

The HEPiX IPv6 working group

David Kelsey (STFC-RAL)

HEPiX meeting, Bologna

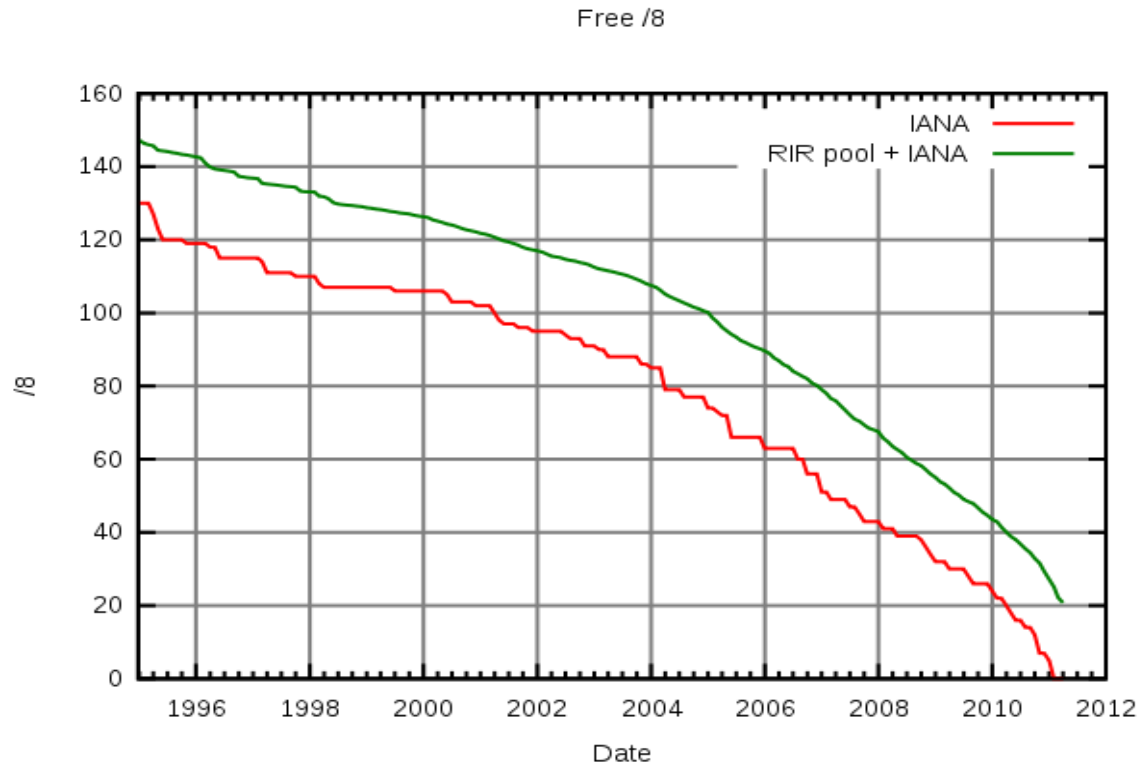
17 Apr 2013

Outline

- Update since Beijing HEPiX (Oct 2012)
- IPv4 address exhaustion – current status
- What is new?
- HEPiX IPv6 testbed
 - File transfer tests
- Software and tools survey
- Testing plans for 2013

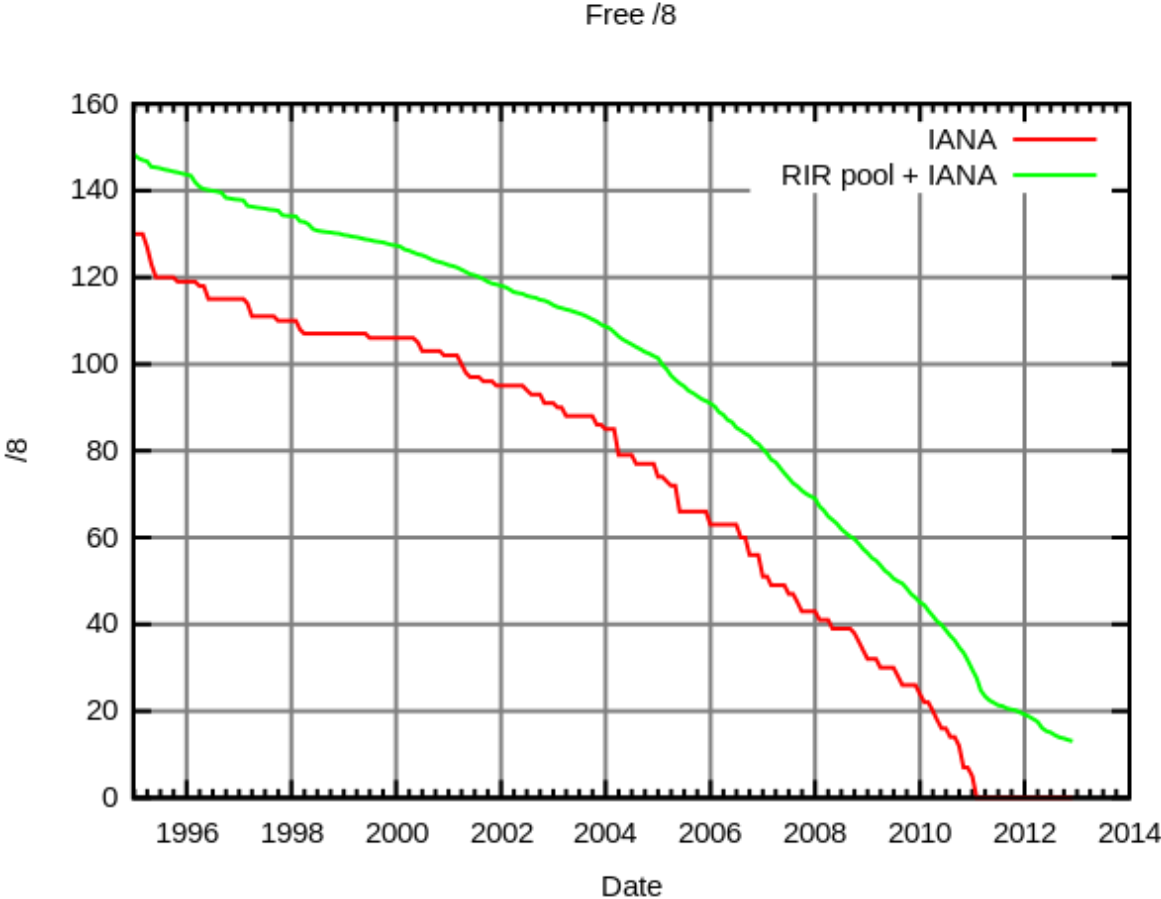
IPv4 address exhaustion ...

IPv4 Free Addresses (/8 blocks) – one year ago



<http://en.wikipedia.org/wiki/File:Ipv4-exhaust.svg>

IPv4 free addresses - now



IPv4 Addresses

- From Geoff Huston (<http://ipv4.potaroo.net>)
- IANA Unallocated Address Pool (Global)
Exhaustion happened: **03-Feb-2011**
- Projected Regional (RIR) Address Pool Exhaustion Dates:
 - APNIC: **19-Apr-2011** (Asia Pacific - happened)
 - RIPENCC: **14-Sep-2012** (Europe - happened)
 - ARIN: **02-Apr-2014** (North America)
 - LACNIC: **26-Aug-2014** (South America)
 - AFRINIC: **24-Jul-2020** (Africa)

HEPiX IPv6 WG

– what is new? ...

IPv4 shortage and CERN

24 January 2013

edoardo.martelli@cern.ch

- IPv4 shortage will soon hit CERN
- Applications will have to live either with private ipv4 addresses or ipv6-only stacks
- Use of IPv6 in the WLCG have to start as soon as possible

Timetable for HEP IPv6 transition

- Last year we said
 - Support for IPv6-only clients *not before* Jan 2014
- Now turning into
 - Support for IPv6-only WNs is required *by* 2014
 - Or soon after
- This will be a challenge!
 - We need to focus carefully
 - And we will need more resources
 - People and equipment

What else is new?

- We had two F2F meetings plus video calls
- New testbed sites
 - IHEP (CN), Glasgow, Imperial London, PIC, USLHCNet Caltech (Chicago)
 - Others in process of joining
- LHC Experiments
 - CMS (in from the start)
 - LHCb was next
 - ATLAS and ALICE now also on board

What is new? (2)

- New group members
 - ATLAS (+2), CMS (+1), PIC (+2), IHEP/CN (+2), CERN/IT/DM (+1), Imperial (+2)
 - More GridPP (UK) sites planning to join soon
- During 2013, we need to include more sites
 - All Tier 1s
 - But more Tier 2s too
- Ixplus at CERN is deploying some IPv6 nodes

The IPv6 testbed ...

The HEPiX IPv6 Testbed

- We have deployed a distributed testbed
- Connected to IPv6 and IPv4 networks
 - IPv6-only/IPv4-only names also registered in DNS
 - e.g. hepix-v6.desy.de & hepix-v4.desy.de
- <https://w3.hepik.org/ipv6-bis/doku.php?id=ipv6:testbed>
- A perl script (on wiki) validates configuration
 - Checks all DNS entries
 - runs ping and ping6 to all nodes

IPv6 Testbed

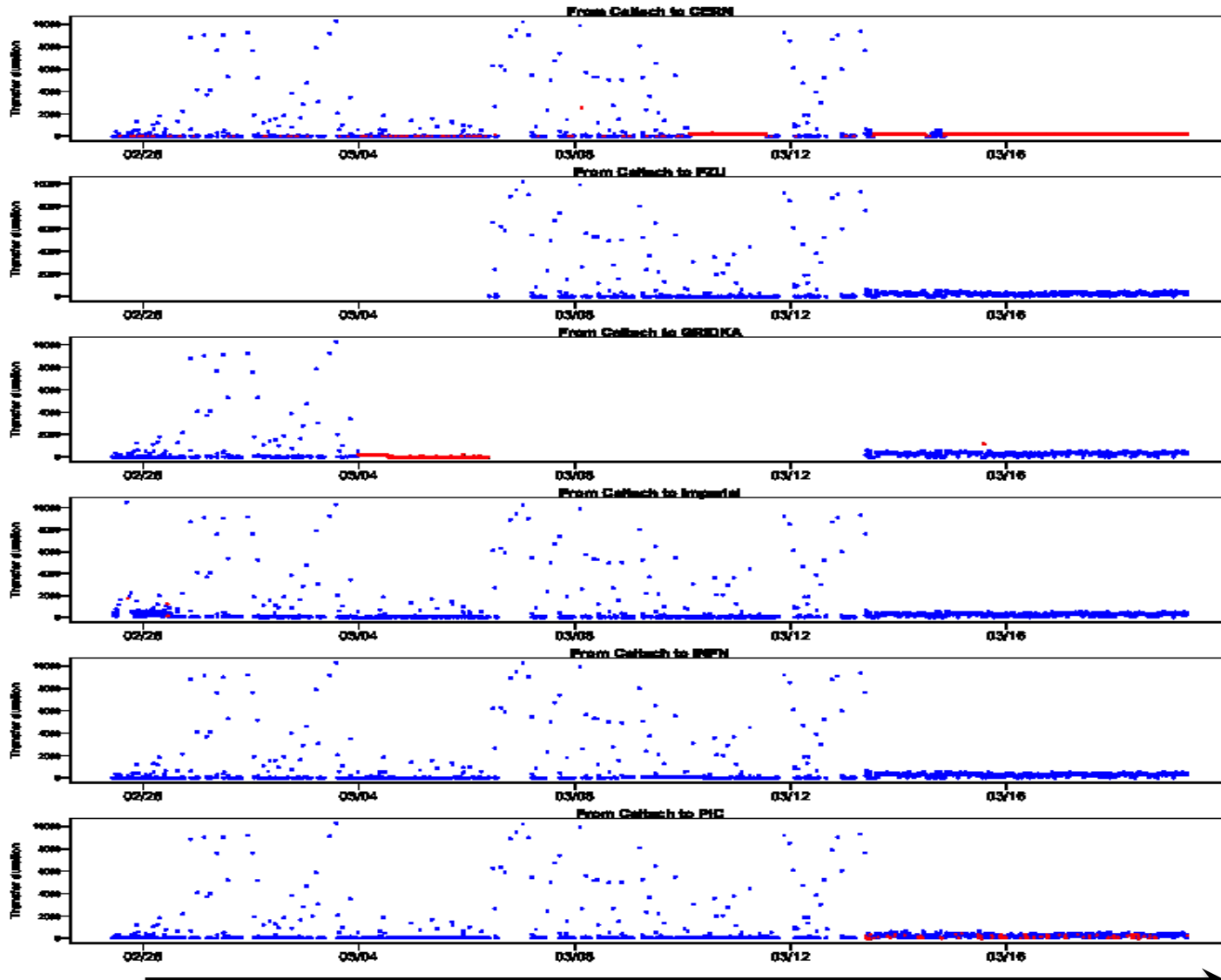
| Location | Institution | Contact | IP name | IPv6 name and address | IPv4 name and address | Installed OS | Active services | Comments |
|---|-------------------------------|---|--------------------------------|--|--|---------------------|--|------------------------|
| Milan, Italy (45.475773 9.231127 E) | INFN | Francesco Prelz Francesco.Prelz@mi.infn.it | seipersei.mi.infn.it | seipersei-v6.mi.infn.it (2001:760:4210:1::d) | seipersei-v4.mi.infn.it (192.84.138.225) | SLC x86_64 5.7 | - gridftp server - FTS | |
| Geneva, Switzerland | CERN | Edoardo Martelli edoardo.martelli@cern.ch | lx5v6hepixon1.cern.ch | lx5v6hepixon1-v6.ipv6.cern.ch (2001:1458:301:a717:215:5dff:feff:5d50) | lx5v6hepixon1- v4.ipv6.cern.ch (128.142.32.26) | SLC x86_64 5 | gridftp server | |
| Geneva, Switzerland | CERN | Edoardo Martelli edoardo.martelli@cern.ch | v6hepixon.cern.ch | v6hepixon-v6.cern.ch (2001:1458:301:a7bc::100:f) | v6hepixon-v4.cern.ch (128.142.223.240) | SLC x86_64 6.3 | gridftp server | |
| Hamburg, Germany | DESY | Kars Ohrenberg kars.ohrenberg@desy.de | hepixon01.desy.de | hepixon01-v6.desy.de (2001:638:700:1004::1:be) | hepixon01-v4.desy.de (131.169.4.190) | SLD x86_64 5.7 | icmp gridftp | |
| Karlsruhe, Germany | KIT | Bruno Hoefft bruno.hoefft@kit.edu | hepixon01.gridka.de | hepixon01-v6.gridka.de (2a00:1398:104:6167::159) | hepixon01-v4.gridka.de (192.108.45.159) | SLC x86_64 5.5 | gridftp | eth0:IPv6 eth1:IPv4 |
| Geneva, Switzerland | USLCHNet Caltech | Ramiro Voicu ramiro.voicu@cern.ch | hermes-gva.uslhcnnet.org | hermes-gva-v6.uslhcnnet.org (2001:67c:2c4:34::141) | hermes-gva- v4.uslhcnnet.org (192.65.196.141) | SLC x86_64 5.9 | gridftp | |
| Chicago, US | USLCHNet Caltech | Ramiro Voicu ramiro.voicu@cern.ch | ml-chi.uslhcnnet.org | ml-chi-v6.uslhcnnet.org (2001:67c:2c4:32::2) | ml-chi-v4.uslhcnnet.org (192.65.196.72) | SL x86_64 6.3 | gridftp | |
| Rome, Italy | GARR | Mario Reale mario.reale@garr.it | hepixon-ui.dir.garr.it | hepixon-ui-v6.dir.garr.it (2001:760:0:106::80) | hepixon-ui-v4.dir.garr.it (193.206.106.80) | SLC x86_64 5.7 | gridFTP server, glite UI 3.2 | |
| Prague, Czech Republic | FZU | Tomas Kouba koubat@fzu.cz | ui.ipv6.farm.particle.cz | ui-v6.ipv6.farm.particle.cz (2001:718:1e01:1725:5054:ff:fe3f:6545) | ui- v4.ipv6.farm.particle.cz (147.231.25.65) | SL x86_64 5.7 | gridftp | eth0:IPv6 eth1:IPv4 |
| Barcelona, Spain | PIC | Fernando Lopez Munoz flopez@pic.es | hepixon01.pic.es | hepixon01-v6.pic.es (2001:67c:1148:500::6) | hepixon01-v4.pic.es (193.109.175.122) | SL x86_64 5.9 | gridftp | eth1:IPv6 eth0:IPv4 |
| London, UK | Imperial College London | lcg-site- admin@imperial.ac.uk | hepixon00.grid.hep.ph.ic.ac.uk | hepixon00-v6.grid.hep.ph.ic.ac.uk (2001:630:12:580:216:3eff:fe7f:144) | hepixon00- v4.grid.hep.ph.ic.ac.uk (146.179.247.167) | CentOS6.3 x86_64 | gridftp | |
| Beijing, China | IHEP | qfz@ihep.ac.cn | ui01-hepixon.ihep.ac.cn | ui01-hepixon-v6.ihep.ac.cn (2401:de00::9998) | ui01-hepixon-v4.ihep.ac.cn (202.122.32.172) | SLC5.5 x86_64 | gridftp | |

IPv6 file transfer tests

- Tony Wildish (CMS)
- Simple data transfers between all nodes in the testbed (over IPv6 channels) - simultaneously
- Transfers a 1 GB file using GridFTP
 - Measures time to transfer
 - Records any errors
- Uses UberFTP to confirm arrival and then delete
- Then starts again
- Very useful for checking ongoing status
- Also for spotting and debugging problems

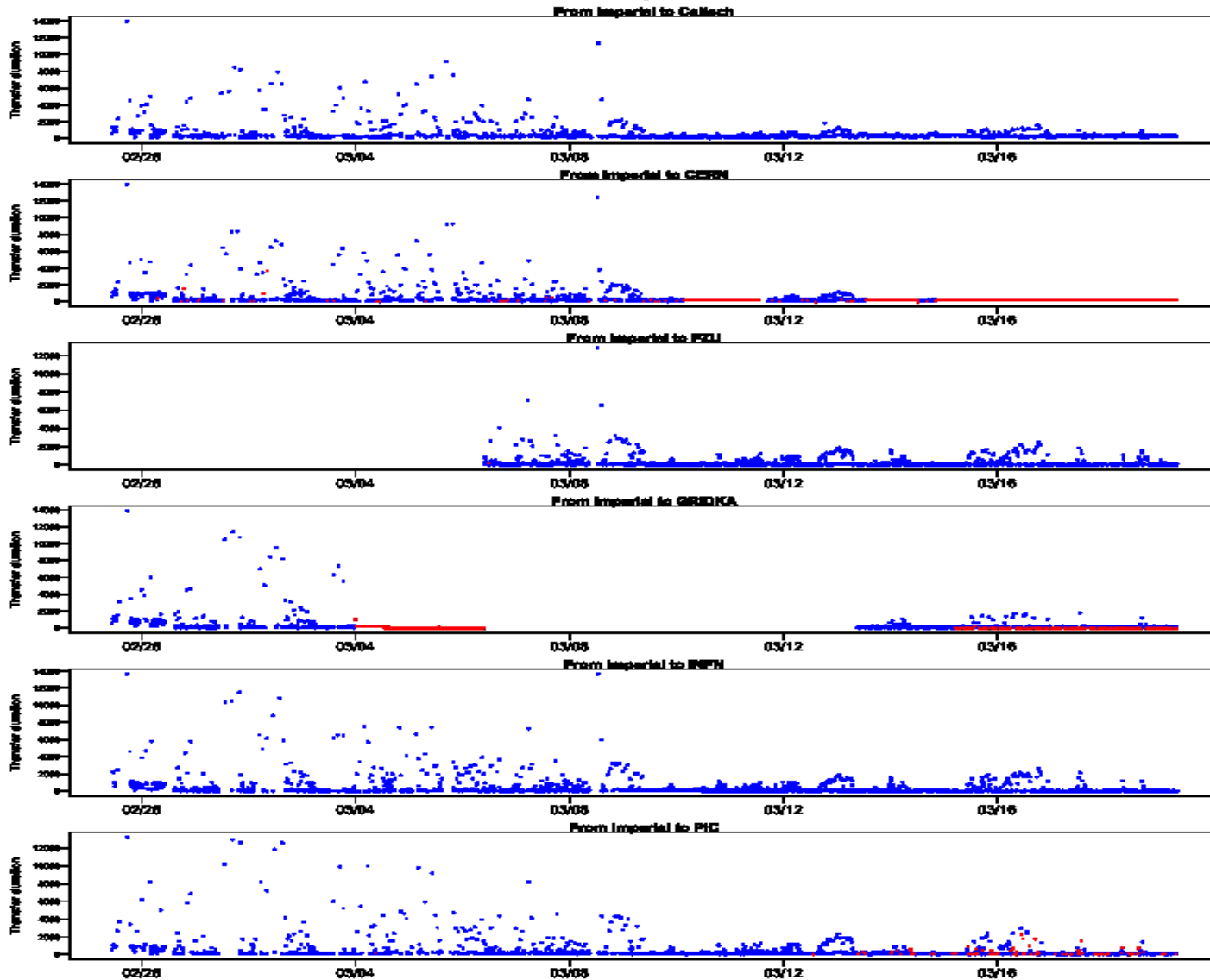
File transfers – Caltech to other sites

Time to transfer (secs)

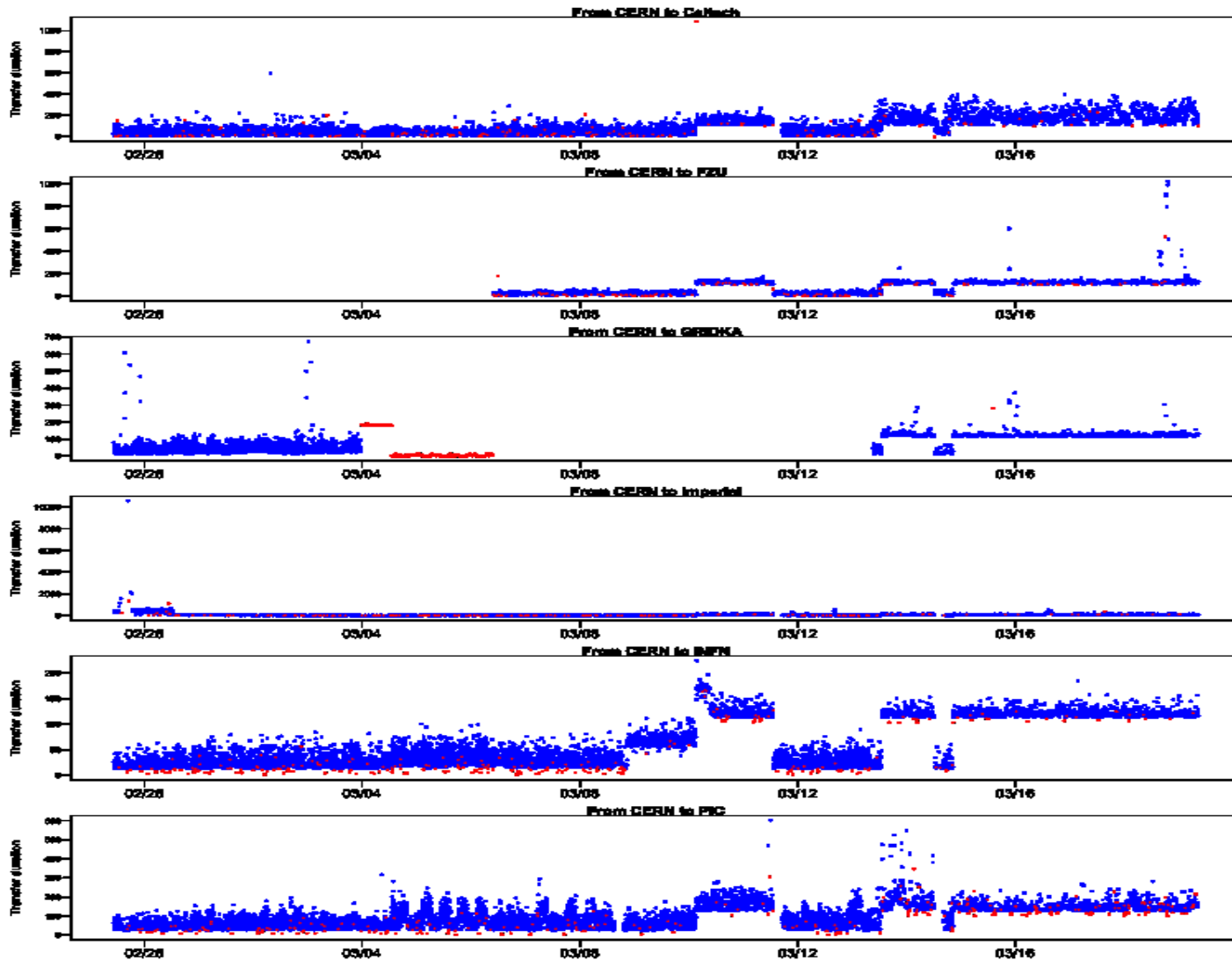


Success
Error

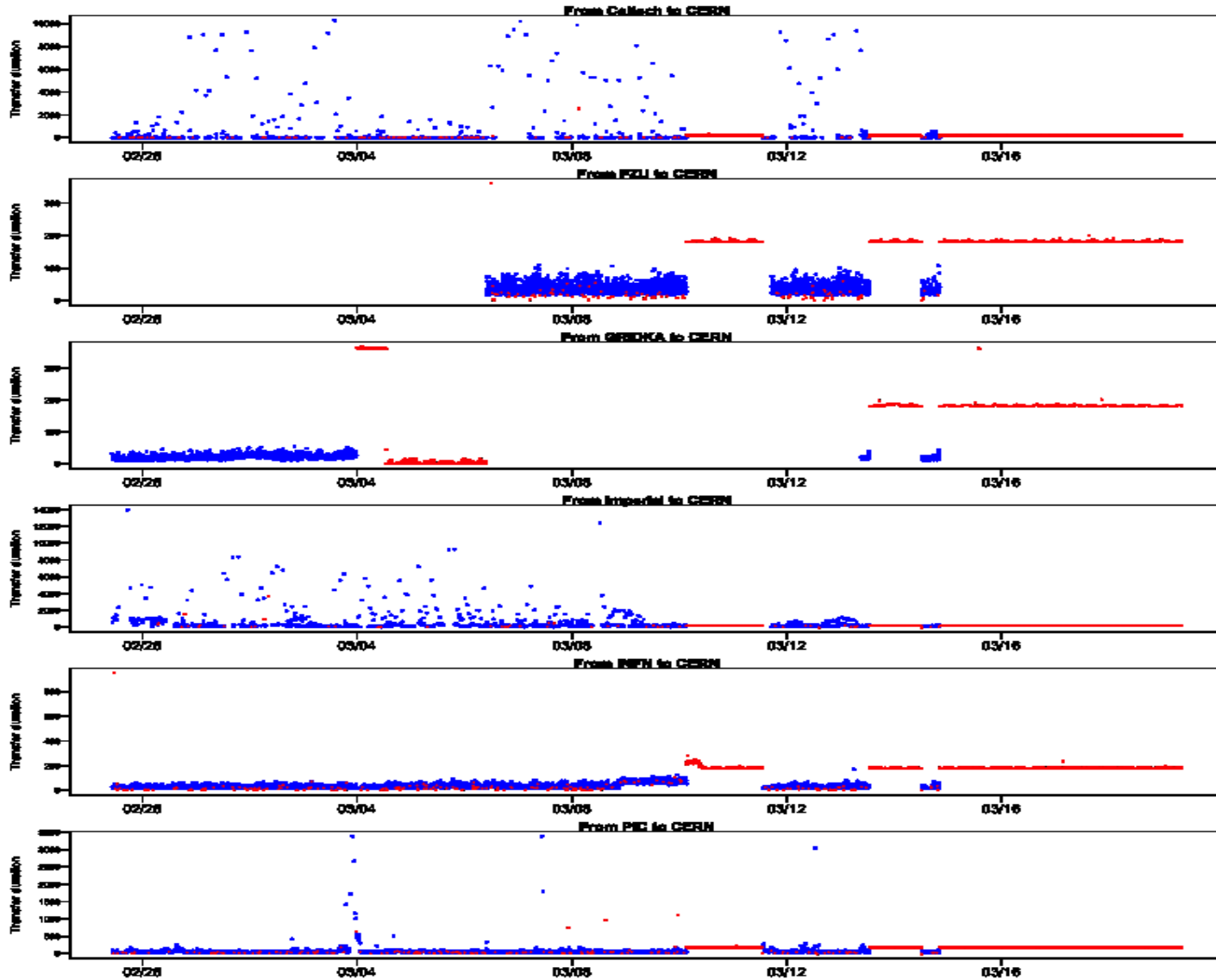
File transfers – Imperial to other sites



File transfers – CERN to other sites



File transfers – sites to CERN



Software & Tools IPv6 Survey

- An “Asset” survey is still underway
 - Spreadsheet to all sites and the LHC experiments
 - Includes **all** applications, middleware and tools
- If IPv6-readiness is known, can be recorded
- Otherwise we will need to investigate further
 - Ask developer and/or supplier
 - Scan source code or look for network calls while running
 - Test the running application under dual stack conditions

IPv6 problems

- See talk by Francesco Prelz – following this
- Batch systems
- OpenAFS (see talk by Arne Wiebalck)
- activemq (used in FTS3)
- dCache is being worked on
- This is not a complete list
- up to date information will be on our wiki site
 - One of our objectives for next 3 months

Batch systems

- News from EGI IPv6 testing
 - Barbara Krasovec (Arnes, Slovenia)
- Tested SGE, Slurm and Torque
- SGE had no support, current status unknown
- Slurm had no support for IPv6 and the status remains the same
- PBS had support for the server side
 - none for the client
 - EGI has tested versions 2.5.7 and 4.x

Future plans and next steps ...

CMS data transfer tests

- Small-scale testbed – one machine per site
- Agreed that this is still very useful
 - To show that data transfers can be sustained
 - Useful for debugging site issues
- Small scale is good to involve many sites
 - Helps learn about IPv6 and check site network
- Not only the current GridFTP mesh
- Add PhEDEx, FTS and Storage elements
 - Starting with DPM

IPv6 on SL5 or SL6?

- Most testing should continue on SL5
 - use same software as current production
 - If you want to use SL6 that is fine – your choice
- Let others debug general SL6 issues
 - When used in production we will move to it

Larger scale testing 2013

- Some sites report they have tried dual-stack and it works
- The group is reluctant to use the WLCG production infrastructure until fully convinced it will not break things
- **To do this we need more extensive testing**
- Glasgow has a larger IPv6 mini-cluster ready for use
- KIT has plans to install one in the coming months
- We will need some more sites
 - CERN – yes (but need effort from WLCG)
 - “Across the ocean” sites good to test long distance behaviour
 - IHEP (Beijing)?, FNAL and/or BNL?
- Others will also want to join
 - DESY for one

Testing: use cases

- Simplest in terms of needs
 - Production Monte Carlo
 - Use case could be IPv6-only machines in opportunistic Cloud resources
 - Or IPv6only worker nodes at CERN
- Next is Production Reconstruction
 - What services does that need?
- Most complex in terms of requirements is general user analysis
 - E.g. requirement (?) to connect to OpenAFS

Production Monte Carlo

- Start with real Worker Nodes (IPv6 only)
 - rather than VMs
- Required network access
 - Some form of workload management
 - To get the work into the WN
 - Output from the job
 - Presumably needs to write to an SE?
- Experiments to specify the details

Other testing plans – next quarter

- CMS - DPM endpoints at Glasgow and FZU
- LHCb (ScotGrid, Imperial, RAL)
 - Workload management and CVMFS
- ALICE – waiting for release of IPv6 in xrootd (V4 soon)
- ATLAS – plans not clear yet
 - Glasgow and CERN to look at production monte-carlo use case
- dCache on IPv6 testing
 - DESY (testbed and dCache team), PIC, KIT, NDGF
- USLHCnet – testing MonALISA
- INFN will work on IPv6 and CREAM CE
- CNAF Tier 1 joining – (STORM?)
- Then decide when we can plan test on production infrastructure

Further info

- HEPiX IPv6 wiki

<https://w3.hepix.org/ipv6-bis/>

- Working group meetings

<http://indico.cern.ch/categoryDisplay.py?categId=3538>

Summary

- During 2013, we must
 - Increase participation in the working group
 - Include all Tier 1s
 - When do we need all Tier 2s to be IPv6 capable?
 - Agree on what services need to be dual stack
 - For access from IPv6-only WN at CERN
 - Test these services
 - First on the testbed
 - Then on production infrastructure
- Next face to face IPv6 meeting is at CERN
 - Probably 4/5 July 2013 (to be confirmed)
- VOLUNTEERS always welcome (please contact me)!

Questions?