



Digital Divide for Global e-Science: Taiwan in Asia Context

Simon C. Lin (林誠謙)

Academia Sinica Computing Centre, Taipei

sclin@sinica.edu.tw

24 May 2005



What is e-Science?

- “e-Science is about global collaboration in key areas of big science and the next generation of infrastructure that will enable it”
-

by Dr John Taylor

Director General of the UK Research Councils

e-Science reflects growing importance of international laboratories, satellites and sensors and their integrated analysis by distributed teams



- **Scale of the problems**
 - Science increasingly done through distributed global *collaborations* enabled by the internet
- **Grids provide access to:**
 - Very large data collections
 - Terascale computing resources
 - High performance visualisation
 - Connected by high-bandwidth networks
- **e-Science is more than Grid & WSRF**

It provides the unprecedented linkage between People, Computers, Data and Instruments that will enable a Paradigm Shift!

Water, water, everywhere,
nor any drop to drink

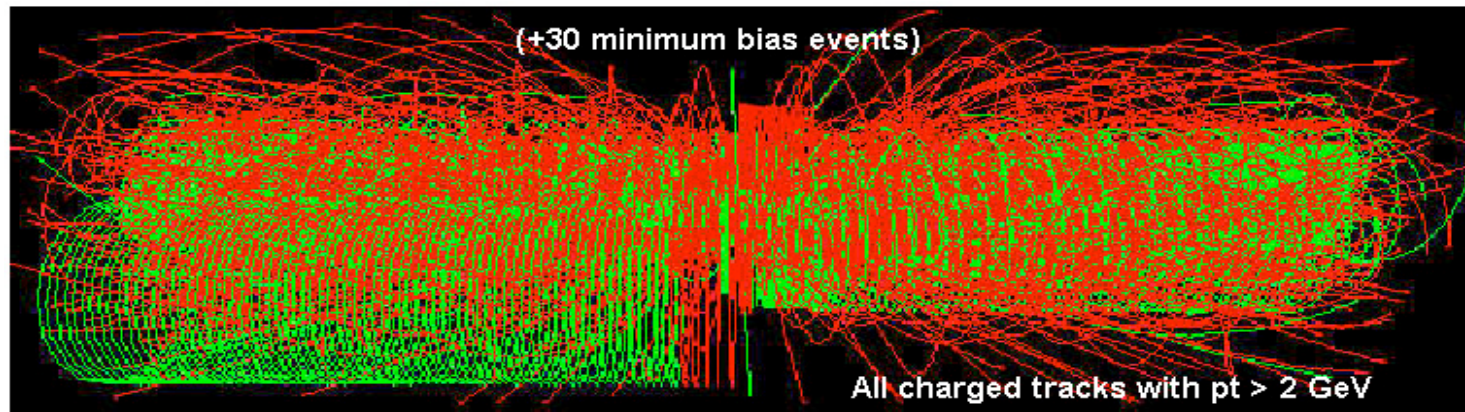
S.T. Coleridge, 1797



The LHC Data Challenge

Selectivity: 1 in 10^{13}

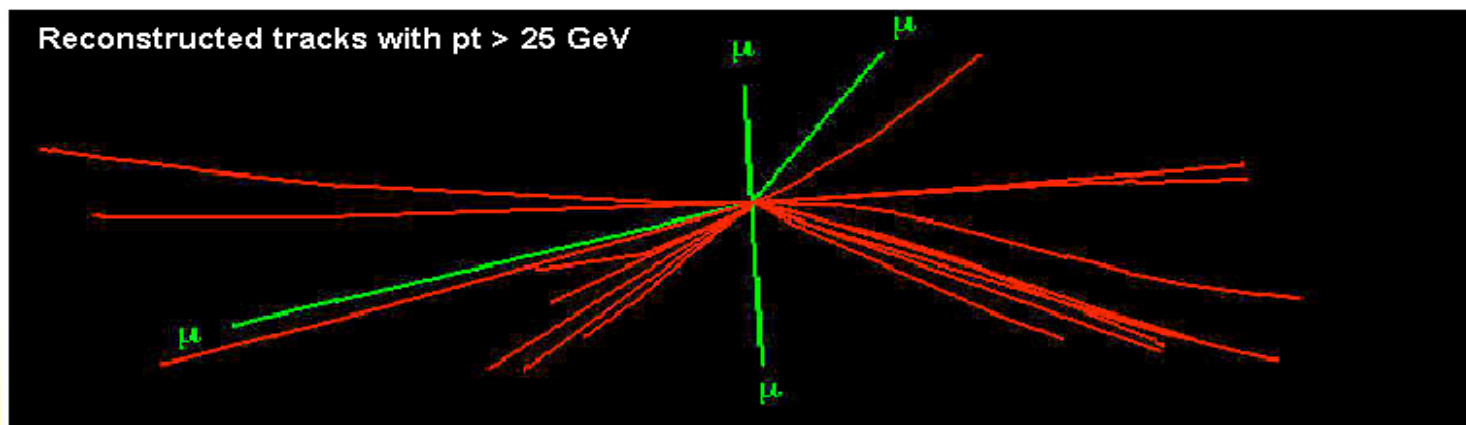
- Starting from this event



Like looking for 1 person in a thousand world populations!

Or for a needle in 20 million haystacks!

- You are looking for this “signature”



Src: CERN
5





The Computing Needs will grow

- It may grow to the scale of **ExaBytes** of Data and **PetaFlops** Computing by **2015**, in particular, the luminosity will be enhanced even in early stage
- The largest commercial database currently can only handle tens of TeraBytes
- The fastest stand-alone computer now is only capable of delivering 70 TeraFlops peak

- **RAM time to move**
 - 15 minutes
 - **1Gb WAN move time**
 - 10 hours (\$1000)
 - **Disk Cost**
 - 7 disks = \$5000 (SCSI)
 - **Disk Power**
 - 100 Watts
 - **Disk Weight**
 - 5.6 Kg
 - **Disk Footprint**
 - Inside machine
- **RAM time to move**
 - 2 months
 - **1Gb WAN move time**
 - 14 months (\$1 million)
 - **Disk Cost**
 - 6800 Disks + 490 units + 32 racks = \$7 million
 - **Disk Power**
 - 100 Kilowatts
 - **Disk Weight**
 - 33 Tonnes
 - **Disk Footprint**
 - 60 m²

May 2003 Approximately Correct *Distributed Computing Economics*

Jim Gray, Microsoft Research, MSR-TR-2003-24



Specific Challenges in Asia

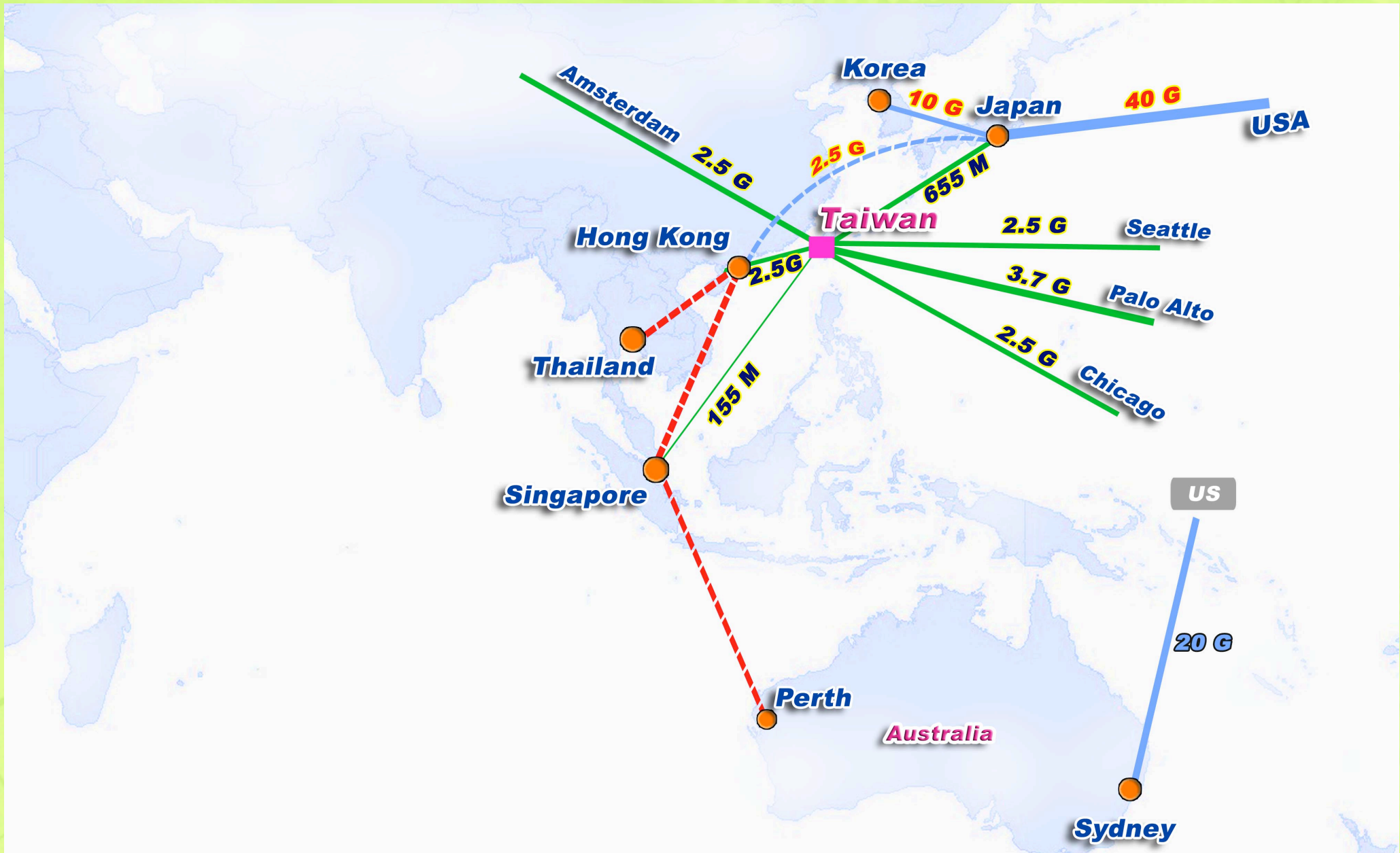
- US-centric Network connectivity
- Weaker collaborative scientific tradition within the region, thus, the culture for collaboration to be built
- Legal context of collaboration to be drafted
- Insights in scientific and technological policy to address its particular circumstances still to be developed





Digital Divide Status of Taiwan in Asia

7 Points of Presence from Taiwan





Taiwan Link Status to LCG

- ASNet* runs one STM16 IPLC from Taipei to Amsterdam
- Two GE Local loop from Amsterdam to Geneva (via NetherLight LightPath Service)
- Plan to double the IPLC to Amsterdam on Summer 2005
- ASCC LCG facilities now have multiple GE uplinks to ASNet's core router; will be replaced by multiple 10GEs this summer
- ASNet connects (10GE*2 + STM64*n) to domestic backbone, a.k.a. TANet/TWAREN joint backbone, to reach T2's in Taiwan.
- Every T2 in Taiwan has it's own 10GE link to the domestic backbone

* ASNet (Academic Services Network, as#9264) is the network division and is also the network name that registered in APNIC.



Services of ASNet / APAN-TW

- Dark fibre and High Speed (GE, 10GE, STM16, STM64) IPv4/IPv6 links for RENOs in the capital city of Taiwan (the TaipeiGigaPoP Metropolitan Network System, exchanging 50 Gbps traffic)
- IPv6 transition help and consultancy for RENOs and ISPs
- H.323 conference hosting
- International transit and links
- Managed international layer-2 circuits (the “EzLink” service for APAN members)

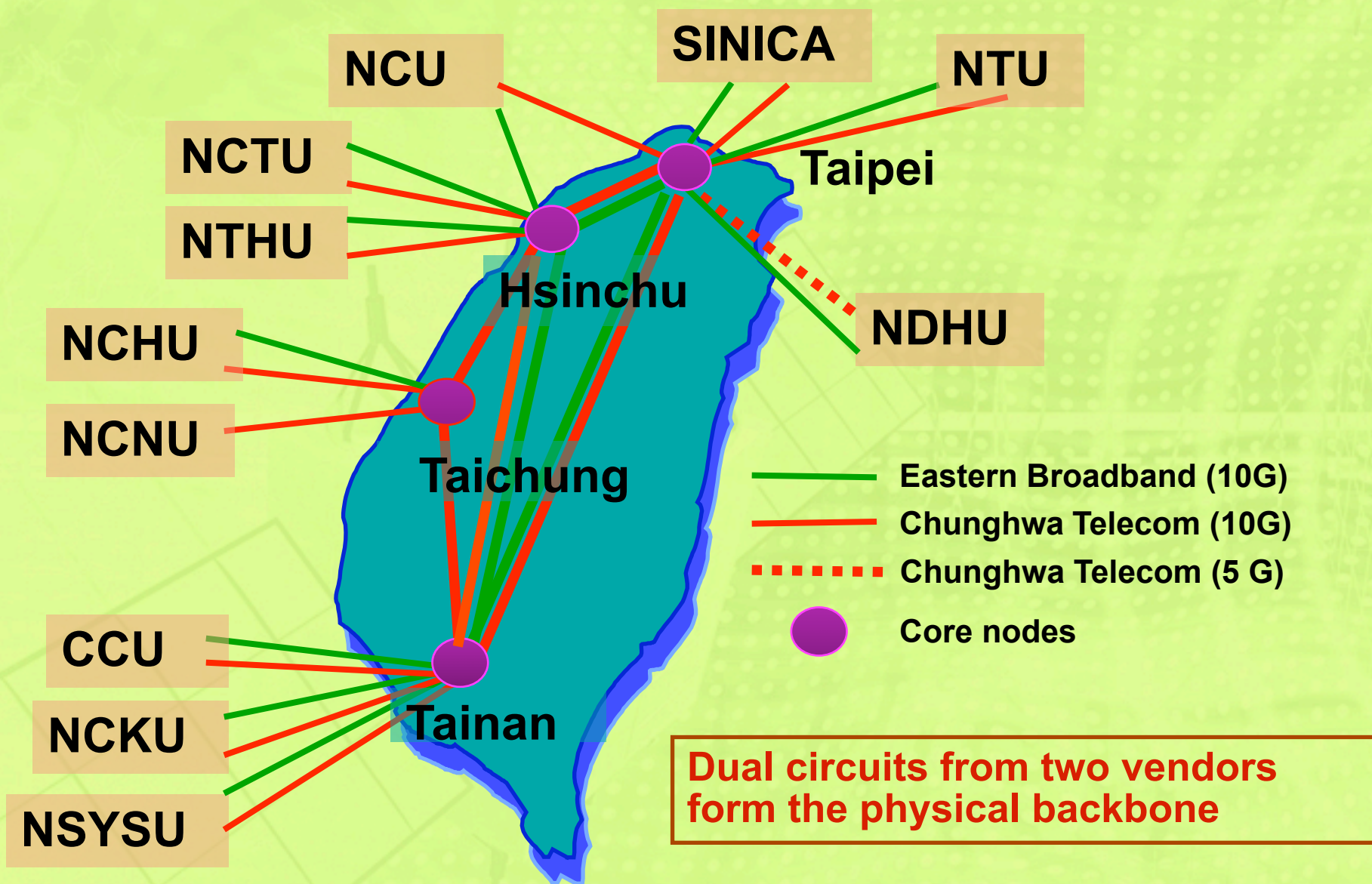


TaipeiGigaPoP

- “**TaipeiGigaPoP**” is the first pure dark fiber system in Taipei City to promote and provide IP-over-fiber service for research and educational communities since 1999
- The successful story of “**TaipeiGigaPoP**” encourages TAnet and ISPs in Taiwan to deploy metro-network services
- Today, IP-over-fiber for metropolitan area proved to be “inexpensive to build”, “easy to maintain”, and “mature technology”



Domestic R&E Backbone --TWAREN Topology





Research Oriented Grid

Applications

- Scientific Computing over Grid (AS)
- High Energy Physics: ATLAS, CMS, LCG, CDF, Belle (NTU, NCU, AS)
- Bioinformatics: CRASA, mpiBLAST-g2 (AS, YMU)
- Digital Archives: DataGrid,
- Atmospheric Science (NTU, NCU, TNU, AS)
- Biodiversity Informatics (AS, NTU, NMNS, etc), GBIF
- HealthGrid (TMU, AS)
- Geospatial Information Infrastructure (AS)
- Earthquake Data Center (AS)
- Earth Science (NCU, AS)
- EcoGrid and Sensor Grid (NCHC)



 **CCLRC** **Grid Operations Centre at RAL**

User Support Centre at FZK 

A second operations & support centre in Taipei

 **ASCC**
 ACADEMIA Sinica Computing Centre  中央研究院計算中心



LCG/EGEE Deployment Milestones

- Started to work with CERN on LCG: **Sept 2002**
- LCG-0 deployed: **March 19, 2003**
- EDG testbed deployed: **March, 2003**
- ASGCCA approved: **June 12th, 2003**
- LCG-1 testbed Ready: **July 30, 2003**
- LCG-2 deployed: **Feb 2, 2004**
- GOC operational in Taiwan: **Apr 2, 2004**
- GGUS operation in Taiwan: **May 2, 2004**
- Mass Storage Service Installed: **July 15, 2004**
- Regional Operation Center Services: **Jan. 2005**
- EGEE Testbed Installed: **March 2005**
- Service Challenge: **May 2005**



Contributions of ASCC in LCG/EGEE

- Collaborate ATLAS & CMS Teams in Taiwan
 - Acting as Tier-1 and Tier-2 Site
- LCG/EGEE
 - Production CA Services
 - LCG Core Site
 - 2nd Grid Operation Center (GOC)
 - 2nd Global Grid User Support (GGUS) Center
 - Participate LCG Technology Development
 - Data Management
 - Grid Technology
 - Certification & Testing
 - Application Software
 - ARDA
 - 3D (Distributed Deployment of Database)
 - Operation and Management



EGEE Participation of ASCC

- Joined EGEE as a non-funded Partner since 2004
- NA3: Training and Induction
- NA4: Application
 - HealthGrid
 - Diligent
- SAI: Support and Operation Management
 - CIC/ROC
- gLite Testbed



Grid3/OSG Collaboration of Taiwan

- Becoming a formal Grid3 site in July 2004
 - For CMS: NTU
 - For Grid3 Regional Center and General Support: ASCC
- Grid3/OSG Testbed
- Grid3/OSG Deployment
- Operation and Management
- Interoperation



Education and Outreach

- To encourage more research groups to join Grid Computing research and application development
- Host Grid Technology and Application Workshop, and related tutorial in the Northern, Southern, Middle and Eastern Taiwan, at least once a year.
 - Around 400 attendants in the past 6 events from May 2004
- Will host Workshop and Tutorial for Asian Countries in cooperation with EGEE NA2

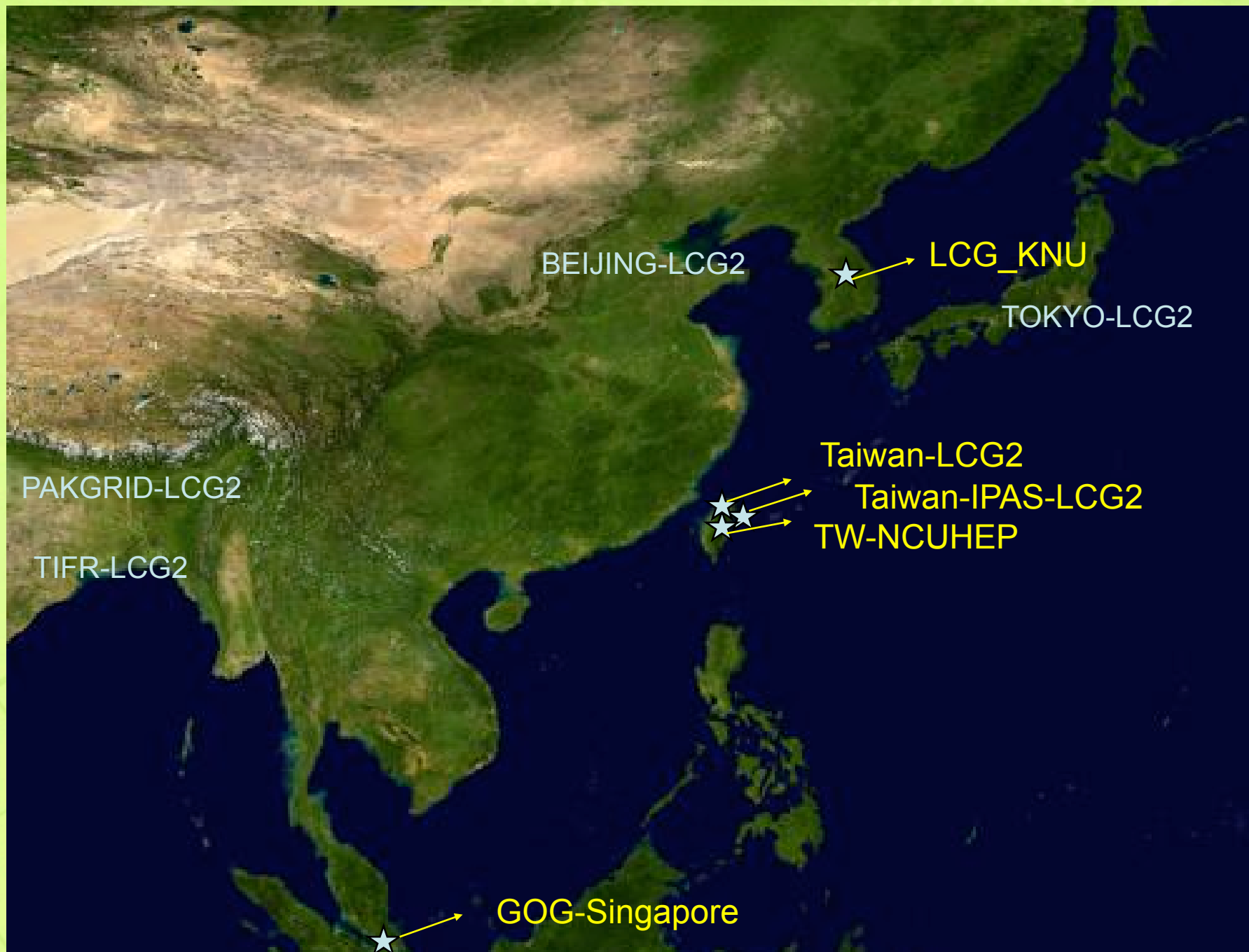


Help Fostering Grid Computing in Asia

- Coordinate LCG Community for better support and cooperation
- Education: Tutorial Program at least once a year
- Hosting LCG Asia Workshop and ISGC every year
- Extending Grid Community
- Facilitate the Formation of Application Driven Community
- System installation, management and consultant service



Asia Pacific Resource Centres



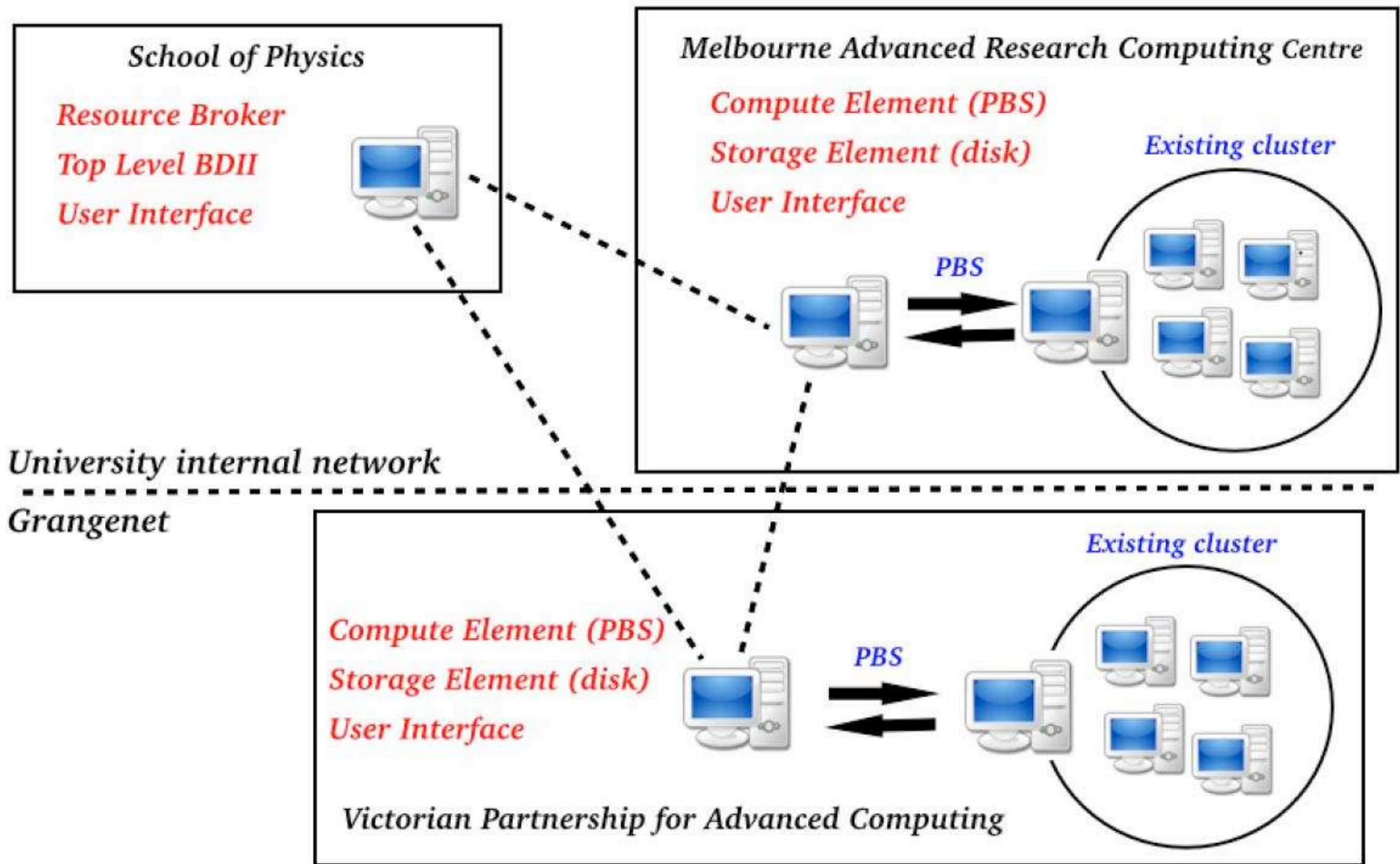


Asia Pacific Resource Centres

- TOKYO-LCG2: International Center for Elementary Particle Physics, University of Tokyo.
- BEIJING-LCG2: Institute of High Energy Physics, Chinese Academy of Sciences
- PAKGRID-LCG2: Pakistan Atomic Energy Commission
- TIFR-LCG2: TATA INSTITUTE OF FUNDAMENTAL RESEARCH
- GOG-Singapore: National Grid Office, HP and various institutes at Singapore.
- LCG_KNU: Department of Physics, Kyungpook University, Korea.
- Taiwan-LCG2: Academia Sinica Computing Centre (ASCC).
- Taiwan-IPAS-LCG2: Institute of Physics, Academia Sinica.
- TW-NCUHEP: High Energy Physics Group at National Central University, may combine with NTU to form a **federation!**
- **Taiwan-NTU-Phys in Grid 3**
- **Australia-Melbourne was in NorduGrid but is ready to join as a node of LCG!**

LCG deployment status in Australia

Current State / Architecture of LCG Deployment (April 19 2005)





Resources from RCs

| | Taiwan- LCG2 | Taiwan- IPAS- LCG2 | TW- NCUHEP | GOG- Singapore | LCG- KNU | PAKGRI D-LCG2 | TIFR- LCG2 | BEIJING- LCG2 | Tokyo- LCG2 |
|--------------|--|--------------------------|---------------|-------------------------|-------------|------------------|---------------|-------------------------|-----------------|
| # CPU | 138 | 22 | 32 | 90 | 6 | 2 | 26 | 40+ 30 | 84 |
| Disk (TB) | 30 | 0.3 | 1.3 | 0.08 | 0.05 | 0.05 | 0.05 | 3.00 | 0.87 |
| VO | Dteam, Alice, Atlas, CMS, BioMed | Dteam, Atlas | Dteam, CMS | Dteam, Atlas, CMS | Dteam | Dteam, CMS | Dteam, CMS | Dteam, CMS, Atlas | Dteam, Atlas |
| | | | | | | | | | |

* AU-ATLAS around 30 (+39) CPUs not included, another 32 CPUs for KNU-Grid3



Resources from RCs

- Asia Pacific in numbers:
 - 10 sites: Taiwan(3), Korea(1), Singapore(1), Japan(1), Pakistan(1), India(1), China(1), Australia(1)
 - Over 500 CPUs on LCG-Asia
 - ASCC will add 400 - 500 CPUs this summer
- Most supported VOs
 - Dteam
 - Atlas
 - CMS
 - BioMed



Current Services

- Asia Pacific Regional Operation Centre (APROC) was formed to provide LCG/EGEE operations support for sites in the Asia Pacific. Services provided by the APROC right now as shown below:
 - Deployment support
 - document installation/administration issues encountered
 - new site startup services:
 - registration
 - certification
 - coordination of region wide upgrades, changes and installation
 - Testbed Operations
 - certification testbed - to verify new m/w installation for region
- Dedicated VO for AP Services (in progress)



Current Supporting Level

- Because the policies are different for each site, the supporting level from APROC varies.
- Taiwan-LCG2: CA services, operation support, SFT, certification, deployment, registration, and user support.
- TW-IPAS-LCG2, TW-NCHHEP: CA services, SFT, and some other monitoring services.
- GOG-Singapore: CA service, coordination, registration, deployment, operation support, certification, and monitoring services (SFT)
- LCG-KNU: certification, registration, operation support, troubleshooting, SFT, and other monitoring services.



More Sites to Join

- There are some sites willing to join EGEE (not necessary LCG) in the near future, such as:
- NCKU: National Cheng Kung University, Taiwan
- NCCU: National Chung Cheng University, Taiwan.
- NCUCC: Computing Center, National Central University.
- We would like to provide extra services for the sites which deploy LCG by themselves within Asia Pacific.
- We also have interoperable issue in AP with Grid3 and NorduGrid.



What We Plan to Do ...

- Most urgent is some (technical) Coordination Structure within the region!
- Communication channels among technical people of the Tier2s in AP must be built
- We must meet regularly, better if more frequently
- May have joint workshop with ISGC or APAN (August 2005) or other relevant meetings
- Will run more training and tutorial sessions in the future
- Will help to coordinate Service Challenge for Tier2 centres in Asia Pacific



Key Factors for Alliance to Succeed

- Complimentary
- Goal alliance
- Risk sharing
- Long term
- Trust

Dick Cheney,
Keynote in Global Telecom Summit,
September 1997, Hong Kong

Conclusion

- Collaboration and Data are the two keywords
- We are about to witness Data Deluge in all disciplines of e-Sciences
- Unprecedented way to collaborate on day-to-day basis will change the sociology of academia life, eco-system of business world and eventually every one in the society
- It is about a new paradigm of collaboration, data will outgrow each of us, not really about competition between us!
- Together, we will achieve goals not possible individually!
- You are all welcome to attend 23-27 August APAN meeting in Taipei and 1st EGEE Workshop in Asia!