Chapter 13

Cloud Types and Services

Hai Jin[1], Shadi Ibrahim[1], Tim Bell[2], Wei Gao[1], Dachuan Huang[1] and Song Wu[1]
[1]Services Computing Technology and System Lab,
Cluster and Grid Computing Lab
Huazhong University of Science and Technology, Wuhan, CHINA
[2]Department of Computer Science and Software Engineering,
University of Canterbury, Christchurch, NEW ZEALAND

1. Introduction

The increasing popularity of Internet services such as the Amazon Web Services, Google App Engine and Microsoft Azure have drawn a lot of attention to the Cloud Computing paradigm. Although the term "Cloud Computing" is new, the technology is an extension of the remarkable achievements of grid, virtualization, Web 2.0 and Service Oriented Architecture (SOA) technologies, and the convergence of these technologies. Moreover, interest in Cloud Computing has been motivated by many factors such as the prevalence of multi-core processors and the low cost of system hardware, as well as the increasing cost of the energy needed to operate them. As a result, Cloud Computing, in just three years, has risen to the top of the IT revolutionary technologies as shown in Figure 1, and has been announced as the top technology to watch in the year 2010 [1].

The name "Cloud Computing" is a metaphor for the Internet. A Cloud shape is used to represent the Internet in network diagrams to hide the flexible topology and to abstract the underlying infrastructure. Cloud Computing uses the internet to deliver different computing services including hardware, programming environments and software while keeping users unaware of the underlying infrastructure and location.

Despite the popularity and interest in cloud computing, a lot of confusion remains about what it is, and there is no formal definition of Cloud Computing. Two of the main definitions that are being used by the Cloud community have been provided by Ian Foster and Jeff Kapalan. Ian Foster gives a detailed definition of the term Cloud Computing: "A large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet".[2]

Jeff Kaplan views cloud computing as "a broad array of web-based services aimed at allowing users to obtain a wide range of functional capabilities on a 'pay-as-

you-go' basis that previously required tremendous hardware/software investments and professional skills to acquire. Cloud computing is the realization of the earlier ideals of utility computing without the technical complexities or complicated deployment worries." [3]

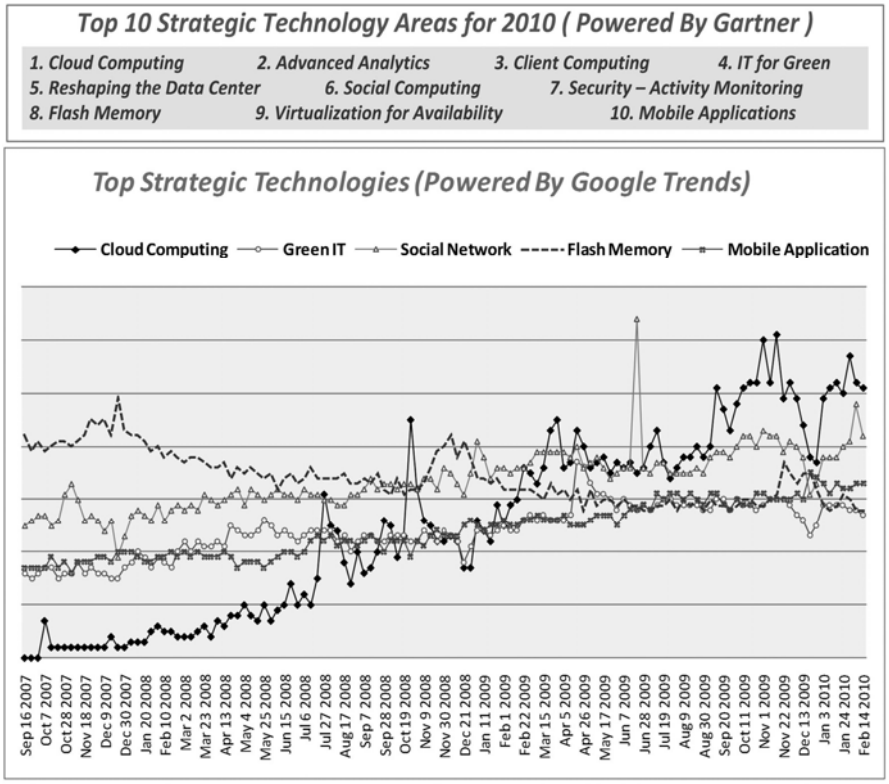

Figure 1. Top 10 strategic technology areas for 2010 and their evolution for the last three Years (based on Gartner [1] and Google Trends [4])

Cloud Computing enables users to access various computing resources simply, including computing cycles, storage space, programming environments and software applications (all you need is a web browser). Moreover, Cloud computing promises to provide other benefits:

- *Less investment*. Clouds provide affordable solutions that handle peaks, or scale easily at a fraction of the traditional costs of space, time and financial investment.
- *Scale*. Cloud vendors have vast data centers full of tens of thousands of server computers, offering computing power and storage of a magnitude never before available - cloud computing promises virtually unlimited resources.

- *Manageability*. The user experience is simplified as no configuration or back-up is needed

However, Cloud Computing also raises many concerns, mainly about security, privacy, compliance and reliability. When users move their data to the service provider data center, there is no guarantee that nobody else has access to this data. If the data is being stored in a different country, there can also be issues about jurisdictions for legal rights, and control of the data. Moreover, to date, there are no clearly defined Service Level Agreements (SLA) offered by the cloud providers.

There has been relatively little unification of the Cloud Computing concept. Consequently, it is useful to take a step back, consider the variety of Clouds offered by leading vendors, and describe them in a unified way, putting the different use and types of clouds in perspective; this is the main purpose of this book chapter.

The rest of this chapter is organized as follows: In section 2 we present the different types of Clouds. Then we briefly introduce the three main Cloud service categories. We then describe the variety of Clouds in leading projects in each category, IaaS, PaaS and SaaS, in sections 4, 5, and 6 respectively. To complete our survey, we present the Amazon cloud family and the different enterprises that are using the Amazon infrastructure, in section 7. A conclusion is provided in Section 8.

2. Cloud Types

Clouds can be classified in terms of who owns and manages the cloud; a common distinction is Public Clouds, Private Clouds, Hybrid Clouds and Community Clouds (see Figure 2).

2.1 Public Cloud

A public cloud, or external cloud, is the most common form of cloud computing, in which services are made available to the general public in a pay-as-you-go manner. Customers - individual users or enterprises - access these services over the internet from a third-party provider who may share computing resources with many customers. The public cloud model is widely accepted and adopted by many enterprises because the leading public cloud vendors as Amazon, Microsoft and Google, have equipped their infrastructure with a vast amount of data centers, enabling users to freely scale and shrink their rented resources with low cost and little management burden. Security and data governance are the main concern with this approach.

2.2 Private Cloud

A Private Cloud, or internal cloud, is used when the cloud infrastructure, proprietary network or data center, is operated solely for a business or organization, and serves customers within the business fire-wall. Most of the private clouds are large

company or government departments who prefer to keep their data in a more controlled and secure environment. Table 1 presents a comparison between public and private clouds.

Table 1. Public *vs.* Private Cloud

	Public Cloud	**Private Cloud**
Infrastructure Owner	Third party (Cloud provider)	Enterprise
Scalability	Unlimited and On-Demand	Limited to the installed Infrastructure
Control and Management	Only manipulate the virtual machines, resulting in less management burden	High level of control over the resources, and need more expertise to mange them.
Cost	Lower cost	High cost including: space, cooling, energy consumption and hardware cost
Performance	Unpredictable multi-tenant environment makes it hard to achieve guaranteed performance	Guaranteed performance
Security	Concerns regarding data privacy	Highly secure

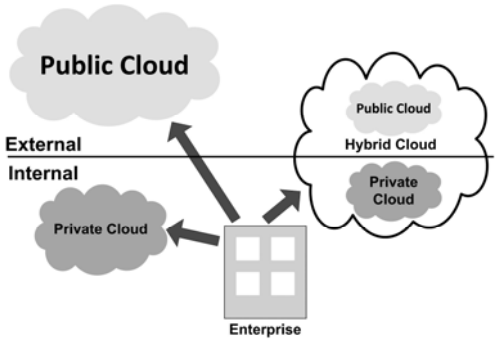

Figure 2. Cloud Types: Public, Private and Hybrid Clouds

2.3 Hybrid Cloud

A composition of the two types (private and public) is called a Hybrid Cloud, where a private cloud is able to maintain high services availability by scaling up their system with externally provisioned resources from a public cloud when there are rapid workload fluctuations or hardware failures. In the Hybrid cloud, an en-

terprise can keep their critical data and applications within their firewall, while hosting the less critical ones on a public cloud.

2.4 Community Cloud:

The idea of a Community Cloud is derived from the Grid Computing and Volunteer Computing paradigms. In a community cloud, several enterprises with similar requirement can share their infrastructures, thus increasing their scale while sharing the cost [5]. Another form of community cloud may be established by creating a virtual data center from virtual machines instances deployed on underutilized users machines [6].

3. Cloud Services and Cloud Roles

A Cloud is essentially a class of systems that deliver IT resources to remote users as a service. The resources encompass hardware, programming environments and applications. The services provided through cloud systems can be classified into Infrastructure as a service (IaaS), Platform as a Service (PaaS) and Software as a service (SaaS).

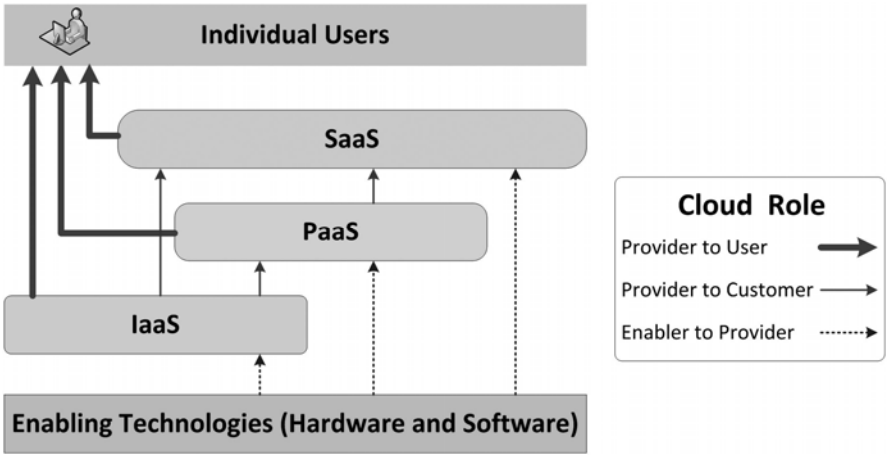

Figure 3. Cloud Services and Cloud Roles

Different enterprises play different roles in building and using cloud systems (Figure 3). These roles range from cloud technology enablers (enabling the underlying technologies used to build the cloud, such as hardware technologies, Virtualization technology, web services and so on), to cloud providers (delivering their infrastructure and platform to customers), to cloud customers (using the providers' services to improve their web applications), and users (who use the web applications, possibly unaware that it is being delivered using cloud technologies).

4. Infrastructure as a Service

Infrastructure as a Service (IaaS) is one of the "Everything as a Service" trends. IaaS is easier to understand if we refer it as Hardware as a Service (i.e. instead of constructing our own server farms, a small firm could consider paying to use infrastructure provided by professional enterprises). Companies such as Google, Microsoft and IBM are involved in offering such services. Large-scale computer hardware and high computer network connectivity are essential components of an effective IaaS.

The IaaS is categorized into: (1) Computation as a Service (CaaS), in which virtual machine based servers are rented and charged per hour based on the virtual machine capacity – mainly CPU and RAM size, features of the virtual machine, OS and deployed software; and (2) Data as a Service (DaaS), in which unlimited storage space is used to store the user's data regardless of its type, charged per GByte for data size and data transfer.

In this section we will describe some popular IaaS systems such as Amazon EC2 [7], GoGrid [8], Amazon S3 [9] and Rackspace[10]. We then compare three widely used CaaS systems (Table 2).

4.1 Amazon Elastic Compute Cloud (EC2)

Amazon has provided a popular universal and comprehensive solution to Cloud Computing, called the Amazon Elastic Compute Cloud (EC2) [7]. This solution was released as a limited public beta on August 25, 2006, but grew rapidly in the following years. After Amazon added many important and powerful features to EC2, it dropped the beta label on October 23, 2008. Today EC2 provides complete control over a customer's computing resources, so new server instances can be set up and booted in minutes, and their capacity can be scaled quickly through a simple web service interface.

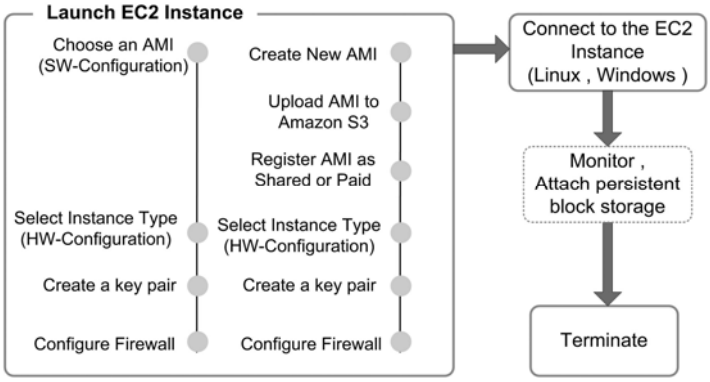

Figure 4. Lifecycle of Amazon Machine Image

EC2 provides many useful features for customers, including a mature and inexpensive billing system able to charge for computing at a very fine-grained level (memory usage, CPU usage, data transfer, etc.), deployment between multiple locations, elastic IP addresses, connection to a customer's existing infrastructure through a Virtual Private Network, monitoring services by Amazon CloudWatch, and elastic load balancing. EC2 has deployed such fine granularity and precision that it has become a benchmark and model in cloud computing.

Amazon's EC2 provides virtual machine based computation environments. It uses the Xen hypervisor [11] to manage their Amazon Machine Image (AMI) instance. AMI [7] is "an encrypted machine image that contains all information necessary to boot instances of your software". Using simple web service interfaces, users can launch, run, monitor and terminate their instances as shown in figure 4. Moreover they can, on the fly, add any of the abovementioned features to their configuration as they desire.

4.2 GoGrid

GoGrid [8] shares many common characteristics with Amazon in the classic cloud computing areas: supporting multiple operating systems through its own image management, and supporting load balancing, cloud storage, and so on. In addition, GoGrid provides customers with a user-friendly web service interface, easy-to-understand video demonstrations, and a strict but inexpensive billing system. Thus both EC2 and GoGrid provide basic and common features of cloud computing. The difference between the services they provide mainly derives from their respective business models. For example, GoGrid provides free cloud and persistent storage, slightly different from Amazon.

GoGrid also provides Hybrid Hosting, which is a distinguishing feature. Many applications simply don't run well in a pure multi-tenant server environment. Databases perform better on a dedicated server where they don't have to compete for input/output resources, and the situation is similar with web server applications. GoGrid provides these special applications with dedicated servers that also have high security assurance.

4.3 Amazon Simple Storage Service (S3)

The Amazon Simple Storage Service [9] (S3) is an online storage web service offered by Amazon Web Services. S3 is accessible to users through web services, REST-style HTTP interfaces[1], or by involving a SOAP interface. Like other cloud computing services, users can request small or large amounts of storage on the fly, providing a highly scalable storage system.

[1] REST stands for "Representational state transfer", a software architecture for distributed hypermedia.

Amazon S3 organizes the storage space into many “buckets”, with each bucket being given a global unique namespace to help locate data addresses, identify the user account for payments, and gathering usage information [12]. S3 deals with all type of data as objects and stores them with their metadata into the bucket chosen by the data owner. An object can be accessed through a URL composed of its key and version ID with its bucket namespace as the prefix.

Amazon S3’s users are spread across countless fields, for example, SmugMug [13], Slideshare [14] and Twitter [15] use Amazon S3 to host images, Apache Hadoop [16] uses S3 to store computation data, and online synchronization utilities such as Dropbox [17] and Ubuntu One [18] use S3 as their storage and transfer facility.

4.4 Rackspace Cloud

Rackspace [10] Cloud was originally launched on March 4, 2006 under the name “Mosso”. In the following three years, it has changed his name from “Mosso LLC” to “Mosso: The Hosting Cloud”, and then finally “Rackspace Cloud” on June 17, 2009. This company provides services including a cloud server, cloud files, and cloud site.

The cloud files service is a cloud storage service providing unlimited online storage and a Content Delivery Network (CDN) for media on a utility computing basis. In addition to the online control panel, this company provides an API service that can be accessed over a RESTful API with open source client code. Rackspace solves the security problem by replicating three full copies of data across multiple computers in multiple zones, with every action protected by SSL.

Table 2. A comparison of three widely used CaaS (prices from Feb 2010)

CaaS	**Amazon EC2**	**GoGrid**	**Rackspace (Cloud Server)**
Virtualization	Xen	Xen	VMware
OS support	Linux , Windows	Linux , Windows	Linux , Windows
Server RAM	1.7 GB and going up to 68.4 GB	0.5 GB and going up to 8GB	256 MB and going up to 16 GB
Load Balancer	Amazon Elastic Load Balancer	Free F5 Load Balancer	No
Persistent Block Storage	Yes	Yes	No
Hybrid Hosting	No	Yes	Yes
24/7 Support	No	Yes	Yes
Pricing	Billed $0.085 – $3.18 per hour (vary for different	Billed $0.19 per GB of deployed RAM per hour	Billed $0.06 per GB of deployed RAM per Hour

	Instance and Regions). The Data Transfer rates vary based on where the data goes out to and comes in from with pricing between $0.00 to $0.15 per GB transferred.	and 60GB of disk, $0.50 per GB of outbound data transferred, and all inbound data transfer is free.	and 40GB of disk, $0.05 per GB of inbound data transfer and $0.22 per GB of outbound data transfer.

5. Platform as a Service

Platform as a Service (PaaS) cloud systems provide a software execution environment that application services can run on. The environment is not just a pre-installed operating system but is also integrated with a programming-language-level platform, which users can be used to develop and build applications for the platform. From the point of view of PaaS clouds' users, computing resources are encapsulated into independent containers, they can develop their own applications with certain program languages, and APIs are supported by the container without having to take care of the resource management or allocation problems such as automatic scaling and load balancing. In this section we introduce three typical PaaS : Google App Engine [19], Microsoft Azure [20], and Force.com [21], and then we compare them in Table 3.

5.1 Google App Engine

Google App Engine (GAE)'s main goal is to efficiently run users' web applications. As shown in Figure 5, it maintains Python and Java runtime environments on application servers, along with some simple APIs to access Google services.

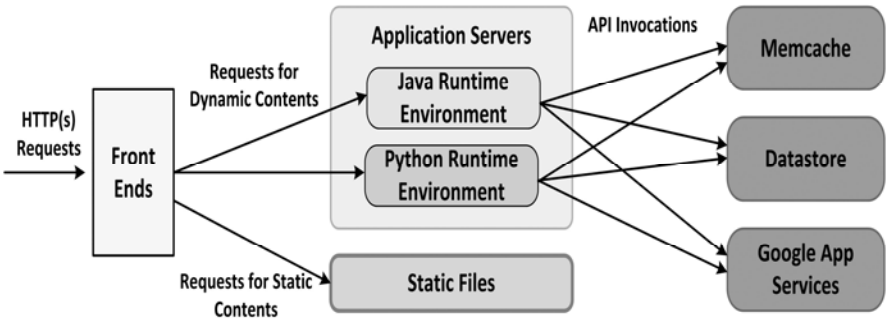

Figure 5. The Architecture of Google App Engine

The front ends spread HTTP requests with load balancing and routing strategies based on the contents. Runtime systems running on application servers deal with the logic processing of applications and provide dynamic web content, while static pages are served by the shared Google infrastructure [22]. To decouple the persistent data from application servers, GAE puts them into the Datastore instead of a local file system. Applications can integrate data services and other Google App Services, such as email, image storage and so on through APIs provided by the GAE.

In addition to the services, Google also provides some tools for developers to help them build web applications easily on GAE. However, since they are tightly connected to the Google infrastructure, there are some restrictions that limit the functionality and portability of the applications.

5.2 Microsoft Azure

Microsoft's cloud strategy is to construct a cloud platform that users can move their applications to in a seamless way, and ensure its managed resources are accessible to both cloud services and on-premises applications. To achieve this, Microsoft introduced the Windows Azure Platform (WAP), which is composed of a cloud operating system named Windows Azure, and a set of supporting services, as shown in Figure 6.

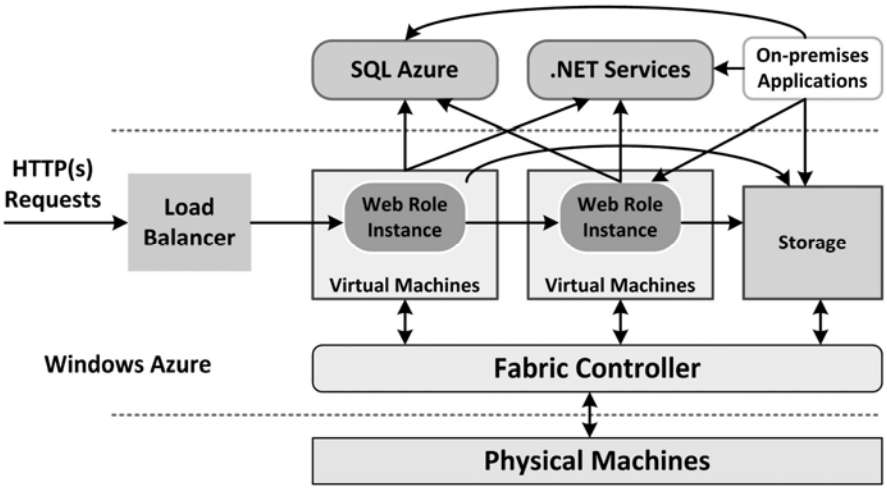

Figure 6. The Architecture of the Windows Azure Platform

Windows Azure is the main part of the WAP. It employs virtual machines as its runtime environments. The applications in Microsoft's cloud offerings are divided into two types: Web role instances, which can serve web requests via the internet information services (IIS); and Worker role instances, which can only receive

messages from other Web role instances or on-premises applications [23]. Windows Azure employs a "fabric controller" to manage all virtual machines and storage servers on the physical machines in a Microsoft data center [24]. Similar to the GAE's datastore, WAP also provides a database service called SQL Azure, to store data in the cloud. One feature of SQL Azure is that it provides a tool for data synchronization across on-premises and/or off-premises databases. Infrastructure services supported by WAP through .NET services currently include access control and exposing services. Both are available for cloud and on-premises applications.

5.3 Force.com

Force.com is an enterprise cloud computing platform offered by Salesforce. It helps service venders develop and deliver stable, secure and scalable applications. Two key enabling technologies of Force.com are multi-tenancy and metadata [25]. The multi-tenancy approach allows different users to share application templates on a public physical computing resource pool, while the application instances are independent from each other. For customized applications, a metadata driven architecture that generates application components according to its own description has been proposed [26]. Other technologies and services of the Force.com platform include service delivery infrastructure, a database, logic services, user interfaces, and developer tools [25].

The idea of the Force.com cloud solution is that it should take care of all common underlying requirements so that users need only focus on the design of their applications. One potential problem of this is that the applications rely heavily on the infrastructure and services of Force.com, which compromises their portability.

Table 3. Survey of four PaaS provides

PaaS provider	**Programming environment**	**Infrastructure**	**Hosted Application**
Google	Python and Java	Google Data Center	Socialwok [27],Gigapan [28], LingoSpot[29]
Azure	.Net (Microsoft Visual Studio)	(Virtual Machine Based) Microsoft Data Centers	Microsoft Pinpoint[30]
Force.com	Apex Programming and Java1	Saleforce Data Center	EA [31],Author Solutions [32],The Wall Street Journal[33]
Heroku	Ruby	Amazon EC2 and S3	Übermind[35,

[34]	Kukori.ca[36], act.ly[37], Cardinal Blue [38]

6. Software as a Service

Software-as-a-Service (SaaS) is based on licensing software use on demand, which is already installed and running on a cloud platform. These on-demand applications may have been developed and deployed on the PaaS or IaaS layer of a cloud platform. SaaS replaces traditional software usage with a Subscribe/Rent model, reducing the user's physical equipment deployment and management costs. The SaaS clouds may also allow users to compose existing services to meet their requirements. This section presents some SaaS clouds and applications.

6.1 Desktop as a Service

Desktop as a Service is a special variant of Software as a Service that provides a virtualized desktop-like personal workspace, and sends its image to the user's real desktop. Instead of a local desktop, the user can access their own desktop-on-the-cloud from different places for convenience, and receive the benefit of SaaS at same time.

The "Global Hosted Operating SysTem" (G.ho.st) [39] is a free and complete Internet-based Virtual Computer (VC) service suite including a personal desktop, files and applications [40]. It offers users an operating system image simulated in a web browser using Flash and Javascript, which can be accessed through any browser. The G.ho.st application services are hosted by the Amazon Web Services (AWS) platform, so users can utilize EC2 and S3 resources through their G.ho.st desktops. One limitation of G.ho.st is, as a lightweight desktop service, it only supports on-line applications, and users cannot run legacy programs on it.

Unlike G.ho.st's browser-based desktop image, the Desktone Virtual-D Platform [41] implements a desktop as a service by encapsulating a virtual machine based desktop, called Virtual Desktop Infrastructure (VDI), into a service. The advantage of VDI is that it can offer the same environment as a native operating system, and allows users to install their own software. The Desktone Virtual-D Platform integrates all desktop virtualization layers and simplifies desktop management, improving security and compliance [42]. This solution delivers desktops as a cost-effective subscription service deployed on cloud.

6.2 Google Apps

Google Apps [43] is a typical SaaS implementation. It provides several Web applications with similar functionality to traditional office software (word processing, spreadsheets etc.), but also enables users to communicate, create and collaborate easily and efficiently. Since all the applications are kept online and are accessed

through a web browser, users can access their accounts from any internet-connected computer, and there is no need to install anything extra locally.

Google Apps has several components. The communication components consist of Google mail and Google Talk, which allow for communication through email, instant messaging and voice calls. The office components include docs and spreadsheets, through which users can create online documents that also facilitate searching and collaboration. Google Calendar is a flexible calendar application for organizing meetings and events. With Google's "Web Pages", administrators can easily publish web pages, while "Start Pages" provide users with a rich array of content and applications that can be personalized.

Google Apps has several significant features. First, it provides an easy-to-use control panel which facilitates the most common administration tasks such as enabling/disabling applications, managing accounts, and customizing interfaces. Second, although hosted on Google, the user can control the branding on all interfaces - email addresses will have only the user's domain name with no mention of Google in the message body [44], and users can customize their web interfaces, layouts and colors on web and start pages. Third, administrators can integrate with existing platforms as well as extend the functionality of the core Google Apps applications with the Application Programming Interfaces (APIs) that are offered. There are APIs available for provisioning, reporting, and migration, as well as manipulating data in Calendar and Spreadsheets, and integrating with Single Sign On (SSO) systems [45].

6.3 Salesforce

Salesforce [46] is a business SaaS cloud platform that provides customizable applications, mostly Customer Relationship Management (CRM) services, to consumers. There are two major products presented by Salesforce. Sales Cloud is a group of comprehensive applications to improve the convenience and efficiency of business activities; and Service Cloud is provided to integrate social network applications like Facebook and Twitter, to construct a users' customer service community.

Saleforce CRM services are deployed on the Force.com cloud platform, which operates a multi-tenancy oriented metadata-driven architecture [26]. Multi-tenancy enables sharing the same version of an application among many users, but each user can only access their own private data, which keeps their activities isolated. All applications' functionalities and configurations are described with metadata, so users can customize applications as they want.

Although the shared application model could cause interference between users, the Salesforce SaaS cloud has these advantages: (1) service providers can develop only one version of application, and don't need to worry about heterogeneous execution environments; (2) the sharing of the physical computing resource, operating

system and runtime environment lowers the cost of the application service; and (3) service consumers are free to choose their preferred version of the application and customize it to fit their business.

6.4 Other software as service examples

As cloud computing technology spreads, more and more Software as a Service implementations have been released. Table 4 gives some other SaaS examples. The services cover many fields in addition to personal file processing and business administration.

Table 4. Some SaaS examples.

SaaS Provider	Important Services
A2Zapps.com [47]	Marketing Automation, School Automation (ERP)
Envysion.com [48]	Video Management
Learn.com [49]	Training, HR, Online Courses
Microsoft [50]	Office Live Meeting, Dynamics CRM, SharePoint
OpenID [51]	Log in Identification
Zoho [52]	Mail, Docs, Wiki, CRM, Meeting, Business

7. The Amazon Family

Since the early stage of Cloud Computing, Amazon has dominated the Cloud market by providing scalable on-demand infrastructure, in particular EC2 and S3, making it easy for enterprises to obtain computing power and storage as a service. [53].

One of the first success stories about the effectiveness of cloud computing for providing low-cost and fast solutions on-demand for an enterprise was the case of the New York Times. In order to make their articles from 1851- 1922 available to the public, they were able to create PDF versions of their archives using 100 EC2 instances. This was a large job – in some cases they needed to take several TIFF images and scale and glue them to create one PDF file in less than 24 hours. Once the archive was created, they stored it in the S3, using 4 TB of storage [54].

Recently Amazon has equiped their IT infrastructure services with new services as shown in Table 5, motivating many businesses, enterprises and academia to join the Amazon Web Services using the always-improving IT infrastructure services suite to build their business and applications (Figure 7). Furthermore, some enterprises use the Amazon cloud services to provide new cloud services, including RightScale [55] providing IaaS, Heroku [34] providing PaaS, Animoto [56] and G.ho.st[39] providing SaaS.

Table 5. The Amazon Web services (prices from Feb 2010)

Amazon Web Service	**Brief description**	**Geographical Regions[1]**	**Pricing Range**
Amazon Elastic Compute Cloud (Amazon EC2)	"Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud."[7]	US – N. Virginia US – N. California EU – Ireland	$0.085 – $3.18 per hour (vary for different Instance and regions)
Amazon Simple Storage Service (Amazon S3)	"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."[9]	US – N. Virginia US – N. California EU – Ireland	0.055 – 0.150 Per GB (vary for different regions)
Amazon SimpleDB [57]	"Amazon SimpleDB is automatically indexing your data and providing a simple API for storage and access and requiring no schema."[57]	US – N. Virginia US – N. California EU – Ireland	$0.140- $0.154 per Hour (vary to different regions)
Amazon CloudFront [58]	"Amazon CloudFront is a web service for content delivery. It delivers the static and streaming content using a global network of edge locations. By routing the requests for any objects to the nearest edge location, so content is delivered with the best possible performance. Amazon CloudFront works seamlessly with (Amazon S3)." [58]	United States Europe Hong Kong Japan	$0.050 - $0.221 per GB (vary for different Data transfer per month and regions)
Amazon Simple Queue Service (Amazon SQS) [59]	"Amazon Simple Queue Service (Amazon SQS) offers a reliable, highly scalable, hosted queue for storing messages as they travel between computers. Amazon SQS makes it easy to build an automated workflow, working in close conjunction with the	US – N. Virginia US – N. California EU – Ireland	$0.01 per 10,000 Amazon SQS Requests ($0.000001 per Request)

[1] Amazon web services provide multiple regions and "availability zones" so customers can connect to the most convenient service, and also choose services in multiple zones to maximize failure-independence.

	Amazon EC2."[59]		
Amazon Elastic MapReduce [60]	"Amazon Elastic MapReduce is a web service that enables businesses, researchers, data analysts, and developers to easily and cost-effectively process vast amounts of data. It utilizes a hosted Hadoop framework running on the web-scale infrastructure of Amazon EC2 and Amazon S3."[60]	US – N. Virginia US – N. California EU – Ireland	$0.015 - $0.42 per hour (vary for different instance and regions)
Amazon Relational Database Service (Amazon RDS) [61]	"Amazon Relational Database Service (Amazon RDS) gives you access to the full capabilities of a familiar MySQL database. This means the code, applications, and tools you already use today with your existing MySQL databases work seamlessly with Amazon RDS." [61]	US Region	$0.11 - $3.10* (vary for different instances)

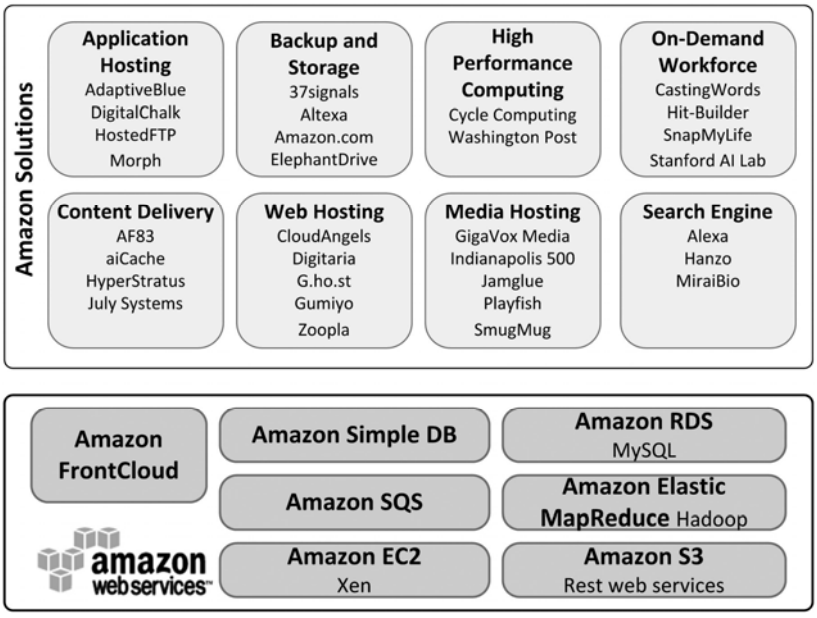

Figure 7. The Amazon Family: Amazon Web services and their Different Solutions and customers [62]

7.1 RightScale: IaaS based on AWS

Rightscale [55] is a web based solution for deploying and managing services on top of IaaS cloud providers such as Amazon and GoGrid. RightScale enables users to simply build, monitor and auto-scale their virtual infrastructure. It has been listed among the top 10 companies to watch in cloud computing [63] especially after its huge success in managing the Amazon EC2 services for companies such as Animoto [55] and G.ho.st [39]. RightScale also provides management tools for deployments across multiple clouds, providing flexibility in the choice of all kinds of services.

Recently, many enterprises have joined Rightscale, running different applications such as scalable websites (ShareThis [64]), grid applications (Animoto[55]), Test & Development, and Social Gaming Applications (PlayFish [65]).

7.2 HeroKu: platform as a service using Amazon web Service

Heroku [34] is a Ruby platform-as-a-service, which offers an in-browser Ruby on Rails multi-tenant development environment associated with cloud-based hosting services. Heroku's platform is entirely based on Amazon Web Services such as EC2 and S3. Thus, they can scale their infrastructure to satisfy their customers' demands at fraction of the traditional cost.

In Heroku, the users' code is compiled into self-contained, read only "slugs", which are then run inside a number of "dynos", depending on the application's need. Furthermore, to scale up an application, new dynos can be started in under two seconds for most apps. A Dyno is an independent process spread across multiple servers. Recently, Heroku has been used by many developers; at the time of writing it was hosting more than 45,000 applications, including websites and facebook applications.

7.3 Animoto Software as service using AWS

Animoto [55] is a web application that automatically generates fast, free (for video up to 30 seconds), and unique video pieces from users' photos, video clips and music. It is based around their own patent-pending technology and high-end motion design. Their system is built on top of Amazon Web Services, namely EC2, S3 and SQS. EC2 is used for web servers, application servers, upload servers, 'director' servers, and database servers. All the music and photos are stored and served by Amazon S3. Amazon's SQS is used to connect all the operations during the video creation process.

Earlier, Animoto were regularly using a 50 virtual machine instance of EC2, but after the huge success of their Facebook application, they scaled up to 3,500 instances using RightScale, within just three days (at its peak RightScale was launching and configuring 40 new instances per minute [66]).

7.4 SmugMug Software as service using AWS

SmugMug [13] is a photo sharing company that offers unlimited storage using the Amazon S3. In early 2006, SmugMug (with 15 employees and 1 programmer), moved its storage to S3, and became fully operational on Amazon S3 in one week, with around 100 Terabytes of customer photos (70,000,000 original images and six display copies of each). This saved them roughly $500,000 compared with increasing the space in their data center.

8. Conclusion

Cloud computing is a very flexible paradigm for delivering computational power. It will mean many things to many people. For some it means being able to set up a new start-up company knowing that initial resources will be inexpensive but a sudden increase in demand from users won't make the company a victim of its own success, as has happened in some cases in the past where servers have been unable to cope with demand, and the company loses clients as they become unhappy with poor response times. For other people, cloud computing means easier administration, with issues such as licensing, backup and security being taken care of elsewhere. In other cases, cloud computing means having a powerful computational environment available anywhere that the user can access a web browser.

With this flexibility, scalability and ease of maintenance, it is little wonder that cloud computing is being touted as a technology to watch. Of course, there are issues: privacy of data can be a concern, good internet connectivity is required, and some organizations may wish to maintain control over their own resources. However, these problems can usually be addressed, and using a cloud remains a very attractive way to set up a powerful system very quickly.

The various forms of service – infrastructure, platform, and software as a service – provide exciting ways to deliver new products that innovators might come up with. Already there are examples of widely used products and web sites that have sustained remarkable growth because creative ideas could be implemented quickly, and because the subsequent demand could be met easily through the flexibility of cloud computing.

The future seems to be limited only by the imaginations of innovators who can think of applications that will help people communicate, store and process vast quantities of information, whether it is millions of individuals with small collections of personal information, or a single large organization with large collections of data to be processed.

Acknowledgement

This work is supported by National 973 Key Basic Research Program under Grant No. 2007CB310900, NSFC under Grant No.60673174, Program for New Century

Excellent Talents in University under Grant NCET-07-0334, Information Technology Foundation of MOE and Intel under Grant MOE-INTEL-09-03NSFC, and National High-Tech R&D Plan of China under Grant 2006AA01A115.

References

1. Gartner –Gartner Newsroom, http://www.gartner.com/it/page.jsp?id=1210613 , 2010.
2. I. Foster, Y. Zhao, I. Raicu and S. Y. Lu, "Cloud Computing and Grid Computing 360-Degree Compared" Proceedings of the Grid Computing Environments Workshop (GCE '08), 2008.
3. "Twenty Experts Define Cloud Computing", SYS-CON Media Inc, http://cloudcomputing.sys-con.com/read/612375_p.htm, 2008.
4. Google Trends, http://www.google.com/trends, feb 23rd 2010.
5. Wikipedia - Cloud Computing, http://en.wikipedia.org/wiki/Cloud_computing ,2010.
6. G. Briscoe, A. Marinos, "Towards Community Cloud Computing" , Proceedings of the IEEE Digital EcosystemS and Technologies (DEST 2009), 2009.
7. Amazon Elastic Compute Cloud (Amazon EC2), http://aws.amazon.com/ec2, 2010.
8. GoGrid homepage, http://www.gogrid.com/index.v2.php, 2010.
9. Amazon Simple Storage Service (Amazon S3), http://aws.amazon.com/s3, 2010.
10. The Rackspace Cloud , http://www.rackspacecloud.com/, 2010.
11. Xen Hyper visor, http://xen.org/, 2010.
12. Amazon Simple Storage Service Developer Guide, http://docs.amazonwebservices.com/AmazonS3/latest/index.html?Introduction.html, 2010.
13. Smugmug homepage, http://www.smugmug.com/, 2010.
14. Slideshare homepage, http://www.slideshare.net/ , 2010.
15. Twitter homepage, http://twitter.com/ , 2010.
16. Apache Hadoop, http://hadoop.apache.org/, 2010.
17. Dropbox homepage, http://www.dropbox.com/, 2010.
18. Ubunto One, https://one.ubuntu.com/, 2010.
19. Google App Engine, http://code.google.com/appengine/, 2010.
20. Windows Azure platform, http://www.microsoft.com/windowsazure/, 2010.
21. Force.com homepage, www.salesforce.com/platform/, 2010.
22. L. Alon, "From Spark Plug to Drive Train: Life of an App Engine Request" Google I/O Developer Conference. http://code.google.com/events/io/2009/, 2009.
23. C. David, " Introducing the Windows Azure Platform", http://view.atdmt.com/action/mrtyou_FY10AzurewhitepapterIntroWindowsAzurePl_1, 2009.
24. C. David, "Introducing Windows Azure", http://view.atdmt.com/action/mrtyou_FY10AzureWhitepaperIntroWindowsAzureSec_1, 2009.

25. Force.com Whitepaper, “A Comprehensive Look at the Force.com Cloud Platform”, http://wiki.developerforce.com/index.php/A_Comprehensive_Look_at_the_Force.com_Cloud_Platform, 2009.
26. Force.com Whitepaper, “The Force.com Multitenant Architecture”, http://www.apexdevnet.com/media/ForcedotcomBookLibrary/Force.com_Multitenancy_WP_101508.pdf, 2009.
27. Socialwok homepage, http://www.socialwok.com/, 2010.
28. Gigapan homepage, http://www.gigapan.org/, 2010.
29. LingoSpot homepage,http://www.lingospot.com/, 2010.
30. Microsoft Pinpoint homepage, http://pinpoint.microsoft.com/en-US/, 2010.
31. EA Video Games - Electronic Arts (EA) homepage, http://www.ea.com/, 2010.
32. Author Solutions homepage,http://www.authorsolutions.com/, 2010.
33. The Wall Street Journal homepage, http://online.wsj.com, 2010.
34. Heroku, Ruby Cloud Platform as a Service, http://heroku.com/, 2010.
35. Übermind homepage, http://www.ubermind.com/, 2010.
36. Kukori.ca homepage, http://kukori.ca/, 2010.
37. act.ly homepage, http://act.ly/, 2010.
38. Cardinal Blue homepage, http://cardinalblue.com/, 2010.
39. G.ho.st Homepage, http://g.ho.st/, 2010.
40. G.ho.st Whitepaper, “The G.ho.st Virtual Computer”, http://g.ho.st/HomePageResources/PDF/GhostWhitePaper.pdf, 2009.
41. The Desktone Virtual-D Platform homepage, http://www.desktone.com/platform/index.php, 2010.
42. The Desktone Virtual-D Platform - Fact Sheet, http://www.desktone.com/downloads/public/collateral/Virtual-D-Fact-Sheet.pdf , 2009.
43. Google App, http://www.google.com/apps/intl/en/business/index.html, 2010.
44. Wikipedia, Google App, http://en.wikipedia.org/wiki/Google_Apps, 2010.
45. Google Apps Products Overview. http://services.google.com/apps/resources/overviews_breeze/Apps/index.html, 2010.
46. Salesforce homepage, http://www.salesforce.com/crm/, 2010.
47. A2Zapps.com homepage, http://www.a2zapps.com, 2010.
48. Envysion.com, http://www.envysion.com, 2010.
49. Learn.com, http://learn.com, 2010.
50. Microsoft Online Services, http://www.microsoft.com/online/products.mspx, 2010.
51. OpenID homepage, http://openid.net/, 2010.
52. Zoho homepage, http://www.zoho.com/, 2010.
53. The network world - 10 cloud computing companies to watch, http://www.networkworld.com/supp/2009/ndc3/051809-cloud-companies-to-watch.html

54. The New york Times Blog, -service, Prorated Super Computing Fun!, http://open.blogs.nytimes.com/2007/11/01/self-service-prorated-super-computing-fun/, 2010.
55. Cloud Computing Management Platform by RightScale, http://www.rightscale.com/, 2010.
56. Animoto homepage, http://animoto.com/, 2010.
57. Amazon SimpleDB, http://aws.amazon.com/simpledb/, 2010.
58. Amazon CloudFront, http://aws.amazon.com/cloudfront/, 2010.
59. Amazon Simple Queue Service (SQS), http://aws.amazon.com/sqs/, 2010.
60. Amazon Elastic MapReduce, http://aws.amazon.com/elasticmapreduce/, 2010.
61. Amazon Relational Database Service (RDS), http://aws.amazon.com/rds/, 2010.
62. Amazon AWS – AWS solutions, http://aws.amazon.com/solutions/aws-solutions/, 2010.
63. Appistry/CloudCamp Survey, "Inside the Cloud", www.appistry.com/go/inside-the-cloud, 2009.
64. ShareThis homepage, http://sharethis.com/, 2010.
65. PlayFish homepage, http://www.playfish.com/, 2010.
66. Rightscale Blog - Animoto's Facebook scale-up, http://blog.rightscale.com/2008/04/23/animoto-facebook-scale-up/, 2010.

Index terms (alphabetically):

Amazon CloudFront
Amazon Elastic Compute Cloud (Amazon EC2)
Amazon Elastic MapReduce
Amazon Machine Image (AMI)
Amazon Relational Database Service (Amazon RDS)
Amazon Simple Queue Service (Amazon SQS)
Amazon Simple Storage Service (Amazon S3)
Amazon SimpleDB
Amazon Web Service (AWS)
Cloud Computing
Community Clouds
Computation as a Service (CaaS)
Content Delivery Network (CDN)
Customer Relationship Management (CRM)
Data as a Service (DaaS)
Desktop as a Service
Global Hosted Operating SysTem (Gh.os.t)
Google App Engine (GAE)
Hybrid Clouds
Infrastructure as a service (IaaS)
Internet Information Services (IIS)

pay-as-you-go
Platform as a Service (PaaS)
Private Clouds
Public Clouds
Service Level Agreement (SLA)
Service Oriented Architecture (SOA)
Software as a service (SaaS)
Virtual Computer (VC)
Virtual Desktop Infrastructure (VDI)
Virtual Machine
Virtual Private Network (VPN)
Windows Azure Platform (WAP)