Experience from EUAsiaGrid

Marco Paganoni
INFN, Italy

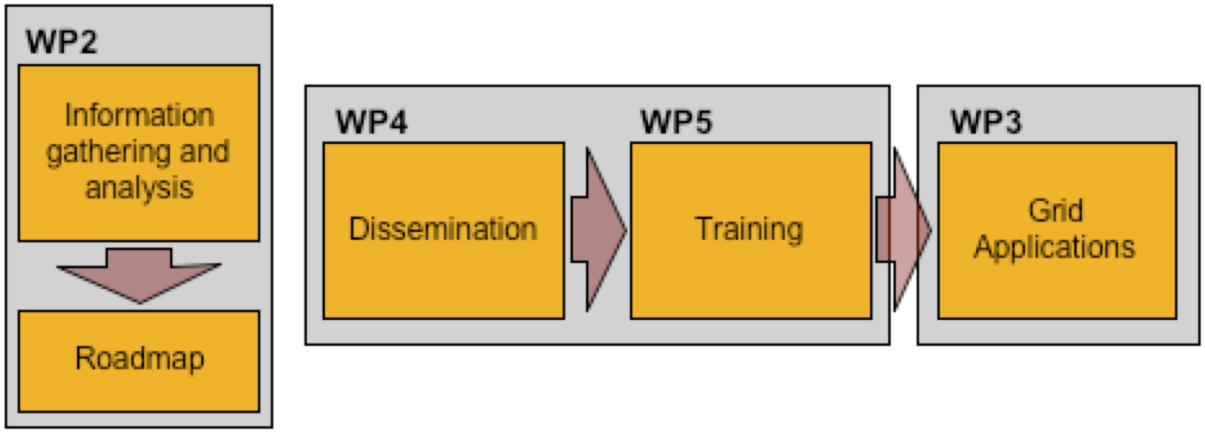
EGI InSPIRE Asia-Pacific workshop
Taipei, Mar 9th , 2010

Partners

1. Istituto Nazionale di Fisica Nucleare (Italy) (coordinator)
2. CESNET (Czech Republic)
3. University of Manchester (United Kingdom)
4. HealthGrid (France)
5. Ateneo de Manila University (Philippines)
6. Australia National University (Australia)
7. Academia Sinica (Taiwan)
8. Advanced Science and Technology Institute (Philippines)
9. Hydro and Agro Informatics Institute (Thailand)
10. Infocomm Development Authority (Singapore)
11. Institute of Information Technology (Vietnam)
12. Institute Teknologi Bandung (Indonesia)
13. National Electronics and Computing Technology Center (Thailand)
14. University Putra Malaysia (Malaysia)
15. MIMOS Berhad (Malaysia)
16. Institut de la Francophonie pour l'Informatique (Vietnam)
17. National University of Singapore (Singapore)

Breakdown in work packages

Work Package	Title	WP Coordinator
WP1	Project management	INFN (M. Paganoni)
WP2	Requirement capture and coordination policy definition	UNIMAN (A. Voss)
WP3	Support of scientific applications	CESNET (L. Matyska)
WP4	Dissemination	ASGC (V. Huang)
WP5	Training	INFN (M. Fargetta)



Challenges of the Project

First Grid project targeting Asia-Pacific region

- ✓ Geographically large and culturally diverse area
- ✓ Uneven levels of adoption of Grids
- ✓ Interoperability with volunteering computing

Wide range of scientific domains addressed

- ✓ Consolidate on traditional areas and engage new communities
- ✓ Dissemination, Training and Support for applications must create a virtuous cycle
- ✓ Focus on the benefit of international collaboration (EU – AP)

Sustainability of the e-infrastructure

- ✓ Define a roadmap for the follow up of the project
  - **MoU to manage the infrastructure after June 30th, 2010**
  - **CHAIN project**
- ✓ Connect end users and stake holders in a stable way

Project impact in the 1st year

Data collection from SPR and Survey

- ✓ First assessment in the AP region of requirements from scientific communities and resources from project partners

EUAsia VO allows wide area outreach

- ✓ Crossing the borders between nations and disciplines

Different groups of people addressed with training

- ✓ Events directly organized by the project and steering of local events
- ✓ "Train the trainers" concept followed
- ✓ "Keep an eye" on trained scientists

Wide audience reached for dissemination

- ✓ iSGTW as main EUAsiaGrid media partner
- ✓ Publication on peer-reviewed journals and proceedings of non-ICT conferences

Starting point for 2nd year

Data challenges achieved

- ✓ Avian Flu DC2 refinement & Dengue Fever drug discovery
- ✓ Earthquake mitigation
- ✓ Climate change mitigation

Complex workflow supported

- ✓ Addressing rare and neglected deseases
- ✓ Mitigating natural disaster human and economic effects

Interesting new scientific domains joining the effort

- ✓ Digital Culture and Heritage
- ✓ Social Science simulation

Concentrate on applications

- ✓ application repository
- ✓ application porting

EUAsia VO

- Generic, application neutral, regional VO for AP region
- Starting from ASGC and UPM, now also EU resources from CESNET, INFN
- Based on "catch-all" approach established in EGEE for researchers from any discipline with a simplified registration procedure
- Resources in EUAsia VO (~ 600 cpu and 65 TB)
- About 200 users
- All partners have UI to access resources
- All partners have set up RA/CA

Name of Institution	CPU(Cores)	Disk(GB)
IT_INFN-CT	100	460
CZ_CESNET	72	8754
ID_ITB	1	15
MY_UPM	126	983
MY_UM	24	134
MY_MIMOS	140	1246
PH_ASTI	56	10
PH_ADMU	1	30
TW_AS	100	51200
TH_HAII	2	1083
TH_NECTEC	4	1474
VN_IFI	8	733
VN_IOIT	4	733

Dengue Fever Data Challenge

- **Objective**
  - ✓ Reduce time and money for drug discovery by Grid in silico simulation
  - ✓ 2.5B people are at risk, 50M cases/year in more than 100 countries
  - ✓ 95% cases are children < 15 yrs in South-East Asia
- **Milestones**
  - *Validate the Pilot run – M15*
  - *Data Challenge on Dengue Fever (NS3 protease structure with ZINC compound library) – M18*
  - *Pilot new designed target structure for in silico simulation – M18*
  - *Post data analysis for the NS3 target – M20*
  - *1st step wet-lab study – M22*
  - *Deploy the VSS on Asia@home – M18*
  - *Enhance the GVSS with APBS – M22*

Dengue Fever Data Challenge /2

- **Grid Computing environment**
  - *EuAsia VO*
  - *GAP Virtual Screening Service (GVSS)*
  - *ASGC provides the core services: GAP-VQS, DIANE2, AMGA, SE*
- **User requirements**
  - *CA*
  - *EuAsia VO*
  - *GAP VQS account*
  - *Download and install GVSS client package*
  - *Registered CDI compounds' data sets*

Drug Discovery - DC

1. Avian Flu - DC2 refined
   - 8 avian-flu mutant targets from EGEE DC2
   - 20,000 highest scored ligands from EGEE DC2 results
2. Dengue fever
   - Dengue NS3 target
   - 300,000 ligands which are prepared from ZINC.

	Number of dockings	CPU time	CPU-cores used on EUAsia VO	Status
1	20,000	3 years	125	Completed
2	300, 000	12 years	125	Completed

Grid-based International Network for Flu Obervation (g-INFO)

Institut de la Francophonie pour l'Informatique (IFI), Hanoi, Vietnam

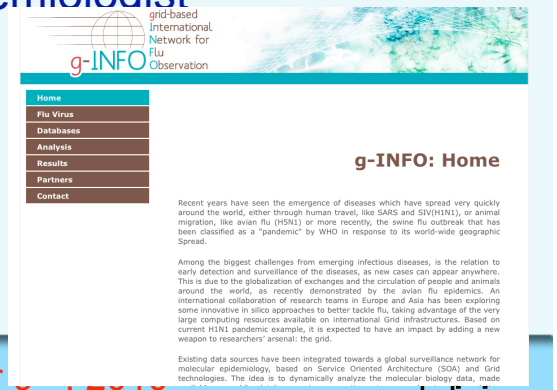
Goal: Dynamically analyze the influenza molecular biology data, made available on public databases using computing, storage and automatic updating services offered by grid technology.

Status:

- ✓ The Information System based on AMGA has been rebuilt to automatically synchronize public influenza databases like NCBI with grid resources.
- ✓ Definition and deployment through the WISDOM Production Environment of a workflow to analyse new virus sequenced: 1*) Alignment 2) Curation*

3) Phylogeny 4) Tree

- ✓ Capture of user requirements, close work with an epidemiologist
- ✓ Website: http://g-info.healthgrid.org/



Plans:

- ✓ Upgrade the pipeline
- ✓ Upgrade the web portal with friendly visualization

Earthquake Simulation

Objectives

- ✓ Dense sensor networks + historical data + simulation model
  → accurate simulation
- ✓ Seismic wave propagation analysis to reduce potential impacts
- ✓ Planning for real time hazard mitigation

Methodology

- ✓ Realize regional seismic wave simulation model over gLite mw
- ✓ Validate accurate wave simulation and implement case studies on disaster mitigation, country by country,
- ✓ Seismic sensor networking and data federation (TW, PH, VN, ID)

Roadmap

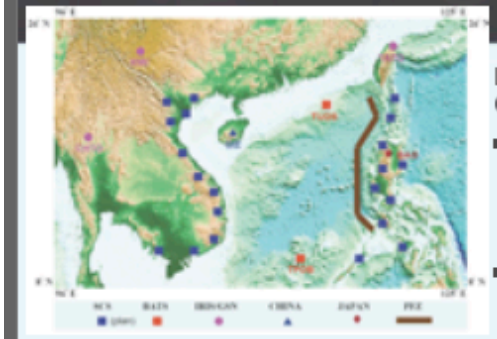
- ✓ Porting on Grid of seismic wave propagation analysis – M18
- ✓ Case Studies – M22
- ✓ Feasibility Analysis for tsunami disaster mitigation – M23

Earthquake Simulation / collaboration

	Sensor Network	Seismic Wave Prop. Analysis	Seismic Data Center
Partners	VN, PH, ID, TW		
User Community	Philippine Institute of Volcanology and Seismology (PHIVOLCS), Vietnamese Academy of Science and Technology, The Incorporated Research Institutions for Seismology (IRIS), Global Seismic Network (GSN), Institute of Earth Science & National Central University, Taiwan, Local, Regional, and Global Disaster Mitigation Organization.		
Tech Maturity	TW- Most dense; VN- Sensor Stn ready; PH- expanding	Analysis Model and knowledge available; Cluster and gLite Resources in place; From Global model toward higher resolution regional/country model;	SeisGrid@TW, IRIS, GSN
Exemplar	Integrated Sensor Network by VN, PH and TW	Without local geological data, accurate analysis is not achievable. High resolution historical TW earthquake data sets.	Federation of available Data Centers

e-Science for Earthquake Disaster Mitigation

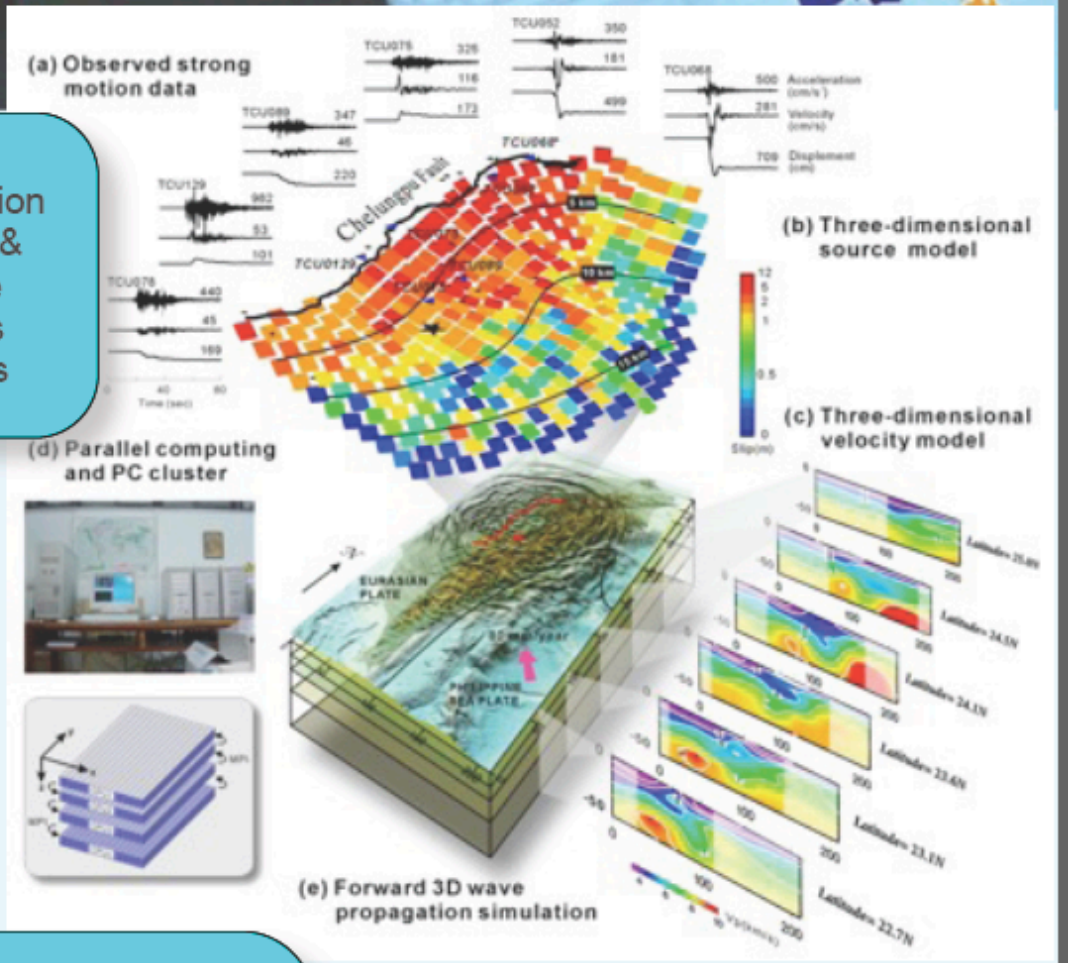
Seismic Sensor Networks



Local Sensor & Observation Data

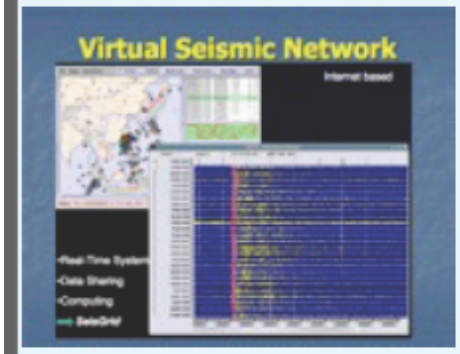
Global/Regional Sensor Data

High Resolution Source & Rupture Process Analysis



Fast Reporting System

Archive



Earthquake Data Center (SeisGrid)

Ref. Historical Events Data

Archive

Forward Simulation & Event Construction on Grid

Risk Analysis & Reduction

FP7-INFRA-223791

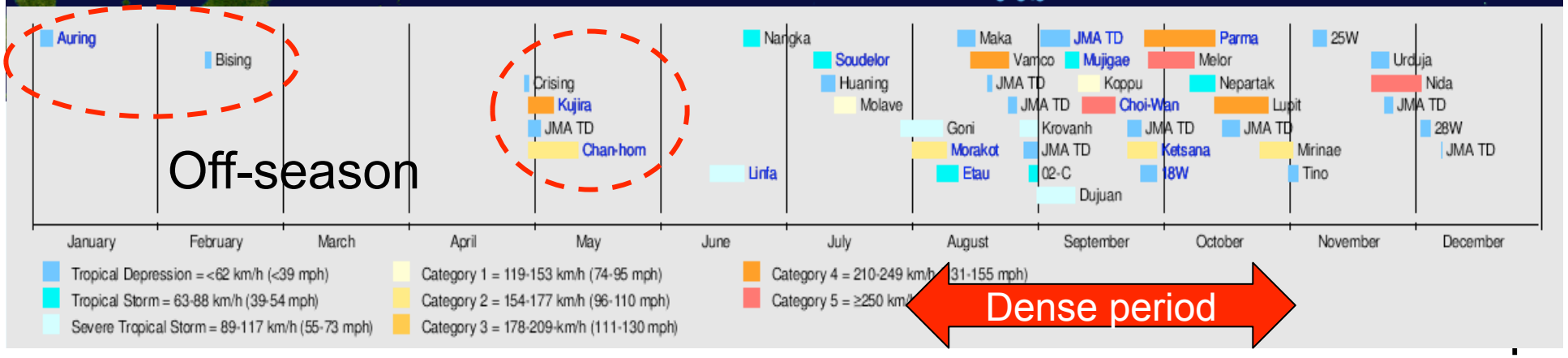
Climate change mitigation

- ✓ Climate change causes more frequent and severe weather extremes
  "Unusual" events are becoming "usual" !!
- ✓ Assessment should be based on deterministic instead of stochastic approach
- ✓ Risk management must be based on possible maximum and minimum rather than average values
- ✓ Collaboration between HAII and University of Cantabria under way and very promising

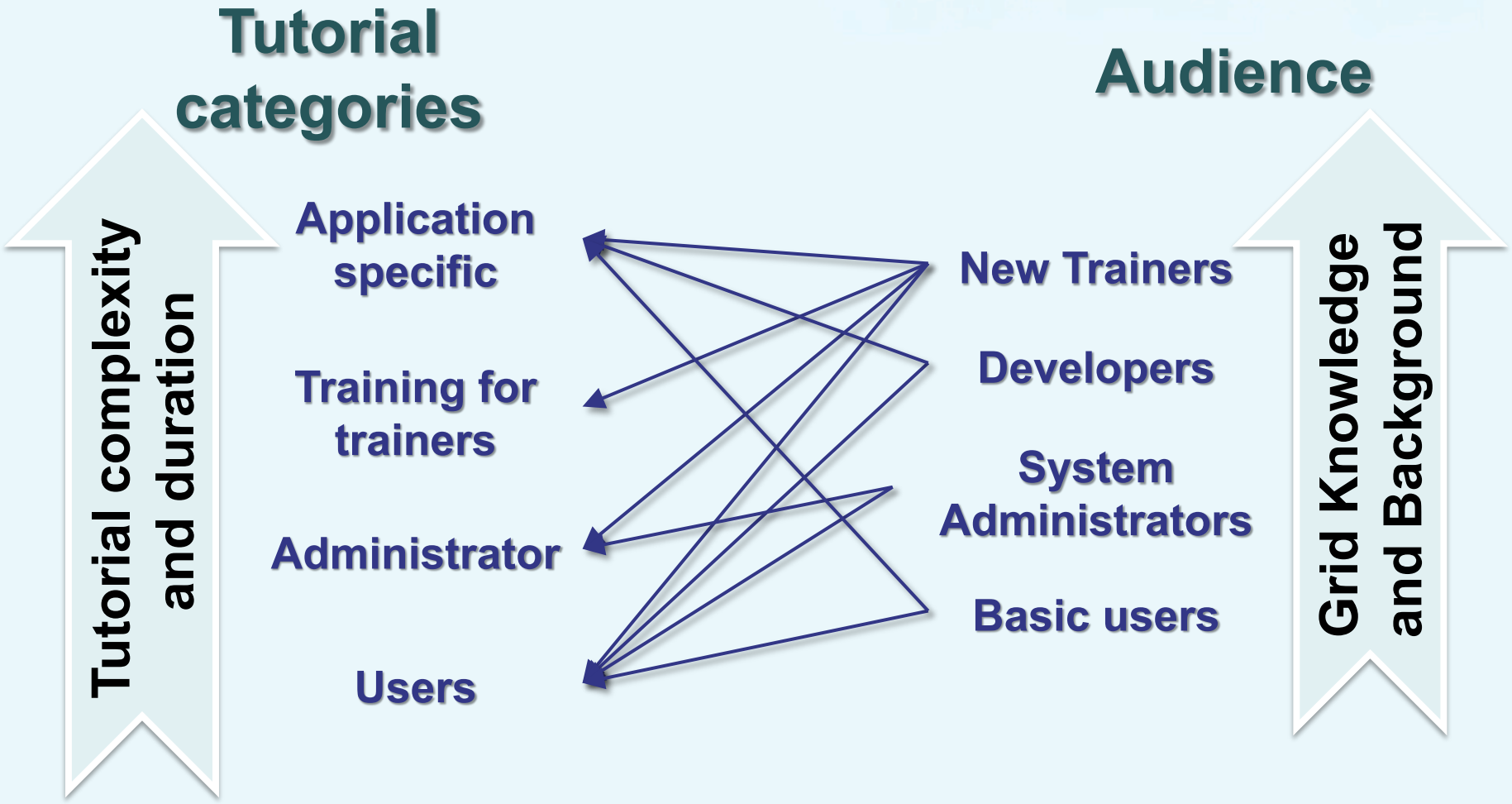
Pacific Typhoons in 2009

Total depressions:	41
Total storms:	22 official, 3 unofficial
Typhoons:	13 official, 2 unofficial
Super typhoons:	5 (unofficial)
Total fatalities:	>2301 total, 240 missing
Total damage (2009):	$10.836 billion

Considered the worst in decades!
(e.g. Ketsana, Parma)



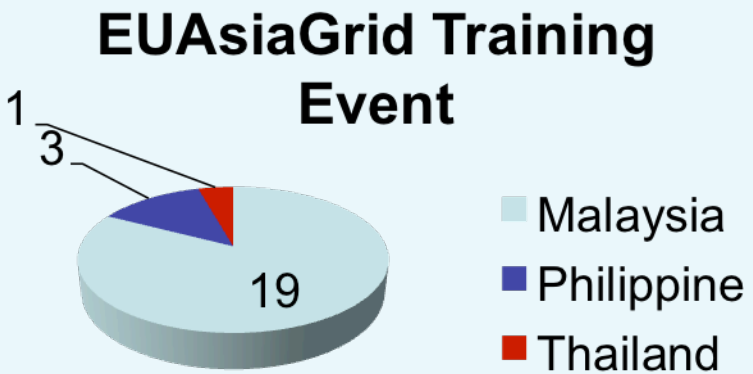
Training tailored to audience

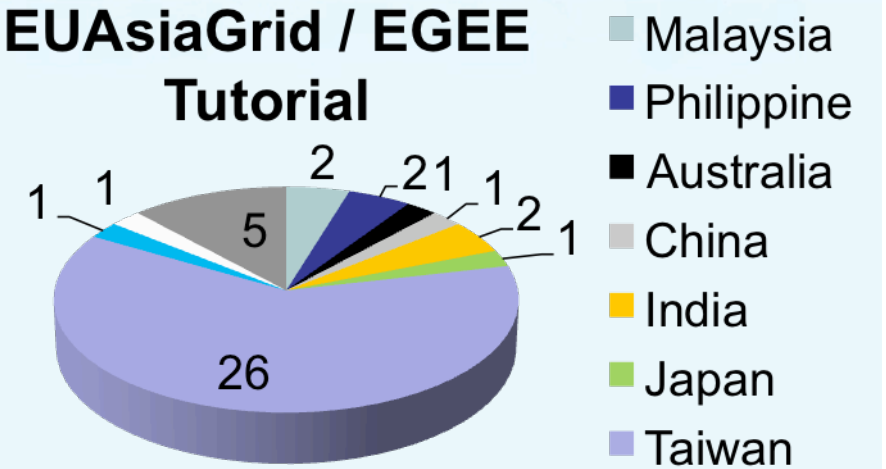


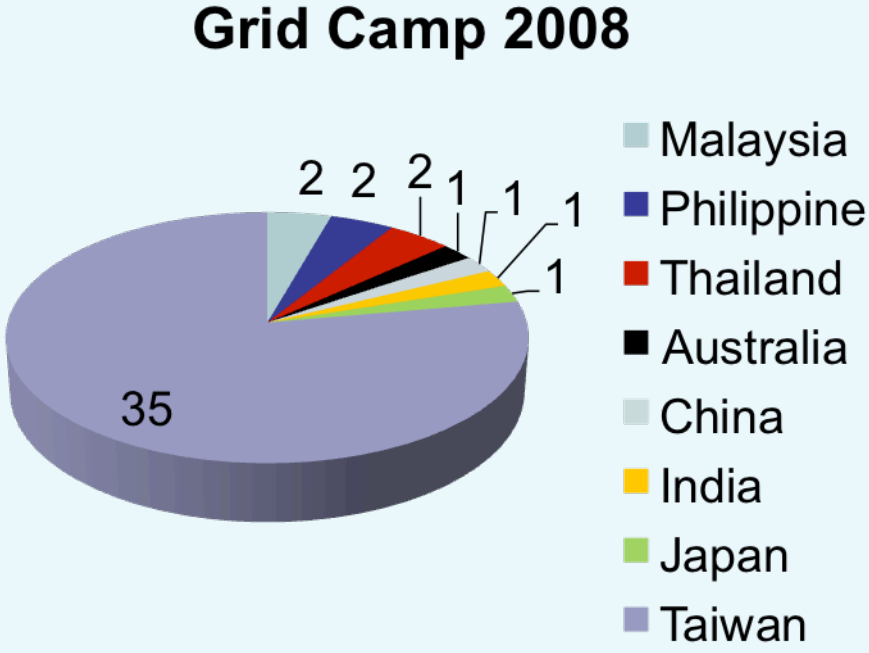
Training events

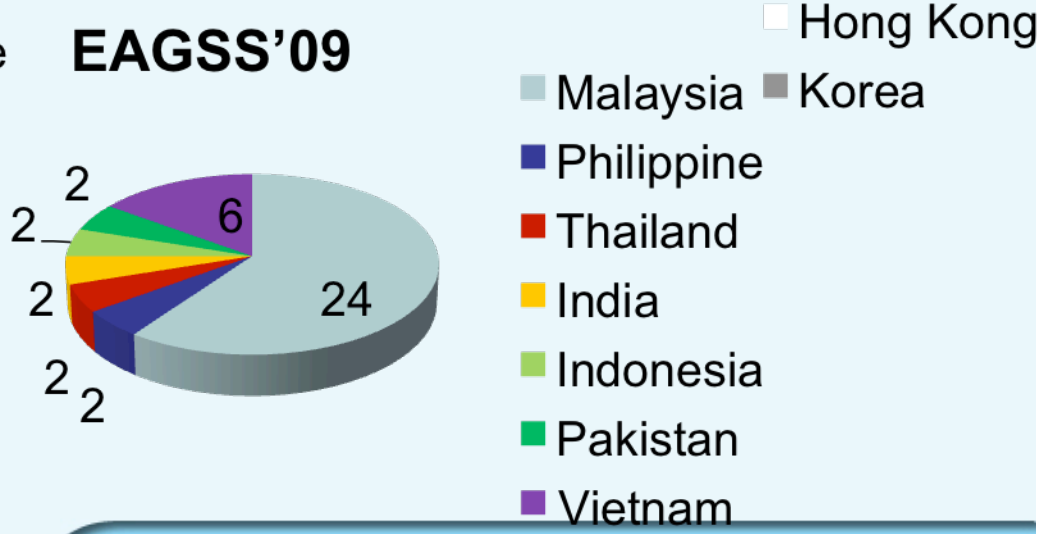
- **EUAsiaGrid Training Event at MIMOS**
  Kuala Lumpur, Malaysia, 28-31 Jul 2008
- **Grid Camp 2008 at ASGC**
  Taiwan, Taipei: 18-24 Oct 2008
- **EUAsiaGrid / EGEE tutorial at ASGC**
  Taiwan, Taipei: 18 Apr 2009
- **EUAsiaGrid Summer School**
  Kuala Lumpur, Malaysia, 27 Jul - 8 Aug 2009
- **EUAGSS09 (with CNRS): sysadmins + new grid users**
  Hanoi, Vietnam, 27 Sep - 03 Oct 2009
- **Train the trainers at ASGC**
  Taipei, Taiwan: 05-09 October 2009
- **ACGRID II (with CNRS): applications + train the trainers**
  Kuala Lumpur, Malaysia, 02-14 Nov 2009
- **gLite training for trainers at ASGC**
  Taipei, Taiwan, 5-7 March 2010
- **Social Simulation Tutorial**
  Taipei, Taiwan, 7 March 2010

Participation to training events









A vision for an Asia-Pacific e-Infrastructure for Research and Education

EUAsiaGrid

Roadmap

e-infrastructure

Roadmap document

- ✓ Organizational and technical roadmap towards a robust, scalable and sustainable e-Infrastructure in Asia-Pacific integrated with EU worldwide e-Infrastructure
- ✓ It addresses:
  - **Policy Makers and Funding Agencies**
  - **Providers of the computing resources**
  - **Scientific Communities**
- ✓ Building on the experience of EELA-2 and SEE-GRID
- ✓ EUAsiaGrid key partners are forming a JRU to participate into CHAIN, a proposal submitted to EU for worldwide support for e-Infrastructure
- ✓ Envisaging a gradual approach moving from JRU to NGIs
- ✓ Profit of good public-private collaboration in the region

Thanks for your attention!

marco.paganoni@mib.infn.it

http://www.euasiagrid.eu/