Cloud Computing Research Around the Globe in Different Universities: An Overview
Mayank Yuvaraj and Ajay Pratap Singh


ABSTRACT
The first decade of the twenty-first century has not witnessed a world without ubiquitous and virtualized cloud computing power. The study analyses the pattern of research initiatives carried around the globe in different universities as well as briefly elucidates the notion of Cloud Computing (CC). The study will help the researchers to define their research problems and understand the concept at its optimum.


INTRODUCTION
Tremendous amounts of research effort have been devoted to studying different issues and problems of CC. Although, researchers have recognized CC as a legitimate area of research yet CC research is scattered across a wide range of disciplines and published sources. The present paper aims to provide an overview of CC research around the globe in different universities. Overall, researchers have addressed different issues and problems of CC as well as proposed many solutions, while some are still being investigated. Moreover, it is impossible to provide information about all the research activities that have taken place so far. So, avoiding any disingenuous labor we have put forth the concept as well as the relevant CC research initiatives carried out in universities.

CLOUD COMPUTING: DEATH OF HARDWARE
In the recent years Information Technology sector has witnessed technological turmoil in the form of thunders of CC. The concept CC has been defined as an integrated package of computing services and applications on web, offered as a utility ensuring throughout abstraction shown in figure 1. It provides an alternative to traditional computing where the operating systems, hardware and software are rented over the Internet. Unlimited ICT capacity and applications are offered as an online service through CC. Unlike traditional computing environment in a CC niche the cloud service provider looks after the resources such as: hardware, maintenance, updating, security, and other essential resources. The customer is independent of IT bothering and he has to only find the right services which can enable the customer to improve the computing performance. Although the concept was in use long time back but the term was formally announced by Google in 2010 which most of the scholars feel as a marketing strategy.

The environment of academic libraries has undergone a drastic change in the last few years. With the emergence of Internet era libraries are facing a big question to justify their services.
Moreover, academic libraries have discovered a sharp fall in the number of gate counts as well as reduction in the circulation of traditional library materials. Some researches indicate that a rapidly growing percentage of the use of electronic resources occurs outside of the library. Where once students and faculty turned to libraries, they now turn to their personal computers, tablets or smart phones when they need to find information. Under such circumstances, CC is an approach to computing that could be worthwhile to academic libraries. Academic libraries can choose to how to get them connected to become a part of the cloud by opting from variations of software, systems, and hardware services. Whatever they choose CC could help the academic libraries save time, money and resources if servers and software were not needed on their premises as they are today. Hopefully libraries could focus more directly on services and materials for patrons if their computer hardware and software were handled by IT companies of the cloud. There are many reasons why CC is being common16. Technologically, we use CC because we can and it is convenient to use. Economically, it is cost effective and pocket friendly, and finally it makes interactivity easier to achieve with targeted audiences.

**CHARACTERISTICS OF CLOUD COMPUTING**
- **Self-seeking on-demand service in cloud**
- **Wider accessibility to networks**
- **Metered service depending on the usage of cloud based services**
- **Flexibility or resiliency in cloud**
- **Pooling of resources available in cloud**
- **Virtualization of computers**
- **Green Computing**
- **Means to deliver IT to the end-users.**

**Flairs of Cloud Computing**

Cloud Computing offers various services as shown in figure 2 which has been illustrated below:

**Software as a Service (SaaS)**

In SaaS model, the applications or software is delivered as a service to the end user, who can access the program online using a web browser or any other suitable users. SaaS services may be available on rental basis or on peruse basis. For example email and few other services offered by google, hotmail, yahoo, sify, skype etc. CC offers the ability of libraries to use online software to handle a task like video chat through gmail video chat or through skype8. Both of these are free services though there is little customization or control available with these applications11. Since the services and application interfaces are often familiar with users, it will decrease the learning curve for library staff and users8.

**Platform as a Service (PaaS)**

In PaaS model, a computing platform supplies tools and a development environment to help organizations to build, test, and deploy Web-based applications. It helps organizations not to make investment in the infrastructure required for building Web and mobile applications, but can rent the use of platforms. PaaS is based subscription model so users only pay for what they use. Special platform or application infrastructure is also being provided to the clients. Moreover, client does not require knowing programming language, database management systems, etc. to run applications. Windows azure, google app engine and force.com are the finest examples of PaaS. Libraries can create applications in online environment where they can build, test, and deploy web based applications11. PaaS gives the library the freedom to explore development options without having to purchase and maintain the required infrastructure8.

**Infrastructure as a Service (IaaS)**

In IaaS model, computer infrastructure is offered as a platform in a virtualization environment which can be used as a service along with raw storage and networking. IaaS is priced on a pay-as-you-go model enabling clients to scale up or down the operations depending on their needs at any given time and pay only for what they use. In the simple way, hardware services such as processors, memory, networks etc on agreed basis for specific duration and price. In IaaS, clients are being offered with storage, networking and processing of data. Amazon’s Elastic Compute Cloud (EC2) or Simple Storage Service (S3), VMWare vCloud are some of its examples. By using IaaS, a library can purchase server space and computing power8. A library does not need to purchase a server which is underutilized but costs the same to purchase and maintain as if it were using all of its resources at all times. By using an IaaS, a library gains the benefit of only paying for the resources you actually use14.

Any organizations can deploy CC applications on public, private, hybrid and community clouds model.

Craggs8 has put forth the main categories of CC deployment:

- **The Public Cloud**: IT resources and services are owned by a third party, located off-premise and made available to anyone on a commercial basis as metered services.
- **The Community Cloud**: IT resources and services are owned and operated on behalf of a community of organizations.
- **The Private Cloud**: IT resources and services are owned/leased by a single company for its own use.
The Internal Cloud: A private cloud where all resources remain on-premise.

The Hybrid Cloud: A combination of two or more cloud models.

Research Initiatives in Cloud Computing

Research initiatives are an inventory of research processes undertaken by various institutions to solve a research problem. Some of the notable research initiatives to the cause of CC is enlisted below:

- Boston University (BU) Cloud Initiative, U.S.A.
  Funded by National Science Foundation grants BU has taken initiatives in distributing/CC that aims to offer an open marketplace to independent, rational parties interested in setting up their own applications. They are exploring the competencies of Colocation Games to develop an economically-sound framework upon which emerging cloud architecture could be implemented.

- Carnegie Mellon University (CMU) Cloud Initiative, U.S.A.
  CMU is host to several CC research programs. CMU has joined hands with silicon valley based researchers to address the need for industry-wide, globally accepted measures for calculating the benefits and risks of CC services. Development of standards and measures to determine the cost, risks, quality and performance for CC services is also on top priority. Further, CMU in Qatar along with IBM (International Business Machines Corporation) is working on Cloud Computing Lab research initiative from 2009 to bring CC to the middle east. It focuses on performance analysis of scientific workloads, with regional relevance, on the cloud as well as solutions to overcome current limitations of the existing cloud paradigm.

Also, Intel Science and Technology Center for Cloud Computing (ISTC-CC) has been formed as an open community to devise critical new underlying technologies for the cloud computing of the future. It is headquartered at CMU that includes researchers from GeorgiaTech, Intel, Princeton and UC-Berkeley. They focus on future clouds and its applications that is carved on four inter related pillars shown in figure 3.

- Duke University (DU) Cloud Initiative
  Funded by National Science Foundation in collaboration with NCSU, UNC Chapel Hill, and NCAT State University to explore and test Trustworthy Virtual CC Duke University has created a trustworthy virtual cloud super-computing environment for researchers. DU's Scalable Computing Support Center has developed two options for CC:

  - Private Duke Cloud, where researchers can purchase high-priority cycles as needed from a pool of machines and pay per CPU hour.
  - Economy Cloud, where researchers can purchase low-priority cycles only, for a lower per CPU rate.

Researchers pool existing resources and share access to greater computing power than they could afford on their own, without worrying about the 'care and feeding’ of the equipment, since DU provides the physical space, systems administration, programming assistance, power and cooling for the machines.

- Florida International University (FIU) Cloud Initiative
  FIU researchers are leveraging CC to analyze aerial images and objects to help support disaster mitigation and environmental protection. Prominent researches carried in the CC domain include creation of a virtualized infrastructure systems and applications laboratory and Verification based Integrity Assurance Framework (VIAF).

- Indian Institute of Technology (IIT) Cloud Initiative, Delhi
  In India, IIT has played a pioneering role and is the earliest adopter of CC solutions. IIT Delhi has developed Baadal in 2011, which is an indigenously developed cloud orchestration and virtualization management software that can work with multiple virtualization technologies like KVM, Xen and VmWare. Further, IIT Delhi has commissioned ownCloud, a file and document sharing utility similar to the popular dropbox, for use by the IIT Delhi community. The utility supports storing and sharing of files, images, music and documents, contacts, calendar, tasks etc. It also supports version control and syncing with Windows/Linux/Mac desktops and Android and iOS based devices.

Researchers at CMU are using CC for multi-tier indexing of Web search engines. They are using CC to characterize the topicality of web content to more effectively process web searches. Routing searches topically requires less effort than traditional searches, enabling significant computational and financial savings. The project is using the Google/IBM cluster to ‘crawl’ the web and perform the data cleansing and pre-processing necessary to develop a web dataset of 1 billion documents to support the research. The web dataset is also being made available to the larger information retrieval community to multiply the impact of the project on that discipline.

Fig. 3: Focus Areas of Intel Science and Technology Center for CC
- 50 TB of virtualized storage based on a NetApp 3210V NAS and HP EVA6400 SAN with FC disks.
- Open source virtualization technology based on KVM.

**Indiana University (IU) Cloud Initiative**
The researchers at IU are working on several CC projects with grants from the National Science Foundation (NSF) and the National Institute of Health (NIH). Their research includes Large-Scale Distributed Scientific Experiments on Shared Substrate\(^5\), Exploring the use of cloud techniques to overcome current medical computing obstacles such as long computation time and large memory requirements; and The FutureGrid project that will provide an experimental platform that accommodates batch, grid and CC.

**Massachusetts Institute of Technology (MIT) Cloud Initiative**
MIT\(^5\) in collaboration with Yale University (YU) and the University of Wisconsin at Madison are working on comparative study of approaches to cluster-based, large-scale data analysis funded by NSF. In addition they are also independently studying CC Infrastructure and Technology for Education.

**North Carolina Agricultural & Technical State University (NCATSU) Cloud Initiative**
The team at the NCATSU is conducting research, funded by a NSF, in collaboration with NCSU, DU, and the University of NC at Chapel Hill to explore and test Trustworthy virtual CC.

**North Carolina State University (NCSU) Cloud Initiative**
NCSU and IBM have developed VCL (Virtual Computing Lab) shown in figure 4 to provide every student with most advanced educational resources. Through cloud infrastructure students will be provided accessibility to most advanced educational materials, select software applications and computing and storage resources.

The VCL solution allows users to remotely access a desired set of applications and environments over the Internet using a personal computer, laptop or mobile device from anywhere, at any time. To VCL users, even the most demanding software applications, operating systems and environments are easily accessible through license-honoring technology in a click of the mouse. Access is instant, and offers a range of options from single desktops to classroom-sized labs, to collections of servers and storage, to high-performance computing clusters.

Further, NCSU\(^13\) is engaged in several CC projects funded by the NSF which includes two collaborative studies on Trustworthy Virtual CC and Hybrid Opportunistic Computing for Green Clouds.

**Purdue University (PU) Cloud Initiative**
The project at Purdue University is investigating linguistic extensions to MapReduce abstractions for programming modern, large-scale systems, with special focus on applications that manipulate large, unstructured graphs. This will impact a broad class of scientific applications. Graphs have important utility in the social sciences (social networks), recommender systems, and business and finance (networks of transactions), among others. The specific case study targeted by the research is a comparative analysis of graph-structured biochemical networks and pathways which underlie many important problems in biology. They are also providing a CC testbed called Wispy to TeraGrid users.

**Tsinghua University (TU) Cloud Initiative, China**
TU\(^3\) is the first prestigious university in mainland China to join Google’s CC related programs and to work with Google on the offering of data processing courses and the research on CC. They have developed TU Cloud which is an all in one CC solution developed at the Grid Computing Division of TU in China that comprises of three components: Nova (virtual computation system: computing cloud), Carrier (distributed file system) and Corsair (distributed file manager based on carrier: storage cloud), which can be utilized independently or in combination.

---

![Fig. 4: Virtual CC lab of NCSU](image-url)
University of California (UC) Cloud Initiative, Irvine
Funded by NSF, UC Irvine is working on a project to provide support for efficient fuzzy queries on large text repositories in the cloud. Supporting fuzzy queries can ultimately help applications mitigate their data quality issues because entities with different representations can be matched for example ‘PO Box’ vs ‘P.O. Box’.

University of California (UC) Cloud Initiative, San Diego
Researchers at the UC, San Diego are studying how to manage and process massive spatial data sets on large-scale compute clusters. This research will use the LiDAR topography data hosted by OpenTopography as a test case and will focus on how CC can aid the management and processing of massive spatial data sets.

University of California (UC) Cloud Initiative, Sant Barbara
The UC, Sant Barbara is actively pursuing several advancements in CC. Their Massive Graphs in Clusters (MAGIC) project is focused on developing software infrastructure that can efficiently answer queries on extremely large graph datasets. They have also designed an open-source implementation of the Google AppEngine interface.

University of Maryland (UM) Cloud Initiative, College Park
The research team at the UM, Cloud Computing Center at College Park is working on a range of projects funded by the NSF. They include a Hadoop Toolkit for Distributed Text Retrieval, Data Intensive Text Processing, Commodity Computing in Genomic Research, and a series of other independent studies.

University of Massachusetts (UM) Cloud Initiative
CC research in University of Massachusetts involve the Google/IBM cloud to learn more about world relationships which is funded by NSF.

University of Melbourne (UM) Cloud Initiative
The Cloud Computing and distributed systems laboratory at UM is actively engaged in the design and development of next-generation computing systems and applications that aggregate or lease services of distributed resources depending on their availability, capability, performance, cost, and users’ quality-of-science requirements. The lab is working towards realizing this vision through its two flagship projects: Gridbus and Cloudbus.

University of Minnesota (UM) Cloud Initiative
Researchers at UM are developing a cloud proxy network that allows optimized and reliable data-centric operations to be performed at strategic network locations.

University of North Carolina (UNC) Cloud Initiative, Chapel Hill
In collaboration with NCSU, DU, and NCAT State University, researchers at the UNC are trying to explore and test trustworthy virtual CC environment.

University of Utah (UU) Cloud Initiative
The researchers at the UU are working jointly with University of Washington(UW) on building a new infrastructure for computational oceanography that uses the Google/IBM cloud to allow ad hoc, longitudinal query and visualization of massive ocean simulation results at interactive speeds.

University of Virginia (UV) Cloud Initiative
The team at the UV is working on several CC projects funded by NSF. They include, feedback controlled management of virtualized resources for predictable esicience and image super-resolution using trillions of examples.

University of Washington (UW) Cloud Initiative
The researchers at the UW are working jointly with University of Utah on building a new infrastructure for computational oceanography that uses the Google/IBM cloud to allow ad hoc, longitudinal query and visualization of massive ocean simulation results at interactive speeds. Further, their Astronomy Survey Group is conducting research to scale the sky with MapReduce/Hadoop.

University of Wisconsin (UW) Cloud Initiative, Madison
The UW, Madison have designed the Hierarchically Redundant, Decoupled storage project (HaRD) to investigate the next generation of storage software for hybrid Flash/disk storage clusters. They are also working with MIT and Yale University (YU) on a comparative study, funded by a NSF CluE grant (Cluster Exploratory Grant), of approaches to cluster-based, large-scale data analysis.

Wayne State University (WSU) Cloud Initiative
Wayne State University is working on developing a unified learning approach, namely URL, to automate the configuration processes of virtualized machines and applications running on the virtual machines and adapt the systems configuration to the dynamics of cloud. Further, they have started Wireless health initiatives that attempts at applying sensors, wireless communications and CC to healthcare applications.

Yale University (YU) Cloud Initiative
The team at YU is working in collaboration with MIT and the UW at Madison on a comparative study, funded by a NSF CluE grant, of approaches to cluster based, large-scale data analysis.

CONCLUSION
Cloud Computing is being considered as the future of computing technology. From the study it is clear that there is an acute dearth of research initiatives in developing countries like India. Although it has many merits and is an end to painless works of updating, configuring and installing software educational organizations should look forward for CC solutions. The implementation of CC will not be succeeded if universities merely view as an online applications. Although, some of the universities have got discovery services, e-mail service on cloud and are further moving ahead step-by-step sharing resources among libraries however moving in patches to cloud is risky and not going to be fruitful in the long run. There is a need of proper training and plan well before introducing CC in the educational setup supplemented by well planned research initiatives.

REFERENCES


17. WHI: Wireless Health Initiative at Wayne State University. (Accessed on 02.08.2012), http://wirelesshealth.wayne.edu/