

Chapter 1

Cloud Computing Fundamentals

Borko Furht

1.1 Introduction

In the introductory chapter we define the concept of cloud computing and cloud services, and we introduce layers and types of cloud computing. We discuss the differences between cloud computing and cloud services. New technologies that enabled cloud computing are presented next. We also discuss cloud computing features, standards, and security issues. We introduce the key cloud computing platforms, their vendors, and their offerings. We discuss cloud computing challenges and the future of cloud computing.

Cloud computing can be defined as a new style of computing in which dynamically scalable and often virtualized resources are provided as a services over the Internet. Cloud computing has become a significant technology trend, and many experts expect that cloud computing will reshape information technology (IT) processes and the IT marketplace. With the cloud computing technology, users use a variety of devices, including PCs, laptops, smartphones, and PDAs to access programs, storage, and application-development platforms over the Internet, via services offered by cloud computing providers. Advantages of the cloud computing technology include cost savings, high availability, and easy scalability.

Figure 1.1, adapted from Voas and Zhang (2009), shows six phases of computing paradigms, from dummy terminals/mainframes, to PCs, networking computing, to grid and cloud computing.

In phase 1, many users shared powerful mainframes using dummy terminals. In phase 2, stand-alone PCs became powerful enough to meet the majority of users' needs. In phase 3, PCs, laptops, and servers were connected together through local networks to share resources and increase performance. In phase 4, local networks were connected to other local networks forming a global network such as the Internet to utilize remote applications and resources. In phase 5, grid computing provided shared computing power and storage through a distributed computing

B. Furht (✉)

Department of Computer & Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA

e-mail: bfurht@fau.edu

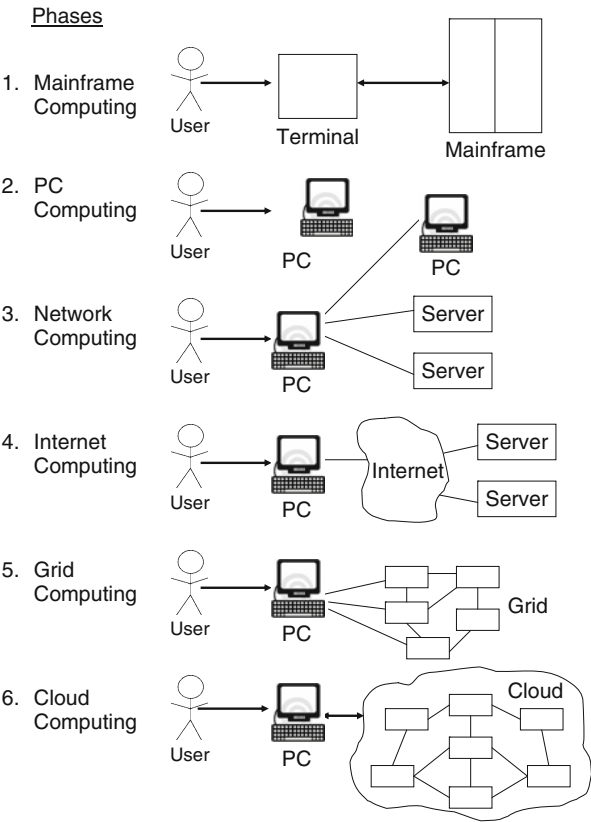


Fig. 1.1 Six computing paradigms – from mainframe computing to Internet computing, to grid computing and cloud computing (adapted from Voas and Zhang (2009))

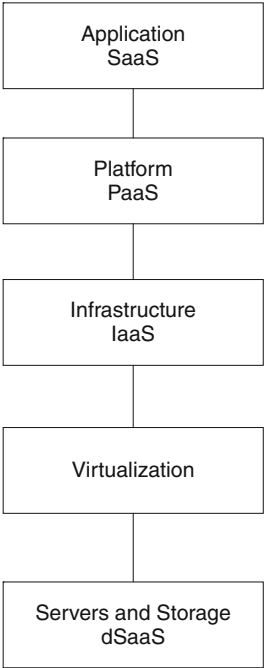
system. In phase 6, cloud computing further provides shared resources on the Internet in a scalable and simple way.

Comparing these six computing paradigms, it looks like that cloud computing is a return to the original mainframe computing paradigm. However, these two paradigms have several important differences. Mainframe computing offers finite computing power, while cloud computing provides almost infinite power and capacity. In addition, in mainframe computing dummy terminals acted as user interface devices, while in cloud computing powerful PCs can provide local computing power and caching support.

1.1.1 Layers of Cloud Computing

Cloud computing can be viewed as a collection of services, which can be presented as a layered cloud computing architecture, as shown in Fig. 1.2 [Jones XXXX]. The

Fig. 1.2 Layered architecture of Cloud Computing (adapted from Jones)



services offered through cloud computing usually include IT services referred as to SaaS (Software-as-a-Service), which is shown on top of the stack. SaaS allows users to run applications remotely from the cloud.

Infrastructure-as-a-service (IaaS) refers to computing resources as a service. This includes virtualized computers with guaranteed processing power and reserved bandwidth for storage and Internet access.

Platform-as-a-Service (PaaS) is similar to IaaS, but also includes operating systems and required services for a particular application. In other words, PaaS is IaaS with a custom software stack for the given application.

The data-Storage-as-a-Service (dSaaS) provides storage that the consumer is used including bandwidth requirements for the storage.

An example of Platform-as-aService (PaaS) cloud computing is shown in Fig. 1.3 [“Platform as a Service,” <http://www.zoho.com/creator/paas.html>]. The PaaS provides Integrated Development Environment (IDE) including data security, backup and recovery, application hosting, and scalable architecture.

According to Chappell (2008) there are three categories of cloud services, as illustrated in Fig. 1.4. Figure 1.4a shows the cloud service SaaS, where the entire application is running in the cloud. The client contains a simple browser to access the application. A well-known example of SaaS is salesforce.com.

Figure 1.4b illustrates another type of cloud services, where the application runs on the client; however it accesses useful functions and services provided in the cloud. An example of this type of cloud services on the desktop is Apple’s iTunes.

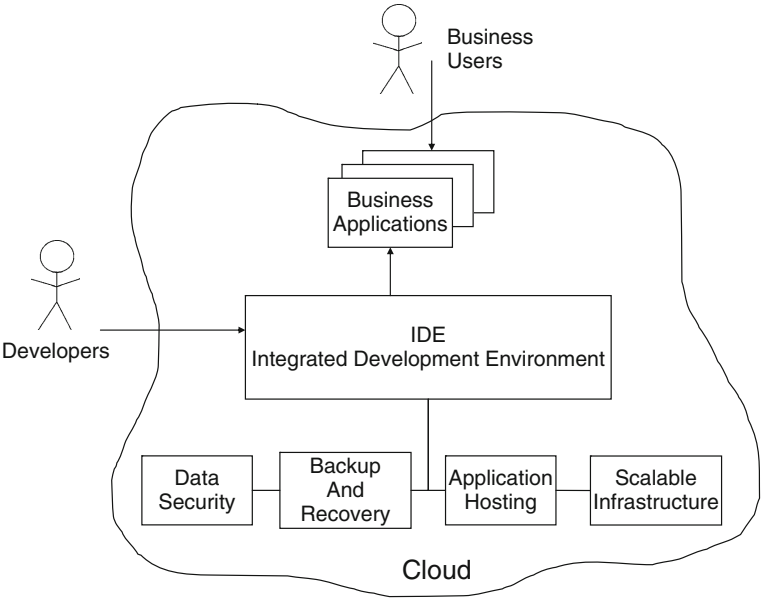


Fig. 1.3 The concept of Platform-as-a-Service, Zoho Creator (adapted from “Platform as a Service,” <http://www.zoho.com/creator/paas.html>)

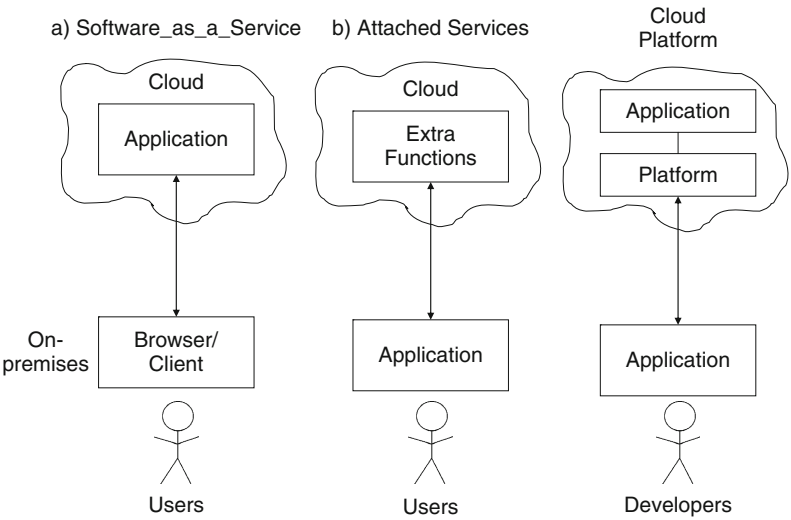


Fig. 1.4 The categories of cloud services (adapted from Chappell (2008))

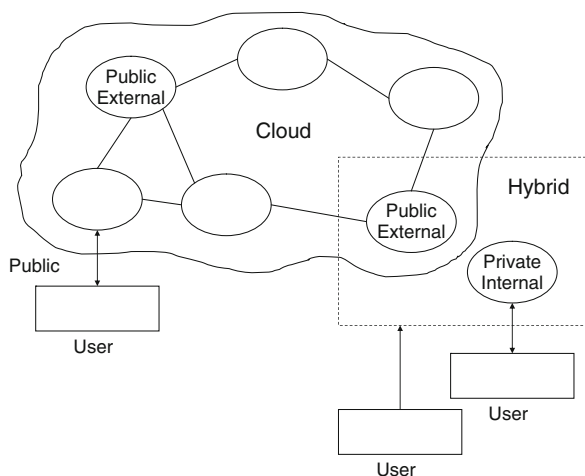
The desktop application plays music, while the cloud service is used to purchase a new audio and video content. An enterprise example of this cloud service is Microsoft Exchange Hosted Services. On-premises Exchange Server is using added services from the cloud including spam filtering, archiving, and other functions.

Finally, Fig. 1.4c shows a cloud platform for creating applications, which is used by developers. The application developers create a new SaaS application using the cloud platform.

1.1.2 Types of Cloud Computing

There are three types of cloud computing (“Cloud Computing,” Wikipedia, http://en.wikipedia.org/wiki/Cloud_computing): (a) public cloud, (b) private cloud, and (c) hybrid cloud, as illustrated in Fig. 1.5.

Fig. 1.5 Three types of cloud computing



In the public cloud (or external cloud) computing resources are dynamically provisioned over the Internet via Web applications or Web services from an off-site third-party provider. Public clouds are run by third parties, and applications from different customers are likely to be mixed together on the cloud’s servers, storage systems, and networks.

Private cloud (or internal cloud) refers to cloud computing on private networks. Private clouds are built for the exclusive use of one client, providing full control over data, security, and quality of service. Private clouds can be built and managed by a company’s own IT organization or by a cloud provider.

A hybrid cloud environment combines multiple public and private cloud models. Hybrid clouds introduce the complexity of determining how to distribute applications across both a public and private cloud.

1.1.3 Cloud Computing Versus Cloud Services

In this section we present two tables that show the differences and major attributes of cloud computing versus cloud services (Jens, 2008). Cloud computing is the IT foundation for cloud services and it consists of technologies that enable cloud services. The key attributes of cloud computing are shown in Table 1.1.

In Key attributes of cloud services are summarized in Table 1.2 (Jens, 2008).

Table 1.1 Key Cloud Computing Attributes (adapted from Jens (2008))

Attributes	Description
Infrastructure systems	It includes servers, storage, and networks that can scale as per user demand.
Application software	It provides Web-based user interface, Web services APIs, and a rich variety of configurations.
Application development and deployment software	It supports the development and integration of cloud application software.
System and application management software	It supports rapid self-service provisioning and configuration and usage monitoring.
IP networks	They connect end users to the cloud and the infrastructure components.

Table 1.2 Key Attributes of Cloud Services (adapted from Jens (2008))

Attributes	Description
Offsite. Third-party provider	In the cloud execution, it is assumed that third-party provides services. There is also a possibility of in-house cloud service delivery.
Accessed via the Internet	Services are accessed via standard-based, universal network access. It can also include security and quality-of-service options.
Minimal or no IT skill required	There is a simplified specification of requirements.
Provisioning	It includes self-service requesting, near real-time deployment, and dynamic and fine-grained scaling.
Pricing	Pricing is based on usage-based capability and it is fine-grained.
User interface	User interface include browsers for a variety of devices and with rich capabilities.
System interface	System interfaces are based on Web services APIs providing a standard framework for accessing and integrating among cloud services.
Shared resources	Resources are shared among cloud services users; however via configuration options with the service, there is the ability to customize.

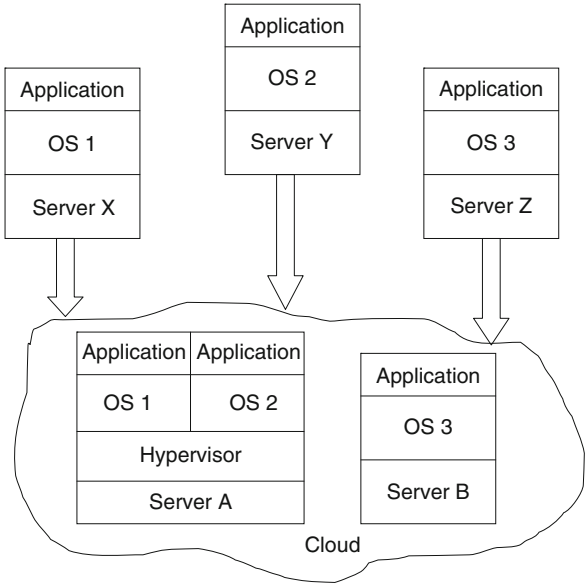
1.2 Enabling Technologies

Key technologies that enabled cloud computing are described in this section; they include virtualization, Web service and service-oriented architecture, service flows and workflows, and Web 2.0 and mashup.

1.2.1 Virtualization

The advantage of cloud computing is the ability to virtualize and share resources among different applications with the objective for better server utilization. Figure 1.6 shows an example Jones]. In non-cloud computing three independent platforms exist for three different applications running on its own server. In the cloud, servers can be shared, or virtualized, for operating systems and applications resulting in fewer servers (in specific example two servers).

Fig. 1.6 An example of virtualization: in non-cloud computing there is a need for three servers; in the cloud computing, two servers are used (adapted from Jones)



Virtualization technologies include virtual machine techniques such as VMware and Xen, and virtual networks, such as VPN. Virtual machines provide virtualized IT-infrastructures on-demand, while virtual networks support users with a customized network environment to access cloud resources.

1.2.2 Web Service and Service Oriented Architecture

Web Services and Service Oriented Architecture (SOA) are not new concepts; however they represent the base technologies for cloud computing. Cloud services are typically designed as Web services, which follow industry standards including WSDL, SOAP, and UDDI. A Service Oriented Architecture organizes and manages Web services inside clouds (Vouk, 2008). A SOA also includes a set of cloud services, which are available on various distributed platforms.

1.2.3 Service Flow and Workflows

The concept of service flow and workflow refers to an integrated view of service-based activities provided in clouds. Workflows have become one of the important areas of research in the field of database and information systems (Vouk, 2008).

1.2.4 Web 2.0 and Mashup

Web 2.0 is a new concept that refers to the use of Web technology and Web design to enhance creativity, information sharing, and collaboration among users (Wang, Tao,

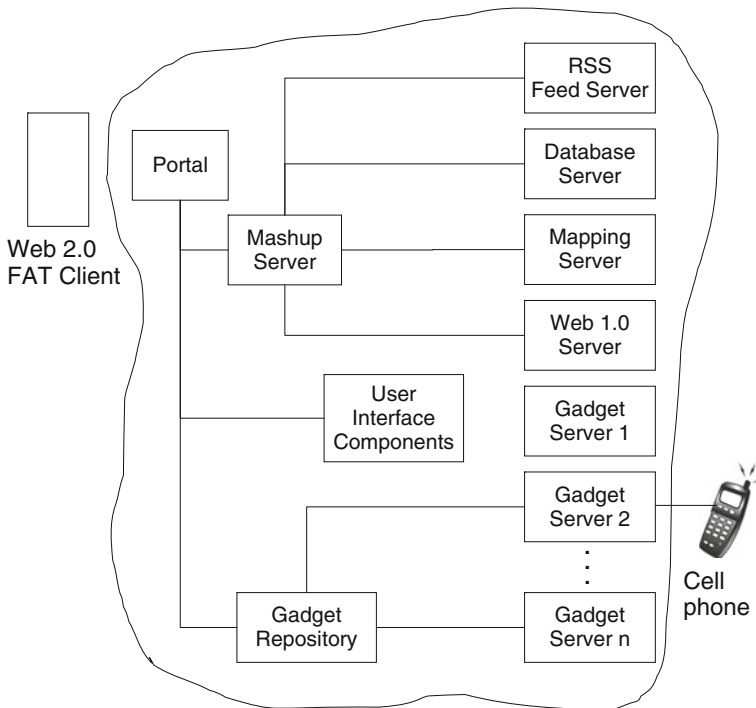


Fig. 1.7 Cloud computing architecture uses various components at different levels (adapted from Hutchinson and Ward (2009))

& Kunze, 2008). On the other hand, Mashup is a web application that combines data from more than one source into a single integrated storage tool. Both technologies are very beneficial for cloud computing.

Figure 1.7 shows a cloud computing architecture, adapted from Hutchinson and Ward (2009), in which an application reuses various components. The components in this architecture are dynamic in nature, operate in a SaaS model, and leverage SOA. The components closer to the user are smaller in nature and more reusable. The components in the center contain aggregate and extend services via mashup servers and portals. Data from one service (such as addresses in a database) can be mashed up with mapping information (such as Yahoo or Google maps) to produce an aggregated view of the information.

1.3 Cloud Computing Features

Cloud computing brings a number of new features compared to other computing paradigms (Wang et al., 2008; Grossman, 2009). There are briefly described in this section.

- **Scalability and on-demand services**
Cloud computing provides resources and services for users on demand. The resources are scalable over several data centers.
- **User-centric interface**
Cloud interfaces are location independent and can be accessed by well established interfaces such as Web services and Internet browsers.
- **Guaranteed Quality of Service (QoS)**
Cloud computing can guarantee QoS for users in terms of hardware/CPU performance, bandwidth, and memory capacity.
- **Autonomous system**
The cloud computing systems are autonomous systems managed transparently to users. However, software and data inside clouds can be automatically reconfigured and consolidated to a simple platform depending on user's needs.
- **Pricing**
Cloud computing does not require up-front investment. No capital expenditure is required. Users pay for services and capacity as they need them.

1.3.1 Cloud Computing Standards

Cloud computing standards have not been yet fully developed; however a number of existing typically lightweight, open standards have facilitated the growth

Table 1.3 Cloud computing standards (“Cloud Computing,” Wikipedia, http://en.wikipedia.org/wiki/Cloud_computing)

Applications	Communications: HTTP, XMPP Security: OAuth, OpenID, SSL/TLS Syndication: Atom
Client	Browsers: AJAX Offline: HTML5
Implementations	Virtualization: OVF
Platform	Solution stacks: LAMP
Service	Data: XML, JSON Web services: REST

of cloud computing (“Cloud Computing,” Wikipedia, http://en.wikipedia.org/wiki/Cloud_computing). Table 1.3 illustrates several of these open standards, which are currently used in cloud computing.

1.3.2 Cloud Computing Security

One of the critical issues in implementing cloud computing is taking virtual machines, which contain critical applications and sensitive data, to public and shared cloud environments. Therefore, potential cloud computing users are concerned about the following security issues (“Cloud Computing Security,” Third Brigade, www.cloudreadysecurity.com).

- Will the users still have the same security policy control over their applications and services?
- Can it be proved to the organization that the system is still secure and meets SLAs?
- Is the system complaint and can it be proved to company’s auditors?

In traditional data centers, the common approaches to security include perimeter firewall, demilitarized zones, network segmentation, intrusion detection and prevention systems, and network monitoring tools.

The security requirements for cloud computing providers begins with the same techniques and tools as for traditional data centers, which includes the application of a strong network security perimeter. However, physical segmentation and hardware-based security cannot protect against attacks between virtual machines on the same server. Cloud computing servers use the same operating systems, enterprise and Web applications as localized virtual machines and physical servers. Therefore, an attacker can remotely exploit vulnerabilities in these systems and applications. In addition, co-location of multiple virtual machines increases the attack surface and risk to MV-to-VM compromise. Intrusion detection and prevention systems need to be able to detect malicious activity in the VM level, regardless of the location of the

VM within the virtualized cloud environment (“Cloud Computing Security,” Third Brigade, www.cloudreadysecurity.com).

In summary, the virtual environments that deploy the security mechanisms on virtual machines including firewalls, intrusion detection and prevention, integrity monitoring, and log inspection, will effectively make VM cloud secure and ready for deployment.

1.4 Cloud Computing Platforms

Cloud computing has great commercial potential. According to market research firm IDC, IT cloud services spending will grow from about \$16B in 2008 to about \$42B in 2012 and to increase its share of overall IT spending from 4.2% to 8.5%.

Table 1.4 presents key players in cloud computing platforms and their key offerings.

Table 1.4 Key Players in Cloud Computing Platforms (adapted from Lakshmanan (2009))

Company	Cloud computing platform	Year of launch	Key offerings
Amazon.com	AWS (Amazon Web Services)	2006	Infrastructure as a service (Storage, Computing, Message queues, Datasets, Content distribution)
Microsoft	Azure	2009	Application platform as a service (.Net, SQL data services)
Google	Google App. Engine	2008	Web Application Platform as a service (Python run time environment)
IBM	Blue Cloud	2008	Virtualized Blue cloud data center
Salesforce.com	Force.com	2008	Proprietary 4GL Web application framework as an on Demand platform

Table 1.5 compares three cloud computing platforms, Amazon, Google, and Microsoft, in terms of their capabilities to map to different development models and scenarios (“Which Cloud Platform is Right for You?,” www.cumulux.com).

1.4.1 Pricing

Pricing for cloud platforms and services is based on three key dimensions: (i) storage, (ii) bandwidth, and (iii) compute.

Storage is typically measured as average daily amount of data stored in GB over a monthly period.

Table 1.5 Cloud Computing Platforms and Different Scenarios (adapted from “Which Cloud Platform is Right for You?,” www.cumulux.com.)

(1) Scenario	On-premise application unchanged in the cloud
Characteristics	Multiple red legacy, java or .NET based application
Amazon	Threat the machine as another server in the data center and do the necessary changes to configuration
Google	Needs significant refactoring of application and data logic for existing Java application
Microsoft	If existing app is ASP.NET application, then re-factor data, otherwise refactoring effort can be quite significant depending on the complexity
(2) Scenario	Scalable Web application
Characteristics	Moderate to high Web application with a back-end store and load balancing
Amazon	Threat the machine instance as another server in the data center and do the necessary changes to configuration. But scalability and elasticity is manual configuration
Google	Use dynamically scalable features of AppEngine and scripting technologies to build rich applications
Microsoft	Build scalable Web applications using familiar .NET technologies. Scaling up/down purely driven by configuration.
(3) Scenario	Parallel processing computational application
Characteristics	Automated long running processing with little to no user interaction.
Amazon	Need to configure multiple machine instances depending on the scale needed and manage the environments.
Google	Platform has minimal built-in support for building compute heavy applications. Certain application scenarios, such as image manipulation, are easier to develop with built-in platform features.
Microsoft	With worker roles and storage features like Queues and blobs, it is easy to build a compute heavy application that can be managed and controlled for scalability and elasticity.
(4) Scenario	Application in the cloud interacts with on-premise data
Characteristics	Cloud based applications interacting with on-premise apps for managing transactions of data
Amazon	Applications in EC2 server cloud can easily be configured to interact with applications running on premise.
Google	No support from the platform to enable this scenario. Possible through each app using intermediary store to communicate.
Microsoft	From features like Service Bus to Sync platform components it is possible to build compelling integration between the two environments.
(5) Scenario	Application in the cloud interacts with on-premise application
Characteristics	On-premise applications
Amazon	Applications in EC2 server cloud can easily be configured to interact with applications running on premise.
Google	No support from the platform to enable this scenario. Possible through each app using intermediary store to communicate.
Microsoft	From features like Service Bus to Sync platform components it is possible to build compelling integration between the two environments.

Bandwidth is measured by calculating the total amount of data transferred in and out of platform service through transaction and batch processing. Generally, data transfer between services within the same platform is free in many platforms.

Compute is measured as the time units needed to run an instance, or application, or machine to servicing requests. Table 6 compares pricing for three major cloud computing platforms.

In summary, by analyzing the cost of cloud computing, depending on the application characteristics the cost of deploying an application could vary based on the selected platform. From Table 1.6, it seems that the unit pricing for three major platforms is quite similar. Besides unit pricing, it is important to translate it into monthly application development, deployments and maintenance costs.

Table 1.6 Pricing comparison for major cloud computing platforms (adapted from “Which Cloud Platform is Right for You?,” www.cumulux.com.)

Resource	UNIT	Amazon	Google	Microsoft
Stored data	GB per month	\$0.10	\$0.15	\$0.15
Storage transaction	Per 10 K requests	\$0.10		\$0.10
Outgoing bandwidth	GB	\$0.10 – \$0.17	\$0.12	\$0.15
Incoming bandwidth	GB	\$0.10	\$0.10	\$0.10
Compute time	Instance Hours	\$0.10 – \$1.20	\$0.10	\$0.12

1.4.2 Cloud Computing Components and Their Vendors

The main elements comprising cloud computing platforms include computer hardware, storage, infrastructure, computer software, operating systems, and platform virtualization. The leading vendors providing cloud computing components are shown in Table 1.7 (“Cloud Computing,” Wikipedia, http://en.wikipedia.org/wiki/Cloud_computing).

Table 1.7 The leading vendors of cloud computing components

Cloud computing components	Vendors
Computer hardware	Dell, HP, IBM, Sun
Storage	Sun, EMC, IBM
Infrastructure	Cisco, Juniper Networks, Brocade Communication
Computer software	3tera. Eucalyptus. G-Eclipse. Hadoop
Operating systems	Solaris, AIX, Linux (Red Hat, Ubuntu)
Platform virtualization	Citrix, VMWare, IBM, Xen, Linux KVM, Microsoft, Sun xVM

1.5 Example of Web Application Deployment

In this section we present an example how the combination of virtualization and on of self service facilitate application deployment (Sun Microsystems, 2009). In this example we consider a two-tier Web application deployment using cloud, as illustrated in Fig. 1.8.

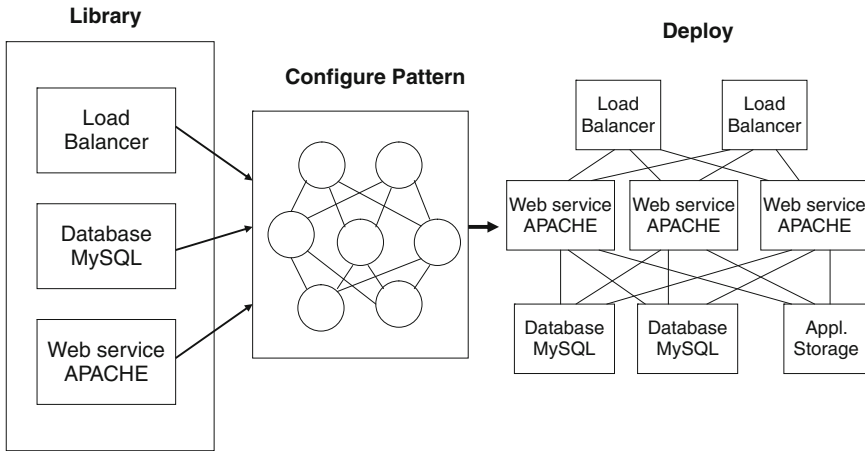


Fig. 1.8 Example of the deployment of an application into a two-tier Web server architecture using cloud computing (adapted from Sun Microsystems (2009))

The following steps comprise the deployment of the application:

- The developer selects a load balancer, Web server, and database server appliances from a library of preconfigured virtual machine images.
- The developer configures each component to make a custom image. The load balancer is configured, the Web server is populated with its static content by uploading it to the storage cloud, and the database server appliances are populated with dynamic content for the site.
- The developer then layers custom code into the new architecture, in this way making the components meet specific application requirements.
- The developer chooses a pattern that takes the images for each layer and deploys them, handling networking, security, and scalability issues.

The secure, high-availability Web application is up and running. When the application needs to be updated, the virtual machine images can be updated, copied across the development chain, and the entire infrastructure can be redeployed.

In this example, a standard set of components can be used to quickly deploy an application. With this model, enterprise business needs can be met quickly, without the need for the time-consuming, manual purchase, installation, cabling, and configuration of servers, storage, and network infrastructure.

Small and medium enterprises were the early adopters to cloud computing. However, there are recently a number of examples of cloud computing adoptions in the large enterprises. Table 1.8 illustrates three examples of cloud computing use in the large enterprises (Lakshmanan, 2009).

Table 1.8 Cloud computing examples in large enterprises

Enterprise	Scenario	Usage	Benefits
Eli Lilly	R&D High Performance Computing	Amazon server and storage cluster for drug discovery analysis and modeling.	Quick deployment time at a lower cost.
New York Times	Data Conversion	Conversion of archival articles (3 million) into new data formats using Amazon elastic compute services.	Rapid provisioning and higher elasticity on the infrastructure resources.
Pitney Bowes	B2B Application	Hosted model mail printing application for clients. Uses MS Azure.net and SQL services for the hosted model option (2009 Go live).	Flexibility at a lower cost and new business opportunity.

1.6 Cloud Computing Challenges

In summary, the new paradigm of cloud computing provides a number of benefits and advantages over the previous computing paradigms and many organizations are adopting it. However, there are still a number of challenges, which are currently addressed by researchers and practitioners in the field (Leavitt, 2009). They are briefly presented below.

1.6.1 Performance

The major issue in performance can be for some intensive transaction-oriented and other data-intensive applications, in which cloud computing may lack adequate performance. Also, users who are at a long distance from cloud providers may experience high latency and delays.

1.6.2 Security and Privacy

Companies are still concerned about security when using cloud computing. Customers are worried about the vulnerability to attacks, when information and

critical IT resources are outside the firewall. The solution for security assumes that cloud computing providers follow standard security practices, as described in Section 1.3.2.

1.6.3 Control

Some IT departments are concerned because cloud computing providers have a full control of the platforms. Cloud computing providers typically do not design platforms for specific companies and their business practices.

1.6.4 Bandwidth Costs

With cloud computing, companies can save money on hardware and software; however they could incur higher network bandwidth charges. Bandwidth cost may be low for smaller Internet-based applications, which are not data intensive, but could significantly grow for data-intensive applications.

1.6.5 Reliability

Cloud computing still does not always offer round-the-clock reliability. There were cases where cloud computing services suffered a few-hours outages.

In the future, we can expect more cloud computing providers, richer services, established standards, and best practices.

In the research arena, HP Labs, Intel, and Yahoo have launched the distributed Cloud Research Test Bed, with facilities in Asia, Europe, and North America, with the objective to develop innovations including cloud computing specific chips. IBM has launched the Research Computing Cloud, which is an on-demand, globally accessible set of computing resources that support business processes.

1.7 Cloud Computing in the Future

In summary, cloud computing is definitely a type of computing paradigm/architecture that will remain for a long time to come. In the near future, cloud computing can emerge in various directions. One possible scenario for the future is that an enterprise may use a distributed hybrid cloud as illustrated in Fig. 1.9.

According to this scenario, the enterprise will use the core applications on its private cloud, while some other applications will be distributed on several private clouds, which are optimized for specific applications.

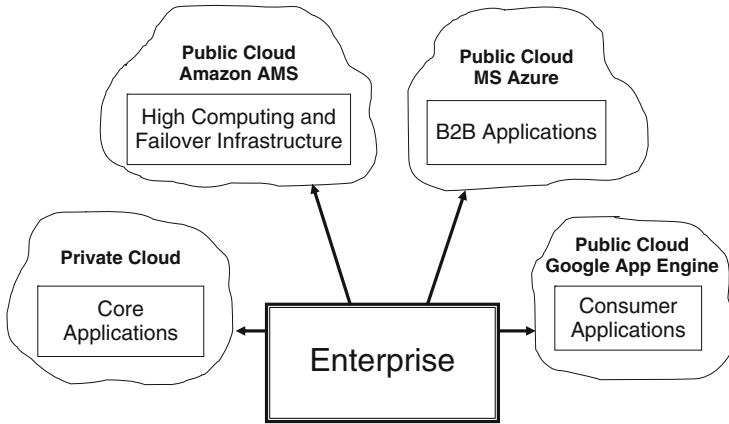


Fig. 1.9 Distributed hybrid cloud architecture (adapted from Lakshmanan (2009))

References

- Chappell, D. (August 2008). *A short introduction to cloud platforms: An enterprise-oriented view*. San Francisco, CA: Chappel and Associates.
- Grossman, R. L. (March/April 2009). The case for cloud computing. *IEEE ITPro*, 23–27.
- Hutchinson, C., & Ward, J. (March/April 2009). Navigation the next-generation application architecture. *IEEE ITPro*, 18–22.
- Jens, F. (September 2008). *Defining cloud services and cloud computing*. <http://blogs.idc.com/ie/?p=190>.
- Jones, M. T. *Cloud computing with linux*. www.ibm.com/developerworks/linux/library/l-cloud-computing.
- Lakshmanan, G. (April 2009). *Cloud computing – Relevance to enterprise*. Infosys White Paper.
- Leavitt, N. (January 2009). Is cloud computing really ready for prime time? *IEEE Computer*, 15–20.
- Sun Microsystems (June 2009). *Introduction to cloud computing architecture*. White Paper, Sun Microsystems.
- Voas, J., & Zhang, J. (March/April 2009). Cloud computing: New wine or just a new bottle? *IEEE ITPro*, 15–17.
- Vouk, M. A. (June 2008). Cloud computing – Issues, research and implementations. *Proceedings of the ITI 30th International Conference on Information Technology Interfaces, Cavtat, Croatia*, 31–40.
- Wang, L., Tao, J., & Kunze, M. (2008). Scientific cloud computing: Early definition and experience. *Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications, Austin, TX*, 825–830.