

# IPv6

## Ein Überblick



**Kars Ohrenberg**

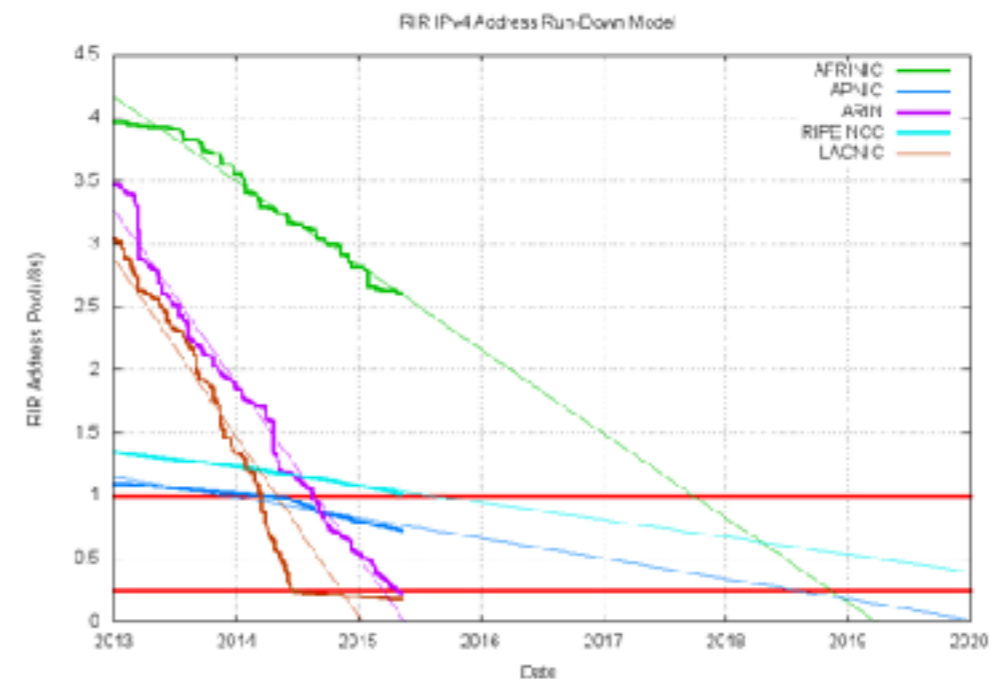
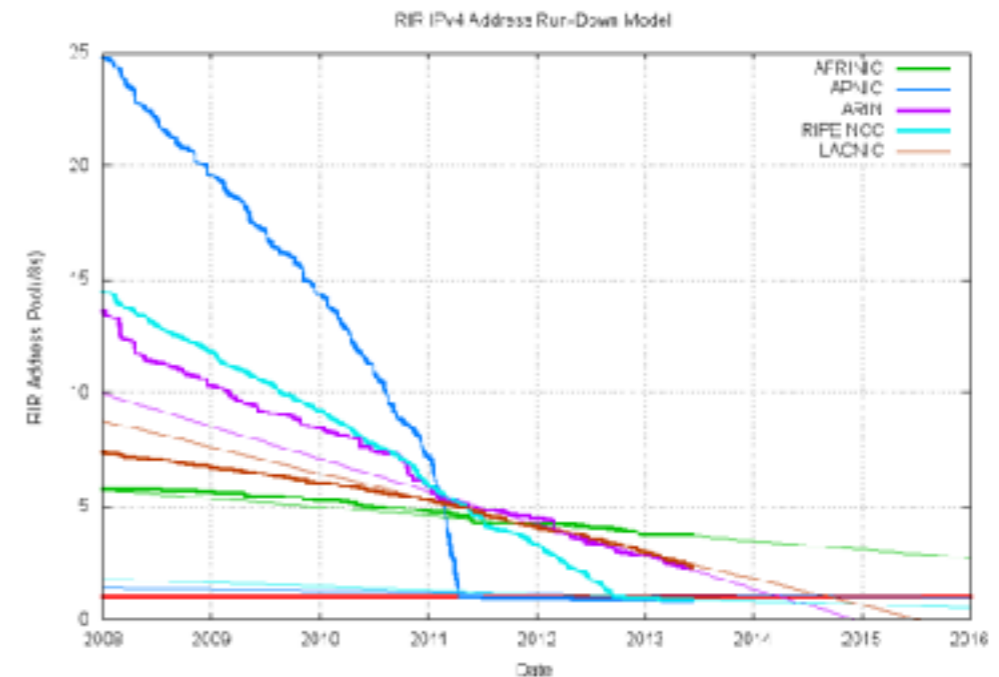
IT Seminar

Hamburg, 12.05.2015

# IPv4 Address Pool Exhaustion

> Am 3.2.2011 wurden die letzten IPv4 Adressen an die RIRs (Regional Internet Registry) vergeben

- Die letzten 5 freien /8-Bereiche wurde dabei an RIPE NCC, ARIN, APNIC, LACNIC und AfriNIC vergeben
- Am 15. April 2011 teilte APNIC die letzten „frei“ zu vergebenden Adressen für die Region Südostasien zu
- Am 14. September 2012 folgte dann RIPE NCC für die Region Europa/Naher Osten
- Am 23. April 2014 dann auch ARIN für die Region Nordamerika



# Weitere Gründe für IPv6

## > Grundsätzliche Optimierungen am Protokoll

- Effizientere IP Header (z. B. 64-Bit Alignment)
- Zustandslose, automatische Konfiguration von IPv6-Adressen
- Multicast
- Mobile IP
- Renumbering und Multihoming
- ...

## > Reduktion der Routing-Tabellen im Internet

- Die vollständige v4-Routing-Tabelle des Internet besitzt z. Zt. ca. 550.000 Einträge

## > Kein NAT mehr benötigt

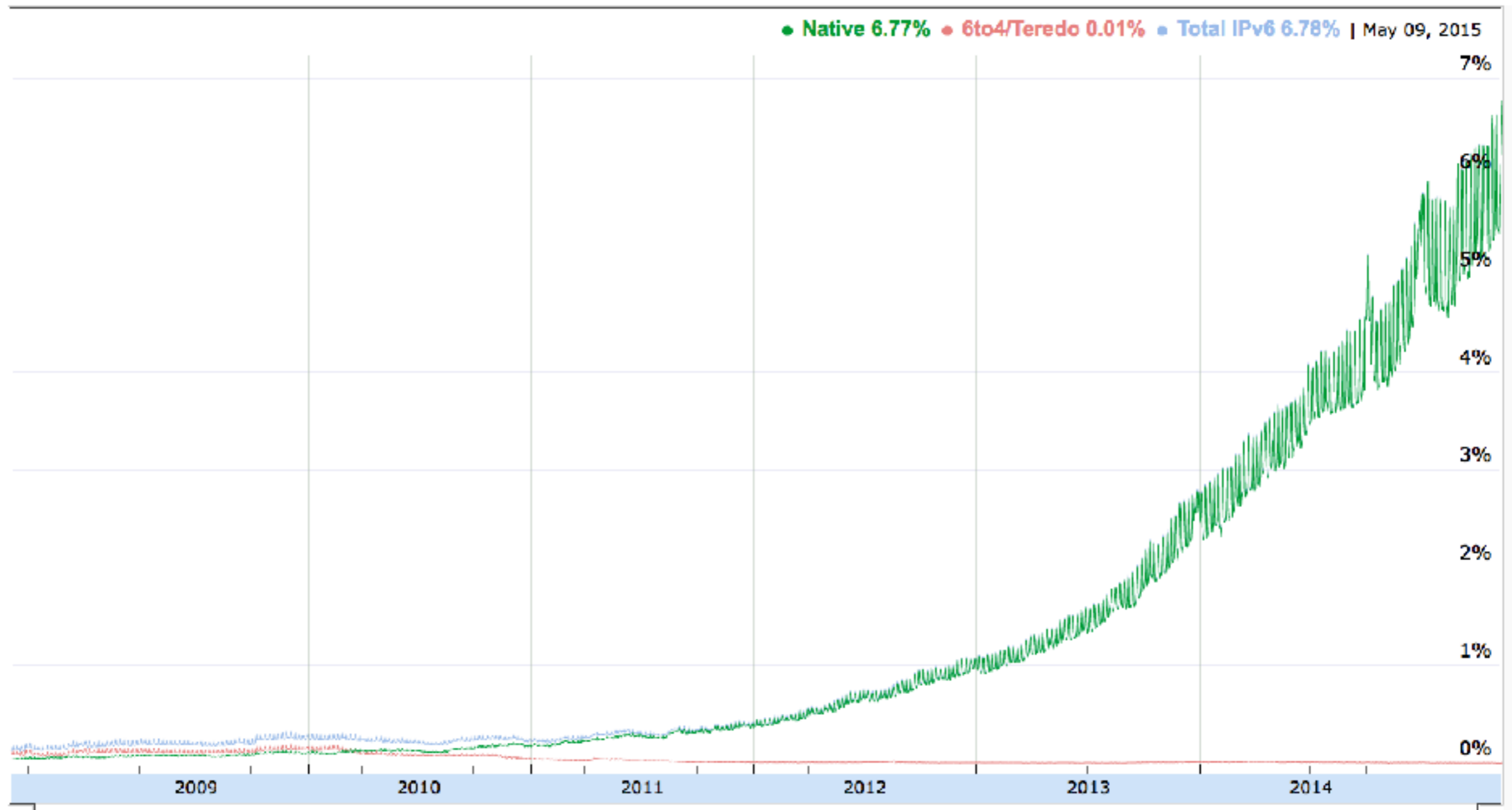


# Geschichte von IPv6

- > 1992: Erste Vorhersagen über die Erschöpfung des IPv4-Adressenraums (zwischen 2005 – 2011)
- > 1995: Erste Spezifikation im RFC 1883
- > 1996: Start des 6bone
- > 1998: IPv6 wird mit RFC 2460 zum Nachfolger von IPv4
- > 1999: RIRs vergeben IPv6-Präfixe
- > 2000: Die wichtigsten Hersteller fangen an IPv6 in ihren Produkten zu berücksichtigen
- > 2001: DESY bekommt IPv6-Präfix (19. Februar)  
(2001:638:700::/48)



# Google IPv6 Usage Statistics



# Der IPv6-Adressraum

- > IPv4-Adresse ist 32 Bit lang
- > IPv6-Adressen sind 128 Bit (16 Bytes) lang
  - Es sind nicht 4-mal soviel Adressen sondern es ist die 4-fache Adresslänge und somit  $2^{96}$ -fache Vergrößerung des Adressbereiches
  - 340 282 366 920 938 463 463 374 607 431 768 211 456 adressierbare Geräte ( $3,4 \cdot 10^{38}$ )
  - $\sim 5 \cdot 10^{29}$  Adressen für jeden auf der Erde lebenden Menschen
  - ‚Das genügt, um jeden Quadratmillimeter der Erdoberfläche inklusive Ozeanen mit rund 600 Billionen Adressen zu bepflastern‘ (c't 3/2007)
  - Hätte man gleichzeitig mit der Entstehung der Erde begonnen mit einer Rate von 1 Milliarde Adressen pro Sekunde IPv6 Adressen aus dem vorhandenen Adressraum zu verteilen, wäre bis heute nur weniger als ein Billionstel des vorhandenen Adressraums vergeben.
- > Üblich ist die Aufteilung in 64-Bit-Präfix für die Netzadresse und der verbleibenden 64-Bit für den Interface Identifier (Host-Adressierung)
  - Das gesamte bisherige IPv4-Internet hätte problemlos in einem Subnetz Platz



# Adressnotation (RFC 4291)

## > Hexadezimale Notation

- Zahl wird in 8 Blöcke zu je 16 Bit (4 Hex-Ziffern) unterteilt, Trennung durch Doppelpunkte

> 2001:0db8:85a3:08d3:1319:8a2e:0370:7344

## > Führende Nullen innerhalb eines Blocks dürfen entfallen

- 2001:0db8:0000:08d3:0000:8a2e:0070:7344 ->

2001:db8:0:8d3:0:8a2e:70:7344

## > Aufeinander folgende 0er-Blöcke dürfen ausgelassen und durch Doppelpunkt ersetzt werden

- 2001:0db8:0:0:0:0:1428:57ab -> 2001:db8::1428:57ab

- Diese Regel darf nur einmalig angewendet werden!

## > Für die letzten vier Bytes der Adresse darf die herkömmliche dezimale Notation verwendet werden

- ::ffff:7f00:1 kann als ::ffff:127.0.0.1 geschrieben werden



# Adresstypen - Unspecified, Loopback

## > Unspecified: 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 oder ::

- Darf nicht zugewiesen werden, zeigt das Fehlen einer Adresse an
  - Eingesetzt z. B. für initialen DHCP Request oder Duplicate Address Detection
- Ähnlich der 0.0.0.0 in IPv4

## > Loopback: 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 : 1 oder ::1

- Adresse des eigenen Standortes, in der Regel mit localhost verknüpft
- Zur Überprüfung des eigenen IPv6-Stacks: `ping6 ::1`
- Wie 127.0.0.1 in IPv4





# Adresstypen - Scoped Address (neu in v6)

## > Link Local Unicast

- Werden von Routern nicht weitergeleitet und sind daher nur im gleichen Teilnetz erreichbar
- Wird automatisch für jedes Interface konfiguriert (basierend auf der MAC-Adresse)
- Format: `FE80:0:0:0:<interface identifizier>`

## > Unique Local Unicast

- Für private Adressen gibt es die Unique Local Addresses (ULA), RFC 4193
  - Site Local Unicast (`fec0::/10`) sind veraltet (RFC 3879)
- Ähnlich den privaten IPv4-Adressen
- Format: `FC00::/7`, `FC<40-Bit-Site-ID>:<16-Bit-Subnet>::/64`
  - `fd00:/8` für lokal generierte ULA
  - `fc00:/8` für global zugewiesene eindeutige ULA



# Adresstypen - Multicast

- > Kein Broadcast mehr in IPv6, stattdessen Multicast
- > Format: `FF00::/8`
  - `FF<Flag><Scope>::<Multicast Group>`
  - 4 Bits für Flag
    - 0: Permanent definierte, bekannte Multicast-Adressen (von IANA zugewiesen)
  - 4 Bits für Scope
    - 1: interfacelokal, diese Pakete verlassen die Schnittstelle nie. (Loopback)
    - 2: link-lokal, werden von Routern grundsätzlich nie weitergeleitet und können deshalb das Teilnetz nicht verlassen.
    - 5: sitelokal, dürfen zwar geroutet werden, jedoch nicht von Border-Routern.
    - e: globaler Multicast
- > All nodes on local link: „`ping6 FF02::1`“ (Broadcast)
- > All router on local link: „`ping6 FF02::2`“



## > IPv6 Neighbor Solicitation

- Sent by a node to determine link-layer address of a neighbor
- =~ ARP request
- Packet description
  - Source Address = link-local address
  - Destination = solicited-node mulitcast address
  - Data contains link-layer address of source
  - Query is: give me your link-layer address?

## > IPv6 Neighbor Advertisement

- Response to a Neighbor Solicitation
- =~ ARP response
- Includes my MAC address so you can send me information
- Packet description
  - Source address = link-local address of source
  - Destination = destination address
  - Data contains linke-layer address of me

# Zusammenfassung IPv6 Adressierung

## > IPv6

- hat einen viel größeren Adressraum
- hat spezielle Adressformatierung
- besitzt einen neuen Adressierungsmechanismus (Scoped Address)

## > IPv6-Knoten kann eine Vielzahl von IPv6-Adressen besitzen und hat die jeweils richtige zu wählen

```
[mac-book-kars] ~ $ ifconfig
lo0: flags=8049<UP,LOOPBACK,RUNNING,MULTICAST> mtu 16384
    inet6 ::1 prefixlen 128
    inet6 fe80::1%lo0 prefixlen 64 scopeid 0x1
    inet 127.0.0.1 netmask 0xff000000
gif0: flags=8010<POINTOPOINT,MULTICAST> mtu 1280
stf0: flags=0<> mtu 1280
en0: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
    ether c4:2c:03:39:42:91
    inet6 fe80::c62c:3ff:fe39:4291%en0 prefixlen 64 scopeid 0x4
    inet6 2001:638:700:20b9::1:b9 prefixlen 64
    inet 131.169.185.185 netmask 0xfffff00 broadcast 131.169.185.255
    media: autoselect (1000baseT <full-duplex>)
    status: active
```



# IPv6 - Adressstruktur am DESY

- > 2001:638:700:1028::1:10c8
- > 48-Bit DESY-Präfix (wie 131.169.0.0/16 in IPv6)
  - 2001::/16 IANA (Adress allocation started on July 1999, to providers only)
  - /23 to RIRs
  - /35 to ISPs (if /35 exhausts get /29) -> give /48 to Organizations
- > 16 Bit DESY Structure, davon 4 Bit für DESY Network + 12 Bit für Vlan-ID
  - 2001:638:700:0xxx::/52 -> Backbone
  - 2001:638:700:1xxx::/52 -> RZ, eg. 2001:638:700:1028::1:c8
  - 2001:638:700:2xxx::/52 -> Office
    - 3xxx - dxxx -> weitere Wolken, z.B. XFEL
  - 2001:638:700:exxx::/52 -> Outside
  - 2001:638:700:fxxx::/52 -> Zeuthen



# IPv6 - Adressstruktur - Beispiel

> 2001:638:700:1028::/64 -> RZ + Vlan 40  
(131.169.40.0/24)

- # public

> 2001:638:700:137F::/64 -> RZ + Vlan 895  
(192.168.86.0/24)

- # private, IPv6 geroutet und IPv4 nicht geroutet

> FD00:638:700:1028::/64 -> RZ + Vlan 40  
(131.169.40.0/24)

- # private, nicht geroutet, sowohl IPv6 als auch IPv4

> Dieses Schema gilt auch, wenn man IPv6 only Netze oder 10.x.x.x Netze einrichtet



# IPv6 - Adressstruktur - Host Adressen

> 64-Bit Interface Identifier („Host Address“)

> 2001:638:700:1028::1:10c8

- `::/112` reserviert für NOC (`x:x:x:x:x:x:x:0000 - x:x:x:x:x:x:x:ffff`)
- Hosts (Dual-Stack) ab `x:x:x:x:x:x:xxx1:0000`
- Hosts (IPv6-only) ab `x:x:x:x:x:x:xxx2:0000`

> Default Gateway stets `x:x:x:x:x:x:x:00ff` bzw. `fe80::x:x:x:x`



# Beispiele für Host Adressen

> 131.169.40.200 wird zu 2001:638:700:1028::1:c8

- Nicht für NOC reservierte, öffentliche IPv4-Adresse im RZ
- 40 (Dezimal) = 28 (Hex), 200 (Dezimal) = c8 (Hex)

> 131.169.185.185 wird zu 2001:638:700:20b9::1:b9

- Nicht für NOC reservierte, öffentliche IPv4-Adresse im Office

> 192.168.184.185 wird zu

2001:638:700:34b9::1:b9 o. fd00:638:700:34b9::1:b9

- private vlan 1209, nicht für NOC reserviert, Adresse im XDAQ

> 131.169.197.5 wird zu 2001:638:700:00c5::0:5

- öffentliche NOC-Adresse im Backbone





## > Neue Record-Typen

- AAAA (RFC 3596)
  - Equivalent zu IPv4 A-Record
- A6 (obsolet mit RFC 3363)
- PTR Record (reverse lookup)
  - Selber Record wie bei IPv4
  - Neue TL-Domain für den IPv6-Raum: war ip6.int, ip6.arpa hat sich durchgesetzt (IPv4 in-addr.arpa)
  - Möglicher neuer Record-Typ: DNAME

## > DESY Nameserver unterstützen seit langem AAAA Records und haben selber eine IPv6-Adresse

- milky.desy.de (AAAA) -> 2001:638:700:1028::1:c8
- way.desy.de (AAAA) -> 2001:638:700:10c2::1:c8



# Dual Stack

- > Um einen einfachen Übergang zu ermöglichen erfolgt in der Regel ein Parallelbetrieb von IPv6 und IPv4, d.h. IPv6 wird hinzugeschaltet ohne IPv4 abzuschalten
- > Die v6-Adresse wird gegenüber der v4-Adresse bevorzugt
  - RFC 3484, „Default Address Selection for Internet Protocol version 6“



# Adresszuweisung - SLAAC

- > Stateless Address Autoconfiguration (SLAAC), RFC 2462
- > Ist grundsätzlich aktiviert, Host empfängt auf dem Local Link die sogenannten Router Advertisement (RA)
  - Enthalten den Präfix, den Default Router, Lebensdauer etc.
- > Host generiert die IP-Adresse aus dem RA und seiner EUI-64 Interface Adresse (IPv6 over Ethernet)
  - Ableitung einer EUI-64 Interface Adresse aus einer 48-Bit MAC-Adresse:
    - MAC = 00:90:27:17:fc:0f
    - Adresse wird ‚geteilt‘ und ff:fe eingefügt => 00:90:27:ff:fe:17:fc:0f
    - Zweites Bit des ersten Byte wird auf 1 gesetzt => 02:90:27:FF:FE:17:FC:0F
  - Verwendbarkeit der Adresse wird mittels Duplicate Address Detection geprüft
- > Keine weiteren Optionen (DNS, Bootfiles, ...) möglich



# Adresszuweisung - DHCPv6

- > Stateful DHCPv6, RFC 3315
- > Ähnlich wie DHCP für IPv4
  - Erlaubt die Übergabe weitere Parameter
  - Hostidentifizierung nicht mehr über MAC-Adresse sondern nur noch über den (auch in DHCPv4 möglichen) DHCPv6 Unique Identifier (DUID)
  - Default Gateway kann nicht übermittelt werden, muss über RA erfolgen
- > Ist gegenüber SLAAC zu bevorzugen da eine Protokollierung der Adressvergabe erfolgen kann.



# DHCPv6 am DESY

- > Es gibt generelle Probleme im Zusammenspiel von SLAAC, DHCP und DNS
  - SLAAC enthält keine Infos über DNS-Server etc.
  - DHCPv6 kann kein Default Gateway übermitteln
- > Lösungen:
  - Setzen des ‚Managed-Flags‘ in Router-Advertisements erzwingt Nachfrage bei DHCPv6 Server nach weiteren Konfigurationen
  - Übergabe des Default Gateways über die Router Advertisements
  - Durch ‚Unterdrückung‘ des Präfixes in den Router Advertisements stehen diese nicht für die Autokonfiguration zur Verfügung
- > DHCPv6 ....
  - DUID teilweise nur schwer zu ermitteln
  - ISC DHCP-Server kann IPv6-Adressen auch basierend auf MAC-Adresse vergeben, QIP folgt ...



# Was geht?

> Switching ✓

> Routing ✓

- LAN ✓

- Aktiv auf einzelnen Subnetzen, wird nach Bedarf und Abspreche erweitert werden

- WAN ✓

- Aktiv seit 22.3.2011

> DNS ✓

> DHCP (✓)

- Für Tests anfragen an NOC

> Applications ...



# HEPiX IPv6 Working Group

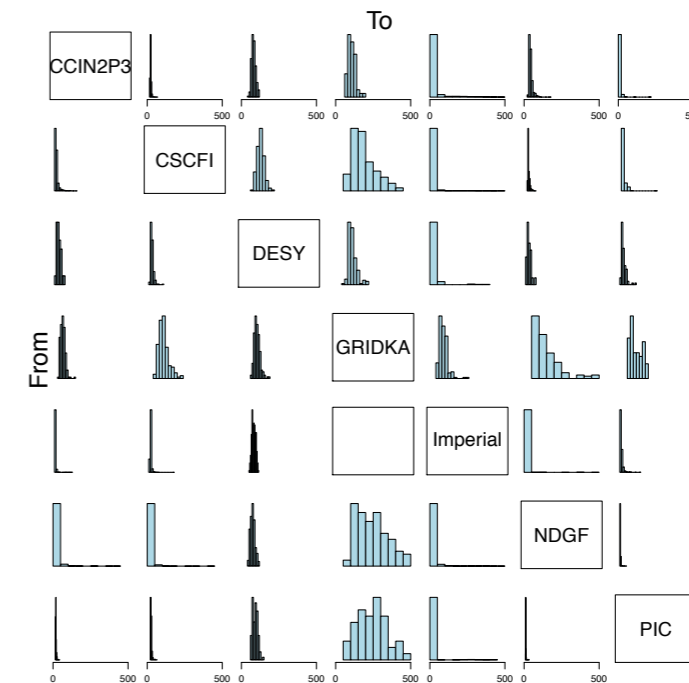
- > The working group was created in April 2011 (<http://hepixon-ipv6.web.cern.ch>)
- > IPv6 compliance of WLCG applications
- > HEPiX IPv6 Distributed Testbed for FTS3 Data Transfers
- > First Tier-2 run production dual-stack
  - Imperial College London, University of Nebraska
- > First Tier-2 will run production dual-stack soon
  - NDGF, PIC



Home

## IPv6 compliance of WLCG applications

Software Component	Type	Used by Experiment	Version	IPv6 Compliance
AliEn	LHC Experiment Application	ALICE		
ARC GE	Middleware	ATLAS, CMS		YES
ARCUR	Middleware	ALICE, ATLAS, CMS, LHCb		Unknown
BOI	Middleware	ATLAS, CMS, LHCb	EM 2	YES
DETMAN	Middleware	ATLAS, CMS		YES
CASTOR	Middleware	ALICE, ATLAS, CMS, LHCb		NO
elengw	Monitoring			Unknown
CMS Big Collector	LHC Experiment Application	CMS		Unknown
CMSST	LHC Experiment Application	CMS		Unknown
cmsweb	LHC Experiment Application	CMS		YES
ORNL 2	LHC Experiment Application	CMS		Unknown



# Status und nächste Schritte

- > Statisch konfigurierte IPv6-Adressen können zum Testen sofort eingesetzt werden
- > DHCPv6 über DUID verfügbar, über MAC-Adresse hoffentlich in den nächsten Tagen
- > Aktivierung von IPv6 für alle eduroam-Clients am DESY
- > IPv6 im LHCONE z. Zt. noch nicht aktiv, Gespräche mit DFN in den nächsten Wochen

