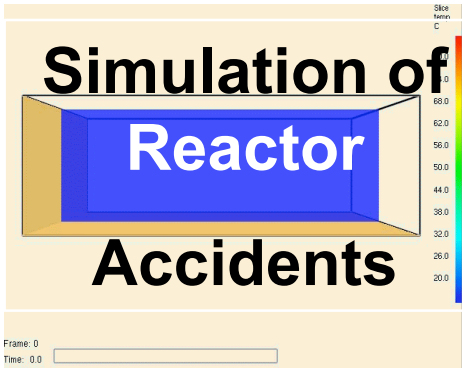
GRID INITIATIVES IN DAE

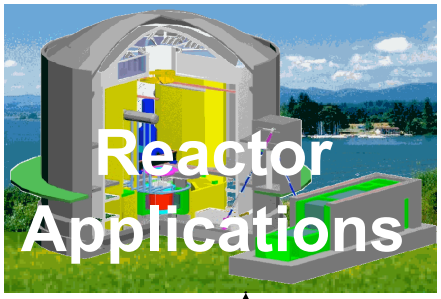
P.S.Dhekne

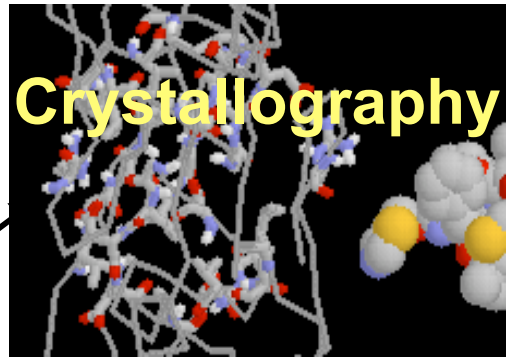
Computer Division, BARC

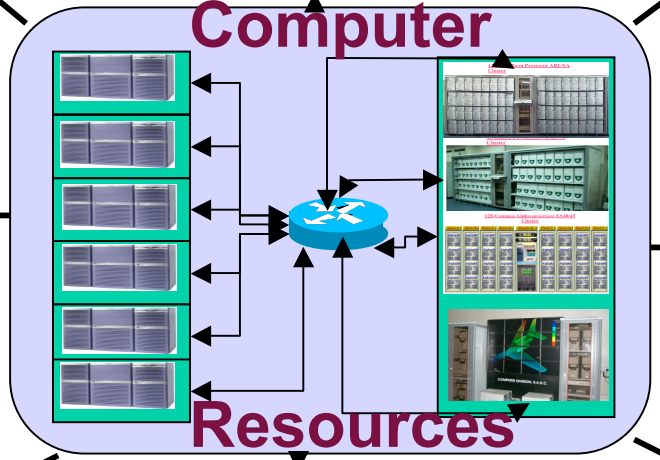
dhekne@barc.ernet.in

High-end Computing Applications

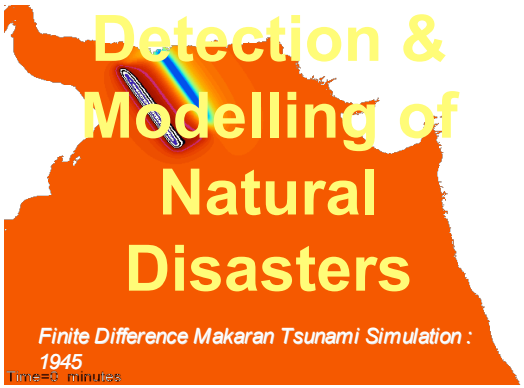




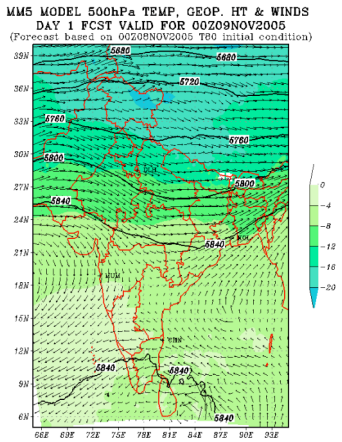




Physics Applications

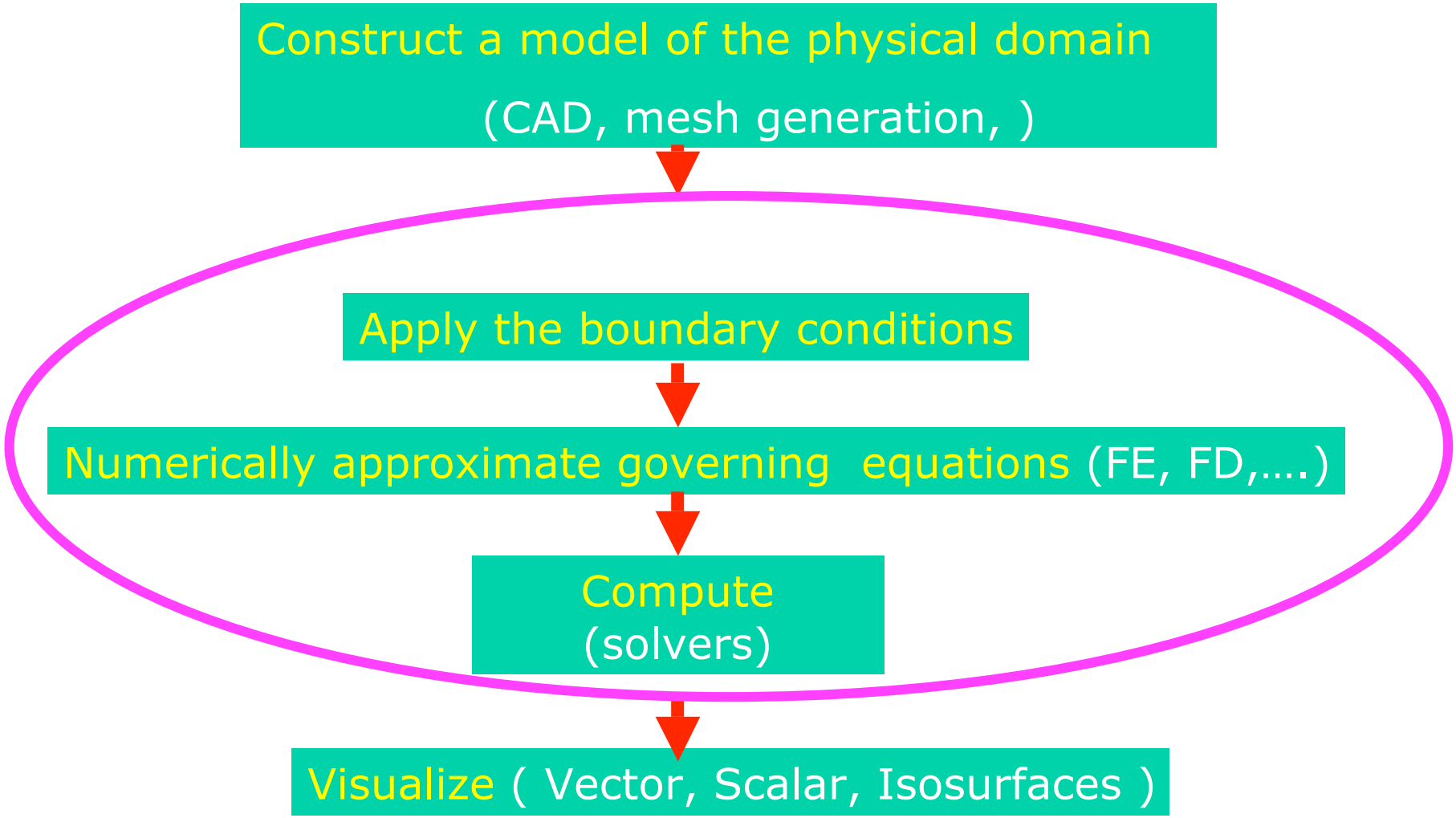


Material Characterisation



Weather Forecasting

Computational Pipeline



Growth of Computing power in BARC

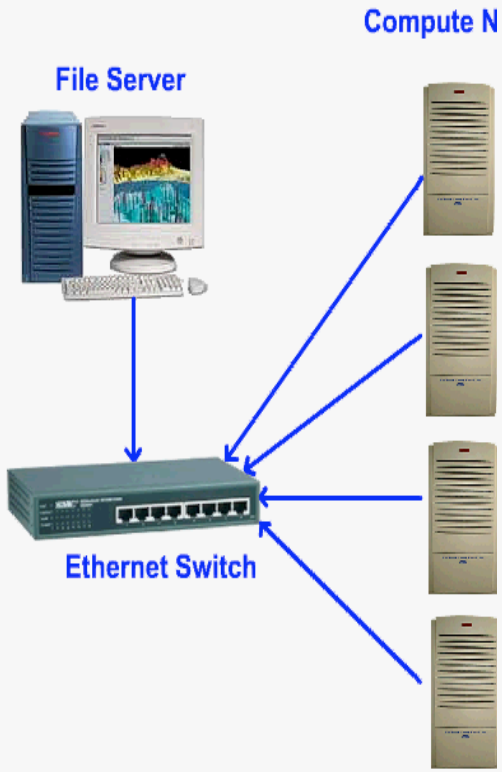
Parallel Cluster Computers

Use locally available commodity hardware

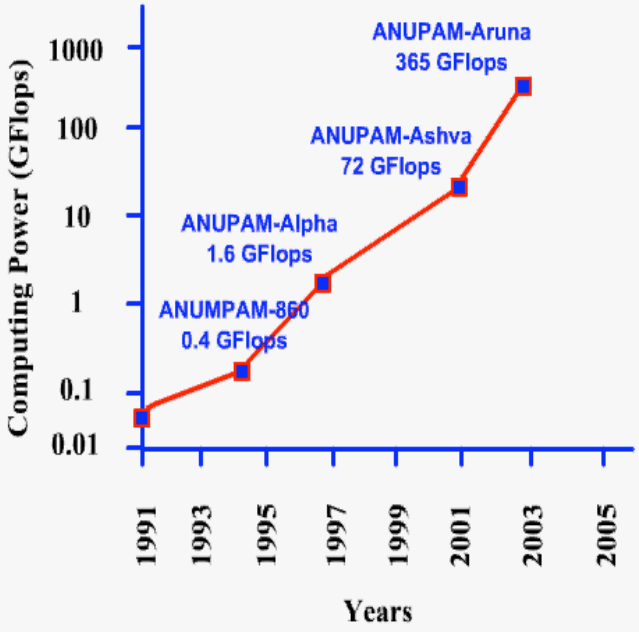
Open source software

Inhouse monitoring management and development tools

Cheaper alternative for supercomputing



ANUPAM Series of BARC Supercomputers



Computing Power increased by order of 1000 from ANUPAM-860 to ANUPAM-ARUNA

ANUPAM-Ashva 72 GFlops

ANUPAM-Aruna 365 GFlops

Performance : Parallel Computers

- Earth Simulator, Japan: 35.86 TFLOPS
- ASCI Q, LANL, USA: 7.22 TFLOPS
- TeraScale Cluster, Virginia Tech University: 11 TFLOPS
- IMSc, Chennai: KABRU 1 TF
- C-DAC Param: 535 GFLOPS
- BARC Anupam: > 2 TFLOPS
- HRI Cluster 480 GFlops

- **All Top 500 Supercomputers nowadays are parallel computers**

NEED FOR VISUALIZATION

- Computational methods are being used for simulation & modeling as an alternative to experimental methods

- Scientific data collected from experiments, data acquisition, computations or sensor networks can be fairly huge in volume & is impossible to be analyzed using conventional methods

- Object oriented graphics or visual computing thus has emerged

- This requires high graphics resolution with wall screen display and very high graphics rendering speed

Present shortcomings

- Limited resolution

Complex geometric models with millions of triangles are rendered into a window of 500x500 pixels then at most 25 % of those triangles could possibly contribute to the final image. By increasing display window resolution more triangles can contribute to the rendered image, thereby adding further details and making rendered image more realistic.

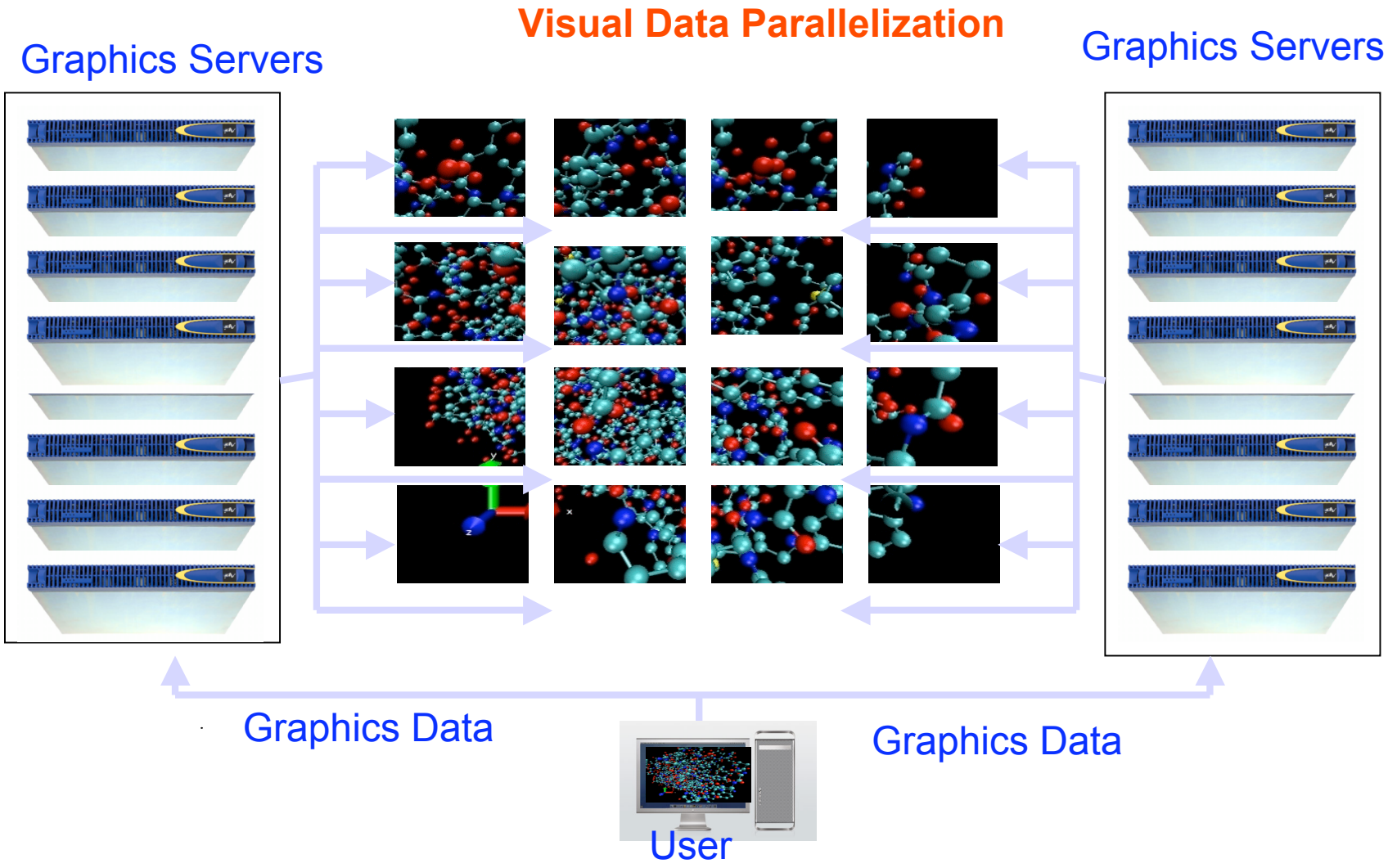
- Rendering power

Many graphics applications still can not run at acceptable rate. Rendering power available is not sufficient to drive high resolution display at interactive rates. Rendering is compute intensive.

- The display resolution is lagging far behind

The resolutions of the displays have been increasing only at an annual rate of 5% for the last two decades and also graphics accelerator cards.

20-MegaPixel Tiled Display System

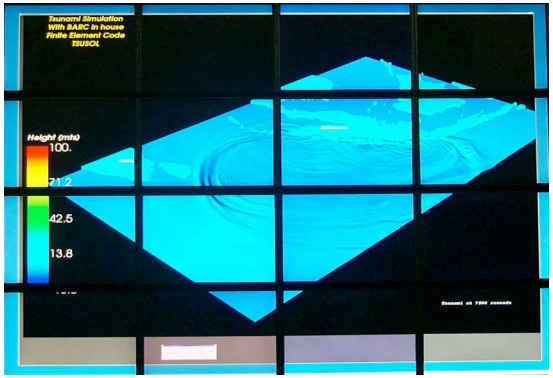


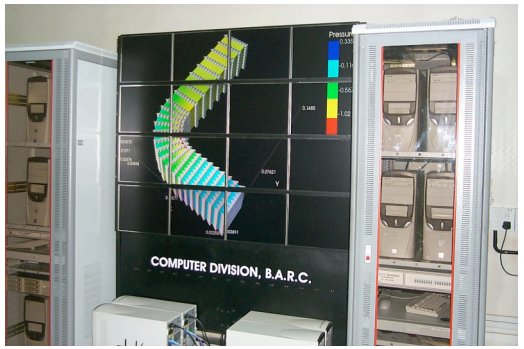
High Performance Visualization Cluster

First of its kind in the country

Tiled display giving very high resolution (20Mpixel), high-speed rendering needed for scientific visualization

- Tiled 4x4 LCD Panels
- 1 Master Client, 16 Graphics Servers
- 5120 x 4096 total resolution
  1280x1024 per LCD 20 Million Pixels
- 16 Times rendering speed





Exploiting parallelism for

- Processing (parallelization of computation)
- I/O (parallel file system)
- Visualization (parallelized graphic pipeline/ Tile Display Unit)

Has given high compute and high throughput Computational system

Clusters :Primary IT infrastructure

Clustering is replacing all traditional Computing platforms and can be configured depending on the method and applied areas

- **LB Cluster -** *Network load distribution and LB*
- **HA Cluster -** *Increase the Availability of systems*
- **HPC Cluster (Scientific Cluster) -** *Computation-intensive*
- **Web farms -** *Increase HTTP/SEC*
- **Rendering Cluster** – *Increase Graphics speed*

HPC : High Performance Computing HA : High Availability LB : Load Balancing

OSS:New Direction in Cost Saving

- The ideas behind “open source” and “free software” are not new
  - Lots of software has been free from the beginning
- What is new is the emergence of business and/or political movements to define, advocate, and institutionalize these notions and OSS is the only way
- Dedicated efforts are needed to development software
  - Auto installation, Operation, Monitoring, Debugging
  - Versioning, Performance monitoring Tools
- Software evolves slowly and is manpower intensive

National Collaborations

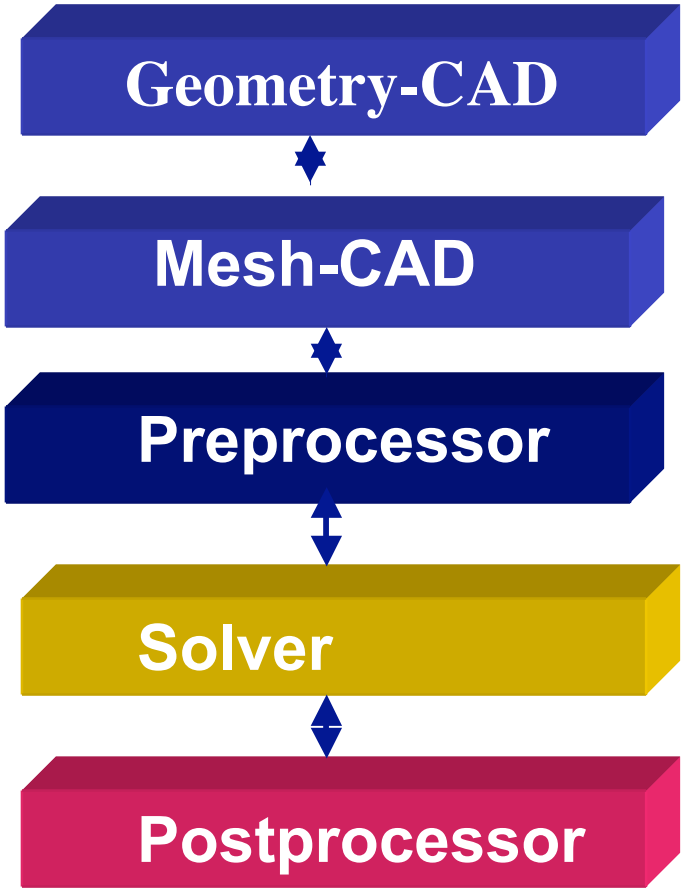
- MOU with NCMRWF, New Delhi for developing parallel version of weather models on ANUPAM
- MOU with NIC & VSSC for Collab-CAD software development
- MOU with VJTI, Mumbai (Solve CFD related computational problems)

Projects through BRNS

- IIT, Mumbai (Parallel Data Base, Parallel File System)
- IIT, Kanpur (Parallel CFD Solver)
- IISC, Banglore (Wormhole Routing, Parallel Compiler)
- SNDT, Mumbai (Parallelization of Algorithms)

Consortium approach for Software

Collaborative Software Architecture based on Free & Open Source Software components (FOSS)



Collab-CAD, NIC, DIT

IITB

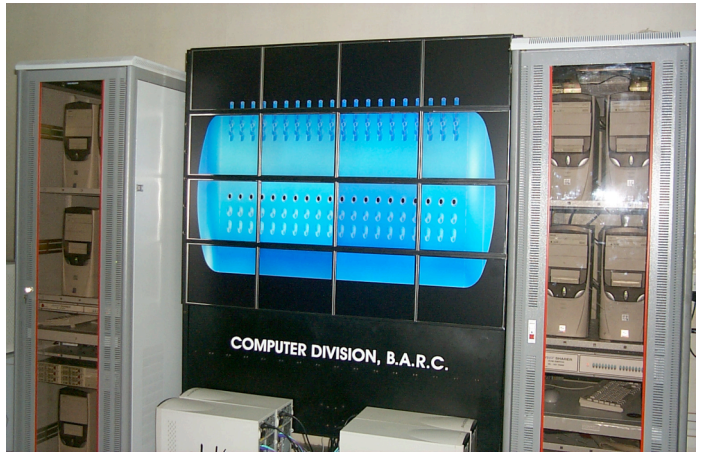
IITK

CD, BARC

Use Compatible Interfaces

- CGNS
- IGES

Snapshots of CollabCAD, 3D Java Viewer



Scientist also now need Pre & Post for scientific applications

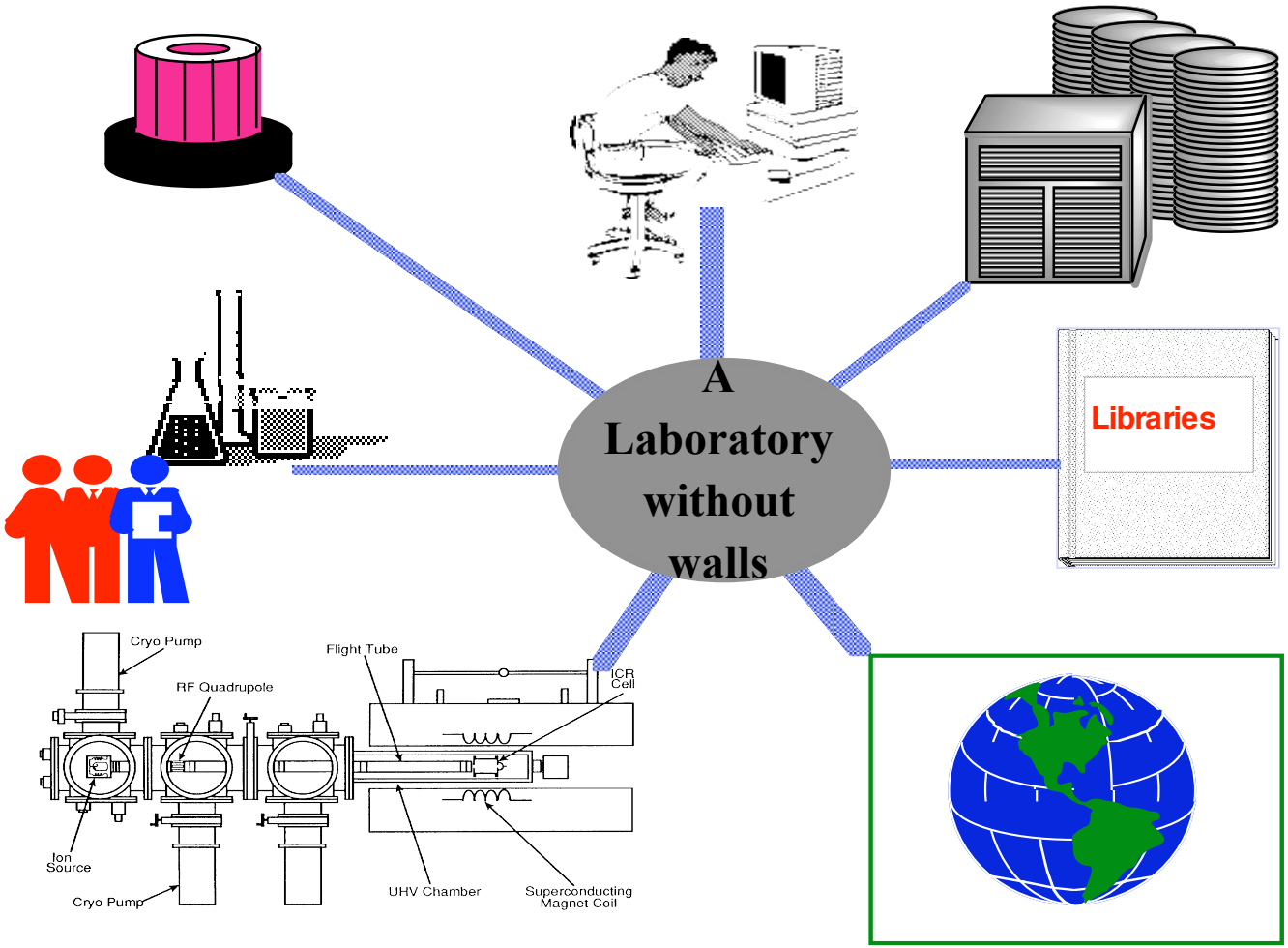
Issues in operating large clusters

- Space Management
  - Node form factor (Desktop, rack mounted 1U/2U)
  - Layout of the nodes
  - Cable routing and weight
- Power Management (new motherboards)
- Cooling arrangements
- Cost of ownership is not really low
- Performance/Price/Power Consumption

INTERNET COMPUTING

- Today you can't run your jobs on the Internet
- Internet Computing using idle PC's, is becoming an important computing platform (Seti@home,Napster,Gnutella,Freenet, KaZak)
  - www is now a promising candidate for core component of wide area distributed computing environment.
  - Efficient Client/server models & protocols
  - Transparent networking, navigation & GUI with multimedia access & dissemination for data visualization
  - Mechanism for distributed computing such as CGI.Java
- With improved performance (price/performance) & the availability of Linux, Web Services (SOAP, WSDL, UDDI,WSFL), COM technology it is easy to develop loosely coupled distributed applications

Collaboratory: Working Together Apart



Tools

- Chat
- Email
- Video Conferencing
- VR
- White Boards
-- Web Portal

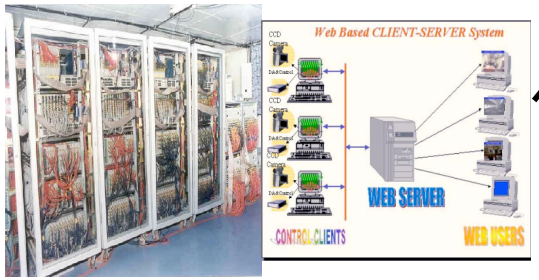
Web based remote monitoring & Control

TELESCIENCE for NRL

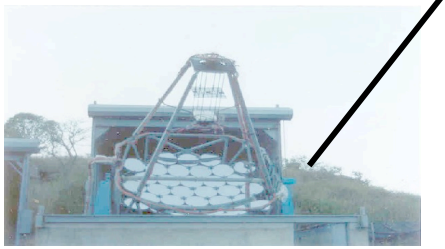
ANUNET with max of 512 Kbps Bandwidth

NRL Laboratory

Simulations
On ANUPAM at
BARC, Mumbai

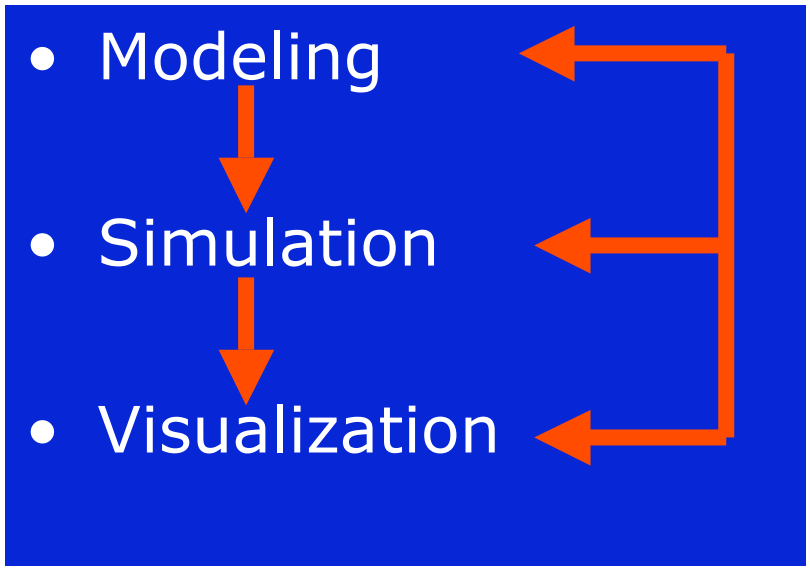


Instrumentation



Remote Control & Mon

Computational Pipeline



Integration

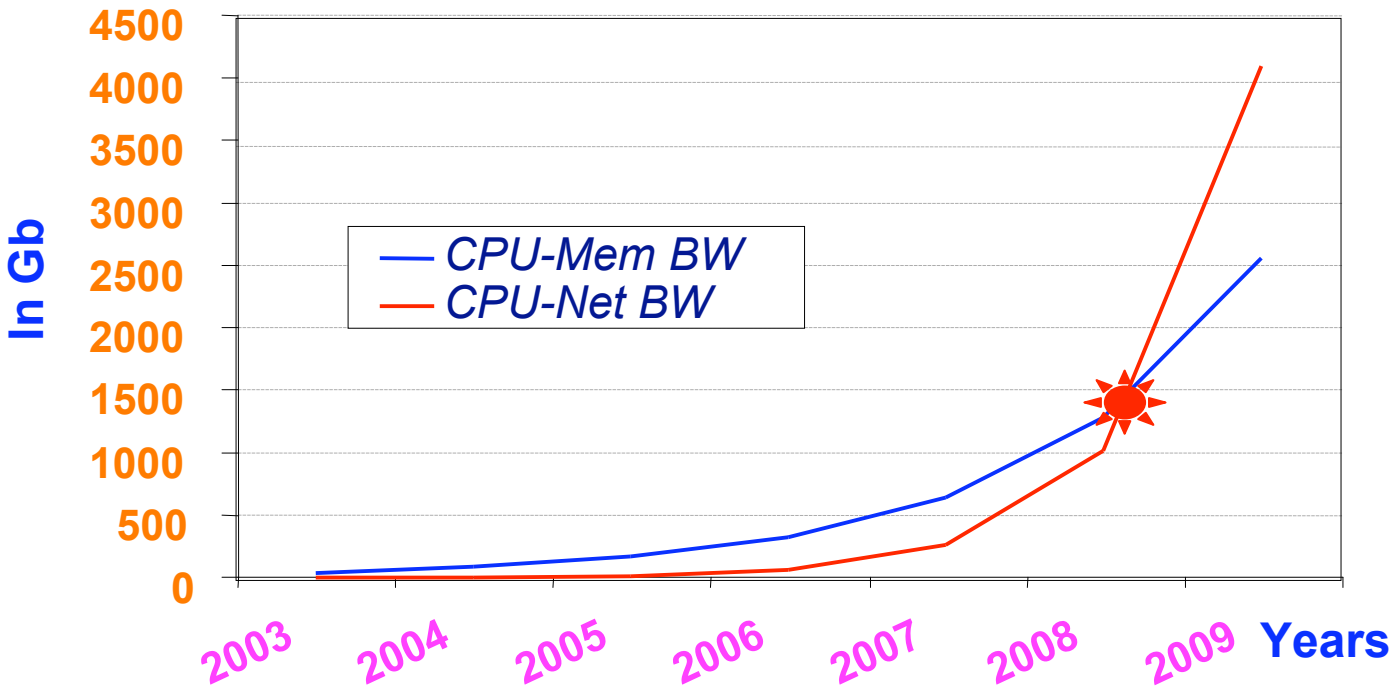
Three different compute systems, different application software for each function

But now user wants integrated solution

Compute-Communication Cross over

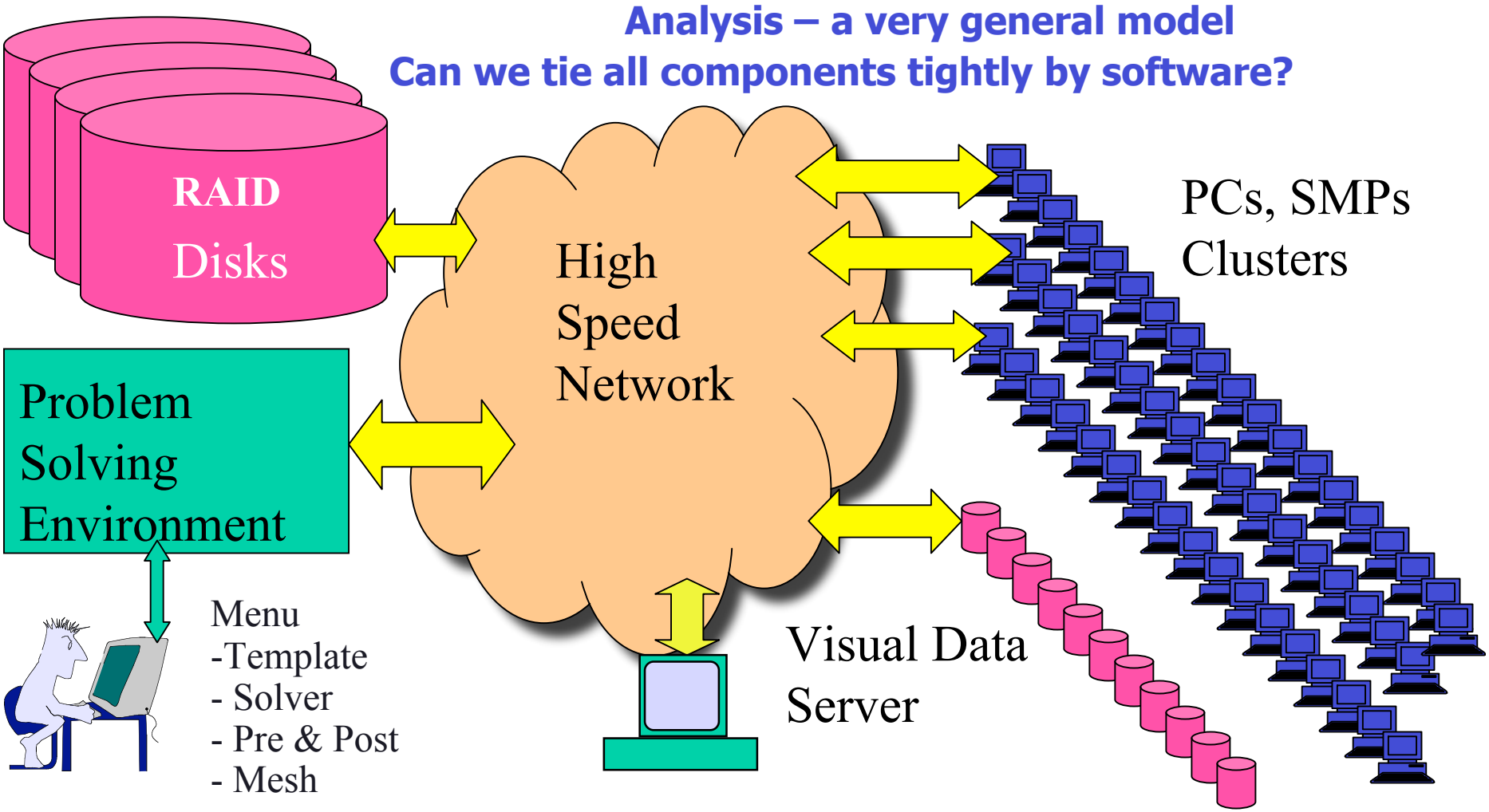
"When the network is as fast as the computer's internal links, the machine disintegrates across the net into a set of special purpose appliances" (George Gilder)

Trend in Technological Change for CPU & Network



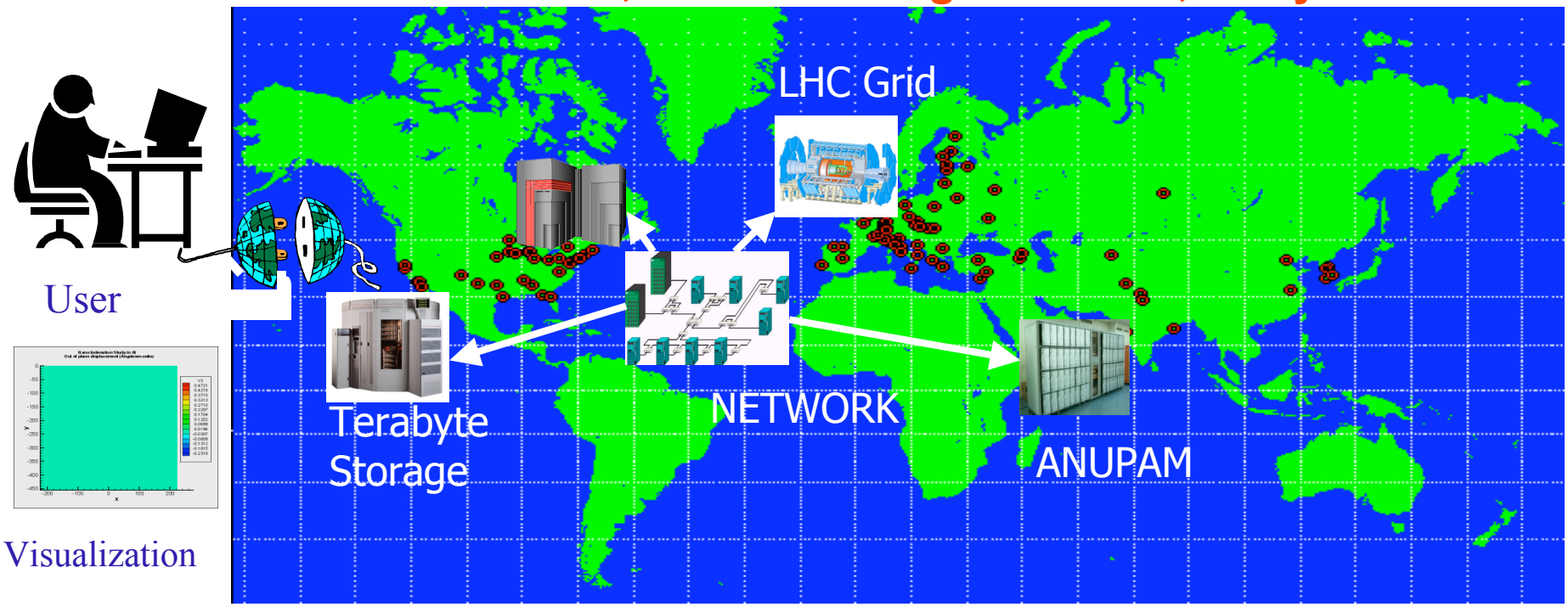
How long Moor's law is going to sustain ?

Computer Assisted Science & Engineering



Virtual Organization : Grids

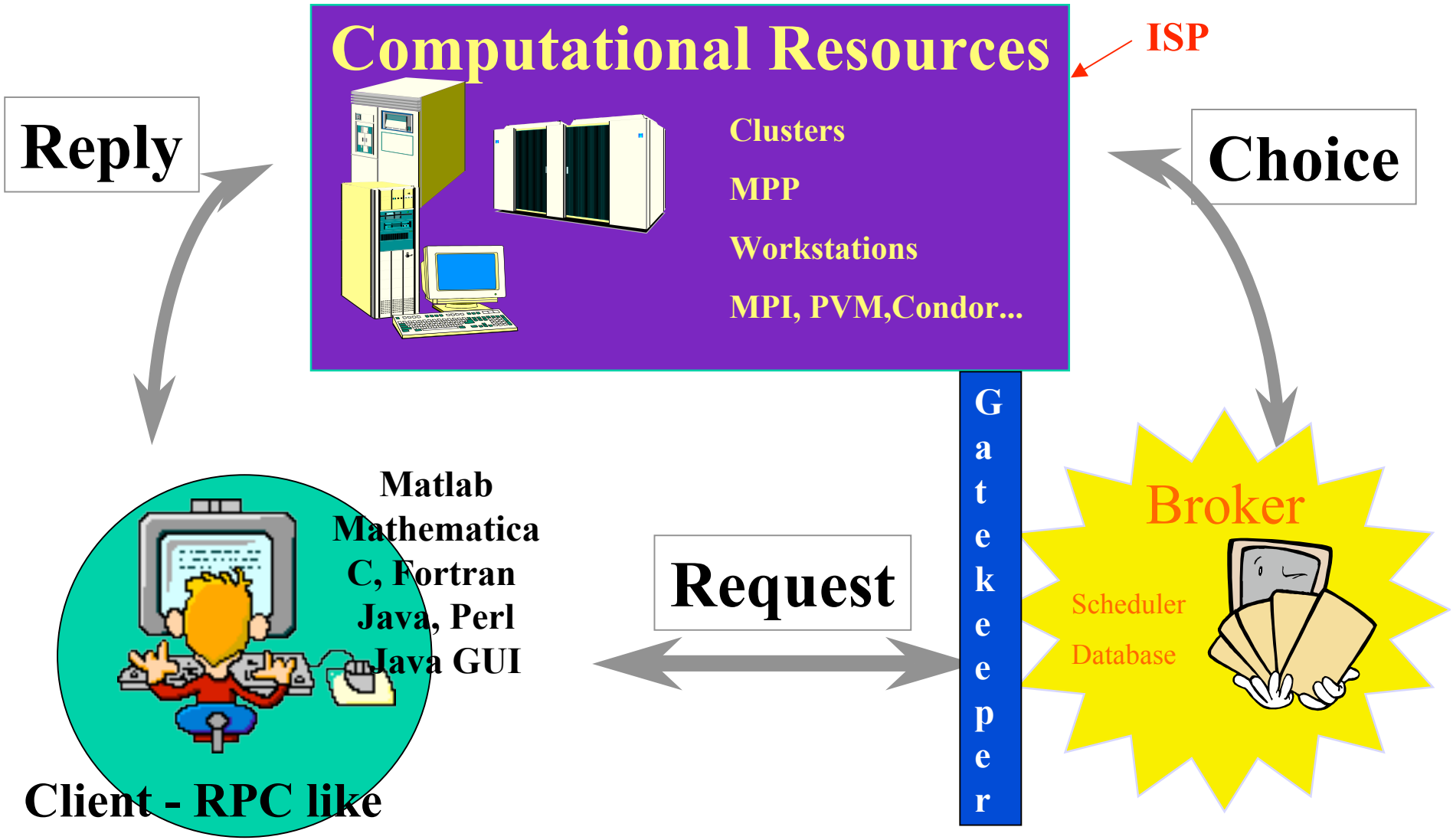
Resource sharing and coordinated problem solving in dynamic, multiple R&D units : Millions of users, Thousand Organizations, Many Countries



Making Information Technology (IT) as easy to use as plugging into electrical or TV socket

Simple philosophy:
Seamless integration of distributed environments to aggregate Computing power, Storage Capacity, Software and Visual Capability

What is needed?



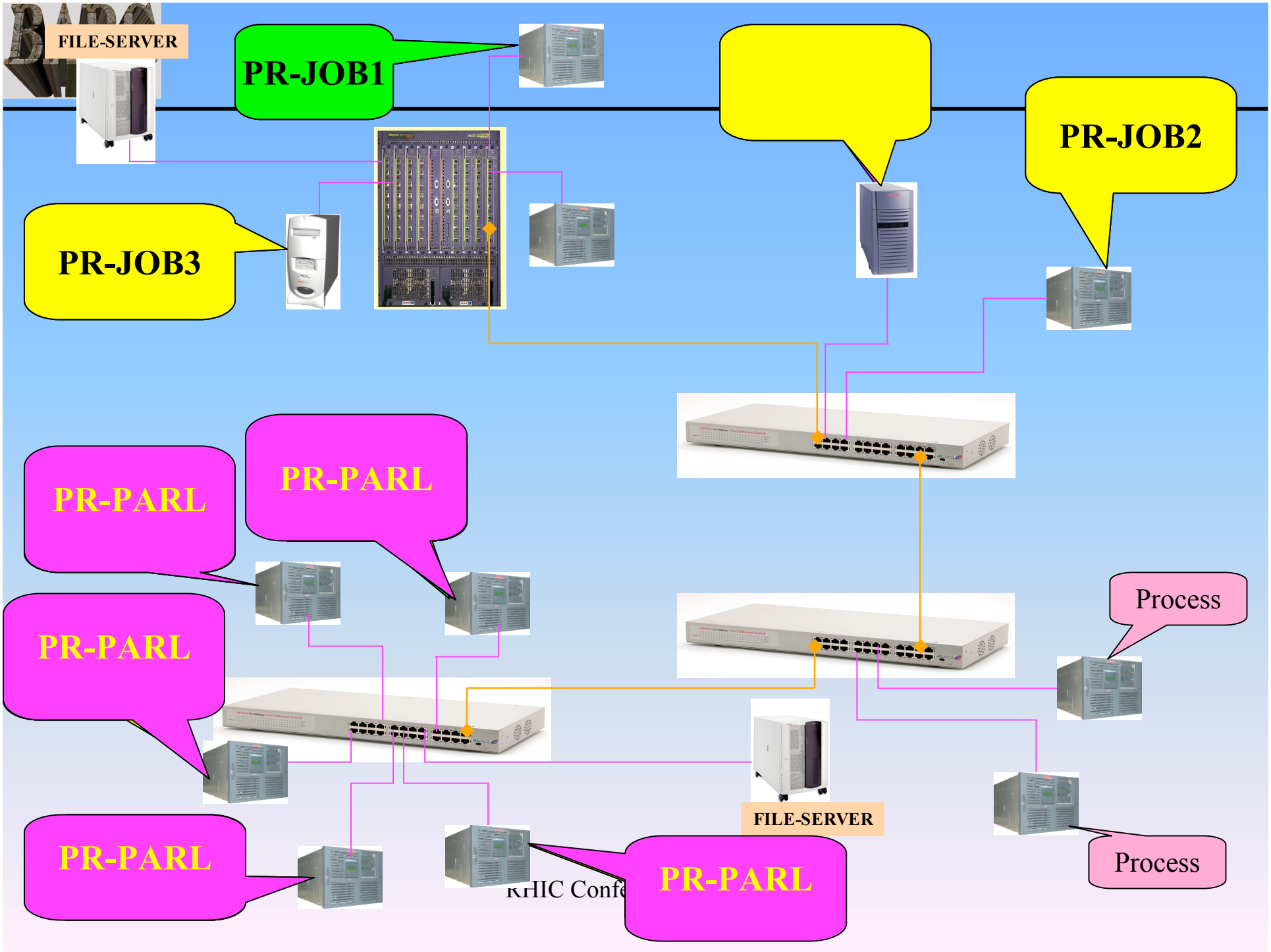
GRID CONCEPT

User Access Point

Result

Resource Broker

Grid Resources


FILE-SERVER
PR-JOB1
PR-JOB2
PR-JOB3
PR-PARL
PR-PARL
PR-PARL
PR-PARL
PR-PARL
Process
Process
FILE-SERVER

LHC Computing

- LHC (Large Hadron Collider) will begin taking data in 2006-2007 at CERN.
- Data rates per experiment of >100 Mbytes/sec.
- >1 Pbytes/year of storage for raw data per experiment.
- Computationally problem is so large that can not be solved by a single computer centre
- World-wide collaborations and analysis.
  - Desirable to share computing and analysis throughout the world.

Data Grids for HEP

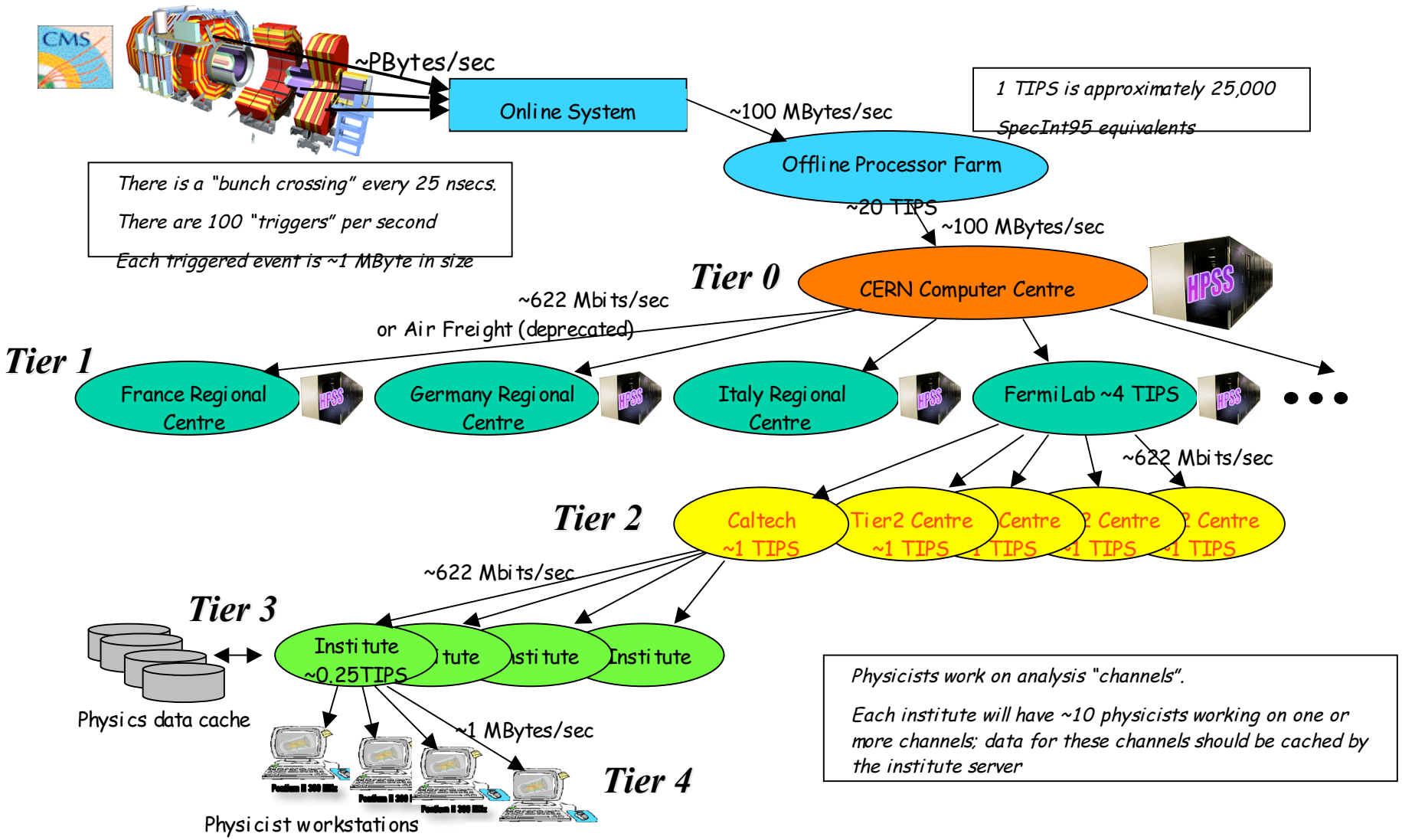


Image courtesy Harvey Newman, Caltech

Why was it not done before?

- Use of Internet as infrastructure has become attractive now
  - **Increasing bandwidth, advanced web services on Internet II and NGI**
- Highly user friendly & mature Web Tech. : **no training needed**
- Provide access via "The Grid" to scarce remote computing resources of millions of PCs from anywhere to anybody
- Computing Problems are getting bigger **(2x each year); need MPP**
- Harnessing the power of Internet to aggregate and share resources spread across the globe is very challenging and highly cost effective and **can give you unlimited computing capability**
- Multiple ways are possible to implement Grid applications

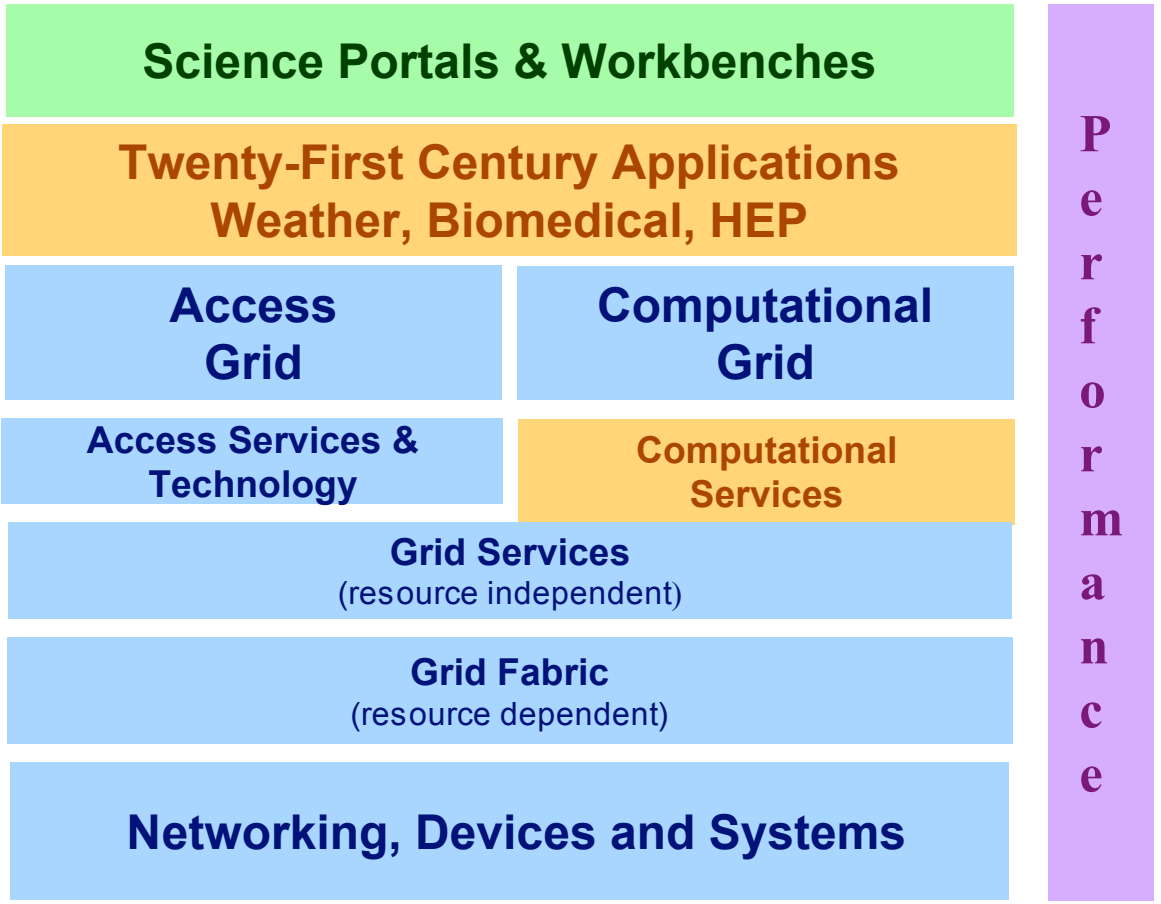
Globus Tools

- Application projects include
  - GriPhyN, PPDG, NEES, EU DataGrid, ESG, Fusion Collaboratory, etc., etc.
- Infrastructure deployments include
  - DISCOM, NASA IPG, NSF TeraGrid, DOE Science Grid, EU DataGrid, etc., etc.
  - UK Grid Center, U.S. GRIDS Center
- Technology projects include
  - Data Grids, Access Grid, Portals, CORBA, MPICH-G2, Condor-G, GrADS, etc., etc.

Grid Services ("Middleware")

- Standard services that
  - Provide uniform, high-level access to a wide range of resources (including networks)
  - Address inter-domain issues of security, policy, etc.
  - Permit application-level management and monitoring of end-to-end performance
- By integration of Distributed Technologies from Web Services, Dbase, Clustering, P2P, Collaboratories, High End Graphics to PKI based Security, you can provide these services
- Middleware-level and higher-level APIs and tools targeted at application programmers
  - Map between application and Grid

Grid Computing Layers



The Internet and burgeoning wired and wireless provide universal connectivity

On Going Projects

- **Globus - Sum of services architecture**
- **Legion - Object based & Parallel Env.**
- **Condor Group – Cluster Computing Env.**
- **Others**

  - CERN, 3 Tier Data Grid

  - Harness Geist

  - Dongarra

  - Grid Forum

  - Particle Physics Data Grid ..... & many more

Major Data Grid Projects

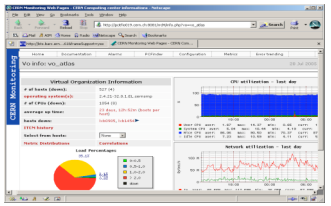
- **Earth System Grid (DOE Office of Science)**
  - **DG technologies, climate applications**
- **European Data Grid (EU)**
  - **DG technologies & deployment in EU**
- **GriPhyN (NSF ITR)**
  - **Investigation of "Virtual Data" concept**
- **Particle Physics Data Grid (DOE Science)**
  - **DG applications for HENP experiments**

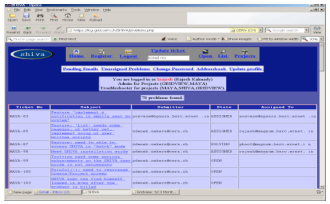
International Collaboration

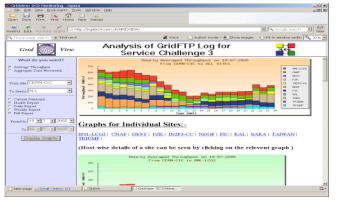
- India became a CERN observer state in 2002
- Large Hadron Collider (LHC) Grid Software Development, DAE-CERN Protocol agreement on computing for LHC data analysis, a DATA Grid called LCG

  ~10 people working in India for 5 years amounting to 7.5 MSWF
- BARC developed software is deployed at LCG, CERN

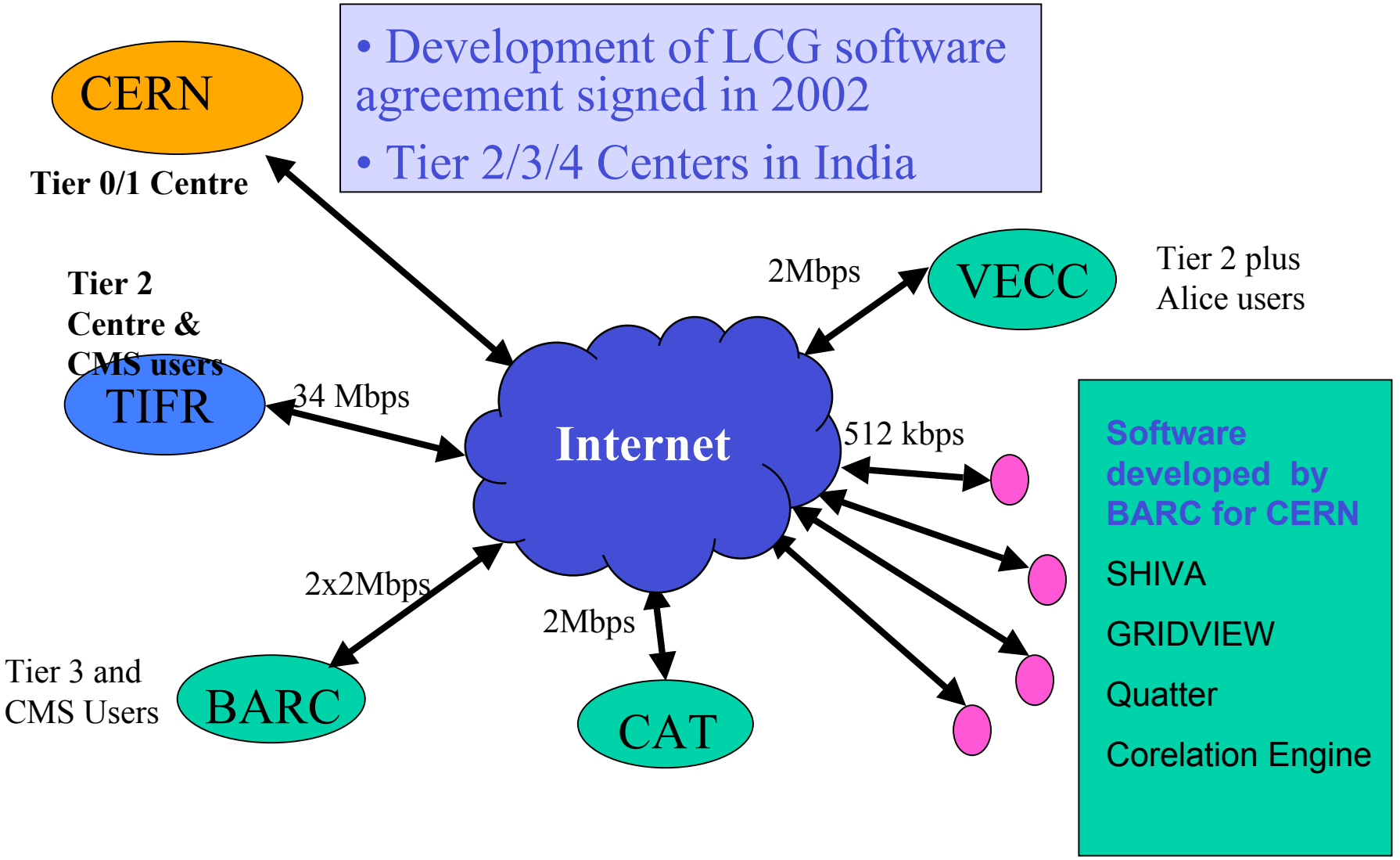
  - Co-relation Engine, Fabric management
  - Problem Tracking System (SHIVA)
  - Grid Operations (GRID VIEW)
  - Quattor a system administration toolkit





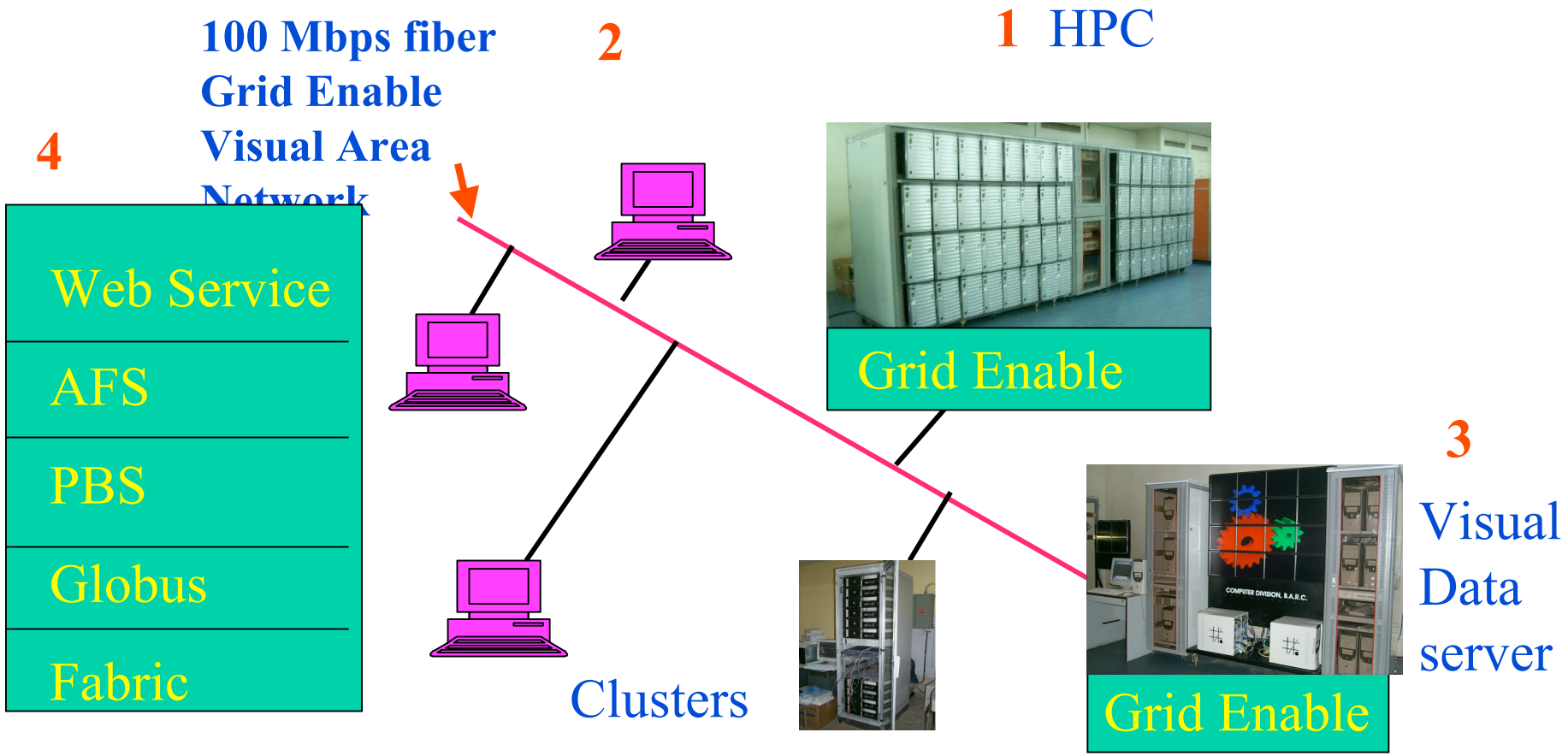


LHC: CERN-DAE COLLABORATION

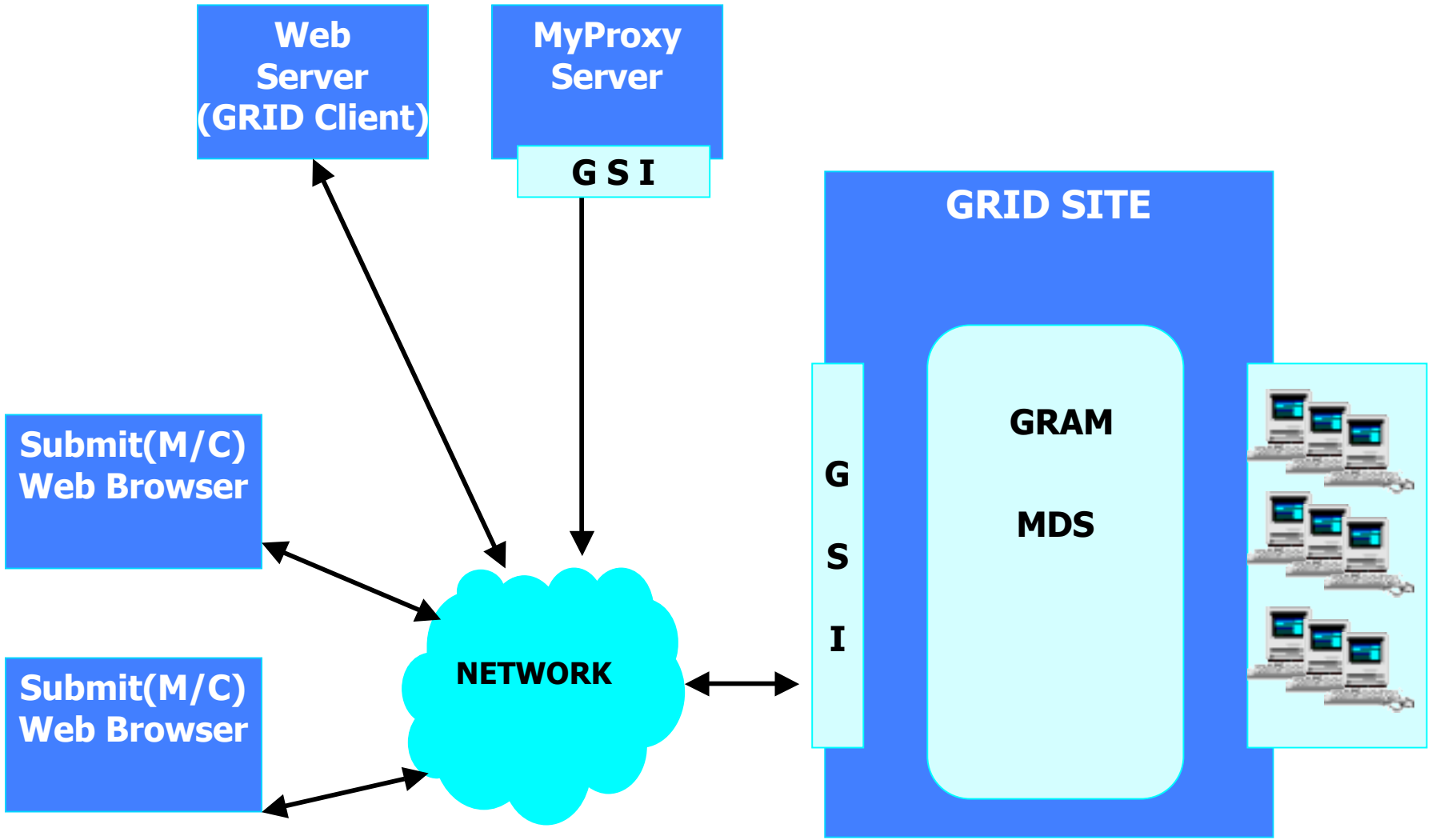


Computing Grid at BARC

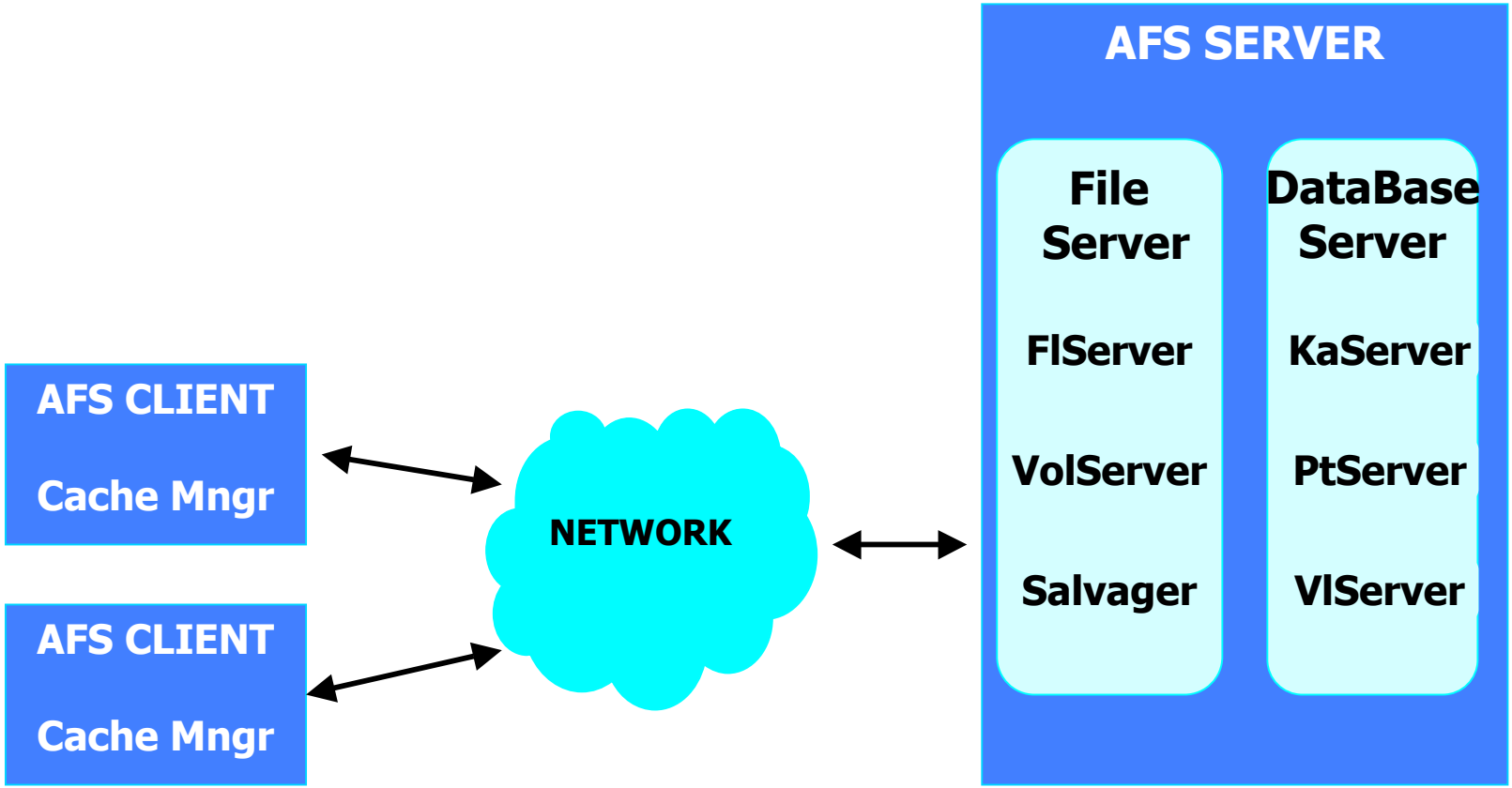
- Computing Grid system has been set up as a Test-Bed using existing Grid Technology Components



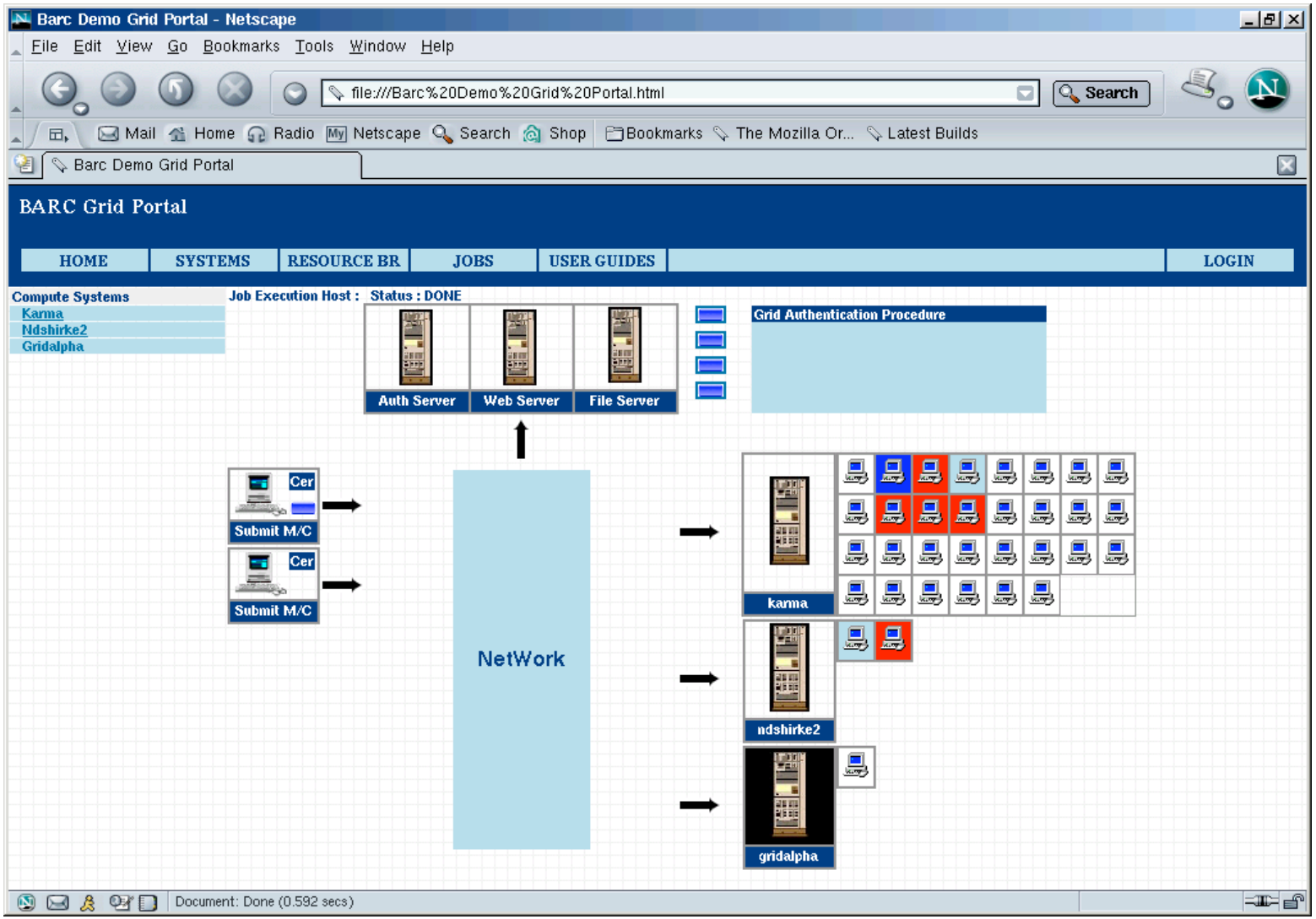
GRID PORTAL ARCHITECTURE



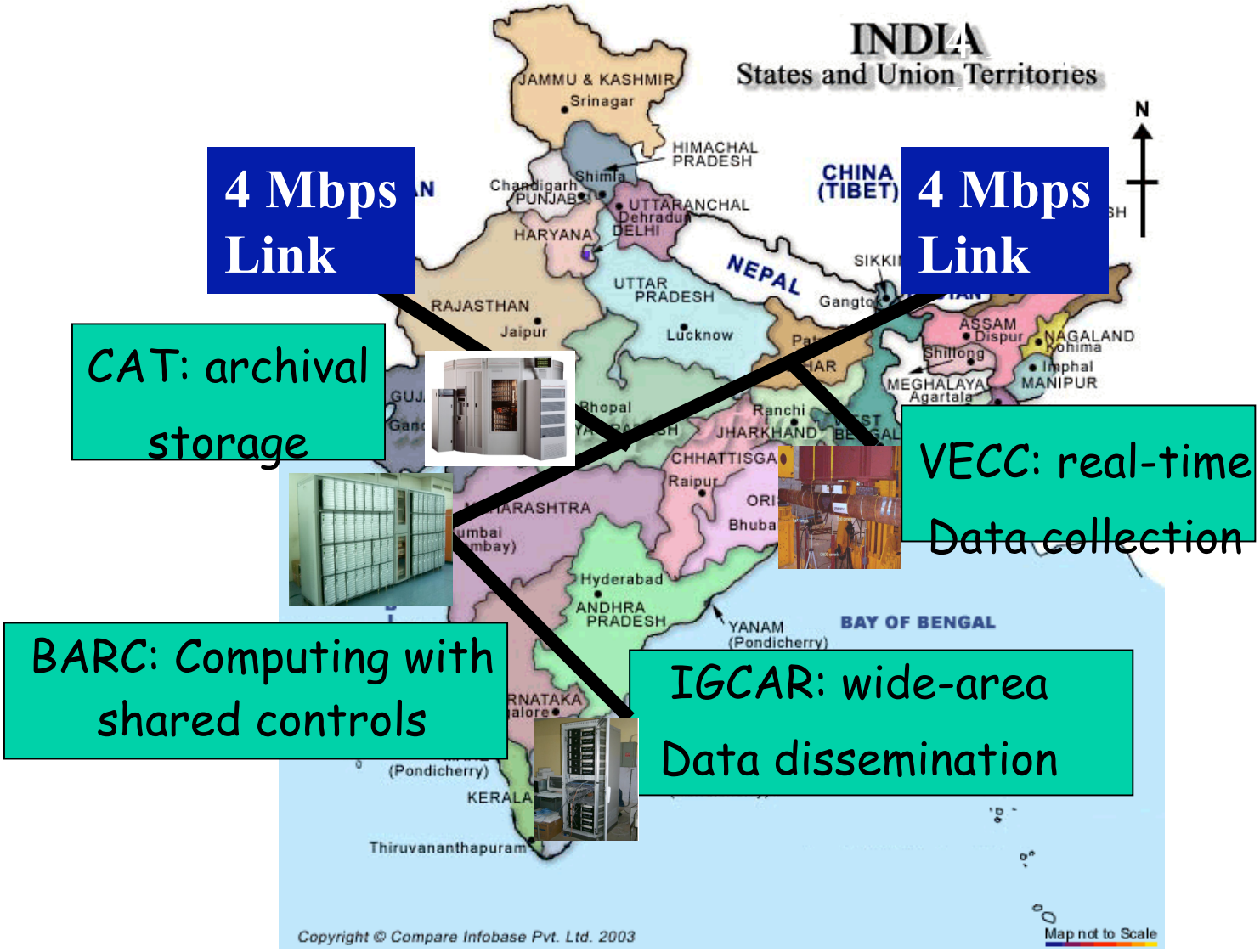
AFS ARCHITECTURE



Snap (Showing Job Location)



DAE Grid



4 Mbps Link

4 Mbps Link

CAT: archival storage

VECC: real-time Data collection

BARC: Computing with shared controls

IGCAR: wide-area Data dissemination

Resource sharing and coordinated problem solving in dynamic, multiple R&D units

Real Challenge

- Future HEP experiments require massive amounts of computing, including data collection and storage, data access, database access, computing cycles, etc.
- How to provide seamless and transparent access to Grid Services

  - without compromising on security & convenience
- Use of network bandwidth on demand basis
- How to do it inexpensive way

Issues Involved?

- Relatively New & Immature Technology
- Few standards & absence of good Tools
- Lack of expertise of users, developers & system support people
- Portability & Scalability to be resolved
- Added complexity due to WAN connectivity


Thank You

Computing Trends

- It is fully expected that the substantial and exponential increases in performance of IT will continue for the foreseeable future (at least next 50 years) in terms of
  - CPU Power (2X – every 18 months)
  - Memory Capacity (2X – every 18 months)
  - LAN/WAN speed (2X – every 9 months)
  - Disk Capacity (2X – every 12 months)
- It is expected that all computing resources will continue to become cheaper and faster, though not necessarily faster than the computing problems we are trying to solve.

Computations of future?

- **PC of 2004 is much more fast as supercomputer of 1990's**

  Still not adequate for many present here

- **PC now comes with over 80 Gigabyte of disk as much as entire 1990's computer centre (Storage Capacity doubles every 12 months)**

  By 2006 several physics projects like CERN, Biology, Astrophysics and sensor networks will produce multiple of Petabytes of data per year

- **To work with a colleague even across campus LAN on Petabyte scale data set, requires ultra fast network**

  Even though computing power, data storage & communication continues to improve, computing resources are failing to satisfy users demand

- **Unprecedented computing requirements for next generation scientific & engineering applications**

Limitations of Parallel Computing

- Programming so many nodes concurrently remains a major barrier for most applications
  - Source code should be known & parallisable
  - Scalable algorithm development is not easy
  - All resources are allotted for a single job
  - User has to worry about message passing, synchronization and scheduling of his job
  - 15% users only require these solutions, rest can manage with normal PCs
- Fortunately lot of free MPI codes and even parallel solvers are now available

GRID CONCEPT

User Access Point

Resource Broker

Grid Resources