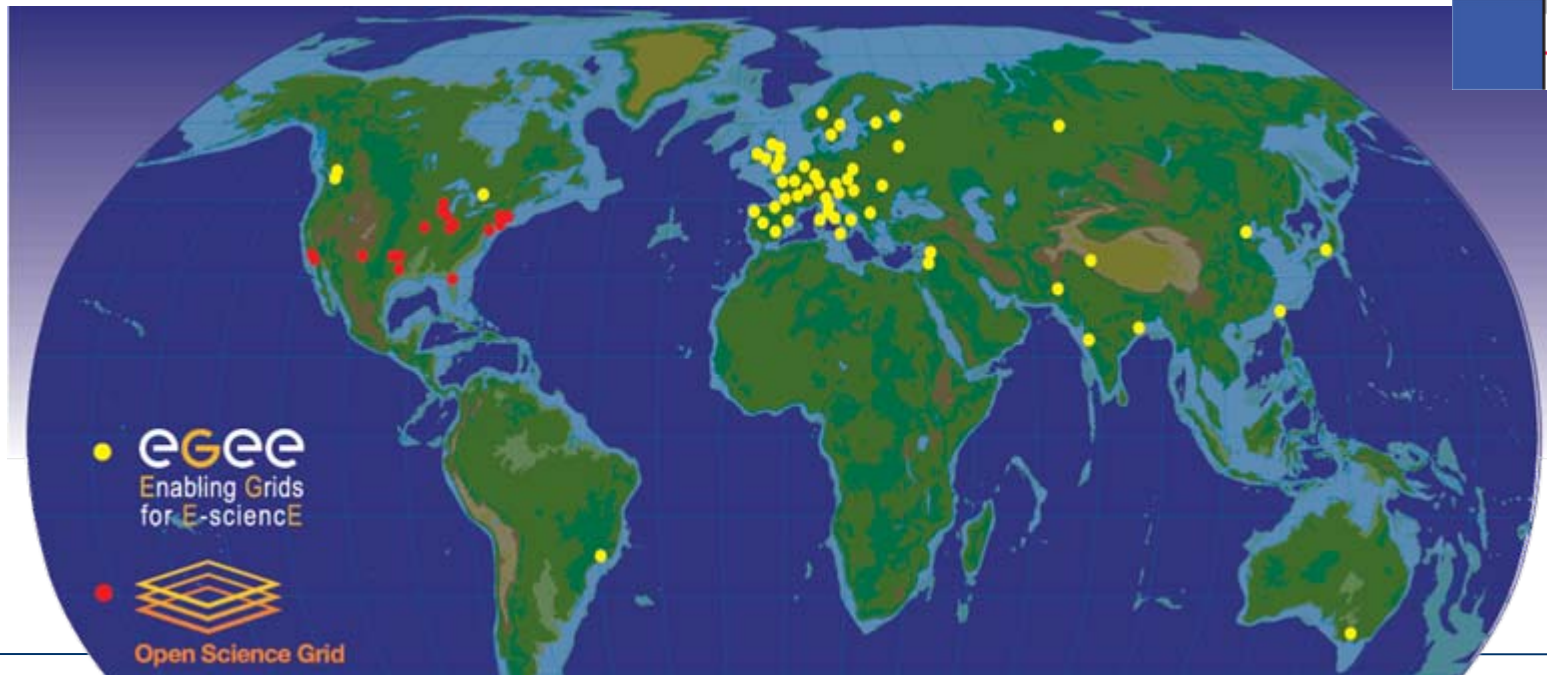




Australia-ATLAS: An LHC site in the grid outback

Tim Dyce – Australia ATLAS
EPP Group, The School of Physics
tjdyce@unimelb.edu.au

- The Worldwide LHC Computing Grid (WLCG) was established within the EGEE and OSG grids to process and store the LHC data
- Australia's only site within WLCG is located at Melbourne, Australia-ATLAS



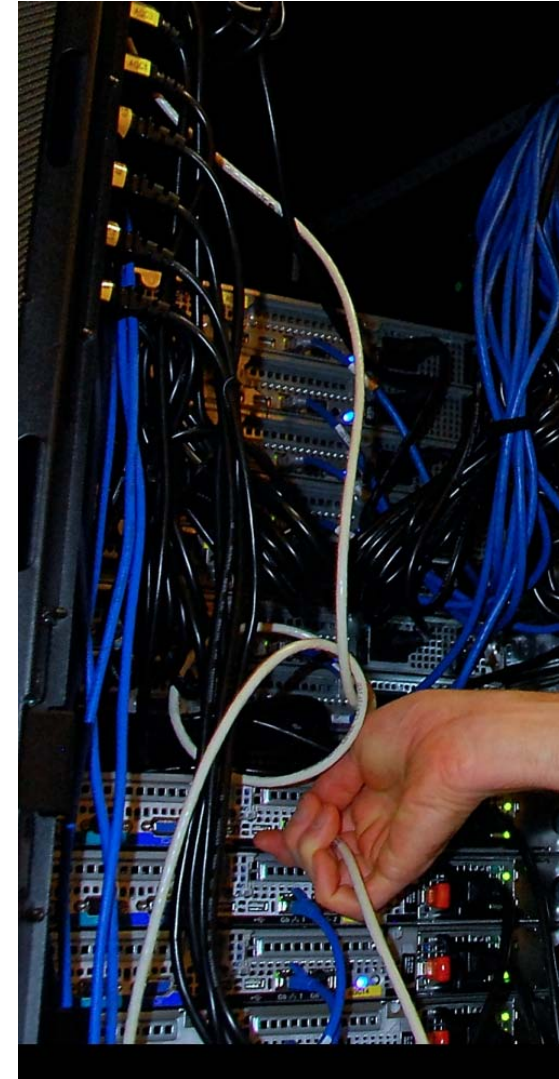
- The ATLAS computing model breaks the WLCG down into clouds and TIERS
 - Data flows from CERN (TIER0) to the TIER1s then the TIER2s
- Our Cloud and TIER1 is currently Taiwan
- But our data flows across the Pacific in order to attain the required bandwidth
- Data also flows to us from The U.S. and Canadian sites



- The LHC will produce ~3.2 Petabytes of data per year
- For Australia-ATLAS our subset of the data requires a minimum transfer speed of ~22 Megabytes a second (24x7)
- The data comes in waves
- The faster we get the data;
 - the quicker our users can access it
 - The less time is wasted

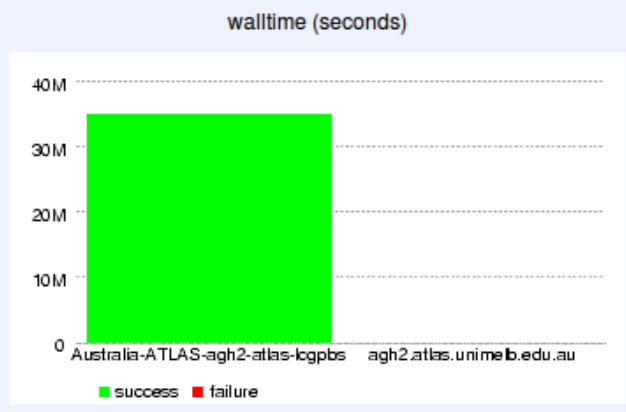
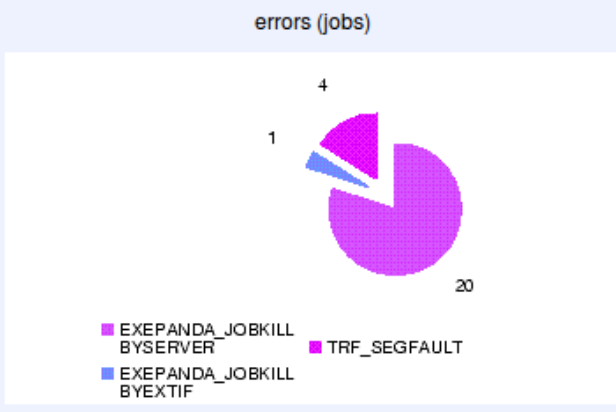
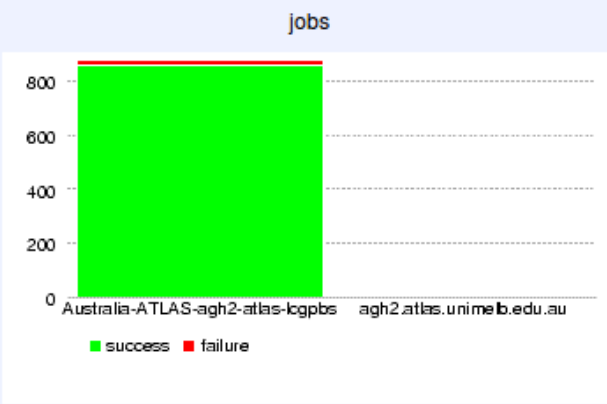
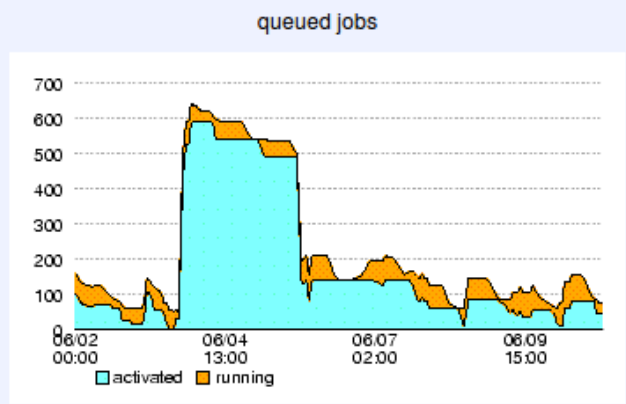
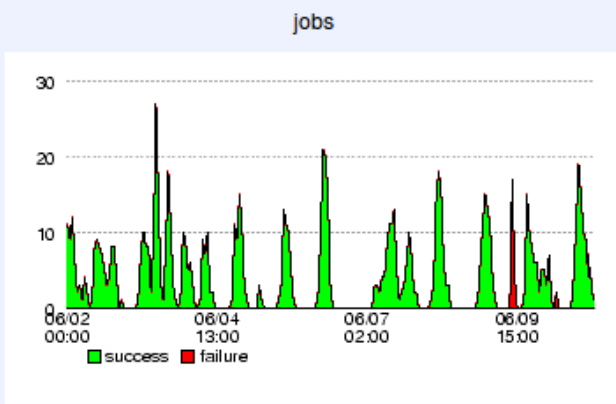


- Currently:
 - 250T of storage
 - 250 cores
- We just completed a round of purchasing
- The resources will approximately double each year



Resource Challenge – June 2009 (STEP09)

2009-06-02 00:00:00 — 2009-06-10 23:59:59

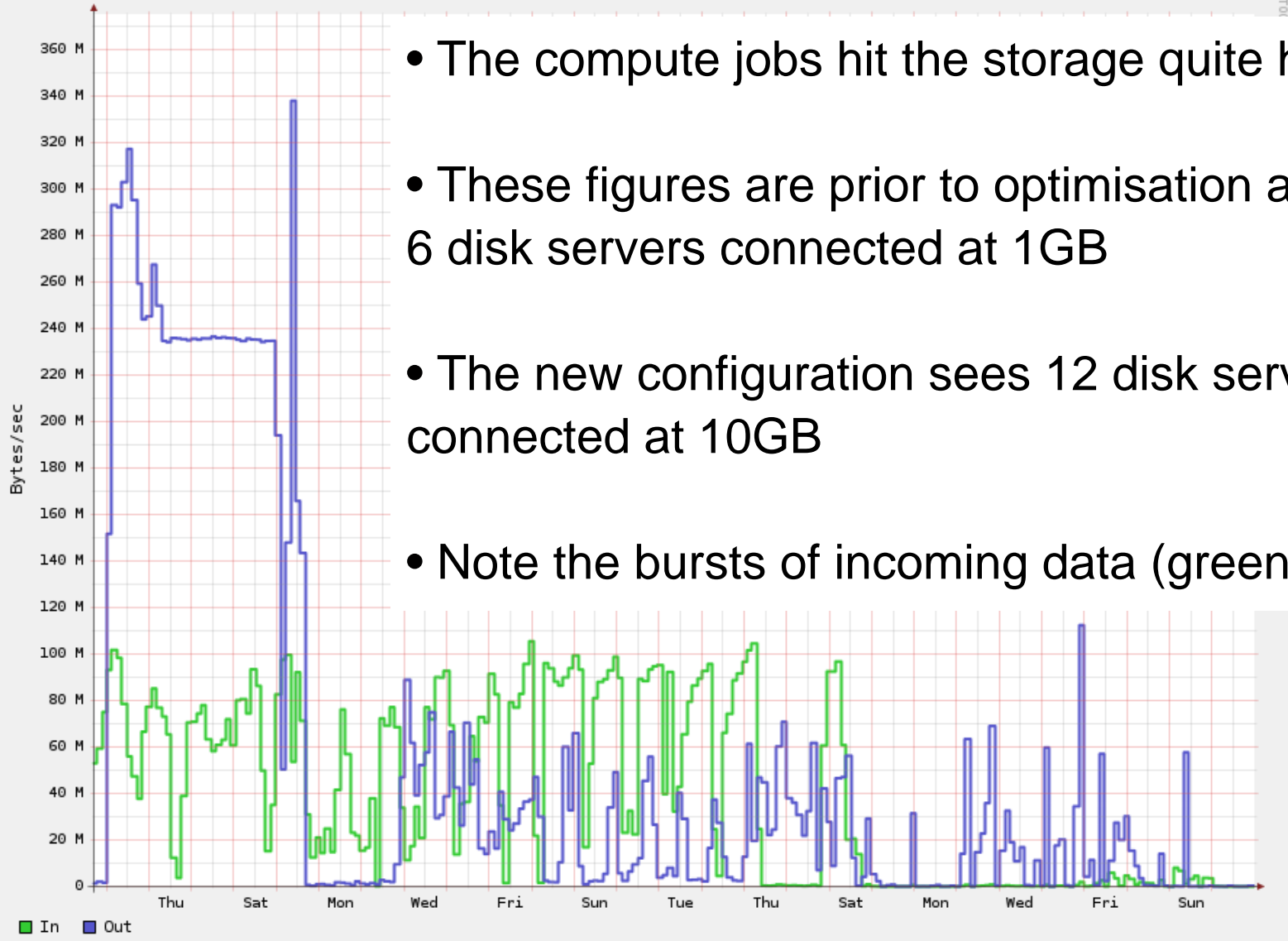


cluster	defined	assigned	waiting	activated	running	holding	transferring	success	failure	efficiency
✘ Australia-ATLAS-agh2-atlas-lcgpbs	0	0	0	0	0	8	0	855	24	97.3%
✘ agh2.atlas.unimelb.edu.au	0	0	0	0	0	0	0	0	1	0%
total	0	0	0	0	0	8	0	855	25	97.2%

CRITICAL
WARNING
NORMAL
GOOD
NO_ACTIVITY



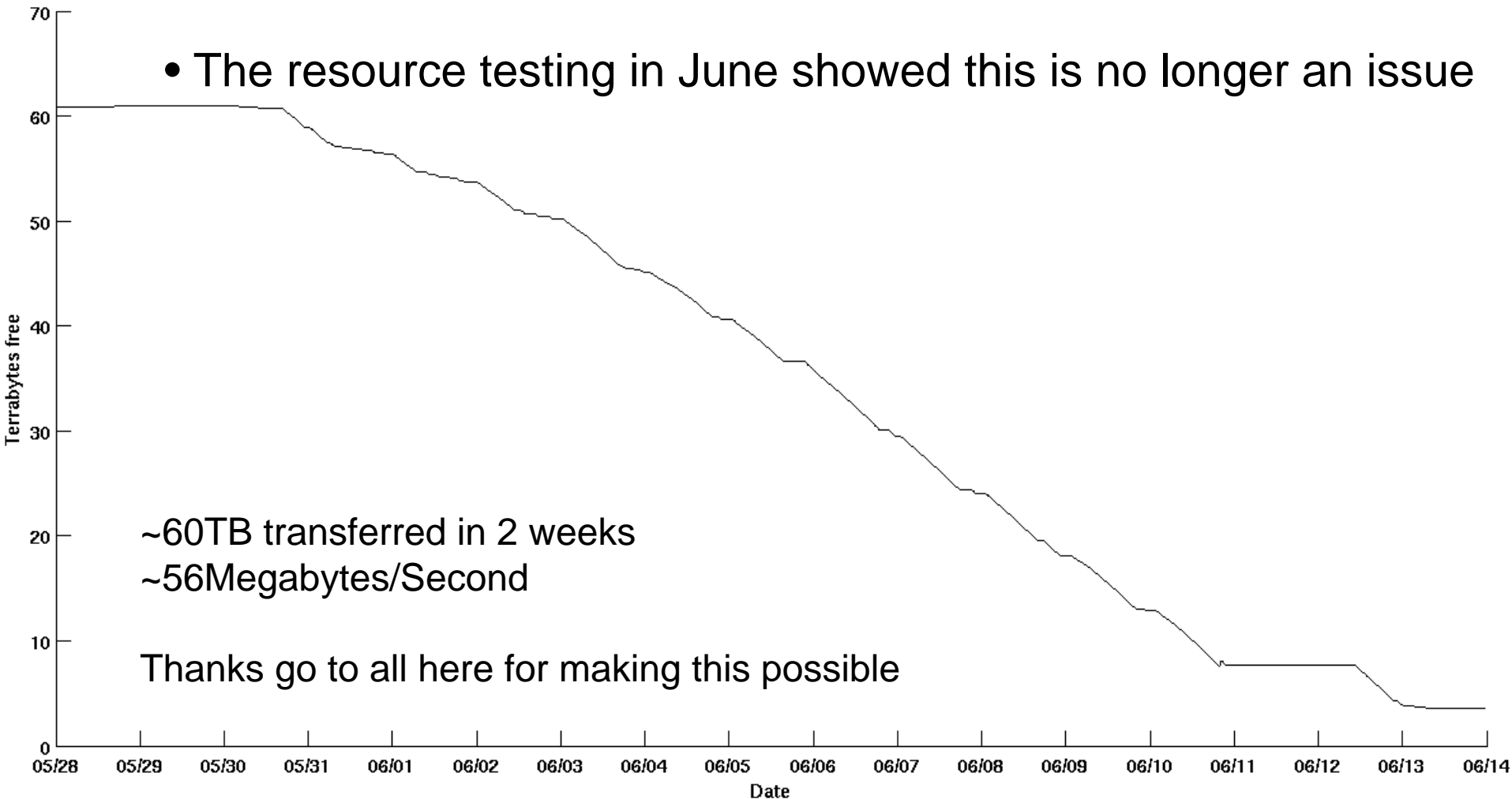
Disk Servers Cluster Network last month



- The compute jobs hit the storage quite hard (blue)
- These figures are prior to optimisation and with only 6 disk servers connected at 1GB
- The new configuration sees 12 disk servers connected at 10GB
- Note the bursts of incoming data (green)

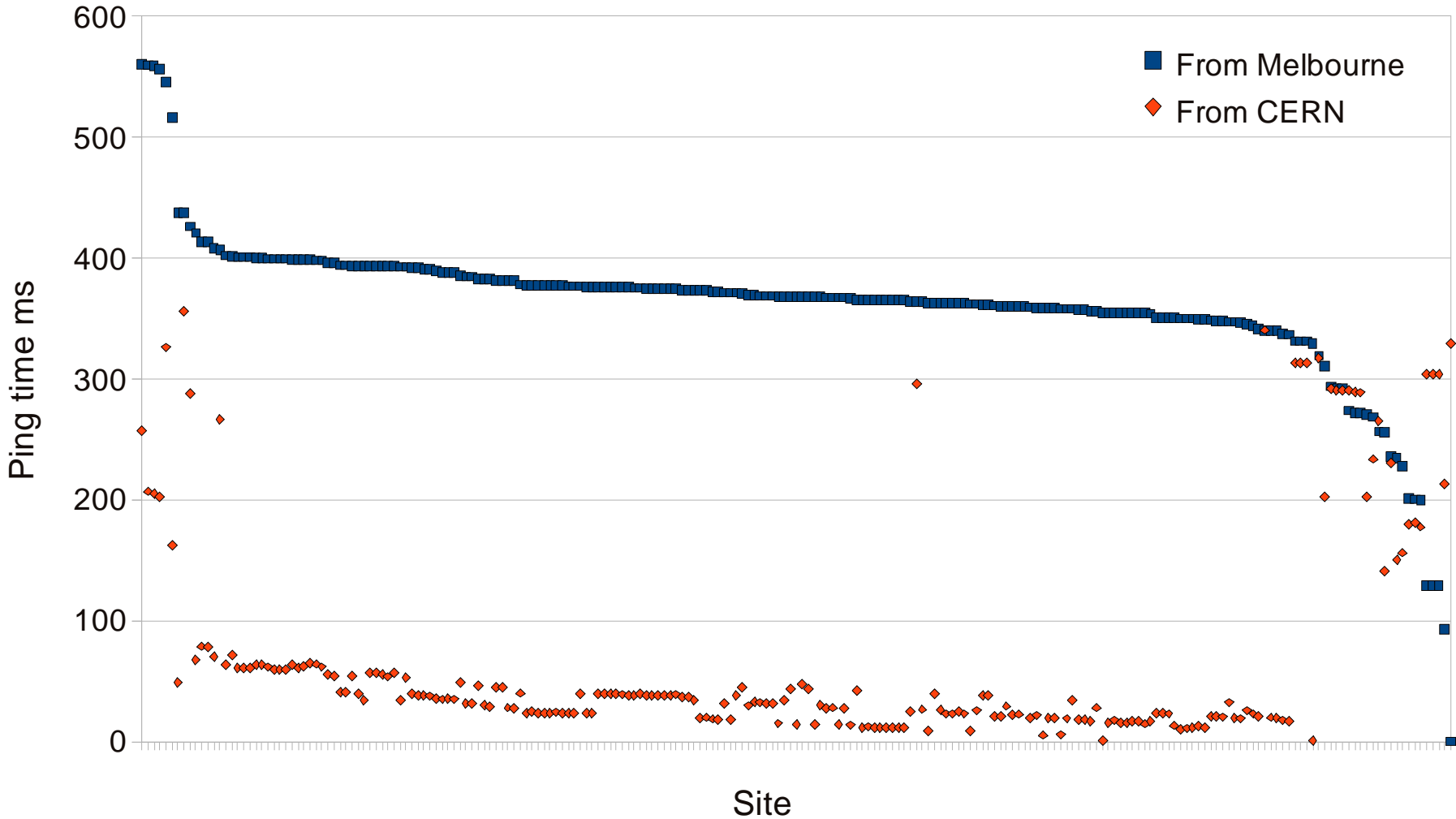
- Sustaining the data rates required to move the experimental data into Melbourne were a real concern

- The resource testing in June showed this is no longer an issue



- Our high latency has rendered us a canary for issues in the grid middleware
- The middleware components that involve message passing tend to struggle at high latency
- In some ways this has worked to our advantage
- We are now more closely involved with the developers, and are working with them to fix the issues
- Routing changes/fibre cuts also create interesting problems
 - Some European sites become invisible

BDII ping times

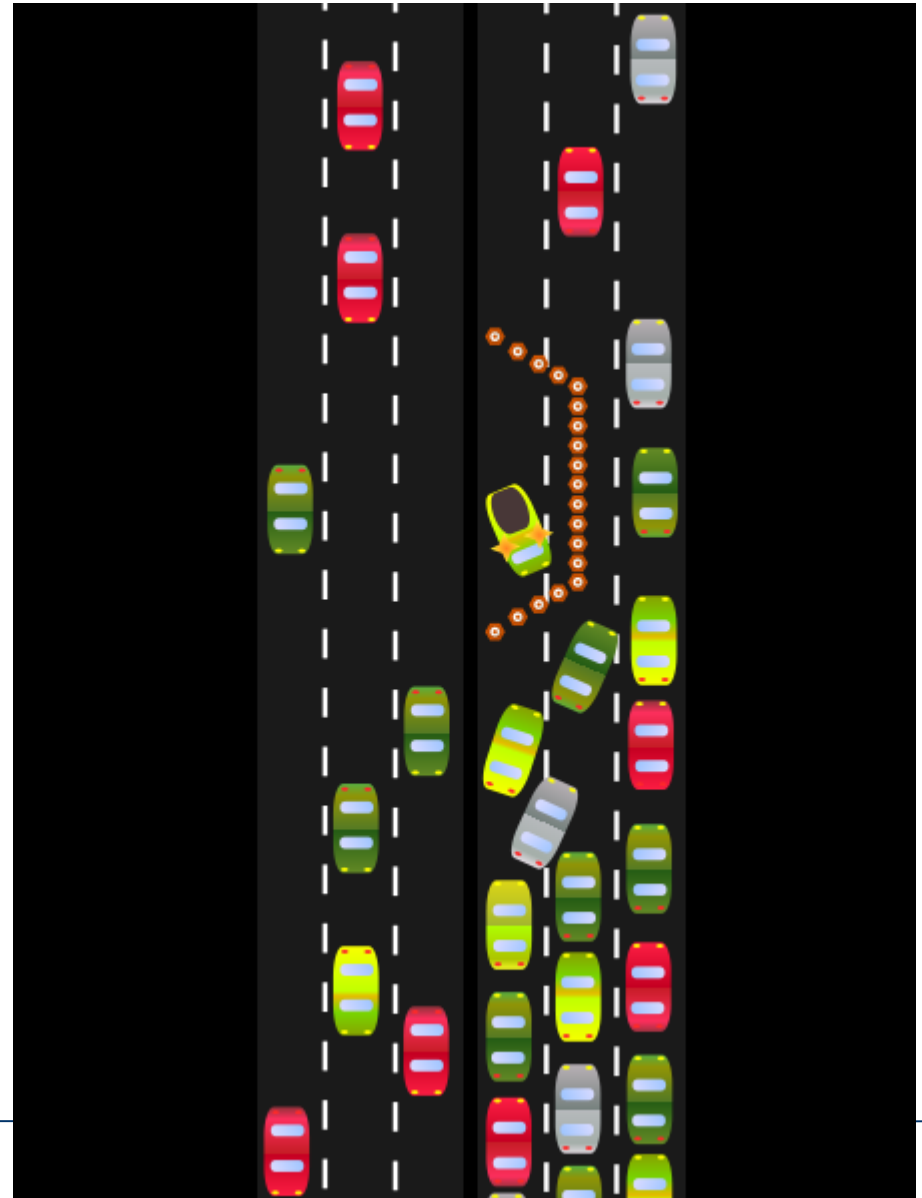


- Other problem components....
- The job accounting system (RGMA)
 - Some accounting weirdness observed.....
- LFC/DQ2 file registrations
 - When a file is transferred to the site it is registered into the file catalog, in Taiwan
 - Each registration is a multiple transaction process
 - 1000's of files.....
 - This was causing large delays, but has since been fixed
 - Files are now registered in bulk into the catalog



What Are We Missing?

- 10Gigabit
- 10Gigabit
- 10Gigabit

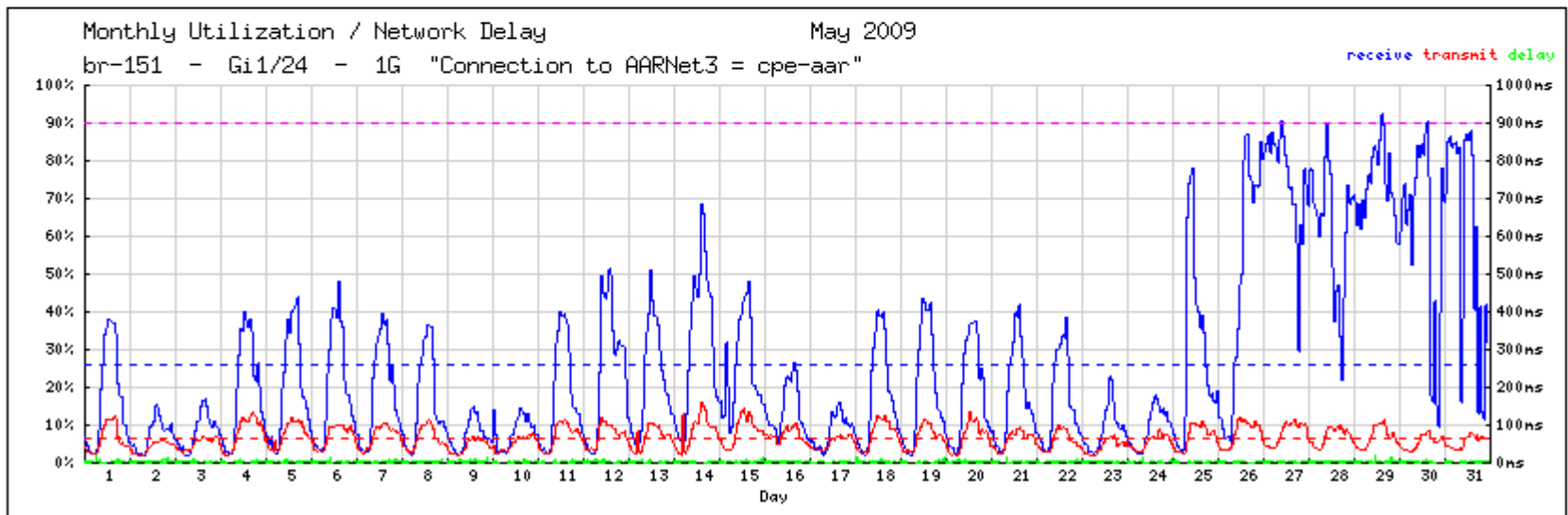


What Are We Missing?

br-151 Interface: Gi1/24 SNMP Index: 24 "Connection to AARNet3 = cpe-aar"

[ssh /telnet](#) to 192.101.254.17

Monthly Counts: May 2009						
	Average	Bytes	Frames	Errors	Discards	FECN / BECN
Receive	25.64%	85.5T	82.7G	0	0	BECN 0
Transmit	6.34%	21.2T	57.1G	0	705.1K	FECN 0
Delay	0 ms					
Specified Range: 8am - 6pm Mon, Tue, Wed, Thu, Fri						
Receive	40.81%	38.4T	37.0G	0	0	BECN 0
Transmit	9.31%	8.8T	25.3G	0	639.7K	FECN 0
Delay	0 ms					

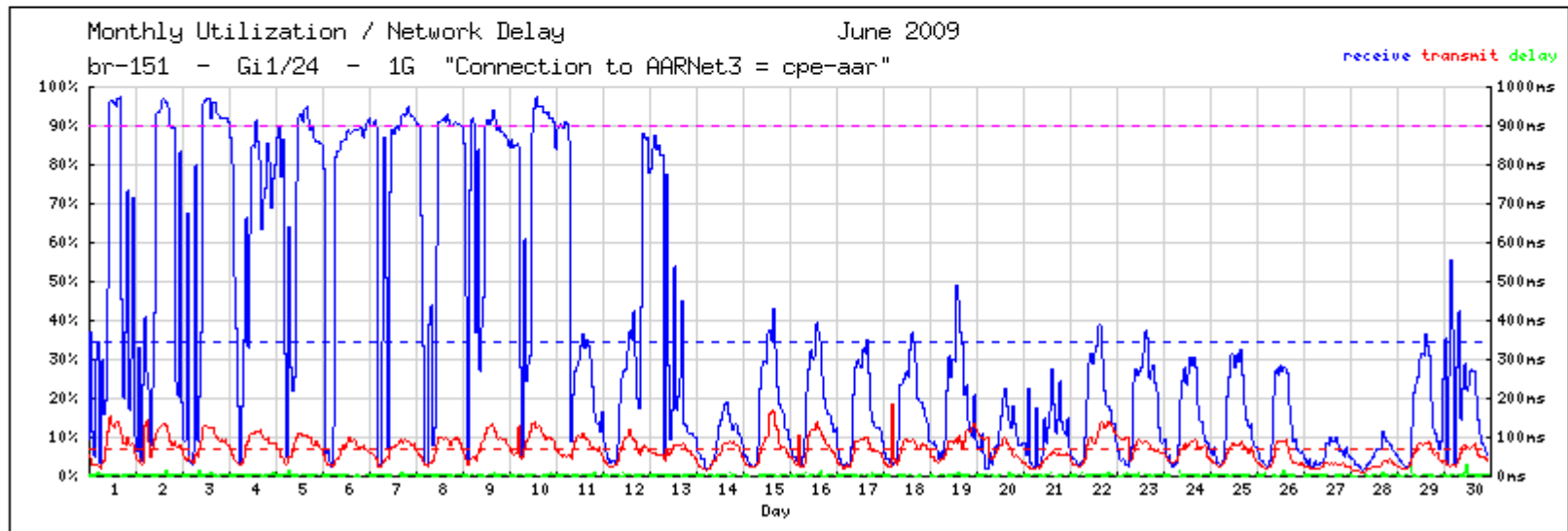


What Are We Missing?

br-151 Interface: Gi1/24 SNMP Index: 24 "Connection to AARNet3 = cpe-aar"

[ssh / telnet](#) to 192.101.254.17

Monthly Counts: June 2009						
	Average	Bytes	Frames	Errors	Discards	FECH / BECH
Receive	34.06%	110.0T	96.6G	0	0	BECH 0
Transmit	6.51%	21.0T	64.9G	0	1.9M	FECH 0
Delay	0 ms					
Specified Range: 8am - 6pm Mon, Tue, Wed, Thu, Fri						
Receive	43.28%	42.8T	39.2G	0	0	BECH 0
Transmit	9.00%	8.9T	27.3G	0	1.7M	FECH 0
Delay	0 ms					



- LHC startup slated for late this month (November)
 - The production data transfers will begin in earnest
- We have setup PerfSonar
 - We are talking with partners about automated testing of the network
 - We plan to schedule and monitor these tests via our Nagios monitoring system
- Jumbo frames
 - Currently enabled at all points on our path except the disk servers of our data sources
- Prior to machine startup we will most likely do another large data transfer (100T+) to make sure we are ready for production data flows



THE UNIVERSITY OF

MELBOURNE