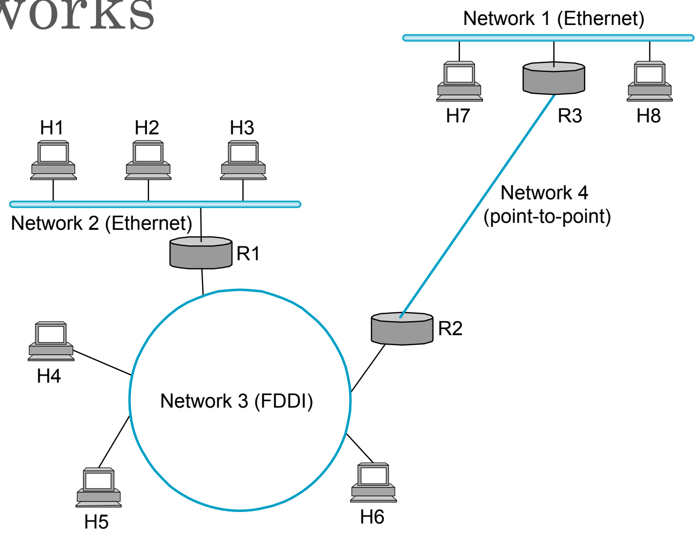
EPL606

Internetworking

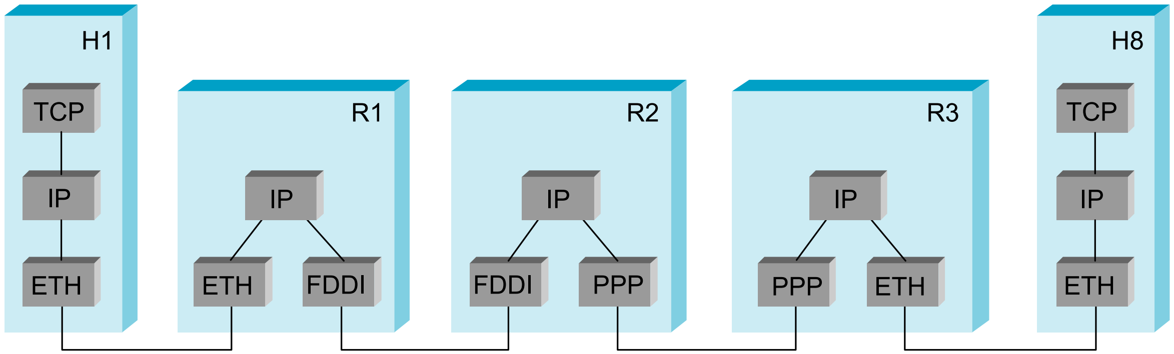
Part 2c

IP Internet

- Concatenation of Networks



- Protocol Stack



Datagram Forwarding

- Strategy
  - every datagram contains destination's address
  - if connected to destination network, then forward to host
  - if not directly connected, then forward to some router
  - forwarding table maps network number into next hop
  - each host has a default router
  - each router maintains a forwarding table

- Example (R2)

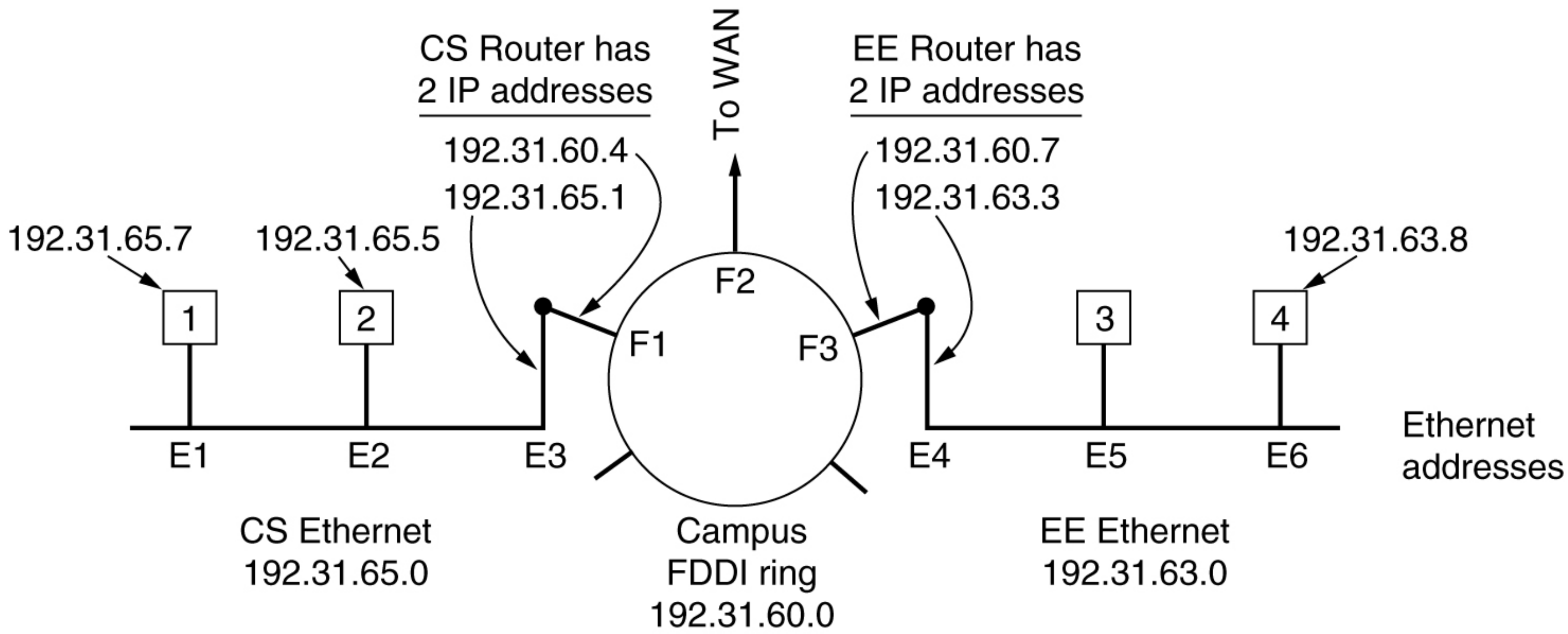
Network Number	Next Hop
1	R3
2	R1
3	interface 1
4	interface 0

Address Translation

- Map IP addresses into physical addresses
  - destination host
  - next hop router
- Techniques
  - encode physical address in host part of IP address
  - table-based
- ARP
  - table of IP to physical address bindings
  - broadcast request if IP address not in table
  - target machine responds with its physical address
  - table entries are discarded if not refreshed

ARP– The Address Resolution Protocol

Three interconnected /24 networks: two Ethernets and an FDDI ring.



ARP Details

- Request Format
  - HardwareType: type of physical network (e.g., Ethernet)
  - ProtocolType: type of higher layer protocol (e.g., IP)
  - HLEN & PLEN: length of physical and protocol addresses
  - Operation: request or response
  - Source/Target-Physical/Protocol addresses
- Notes
  - table entries timeout in about 10 minutes
  - update table with source when you are the target
  - update table if already have an entry
  - do not refresh table entries upon reference

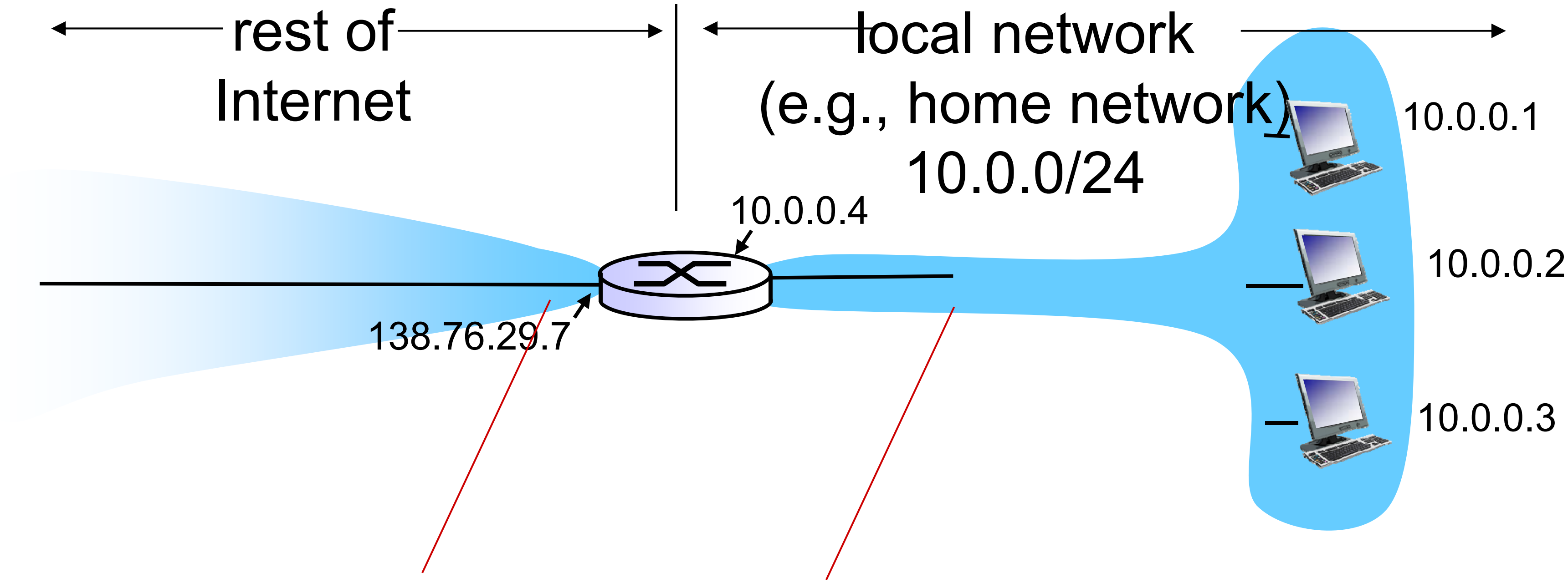
ARP Packet Format

<table>
<tr><th>0</th><th>8</th><th>16</th><th>31</th></tr>
<tr><td colspan="2">Hardware type = 1</td><td colspan="2">ProtocolType = 0x0800</td></tr>
<tr><td>HLen = 48</td><td>PLen = 32</td><td colspan="2">Operation</td></tr>
<tr><td colspan="4">SourceHardwareAddr (bytes 0–3)</td></tr>
<tr><td colspan="2">SourceHardwareAddr (bytes 4–5)</td><td colspan="2">SourceProtocolAddr (bytes 0–1)</td></tr>
<tr><td colspan="2">SourceProtocolAddr (bytes 2–3)</td><td colspan="2">TargetHardwareAddr (bytes 0–1)</td></tr>
<tr><td colspan="4">TargetHardwareAddr (bytes 2–5)</td></tr>
<tr><td colspan="4">TargetProtocolAddr (bytes 0–3)</td></tr>
</table>

Internet Control Message Protocol (ICMP)

- Echo (ping)
- Redirect (from router to source host)
- Destination unreachable (protocol, port, or host)
- TTL exceeded (so datagrams don't cycle forever)
- Checksum failed
- Reassembly failed
- Cannot fragment

NAT: network address translation



all datagrams *leaving* local network have *same* single source NAT IP address: 138.76.29.7,different source port numbers

datagrams with source or destination in this network have 10.0.0/24 address for source, destination (as usual)

NAT: network address translation

motivation: local network uses just one IP address as far as outside world is concerned:

- range of addresses not needed from ISP: just one IP address for all devices
- can change addresses of devices in local network without notifying outside world
- can change ISP without changing addresses of devices in local network
- devices inside local net not explicitly addressable, visible by outside world (a security plus)

NAT: network address translation

implementation: NAT router must:

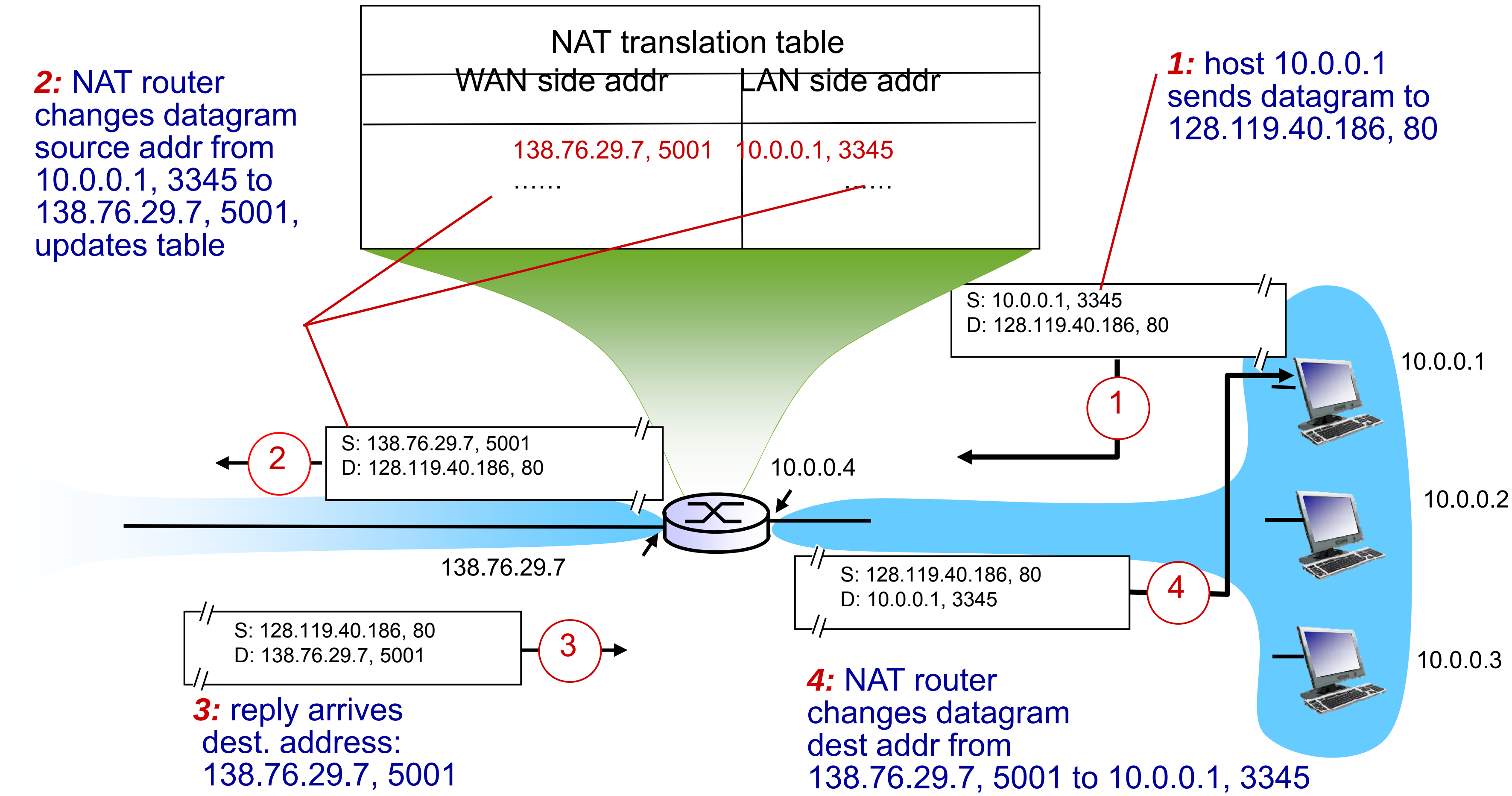
- *outgoing datagrams: replace* (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)

  . . . remote clients/servers will respond using (NAT IP address, new port #) as destination addr

- *remember (in NAT translation table)* every (source IP address, port #) to (NAT IP address, new port #) translation pair

- *incoming datagrams: replace* (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table

NAT: network address translation

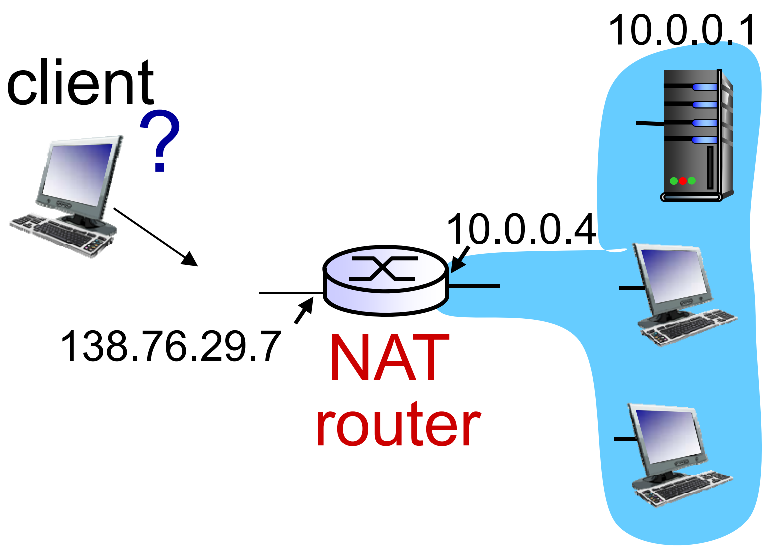


NAT: network address translation

- 16-bit port-number field:
  - 60,000 simultaneous connections with a single LAN-side address!
- NAT is controversial:
  - routers should only process up to layer 3
  - violates end-to-end argument
    - NAT possibility must be taken into account by app designers, e.g., P2P applications
  - address shortage should instead be solved by IPv6

NAT traversal problem

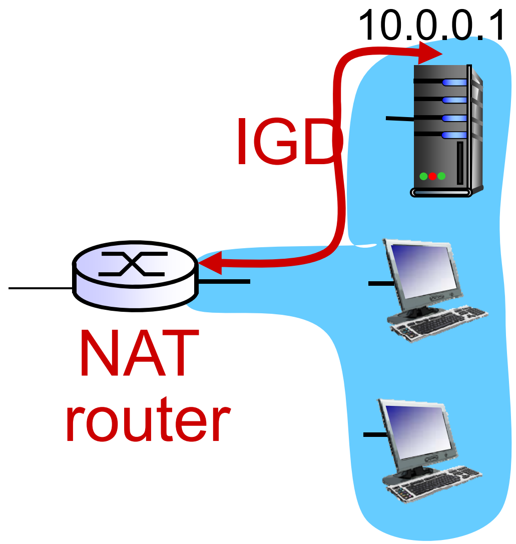
- client wants to connect to server with address 10.0.0.1
  - server address 10.0.0.1 local to LAN (client can't use it as destination addr)
  - only one externally visible NATed address: 138.76.29.7
- *solution1:* statically configure NAT to forward incoming connection requests at given port to server
  - e.g., (123.76.29.7, port 2500) always forwarded to 10.0.0.1 port 25000



NAT traversal problem

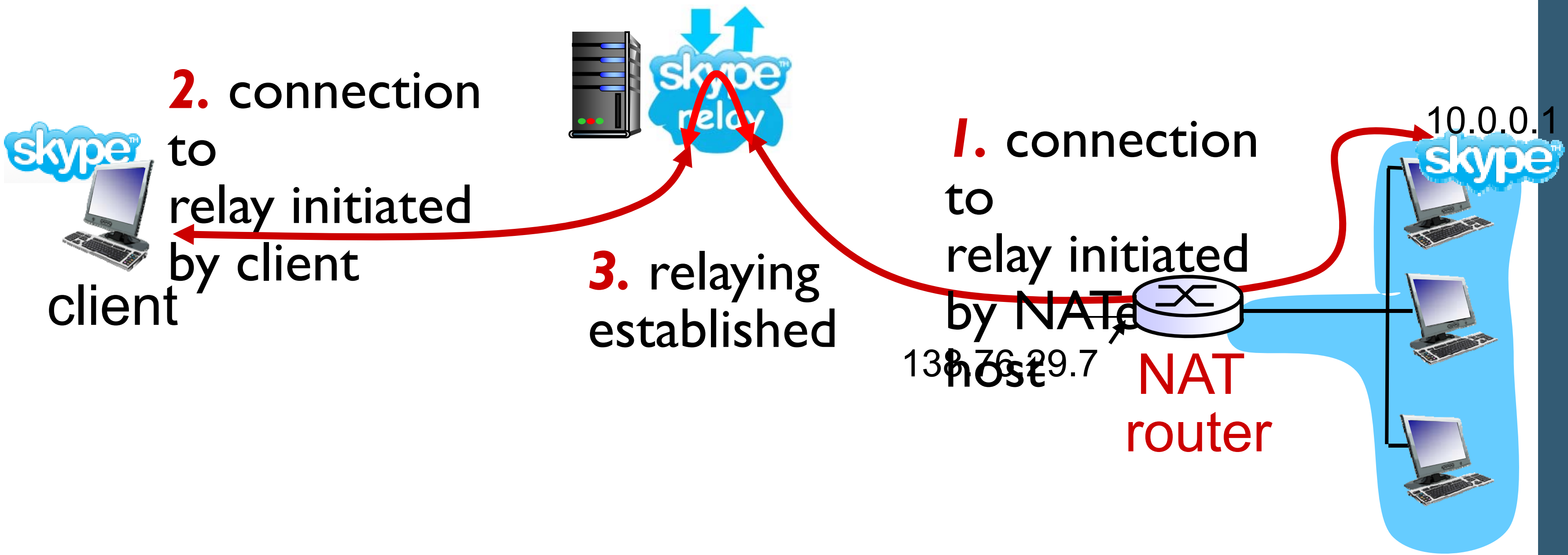
- *solution 2:* Universal Plug and Play (UPnP) Internet Gateway Device (IGD) Protocol. Allows NATed host to:
  - learn public IP address (138.76.29.7)
  - add/remove port mappings (with lease times)

i.e., automate static NAT port map configuration



NAT traversal problem

❖ *solution 3:* relaying (used in Skype)
  - NATed client establishes connection to relay
  - external client connects to relay
  - relay bridges packets between to connections



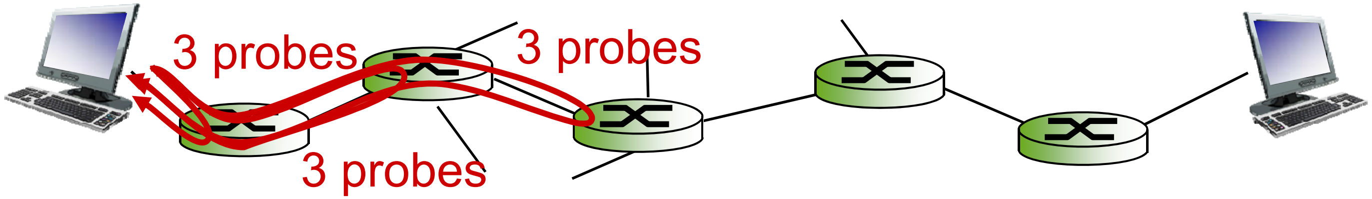
ICMP: internet control message protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP msgs carried in IP datagrams
- ICMP message: type, code plus first 8 bytes of IP datagram causing error

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

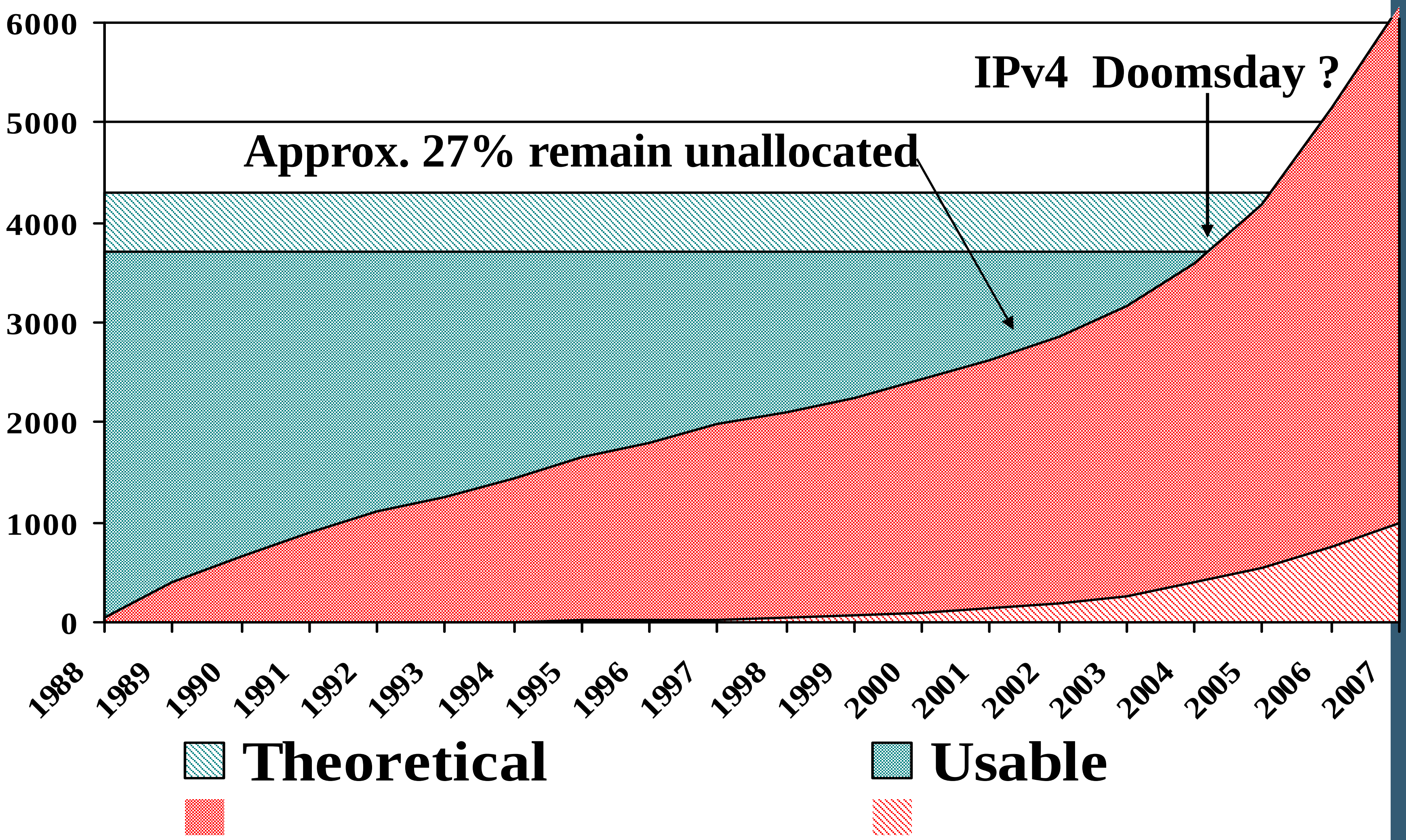
Traceroute and ICMP

- source sends series of UDP segments to dest
  - first set has TTL =1
  - second set has TTL=2, etc.
  - unlikely port number
- when nth set of datagrams arrives to nth router:
  - router discards datagrams
  - and sends source ICMP messages (type 11, code 0)
  - ICMP messages includes name of router & IP address
- when ICMP messages arrives, source records RTTs
- stopping criteria:
  - UDP segment eventually arrives at destination host
  - destination returns ICMP "port unreachable" message (type 3, code 3)
  - source stops



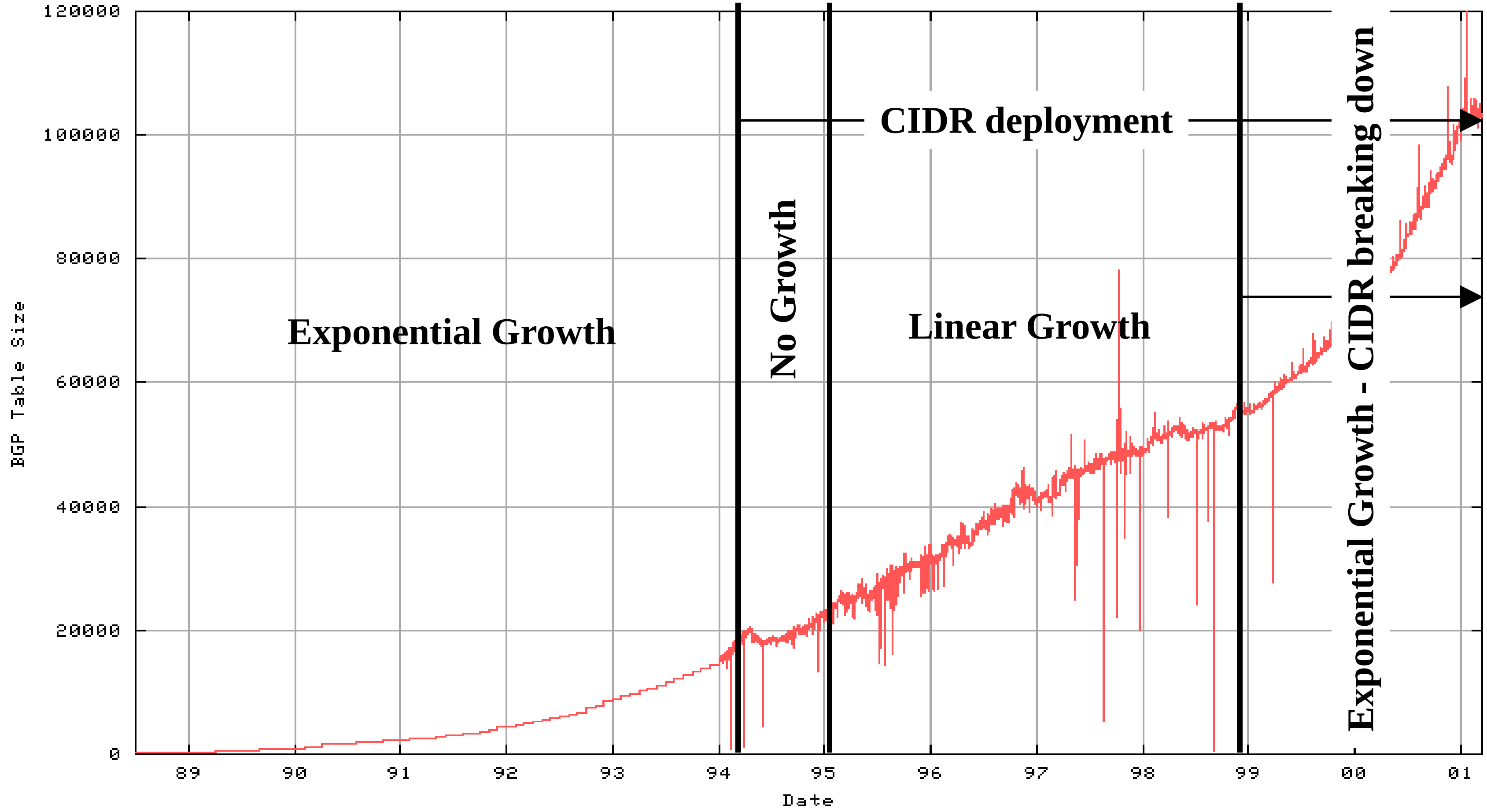
Next Generation IP (IPv6)

Size of the Internet



Distribution Statement A: Cleared for Public Release; Distribution is unlimited.

Internet BGP Routing Table



http://www.telstra.net/ops/bgptable.html

Distribution Statement A: Cleared for Public Release; Distribution is unlimited.

What about technologies & efforts to slow the consumption rate?

- Dial-access / PPP / DHCP
  - Provides temporary allocation aligned with actual endpoint use.
- Strict allocation policies
  - Reduced allocation rates by policy of ‘current-need’ vs. previous policy based on ‘projected-maximum-size’.
- CIDR
  - Aligns routing table size with needs-based address allocation policy. Additional enforced aggregation actually lowered routing table growth rate to linear for a few years.
- NAT
  - Hides many nodes behind limited set of public addresses.

What did intense conservation efforts of the last 5 years buy us?

- Actual allocation history
  - 1981 – IPv4 protocol published
  - 1985 ~ 1/16 total space
  - 1990 ~ 1/8 total space
  - 1995 ~ 1/4 total space
  - 2000 ~ 1/2 total space

- The lifetime-extending efforts & technologies delivered the ability to absorb the dramatic growth in consumer demand during the late 90's.

In short they bought – TIME –

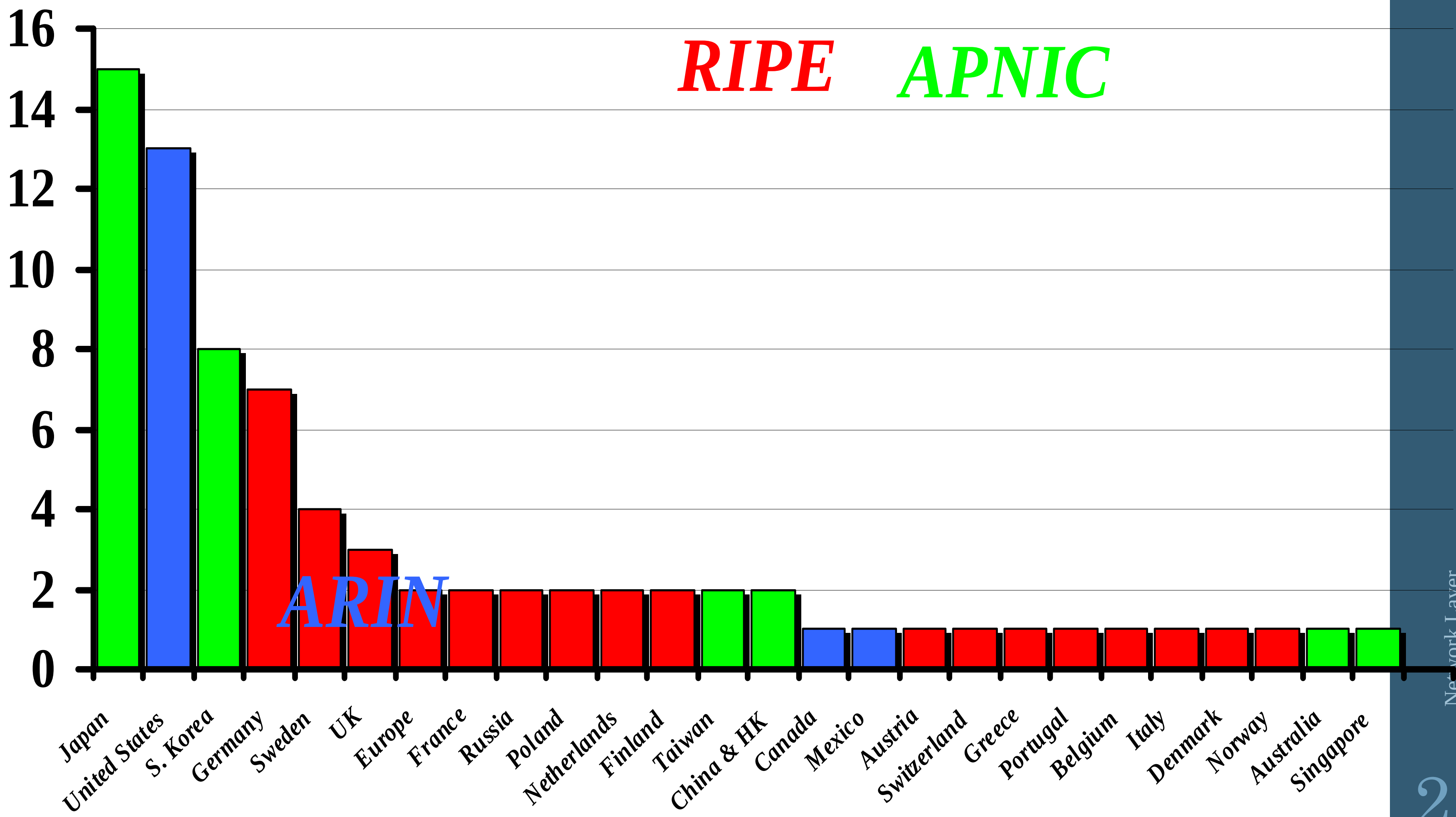
Would increased use of NATs be adequate?

NO!

- NAT enforces a ‘client-server’ application model where the server has topological constraints.
  - ➢ They won’t work for peer-to-peer or devices that are “called” by others (e.g., IP phones)
  - ➢ They inhibit deployment of new applications and services, because all NATs in the path have to be upgraded BEFORE the application can be deployed.

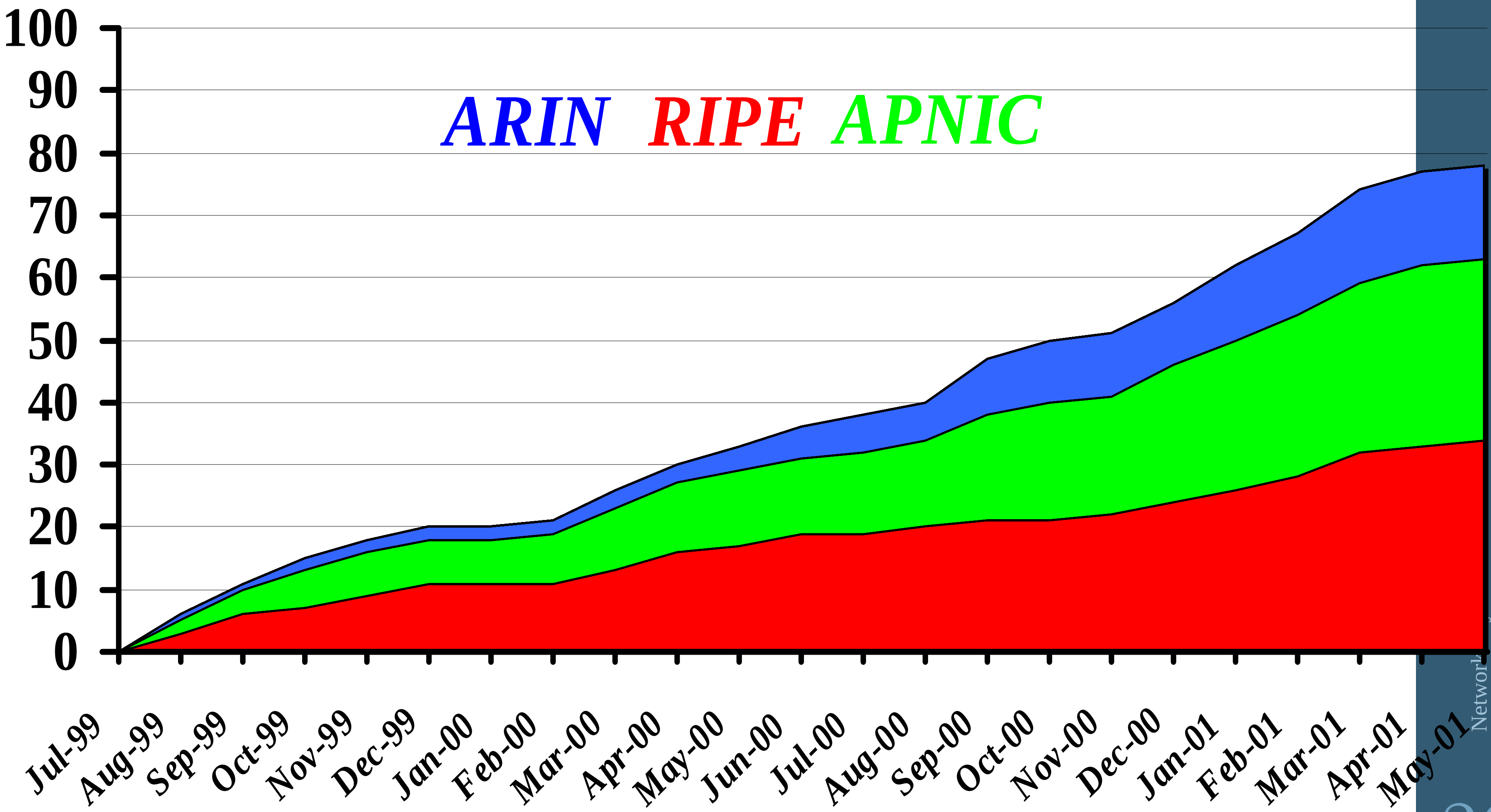
- NAT compromises the performance, robustness, and security of the Internet.
- NAT increases complexity and reduces manageability of the local network.
- Public address consumption is still rising even with current NAT deployments.

78 Top Level IPv6 ISPs in 26 Countries



Distribution Statement A: Cleared for Public Release; Distribution is unlimited.

78 Top Level IPv6 ISPs in 22 months



Distribution Statement A: Cleared for Public Release; Distribution is unlimited.

What Ever Happened to IPv5?

0	IP (deprecated)	March 1977 version	
1	IP	January 1978 version	(deprecated)
2	IP	February 1978 version A	(deprecated)
3	IP	February 1978 version B	(deprecated)
4	IPv4	September 1981 version	(current widespread)
5	ST	Stream Transport	(not a new IP, little use)
6	IPv6	December 1998 version	(formerly SIP, SIPP)
7	CATNIP	IPng evaluation	(formerly TP/IX; deprecated)
8	Pip	IPng evaluation	(deprecated)
9	TUBA	IPng evaluation	(deprecated)
10-15		unassigned	

Benefits of 128 bit Addresses

- Room for many levels of structured hierarchy and routing aggregation
- Easy address auto-configuration
- Easier address management and delegation than IPv4
- Ability to deploy end-to-end IPsec
  (NATs removed as unnecessary)

Incidental Benefits of New Deployment

- Chance to eliminate some complexity in IP header
  - improve per-hop processing
- Chance to upgrade functionality
  - multicast, QoS, mobility
- Chance to include new features
  - binding updates

IPv6 Enhancements (1)

- Expanded address space
  - 128 bit
- Improved option mechanism
  - Separate optional headers between IPv6 header and transport layer header
  - Most are not examined by intermediate routes
    - Improved speed and simplified router processing
    - Easier to extend options
- Address autoconfiguration
  - Dynamic assignment of addresses

IPv6 Enhancements (2)

- Increased addressing flexibility
  - Anycast - delivered to one of a set of nodes
  - Improved scalability of multicast addresses
- Support for resource allocation
  - Replaces type of service
  - Labeling of packets to particular traffic flow
  - Allows special handling
  - e.g. real time video

Summary of Main IPv6 Benefits

- Expanded addressing capabilities
- Structured hierarchy to manage routing table growth
- Serverless autoconfiguration and reconfiguration
- Streamlined header format and flow identification
- Improved support for options / extensions

Types of address

- Unicast
  - Single interface
- Anycast
  - Set of interfaces (typically different nodes)
  - Delivered to any one interface
  - the "nearest"
- Multicast
  - Set of interfaces
  - Delivered to all interfaces identified

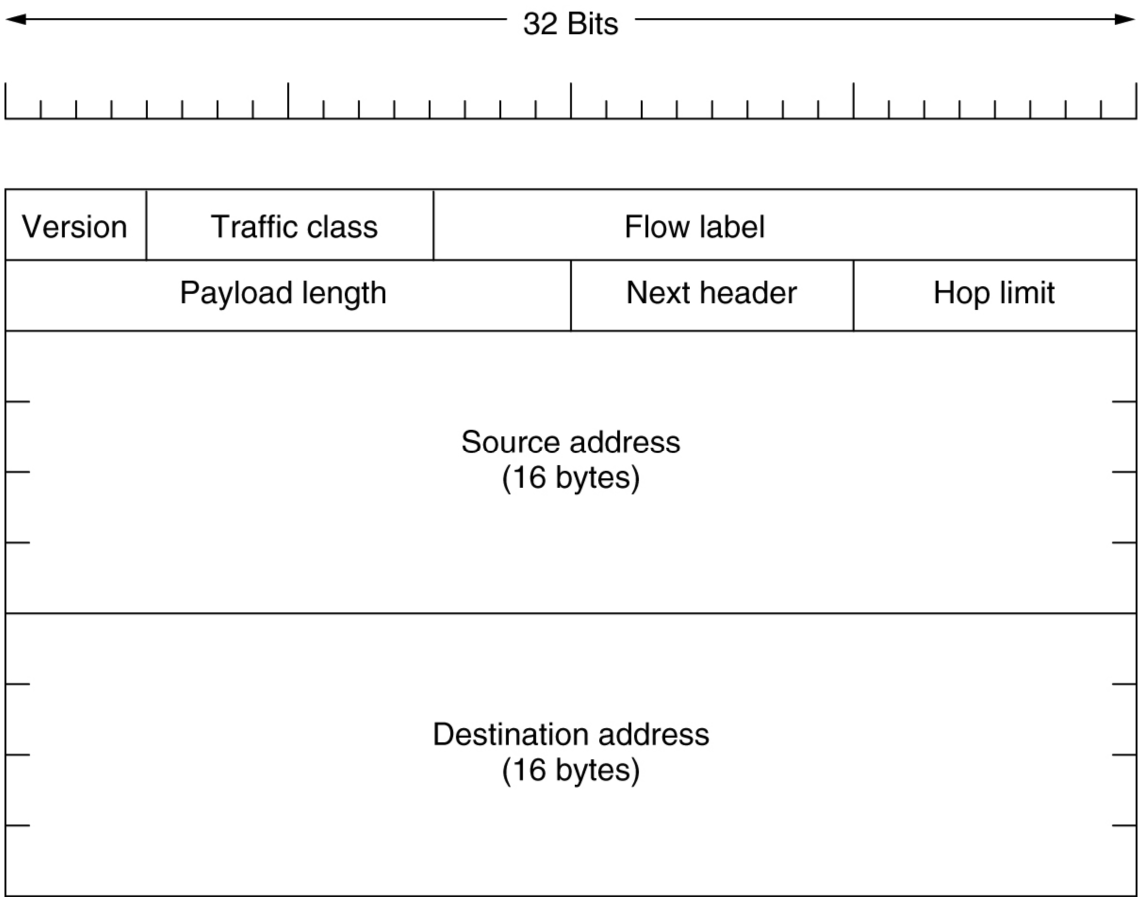
IPv6 Addressing

	n bits	m bits	o bits	p bits	(125-m-n-o-p) bits
010	Registry ID	Provider ID	Subscriber ID	Subnet ID	Interface ID

- Classless addressing/routing (similar to CIDR)
- Notation: x:x:x:x:x:x:x:x (x = 16-bit hex number)
  - contiguous 0s are compressed: 47CD::A456:0124
  - IPv6 compatible IPv4 address: ::128.42.1.87
- Address assignment
  - provider-based (can't change provider easily)
  - geographic

The Main IPv6 Header

- The IPv6 fixed header (required).



IPv6 Header (Cont)

Priority: identify priority among datagrams in flow
Flow Label: identify datagrams in same "flow."
(concept of"flow" not well defined).
Next header: identify upper layer protocol for data

Changes from IPv4

Checksum: removed entirely to reduce processing time at each hop
Options: allowed, but outside of header, indicated by "Next Header" field
ICMPv6: new version of ICMP
- additional message types, e.g. "Packet Too Big"
- multicast group management functions

Extension Headers

Extension header	Description
Hop-by-hop options	Miscellaneous information for routers
Destination options	Additional information for the destination
Routing	Loose list of routers to visit
Fragmentation	Management of datagram fragments
Authentication	Verification of the sender's identity
Encrypted security payload	Information about the encrypted contents

Transition from IPv4 to IPv6

- not all routers can be upgraded simultaneously
  - no "flag days"
  - how will network operate with mixed IPv4 and IPv6 routers?
- tunneling: IPv6 datagram carried as payload in IPv4 datagram among IPv4 routers