In Defence of NATs

Geoff Huston

APNIC

IEEE Global Internet Symposium, May 2017

The Architecture of the 1990 Internet

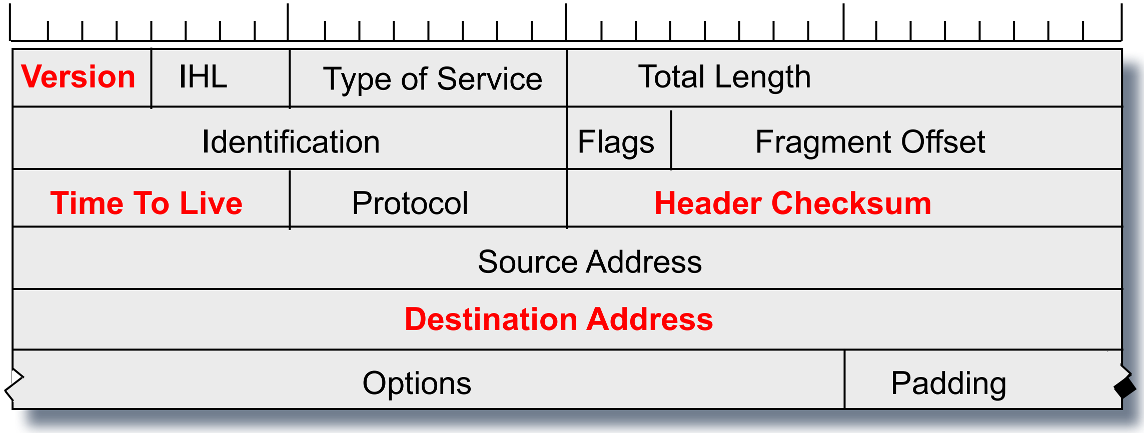
“Dumb Network, Smart Hosts”

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 Internet

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



Version IHL Type of Service Total Length

Identification Flags Fragment Offset

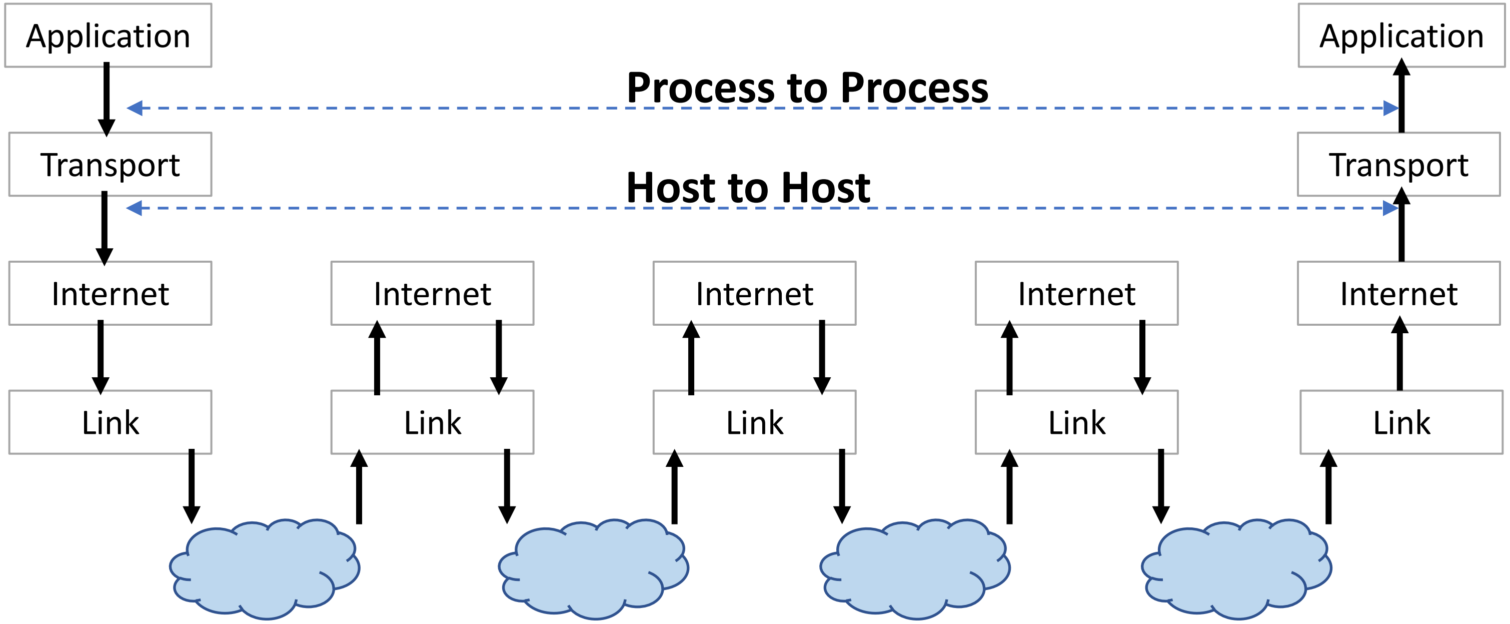
Time To Live Protocol **Header Checksum**

Source Address

Destination Address

Options Padding

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

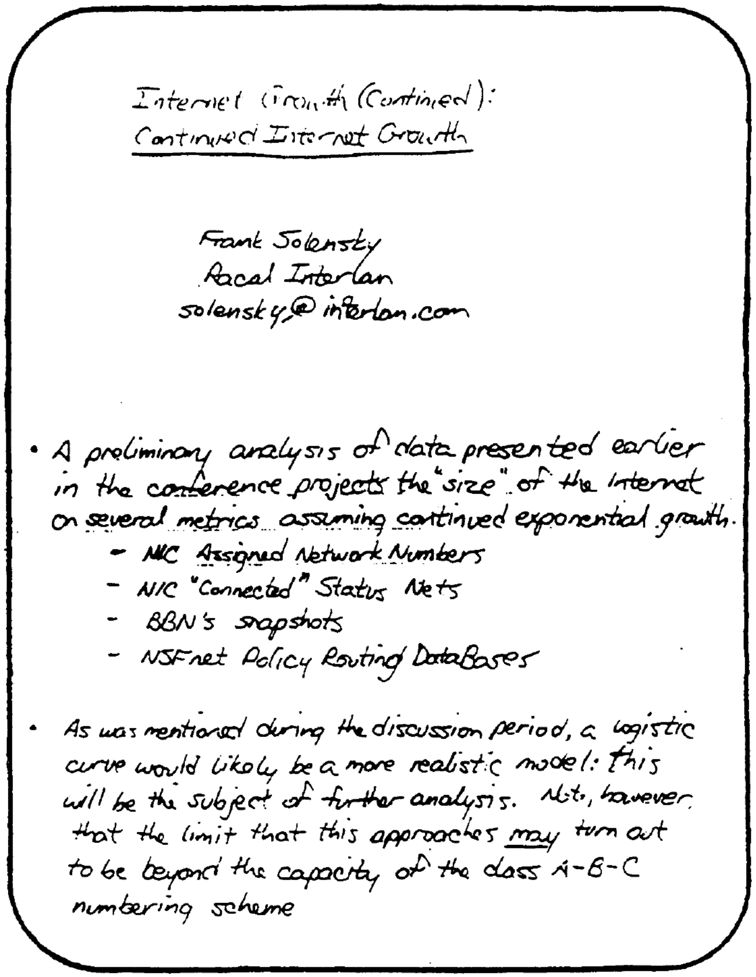
- There is no “setup” and “tear down” of network state
- There is no requirement for symmetry between forward and reverse packet flows
- While it is preferred that the network maintain the order of packets, it’s not a strict requirement
- End-to-End Transport and Hop-by-Hop Forwarding are decoupled: 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 18th IETF Meeting

August 1990



Internet Growth (Contined):

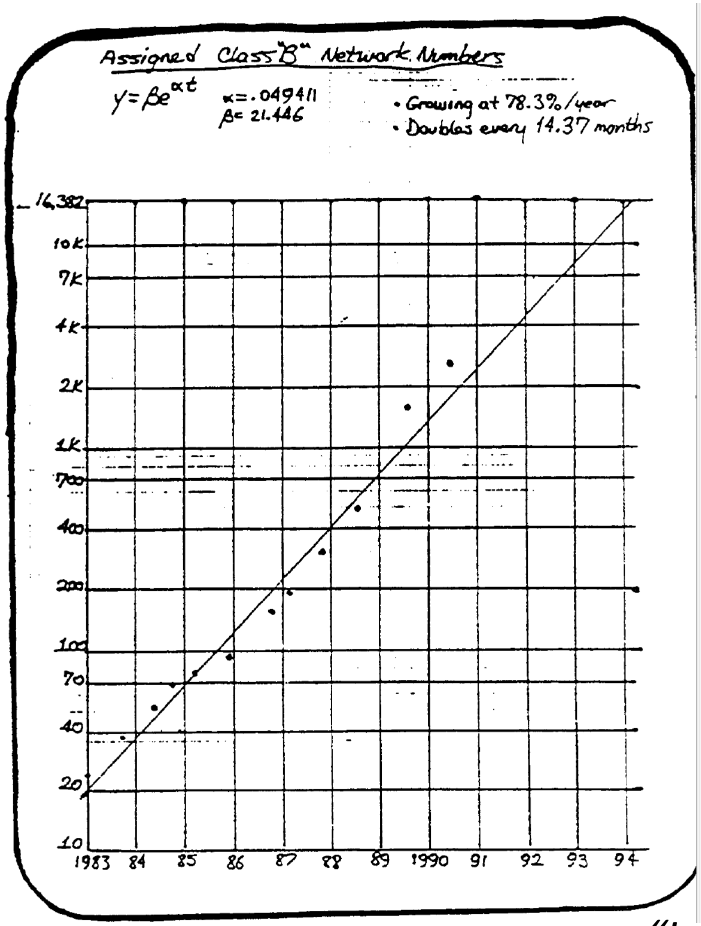
<u>Continued Internet Growth</u>

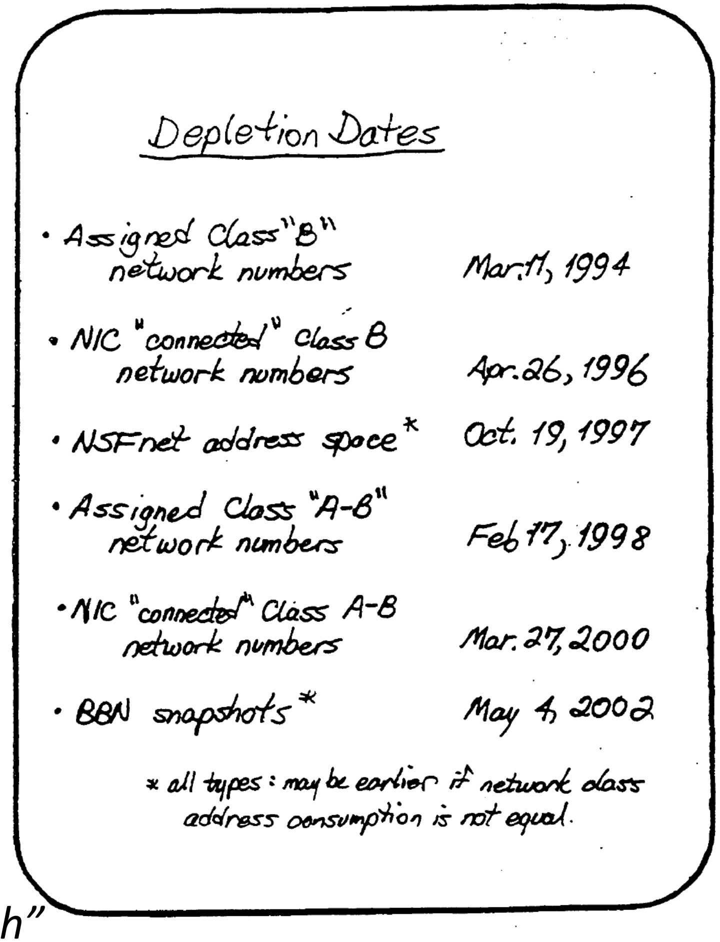
Frank Solensky
Racal Interlan
solensky@interlan.com

- 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 <u>may</u> turn out to be beyond the capacity of the class A-B-C numbering scheme

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

Proceedings of the 18th IETF Meeting





Depletion Dates

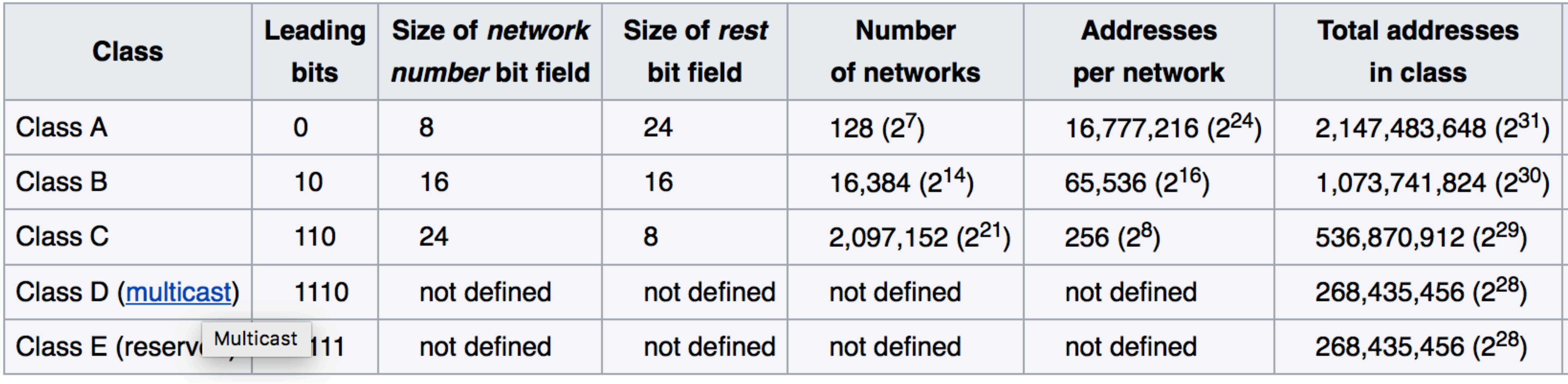
• Assigned Class "B" network numbers	Mar. 11, 1994
• NIC "connected" class B network numbers	Apr. 26, 1996
• NSFnet address space*	Oct. 19, 1997
• Assigned Class "A-B" network numbers	Feb 17, 1998
• NIC "connected" Class A-B network numbers	Mar. 27, 2000
• BBN snapshots*	May 4, 2002

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

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

Address “Classes”

- Each IP address has a network part and a host part
- The original IPv4 address plan divided the IPv4 address pool into three “Classes”



Class	Leading bits	Size of *network number* bit field	Size of *rest* bit field	Number of networks	Addresses per network	Total addresses in class
Class A	0	8	24	128 (2^{7})	16,777,216 (2^{24})	2,147,483,648 (2^{31})
Class B	10	16	16	16,384 (2^{14})	65,536 (2^{16})	1,073,741,824 (2^{30})
Class C	110	24	8	2,097,152 (2^{21})	256 (2^{8})	536,870,912 (2^{29})
Class D (multicast)	1110	not defined	not defined	not defined	not defined	268,435,456 (2^{28})
Class E (reserv	111	not defined	not defined	not defined	not defined	268,435,456 (2^{28})

https://en.wikipedia.org/wiki/Classful_network

Removing 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

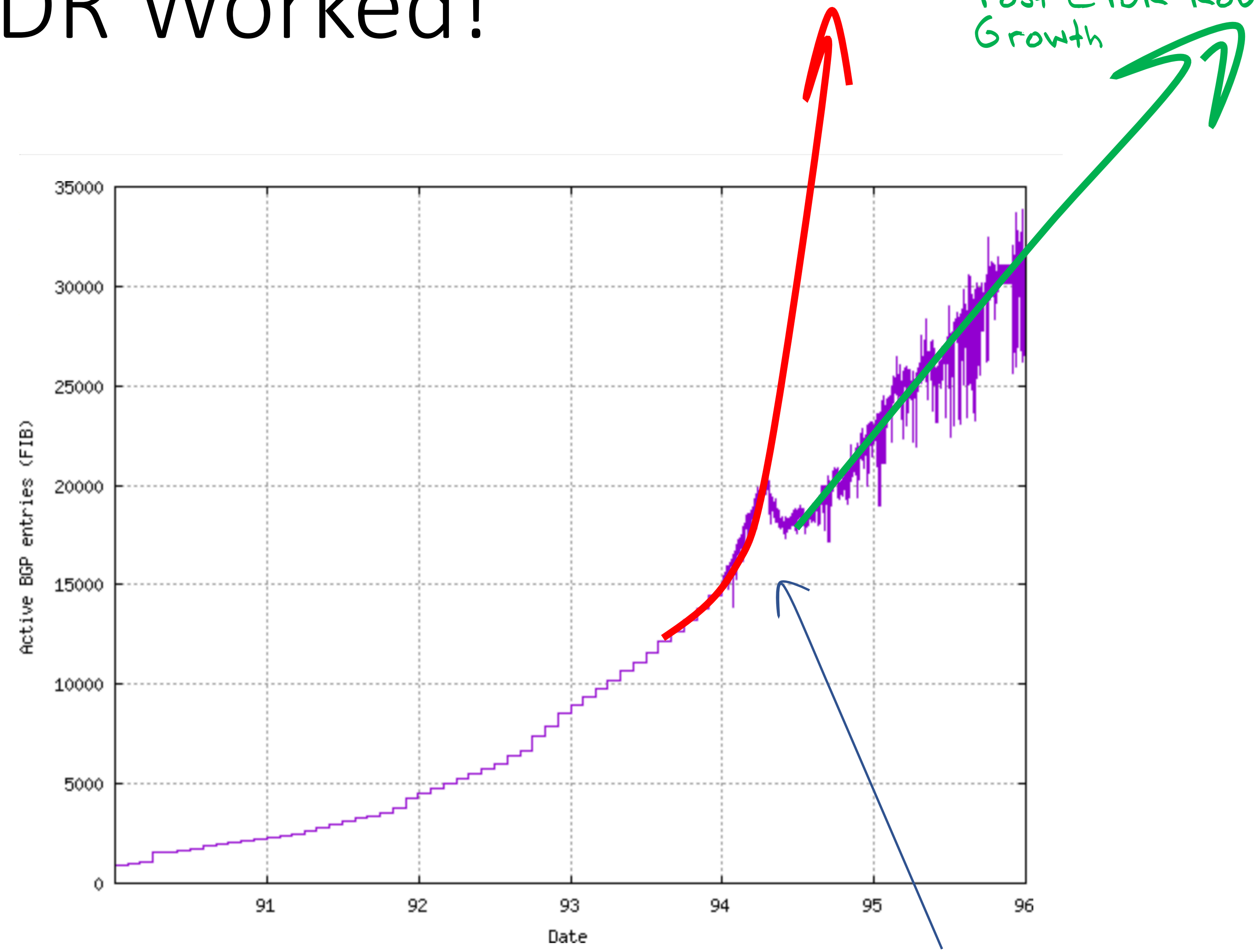
E.g. 192.0.2.0/24

Length of network prefix in bits

CIDR Worked!

Pre-CIDR Routing Table Growth

Post-CIDR Routing Table Growth



March 1994 – deployment of CIDR in BGP-4

“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, the mobile devices 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

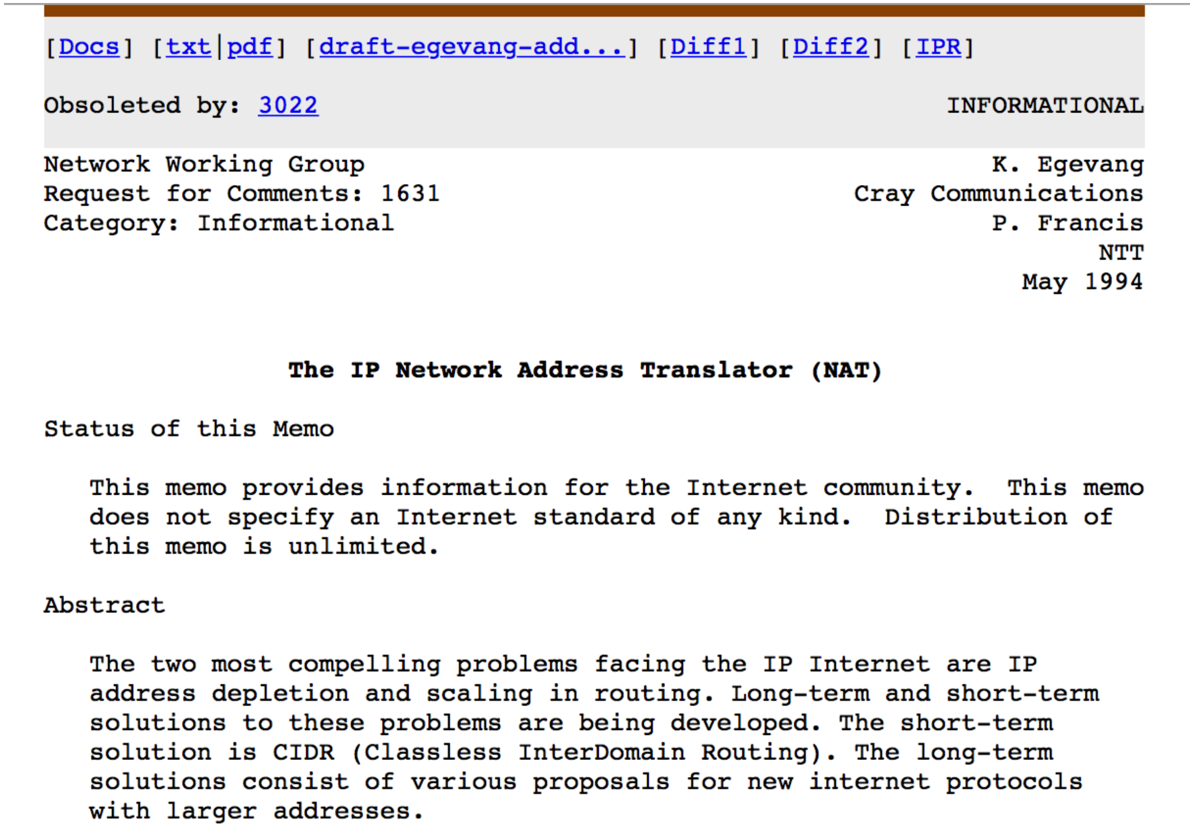
- The answer that was adopted by the IETF was IPv6

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

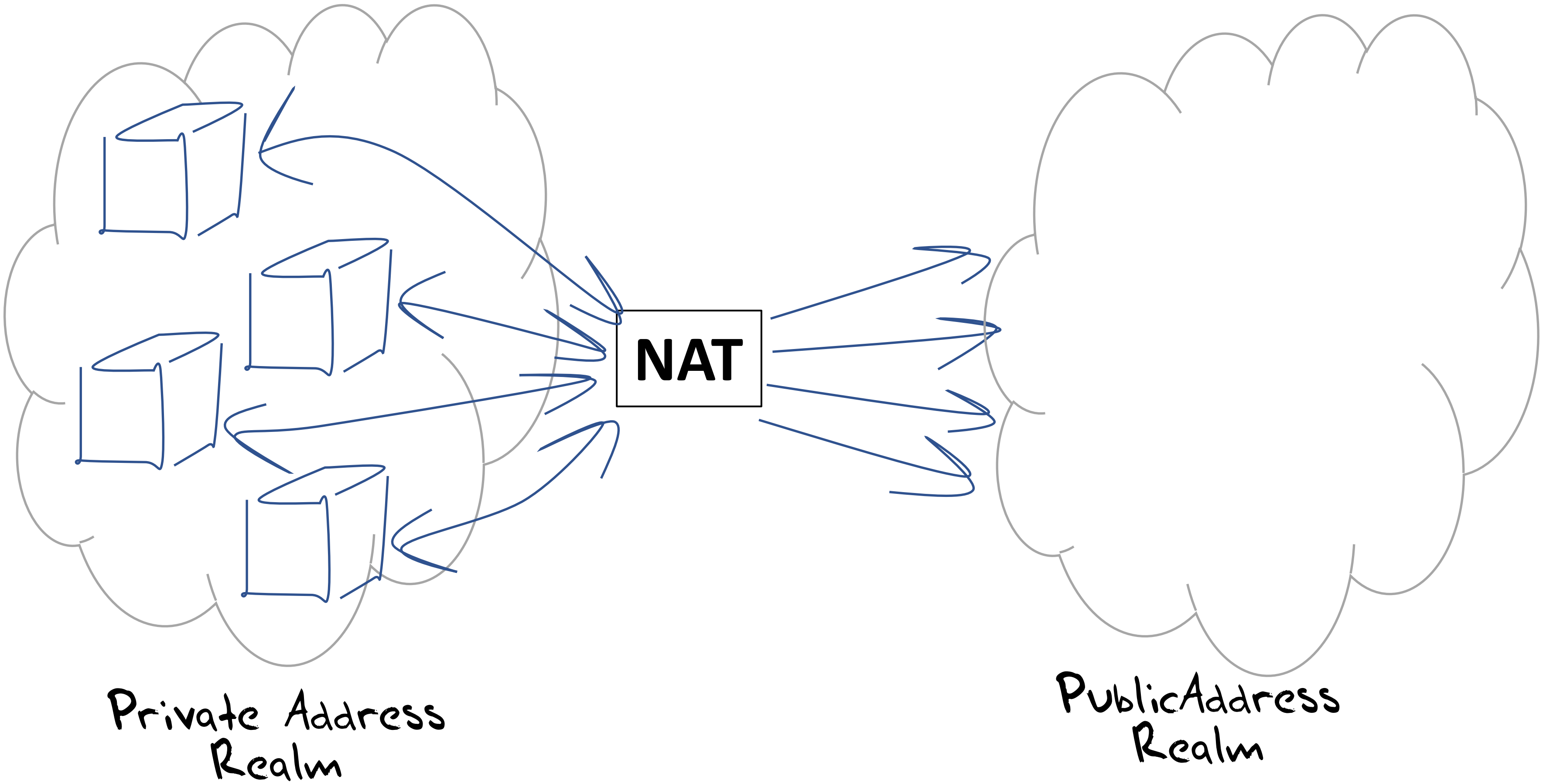
There were many other ideas that were aired at the time

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

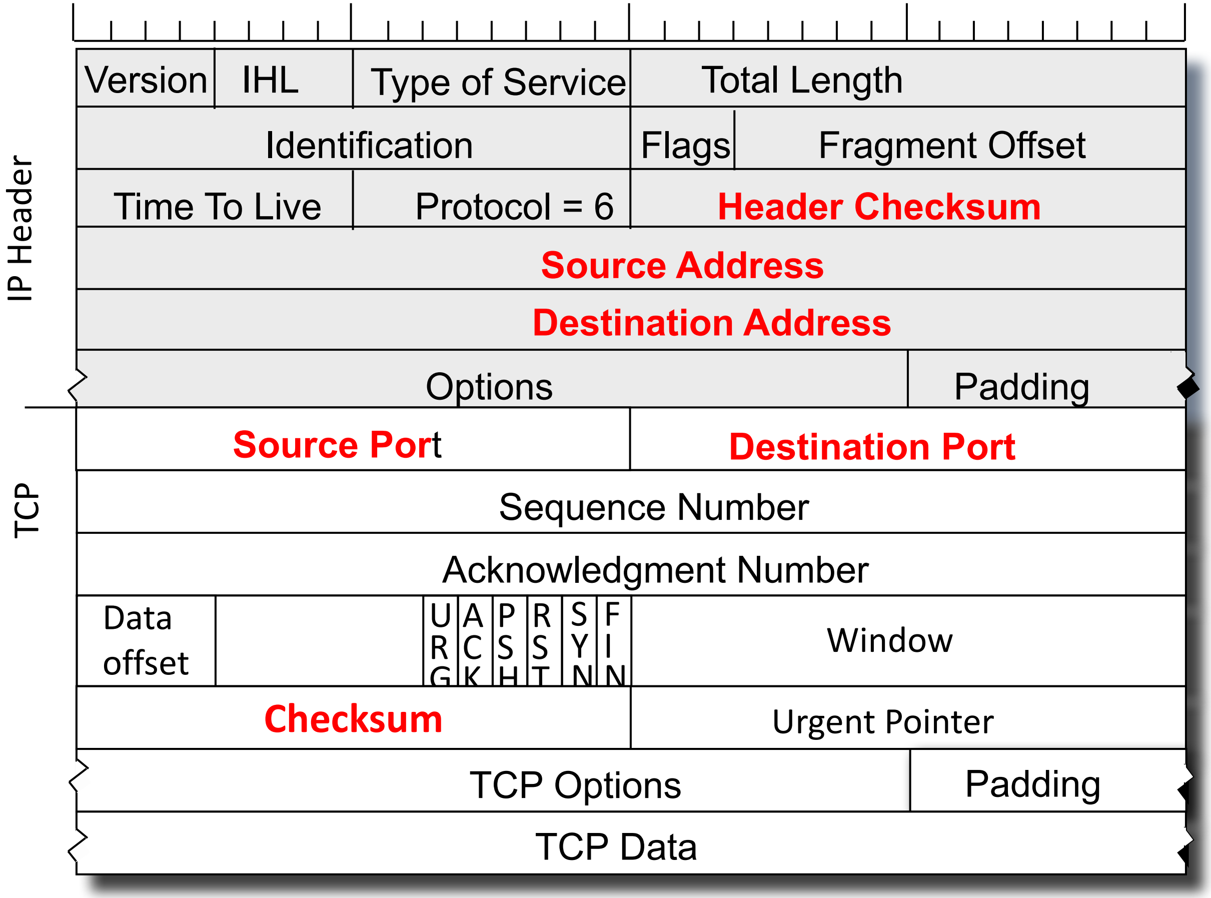


```
[Docs] [txt|pdf] [draft-egevang-add...] [Diff1] [Diff2] [IPR]

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



IP Header

Version
IHL
Type of Service
Total Length

Identification
Flags
Fragment Offset

Time To Live
Protocol = 6
Header Checksum

Source Address

Destination Address

Options
Padding

TCP

Source Port
Destination Port

Sequence Number

Acknowledgment Number

Data offset
URG
ACK
PSH
RST
SYN
FIN
Window

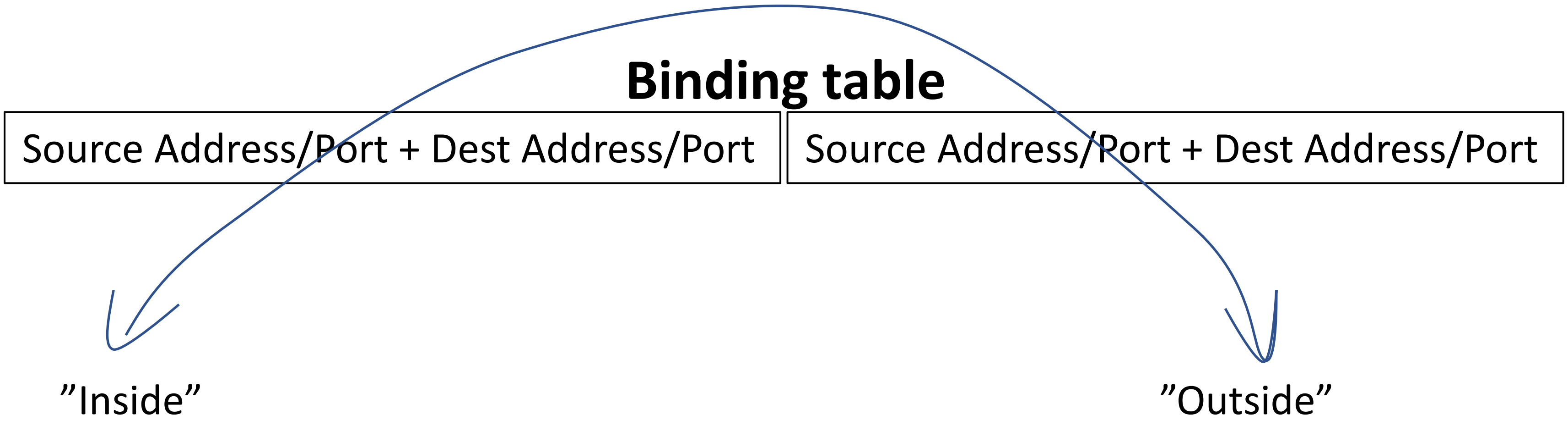
Checksum
Urgent Pointer

TCP Options
Padding

TCP Data

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

NAT



Binding table

Source Address/Port + Dest Address/Port

Source Address/Port + Dest Address/Port

”Inside”

”Outside”

NATs take the 96-bit 4-tuple of Protocol+Address+Ports and replace 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 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” and 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, TEREDO,…)

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

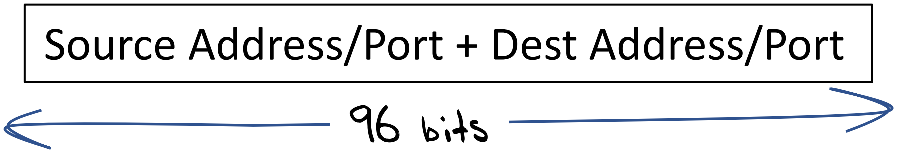
How far can we push NATs?

> 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



Source Address/Port + Dest Address/Port

96 bits

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?

- *Backward Compatibility with IPv4?*
  - Nope!

- *More address bits?*
  - 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

- *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

- *Less 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

NATs are under-appreciated!

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

NATs are under-appreciated!

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.

NATs are under-appreciated!

This certainly appears to be an interesting 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

In Defence of NATs

- It is 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

In Defence of NATS

- 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?