INTRODUCTION TO CLOUD COMPUTING

CIT 3400

LECTURE 1

DR. AMOS CHEGE, PH.D.

LECTURER COMPUTER SCIENCE

"Computation may someday be organized as a public utility."

- John McCarthy, 1961

CLOUD COMPUTING

- No longer the next big thing the current big thing
 - Began in 2007 IBM and Google "Blue Cloud"
 - Name cloud inspired by cloud symbol representing internet in diagrams
 - Amazon popularized idea of the cloud

QUESTIONS TO ANSWER

• What clouds have you used today (yesterday)?

• What is a cloud?

CLOUD DEFINITION

Cloud computing is a set of service-oriented architectures, which allow users to access a number of resources in a way that is elastic, cost-efficient, and on-demand.

CLOUD DEFINITION

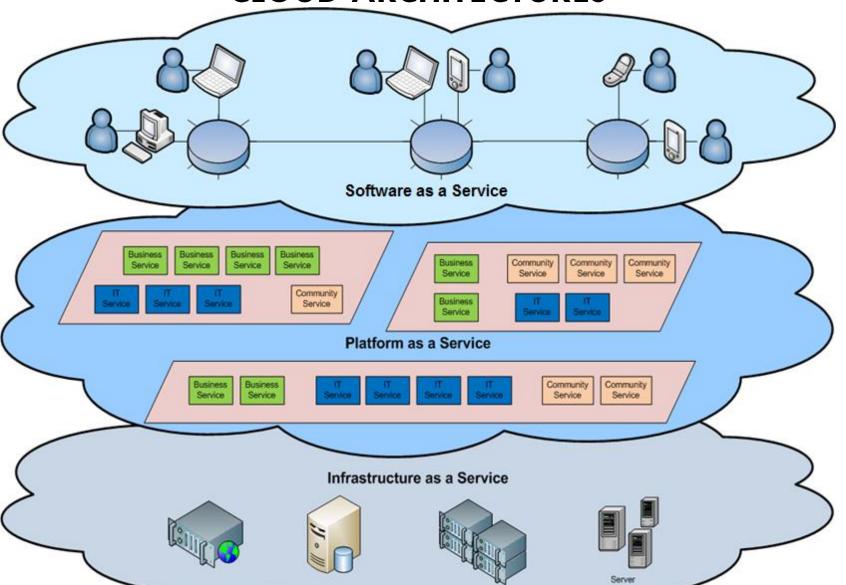
The U.S. National Institute of Standards and Technology (NIST) defines cloud computing as:

• Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

CLOUD ARCHITECTURES

- Scalable resource allocation
- Tailored services
 - Software as a Service (SaaS)
 - Platform as a Service (PaaS)
 - Infrastructure as a Service (laaS)

CLOUD ARCHITECTURES



CLOUD COMPUTING LAYERS

- Application Service (SaaS)
 - MS Live/Exchange, Google Docs, Salesforce.com, Quicken Online, Jupyter
- Application Platform (PaaS)
 - Google App Engine, Heroku, AWS
- Server Platform (laaS)
 - Google Compute Engine, Amazon EC2, OpenStack, Eucalpytus

CLOUD COMPUTING LAYERS

	Services	Description
	Services	Services – Complete business services such as PayPal, OpenID, OAuth, Google Maps, Alexa
Application Focused	Application	Application – Cloud based software that eliminates the need for local installation such as Google Apps, Microsoft Online
	Development	Development – Software development platforms used to build custom cloud based applications (PAAS & SAAS) such as SalesForce
Infrastructure Focused	Platform	Platform – Cloud based platforms, typically provided using virtualization, such as Amazon ECC, Sun Grid
	Storage	Storage – Data storage or cloud based NAS such as iCoud, Dropbox, CloudNAS
	Hosting	Hosting – Physical data centers such as those run by IBM, HP, Amazon, etc.

CLOUD SUMMARY

- Cloud computing is an umbrella term used to refer to Internet based development and services.
- Characteristics of cloud data, applications, services, and infrastructure:
 - Remotely hosted: Services and data are hosted on remote resources.
 - Ubiquitous: Services and data are available from anywhere.
 - **Commodified:** The result is a utility computing model similar to traditional utilities such as electricity and water.
 - You pay for what you use!

CLOUD COMPUTING

Everyone has an opinion on what to use a cloud for

- Applications on the internet email, tax prep
- Storage for business, personal data
- Web services for photos, maps, GPS
- Rent a virtual server, load software on it, turn it on /off, clone it if sudden workload demand increases
- Store, secure data for authorized access (really?)
- Use a platform including OS, Apache, MySQL, Python, PHP

CLOUD COMPUTING CHARACTERISTICS

• So what are its characteristics?

- Described as: On-demand computing, pay as you go, software as a service,
 utility computing
- Usually costs, but cost-effective
- Emphasizes availability
- Virtualization
- Scalable (expand on current hardware)
- Elastic (dynamically add hardware as needed by application/user)
- Distributed and highly parallel approach
- Replication, replication, replication ...

CHARACTERISTICS OF CLOUD COMPUTING

On-demand self service:

• Cloud computing resources can be provisioned on-demand by the users, without requiring interactions with the cloud service provider. The process of provisioning resources is automated.

• Broad network access:

• Cloud computing resources can be accessed over the network using standard access mechanisms that provide platform-independent access through the use of heterogeneous client platforms such as workstations, laptops, tablets and smartphones.

CHARACTERISTICS OF CLOUD COMPUTING

• Resource pooling:

• The computing and storage resources provided by cloud service providers are pooled to serve multiple users using multi-tenancy. Multi-tenant aspects of the cloud allow multiple users to be served by the same physical hardware.

Rapid elasticity:

• Cloud computing resources can be provisioned rapidly and elastically. Cloud resources can be rapidly scaled up or down based on demand.

CHARACTERISTICS OF CLOUD COMPUTING

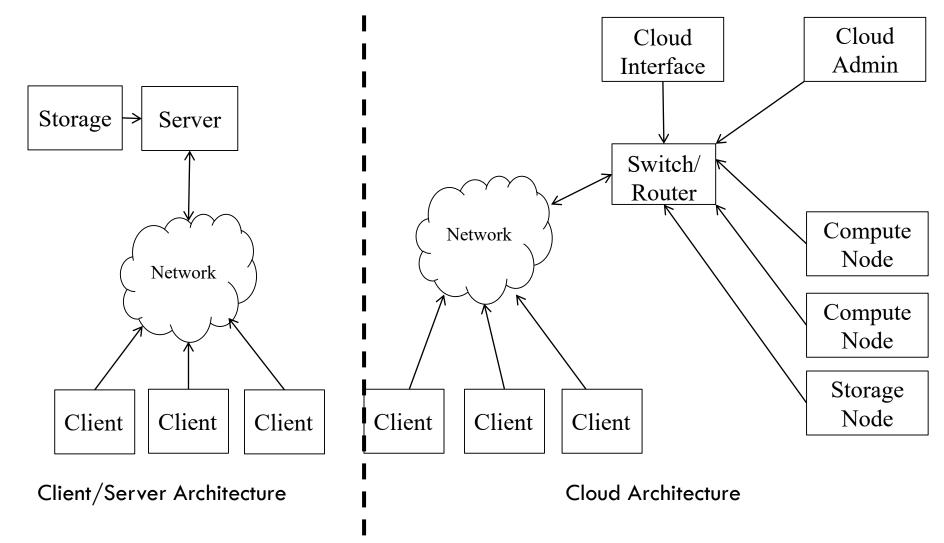
• Reliability:

• Applications deployed in cloud computing environments generally have a higher reliability since the underlying IT infrastructure is professionally managed by the cloud service.

Multi-tenancy:

- The multi-tenanted approach of the cloud allows multiple users to make use of the same shared resources.
- In virtual multi-tenancy, computing and storage resources are shared among multiple users.
- In organic multi-tenancy every component in the system architecture is shared among multiple tenants

CLIENT/SERVER VS. CLOUD ARCHITECTURE



TYPES OF CLOUDS

Public Cloud

- Marketed based on
 - Resources offered, availability, security, price

Local/Private Cloud

• Cloud architectures tailored to an organization's needs.

Hybrid Cloud

Combination of public and local cloud resources.

SUPER CLOUDS









Infrastructure Services

Storage

- Amazon S3 Amazon EBS CTERA Portal
- Mosso Cloud Files Nirvanix

Compute

- Amazon EC2 Serve Path GoGrid Elastra
- Mosso Cloud Servers Joyent Accelerators
- AppNexus Flexiscale
- Elastichosts Hosting.com CloudNine
- Terramark GridLayer
- ITRICITY LayeredTech

Services Management

- RightScale enStratus Scalr
- CohesiveFT Kaavo
 - CloudStatus Ylastic
- Dynect CloudFoundry
- NewRelic
 - Cloud42

Cloud Software

Globus Toolkit Xeround

Oracle Coherence Gemstone Gemfire Apache CouchDb

10Gen MongoDB

Data

- Apache HBase Hypertable TerraCotta
- Tokyo Cabinet Cassandra memcached -

Appliances

PingIdentity -Symplified rPath -Vordel .

Compute

- Beowulf Sun Grid Engine Hadoop OpenCloud Gigaspaces DataSynapse Xeround
 - File Storage EMC Atmos -
 - ParaScale -Zmamda -CTERA

Cloud Management

3Tera App Logic OpenNebula Open.ControlTier **Enomaly Enomalism** Altor Networks VMware vSphere OnPathTech CohesiveFT VPN Cubed Hyperic Eucalyptus Reductive Lbs Puppet

OpenQRM

Appistry

CLOUD TAXONOMY

Platform Services

General Purpose - Force.com

- Etelos LongJump
- AppJet . Rollbase
- Bungee Labs Connect
- Google App Engine Engine Yard Caspio
- Qrimp MS Azure Services Platform

- Mosso Cloud Sites

Business Intelligence

- Aster DB Quantivo Cloud9 Analytics Blink Logic **K2** Analytics

LogiXML Oco Panorama PivotLink Sterna ColdLight Neuron

Infobright

Vertica

Integration

- Amazon SQS MuleSource Mule OnDemand Boomi SnapLogic OpSource Connect Cast Iron Microsoft BizTalk Services

Appian Anywhere

HubSpan

- Informatica

On-Demand

SnapLogic SaaS Solution Packs

-FathomDB - Microsoft SDS

Development & Testing

- Keynote Systems Mercury SOASTA SkyTap Aptana LoadStorm Collabnet

Database

- Dynamsoft

- Google BigTable Amazon SimpleDB

Software Services

Billing

Aria Systems eVapt OpSource Redi2 Zuora Financials Concur -Xero -Workday -Beam4d.

Legal DirectLaw -Advologix -Fios -Sertifi

Backup &

JungleDisk-

Zmanda Cloud -

Recovery

Mozy

Backup

OpenRSM -

Syncplicity -

Sales Productivity Xactly LucidEra StreetSmarts Success Metrics

IBM Lotus Live Google Apps Desktoptwo -Parallels -ClusterSeven -

Desktop

Zoho -

Human Resources

Management Taleo -Workday -ICIMS_

Clickability -SpringCM -CrownPoint -

Content

Social Networks Collaboration

Ning -Box.net -DropBox -Zembly.

Amitive -

CRM NetSuite -Parature -Responsys -

Rightnow Salesforce.com -LiveOps MSDynamics -

Oracle On _

Demand

Management NetDocuments -Questys DocLanding

Document

Aconex Xythos Knowledge TreeLive SpringCM



WHAT MOTIVATED CLOUD COMPUTING

Initial motivation:

Web-scale problems

Solutions:

Large data centers

How to access:

Highly-interactive Web applications (thin client)

INITIAL MOTIVATION: WEB-SCALE PROBLEMS

• Characteristics:

- Definitely data-intensive
- May also be processing intensive

• Examples:

- Crawling, indexing, searching, mining the Web
- "Post-genomics" life sciences research
- Other scientific data (physics, astronomers, etc.)
- Sensor networks
- Web 2.0 applications
- SmartThings/home integration

HOW MUCH DATA?

- Google processes over 24 PB a day (24k terabytes)
- CERN's LHC generates 25 PB a year
- "all words ever spoken by human beings" ~ 5 EB (5m terabytes)
- Amount of data that exists in the digital universe 3+ ZB (3b terabytes)
- Brain Research through Advancing Innovative Neurotechnologies (BRAIN) project est: multiple yottabyes (trillions of terabytes)
- LARGE data is the next frontier
- How do we store this amount of data?
 - HDD density
 - SDD density

• How do we filter/access useful information?

APPLICATIONS

• What does cloud computing actually do?

- Consider applications you may currently be running on laptop, desktop, phone, server
- Cloud has them also, or can potentially bring them to you
- Brings applications, views, manipulates, shares data

CLOUDS

 Allow access to applications other than on local computer or internet connected device

But

Only as long as have internet connection

- Instead, company hosts your application Advantages?
 - No more licenses, service packs, etc.
 - Less hardware, etc.
 - Can access anywhere

POTENTIAL PROBLEMS

Internet connection

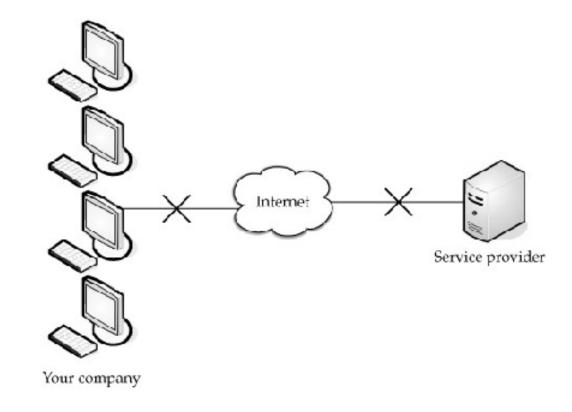
Completely dependent on network

Cloud site failure

Back-end server/network failure may
 Result in inaccessible data

Sensitive information

• How much do you trust the public cloud vendor?



Application integration – (exchange info when local and on cloud)

CLOUD COMPONENTS

•3 components

- Clients
- Datacenter
- Distributed servers

CLOUD COMPONENTS

Clients

Mobile

SmartPhones, Tablets, Service Hubs

• Thin

• no internal hard drives, lets servers do all work, displays info

• Thick

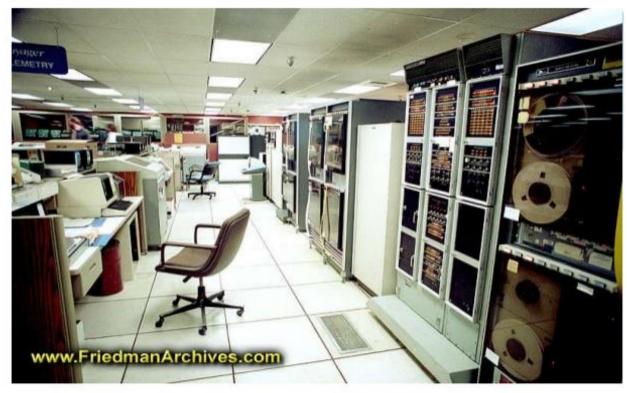
Laptops, desktop computers

• Which is the best?

Thin - lower costs, security, power consumption, easy to replace, less noise

DATA CENTER

 Data Center – facility used to house computer systems and associated components





DISTRIBUTED SERVERS

Servers host the resources needed by cloud users

Compute nodes

- Provides CPU, Memory, Scratch Storage, and Networking resources through virtualized interfaces.
- Hosts guest operating systems (Virtual Machines) using one or more VM hypervisors
- Resource interface depends on the type of cloud (horizontal/vertical cloud)

Storage nodes

- Compute nodes only provide temporary storage space for users/applications
- Storage nodes provide long term data storage solutions
- Can be mapped to specific processes running on compute nodes, users, interface applications, etc.

Administrative nodes

Provides "hidden" back-end services such as resource load balancing, administrative/resource databases, security/firewalls, cloud macromanagement

IMPROVEMENTS SINCE '80S

Disk capacity

- From 10s MB to several TB orders of magnitude
- IBM built 120PB storage array

Bandwidth

• 1-10Mbps to 100s of Gbps



IMPROVEMENTS SINCE '80S

- CPU Improvements:
 - Transistor shrink, increased clock rate, advanced instruction pipeline, cache memory, multicores, faster bus interconnections, lower energy consumption...
- https://www.top500.org/lists/ lists the top 500 highest (known)
 performance super clusters
 - Current leader: Sunway TaihuLight (China)
 - 10.6 million CPU cores
 - 1.3PB RAM
 - Power consumed while fully operational:
 - 15.3mW
 - Linpack: 93PFlop/s (.093EFlop/s) or
 - 93,000,000,000,000 floating point operations per second.
 - By comparison, your laptop: ~125 Gflop/s, Smartphone: ~400 Mflop/s

SOLUTION FOR COMPUTATIONAL GROWTH: LARGE DATA CENTERS

- Web-scale problems? Throw more machines at it!
- Decades ago computing power in mainframes in computer rooms
- Personal computers changed that
- Now, network data centers with centralized computing are back in vogue
- In the future businesses will not need to invest in a data center
- How can we easily access datacenter resources to fit our needs?