Malaysia's Computational Journey Towards Science DMZ

Suhaimi Napis
Technical Advisory Committee (Research Computing) MYREN-X
Universiti Putra Malaysia
suhaimi@upm.my
In the Beginning...

- Research on parallel/distributed computing since 1990s
- A proposal for National Centre for Scientific Computing (NCSC) championed by UM was not able to secure support by the then Ministry of Science, Technology and Environment (MOSTE) in 2000
- Early initiative called EMAS-GRID linking small clusters at USM, UKM and UM carrying out bioinformatics research in 2003-4
- Another project, TEMAN (Testbed Environment for Multimedia Application and Networking (2001-2) was established and shifted to TEMAN II initiative which eventually becomes MYREN (2005)
Malaysian researchers have for a long time recognised the importance of HPCs and Grid computing.

In the early 2005, around 90 researchers from universities and research institutes had a workshop to re-initiate the development of advanced computing.

The outcome of the workshop was a National Grid Computing Initiative Roadmap and Terms of Reference (Version 1.0).

The Technology Roadmap was revised with major changes of the original document and implemented multi-vendor, multi-technology, multi-platform, multi-"x"; which was called KnowledgeGrid Malaysia with several applications: Automotive Grid, Animation Grid, Bioinformatics Grid, etc.

In the Beginning...cont.
The Governance committee eventually revised the National Grid Computing Initiative Terms of Reference back to its original intent (Revised version 2.0)

The RM132mil (USD35Mil) facility was decommissioned at the end September 2012... :( :( 
Malaysian researchers have for a long time recognised the importance of HPCs and Grid computing.

First meeting to discuss about the formation of Academic Grid among institutions of higher learning under Ministry of Higher Education was held on 28 September 2007.

Representatives from several universities attended the meeting and gave updates of Grid activities in respective universities.

Agreed unanimously on the formation of Academic Grid to support teaching and learning of Grid and sharing of knowledge and resources among universities.

Recognise the two types of users:
1) Application users and
2) Technology developers
In 2009, a group of universities lead by UPM have taken the initiatives to create a more distributed grid infrastructure.

A VO was formed for operationalizing the grid as a true sharing of distributed resources where resources are really distributed across the network.

The VO was known as Academic Grid with 5 participating universities; UPM, UM, UTM, USM and UUM.

Some of the university sites participates in VOs outside Malaysia under the EUAsia Grid Project where Resources at Academic Grid Malaysia received jobs from international partners and vice versa.
Academic Grid Malaysia continues to function until today but clusters at all sites are "independent" of each other.

With the establishment of National Centre for Particle Physics at the University of Malaya, new resource has been procured at the university.

Data Intensive Computing Centre (DICC), of the University of Malaya, currently has a 11-node, 702 processor-core and 2.7 TB RAM Linux cluster. It consists of 11 nodes of CPU unit and 2X server for cloud services and computational, also GPU farm consists of 4 servers with K10, K20, K40 and Titan cards. Yahoo donated servers currently being configured to be added to the infrastructure.
Examples of Current Research and Needs

- Tier 2 Site Complex Muon Solenoid Expt (National Centre for Particle Physics, University of Malaya)
- Collaboration for data transfer, storage and analysis of genomic projects; terabytes of data produced from next-generation sequencing (UM, UKM, UPM, USM, UITM, etc);
- Square Kilometer Array, Radio Telescope and Astronomy (UPM, UKM)
- New HPC at Meteorological Department, Malaysia dedicated for weather and disaster prediction and simulation. (Unfortunately METmy is not connected to MYREN yet)
- Others including national highspeed innovation platform and testing for new devices
The Science Cloud: Research Computing Services

- High Performance Computing (HPC) and Federated Cloud Computing that caters for researchers’ requirements for:
  - General Purpose Compute Resource
  - RAM-intensive Compute Resource
  - GPU-Based Compute Resource
  - Desktop Compute Resource
  *For Science and Engineering Research, Animation Rendering, Big Data, etc*

- Virtual Servers (VMs) and Storage provisioning for researchers

- Virtual Workbench for various domain specific analyses

- Common Storage Facility for Research Clearing House (Research Repositories eg. Databases for DNA/Protein, Biodiversity, etc)

- Adopting Science DMZ model; A Private Community Cloud that will generate contents and traffic for MYREN-X
Some Expected Benefits

- MYREN-X is now an ISP and as of 1st April, 2018 it will be under a company established under Consortium of Public Universities. It is for all research and education entities both public and private. Connectivity to all university campuses and also to all polytechniques and community colleges will be from 10 to 100Gbps.

- Not just a network but will also provide research and education computing services

- Broader use of cyberinfrastructure within the research community including the sciences, medicine, arts and humanities and social science

- Increased capability, expertise and effective use of cyberinfrastructure

- Enhanced and easier to use Distributed Computing Infrastructure (Grid/High Performance Computing and Federated Cloud) that is production ready with application softwares and scientific gateway portal
Some Expected Benefits

- New ways of retrieving, processing, visualising and archiving data, opening up new areas of research and expanding existing ones whilst allowing the results of research to be shared by a broader community.

- With trust framework afforded by SIFULAN it is hoped that the spirit of sharing can be enhanced simply because we know who access what, at what time and for what purpose (i.e. auditable) because every “transaction” must go through the Authentication, Authorization and Audit (AAA).

- Let researchers do what they do best (i.e. research) and not be bogged down to worrying about the computing requirement; worst still having to develop it on their own!
Concluding Remarks

- Research computing requirements of researchers must be addressed from both users and developers.

- Move towards Shared "decentralised centralised" cyberinfrastructure for optimal use and cost-saving.

- Trust Framework by SIFULAN will pave the way for greater spirit of sharing.

- Cyberinfrastructure requirements for research computing may be included in MYREN-X Project using Science DMZ as a model.

- Present regional and global developments in HPC ecosystem and next-generation interconnects of Asia Pacific Research Platform (APRP) is timely for future development of cyberinfrastructure.