



# International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: [www.ijarcsse.com](http://www.ijarcsse.com)

## Logical Data Model For Cloud Computing

**Nidhi Khurana**Research Scholar , Singhania University  
Faculty,Guru Nanak Institute of Management,New Delhi, India.**Dr Rattan Datta**Director, Merit Institute  
New Delhi ,India.

*Abstract--Cloud computing creates virtual Computing Environment in typical network computing pattern for sharing data and computation. Data is huge and multiple data center need collaboration in deployment and implementation of data intensive application because it is inevitable to transmit data across the different data centers. This Paper introduce cloud computing concept and highlight the various data models supported by existing cloud base storage.Cloud Data Management has to address several challenges,especially when replicated data are concurrently updated at different sites or when the system workload and the resources requested by clients change dynamically.*

*Keywords: Virtualization, Data Grid , Logical data model, Big data Processing ,multi-tenancy.*

### I. Cloud Computing

"Cloud computing" in common words can be described as "renting" or borrowing online software instead of actually purchasing and installing it on their own computers. It is the same business model as people using Gmail or Yahoo mail services, except that cloud computing goes much further. Cloud computing is where entire businesses and thousands of employees will run their computer tools as online rented products. All of the processing work and file saving will be done "in the cloud" of the Internet, and the users will plug into that cloud every day to do their computer work. Software and Platform as a Service describe the business model of users logging into a centralized hub to access their software products. Users open their files and software only while online, using only their web browser and passwords. It is similar to the idea of mainframe terminals, but cloud computing involves a much larger "cloud" (network) of processing computers at the center.

#### A. Benefits Cloud Computing:

The primary benefit of cloud computing is reduced cost for everyone involved. Software vendors do not have to spend thousands of hours supporting users over the phone... they would simply maintain and repair a single central copy of the product online. Conversely, users wouldn't have to shell out the large up-front costs of fully purchasing word processing, spreadsheet, or other end user products. Users would instead pay nominal rental fees to access the large central copy.

#### B. The Downsides of Cloud Computing:

The risk of cloud computing is that the users must place a high level of trust into the online software vendors that they will not disrupt the service. In a way, the software vendor holds its customers "hostage" because all of their documentation and productivity is now in the vendor's hands. Security and protection of the file privacy becomes even more necessary, as the massive Internet is now part of the business network. When a 600-employee business switches to cloud computing, they must choose their software vendor carefully. There will be dramatically-reduced administration cost to use cloud computing software. But there will be an increase in the risks of service disruption, connectivity, and online security.

### II. Inside Cloud Computing

According to the U.S. National Institute of Standards and Technology, cloud computing consists of five essential characteristics, three distinct service models, and four deployment Models

#### A. FIVE Essential Characteristics

- On-Demand Self-Service
- Resource Pooling
- Rapid Elasticity
- Measured Service
- Broad Network Access

#### B. Three Distinct Service Models

- Software-as-a-Service

This term generally refers to applications that are delivered to end users over the Internet or broad band access. There are hundreds of SaaS providers covering a wide variety of applications. Oracle CRM On Demand, Salesforce.com, and Google Apps are examples of the SaaS model.

- Infrastructure-as-a-Service

Refers to computing hardware (servers, storage, and network) delivered as a service. This typically includes the associated software as well, including operating systems, virtualization, clustering, and so on. Amazon Web Services, for example, offers their Elastic Compute Cloud (EC2) for compute servers, SimpleDB for database and Simple Storage Service (S3) for storage.

- Platform-as-a-Service

This model conveys how an application development and deployment platform can be delivered as a service to developers, allowing them to quickly build and deploy an SaaS application for end users. These platforms typically include database and middleware, and are often specific to a language or API. For example, Google AppEngine is based on Java and Python, Engine Yard.

### **C. Four Deployment Models**

**Public Cloud** A public cloud is one based on the standard cloud computing model, in which a service provider makes resources, such as applications and storage, available to the general public over the Internet. Public cloud services may be free or offered on a pay-per-usage model.

The main benefits of using a public cloud service are:

- Easy and inexpensive set-up because hardware, application and bandwidth costs are covered by the provider.
- Scalability to meet needs.
- No wasted resources because you pay for what you use.

Examples of public clouds include Amazon Elastic Compute Cloud (EC2), IBM's Blue Cloud, Sun Cloud, Google AppEngine and Windows Azure Services Platform.

### **Private Cloud**

Private cloud (also called internal cloud or corporate cloud) is a marketing term for a proprietary computing architecture that provides hosted services to a limited number of people behind a firewall. Advances in virtualization and distributed computing have allowed corporate network and datacenter administrators to effectively become service providers that meet the needs of their "customers" within the corporation. Marketing media that uses the words "private cloud" is designed to appeal to an organization that needs or wants more control over their data than they can get by using a third-party hosted service such as Amazon's Elastic Compute Cloud (EC2) or Simple Storage Service (S3).

### **Hybrid Cloud**

A hybrid cloud is a composition of at least one private cloud and at least one public cloud. A hybrid cloud is typically offered in one of two ways: a vendor has a private cloud and forms a partnership with a public cloud provider, or a public cloud provider forms a partnership with a vendor that provides private cloud platforms. A hybrid cloud is a cloud computing environment in which an organization provides and manages some resources in-house and has others provided externally. For example, an organization might use a public cloud service, such as Amazon Simple Storage Service (Amazon S3) for archived data but continue to maintain in-house storage for operational customer data. Ideally, the hybrid approach allows a business to take advantage of the scalability and cost-effectiveness that a public cloud computing environment offers without exposing mission-critical applications and data to third-party vulnerabilities. This type of hybrid cloud is also referred to as hybrid IT.

### **Community Cloud**

A community cloud is a multi-tenant infrastructure that is shared among several organizations from a specific group with common computing concerns. Such concerns might be related to regulatory compliance, such as audit requirements, or may be related to performance requirements, such as hosting applications that require a quick response time, for example. The goal of a community cloud is to have participating organizations realize the benefits of a public cloud -- such as multi-tenancy and a pay-as-you-go billing structure -- but with the added level of privacy, security and policy compliance usually associated with a private cloud. The community cloud can be either on-premises or off-premises, and can be governed by the participating organizations or by a third-party managed service provider (MSP).

## **III. CLOUD STORAGE**

Cloud storage is a model of networked online storage where data is stored on virtualized pools of storage which are generally hosted by third parties. Hosting companies operate large data centers; and people who require their data to be hosted buy or lease storage capacity from them and use it for their storage needs. The data center operators, in the background, virtualizes the resources according to the requirements of the customer and expose them as storage pools, which the customers can themselves use to store files or data objects. Physically, the resource may span across multiple servers. Cloud storage services may be accessed through a web service application programming interface (API), or through a Web-based user interface.

Cloud storage is made up of many distributed resources, but still acts as one highly fault tolerant through redundancy and distribution of data, highly durable through the creation of versioned copies, typically eventually consistent in regards to data replicas.

### **A. CLOUD STORAGE ADVANTAGES**

Companies need only pay for the storage they actually use as it is also possible for companies by utilizing actual virtual storage features like thin provisioning. Companies do not need to install physical storage devices in their own datacenter or offices, but the fact that storage has to be placed anywhere stays the same (maybe localization costs are lower in offshore locations). Storage maintenance tasks, such as backup, data replication, and purchasing additional storage devices are offloaded to the responsibility of a service provider, allowing organizations to focus on their core business, but the fact stays the same that someone has to pay for the administrative effort for these tasks.

### **B. DATA MODEL**

It is also important to differentiate between cloud databases which are relational as opposed to non-relational or NoSQL:

- **SQL DATABASES**

Such as Oracle Database, Microsoft SQL Server and MySQL, are one type of database which can be run on the cloud (either as a Virtual Machine Image or as a service, depending on the vendor). SQL databases are difficult to scale, meaning they are not natively suited to a cloud environment, although cloud database services based on SQL are attempting to address this challenge.

- **NOSQL DATABASES**

Such as Apache Cassandra, CouchDB and MongoDB, are another type of database which can run on the cloud. NoSQL databases are built to service heavy read/write loads and are able to scale up and down easily, and therefore they are more natively suited to running on the cloud. However, most contemporary applications are built around an SQL data model, so working with NoSQL databases often requires a complete rewrite of application code.

### **C. Deployment MODELS OF CLOUD**

There are two primary methods to run a database on the cloud:

- **VIRTUAL MACHINE IMAGE**

Cloud platforms allow users to purchase virtual machine instances for a limited time. It is possible to run a database on these virtual machines. Users can either upload their own machine image with a database installed on it, or use ready-made machine images that already include an optimized installation of a database. For example, Oracle provides a ready-made machine image with an installation of Oracle Database 11g Enterprise Edition on Amazon EC2.

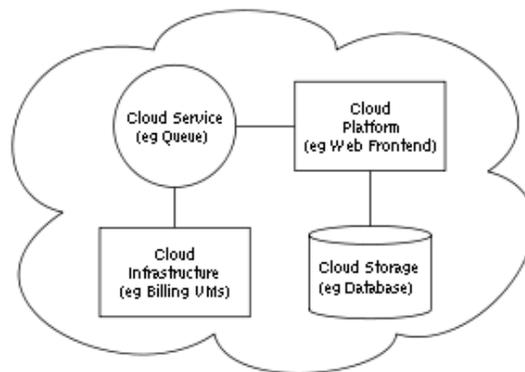
- **DATABASE AS A SERVICE**

Some cloud platforms offer options for using a database as a service, without physically launching a virtual machine instance for the database. In this configuration, application owners do not have to install and maintain the database on their own. Instead, the database service provider takes responsibility for installing and maintaining the database, and application owners pay according to their usage. For example, Amazon Web Services provides two database services as part of its cloud offering, SimpleDB which is a NoSQL key-value store, and Amazon Relational Database Service which is an SQL-based database service with a MySQL interface.

- **MANAGED DATABASE HOSTING** on the cloud, where the database is not offered as a service, but the cloud provider hosts the database and manages it on the application owner's behalf. For example, cloud provider Rackspace offers managed hosting for MySQL databases.

## **IV. Cloud Architecture**

The Cloud architecture requirements for large platforms designed for heterogeneous service types tend to be very different from requirements for application-specific platforms. The very basic architecture of cloud consists of cloud storage, cloud service, cloud platform and cloud infrastructure which can be seen as follows.



**Figure 1 : Cloud Computing Architecture**

### **A. Cloud Models and their focus area**

Different cloud models focus on different architectural areas. Some important cloud architecture models and their key focus areas are listed in the following table.

Cloud architecture models	Key architectural focus areas
Amazon EC2	Simple storage and relational database services for enterprises
Cloud Security Alliance	Enterprise security
Cisco	IT data centers and networks
IBM	Cloud service management
Storage Networking Industry Association	IT cloud storage
Windows Azure	Microsoft data center applications

Table 1. Cloud Architecture Models and their key focus areas.

### V. Actors In Cloud Computing

While making a cloud computing to work following are the actors who play various role. They may be divided into those who actually use and control the content.. All the actors are linked with each other and perform different tasks assigned to them. The following diagram shows how the actors are connected to each other.

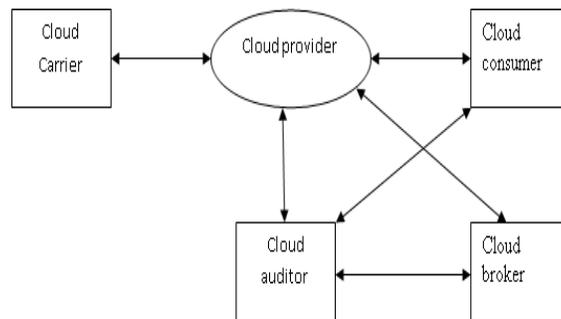


Fig 2: Actors in cloud computing

- A. Cloud Consumer:** person or organization that maintains a business relationship with, and uses service from, Cloud Providers
- B. Cloud Provider:** person, organization, or entity responsible for making a service available to interested parties
- C. Cloud Auditor:** party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation
- D. Cloud Broker:** entity that manages the use, performance and delivery of cloud services, and negotiates relationships between Cloud Providers and Cloud Consumers
- E. Cloud Carrier:** intermediary that provides connectivity and transport of cloud services from Cloud Providers to Cloud Consumers

### VI. E-R Model

Entity Relationship Model is widely used conceptual level data model. An Entity Relationship Diagram (ERD) is a visual representation of different data using conventions that describe how these data are related to each other. ER diagrams are frequently used during the design stage of a development process in order to identify different system elements and their relationships with each other

There are three basic elements in an ER Diagram: entity, attribute, relationship.

#### A. Entity

An entity can be a person, place, event, or object that is relevant to a given system.

#### B. Attribute

An attribute is a property, trait, or characteristic of an entity, relationship, or another attribute

#### C. Relationship

A relationship describes how entities interact.

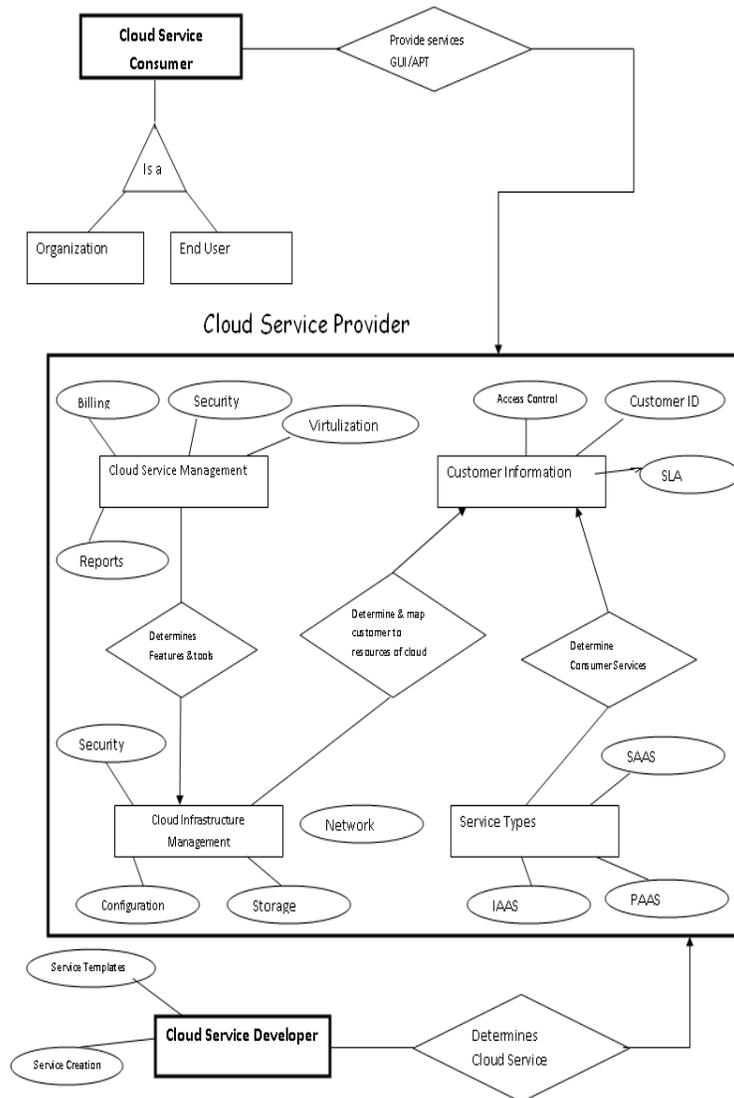
ER diagrams constitute a very useful framework for creating and manipulating databases. First, ER diagrams are easy to understand and do not require a person to undergo extensive training to be able to work with it efficiently and

accurately. This means that designers can use ER diagrams to easily communicate with developers, customers, and end users, regardless of their IT proficiency. Second, ER diagrams are readily translatable into relational tables which can be used to quickly build databases. In addition, ER diagrams can directly be used by database developers as the blueprint for implementing data in specific software applications. Lastly, ER diagrams may be applied in other contexts such as describing the different relationships and operations within an organization. Conceptual modeling forces all participants to critically evaluate the current (and potential future) business requirements of an organization. An effective conceptual model will also provide an organization with a solid schematic of its data needs, independent of any other considerations (i.e. application programming; vendor-specific database management systems).

### VII. E-R Model Of Cloud Reference Architecture

We propose an ER Model that logically depicts the architecture of cloud computing as it is the simple way to represent the data more specifically.

There are three main Entities for cloud Architecture which acts as follows



**Figure 3 : E-R Diagram Logical data model for cloud reference architecture**

#### A. Cloud Service Provider

- The cloud service provider can customize one or more cloud services for each consumer organization.
- The service provider maintains the profiles, SLAs, interfaces, access controls, and cost structure for the customized service.
- The cloud service provider can activate or deactivate the customized services for each consumer organization.
- The cloud service provider maintains a unique consumer ID for each cloud service consumer.

## **B. Cloud Service Consumer**

- The cloud consumer is the principal stakeholder for the cloud computing service.
- A cloud consumer represents a person or organization that maintains a business relationship with, and uses the service from a cloud provider.
- A cloud consumer browses the service catalog from a cloud provider, requests the appropriate service, sets up service contracts with the cloud provider, and uses the service.
- The cloud consumer may be billed for the service provisioned, and needs to arrange payments accordingly.

## **C. Cloud Service Developer**

As cloud computing evolves, the integration of cloud services can be too complex for cloud consumers to manage. A cloud consumer may request cloud services from a cloud broker, instead of contacting a cloud provider directly. A cloud developer is an entity that provides various services of cloud.

- A Cloud Service Developer will be able to create services that determine which service cloud will need to develop.
- It creates a template that feeds the Cloud service provider so that the consumer can use that template.

## **VIII. Conclusion**

Various benefits of cloud technology, in specific reference to data models were explored in the paper. Technology architecture considerations and limitation of current data model were discussed. Specifically, technology innovation and rethinking required in cloud data model and workflows were brought together and proposed in the form of an entity relationship model of cloud computing. This model can be helpful to many organizations who want to move into cloud. The ER model logically explains how actors are linked to each other and how they work for cloud. Further, we can extend the modeling to a data cube model.

## **References**

- [1] Storm warning for cloud computing <http://news.bbc.co.uk/2/hi/7421099.stm> last updated, Tuesday, 27 May 2008.
- [2]) Pengcheng Xiong Sch. of Comput. Sci., Georgia Inst. of Technol., Atlanta, GA, USA Yun Chi ; Shenghuo Zhu ; Hyun Jin Moon ; Calton Pu ; Hacigumus, H. , Intelligent management of virtualized resources for database systems in cloud environment, Data Engineering (ICDE), 2011 IEEE 27th International Conference on Date of Conference: 11-16 April 2011 **Page(s):** 87 – 98
- [3] S. Das, D. Agrawal, and A. El Abbadi. G-Store: A Scalable Data Store for Transactional Multi key Access in the Cloud. In ACM SOCC, 2010
- [4] P. Agrawal, A. Silberstein, B. F. Cooper, U. Srivastava, and R. Ramakrishnan. Asynchronous view maintenance for VLSD databases. In SIGMOD Conference, pages 179–192, 2009.
- [5] Gens, F.: IT Cloud Services User Survey, part 2: Top Benefits and Challenges (2008)
- [6] Johnston, S. (2009) Cloud Computing Types: Public Cloud, Hybrid Cloud, Private Cloud. Random rants about technology. Available at: <http://samj.net/2009/03/cloud-computing-types-public-cloud.html>.
- [7] Kevin Fogarty, Top Cloud Computing Security Risk: One Company Gets Burned, Network World, July 14, 2010, <http://www.networkworld.com/news/2010/071410-topcloud-computing-security-risk.html>.
- [8] John Markoff, Cyber attacks on Google Said to Hit Password System, NY Times, June 28, 2010, available at <http://www.nytimes.com/2010/04/20/technology/20google.html?sudsredirect=true>.
- [9] John D. Sutter, Twitter Hack Raises Questions About “Cloud Computing,” CNN, July 16, 2009, <http://www.cnn.com/2009/TECH/07/16/twitter.hack/index.html>.
- [10] Amazon. Amazon Elastic Compute Cloud (EC2), 2010 [/http://www.amazon.com/ec2/S](http://www.amazon.com/ec2/S).
- [11] Attanasio CR. Virtual machines and data security. In: Proceedings of the workshop on virtual computer systems. New York, NY, USA: ACM; 1973. p. 206–9.
- [12] Auger R. SQL Injection, 2009 [/http://projects.webappsec.org/SQL-InjectionS](http://projects.webappsec.org/SQL-InjectionS).
- [13] Basta A, Halton W. Computer security and penetration testing. Delmar Cengage Learning 2007.
- [14] Bernard Golden. Defining private clouds, 2009 [/http://www.cio.com/article/2695/Defining\\_Private\\_Clouds\\_Part\\_OneS](http://www.cio.com/article/2695/Defining_Private_Clouds_Part_OneS) [accessed on: 11 January 2010].
- [15] Boss G, Malladi P, Quan D, Legregni L, Hall H. Cloud computing, 2009, p. 4 [/http://www.ibm.com/developerswork/websphere/zones/hipods/library.htmlS](http://www.ibm.com/developerswork/websphere/zones/hipods/library.htmlS)].
- [16] Clavister. Security in the cloud, Clavister White Paper [/http://www.it-wire.nu/members/cla69/attachments/CLA\\_WP\\_SECURITY\\_IN\\_THE\\_CLOUD.pdfS](http://www.it-wire.nu/members/cla69/attachments/CLA_WP_SECURITY_IN_THE_CLOUD.pdfS).
- [17] Y Zhang, M. Roughan, N. Duffield, and A. Greeberg, “Fast accurate computation of large-scale IP traffic matrices from link loads,” in ACM SIGMETRICS, 2003, pp. 206–217.
- [18] Yin Zhang, Matthew Roughan, Nick Duffield, AT&T Labs, Shannon Laboratory, Florham Park, NJ, Proceedings of the 2003 ACM SIGMETRICS international conference on Measurement and modeling

of computer systems Pages 206 - 217 : Fast accurate computation of large-scale IP traffic matrices from link loads.

- [19]. T. Kraska, M. Hentschel, G. Alonso, and D. Kossmann. Consistency Rationing in the Cloud: Pay only when it matters. PVLDB, 2(1):253–264, 200
- [20] J. Dean and S. Ghemawat. MapReduce: simplified data processing on large clusters. In OSDI, pages 137–150, 2004.
- [21] B. Reinwald. Database support for multi-tenant applications. In IEEE Workshop on Information and Software as Services, 2010.
- [22] H. T. Vo, C. Chen, and B. C. Ooi. Towards elastic transactional cloud storage with range query support. PVLDB, 3(1):506–517, 2010.
- [23] Elmasri, R. & Navathe, S. B. (2003) Fundamentals of Database Systems. Boston: Addison Wesley.
- [24] Chen, P. P. (1976) The Entity Relationship Model - Toward a Unified View of Data. ACM Transactions Database Systems, 1(1), 9-36.