© 2019 IJSRST | Volume 6 | Issue 5 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology
DOI : https://doi.org/10.32628/IJSRST

A Comprehensive Review on Data Storage

Sumalatha Sriramoju
PG Scholar, Master of Science in Information Technology, University of Management and Technologies, USA

ABSTRACT

Data storage is the collective methods and technologies that capture and retain digital information on electromagnetic, optical or silicon-based storage media. Storage is a key component of digital devices, as consumers and businesses have come to rely on it to preserve information ranging from personal photos to business-critical information. Modern storage systems require enhanced capabilities to allow enterprises to apply machine learning-enabled artificial intelligence (AI) to capture this data, analyze it and wring maximum value from it. This paper provides an overview of data storage technologies, evaluation of storage heirarchy and importance of data storage.

Keywords : Data Storage, Technologies, Storage Hierarchy

I. INTRODUCTION

The term storage may refer both to a user's data generally and, more specifically, to the integrated hardware and software systems used to capture, manage and prioritize the data. This includes information in applications, databases, data warehouses, archiving, backup appliances and cloud storage. Digital information is written to target storage media using software commands. The smallest unit of measure in a computer memory is a bit, described with a binary value of 0 or 1, according to the level of electrical voltage contained in a single capacitor. Eight bits make up one byte.

Data storage capacity requirements define how much storage is needed to run an application, a set of applications or data sets. Capacity requirements consider the types of data. For instance, simple documents may only require kilobytes of capacity, while graphic-intensive files, such as digital photographs, may take up megabytes, and a video file can require gigabytes of storage. Computer applications commonly list the minimum and recommended capacity requirements needed to run them.

On an electromechanical disk, bytes store blocks of data within sectors. A hard disk is a circular platter coated with a thin layer of magnetic material. The disk is inserted on a spindle and spins at speeds of up to 15,000 revolutions per minute (rpm). A sector on a standard disk is 512 bytes. Recent advances in disk include shingled magnetic recording, in which data writes occur in overlapping fashion to boost the platter's areal density.

On solid-state drives (SSDs), data is written to pooled NAND flash, designed with floating gate transistors that enable the cell to retain an electrical charge. An SSD is not technically a drive, but it exhibits design characteristics similar to an integrated circuit, featuring potentially millions of nano transistors placed on millimeter-sized silicon chips. Backup data copies are written to disk appliances with the aid of a hierarchical storage management system. And although less commonly practiced than in years past, the tactic of some organizations remains to write disk-

based backup data to magnetic tape as a tertiary storage tier. This is a best practice in organizations subject to legal regulations.

A virtual tape library (VTL) uses no tape at all. It is a system in which data is written sequentially to disks but retains the characteristics and properties of tape. The value of a VTL is its quick recovery and scalability.

Larger application scripts and real-time database analytics have contributed to the advent of highly dense and scalable storage systems, including high-performance computing storage, converged infrastructure, composable storage systems, hyper-converged storage infrastructure, scale-out and scale-up network-attached storage (NAS) and object storage platforms.

II. METHODS AND MATERIAL

EVALUATION OF STORAGE HIERARCHY

Organizations increasingly use tiered storage to automate data placement on different storage media, based on an application's capacity, compliance and performance requirements.

Enterprise data storage is often classified as primary and secondary storage, depending on how the data is used and the type of media it requires. Primary storage handles application workloads central to a company's day-to-day production and main lines of business.

Primary storage is occasionally referred to as main storage or primary memory. Data is held in random access memory (RAM) and other built-in devices, such as the processor's L1 cache. Secondary storage encompasses data on flash, hard disk, tape and other devices requiring I/O operations. Secondary storage media is often used in backup and cloud storage.

Primary storage generally provides faster access than secondary storage due to the proximity of storage to the computer processor. On the other hand, secondary storage can hold much more data than primary storage. Secondary storage also replicates inactive data to a backup storage device, yet keeps it highly available in case it is needed again.

Digital transformation of business has prompted more and more companies to deploy multiple hybrid clouds, adding a remote tier to buttress local storage.

TYPES OF DATA STORAGE DEVICES/MEDIUMS

Data storage media have varying levels of capacity and speed. These include cache memory, dynamic RAM (DRAM) or main memory; magnetic tape and magnetic disk; optical disc, such as CDs, DVDs and Blu-ray disks; flash memory and various iterations of in-memory storage; and cache memory.

The main types of storage media in use today include hard disk drives (HDDs), solid-state storage, optical storage and tape. HDDs are widely used storage in personal computers, servers and enterprise storage systems, but SSDs are starting to reach performance and price parity with disk.

Optical data storage is popular in consumer products, such as computer games and movies, and is also used in high-capacity data archiving systems. Intel optane memory is a smart technology that accelerates computer’s responsiveness. It accesses your computer’s frequently used documents, pictures, videos, and applications quickly and remembers them after you power off enabling you to create, game, and produce with less waiting. This increases productivity across the board tailored to an individual’s everyday usage.

IMPORTANCE OF DATA STORAGE

Underscoring the importance of storage is a steady climb in the generation of new data, which is attributable to big data and the profusion of internet of things (IoT) devices. Modern storage systems require enhanced capabilities to allow enterprises to apply machine learning-enabled artificial intelligence (AI) to capture this data, analyze it and wring maximum value from it. Larger application scripts and real-time database analytics have contributed to the advent of highly dense and scalable storage systems, including high-performance computing storage, converged infrastructure, composable storage systems, hyper-converged storage infrastructure, scale-out and scale-up network-attached storage (NAS) and object storage platforms.

In these days the quantity of data created per second is very large. Data stream real time analysis is required to manage this large data. Through proper analysis we can get crucial data, through this we can predict network traffic, intrusion related activity weather log records or click systems in web exploring manufacturing process, call details records, email, blogging, twitter posts and others. Data generated from stream is just snapshot of stream data. Snapshot is based on time interval. The main of algorithms is usage of resources. Resources can be memory or time. In stream database, to perform stream mining we must consider accuracy, amount of space, time required to learn from training examples for getting prediction. Data is large and growing. There are important patterns and trends in the data. We don't fully know where to look or how to find them. Big data analysis is most important because the data is continuously changing based on interval of time to store Big data most of companies are using cloud setup.

Big data storage space is a storage framework that is designed especially to store, take care of as well as retrieve massive quantities of information, or huge data. Big data storage enables the storage space and sorting of huge data as if it can easily be accessed, used as well as refined by applications as well as services dealing with large information. Big Data has actually gotten much focus from the academic community and also the IT market. Huge data has obtained much attention from the academia and also the IT industry. Currently, over 2 billion people globally are connected to the Net and also over 15 billion individuals very own smart phones. By 2020, 50 billion tools are expected to be linked to the internet. At this moment, predicted information manufacturing will be 44 times higher than that in 2015. As info is transferred and shared at light speed on optic fiber as well as wireless networks, the quantity of information and rate of market development boost. Nonetheless, the fast development rate of such big information generates many difficulties such as the quick growth of information, transfer speed, diverse data as well as safety and security.

III. RESULTS AND DISCUSSION

MERITS AND DEMERITS OF ONLINE DATA STORAGE

Merits of Online Data Storage

- **Data storage saving:** By storing your data online you are reducing the burden of your hard disk, which means you are eventually saving disk space.
- **Worldwide accessibility:** This is the main advantage of online data storage. You can access your data anywhere in the world. You don't have to carry your hard disk, pen drive or any other storage device.
- **Data safety:** You cannot trust your HDD and storage device every time because it can crash anytime.

In order to make your data safe from such hazards you can keep it online.

- **Security:** Most of the online storage sites provide better security.
- **Easy sharing:** you can share data faster, easy and secure manner.
- **Data recovery:** online data storage sites provide quick recovery of your files and folders. This makes them more safe and secure.
- **Automatic backup:** you can even schedule automatic backup of your personal computer in order to avoid manual backup of files.

Demerits of Online Data Storage

Online data storage has its negative aspects however if managed well you can definitely prevent them. Several of them are as complies with:

- Inappropriate handing can cause trouble: You have to require your user-id and password safe to safeguard your information as if someone knows or perhaps presume your qualifications, it might lead to loss of information. Use facility passwords and also try to prevent storage them in your personal storage devices such as pen drive and HDD.
- Select trustworthy resource to avoid any type of danger. To access your files almost everywhere the only thing you need is net link. If you do not obtain internet connection someplace after that you will certainly end up with no gain access to of information despite the fact that it is securely stored online.

HIGHLIGHTS OF THE DATA STORAGE TECHNOLOGIES

• **Redundant Array of Independent Disks (RAID):** This technology was developed to address the cost, performance and availability requirements of data. It continues to evolve today, and is used in all storage architectures such as DAS, SAN, etc. (See below for RAID Chart.)

• **Direct-Attached storage (DAS):** This type of storage connects directly to the server (host) or a group of servers in a cluster. Storage can be either internal or external to the server. External DAS alleviated the challenges of limited internal storage capacity.

• **Storage Area Network (SAN):** This is a dedicated, high-performance Fiber Channel (FC) network to facilitate block-level communications between servers and storage. Block storage is just that: evenly sized blocks of data. Database servers can often take advantage of block storage systems. Storage is partitioned and assigned to a server for accessing its data. SAN offers scalability, availability, performance and cost benefits compared to DAS.

• **Network-Attached Storage (NAS):** This is dedicated storage for file servicing applications. Unlike SAN, it connects to an existing communication network (LAN) and provides file access to heterogeneous clients. This is the most familiar kind of storage – it's what we interact with most on a daily basis. Users of file storage have access to files and can read and write to either the whole file or a part of it. File systems are what operating systems provide on all of our personal computers. In a shared environment file storage is often seen as a network drive. Because it's purposely built for providing storage to file server applications, it offers higher scalability, availability, performance and cost benefits compared to the general-purpose file servers.

• **Internet Protocol SAN (IP-SAN):** IP-SAN is a convergence of technologies used in SAN and NAS. IP-SAN provides block-level communication across a local or wide area network (LAN or WAN), resulting in greater consolidation and availability of data.

Object Storage

Object storage is possibly the least familiar sort of storage to the majority of people. Object storage space doesn't give accessibility to raw blocks of data and also does not supply file-based access. Object storage space provides accessibility to whole objects, or balls of data as well as normally does so with an API particular to that system. Unlike documents storage, object storage typically does not permit the ability to write to one part of a file. Objects have to be upgraded all at once system. Three of the most usual business things storage systems are Amazon's S3, Google's Cloud and Microsoft's Azure. Perfect use cases consist of backups, archiving as well as static Web web content like images and also manuscripts. One of the major benefits of object storage space systems is their capacity to dependably store a large amount of data at reasonably inexpensive.

IV. CONCLUSION

Data storage is the recording (storing) of information (data) in a storage medium. DNA and RNA, handwriting, phonograp hic recording, magnetic tape, and optical discs are all examples of storage media. Recording is accomplished by virtually any form of energy. Electronic data storage requires electrical power to store and retrieve data. In this paper, a comprehensive overview is provided on data storage.

V. FUTURE SCOPE

I recommend using online data storage such as Cloud Data Storage which provides all the advantages of Online Data Storage. To get the highest benefit of Online Data Storage take the necessary security measures and precautions. Like Amazon Webservices and Azure who implement Cloud Technology to store data.

VI. REFERENCES

[1]. https://en.wikipedia.org/wiki/Data_storage

[2]. https://searchstorage.techtarget.com/definition/storage

[3]. https://beginnersbook.com/2013/04/advantages-and-disadvantages-of-online-data-storage/

[4]. http://dataarchivecorp.com/evolution-of-data-storage/

[5]. Patterson, David A.; Hennessy, John L. (2005). Computer Organization and Design: The Hardware/Software Interface (3rd ed.). Amsterdam: Morgan Kaufmann Publishers. ISBN 1-55860-604-1. OCLC 56213091.

[6]. Anusha Medavaka, “An Overview of Security Mechanisms Towards Different Types of Attacks” in “International Journal of Scientific Research in Science and Technology”, Vol. 4, Issue No. 10, October-2018ISSN : 2395-602X]

[7]. Storage as defined in Microsoft Computing Dictionary, 4th Ed. (c)1999 or in The Authoritative Dictionary of IEEE Standard Terms, 7th Ed., (c) 2000.

[8]. Anusha Medavaka,“A study on the process of hiding protective information from the big data processing databases” in “International journal of basic and applied research”, Vol. 9, Issue No. 6, June-2019ISSN : 2278-0505]

[9]. "Primary Storage or Storage Hardware" (shows usage of term "primary storage" meaning "hard disk storage") Archived 10 September 2008 at the Wayback Machine. Searchstorage.techtarget.com (2011-06-13). Retrieved on 2011-06-18.

[10]. Anusha Medavaka,“K-Means Clustering Algorithm to Search into the Documents ContainingNatural Language” in “International Journal of Scientific Research in Science and

Technology", Vol. 3, Issue No. 8, Dec-2017ISSN : 2395-602X]

[11]. Anusha Medavaka,Siripuri Kiran, "Implementation of dynamic handover reduce function algorithm towards optimizing the result in reduce function" in "International Journal of Academic Research and Development", Vol. 4, Issue No. 4, July-2019ISSN : 2455-4197]

Cite this article as :

Sumalatha Sriramoju, " A Comprehensive Review on Data Storage", International Journal of Scientific Research in Science and Technology(IJSRST), Print ISSN : 2395-6011, Online ISSN : 2395-602X, Volume 6, Issue 5, pp. 236-241, September-October-2019. Journal URL : http://ijsrst.com/IJSRST196523