Global Knowledge ®

Expert Reference Series of White Papers

A Practical View of NIST's Cloud Definition

A Practical View of NIST's Cloud Definition

Vince Lo Faso, Global Knowledge Instructor, Cloud Essentials Professional, ITIL® Expert

Introduction

The National Institute of Standards and Technology (NIST) has created a robust, comprehensive cloud definition that has been well-accepted across the IT industry. It covers five essential cloud characteristics, three service models, and four deployment models. Spanning two pages of text, it initially seems overwhelming. Yet this cloud definition is very effective in establishing clear boundaries and scope for cloud computing. It can be used to filter the overly hyped cloud marketing literature to better understand the business value of true cloud services. This white paper examines NIST's cloud definition in detail with real world case study examples to illustrate how it is applicable to today's cloud market landscape.

At the request of the federal CIO Vivek Kundra, NIST was mandated to assist government agencies to adopt cloud computing for their IT operations. As part of their mandate, NIST created multiple working groups to define cloud computing, its architecture, and requirements. In this paper we explore the center core of NIST's cloud definition (document Special Publication 800-145), which has been well accepted throughout the IT industry across vendors, service suppliers, IT organizations, and customers.

Overview of Cloud Definition

The NIST cloud definition is a comprehensive description of the essential defining quality of cloud computing. They define it as:

"Cloud computing is 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 released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models."

This paper will elaborate and further illustrate how this is applicable in the current environment of the cloud computing industry. In the next several sections, we will state the NIST definition, elaborate on the key principles, and provide some case study examples.

Cloud Characteristics

There are five key attributes of a true cloud service. While there may be some variations in certain cases or environments, a cloud service should adhere to these traits.

NIST's five essential characteristics are:

1. **On-demand self-service**
   NIST defines this as:

   "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."

   This is a key attribute of a true cloud service. A customer must be able to request the usage of a cloud service through an automated interface (such as a web portal, kiosk, mobile app, etc.) without the need to speak with a middleman or sales person. The consumer can request this at any time. This feature should also enable the consumer to cancel the usage of a cloud service at any time. From the consumer's perspective, engaging a cloud service and releasing a cloud service should be as convenient and hassle free as possible. For example, there should be no need to speak with a call center representative or request/release a cloud service only during working hours.

2. **Broad network access:**
   NIST defines this as:

   "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)."

   A true cloud service must be accessible and usable through a broadly available communication network. Generally speaking, it means that as long as a consumer has Wi-Fi, broadband, or landline network connectivity, then he/she can utilize the cloud service. There should be no location dependency for the cloud service. Furthermore, a cloud service should be accessible with minimal dependency on the device used for accessing the cloud service.

3. **Resource pooling:**
   NIST defines this as:

   "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. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth."

   The underlying resources in a cloud service are shared across multiple customers. This multi-tenancy model has certain privacy and security concerns that is shared by all cloud users, therefore, all users must take the necessary precautions and risk-management activities for protecting and guarding their assets, be it data or otherwise.

4. **Rapid elasticity:**
   NIST defines this as:

   "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 to be unlimited and can be appropriated in any quantity at any time."

   A powerful attribute of a cloud service is that it can scale up or down as required automatically and in real-time (or near real-time). This means that varying workloads will be met with the right level of resource capacity (CPU power, storage, network bandwidth, etc.), adjusting to real-world demands from end users.

5. **Measured service:**
   NIST defines this as:

   "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."

   A measured cloud service provides the underpinning for the pay-as-you-go model. This allows a cloud provider to charge consumers for only the resources or services that are actually consumed by the customer. The old model of having a fixed IT budget that pays for IT resources regardless of whether they are underutilized or over utilized, which no longer applies in cloud computing.

Cloud Service Models

NIST defines three general yet distinctive cloud service models. In practice, there are other service models available. However, even these additional service models are a variation or combination of these three basic service models. For each service model, we state the NIST definition, elaborate on key principles, and illustrate the service model with three real-world case studies.

SaaS Model

The SaaS service model is defined as:

"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."

The SaaS service model generates the most interest to business users and managers. Through SaaS-based applications, IT and business units can focus on supporting and enabling business operations and functions. The SaaS model manages the underlying software and IT infrastructure. IT is released from the day-to-day activities of running a data center, IT operations, and maintenance.

1. **Salesforce.com**
   Salesforce.com is one of the pioneers of enterprise-quality SaaS cloud providers. Its service offerings displace the traditional data center based CRM application such as Siebel and PeopleSoft. Headquartered in San Francisco, Salesforce.com has revenue of over $4.07 billion. Initially started by four former Oracle executives, Salesforce.com has grown to 13,400 employees (as of May 2014). Currently, it has 2,100,000 service subscribers across 104,000 customers (as of 2011). Its subscription model is based on a per-user basis, per month. Presently, there are five subscription levels: Contact Manager level ($5/user/month); Group level ($25/user/month); Professional level ($65/user/month); Enterprise level ($125/user/month); and Performance level ($300/user/month). All subscriptions are billed annually.

2. **Gmail**
   Gmail is a free cloud-based email service from Google. It was initially launched as an invitation only service in April 1, 2004. It became a production quality service available to the general public on Feb 7, 2007. As of June 2014, it boasts over 500 million users. Its distinguishing feature has been its ever growing space for email users. When it launched, Gmail stunned the world by offering 1GB of email space. The competition at the time (such as Hotmail and Yahoo!) quickly followed suit by increasing their email space from an initial storage space of only 2-4MB of email. Today, Gmail has evolved into a critical component of Google's overall office productivity SaaS offering, which now includes 30GB of disk storage and can support email file attachments of 25MB.

3. **Intuit - QuickBooks**
   Intuit is an example of a software provider that offers both conventional software (purchased software licenses for self-hosting) and SaaS-based online offerings. Its beginnings started as a traditional software provider in 1983 in Palo Alto, CA by its two founders, Scott Cook and Tom Proulx. Today, Intuit has over 8,200 employees with over $4.1 billion in revenue (2013). It offers a range of financial accounting and tax software. Its flagship product is QuickBooks, a software suite for managing business financial operations. As a SaaS offering is it available in three different levels. Each level offers an increasing set of functions and features. The three levels and subscription pricing are: Simple Start ($9.95/month for 1 user); Essentials ($14.95/month for 3 users); and Plus ($24.95/month for 5 users). Subscriptions are billed monthly. Intuit's overall SaaS offerings now represent $1.5 billion of its 2013 revenue.

PaaS Model

The PaaS service model is defined as:

"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."

The PaaS service model provides a powerful development platform for software developers. Most PaaS providers support a range of programming languages for developers to use. All leverage Web Services (i.e., web-based API) to create cloud-based application. The underlying infrastructure is opaque to the developer as it is fully managed by the PaaS provider.

1. Google App Engine

As a platform as a service, App Engine allows software developers to write applications in one of several languages (Python, Java, and its derivatives, Go and PHP) and are run over a scalable, elastic infrastructure. Each Google application is executed in a "sandbox" environment, protecting it and other Google applications from interfering with each other or monopolizing the underlying infrastructure resources. As a platform, it was released initially on April 7, 2008 in "preview" mode, and then came out of preview mode on September 2011.

Google does not charge developers for writing software on App Engine; they charge only for the resources that applications consume when they are executed and running on Google's infrastructure. Google tracks, monitors, and invoices applications based on resources such as storage, virtual servers, network traffic, and API calls.

App Engine offers a free tier, which means that an application can run on App Engine as long as it stays within limits of infrastructure resource utilization. For example, a Google app can use up to 1GB of data storage. Storage above that limit will incur expenses. For paid applications, there are no functional limits. Hard limits, however, do exist to prevent misuse (intentionally or unintentionally) of resources. For paid application usage, Google provides an uptime guarantee of 99.95% on a monthly basis. Some key customers include Rovio and Khan Academy.

2. Azure

Microsoft presents its Azure platform as a tightly grouped IaaS and PaaS offering that is integrated with Microsoft's traditional product line of Windows Servers, SQL, SharePoint, Active Directory, and BizTalk. It was initially released on February 1, 2010 and was called "Windows Azure." Recently, it was renamed to Microsoft Azure (March 25, 2014). The PaaS development offering supports several programming languages: .NET, Java, PHP, Node.js, Python, and Ruby.

The scope of services within the Azure PaaS offering is broad, and includes services that other cloud providers would consider IaaS. More specifically, the Azure app service (PaaS) includes:

Media Services, Service Bus, Notification Hubs, Scheduler, Automation, BizTalk Services, Visual Studio Online, Active Directory, Multi-Factor Authentication, CDN, API Management, and RemoteApp. For the underlying infrastructure, Microsoft uses a customized Windows OS and Hyper-V hypervisor called Microsoft Azure and Microsoft Azure Hypervisor, respectively. Scalability is managed through a fabric layer called Microsoft Azure Fabric Controller.

The subscription model is billed on the app service that is used. It offers a pay-as-you-go monthly plan, as well as a prepaid six-month and 12-month plan. The longer the term, the greater discount Microsoft provides (up to 32%). Azure provides an SLA of 99.9% per month for its PaaS services. Key customers include BMW, NBC Sports, and HarperCollins.

3. Force.com

Force.com is a PaaS offering from Salesforce.com. It was the underlying foundation upon which the Salesforce.com software was developed and hosted. With its 2013 acquisition of ExactTarget (marketing SaaS) and a previous acquisition of Heroku (PaaS provider) in 2010, the three environments are in the process of merging and are collectively referred to as Salesforce1 Platform. This merged platform is a comprehensive integrated environment for building enterprise-quality applications that are primarily focused on supporting business sales and marketing operations and have greater explicit support for mobile platforms. Their main website is developer.salesforce.com.

The subscription model for the Force.com-specific platform has three levels. Unlike Google App and Azure, the units of measure are less focused on infrastructure consumption and more on programming objects and interfaces. Their three subscriptions levels are: Enterprise App level ($25/user/month); App Bundle ($80/user/model); and Unlimited App level ($150/user/month). All subscriptions are billed annually. Each level has increasing usage of application objects, supported apps, API calls, and so forth.

Key users and clients of Force.com that have integrated their own applications with Saleforce.com software via an integration application include Evernote, DocuSign, and MailChimp.

IaaS Model

The IaaS service model is defined as:

"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)."

The IaaS model allows IT administrators to operate and manage traditional data center resources in the cloud. All of the traditional IT infrastructure layers (servers, storage, network, and so forth) are available virtually from a cloud IaaS provider. Customers are responsible for managing the OS layer and higher, while the cloud provider focuses on the hypervisor layer and below (server hardware, network connectivity, power/cooling, HVAC, etc).

Case Study Examples of IaaS Cloud Providers

1. Amazon Web Services

Amazon Web Services (AWS) is an operation of Amazon.com that provides cloud services (primarily IaaS and PaaS) to the general public. AWS has been a pioneer in the commercial cloud computing space and is currently the market leader. Recently Gartner positioned AWS at the highest level in its Magic Quadrant for Cloud Computing (June 2014).

AWS was launched in 2006. Its first major service offering was EC2 (elastic cloud compute), a virtual server offering. It launched on Aug 25, 2006 (Linux-based version). It was initially in public-beta mode, and then went into production mode on Oct 23, 2008 with formal SLAs.

AWS currently has a large selection of IaaS component offerings. Some of the major services include S3 (file object storage), EBS (block storage), ELB (load balancer), Route 53 (DNS service), VPC (virtual private cloud–networking and subletting), Glacier (tape archival).

AWS global infrastructure covers 10 regions across the world, each one with multiple data centers (Availability Zones) to provide redundancy and resource distribution. Due to security concerns, AWS provides a dedicate cloud environment uniquely for US government agencies and partners called GovCloud.

Each IaaS component has its own pricing model with a variety of tiers. Generally, these are all based on resources consumed (on a pay-as-you-go basis) or for a guaranteed level of performance. For example, the EC2 virtual server service has price ranges from the low end (t2.micro) for $0.013/hour to the high end (c3.8xlarge) at $1.68/hour.

Some of AWS major customers include: Netflix, Dropbox, Samsung, NY Times, Washington Post, Newsweek, and Adobe.

2.   Rackspace

Rackspace is an IaaS cloud provider based in San Antonio, TX. It started out as a traditional web-hosting company and has grown into a leading cloud IaaS provider. Rackspace was founded in October 1998 by three individuals. Currently, it employs over 5,700 employees and has revenue over $1.5 billion 2013.

Its underlying architecture is based on Xen open source virtualization and OpenStack. Rackspace has been a key contributor to the OpenStack foundation and was a co-founder with NASA.

Rackspace IaaS offerings cover the following components: cloud servers, block storage, databases, file storage, load balancers, backups, and cloud monitoring. Each has its own pricing model with a variety of tiers, and support both pay-as-go-you and long-term discounted plans. For example, a low end cloud server starts at $0.04/hour and a high end server is priced at $5.44/hour. Key customers of Rackspace include: BitHub, Domino's, KarmaCRM

3.   CenturyLink

CenturyLink is a major telecommunication company that has entered into the cloud computing industry via two key acquisitions, Savvis in April 2011 and Tier 3 in Nov 2013. It recently (Jan. 2014) rebranded these entities into CenturyLink Cloud (CLC).

CenturyLink Cloud's IaaS offerings span a wide range of IT components from servers and storage to network, firewalls and Content Distribution Network (CDN). Pricing is based on a per-as-you-go model. For long term plans, a sales quote from their sales team is required. For a basic virtual Linux server, the cost can be as low as $0.01/hr. Some key customers of CenturyLink Cloud include exterro, Obeo, and XSP.

Cloud Deployment Model

NIST outlines four cloud deployment models. For each deployment model, we state NIST's definition, elaborate on key principles, and provide two case study examples. In all four models, there are three determining factors: who controls security, who has access to the data, and whether the resources are shared or dedicated.

Public Cloud

A Public cloud deployment is defined as:

"The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider."

For public cloud services, it is assumed that it is multi-tenant and the underlying resources are shared among multiple customers. The public cloud provider owns and controls the security and protection of data between one customer and another customer.

Case Study Examples of Public Cloud Deployment

All of the case study examples from the Service Model section (Salesforce.com, Gmail, Intuit, App Engine, Azure, force.com, Amazon Web Services, Rackspace, CenturyLink) are examples of public cloud providers.

Private Cloud

A Private cloud deployment is defined as:

"The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises."

In private cloud deployment, the IT assets are fully dedicated to a single company, with no sharing of resources outside of the corporate entity. Security and data protection is owned by the same business entity.

Case Study Examples of Private Cloud Deployment

1) OpenStack

OpenStack is an open source cloud platform that provides private cloud computing services for IaaS model. Some companies use it to deliver IaaS services internally to their business units. As a cloud technology, it can also be used by organizations to deliver IaaS to the public.

OpenStack started in July 2010 as a joint project by NASA and Rackspace. Ownership and management of its development was transitioned in September 2012 to the OpenStack Foundation, a non-profit entity. Support for OpenStack has grown tremendously and numerous vendors committed their support for its ongoing development. Some vendors include: VMware, Red Hat, HP, IBM, EMC, and Oracle.

OpenStack uses a very modular structure in which each component delivers a specific IaaS resource. Some of these are: Compute (Nova), Object Storage (Swift), Block Storage (Cinder), Networking (Neutron), Orchestration (Heat), and Database (Trove). Some key companies who use OpenStack internally include Intel, Argonne National Laboratory, CERN, and NeCTAR.

2) vCloud

vCloud Suite is a private cloud technology platform from VMware. Customers use the vCloud Suite to implement a private cloud computing IaaS type environment. The vCloud Suite consists of several core products: vSphere, vCenter Site Recovery Manager, vCloud Network and Security, Automation Center, Operations Management Suite, and vCloud Director. Together, these key VMware products enable a cloud-based virtual data center. IT organizations can deliver cloud services to their internal business units with the same scalability and pay-as-go-go capabilities that public cloud providers deliver.

The vCloud Suite is available in three configurations: Standard, Advanced, and Enterprise. Key customers of vCloud include: Columbia Sportswear, Catholic Health Initiatives, and Microstrategy.

Community Cloud

A community cloud deployment is defined as:

"The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises."

This deployment is effective for consortium groups and special interest user groups. Generally, security and data access between members of a consortium or user group is permitted. However, outside of the consortium or user group, access is restricted.

Case Study Examples of Community Cloud Deployment

1) Facebook

Facebook is the largest social networking service with a community user base of over 1.28 billion (as of March, 2014). It was started by Mark Zuckerberg and four other Harvard college mates in February 2004. It is now headquartered in Menlo Park, CA and employs over 6,800 employees. As a community cloud service, there is no subscription fee to join Facebook. All revenue for Facebook is through advertising. As of 2012, revenue was $7.87 billion.

A major challenge with Facebook, from a community perspective, is privacy. General Internet users who are not subscribed to Facebook do not have viewable access to the information in Facebook. Subscribed Facebook users have varying levels of access to data about other Facebook users. While users have control over their privacy settings, the default settings and types of control level are sometimes changed independently by Facebook. Issues about ownership of photos and information about one's self have caused concerns among some Facebook users. Most recently (July 2014), news about Facebook conducting psychological experiments and studies on Facebook subscribers, without direct explicit consent, caused significant outcry from the public.

2) LinkedIn

LinkedIn is a professional business social network service. It was launched on May 5, 2003 by five founding members. It is based in Mountain View, CA and has 5,000 employees. As of June 2013 there are over 259 million users across 200 plus countries.

Subscription to LinkedIn is free for all users. There are paid subscription levels that provide additional features and access to more information. In 2013, LinkedIn had revenue of $1.52 billion. There are three levels of paid subscriptions: Business level ($29.99/month); Business Plus level ($59.99/month); and Executive level ($99.99/month). If users pay annually, LinkedIn provides a 25% discount. Other sources of revenue include paid advertising, directory services, and recruiting services.

Hybrid Cloud

A Hybrid cloud deployment is defined as:

"The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds)."

There are several use cases or cloud configurations that employ a hybrid cloud deployment model. One is cloud bursting, which means going from private cloud to public cloud. Another is backup and disaster recovery, also going from private cloud to public cloud.

Case Study Examples of Hybrid Cloud Deployment

1) AWS - Virtual Private Cloud (VPC)

Amazon Web Services (AWS) provides a cloud service called Virtual Private Cloud (VPC), which allows a customer to extend their data center into AWS' cloud infrastructure. For example, with an AWS VPC, a customer can run their application servers in AWS and have all their data stored in their own data center and storage devices. This hybrid configuration gives the client control and ownership of data security and data protection while enabling full scalability of server resources for the application servers running in the AWS cloud.

2) Bluelock

Bluelock's cloud services is another example of how an IT organization can extend their data center resources and operations to the public cloud.

Bluelock is a VMware partner that enables IT organizations that are VMware based to extend their current data center into a VMware technology-based public cloud provider. Bluelock can provide backup in the cloud, recovery to the cloud, or recovery from a cloud provider through their hybrid cloud service model.

Conclusion

With case study examples, we have reviewed NIST's definition of cloud computing in this white paper. Although the definitions may at times appear verbose, they clearly delineate the attributes and features that make cloud computing a true game-changer for IT organizations and businesses. It provides clear boundaries and scope to the cloud computing paradigm and is very effective in sorting out the marketing hype that currently surrounds cloud computing.

Learn More

Learn more about how you can improve productivity, enhance efficiency, and sharpen your competitive edge through training.

Telecom Architectures and Information Technologies

Cloud Essentials

Visit **www.globalknowledge.com** or call **1-800-COURSES (1-800-268-7737)** to speak with a Global Knowledge training advisor.

About the Author

Vince Lo Faso is the Managing Director of Cloud Service Management at Navigo Technologies, LLC. He is an IT Service and Cloud Management professional with more than 24 years of IT industry experience. He is ITIL® V3 Expert certified; Cloud Essentials™ Professional (CEP) certified; and AWS Partner Business and Technical Professional accredited. Vince holds a master's degree in computer science and has spoken as conferences such as VMworld User Conference, HP Universe, and local user groups. In addition to having worked as a consultant and practice manager for several HP VARs, Vince Lo Faso has held IT positions with Kraft Canada, Sprint Paranet, and Concordia University.