Topics in Cloud Computing

Salvatore Filippone, PhD

School of Aerospace, Transport and Manufacturing
salvatore.filippone@cranfield.ac.uk

Cloud Security

Areas of cloud computing that are uniquely troublesome:

- Auditing
- Data integrity
- e-Discovery for legal compliance
- Privacy
- Recovery
- Regulatory compliance

Security, Privacy, Trust

Trust in the Cloud Do components/guests trust each other

Trust on the Cloud Why do we trust provider with our data/comp?

Huge barrier to many use cases

- Hacking?
- Leak of info to co-hosted foreign software?
- What if cloud provider is malicious?

Two distinct problems here

- Can I trust remote system?
- If not, can I still use cloud computing?
- Solution- Encryption?

Trust by Cloud

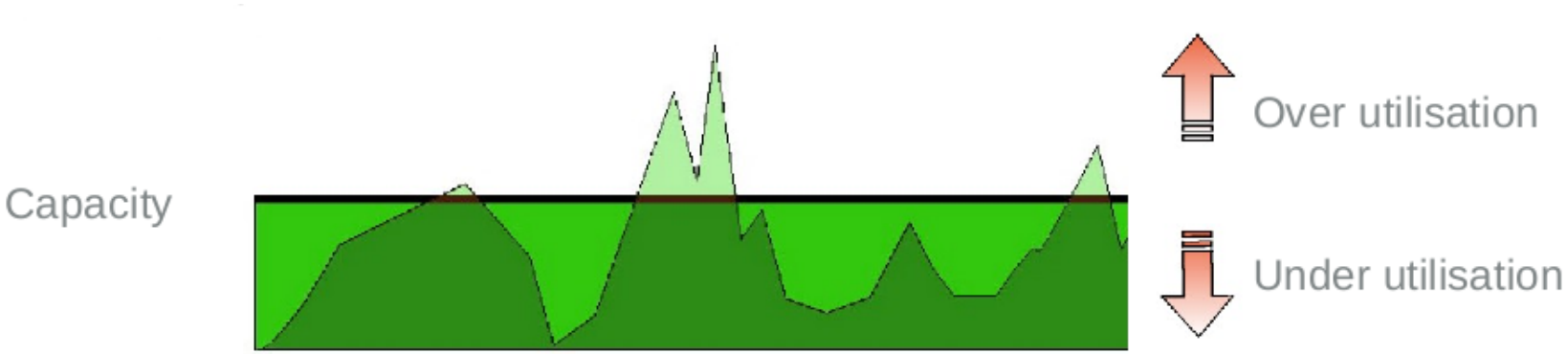
- IP is anonymous
- No accountability:
  - How do you know who a client is
  - Where a client is (GeoMapping)
  - Is done as part of targeted advertising
- Sybil attacks, DDoS, Botnet etc etc etc

Under and Over Provisioning

- When specifying an IT system, a key characteristic is the expected workload for that system
- This influences both the physical hardware that must be purchased and the choice of software architecture used to implement the system
- An ability to react to changing workloads is a key design characteristic, but can be difficult to implement in practice
- Equally important is the up-front (capital) costs of implementing a system

Under and Over Provisioning

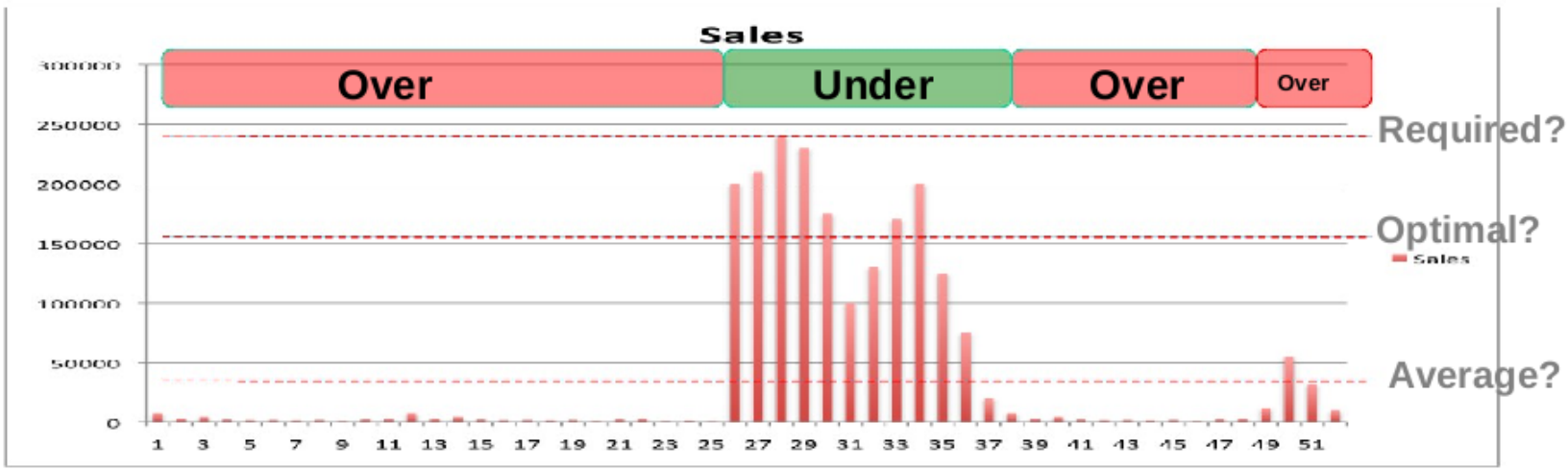
- Traditionally, a computer runs one operating system
- May be underused (over provisioned) — waste of computational resource
- Or overused (under provisioned) — performance bottlenecks



Under and Over Provisioning

Example — on-line ticketing application

- Demands for tickets will depend upon the events being offered
- Other factors may be important, such as the time of year
- Key design issue is that the demand cannot be predicted
- Resource allocation is problematic — may lead to poor user experience



Economic implications

Over Provisioning

Historically, IT departments have had to build systems that were big enough to handle peak demand. But that's not how you provision in the cloud. In the cloud, you provision based around average consumption, and then allow the cloud vendor to deal with the occasional spike.

Economic implications

Data Centre

Data centre requires significant up-front capital expenditure

- Physical location
- Equipment
- Installation
- How is peak capacity calculated?
- Need to predict computational requirements
- What is expected growth in computational requirements?

Economic implications

Cloud Computing

Presents a different economic model to existing systems

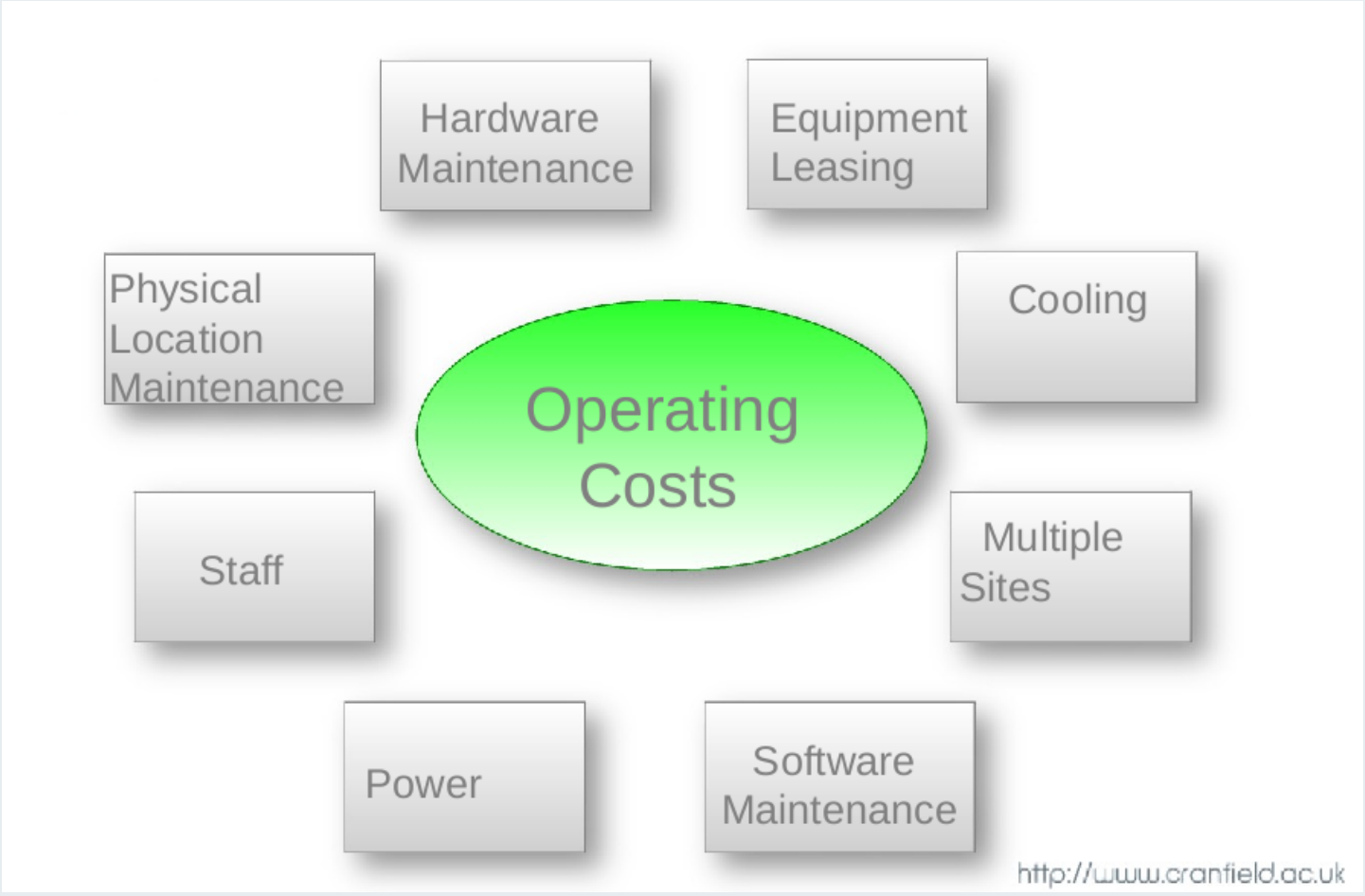
- Reduction in initial capital expenditure
- Transfer of costs to operating expenditure

Some important financial questions

- Is your current infrastructure under/over provisioned?
- Is your workload forecast uncertain?
- Are your computational requirements ad-hoc?
- Over a finite time-scale, can you justify the usage costs of the Cloud?

Economic Implications

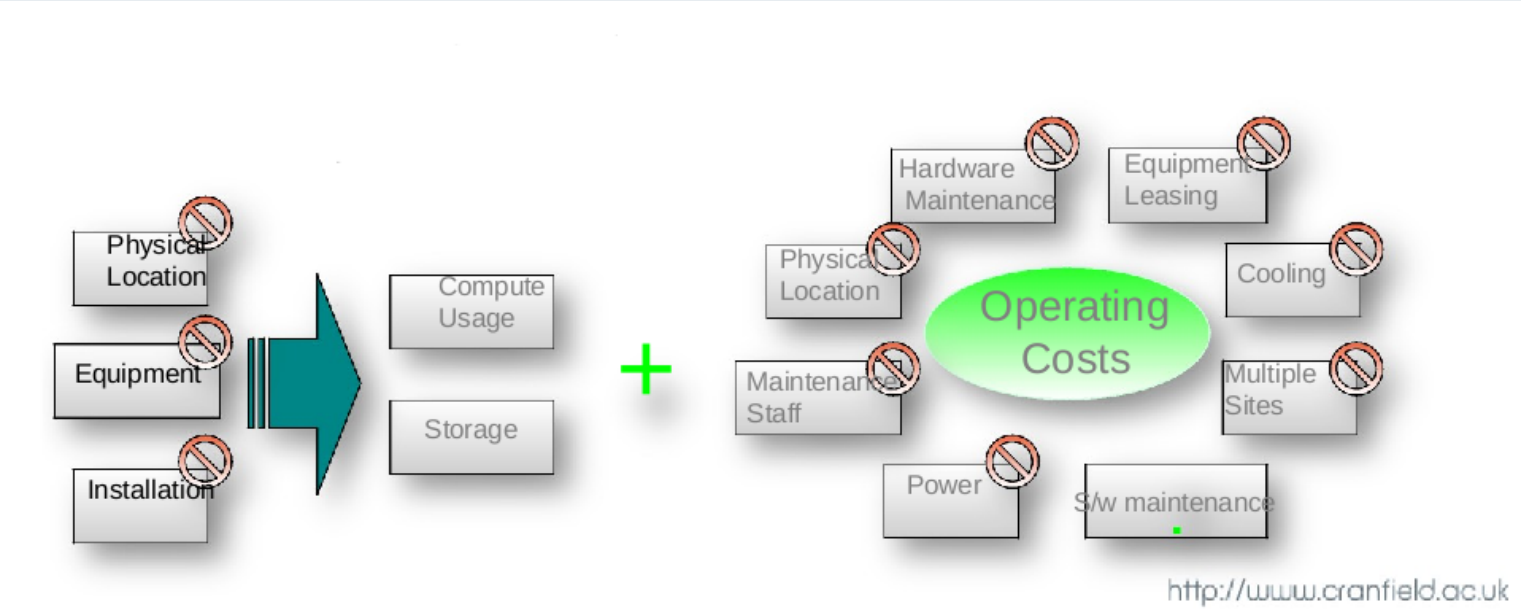
Data Centre



Economic Implications

Cloud Computing

- Convert capital costs to operating costs ⇒ Some operating costs can be reduced or removed
- Pay for usage ⇒ Computational, Storage, Communications



Legal Issues

Legal Issues are an important facet of risk management

- Liabilities must be understood and mitigated
- In Cloud Computing, these can be due to national (or international) laws
- Determines specific security requirements
- These must take priority over functional requirements
- A key legal issue that has not been fully addressed is that of data sovereignty

Data Security and Sovereignty

- Fundamental problem of Cloud Computing
- Can we trust the vendor to process the data securely
- Where is the data stored and/or processed?
- Is the data backed up?
- Specific issues regarding export of personal data (ie UK Data Protection Act)

Risk Management

- Dependent upon the cloud service model chosen, type of cloud
- Evaluate your risks– perform the following analysis:
  1. Which resources (data, services, or applications) to move to cloud.
  2. Sensitivity of the resource to risk. - Loss of privacy, unauthorized access by others, loss of data, interruptions in availability.
  3. Risk associated with type of cloud - public, private, hybrid, and shared community types– Where data and functionality will be maintained.
  4. Which model? IaaS, SaaS, and PaaS – security levels, boundaries
  5. For a provider- how data is transferred, where it is stored, and how to move data both in and out of the cloud.

Risk Management

Organizations using cloud computing must consider:

- Sensitive data - Sensitive data processed outside the enterprise brings with it an inherent level of risk, because outsourced services bypass the controls used in-house.
- Regulatory compliance - Customers are ultimately responsible for the security
- and integrity of their own data, even when it is held by a cloud computing provider.
- Data location - Do you have regulatory or contractual obligations to store data in specific jurisdictions — Can your data by stored outside the U.S., for example?
- Data segregation - It is critical that your data is segregated from other customers.
- Recovery - Cloud providers must have plans to restore your data and service in case of a disaster.
- Investigative support - Investigating inappropriate or illegal activity may be impossible in cloud computing.
- Long-term viability - What is the long term viability of your cloud computing vendor?

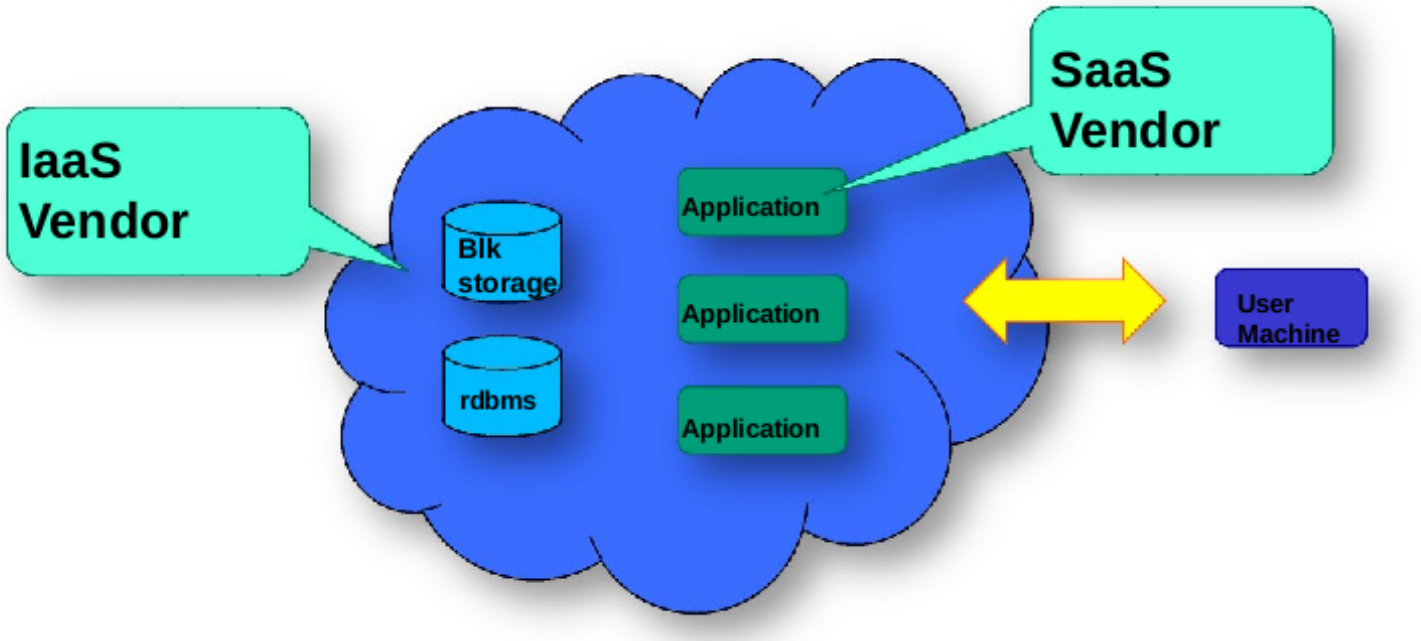
Risk Management

- Have "golden" system image references
- Take a system image off-line and analyse the image for vulnerabilities or compromise
- The compromised image is a primary forensics tool.
- Many cloud providers offer a snapshot feature that can create a copy of the client's entire environment
- Includes machine images, applications and data, network interfaces, firewalls, and switch access.
- If you feel that a system has been compromised, you can replace that image with a known good version

Data Security and Sovereignty

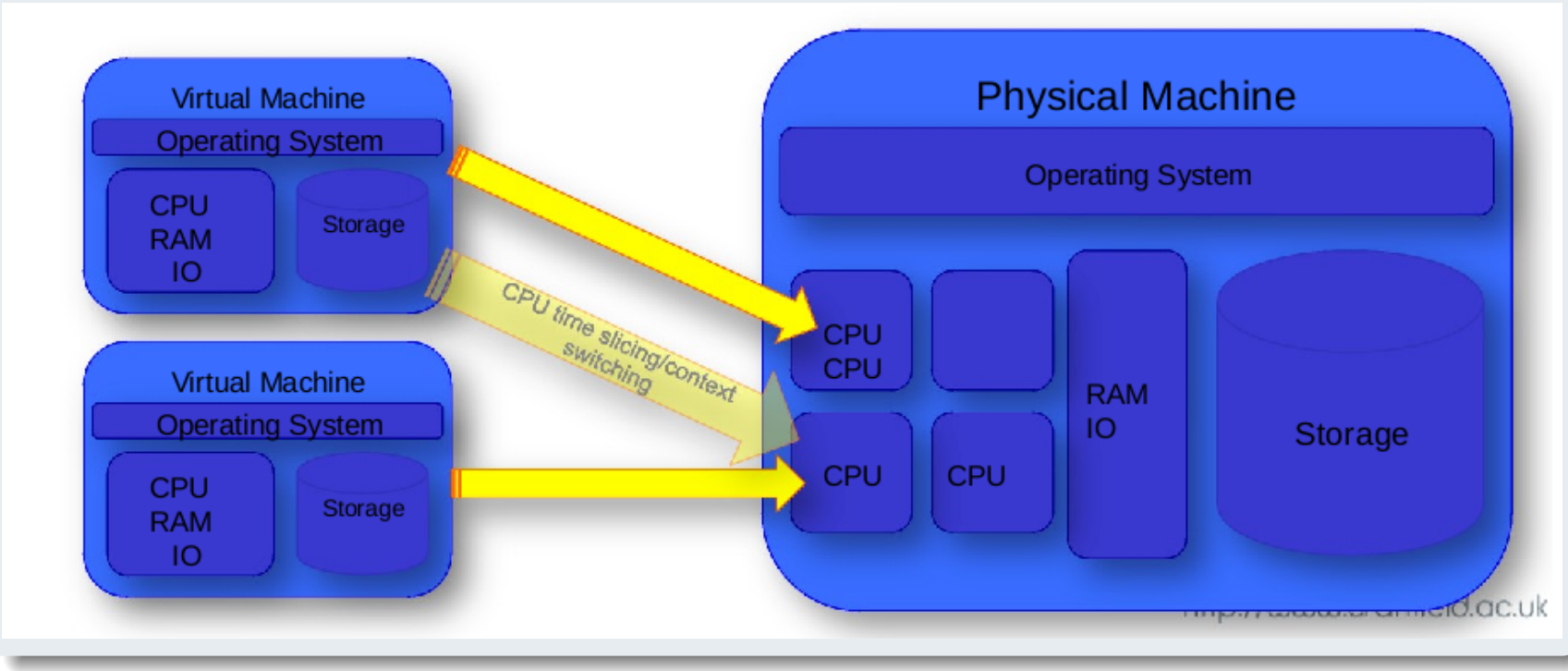
Can we trust the vendor?

- Is the data backed up regularly, and if so, how regularly?
- Is the data secure from access by unauthorised parties?
- How can we be sure that our data will not be exposed?
- Are we talking about our SaaS vendor, or the IaaS/PaaS vendor that they use?



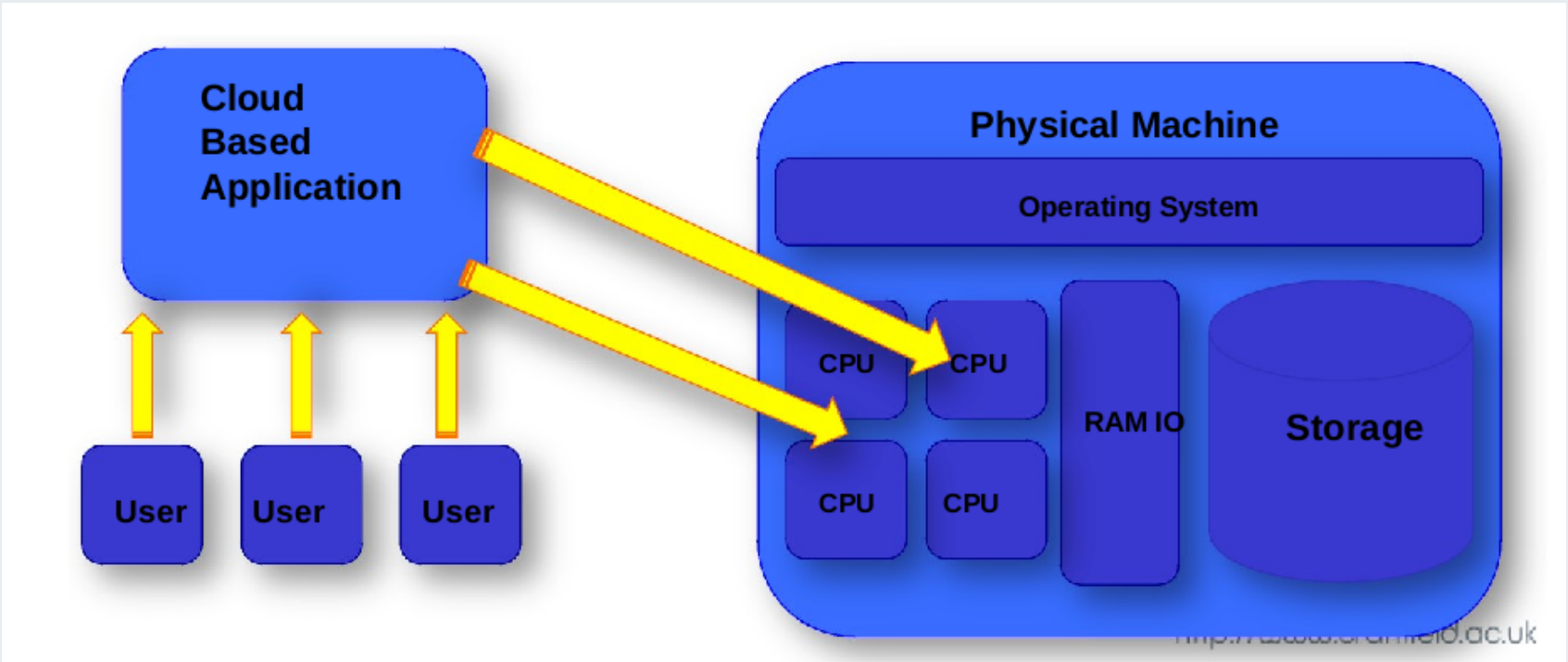
Multi-tenancy

Hosting applications from multiple customers on the same hardware.



Multi-tenancy

can also refer to hosting a single instance of an application for several different users



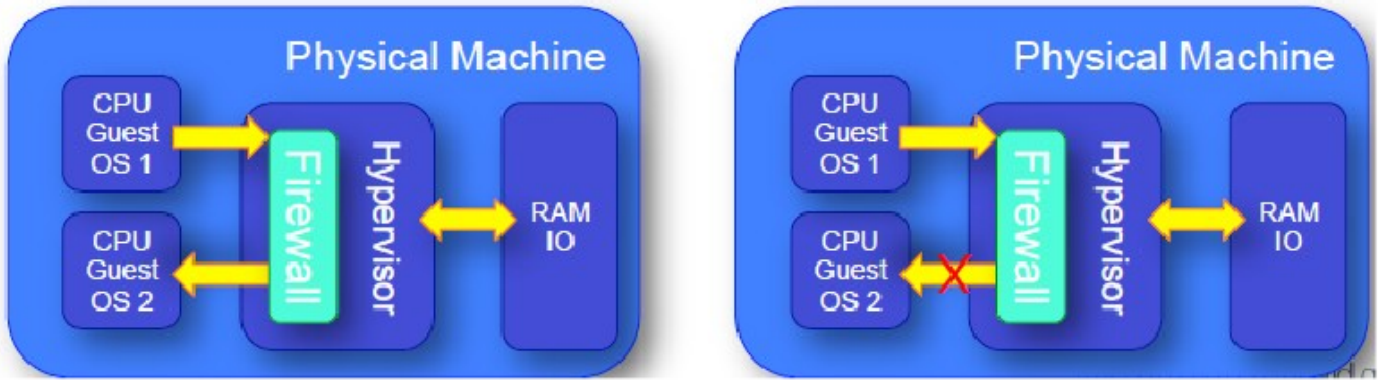
Data Security and Sovereignty

WS Solution to data security in multi-tenanted systems

- Using Xen hypervisor and para-virtualised guest OS — limits access to physical resources

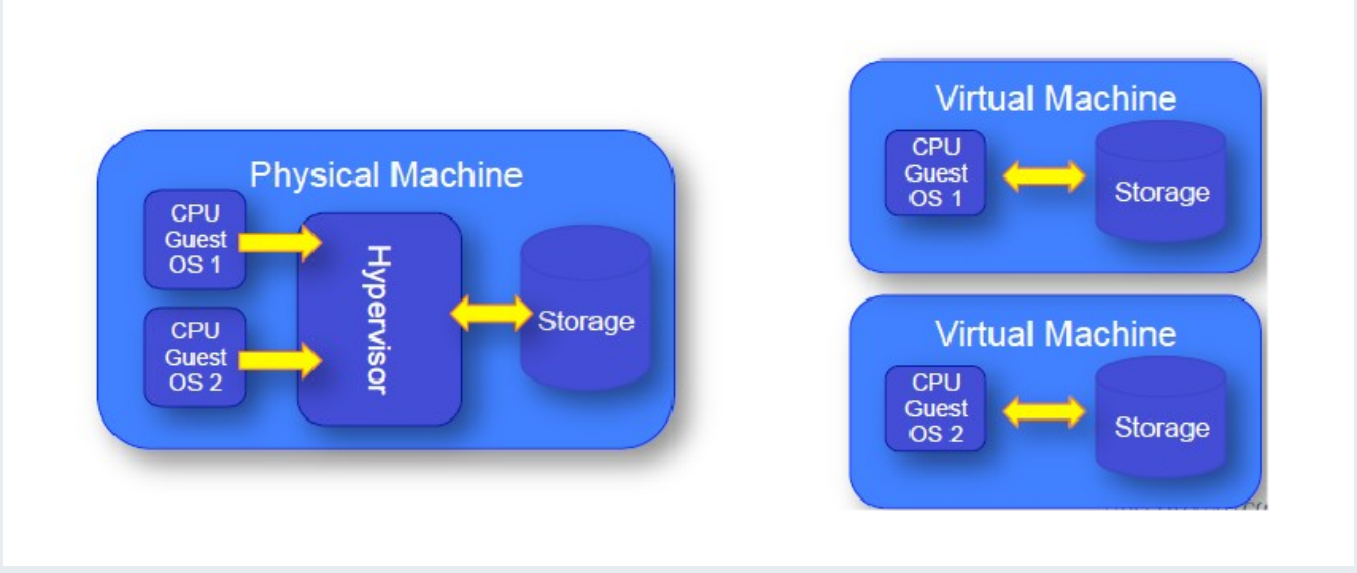
Multiple instances are logically isolated from each other A

- Internal hypervisor firewall limits opportunities to communicate with other guest systems — treated as separate physical hosts
- System RAM is separated using similar mechanisms



AWS Solution to data security in multi-tenanted systems

- No access to physical disk devices — access limited to virtualised devices
- Specific mechanism to wipe data from physical disk once it has been used
- Recommends using an encrypted file system on top of the virtual block storage device



Data Security and Sovereignty

How secure is this environment?

- If the hypervisor controls security, could we make use of an exploit/ undocumented feature to access data or other instances?
- Possibly! No system can be guaranteed to be 100% secure...

What measures can be taken?

- Ensure that the hypervisor is fully patched and up to date
- Larger vendors may have an interest in the development of the hypervisor software
- For example, Amazon uses a highly customised version of the Xen hypervisor, using in-house expertise to ensure security is optimal
- Limit the kernels available to users — reduces the risk of custom kernels circumventing security

Data Security and Sovereignty

Understanding vendor security policies

- Do they have a security policy and is it available for inspection?
- How is it implemented and policed?
- Are independent audits carried out?

Existing auditing standards can be used

- Auditing Standards No. 70 (SAS70) Type II Audit
- Not specific to Cloud Computing security
- "SAS70 certifies that a service organization has had an in-depth audit of its controls (including control objectives and control activities), which in the case of AWS relates to operational performance and security to safeguard customer data" — `http://aws.amazon.com/security/`
- But are we confident that these controls are themselves comprehensive and correct?

Data Security and Sovereignty

Some attempts at standardising — Open Cloud initiative

- Members include IBM, Cisco, SAP, EMC
- Members do not include Amazon, Google, Microsoft, Salesforce
- A young market — may take time for standards to appear
- In the mean time, consumers of Cloud services will need to independently check each vendor

Data Security and Sovereignty

Some best practices in security

- Protect in-transit data — use SSL on instances, or create a VPN
- Storing data — use encryption to protect data before storing it on the Cloud
- Protect AWS credentials — these are the access keys to your Cloud data
- Implement security groups — isolate access to instances from different user groups
- Disable password based security and use certificates for access to instances
- Ensure your application is secure —- patch your instance OS and applications
- Migrate security mechanisms from existing applications to the Cloud

AWS Security Documents —

`http://media.amazonwebservices.com/Whitepaper_Security_Best_Practices_2010.pdf`

`http://s3.amazonaws.com/aws_blog/AWS_Security_Whitepaper_2008_09.pdf`

Data Security and Sovereignty

What is data sovereignty?

- Most countries have legislation to ensure that any personal data about their citizens is secure
- Each country will have a different set of laws, which need to be implemented by companies operating there
- Data sovereignty refers to the set of rules that are applicable to stored data within a particular country
- For example, data held in the UK is subject to UK (EU) data protection laws
- Typical rules include
  - Ensuring data is not given to third parties without consent
  - Personal data held by an organisation is available to the citizen on request
  - Processing of data is regulated according to common rules

Data Security and Sovereignty

Some examples of data regulated by UK law

- Images from CCTV cameras
- Government database of information about all children in the UK
- A company that provides personal data to a third party in order for them to carry out a sub-contracted activity (ie. Call centres or payroll)
- A company enters administration and sells assets to third parties in order to recover funds. If those assets include customer data, UK data protection would apply

See also `http://www.gov.uk/data-protection`

Data Protection Principles

Personal data must be:

- used fairly, lawfully and transparently
- used for specified, explicit purposes
- used in a way that is adequate, relevant and limited to only what is necessary accurate and, where necessary, kept up to date kept for no longer than is necessary
- handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction or damage

Data Protection Principles

There is stronger protection for sensitive data:

- race
- ethnic background
- political opinions
- religious beliefs
- trade union membership
- genetics
- biometrics (where used for identification)
- health
- sex life or orientation

Data Security and Sovereignty

Some examples of the problem of data sovereignty

- In the EU, data can be transferred between member states
- However, it is illegal to export data outside the EU without consent
- This might occur when transferring customer data to a call centre
- If the Cloud vendor facility is outside the EU, storing restricted data in the Cloud would be a breach of the law

Data Security and Sovereignty

European Data Protection Directive

- Controls transfer of information outside of EU or EEA
- Provision for data transfer given to countries whose local laws are considered adequate
- Currently includes Canada, Switzerland, Argentina

If country is not on EU list of acknowledged countries

- Use EU approved contractual clauses
- Ensure data is only used for explicit purposes
- A procedure should allow incorrect data to be identified and corrected in the 3rd party country

Source: "FAQ Relating to Transfers of Personal Data from the EU/EEA to Third Countries"

http://ec.europa.eu/justice/policies/privacy/docs/international_transfers_faq/international_transfers_faq.pdf

Data Security and Sovereignty

Exceptions to EU Data Sovereignty

- Insisting on specific guarantees for data protection can seriously limit trade
- Individual agreements have been signed with other countries
- The US and EU have agreed a Safe Harbour commitment
- This allows US companies to trade with EU member states without fear of prosecution even if data protection policies do not fulfill EU legislation requirements

Data Security and Sovereignty

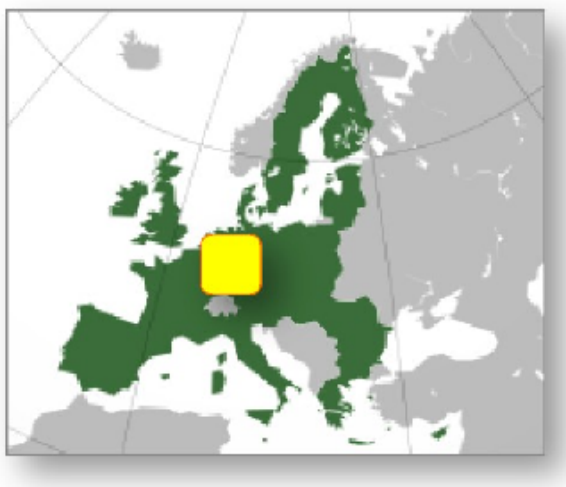
Still a legally contentious area!

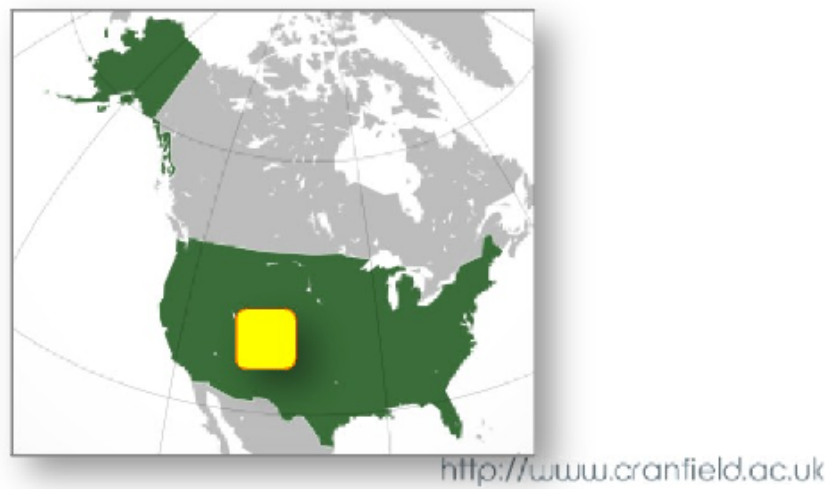
- Individual member states, and even constituencies within each state, may have differing interpretations
- German state of Schleswig-Holstein printed a declaration of the lawfulness of using a Cloud resource outside of the EU
- Paraphrased on a blog: " ... clouds located outside the European Union are per se unlawful, even if the EU Commission has issued an adequacy decision in favor of the foreign country in question (for example, Switzerland, Canada or Argentina)..." Source: Hunton & Williams LLP
- This is because the EU does not confer agent status on an external country, and therefore moving data requires a transfer operation to a third party. This is not legal under EU data protection law. Workarounds are possible!

Data Security and Sovereignty

Some solutions to the problem of data sovereignty

- Create data processing facilities in specific geographical zones
- The zones are strategically chosen to allow better performance in important markets
- Data should be kept in the same zone as the country of origin





Data Security and Sovereignty

Some examples of the problem of data sovereignty

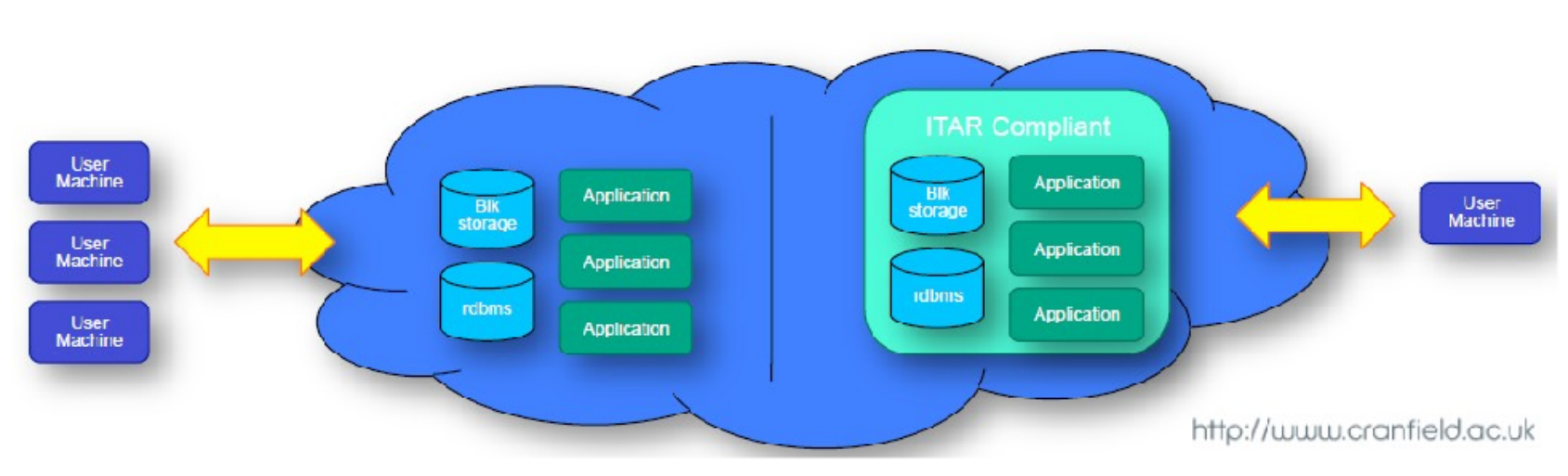
- The transfer of restricted information outside of a country's borders
- ITAR (International Traffics in Arms Regulations)
- For example, storing defense related data outside of a country may require a defense export license
- Storing such material on the Cloud could constitute a serious breach of the law

Data Security and Sovereignty

Some solutions to the problem of data sovereignty

For the US government there is the ITAR implementation:

- All physical computer systems involved in Cloud are ITAR compliant
- Equipment is physically isolated from other Cloud infrastructure
- All support personnel are US citizens with ITAR certification
- Entire processing chain is encrypted



Available from IBM and AWS.

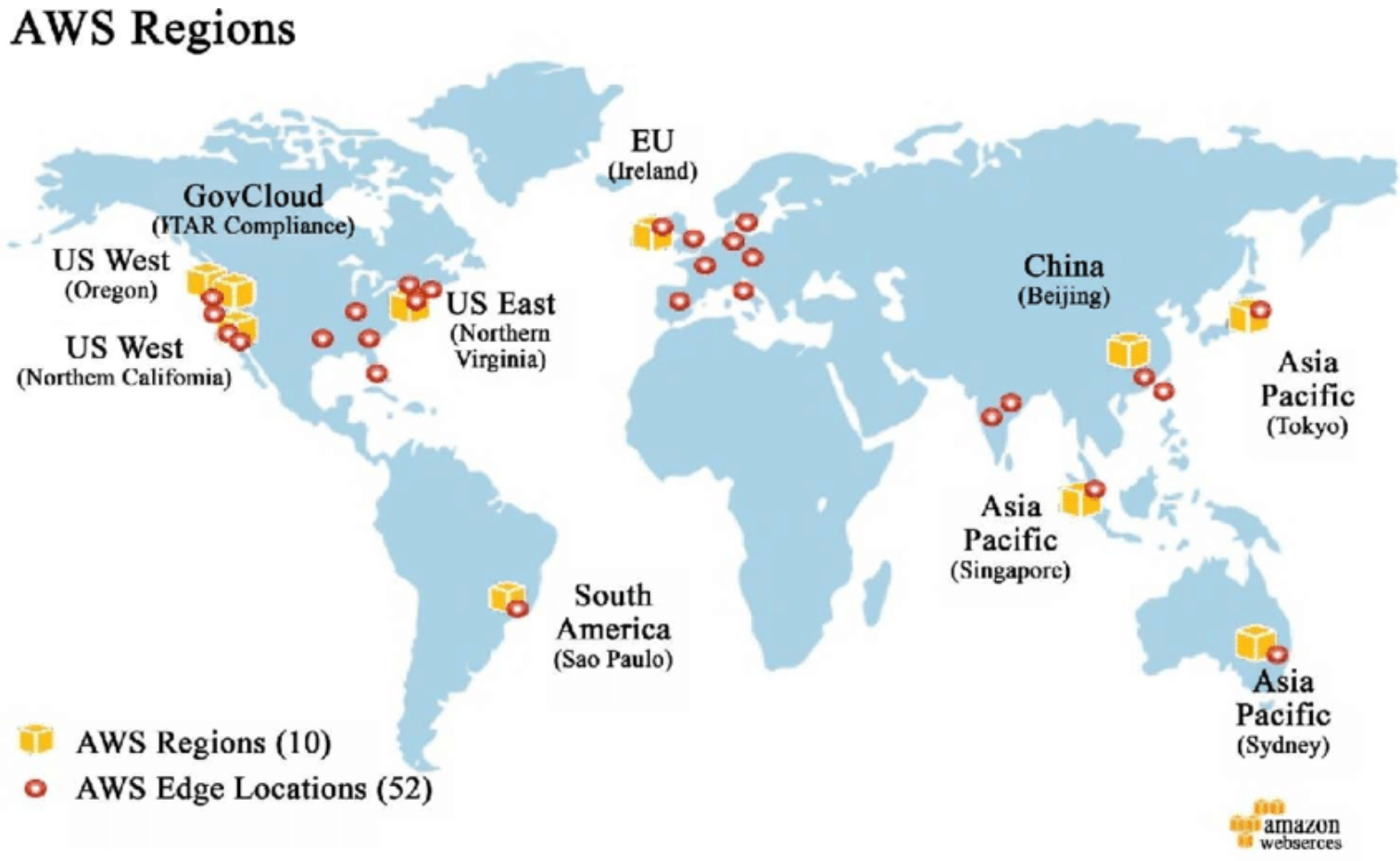
Data Security and Sovereignty

Some solutions to the problem of data sovereignty

AWS has built data centre facilities in the major regions

- the US (North Virginia and California),
- South America (Brazil)
- the EU (Ireland, Germany),
- APAC (Singapore, Bejing, Tokyo, Seoul, Sidney))

Data Security and Sovereignty



Risk Management

Licensing Agreements

- It is possible to use commercial software on Cloud infrastructure
- How can usage be monitored and controlled
- One potential solution — Licensing Servers

Software as a Service

- Terms and Conditions of Software Companies
- The basic problem of restricting unauthorised usage (existing licensing models — per-seat and per-user)
- Usage control can be built into the software
- Can we identify users? — Certification
- SaaS based on metered usage — payments system built into implementation

A brief word on Service Level Agreements (SLA)

- A contractual obligation on behalf of the vendor to supply services at a specified level
- Penalties for breaking SLA
- Should be legally binding, although this may be difficult to enforce in practice
- Multiple tiers of technology make the causes of SLA failure difficult to identify
- Conversely, the technology of the Cloud also makes it easier to track usage and availability
- Moving away from product-oriented economy (warranties) to service-oriented economy (SLAs)
- On-going research into embedding SLA aware infrastructures into Cloud, SOA and Grid environments

Risk Management — Cloud SLAs Key Facts

1. Service Availability

- Is it measured monthly or annually?
- The proportion of time that it will actually be working. Usually expressed as 99.99% availability - which equates to less than five minutes of downtime per month - and if the service provider fails to meet that measure then penalties apply.

2. Planned downtime

- How much is involved, and can you skip it?
- 99.99% availability during 'scheduled uptime,' but not during 'planned downtime' when the provider carries out regular maintenance and upgrades.

Risk Management — Cloud SLAs Key Facts

3. Service interruptions

- Who starts the clock when a service is disrupted?
- Another reason that you may not get 99.99% availability you expect is because of how it is measured. If a problem occurs then it may take a few minutes before you notice that anything is amiss, and a few more checking your own systems and network connectivity before you identify that the problem is at the service provider end.

4. Liability

- What's the limit?
- Aside from availability, most SLAs deal with the penalties that apply if the service provider breaches the agreement. In the form of service credits or reductions from the monthly bill, not actual payments.

Environmental Issues

Data centre costs to the environment

- Electrical power requirements
- Cooling

Traditional data processing has a large impact on the environment

- Can amount to 30% of a company's energy usage
- IT generates as much CO2 as the world's airlines — about 2%
- Using a PC for 15 minutes can generate up to 7g of CO2
- Boiling a kettle uses about 15g of CO2

Environmental Issues

Google has some statistics

CO2 Emissions	Google Searches
CO2 emissions of an average daily newspaper (100% recycled paper)	850
A glass of orange juice	1,050
One load of dishes in an EnergyStar dishwasher	5,100
A five mile trip in the average U.S. automobile	10,000
A cheeseburger 7 15,000	
Electricity consumed by the average US household in one month	3,100,000

Source: Google

`http://www.google.com/corporate/green/datacenters/`

Uspenski Cathedral, Helsinki

Large data centre being built underneath Uspenski Cathedral

- Heat from the computer nodes used to heat local business and homes
- Sufficient energy to heat 500 homes
- Saves on cooling costs
- Offsets electricity consumption against heating provision — uses about half the energy of a typical data centre
- Project energy cost savings: 375,000 EUR / year

Data Mobility

Proprietary Vendor Lock-in:

- Different standards, resources, architectural abstractions (IaaS, PaaS, SaaS)
- Different frameworks and programming environments (.Net, Java)
- Scenarios: Price increases, provider bankruptcy, change of service offering , functionality) . . .
- If a customer uses a service for long term, he invests in this service:
  - The owned business model is focused on the service
  - Employees are trained
  - Services are refined
- Migration costs if you needed to walk away from your current service provider.
- Reasons for walk away — Technology improvement elsewhere, reduced prices, mergers or corporate purchases
- financial cost to “hop” clouds

Data Mobility

Some attempts at standardization

- AWS API becoming a de facto standard in some areas
- Even in similar environments (Eucalyptus and AWS), hard to achieve because of mismatch between offered resources (e.g. Walrus and S3)
- OpenNebula

Will Cloud vendors participate in open standards?

- Vendors have invested heavily in infrastructure and need to ensure a return on this expenditure $\Rightarrow$ No motivation to make data mobility easy, as it makes it easier to move to another vendor

Costs associated with moving between vendors

- Moving data off Cloud can be expensive, as transactions and data bandwidth are both chargeable.
- If your application generates large amounts of data, it may be cheaper to keep it on the Cloud than transfer it to local storage (within a reasonable amount of time)

Cloud Security Body

Cloud Security Alliance (CSA) is a not-for-profit organization with a mission to "promote the use of best practices for providing security assurance within Cloud Computing, and to provide education on the uses of Cloud Computing to help secure all other forms of computing".

Gartner Hype Cycle 2015