EP2120 Internetworking

IK2218 Protocols and Principles of the Internet

Lecture 4

György Dán
KTH/EE/LCN

Literature:

Forouzan, TCP/IP Protocol Suite
(3^{ed} Ch 6,8, 27.1, 27.3)(4^{ed} Ch 6,7,27)

Internet Protocol (IP)

IP

Basic functionality
and
IPv4 and IPv6 packet headers

IP Service and Functionality

- Best effort service - *unreliable* and *connectionless*
  - Application or Transport layer handles e2e reliability
- Design follows the end-to-end argument
  - Implements only the absolutely necessary functionality
- IPv4 and IPv6 functionality – network layer functionality
  - Addressing
    - Globally unique addresses
    - Aggregation, subnet identification
  - Forwarding
  - Fragmentation
  - Multiplexing/Demultiplexing
  - Routing protocols
  - Error handling and diagnosis

Protocol specifications

(IPv4 and IPv6)

IP

Versioning - Upgradability

- Version 3 (IEN 21, 1 February 1978)
  - Stems from when NCP was being split into one component handling hop-by-hop communication (IP) and one component handling end-to-end communication (TCP).
- Version 4 (RFC 791) – IPv4
  - exclusively used since 1983.01.01
- Version 5 (IEN119/RFC 1190/RFC1819)
  - ST-II - Internet Stream Protocol (ST) – multimedia streaming
    - Connection oriented with resource reservation
- Version 6 (RFC 2460) – IPv6
  - coming (?!)

IP

Needed for packet processing

IPv4

- Header Length (4 bits)
  - Size of IPv4 header including options (20-60 bytes)
  - Granularity: 4 bytes
    - 5≤HLEN≤15
    - limits header size (20 ≤HS ≤60)!!!
- Total Length (16 bits)
  - Total length of datagram including header (20-65535 bytes, practice ≤ 8KB)
  - If datagram is fragmented: length of fragment
  - Granularity: 1 byte

IPv6

- Payload Length (16 bits)
  - Total length of payload excluding base header (0-65535 bytes)
  - If datagram is fragmented: length of fragment
  - Granularity: 1 byte
  - Extendable: Jumbo payload up to 4GB (see later)

Addressing

- Packet contains source and destination addresses

IPv4

- 32 bit addresses

IPv6

- 128 bit addresses

Multiplexing/Demultiplexing

- Interface to higher layers

IPv4

- Protocol type field
  - 8 bits

IPv6

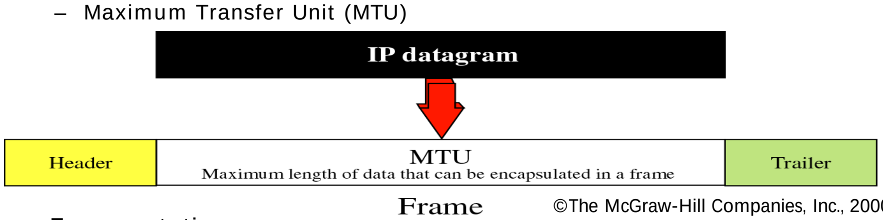
- Next header field
  - 8 bits

- Assigned by IANA
  - Internet Assigned Numbers Authority
- ~137 assigned

decimal	keyword	protocol
1	ICMP	Internet Control Message
4	IP	IP in IP (encapsulation)
6	TCP	Transmission Control
17	UDP	User Datagram
41	IPv6	IPv6 in IPv4
46	RSVP	Reservation Protocol

Fragmentation – MTU

- Adaptation to capabilities of the link layer
- Maximum payload size of a link
  - Maximum Transfer Unit (MTU)



- Fragmentation
  - Datagram size > MTU ⇒ divide datagram into fragments
- Questions
  - Who should fragment?
  - How to route the fragments?
  - Who should reassamble?
  - What if a fragment is missing?
  - What information is needed?

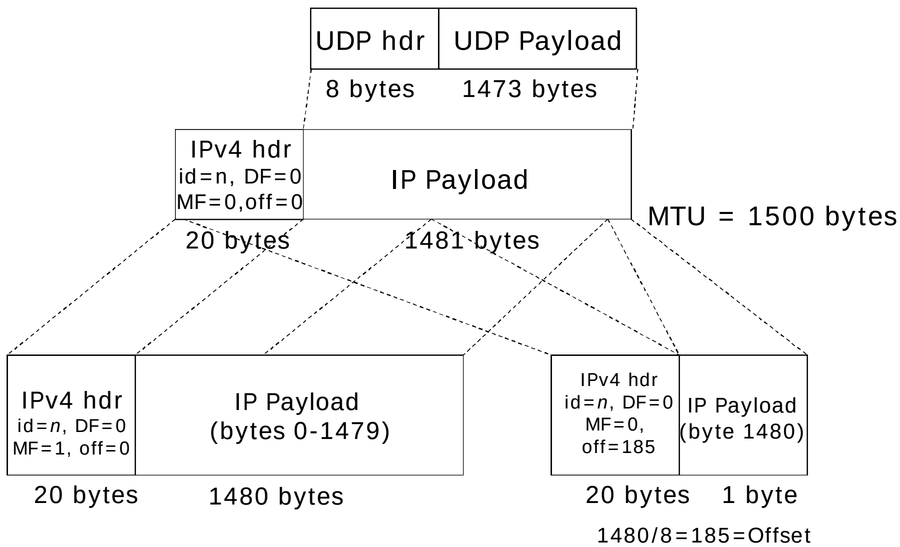
Fragmentation in IPv4 vs. IPV6

	IPv4	IPv6
Who should fragment?	Hosts and routers (unless DF bit set)	Hosts only (router discards and notifies sender)
How to route fragments?	Independently	
Who should reassemble?	Destination host	
Lost fragment?	Discard entire datagram	
Minimum link MTU	68 bytes/576 bytes (rfc791)	1280 bytes (RFC 2460)
Where to store the information?	IPv4 header	Fragmentation extension header

Fragmentation Fields in IPv4

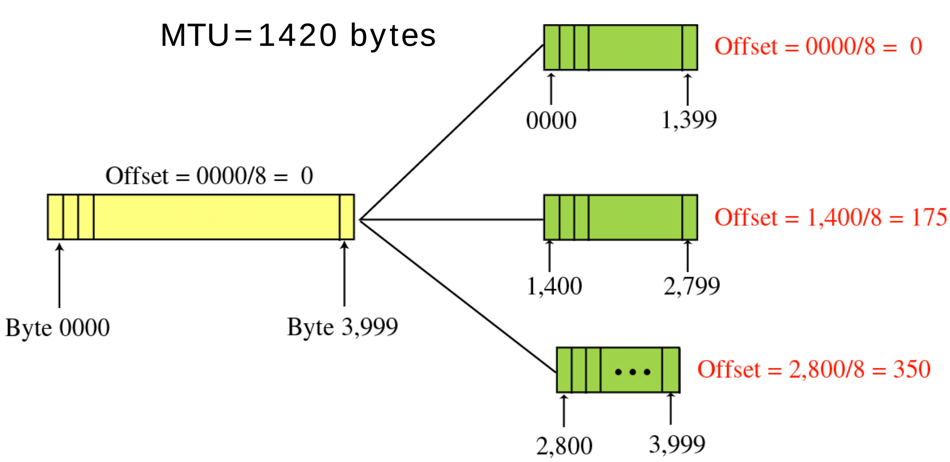
- Identification: 16 bits
  - Datagram uniquely identified by: src IP, dest IP, ID and protocol
  - The ID is copied to all fragments of a datagram upon fragmentation
- Flags: 3 bits
  - RF (Reserved Fragment) – for future use (set to 0)
  - DF (Dont Fragment).
    - Set to 1 if datagram should not be fragmented.
    - If set and fragmentation needed, datagram will be discarded and an error message will be returned to the sender
  - MF (More Fragments)
    - Set to 1 for all fragments, except the last.
- Fragmentation Offset: 13 bits
  - 8-byte units: (ip→ip_frag << 3)
  - Shows relative position of a fragment with respect to the whole datagram

IPv4 Fragmentation Example



IP

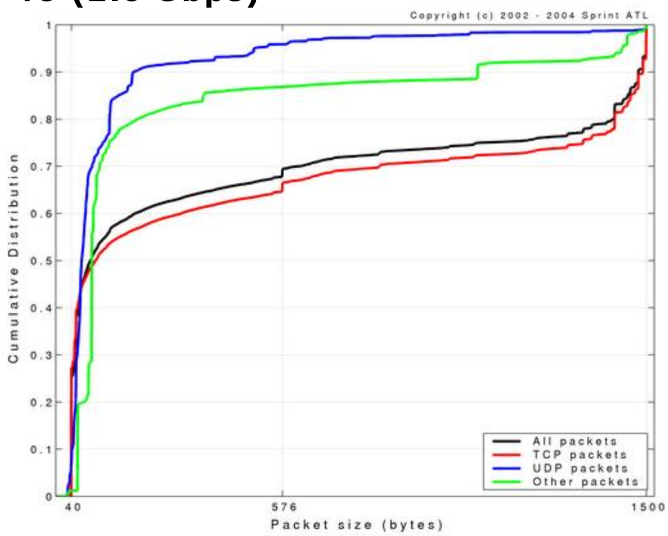
IPv4 Fragmentation Example
Multiple Fragments



IP

Example – Packet sizes in the Internet

- **OC-48 (2.5 Gbps)**



IP

Quiz

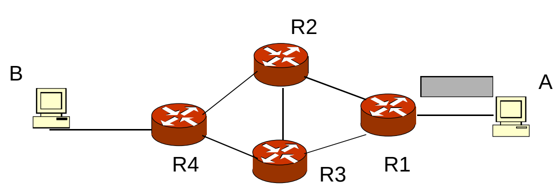
- Consider an IPv4 internetwork. Host A wants to send 56 bytes of data to Host B. Assuming fragments do not get lost, in the worst case the number of fragments received by Host B is

  1

  2

  3

  7



IP

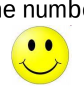
Infinite Loop Avoidance

"Anything that can go wrong will go wrong"
Murphy's law

- Limit the "lifetime" of every datagram
  - Maintain "down counter" in packet header
  - Router upon reception
    - Drops packet if down counter ≤ 1
      - Notifies sender about the expiration of the packet (how?)
    - Decrements counter and forwards the packet otherwise
  - Default initial value: 64

IPv4

- Time To Live (TTL): 8 bits
  - Every router holding a datagram for more than 1 second should decrement the TTL by the number of seconds



IPv6

- Hop Limit: 8 bits

Error Checking (in IPv4)

"Anything that can go wrong will go wrong"
Murphy's law

- Detect bit errors in the IPv4 header
  - Covers the header only (not the payload)
- Operation
  - Calculated hop-by-hop (not end-to-end)
  - Calculation (Internet Checksum Algorithm, RFC 1071)
    - Header = sequence of 16 bit words
    - Add words together using one's complement
    - One's complement of the result ⇒ Checksum

```
  1000 0110 0101 1110
+ 1010 1100 0110 0000
1 0011 0010 1011 1110
  0011 0010 1011 1111
  1100 1101 0100 0000
```

- Other checksums
  - L2 checksums (Link layer checksum, hop-by-hop)
  - L4 checksums (TCP/ICMP/UDP checksums end-to-end)
- No checksum in IPv6

Quality of Service

"Everyone is equal, but some are more equal than others"
(G. Orwell, Animal Farm)

- Network used by applications with different requirements
  - Bulk data transfer
  - Real-time multimedia communications (VoIP, IPTV)
  - Enterprise communications (VPNs)

- Preferential treatment for more sensitive data
  - How should the treatment be defined?
    - Packets, flows, aggregates…
  - What support is needed in the packet header?
  - Compatible with end-to-end argument?
  - Access control?

QoS support in IPv4 and IPv6

- Fields in IPv4
  - Type of Service (ToS): 8 bits
    - RFC 791, 1122, 1349, 1455, *2474 (DiffServ), 3168 (ECN)*
- Fields in IPv6
  - Traffic Class: 8 bits
    - RFC *2474 (DiffServ), 3168 (ECN)*
  - Flow Label: 20 bits
    - Unique to flow between two interfaces
      - Helps to speed up packet processing
    - Value 0 if not needed (e.g., short lived flow)
- Few applications set the TOS/Traffic Class field
  - Typically the ISPs...

The ToS Byte – Original proposal

Bit 0 … Bit 7

Precedence	TOS		

- Original Proposal – RFC 791
  - Specify QoS on a per-packet basis
- Bits 0-2: Precedence
  - Defines priority, e.g., when packets must be dropped
- Bits 3-6: TOS
  - Bit 3: 0 = Normal Delay, 1 = Low Delay
  - Bit 4: 0 = Normal Throughput, 1 = High Throughput
  - Bit 5: 0 = Normal Reliability, 1 = High Reliability
  - Bit 6: 0 = Normal Cost, 1 = Minimize Cost.
- RFC 1122, 1349, 1455 modified this meaning of the ToS field

DSField – Current Proposal

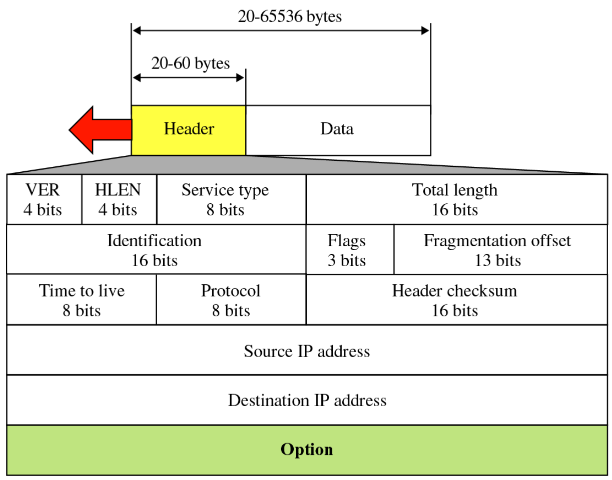
Bit 0 … Bit 7

DSCP	ECN

- Differentiated Services (DiffServ) - 6 bits DS field
  - RFC 2474
  - Bits 0-6: Differentiated Services CodePoint (DSCP)
- Treatment of aggregates
  - When entering an *area* : set DSCP, traffic conditioning
  - Determines the QoS handling of the IP datagram in the routers within that area
  - Per hop behavior (PHB) in the area
    - Scheduling, Dropping
- Explicit Congestion Notification (ECN) – signal congestion
  - ECN Capable Transport (ECT)
  - Congestion Experienced (CE)

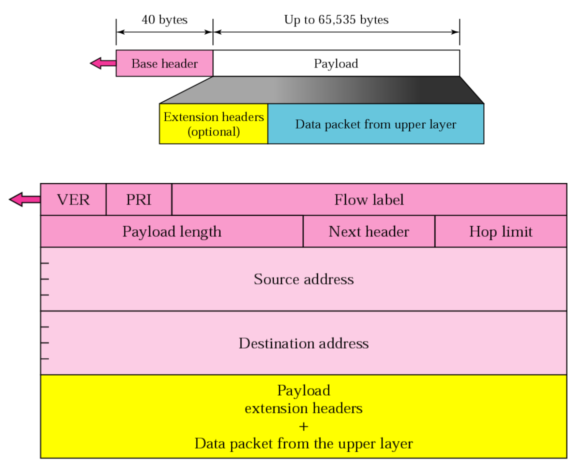
IPv4 Header – RFC 791

- Version
- HLEN – Header Length
- Total Length
  - Header + Payload
- IP Addresses
  - Source, Destination
- Fragmentation
  - ID, Flags, Offset
- Protocol
  - Higher level protocol-Demux
- Type of Service
- TTL – Time To Live
  - Limits lifetime
- Header checksum
- Options



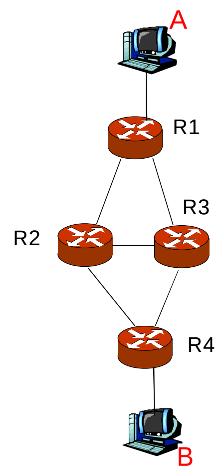
IPv6 Header – RFC 2460

- Version
- Payload Length
- IP Addresses
  - Source, Destination
- Next header
  - Mux/demux
  - Extension headers
- PRI – Traffic class
- Flow label
- Hop Limit



Network layer functions and IP

- Functions provided by IP
  - Logical addressing
    - Locating hosts
  - Routing
    - Path determination
  - Forwarding
    - Move packet from input to output of the routers
  - Fragmentation
    - Adaptation to lower layer
  - Multiplexing/demultiplexing
    - Many transport layer protocols
  - Protocol specifications
- Functions not provided by IP
  - Connection establishment and termination
  - Resource management



Summary

- IP network layer basics
  - Delivery: Connectionless (datagram service)
  - Forwarding: next-hop routing (almost exclusively)
  - Routing tables: RIBs and FIBs
  - Longest prefix matching
  - Datagram header fields
    - Related functionality

Next time

- ARP
- More IP (options+)
- ICMP