

Problems of network data storage and cloud databases

Mubariz¹ Khalilov^{1*}

¹Information science/Baku State University, Baku, Azerbaijan
*Corresponding Author and ⁺Speaker: khalilov_mubariz@mail.ru
Presentation/Paper Type: Oral / Full Paper

Abstract – Information becomes one of the most important assets of an enterprise and determines its competitiveness. One of the challenge of IT managers face is the huge amount of distributed data circulating in the information system. The inability to effectively manage them can negatively affect the profitability of the enterprise and limit its ability to grow. Therefore, nowadays, transfer, storage, protection and management of data in the context of the steady increase of the structure volume and complexity have become the major problems for business. Thus, the actual problems of network and cloud storage of data have been investigated.

Keywords - network, clouds, storage, information, data

I. INTRODUCTION

Computer information technology has enormously changed humanity. However, every year electronics is increasingly being integrated into our daily lives; the role of computer technology is indispensable. It is fact that with the advent of the Internet, most people on the planet have access to a huge database of information. Additionally, each year the number of participants and information is greatly increasing. Information can be accumulated on data carriers that can be easily transported through floppy disks, compact disks, flash media, hard disks. Nevertheless, there are cases when it is not impossible to take a certain medium with you or lack of amount of memory not allow to record the necessary information.

Many users think that information should be stored on the PC's hard drive, and if there is necessity to take it away, it can be copied to portable media. Meanwhile, with the advent of network and cloud storages, this approach is not considered useful: it is more convenient to place files in the cloud and change a location, knowing that at any time there is an access to them via a channel on the Internet.

Using the cloud is rather convenient to store backups and exchange information with other people as well. It is sufficiently to send certain link to the addressee, and he will receiver could upload these files at any suitable time.

Finally, in order to avoid the loss of important information (archiving issue) for instance, a hard disk fails, the requirement is coping it to the cloud and updating periodically.

Network storage system architectures were developed in the 1990s, and their task was to address the main shortcomings of DAS systems. In generally, in the field of storage systems network solutions have to implement three tasks: reduce costs and complexity of data management, decrease local network traffic, increase data availability and overall performance. At the same time, NAS and SAN architectures solve various aspects of a common problem. The result was the simultaneous coexistence of two network architectures, each of which has its own advantages and functionality.

A. NAS Attached Storage Devices

The primary goal of NAS systems is to simplify a file sharing. At a basic level, NAS devices are equipment that connects directly to the local network. This is their main difference from systems with individual servers with direct connection of isolated drives. NAS devices, often referred to as filers, consist of a single head unit that performs data processing and network connectivity to a disk chain. NAS devices allow to use storage systems on Ethernet networks, the TCP / IP protocol to organize file sharing. These devices admit clients to share files, even if the client systems are running various OSs. Unlike DAS architecture, there is not necessity for NAS systems to take servers offline to increase overall capacity; disks can be added to the NAS structure by simply connecting the device to the network.

B. NAS Storage Area Networks (SAN)

These solutions were the answer to the shortcomings of the DAS and NAS systems, and essentially, to the problems of communication channel overload and delays in local IP networks (10/100-Mbit/s) as well. The SAN concept was first proposed in 1998. Like many other modern computer technologies, it was borrowed from the mainframe world, that was used, for example, in data centers to connect computers to storage systems and distributed networks. The major difference between a SAN and a NAS is the way data is exchanged between storage devices and servers. In general, the SAN architecture is aimed at resolving problems caused by intensive backup and data exchange procedures by moving the entire system to a dedicated subnet. Fiber Channel protocol-based SANs allow to broadly vary storage capacity and guarantee higher bandwidth within a dedicated subnet (disk arrays and tape libraries not equipped with Fiber Channel interfaces can be connected to the SAN using Fiber Channel-SCSI routers)

C. Cloud data storage

Cloud data storage - an online storage model in which data is stored on numerous servers distributed on the network and provided for use by customers, mainly by a third party. In

contrast to the model for storing data on its own, dedicated servers, purchased or leased specifically for such purposes, the number or any internal structure of the servers is generally not visible to the client. Data is stored, as well as processed, in the so-called cloud, which represents, from the point of view of the client, one large, virtual server. Physically, such servers can be located remotely from each other geographically, up to a location on different continents. For active users that are constantly on the move, cloud storage is becoming an indispensable service. The very concept of cloud storage suggests that the user does not need to know exactly where his data is located: physically, they can be located not only on various servers, but the servers themselves can be on different continents. Nevertheless, the user sees the storage as a regular drive (the volume is determined by the service provision policy), and in some cases it can even include a cloud disk in the system and watch it as a local volume [1].

D. Cloud Storage Benefits

The advantages over local drives are obvious. The first one is its economic feasibility. Placing data in the clouds is significantly cheaper compared to standard storage. The second issue is its data availability. Since, the data is generally available to customers regardless of their location, however, it may depend on users Internet channel. Third aspect is the speed of content delivery. Almost all cloud storage providers offer to their customers CDN (Content Delivery Network or Content Distribution Network - a geographically distributed network infrastructure that allows optimizing the delivery and distribution of content to end users on the Internet). The fourth and very relevant advantage is the reduction in server load. However, the processes of the web server are not busy with the return of static, the load on the disk is reduced [2].

Data loss protection issue have always been very relevant, especially in contrast with the past. Files will never be lost if you be chosen a tariff plan with the appropriate guarantee and protection system. The client pays only for the storage space that he actually be in using, but not for renting a server, in other words, all the resources. There is not necessity for client to engage in the acquisition, support and maintenance of its own data storage infrastructure, which ultimately reduces the overall cost of production. All procedures for backing up and maintaining data integrity are performed by the cloud center provider and not involve the client in this process. It is feasible to work with cloud storages both for a company and an individual, anyone can start working, even not particularly versed in information technology [3].

E. Potential issues

Despite the many advantages, there are malfunctions in such a complex network. In most cases, this is due to a disconnection in a certain area, that leads to delays in work and loss of time and resources. Moreover, in addition to this, topics related to information security are remained as actual issues.

Security in the process of storing and sending data is one of the most basic issue when working with the cloud, in relation to confidential, private data, especially with the current level of cybercrime.

The reliability and timeliness of receiving and accessing data in the cloud significantly depends on many various intermediate parameters, mainly such as data transmission channels on the way from the client to the cloud, the last mile

question, the question of the proper quality of the client's Internet service provider, the question of the cloud's availability in given point in time.

Building a large "cloud", on which all the company's business systems work, it is necessary to ensure high fault tolerance, constantly back up data so that when the server crashes, instead of the clouds simultaneously disconnection, "moving" to another server, or (within a reasonable time) restoring from backups is emerged. All these leads to higher costs for creating of a fault-tolerant architecture. Investments in maintaining and backing up "clouds" are increasing. Compared to traditional data warehouses, "cloud" requires highly qualified employees - for the correct administration, installation, configuration and support of these solutions. Therefore this increases the cost of ownership of data warehouses on the "cloud".

F. Theoretical part

Currently, there are many Internet services that provide the maintenance of "Cloud storage" with the various capabilities. The only thing that is observed in the way of information processing is similarity. It is important to emphasize that modern "repositories" require mandatory registration (there are exceptions) and there is not matter if the services are used for a fee or not.

There are various criterias through user is able ~~can~~ and should choose a storage - the quantity of allocated memory, the bandwidth of the channel (both incoming and outgoing traffic). The paid service, the convenience and availability of them could be implemented via the web interface and mobile platforms with special applications. The feature like streaming work with files, for example, listening to a music online from personal "cloud" is also a privilege. Moreover, the ability of working with files through a proxy could be rather beneficial for IT companies.

G. Practical part

In this part of the study, the most popular existing solutions on a global scale are considered. Some services can be classified as for domestic use, and others for large enterprises as well.

Dropbox will be considered as one of the leaders amongst cloud data storages that allows users to retain their data on servers in the cloud. The service has not newly become popular and it has several advantages, for instance, such as:

- Online synchronization of files of any size and type in automatic mode between PCs, work is possible on Windows, Mac and Linux, and these are the most popular computer systems in present time.
- Invite friends to the system allows users to obtain more space.
- Users' files will remain with them even in the absence of Internet connection.
- Ability of manually setting the Internet connection speed, etc.

REFERENCES

- [1] [http://www.tadviser.ru/index.php/Статья:NAS_\(Network_Attached_Storage\)](http://www.tadviser.ru/index.php/Статья:NAS_(Network_Attached_Storage))
- [2] <http://www.nstor.ru/ru/catalog/76/77.html>
- [3] http://www.gazeta.ru/tech/2013/04/19_a_5261993.shtml
- [4] locnet.narod.ru/resheniya/nas.htm