Co-ordination & Harmonisation of Advanced e-Infrastructures for Research and Education Data Sharing

Cloud and HPC examples for Research & Education
Dr. Ognjen Prnjat, GRNET

ISGTC 2013, Taipei, March 2013

Research Infrastructures – Grant Agreement n. 306819
Outline

- State of the art in eInfrastructures in Europe
- Worldwide Grid collaborations
- CHAIN-REDS: Example of HPC collaboration in SEE
- CHAIN-REDS: Example R&E cloud platform
Outline

- State of the art in eInfrastructures Europe
- Worldwide Grid collaborations
- Towards CHAIN-REDS: Example of HPC collaboration in SEE
- Towards CHAIN-REDS: Example R&E cloud platform
The Research **Network infrastructure** provides fast interconnection and advanced services among Research and Education institutes of different countries’

The Research **Distributed Computing Infrastructure** provides a distributed environment for sharing computing power, storage, instruments and databases through the appropriate software (middleware) in order to solve complex application problems.

This integrated networking & DCI environment is called **electronic infrastructure (eInfrastructure)** allowing new methods of global collaborative research - often referred to as **electronic science (eScience)**.

The creation of the eInfrastructure is a key objective of the **European Research Area**.
Network: GEANT

GÉANT: the pan-European research and education network
Transforming the way users collaborate

Backbone topology as at March 2012. GÉANT is operated by DANTE on behalf of Europe's NRENs.

GÉANT is co-funded by the European Commission within its 7th RTD Framework Programme.
Grid: European Grid Infrastructure

30/05/2012 Project Presentation – May 2012
## Grid: European Grid Infrastructure

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Value (Yearly increase)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Installed Capacity</strong></td>
<td></td>
</tr>
<tr>
<td>Logical CPUs</td>
<td>EGI-InSPIRE and EGI Council members</td>
</tr>
<tr>
<td></td>
<td>270,800 (+30.7%)</td>
</tr>
<tr>
<td></td>
<td>Including integrated RPs</td>
</tr>
<tr>
<td></td>
<td>399,300</td>
</tr>
<tr>
<td>HEP-SPEC 06</td>
<td>EGI-InSPIRE and EGI Council members and integrated RPs</td>
</tr>
<tr>
<td></td>
<td>2.96 Million (+49.5%)</td>
</tr>
<tr>
<td>Storage</td>
<td>Disk (PB)</td>
</tr>
<tr>
<td></td>
<td>139 PB (+31.4%)</td>
</tr>
<tr>
<td></td>
<td>Tape (PB)</td>
</tr>
<tr>
<td></td>
<td>134.3 PB (+50%)</td>
</tr>
<tr>
<td><strong>Resource Centres</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EGI-InSPIRE and EGI Council members</td>
</tr>
<tr>
<td></td>
<td>326</td>
</tr>
<tr>
<td></td>
<td>Including integrated RPs</td>
</tr>
<tr>
<td></td>
<td>352</td>
</tr>
<tr>
<td></td>
<td>Supporting MPI</td>
</tr>
<tr>
<td></td>
<td>90</td>
</tr>
<tr>
<td><strong>Countries</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EGI-InSPIRE and Council members</td>
</tr>
<tr>
<td></td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Including integrated RPs</td>
</tr>
<tr>
<td></td>
<td>54</td>
</tr>
<tr>
<td><strong>Performance</strong></td>
<td>Monthly Availability/Reliability</td>
</tr>
<tr>
<td></td>
<td>94.50%/95.42%</td>
</tr>
<tr>
<td><strong>Utilization</strong></td>
<td>HEP-SPEC 06 Hours</td>
</tr>
<tr>
<td></td>
<td>10.5 Billion (+52.91%)</td>
</tr>
<tr>
<td></td>
<td>Jobs</td>
</tr>
<tr>
<td></td>
<td>492.5 Million Jobs/year</td>
</tr>
<tr>
<td></td>
<td>1.35 Million Jobs/day (+46.42%)</td>
</tr>
</tbody>
</table>
HPC: PRACE

- 24 countries
- 5 Tier0 systems currently deployed
- ~9 Pflops total power (Tier-0)
- 15+ Tier 1 systems
- PRACE PP + phases of implementation projects, total EC investment ~ 70MEuro
- Access by European peer review
Pan-European Cloud initiatives

- No core initiative available yet
- StartusLab - GRNET leads operations, IaaS
- EGI cloud activity - providing virtualised Grid sites on cloud; will start moving towards cloud-based models
- Future CHAIN possibilities: GRNET and other clouds to be accessible directly or via CHAIN scientific gateway
Outline

- State of the art in eInfrastructures in Europe
- **Worldwide Grid collaborations**
- CHAIN-REDS: Example of HPC collaboration in SEE
- CHAIN-REDS: Example R&E cloud platform
Regional Grid Infrastructures

- CNGrid
- Garuda
- EUAsiaGrid
- SAGrid & SANREN
- GISELA
- EUMedGrid
- EUChinaGrid
- EUIndiaGrid
- NAREGI
- CNGrid
- DEIS
- KnowARC
- BalticGrid
- TeraGrid
- OSG
- EELA
- GISELA

Projects with other funding

European Commission co-funded projects
CHAIN: global coverage

Coordination & Harmonisation of Advanced eINfrastructures
Mediterranean Grid (EUMEDGRID)

39 sites (30 in production) in 16 countries
- **CNGrid Environment**
  - 14 sites
  - One OP Center
  - Some domain app. Grids
  - CNGrid GOS

- **Applications**
  - Resource & Environment
  - Manufacturing
  - Services

- **Sites with gLite/EMI (IHEP, PKU) in WLCG**

- **HPC Systems**
  - Two 100 Tflops
  - 3 PFlops
India National Knowledge Network

- A state-of-the-art multi-gigabit pan-India network for providing a unified high speed network backbone for all knowledge related institutions in the country
- An ultra-high speed CORE, multiples of 10 Gbps
**GARUDA**

- Connects 45 C-DAC centres and partners
- Over 6000 CPUs (~ 70 teraflops) and terabytes of mass storage to provide distributed data.

**GARUDA Grid is now powered by the NKN**
Red CLARA connects almost all the countries in Latin America and is managed by CLARA.

CLARA is interested in managing also computing initiatives.

Currently ROC-LA managing Grid operations.
15 Countries (11 in Latin America)
19 Partners (14 in Latin America)
12 Third Parties (9 in Latin America)
CHAIN: global coverage

Coordination & Harmonisation of Advanced eInfrastructures
CHAIN contributions

- State of the art in eInfrastructures analysis around the world
- Recommendations and guidelines
- Roadmaps
- Support to Intercontinental VRC
- Cross-platform access through Science Gateway

CHAIN-REDS activity on interoperations of the intercontinental infrastructures
Outline

- State of the art in eInfrastructures in Europe
- Worldwide Grid collaborations
- **CHAIN-REDS: Example of HPC collaboration in SEE**
- CHAIN-REDS: Example R&E cloud platform
High-Performance Computing in SEE

- 120 Tflops aggregate
- 2 BlueGene machines
- Bulgaria, Romania, Serbia, Hungary, FYRoM offering resources
- Procurements coming - Greece and Serbia
- 26 applications in 3 VRCs
- Envisaged as bridge to PRACE
- Joint operations centre studied and assessed
HPC Systems in the region
Flexible Access Mechanisms

- Pilot call for access to resources
  - Resources offered: 4.6M Core hours, 1.8 M GPU hours
  - Allocations for 1 year – starting December 2012
  - Peer review based
  - Access to the resources from all countries of the region
  - Access to LinkSceem applications

- Fast track access mechanism
  - Limited resources provided
  - 2 Month allocation period
  - Suitable for: New user communities – Non experienced users; 7 new applications so far
State of the art in eInfrastructures in Europe
Worldwide Grid collaborations
CHAIN-REDS: Example of HPC collaboration in SEE
CHAIN-REDS: Example R&E cloud platform
GRNET ~okeanos service

- ~okeanos is set to deliver *Infrastructure*
  - Compute (Virtual Machines)
  - Network (Virtual Networks, private Ethernets (L2) or public IPv4/6)
  - and Storage (Virtual Disks)

as a *Service*

- Alpha2: from March 2012 - 3000 VMs - 2000 alpha users; Beta version now
- Target group: GRNET’s customers
  - direct: IT depts of connected institutions
  - indirect: university students, researchers in academia
okeanos service
~okeanos service
- **Compute**: Cyclades
- **Files**: Pithos+
- **Images**: Plankton
- **Identity**: Astakos

- **Volumes**: Archipelago
- **Accounting/Billing**: Aquarium
Cloud for R&E: the process

- **Vision:** flexible, production-quality cloud services for Greek R&E community

- **Rationale:**
  - Step beyond Grid in terms of flexibility and availability
  - Economies of scale for the community; solving understaffing problems, poor service, low maintenance
  - Minimizes the investment in equipment and support contracts

- **Policy background:** existing MoU in place for Grid computing, expanding for HPC as well

- **Strategy:**
  - Technical workshops and requirements capture meetings
  - Gradual offering of services, starting with storage, moving to VM on demand, IaaS, and then SaaS
  - Paving the way to public sector

- **Funding:** in the context of GRNET4 project, 2.2ME DCs; 4.5ME s/w and services)
Commercial IaaS vs own IaaS

Commercial IaaS
- Amazon EC2 not an end-user service
- Need to develop custom UI, AAI layers
- Vendor lock-in
- Unsuitable for IT depts
  - persistent, long-term servers, custom networking requirements

Gain know-how, build on own IaaS → reuse for own services
~okeanos: Platform design

Web Client

CLI Client

Web Client 2

admin@home

OpenStack Compute API v1.1

GRNET Proprietary

Synnefo cloud management software

Google Ganeti

KVM

Virtual Hardware

Debian
~okeanos: up and running!

- Want you own cloud platform?
  - GRNET available to provide the open-source synnefo software

- Want to try GRET IaaS?
  - IaaS is open as trial for collaborating communities
  - 1 rack of 1000 VMs available via EduGain
  - Will be further articulated via CHAIN-REDS project
Diverse Grid solutions from around the world

CHAIN provided a platform for integration of Europe and other continents; CHAIN-REDS will continue this support

CHAIN-REDS will expand towards HPC and clouds, some examples given here
Co-ordination & Harmonisation of Advanced e-Infrastructures for Research and Education Data Sharing

Thank you

Research Infrastructures – Grant Agreement n. 306819