"From IPv4 to eternity": the HEPiX IPv6 working group

Edoardo.Martelli@cern.ch

CHEP2012, New York

21 May 2012

On behalf of the co-authors

Dave Kelsey (RAL), Bob Cowles (SLAC), Marek Elias (FZU), Thomas Finnern (DESY), Lars Fischer (NORDUnet), David Foster (CERN), Bruno Hoeft (KIT), Tomas Kouba (FZU), Simon Leinen (SWITCH), Mark Mitchell (Univ Glasgow), Kars Ohrenberg (DESY), Andreas Pfeiffer (CERN), Francesco Prelz (INFN), Mario Reale (GARR), Sandor Rozsa (Caltech), Sabah Salih (Univ Manchester), Luuk Uljee (SARA), Ronald van der Pol (SARA), Ramiro Voicu (Caltech), Mattias Wadenstein (Univ Umea), Tony Wildish (Princeton University)

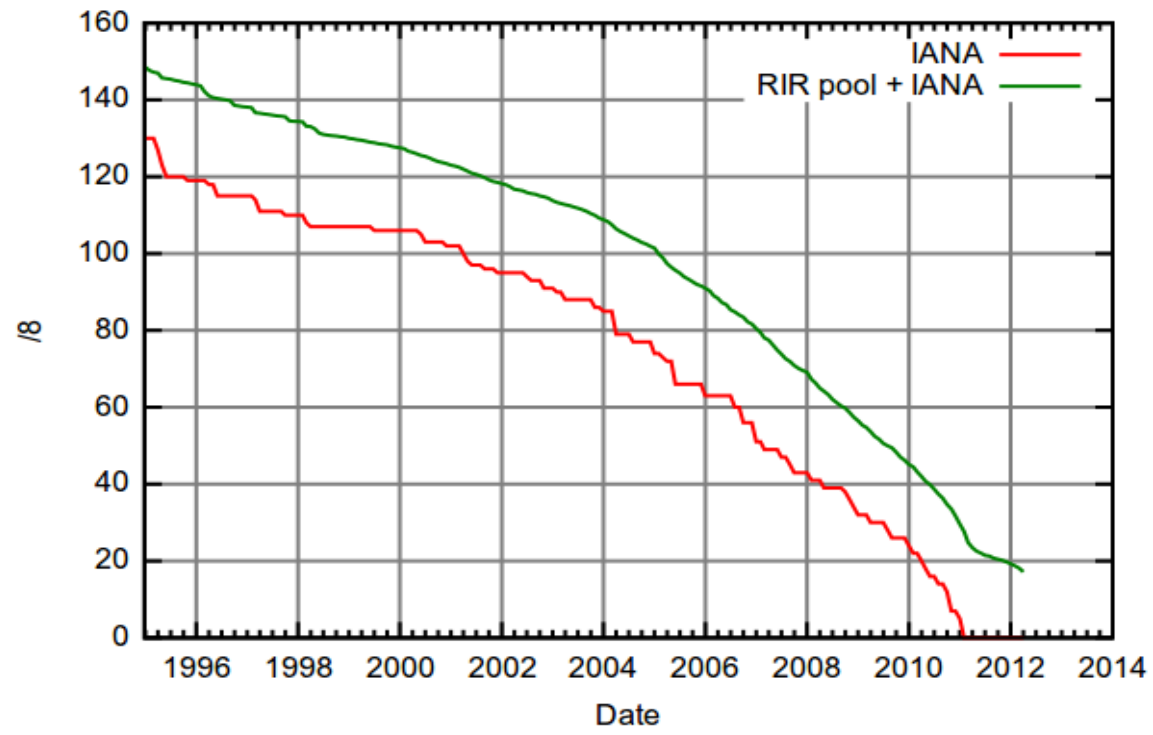And several others who have participated in the past

Outline

- Background – why move to IPv6?

- The HEPiX IPv6 working group

- HEPiX IPv6 Testbed

- Recommendations and future plans

Why Move to IPv6?

IPv4 address space depletion

Remaining IPv4 Free Addresses (/8 blocks):

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

IPv4 Addresses exhaustion dates

- 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-Aug-2012** (Europe)
  - ARIN: **20-Jun-2013** (North America)
  - LACNIC: **29-Jan-2014** (South America)
  - AFRINIC: **05-Nov-2014** (Africa)

(source: http://www.potaroo.net/tools/ipv4/index.html   )

World IPv6 Launch Day

- http://www.worldipv6launch.org/

- 6 June 2012 "The Future is Forever"

- ISPs, home routing equipment vendors, web companies all coming together

- Permanently enable IPv6 by 6th June 2012

Ready to go?

- IPv6 *is* coming to stay!
- And it will bring a lot of new opportunities
- The IPv4 software legacy will keep us busy for a while.
- Are the HEP and WLCG communities ready?

Hepix IPv6 Working Group

WG membership

- Chaired by Dave Kelsey (RAL)
- Active members:
  - CERN, DESY, FNAL, FZU, GARR, Glasgow, INFN, KIT, Manchester, RAL, SLAC, USLHCnet (Caltech)
  - CMS, ALICE and LHCb (ATLAS to come)
- Nearly 50 on the mail list
- Regular video and face-to-face meetings

HEPiX IPv6 WG

Created in April 2011 with aims:

- Consider how IPv6 should be deployed in HEP and especially in WLCG

- Readiness and Gap analysis

- HEP applications, Middleware, Security issues, System management and monitoring tools, End to end network monitoring tools

- Run a distributed HEP testbed to verify compliance in reality

- Share experiences

IPv6 and WLCG

Currently it's not known when WLCG will need to deploy IPv6-capable services.

But:

- up rise of VMs and Cloud services may be impaired by the IPv4 shortage

- some sites in the Far East already prefers IPv6

IPv6 and WLCG

To get there takes time and effort!

- It's not only a network deployment challenge
- software and tools must be able to use IPv6.
- Need operational monitoring, security and tools
- IPv6 performance must be as good as IPv4

WG activity:
Implementation check list

List of steps to take and recommendations to design and implement a dual stack network

Based on the real experience of the members

WG activity:
Software & Tools IPv6 Survey

- An "Asset" survey is now underway
  - A spreadsheet to be completed by all sites and the LHC experiments
  - Includes **all** applications, middleware and tools
  - Tickets to be entered for all problems found
- 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

WG activity:
Distributed Dual Stack Testbed

A place where to gain real experience

Implemented on real networks, in a distributed environment as close as possible to production

Open to anyone in WLCG

To test applications over IPv6 but also in the dual-stack cohabitation

Software with IPv6 problems

- Need to check many things
  - Break when installed on a dual-stack node?
  - Does it bind to both stacks?
  - Is IPv6 preferred?
  - Can it be configured to prefer V4 or V6?
- Already found a few problems

WG activity:
Attention to Security

New protocol => many new weaknesses to exploit

Everything to be verified again

Share recommendations and best practices

Working with others

- Use tools developed by EGEE
- EGI recently started IPv6 activities
  - Work together
  - Avoid duplication
- EMI testing nightly builds for IPv6 compliance
  - Will work together
  - Allow EMI developers access to our testbed

EGEE IPv6 tools

- Source code checker
  - A bash script looking for non compliant function calls and address data structures

- Dynamic Code Checker (IPV6 CARE tool)
  - A tool based on the LD_PRELOAD mechanism to intercept calls to non compliant functions in the dynamically linked libraries

- Analysis of all gLite code was performed
  - And code was modified to fix problems
- *https://twiki.cern.ch/twiki/bin/view/EGEE/IPv6FollowUp*

HEPiX IPv6 Testbed Experiences

The IPv6 Testbed

- Deployed a distributed testbed: CERN, DESY, FZU, GARR, INFN, KIT and USLHCnet

- All installations have uniform architecture (x86_64) and uniform OS (Scientific Linux 5) for full support of `WLCG' applications and middleware.

- Each node with at least 1 Gbps network connection.

- All running at least one GRIDFTP server, giving access only to the ipv6.hepix.org VO.

- https://w3.hepix.org/ipv6-bis/doku.php?id=ipv6:testb

GridFTP tests

- Successfully installed and tested GridFTP clients and servers on all nodes

- Virtual organization ipv6.hepix.org

- Full mesh of data transfers (globus_url_copy): Tested and works

- CMS now performing continuous data transfers between pairs of nodes

The CMS file transfer tests

- Reliability test - not a stress/performance test
- Single 2000 MB file from IPv6 VM at CERN transfered to 4 systems
- globus_url_copy and uberftp to confirm file arrived then delete
- Tests have been running continuously since February 2012
- Statistics since April 20th:

Site	#_of_transfers	Failed_transf.	Average_duration	Duration_range
DESY	390	13 (3.3 %)	66s (~30 MB/s)	41 - 425s
Gridka	780	29 (3.7 %)	130s (~15 MB/s)	110 - 439s
INFN	1299	43 (3.3 %)	66s (~30 MB/s)	34 - 549s
Uslhcnet	1299	28 (2.2 %)	81s (~25 MB/s)	38 - 549s

Can still conclude: no show-stoppers. CMS PhEDEx should work.

FTS

- gSOAP supports IPv6
  - on TCP since version 2.5 (2005)
  - on UDP since version 2.7.2 (still 2005)
- BUT compiled without the "WITH_IPv6" flag.
- Oracle IPv6-enabled from version 11g rel 2, but FTS transfer agent libraries in EMI-1 still carry a hard dependency on Oracle V10
- Transfer agents (Tomcat/Axis servlets) can be invoked on dual stack hosts and from dual stack clients
- but 'urlcopy' agent still uses IPv4 for file transfer
- As in the globus-url-copy command, IPv6 resolution in the Globus FTP client needs to be *explicitly* enabled

UberFTP

- UberFTP is an interactive GridFTP client tool

- No IPv6 support

- Francesco Prezl has written an IPv6 patch and sent it to the author for integration

Testbed results so far

- Time consuming

- Several fake "IPv6 ready" exposed

- Necessary to develop and test IPv6 patches

More results and Future plans

Managing IPv6 at large sites

- Best practices are still far from clear!
- Large sites (e.g. CERN and DESY) wish to manage the allocation of addresses
  - Do not like autoconfiguration (SLAAC)
- Wish to filter out Router Advertisements
- DHCPv6 very attractive
  - BUT IETF still discussing
  - Will the 'route' options be there or not?

IPv6 security

- Are operational security teams ready for IPv6?   No!
- Challenges include
  - Address format has multiple forms, many addresses per host and addresses difficult to remember
  - IPv6 standards contain many suggestions - implementation optional
  - Required security features, like RAGuard and SEND, are a long way from full deployment
  - Incomplete and immature implementations
  - Many vulnerabilities expected
  - Log parsing tools must all change
  - Dual stack and tunnels cause problems – e.g. packet inspection
- Must test that things which are supposed to be blocked really are.

Future plans

* Aim to implement Dual Stack on all WLCG services
  - Avoid complications of tunnels, proxies, gateways etc.
- Perform full asset survey (Spring 2012)
  - Identify show-stoppers & quantify effort and resources required to fix
- Expand testbed gradually during 2012
  - work with EGI and EMI
  - Considering merging of EGI and HEPiX testbeds later this year
  - All WLCG services
  - Perform more extensive functionality and performance tests
- Must consider operational impact
  - including security and monitoring

Future plans (2)

- Review status at end of 2012

- Produce implementation plans for 2013 and/or later

- Need to perform tests on the production infrastructure

  – involve WLCG Tier 1 centres

- Plan several HEP IPv6 "Days" (for 2013?)

  – turn on dual stack for 24 hours on production infrastructure and test/observe

Conclusion

Conclusions

- The HEPiX IPv6 working group is functioning well

- *MUCH* work still to be done during the next year or two & effort is difficult to find (new volunteers are welcome to join)

- IPv6 won't be easy to exploit in the beginning, but then it will bring new functionalities and opportunities

Further info

- HEPiX IPv6 wiki

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

- Working group meetings

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