Gurudatt Kulkarni, Ramesh Sutar, Jayant Gambhir / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 1,Jan-Feb 2012, pp.945-950

"Cloud Computing-Storage as Service"

Gurudatt Kulkarni[1], Ramesh Sutar[2], Jayant Gambhir[3]

1, 2, 3 Lecturer in Marathwada Mitra Mandal's Polytechnic, Thergoan, Pune

Abstract-Storage as a Service is a business model in which a large company rents space in their storage infrastructure to a smaller company or individual.In the enterprise, SaaS vendors are targeting secondary storage applications by promoting SaaS as a convenient way to manage backups. The key advantage to SaaS in the enterprise is in cost savings -- in personnel, in hardware and in physical storage space. For instance, instead of maintaining a large tape library and arranging to vault (store) tapes offsite, a network administrator that used SaaS for backups could specify what data on the network should be backed up and how often it should be backed up. His company would sign a service level agreement (SLA) whereby the SaaS provider agreed to rent storage space on a cost-per-gigabyte-stored and cost-per-data-transfer basis and the company's data would be automatically transferred at the specified time over the storage provider's proprietary wide area network (WAN) or the Internet. If the company's data ever became corrupt or got lost, the network administrator could contact the SaaS provider and request a copy of the data. It covers the key technologies in Cloud Computing and Cloud Storage, several different types of clouds services, and describes the advantages and challenges of Cloud Storage after the introduction of the Cloud Storage reference model.

Keywords-Storage, Public Cloud, S3, SLA

1.0 Introduction

Storage as a Service is generally seen as a good alternative for a small or mid-sized business that lacks the capital budget and/or technical personnel to implement and maintain their own storage infrastructure. SaaS is also being promoted as a way for all businesses to mitigate risks in disaster recovery, provide long-term retention for records and enhance both business continuity and availability.Cloud storage is a service model in which data is maintained, managed and backed up remotely and made available to users over a network (typically the Internet).

There are three main cloud storage models:[2]

- Public cloud storage services, such as Amazon's Simple Storage Service (S3), provide a multi-tenant storage environment that's most suitable for unstructured data.
- Private cloud storage services provide a dedicated environment protected behind an organization's firewall. Private clouds are appropriate for users who need customization and more control over their data.
- Hybrid cloud storage is a combination of the other two models that includes at least one private cloud and one public cloud infrastructure. An organization might, for example, store actively used and structured data in a private cloud and unstructured and archival data in a public cloud.

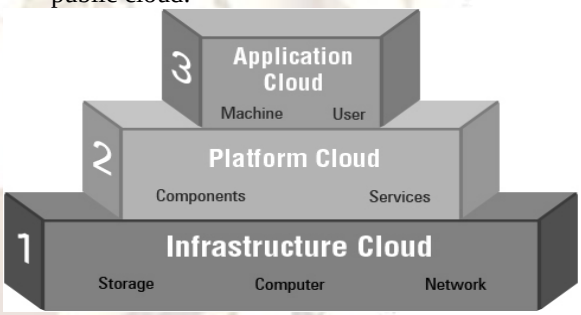

Figure1.0:- Cloud Computing Steps

Cloud computing is about moving services, computation or data—for cost and business advantage— off-site to an internal or external, location-transparent, centralized facility or contractor. By making data available in the cloud, it can be more easily and ubiquitously accessed, often at much lower cost, increasing its value by enabling opportunities for enhanced collaboration, integration, and analysis on a shared common platform. Depending on the type of provided capability, there are four scenarios where Clouds are used as showed in Fig.1.1

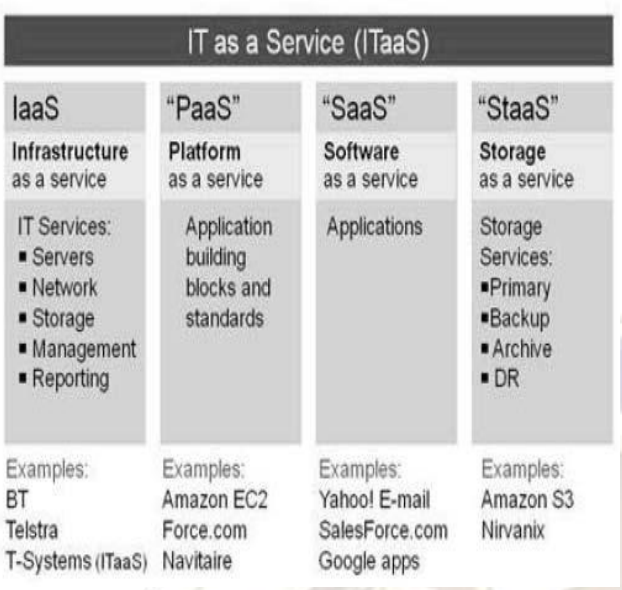

Figure 1.1 Cloud Computing Service Types With Examples [1]

1) **Infrastructure as a Service**
IPs manages a large set of computing resources, such assorting and processing capacity. Through virtualization, they are able to split, assign and dynamically resize these re-sources to build ad-hoc systems as demanded by customers, the SPs. They deploy the software stacks that run their ser-vices. This is the Infrastructure as a Service (IaaS) scenario.

2) **Platform as a Service**
Cloud systems can offer an additional abstraction level: instead of supplying a virtualized infrastructure, they can provide the software platform where systems run on. The sizing of the hardware resources demanded by the execution of the services is made in a transparent manner. This is denoted as Platform as a Service (PaaS). A well-known example is the Google Apps Engine.

3) **Storage as a Service [1]**
Commonly known as Storage as a Service (StaaS),it facilitates cloud applications to scale beyond their limited servers. StaaS allows users to store their data at remote disks and access them anytime from any place. Cloud storage systems are expected to meet several rigorous requirements for maintaining users' data and information, including high availability, reliability, performance, replication and data consistency; but because of the conflicting nature of these requirements, no one system implements all of them together.
4) Software as a Service

Finally, there are services of potential interest to a wide variety of users hosted in Cloud systems. This is an alter-native to locally run applications. An example of this is the online alternatives of typical office applications such asword processors. This scenario is called Software as a Ser-vice (SaaS).

2. 0 Cloud Storage [6]

Cloud storage is amorphous today, with neither a clearly defined set of capabilities nor any single architecture. Choices abound, with many traditional hosted or managed service providers (MSP) offering block or file storage, usually alongside traditional remote access protocols or virtual or physical server hosting. Other solutions have emerged, typified by the Amazon S3 service, that resembles flat databases designed to store large objects. The Taneja Group defines cloud storage as a specific category within the larger field of "storage in the clou" solutions. Storage in the cloud encompasses traditional hosted storage, including offerings accessed by FTP, WebDAV, NFS/CIFS, or block protocols either remotely [1]

or from within a hosted environment. Cloud storage is an evolution of this hosted storage technology that wraps more sophisticated APIs, namespaces, file or data location virtualization, and management tools, around storage. Figure 2 shows the evolution of Cloud Storage based on traditional network storage and hosted storage. There are hundreds of different cloud storage systems. Some have a very specific focus, such as storing Web email messages or digital pictures. Others are available to store all forms of digital data. Some cloud storage systems are small operations, while others are so large that the physical equipment can fill up an entire warehouse. The facilities that house cloud storage systems are called data centers. At its most basic level, a cloud storage system needs just one data server connected to the Internet. A client (e.g., a computer user subscribing to a cloud storage service) sends copies of files over the Internet to the data server, which then records the information. When the client wishes to retrieve the information, he or she accesses the data server through a Web-based interface. The server then either sends the files back to the client or allows the client to access and manipulate the files on the server itself.

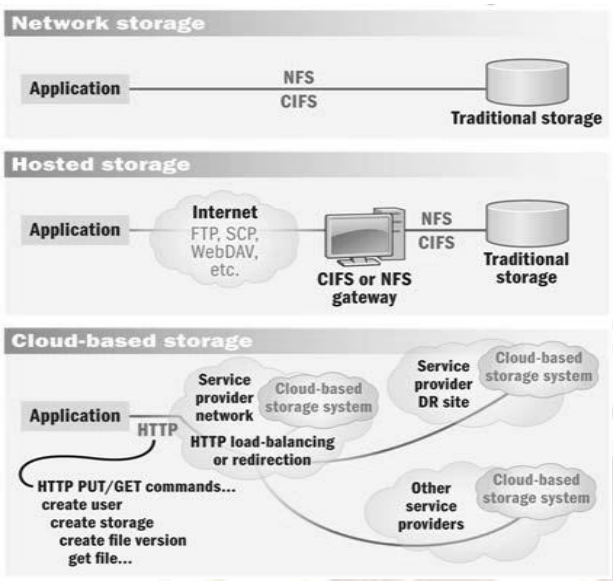

Figure 2.0 Evolution of Cloud Storage

Comedian George Carlin has a routine in which he talks about how humans seem to spend their lives accumulating "stuff". Once they've gathered enough stuff, they have to find places to store all of it. If Carlin were to update that routine today, he could make the same observation about computer information. It seems that everyone with a computer spends a lot of time acquiring data and then trying to find a way to store it.

3.0 CLOUD STORAGE REFERENCE MODEL [1]

The appeal of cloud storage is due to some of the same attributes that define other cloud services: pay as you go,
the illusion of infinite capacity (elasticity), and the simplicity of use/management. It is therefore important that any interface for cloud storage support these attributes, while allowing for a multitude of business cases and offerings, long into the future. The model created and published by the Storage Networking Industry Association™ shows multiple types of cloud data storage interfaces able to support both legacy and new applications. All of the interfaces allow storage to be provided on demand, drawn from a pool of resources. The capacity is drawn from a pool of storage capacity provided by storage services. The data services are applied to individual data elements as determined by the data system metadata. Metadata specifies the data requirements on the basis of individual data elements or on groups of data elements (containers). [1]

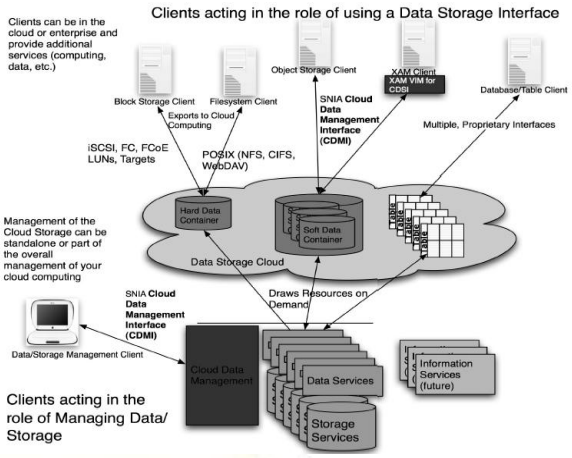

Figure 3.0 .Cloud Storage Reference Model

As shown in Fig 3.0, the SNIA Cloud Data Management Interface (CDMI) is the functional interface that applications will use to create, retrieve, update and delete data elements from the cloud. As part of this interface the client will be able to discover the capabilities of the cloud storage offering and use this interface to manage containers and the data that is placed in them. In addition, metadata can be set on containers and their contained data elements through this interface. It is expected that the interface will be able to be implemented by the majority of existing cloud storage offerings today. This can be done with an adapter to their existing proprietary interface, or by implementing the interface directly. In addition, existing client libraries such as XAM can be adapted to this interface as show in Figure 3.0.This interface is also used by administrative and management applications to manage containers, accounts, security access and monitoring/billing information, even for storage that is accessible by other protocols. The capabilities of the underlying storage and data services are exposed so that clients can understand the offering. Conformant cloud offerings may offer a subset of either interface as long as they expose the limitations in the capabilities part of the interface.

4.0 Amazon S3

Amazon S3 is storage for the Internet. It is designed to make web-scale computing easier for developers. Amazon S3 provides a simple web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the web. It gives any developer access to the same highly scalable, reliable, secure, fast, inexpensive infrastructure that Amazon uses to run its own global network of web sites. The service aims to maximize benefits of scale

and to pass those benefits on to developers. According to the Spring 2010 Storage magazine/Search Storage Purchasing Intentions survey, 14% of respondents said they're using cloud storage now, with the largest numbers using cloud storage for disaster recovery (6%). But 4% are using it to hold primary data from their data centers, and an equal number are using it for near line data storage.But before you take plunge and sign up with a cloud storage service provider, there are some things you need to know. Is cloud storage secure? How much will it cost? What services are best for SMBs? In our cloud storage services guide for beginners, we've collected our top tips and expert advice in one place so you can get answers to your most important questions. Learn about cloud backup, cloud archiving, cloud disaster recovery, and using the cloud for primary storage.

4.1 Amazon S3 Functionality [3]

Amazon S3 is intentionally built with a minimal feature set.

- Write, read, and delete objects containing from 1 byte to 5 terabytes of data each. The number of objects you can store is unlimited.
- Each object is stored in a bucket and retrieved via a unique, developer-assigned key.
- A bucket can be stored in one of several Regions. You can choose a Region to optimize for latency, minimize costs, or address regulatory requirements. Amazon S3 is currently available in the US Standard, US West (Oregon), US West (Northern California), EU (Ireland), Asia Pacific (Singapore), Asia Pacific (Tokyo), South America (Sao Paulo), and GovCloud (US) Regions. The US Standard Region automatically routes requests to facilities in Northern Virginia or the Pacific Northwest using network maps.
- Objects stored in a Region never leave the Region unless you transfer them out. For example, objects stored in the EU (Ireland) Region never leave the EU.
- Authentication mechanisms are provided to ensure that data is kept secure from unauthorized access. Objects can be made private or public, and rights can be granted to specific users.
- Options for secure data upload/download and encryption of data at rest are provided for additional data protection.
- Uses standards-based REST and SOAP interfaces designed to work with any Internet-development toolkit.
- Built to be flexible so that protocol or functional layers can easily be added. The default download protocol is HTTP. A Bit Torrent™ protocol interface is provided to lower costs for high-scale distribution.
- Includes options for performing recurring and high volume deletions. For recurring deletions, rules can be defined to remove sets of objects after a pre-defined time period. For efficient one-time deletions, up to 1,000 objects can be deleted with a single request.

4.2 Protecting Your Data

Data stored in Amazon S3 is secure by default; only bucket and object owners have access to the Amazon S3 resources they create. Amazon S3 supports multiple access control mechanisms, as well as encryption for both secure transit and secure storage on disk. With Amazon S3's data protection features, you can protect your data from both logical and physical failures, guarding against data loss from unintended user actions, application errors, and infrastructure failures. For customers who must comply with regulatory standards such as PCI and HIPAA, Amazon S3's data protection features can be used as part of an overall strategy to achieve compliance. The various data security and reliability features offered by Amazon S3 are described in detail below.

4.3 Data Security Details

Amazon S3 supports several mechanisms that give you flexibility to control who can access your data as well as how, when, and where they can access it. Amazon S3 provides four different access control mechanisms: Identity and Access Management (IAM) policies, Access Control Lists (ACLs), bucket policies, and query string authentication. IAM enables organizations with multiple employees to create and manage multiple users under a single AWS account. With IAM policies, you can grant IAM users fine-grained control to your Amazon S3 bucket or objects. You can use ACLs to selectively add (grant) certain permissions on individual objects. Amazon S3 Bucket Policies can be used to add or

deny permissions across some or all of the objects within a single bucket. With Query string authentication, you have the ability to share Amazon S3 objects through URLs that are valid for a predefined expiration time. User can securely upload/download your data to Amazon S3 via the SSL encrypted endpoints using the HTTPS protocol. Amazon S3 also provides multiple options for encryption of data at rest. Amazon S3 also supports logging of requests made against your Amazon S3 resources.

4.4 Data Durability and Reliability

Amazon S3 provides a highly durable storage infrastructure designed for mission-critical and primary data storage. Objects are redundantly stored on multiple devices across multiple facilities in an Amazon S3 Region. To help ensure durability, Amazon S3 PUT and COPY operations synchronously store your data across multiple facilities before returning SUCCESS. Once stored, Amazon S3 maintains the durability of your objects by quickly detecting and repairing any lost redundancy. Amazon S3 also regularly verifies the integrity of data stored using checksums. If corruption is detected, it is repaired using redundant data. In addition, Amazon S3 calculates checksums on all network traffic to detect corruption of data packets when storing or retrieving data.Amazon S3's standard storage is:

- Backed with the Amazon S3 Service Level Agreement.
- Designed to provide 99.999999999% durability and 99.99% availability of objects over a given year.
- Designed to sustain the concurrent loss of data in two facilities.

Amazon S3 provides further protection via Versioning. You can use Versioning to preserve, retrieve, and restore every version of every object stored in your Amazon S3 bucket. This allows you to easily recover from both unintended user actions and application failures. By default, requests will retrieve the most recently written version. Older versions of an object can be retrieved by specifying a version in the request. Storage rates apply for every version stored.

4.5 Reduced Redundancy Storage (RRS)

Reduced Redundancy Storage (RRS) is a storage option within Amazon S3 that enables customers to reduce their costs by storing non-critical, reproducible data at lower levels of redundancy than Amazon S3's standard storage. It provides a cost-effective, highly available solution for distributing or sharing content that is durably stored elsewhere, or for storing thumbnails, transcoded media, or other processed data that can be easily reproduced.

4.6 Content Storage and Distribution

Amazon S3 provides a highly durable and available store for a variety of content, ranging from web applications to media files. It allows you to offload your entire storage infrastructure onto the cloud, where you can take advantage of Amazon S3's scalability and pay-as-you-go pricing to handle your growing storage needs. You can distribute your content directly from Amazon S3 or use Amazon S3 as an origin store for pushing content to your Amazon Cloud Front edge locations. For sharing content that is either easily reproduced or where you're storing an original copy elsewhere, Amazon S3's Reduced Redundancy Storage (RRS) feature provides a compelling solution.

4.7 Storage for Data Analysis

Whether you're storing pharmaceutical data for analysis, financial data for computation and pricing, or photo images for resizing, Amazon S3 is an ideal location to store your original content. You can then send this content to Amazon EC2 for computation, resizing, or other large scale analytics – without incurring any data transfer charges for moving the data between the services. You can then choose to store the resulting, reproducible content using Amazon S3's Reduced Redundancy Storage feature

4.8. Backup, Archiving and Disaster Recovery

The Amazon S3 solution offers a highly durable, scalable, and secure solution for backing up and archiving your critical data. You can use Amazon S3's Versioning capability to provide even further protection for your stored data. If you have data sets of significant size, you can use AWS Import/Export to move large amounts of data into and out of AWS with physical storage devices. This is ideal for moving large quantities of data for periodic backups, or quickly retrieving data for disaster recovery scenarios.

CONCLUSION

Cloud Storage with a great deal of promise, aren't designed to be high performing file systems but

rather extremely scalable, easy to manage storage systems. They use a different approach to data create a very scalable storage system. Typically cloud computing provides cost effective redundancies in storage hardware. This translates into uninterrupted service during a planned or unplanned outage. This is also true for hardware upgrades which for the end user will no longer be visible
resiliency, Redundant array of inexpensive nodes, coupled with object based or object-like file systems and data replication (multiple copies of the data), to

REFERENCES

1. Jiyi WU1,2, Lingdi PING1, Xiaoping GE3,Ya Wang4, Jianqing FU1, 2010 International Conference on Intelligent Computing and Cognitive Informatics, "Cloud Storage as the Infrastructure of Cloud Computing"
2. http://searchsmbstorage.techtarget.com/feature/Understanding-cloud-storage-services-A-guide-for-beginners
3. http://aws.amazon.com/s3/
4. Storage Networking Industry Association. Cloud Storage Reference Model, Jun.2009.
5. http://www.business.att.com/enterprise/Service/hosting-services/cloud/storage/
6. Storage Networking Industry Association. Cloud Storage for Cloud Computing, Jun.2009