

Table of Contents

Summary of pre-GDB on IPv6, June 11, 2014 (CERN).....	1
Agenda.....	2
IPv6 1010 - E. Martelli.....	3
IPv6 - IPv4 Semantics Differences - F. Prezl.....	4
IPv6 Deployment at CERN - E. Martelli.....	5
Tests - T. Wildish.....	6
IPv6 Readiness and Configuration Guidance.....	7
WLCG SW and Tool Survey - D. Kelsey.....	7
DPM - S. Skipsey.....	7
dCache - U. Tigerstedt.....	7
StoRM - C. Walker.....	7
FTS3 - M. Salichos.....	8
Xrootd - L. Janyst.....	8
Batch Systems - F. Prezl.....	8
Experiment Plans.....	10
ATLAS, CMS and LHCb - A. Dewhurst.....	10
ALICE - C. Grigoras.....	11
Monitoring IPv6.....	12
perfSonar - D. Rand.....	12
Nagios - M. Elias.....	12
Site Status and Plans.....	13
Site Survey - D. Kelsey.....	13
UK IPv6 Status - C. Walker.....	13
IPv6 at FZU (Prague) - M. Elias.....	13
IPv6 at BITP (Urainia) - O. Shadura.....	14
Wrap-Up - What Next?.....	15

Summary of pre-GDB on IPv6, June 11, 2014 (CERN)

Agenda

<https://indico.cern.ch/event/313194>

IPv6 1010 - E. Martelli

See slides, in particular for address formats.

No more broadcast addresses: replaced by special multicast addresses

Fragmentation/reassembly no longer permitted in routers: have to be done at source and destination.

ARP replaced by NDP (Network Discovery Protocol)

- Based on 5 types of ICMP packers: RS/RA, NS/NA, Redirect

SLAAC : address autoconfiguration

- Require a routed IPv6 network
- Based on ICMPv6 RD/RA packets + EUI-64 for generating local part of the address (derived from MAC address)

Several routing protocols

- IPv4 ones have a new version for IPv6 only
- ISIS : IPv4 and IPv6
- BGP : IPv4 and IPv6

Selecting IPv4 vs. IPv6

- Client decision based on DNS reply
- Generally one tried first (based on app config) and the other one tried if the first one is not replying: users will experience a delay (30 to 180s) if they have to wait for the timeout
- Some apps (e.g. some browsers) start to try both at the same time and use the one which answers first to avoid this problem

IPv6 - IPv4 Semantics Differences - F. Prelz

Every network endpoint is ALWAYS associated to multiple active network address: at least a "link-local" and one or more global addresses

- Origin and destination addresses have to be chosen for each connection: done with new system configuration parameters
 - ◆ `/etc/gai.conf` and `ip -6 addrelabel`)
 - ◆ Except for Mobile IP extension that defines the concept of a "Home Address"
- Deciding between IPv4 and IPv6 first is just a special case...
- Main RFC related to this issue: RFC6724
- Consistencies issues seen when upgrading from SL5 to SL6
- This difference is probably the main challenges in making apps IPv6-ready: most apps are not prepared to deal with a potentially large set of addresses

IP route assignment: only possible via a static route definition or via "Router Advertisements" which are multicast messages

- Default route cannot be assigned by dhcpv6

IP address and DNS server assignment

- DNS automatic update through Router Advertisements in stateless configuration should allow painless DNS configuration... but the RFC is not supported by some major vendors (e.g. CISCO)

Security potential implications

- Handling of header extensions optional in IPv4 are required in IPv6: good portion of new codes that may be a good target for vulnerability attacks
- Large number of optional protocol headers may mean that the upper-layer payload useful for filtering/firewalls may not be in the first fragment, requiring packet reassembly
- ICMPv6 is now required, at least on a network segment

IPv6 Deployment at CERN - E. Martelli

Dual stack done

- Network DB: IPv6 is now the main navigation data, IPv6 address tables fully populated
- Same routing infrastructure: BGP + OSPF

Identical perfs for IPv4 and IPv6

- All production network can forward IPv6 packets at wire speed
- Only exception: policy based routing for stateful firewall bypass but not a showstopper currently as IPv6 traffic is low

Common provisioning tools for IPv4 and IPv6: no change for users and network admins

- IPv6 info fully exposed
- Users can define their devices as IPv6-ready (connectivity ok, apps listening on both IPv4 and IPv6 addresses)

Same network services: dns, dhcp, firewall...

- firewall: most ACLs translated automatically for IPv6, a few additional rules IPv6-only

Challenges

- Size of routing tables and ACLs have doubled
- dhcpv6 still in an early stage
 - ◆ RAs necessary to configure default gateway and prefix length but no predictive load balancing in multi-router subnets: all prefixes exposed
 - ◆ MAC address authentication not always work with IPv6: no requirement to use the MAC address of the interface the packet is sent through. Management of UUID is required.
- New security threat to take into account
- New issues to be solved by support lines

Lessons learnt

- dhcpv6 is definitely not dhcpv4
- Do a staged deployment, involve users of different profiles

See <http://cern.ch/ipv6>

Tests - T. Wildish

3.4 PB transferred since the beginning of test 1 year ago

- Not a heavy traffic
- Everybody can play: only need an IPv6-enabled gridftp and an IPv4 uberftp
 - ◆ 20 GB of disk required
- Best effort support

gridftp dashboard useful to identify problems

- see slides

Future plan: moved to a FTS3-based testbed

- Currently PHEDEX-based

IPv6 Readiness and Configuration Guidance

WLCG SW and Tool Survey - D. Kelsey

<http://hepiv6.web.cern.ch/wlcg-applications>

- Good progress in coverage but still several important apps not covered (e.g. ARC CE and CREAM CE)

About to start looking at storage and batch systems

DPM - S. Skipsey

DPM itself is IPv6-compliant out of the box some dependencies requiring specific steps/versions

- Globus libs need changes to standard `/etc/gai.conf` to bind to IPv6
- MySQL v5.5+ required
 - ◆ Can use localhost alternatively
- YAAM needs hacks as it relies on IPv4 behaviour
- BDII must be dual-stacked for using IPv6 DPM
- fetch-crl sometimes have problems if binding to IPv6

DPM protocols

- gridftp: dependant on Globus libs, third-party gridftp needs some standard agreements
- xrootd v4 is IPv6-compliant but not xrootd v3 API compatible: new DPM plugin currently being tested
- http: no known issue

dCache - U. Tigerstedt

IPv6 support ready since 2.9 but not enabled by default

gsiftp: changed in 2.9.4+ to comply with globus-url-copy and DPM gsiftp rather than the standard...

xrootd: since 2.9.4+, IPv6 ready, compatible with xrootd v4 client

StoRM - C. Walker

Not really tested by developers but supposed to work.

QMUL tested it and found no issue except the incorrect log of IPv6 addresses to be fixed in GT6

- Also MySQL requires a minimum version to use IPv6 if wanted
- `/etc/gai.conf` needs to be updated to prefer IPv6 in a dual-stack environment
- Backend not starting correctly on reboot: being investigated
- Configuration adjustment needed to use IPv6 addresses instead of host names in SURF

FTS3 - M. Salichos

Problems identified with some dCache endpoints but no time to troubleshoot them yet: may be related to configuration problems or dCache version before 2.9

- Only problems in the gridftp transfer phase
- Presently prevents enabling IPv6 in production
 - ◆ Decision should be taken by developers and experiments
- Tests done last Friday... more investigation needed
- Imperial College used to run a dual-stack FTS3 instance and saw no problem with any site...

Known issues in dependencies

- ActiveMQ: IPv6-ready but the C client used by FTS3 was not until a very recent version
- MySQL v5.5+ is required
- GFAL2: IPv6 must be explicitly enabled in `~/etc/gfal2.d/gsiftp_plugin.conf`

Xrootd - L. Janyst

xrootd v4 client is IPv6-ready: old client is deprecated

- Will prefer IPv6 if available
- No configuration necessary except if wanting a different default binding strategy
- Old clients will work perfectly with xrootd v4 servers in IPv4

xrootd v4 server is also IPv6-read

- As for the client no configuration necessary if default configuration (dual-stack) is suitable

Federation (4.1): no longer any requirement for dual-stack servers

- Redirectory will attempt to match a downstream server with the same capacity as the client if the client is IPv4 or IPv6 only
- If both are supported by the client, both will be tried and the less loaded (the one answering the more quickly) will be used

Batch Systems - F. Prelz

Not a very critical issue: CE acts as a protocol translator, no direct connection to the batch system from out of the LAN

LSF: supposed to be IPv6-ready and dual-stack friendly for a very long time, v9 documents how to enable/configure it

- Cannot mix IPv4-only and IPv6-only nodes in the same cluster
- Not aware of any dual-stack validation in the community

Condor: single-stack pool operation available in v8 but dual-stack pools supposed to be rolled out in 8.3

- Problem with multiple addresses per endpoint)

Grid Engine: IPv6 support in UGE claimed for 8.3 (currently 8.1)

- A plan since 10 years...

Torque

- A SVN ipv6 branch started in 2007 and died since then...
 - ◆ Network code spread everywhere...
- Reports from different groups that a dual-stack Torque fails
- No plan for PBS Professional too

SLURM: claim that adding the IPv6 support should be straightforward but no real work started

- Impacted code seems to be concentrated in a few places...

Experiment Plans

ATLAS, CMS and LHCb - A. Dewhurst

Current VO focus is SW for Run2

- Hopefully this SW will make IPv6 testing easier
- Expect IPv6 support rollout to take many years...

When possible, experiments plan to make their central services dual-stack asap

- Medium term: SE with dual-stacked protocols
 - ◆ When enough available, IPv6-only CE/SE would be feasible
- ATLAS AMI: based on an Oracle back-end hosted at CCIN2P3, plan to dual-stack web frontends if possible

Workload Management Systems

- PanDA: currently migrating to BigPanda, exclusively using http so should work, Pilot factory had some initial evaluations with dual-stack
 - ◆ Imperial is planning a dual-stack BigPanDA test instance
- CMS: validation not yet started but no problem anticipated (rely on CondorG + glideinWMS that are supposed to work)
 - ◆ Dual-stack glideinWMS pilot planned soon
 - ◆ Also need to test CRAB3 used by users to submit analysis jobs
- DIRAC: recently tested from a dual-stack client, no issue found
 - ◆ Next test: IPv6-only WN

Data management: all expts relying on gridftp, xrootd and http

- Mainly through FTS3
- See previous discussion for exact plans
- ATLAS has setup a testing infrastructure for Rucio validation (stress tests): could be reused for IPv6 tests
- CMS already validated IPv6 xrootd: next step is to validate AAA
 - ◆ NEbraska will enable IPv6 to their xrootd servers from US redirector next week
 - ◆ CMSSW: a new release needed to upgrade the xrootd client
- LFC: not used or about to be abandoned in the next year, no plan to test

Databases

- Frontier/Squid still relying on Squid 2.x which is not IPv6 compliant
 - ◆ Squid 3.x is IPv6 ready but has long running problems that may impact significantly performances (lack of collapsed forwarding, now fixed and headers not updated)

CVMFS should work fine

- Has no problem with Squid 3.x

Discussion

- A. Hanuchevsky: was there a study on the impact of IPv6, apart from enabling the stack (network infrastructure, router replacement...)

- ◆ Dave: no. Difficult to do as it is very site dependent.

ALICE - C. Grigoras

All central services are IPv6-ready except the API service based on Xrootd

- All dual-stacked a while ago
- DNS load balancing includes both IPv4 and IPv6 addresses
- No problem found... but almost no IPv6 requests

Sites services: several known issues

- VOBOX perl http server used by CE, Packman and CMReport
 - ◆ Need to upgrade to perl 5.14+ to get native IPv6 support: also doable with a backport to 5.10
- Xrootd/EOS
- CASTOR
- ApMon
- Need to upgrade AliEN xrootd clients to v4 and v4 API
- 4 sites have deployed dual-stack VOBOX: more welcome

Monitoring: work to be done to integrate IPv6 network information in testing and network path calculation

Monitoring IPv6

perfSonar - D. Rand

perfSonar can run bwctl and owamp tests in IPv4 or IPv6, specified on a node by node basis.

- -v4 or -v6 suffix added to hostname

A Mash dashboard instance setup for IPv6 tests at Imperial College

- Asymmetry shown on several routes: to be investigated
- Allow a comparison with IPv4 results
- Open to new sites who had like to benefit from IPv6 testing with perfSonar

Nagios - M. Elias

Access to the service through http: ok

- Several plugins available for basic IPv6 testing

Livestatus + check_mk multisite: multisite not IPv6-ready in current version

Sensors

- Ping: by default selects IPv4 or IPv6. Would be good to run both
- NRPE: ok in SL6 (not in SL5)
- NSCA: no support for IPv6
 - ◆ Hopefully a future version will support it
- port checks: need to run both for IPv4 and IPv6
 - ◆ check_46 plugin can help

Other monitoring tools

- munin: no problem with IPv6
- Ganglia: basically works but a few configuration issues

Site Status and Plans

Site Survey - D. Kelsey

Run end of May

- See https://www.gridpp.ac.uk/wiki/2014_IPv6_WLCG_Site_Survey for results
 - ◆ Live table: sites should continue to answer
 - ◆ Considering opening a ticket against sites who will not answer after a reminder
- Got T0, 11/13 T1s, 85/155 T2s

T0 and T1s

- Only CERN foreseeing a lack of IPv4 addresses
- 7 T1 with partial IPv6 readiness, 4 without
 - ◆ 6 with plans for complete basic readiness in the next year

T2s

- 5 sites already with IPv4 address shortage
 - ◆ Private IPv4 WNs: not used at most sites
- 16 with full IPv6 connectivity, 9 with partial
 - ◆ 60 without IPv6 connectivity: 45 with no plans

UK IPv6 Status - C. Walker

Imperial

- All services except dCache dual stack
- Some IPv6 only node
- DPM and StoRM tested
- Failed attempt with OpenStack Havana

QMUL, Brunel, Oxford

- perfSonar + RIPE probe
- Brunel and Oxford: test DPM
- QMUL: test StoRM, in production soon

Main questions are operational ones: address allocation policy, dhcpv6 vs. SLAAC, firewall (ip6tables)

- In particular dhcpv6 doesn't configure router: how to do it? RA?

IPv6 at FZU (Prague) - M. Elias

IPv4 shortage: just one C class

- Started to investigate IPv6 in 2011

Testbed started to test current tools and administration process

- PXE: no IPv6 support

- New config management tools: Puppet
- Integrated with Nagios monitoring

Perfs: reached 21 Gb/s consolidated for DPM traffic

IPv6 at BITP (Urainia) - O. Shadura

See slides

Wrap-Up - What Next?

Next testing

- Continue file transfer testbed with more storage options
 - ◆ Involve all t1s?
 - ◆ Invite T2s?
- FTS3 IPv6 pilot: taken forward by FTS3 dev team, experiments and OpsCoord
- Atlas AMI
- IPv6-only WNs

Other issues

- Squid v3 issues preventing its use

Monitoring: important to set up the appropriate monitoring as we are moving to production

- perfSonar is an important piece and a good way to start for a site

Objective for next year: foster adoption of dual stack for production services at several/many sites

- WG could help to coordinate the move to avoid every site running into the same problems at the same time
 - ◆ In fact sharing of experience worked well so far...
 - ◆ Sites should start by setting up their IPv6 basic network infrastructure

This topic: LCG > GDBMeetingNotes20140610

Topic revision: r1 - 2014-07-10 - MichelJouvin



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

or Ideas, requests, problems regarding TWiki? use Discourse or Send feedback