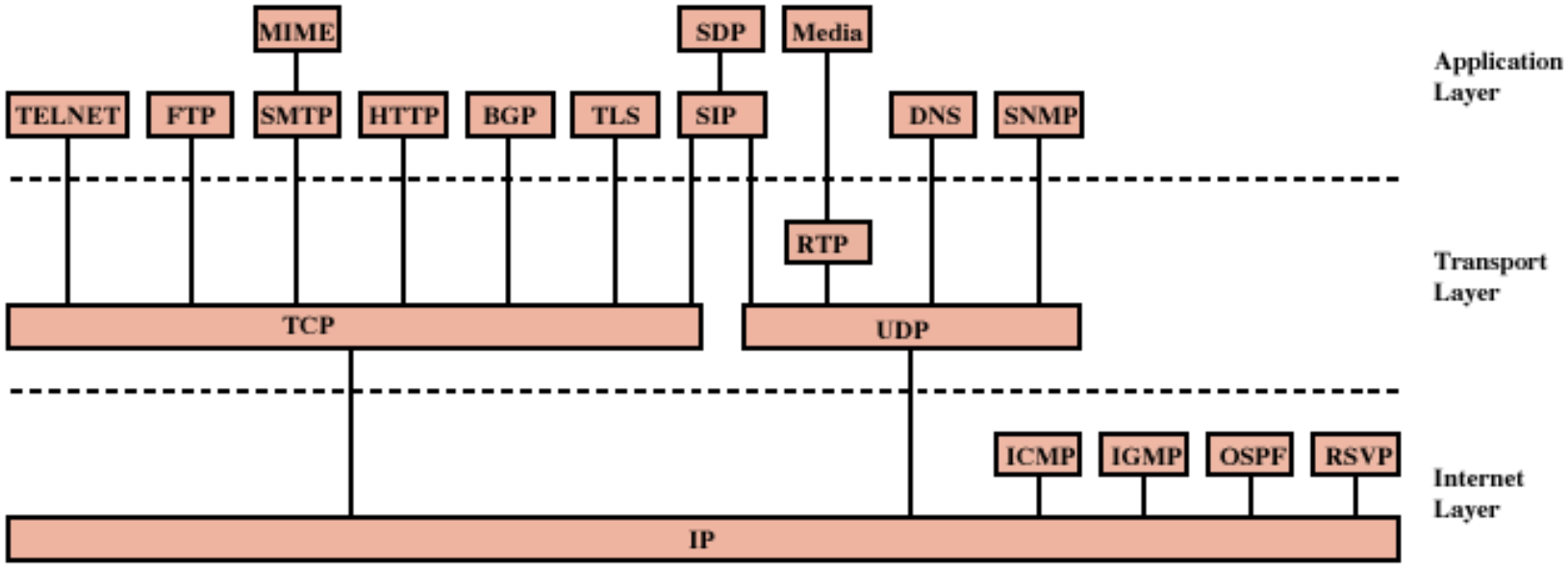
INTERNET PROTOCOL (IP)

Protocols of TCP/IP Protocol Suite



BGP	=	Border Gateway Protocol
DNS	=	Domain Name System
FTP	=	File Transfer Protocol
HTTP	=	Hypertext Transfer Protocol
ICMP	=	Internet Control Message Protocol
IGMP	=	Internet Group Management Protocol
IP	=	Internet Protocol
MIME	=	Multi-Purpose Internet Mail Extension
OSPF	=	Open Shortest Path First
RSVP	=	Resource ReSerVation Protocol
RTP	=	Real-Time Transport Protocol
SDP	=	Session Description Protocol
SIP	=	Session Initiation Protocol
SMTP	=	Simple Mail Transfer Protocol
SNMP	=	Simple Network Management Protocol
TCP	=	Transmission Control Protocol
TLS	=	Transport Layer Security
UDP	=	User Datagram Protocol

Internet Protocol (IP)

- Internet Protocol is one of the major protocols in the TCP/IP protocols suite.
- This protocol works at the network layer of the OSI model and at the Internet layer of the TCP/IP model.
- IP provides connectionless (datagram) service ; Each packet treated separately
- Datagrams
  - Packets in the network (internet) layer are called *datagrams*.
  - A datagram is a variable-length packet consisting of two parts: header and data.
  - The header is 20 to 60 bytes in length and contains information essential to routing and delivery.
  - It is customary in TCP/IP to show the header in 4-byte sections.

- Network layer protocol common to all routers
- This protocol has the responsibility of identifying hosts based upon their logical addresses and to route data among them over the underlying network.
- IP provides a mechanism to uniquely identify hosts by an IP addressing scheme.
- IP uses best effort delivery, i.e. it does not guarantee that packets would be delivered to the destined host, but it will do its best to reach the destination.
- Internet Protocol version 4 uses 32-bit logical address.

Design Issues

- Routing
- Datagram lifetime
- Fragmentation and re-assembly
- Error control
- Flow control
- Addressing

Routing

- End systems and routers maintain routing tables
  - Indicate next router to which datagram should be sent
  - Static
    - Tables do not change but may contain alternative routes
  - Dynamic
    - If needed, the tables are dynamically updated
    - Flexible response to congestion and errors
    - status reports issued by neighbours about down routers
- Source routing
  - Source specifies route as sequential list of routers to be followed
  - useful, for example, if the data is top secret and should follow a set of trusted routers.
- Route recording
  - routers add their address to datagrams
  - good for tracing and debugging purposes

Datagram Lifetime

- Datagrams could loop indefinitely
  - Not good
    - Unnecessary resource consumption
    - Transport protocol needs upper bound on datagram life
- Datagram marked with lifetime
  - Time To Live (TTL) field in IP
  - Once lifetime expires, datagram is discarded (not forwarded)
  - Hop count
    - Decrement time to live on passing through each router
  - Time count
    - Need to know how long since last router
    - global clock is needed

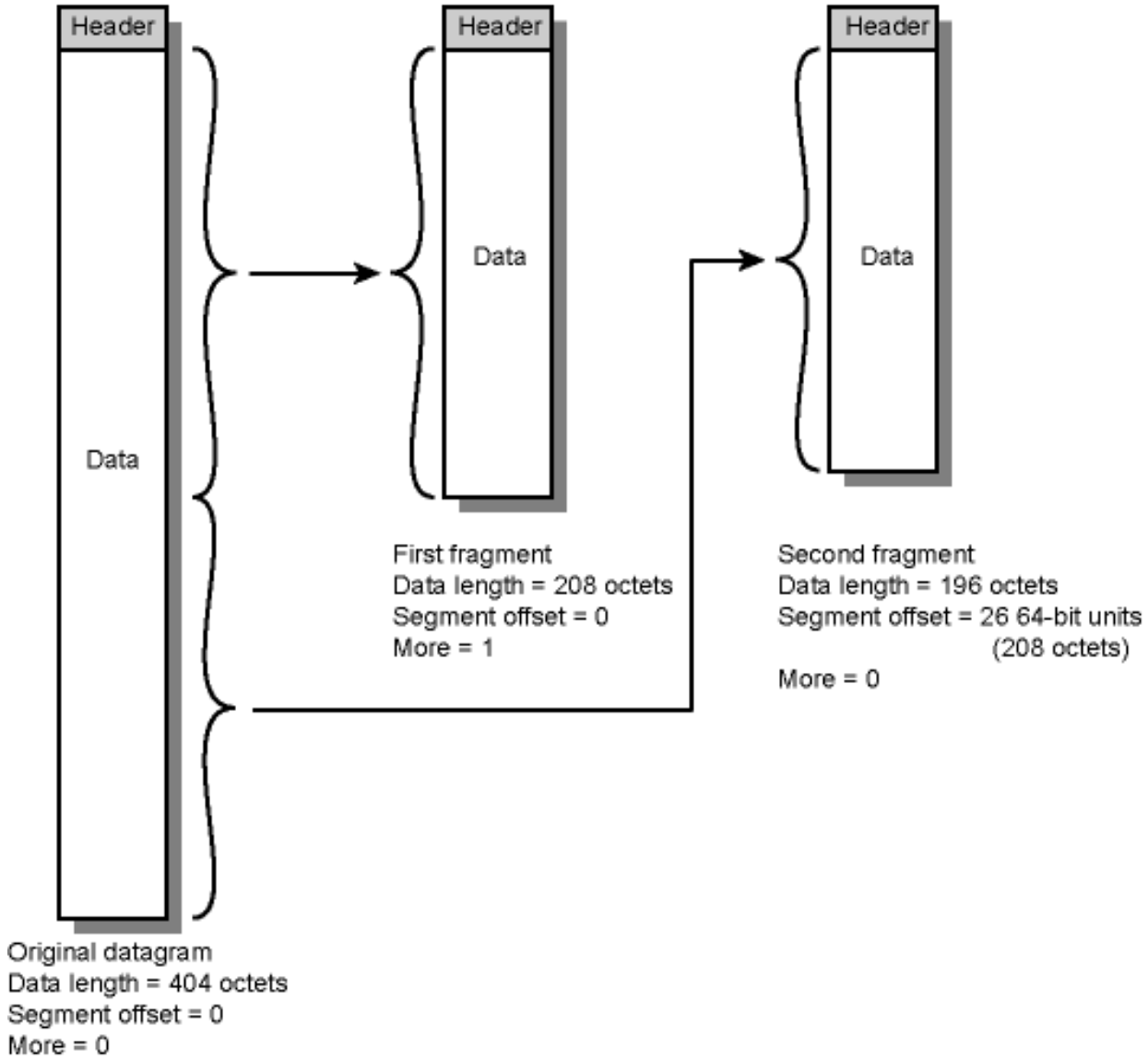
Fragmentation and Re-assembly

- Different maximum packet sizes for different networks
  - routers may need to split the datagrams into smaller fragments
- When to re-assemble
  - At destination
    - Packets get smaller as data travel
      - inefficiency due to headers
  - Intermediate reassembly
    - Need large buffers at routers
    - All fragments must go through same router
      - Inhibits dynamic routing

IP Fragmentation

- In IP, reassembly is **at destination** only
- Uses fields in header
  - Data Unit Identifier –
  - In order to uniquely identify datagram – all fragments that belong to a datagram share the same identifier
    1. Source and destination addresses
    2. Upper protocol layer (e.g. TCP)
    3. Identification supplied by that layer
  - Data length
    - Length of user data in octets (if fragment, length of fragment data)
    - Actually header contains total length but data length can be calculated
  - Offset
    - Position of fragment of user data in original datagram (position of the first byte of the fragment)
    - In multiples of 64 bits (8 octets)
  - *More* flag
    - Indicates that this is not the last fragment (if this flag is 1)

Fragmentation Example



Dealing with Failure

- Reassembly may fail if some fragments get lost
- Need to detect failure to free up the buffers
- One solution: Reassembly time out
  - Assign a reassembly lifetime to the first fragment
  - If timer expires before all fragments arrive, discard partial data

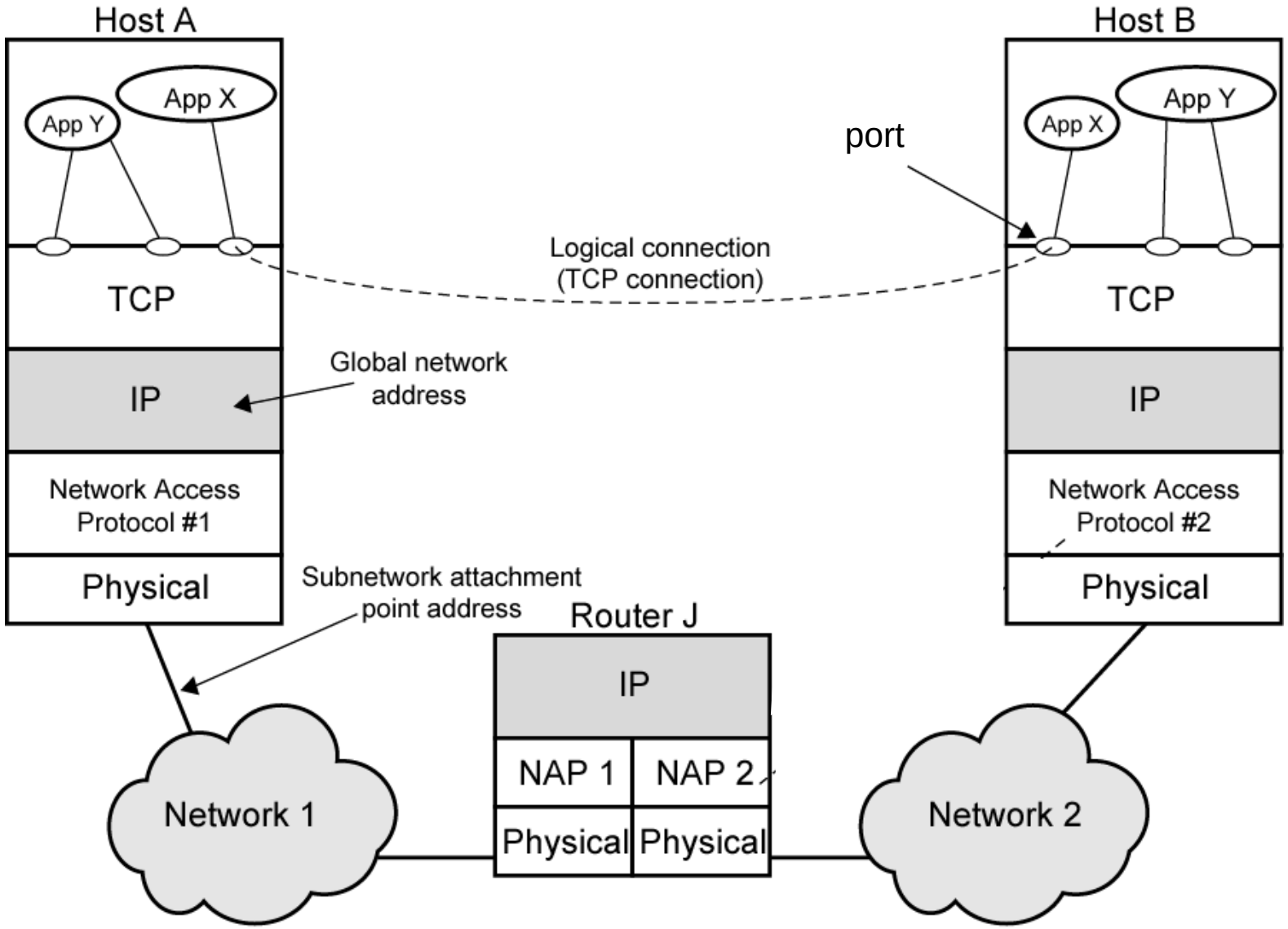
Error Control

- In IP, delivery is not guaranteed
- Router may attempt to inform source if packet discarded, if possible
  - specify the reason of drop, e.g. for time to live expiration, congestion, bad checksum (error detected)
- Datagram identification needed
- When source receives failure notification, it
  - may modify transmission strategy
  - may inform high layer protocol
- such a failure notification is not guaranteed

Flow Control (in IP layer)

- Allows routers and/or stations to limit rate of incoming data
- In connectionless systems (such as IP), mechanisms are limited
- Send flow control packets requesting reduced flow
  - e.g. using *source quench* packet of ICMP

Addressing in TCP/IP



Addressing Level

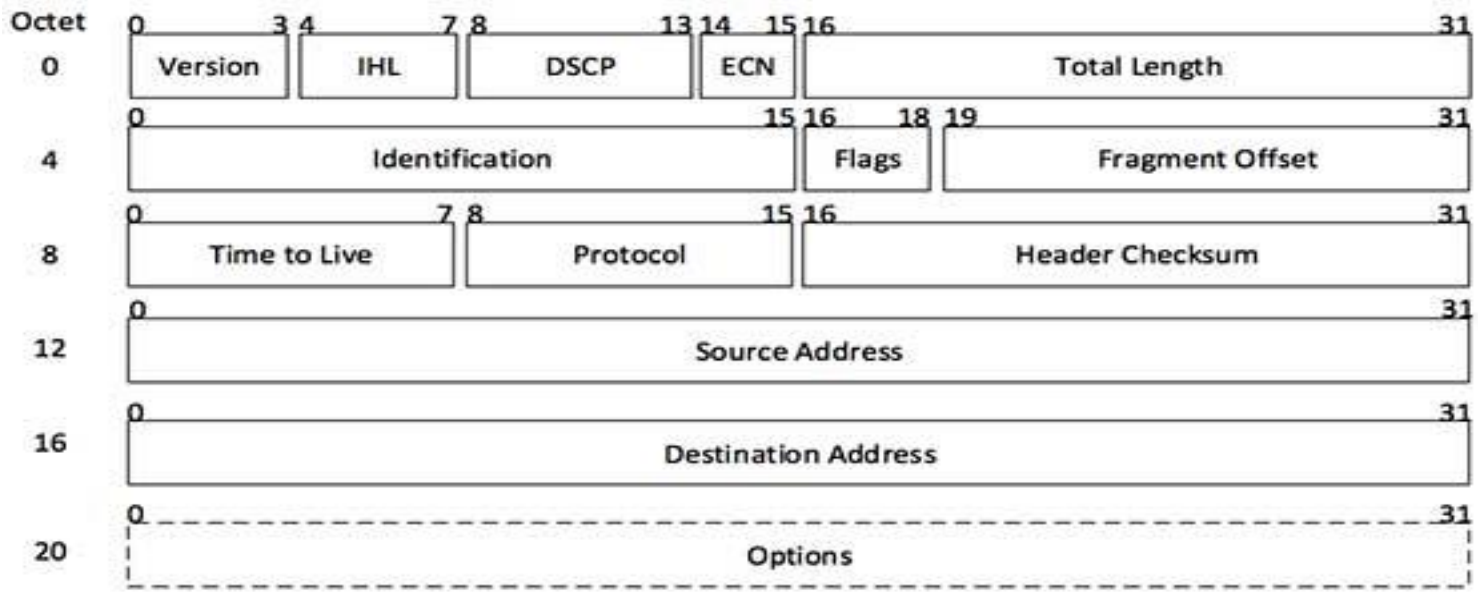
- Level in communication architecture at which entity is named
- Unique address for each end system
  - e.g. workstation or server
- And each intermediate system
  - (e.g., router)
- Network-level address
  - IP address or internet address
  - OSI - network service access point (NSAP)
  - Used to route PDU through network
- At destination data must routed to some process
  - Each process assigned an identifier
  - TCP/IP port
  - Service access point (SAP) in OSI

IP Datagram

- Internet Protocol being a layer-3 protocol (OSI) takes data Segments from layer-4 (Transport) and divides it into packets. IP packet encapsulates data unit received from above layer and add to its own header information.
- The encapsulated data is referred to as IP Payload. IP header contains all the necessary information to deliver the packet at the other end.

IP Header	Layer – 4 Data

(IP Encapsulation)



[Image: IP Header]

Header Fields

- Version
  - Version no. of Internet Protocol used (e.g. IPv4).
- Internet header length
  - Length of entire IP header.
  - Unit is 32 bit words
  - Including options
  - minimum 5 (means 20 octets)
- DSCP Differentiated Services Code Point
  - Earlier it is “Type of Service”
  - Used for QoS support
- ECN (Explicit Congestion Notification)
  - It carries information about the congestion seen in the route.

Header Fields

- Total length
  - Length of entire IP Packet (including IP header and IP Payload) in octets
- Identification
  - If IP packet is fragmented during the transmission, all the fragments contain same identification number to identify original IP packet they belong to.
  - Sequence number
  - Used with addresses and user protocol to identify datagram uniquely
- Flags
  - As required by the network resources, if IP Packet is too large to handle, these ‘flags’ tells if they can be fragmented or not.
  - More bit
  - Don’t fragment

- Fragmentation offset
  - This offset tells the exact position of the fragment in the original IP Packet.
- Time to live
  - To avoid looping in the network, every packet is sent with some TTL value set, which tells the network how many routers (hops) this packet can cross. At each hop, its value is decremented by one and when the value reaches zero, the packet is discarded.
- Protocol
  - Next higher layer to receive data field at destination
  - Tells the Network layer at the destination host, to which Protocol this packet belongs to, i.e. the next level Protocol. For example protocol number of ICMP is 1, TCP is 6 and UDP is 17.

Header Fields

- Header checksum
  - This field is used to keep checksum value of entire header which is then used to check if the packet is received error-free.
  - Verified and recomputed at each router
- Source address
  - 32-bit address of the Sender (or source) of the packet.
- Destination address
  - 32-bit address of the Receiver (or destination) of the packet.

- Options
  - This is optional field, which is used if the value of IHL is greater than 5. These options may contain values for options such as Security, Record Route, Time Stamp, etc.

Option	Description
Security	Specifies how secret the datagram is
Strict source routing	Gives the complete path to be followed
Loose source routing	Gives a list of routers not to be missed
Record route	Makes each router append its IP address
Timestamp	Makes each router append its address and timestamp

- Padding
  - To fill to multiple of 32 bits long

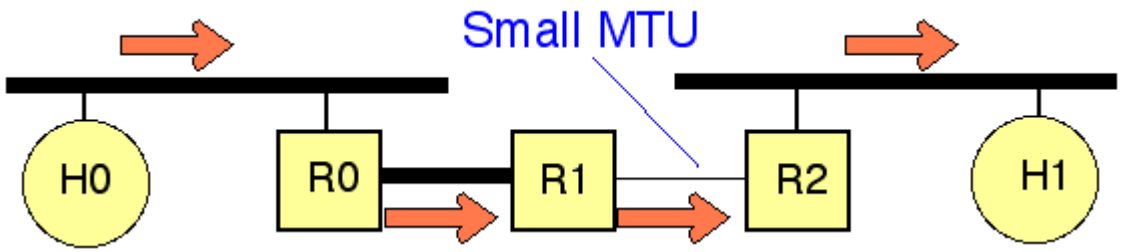
Data Field

- User (upper layer) data
- any octet length is OK
  - But max length of IP datagram (header plus data) is 65,535 octets

IP Fragmentation

Maximum Transmission Unit (MTU)

- An IP packet that is larger than the Maximum Transmission Unit (MTU) of an interface, is too large for transmission over that interface.
- The packet must either be fragmented, or discarded (and an ICMP error message returned to the sender).
- In either case, the original data will be fragmented into smaller packets (less than the smallest MTU) in order to allow it to be received by the final destination system.



- Two Approaches
- IP Router Segmentation - performing the fragmentation in the routers
- IP Path MTU Discovery - forcing the sender to perform the fragmentation

IP Fragmentation processing at a Router/ IP Router Segmentation

- In this simple approach, the sender simply has to ensure that each packet is less than the MTU of the link on which it is sent. (The router always knows this from the link interface configuration information).
- The network layer then has to arrange to cut packets up into smaller fragments whenever a router encounters a link with an MTU smaller than the received IP packet size.
- All the fragments of an IP packet carry the same ID in the IP packet header (allowing the final receiver to reassemble the fragmented parts into the original PDU). This is called "IP fragmentation" or "IP segmentation".

- The problem is, this offloads a lot of work on to routers, and in the worst case, can also result in packets being segmented by several IP routers one after another, resulting in very peculiar fragmentation.

- IP Router fragmentation is not recommended in the modern Internet

IP Fragmentation processing at a Sender/ IP Path MTU Discovery

- Path MTU Discovery allows a sender to fragment/segment a long internet packet, rather than relying on routers to perform IP-level fragmentation.
- This is more efficient and more scalable.
- It is therefore the recommended method in the current Internet.

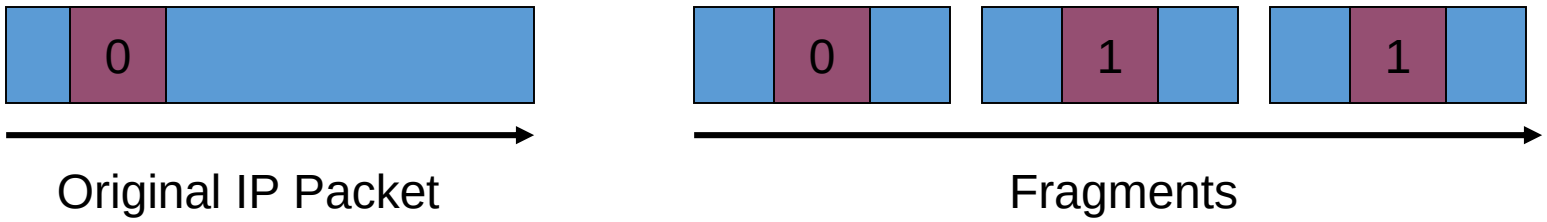
IP Reassembly processing at the Receiving End System

- IP fragmentation and reassembly employs updating and using the values in the second 32 bits of the IPv4 packet header.
- An end system that accepts an IP packet (with a destination IP address that matches its own IP source address) will also reassemble any fragmented IP packets before these are passed to the next higher protocol layer.

- The system stores all received fragments (i.e., IP packets with a more-fragments flag (MF)set to one, or where the fragment offset is non-zero), in one of a number of buffers (memory space).
- Packets with the same 16-bit Identification value are stored in the same buffer, at the offset specified by the fragment offset field specified in the packet header.
- Packets which are incomplete remain stored in the buffer until either all fragments are received, OR a timer expires, indicating that the receiver does not expect to receive any more fragments.
- Completed packets are forwarded to the next higher protocol layer.

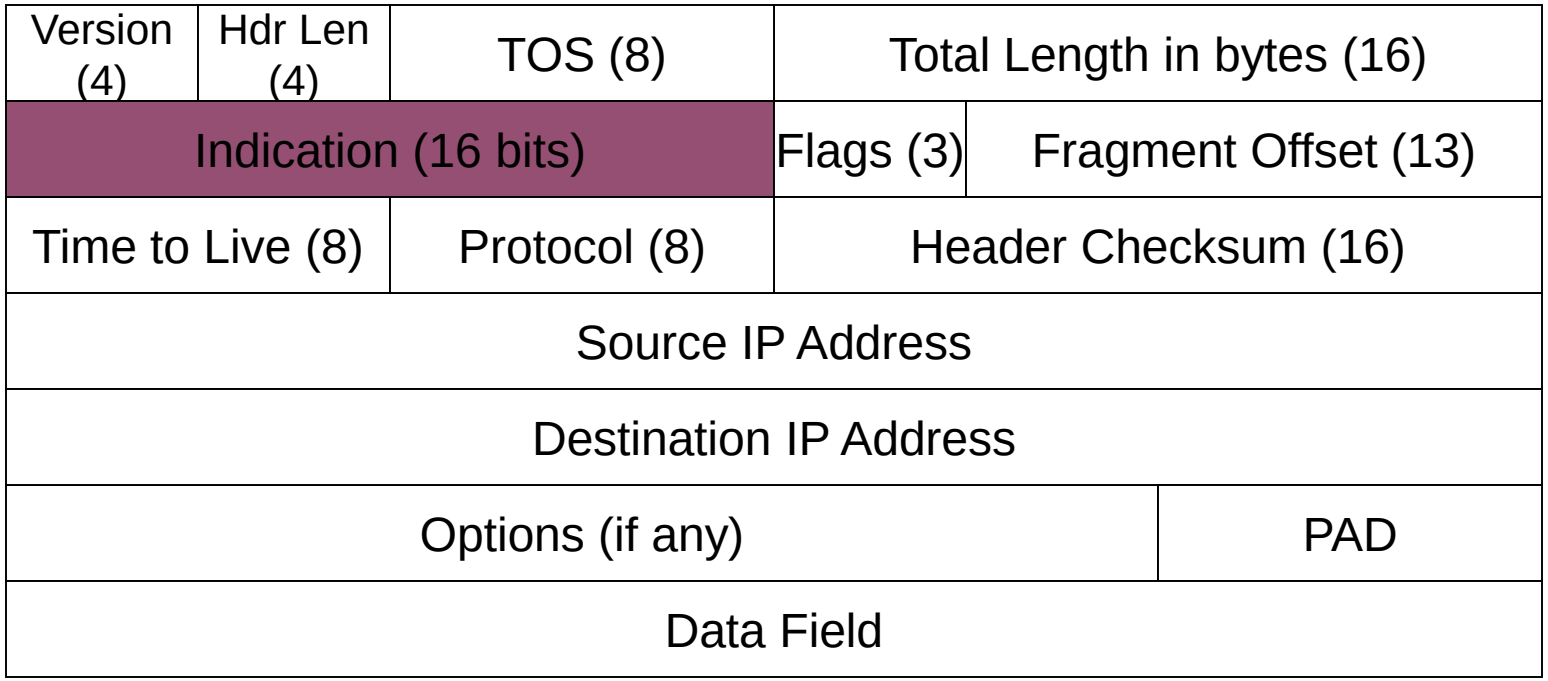
Fragmentation and IP Fields

- *More Fragments* field (1 bit)
  - 1 if more fragments
  - 0 if not
  - Source host internet process sets to 0
  - If router fragments, sets More Fragments field in last fragment to 0
  - In all other fragments, sets to 1



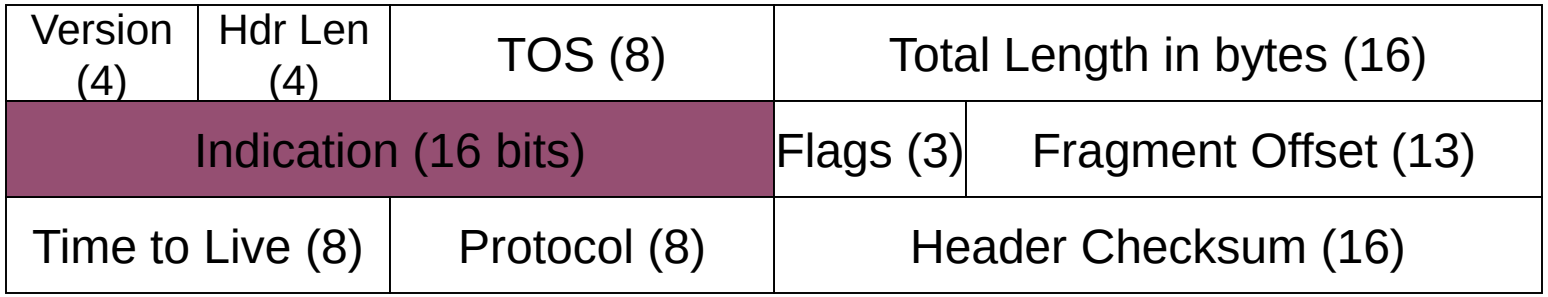
Identification Field

- IP packet has a 16-bit *Identification* field



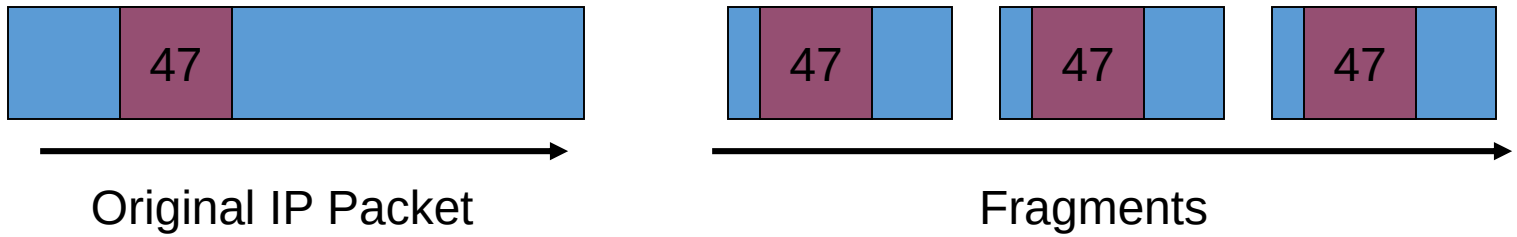
Identification Field

- IP packet has a 16-bit *Identification* field
  - Source host internet process places a random number in the Identification field
  - Different for each IP packet



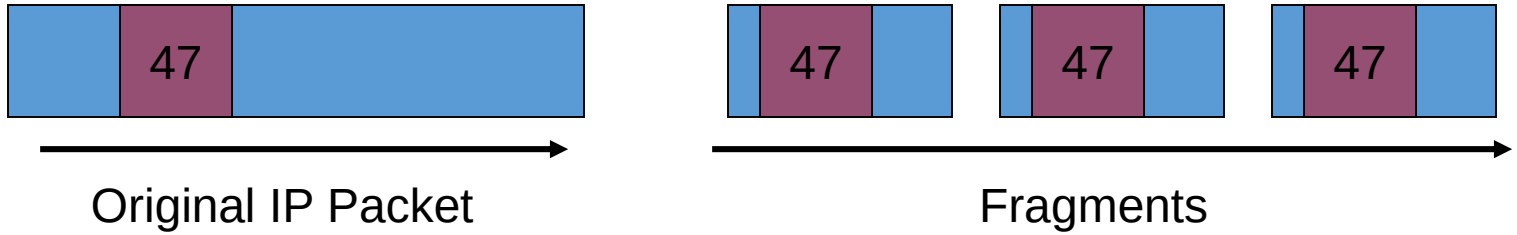
Identification Field

- IP packet has a 16-bit Identification field
  - If router fragments, places the original Identification field value in the Identification field of each fragment



Identification Field

- Purpose
  - Allows receiving host's internet layer process know what fragments belong to each original packet
  - Works even if an IP packet is fragmented several times



Fragment Offset Field

- Fragment offset field (13 bits) is used to reorder fragments with the same Identification field
- Contains the data field's starting point (in octets) from the start of the data field in the original IP packet

Fragment Offset Field

- Receiving host's internet layer process assembles fragments in order of increasing fragment offset field value
- This works even if fragments arrive out of order!
- Works even if fragmentation occurs multiple times