Impact of cloud computing and Cost effectiveness in R&D libraries

Sabita Kundu

Batanagar Institute of Engineering, Management and Science
Batanagar, Kolkata- 700 141
bitu130506@yahoo.com

Abstract: Cloud Computing (CC) is a new technique of computing that is extensively used in today's industry as well as society or by Academic Scholar. After promotion in IT computing from traditional library, operational expenditure is a major concern. Started systems modeling, clustering, and virtualization with Cloud Computing reduces computing cost significantly for both small users and large enterprises. Due to a subscription-based service Cloud Computing can obtain networked storage space and computer resources. Cloud Computing enables operators a great advantage since it can extends service to external customers using Internet Technologies. Cloud services can be Software as a Service (SaaS), Infrastructure as a Service (IaaS), and Platform as a Service (PaaS). The vision of using cloud computing in R&D libraries is to deliver library resources, services and expertise at the point of need, within user workflows. It can also create a powerful, unified presence for research libraries on the web and give user an indigenous facility, assembly group and comprehensive reach. This paper describes the current status of user models in R&D libraries through Cloud Computing.

Keywords: R&D libraries, Cloud Computing, Operational Expenditure, SaaS, IaaS, PaaS.

1. Introduction:

Use of modern technology in library has great relevance in the context of fourth law of Library Science. "Save the time of the Reader/user", in which the internal efficiency of the Library develops. IT facilitates in every steps in library that means collection, storage, organization, processing analysis, presentation, and detonation availability of information in machine readable form and multi use of machine readable records by users for any number of time and users can access information irrespective of time and location of the information source.

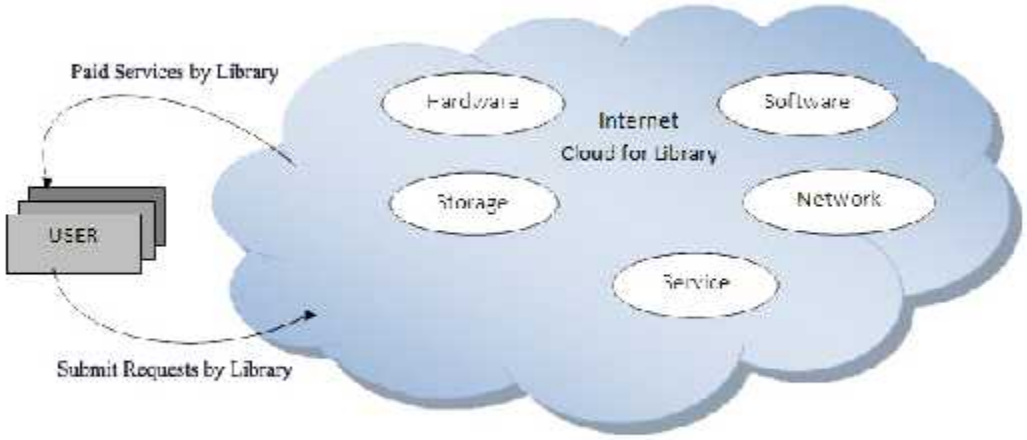
After promotion in IT computing from traditional library, have to face operational expenditure. It was started systems modeling, clustering, and virtualization with Cloud Computing. It has reduced computing cost significantly for both small users and large enterprises. Due to a subscription-based service it can obtain networked storage space and computer resources. This is a new technique of computing that is extensively used in today's industry as well as society or by academic scholar. It is a contemporary model of ICT and networking. The basic principle of cloud computing entails the reduction of in-house data centre's and the delegation of a portion or all of the information Technology infrastructure capability to others. The paper describes the current status of user models in R&D libraries. Then it proposed to improve current user service model with cloud computing. Cloud computing also explores some of the security issues surrounding data location, mobomility and availability and also explores the application of cloud computing in special libraries.

The notion of cloud computing obtain attention in 2006, when Amazon launched their new business. From the beginning of 2010 there was a recognizable upward trend in search queries for cloud computing.

There are two types of clouds available. The data centre- hardware and software is what we will call a Cloud. When a cloud is made accessible in a pay-as-you-go manner to the general public, we call it a public Cloud; the service being sold is Utility Computing. We use term Private Cloud to refer to internet data centre of a business or other organization, not made available to the general public. Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centers that provide those services

2. Cloud Computing technologies for libraries:

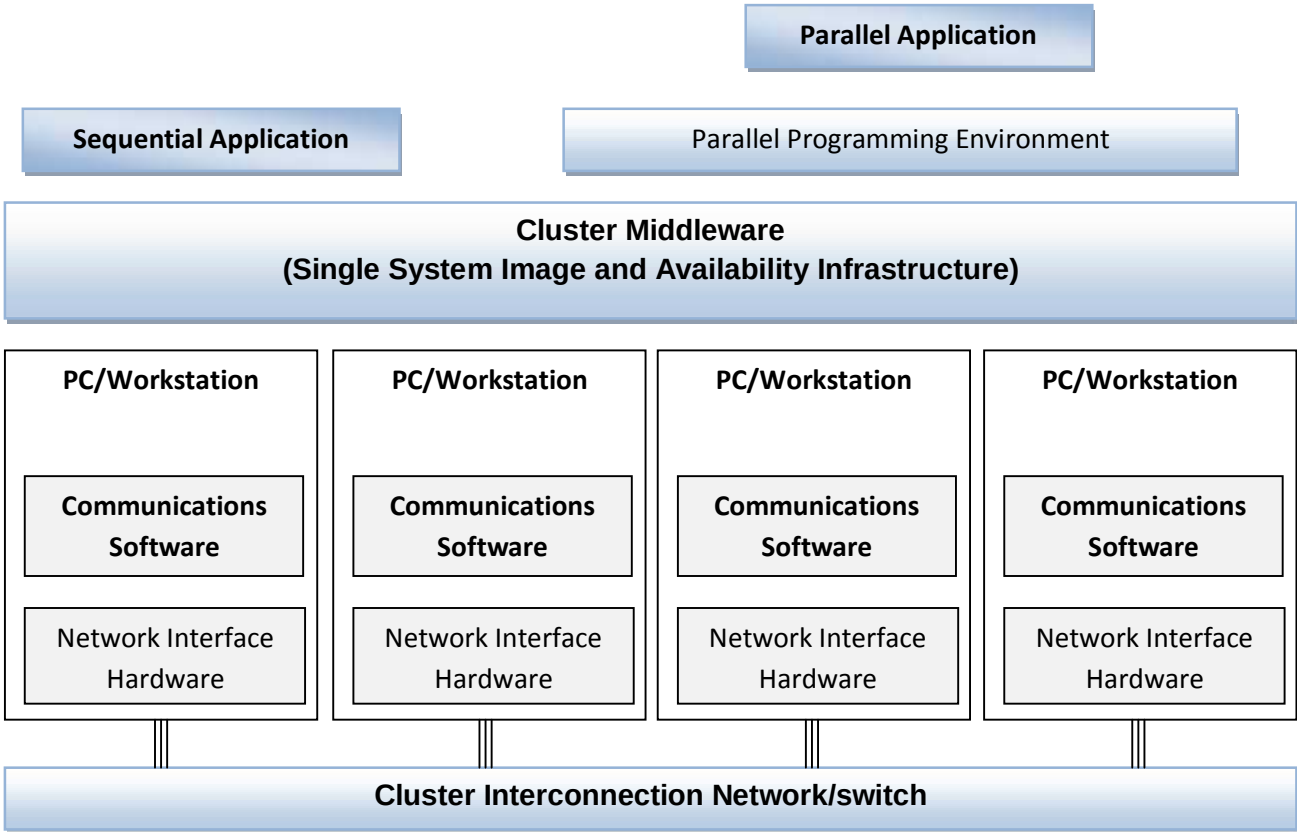
Institute of higher learning such as colleges universities are the core of innovation throughout their advanced research and development. Subsequently, higher institutions may benefit greatly by harnessing the power of cloud computing. Cloud computing has large potential for libraries. Libraries may put more and more content into the cloud. This paper explores an overview of cloud computing, types and development models of cloud computing. It also discusses the application of cloud computing in libraries and some of the cloud computing initiatives taken over the globe. The establishment of shared public cloud can save manpower and material resources greatly among academic libraries. (T, 2012)



However, by understanding the type of service offered by Cloud Computing, one begins to understand what this new approach is all about. There are following three main types of services that can be offered by the Cloud.

3. Cloud computing: an ICT innovation:

In ICT domain, things change fast as the researchers all over world are constantly working on making life easier and more exciting making it most cost effective for users. The technological development is a constant process, a whole new breed of technologies are coming up to reform the ICT industry and other user. The flexibility of IT operations, tools and management would immensely expand the services naturally adapted for cloud environments. Hence, the next generation data centers are being designed differently. They are integrated, flexible and much more exposed to new technologies as it becomes available. The cloud has plenty of benefits for users apart from creativities. The basic idea is to assess cloud service for the library and also to present an overview on cloud computing models and its application for library services.



4. A Basic Cluster Architecture: drawing architecture

The basic architecture of a computer cluster over PCs has been posturized. Workstations are a simple cluster of computers built with commodity components and fully supported with desired SSI features and HA capability. The processing nodes are commodity workstations, PCs, or servers. These commodity nodes are easy to replace or upgrade with new generations of hardware. The node operating systems should be designed for multiuser, multitasking, and multithreaded applications. The nodes are interconnected by one or more fast commodity networks. These networks use standard communication protocols and operate at a speed that

should be two orders of magnitude faster than that of the current TCP/IP speed over Ethernet. (H. 2012)

The network interface card is connected to the node's standard I/O bus (e.g., PCI). When the processor or the operating system is changed, only the driver software needs to change. We desire to have a platform-independent cluster operating system, sitting on top of the node platforms. But such a cluster OS is not commercially available. Instead, we can deploy some cluster middleware to glue together all node platforms at the user space. An availability middleware offers HA service. An SSI layer provides a single entry point, a single file hierarchy, a single point of control, and a single job management system. Single memory may be realized with the help of the compiler or a runtime library. A single process space is not necessarily supported.

An idealized cluster for supports:

In general, an idealized cluster is supported by three subsystems. First, conventional databases and OLTP monitors offer users a desktop environment in which to use the cluster. In addition to running sequential user programs, the cluster supports parallel programming based on standard languages and communication libraries using PVM, MPL, or Open MP. The programming based environment also includes tools for debugging, profiling, monitoring, and so forth. A user interface subsystem in needed to combine the advantage of the web interface and the Windows GUI. It should also provide user-friendly links to various programming environments, job management tools, hypertext, and search support so that users can easily get help in programming the computer cluster. (CUDA, 2014)

Cloud providers will create their own APIs (application programming interfaces) so that software developers can use them to create client applications in order to access that functionality. Currently, some of those API are proprietary; an issue which will be revisited later when examining some of the limitations of cloud computing.

5. Cloud computing in R&D Library:

For application of cloud technology, librarians need to know the use of cloud infrastructure, cloud applications and developing and deploying applications into a cloud environment. Application of CC for libraries will increase the efficiency and opportunity for collaboration, decrease in-house technical expertise, cost reduction, data security, privacy, and ownership, long term stability.

Libraries have been adopting cloud based solutions for different services including storage or data back-up services, electronic journal access management, library bibliographic database, institutional repositories and digital libraries, knowledge management, mobile applications, E-learning, statistics tracking, Integrated Library System (LIS) hosting and cloud-based sharing applications.

6. Cloud services are modular:

Cloud services are modular; library can pick and choose one or two cloud services for less perceptive computing needs and keeps the rest in-house until they are more contented with the prospect. With could-based computing, work done in one place and amenities are available to all

from anywhere. When library resources are stored in the cloud it offers several benefits. Universal data can be easily shared surrounded by library users. The need for local storage, maintenance and backups is eliminated. And the end libraries can achieve web scale when they extremely aggregate data and users, something a cloud environment makes possible. Another great benefit of resources stored in the cloud is the opportunity for collaboration and cooperative intelligence. Library can agree to share pools of data for cooperative collection building, cooperative cataloguing, cooperative preservation or digitization, cooperative sharing of materials etc.

6.1 Infrastructure as a Service (IaaS): Product offered via this mode includes the remote delivery (through the Internet) of a full computer infrastructure (e.g., virtual computers, service, storage device, etc.)

6.1.1 IaaS in R&D Library:

libraries have started jumping on board with new style of computing and have been adopting laas to host their websites and store their data, to alleviate IT departments, and increase efficiency. IaaS environments allow users to provision servers, storage space, and networking components to meet their computing needs in an IaaS environment, the organization is responsible for starting and sizing a server & managing its network access. **(P,2012.)**

6.2 Platform as a Service (PaaS): To understand Cloud Computing layer one needs to remember the traditional computing model where each application managed locally required hardware, an operating system, a database middleware, web servers, and other software. One also needs to remember the team of network, database, and system management experts that are needed to keep everything up and running. With cloud computing, these services are now provided remotely by cloud providers under this layer.

6.2.1. PaaS in R&D Library:

Libraries are beginning to develop using the services to easily create their own Webbased and mobile applications without having to create the infrastructure for a development environment within their libraries. PaaS solutions focus on providing a hosted platform on which a specific application can be deployed. This platform is often some provisioned space and computing resources from a hosting company running a pre-configured set of tools. Organizations can deploy a locally developed or managed application on the platform but do not manage the underlying server infrastructure. (P,2012.)

6.3. Software as a Service (SaaS): Under this layer, applications are delivered through the medium of the internet as a service. Instead of installing and maintaining software, we simply access it via Internet, freeing ourselves from complex software and hardware management. This type of cloud server offers a complete application functionality that ranges from productivity (e.g. office-type) applications to programs such as those for Customer Relationship management (CMR).

6.3.1 SaaS in R&D Library:

Libraries are making use of both user and business grade cloud applications to provide new services to patrons, market the library, save money as cost effective by replacing traditional

software products, collaborate with staff, much more. In a SaaS environment, organizations use an application via a hosted service. They do not have access to the underlying infrastructure, and are not responsible for managing the underlying software. A common R&R library example is electronic journal subscription systems. (P, 2012.)

7. Reduce Cost than traditional IT Computing:

In traditional IT computing, users must acquire their own computer and peripheral equipment as capital expenses. In addition, they have to face operational expenditures in operating and maintaining the computer systems, including personnel and service cost. Figure 1 shows the addition of variable operational cost on top of fixed capital investments in traditional IT. Where the fixed cost is the main cost, and that could be reduced slightly as the number of users increases. However, the operational cost may increase sharply with a larger number of users. Therefore, the total cost escalates quickly with massive number of users. On the other hand, cloud computing applies a pay-per-use business model, in which user jobs are outsourced to data centers. (C. 2010)

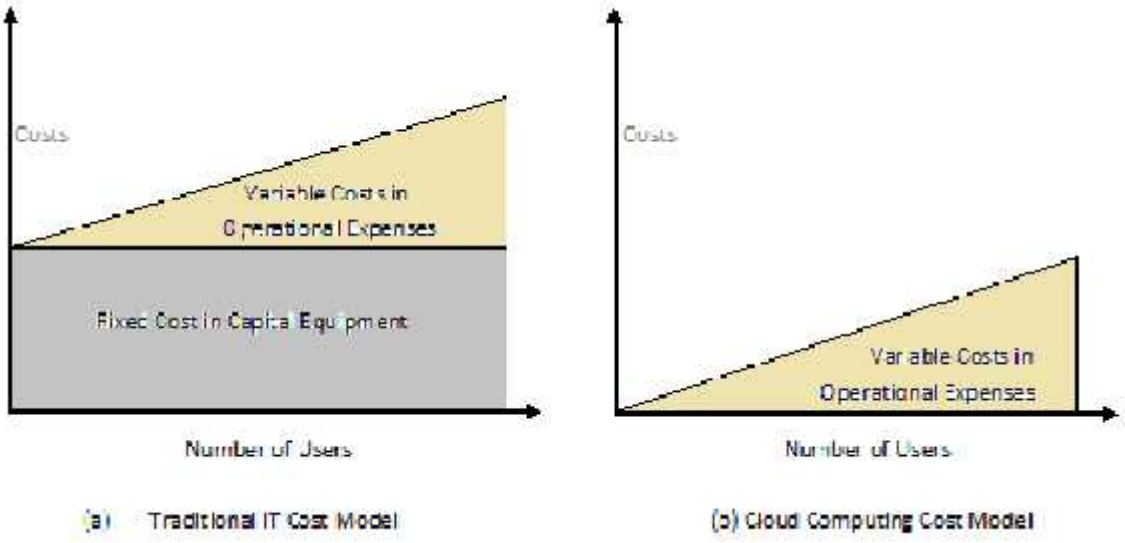


Figure: 1

Overall, cloud computing will reduce computing costs significantly for both small users and large enterprises. Computing economics does show a big gap between traditional IT users and cloud users. The savings in acquiring expensive computers up front releases a lot of burden for startup libraries. The fact that cloud users only pay for operational expenses and do not have to invest in permanent equipment is especially attractive to massive numbers of small users. This is a major driving force for cloud computing to become appealing to most enterprises and heavy computer users. In fact, any IT users whose capital expenses are under more pressure thn their operational expenses should consider sending their overflow work to utility computing or cloud service providers. (C. 2010).

8. Basic Cloud Security:

Three basic cloud security enforcements are expected. First, facility security in data in data centers demands on-site security year round. Biometric readers, CCTV (close-circuit TV), motion detection, and man traps are often deployed. Also, network security demands fault-tolerant external firewalls, intrusion detection systems (IDSes), and third-party vulnerability assessment. Finally, platform security demands SSL and data decryption, strict password policies, and system trust certification. In the mapping of cloud models, a special security measures are deployed at various cloud operating levels.

Servers in the cloud can be physical machines or VMs. User interface are applied to request services. The provisioning tool carves out the systems from the cloud to satisfy the requested service. A security-aware cloud architecture demands security enforcement. Malware-based attacks such as network worms, viruses, and DDoS attacks exploit system vulnerabilities. These attacks compromise system functionality or provide intruders unauthorized access to critical information. (H, 2010)
Thus, security defense are needed to protect all cluster servers and data centers. Here are some cloud components that demands special security protection:

a) Protection of servers from malicious software attacks such as worms, viruses, and malware.
b) Protection of hypervisors or VM monitors from software-based attacks and vulnerabilities.
c) Protection of VMs and monitors from service disruption and DoS attacks.
d) Protection of data and information from theft, corruption, and natural disasters.
e) Providing authenticated and authorized access to critical data and service.

9. The new workflows for LIS create by cloud computing:

Cloud Computing offers the opportunity for a cooperative platform for libraries to build on. They are four key principles of a cooperative platform:-

- Directness, means the services and data are made available to support greater inters operability, not only cloud services, but also with within and between library developed and third party applications;
- Extensibility means the platform can easily accommodate the addition of new services and applications, developed either by the service provider or by member of the community;
- Data prosperity, means library can interact with and expose a wide variety of information about purchased, licensed, and digital content through this platform; and
- Teamwork, means libraries can harness the collective power of the community of libraries to innovate and share solutions

10. Advantage of Cloud Computing in R&D Library:

Cloud Computing is very helpful in resource sharing in between libraries.

10.1 Space Sharing:

A Common Scheme is to assign higher priorities to short, interactive job in daytime and during evening hours using tiling. In this space-sharing mode, multiple jobs can run on disjointed partitions (groups) of nodes simultaneously. At most, one process is assigned to a node at a time. Although a partition of nodes is dedicated to job, and the I/O subsystem may be shared by all jobs. Space Sharing must solve the tilling problem and the large-job problem. (H, 2012)

10.2 Time Sharing:

In the dedicated or space-sharing model, only one user process is allocated to a node. However, the system processes or daemons are still running on the same node. In the time-sharing mode, multiple user processes are assigned to the node. Time sharing introduces the following parallel scheduling policies. (H, 2012)

10.3. Load Sharing Facility (LSF) for Cluster Computing:

LSF is a commercial workload management system from Platform Computing. LSF emphasizes job management and load sharing on both parallel and sequential jobs. In addition, it supports cheek pointing, availability, load migration, SSI. LSF is highly scalable and can support a cluster of thousands of nodes. LSF has been implemented for various UNIX and Windows/NT platforms. Currently, LSF is being used not only in clusters but also in grids and clouds. (H, 2012)

11. Challenge 2 – Data Privacy and Security Concerns:

Current cloud offerings are essentially public (rather than private) networks, exposing the system to more attacks. Many obstacles can be overcome immediately with well-understood technologies such as encrypted storage, virtual LANs, and network middle boxes (e.g., firewalls, packet filters). For example you could encrypt your data before placing it in a cloud. Many nations have laws requiring SaaS providers to keep customer data and copyrighted material within national boundaries.

Traditional network attacks include buffer overflows, DoS attacks, spyware, malware, rootkits, Trojan horses, and worms. In a cloud environment, newer attacks may result from hypervisor malware, guest hopping and hijacking, or VM rootkits. Another type of attack is the man-in-the-middle attack for VM migrations. In general, passive attacks steal sensitive data or passwords. Active attacks may manipulate kernel data structures which will cause major damage to cloud servers. We will study all of these security and privacy problems on clouds. (T, 2009)

12. Conclusion:

India already has over 200 million internet users and every month add 4 million new ones. By the end of 2015, India will have more internet user than the United States. It is promising market for Indian R&D Libraries. The internet is making it cheaper than ever to efficiently market to new customers, both within India and abroad. And increase the user of libraries means R&D

Libraries. The cloud / grid systems provide latest developments; explain how to create high-performance cluster, scalable networks, automated data centers and high-throughput.

References:

1. Distributed and Cloud Computing: from parallel processing to the internet of thing by Kai Hwang; Geoffrey C. Fox; Jack J. Dongarra of London: Elsvier, 2012.
2. Understanding the Cloud Computing by Sunil Tyagi in the Proceeding "Innovative Challenges in Information Services" National Seminar conducted by Dept. of Library of University of Jammu held on March 2012.
3. Using cloud computing in Libraries by O.S.S. in the Proceeding "Innovative Challenges in Information Services" National Seminar conducted by Dept. of Library of University of Jammu held on March 2012.
4. T. Chou, Introduction to Cloud Computing: Business and Technology, Active Book Press, 2010.
5. T. Mather, et al., Cloud Security and Privacy: An Enterprise Perspective on Risk and Compliance, O'Reilly Media, Inc., 2009.
6. K.Hwang, & D. Li, Trusted cloud computing with secure resources and data coloring, IEEE Internet Comput. (2010)
7. IBM, http://www.redbooks.ibm.com/redpapers/pdfs/red.
8. wikipedia.http://en.Wikipedia.org/wiki/CUDA access 15/07/14