CPET 581/499 Cloud Computing: Technologies and Enterprise IT Strategies

Lecture 6
Cloud Platform Architecture over Virtualized Data Centers
Part -1: Cloud Computing and Service Models

Text Book: Distributed and Cloud Computing, by K. Hwang, G C. Fox, and J.J. Dongarra, published Elsevier/Morgan Kaufmann, 2012.

Spring 2015

A Specialty Course for Purdue University's M.S. in Technology Graduate Program: IT/Advanced Computer App Track
Paul I-Hai Lin, Professor
Dept. of Computer, Electrical and Information Technology
Purdue University Fort Wayne Campus

Ch. 4 - Topics of Discussion

- Cloud Computing and Service Models
- Data-Center Design and Interconnection Networks
- Architectural Design of Computer and Storage Clouds
- Public Cloud Platforms: Google App Engine, Amazon Web Services and Microsoft Window Azure
- Inter-Cloud Resource Management
- Cloud Security and Trust Management

Cloud Computing & Service Models

- Cloud Services
  - Google App Engines (GAE) – Google Cloud Platform, https://developers.google.com/appengine/
  - https://cloud.google.com/appengine/docs
    - PaaS: Python, PHP, Java, GO
  - Amazon Web Services (AWS), http://aws.amazon.com/
    - AWS Elastic Beanstalk (PaaS), http://aws.amazon.com/elasticbeanstalk/
  - OpenShift – by Red Hat, https://www.openshift.com/
  - Microsoft Azure, http://www.windowsazure.com/en-us/
  - IBM Cloud, http://www.ibm.com/cloud-computing/us/en/paas.html
  - Oracle Cloud, https://cloud.oracle.com/home
    - Applications (SaaS), Platform (PaaS)
  - SAP HANA Cloud Platform, http://hcp.sap.com/platform.html
- Private and Hybrid Clouds

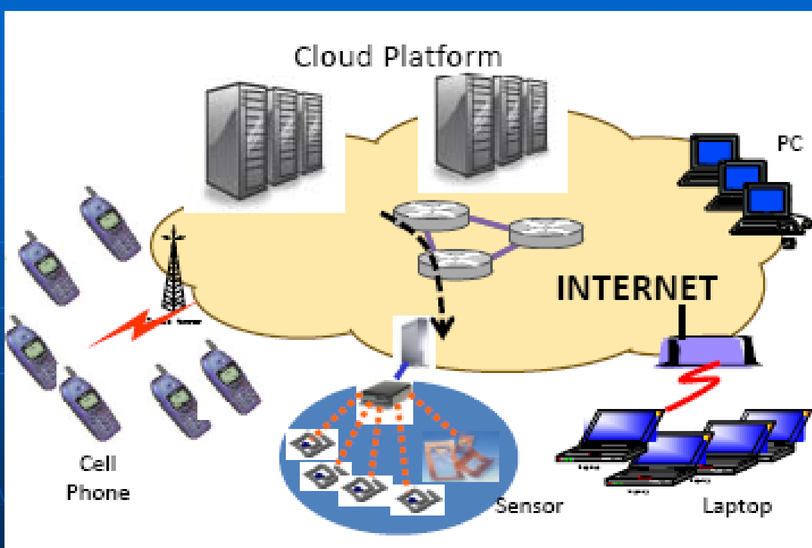
Cloud Computing: Info & References

- Conferences
  - IBM InterConnectGO, http://www.ibm.com/cloud-computing/us/en/interconnect/
  - Oracle Cloud Conference, http://www.oracle.com/us/corporate/events/cloud-conference-1405077.html
  - Microsoft
- Journals, Magazines
- Private and Hybrid Clouds

- Historical roots in today's Internet apps
  - Search, email, social networks
  - File storage (Live Mesh, Mobile Me, Flicker, …)
- A cloud infrastructure provides a framework to manage scalable, reliable, on-demand access to applications
- A cloud is the "invisible" backend to many of our mobile applications
- A model of computation and data storage based on "pay as you go" access to "unlimited" remote data center capabilities

(Courtesy of Geoffrey Fox, 2012)

Cloud Roles in The Future Internet



(Courtesy of G. Xie and Z. Li 2012)

Public, Private and Hybrid Clouds:
Functional Architecture and Connectivity

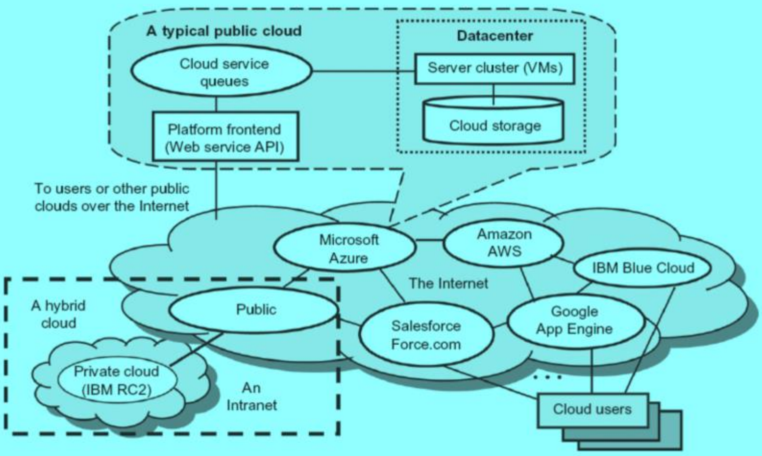
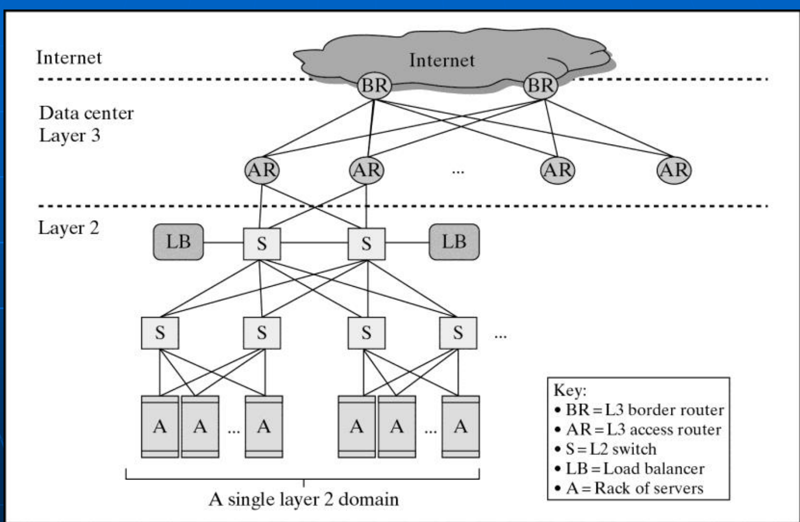


FIGURE 4.1
Public, private, and hybrid clouds.

Figure 4.2 Standard data-center networking for the cloud to access the Internet (Courtesy of Dennis Gannon, 2010 [26])



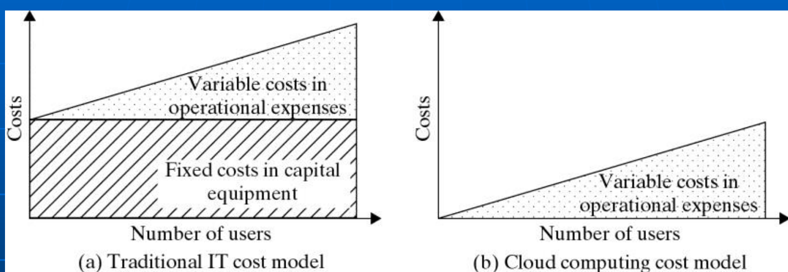
Three Aspects in Hardware that can benefit from Cloud Computing

1. The illusion of infinite computing resources available on demand, thereby eliminating the need for cloud users to plan far ahead for resource provisioning.
2. The elimination of an up-front commitment by cloud users, thereby allowing companies to start small and increase the hardware resources when needed in the future.
3. The ability to pay the costs of computing resources on a short-term basis as needed (e.g., processors by the hour and storage by the day) and release them after done and thereby rewarding resource conservation.

(Courtesy of M. Ambrust, et al 2009)

Cost Models Comparison

Figure 4.3 Computing economics between traditional IT users and cloud users



Cost-Effectiveness in Cloud Computing vs. Datacenter Utilization

(Courtesy of M. Ambrust, et al 2009)

$$\text{UserHours}_{\text{cloud}} \times (\text{revenue} - \text{Cost}_{\text{cloud}}) \geq$$

$$\text{UserHours}_{\text{datacenter}} \times \left(\text{revenue} - \frac{\text{Cost}_{\text{datacenter}}}{\text{Utilization}}\right)$$

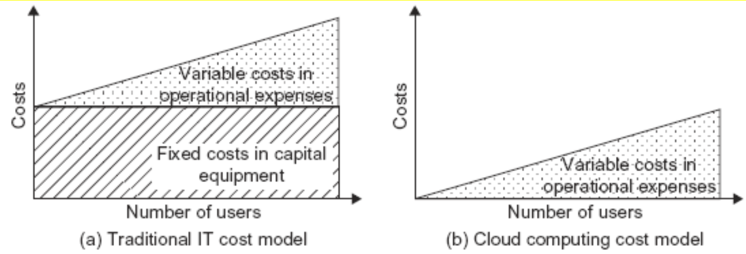


FIGURE 4.3

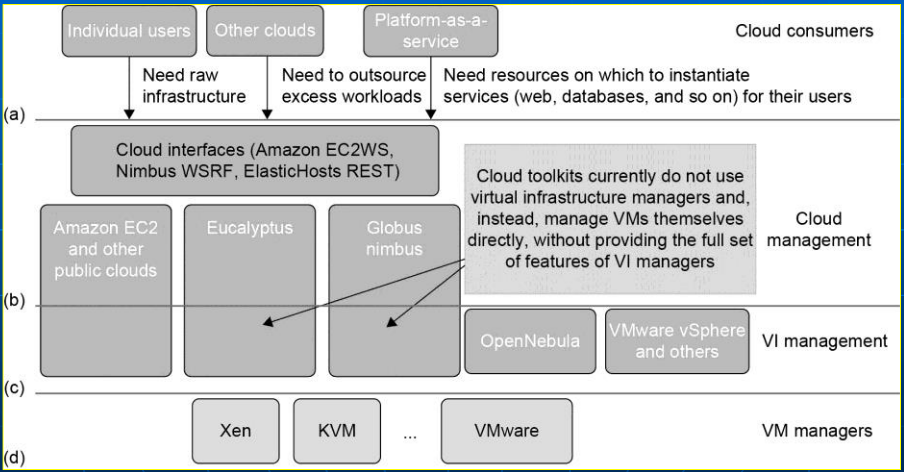
Computing economics between traditional IT users and cloud users.

source: Cloud Computing: What IT Professionals Needs to Know, Microsoft White Paper

Cloud Role Evolution

	Software as a Service	Platform as a Service	Infrastructure as a Service
Microsoft Products	Office 365 Microsoft Dynamics CRM Online Windows Intune	Windows Azure SQL Azure	Windows Server 2008 R2 System Center
On-Premise Roles	System Admin IT Generalist Enterprise Admin Security Specialist	Web Developer Database Admin Windows Developer Business Intelligence Developer	Enterprise Admin Virtualization Admin
Cloud Skills	Migration Management	Security Integration Identity Management	Planning Design Provisioning
Cloud Roles	Cloud Service Manager	Cloud Developer/ Data Steward	Datacenter Operator

Figure 4.4 Cloud ecosystems for building private clouds: 4-levels
(a) Consumers demand a flexible platform; (b) Cloud Manager provides virtual resources over an IaaS platform, (c) VI manager allocates VMs; (d) VM managers handle VMs installed on servers.



The IaaS, PaaS, and SaaS Cloud Service Models at Different Service Levels

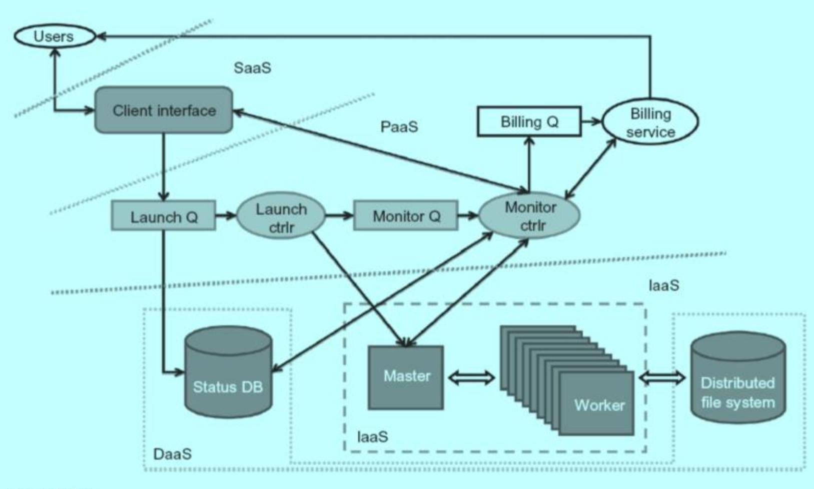


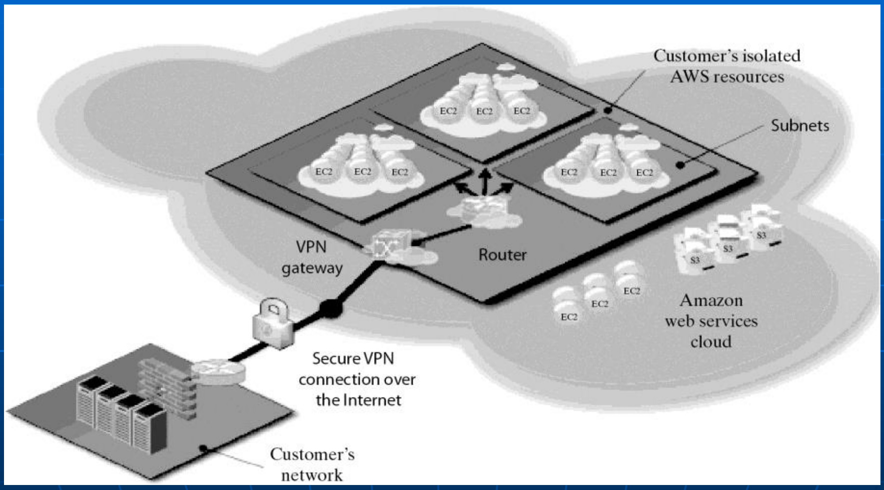
FIGURE 4.5

The IaaS, PaaS, and SaaS cloud service models at different service levels.

Example 4.1 Amazon VPC (Virtual Private Cloud) for Multiple Tenants

- Amazon VPC, http://aws.amazon.com/vpc/
- Additional EC2 instances or more storage (S3)
- Services
  - Resources from multiple data centers globally distributed
  - CLI (Command Line Interface), Web services (SOAP and Query), web-based console UI …
- Multiple Connectivity Options
  - Connect directly to the Internet (public subnets)
  - Connect to the Internet using Network Address Translation (private subnets)
  - Connect securely to your corporate datacenter
  - Combine connectivity methods to match the needs

Example 4.1 Amazon VPC (Virtual Private Cloud) for Multiple Tenants



Some IaaS Offerings from Public Clouds :

Table 4.1 Public Cloud Offerings of IaaS [10,18]

Cloud Name	VM Instance Capacity	API and Access Tools	Hypervisor, Guest OS
Amazon EC2	Each instance has 1–20 EC2 processors, 1.7–15 GB of memory, and 160–1.69 TB of storage.	CLI or Web Service (WS) portal	Xen, Linux, Windows
GoGrid	Each instance has 1–6 CPUs, 0.5–8 GB of memory, and 30–480 GB of storage.	REST, Java, PHP, Python, Ruby	Xen, Linux, Windows
Rackspace Cloud	Each instance has a four-core CPU, 0.25–16 GB of memory, and 10–620 GB of storage.	REST, Python, PHP, Java, C#, .NET	Xen, Linux
FlexiScale in the UK	Each instance has 1–4 CPUs, 0.5–16 GB of memory, and 20–270 GB of storage.	Web console	Xen, Linux, Windows
Joyent Cloud	Each instance has up to eight CPUs, 0.25–32 GB of memory, and 30–480 GB of storage.	No specific API, SSH, Virtual/Min	OS-level virtualization, OpenSolaris

4 - 18

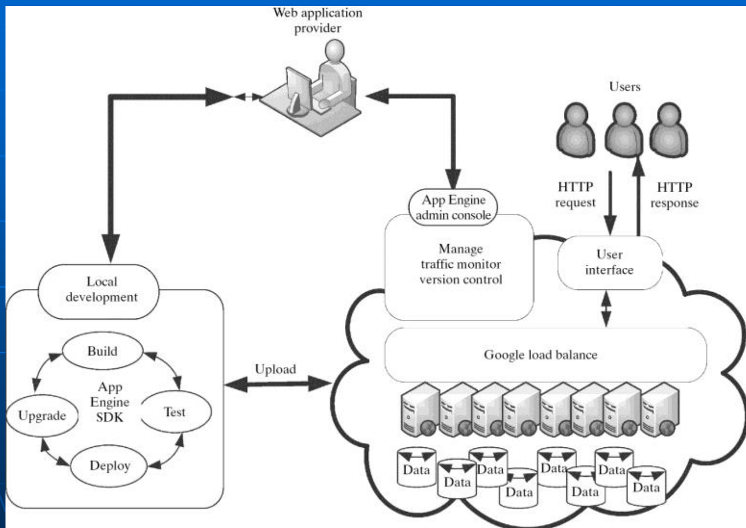
PaaS Offerings from Public Clouds

Table 4.2 Five Public Cloud Offerings of PaaS [10,18]

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.com's Force.com	Apex, Eclipse-based IDE, Web-based Wizard	Workflow, Excel-like formula, Web programming on demand	Business applications such as CRM
Microsoft Azure	.NET, Azure tools for MS Visual Studio	Unrestricted model	Enterprise and Web applications
Amazon Elastic MapReduce	Hive, Pig, Cascading, Java, Ruby, Perl, Python, PHP, R, C++	MapReduce	Data processing and e-commerce
Aneka	.NET, stand-alone SDK	Threads, task, MapReduce	.NET enterprise applications, HPC

4 - 19

Example 4.3 Google App Engine for Paas Applications (Figure 4.7)



Software as a Services (SaaS)

- Google Gmail and docs
- Microsoft SharePoint
- CRM from Salesforce.com
- 10 SaaS delivery companies to watch, June 4, 2012, Networkworld, by Christine Bums, http://www.networkworld.com/news/2012/060412-ecs-saas-companies-259409.html
  - AppDirect NetSuite
  - Concur Parallels
  - Ingram Micro Salesforce.com
  - Jamcrakr Standing Cloud
  - Ospero Workday

Cloud Computing as A Service

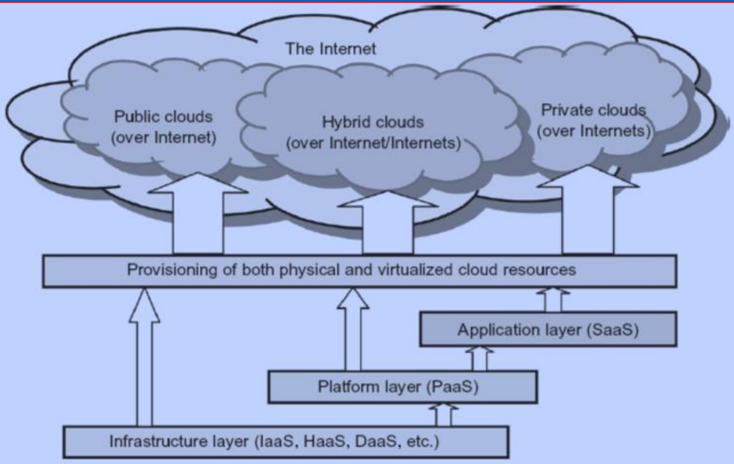


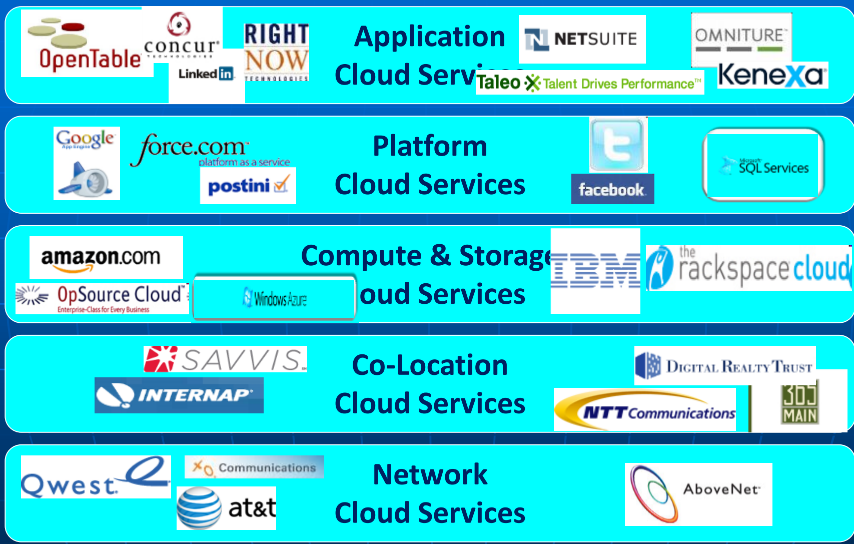
FIGURE 4.15

Layered architectural development of the cloud platform for IaaS, PaaS, and SaaS applications over the Internet.

Public Clouds vs. Private Clouds :

Characteristics	Public clouds	Private clouds
Technology leverage and ownership	Owned by service providers	Leverage existing IT infrastructure and personnel; owned by individual organization
Management of provisioned resources	Creating and managing VM instances within proprietary infrastructure; promote standardization, preserves capital investment, application flexibility	Client managed; achieve customization and offer higher efficiency
Workload distribution methods and loading policies	Handle workload without communication dependency; distribute data and VM resources; surge workload is off-loaded	Handle workload dynamically, but can better balance workloads; distribute data and VM resources
Security and data privacy enforcement	Publicly accessible through remote interface	Access is limited; provide pre-production testing and enforce data privacy and security policies
Example platforms	Google App Engine, Amazon AWS, Microsoft Azure	IBM RC2

Today's Cloud Services Stack



(Courtesy of T. Chou, 2010)

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

Challenges in Cloud Computing [1]

- Concerns from The Industry (Providers)
  - Replacement Cost
    - Exponential increase in cost to maintain the infrastructure
  - Vendor Lock-in
    - No standard API or protocol can be very serious
  - Standardization
    - No standard metric for QoS is limiting the popularity
  - Security and Confidentiality
    - Trust model for cloud computing
  - Control Mechanism
    - Users do not have any control over infrastructures

Challenges in Cloud Computing [2]

- Concerns from Research Community :
  - Conflict to legacy programs
    - With difficulty in developing a new application due to lack of control
  - Provenance
    - How to reproduce results in different infrastructures
  - Reduction in Latency
    - No specially designed interconnect used
    - Very low controllability in layout of interconnect due to abstraction
  - Programming Model
    - Hard to debug where programming naturally error-prone
    - Details about infrastructure are hidden
  - QoS Measurement
    - Especially for ubiquitous computing where context changes

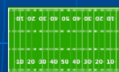
Cloud is built on Massive Datacenters

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

Economies 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

(Courtesy of Dennis Gannon, 2010)

Top 10 Obstacles and Opportunities for Cloud Computing

Table 6: Top 10 Obstacles to and Opportunities for Adoption and Growth of Cloud Computing.

	Obstacle	Opportunity
1	Availability of Service	Use Multiple Cloud Providers to provide Business Continuity; Use Elasticity to Defend Against DDOS attacks
2	Data Lock-In	Standardize APIs; Make compatible software available to enable Surge Computing
3	Data Confidentiality and Auditability	Deploy Encryption, VLANs, and Firewalls; Accommodate National Laws via Geographical Data Storage
4	Data Transfer Bottlenecks	FedExing Disks; Data Backup/Archival; Lower WAN Router Costs; Higher Bandwidth LAN Switches
5	Performance Unpredictability	Improved Virtual Machine Support; Flash Memory; Gang Scheduling VMs for HPC apps
6	Scalable Storage	Invent Scalable Store
7	Bugs in Large-Scale Distributed Systems	Invent Debugger that relies on Distributed VMs
8	Scaling Quickly	Invent Auto-Scaler that relies on Machine Learning; Snapshots to encourage Cloud Computing Conservationism
9	Reputation Fate Sharing	Offer reputation-guarding services like those for email
10	Software Licensing	Pay-for-use licenses; Bulk use sales

(Courtesy of M. Ambrust, et al 2009)

Table 2: Economies of Scale in 2006 for medium-sized datacenter (1000 servers) vs. very large datacenter (50,000 servers). [24]

Technology	Cost in Medium-sized DC	Cost in Very Large DC	Ratio
Network	$95 per Mbit/sec/month	$13 per Mbit/sec/month	7.1
Storage	$2.20 per GByte / month	$0.40 per GByte / month	5.7
Administration	≈140 Servers / Administrator	>1000 Servers / Administrator	7.1

Table 3: Price of kilowatt-hours of electricity by region [7].

Price per KWH	Where	Possible Reasons Why
3.6¢	Idaho	Hydroelectric power; not sent long distance
10.0¢	California	Electricity transmitted long distance over the grid; limited transmission lines in Bay Area; no coal fired electricity allowed in California.
18.0¢	Hawaii	Must ship fuel to generate electricity

(Courtesy of M. Ambrust, et al 2009)

Cost Considerations : Power, Cooling, Physical Plant, and Operational Costs

Table 5: We update Gray's costs of computing resources from 2003 to 2008, normalize to what $1 could buy in 2003 vs. 2008, and compare to the cost of paying per use of $1 worth of resources on AWS at 2008 prices.

	WAN bandwidth/mo.	CPU hours (all cores)	disk storage
Item in 2003	1 Mbps WAN link	2 GHz CPU, 2 GB DRAM	200 GB disk, 50 Mb/s transfer rate
Cost in 2003	$100/mo.	$2000	$200
$1 buys in 2003	1 GB	8 CPU hours	1 GB
Item in 2008	100 Mbps WAN link	2 GHz, 2 sockets, 4 cores/socket, 4 GB DRAM	1 TB disk, 115 MB/s sustained transfer
Cost in 2008	$3600/mo.	$1000	$100
$1 buys in 2008	2.7 GB	128 CPU hours	10 GB
cost/performance improvement	2.7x	16x	10x
Cost to rent $1 worth on AWS in 2008	**$0.27–$0.40** ($0.10–$0.15/GB × 3 GB)	**$2.56** (128× 2 VM's@$0.10 each)	**$1.20–$1.50** ($0.12–$0.15/GB-month × 10 GB)

(Courtesy of M. Ambrust, et al 2009)

Conclusion and Summary