Garuda: The National Grid Computing Initiative - the platform for collaboration and innovation

Subrata Chattopadhyay
CDAC Knowledge Park, Bangalore
subratac@cdac.in
Outline

• Introduction on Garuda
• NKN – highlights
• Tools and Services - GarudaWare
• How to Join and Access
• Major Applications
• Collaborations
• Conclusion
Global Access to Resources Using Distributed Architecture
The Foundation Phase

• Migration to SOA.
  – GT4.x based Grid Middleware stack

• Open Source Grid meta scheduler
  – Gridway

• Open Source Data Grid solution
  – SRM

• Viability of commercial applications / services
  – CAE (Zeus Numerix) and
  – Molecular Docking (C-DAC & Jubilant BioSys) on SOA GARUDA

• IGTF accredited Certification Authority
  – IGCA was setup - First CA in India to address security issues of grid

• Demonstrate Performance benefits for selected pilot applications
  – Improved processing time from 6.1hrs to 54 mins for processing one set of Radar data (9GB) for the DMSAR application using the new resources of GARUDA (272 cores)
  
  – Winglet applications for processing 6000 winglets taking nearly 30 days to complete on ordinary machines were able to complete in 3 hours by harnessing the large computational power offered by GARUDA.
Operational Phase

Currently in the Operational Phase

Objective -

To provide stable, robust, and efficient environment with guaranteed quality of service for various applications, which is of national importance.

- 2010 - 2011 - 2012
  - Architecture & Framework
  - Middleware & Tools
  - Migration to NKN
  - Grid Enablement Support
  - Stabilize & Enhance - NKN, middleware, tools
  - Participation - Intl Commn, Stds Forum
  - Resource Aggregation & Augmentation
  - Deploy Operational Applications
  - Next Gen Services
  - Knowledge Dissemination
  - Resource Aggregation & Augmentation

- 2010 - 2011 - 2012
  - Architecture & Framework
  - Middleware & Tools
  - Migration to NKN
  - Grid Enablement Support
  - Stabilize & Enhance - NKN, middleware, tools
  - Participation - Intl Commn, Stds Forum
  - Resource Aggregation & Augmentation
  - Deploy Operational Applications
  - Next Gen Services
  - Knowledge Dissemination

- 2010
  - Architecture & Framework
  - Middleware & Tools
  - Migration to NKN
  - Grid Enablement Support
  - Stabilize & Enhance - NKN, middleware, tools
  - Participation - Intl Commn, Stds Forum
  - Resource Aggregation & Augmentation
  - Deploy Operational Applications
  - Next Gen Services
  - Knowledge Dissemination

- 2011
  - Architecture & Framework
  - Middleware & Tools
  - Migration to NKN
  - Grid Enablement Support
  - Stabilize & Enhance - NKN, middleware, tools
  - Participation - Intl Commn, Stds Forum
  - Resource Aggregation & Augmentation
  - Deploy Operational Applications
  - Next Gen Services
  - Knowledge Dissemination

- 2012
  - Architecture & Framework
  - Middleware & Tools
  - Migration to NKN
  - Grid Enablement Support
  - Stabilize & Enhance - NKN, middleware, tools
  - Participation - Intl Commn, Stds Forum
  - Resource Aggregation & Augmentation
  - Deploy Operational Applications
  - Next Gen Services
  - Knowledge Dissemination
Garuda on MPLS based NKN

Legend
H Head Node
G Gateway
High Capacity, Highly Scalable Backbone
Provide Quality of Service (QoS) and Security
Wide Geographical Coverage
Common Standard Platform
Bandwidth from Many NLD’s
Highly Reliable & Available by Design
Test beds (for various implementation)
Dedicated and Owned.
Garuda High Level System Components

- GARUDA – enabled Applications
- WSRF+GT4 + other Services + Cloud S/W (Nimbus/VMware)
- Virtualization support
- Grid Security and High-Performance Grid Networking
- NKN

Computing Resources and Virtual Organizations

- CDAC Resource centers
- Non-Research Organizations
- Research Organizations
- Educational institutions Computing Centers

Resources
Security
Middleware
Resource Management
User Environments
Programming Environments
Data Grid
Grid Applications
GARUDA Middleware components

**Management, Monitoring & Accounting**
- Paryaveekshanam
- GARUDA Accounting
- GARUDA Information Service
- Web MDS

**GARUDA Resources**
- Compute, Data, Storage,
- Scientific Instruments,
- Software,

**Access Methods**
- Access Portal
- Problem Solving Environments
- Cmd line interface
- Visualization gateways
- workflows

**Security Framework**
- IGCA Certificates
- VOMS
- MyProxy

**Resource Mgmt & Scheduling**
- Resource Reservation
- GridWay Meta-scheduler
- Torque, Load Leveler
- Globus 4.x (WS Components)

**Legend**
- CDAC Development
- Customization
- Open Source
Garuda Middleware tools

- Advanced Reservation system
- Virtual Organisation Management Service
- Integrated Login service
- Compilation service
- Accounting service
- Quality of Service (ongoing)
- Resource Administration Tool
- Garuda Information Registry
Other Tools on Garuda

- Garuda Access portal
  - Job Resubmission through Mobile
  - Integrated with visualisation Gate way

- Automatic Grid Service Generator (AGSG)

- GSRM a Data grid solution for GARUDA
  - Integrated with VOMS

- Resource monitoring through Paryavekshmanam
  - SMS (text) alerts through Mobile for Grid Administrators

- Hub Zero based Garuda Visualization gateway
  - Visualization tools for different application domains
- **Indian Grid Certification Authority** located at C-DAC, Knowledge Park, Bangalore, India.
- IGCA is the **accredited** member of APGridPMA.
- Issues **X.509 Certificates** to support the secure environment in Grid. (for GARUDA, institutes that do research in grid from India and foreign institutes that collaborates with GARUDA).
- [http://ca.garudaindia.in](http://ca.garudaindia.in)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Valid User Certificates</strong></td>
<td>285</td>
</tr>
<tr>
<td><strong>Number of Valid Host Certificates</strong></td>
<td>54</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>339</td>
</tr>
</tbody>
</table>
GARUDA Short Live Certificate

GARUDA SLCS provides grid users an instant access to GARUDA grid for a trial period of 30 days.

**Highlights:**
- Hassle free registration
- Get an access in less than 5mins.
- Service over the internet.

**Features:**
- GARUDA Job submission portal
- GARUDA Compiler Service

Website: http://labs.garudaindia.in
GARUDA Resources

- CDAC Resource:
  - 4TF HPC clusters each at Bangalore, Chennai & Hyderabad
  - PARAM Yuva at Pune and PARAM Padma at Bangalore
- Fourteen of the partner institutions are also contributing resources including satellite terminals.
- Total computing power is more than 5500 CPUs equivalent to 65TF
- Storage space 220 TB
<table>
<thead>
<tr>
<th>Institution</th>
<th>Location</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space Application Centre</td>
<td>Ahmedabad</td>
<td>VSAT Terminal - 2 Nos.</td>
</tr>
<tr>
<td>Indian Institute of Science</td>
<td>Bangalore</td>
<td>64 cpu; POWER5; Linux</td>
</tr>
<tr>
<td>Raman Research Institute</td>
<td>Bangalore</td>
<td>32 cpu; Opteron; Linux</td>
</tr>
<tr>
<td>Institute of Mathematical Sciences</td>
<td>Chennai</td>
<td>24 cpu; Opteron cluster (Cray XD1)</td>
</tr>
<tr>
<td>Madras Institute of Technology</td>
<td>Chennai</td>
<td>16 cpu; P4; Linux</td>
</tr>
<tr>
<td>Indian Institute of Technology</td>
<td>Delhi</td>
<td>32 cpu; Opteron; Linux</td>
</tr>
<tr>
<td>Jawaharlal Nehru University</td>
<td>Delhi</td>
<td>32+16+16 cpu; Opteron, Opteron, Itanium; Linux</td>
</tr>
<tr>
<td>Institute of Genomics and Integrative Biology</td>
<td>Delhi</td>
<td>48 cpu; Xeon; Linux</td>
</tr>
<tr>
<td>Indian Institute of Technology</td>
<td>Guwahati</td>
<td>128 cpu; Opteron; Linux</td>
</tr>
<tr>
<td>University of Hyderabad</td>
<td>Hyderabad</td>
<td>32 way SMP; POWER4, AIX</td>
</tr>
<tr>
<td>Indian Institute of Technology</td>
<td>Kharagapur</td>
<td>16+16 cpu; Power PC2, Xeon; AIX, Linux</td>
</tr>
<tr>
<td>Physical Research Laboratory</td>
<td>Ahmedabad</td>
<td>320 cpus; 64bit AMD</td>
</tr>
<tr>
<td><strong>CDAC</strong></td>
<td><strong>Bangalore</strong></td>
<td>64 cpu Power 5; 320 cpu Xeon Linux</td>
</tr>
<tr>
<td><strong>CDAC</strong></td>
<td><strong>Hyderabad</strong></td>
<td>320 cpu Xeon Linux</td>
</tr>
<tr>
<td><strong>CDAC</strong></td>
<td><strong>chennai</strong></td>
<td>320 cpu Xeon Linux</td>
</tr>
<tr>
<td><strong>CDAC</strong></td>
<td><strong>Pune</strong></td>
<td>32 cpu Xeon Linux: 4068 CPU Linux</td>
</tr>
</tbody>
</table>
GARUDA Operations & Management

- Looks after deployment of middleware and network
- Operates from CDAC KP Bangalore
- Operation Centre with High resolution, scalable display wall
- Conduct Regular Monday meetings among administrators to maintain Garuda health

gridsupport@garudaindia.in rt@cdacb.ernet.in
GARUDA Partners

• Motivation
  – To Collaborate on Research and Engineering of Technologies, Architectures, Standards and Applications
  – To Contribute to the aggregation of GARUDA resources

• Participation
  – 36 research & academic institutions in the 17 cities
  – 8 centres of C-DAC
  – Total of 45 institutions
  – Additional 10 labs with LOI
<table>
<thead>
<tr>
<th>Group Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bioinformatics</strong></td>
<td>application of statistics and computer science to the molecular biology</td>
</tr>
<tr>
<td><strong>ClimateModelling</strong></td>
<td>Deals with the dynamics of the climate system.</td>
</tr>
<tr>
<td><strong>OSDD</strong></td>
<td>Community dedicated to develop drugs for tropical infectious diseases like malaria, tuberculosis</td>
</tr>
<tr>
<td><strong>GeoPhysis</strong></td>
<td>Study related to physics of the Earth and its environment in space</td>
</tr>
<tr>
<td><strong>CAE</strong></td>
<td>usage of computer software to solve engineering problems</td>
</tr>
<tr>
<td><strong>IndianHeritage</strong></td>
<td>Focused on technology products for preserving &amp; processing Heritage texts</td>
</tr>
<tr>
<td><strong>HealthInformatics</strong></td>
<td>Focused on utilizing compute power for health informatics</td>
</tr>
<tr>
<td><strong>MaterialScience</strong></td>
<td>interdisciplinary field applying the properties of matter to science and engineering</td>
</tr>
<tr>
<td><strong>Euindia</strong></td>
<td>The vision of a worldwide Grid for Research by both Europe and India</td>
</tr>
<tr>
<td><strong>ToolsDeveloper</strong></td>
<td>Forum to communicate and collaborate on developing Garuda Tools</td>
</tr>
<tr>
<td><strong>GarudaAdmin</strong></td>
<td>Meant for administrators from resource providers &amp; Garuda Operation team members</td>
</tr>
</tbody>
</table>
How to Join Garuda

**Model 1**
- By signing “Memorandum of understanding (MOU) ”
  - by paying annual fee or
  - contributing Resources to GARUDA grid.

**Model 2**
- By signing “Letter of Exchange (LOE) ”.
  - The validity is for 6 months and this is free of cost.
  - Intended for new users.
  - On expiry partnership can be continued through MOU.
  - The MOU and LOE to be signed by **Head of the Institute** like / Director/Vice chancellor/or equivalent.
How to Access Garuda

- CLI - Command Line Interface for advance users
- Standard based Job Submission Portal
- PSE – Problem Solving Environment for individual application domain
- Workflows
- Hand held devices like Mobile
- Cloud interface through which IaaS an SaaS kind of requests will be addressed
Applications on GARUDA
Bioinformatics: Open Source Drug Discovery

Project Team: OSSD community

- **26** OSSD users given access to Garuda through **OSDD VO**
- HPC clusters to run drug discovery problems
- Users connected through both **NKN** and **Internet**
- **Galaxy Workflow** for genomics proteomics applications
- Distributed job execution through **Gridway**

Grid Enabled Bioinformatics tools useful in drug discovery pipeline
- Genetic Algorithm (GA) to predict tertiary structure from primary structure of protein sequence
- Grid-enabled PSP uses distributed approach for jobs submission across Garuda
- Protein sequence is divided into multiple overlapping regions and structure of each part is predicted using GA on individual Grid resources
- Results are combined on Grid head node and final structure is predicted on any of the Grid resource.
Bioinformatics: Taxogrid

Project Team: CDAC Pune

Resources
- gg-hyd.cdac.org.in
- gg-che.local
- gg-blr.tfg
- xn00.npsf.cdac.ernet.in
- yuva.pnq-cdac-hq-100.garuda

Software Versions
- Globus – 4.0.7
- GridWay – 5.4
- DRMAA - 1

Dataset
- Database: UniProtKB
- Mycobacterium tuberculosis H37Rv
- No. of sequences: 3988
- Size of input dataset: 1.65 MB
- Size of output file: 10.4
- Method: Maximum Likelihood

TaxoGrid Benchmark (Time Scale)

Time for speedup at value ‘1’ : 1144 minutes
Time for speedup at value ‘3.31’ : 346 minutes
CAE: Aeroacoustics Optimization

Project Team: Zeus Numerix

Aim:
Optimize the noise generated by a 3-D wing with flaps in landing configuration by variation of flap location and orientation.

- Uses **Kepler** workflow Framework integrated with native **Globus** job submission routines
- Optimization Module uses **OPT4J** framework
- Optimization module includes **AFFG** (Adaptive Fuzzy Fitness Granule) routine which can reduce the number of fitness function evaluations up to 50%.
- Completion successful 40 simultaneous simulations (parallel + serial)
CAE: Wing Optimization

Project Team: Zeus Numerix

Aim: Optimise wing by variation of winglet design parameters using Genetic algorithm optimizer

Likely Impact: The Design Methodology will be used in Regional Transport Aircraft being designed by India

Highlights: For the first time in INDIA, a successful demonstration of automated Genetic Algorithm based High-fidelity and Multi-disciplinary shape optimization has been carried out
Medical and Health Care

Project Team: CDAC Pune

- Built the Adverse Event Reporting System (caAERS) and caTissue locally and studied its design and architecture.
- Built local testbed for porting and testing web applications on Grid.
- Studied caBIG® Data sharing and Security Frameworks.
- Networking with the potential users of caBIG Tools through workshop conducted at Pune during December 2010.
- Deploying & testing caBIG tools, caAERS and caTissue on local GRID testbed.
- Re-factoring the tools and building the Grid services for caTissue and caAERS web application.
- Deployment of caBIG tools over the GARUDA.
Collaborative Class Room

Website: http://ccr.garudaindia.in

Project Team: CDAC Hyderabad

Supported Features:

- Interface to Access grid
- GSRM based data storage for maintaining course repositories
- Indexing of course material based on key words
Collaborative Projects

– Grid-enabled aerosol modeling system for climate change studies (IIT, Bombay)
  • To design a Grid enabled modeling system for climate change studies using regional aerosol chemistry models with coupling to regional atmospheric models

– Middleware for Online Remote Visualization of Weather Applications (IISc, Bangalore)
  • to develop middleware frameworks for large-scale parameter explorations of critical weather applications on grids

– End Point Security for Grid Garuda (C-DAC Hyderabad)
  • To provide end system security solutions to grid partners to secure their networks
Interoperability with International Grids

• Integrating technological components of Garuda and EGI
  – Glite and Globus
  – Customizing Gridway meta-scheduler
  – To run real life application across both Infrastructures

• Collaboration between CaBig and Garuda
  – Interoperation of technological service among these grids
  – Cancer Research application portability
  – Contribution to standards for using distributed computing in Health care
Community Building & Dissemination

- Workshop on Garuda at IFIP networking 2010 – IIT, Chennai
- Garuda Boot Camp – 8-9 July 2010 at IMSc Chennai
- GARUDA Boot Camp - 23-24 Sept 2010 at NCRA Pune
- Material Science Virtual Community meet – 25 Sept 2010 at Pune
- GARUDA Boot Camp - 23-24 Nov 2010 at UoH, Hyderabad
- GARUDA Boot Camp - 28-29 Dec 2010 at MCC, Calicut
- Emerging technologies for collaborative Health care 13-14 May 2011 at SGPGI Lucknow
International Visibility

ISGTW INTERNATIONAL SCIENCE GRID THIS WEEK

Current Issue: 30 May 2007

Feature - EU-IndiaGrid: Building a Partnership Across Hemispheres

Over the last 30 years, India has posted an average annual growth rate of more than 7%. India’s success in the IT field has been especially dazzling. It is already a major exporter of software services and software workers. Despite these advances, the nation still faces pressing problems and currently lacks the computing infrastructure required to take full advantage.

Image of the Week - India’s National Grid Computing Initiative

GARUDA—which takes its name from a large, bird-like creature in Hindu and Buddhist mythology—is a collaboration of researchers and institutions aiming to establish a national grid in India. Currently in its Proof of Concept phase, GARUDA will link centers for computing, mass storage and scientific instruments, to enable data and computing intensive science in India for the 21st century.

The GARUDA high-speed network will connect 45 institutions in 17 cities at 10-100 megabytes per second bandwidth. To access an interactive version of this map visit the Garuda collaboration’s web site.

The GARUDA project is coordinated by the Center for Development of Advanced Computing, CDAC. It is a partner of EUnetGrid project. Interoperability between GARUDA and EUnetGrid infrastructure is a main goal of EUnetGrid.
Challenges!

• Heterogeneous Resources
• Policy across different administrative domains
• Diverse requirements and priorities of Application
• Funding and sustainability
• Killer Apps - to build vibrant community
• Interoperation among international grids
Plans for future

- Expand Garuda offerings - Users and Applications
- GVG-Hub for scientific visualization tool
- Scilab
- Scientific application as service on Grid-Cloud interoperable resources
- Upgrade GARUDA resources
- Scientific Equipments as Resource
- Announce Challenge competition for Student Participation
Thank you!