

# Part I

## Examining the Value Proposition

### IN THIS PART

**Chapter 1**

Defining Cloud Computing

**Chapter 2**

Assessing the Value Proposition

**Chapter 3**

Understanding Cloud Architecture

**Chapter 4**

Understanding Services and Applications by Type



# Defining Cloud Computing

**C**loud computing refers to applications and services that run on a distributed network using virtualized resources and accessed by common Internet protocols and networking standards. It is distinguished by the notion that resources are virtual and limitless and that details of the physical systems on which software runs are abstracted from the user.

In an effort to better describe cloud computing, a number of cloud types have been defined. In this chapter, you learn about two different classes of clouds: those based on the deployment model and those based on the service model. The deployment model tells you where the cloud is located and for what purpose. Public, private, community, and hybrid clouds are deployment models.

Service models describe the type of service that the service provider is offering. The best-known service models are Software as a Service, Platform as a Service, and Infrastructure as a Service—the SPI model. The service models build on one another and define what a vendor must manage and what the client's responsibility is.

Cloud computing represents a real paradigm shift in the way in which systems are deployed. The massive scale of cloud computing systems was enabled by the popularization of the Internet and the growth of some large service companies. Cloud computing makes the long-held dream of utility computing possible with a pay-as-you-go, infinitely scalable, universally available system. With cloud computing, you can start very small and become big very fast. That's why cloud computing is revolutionary, even if the technology it is built on is evolutionary.

Not all applications benefit from deployment in the cloud. Issues with latency, transaction control, and in particular security and regulatory compliance are of particular concern.

## IN THIS CHAPTER

**Defining cloud computing**

**Learning about cloud types**

**Understanding the paradigm shift that is cloud computing**

**Comparing the benefits and disadvantages of cloud systems**

# Defining Cloud Computing

---

Cloud computing takes the technology, services, and applications that are similar to those on the Internet and turns them into a self-service utility. The use of the word “cloud” makes reference to the two essential concepts:

- **Abstraction:** Cloud computing abstracts the details of system implementation from users and developers. Applications run on physical systems that aren’t specified, data is stored in locations that are unknown, administration of systems is outsourced to others, and access by users is ubiquitous.
- **Virtualization:** Cloud computing virtualizes systems by pooling and sharing resources. Systems and storage can be provisioned as needed from a centralized infrastructure, costs are assessed on a metered basis, multi-tenancy is enabled, and resources are scalable with agility.

Computing as a utility is a dream that dates from the beginning of the computing industry itself. A set of new technologies has come along that, along with the need for more efficient and affordable computing, has enabled an on-demand system to develop. It is these enabling technologies that are the focal point of this book.

Many people mistakenly believe that cloud computing is nothing more than the Internet given a different name. Many drawings of Internet-based systems and services depict the Internet as a cloud, and people refer to applications running on the Internet as “running in the cloud,” so the confusion is understandable. The Internet has many of the characteristics of what is now being called cloud computing. The Internet offers abstraction, runs using the same set of protocols and standards, and uses the same applications and operating systems. These same characteristics are found in an intranet, an internal version of the Internet. When an intranet becomes large enough that a diagram no longer wishes to differentiate between individual physical systems, the intranet too becomes identified as a cloud.

Cloud computing is an abstraction based on the notion of pooling physical resources and presenting them as a virtual resource. It is a new model for provisioning resources, for staging applications, and for platform-independent user access to services. Clouds can come in many different types, and the services and applications that run on clouds may or may not be delivered by a cloud service provider. These different types and levels of cloud services mean that it is important to define what type of cloud computing system you are working with.

To help clarify how cloud computing has changed the nature of commercial system deployment, consider these three examples:

- **Google:** In the last decade, Google has built a worldwide network of datacenters to service its search engine. In doing so Google has captured a substantial portion of the world’s advertising revenue. That revenue has enabled Google to offer free software to users based on that infrastructure and has changed the market for user-facing software. This is the classic Software as a Service case described in Chapter 8.

- **Azure Platform:** By contrast, Microsoft is creating the Azure Platform. It enables .NET Framework applications to run over the Internet as an alternate platform for Microsoft developer software running on desktops, which you will learn about in Chapter 10.
- **Amazon Web Services:** One of the most successful cloud-based businesses is Amazon Web Services, which is an Infrastructure as a Service offering that lets you rent virtual computers on Amazon's own infrastructure. AWS is the subject of Chapter 9.

These new capabilities enable applications to be written and deployed with minimal expense and to be rapidly scaled and made available worldwide as business conditions permit. This is truly a revolutionary change in the way enterprise computing is created and deployed.

## Cloud Types

---

To discuss cloud computing intelligently, you need to define the lexicon of cloud computing; many acronyms in this area probably won't survive long. Most people separate cloud computing into two distinct sets of models:

- **Deployment models:** This refers to the location and management of the cloud's infrastructure.
- **Service models:** This consists of the particular types of services that you can access on a cloud computing platform.

This is a very useful demarcation that is now widely accepted.

## The NIST model

The United States government is a major consumer of computer services and, therefore, one of the major users of cloud computing networks. The U.S. National Institute of Standards and Technology (NIST) has a set of working definitions (<http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc>) that separate cloud computing into service models and deployment models. Those models and their relationship to essential characteristics of cloud computing are shown in Figure 1.1.

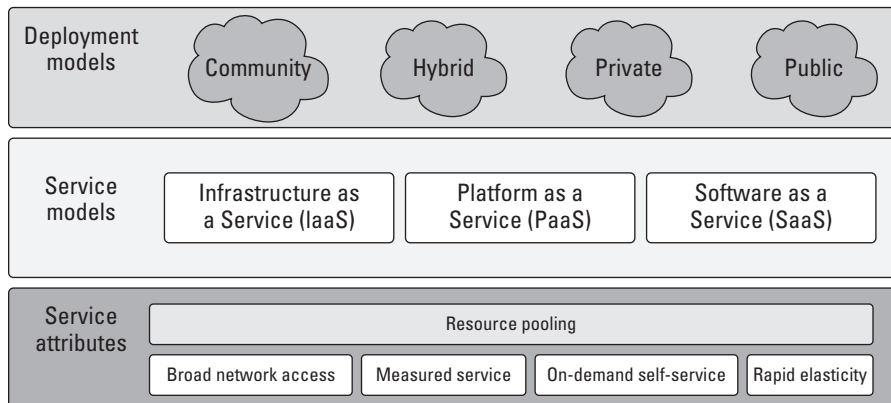
The NIST model originally did not require a cloud to use virtualization to pool resources, nor did it absolutely require that a cloud support multi-tenancy in the earliest definitions of cloud computing. Multi-tenancy is the sharing of resources among two or more clients. The latest version of the NIST definition does require that cloud computing networks use virtualization and support multi-tenancy.

## Part I: Examining the Value Proposition

---

**FIGURE 1.1**

The NIST cloud computing definitions



Because cloud computing is moving toward a set of modular interacting components based on standards such as the Service Oriented Architecture (described in Chapter 13), you might expect that future versions of the NIST model may add those features as well. The NIST cloud model doesn't address a number of intermediary services such as transaction or service brokers, provisioning, integration, and interoperability services that form the basis for many cloud computing discussions. Given the emerging roles of service buses, brokers, and cloud APIs at various levels, undoubtedly these elements need to be added to capture the whole story.

## The Cloud Cube Model

The Open Group maintains an association called the Jericho Forum (<https://www.opengroup.org/jericho/index.htm>) whose main focus is how to protect cloud networks. The group has an interesting model that attempts to categorize a cloud network based on four dimensional factors. As described in its paper called "Cloud Cube Model: Selecting Cloud Formations for Secure Collaboration" ([http://www.opengroup.org/jericho/cloud\\_cube\\_model\\_v1.0.pdf](http://www.opengroup.org/jericho/cloud_cube_model_v1.0.pdf)), the type of cloud networks you use dramatically changes the notion of where the boundary between the client's network and the cloud begins and ends.

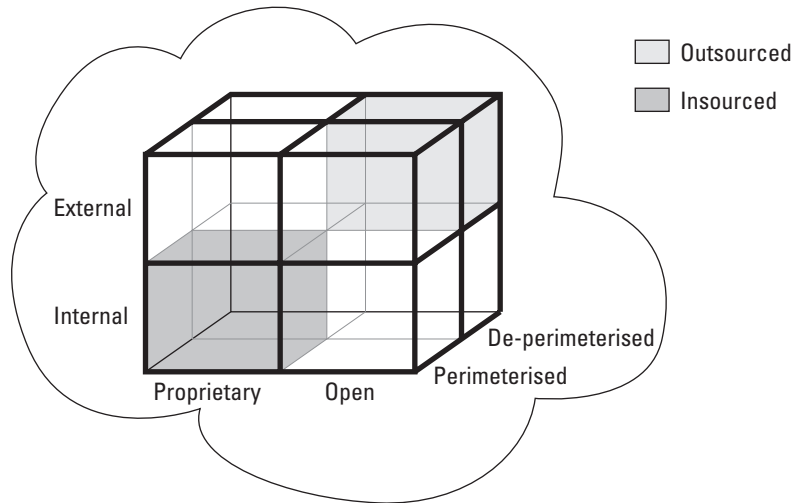
The four dimensions of the Cloud Cube Model are shown in Figure 1.2 and listed here:

- **Physical location of the data:** Internal (I) / External (E) determines your organization's boundaries.
- **Ownership:** Proprietary (P) / Open (O) is a measure of not only the technology ownership, but of interoperability, ease of data transfer, and degree of vendor application lock-in.

- **Security boundary:** Perimeterised (Per) / De-perimeterised (D-p) is a measure of whether the operation is inside or outside the security boundary or network firewall.
- **Sourcing:** Insourced or Outsourced means whether the service is provided by the customer or the service provider.

**FIGURE 1.2**

The Jericho Forum's Cloud Cube Model



Taken together, the fourth dimension corresponds to two different states in the eight possible cloud forms: Per (IP, IO, EP, EO) and D-p (IP, IO, EP, EO). The sourcing dimension addresses the deliverer of the service. What the Cloud Cube Model is meant to show is that the traditional notion of a network boundary being the network's firewall no longer applies in cloud computing.

## Deployment models

A deployment model defines the purpose of the cloud and the nature of how the cloud is located.

The NIST definition for the four deployment models is as follows:

- **Public cloud:** The public cloud infrastructure is available for public use alternatively for a large industry group and is owned by an organization selling cloud services.
- **Private cloud:** The private cloud infrastructure is operated for the exclusive use of an organization. The cloud may be managed by that organization or a third party. Private clouds may be either on- or off-premises.
- **Hybrid cloud:** A hybrid cloud combines multiple clouds (private, community of public) where those clouds retain their unique identities, but are bound together as a unit. A

## Part I: Examining the Value Proposition

---

hybrid cloud may offer standardized or proprietary access to data and applications, as well as application portability.

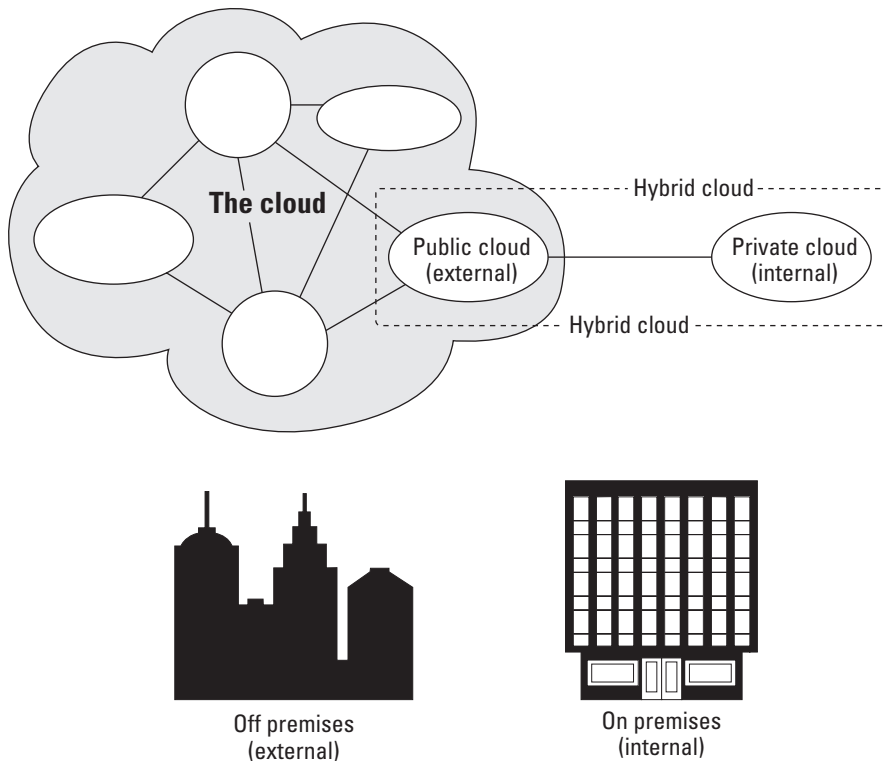
- **Community cloud:** A community cloud is one where the cloud has been organized to serve a common function or purpose.

It may be for one organization or for several organizations, but they share common concerns such as their mission, policies, security, regulatory compliance needs, and so on. A community cloud may be managed by the constituent organization(s) or by a third party.

Figure 1.3 shows the different locations that clouds can come in. In the sections that follow, these different cloud deployment models are described in more detail.

**FIGURE 1.3**

Deployment locations for different cloud types



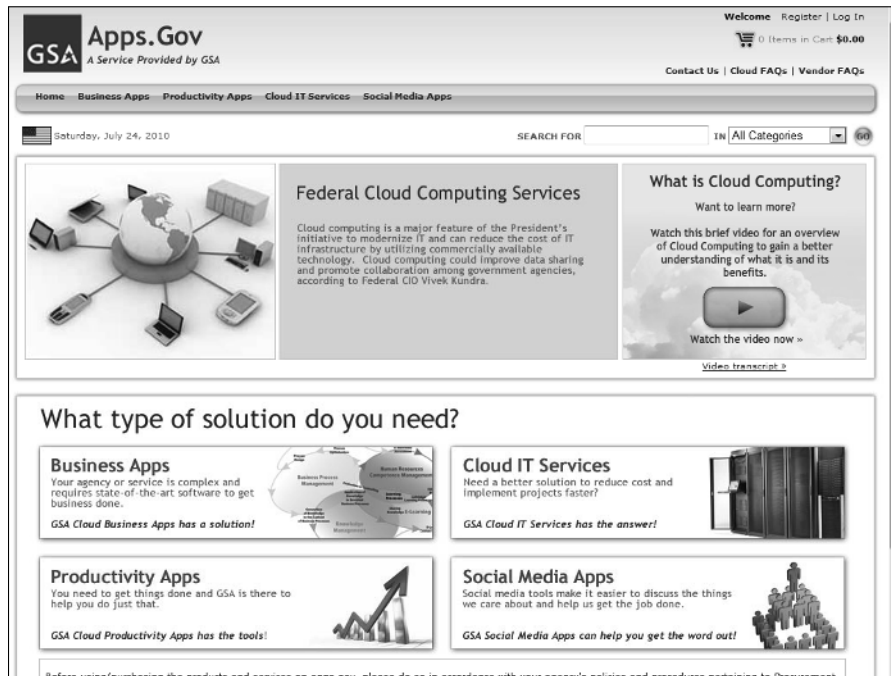
The United States Government, under the auspices of the General Services Administrator (GSA), launched a cloud computing portal called Apps.gov, as shown in Figure 1.4, with the purpose of providing cloud services to federal agencies. Described under the “U.S. Federal Cloud Computing



Initiative (<http://www.scribd.com/doc/17914883/US-Federal-Cloud-Computing-Initiative-RFQ-GSA>), the goal of the initiative is to make large portions of the federal government's apparatus available under a cloud computing model. This is a good example of a community cloud deployment, with the government being the community.

**FIGURE 1.4**

Apps.gov is the U.S. government's cloud computing system for its various agencies.



Apps.gov is also making available connections to free media services from its cloud, such as Twitter and YouTube. An example of this connection in practice is the YouTube channel created by the White House for citizens' outreach. You can find the White House channel at <http://www.youtube.com/whitehouse> and the general U.S. Government YouTube channel at <http://www.youtube.com/usgovernment>. You can see YouTube in action when you visit WhiteHouse.gov and click the video link that usually appears on that home page.

## Service models

In the deployment model, different cloud types are an expression of the manner in which infrastructure is deployed. You can think of the cloud as the boundary between where a client's network, management, and responsibilities ends and the cloud service provider's begins. As cloud

## Part I: Examining the Value Proposition

---

computing has developed, different vendors offer clouds that have different services associated with them. The portfolio of services offered adds another set of definitions called the service model.

There are many different service models described in the literature, all of which take the following form:

XaaS, or “<Something> as a Service”

Three service types have been universally accepted:

- **Infrastructure as a Service:** IaaS provides virtual machines, virtual storage, virtual infrastructure, and other hardware assets as resources that clients can provision.

The IaaS service provider manages all the infrastructure, while the client is responsible for all other aspects of the deployment. This can include the operating system, applications, and user interactions with the system.

- **Platform as a Service:** PaaS provides virtual machines, operating systems, applications, services, development frameworks, transactions, and control structures.

The client can deploy its applications on the cloud infrastructure or use applications that were programmed using languages and tools that are supported by the PaaS service provider. The service provider manages the cloud infrastructure, the operating systems, and the enabling software. The client is responsible for installing and managing the application that it is deploying.

- **Software as a Service:** SaaS is a complete operating environment with applications, management, and the user interface.

In the SaaS model, the application is provided to the client through a thin client interface (a browser, usually), and the customer’s responsibility begins and ends with entering and managing its data and user interaction. Everything from the application down to the infrastructure is the vendor’s responsibility.

The three different service models taken together have come to be known as the SPI model of cloud computing. Many other service models have been mentioned: StaaS, Storage as a Service; IdaaS, Identity as a Service; CmaaS, Compliance as a Service; and so forth. However, the SPI services encompass all the other possibilities.

It is useful to think of cloud computing’s service models in terms of a hardware/software stack. One such representation called the Cloud Reference Model is shown in Figure 1.5. At the bottom of the stack is the hardware or infrastructure that comprises the network. As you move upward in the stack, each service model inherits the capabilities of the service model beneath it. IaaS has the least levels of integrated functionality and the lowest levels of integration, and SaaS has the most.

Examples of IaaS service providers include:

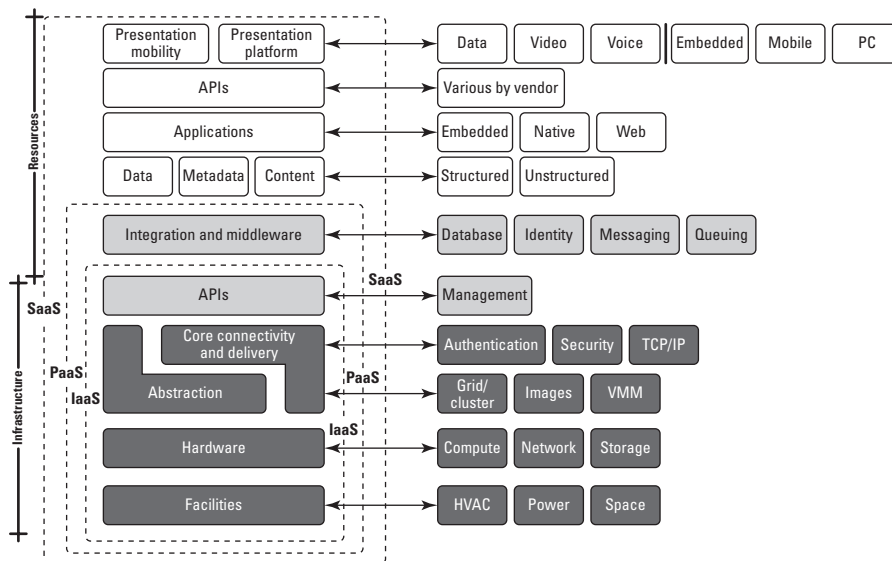
- Amazon Elastic Compute Cloud (EC2)
- Eucalyptus
- GoGrid

- FlexiScale
- Linode
- RackSpace Cloud
- Terremark

All these vendors offer direct access to hardware resources. On Amazon EC2, considered the classic IaaS example, a client would provision a computer in the form of a virtual machine image, provision storage, and then go on to install the operating system and applications onto that virtual system. Amazon has a number of operating systems and some enterprise applications that they offer on a rental basis to customers in the form of a number of canned images, but customers are free to install whatever software they want to run. Amazon's responsibilities as expressed in its Service Level Agreement, which is published on Amazon's Web site, contractually obligates Amazon to provide a level of performance commensurate with the type of resource chosen, as well as a certain level of reliability as measured by the system's uptime.

**FIGURE 1.5**

The Cloud Reference Model



A PaaS service adds integration features, middleware, and other orchestration and choreography services to the IaaS model. Examples of PaaS services are:

- Force.com
- GoGrid CloudCenter

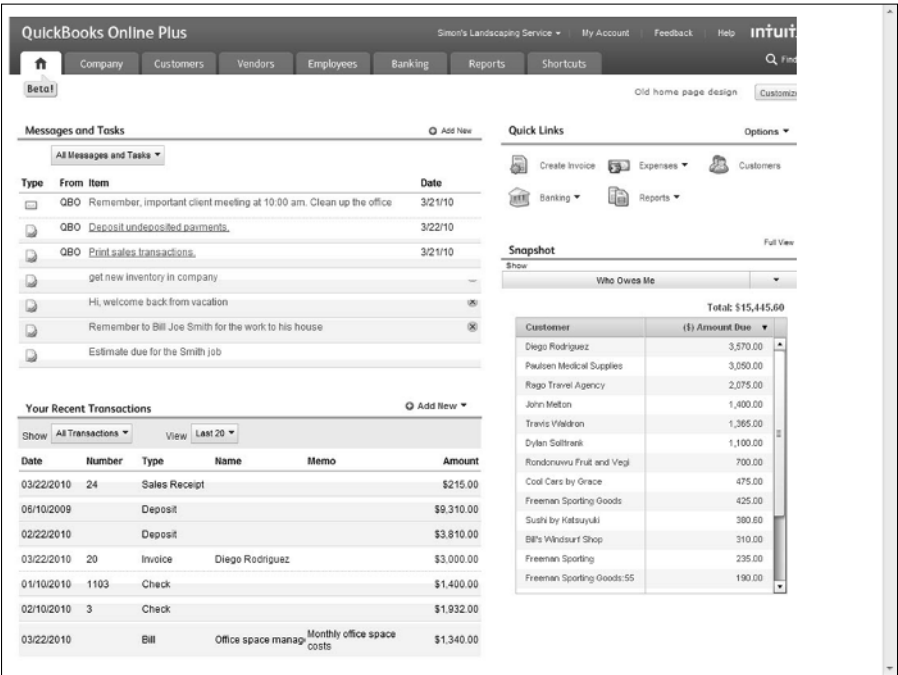
# Part I: Examining the Value Proposition

- Google AppEngine
- Windows Azure Platform

When a cloud computing vendor offers software running in the cloud with use of the application on a pay-as-you-go model, it is referred to as SaaS. With SaaS, the customer uses the application as needed and is not responsible for the installation of the application, its maintenance, or its upkeep. A good example of an SaaS offering is an online accounting package, with the online versions of Quicken and Quickbooks a prime example. Figure 1.6 shows a home page for QuickBooks Online plus on the Intuit.com Web site.

**FIGURE 1.6**

A home page for a Quickbooks customer on the Intuit.com Web site is an example of an SaaS service.



A client using an SaaS service might—as is the case for Quickbooks online—log into the service from his browser, create an account, and enter data into the system. Intuit.com has a service agreement that not only covers the performance of the hardware and software, but extends to protecting the data that they store for clients, and other fundamental characteristics.

Other good examples of SaaS cloud service providers are:

- GoogleApps
- Oracle On Demand
- Salesforce.com
- SQL Azure

These service model classifications start to get confusing rather quickly when you have a cloud service provider that starts out offering services in one area and then develops services that are classified as another type. For example, Salesforce.com started out as a Customer Relationship Management SaaS platform that allowed clients to add their own applications. Over time Salesforce.com opened an API called the Force API that allowed developers to create applications based on the Salesforce.com technologies. Force.com is thus their PaaS service.

As another example, take the PaaS offering that is the Windows Azure Platform. Windows Azure Platform allows .NET developers to stage their applications on top of Microsoft's infrastructure so that any application built with the .NET Framework can live locally, in Microsoft's cloud network, or some combination thereof. As Microsoft adds enterprise applications to its cloud service portfolio, as it has in the case of SQL Azure (and many other enterprise applications to come), these offerings fall under the rubric of being an SaaS service model.

Because a discussion of service models forms the basis for Chapter 4, I refer you to that chapter for a more in-depth discussion of this topic.

## Examining the Characteristics of Cloud Computing

---

Cloud computing builds on so many older concepts in computer technology that it can be hard for people newly introduced to the concept to grasp that it represents a paradigm shift in computing. It's an evolutionary change that enables a revolutionary new approach to how computing services are produced and consumed.

### Paradigm shift

When you choose a cloud service provider, you are renting or leasing part of an enormous infrastructure of datacenters, computers, storage, and networking capacity. Many of these datacenters are multi-million-dollar investments by the companies that run them. To give you some sense of scale, it has been estimated that a state-of-the-art microchip fabrication facility can cost anywhere from \$2 to \$5 billion. By comparison, a state of the art cloud computing datacenter can run in the range of \$100 million. Most of the large cloud computing service providers have multiple datacenters located all over the world. An accurate count can be difficult to obtain, but in Chapter 9

## Part I: Examining the Value Proposition

---

the location of some 20 datacenters in Amazon Web Service's cloud are detailed. Google's cloud includes perhaps some 35 datacenters worldwide.

In the 1960s, military initiative aimed at miniaturizing electronics funded many of the semiconductor production lines that led to advanced microprocessors, dense memory arrays, and the sophisticated integrated circuit technology that makes computers, mobile devices, and so much more possible today. In the 1990s, the commercialization of the Internet gave rise to some very large companies that were forced to build very large computing infrastructures to support their businesses.

Amazon.com's infrastructure was built to support elastic demand so the system could accommodate peak traffic on a busy shopping day such as "Black Monday." Because much of the capacity was idle, Amazon.com first opened its network to partners and then as Amazon Web Services to customers.

Google's business has also grown exponentially and required the building of datacenters worldwide. One of its datacenters in Dalles, Oregon, built in 2006 on the banks of the Columbia River, is shown in Figure 1.7. It is the size of an American football field.

**FIGURE 1.7**

The Google Dalles, Oregon, datacenter shown in Google Earth is an industrial-sized information technology utility.



As these various datacenters grew in size, businesses have developed their datacenters as “green-field” projects. Datacenters have been sited to do the following:

- Have access to low cost power
- Leverage renewable power source
- Be near abundant water
- Be sited where high-speed network backbone connections can be made
- Keep land costs modest and occupation unobtrusive
- Obtain tax breaks
- Optimize the overall system latency

These characteristics make cloud computing networks highly efficient and capture enough margin to make utility computing profitable.

It has been estimated that the Internet consumes roughly 10 percent of the world’s total power, so these companies are very big energy consumers. In some cases, such as Google, these companies may also become some of the major energy producers of the 21st century. Essentially what has happened is that the Internet has funded the creation of the first information technology utilities. That’s why cloud computing is such a big deal.

According to the research firm IDC, the following areas were the top five cloud applications in use in 2010:

- Collaboration applications
- Web applications/Web serving
- Cloud backup
- Business applications
- Personal productivity applications

The last five years have seen a proliferation of services and productivity applications delivered online as cloud computing applications. Examples of the impact of cloud computing abound in your everyday life, although many people do not make the connection to what was once a straightforward client/server Internet deployment. Movement of these applications to the cloud has been transparent, and in many cases the older on-premises deployment is supported by the same applications hosted in the cloud.

For example, many people have used ChannelAdvisor.com for their auction listings and sales management. That site recently expanded its service to include a CRM connector to Salesforce.com. One of the largest call center operations companies is a cloud-based service, Liveops.com. Figure 1.8 shows the Liveops home page.

Cloud computing has shifted the economics of software delivery in a manner similar to the way that music downloads have shifted the delivery of commercial music. The cost advantages of cloud

## Part I: Examining the Value Proposition

computing have enabled new software vendors to create productivity applications that they can make available to people at a much smaller cost than would be possible for shrink-wrapped software. Given the general demise of the big-box computer store along with many other traditional retail models, it has become increasingly difficult for vendors to get shelf space. You can visit your local Wal-Mart to get some sense of this issue.

In Chapter 16, “Working with Productivity Software,” some of these applications are described. This new model of computer application delivery has allowed vendors like Google to offer complete office suites to individuals for free, supported by its advertiser subscription model. Even Google’s business offerings have had some major successes against industry leader Microsoft Office. Last year, Los Angeles County switched to Google Docs.

**FIGURE 1.8**

Liveops.com is a cloud computing call center service.



## Benefits of cloud computing

“The NIST Definition of Cloud Computing” by Peter Mell and Tim Grance (version 14, 10/7/2009) described previously in this chapter (refer to Figure 1.1) that classified cloud computing into the three SPI service models (SaaS, IaaS, and PaaS) and four cloud types (public, private, community, and hybrid), also assigns five essential characteristics that cloud computing systems must offer:



- **On-demand self-service:** A client can provision computer resources without the need for interaction with cloud service provider personnel.

- **Broad network access:** Access to resources in the cloud is available over the network using standard methods in a manner that provides platform-independent access to clients of all types.

This includes a mixture of heterogeneous operating systems, and thick and thin platforms such as laptops, mobile phones, and PDA.

- **Resource pooling:** A cloud service provider creates resources that are pooled together in a system that supports multi-tenant usage.

Physical and virtual systems are dynamically allocated or reallocated as needed. Intrinsic in this concept of pooling is the idea of abstraction that hides the location of resources such as virtual machines, processing, memory, storage, and network bandwidth and connectivity.

- **Rapid elasticity:** Resources can be rapidly and elastically provisioned.

The system can add resources by either scaling up systems (more powerful computers) or scaling out systems (more computers of the same kind), and scaling may be automatic or manual. From the standpoint of the client, cloud computing resources should look limitless and can be purchased at any time and in any quantity.

- **Measured service:** The use of cloud system resources is measured, audited, and reported to the customer based on a metered system.

A client can be charged based on a known metric such as amount of storage used, number of transactions, network I/O (Input/Output) or bandwidth, amount of processing power used, and so forth. A client is charged based on the level of services provided.

While these five core features of cloud computing are on almost anybody's list, you also should consider these additional advantages:

- **Lower costs:** Because cloud networks operate at higher efficiencies and with greater utilization, significant cost reductions are often encountered.
- **Ease of utilization:** Depending upon the type of service being offered, you may find that you do not require hardware or software licenses to implement your service.
- **Quality of Service:** The Quality of Service (QoS) is something that you can obtain under contract from your vendor.
- **Reliability:** The scale of cloud computing networks and their ability to provide load balancing and failover makes them highly reliable, often much more reliable than what you can achieve in a single organization.
- **Outsourced IT management:** A cloud computing deployment lets someone else manage your computing infrastructure while you manage your business. In most instances, you achieve considerable reductions in IT staffing costs.

## Part I: Examining the Value Proposition

---

- **Simplified maintenance and upgrade:** Because the system is centralized, you can easily apply patches and upgrades. This means your users always have access to the latest software versions.
- **Low Barrier to Entry:** In particular, upfront capital expenditures are dramatically reduced. In cloud computing, anyone can be a giant at any time.

This very long list of benefits should make it obvious why so many people are excited about the idea of cloud computing. Cloud computing is not a panacea, however. In many instances, cloud computing doesn't work well for particular applications.

### Disadvantages of cloud computing

While the benefits of cloud computing are myriad, the disadvantages are just as numerous. As a general rule, the advantages of cloud computing present a more compelling case for small organizations than for larger ones. Larger organizations can support IT staff and development efforts that put in place custom software solutions that are crafted with their particular needs in mind.

When you use an application or service in the cloud, you are using something that isn't necessarily as customizable as you might want. Additionally, although many cloud computing applications are very capable, applications deployed on-premises still have many more features than their cloud counterparts.

All cloud computing applications suffer from the inherent latency that is intrinsic in their WAN connectivity. While cloud computing applications excel at large-scale processing tasks, if your application needs large amounts of data transfer, cloud computing may not be the best model for you.

Additionally, cloud computing is a stateless system, as is the Internet in general. In order for communication to survive on a distributed system, it is necessarily unidirectional in nature. All the requests you use in HTTP: PUTs, GETs, and so on are requests to a service provider. The service provider then sends a response. Although it may seem that you are carrying on a conversation between client and provider, there is an architectural disconnect between the two. That lack of state allows messages to travel over different routes and for data to arrive out of sequence, and many other characteristics allow the communication to succeed even when the medium is faulty. Therefore, to impose transactional coherency upon the system, additional overhead in the form of service brokers, transaction managers, and other middleware must be added to the system. This can introduce a very large performance hit into some applications.

If you had to pick a single area of concern in cloud computing, that area would undoubtedly be privacy and security. When your data travels over and rests on systems that are no longer under your control, you have increased risk due to the interception and malfeasance of others. You can't count on a cloud provider maintaining your privacy in the face of government actions.

In the United States, an example is the National Security Agency's program that ran millions of phone calls from AT&T and Verizon through a data analyzer to extract the phone calls that matched its security criteria. VoIP is one of the services that is heavily deployed on cloud computing systems. Another example is the case of Google's service in China, which had been subject to a

filter that removed content to which the Chinese government objected. After five years of operation, and after Google detected that Chinese hackers were accessing Gmail accounts of Chinese citizens, Google moved their servers for Google.ch to Hong Kong.

So while the cloud computing industry continues to address security concerns, if you have an application that works with sensitive data, you need to be particularly aware of the issues involved. Chapter 12, “Understanding Cloud Security,” expands upon these points in more detail.

These days most organizations are faced with regulatory compliance issues of various kinds. In the United States, companies must comply with the accounting requirements of the Sarbanes-Oxley Act; health care providers comply with the data privacy rules of HIPAA, and so on. In Europe, the European Common Market has a raft of its own legislation for companies to deal with. Rules apply to data at rest, and different rules may apply to data in transit. If you stage your cloud computing deployment across states and countries, the bad news is that you may end up having to comply with multiple jurisdictions. Don’t expect much support from the cloud system provider or from the governments involved. The laws of most regulatory agencies place the entire burden on the client. So when it comes to compliance, cloud computing is still the “Wild West” of computing.

## Assessing the Role of Open Standards

---

When you consider the development of cloud computing to date, it is clear that the technology is the result of the convergence of many different standards. Cloud computing’s promise of scalability completely changes the manner in which services and applications are deployed. Without standards, the industry creates proprietary systems with vendor lock-in, sometimes referred to as “stovepipe” clouds. Because clients do not want to be locked into any single system, there is a strong industry push to create standards-based clouds.

The cloud computing industry is working with these architectural standards:

- Platform virtualization of resources
- Service-oriented architecture
- Web-application frameworks
- Deployment of open-source software
- Standardized Web services
- Autonomic systems
- Grid computing

These standards help to enable different business models that cloud computing vendors can support, most notably Software as a Service (SaaS), Web 2.0 applications, and utility computing. These businesses require open standards so that data is both portable and universally accessible.

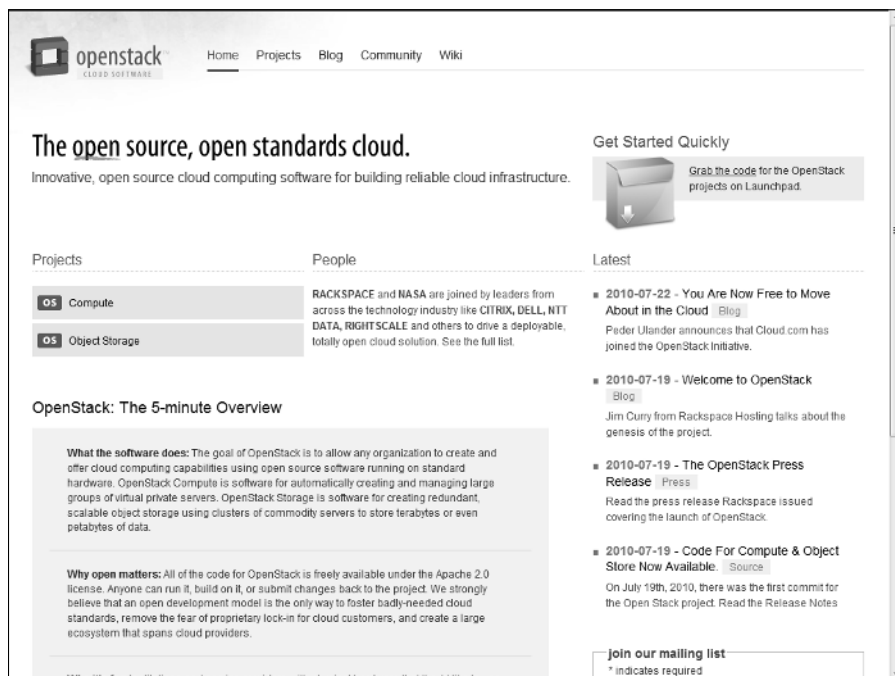
## Part I: Examining the Value Proposition

The race to create the first generation of open cloud platform technologies that will compete with proprietary technologies offered by companies such as Microsoft (Azure Platform) and VMware (vSphere) is already underway. Rackspace.com, one of the large IaaS cloud service providers, announced in July 2010 that it is initiating an open-source project called OpenStack that will begin with the code used to run its Cloud Files and Cloud Servers technologies. NASA has also donated some of the Nebula Cloud Platform technology that it developed. The software developed will be released under the Apache 2.0 license. Founding members of this project include AMD, Citrix, Dell, Intel, NTT Data, and several other cloud service providers. OpenStack.org's home page (<http://www.openstack.org/>) is shown in Figure 1.9.

The first two deliverables of the project are a distributed object store based on Rackspace Cloud Files and a scalable machine provisioning technology based on NASA Nebula and Rackspace Cloud Servers. OpenStack Compute software will automatically create large groups of virtual private servers on industry-standard systems. OpenStack Storage is the software that will create redundant object-based storage using clusters of commodity servers and storage systems.

**FIGURE 1.9**

OpenStack.org is an industry group seeking to create open cloud standards based on Rackspace.com and NASA technologies.



Eucalyptus (<http://open.eucalyptus.com/>) is a Linux-based software platform for creating cloud computing IaaS systems based on computer clusters. The project has an interface that can connect to Amazon's compute and storage cloud systems (EC2 and S3), and it maintains a private cloud as a sandbox for developers to work in. Eucalyptus works with a number of technologies for system virtualization, including VMware, Xen, and KVM. Eucalyptus is an acronym taken from the expression "Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems." Most of the major Linux vendors support this project, which is based on the original work of Rich Wolski at the University of California at Santa Barbara. The company Eucalyptus Systems was formed in 2009 to support the commercialization of the Eucalyptus Cloud Computing Platform.

OpenStack and Eucalyptus are by no means unique; several other projects are underway to create open-source cloud platforms. There also are numerous research projects in the area. The IEEE Technical Committee on Services Computing (<http://tab.computer.org/tcsc/>) sponsors a conference in this area called CLOUD and has some working groups and publications in this area. Figure 1.10 shows the home page of TCSC.

**FIGURE 1.10**

The home page of the IEEE Technical Committee on Services Computing (<http://tab.computer.org/tcsc/>), which works with the cloud computing community to create standards



### Summary

---

In this chapter, you learned what is meant by the term “cloud computing.” Cloud computing defines systems that are virtualized and where resources are pooled so that customers can provision computing resources as needed.

To understand cloud computing and fully appreciate its subtleties, you need to categorize the different cloud types. The two types of cloud models are those based on where the cloud is deployed and those based on the types of services that clouds offer. In this chapter, you were introduced to these cloud types.

Cloud computing has a number of benefits and has attracted great industry and general interest. Cloud computing allows systems to be created cheaply with little upfront costs and to be scaled to massive sizes, when needed. Not all applications and services benefit from cloud computing, and I presented some of the factors that help you differentiate between successful deployments and those that are not.

In Chapter 2, “Assessing the Value Proposition,” the economics of cloud computing are discussed.