Delivering order-of-magnitude improvements to research data throughput

Internet2 Global Summit, Washington DC
27 April 2015
Typical AARNet Office
Supporting the AARNet Portfolio

It used to be about the plumbing, now it’s so much more…
Three Pivotal Activities

Innovation Incubator

Collaboration Enabler

Community Developer

It used to be about the plumbing, now it’s **so much more**...
Three Pivotal Activities

Collaboration
Enabler
The project: RDSI
eResearch infrastructure

www.rdsi.edu.au
RDSI: cloud storage for researchers

“Enabling researchers to use, reuse and manipulate significant data collections”

https://www.rdsi.edu.au/node-statuses
RDSI: cloud storage for researchers

“Enabling researchers to use, reuse and manipulate significant data collections”

https://www.rdsi.edu.au/node-statuses
The New World of Research

Research Drivers: “Really Big Data”

- Large Hadron Collider
- Square Kilometre Array
- Climate Change Data
- Genomics Data
Other research projects

Catlin Global Reef Record

Disease transmission among lizards

Global research on greenhouse gases

Historical newspaper record

https://www.rdsi.edu.au/use-cases
The problem: data sharing
Data sharing

Problems to solve

Data

Resources

Users
The problem:

A small amount of packet loss makes a huge difference in TCP performance.

Throughput vs. Increasing Latency with .0046% Packet Loss

- Brisbane-Melbourne (30ms)
- Perth-Sydney (50ms)
- Perth-LAX (160ms)
- NYC-London (75ms)

Image: ESNet
The solution: Science DMZ
The solution:

“Friction free networking”

1. **Network devices** selected for high performance applications

2. Dedicated systems for **data transfer**

3. Performance and network **measurement tools**

4. **Security mechanisms** designed for high performance
The solution:

“Friction free networking”
Science DMZ and RDSI
Science DMZ

Standard campus connection

AARNet

To other R&E destinations or the internet

University campus

Border router

Enterprise firewall

Site/Campus LAN

- Staff
- Students
- Instruments
- Compute/storage
- etc

Data-intensive resources

All flows pass through firewall, impacting performance

To other R&E destinations or the internet
Campus connection with Science DMZ

- To other R&E destinations or the internet
- University campus
  - Border router
  - Enterprise firewall
  - Site/Campus LAN
  - Site/campus access to Science DMZ resources
  - Per-service security policy control points
  - Data-intensive resources
- Science DMZ
  - Border Edge Switch/router
  - Data Transfer Node with high-speed storage
  - Cleaner, high-bandwidth path to/from WAN
- Data transfer node with high-speed storage
- Data-intensive resources
- To other R&E destinations or the internet
Science DMZ deployment

AARNet backbone network

APL NAP #1

APL NAP #2

Inter-Node VPN

APL

Node

RDSI

Node

Network infrastructure

Node Storage

DTN 1

DTN 2

BES-1

BES-2

PerfSonar 1

PerfSonar 2

PE1

APL-NTU1

PE2

APL-NTU2

APL

Node

Science DMZ
RDSI: high speed network connectivity

- **AARNet 4 Routed Access**
- **Inter-node VPN**

Primary Node
Secondary Node
Components

Data transfer nodes:
Dell Poweredge 720XD, 64GB RAM, 2 x SSD, 2 x 3.4GHz 8-core CPU
3 x Mellanox Connect3 VPI card (10/40 ethernet 10/40/56 IB)

PerfSonar:
2-blade Supermicro + Mellanox NICs
PerfSonar Toolkit; MadDash dashboard

Data transfer tools: Aspera, Globus/GridFTP

Border edge switch:
Juniper EX4550; Cisco 6800/NX6K/NX5K; Dell Force10
DTN: Aspera + Globus Online

Bandwidth

<table>
<thead>
<tr>
<th></th>
<th>INBOUND</th>
<th>OUTBOUND</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>2.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>3.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>4.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>5.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>6.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>7.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>8.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
<tr>
<td>9.0</td>
<td>0.0 MB/s</td>
<td>0.0 MB/s</td>
</tr>
</tbody>
</table>

Australian Academic and Research Network

Science DMZ

29
WARNING

No petrol (lead replacement or unleaded) available for 815 km between Yulara and Tjukayirla Roadhouse

Diesel and Avgas are available at some communities
Avgas is unsuitable for use in unleaded vehicles
Challenges

Latency

Solution: tightly manage the DTN:

NIC drivers, OS updates …

Careful choice of TCP congestion avoidance algorithm
Compatibility of storage architectures & vendors
Compatibility of storage architectures & vendors

Challenges

AARNet 4 Routed Access

Inter-node VPN

EMC²

NetApp

Hitachi

DataDirect Networks

Dell

IBM

HP

Teradata

Cisco

ASI Solutions

Xenon

Amazon Web Services

Datapod

Actifio

SGI
Compatibility of storage architectures & vendors

Solution: Science DMZ architecture
+ L3VPN for inter-Node traffic
User tools & familiarity

Command line tools

Typical Australian researcher
User tools & familiarity
Solution: keep it simple

-> Focus on one tool (eg. Aspera)

-> Many users just want a mapped drive

... although non-web authentication an issue

-> Partner with eResearch organisations for training and outreach
Moving forward
Future plans

Improved integration with other projects

nectar

RDSI Research Data Storage Infrastructure

cloudstor

AARNet’s FTP & HTTP Mirror
Future plans

Improved integration with other projects

![Cloudstor Logo]

**Total uploaded bytes**

- Week 14: 176 T
- Week 15: 178 T
- Week 16: 179 T
- Week 17: 180 T

Total uploaded bytes: 180308397628469 bytes

**Total users**

- Week 14: 18 k
- Week 15: 18 k
- Week 16: 18 k
- Week 17: 19 k

Total user count: 18637 users
Better integration with cloud providers
Add SDN for more flexibility…

Future plans

University campus

Border router
Enterprise firewall

Dedicated AARNet4 L2/L3 VPN service/s

Per-service security policy control points

Data-intensive resources

Science DMZ

Border Edge Switch/router

Data Transfer Node with high-speed storage

Site/Campus LAN

- Staff
- Students
- Instruments
- Compute/storage
- etc

To other R&E destinations or the internet

Private connectivity eg.
- RDSI Node,
- cloud provider,
- remote campus
- research partner

AARNet

PE

APL-NTU

AARNet

R
d
s
i
n

d
f
n
i
l
i
b
i
l

D
f
s
n
f
l
b
i
l

42
Future plans

Increased bandwidth

100Gbps Optical Network

KEY
- AARNet POP
- < 1 Gbps
- < 2.5 Gbps
- < 10 Gbps
- WDM Transmission
Future plans

Increased bandwidth

100Gbps Optical Network

KEY
- AARNet POP
- < 1 Gbps
- < 2.5 Gbps
- < 10 Gbps
- WDM Transmission

Australia Academic and Research Network (AARNet)