# Role of Spatial Database in Virtual Networking - Cloud Computing

*DR. NEERAJ BHARGAVA*

**Associate Professor, Department of Computer Science**
**School of Engineering & System Sciences, MDS University, Ajmer**

*drneerajbhargava@yahoo.co.in*

***ABSTRACT*:** *During the past few years, Cloud computing has become a key IT buzzword. This research paper gives an introduction of Cloud computing & reviews its definitions. This paper explores some of the basics of Cloud computing such as its components and architecture. The subsection of the paper compares some major properties of representative companies of cloud computing. GIS is a field of science, which could use Cloud Computing for distributed parallel processing of large set of data, and store and share the end results with users from around the world. In this paper, we propose architecture for GIS data processing using Cloud Computing technologies.*
*In conclusion readers will discover how any enterprise/business GIS can use Cloud computing to foster innovation and reduce IT costs.*

***Keywords— - Cloud computing, utility computing, autonomic computing, GIS***

## I: INTRODUCTION

Cloud computing as a technology trend is capturing the imagination of many. There are several definitions for this term to be inclusive of "everything" as a service-software to flexible computing infrastructure.

Ease of solution deployment is attractive to IT solution developers. Dramatic cost advantages are truly compelling for the users of flexible computing infrastructure. Cloud computing also represents a paradigm shift from owned IT infrastructure to using software applications, computing and storage as services over a wide-area distribution network such as the internet. This trend is being enabled by quantum performance improvement in the IT infrastructure components combined with enabling technologies such as virtualization, distributed computing, multi-core processors, service-oriented architectures and key innovations in the way cloud computing infrastructure is built.

## II: DEFINITION

Cloud computing is a pool of scalable IT-enabled capabilities which can be utilized over the internet (cloud) as a service.

The idea of cloud computing is based on a very fundamental principal of `reusability of IT capabilities`. The difference that cloud computing brings compared to traditional concepts of "grid computing", "distributed computing", "utility computing", or "autonomic computing" is to broaden horizons across organizational boundaries.

"Cloud computing is a computing paradigm in which tasks are assigned to a combination of connections, software and services accessed over the Internet. This network of servers and connections is collectively known as "the cloud." Computing at the scale of the cloud allows users to access supercomputer-level power. Using a thin client or other access point, like an iPhone, BlackBerry or laptop, users can reach into the cloud for resources as they need them."[1]



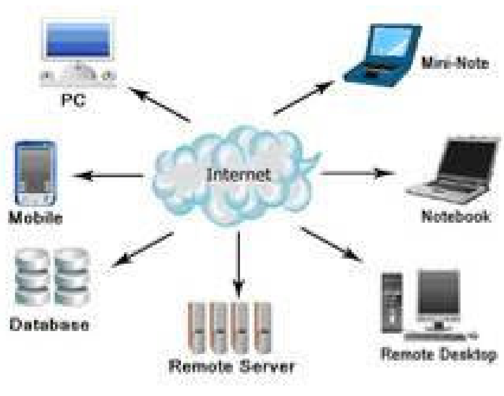


Figure 1:-Cloud Computing

Wikipedia says; "Cloud computing refers to computing resources being accessed which are typically owned and operated by a third-party provider on a consolidated basis in data center locations. Consumers of cloud computing services purchase computing capacity on-demand and are not generally concerned with the underlying technologies used to achieve the increase in server capability. There are however increasing options for developers that allow for platform services in the cloud where developers do care about the underlying technology."[2]

Though many cloud computing architectures and deployments are powered by grids, based on autonomic characteristics and consumed on the basis of utilities billing, the concept of a cloud is fairly distinct and complementary to the concepts of grid, SaaS, Utility Computing etc. In theory, cloud computing promises availability of all required hardware, software, platform, applications, infrastructure and storage with an ownership of just an internet connection.

It is called cloud computing because the data and application exist on a "Cloud" of Web server.

GIS applications are generally both compute and data intensive in nature. GIS technology has been around for decades. These mature technologies are increasingly using geospatial and non-spatial data. Advance data collection technologies have facilitated large amounts of rich data to be collected at diverse data sources. In time, size of these data would grow to be large enough to restrict any single organization to maintain and handle these data. In addition, GIS functions and services that operate on these data are geographically and logically distributed due to the source of data, location of computing facilities and organizations. The spatial analysis on large amount of data is complex and computationally intensive.

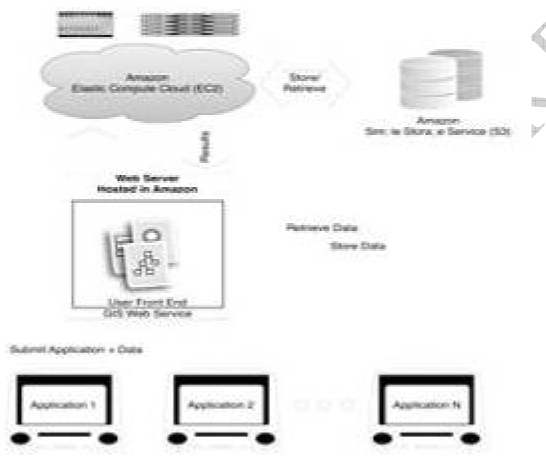


Fig 2: GIS Applications using Cloud computing infrastructure

### III: Clouds in Existence Today

1. Amazon External Internet Cloud:
- Simple Storage Service(S3)- $0.15c/month per a gig
- Elastic Computing Cloud(EC2)- Pay per use via on demand
- VMs- $0.10 VM instance/hour

2. Google's Proprietary Internal Cloud
- Estimated to harness 100,000's of servers.

Google is also said to be preparing to offer an external storage cloud.

3. IBM "Blue Cloud" offering for Enterprise

Data Center Cloud Creation:

Combines data-intensive Grid Virtualization (via IBM offering), and Elastic Computing (via Tivoli).

## Other examples of Cloud Computing:

| **System Property** | **Amazon Elastic Compute Cloud (EC2)** | **Google App Engine** | **Microsoft Live Mesh** | **(Sun Grid)** |
|---|---|---|---|---|
| Focus | Infra-structure | Platform | Infrastructure | Infrastructure |
| Service Type | Compute, Storage (Amazon S3) | Application Container | OS level | Compute |
| Virtualization | OS Level running on a Xen hypervisor | Application Container | OS level | Job management system (Sun Grid Engine) |
| Dynamic Negotiation of QoS Para-meters | None | None | None | None |
| User Access Interface | Amazon EC2 Command-line Tools | Web-based Administration Console | Web-based Live Desktop and any devices with Live Mesh installed | Job submission scripts, Sun Grid Web Portal |
| Web APIs | Yes | Yes | Unknown | Yes |
| Value-added Service Providers | Yes | No | No | Yes |
| Programming Frame-work | Customiz-able Linux-based Amazon Machine Image (AMI | Python | Not applicable | Solaris OS, Java, C, C++, FORTRAN |

Table 1: Comparisons of some representative of Cloud Platform

### IV: COMPONENTS OF A CLOUD

There are different clouds present in the computing space today which could be classified into the following components:

**Infrastructure:**

Cloud Infrastructure is the concept of providing `hardware as a service` i.e. shared/reusable hardware for a specific time of service. Example includes virtualization, grid computing, and para virtualization. This service helps reduce maintenance and usability costs, considering the need for infrastructure management & upgrade

**Storage:**

Cloud Storage is the concept of separating data from processing and storing in a remote place. Cloud Storage also includes database services. Examples are Google's BigTable, Amazon's Simple DB etc.

**Platform:**

A Cloud Platform is a service for application deployment and managing the required hardware & software needs. This could be a single service platform or a solution stack. Examples include Web application frameworks, Web hosting etc.

| SETI | Forums | Bit Torrent | Blog-space | Skype |
|---|---|---|---|---|
| Web site Hosting | Groove | Salesforce | Linkdln | VPS |
| Secondlife | FaceBook | YouTube | Yahoo | Qualys |

Table 2: Cloud computing Platform

**Application:**
A Cloud Application is an offering of software architecture that eliminates the need to install, run and maintain an application at the user's desktop/device. A cloud application eliminates the cost/resources required to maintain and/or support applications.

**Services:**
A Cloud Service is an independent piece of software which can be used in conjunction with other services to achieve an interoperable machine-to-machine interaction over the network. Examples include Amazon's Simple Queue Service, Google maps, Amazon's flexible payment service etc.

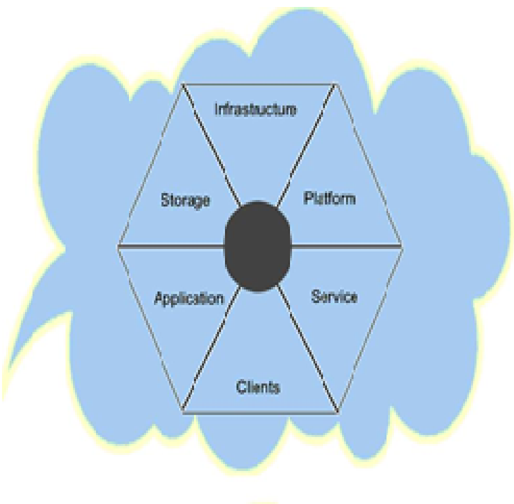


Fig 3: Components of a Cloud.

**Client:**
Cloud Client is a requester software or hardware device which tries to utilize cloud computing services over the network. The client device could be a Web browser, PC, laptop or mobile etc.

## V: CLOUD COMPUTING ARCHITECTURE

Cloud computing system, can be divided into two sections: the front end and the back end. They connect to each other through a network, usually the Internet[3] . The front end is the side the computer user, or client, sees. The back end is the "cloud" section of the system.

The front end includes the client's computer (or computer network) and the application required to access the cloud computing system. Not all cloud computing systems have the same user interface. Services like Web-based e-mail programs leverage existing Web browsers like Internet Explorer or Firefox. Other systems have unique applications that provide network access to clients.

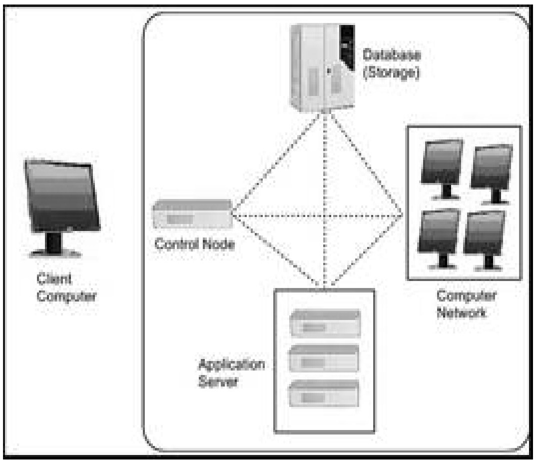


Fig 4: Cloud Computing Architecture

On the back end of the system are the various computers, servers and data storage systems that create the "cloud" of computing services. In theory, a cloud computing system could include practically any computer program you can imagine, from data processing to video games. Usually, each application will have its own dedicated server.

A central server administers the system, monitoring traffic and client demands to ensure everything runs smoothly. It follows a set of rules called protocols and uses a special kind of software called middleware. Middleware allows networked computers to communicate with each other.

If a cloud computing company has a lot of clients, there's likely to be a high demand for a lot of storage space. Some companies require hundreds of digital storage devices. Cloud computing systems need at least twice the number of storage devices it requires to keep all its clients' information stored. That's because these devices, like all computers, occasionally break down. A cloud computing system must make a copy of all its clients' information and store it on other devices. The copies enable the central server to access backup machines to retrieve data that otherwise would be unreachable. Making copies of data as a backup is called redundancy.

Cloud computing has increases profitability by improving resource utilization. Costs are driven down by delivering appropriate resources only for the time those resources are needed [14] .

GIS applications running on user's end submit application and data to the front-end web-service. This web service in turn using a suitable middleware to distributed the application tasks to virtualized nodes. This distribution of tasks enabled parallel processing of user tasks. Similarly, for the large amount of data, the web-service can directly use Amazon S3 API to store the data in the S3. When the analysis is being performed on Amazon EC2, data stored in S3 can be retrieved at real-time and processed. The results obtained after computation and analysis can also be stored in S3 for future reference. The users can then be notified asynchronously about the availability of the results.

Cloud based GIS application execution and analysis has several advantages over traditional approaches, as listed in Table 3

| **Parameters** | **Traditional Approach** | **Using Cloud Computing** |
|---|---|---|
| Availability | Clusters are normally restricted in size and availability | Cloud uses virtualized compute nodes, hence it can provide instant access to large and unrestricted number of nodes upon request |
| Dynamic Configurat-ion | Clusters cannot be reconfigured for each application | Virtual Machines can be configured according to the requirements of a GIS application user |
| On Demand Access | Clusters need to be reserved for usage | Cloud Computing provides nodes immediately upon request. It's a pay-as you-go model. |
| Storage Space | Clusters are limited in storage space both locally and in network share modes | Cloud providers deliver large space for storing GIS application data. |
| Global Access (Collaborat-ion between GIS researchers) | Clusters are generally behind firewall and are not accessible across domains | Virtual Machines in Clouds have public IPs and are accessible from anywhere around the world |
| Cost | Setting up of clusters need large amount of initial investment | Clouds provide payas-you-go model for using its services. |
| Failures/ Maintainence | Failure rates of cluster nodes are very high given their use by large number of people (sharing mode). Maintenance can be very expensive and time consuming for real-time GIS Applications. | Virtual Machines are hosted in Cloud hardware that are constantly upgraded by the service provider. VMs can be migrated online to another host upon hardware failures. |
| GIS Applications | Too much overheads | Easy access for low Cost |

Table 3: Comparison of Traditional and Cloud computing approach

## VI: PROS AND CONS OF CLOUD COMPUTING IN GIS

**Pros:**

a) Lower operational costs

b) Quicker development times

c) Device independence

d) Enables heavy duty data crunching to better process and explore internet information pools

e) Pay for usage reduce fixed expenses on hardware, software, maintenance and support

**Cons:**

a) Data and processing is at the mercy of service provider and reliable internet connection.

b) Capabilities limited by marketplace demand, standardization and providers incentives.

c) Security concerns, liabilities, legal position and data processing ownership/ responsibility


Fig 5 Cloud computing in GIS

## VII: CONCLUSION

In today's global competitive market, companies must innovate and get the most from its resources to succeed. This requires enabling its employees, business partners, and users with the platforms and collaboration tools that promote innovation. Cloud computing infrastructures are next generation platforms that can provide tremendous value to companies of any size. They can help companies achieve more efficient use of their IT hardware and software investments and provide a means to accelerate the adoption of innovations

**REFRENCES**

[1] sprythink.com/glossary.html
[2]en.wikipedia.org/wiki/Cloud_computing
[3]http://itresources.whatis.com/document;5133970/tech-research.htm
[4]http://www.emqus.com/index.php?/emq/article/cloud_computing_and_saas_new_data_mining_tools_for_the_irs_998
[5]http://www.thesecurityconsortium.net/issa
[6]chrisichardson.blogcity.com/a_great_definition_of_cloud_computing.htm
[7]http://jobsearchtech.about.com/od/historyoftechindustry/a/cloud_computing.html)
[8]http://cloudcomputing.syson.com/node/927844
[9]Gruman, Galen (2008-04-07)."What cloud computing really means". InfoWorld.
[10] Amazon Elastic Compute Cloud (EC2), http://www.amazon.com/ec2/
[11]GoogleAppEngine,http://appengine.google.com
[12]MicrosoftLiveMesh,http://www.mesh.com
[13] Sun network.com (Sun Grid),http://www.network.com
[14] Won Kim: "Cloud Computing: "Status and Prognosis", in Journal of Object Technology, vol. 8, no.1, January-February 2009, pp. 65-72 http://www.jot.fm/issues/issue_2009_01/column4/
[15] Rajkumar Buyya, Suraj Pandey, and Christian Vecchiola, Cloudbus Toolkit for Market-Oriented Cloud Computing, Proceeding of the 1st International Conference on Cloud Computing (CloudCom 2009, Springer, Germany), Beijing, China, December 1-4, 2009.
[16] Rajkumar Buyya, Chee Shin Yeo, and Srikumar Venugopal, Market- Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities, Keynote Paper, Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC 2008, IEEE CS Press, Los Alamitos, CA, USA), Sept. 25-27, 2008, Dalian, China