Intro to IPv6

NANOG 78

Alex Latzko Lead Network Architect
Server Central Turing Group

Target Audience (that's you!)

Are you comfortable with IPv4 and the concepts of routing and CIDR notation?

- Yes, you're in the right place
- No, you might get lost.

Why we're here

Not an agenda, but we're going here.

- A bit of IPv6's timeline
- Addressing concepts
- Differences between IPv4 and IPv6
- Some philosophical issues

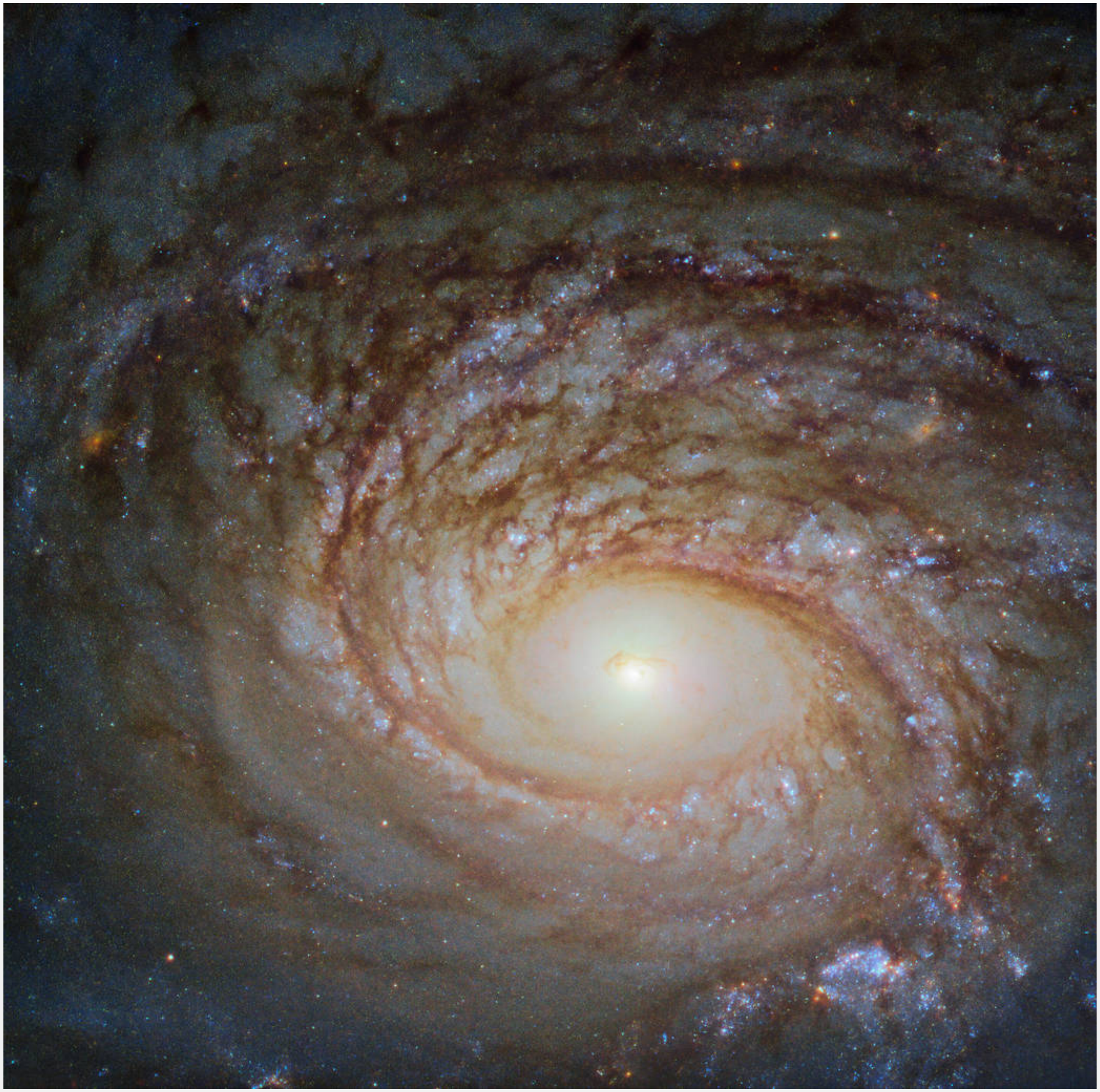
Prelude

- It’s probably the most used network protocol?
- IPv4 addresses are normally displayed as a dotted quad

255.255.255.255

What is IPv6 and what isn't it?

- IPv6 is the next generation transit protocol, and it's new and wonderful!
  - The first "real" IPv6 RFC was published in December 1998
  - The first public blocks were allocated to APNIC and ARIN on 1 July 1999.
- Big features
  - Designed with security in mind
  - Implemented in hardware it's just as fast if not faster than IPv4.
  - The extensible nature of IPv6 headers can do interesting things
  - There are several migration aids including delivering IPv6 datagrams encapsulated in IPv4 datagrams
- The biggest features is yet to come…



The biggest feature.

- Every IPv6 intro has some metaphor for the size of address space
  – The mass of the Earth is 5.972 × 10^27g. If you assume each IPv6 address weighs 1g, then the IPv6 address space weighs in at roughly 57 Billion Earths.
- IPv6 addresses are 128 bits long, so 2^128 is the total space in IPv6. In case you are curious, in base 10 that's

340,282,366,920,938,463,463,374,607,431,768,211,456

Aka ~3.403 * 10^{38} or 340.3 Undecillion

Format of an IPv6 address

All IPv6 addresses are represented as eight groups of four hex digits separated by colons. Here's an example:

2001:0db8:0000:0000:0000:0000:0000:0001

Most people will never have to type a numeric IPv6 address. DNS will do all the dirty work for them with a protocol specific address record, the AAAA (it's pronounced quad A)

CIDR notation in IPv6 is the same as IPv4, only with more numbers.

IPv6 address .gzip

2001:0db8:0000:0000:0000:0000:0000:0001

- To improve readability the alpha component of IPv6 addresses should be in lower case letters.
- There might be lots of zeros in an IPv6 address, and there are some rules about getting rid of them.
  - Leading zeros in a block not only can be suppressed, they MUST be suppressed.
  - Any one contiguous block of 0 can be replaced by a “::”.
    - The symbol “::” MUST be used to its maximum capacity.
    - The symbol “::” MUST NOT be used to replace just one 16-bit 0 field.
    - Use the symbol “::” to replace the longest consecutive run set of 16 bit 0 fields.

2001:db8::1 [2]

IPv4 syntax embedded in IPv6 addresses

- An alternative form that is sometimes more convenient when dealing with a mixed environment of IPv4 and IPv6 nodes is x:x:x:x:x:x:d.d.d.d, where the 'x's are the hexadecimal values of the six high-order 16-bit pieces of the address, and the 'd's are the decimal values of the four low-order 8-bit pieces of the address (standard IPv4 representation).

Examples:

0:0:0:0:0:0:13.1.68.3 (compressed ::13.1.68.3)

0:0:0:0:0:FFFF:129.144.52.38 (compressed ::ffff:129.144.53.38)

Embedded IPv4 addresses have several common uses.

Examples:

- As a migration aid when you are setting up dual stacking
- As the target when using a specific flavor of IPv6 inside IPv4 tunnel

IPv6 address types

There are three address types:

- Unicast: An identifier for a single interface,
  - each interface physical can have many unicast addresses.
- Anycast: An identifier for a set of interfaces.
  - Typically they identify different nodes.
  - The datagram is sent to the topologically nearest address (kind of like IPv4)
- Multicast: addresses also identify groups of interfaces,
  - Typically they belong to different nodes.
  - Traffic is sent to all accessible nodes in the multicast group.

A bit of a curve ball versus IPv4

- There are no broadcast addresses in IPv6.
- Their function is superseded by multicast.
  - There is no ARP in IPv6, the IP to MAC mapping is done via a totally different mechanism we'll cover in a few minutes.

IPv6 address ranges, a quick reference

- The high-order (left-most) bits of an IPv6 address are used to identify its type, as shown here:

Address type	Binary Prefix	Hex Prefix
Unspecified	0000…0 (128 bits)	::/128
Loopback	000..01 (128 bits)	::1/128
IPv4 Mapped	00..01111111111111111(96 bits)	::ffff/96
Multicast	11111111	ff00::/8
Link-Local Unicast	1111111010	fe80::/10
Unique Local Address	1111110	fc00::/7
Site Local Address	1111111011	fec0::/10
Global Unicast	Everything else	

- Anycast addresses are taken from the global unicast pool. Anycast and unicast addresses cannot be distinguished based on format.
- RFC for Multicast addressing: 2375/3306/3307/3180/6034 etc

Some things worth knowing

- You have been seeing the IPv6 address 2001:db8::/32 used extensively. It is a magic address reserved for writing documentation and you should never accept it.

- Three fundamental differences are
  - There is no broadcast in IPv6
  - Packet fragmentation can only occur at the logical ends of a routed path, and the minimum MTU is 1280
  - There is a magic non-routable address called “link local” used exclusively inside the local broadcast domain.
  - The address format is variable, so parsing it takes some effort.

EUI-64 An important building block

- EUI-64 is used as part of many IPv6 functions and for Ethernet. It's derived from the 48 bit MAC address burned into the hardware at the factory (or overridden by software).

- There is a magic conversion which makes the 48 bit MAC address into slightly modified EUI-64 format:
  - Split the 48 bit MAC and pad it with FF:FE out to 64 bits
  - Invert the 7th most significant bit from 0 to 1
    - For example MAC 00:0c:29:0c:47:d5 would become EUI-64 02:0c:20:FF:FE:0c:47:d5

A very special IPv6 address. Link Local

There is a default magic address called the link local address which will let you immediately talk to a neighbor on a point to point or local ethernet connection

- Take the link local prefix of fe80::/64 (see the table is useful)
- Concatenate it with the EUI-64
- We get a link local address of

fe80::02:0c:20:FF:FE:0c:47:d5/64

- That’s very useful if not globally routable

It’s all about the prefix length.

- The smaller the number in the prefix length, the larger the block.

- Let’s build upwards in this case
  - /48 which is the smallest block that can be injected into global routing tables.
  - /32 is the default allocation for an ISP by ARIN.
    - We’re in North America so ARIN is the RIR and they have separate rules for end user and ISP initial allocations That’s 65,536 /48 subnets for you to assign so you don’t need more space until you’re quite big.
    - You are assigning each customer a /48, aren’t you?

- Quick note: currently assigned IPv6 public space lives in 2000::/3 which is still a seriously huge amount of space.

Common prefix lengths

- Unlike IPv4 space, IPv6 is easy to get and you should be allocating sufficiently large blocks of IPv6 space.

- The basic subnets lots of people use are
  - /48 which is the smallest block that can be injected into global routing tables.
  - /56 & /60 are often used for summarization or as a base customer allocation.
  - /64 which is the normal LAN allocation. Since there are 64K of them in a /48 this shouldn't be an issue anywhere.
  - /126 and /127 for Point to Point links. The RFC says you should allocate a /64 and provision a /127 for PTP, it's still common to see providers provision a /126 (and *gasp* dense allocated inside a /64) for PTP links
  - /128 which is a loopback address.

Hitting the curve ball that was tossed earlier

- Think back, way back, probably fifteen minutes, and you will remember that IPv6 doesn’t have the concept of broadcast
- ARP is a broadcast mechanism.
  - Pretty simple one really, you broadcast a “Hi, I’m IP and I’m looking for MAC”   If all works, either MAC responds if local or a router responds if MAC isn’t local.
- There is a new mechanism called Neighbor Discovery that replaces ARP

Neighbor Discovery Protocol (NDP)

- What we get instead is our new friend: Neighbor Discovery Protocol, with its humble servants
  - Router Solicitation
  - Router Advertisement
  - Neighbor Solicitation
  - Neighbor Advertisement
  - Redirect
  - Duplicate Address Detection
  - Parameter Discovery, link MTU, hop limits.

Neighbor Solicitations/Advertisements

- This is what replaces IPv4's ARP, but it's a directed multicast rather than a broadcast.
- Neighbor Solicitation
  - Send out a LAYER 3 multicast request to the defined prefix FF02::1:FF00:0/104 with a host address of the last 24 bits of the MAC address. (if you care: RFC4291 Page 15 Solicited-Node Address)
  - Hopefully get a Neighbor Advertisement back with the routable IPv6 address.
- Duplicate Address Detection
  - What it sounds like, checking to make sure an auto-generated address is unique.

Router Advertisements/Solicitations

- Once enabled, IPv6 routers will send out Router Advertisements on a periodic basis.
  - Default is about every three minutes, it's tunable.
  - The advertisement will contain advertisements for
    - All the /64 prefixes configured on a given interface
    - Link MTU
    - Hop limits
    - Whether or not it is available as a default router
- Once an IPv6 host has received a Router Advertisement it can than magically auto configure a routable IPv6 address using the /64 advertisement and the EUI-64 [danger lurks, as we'll see later]
- Router Solicitations are a multicast to the all-routers multicast address to request an advertisement.
- Redirect will tell you if there is a better router for a given target.

Ways of Addressing

- Primus: configure static addresses and routes with IPv6 just the same way you can with IPv4

- Secundus: just let the IPv6 client use Stateless Address Autoconfiguration (SLAAC)

- Tertius: use DHCPv6 in a very similar fashion to IPv4 to statefully configure a single client.

- DHCPv6 has a lovely addition called prefix delegation (DHCP-PD)
  - DHCP-PD will
    - provide an IPv6 prefix which can than be routed upstream.
    - default to whatever is configured
    - has a feature called 'hints' which will let you request a different size delegation up to whatever maximum is configured in the DHCP server. You could for instance request a /56 so you can locally subnet into a group of /64

- Using DHPC-PD to request/assign subnets and SLAAC locally is a very popular combination.

Stateless Address Autoconfiguration (SLAAC)

- SLAAC is a very simple thing in concept and implementation
  - Determine the EUI-64 address of the interface in question
    - MAC 18-66-da-16-70-eb (e.g. a dell desktop)
    - EUI-64 1a66:daff:fe16:70eb
  - Find the /64 that covers the ethernet segment using router solicitation or promiscuous eavesdropping
    - /64 prefix 2001:db8:dead:beef::/64
  - Slam them together
    - 2001:db8:dead:beef:1a66:daff:fe16:70eb/64
  - Conversely, we automatically get the link local address.
    - Link Local fe80::1a66:daff:fe16:70eb/64

Potential dangers of SLAAC

- It's entirely possible you can speak IPv6 without knowing it!
  - All you need to speak IPv6 is kernel support and a willing router sending RA
  - Since SLAAC addresses are built from RA + EUI-64 it's very easy to track a machine across networks. This could be a real issue in public spaces
  - Privacy addressing (RFC4941/7271/7721) are a first pass at resolving these issues
  - RFC 8084 is the most recent update and worth reading.
    - Truthfully, unless you are an implementer it's not necessary and out of scope here

The Gorilla in the room.

- For local network use host address NAT is not only not a requirement, it's not possible
  - Nat 66 aka prefix NAT can be used to solve some issues such as non-routed multi-homing.
  - Local NAT is not the same as the various IPv4/IPv6 mechanisms which are way out of scope here.

- Use a good stateful firewall if you provide inbound services, and a simple blocking firewall if you are not.
  - PCI just wants you to obscure the internal network and that's easy to do with proper stateful firewalling and load balancing.
  - ICMP is a requirement for IPv6 to work properly, so take care in blocking it.

- One reason people would wrongly assign a small subnet (say a /124) was to mitigate a common DOS attack based on ND where a switches memory is totally consumed. This has been rectified by rate limiting and memory management at the silicon level in current products.

At a glance: IPv4 vs IPv6

Description	IPv4	IPv6
Address Length	32 bits	128 bits
Address Representation	Commonly dotted quads (xxx.xxx.xxx.xxx) where xxx is from 0-255	8 groups of 4 hex digits separated by colons, well known compressed fomat
Address types	unicast, multicast, broadcast	unicast, multicast, anycast
Packet Header	20 bytes	40 bytes
Configuration	Manual, DHCP	Manual, SLAAC, DHCP
IPSec Support	Optional	Baked in

Routing stuff and a mindset change.

- Some realities
  - IPv4 has very limited address space available, the RIR are functionally out and this has raised an interesting secondary market
  - IPv6 has functionally unlimited address space, it’s easy to get and you need to change your mindset!
  - Use proper hierarchy, there is no need to conserve IPv6 space.
  - You may need to pass IPv4 over an IPv6 only link (think LTE cellphones, they’re native IPv6.) This is called tunneling and beyond the scope of this presentation.
- Did I leave out huge amounts, you bet I did. You can find complete samples on the interwebz ™ if you use your favorite search engine.

What routing protocols support IPv6

- Routing in IPv6 is just like IPv4 except where it isn't
  - Static routes are static routes.
    - Instead of a /30 or /31 we use a /127 for the PTP
  - RIP v1 and RIP v2 don't understand IPv6 at all
  - RIPng (RFC2080) adds support for IPv6
  - OSPFv2 doesn't understand IPv6 at all
  - OSPFv3 speaks IPv6 and lots of people will run OSPFv2 and OSPFv3 rather than use OSPFv3 realms
    - You wind up having to run parallel routing protocols most of the time
  - ISIS doesn't care about life at that level so IPv4 and IPv6 are treated equally and transparently.
  - BGP speaks IPv4 and IPv6 fluently and can be configured in lots of ways to do interesting things
    - looks just like IPv4 only with IPv6 addresses and filters.
    - Instead of a /32 use a /128 for the loopback for IBGP

Evil! Pure and simple from the Eighth Dimension [3]

- A bit back, Geoff Houston of APNIC did a presentation on connection failure rates and the difference between IPv4 and IPv6. The basic outcome was stated above.
- Sometimes, you will need to pass IPv6 traffic over an IPv4 only link due to hardware or provider limitations.
- Ok, maybe not that bad, but close:
  - Tunneling protocols
    - 6to4
    - 6RD
    - Teredo
    - ISATAP
  - Translation
    - 464XLAT
    - NAT64/DNS64

Tidbits and leftovers

- If you are specifying an IPv6 address in a URL you need to embed it into square brackets like this
  - https://[2001:db8:dad:17]/my_website_index.html
- Most PING and TRACEROUTE applications will default to IPv4 unless fed an IPv6 literal address, or the only response to a DNS query is an AAAA record
  - ping6 or ping -6 or some variant will normally force an IPv6 request

Q&A and Footnotes

- [1] Douglas Adams Hitchhikers Guide to The Galaxy
- [2] Aston Martin never released a DB8, it would have been in the timeframe of 2001 making for a great geek joke.
- [3] Buckaroo Banzai reputed to have been spoken this at a conference regarding the oscillation overthruster when his presentation was interrupted by Red Lectroids intent on stealing the working prototype
- Look at the APNIC heatmap, the current one is at https://stats.labs.apnic.net/ipv6
- The ARIN first_request guide is https://www.arin.net/resources/guide/ipv6/first_request/
- There are a set of router configuration samples as an appendix, the details are better gone over in small print and text.

Router Configurations Appendix 1

I don’t think anyone really knows the true market share numbers; however, Cisco and Juniper are the two biggest commercial router vendors, so I’ll show samples of their configurations later in the presentation after some more underpinning.

If you are using one of the smaller brands, please refer to vendor documentation augmented by web searching for samples configurations.

If you are using F/OSS routing solutions, please refer to the source.

Routing Sample for Cisco Static Route

- Cisco routers require you to globally enable ipv6
- You need to tell IOS it's speaking IPv6 on a command line

```
R1(config)#ipv6 unicast-routing
R1(config)#ipv6 cef
R1(config)#int Gi0/0
R1(config)#ipv6 address 2001:db8:AABB:1234::/64 eui-64 <- autogen address
R1(config)#ipv6 address 2001:db8:AABB:1234::1/64 <- manual address
R1(config-if)#exit
R1(config)#int Gi0/1
R1(config)#ipv6 address 2001:db8::1/127
R1(config)# ipv6 route ::0 2001:db8::0
R1(config-if)#exit
```

Routing Sample for Cisco OSPF

- Cisco routers require you to globally enable ipv6
- You need to configure OSPF inside the interface statement
- You need to tell IOS it's speaking IPv6 on a command line

```
R1(config)#ipv6 unicast-routing
R1(config)#ipv6 cef
R1(config)#int Gi0/0
R1(config)#ipv6 router opsf 1
R1(config-router)#router-id 1.1.1.1
R1(config-router)#exit
R1(config)#int lo 0
R1(config-if)# ipv6 address 2001:db8::127/128
R1(config-if)# ipv6 ospf 1 area 0
R1(config-if)# exit
R1(config-int) ipv6 address 2001:db8::dead:beef::127/128
R1(config-if)#ipv6 ospf 1 area 0
R1(config-if)#ipv6 ospf cost <metric>
R1(config-if)#ipv6 ospf hello-interval <second>
R1(config-if)#ipv6 ospf dead-interval <seconds>
R1(config-if)#ipv6 ospf network point-to-point
R1(config)#ipv6 address 2001:db8:AABB:1234::/64 eui-64 <- autogen address
R1(config)#ipv6 address 2001:db8:AABB:1234::1/64 <- manual address
R1(config-if)#ipv6 ospf 1 area 0
R1(config-if)#ipv6 ospf cost <metric>
R1(config-if)#ipv6 ospf hello-interval <second>
R1(config-if)#ipv6 ospf dead-interval <seconds>
R1(config-if)#ipv6 ospf network point-to-point
R1(config-if)#exit
```

Routing Sample for Juniper Static Route

- Juniper routers default to IPv6 enabled
- A static route on a Juniper in IPv6 looks very similar to an IPv4 static route you must put it in the IPv6 RIB.
- You still need to configure the IPv6 address on the interface
- The sample will be in Juniper set format

```
set interface xe-0/0/1 unit 0 family inet6 address 2001:db8:aabb:1234::1/64 <- manual address
set interface xe-0/0/1 unit 0 family inet6 address 2001:db8:dead:beef::/64 eui-64 <- autogen address
set interface xe-0/0/2 unit 2 family inet6 address 2001::db8::1/127
set routing-options rib inet6.0 static route ::0 next-hop 2001:db8::0
```

Routing Sample for Juniper OSPF Route

- Juniper routers default to IPv6 enabled
- An OSPF route on a Juniper in IPv6 looks very similar to an IPv4 ospf route you must put it in the OSPFv3 protocol region
- You still need to configure the IPv6 address on the interface
- The sample will be in Juniper set format

```
set interface lo0 unit 0 family inet6 address 2001:db8::127/128 primary preferred
set interface lo0 unit 0 family inet6 address ::1/128
set interface xe-0/0/1 unit 0 family inet6 address 2001:db8:aabb:1234::1/64 <- manual address
set interface xe-0/0/1 unit 0 family inet6 address 2001:db8:dead:beef::/64 eui-64 <- autogen
address
set interface xe-0/0/2 unit 0 family inet6 address 2001::db8::1/127
set protocols ospf3 area 0.0.0.0 interface lo0.0
set protocols ospf3 area 0.0.0.0 interface xe-0/0/1.0 metric 10 interface-type p2p
set protocols ospf3 area 0.0.0.0 interface xe-0/0/2.0 metric 10 interface-type p2p
```