IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 10, October 2019

Storage Architecture Purpose the approach and Parameters

Priyanka Mukund Kamble

M.Tech Student, Department of Computer Engineering, Bharti Vidyapeeth College Of Engineering, Pune, India

Abstract: Analysis of automation that have near large data storage and storage networking. Data centers and cost reduction have driven the steady adoption of virtual server technologies. It centers on collating the system framework of the different building blocks of storage networks. With the assist of SAN,DAS and NAS, storage can be meticulously applied. Both technologies alternative the need to eliminate direct attached storage to aid more elastic storage access. Induced by peoples and software which is running applications and in this used different types of data format which is test, video, audio and images. DAS uses the IDE, SCSI, SATA, etc protocols and SAN uses the Fiber Channel, iSCSI, FoE, etc protocols to provide the storage facility. NAS is a device that is based on a network and it is used for sharing non-storage traffic.NAS is shareable .NAS provides some features such that provisioning, replication and snapshots .NAS can be costly.SAN provide the dedicated network for storage devices and the processors that access those devices .SAN used fiber channel technology for sharing data fatly .Replication can be done in SAN storage. The NAS and SAN can be used in industry to providing network protocols for managing storage. Solid State Disks is a type of non-volatile storage media that stores persistent data on solid-state flash memory. It optimized to deliver high read and write performance for both sequential and random data request. It refers to flash drive or solid-state disks.

Keywords: DAS, NAS and SAN Protocols Compare with SSD

I. INTRODUCTION

There are many problems to the current available networks. This problem increases largely as the network grows. So, in this the data grow largely day by day and storage handles the problems. Such that accessibility, speed, protection, capacity, I/O transfer rate, information access time, storage, data sharing. The advancement of networking technology for both LAN and WAN and that's create new application and generate large amount of data storage.
A modern storage system relies on complex software and interconnects to provide scalable, reliable access to large amounts of data across multiple machines. In conventional, disk-based storage systems the overheads from file systems, remote block device protocols (e.g., iSCSI and Fiber Channel), and network stacks are tiny compared to the storage media access time, so software overheads do not limit scalability or performance.

The emergence of non-volatile, solid-state memories (e.g., NAND flash, phase change memory, and spin-torque MRAM) changes this landscape completely by dramatically improving storage media performance.

Following applications of storage areas:

1. Capacity
2. Performance
3. Availability
4. Scalability
5. Cost

For the networking technology , the organizations can be used centralized computers and information storage devices in their data center .There is a evolution of storage architecture is open system which is affordable and easy to deploy and used it for business and department application to have their own storages and servers .There are two types of storage architectures:

1. Server centric storage architecture
2. Information centric storage architecture

1. **Server centric storage architecture:**

In this architecture, each server has limited number of storage devices and any administrative task, such as maintenance of the server or increasing storage capacity, might result in unavailability of information.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 8, Issue 10, October 2019

The growth of the departmental servers in an industry level result is unprotected, unmanaged fragmented islands of information and increased capital operating expenses.

2. **Information centric storage architecture**:

To overcome above challenges ,we use for storage is information centric storage architecture, which is converted server centric storage architecture to information centric storage architecture .In this architecture ,storage devises managed or maintained by centrally and it is independent of other servers. This is shared with multiple servers.

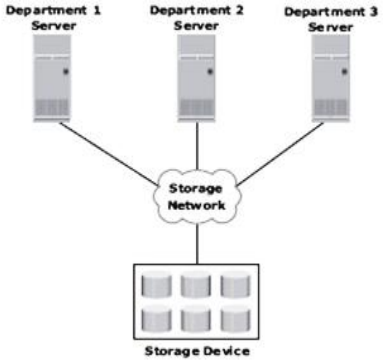


Fig. Information centric storage architecture

II TYPES OF STORAGE TECHNOLOGY

There are three categories of data storage technology which is include Direct Attached Storage(DAS),Storage Area Network (SAN),Network Attached Storage (NAS) between this three technology define its framework, similarity and also benefits of sharing storage resources by the use of network and also define how different schemes can be used to handle these three.

1. DAS:

- DAS is block device which is generally confined to access by a single host.
- It is simplest and very useful for single computer, PC and servers.
- This technology includes some protocols they are IDE, ATA, SATA, and SCSI etc.

2. SAN:

- SAN provide block oriented technique to handle the IO operations.
- This is used for faster transmission of data.
- In this technology include protocols fiber channel, iSCSI etc.

3. NAS:

- NAS is shareable block device over the network.
- NAS technology used for file system related.
- In this technology include protocols are NFS, CIFS, AFS etc.

III. DIRECT ATTACHED STORAGE TECHNOLOGY

Direct Attached Storage (DAS) is digital storage directly attached to the computer and accessing it, as opposed to storage accessed over a computer network (i.e. Network Attached Storage) .Example of DAS include hard drive, solid state drive, optical disc drives and storage on external drives. The name “DAS” is a retronym to contrast with storage area network (SAN) and the network attached storage (NAS).A typical DAS system is made of a data storage device (for example enclosures holding a number of hard disk drives) connected directly to a computer through a host adapter (HBA).Between these network device (like hub, switch or router) and this is the main characteristics of DAS. The DAS use this important protocol for connections such that ATA, SATA, SCSI, Fiber Channel, SAS, USB, USB3.0.

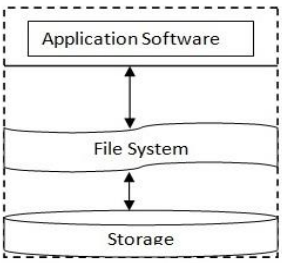


Fig. DAS storage

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 8, Issue 10, October 2019

The fig. explains the DAS network. It is a block level storage system and it is combination of both SAN and NAS. As per the fig. the following content include in DAS storage.

- DAS is directly attached to the server or host. In between this does not require database.
- It is provide highest performance among the three as it interfaces directly with storage.
- DAS service is limited to particular machine or server and it cannot use another host until it is made available to them using either SAN or NAS.
- DAS controller unit allows up to maximum of four servers to have access with same logic storage device.

IV. NETWORK ATTACHED STORAGE (NAS)

Network Attached Storage (NAS) which is directly attached to the computer n/w (LAN) through the network file systems they are NFS (Network File System) and CIFS (Common Internet File System) .Today , network is used Ethernet LAN , but in that the network is support IP based protocols that is NAS uses . NAS is generally uses the disk storage and sometimes CD-ROM, in an integrated package. Tape drives are concept for a data storage device that reads and writes data on a magnetic tape.

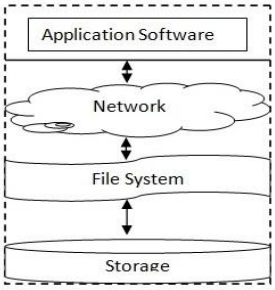


Fig. NAS architecture

In the NAS, shown in fig storage shared on Local area Network,

- It uses standard network and file sharing protocols such as FTP, SMB & CIFS.
- Communication is done by the file level storage in NAS.
- NAS is like a home or corporate network which allows access of data from anywhere with great amount of privacy

There is no. of options or reasons to provide which technology is better.
Some organizations may implement combination of SAN, DAS, NAS solutions consider the following key points for specific area to use n/w storage.

1] Easy installation –

NAS storage is easy to installed than the SAN , NAS is installed on the LAN or WAN.NAS Storage is sharable block device so the Host can easily and fast access data , without using volume manager or any another devices or drivers .
SAN is opposite to NAS which take more time and perfect planning to, like in SAN transmission use fiber channel network and that network installed and managed by the SAN software.

2] Backup- In the NAS communication done sharable block device so, in their number of copies of data or replicas available which is on the tape drive .In SAN this facility or property provides only selected some packages.

3] File sharing-

NAS provide file sharing but SAN provide better performance than NAS .only installation and its managing is easy in NAS.

4] Performance-

SAN provides better and faster performance than NAS.

- In SAN use dedicated high-speed network and sub network that interconnects to the shared pools of storage devices to multiple servers.
- SAN uses FC for fast transmission of data.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 8, Issue 10, October 2019

5] Resource pulling-

NAS is shared block device so the NAS shares resources to the pool but the SAN having better performance and capacity to provide resource in the pool and it is fast accesses. Also SAN use tape drives and connected multiple disk to the pool so capacity of SAN pool is day by day increases.

V. STORAGE AREA NETWORK (SAN)

Storage Area Network which is the latest technology to stored large amounts of data. In this technology stored the enterprises, industries and organizations of large data sets. This is a computer network which provides accessed to consolidated block level data storage. SANs are primarily used to handle such devices like disk arrays & tape libraries .SANs are typically composed of hosts , switches, storage elements , and storage devices that are interconnected using a variety of technologies , topologies and protocols. SANs may also span multiple sites. A SAN presents storage devices to a host such that the storage appears to be locally attached.

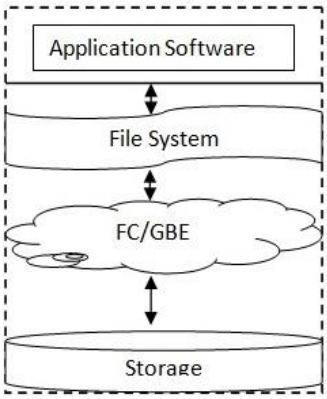


Fig. SAN architecture

The architecture of SAN storage which is provided block oriented to system in between computer systems and the targets disk systems. SAN use fiber channel or Ethernet for connectivity between hosts and storage. In this architecture use, dedicated SAN network connecting multiple application serves, database servers, NAS filters on one side and a number of disk systems and tape drive system on the other. In the servers and storage devices are connected each other by the SANs as peers. Following benefits of SAN charges.

1. Access - In this the data can be accessed by processor any time. High availability of access data and with this improves performance of the data. For access data fast use fiber channel cable.
2. Consolidation - Sharing multiple device or replace devices with each other and share the capacity in the tape or disk drive.
3. Protection- In this security refers to the collective measure, processes, tools and technologies that enable the securing of a SAN infrastructure .It is a broad process that ensures the SAN infrastructure operates securely and is protected from any vulnerability. IN SAN infrastructure include encrypting data at rest when stored on a SAN infrastructure or storage drives.

VI. SAN PROTOCOLS

Fiber Channel Protocol (FCP) - The most widely used SAN or block protocol, deployed in 70% to 80% of the total SAN market. FCP uses Fiber channel transport protocols with embedded SCSI commands.

- Internet small computer system interface (iSCSI) - The next largest SAN or bock protocol, with approximately 10% to 15% of the market. iSCSI encapsulates SCSI commands inside an Ethernet frame and then uses an IP Ethernet network for transport.
- SCSI – Is the dominant block level access method for disk in the data center scs1 itself is still widely used in its native format but it has also been encapsulated into other protocols for use within storage network for consolidated storage.
- SAS – serial Attached scs1 is a point to point serial protocol that moves data to and from computer storage devices such as hard drives and tape drives.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 8, Issue 10, October 2019

1) Fabric. NAS uses TCP/IP networks, most commonly Ethernet. Traditional SANs typically run on high speed Fibre Channel networks, although more SANs are adopting IP-based fabric because of FC's expense and complexity. High performance remains a SAN requirement and flash-based fabric protocols are helping to close the gap between FC speeds and slower IP.

2) Data processing. The two storage architectures process data differently: NAS processes file-based data and SAN processes block data. The story is not quite as straightforward as that of course: NAS may operate with a global namespace, and SANs have access to a specialized SAN file system. A global namespace aggregates multiple NAS file systems to present a consolidated view. SAN file systems enable servers to share files. Within the SAN architecture, each server maintains a dedicated, non-shared LUN. SAN file systems allow servers to safely share data by providing file-level access to servers on the same LUN.

3) Protocols. NAS connects directly to an Ethernet network via a cable into an Ethernet switch. NAS can use several protocols to connect with servers including NFS, SMB/CIFS, and HTTP. On the SAN side, servers communicate with SAN disk drive devices using the SCSI protocol. The network is formed using SAS/SATA fabrics, or mapping layers to other protocols such as Fibre Channel Protocol (FCP) that maps SCSI over Fibre Channel, or iSCSI that maps SCSI over TCP/IP.

4) Performance. SANs are the higher performers for environments that need high-speed traffic such as high transaction databases and ecommerce websites. NAS generally has lower throughput and higher latency because of its slower file system layer, but high-speed networks can make up for performance losses within NAS.

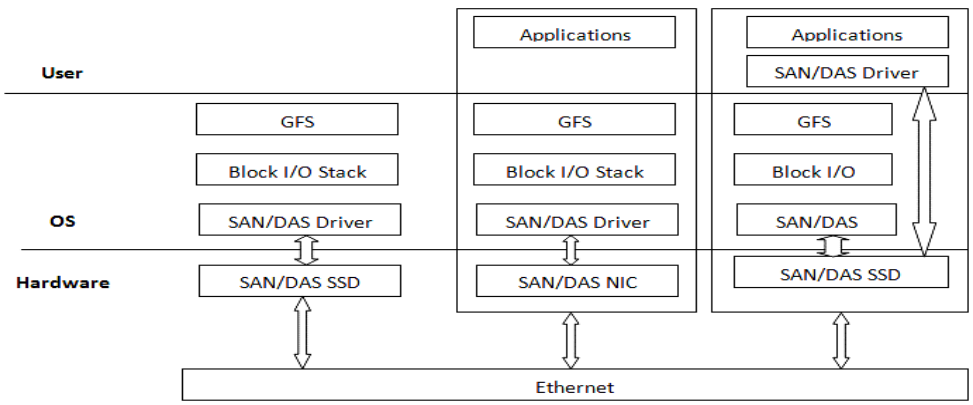
5) Scalability. Entry level and NAS devices are not highly scalable, but high-end NAS systems scale to petabytes using clusters or scale-out nodes. In contrast, scalability is a major driver for purchasing a SAN. Its network architecture enables admin to scale performance and capacity in scale-up or scale-out configurations.

6) Price. Although a high-end NAS will cost more than an entry-level SAN, in general NAS is less expensive to purchase and maintain. NAS devices are considered appliances and have fewer hardware and software management components than a storage area network. Administrative costs also figure into the equation. SANs are more complex to manage with FC SANs on top of the complexity heap. A rule of thumb is to figure 10 to 20 times the purchase cost as an annual maintenance calculation.

7) Ease of management. In a one-to-one comparison, NAS wins the ease of management contest. The device easily plugs into the LAN and offers a simplified management interface. SANs require more administration time than the NAS device. Deployment often requires making physical changes to the data center, and ongoing management typically requires specialized admin. The exception to the SAN-is-harder argument is multiple NAS devices that do not share a common management console.

Fast SSDs:

Fast non-volatile memories are already prestige the design of storage devices in the industry. As faster memories become available the cost of software and hardware overheads on storage accesses will become more pronounced. Commercially available NAND flash-based SSDs offer access latencies of tens of microseconds, and research prototypes targeting more advanced memory technologies have demonstrated latencies

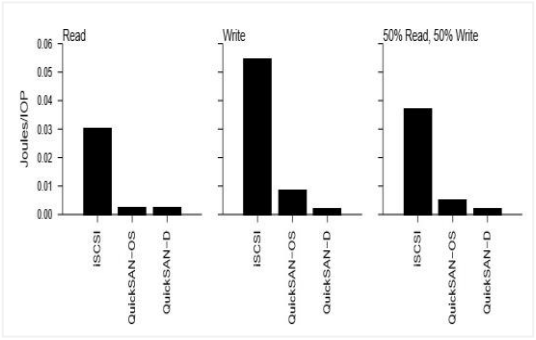


IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 8, Issue 10, October 2019

SAN/DAS configurations supports multiple storage topologies and two software interfaces. At left, a singlemachine hosts multiple SAN/DAS SSDs, acting a central block server. The center machine hosts a SAN/DAS NIC that provides access to remote SSDs. The machine at right hosts a single SSD and is poised to access its local (for maximum performance) or remote data via the users interface

Workload alliance:

An advantage of amalgamate DAS, NAS and SAN over distributed storage systems that spread data across many hosts is that shutting down hosts to save power does not make any data unreachable. Distributed storage systems, however, provide fast access to local data, improving performance. The DAS, NAS and SAN SSD's ability to service requests from remote nodes, even if their host machine is powered down can achieve the best of both worlds: Fast local access when the cluster is under heavy load and global data visibility to support migration. To explore this application space, we use four servers running a persistent key-value store (MemCacheDB). During normal, fully-loaded operation each server runs a single instance of the MemCacheDB. We use four machines running to drive the key-value stores on the four machines under test. As system load drops, we can migrate the key-value store to another server. On their new host, they can access the same data and transparently resume operation. We assume that a front-end steering mechanism redirects traffic as necessary using DNS. We can perform the same migration in both the centralized DAS and SAN and centralized iSCSI configurations.



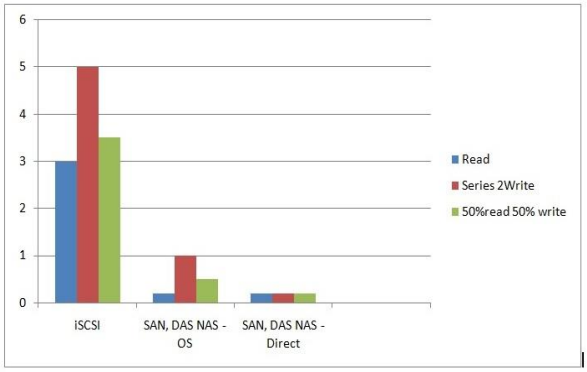


Figure: DAS and SAN energy efficiency Depending on the software stack and block transport mechanism, the energy cost of storage access varies. Removing the high software overheads of iSCSI account for most of the gains, but the user space interface saves 76% of energy for writes compared to the kernel version.

We have implemented emulation in SAN,NAS and DAS to improve data probity and availability. When emulation is enabled, SAN,NAS and DAS allocates half of the storage capacity on each SAN,NAS and DAS SSD as a emulation of the primary portion of another SSD. SAN,NAS and DAS transparently issues two write requests for each write. SAN,NAS and DAS can also select from any replicas to service remote read requests using a round-robin scheme, although it always selects a local replica, if one is available. With this scheme, SAN,NAS and DAS can tolerate the failure of any one SSD.

CONCLUSION

It touches on the fundamental ideas underlying each storage models, the prevailing protocols and the trends of technology evolution. It also describes how networks are put together with various network elements. DAS, NAS and SAN accommodate network functionality into a high-performance SSD to allow access to remote data without operating system intervention. DAS, NAS and SAN reduces software and block transport overheads by between 82 and 95% compared it Fiber Channel and iSCSI-based DAS, NAS and SAN implementations and can improve bandwidth for small requests by up to 167×. We also demonstrated that DAS and SAN can improve energy efficiency by 58% compared to an iSCSI-based DAS and SAN. DAS and SAN illustrate the ongoing need to redesign computer system architectures to make the best use of fast non-volatile memories.

Storage technology has clearly become more varied and sophisticated, and accordingly decisions have become more complex than ever. Choice means flexibility and that's good, but which choice to make is not always clear.

REFERENCES

[1]. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.531.453&rep=rep1&type=pdf
[2]. http://bnrg.cs.berkeley.edu/~randy/Courses/CS294.S13/12.1.pdf
[3]. http://aad.tpu.ru/practice/EMC/Information%20Storage%20and%20Management-v.2.pdf
[4]. Desai, P. and Jayakumar, N., AN EXTENSIBLE FRAMEWORK USING MOBILITYRPC FOR POSSIBLE DEPLOYMENT OF ACTIVE STORAGE ON TRADITIONAL STORAGE ARCHITECTURE.
[5]. Jayakumar, N. and Kulkarni, A., 2017. A Simple Measuring Model for Evaluating the Performance of Small Block Size Accesses in Lustre File System. *Engineering, Technology & Applied Science Research*, *7*(6), pp.2313-2318.
[6]. Rishikesh Salunkhe, N.J., 2016. Query Bound Application Offloading: Approach Towards Increase Performance of Big Data Computing. *Journal of Emerging Technologies and Innovative Research*, *3*(6), pp.188-191.
[7]. Salunkhe, R., Kadam, A.D., Jayakumar, N. and Joshi, S., 2016, March. Luster a scalable architecture file system: A research implementation on active storage array framework with Luster file system. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)* (pp. 1073-1081). IEEE.
[8]. Jayakumar, N., Iyer, M.S., Joshi, S.D. and Patil, S.H., A Mathematical Model in Support of Efficient offloading for Active Storage Architectures.
[9]. Naveenkumar, J. and Joshi, S.D., 2015. Evaluation of Active Storage System Realized through MobilityRPC.
[10]. IBM Storage networking 2011
[11]. Jayakumar, N., Bhardwaj, T., Pant, K., Joshi, S.D. and Patil, S.H., A Holistic Approach for Performance Analysis of Embedded Storage Array.
[12]. Naveenkumar, J., Makwana, R., Joshi, S.D. and Thakore, D.M., 2015. Performance Impact Analysis of Application Implemented on Active Storage Framework. *International Journal*, *5*(2).
[13]. Naveenkumar, J. and Joshi, S.D., 2015. Evaluation of Active Storage System Realized Through Hadoop.
[14]. Zaeimfar, S.N.J.F., 2014. Workload Characteristics Impacts on file System Benchmarking. *Int. J. Adv*, pp.39-44.
[15]. Jayakumar, D.T. and Naveenkumar, R., 2012. SDjoshi,". *International Journal of Advanced Research in Computer Science and Software Engineering," Int. J*, *2*(9), pp.62-70.
[16]. Suryawanshi, A.U. and Naveenkumar, J., PRIVACY-PRESERVING FOR A SECURE DATA STORAGE ON CLOUD USING PUBLIC AUDITING TECHNIQUE.
[17]. Suryawanshi, A.U. and Kumar, J.N., A Multiuser Model of Privacy-Preserving Auditing for Storing Data Security in Cloud Computing.
[18]. Naveenkumar J and Raval K.S, Clouds Explained Using Use-Case Scenarios.
[19]. Sagar S Lad, S.D Joshi, N.J, Comparison study on Hadoop's HDFS with Lustre File System. International Journal of Scientific Engineering and Applied Science, 1(8), pp.491–494, 2015.
[20]. Jayakumar, N., Bhardwaj, T., Pant, K., Joshi, S.D. and Patil, S.H., 2015. A Holistic Approach for Performance Analysis of Embedded Storage Array. *Int. J. Sci. Technol. Eng*, *1*(12), pp.247-250.
[21]. Kumar, N., Kumar, J., Salunkhe, R.B. and Kadam, A.D., 2016, March. A Scalable Record Retrieval Methodology Using Relational Keyword Search System. In *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies* (p. 32). ACM

BIOGRAPHY

Priyanka Mukund Kamble
M.Tech Student, Department of Computer Engineering, Bharti Vidyapeeth College Of Engineering, Pune, India