Cloud computing
From Wikipedia, the free encyclopedia

Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services),[1][2] which can be rapidly provisioned and released with minimal management effort. Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in either privately owned, or third-party data centers[3] that may be located far from the user—ranging in distance from across a city to across the world. Cloud computing relies on sharing of resources to achieve coherence and economy of scale, similar to a utility (like the electricity grid) over an electricity network.

Advocates claim that cloud computing allows companies to avoid up-front infrastructure costs (e.g., purchasing servers). As well, it enables organizations to focus on their core businesses instead of spending time and money on computer infrastructure.[4] Proponents also claim that cloud computing allows enterprises to get their applications up and running faster, with improved manageability and less maintenance, and enables Information technology (IT) teams to more rapidly adjust resources to meet fluctuating and unpredictable business demand.[4][5][6] Cloud providers typically use a "pay as you go" model. This will lead to unexpectedly high charges if administrators do not adapt to the cloud pricing model.[7]

In 2009, the availability of high-capacity networks, low-cost computers and storage devices as well as the widespread adoption of hardware virtualization, service-oriented architecture, and autonomic and utility computing led to a growth in cloud computing.[8][9][10] Companies can scale up as computing needs increase and then scale down again as demands decrease.[11] In 2013, it was reported that cloud computing had become a highly demanded service or utility due to the advantages of high computing power, cheap cost of services, high performance, scalability, accessibility as well as availability. Some cloud vendors are experiencing growth rates of 50% per year,[12] but being still in a stage of infancy, it has pitfalls that need to be addressed to make cloud computing services more reliable and user friendly.[13][14]

Contents
History

Origin of the term

The origin of the term *cloud computing* is unclear. The word "cloud" is commonly used in science to describe a large agglomeration of objects that visually appear from a distance as a cloud and describes any set of things whose details are not further inspected in a given context. Another explanation is that the old programs that drew network schematics surrounded the icons for servers with a
circle, and a cluster of servers in a network diagram had several overlapping circles, which resembled a cloud.[16] In analogy to the above usage, the word cloud was used as a metaphor for the Internet and a standardized cloud-like shape was used to denote a network on telephony schematics. Later it was used to depict the Internet in computer network diagrams. With this simplification, the implication is that the specifics of how the end points of a network are connected are not relevant for the purposes of understanding the diagram. The cloud symbol was used to represent networks of computing equipment in the original ARPANET by as early as 1977,[17] and the CSNET by 1981[18]—both predecessors to the Internet itself.

The term cloud has been used to refer to platforms for distributed computing. In Wired's April 1994 feature "Bill and Andy's Excellent Adventure II" on the Apple spin-off General Magic, Andy Hertzfeld commented on General Magic's distributed programming language Telescript that:

"The beauty of Telescript ... is that now, instead of just having a device to program, we now have the entire Cloud out there, where a single program can go and travel to many different sources of information and create sort of a virtual service. No one had conceived that before. The example Jim White [the designer of Telescript, X.400 and ASN.1] uses now is a date-arranging service where a software agent goes to the flower store and orders flowers and then goes to the ticket shop and gets the tickets for the show, and everything is communicated to both parties."

—[19]

References to "cloud computing" in its modern sense appeared as early as 1996, with the earliest known mention in a Compaq internal document.[20] The popularization of the term can be traced to 2006 when Amazon.com introduced its Elastic Compute Cloud.[21]

1970s

During the 1960s, the initial concepts of time-sharing became popularized via RJE (Remote Job Entry);[22] this terminology was mostly associated with large vendors such as IBM and DEC. Full time-sharing solutions were available by the early 1970s on such platforms as Multics (on GE hardware), Cambridge CTSS, and the earliest UNIX ports (on DEC hardware). Yet, the "data center" model where users submitted jobs to operators to run on IBM mainframes was overwhelmingly predominant.

1990s

In 1990s, telecommunications companies, who previously offered primarily dedicated point-to-point data circuits, began offering virtual private network (VPN) services with comparable quality of service, but at a lower cost. By switching traffic as they saw fit to balance server use, they could use overall network bandwidth more effectively. They began to use the cloud symbol to denote the demarcation point between what the provider was responsible for and what users
were responsible for. Cloud computing extended this boundary to cover all servers as well as the network infrastructure.[23] As computers became more diffused, scientists and technologists explored ways to make large-scale computing power available to more users through time-sharing. They experimented with algorithms to optimize the infrastructure, platform, and applications to prioritize CPUs and increase efficiency for end users.[24]

2000s

Since 2000, cloud computing has come into existence. In early 2008, NASA's OpenNebula, enhanced in the RESERVOIR European Commission-funded project, became the first open-source software for deploying private and hybrid clouds, and for the federation of clouds.[25] In the same year, efforts were focused on providing quality of service guarantees (as required by real-time interactive applications) to cloud-based infrastructures, in the framework of the IRMOS European Commission-funded project, resulting in a real-time cloud environment.[26][27] By mid-2008, Gartner saw an opportunity for cloud computing "to shape the relationship among consumers of IT services, those who use IT services and those who sell them"[28] and observed that "organizations are switching from company-owned hardware and software assets to per-use service-based models" so that the "projected shift to computing ... will result in dramatic growth in IT products in some areas and significant reductions in other areas."[29]

In August 2006 Amazon introduced its Elastic Compute Cloud.[21] Microsoft Azure was announced as "Azure" in October 2008 and was released on 1 February 2010 as Windows Azure, before being renamed to Microsoft Azure on 25 March 2014.[30] For a time, Azure was on the TOP500 supercomputer list, before it dropped off it.[31] In July 2010, Rackspace Hosting and NASA jointly launched an open-source cloud-software initiative known as OpenStack. The OpenStack project intended to help organizations offering cloud-computing services running on standard hardware. The early code came from NASA's Nebula platform as well as from Rackspace's Cloud Files platform. On March 1, 2011, IBM announced the IBM SmartCloud framework to support Smarter Planet.[32] Among the various components of the Smarter Computing foundation, cloud computing is a critical part. On June 7, 2012, Oracle announced the Oracle Cloud.[33] While aspects of the Oracle Cloud are still in development, this cloud offering is poised to be the first to provide users with access to an integrated set of IT solutions, including the Applications (SaaS), Platform (PaaS), and Infrastructure (IaaS) layers.[34][35][36]

Similar concepts

Cloud computing is the result of the evolution and adoption of existing technologies and paradigms. The goal of cloud computing is to allow users to take benefit from all of these technologies, without the need for deep knowledge about or expertise with each one of them. The cloud aims to cut costs, and helps the users focus on their core business instead of being impeded by IT obstacles.[37] The main enabling technology for cloud computing is virtualization. Virtualization software separates a
physical computing device into one or more "virtual" devices, each of which can be easily used and managed to perform computing tasks. With operating system–level virtualization essentially creating a scalable system of multiple independent computing devices, idle computing resources can be allocated and used more efficiently. Virtualization provides the agility required to speed up IT operations, and reduces cost by increasing infrastructure utilization. Autonomic computing automates the process through which the user can provision resources on-demand. By minimizing user involvement, automation speeds up the process, reduces labor costs and reduces the possibility of human errors.[37] Users routinely face difficult business problems. Cloud computing adopts concepts from Service-oriented Architecture (SOA) that can help the user break these problems into services that can be integrated to provide a solution. Cloud computing provides all of its resources as services, and makes use of the well-established standards and best practices gained in the domain of SOA to allow global and easy access to cloud services in a standardized way.

Cloud computing also leverages concepts from utility computing to provide metrics for the services used. Such metrics are at the core of the public cloud pay-per-use models. In addition, measured services are an essential part of the feedback loop in autonomic computing, allowing services to scale on-demand and to perform automatic failure recovery. Cloud computing is a kind of grid computing; it has evolved by addressing the QoS (quality of service) and reliability problems. Cloud computing provides the tools and technologies to build data/compute intensive parallel applications with much more affordable prices compared to traditional parallel computing techniques.[37]

Cloud computing shares characteristics with:

- **Client–server model**—*Client–server computing* refers broadly to any distributed application that distinguishes between service providers (servers) and service requestors (clients).[38]
- **Computer bureau**—A service bureau providing computer services, particularly from the 1960s to 1980s.
- **Grid computing**—"A form of distributed and parallel computing, whereby a 'super and virtual computer' is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks."
- **Fog computing**—Distributed computing paradigm that provides data, compute, storage and application services closer to client or near-user edge devices, such as network routers. Furthermore, fog computing handles data at the network level, on smart devices and on the end-user client side (e.g. mobile devices), instead of sending data to a remote location for processing.
- **Dew computing**—In the existing computing hierarchy, the Dew computing is positioned as the ground level for the cloud and fog computing paradigms. Compared to fog computing, which supports emerging IoT applications that demand real-time and predictable latency and the dynamic network reconfigurability, Dew computing pushes the frontiers to computing
applications, data, and low level services away from centralized virtual nodes to the end users.[39]

- **Mainframe computer**—Powerful computers used mainly by large organizations for critical applications, typically bulk data processing such as: census; industry and consumer statistics; police and secret intelligence services; enterprise resource planning; and financial transaction processing.
- **Utility computing**—The "packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity."[40][41]
- **Peer-to-peer**—A distributed architecture without the need for central coordination. Participants are both suppliers and consumers of resources (in contrast to the traditional client–server model).
- **Green computing**
- **Cloud sandbox**—A live, isolated computer environment in which a program, code or file can run without affecting the application in which it runs.

### Characteristics

Cloud computing exhibits the following key characteristics:

- **Agility** for organizations may be improved, as cloud computing may increase users’ flexibility with re-provisioning, adding, or expanding technological infrastructure resources.
- **Cost reductions** are claimed by cloud providers. A public-cloud delivery model converts capital expenditures (e.g., buying servers) to operational expenditure.[42] This purportedly lowers barriers to entry, as infrastructure is typically provided by a third party and need not be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is "fine-grained", with usage-based billing options. As well, less in-house IT skills are required for implementation of projects that use cloud computing.[43] The e-FISCAL project’s state-of-the-art repository[44] contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.
- **Device and location independence**[45] enable users to access systems using a web browser regardless of their location or what device they use (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect to it from anywhere.[43]
- **Maintenance** of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places (e.g., different work locations, while travelling, etc.).
- Multitenancy enables sharing of resources and costs across a large pool of users thus allowing for:
  - centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
  - peak-load capacity increases (users need not engineer and pay for the resources and equipment to meet their highest possible load-levels)
  - utilisation and efficiency improvements for systems that are often only 10–20% utilised.[46][47]

- Performance is monitored by IT experts from the service provider, and consistent and loosely coupled architectures are constructed using web services as the system interface.[43][48][49]

- Productivity may be increased when multiple users can work on the same data simultaneously, rather than waiting for it to be saved and emailed. Time may be saved as information does not need to be re-entered when fields are matched, nor do users need to install application software upgrades to their computer.[50]

- Reliability improves with the use of multiple redundant sites, which makes well-designed cloud computing suitable for business continuity and disaster recovery.[51]

- Scalability and elasticity via dynamic ("on-demand") provisioning of resources on a fine-grained, self-service basis in near real-time[52][53] (Note, the VM startup time varies by VM type, location, OS and cloud providers[52]), without users having to engineer for peak loads.[54][55][56] This gives the ability to scale up when the usage need increases or down if resources are not being used.[57]

- Security can improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than other traditional systems, in part because service providers are able to devote resources to solving security issues that many customers cannot afford to tackle or which they lack the technical skills to address.[58] However, the complexity of security is greatly increased when data is distributed over a wider area or over a greater number of devices, as well as in multi-tenant systems shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.

The National Institute of Standards and Technology's definition of cloud computing identifies "five essential characteristics":

https://en.wikipedia.org/wiki/Cloud_computing
**On-demand self-service.** A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

**Broad network access.** Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).

**Resource pooling.** The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.

**Rapid elasticity.** Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear unlimited and can be appropriated in any quantity at any time.

**Measured service.** Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

— National Institute of Standards and Technology[2]

### Service models

Though service-oriented architecture advocates "everything as a service" (with the acronyms EaaS or XaaS or simply aas),[59] cloud-computing providers offer their "services" according to different models, of which the three standard models per NIST are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).[2] These models offer increasing abstraction; they are thus often portrayed as a *layers* in a stack: infrastructure-, platform- and software-as-a-service,[60] but these need not be related. For example, one can provide SaaS implemented on physical machines (bare metal), without using underlying PaaS or IaaS layers, and conversely one can run a program on IaaS and access it directly, without wrapping it as SaaS.

The NIST's definition of cloud computing defines the service models as follows:[2]

**Software as a Service (SaaS).** The capability provided to the consumer is to use the provider's applications
running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

*Platform as a Service (PaaS).* The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

*Infrastructure as a Service (IaaS).* The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls).
According to the Internet Engineering Task Force (IETF), the most basic cloud-service model is that of providers offering computing infrastructure – virtual machines and other resources – as a service to subscribers. Infrastructure as a service (IaaS) refers to online services that abstract the user from the details of infrastructure like physical computing resources, location, data partitioning, scaling, security, backup etc. A hypervisor, such as Xen, Oracle VirtualBox, Oracle VM, KVM, VMware ESX/ESXi, or Hyper-V, runs the virtual machines as guests. Pools of hypervisors within the cloud operational system can support large numbers of virtual machines and the ability to scale services up and down according to customers' varying requirements. Linux containers run in isolated partitions of a single Linux kernel running directly on the physical hardware. Linux cgroups and namespaces are the underlying Linux kernel technologies used to isolate, secure and manage the containers. Containerisation offers higher performance than virtualization, because there is no hypervisor overhead. Also, container capacity auto-scales dynamically with computing load, which eliminates the problem of over-provisioning and enables usage-based billing.[61] IaaS clouds often offer additional resources such as a virtual-machine disk-image library, raw block storage, file or object storage, firewalls, load balancers, IP addresses, virtual local area networks (VLANs), and software bundles.[62]

IaaS-cloud providers supply these resources on-demand from their large pools of equipment installed in data centers. For wide-area connectivity, customers can use either the Internet or carrier clouds (dedicated virtual private networks). To deploy their applications, cloud users install operating-system images and their application software on the cloud infrastructure.[63] In this model, the cloud user patches and maintains the operating systems and the application software. Cloud providers typically bill IaaS services on a utility computing basis: cost reflects the amount of resources allocated and consumed.[64][65][66][67]

Platform as a service (PaaS)

PaaS vendors offer a development environment to application developers. The provider typically develops toolkit and standards for development and channels for distribution and payment. In the PaaS models, cloud providers deliver a computing platform, typically including operating system, programming-language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers. With some PaaS offers like Microsoft Azure and Google App Engine, the underlying computer and storage resources scale automatically to match application demand so that the cloud user does not have to allocate resources manually. The latter has also been proposed by an architecture aiming to facilitate real-time in cloud environments.[68] Even more specific application types can be provided via PaaS, such as media encoding as provided by services like bitcodin.com[69] or media.io.[70]

Some integration and data management providers have also embraced specialized applications of PaaS as delivery models for data solutions. Examples include iPaaS (Integration Platform as a Service) and dPaaS (Data Platform as a Service). iPaaS enables customers to develop, execute and govern integration flows.[71] Under the
 iPaaS integration model, customers drive the development and deployment of integrations without installing or managing any hardware or middleware.[72] dPaaS delivers integration—and data-management—products as a fully managed service.[73] Under the dPaaS model, the PaaS provider, not the customer, manages the development and execution of data solutions by building tailored data applications for the customer. dPaaS users retain transparency and control over data through data-visualization tools.[74] Platform as a Service (PaaS) consumers do not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but have control over the deployed applications and possibly configuration settings for the application-hosting environment.

A recent specialized PaaS is the Blockchain as a Service (BaaS), that some vendors such as Microsoft Azure have already included in their PaaS offering.[75]

Software as a service (SaaS)

In the software as a service (SaaS) model, users gain access to application software and databases. Cloud providers manage the infrastructure and platforms that run the applications. SaaS is sometimes referred to as "on-demand software" and is usually priced on a pay-per-use basis or using a subscription fee.[76] In the SaaS model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients. Cloud users do not manage the cloud infrastructure and platform where the application runs. This eliminates the need to install and run the application on the cloud user's own computers, which simplifies maintenance and support. Cloud applications differ from other applications in their scalability—which can be achieved by cloning tasks onto multiple virtual machines at run-time to meet changing work demand.[77] Load balancers distribute the work over the set of virtual machines. This process is transparent to the cloud user, who sees only a single access-point. To accommodate a large number of cloud users, cloud applications can be multitenant, meaning that any machine may serve more than one cloud-user organization.

The pricing model for SaaS applications is typically a monthly or yearly flat fee per user,[78] so prices become scalable and adjustable if users are added or removed at any point.[79] Proponents claim that SaaS gives a business the potential to reduce IT operational costs by outsourcing hardware and software maintenance and support to the cloud provider. This enables the business to reallocate IT operations costs away from hardware/software spending and from personnel expenses, towards meeting other goals. In addition, with applications hosted centrally, updates can be released without the need for users to install new software. One drawback of SaaS comes with storing the users' data on the cloud provider's server. As a result, there could be unauthorized access to the data. For this reason, users are increasingly adopting intelligent third-party key-management systems to help secure their data.

Mobile "backend" as a service (MBaaS)
In the mobile "backend" as a service (m) model, also known as backend as a service (BaaS), web app and mobile app developers are provided with a way to link their applications to cloud storage and cloud computing services with application programming interfaces (APIs) exposed to their applications and custom software development kits (SDKs). Services include user management, push notifications, integration with social networking services[80] and more. This is a relatively recent model in cloud computing,[81] with most BaaS startups dating from 2011 or later[82][83][84] but trends indicate that these services are gaining significant mainstream traction with enterprise consumers.[85]

Serverless computing

Serverless computing is a cloud computing code execution model in which the cloud provider fully manages starting and stopping virtual machines as necessary to serve requests, and requests are billed by an abstract measure of the resources required to satisfy the request, rather than per virtual machine, per hour.[86] Despite the name, it does not actually involve running code without servers.[86] Serverless computing is so named because the business or person that owns the system does not have to purchase, rent or provision servers or virtual machines for the back-end code to run on.

Cloud clients

Users access cloud computing using networked client devices, such as desktop computers, laptops, tablets and smartphones and any Ethernet enabled device such as Home Automation Gadgets. Some of these devices—cloud clients—rely on cloud computing for all or a majority of their applications so as to be essentially useless without it. Examples are thin clients and the browser-based Chromebook. Many cloud applications do not require specific software on the client and instead use a web browser to interact with the cloud application. With Ajax and HTML5 these Web user interfaces can achieve a similar, or even better, look and feel to native applications. Some cloud applications, however, support specific client software dedicated to these applications (e.g., virtual desktop clients and most email clients). Some legacy applications (line of business applications that until now have been prevalent in thin client computing) are delivered via a screen-sharing technology.

Deployment models

Private cloud

Private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party, and hosted either internally or externally.[2] Undertaking a private cloud project requires a significant level and degree of engagement to virtualize the business environment, and requires the organization to reevaluate decisions about existing resources. When done right, it can improve business, but every step in the project raises security issues that must
be addressed to prevent serious vulnerabilities. Self-run data centers\[^{87}\]\ are generally capital intensive. They have a significant physical footprint, requiring allocations of space, hardware, and environmental controls. These assets have to be refreshed periodically, resulting in additional capital expenditures. They have attracted criticism because users "still have to buy, build, and manage them" and thus do not benefit from less hands-on management,\[^{88}\]\ essentially "[lacking] the economic model that makes cloud computing such an intriguing concept".\[^{89}\][^90]\]

**Public cloud**

A cloud is called a "public cloud" when the services are rendered over a network that is open for public use. Public cloud services may be free.\[^{91}\]\ Technically there may be little or no difference between public and private cloud architecture, however, security consideration may be substantially different for services (applications, storage, and other resources) that are made available by a service provider for a public audience and when communication is effected over a non-trusted network. Generally, public cloud service providers like Amazon Web Services (AWS), Microsoft and Google own and operate the infrastructure at their data center and access is generally via the Internet. AWS and Microsoft also offer direct connect services called "AWS Direct Connect" and "Azure ExpressRoute" respectively, such connections require customers to purchase or lease a private connection to a peering point offered by the cloud provider.\[^{43}\]\]

**Hybrid cloud**

Hybrid cloud is a composition of two or more clouds (private, community or public) that remain distinct entities but are bound together, offering the benefits of multiple deployment models. Hybrid cloud can also mean the ability to connect colocation, managed and/or dedicated services with cloud resources.\[^{2}\]\ Gartner, Inc. defines a hybrid cloud service as a cloud computing service that is composed of some combination of private, public and community cloud services, from different service providers.\[^{92}\]\ A hybrid cloud service crosses isolation and provider boundaries so that it can’t be simply put in one category of private, public, or community cloud service. It allows one to extend either the capacity or the capability of a cloud service, by aggregation, integration or customization with another cloud service.

Varied use cases for hybrid cloud composition exist. For example, an organization may store sensitive client data in house on a private cloud application, but interconnect that application to a business intelligence application provided on a
public cloud as a software service.[93] This example of hybrid cloud extends the capabilities of the enterprise to deliver a specific business service through the addition of externally available public cloud services. Hybrid cloud adoption depends on a number of factors such as data security and compliance requirements, level of control needed over data, and the applications an organization uses.[94]

Another example of hybrid cloud is one where IT organizations use public cloud computing resources to meet temporary capacity needs that can not be met by the private cloud.[95] This capability enables hybrid clouds to employ cloud bursting for scaling across clouds.[2] Cloud bursting is an application deployment model in which an application runs in a private cloud or data center and "bursts" to a public cloud when the demand for computing capacity increases. A primary advantage of cloud bursting and a hybrid cloud model is that an organization pays for extra compute resources only when they are needed.[96] Cloud bursting enables data centers to create an in-house IT infrastructure that supports average workloads, and use cloud resources from public or private clouds, during spikes in processing demands.[97] The specialized model of hybrid cloud, which is built atop heterogeneous hardware, is called "Cross-platform Hybrid Cloud". A cross-platform hybrid cloud is usually powered by different CPU architectures, for example, x86-64 and ARM, underneath. Users can transparently deploy and scale applications without knowledge of the cloud's hardware diversity.[98] This kind of cloud emerges from the raise of ARM-based system-on-chip for server-class computing.

Others

Community cloud

Community cloud shares infrastructure between several organizations from a specific community with common concerns (security, compliance, jurisdiction, etc.), whether managed internally or by a third-party, and either hosted internally or externally. The costs are spread over fewer users than a public cloud (but more than a private cloud), so only some of the cost savings potential of cloud computing are realized.[2]

Distributed cloud

A cloud computing platform can be assembled from a distributed set of machines in different locations, connected to a single network or hub service. It is possible to distinguish between two types of distributed clouds: public-resource computing and volunteer cloud.

- Public-resource computing—This type of distributed cloud results from an expansive definition of cloud computing, because they are more akin to distributed computing than cloud computing. Nonetheless, it is considered a sub-class of cloud computing, and some examples include distributed computing platforms such as BOINC and Folding@Home.
- **Volunteer cloud**—Volunteer cloud computing is characterized as the intersection of public-resource computing and cloud computing, where a cloud computing infrastructure is built using volunteered resources. Many challenges arise from this type of infrastructure, because of the volatility of the resources used to build it and the dynamic environment it operates in. It can also be called peer-to-peer clouds, or ad-hoc clouds. An interesting effort in such direction is Cloud@Home, it aims to implement a cloud computing infrastructure using volunteered resources providing a business-model to incentivize contributions through financial restitution.[99]

**Intercloud**

The Intercloud[100] is an interconnected global "cloud of clouds"[101][102] and an extension of the Internet "network of networks" on which it is based. The focus is on direct interoperability between public cloud service providers, more so than between providers and consumers (as is the case for hybrid- and multi-cloud).[103][104][105]

**Multicloud**

Multicloud is the use of multiple cloud computing services in a single heterogeneous architecture to reduce reliance on single vendors, increase flexibility through choice, mitigate against disasters, etc. It differs from hybrid cloud in that it refers to multiple cloud services, rather than multiple deployment modes (public, private, legacy).[106][107][108]

**Architecture**

Cloud architecture,[109] the systems architecture of the software systems involved in the delivery of cloud computing, typically involves multiple *cloud components* communicating with each other over a loose coupling mechanism such as a messaging queue. Elastic provision implies intelligence in the use of tight or loose coupling as applied to mechanisms such as these and others.

**Cloud engineering**

Cloud engineering is the application of engineering disciplines to cloud computing. It brings a systematic approach to the high-level concerns of commercialization, standardization, and governance in conceiving, developing,
operating and maintaining cloud computing systems. It is a multidisciplinary method encompassing contributions from diverse areas such as systems, software, web, performance, information, security, platform, risk, and quality engineering.

Security and privacy

Cloud computing poses privacy concerns because the service provider can access the data that is in the cloud at any time. It could accidentally or deliberately alter or even delete information. Many cloud providers can share information with third parties if necessary for purposes of law and order even without a warrant. That is permitted in their privacy policies, which users must agree to before they start using cloud services. Solutions to privacy include policy and legislation as well as end users' choices for how data is stored. Users can encrypt data that is processed or stored within the cloud to prevent unauthorized access. According to the Cloud Security Alliance, the top three threats in the cloud are Insecure Interfaces and API’s, Data Loss & Leakage, and Hardware Failure—which accounted for 29%, 25% and 10% of all cloud security outages respectively. Together, these form shared technology vulnerabilities. In a cloud provider platform being shared by different users there may be a possibility that information belonging to different customers resides on same data server. Therefore, Information leakage may arise by mistake when information for one customer is given to other. Additionally, Eugene Schultz, chief technology officer at Emagined Security, said that hackers are spending substantial time and effort looking for ways to penetrate the cloud. "There are some real Achilles' heels in the cloud infrastructure that are making big holes for the bad guys to get into". Because data from hundreds or thousands of companies can be stored on large cloud servers, hackers can theoretically gain control of huge stores of information through a single attack—a process he called "hyperjacking". Some examples of this include the Dropbox security breach, and iCloud 2014 leak. Dropbox had been breached in October 2014, having over 7 million of its users passwords stolen by hackers in an effort to get monetary value from it by Bitcoins (BTC). By having these passwords, they are able to read private data as well as have this data be indexed by search engines (making the information public).

There is the problem of legal ownership of the data (If a user stores some data in the cloud, can the cloud provider profit from it?). Many Terms of Service agreements are silent on the question of ownership. Physical control of the computer equipment (private cloud) is more secure than having the equipment off site and under someone else's control (public cloud). This delivers great incentive to public cloud computing service providers to prioritize building and maintaining strong management of secure services. Some small businesses that don't have expertise in IT security could find that it's more secure for them to use a public cloud. There is the risk that end users do not understand the issues involved when signing on to a cloud service (persons sometimes don't read the many pages of the terms of service agreement, and just click "Accept" without reading). This is important now that cloud computing is becoming popular and required for some services to work, for example for an intelligent personal assistant (Apple's Siri or
Google Now). Fundamentally, private cloud is seen as more secure with higher levels of control for the owner, however public cloud is seen to be more flexible and requires less time and money investment from the user.[115]

Limitations and disadvantages

According to Bruce Schneier, "The downside is that you will have limited customization options. Cloud computing is cheaper because of economics of scale, and — like any outsourced task — you tend to get what you get. A restaurant with a limited menu is cheaper than a personal chef who can cook anything you want. Fewer options at a much cheaper price: it’s a feature, not a bug." He also suggests that "the cloud provider might not meet your legal needs" and that businesses need to weigh the benefits of cloud computing against the risks.[116] In cloud computing, the control of the back end infrastructure is limited to the cloud vendor only. Cloud providers often decide on the management policies, which moderates what the cloud users are able to do with their deployment.[117] Cloud users are also limited to the control and management of their applications, data and services.[118] This includes data caps, which are placed on cloud users by the cloud vendor allocating certain amount of bandwidth for each customer and are often shared among other cloud users.[119]

Emerging trends

Cloud computing is still as much a research topic, as it is a market offering.[120] What is clear through the evolution of cloud computing services is that the chief technical officer (CTO) is a major driving force behind cloud adoption.[121] The major cloud technology developers continue to invest billions a year in cloud R&D; for example: in 2011 Microsoft committed 90% of its US$9.6bn R&D budget to its cloud.[122] Centaur Partners also predict that SaaS revenue will grow from US$13.5B in 2011 to $32.8B in 2016.[123] This expansion also includes Finance and Accounting SaaS.[124] Additionally, more industries are turning to cloud technology as an efficient way to improve quality services due to its capabilities to reduce overhead costs, downtime, and automate infrastructure deployment.[125]

See also

- Category: Cloud computing providers
- Category: Cloud platforms
- Cloud computing security
- Cloud computing comparison
- Cloud management
- Cloud research
- Cloud storage
- Edge computing
- eScience
- Mobile cloud computing
- Personal cloud
- Robot as a Service
- Service-Oriented Architecture
- Ubiquitous computing
- Web computing

References

12. "The economy is flat so why are financials Cloud vendors growing at more than 90 percent per annum?". FSN. March 5, 2013.

17. "Internet History 1977".


31. "Faenov - Cluster Platform SL230s Gen8, Xeon E5-2670 8C 2.600GHz, Infiniband QDR - TOP500 Supercomputer Sites".


44. "e-FISCAL project state of the art repository".
46. "Jeff Bezos' Risky Bet". Business Week.


70. "cloud.ca".
71. "media.io" (http://www.media.io/)
97. "How Cloudbursting "Rightsizes" the Data Center".
Further reading

- Armbrust, Michael; Stoica, Ion; Zaharia, Matei; Fox, Armando; Griffith, Rean; Joseph, Anthony D.; Katz, Randy; Konwinski, Andy; Lee, Gunho; Patterson, David; Rabkin, Ariel (1 April 2010). "A view of cloud computing". *Communications of the ACM*. 53 (4): 50. doi:10.1145/1721654.1721672.

**External links**


**Categories:** Cloud computing | Cloud infrastructure

- This page was last modified on 30 January 2017, at 05:20.
- Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.