

# Cloud Computing

Prof. Tahar Kechadi

UCD School of Computer Science

Insight Centre for Data Analytics

Science East (O'Brien Building), 3<sup>rd</sup> Floor, E3.42



## Important Orientation Information

### ● Applications for Extenuating Circumstances

- These refer to very grave issues that occasionally arise such as
  - Serious illness, hospitalisation, an accident
  - Family bereavement (parent, sibling)
  - Ongoing serious personal or emotional circumstances
- Extenuating Circumstances **do not** cover events which are **foreseen** (e.g. 21<sup>st</sup> party, Debs ball, wedding etc.)

### ● Minor Circumstances (absent for a few days)

- These situations should be handled locally by making direct contact with the lecturer/relevant School.
- Extenuating Circumstances do NOT apply in these cases

### ● Missing a Lecture, Lab session or Tutorial

- ...

### ● Late submission of Coursework

- Where coursework is submitted late due to unanticipated exceptional or extenuating circumstances, students should follow procedures under the **Policy on Late Submission of Coursework**:  
[http://www.ucd.ie/registry/academicsecretariat/docs/latesub\\_po.pdf](http://www.ucd.ie/registry/academicsecretariat/docs/latesub_po.pdf)
- An application for Extenuating Circumstances is not appropriate in this case.

## Plagiarism & UCD Computer Science

- Plagiarism is a serious academic offence
  - [Student Code, section 6.2] or [UCD Registry Plagiarism Policy] or [CS Plagiarism policy and procedures]
- Our staff and demonstrators are proactive in looking for possible plagiarism in all submitted work
- Suspected plagiarism is reported to the CS Plagiarism subcommittee for investigation
  - Usually includes an interview with student(s) involved
  - 1<sup>st</sup> offence: usually 0 or NG in the affected components
  - 2<sup>nd</sup> offence: referred to the University disciplinary committee
- Student who enables plagiarism is equally responsible
 

[http://www.ucd.ie/registry/academicsecretariat/docs/plagiarism\\_po.pdf](http://www.ucd.ie/registry/academicsecretariat/docs/plagiarism_po.pdf)

[http://www.ucd.ie/registry/academicsecretariat/docs/student\\_code.pdf](http://www.ucd.ie/registry/academicsecretariat/docs/student_code.pdf)

<http://libguides.ucd.ie/academicintegrity>

## Organisation

- Lectures
  - Monday: 15:00 – 16:00
  - Wednesday: 15:00 – 16:00
- Practical
  - Thursday: 15:00 – 17:00
  - This slot is available for students who are registered to COMP41110 and COMP30520.
  - Course material: [Moodle](#) (Virtual Learning Environment)
- Project and Practical work
  - Tutorial and practical sessions will be organised to support the material presented in the lectures, to help you in your project work, or to give some more details about certain topics
  - Projects will be posted on Moodle
  - The projects will be submitted by the end of the semester (the last Friday of the term)

## Assessment

- **Continuous Examination**

- Worth 30% of the overall mark
  - 3 MCQs of 30 minutes each
  - MCQ 1 : will take place on October 04<sup>th</sup> (in Practical Session)
  - MCQ 2 : will take place on October 25<sup>th</sup> (in Practical Session)
  - MCQ 3 : will take place on November 22<sup>nd</sup> (in Practical Session)

- **Tutorial/Practical Sessions**

- Worth 30% of the overall mark
  - COMP41110 → Tutorial/Practical submissions + Project
  - COMP30520 → Tutorial/Practical submissions

- **Final Exam**

- Worth 40%

- **Marking Scheme**

- The standard UCD Grading Scheme is used for this module.

## Course Outline

- Motivation: Why Cloud Computing?
- What is cloud computing?
- Virtualisation
- IaaS, PaaS, and SaaS
- Data storage in the Cloud
- Hadoop and MapReduce
- Case Studies: AWS, EC2, S3, etc.
- Migration in the cloud
- Cloud Security
- Advanced Topics

## References

- **Cloud Computing: SaaS, PaaS, IaaS, Virtualisation,...**
  - Kris Jamsa, Jones & Barnett Learning, 1<sup>st</sup> Edition, 2013
- **Cloud Computing Explained: Impl. Handbook for Enterprises**
  - J. Rhoton, 2013 Edition
- **Cloud Computing & SOA Convergence in Your Enterprise: A Step-by-Step Guide,**
  - D.S. Linthicum, Addison Wesley, 2010
- **Cloud Application Architectures: Building Applications & Infrastructure in the Cloud**
  - G. Reese, O'REILLY, 2009
- **Cloud Security & Privacy: An Enterprise Perspective on Risks & Compliance,**
  - T. Mather, S. Kumaraswamy, S. Latif, O'REILLY, 2009
- **Enterprise Cloud Computing: A strategy Guide for Business & Technology Leaders,**
  - A. Mulholland, Meghan-Kiffer Press, 2009

## Why Cloud Computing?

- **Understand the cloud principles**
  - Very different environment and many very interesting algorithms
  - How does it work? What are its strengths & weaknesses?
- **Learn new technologies**
  - MapReduce, Ajax, NoSQL, DHT, etc.
  - How to build a system such as Facebook?
- **Build complex systems**
  - Need scalable, efficient, robust and secure systems

## Why Cloud Computing?

- **Anticipate the technology of the future**
  - Need to understand the current state of the technology
  - Anticipate what's possible in the future
- **Understand the cloud impact on society**
  - Vulnerability, privacy concerns, data survivability, etc.
- **Hands-on on cloud technologies**
  - The best way to understand the technology
  - Gain practical experience and skills in cloud technologies

## Need for Cloud Computing

- **The Big Data Problems**
  - Web-scale problems
  - Very large data centres
  - Different models of computing
- **How big the data is?**
  - Google processes 3.5Billion requests per day. 1.2 Trillion per year.
  - 2.5 EB of data created per day in 2015
  - 5 EB is the number of all words even spoken by human beings
  - CERN's LHC generates 15 PB per year.
- **Process & store exabytes of data**
  - Banks, financial firms, academia, governments, military, etc.

## Big Data Problems

- **Web-scale Problems**

- Data intensive
- Processing intensive

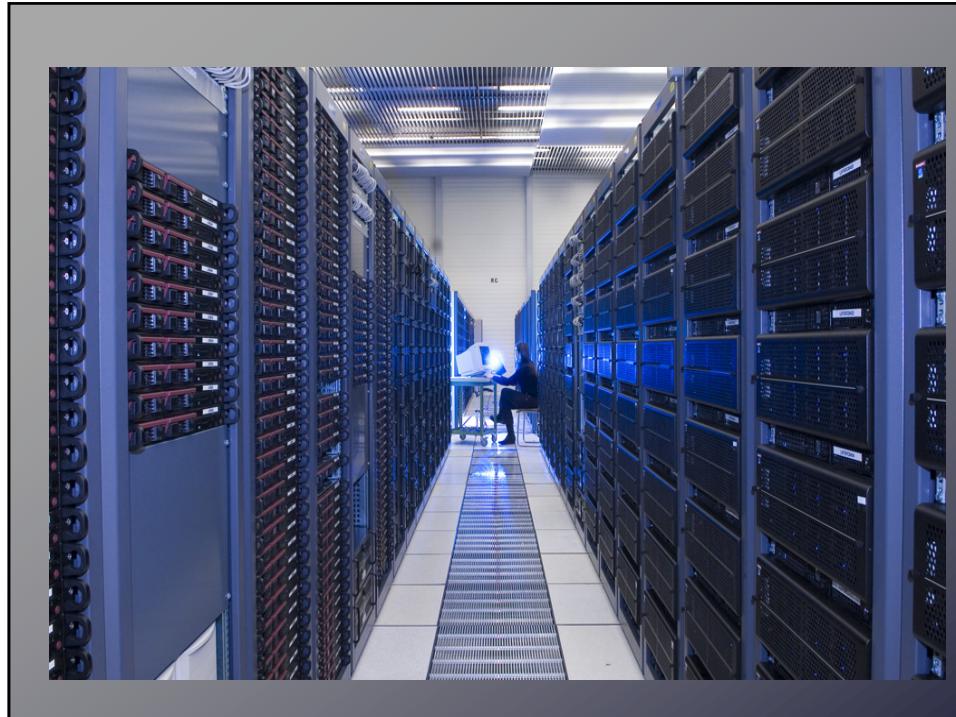
- **Examples**

- Crawling, indexing, searching, mining
- Scientific data (astronomy, particle physics, etc.)
- Sensor networks
- Web 2.0 applications

## Big Data Problems

- **Large Data centres**

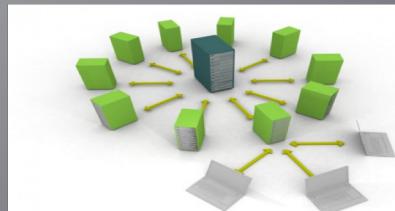
- Redundancy
- Reliability
- Efficiency
- Utilisation
- Transparency
- Management
- Centralised computing resources !!



## Centralised vs Distributed

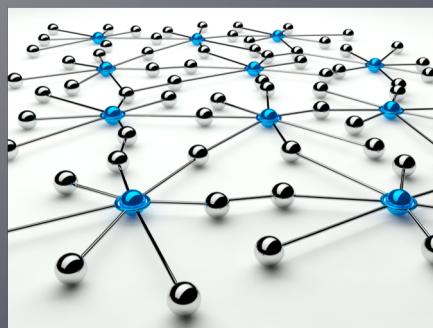
### Centralised system

- Early computing was performed on a single machine (1 cpu, array processors, super-computers); called centralised computing
- Have the overall supervision of computations



### Distributed system

- Collection of independent computers, interconnected via a network, to share resources and capable of collaborating on a task.

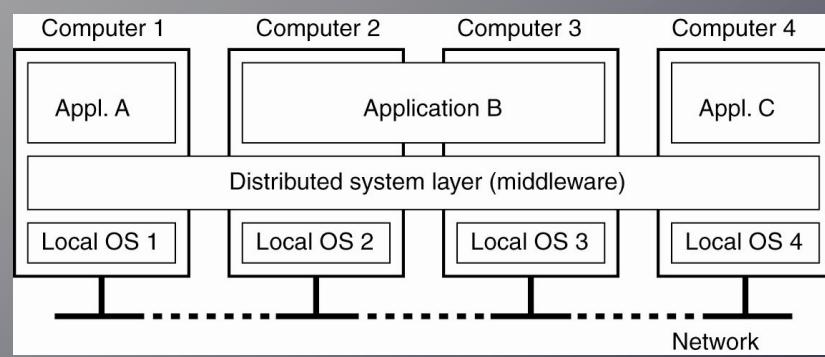


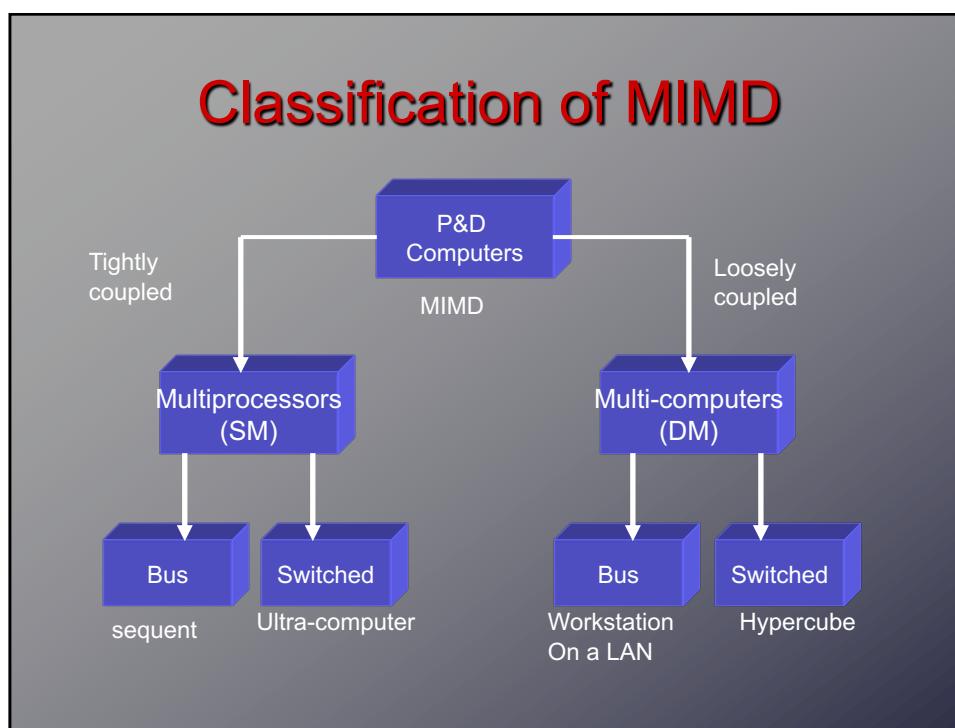
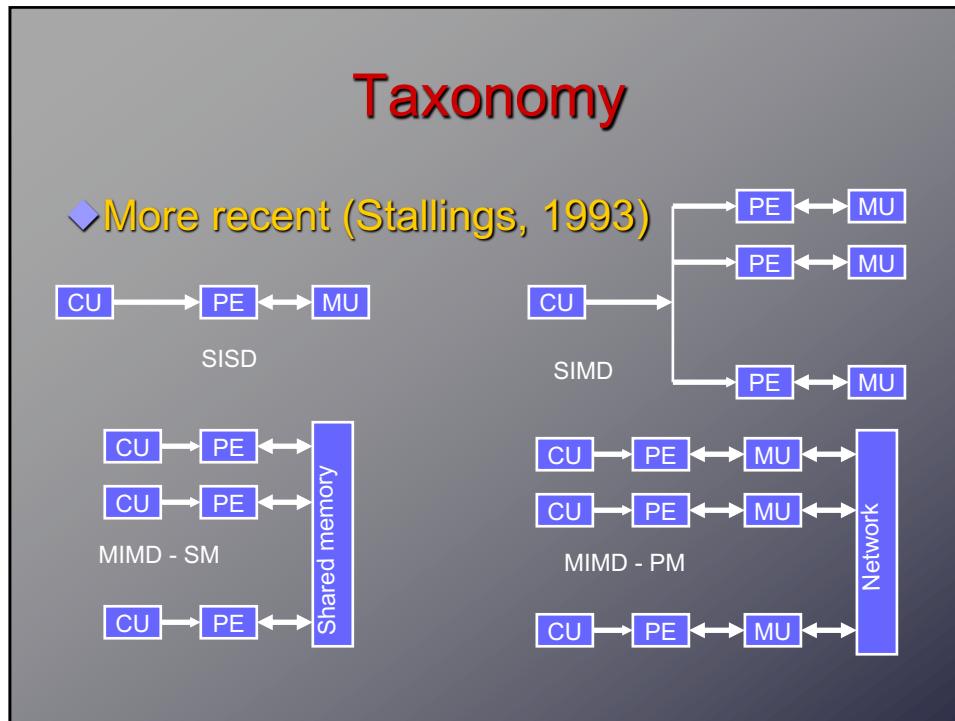
## Centralised vs. Distributed

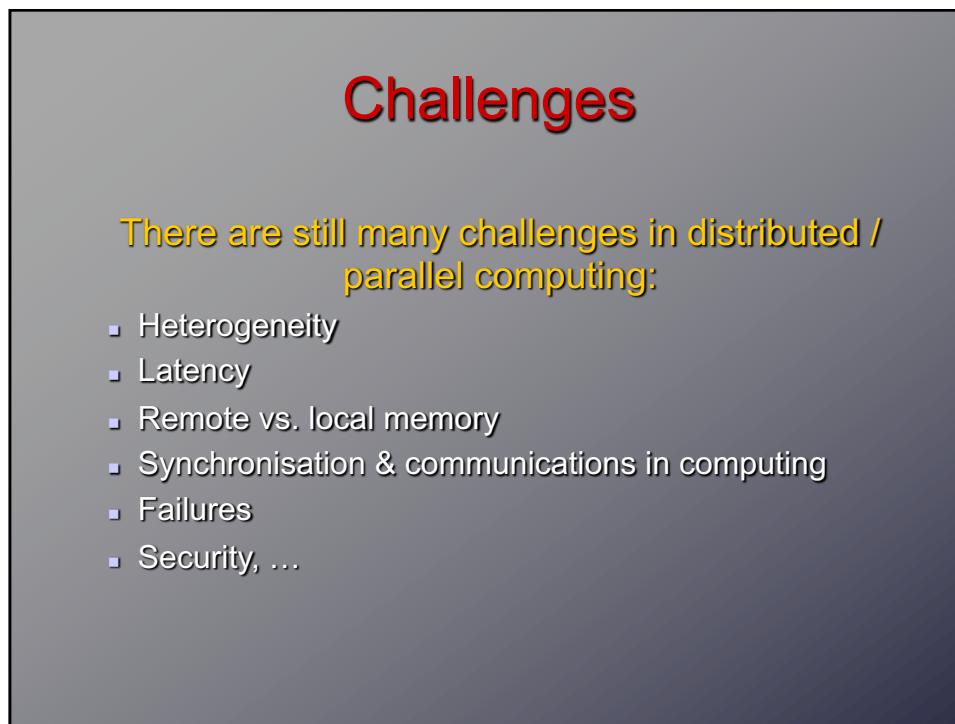
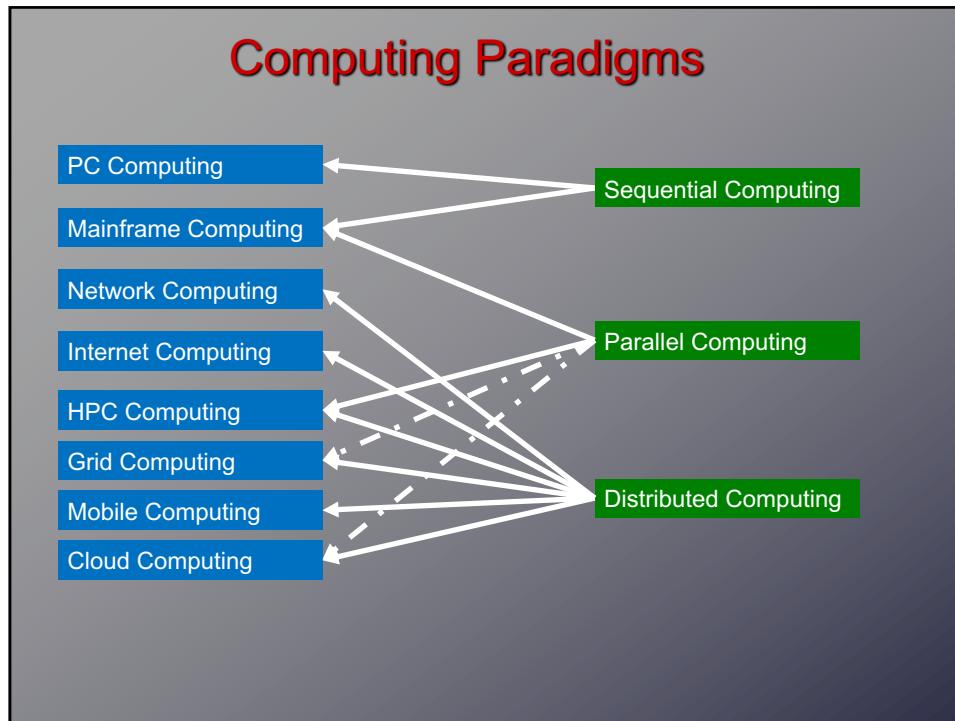
- *Distributed computing* has become increasingly common for several reasons:
  - Bigger challenges in science and industry
  - Performance (Task farming / Parallelisation)
  - Resource sharing
  - Scalability
  - Availability
  - DS are much cheaper than mainframes

## Components

- Middleware ~ Distributed OS
  - Middleware technologies based on threads / sockets, RPC, Java RMI, SOAP/Web Services, etc.







# Cloud Computing

## ● Definition

- Cloud Computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimum management effort or service provider interaction.
- Examples of computing resources include networks, servers, storage, applications, and services
  - National Institute of Standards & Technology (NIST)

# CC Characteristics

## ● On-demand self-service

- Provision of computing resources as needed automatically without requiring human interaction with service provider

## ● Broad network Access

- Resources are available over the network and accessed through standard mechanisms

## ● Elastic resource pooling

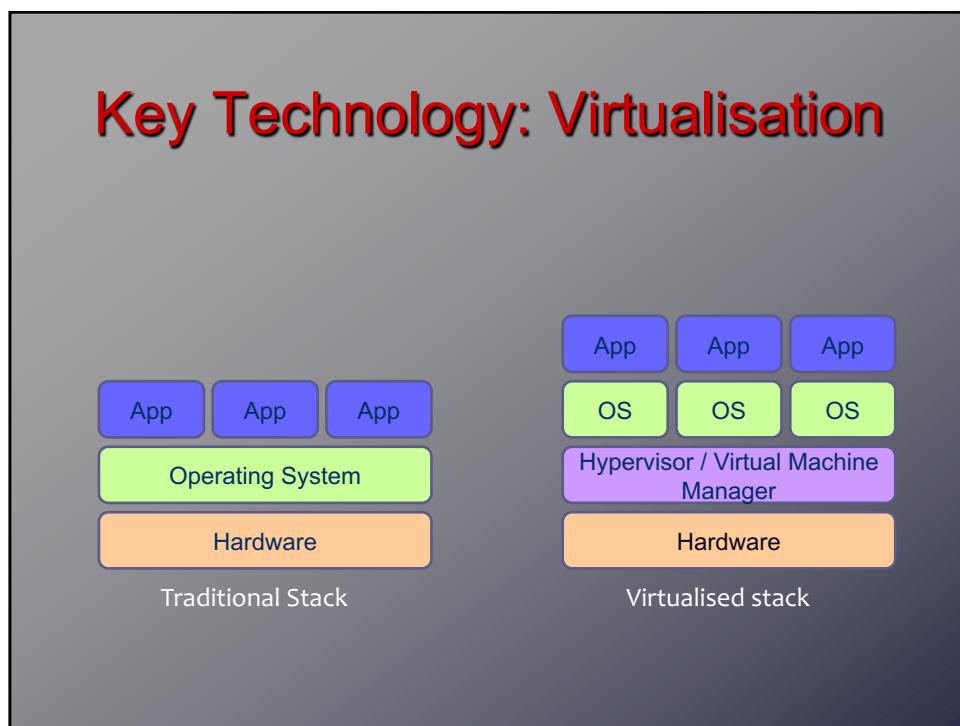
- Computing resources are pooled to serve multiple users using a multitenant model

## ● Rapid elasticity

- Resources can be rapidly and elastically provisioned to the users

## ● Measured service

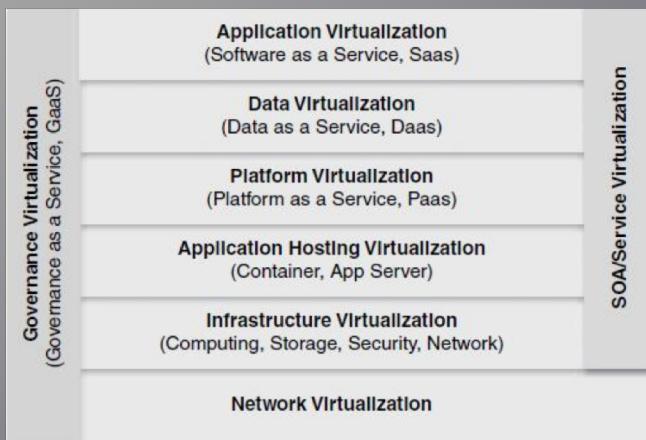
- Cloud controls and optimises resource use automatically



## Why virtualised resources?

- Cost reduction
- Isolation
- Testing / evaluation of OS, kernel, or applications
- Easy duplication
- Running applications not supported by the host
- Green IT!

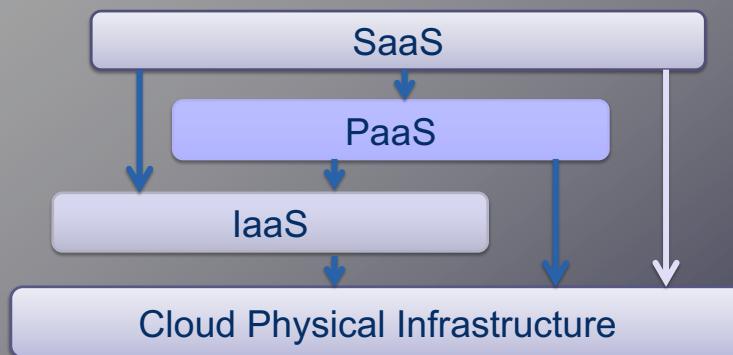
## Cloud Computing Model



## Different delivery Models

- Infrastructure as a Service (IaaS) / Utility computing
  - Why buy machines when you can rent cycles?
  - Examples: Amazon's EC2, GoGrid, Rackspace, etc.
- Platform as a Service (PaaS)
  - Give me nice API and development infrastructure
  - Example: Google App Engine, Cloud Foundry
- Software as a Service (SaaS)
  - Just run it for me!
  - Example: Gmail, Salesforce

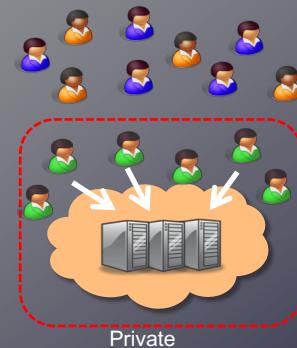
## Different delivery Models: Hierarchy



## Different types of Cloud

### ● Private cloud

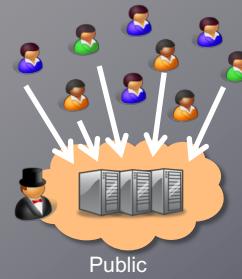
- only users from your organisation can use the cloud



## Different types of Cloud

### ● Public Cloud

- APIs allow any user to launch virtual machines or services remotely on the cloud



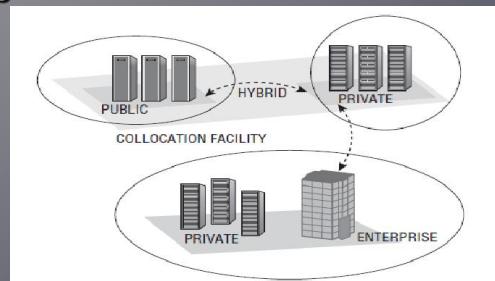
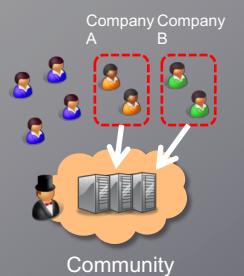
## More types ...

- **Community cloud**

- Only users from a community can use the cloud

- **Hybrid cloud**

- Infrastructure built from mix of public, community and private clouds



## Services!!!

- Storage
- Database
- Information
- Processing
- Applications
- Platforms
- Security
- Management / Governance

# Storage

- **Cloud Storage**

- One of the most common uses of the clouds.
- Amazon's S3 holds more than a trillion objects (about 5 TB in size each), and processes up to 650,000 requests per second at peak times! (Q1, 2012)

- **Storage as a Service**

- This is becoming prevalent in the cloud such as Database as a Service or Data Warehouse as a Service

- **Examples**

- Amazon's SimpleDB, RDS, Redshift...
- Microsoft's SQL Server Data Services (SSDS)
- Oracle's Database 11g, Fusion Middleware...

# Cloud Storage Challenges

- **Security**
- **Reliability**
- **Outages**
- **Theft**

