

IPv6 dual-stack deployment for the distributed computing center

Kotliar¹, V. Kotliar^{1*}NRC "Kurchatov institute HEP, RU-142281, Protvino, Moscow region, Russia E-mail: {Anna.Kotliar Viktor.Kotliar}@ihep.ru

Corresponding author



Computing Center of the Institute for High Energy Physics in Protvino provides computing and storage resources for various HEP experiments (Atlas, CMS, Alice, LHCb) and currently operates more than 150 working nodes with around 3000 cores and provides near 2PB of disk space. All resources are connected through two 10Gb/s links to LHCONE and other research networks. IHEP computing center has IPv4 address space limited to one C-sized and all working nodes are installed behind the NAT which has some drawbacks for production use. To optimize routing, switching and to get higher network throughput for data transfer the IPv6 dual-stack deployment was made for the computing farm.

In this work the full cycle of the real IPv6 dual-stack deployment from zero to production will be shown. This work can be used by other WLCG centers and all other data centers for distributed computing as an example of necessary steps and configurations which have to be made.



Introduction

- 3044 CPU, 26875 HEP-SPEC06;
- 2183 TB;
- 2x10Gb/s Internet channels IPv6 dual stack;







Current status: IHEP external network





Current status: IHEP internal network



Current status: IHEP external network2



2016

2021

2016->2021

- IPv6 added (change site configuration for compute)
- more connections for RU-LHCONE
- No major changes in resources (storage, compute)





IPv6 introduction

- IPv4 addresses are sized at 32-bits and expressed in decimal octets:
 - 192.0.2.1 (address)
 - 192.0.2.1/255.255.255.0 (address and netmask)
 - 192.0.2.1/24 (address and prefix length)
- IPv6 addresses are sized at 128-bits and expressed in hexadecimal words:
 - 2001:0db8:0000:0000:0000:0000:0000:0001 (address)
 - 2001:0db8:0000:0000:0000:0000:0001/64 (address and prefix length)
 - 2001:0db8::1/64 ٠
- IPv6 Router Advertisements
 - IPv6 routers can be configured to send advertisements, both periodically • and on request, with the following information:
 - The IPv6 prefix/prefix length in use on this link •
 - The IPv6 address of the router •
 - Various 'flags' (or hints) which tell hosts how to behave
- IPv6 Dynamic Address Assignment
 - Hosts can auto-generate their own addresses using SLAAC (State-Less Auto-Address Configuration)
 - Or via DHCPv6 through use of the 'A' or 'M' flags.
- DNS

[tez@tetris] ~]\$ host host.example.wlcg

host.example.wlcg has address 2001:db8::1

[tez@tetris] ~]\$ host 2001:db8::1

namepointer host.example.wlcg. 05.09.2021



IPv6 introduction

- Device Unique IDentifier [DUID]
 - DHCPv6 does not necessarily identify machines by MAC address
 - DUIDs can be one of
 - Link-layer address plus time [DUID-LLT]
 - Vendor-assigned UID based on Enterprise Number [DUID-EN]
 - Link-layer address [DUID-LL]UUID-based DUID [DUID-UUID]
 - Not all DHCPv6 clients use the same DUID type by default
- Packet Fragmentation
 - IPv4 supports packet fragmentation
 - An intermediate router can break a large IPv4 packet into several smaller ones prior to forwarding
 - IPv6 does not support packet fragmentation
 - This is not permitted in IPv6; so, a router drops the packet and responds to the sender with ICMPv6 Message Too Big along with the size of packet that it will accept and forward
 - The sender receives the response and transmits a smaller packet
 - Repeat ad-infinitum along the traffic path until packet
 - successfully reaches the destination
 - Packet fragmentation hides broken network paths



IPv6 why at IHEP

- There is no IPv4 space left (PA too expensive)
- Use of NAT with IPv4 reduces performance (NAT gateways)
- IPv6 restores end-to-end host connectivity to all
 - Allow to growth computing cluster
- Some benefits for big data transfers (major traffic for IHEP)
- Possibility to change IP addresses assignments for GPN
- Restriction from WLCG to connect storages



Step by step IHEP IPv6

атус:	Открыта		Начата:	06.02.2013	
v6 поддержка	в ИФВЭ				
Связанные за	адачи				
связана с clc - Зад Cluster Network external ipmlemen	д <u>ачи #1145</u> : Internet nt IPv6	Закрыта	03.10.2016	06.02.2019	
связана с ЛПОИТ <u>#1178</u> : IPv6 бумаж	[:] - Бумажки кка от НИЦ КИ	Закрыта	23.11.2016	02.12.2016	
связана с ЛПОИТ <u>#1677</u> : network е. 194.190.160.0/21 / "РосНИИРОС"	- <mark>Бумажки</mark> xternal AHO	Закрыта	10.07.2018	06.08.2018	
связана с ЛПОИТ network IPv6 tes	⁻ - <u>Задачи #1803</u> : sting setup	Закрыта	26.11.2018	24.12.2018	
связана с ЛПОИТ - <u>Задачи #1762</u> : network external&internal IPv4 IPv6 IPAM system		Открыта	19.10.2018		
связана с ЛПОИТ network IPv6 ne	⁻ - Задачи #1831 : tflow v9 collector	Закрыта	24.12.2018	17.01.2019	
связана с clc - <u>За</u> Cluster Network ipmlement IPv6 Задачи #1865: clu internal add dhcp	gayu #1852: Internat internal Internat network Internation of the state of th	Закрыта	30.01.2019	01.03.2019	
связана с ЛПОИТ network internal peering CORE1-C0	- <u>Задачи #1857</u> : setup IPv6 bgp ORE2	Закрыта	06.02.2019	06.02.2019	
связана с ЛПОИТ network external IPv6 add IPv6 mr	- <u>Задачи #1883</u> : monitoring tg g	Закрыта	15.03.2019	15.03.2019	



Step by step IHEP IPv6: 0 preparations

- 1. You need to understand a concrete reason for YOUR IPv6 setup
- 2. Check hardware stack for IPv6 support (usually no problem here)
- 3. Check used OS support (usually no problem here)
- 4. Check application support (depends on software)
 - System applications
 - DNS, mail, DHCP, NTP, netflow...
 - Used application software
 - Everything will be affected (decide default IP protocol for concrete software if dual-stack)
- 5. Check possible security issues (how you will firewall IPv6)
- 6. Decide IPv6 address range to use (usually IPv4 /24 -> IPv6/64)
 - Think about addressing plan depends on your network topology
 - Many different recommendations (VLAN, IPv4, usage) IPv6 Address Planning Tom Coffeen(O`Reilly)

Step by step IHEP IPv6: 0 preparations



Figure A-4. Intrasite (/48 to /64) Hierarchy Planner

Site	Prefix:	200	1:678	:7d	8::/	/48]				200	1:6	78:	7d8	:21a5	::179/6 [,]
Site	CIDR	Subnet	Description	Bernent	CIDR	Subnet	Description	Dement	CIDR	Subnet	Description	Element	CIDR	Subnet	Description	
	/52	:0000::/52	reserved		/56	:1000::/56	reserved		/60	:1100::/60	reserved		/64	:1110::/64	reserved	
	/52	:1000::/52			/56	:1100::/56	Serve	ər	/60	:1110::/60			/64	:1111::/64	<u> </u>	
	/52	:2000:752	GRIL	/	/56	:1200::/56		ŦI.	/60	:1120::/60			/64	:1112:/64		
	/52	:4000::/52			/56	:1400::/56		11	/60	:1140:/60			/64	:1114:/64	<u> </u>	
	/52	:5000::/52			/56	:1500::/56		11	/60	:1150::/60			/64	:1115::/64	<u> </u>	
	/52	:6000::/52			/56	:1600::/56		11	/60	:1160:: 60		ih L	/64	:1116::/64	D_{VI}	dt hyta
(40	/52	:7000::/52		160	/56	:1700::/56		150	/60	:1170::/00			/64	:1117::/64	V4 10	ຊເ ມັງເອ
/48	/52	:8000::/52		/62	/56	:1800::/56		/00	/60	:1180::/60		/00	/64	:1118::/64		
	/52	:9000::/52			/56	:1900::/56			/60	:1190::/60			/64	:1119::/64		
	/52	:a000::/52			/56	:1a00::/56		1	/60	:11a0::/60			/64	:111a::/64		
	/52	:b000::/52			/56	:1b00::/56			/60	:11b0::/60			/64	:111b::/64		
	/52	:0000::/52			/56	:1c00::/56			/60	:11c0::/60			/64	:111c::/64		
	/52	:d000::/52			/56	:1d00::/56		1	/60	:11d0::/60			/64	:111d::/64		
	/52	:e000::/52			/56	:1e00::/56			/60	:11e0::/60			/64	:111e::/64		
	/52	:1000::/52	reserved		/56	:1100::/56	reserved		/60	:11f0::/60	reserved		/64	:111f::/64	reserved	



Step by step IHEP IPv6: 1 one computer setup

1. Get IPv6 address space from your provider or LIR (PI recommended)

inet6num: 2001:678:7d8::/48 country: RU

- 2. Setup you internal network infrastructure and connect one host to IPv6 network
 - ping -6 should be ok internally between routers and IPv6 testing computer
- 3. Setup external routing for IPv6 (ripe&providers)

route6:	2001:678:7d8::/48
origin:	AS2643
descr:	IHEP-IPv6

- ping -6 should be ok externally between testing computer and Internet inside->outside and outside->inside
- 4. Setup DNS to work direct and back zones

- nserver: ns.ihep.su
- nserver: ns.psn.ru
- mnt-by: AS2643-MNT
- ping -6 by name should be ok outside->inside

Step by step IHEP IPv6: 2 monitoring& accounting + firewall

- 1. Setup IPv6 accounting (netflow)
 - Flow DB (posgresql)
 - Flow collector (nfacctd for netflow v9)
 - IPv6 flow client tools
 - Adapt something already used (flox)
 - Add new tools (pmacct plug-in for cacti)
 - Setup export for netflow v9 on routers
- 2. Add or adapt links monitoring to see IPv6 traffic on interfaces
- 3. Add IPv6 peers for monitoring
- 4. Add IPv6 basic firewall rules to border firewalls



Step by step IHEP IPv6: 3 add distributed cluster

Cluster: storages, compute nodes, servers (dual stack statefull IPv6 configuration)

- 1. Storages
 - add IPv6 for static configuration on servers through cluster management system
 - add DNS records
 - Configure storage software to work with IPv4 and IPv6(prefer)
- 2. Compute
 - Add IPv6 DHCP server for MAC->IPv6 binding (not a part of IPv6 standard)
 - Full IP control
 - Easy to use already used IPv4 dhcp rules by MAC
 - Configure network topology to use IPv6 direct connection from nodes to outside (IPv4 works behind NAT)
 - Configure system and software on nodes prefer IPv6 in dual stack
- 3. Management servers
 - Configure static IPv6 for dual stack
 - One by one configure all software to use IPv6 instead of IPv4 if possible
- 4. Configure firewall rules everywhere for new servers

Step by step IHEP IPv6: 4 add IHEP LAN

- 1. Should we use IPv6 only (with IPv6 to IPv4 gates) or dual stack
- 2. Decide how to use DHCP (state full or state less)
- 3. User can not remember IPv6 address
 - Need a tool for IPAM (IP Address Management) for IPv4&IPv6
- 4. Network complexity will increase -> more complex troubleshooting
- 5. Many different OS need to be considered (Windows 7,8,10...,Linux, IOS, Androids)
- 6. Many different switches MRV, HP, Cisco, WiFi routers, ADS Modems and so on
- 7. Some old software with no support

Still in progress.....



Current status: IHEP IPv6

- 1. External connectivity V
- 2. Computing centre use IPv6 dual-stack (only were needed) V
- 3. Still in progress for IHEP GPN (deploy were needed)
- 4. No plans to switch everything to IPv6 only or everything to IPv4&IPv6 (still use only IPv4 devices)





Thank you!

Any questions?