NCRIS Investment Plan

Platforms for Collaboration

Facilitator
Rhys Francis
Commonwealth Government spend of about $540M over the five years: 2006-2011

- Evolving bio-molecular platforms and informatics
- Integrated biological systems
- Characterisation
- Fabrication
- Biotechnology products
- Networked biosecurity framework
- Optical and radio astronomy
- Integrated marine capability
- Structure and evolution of the Australian continent
- Population health and clinical data linkage
- Terrestrial ecosystem research network
- Platforms for Collaboration (allocated $75M)
Timeline for PfC

- Planning commenced 2006
- Facilitator engaged part time May
- Facilitator full time by September
- Consultations to December
- Options and plan development completed by March
- Presentation to NCRIS April 13
- Decision:
  - Implementation to proceed
Reference Group

Tom Cochrane: DVC QUT,
James Dalziel: Prof Maquarie
Peter Nissen: Nat Brdbnd Adv for Edu
Linda O’Brien: VP (Information) UoM
John Sheridan: Prof Monash (ex MAS)
John Shipp: University Librarian USyd
Nick Tate, Director ITS, AUScert, UQ
Ah Chung Tsoi, eResearch Director, Monash
   now at University of Hong Kong
Nigel Ward: Exec Dir ADL
Starting Point – NCRIS Roadmap

- Data access and discovery, storage and management
- Grid enabled technologies and infrastructures
- Technical expertise
- High performance computing
- High capacity communication networks
Related Issues

- Sharing identities across the system
- Allowing researchers to specify who can do what with their resources
  - At a national and international level
- Expanding e-Research beyond the builder-users of cyber-infrastructure
  - Tools and service commoditisation
- Engaging the broader community
  - AREN and AAF reach and inclusiveness
PfC investment principles

• Build on co-operative arrangements

• Reduce barriers to adoption, support new as well as expert e-Researchers
  - Simple tools for simple things

• Provide robust and enduring services
  - From providers where that is the business

• Prefer services which are of value to multiple communities
## Main investment decisions

| AREN          | Harmonize AREN with regional and campus networks & policies  
<table>
<thead>
<tr>
<th></th>
<th>Foreshadow investment next period</th>
</tr>
</thead>
</table>
| AAF           | Assist the establishment of the AAF  
|               | Adopt for PfC services                                   |
| Compute       | Extend concept to include several application oriented sites as well as sustaining the (generic) peak facility |
| Inter-operation | Extend regional service providers to support 3rd party resources/data |
| Data          | Develop roles and missions in data  
|               | Create ANDS for some key services                        |
Three layers

Foundation services

Tools and Applications

Data

Interoperation

Compute
Two fundamental needs

- **System wide networking**
  - From desktop to national and global facilities
  - All desktops to all facilities
  - High speed
  - Subscription based (no volume charging)

- **System wide access control**
  - Home institution identification
  - Researcher controlled authorisation
  - Single sign on
  - Community wide roles and rights

Foundation services
Three further needs

• A federated data environment
  - National reach (location independence)
  - Common authentication/authorisation/search

• A national interoperation partnership
  - Compute and Data facilities
  - Large instruments
  - Sensor networks

• A distributed compute infrastructure
  - Relatively tightly integrated (SSO)
  - Able to include shared and dedicated facilities
The builders & the users

• Perhaps where e-Research happens now ...
  Builders = Users

• The challenge is to separate them
  - Services, Tools, Environments, Components, Software for the builders
  - Services and Tools for the users

• Expertise building
  - Training for the builders
  - User and consulting support for the users

Tools and Applications
Further thinking

Consulting and user support

Services (Tools) ((Software))

Middleware

Interoperability

Data Federation

Sensors

Compute infrastructure capability/capacity

Instruments

AAF (trust)

AREN (connectivity)
What is the problem?

Interoperation

Data
Grid
Compute

Content
Tools
Systems

Semantics
Adaptability
Virtualisation
Proposed arrangement

Researchers and research communities
Resource owners and operators
Users and User-Builders

Data Tools & Discipline Services

Interoperation & Collaboration Infrastructure

Compute Tools & Discipline Services

New tools, services & expertise

Foundation Services: AREN and AAF
Main PfC Activities

- **Data Management: ANDS**
  - Federation services, describing, finding, accessing
  - Stewardship for specific collections
  - Outreach for researchers and institutions

- **Interoperation: ICI**
  - Workflow, portals, collaboration services
  - Data management and data movement services
  - Job submission, VMs, AAA services

- **Computing: APAC**
  - Capability computing, National Facility
  - Shoulder and application oriented systems
  - Techniques and outreach for researchers
### Main PfC Activities

<table>
<thead>
<tr>
<th>Category</th>
<th>Code</th>
<th>Description</th>
<th>Code</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Management: ANDS</strong></td>
<td>21</td>
<td>Federation services</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stewardship</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outreach</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td><strong>Interoperation: ICI</strong></td>
<td>20</td>
<td>Workflow, portals, collaboration</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data and data movement</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Job submission, VMs, AAA</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td><strong>Computing: APAC</strong></td>
<td>26</td>
<td>Capability computing</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shoulder application systems</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Techniques, outreach, mgt</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>
ARIC

Australian e-Research Infrastructure Council

- An independent Chair
- An Executive Director
- A DEST Principal Adviser
- ~8 members drawn from providers and users of e-Research capabilities
- A small secretariat to assist programme management

Determine strategic directions and component variations
Convene an e-Research community forum for advice
Approve the annual business plan prepared by the Executive Director and supported activities
Make recommendations to DEST and NCRIS
Overall Arrangement

Computational Infrastructure
Interoperation & Collaboration Infrastructure
Data Management Infrastructure

ARIC
Executive Director
Secretariat funding

Host
Host
Host

AARNet
AAF
New entities/arrangements

Researchers and research communities
Resource owners and operators
Users and User-Builders

Data Tools & Discipline Services
Interoperation & Collaboration Infrastructure
Compute Tools & Discipline Services

Foundation Services: AREN and AAF

NCRIS contracts
Potentially a new entity
Researchers and research communities
Resource owners and operators
Users and User-Builders

New tools, services & expertise

Data Tools & Discipline Services
Interoperation & Collaboration Infrastructure
Compute Tools & Discipline Services

Foundation Services: AREN and AAF

More services in everyday operation
Three major additions

★ AAF - the Australian Access Federation and the deployment and use of its services across NCRIS

★ ANDS - which can assisting institutions develop and implement strategies for research data

★ COMPUTING - a tier of compute infrastructure focussing on capability needs rather than generic needs
Three major changes

★ ICI - repurposing the grid to researchers across the spectrum, using data and web services

★ RNSPs - new role for regional service providers to include linking major research resources

★ NEAT - an e-Research architecture development process to align research communities, middleware and operational services
Three contextual issues

• Harmonisation of campus, regional and national network and authorisation infrastructures
• Policies and environments that enhance e-Research and collaboration
• Support for the growth, enhancement and sharing of e-Research expertise; rewards, incentives and recognition
Computation
Capability Computing

- Peak system a strategic requirement
- However
  - Demands high levels of user expertise
  - Capability is for few users
- Decision points
  - Need $4M per annum to buy capability
  - Double that to operate, add outreach etc
  - So $9-10M pa - but can share costs
  - Some institutions will buy large shares
  - Expertise is rare, currently works well
  - Has international standing and reputation
  - ANU have agreed to continue to operate
Shoulder systems

• Shoulder/Application/Regional Systems
  - Support applications and tools for a broader community with less expert users
  - Locate near investments and expertise in applications and related research
  - Contenders: Bio, Nano, Image processing, Physics, Chemistry, Materials, Engineering

• Issues
  - Need (applications/tools/communities) mix
  - Need to lift strength in regional providers
  - Could see 8-10 but do 3 (readiness?)
Investment positioning

<table>
<thead>
<tr>
<th>Function</th>
<th>NCRIS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capability</td>
<td>$18M</td>
<td>$38M</td>
</tr>
<tr>
<td>Shoulders</td>
<td>$4M</td>
<td>$12M</td>
</tr>
<tr>
<td>Tools</td>
<td>$2M</td>
<td>$4M</td>
</tr>
<tr>
<td>Outreach</td>
<td>$2M</td>
<td>$2M</td>
</tr>
</tbody>
</table>

Ongoing refinement

- Peak versus capability in National Facility...
- Generic vs application focussed systems...
Inter-operation
Current Infrastructure

Portal Tools:
- GridSphere

Workflow Tools:
- Kepler

Security:
- APAC CA
- MyProxy
- VOMRS

Systems:
- Gateways
- Partners’ Facilities

Network:
- GrangeNet
- APAC VPN (AARNet)

A virtual system of computing, data storage and visualisation facilities
ICI activities

Operational aspect -

• Information, Portals & Collaboration Services
• Data management and Data Transport Services
• Job submission, Gateway VMs, AAA
• Help desk

Development aspect -

• Next generation tools and services
• NEAT determined
ICI – goal is greater reach

ICI Partners operate the core infrastructure
ICI Affiliates provide/access resources
Ongoing refinement

- New Partners – some self funding
- New Affiliates – more work
- Integration of new services
- Integration of new resources
- In-house expertise development
Data Management
Some of the consultations...

<table>
<thead>
<tr>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>APAC National Facility</td>
</tr>
<tr>
<td>ATLAS Experiment</td>
</tr>
<tr>
<td>Australian Bureau of Statistics</td>
</tr>
<tr>
<td>Australian Science and Technology Heritage Centre, University of Melbourne</td>
</tr>
<tr>
<td>BlueNet</td>
</tr>
<tr>
<td>Bureau of Meteorology</td>
</tr>
<tr>
<td>Centre for Cultural Research at the University of Western Sydney</td>
</tr>
<tr>
<td>Centre for Public Culture and Ideas, Griffith University</td>
</tr>
<tr>
<td>Council of Australian University Librarians</td>
</tr>
<tr>
<td>CSIRO</td>
</tr>
<tr>
<td>Geoscience Australia/Corporate Information Management and Access</td>
</tr>
<tr>
<td>Institute of Molecular Bioscience – The University of Queensland</td>
</tr>
<tr>
<td>Melbourne Health – Molecular Medicine Informatics Model</td>
</tr>
<tr>
<td>Molecular and Materials Structure Network</td>
</tr>
<tr>
<td>Monash University</td>
</tr>
<tr>
<td>National Library of Australia</td>
</tr>
<tr>
<td>Radio and Optical Astronomy Research Capability</td>
</tr>
<tr>
<td>Reserve Bank of Australia</td>
</tr>
<tr>
<td>Securities Industry Research Centre of Asia-Pacific</td>
</tr>
<tr>
<td>Sesame Scientific Data Management System, Bio21 Molecular Science and Biotechnology Institute</td>
</tr>
<tr>
<td>Tasmanian Partnership for Advanced Computing</td>
</tr>
<tr>
<td>The Australian Social Science Data Archive</td>
</tr>
<tr>
<td>University of Melbourne</td>
</tr>
<tr>
<td>Victorian Partnership for Advanced Computing</td>
</tr>
</tbody>
</table>
Data missions

<table>
<thead>
<tr>
<th>Community agreed</th>
<th>Services: Find, Mine, Access, Authorise</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRCs MNRFs Centres; Data may be used, shared or published</td>
<td></td>
</tr>
<tr>
<td>Collections; Data published, local repositories</td>
<td></td>
</tr>
<tr>
<td>Independent research; Data used and retained</td>
<td></td>
</tr>
<tr>
<td>F Services: stores, repositories, tools, expertise</td>
<td></td>
</tr>
</tbody>
</table>
Investment Positioning

Ongoing refinement
- Services require an entity to operate
- New non-expert data services are vital
- ANDS needs to be more like a co-operative than an enterprise

<table>
<thead>
<tr>
<th></th>
<th>NCRIS</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federated services</td>
<td>$2M</td>
<td>$0M</td>
<td>$2M</td>
</tr>
<tr>
<td>New services (NEAT)</td>
<td>$5M</td>
<td>$10M</td>
<td>$15M</td>
</tr>
<tr>
<td>Stewardship</td>
<td>$7M</td>
<td>$7M</td>
<td>$14M</td>
</tr>
<tr>
<td>Outreach</td>
<td>$7M</td>
<td>$7M</td>
<td>$14M</td>
</tr>
</tbody>
</table>
Change Vectors

- Harmonising national, regional and campus bandwidth, authorisation and charging policies rather than ‘more backbone’

- Provisioning for less expert users in computational systems through ‘application’ focused computational systems/services

- Development of regional service providers that do bring 3rd party resources into the interoperation infrastructure

- Understanding the roles and missions in data, develop federating and collections services
Thank you

pfc.org.au