

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

WORKLOAD AND PERFORMANCE ANALYSIS OF DATA IN DISTRIBUTED STORAGE AREA NETWORK

Aman Singh¹, Roshan Srivastava²

¹M.Tech, Lovely Professional University, Phagwara
er.aman003@gmail.com

² Lovely Professional University, Phagwara
roshan.16876@lpu.co.in

Abstract: Storage area network is defined as a network whose main purpose is to transfer the data from computer systems to storage elements. It consists of a communication infrastructure, this infrastructure helps to provide the physical connections. In the DSAN, there is a management layer, which helps to organize the connections, storage elements, and computer systems. Our aim is to enhance the performance and to decrease the workload in distributed storage area network. There are lots of activities in DSAN, like to store information in DSAN, to mining information from DSAN. Now if we concentrate on mining information from DSAN then to make it easy we need a searching option. Now days we use prefix search for mining information from storage area network.

Keywords: DSAN, Storage, Performance enhancement, LAN.

1. INTRODUCTION

A computer network is simply a collection of large number of devices. These devices are generally known as the computers. These devices are interconnected with each other by communication channels. The communication channels allow the network, for efficient sharing of resources, services, and information. A storage area network is a type of local area network that designed to handle the large data transfers. It supports the data storage, retrieval and replication on business networks. It uses the high end servers; multiple disk arrays and Fiber Channel inter connection technology. Storage area network is also defined as a network whose main purpose is to transfer the data from computer systems to storage elements. It consists of a communication infrastructure, this infrastructure helps to provide the physical connections. In the DSAN, there is a management layer, which helps to organize the connections, storage elements, and computer systems. Storage area network is identified with block input/output services. DSAN has its own network of storage devices, these storage are not accessible through the local area network by other devices. DSAN is a high speed network that attaches servers and storage devices. Sometimes DSAN is referred as the network behind the servers. DSAN

allows the any to any (ATA) connection across the whole network. For this purpose it uses the interconnect elements, such as switches and directors. DSAN helps to eliminate the traditional connection between a server and storage. Server helps to owns and manages the storage devices. It also eliminates the restriction to the amount of data that a server can access [1] DSAN introduces the flexibility of the networking to enable one or many heterogeneous servers to share a common storage utility. A network might include many storage devices, including disk, tape, and optical storage. In the DSAN, the storage utility may be located far from the servers that it uses. In the DSAN, the file systems built on top of DSANs, which provide file level access.

In previous years, the data centre is created islands of SCSI disk arrays. These arrays are known as direct-attached storage (DAS). DSAN consolidates such storage islands together using a high-speed network. The Operating systems maintain their own file systems.[2] DSAN help to increase storage capacity utilization, since multiple servers consolidate their private storage space onto the disk arrays. Common uses of a DSAN include provision of transactional accessed data that require high-speed block-level access to the hard drives such as email servers, databases, and high usage file servers.

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

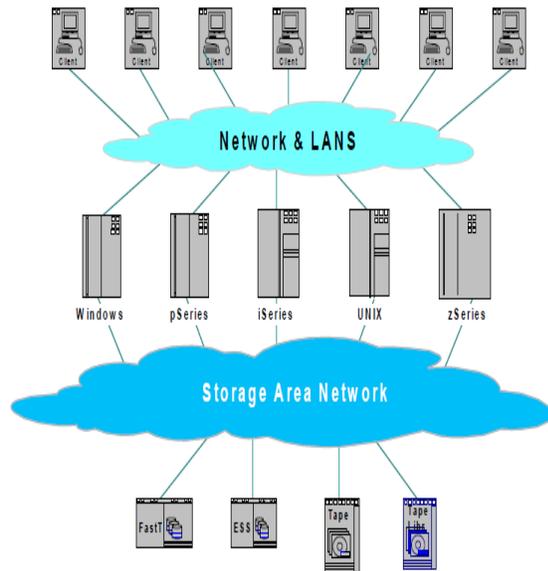


Figure 1: Storage Area Network

In the figure 1, we discuss about the storage area network, here in this figure there are several operating systems that take the information from the networks and stored in the storage area network.

Architecture of DSAN:

The DSAN overlay network consists of all the participating peers as network nodes. There are links between any two nodes that know each other: i.e. if a participating peer knows the location of another peer in the DSAN network, then there is a directed edge from the former node to the latter in the overlay network. Based on how the nodes in the overlay network are linked to each other, we can classify the DSAN networks as unstructured or structured.

Structured Storage Network:

Structured DSAN network employ a globally consistent protocol to ensure that any node can efficiently route a search to some peer that has the desired file, even if the file is extremely rare. Such a guarantee necessitates a more structured pattern of overlay links. By far the most common type of structured DSAN network is the distributed hash table (DHT), in which a variant of consistent hashing is used to assign ownership of each file to a particular peer, in a way analogous to a traditional hash table's assignment of each key to a particular array slot.

Unstructured Storage Area Network:

An unstructured DSAN network is formed when the overlay links are established arbitrarily. Such

networks can be easily constructed as a new peer that wants to join the network can copy existing links of another node and then form its own links over time. In an unstructured DSAN network, if a peer wants to find a desired piece of data in the network, the query has to be flooded through the network to find as many peers as possible that share the data. The main disadvantage with such networks is that the queries may not always be resolved. Popular content is likely to be available at several peers and any peer searching for it is likely to find the same thing. But if a peer is looking for rare data shared by only a few other peers, then it is highly unlikely that search will be successful. Since there is no correlation between a peer and the content managed by it, there is no guarantee that flooding will find a peer that has the desired data. Flooding also causes a high amount of signaling traffic in the network and hence such networks typically have very poor search efficiency. Most of the popular DSAN networks are unstructured.

Security Risks in DSAN:

Users of DSAN file-sharing systems face many of the same security risks as other Internet users. Just as in other applications, DSAN users must take care to only run programs from sources that they trust, and should be careful to check for viruses. They should safeguard their computer from attack when online. File sharing adds an extra dimension to these concerns due to the quantity and frequency of files traded the relatively unsophisticated user base, and the rise of self-help systems to prevent copyright infringement. At this time, DSAN file-sharing applications are not known to be any less -- or any more -- securing than Internet applications on the market in other areas.

Viruses:

Storage Area file sharing networks enable files to be transferred among millions of computers most of which are owned and operated by total strangers[4] there is an ever-present risk that files downloaded from a Storage Area file sharing network could carry various kinds of malicious software like viruses and "worms." mlt is of course, possible to receive a dangerous file in numerous ways, such as over the Web or by e-mail. The best protection against viruses continues to be the use of up-to-date anti-virus software. 100% protection can never be achieved, but users should be aware that to download files without adequate protection opens them up to substantial risks.

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

Online Attacks:

When Storage Area networks identify shared files to millions of users, they also identify the location of a user's computer, and could even target that computer's IP address (Internet Protocol address) with attempts to gain access. This is not a risk unique to Storage Area file sharing networks; all Internet communications involve an exchange of IP addresses. But because Storage Area file sharing networks search millions of computers, they can provide access to millions of IP addresses.

Self-Help Attacks:

A new form of security threat may be growing for Storage Area users in the rise of self-help techniques by copyright holders concerned about infringement on file-trading networks. More benign versions flood DSAN networks with bogus copies of copyrighted works in order to fool people into downloading or storing them.

2. LITERATURE REVIEW

Vagelis Hristidis, et.al discuss that the keyword search is the most popular information discovery method because the user does not need to know either a query language or the underlying structure of the data. When a set of keywords is provided by the user, the search engine returns all documents that are associated with these keywords. Two keywords and a document are associated when the keywords are contained in the document and their degree of associability is often their distance from each other. In this paper author introduces the minimal join network.

Steve Chidlow, discuss about the various DSAN technologies. A Storage Area Network is a separate network dedicated to storage devices and at minimum consists of one large banks of disks mounted in racks that provide for shared storage space which is accessible by many systems. The robotic tape libraries may be attached to the DSAN. Data storage resides on hard disks that are locally attached to individual servers. This is known as Direct Attached Storage. A DSAN functions as a high-speed network similar to a conventional local area network and establishes a direct connection between storage resources and the file server infrastructure.

The DSAN effectively acts as an extended storage bus using the same networking elements of a LAN including routers, bridges, hubs and switches. Thus,

servers and storage can be de coupled allowing the storage disks to be located away from their host servers. The DSAN is effectively transparent to the server operating system, which sees the DSAN attached disks as if they were local SCSI disks. A dedicated DSAN carries only storage data. This data can be shared with multiple servers without being subject to the bandwidth constraints of the normal network. DSAN allows for data to be managed centrally and to assign storage chunks to host systems as required. DSAN and NAS act as competitors, but in reality they are complementary technologies, DSAN delivering effective block-based input/output, whilst NAS excels at file based input/output. A hybrid device called a NAS Head or a NAS Gateway has storage that resides in the storage arrays attached to a DSAN whilst still delivering file systems over the LAN. A combination of a DSAN with NAS Gateways may be an effective way for sites to deliver file-based functionality e.g. for user home directories.[3] In fact, DAS still has an ongoing use for many purposes, the cost of connecting servers to the DSAN can be high and for systems like DNS servers DSANs have been mainly based on Fibre Channel technology, new IP based options using more commodity, like components are a possibility in the future.

Le-Shin Wu present and evaluate learning techniques to improve local query routing. Authorn validate prototypes of the 6S network via simulations with 70, 500 model users based on actual Web crawls. Author find that the network topology rapidly converges from a random network to a small world network, with clusters emerging from user communities with shared interests.

3. PURPOSED WORK

A storage area network is a dedicated network that provides access to consolidate and blocks level data storage. DSANs are primarily used to make storage devices, such as disk arrays, tape libraries, and optical jukeboxes, accessible to servers so that the devices appear like locally attached devices to the operating system.

A DSAN typically has its own network of storage devices that are generally not accessible through the local area network by other devices. Now our aim is to enhance the performance and to decrease the workload in distributed storage area network. There are lots of activities in DSAN, like to store information in DSAN, to mining information from

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

DSAN. Now if we concentrate on mining information from DSAN then to make it easy we need a searching option. Now days we use prefix search for mining information from storage area network. It will consume time for searching because it follows prefix pattern for search and for it, almost searches through all files which increase the workload.

- [4] Introduction to Storage Area Network, DSAN, International Technical Support Organization, 1999
- [5] J. Tate, A. Bernasconi, P. Mescher, F. Scholten, Introduction to Storage Area Networks, IBM Red Book, 2003.
- [6] Ravi Kumar Khattar, *et.al* Introduction to Storage Area Network, DSAN, SG 5470-00.

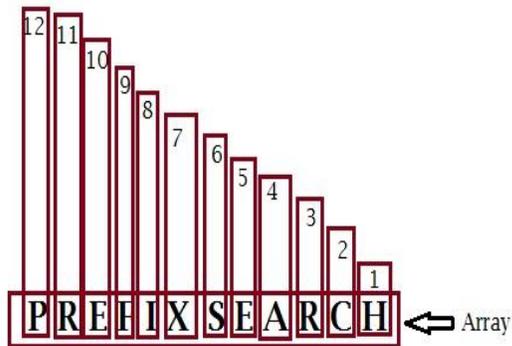


Figure 2: prefix search

In the figure 2, prefix searching is defined. Now to overcome this problem we can use prefix searching. Prefix searching can do search in less time as compare to suffix search because prefix search starts searching from the last element in the array and it will provide the more accurate result in less time. For prefix it does not increase the workload and it will give accurate results in less time. Means if we use prefix search in DSAN then we can increase the performance and decrease the workload of distributed storage area network easily.

REFERENCES

- [1] Le-Shin Wu, "ADAPTIVE PEER TO PEER SOCIAL NETWORKS FOR DISTRIBUTED CONTENT BASED WEB SEARCH".
- [2] JISC Technology and Standards Watch Report STORAGE AREA NETWORKS, JISC Technology and Standards Watch Report: Storage Area Networks November 2003
- [3] DISCOVER: Keyword Search in Relational Databases, Proceedings of the 28th VLDB Conference, Hong Kong, China, 2002