Research Data Storage Infrastructure (RDSI) Project

Transforming Australia's Research Data Storage

Dr Nick Tate
Director, RDSI

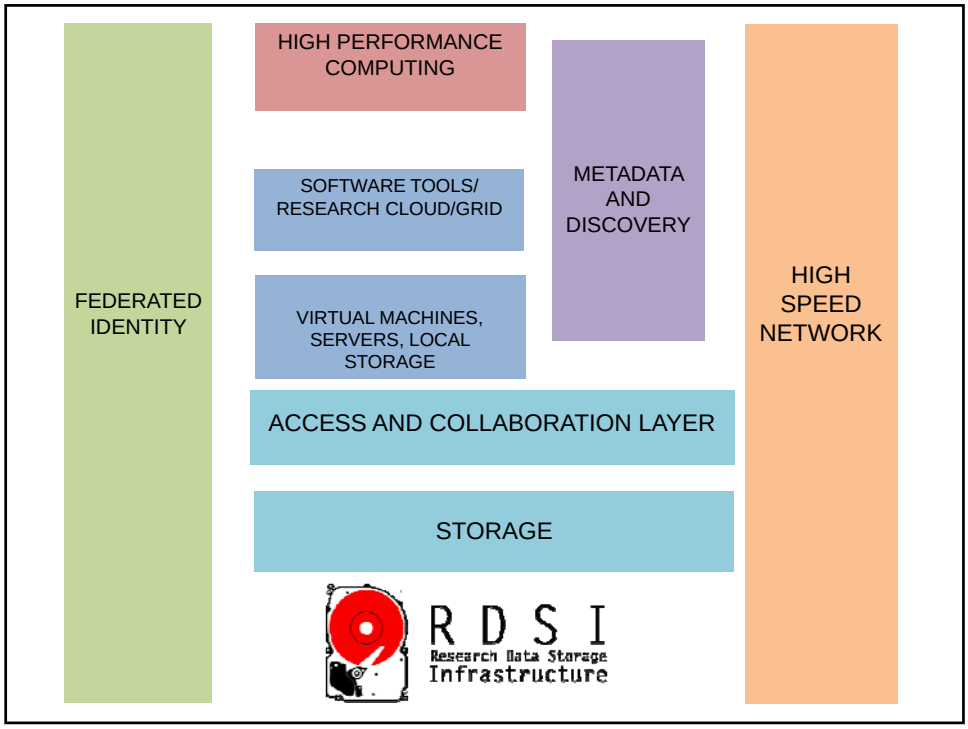
13th July 2011

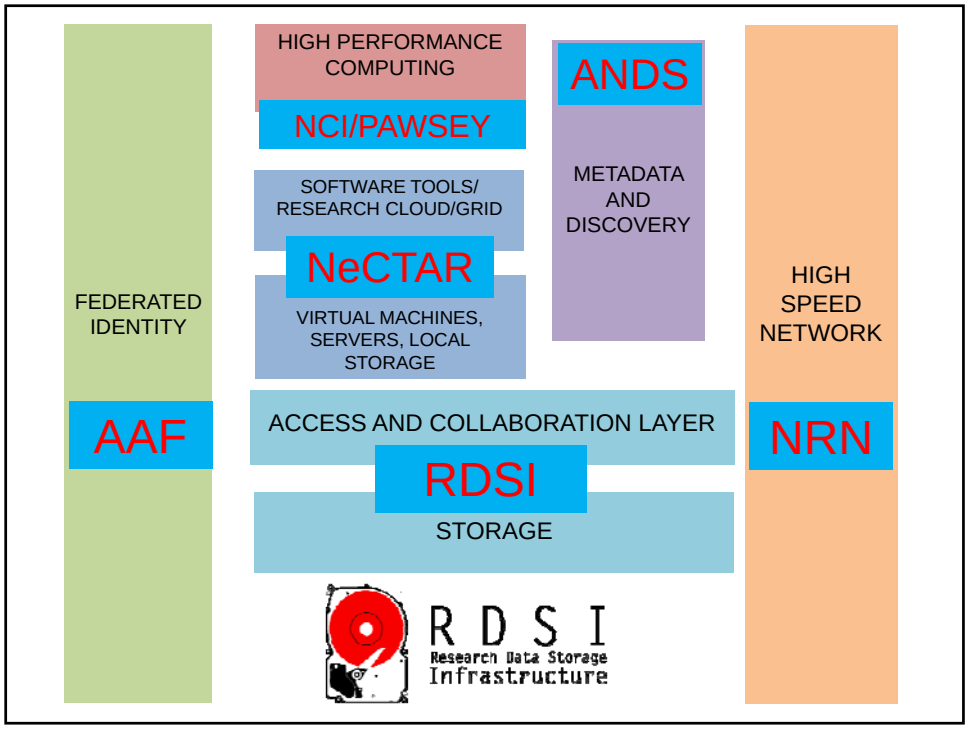
R D S I
Research Data Storage
Infrastructure

Research Data Storage Infrastructure (RDSI) Project

The goal is that researchers will be able to use and manipulate significant collections of data that were previously either unavailable or difficult to access and that there will be a consistent means of accessing this data

R D S I
Research Data Storage
Infrastructure


FEDERATED IDENTITY
HIGH PERFORMANCE COMPUTING
SOFTWARE TOOLS/ RESEARCH CLOUD/GRID
VIRTUAL MACHINES, SERVERS, LOCAL STORAGE
METADATA AND DISCOVERY
HIGH SPEED NETWORK
ACCESS AND COLLABORATION LAYER
STORAGE
R D S I
Research Data Storage Infrastructure


FEDERATED IDENTITY
AAF
HIGH PERFORMANCE COMPUTING
NCI/PAWSEY
SOFTWARE TOOLS/ RESEARCH CLOUD/GRID
NeCTAR
VIRTUAL MACHINES, SERVERS, LOCAL STORAGE
ANDS
METADATA AND DISCOVERY
HIGH SPEED NETWORK
NRN
ACCESS AND COLLABORATION LAYER
RDSI
STORAGE
R D S I
Research Data Storage Infrastructure

Funding and Structure

- The University of Queensland (UQ) has been appointed lead agent for this project
- DIISR is providing $50m of EIF funding for the project to UQ
- This type of funding is for infrastructure and may not be used for operations
- Co-investment is therefore required

R D S I
Research Data Storage Infrastructure

RDSI Project Board

Independent Chair – Brian Anker

Board Members
- Professor Max Lu (UQ)
- Professor Jill Trewhella (Sydney)
- Professor Doug McEachern
- Peter Nikoletatos (Curtin)
- John Shipp (Sydney)

Further members to be appointed

Project Director attends board meetings

R D S I
Research Data Storage Infrastructure

RDSI Board Committees

Infrastructure Advisory Panel

Chair: Peter Nikoletatos
Members: Mary Sharp
Rick Van Haeften

Other Members to be appointed

Resource Allocation Panel

Chair: John Shipp
Members: Professor Doug McEachern

Other Members to be appointed

RDSI
Research Data Storage Infrastructure

RDSI Project Team

Project Director - Dr Nick Tate (UQ)
Office Manager - Ms Alice Stokes (UQ)
Project Manager - Ms Viviani Paz (UQ)

Programme Team - Dr Rodney McDuff (UQ)
Dr Rob Bell (CSIRO)
Dr Markus Buchhorn

Legal Team (0.5FTE) Ms Elise Neilson (UQ)
Ms Cristy Sun (UQ)

RDSI
Research Data Storage Infrastructure

RDSI Vision

- Significant Research data collections on a small number of scalable nodes
- Highly connected
- Common interfaces (Architecture)
- Connecting to data held by others
- Vendor Panel

R D S I
Research Data Storage Infrastructure

RDSI Programmes

- Node Development (NoDe)
- Research Data Services (ReDS)
- Data Sharing (DaSh)
- Vendor Panel (VePa)

R D S I
Research Data Storage Infrastructure

Sourcehttp://en.wikipedia.org/wiki/File:Thomas_J_Watson_Sr.jpg

RDSI Programmes

- Node Development (NoDe) $10,628,480
- Research Data Services (ReDS) $28,389,978
- Data Sharing (DaSh) $7,695,000
- Vendor Panel (VePa)

RDSI
Research Data Storage Infrastructure

RDSI Programmes

- NoDe and ReDS programmes by open call
- Multiple rounds in NoDe and ReDS
- "Call for proposals" to include minimum requirements (from contract)
- Strawman/Tinman consultation used to refine additional requirements
- All nodes must sub-contact to UQ

RDSI
Research Data Storage Infrastructure

Source: http://news.cnet.com/2300-10805_3-10001679.html?tag=mncol

NoDe Programme

Primary Nodes

- Hold research data sets of lasting value and importance which support research on a multi-institutional basis
- be connected to the Australian Research and Education Network (AREN) optical backbone or be in the process of being connected to it.

Up to 6 Primary Nodes across Australia

R D S I
Research Data Storage Infrastructure

NoDe Programme

Additional Nodes

- hold research data sets of value and importance which are required to be held in the location of the proposed Additional Node. This could be because of jurisdictional restrictions or because of conditions of investment;
- be connected to the AREN or be in the process of being connected to it.

Up to 4 Additional Nodes across Australia

R D S I
Research Data Storage Infrastructure

NoDe Programme

All Node Operators

- Have access to necessary infrastructure and staff
- Be a substantive organisation
- Demonstrated commitment to data retention and sharing
- Support agreed service levels
- Implement DaSh Technical Architecture
- Participate in DaSh Network
- Purchase through the Vendor Panel (for RDSI funds)
- Implement RDSI Service Metrics Reporting

RDSI
Research Data Storage Infrastructure

NoDe Programme

Primary Node Operators will also

- have a governance structure for node-related policies
- contribute towards RDSI project operational costs

It is expected that 90% of NoDe/ReDS funding will go to Primary Nodes

RDSI
Research Data Storage Infrastructure

NoDe Programme

The Node Proposition

- Nodes receive funding under the ReDS programme
- Nodes received funding from the NoDe programme
- Nodes are able to participate in DaSh high speed network
- Nodes are positioned well for future investment

- Nodes must co-invest by funding operational costs
- Nodes must make contribution to some RDSI operations costs

R D S I
Research Data Storage Infrastructure

RDSI Programmes

Research Data Services (ReDS)

- The ReDS programme will identify research data holdings of lasting value and importance and contribute funding to their development at the most appropriate Node(s)

R D S I
Research Data Storage Infrastructure

ReDS Programme

- ReDS funding goes to nodes
- Merit Allocation process to be developed
- Resource Allocation panel will review proposals
- Ongoing discussion about “who” applies
- ReDS Strawman published
- ReDS Strawman workshop in June

R D S I
Research Data Storage Infrastructure

RDSI Programmes

Data Sharing (DaSh)

- The DaSh Programme will build capability to support the sharing and re-use of research data and, as a result, is aimed at expanding the scale and scope of problems that Australian researchers may seek to address

R D S I
Research Data Storage Infrastructure

DaSh Programme

DaSh Collaboration Network

- Interconnection between primary nodes at 40/100 Gbps
- Additional nodes at 1/10 Gbps
- NRN project

DaSh Technical Architecture

- Under discussion
- Federated vs. distributed
- Strawman released today
- Strawman workshop held on 17th June

R D S I
Research Data Storage Infrastructure

DaSh Programme

Extending the reach of RDSI

- Affiliated organisations
- Public storage gateway

Service Management

- Service through transparency
- Nodes self-reporting with audit by RDSI

R D S I
Research Data Storage Infrastructure

Source: http://www.urbanspoon.com/u/photo_list/1544244?photo_id=345322

Source: http://www.urbanspoon.com/u/photo_list/1544244?photo_id=423246

RDSI Programmes

Vendor Panel (VePa)

- The Project Director will work with relevant parties, including the Council of Australian University Directors of Information Technology (CAUDIT) and the RDSI Infrastructure Advisory Panel to negotiate sector wide storage agreements and establish a panel of vendors, who will provide cost effective storage infrastructure for the sector as a whole

RDSI
Research Data Storage Infrastructure

Activities to Date

- NoDe Strawman Developed
- NoDe Strawman Workshop
- ReDS Strawman Published
- ReDS Strawman workshop
- DaSh Strawman published
- DaSh Strawman Workshop
- NoDe Tinman Workshop/publication
- Website established
- Initial team in place
- Project Plan developed and approved

RDSI
Research Data Storage Infrastructure

Workshops/Consultations to date

Date	Workshop/Consultation	Where
Thursday, 14 April 2011	NoDe Strawman Workshop	Brisbane
Friday, 3 June 2011	NoDe Tinman Workshop	Brisbane
Tuesday, 7 June 2011	NoDe Tinman Consultation	Perth
Thursday, 9 June 2011	NoDe Tinman Consultation	Sydney
Tuesday, 14 June 2011	NoDe Tinman Consultation	Brisbane
Wednesday, 15 June 2011	NoDe Tinman Consultation	Melbourne
Thursday, 16 June 2011	***NoDe Tinman Consultation***	***Hobart***
Friday, 17 June 2011	DaSh Strawman Workshop	Brisbane
Monday, 20 June 2011	NoDe Tinman Consultation	Canberra
Thursday, 23 June 2011	NoDe Tinman Consultation	Darwin
Monday, 27 June 2011	ReDS Strawman Workshop	Brisbane
Thursday, 30 June 2011	NoDe Tinman Consultation	Adelaide
Friday, 1 July 2011	NoDe Tinman Consultation	Hobart

RDSI
Research Data Storage Infrastructure

Next Steps
Constructing the Vendor Panel

- Vendor Briefing being developed
- Likely to be held in July

RDSI
Research Data Storage Infrastructure

Next Steps
Call for Node Proposals

- Likely to be 18th to 22nd July
- Multiple rounds being envisaged

Next Steps
Assessing the Responses

- **Project team will undertake initial analysis**
- **Infrastructure Advisory Panel (IAP) reviews responses and develops recommendations to RDSI Board**
- **RDSI Board considers recommendations from IAP and recommends Nodes to DIISR**
- **If DIISR approves a Node recommendation, UQ contracts with Node operator**

Contact

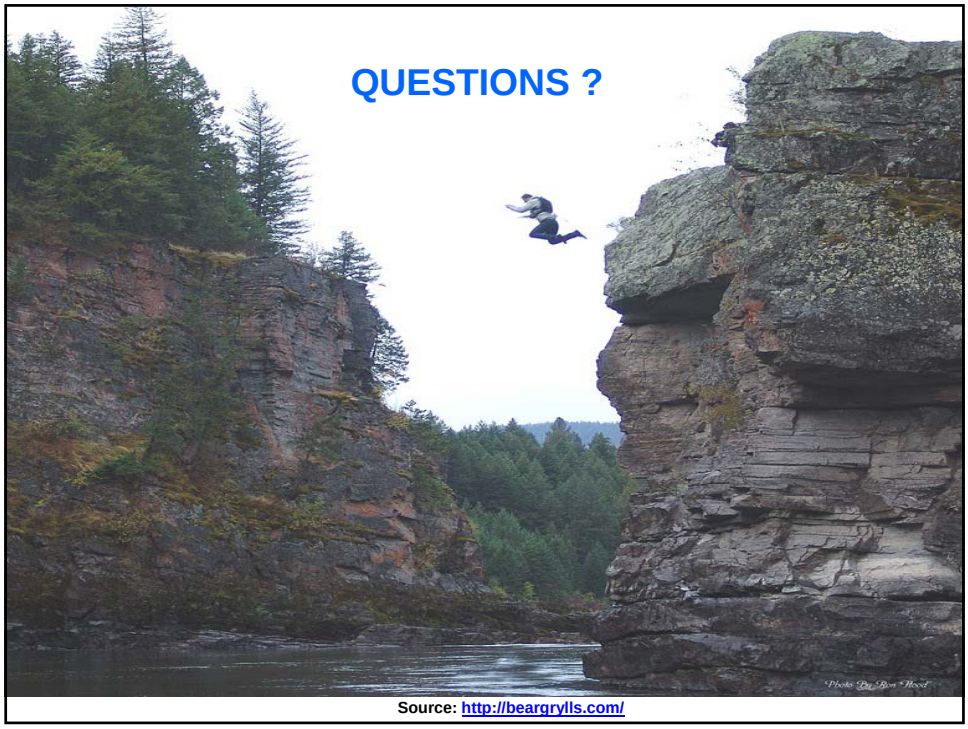
Project Director: Dr Nick Tate

Email: N.Tate@rdsi.uq.edu.au

Phone: (07) 3365 2019

Website: www.rdsi.uq.edu.au

RDSI
Research Data Storage Infrastructure



Source: http://beargrylls.com/