Cloud Services: Drivers and Challenges

Documents, Services and Data on the Web

Sandra Sampaio, Norman Paton
School of Computer Science
University of Manchester

Location in Workshop

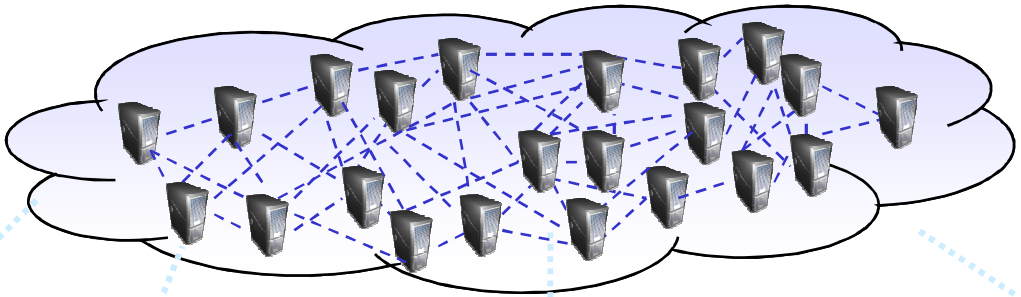
- Cloud computing: key features.
- Cloud delivery models.
- Issues and challenges.
- Cloud deployment models.
- Cloud economics.

A Definition

- National Institute of Standards and Technology (NIST):
  - http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf
  - "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."

What is Cloud Computing?

- It is an emerging IT infrastructure paradigm where data and services reside in massively scalable data centres and can be ubiquitously accessed from any connected devices over the internet.



Businesses, from startups to enterprises

4+ billion phones *by 2010 [Source: Nokia]*

Web 2.0-enabled PCs, TVs, etc.

Slide adapted from the presentation: IBM Google Announcement on Internet-Scale Computing ("Cloud Computing Model")

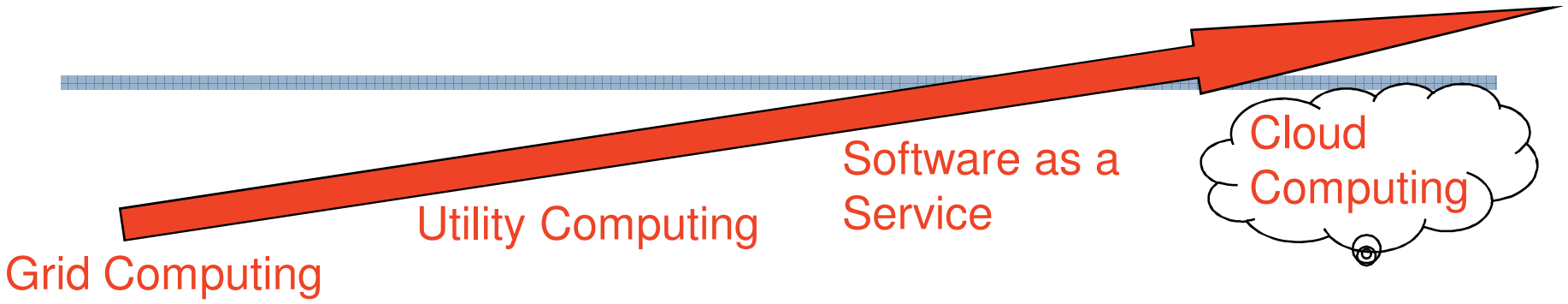
What is a Cloud Service?

- It is any IT resource that is made remotely accessible via a cloud. A cloud refers to a distinct IT environment that is designed for the purpose of remotely provisioning scalable and measured IT resources.

- Within the context of cloud computing, a cloud service is a broad concept, ranging from a simple web service to a remote access point for administrative tools or larger environments and other IT resources.

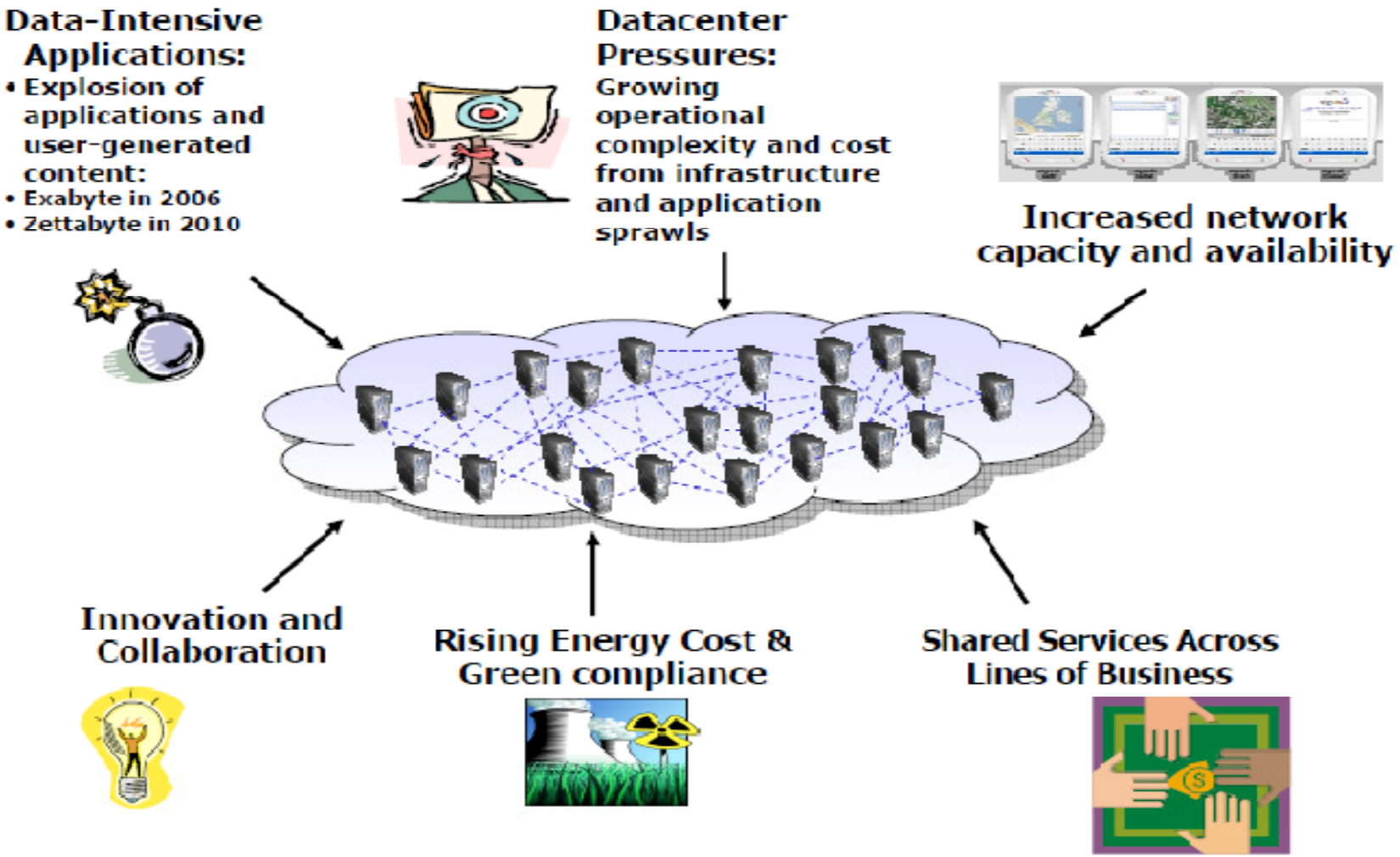
Slide adapted from the presentation: IBM Google Announcement on Internet-Scale Computing ("Cloud Computing Model")

Infrastructure Evolution Towards Ubiquitous Computing



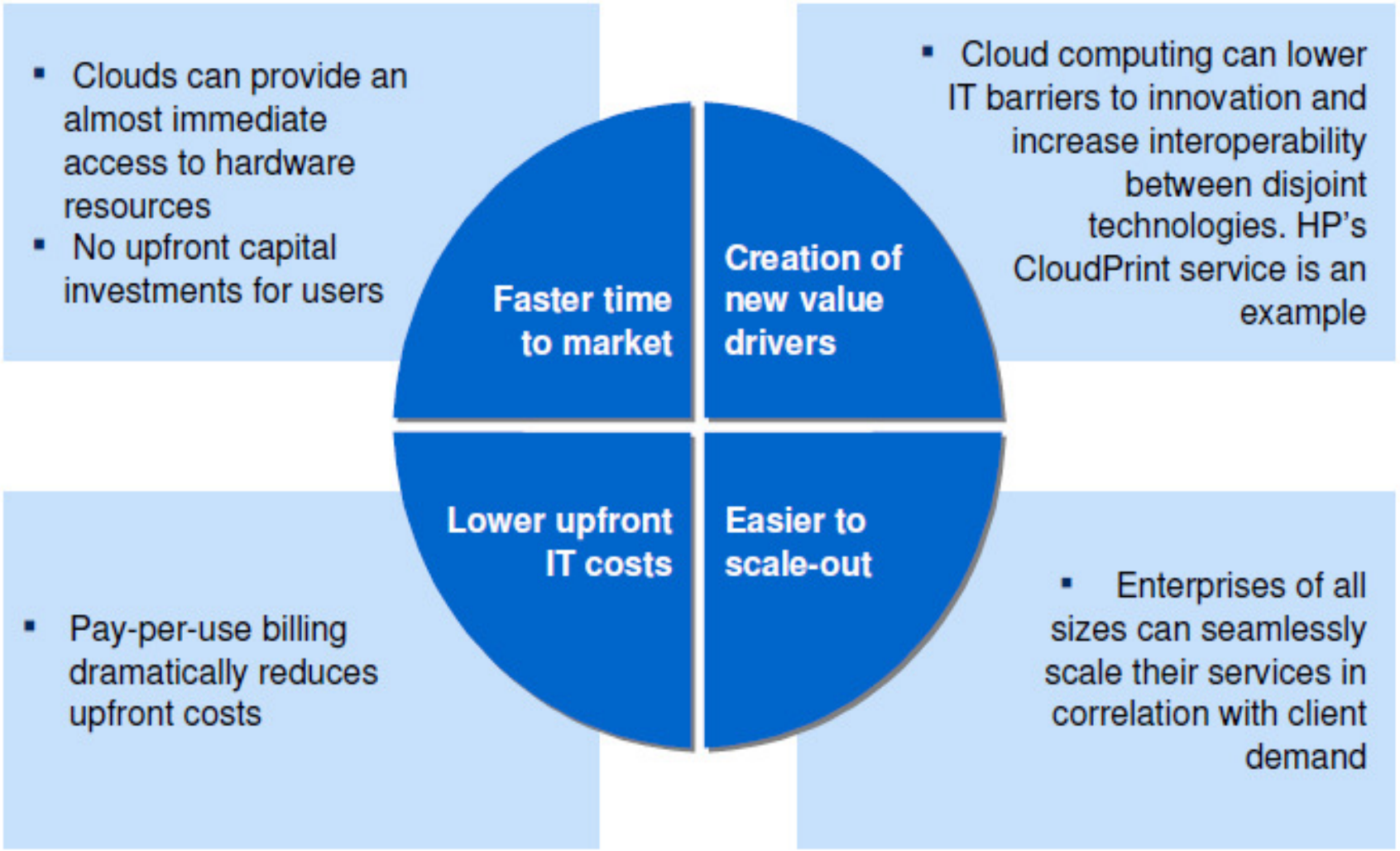
Grid Computing	Utility Computing	Software as a Service	Cloud Computing
The application of several computers to a single problem at the same time - usually to a scientific problem that requires a great number of computer processing cycles or access to large amounts of data	Packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility	Model of software deployment where an application is licensed for use as a service provided to customers on demand.	Style of computing in which resources are provided “as a service” over the Web to users who need not have knowledge of, expertise in, or control over the technology infrastructure

Forces Driving Cloud Computing



Slide adapted from the presentation: IBM Google Announcement on Internet-Scale Computing ("Cloud Computing Model")

Cloud Computing Value Drivers

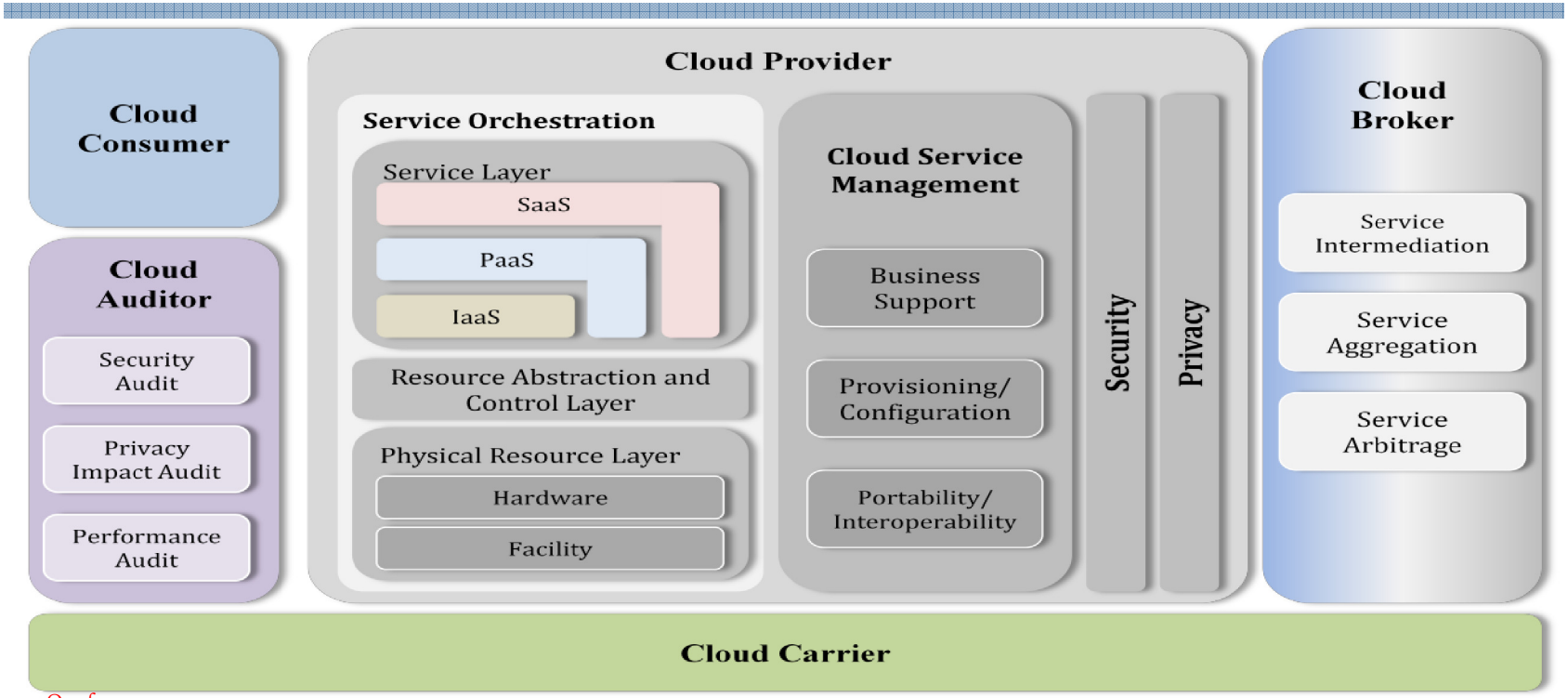


Slide extracted from the presentation "Clearing the Air on Cloud Computing" by McKinsey & Company, March 2009

Characteristics of Cloud Computing

- **Virtual** – Physical location and underlying infrastructure details are transparent to users.
- **Scalable** – Able to break complex workloads into pieces to be served across an incrementally expandable infrastructure.
- **Elastic** – allows the use of as many resources as necessary to optimally respond to the cost and timing constraints of an application.
- **Efficient** – Service-oriented-architecture for dynamic provisioning of shared computational resources reduces systems administration overheads.
- **Flexible** – Can serve a variety of workload types – both consumer and commercial.

The Conceptual Reference Model



Our focus

Actor	Definition
Cloud Consumer	A person or organization that maintains a business relationship with, and uses service from, *Cloud Providers*.
Cloud Provider	A person, organization, or entity responsible for making a service available to interested parties.
Cloud Auditor	A party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation.
Cloud Broker	An entity that manages the use, performance and delivery of cloud services, and negotiates relationships between *Cloud Providers* and *Cloud Consumers*.
Cloud Carrier	An intermediary that provides connectivity and transport of cloud services from *Cloud Providers* to *Cloud Consumers*.

Figures source: http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=909505

Location in Workshop

- Cloud computing: key features.
- Cloud delivery models.
- Issues and challenges.
- Cloud deployment models.
- Cloud economics.

Cloud Delivery Models

- A cloud delivery model represents a specific, pre-packaged combination of IT resources offered by a cloud provider. The following cloud delivery models have become widely established:
  - Infrastructure-as-a-Service (IaaS)
  - Platform-as-a-Service (PaaS)
  - Software-as-a-Service (SaaS)

Infrastructure-as-a-Service (IaaS)

- Consists of infrastructure-centric IT resources that can be accessed and managed via cloud service-based interfaces and tools. Examples include hardware, network, connectivity, operating systems, drivers, etc.

- They are typically virtualized and packaged into bundles that simplify up-front runtime scaling and customization.

Platform-as-a-Service (PaaS)

- It is a pre-defined “ready-to-use” environment typically comprising already deployed and configured IT resources.
- Relies on the usage of a ready-made environment that establishes a set of pre-packaged products and tools used to support the entire delivery lifecycle of custom applications.
- Gives the capability to deploy consumer-created or acquired applications using programming languages and tools supported by the provider.
- In other words, it represents the software building blocks for developing application software in the cloud.

Software-as-a-Service (SaaS)

- Represents the positioning of software programs as shared cloud services, made available as a “product” or generic utility.
- Typically used to make a reusable cloud service widely available to a range of cloud consumers. A marketplace exists around SaaS products that can be leased and used for different purposes.

Examples of Available Services

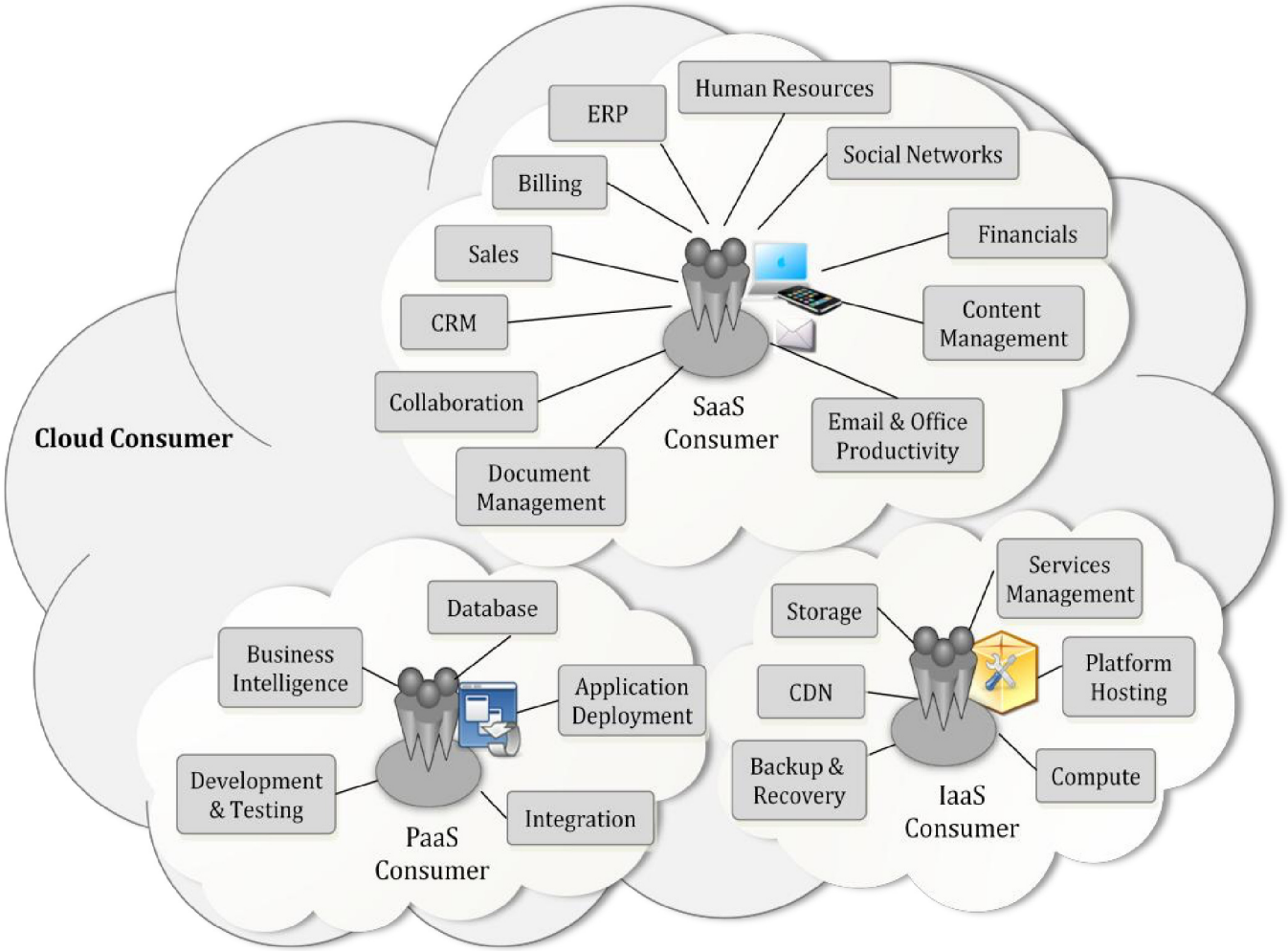
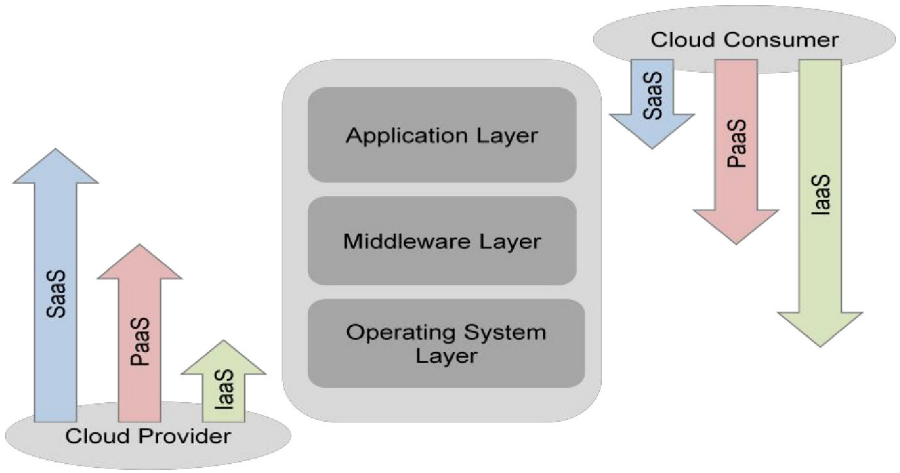


Figure source: http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=909505

Scope of Controls

- Applications at the Application Layer are used by SaaS consumers, installed/managed/maintained by SaaS providers, or PaaS consumers and IaaS consumers.
- Software building blocks at the Middleware Layer are used by PaaS consumers, installed/management/maintained by PaaS providers or IaaS consumers, and hidden from SaaS consumers.
- Operating system and drivers at the Operating System Layer are hidden from both SaaS and PaaS consumers, and are controlled by both the IaaS provider and consumers.



Activity 1: Cloud Delivery Models

- Look at *Amazon Web Services* (AWS) as an example of a cloud service provider, and classify the services they provide as infrastructure, platform or software as-as-service.
- Please start with the following (http://aws.amazon.com/products/), specifically:
  - EC2.
  - S3.
  - EBS.
  - RDS.
  - EMR.
  - CloudSearch.
- Now look at *Microsoft Azure*: https://azure.microsoft.com/en-us/solutions/.
  - For each of the services you looked at in AWS, look for a corresponding service in Azure.

Location in Workshop

- Cloud computing: key features.
- Cloud delivery models.
- Issues and challenges.
- Cloud deployment models.
- Cloud economics.

Cloud Consumers' Concerns

- **Portability:** ability to communicate/transfer data or applications between multiple clouds at low cost and minimal disruption.
- **Interoperability:** ability to use data and services across multiple cloud providers with a unified management interface.
- **Security:** in the cloud, it ranges from physical security to application security.
  - Cloud-based systems still need to address security requirements such as authentication, authorization, availability, confidentiality, identity management, integrity, audit, security monitoring, incident response, and security policy management.
  - Security is a shared responsibility between cloud providers and consumers, as they collaboratively design, build, deploy, and operate cloud- based systems.

Location in Workshop

- Cloud computing: key features.
- Cloud delivery models.
- Issues and challenges.
- Cloud deployment models.
- Cloud economics.

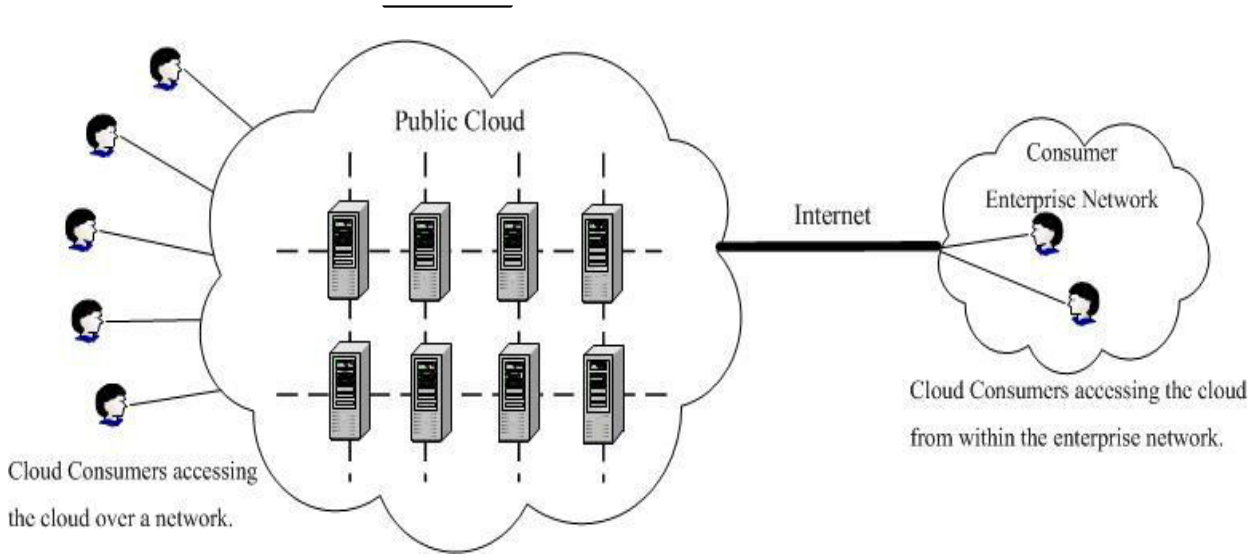
Cloud Deployment Models

- A cloud infrastructure may be operated in one of the following deployment models:
  - Public Cloud
  - Private Cloud
  - Community Cloud
  - Hybrid Cloud

- The differences are based on **how exclusive** the computing resources are made to a cloud consumer.

Public Cloud

- A **publicly accessible** cloud environment owned by a third-party cloud provider.
- The cloud provider is responsible for the creation and on-going maintenance of the cloud and its IT resources.
- Cloud services are sold to a diverse pool of cloud consumers over a public network.

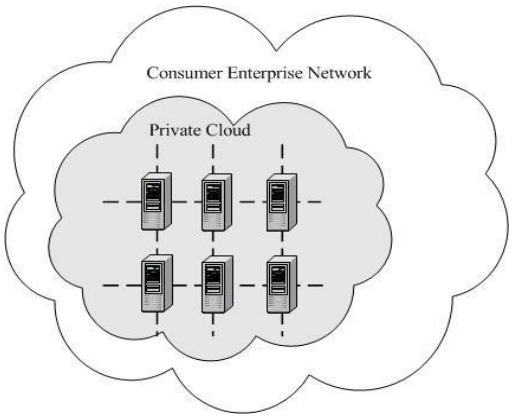


Public Cloud

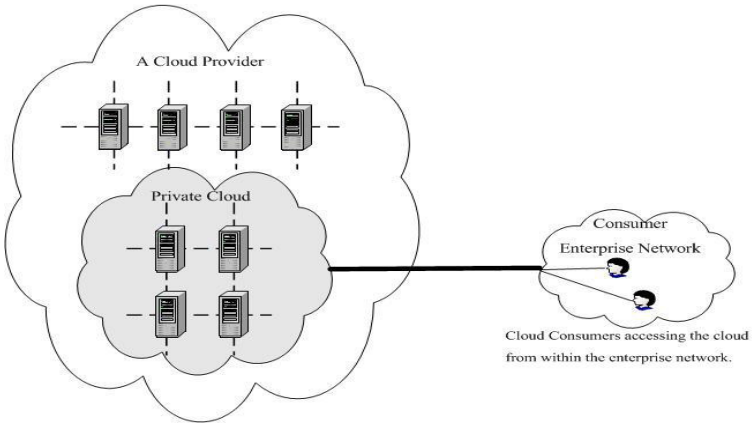
Figure source: http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=909505

Private Cloud

- Owned by a **single cloud consumer organization**, which has the exclusive access to and usage of the infrastructure and computation resources.
- It may be managed either by the cloud consumer organization or by a third party, and may be hosted on the organization's premises (i.e. on-site private clouds) or outsourced to a hosting company (i.e. outsourced private clouds).



On-site Private Cloud

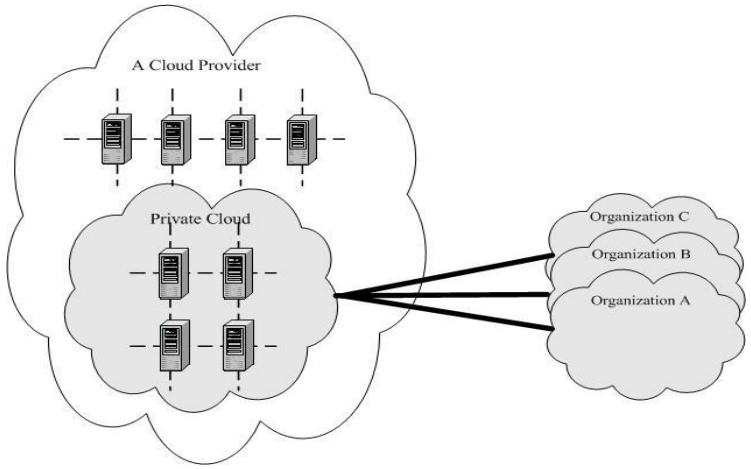


Out-sourced Private Cloud

Figure source: http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=909505

Community Cloud

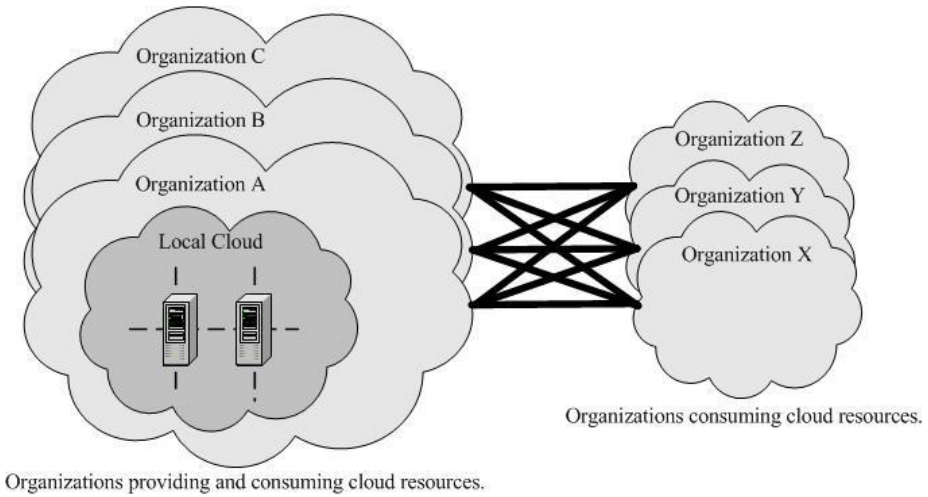
- Similar to a public cloud, except that its access is limited to a **specific community of cloud consumers**.
- The cloud may be jointly owned by the community members or by a third-party provider that provisions the cloud. And so, it may be implemented on customer premise or outsourced to a hosting company.
- The member cloud community typically shares the responsibility for defining and evolving the community cloud.



Outsourced Community Cloud comprised by a number of participating organizations.

Figure source: http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=909505

Community Cloud



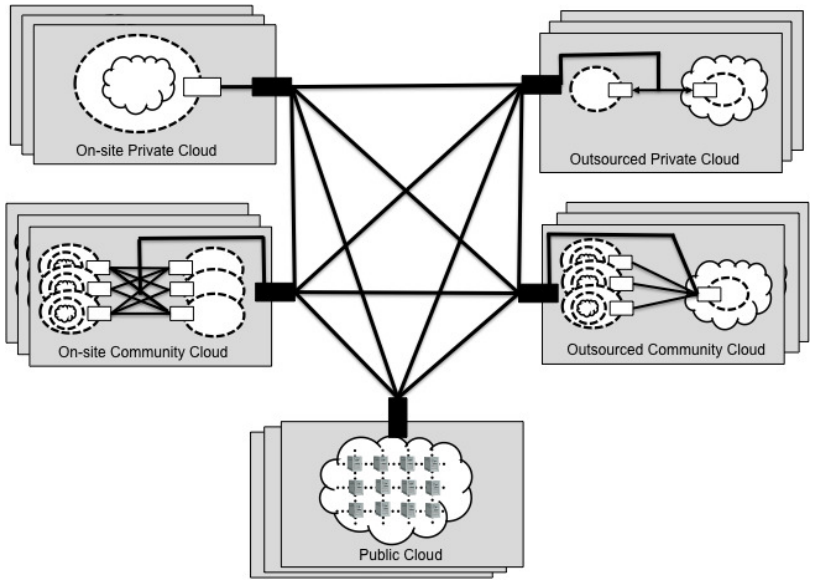
On-site Community Cloud comprised of a number of participating organizations.

A cloud consumer can access the local cloud resources, and also the resources of other participating organizations through the connections between the associated organizations.

Figure source: http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=909505

Hybrid Cloud

- A cloud environment comprised of **two or more different cloud deployment models** that remain as distinct entities, but are bound together by standardized or proprietary technology that enables data and application portability.
- This deployment model is useful, for example, in situations where a cloud consumer chooses to deploy cloud services processing sensitive data to a private cloud and other, less sensitive cloud services, to a public cloud.



Hybrid Cloud built with a set of clouds in all deployment model variants.

.

Location in Workshop

- Cloud computing: key features.
- Cloud delivery models.
- Issues and challenges.
- Cloud deployment models.
- Cloud economics.

Cloud Computing Economics

- Cloud computing users can avoid capital expenditure on hardware, software and services, rather paying a provider only for what they use.
- Consumption is billed on a utility (e.g. resources consumed, like electricity) or subscription (e.g. time based, like a newspaper) basis with little or no upfront cost.
- Other benefits of this time sharing style approach are low barriers to entry, shared infrastructure and costs, low management overhead and immediate access to a broad range of applications. Users can generally terminate the contract at any time (thereby avoiding return on investment risk and uncertainty) and the services are often covered by service level agreements with financial penalties.

Activity 2: Naïve IaaS Economics

- Look up the specification of the least expensive Dell Server at www.dell.co.uk.
- Compare the cost of this with a reserved instance in AWS:
  - What are the pros and cons?
  - How should the capital cost be related to the rental cost?

What are the costs?

- The decision as to whether to use a cloud or local computational resources may well not be straightforward. Financial decisions must take into account direct costs:
  - Capital Expenditure: the cost of procurement.
  - Operational Expenditure: the costs of licensing, maintenance, etc.
- The benefits of the cloud approach for infrastructure and platform services may also be quite difficult to quantify, with some of the measures not necessarily easy to relate directly to financial value.
- The following slides look at some metrics that can be applied to inform the answering of the question "should we use a cloud"?

Total Cost of Ownership (TCO)

- Definition [Williams 2011]:
  - "Total cost of ownership (TCO) is simply the sum total of all associated costs relating to the purchase, ownership, usage, and maintenance of a particular product".
- *Non-computing example*: The TCO of a car will include its purchase cost (perhaps accounted for as depreciation), along with costs for insurance, fuel, servicing, maintenance, etc.
- *TCO for a computing purchase* include: initial cost of hardware and software (accounted for as depreciation), hardware and software maintenance, hosting costs (machine room, energy costs, networking costs), systems and software administration and support.
- *TCO for a complex infrastructure* may be difficult to compute: the lifespan of a machine room (building, cooling, electricity supply) may be longer than that of the networking infrastructure, which may be longer than that of the backup infrastructure, which may be longer than that of the servers, which may be longer than that of the storage.

Activity 3: Total Cost of Ownership (Procure)

- A server has a purchase price of £100,000, and has an operational lifetime of 3 years. There is an annual hardware maintenance charge of 10% of the purchase price each year. The systems and database software costs £10,000, and has a support/maintenance cost of 20% of the purchase price each year. The energy costs are £1000 per year. The server needs 10% of a systems administrator and 10% of a database administrator, who each cost £80,000 per year. What is the 3 year TCO?

Item	Annual Charge	3 Year Charge
Hardware		
Hardware Maintenance		
Software		
Software Maintenance		
Energy		
Staff		
Total		

Activity 4: Total Cost of Ownership (Cloud)

- Activity 3 provided calculations for the total cost of ownership for a database server hosted within an organisation. Let's assume that the previous procurement assumed a 20% growth in storage needs each year during the 3 year period, and over-provisioned. We assume an initial storage cost of £20,000, which grows elastically. Let's also assume that the database software costs grow with scale, from £10,000, and that we need only 10% of one administrator. Complete the table.

Item	Annual Charge	3 Year Charge
Storage		
Software Service		
Staff		
Total		

Comment on Activity 4

- Please note that these are *made up figures*! The actual costs for any given scenario need to be looked at in detail. What are the expectations in practice?
  - In a stable and well established setting, the TCO for a procurement may be expected to be less than for a cloud solution.
  - In an unpredictable setting with the potential for rapid growth, the TCO for a procurement may involve over provisioning (or the risk of churn). In addition, for a growing organisation large up-front capital costs may be difficult to meet.
- Thus the real value of moving to cloud may not be in direct financial savings, but in the ability to provision flexibly and dynamically, both:
  - To increase the available resources when there is growing demand, or there are short term peaks.
  - To decommission resources flexibly when there is reduced demand ("deprovision").

Additional Not-Quite-Financial Metrics

- Knowing the TCO is not in itself especially helpful for indicating the cost-effectiveness of a purchase. Additional metrics that may cast light on the effectiveness of an investment may include:
  - *Availability*: The fraction of a time when a service is available (typically measured as a number of "9"s: 0.9999 = 4 nines.
    - 4 nines is 52 minutes of downtime in a year.
    - The direct financial cost of non-availability in lost revenue, wasted time, may or may not be straightforward to compute.
  - *Time to market*: The period that it takes to make a new service available. If hardware/software procurement is on the critical path to service deployment, then this could have a significant impact on the ability of an organisation to bring products to market ahead of competitors.
  - *Churn rate*: The number of customers lost within a given period. The ability to reduce the churn rate may depend on ensuring that servers are not overloaded for a prolonged period.

Return on Investment (ROI)

- The ROI is a straightforward measure of the benefit that might be expected to result from a capital expenditure.
- ROI = (Gain from investment – Cost of investment) / Cost of Investment
- Example:
  - Assume that we make an investment of £100,000 that yields an income of £40,000 per year for 3 years (note it may be easier to estimate the cost than the gain).
  - ROI = (120,000 – 100,000) / 100,000
  - ROI = 0.2 (or 20%)
- Note that this model takes no account of the fact that the money might be expected to yield a return from some other source (e.g. by investing it in another project, or by putting it in the bank).

Net Present Value (NPV)

- The NPV provides a mechanism for taking into account the expected level of return from an investment before it is considered worthwhile.
- If an organisation expects to make an investment only when it yields a return of (say) 10% a year, then the NPV can be calculated for a period of years by discounting the future return down to its present value.
- The value of the return in year *t* is given by the following formula, where R_t is the income in the year and *i* is the discount rate:
  - $R_t / (1 + i)^t$
- For the example from the previous slide:
  - NPV = - Cost of Investment +

    (Income in year 1 / 1.1) +

    (Income in year 2 / 1.1^2) +

    (Income in year 3 / 1.1^3)
  - NPV = -100,000 + (40,000 / 1.1) + (40,000 / 1.21) + (40,000 / 1.331)
  - NPV = -100,000 + 99,474 = -525.

Conclusions

- Cloud services exist that provide a wide range of functionality, from the leasing of infrastructure, through pay-as-you-go access to technical platforms, to complete application services.
- Clouds allow outsourcing of different aspects of the computational requirements of an organisation, along with the staff required to administer local solutions.
- The question as to whether or not an organisation should outsource to a cloud is complex; cost clearly plays a part in any answer, but perhaps the greatest potential for benefits comes from elasticity.
- In the next workshop, we will obtain some experience with AWS, for data storage and for running map / reduce jobs.

Suggested Reading Material

- The Economics of Cloud Computing, Bill Williams, Cisco Press, 2011.
- Cloud Computing, Concepts, Technology & Architecture, T. Erl, Z. Mahmood, and R. Puttini, Prentice Hall, 2013.