CompTIA

THE CLOUD COMPUTING ECOSYSTEM

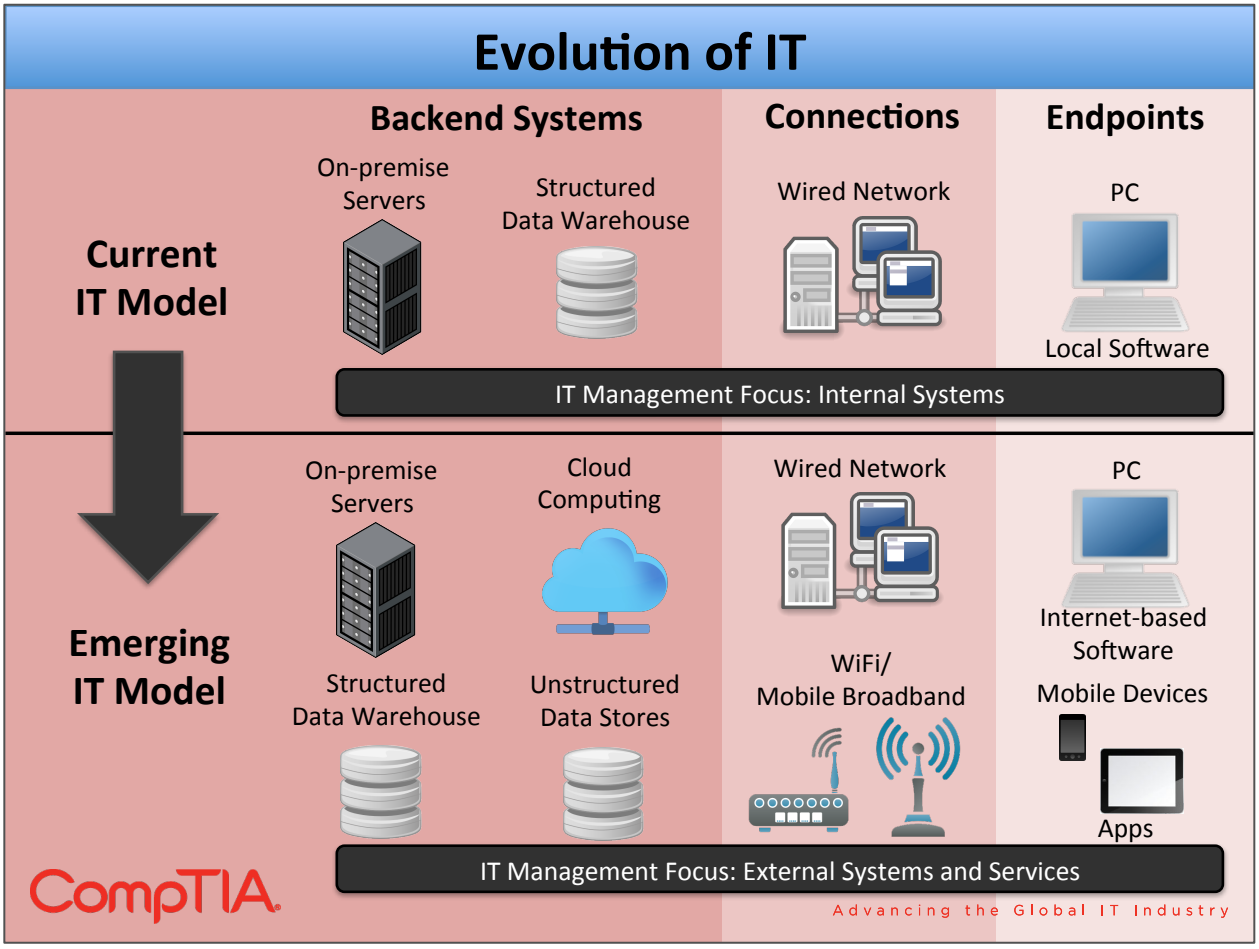
RESEARCH BRIEF

RESEARCH

www.comptia.org

The Changing IT Landscape

Although technology is in a constant state of flux, there are few occasions when the disruptions are so great that entire business models are transformed. The introduction of the PC was one of these occasions; providing compute ability to a broad population allowed for much greater automation and collaboration, and workflows adjusted accordingly.



Several different trends have combined recently to produce a similar disruption. After a period of early experimentation, it is clear that these trends will reach mass adoption, causing today's IT landscape to look markedly different from the landscape of ten years ago. There are many more options available for building solutions, and many of the options have come from the consumer space rather than being developed specifically for enterprise use.

As businesses adjust to these new models, a foundational level of knowledge will be required. Without at least a basic understanding of unique characteristics, it will be impossible to map a strategy for using new technology to achieve business objectives and separate from the competition. Cloud computing is perhaps the most significant of the new trends changing the approach to IT.

A New Model for IT Infrastructure

For most industry observers, the cloud era began when Amazon Web Services offered their first service to the public on March 13, 2006. This first service (Simple Storage Service, or S3) was quickly followed with another offering (Elastic Compute Cloud, or EC2), and cloud computing became one of the hottest trends in technology.

From a certain vantage point, cloud computing could be viewed as nothing particularly new. It certainly shares many characteristics of existing models, such as time-sharing from the early mainframe days or hosted systems that became prevalent as data centers were built out to provide off-premise infrastructure for multiple customers. This viewpoint, though, disregards the ways in which cloud computing is novel, namely the public nature of many services and the ease and speed with which users can access cloud resources.

Why Amazon?

As with most disruptive technologies, the cloud model as it exists today did not come from an established IT firm. In building out the IT architecture required to support their goals in the retail space, Amazon was forced to create many innovations in-house. They realized that these concepts would be valuable to a larger audience, and having a business built on slim profit margins enabled them to construct a plan for offering affordable IT resources on-demand. In a very short amount of time, Amazon went from a retail powerhouse to an overall technology giant.

Another critique of cloud is that it is only a delivery mechanism, and that the real value is in building services for end users on top of cloud components. This perspective has some merit—with services accounting for 57% of the overall US IT and telecom industry, there is less opportunity directly related to whatever hardware and software provide the underlying structure.

Still, those services must be built on a foundation of knowledge. Assuming that there is nothing new to cloud computing or assuming that the details are insignificant can lead to a position where it is difficult to provide high-value services. Through the early years of the cloud era, definitions have shifted and many firms have moved quickly into more advanced stages of implementation, but a basic description of cloud computing can be developed and can act as a starting point for deeper discussions.

Defining the Cloud

Attempting to build consensus around a cloud definition can be a challenging exercise. Fatigue from cloud hype and generalizations about cloud characteristics can cloud the discussion. However, the definition from the National Institute of Standards and Technology (NIST) has become generally accepted and provides insight into what makes cloud computing so revolutionary. NIST identifies five characteristics that should be present in an offering for it to be considered a cloud service:

- **Rapid elasticity:** Dynamic, unpredictable workloads require companies to build excess capacity that goes unused in off-peak times. The ability to automatically scale resources as needed can allow businesses to avoid constraints and grow as needed.

- **On-demand self-service:** Along with rapid elasticity, this is perhaps the primary defining characteristic of cloud computing. Access to systems is simplified to the point that any user can quickly provision the resources they need, removing much of the overhead from a more traditional setup.
- **Broad network access:** Ubiquitous, affordable broadband allows cloud resources to be accessed over the public Internet. Standard Internet access also allows users to access resources with common interfaces. Except for the most demanding applications, bandwidth and latency are non-issues, and geographically diverse cloud data centers are beginning to address even the needs of the demanding applications.
- **Measured service:** The many benefits of cloud computing would not be achievable without some form of metering to allow users to pay for only the resources they are consuming. Large providers are able to spread their fixed costs across many users to keep per-use charges low, and built-in monitoring simplifies the billing system.
- **Resource pooling:** Perhaps the least unique characteristic of cloud systems (especially in a private cloud model), resource pooling is still an important piece of many cloud installations and requires end users to consider the security and reliability implications of a multi-tenant model.

With many of these concepts also applying to hosted systems, it is easy to see how the lines have become blurred between the two models. For some end users, the distinction between hosted systems and cloud systems is irrelevant as a hosted system that does not feature elasticity or self-service meets all their current needs. For this reason and because cloud computing has such high public awareness, the term "hosted" is fading as a descriptor. Those qualities introduced by true cloud systems, though, will continue to become more widespread and change businesses' approach to IT.

The Cloud Stack

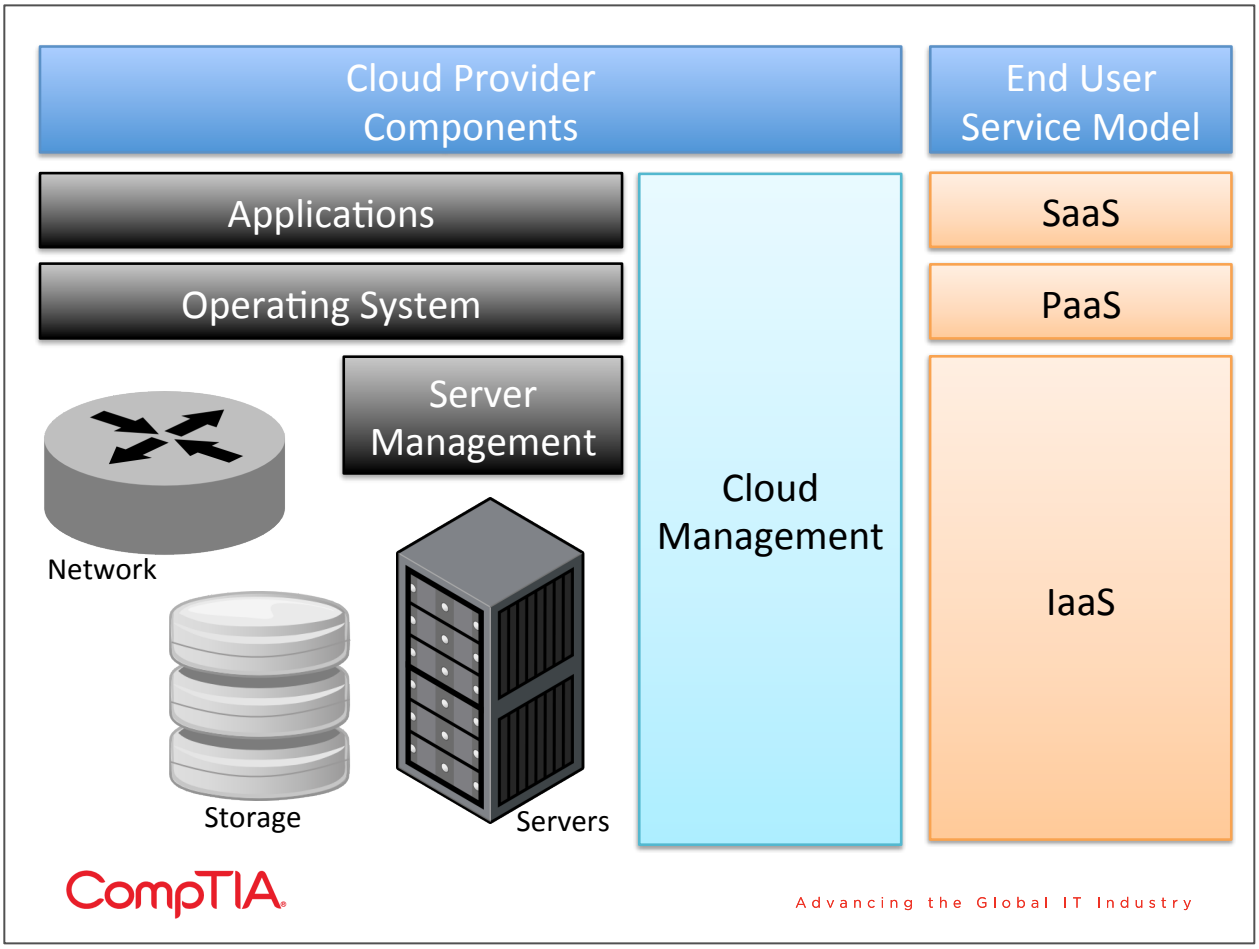
A second piece of the NIST definition is the different service models that can be offered by cloud providers. These service models are based on the typical components found in any data center, with one important component that creates the differentiation and value.

Infrastructure as a Service (IaaS) is the most foundational offering. Here, the intent is to give customers direct use of resources (compute and/or storage) and allow them to build their own platforms and applications. The most basic components needed are obviously the physical pieces—the servers and storage that will be used along with the network components that allow resources to communicate with each other inside a provider setting and also be accessed over the Internet.

The next component is the server management. In most cases, this is a hypervisor that creates virtual machines on top of physical servers. This hypervisor is essentially the same as the one that would be used by any server administrator to create virtual resources. In some cases, the server management is software that mimics the function of a hypervisor across a large pool of physical machines without creating virtual instances. Also known as a bare-metal cloud, this method removes any performance overhead introduced by a hypervisor.

Turning these components into an Infrastructure as a Service offering requires one final piece: cloud management software. Sometimes referred to as a cloud operating system, this is the software that automatically manages resources to provide elasticity, quickly allocates resources based on simple

requests from end users, and dynamically controls network settings to ensure consistent operation. Some providers, such as Amazon and Microsoft, build proprietary cloud management systems. Others choose to use third party software, which could be open source (such as OpenStack or CloudStack) or licensed (from companies such as VMWare or Eucalyptus).



From an IaaS starting point, building out Platform as a Service (PaaS) or Software as a Service (SaaS) is relatively straightforward. For PaaS, a provider will install an operating system on top of the virtual machines that get created, possibly also adding software development tools. This enables developers to quickly begin their work without worrying about the details of the virtual machines themselves. For SaaS, applications are fully built out and provided to end users.

In both of these final service models, the components that the provider adds—operating system, development tools, or applications—may be mostly similar to the version of those components in a hosted environment. The key difference is that they are built on cloud infrastructure and include the appropriate hooks to utilize that infrastructure. This gives the platform or application elasticity, and it also enables the provider to explore different business models, such as a departure from traditional software licensing.

Deployment Models

The final element of the NIST definition is a description of the various deployment models used for cloud computing. Along with understanding the definition of the different models, it is worthwhile thinking about the different challenges that may exist with each model.

Public clouds were the primary form of clouds used in the early stages of adoption. Here, a single provider makes an offering accessible to the general public, and many different users share a single pool of resources. Especially for a user transitioning from on-premise resources, this introduces a number of concerns. First, security is typically handled differently by the cloud provider than the end user might prefer. This leads to a new paradigm where security is a shared responsibility, and the end user must determine how to enhance security beyond what the provider has in place. Second, a cloud provider can experience an outage, forcing end users to build new forms of redundancy into their systems. Finally, multi-tenant architecture means that one user may constrain a given physical resource, causing issues for the other users sharing that resource.

Many companies explore private clouds as a way to address these concerns. With a private cloud, resources are reserved for a single user. A company can build a private cloud using infrastructure it owns, or it could contract with a third party to build a private cloud at a colocation facility. In the case of a company using its own infrastructure, it becomes the cloud provider in a sense and needs to consider everything from the physical components to the cloud management system. Some public cloud providers are offering virtual private clouds, but these are still multi-tenant models where the end user is able to segregate virtual instances and shield them from direct Internet access.

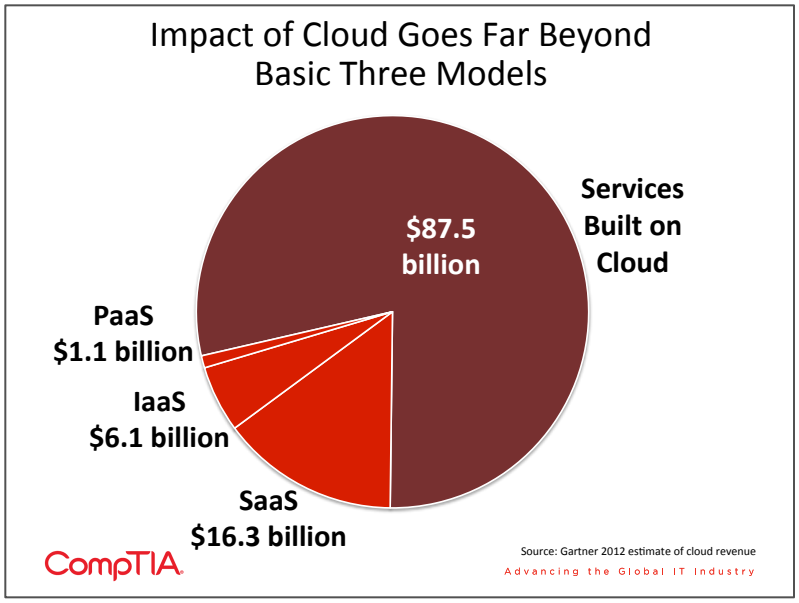
With private clouds, the primary concerns are cost and skill. The average costs of a private cloud will generally exceed those of a public cloud, since a single entity is bearing the burden of upkeep and unused capacity. Cost may not be a primary driver for a cloud transition if a company is more focused on agility or creating new offerings, but it is still a factor to consider. The proper skills are also more critical for private clouds. A public cloud provider has more ability to draw top talent, and a private cloud user must ensure that the appropriate skills exist either in-house or with their private cloud provider.

NIST defines a community cloud as a pool of resources shared by a limited group of consumers, but this model has not been prevalent in the industry up to this point. That leaves hybrid clouds as the final major model. Simply put, this is a combination of public and private clouds put to use by a single organization. The NIST definition focuses on the ability to port data between these two models, implying a single application being addressed. The term “multi-cloud” has become common to describe the more general situation of companies using different cloud models for different applications.

In general, the multi-cloud or hybrid cloud model is likely the model that most companies will use moving forward. Application needs will dictate usage, and different providers may specialize in different areas or provide different features and benefits. In addition, on-premise systems will still serve a purpose for many organizations, especially in the near future. The challenge that is created, then, is managing these disparate systems and understanding how to migrate between them efficiently.

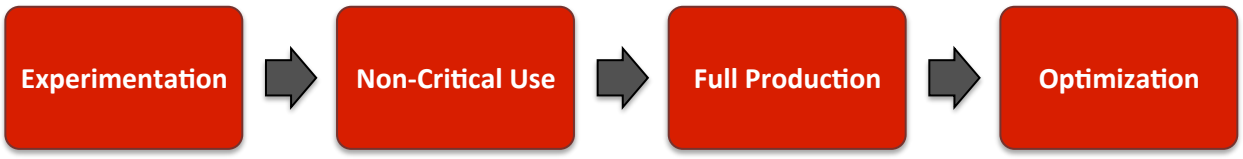
Cloud's Broad Impact

Characteristics, service models, and deployment models may provide a description of what cloud is, but the industry is finding that there are many more facets involved when trying to understand what cloud does. Beyond IaaS, PaaS, and SaaS, Gartner defines two other categories of cloud models—cloud advertising and Business Process as a Service (BPaaS). Both of these categories are similar, referring to new business models that utilize cloud infrastructure to provide higher-value products and services for end users. According to Gartner's calculations of 2012 cloud revenue, these higher-level services dwarf direct usage of the three primary cloud service models.



CompTIA research also finds that the services enabled by cloud are likely to be the areas of greatest potential moving forward. Channel firms that provide cloud offerings find than traditional offerings backed by cloud—such as email, help desk, or storage—are more viable that direct offerings related to IaaS or PaaS. These firms are also finding new services that may not have been easily offered without cloud architecture, such as virtual desktops.

As these higher-level services become widely adopted, businesses are finding that they must change their operational models to account for more agile technology that can be easily procured without the oversight of the IT department. Similarly, workflows must be changed as applications move to the cloud and potentially get rebuilt to take advantage of cloud architecture. In a standard adoption progression from early experimentation to full optimization, companies will have different needs at every stage. As companies move forward in their cloud adoption, they find that these business transformations are just as challenging as the technical hurdles involved in cloud migrations.



Cloud, then, is much more than simply a new delivery model for existing applications or a new item inside a solution provider's portfolio. The transformative potential is tremendous, and the IT industry will likely go through many changes over the next decade or more. Understanding the basic concepts behind this revolutionary model will allow companies to build the services, processes, and products required to compete in a new environment.