

G54CCS – Connected Computing at Scale

Principles of “Cloud Computing”

Last week...

- Dr Martin Flintham
 - mdf@cs.nott.ac.uk
- Dr Richard Mortier
 - rmm@cs.nott.ac.uk
- Course material
 - <http://moodle.nottingham.ac.uk/course/view.php?id=14450>

- What is “The Cloud”?

What is “The Cloud”?

- “It’s a trap. It’s worse than stupidity: it’s marketing hype. Somebody is saying this is inevitable – and whenever you hear that, it’s very likely to be a set of businesses campaigning to make it true”
 - Richard Stallman, FSF, 2008
- “It’s nothing new... we’ve redefined cloud computing to include everything we do... I don’t understand what we would do differently... other than change the wording of some of our ads”
 - Larry Ellison, CEO Oracle, 2008

What is “The Cloud”?

- It is a buzzword
- It is a philosophy
- It is about delivering **services** to users via the Internet
 - End users, developers, companies
- No single technology is cloud computing
 - Makes use of many elements of computer science
 - User interfaces, Programming, Networking, Databases

Personal Computing

- You have a laptop or a pc
 - Local software installation
 - Local operating system maintenance
 - Local system maintenance
- Obvious hardware / software limitations
 - CPU, memory, network
- Very low utilization
 - How many CPU cycles are you actually using?
- Very high up-front cost
 - No **economy of scale**
- The “computing” happens on your desktop



Software

Libraries / Runtimes

Operating System

Hardware

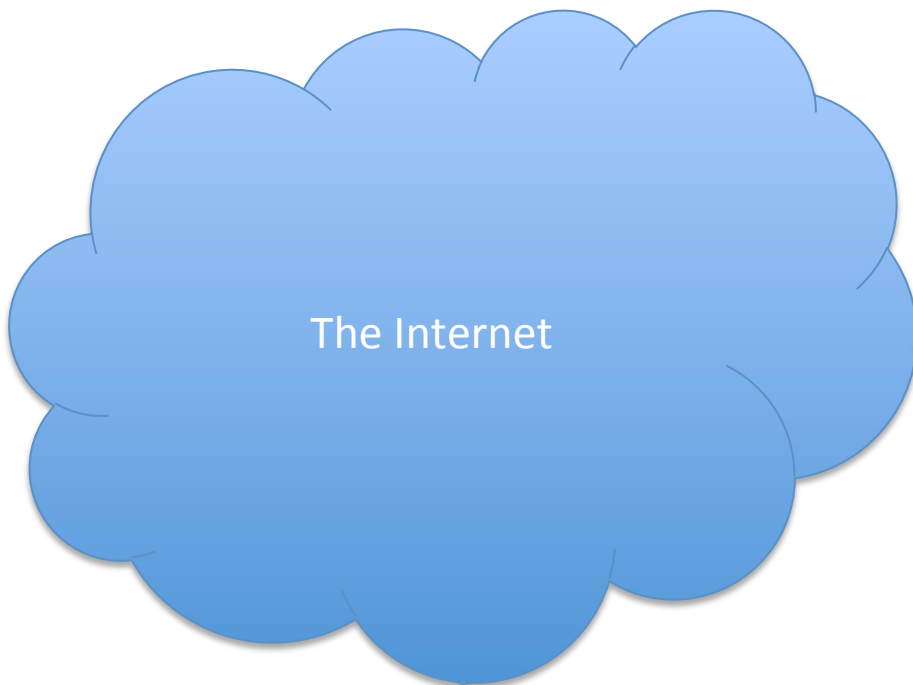


Software

Libraries / Runtimes

Operating System

Hardware



No Software

Any Operating System

Any Hardware

The Internet



Hosting a simple blog?

My Blog

Wordpress software

PHP Runtime

Apache Webserver

SQL Database

Linux

Hardware

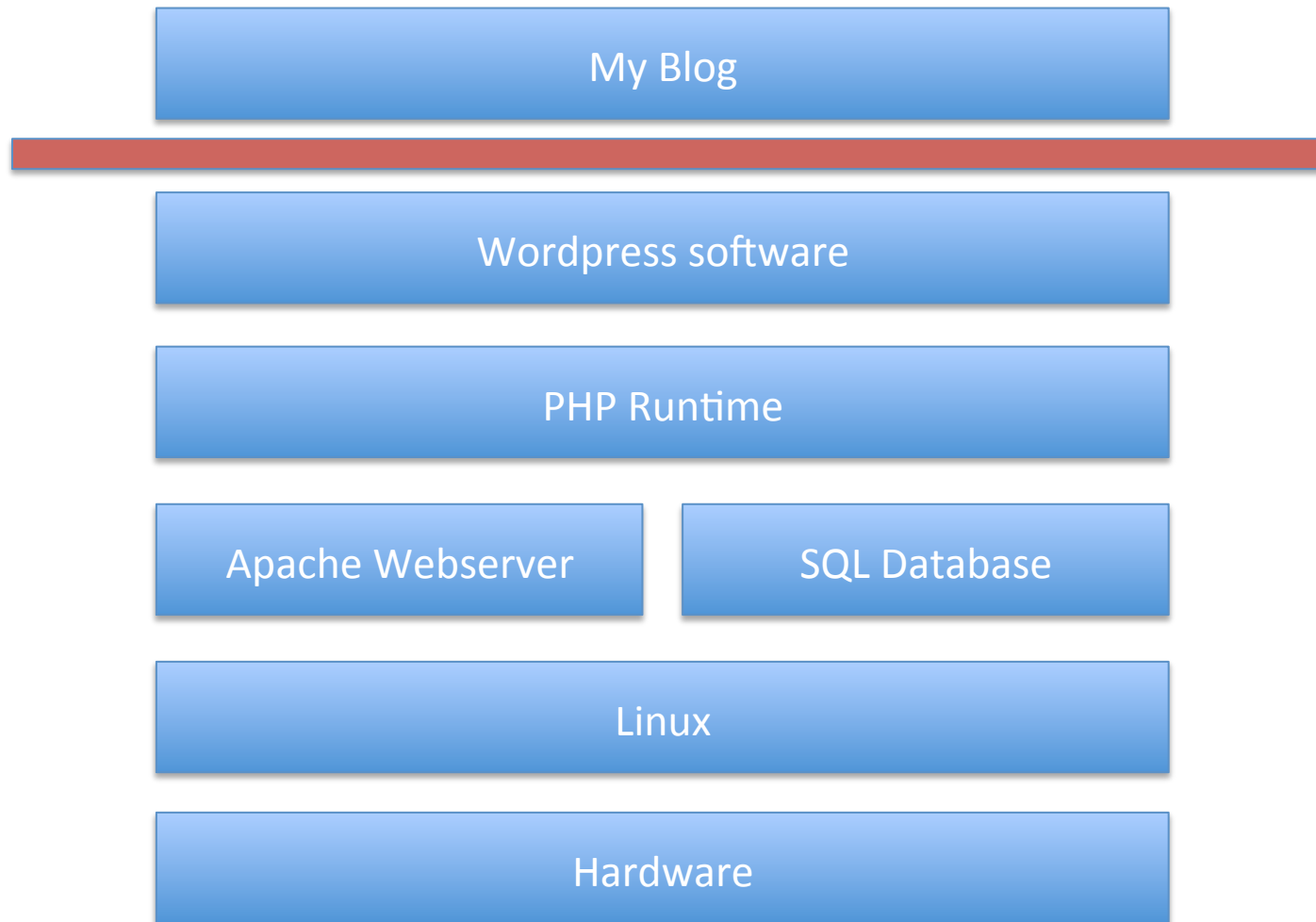
A Data Rich Society

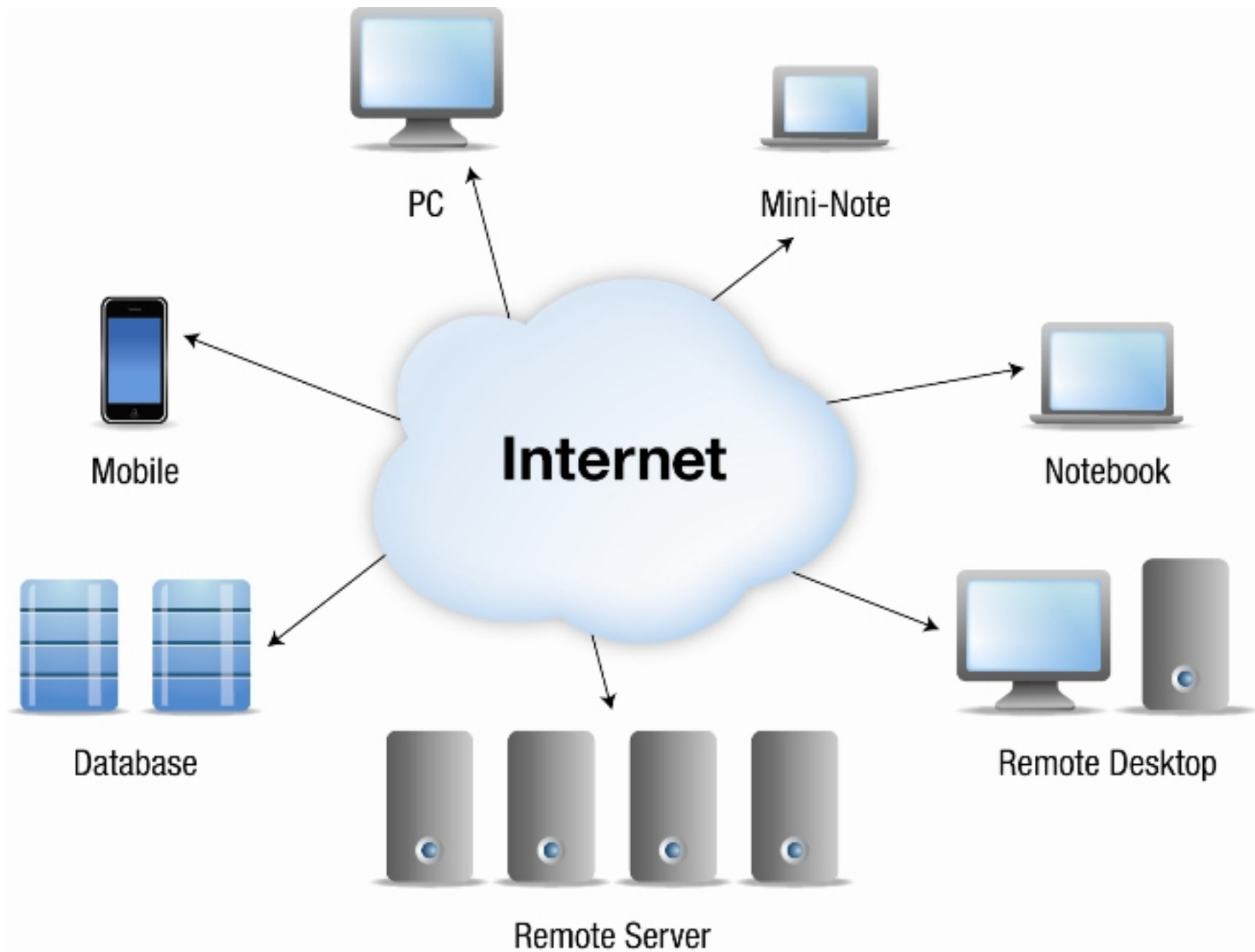
- “Web scale” problems and utilities
 - “640K ought to be enough for anybody”
- Wayback machine - 2 PB data + 20 TB/month (2006)
- Google - 24 PB data a day (2009)
- Large Hadron Collider - 15 PB data a year (2008)
- BBC iPlayer – 7PB data transferred per month (2013)
- Hotmail – 150PB data (2013)
- Facebook – 100PB disk storage (2013)
- Flickr, Twitter, Youtube, Google Docs, Dropbox...

Datacenter as Computer

- Computing paradigm shift
- “Program”
 - Web search, document editing, email, maps, mass data processing
- “Computer”
 - An abstraction, **you** don’t need to worry about the details
 - 1000s of computers, large storage, fast networks
 - £100,000,000 warehouse-sized facilities and workloads
 - Started as a byproduct of dot-com development – Amazon, Google
- The computing happens “out there in the cloud”
 - Somewhere on the Internet

Hosting a simple blog?

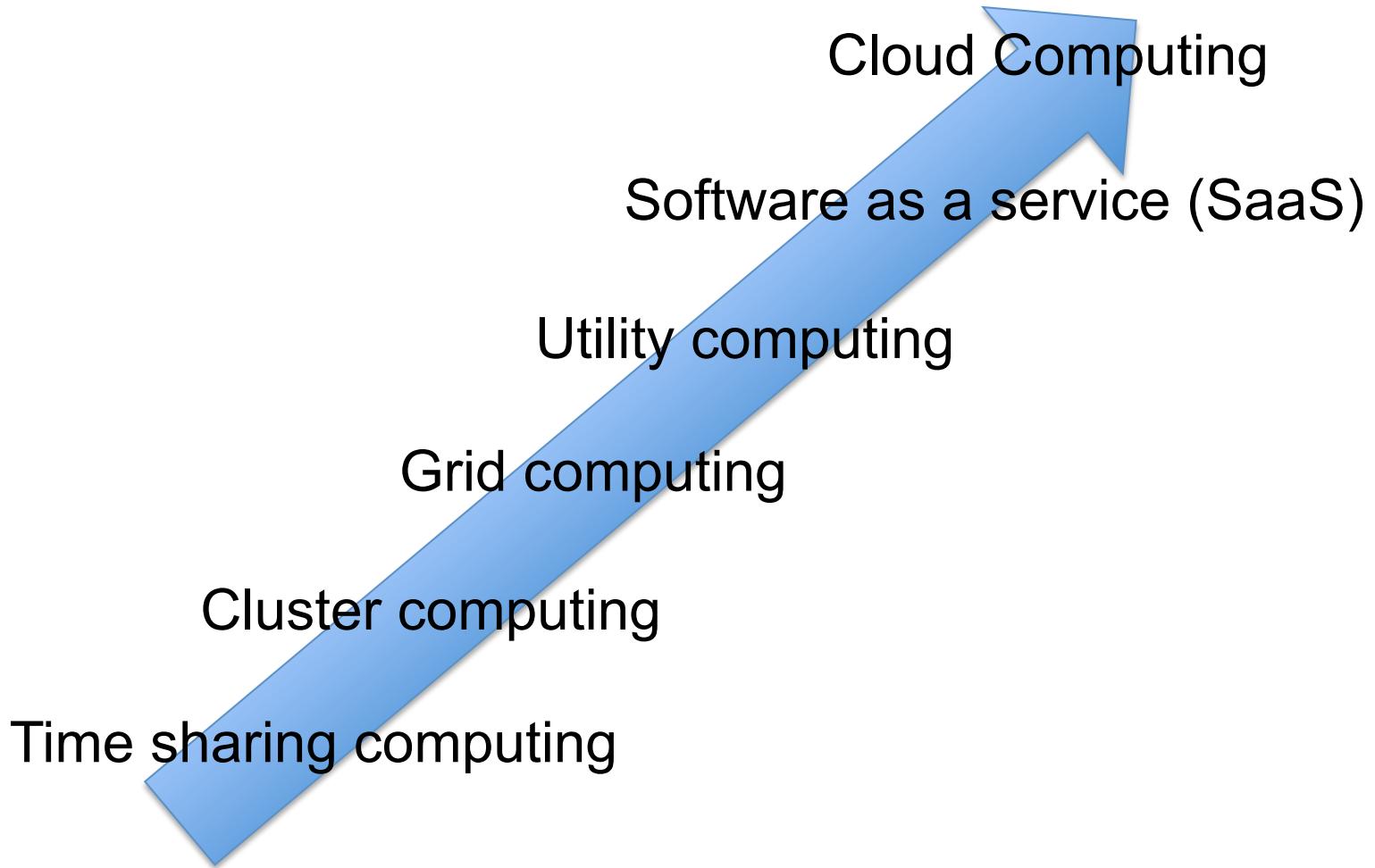








Evolution of Cloud Computing



Grid and Cluster Computing

- Cluster Computing
 - Tightly coupled computers
 - Single system image
 - Centralised job management and scheduling
 - Since 1987
- Grid Computing
 - Resources owned by multiple orgs, coordinated to solve a common problem
 - Loosely coupled computers
 - No single system image
 - Distributed job management and scheduling
 - Since early 1990s



Utility Computing

- Water, Gas, Electricity
 - Provided as commodity services
 - Connected to utility companies’ “public” infrastructure
 - You get them “on demand”
 - You “pay as you use”
- Computing resources as utility
 - CPU power, bandwidth, storage, enterprise level software
- “Computing may someday be organised as a public utility just as the telephone system is a public utility”
 - John McCarthy, MIT Centennial in 1961

Characteristics

- Autonomic computing
 - Self-managing
- Client-server computing
 - Suppliers and customers
- Grid and Cluster computing
 - Networked computers acting together to complete a complex task
- Mainframe computing
 - High-power computing
- Utility computing
 - Computing as a metered resource
- Peer-to-peer computing
 - Participants are suppliers and consumers
- Service-oriented computing
 - Software as a service

3 Key Aspects of Cloud Computing (Amazon)

- The illusion of infinite computing resources available on demand
 - We can grow forever
 - “Your Gmail account will never be full”
- The elimination of upfront commitment by cloud users
 - We don’t need to spend \$100M to get started
 - Capital expenditure vs operating expenditure
- The ability to pay for use of computing resources on a short-term basis as needed
 - Pay as you go
 - Elastic

...or 5 (NIST)

- Rapid elasticity
 - The ability to scale resources up and down as required
- Measured service
 - Controlled and monitored by the cloud provider
 - Billing, resource optimization, planning
- On demand self service
 - Consumers can use service with little human interaction with the provider
- Ubiquitous network access
 - Capabilities available over the network
 - Standard access mechanisms
- Resource pooling
 - Multiple tenancies
 - Location independence

<http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>

Elasticity



Elasticity

- Unpredictable workload spikes are more frequent and significant
 - Death of Michael Jackson:
 - 22% of tweets
 - 20% of Wikipedia traffic
 - Google thought they are under attack
 - Obama inauguration day: 5x increase in tweets
- Cyclical demand curves
 - Daily, weekly, yearly
- “The Cloud” should
 - Automatically handle this
 - Allow us to just ask for more, or less

Sign up Log in



HOW IT WORKS | HOW IT'S USED | FEATURES | PRICING | FOR BUSINESS

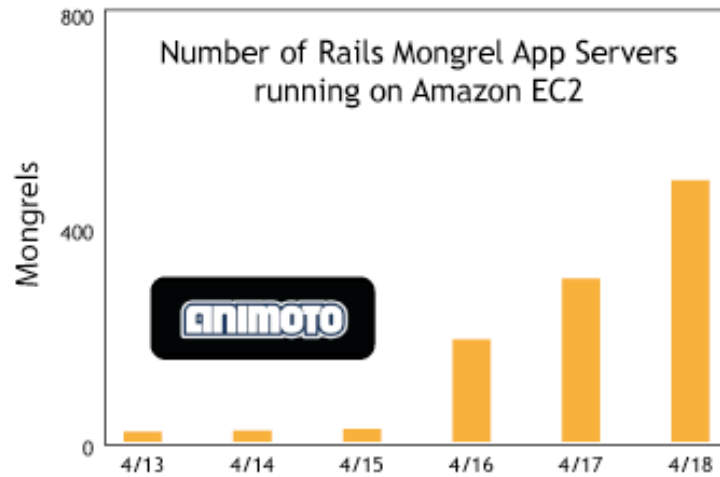
Create videos for home or business

Welcome to the easiest way to make beautiful videos online. Try it for free!

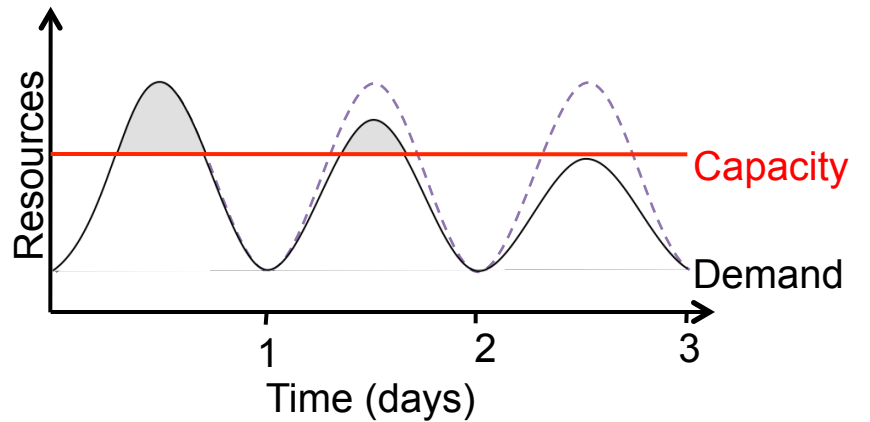
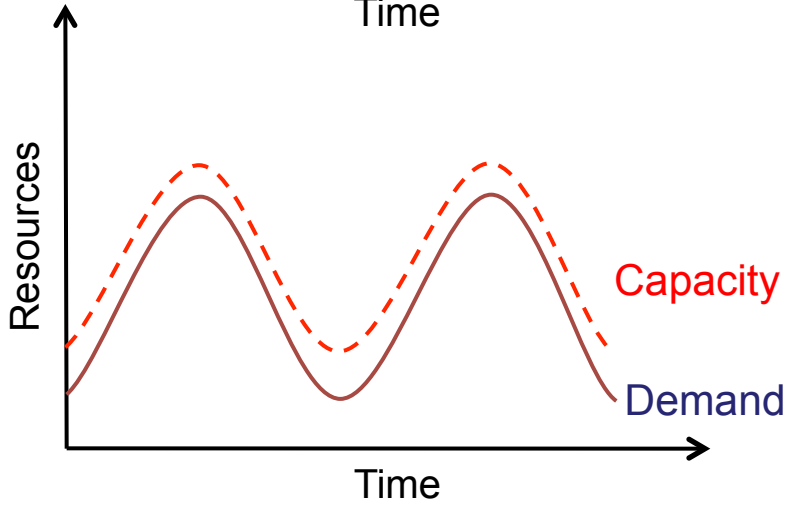
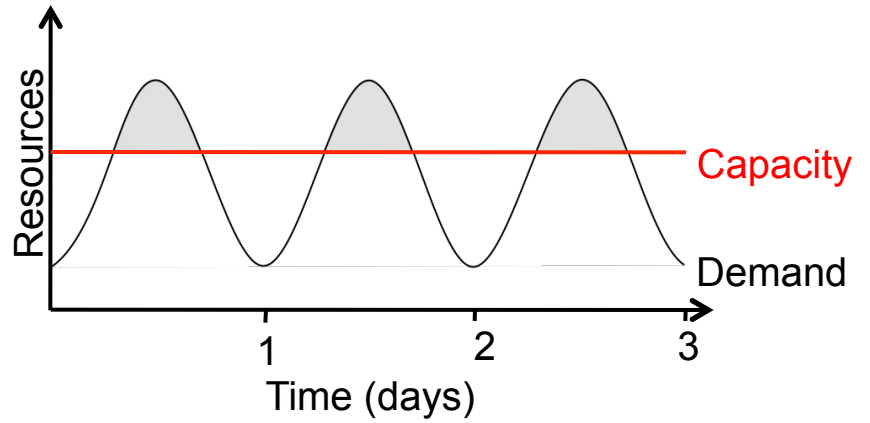
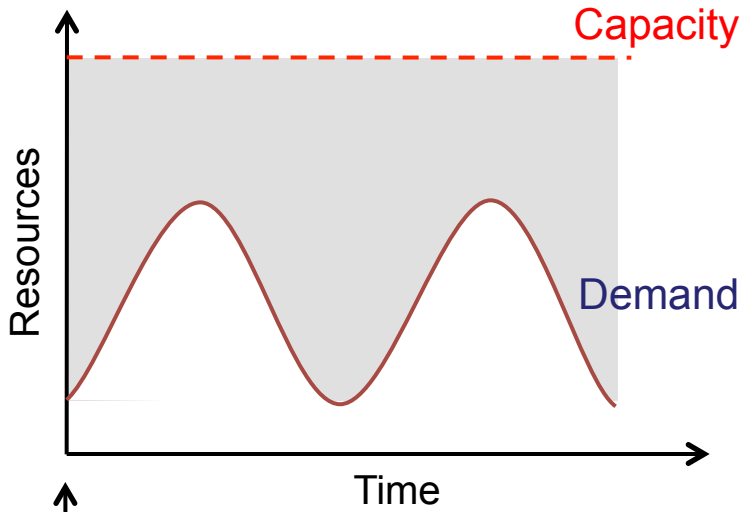
Get started

Need videos for your business or studio?
[Find out more](#)

Animoto



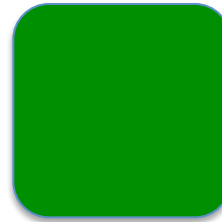
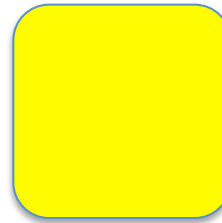
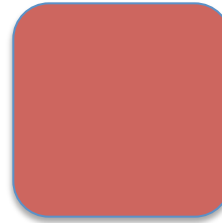
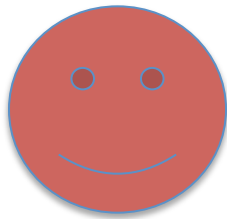
Elasticity



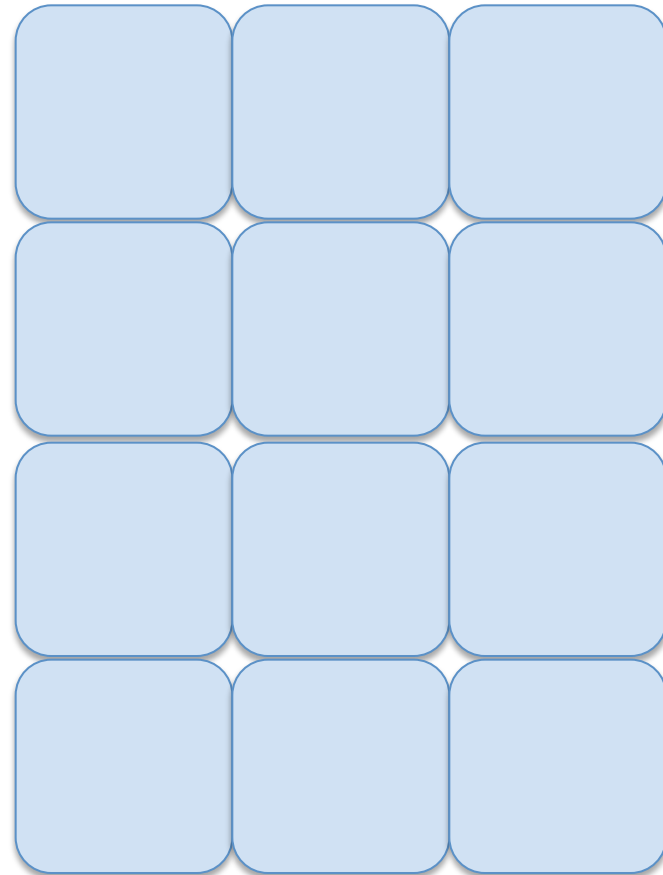
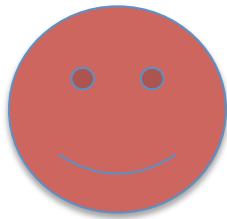
Multi Tenancy / Resource Pooling

- How does a cloud provider support this illusion of elasticity?
- Pool available resources
 - Economy of scale again, have a big data center
 - Make the best use of our resources
- Users (tenants) share the underlying resource pool
 - Dynamically allocated resources to tenants as required
 - Tenants have no control of how resources are allocated, and even physically exactly where they are
 - Tenants should not know that they are sharing resources

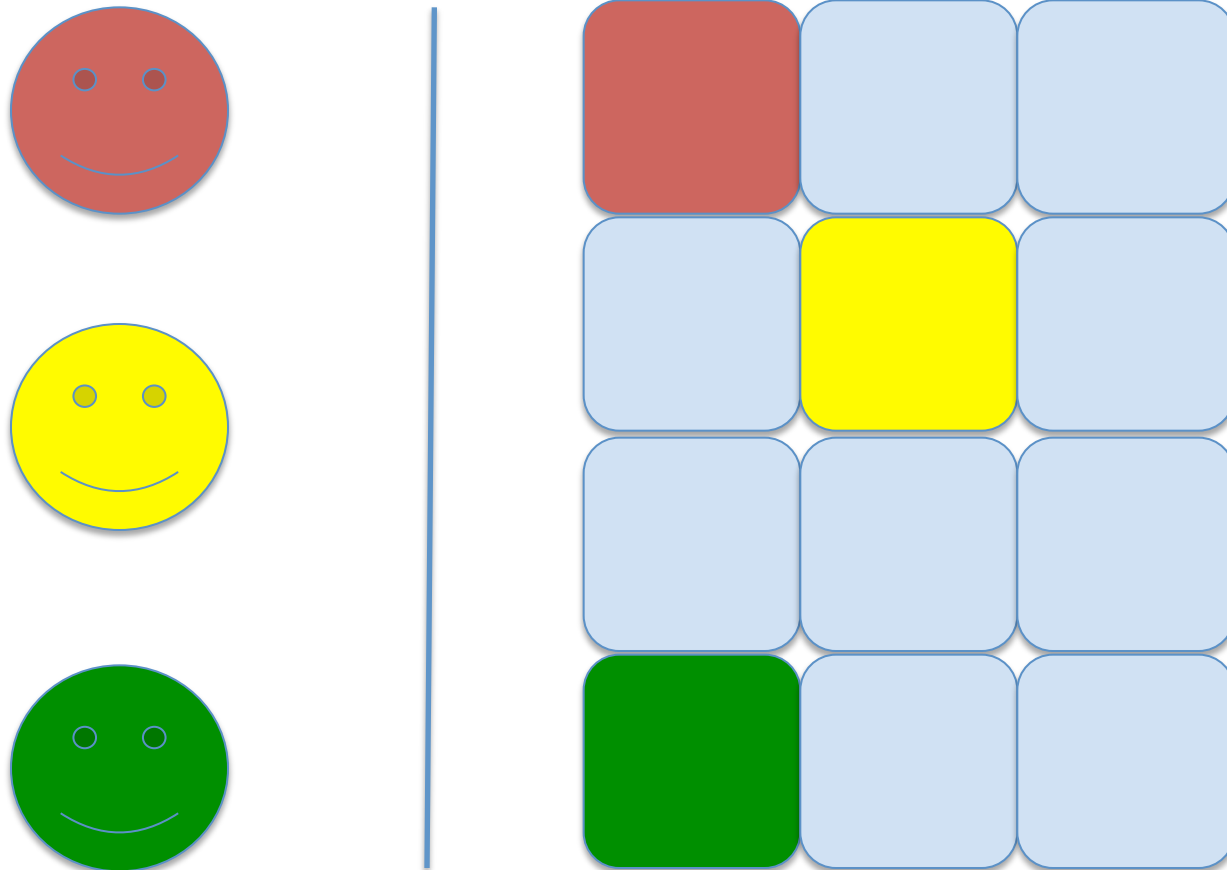
Multi Tenancy / Resource Pooling



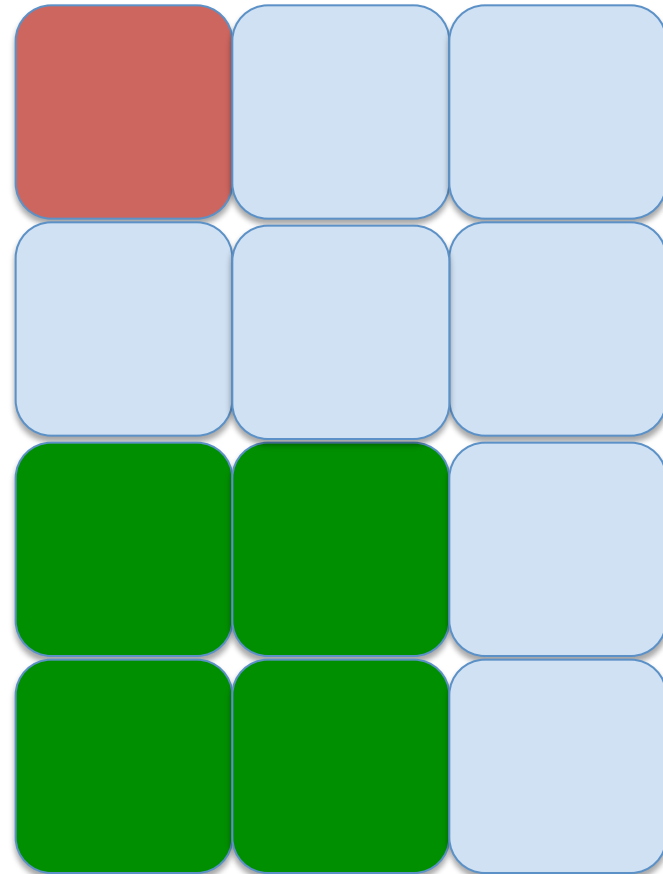
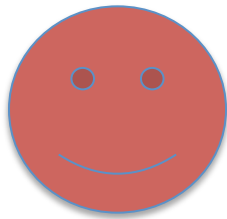
Multi Tenancy / Resource Pooling



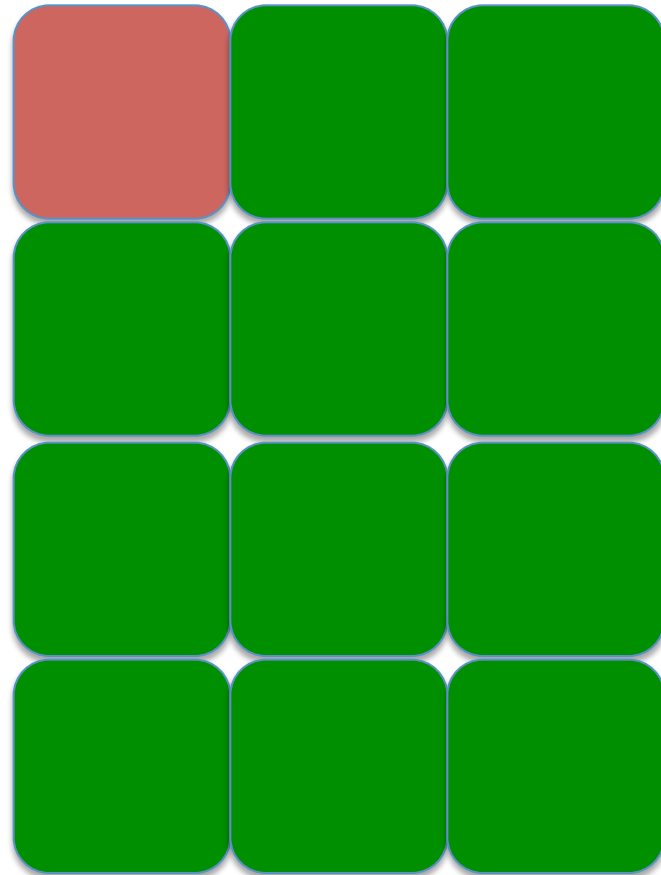
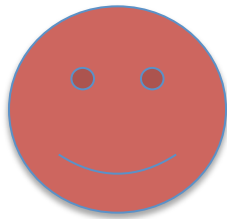
Multi Tenancy / Resource Pooling



Multi Tenancy / Resource Pooling



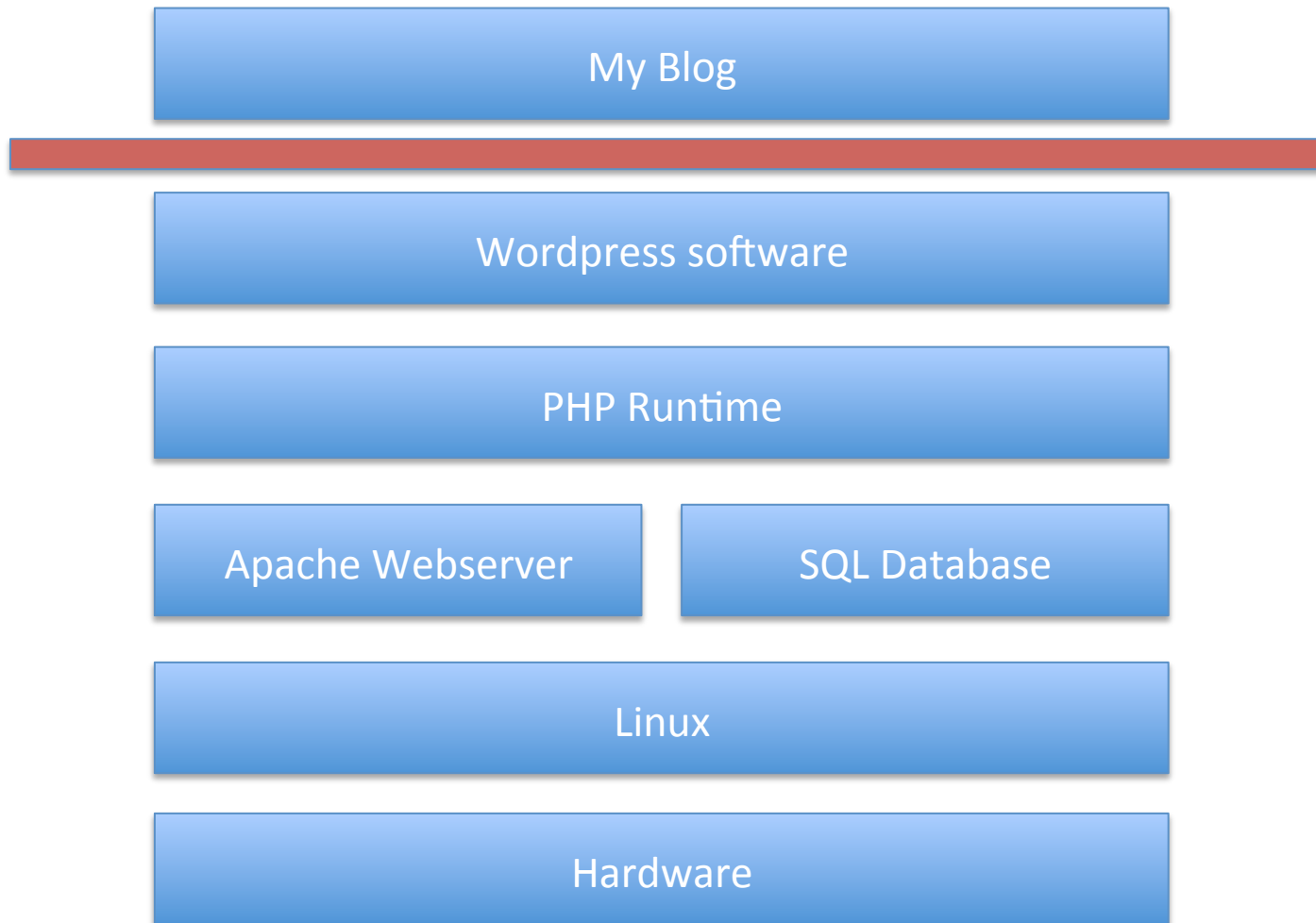
Multi Tenancy / Resource Pooling



Three Cloud Computing Services

- Software as a Service (SaaS)
 - Apps through a browser
 - “Just run it for me”
- Platform as a Service (PaaS)
 - Delivery of a computing platform for custom software development as a service
 - “Give me a nice API and take care of the implementation”
- Infrastructure as a Service (IaaS)
 - Deliver computing infrastructure as a service
 - CPU cycles, storage, network utilities
 - “Why buy expensive machines when you can rent CPU cycles?”

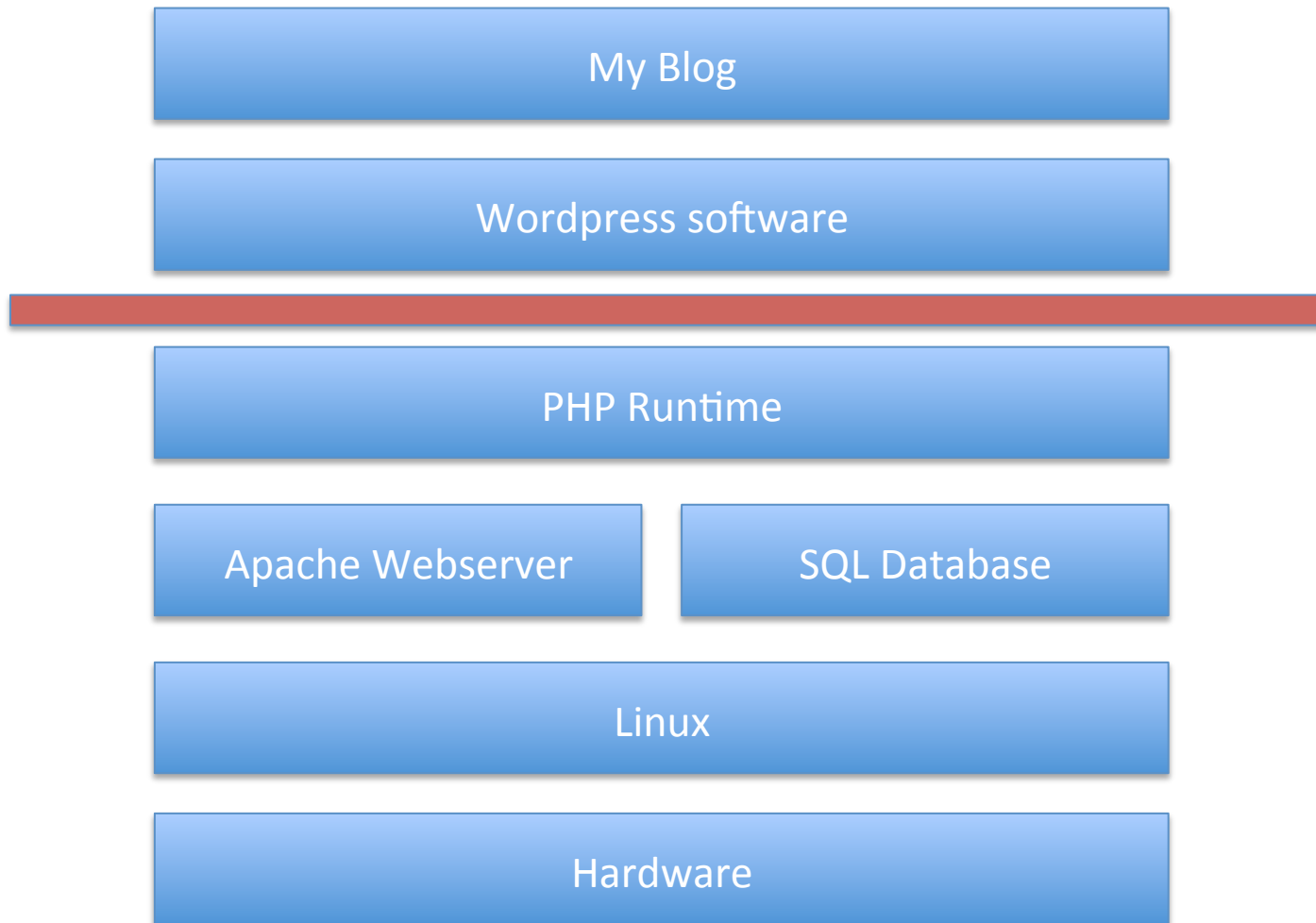
Software as a service



Software as a Service (SaaS)

- Wordpress, Flickr, Google Docs, Gmail, Dropbox, Office365
- The entire software / hardware stack is in the cloud
 - Delivered via “thin client” – web browser, mobile app
 - Application is licensed to a customer as a service on demand
 - Subscription model, pay per transaction, per customer customisation
- Advantages
 - Accelerated feature delivery
 - Central hosting, single configuration, access to (user) data

Platform as a service



Platform as a Service (PaaS)

- Google App Engine, Windows Azure, Hadoop
- Delivery of an integrated computing platform as a service
 - Build / test / deploy custom applications
- Deploy applications in the cloud
 - Abstract the “web server”
 - Cluster of web servers can handle millions of hits
 - Cloud provider manages underlying hardware and platform layers, hosting
 - Multi-tenancy, scalable

Platform as a Service (PaaS)

- Not just a cloud web server
- Scalability / Abstraction
 - Google App Engine request model
 - Large distributed databases / NoSQL databases
 - Amazon SimpleDB
 - Google App Engine datastore
 - Azure AppFabric
 - Service bus
 - Caching

Platform as a Service (PaaS)

```
import webapp2
```

```
class MainPage(webapp2.RequestHandler)
```

```
    def get(self):
```

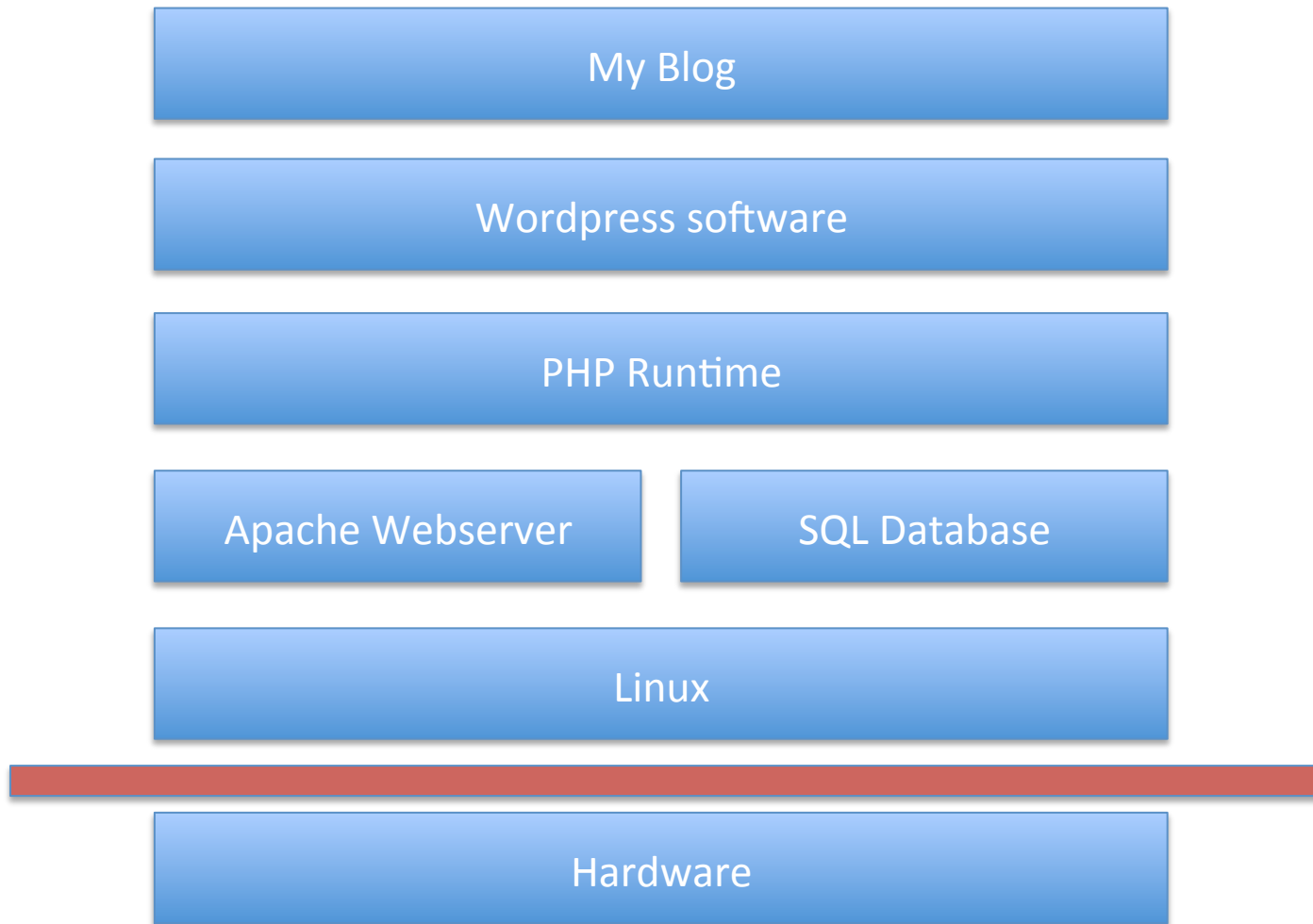
```
        self.response.headers['Content-Type'] = 'text/plain'
```

```
        self.response.write('Hello World!')
```

```
app = webapp2.WSGIApplication([('/', MainPage)], debug=True)
```

- Google App Engine takes care of
 - Scalable data storage
 - Scalable application hosting
 - Billing

Infrastructure as a service



Infrastructure as a Service (IaaS)

- Delivery of computing infrastructure (hardware) as a service
 - Rent virtual computers / servers by the hour
 - Rent disk storage by the byte / block
- Amazon Elastic Compute Cloud (EC2)
- Abstraction from real hardware
 - Looks like a CPU, disk, network interface
 - Doesn't matter if it isn't really
- Customer is responsible for everything else
 - Operating system
 - Application platform (web server, database)
 - Software

Infrastructure as a Service (IaaS)

- Again, use multi-tenancy and resource pooling
- Virtualisation is a key technology
 - Run multiple virtual computers on one physical machine
 - RAM disk / connect a virtual block storage volume
 - Move a running virtual computer to a different physical machine on hardware failure
 - Assign and access via a real IP address
 - Quickly create new virtual computers from templates / snapshots on demand
- Self service / programmatically controlled
- Instagram, Reddit, Netflix
 - Big online services as cloud “customers”

Amazon EC2 Amazon Elastic MapReduce Amazon CloudFront

Navigation

Region: US East

- > EC2 Dashboard
- INSTANCES
 - > Instances
- IMAGES
 - > AMIs
 - > Bundle Tasks
- ELASTIC BLOCK STORE
 - > Volumes
 - > Snapshots
- NETWORKING & SECURITY
 - > Elastic IPs
 - > Security Groups
 - > Key Pairs

Amazon EC2 Console Dashboard

Getting Started

To start using Amazon EC2 you will want to launch a virtual server, known as an Amazon EC2 instance.

[Launch Instances](#)

Note: Your instances will launch in the US East (N. Virginia) region.

Service Health

Current Status	Details
✔ Amazon EC2 (US East - N. Virginia)	Service is operating normally View complete service health details

My Resources

You are using the following Amazon EC2 resources in the US East (N. Virginia) region: [Refresh](#)

- 0 Running Instances
- 0 Elastic IPs
- 0 EBS Volumes
- 0 EBS Snapshots
- 3 Key Pairs
- 1 Security Group

Related Links

- > Documentation
- > All EC2 Resources
- > Forums
- > Feedback
- > Report an Issue

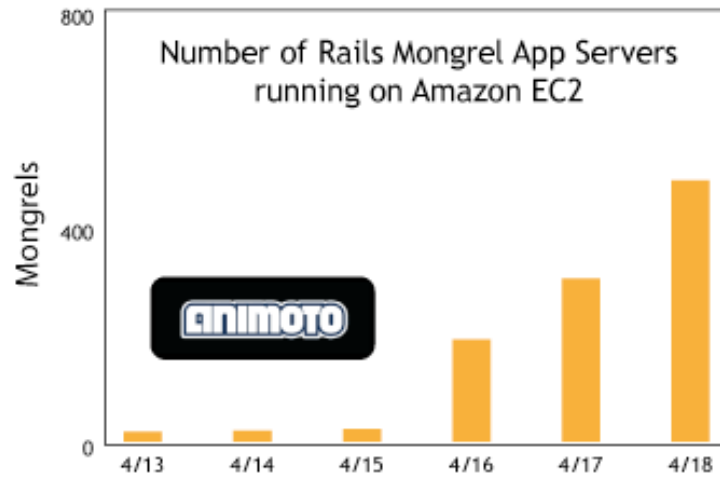
My Instances

[Launch Instance](#) Instance Actions Reserved Instances

Viewing: All Instances All Instance Types

	Name	Instance	AMI ID	Root Device	Type	Status	Lifecycle	Security Groups	Key Pair Name	Monitoring	Virtualization
<input checked="" type="checkbox"/>		i-1f4d7f75	ami-1634de7f	ebs	t1.micro	● running	normal	ciab-edu	ciabkey	disabled	paravirtual
<input type="checkbox"/>		i-597aec32	ami-7fa44a16	ebs	m1.large	● stopped	normal	ciab-portal	ciabkey	disabled	paravirtual
<input type="checkbox"/>		i-b56752de	ami-3b16ff52	ebs	c1.medium	● stopped	normal	ciab-portal	ciabkey	disabled	paravirtual
<input type="checkbox"/>		i-0655586d	ami-714ba518	ebs	c1.medium	● stopped	normal	ciab-portal	ciabkey	disabled	paravirtual
<input type="checkbox"/>		i-ce9585a5	ami-714ba518	ebs	c1.medium	● stopped	normal	ciab-portal	ciabkey	disabled	paravirtual
<input type="checkbox"/>		i-c8eef7a3	ami-4b4ba522	ebs	m1.large	● stopped	normal	Spice	ciabkey	disabled	paravirtual
<input type="checkbox"/>		i-7bbef911	ami-f11ff098	ebs	c1.medium	● stopped	normal	ciab-edu	ciabkey	disabled	hvm

Animoto



Everything as a Service?

- Desktop as a Service
- Communication as a Service
- Monitoring as a Service
- Government as a Service
- **Workers** as a Service
- ...
- SaaS, PaaS, IaaS (SPI) are the most commonly used models

The Cloud or A Cloud?

- Public (external) clouds
 - Open market for on-demand computing
 - Google, Amazon, Microsoft
- Private (local)
 - For corporations with large scale IT
 - Nottingham University has “a cloud”
 - VMWare
- Hybrid
 - Extend private cloud by connecting to external (public) cloud vendors / services
- Burst
 - Use the local cloud, burst into the public cloud when resources are needed

Advantages

- Flexibility
 - Software
 - multiple software platforms
 - Access
 - access resources from any connected machine
 - Deployment
 - from anywhere at anytime
 - Infrastructure is controlled by software
- Scalability / Agility
 - Instant
 - Control via software
 - Start small, scale resources up / down as required

Advantages

- Maintenance
 - ... is the responsibility of the vendor
 - Software updates
 - Hardware health / replacement
 - Reduce the size of the IT department
- Utilization
 - Consolidation of large numbers of resources
 - Economy of scale
 - Higher utilization
 - Reduced power usage, CO2 footprint

Advantages

- Cost
 - Pay as you go model
 - Tap the infrastructure of corporate giants
 - Google, Amazon, Microsoft
 - No upfront cost
 - Time to service / market reduced

Disadvantages

- Security
 - Am I at risk from other tenants?
 - What if other tenants do bad things?
 - Am I at risk from network traffic snooping?
 - What service-level-agreement do I have? What happens if it breaks?
- Privacy
 - Can the vendor look at my data?
 - If I don't know where it's located, which laws apply?
- Vendor lock-in / migration costs
 - All of my code is written for Google App Engine
 - It may cost a lot to move from legacy systems
- Network-dependent
 - Internet connectivity in some countries may be unreliable

Make Money by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

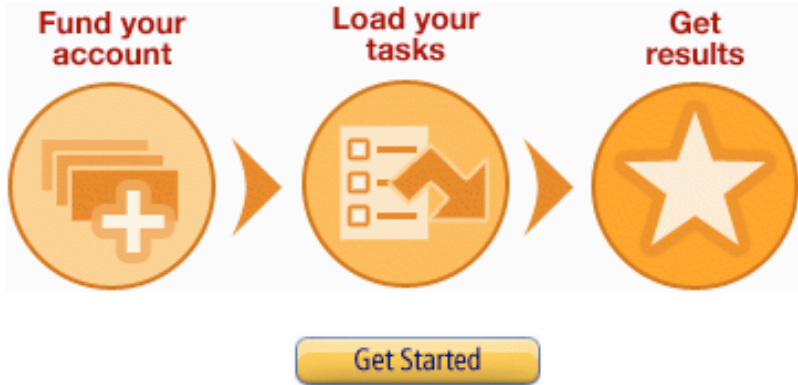


Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Get started.](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results





Summary

- The cloud == making it someone else problem
- Characteristics of the cloud
 - Utility, elasticity, illusion of scale, economy of scale...
- Cloud ...as a service
- Types of cloud
- Advantages, disadvantages