I. Introduction

Since its introduction, cloud computing has rapidly grown to become an important component of the information technologies that businesses use, and although spending on cloud computing constitutes less than five percent of overall spending on information technology, spending on cloud computing has grown at 4.5 times the rate of general information technology spending since 2009, a figure that is projected to rise by a multiple of six in the near future. A 2016 survey of chief information officers revealed that 16 percent of data processing workloads run in the public cloud, which will grow to 41 percent by 2021. Of the Fortune Global 50 companies, 48 have announced plans to adopt cloud computing. In other words, the importance of cloud computing to businesses is likely to continue growing.

The objective of this chapter is to provide a brief overview of the technology that enables cloud computing, the general characteristics and types of cloud computing, the models that can be used to deploy them, and some advantages and disadvantages of cloud computing.

II. Technical Elements of Cloud Computing

An oft-cited definition of cloud computing is one developed by the National Institute of Standards and Technology (NIST): "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and..."
Cloud 3.0: Drafting and Negotiating Cloud Computing Agreements

released with minimal management effort or service provider interaction. Although the NIST Definition is helpful in describing the characteristics of cloud computing, it is not quite as helpful in explaining exactly what the cloud is.

Perhaps another useful starting point is the history of the term “cloud.” Well before the advent of cloud computing, computer network designers used a pictogram of a cloud in networking diagrams to identify a group of networked computers connected to the network depicted. The purpose of using the pictogram was to keep the diagram simple—the cloud was intended to depict the abstract notion of computing resources that are linked to a network where the details were either unknown or not particularly important. The notion of abstraction is particularly appropriate when it comes to cloud computing. At the conceptual level, cloud computing can perhaps best be thought of as the abstraction of computing resources. The following sections provide further detail on how that abstraction is achieved.

A. Hardware

Cloud computing has always been (and likely always will be) powered by computers and related equipment in some shape or form. For cloud computing, this will typically consist of hundreds or thousands of computers densely packed into data centers, networked together along with storage devices with connectivity to the internet.

B. Network Connectivity

It should come as no surprise that one of the enabling technologies for cloud computing is network connectivity. Although cloud computing services can in theory be provided through various types of network connections, most if not all current cloud computing services are delivered through the internet. As cloud computing often requires the rapid transmission of significant amounts of data between the cloud computing service provider and its users, the rise of broadband and mobile internet connectivity over the past decade has been instrumental in enabling the corresponding rise in cloud computing.

---


8 Readers should be cautioned, however, that “typically” does not equate to “necessarily”—the term “cloud” can be used generically to describe any remote computing resources, including far less robust arrangements. Thus, when prospective service providers indicate that they utilize “the cloud,” users would be well advised to enquire exactly what that entails.

C. Communication Standards

The development and rapid adoption of web, internet, and related technology standards has also played an important role in enabling the rise of cloud computing by setting out standardized formats and protocols for receiving, transmitting, and displaying data and information. Service providers can conform to these uniform standards, knowing that a broad base of users can readily access them, including through browsers that are already installed on their computers, obviating the need to install specialized software.

A subset of such standards that is particularly important for cloud computing are the service standards and protocols that are used to enable applications on two different computers to communicate and interact with each other.10

D. Virtualization

There is of course more to cloud computing than just computers. One of the key enabling technologies for cloud computing is virtualization technology. This technology enables simulation of the hardware of an actual computer in software, commonly referred to as a “virtual machine.” A somewhat rough analogy11 is a flight simulator game: the game simulates and presents the experience of flying an airplane, including responding and reacting to inputs in a manner similar to a real airplane. Virtualization software (sometimes referred to as a “hypervisor”) does something similar, except that instead of an airplane, it simulates (or “virtualizes”) a computer and does so very accurately—so accurately that almost any software that can be installed on a physical computer can also be installed and operated within that virtual machine. Virtual machines mimic, in software, all the elements of a physical computer, including a central processing unit, memory, storage, and networking interfaces, subject, of course, to the limitations of the physical computer (the “host”) on which the hypervisor is installed. As a result, one can create a virtual machine running the Linux operating system, for example, even though the underlying host is running Windows.12 One can even install multiple virtual machines on a single host, each running a different operating system (again, subject to the limits of the physical computer). Moreover, any software running within a virtual machine is isolated within

10 These include technologies such as Simple Object Access Protocol, representational state transfer, Web Service Description Language, and Universal Description, Discovery, and Integration, among others.
11 It should perhaps be emphasized that this is a conceptual rather than a technical analogy: the design and architecture of a flight simulator bears little to no resemblance to virtualization software. That being said, one of the earliest implementations of virtual machines on personal computers was developed specifically for a text-based adventure game called Zork, to reduce its size so that it could be run on personal computers (as it had originally been developed on a much larger mainframe computer), and to enable portability across different hardware platforms. See P. David Lebling, Zork and the Future of Computerized Fantasy Simulations, 5 BYTE 12, 172 (Dec. 1980), https://archive.org/details/byte-magazine-1980-12.
12 There are a number of software packages that enable one to experience virtualization firsthand on standard desktop computers, such as VirtualBox, https://www.virtualbox.org/ (free and available for Windows, Linux and Mac) or Parallels Desktop, https://www.parallels.com/ca/products/desktop (paid, for Mac only). Microsoft Hyper-V can also be enabled on certain versions of Windows 8 and 10. Anthony Bartolo, Step-By-Step: Enabling Hyper-V for Use on Windows 10, CANITPRO.NET, Sept. 8, 2015, https://blogs.technet.microsoft.com/canitpro/2015/09/08/step-by-step-enabling-hyper-v-for-use-on-windows-10/.
that virtual machine and cannot communicate with or affect either the physical computer
or other virtual machines installed on that physical computer.13

Because virtual machines are created entirely in software, they provide all of the ben-
fits of software, many of which are key to enabling cloud computing. These include:

1) **Portability and replicability.** Given that a virtual machine is created in software,
it can be easily saved as a file and copied or moved to and operated on another
host with virtualization software installed, without the need to reinstall and re-
configure the operating system and all applications installed on the virtual ma-
chine, making it much faster and easier to move to another host, or to provision
a new server by simply copying an existing virtual machine that has already been
configured.

2) **Reliability and availability.** Some hypervisors can be configured to automatically
move a virtual machine to another host under certain conditions (for example, if
it detects a possible failure of the host)14 or to maintain an identical running copy
of the virtual machine on another host and automatically switch to that copy, in
both cases enhancing availability and reliability, and enabling easier disaster re-
cover.

3) **Greater hardware flexibility.** Given that virtualization enables virtual machines
to have characteristics independent of the hardware on which they run, there is
typically more flexibility when selecting and purchasing physical computers so
long as they are compatible with the applicable virtualization software.

4) **Recovery.** Given that multiple images of virtual machines can be saved at any
time, periodic “snapshots” can be created. If a virtual machine fails for any reason
(for example, after a failed upgrade or patch to the operating system), it can sim-
ply be “rolled back” to an earlier snapshot that was working properly.

5) **Security and stability.** Given that virtual machines are isolated from both each
other and the host computer, they can be used to enhance security, for example
by creating a separate virtual machine to run test environments for new or exper-
imental applications so that failures or crashes of such applications will impact
only that separate virtual machine.15

6) **Efficiency.** Physical servers are often underutilized. Virtualization can enable
greater efficiency by replacing physical computers with virtual machines and con-
solidating those virtual machines onto fewer physical computers, reducing both
hardware and operating costs.

---

13 This is to some extent an overgeneralization. A hypervisor can be configured to enable interaction
between a given virtual machine and its host or other virtual machines by, for example, creating virtual
network connections between them. Certain types of virtualization technology also allow for (or require)
some limited awareness of the underlying physical host within the guest operating system. In addition,
vulnerabilities occasionally are discovered that could be used to access the host system from within
a virtual machine, typically referred to as virtual machine escape. See, e.g., Dan Goodin, *Extremely
arstechnica.com/information-technology/2015/05/extremely-serious-virtual-machine-bug-
threatens-cloud-providers-everywhere/.

14 See also Part II.E, infra.

15 Security is thereby enhanced as compared to running such applications on the same computer with-
out the use of virtualization. Security would not be improved as compared to running such applications
on a separate physical computer, although of course doing so would entail additional cost.
One notable variation of virtualization is containerization. In contrast to “full” virtualization, which runs a completely separate operating system within a virtual machine, a container utilizes some elements of the operating system installed on the host, but still remains isolated from the host. Containers share some traits with full virtualization, such as ease of porting or replicating. However, containers typically require significantly less overhead than full virtualization due to their reliance on elements of the underlying operating system, resulting in lower resource usage and faster creation and execution. Thus, a given physical computer can host a much larger number of containers than virtual machines. Lastly, some container technologies allow for portability across different hosts.

E. Multitenant Technology

“Multitenant technology” is a term used to describe an approach to application development that enables multiple users (or multiple groups of users) to each access the same instance of an application in such a way that the application appears to be customized for each such user and logically isolated from every other user.

Although multitenant architecture can achieve the same result as virtualization, the two technologies are distinct from one other. For example, a service provider could service multiple users by creating a distinct virtual machine for each such user and installing the same non-multitenant application in each virtual machine. Given that virtual machines are isolated from one another, each user would have its own uniquely configured application. In contrast, if the same application were developed using multitenant architecture, each user would access a single instance of the application. The data and configuration stored for each user would be distinct, however, and accessible only by that user through the use of logical controls within the application. Multitenant architecture can also be combined with virtualization. For example, a single instance of a multitenant application could be installed and operated within a virtual machine to serve multiple users. Alternatively, there could be multiple instances in multiple virtual machines, with each instance serving multiple users.

Although multitenant architecture is typically much more complex (and therefore more costly) to develop, it can result in cost savings as it enables resources to be shared more efficiently across users, as compared to creating an individual virtual machine and individual instance of the application for each user. However, it may reduce the flexibility afforded to end-users. For example, given that all users of a multitenant application are accessing the same instance of that application, they typically must all use the same version of the application.

---

16 This is sometimes called container-based virtualization or operating system level virtualization.
17 This was perhaps the primary benefit of Docker (a particular type of container technology; see What Is Docker, Docker, https://www.docker.com/what-docker) as compared to other container technologies, such as Linux containers, which are to some degree dependent on the configuration of the underlying host. See Chenxi Wang, What is Docker? Linux Containers Explained, InfoWorld, June 29, 2017, https://www.infoworld.com/article/3204171/linux/what-is-docker-linux-containers-explained.html.
18 In this context, a “user” could be an organization with multiple individual users, all of whom could access the data and configuration for that organization.
F. Resource-Pooling, Allocation, and Management Technologies

Another group of technologies that facilitate cloud computing are those that are used to pool, allocate, and manage the underlying physical resources, such as processing power, network communications, and data storage for use by individual users. Such physical resources are typically comprised of hundreds or thousands (or more) computers located in one or multiple data centers, each connected through networking equipment.

One example of such technology is virtual infrastructure managers (VIMs). VIMs are used to manage virtual machines across physical servers. For example, assume that a (quite small) cloud service provider uses two computers. Each has a hypervisor installed on it, which in turn hosts multiple virtual machines. If one computer is nearing the limits of its capacity and a request for a new virtual machine comes in, a VIM can be used to allocate that virtual machine to the second computer. VIMs can also be used to move virtual machines from one physical computer to another, in some cases even while the virtual machine continues to operate.

Along similar lines, various types of cloud storage technologies can be used to allocate, move, and manage storage capacity among multiple virtual machines. Although there are various ways of implementing this approach, such technologies essentially virtualize storage. For example, if a user creates a “virtual drive” within a virtual machine and writes data to it, storage technology will take that data, break it up into chunks, and write it to multiple physical storage devices, such as hard drives, which can be on multiple computers. Even though the chunks are spread across multiple computers and multiple drives, the storage technology keeps track of exactly where each chunk is so that, to an end user, those disparate chunks appear to be on one drive. In addition to keeping track of those chunks, such technologies will also either record multiple copies of the data, or apply an algorithm that generates additional data so that if one or more chunks are lost (for example, due to the failure of a physical hard drive), the data can still be recovered.

Along similar lines, functions that would otherwise be served by physical networking devices, such as firewalls, routers, and switches, can be replicated in software.

20 Traditional approaches include, for example, storage area networks—typically proprietary hardware consisting of one or more devices, each containing numerous drives. More recently, software solutions have been developed to implement the same (or better) functionality using generic, nonvendor-specific hardware, some of which is open source. These include, for example, Ceph (http://ceph.com/) and Gluster (https://www.gluster.org/), both open source, as well as proprietary technologies such as VMware vSAN (https://www.vmware.com/products/vsan.html) and a subset of the functionality in Nutanix (https://www.nutanix.com/).

21 Examples of such technologies include virtual switch technology built directly within hypervisors to enable virtual machines on a given host to communicate between each other and to the host machine and elsewhere (for example, virtual switches in VMware’s vSphere, https://www.vmware.com/support/ws55/doc/ws_net_component_vswitch.html, or libvirt virtual network switches in KVM, https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Virtualization_Administration_Guide/chap-Virtualization_Administration_Guide-Virtual_Networking.html, as well as software-defined networking technologies, which are typically (but not always) implemented in dedicated physical networking devices and shift control of certain networking functions (typically referred to as the control plane) so that they are programmatically controlled by software and, in some cases, by virtualizing specific network functions such as load balancers or firewalls. Example of the latter include OpenDaylight, https://www.opendaylight.org/, Open vSwitch, http://openvswitch.org/, and VMware NSX, https://www.vmware.com/ca/products/nsx.html.
Perhaps due in part to the burgeoning popularity of cloud computing, a number of groups and companies have created collections of software tools (or “stacks”) that can be used to comprehensively implement a cloud computing environment. These include, for example, OpenStack, Apache CloudStack, and vCloud Suite.

G. Recent Developments

As with most other types of information technology, cloud computing continues to develop and evolve at a rapid pace. The following are some examples of recent developments in cloud computing.

1) **Cloud-native applications.** The term “cloud native” is generally used to refer to applications that have been designed and written to run on cloud services, meaning that they can run on various cloud platforms, can be easily copied or moved, and are scalable. They typically will also take a “microservice” approach: instead of one large application performing many functions, those functions are divided into numerous individual applications (or microservices) that can interface with each other, making each easier to maintain and allowing for a more modular approach in that end users can combine microservices as appropriate for the task at hand. One of the primary enabling technologies for cloud-native applications is container technology. Conceptually, cloud-native applications can be thought of as abstraction at the application level in the same way that physical machines are abstracted into virtual machines.

2) **Serverless computing.** A somewhat related concept is that of “serverless computing” where the cloud provider assumes responsibility for dynamically providing whatever resources are required by the application, rather than the end user manually allocating additional resources as required. For example, with a traditional approach, a user would create an application, select an appropriately sized virtual machine, and run it, paying for the use of that virtual machine. If the workload grows, then the user would increase the size of that virtual machine. In contrast, if the same application were implemented using a serverless approach, the end user would pay only for the resources actually used by the application. If, for example, the application were not used for a given period of time, no resources would be used; therefore, no usage fees would be paid.

---

22 https://www.openstack.org/.
23 https://cloudstack.apache.org/.
26 For a more detailed description of microservices architecture, see MuleSoft, Microservices vs Monolithic Architecture, https://www.mulesoft.com/resources/api/microservices-vs-monolithic.
27 See Part II.D., supra.
28 Clark & Hinkle, supra note 25, at 7.
3) **Edge computing.** To date, most cloud computing has been enabled through the use of large, centralized data centers hosting thousands of computers. This has resulted in a significant number of users being rather distant from the physical location where processing actually occurs, causing small delays in the processing of information (sometimes referred to as latency). Although in many cases the degree of latency is not critical, for certain applications, such as autonomous vehicles and robotics, it may be. As such applications become more widely implemented, there will be an increasing demand for low-latency cloud computing, resulting in the increased use of decentralized data centers that are more geographically dispersed closer to end users and therefore lower latency.30 Interestingly, some cloud providers have developed solutions to “extend” their cloud services to operate on an end user’s equipment to achieve this objective.31

4) **Big computing.** Artificial intelligence and, in particular, deep learning typically require significant amounts of computing power, but for limited periods of time. As such technologies become more widely used, the demand for high-performance computing will increase. As a result, cloud providers have begun to develop “big compute” cloud services to enable end users to utilize deep learning without the need to procure a high-performance data center.32

5) **Decentralized cloud computing.** At present, Infrastructure as a Service (IaaS)33 cloud computing is generally dominated by a small number of large companies, each of which owns and operates multiple, large-scale data centers with millions of computers. This is perhaps to be expected, given the advantages afforded by economies of scale. However, there are still millions of computers and computing devices outside of those owned by cloud providers that, for the most part, are not used to their full capacity.34 Technologies to harness such unutilized capacity have been used for more than two decades with a fair degree of success, but until recently have largely been limited to volunteer efforts for scientific research.35 However, a

---

33 For an explanation of Infrastructure as a Service, see Part III.B., infra.
35 The first such effort was the Great Internet Mersenne Prime Search, https://www.mersenne.org/, launched in 1996 to (as the name suggests) search for Mersenne prime numbers, which takes a significant amount of computing power. The project has discovered 15 prime numbers to date. In 2017, slightly over 1.5 million CPUs participated in the project. Other similar projects include Distributed.net, http://www.distributed.net/Main_Page, which undertakes research projects primarily related to cryptography; SETI@home, https://setiathome.berkeley.edu/, which uses computing power to analyze radio telescope data for signs of extraterrestrial intelligence; and Folding@home, http://folding.stanford.edu/, which focuses on disease research by simulating the folding of proteins and other elements.
number of ventures have recently been launched to offer general IaaS cloud computing services powered by decentralized computing resources, using blockchain technology to assist in accounting for the usage and provisioning of resources. Such an approach has some possible advantages as compared with traditional cloud computing, such as lower cost and higher responsiveness, although it may too early to assess whether those advantages will be realized at any significant scale.

III. Characteristics and Models

While Part II above focused on the elements of technology that are used to create cloud computing solutions, Part III outlines the general characteristics of cloud computing from the perspective of the user, and describes how cloud computing services are characterized both in terms of the nature of the services being provided (the “service model”) and who is entitled to use such services (the “deployment model”).

A. Characteristics of Cloud Computing Services

Application of the technologies described above results in the availability of computing services that have a specific set of characteristics. The NIST Definition identifies five such characteristics:

1) **On-demand self-service.** End users can easily obtain the computer services through automated means (such as a website) and without human assistance.

2) **Broad network access.** Services are provided through standardized means (for example, the internet and web browsers) and can be accessed and used utilizing a variety of devices, such as mobile phones, tablets, or desktop computers.

3) **Resource-pooling.** Computing resources are pooled together and dynamically assigned (and reassigned) for use by multiple users, generally without knowledge or control on the part of the user regarding where those resources are located.

4) **Rapid elasticity.** The quantity of resources can easily and quickly (and in some cases automatically) increase or decrease on demand in almost any quantity and at any time.

5) **Measured service.** Billing is metered and delivered as a utility service.

B. Service Models

The NIST Definition also sets out three service models under which cloud services are categorized based on the nature of the services provided:

36 Examples include Blockchain-Based Decentralized Cloud Computing, iExec, http://iexec.ec/, and DFINITY, https://dfinity.network/, both of which offer (or plan to offer) general IaaS services. STORJ, https://storj.io/, focuses on distributed cloud storage.


38 Paraphrased from MELL & GRANCE, supra note 6, at 2.

39 Paraphrased from id. at 2–3.
1) **Software as a Service (SaaS).** The provision of application services running on a cloud infrastructure, typically accessed by end users through a web browser, without any ability on the part of the end user to manage or control the infrastructure used to provide such services. Examples include e-mail services (such as Google Gmail) and customer relationship management applications (such as Salesforce). SaaS generally comprises the largest proportion of spending on cloud computing.

2) **Platform as a Service (PaaS).** The provision of programming languages and tools to enable end users to create applications and deploy those applications on a cloud platform for use by the user or by others. As with SaaS, the end user does not manage or control the underlying infrastructure. Examples include Heroku and Google App Engine.

3) **Infrastructure as a Service (IaaS).** The provision of fundamental computing resources such as processing, storage, and networking to enable users to run any software they choose, including operating systems or applications, but with no control or management of the underlying infrastructure other than the specific computing resources provisioned. Examples include Amazon EC2 (computing), Amazon S3 (storage), and Microsoft Azure.

A given cloud service provider does not necessarily own or control all the resources used to provide its services. For example, a SaaS vendor may decide to implement its application using the services provided by a separate IaaS vendor so that the SaaS vendor can focus on its strengths (namely, the development of the application) rather than non-core competencies, such as building and managing data centers and physical servers. That IaaS vendor may, in turn, decide to use a third-party vendor to provide data centers and hardware.

---

40 The NIST Definition defines a cloud infrastructure as one that meets the five characteristics described in Part III.A., supra.
41 https://www.google.com/gmail/about/.
42 https://www.salesforce.com/.
43 SaaS represented approximately 73 percent of worldwide spending on public cloud computing in 2015. IaaS represented 16 percent, and PaaS represented 11 percent. See Gantz & Miller, supra note 2, at 4.
44 https://www.heroku.com/.
45 https://cloud.google.com/appengine/.
46 https://aws.amazon.com/ec2/.
47 https://aws.amazon.com/s3/.
48 https://azure.microsoft.com/. Both Amazon AWS and Microsoft Azure include many different offerings, some of which are IaaS in nature, whereas others are more properly described as PaaS or SaaS.
49 See, e.g., Barb Darrow, *Welcome to the Era of Great Data Center Consolidation*, FORTUNE, Feb. 15, 2017, http://fortune.com/2017/02/15/data-center-consolidation-cloud/. The article explains that software developers are increasingly relying on IaaS vendors to deliver the functionality of their software. Jeff Bezos, CEO of Amazon, has compared this to the evolution of the power industry: “You go back in time a hundred years, if you wanted to have electricity, you had to build your own little electric power plant, and a lot of factories did this. As soon as the electric power grid came online, they dumped their electric power generator, and they started buying power off the grid. It just makes more sense. And that’s what is starting to happen with infrastructure computing.” Brad Stone, *The Everything Store: Jeff Bezos and the Age of Amazon* (2013).
50 Most IaaS vendors, however, own and control their own physical infrastructure.
C. Deployment Models

The NIST Definition also sets out four different deployment models that describe who is permitted to use the infrastructure powering the cloud service:\textsuperscript{51}

1) **Private cloud.** The infrastructure can be used by only a single organization (which may be comprised of multiple end users). It may be owned, managed, or operated by the organization itself, a third party, or a combination thereof, and the physical infrastructure may be located on the premises of that organization or elsewhere. Private clouds are typically implemented where isolation of the physical resources is required—for example, to facilitate higher security or for regulatory compliance.\textsuperscript{52} The trade-off is that private clouds are more costly and less flexible compared to public clouds.\textsuperscript{53} The end user typically will pay for all the dedicated physical resources comprising the private cloud and, if the capacity of the private cloud is reached, must pay for additional physical resources to be installed and added to the cloud.\textsuperscript{54}

2) **Community cloud.** Similar to private cloud except that use is limited to a specific community whose members have shared concerns, interests, or needs.

3) **Public cloud.** A cloud service that is available for use by the general public.\textsuperscript{55} The infrastructure may be owned, managed, and operated by a business, governmental, or other organization (or a combination thereof). The physical infrastructure is typically located on the premises of the provider.

4) **Hybrid cloud.** A combination of two or more of the above, with technology that enables data and applications to be moved or shared between or among them. Although the NIST Definition only contemplates the use of this term to reference combinations of different cloud types, the term “hybrid” also is often used to describe offerings that enable on-premise computing resources to access cloud services, such as additional storage. End users can install either an appliance or software on premises, which will then enable them to access cloud storage as if it were an on-site physical storage device.\textsuperscript{56}

\textsuperscript{51} Paraphrased from MELL & GRANCE, supra note 6, at 3.

\textsuperscript{52} Some vendors use the term “private cloud” to mean something different. In some cases, offerings marketed as private clouds do, in fact, involve shared resources that are logically (but not physically) segregated. Care should be taken in investigating offerings described as private clouds to ensure the degree and nature of isolation meets the requirements of the end user.

\textsuperscript{53} For a useful summary of trade-offs between private and public clouds, see Comparing Private and Public Clouds, RACKSPACE, https://www.rackspace.com/library/cloud-computing-difference.

\textsuperscript{54} Some vendors have begun to offer utility pricing (i.e., pricing based on use) for private clouds—see, e.g., Arturo Suarez, Introducing Utility Pricing to Your Managed Private Cloud, UBUNTU, July 9, 2015, https://insights.ubuntu.com/2015/07/09/introducing-utility-pricing-to-your-managed-private-cloud/, but such pricing is only for management and operation services and does not include the underlying hardware, which the end user must purchase separately.

\textsuperscript{55} Availability to the general public does not necessarily mean that the public cloud services purchased by a given user are made available to the general public, or that they are made freely available to the general public.

Some cloud providers have developed slight variations on the above models. One example is a virtual private cloud, a pool of computing resources created within a public cloud infrastructure but subject to a higher degree of isolation than usual to ensure the pool is available to only the user creating the virtual private cloud.\(^57\)

Virtual private clouds should not be confused with virtual private servers. Despite the similar name and the use of some of the same underlying technology, virtual private servers are distinct from virtual private clouds, or cloud computing generally for that matter. A virtual private server is essentially a virtual machine created and hosted on a computer that is allocated for the exclusive use of a given user. In other words, the end user is purchasing a portion of the resources of a physical computer that has been divided up into multiple virtual machines. In contrast to cloud computing, there are typically limits to the scalability of a given virtual private server, and if there is a failure of the host computer, all virtual private servers on that computer will also fail.

Another variation of the above models is a multicloud approach, where the services of two or more cloud service providers are used in tandem. For example, an end user could configure a primary virtual server with one service provider and a backup server with another, and then configure them to allow the backup server to automatically take over in the event the primary virtual server fails. Some organizations have developed solutions that facilitate the implementation of and help manage multicloud approaches.\(^58\)

In addition, some organizations have begun working toward the development of cloud interoperability—an approach sometimes referred to as “intercloud.” In contrast to a multi-cloud approach, where the end user or its service provider assumes the responsibility for developing solutions to address the specific standards of each different cloud service vendor (which may change over time), an intercloud approach focuses on the development of standardized interfaces to be adopted by cloud service providers to enable interoperability among themselves, and to allow users to easily access multiple cloud service providers though that same interface.\(^59\) To this end, the Institute of Electrical and Electronics Engineers has begun developing standards to help facilitate the adoption of intercloud computing.\(^60\) However, to date, intercloud does not appear to have gained much traction with cloud providers.\(^61\)

---

\(^57\) This is typically accomplished through the use of a separate virtual network along with additional security measures. See, e.g., *Amazon Virtual Private Cloud*, Amazon, https://aws.amazon.com/vpc/.


\(^59\) *Cloud Strategy Partners, LLC, IEEE Intercloud Interoperability and Federation Framework 6–8*, https://cloudcomputing.ieee.org/images/files/education/studygroup/IEEE_Intercloud_Interoperability_and_Federation_Framework.pdf (suggesting that cloud computing interoperability is inevitable, following similar trends toward uniform standards in various other industries, such as electric power, financial markets, telephony, and the internet).


\(^61\) For example, Cisco shut down its intercloud offering in early 2017. Simon Sharwood, *Cisco to Kill Its Intercloud Public Cloud on March 31, 2017*, *Register*, Dec. 13, 2016, https://www.theregister.co.uk/2016/12/13/cisco_to_kill_its_intercloud_public_cloud_on_march_31st_2017/. Although one possible advantage of intercloud would be to enable smaller cloud providers to present themselves collectively as a much larger network, it is less clear why market-dominant cloud providers would be interested in intercloud. Some vendors, however, do continue to market intercloud connectivity. See, e.g., *Multi-Cloud Connectivity with RackConnect Global*, RACKCONNECT, https://www.rackspace.com/hybrid/rackconnect/global.
IV. Advantages of Cloud Computing

Although the technology that enables cloud computing is critical, business considerations are of equal importance. Innovative technology is unlikely to be adopted if doing so does not make business sense. This section outlines some of the business advantages of cloud computing.

A. Costs and Cash Flow

Many cloud providers, particularly those offering IaaS, price their services on a utility basis: there are no up-front costs, and fees are charged only as and when the services are used.\(^62\) This enables end users to more precisely align their expenses with their requirements.\(^63\) For example, a small business can begin with a small virtual machine provisioned at a low cost. As its business grows, it can increase the size of the virtual machine in small increments or purchase additional virtual machines fairly quickly and easily. If there is a downturn in business, the company can also quickly scale down its usage, and therefore its costs, accordingly. If there is a significant project that requires an unusually high level of computing resources, it can temporarily procure as much as it may require through the cloud provider. Given that costs incurred are based on the needs of the business at a given point in time, expenditures are more likely to align with revenue.

Compare this to the more traditional model in which an organization must purchase a server and license software at a significant up-front cost. Given that the quantity of resources purchased must be sufficient for peak usage, and given that those resources (depending on the business) may not be used on a 24/7 basis, the resources will typically be underutilized. As its business grows, it must buy a minimum of one additional server, even if it does not require all of the capacity of an additional server. It must also factor in room for growth because the process of purchasing, installing, and configuring an additional server takes substantial time. If there is a downturn in business, the company is unlikely to realize much value when it disposes of unneeded servers. Given that the initial capital outlays do not necessarily align with revenue, financing may also be required. On-premise hardware will also require other costs, such as supporting hardware (for example, uninterruptible power supplies), space, staff, electricity, and possibly support and maintenance contracts requiring long-term financial commitments.\(^64\)

B. Efficiency—Economies of Scale and Commodification

Of course, hardware must still be purchased to enable cloud computing, and although end users who purchase cloud services are relieved of those capital expenditures, cloud

---

\(^62\) This is less likely to be the case for SaaS providers. Although some SaaS providers will allow for a degree of flexibility in increasing or decreasing usage, there will typically be some limitations, such as a minimum subscription period or minimum usage requirements, and the offering will therefore not be as flexible as most public cloud IaaS offerings.

\(^63\) This may not be true for private cloud deployments. See Part III.C., supra.

\(^64\) For a more in-depth economic analysis of cloud versus on-premise equipment, see Cloud Economics—Are You Getting the Bigger Picture?, CLOUD TECHNOLOGY PARTNERS, https://www.cloudtp.com/doppler/cloud-economics-getting-bigger-picture/.
providers (and in particular IaaS providers) are not. Even though they face the same challenges as end users regarding capital expenditures, cloud providers have been able to maintain competitive pricing by becoming more efficient in their delivery of services in a number of ways:

1) **Hardware Costs.** Cloud providers can rely on economies of scale to realize savings due to the volume of their purchases. In addition, rather than purchasing the brand-name, premium-priced equipment that enterprises typically purchase, cloud providers purchase lower-priced commodity hardware because resiliency and redundancy is built into the enabling technology they use. Some cloud providers have developed their own custom hardware designs to maximize the cost efficiency of their hardware purchases.

2) **Software Costs.** Many IaaS cloud providers make extensive use of open source software, thus eliminating (or greatly reducing) the cost of software. In addition, certain information technology functions that were traditionally performed by hardware are now handled through the use of software.

3) **Efficient Resource Usage.** An inherent function of the technology used to enable the provision of cloud computing is efficient resource allocation and utilization across physical resources. IaaS providers can also balance peaks and valleys in resource usage over time and geography across their customer base. The net result is that IaaS providers can operate much more efficiently than enterprises do—by some estimates up to three times more efficiently.

**C. Elasticity and Scalability**

IaaS users can provision one or many virtual server in minutes and can add or remove resources within similar timeframes. In many cases, resources can be configured to

---


68 See Part II.F., supra.

69 See Part II.D., supra.

automatically scale based on workloads. Compare this to a typical process to procure a physical server, which would likely involve several days at a minimum. Unlike physical hardware, there is no need to install or remove servers, nor are there concerns about hardware failures. In brief, cloud computing is more flexible and convenient for end users.

D. Increased Functionality and Reliability

One aspect of cloud computing that has often been cited by Amazon in relation to its cloud offerings is what it has dubbed the “flywheel effect” or the “virtuous circle.” Cloud computing features significant economies of scale. As a cloud provider signs on more customers, there is more usage; more usage leads the cloud provider to roll out more infrastructure; more infrastructure leads to greater economies of scale, resulting in lower infrastructure costs; lower infrastructure costs lead to additional innovations and reduced prices, both of which in turn lead to more customers, and so on. The net result for Amazon has been an exponential increase in new features over time—for example, it launched a total of 24 new services in 2008, but in 2015 it launched 516. Amazon’s early emphasis on developing its cloud computing business has given it a substantial lead in this efficiency race, which some believe is unlikely to be matched by any other provider.

The adoption of a SaaS as a delivery mechanism, as compared to more traditional software licensed for on-premise use, has given SaaS providers unprecedented insight into the usage of their applications. SaaS vendors can gather great quantities of data on how their applications are used, who is using them, when they are used, what features are used (and which are not), problems encountered, and so on, all in real time. As a result, SaaS providers are able to identify and fix bugs sooner and identify and improve the features and functionality of their application more rapidly. In addition, SaaS users have fewer concerns about patches, version management, and compatibility and interoperability with hardware, operating systems, and other software.

E. Additional Revenue Streams

Cloud computing approaches also open up the possibility of realizing additional revenue streams for providers, particularly for SaaS providers. Although this is of course primarily of benefit to providers, it may indirectly benefit users by enabling the cloud provider to lower fees charged to end users for the cloud provider’s services. The following are a few examples:

72 Id.
73 Collection of usage information is also possible with on-premise software (often referred to as “telemetry”) and has become more prevalent over time. However, it is subject to certain challenges, e.g., the blocking of telemetry transmissions by an end user’s security measures, such as firewalls.
1) **Data monetization.** In contrast to on-premise software, with SaaS the service provider will typically have some degree of access to the data of its users. This opens up the possibility of utilizing the data for purposes other than the provision of services to the end user in order to generate additional revenue. Although the topic is not without controversy and often raises the hackles of prospective SaaS users, data monetization is not a novel concept and is a strategy that has been employed by other types of service providers since before the advent of cloud computing.\(^{75}\) In some cases, the potential value in data monetization could exceed the revenue generated from the service itself.\(^{76}\) It is not difficult to envision how this could be applied to various types of SaaS applications across numerous industries.\(^{77}\) Of course, end users may have concerns regarding privacy, confidentiality, and security of their data, particularly in the case of highly sensitive data such as personal health information.

2) **Networks, ecosystems, and marketplaces.** Although not necessarily exclusive to cloud offerings, cloud providers are to some extent better positioned to enable integration with third-party software or solutions, due in part to the use of industry standard interfaces and protocols.\(^{78}\) Cloud providers can further encourage this by developing interfaces and protocols to enable third parties to easily integrate with the cloud provider’s offerings, or offer PaaS solutions to enable the development or migration of their applications.\(^{79}\) Doing so allows the cloud provider to earn additional revenue by charging for the privilege of integrating with the cloud provider’s platform, for example, or acting as a reseller or distributor and earning

---

\(^{75}\) For example, ADP, a payroll service provider, leverages the information it collects in the course of providing such services to develop analytics and statistical reports and related services. See Tom Davenport, *How ADP Gives Data Value Back To Its Customers*, FORBES, Mar. 12, 2018, https://www.forbes.com/sites/tomdavenport/2018/03/12/how-adp-gives-data-value-back-to-its-customers/?sh=5732c. American Express has used transaction information from its 90 million cards to create Business Insights, “a global information, analytics and consulting organisation that combines real behavioural information—based on actual aggregated purchasing data—with sophisticated analytics to reveal unique insights about your customers, competitive set and marketplace.” *Attract Customers/Business Insights*, AMERICAN EXPRESS, https://www.americaexpress.com/uk/content/merchant/business-insights.html. Financial institutions and credit bureaus have a long history of sharing and aggregating data to produce credit reports.

\(^{76}\) One possible example, albeit not from the SaaS world, is Nest, a maker of smart thermostats. Nest collects and aggregates information on energy consumption from those thermostats to provide aggregated energy consumption data to energy companies to enable them to better plan their energy production. The founder of Nest believes that the sale of such data will eventually earn more than the sales of the thermostats. See Matthew Mobrea, *Google’s Real Plan Behind the Purchase of the Nest Thermostat*, ITWORLD, Apr. 25, 2014, https://www.itworld.com/article/2833423/big-data/google-s-real-plan-behind-the-purchase-of-the-nest-thermostat.html.

\(^{77}\) For example, an online sales or e-commerce platform (e.g., Shopify) could develop analytics on retail sales and pricing strategy; an online accounting services provider could develop analytics on financial information and lending. See also Leo Polovets, *The Value of Data, Part 3: Data Business Models*, CODING VC, Mar. 12, 2015, https://codingvc.com/the-value-of-data-part-3-data-business-models. The article discusses other companies that currently use such strategies.


Overview of Cloud Computing Technology

a margin when its users purchase third-party solutions. The availability of an ecosystem comprised of the cloud provider’s offerings integrated with third-party solutions may be attractive to end users insofar as it offers one-stop shopping and may reduce the efforts otherwise needed to integrate disparate solutions. This, of course, may make it difficult for end users to transition to another cloud provider—a benefit for the cloud provider, but perhaps a detriment to the end user.

3) Advertising. Cloud providers can also potentially earn revenue through the placement of advertisements within their offerings, but at the risk of alienating users. In contrast to passive, content-only websites, SaaS providers may be in a position to leverage the context of the offered services or end user data to target ad placements much more precisely. For example, users of an online accounting solution could be presented with advertisements for loans or tax-return preparation.

V. Disadvantages of Cloud Computing

Cloud computing entails certain disadvantages and risks. The following are some of the more typical concerns:

1) Reliance on internet connectivity. Given that cloud services are almost always accessed through the internet, continued access will be dependent on the reliability of both the internet and the end user’s internet service provider. If either fails, cloud services cannot be used. Although the risk of a failure of a given internet service provider may be mitigated by using multiple service providers, doing so would come at an additional cost and does not address widespread internet outages. Lastly, given the nature of cloud services, additional bandwidth will likely be required, which will increase costs.

2) Privacy and security. The use of cloud services will in almost all cases result in the transfer of data to the cloud service provider, which may raise concerns about the cloud provider’s ability to protect against data breach. Cloud providers will typically implement a standard set of security measures, which may or may not meet the requirements of a given user. Although larger providers may implement a fairly robust set of security measures, users must nonetheless evaluate the extent

---

80 One example of this is Amazon’s AWS Marketplace, https://aws.amazon.com/marketplace/, which offers thousands of software options that can be deployed on its cloud platform, including operating systems, security software, databases, and business intelligence applications. At the SaaS level, Salesforce’s AppExchange, https://appexchange.salesforce.com, offers thousands of applications and services that can be used seamlessly within its platform. Note that the possibilities are not limited to software applications. For example, a SaaS provider offering supply-chain management applications could integrate with providers of shipping or warehousing services. Nor are they limited to providers. For example, a SaaS provider could offer sales solutions for insurance agents as well as services for insurance companies, both of whom would benefit from the interoperability of a common platform.

81 Perhaps one of the earliest examples of this was introduction of advertisements in Google’s Gmail. Google displayed targeted advertising by scanning e-mail contents, although only for users of the free consumer version. However, this practice was not without controversy, and ultimately Google decided to discontinue the practice. See Daisuke Wakabayashi, Google Will No Longer Scan Gmail for Ad Targeting, N.Y.Times, June 23, 2017, https://www.nytimes.com/2017/06/23/technology/gmail-ads.html.

to which such measures are adequate for their specific needs, and the extent to which providers contractually commit to responsibility for such measures. The question of security should also be considered in comparison to what security measures a given end user would take in the alternative—for example, if an end user were to set up its own on-premises equipment instead of using a cloud provider, would it implement security measures that are more robust than those implemented by the cloud provider?83 Lastly, end users may wish to consider adding their own security measures to those offered by the cloud service provider, such as encrypting data before transmitting it to the cloud.84

3) **Reliance on the cloud provider.** The use of a cloud provider will necessarily result in some degree of reliance on the provider. If there is a failure in the cloud provider’s service, an end user’s business operations will likely be impacted. This can be mitigated to some extent by using multiple redundant service providers, though typically only for IaaS and at additional cost and complexity. However, as with security considerations, this question should also be considered in comparison with the alternatives. For example, if an end user were to set up its own on-premise equipment instead of using a cloud provider, it would still be reliant on the vendor of that equipment to address a failure.

VI. **Conclusion**

This chapter is an admittedly cursory overview of the more significant technologies used in cloud computing. It is by no means comprehensive, but will equip business lawyers with a basic technological understanding of cloud computing to counsel more effectively when advising on cloud computing transactions.

83 Some have argued that, in general terms, cloud providers will be better positioned to implement robust security than their end users due, in part, to economies of scale. For example, cloud providers can design and build physically secure data centers and employ teams of security experts and spread that cost across their user base; given that they can monitor a much greater amount of internet traffic, they may be able to detect malware and similar threats sooner, and the larger scale of a cloud provider’s infrastructure may be less susceptible to distributed denial of service attacks. See Andrew Froehlich, *Why Cloud Security Beats Your Data Center*, InformationWeek, July 21, 2015, https://www.informationweek.com/cloud/infrastructure-as-a-service/why-cloud-security-beats-your-data-center/d-id/1321354; Brandon Butler, *Public Cloud vs. On-Premises, Which Is More Secure?*, NetworkWorld, Dec. 17, 2015, https://www.networkworld.com/article/3016673/public-cloud/public-cloud-vs-on-premises-which-is-more-secure.html. It is notable that U.S. intelligence agencies concluded that Amazon Web Services could maintain adequate security when they awarded it a contract for the provision of cloud services, see Frank Konkel, *The Details About the CIA’s Deal With Amazon*, ATLANTIC, July 17, 2014, https://www.theatlantic.com/technology/archive/2014/07/the-details-about-the-cias-deal-with-amazon/374632/, although that, of course, is not necessarily indicative of the security measures that all end users can expect. Lastly, many recent data breaches that involve cloud services resulted from end user misuse or misconfiguration. See, e.g., Dan O’Sullivan, *The RNC Files: Inside the Largest US Voter Data Leak*, UpGuard, Dec. 20, 2017, https://www.upguard.com/breaches/the-rnc-files.

84 See Stephen Lawton, *Cloud Encryption: Using Data Encryption in the Cloud*, Tom’s IT Pro, Apr. 30, 2015, http://www.tomspitpro.com/articles/cloud-data-encryption.2-913.html. User-controlled security measures may be more easily implemented for IaaS than for SaaS because end users will have a greater degree of control over infrastructure for the former. Some vendors have developed resources that enable end users to control encryption in connection with SaaS services. See, e.g., Vaultive’s solution for Microsoft Office 365, https://vaultive.com/for-your-technology/office-365-security/.