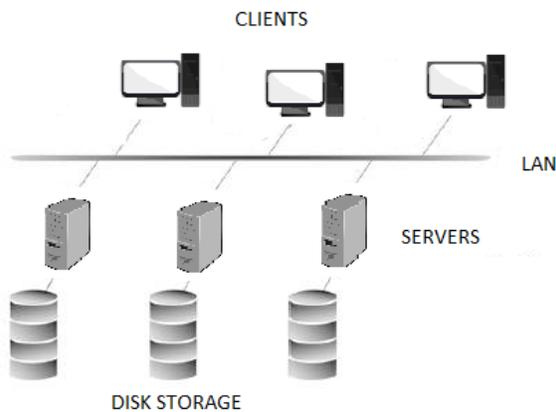


Figure 4: Direct Attached Storage

The Direct Attached Storage is the simplest. The DAS been the basic storage model acts as a building block for the Network Attached Storage and the Storage Area Network (Marshall,

Beaver, & McCarty, 2009). Therefore, the performance of a NAS or SAN will be determined by the DAS.

DAS systems provide the flexibility and scalability needed in an enterprise. It is a simple yet powerful storage solutions which are ideal for applications requiring an efficient yet simple array to maintain availability. A DAS controller allows a maximum of four servers to access the same logic storage unit (Marshall, Beaver, & McCarty, 2009).

DAS is easy to understand and install, and it also presents reduced cost in comparison to SAN and NAS. Given its simplicity, it should also be noted that growth potential is very limited and beyond a certain point, it could become costly and difficult to administer (Kitta, 2011). So it will be important to note that the Direct Attached Storage is used where data sharing is limited to particular departments and where centralized management is not an issue.

The storage disk in a direct attached storage is usually managed by the operating system at the client side (Marshall, Beaver, & McCarty, 2009). As described by Kitta (2011), DAS configuration can be done in two ways, either the disk is directly installed on the server, or it is installed in a separate enclosure which is the attached to the server.

2.1.1 Advantages of Direct Attached Storage (DAS)

- Simpler to setup and configure when compared to NAS and SAN which makes it more suitable for small businesses.
- Low cost to set up.
- With DAS, networks are not needed. So it doesn't require IP addresses
- It is faster and has a better latency in comparison to NAS and SAN (Marshall, Beaver, & McCarty, 2009) (Kitta, 2011)

2.1.2 Disadvantages of Direct Attached Storage (DAS)

- DAS has poor capacity utilization. Capacity is limited by the physical structure of the server and external enclosures.
- It has limited availability since storage access is directly connected to the server. Data accessibility can be impossible if the server is down.
- DAS is faced with scalability issues.

- Each server requires its own backup solution (Marshall, Beaver, & McCarty, 2009).

2.2 Storage Area Network (SAN)

The Storage Networking Industry Association (SNIA) defines SAN as a network which allows communication between storage devices and computer systems and also storage devices with other storage devices; it is a network made up of interconnected computers network and storage devices (Barker & Massiglia, 2002). Tate (2016) indicated that storage area network takes away storage resources from the shared network (the user network) and re-organizes them to an independent network with higher performance. This enables each server to obtain access to shared storage as if the storage was attached directly to the server. To access a storage device, on an SAN, a host will send out a block-based request, and it indicates storage device which is accepted by the host's HBA (Marshall, Beaver, & McCarty, 2009). SAN can be described as a high-speed network connecting storage devices and servers together. SAN permits any-to-any connection in the network, as it uses interconnect elements such routers, switches, etc., (Tate, 2016). With the SAN network, the traditional dedicated connection from the storage device to the server is eliminated, and also, the concept of the server owning and managing the storage devices is also eliminated. SAN uses Fibre Channel or Ethernet to provide a connection between storage and host (Tate, 2016). A SAN network can be put together using three principle components: fabric, host bus adapters (HBAs) and switches (Marshall et al., 2009).

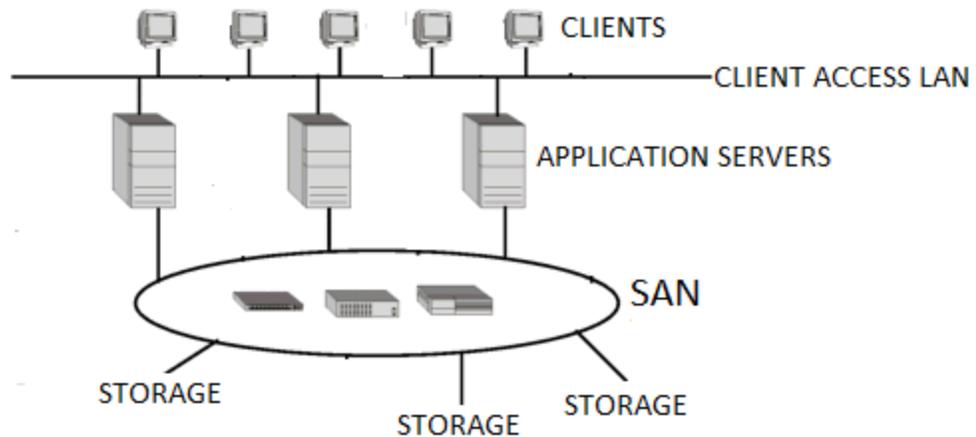


Figure 5: Storage Area Network

In the earlier discussion of direct attached storage (DAS), the pros and cons were noted. Some of the pros are that DAS storage is accessed over very short, dedicated, low latency, contention free interconnects (SCSI bus). This guarantees quick and reliable access to devices. However, some of the cons of DAS include increased management, low utilization capacity islands and a limited number of devices such as disk drives. The idea behind SAN was to keep the pros but overcome the cons of DAS storage. As such, SAN technology provides a high-speed, low-latency storage network that is relatively contention free. SANs also get rid of capacity islands by allowing large numbers of drives to be pooled and shared between multiple hosts (Tate, 2016). They also simplify management by reducing the number of management points.

2.2.1 Advantages of Storage Area Network

- SAN is highly scalable. It can grow as the required space grows. It is not physically limited.
- High performance is achieved as it is on its network and is not affected by Ethernet or local disk throughput bottlenecks. SAN can be shared by all since it is not directly attached to a network or server.
- Long distance connectivity is possible with SAN. This enables the setup of storage in a remote location.
- SAN is versatile

- SAN enables centralized management from a single interface (Barker & Massiglia, 2002).

2.2.2 Disadvantages of Storage Area Network

- SAN is very expensive to deploy and maintain.
- Managing a SAN system can be quite cumbersome at times (Vacca, 2013).

2.2.3 Storage Area Network (SAN) Components

A SAN is made up of one or several servers that are attached to a storage array, and this connection is achieved by using SAN (Vacca, 2013). SAN comprises of three components

- Host components
- Fabric components
- Storage components

Host Components

The host part of a SAN is made up of servers and all other elements that allow physical connection between the server and the SAN itself (Vacca, 2013). These components are:

- Host Bus Adapters (HBAs): The HBAs are located inside the servers. The host is connected to the fabric through its HBAs (Vacca, 2013). The HBAs are used to connect the servers to the fabric part of a SAN, and HBA permits the server's operating system to communicate with the external storage arrays in the SAN (Poelker & Nikitin, 2009).
- Host Bus Adapter Drivers: The HBA drivers run on the server, and its function is to enable the operating system of the server to have communication with the HBA and as a result communication is ensured with the devices in the SAN (Poelker & Nikitin, 2009).

Fabric Components

The fabric part of the SAN is the actual network part of the SAN (Vacca, 2013). All the fabric components of the SAN are responsible for moving data from a source component, for example, the HBA of a server to a receiving component, for example, a port on a storage device (Poelker & Nikitin, 2009).

The Fabric components include:

- **SAN Switches:** these switches connect different elements of the SAN network (Vacca, 2013). For example, the connection of the hosts to the storage arrays. SAN switches serve as the connection point for the SAN fabric connecting storage devices, servers and also, other switches. In the event of a path failure, SAN switches allow for path redundancy if this happens. The type of SAN switch, the kind of design features it has, and also the capacity of its port all have an effect on the overall performance, how much fault tolerance it possesses and its overall capacity (Vacca, 2013). The fabric topology is determined by the number, types and manner of switch connection (Vacca, 2013). The standard SAN switches in smaller SAN connection are usually called modular switches.
- **Data Routers:** Data routers are usually set up between Small Computer System Interface (SCSI) devices and Fiber Channel (FC) devices in the SAN (Tate, 2016). Servers in the network (SAN) normally have access to the SCSI disk in the SAN network using the data routers (Poelker & Nikitin, 2009).
- **Cables:** SAN cables are the physical links that connect all of SAN together. SAN cables are special fiber optic cables (Poelker & Nikitin, 2009). The maximum distance between SAN components is usually determined by the type of SAN cable used in a SAN network. This distance will contribute to the total bandwidth rating of the SAN network itself. SAN uses any of three types of fiber optic cables where the difference between them is usually the type of glass fiber core used in making them (Poelker & Nikitin, 2009).
- **Communications Protocol:** Communication among fabric is achieved through the use of the Fiber Channel (FC) protocol. Many of the deployed storage area networks use fiber channel as the preferred communication protocol (Tate, 2016). This protocol uses serial I/O bus cable with very high speed for data transmission between two ports, and fiber channel supports point-to-point topology, arbitrated loop topology, and switched fabric topology. Switched fabric topology is the basis for most of today's SAN (Vacca, 2013).

Storage Components

The storage components refer to the storage arrays in a SAN (Vacca, 2013). Storage arrays are made up of the storage processors which are responsible for communicating with the disk array and also provide the RAID functionality as well as LUN (EMC World, 2013). RAID and LUN were discussed in detail in the previous chapter.

- Storage processors: They are the mechanism for data communication (EMC World, 2013)
- Storage devices: The actual data is stored in disk arrays or tape devices and in some cases on both (Shore, 1988). Disk arrays are multiple disks devices grouped together. This grouping is done through RAID technology, which in turn increases performance but also redundancy. The disk drives are always inside intelligent storage arrays in a SAN (Poelker & Nikitin, 2009).
- Tape storage devices: Tape storage devices also form part of the storage component of SAN. Tape storage possesses large capacity for data storage which is a major advantage and this large capacity for data storage makes it cheaper when compared to the cost of hard disk storage (Henry-Stocker & Marks, 2000). However, it must be noted that tape storage stores data in sequence and accessing a specific data can be time-consuming because the tape will be rolled back each time (Stair, & Reynolds, 2008).

2.2.4 Storage Area Network Connectivity

SAN connectivity is made up of components, both hardware and software that connect storage devices and servers. In discussing connectivity in SAN, we look at the role of Fibre Channel (FC), Ethernet and Internet Small Computer System Interface on the connectivity of a storage area network. Fibre Channel is the primary architecture for most SANs in use (Poelker & Nikitin, 2009). These protocols are discussed extensively in the next chapter.

2.3 Network Attached Storage (NAS)

Network-attached storage (NAS) is a storage device which is connected to a network, and it is designed to enable users on the network with proper authorization to store and retrieve data which are remotely stored (Snedaker, 2007). As the name suggests, it is the storage attached to the network. This network refers to the typical Ethernet network. A NAS often integrates an

optimized operating system dedicated to file sharing (Snedaker, 2007). NAS devices are easy to implement, and it offers storage and file sharing over a standard Ethernet network (Ganger & Femling, 2008). NAS devices are quite scalable, which means that as additional storage is required, the available storage can be easily extended for more.

When comparing NAS to SAN, NAS does not require specific host adapter to connect to the network since it uses standard network protocols. When considering small businesses, several considerations such as, requiring storage with lower cost, ease of operation because small-scale businesses do not have dedicated IT staffs, and the need for a storage solution which is capable of growing alongside their storage needs makes the NAS a best fit for small businesses. NAS devices are rapidly becoming popular in many industries as an effective, and low-cost solution for storage needs.

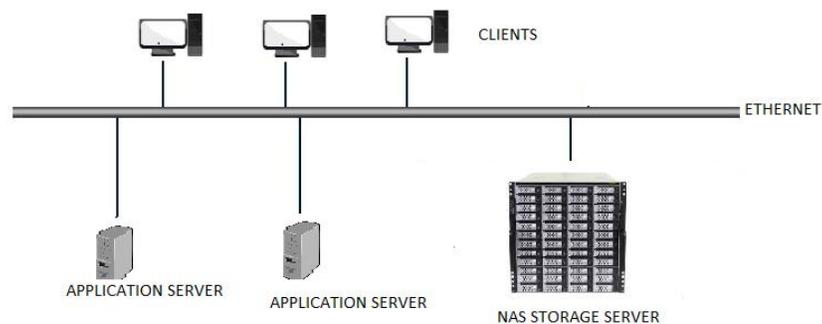


Figure 6: Network Attached Storage

2.3.1 NAS Protocol

When considering communication in NAS, clients use higher level protocols, usually the application protocol in the OSI model. There are two protocol commonly used with NAS; these are NFS and CIFS/SMB (Wu & Buyya, 2015). Both of these protocols operate using the client/server model just TCP/IP protocol (Kovalick, 2013).

NFS (Network File System)

The NFS was first developed by Sun Microsystems in 1980 (Kovalick, 2013). NAS uses the NFS to share files with authorized users over a network (Wahl & Pantol, 2014). Initially, NFS was developed for the UNIX environment, but progressive works later expanded to include non-

UNIX systems; however, most NFS clients are UNIX-based (Wu & Buyya, 2015). Two versions (v2 and v3) are popularly used.

CIFS (Common Internet File System)

The CIFS also known as Server Message Block (SMB) was originally developed by IBM and Microsoft (Arzensek, 2014). It was designed to support file sharing in DOS and CIFS implements file sharing support with file locking mechanisms (Smith, Hertel, & Iverson, 2004). Locking is implemented to permit local buffering and exclusive access. CIFS also supports Unicode file naming (Javvin, 2007).

HTTP - Hyper Text Transfer Protocol

Hypertext Transfer Protocol (HTTP) is supported by many NAS systems and clients can easily download files from web browser provided the NAS system supports HTTP (Kovalick, 2013).

There are other protocols from other vendors competing with the NFS and CIFS. These include Netware communication protocol from Novell and AppleTalk from Apple.

2.3.2 Advantages of Network Attached Storage

- Reduced cost as NAS devices utilizes the available IP networks for its connection
- It provides comprehensive data access and file sharing
- File locking provides data security.
- Flexibility of storage functionalities
- There is increased efficiency with centralized storage
- Ease of scalability of storage capacity
- Simplified management
- There is the flexibility to support UNIX and Windows clients (Snedaker, 2007)

2.3.3 Disadvantages of Network Attached Storage (NAS)

- The performance of NAS is dependent on the available bandwidth as it shares the network with other computing counterparts.
- Again, been in a shared environment, it may not offer guaranteed services for critical operations unlike in SAN.

- Without properly quota allocation, a few users may take over all storage space since it is a shared storage.
- Data transfer is also inefficient as NAS is implemented over TCP/IP protocol unlike in SAN, which has a protocol specifically for data transfer (Marshall et al., 2009).

CHAPTER THREE: STORAGE PROTOCOLS

A storage protocol is the communication channel between the storage devices and storage controllers and this storage protocol defines the parameters for communication (Wu & Buyya, 2015). File-based storage and block-based storage are the two most popular storage systems (Wu & Buyya, 2015). Network Attached Storage (NAS) systems deploy file-based storage while block-based storage is popular with SAN. There are two storage protocols; file-oriented protocol and block-oriented protocol

3.1 File Level Storage

This is the storage technology deployed in NAS systems. Hard drives also use file-based storage. The file-level storage uses protocols such as NFS and SMB/CIFS (Stern et al., 2001).

3.1.1 Features of file level storage

- It is cheaper than block-based storage
- Widely used by Network attached storage systems.
- Simple to use (StoneFly, 2016).

3.2 File-Oriented Protocol

The file-oriented protocol is the protocol utilized by a client to access a file server or to access a network attached storage (Wu & Buyya, 2015). A file-oriented protocol works by reading and writing variable-length files and segments into blocks before they are stored on a disk or a tape; and the the two common file-oriented protocols implemented in NAS are NFS and CIFS (Stern et al., 2001).

3.2.1 Network File System (NFS)

Sun Microsystems initially developed Network File System (NFS) in the early eighties (Langbehn, 2005). The initial intention behind the development of NFS was to eliminate the distinction between a local and a remote file (Kozierok, 2005). After the required setup is performed, a file hosted on a remote computer can be used conveniently as if it were on a hard disk on the user's local machine. Network File System (NFS) when deployed in NAS, it enables the NAS system to share files and directories over a network with authorized users. NFS was first developed for the UNIX environment, but progressive works later expanded to include non-

UNIX systems; however, most NFS clients available are UNIX-based (Langbehn, 2005). NFS is available in versions two, three and four. Version one did not exist; it was only a prototype. NFS (v2 and v3) are popularly used. These two versions are popular in the IT industry and used as the standard protocol for file sharing under the UNIX environment. When considering both NFSv2 and NFSv3, there are three major differences between them:

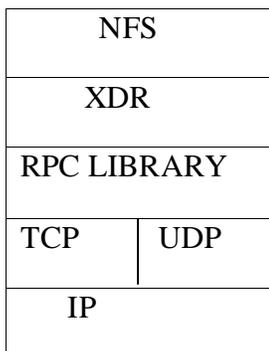
- File support size: A maximum length of four gigabytes is supported by version 2 and the actual implementations in this version is only limited to about two gigabytes file size while version 3 can support up to 264 -1 bytes in length (Stern, Eisler, & Labiaga, 2001).
- File transfer size: NFSv2 is limited to only 8192 bytes for each NFS read/write request (Stern et al., 2001). With NFS Version 3, the client and server can negotiate a limit acceptable to both (Stern et al., 2001). The problem posed with version 2 specifying a limit is that when a file larger than the limit is transferred, it needs to be fragmented and later reassembled. NFSv3 eliminates this fragmentation and reassembly problem and challenges that go with it.
- Throughput: With the NFSv2, the server needs to wait for data to be committed to stable storage (i.e. magnetic disk) before it can send a success reply to a write request (Langbehn, 2005). This will have a negative impact on the write throughput of the NFS clients. With the NFSv3, the clients can state that the write does not need to be committed to stable storage and this, in turn, enables the NFS server to respond quickly to such write request (Stern et al., 2001). However, the NFSv3 has a procedure called commit. The commit procedure also called procedure 21 indicates to the NFS server that before a success reply is returned, it needs to write uncommitted data to stable storage (Callaghan, 1999). This greatly helps to increase throughput because the NFS server can commit data to stable storage and at the same time the client can also be performing some other functions before the client sends the commit procedure.

The NFS, similarly to the TCP/IP uses a client/server model of operation (Kozierok, 2005). A computer may act as a server or as a client, and it may even act as both. This model of operation

involves the computer sharing some of its resources and accessing resources from other computers.

The architecture used by NFS consists of three main components. These are the External Data Representation (XDR), the Remote Procedure Call (RPC), and a set of NFS procedures and operations (Kozierok, 2005). The XDR standard defines the way data is represented when they are exchanged between clients and servers (Kozierok, 2005). The RPC protocol is used to call procedures on remote machines. All NFS operations are implemented as RPC procedures (Dong, 2007). A summary of NFS procedures and their function is shown in Table 4

The implementation of Transmission Control Protocol (TCP) over IP can be used by all the various versions of NFS (Kozierok, 2005). NFSv2 and NFSv3 can also UDP over IP but NFSv4 implements TCP over IP only (Kozierok, 2005). When NFSv2 or NFSv3 implements UDP over IP rather than TCP over IP, there is less protocol overhead. A reduced protocol overhead will provide better performance. Of course, this is only when the network is non-congested. When congestion occurs, for example, the server goes down, and the UDP clients continue to send requests to the server leading to network saturation, or if a frame is lost, the entire RPC request will have to be retransmitted. These problems are not present when TCP over IP is implemented, so TCP should be used with NFS server. TCP is a stateful protocol while NFS is a stateless protocol (Stern et al., 2001). Then how is TCP which is a stateful protocol a preferred choice when connecting to a stateless NFS? Though at first it sounds contradictory, a careful observation of the NFS protocol stack shows that RPC is in between NFS and TCP as shown below (Stern et al., 2001). In this model, RPC is responsible for hiding issues of TCP from NFS. In the event of a server crashing, the RPC layer simply re-establishes the connection, and the client is unaware it crashed (Stern et al., 20019).



DATALINK
PHYSICAL

Table 2: NFS protocol stack (Stern et al., 20019).

Procedures	Functions
Procedures 0	NULL - Do nothing
Procedures 1	GETATTR - Get file attributes
Procedures 2	SETATTR - Set file attributes
Procedures 3	LOOKUP - Lookup filename
Procedures 4	ACCESS - Check Access Permission
Procedures 5	READLINK - Read from symbolic link
Procedures 6	READ - Read From file
Procedures 7	WRITE - Write to file
Procedures 8	CREATE - Create a file
Procedures 9	MKDIR - Create a directory
Procedures 10	SYMLINK - Create a symbolic link
Procedures 11	MKNOD - Create a special device
Procedures 12	REMOVE - Remove a File
Procedures 13	RMDIR - Remove a Directory
Procedures 14	RENAME - Rename a File or Directory
Procedures 15	LINK - Create Link to an object
Procedures 16	REaddir - Read From Directory
Procedures 17	REaddirPLUS - Extended read from directory
Procedures 18	FSSTAT - Get dynamic file system information
Procedures 19	FSINFO - Get static file system Information
Procedures 20	PATHCONF - Retrieve POSIX information
Procedures 21	COMMIT - Commit cached data on a server to stable storage

Table 3: Summary of NFS procedures

(Callaghan, 1999)

NFS RPC requests are usually sent one at a time by the NFS client to the server (Jia & Zhou, 2005). RPC call cannot be issued while another RPC call is ongoing, this is similar to what happens in system calls. RPC call in progress must be completed and also acknowledged before a client process issues another RPC (Stern et al., 2001). When an RPC request is made by the NFS client, a timeout period is usually set during which the server will have to complete and acknowledge the request (Jia & Zhou, 2005). If a request is lost in transit, the client continues to retransmit the request until it gets completed either with an acknowledgement or an error from the RPC.

3.2.2 Common Internet File System / Server Message Block (CIFS/SMB)

SMB is a network protocol implemented in windows for file-sharing (Vernooij, 2008). This protocol is usually seen in local area network. Various versions of SMB are called dialect and the Common Internet File System (CIFS) Protocol is a dialect of SMB (Hertel, 2003). However, both SMB and CIFS are mostly used interchangeably and also here in this report. Though popularly implemented in the windows environment, CIFS/SMB is also available on some other operating systems like UNIX.

SMB is known as the protocol behind the "Network Neighbourhood" and remote printing in Windows (Vernooij, 2008). This protocol allows clients to access remotely stored files locally (Eckstein, 2007). All operations i.e. read, write, rename, etc. are all supported, and everything appears to the client locally although files are remotely stored on a server. The primary function of the SMB is file sharing; however, it can also be used for browsing, authentication, and printer services (Eckstein, 2007).

The client sends basic requests such as read file to the server using CIFS protocol (Hertel, 2003). Once the request is received at the server end, the server will have to check both authentication and authorization and then go ahead to execute the request and forward a response packet to the client, and once the client receives the packet, it parses the received packet and then determines if request was successful or unsuccessful (Hertel, 2003).

Properties of SMB/CIFS Protocol

Protocol Dialects/Negotiation: Dialect refers to the various versions of the SMB protocol. The dialect is identified using a unique string and an example is "PC NETWORK PROGRAM 1.0" (Hertel, 2003). At the initial stage when a client needs to access a remote server, it sends its first packet referred to as a 'negotiate protocol packet', and specifies the list of dialects understandable to the client (Vernooij, 2008). The server sends a response packet back to the client indicating its preferred dialect for communication. This way, both the client and server can negotiate which to use.

Command batching: Another property is piggybacking other CIFS packets (Hertel, 2003). This helps in the reduction of response latency, hence giving an enhanced bandwidth utilization.

Client/server model: As noted earlier with file-oriented protocol, the CIFS architecture is based on a client sending a request, and a server replies to each request sent (Kozierok, 2005).

A typical CIFS request and response header is shown below. The various fields are also described below it.

0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0xFF								S								M								B							
Command								Error Class								0								Error Code							
Error Code								Flags								Flags2															
Pad or security signature																															
Tree ID (TID)																Process ID (PID)															
User ID (UID)																Multiplex ID (MID)															
WordCount								ParameterWords (size is determined by WordCount)																							
ByteCount																															
Buffer (size is determined by ByteCount)																															

Table 4: CIFS packet header

(CodeFX, 2001)

CIFS packet header fields described below:

- Header: Start of CIFS packet and 4-byte in length. Shown above, the first byte is 0xFF, the second, third and fourth bytes are the ASCII representations S, M, and B
- Command: Indicates CIFS packet type and it is one byte in size.
- Error Class: This field indicates if a request was success or not, a zero value indicates success
- Error code: This states the type of error that occurred.
- Flags: This is used to specify a particular option depending on the value indicated.
- Flags2: This 16-bit field specifies more options.
- Pad/security signature: set to zero.
- Tree ID (TID): This field states which resources the given CIFS packet is referring to
- Process ID (PID): This is the process ID of the process sending CIFS request to the client.
- User ID (UID): This is at the client side, and it states the ID of the user sending a CIFS request
- Multiplex ID (MID): This field is a 16-bit value used to uniquely identify each outstanding client requests. This field ensures response is sent to the appropriate request when there are several other requests waiting.
- WordCount: This field specifies the number of 16-bit words that the ParameterWord field will hold.
- ParameterWords: This is a variable length field, its size is determined by the WordCount field and it is used to hold command-specific data.
- ByteCount: Its function is similar to that of WordCount, it specifies the number of bytes of data the buffer field will hold.
- Buffer: The size of the buffer field is determined by the ByteCount field. This field holds variable length of data, usually large amount of raw data (CodeFX, 2001).

CodeFX (2001) stated that what makes the parameter data section differ from the buffer section is usually the type of data they each store. Whereas the parameter section holds a small number of packets, the data buffer section holds large amount of data, usually raw data (CodeFX, 2001).

3.3 Block level storage

This storage involves the creation raw data volumes as blocks which can be controlled as an individual hard drive by the server, and they can also be individually formatted (Long, 2006).

3.3.1 Features of block level storage

- Common in SAN systems.
- High efficiency in the transportation of data storage.
- Uses iSCSI and FCoE protocols for transfer of data and it also uses SCSI as communication interface between the initiator and target (StoneFly, 2016).

3.4 Block-oriented protocol

Block oriented protocol refers to block-level protocol which reads and writes individual fixed-length blocks of data (Long, 2006). Block-level storage is usually deployed by large enterprises using the storage area networks (SANs) or by businesses using storage systems that are large-scaled. Each block in a block-level storage system are managed by the server operating system and can be managed as an individual hard drive (StoneFly, 2016).. For storage blocks to be visible and accessible by the server's operating system, block level storage employs protocols like fiber channel, iSCSI and Fiber Channel over Ethernet (FCoE) (Long, 2006).

3.4.1 Internet Small Computer System Interface (iSCSI)

Internet Small Computer System Interface acts as a transport layer protocol used to enable data transfers over intranets (Rouse, 2014). It is also used to manage storage in long distances as stated previously in the deployment of SAN. One of the benefits of iSCSI is the reduction of cost as hardware purchase are not necessary, as such, cost is reduced significantly yielding high-performance SAN (Long, 2006). Packets are usually sent over the network through a point-to-point connection (Rouse, 2014). Received packets are disassembled, and the iSCSI separates the SCSI commands, this way the storage appears local to the operating system (Rouse, 2014). iSCSI is popular among small to midsized businesses. The iSCSI protocol is used to facilitate communication between a cluster of hosts nodes and the storage pool.

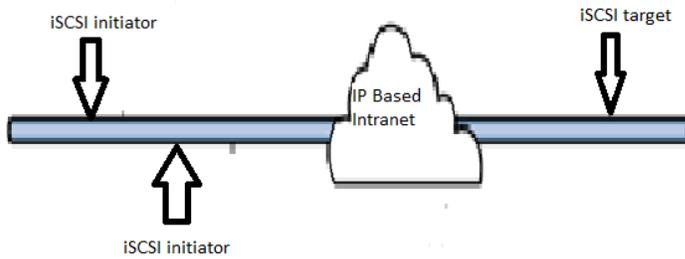


Figure 7: iSCSI Initiators and an iSCSI Target Connected via an IP-based Network

(Lee, 2014)

Benefits of iSCSI

- Simplified storage.
- It offers lower cost of operation
- iSCSI uses TCP/IP, so no expert knowledge required unlike fiber channel
- iSCSI utilizes Gigabit Ethernet network components, this in turn significantly simplifies the storage network environments
- Usage of Ethernet ensures reduced deployment cost (Lee, 2014).

3.4.2 Fiber Channel

Fibre Channel (FC) serves as an interface for connecting computers and storage systems. Predominantly used for today's SANs, the fibre Channel technology help to meet the requirements needed to meet the needs of complex SAN (Norman, 2001).

Fibre Channel is particularly beneficial in that it employs both the transport properties of an I/O bus with the distance characteristics and flexibility traditional network (Tate, 2016). There are three major fibre channel topologies: Point-To-Point, the Loop and the Fabric topology. The Point-To-Point Connection is the most basic type of connection which can be designed to be by itself or as a subsection in Fabric Loop (Norman, 2001). However, it should be noted that the point-to-point connection is limited in connectivity. In the Arbitrated Loop topology, all devices are in a loop or ring form just like a token ring network connection (Long, 2006). With this topology, removal of a single device will cause the entire loop to breakdown similar to what is experienced in a token ring network. Arbitrated Loops are easy to setup and manage and are relatively cheap as well. The third topology which is the Fabric topology, all devices are

connected to fibre channel switches and this fabric topology is taken to be the "ideal" Fibre Channel topology (Norman, 2001). This topology presents several advantages: Optimized connection is made possible due to the state of the fabric been managed by the switches (Marshall et al., 2009). Fabric topology enables multiple pairs of ports to be able to communicate at the same time in fabric, and failure in a single port does not affect other port as seen with the arbitrated loop topology.

Fibre Channel employs the usage of multilayer protocol architecture similar to what is seen with the Seven Layer OSI Model, but it has five layers (Norman, 2001). Fibre channel layers are: FC-0, FC-1, FC-2, FC-3, and FC-4 (Long, 2006). All five layers of the Fibre Channel layers corresponds to the physical, data link, network, and transport layer of the OSI layer (Norman, 2001). The figure below FC layers and OSI layers.

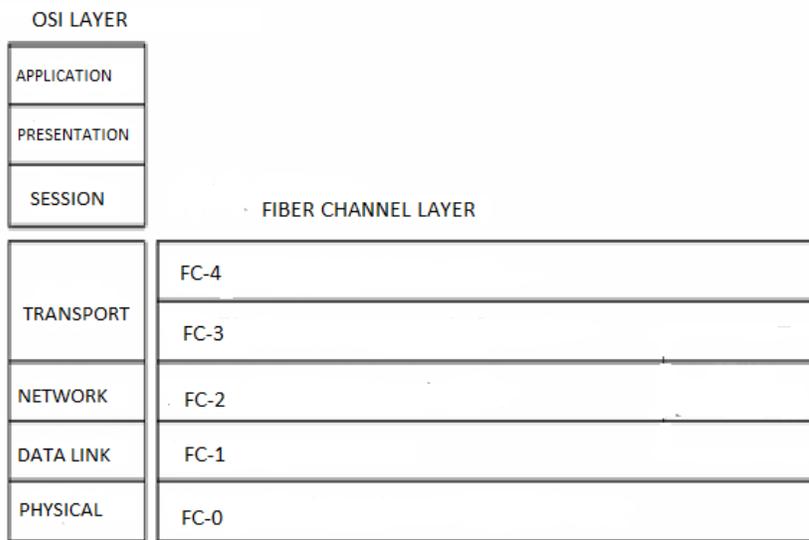


Figure 8: Fiber Channel layers vs OSI layers

FC-0 layer: at this layer, physical interface properties just like in OSI model are described. At the physical layer, the receiver gets a stream of bits which are sent over the media (Norman, 2001). The streams of bits are received at the output end.

FC-1 layer: The FC-1 level gives detail description of data encoding at the transmission end and its subsequent decoding at the receiver end (Norman, 2001). For proper encoding and decoding of transmitted data, the FC-1 level is in charge of transmission protocol, error control and any particular encoding and decoding rules that should govern transmitted data (Long, 2006).

Norman (2001) stated that this 8/10 bit encode/decode scheme implemented by the FC-1 layer helps to minimize errors that take place during encoding such as when a valid transmission character received was not the transmitted character (Norman, 2001). This makes this layer responsible for bits and frame synchronization.

FC-2 layer: is the transport core of fibre channel and this layer encapsulates user data and has it transmitted to the right receiver (Norman, 2001). This layer being involved with transport is a very complex layer. Its primary functions are data encapsulation, error and flow control and also class of service (Norman, 2001). The flow control implemented at this level ensures that buffers are not overrun. This flow control can either be an end to end flow control or buffer to buffer flow control (Norman, 2001).

FC-3 layer: it is used to support "common services"; such as multicast and hunt groups. (Norman, 2001).

FC-4 layer: Upper layer protocols are mapped to Fiber channel data structures in this (Tate, 2016) (Norman, 2001)

Fibre channel has evolved tremendously since its inception, and the aspect of speed and performance has been significantly improved.

3.4.3 Fiber Channel over Ethernet (FCoE)

Fibre Channel over Ethernet (FCoE) can be described as a storage protocol that allows communications from Fibre Channel to run directly over Ethernet ((Lemara & Gai, 2007). Hence the name FC over Ethernet. With this protocol, Ethernet becomes the Fibre Channel physical interface, and there are no fiber channel frames, just Ethernet frames. No separate protocol is required for I/O solution as fibre channel is used as the main storage protocol (Lemara & Gai, 2007). FCoE reduces switch complexity and also helps in the reduction of cable count and also reduced interface card count. Fibre Channel over Ethernet is a natural extension of the fibre channel.

FCoE enables the movement of fibre channel traffic over available Ethernet infrastructure. FCoE uses a lossless Ethernet fabric. (EMC World, 2013).

CHAPTER FOUR: COMPARISON OF STORAGE TOPOLOGIES AND TRENDS

4.1 Pros and Cons of various storage Topologies

The different storage topologies were discussed in detail in chapter two. The ultimate choice of what storage topology to follow is a choice to be made by an organization after careful consideration of the storage needs, organization's size, and funds available for storage deployment, management and maintenance. To also make this decision, the various benefits and drawbacks of the different topologies must be considered. These pros and cons of DAS, NAS and SAN, are highlighted in this section.

Advantages of Direct Attached Storage (DAS)

- Simpler to set up and configure when compared to NAS and SAN, which makes it more suitable for small businesses.
- Low cost to set up.
- With DAS, networks are not needed. So it doesn't require IP addresses
- It is faster and has a better latency in comparison to NAS and SAN (Kitta, 2011).

Disadvantages of Direct Attached Storage (DAS)

- DAS has poor capacity utilization. Capacity is limited by the physical structure of the server and external enclosures.
- It has limited availability since storage access is directly connected to the server. Data accessibility can be impossible if the server is down.
- DAS is faced with scalability issues.
- Each server requires its backup solution (Marshall, Beaver, & McCarty, 2009).

Advantages of Network Attached Storage

- Reduced cost as NAS devices utilizes the available IP networks for its connection
- It provides comprehensive data access and file sharing
- File locking gives some security
- Flexibility of storage functionalities
- There is increased efficiency with centralized storage

- Ease of scalability of storage capacity
- Simplified management
- There is the flexibility to support UNIX and Windows clients (Snedaker, 2007)

Disadvantages of Network Attached Storage (NAS)

- The performance of NAS is dependent on the available bandwidth as it shares the network with other computing counterparts.
- Again, been in a shared environment, it may not offer guaranteed services for critical operations unlike in SAN.
- Without properly quota allocation, a few users may take over all storage space since it is a shared storage.
- Data transfer is also inefficient as NAS is implemented over TCP/IP protocol unlike in SAN, which has a protocol for specific for data transfer (Marshall et al., 2009).

Advantages of Storage Area Network

- SAN is highly scalable. It can grow as the required space grows. It is not physically limited.
- High performance as it is on its network and is not affected by Ethernet or local disk throughput bottlenecks. SAN can be shared by all since it is not directly attached to a network or server.
- Long distance connectivity is possible with SAN. This enables the setup of storage in a remote location.
- SAN enables centralized management from a single interface.
- SAN is versatile (Barker & Massiglia, 2002).

Disadvantages of Storage Area Network

- SAN is very expensive to deploy and maintain.
- Managing a SAN system can be quite cumbersome at times (Vacca, 2013).

4.2 Traditional storage vs. cloud storage

In recent times, there has been rapid development in information technology. The industry is seeking for more reliability and scalability of their infrastructure. This has consequently led IT to become more cloud-based these days. Cloud storage is taking the IT industry by storm. While cloud storage continues to grow and expand as more and more organizations opt for this kind of storage, the traditional storage model is still used by some organizations. The choice of cloud storage or traditional storage by an organization will depend heavily on the organization's requirements for data storage, their size, cost involved and other considerations (Jiehui et al., 2011). To be able to make informed decisions regarding what kind of storage to opt for, it is necessary to know the difference between the two. What they entail and what they can offer to the users, the various limitations, etc.

As earlier discussed, cloud storage involves the storage of data in an off-site data centers managed by a cloud storage providers, where these cloud storage providers are third party companies. In traditional storage, data is usually stored in computers and servers which are directly connected to local computers via LAN or WAN. The data information is typically stored in disks (Shore, 1988). With the traditional storage, higher demand for storage space will require secondary backup devices and in some cases a third party website.

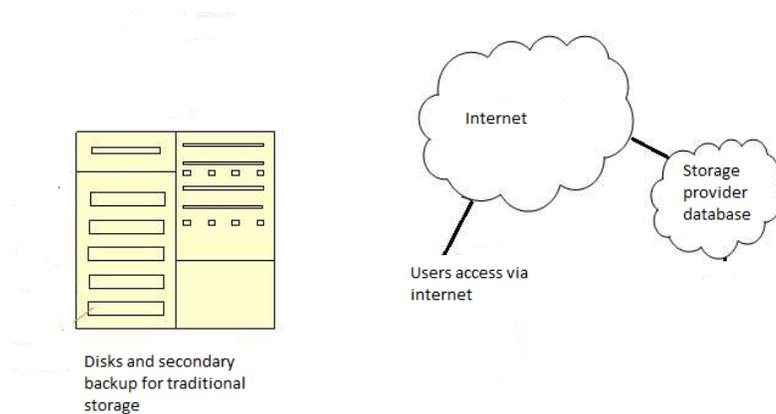


Figure 9: Traditional vs cloud storage

4.2.1 Comparing traditional storage with cloud-based storage

Data accessibility: One major advantage of cloud storage over traditional storage is that stored data can be accessed from anywhere in the world as long as there is an internet connection. With traditional storage, physical devices such as servers are required. For example, an organization that has branches in several cities in a country can have all branches access stored data via the internet whereas, this same organization if implementing traditional storage will require each of their branches to have their servers and there will also be a need for carefully designed systems to access shared data.

Scalability: Regarding scalability, cloud storage is a clear winner. Companies can store as much or as little data as they want from month to month. With the traditional storage model, scalability is a major limitation faced by organizations who choose to go with this model. There is the issue of the wrong estimation of the size of storage needed by a user. If more is needed, there is the need to purchase additional disk drives. Then what happens after a month when the user requires less storage, of course, the extra disk drives cannot be returned for a refund. Cloud storage solve this problem because you pay only for the capacity that you are using, and you can add more capacity as needed as it works by pay by the use of time, space or combination of both (Jiehui et al., 2011).

Cost: Implementation of cloud storage is more cost-effective in comparison to traditional storage. In traditional storage, implementing a file server can run into thousands of dollars (Lowe, 2013). For small to mid-sized business, it can be quite expensive to maintain even. For this kind of businesses, it can also be very cumbersome to manage a traditional in-house storage.

Security: The issue of security is subjective. While for some handing over their data to a third party raises concerns about security, for others, cloud storage can be viewed as more secure compared to traditional storage because with traditional storage, theft and intrusion are common. However, there have been arguments that traditional storage is more secured as there is no total privacy of stored data as long as your data is stored with third parties.

While comparing the cloud storage model to traditional storage, mostly the benefits of cloud storage was highlighted and, of course, this is why recent trends have seen more enterprises opting for storage in the cloud. However, cloud storage has some drawbacks which may lead

some enterprise to stick with traditional storage. It should be noted that the choice of what kind of storage to adopt will depend greatly on the needs and requirements of the enterprise.

Looking at some of the drawbacks of cloud storage, the issue of access problems comes right to mind. Having highlighted previously that one of the benefits of cloud storage lacking in traditional storage is anywhere and everywhere access as long as there is an internet connection, but what if the internet connection goes down either at client's side or that of the storage provider?. This presents a serious problem getting access quickly to data or worse still; one could be faced with latency which could delay data retrieval. Another drawback with the cloud storage is reliability. The issue of reliability can be argued from both sides but regarding the drawback with cloud storage, there could be a reliability problem. This is because there is a dependency on storage providers and if they are faced with any form of technical difficulties that gets them offline, then accessing data when needed becomes impossible. An example of this was when Amazon had a publicized outage in February 2008 (Miller, 2009)

4.3 New market trends and direction of storage technologies

Storage technology is undergoing rapid changes in recent years. The deployment, provisioning and management of storage technologies is experiencing massive changes in the last couple of years. This is 2016 and there are various trends in storage which are significantly affecting data centers. Discussed in this section are some of the recent trends going on with storage technologies.

Recently, one of the hottest trends in the storage industry is the evolution of flash as primary storage is rapidly turning to flash territory (Shoshani & Rotem, 2010). The use of flash storage is increasingly becoming more popular as several organizations are looking towards this option. This increase in demand for flash storage in comparison to traditional storage is its ability to handle data at a much faster rate, and this fast performance feature tends to meet the requirements for today's storage needs.

Also, another big trend in the storage industry is the emergence of object storage systems which stores data in a flat namespace with unique identifiers which allows data to be retrieved without the server having to know the actual data location (Minx, 2014). Options presented by object storage has been limited in previous years but recently, organizations are using this technology for long-term data retention and backup. Enterprises have been able to make use of object

storage and flash to improve on performance and capacity issues. Also, software-defined storage is also making wave as it is used as it provides a high level of scalability (Tate, 2015). The allure for software-defined storage is also around the flexibility it offers and the low overall cost of storage (Quintero, 2015).

As storage systems continue to evolve, the storage industry is looking more into convergence as converged infrastructure enables efficiency in operations, applications and service management. Also, there will be continuous development of more scalable storage systems with higher performance.

CHAPTER 5: CONCLUSION

As discussed in this report, the storage industry is continuously growing and storage in the cloud is not left behind. In discussing storage, the prevalent issues to consider will always have to do with capacity, speed, cost and size, and cloud storage is better than the traditional storage methods when putting all these factors into consideration.

Effective storage is the core of any datacenter and to meet the requirements of large enterprises of today, there is a need to develop higher performance and efficient storage model. As discussed in the previous section, the recent trend in flash-based storage will continue to evolve to effective support to datacenters.

As more businesses join in the wagon of cloud storage, there is more to expect in the future.

REFERENCES

- Alhamad, M., Dillon, T., & Chang, E. (2010). Conceptual SLA framework for cloud computing. *2010 4th IEEE International Conference on Digital Ecosystems & Technologies (DEST)*, , 606. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=81351155&site=eds-live&scope=site>
- Arzensek, J. (2014). *Implementing the IBM storwize V7000 unified disk system. [electronic resource]* Poughkeepsie, NY : IBM Corp., International Technical Support Organization, 2014; 2nd ed. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7013133&site=eds-live&scope=site;> <http://proquestcombo.safaribooksonline.com/?uiCode=MacEwan&xmlId=0738439185>
- Barker, R., & Massiglia, P. (2002). *Storage area network essentials. [electronic resource] : A complete guide to understanding and implementing SANs* New York : Wiley, c2002 (Norwood, Mass. : Books24x7.com generator]). Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7052387&site=eds-live&scope=site;> <http://www.books24x7.com/marc.asp?bookid=7247>
- Buyya, R., Broberg, J., & Goscinski, A. (2011). *Cloud computing. [electronic resource] : Principles and paradigms* Hoboken, N.J. : John Wiley & Sons, c2011 (Norwood, Mass. : Books24x7.com generator]). Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?dir>

[ect=true&db=cat03710a&AN=alb.7058968&site=eds-live&scope=site;
http://www.books24x7.com/marc.asp?bookid=40612](http://www.books24x7.com/marc.asp?bookid=40612)

Callaghan, B. (1999). *NFS illustrated*. Addison-Wesley.

Carswell, R. (2015). *Test drive the microsoft private cloud* Boston, MA : Cengage Learning, PTR, 2015]. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.6784543&site=eds-live&scope=site;>
<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9781285874456;>
<http://proquestcombo.safaribooksonline.com/?uiCode=MacEwan&xmlId=9781285874456>

CodeFX, s. (2001). **CIFS explained**. Retrieved from

http://www.codefx.com/CIFS_Explained.htm

Collings, T., & Wall, K. (2002). *Red hat linux networking and system administration*. [electronic resource] New York, N.Y. : Hungry Minds, Inc., c2002 (Norwood, Mass. : Books24x7.com generator]). Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7054078&site=eds-live&scope=site;>
<http://www.books24x7.com/marc.asp?bookid=4830>

Coyne, L., Shivaramakrishnan, G., & Sing, J. (2014). *IBM Private, Public and Hybrid Cloud Storage Solutions*. [electronic resource] : Data ONTAP 8.1 operating in 7-mode

Poughkeepsie, NY : IBM Corp., International Technical Support Organization, c2014.

Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5978998&site=eds-live&scope=site;>

<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=0738436895>

Daniel, E. D., Mee, C. D., Nielsen, H., Engel, F. K., Gooch, B. R., Watkinson, J. R., . . . Noble, D. L. (1999). *Magnetic recording: The first 100 years* IEEE Press. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=htm&AN=XHTE1997912-H&site=eds-live&scope=site>

Dhamdhere, S. N. (2014). *Cloud computing and virtualization technologies in libraries*.

Hershey, PA: IGI Global. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=647108&site=eds-live&scope=site>

Dong, J. (2007). *Network dictionary. [electronic resource]* Saratoga, Calif. : Javvin

Technologies, Inc., c2007 (Norwood, Mass. : Books24x7.com generator]). Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7055495&site=eds-live&scope=site;>

<http://www.books24x7.com/marc.asp?bookid=23111>

Eckstein, R. (2007). *Using samba. [electronic resource]* Sebastopol, Calif. : O'Reilly, 2007; 3rd ed. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5573563&site=eds-live&scope=site;>

<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=0596007698>

Ganger, D. L., & Femling, R. (2008). *Mastering system center data protection manager 2007*.

[*electronic resource*] Indianapolis, Ind. : Wiley Pub., c2008 (Norwood, Mass. :

Books24x7.com generator]). Retrieved from

[http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7055726&site=eds-live&scope=site;](http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7055726&site=eds-live&scope=site;http://www.books24x7.com/marc.asp?bookid=24410)
<http://www.books24x7.com/marc.asp?bookid=24410>

Goda, K., & Kitsuregawa, M. (2012). *The history of storage systems* Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edswsc&AN=000309838000030&site=eds-live&scope=site>

Henry-Stocker, S., & Marks, E. R. (2000). *Solaris solutions for system administrators*.

[*electronic resource*] : *Time-saving tips, techniques, and workarounds* New York : John

Wiley & Sons, c2000 (Norwood, Mass. : Books24x7.com generator]). Retrieved from

[http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7050714&site=eds-live&scope=site;](http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7050714&site=eds-live&scope=site;http://www.books24x7.com/marc.asp?bookid=2297)
<http://www.books24x7.com/marc.asp?bookid=2297>

Hertel, C. (2003). *Implementing CIFS: The common internet file system*.

Hoover, J. N. (2013). *Compliance in the ether: Cloud computing, data security and business*

regulation [notes] Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edshol&AN=hein.journals.jobtela8.19&site=eds-live&scope=site>

Hoskins, J. (2005). *Exploring IBM server & storage technology. [electronic resource] : A layman's guide to the IBM eServer and TotalStorage families, sixth edition* Gulf Breeze, Fla. : Maximum Press, c2005 (Norwood, Mass. : Books24x7.com generator)]; 6th ed. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7055492&site=eds-live&scope=site;>
<http://www.books24x7.com/marc.asp?bookid=23108>

Jang, M. H. (2003). *Mastering red hat linux 9. [electronic resource]* San Francisco ; London : Sybex, 2003. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7167275&site=eds-live&scope=site;>
<http://site.ebrary.com/lib/rdcollege/Doc?id=10149952;>
<http://site.ebrary.com/lib/macewan/Doc?id=10149952;>
<http://ezproxy.agpc.talonline.ca/login?url=http://site.ebrary.com/lib/gprc/Doc?id=10149952;>
<http://site.ebrary.com/lib/concordiaab/Doc?id=10149952;>
<http://site.ebrary.com/lib/norquest/Doc?id=10149952;>
<http://site.ebrary.com/lib/oldscollge/Doc?id=10149952>

Javvin. (2007) *Network protocols handbook. [electronic resource]* (2007). Saratoga, Calif. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5572520&site=eds-live&scope=site;>
<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9781602670020>

Jia, W., & Zhou, W. (2005). *Distributed network systems. [electronic resource] : From concepts to implementations* New York : Springer, c2005. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.4306795&site=eds-live&scope=site;>

<http://www.springerlink.com/openurl.asp?genre=book&isbn=978-0-387-23839-5;>

<http://www.mylibrary.com/?id=19035>

Jiehui, J. U., Jiyi, W. U., Jianqing, F. U., & Zhijie, L. I. N. (2011). A survey on cloud storage.

Journal of Computers, (8), 1764. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsdaj&AN=c353b9b1be9ec78a8c3ff6851c4ca9bd&site=eds-live&scope=site>

Kitta, T. (2011). *SharePoint server 2010 enterprise content management*. [electronic resource]

2011, Indianapolis, Ind. : J. Wiley & Sons, c2011

Kovalick, A. (2013). *Video systems in an IT environment : The basics of networked media and file-based workflows* Burlington, MA : Focal Press, 2013; Second edition. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.6809413&site=eds-live&scope=site;>

<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9780240810423>

Kozierok, C. M. (2005). *The TCP/IP guide*. [electronic resource] : A comprehensive, illustrated internet protocols reference San Francisco : No Starch Press, c2005; 1st ed. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5575364&site=eds-live&scope=site;>

<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9781593270476>

Langbehn, T. (2005). NFS Version 4 - uni-potsdam.de. (n.d.). Retrieved from

<http://www.dcl.hpi.uni-potsdam.de/teaching/osAdminSem05/NFSv4.pdf>

- Lee, G. G. (2014). *Cloud networking. [electronic resource] : Understanding cloud-based data center networks* Waltham, MA : Morgan Kaufmann, c2014 (Norwood, Mass. : Books24x7.com generator]). Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7062387&site=eds-live&scope=site;http://www.books24x7.com/marc.asp?bookid=69992>
- Long, J. (2006). *Storage networking protocol fundamentals* Indianapolis, Ind. : Cisco Press, c2006. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.3593851&site=eds-live&scope=site>
- Lowe, D. (2013). *Networking all-in-one for dummies. [electronic resource]* Hoboken, N.J. : Wiley, c2013; 5th ed. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.6734292&site=eds-live&scope=site;http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9781118381007>
- Marshall, D., Beaver, S. S., & McCarty, J. W. (2009). *VMware ESX essentials in the virtual data center. [electronic resource]* Boca Raton, FL : Auerbach Publications, c2009 (Norwood, Mass. : Books24x7.com generator]). Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7057053&site=eds-live&scope=site;http://www.books24x7.com/marc.asp?bookid=30503>

- May, J. M. (2001). *Parallel I/O for high performance computing* San Francisco, CA : Morgan Kaufmann Publishers, 2001. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.2458961&site=eds-live&scope=site>
- Mell, P., & Grance, T. (2010). The NIST definition of cloud computing. *Communications of the ACM*, 53(6), 50-50. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=51189663&site=eds-live&scope=site>
- Miller, M. (2009). *Cloud computing. [electronic resource] : Web-based applications that change the way you work and collaborate online* Indianapolis, Ind. : Que, 2008, c2009. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5570462&site=eds-live&scope=site>;
<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9780768686227>
- Minx. (2014). *Data storage networking : Real world skills for the comptia storage+TM certification* Indianapolis, Indiana : Sybex, 2014. Retrieved from <http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7224497&site=eds-live&scope=site>;
<http://site.ebrary.com/lib/rdcollege/Doc?id=10849294>;
<http://site.ebrary.com/lib/oldcollege/Doc?id=10849294>;
<http://site.ebrary.com/lib/norquest/Doc?id=10849294>;
<http://site.ebrary.com/lib/macewan/Doc?id=10849294>;

[http://ezproxy.agpc.talonline.ca/login?url=http://site.ebrary.com/lib/gprc/Doc?id=10849294;](http://ezproxy.agpc.talonline.ca/login?url=http://site.ebrary.com/lib/gprc/Doc?id=10849294)
<http://site.ebrary.com/lib/concordiaab/Doc?id=10849294>

Norman, D. (2001). Fibre channel technology for storage area networks. Retrieved from

https://www.rivier.edu/faculty/vriabov/CS553a_Paper_DNorman.pdf

Noyes, T., & Dickinson, W. E. (2000). The random-access memory accounting machine. *IBM*

Journal of Research & Development, 44(1), 16. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=3149567&site=eds-live&scope=site>

Patterson, D. A., & Hennessy, J. L. (2012). *Computer organization and design. [electronic*

resource] : *The hardware/software interface* Waltham, MA : Morgan Kaufmann, c2012; 4th

ed. Retrieved from

[http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5552827&site=eds-live&scope=site;](http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5552827&site=eds-live&scope=site)
<http://lib.mylibrary.com/?id=331090>

Paulsen, K. (2011). *Moving media storage technologies. [electronic resource]* : *Applications &*

workflows for video and media server platforms Burlington, MA : Focal Press, 2011.

Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?dir>

[ect=true&db=cat03710a&AN=alb.4980879&site=eds-live&scope=site;](#)

<http://www.sciencedirect.com/science/book/9780240814483>

Poelker, C., & Nikitin, A. (2009). *Storage area networks for dummies*. [electronic resource]

Hoboken, N.J. : Wiley, c2009; 2nd ed. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?dir>

[ect=true&db=cat03710a&AN=alb.5575315&site=eds-live&scope=site;](#)

<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9780470385135>

Posey, B. (2013). Five LUN management basics - SearchStorage. (n.d.). Retrieved from

<http://searchstorage.techtarget.com/tip/Five-LUN-management-basics>

Quintero, D. (2015). *IBM software defined infrastructure for big data analytics workloads*

United States] : International Business Machines Corporation, International Technical

Support Organization, 2015]; First edition (June 2015). Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?dir>

[ect=true&db=cat03710a&AN=alb.7116393&site=eds-live&scope=site;](#)

<http://www.books24x7.com/marc.asp?bookid=93084>

Reilly, E. D. (2003). *Milestones in Computer Science and Information Technology*.

Rouse, M. (2014). **iSCSI (internet small computer system interface)**. Retrieved from

<http://searchstorage.techtarget.com/definition/iSCSI>

Schulz, G. (2012). *Cloud and virtual data storage networking*. [electronic resource] : Your

journey to efficient and effective information services Boca Raton, Fla. : CRC Press, c2012

(Norwood, Mass. : Books24x7.com generator]). Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7060160&site=eds-live&scope=site;>
<http://www.books24x7.com/marc.asp?bookid=46424>

Shore, B. (1988). *Introduction to computer information systems* New York : Holt, Rinehart & Winston, c1988. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.848630&site=eds-live&scope=site>

Shoshani, A., & Rotem, D. (2010). *Scientific data management : Challenges, technology, and deployment* Boca Raton : CRC Press, c2010. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.4743274&site=eds-live&scope=site;> <http://bvbr.bib-bvb.de:8991/F?func=service&doc%5Flibrary=BVB01&doc%5Fnumber=018927684&line%5Fnumber=0001&func%5Fcode=DB%5FRECORDS&service%5Ftype=MEDIA>

Smith, R. W., Hertel, C. R., & Iverson, W. (2004). *The definitive guide to samba 3. [electronic resource]* Berkeley, Calif. : Apress, c2004 (Norwood, Mass. : Books24x7.com generator)].

Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7053023&site=eds-live&scope=site;>
<http://www.books24x7.com/marc.asp?bookid=8885>

Snedaker, S. (2007). *Business continuity & disaster recovery for IT professionals*. Burlington, MA: Syngress. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=211412&site=eds-live&scope=site>

Sosinsky, B. A. (2009). *Networking bible*. [electronic resource] Indianapolis, Ind. : Wiley Pub., c2009. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5567016&site=eds-live&scope=site>;
<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9780470431313>

Stair, R. M., & Reynolds, G. W. (2008). *Fundamentals of information systems : A managerial approach* Boston, Mass. : Thomson Course Technology, c2008; 4th ed. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.3971511&site=eds-live&scope=site>

Stern, H., Eisler, M., & Labiaga, R. (2001). *Managing NFS and NIS*. [electronic resource]

Sebastopol, CA : O'Reilly & Associates, c2001; 2nd ed. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5559121&site=eds-live&scope=site>;
<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=1565925106>

StoneFly, I. (2016). **What is file level storage vs. block level storage?**. Retrieved from

<http://www.stonefly.com/resources/File-Level-Storage-vs-Block-Level-Storage.asp>

Tate, J. (2015). *VersaStack solution by cisco and IBM with SQL, spectrum control, and spectrum protect* Poughkeepsie, NY : IBM Corporation, International Technical Support Organization, 2015; First edition. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7149064&site=eds-live&scope=site;>

<http://proquestcombo.safaribooksonline.com/?uiCode=MacEwan&xmlId=9780738441078>

Tate, J. (2016). *Introduction to storage area networks* Poughkeepsie, NY : IBM Corporation,

International Technical Support Organization, 2016; Seventh edition. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7333938&site=eds-live&scope=site;>

<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9780738441436;>

<http://proquestcombo.safaribooksonline.com/?uiCode=MacEwan&xmlId=9780738441436>

Vacca, J. R. (2013). *Computer and information security handbook*. [electronic resource]

Waltham, Mass. : Morgan Kaufmann, c2013; 2nd ed. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.6378355&site=eds-live&scope=site;>

<http://proquest.safaribooksonline.com/?uiCode=ualberta&xmlId=9780123943972>

Vernooij, J. R. (2008). *The Official Samba-4 HOWTO*,

Wahl, C., & Pantol, S. (2014). *Networking for VMware administrators*. [electronic resource]

Upper Saddle River, NJ : VMware Press, 2014. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.7013080&site=eds-live&scope=site;>

<http://proquestcombo.safaribooksonline.com/?uiCode=MacEwan&xmlId=9780133511260>

Wieder, P. (2011). *Service level agreements for cloud computing*. [electronic resource] New York, NY : Springer, c2011. Retrieved from

<http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat03710a&AN=alb.5422737&site=eds-live&scope=site;>

<http://dx.doi.org/10.1007/978-1-4614-1614-2>

Wu, C., & Buyya, R. (2015). *Cloud Data Centers and Cost Modeling: A Complete Guide to Planning, Designing and Building a Cloud Data Center*. Morgan Kaufmann.

Wu, C. H. J., & Irwin, J. D. (2013). *Introduction to computer networks and cybersecurity*. CRC Press.