How to Build Cloud Computing Infrastructure

Sachchidanand Singh
IBM Software Lab
Pune (India)

ABSTRACT

In this paper we explain the concept of Cloud Computing and how to build your own Cloud Computing Infrastructure. This paper covers the different types of Cloud deployment models, service models, Cloud testing strategy, entry & exit criteria for vendor and product selection, Cloud adoption trends, customer success stories, summary and conclusion.

General Terms

Management, Performance, Design & Security.

Keywords

Managed Service Provider (MSP), Disaster Recovery (DR).

1. INTRODUCTION

A type of computing where IT-related capabilities are provided as service to external customer using internet is called Cloud Computing.

The use of computers in the cloud can make lots of sense to the end users and customers. While thinking about Cloud infrastructure a simple question comes to mind that instead of buying and maintaining our own machines, why not use the internet accessible servers on offer by the vendors in the market.

The term Cloud Computing includes the applications, services and the infrastructure.

2. WHY CLOUD COMPUTING?

The IT capacity can be adjusted quickly and easily to suite changes in the demand using Cloud Computing. The mobile and desktop applications can use services available on the cloud to synchronize information across many systems.

It is expected that by 2012, 80% of Fortune 1000 enterprises will be using some form of software-as-a-service. An organization can lease the resource on need basis by deploying the IT infrastructure and services over the network and it can avoid the capital costs of software and hardware.

3. FEATURES OF CLOUD COMPUTING

The important characteristics of Cloud Computing are listed in the following lines-

On demand access- It can be used on need basis.

Scalability- It is able to grow as per the application demand and its user base.

Elasticity- It means ability to adapt & adjust readily to different real-time conditions. It should support large number of users without affecting the speed and performance.

Cost reduction- It reduces the hardware and software cost incurred in building the applications.

Minimum management effort- It involves minimum efforts from management since support, infrastructure and operations costs are the head-ache of service providers.

Device or location independence- The applications hosted on Cloud can be accessed from any geographical location.

4. CLOUD DELOYMENT MODELS

There are 04 types of Cloud Computing deployment models available-

4.1 Public Cloud

Here vendor makes resources like applications and storage available to the public on internet. The service can be free or pay-per-usage basis.

4.2 Private Cloud

Here the infrastructure is run only for a single organization. It can be managed internally or by third-party.

4.3 Hybrid Cloud

Here the organization provides and manages some resources in-house and others are provided externally. It consists of at least one private cloud and at least one public cloud. The hybrid approach allows a business to take advantage of the scalability and cost effectiveness.

For example- An organization can use public cloud service like Amazon Simple Storage Service (Amazon S3) for archived data but can maintain in-house storage for the operational customer data.

4.4 Community Cloud

It has a multi-tenant type of infrastructure which can be shared among multiple organizations.

It can be either on-premises or off-premises and maintained by the participating organizations or by a third-party.

5. CLOUD SERVICE MODELS

There are 03 Cloud Computing service models as given below-

5.1 Software as a Service (SaaS)

Here software and its associated data are hosted centrally and is accessed by the users using a thin client.

5.2 Platform as a Service (PaaS)

It allows customers to rent hardware, operating system, storage and network capacity over the internet. This service delivery model provides a computing platform and a solution stack as a service.

5.3 Infrastructure as a Service (IaaS)

Here organizations can outsource the equipment used to support operations like hardware, storage, servers and

networking components. Here the service provider is responsible to operate and maintain these equipments.

6. CLOUD TESTING TYPES

There are several types of testing which can be performed on Cloud Infrastructure-

Load/Performance Testing- It is used to check response time, max load, other performance numbers & scalability.

Functional Testing- It is used to test feature & functionality.

Security Testing- It is performed to check whether there is any information leakage by encrypting the application or using other software, hardware & firewall.

Negative Testing- The developed application should handle invalid input or unexpected user behavior properly.

Compatibility Testing- It checks the compatibility of the developed application with the computing environment.

Compliance Testing- It is an audit of a system carried out against a known criterion.

Latency Testing- The latency testing is performed to check network response in different geographical time zones.
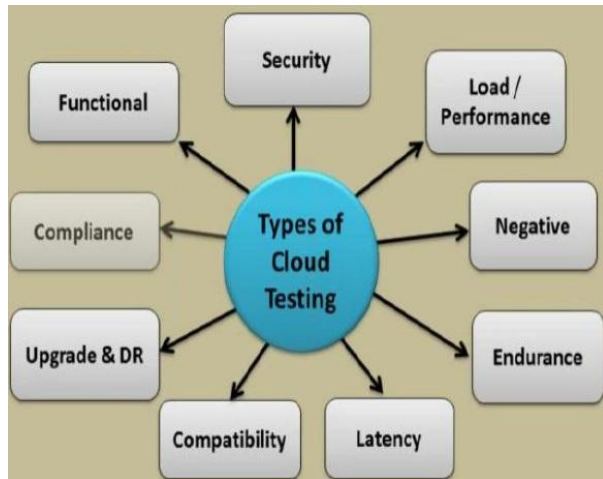

Figure 1. Types of Cloud Testing

7. BUILD CLOUD INFRASTRUSTURE

There many vendors offering Cloud services in the market like Amazon, IBM, LoadStorm, Microsoft, Google & HP etc. The IBM offers SmartCloud, a set of technologies that serve as the underpinnings of cloud computing. IBM has 13 cloud computing centers across the globe. IBM cloud offering mainly comes under three umbrellas: SmartCloud Foundation, SmartCloud Services and SmartCloud Solutions.

In a broader sense, a cloud application platform provides developer accessible services to create applications on the cloud environment.

We will build Cloud infrastructure using Microsoft Windows Azure service. It is a cloud computing platform which can be used to build and host web applications. Microsoft Windows Azure is a group of cloud technologies and each one provides a specific set of service to the application developers. It acts as a platform in order to run a windows applications and store application data on the cloud.

The Windows Azure provides a browser accessible portal to the customers so that they can create, configure and monitor their applications. A customer needs to provide a Windows

Live ID first and then he can select options either to create a hosting account to run a application or a storage account to store the data.

The Windows Azure Storage services include Blob Service, Queue Service, Azure Table Service & Windows Azure Drives.

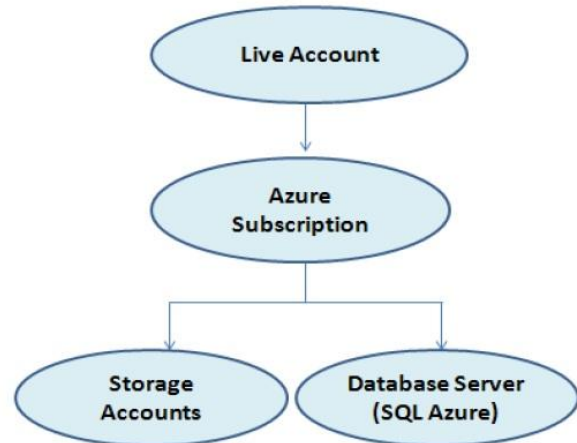Figure 2. Getting started with Windows Azure

7.1 Blob Storage Service

A BLOB (Binary Large Object) service is used to stores large unstructured text or binary data such as video, audio and images. The blob service provides following 03things: the storage account, containers and blobs. The container provides a way to organize a set of blobs within the storage account. The REST (REpresentational State Transfer) API for blob service provides a mechanism to work with the container and blob resources.

We can specify a blob type when blob is created. Once the blob has been created, its type cannot be changed, and it can be updated only by using operations appropriate for that blob type, i.e., writing a block or list of blocks to a block blob, and writing pages to a page blob. All blobs reflect committed changes immediately. Each version of the blob has a unique tag, called an ETag.

The storage service offers two types of blobs in which text and binary data can be stored-

Block blob- It is optimized for streaming.

Page blob- It is optimized for random read or write operations.

Any blob can be leased for exclusive write access. When a blob is leased, only calls that include the current lease ID can modify it. Here is a sample code for blob storage service-

//Point to Storage Account

CloudStorageAccount   storageAccount   =
CloudStorageAccount.Parse(@"DefaultEndpointsProtocol=htt
p;AccountName=ssingh;AccountKey=...");

//It creates a Blob Client

CloudBlobClient   blobclient   =
storageAccount.CreateCloudBlobClient();

//It creates a Container

```
CloudBlobContainer        blobContainer        =
blobClient.GetContainerReference("pics");
```

```
blobContainer.CreateIfNotExist();
```

```
//Upload
```

```
CloudBlockBlob        blockBlob        =
blobContainer.GetBlockBlobReference("pic1.jpg");
```

```
blockBlob.UploadFile(@"c:\ssingh\pics\pics1.jpg");
```

```
//Download
```

```
blockBlob.DownloadToFile(@"c:\ssingh\pics\pics1_backup.jpg");
```

7.2 Queue Storage Service

A Queue is a service for storing large numbers of messages that can be accessed from anywhere in the world via authenticated calls using HTTP or HTTPS. A single queue message can be up to 64KB in size, a queue can contain millions of messages.

Queue Services are used to store messages that can be read by any client who has access for that storage account. A sample code for Queue storage service-

```
//Point to Storage Account
```

```
CloudStorageAccount        storageAccount        =
CloudStorageAccount.Parse(@"DefaultEndpointsProtocol=http;AccountName=ssingh;AccountKey=...");
```

```
//Create a Queue Client
```

```
CloudQueueClient        queueClient        =
storageAccount.CreateCloudQueueClient();
```

```
//Create a Queue
```

```
CloudQueue        queue        =
queueClient.GetQueueReference("uiactions");
```

```
queue.CreateIfNotExist();
```

```
//Create Message
```

```
CloudQueueMessage        msg        =        new
CloudQueueMessage("message");
```

```
//Add Message
```

```
queue.AddMessage(msg);
```

```
//Read Message
```

```
CloudQueueMessage readmsg = queue.GetMessage();
```

7.3 Azure Table Storage

It provides ability to store the structured and non-relational entities in the Windows Azure Storage service.

Azure tables store data in the form of entity which is similar to the rows. An entity is a set of properties and a property is a name-value pair similar to the column. An entity can be up to 1 MB while Table up to 100 TB in the size.

An entity can have up to 255 properties, including 03 system properties. Every entity can have 03 system properties-

PartitionKey property

RowKey property

Timestamp property

The keys used in Windows Azure Tables facilitate efficient querying. And the PartitionKey is used for load balancing when the table service decides it's time to spread a table over multiple servers.

In Windows Azure Tables, the string PartitionKey and RowKey properties work together as an index for our table, so when defining them, we must consider how our data is queried. Together, the properties also provide for uniqueness, acting as a primary key for the row. Each entity in a table must have a unique PartitionKey/RowKey combination.

A Windows Azure table does not have a specified schema. In fact it is a structured container of rows or entities. Using Windows Azure service we design classes and a container table to which one or more classes belongs to.
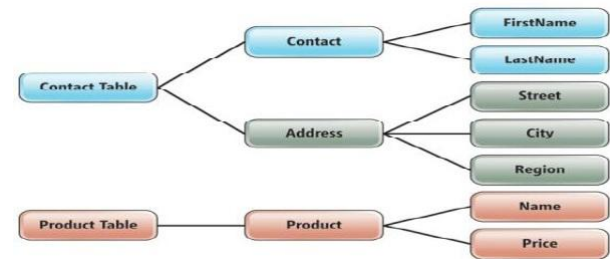

Figure 3. Rows of Windows Azure Table can represent similar or different entity

7.4 Windows Azure Drives

Windows Azure Storage allows application developers to store their data in the cloud. An application can access its data from anywhere and store any amount of data for any period of time.

It acts as a local NTFS volume which is mounted on the server's file system. The Windows Azure drive provides Windows Azure applications an ability to use a durable drive to store and access data using Windows NTFS APIs.

The data written to a Windows Azure drive is stored in a page blob defined within the Windows Azure Blob service. Since the data written to the drive is stored in a page blob so the data is maintained even if the role instance is recycled. Therefore a Windows Azure drive can be used to run an application that must maintain state.

Windows Azure Drive provides a facility of local on-disk cache of the drive's data. Using local on-disk cache is useful to reduce the transaction cost of accessing the blob service.

A drive is a page blob formatted as a Fixed NTFS VHD between 16 MB and 1TB. And a page blob can only be mounted by only one VM at a time for read or write access.

7.4.1 Windows Azure VM Size

Windows Azure provides four different VM sizes, named small, medium, large, and extra large. The local drive space allocated to each of these VM sizes is 250GB, 500GB, 1000GB, and 2000GB, respectively. The local on-disk cache space used by a Windows Azure Drive will be allocated from this drive space.

A VM can dynamically mount up to 16 drives. Our application will choose which size to use with a setting in the role's service *.csdef definition file:

```
<ServiceDefinition        name="<name-of-service>"
upgradeDomainCount="<number-of-upgrade-domains>">
```

```
<WebRole                    name="<name-of-web-role>"
vmsize="Small|Medium|Large|ExtraLarge">
```

The Windows Azure Managed Library provides the CloudDrive class for mounting and managing Windows Azure drives. The CloudDrive class is part of the Microsoft.WindowsAzure.StorageClient namespace. Once a Windows Azure drive has been mounted, we can access it via existing NTFS APIs.

The Windows Azure service can read from and write to the Windows Azure drive via a mapped drive letter (e.g., Y:\).

7.5 SQL Azure

It is a cloud based database service provided by Microsoft. It offers data storage capability as a part of the Azure Services platform. It allows users to make relational queries against stored data which can be structured/semi-structured/unstructured documents.

The SQL Azure uses Microsoft SQL Server as its backend. It provides high availability by storing multiple copies of databases. It is well suited for transactional query-based business application.

The high availability and fault tolerance is built-in feature and no physical administration is required. The SQL Azure is operated by Microsoft and has a 99.9% monthly SLA.

A customer can provision multiple SQL Azure databases to run an application and can use either the Web or Business Edition offering. The Web Edition supports up to a 5 GB maximum T-SQL based relational database. While the Business Edition supports up to a 150 GB maximum size T-SQL-based relational database.

8. CLOUD ADOPTION TRENDS

As per Gartner report Cloud Computing has the potential for broad long-term impact in most of the industries. The Cloud market will see the full range of large enterprise providers fully engaged in delivering a range of offerings to build cloud environments and deliver cloud services.
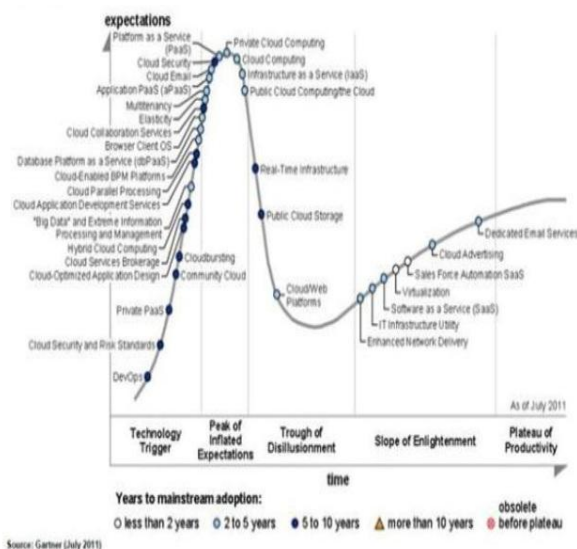

Figure 4. Gartner Hype Cycle for Cloud Computing

9. VENDOR & PRODUCT SELECTION

In this section we will define the entry and exit criteria for vendor and product selection. The following are the main criteria for selection of vendor and product-

-Scalability supports

-Security models

-Pricing

-Integration with existing systems

-Adherence to regulatory and legal compliance

-Vendor's financial stability

-Cloud Provider Integration

-Hybrid Cloud Suitability

-Vendor comparison based on the selected criteria

10. BEST PRACTICES FOR CLOUD

In this section will focus on the best practices to be followed while working with cloud infrastructure. The important points are listed below-

-Build understanding of Cloud & its offerings may be by self study, consult or involve expert

-Start small & take baby steps. Build the confidence gradually

-Plan optimal security

-Identify the right strategy

-Do risk management

-Do proof of concept (PoC's) & pilot projects

-Provision for surprises & changes

11. CUSTOMER SUCCESS STORY

11.1 Startups and Gaming Companies

-Zynga, a developer of social games with popular titles as Farmville, Mafia Wars, Cafe World, etc is leveraging the RightScale Cloud Management Platform to manage its growth, maintain high availability and keep cost in check.

-Hungama Digital Media Entertainment, the largest publisher of Bollywood movies content, started with traditional model which caused delay in implementation, effecting the time-to-market. It has now migrated about 80% of the infrastructure to the cloud using Amazon AWS cloud platform.

11.2 Small Business

-Indian independent software vendors (ISVs) such as Ramco is offering a SaaS version of their ERP offering called Ramco OnDemand ERP with several customers such as Hindustan Motors & KVK Energy.

11.3 Enterprise

-The New York Times has made all the public domain articles from 1851-1922 available free of charge with the help of Amazon Elastic Compute Cloud (EC2) and Simple Storage Service (S3).

-The event CONCERTO DE IDÉIAS conducted by Association for Promotion of Brazilian Software Excellence (SOFTEX) was hosted by the Cloud Center in Sao Paolo using IBM Idea Factory which was a Web 2.0-based cloud computing platform.

12. CLOUD COMPUTING BENEFITS

Running applications and storing their data in the cloud can have clear benefits. Rather than buying, installing, and operating its own systems, for example, an organization can rely on a cloud provider to do this for them. Also, customers pay just for the computing and storage they use, rather than maintaining a large set of servers only for peak loads. And if they're written correctly, applications can scale easily, taking advantage of the enormous data centers that cloud providers offer.

There are following benefits of Cloud Computing-

-Reduction in capital expenditure

-Optimal utilization of resources hardware & software

-Pay per use model

-Fast and flexible deployment & provisioning

-Scalability and Elasticity

-Broad network access

-Resource pooling

-Ideal for virtual teams spreads across different geographical location

13. CHALLENGES OF CLOUD

In this section we will highlight the challenges faced while working on the cloud computing infrastructure. The main challenges are listed below-

-Dependency on internet with connectivity, latency, bandwidth and availability issues

-High initial setup or migration cost in some cases

-Lack of standards

-Security in Public Cloud

-Supportability of legacy systems

-Lack of control

-Limitations in terms of services available with Cloud

-Lack of tried and tested solutions and guidelines

14. SUMMARY AND CONCLUSION

Cloud Computing Infrastructure can be built using Microsoft Windows Azure storage services like blob, queue, azure table and azure drives. As per Gartner report the Cloud Computing is among one of the top 10 strategic technologies for 2012 and it has the potential for the broad long-term impact.

The research undertaken for Microsoft by International Data Corporation (IDC) indicates that the cloud will support creation of 14 million jobs worldwide.

15. ACKNOWLEDGMENTS

16. REFERENCES

[1] Gartner Hype Cycle for Cloud Computing, http://www.gartner.com/technology/research/hype-cycles/

[2] Gartner's 2011 Hype Cycle Special Report Evaluates the Maturity of 1,900 Technologies, http://www.gartner.com/it/page.jsp?id=1763814

[3] Gartner Hype Cycle for Cloud Computing, http://www.gartner.com/hc/images/215650_0001.gif/

[4] Storage- BLOB (Binary Large Object) storage, http://www.windowsazure.com/en-us/home/features/storage/

[5] How to Use the Blob Storage Service from Java, http://www.windowsazure.com/en-us/develop/java/how-to-guides/blob-storage/

[6] How to Use the Queue Storage Service from Java, http://www.windowsazure.com/en-us/develop/java/how-to-guides/queue-service/

[7] IBMSmartCloud. IBM cloud computing- Rethink IT-Reinvent business, http://www.ibm.com/cloud-computing/us/en/

[8] Windows Azure Table Storage – Not Your Father's Database, http://msdn.microsoft.com/en-us/magazine/ff796231.aspx

[9] Using the Windows Azure Storage Services, http://msdn.microsoft.com/en-us/library/windowsazure/ee924681.aspx

[10] Data Management, http://www.windowsazure.com/en-us/home/features/sql-azure/

[11] Cloud Services, http://www.windowsazure.com/en-us/home/scenarios/saas/

[12] Cloud Computing, Hands On, https://manual.cs50.net/images/b/b1/Cloud-computing-hands-on.pdf

[13] Gartner Identifies the Top 10 Strategic Technologies for 2012, http://www.gartner.com/it/page.jsp?id=1826214