APSR Colloquium
“National Perspectives on Sustainable Repositories”
10 June 2005

APAC:
A National Research Infrastructure Program

John O’Callaghan
Executive Director
Australian Partnership for Advanced Computing
Australian Partnership for Advanced Computing

“providing advanced computing and grid infrastructure for e-Research”

Partners:
• Australian Centre for Advanced Computing and Communications (ac3) in NSW
• CSIRO
• Queensland Parallel Supercomputing Foundation (QPSF)
• Interactive Virtual Environments Centre (IVEC) in WA
• South Australian Partnership for Advanced Computing (SAPAC)
• The Australian National University (ANU)
• The University of Tasmania (TPAC)
• Victorian Partnership for Advanced Computing (VPAC)
National Role of APAC

- **Computing Infrastructure**
  - peak computing facilities

- **Information Infrastructure**
  - support community-based data collections
    - cf. institutional, reference collections

- **Grid Infrastructure**
  - seamless access to the computing and information infrastructure
  - new collaborative services for research groups
    - collaborative visualisation
    - computational steering
    - tele-presence
    - virtual organisation support
  - support for participation in international research programs
    - eg., astronomy, high-energy physics, earth systems, geosciences
• 10 Gbps
• IPv6
• Multicast

Australian Backbones

Darwin

Perth
IVEC
CSIRO

Canberra
ANU

Brisbane
QPSF

Melbourne
VPAC
CSIRO

Sydney
ac3

APAC
National Facility

GrangeNet Backbone

Centie/GrangeNet Link

AARNet Links

Internet2
Canarie
Geant
APAN
APAC National Facility

- **Usage**
  - mainly biology, chemistry, physics
  - currently 247 projects and 722 users (27 universities)

- **Computing Systems**
  - SGI Altix 3700 Bx2 system (mid-2005)
    - 1680 processors
  - HP AlphaServer SC ES45
    - 508 processors
  - Dell Linux cluster
    - 150 processors

- **Mass Data Storage System (MDSS)**
  - Hierarchical Storage Management
  - Storagetek (robotic silo) tape library
    - Capable of a petabyte ($10^{15}$ bytes) of storage

- **Visualisation Systems**
  - Virtual reality systems
  - Access Grid rooms

- **Staff (15.5 efts)**
  - User support
  - Systems support
  - Computational tools and techniques
  - Large-scale data collection management

http://nf.apac.edu.au
• Basic Services
  – *single ‘sign-on’ to the facilities*
  – *portals to the computing and data systems*
  – *access to software on the most appropriate system*
  – *resource discovery and monitoring*
APAC Grid Services

data centres
instruments
research teams
sensor networks
other grids: institutional, national, international
grid-based portals
distributed computation
federated data access
remote control
collaboratories
Context of APAC Grid

Astronomy Grids

HEP Grids

Bio-Grids

Earth Systems Grids

Geosciences Grids

APAC National Grid Infrastructure

Institutional Infrastructure

International Infrastructure
APAC Grid Projects

• APAC Grid Applications Support
  – Astronomy
  – High-Energy Physics
  – Bioinformatics
  – Computational Chemistry
  – Geosciences
  – Earth Systems

• APAC Grid Infrastructure
  – Computing Infrastructure
    • Globus-based middleware (VDT)
    • certificate authority
    • system monitoring and management (grid operation centre)
  – Information Infrastructure
    • resource broker (SRB)
    • metadata management support
    • resource discovery
  – User Interfaces and Visualisation Infrastructure
    • portals to application software
    • workflow engines
    • visualisation tools
APAC Initiative
Data Intensive Projects

- assist management of ‘community-based data collections’

- support for projects which are fundamentally data intensive
  - require little if any computational resources
  - data sets of ‘national significance’
  - allow easier distributed access to large-scale data sets
  - appropriate to be managed on the APAC National Facility

- support is in the form of:
  - a grant of storage capacity on the MDSS
  - advice on the access and management of the data
  - limited computational resource on systems in NF

- complements institutional repositories (APSR)

Data Intensive Projects

• Projects currently supported
  – Research Community projects
    • Paradisec, Reciprocal
  – APAC Grid projects
    • astronomy, earth systems, high-energy physics

• Resources
  – APAC National Facility: Mass Data Storage System
  – Specialist support
    • Stuart Hungerford (ANU)
  – APAC Grid Information Infrastructure team
    • Markus Buchhorn (ANU)

• Future
  – other data services: streaming media, images, animations
  – consistent access policies for information infrastructure
  – long-term support for collections and preservation environments
  – planning workshop on requirements and possible support
    • Oct 2005
APAC Grid Architecture

- PORTALS
- WORKFLOW
- APPLICATION
- APPLICATION
- GRID MIDDLEWARE
- GENERIC
- GRID MIDDLEWARE
- SECURITY
- COMPUTING, STORAGE, VISUALISATION, SYSTEMS
- NETWORKS