Comparison of Connectionless Network Layer Protocols

http://webct.rpi.edu/
Or
http://www.ecse.rpi.edu/Homepages/shivkuma/

Shivkumar Kalyanaraman
Rensselaer Polytechnic Institute
shivkuma@ecse.rpi.edu

Based in part upon slides of I. Stoica (UCB)

Forwarding Models

- Connection-oriented:
  - ATM, X.25, frame-relay…
- Connection-less:
  - IP, IPv6
  - CLNP
  - IPX, IPX+
  - Decnet
  - Appletalk
  - Major differences in addressing and related issues: allocation, configuration, resolution, hierarchy…
  - Minor differences in formats/encoding, TTL/hop count, fragmentation etc

Addressing Differences

- Node or interface:
  - IP, IPX, IPv6, Appletalk address *interfaces*
  - CLNP, Decnet: addresses for *nodes*. Nodes w/ multiple interfaces in same area can have single address
- Hierarchy: fixed or variable boundaries
  - *Locator* (network ID) + Host ID
  - IP, IPX, CLNP: arbitrary number of levels
  - Classful IP: fixed boundaries
- Owning vs Renting addresses:
  - Original IP model: own address
  - DHCP, Provider-based addressing, IPv6 address lifetime: rent addresses
  - Rent => renumbering overhead. NAT helps
- Configuration ease: facilitates stateless, easy address resolution/neighbor discovery?

Recall: 7 Things to (auto-) configure…

- 1. End systems need Layer 3 address, names, masks
- 2. Router finds Layer 3 addresses of end systems
- 3. Router finds Layer 2 addresses of end systems
- 4. End systems find a (default) router, name server
- 5. End nodes on the same LAN discover that they can send directly to each other
- 6. End systems find the best router for exit traffic
- 7. End systems communicate on a router-less LAN

- Typically end systems only know their hardware (IEEE 802) address…

Address structures: IP

- 4 bytes, subnet/CIDR mask for flexible boundaries, arbitrary levels. Original: classful; Current: classless
- ARP for address resolution. Small IP address => cannot derive Ethernet address from IP address
- BOOTP/DHCP (stateful configuration). No stateless auto-configuration features
- Addresses centrally assigned; then moved to provider-based + private/NAT model in mid-90s

Flexible boundary: *decided by mask.*
CIDR/supernet-mask used by provider for netID
Subnet mask for intra-AS assignment

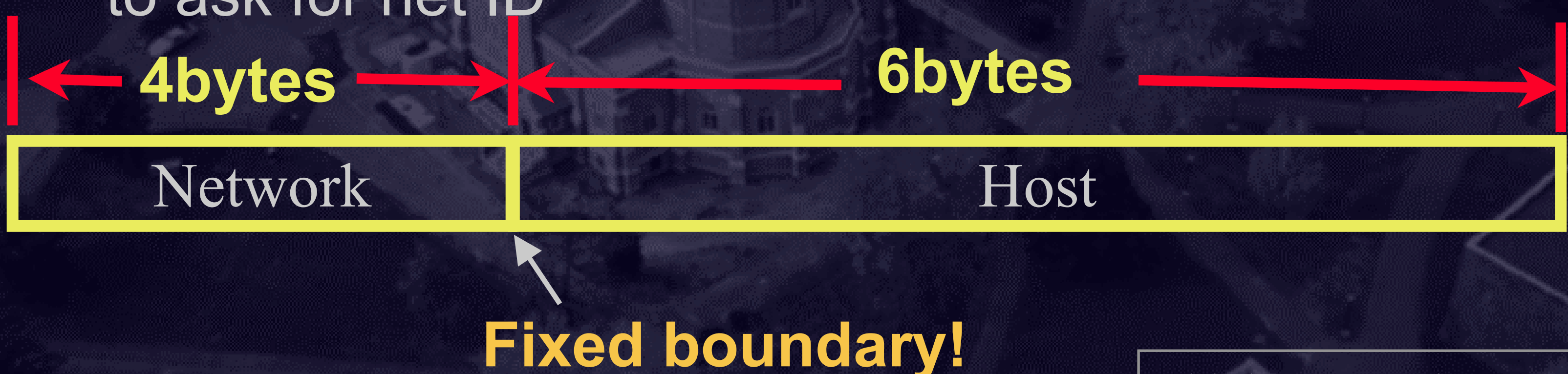
IP Configuration

- 1. End systems: Layer 3 address, names, masks: DHCP
- 2. Router finds Layer 3 addresses of end systems: Same network ID (I.e. IP prefix)
- 3. Router finds Layer 2 addresses of end systems: ARP
- 4. End systems find a default router, name server: DHCP
- 5. End nodes on the same LAN discover that they can send directly to each other: Same network ID + ARP
- 6. End systems find the best router for exit traffic: ICMP Router Redirect
- 7. End systems communicate on a router-less LAN: need a DHCP server at least. Same prefix => same LAN; ARP

- Bottom-line: server necessary for IP auto-configuration on LAN. Server-less not possible.

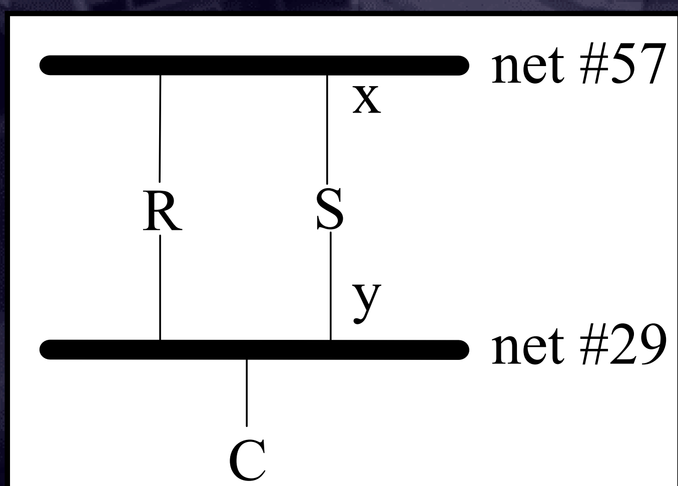
Address Structure: IPX, IPX+

- Internetworking Packet Exchange (IPX)
- IPX: 10 bytes. IPX+: 16 bytes => larger than IP
- Simple structure:
  - IPX: 4B NetID + 6B Node ID.
  - IPX+: Adds 6B Domain ID
  - 6 byte NodeID = IEEE link address => no ARP needed! Address resolution w/o traffic overhead or delays
  - Plug-n-play: Node boots with LAN address, broadcasts to ask for net ID

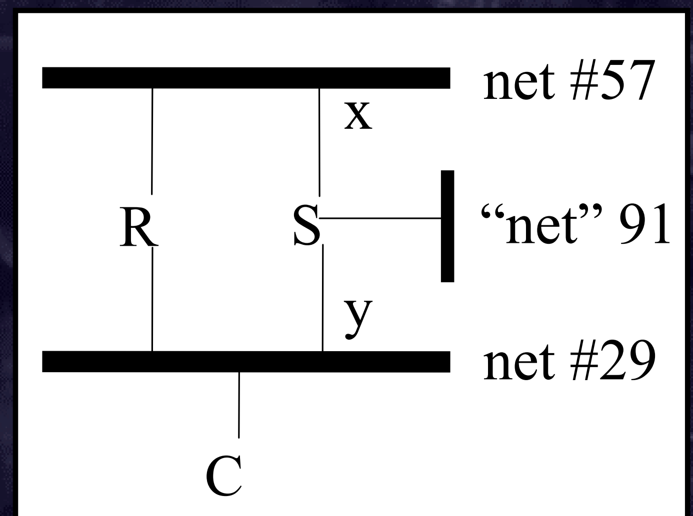


IPX

- No registry => many little IPX nets, non-unique assignments
- Internal network number: servers deplete netIDs to get better routes. Adds configuration overhead. Lousy feature.



=>



Internal network number example for IPX

IPX+

- IPX+: Adds "domain number" in an expanded header
  - Intra-domain routers need not be upgraded
  - NetID FFFC reserved to reach domain boundary
  - Boundary routers then uses expanded header

# of octets	
2	checksum
2	packet length
1	"transport control" (hop count)
1	packet type
4	destination net
6	destination node
2	destination socket
4	source net
6	source node
2	source socket

Figure 10.2 IPX header format

# of octets	
30	regular IPX header, but destination net = "interdomain" and destination socket = "IPX+"
6	destination domain #
6	source domain #
4	real destination network #
2	real destination socket #

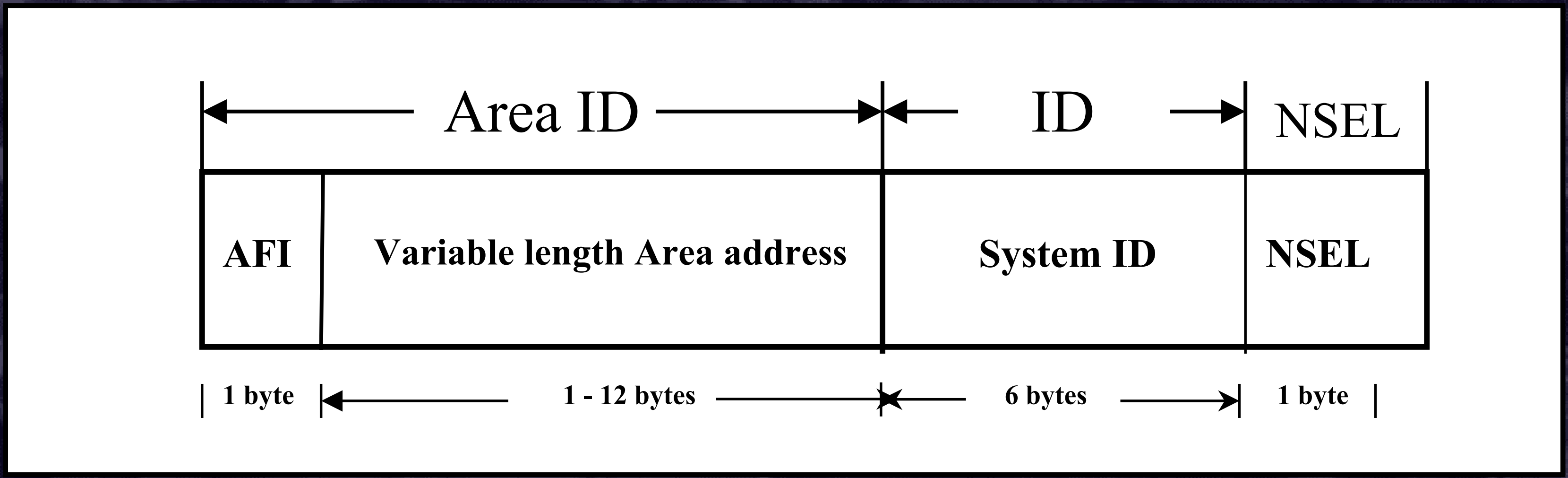
Figure 10.3 IPX+ format

IPX, IPX+ Auto-configuration

- 1. End systems acquire link prefix: snoop/solicit for router advts. L3 address = prefix + IEEE address
- 2. Router finds L3 addr of end systems: Same network ID
- 3. Router finds L2 addr of end systems: nodeID in addr!
- 4. End systems find a default router: solicit for advt
- 5. End nodes on the same LAN send directly to each other: Same network ID=> direct; nodeID gives LAN addr
- 6. End systems find the best router for exit traffic: End node asks for best router before transmission. (weak!)
- 7. End systems communicate on a router-less LAN: Same prefix => same LAN; nodeID = LAN addr; default prefix = 0 also works

- IPX has the simplest server-less auto-configuration solution.

CLNS Addressing: NSAP Format



- NSAP format has 3 main components
  - Area ID: *globally defined* locator
  - System ID: maps to IEEE 802 LAN address usually
  - N-Selector (NSEL): like UDP ports
- Variable length with 20-byte maximum
  - Pkt format needs an address length field!

Address Structure: CLNP…

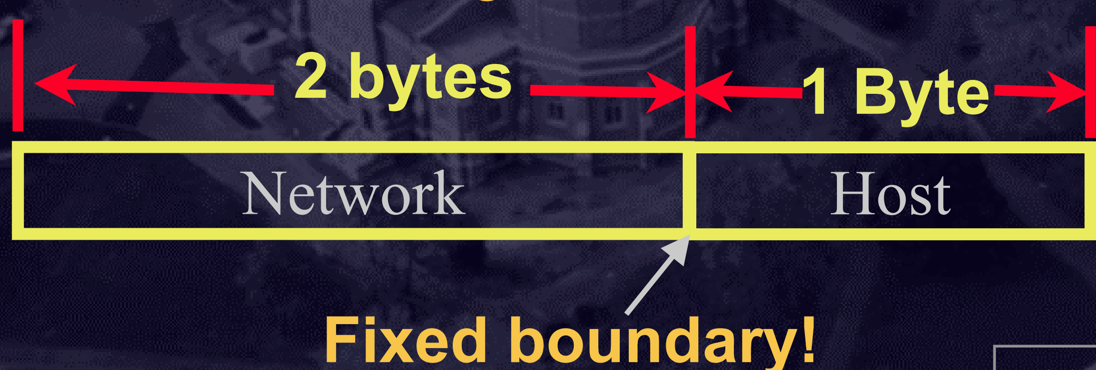
- Between areas, Level 2 routing operates. Many levels of hierarchy possible, just like IP-CIDR.
  - Longest prefix match like IP
- Area larger than single link, all nodes in the area share the same area prefix.
  - Within an area, cannot tell which link (subnet) a node is on, because address is a node-address
  - Advantage: a node can move within area and retain address
  - I.e., no hierarchy in ID field => flat, no topological significance
  - Originally ID: 6 bytes, maps to IEEE address like IPX. But ISO allows this to be variable length too (0-8 bytes)
    - Level 1 routing operates here based upon exact match
    - Bridging in IP provides similar function to level 1 routing
  - Unlike IP cannot use netID or prefix match to decide if destination directly connected => need ES-IS protocol
- Can do cool things like embedding X.25 DTE addresses in area part, and inferring phone-numbers from CLNP addreses!

CLNP Auto-configuration: ES-IS Protocol

- **1,4.** End-node acquire L3 address, and find default router by listening/querying for an hello from routers (IS-Hello).
  - Address = area prefix from router + IEEE address
- **2,3.** Router finds end-node's L3 & L2 address by having end-nodes advertise a ES-hello as part of ES-IS.
  - Unlike IP it cannot look at area-ID and assume direct connectivity
- **5,6.** End-nodes cannot figure out if they are directly connected.
  - So routers send a redirect after forwarding first packet.
  - Redirects are also used to get best exit router.
  - Router, Destination, Neighbor caches like IPv6
- **7.** Routerless LAN: if no router, data packet (not a special ARP-like message) is multicast.
  - Destination replies with LAN address.

Address structure: Appletalk

- Address: 3 bytes long: 2 bytes net ID, 1 byte host
- LAN can have a range of net IDs
  - **Similar to subnet mask, but more flexible. Ranges can start and end on any number, not a power of 2**
  - **Direct connectivity: Don't do AND operation with mask => check if address in range**
- Hosts snoop on received packets to learn best exit router for destinations: no redirects.
- Appletalk does no fragmentation/reassembly



Appletalk Auto-configuration

- 1. End-node acquires L3 address:
  - **Discover router and netID range by snooping** for RIP-like messages or by broadcasting a **query** for one.
  - **Host ID: Randomly choose** an address in range! (cool!)
  - Send message to address hoping ***not*** to get a reply!
- 2. Router finds L3 address of end-node: **same net-ID**
- 3. Router finds L2 address of end-node: **ARP**
- 4. End-nodes find router: **solicit/listen for router traffic**
- 5. End-nodes send directly to each other: **in range => direct**
- 6. Best router discovery: **snoop on received traffic**
- 7. Router-less LAN: **same range => direct. Else default range.**
- **Miscl: Zone** concept to limit name resolution broadcasts
  - Routers on LAN learn range from **seed router** in LAN
- Cutest solution to auto-configuration, and done with short address space!

DECnet Phase IV

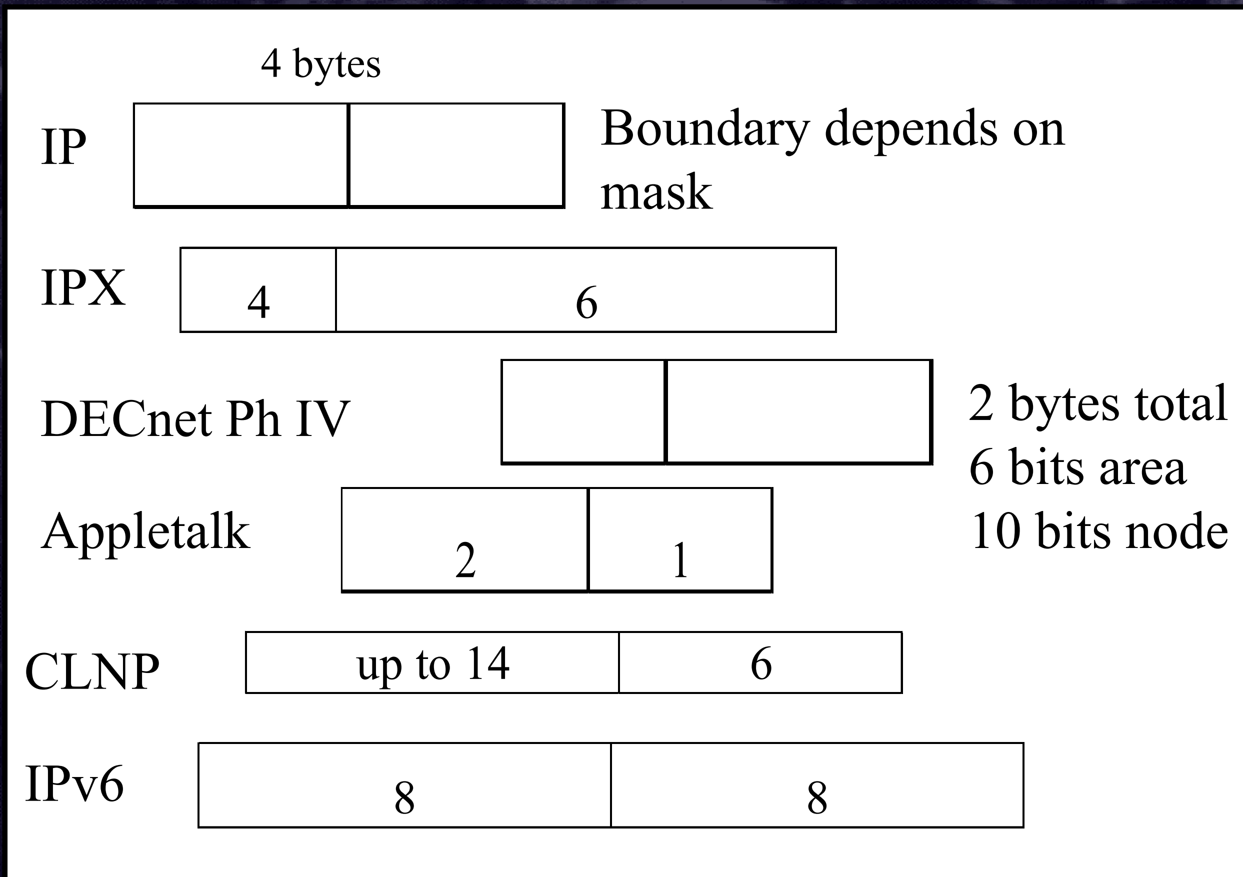
- Was meant as a transition protocol, but CLNP delayed
- 2-byte addresses: 6-bits area, 10-bits node
  - Shortest L3 address among all L3 protocols seen…
- Bold auto-configuration hack:
  - Directly compute 6-byte IEEE address from 2-byte DECnet address!!
  - DEC OUI + 0-byte = AA-00-04-00 (aka HIORD)
  - Program ethernet chips to ignore hardware address and listen to HIORD+DECnet address instead!!
- Like CLNP, address refers to node (not I/f) within area
  - Intra-LAN bit in header to inform receivers of direct connectivity
  - Else one hop through router even for direct case

DECnet auto-configuration

- 1. End nodes get L3 address: manually configured (ugh!)
- 2. Router finds L3 address of end-node: ES-hellos like in CLNP
- 3. Router finds L2 address of end-node: HIORD+L3 address! Bold!
- 4. End-nodes find a router: router (IS) hellos like CLNP
- 5. End-nodes send directly: intra-LAN bit in rcvd packets
- 6. Best-exit router: Learn from rcvd traffic like Appletalk
- 7. Router-less LAN: No problem! HIORD + L3 address!

- Bold solution, with smallest address size.
- Penalty: end-nodes need manual configuration.

Comparison of Address Formats



(Auto-) configuration Techniques

- Manually configure hosts and routers {DECnet}
- Manually configure routers only {IP, IPv6, IPX, Appletalk (seed router), CLNP}
- DHCP server {IP, IPv6 (optional)}
- ARP {IP, Appletalk}
- IEEE address embedded in host-ID {IPX,CLNP,IPv6 (EUI)}
- LAN addr = HIORD + L3 addr {DECnet}
- ES-Hellos and IS-Hellos {CLNP, DECnet}
- Snoop on RIP traffic for router info {Appletalk, IPX}
- Best-exit inferred from rcvd traffic {DECnet, Appletalk}
- Redirects for best-router only (IP, IPv6, IPX)
- Redirects for best-router and direct end-node (CLNP)
- Intra-LAN flag for direct end-node (DECnet)

Packet Formats

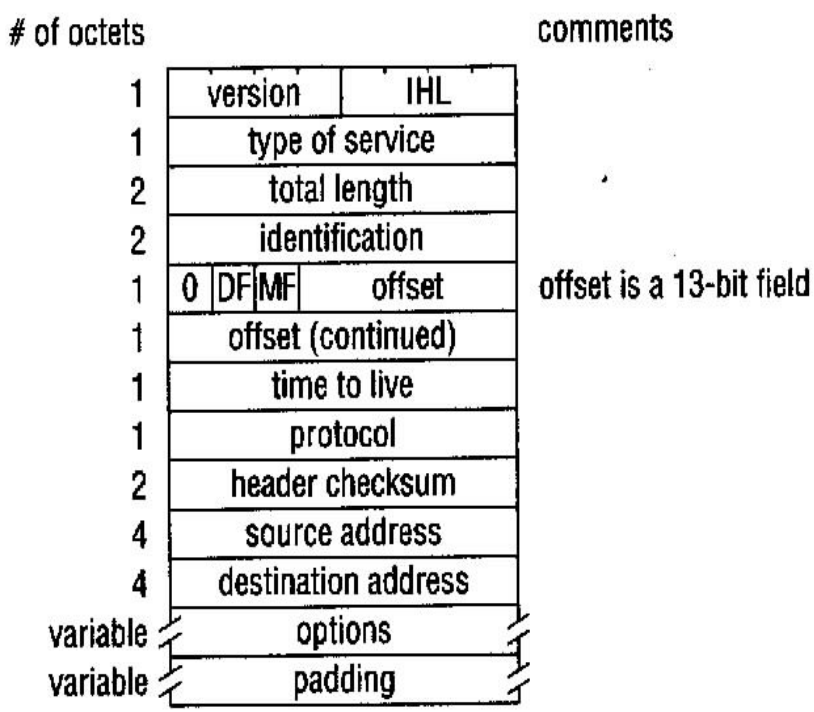


Figure 10.1 IP header format

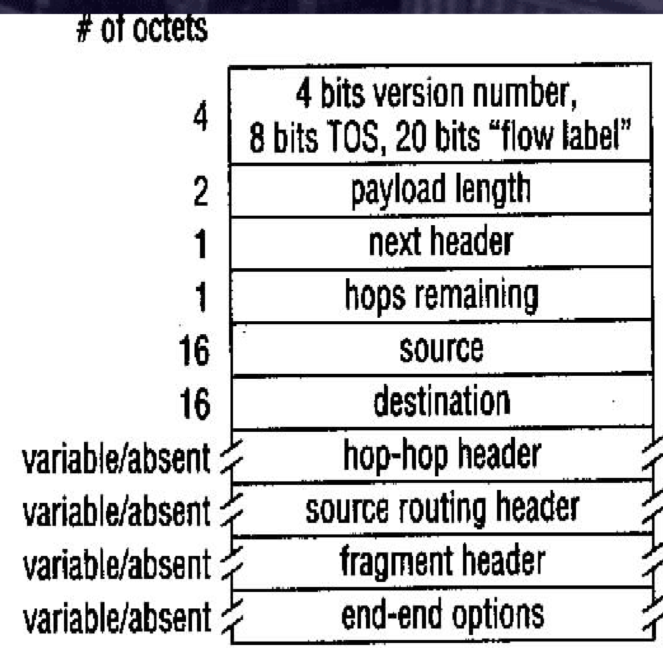


Figure 10.6 IPv6 header format

IP **IPv6**

Similarity: Same core methods

Packet Formats

# of octets	
1	network layer protocol identifier
1	length indicator
1	version/protocol ID extension
1	lifetime
1	SP | MS | E/R | type
2	segment length
2	header checksum
1	destination address length indicator
1-20	destination address
1	source address length indicator
1-20	source address
2 (or not present)	data unit identifier
2 (or not present)	segment offset
2 (or not present)	total length
variable	options

Figure 10.9 CLNP

# of octets	
1	p | v | | rts | rqr | format
2	destination
2	source
1	c | reserved | visit count
1	protocol type

Figure 10.7 DECnet Phase IV short format

# of octets	
1	p | v | i-l | rts | rqr | format
8	destination
8	source
1	reserved | C
1	visit count
2	reserved
1	protocol type

Figure 10.8 DECnet Phase IV long format

CLNP

DECnet, Phase IV

Similarity: Address refers to ***node***

Packet Formats (Contd)

# of octets	
2	checksum
2	packet length
1	"transport control" (hop count)
1	packet type
4	destination net
6	destination node
2	destination socket
4	source net
6	source node
2	source socket

Figure 10.2 IPX header format

# of octets	
30	regular IPX header, but destination net = "interdomain" and destination socket = "IPX+"
6	destination domain #
6	source domain #
4	real destination network #
2	real destination socket #

Figure 10.3 IPX+ format

# of octets	
1	hop count (middle 4 bits)
1	packet length
2	checksum
2	destination net
2	source net
1	destination node
1	source node
2	destination socket
2	source socket
2	protocol type

Figure 10.5 AppleTalk header format

Appletalk

Similarity: Address = ***interface***

Cool auto-configuration

IPX, IPX+

Header Design Issues

- Non-adjacent address components (IPX, IPX+, Appletalk)
- TTL:
  - time (CLNP) vs hop count (IP, IPv6)
  - Counts up (IPX,vs counts down (IP, CLNP)
- UDP-like port space in L3 header vs L4 header
- Small diffs in fragmentation/reassembly (IP, IPv6, CLNP)
  - Don’t care about fragmentation/reassembly (Appletalk, DECnet)
- ICMP functions requested (CLNP, DECnet)
  - ICMP separate protocol (IP, IPv6)
  - No error reporting (IPX, Appletalk)
- Fixed vs Variable length header/fields
- Header checksum (CLNP different algorithm)

Summary

- Addressing and auto-configuration are primary differences in connectionless protocols
- Minor differences in other aspects of header design and forwarding-plane operation