Aalto University

Open Cloud Platforms & Cloud Services

Zhonghong Ou
Post-doc researcher
Data Communications Software (DCS) Lab,
Department of Computer Science and Engineering,
Aalto University

Cloud technology videos

- http://www.youtube.com/watch?v=txvGNDnKNWw&feature=related

- http://www.youtube.com/watch?v=QJncFirhjPg

Open cloud platform

- Eucalyptus
- Open Cirrus
- OpenNebula
- OpenStack
- …
- Mesos

Eucalyptus

Basics

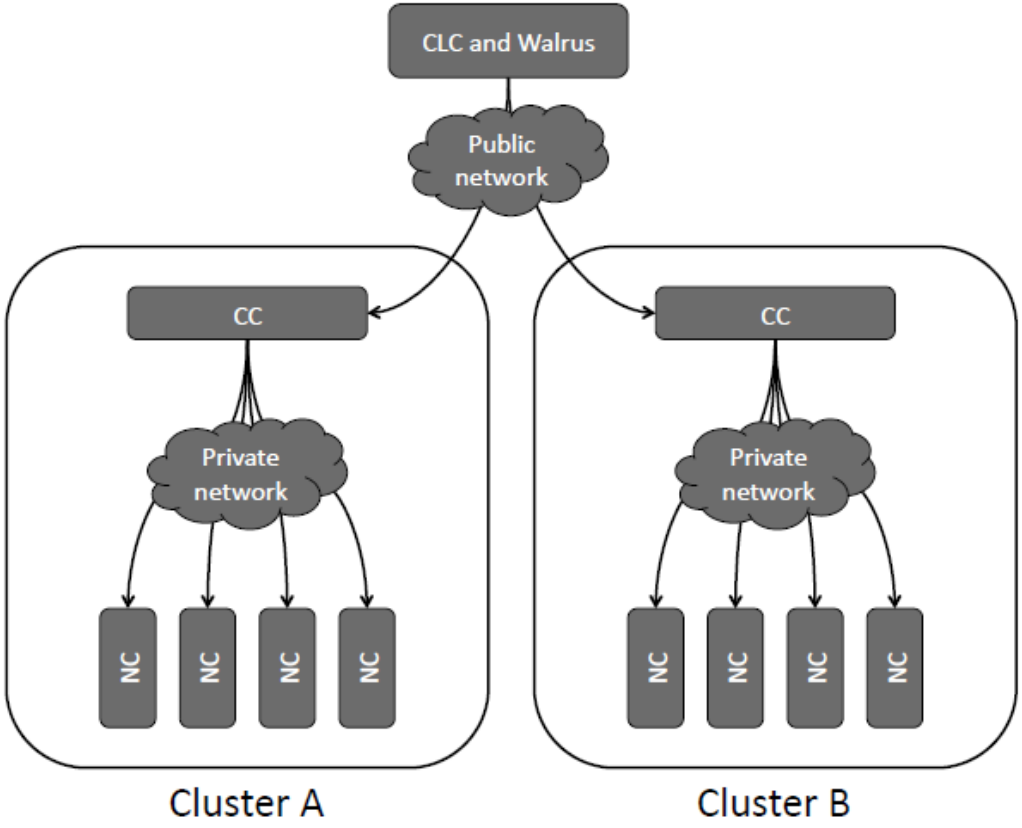
- Eucalyptus: Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems.
- Users can run and control entire virtual machine instances deployed across physical resources, so-called Infrastructure as a Service (IaaS).
- Suitable for heterogeneous infrastructure usually found in academic settings.
- Hypervisor agnostic, supporting Xen and KVM currently.

Cloud Computing vs. Grid Computing

- Grid systems mainly focus on providing individual user requests with large fractions of the total resource pool, while cloud systems often focus on scaling to large number of users, each cousuming tiny fraction of the resource capacity.

- Grid computing takes a middleware-based approach to promote resourse federation among cooperating, but seperate, administrative domains, while cloud service venues are unfederated.

Platform


CLC and Walrus
Public network
CC
CC
Private network
Private network
NC
NC
NC
NC
NC
NC
NC
NC
Cluster A
Cluster B

Components

- Node Controller (NC)
  - controls the execution, inspection, and terminating of VM instances on the host where it runs.
- Cluster Controller (CC)
  - gathers information about and schedules VM execution on specific node controllers, as well as manages virtual instance network.
- Storage Controller (Walrus)
  - a put/get storage service that implements Amazon's S3 interface, providing a mechanism for storing and accessing virtual machine images and user data.
- Cloud Controller (CLC)
  - the entry-point into the cloud for users and administrators. It queries node managers for information about resources, makes high-level scheduling decisions, and implements them by making requests to cluster controllers.

Node controller (NC)

- Queries and controls the system software on its node (i.e., the host operating system and the hypervisor) in response to queries and control requests from its Cluster Controller;
- Makes queries to discover the node's physical resources – the number of cores, the size of memory, the available disk space – as well as to learn about the state of VM instances on the node;
- Upon receiving requests from Cluster controller, an NC verifies the authorization (only the owner and administrator are authorized) and confirms the resource availability, if both are fine, then it will execute the request on the assistance of the hypervisor.
- To start an instance, the NC makes a node-local copy of the instance image files (*the kernel, the root file system, and the ramdisk image*), either from a remote image repository or from the local cache, creates a new endpoint in the virtual network overlay, and instructs the hypervisor to boot the instance.
- To stop an instance, the NC instructs the hypervisor to terminate the VM, tears down the virtual network endpoint, and cleans up the files associated with the instance (the root file system is not preserved after the instance terminates).

Cluster controller (CC)

- Cluster Controllers control VM instances on a node by making *runInstance* and *terminateInstance* requests to the node's NC.
- Three functions:
  - Schedule incoming instance run requests to specific NCs;
  - Control the instance virtual network overlay;
  - Gather/report information about a set of NCs.

- When a CC receives a set of instances to run, it contacts each NC component through its *describeResource* operation and sends the *runInstances* request to the first NC that has enough free resources to host the instance.

- When a CC receives a *describeResources* request, it also receives a list of resource characteristics (cores, memory, and disk) describing the resource requirements needed by an instance (termed a *VM "type"*). With this information, the CC calculates how many simultaneous instances of the specific "type" can execute on its collection of NCs and reports that number back to the CLC.

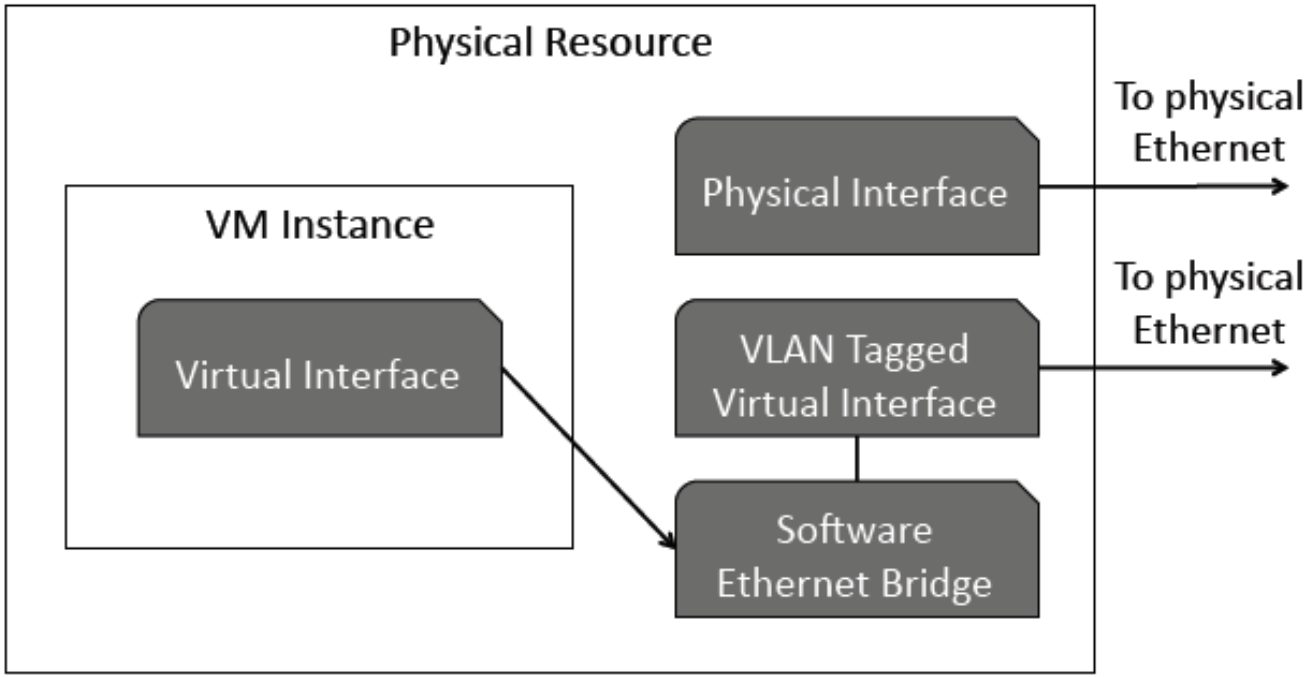
Virtual network overlay

- Problems:
  - Connectivity
    - Each VM should have connectivity with each other;
    - Some VM should be accessible externally.
  - Isolation
    - Super-user access might cause conflict in the uderlying physical resources, e.g. IP network and MAC addresses;
    - VMs belonging to a single cloud allocation (e.g. a single user) should be able to communicate, VMs belonging to different allocatoins should be isolated.
  - Performance.
    - Virtulization should provide comparable performance as native physical resources.

Virtual network overlay (Cont.d)

- Three configuration modes:
  - Attach the VM's interface directly to a software Ethernet bridge connected to the real physical machine's network, allowing the administrator to handle VM network DHCP requests the same way they handle non-EUCALYPTUS component DHCP requests;
  - Define static Media Access Control (MAC) and IP address tuples. In this mode, each new instance created by the system is assigned a free MAC/IP tuple, which is released when the instance is terminated.
  - Fully manages and controls the VM networks, providing VM traffic isolation, the definition of ingress rules (configurable firewalls) between logical sets of VMs, and the dynamic assignment of public IP addresses to VMs at boot and run-time.
  - The first two provide near-native inter-VM performance, but don't provide inter-VM network isolation.

Virtual network overlay (Cont.d)

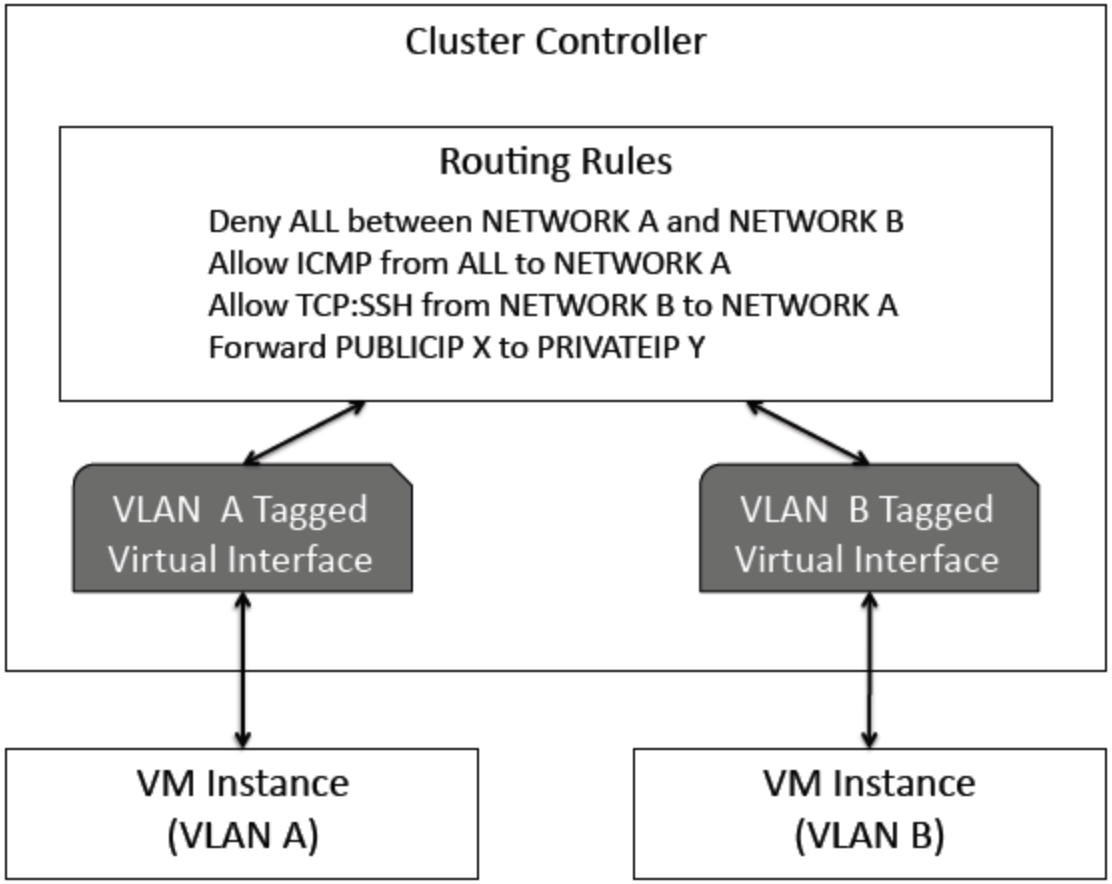


Each EUCALYPTUS VM instance is assigned a virtual interface that is connected to a software Ethernet bridge on the physical machine, to which a VLAN tagged interface is further connected.

Virtual network overlay (Cont.d)

- Users are allowed to attach VMs, at boot time, to a “network” that is named by the user.
- Each “network” is assigned a unique VLAN tag by Eucalyptus as well as a unique IP subnet from a range specified by the administrator of Eucalyptus in advance.
- The CC acts as a router between VM subnets.
- The CC uses the Linux iptables packet filtering system to implement and control VM network ingress rules.

Linux iptables packet filtering



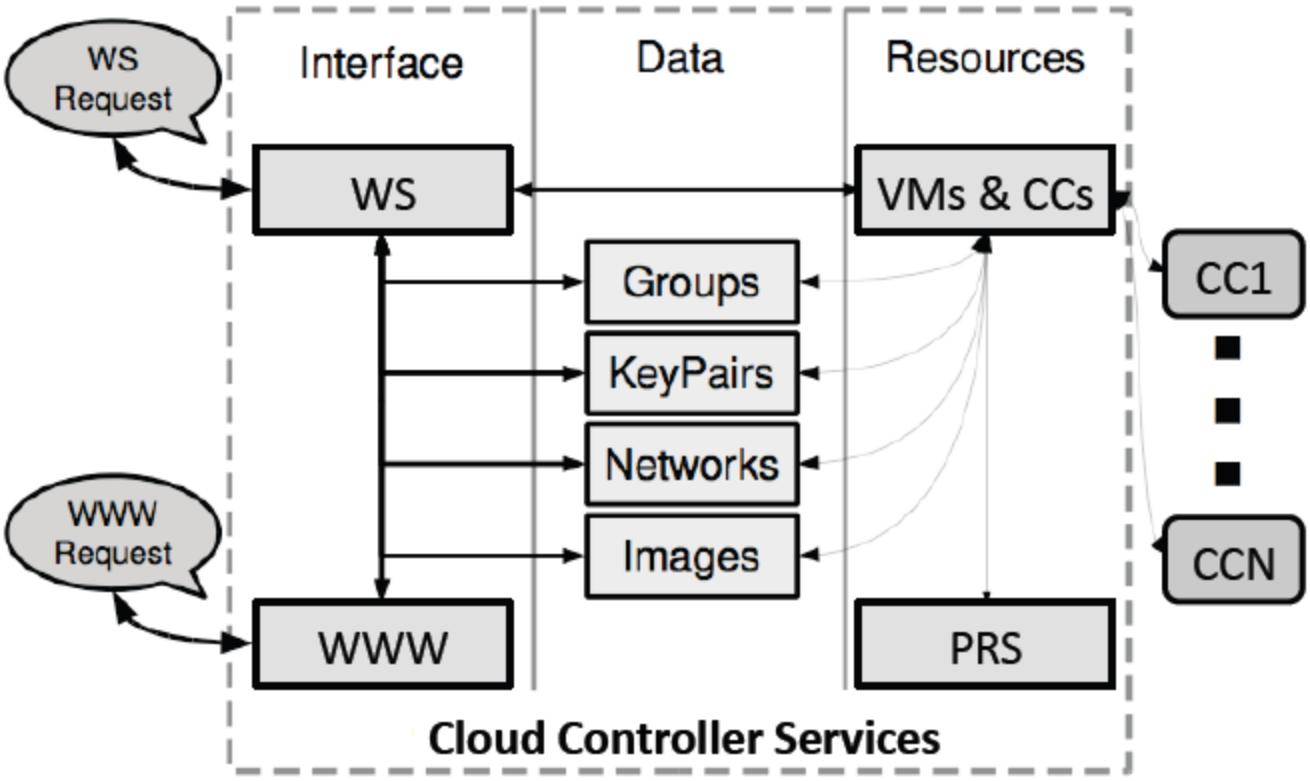
Storage controller (Walrus)

- Interfaces (compatible with Amazon S3)
  - REST (via HTTP), so-called ”Query” interface;
  - SOAP;
- Functions:
  - stream data into/out of the cloud as well as from instances that they have started on nodes;
  - acts as a storage service for VM images. Root file-system as well as kernel and ramdisk images used to instantiate VMs on nodes can be uploaded to Walrus and accessed from nodes.
- Supports concurrent and serial data transfers.
- Does not provide locking for object writes.
- When a node controller (NC) requests an image from Walrus before instantiating it on a node, it sends an image download request that is authenticated using an internal set of credentials. Then, images are verified and decrypted, and finally transferred.

Cloud controller (CLC)

- Resource Services
  - perform system-wide arbitration of resource allocations, let users manipulate properties of the virtual machines and networks, and monitor both system components and virtual resources.
- Data Services
  - govern persistent user and system data and provide for a configurable user environment for formulating resource allocation request properties.
- Interface Services
  - present user-visible interfaces, handling authentication & protocol translation, and expose system management tools providing.

Cloud controller (CLC) (Cont.d)



Glossary

- **Cloud** - A federated set of physical machines that offer computing resources through virtual machines, provisioned and recollected dynamically.
- **Cloud Controller (CLC)** - Eucalyptus component that provides the web UI (an https server on port 8443), and implements the Amazon EC2 API. There should be only one Cloud Controller in an installation of UEC. This service is provided by the Ubuntu *eucalyptus-cloud* package.
- **Cluster** - A collection of nodes, associated with a Cluster Controller. There can be more than one Cluster in an installation of UEC. Clusters are sometimes physically separate sets of nodes. (e.g. floor1, floor2, floor2).
- **Cluster Controller (CC)** - Eucalyptus component that manages collections of node resources. This service is provided by the Ubuntu *eucalyptus-cc* package.
- **EBS** - Elastic Block Storage. http://aws.amazon.com/ebs/
- **EC2** - Elastic Compute Cloud. Amazon's pay-by-the-hour, pay-by-the-gigabyte public cloud computing offering.
- **EKI** - Eucalyptus Kernel Image.
- **EMI** - Eucalyptus Machine Image.
- **ERI** - Eucalyptus Ramdisk Image.
- **Eucalyptus** - **E**lastic **U**tility **C**omputing **A**rchitecture for **L**inking **Y**our **P**rograms **T**o **U**seful **S**ystems. An open source project originally from the University of California at Santa Barbara, now supported by Eucalyptus Systems, a Canonical Partner.
- **Front-end** - Physical machine hosting one (or more) of the high level Eucalyptus components (cloud, walrus, storage controller, cluster controller).
- **Node** - A node is a physical machine that's capable of running virtual machines, running a node controller. Within Ubuntu, this generally means that the CPU has VT extensions, and can run the KVM hypervisor.
- **Node Controller (NC)** - Eucalyptus component that runs on nodes which host the virtual machines that comprise the cloud. This service is provided by the Ubuntu package *eucalyptus-nc*.
- **S3** - Simple Storage Service. Amazon's pay-by-the-gigabyte persistent storage solution for EC2. http://aws.amazon.com/s3/
- **Storage Controller (SC)** - Eucalyptus component that manages dynamic block storage services (EBS). Each 'cluster' in a Eucalyptus installation can have its own Storage Controller. This component is provided by the 'eucalyptus-sc' package.
- **UEC** - Ubuntu Enterprise Cloud. Ubuntu's cloud computing solution, based on Eucalyptus.
- **VM** - Virtual Machine.
- **VT** - Virtualization Technology. An optional feature of some modern CPUs, allowing for accelerated virtual machine hosting.
- **Walrus** - Eucalyptus component that implements the Amazon S3 API, used for storing VM images and user storage using S3 bucket put/get abstractions.

Open Cirrus
-A Global Cloud Computing Testbed

Motivation

- Applications researchers in areas such as machine learning and scientific computing can get access to large-scale cluster resources, e.g. data centers provided Amazon, Microsoft, Yahoo!, Google, and IBM.

- System researchers, who are developing the techniques and software infrastructure to support cloud computing, still have trouble obtaining low-level access to such resources.

- Open Cirrus aims to address this problem by providing a single testbed based on a range of heterogeneous distributed data centers for systems, applications and services.

Participants

HP Labs

Intel Research

Yahoo!

University of Illinois at Urbana Champaign

Karlsruhe Institute of Technology, Germany

Infocomm Development Authority, Singapore

Electronics and Telecommunications Research Institute

Malaysian Institute for Microelectronic Systems

Russian Academy of Sciences

Carnegie Mellon University

ChinaMobile

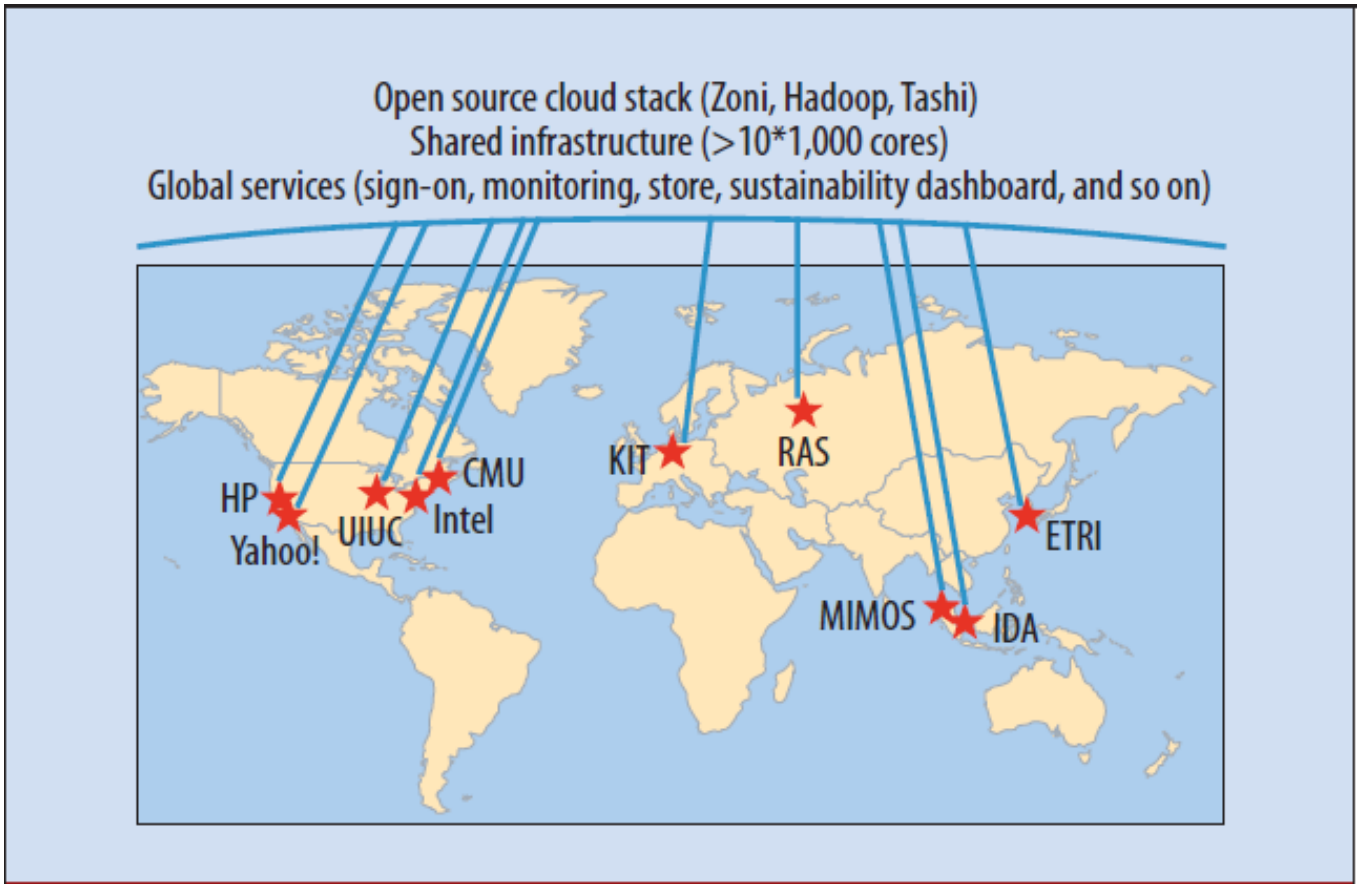
CESGA

ChinaTelecom

GaTech

Computer Network information Center
Chines Academy of Sciences

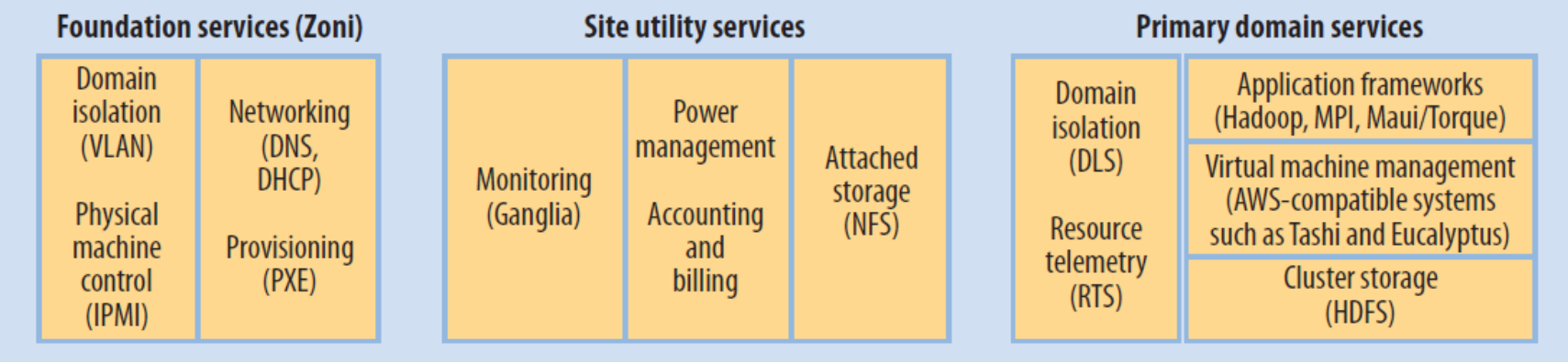
Geo-distribution



High-level architectural choices

- Systems versus application-only research.
  - Open Cirrus enables research using physical machines in addition to virtualized machines.
- Federated versus unified sites
  - Open Cirrus federates numerous sites with various hardware, services, and tools, in contrast to a unified architecture such as PlanetLab.
- Data-center focus versus centralized homogeneous infrastructure.
  - Open Cirrus revolves around multiple data centers, compared to a centralized approach such as Emulab.

Service stack architecture

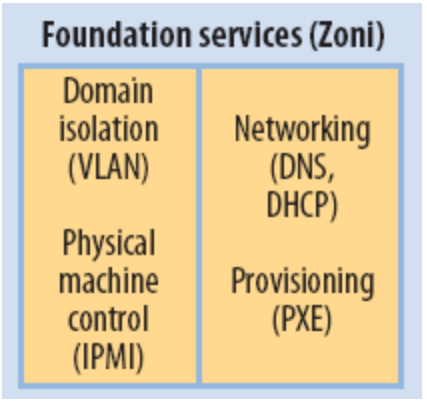


- PXE: Preboot Execution Environment
- IPMI: Intelligent Platform Management Interface

Service stack architecture
-Zoni

- Responsible for managing physical resources in the cluster and is crucial to providing users with bare-metal server access to conduct software system research.

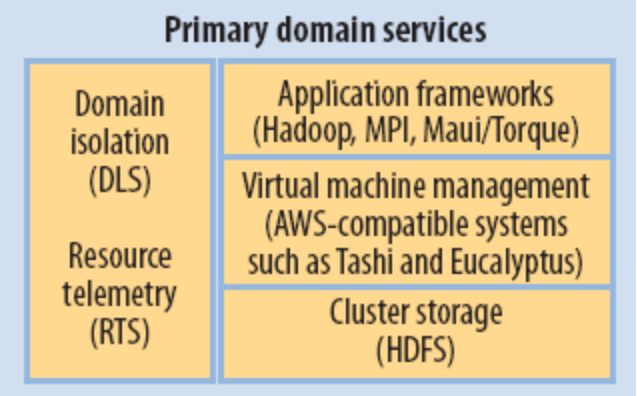
- Provides five key functions:
  - allocation of server nodes;
  - isolation of node groups, called domains;
  - provisioning of key software in a domain;
  - out-of-band server management; and
  - debugging of allocated nodes.



Service stack architecture
-Primary domain services

- To support users working with very large data sets, a cluster storage system, in particular the Hadoop distributed file system (HDFS), is used to aggregate the storage of all the nodes in the domain.

- To support a diverse set of user needs, the recommended primary domain services include a virtual machine management (VMM) layer, which provides a convenient mechanism for allocating resources to various users and services.
  - Hadoop
  - Maui + Torque
  - MPI



- Different sites may select any VMM service as long as it supports the EC2 interface from Amazon Web Services (AWS).
  - Tashi
  - Eucalyptus

- Data Location Service (DLS)
  - a clearinghouse for data location information independent of a storage mechanism,
- Resource Telemetry Service (RTS)
  - provides a means to obtain an abstract distance measure between two location identifiers.

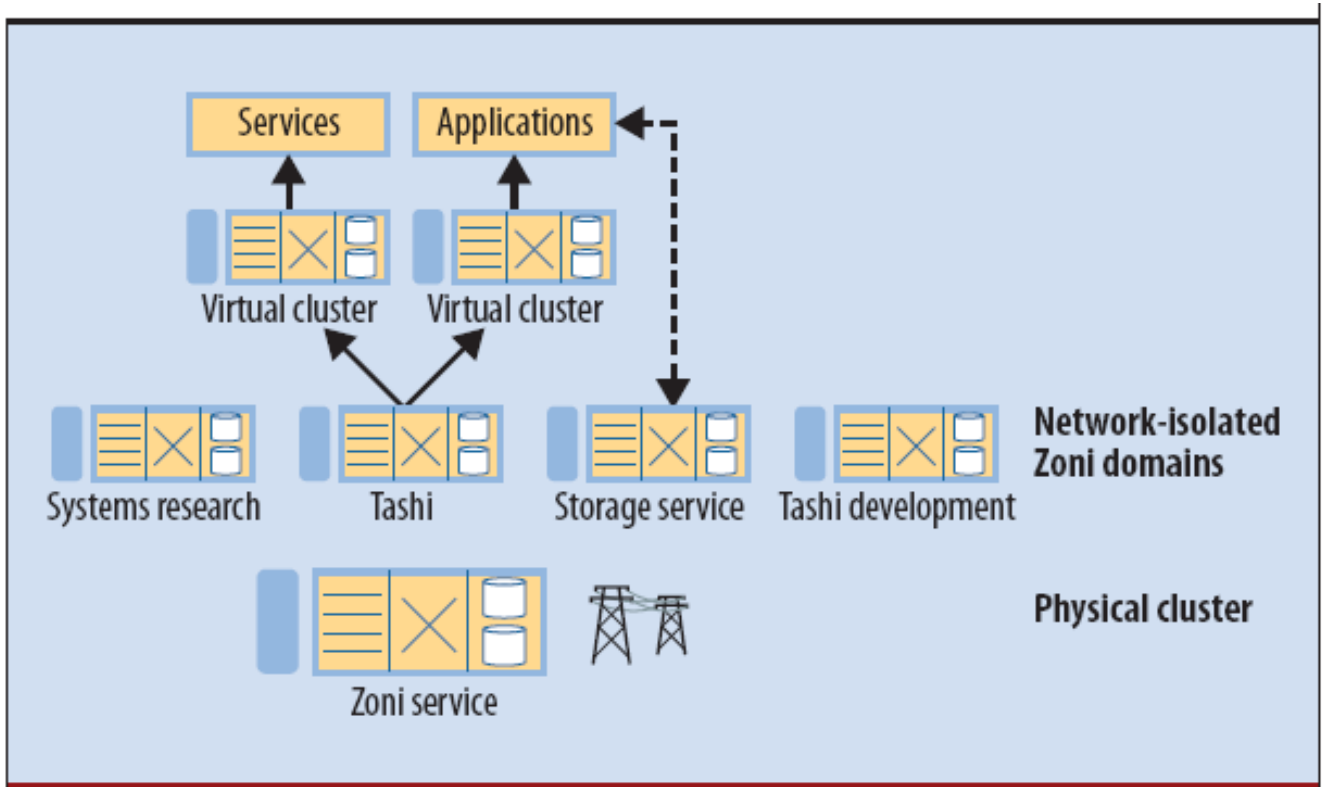
Tashi

- The Tashi project aims to build a software infrastructure for cloud computing on massive Internet-scale datasets (what is called Big Data). The idea is to build a cluster management system that enables the Big Data that are stored in a cluster/data center to be accessed, shared, manipulated, and computed on by remote users in a convenient, efficient, and safe manner.

- While Tashi is similar to other systems that manage logical clusters of VMs, it was developed to support research in coscheduling computation, storage, and power.

http://incubator.apache.org/tashi/

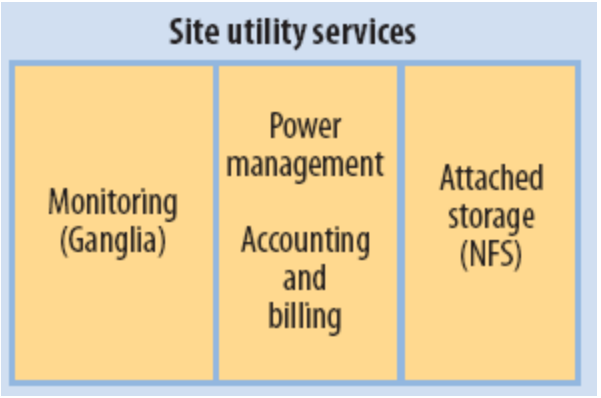
Example service in Open Cirrus



Service stack architecture
-Site utility services

- A monitoring service (such as Ganglia) not only enables the site administrator to monitor the cluster's health, it also facilitates collection of cluster operational data that may inform future research projects.

- Some conventional network file system storage is convenient for storing user scripts, small data sets, and small output files.



- Site utilities also include facilities for tracking resources consumed by users and managing the cluster's power consumption.

basic characteristics of the current Open Cirrus sites

Site	Cores	Servers	Public partition (servers)	Memory size (Tbytes)	Storage size (Tbytes)	Spindles	Network data rate	Focus
CMU	1,165	159	50	2.40	892	784	1 Gbps	Tashi, distributed file systems, applications/data sets
ETRI	1,024	256	200	0.50	128	256	1 Gbps	Large data sets, cloud infrastructure
HP	1,024	256	178	3.30	632	1,152	10 Gbps internal; 1 Gbps x-rack	Networking, federation
IDA	2,400	300	100	4.80	59+	600	1 Gbps	Applications based on Hadoop, Pig
Intel	1,364	198	198	1.77	610	746	1 Gbps	Tashi, Zoni, MPI, Hadoop
KIT	2,656	232	128	10.00	280	192	40 Gbps	Applications with high throughput
MIMOS	1,024	167	16	0.50	36	258	1 Gbps	Platform, tools, testing, security
UIUC	1,024	128	64	2.00	524	258	1 Gbps	Data sets, cloud infrastructure
RAS	1,136	142	600	9.10	36	142	1 Gbps	Hadoop, Tashi, Zoni, Pig, MPI
Yahoo!	3,200	480	400	2.40	1,200	1,600	1 Gbps	Hadoop, Pig

Approximately 100 research projects at 10 sites use Open Cirrus at the systems and applications levels.

Open Cirrus economic model

- Single site
  - Suppose a medium-sized company which needs the same resource as UIUC Open Cirrus site: 128 servers (1,024 cores) and 524 Tbytes.
  - AWS rates: US$0.12 per Gbyte/month and $0.10 per CPU-hour.
  - Renting a cloud:
    - Monthly storage cost : 524 × 1,000 × $0.12=$62,880
    - Total monthly cost: $62,880 + 1,024 × 24 × 30 × $0.10 = $136,608
  - Owning a cloud:
    - Amortized monthly costs: hardware (45%) + power (40%) + network (15%)
    - Service lifetime: M months
    - Monthly storage cost (assuming $300 1-Tbyte disks) and scaling for power and networking: 524 × $300/0.45/M=$349,333/M
    - Total monthly cost: $700,000/0.45/M + $7,500=$1,555,555/M + $7,500
  - The break-even point
    - Storage: $349,000/M < $62,880, or M > 5.55 months
    - Overall: $1,555,555/M + $7,500 < $136,608, or M > 12 months.

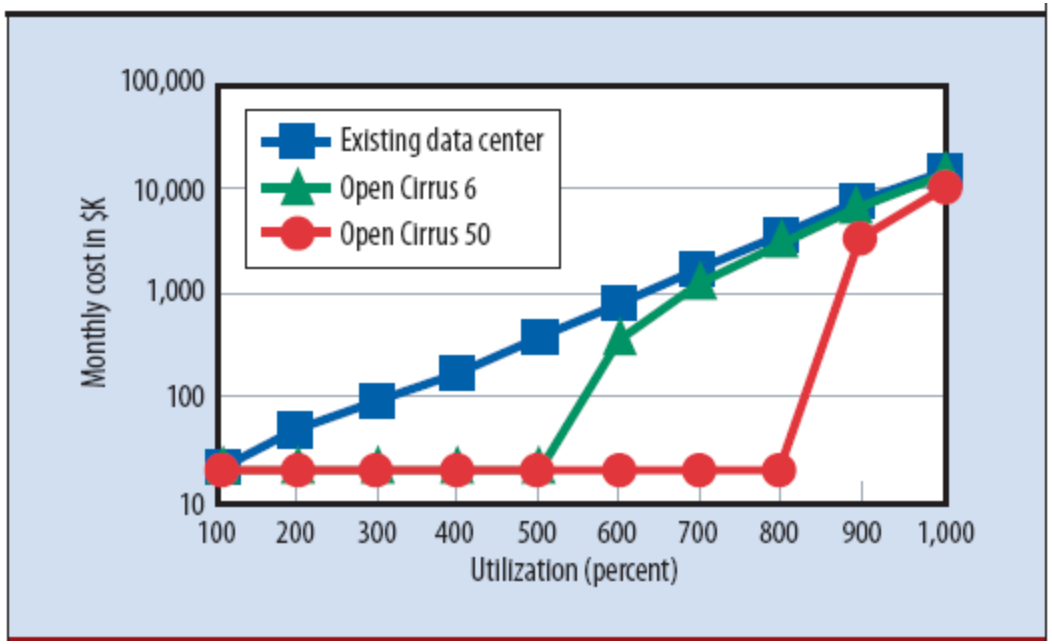
Conclusion:
If the service runs for more than 12 months, owning the cloud infrastructure is preferable to renting it.
Similarly, it's better to own storage if you use it for more than 6 months.

Open Cirrus economic model (Cont.d)

- Single site
  - Underutilization
  - With X percent resource utilization, the break-even time becomes 12 × 100/X months.
  - Given the typical hardware lifetime of 36 months, the break-even resource utilization is 12 × 100/X < 36, or X > 33.3%.

Conclusion:
Even at the current 20% CPU utilization rates observed in industry, storage utilization greater than 47% would make ownership preferable, as storage and CPU account evenly for costs.

Open Cirrus economic model (Cont.d) -Federated sites



Costs incurred by a single under-provisioned cloud for three options: offloading only to Amazon Web Services (existing data center), offloading to five federated clouds (Open Cirrus 6) and AWS, and offloading to 49 federated clouds (Open Cirrus 50) and AWS.

Comparison of cloud computing testbed

Characteristics	Open Cirrus	Google/ IBM cluster	TeraGrid	PlanetLab	Emulab	Open Cloud Testbed	Amazon EC2	LANL cluster
Type of research	Systems and services	Data-intensive applications	Scientific applications	Systems and services	Systems	Interoperability across clouds using open APIs	Commercial use	Systems
Approach	Federation of heterogeneous data centers	Cluster supported by Google and IBM	Multisite heterogeneous clusters for super-computing	Nodes hosted by research institution	Single-site cluster with flexible control	Multisite heterogeneous clusters	Raw access to virtual machines	Reuse of LANL's retiring clusters
Participants	CMU, ETRI, HP, Intel, IDA, KIT, MIMOS, RAS, UIUC, Yahoo!	Google, IBM, MIT, Stanford Univ., Univ. of Washington	Many universities and organizations	Many universities and organizations	Univ. of Utah	Four centers	Amazon	CMU, LANL, NSF
Distribution	10 sites	Central-ized—one data center in Atlanta	11 partners in US	More than 700 nodes worldwide	More than 300 machines	480 cores distributed in four locations	Several unified data centers	Thou-sands of older but still useful nodes at one site

Cloud services

- Cloud storage
  - Box.net
  - MobileMe (Apple)
  - Ovi store (Nokia)
- Cloud applications (Software as a service, SaaS)
  - Google App Engine
  - Salesforce
  - 37Signals
  - ZOHO
- Public cloud service
  - Amazon Web Services (AWS)
  - Microsoft Azure
  - Rackspace
- Cloud appliances
  - Pogoplug
  - Ctera
  - Tonidoplug

Cloud service

Cloud Applications
(Apps-as-a-Service)

Cloud (Applications) Platforms
(Platform-as-a-Service)

Cloud Infrastructure
(Infrastructure-as-a-Service)

Cloud services-Storage

- Box.net
  - cloud service that provides on line access to all your files and content, can not only store files and data, it has other services including document management, project management, plus FTP and other file transfers.
  - Billing
    - Free personal option for 1GB of online storage;
    - 10GB of data $9.99/month;
    - 15GB $19.99/month;
    - Business: $15/mo/user product for storing up to 15GB/user.

- Apple's MobileMe
  - Cloud computing product.
  - Store your email, contacts, and calendar and sync them to your iPhone, PC, and iPad.
  - A free 60 day trial, after 60 days, MobileMe will run you $99/year.
- Ovi Store

Cloud services-Cloud applications (Software as a Service, SaaS)

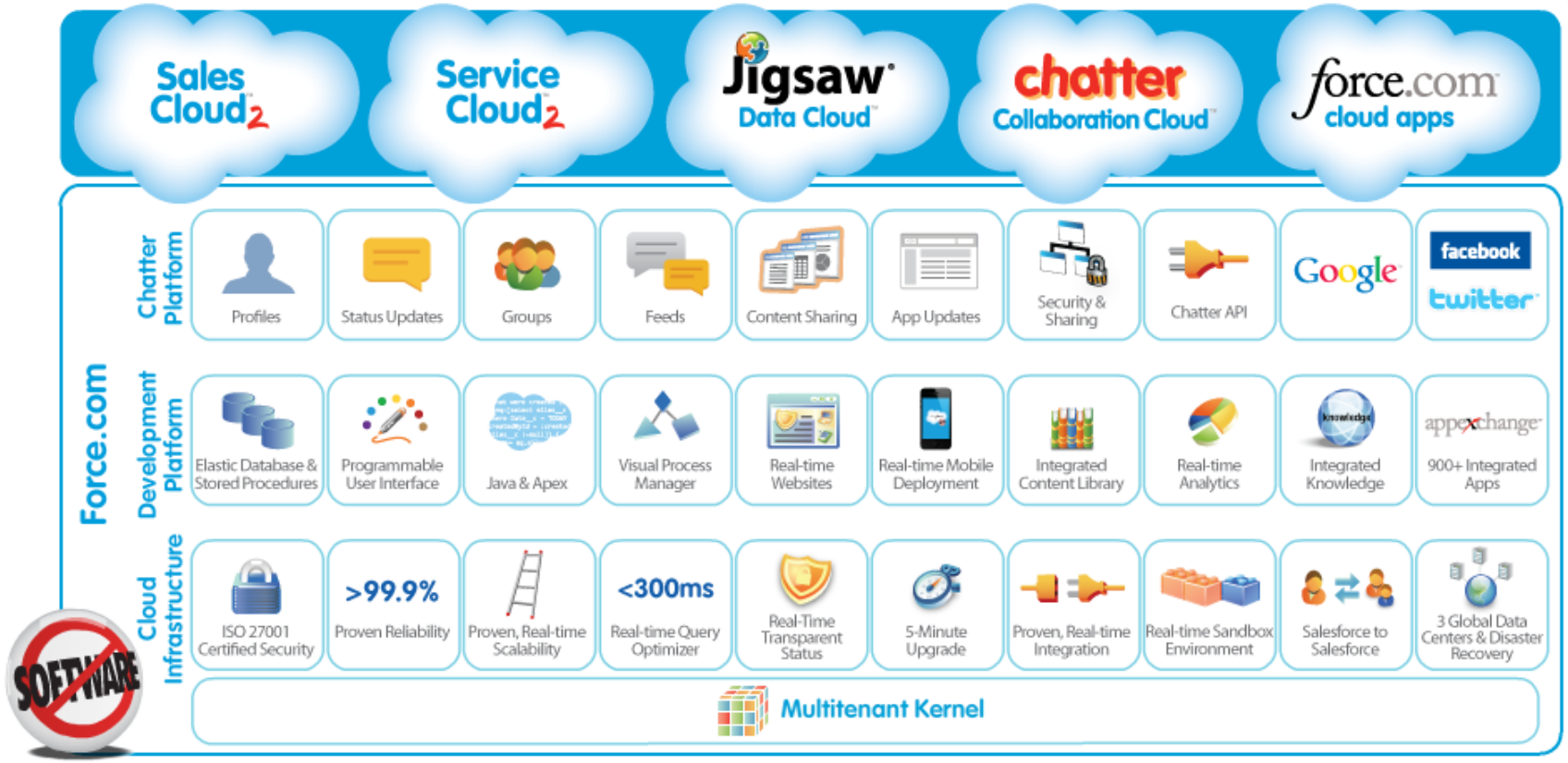
- Google
  - The GoogleApps suite for business provides email, calendaring, documents and other software for $50 per user per year.
    - It eliminates the need for installing and maintaining office applications like Outlook, Excel, Power Point and Word, and provides on line storage for emails and files.
  - Google App Engine

    - Enables you to build and host web apps on the same systems that power Google applications.
    - The sandbox isolates your application in its own secure, reliable environment.
    - Each application costs $8 per user, per month up to a maximum of $1000 a month.
    - SDK for JAVA, SDK for Python, Plugin for Eclipse.

Cloud services-Cloud applications (Salesforce)

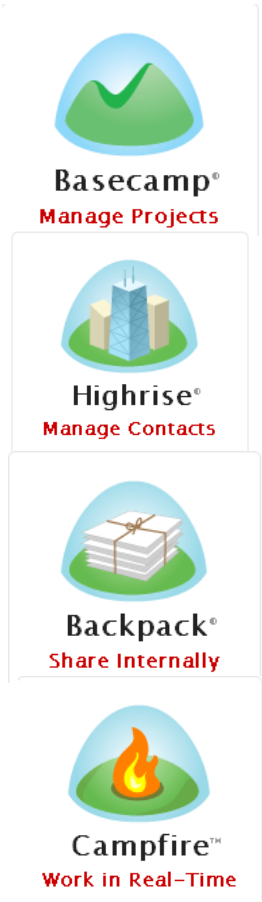
- Provides cloud Customer Relationship Management (CRM) software solutions.
- Sales Cloud
  - Has an amount of features, including a customer database, sales lead tools, workflow, integration to desktop applications (like Office), search tools, reporting, and access to other cloud applications.
  - Is priced at $5 to $250 per user per month depending on the features selected.
- Service Cloud
  - Includes customer trouble reporting and management tools, integration to social media sites like Twitter and Facebook, and other services to care for your customers.
  - Is priced at $65 to $265 per user per month, depending on the features selected.
- Chatter
  - Updates on people, groups, documents, and your application data come straight to you in your real-time feeds.
- Force.com
  - Give developers a platform to create rich, collaborative custom apps fast-without buying hardware or installing software.

Salesforce (Cont.d)



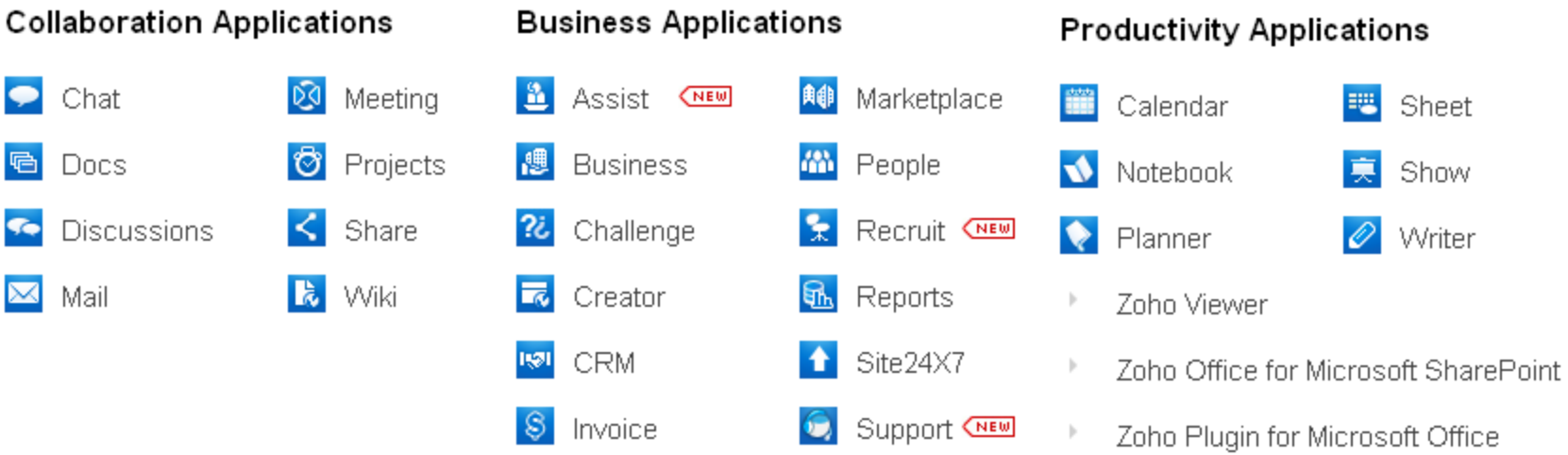
Cloud services-Cloud applications (37Signals)

- The “sole investor” in 37Signals is Jeff Bezos of Amazon.com.
- Basecamp
  - Project management and collaboration.
  - Is priced from $49/month to $149/month.
- Highrise
  - Contact and customer management.
  - Is priced from $24/month to $99/month.
- Backpack
  - Share information with the team, internal communication.
  - Is priced from $24/month to $149/month.
- Force.com
  - Team collaboration with real time chat.
  - Is priced from $12/month to $99/month.

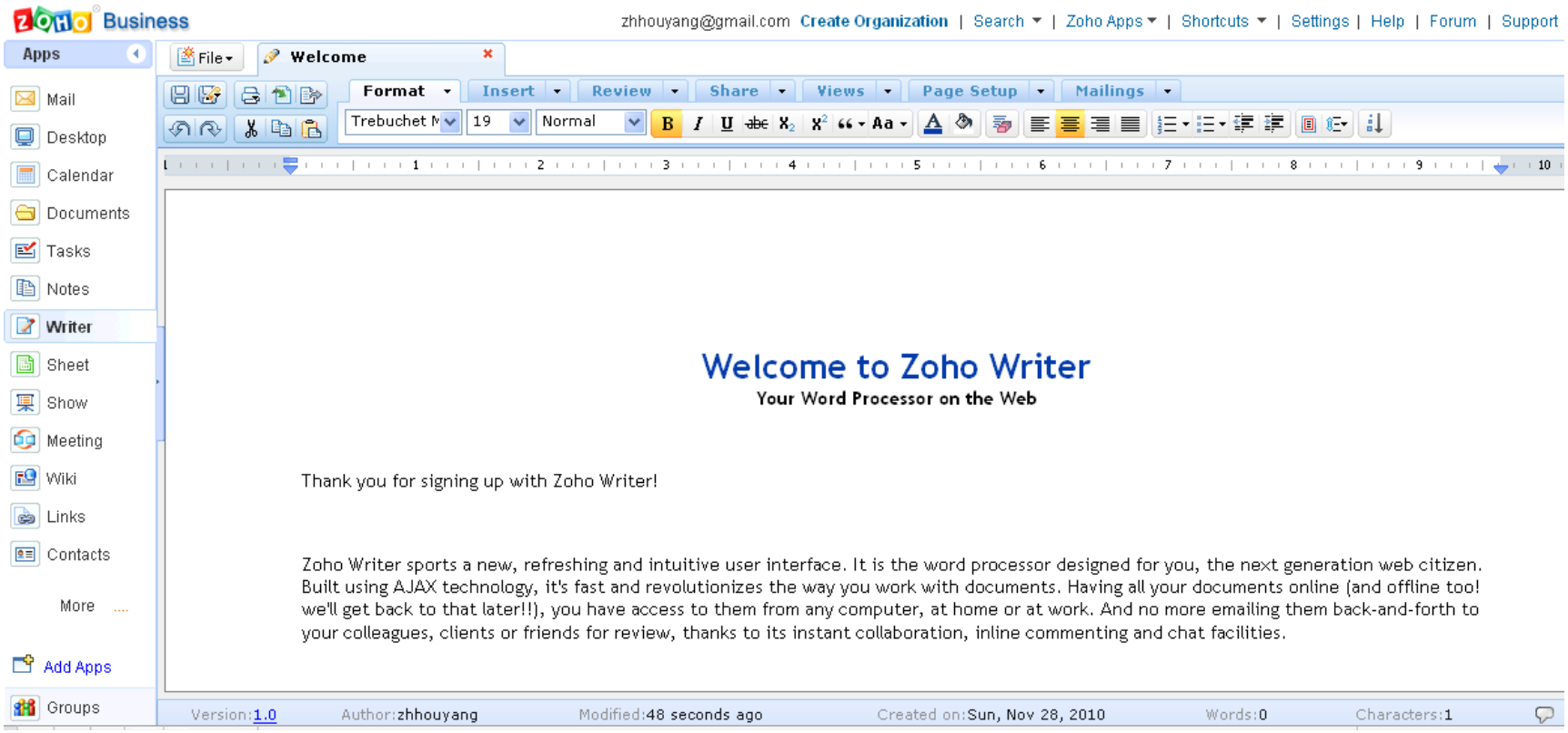


Cloud services-Cloud applications (ZOHO, SaaS)

- Zoho provides a wide, integrated portfolio of rich online applications for businesses.
- Services include:



ZOHO (Cont.d) Screenshot

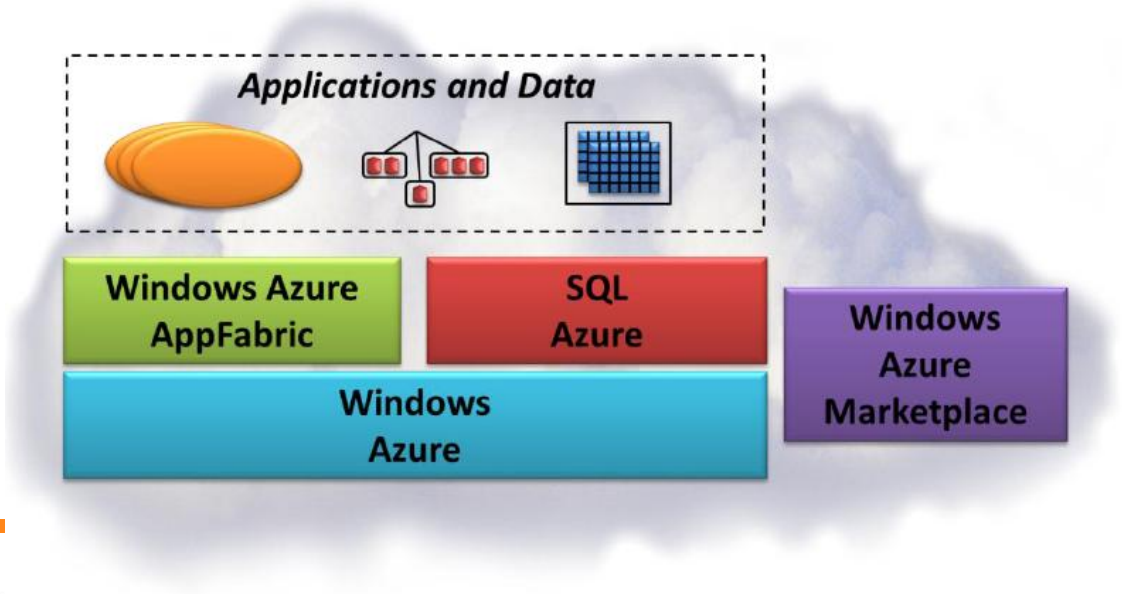


Public cloud service(AWS)
Infrastructure as a Service (IaaS)

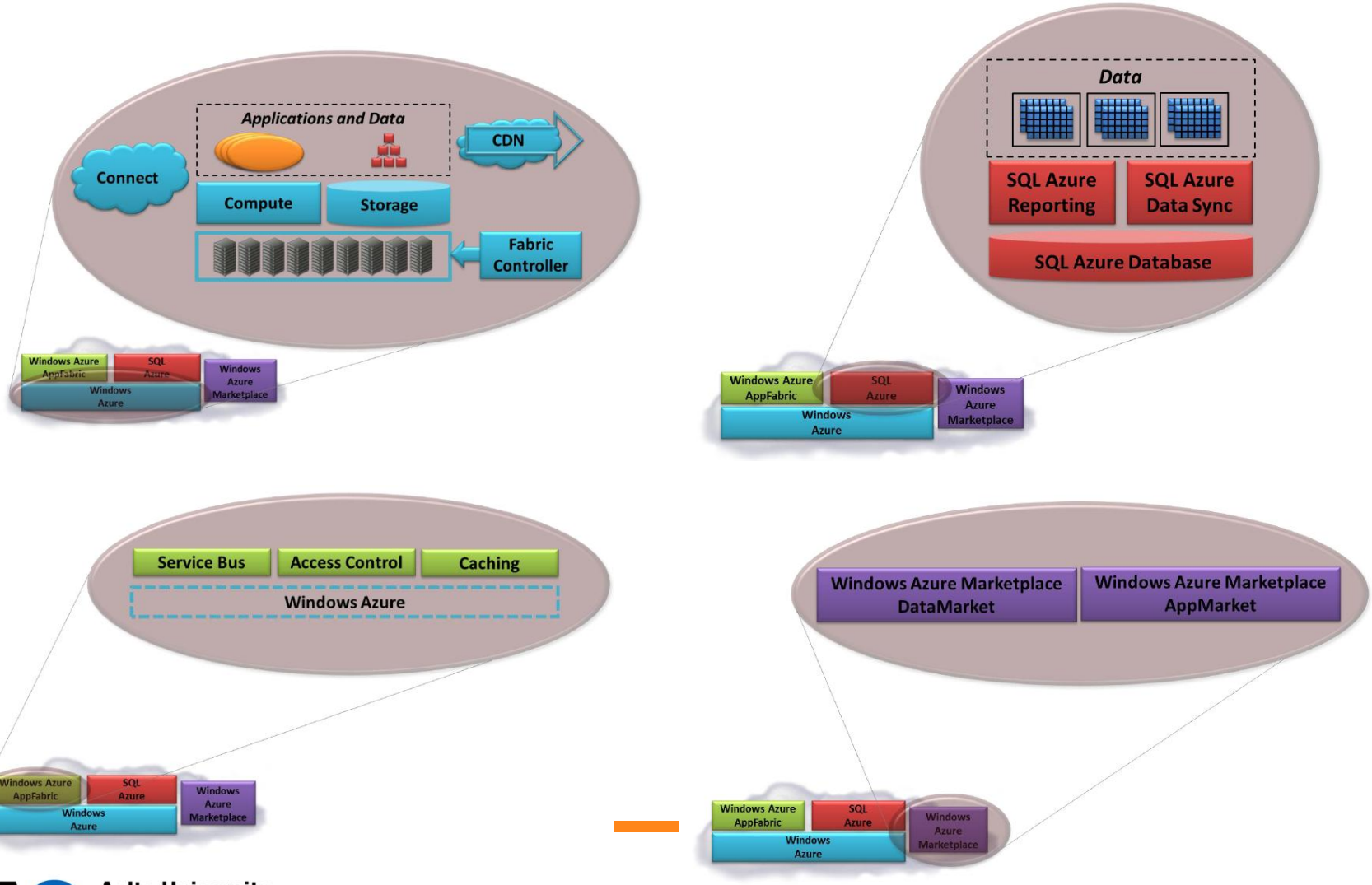
- One of the biggest public cloud providers.
- Has an incredible array of cloud computing services, called Amazon Web Services (AWS), including:
  - Amazon S3 (Simple Storage Service): cloud storage
  - Amazon EC2 (Elastic Compute Cloud): cloud computing
  - Amazon VPC (Virtual Private Cloud): secure bridge between private cloud and public cloud
  - Amazon Elastic MapReduce: processing data-intensive tasks
  - Amazon CloudFront: content delivery
  - Amazon RDS (Relational Database Service): cloud database
  - Amazon SNS (Simple Notification Service): cloud notification
- A basic Linux server starts at $.085 per hour and a Windows server at $.12 per hour.
- For Amazon S3 storage, $.15 per GB/month.
- Amazon has a solution for huge volume of storage with their Import/Export service, which allows for secure shipping of a USB or SATA drive to Amazon to copy data into the cloud. The data should be encrypted prior to shipping to keep it secure.

Public cloud service(Windows Azure) Platform as a Service (PaaS)

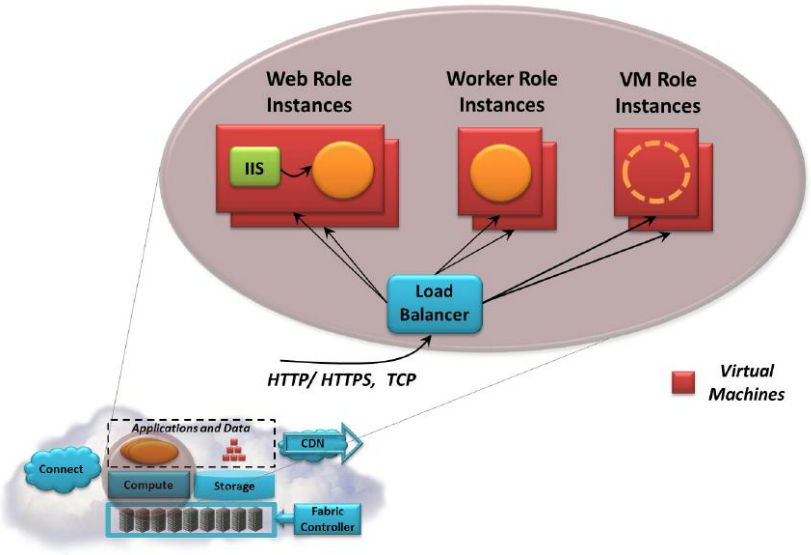
- Windows Azure provides what's commonly called Platform as a Service (PaaS).
  - It provides a platform that lets customers run applications without worrying about administering the environment they run in.
- A simple Windows server on Microsoft's cloud is $.12 per hour. Storage, as with Amazon is $.15 per GB/month.
- http://www.microsoft.com/windowsazure/resources/default.aspx?pmc=NO-CARE-01

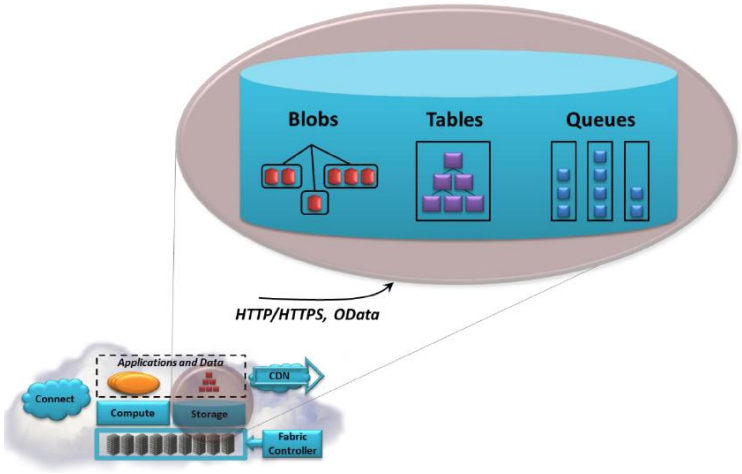


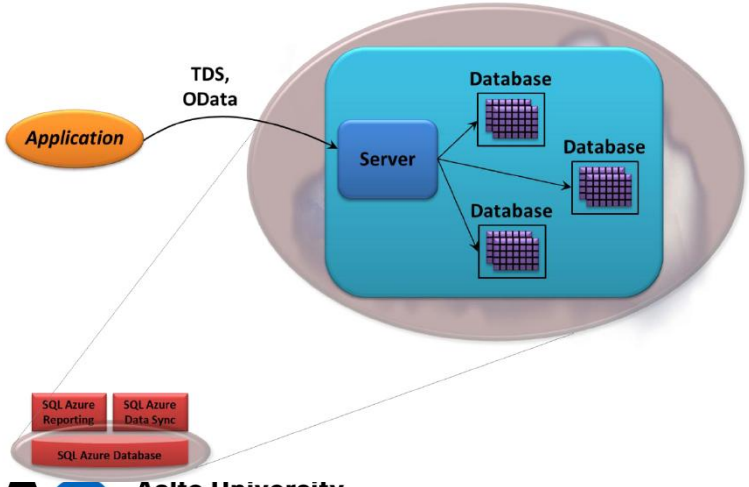
Windows Azure (Cont.d)

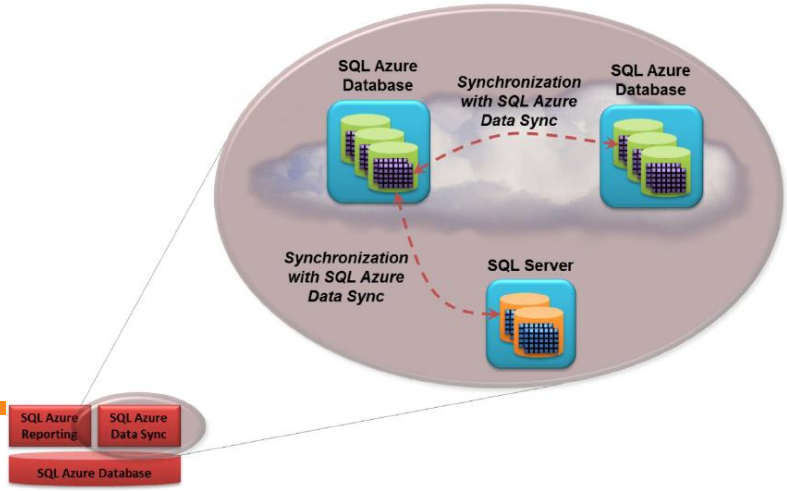


Windows Azure (Cont.d)

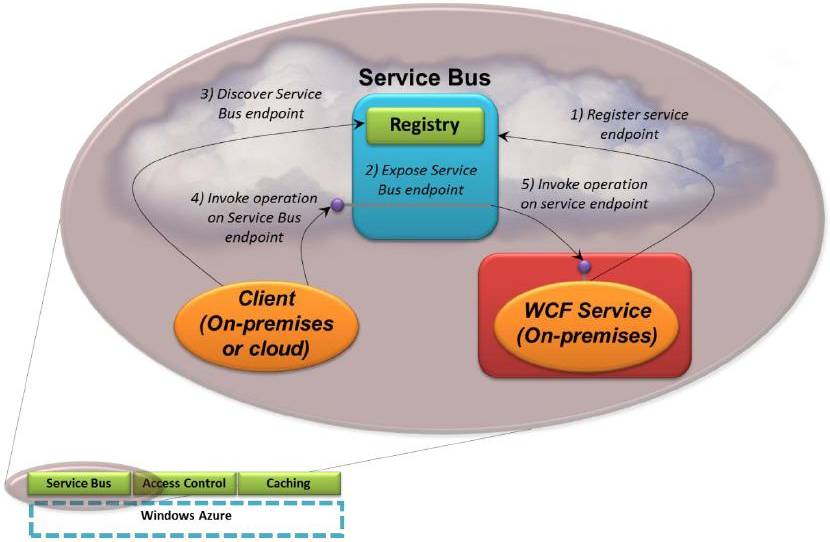


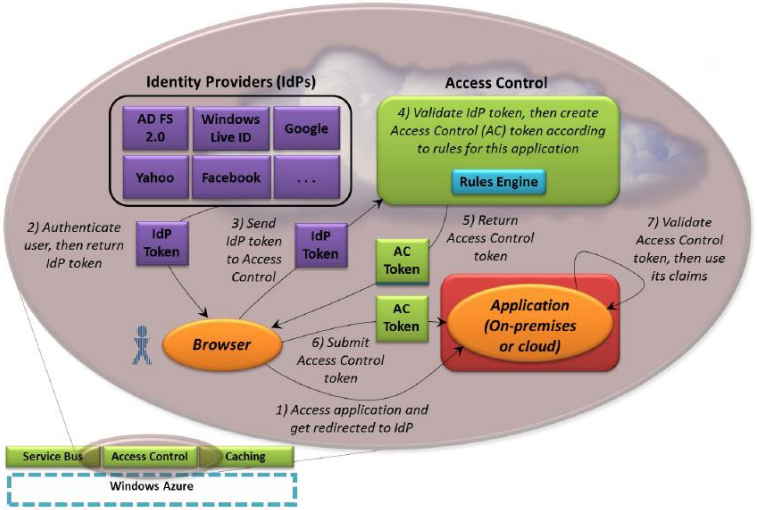


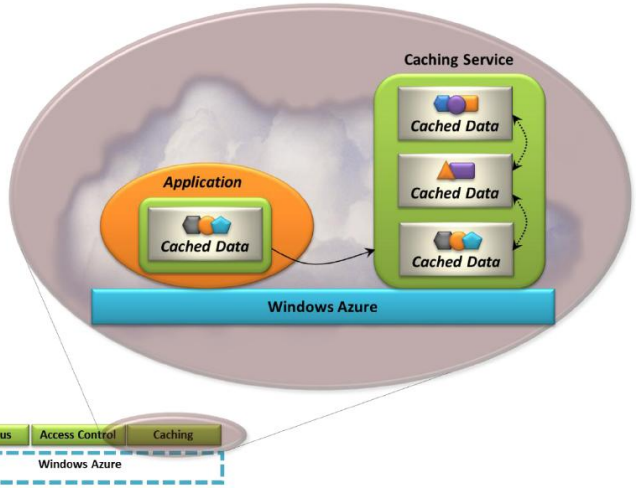


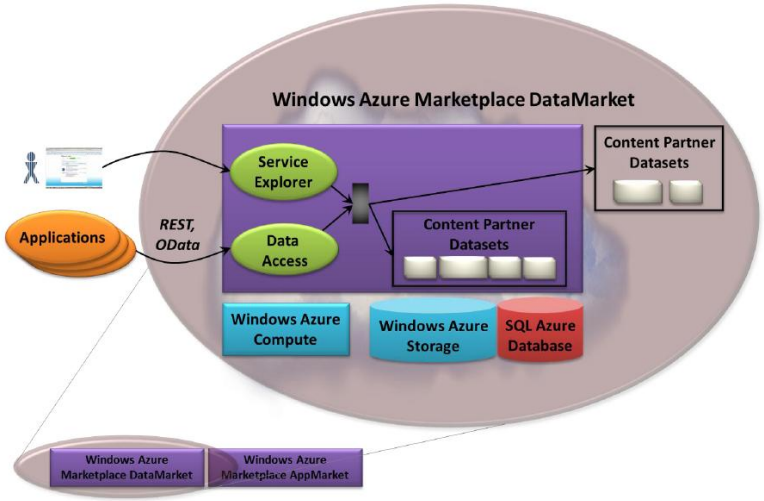


Windows Azure (Cont.d)









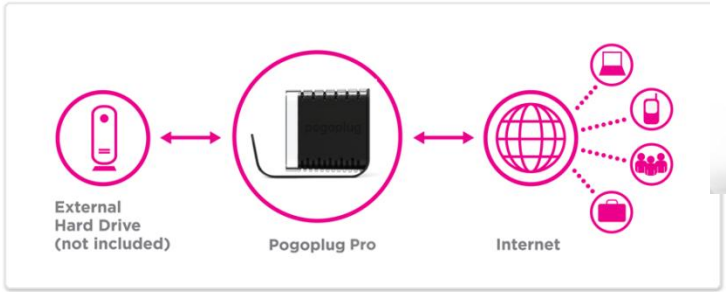
Public cloud services (Rackspace)

- Has been in the hosting business since 1998 and has 9 data centers throughout the world.

- Its cloud services include:
  - Cloudserver
    - Similar cloud service as Amazon EC2.
    - Starting at $0.015/ hour ($10.95 / month).
  - Cloudsites
    - Host scalable and reliable websites
    - Starting at $149 / month.
  - Cloudfiles
    - Provides unlimited file storage & hosting.
    - Similar cloud storage service as Amazon S3.
    - Starting at $0.15/GB.

Cloud appliances (Pogoplug)

- You connect one or more USB drives to the physical Pogoplug device. The Pogoplug software allows you to access the files from anywhere in the cloud, either on your network or over the Internet, with either a PC or a PDA.
- Products include:
  - Pogoplug
    - £99 / 99€
  - Pogoplug Pro
    - $99
    - Exclusively available now at U.S.
  - Pogoplug Biz
    - $299/£249/€249
    - Share massive amounts of content with clients and co-workers.



Cloud appliances (CTERA)

Enter the Age of Cloud Attached Storage™

- Provides storage and data protection for SMBs (Small and Medium Businesses) and enterprise branch offices, by combining cloud storage services with on-premises storage appliances.
- Access to more than more than 20,000 VARs (Value-Added Resellers) and MSPs (Managed Service Providers)
- Products include:
  - CloudPlug
    - Converts any external USB/eSATA drive into Network Attached Storage with secure cloud backup, remote access and collaboration services, and allows to share and synchronize files on your local network.
    - Approximately $200.
  - CTERA C200
    - Cloud Attached Storage appliance
    - Data is synchronized between individual PCs on the network and the C200 drives, then backed up using CTERA's integrated online backup service.
    - Approximately $371.
  - CTERA C400
    - Boasts up to 8TB of local storage space, with RAID5/6 capability and four hot-swappable drive bays.
    - Retails for a price of $1,499.

Cloud appliances (TonidoPlug)

- TonidoPlug is a tiny, low-power, low-cost personal home server and NAS device powered by Tonido software that allows you to access your files, music and media from anywhere via a web browser (Powered by Tonido ® software).

- Like the PogoPlug and Ctera C200, it is a physical device and requires supplying and connecting a USB drive for storage.

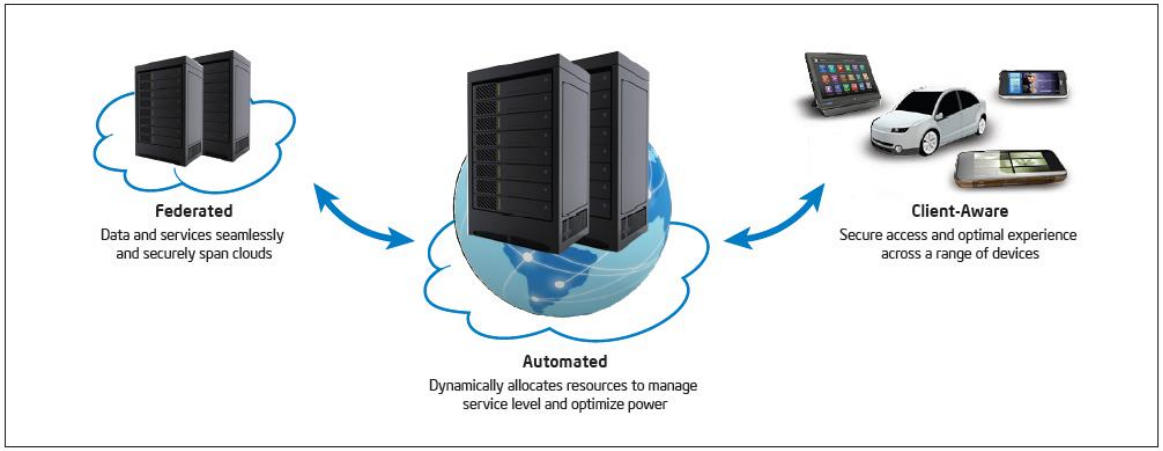
- Running on top of embedded Ubuntu Jaunty Linux OS, based on GHz ARM processor.

- Price: $99.

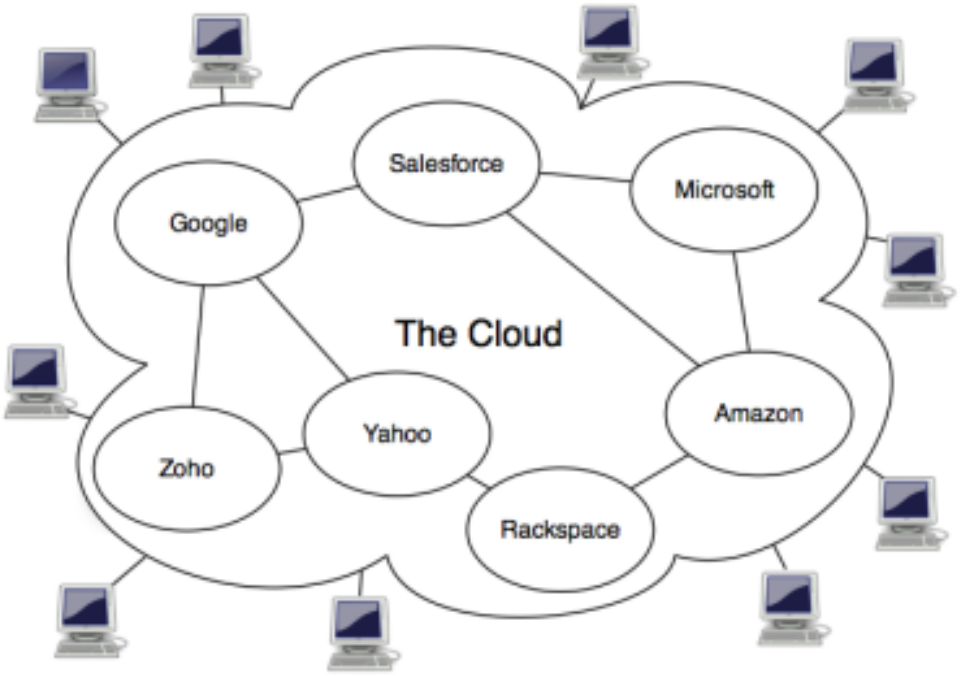
Future of Cloud computing

- According to MarketsandMarkets, the lucrative sector will increase from $37.8 billion in 2010 to $121.1 billion in 2015 at a CAGR of 26.2% from 2010 to 2015.

- Intel's cloud 2015 vision
  - Federated
  - Automated
  - Client-aware



Future of cloud computing (Cont.d)



References

- Virtualization for dummies. Bernard Golden. Wiley Publishing, Inc. ISBN: 978-0-470-14831-0.
- Wikipedia.
- http://www.smallcloudbuilder.com/everything-else/articles/123-introduction-to-small-cloud-computing-part-2
- David Chappell & associates. Introducing the Windows Azure platform.