

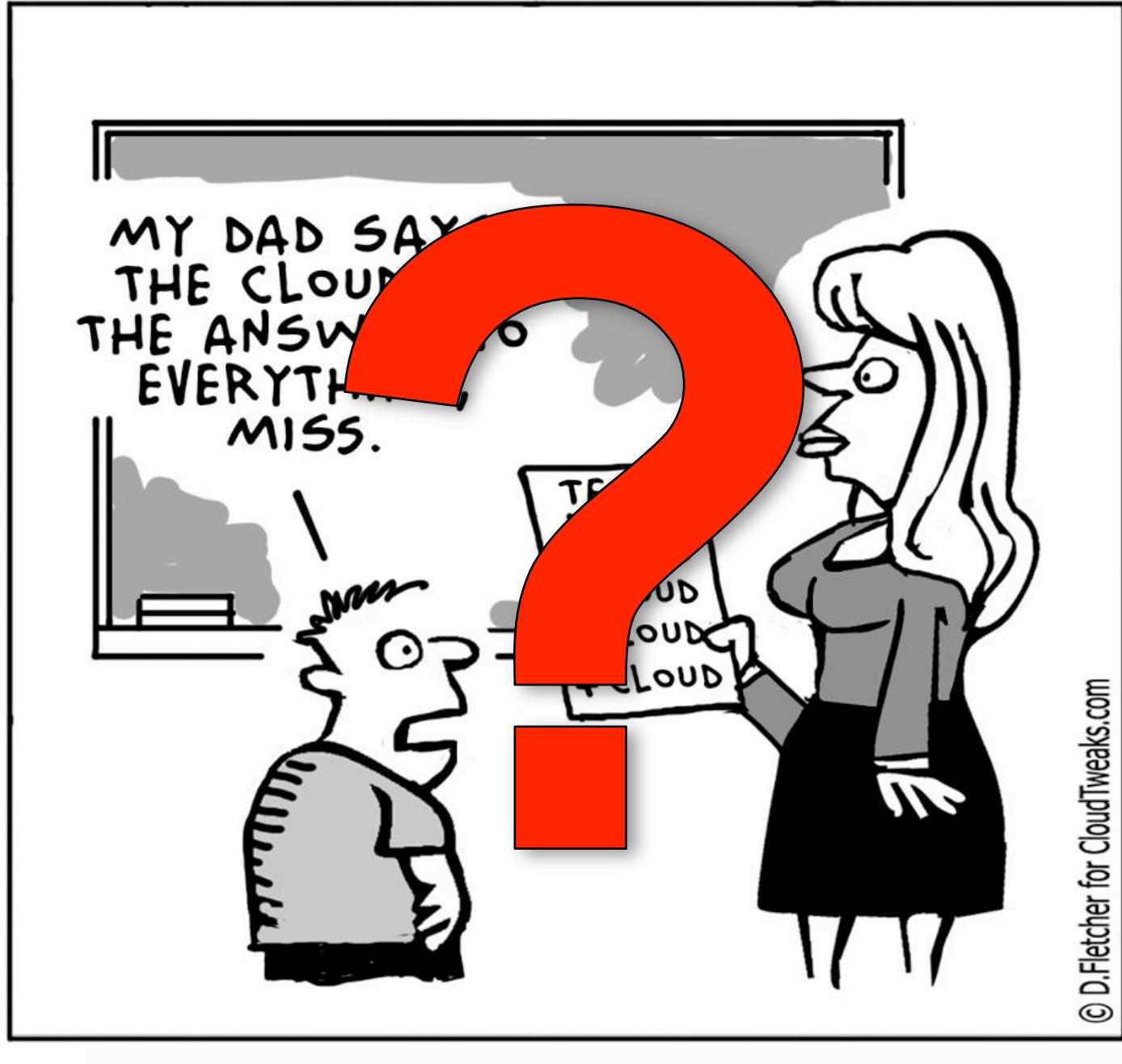
# *Cluster and Cloud Computing – Lecture 1*

Professor Richard O. Sinnott

Director, eResearch  
University of Melbourne

2<sup>nd</sup> March 2017

[rsinnott@unimelb.edu.au](mailto:rsinnott@unimelb.edu.au)



© D.Fletcher for CloudTweaks.com

*Director,  
eResearch  
University of  
Melbourne*

*CEO Own Company  
(real time  
systems/telecoms)*

*PhD Distributed  
Systems*

*Multiple PhD/MSc  
supervised in this  
area*

**Richard**



*Technical Director,  
Bioinformatics  
Research Centre  
University of Glasgow*

*MSc Software  
Engineering*

*Chair in Applied  
Computing Systems,  
University of  
Melbourne*

*BSc Theoretical  
Physics*

*Lecturer*

*Technical Director  
National e-  
Science Centre,  
University of  
Glasgow*

*Post-doc  
GMD Fokus  
Berlin*

*Distributed  
Systems  
Standards  
creator*



Luca



*MSc Statistics*

*IT Consultant for the  
World Bank, United  
Nations, European  
Union, Asian  
Development Bank*

*Committer at the  
Apache Software  
Foundation*

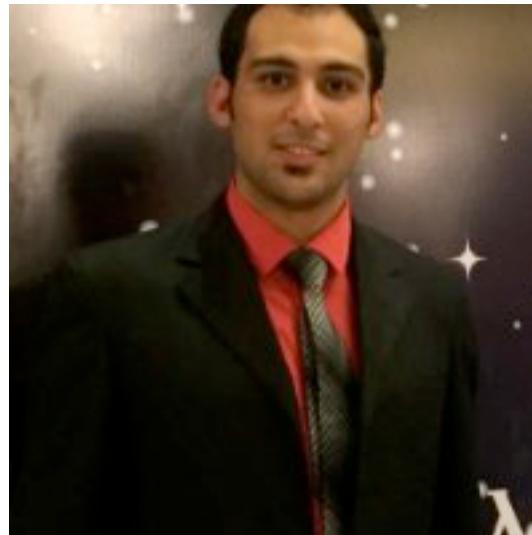
*GIS Software  
Engineer, ESRI*

*Data Architect,  
AURIN Project*

# Farzad

University of  
Melbourne  
Doctor of  
Philosophy (PhD)

*Internet of Things  
Big Data  
Cloud Computing*



Software  
Engineer  
(Big Cloud/  
Twitter!!!)

Sharif University of Technology,  
Master of Science,  
Information Technology

Shiraz University  
Bachelor of Science  
IT

## Completed

- National e-Science Centre (I, II, III)
- Dynamic Virtual Organisations for e-Science Education
- Biomedical Research Informatics Delivered by Grid Enabled Services
- GridNet, GridNet-2
- Grid Enabled Microarray Expression Profile Search
- Glasgow early adoption of Shibboleth
- Joint Data Standards Survey
- ESP-Grid
- HPC Compute cluster award // Sun industrial sponsorship
- OGC Collision
- OMII-Security Portlets // OMII-RAVE
- Integrating VOMS and PERMIS for Superior Grid Authorization
- NCeSS
- CESSDA PPP
- Pharming of Therapeutic RNA
- Grid Enabled Occupational Data Environment
- Towards an e-Infrastructure for e-Science Digital Repositories
- Grid enabled Biochemical Pathway Simulator
- Virtual Organisations for Trials and Epidemiological Studies
- A European e-Infrastructure for e-Science Repositories
- Modelling, Inference and Analysis for Biological Systems up to the
- Drug Discovery Portal
- Parliamentary Discourse
- Scots Words and Placenames
- Qvolution stress management survey system
- Advanced Grid Authorisation through Semantic Technologies Shir
- AlstromUK VRE
- Grid-enabled Virtual Organisations
- Clinical Streaming
- Enhancing Repos
- Proxy Credential A
- Scottish Bioinform
- Generation Scotla
- Breast Cancer Tis
- Data Management
- Meeting the Design
- EU FW7 AvertIT
- EU FW7 EuroDSL
- NeSC Research P
- NeSC Information
- ESF Network for Study of Adrenal Tumors
- Scottish Health Informatics Platform for Research (SHIP)
- National E-Infrastructure for Social Simulation (NeISS)
- EU R4SME Diagnosis of Parkinsons Disease (DiPAR)
- Automating River Pollution Detection (CAPIM)
- Endocrine genomics Virtual Laboratory (endoVL)
- DSDNetwork Australasia

## Project Portfolio

## On-Going

- EU European Platform for Study of Wolfram, Alstrom, Bardet Biedl (EuroWABB)
- Multicenter prospective study of biochemical profiles of monoamine-producing tumors (PMT Study)
- European Society of Hypertension Study on Pheo/PGL
- International DSD
- EU FW7 European Network for Study of Adrenal Tumors Cancer Research Platform (ENSAT-CANCER)
- VicHealth Health Indicators and Spatial Objective Data
- Jury Research Platform
- Research Infrastructure Network (AURIN)
- Study portal
- Study of environmental factors on onset of T1D
- es Data Network (ADDN)
- mann-Pick A, B and C Registry
- Data Journalism in the Big Data Era
- ied 18F-fluorodeoxyglucose positron emission tomography and 123I-imaging for Adrenal Neoplasia
- nics Health Alliance (variant DB)
- ncryption/Decryption and Secure Deletion
- on of Pancreatic Beta Cells
- eric Physics and Climate Research)
- nd Urban Environments
- cs-based strategies for improved diagnosis and treatment of endocrine
- umption database and mobile app
- tistics Research Environment
- Clinical Research Network
- ustralian Adrenal Alliance
- earch Software Solutions (PRESS)
- Mobile applications for the Environmental Determinants of Islet Autoimmunity



# Examples used throughout

- Twitter data analytics for business
- Mobile Applications for Patients with Neuroendocrine Tumours
- Systems Genomics Support Platform
- SWARM: Smartly-aggregated Wiki-style ARgument Marshalling (SWARM)
- ORCA Cognitive Assessment Platform
- VicSpin Victoria-wide Flu Surveillance System
- ElectraNet LIDAR
- VectorNZ Lidar

# Course Contents

- Lectures 1 & 2 – 2<sup>nd</sup> March
  - Information Session & How we got here (Distributed Systems, Grid...)
    - Richard Sinnott
- Lectures 3 & 4 – 9<sup>th</sup> March
  - Domain Drivers – tour of some big data projects
    - Richard Sinnott
- Lectures 5 & 6 – 16<sup>th</sup> March
  - Parallel Systems, Distributed Computing and HPC/HTC
    - Richard Sinnott
- Lectures 7 & 8 – 23<sup>rd</sup> March
  - HPC @ UniMelb and Practicalities of HPC/HTC
    - Richard Sinnott, [Lev Lafayette](#) & [Farzad Khodadadi](#)
    - [Linux and HPC practicalities](#)
    - [Using mpi4py workshop](#)

Programming Assignment handed out (24<sup>th</sup> Mar – due 6<sup>th</sup> April)

# Course Contents

- Lectures 9 & 10 – 30<sup>th</sup> March
  - Cloud Computing – Programming Clouds: Getting to grips with NeCTAR
    - Richard Sinnott & Farzad Khodadadi
    - Introduction to Cloud Computing
    - Getting to grips with OpenStack/NeCTAR
    - Scripting for the Cloud (Introduction to Boto & Ansible demonstration) workshop

Second Programming Assignment handed out (7<sup>th</sup> April – 10<sup>th</sup> May)

- Lectures 11 & 12 – 6<sup>th</sup> April
  - Big Data and Related Technologies
    - Luca Morandini (Data Architect, AURIN)
    - Big Data V-challenges, Cap Theorem and noSQL technologies
    - CouchDB workshop
- Lectures 13 & 14 – 13<sup>th</sup> April
  - Service-oriented architectures & Other Things Needed for Assignment II
    - Farzad Khodadadi & Luca Morandini
    - SOA & SOAP vs ReST
    - Hands-on examples of coding/demonstrating SOAP, ReST
    - Code versioning systems and GitHub (Git) workshop

Easter Break 17th-21st April

# Course Contents... ctd

- Lectures 15 & 16 – 27<sup>th</sup> April
  - Big Data Analytics
    - Luca Morandini
    - Big Data Technologies – Hadoop, HDFS, Spark, ...
    - Hadoop cluster on Cloud and practical application workshop
- Lectures 15 & 16 – 4<sup>th</sup> May
  - Cloud Underpinnings and Other Things
    - Richard Sinnott & Farzad Khodadadi & Luca Morandini
    - Virtualisation background
    - Compare and Contrast EC2 with NeCTAR Research Cloud
    - Microhosting Environments (Docker)
    - Discussion of 2<sup>nd</sup> Programming Assignment
- Lecture 17 & 18 – 11<sup>th</sup> May
  - Really Big Compute and You
    - Richard Sinnott & Christopher Samuels & You
    - VLSCI – Christopher Samuels
    - “Some” teams randomly chosen to present their assignment II
    - 15minutes each

# Course Contents... ctd

- Lecture 21 & 22 – 18<sup>th</sup> May
  - Security and Clouds & demonstration of assignments
    - Richard Sinnott & You
    - “Some” more teams randomly chosen to present their assignment II
    - 15minutes each
- Lecture 23 & 24 – 25<sup>th</sup> May
  - Subject Review and Working Through Past Papers
  - Feedback and SES
    - Richard Sinnott

# At the end of the course....

- You will...
  - Understand more about the history of cluster and Cloud computing and the current state of the art
  - Know more of the domain drivers that are shaping this area – especially the current flavour of the month “big data”
  - Understand more on parallel systems, multi-core software development
  - Be able to use HPC/HTC systems
  - Be able to use Cloud resources
  - Be able to develop applications running on the Cloud
  - Learning about next generation data management systems
  - Be savvy with web based systems development
  - Work on a non-trivial software system development as part of a team (=how it nearly always works in the real world!)
  - Learn more the pulse of cities!
  - Have software experiences that are in great demand  
    >30 of my ex-students have jobs in this space (reference machine!!!)
  - Have visited a data centre

# Lecture Slides

- On the web before lecture
- They may be updated slightly before the lecture to reflect recent developments
  - Or cover materials that folk want more details on from previous lectures
- Other sources of materials
  - Key papers
  - Interesting articles
  - Snippets from the web
  - Crib sheets for Cloud Computing
  - Crib sheets for HPC

# Course Assessment

- Assignments
  - During semester worth 50%
  - Expected to take about 40+ hours
- Written examination
  - A written examination (two hours) at the end of the semester worth 50%
- All components must be completed satisfactorily to pass the subject
  - At least 50% in all assignments + exam

# Assessments

- Assignment 1 – 10 marks
  - HPC exercise (search/analyse large text file)
- (Team) Assignment (40 marks)
  - Twitter analytics on the Cloud
  - Team-based approach
  - Self organise vs be organised
    - Importance of being a team player!
    - Peer review...
- Exam – 50 marks
- Note
  - Importance of original work
    - Collusion “seriously frowned” upon
    - Use of Turnitin & MOSS

# Computational Resources

- Departmental Computing Resources:
  - You all have access to PCs in the lab with Linux and Eclipse installed
  - Any local issues then talk to CIS technical staff
- HPC systems
  - SPARTAN cluster – general purpose cluster for UniMelb
    - Provisioned last year
    - Reasonable grunt, but... UniMelb!
- Cloud resources
  - NeCTAR research cloud – [www.nectar.org.au](http://www.nectar.org.au)
  - 30,000 servers across Australia (availability zones)
    - Small (4Gb), medium, large, xlarge, xxlarge (64Gb)
    - MUCH more on NeCTAR later

# Student Evaluation Survey

- (Told I should provide this info!!!)

Subject:	COMP90024_U2016SM11 - Cluster and Cloud Computing	Department:	Comp & Inf
		Resp. Rec'd / Expected:	78 / 123

Standard Questions		COMP90024_U2016SM11 -								
		Responses (%)					Subject			
		[SD]	[D]	[N]	[A]	[SA]	S.D.	Std. Err.	N	Mean
Q1	Overall, this subject has been intellectually stimulating	1	1	9	41	47	.79	.09	78	4.3
Q2	Overall, this subject has been well co-ordinated	1	4	10	42	43	.87	.10	77	4.2
Q3	Overall, this subject has been supported by useful learning resources	1	5	10	40	44	.91	.10	78	4.2
Q4	Overall, this subject has been well-taught	0	5	17	42	36	.85	.10	78	4.1
Q5	Focusing on my own learning in this subject, I have been required to work at a high standard	1	1	4	32	62	.75	.08	78	4.5
Q6	Focusing on my own learning in this subject, I found the assessment tasks useful in guiding my study	1	1	4	42	51	.74	.08	78	4.4
Q7	Focusing on my own learning in this subject, I received valuable feedback on my progress	3	4	5	38	51	.93	.11	77	4.3
Q8	Focusing on my own learning in this subject, I learnt new ideas, approaches and/or skills	1	1	3	35	60	.73	.08	78	4.5
Q9	Focusing on my own learning in this subject, I learnt to apply knowledge to practice	1	1	4	36	58	.75	.08	78	4.5
Q10	Focusing on my own learning in this subject, I have been part of a group committed to learning	1	3	4	35	57	.80	.09	77	4.4

Responses: [SD] Strongly Disagree=1 [D] Disagree=2 [N] Neither=3 [A] Agree=4 [SA] Strongly Agree=5

Questions?

# The Buzz

## Top 10 Strategic Technology Trends for 2013

1. Mobile Devices Battles
2. Mobile Applications & HTML5
3. Personal Cloud
4. Internet of Things
5. Hybrid IT & Cloud Computing
6. Strategic Big Data
7. Actionable Analytics
8. Mainstream In-Memory Computing
9. Integrated Ecosystems
10. Enterprise App Stores



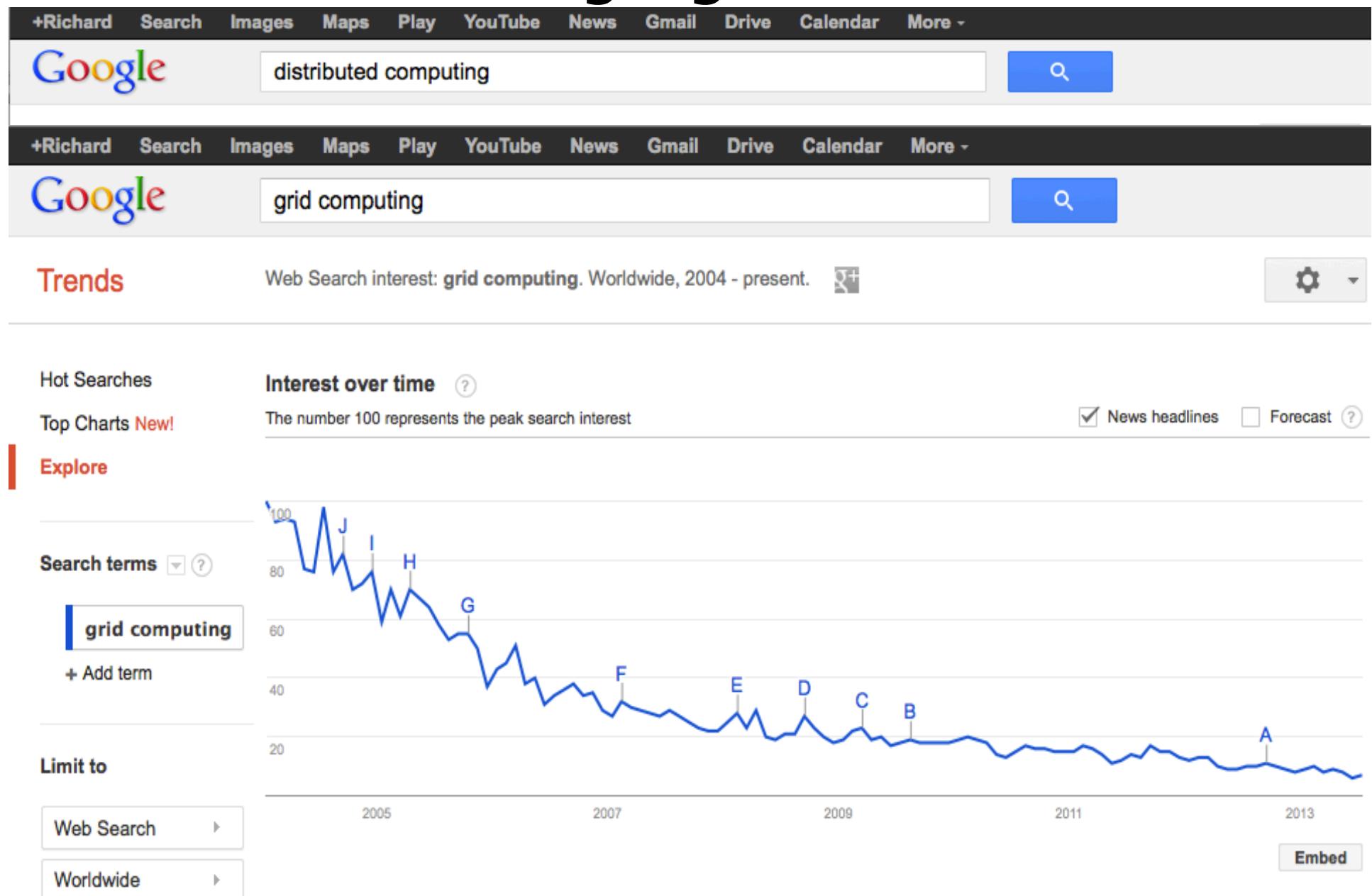
Big Data Drives Rapid Changes in Infrastructure and \$232 Billion in IT Spending Through 2016

12 October 2012

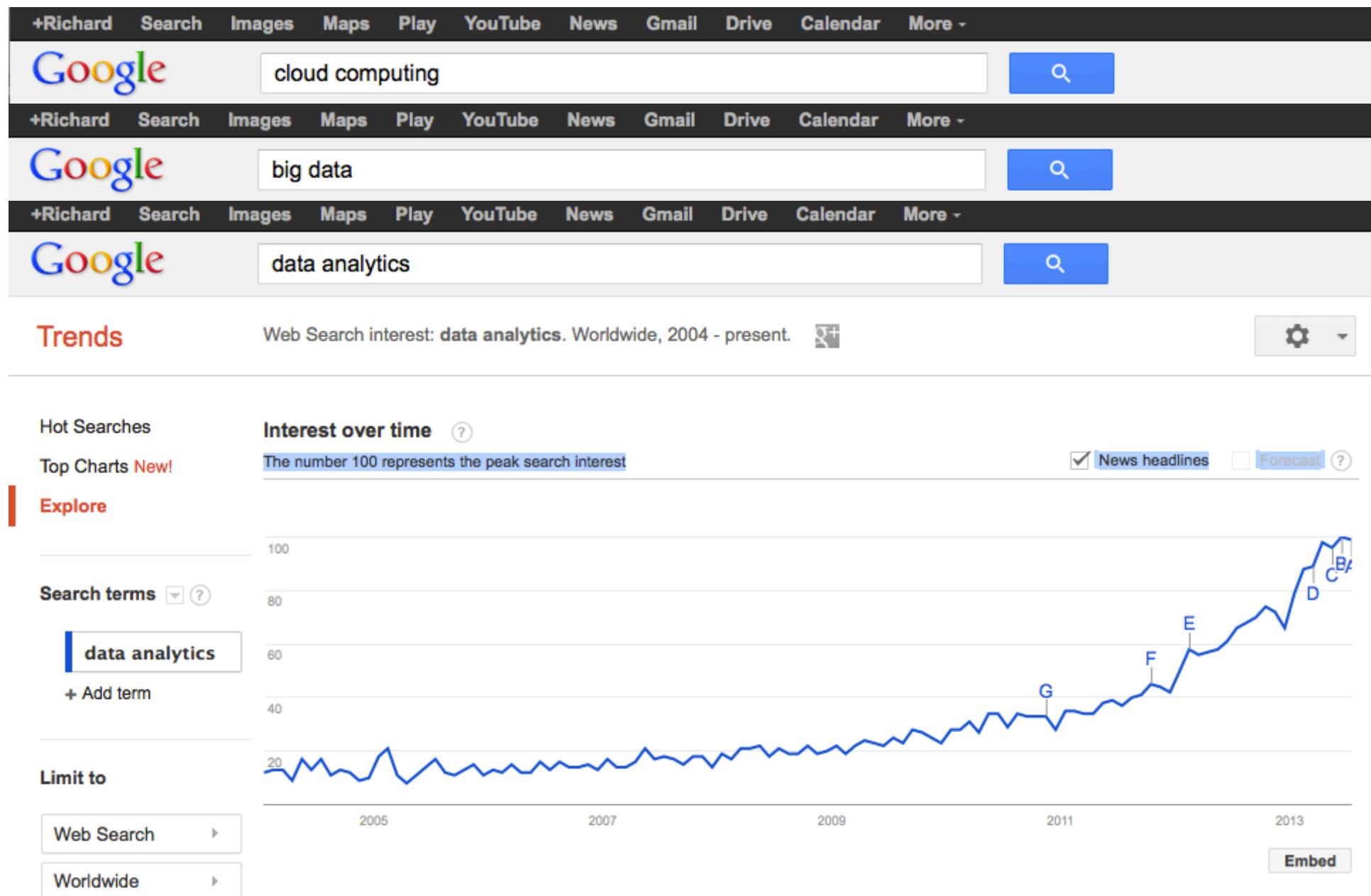
Mark A. Beyer John-David Lovelock Dan Sommer Merv Adrian

Big data has become a major driver of IT spending. The benefits to organizations for adding big data to their information management and analytics infrastructure will force a more rapid cycle of replacing existing solutions.

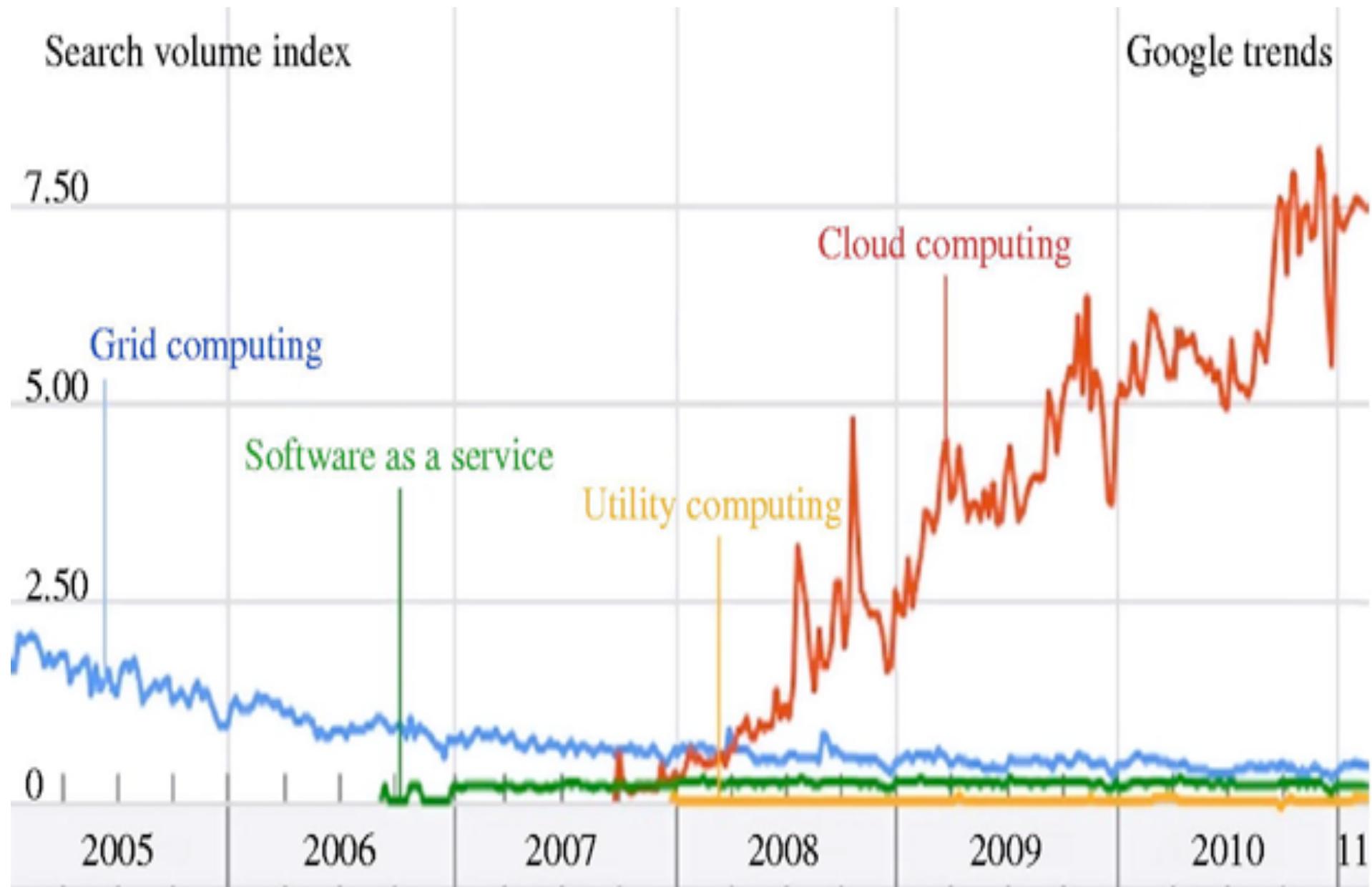
# The not so long ago buzz...



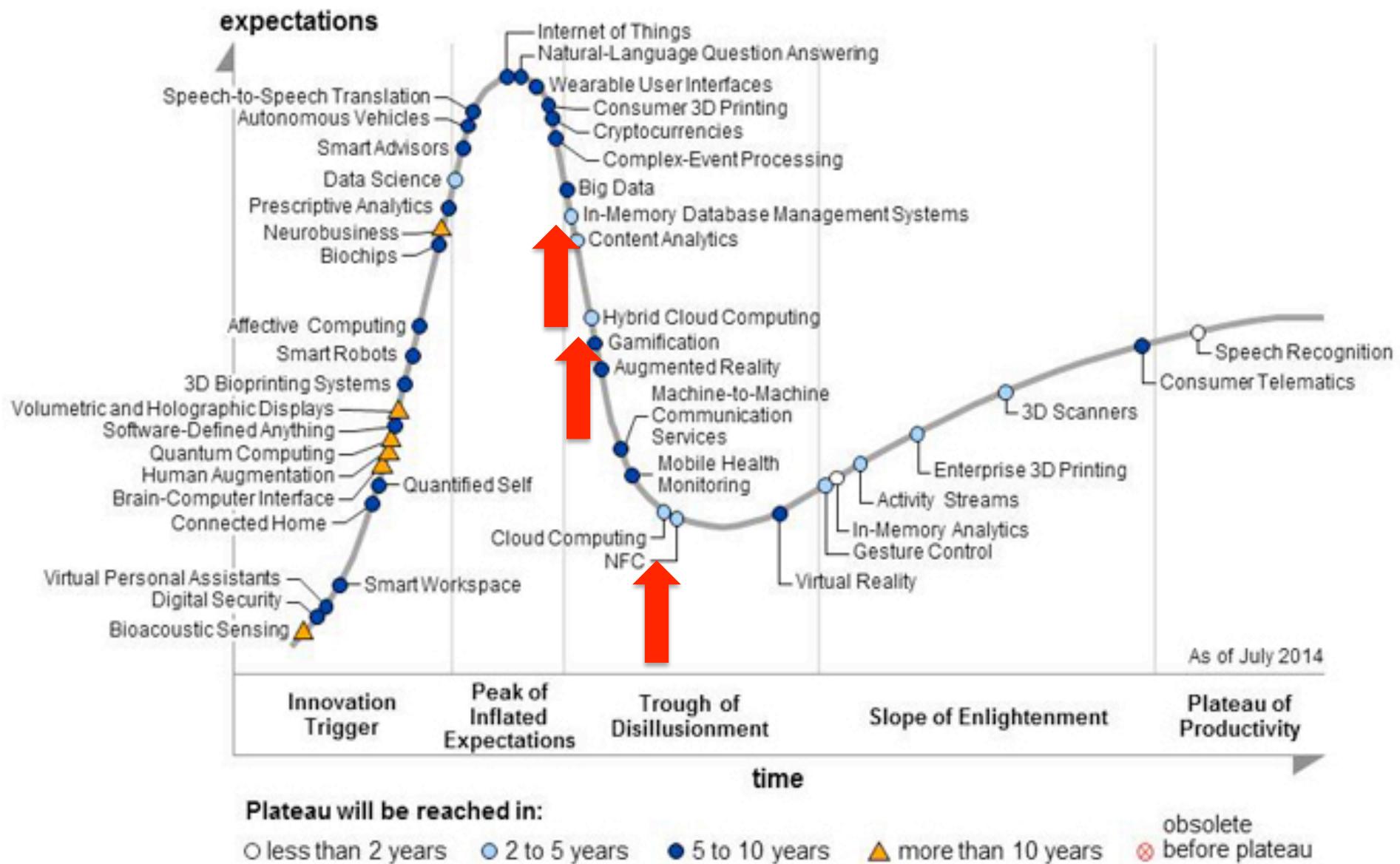
# The latest buzz...



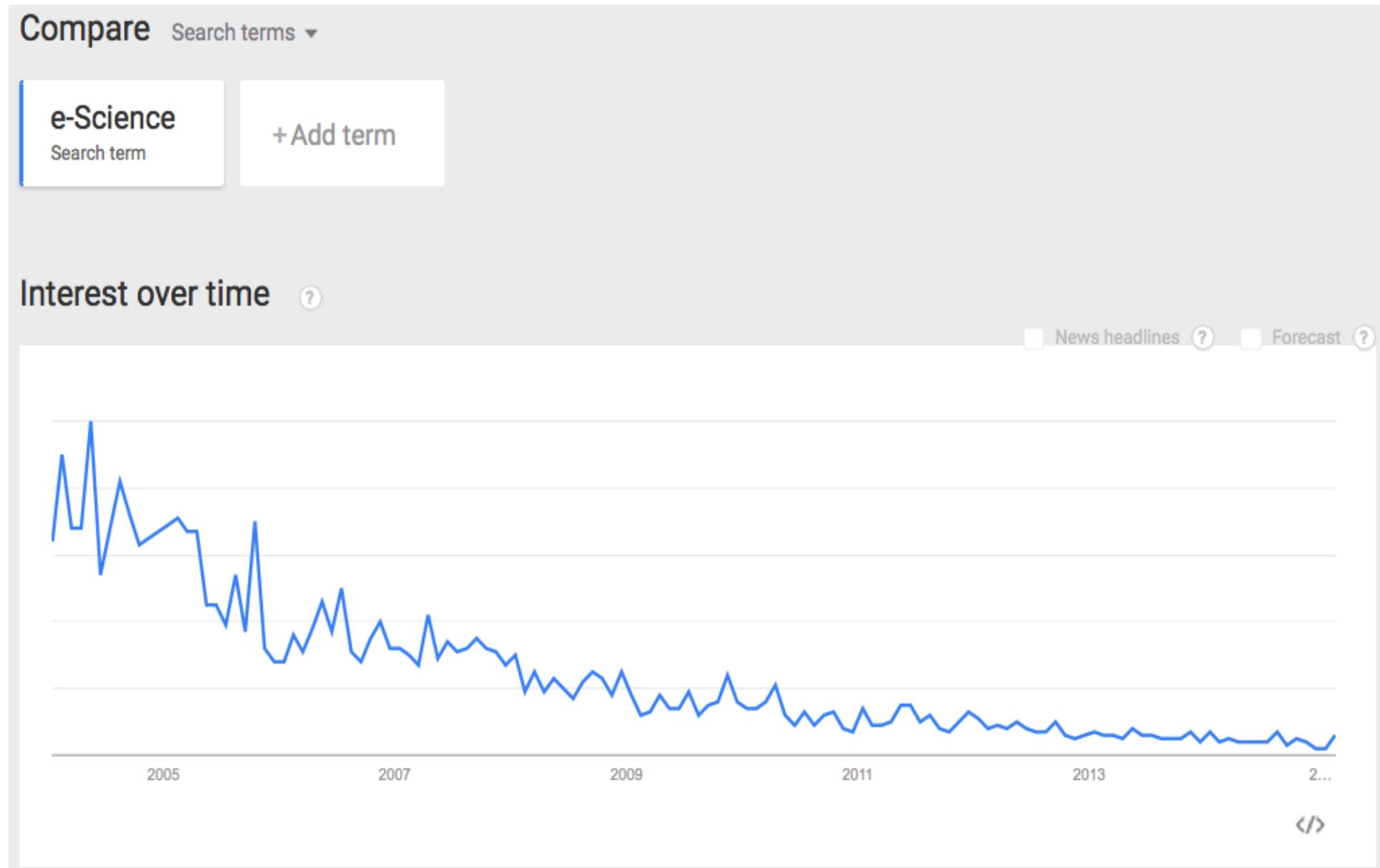
# The cycles we go through...



# The Hype Cycle... (Gartner 2015)



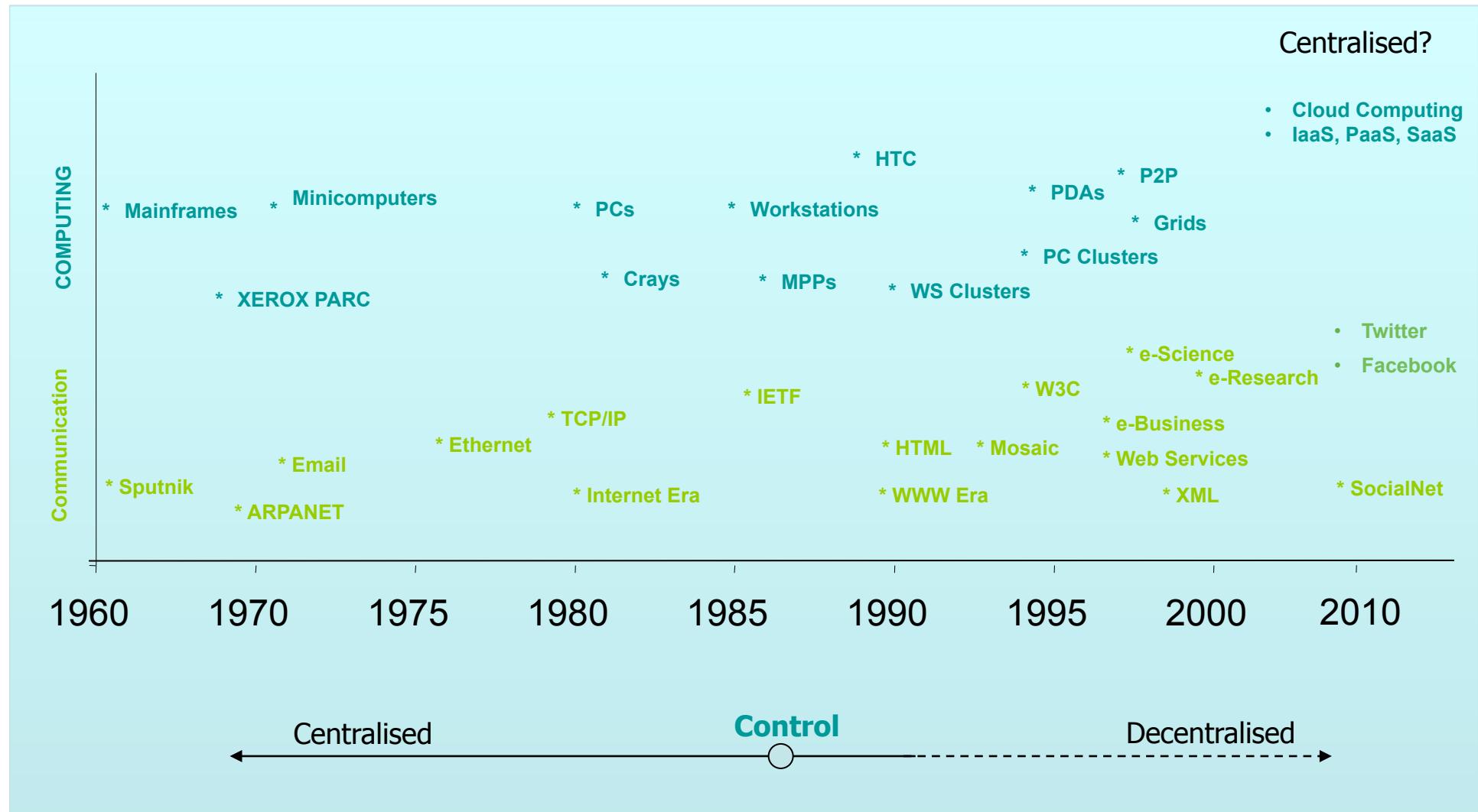
# Not just IT trends...



# Lessons...

- Don't be fooled by the hype
  - Marketing and business drivers
- The basic principles of software engineering persist
  - NeSC -> MeG experiences
    - Data processing,
    - Distributed systems,
    - Security ...
- Learning from experiences is key
  - But obviously working on/with the bleeding edge can do no harm (for grants or on CVs)

# Computing and Communication Technologies (r)evolution: 1960-...!



# What's in a name...

- Centralised system
  - *Single physical (centralised) system. All resources (processors, memory and storage) fully shared and tightly coupled within one integrated OS*
- Parallel system
  - *All processors either tightly coupled with centralised shared memory or loosely coupled with distributed memory. Inter-process communication through shared memory or through some form of message passing*
- Distributed system
  - *Multiple autonomous computers with their own private memory, communicating through some form of message passing over a computer network*

# Cloud Computing....

- ...is a colloquial expression used to describe a variety of different types of computing concepts that involve a large number of computers that are connected through a real-time communication network (typically the Internet). *Cloud computing is a jargon term without a commonly accepted non-ambiguous scientific or technical definition.* In science, cloud computing is a synonym for distributed computing over a network and means the ability to run a program on many connected computers at the same time. The popularity of the term can be attributed to its use in marketing to sell hosted services in the sense of application service provisioning that run client server software on a remote location.

# Cloud Computing....

- ...Proponents claim that cloud computing allows companies to avoid upfront infrastructure costs, and focus on projects that differentiate their businesses instead of on infrastructure. Proponents also claim that cloud computing allows enterprises to get their applications up and running faster, with improved manageability and less maintenance, and enables IT to more rapidly adjust resources to meet fluctuating and unpredictable business demand. Cloud providers typically use a "pay as you go" model. This can lead to unexpectedly high charges if administrators do not adapt to the cloud pricing model.

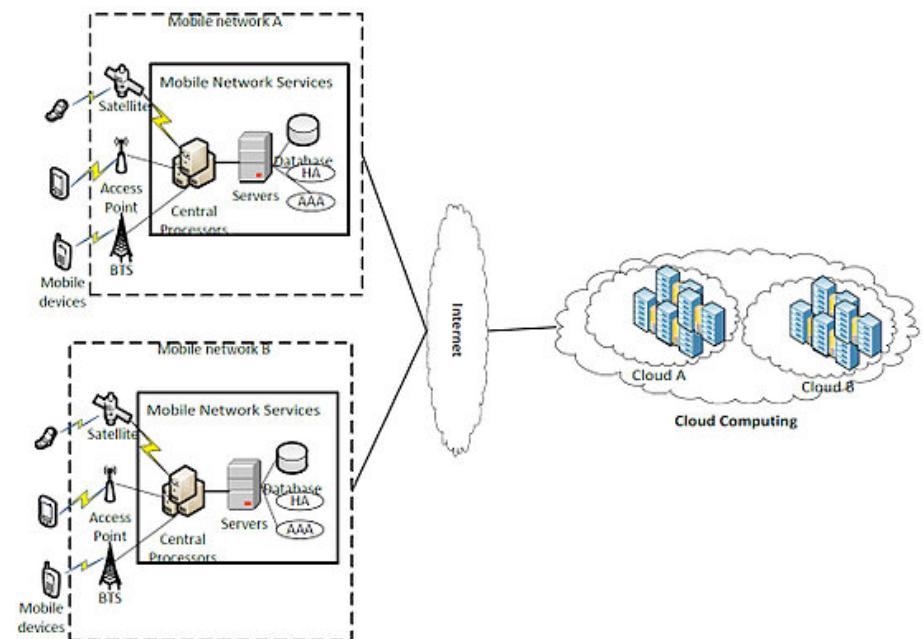
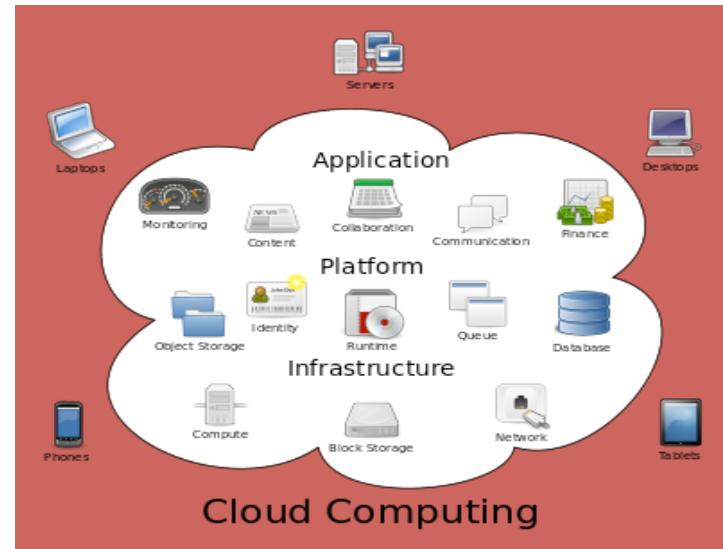
[Wikipedia 2016](#)

# Cloud Characteristics

- Five essential characteristics:
  - *On-demand self-service*. A consumer can provision computing capabilities as needed without requiring human interaction with each service provider.
  - *Networked access*. Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous client platforms.
  - *Resource pooling*. The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model potentially with different physical and virtual resources that can be dynamically assigned and reassigned according to consumer demand.
  - *Rapid elasticity*. Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly upon demand.
  - *Measured service*. Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service.
    - National Institute of Standards and Technology

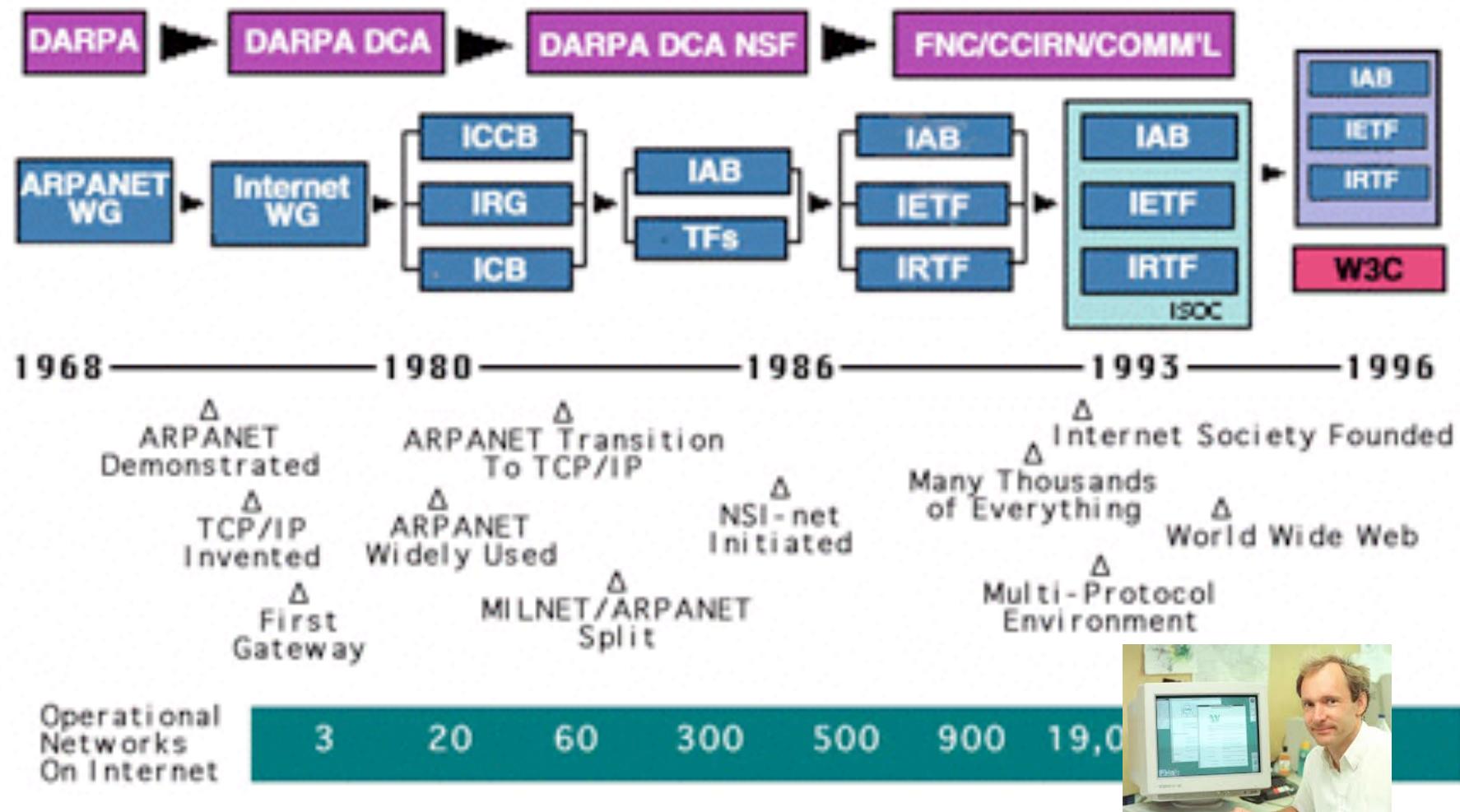
# Cloud Computing....

- Flavours
  - Compute clouds
    - Amazon Elastic Compute Cloud
    - Azure
    - ...
  - Data clouds
    - Amazon Simple Storage Service
    - Google docs
    - iCloud
    - Dropbox
    - ...
  - Application clouds
    - App store
    - Virtual image factories
    - Community-specific
  - Private, public, hybrid, mobile, health, ... clouds
    - ...



Break...?

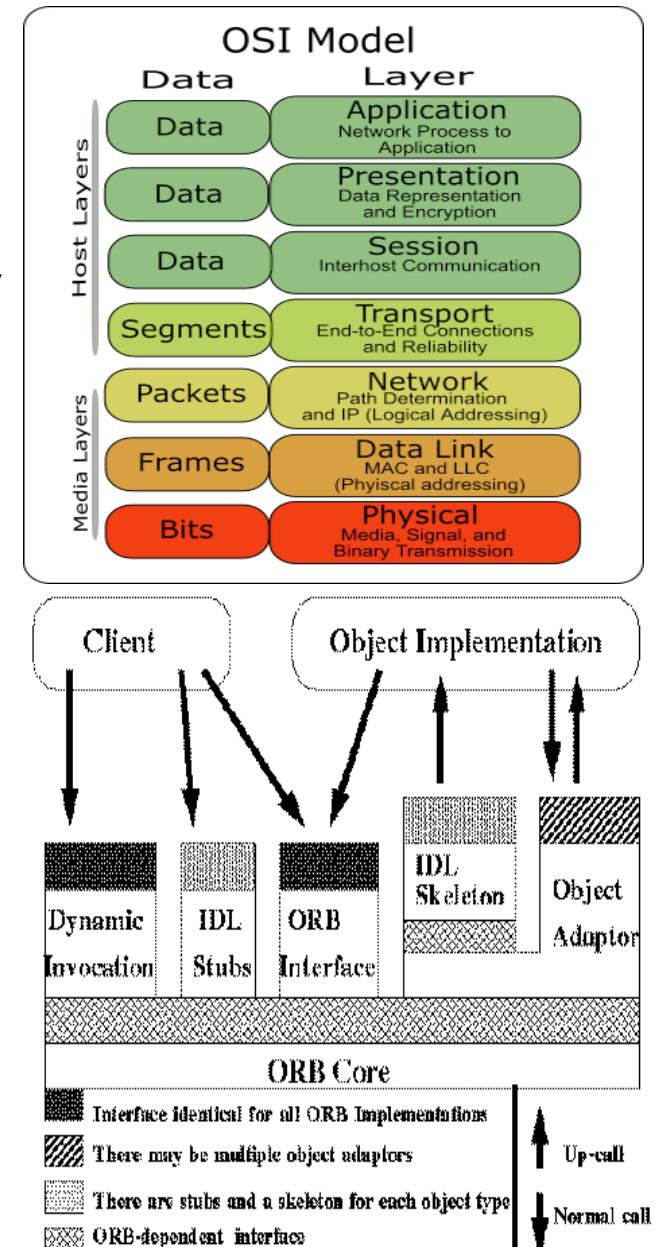
# In the beginning....



- <http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet>

# Distributed Systems - A Very Brief History

- Once upon a time we had standards
  - With very detailed conformance, consistency and compliance demands
    - Services, protocols, inter-operability, ...
- Then we had more standards
  - Open distributed processing
  - With slightly less rigorous compliance demands
    - OMG Common Object Request Broker Architecture (CORBA)
    - Distributed Computing Environment
    - Multiple technologies
      - Client server, remote procedure call, ...



# Key distributed systems focus mid-90s

- Transparency and heterogeneity of computer-computer interactions
  - finding/discovering resources (trader!),
  - binding to resources in real time,
  - run time type checking,
  - invoking resources
  - ...

## Computer-computer interaction focus

Client.java

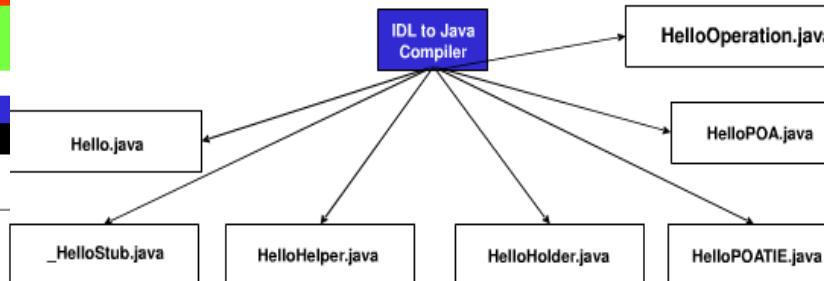
```
public class Client {  
    public static void main(String[] args) {  
        String iorFile = "week1.ior";  
        try {  
            File file = new File(iorFile);  
  
            if (!file.exists()) {  
                System.err.println("Error: File " + iorFile + " does not exist!!");  
                System.exit(1);  
            }  
            ORB orb = ORB.init(args, null);  
  
            BufferedReader reader = new BufferedReader(new FileReader(file));  
            String string_ref = reader.readLine();  
            reader.close();  
            org.omg.CORBA.Object obj = orb.string_to_object(string_ref);  
  
            Hello server = HelloHelper.narrow(obj);  
            String response = server.getTime();  
            System.out.println(response);  
            ...  
        }  
    }  
}
```

(Simplified)

```
public static void main(String[] args) {  
    String iorFile = "week1.ior";  
    org.omg.CORBA.ORB orb = org.omg.CORBA.ORB.init(args, null);  
  
    // Get reference to the root POA  
    POA rootPoa =  
        ORBHelper.narrow(orb.resolve_initial_references("RootPOA"));  
    // Activate the POA Manager - else no requests will be processed!  
    rootPoa.the_POAManager().activate();  
  
    HelloImpl servant = new HelloImpl();  
  
    // Create a CORBA reference for the servant  
    org.omg.CORBA.Object obj = rootPoa.servant_to_reference(servant);  
  
    PrintWriter writer = new PrintWriter(new FileWriter(iorFile));  
    writer.println(orb.object_to_string( obj ));  
    writer.flush();  
    writer.close();  
  
    // OK, now just sit back and wait for the action...  
    orb.run();  
    ...  
}
```

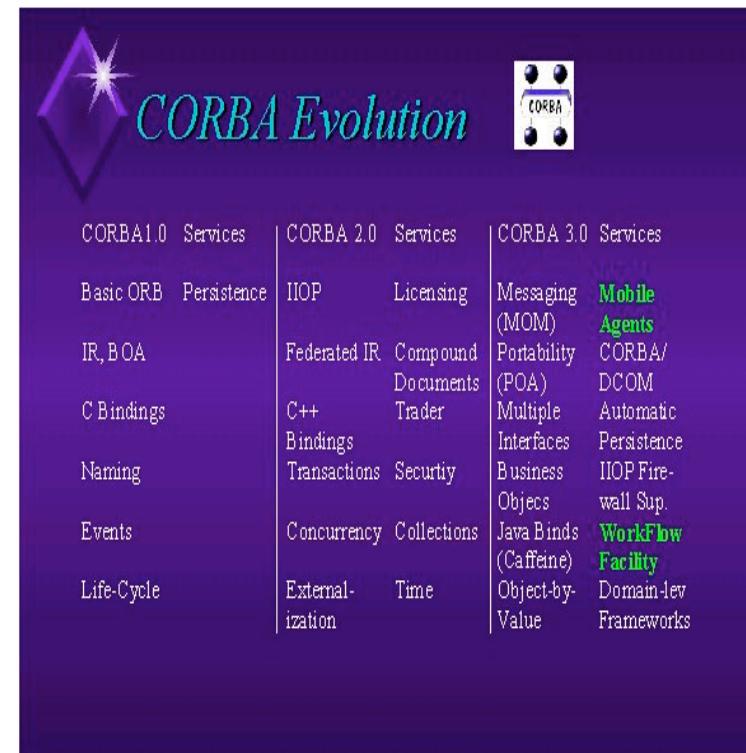
## 2. Compiling Hello IDL

\$ idl -ir -d generated hello.idl



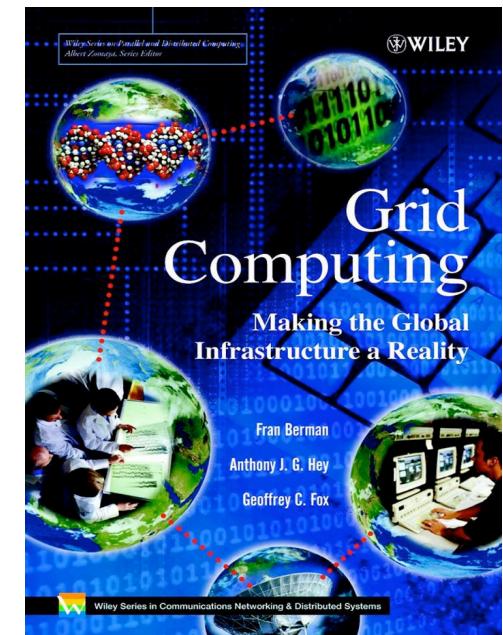
# Challenges of earlier distributed (networked) system implementations

- Complexity of implementations
  - Middleware bloat and lock-in
  - Vision and challenges of reality
- Vendor specific solutions
  - AnsaWare
  - IBM Distributed SOM / Component Broker
  - Microsoft Network OLE
  - IONA Orbix
  - Visigenic Visibroker
  - BEA ObjectBroker
  - HP ORBPlus and Distributed SmallTalk
  - Expersoft PowerBroker
  - Other ORB's
    - Less mature standards
- Scale of the problem area
  - Telecoms, banking, ...
  - The growth (boom!) of the web



# Distributed Systems History...ctd

- Enter the web era
  - My first ftp 1993 put/get files to/from Australia
  - Then the web pretty much exploded
- Peer-peer processing
  - File sharing ...
- Scaling of...
  - machines,
  - people,
  - domains of application
- Grid computing
  - From computer-computer focus
  - To organisation-organisation focus



# e-**Research** and the Grid

'e **Research** is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it.'

E- **Research** will change the dynamic of the way science is undertaken.'



John Taylor  
Director General of Research Councils  
Office of Science and Technology

## Grid is infrastructure used for e-Science

- Power Grid Metaphor: compute and data resources on demand
- Major investment by UK Govt (£250m+) to realise this vision
- Across ALL Research Councils (+ EU + industry + ...)
  - Cyber-infrastructure, European Grid Initiative, Superscience,

From presentation by Tony Hey

# The Grid Metaphor



# Overcoming heterogeneity...

## Guide to International Plugs & Sockets



[www.interpower.com](http://www.interpower.com)  
formerly Panel Components Corporation

P.O. Box 115, Oskaloosa, IA 52577 (USA)  
Call: (641) 673-5000 Fax: (641) 673-5100  
E-mail: [info@interpower.com](mailto:info@interpower.com)

Toll-free (U.S./Can./P.R./V.I.)  
Call toll-free: (800) 662-2290  
Fax toll-free: (800) 645-5360

Rev. 8-05

### Continental European

For use in Germany, Austria, Finland, the Netherlands, Norway, Sweden, France/Belgium, and other countries utilizing the CEE 7 standard. (Note: Sweden requires use of socket with separate mounting bracket. France/Belgium require use of different socket types.) The CEE 7/7 is non-polarized.



For black (88010920) or brown (88010620) plug

<b>Plug with Strain Relief</b>
Part Number <b>88010801</b>
Type CEE 7/7
Color Gray
Rating 16A/250VAC
Approvals

### Europlays

Interpower Europlays are wirable. Once shut, they cannot be re-opened. For use in Germany, Austria, Finland, the Netherlands, Norway, Sweden, France/Belgium, and other countries utilizing the CEE 7 standard.



<b>Plug</b>
Part Number <b>88040030</b>
Type CEE 7/16
Color Black
Rating 2.5A/250VAC
Approvals

### Australian

Plugs and in-line sockets are also acceptable for use in New Zealand according to the New Zealand Ministry of Energy. Interpower Corporation also offers hospital-grade and higher average re-wireable plugs and sockets. See [www.interpower.com](http://www.interpower.com) for more information.



<b>Plug</b>
Part Number <b>88010713</b>
Type AS/NZS 3112
Color White
Rating 10A/250VAC
Approvals Department of Fair Trading

### Danish

Use a Danish plug on products to be sold in Denmark. Although the Continental European CEE 7-style plug fits into the Danish socket, the plug will not be grounded, presenting an extreme safety hazard. Medical and data versions also available. See [www.interpower.com](http://www.interpower.com).



<b>In-Line Socket</b>
Part Number <b>88010411</b>
Type AS/NZS 3112
Color White
Rating 10A/250VAC
Approvals EANSW

### Panel-Mount Socket

Part Number <b>88010300</b>
Type CEE 7
Color Gray
Rating 16A/250VAC
Approvals

IP 54 "splashproof"

### French/Belgian

France and Belgium use a variation of the CEE 7 socket which contains a male grounding pin and shuttering. The CEE 7/7 plug used throughout Continental Europe has a receptacle for this pin and can be used in France and Belgium. France and Belgium now require safety shuttering on all of their sockets.



<b>Panel-Mounting Socket</b>
Part Number <b>88010550</b>
Type CEE 7
Color Ivory
Rating 16A/250VAC
Approvals —

<b>Panel-Mounting Socket</b>
Part Number <b>88010321</b>
Type CEE 7
Color Gray
Rating 16A/250VAC
Approvals

IP X4 "splashproof"

### Italian



Note: This plug is non-polarized.

<b>Plug</b>
Part Number <b>88010771</b>
Type CEI 23-16/VII
Color Gray/Charcoal
Rating 10A/250VAC
Approvals

<b>Panel-Mounting Socket</b>
Part Number <b>88010572</b>
Type CEI 23-16/VII
Color Silver/Metallic
Rating 10A/250VAC
Approvals

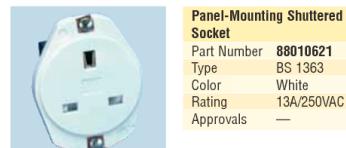
### United Kingdom/Ireland

Line and neutral contacts of socket are shuttered to minimize accidental entry of an object other than mating plug.

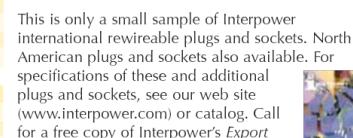


### Japanese Grounding Adapter

When exporting products to Japan or other parts of the world (such as Central America) where 2-wire, non-polarized mains power is typical, include a grounding adapter to ensure that products will be grounded. Shown below is a 3-wire to 2-wire adapter with non-polarized blades.



Line and neutral contacts are shuttered. For black (88010980)



### Order direct from our warehouse in Oskaloosa, IA

- No minimum order requirement
- Same day shipment available for stock item orders received by 6:00 p.m. Central Time
- Most items available from stock
- Free evaluation samples
- Assembly services available
- Technical support available



### Indian/South African

Old British Standard (BS 546) used in India and parts of Southern Africa.



For black (88010230) plug



For black (88010220) socket

<b>Plug</b>
Part Number <b>88010763</b>
Type BS 546
Color White
Rating 15A/250VAC
Approvals —

<b>Panel-Mounting Socket</b>
Part Number <b>88010561</b>
Type BS 546
Color Ivory
Rating 15A/250VAC
Approvals —

### Israeli



For black (88010220) socket

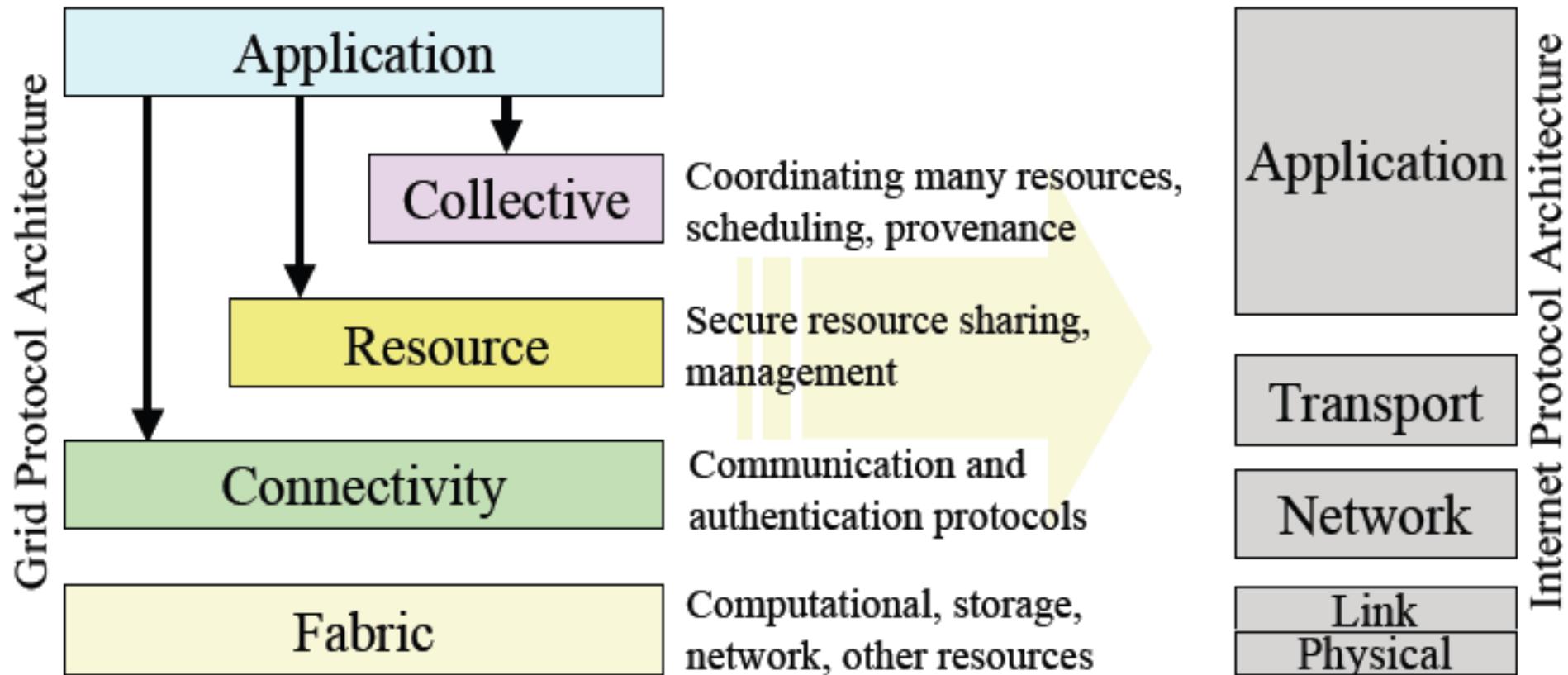


<b>Plug</b>
Part Number <b>88010780</b>
Type SI 32
Color Gray
Rating 13A/250VAC
Approvals

<b>Panel-Mounting Socket</b>
Part Number <b>88010580</b>
Type SI 32
Color White
Rating 16A/250VAC
Approvals

IP X4 "splashproof"

# Typical (Compute) Grid Architecture



All sorted...?

# Grid Standards

- Even increasing numbers and ever decreasing understanding and acceptance/adoption by implementers
  - Global Grid Forum
  - Open Grid Forum
  - OASIS
  - IETF
  - W3C
  - ...



I E T F®



+ many more...

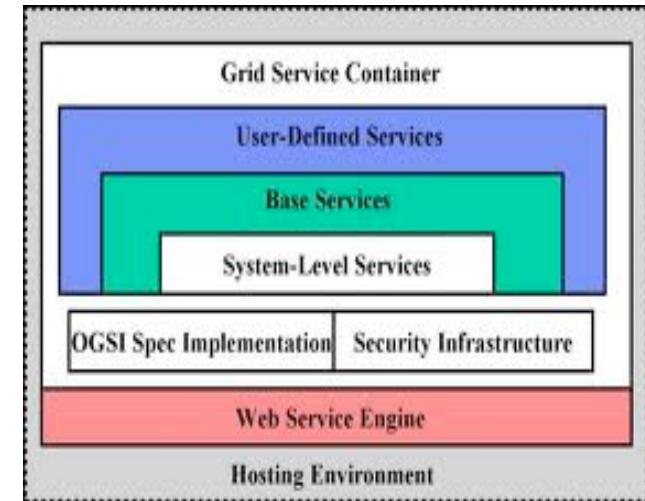
# Grid Technologies... (~2002+)

- Globus Toolkit Project – [www.globus.org](http://www.globus.org)
  - GT2 - Complex software system for large, scale distributed software systems development
    - Physiology of the Grid (Foster et al)  
[www.globus.org/alliance/publications/papers/ogsa.pdf](http://www.globus.org/alliance/publications/papers/ogsa.pdf)
      - *The Open Grid Services Architecture*
    - The Anatomy of the Grid: Enabling Scalable Virtual Organizations (Foster et al)  
<http://dl.acm.org/citation.cfm?id=1080667>
  - MANY MB of source code
    - Many software engineers worked in making this
      - and many more in making it work!!!



# Grid Technologies... (~2004+)

- Move to service-based approach
  - The Open Grid Services Infrastructure (OGSI)
  - [www.globus.org/toolkit/draft-ggf-ogsi-gridservice-33 2003-06-27.pdf](http://www.globus.org/toolkit/draft-ggf-ogsi-gridservice-33_2003-06-27.pdf)
- GT3 – core technologies re-factored as “Grid Services”
  - stateful Web services
  - extension of Web services interfaces
  - asynchronous notification of state change
  - references to instances of services
  - collections of service instances
  - service state data augmenting constraints of XML Schema definition



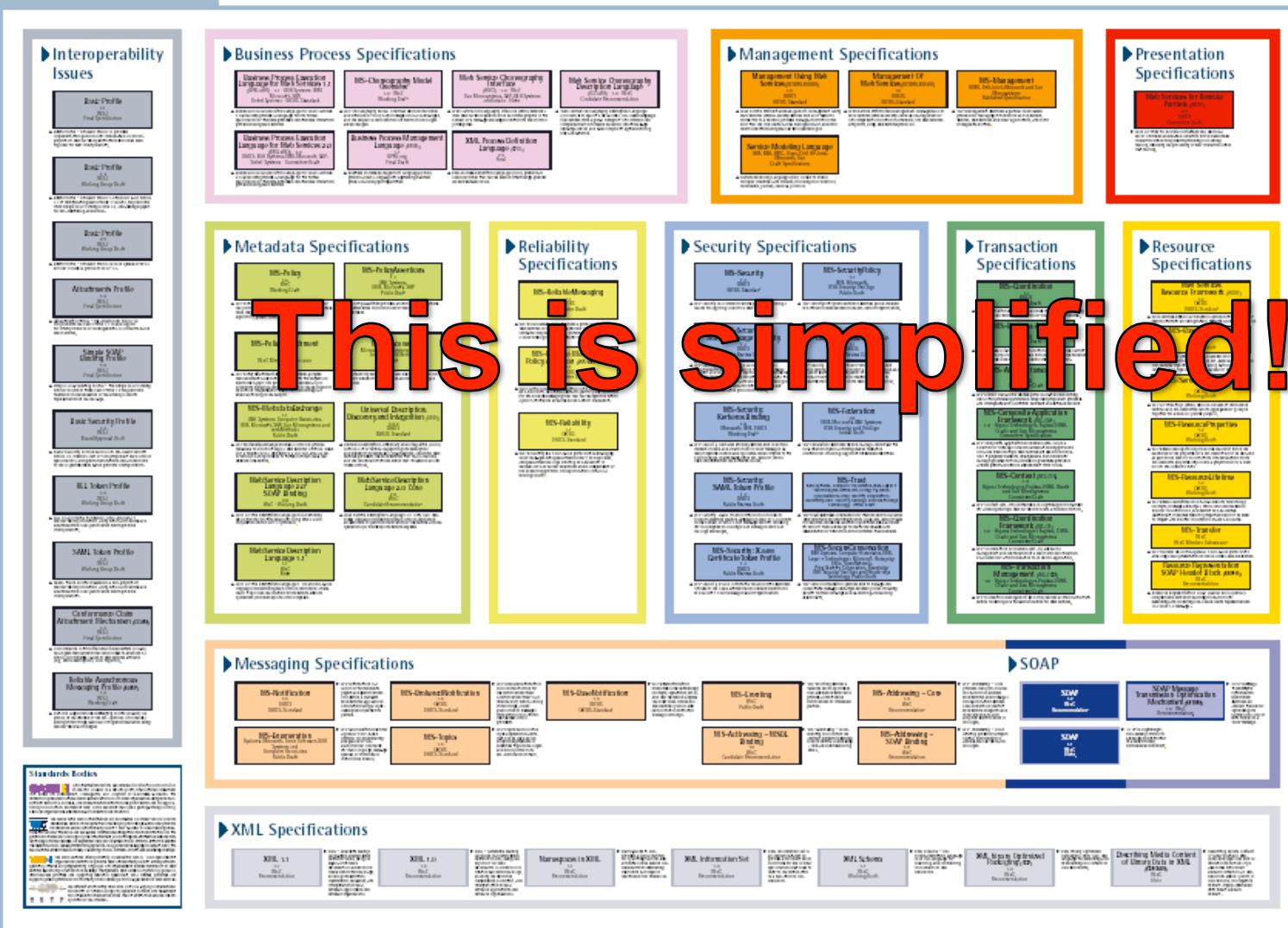
# Grid Technologies... (~2009+)

- Complete reassessment of OGSI approach to be “purer” web services
  - Web service resource framework (WSRF)
  - GT4 implemented this
    - Many software engineers hardened their skill sets using this  
(is one nice way of putting it!!!!)  
*The only way to make software secure, reliable, and fast is to make it small*  
AS Tanenbaum
- Also MANY other standards and efforts ...
  - Business and commercial drivers
    - Vendors shaping standards to their commercial advantage



# Flux of Web Service Standards

# Web Services Standards Overview



innoQ

*The nice thing about standards is that you have so many to choose from...*

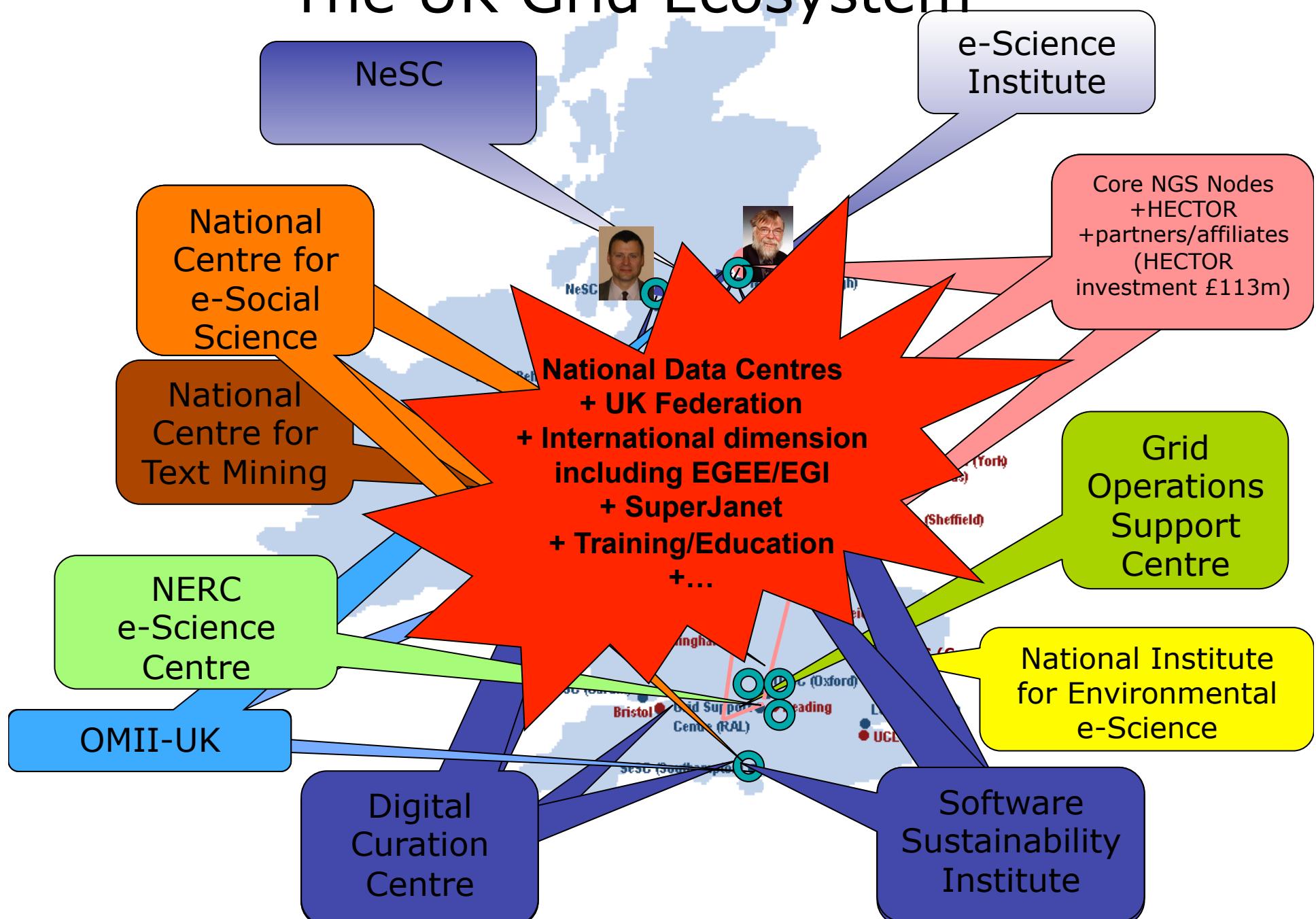
*The development of robust Grid security infrastructures is very much dependent upon agreements on technologies and practices. Standardisation plays an extremely important role in this regard. With the move of the Grid community towards web services and service-oriented architectures, web service security standards and their associated implementations are crucial. Unfortunately it is the case that a multitude of specifications and proposals for web service standards have been promised and put forward, or merely promised. There are often cases of web service standards covering similar topics resulting in multiple competing specifications such as WS-Notifications and WS-Eventing; WS-ReliableMessaging and WS-Reliability; WS-Orchestration, WS-Co-ordination and WS-Choreography, along with the many varieties of workflow or business process languages that have been put forward to name but a few examples of the issues in the proliferation of web service standards. It is also the case that at the time of writing, many web services standards are only in working draft or draft status, often with no associated implementations or acknowledged conformance or interoperability definitions. Claiming conformance or compliance to a particular web service standard is thus often not possible (or meaningful!).*

*It is also apparent that although many standards use the common prefix “WS-”, this does not mean that there is an agreed WS-Architecture. This stems from a variety of reasons: vendor and commercial issues; political aspects and also the different bodies involved. For example the Internet Engineering Task Force (IETF) ([www.ietf.org](http://www.ietf.org)); the World Wide Web Consortium (W3C) ([www.w3.org](http://www.w3.org)); the Organization for the Advancement of Structured Information Standards (OASIS) ([www.oasis-open.org](http://www.oasis-open.org)); and the Web Services Interoperability Organization (WS-I) ([www.ws-i.org](http://www.ws-i.org)) are some of the most prominent bodies. The consequence of this profusion of standards and standards making bodies, and the lack of consensus on the core web service architecture, impacts directly upon development of Grid standards, architectures and associated implementations and middleware.*

**Me: Grid Security : Practices, Middleware and Outlook**

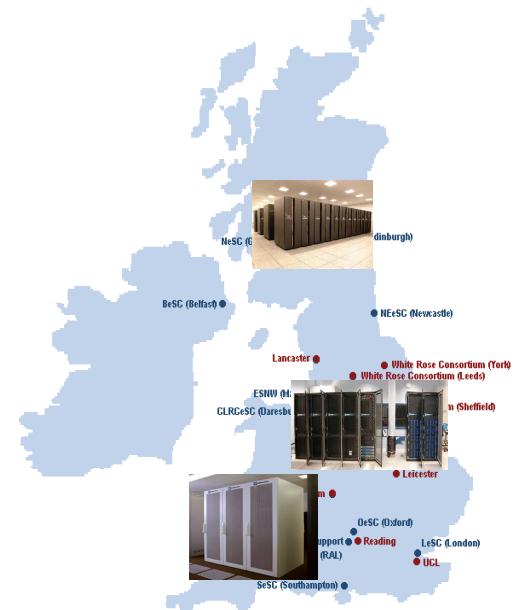
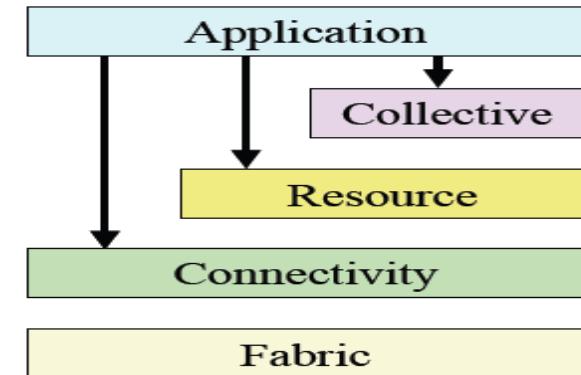
# The UK Grid Ecosystem

*Circa 2004*



# How hard can a compute Grid be...?

- Information Systems
  - What resources are available
    - Servers, CPUs, memory, storage, queues, OS, applications, databases, ...
- Monitoring and Discovery Systems
  - What is the status of those resources
    - Queues empty/very full
    - Machines running for a week
    - Hard to tell for some applications...?
- Job scheduling/resource brokering
  - Please run these {jobs}
    - Fastest, most secure/reliable, cheapest ...
      - Jobs need inter-process communication...?
        - » never get resources vs resource starvation
        - » Physics pilot jobs workaround
  - Virtual organisation support
    - I'm a chemist and want to run my simulation



# Fragility of it all...

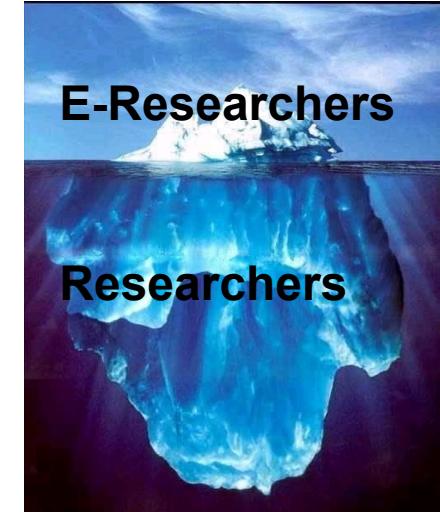
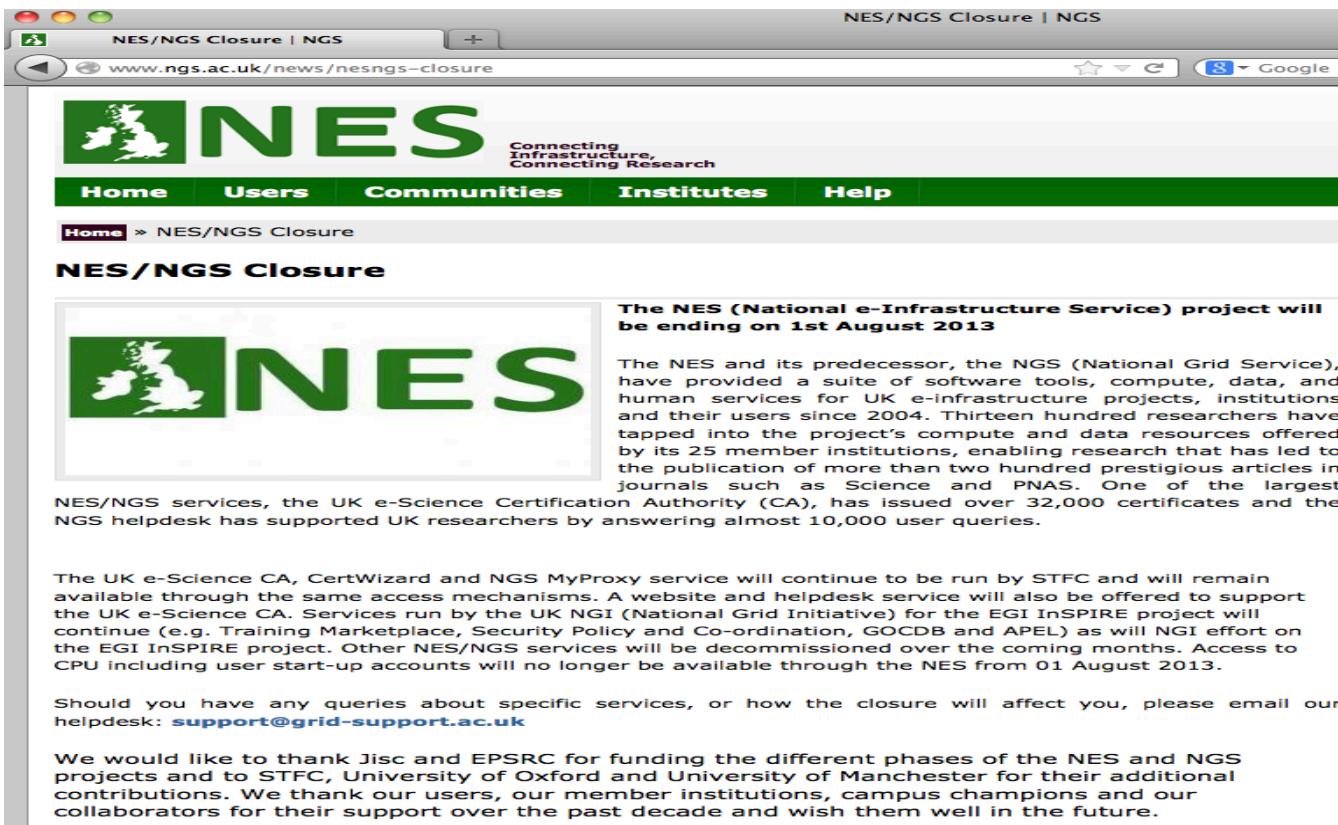
- How I broke the UK e-Science Grid...



# Other Challenges

- Security
    - Public Key Infrastructures
      - *\$> openssl pkcs12 -in cert.p12 -clcerts -nokeys -out usercert.pem!*

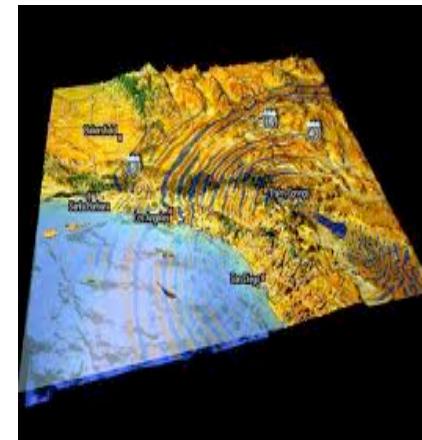
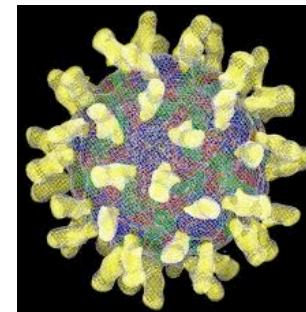
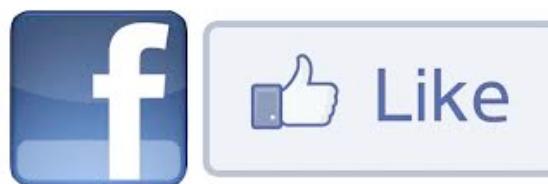
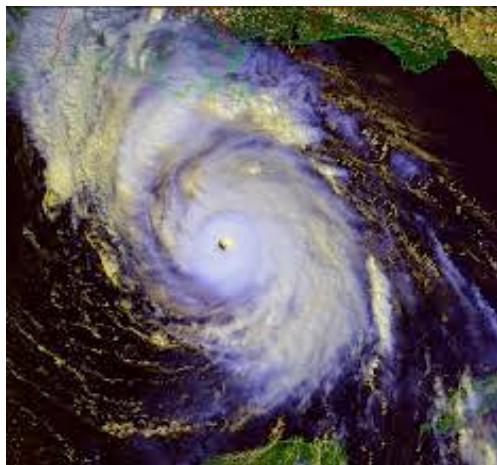
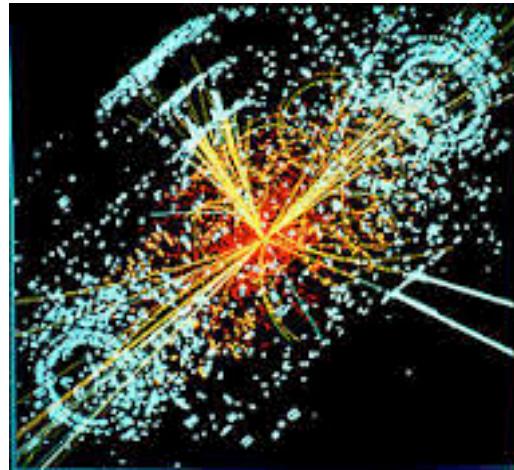
# ???? (more in later lectures)



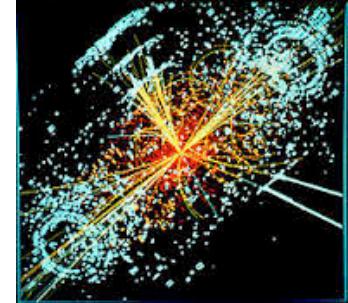
(NOT TO SCALE)

# Doom and gloom...!?!?

- But...



# High Energy Physics?



- Why did/does Grid work for them...?
- Size, scale and resourcing
  - GridPP
  - Enabling Grids for e-Science (EGEE)
  - European Grid Initiative
    - Huge numbers of admins making systems tick over  
不均匀性 - (=not searching for Higgs Boson)
    - Heterogeneity vs homogeneity 同质
      - Specific versions of operating system
        - Scientific Linux v4, v5
      - Specific versions of middleware
        - Unilateral updates/deployments
          - » globus-\*, gLite-wms-\*,
          - » edg-job-\*, lcg-job-\*, GANGA
    - Well-defined problem...



# My philosophy since...

- Less focus on complex middleware stacks
  - FAR too much focus on technology issues
    - More on customer needs and making a difference
- And to inspire you (and make you think what you do with social media!!!)
  - A snapshot of what you will be doing with Twitter analytics!



Questions?