Prof. Dr. Hans Peter Großmann mit M. Rabel sowie
H. Hutschenreiter und T. Nau | Sommersemester 2012 |
Institut für Organisation und Management von
Informationssystemen

Thomas Nau, kiz

Lecture Computer Networks

Internet Protocol Version 6

Problems with IPv4

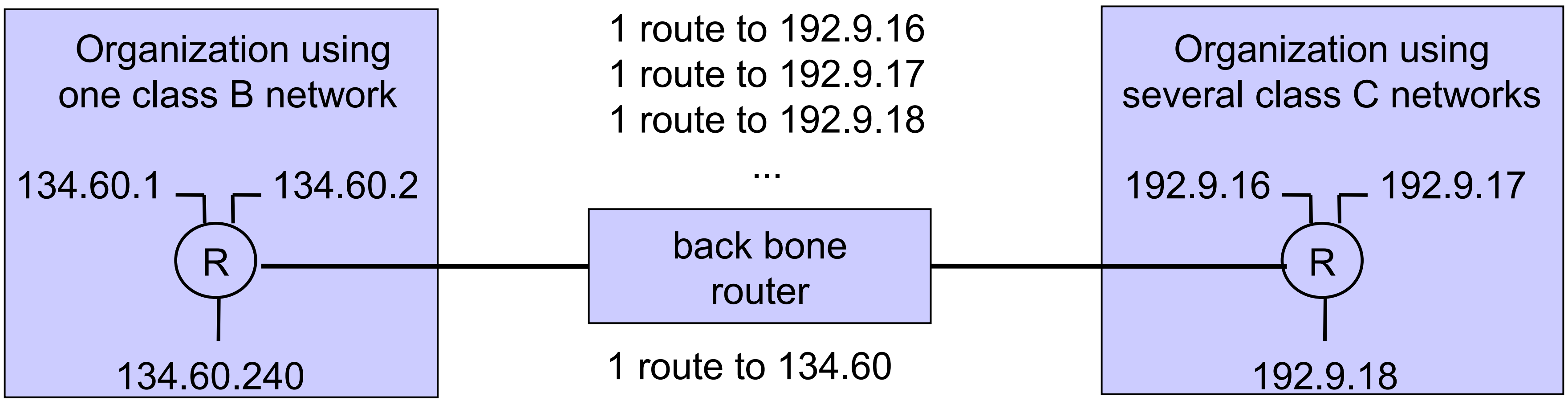
- Running out of 32-bit IP addresses
  - many of the available class B networks are already assigned
- Router problem
  - the backbone routing tables increase in size because of the large number of new networks

Interim Solutions

Classless Inter-Domain Routing (CIDR)

- allocates several class C networks in-order
- uses address/netmask entries for routing to reduce the size of the routing tables
- requires a supporting backbone routing protocol to exchange the necessary information: BGP-4

Without CIDR



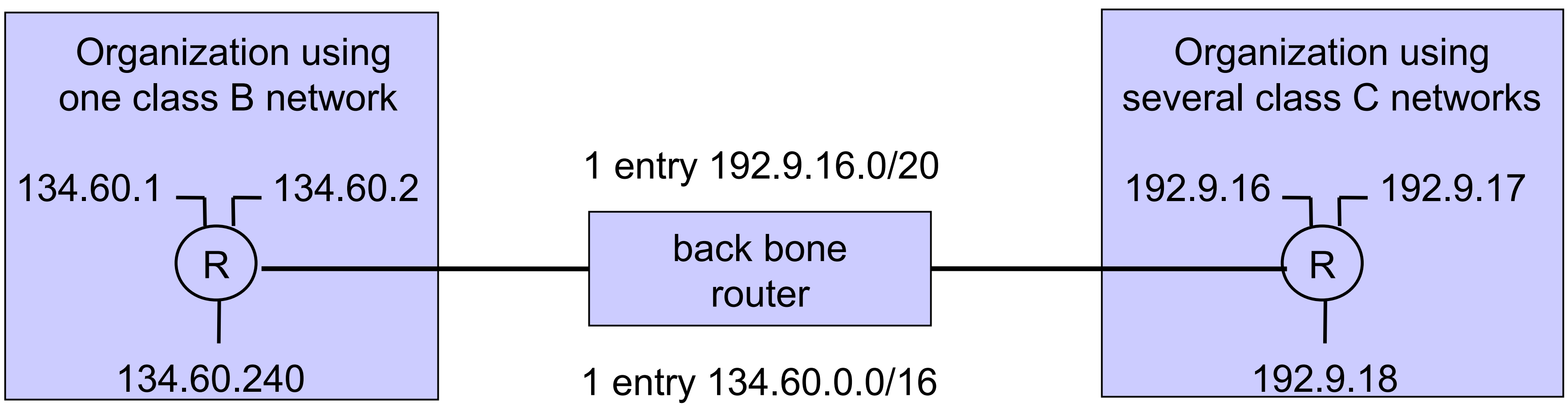
CIDR

Address	Binary
192.9.16 192.9.17 192.9.18 ... 192.9.31	1100 0000 0000 1001 0001 0000 1100 0000 0000 1001 0001 0001 1100 0000 0000 1001 0001 0010 1100 0000 0000 1001 0001 1111
Mask	1111 1111 1111 1111 1111 0000 f f f f f 0

CIDR (cont.)

- Number of entries reduced from 16 to 1
  - 192.9.16.0/0xFFFFF000 or 192.9.16.0/20
- No more implied classes A, B and C
  - classless routing
- Search algorithm uses longest match
- Routing tables used internally by organizations do not change

CIDR (cont.)


Organization using
one class B network
134.60.1
134.60.2
R
134.60.240
1 entry 192.9.16.0/20
back bone
router
1 entry 134.60.0.0/16
Organization using
several class C networks
192.9.16
192.9.17
R
192.9.18

Real Solutions

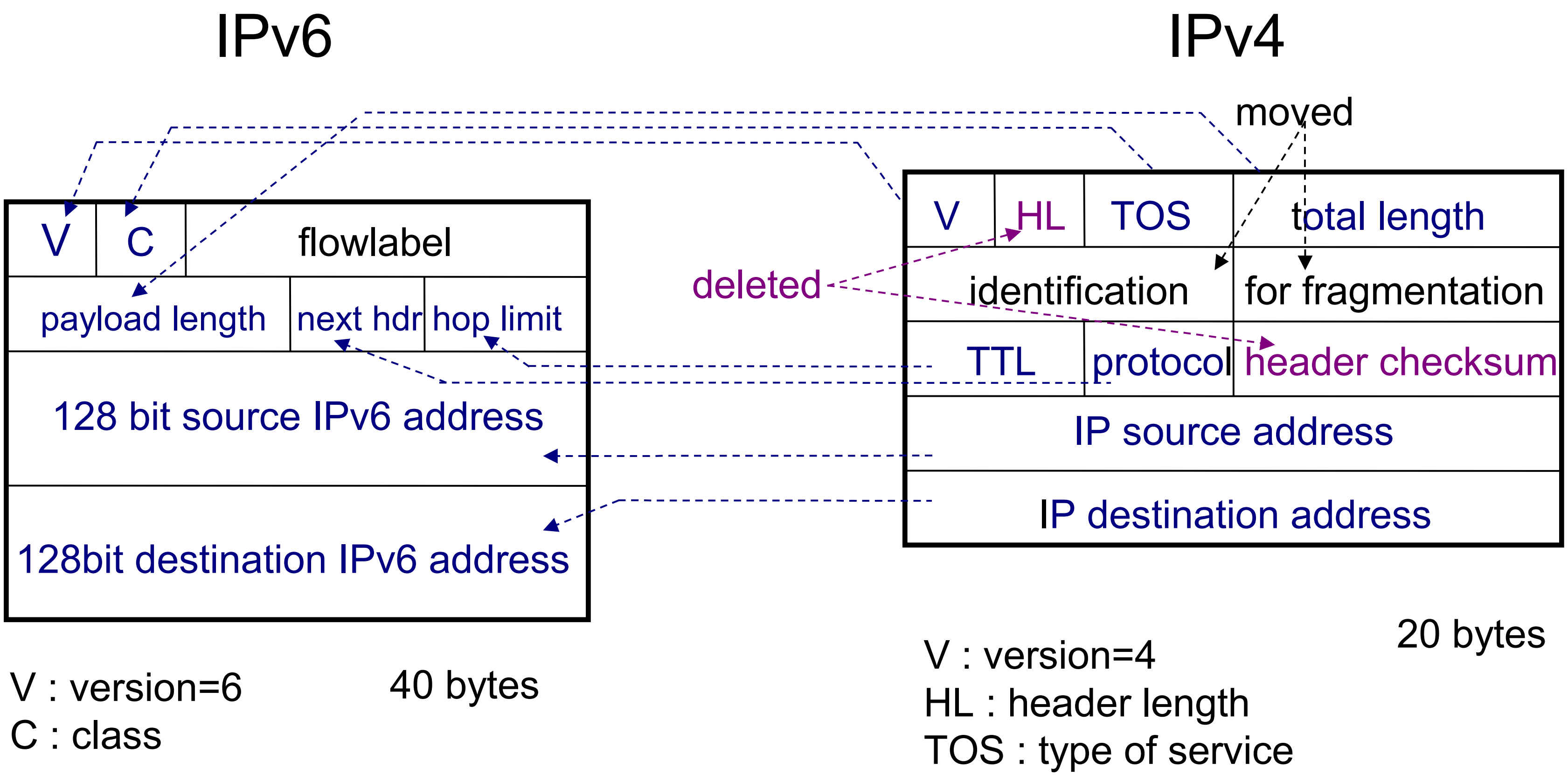
A solution must support:

- a large address space because the number of nodes attached to the Internet doubles about every 12 months
- nomadic computing devices such as laptops, PDAs or cell phones
- security and authentication
- sophisticated routing
- auto-configuration

Major Changes from IPv4 to IPv6

- 128 bit addresses
- Simple, extensible and structured headers (no checksum)
- No broadcasting; multicasting is mandatory
- Only end-to-end fragmentation
- Security (encryption) and authentication options
- Support for plug-and-play, jumbograms and quality of service (flows)

Header Comparison



IPv6 Header Fields

No checksum field

- no need for a checksum because most fields are included in upper-layer pseudo header
  - providing a checksum is mandatory for all upper-layers
- routers do not need to recalculate the checksum
  - results in speed improvement
- addresses are 64-bit aligned
  - speed improvement for modern CPUs which best access memory at 64 bit boundaries

IPv6 Header Fields (cont.)

Payload length

- length of data following the standard 40 byte header
- zero for jumbograms; length has to be specified in an extension header

Hop limit

- named "time-to-live" in v4
- decremented by each forwarding node
- datagrams have to be discarded by forwarding node when counter reaches zero

IPv6 Header Fields (cont.)

Next header

- related to “protocol” in v4
- used to (de)multiplex upper layer protocols and to implement IP options which used to be part of the IPv4 header
- some protocol numbers stay the same (TCP=6, UDP=17)
- also new ones introduced (ICMPv6=58, v4=1) due to the number of changes made to the protocol

IPv6 Header Fields (cont.)

Flow label

- 24 bit flow label is setup by the source
- zero is default
- cannot be changed by destination
- might be used with a resource reservation protocol
- still to be considered an experimental feature

IPv6 Header Fields (cont.)

Class, specified by source

- experimental too
- high-order bit reflects delay sensitivity such as for keystrokes, mouse movements or VoIP data

IPv6 Header Fields (cont.)

Extension headers evaluated by every node

- "hop-by-hop" options
  - must immediately follow IPv6 header to minimize time required to walk through linked list
  - e.g. routing header (similar to IPv4 source routing option) or "Jumbo Payload"

Jumbo Payload Option Header

0	8	16	31
Next header	0	Type=194	Length=4
Jumbo Payload Length			

- “Jumbo Payload” option is a hop-by-hop option
- Upper two bits of “type” define what to do if option is not supported by forwarding node
- Jumbograms require the MTU (maximum transfer unit) to be larger or equal to 2^16 (65536)
- Fragmentation is not supported for jumbograms
- Length field of IPv6 header must be set to zero

IPv6 Header Fields (cont.)

Extension headers evaluated by destination only

- fragmentation header
- authentication header
- security encapsulation header
- destination options header
- upper layer headers (TCP, UDP, ICMPv6)

Fragmentation Header

0	8	16	28	31
Next header	Reserved	Offset	res	M
Identification (32bit, v4 uses 16 bits)				

- Fragmentation in IPv6 is done on an end-to-end basis only
- Routers use ICMPv6 messages to handle oversized packages
  - “package too big” sent to transmitting station including outgoing MTU
- All but the last fragment are a multiple of 8 bytes
- M (“more fragments to come”) bit is set for all but the last fragment

Authentication and Security Headers

Authentication header provides:

- data integrity; modifications done to the packets can be detected (e.g. using MD5 algorithm)
- using digital signatures the receiver can prove that other station did sent data

Security encapsulation header provides:

- confidentiality using cryptography e.g. AES or DES algorithms
- integrity and authentication if used in combination with authentication header

IPv6 Addresses

- DNS type “AAAA” (referred to as “quad-A”) used for 128bit addresses
- BIND (“the standard” DNS server application) supports IPv6
- Reverse lookups are done in the ip6.int domain
  (v4: in-addr.arpa)
- DNS host does not need an IPv6 network stack, just an IPv6 resolver library

IPv6 Addresses (cont.)

- 128bit fields result in
  - 340282366920938463463374607431768211456 = 340×10^{36} possible addresses

- Allocation is not 100% efficient
  - optimistic estimation: approximately 700×10^{21} /m^2
  - pessimistic (RFC 1715): about 1700/m^2 earth surface

IPv6 Addresses (cont.)

Representation

- eight 16bit hex numbers separated by colons
  - e.g. 5800:0:0:0:0:0:56:78
- several zeros in a row may **once** be collapsed
  - 5800::56:78

IPv6 Addresses (cont.)

High-order bits are format prefix (variable length)

0000 0000	reserved (loopback addresses, ...)
0000 001	reserved for NSAP
0000 010	reserved for IPX
001	aggregatable global unicast
1111 1110 10	link-local
1111 1110 11	site-local
1111 1111	multicast

IPv6 Addresses (cont.)

Special addresses

- unspecified address; used whenever own address isn't known yet e.g. as source address during bootstrap
  - 0::0 equal to ::
- loopback address (IPv4: 127.0.0.1)
  - 0::1 equal to ::1

IPv6 Addresses (cont.)

Link-local address

- does never change
- will never be forwarded
- can be configured without additional data
- FE8::<64 bit interface ID>
  - interface ID is typically created from 48 bit MAC address filled up with zeros

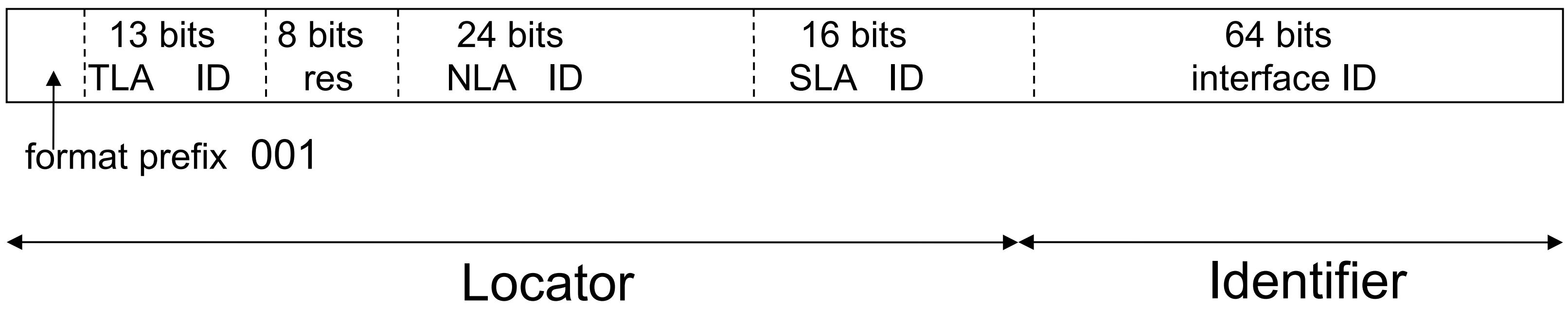
IPv6 Addresses (cont.)

Site-local address

- does also never change
- used by sites not or not yet connected to others
- will never be forwarded by organisational gateway
- routers need an appropriate configuration
- can also be configured without additional data
- FEC::<16 bit subnet><64 bit interface ID>

Aggregatable Global Unicast Address

- Additional prefixes can be assigned out of the reserved bits if the number of TLAs is exhausted



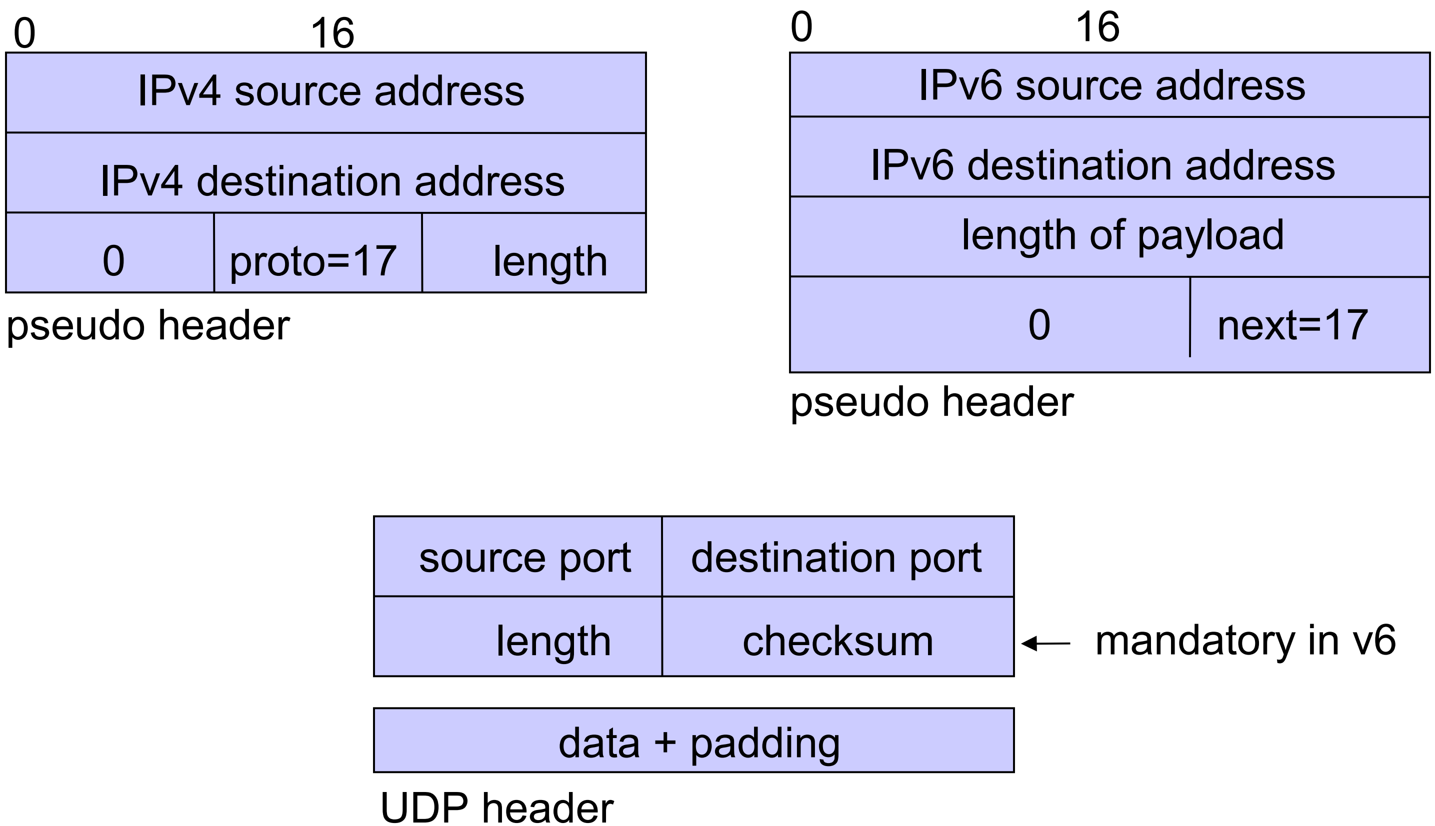
IPv6 Addresses (cont.)

IPv4 related addresses

- IPv4-mapped IPv6 address
  - created on-the-fly by DNS servers if client asks for IPv6 address and no v6 DNS entry exists
  - used for v4-only hosts e.g.: ::FFFF:134.60.1.111.

- IPv4-compatible IPv6 address (now deprecated)
  - v4 address of a v4/v6 node e.g.: ::134.60.1.111.
  - DNS server must be configured appropriately only if no v6 capable neighbor router exists; routing will be done using IPv6-in-IPv4 tunnels
  - used for automatic tunneling; network stack will encapsulate packets automatically

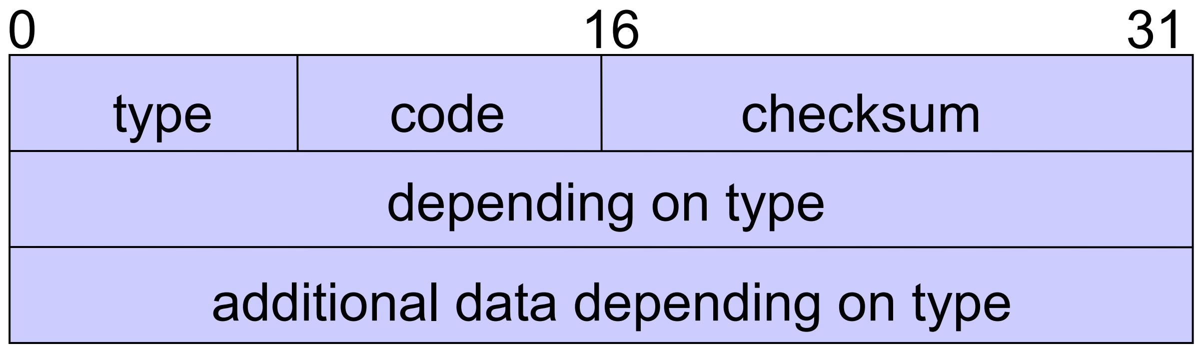
UDP Issues (former optional checksumming)



ICMPv6

- ICMPv6 replaces ICMPv4, IGMPv4 and ARPv4
- Large number of changes has lead to a new “next header” value: 58
- Returns as much as possible of the offending packet without exceeding 576 bytes (minimum link MTU)
- Uses the same pseudo header as UDP and TCP

ICMPv6 (cont.)



Type

0..127:	error messages
128..255:	information only

ICMPv6 (cont.)

Type	Code	Description
1		destination unreachable
	0	no route
	1	administratively prohibited (firewall)
	2	not a neighbour
	3	address unreachable
	4	port unreachable
2	0	packet too big
3		time exceeded
	1	hop limit exceeded
	2	fragment reassembly time exceeded

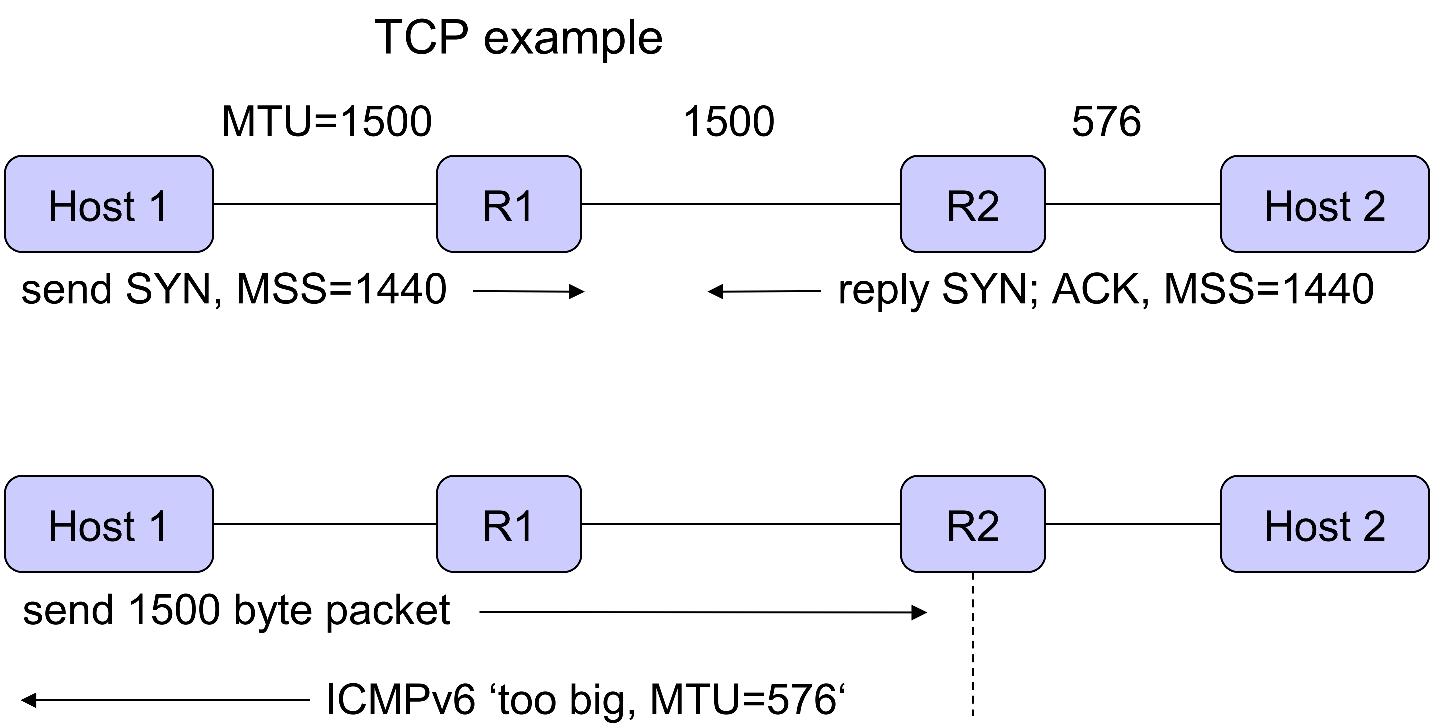
ICMPv6 (cont.)

Type	Code	Description	
128	0	echo request	ping
129	0	echo reply	
130	0	group membership query	multicast
131	0	group membership report	
132	0	group membership reduction	
135	0	neighbour solicitation	ARP
136	0	neighbour advertisement	

Path MTU Discovery

- MTU discovery is mandatory because fragmentation is done on end-to-end basis
- Routers send ICMPv6 message “packet-too-big” if packet size is too large for outgoing link
- MTU size of outgoing link is returned as part of the error message

Path MTU Discovery (cont.)



Solicited Node-Multicast Address

- Node must join the solicited-node multicast group which is used as ARP replacement
- Must do so for every assigned unicast and anycast address

format:

FF02::01:<32 low order bits of interface ID>

multicast

well-known

link-local

Recognized Addresses of v6 Node

Every v6 node must recognize:

router only

loopback address	::1
all-nodes multicast	FF02::1
all-routers multicast	FF02::2
assigned unicast	<subnet>:<Interface ID>
subnet-router anycast	<subnet>::
solicited-node multicast	FF02::1:<32low-order bits NIC ID>
link-local	FE80::<Interface ID>
multicast group	FF<flag+scope>::<group ID>
anycast	like unicast

Plug and Play/Autoconfiguration

IPv6 features using multicast and ICMPv6:

- address autoconfiguration
- detection of duplicate addresses
- address resolution (ARP)
- locator and parameter discovery (subnet, MTU, ...)
- router discovery
- mobile computing

Booting an IPv6 Node

Steps:

- initialize interface and get its hardware address
- setup link-local address using the MAC address
- join solicited-node multicast group
- join all-nodes multicast group
- send ICMPv6 neighbor-solicitation (replaces ARP) for own link-local address
- abort if receiving reply within a second as this means someone is using the same IPv6 link-local address

Booting an IPv6 Node (cont.)

- send router-solicitation to all-routers multicast address
- extract prefix and hierarchical structure from the received router-advertisement packet
- assign unicast address to interface
- default router and the locator are known now
- DHCP (dynamic host configuration protocol) servers can be used to provide add on information beyond the network stack's scope such as IP addresses of DNS servers

Mobile Computing

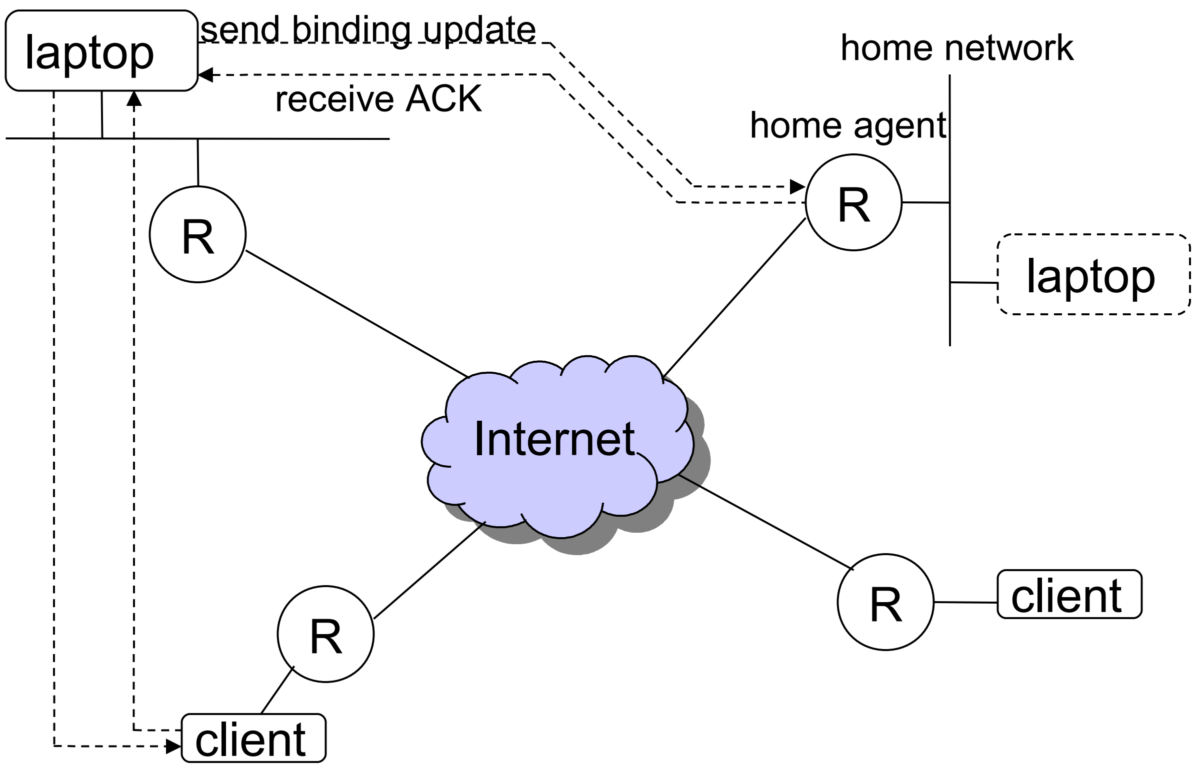
Goal:

- exchanging packets with new location must be transparent

Solution:

- assign additional temporary address using IPv6 autoconfiguration feature (“care-off” address)
- send binding-update to home-agent, using “new-IPv6-destination option”; wait for ACK
- optionally send address update to clients too

Mobile Computing (cont.)



Mobile Computing (cont.)

Updated clients

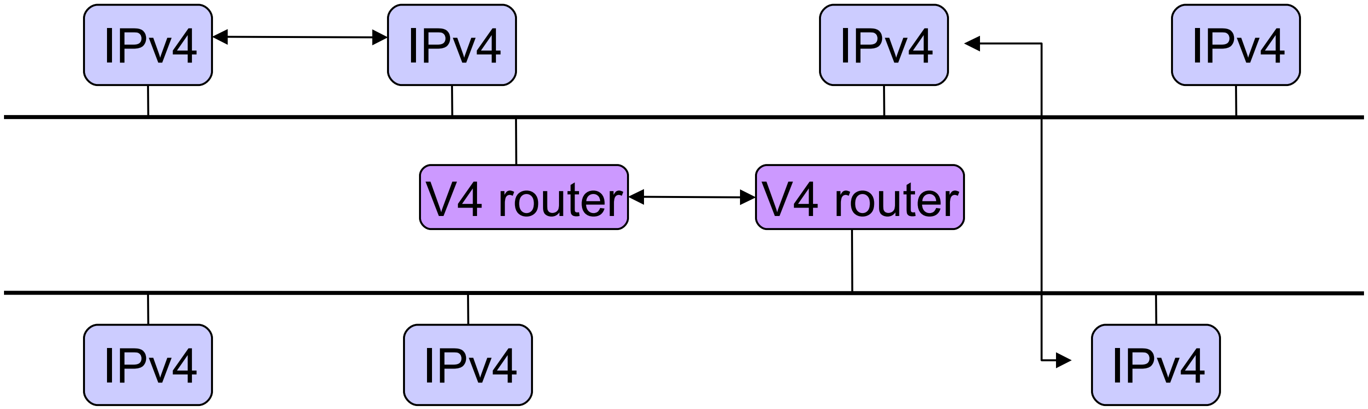
- send packets to home address of laptop using loose source route option via care-of address

Others

- still send packets to home address
- home-agent intercepts and sends them to care-of address using IPv6-in-IPv6 tunnel
- replies are sent to clients directly

Transition Mechanisms

Today's Internet: IPv4 hosts and application



Transition Mechanisms (cont.)

- many more options for different stack combinations but beyond the scope of this lecture
  - dual-stack
  - NAT
  - automatic tunneling
  - ...

Transition Mechanisms (cont.)

Adding IPv6 stacks and application