



From High Performance to Distributed to Cloud:

The Need for “Trusted” Computing

Suhaimi Napis, PhD

CTO and Director, InfoComm Development Centre

Research Director (Acting) and E-Research Area Malaysian Research and Education Network (MYREN)
Vice Chair, ASEAN Science and Technology Research and Education Network Alliance (ASTRENA)
Director, Natural Resource Area; Co-Chair, Agriculture Working Group, Asia Pacific Advanced Network (APAN)
Member, TaskForce for TransEurAsian Information Network Phase 4 (TEIN4)
Member, Steering Committee on ICT, Ministry of Higher Education

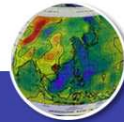
suhaimi@putra.upm.edu.my



www.euasiagrid.eu



www.euasiagrid.org



Outline

- Introduction
- Academic Grid Malaysia and BIRUNI GRID
- APAN, TEIN3, and Other Networks
- HPC, Grid, Volunteer and Cloud Computing
- Future Plans
- Concluding Remarks

FP7-INFRA-223791



Introduction



- **Motivation**

- Academic staffs, research officers/assistants, and post-graduate students **made up of over 80% of the entire research community in Malaysia!** (Not including undergraduates)
- Working on Grid computing and its application in various disciplines
- Requires minimal facilities in their labs of respective universities to carry out their R&D
- Need something that they can “**play around**” and should anything happens to the facilities (eg. Nodes crashing, hanging, etc) they can simply poweroff and restart!

FP7-INFRA-223791



Introduction-cont



- **Basically two types of users:**
 - Researchers/students doing research on certain aspects of GRID
 - Researchers/students using GRID for computation; (Bioinformatics, Cheminformatics, Medical Informatics)
- **Need basic Grid facilities and have total control of the entire Grid Ecosystem (all aspects of Grid)**
- **After acquiring certain level of expertise or if they want more compute power, then they can tap on other production facilities**
- **Hence the need for a learning and discovery Grid**

FP7-INFRA-223791



The Academic Grid (A-Grid)



- Generally, some universities have long been involved in Grid R&D and its application in several domains (Lifesciences, Pharmacy, Chemistry, Physics, Mathematics)
- But some universities are still at very initial stage and have only a few clustered PCs BUT they are very enthusiastic researchers and students
- Some universities have already assembled a number of small clusters of HPCs and SMPs for internal use but they are not connected together
- Some managed to connect together and created Campus Grid
- Researchers are reluctant to share the resources; even among research groups within a university
- Server hugging!!

FP7-INFRA-223791



The Academic Grid (A-Grid)



- Each University should form Campus Grid linking clusters of HPCs, SMPs and PCs.
- Campus Grids will be connected together via MYREN (MYREN Phase 2 is on Tender Process and expected to be fully implemented by Q4 or Q1-2010)
- A-Grid Operation Centre (A-GOC) will be formed (USM already received Provisional Certification Authority (CA) status and UPM is in the process of getting CA (considering CA for A-Grid and Malaysia)
- Some of the A-Grid facilities will be on production Grid and can peer with facilities at MIMOS or other Grid partners internationally via TEIN3, INTERNET2, and other networks
- Addressing redundancy

FP7-INFRA-223791



BIRUNI GRID



- UPM has just commissioned BIRUNI GRID with the hope of making it the Nucleus HPC Clusters for A-Grid
- Hardware Racking and Power up were done by suppliers while the rest of the configuration and deployment was done by UPM Grid Team
- Excellent learning opportunities for the staff
- Many lessons learned; some are good and some are not so good but in the end the staff learned by doing and got their training the hard way.
- UPM gets help from Academia Sinica Grid Computing and other international partners

FP7-INFRA-223791



BIRUNI GRID-cont



- BIRUNI GRID consists of three blade-based clusters:
 - Khaldun Sandbox Cluster (10 worker nodes with GB)
 - Razi Cluster (28 worker nodes with GB)
 - Haitham Cluster (8 worker nodes with Infiniband)
- Khaldun Sandbox Cluster is used for Experimental Grid (Development, and Learning purposes) while Razi and Haitham Clusters are designated as Production Grid.
- At any time, the Haitham Cluster and Khaldun Sandbox Cluster worker nodes can support Razi Cluster worker nodes on demand

FP7-INFRA-223791



BIRUNI GRID-cont



- **Middleware deployment:**
 - Scientific Linux 4.7 installed on each server, with an option to run virtualization on blade servers to support multilayer grid computing.
 - Lustre Filesystem 1.6 installed on two storage nodes with SAN as main storage target.
 - Head node provides grid operational application (software repository, routing into another network/internet, DNS and monitoring)
 - Two Blade Servers serve as Computing Element, VO manager, Information System, and gLite Middleware Operation.

FP7-INFRA-223791



BIRUNI GRID-cont



FP7-INFRA-223791



Network Connectivity

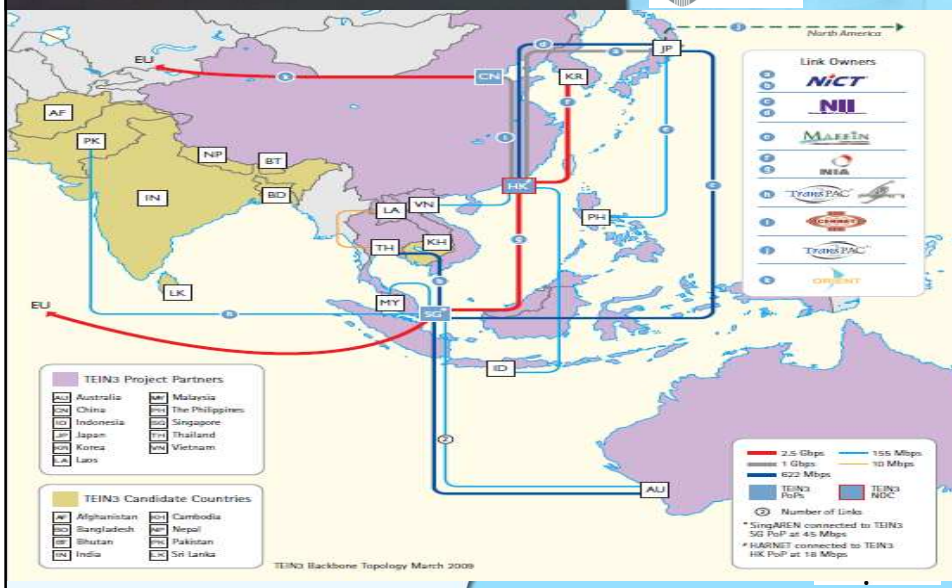


- Major Malaysian Universities, Institutions of Higher Learning and Research Institutions are connected to Malaysian Research and Education Network (MYREN)
- UPM was one of the co-founding proposer
- Connected to TransEurAsian Information Network (TEIN) now in Phase 3 with link of 155Mbps for Malaysia
- Internet2 and others link to the US

16/04/09
FP7-INFRA-223791

FP7-INFRA-223791

TEIN3 – Launched January 2009



TEIN3

ASTRENA Framework

Network – with regional nets

Exchange Point
Access Point
Current status

16/04/09
FP7-INFRA-223791

U.P.M. UNIVERSITI PUTRA MALAYSIA
EUAsiaGrid
Tein
DANTE
GEANT2 LAUNCH EVENT 14-15 June 2005
ASEAN SCIENCE & TECHNOLOGY RESEARCH & EDUCATION NETWORK ALLIANCE
AGRID MALAYSIA
MALAYSIA GRID FORUM

Future Plans

- Efforts to have Certification Authority for UPM and hopefully for Academic Grid
- Grid Application deployment for Bioinformatics/ Cheminformatics and Medical Informatics and Graphic/Image Rendering
- Promote greater awareness on Grid Computing and shared resources among researchers and stakeholders of Institutions of Higher Learning in Malaysia through the Ministry of Higher Education which will soon be the custodian of MYREN Phase 2
- Registration of Malaysia Grid Forum Society to be affiliated with OGF, launching of Grid Journal, yearly scientific conference on Grid

FP7-INFRA-223791

U.P.M. UNIVERSITI PUTRA MALAYSIA
EUAsiaGrid
ACADEMIC GRID MALAYSIA
MALAYSIA GRID FORUM

HPC, Grid, Volunteer, Cloud



- **HPC = High Performance Computing Cluster(s)**
Resources in a single site
- **Grid = Distributed Computing; Resources from Cluster of Clusters**
- **Volunteer = Computing Resources provided by people “donation”** like BOINC, Asia@Home, Folding@Home, LHC@Home projects
- **Cloud Computing: Resources provided as a service over the Internet**

FP7-INFRA-223791



Cloud Computing 101



- **Infrastructure as a Service (IaaS): All computer infrastructure provided as a service; complete outsource**
- **Platform as a Service (PaaS): Web applications and services available entirely over the internet**
- **Software as a Service (SaaS): Licensed application as a service on-demand**

- **Simple concepts but there are advantages and disadvantages**

FP7-INFRA-223791



Cloud Computing 101



Pros:

- Avoids capital expenditures by renting usage, pay only when used, few or no ICT staff, minimise risks on return of investment (ROI)

Cons:

- Total dependencies to the services provided; Limited to what's on offer; Data storage issues, Security and privacy issues; may sacrifice personal data to third parties

16/04/09
FP7-INFRA-223791

FP7-INFRA-223791



Cloud Computing 101



• Issues:

- Service Level Agreement (SLA)
- Quality of Services (QoS)
- Provenance (who, when, where it comes from ie origin of jobs, raw data, etc)

• Risks:

- Privileged user access (vetting of administrators)
- Regulatory compliance (external audits/security certs)
- Recovery of Data in case of disaster
- Data location, segregation and availability
- Can the provider be trusted??

FP7-INFRA-223791



Concluding Remarks

- Promote **TRUST** and **CONFIDENCE** and encourage sharing for example by implementing and acculturating Federation (eg VOMS or GridShib, Digital Certificates) and developing policies for harmonization of Raw and Clean Data coming from various disciplines
- Hide the complexity and technical details of the Grid from end users and application developers
- **Risks and issues (plenty) = Opportunities** for University-Industry R&D initiatives
- **MDeC shall play its role**

FP7-INFRA-223791




“Trusted” Computing

NO TRUST....
NO GRID!



FP7-INFRA-223791



Upcoming Events



28th Asia Pacific Advanced Network (APAN) Meeting

20-24th July 2009, Berjaya Timesquare, Kuala Lumpur
www.apan.net

GridDotMy2009 (Grid Scientific Local Conference)

21 and 22 July 2009, Colocated at the same venue as APAN)
email: syazwina@putra.upm.edu.my

EUAsiaGrid Summer School (Grid Training)

27 July to 7 August, 2009, iDEC, UPM (Two-week hands-on)
<http://indico.twgrid.org/confRegistrationFormDisplay.py?confId=703>

ALL OPEN FOR REGISTRATION

FP7-INFRA-223791