

IPv6 Deployment at FZU in Prague

Jiří Chudoba, Marek Eliáš, Lukáš Fiala, Tomáš Kouba
elias@fzu.cz

HEPiX Spring 2014 Workshop, 19-23 May 2014

Institute of Physics AS CR, v. v. i. (FZU)

Outline:

1. production DPM on dual-stack
2. setup of WNs
3. setup of DPM head/disk nodes
4. monitoring setup
5. other activities in IPv6 WG

HEPiX IPv6 WG

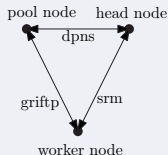
- ▶ we participate in the HEPiX IPv6 Working Group
- ▶ communication and information exchange helped us a lot

Currently on dual-stack:

- ▶ dpm headnode
- ▶ all production disk nodes
- ▶ all but 2 subclusters of worker nodes

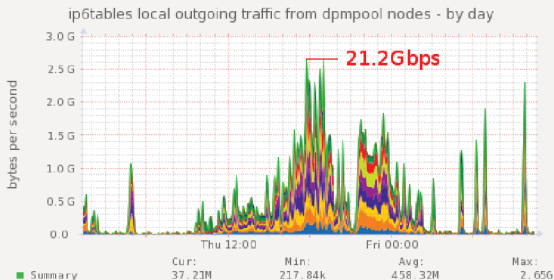
What goes over IPv6:

- ▶ dpns between disk nodes and head node
- ▶ srm between WNs and headnode
- ▶ actual data transfer via gridftp
- ▶ we tested webdav access: curl and aria2c

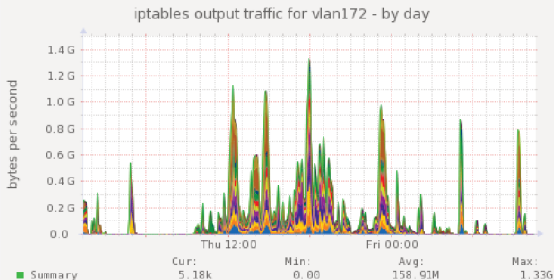


What does not:

- ▶ xrootd goes via IPv4; we expect new version will change this

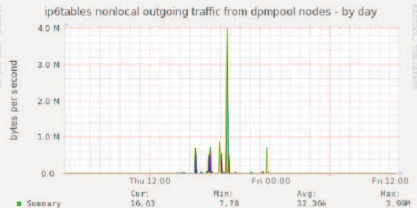
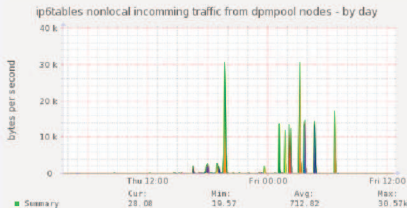


DPM → WNs:
IPv6 traffic

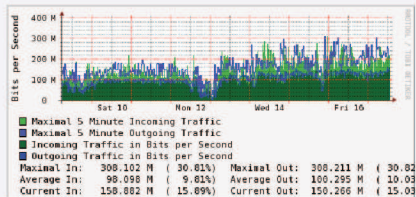


DPM → WNs:
IPv4 traffic

Just DPM disk nodes:



IPv6 traffic from our testbed:



ifcfg-ethX, sysctl.conf, dhclient.conf

- ▶ IPV6_* variables in ifcfg files do not work reliably
- ▶ sysctls are still needed to disable SLAAC and accept RA
- ▶ need to setup dhclient to use DUID-LL

Firewall ACLs

- ▶ anything out, nothing in

Conclusion:

- ▶ added IPv6 connectivity did not break anything
- ▶ added AAAA and IPv6 PTR records did not break anything
⇒ we started setting IPv6 for whole subclusters

Our steps:

- ▶ firstly one dual-stack disk node: did not break anything
- ▶ dual-stack head node
- ▶ adding another disk node was then complicated:
 - ▶ add AAAA and PTR records for the disk node
 - ▶ clear the nsd host cache on the head node
 - ▶ add IPv6 address to the disk node

Probably the most easy way:

- ▶ dualstack all the disk nodes first: add IPv6 address and AAAA and PTR records
- ▶ clear the nsd host cache on the head node (otherwise head node may refuse to respond dpns queries from the disk node)
- ▶ add IPv6 address and AAAA + PTR records to the head node

Other issues:

- ▶ SL5: we don't like dhcp6c
 - ▶ no possibility to choose the DUID type
 - ▶ not tested enough in our testbed

⇒ we use static IPv6 address setup on SL5 head node and disk nodes
- ▶ SL6: problems with getaddrinfo preference, custom `gai.conf` should be installed — see article by Francesco Prezl in <http://hepixon-ipv6.web.cern.ch/knowledge-base>

- ▶ using lcg-cp, the actual data transfer used just IPv4
- ▶ more steps needed to enable IPv6 GridFTP transfers

Updated system

- ▶ SL6.5
- ▶ lcg-utils 1.6
- ▶ gfal 1.6

Environment variables

- ▶ `GLOBUS_IO_IPV6=TRUE; GLOBUS_FTP_CLIENT_IPV6=TRUE`
- ▶ advised by Duncan Rand from Imperial College
- ▶ we simply put this to `/etc/profile` on all WNs
- ▶ did not break IPv4-only WNs

- ▶ DHCPv6, DUID-LL
- ▶ router RA setup: does not advertise prefix, just a route
⇒ to prevent clients to do SLAAC
- ▶ differences in IPv6 ACL on our C6500 router: no object-groups
- ▶ more details:
www.farm.particle.cz/twiki/bin/view/VS/VsIPv6Pub

IPv6 distribution in VLANs

- ▶ IPv4 production public C-class range: no IPv6, MTU=1500
- ▶ IPv4 production private network: public IPv6, MTU=9000
- ▶ special IPv6 network with MTU=1500: to connect production nodes not supporting jumboframes
- ▶ testbed network: IPv6 only, MTU=9000

Nagios sensors

- ▶ DHCPv6 lease sensor
- ▶ IPv6 default route sensor (routes are setup by RA and can expire)
- ▶ monitoring node currently IPv4-only
- ▶ ping6 sensor still missing

Network traffic monitoring

- ▶ our netflow monitoring not ready yet
- ▶ we work to replace flow-tools by nfdump+nfsen
- ▶ this requires switch from netflow v5 to v9
- ▶ ipt_NETFLOW kernel module already tested to report IPv6 traffic
- ▶ nowadays using just iptables munin plugin

Participation in activities inside the IPv6 WG

- ▶ GridFTP testing mesh of the IPv6 WG by Tony Wildish
- ▶ HEPiX IPv6 perfsonar mesh by UK sites (more by Chris Walker in the next talk)
- ▶ our testbed dpm head node monitored by the dual-stack SAM Nagios at Imperial

Grid certificates

- ▶ testing accessibility of CRLs of CAs from lcg-CA bundle
- ▶ 25 out of 101 CRLs are accessible via IPv6
- ▶ Beijing meeting 2012: 21
- ▶ Vancouver meeting 2011: 6
- ▶ <http://www.particle.cz/farm/admin/IPv6EuGridPMACrlChecker/>

Thank You

Marek Eliáš

elias@fzu.cz

<http://www.farm.particle.cz>

Work partially supported by CESNET, z. s. p. o.
project number 482/2013.

