EE 542: USC Fall 2017
Internet and Cloud Computing

Lectures 3 ~ 4 : Aug.28 and 30, 2017

Cloud Architecture and Service Models (Chapter 4)

Prof. Kai Hwang, EEB 212
kaihwang@usc.edu

Some Representative Public Clouds

Table 4.1
Five public cloud platforms and their service offerings (2016)

Public Clouds	Platform Model(s)	Typical Service Offerings	Website and Coverage in Book
Amazon Web Service (AWS)	IaaS, PaaS	EC2, S3, SQS, EMR, VPC, EBS, SNS, CloudFront, etc.	http://aws.amazon.com/, Section 4.3
Google App Engine	PaaS, SaaS	Gmail, Docs, GFS, BigTable, Chubby	https://developer.google.com, Section 4.4.1~2
Microsoft Azure	PaaS, SaaS	Live, SQL, Office 365, Dynamic CRM	http://www.windowsazure.com, Section 4.4.3
IBM SmartCloud	PaaS, SaaS, IaaS	Compute, Storage, Backup, Networking, Virtualization	http:/www.ibm.com/cloud-computing, Section 4.5.2
SalesForce Clouds	SaaS, PaaS	CRM, Sales, Marketing, Apex, Visual force	https://salesforce.com, Section 4.5.1

Enabling Technologies for The Clouds

Table 4.3 Cloud-Enabling Technologies in Hardware, Software, and Networking

Technology	Requirements and Benefits
Fast platform deployment	Fast, efficient, and flexible deployment of cloud resources to provide dynamic computing environment to users
Virtual clusters on demand	Virtualized cluster of VMs provisioned to satisfy user demand and virtual cluster reconfigured as workload changes
Multitenant techniques	SaaS for distributing software to a large number of users for their simultaneous use and resource sharing if so desired
Massive data processing	Internet search and Web services which often require massive data processing, especially to support personalized services
Web-scale communication	Support for e-commerce, distance education, telemedicine, social networking, digital government, and digital entertainment applications
Distributed storage	Large-scale storage of personal records and public archive information which demands distributed storage over the clouds
Licensing and billing services	License management and billing services which greatly benefit all types of cloud services in utility computing

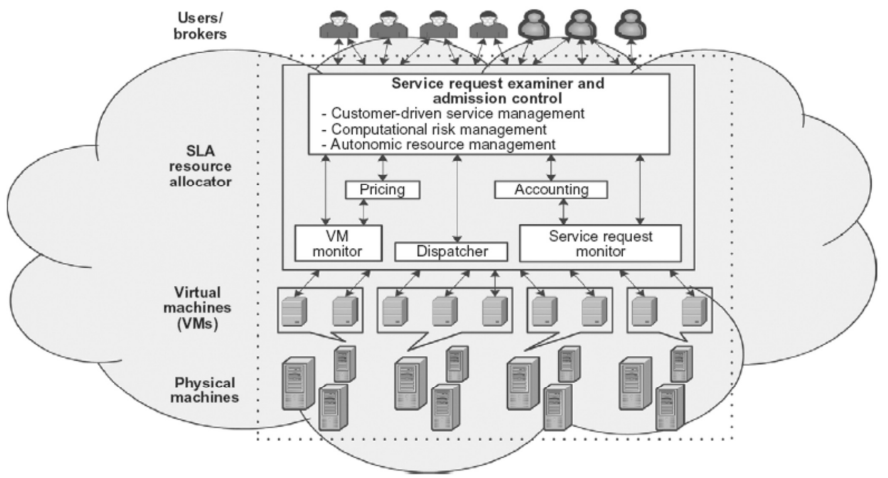


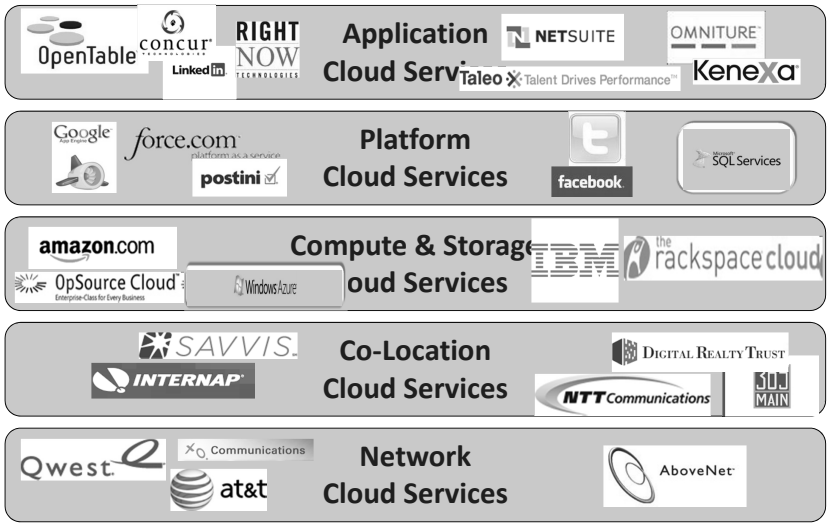
FIGURE 4.16
Market-oriented cloud architecture to expand/shrink leasing of resources with variation in QoS/demand from users.
(Courtesy of Raj Buyya, et al. [11])

Cloud Categories Based on Applications

Table 4.2
Five cloud service categories and their representative providers

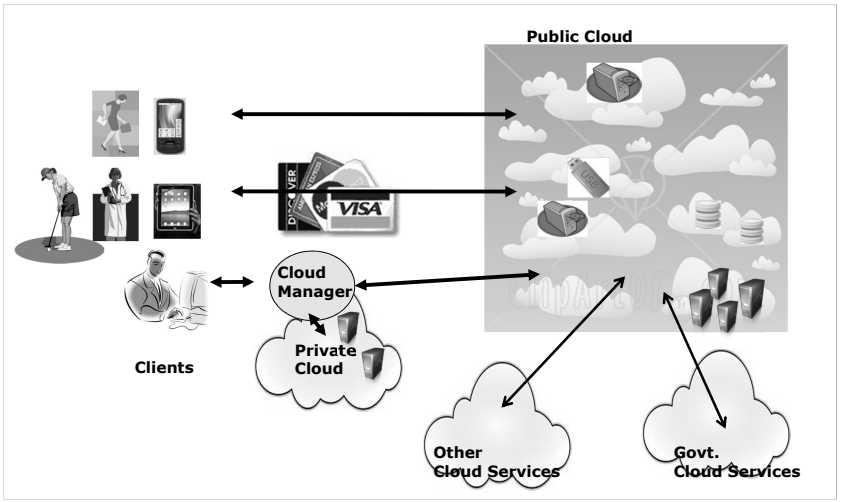
Cloud Categories	Cloud Service Providers
Application Clouds	OpenTable, Kenexa, NetSuite, RightNow, WebEx, Blackbaud, Concur Cloud, Telco, Omniture, Vocus, Microsoft OWA (Office 365), Google Gmail, Yahoo!, Hotmail
Platform Clouds	Force.com, Google App Engine, Facebook, IBM Blue Cloud, Postini, SQL Server, Twitter, Microsoft Azure, SGI Cyclone, Amazon EMR
Compute and Storage Clouds	Amazon AWS, Rackspace, OpSource, GoGrid, MeePo, FlexiScale, HP Cloud, Banknorth, VMware, XenEnterprise, iCloud
Colocation Clouds	Savvis, Internap, Digital Realty, Trusted Advisor, 365 Main
Network Clouds	AboveNet, AT&T, Qwest, NTT Communications

Today's Cloud Services Stack

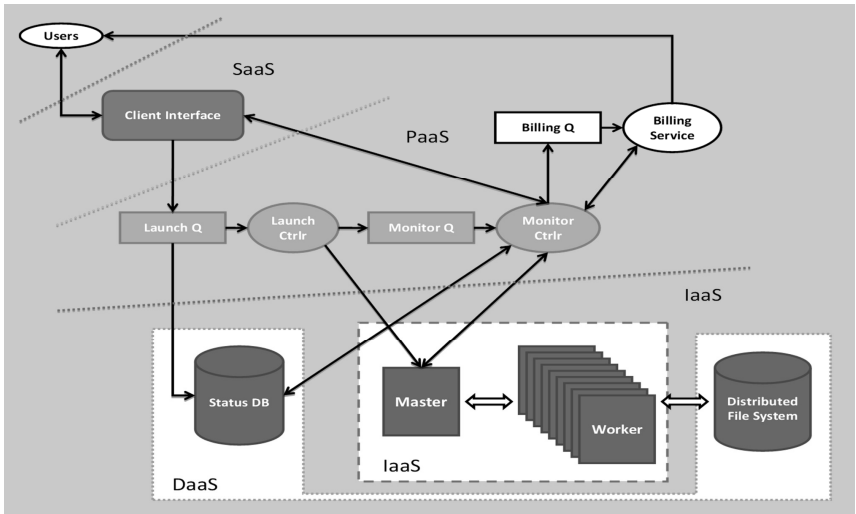


(Courtesy of T. Chou, 2010)

Subscription-Oriented Cloud Services
X (compute, apps, data, ..) as a Service (.. aaS)



Extended Cloud Service Models :
DaaS, HaaS, SecaaS, NaaS, CoLaaS, etc.



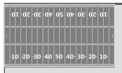
Cloud is built on virtualized datacenters

Range in size from "edge" facilities to megascale (100K to 1M servers)

Economics of scale

- Approximate costs for a small size center (1K servers) and a larger, 400K server center.

Technology	Cost in small-sized Data Center	Cost in Large Data Center	Ratio
Network	$95 per Mbps/ Month	$13 per Mbps/ month	7.1
Storage	$2.20 per GB/ Month	$0.40 per GB/ month	5.7
Administration	~140 servers/ Administrator	>1000 Servers/ Administrator	7.1



This data center is
11.5 times
the size of a football field

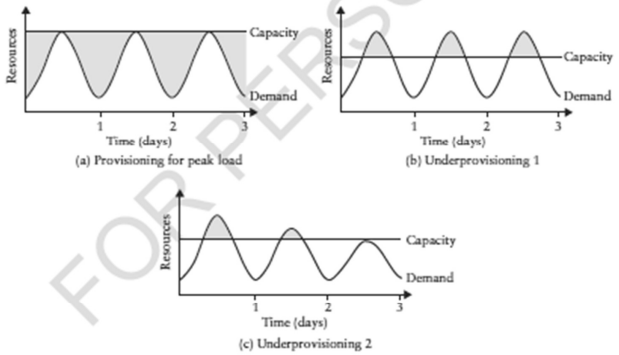


Figure 4.7
Three cloud resource provisioning cases: (a) heavy waste due to over-provisioning of resources, (b) under-resources provisioning, and (c) under-then over-provisioning.

Table 4.5
Open source virtual infrastructure managers and cloud operating systems

Manager/OS, Platforms, License	Resources Being Virtualized, Web Link	Client API, Language	Hypervisors Used	Public Cloud Interface	Special Features
Nimbus, Linux, Apache v2	VM creation, virtual cluster, http://www.nimbusproject.org/	EC2 WS, WSRF, CLI	Xen, KVM	EC2	Virtual networks
Eucalyptus, Linux, BSD	Virtual networking (Example 3.23) http://www.eucalyptus.com/	EC2 WS, CLI	Xen, KVM	EC2	Virtual networks
OpenNebula, Linux, Apache v2	Managing VM, host, virtual network, scheduling tools http://www.opennebula.org/	XML-RPC, CLI, Java	Xen, KVM	EC2, ElasticHosts	Virtual networks, Dynamic provisioning
vSphere/6, Linux, Windows, Proprietary	Virtualizing OS for data centers http://www.vmware.com/products/vsphera/ (Example 4.7)	CLI, GUI, Portal, WS	VMware ESX, ESXi	VMware vCloud partners	Data protection, vStorage, VMFS, high availability

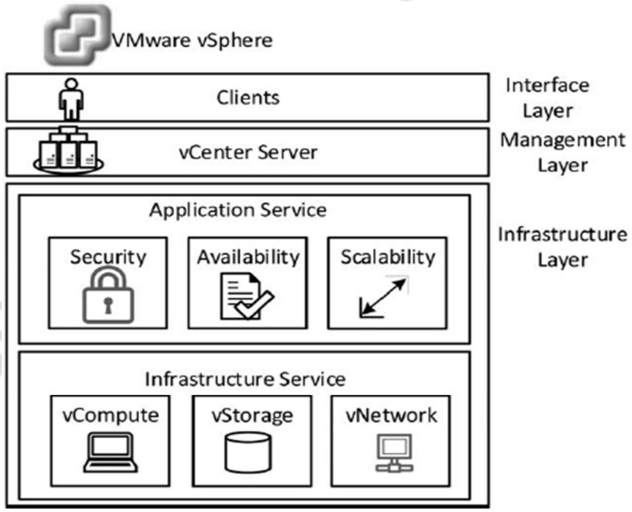


Figure 4.12
Functional modules in VMware vSphere 6 operating system for private/hybrid clouds.

Table 4.6
Comparison of three cloud platform architectures

Cloud System Features	Amazon Web Service (AWS): Public Cloud	OpenStack Systems: Private Cloud	VMWare Systems: Hybrid Cloud
Service Model(s)	IaaS, PaaS	IaaS	IaaS, PaaS
Developer/Provider and Design	Amazon (Sec.4.3)	Rackspace/NASA and Apache (Section 2.3.4)	VMWare (Section 2.3.5) Proprietary
Architecture Packages and Scale	Data centers distributed as availability zones in many global regions (Figure 2.14)	Small cloud at owner sites, licensed thru Apache (Figure 2.15)	Private clouds interacting with public clouds (Figure 2.18)
Cloud OS/ Software Support	Supporting both Linux and Windows machine instances with autoscaling and billing	Open source, extending from Eucalyptus and OpenNebula	vSphere and vCenter, supporting x-86 servers with NSX and vSAN
User Spectrum	General public: enterprises and individual users	Research centers or small businesses	Enterprises and large organizations

Some IaaS Offerings from Public Clouds :

Table 4.7
Public IaaS clouds and their VM instance configurations (August 2015)

Cloud Name	Virtual Machine Instance Configurations	API and Access Tools
Amazon EC2	Each instance has 1–20 EC2 processors, 1.7–15 GB memory, and 160 TB storage	CLI or Web Service (WS) porta
GoGrid	Each instance has 1–6 CPUs, 0.5–8 GB memory, and 30–480 GB storage	REST, Java, PHP, Python, Ruby
Rackspace Cloud	Each instance has a 4-core CPU, 0.25–16 GB memory, and 10–620 GB storage	REST, Python, PHP, Java, C#, .I
FlexiScale in UK	Each instance has 1–4 CPUs, 0.5–16 GB memory, and 20–270 GB storage	Web console

The Amazon Web Service (AWS):

The Most popular Public Cloud in Use Today

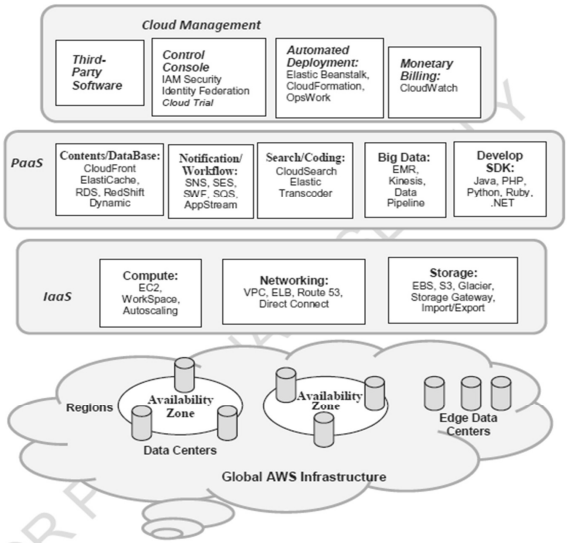


Figure 4.14
The AWS public cloud consisting of the top management layer, PaaS, and IaaS platforms, and the global infrastructure built over data centers in availability zones located in various regions.

Table 4.9
Compute, storage, database, and networking services in AWS cloud

Category	Offering	Service Modules or Short Description
Compute	EC2	Virtual servers in the AWS cloud
	Lambda	Run code in response to events
	EC2 Container Service	Run and manage Docker containers
Storage & Content Delivery	S3	Scalable storage in the AWS cloud
	Elastic File System	Fully managed file system for EC2
	Storage Gateway	Integrate on-premises IT facilities with cloud storage
	Glacier	Archive storage in the AWS cloud
	CloudFront	Global content delivery network
Database	RDS	MySQL, PostgreSQL, Oracle, SQL Server
	DynamicDB	Predictable and scalable NoSQL data store
	ElastiCache	In-memory cache
	RedShift	Managed petabyte-scale warehouse service
Networking	VPC	Virtual private cloud as isolated cloud resources
	Direct Connect	Dedicated network connection to AWS
		Scalable DNS and domain name registration

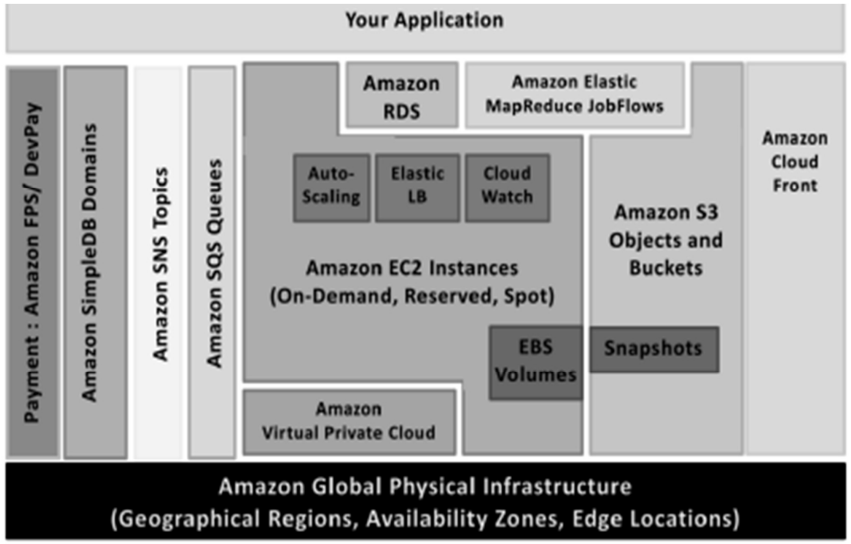
Compute, Storage, Database and Networking Services in The AWS Cloud in 2015

Category	Offering	Service Modules or Short Description
Compute	EC2	Virtual Servers in the AWS Cloud
	Lambda	Run Code in Respoinse to Events
	EC2 Container Service	Run and Manage Docker Containers
Storage & Content Delivery	S3	Scalable Storage in the AWS Cloud
	Elastic File System	Fully Management File System for EC2 (Preview)
	Storage Gateway	Integrate On-Premises IT Facilities with Cloud Storage
	Glacier	Archive Storage in the Cloud
	CloudFront	Global Content Delivery Network
Database	RDS	MySQL, Postgres, Oracle, SQL Server, and Amazon
	DynamicDB	Predictable and Scalable NoSQL Data Store
	ElastiCache	In-Memory Cache
	Redshift	Managed Petabyte-Scale Warehouse Service
Networking	VPC	Virtual Private Cloud as Isolated Cloud Resources
	Direct Connect	Dedicated Network Connection to AWS
	Route S3	Scalable DNS and Domain Name Registration

IaaS Clouds and Their Infrastructures and Service Offerings (Aug. 2015)

Cloud Name	Virtual Machine Instance Configurations	API and Access Tools
Amazon EC2	Each instance has 1-20 EC2 processors, 1.7-15 GB memory, and 160 TB storage	CLI or Web Service (WS) portal
GoGrid	Each instance has 1-6 CPUs, 0.5-8 GB memory and 30-480 GB storage	REST, Java, PHP, Python, Ruby
Rackspace Cloud	Each instance has a 4-core CPU, 0.25-16 GB memory and 10-620 GB storage	REST, Python, PHP, Java, C#, .NET
Flexiscale in UK	Each instance has 1-4 CPUs, 0.5-16 GB memory, and 20-270 GB storage	Web console

AWS Architecture



Important AWS Links To Visit and Documented Reports/Articles To Read

1. The main AWS site: http://aws.amazon.com
2. Just use http://aws.amazon.com/ec2/ to access ec2, or http://aws.amazon.com/s3/ for S3, or …./sqs/, ../sns/, or ../simpledb/, ……/sdk/, ……/fps/, etc.
3. Many PDF reports and articles, application examples, SDK tools, etc can be found in the above web sites. You do not need an account to access these sites.
4. You need to establish your team accounts on AWS ASAP. You need to use the AWS in doing some problems in HW#2 and to do the Team project as well.

EC2 Features on The AWS Platform

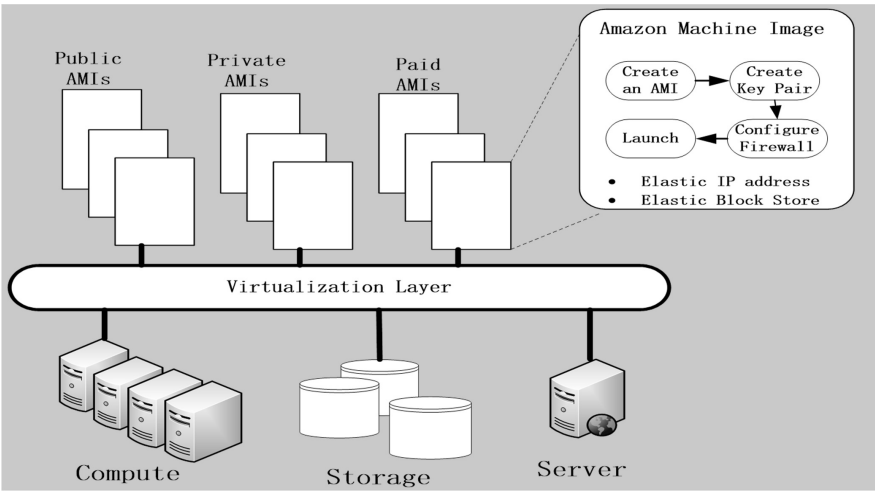
- Virtual computing environments, known as *instances*
- Pre-configured templates for your instances, known as *Amazon Machine Images (AMIs)*, that package the bits you need for your server (including the operating system and additional software)
- Various configurations of CPU, memory, storage, and networking capacity for your instances, known as *instance types*
- Secure login information for your instances using *key pairs* (AWS stores the public key, and you store the private key in a secure place)
- Storage volumes for temporary data that's deleted when you terminate your instance, known as *instance store volumes*
- Persistent storage volumes for your data using Amazon Elastic Block Store (Amazon EBS), known as *Amazon EBS volumes*
- Multiple physical locations for your resources, such as instances and Amazon EBS volumes, known as *regions* and *Availability Zones*
- A firewall that enables you to specify the protocols, ports, and source IP ranges that can reach your instances using *security groups*
- Static IP addresses for dynamic cloud computing, known as *Elastic IP addresses*
- Metadata, known as *tags*, that you can create and assign to your Amazon EC2 resources
- Virtual networks you can create that are logically isolated from the rest of the AWS cloud, and that you can optionally connect to your own network, known as *virtual private clouds* (VPCs)

Amazon Machine Images (AMI)

Image Type	Definition
Private	Images created by you, which are private by default. You can grant access to other users to launch your private images.
Public	Images created by users and released to the Amazon Web Services community, so anyone can launch instances based on them and use them any way they like. The Amazon Web Services Developer Connection Web site lists all public images.
Paid	You can create images providing specific functions that can be launched by anyone willing to pay you per each hour of usage on top of Amazon charges.

- **AMI is a packaged server environment in EC2, based on Linux running any user software or application. AMIs are the templates for VM instances.**
- **Elastic IP address is specially reserved for EC2. Elastic Block Store offers persistent storage for EC2 instances.**

Amazon EC2 Execution Environment



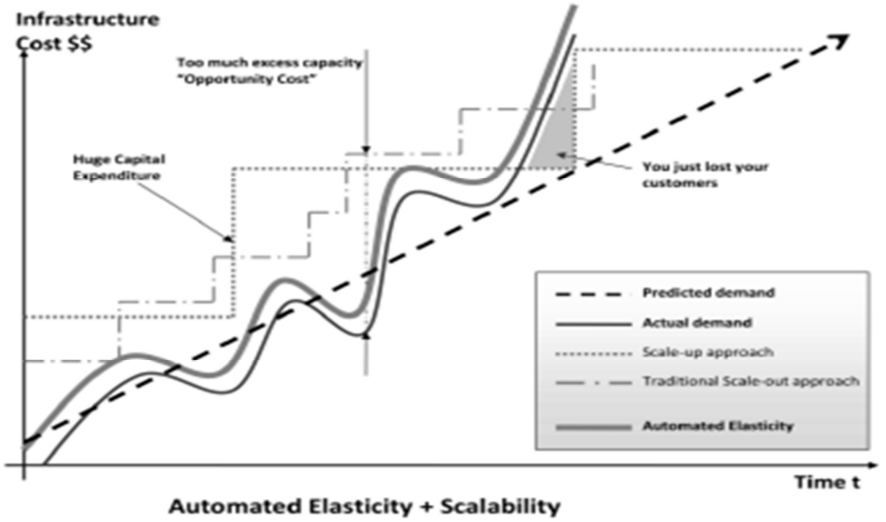
Amazon EC2 Instances:

Family	Description
Cluster Compute	Have a very large amount of CPU coupled with increased networking performance. They're well-suited for High Performance Compute (HPC) applications and other demanding network-bound applications. For more information, see Overview (p. 121).
Cluster GPU	Provide general-purpose graphics processing units (GPUs), with proportionally high CPU and increased network performance for applications that benefit from highly parallelized processing. They're well-suited for HPC applications as well as rendering and media processing applications. For more information, see Overview (p. 121).
High CPU	Have proportionally more CPU resources than memory (RAM). They're well-suited for compute-intensive applications.
High I/O	Provide tens of thousands of low-latency, random I/O operations per second (IOPS) to an application. They're well-suited for NoSQL databases, clustered databases, and OLTP (online transaction processing) systems. For more information, see High I/O Instances (p. 117).
High Memory	Have proportionally more memory resources. They're well suited for high-throughput applications, such as database and memory caching applications.
High-Memory Cluster	Have large amounts of memory coupled with high CPU and network performance. These instances are well suited for in-memory analytics, graph analysis, and scientific computing applications.
High Storage	Provide very high storage density and high sequential read and write performance per instance. They are well-suited for data warehousing, Hadoop/MapReduce, and parallel file systems. For more information, see High Storage Instances (p. 119).
Micro	Provide a small amount of consistent CPU resources and enable you to burst CPU capacity when additional cycles are available. They're well-suited for lower throughput applications and websites that consume significant compute cycles periodically. For more information, see Micro Instances (p. 110).
Standard	Have memory-to-CPU ratios suitable for most general-purpose applications.

Auto Scaling Features on Amazon EC2

- Auto Scaling allows you to scale your Amazon EC2 capacity up or down automatically according to conditions you define.
- With Auto Scaling, you can ensure that the number of Amazon EC2 instances you're using increases seamlessly during demand spikes to maintain performance, and decreases automatically during demand lulls to minimize costs.
- Auto Scaling is particularly well suited for applications that experience hourly, daily, or weekly variability in usage.
- Auto Scaling is enabled by Amazon CloudWatch and available at no additional charge beyond Amazon CloudWatch fees.

Understanding Elasticity in Cloud Resources



Lecture 4: Cloud Platforms and Services Offerings, August 30, 2017

S3: Simple Storage Service
Storage Cloud Service: Amazon

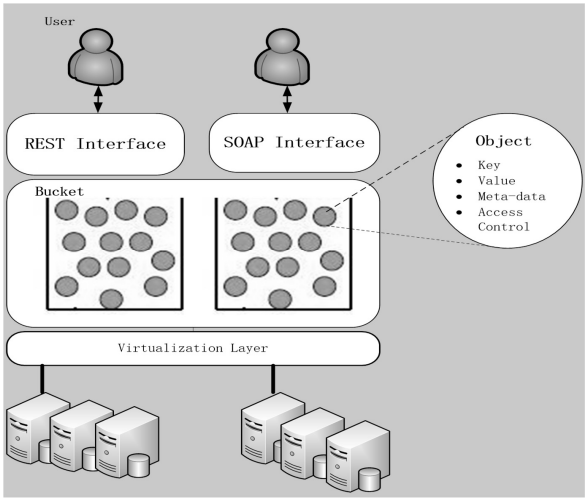
- Object-Based Storage
- 1 B – 5 GB / object
- Redundant thru geographic dispersion
- 99.99% Availability Goal
- Private or Public
- Per-object URLs & ACLs
- BitTorrent Support

$.15 per GB per month storage

$.01 for 1000 to 10,000 requests

$.10 - $.18 per GB data transfer

Amazon S3 for Storage Provisioning



- Object is the basic unit of data
- Bucket for storing objects
- Key for data object retrieval
- Object is attributed to value, metadata, and access control

Key Features in Amazon S3:

- Write, read, and delete objects containing from 1 byte to 5 TB of data each. The number of objects you can store is unlimited.
- Each object is stored in a bucket and retrieved via a unique, developer-assigned key.
- A bucket can be stored in one of several Regions. You can choose a Region to optimize for latency, minimize costs, or address regulatory requirements.
- The US Standard Region automatically routes requests to facilities in Northern Virginia or the Pacific Northwest using network maps.
- Objects stored in a Region never leave the Region unless you transfer them out. For example, objects stored in the EU (Ireland) Region never leave the EU.

Steps to Use Amazon S3

- Create a Bucket to store your data. You can choose a Region where your bucket and object(s) reside to optimize latency, minimize costs, or address regulatory requirements.
- Upload Objects to your Bucket. Your data is durably stored and backed by the Amazon S3 Service Level Agreement.
- Optionally, set access controls. You can grants others access to your data from anywhere in the world.

Application, Mobile and Analytics Services in the AWS Cloud

Category	Offering	Service Modules or Short Description
Application Services	SQS	Message Queue Services
	SWF	Workflow for Coordinating App Components
	AppStream	Low Latency Application Streaming
	Elastic Transcoder	Easy-To-Use Scalable Media Transcoding
	SES	Email Sending Service
	CloudSearch	Managed Search Service
	API Gateway	Build, Deploy and Mange APIs
Mobile Services	Cognito	User Identity and App Data Synchronization
	Device Farm	Test Android and iOS apps on Mobile Devices
	Mobile Analytics	Collect, View and Export App Analytics
	SNS	Simple Push Notification Service
Analytics Services	EMR	Managed Elastic Hadoop Framework
	Kinesis	Real-Time Processing of Streaming Big Data
	Data Pipeline	Orchestration for Data-Driven Workflows
	Machine Learning	Build Smart Applications Quickly and Easily

Amazon Simple Queue Service (Amazon SQS)

- Amazon SQS is a fast, reliable, scalable, fully managed queue service. SQS makes it simple and cost-effective to decouple the components of a cloud application. You can use SQS to transmit any volume of data, at any level of throughput, without losing messages.
- Amazon SQS can be used with other AWS Services such as Redshift, DynamoDB, RDS, EC2, and S3, to make distributed applications more scalable and reliable. Some common queue designs follows:
  - Work Queues: Decouple components of a distributed application that may not all process the same amount of work simultaneously.
  - Buffer and Batch Operations: Add scalability and reliability to your architecture, and smooth out temporary volume spikes without losing messages or increasing latency.
  - Request Offloading: Move slow operations off of interactive request paths by enqueing the request.
  - Fanout: Combine SQS with Simple Notification Service (SNS) to send identical copies of a message to multiple queues in parallel for simultaneous processing.

Amazon Simple Notification Services (SNS) :

- The SNS is a fast, flexible, fully managed push messaging service. SNS makes it simple and cost-effective to push to mobile devices such as iPhone, iPad, Android, Kindle Fire, and internet connected smart devices, as well as pushing to other distributed services.
- Besides pushing directly to mobile devices, SNS can also deliver notifications by SMS text message or email, to Simple Queue Service (SQS) queues, or to any HTTP endpoint.
- Amazon SNS lets you push messages to mobile devices or distributed services, via API or an easy-to-use management console. You can seamlessly scale from a handful of messages per day to millions of messages or higher.
- With SNS you can publish a message once, and deliver it one or more times. So you can choose to direct unique messages to individual Apple, Google or Amazon devices, or broadcast deliveries to many mobile devices with a single publish request.

Amazon SimpleDB Service

- Provide a simplified relational database model. Structured data are organized into domains in terms of relational tables.
- Data items correspond to rows in the table. Specific attributes form the columns.
- Cells in the table carry the values of attributes. An attribute may have multiple values.

Amazon Elastic Block Service (EBS)

- EBS provides volume block interface from 1 GB to 1 TB for saving and restoring VM instances used in EC2.
- User can also use EBS to save persistent data to be used in various EC2 instances.
- This is different from S3, which is a messaging interface. EBS acts more like a distributed file system. in terms of relational tables.

Administraion, Security, Enterprise and Deployment Services in the AWS Cloud

Category	Offering	Service Modules or Short Description
Administration & Security	Directory Service	Managed Directory in the AWS Cloud
	Identity/Access Manager	Access Control and Key Management
	Trusted Advisor	AWS Cloud Optimization Expert
	Cloud Trail	User Activity and Change Tracking
	Configuration	Resource Configurations and Inventory
	CloudWatch	Resource and Application Monitory
	Service Catalog	Personalized Catalog of AWS Resources
Enterprises Applications	Workplaces	Destops in the AWS Cloud
	WorkDocs	Secure Enterprise Storage and Sharing Service
	WorkMail	Secure Email and Calendaring Service
Deployment and Management	Elastic Beanstalk	AWS Application Containers
	OpsWorks	DevOps Application Managemnt Service
	CloudFormation	Templated AWS Resource Creation
	CodeEeploy	Automated Deployments
	CodeCommit	Managed Git Repositories
	Code Pipeline	Continuous Delivery of Codes

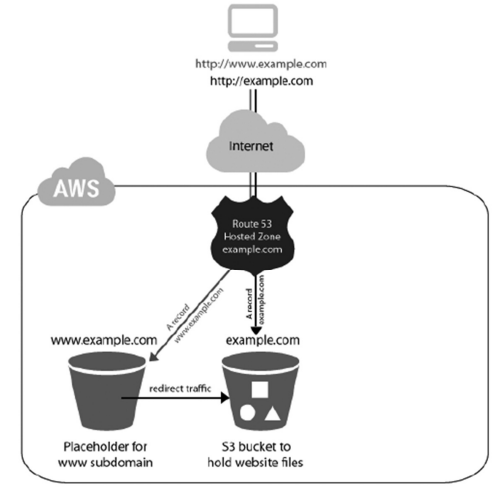


Figure 4.18
Construction of a static website hosted on the AWS S3.

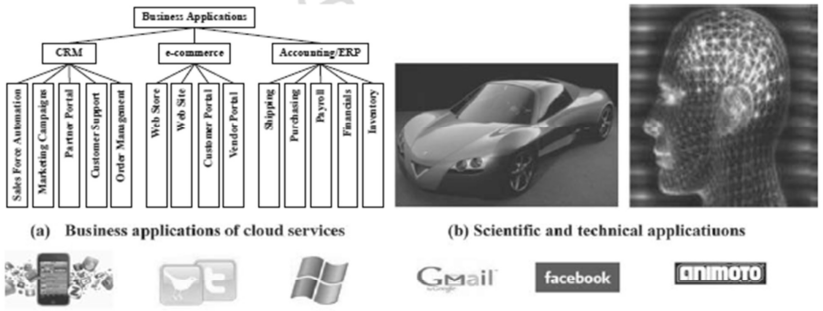


Figure 4.19
Typical cloud applications in business, science/technical, and social media areas.

Other Public Clouds from Google, Microsoft, IBM, SalesForce, HP, SGI, etc.

Table 4.8
Public clouds offering PaaS services (August 2015)

Cloud Name	Languages and Developer Tools	Programming Models Supported by Provider	Target Applications and Storage Option
Google App Engine	Python, Java, and Eclipse-based IDE	MapReduce, Web programming on demand	Web applications and BigTable storage
Salesforce, Force.com	Apex, Eclipse-based IDE, web-based Wizard	Workflow, Excel-like, Web programming on demand	CRM and add-on app development for business
Microsoft Azure	.NET, Azure tools for MS Visual Studio	Dryad, Twister, .NET Framework	Enterprise and web applications
Amazon Elastic MapReduce	Hive, Pig, Cascading, Java, Ruby, Perl, Python, PHP, R, and C++	MapReduce, Hadoop, Spark	Data processing, e-mail, e-commerce, S3, and WorkDocs

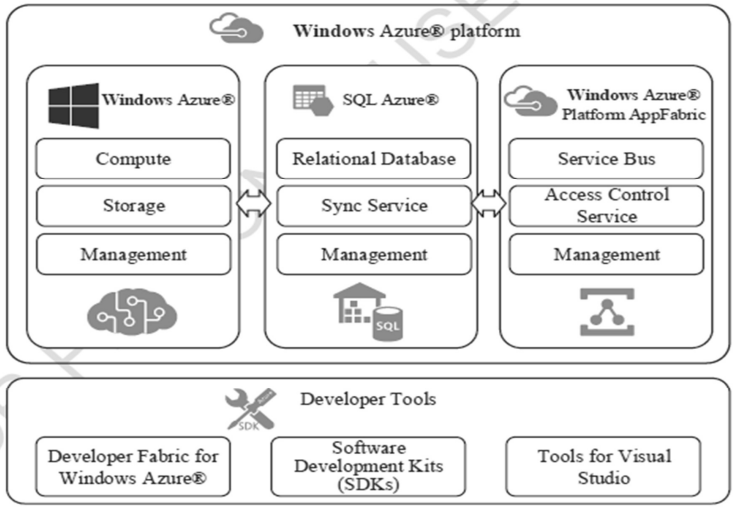


Figure 4.25
Microsoft Windows Azure platform for cloud computing. Courtesy of Microsoft, http://www.microsoft.com /windowsazure, 2016.

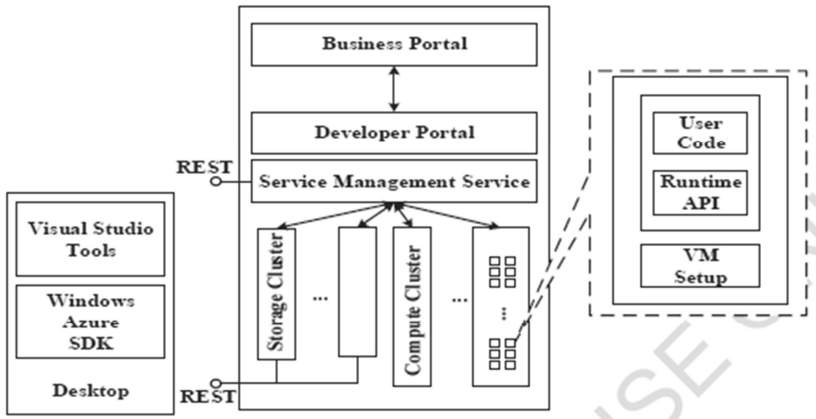


Figure 4.26
Microsoft Azure and its interfaces with business, developer, and client applications.

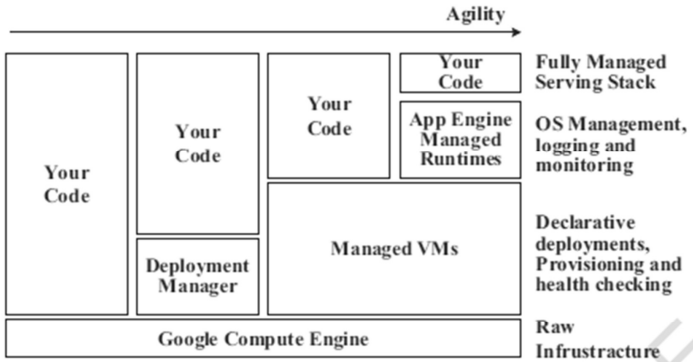


Figure 4.20
Building Google App Engine with increasing agility from raw infrastructure to Compute Engine, using OS support and managed VMs and runtimes to reduce user coding efforts. Courtesy of Google website, http://www.google.com, retrieved October 2016.

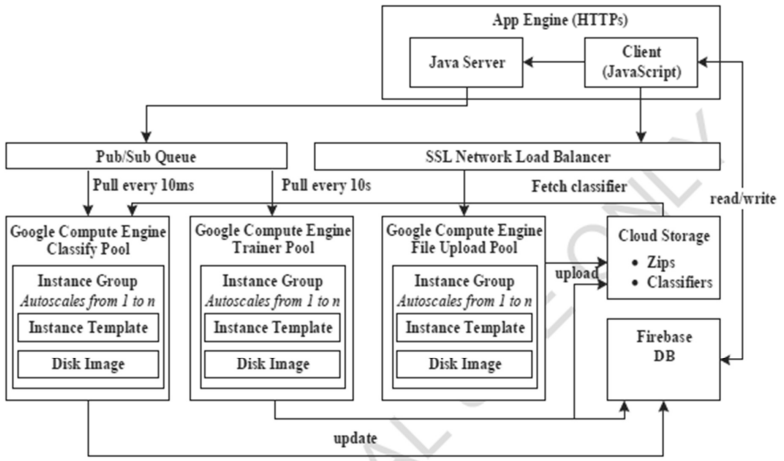


Figure 4.21
The architecture of Google Compute Engine as three functional pools and cloud storage to support Google App Engine applications development and executions.

SaaS Clouds and Service Offerings

Table 4.12
Four SaaS cloud platforms and their service offerings (August 2015)

Model	Amazon AWS	Google App Engine	Microsoft Azure	Salesforce
Platform Support	AWS EC2, S3, EMR, SNS, etc.	GAE, GFS, BigTable, MapReduce, etc.	Azure, .NET service, Dynamic CRM,	Force.com, Online CRM, Gifttag
SaaS Offerings	Elastic Beanstalk, CodeDeploy, OpsWorks, CodeCommit, Code-Pipeline, Mobile Analytics	Gmail, Docs, YouTube, WhatsApp	Live, SQL, Office 365 (OWA), Hotmail	Sales, service, market, data, collaboration, analytics
Security Features	CloudWatch, Trusted Advisor, Identity/Access Control	Chubby locks for security enforcement	Replicated data, rule-based access control	Admin./record security, use metadata API
APIs and Languages	API Gateway, Latin Pig	Web-based admin. console, Python	Azure portal, .NET Framework	Apex, Visualforce, AppExchange, SOSL, SOQL

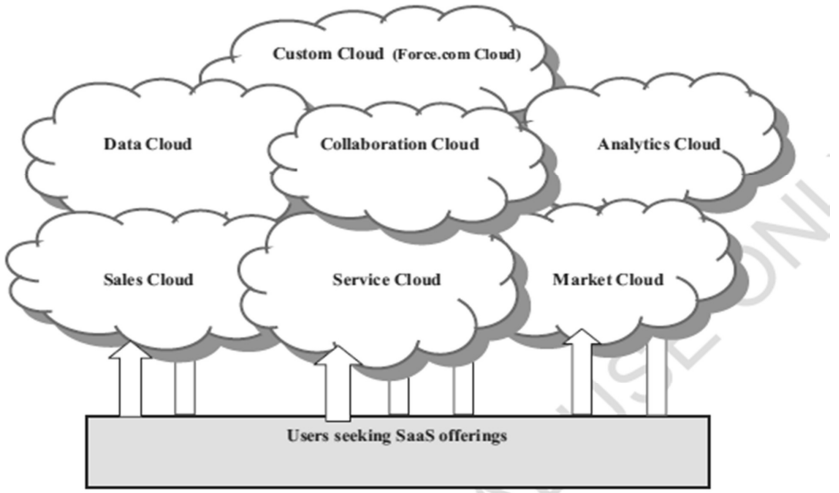


Figure 4.28
Seven Salesforce cloud service offerings: all for SaaS applications except the custom cloud offering PaaS applications.

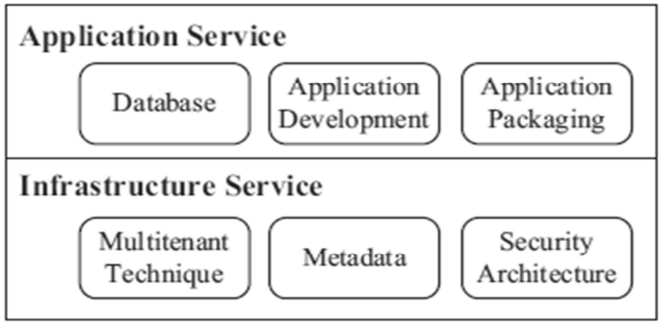


Figure 4.27
Salesforce Force.com cloud platform supporting both PaaS and SaaS applications.

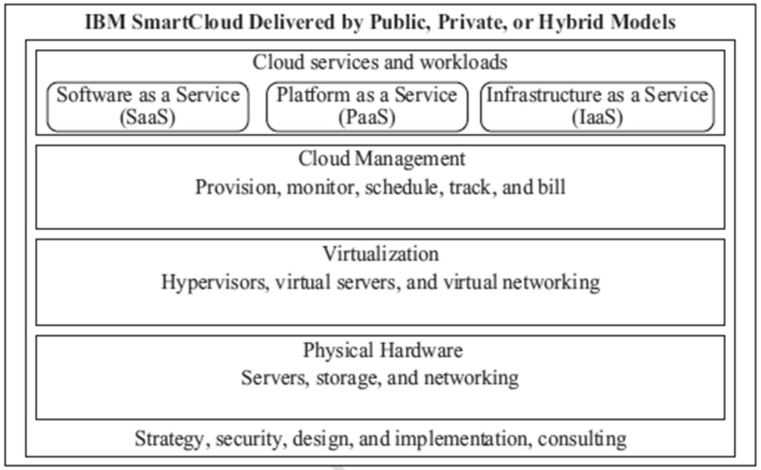


Figure 4.29
Architecture of an IBM SmartCloud offered as public, private, and hybrid platforms based on IBM Tivoli, DB2, WebSphere, and Xen-based virtualization.

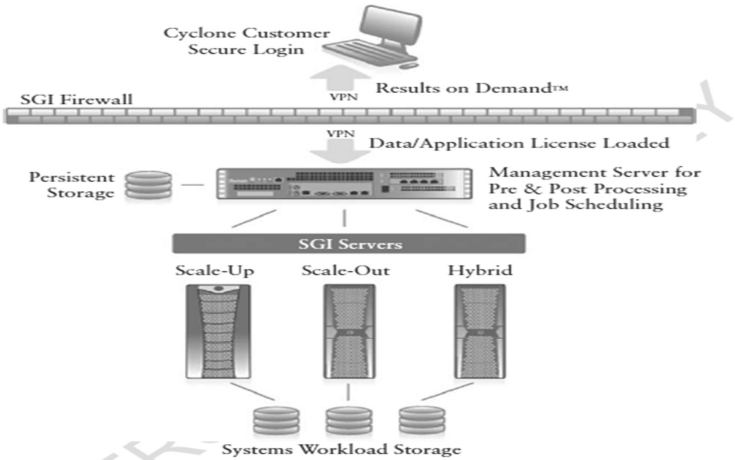


Figure 4.30
SGI Cyclone HPC cloud for enabling SaaS and IaaS applications.

The Nebula Cloud
Built at NASA Ames Research Center

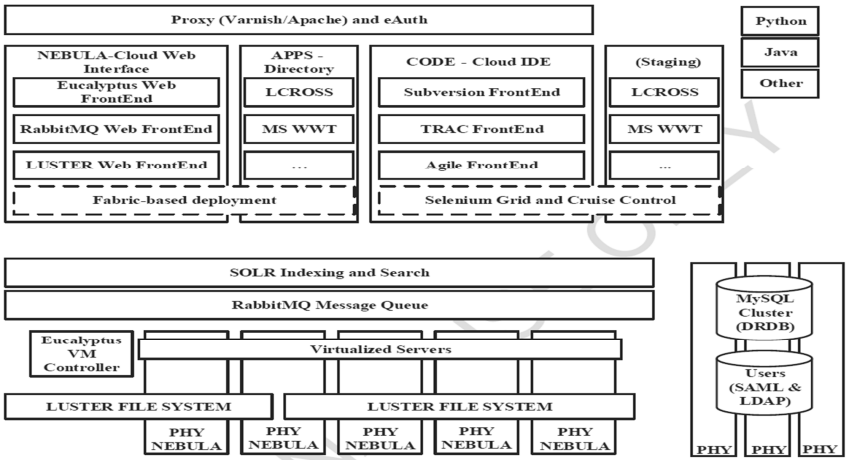


Figure 4.31
The Nebula cloud project at NASA, http://nebula.nasa.gov.

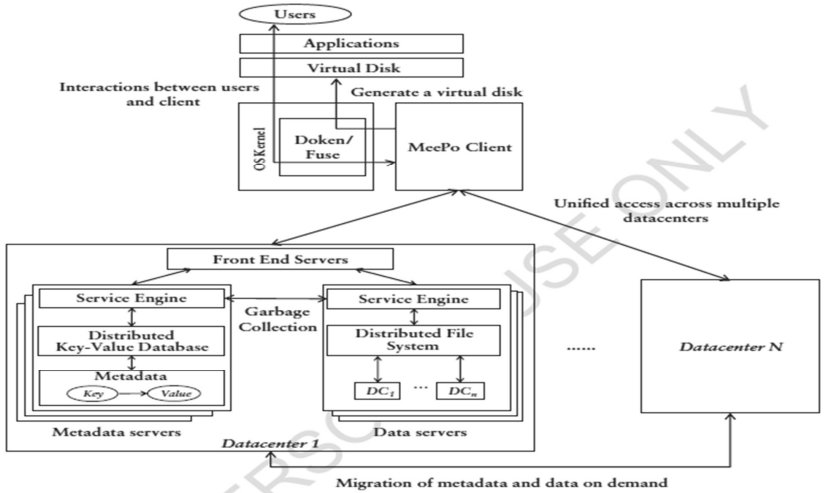


Figure 4.8
MeePo: A community cloud built at Tsinghua University over multiple data centers, where DC stands for data chunks. Courtesy of Wu et al., "Associative Big Data Sharing in Community Clouds—The MeePo Approach," *IEEE Cloud Computing Magazine* (January 2016).

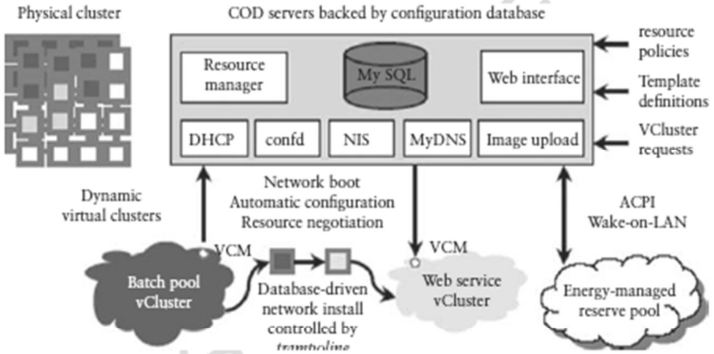


Figure 4.9
Cluster-on-Demand (COD) partitions a physical cluster into multiple virtual clusters. Courtesy of J. Chase, et al., "Dynamic Virtual Clusters in a Grid Site Manager," IEEE 12th Symposium on High-Performance Distributed Computing (HPDC), 2003.

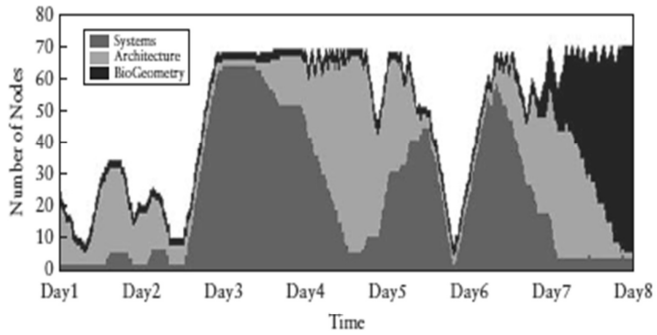


Figure 4.10
The cluster size variation in the COD for eight days at Duke University. Courtesy of J. Chase, et al., "Dynamic Virtual Clusters in a Grid Site Manager," IEEE 12th Symposium on High-Performance Distributed Computing (HPDC), Washington, DC, June 2003.

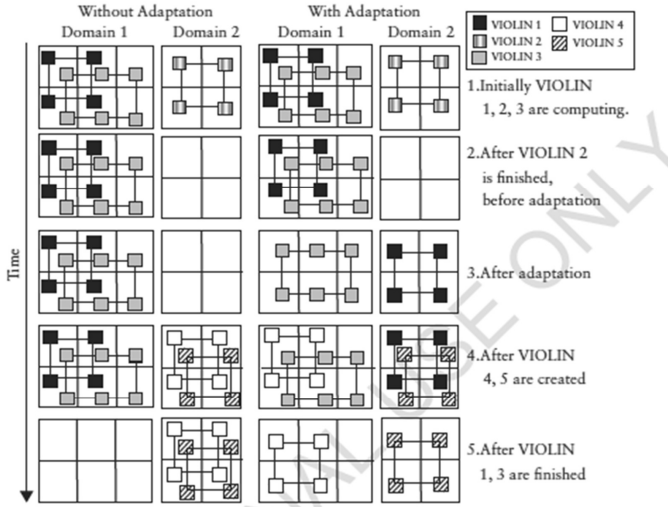


Figure 4.11
VIOLIN adaptation scenario of five virtual clusters on two hosted clusters. There are less idle nodes after the adaptation. Courtesy of P. Ruth et al., "Automatic Live Migration of Virtual Computational Environments in a Multi-Domain Infrastructure," *Technical Report*, Purdue University (2006).