Chapter

2

Cloud Computing

CHAPTER AUTHORS

Boa Ho Man

Goh Hao Yu Gerald

Tan Wei Hao Benjamin

CONTENTS

1 INTRODUCTION TO CLOUD COMPUTING

Fujitsu will invest $537 million in cloud computing for 2011. -- The Nikkei Daily

Microsoft to spend 90% R&D budget on Cloud. -- The Times of India

What is cloud computing? In this chapter, we will give you an introduction to cloud computing, share with you some information regarding the different aspects of cloud computing and cover information security in cloud computing.

We will fill you in on what cloud computing is, the principles behind cloud computing and the different types of cloud computing service providers in this section.

1.1 What is Cloud Computing

Cloud computing is Internet-based computing, whereby hardware and software resources are provided to users on-demand. It is a by-product and consequence of the ease-of-access to remote computing sites provided by the Internet.[i] Through cloud computing, you are able to use software delivered through the Internet on the browser without any installation, host an application on the Internet, set up your own remote file storage and database system and more.

1.2 The Five Principles of Cloud Computing

Cloud computing is different from your traditional web service because of the principles[ii] behind cloud computing. These principles are

- ***Resource pooling:*** Cloud computing providers harness large economies of scale through resources pooling. They put together a vast network of servers and hard drives and apply the same set of configurations, protection and the works for them.
- ***Virtualization:*** Users do not have to care about the physical states of their hardware nor worry about hardware compatibility.
- ***Elasticity:*** Addition of more hard disk space or server bandwidth can be done with just a few clicks of the mouse on-demand. Geographical scalability is also available in cloud computing - one can choose to replicate data to several data centres around the world.
- ***Automatic resource deployment:*** The user only needs to choose the types and specifications of the resources he require and the cloud computing provider will configure and set them up automatically.
- ***Metered billing:*** Users are charged for only what they use.

These principles allow cloud computing to bring more cost-savings, automation and flexibility to the users, compared to using a traditional web service provider.

1.3 Types of Cloud Computing Providers

There are three kinds of cloud computing providers - Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS).

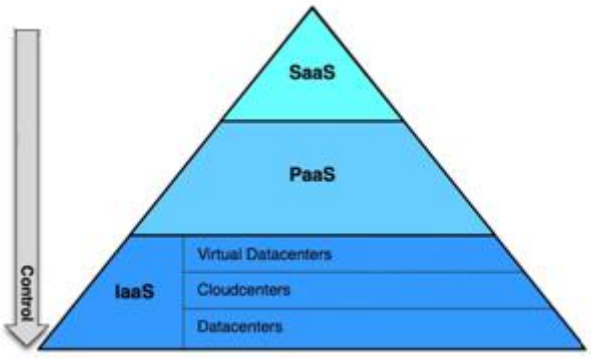

Figure 1 - The XaaS pyramid[iii]

IaaS providers offer hardware and the bare minimum software for users to develop on e.g. virtual servers, hard disk space. Users have broad control over the services offered by IaaS providers as they are able to configure settings to a very large extent and they are free to implement and utilize any form of software and programming environment on top of the services. An example of an IaaS provider is Rackspace (see Figure 2 - Rackspace is an *IaaS that* **allows users to start up a server and operating system of their choice[]**).

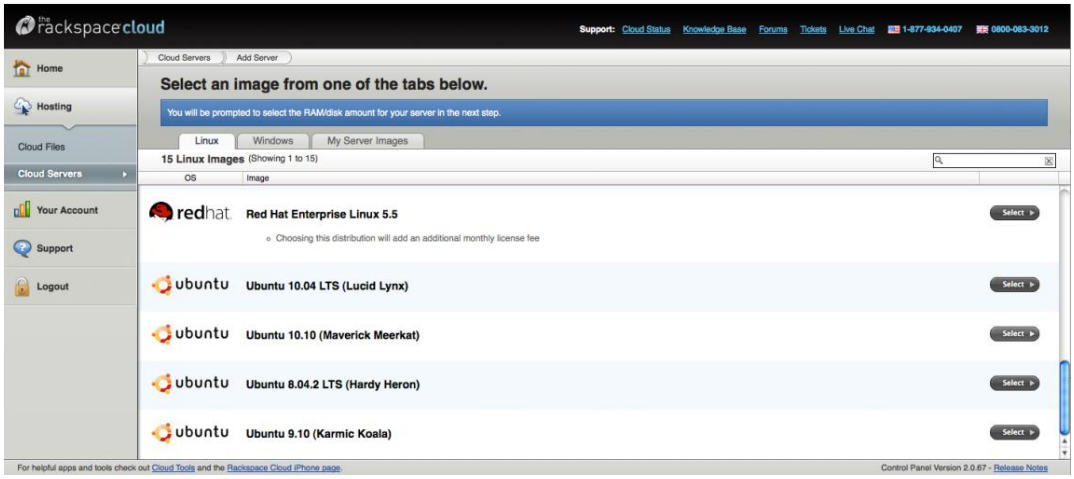

Figure 2 **- Rackspace is an IaaS that allows users to start up a server and operating system of their choice[iv]**

For services offered by PaaS providers, users have very little control over their software and programming environment. This is because PaaS providers implement a software layer over the hardware they offer, forcing users to work with the providers' software layer. This is not necessarily a bad thing as PaaS providers reduce the technical expertise needed for users to create their own web application. An example of a PaaS provider is Google App Engine (see *Figure 3*). One can make use of Google App Engine to host an application on the web and make use of the database and file storage services provided by Google App Engine to complement their application. However, As Google App Engine is a PaaS, only Java and Python can be used to code the application.

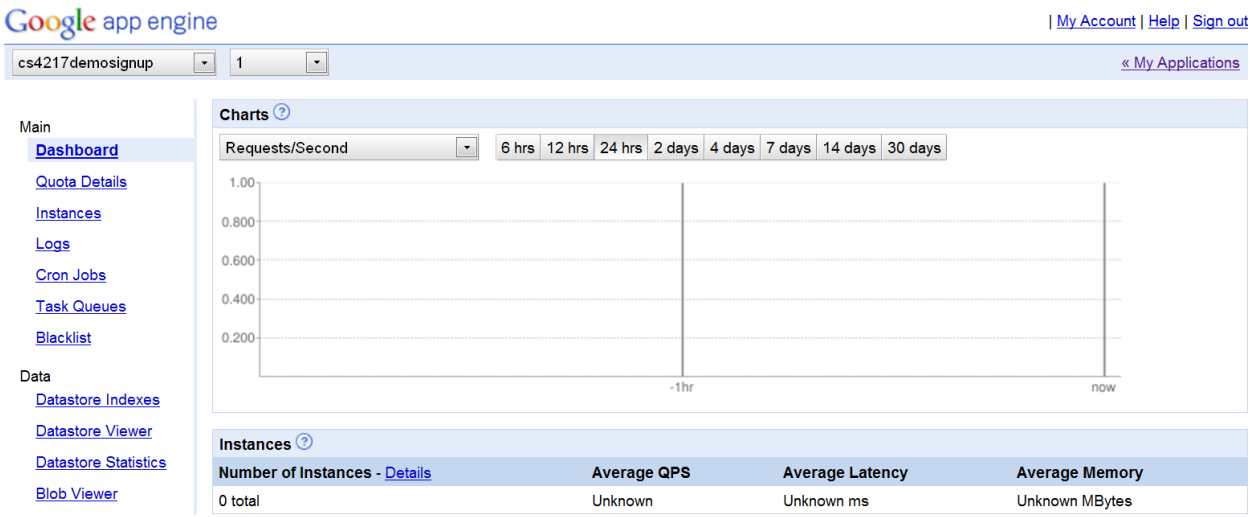

Figure 3 - **Google App Engine comes with a server, database and file storage[v]**

SaaS providers basically offer software for users over the Internet through a web browser. An example of a SaaS is Google Docs, where one can edit documents through the software delivered over the Internet. The main advantage of using a SaaS is that you do not have to worry about installation, storage space, data loss due to PC crashes or patches - the cloud computing provider handles all that and simply delivers the software to you in your web browser.

2 SERVERS IN THE CLOUD

Servers are the backbone of cloud computing. In this section, we will look at platform virtualization, a technology used by cloud computing providers to offer servers to users and some concepts that you will need to know when you begin utilizing cloud computing servers.

2.1 Platform Virtualization

We will go through what exactly is platform virtualization and its features.

2.1.1 What is Platform Virtualization

Platform virtualization is a technology to abstract physical hardware resources of a single server into a number of virtual computing environments, allowing multiple operating systems to be installed into each of these environments.

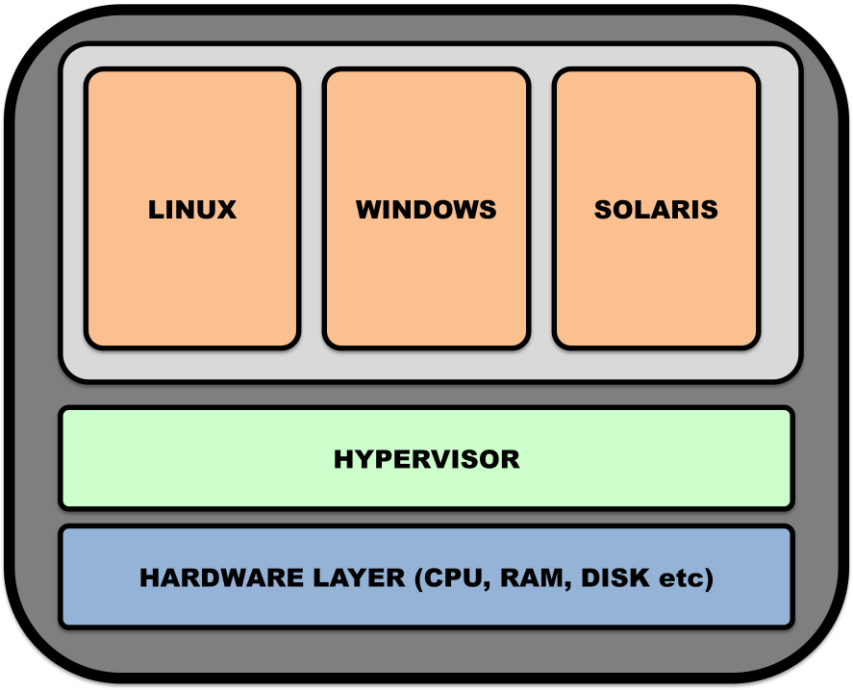

Figure 4 - Platform Virtualization

The heart of the virtual machine is the *hypervisor*. This is the software that sits between the hardware and operating systems. Its main role is allocate system resources. Each of these individual operating systems behaves as if it has the resources of the whole server to itself. This is illustrated in *Figure 4*. Each of the three operating systems is in fact three different virtual machine image. Each of these machine images is a snapshot of an operating system, and these images are then loaded onto the virtual machine environment.

It is important to realize that you do not pay for an entire server. Instead, cloud providers provide you with a virtual machine image, one of many that is loaded on the virtual computing environment.

2.1.2 Features of Platform Virtualization

Through the features of platform virtualization, cloud computing service providers are able to offer servers that are cost-effective, rapidly deployable and scalable. These features are namely,

Increased server utilization: The average server in an enterprise data center without platform virtualization has a typical utilization of 5 to 10 percent. Even at peak loads, the maximum is around 20 percent. Cloud computing providers leverage on this fact by running multiple virtual servers on one physical server through platform virtualization and are thus able to lease out more servers (virtual) with the same amount of hardware resources.

Removal of hardware dependencies: Cloud computing service providers and users no longer need to bother about hardware compatibility issues. They just need to focus their software applications and service-level agreements. Furthermore, developers only need to worry about the operating system they need, and how much performance they require.

Removal of software dependencies: With virtualization, there is no need to worry about device driver dependencies as they are handled by the virtual machine environment.

Quick Deployment and Teardown: Virtualization makes it extremely easy to deploy a "server" – simply load the virtual machine image. It is also easy to tear down a machine image.

2.2 Cloud Computing Server Concepts

Starting your first cloud-computing server requires some knowledge of the configurations that are offered. These configurations are common across cloud computing service providers.

Machine image: A machine image is a snapshot of an operating system. The image contains an entire operating system and pre-installed software, if any. Most providers offer default Linux images of various distributions and machine architectures (32/64-bit). Windows Server images are starting to be more commonplace too. You can choose to customize your own server image, which would facilitate deployment, especially if you have a fixed set of software tools that your environment requires.[vi]

Cloud instance: A cloud instance is an instantiation of the virtual machine image in a cloud infrastructure. When an instance is launched, the hardware in the cloud's infrastructure is provisioned for you. However, it is up to the cloud provider to decide the exact hardware specifications, and most of the time, this consumer has no idea of the underlying hardware infrastructure.

Elastic IP address: This is just a static IP address. The only difference is that you can programmatically map this address to any virtual machine instance without a network administrator's help, and without waiting for the DNS to propagate the change.[vii]

Availability Zone: Think of this as the location of the data-center. Amazon currently has availability zones in the US and Singapore. Microsoft's Windows Azure has its cloud infrastructure in the US, Europe (Ireland, Netherlands) and Asia (Hong Kong, Singapore).

3 File Storage in the Cloud

Following sections will introduce the key concepts of cloud storage, show their various benefits and discuss some of the current concerns about cloud storage.

3.1 General Architecture

Most cloud storage providers generally follow a three-layer architecture. In the figure below, you can see an illustration of the general architecture and some of the characteristics[viii] that are tied to this architecture in current cloud storage.

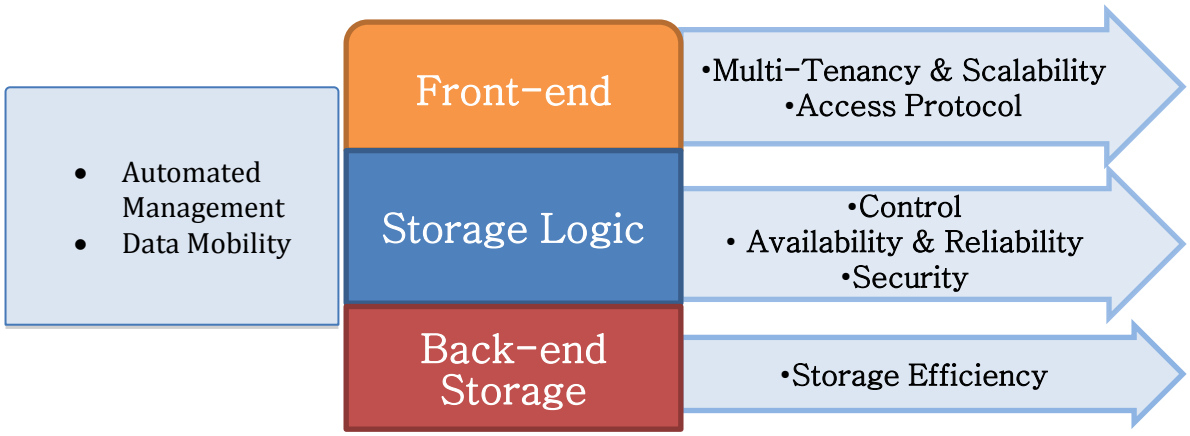

Figure 5 - Generic cloud storage architecture[ix]

- **The front end** is in charge of the communication between the clients and the servers. There will be different APIs to access the actual storage. This layer is also about achieving results such as *multi-tenan*cy, a term we will explain in the next chapter. In addition, it provides the means for different types of scalability through various methods.
- **The storage logic layer** handles a variety of features, and is in charge of certain administrative procedures such as ensuring a high level of availability and reliability for instance. It is also a form of security perimeter. Furthermore, it acts like a controller for cloud storage.
- **The back-end** focuses on the actual implementation of the physical storage of data with protocols such as the GFS (Google File System). It involves the use of various ways to increase storage efficiency and in a way to drive the infrastructure costs down.

3.2 Defining Characteristics

In this chapter, we will go into more details in some of these characteristics listed in the diagram in the previous chapter.

Multi-tenancy, which refers to the ability for a single instance of services to serve multiple clients or tenants, also applies to several different layers of the cloud storage stack and this allows numerous clients to subscribe to the same cloud computing capabilities while retaining privacy and security over their sensitive data.

Automated Management is an important quality of the cloud storage. Generally, cost can be divided into two categories: the cost of the physical storage infrastructure itself and the cost of managing it. The management cost is hidden but is really a substantial component of the overall cost in the long run. The cloud storage must be able to add new storage and automatically configure itself to accommodate it and to find errors automatically. Automated management is relatively critical to cloud storage because what cloud computing is selling is essentially convenience.

Consistency in performance around the globe is one of the core reasons to choose cloud storage over traditional file hosting. With traditional file hosting, files are typically stored on one server hence clients who are far away from that server will suffer from bad performance. With cloud storage, there are 2 levels of geographical scalability. Firstly, the file is distributed around multiple servers in the region where your original data is stored at. Secondly, there are on-demand CDNs (content delivery networks). These are networks that have servers distributed globally to allow fast content delivery to clients anywhere in the world. By using CDNs, cloud storage can also achieve the same high level of consistency in performance all around the world and also make your data more mobile because it is available and highly accessible at all parts of the world.

Unique access methods are also one of the main differences between cloud storage and traditional storage. Many cloud storage providers now implement multiple access methods but the most prevalent one is still the Web-Service API. These are implemented by following the REST (Representational State Transfer) architecture. The architecture is used to develop protocols over the layer of HTTP to harness HTTP as a transport utility. By following this architecture, APIs are stateless and therefore relatively efficient. Bigger cloud storage providers such as Amazon (S3) and Microsoft Azure are both currently using this approach. There are also other forms of access methods such as file based APIs such as NFS and FTP and these two APIs are adopted by IBM Smart Business Storage Cloud.

High Reliability is one of the cornerstones of cloud storage. One might think that with the technological advances today, hard-disk failures and mass information losses are no longer common. On the contrary, hardware failures are inevitable and could be devastating if backups were not adequate. Cloud providers generally use two different approaches to ensure reliability.

***Replication*:** Big cloud service providers generally have the same information stored on multiple machines. In the case of Google, their cloud back-end storage is typically split into huge clusters and entirely broken into chunks of 64mb each. Each of these chunks is uniquely identifiable and they are replicated to multiple servers in their data centers. Furthermore, these machines are run on different power supplies. That way, even if one power supply fails, clients will still have access to their data.

***Reconstruction*:** Some service providers also use data-reconstruction algorithms to help with lost or damaged data. One of these algorithms is IDA (Information Dispersal Algorithm). This algorithm is able to construct a full set of data from multiple parts of the data that has been distributed before-hand. For example, if the data is divided into 4 parts, it can still be reconstructed if one site holding one part of the data is lost. Different ratios are possible to implement as well. E.g. 20 parts will allow 8 failed sites. These pieces of data are usually distributed at different geographical locations to reduce the chances of all parts of the data being lost at one time.

Good cost-to-storage ratio is another characteristic of cloud storage that is worth mentioning. To reduce cost, more data must be stored with the same hardware resources. One common way to do this is to use data-reduction algorithms to reduce the resources data take up. There are notably 2 different approaches to this: *compression*- the encoding of the data in another more economical representation to achieve data reduction, *de-duplication*- the removal of any identical copies of data found through the scanning of data signatures.

High levels of security are essential to cloud storage, in particular, when storing sensitive data on the cloud. This topic will be further discussed in Section 5.

3.3 Concerns about Cloud Storage

The integration of cloud computing technologies to existing IT infrastructure and the performance and latency of cloud storage are two concerns that are specific to cloud storage.

3.3.1 Integration

Before utilizing the cloud storage, an organization will need to integrate the cloud storage into their existing work-flow or other forms of offline storage facilities. The fact is normal file servers and cloud storage services do not use the same file access protocols. Servers use block protocol access to their storage, but cloud storage services generally only provides web protocol access such as REST-based APIs, SOAP-based APIs which are APIs designed on-top of the HTTP protocol to provide access with better efficiency. Each of the major providers has their separate set of APIs to handle the operations. This complicates things a little.

Mature organizations generally have more complicated existing file storage workflows. A considerable amount of time, money and attention will have to be spent to integrate the use of cloud storage into their existing workflows. On the other hand, a younger organization with a less complex infrastructure will not face the same problem because it will certainly be easier to integrate cloud storage into a workflow that is not so developed yet.

3.3.2 Performance and Latency

Cloud storage may be used by organizations for periodic backups of massive amounts of data. These back-up operations will involve sending data to a geographically distant location. This will inevitably be slower compared to offline storage solutions. While cloud storage is more

convenient to use, is immediately scalable for organizations and more reliable, but unfortunately speed-wise it still trails behind offline storage solutions.

In general, cloud storage today is targeted at less performance demanding operations. Organizations should generally leave the operations having a stringent requirement for performance outside of cloud-storage. These include real-time transactions in banks for example.

4 DATABASE IN THE CLOUD

Database is a key component in most computing infrastructures. Database allows users to store data in an organized manner and retrieve them easily. In this section, we will discuss a new type of database that is gaining popularity especially in cloud computing, the non-relational database (NOSQL), and compare it with the relational database, from a cloud perspective. We will also look at some common database architectures that cloud computing providers employ and issues regarding programming environments that come with the different types of database.

4.1 Non-Relational

Non-relational databases are commonly referred to by the term "NOSQL"[x] (pronounced “No S Q L”). They are made up of individual tables and these tables cannot have defined relationships between them, unlike in relational databases. For example, in the database schema as shown in *Figure 6*, one can retrieve the account balance of a specific Customer given the Customer's name through table joins using SQL due to the Primary Key/Foreign Key relationship. In a non-relational database for the same schema, without the relationship, the developer has to use application code to obtain the Customer's account number and then access the Account table and match the account number obtained previously to retrieve the balance.

Table: Customer

Fields: `name PRIMARY KEY, address, accountNumber`

`FOREIGN KEY accountNumber REFERENCES Account.number`

Table: Account

Fields: `number PRIMARY KEY, pinCode, balance`

Figure 6 - Primary Key/Foreign Key relationship

4.2 Relational vs. Non-Relational

Many cloud computing providers offer users both relational and non-relational databases. Both types of database are scalable in the cloud and can be highly available. In terms of speed, we have done up our own series of benchmarking tests for some relational and non-relational databases cloud computing services. Our findings[xi] indicate that relational cloud databases perform create, update and delete operations faster than non-relational cloud databases. However, for read operations, non-relational cloud databases do perform better than relational cloud databases. Usage-wise, both relational and non-relational cloud databases are as easy to use as the other. This is because cloud computing providers take on most of the burden of database administration, especially for relational databases, as relational databases usually come with heavy database administration workload compared to non-relational databases.

4.3 Architectures

Cloud databases providers often let users choose from multiple database architectures. Since these different architectures have different levels of database consistency, latency and costs, you need to understand the architectures to have a better idea of which service suits your application's needs. We will discuss two different architectures which are being used by major cloud service providers here - the Master/Slave architecture [xii] and the architecture based on the Paxos algorithm [xiii].

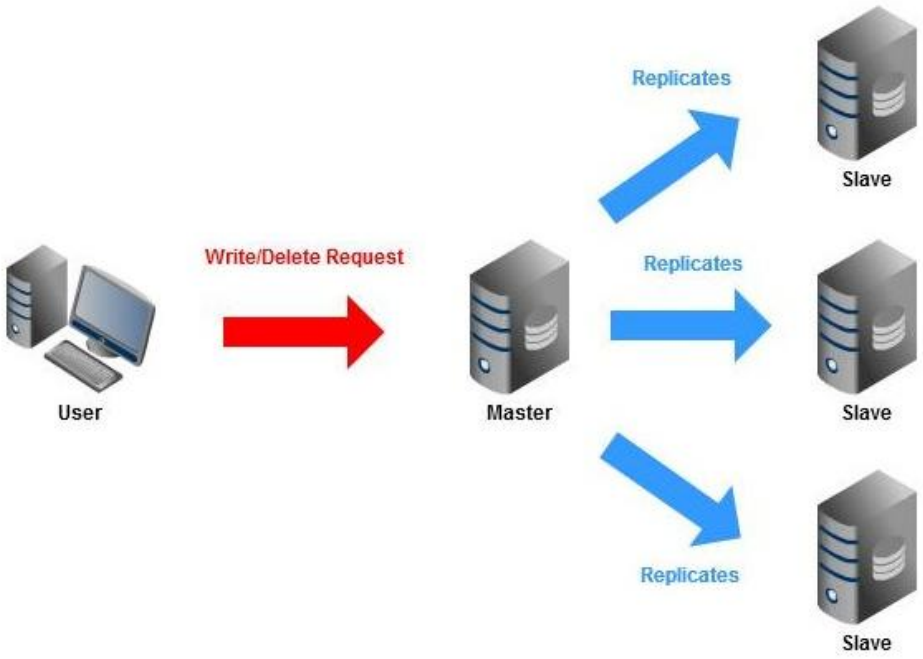

Figure 7 - **The Master/Slave database architecture**

In the Master/Slave database architecture (see Figure 7), a database server acts as the *Master*. When the user sends in a write/delete request to his database, the request goes to the *Master* database server. The *Master* database server checks against and updates its own database and then asynchronously replicates the update in other *Slave* database servers.

Figure 8 - Paxos architecture

For the Paxos architecture (see Figure 8), when the user sends a write/delete request, this request goes to a network of several database servers. The different database servers will check the requests against their own databases and states and then communicate with each other to affirm the request.

There are numerous pros and cons of using one database architecture over the other.

- ***Master/Slave architectures uses lesser write/delete CPU time:*** Databases built on the Paxos architecture use more write CPU time than databases built on the Master/Slave architecture due to the servers needing to communicate with each other to affirm the write/delete request, unlike in the Master/Slave database architecture, where the *Master* itself affirms the request and sends the affirmed changes to the *Slaves*.
- ***Master/Slave architecture has lower write/delete latency:*** It is higher for the databases built on the Paxos architecture as the affirming of requests between the various servers takes time.
- ***Master/Slave architectures have stronger query consistency:*** Query consistency of databases built on Paxos architecture is "eventual" as they require time to process certain tricky requests among the data centres - a read request might come in before the processing of a previous write/delete request can be completed, resulting in the read request not getting the most up-to-date results.
- ***Paxos architectures have higher availability and reliability:*** Databases built on Paxos architecture do not suffer from downtimes like their counterparts built on Master/Slave architecture. For example, if the *Master* data centre for a certain database built on Master/Slave architecture goes down for maintenance, write/delete requests will not be processed. However for databases built on the Paxos architecture, the user's database can still be updated even if a data centre goes down, as long as there are other data centres that remain operational, since any of the data centres can process the write/delete request.

The overview of the two different architectures should now give you a better understanding of why some databases offered by cloud computing providers are more costly than others or why there is higher consistency for some types of database over the others for example. Moreover, the two types of architectures covered can also serve as examples for you to make use of infrastructures offered by cloud computing providers to model and build your very own cloud database architecture. For example in *Figure 9*, the Amazon Read Replica instances act as the slave databases in a *Master/Slave* architecture.

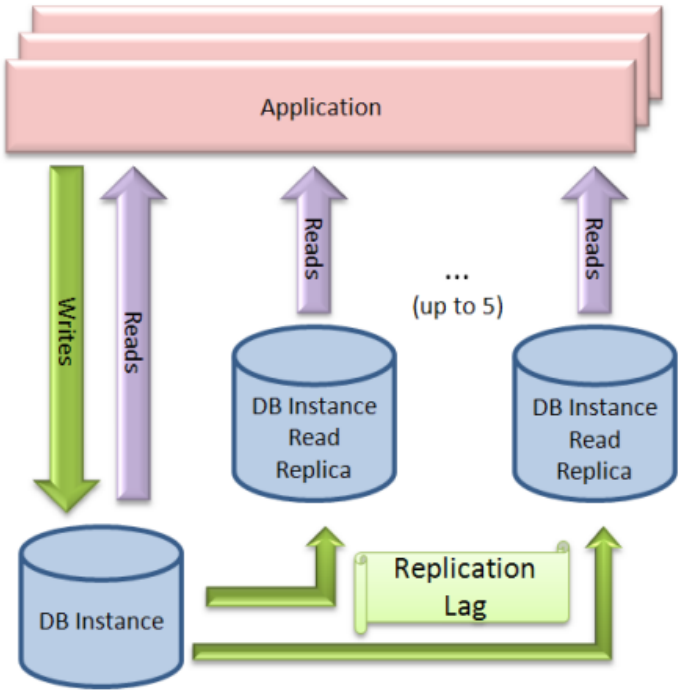

Figure 9 - **Build a Master/Slave database using Amazon RDS and its Read Replica complement[xiv]**

4.4 Programming Environment

Aside from deciding whether to use a relational or non-relational database what architecture you want your database to be built upon, you should be concerned with the programming environments that come with the database. This is because the programming environment contributes to the perceived speed of database operations from the client side and the migration possibility of your database.

Different databases can be accessed by only certain programming languages and their APIs. For example, if you use Google App Engine's non-relational database, you can only use Java or Python to access it. However, if you choose to use MySQL hosted on Amazon Web Services, you are able to use a myriad of programming languages such as C#, Visual Basic and Java. The runtime of programs coded in these various languages differ, impacting the end user's experience, because the information interchange of client-server and server-database depend mostly on the programming environment. The migration possibility of your database from one cloud-computing provider to another, or even from a cloud-computing provider to your own server, matters. There might be unexpected circumstances that occur, forcing you to drop your current cloud-computing provider. Therefore, before you actually settle on a particular database from a particular cloud-computing provider, consider if you can easily port your application and its database code after you have implemented it.

5 Information Security in Cloud Computing

Currently, there is no real security standard associated with cloud computing yet. Cloud service providers implement their individual proprietary standards and security technologies. In a vendor cloud model, it is ultimately up to clients to ensure that security in the cloud meets their own security standards and agree with their policies through requirements gathering and provider risk assessments. Due to the nature of this topic, much of this information will be provided in the perspective of SMBs (Small-medium Businesses) to large organizations. We will

look at security features offered by cloud computing providers, the security risks involved in cloud computing and some recommendations regarding security in the cloud.

5.1 Security Features in Cloud Computing

There are a number of security benefits of using cloud computing. However, the level of security depends on the provider. In this chapter we have sourced out some of the key security features for you from various sources [xv][xvi] as a guide to some of the security features and benefits of cloud computing.

- **Economies of Scale**
  The pooling of resources on a large scale translates to security benefits in two areas.
- ***Costs:*** Security measures are going to be cheaper when implemented on a large scale. The same amount of investment could buy better security protection such as packet filtering, patch management and the hardening of virtual machine instances for example.
- ***Expertise*****:** Major cloud service providers such as Amazon, Google, Azure, and Rackspace all have massive resources and expertise in security fields at their disposal. Because of the concentrated effort at security, they may be able to provide better security measures than SMBs (small-medium businesses) can possibly achieve. It can also simplify the process for large organizations.
- **Centralized Data**
  Using a centralized data model makes securing and managing of data easier.
- ***Reduced physical data leakage*****:** There are many ways internal data can be leaked through physical means, loss of company thumb-drives, laptops and backup disks all contribute to this. With cloud computing, we are able to avoid this problem because data is now stored in the cloud away from the physical devices.
- ***Monitoring benefits*****:** With centralized storage, it is easier to control and monitor. It is both easier and cheaper to implement security controls on centralized data than for individual clients because there is only place where targeted attacks can happen. The resources can then be reallocated quickly for filtering, traffic controlling, verification, encryption and other security measures and this will improve resilience against security threats.
- **Incident Investigation**
  Cloud storage helps speed up the process incident investigation in two ways.
- ***Forensic readiness:*** With Infrastructure as a Service (IaaS) providers, a dedicated forensic server image in the same cloud can be built and placed offline. When a security incident happens, you can have the server up absolutely instantly! In using cloud computing, you are avoiding all the clunky hardware provisioning that is required and you are also able to have your server as and when you need it.
- ***Decreased evidence acquisition time:*** Data transfer between 2 servers in the same cloud is extremely fast so the evidence transfer time between the compromised server and forensics server will be reduced. This will help speed up the forensics process and the quicker the security forensics process can be done, the quicker the server can be back up and running.
- **Logging**
  Cloud storage abstracts away clunky provisioning for logging.
- Automatic logging: Logs are an important link in security investigations and setting up new defensive infrastructures. However, logging is often an afterthought in many organizations. Because of clunky resource allocation, there are often little or no logs at all. Cloud storage changes all of this by introducing automatic logging for your applications deployed in the cloud.
- **Gold Images**
  Cloud storage helps with the deployment and management of gold images.

- ***Ease of Management*:** A gold-image is a unit or an instance that has been fully subjected to proper stability and vulnerability tests and is ready for public deployment. Now we would like to deploy this to multiple units or instances. In normal systems, all the units would be running on separate platforms and environment which makes it really difficult to track each individual setting. On the other hand, because of the uniformity of cloud computing's instances in platforms and environment, it makes it a lot easier to deploy and manage each of these gold images.

5.2 Top Security Risks and Recommendations

Despite the many benefits that cloud computing can potentially offer to an organization, on the other hand, there are unique attributes in cloud computing that require special risk assessment in areas such as data integrity, recovery, and privacy. In this section, a few key areas of concerns raised by different sources will be mentioned.[xvii][xviii][xix] This section should serve as a starting guide to some security concerns that organizations may have with cloud computing and these factors should always be considered prior to starting a relationship with cloud computing.

- ***Privileged user access:*** Sensitive data processed outside of the internal networks of the organizations bring with it an inherent risk because cloud services bypass the physical control that organizations are able to have otherwise. We do not know who have direct access to your data. It could be high executives of the cloud providers or system administrators or even employees in-charge of hardware. These roles in cloud architectures are unavoidable and they present possible risks. Carelessness and malicious insiders could lead to data exposure from the inside.
- ***Data protection:*** Data protection refers to the physical protection of data through access methods and encryption. This is an area of security risk because many cloud providers are still using the classic authentication method which is the username/password model which is a weak secure model. Moreover, it does not provide any level of granularity where different levels of access are given to different people.
- ***Data isolation:*** This risk category covers the failure of mechanisms separating storage, memory or other resources between different clients. However, attacks on resource isolation mechanisms are still less prominent and in essence very difficult for an attacker to put in practice compared to attacks on traditional OSs.
- **Data sanitization**: Sanitization refers to the proper removal of data from a device once it is not used. There can be two scenarios that can be problematic in this area.

  The first scenario is when the hardware is removed due to failure. When the commodity hardware fails and it is thrown away, your data risk being retrieved from the unwanted hardware if it is not properly sanitized before it is thrown away.

  The second scenario is when the clients terminate their service with a particular cloud provider. Residual data that is left with the cloud are often not deleted and kept on purpose. This can pose a problem because data isolation mechanisms are not fool-proof and this could lead to other clients accessing that block of data by accident after your termination.

- ***Data location:*** Because of the distributed nature of the cloud, you probably won't know exactly where its data is hosted at. You might not even know what country it will be stored in. You need to ask providers if they will commit to storing and processing data in specific location and consequently jurisdictions, and whether it is in your contract to allow them to obey local privacy requirements on behalf of their clients. There is a need to go through this process because different countries may have vastly different jurisdiction and procedures that could complicate things a little when security incidents arise.

- ***Data loss and recovery:*** A cloud provider should be transparent about what will happen to the data and service in case of data loss. While the chances of a total loss are low because of the high levels of redundancy, things can go wrong. There is a possibility of multiple servers losing your data at the same time. Therefore, it is crucial to watch out for agreements with a cloud service provider to see if they are clear about their data restoration processes, how long it would take, and how much it would cost for example.
- ***Investigative support:*** In the previous chapter, we have introduced some benefits to IT forensics when using cloud computing but investigating inappropriate or illegal activity may be difficult due to some implementation details of cloud service providers as well. If you cannot come to an agreement with the providers to support you in specific forms of investigation, then the only safe assumption is that investigation and discovery requests will be highly difficult if not impossible. It is imperative for any organization to read the fine-prints about investigative support before starting the service.
- ***Incidence response:*** Even though a cloud provider may be willing to help with investigative procedures, they might not be capable of doing so in a quick and effective way. The complexity of the cloud can often obscure this procedure. For example, it reportedly took one IaaS provider approximately eight hours to recognize and begin taking action on a Denial of Service attack. Understanding and negotiating the procedures for incident response should be done before entering a contract with any cloud providers.

5.3 Defense Recommendations

The security risks and challenges that cloud computing presents to IT professionals are formidable and many, especially for public clouds whose infrastructure and computational resources are owned by an outside party that sells those services to the general public. Nonetheless, certain measures can be taken and below we have taken some of the guidelines that NIST (National Institute of Standards and Technology) of the United States has recommended [xx] for federal agencies and departments whose requirement for security is as stringent as large organizations. We have also added some things that may be relevant to the general public.

- Extend organizational practices to include policies, procedures, and standards specific to the cloud. These could include the design, implementation, testing, and monitoring of deployed or engaged services. Audit mechanisms and tools should be implemented to ensure organizational practices are followed throughout the system lifecycle.
- Understand the various types of laws and regulations that impose security and privacy obligations on the organization particularly those involving data location, privacy and security controls. There is a need to also review and assess cloud provider's offerings with respect to the organizational requirements to be met and ensure that terms adequately meet the requirements set.
- Encourage contract transparency that allows visibility into the security and privacy controls and processes employed by cloud providers and their performance continuously evolving and shifting risk landscape.
- Understand cloud security such as access controls, encryption methods and data handling procedures that the cloud provider employs and assess the risks involved and most importantly avoid placing sensitive data in the cloud.
- Understand and negotiate (if possible) the contract provisions and procedures for incident response required by the organization to find the fastest way to resolve an incident.

References

[i] Cloud Computing News Topics http://www.techeye.net/topic/cloud-computing

[ii] The Cloud at Your Service, Page 3 – 6 (Manning Publications).

[iii] Web Computing Resource http://www.flukkytom.com/

[iv] Rackspace http://www.rackspace.com/

[v] Google App Engine https://appengine.google.com/

[vi] Amazon Machine Images http://aws.amazon.com/amis?_encoding=UTF8&jiveRedirect=1

[vii] Amazon EC2 Elastic IP Addresses
http://aws.amazon.com/articles/1346?_encoding=UTF8&queryArg=searchQuery&x=0&fromSearch=1&y=0&searchPath=all&searchQuery=elastic%20ip

[viii] Anatomy of a cloud storage infrastructure by M. Tim Jones
http://public.dhe.ibm.com/software/dw/cloud/library/cl-cloudstorage-pdf.pdf

[ix] Cloud Storage http://www.ibm.com/developerworks/cloud/library/cl-cloudstorage/figure1.gif

[x] NOSQL Databases http://nosql-database.org/

[xi] Cloud Database Benchmarking https://sites.google.com/site/cs4217jan2011team5/database

[xii] A Primer on Database Clustering Architectures
http://www.scaledb.com/pdfs/ArchitecturePrimer.pdf

[xiii] Leslie Lamport (2001). Paxos Made Simple.

[xiv] Amazon Web Services Blog
http://aws.typepad.com/aws/2010/10/amazon-rds-announcing-read-replicas.html

[xv] Assessing the Security Benefits of Cloud Computing by Craig Balding
http://cloudsecurity.org/tags/forensics.html

[xvi] Benefits, Risks and Recommendations for Information Security by ENISA 09/2009
http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment/at_download/fullReport

[xvii] Article on Gartner Report: 7 Cloud Computing Security Risks
http://www.cio.com/article/423713/Gartner_Seven_Cloud_Computing_Security_Risks

[xviii] Centre For the Protection Of National Infrastructure (CPNI) Info Security Briefing 01/2010
http://www.cpni.gov.uk/Documents/Publications/2010/2010007-ISB_cloud_computing.pdf

[xix] Cloud Security and Privacy : Data Security and Storage
http://mscerts.net/programming/cloud%20security%20and%20privacy%20%20%20data%20security%20and%20storage.aspx

[xx] NIST's Guidelines on Security and Privacy in Public Cloud Computing (SP 800-144)
http://csrc.nist.gov/publications/drafts/800-144/Draft-SP-800-144_clou§d-computing.pdf