15-441 15-641 Computer Networking

Lecture 6 – The Internet Protocol

Peter Steenkiste

Fall 2016

www.cs.cmu.edu/~prs/15-441-F16

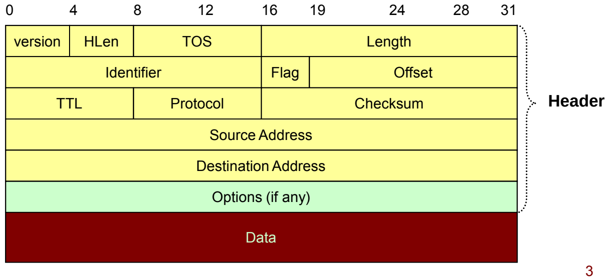
Outline

- The IP protocol
  - IPv4
  - IPv6
- IP in practice
  - Network address translation
  - Address resolution protocol
  - Tunnels

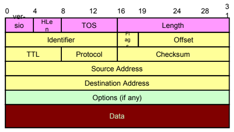
IP Service Model

- 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



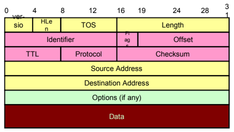
IPv4 Header Fields



- 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

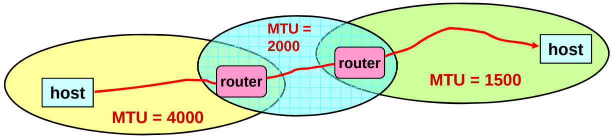
- Identifier, flags, fragment offset → used for fragmentation
- Time to live
  - Must be decremented at each router
  - Packets with TTL=0 are thrown away
  - Ensure packets exit the network
- Protocol
  - Demultiplexing to higher layer protocols
  - TCP = 6, ICMP = 1, UDP = 17…
- Header checksum
  - 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
  - Informing the sender if packet does not make it
  - Delivery of packets in same order as they were 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

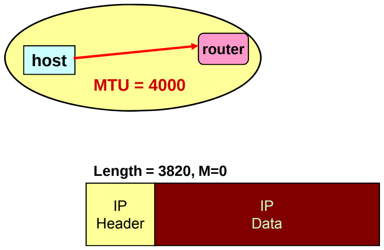


- Every network has own Maximum Transmission Unit (MTU)
  - Largest IP datagram it can carry within its own packet frame
    - E.g., Ethernet is 1500 bytes
  - Don't know MTUs of all intermediate networks in advance
- IP Solution
  - When hit network with small MTU, router fragments packet
  - Destination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

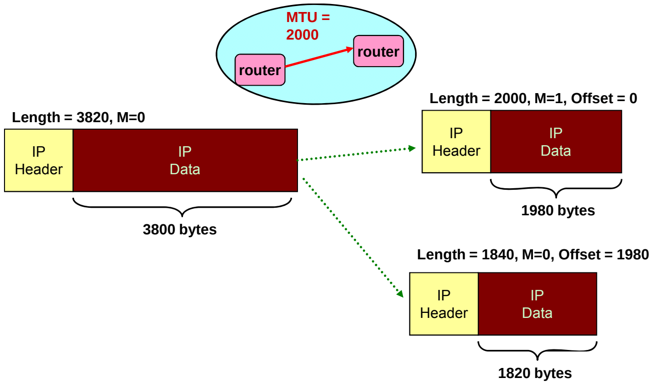
Fragmentation Related Fields

- 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



IP Fragmentation Example #2



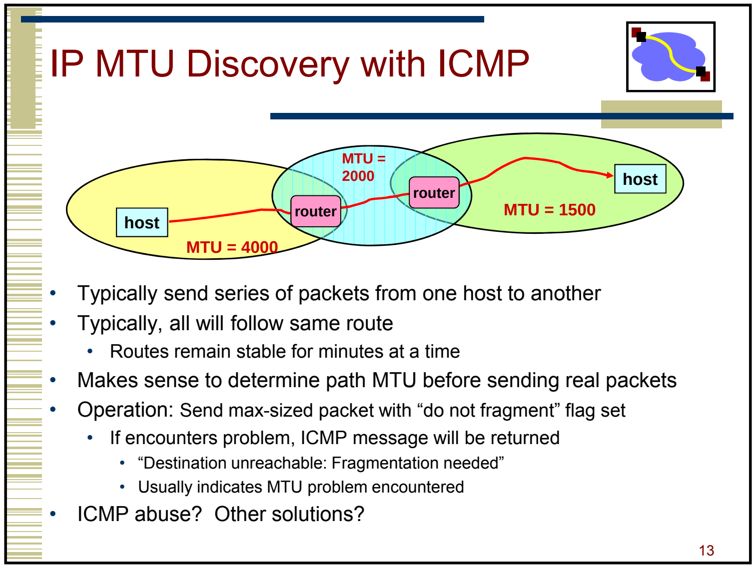
Fragmentation is Harmful

- 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

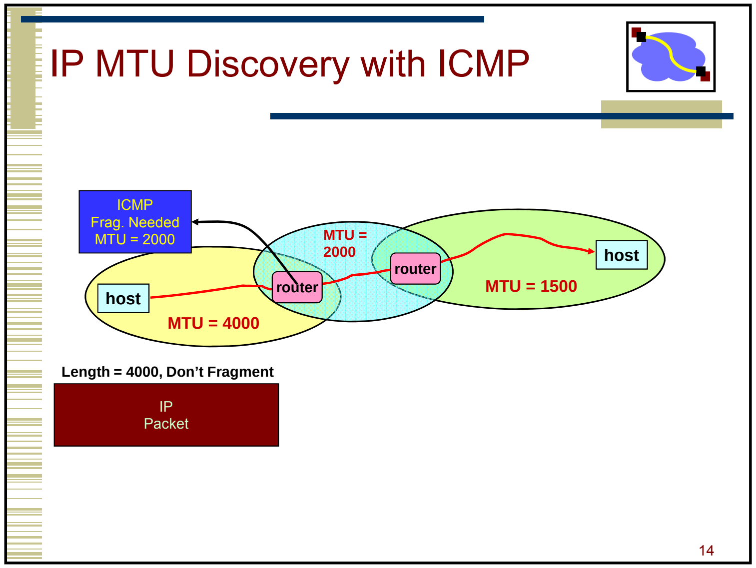
Internet Control Message Protocol (ICMP)

- Short messages used to send error & other control information
- Some functions supported by ICMP:
  - Ping request /response: check whether remote host reachable
  - Destination unreachable: Indicates how packet got & why couldn't go further
  - Flow control: Slow down packet transmit 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
- How useful are they functions today?

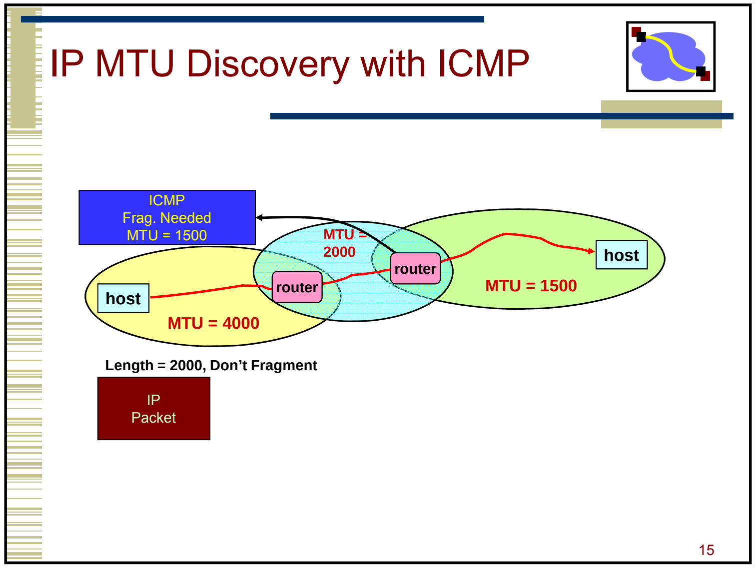


IP MTU Discovery with ICMP

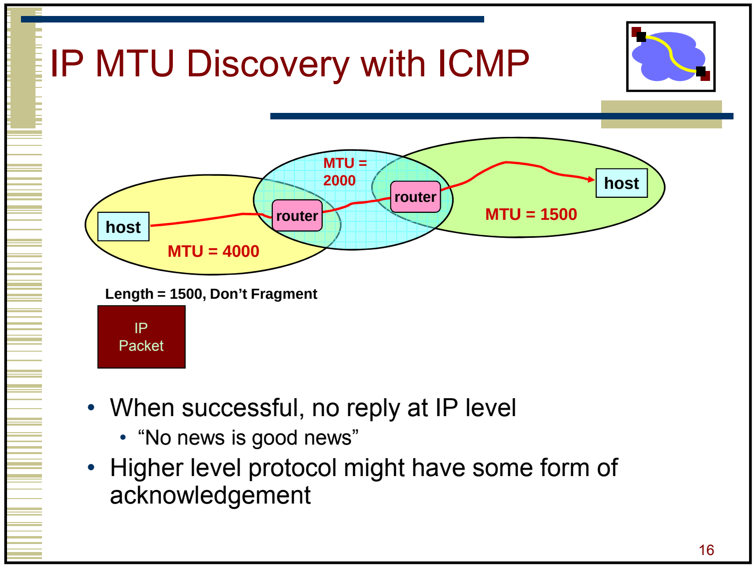
- Typically send series of packets from one host to another
- Typically, all will follow same route
  - Routes remain stable for minutes at a time
- Makes sense to determine path MTU before sending real packets
- Operation: Send max-sized packet with "do not fragment" flag set
  - If encounters problem,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



IP MTU Discovery with ICMP



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

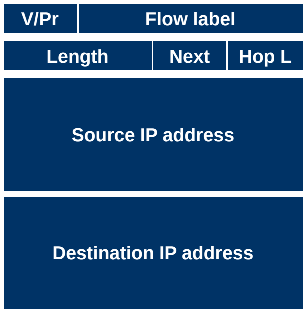
Outline

- The IP protocol
  - IPv4
  - IPv6

- IP in practice
  - Network address translation
  - Address resolution protocol
  - 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 (really?)
- Support for guaranteed services: priority and flow id
- Options handled as “next header”
  - reduces overhead of handling options



IPv6 Address Size Discussion

- Do we need more addresses? Probably, long term
  - Big panic in 90s: “We’re running out of addresses!”
  - Big worry: Devices. Small devices. Cell phones, toasters, everything.
- 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.

IP Router Implementation: 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 are in trouble!

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.
  - No “flag day”
- Fundamentally hard because a (single) IP protocol is critical to achieving global connectivity across the internet
- Process uses a 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!
- 20 years later, this is still a major challenge!