Grid Initiatives in India

From

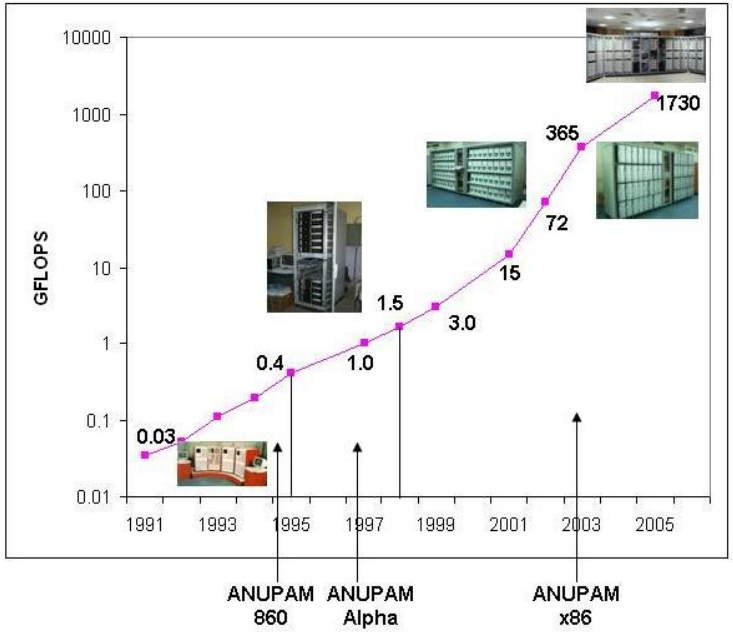
P.S.Dhekne

Presented by

Rajesh Kalmady, BARC, India

Parallel Computing at BARC

- BARC started development of Parallel & Cluster Computing to meet computing demands of in-house users with the aim to provide inexpensive high-end computing since 1990-91



- Have built so far 16 different models using varying CPU and networking technologies
- Today Clusters are the primary IT infrastructure

Cluster based Systems

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*

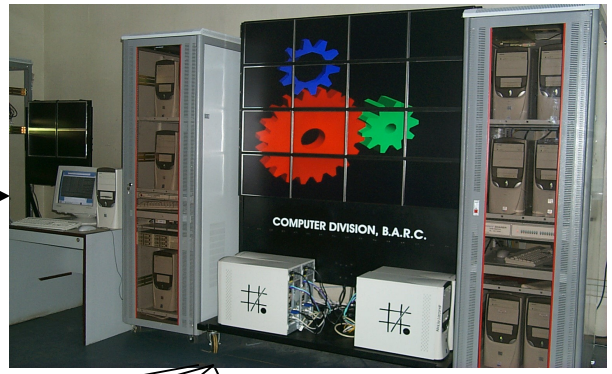
HPC : High Performance Computing HA : High Availability LB : Load Balancing

HPC Environment at BARC

Pre-processing

Solver

Post-processing

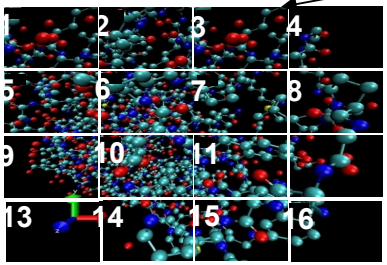


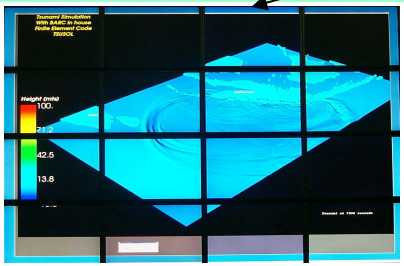
Front-end

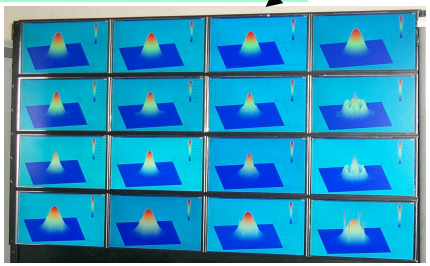
Supercomputing Cluster

1.7 TF HPL

Multiple Graphics HW







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

Software Development

- **Program Development Tools**
  - **Libraries, Debuggers, Profilers etc.**
- **System Software**
  - **Communication drivers, monitors, firmware**
  - **Job submission and queuing system**
  - **Cluster file system**
- **Management and Monitoring tools**
  - **Automatic installation**
  - **Cluster Monitoring System**
  - **Accounting System**
  - **SMART – Self Diagnosis and Correction system**

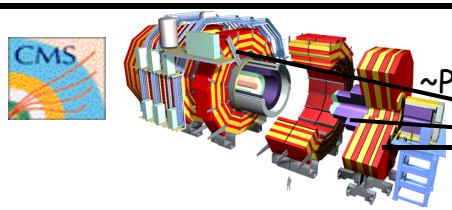
Difficulties in today's systems

- Major organizations have their own computer systems, thus idle when no load but not available to outsiders
- For operating computer centre 75 % cost come from environment upkeep, staffing, operation and maintenance; why every one should do this?
- As digital data is growing high-speed connectivity is essential; bandwidth & data sharing is an issue
- Supercomputers, Visual systems and Networks are not tightly coupled by software; difficult for users to use it

e-Science" and "e-Research

- Collaborative research that is made possible by sharing across the Internet of resources (data, instruments, computation, people's expertise...)
  - Crosses organisational boundaries
  - Often very compute intensive
  - Often very data intensive
  - Sometimes large-scale collaboration
- Owning Vs Sharing the resources
- Today you can't submit jobs on the Internet

Data Grid at CERN



~PBytes/sec

Online System

~100 MBytes/sec

1 TIPS is approximately 25,000 SpecInt95 equivalents

There is a "bunch crossing" every 25 nsecs.
There are 100 "triggers" per second
Each triggered event is ~1 MByte in size

Offline Processor Farm
~20 TIPS

~100 MBytes/sec

Tier 0

CERN Computer Centre

~622 Mbits/sec
or Air Freight
(deprecated)

Tier 1

France Regional Centre

Germany Regional Centre

Italy Regional Centre

FermiLab ~4 TIPS

~622 Mbits/sec

Computing power & data bandwidth shared by all collaborators

Tier 2

Caltech
~1 TIPS

Tier2 Centre
~1 TIPS

~622 Mbits/sec

Tier 3

Institute
~0.25TIPS

Physics data cache

~1 MBytes/sec

Tier 4

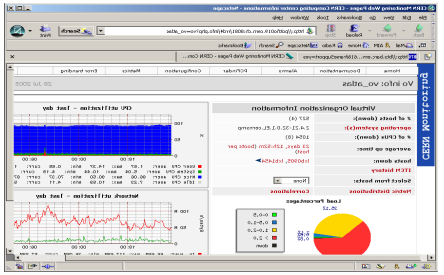
Physicist workstations

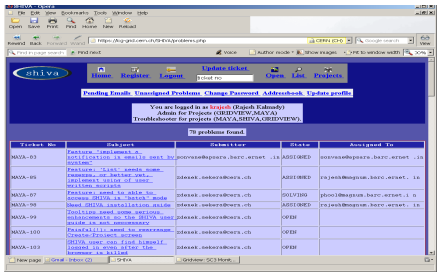
Physicists work on analysis "channels".
Each institute will have ~10 physicists working on one or more channels; data for these channels should be cached by the institute server

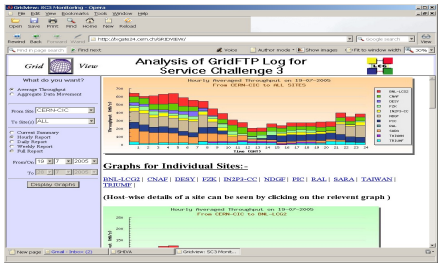
Image courtesy Harvey Newman, Caltech

DAE-CERN Collaboration

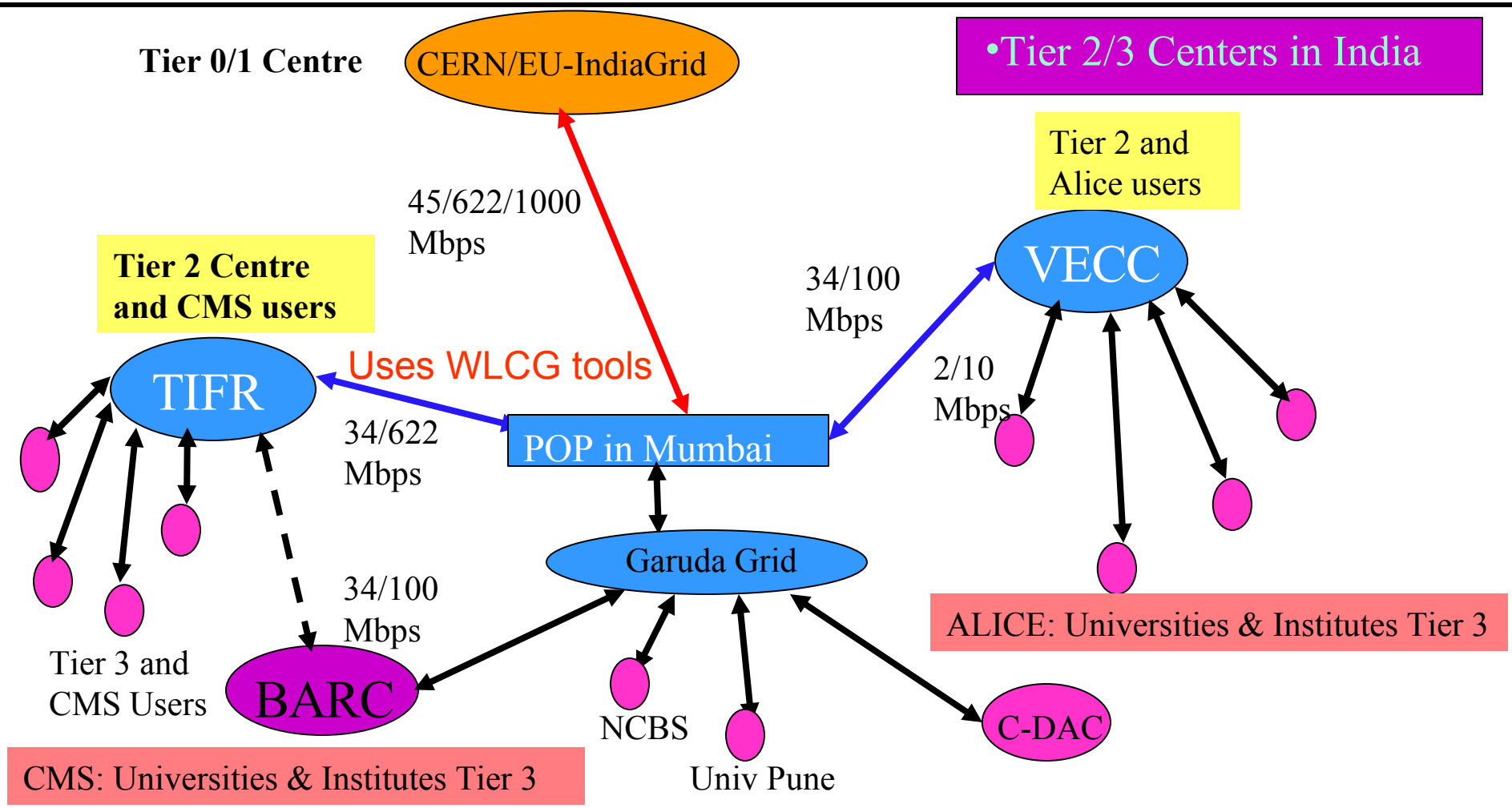
- DAE-CERN Protocol agreement on Grid computing for software development for WLCG.
- DAE developed software is deployed at WLCG, CERN
  - Co-relation Engine, Fabric management
  - Problem Tracking System (SHIVA)
  - Grid Monitoring (GRID VIEW)
  - Quattor toolkit enhancements
  - Data Base Management
  - Fortran Library conversion







Regional LCG Tier-2 in India



DAE/DST/ERNET: Geant link operational since August 2006

GridView: Visualization and Monitoring Tool for LCG - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites Media History Mail Print Skype

Address http://gridview.cern.ch/GRIDVIEW/

Google Go Bookmarks 161 blocked Check AutoLink AutoFill Send to Settings

Search Web Mail My Yahoo! Personals Entertainment Games

GridView: Visualization and Moni... Add Tab

Grid View

Monitoring and Visualization Tool for LCG

Data Transfer | Job Status | Service Availability

(Version: gridview-3.0.2, Installation Date: Jan 08, 2007)

<< ABOUT ABOUT >>

LCG

What do you want ?

- (•) VO-wise graph
- ( ) Site-wise graph

- (•) Average Throughput
- ( ) Aggregate Data X-fer

VO: All / Alice / Atlas

Source Site(s): CERNCI

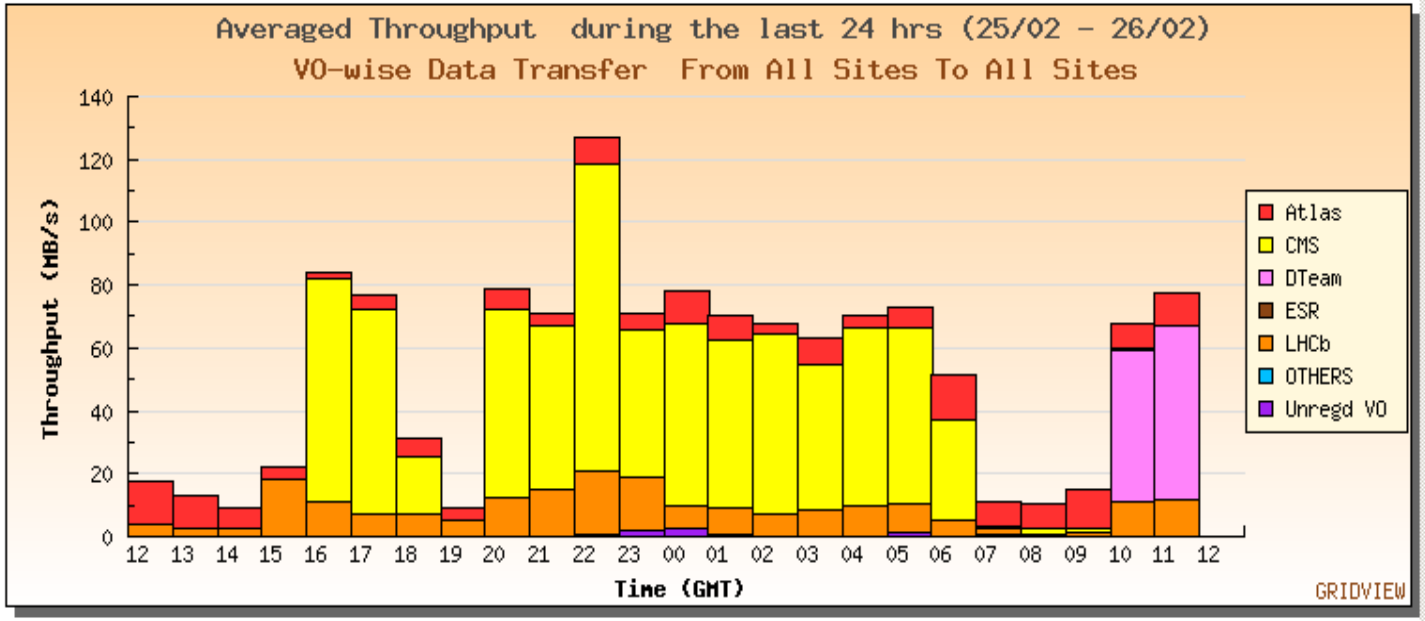
Dest Site(s): ALL

- (•) Current Summary
- ( ) Hourly Report
- ( ) Daily Report
- ( ) Weekly Report
- ( ) Monthly Report

From 26 2 2007

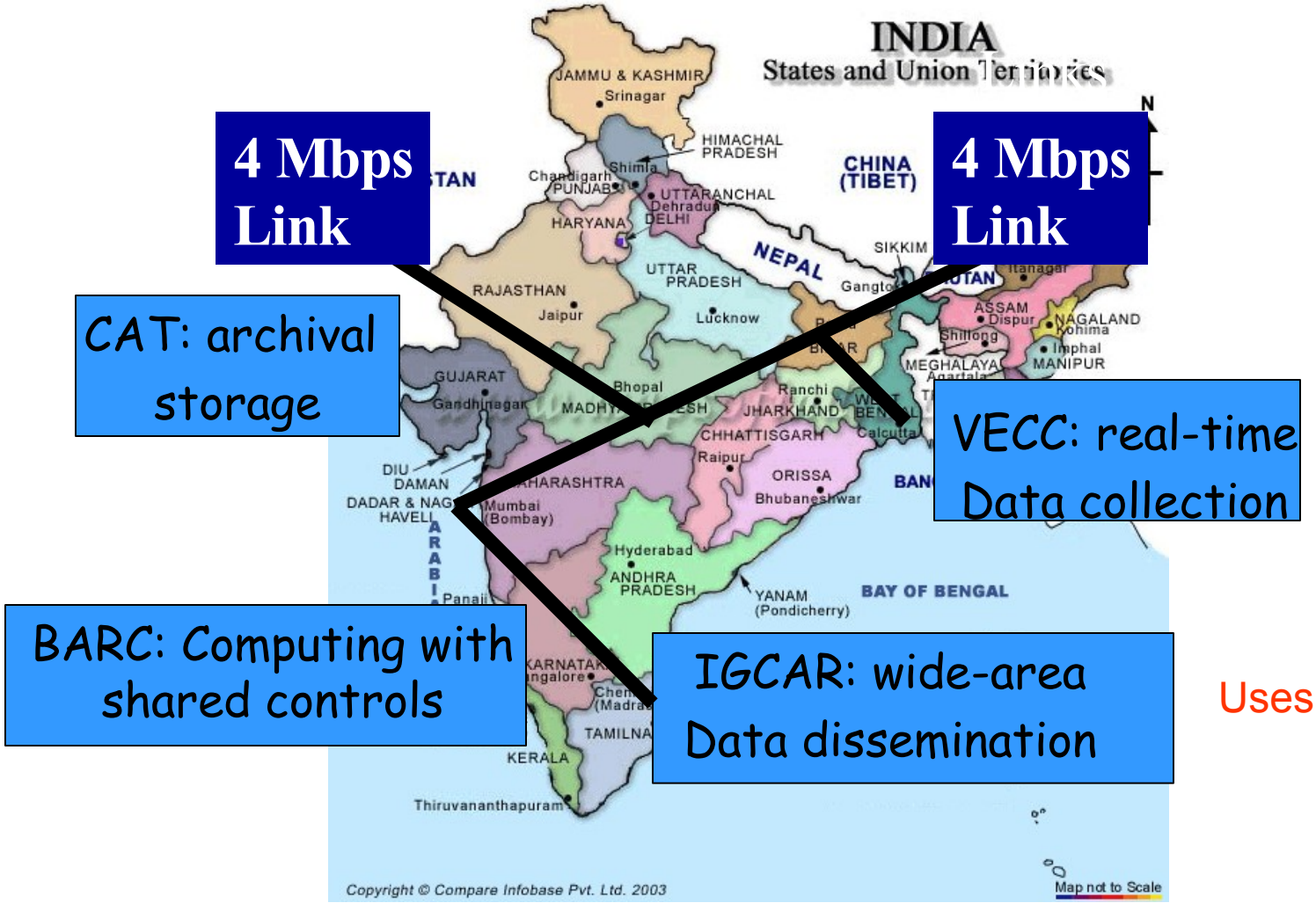
Current Status

(VO-wise Data Transfer From All Sites To All Sites)



(**OTHERS:** VOs giving throughput less than 1% of max, click here for names)

DAE Grid (Private)



Uses WLCG tools

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

Welcome to the BARC GRID Portal - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://gridportal.barc.daegrid.gov.in/

Scientific Linux Distros

Welcome to the BARC GRID Portal | Welcome to the GENIUS INFN GR...

BARC Grid Portal

Grid Services

- File Services
- Security Services
- Job Services
- Data Services
- Info Services
- Interactive Services
- Grid Settings
- Set VO
- Current VO Services
- Statistics
- Logout

based on Genius
INFN GridPortal

Welcome to BARC Grid Portal

New Computer Centre, BARC

Done gridportal.barc.daegrid.gov.in

1 2 3 4 Welcome to the BARC GRID 16:13

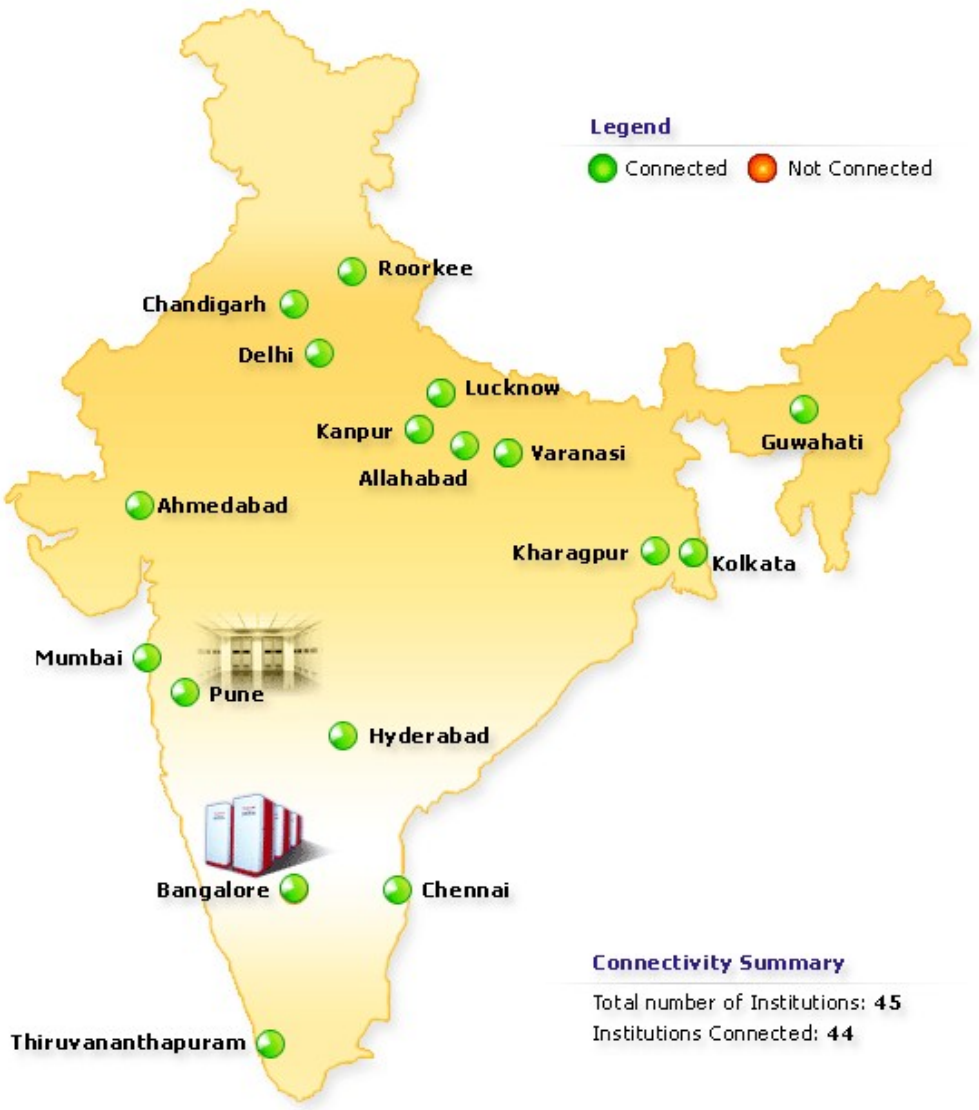
National Grid Initiative - GARUDA

- Department of Information Technology (DIT), Govt. of India, has funded C-DAC (Centre for Development of Advanced Computing) to deploy nationwide computational grid named GARUDA
- Currently in its Proof of Concept phase.
- It will connect 45 institutes in 17 cities in the country at 10/100 Mbps bandwidth.

Garuda - Deliverables

- Grid tools and services to provide an integrated infrastructure to applications and higher-level layers
- A Pan-Indian communication fabric to provide seamless and high-speed access to resources
- Aggregation of resources including compute clusters, storage and scientific instruments
- Creation of a consortium to collaborate on grid computing and contribute towards the aggregation of resources
- Grid enablement and deployment of select applications of national importance requiring aggregation of distributed resources

Garuda – Network Connectivity



The EU-IndiaGrid Project Joining European and Indian grids for e-science

- To support the interconnection and interoperability of the prominent European Grid infrastructure (EGEE) with the Indian Grid infrastructure for the benefit of eScience applications
- Two year project started from Oct 2006 with BUDGET of 1208 k-EUR total fund out of which 1015.9 k-EUR from European Commission **(5 Europe & 9 Indian partners)**
- Person months
  - **353.3 PM total**
  - **226.4 PM funded from European Commission**
- First kickoff meeting in ICTP Italy during 18-20 Oct, 2006
- Workshop WLCG & EU-IndiaGrid at TIFR during 1- 4 Dec, 2006
- Belief Conference in New Delhi from 13-15 Dec 2006

PARTNERS

EUROPE

- **INFN (project coordinator),**
- **Metaware SpA,**
- **Italian Academic and Research Network (GARR)**
- **Cambridge University**

INTERNATIONAL

- **Abdu Salam International Centre for Theoretical Physics (ICTP)**

INDIA

- **Indian Education and Research Network (ERNET),**
- **University of Pune,**
- **SAHA Institute of Nuclear Physics, Kolkata (SINP) & VECC,**
- **Centre for Development of Advanced Computing (C-DAC),**
- **Bhabha Atomic Research Centre, Mumbai (BARC)**
- **TATA Institute for Fundamental Research (Mumbay) (TIFR)**
  National Centre for Biological Sciences, Bangalore (NCBS)

EU-IndiaGrid Status

GEANT-ERNET Milan-Mumbai 45 Mb/s link opened since August 2006

– WLCG Tier-II CMS & ALICE centres and 10 Universities are interconnected

– 50 Research laboratories and educational institutes situated in 17 major Indian cities are interconnected within the GARUDA National Grid Initiative through ERNET from Jan 2007

Key issues,

– Certification Authority

– Creation of a pilot test bed

– Interoperability between gLite & GT

Cooperation with Academia Sinica (Regional Operation Centre for Asia) allowed:

– To established Registration Authorities at each site in order to ensure immediate grid access worldwide to Indian users

– Set the necessary steps for an Internationally recognized Indian Certification Authority (Responsibility taken by C-DAC)

Other Grids in India

- ERNET's partnership with GÉANT2 is supported by the EUIndiaGrid initiative, a project that aims to interconnect European Grid infrastructures with related projects in India.
- BARC MOU with INFN, Italy to establish Grid research Hub in India and Italy
- 11th five year plan proposals for E_infrastructure and Grid for S&T applications submitted to GOI with possibility for Weather, Bio and e-Governance

Summary

- Grids presents a single & unified resources for solving large-scale compute & data-intensive appln.
  - Owning Vs Sharing
- Providing affordable (free) access to national as well as international resources to the Universities, Schools & Colleges may bring innovative leadership to India
- India can make great contribution in developing global HPC applications (modeling & Simulation software), model verification & validations, tools development, evaluations & testing etc

THANK YOU