NIST Time   NIST Home   About NIST   Contact Us   A-Z Site Index | Search |

Information Technology Laboratory

About ITL   Publications   Topic/Subject Areas   Products/Services   News/Multimedia   Programs/Projects

NIST Home > ITL > Computer Security Division > Final Version of NIST Cloud Computing Definition Published

Final Version of NIST Cloud Computing Definition Published

From NIST *Tech Beat*: October 25, 2011

Select Language ▼ *
Powered by Google Translate

🔗 SHARE f ✏ ✉ …

Contact: Evelyn Brown
301-975-5661

After years in the works and 15 drafts, the National Institute of Standards and Technology's (NIST) working definition of cloud computing, the 16th and final definition has been published as *The NIST Definition of Cloud Computing* (NIST Special Publication 800-145).

Cloud computing is a relatively new business model in the computing world. According to the official NIST definition, "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."

The NIST definition lists five essential characteristics of cloud computing: on-demand self-service, broad network access, resource pooling, rapid elasticity or expansion, and measured service. It also lists three "service models" (software, platform and infrastructure), and four "deployment models" (private, community, public and hybrid) that together categorize ways to deliver cloud services. The definition is intended to serve as a means for broad comparisons of cloud services and deployment strategies, and to provide a baseline for discussion from what is cloud computing to how to best use cloud computing.

"When agencies or companies use this definition," says NIST computer scientist Peter Mell, "they have a tool to determine the extent to which the information technology implementations they are considering meet the cloud characteristics and models. This is important because by adopting an authentic cloud, they are more likely to reap the promised benefits of cloud—cost savings, energy savings, rapid deployment and customer empowerment. And matching an implementation to the cloud definition can assist in evaluating the security properties of the cloud."

While just finalized, NIST's working definition of cloud computing has long been the de facto definition. In fact before it was officially published, the draft was the U.S. contribution to the InterNational Committee for Information Technology Standards (INCITS) as that group worked to develop a standard international cloud computing definition.

The first draft of the cloud computing definition was created in November 2009. "We went through many versions while vetting it with government and industry before we had a stable one." That one, version 15, was posted to the NIST cloud computing website in July 2009. In January 2011 that version was published for public comment as public draft SP 800-145.

Researchers received a large amount of feedback, which mainly dealt with interpretations. The definition from draft to final remained substantively the same and only a modest number of changes were made to ensure consistent interpretations. *The NIST Definition of Cloud Computing* (SP 800-145) is available at http://csrc.nist.gov/publications/PubsSPs.html#800-145.

The National Institute of Standards and Technology (NIST) is an agency of the U.S. Department of Commerce.

Privacy Policy / Security Notice / Accessibility Statement / Disclaimer / Freedom of Information Act (FOIA) / Environmental Policy Statement / No Fear Act Policy / NIST Information Quality Standards / Scientific Integrity Summary

✉ Sign Up for NIST E-mail alerts:
Enter email address | Go

Date created: October 25, 2011 | Last updated: October 25, 2011   Contact: **Webmaster**

NIST

National Institute of
Standards and Technology
U.S. Department of Commerce

The NIST Definition of Cloud Computing

Recommendations of the National Institute of Standards and Technology

Peter Mell
Timothy Grance

NIST Special Publication 800-145

The NIST Definition of Cloud Computing

Peter Mell
Timothy Grance

C O M P U T E R   S E C U R I T Y

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analysis to advance the development and productive use of information technology. ITL's responsibilities include the development of technical, physical, administrative, and management standards and guidelines for the cost-effective security and privacy of sensitive unclassified information in Federal computer systems. This Special Publication 800-series reports on ITL's research, guidance, and outreach efforts in computer security and its collaborative activities with industry, government, and academic organizations.

Acknowledgements

Errata

The following changes have been incorporated into Special Publication 800-145, as of the date indicated in the table.

DATE	TYPE	CHANGE	PAGE NUMBER
4/27/2012	Editorial	Corrected page number from "2" to "1"	1

1. Introduction

1.1 Authority

The National Institute of Standards and Technology (NIST) developed this document in furtherance of its statutory responsibilities under the Federal Information Security Management Act (FISMA) of 2002, Public Law 107-347.

NIST is responsible for developing standards and guidelines, including minimum requirements, for providing adequate information security for all agency operations and assets; but such standards and guidelines shall not apply to national security systems. This guideline is consistent with the requirements of the Office of Management and Budget (OMB) Circular A-130, Section 8b(3), "Securing Agency Information Systems," as analyzed in A-130, Appendix IV: Analysis of Key Sections. Supplemental information is provided in A-130, Appendix III.

This guideline has been prepared for use by Federal agencies. It may be used by nongovernmental organizations on a voluntary basis and is not subject to copyright, though attribution is desired.

Nothing in this document should be taken to contradict standards and guidelines made mandatory and binding on Federal agencies by the Secretary of Commerce under statutory authority, nor should these guidelines be interpreted as altering or superseding the existing authorities of the Secretary of Commerce, Director of the OMB, or any other Federal official.

1.2 Purpose and Scope

Cloud computing is an evolving paradigm. The NIST definition characterizes important aspects of cloud computing and is intended to serve as a means for broad comparisons of cloud services and deployment strategies, and to provide a baseline for discussion from what is cloud computing to how to best use cloud computing. The service and deployment models defined form a simple taxonomy that is not intended to prescribe or constrain any particular method of deployment, service delivery, or business operation.

1.3 Audience

The intended audience of this document is system planners, program managers, technologists, and others adopting cloud computing as consumers or providers of cloud services.

2. The NIST Definition of Cloud Computing

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.

Essential Characteristics:

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. 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.

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 to be 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[1] 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.

Service Models:

Software as a Service (SaaS). The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure[2]. 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

---

[1] Typically this is done on a pay-per-use or charge-per-use basis.

[2] A cloud infrastructure is the collection of hardware and software that enables the five essential characteristics of cloud computing. The cloud infrastructure can be viewed as containing both a physical layer and an abstraction layer. The physical layer consists of the hardware resources that are necessary to support the cloud services being provided, and typically includes server, storage and network components. The abstraction layer consists of the software deployed across the physical layer, which manifests the essential cloud characteristics. Conceptually the abstraction layer sits above the physical layer.

languages, libraries, services, and tools supported by the provider.[3] 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).

Deployment Models:

Private cloud. 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.

Community cloud. 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.

Public cloud. 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.

Hybrid cloud. 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).

---

[3] This capability does not necessarily preclude the use of compatible programming languages, libraries, services, and tools from other sources.