CAB432 Cloud Computing
Lecture 1 - Introduction

Faculty of Science

Cloud is many things, but think about *elasticity* and *scale*.

WHAT IS CLOUD?

XaaS, Public, Private,

THE BASICS

What is Cloud Computing?

Elastic Utility Computing at Scale:

Utility computing

- IT infrastructure and services 'piped in'
- We turn on the tap as needed.
- Pay only for what we consume

Elastic provisioning

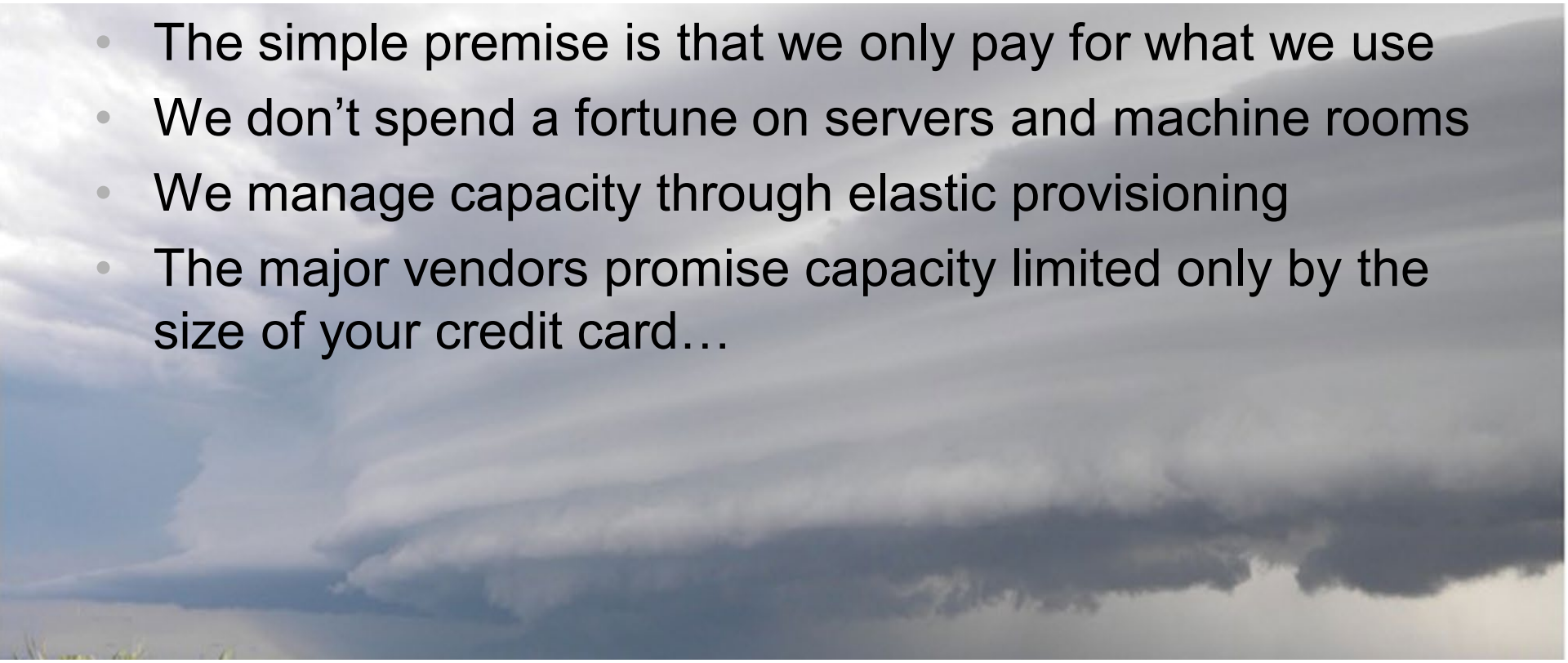
- Scale the services as we need them
- Automatically adjusted with load

Modern cloud computing operates globally and at *massive scale*.

The Origins of Cloud

- Amazon had an online bookstore… and they were better than most at large scale distributed computing
- Moving large scale IT from CAPEX to OPEX
- AWS was launched in 2002 and offered initial Compute (EC2) and Storage (S3) services in 2006
- Microsoft and Google followed soon afterwards
- They were good at distributed computing too…

The Business Case for Cloud



- The simple premise is that we only pay for what we use
- We don't spend a fortune on servers and machine rooms
- We manage capacity through elastic provisioning
- The major vendors promise capacity limited only by the size of your credit card…

Public, Private and Hybrid Clouds

PUBLIC CLOUD	PRIVATE CLOUD	HYBRID CLOUD
• Services offered by the major vendors • Available globally on a commercial basis • 'Limited only by the size of the credit card'	• Services hosted and managed by large companies or government • Same elastic service model, but smaller scale • Usually more limited service offerings	• Services based on a mix of public and private clouds • Often used to manage regulatory requirements and client concerns over location of sensitive data sets. • Government services, Major corporates

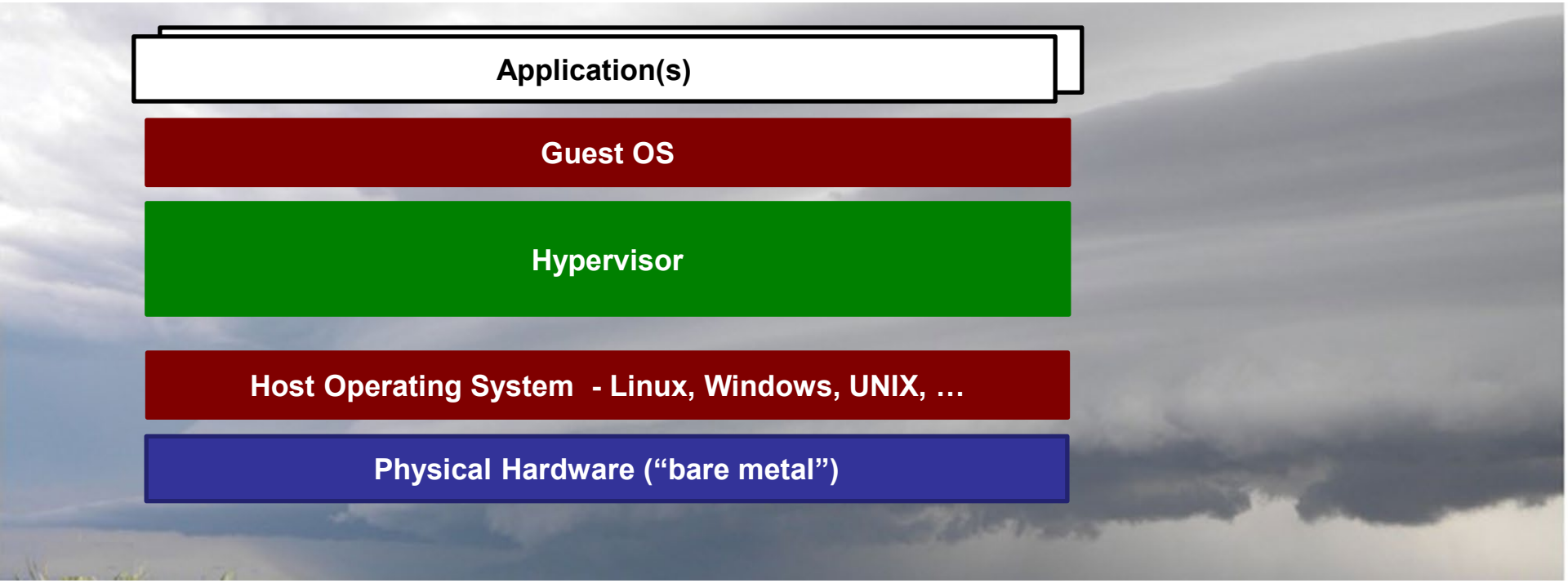
Pre-Requisites for Cloud Usage

- Anyone can use and combine a basic cloud service
  - Phone client and gmail, Flickr, Outlook, Facebook, Twitter
- Business migration requires a fat, bi-directional pipe
  - IT departments have been happy to allow a fat, if constrained, outward pipe for customers to access company sites
  - But the inward pipe has been a trickle, tightly guarded and limited to the barest essentials.
  - But with cloud, *all* the corporate data lies elsewhere, so the network changes fundamentally

Pre-Requisites for Cloud

VIRTUALISATION	ELASTICITY	SCALE
• **Costly to maintain lots of small machines** • **Better to maintain lots of Virtual Machines in the one place**	• Measure load and automatically scale as needed • Scale ***out*** when people are waiting. Scale ***in*** when we have too many machines	• Cloud data centres are huge • Major vendors have nodes around the world and their own undersea cables.

A Quick Guide to Virtualisation



But there are important differences

- Cloud is neither mainframe nor simple virtualised hosting
- As always, the boundaries become somewhat blurred
- One person's small cloud is another's mainframe
  - Variations discussed: http://en.wikipedia.org/wiki/Mainframe_computer
  - But at one level the mainframe remains a single server
- But modern utility computing is fundamentally distinct
- Server farms based on thousands of commodity components, with distributed processing and redundancy
- Racks are being replaced by modular components

The Forecast is Cloudy

- Cloud used to mean *scalable* virtualisation
  - Mainly compute and storage services
  - IaaS and low-end PaaS
- And the on-demand provision of really bad puns
  - plus ça change: “You cannot be Cirrus” (The Economist)
- Cloud has changed fundamentally in recent years
  - Vastly richer offerings in the PaaS and SaaS space
  - Managed services have much higher margins, so no surprises
  - We will consider these in detail later

Utility Computing is a very old Idea

- IT offsite - share a big machine instead of maintaining lots of small ones
- IT as a service – rent rather than buy
- But modern Cloud is *very* different

- http://en.wikipedia.org/wiki/Electronic_Data_Systems
- https://en.wikipedia.org/wiki/IBM_System/370

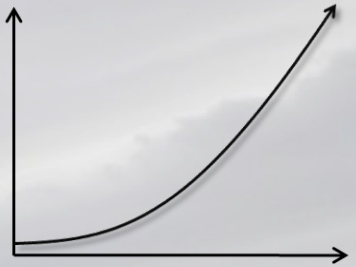
IBM®

EDS an HP company

Elasticity and Cloud Workloads

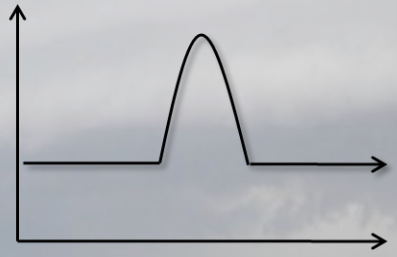
On and Off

Rapid Growth



Predictable Bursting

Unpredictable Bursting



Elasticity means that we can provide resources that match the workload at the time.

And we are able to scale automatically.

As long as we have credit available…

The Scale Speaks for Itself

Source: AWS web pages

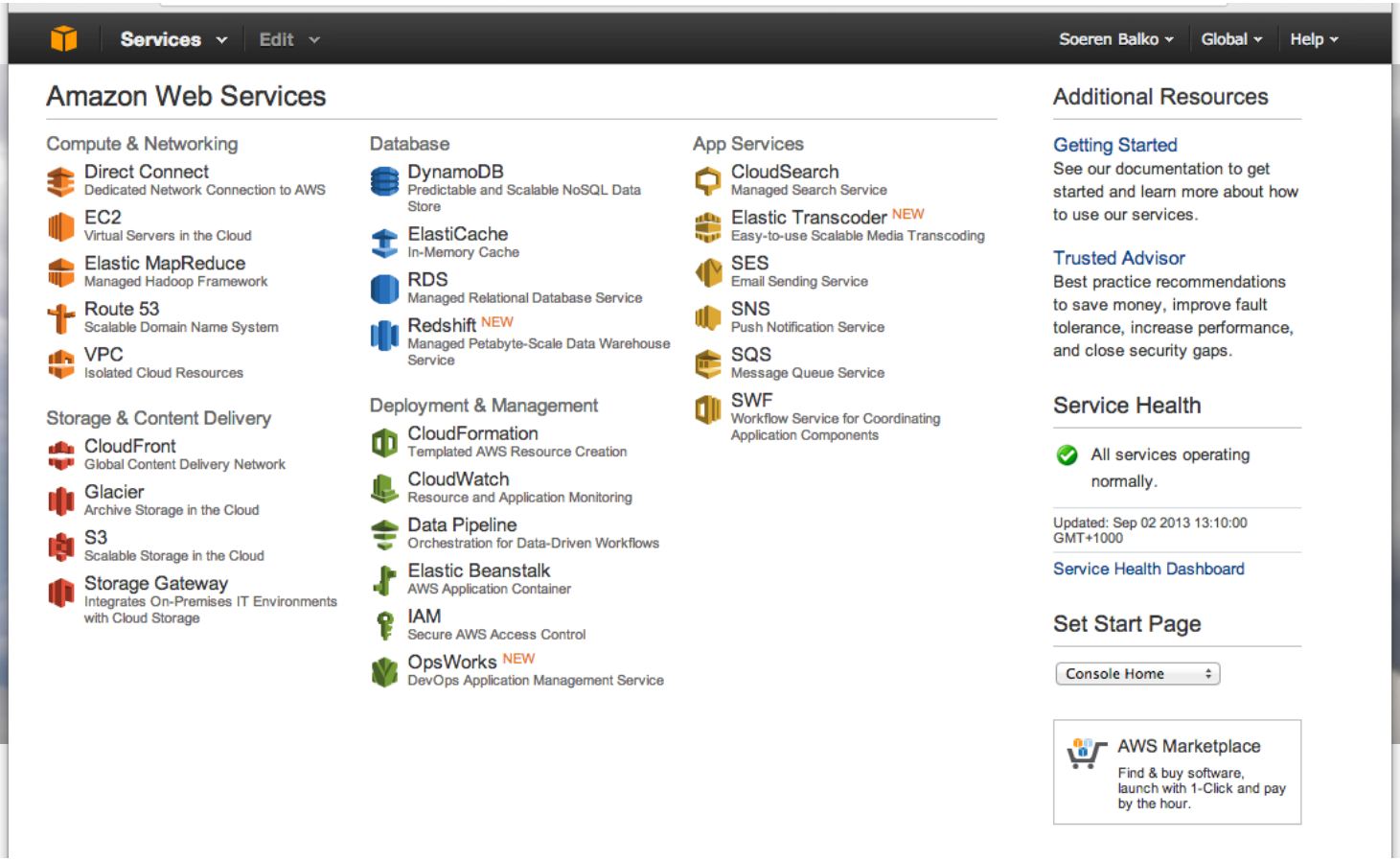
Types of Cloud Services

- The earliest cloud services were ‘close to the metal’
- Service offerings have become far richer and varied
- We will consider three standard service categories:
  - IaaS – Infrastructure as a Service
  - PaaS – Platform as a Service
  - SaaS – Software as a Service
- Later we will talk about containers and CaaS

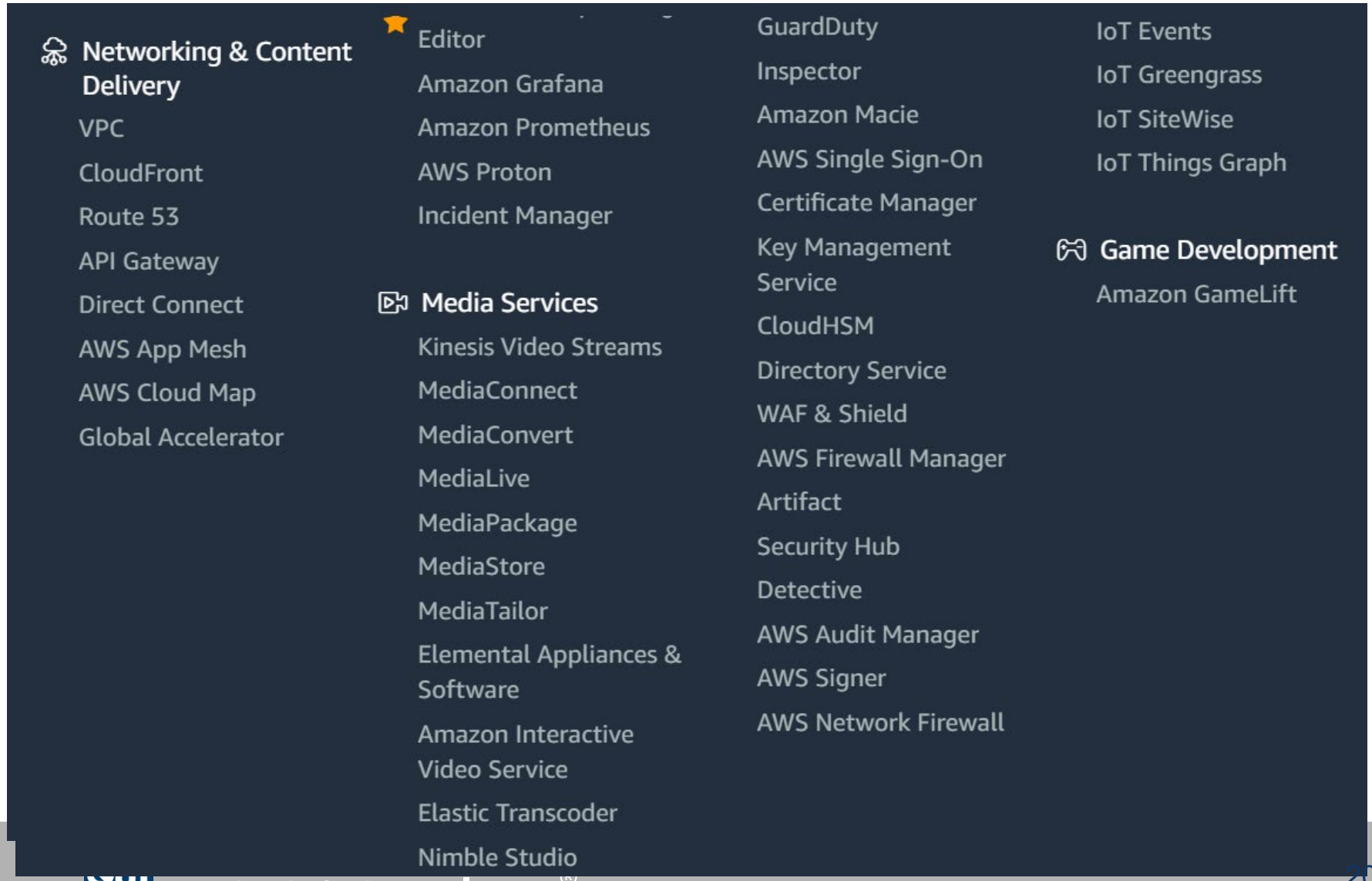
XaaS – the Cloud Service Hierarchy

Infrastructure as a Service	Platform as a Service	Software as a Service
• Virtual Machine + (Linux) OS • Pick your size • Storage – entity storage, SQL, NoSQL, archive • Pick your favourite DB	• Not just the OS • Pre-configured software stack on each VM • Managed provisioning and scaling	• Application hosting in the Cloud • Subscription-based • All hosting and management is done for you – for a fee • Salesforce is the usual example here

Services: AWS 2014



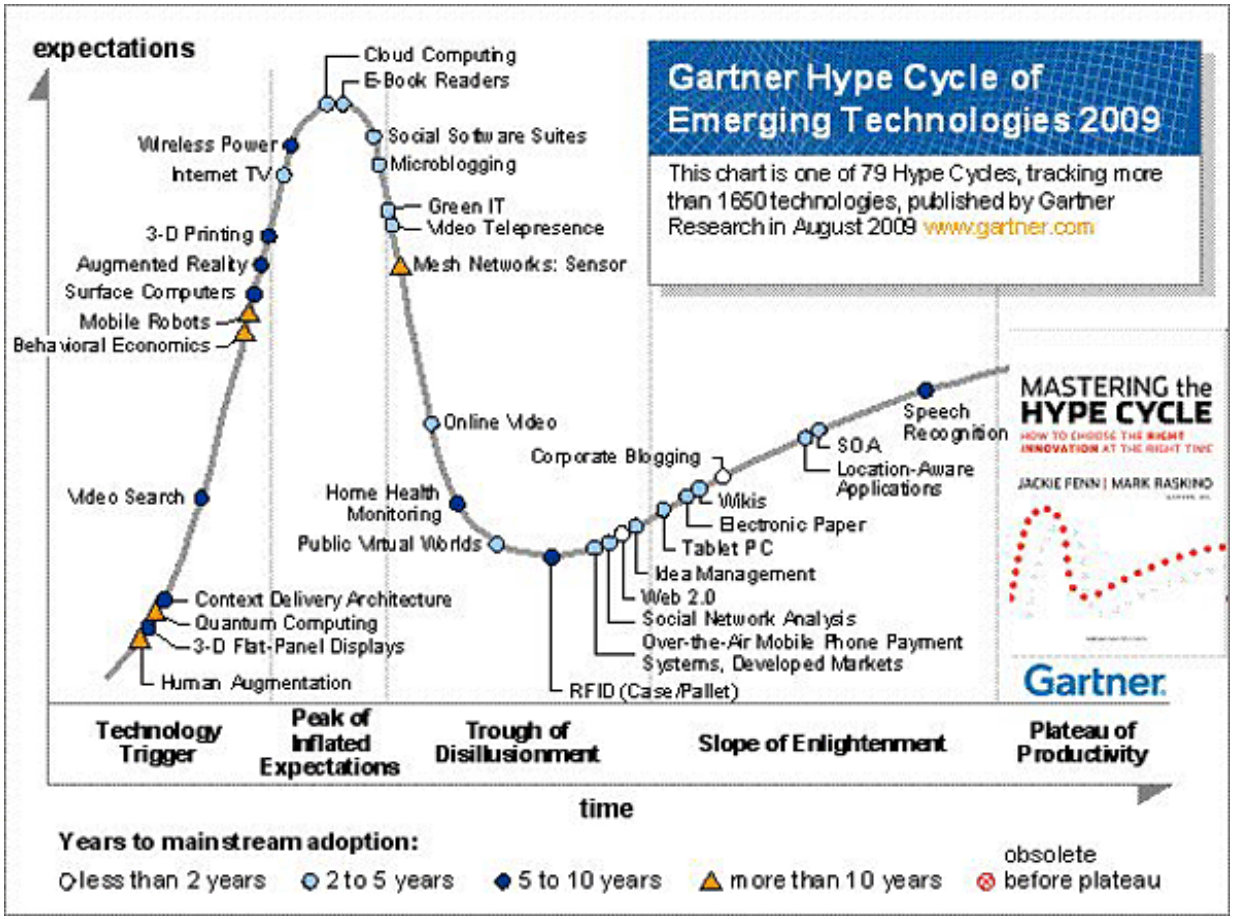
Services: AWS 2021



There might have been some hype

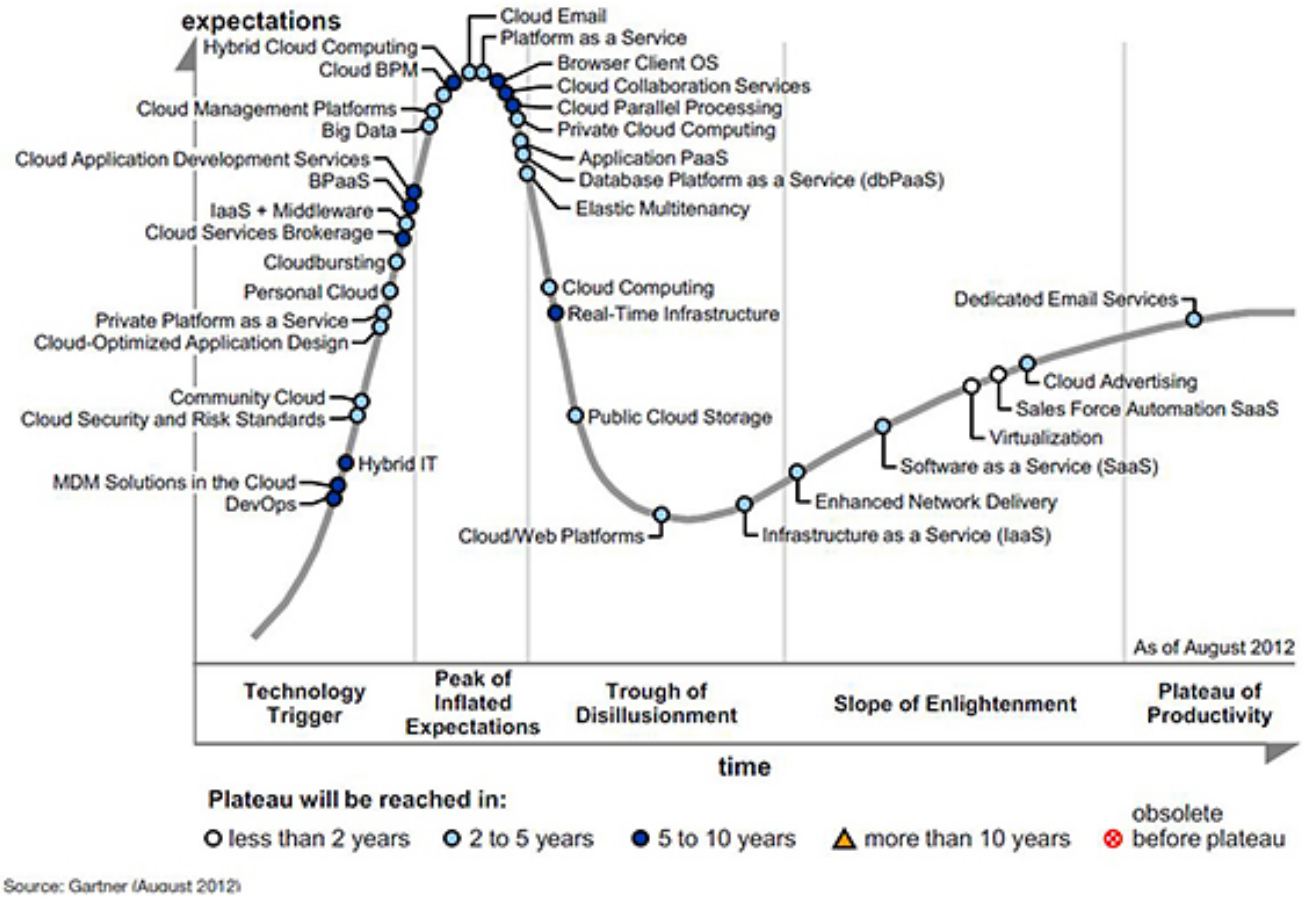
HOW IT BECAME CLOUDY

Cloud in 2009



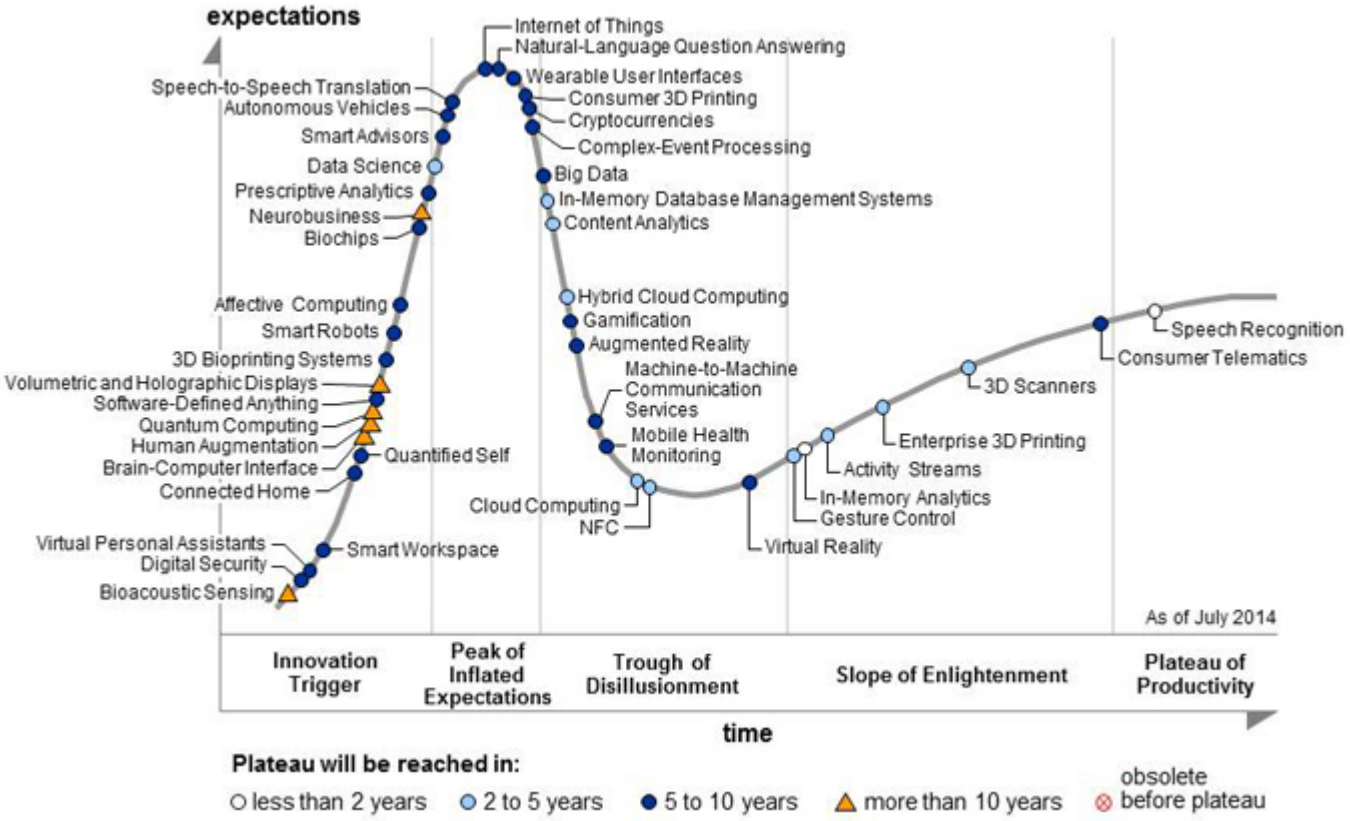
The Progress of Cloud

Figure 1. Hype Cycle for Cloud Computing, 2012



Source: Gartner (August 2012)

And more (2014)



Probably the last year for which it is worth looking at Cloud as such