Everything you always wanted to know about the Grid and never dared to ask

Tony Hey and Geoffrey Fox

Outline

- Lecture 1: Origins of the Grid – The Past to the Present (TH)
- Lecture 2: Web Services, Globus, OGSA and the Architecture of the Grid (GF)
- Lecture 3: Data Grids, Computing Grids and P2P Grids (GF)
- Lecture 4: Grid Functionalities – Metadata, Workflow and Portals (GF)
- Lecture 5: The Future of the Grid – e-Science to e-Business (TH)

Lecture 1

Origins of the Grid –

The Past to the Present

[Grid Computing Book:

Chs 1,2,3,4,5,6,36]

Lecture 1

1. Licklider and the ARPANET
2. Technology Trends: Moore's Law and all that
3. The Imminent Data Deluge
4. Grids, e-Science and Cyberinfrastructure
5. Early Attempts at Building Grids
6. The NASA Information Power Grid
7. Grids Today

J.C.R. Licklider and the ARPANET

- Licklider was an experimental psychologist who was recruited to head up two groups - ‘Behavioural Sciences’ and ‘Command and Control’ – at the US Advanced Research Projects Agency ARPA in 1962
- Had realized that in his research he had to spend most of his time organizing and manipulating data before he could do research so he brought with him to ARPA a vision

Origins of the Internet

- ‘Lick’ described his vision as follows:

  ‘If such a network as I envisage could be brought into operation, we could have at least four large computers, perhaps six or eight small computers, and a great assortment of disk files and magnetic tape units … all churning away’

- He had no idea how to build it so he funded Computer Science research at MIT, CMU, Stanford and Berkeley: these groups became the core of his ‘Intergalactic Computer Network’

Technology Trends

- Moore's Law - capacity doubles every 18 months - set to continue for 10 years
- 0.1 micron feature size by 2005: X-ray lithography for smaller features?
- By 2010 or so will have 128 Gbyte DRAMs 'with as many transistors as there are stars in our galaxy'

Moore's Law

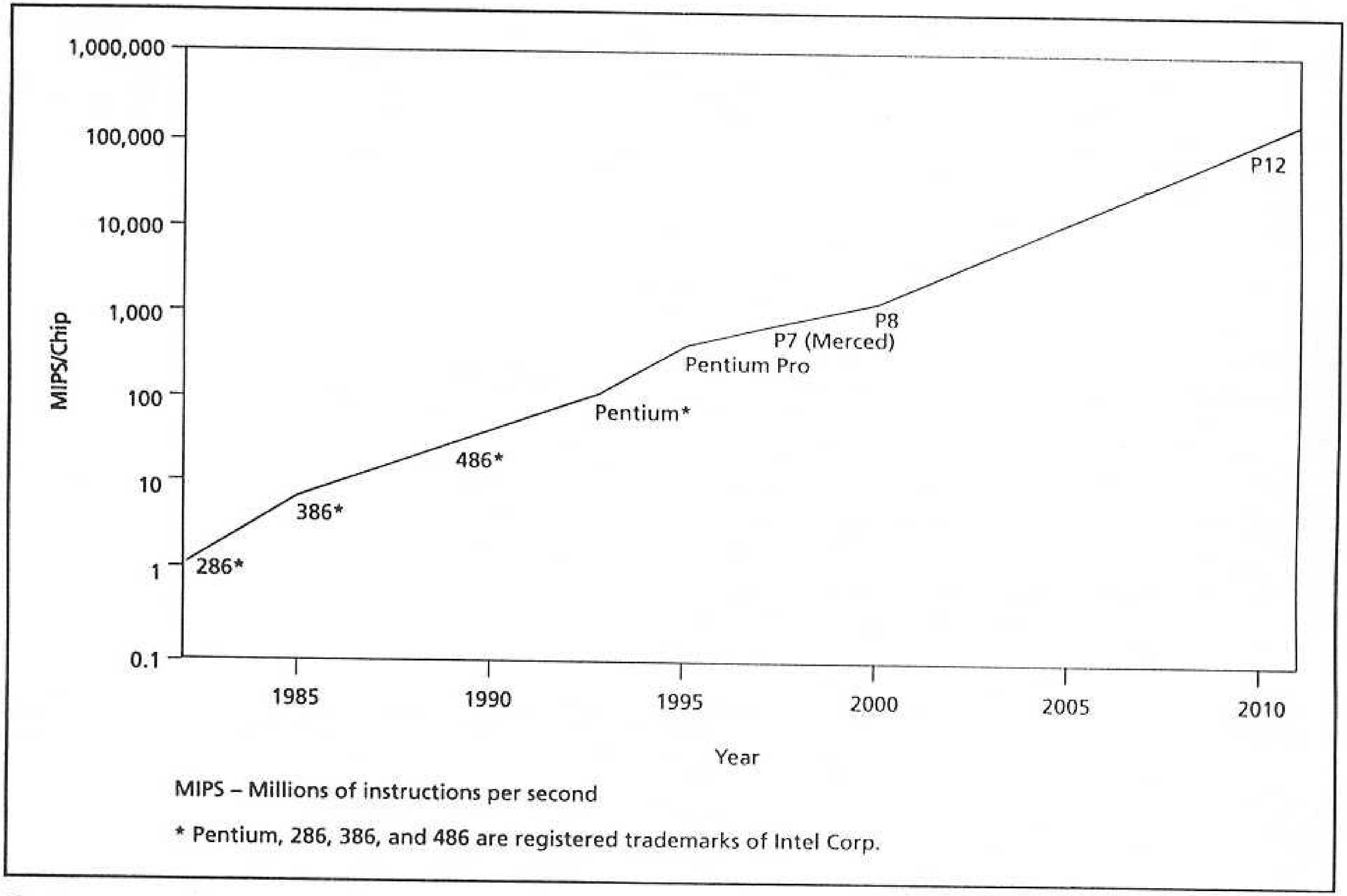


Figure 4.
Increase in MIPS per chip.

Growth of DRAM Use

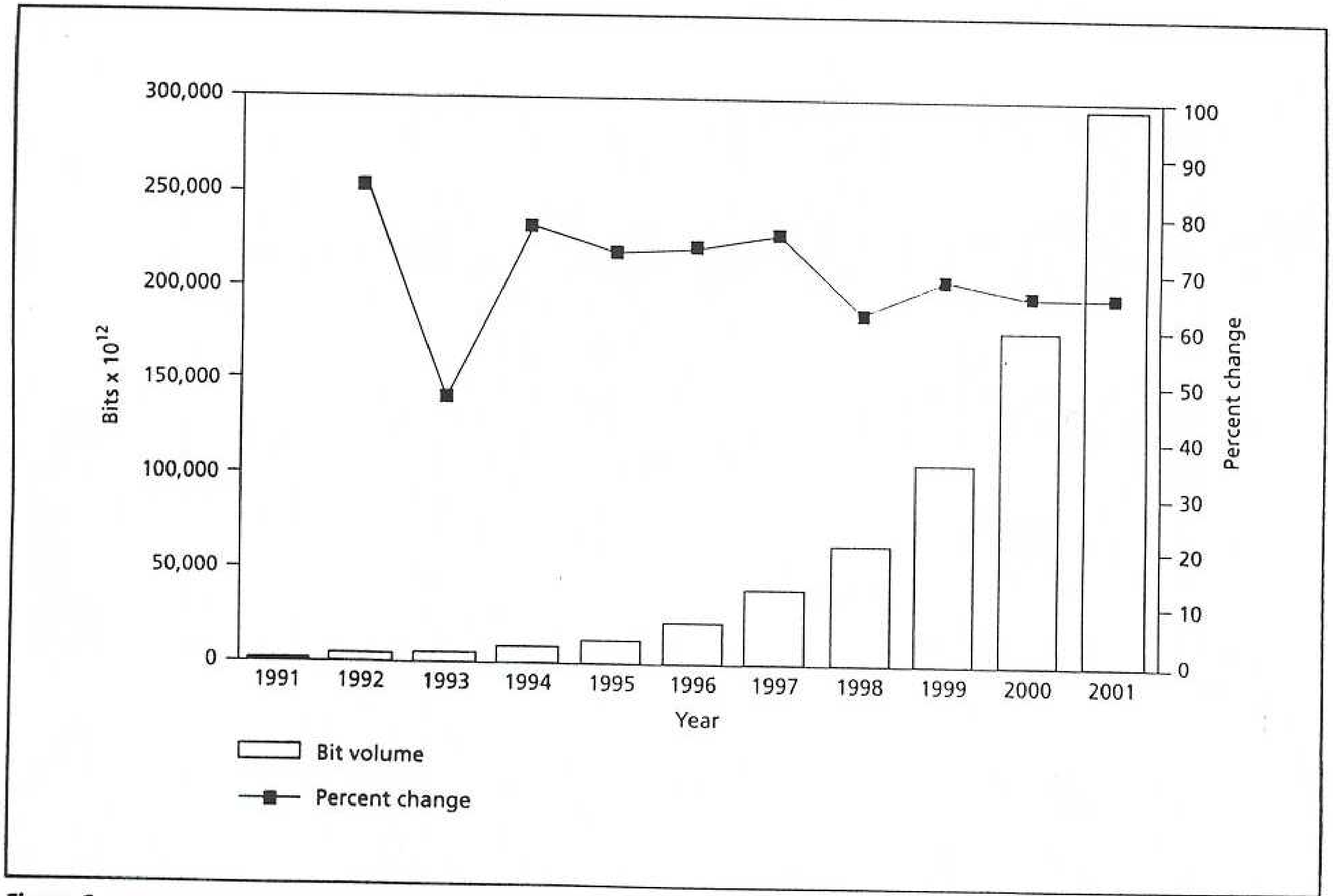


Figure 6.
Astounding worldwide growth of DRAM use and percent change per year.

Optical Amplifiers Transform Long-Distance Lightwave Telecommunications

REFERENCES

[1] R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "Low-noise Erbium-doped fiber amplifier operating at 1.54 μm," *Electron. Lett.*, vol. 23, pp. 1026–1028, Sept. 1987.

WDM NETWORKS: A U.S. PERSPECTIVE

Herwig Kogelnik

Lucent Technologies, Inc., Bell Labs
Crawford Hill Laboratory
Holmdel, NJ 07733, USA

INTRODUCTION

For America's long-distance networks, this is the year of WDM, of Wavelength-Division Multiplexing. After years of R&D preparation, suppliers have started volume production of WDM systems in the fall of 1995, and in 1996 the long-distance network operators have begun a large-scale deployment of WDM systems involving thousands of fiber amplifiers and hundreds of WDM SONET terminals. This large deployment is making a significant contribution to the evolution of a mature WDM network technology and places WDM network elements and components on an accelerated cost/technology learning curve. These include WDM sources and a progression of bit-rate transparent components from fibers, to fiber amplifiers, WDM routers and WDM cross-connect fabrics. Coupled with this we are witnessing a progression of WDM networking R&D from point-to-point systems, to ...ings, more complex WDM architectures and WDM network management ...xploit the new degree of freedom (in addition to space and time) ...n greater flexibility for network design.

2.2. Erbium Doped Fibre Amplifiers for WDM transmission

Conventional Erbium Doped Fibre Amplifiers (EDFA) are not very suitable for WDM trans... because of the non uniform gain spectrum. For these types of amplifiers, WDM can be used only if the wavelengths correspond to a particular ...sition of equal gain of the amplifier. This imposes severe practical constraints on wavel... ...flat gain amplifiers are thus by far preferable. Let us mention two ...flattening. The first one consists in introducing within ... ...iform gain [5]. A second approach consists i... flatter gain curve [6]. This is obta... (Erbium Doped Fluoride Fibre Am... used for WD... transmission: 1 dl...

Currently, third-generation trans-oceanic lightwave systems are being deployed that operate at 5.0 Gbits/s per fiber pair and use erbium-doped fiber amplifiers (EDFA) in repeaters to boost signals. Similar technology has enabled undersea non-repeatered systems to span distances well beyond 250 kilometers.

...with this we are ... All three experiments were conducted at 1550 nm, the wavelength of optical fiber's minimum attenuation and the only one at which erbium-doped fiber amplifiers function. Broadband optical amplifiers are key to WDM systems since they amplify the entire spectrum of interest in a single ... Repeaters that regenerate pulse ...lify them, admit-...

Conclusions

New undersea lightwave systems now being installed have just begun to tap the enormous potential ...red by the EDFA. A variety of new technologies will ...undersea systems with not only vastly greater ...ies, but also with the ability to implement com...works. A key technology for these advance... WDM, which enables the same EDFA to amplify ...nals at numerous different wavelengths. This ...y digital signals supports both significantly ...capacity and flexible traffic routing. Where ...ity is the focus, higher-bit-rate systems ...a single wavelength are also feasible. Both ...n and higher-rate single-wavelength sys-... ...possible with the NRZ signaling used in the cur-...ent generation of undersea optical amplifier systems.

2.6 Terabit/s WDM Transmission Experiment using Optical Duobinary Coding

...no, T. Ono, K. Fukuchi, T. Ito, H. Yamazaki, M. Yamaguchi, and K. Emura
Opto-Electronics Research Laboratories, NEC Corporation
4-1-1, Miyazaki, Miyamae, Kawasaki 216, Japan

Future Directions for Undersea Communications

...W. Korfoot ...tor K. Runge

Technologies used in undersea systems have changed radically in the las... five years, driven primarily by the emergence of the erbium-doped fiber amplifier (EDFA). EDFAs have enabled long, transoceanic systems to offer significantly higher capacities (5 gigabits per second [Gbits/s] per fiber pair) and much longer lengths in repeaterless systems. The opportunities offered by EDFAs, combined with other technological advances, extend far beyond those used in currently planned systems. This paper describe[s] the new technologies that are expected to play a role in future undersea ...and the opportunities for applying those technologies.

WDM breaks the Tb/s barrier

Just as ADSL promises to extend th[e] usefulness of existing wire plant, wave length-division multiplexing (WDM) promises to extend the usefulness of existing fiber. WDM is basically old-fashioned frequency-division multiplexing at optical frequencies. Researchers at each of three laboratories have used it to attain a sort of Holy Grail of digital communications: data transmission at 1 terabit a second—yes, a full 1000 Gb/s.

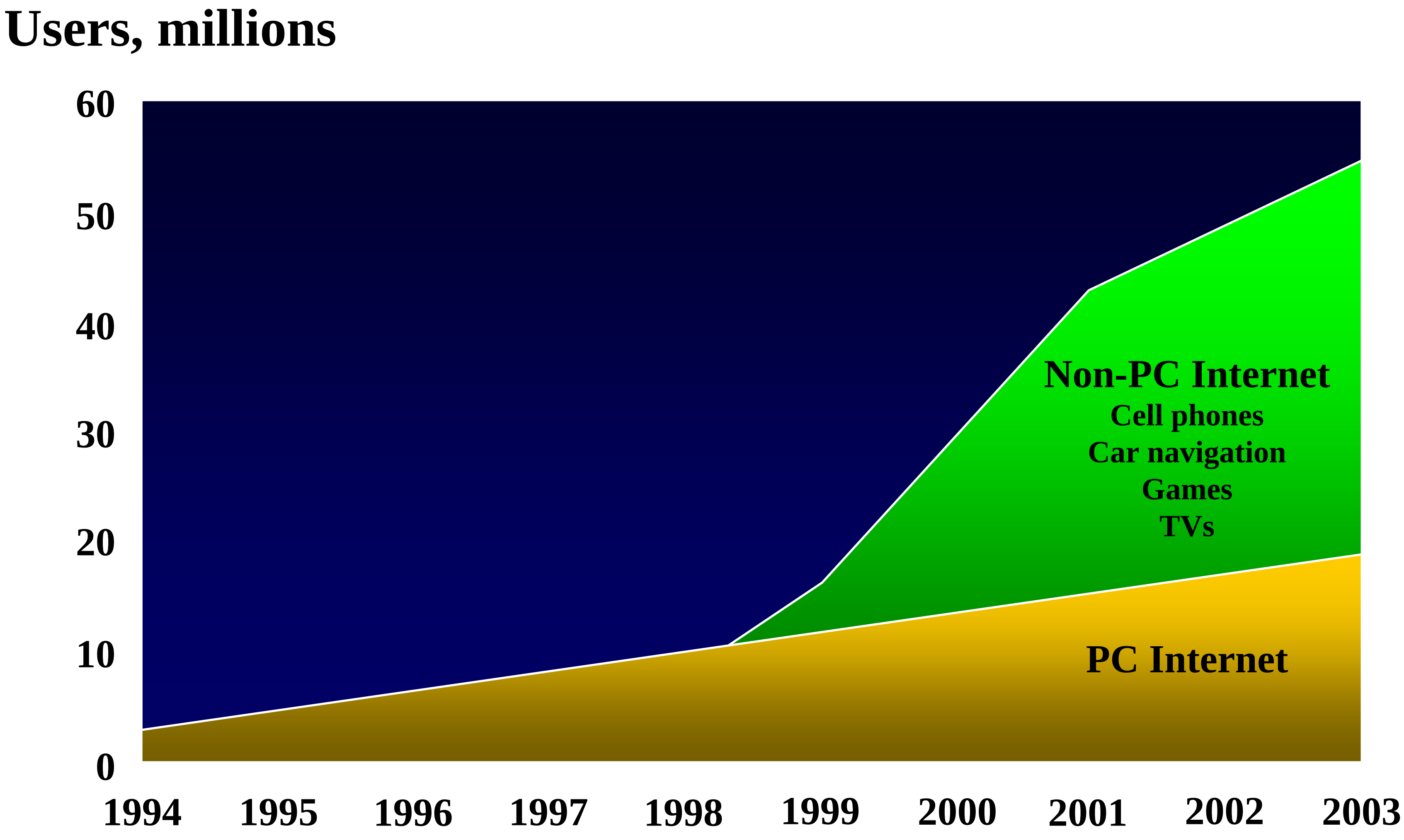
Gilder's Law

- Proposed by George Gilder, prolific author and prophet of the new technology age:

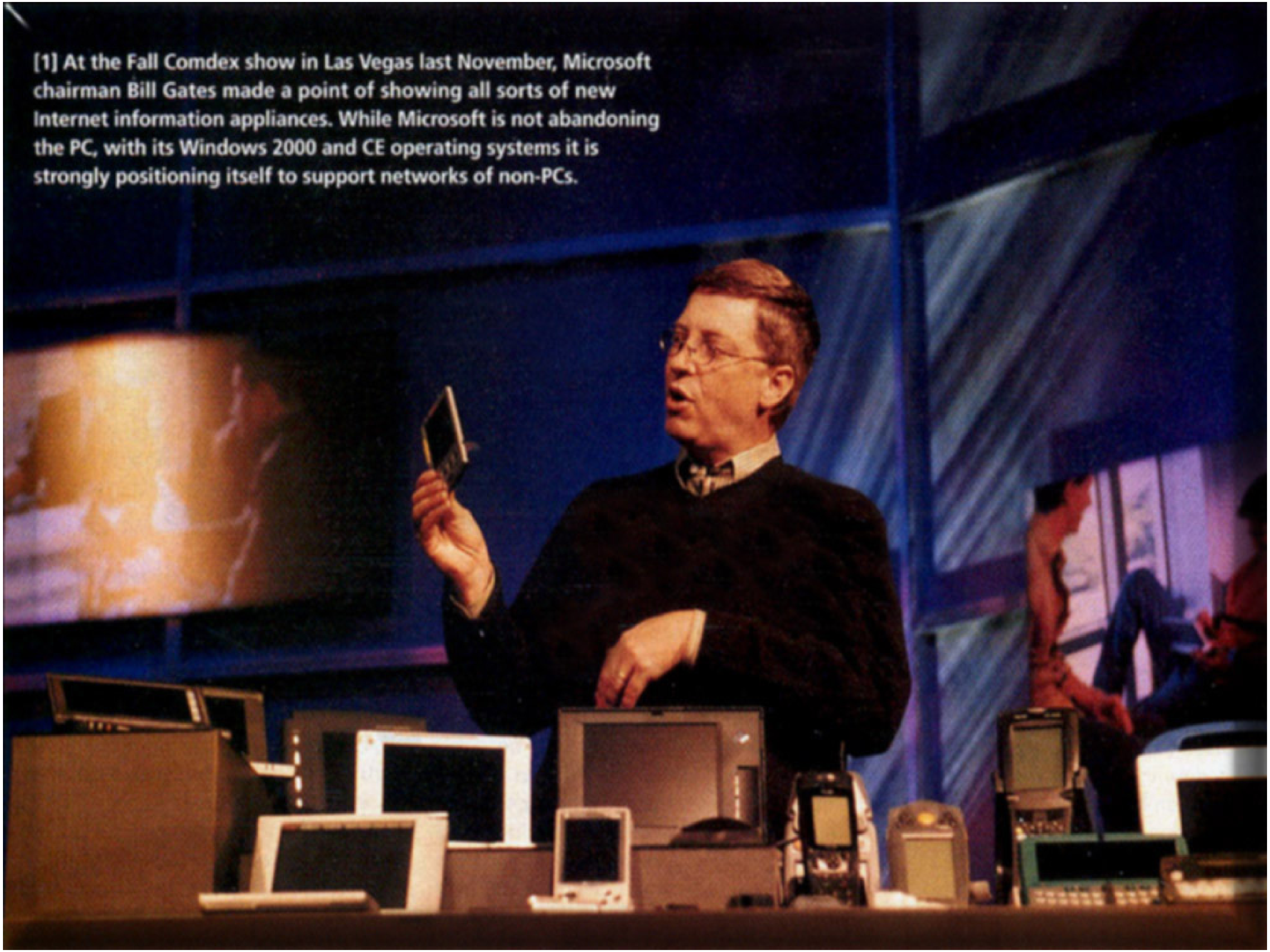
  'The total bandwidth of communication systems triples every twelve months'

- Seems to continue to be true for backbone networks but end-to-end performance is likely to be limited by the MANs and LANs

Dawn of the Non-PC Internet



Source: Matsushita Electric



[1] At the Fall Comdex show in Las Vegas last November, Microsoft chairman Bill Gates made a point of showing all sorts of new Internet information appliances. While Microsoft is not abandoning the PC, with its Windows 2000 and CE operating systems it is strongly positioning itself to support networks of non-PCs.

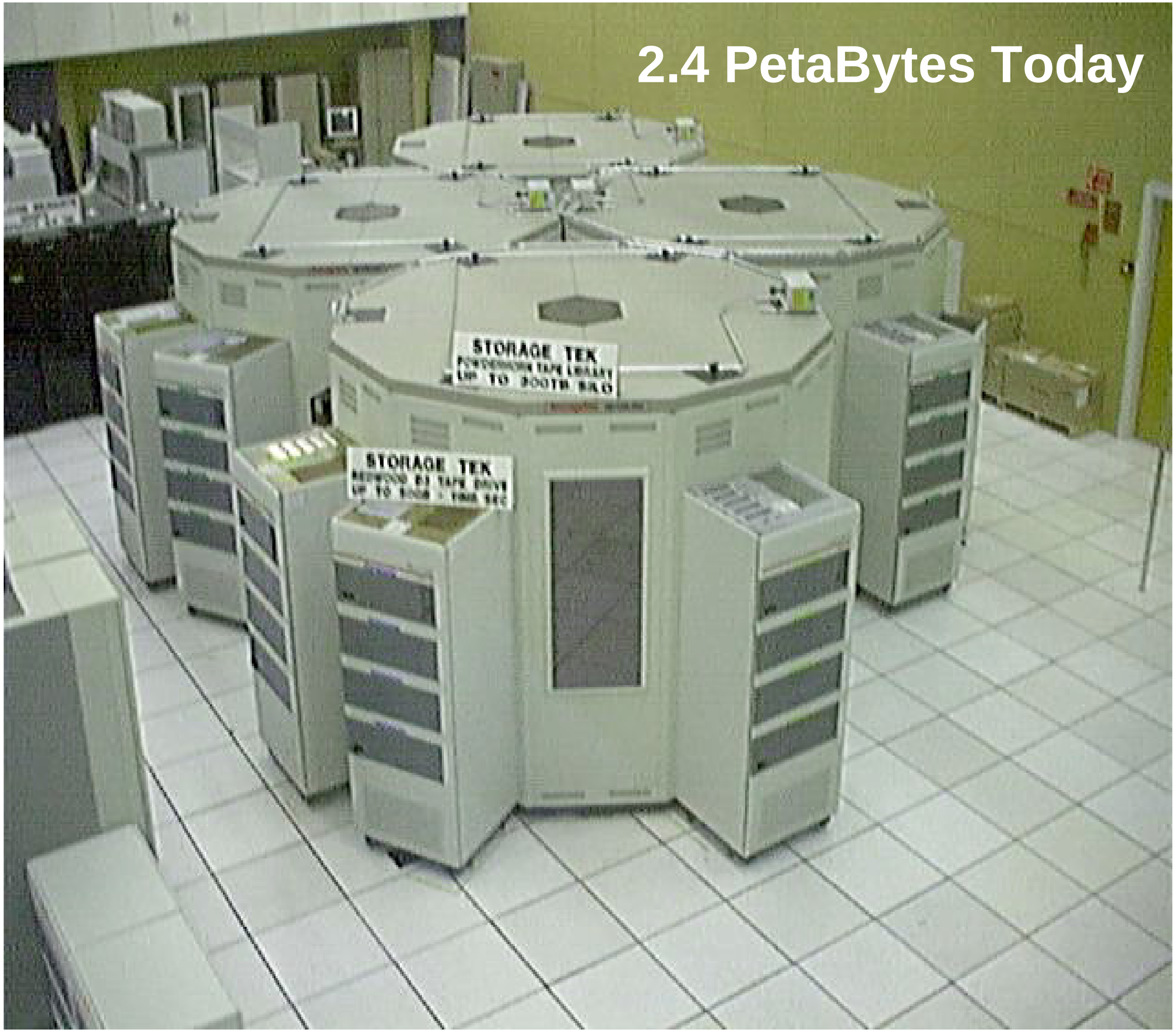
The Imminent Data Deluge

- How large is a Petabyte?

  - Big enough to store a few high-resolution images for every human being on Earth

  - Would fill enough high-capacity magnetic tapes to fill a railway carriage

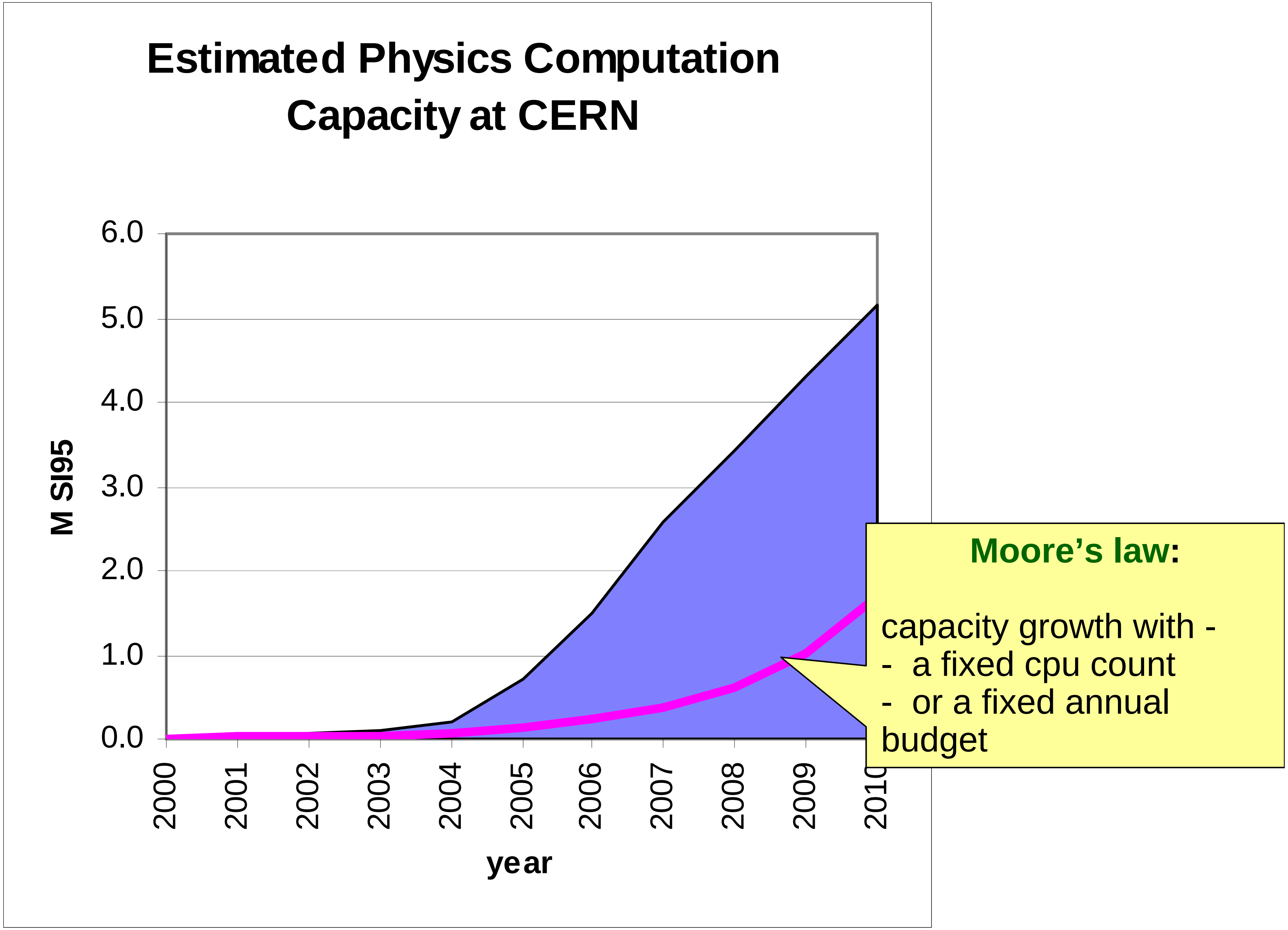
- To stream a Petabyte through a single high-speed tape drive would take 3 years


2.4 PetaBytes Today
STORAGE TEK
STORAGE TEK

Particle Physics – Hunting the Higgs

- 2006/7: First pp collisions at TeV energies at the Large Hadron Collider at CERN in Geneva
- ATLAS/CMS Experiments involve 2000 physicists from 200 organizations in US, EU, Asia
- Need to store,access, process, analyse 10 Petabytes/yr with 200 Teraflop/s distributed computation
- Building hierarchical Grid infrastructure to distribute data and computation
- Many 10’s of million $ funding for global particle physics Grid – GryPhyN, PPDataGrid, iVDGL, EU DataGrid, EU DataTag, GridPP projects

➢ Need Exabytes and Petaflop/s by 2015



Estimated Physics Computation Capacity at CERN
M SI95
6.0
5.0
4.0
3.0
2.0
1.0
0.0
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
year
Moore's law:
capacity growth with -
- a fixed cpu count
- or a fixed annual budget

CERN's Users in the World

Europe: 267 institutes, 4603 users
Elsewhere: 208 institutes, 1632 users

Astronomy: History of the Universe

- Virtual Observatories – NVO, AVO, AstroGrid
  - Store all wavelengths, need distributed joins
  - NVO 500 TB/yr from 2004
- Laser Interferometer Gravitational Observatory
  - Search for direct evidence for gravitational waves
  - LIGO 250 TB/yr, random streaming from 2002
- VISTA Visible and IR Survey Telescope in 2004
  - 250 GB/night, 100 TB/yr, Petabytes in 10 yrs

➢ New phase of astronomy, storing, searching and analysing Petabytes of data

Engineering, Environmental and Medical Applications

- Real-Time Industrial Health Monitoring
  - UK DAME project for Rolls Royce Aero Engines
  - 1 GB sensor data/flight, 100,000 engine hours/day
- Earth Observation
  - ESA satellites generate 100 GB/day
  - NASA 15 PB by 2007
- Medical Images to Information
  - UK IRC Project on mammograms and MRIs
  - 100 MB/mammogram, UK 3M/yr, US 26M/yr
  - 200 MB/patient, Oxford 500 women/yr

➢ Many Petabytes of data of real commercial interest

Importance of Metadata

- Metadata is ‘data about data’

  e.g. catalogues, indices, directory structures

- Librarians work with books which have same basic ‘schema’

  e.g. title, author(s), publisher, date, etc

- Need for hierarchical, community-based approach to defining metadata and schemas

  e.g. CML, AML, …

➢ Metadata important for interoperability of databases/federated archives, and for construction of intelligent search agents

Simulation Output as Digital Library

- Digital Libraries usually for archiving of text, audio and video data
- Scientific data require transformation, data-mining and visualisation tools
- For distributed collaborations need simulation output to be available as new kind of digital library, complete with catalogues and finding aids as well as data itself

Emergence of a new research methodology?

- Traditional scientific methodologies are theory and experiment
- Last half of 20th century saw emergence of scientific simulation as a third methodology
- This century will see emergence of a fourth methodology - collection-based research
  - scientists will reduce, mine and sift data 'published' in ways not possible with paper journals with their tables and graphs

Data Curation and Preservation

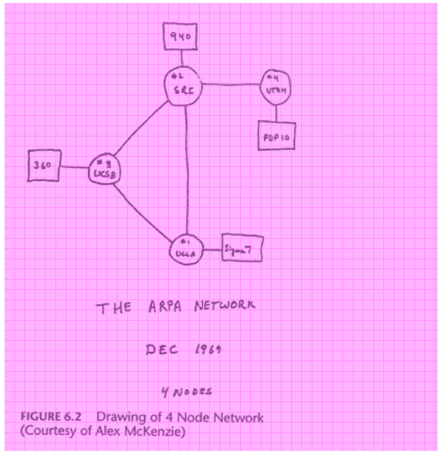
- In next 5 years e-Science projects will produce more scientific data than has been collected in the whole of human history
- In 20 years can guarantee that the operating and spreadsheet program and the hardware used to store data will not exist

➢ Need to research and develop technologies and best practice for curating digital data

➢ Develop best practice for data refreshment, replication, repackaging and transformation

Current Trends: Summary (1)

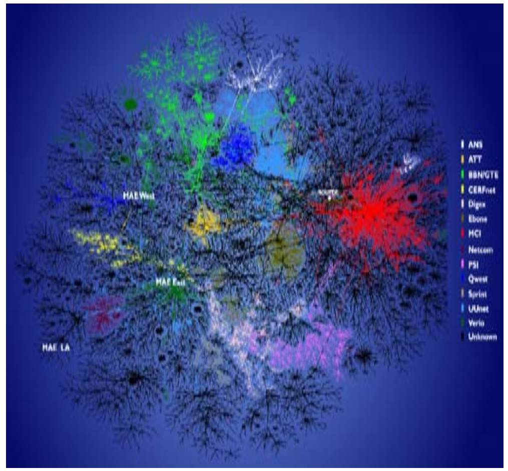
- Proliferation of resources
  - Everyone has computers
  - Multiple IP addresses per person



Arpanet 1969

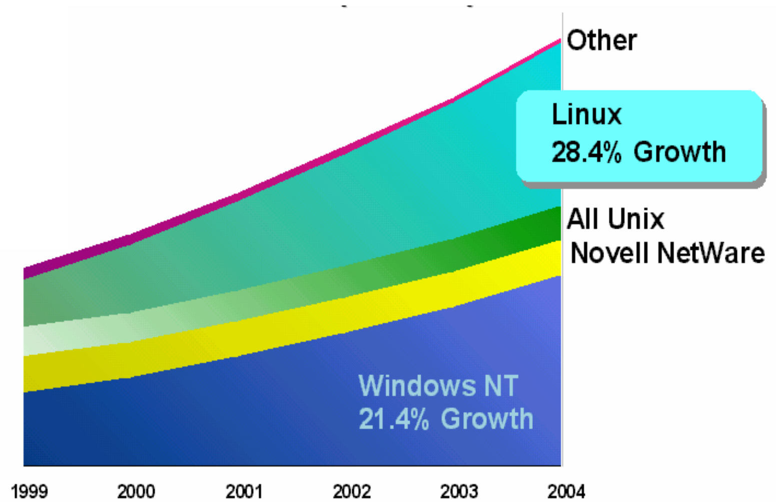
- Increasing Application Complexity
  - Multi-scale
  - Multi-disciplinary
  - Immense amounts of data

Internet 2002



Current Trends: Summary (2)

- Coordination/collaboration is default mode of interaction
  - The Internet
  - Globalization, virtualization
  - Open source movement



Source: Dave Turek, IBM

from IDC "Server Operating Environments Market Forecast & Analysis"
January 2001

- At all scales, heterogeneity is a fact of life

e-Science and the Grid

‘e-Science is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it.’

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

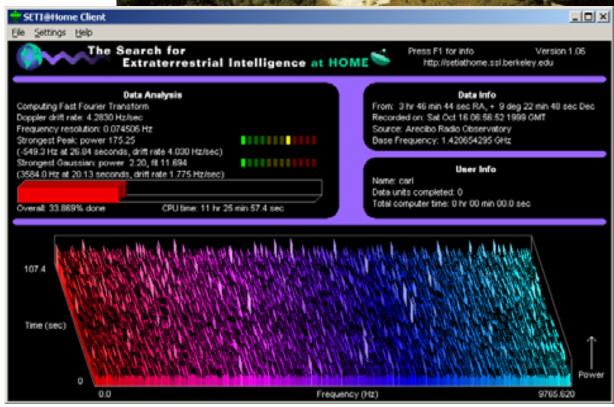
The Grid as an Enabler for Virtual Organisations

- Ian Foster, Carl Kesselman and Steve Tuecke
- ‘The Grid is a software infrastructure that enables flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources’

  \- includes computational systems and data storage resources and specialized facilities

The Grid is NOT just SETI@home

- SETI@home
  - 3.8M users in 226 countries
  - 1200 CPU years/day
  - 38 TF sustained (Japanese Earth Simulator is 40 TF peak)
  - 1.7 ZETAflop over last 3 years (10^21, beyond peta and exa …)
  - Highly heterogeneous: >77 *different* processor types

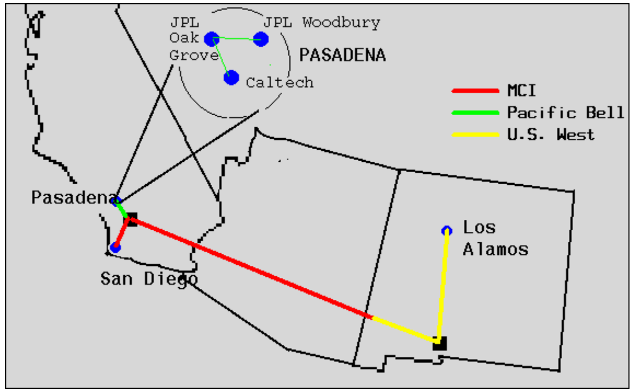


Two commercial views of the Grid

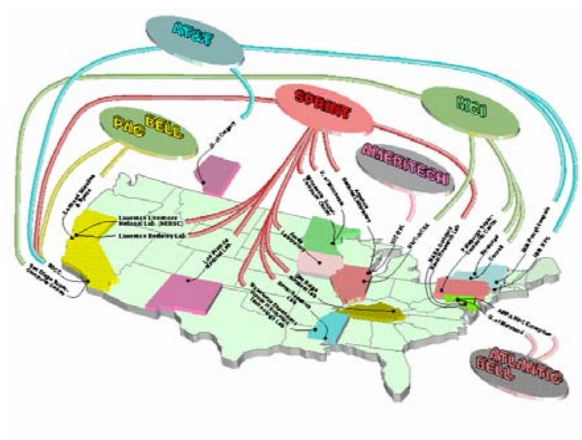
- "A new technology infrastructure, which we call the GRID, will emerge to supersede the Web, and form the bedrock of the new economy" (Deutsche Bank)
- "Web and computing services shall go the way of the electric power grid – always available, easily accessible, and predictably expandable" (PricewaterhouseCoopers)

A short history of the Grid

- "Science as a team sport"
  - Grand Challenge Problems of the 80's
- Parallel computation
  - First serious study of program coordination
- Gigabit Testbed program
  - Focus on applications for the local to wide area
- I-Way at SC '95
  - First large-scale grid experiment
  - Provided the basis for modern grid infrastructure efforts

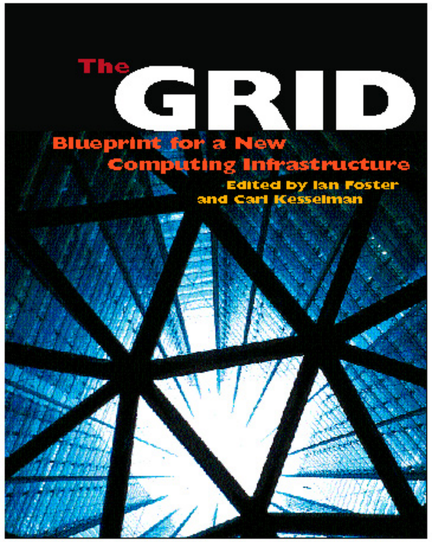


CASA Gigabit Testbed



1995 – 2001: Maturing the Grid

- "Grid book" gave a comprehensive view of the state of the art
- Important infrastructure and middleware efforts initiated
  - Globus, Legion, Condor, NWS, SRB, NetSolve, AppLes, etc.
- 2000+: Beginnings of a Global Grid
  - Evolution of the Global Grid Forum
  - Some projects evolving to de facto standards (e.g. Globus, Condor, NWS)

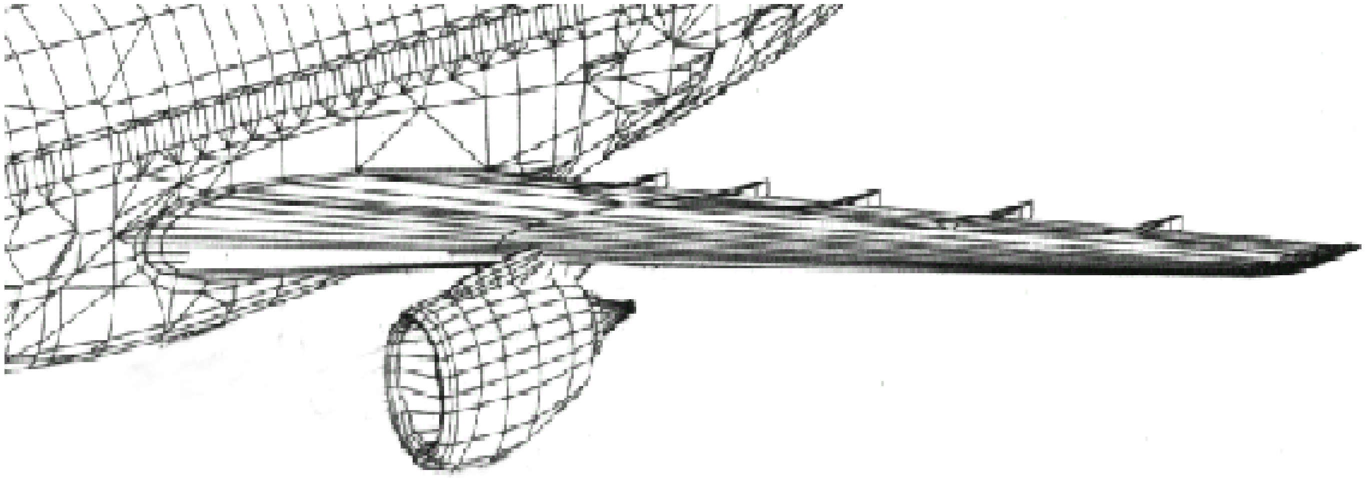


NASA's IPG

Vision for the Information Power Grid is to promote a revolution in how NASA addresses large-scale science and engineering problems by providing persistent infrastructure for:

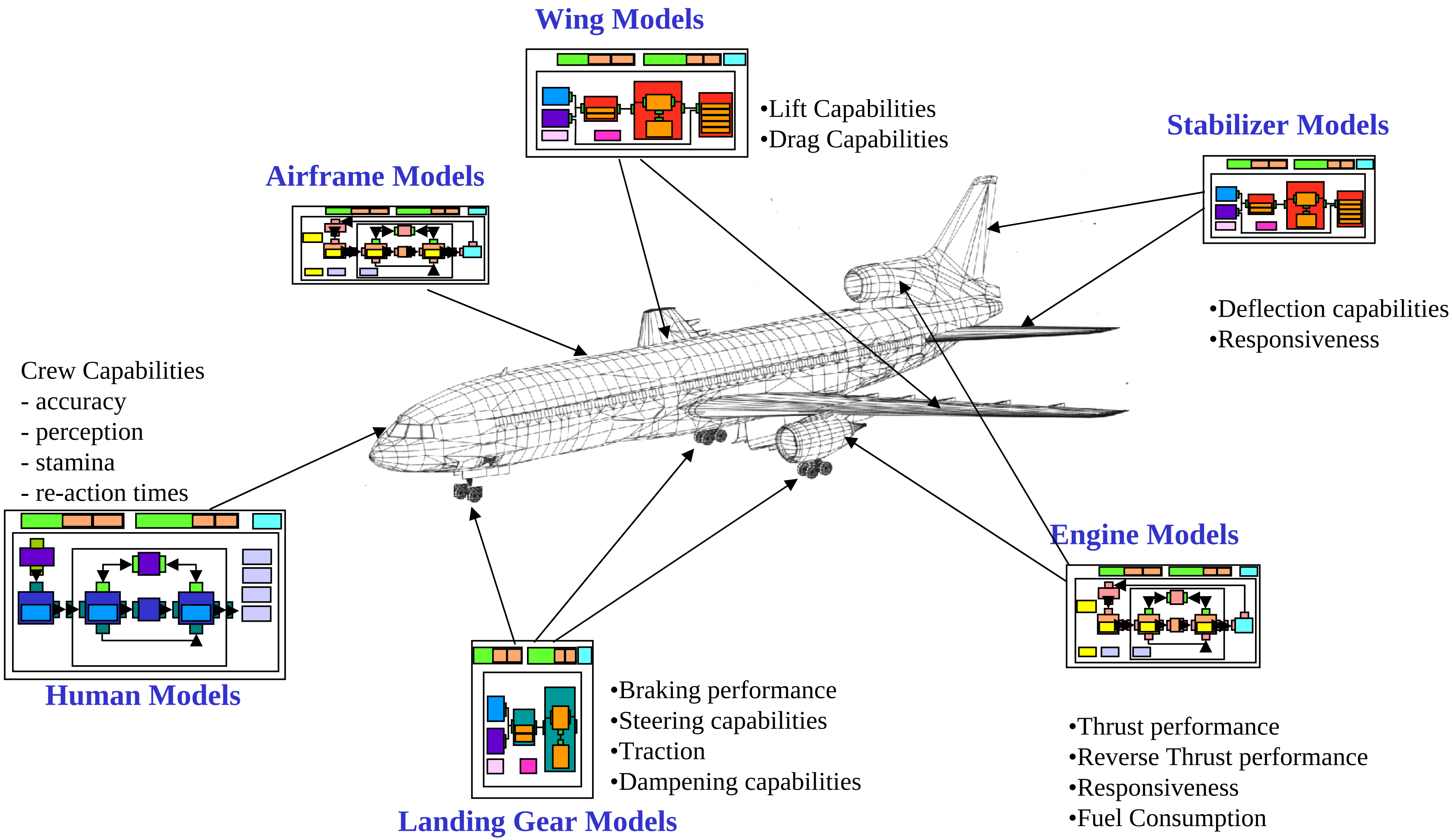
- "highly capable" computing and data management services that, on-demand, will locate and co-schedule the multi-Center resources needed to address large-scale and/or widely distributed problems
- the ancillary services that are needed to support the workflow management frameworks that coordinate the processes of distributed science and engineering problems

Multi-disciplinary Simulations



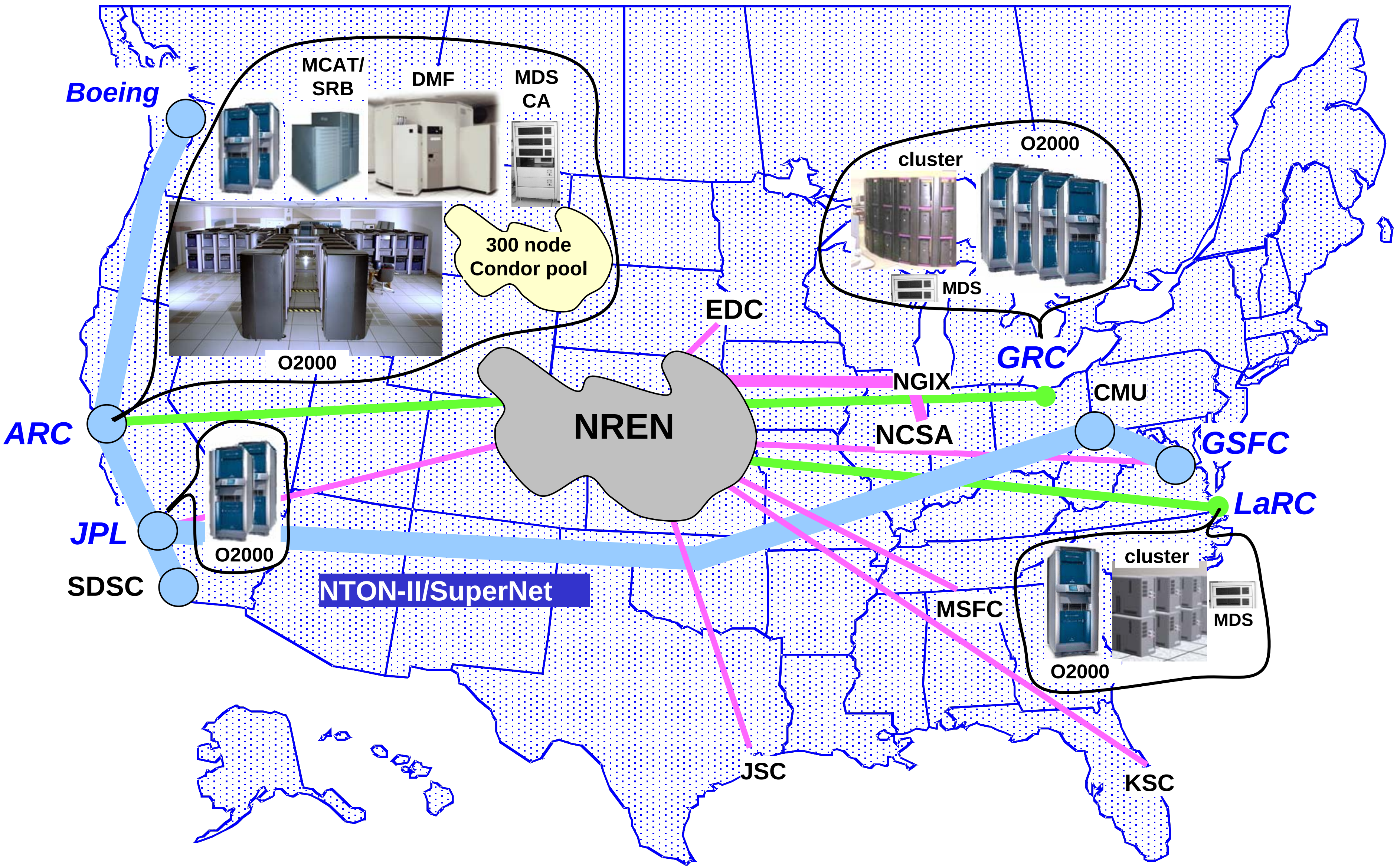
Multiple sub-systems, e.g. a wing lift model operating at NASA Ames and a turbo-machine model operating at NASA Glenn, are combined using an application framework that manages the interactions of multiple models and uses IPG services to coordinate computing and data storage systems across NASA.

Multi-disciplinary Simulations

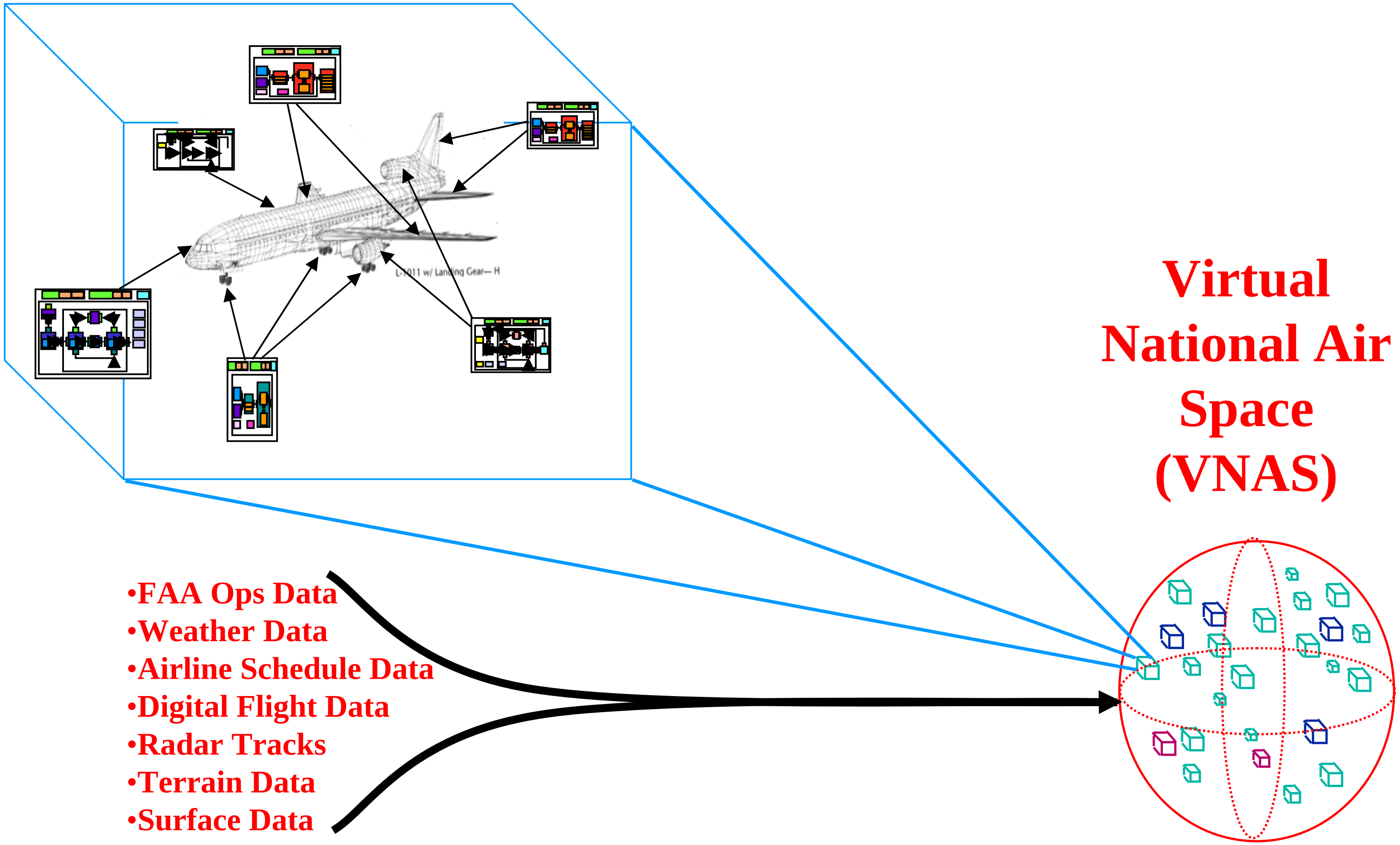


Whole system simulations are produced by coupling all of the sub-system simulations

IPG Baseline System



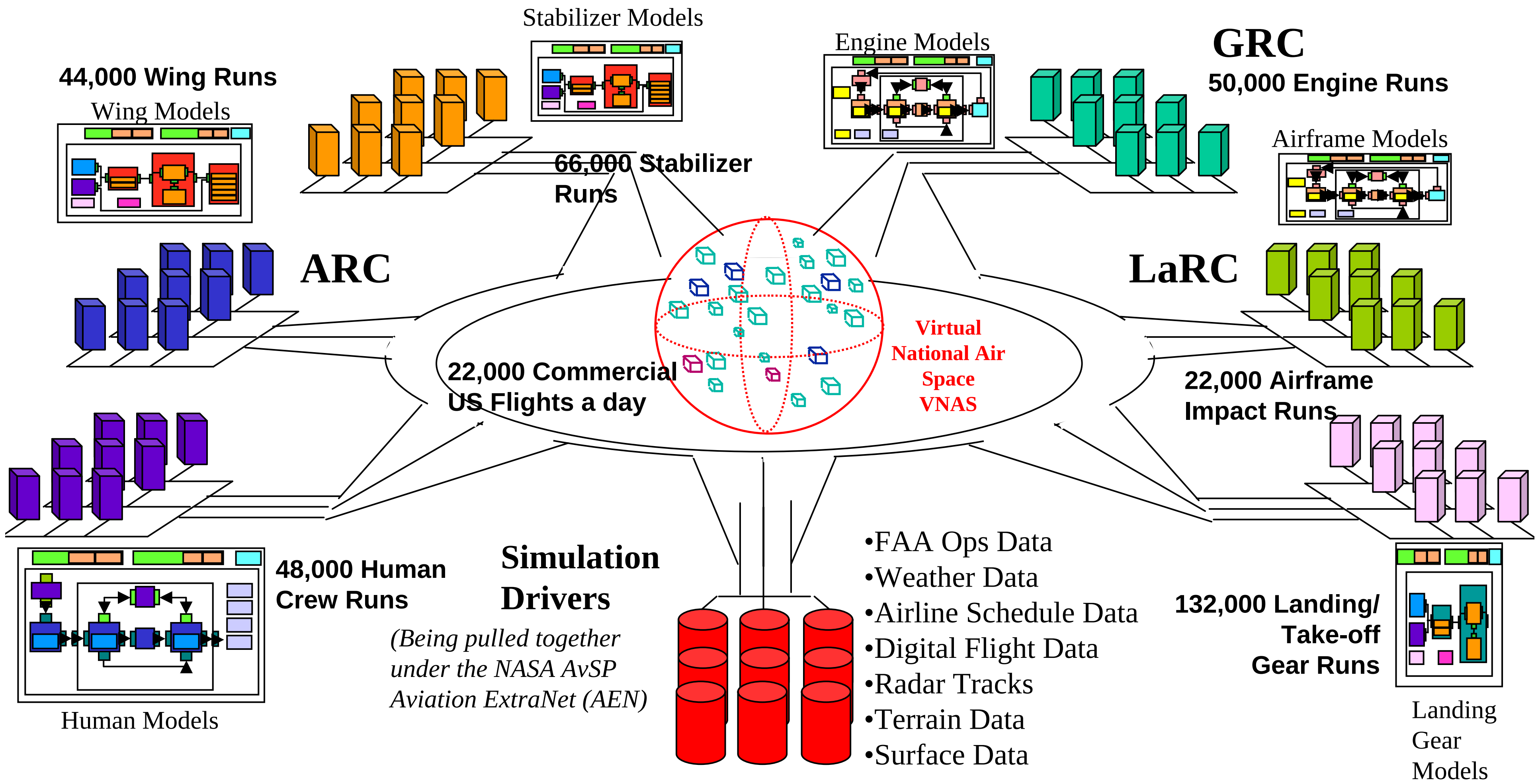
Multi-disciplinary Simulations



Simulated aircraft are inserted into a realistic environment, which requires adding many types of operations data to the systems simulation

Multi-disciplinary Simulations

National Air Space Simulation Environment



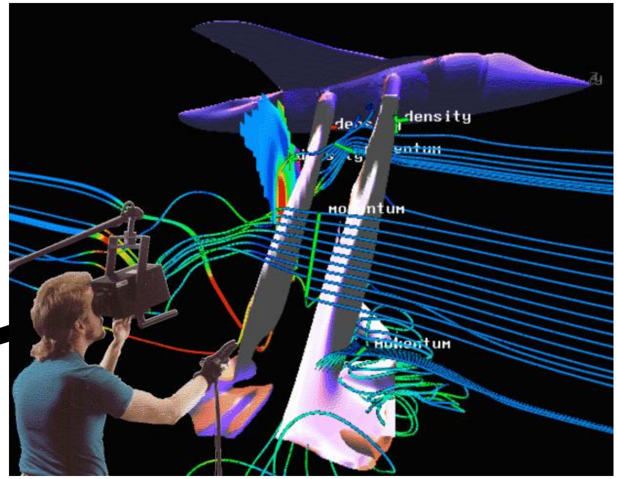
Many aircraft, flight paths, airport operations, and the environment are combined to get a virtual national airspace

Online Instrumentation

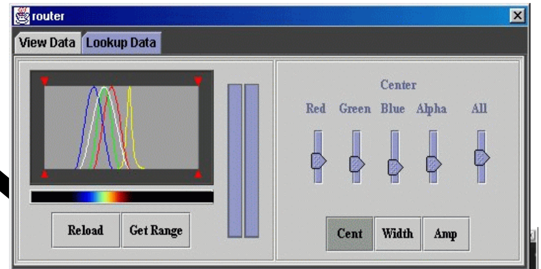
Unitary Plan Wind Tunnel

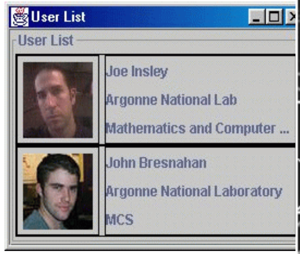
real-time collection

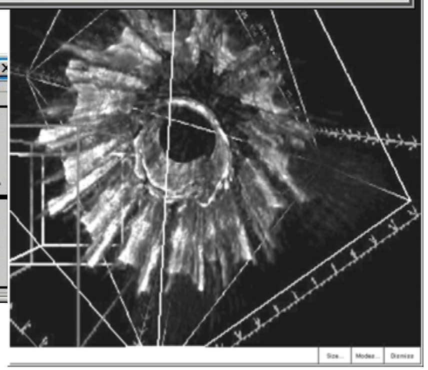
multi-source data analysis, desktop & VR clients with shared controls



real-time experiment control







computer simulations

archival storage

NASA IPG Resources

- ***Computing resources***:
  - ~800 CPU nodes in half a dozen SGI Origin 2000s and several workstation clusters at Ames, Glenn, and Langley, and ~200 nodes in a Condor pool

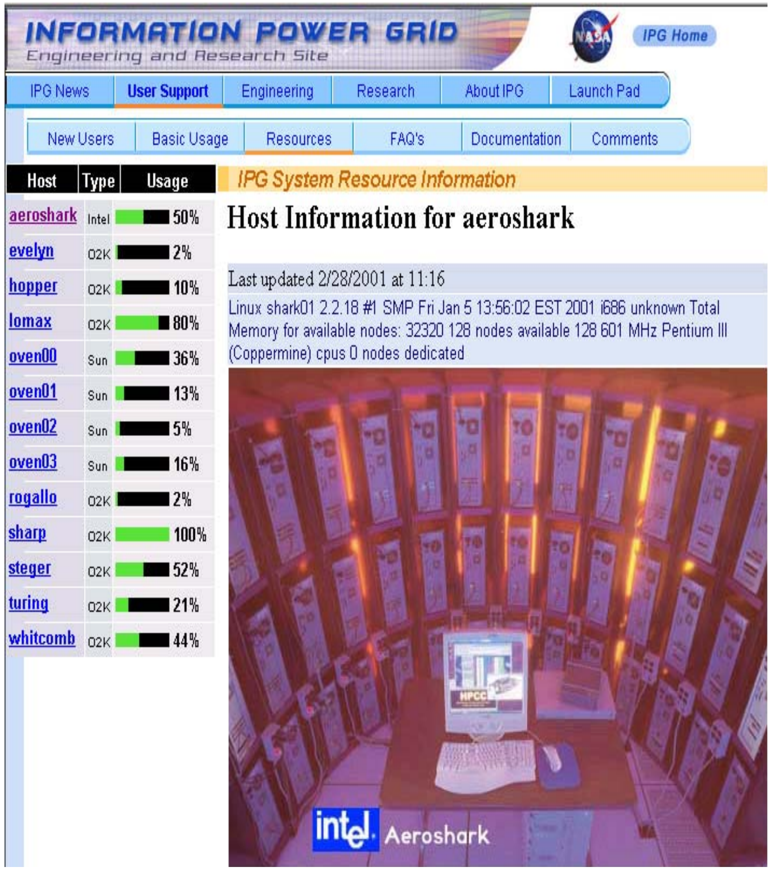
- ***Storage resources***:
  - 50-100 Terabytes of archival information/data storage uniformly and securely accessible from all IPG systems via MCAT/SRB and GSI/Gridftp

➢ **Globus provides Grid common services**

Grid portals used to provide resource access and information

Grid portals provide:

- submission, tracking, and management of jobs running on IPG resources
- integration of IPG services with the user desktop environment,
- access to persistent user profiles.



Grid Computing Today

DISCOM

SinRG

APGrid

IPG …

US Grid Projects (1)

- NASA Information Power Grid
- DARPA CoABS Grid
- DOE Science Grid
- NSF National Virtual Observatory
- NSF GriPhyN
- DOE Particle Physics Data Grid
- NSF DTF TeraGrid
- DOE ASCI DISCOM Grid

US Grid Projects (2)

- DOE Earth Systems Grid
- DOE FusionGrid
- NEES Grid
- NIH BIRN
- NSF iVDGL
- NSF GEON
- NSF NEON
- …..

EU Grid Projects (1)

- DataGrid (CERN, ..)
- EuroGrid (Unicore)
- Damien (Metacomputing)
- DataTag (TransAtlanticTestbed, …)
- Astrophysical Virtual Observatory
- GRIP (Globus/Unicore)
- GRIA (Industrial applications)
- GridLab (Cactus Toolkit, ..)
- CrossGrid (Infrastructure Components)
- EGSO (Solar Physics)

EU Grid Projects (2)

- GridStart (Coordination, …)
- FLOWGRID (CFD VO)
- OPENMOL (Chemistry, Pharma, ..)
- GRACE (Distributed Search, …)
- COG (Industrial ontolgies, ..)
- MOSES (Knowledge Grid)
- BIOGRID (Biotechnology industry)
- GEMSS (e-Healthcare, ..)
- SELENE (Metadata, P2P, ..)

National Grid Projects (1)

- UK - e-Science Grid
- Netherlands – VLAM-G, DutchGrid
- Germany – UNICORE Grid, D-Grid
- France – Etoile Grid
- Italy – INFN Grid
- Eire – Grid-Ireland
- Scandinavia - NorduGrid
- Poland – PIONIER Grid
- Hungary – DemoGrid
- ….

National Grid Projects (2)

- Japan – JpGrid, ITBL
- South Korea – N*Grid
- Australia – Nimrod-G, ….
- Thailand
- Singapore
- AsiaPacific Grid ?
- Pragma?
- …

Community Grid Model

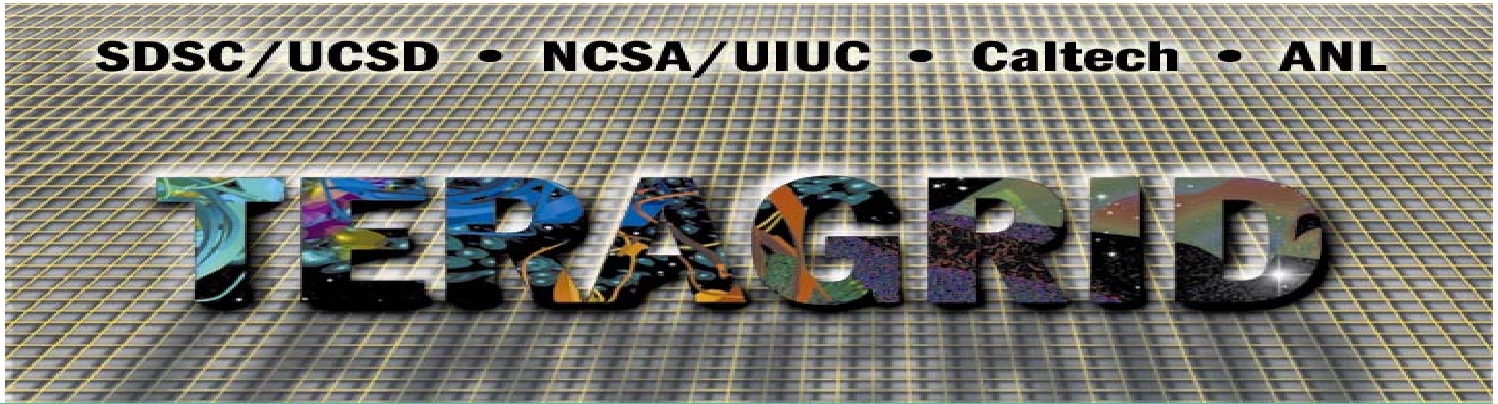
Grid Applications

User-focused and targeted grid middleware, tools, and services

Common Infrastructure layer (NMI, GGF standards, OGSA etc.)

Grid Resources

The NSF TeraGrid Project



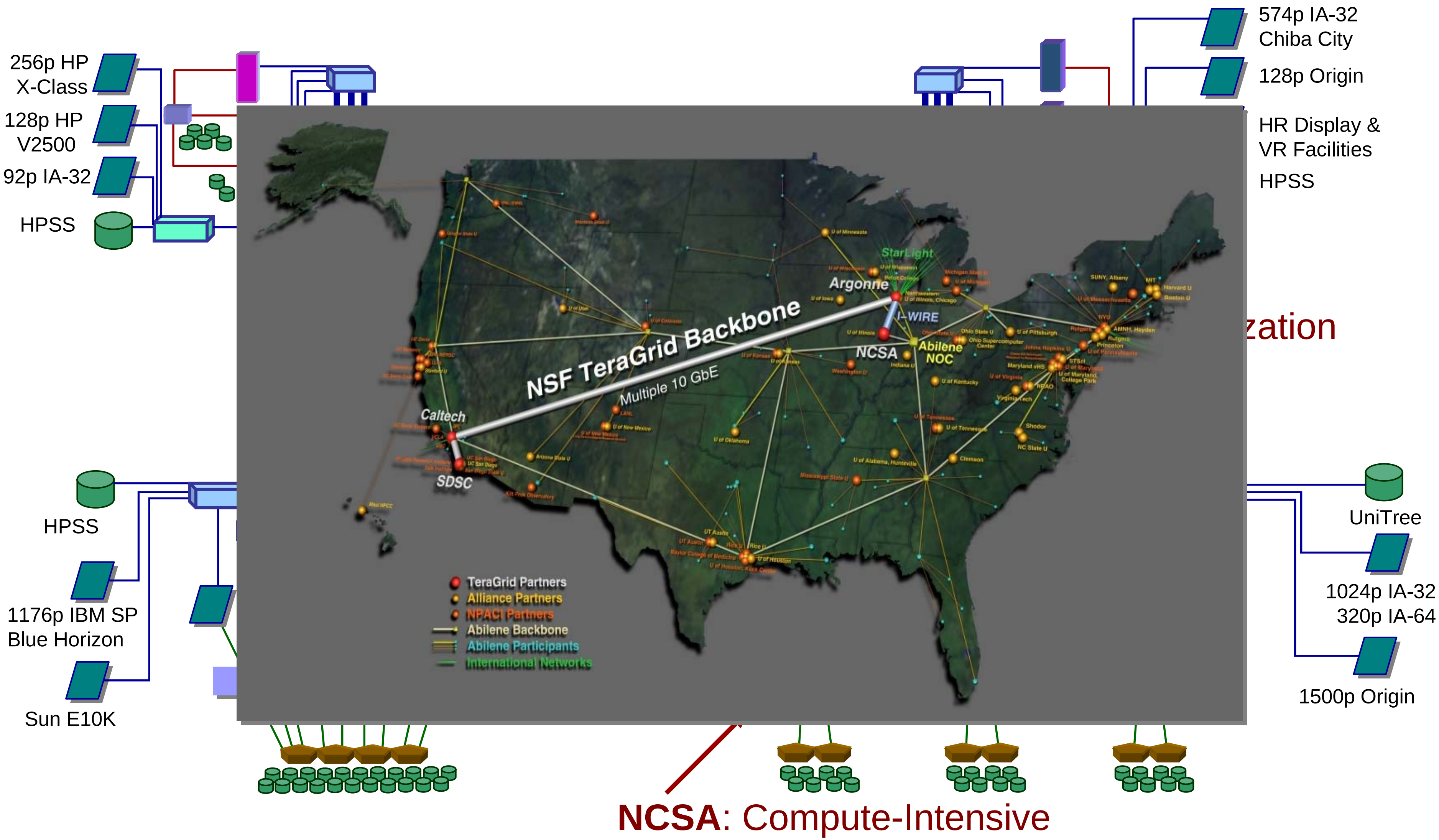
TeraGrid will provide in aggregate

- *13.6 trillion calculations per second*
- *Over 600 trillion bytes of immediately accessible data*
- *40 gigabit per second network speed*

TeraGrid will provide a new paradigm for data-oriented computing

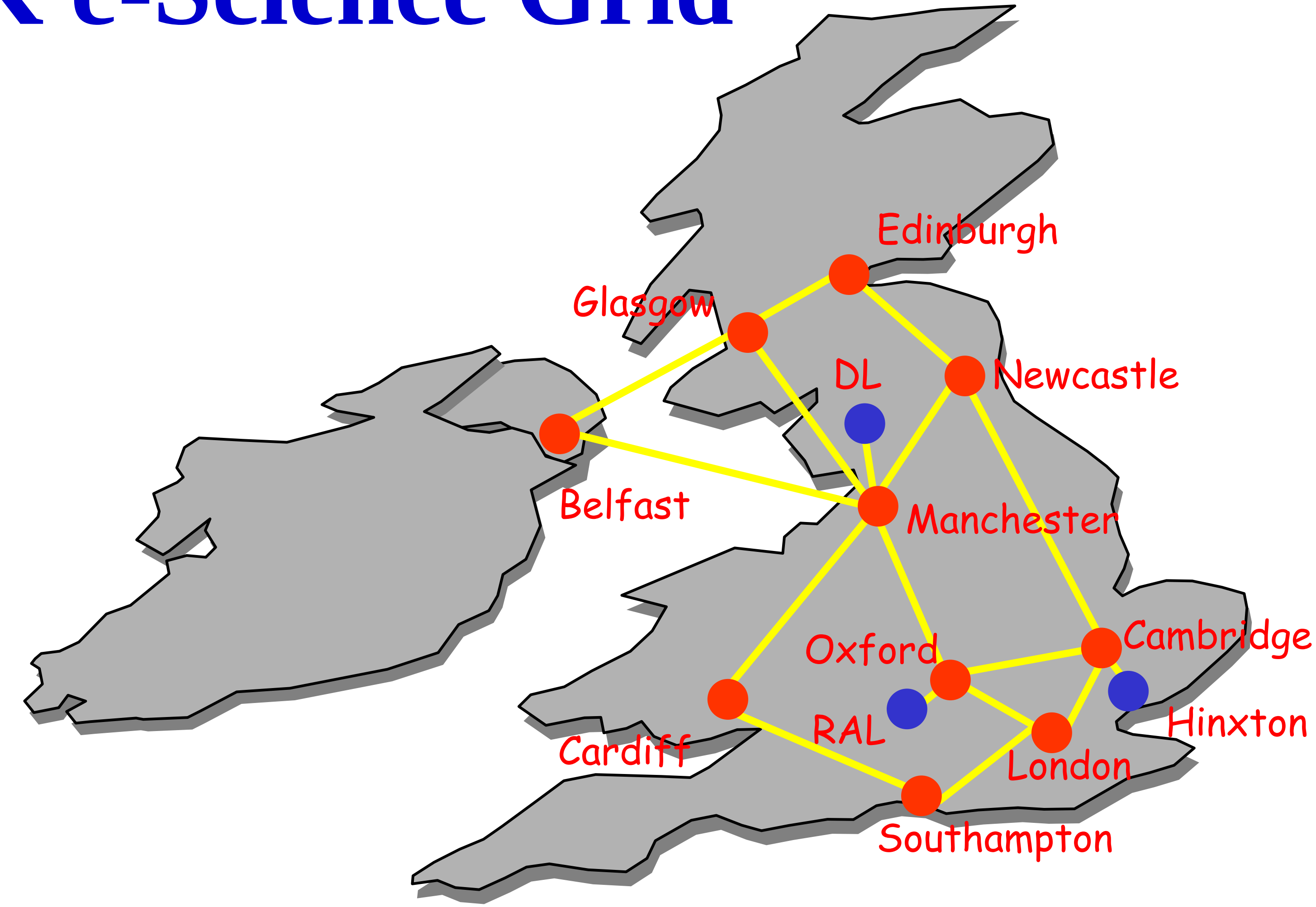
Critical for disaster response, genomics, environmental modeling, …

TeraGrid

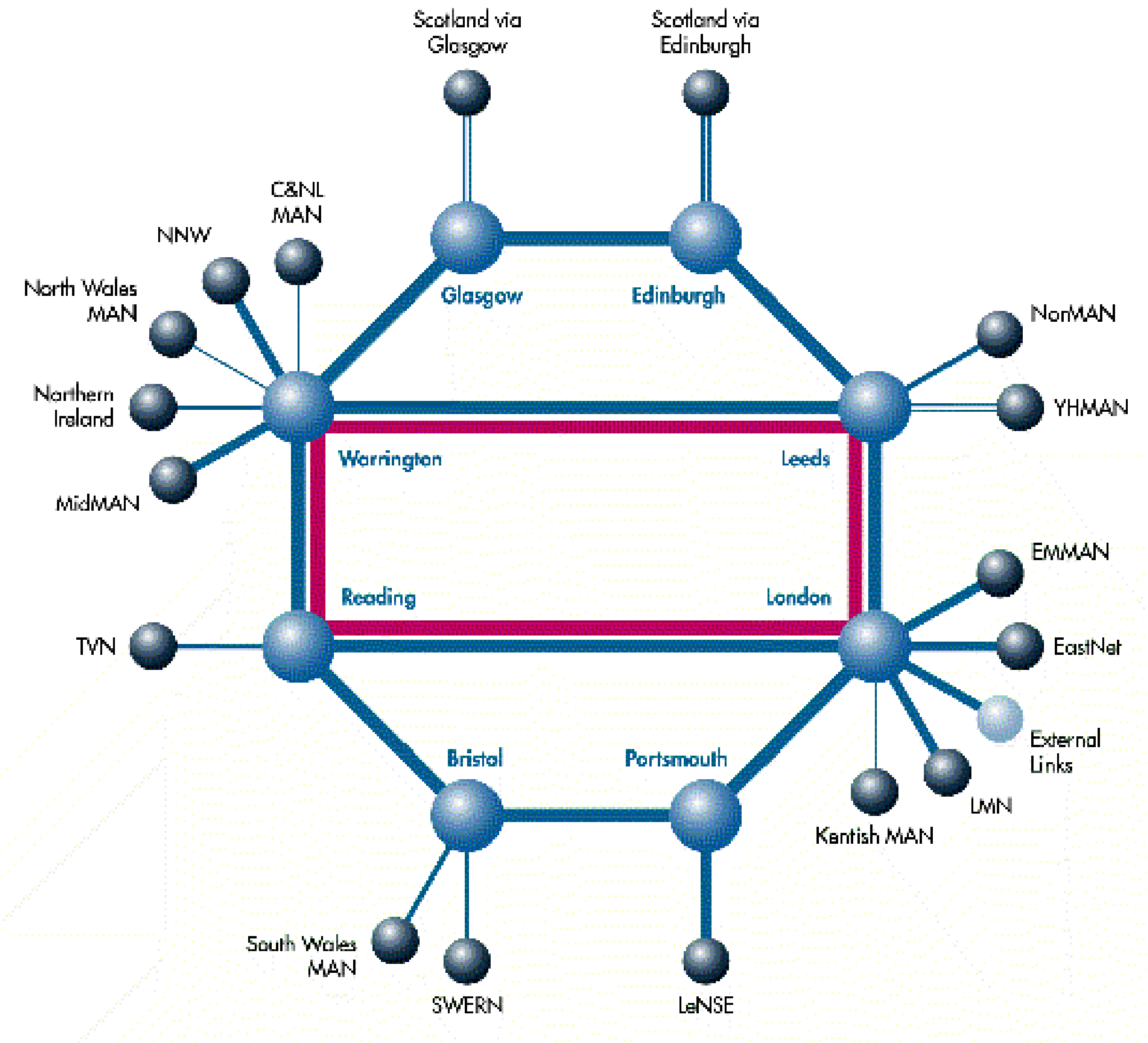


256p HP X-Class
128p HP V2500
92p IA-32
HPSS
574p IA-32 Chiba City
128p Origin
HR Display & VR Facilities
HPSS
zation
HPSS
1176p IBM SP Blue Horizon
Sun E10K
UniTree
1024p IA-32 320p IA-64
1500p Origin
NSF TeraGrid Backbone
Multiple 10 GbE
Caltech
SDSC
Argonne
StarLight
I-WIRE
NCSA
Abilene NOC
TeraGrid Partners
Alliance Partners
NPACI Partners
Abilene Backbone
Abilene Participants
International Networks
NCSA: Compute-Intensive

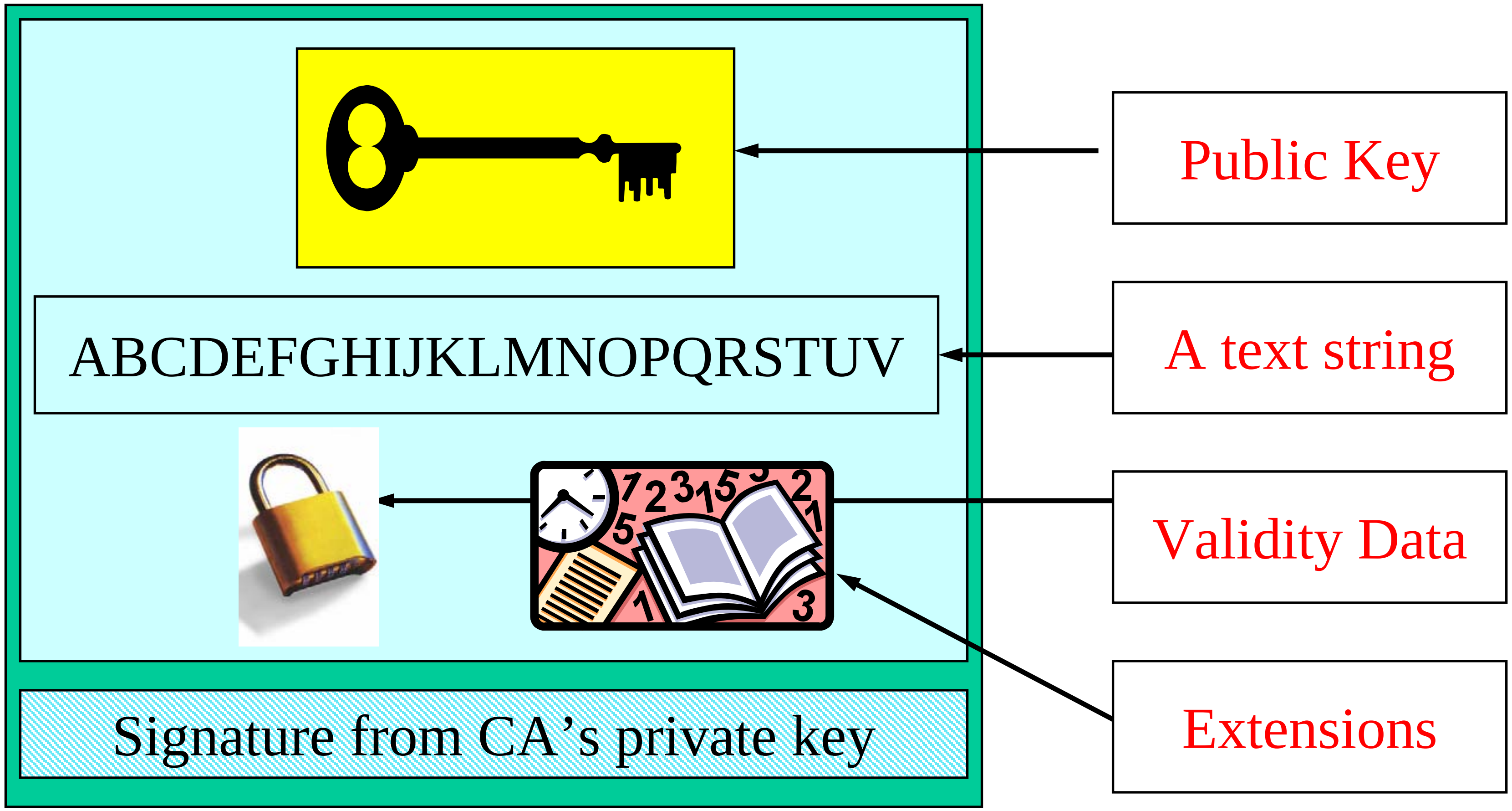
UK e-Science Grid



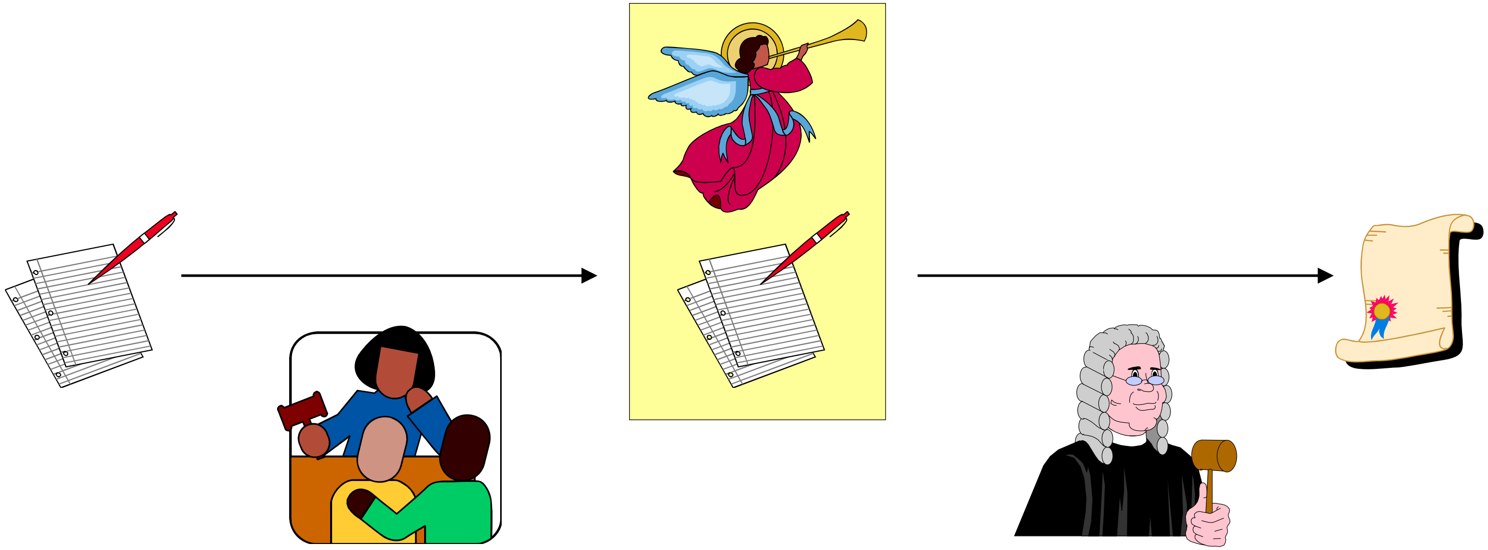
SuperJANET4



Anatomy of a Digital Certificate

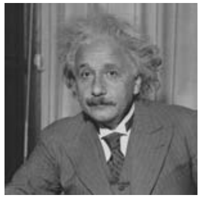


How a certificate is issued

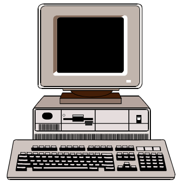


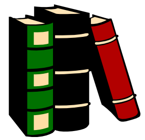
- The Registration Authority (RA) approves a request for a certificate. The RA is local to the users.
- The CA then issues the corresponding certificate.

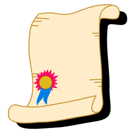
How does it work?



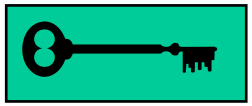
1. Scientist wishes to access a resource, so he sends a copy of the certificate to the resource







2. Resource says: prove it's your certificate



Private Key

Challenge

Response

3. Scientist proves that he has the corresponding private key

4. Resource is convinced that scientist is who he claims to be and decides to give him access

Service based architecture

Find them

Publication, registration, discovery, matchmaking, deregistration

Run them

Execution, monitoring, exception handling

Organise them

Interoperation, composition, substitution

Web Services

- Loosely Coupled Distributed Computing
  - Think Java RMI or C remote procedure call
- Text Based Serialization
  - XML: “Human Readable” serialization of objects
- IBM and Microsoft Lead
  - Initial draft with joint copyright.
  - W3C Standardization.
- Three Parts
  - Messages (SOAP: Service Oriented Access Protocol)
  - Definition (WSDL: Web Services Description Language)
  - Discovery (UDDI: Universal Description, Discovery, and Integration)

Open Grid Services Architecture

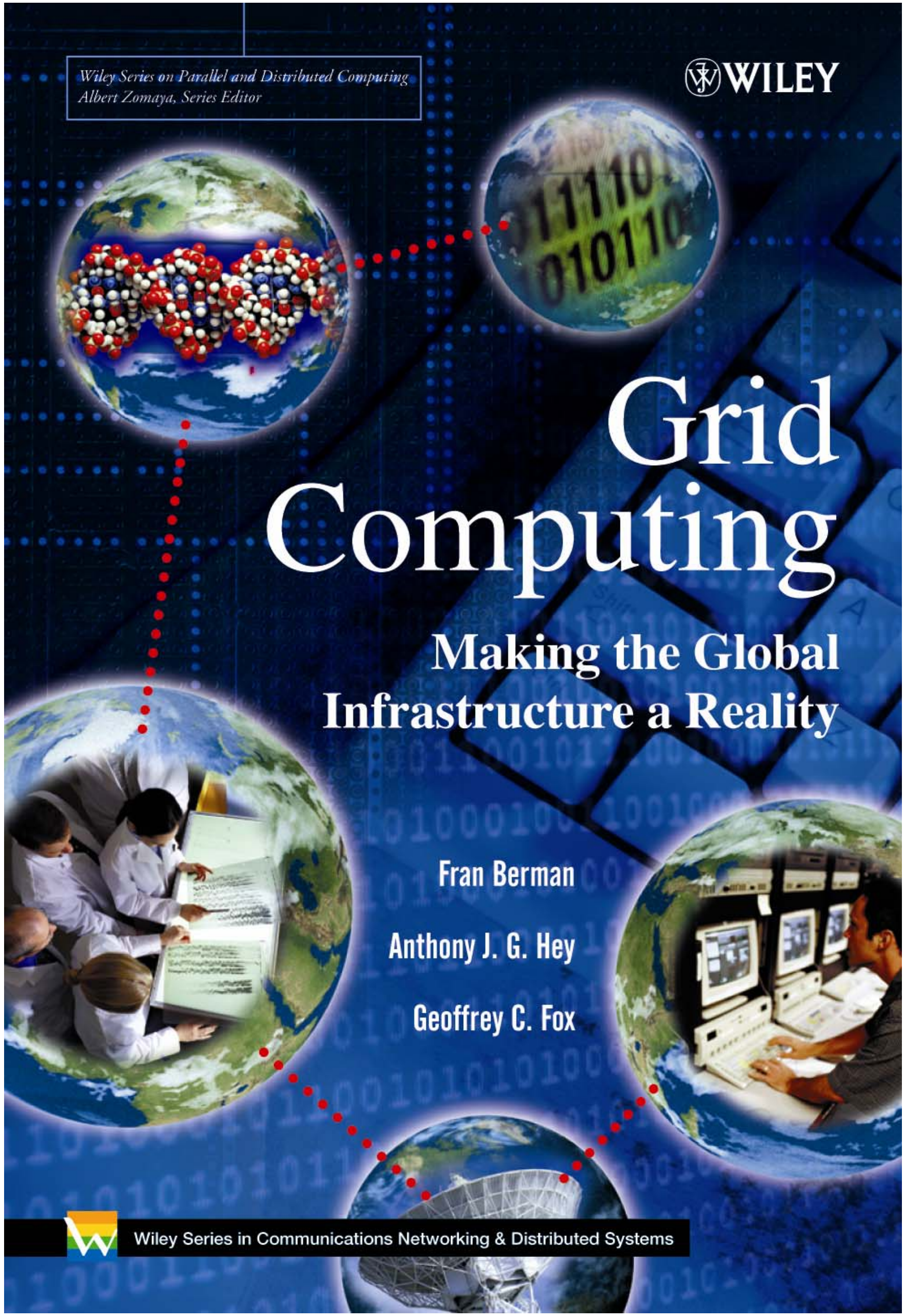
- Development of Web Services with Globus
- OGSI will provide low level services:

  Naming /Authorization / Security / Privacy/…

➢ Projects looking at higher level services:

  Workflow/Transactions/DataMining/Knowledge Discovery/…

➢ Exploit Synergy: Commercial Internet with Grid Servic


Wiley Series on Parallel and Distributed Computing
Albert Zomaya, Series Editor
WILEY
Grid Computing
Making the Global Infrastructure a Reality
Fran Berman
Anthony J. G. Hey
Geoffrey C. Fox
Wiley Series in Communications Networking & Distributed Systems