



UNIVERSITY OF
CAMBRIDGE

Cloud Computing

Introduction

Eva Kalyvianaki
ek264@cam.ac.uk

My Background and Contact Details

- Dr **Ev**angelia Kalyvianaki, Associate Professor
- @CST since Oct. 2017
- PhD from the CL, Cambridge University (srg, netos group)
- I like building systems, working in Cloud computing, distributed systems management, data centre and AI scheduling

- Office in FN15
- ek264@cam.ac.uk
- <http://www.cst.cam.ac.uk/~ek264/>

Course Logistics

- <http://www.cl.cam.ac.uk/teaching/2526/CloudComp/>
- Moodle page: tutorials, coursework information and submission areas
- Research papers (will be given per lecture)

- Coursework project in groups
- Project assessment via report and code testing
- 26 students: 6 groups of 3 students and 2 groups of 4 students.
- Details of the coursework in Tutorials
- Teaching Assistants:
 - Andreas Grammenos, ag926@cam.ac.uk
 - Simon Virag, siv24@cam.ac.uk

email Eva and TAs by Wednesday 28/1 at 10am

your groups: names, emails, and one public key per group

Course Contents

1. Introduction to Cloud Computing, 23/1
2. MapReduce Batch Processing, 26/1
3. MapReduce in Heterogeneous Environments, 28/1
4. Large-Scale Resource Management I, 30/1
5. Tutorial I, 2/2
- 6. Tutorial II (coursework I 4/2 – 25/2)**
7. Tutorial III, 6/2
8. Large-Scale Resource Management II, 9/2
9. Data Center Networking (guest lecture Dr Paolo Costa, MSR), 11/1
10. Real-Time Data Stream Processing, 13/2
11. Cloud Storage, 16/2
12. Virtualization I 18/2
13. Virtualization II 20/2
- 14. Tutorial IV (coursework II 23/2 – 9/3)**

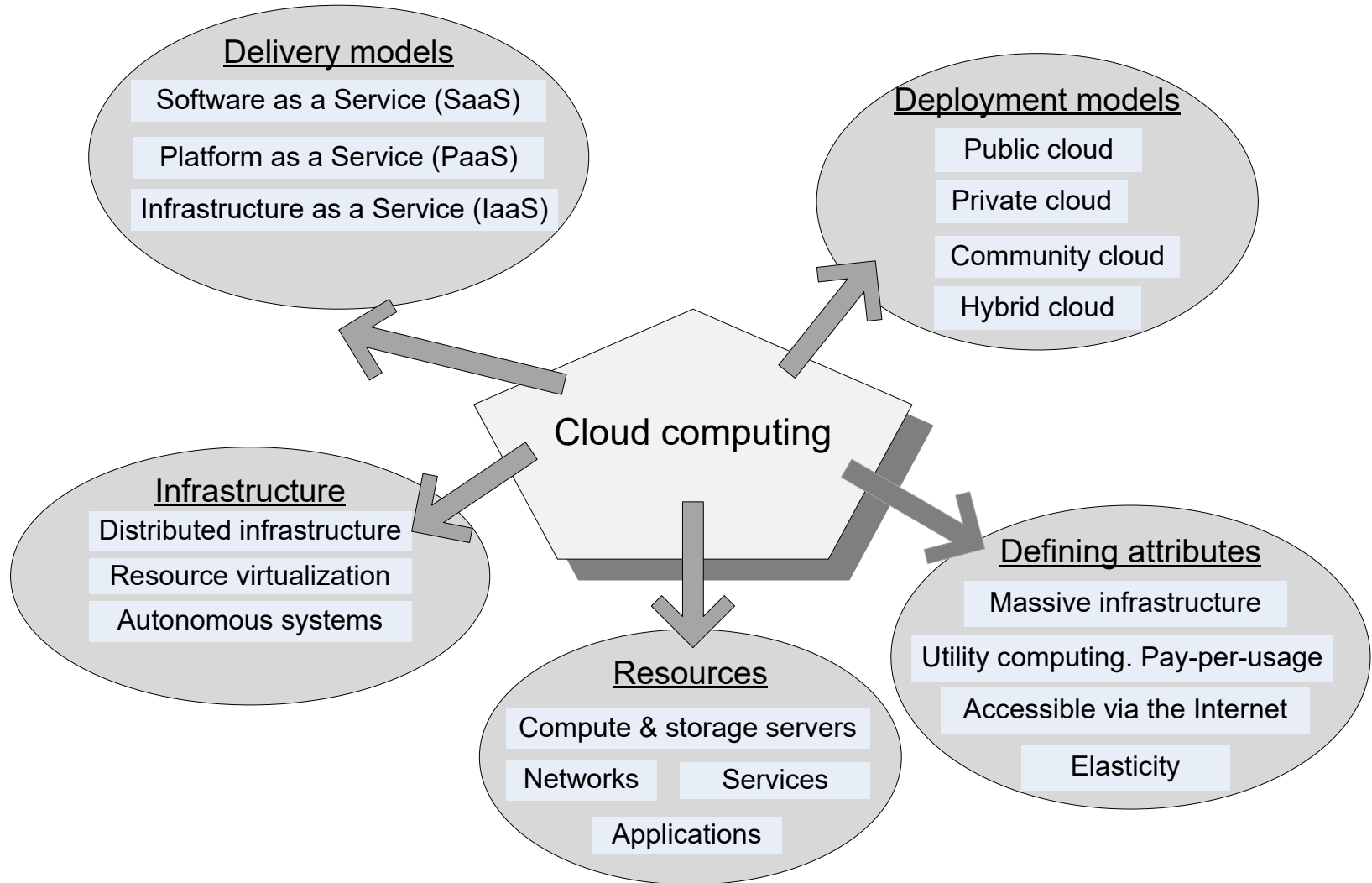
Lecture Contents

- What is Cloud Computing?
- Early models of Cloud Computing.
- Delivery models and services.
- Ethical issues in Cloud Computing.
- Cloud vulnerabilities.
- Parallel Computing.
- Distributed Systems.

What is Cloud Computing?

- What do you think?
- **"Cloud computing** is an information technology (IT) paradigm that enables ubiquitous access to shared pools of configurable system resources and higher-level services that can be rapidly provisioned with minimal management effort, often over the Internet. Cloud computing relies on sharing of resources to achieve coherence and economies of scale, similar to a public utility." https://en.wikipedia.org/wiki/Cloud_computing
- "Simply put, cloud computing is the delivery of computing services – servers, storage, databases, networking, software, analytics and more – over the Internet ("the cloud"). Companies offering these computing services are called cloud providers and typically charge for cloud computing services based on usage, similar to how you're billed for gas or electricity at home." <https://azure.microsoft.com/en-gb/overview/what-is-cloud-computing/>

Cloud Computing Models, Resources, Attributes



Early Models of Cloud Computing

- Basic reasoning: information and data processing can be done more efficiently on large farms of computing and storage systems accessible via the Internet.

Two early models:

1. Grid computing – National Labs in 90s for scientific computing.

- *"Grid computing is the collection of computer resources from multiple locations to reach a common goal. The grid can be thought of as a distributed system with non-interactive workloads that involve a large number of files."* from Wikipedia

2. Utility computing –IT companies in 2005 for enterprise computing.

- *"Utility computing is a service provisioning model in which a service provider makes computing resources and infrastructure management available to the customer as needed, and charges them for specific usage rather than a flat rate."*from Wikipedia

Cloud computing - Characteristics

"Cloud Computing offers on-demand, scalable and elastic computing (and storage services). The resources used for these services can be metered and users are charged only for the resources used." from the Book

Shared Resources and Resource Management:

1. Cloud uses a shared pool of resources
2. Uses Internet techn. to offer **scalable** and **elastic** services.
3. **Elastic computing** → **dynamically** and **on-demand** resources.
4. Resources are metered and users are charged accordingly.
5. More cost-effective due to **resource-multiplexing**
 - Lower costs for the cloud service provider are passed to the cloud users.

Cloud computing (cont' d)

Data Storage:

6. Data is stored:

- in the “cloud”, in certain cases closer to the site where it is used.
- appears to the users as if stored in a location-independent manner.

7. The data storage strategy can increase reliability, as well as security, and can lower communication costs.

Management:

8. Maintenance and security are operated by service providers.

9. Service providers operate more efficiently due to specialisation and centralisation.

Cloud Computing Advantages

1. Resources (CPU, storage, network) are **shared**.
2. When multiple applications share a system, their peak demands for resources are not synchronised thus, **multiplexing** leads to a higher resource utilization.
3. Resources can be **aggregated** to support data-intensive applications.
4. Data sharing facilitates **collaborative** activities. Many applications require multiple types of analysis of shared data sets and multiple decisions carried out by groups scattered around the globe.

Cloud Computing Advantages

- 5. Eliminates the **initial investment costs** for a private computing infrastructure and the maintenance and operation costs.
- 6. **Cost reduction:** concentration of resources creates the opportunity to pay as you go for computing.
- 7. **Elasticity:** the ability to accommodate workloads with very large peak-to-average ratios.
- 8. **User convenience:** virtualization allows users to operate in familiar environments rather than in idiosyncratic ones.

Types of clouds

1. **Public Cloud** - the infrastructure is made available to the general public or a large industry group and is owned by the organization selling cloud services.
2. **Private Cloud** – the infrastructure is operated solely for an organization.
1. **Hybrid Cloud** - composition of two or more Clouds (public, private, or community) as unique entities but bound by a standardised technology that enables data and application portability.
2. **Other types: e.g., Federated Cloud, Edge** - the infrastructure is shared by several organizations and supports a community that has shared concerns.

Why cloud computing is successful when others failed?

- It is in a better position to exploit recent advances in software, networking, storage, and processor technologies promoted by the same companies who provide Cloud services.

- Economical reasons: It is used for enterprise computing; its adoption by industrial organizations, financial institutions, government, and so on has a huge impact on the economy.

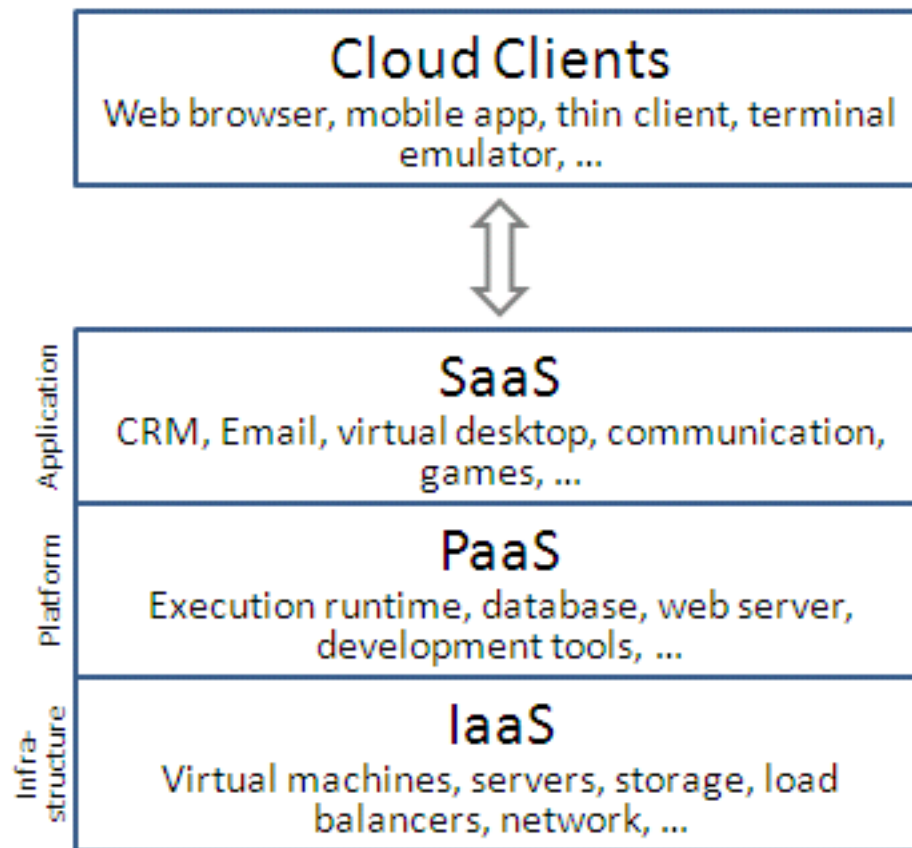
- Infrastructures Management reasons:
 - A single Cloud of mostly homogeneous (now more heterogeneous) set of hardware and software resources.
 - Single administrative domain (AD). Security, resource management, fault-tolerance, and quality of service are less challenging than in a heterogeneous environment with resources in multiple ADs.

Challenges (ongoing) for cloud computing

1. **Availability of service:** when the service provider cannot deliver?
2. Diversity of services, data organization, user interfaces available at different service providers **limit mobility** → once a customer is hooked to one provider it is hard to move to another.
3. **Data transfer bottleneck** → applications are data-intensive.
4. **Data confidentiality and auditability, security** major concern for sensitive applications, e.g., healthcare applications.
5. **Performance unpredictability**, because of resource sharing. How to provide QoS guarantees? How to support elasticity?
6. **Resource management:** A big challenge to manage different workloads running on large data centers.

Cloud Delivery Models

1. **Software as a Service (SaaS)** (high level)
2. **Platform as a Service (PaaS)**
3. **Infrastructure as a Service (IaaS)** (low level)



The Three delivery models of Cloud Computing

Cloud Service Models

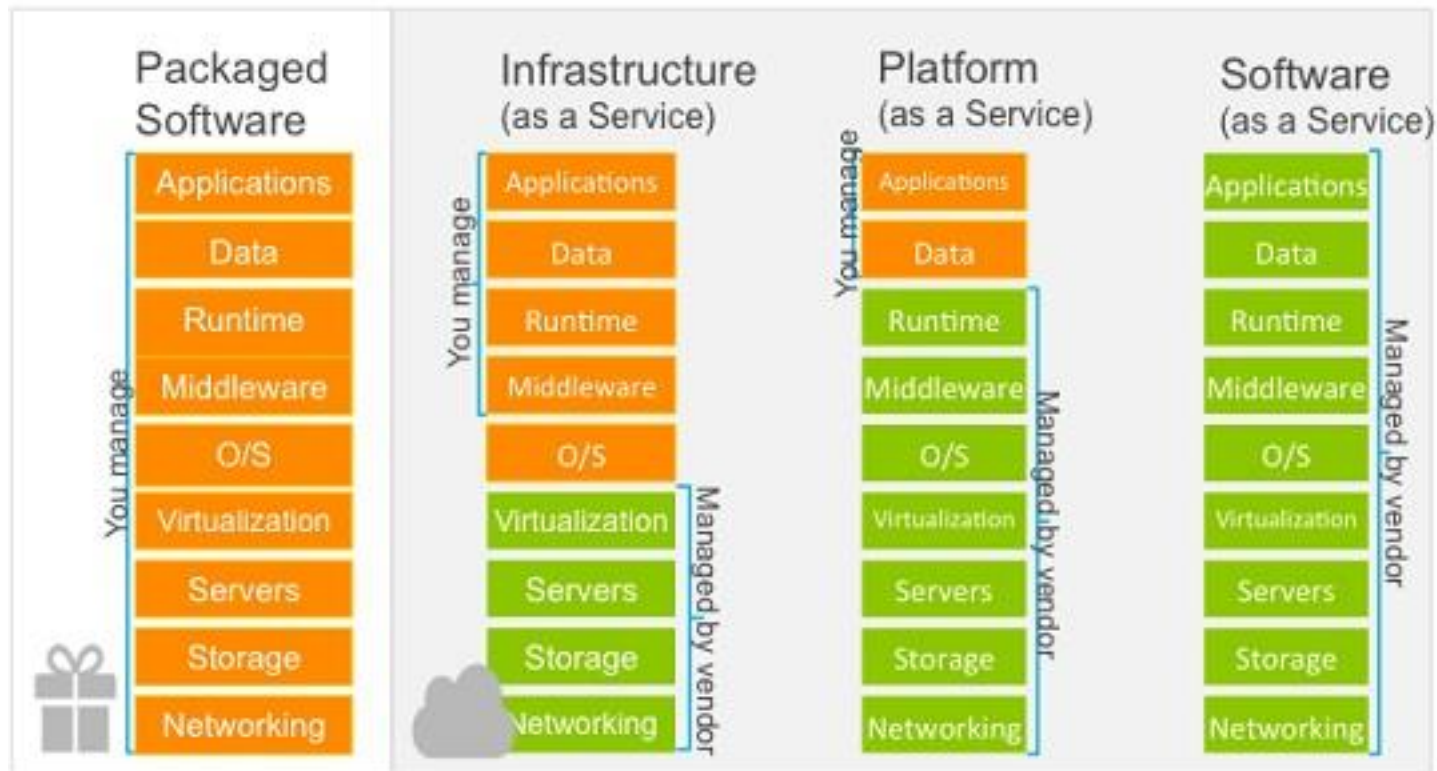


Figure 1.

Source: Microsoft Azure

Infrastructure-as-a-Service (IaaS)

- Infrastructure is resources: CPU, VMs, storage, etc
- Users can deploy and run arbitrary software: OS or apps.
- No control of the underlying Cloud infrastructure.
- Users have control over OSes, storage, deployed applications.
- Services offered by this delivery model include: server hosting, storage, computing hardware, operating systems, virtual instances, load balancing, Internet access, and bandwidth provisioning.
- Example: Amazon EC2

Platform-as-a-Service (PaaS)

- Users can deploy consumer-created or acquired applications using PLs and tools by the service provider.
- Users control the deployed applications and, possibly, application hosting environment configurations.
- Users do not manage or control the underlying Cloud infrastructure.
- Not very useful when:
 - The application must be portable.
 - Proprietary programming languages are used.
 - The hardware and software must be customized.
- Examples: Google App Engine, Windows Azure

Software-as-a-Service (SaaS)

- Applications are supplied by the service provider.
- Users do not control the underlying Cloud infrastructure or individual application.
- Services include:
 - Enterprise services: workflow management, communications, digital signature, customer relationship management (CRM), desktop software, financial management, geo-spatial, ...
- Not good for real-time applications or when data is not allowed to be hosted externally.
- Examples: Gmail, Salesforce

Cloud activities

- Service management and provisioning including:
 - Virtualization.
 - Service provisioning.
 - Call center.
 - Operations management.
 - Systems management.
 - QoS management.
 - Billing and accounting, asset management.
 - SLA management.
 - Technical support and backups.

Cloud activities (cont' d)

- Security management including:
 - ID and authentication.
 - Certification and accreditation.
 - Intrusion prevention.
 - Intrusion detection.
 - Virus protection.
 - Cryptography.
 - Physical security, incident response.
 - Access control, audit and trails, and firewalls.

Cloud activities (cont' d)

- Customer services such as:
 - Customer assistance and on-line help.
 - Subscriptions.
 - Business intelligence.
 - Reporting.
 - Customer preferences.
 - Personalization.
- Integration services including:
 - Data management.
 - Development.

Other topics

Ethical issues:

- The control is relinquished to third party services.
- Data is stored on multiple sites run by several organizations.
- Multiple services interoperate across the network.
- Unauthorised access, data corruption, ...

De-perimeterisation:

- Systems can span multiple organisations and cross security borders.
- The complex structure of Cloud services can make it difficult to determine who is responsible in case something undesirable happens.

Privacy issues:

Cloud service providers operate on petabytes of sensitive personal information stored in data centers around the world: privacy issues addressed by these companies and the countries where the data centers are located.

Cloud Vulnerabilities

- Clouds are affected by malicious attacks and failures of the infrastructure, e.g., power failures.
- Such events can affect the Internet domain name servers and prevent access to a Cloud or can directly affect the Clouds:
 - in 2004 an attack at Akamai caused a domain name outage and a major blackout that affected Google, Yahoo, and other sites.
 - in 2009, Google was the target of a denial of service attack which took down Google News and Gmail for several days;
 - in 2012 lightning caused a prolonged down time at Amazon.

Back to Basics -- Parallel Computing

- *"**Parallel computing** is a form of computation in which many calculations are carried out simultaneously, operating on the principles that large problems can often be divided into smaller ones, which are then solved concurrently (in parallel)."*Wikipedia
- Hardware and software systems allow us to:
 - Solve problems demanding resources not available on a single system.
 - **Reduce the time required to obtain a solution.**

Parallel Computing – Amdahl's Law

The speedup S measures the effectiveness of parallelization

$$S(N) = T(1) / T(N)$$

$T(1) \rightarrow$ the execution time of the sequential computation.

$T(N) \rightarrow$ the execution time when N parallel computations are executed

Amdahl's Law: if α is the fraction of running time a sequential program spends on non-parallelisable segments of the computation then:

$$S \approx 1 / \alpha$$

This is a theoretical upper bound on the best speedup we can get from parallelising a certain program.

Back to Basics -- Distributed systems

- Collection of autonomous computers, connected through a network and distribution software (often) called middleware which enables computers to coordinate their activities and to share system resources for a common goal.

- Characteristics:
 1. Users perceive the system as a single, integrated computing facility.
 2. Components are autonomous.
 3. Scheduling and security policies are implemented by each system.
 4. Multiple points of control and multiple points of failure.
 5. Resources may not be accessible at all times.
 6. Can be scaled by adding additional resources.
 7. Can maintain availability even at low levels of hw, sw, net reliability.

Summary

- What is Cloud Computing?
- Early models of Cloud Computing.
- Delivery models and services.
- Ethical issues in Cloud Computing.
- Cloud vulnerabilities.
- Parallel Computing and Distributed Systems (brief)

email Eva and TAs by Wednesday 28/1 at 10am
your groups: names, emails and one public key per group

→ email from Eva to follow