Outline

• IP protocol

• IPv6

• Tunnels
IP Service Model

- Low-level communication model provided by Internet
- Datagram
  - Each packet self-contained
    - All information needed to get to destination
    - No advance setup or connection maintenance
  - Analogous to letter or telegram

IPv4 Packet Format

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>version</td>
<td>HLen</td>
<td>TOS</td>
<td>Length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifier</td>
<td>Flag</td>
<td>Offset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTL</td>
<td>Protocol</td>
<td>Checksum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source Address</td>
<td>Destination Address</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Options (if any)</td>
<td>Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IPv4 Header Fields

- Version: IP Version
  - 4 for IPv4
- HLen: Header Length
  - 32-bit words (typically 5)
- TOS: Type of Service
  - Priority information
- Length: Packet Length
  - Bytes (including header)
- Header format can change with versions
  - First byte identifies version
- Length field limits packets to 65,535 bytes
  - In practice, break into much smaller packets for network performance considerations
IPv4 Header Fields

- Identifier, flags, fragment offset → used for fragmentation
- Time to live
  - Must be decremented at each router
  - Packets with TTL=0 are thrown away
  - Ensure packets exit the network
- Protocol
  - Demultiplexing to higher layer protocols
  - TCP = 6, ICMP = 1, UDP = 17…
- Header checksum
  - Ensures some degree of header integrity
  - Relatively weak – 16 bit
- Source and destination IP addresses
- Options
  - E.g. Source routing, record route, etc.
  - Performance issues
    - Poorly supported

IP Delivery Model

- **Best effort service**
  - Network will do its best to get packet to destination
- Does NOT guarantee:
  - Any maximum latency or even ultimate success
  - Sender will be informed if packet doesn’t make it
  - Packets will arrive in same order sent
  - Just one copy of packet will arrive
- Implications
  - Scales very well (really, it does)
  - Higher level protocols must make up for shortcomings
    - Reliably delivering ordered sequence of bytes → TCP
  - Some services not feasible (or hard)
    - Latency or bandwidth guarantees
IP Fragmentation

- Every network has its own Maximum Transmission Unit (MTU)
  - Largest IP datagram it can carry within its own packet frame
    - E.g., Ethernet is 1500 bytes
  - Don’t know MTUs of all intermediate networks in advance
- IP Solution
  - When hit network with small MTU, router fragments packet
  - Destination host reassembles the paper – why?

Fragmentation Related Fields

- Length
  - Length of IP fragment
- Identification
  - To match up with other fragments
- Flags
  - Don’t fragment flag
  - More fragments flag
- Fragment offset
  - Where this fragment lies in entire IP datagram
  - Measured in 8 octet units (13 bit field)
IP Fragmentation Example #1

- Host to router
- MTU = 4000

Length = 3820, M=0

IP Header

IP Data

IP Fragmentation Example #2

- Router to router
- MTU = 2000

Length = 3820, M=0

IP Header

IP Data

Length = 2000, M=1, Offset = 0

IP Header

IP Data

1980 bytes

Length = 1840, M=0, Offset = 1880

IP Header

IP Data

1820 bytes
Fragmentation is Harmful

- Uses resources poorly
  - Forwarding costs per packet
  - Best if we can send large chunks of data
  - Worst case: packet just bigger than MTU
- Poor end-to-end performance
  - Loss of a fragment

Path MTU discovery protocol → determines minimum MTU along route
- Uses ICMP error messages
- Common theme in system design
  - Assure correctness by implementing complete protocol
  - Optimize common cases to avoid full complexity

Internet Control Message Protocol (ICMP)

- Short messages used to send error & other control information
- Examples
  - Ping request / response
    - Can use to check whether remote host reachable
  - Destination unreachable
    - Indicates how packet got & why couldn't go further
  - Flow control
    - Slow down packet delivery rate
  - Redirect
    - Suggest alternate routing path for future messages
  - Router solicitation / advertisement
    - Helps newly connected host discover local router
  - Timeout
    - Packet exceeded maximum hop limit
IP MTU Discovery with ICMP

- Typically send series of packets from one host to another
- Typically, all will follow same route
  - Routes remain stable for minutes at a time
- Makes sense to determine path MTU before sending real packets
- Operation: Send max-sized packet with "do not fragment" flag set
  - If encounters problem, ICMP message will be returned
    - "Destination unreachable: Fragmentation needed"
    - Usually indicates MTU problem encountered
- ICMP abuse? Other solutions?

MTU = 4000
MTU = 1500
MTU = 2000

IP Packet
IP MTU Discovery with ICMP
IP MTU Discovery with ICMP

- When successful, no reply at IP level
  - "No news is good news"
- Higher level protocol might have some form of acknowledgement
Important Concepts

- Base-level protocol (IP) provides minimal service level
  - Allows highly decentralized implementation
  - Each step involves determining next hop
  - Most of the work at the endpoints
- ICMP provides low-level error reporting
- IP forwarding → global addressing, alternatives, lookup tables
- IP addressing → hierarchical, CIDR
- IP service → best effort, simplicity of routers
- IP packets → header fields, fragmentation, ICMP

Outline

- IP protocol
- IPv6
- Tunnels
IPv6

- “Next generation” IP.
- Most urgent issue: increasing address space.
  - 128 bit addresses
- Simplified header for faster processing:
  - No checksum (why not?)
  - No fragmentation (?)
- Support for guaranteed services: priority and flow id
- Options handled as “next header”
  - reduces overhead of handling options

### IPv6 Addressing

- Do we need more addresses? Probably, long term
  - Big panic in 90s: “We’re running out of addresses!”
- 128 bit addresses provide space for structure (good!)
  - Hierarchical addressing is much easier
  - Assign an entire 48-bit sized chunk per LAN – use Ethernet addresses
  - Different chunks for geographical addressing, the IPv4 address space,
  - Perhaps help clean up the routing tables - just use one huge chunk per ISP and one huge chunk per customer.
IPv6 Autoconfiguration

- Serverless ("Stateless"). No manual config at all.
  - Only configures addressing items, NOT other host things
    - If you want that, use DHCP.
- Link-local address
  - 1111 1110 10 :: 64 bit interface ID (usually from Ethernet addr)
    - (fe80::/64 prefix)
  - Uniqueness test ("anyone using this address?")
  - Router contact (solicit, or wait for announcement)
    - Contains globally unique prefix
    - Usually: Concatenate this prefix with local ID \(\rightarrow\) globally unique IPv6 ID
- DHCP took some of the wind out of this, but nice for "zero-conf"
  (many OSes now do this for both v4 and v6)

Fast Path versus Slow Path

- Common case: Switched in silicon ("fast path")
  - Almost everything
- Weird cases: Handed to CPU ("slow path", or "process switched")
  - Fragmentation
  - TTL expiration (traceroute)
  - IP option handling
- Slow path is evil in today’s environment
  - "Christmas Tree" attack sets weird IP options, bits, and overloads router.
  - Developers cannot (really) use things on the slow path
    - Slows down their traffic – not good for business
    - If it became popular, they’d be in the soup!
### IPv6 Header Cleanup: Options

- 32 IPv4 options → variable length header
  - Rarely used
  - No development / many hosts/routers do not support
    - Worse than useless: Packets w/options often even get dropped!
  - Processed in “slow path”.
- IPv6 options: “Next header” pointer
  - Combines “protocol” and “options” handling
    - Next header: “TCP”, “UDP”, etc.
  - Extensions header: Chained together
  - Makes it easy to implement host-based options
  - One value “hop-by-hop” examined by intermediate routers
    - E.g., “source route” implemented only at intermediate hops

### IPv6 Header Cleanup: “no”

- No checksum
  - Motivation was efficiency: If packet corrupted at hop 1, don’t waste b/w transmitting on hops 2..N.
  - Useful when corruption frequent, b/w expensive
  - Today: corruption is rare, bandwidth is cheap
- No fragmentation
  - Router discard packets, send ICMP “Packet Too Big” → host does MTU discovery and fragments
  - Reduced packet processing and network complexity.
  - Increased MTU a boon to application writers
  - Hosts can still fragment - using fragmentation header. Routers don’t deal with it any more.
Migration from IPv4 to IPv6

- Interoperability with IP v4 is necessary for incremental deployment.
- Combination of mechanisms:
  - Dual stack operation: IP v6 nodes support both address types
  - Tunnel IP v6 packets through IP v4 clouds
  - IPv4-IPv6 translation at edge of network
    - NAT must not only translate addresses but also translate between IPv4 and IPv6 protocols
  - IPv6 addresses based on IPv4 – no benefit!
  - More on NATs and tunnels in the next lecture

Outline

- IP protocol
- IPv6
- Tunnels
Motivation

There are many cases where not all routers have the same features or consistent state

- An experimental IP feature is only selectively deployed – how do we use this feature end-end?
  - E.g., IP multicast
- A few are using a protocol other than IPv4 – how can they communicate?
  - E.g., incremental deployment of IPv6
- I am traveling with a CMU laptop - how can I keep my CMU IP address?
  - E.g., must have CMU address to use services

Tunneling

- Force a packet to go to a specific point in the network.
  - Cannot rely on routers on regular path
- Achieved by adding an extra IP header to the packet with a new destination address.
  - Similar to putting a letter in another envelope
  - Preferable to IP source routing
- Used increasingly to deal with special routing requirements or new features.
  - Mobile IP...
  - Multicast, IPv6, research, ..
IP-in-IP Tunneling

- Described in RFC 1993.
- IP source and destination address identify tunnel endpoints.
- Protocol id = 4.
  - IP
  - Several fields are copies of the inner-IP header.
    - TOS, some flags, ...
  - Inner header is not modified, except for decrementing TTL.

### Tunneling Example

```
A → B
C → D
E → F
G → F
H → I
J → K
```

```
a → b
e → f
j → k
```

Payload

```
A → K
C → F
A → K
```

```
Payload
```

```
A → K
```

```
Payload
```
Tunneling Applications

- Virtual private networks.
  - Connect subnets of a corporation using IP tunnels
  - Often combined with IP Sec (later)
- Support for new or unusual protocols.
  - Routers that support the protocols use tunnels to “bypass” routers that do not support it
  - E.g. multicast, IPv6 (!)
- Force packets to follow non-standard routes.
  - Routing is based on outer-header
  - E.g. mobile IP (later)

Extending Private Network

- Supporting Road Warrior
  - Employee working remotely with assigned IP address 198.3.3.3
  - Wants to appear to rest of corporation as if working internally
    - From address 10.6.6.6
    - Gives access to internal services (e.g., ability to send mail)
- Virtual Private Network (VPN)
  - Overlays private network on top of regular Internet
Supporting VPN by Tunneling

- Idea: client sets up tunnel to company’s firewall
- Example: client wants to send packet to internal node 10.1.1.1
- Entering Tunnel
  - Add extra IP header directed to firewall (243.4.4.4)
  - Original header becomes part of payload
  - Possible to encrypt it
- Exiting Tunnel
  - Firewall receives packet
  - Strips off header
  - Sends through internal network to destination

Overlay Networks

- A network “on top of the network”.
  - E.g., initial Internet deployment
    - Internet routers connected via phone lines
    - An overlay on the phone network
  - Tunnels between nodes on a current network
  - Examples: IPv6 “6bone”, multicast “Mbone”.
  - But not limited to IP-layer protocols…
    - Peer-to-peer networks, anonymising overlays
    - Application layer multicast
    - Improve routing, e.g. work around route failures
Important Concepts

- IP has a very simple service model
- IPv4 is a simple protocol, but there are issues
  - 32 bit address space is too small
  - Some messy features, e.g., fragmentation
  - Very simple “control” protocol
- NATs change to Internet addressing model
  - Have moved away from “everyone knows everybody” model of original Internet
- Firewalls + NAT hide internal networks
- VPN / tunneling build private networks on top of commodity network