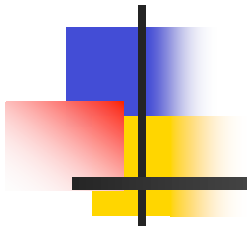
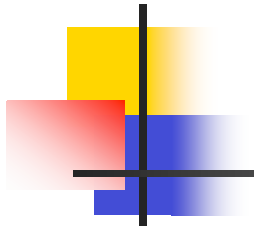


What Do ISOC and the IETF Think About IPv4 Address Space Exhaustion?

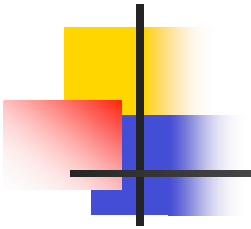
Fred Baker
Member, ISOC Board
Chair, IPv6 Operations WG



ISOC Viewpoint

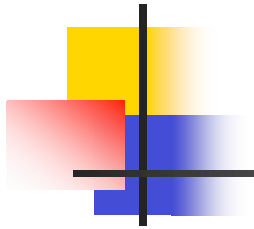
- IETF and NRO say that IPv6 is the next best bet
- Therefore, ISOC promotes training and public policy discussion supporting an IPv6 solution

IETF Viewpoint: The short answer:



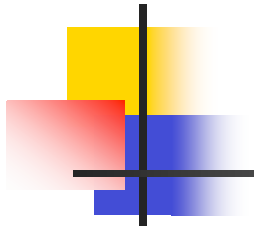
- We thought it might happen sometime…
- We tried to prepare for it
- I dunno, what does the operational community think about it?

The rest of this presentation is the somewhat-longer version…



IETF viewpoint: A little history

- The IETF first worried about this ~1992:
  - Frank Solensky: then-current address allocation policy would have us run out sometime in 1993
  - RIRs/ISPs shifted to CIDR pretty quickly
  - IETF decided it was a good idea ~2 years later (RFC 1519)



History continued

- Some address space was set aside for unconnected networks (RFC 1918)
  - NATs were invented about five minutes later
- IANA started soliciting the return of unused address space, recovering perhaps 15%
- The IETF started thinking about what it called “IPng”

OSI CLNP:
- Callon: TUBA (RFC 1347)

IPv5:
- BBN Stream Transport I and II (RFC 1190)

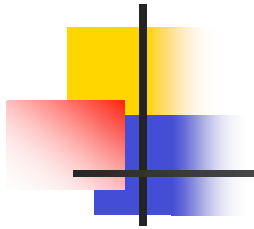
IPv6:
- Hinden/Deering merged proposals (now RFC 2460)

IPv7:
- McGovern/Ullmann: CATNIP (RFC 1707)

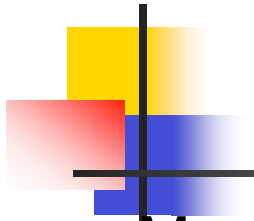
IPv8:
- Paul Francis: PIP (RFCs 1621 and 1622)
- IPng Recommendation: RFC 1752

Routing:
- Chiappa: NIMROD routing architecture (RFCs 1753, 1992, 2102, 2103)

RFC 1752 Comments on Routing Table Size



- In 1995, Route Table growth rate was considered a problem
  - Growth rate faster than memory technology
- Recommendations:
  - ISPs need to aggregate
  - Vendors need to put lots of memory into routers
  - Rigorous Address assignment policy
- Recommendations for address management
  - Renumber networks rather than assigning multiple prefixes
    - Make renumbering easy
  - Multihomed edge networks seen as the limit on the effectiveness of prefix aggregation

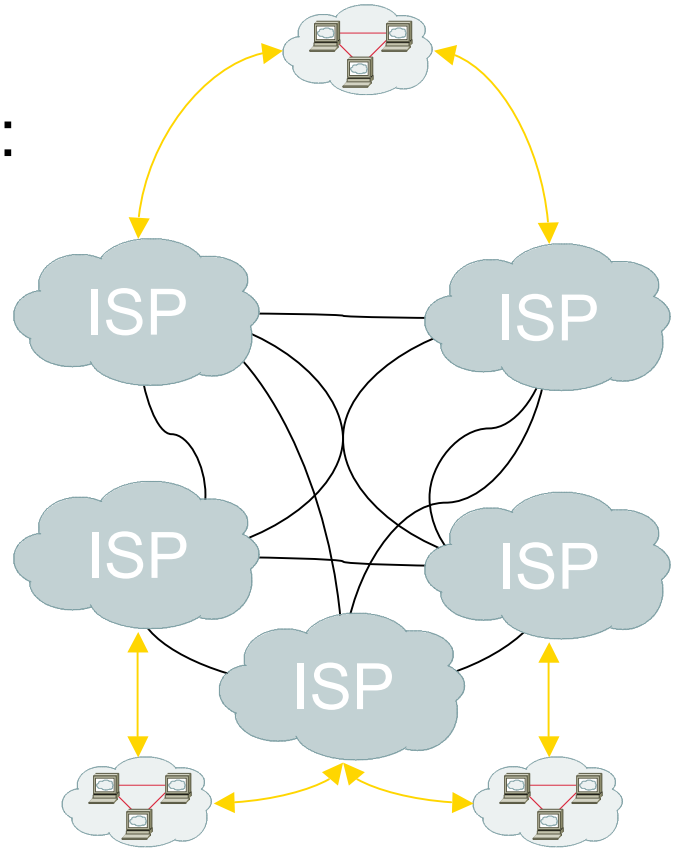


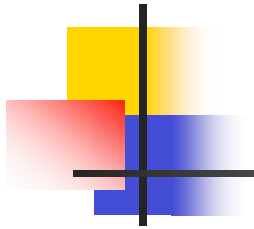
Renumbering

- Much work done:
  - Various preparatory documents
  - RFC 2894:
    - Renumbering Protocol
  - RFC 2874:
    - DNS records to support renumbering
  - RFC 4076:
    - DHCP extensions to manage numbers
  - RFC 4192:
    - Procedures for renumbering IPv6 networks without a flag day
- RFC 4192 Key learning:
  - The issues that make renumbering hard cannot be solved with a protocol
  - Good configuration database tools can help
- Issues
  - Errors in application design
  - Non-use of DNS by applications
  - Configuration of numeric objects requires numbers

Multihoming

- Principal reason for edge network getting an AS number is multihoming:
  - Network viewpoint:
    - Prefixes are assigned to entities whose routing connectivity and size make it advantageous to maintain global knowledge of their routing and who desire additional guarantees of internet connectivity
  - Customer viewpoint:
    - Service is obtained from multiple providers to improve reliability or other characteristics



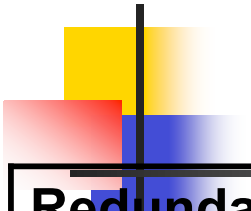


Basic Route Table Math:

- Prefixes for remote ISP AS's
  - 3792 AS's display transit/peering services
- Prefixes for Multihomed Edges
  - 23,824 origin-only AS's of 27,616
  - Of which 11,611 originate a single prefix
  - 18,466 are not even multihomed…
- Prefixes for my own network
  - "More Specific"
- Other prefixes I might choose to use

http://bgp.potaroo.net/as2.0/bgp-active.html 27 Feb 2008

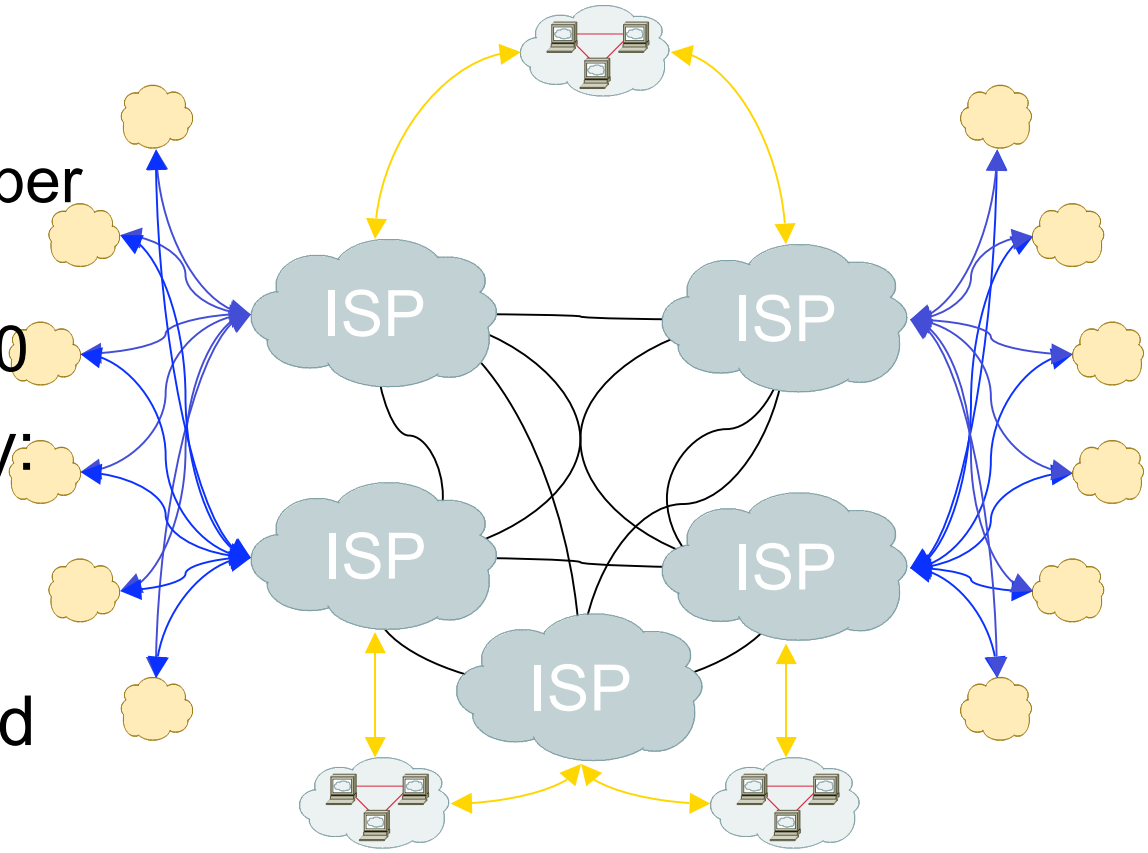
RFC 3582 multihoming requirements



Redundancy	Shields edge from network failures
Address portability	ISP-portable Prefixes
Load sharing	Controlled by edge
Performance	Traffic distributed by edge policy
Policy	Edge network can use any policy
Simplicity	Simple to install/maintain
Transport session survivability	Sessions survive failures
Impact on DNS	No DNS impact
Datagram filtering	Not affected by ISP ingress filtering
Scaling: impact on routers	Route table prefix count
Scaling: impact on hosts	Requires no host changes
Scaling: host/router interaction	No change to Neighbor Discovery etc
Scaling: network management	Simple to monitor/configure
Scaling: ISP cooperation	Requires no ISP cooperation

Present model - PI/PA multihoming

- Current statistics:
  - US: about one multihomed network per 18,000 population
  - World: about 1:50,000
- Expected 2050 density:
  - About 1:1000?
- Implied number of prefixes for multihomed edges:



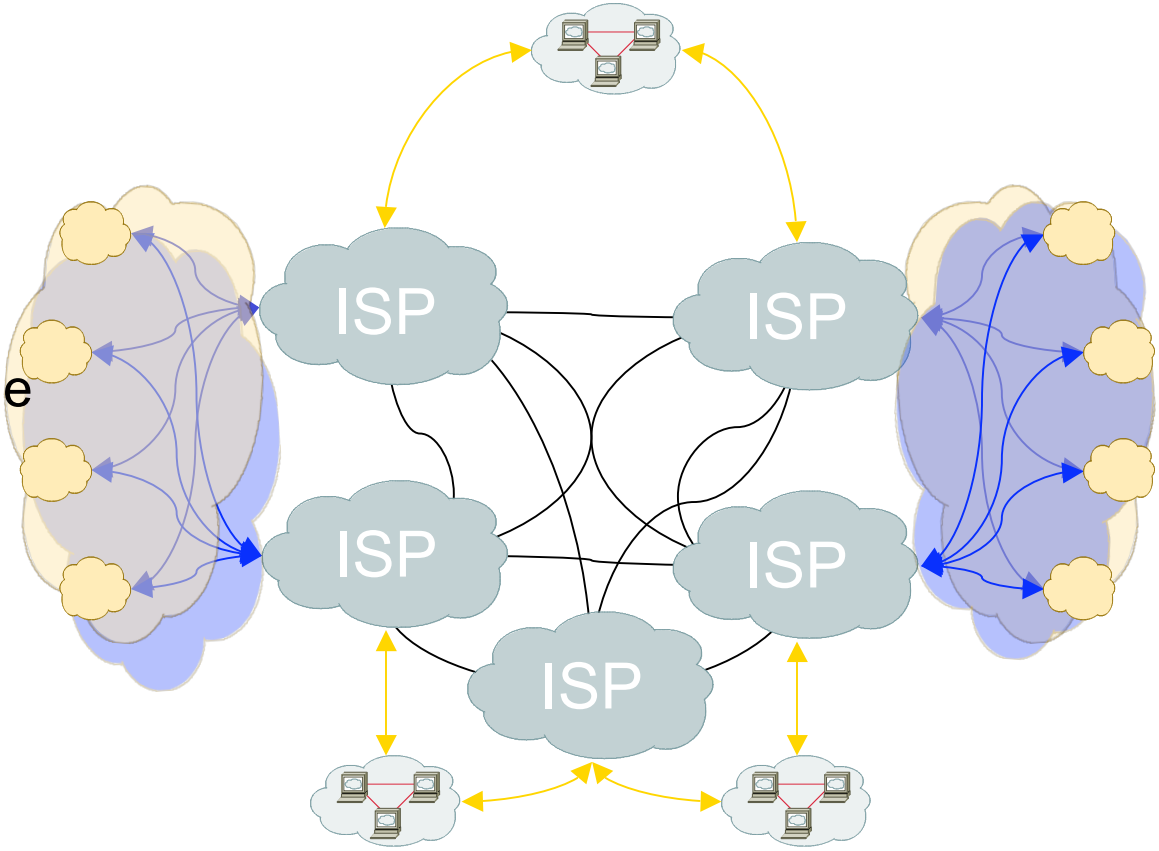
$$\frac{10,000,000,000\ people}{1000\ people/prefix} \approx 10,000,000\ prefixes$$

RFC 3582 analysis of PI/PA multihoming

	PI	PA like PI
Redundancy	✓	✓
Address portability	✓	no
Load sharing	✓	✓
Performance	✓	✓
Policy	✓	✓
Simplicity	✓	✓
Transport session survivability	✓	✓
Impact on DNS	✓	✓
Datagram filtering	✓	Issues
Scaling: impact on routers	$O(10^{7})$ prefixes	$O(10^{7})$ prefixes
Scaling: impact on hosts	✓	✓
Scaling: host/router interaction	✓	✓
Scaling: network management	✓	✓
Scaling: ISP cooperation	✓	Issues

Shim6 viewpoint: PA multihoming

- Premise:
  - ISPs have prefixes
  - Edge networks inherit prefixes from ISPs
  - Only the ISP's prefix is advertised in BGP, not the inherited network prefix
- Prefixes in the internet core:
  - O(tens of thousands of prefixes)
- Prefixes for multihomed edge networks:
  - *ZERO*

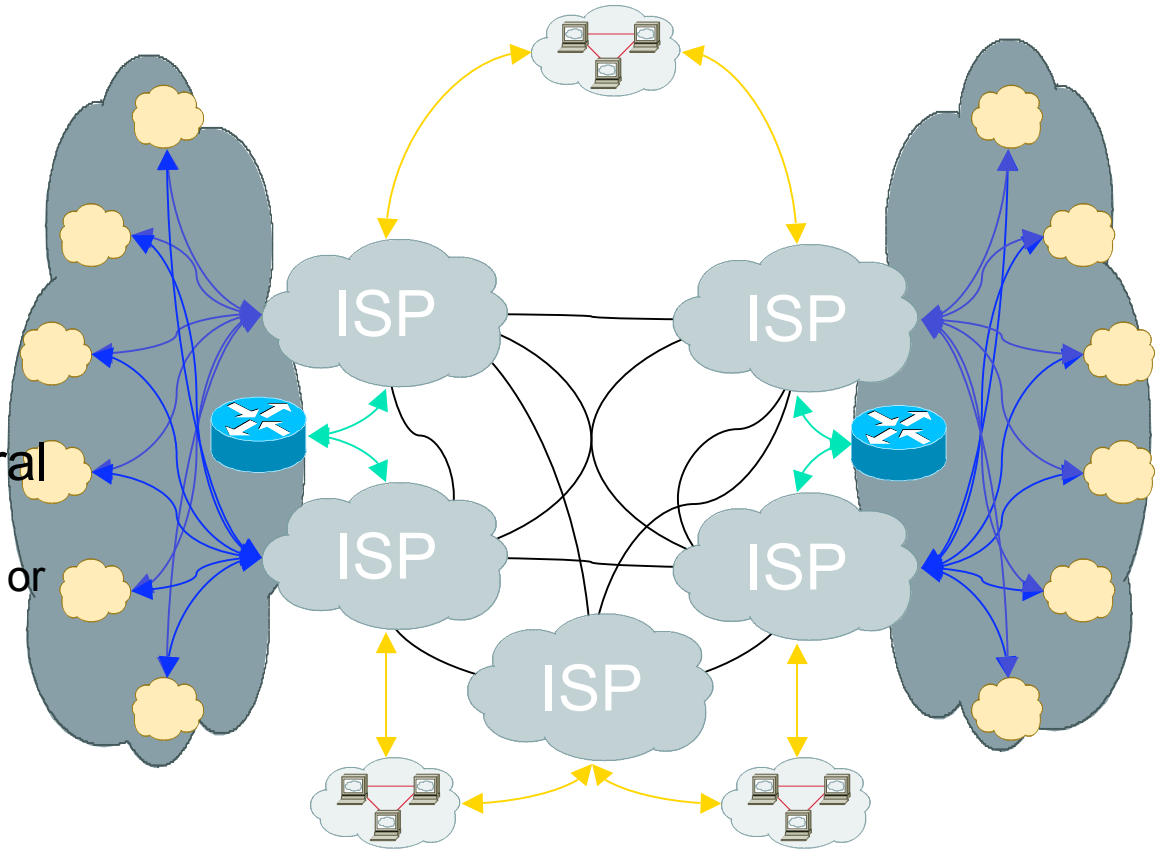


RFC 3582 analysis of shim6 multihoming

Redundancy	Multiple routes
Address portability	Addresses not portable
Load sharing	Host picks route by address pair
Performance	Performance only partially predictable
Policy	Address Pair policy is local
Simplicity	Not as simple as a single prefix
Transport session survivability	SCTP survives; UDP/TCP does not
Impact on DNS	✓
Datagram filtering	Ingress filtering affects routes
Scaling: impact on routers	$O(10^4)$ prefixes
Scaling: impact on hosts	Hosts must select address pair
Scaling: host/router interaction	✓
Scaling: network management	Choice of address pair not controlled in network routing but in host
Scaling: ISP cooperation	✓

Exchange-based multihoming

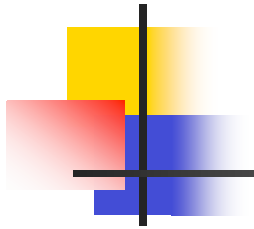
- Imagine:
  - We deploy a prefix for every 1,000,000 people in a regional prefix
    - (Exact number not algorithmically important)
  - Interchange ISP could be government-related or simply an exchange cooperative
- The prefix identifies the general region
  - Delivery is to an ISP's customer or to the regional switch and then to the customer
- Implied number of prefixes for multihomed edges:



$$\frac{10,000,000,000\ people}{1,000,000\ people/exchange} \approx 10,000\ prefixes$$

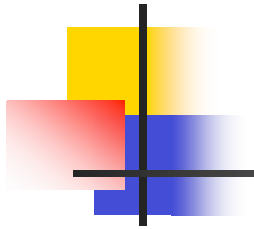
RFC 3582 analysis of exchange-based multihoming

Redundancy	✓
Address portability	Portable within domain
Load sharing	✓
Performance	✓
Policy	✓
Simplicity	✓
Transport session survivability	✓
Impact on DNS	✓
Datagram filtering	✓
Scaling: impact on routers	$O(10^4 - 10^5)$ prefixes
Scaling: impact on hosts	✓
Scaling: host/router interaction	✓
Scaling: network management	✓
Scaling: ISP cooperation	Some form of exchange required



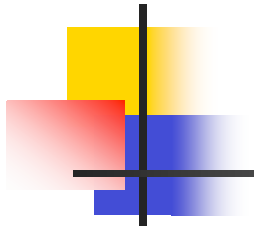
Geographic Addressing

- Inserts latitude and longitude into address
  - Prefixes aggregate with geographic distance, hence route table manageable
  - Business and routing issues not worked out (understatement)
- But see http://tools.ietf.org/html/draft-hain-ipv6-geo-addr



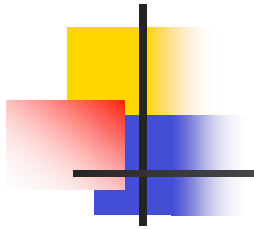
My conclusion

- Nobody at the IETF is saying this
- But it looks like
  - Renumbering is an operational procedure, not a protocol procedure
    - Protocols based on assumptions that aren't valid:
      - "All edge networks use /48 prefixes"
      - "Folks don't reorganize their networks when they renumber them."
  - Operators strongly prefer addressing models that maximize their capital and operational expenditures
    - PA/PI: $O(10^7)$ prefixes in backbone



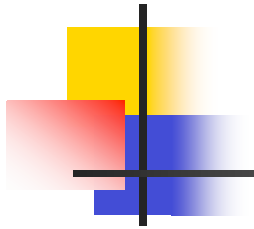
Transition solutions

- IETF ngtrans discussed numerous transition scenarios
  - Dual stack: “bring up IPv6 in your IPv4 network”
  - Overlay: various forms of tunnel mechanisms from static to brokered to implicit, example ISATAP
  - Translation: various mechanisms, example NAT-PT
- Halted discussion and made “Dual Stack” recommendation
  - Reason: discussion was devolving into numerous grad student projects with little operational basis
  - Most solutions with promise published as “Experimental”
  - Some solutions published as “Proposed Standard”
- Now revisiting transition solutions
  - We believe that there is more maturity on the question and nearer term requirement



Comment on NAT-PT

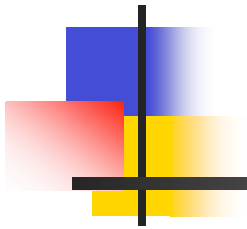
- Historical note on “deprecation”
  - IETF Deprecated RIP with deployment of CIDR
  - Implication: “IETF will not maintain a CIDR-free protocol, we don’t recommend it”
  - That hasn’t stopped anyone from using it, nor did anyone expect it to
- RFC 4966 deprecates NAT-PT and gives rationale
  - Nobody expects that it will therefore not be used
  - The point is that this is not a long term solution and has issues in short term use



Continuing work on translation

- Current work on translation approaches that work better
  - Comcast and NTT comments on transition
  - Various NAT solutions proposed
- Please comment on drafts being developed:
  - draft-bagnulo-v6ops-6man-nat64-pb-statement
  - draft-ietf-v6ops-cpe-simple-security
  - draft-vandevelde-v6ops-cpe-default-route-detection
  - draft-stenberg-v6ops-pd-route-maintenance

  - http:tools.ietf.org/html
  - Comments to v6ops@ops.ietf.org



What Do ISOC and the IETF Think About IPv4 Address Space Exhaustion?

Fred Baker
Member, ISOC Board
Chair, IPv6 Operations WG