Handbook of Cloud Computing

Borko Furht · Armando Escalante
Editors

Handbook of Cloud Computing

Springer

Editors
Borko Furht
Department of Computer and Electrical
Engineering and Computer Science
Florida Atlantic University
777 Glades Road
Boca Raton, FL 33431, USA
bfurht@fau.edu

Armando Escalante
LexisNexis
6601 Park of Commerce Boulevard
Boca Raton, FL 33487, USA
armando.escalante@lexisnexis.com

ISBN 978-1-4419-6523-3 e-ISBN 978-1-4419-6524-0
DOI 10.1007/978-1-4419-6524-0
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010934567

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Cloud computing has become a great solution for providing a flexible, on-demand, and dynamically scalable computing infrastructure for many applications. Cloud computing also presents a significant technology trends, and it is already obvious that it is reshaping information technology processes and the IT marketplace.

This Handbook is a carefully edited book – contributors are 65 worldwide experts in the field of cloud computing and their applications. The Handbook Advisory Board, comprised of nine researchers and practitioners from academia and industry, helped in reshaping the Handbook and selecting the right topics and creative and knowledgeable contributors. The scope of the book includes leading-edge cloud computing technologies, systems, and architectures; cloud computing services; and a variety of cloud computing applications.

The Handbook comprises four parts, which consist of 26 chapters. The first part on *Technologies and Systems* includes articles dealing with cloud computing technologies, storage and fault tolerant strategies in cloud computing, workflows, grid computing technologies, and the role of networks in cloud computing.

The second part on *Architectures* focuses on articles on several specific architectural concepts applied in cloud computing, including enterprise knowledge clouds, high-performance computing clouds, clouds with vertical load distribution, and peer-to-peer based clouds.

The third part on *Services* consists of articles on various issues relating to cloud services, including types of services, service scalability, scientific services, and dynamic collaborative services.

The forth part on *Applications* describes various cloud computing applications from enterprise knowledge clouds, scientific and statistical computing, scientific data management, to medical applications.

With the dramatic growth of cloud computing technologies, platforms and services, this Handbook can be the definitive resource for persons working in this field as researchers, scientists, programmers, engineers, and users. The book is intended for a wide variety of people including academicians, designers, developers, educators, engineers, practitioners, researchers, and graduate students. This book can also be beneficial for business managers, entrepreneurs, and investors. The book can have a great potential to be adopted as a textbook in current and new courses on Cloud Computing.

The main features of this Handbook can be summarized as:

1. The Handbook describes and evaluates the current state-of-the-art in a new field of cloud computing.
2. It also presents current systems, services, and main players in this explosive field.
3. Contributors to the Handbook are the leading researchers from academia and practitioners from industry.

We would like to thank the authors for their contributions. Without their expertise and effort, this Handbook would never come to fruition. Springer editors and staff also deserve our sincere recognition for their support throughout the project.

Boca Raton, Florida

Borko Furht
Armando Escalante

Contents

Contributors

Ankur Agarwal Department of Computer Science and Engineering, FAU, Boca Raton, FL, USA, ankur@cse.fau.edu

Tim Bell Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, New Zealand, tim.bell@canterbury.ac.nz

Paolo Bientinesi AICES, RWTH, Aachen, Germany, pauldj@aices.rwth-aachen.de

Norman Bobroff IBM Watson Research Center, Hawthorne, NY, USA, bobroff@us.ibm.com

Juan Cáceres Telefónica Investigación y Desarrollo, Madrid, Spain, caceres@tid.es

David Chapman Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County, MD, USA, dchapm2@umbc.edu

Jinjun Chen Faculty of Information and Communication Technologies, Swinburne University of Technology, Hawthorn, Melbourne, Australia 3122, jchen@swin.edu.au

Lei Chen Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, leichen@cse.ust.hk

Karim Chine Cloud Era Ltd, Cambridge, UK, karim.chine@polytechnique.org

David Cohen Cloud Infrastructure Group, EMC Corporation, Cambridge, MA, USA, david.cohen@emc.com

Vincenzo D. Cunsolo University of Messina, Contrada di Dio, S. Agata, Messina, Italy, vdcunsolo@unime.it

Kemal A. Delic Hewlett-Packard Co., New York, NY, USA, kemal.delic@hp.com

Mac Devine IBM Corporation, Research Triangle Park, NC, USA, wdevine@us.ibm.com

Salvatore Distefano University of Messina, Contrada di Dio, S. Agata, Messina, Italy, sdistefano@unime.it

Patrick Dreher Renaissance Computing Institute, Chapel Hill, NC, 27517 USA, dreher@renci.org

Kathleen Ericson Department of Computer Science, Colorado State University, Fort Collins, CO, USA, ericson@cs.colostate.edu

Liana Fong IBM Watson Research Center, Hawthorne, NY, USA, llfong@us.ibm.com

Ian Foster Argonne National Laboratory, Argonne, IL, USA; The University of Chicago, Chicago, IL, USA, foster@anl.gov

Marc Eduard Frîncu Institute e-Austria, Blvd. Vasile Parvan No 4 300223, Room 045B, Timisoara, Romania, mfrincu@info.uvt.ro

Borko Furht Department of Computer & Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA, bfurht@fau.edu

Wei Gao Services Computing Technology and System Lab; Cluster and Grid Computing Lab, Huazhong University of Science and Technology, Wuhan, China, gaowei715@gmail.com

Giulio Giunta Department of Applied Science, University of Napoli Parthenope, Napoli, Italy, giulio.giunta@uniparthenope.it

Milt Halem Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County, MD, USA, halem@umbc.edu

Mohammad Mehedi Hassan Department of Computer Engineering, Kyung Hee University, Global Campus, South Korea, hassan@khu.ac.kr

Nathan Henehan Senior Software Developer, NACS Solutions, Oberlin, OH, USA, nhenehan@gmail.com

Juan J. Hierro Telefónica Investigación y Desarrollo, Madrid, Spain, jhierro@tid.es

Dachuan Huang Services Computing Technology and System Lab; Cluster and Grid Computing Lab, Huazhong University of Science and Technology, Wuhan, China, hdc1112@gmail.com

Eui-Nam Huh Department of Computer Engineering, Kyung Hee University, Global Campus, South Korea, johnhuh@khu.ac.kr

Roman Iakymchuk AICES, RWTH, Aachen, Germany, iakymchuk@aices.rwth-aachen.de

Shadi Ibrahim Services Computing Technology and System Lab; Cluster and Grid Computing Lab, Huazhong University of Science and Technology, Wuhan, China, shadi@hust.edu.cn

Hai Jin Services Computing Technology and System Lab; Cluster and Grid Computing Lab, Huazhong University of Science and Technology, Wuhan, China, hjin@hust.edu.cn

Karuna P. Joshi Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County, MD, USA, kjoshi1@umbc.edu

Burton S. Kaliski Jr. Office of the CTO, EMC Corporation, Hopkinton, MA, USA, burt.kaliski@emc.com

Hari Kalva Department of Computer Science and Engineering, FAU, Boca Raton, FL, USA, hari.kalva@fau.edu

Giuliano Laccetti Department of Mathematics and Applications, University of Napoli Federico II, Napoli, Italy, giuliano.laccetti@unina.it

Rui Li Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, cslr@cse.ust.hk

Wen-Syan Li SAP Technology Lab, Shanghai, China, wen-syan.li@sap.com

Geng Lin IBM Alliance, Cisco Systems, San Francisco, CA, USA, gelin@cisco.com

Xiao Liu Faculty of Information and Communication Technologies, Swinburne University of Technology, Hawthorn, Melbourne, Australia 3122, xliu@groupwise.swin.edu.au

Yanbin Liu IBM Watson Research Center, Hawthorne, NY, USA, ygliu@us.ibm.com

Anthony M. Middleton LexisNexis Risk Solutions, Boca Raton, FL, USA, tony.middleton@lexisnexis.com

Raffaele Montella Department of Applied Science, University of Napoli Parthenope, Napoli, Italy, raffaele.montella@uniparthenope.it

Jeff Napper Vrije Universiteit, Amsterdam, Netherlands, jnapper@cs.vu.nl

Phuong Nguyen Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County, MD, USA, phuong3@umbc.edu

Jeffrey M. Nick Office of the CTO, EMC Corporation, Hopkinton, MA, USA, jeff.nick@emc.com

Sangyoon Oh WISE Research Lab, School of Information and Communication Engineering, Ajou University, Suwon, South Korea, syoh@ajou.ac.kr

Shrideep Pallickara Department of Computer Science, Colorado State University, Fort Collins, CO, USA, shrideep@cs.colostate.edu

Sangmi Lee Pallickara Department of Computer Science, Colorado State University, Fort Collins, CO, USA, sangmi@cs.colostate.edu

A.S. Pandya Department of Computer Science and Engineering, FAU, Boca Raton, FL, USA, pandya@fau.edu

Manish Parashar NSF CAC, Rutgers University, Piscataway, NJ, USA, parashar@rutgers.edu

Thomas Phan Microsoft Corporation, Washington, DC, USA, thomas.phan@acm.org

Marlon Pierce Community Grids Lab, Indiana University, Bloomington, IN, USA, mpierce@cs.indiana.edu

Álvaro Polo Telefónica Investigación y Desarrollo, Madrid, Spain, apv@tid.es

Antonio Puliafito University of Messina, Contrada di Dio, S. Agata, Messina, Italy, apuliafito@unime.it

Jeff A. Riley Hewlett-Packard Co., New York, NY, USA, jeff.riley@hp.com

Ivan Rodero NSF CAC, Rutgers University, Piscataway, NJ, USA, irodero@rutgers.edu

Luis Rodero-Merino INRIA-ENS, INRIA, Lyon, France, luis.rodero-merino@ens-lyon.fr

S. Masoud Sadjadi CIS, Florida International University, Miami, FL, USA, sadjadi@cis.fiu.edu

Marco Scarpa University of Messina, Contrada di Dio, S. Agata, Messina, Italy, mscarpa@unime.it

Eric Sills North Carolina State University, Raleigh, NC 27695, USA, edsills@ncsu.edu

Mikael Fernandus Simalango WISE Research Lab, Ajou University, Suwon, South Korea, mikael@ajou.ac.kr

Vivek Somashekarappa Armellini Inc., Palm City, FL, USA, vsomashekar@gmail.com

Luis M. Vaquero Telefónica Investigación y Desarrollo, Madrid, Spain, lmvg@tid.es

David Villegas CIS, Florida International University, Miami, FL, USA, dvill013@cis.fiu.edu

Mladen A. Vouk Department of Computer Science, North Carolina State University, Box 8206, Raleigh, NC 27695, USA, vouk@ncsu.edu

Song Wu Services Computing Technology and System Lab; Cluster and Grid Computing Lab, Huazhong University of Science and Technology, Wuhan, China, wusong@mail.hust.edu.cn

Yun Yang Faculty of Information and Communication Technologies, Swinburne University of Technology, Hawthorn, Melbourne, Australia 3122, yyang@groupwise.swin.edu.au

Yelena Yesha Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County, MD, USA, yeyesha@csee.umbc.edu

Yaacov Yesha Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County, MD, USA, yayesha@umbc.edu

Dong Yuan Faculty of Information and Communication Technologies, Swinburne University of Technology, Hawthorn, Melbourne, Australia 3122, dyuan@groupwise.swin.edu.au

Gaofeng Zhang Faculty of Information and Communication Technologies, Swinburne University of Technology, Hawthorn, Melbourne, Australia 3122, gzhang@groupwise.swin.edu.au

Jinzy Zhu IBM Cloud Computing Center, China, jinzyzhu@cn.ibm.com

About the Editors

Borko Furht is a professor and chairman of the Department of Electrical & Computer Engineering and Computer Science at Florida Atlantic University (FAU) in Boca Raton, Florida. He is also Director of recently formed NSF-sponsored

Industry/University Cooperative Research Center on Advanced Knowledge Enablement. Before joining FAU, he was a vice president of research and a senior director of development at Modcomp (Ft. Lauderdale), a computer company of Daimler Benz, Germany; a professor at University of Miami in Coral Gables, Florida; and a senior researcher in the Institute Boris Kidric-Vinca, Yugoslavia. Professor Furht received a Ph.D. degree in electrical and computer engineering from the University of Belgrade. His current research is in multimedia systems, video coding and compression, 3D video and image systems, wireless multimedia, and Internet and cloud computing. He is presently Principal Investigator and Co-PI of several multiyear, multimillion-dollar projects, including NSF PIRE project and NSF High-Performance Computing Center. He is the author of numerous books and articles in the areas of multimedia, computer architecture, real-time computing, and operating systems. He is a founder and editor-in-chief of the *Journal of Multimedia Tools and Applications* (Springer). He has received several technical and publishing awards, and has consulted for many high-tech companies including IBM, Hewlett-Packard, Xerox, General Electric, JPL, NASA, Honeywell, and RCA. He has also served as a consultant to various colleges and universities. He has given many invited talks, keynote lectures, seminars, and tutorials. He served on the Board of Directors of several high-tech companies.

Armando J. Escalante is Senior Vice President and Chief Technology Officer of Risk Solutions for the LexisNexis Group, a division of Reed Elsevier. In this position, Escalante is responsible for technology development, information systems and operations. Previously, Escalante was Chief Operating Officer for Seisint, a privately owned company, which was purchased by LexisNexis in 2004. In this position, he was responsible for Technology, Development and Operations. Prior to 2001, Escalante served as Vice President of Engineering and Operations for Diveo Broadband Networks, where he led world-class Data Centers located in the U.S. and Latin America. Before Diveo Broadband Networks, Escalante was VP for one of the fastest growing divisions of Vignette Corporation, an eBusiness software leader. Escalante earned his bachelor's degree in electronic engineering at the USB in Caracas, Venezuela and a master's degree in computer science from Stevens Institute of Technology, as well as a master's degree in business administration from West Coast University.

Part I
Technologies and Systems

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

B. Furht, A. Escalante (eds.), *Handbook of Cloud Computing*,
DOI 10.1007/978-1-4419-6524-0_1,

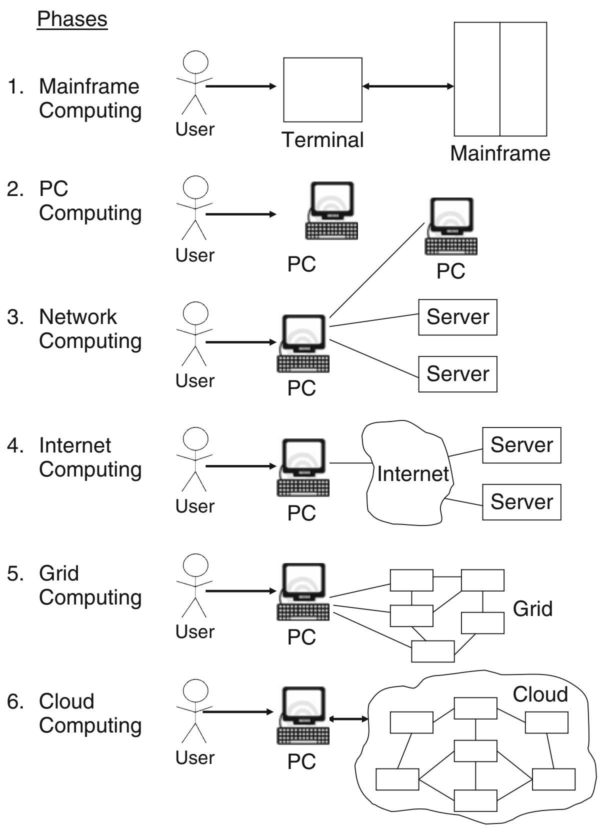

Fig. 1.1 Six computing paradigms – from mainframe computing to Internet computing, to grid computing and cloud computing (adapted from Voas and Zhang (2009))

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 cashing 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

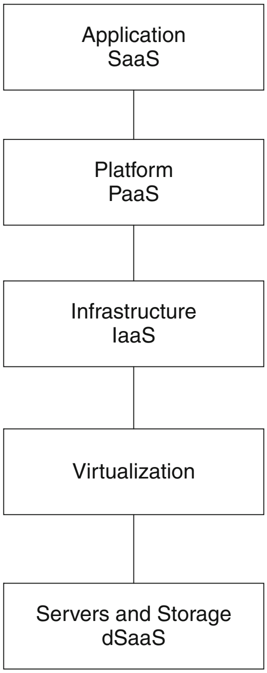

Fig. 1.2 Layered architecture of Cloud Computing (adapted from Jones)

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.

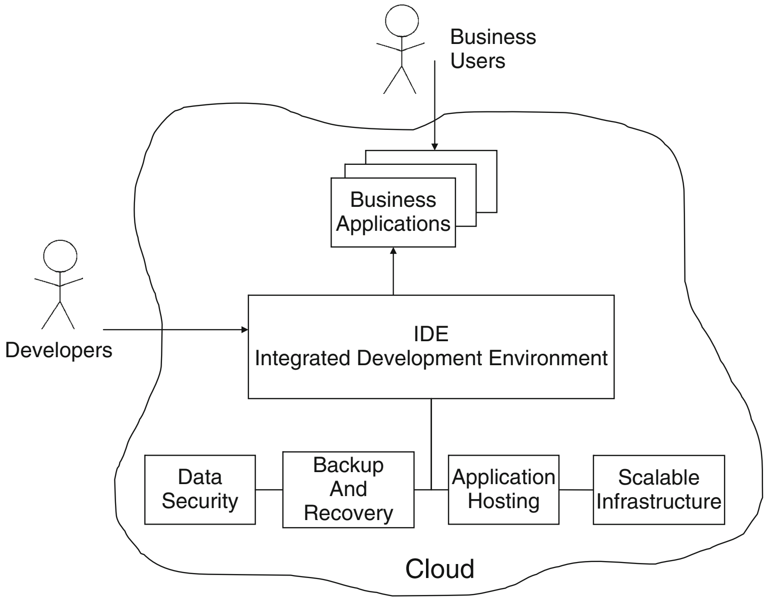

Fig. 1.3 The concept of Platform-as-a-Service, Zoho Creator (adapted from "Platform as a Service," http://www.zoho.com/creator/paas.html)

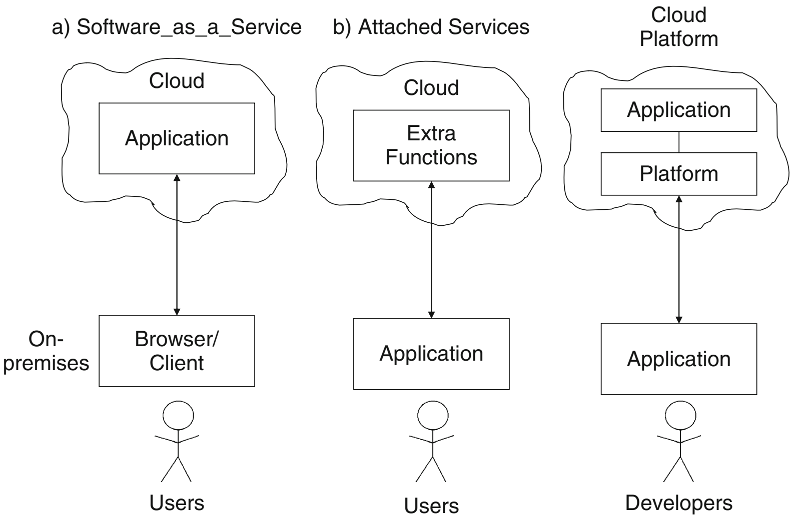

Fig. 1.4 The categories of cloud services (adopted from Chappell (2008))

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.

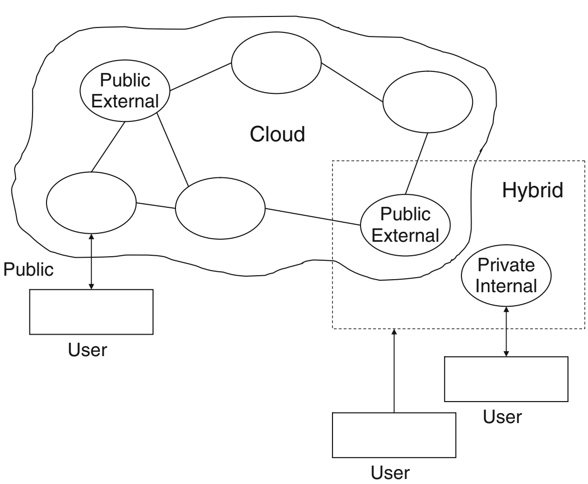

Fig. 1.5 Three types of cloud computing

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).

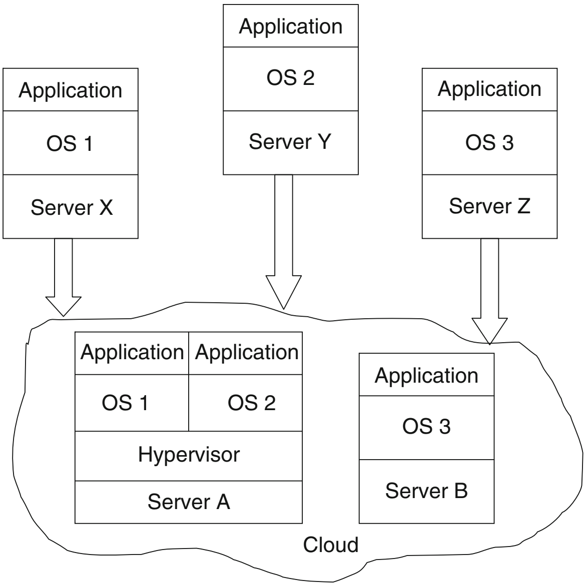

Fig. 1.6 An example of virtualization: in non-cloud computing there is a need for three servers; in the cloud computing, two servers are used (adapted from Jones)

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,

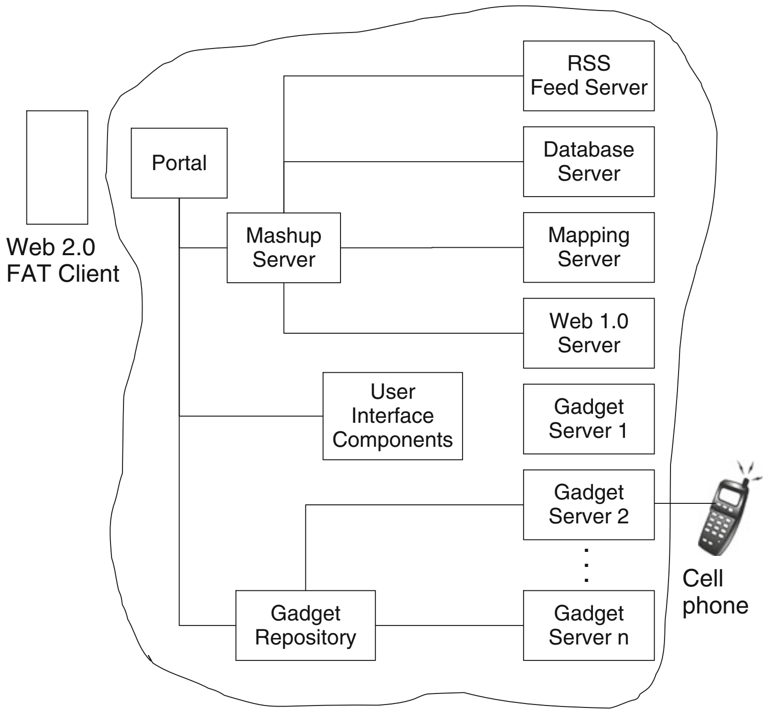

Fig. 1.7 Cloud computing architecture uses various components at different levels (adapted from Hutchinson and Ward (2009))

& 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 accesses by well established interfaces such as Web services and Internet browsers.

- Guaranteed Quality of Service (QoS)
  Cloud computed 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-from 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