

DATA BACKUP ON: CLOUD COMPUTING TECHNOLOGY IN DIGITAL LIBRARIES PERSPECTIVE

Mrs. Saguna Bhargav Vora

Department of Computer Science & Engineering
Changwon National University,
Changwon, South Korea.
shagunvora@gmail.com

Mrs. Joystna Gajanan Anandache

Department of Computer Science & Engineering
Changwon National University,
Changwon, South Korea.
joystna.anandache@gmail.com

Abstract: Latest technological development has brought a dramatic change in library science. Information technology has impacted positively on library, information system and services they provide for users. The libraries have been automated, networked and now moving towards paper less. Now libraries may soon be building and managing their own data centers. The Provision of library and maintenance of infrastructure for Web based digital library faces several challenges.

There are different platforms in Library science for attaining economy in information handling. The basic principle of cloud computing entails the reduction of in-house datacenter's and Information Technology infrastructure capability. Universities and Colleges are the core of innovation through their advanced research and development. Subsequently, Higher Institutions may benefited greatly by implementing the Digital library with cloud computing, including cost cutting.

This paper overviews the basic concept of cloud computing used in libraries and how cloud computing actually works is illustrated in this paper. This paper explores some of the security issues as well as data back up over existing system surrounding data location, mobility and availability.

This paper also includes new method to back up data my installing new server in existing system and can restore in data corruption issue.

INTRODUCTION

There is no alternative of library for the enrichment of human civilization. A library is not only a collection of books, which has treated as the source of knowledge and information but also comes as the reflection of any community. We have seen the rapid advancement in the field of communication, science, and technology, which has a great impact on many issues. The concept of digital library is nothing new, but cloud computing has brought new opportunities for the advancement of digital library. The end user does not have to worry about the resource or disk space in cloud computing. Cloud computing has already proved its necessity, which has considered as the upcoming pattern of computing in coming days. Here, I will emphasize on the issue of cloud computing and its role for digital libraries.

Digital library is a development-oriented hardware and software integration platform. Cloud computing offers

information retrieval systems, particularly digital libraries and search engines, a wide variety of options for growth and reduction of maintenance and encourages efficient resource use. These features are particularly attractive for digital libraries. Cloud infrastructure allows rapid growth in collection size and support a larger user base.

This paper proposed to apply Cloud Computing in digital libraries. By establishing a pubic cloud among many digital libraries, it not only can conserve library resources but also can improve its user satisfaction.

Cloud Computing is a completely new Information Technology and it is known as the third revolution after PC and Internet in IT. The basic principle of Cloud Computing is collecting large quantities of information and resources stored in personal

computers, mobile phones and other equipment, Cloud Computing is capable of integrating them and putting them on the public cloud for serving users.

Today we are living in the age of information. Information technology play very vital role in library science. For collection, Storage, organization, processing, analysis of information. Library filed facing many challenges in the profession due to applications of information technology. The emergence of e-publications, digital libraries, internet usage, web tools applications for libraries leads to the further developments in library profession.

Library is a growing organism. Automation is a technique to make a system automated, i.e. self service. A properly computerized library will help its user with quick and prompt predominantly by computerization. Thus library automation the repetitive and clerical job involved in the function and services of the libraries. Now the computers have become capable of introducing in operations, processes, techniques and methods of library. It is concluded that the system like library though it is fast moving, we can achieve high grade performance

I. WHAT IS CLOUD COMPUTING

"A cloud is a type of parallel and distributed system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resource(s) based on service-level agreements established through negotiation between the service provider and consumer." Figure 1 denotes resources of cloud computing.

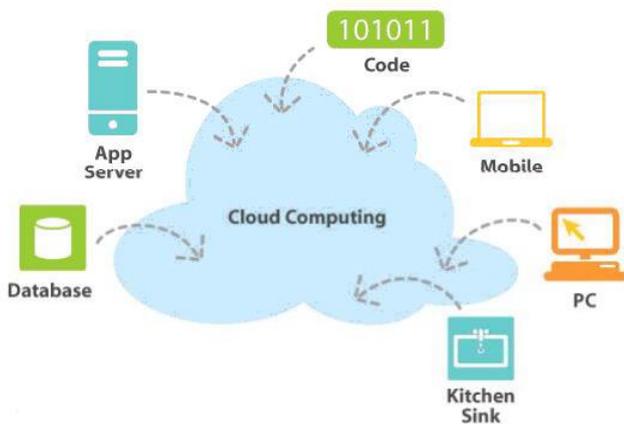


FIGURE1: RESOURCES OF CLOUD COMPUTING

Cloud computing seems to offer some incredible benefits for communicators: the availability of a software applications, quick processing, unlimited storage, and the ability to easily share and process information. All of this is available through your browser any time you can access the Internet.

Three typical kinds of cloud computing services are:

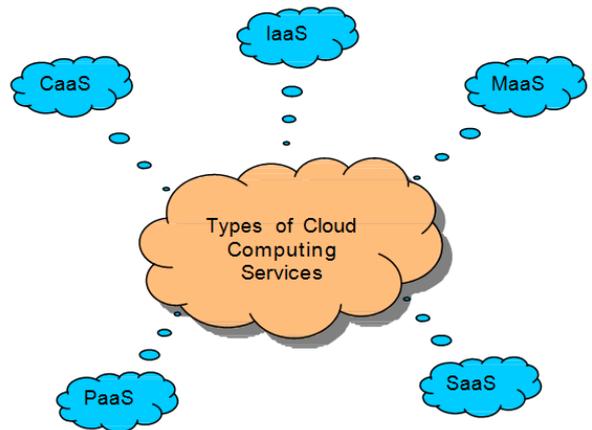


FIGURE2: TYPES OF CLOUD COMPUTING

COMMUNICATION AS A SERVICE (CAAS): Allow for certain messaging tools viz voice over IP (VOIP), Instant Messages (IM) and Video Conferencing.

INFORMATION AS A SERVICE (IAAS):Clouds that provide large and affordable computing resources that run enterprise programs, which is known as Infrastructure as a service(IaaS), Allows customer to maintain their application while loading dataon the IaaS infrastructure.

MONITORING AS A SERVICE (MAAS): Outsourcing of security service to a third party security team.

PLATFORM AS A SERVICE (PAAS): Local Storage Clouds also known as a Platform as a Service(PaaS), Meant for web-based development infrastructure.

SOFTWARE AS A SERVICE(SAAS):Application Clouds also called Software as a Service (SaaS), When a software vendor supplies software over a network as opposed to the typical distribution of installation of individual computers.

II. ROLE OF CLOUD COMPUTING IN LIBRARIES

Cloud computing is a completely new in technology and it is known as third revolution after PC and Internet. Cloud computing is an enhancement of distributed computing, parallel computing, grid computing and distributed databases. Cloud computing has large potential for libraries. Libraries may put more and more content into the cloud. Using cloud computing user would be able to browse a physical shelf of books, CDs or DVDs or choose to take out an item or scan a bar code into his mobile device. All historical and rare documents would be scanned into a comprehensive, easily searchable database and would be accessible to any researcher. Many libraries already have online catalogues and share bibliographic data with Online Computer Library Center. More frequent online catalogues are linked to consortium that share resources.

Data storage cloud is a main function of libraries, particularly those with digital collections, storing large digital files can stress local server infrastructures. The files need to be backed up, maintained, and reproduced for patrons. This can strain the data integrity as well as hog bandwidth. Moving data

to the cloud may be a leap of faith for some library professionals. A new technology and on the surface it is believed that library would have some control over this data or collections. However, with faster retrieval times for requests and local server space it could improve storage solutions for libraries. Cloud computing or IT infrastructure that exists remotely, often gives users increased capacity and less need for updates and maintenance, and has gained wider acceptance among librarians

III. ARCHITECTURE OF CLOUD COMPUTING IN DIGITAL LIBRARY

The architecture behind cloud computing is a massive network of "cloud servers" interconnected as if in a grid running in parallel, sometimes using the technique of virtualization to maximize computing power per server.

The following figure 3 represents the architecture of cloud computing in digital library.

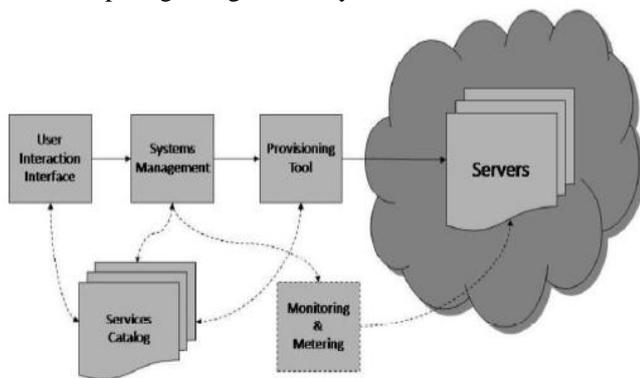


FIGURE3: CLOUD COMPUTING ARCHITECTURE FOR DIGITAL LIBRARY

A front-end interface allows a user to select a service from a catalogue. This request gets passed to the system management which finds the correct resources, and then calls the provisioning services which carves out resources in the cloud. The provisioning service may deploy the requested stack or web application as well.

User interaction interface: This is how users of the cloud interface with the cloud to request services.

Service catalogue: This is the list of services that user can request.

System management: This is the piece which manages the computer resources available and delivers on the requested service. It may also deploy the required images.

Monitoring and metering: This optional piece tracks the usage of the cloud so the resources used can be attributed.

Servers: The servers are managed by the system management tool. They can be either virtual or real.

Figure 4 represents the process flow in digitalization system.

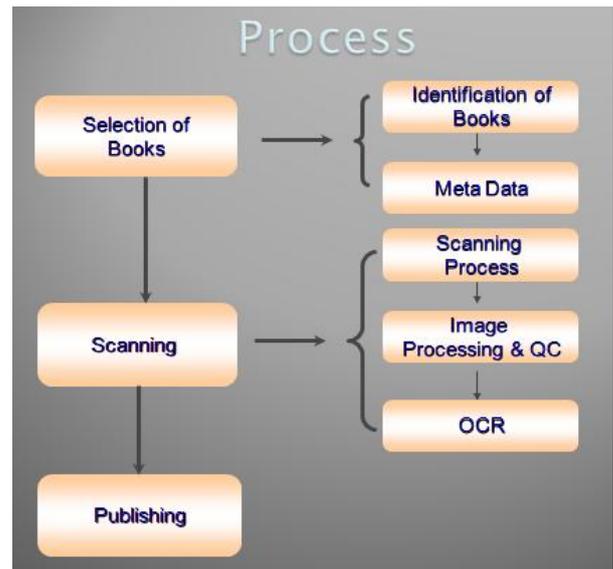


FIGURE 4: PROCESS FLOW OF DIGITAL LIBRARY

IV. CHOOSING THE RIGHT BACKUP METHOD:

The data stored in the cloud may be frequently updated by the user, including insertion, deletion, modification, appending, recording etc, to ensure storage correctness under dynamic data update is hence of paramount importance.

However solution to resolve data lost issue by back up with data storage on cloud or external server.

On-Site Backup Pros:

- **Cost:** Hard drives are inexpensive, so you can purchase plenty of backup space and ensure that you have room to expand without additional costs.
- **Installation:** Hard drives are easy to install, and usually come with software designed to make backing up your data easy and automatic.
- **Speed:** On-site backups are much faster than online storage solutions. This method is ideal for archiving large amounts of data.

On-Site Backup Cons:

- **Security:** A hard drive can easily be stolen. If there is financial or other sensitive data on it that is not encrypted, this can cause significant problems.
- **Damage:** Because it's in the same location as you or your office, should you or your office suffer a natural disaster all of your data may become irretrievable.
- **Virus protection:** If you have an undetected virus on your computer, it will probably get backed up to the

hard drive, thereby keeping it alive for as long as your backup is around.

Cloud-Based Backup Pros:

- **Backups of backups:** Your data is always stored off-site and is redundantly copied to other servers in different locations. If one goes down, your data can be retrieved from elsewhere on the backup network.
- **Security:** Data is encrypted by the backup service's software program on your computer before it is sent to the cloud, so thieves on the Internet can't intercept it.
- **Virus protection:** The backup service software detects any virus or infection before data is sent. If a virus is found, that file is flagged and not copied to the backup service. You will be notified that the corrupted file has not been deleted from your computer, so you won't lose any data, but it won't be backed up online.

Cloud-Based Backup Cons:

- **Cost:** Online backups are more expensive, often requiring monthly or annual fees based on the amount of data stored on the servers.
- **Capacity:** Cloud-based backups are best for sensitive information but not for large backups, such as movies, photos and music files. Since some Internet providers limit the amount of data you can send and receive in a month, you must be careful to avoid large backups that cause you to exceed their limits.
- **Speed:** It can take a long time to back up large files online, even with a broadband connection.

HOW TO RESTORE A CLOUD BACKUP: To update or restore cloud backup, customers need to use the service provider's specific client application or a Web browser interface. Files and data can be automatically saved to the cloud backup service on a regular, scheduled basis, or the information can be automatically backed up anytime changes are made (also known as a "cloud sync").

V. CHOOSING THE BEST BACKUP OPTION

Before you choose a data backup option, assess the advantages and risks of each media, your financial resources, and your needs, such as the amount of data to be backed up, protection for sensitive data (customer data, personally identifiable information, or personal health information), and accessibility of data (permanent archiving, temporary backups, and rolling backups).

Home users storing a relatively small amount of personal data should consider keeping primary files on the hard drive of

their computer, with at least two backup copies on solid-state storage, optical storage (stored in jewel cases), or remote storage.

Individuals or small businesses who want to store large amounts of non-sensitive data should consider keeping working files on their hard drives or servers, with at least two backup copies on separate servers, high-capacity optical media, high-capacity solid-state storage, digital tape systems, or cloud storage. If the stored data is sensitive, be sure to carefully consider the risks of cloud storage, encrypt your data, and keep any storage media physically secure.

Large businesses or organizations should consider keeping one backup copy onsite and another offsite either through a separate data service (such as a cloud service provider or remote server backup) or on the organization's own offsite servers or digital tape system.

3-2-1 rule of backups:

3 – Keep 3 copies of any important file: 1 primary and 2 backups.

2 – Keep the files on 2 different media types to protect against different types of hazards.

1 – Store 1 copy offsite (e.g., outside your home or business facility).

VI. PROPOSED METHOD FOR DATA BACKUP ON LIBRARY SYSTEM.



Conclusion & Future Challenges:

Implementation is expensive, can be overcome this issue by new technology or idea. As IT is the rapid growing technology and each and every organization is trying to get aid of IT to perform better for the growth of the business. Cloud is one of the solutions to enhance the business with cost effective strategy. More and more IT managers are turning to cloud services to accelerate their responsiveness to business needs. Cloud computing services can enable faster time to market and reduced startup costs through faster IT deployments and end-user-self-service. Here, in this paper we have explained now cloud based digital library boosts up the overall business and process of digital library. Furthermore, here we are highlighting some challenges for the future enhancement in cloud computing.

- Security Management

- Vendor Risks
- Layered Cloud Architecture
- SaaS & PaaS Issues.

REFERENCES:

- [1] Wikipedia. Cloud Computing <http://en.wikipedia.org/wiki/cloud-computing>
- [2] Rupesh Sanchati & Gaurav Kulkarni. Cloud computing in Digital and university Libraries: Global Journal of Computer Science and Technology, Global Journals Inc. (USA)
- [3] Cloud Computing and Digital Library <http://mashiat.hubpages.com/hub/Cloud-computing-and-Digital-library>
- [4] Introduction to cloud computing architecture: white paper [Abstract][online]. (June 2009). Available from: http://webobjects.cdw.com/webobjects/media/pdf/Sun_CloudComputing.pdf [Accessed 8 June 2011]
- [5] Google App Engine. <http://code.google.com/appengine>.
- [6] Cloud Computing Architecture in digital library available at: <http://en.wikipedia.org/wiki/cloud-computing>
- [7] [7] An Approach for Data Storage Security in Cloud Computing
- [8] [8] Cloud Computing Data Storage and Security Enhancement
- [9] International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 6, August 2012
- [10] [9] Cloud Computing For Business
- [11] Jaimin N. Undavia, Nikhil P. Shah / International Journal of Engineering Research and Applications