IPv6 deployment at CERN

ISGC, Taipei, 16th March 2016

edoardo.martelli@cern.ch

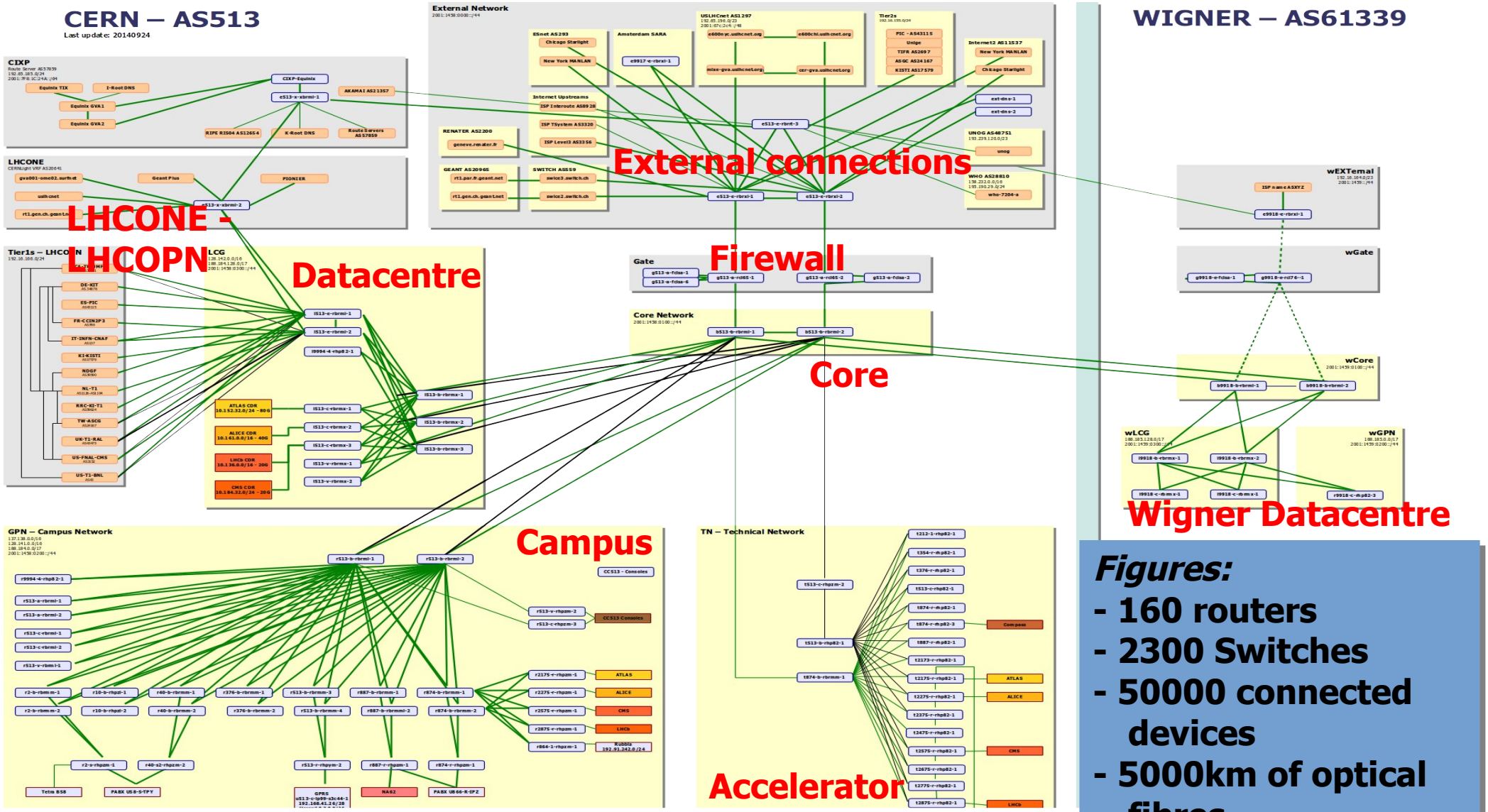
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Agenda

- Introduction: the CERN network
- IPv6 project
- IPv6 deployment outcome
- Challenges and lessons learnt
- IPv4 depletion status and projections
- What's after depletion

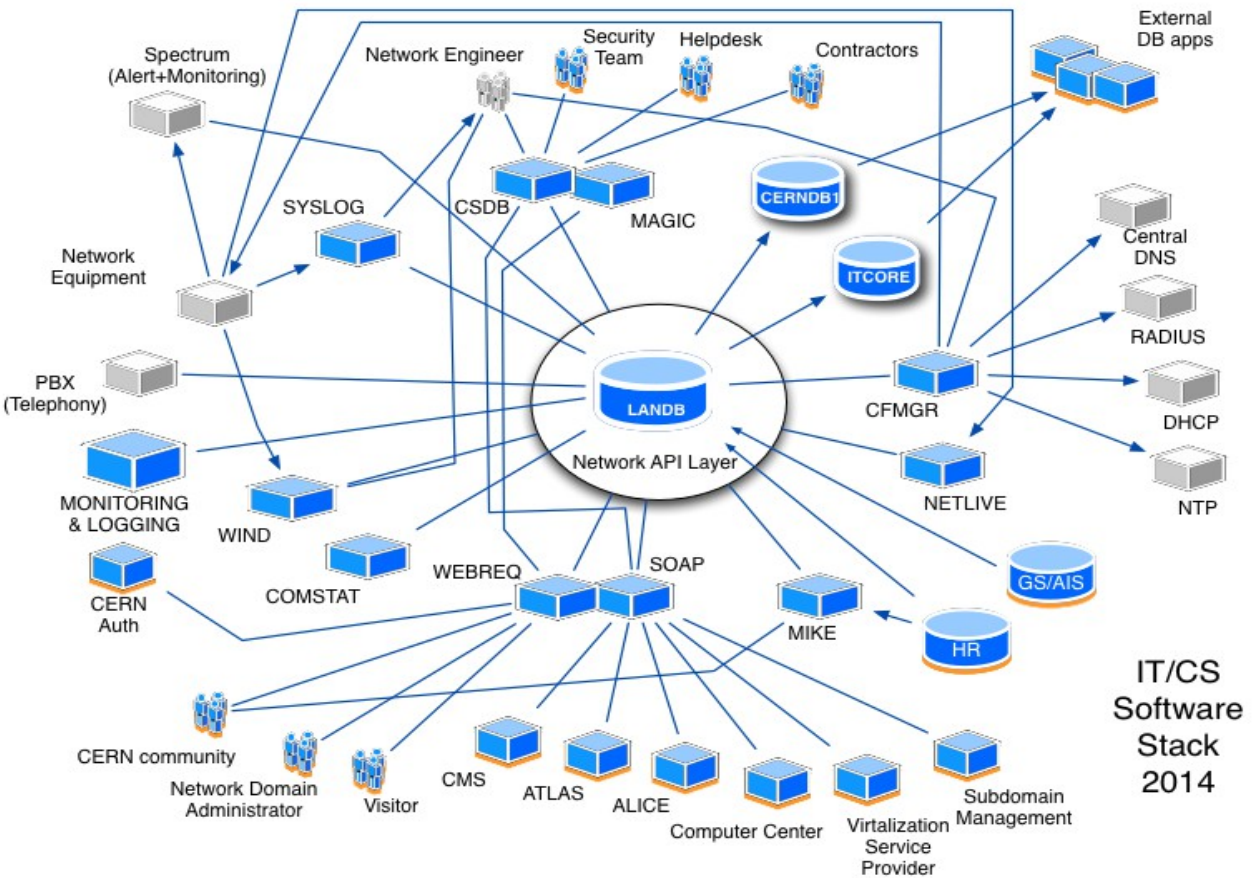
CERN Network

Network domains



Network Provisioning and Management System

- >250 Database tables
- ~200,000 Registered devices
- >1,000,000 lines of codes
- >15 years of development



IPv6 deployment project

Drivers

CERN started playing with IPv6 in 2001, but for many years there was no reason for deploying it on al large scale

Main IPv6 driver:

- Large **Virtual Machines** deployments ramped up in 2010
- It was soon planned to have 130,000 VMs with public IP addresses for LHC physics analyses by 2014

Approval and resources

IPv6 deployment approved by IT management in 2011

Allocated resources:

- *For network design/testing/deployment:*
- 1x Network Engineer FTE for 2 years.
- *For network database and NMS applications:*

  2x Software Developers FTE for 2 years

Initial IPv6 service definition

- Dual Stack configuration
- Every device can be dual-stack (assign at least one IPv6 address for every assigned IPv4 address)
- Identical performance as IPv4, no penalties
- Common provisioning tools (NMS) for IPv4 and IPv6
- Same network services portfolio as IPv4 (DNS, DHCP, NTP, Radius)
- Common security policies for IPv4 and IPv6

Initial workplan

- Testing IPv6 support of existing network devices
- Design and development of Network-DB schema
- Population of IPv6 records of Network-DB
- Development of the NMS tools
- Configuration of network devices
- Network services (DNS, DHCPv6, Radius, NTP)
- Network-DB Web interface for end-users
- Training for Support Lines and Service Managers

To be ready for production in 2013

The IPv6 service today

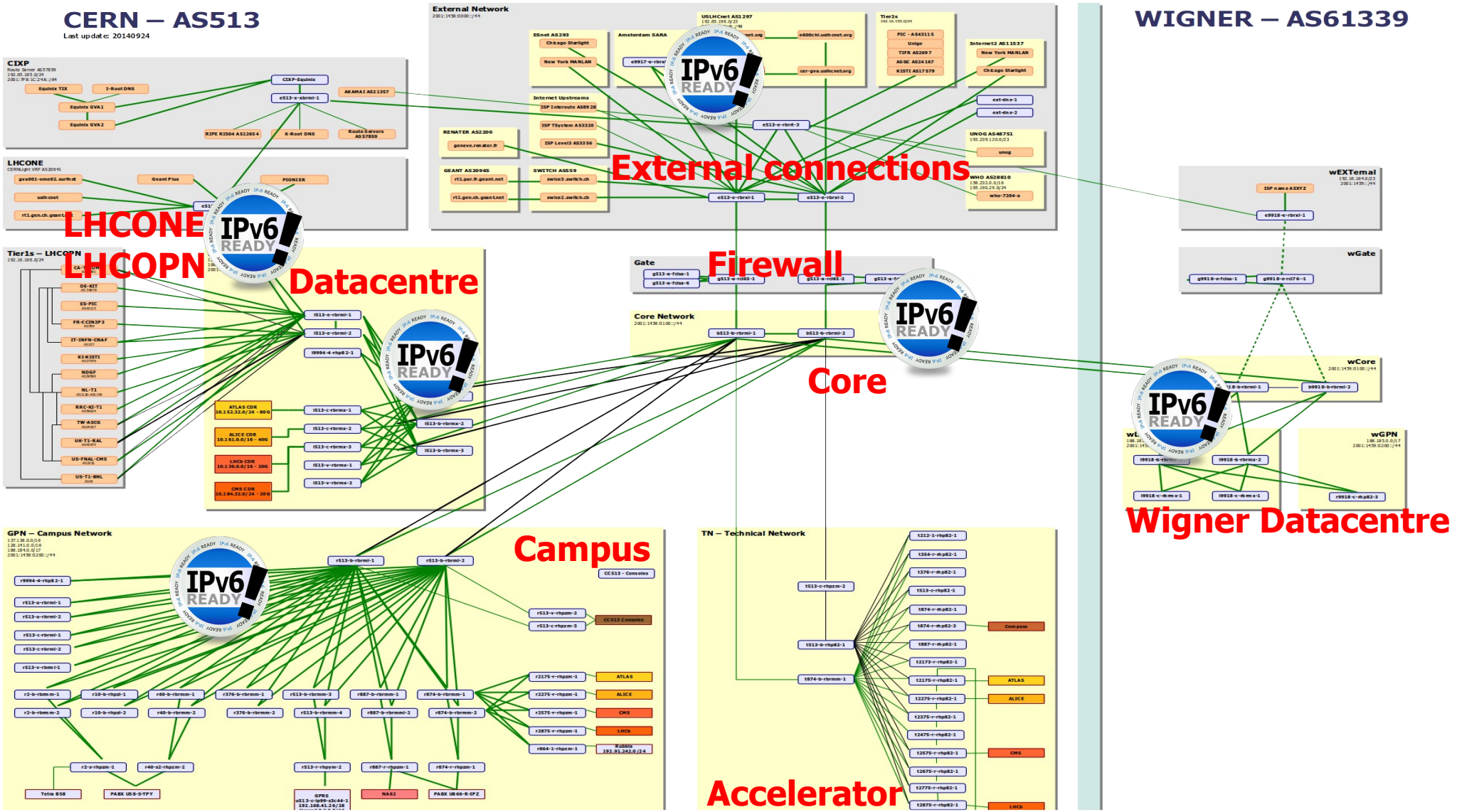
Dual stack network

- Dual stack configuration of all routers and switches in the domains Campus, DataCentre (Geneva and Wigner), Firewall, External, LHCOPN/ONE

- Domains not done because of legacy equipment and protocols: LHC accelerator control network, LHC detectors data acquisition networks

- Same routing architecture (BGP and OSPF)

Dual stack domains



Dual stack network database

- IPv6 now main navigation key (ready to drop IPv4)
- IPv6 records added beside every IPv4 record
- New schema compatible with all legacy queries (no need to rewrite all the applications)
- IPv6 address tables fully populated

Every device can connect dual-stack

- Every device with an IPv4 address has an IPv6 address assigned in the Network DB

- All assigned IPv6 addresses have a name in **ipv6.cern.ch**

```
# host ping.ipv6.cern.ch
ping.ipv6.cern.ch has IPv6 address 2001:1458:201:1c80::100:175

# host TELEPHONE-62470.ipv6.cern.ch
TELEPHONE-62470.ipv6.cern.ch has IPv6 address
fd01:1458:204:27a::100:2e
```

- Dynamic (portable) devices get a name in **dyndns6.cern.ch**

```
# host myiphone.dyndns6.cern.ch
myiphone.dyndns6.cern.ch has IPv6 address
2001:1458:202:180::101:8a26
```

Line rate performance

All production network devices can forward IPv6 packets at wire speed. No penalties to IPv6 adopters

Only exception: policy base routing for statefull firewall bypass (not implemented yet because of low traffic volume)

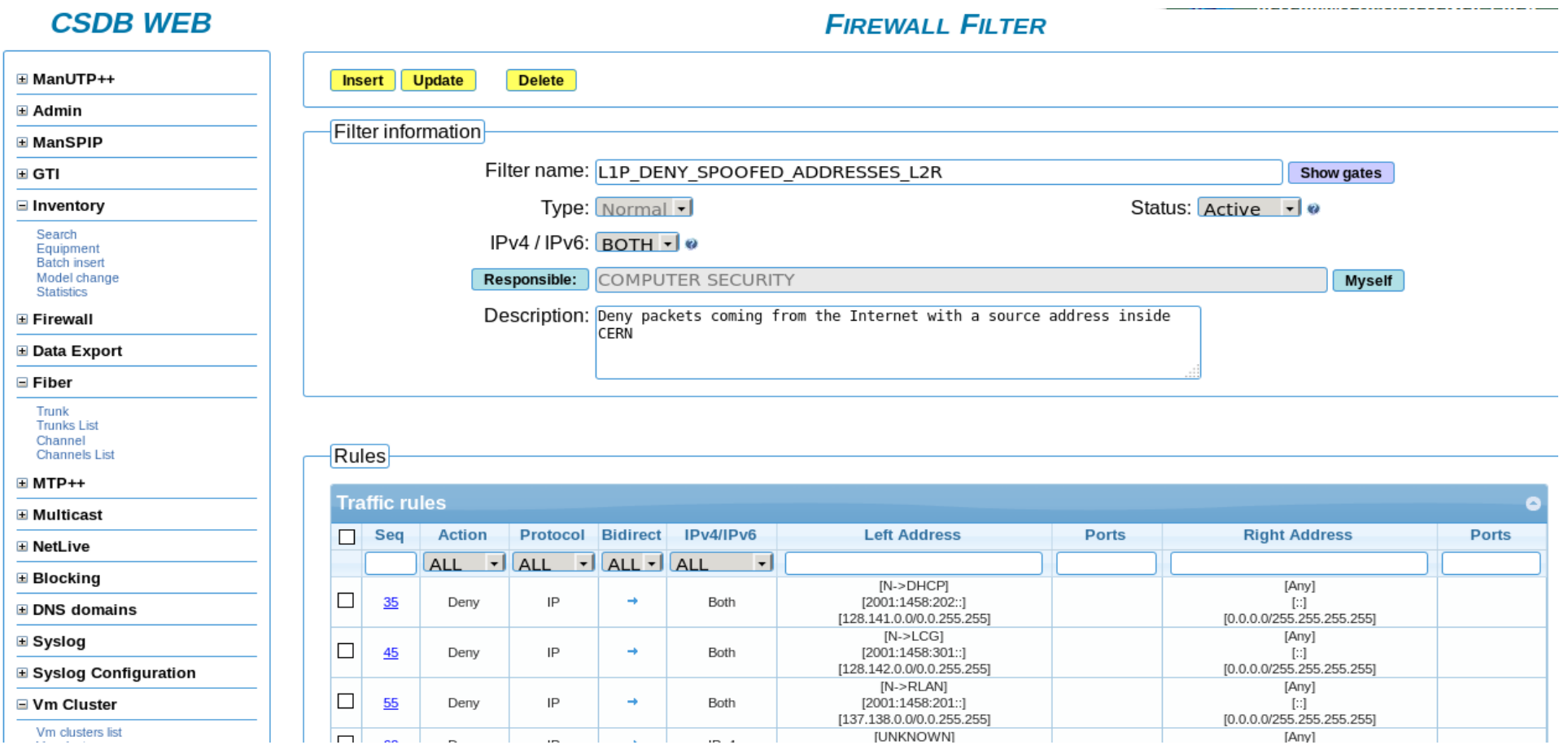
Dual-stack provisioning tools

NMS:
- routers' configuration generators for all the vendors
- DHCPv6 and DNS configurations from Network-DB
- ACLs for firewalls generated from Network-DB

CSDBweb (Network-DB interface for engineers):
- IPv6 everywhere there is IPv4

WebReq (Network-DB interface for end-users):
- All IPv6 info visible together with IPv4, IPv6-ready flag settable

CSDBweb (engineering)



Webreq (end-users)

Device Information

- **Device Name:** **RIPE-ATLAS-PROBE** [Last Operation]
- Location: 0031 S-0012
- Manufacturer: UNKNOWN
- Model/Type: UNKNOWN
- Generic Type: UNKNOWN
- Description: RIPE MEASUREMENT PROBE
- Tag:
- Serial Number:
- Operating System: UNKNOWN Version: UNKNOWN
- CERN Inventory number:
- Network Interface Card(s): 00-20-4A-C8-24-98/ETHER-AUTO-10/100
- Responsible for the device: MARTELLI EDOARDO IT CS EDOARDO.MARTELLI@CERN.CH / Tlf: 72613
- Main User of the device: MARTELLI EDOARDO IT CS EDOARDO.MARTELLI@CERN.CH / Tlf: 72613
- HCP Response: This system **CAN** obtain an IP address automatically [more info]
- IPv6 Ready: This system **IS NOT** IPv6 ready
- Last changed: 21-02-2014 (15:51)

Interface(s) Information

>>Network Service HELP<< >>Network Interface Card(s) HELP<<

Interface Name	IP Address	Service Name	Internet Connectivity
RIPE-ATLAS-PROBE.CERN.CH	137.138.32.177 2001:1458:201:b459::100:3f	S31-S-IP3	Y

Subnet IPv4 Mask: 255.255.255.192 **Default IPv4 Gateway:** 137.138.32.129	**Name IPv4 Servers:** 137.138.16.5, 137.138.17.5 **Time IPv4 Servers:** 137.138.16.69, 137.138.17.69
Subnet IPv6 Netmask: 64 **Default IPv6 Gateway:** 2001:1458:201:b459::1	**Name IPv6 Servers:** 2001:1458:201:1000::5, 2001:1458:201:1100::5 **Time IPv6 Servers:** 2001:1458:201:1040::69, 2001:1458:201:1140::69
IP Aliases: NONE	
Bound Interface Card(s): NONE	

Users can control IPv6 behavior

Users can declare their own devices as “IPv6-ready”

IPv6-ready means:

- IPv6 connectivity is OK
- all running server applications are listening on both v4 and v6 sockets

Consequences in the network:

- Firewall: IPv6 equivalent of existing IPv4 security openings applied to the central firewall
- DNS: DEVICENAME.cern.ch returns A and AAAA records, reverse relsolution returns DEVICENAME.cern.ch (and host certificates can work properly)

Same network services as IPv4

DNS:
- direct and reverse resolution of all assigned addresses
- servers can be queried over IPv6
- announced in the DHCPv6 leases

NTP:
- reachable over IPv6

DHCPv6:
- Static and Dynamic assignments based on the MAC address of the requestor

"dual-stack" security policies

Firewall rules database

- IPv6 policies equivalent of all existing IPv4 policies
- IPv6 specific options supported (e.g. ICMPv6)
- IPv6 only policies created

Firewall management software

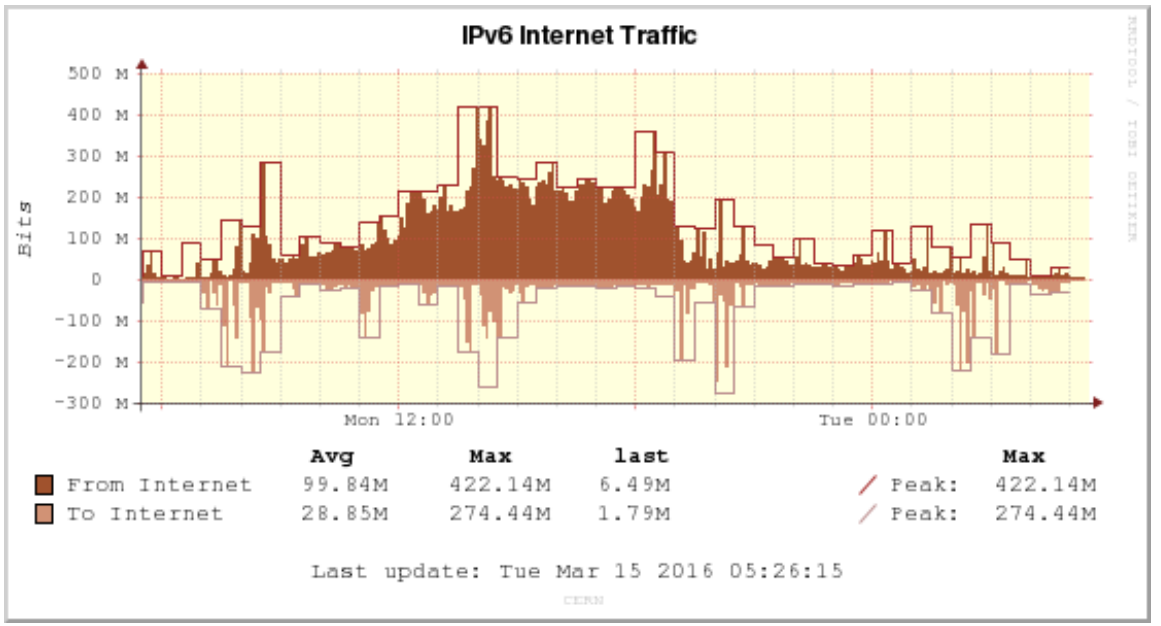
- All firewalls managed by the CERN NMS

IPv6 on a normal day

DHCPv6 active leases: 5000 avg, 10000 peak (55% of DHCPv4)

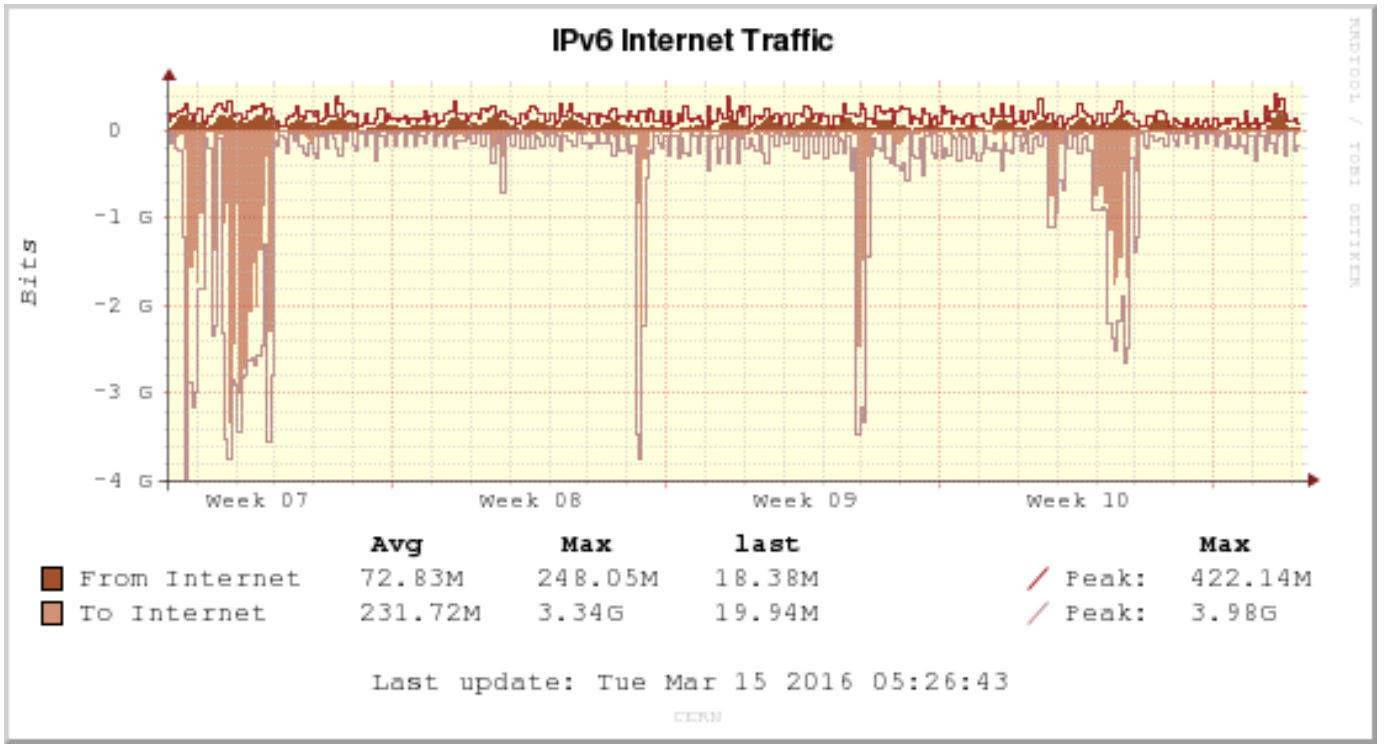
DNS queries over IPv6: 210,000/hour (4% of queries over IPv4)

Internet traffic: 5% of ISP traffic



Growing IPv6 traffic

More and more LHC data transfers happens over IPv6



Project Timeline – early stages

2001: CERN IPv6 testing started

2003, June: public IPv6 prefix assigned to CERN

2003, September: IPv6 deployed in the CERN External Network: CERN prefix announce to NRENs. Direct and Reverse DNS over IPv6.

2003, November: IPv6 Land Speed record in collaboration with Caltech

2009, November: CERN IPv6 prefix visible in the whole IPv6 Internet.

Project Timeline – 2011

2011, January: IPv6 deployment project approved

2011, February: IPv6 address plan issued

2011, March: Development LANDB (Network-DB) schema includes IPv6 information.

2011, July: IPv6 connectivity in part of LCG, CORE and GPN backbones (Brocade routers)

2011, July: Prototype of DNS servers

2011, August: Pilot IPv6 services for LCG and GPN users

Project Timeline – 2012

2012, March: LANDB with IPv6 tables in production

2012, March: CSDWEB (Users LANDB web interface) support of IPv6 information

2012, March: training of Operation and Deployment teams about new CSDB (engineering LANDB web interface)

2012, July: CSDB supports IPv6 for deployment of new network connections

2012, October: cfmgr Brocade and HP routers configuration compilers can generate IPv6 configurations

Project Timeline – 2013

2013, March: all routers in the datacentre of Building 513 support IPv6 for end-users

2013, March: WEBREQ support of IPv6 information (not dispayed to end-users yet)

2013, April: DHCPv6 for static devices

2013, April: All LCG datacentre routers have dual-stack services

2013, June: NTP service ready: ip-time-1.ipv6.cern.ch and ip-time-2.ipv6.cern.ch

2013, September: DHCPv6 for portable devices

Project Timeline – 2013 cont.

2013, September: DNS replies over IPv6 from ip-dns-1.ipv6.cern.ch and ip-dns-2.ipv6.cern.ch

2013, October: Firewallmanagement software completed (LANDB schema and translation of existing IPv4 rules, CSDBWEB, WEBREQ, cfmgr gate update).

2013, October: DNS automatically configured from LANDB information

2013, November: All Campus routers have dual-stack services

2013, November: LANDB IPv6 information available from SOAP interface

2013, November: WEBREQ shows IPv6 information to any user

Project Timeline – 2014

2014, January: Automatic IPv6 configuration in the central firewall for IPv6-ready flagged devices

2014, January: Dynamically leased addresses published in dyndns6.cern.ch

2014, February: IPv6-ready flag fully functional (DNS and Firewall)

2014, February: Training for IT Service desk

2014, February: DHCPv6 leases to any device in the IT buildings

2014, April: DHCPv6 leases to any device in the IT datacentre

Project Timeline – 2014 cont.

2014, May: DHCPv6 leases to any registered device connected to a portable socket or WIFI

2014, May 8th: dual-stack lxplus instance available at lxplus-ipv6.cern.ch

2014, May 12th: imap, pop, smtp, ldap services dual stack

2014, June 3rd: DHCPv6 leases to any static device in GPN; DHCPv6 deployment completed.

All major milestones completed

Challenges and lessons learnt

Benefits

Simplified management of addresses
- one subnet size fits all (/64)
- no-brainer address planning for new deployments
- reduced risk of future renumbering

Future proof
(hopefully)

Challenges

- Size of routing tables and ACLs have doubled in number of entries and quadrupled in memory utilization
- New problems to be solved by Support lines
- DHCPv6 still in an early stage
- New security threats to take into account
- Legacy applications don't understand IPv6, and some will never do

Challenges: DHCPv6

Rationale for using DHCPv6: Network-DB driven address assignment for automatic configuration of DNS and firewall, user traceability, light access control

Drawbacks

- RAs necessary for default-gateway and mask-length: thus two protocols to maintain and control, no predictive load-balancing for multi-router subnets, all available prefixes exposed

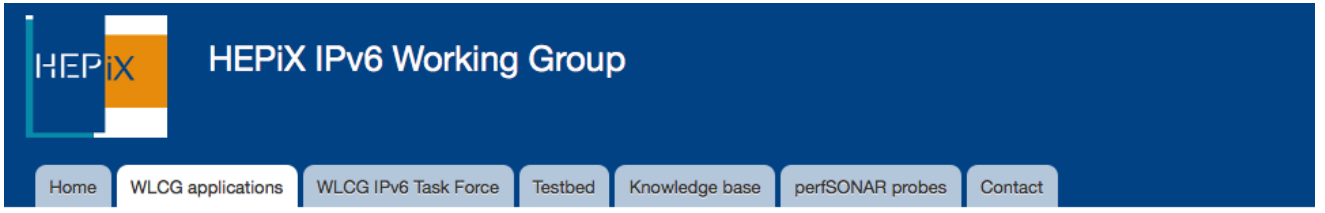
- MAC address authentication not always works: DHCPv6 clients don't have to use the MAC address of the interface they send the request via. Waiting for implementation of RFC6939 to fix it. DUID management is not an option

- Not all devices use DHCPv6 by default (iOS up to v6, old MacOS/Linux/Windows versions, industrial devices…). Android doesn’t support DHCPv6 clients

Challenges: Compliance of applications

Physics Community actively reviewing IPv6 compliance of its applications

The HEPiX IPv6 Working group is pushing developers to code IPv6 support and correct bugs in WLCG applications



Home

IPv6 compliance of WLCG applications

Software Component	Type	Used by Experiment	Version	IPv6 Compliance
AliEN	LHC Experiment Application	ALICE		
ARC CE	Middleware	ATLAS, CMS		YES
ARGUS	Middleware	ALICE, ATLAS, CMS, LHCb		Unknown
BDII	Middleware	ATLAS, CMS, LHCb	EMI 2	YES
BestMAN	Middleware	ATLAS, CMS		YES
CASTOR	Middleware	ALICE, ATLAS, CMS, LHCb		NO
cfengine	Monitoring			Unknown
CMS Tag Collector	LHC Experiment Application	CMS		Unknown

http://hepix-ipv6.web.cern.ch/wlcg-applications

Lessons learnt

Catching up with 20 years of IPv4 experience and development takes a lot of time

The configuration of the network is the easy part. Address management is what took most of the time

DHCPv6 is definitely not like DHCP (v4)

Don't rush. Have a staged deployment with a large variety of early adopters. Poke them: they may not report all the problems

Don’t trust lab tests: only the deployment on the live network will prove it can cope with the two protocols.

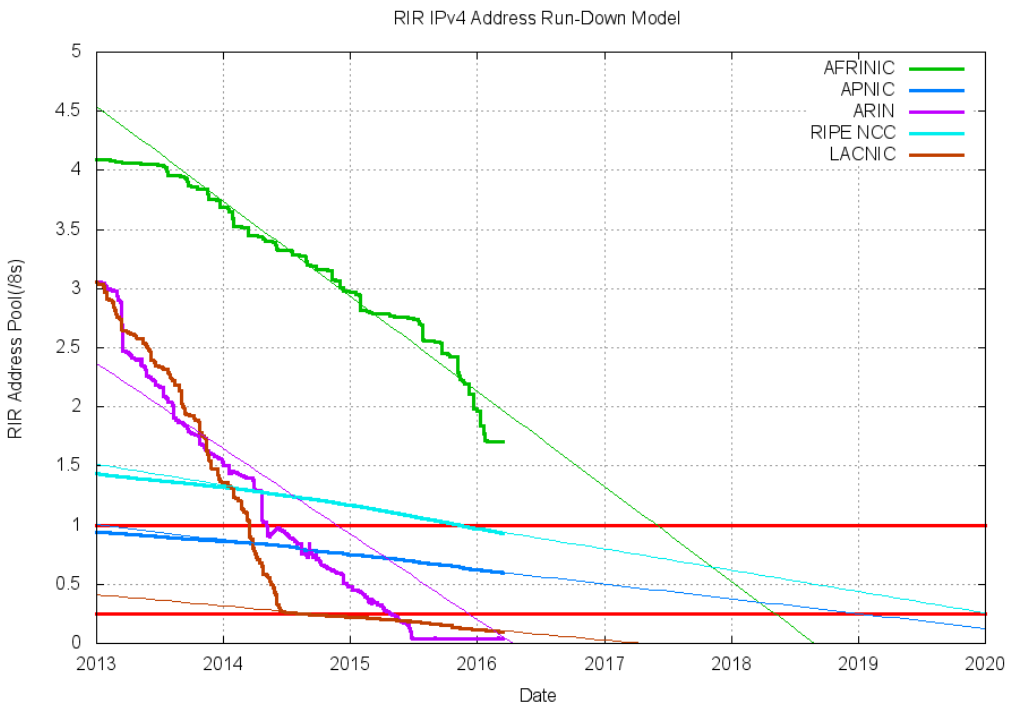
Don't assume applications developers will like IPv6: they have already enough bugs to fix without adding those of IPv6.

IPv4 depletion: status and projections

World IPv4 pools' status

Region	Exhaustion date	Remaining /8 (16M)
Asia-Pacific	19-Apr 2011 (last /8)	0.5981
Europe	14-Sep-2012 (last /8)	0.9298
North America	24-Sep-2015	0
South America	10-Jun-2014 (last /8)	0.0981
Africa	1-May-2018	1.7003

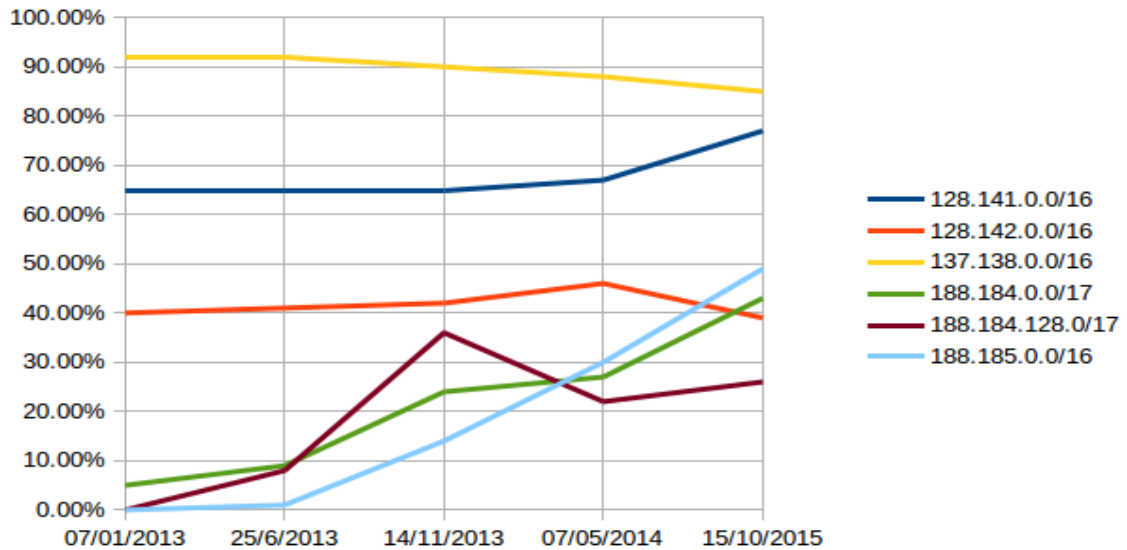
[15th March 2016]



http://www.potaroo.net/tools/ipv4/index.html

CERN IPv4 pools' status

Campus dynamics	75% used
Campus statics	85% used
GPN Data Centre	43% used
LCG servers	39% used
LCG VMs	26% used
Wigner data centre	49% used



[15th of October 2015]

Projections for Campus

- In use today: ~99000
- Plan to free ~28700 with ongoing optimizations
- 16384 (/18) addresses will be added for new WIFI controller
- ~7000 new addresses assigned in the last 18 months

=> Enough addresses for the next 4-6 years

Projections for Geneva datacentre

- ~10000 addresses allocated in the last 12 months

- 16384 (/18) Wigner IP addresses moved to Geneva for new DC infrastructure

- ~72000 addresses left.

=> remaining IPv4 addresses can 5-6 years, unless a new datacentre becomes necessary

Projections for Wigner datacentre

- ~16000 addresses allocated in the last 18 months

- 25% (16384) of Wigner IP addresses moved to Geneva for new DC infrastructure

- ~17000 addresses left.

=> remaining IPv4 addresses may last 12-24 months, probably just enough for all the servers that can be hosted there

What's after IPv4 exhaustion

NAT

***NAT** has been avoided at CERN because it is a performance bottleneck, may cause problems to special applications, may be difficult to track*

Anyway, in case of shortage:

- WIFI users can be easily NATed
- Campus desktops can be NATed too

IPv6-only

Use of IPv6-only for global connectivity is possible: NAT64 allows IPv6-only clients to reach IPv4-only servers with the help of tweaked DNS servers

However:

- NAT64 still doesn't work with special and legacy applications
- Very few WLCG sites have started deploying IPv6 at production level: they will have problems to reach IPv6-only CERN servers

Worst case scenario

IPv4 public addresses assigned only to servers that must be visible outside CERN. All the rest can be NATed

Conclusions

Conclusions

- IPv6 deployment in a large organization requires time and investments

- Deployment in production networks didn't cause any major issue

- When finally available, users adopt IPv6

- Plan in advance for IPv4 address exhaustion

Questions, comments, experiences?


CERN