

APNIC 44

# In Defence of NATs

Geoff Huston,  
Chief Scientist APNIC



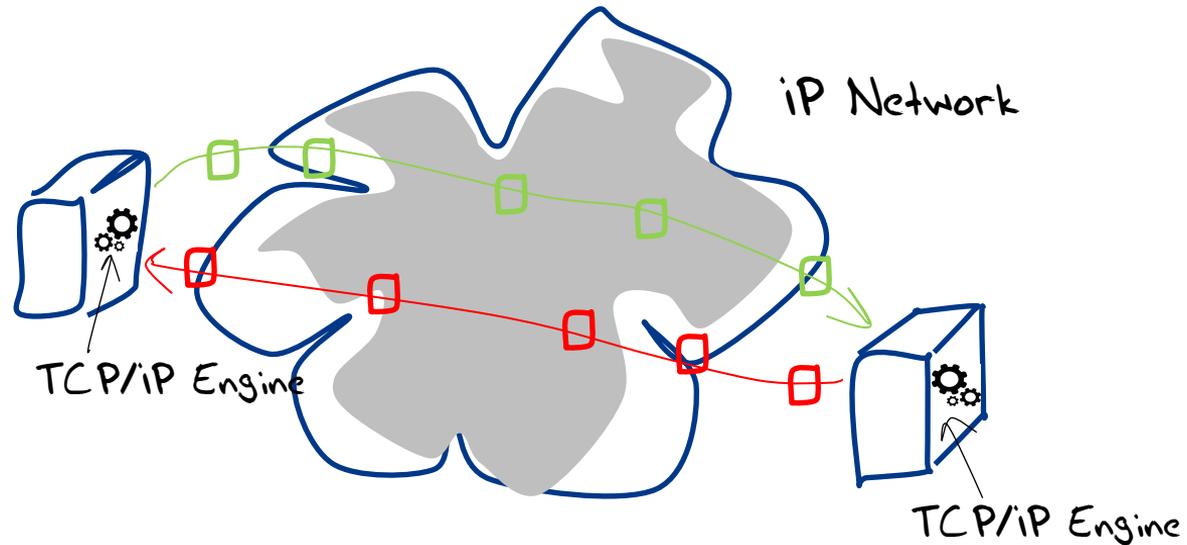
TAICHUNG, TAIWAN

7-14 September 2017

#apnic44

# The Architecture of the 1990's Internet

“Dumb Network, Smart Hosts”



# The Architecture of the 1990's Internet

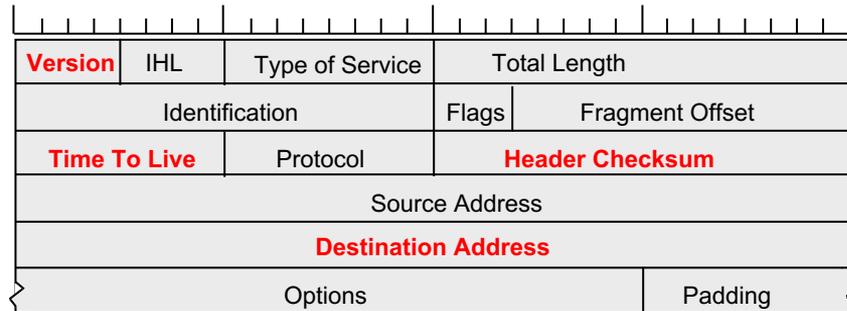
“Dumb Network, Smart Hosts”

Remove all the functionality from the network apart from forwarding and buffering

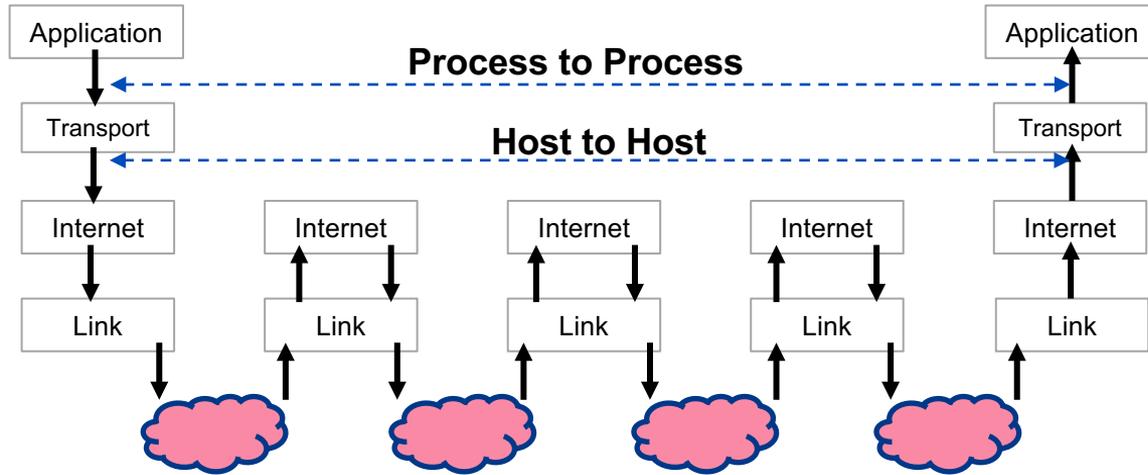
Place all the responsibility for data flow integrity and control into the end host protocol stacks

# The Architecture of the 1990's Internet

Each IP header carries sufficient information to enable a network of stateless forwarders to pass a packet to its intended destination



# Data Flow



The intention was that only the IP header is used by the network, and the entire remainder of the packet, including the transport headers, is an opaque payload

Each packet can be forwarded (optionally fragmented) or discarded

# IP's Strengths

- There is no “setup” and “tear down” of network state
- There is no requirement for symmetry between forward and reverse packet flows
- While it is preferred that the network maintain the order of packets, it's not a strict requirement
- End-to-End Transport and Hop-by-Hop Forwarding are decoupled: the end-to-end transport protocol is a host choice, not a network choice

# Consequences

- IP addresses are globally significant unique identifiers – they are not local virtual circuit identifiers
- All IP routers need to be aware of the relative location of all active IP addresses
- The total capacity of the network is limited by the number of IP addresses and the ability to represent the relative location of all these addresses within every router

# Proceedings of the 18<sup>th</sup> IETF Meeting

August 1990

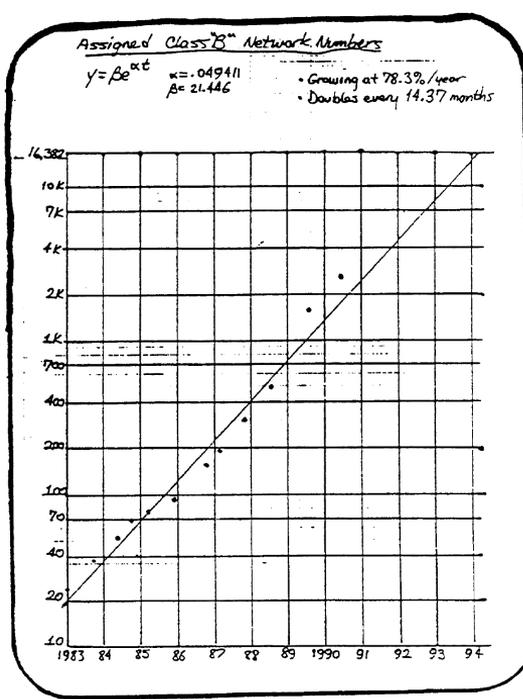
“Internet Growth”  
Frank Solensky,  
Proc. IETF, Aug 1990

Internet Growth (Continued):  
Continued Internet Growth

Frank Solensky  
Rosal Indarlan  
solensky@interlan.com

- A preliminary analysis of data presented earlier in the conference projects the “size” of the Internet on several metrics, assuming continued exponential growth.
  - NIC Assigned Network Numbers
  - NIC “connected” Status Nets
  - BBN’s snapshots
  - NSFnet Policy Routing DataBases
- As was mentioned during the discussion period, a logistic curve would likely be a more realistic model; this will be the subject of further analysis. Note, however, that the limit that this approaches may turn out to be beyond the capacity of the class A-B-C numbering scheme

# Proceedings of the 18<sup>th</sup> IETF Meeting



Depletion Dates

• Assigned Class "B" network numbers	Mar. 11, 1994
• NIC "connected" Class B network numbers	Apr. 26, 1996
• NSFnet address space*	Oct. 19, 1997
• Assigned Class "A-B" network numbers	Feb. 17, 1998
• NIC "connected" Class A-B network numbers	Mar. 27, 2000
• BBN snapshots*	May 4, 2002

\* all types: may be earlier if network class address consumption is not equal.

"Internet Growth" Frank Solensky, Proc. IETF, Aug 1990

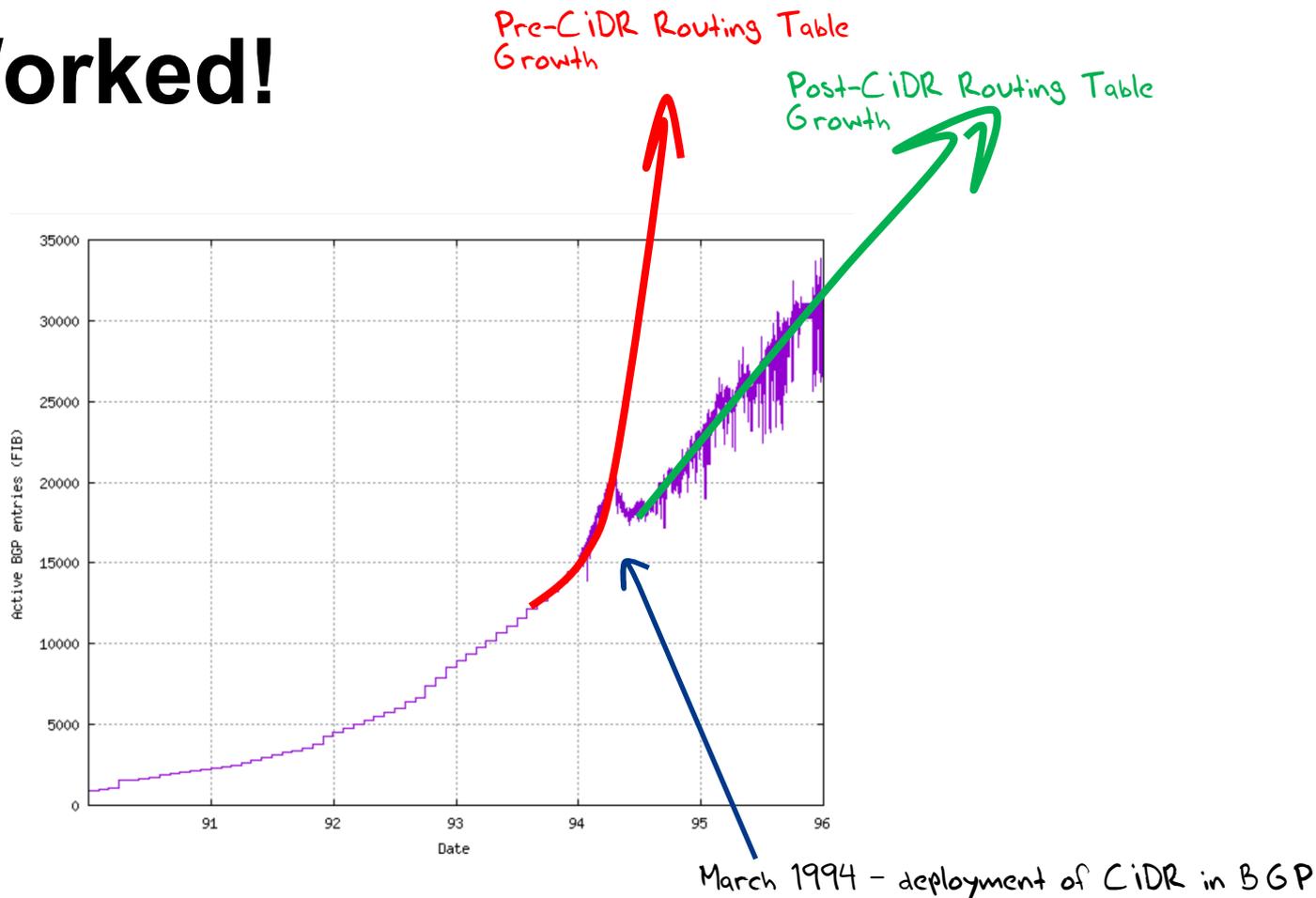
# Removing IP Classes

- The problem was **not** that we were running out of addresses
- The problem was that we were running out of Class B addresses
- The adopted solution was to keep the implicit routing aggregation of the network / host division of addresses, but carry the network/host division point with the address prefix

E.g. 192.0.2.0(24)

← Length of network prefix in bits

# CIDR Worked!



# “Buying Time” was just a stopgap measure

- The underlying issue was the personal computer, which dramatically increased the numbers of connected devices and indirectly fueled the first wave of commercialization of the Internet
- And this was not going to stop – laptops, mobile devices, and embedded ‘things’ have all added to the numbers of connected devices
- The thinking at the time was to change as little as possible, so we thought that enlarging the address fields of the IP header would be sufficient to meet this challenge to scaling

The “answer” that was adopted by the IETF was IPv6

# The origins of IPv6

- IPv6 represented a minimal change to IPv4
- Some aspects of IP were changed:
  - Expanded address fields
  - Altered fragmentation controls
  - Re-formatted options and control fields
- Much was unaltered:
  - UDP and TCP transport protocol behaviour
  - Hop-by-hop destination-based datagram forwarding
- And some was unspecified:
  - IPv6 Address Plan

# There were many other ideas that were aired at the time

And one of them was a mechanism for address "sharing"

---

[\[Docs\]](#) [\[txt|pdf\]](#) [\[draft-egevang-add...\]](#) [\[Diff1\]](#) [\[Diff2\]](#) [\[IPR\]](#)

Obsoleted by: [3022](#) INFORMATIONAL

Network Working Group K. Egevang  
Request for Comments: 1631 Cray Communications  
Category: Informational P. Francis  
NTT  
May 1994

## The IP Network Address Translator (NAT)

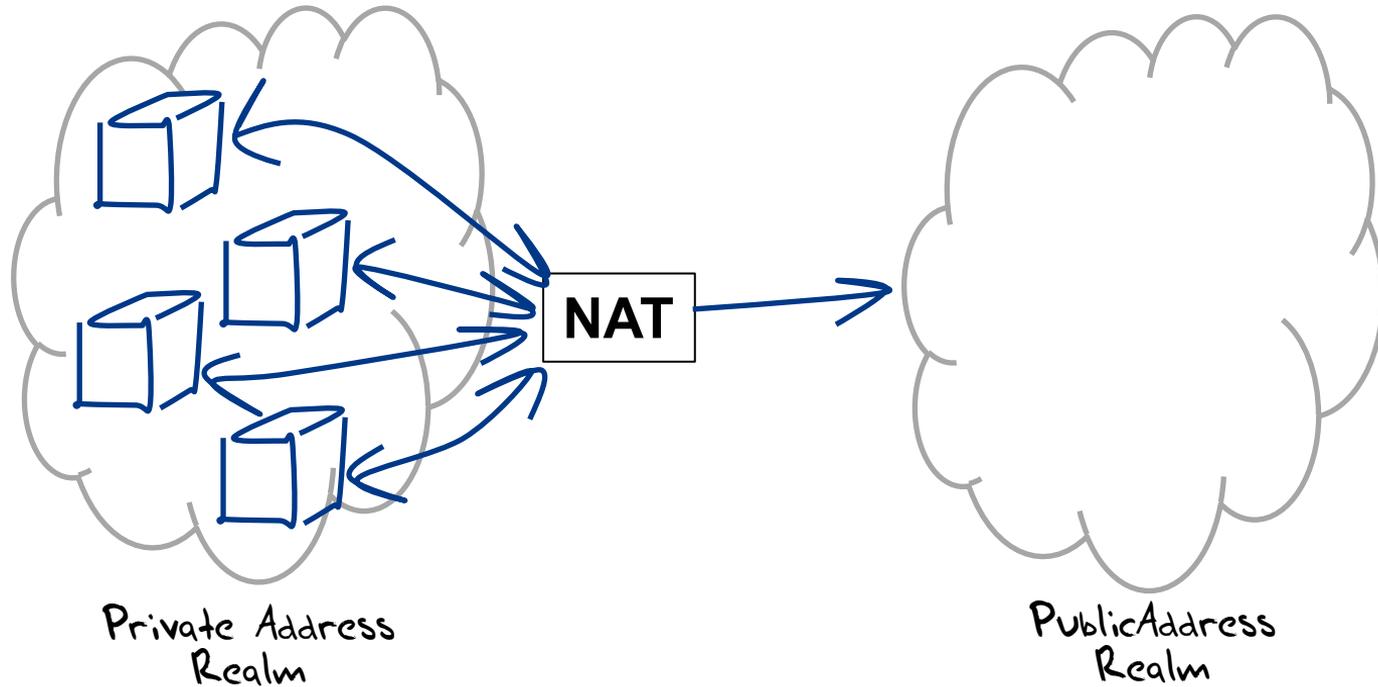
Status of this Memo

This memo provides information for the Internet community. This memo does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

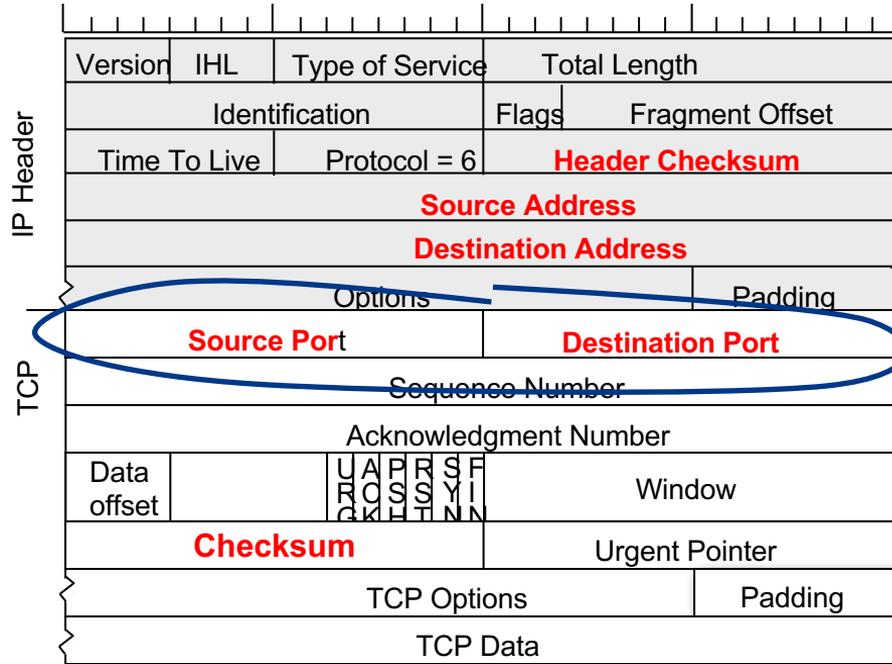
Abstract

The two most compelling problems facing the IP Internet are IP address depletion and scaling in routing. Long-term and short-term solutions to these problems are being developed. The short-term solution is CIDR (Classless InterDomain Routing). The long-term solutions consist of various proposals for new internet protocols with larger addresses.

# Address Sharing

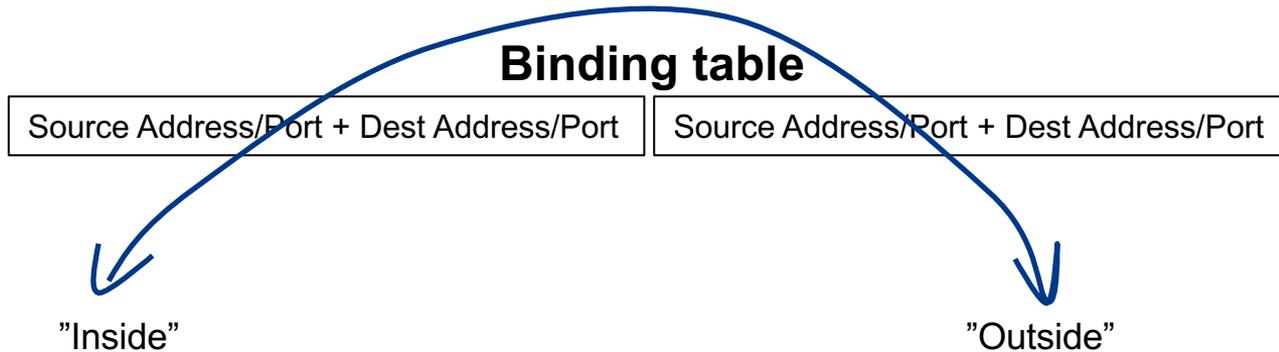


# Port-Translating NATs



These NATs allow address sharing by using the transport protocol's port fields as part of the address 'distinguisher'

# NAT



NATs take the 96-bit 4-tuple of Addresses+Ports and replaces the packet's fields that match one side of the binding table with the fields on the other side

Outbound packets that do not match any binding table generate a new table entry

Inbound packets that are not matched in the binding table are discarded

# NATs:



- Enforce symmetric network paths
- Require session state within the network
- Enforce Client / Server architecture
- Create fragility in the network
- Violate layer integrity
- Motivate the use of proxies and gateways
- Prevent innovation in transport services
- Handle UDP bindings inconsistently

# NATs are Evil!



- These considerations led to a long-held mantra in the IETF that “NATS are Evil”
  - The IETF refused to work on standardizing NAT behaviour for many years
- The consequence was that NATs were developed with idiosyncratic behaviours, particularly in relation to UDP-based binding
- Which led to all kinds of convoluted application level behaviours to discover and adapt to the NATs in the path (STUN, TURN, ICE, TEREADO,...)

# And yet...

# NATs run today's Internet

- 2.8B advertised IPv4 Addresses
- 25B connected devices \*
- The average address sharing ratio appears to be 10:1

NATs are everywhere!

\* <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/>

# Why are NATs so pervasive?

- They are backward compatible with the existing network
- They allow for uncoordinated piecemeal deployment
- They impair open peer to peer connectivity and enforce client behaviours behind the NATs
- They use the port space to enlarge the effective address space
- They isolate the edge networks from the carriage network

# Why are NATs so important?

- They are backward compatible
- They allow for uncoordinated
- They impair open peer to peer behind the NATs
- They use the port space to enlarge the effective
- They isolate the edge networks from the carriage network

This is an extremely important aspect of NATs. It allowed unconstrained innovation in devices (and services) offered in edge networks without meeting in any limitation imposed by the carriage operators

# How far can we push NATs?

I have 7,000 DSL broadband customers behind it. Peak time throughput is hitting up at 4 gbps... I see a little over 100,000 service flows (translations) at peak time

I think each MX104 MS-MIC-16G can able about ~7 million translations and about 7 gbps of cgnat throughput... so I'm good.

I have a /25 for each MX104 outside public address pool (so /24 total for both MX104's)... pretty sweet how I use /24 for ~7,000 customers :)

Aaron Gould, NANOG, <https://mailman.nanog.org/pipermail/nanog/2017-April/090809.html>

# How far can we push NATs?

The NAT binding space is a 96 bits wide 4-tuple



NAT type:	Address Compression Ratio
CGN with per user port blocks	10:1
CGN with per user port blocks + pooled overflow	100:1
CGN with pooled ports	1,000:1
CGN with 4-tuple binding maps	>>10,000:1

# So, in comparison, what does IPv6 offer?

*Is IPv6 Backward Compatibility with IPv4?*

- Nope!

*Does IPv6 have more usable address bits?*

- Yes, but
- The current IPv6 address plans typically use a 48 bit end site prefix
- CGNats potentially push the virtual IPv4 address space to between 42 and 45 bits

*Does IPv6 have more flexibility?*

- It is unclear if there is a clear need to shift back to the overloaded location / identifier semantics of addresses as the client / server model is closely aligned to today's Internet

*Does IPv6 represent lower cost?*

- Well no – for as long as IPv4 remains the dominant protocol then the Internet needs to support IPv4, which implies the use of NATs. It is IPv6 that is the discretionary expenditure item for many service providers

# What is going on?

- Our industry is running to a different script than what we had planned:
  - We had envisaged a transition from IPv4 through Dual Stack to IPv6
  - We are seeing a somewhat different transition from IPv4 to NAT44 to Dual Stack-plus-NATs to eventually end up with IPv6
  - The NAT phase has removed the urgency from the transition by papering over the crunch time of IPv4 supply exhaustion

# But its not just transition

- NATs have bought the Internet around 5 years of additional time in the transition by extending IPv4 addresses
- But they do more than address extension – they also alter the relative roles of names and addresses in the network's architecture

# NATs make addresses temporary

For some time now we have been contemplating what it means to have a name-based data network, where instead of using a fixed relationship between names and IP addresses, we eschew this mapping and perform network transactions by specifying the name of the desired service or resource to the network

NATs are an interesting step in this direction, where IP addresses have lost their static association with particular endpoints, and are used more as **ephemeral session tokens** than endpoint locators

# Name-based Networks

This certainly appears to be an important step in the direction of **named data networking** where addresses lose any permanent association with endpoint identity, and retain only the fixed semantics of a network locator used in routing

The **service name** becomes the endpoint identifier

Addresses are not used to identify service endpoints, but instead are used to distinguish and direct session traffic within the network

# In Defence of NATs

- It's NATs, and only NATs, that have kept the Internet running for the past decade
- In that time the network has grown from around 2B connected devices to ten times that number
- The IPv4 network and its application suite is now built on the assumption of pervasive use of NATs
- These same application and service design parameters are used in the context of IPv6

# IPv6 Addresses

- Do we use “fixed” end addresses in IPv6 anyway?
  - Well, no, not all the time!
  - Clients typically use “privacy” addresses that use a random 64 bit interface identifier
  - So the IPv6 IP end address is now ephemeral for clients
- Services are increasingly using name based hosting
  - There are more levers and control points in the DNS as distinct from IP anycast
- So even IPv6 has eschewed “fixed” end addressing!

# NATs in IPv6

For some this is heresy!

However:

- IPv6 has already dissociated itself from fixed end addresses
- we have already marked off ULAs as private use IPv6 address prefixes (fc00::/7 – RFC4193)
- Persistent private addresses and NATs avoid forced site renumbering
- NATs allow site multi-homing without route de-aggregation
- NATs obscure internal site network details and enforce client obscurity
- As IPv6 gathers momentum it may be the case that network admins will use ULA prefixes plus NATs to re-create the IPv4 NAT architecture in IPv6

# What's the Message Here?

# NATs are Good!



- Maintaining a network infrastructure that uniquely names and numbers every attached device has proved to be economically infeasible in the Internet
- NATs segment the device population into clients and servers - Clients are not uniquely named, nor uniquely addressed
- We may not have planned it this way, but undeniably what keeps today's Internet running as a single cost effective network is NATs.

# NATs as part of the Architecture

- NATs have pushed the network into a mode where IP addresses are ephemeral conversation tokens without lasting significance as an endpoint identifier
- For clients, address uniqueness is a locally scoped property, rather than a globally scoped attribute
- Some services still use unique addresses
- Other services use generic “aggregate” addresses and relay on the application to perform service identification

# Today's Internet Architecture

Is this the model of disambiguation of location and service identity that we've been searching for for the past couple of decades?

Are we over a model of networking where “addresses” uniquely denote points of attachment to a common network?

Are addresses locally scoped elements that provides disambiguation only to the extent that they are necessary?

# Questions?