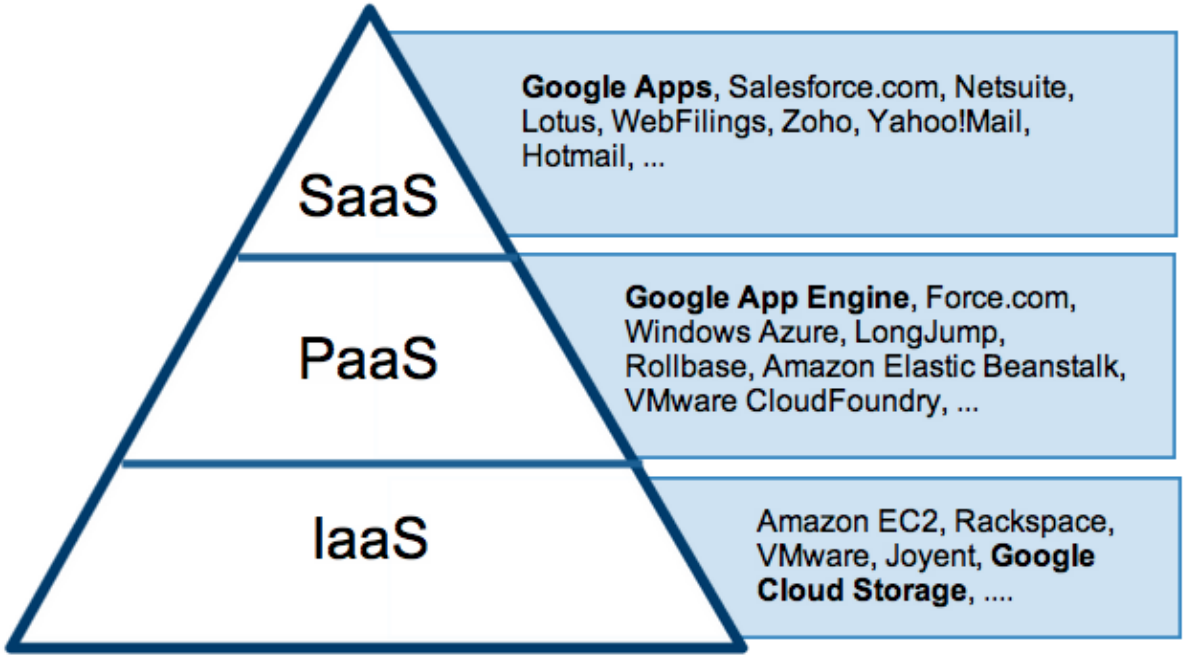
Cloud Computing & Cloud Models



Source: Gartner AADI Summit Dec 2009

Cloud Models

Topics

- Defining cloud computing
- Understanding:
  - Distributed Application Design
  - Resource Management Automation
  - Virtualized Computing Environments
  - High-performance Computing Models
  - Cloud computing technologies

Cloud Computing

- Refers to flexible self-service, network-accessible computing resource pools that can be allocated to meet demand.
  - Note for this class, we will use the definition from NIST.
- Allows resource allocation of to be adjusted on demand.
  - Hardware-independent.
- Services flexible because resources and processing power adjust to meet changes in need without the need for direct IT personnel involvement.

Intro

- Data and services easily available via cell phones, tablets, and other mobile devices.
  - Easily serviced from the cloud.
- Cloud costs billed as an operational expense (like electricity)
  - Converts capital cost to operational cost.
  - Eliminates traditional up-front infrastructure capital cost.
- Cloud allows rapid application deployment
  - Rapid scale-up
  - Eliminates many current infrastructure issues

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

Special Publication 800-125

Guide to Security for Full Virtualization

Cloud Computing Defined

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. --NIST SP125

NIST Cloud Attributes

Five essential characteristics

Three service models

Four deployment models.

NIST Cloud Attributes

Essential Characteristics

1. On-demand self-service.
2. Broad network access.
3. Resource pooling
4. Rapid elasticity.
5. Measured service.

NIST Cloud and Service Models

Three Service Models

1. Software as a Service (SaaS)
2. Platform as a Service (PaaS)
3. Infrastructure as a Service (IaaS)

Four Deployment Models

1. Private Cloud
2. Public Cloud
3. Hybrid Cloud
4. Community Cloud

Selected Cloud Attributes

- Pay Only for What You Use (On demand)
- Cloud Bursting
  - Automatic failover to public cloud services when local resources insufficient
- Resource Pooling

Selected Cloud Attributes Two

- Clouds Virtualize Application Development Cycle
  - Improves business agility by making services immediately available with whatever resources are needed..

Common Cloud Computing Solutions Characteristics

Provider Managed

- Once applications and services move to external cloud, an organization no longer needs to deal with local data center issues (power, space, cooling et. al.).
- Developers need only know whether their applications will be running on one cloud service platform or another.
- Individual hardware characteristics and capacity measures no longer relevant.
  - Similarly, tech refresh and update becomes a for the cloud provider issue.

Resource Pooling

- Capacity and resources available to cloud computing services can be increased or decreased on demand.
  - Organization to spin up a new offering with minimal costs
  - Meet cyclic use patterns with increased capacity, paying for only needed resources.
- In contrast, traditional data centers must always plan and budget for future growth.
- Cloud computing draws resources from a pool as needed.
  - Similar to electric power companies.

Broad Network Access

- Cloud services available via networked devices and technologies, facilitating rapid access by mobile customers and remote locations.
- Provides "anywhere, anytime" service model not possible with traditional data centers, where service downtime and local-area outages in power and networking can impact uptime.
- Cloud computing vendors can be located anywhere in the world.

Sustainable

- Since cloud providers provision resources dynamically, during off-peak times it is possible to reduce power and cooling requirements.
- Flexibility in cloud hosting location allows providers to move data center activity north during summer months to save on cooling costs or transfer operations to areas with excess power production capability.

On Demand Self-service

- After limits for resource availability are configured within the cloud provider's systems, available resource capacity can be automatically expanded or managed by client.

Distributed Application Design

- Distributed design is a fundamental supporting cloud computing technology.
  - In contrast, early software operated on a single system, together with its data and ancillary programs.
- In cloud computing environments, even the location and type of hardware supporting a software application can shift from moment to moment as additional capacity is allocated or services are transferred between cloud provider data centers. …
- Cloud interconnected through standard APIs and XML web service interfaces, allowing developers to rapidly move their applications into the cloud without requiring a completely new set of skills.

Lack of Standardization

- APIs vary from one cloud provider to another, so applications developed under Amazon's EC2 will not be able to directly transfer to Microsoft's Azure

Understanding Resource Management Automation

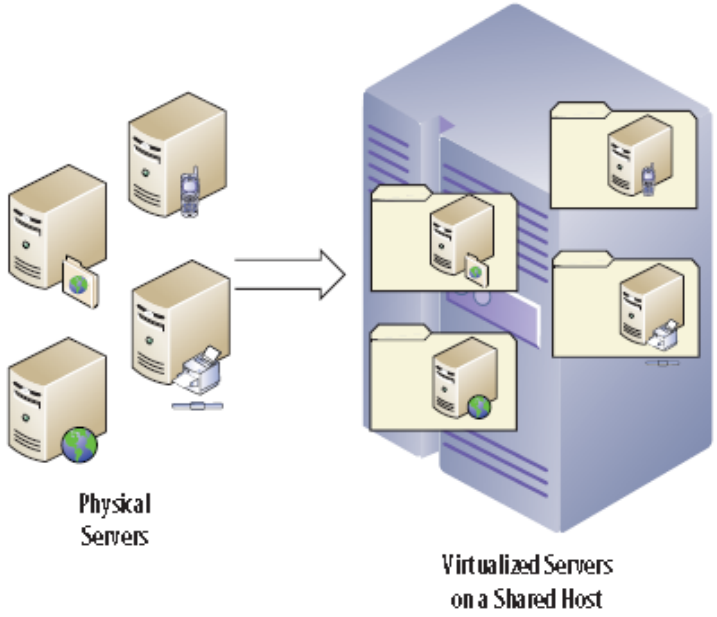
- Automatic management of resources.
- When demand nears capacity, the cloud hosting software is able to identify need and respond by adding resources up to predetermined levels based on an organization's contractual limits or limits configured in the management software.
  - Protects application availability.
- During off-peak periods, automatically reduces resource allocation.

Understanding Virtualized Computing Environments

- Virtualization of storage systems, storage area networks (SANs), and virtualization of computer systems form cloud computing's backbone.
- Because an organization no longer needs to worry about where data is located or what hardware resources are available, organization can focus on increasing business competitiveness.
- Hardware independence.
- Geographic independence.

Virtualized Computing: Multitenancy

- Allows multiple systems to run on a more powerful server, as shown is referred to as multitenancy.
- Allows system resources to be more fully utilized, further reducing operating costs and data center power and cooling requirements.



Understanding High-Performance Computing Models

- Cloud computing also borrows from high-performance computing (HPC) techniques for separating individual procedures into multiple simultaneous processes. (Decentralization)
- Individual results are combined later to provide the complete final result, as shown.

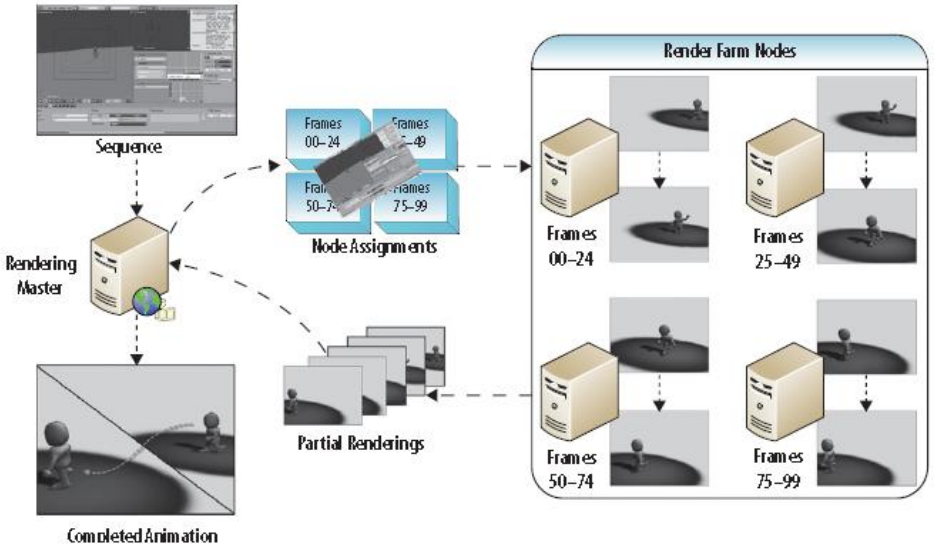


FIGURE 1.3 Rendering a complete video sequence using multiple computers simultaneously

Understanding Cloud Computing Technologies

Three different levels of cloud services:

1. Software as a Service (SaaS)
2. Platform as a Service (PaaS) development environments
3. Infrastructure as a Service (IaaS)

Cloud computing services run atop hosting virtualized hardware servers and are accessed via the network, making them available to many types of clients including:

- Workstations
- Thin clients
- Mobile clients

Cloud Services Becoming Commonplace

- For example, some organizations utilize: Salesforce CRM operating alongside Google Apps for user productivity while Azure SQL database applications power business applications running in Amazon S3 cloud services, with Iron Mountain providing cloud backup and recovery
  - Being accessed using iPads, whose automatic integration with cloud-based Dropbox, Flickr, and social media services improve customer interaction.

Cloud Models

Topics

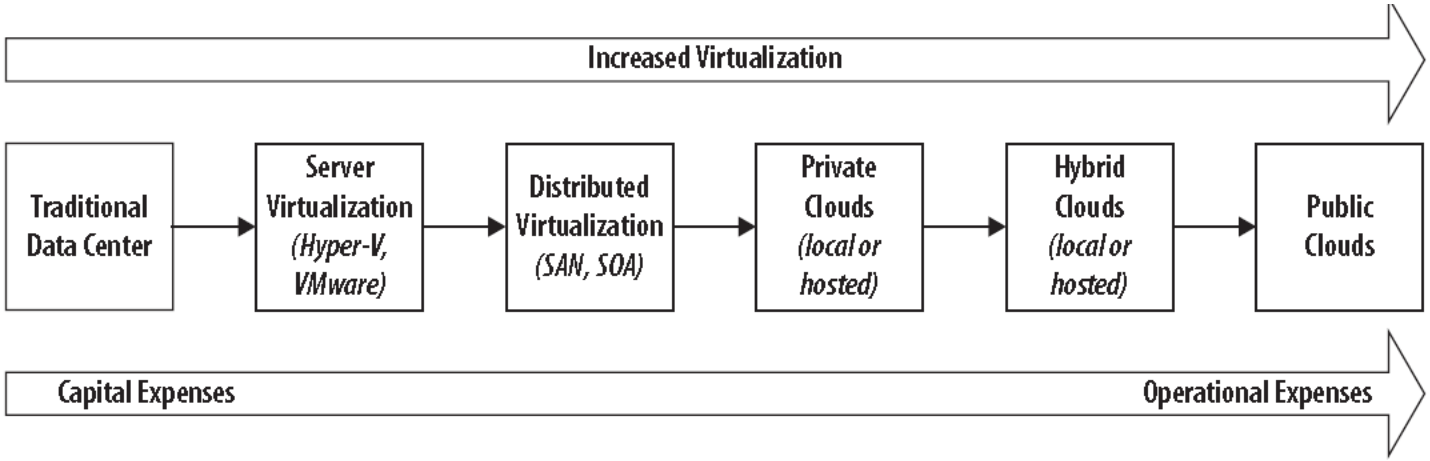
When planning cloud computing deployments, enterprise architects and network planners need to be able to identify expectations for control and management based on the type of cloud and its level categorization.

Topics

- Cloud computing models relating to strategies for extending virtualization into the cloud.
- Evolving from virtualization to the cloud
- Planning cloud organization roles
- Identifying cloud deployment models and scope modifiers
- Future cloud models

Organizational Evolution

- Evolution of traditional data center infrastructure into the cloud process starts with server virtualization, moves through privately hosted and hybrid clouds into fully public cloud infrastructures.



Data Center Virtualization

- Data centers rapidly deploying virtualization technologies
  - Consolidates server resources.
- Allows organizations to concentrate on a smaller number of physical hosts
  - Requires less power and less cooling.
- Hardware independent.
- Improves disaster recovery and business continuity capabilities
  - Virtualized servers can be simply moved to a new host site and brought online to return normal operational capabilities.

Distributed Virtualization

- By extending virtualization to include distributed resources organizations can increase the flexibility of their server infrastructure as well as operational resources available to virtual machines.

Technologies that have improved distributed virtualization flexibility include:

- Virtualization of data storage across distributed storage area network (SAN) infrastructures
- Interoperation of application component services through service-oriented architectural (SOA) integration
- Automatic load-management that can migrate virtualized server instances from one host to another based on total resource load

Private Clouds

- Private cloud resides on hardware located in local data center but running cloud infrastructural software.
  - Enables organization to take advantage of self-service resource allocation and consumption metering for cost recovery billing models.
- Even when hardware remains heterogeneous, cloud software provides a standard platform for application development.
  - Transforming IT toward a utility business model.
  - Allocates resources based on service performance rather than on projections of planned resource needs.

Private Clouds

- While, mild efficiencies of scale can be achieved, costs, both capital and operational, remain.

Hybrid Clouds

- As organizations continue to transition, they can bridge local private clouds to create hybrid clouds
  - Allows better response to peak loads and unanticipated demands.
- Billing continues to develop along the utility model, allowing load to determine cost as operational expenses and internal billing for cost recovery.

Hybrid Clouds

- Allow organizations to retain control over critical data resources while transferring less-sensitive operations to more efficient public cloud.
  - Capital expenses reduced because only key services are retained as local server resources.
- Model for mid-transition between local and public cloud services.
  - Allows developers to test applications using local resources with very low latency and locally controlled high-capacity networking.

Public Clouds

- Public clouds operate like public power production systems
  - Bring industrial-scale/cost efficiencies and hosting location flexibility to an organization
  - Following the utility model transformation from building power generation to the distributed power grid.
- Obviously, the trail from servers to virtualized public cloud computing will process at different rates as individual services are transformed and migrated, so the total elimination of data center resources is not a logical near-term target.

Resistance Due to Perceived Loss of Control

- Perception "if you cannot touch it, you no longer own it."
  - Holdover from mainframe computing where all technologies were held in secure, closed central data centers.
- To meet cloud computing initiatives, infrastructure and operations staff must evolve their skill sets along with the organization's transformation to remain viable in the new configuration.

Cloud Deployment Models

- National Institute of Standards and Technology (NIST) has published a definition of cloud computing that CompTIA uses in their Cloud Essentials exam.
- NIST Special Publication 800-145 documents four models for cloud deployments:

1. Private
2. Community
3. Hybrid
4. Public.

Private Clouds

- Provisioned for use by a single user or group within an organization.
- Owned, managed, and operated by the organization.
- Reside on a private network owned or managed by the organization.
- Often selected when external mandates such as regulations and legislative requirements require a high degree of access accountability, control, and governance.

Community Clouds

- Provisioned for use by a group of related organizations with shared concerns
  - Such as a group of governmental or educational institutions that choose to share a common cloud of services not available to the general public
  - Community clouds may reside as local, private cloud resources for the hosting organization and be accessed remotely as a community cloud by its partner organizations.
- Partitioned public clouds are examples of community clouds, with public cloud services isolated from general consumption through limitations restricting access to specified network address schemes or other forms of access specification.

Public Clouds

- Provisioned for use the general public, public cloud services represent the most thoroughly virtualized cloud infrastructural design, removing data center information resources partially or completely.
- Public clouds reside on hosting data center resources and are accessed via public Internet connectivity by users located anywhere in the world.
- Transparent redirection of public cloud services to data centers in variable locations presents concerns for organizations with regulatory or legislative mandates demanding data accountability and governance.

Hybrid Clouds

- Provisioned using components of private, community, or public clouds, the hybrid cloud provides access to two or more infrastructures bridged by standardized technologies or proprietary cloud services.
- Hybrid clouds are simply a mixture of cloud types, such as a private cloud customer relationship management (CRM) application together with public cloud Google Apps services used to integrate CRM data into an organization's collaboration services.

Cloud Bursting

- A hybrid cloud implementation where local private cloud resources support of an application until a spike in demand exceeds local resource limits, at which point the app "bursts" out of the private cloud into designated public cloud resources to manage the overrun.
- For example, a tax preparation service might experience a volume in when its software is discounted to end users, creating a flood of sudden new clients over a short time and overrunning private cloud capacities available in its organizational data center.

Onsite Private Clouds

- When the traditional data center is extended to include cloud services on site, the organization's traditional network and IT support will continue to be involved in cloud support.
- The cloud services conceal operational details such as workload location and multitenancy on individual host systems, but they can provide enhanced control over resource monitoring and flexibility with dedicated virtualization hosts or physical server hosting scenarios.
- Costs may be high if new data centers are required or data center conversion is required for the new private cloud, and local resource constraints will still be present if not coupled to external services for cloud bursting.

Sharing the Same Box

- Multitenancy refers to a particular hosting server sharing workloads from multiple clients or services, which are separated only by access policies configured on the cloud server software.
- Attacks on one service could overwhelm resources available to an unrelated service if multitenancy planning is not imposed to isolate key services.

Outsourced Private Clouds

- All of the traditional outsourcing security issues factor in, such as network bandwidth mandates and the need for transport security between the organization and the outsourcing host data center.
- All of the same limitations from onsite private clouds are present in outsourced private clouds, save that outsourcing host organizations can typically retain a larger resource pool than is present in the onsite data center and will accomplish tech refresh without intervention by the client organization.
- Data center costs reduced for outsourced private cloud implementations, with higher operational costs for the outsourcing itself.

Onsite Community Clouds

- When a private cloud is expanded to provide services to a community of related organizations, it is termed a community cloud.
- The community cloud operates as a private cloud to the hosting organization but as a remote partitioned public cloud to the other organizations in the same community.
- Allowing only a limited scope of requestors access helps to improve the security of community clouds, but resource limitations and high costs are still retained from the private cloud model. Because the community's networks and resource requirements may vary widely from the hosting organization's standards, they can create variable costs in addition to those of the private cloud model.

Outsourced Community Clouds

- Carry the same issues as their onsite community cloud counterparts and gain the same advantages as their outsourced private equivalents — data center costs will be lower, but the outsourcing operating expenses may be higher than for self-hosted alternatives.
- One change from private community clouds is that all organizations will access the outsourced community cloud as a remote partitioned public cloud because no organization in the community will host the outsourced resources.

Public Clouds

- Public cloud models continue the evolution of virtualization, extending the outsourced community cloud to services available to authorized access from organizational, community, and general public security requestors.
- All access will be remote, while operational details such as workload location and multitenancy are concealed beyond the organization's monitoring scope.
- Public clouds typically carry the lowest up-front costs because they rely on existing data centers, creating very large resource pools.
- Although these provide a high degree of elasticity, they require management to ensure that rising demands do not generate unexpected cost overruns.
- Service-level agreements and other contractual agreements also present challenges for the organization when dealing with public cloud services.

Hybrid Clouds

- Hybrid cloud models can bridge any of the previously mentioned models for cloud computing and will include all of the same limitations and advantages of their component models with the additional requirement for standardization and compatibility between onsite, outsourced, and public components (Figure 2.3).

Hybrid clouds require more management than the other models but can allow an organization or community the ability to align resources with business requirements to gain the best solution to meet all of their various needs.

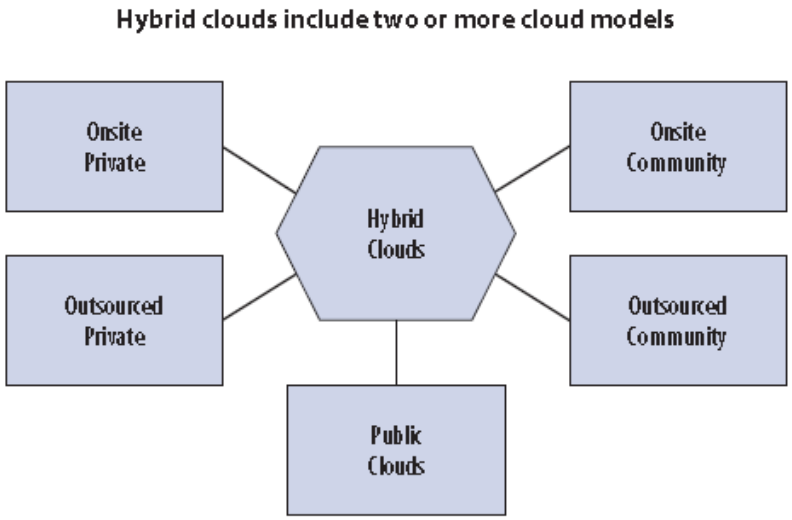


FIGURE 2.3 Hybrid clouds include two or more cloud models and may contain all models in some cases.

Hybrid as a Surface Rather than an Axis

- Hybrid clouds might be built as horizontal hybrid models intended to provide services such as multiple avenues of data access or presentation to different access groups, or they may be constructed as vertical hybrid models that bring together all services required for a particular task, such as the database, web interface, payment application, and shipping management applications supporting an e-purchase service.
- Hybrids can also span both vertical and horizontal hybrid models at once to construct applications layered atop an organization's various needs and products, supporting multiple forms of access and consumption by various groups (private, community, or public).

Future Cloud Models

- As cloud computing matures, additional models will undoubtedly evolve.
- Already, cloud-based services such as disaster recovery and backup are expanding traditional data center and core IT functions into the cloud.
- Users can often access cloud services like Dropbox entirely within their web browsers, bypassing many controls of information provisions in the enterprise environment.
- Because data operations such as workload location and resource pool limits are hidden from common use by cloud infrastructural components, regulatory mandates and legal requirements for accountability and responsibility require additional planning and user training.

Questions???