An analysis of data storage and retrieval of file format system

R. Mariappan and B. Parthasarathy
Dept. of Master of Computer Appl., Mailam Engg. College, Mailam-604 304, Villupuram Dt., Tamil Nadu, India
rm_mca@sify.com

Abstract: The present work focuses on the analysis and examination of different types of file format system, storage devices and storage & retrieval systems which are most frequently used in the various applications. It is focused on the file format to identify the right file to right applications for the fast storing and accessing speed with less memory space along with the suitable file format on specific storage devices, RAM and processors. The analysis evolves based on different types of file format on different storage devices taking into account different research perspectives. The outcome of the work motivates multimedia applications users and developers for choosing appropriate file format, storage devices, RAM and processors to undertake their applications in a successful manner.

Keywords: Data storage, RAM, file format systm

Introduction

Data is a major aspect of the computer field in both hardware and software. The data storage in the system can store data for easy and fast access at any time and anywhere. Data is organized for storage in a hierarchy in the form such as bits, bytes / characters, fields, records, files and databases and data warehouse. The storage media and file formats are all vital importance in order for computer systems to store and organize the various kinds of data in effectively and efficiently.

In the past decade the traditional users and developer's applications that supported with only text and image file formats. These files could be handled by disk operating systems. In the modern GUI world the software and hardware developers are facing a challenging trouble to use different types of file formats and storage devices. ,Storage media are of vital importance in order for computer systems to store and organize information effectively and efficiently (Papadopoulos & Zhang, 2007). Storage media compose a hierarchy with primary storage (registers, cache and main memory) tertiary storage (tapes, optical disks and flash memory) and secondary storage (hard disks).

Several types of data storage exist in most computer systems (Abrahma Silberschatz *et al.*, 2002). They are classified by the speed with which they can access data, by their cost per unit of data to buy the memory, and by their reliability. Among the media available are cache, main memory, flash memory, magnetic disks , optical disk and magnetic tapes.

It is clear that the upload time increases for files stored in data base as file size increases. But download time of files whether stored in database or file system is almost same since it is faster retrieving the search results. The digital entities are images of reality, stored as data (bits) that must be interpreted for display by applying information and knowledge. The preservation of digital entities requires the ability to characterize data, information, and knowledge (Regan W. Moore, 2000).

Data integrity is a fundamental aspect of storage security and reliability. The reliable access to data is prerequisite for most computer system and applications. File system consistency is one of the common ones. Data storage patterns are being employed for efficient retrieval of data from secondary storage devices (Gopalan Sivathanu *et al.*, 2005).

It is meant to operate in a COMOHO (commercial office, mobile office, and home office) environment, i.e. computing anywhere, anytime. The main desktop office machine (CO) holds all files. Since Cyber All will store all personal information, e.g. documents, photos, and videos, this data needs to be valid and hence understood in an *indeterminate* future (Cyber All, 2006. A Personal store for everything).

For many years, when we used a word-processing application, files were stored in a format that was fully understood only by the application we used. WordPerfect stored WordPerfect files, Microsoft Word stored Word files, and so on.

The Chemical Abstracts Service (CAS) has built files which require hundreds of millions of bytes of storage. To cope with this size and variability, CAS has developed an internal standardized storage format and has implemented it on the IBM System/360. This format, which is called Standard File Format Implementation of standardized files, has permitted the CAS system to operate with greater generality and flexibility (Anzelmo, 1971).

Data must be stored and documented in a meaningful way that is absolutely clear (Joseph C. Slater, 2007). Regarding graphics, use EPS or PDF for all graphics that are not a pure photographs. Converting a photograph to a JPEG or TIFF is a simple task on a Mac R. On other platforms there is free software available, or Adobe Acrobat R can be used. Magnetic - based media such as floppy disks and hard drives have a half life of between five and seven years depending on the storage conditions.

Need of the Study

In the fast emergent world, we need a good storage media and storage software. In the earlier days, the data were stored and retrieved in various methods. Those methods needed many storage spaces in both permanent (Floppy, Hard Disk etc.,) and temporary (RAM, VRAM

etc.,) storage and be also took more time to store and retrieve on the devices.

In the recent and emerging trends in information technology has been totally changed for the users needs and requirements of choices are large capacity of storage devices and file format to meet and challenges in their many applications. Computer data storage and retrieval system helps as per the need of the consumers and the users.

In the middle of twentieth century, after invention of electrical, electronic and computer instrument, the requirement of data and database usage also tremendously increased. Then they needed huge amount of data storage media and fastest data store and access method.

Fig.1

Statement of the problem

Data is a major aspect of the computer field in both hardware and software. The computer peripherals are designed based on certain kind of data. Since 1946, in the computer field the data storage media is a major component of hardware in order to store and access the data. Only in 1980's the computer Hardware Manufacturers introduced many kinds of technologies to store huge amount of data in same size of floppy disk and Hard disk etc.

Before the introduction of the windows operating system the data were processed based on Disk Operating System only. In the DOS era, the Data were

constructed by Bits and Bytes. But, after the introduction of GUI, the computer world has come to the peak. Generally, the computer software professionals need to know the apt technology of store and access of the data.
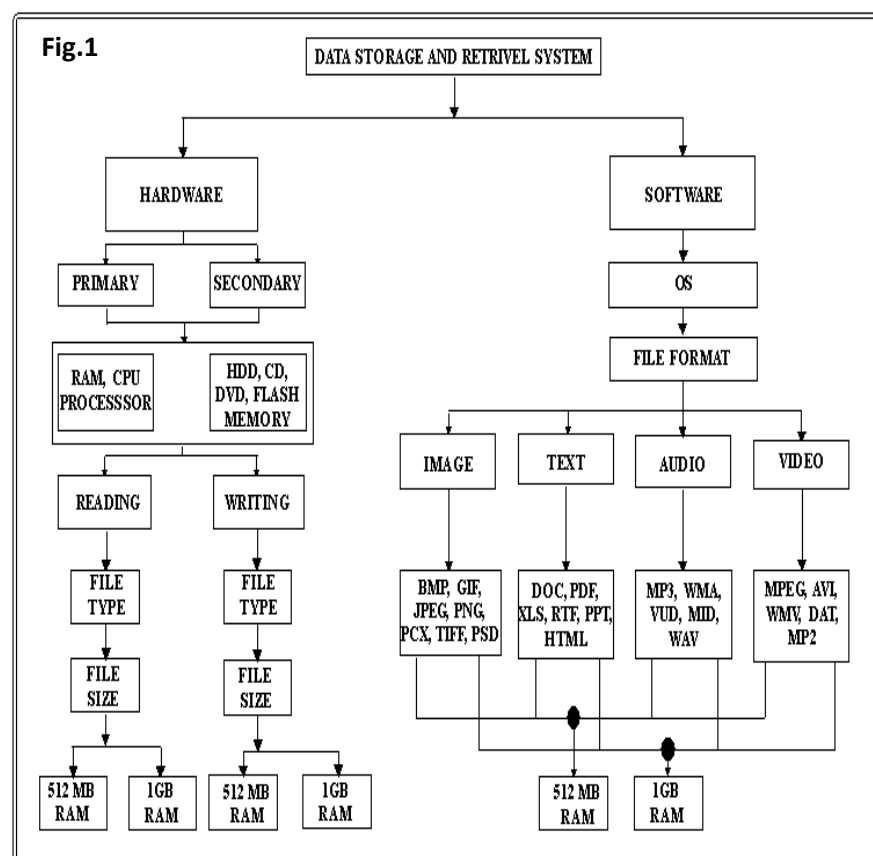
Objectives of the study

The following are the objectives of this research.
- To study and analyze the various file format such as image, text, audio, video and others.
- To understand the suitable file format with less memory space occupied.

Methodology

In the present study and analysis the descriptive research design has been followed. This research proposed to collect the following details about the data, data collection, file format storage device/ media and supporting software and hardware for the

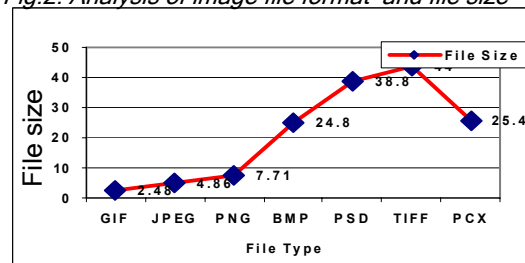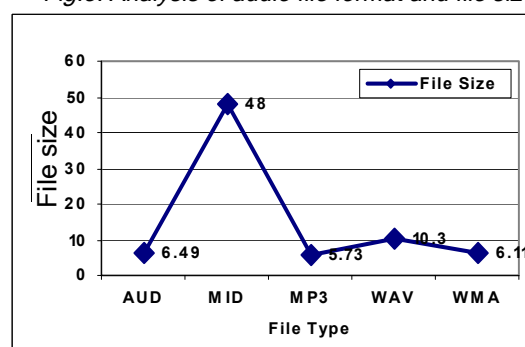Fig.2. Analysis of image file format and file size

Fig.3. Analysis of audio file format and file size

purpose of the analysis of the study.

The following different types of file format, storage devices media are to be considered in the analysis of the researches data storage and retrieval systems
- Image file formats - GIF, JPEG, PNG, BMP, EPS, PSD,TIFF
- Audio File formats - WAV,MID,AUD, WMA, MP3
- Video File format - AVI, DAT, MP3, MPEG, WMV
- Text File format - DOC,XLS, PDF, HTML, RTF,TIFF
- Source File format and Database file format etc.,

Fig.1-5 shows the analysis of Data storage and retrieval system (Flow chart is given).

Results and discussion

Based on the analysis of the study, the following comparison analysis table has been made. The following

table and figures show the result of various types of file format stored in different types of storage devices.
Table 1: Image File format
Table 2: Audio File format
Table 3: Video File format
Table 4: Text File format
From this analysis of the Tables and Figures, it gives various performances of different types of file format, file size and minimum time taken to read and write for access speed needed to store on the various storage devices.

Conclusion

Data storage and retrieval of file format system that can be used for data processing, multimedia and other applications with supportive of software, hardware technology effectively and also efficiently. It gives various results for each and every application, the file format GIF for digital images and well suited for graphics, JPEG for Photograph images on Web applications, PDF for Text on document, MP3 for Audio application, and WMV for Video application and so on. The above file format storage and retrieval of multimedia data applications and constant use on the net and others.

Suggestion and future work

On the basis of the analysis, the suggestion of this result is to give a better idea to the software and hardware professionals and users can easily identify that, which is the best file format to store the relevant data such as image, audio, video, text data file format depend on the storage devices for the reading and writing time of access speed which is

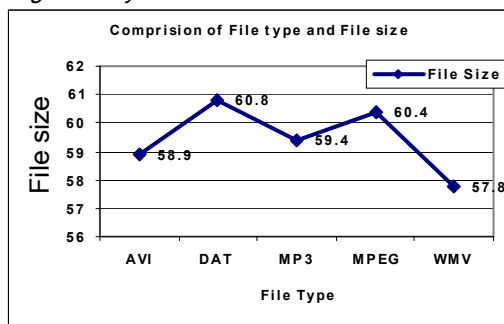Fig.4. Analysis of video file format & file size

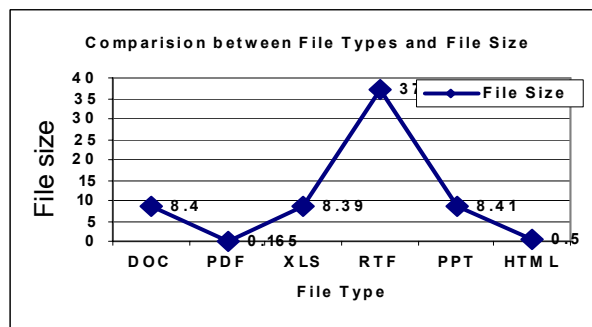Fig 5. Analysis of text file format and file size

Table 1. Analysis of image file format and file size

S.No	System config.	Process speed	File type	file Size
1	P - IV	1.60 G.Hz	GIF	2.48
2	P - IV	1.60 G.Hz	JPEG	4.86
3	P - IV	1.60 G.Hz	PNG	7.71
4	P - IV	1.60 G.Hz	BMP	24.8
5	P - IV	1.60 G.Hz	PSD	38.8
6	P - IV	1.60 G.Hz	TIFF	44
7	P - IV	1.60 G.Hz	PCX	25.4

Table 2. Analysis of audio file format and file size

S.No	System config.	Process speed	File type	File size
1	P - IV	1.60 G.Hz	AUD	6.49
2	P - IV	1.60 G.Hz	MID	48
3	P - IV	1.60 G.Hz	MP3	5.73
4	P - IV	1.60 G.Hz	WAV	10.3
5	P - IV	1.60 G.Hz	WMA	6.11

fast and efficiently and effectively in the selective storage and retrieval of data for multimedia applications.

In future there might be storage devices, file format which will have a large storage capacity and also with quick access time. Let us firmly believe that the evaluation of invention shall take us there.

References
1. Abrahma Silberschatz, Hentry F. Korth and S. Sudharssan (2002) Database system concept. 5th Edition, Tata McGraw Hill.
2. Anzelmo D (1971) *Computers. IEEE Trans. on A Data-Storage Format for Information System Files.* C-20 (1), 39-43.
3. Gopalan Sivathanu, Charles P. Wright, and Erez Zadok (2005) Ensuring data integrity in storage techniques and applications. Stony Brook Univ., Computer Sci. Dept.
4. Joseph C. Slater (2007) How to archive your/Thesis/ Dissertation/ Project Data. Write State Univ.
5. Papadopoulos Marios Hadjieleftheriou AN and Zhang D (2007) Disk storage and basic file structures. Chapter 19 of Handbook of Database Technologies. Hammer J & Schneider M (eds.), CRC Press.
6. Regan W. Moore (2000) The preservation of data, information and knowledge. Univ. of California, Sandiego.

Table 3. Video file format & file size analysis

S.No	System configuration	Process speed	File type	File size
1	P - IV	1.60 G.Hz	AVI	58.9
2	P - IV	1.60 G.Hz	DAT	60.8
3	P - IV	1.60 G.Hz	MP3	59.4
4	P - IV	1.60 G.Hz	MPEG	60.4
5	P - IV	1.60 G.Hz	WMV	57.8

Table 4. Analysis of text file format and file size

S.No	System config.	Process speed	File type	File size
1	P - IV	1.60 G.Hz	DOC	8.4
2	P - IV	1.60 G.Hz	PDF	0.165
3	P - IV	1.60 G.Hz	XLS	8.39
4	P - IV	1.60 G.Hz	RTF	37
5	P - IV	1.60 G.Hz	PPT	8.41
6	P - IV	1.60 G.Hz	HTML	0.5