Assessing the Role of Open Standards

- Open standards determine the format, storage, and exchange of data and enable different organizations and systems to communicate seamlessly.
- Open Standards" are standards made available to the general public and are developed (or approved) and maintained via a collaborative and consensus driven process. "Open Standards" facilitate interoperability and data exchange among different products or services and are intended for widespread adoption.
- Open standards allow others to make compatible products, so they're not locked into just one company's software or hardware. They promote compatibility, interoperability, and innovation.
- **Open standards can help with interoperability**, which is often a requirement for cloud environments.
- Interoperability can be viewed in two dimensions:
- **Horizontal**: Between two different applications
- **Vertical**: Within the components of a single application.

Assessing the Role of Open Standards

- Cloud computing's promise of scalability completely changes the manner in which services an applications are deployed. Without standards, the industry creates proprietary systems with vendor lock-in, sometimes referred to as **"stovepipe" clouds**.

- Clients do not want to be locked into any single system, there is a strong industry push to create standards-based clouds.

The cloud computing industry is working with these **architectural standards:**

- Platform virtualization of resources
- Service-oriented architecture
- Web-application frameworks
- Deployment of open-source software
- Standardized Web services
- Autonomic systems
- Grid computing
- These standards help to enable different business models that cloud computing vendors can support, most notably Software as a Service (SaaS), Web 2.0 applications, and utility computing.
- These businesses require open standards so that data is both portable and universally accessible.

Assessing the Role of Open Standards

- The race to create the first generation of open cloud platform technologies that will compete with proprietary technologies offered by companies such as Microsoft (Azure Platform) and VMware (vSphere) is already underway.

- Rackspace.com, one of the large IaaS cloud service providers, announced in July 2010 that it is initiating an open-source project called OpenStack that will begin with the code used to run its Cloud Files and Cloud Servers technologies.

- NASA has also donated some of the Nebula Cloud Platform technology that it developed. The software developed will be released under the Apache 2.0 license. Founding members of this project include AMD, Citrix, Dell, Intel, NTT Data, and several other cloud service providers.

- openStack.org's home page (http://www.openstack.org/).

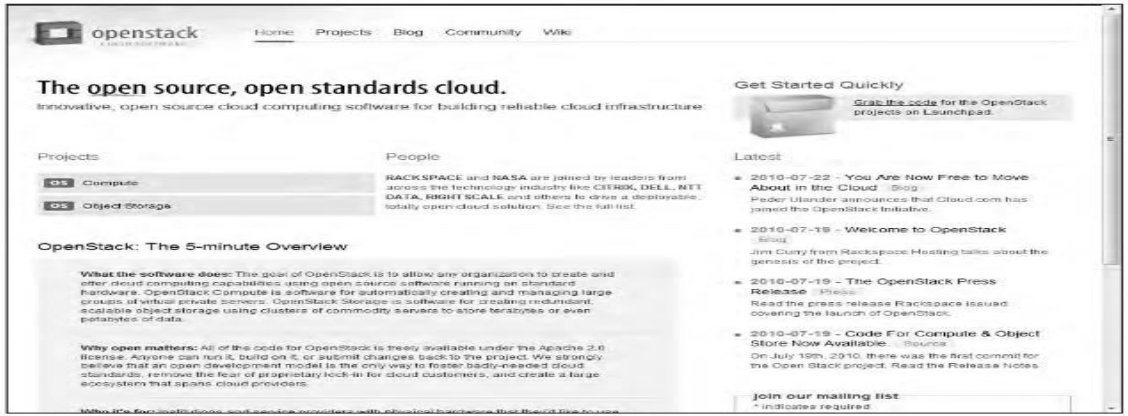
Assessing the Role of Open Standards

- Eucalyptus (http://open.eucalyptus.com/) is a Linux-based software platform for creating cloud computing IaaS systems based on computer clusters.
- The project has an interface that can connect to Amazon's compute and storage cloud systems (EC2 and S3), and it maintains a private cloud as a sandbox for developers to work in. Eucalyptus works with a number of technologies for system virtualization, including VMware, Xen, and KVM. Eucalyptus is an acronym taken from the expression "Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems."
- Most of the major Linux vendors support this project, which is based on the original work of Rich Wolski at the University of California at Santa Barbara.
- The company Eucalyptus Systems was formed in 2009 to support the commercialization of the Eucalyptus Cloud Computing Platform.
- OpenStack and Eucalyptus are by no means unique; several other projects are underway to create open-source cloud platforms. There also are numerous research projects in the area.
- The IEEE Technical Committee on Services Computing (http://tab.computer.org/tcsc/) sponsors a conference in this area called CLOUD and has some working groups and publications in this area.

Assessing the Role of Open Standards

- The first two deliverables of the project are a distributed object store based on Rackspace Cloud Files and a scalable machine provisioning technology based on NASA Nebula and Rackspace Cloud Servers.
- OpenStack Compute software will automatically create large groups of virtual private servers on industry-standard systems.
- OpenStack Storage is the software that will create redundant object-based storage using clusters of commodity servers and storage systems.

OpenStack.org is an industry group seeking to create open cloud standards based on Rackspace.com and NASA technologies.

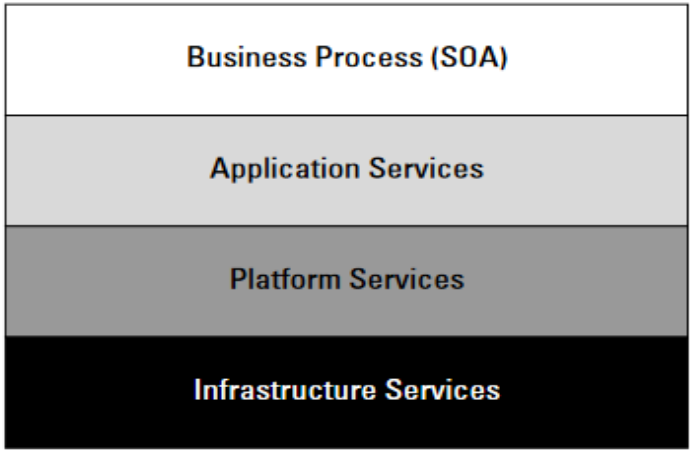


Infrastructure as a Service (IaaS)

- Cloud computing into four layers that form a cloud computing ecosystem.
- The Application layer forms the basis for Software as a Service (SaaS),
- The Platform layer forms the basis for Platform as a Service (PaaS) model.
- Infrastructure as a Service (IaaS) creates what may be determined to be a utility computing model, something that you can tap into and draw from as you need it without significant limits on the scalability of your deployment.
- You pay only for what you need when you need it. IaaS may be seen to be an incredibly disruptive technology, one that can help turn a small business into a large business nearly overnight.
- This is a most exciting prospect; one that is fuelling a number of IaaS start-ups during one of the most difficult recessions of recent memory.

Infrastructure as a Service (IaaS)

The cloud computing ecosystem



Infrastructure as a Service (IaaS)

- **Infrastructure as a Service (IaaS)** is a cloud computing service model in which hardware is virtualized in the cloud. In this particular model, the service vendor owns the equipment: servers, storage, network infrastructure, and so forth.
- The developer creates virtual hardware on which to develop applications and services. Essentially, an IaaS vendor has created a hardware utility service where the user provisions virtual resources as required.
- The developer interacts with the IaaS model to create virtual private servers, virtual private storage, virtual private networks, and so on, and then populates these virtual systems with the applications and services it needs to complete its solution.
- In IaaS, the virtualized resources are mapped to real systems. When the client interacts with an IaaS service and requests resources from the virtual systems, those requests are redirected to the real servers that do the actual work.

Infrastructure as a Service (IaaS)

- The fundamental unit of virtualized client in an IaaS deployment is called a **workload.** A workload simulates the ability of a certain type of real or physical server to do an amount of work.
- The work done can be measured by the number of Transactions Per Minute (TPM) or a similar metric against a certain type of system. A workload has certain other attributes such as Disk I/Os measured in Input/Output Per Second IOPS, the amount of RAM consumed under load in MB, network throughput and latency, and so forth.
- In a hosted application environment, a client's application runs on a dedicated server inside a server rack or perhaps as a standalone server in a room full of servers.
- In cloud computing, a provisioned server called an instance is reserved by a customer, and the necessary amount of computing resources needed to achieve that type of physical server is allocated to the client's needs.
- Figure shows how three virtual private server instances are partitioned in an IaaS stack. The three workloads require three different sizes of computers**: small, medium, and large.**
- A client would reserve a machine equivalent required to run each of these workloads. The IaaS infrastructure runs these server instances in the data center that the service offers, drawing from a pool of virtualized machines, RAID storage, and network interface capacity.

Infrastructure as a Service (IaaS)

- A client would reserve a machine equivalent required to run each of these workloads. The IaaS infrastructure runs these server instances in the data center that the service offers, drawing from a pool of virtualized machines, "**redundant array of inexpensive disks**" or "**redundant array of independent disks (**RAID) storage, and network interface capacity.
- These three layers are expressions of physical systems that are partitioned as logical units. **Logical Unit Number(LUNs)** the cloud interconnect layer, and the virtual application software layer are logical constructs.

- LUNs are logical storage containers, the cloud interconnect layer is a virtual network layer that is assigned IP addresses from the IaaS network pool, and the virtual application software layer contains software that runs on the **physical VM instance(s)** that have been partitioned from physical assets on the IaaS' private cloud.

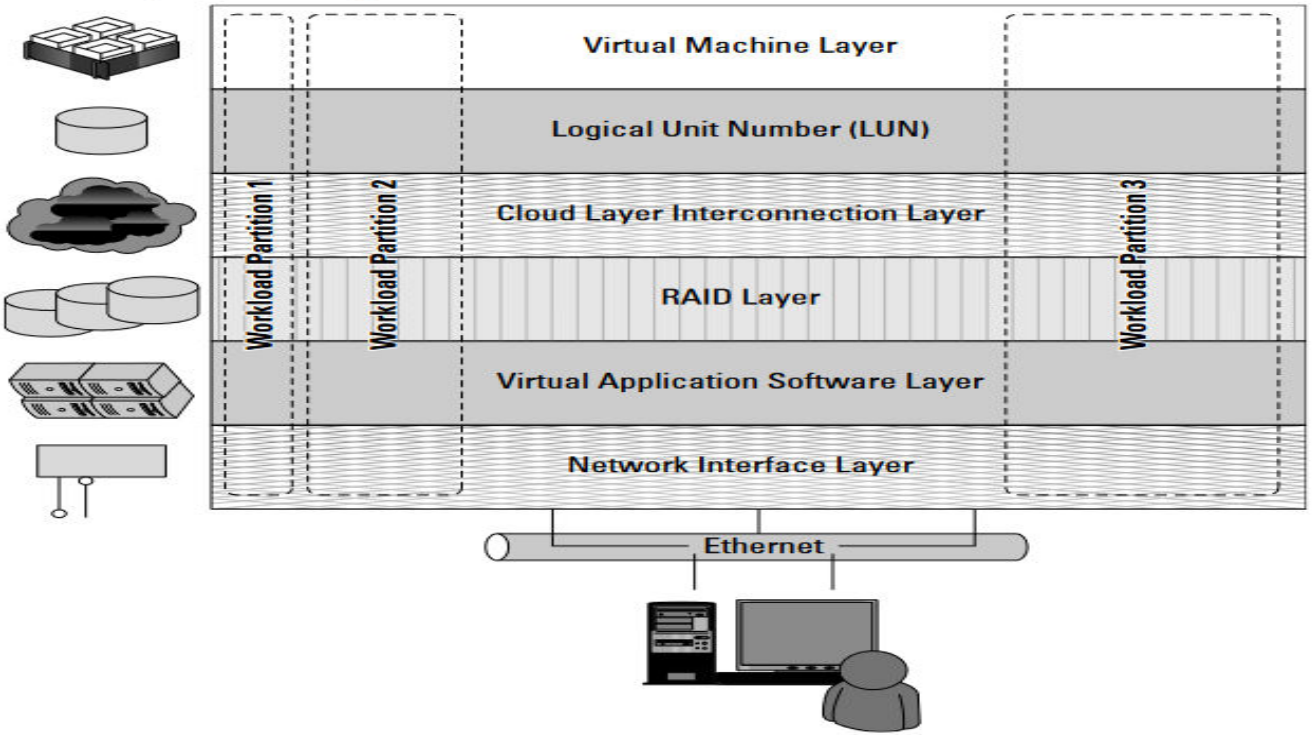
- From an architectural standpoint, the client in an IaaS infrastructure is assigned its own private network. The Amazon **Elastic Computer Cloud (EC2),** behaves as if each server is its own separate network—unless you create your own Virtual Private Cloud (an EC2 add-on feature), which provides a workaround to this problem.

- When you scale your EC2 deployment, you are adding additional networks to your infrastructure, which makes it easy to logically scale an EC2 deployment, but imposes additional network overhead because traffic must be routed between logical networks.

Infrastructure as a Service (IaaS)

-

A virtual private server partition in an IaaS cloud



Infrastructure as a Service (IaaS)

- Amazon Web Service's routing limits broadcast and multicast traffic because Layer-2 (Data Link) networking is not supported. Rackspace **Cloud (http://www.rackspacecloud.com/)** follows the AWS IP assignment model.

- Other IaaS infrastructures such as the one Cloudscaling.com (http://www.cloudscaling.com) offers or a traditional VMWare cloud-assigned networks on a per-user basis, which allows for Level 2 networking options.

- The most prominent Level 2 protocols that you might use are tunneling options, because they enable VLANs.

- Consider a transactional eCommerce system, for which a typical stack contains the following components:

- Web server
- Application server
- File server
- Database
- Transaction engine

- This eCommerce system has several different workloads that are operating: queries against the database, processing of business logic, and serving up clients' Web pages.

Infrastructure as a Service (IaaS)

- The classic example of an IaaS service model is **Amazon.com's Amazon Web Services (AWS).** AWS has several data centers in which servers run on top of a virtualization platform (Xen) and may be partitioned into logical compute units of various sizes.

- Developers can then apply system images containing different operating systems and applications or create their own system images.

- Storage may be partitions, databases may be created, and a range of services such a messaging and notification can be called upon to make distributed application work correctly.

Cross-Ref

- Amazon Web Services offers a classic Service Oriented Architecture (SOA) approach to IaaS.

Pods, aggregation, and silos

- **Workloads support** a certain number of users, at which point you exceed the load that the instance sizing allows. When you reach the limit of the largest virtual machine instance possible, you must make a copy or clone of the instance to support additional users.
- A group of users within a particular instance is called a **Pod**. Pods are managed by a Cloud Control System (CCS). In AWS, the CCS is the AWS Management Console.

Infrastructure as a Service (IaaS)

- Sizing limitations for pods need to be accounted for if you are building a large cloud-based application. Pods are aggregated into pools within an IaaS region or site called an availability zone.
- In very large cloud computing networks, when systems fail, they fail on a pod-by-pod basis, and often a zone-by-zone basis.
- For AWS' IaaS infrastructure, the availability zones are organized around the company's data centers in Northern California, Northern Virginia, Ireland, and Singapore.
- A failover system between zones gives IaaS private clouds a very high degree of availability.
- Figure shows how pods are aggregated and virtualized in IaaS across zones.
- When a cloud computing infrastructure isolates user clouds from each other so the management system is incapable of interoperating with other private clouds, it creates an information silo, or simply a silo.
- Most often, the term silo is applied to PaaS offerings such as Force.com or QuickBase, but silos often are an expression of the manner in which a cloud computing infrastructure is architected.

Infrastructure as a Service (IaaS)

- **Silos** are the cloud computing equivalent of compute islands: They are processing domains that are sealed off from the outside.
- When you create a private virtual network within an IaaS framework, the chances are high that you are creating a Silo.
- **Silos impose restrictions on interoperability** that runs counter to the open nature of build-componentized service-oriented applications.
- However, that is not always a bad thing. A silo can be its own ecosystem; it can be protected and secured in ways that an open system can't be. Silos just aren't as flexible as open systems and are subject to vendor lock-in.

Infrastructure as a Service (IaaS)

Pods, aggregation, and failover in IaaS