

V6 – DON'T DO IT WRONG

15-Feb-2023 NANOG 87 Atlanta GA

Hello!



ALEX LATZKO

PRINCIPAL NETWORK ARCHITECT
ALATZKO@DEFT.COM

Present

- Principal Network Architect at Deft
- Fairly fluent in JunOS and IOS
- Something of an IPv6 evangelist

Past:

- <https://www.linkedin.com/in/alexlatzko/>

What are we doing here?

- Free advice! (It's worth more than you think)
- I've been telling this tale for a long time...
- As a consultant, I usually charge lots of money for the same info...
- 100% discount today!

Agenda (or where are we going?)

- Start with some philosophy.
- Some tidbits about IPv4.
- Advice about IPv6.
- Some more philosophy .

What is the right headspace?

- Keep an open mind about how you do things, what you did twenty years ago is probably badly in need of updating.
- There are more ways to do things than there are engineers doing them, but not all of them are optimal or even correct.
- Automation is our friend, and I'm not going any place near it in this talk. Just ensure you have a back out plan in place in case automation breaks, because it does, and when it does it could be a disaster.
- Discrete OOB is your saving throw when everything else breaks.

Don't believe the hype machine.

- IP address space is a logical construct, nothing real about it, but it does have actual constraints.
- Don't believe people who say IPv4 has four billion addresses! That's a raw number.
- If you don't already have IPv4 space, you're probably going to be paying a stiff price for it, either monthly from your provider, or through a broker since the RIR are out of stock.

Those four billion IPv4 addresses

- Start with 4,294,967,296 IPv4 addresses
- Subtract out all the reserved space and its significantly less.
- Subtract out all the space the USDOD has on ice, and it's even less.
- 10/8 is popular for a reason, but it's still limited. If you run out of RFC1918 space, don't expand into 11/8, go to IPv6!

Some last IPv4 tidbits

- If you were to fully disaggregate IPv4 you would get about 14M /24 routes, which will happily blow up the FIB in any hardware-based router I know about.
- The Default Free Zone is currently about 920K for IPv4 and 160K for IPv6. Do you remember what happened when the RSP720 all fell over? [2]

Speaking of IPv6 – the plan

- IPv6 space is readily available, there is no reason to be stingy, most RIR have it rigged that way.
- Before we do anything, if you're still using a spreadsheet or post-it notes, get a real IPAM that treats IPv4 and IPv6 on an equal basis.
- A very fast review of IPv6 (time me, it should be under a minute).
- How to divvy up large superblocks in what I consider to be a sane fashion.
- Down in the addressing trenches, where the sermon begins

Quick Overview of IPv6

- IPv6 available space is huge, 2^{128} , grep the web for your favorite analogy.
- IANA unicast address assignments are currently limited to the IPv6 unicast address range of $2000::/3$. That's still mind bogglingly big.
- IPv6 addresses are 128 bits normally displayed as 16 hex characters. For practical purposes DNS is your friend just like it is for IPv4.
- There are lots of good introductions to IPv6 out there, I personally like the one linked in the end notes. [4]
- Don't try to use IPv4 heuristics to conserve IPv6 space

Routable blocks (/48 +)

- Odds on if you're an ISP and getting your first block from a RIR it'll be between /36 and /32.
- Before you start slicing it up think about your address plans. Paper exercises are your friend.
- Think about the size of your POP in terms of bit boundaries and how to stack things.
- Think about whether to build for now and expand each POP to effectively utilize your allocation or build big and run out or rejigger.

Salient Points

- The smallest publicly routable IPv6 block is a /48 which is a huge footprint.
- There is no reason to use a /122 or /112 prefix in production, and it will break things like SLAAC and make privacy addressing impossible.
- For the sake of your favorite deity, don't build a billing system around assigning /128 to end users (a popular university trick in the mid 2000s)
- Unlike IPv4 where there really is no base building block the /64 is the base for IPv6 and feel free to use a whole /64 for one host, it's a good idea, since IPv6 is happy with lots of unicast addresses for a single host.

- It's a bit counter intuitive: **the bigger the prefix, the smaller the actual block!**

Base Prefix	Breaks Into	That Breaks Into /64
32	16 /36	16M
32	256 /40	16M
40	16 /48	1M
48	256 /56	64K
56	256 /64	256

Useful Address Thoughts

- Take a /48 for your POP and then allocate /64 for P2P, just provision the link as a /127 since that's the BCP and RFC recommendation. There are known sideband attacks against /126 in P2P. [6]
- Go ahead and put a /64 on each port on a TOR and aggregate them up into a /56. Then use SLAAC and RA to magically configure hosts and still have hard knowledge of what hardware has which IP.
- If you run out of space in your POP, just add another /40 hold down route and keep on going.

Free Advice (it's worth the price)

- Don't like what the RFC says, then get involved in the IETF and see if you can convince some savvy people your way is right.
- RFC 822 says that “+” and “-” are valid characters and case doesn't matter. General case, use a well debugged input validation library, don't roll your own.
- RFC1884 was published in 1995, and you could call that the start of IPv6, so it's been around for a fair amount of time now
- RFC1925 has sage advice for you to follow, here are my favorite two sections:

Sage Advice from RFC1925

- With sufficient thrust, pigs fly just fine. However, this is not necessarily a good idea. It is hard to be sure where they are going to land, and it could be dangerous sitting under them as they fly overhead.
- Some things in life can never be fully appreciated nor understood unless experienced firsthand. Some things in networking can never be fully understood by someone who neither builds commercial networking equipment nor runs an operational network.

References

- [1] There are 14,461,946 publicly routable IPv4 /24 last time I figured it out
- [2] Here's a link, it happened in 2014
 - <https://arstechnica.com/information-technology/2014/08/internet-routers-hitting-512k-limit-some-become-unreliable/>
- [3] There are 592,709,120 reserved IPv4 addresses last time I counted them up. This doesn't consider large US DOD blocks which will probably never see the light of day (yes, I know about 11/8, it was originated from ASN23352 under DOD contract some years ago)
- [4] <https://youtu.be/blabib0mX6s>
- [5] Someone is bound to ask, what happens if you don't have a route and no default, and the answer is normally a simple ICMP destination unreachable.
- [6] RFC6164: Using 127-Bit IPv6 Prefixes on Inter-Router Links



Thank you

15 Feb 2023