A review at storage systems and pipelining

Yogesh Sigar
yogesh.sigar2017@vitbhopal.ac.in
VIT Bhopal University, Bhopal, Madhya Pradesh

Shubham Malhotra
shubham.malhotra2017@vitbhopal.ac.in
VIT Bhopal University, Bhopal, Madhya Pradesh

Soyal Sharma
soyal.sharma2017@vitbhopal.ac.in
VIT Bhopal University, Bhopal, Madhya Pradesh

Shetty Manvith
shetty.manvith2017@vitbhopal.ac.in
VIT Bhopal University, Bhopal, Madhya Pradesh

P. Sanjeevi
sanjeevi.p@vitbhopal.ac.in
VIT Bhopal University, Bhopal, Madhya Pradesh

ABSTRACT

An external storage device is a device that contains all the addressable storage that is not inside computers main memory. It has types including removable and non-removable, permanent and non-permanent. External storage is often used to store information that's accessed less frequently by applications running on a desktop, laptop, server or mobile devices, such as an Android or iOS smartphone or tablet. When moving large amounts of data to the cloud, providers will usually use external storage devices in a practice commonly referred to as cloud seeding. Pipelining is an implementation technique in which more than one instruction is overlapped in execution. The computer pipeline is divided into stages. Each stage performs and finishes a part of an instruction in parallel. The stages are related and connected one to the next to form a pipe - instructions enter at one end, progress through the stages, and exit at the other end. Pipelining does not decrease the time span essential for individual instruction execution. Instead, it enhances instruction throughput. The throughput of the instruction pipeline is estimated and evaluated by how often an instruction exits the pipeline. Under these constraints, the progress rate from pipelining equals the number of pipe stages. Usually, the stages will not be accurately well balanced; besides, the pipelining itself involves some overlook on the inner aspect. Pipelining is usually viewed as reducing the C.P.I. (Cycles Per Interval)

Keywords— *Storage systems, Cloud, RAM, ROM, Magnetic disks, Optical drives, Pipelines, Instruction, Software, graphic, HTTP*

1. INTRODUCTION

An external storage system is made by combining capabilities of external storage devices such as hard drives and adding components of hardware and software to obtain a reliable, high performance and easy to manage storage system [1]. Now an external storage device also known as secondary or auxiliary storage is devices which contain all the addressable data which is not in the computer mainframe storage. It can be removable or non-removable, temporary or permanent and wired or wireless. [2].

Imbedding concurrency in a computer system by implementing it in the form of a pipeline is called pipelining. Which is a configuration of independent autonomous units? All of them are dedicated to performing a subfunction in an overlapped mode with each other. A segment or an autonomous unit of a pipeline is also known as a facility-segment or a pipeline segment. This technique has emerged as a vital aspect of computer architecture especially where scientifically oriented computers are concerned [3].

A data pipeline is a set of data processing element which is connected in series, in this way output of the previous element is input to the next one. Most of the time these elements are executed in a parallel or time-sliced fashion. A little amount of buffer storage can be included in between of the elements.

The computer-related pipeline includes:
- Instruction pipelines
- Graphic pipelines
- Software pipelines
- HTTP pipelines [4]

2. EXTERNAL STORAGE SYSTEM

Computer data storage also termed as storage or memory is defined as the technical knowledge of recording media and computer components that are persistently used to retain digital data and cloud storage systems is commonly used [26-35]. It is the main and most important function of computers. THE CENTRAL PROCESSING UNIT also termed as CPU of a computer is an external device which can plot the data by performing various computations [5].

The first storage device was launched into the market by IBM in the year 1956. Since from that very moment we are able to see a very huge progress in the field of hard disk drive (HDD) and this also known as the base of the industry of storage systems. These systems are being constructed by obtaining a storage device that is HDD and further adding layers of software and hardware in order to get a highly reliable, high performance and easily manageable device. They are also known as storage subsystems. Due to the increase in the modern technologies over the period of time, storage systems have started including advanced technologies which further add a good value to the HDD. They even support a variety of added services, management systems, and interface alternatives. Due to this, they are also considered up as the parts of the storage system [6].

2.1 Cache

It is defined as the storage space or buffer which is a high-speed memory device that keeps high-speed data from the processor before it is being sent to the RAM. RAM works at a very slow rate as compared to the processor that send the data.

2.2 Random Access Memory

It is also called RAM. It is defined as an integrated circuit memory chip that allows the information to be stored or used and the data is being used by the operating system and other programs which our computer system is still in use. RAM can be termed to be said as volatile in nature, meaning that it exists only while the system is switched on and changes along with the systems immediate requirements. It is accessed with high speed of the computer system and software and keeps data and applications along with the programs. The graphic user interface also termed as GUI is stored in the RAM. Once our system is turned off, RAM is cleared and only start up again when our system is rebooted. It holds constructs and caching regions which further constitute a wealth of state related information. Thus, the availability of that data, along with the techniques to further recover it, gives us a new method for an investigation [7].

External storage is further being classified in various parameters that are:

Magnetic Storage

Optical storage which further includes:
- Compact Disc
- DVD
- Blu-Ray
- M-Disc

Flash memory devices which further includes:
- Memory Card
- Memory Stick
- USB Drives [8]

2.3 Magnetic Storage

2.3.1 Magnetic Disk

Recording the data onto a disk has a peak advantage over the access time as the head a readily be moved to an appropriate place of the disk as compared to the tape which needs to be rewound or advanced. There are two types of disk that are floppy and hard disk.

The basic principle of manufacturing and recording on a floppy disk are very similar to that of a magnetic tape that is the same particulate materials on a plastic substrate [9].

2.4 Optical Data Storage

From the last decade, we are able to see the boom in optical techniques for data storage. They attain notably high storage density and low access time further not attainable by conventional recording techniques. At present, inspire of the lack of widespread commercial sources of optical recording, some unique devices have successfully been demonstrated. Video recording is one of the applications of knowledge accumulated through the development of optical data storage. It is observed that optical storage technology will be treated as a competitive contender for next-generation data storage if conventional magnetic memory technology reaches its limit [10]

2.4.1 Compact disc: It is an optical disc data storage which was co-developed by Sony and Philips in the year 1982. It was originally developed to store and play only sound recordings but later. It was adopted for data storage as well. Several various formats of the compact disc also came into picture including write-once audio and data storage (CD-R), rewritable media (CD-RW), video compact disc (V-CD) and many more. Though it was debuted on the market in 1982. They were further able to store up to 74 minutes of the high quality of music and even termed up as CD - Audio. They were supposed to be associated with the music industry and had taken up a good place in the music market. Also, as the disc was supposed to be 4.72-inch plastic disc, even though programmable as well as durable, could only perform the task of read-only [11]

2.4.2 DVD: Also termed as digital versatile disc. It is an optical recording medium with the storage capacity seven times higher than the conventional compact disc. This has only been possible because of the use of optics, shorter wavelength and larger numerical aperture which further reduces the spot diameter by a factor of 1.65 [12]

2.4.3 Blu-ray: It is called the next generation optical disc format jointly developed by the blue ray disc association (BDA), which is the group of world's leading consumer electronics, personal computer, and media Manufacturers. It was developed to create a recording, rewriting and playback of high definition video, and large storing amounts of data [13].

2.5 Flash Memory Devices
Flash memory is defined as a nonvolatile memory chip which is further used for storage as well as transferring of data between a personal computer and a digital device. The main advantage of this is that the data can be electronically programmed as well as erase. This technology is further found in us flash drives, mp3 players and solid-state drives. It can also be termed as a standalone memory storage device such as USB drive. The electronically erasable programmable read-only memory (EEPROM) is a type of data memory device which is further used to erase or write digital data. The flash memory is a distinct type of EEPROM which is further erased and programmed in large blocks. These SSDs are supposed to become up as the primary storage for most of the applications used within the enterprises, with higher latency HDDs being used, and cooler storage. These sorts of memories are more expensive as compared up with the HDDs due to high manufacturing cost and lower storage capacity. There are various cases being seen where flash memory has achieved or approaching the operating cost parity with the HDDs [14].

It further includes the following:
1) Memory card
2) Memory stick
3) USB drives

2.5.1 Flash Memory Types: There are basically two types of flash memory. The two of the types are:
1) NAND Flash memory
2) Nor Flash memory

2.5.1.1 NAND Flash memory: They have a different structure compared to nor memories. These memories can further be used in many block devices such as hard disks. While reading the NAND flash memory, the contents are paged first into a memory-mapped RAM. This further makes the memory management unit essential. Nearly, around 1990, this NAND type memory was introduced and it further brought the change in the interface of the flash device into a block access type where the bytes cannot be addressed themselves or independently. Only and only the blocks of the data also known as a page could be streamed in or out of the NAND flash. Its design and concept further helped the manufacturers to produce NAND flash denser as compared to NOR and further cutting up the cost considerably.

2.5.1.2 NOR Flash memory: NOR Flash memory is capable of reading individual flash memory cells, and it further behaves like a traditional ROM in this mode. Commands are written to the first page of the mapped memory for erasing and writing of the data as being defined by the "common flash interface" created by Intel. It gained very fast its importance and was adopted within the embedded systems as a replacement for EPROMs, further simplicity to program and use [15].

2.5.2 Advantages and Disadvantages of Flash Memory: There are various advantages or pros of using flash memory. Some of them are:

2.5.2.1 Ease of Use: Flash drives, handheld and being lightweight fits easily inside the smallest pockets, purses, and bags. They further don't require any charging or some similar sort of thing like this, they can be taken anywhere, and used without any accessories.

2.5.2.2 Speed and Versatility: Another very major benefit of USB flash drives is their efficiency and versatility. They can relatively transfer files at very high speed especially since they are not dependent on internet connections. They are even able to run various types of applications including the operating system as well.

2.5.2.3 Security: As they seem to be simple and one dimensional in appearance and size, USB flash drives can come equipped with various forms of advanced functions like security protection which is taking the major contribution. Most of the USB is equipped with password restricted access, further requiring a password to open the files within the drive.

2.5.3 There are various disadvantages of flash memory as well. Some of them are:
2.5.3.1 Malware risk: As far as we know flash drives aren't vulnerable to hacking as compared with the cloud storage drives. They are not completely clear from attacks. If the computer is infected with malware, further connecting drive to a computer may cause a risk of infecting the data to be infected. Further putting the data on the drive at risk of corruption or theft.

2.5.3.2 Vulnerability to loss and destruction: Though the compact size of the drive makes them excellent for travel and space saving, it also makes vulnerable the data to loss, destruction or theft.

2.5.3.3 Limited lifespan: The major factor further determining the lifespan of the flash drive is how it is able to erase or write the data. The flash drive can last as long as 10 years or even if it's used enough, then it can have a brief lifespan as short a few months to a year [16].

3. PIPELINING

It is an implementation technique in which multiple instructions overlap while execution. The pipeline used in computers is divided into stages. Every stage completes each part of instructions in parallel. The stages are connected in form of pipe – instruction enter at one end, progress through stages, and exit at another end.

Pipelining does not decrease individual instruction time. Instead, it increases instruction throughput. The throughput of an instruction pipeline is determined by how many times an instruction exits the pipeline. As the pipe stages are intertwined together, all stages must be ready at the same time. The time required to move one step further in the pipeline is called Machine cycle. The length of machine cycle is estimated by the time required by slowest pipe stage [17].

3.1 Pipeline hazards

In some situations, the next instruction can't be executed in the instruction cycle during its own designated cycle. These types of situations are called Hazards. They reduce the speedup performance gained by pipelining.

There are three major types of hazards:
- **Structural hazards:** they are encountered when hardware cannot support all possible combination in the overlapped execution.
- **Data hazards:** when an instruction depends on the result of a previous instruction in a way that is exposed by the overlapping of instructions in the pipeline.
- **Control hazards**: encountered through branches in pipelining and instructions that change the pc.

These hazards in a pipeline can make it necessary for a pipeline to stall and a processor can stall at two major events:
A cache miss can stall all the instruction on pipeline both before and after the instruction which can cause the miss.
Eliminating a hazard can often cause some instructions to delay and while the majority of instructions to pass. And when an instruction is stalled it will stall all the instructions coming after it. The instruction issued earlier will clear [18].

3.2 Basic concept

Pipelining is a very effective way of organizing a concurrent activity in a computer system. The basic idea is extremely simple. It is often encountered in manufacturing plants, where pipelining is usually known as an assembly line operation. An assembly line is mostly used in car manufacturing plants where at one station one part of the car is fixed. Consider this idea of pipelining used in computer systems. The processor executes a program by collecting and executing instructions one by one [19].

3.3 Types of pipelining

3.3.1 Instruction pipelining: In this type, instructions can be executed through three phases of an instruction cycle. This technique increases the throughput of the system.

The three phases are:
- **Fetch:** read the instruction from the memory.
- **Decode:** it decodes the instruction and collects the source operand.
- **Execute:** perform the operation [20]

3.3.2 Software pipelining: Software pipelining is a technique used in optimizing loops. It is parallel in working with hardware pipelining.

it is a type of out of order execution except that the reordering is done by a compiler or assembly code written by the user(programmer) instead of the processor Some architectures have explicit support for software pipelining, notably Intel's IA-64 architecture [20].

3.3.3 HTTP pipelining: It is a technique in which multiple HTTP requests are sent to a single TCP server without waiting for corresponding responses. This results in dramatic improvements in loading time of HTML pages especially over high latency connections such as satellite internet connections. The speedup is less considerate on broadband connections [20].

3.3.4 Graphic pipelining: It is a conceptual model of what needs to be done in a graphical system to render 3d imagery into 2d imagery. It is a process to turn 3d models into what of a computer display. It is usually used in real-time rendering. (Figure 1) [20].

Fig. 1: Graphic pipelining [26]

4. CONCLUSION

In this journal with the help of the journals available, we have tried to research the fields of computer architecture namely external storage devices and pipelining. External storage is an easy way to add additional storage or options to your computer without having to open the computer. In this journal, we have tried to discuss important aspects of both this topic. Pipelines and pipelining also apply to computer memory controllers and moving data through various memory staging places.

5. REFERENCES

[1] R. J. T. Morris B. J. Truskowski, Page 205, IBM Systems Journal, Vol 42, No 2, 2003.

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

[3] C. V. Ramamurthy And K. H. Kim, Page 285, University of California, National Computer Conference, 1974.

[4] https://en.wikipedia.org/wiki/Pipeline_(computing)

[5] en.Wikipedia.org/data storage

[6] R. J. T. Morris B. J. Truskowski, The Evolution of Storage System, IBM Systems Journal, Vol 42, No 2, 2003

[7] Glory V. Udoh., memory management in a computer system, M.Sc., ISSN: 2231-5381

[8] en.wikipedia.org (Assessed on 31-10-2018)

[9] Rees D. Rawlings, Magnetic material, Magnetic drives, 2009.

[10] Di Chan. Zook, An overview of optical data storage technology, Proceedings of the IEEE (Volume: 63 Issue:8 , Aug. 1975)

[11] Rajendra Munoo, Master of Arts (Library and Information Management) University of Sheffield, December 2000

[12] Keeps A. Schumer Immune, SMPTE Periodical - The Digital Versatile Disc (DVD): System Requirements and Channel Coding, SMPTE Journal (Volume: 105, Issue: 8, Aug. 1996)

[13] T. Ravi Kumar, Dr. R. V. Krishnaiah, an optical disk with blue ray technology, IJCEA, JULY-SEP-2013

[14] Tom Coughlin, Roger Hoyt, and Jim Handy, Digital Storage and Memory Technology, November 2017,

[15] Marco A. A. Savio, Frank R. Chu, Anand Kulkarni, and Robert Selinger, NAND flash memory and its role in Storage architectures, Proceedings of the IEEE | Vol. 96, No. 11, November 2008

[16] https://www.cybrary.it/2018/05/pros-and-cons-of-usb-flash-drives/

[17] web.cs.iastate.edu

[18] Carl Hamachi, Zonke Vanesa, Sarwat zaky, Computer organization fifth edition, 2002, 453-460

[19] www.studytonight.com

[20] en.wikipedia.com/types of pipelines

[21] P. Sanjeevi and P. Viswanathan, 'NUTS scheduling approach for cloud data centers to optimize energy consumption', *Computing (Springer)*, Vol. 99, No. 12, pp. 1179-1205, 2017.

[22] P. Sanjeevi and P. Viswanathan, 'Workload Consolidation Techniques to Optimize Energy in Cloud: Review', *Int. J. of Internet Protocol Technology*, Vol. 10, No. 2, pp. 115–125, 2017.

[23] P. Sanjeevi and P. Viswanathan, 'DTCF: Deadline Task Consolidation First for energy minimization in cloud data centers', *International Journal of Networking and Virtual Organizations*, Inderscience, Vol. 19, No. 3, pp. 209–233, 2017.

[24] P. Sanjeevi and P. Viswanathan, 'Employing Smart Homes IoT Techniques for Dynamic Provision of Cloud Benefactors', *Int. J. of Critical Computer-Based Systems*, Inderscience, Vol. 7, No. 3, pp. 209–224, 2017.

[25] P. Sanjeevi and P. Viswanathan, 'A survey on various problems and techniques for optimizing energy efficiency in cloud architecture', *Walailak Journal of Science and Technology*, Vol. 14, No. 10, 2017.

[26] P. Sanjeevi, P. Viswanathan, M. R. Babu, and P. V. Krishna, 'Study and Analysis of Energy Issues in Cloud Computing', *International Journal of Applied Engineering Research*, Vol. 10, No. 7, pp. 16961-16969, 2015.

[27] P. Sanjeevi, G. Balamurugan, and P. Viswanathan, 'The Improved DROP Security based on Hard AI Problem in Cloud', *Int. J. of Internet Protocol Technology*, Vol. 9, No. 4, pp. 207-217, 2017.

[28] P. Sanjeevi and P. Viswanathan, 'A green energy optimized scheduling algorithm for cloud data centers', *IEEE International Conference on Computing and Network Communications, Trivandrum*, pp. 941-945, 2015.

[29] P. Sanjeevi and P. Viswanathan, 'Towards energy-aware job consolidation scheduling in the cloud', *International Conference on Inventive Computation Technologies (ICICT 2016), IEEE Xplore*, pp. 361- 366, 2016.

[30] G. Kesavan, P. Sanjeevi and P. Viswanathan, 'A 24-hour IoT framework for monitoring and managing home automation', *International Conference on Inventive Computation Technologies (ICICT 2016), IEEE Xplore,* pp. 367-371, 2016.