SOLID EARTH AND ENVIRONMENT GRID

SUMMARY OF DAY ONE (29 July2003)

Session 1: SETTING THE SCENE

❖ **Robin Batterham**

- ➢Outlined the NRP's
- ➢Drivers for e - science and rates of change in a digital world
- ➢Described how R & D has changed and highlighted the move to complex inter disciplinary and team oriented research for 21st century;
- ➢The need for new networks of excellence = GRID to facilitate
- ➢Which will produce more opportunity;

KEY MESSAGE
Mode 2 science – networks of excellence ⇒GRID

❖ **Lesley Wyborn - the vision thing**

❖ Definitions:
- ➢*The GRID is an infrastructure that will make access to computing power, scientific data repositoriesand experimental facilities as easy as the current web makes access to information.*
- ➢*The GRID will open up storage and transaction power in the same way theweb opened access to content;*
- ➢**OR** we will go from browsing the web to programming the web
- ➢Data power program does not have to be local
- ➢Lesley then examined some our (solid earth)research methodologies and challenged us to think about in another way and to exploit this immense opportunity that we have:
- ➢She identified he issue that need to taken on and beaten:
  - Language global consistent
  - data structure and standards
  - data pedigree essential
  - not positioned to exploit the capacity.
  - preservation
  - Mathematically based sciences have a language that facilitates ready adoption of the GRID – where ARE IN THE GEOSCIENCE?
- ➢GRID is part of the science infrastructure to undertake and deliver the NRP's.

KEY MESSAGE:
GRID is unstoppable – use it
A number of challenges to get onboard
Australia to power the SEE GRID to take the lead

- **Prof Sue O'Reilly**

- Asked the question:
  - Can the SeeGRID provide a focus for the next quantum leap in
    - Establishing **connectivity** in the Earth Science Community?
    - Understanding of the Australian Continent (new knowledge)?
  - SeeGRID – is it potentially
    - a telescope to focus our efforts;
    - see the same **datascape**
    - BUT add creativity to knowledge base.
  - Reviewed the global context for geoscience and then the relevance to Australia
    - Emphasises the importance of the earth support system to our society and its wellbeing.
  - Need to arrive at what Sue described as
    - Integrated Systemic solutions with sustainable outcomes – Science for 21 st Century in which geology has a key integrating role to play.
  - Australian geoscienceis positioned to capitalise onan excellent base;

KEY MESSAGE:
The science we should tackle needs SEE GRID (whole earth 3D-4D 5D).
Codification of geoscience to facilitate exchange and collaboration.

Session 2: DRIVERS

- **Prof Lawrence Cram – e-Research in Australia**

- e-research in Oz overview and the ARC'srole and aspirations
- outstanding research is often based on international collaboration in a setting supported by first rate:
  - communications technologies
  - computing capability
  - information and database resources
- nation's IT capability has acted like adrenaline to S&T US National Science Board
- diagram of cyberinfrastructurethat captured
  - hardware, software, services
  - personnel – later highlighted the under investment relative to hardware and software
  - organisations
- outlined features of e research
- saw the GRID as having promise in
  - providing a valuable management structure
  - new ICT resources that might foster creativity (virtual labs)
  - will have the anarchy; but will require standards; continuity and consistency and sustainability of the infrastructure crucial
  - Foreshadows several later talks

KEY MESSAGE: E research is an international agenda – Oz need to get on board and tackle the challenges.

- **Paul Kelly: Executive Director of the ANZLIC**

- Spatial data infrastructure for Australia
- Pointed out the problems with accessing and using spatial data
- Explained the role and vision of ANZLIC and what the Australian Spatial data Infrastructure (ASDI) – data products and services available to all users
- Coordination of 10 partners and that they sort to connect the drivers with enablers by occupying those important empowering layers.
- Provided examples
  - maximising the benefits
  - facilitate industry growth
  - rising community expectations for online services
  - globalisation and technology
  - changing societal priorities – environmental issues
  - security, emergency management
- draw a comparison between the GRID and ASDI
  - merging technologies
  - principles based
  - dependent on collaboration and partnerships
  - striving for global and local outcomes
- ASDI provides
  - Access to data
  - Access to investment
  - Services and products
  - Encourage increased use of SD
  - Deliver triple bottom line.

- **Jonathon Doig:CANRI**

- Community access to National Resource Information in a distributed environment
- Demonstrated→ can be done
- 4 yr NSW funded project
- Open platform for information sharing to facilitate decision making
- Off the shelf software
- Must be:
  - Adaptable – extensible – scalable
  - Outlined custodianship principles
  - Outlined technical principles
  - Separate data service from presentation
  - Demonstrated functionality
  - Demonstrated web services:Gif images
  - Open industry standard
  - Web feature services:in GML
  - Webmap composer to create your own application
- Emerging Issues:
  - SDI governance – needs lead agency and a champion

- Quality of service
- Common conceptual model
- Information classes required
- Online registry – standards still emerging
- Web access management:authentication issues

- GRID enabled:limiting factors
  - Sparcity of data requirements (definition)
  - Willingness to collaborate

Session 3 ENABLERS

- **George McLaughlin: AARNET**

- Global cyber infrastructure (infrastructure to support e science)
- Outlined a vision of
  - connecting people to resources
  - engaging unis, gov and industry – affordable network infrastructure, advanced communications and grid services
  - developing new ways of delivering capacity
- Gave us an Ian Foster definition for the GRID
  - "Resource sharing and coordinated problem solving in dynamic multi-institutional virtual organisations"
    - on demand,ubiquitous access to computing data and services
  - outlined the e science world we are moving into
- introduced us to
  - Grangenet,
  - GEON – NSF funded geosciences network in the US
  - Geos and IT people together
  - Reminded us that OZ is developinga$27 m faciity in ACCESS MNRF to provide an Earth Systems Simulator
- Explained and demonstrated the Australian and global networking that exist

- **Prof John O'Callaghan** – Australian Partnership for Advanced Computing

(APAC)

- Reminded us that the way in which science is done is changing;
- Placed the grid centrally in the 3 overlapping balloons of information, computing and communications as the key enabler
- And as the infrastructure layer between the user community and web services
- He gave us the history APAC and its achievements
  - Partnership 8 organisation
  - Had a fair sort of grunt – 63 in top 500
  - Some modest storage – a peta byte online

- And some very nice visualisation capacity including virtual reality
  - Staff at ANU
- Grangenet; ACIGA Data Grid
- Virtual Observatories – eg MACHO
- Supporting Bio informatics
- APAC Initiatives
  - 1) Provide more support for 'data-intensive' computing
    - Bring your large scale datasets to play
  - 2) install and operate an APAC GRID
  - 3) support national and international e-science initiatives.

- **Prof Bill Appelbe:VPAC**

- Building a geoscience repository of downloadable geoscience software – the ACCESS MNRF
  - Software is as poor a shape as the data
  - Underware – not work layer
- Explained VPAC – independent company
- Opportunistic focus:
  - Geoscience
  - Bioinformation and life science
  - Computational engineering
- Collaboration essential
- Geoscience software
  - Hero codes – PhD student (poorly documented)
  - No 'community codes'
  - Can we identify the best geoscience software
  - Open sources
- Repository:exists in beta form
- Want contributions
- Software frameworks:model adaptability
- Geoscience is a real software challenge! in several areas
  - Software
  - Dataset management and curation
  - Limited funding
- Where is the big picture project?Needs collaboration of multi-skilled teams on focussed international scope projects

Session 4 ARCHITECTURE

- **Rob Woodcock: CSIRO and the pmd*CRC**
- Started with a practical scenario
  - Inversion in Earth processes
  - Pointed to the challenge of sparse data sets – an incredibly under sampled environment
  - Highlighted the need to be able to be able to run many computationally intensive scenarios because the parameter space is massive

    - Numerical misfits analysis isnot possible so we need replace it by human appraisal of the misfit which lead to late to the importance of **people ware** in this business echoing some early sentiments
    - Accessible
    - Rob examined a number of use cases
        - Business to reporting agency
        - Lab to business
        - Information services
        - Computational services
        - Value added reseller
    - What was happening in pmd CRC
        - User case need agreements
        - SEE perspective needed the Community of Practice(CoP)
    - To use his own term he handballed off some of the issues by referring to some of tomorrow presentations;
- He pointed us to the challenge of moving into the use of registries and passed the buck to Rob Atkinson
- CoP – people ware
    - Warned us of dangers of **I** and told us we will only succeed if we all succeed

KEY MESSAGE:will only work if the Solid Earth and Environment community come together and get there act together.
Information and computational services can operate on the GRID

- **Simon Cox WFS (Web Feature Services)**

- Feature model
- Pre-conditions for complex interoperability distribution system
    - Language
    - Led us into it through the history of the web
    - Web services:GRID – data that is reusable, ie by systems other than human
    - Computational services – willing to wait – if it works ⇒ service chain
- Messages need to be standardised
- Standards for geoscience data
- Working with internet consortium
- Sensor collection service
- Senor planning service
- Implementing a community language; 2 aspects defined:
    - principal elements
    - vocab of valid property values
- XMML feature catalogue – work to date
- Common encoding – interpoperability with the community
    - Agree on the specific language used on skinny red arrows!

- **Rob Atkinson** – Social Change Online

Data Models for Interoperability in Environmental research.

- The problem 'how do we derive useful information from sparse samples and many possible models of behaviour'
- Different agencies collect and manage the data
- How do we deal with:

- ➢Many possible models
- ➢Reruns
- ➢Models may take time
- ➢Archive each results for audit trail reason
- ➢Huge or trivial amounts of data

- ❖ Multiple motivations
- ❖ Web services concepts
- ❖ Publish-find-bind concept and the issues around the data models
- ❖ Introduced us to Features to objects
- ❖ role of registries
  - ➢central purpose is to tor services and match then with service requests
  - ➢Registries were easy to build and hard to populate

KEY MESSAGE:HAVE AGO – build something and learn!
Who is going to do it so we can learn?

❖ **Frank Brassil**

Reminded us that the great repositories of data in this country – the Gsurveys
Are into preparing for the day

KEY MESSAGE: Open systems approach essential.

❖ SUMMARY

- ❖ GRID exists:Resource sharing and coordinated problem solving in dynamic multi-institutional virtual organisations
- ❖ Geo's on the cusp of having one of the most powerful enabling tools the discipline has ever had at its disposal – opportunity to build a SEE GRID
- ❖ Fix the geoscience software – has to be through international collaboration
- ❖ Hardware software not the issue – people ware is
- ❖ Co P:must get our act together
- ❖ Which is not to underestimate the challenges of standards and models and registries etc.
- ❖ Must have the agreed language
- ❖ Studies of unprecedented scale and complexity for geoscience
- ❖ Lack of codification of aspects of geosciences is one of the great challenges to overcome if we are going to work at the scale we aspire to the complexity facilitated by the GRID
- ❖ How much work to be done before it will be easy and simple to me
- ❖ **Collaboration essential**