Network Layer
Part 3

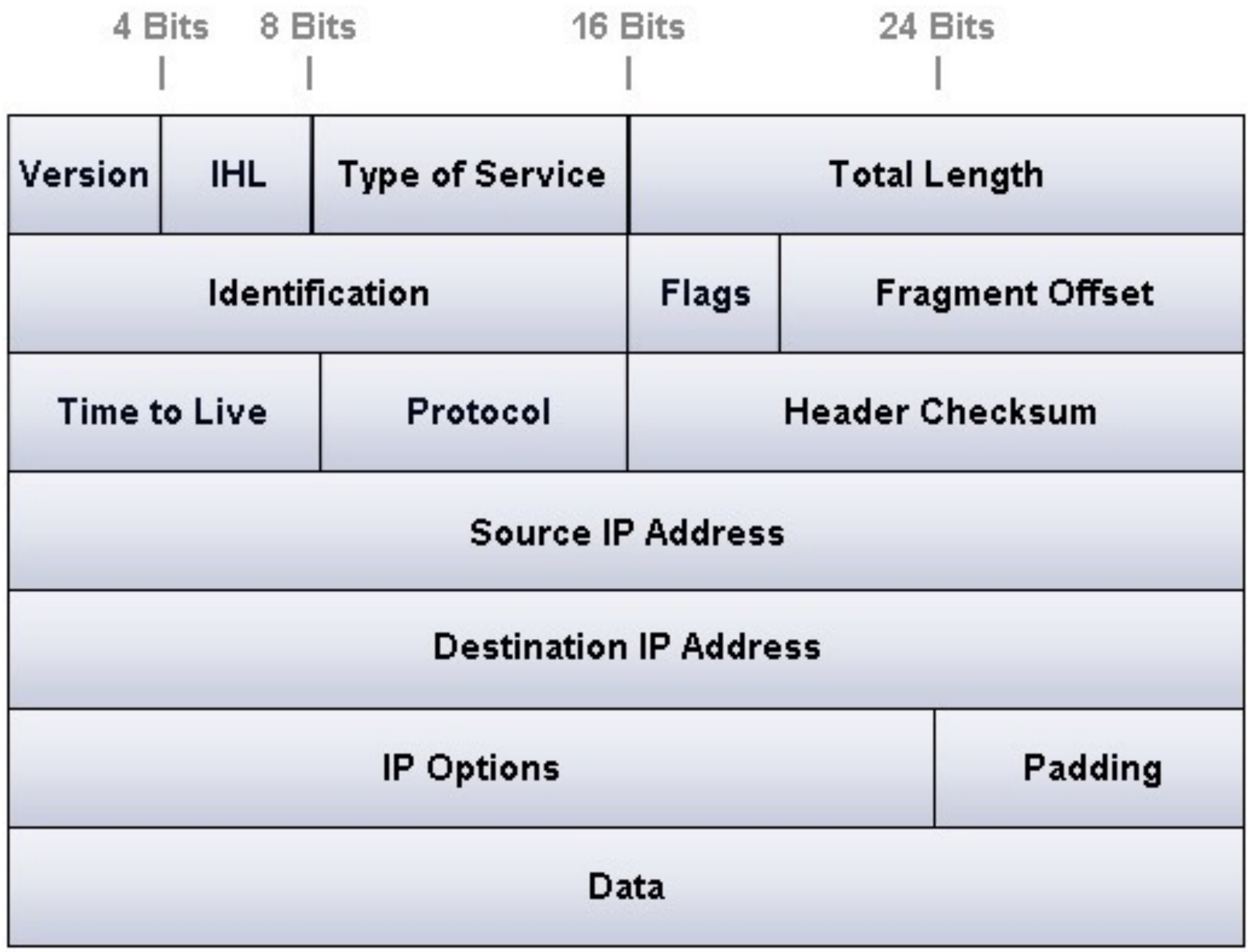
Mark Allman
mallman@case.edu

EECS 325/425
Fall 2018

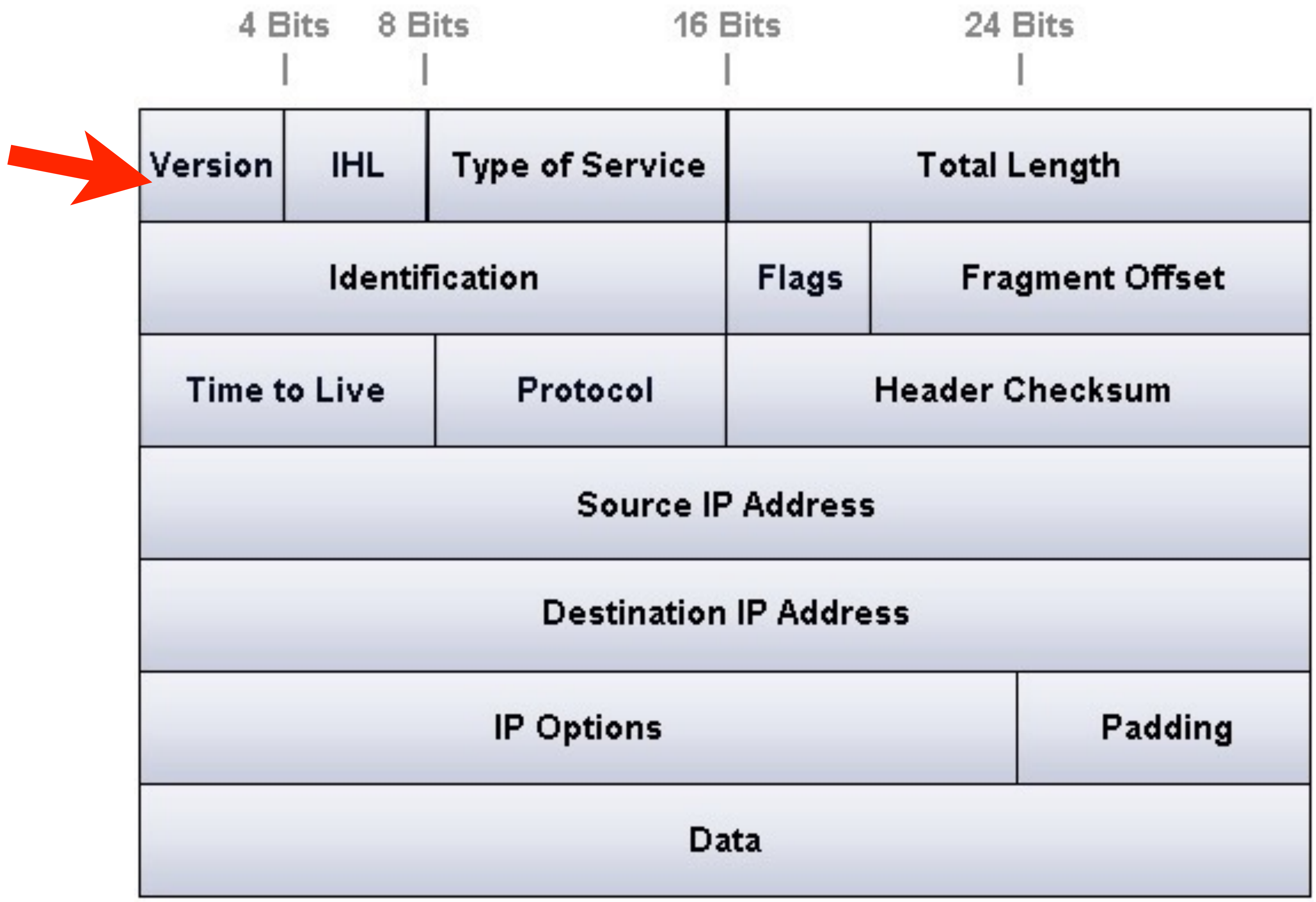
*"Way down South they had a jubilee,
Those Georgia folks they had a jamboree.
They're drinkin' home brew from a wooden cup,
The folks dancin' got all shook up."*

IP datagram format

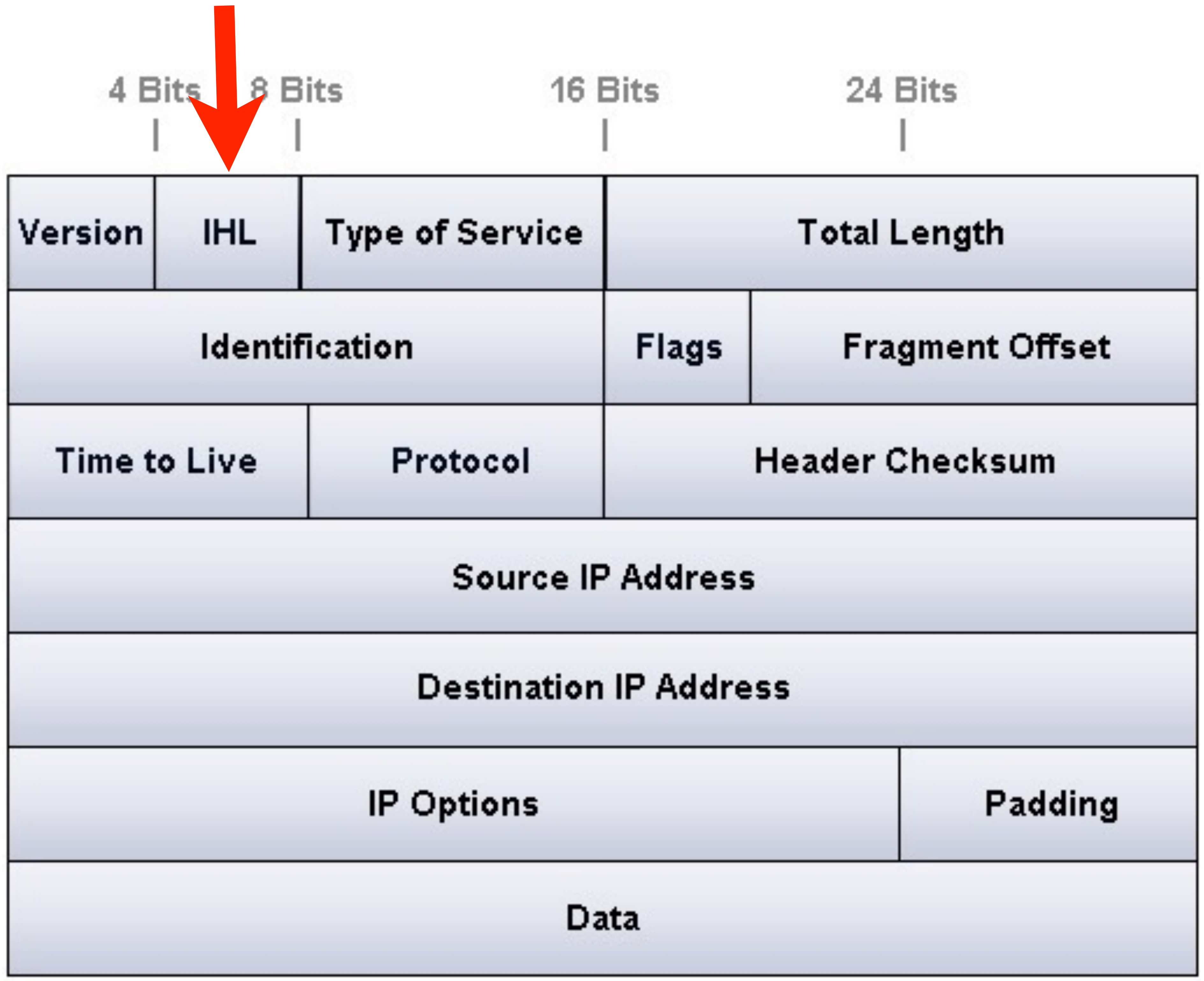
IP datagram format



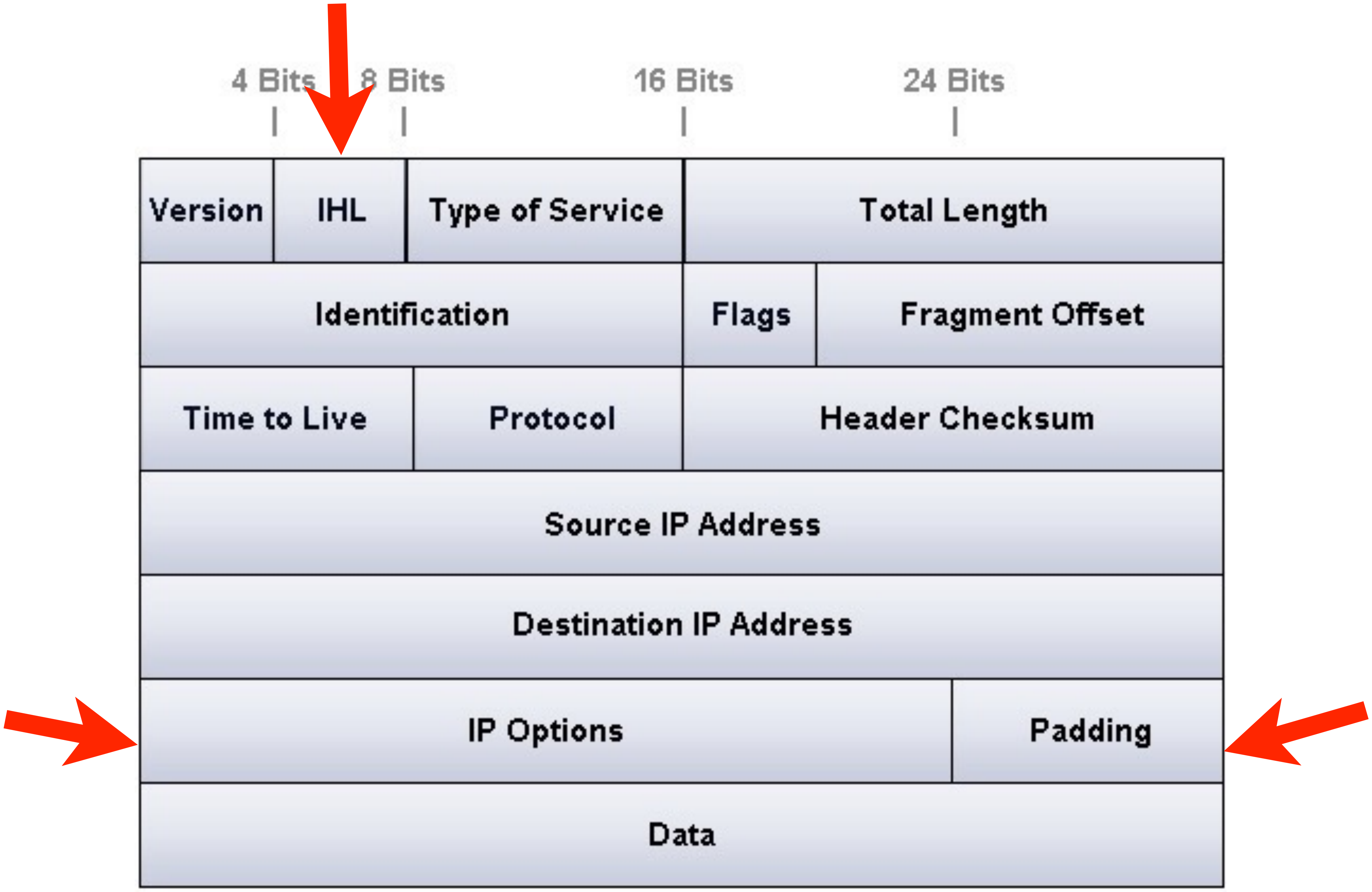
IP datagram format



IP datagram format

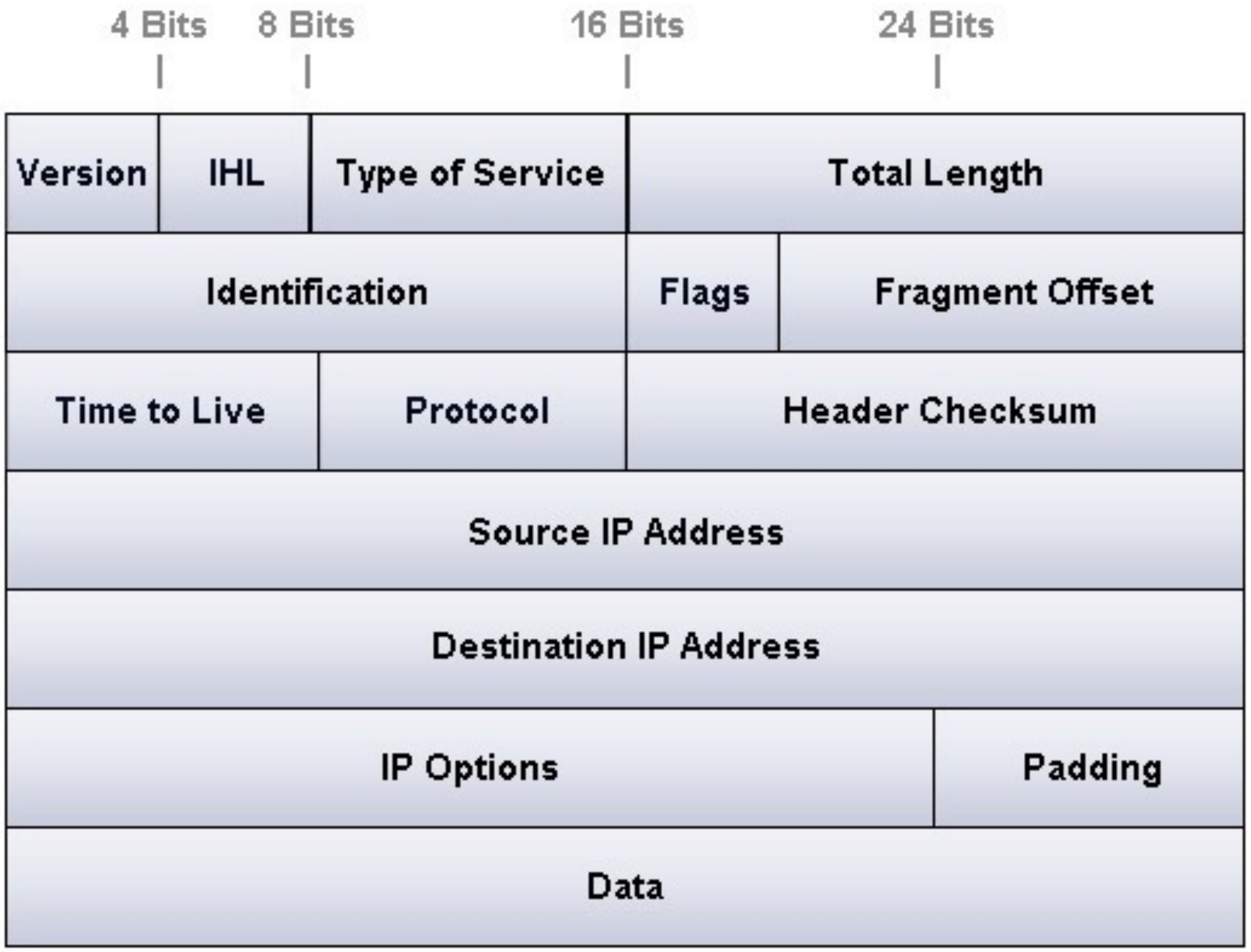


IP datagram format

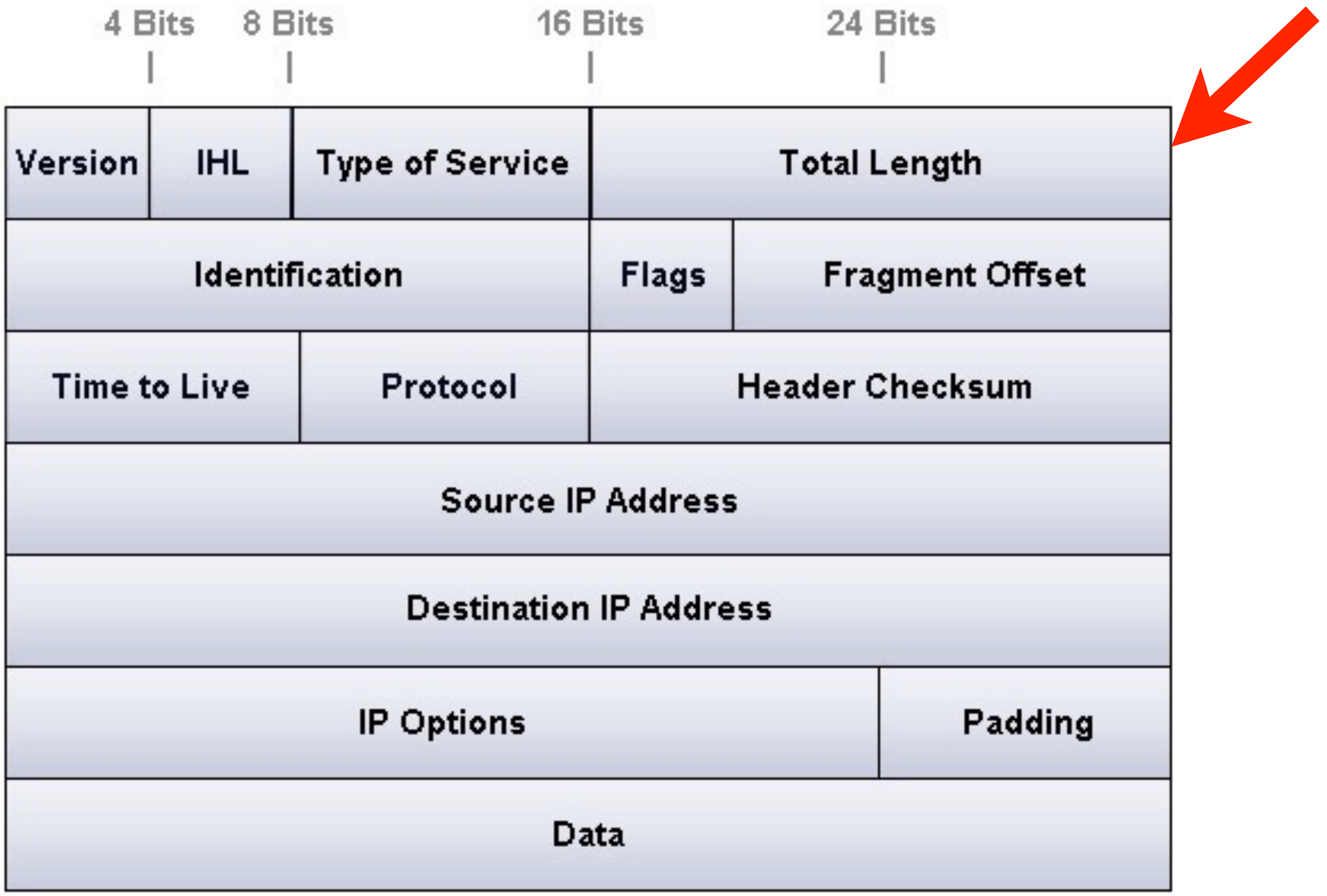


IP datagram format

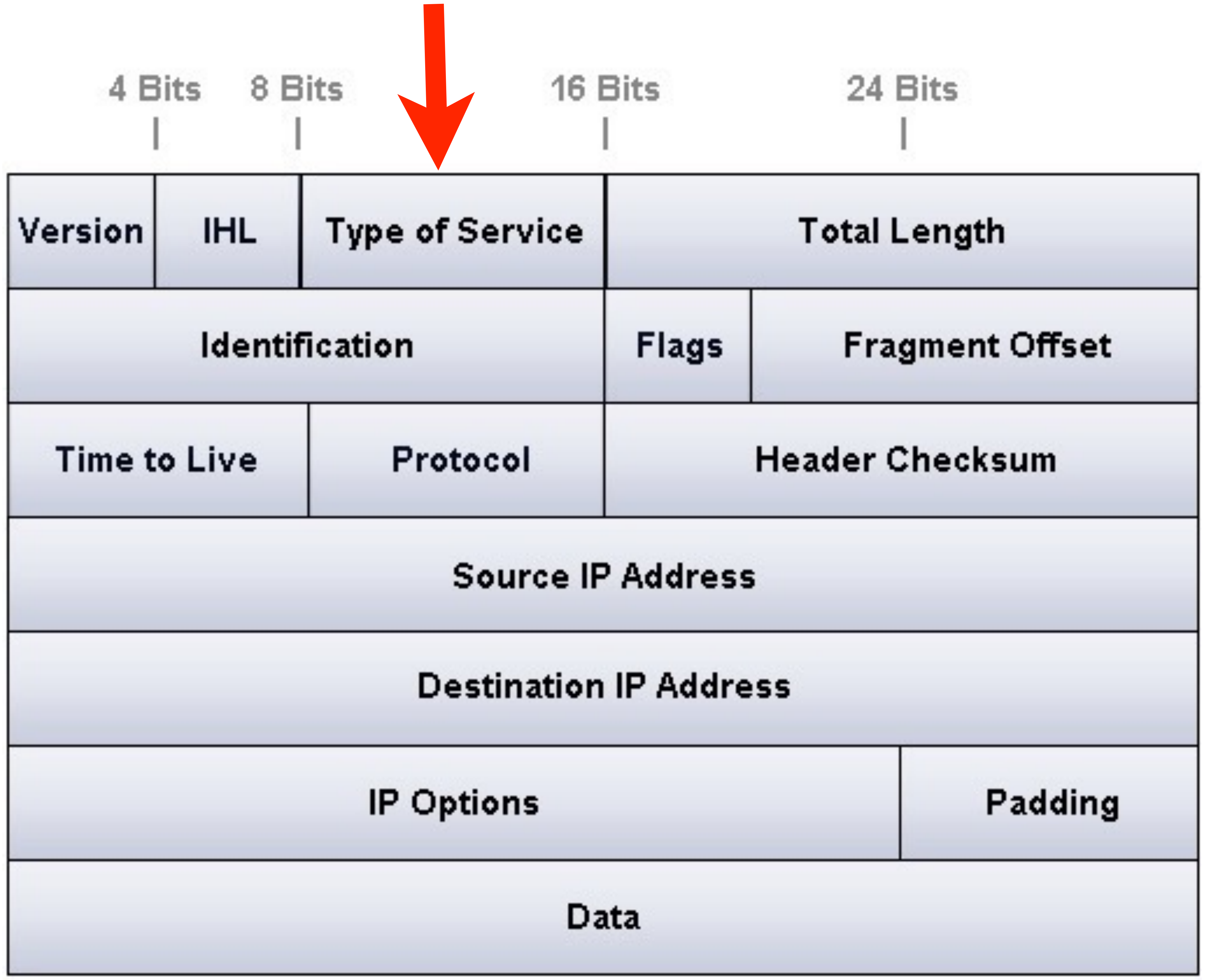
IP datagram format



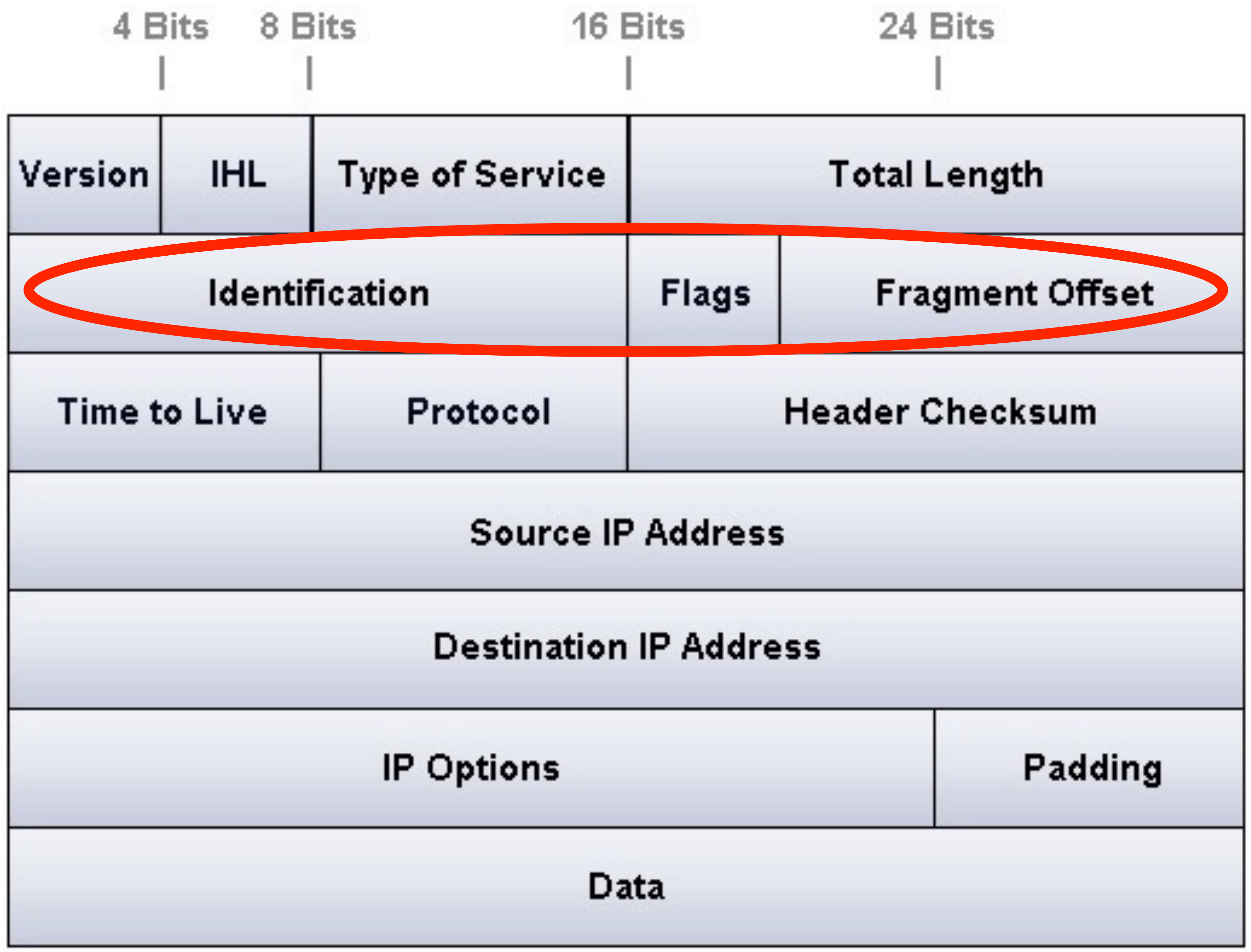
IP datagram format



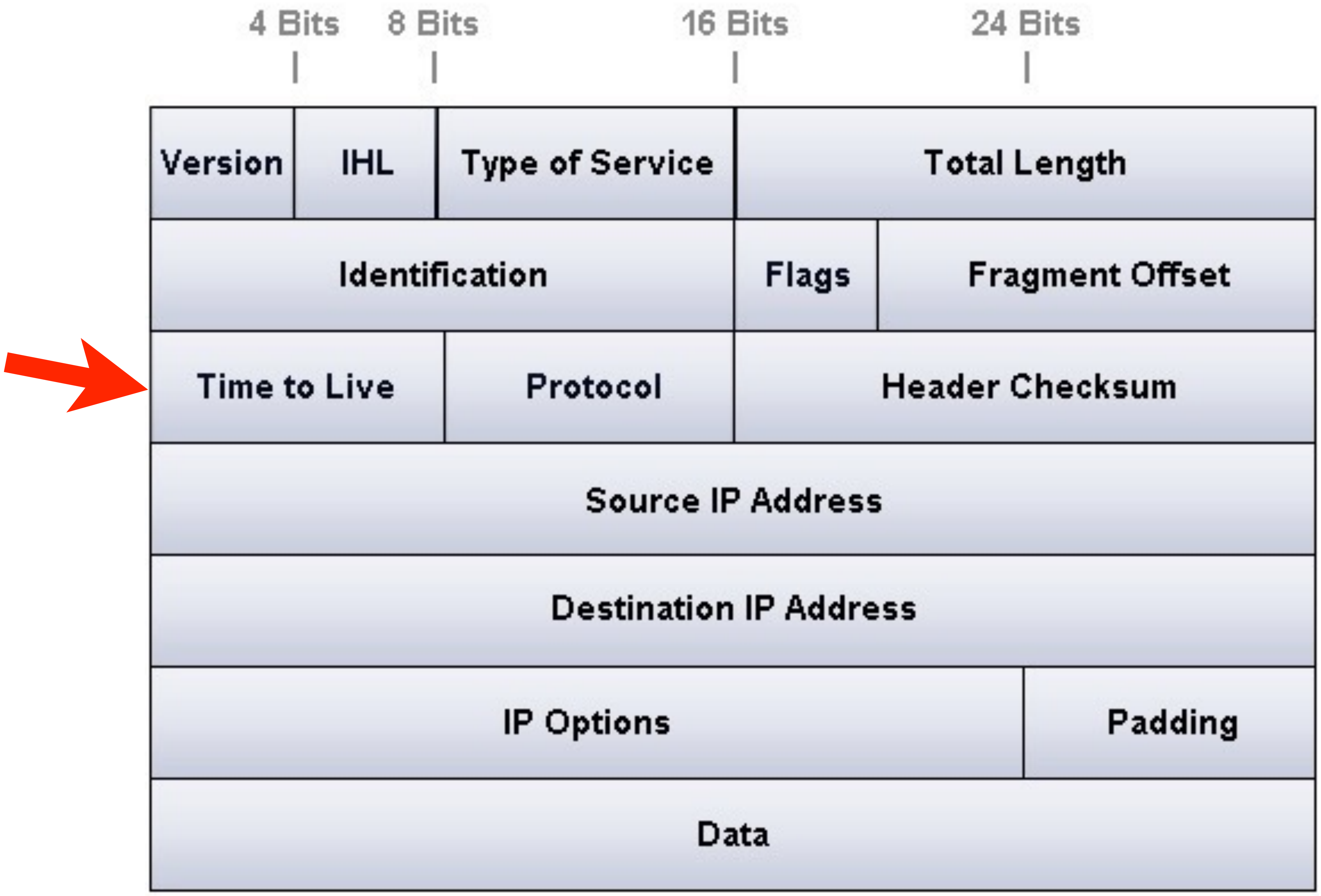
IP datagram format



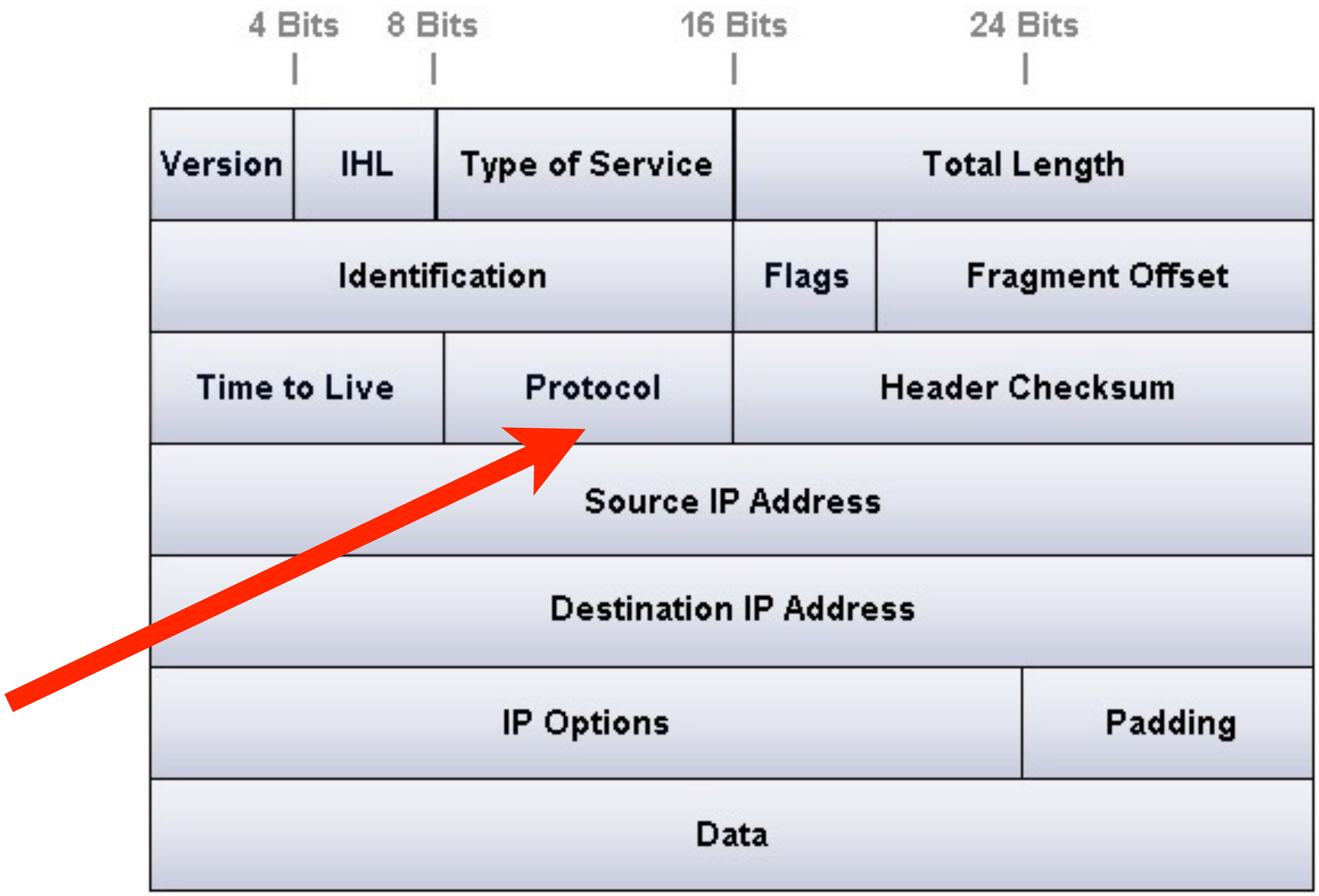
IP datagram format



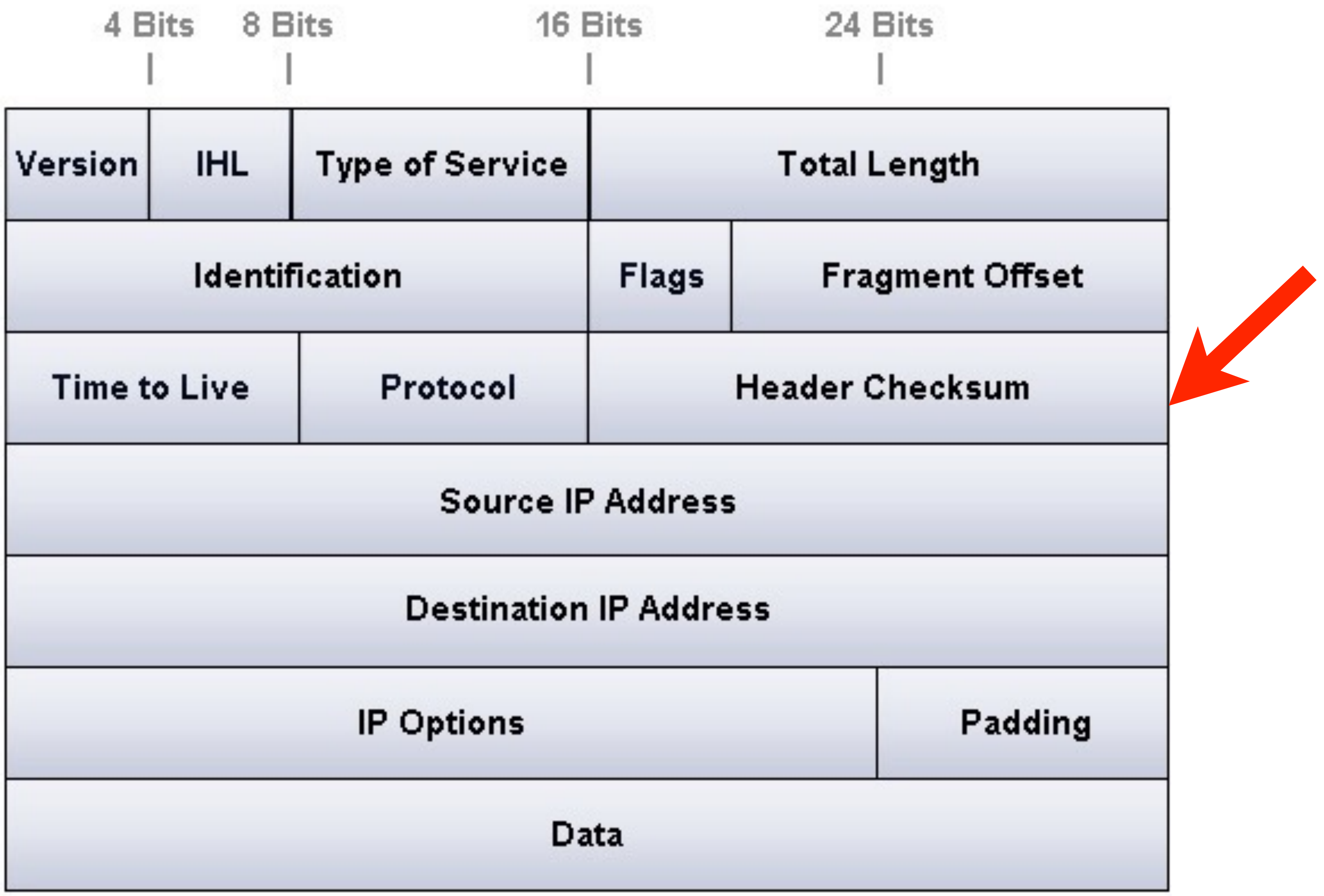
IP datagram format



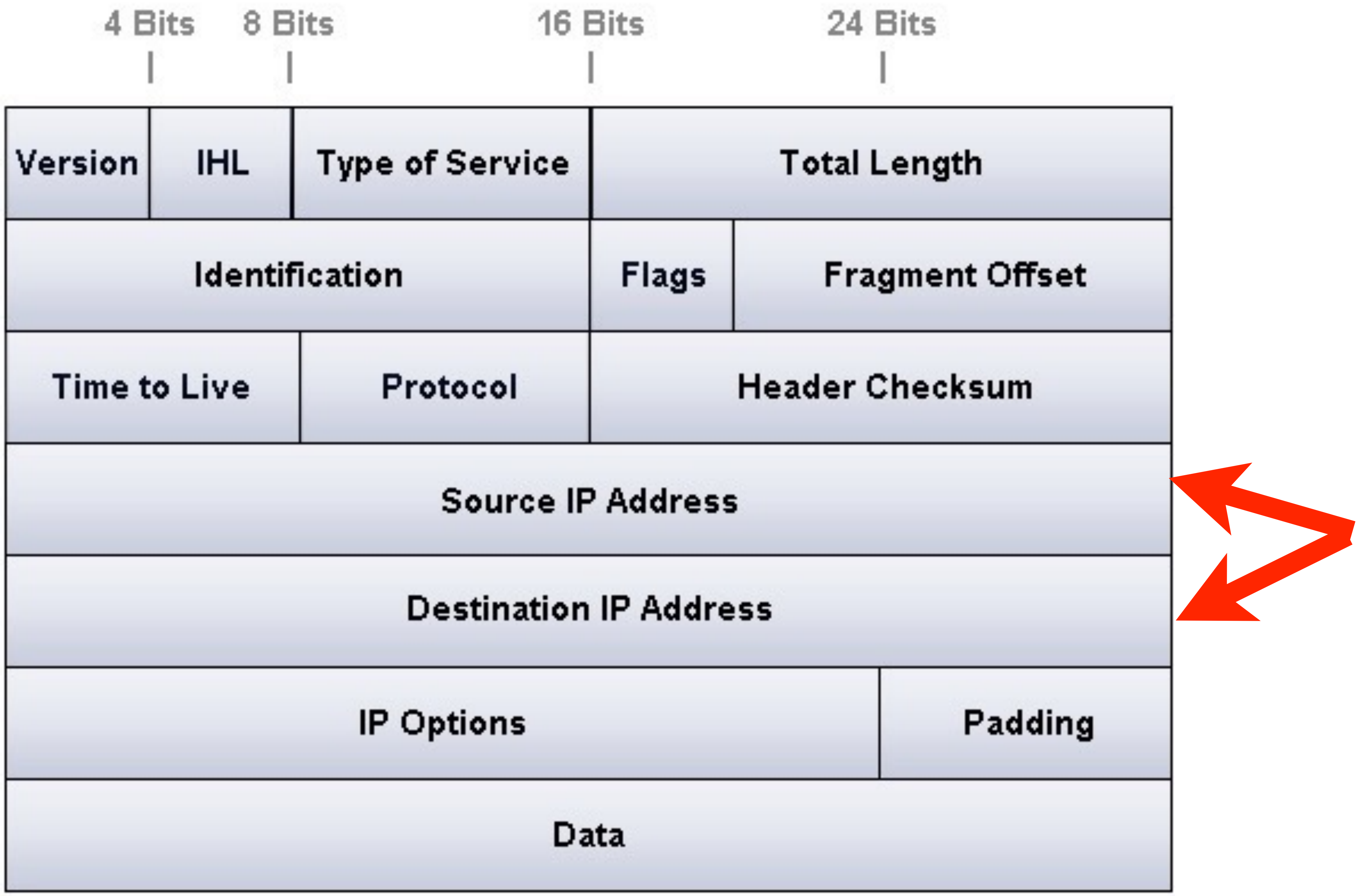
IP datagram format



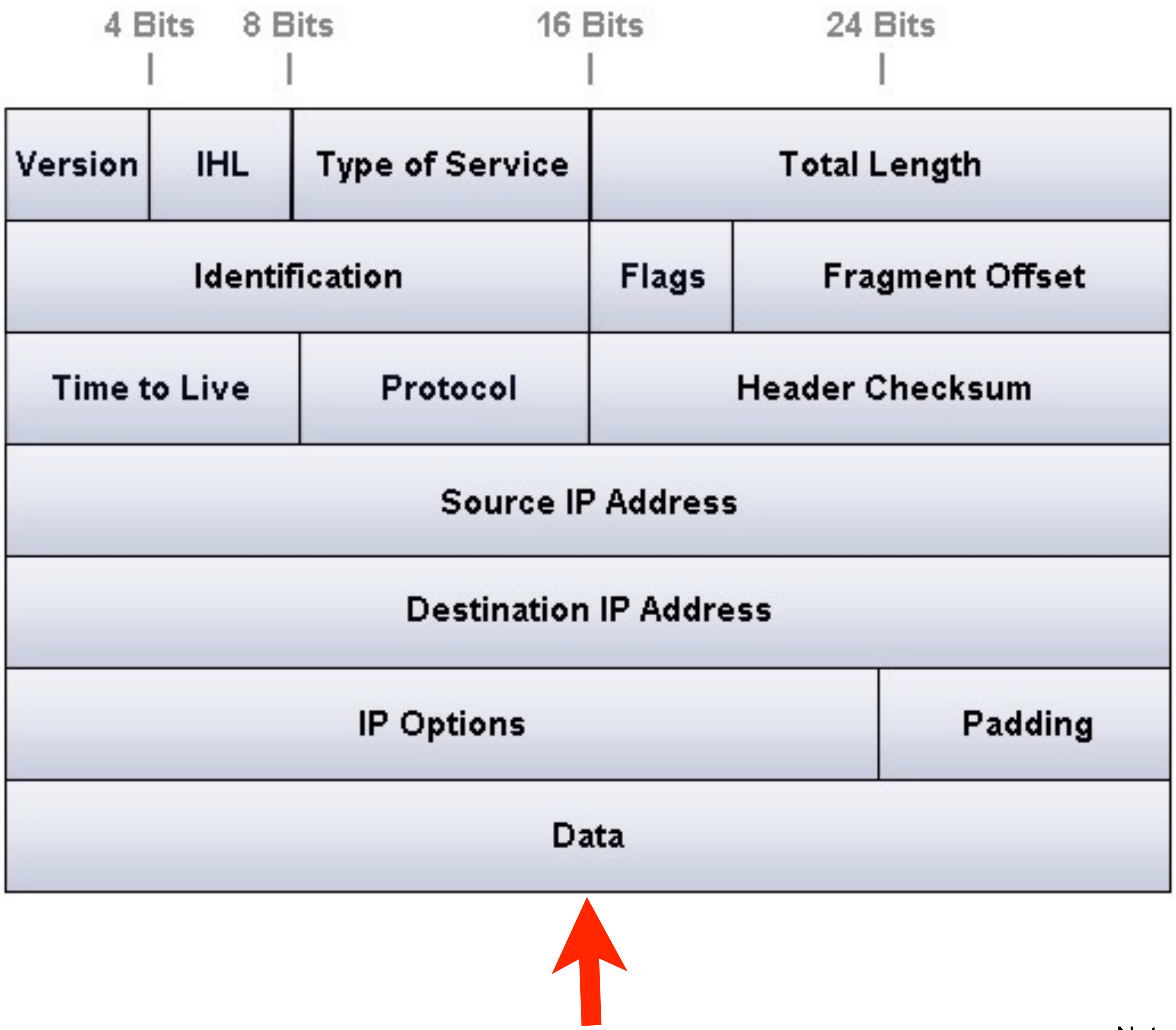
IP datagram format



IP datagram format



IP datagram format



Packet Sizes

Packet Sizes

- How big can an Ethernet frame be?

Packet Sizes

- How big can an Ethernet frame be?
  - 1518 bytes

Packet Sizes

- How big can an Ethernet frame be?
  - 1518 bytes

- How big can an IP packet be?

Packet Sizes

- How big can an Ethernet frame be?
  - 1518 bytes

- How big can an IP packet be?
  - 2^16 = 65,535 bytes

Packet Sizes

- How big can an Ethernet frame be?
  - 1518 bytes

- How big can an IP packet be?
  - 2^16 = 65,535 bytes

- Is this a problem?

Packet Sizes

- How big can an Ethernet frame be?
  - 1518 bytes

- How big can an IP packet be?
  - 2^16 = 65,535 bytes

- Is this a problem?
- Mismatch between LL and IP packet sizes … UGH!

IP Fragmentation & Reassembly

IP Fragmentation & Reassembly

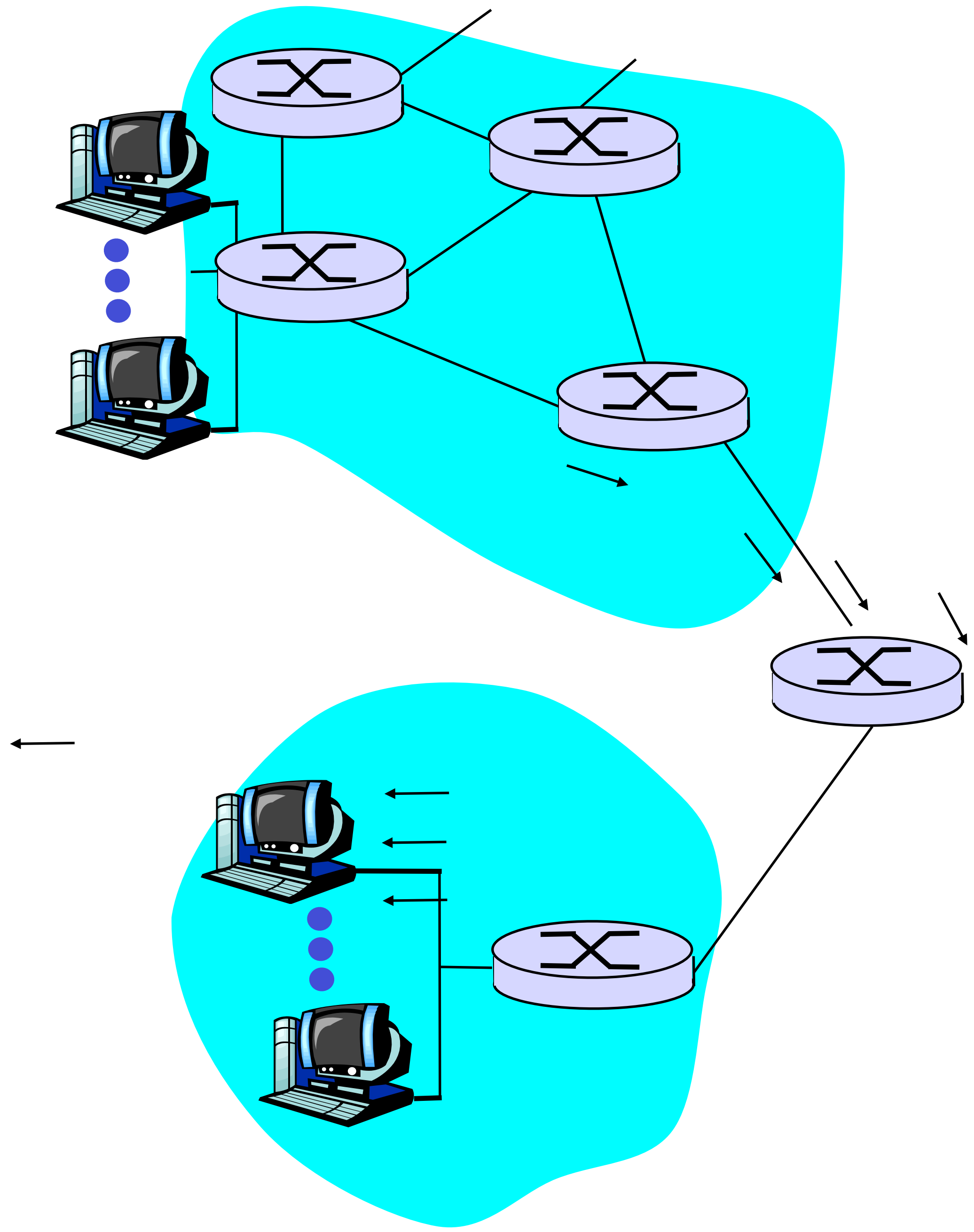
- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs

IP Fragmentation & Reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments

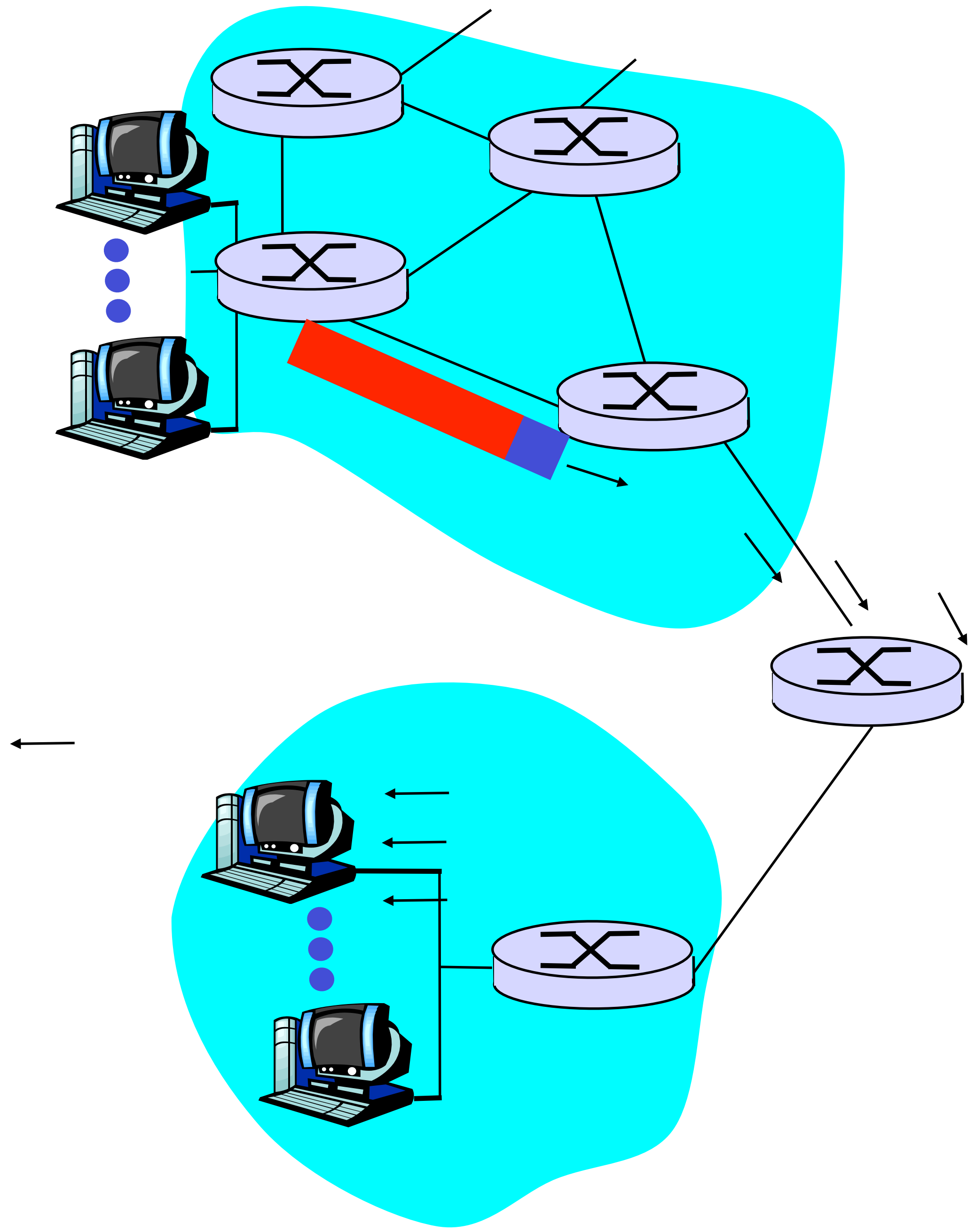
IP Fragmentation & Reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



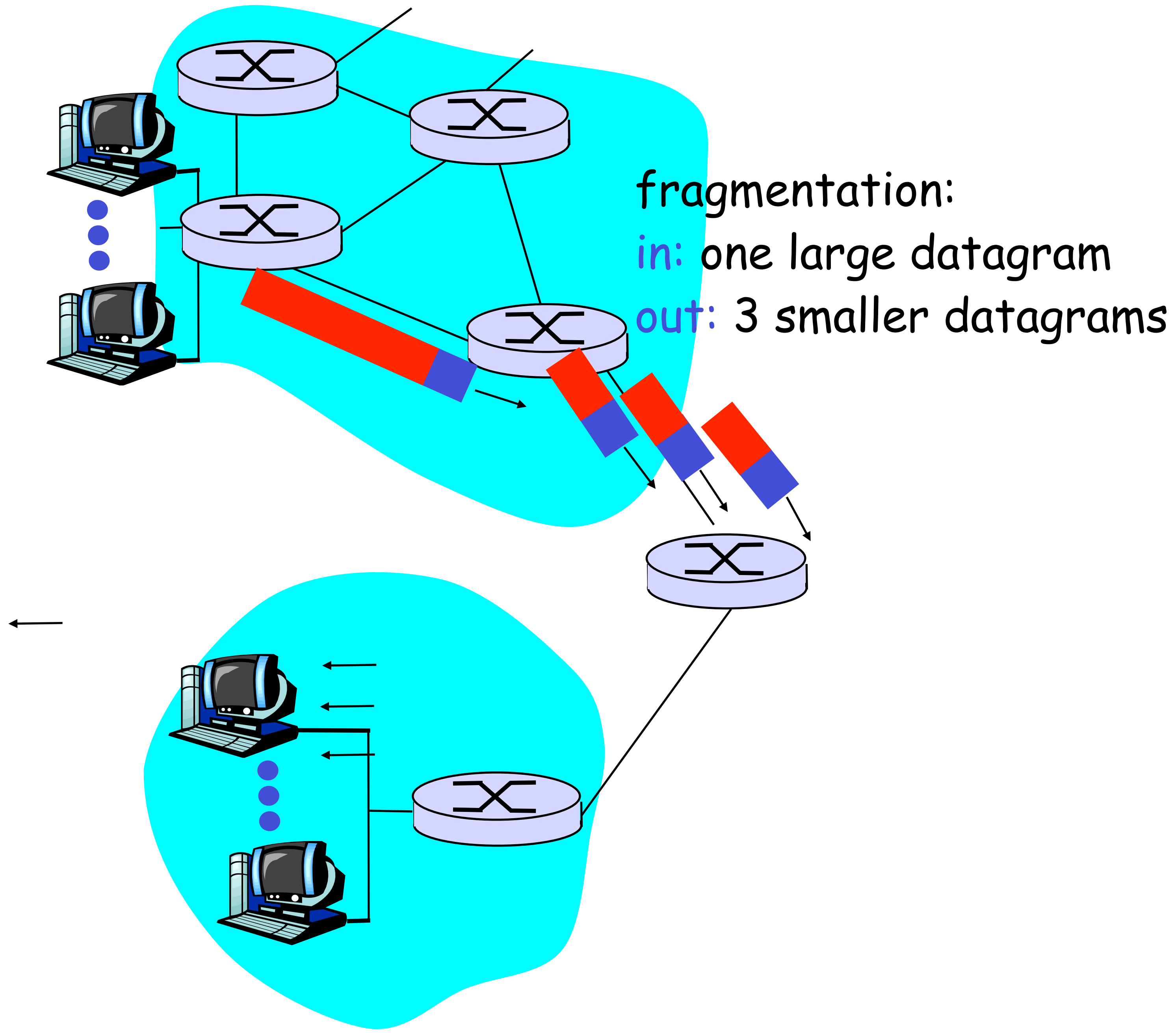
IP Fragmentation & Reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



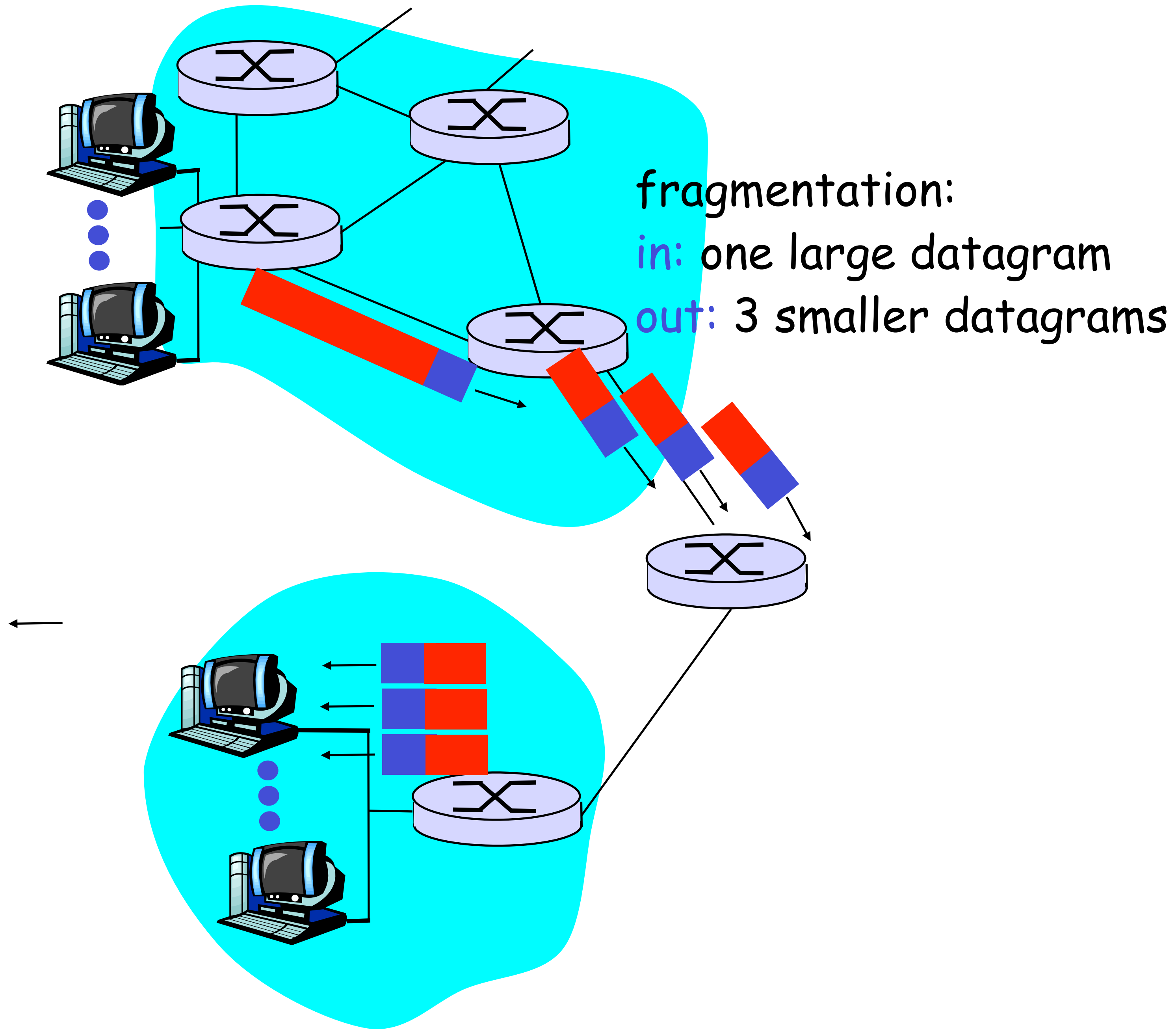
IP Fragmentation & Reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



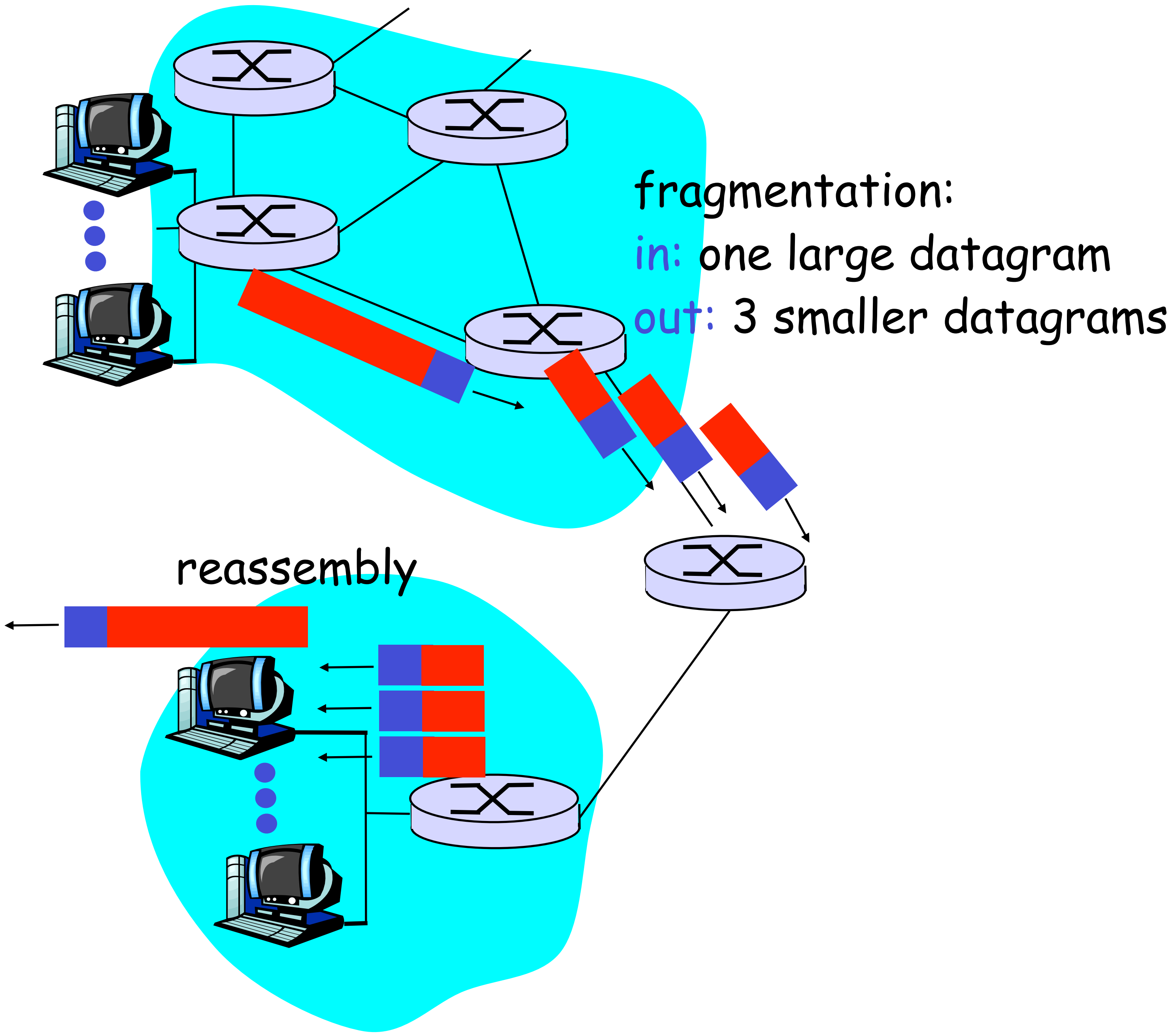
IP Fragmentation & Reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



IP Fragmentation & Reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments

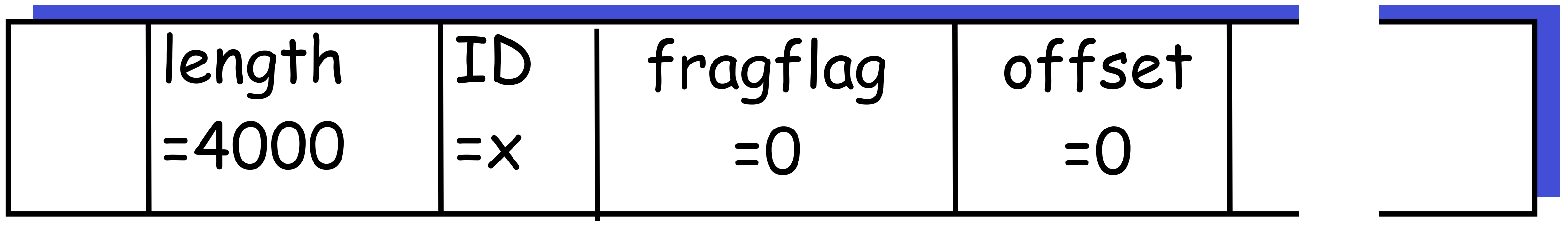


IP Fragmentation and Reassembly

Example

- 4000 byte IP datagram
- MTU = 1500 bytes

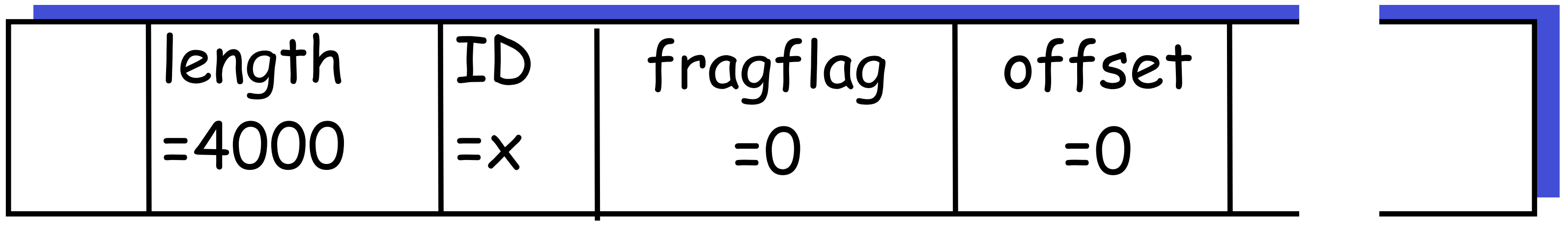
IP Fragmentation and Reassembly



Example

- 4000 byte IP datagram
- MTU = 1500 bytes

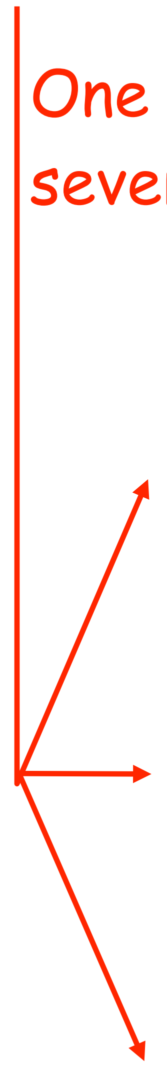
IP Fragmentation and Reassembly



Example

- 4000 byte IP datagram
- MTU = 1500 bytes

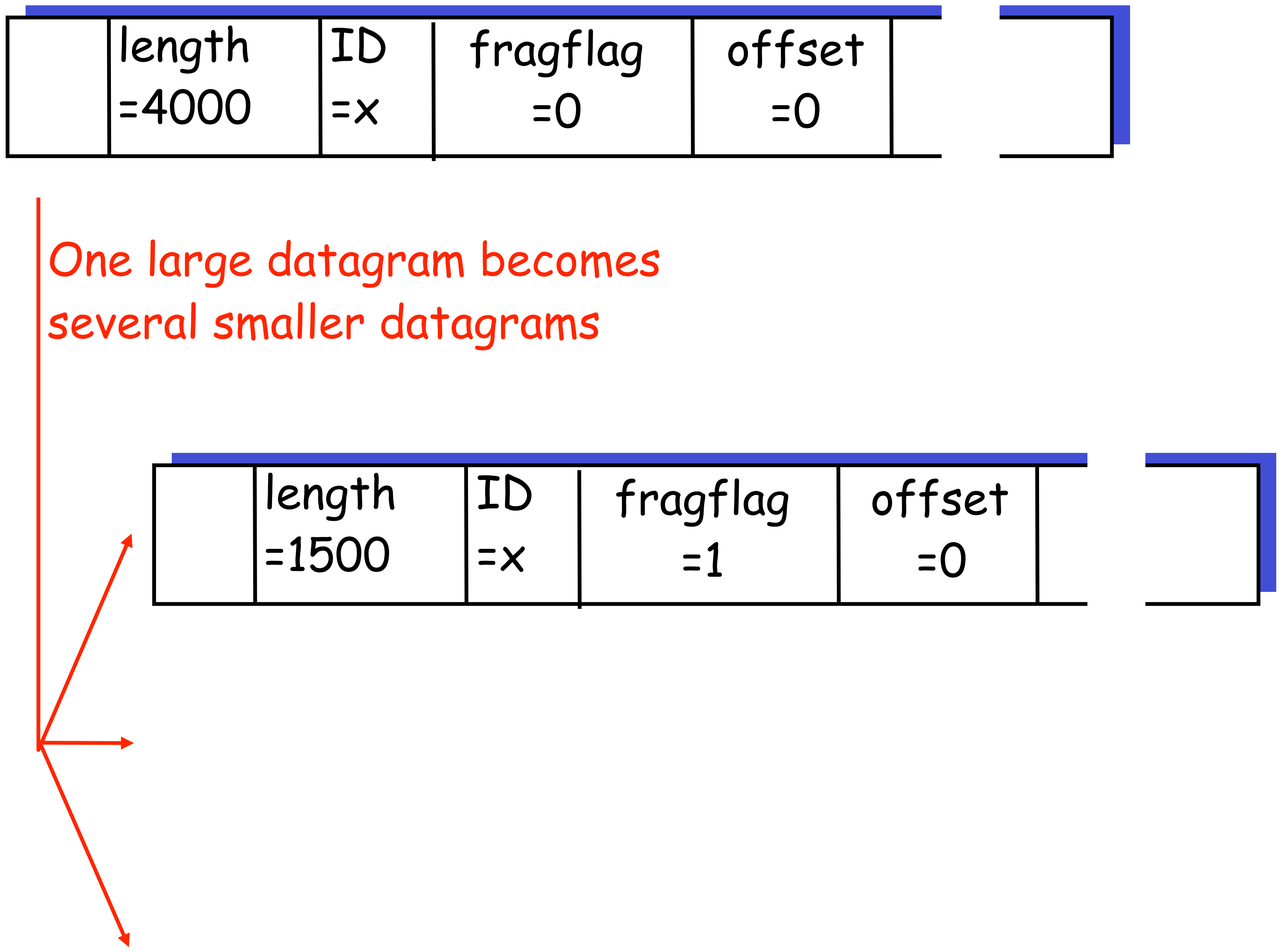
One large datagram becomes several smaller datagrams



IP Fragmentation and Reassembly

Example

- 4000 byte IP datagram
- MTU = 1500 bytes

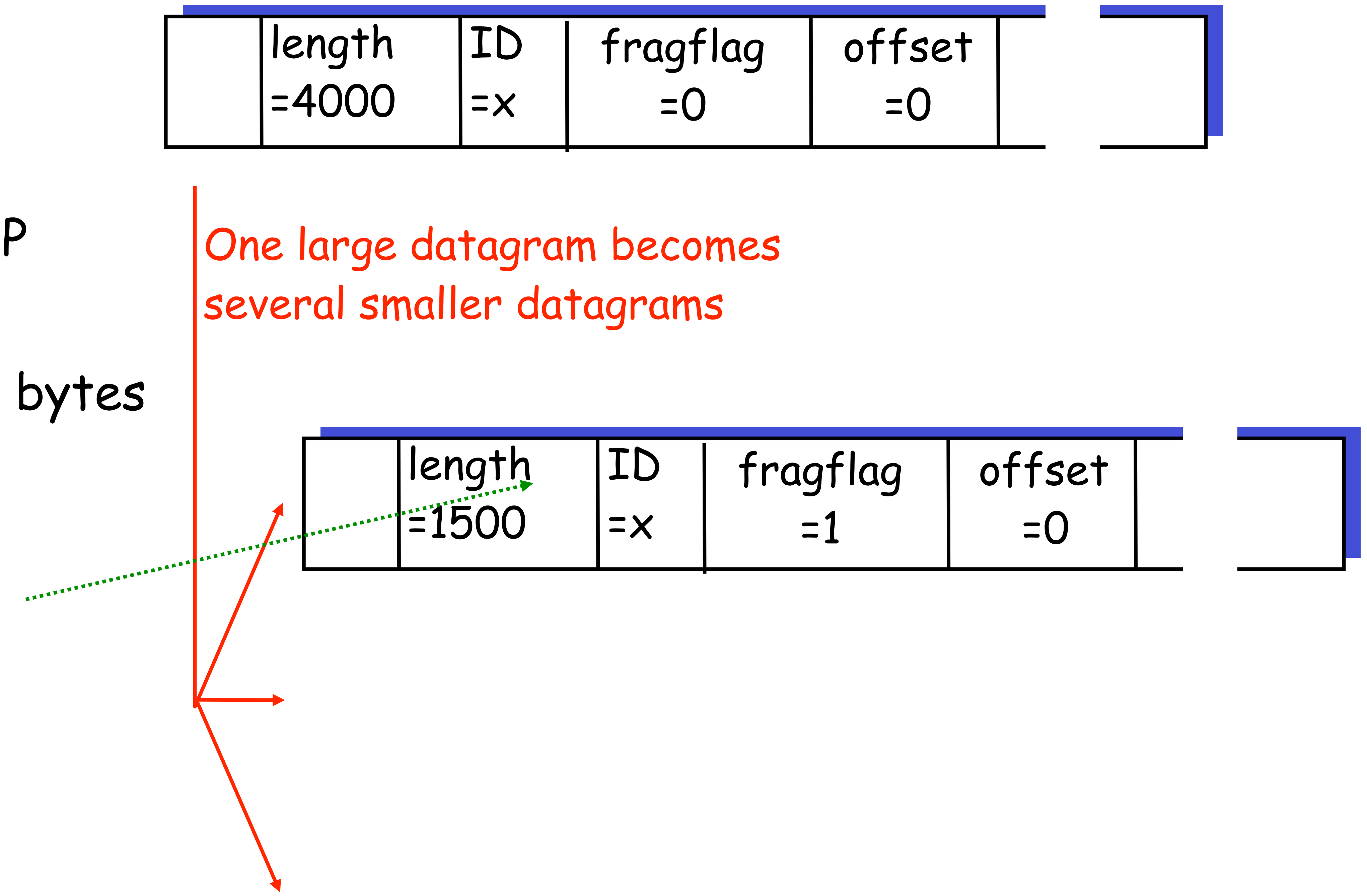


One large datagram becomes several smaller datagrams

IP Fragmentation and Reassembly

Example

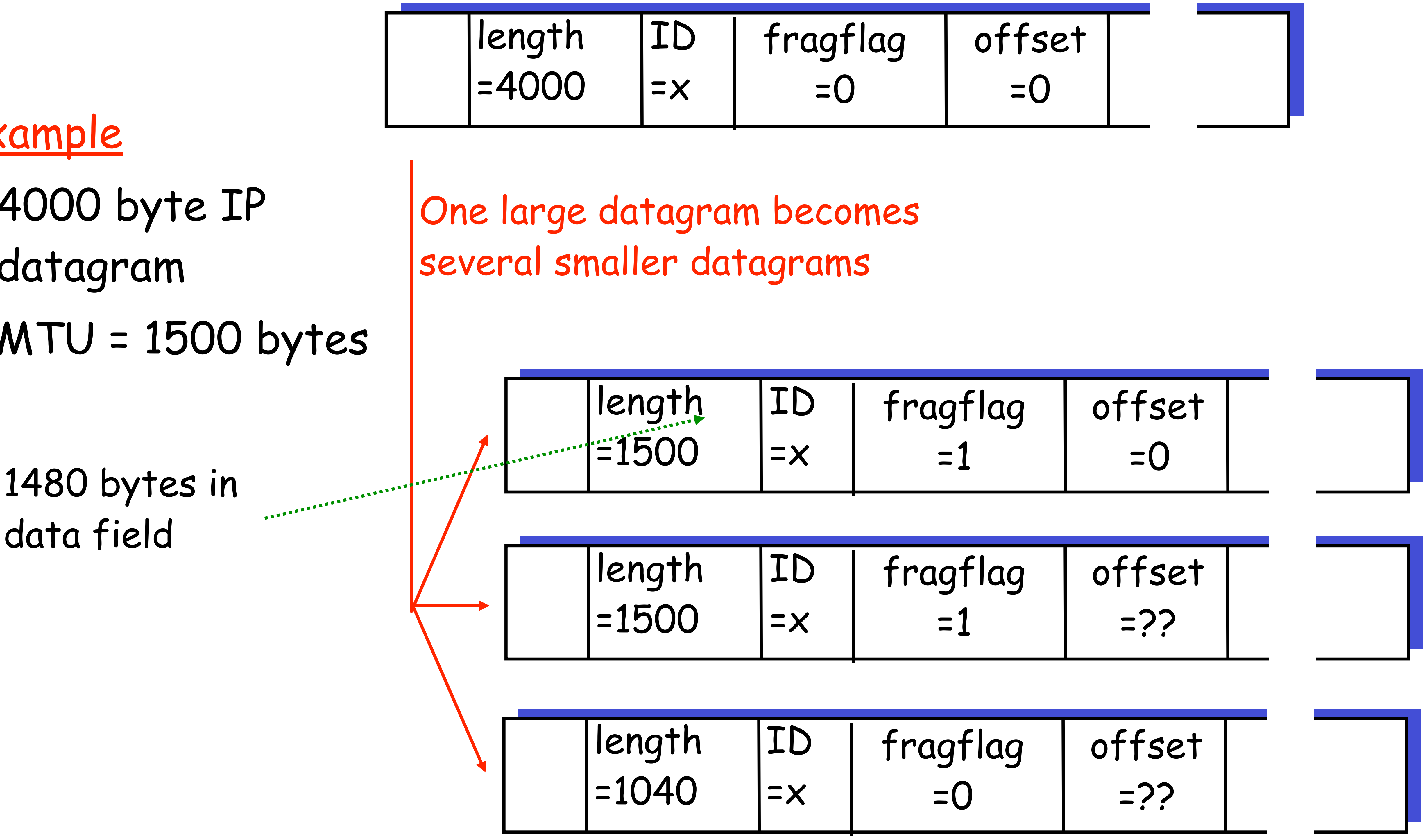
- 4000 byte IP datagram
- MTU = 1500 bytes



IP Fragmentation and Reassembly

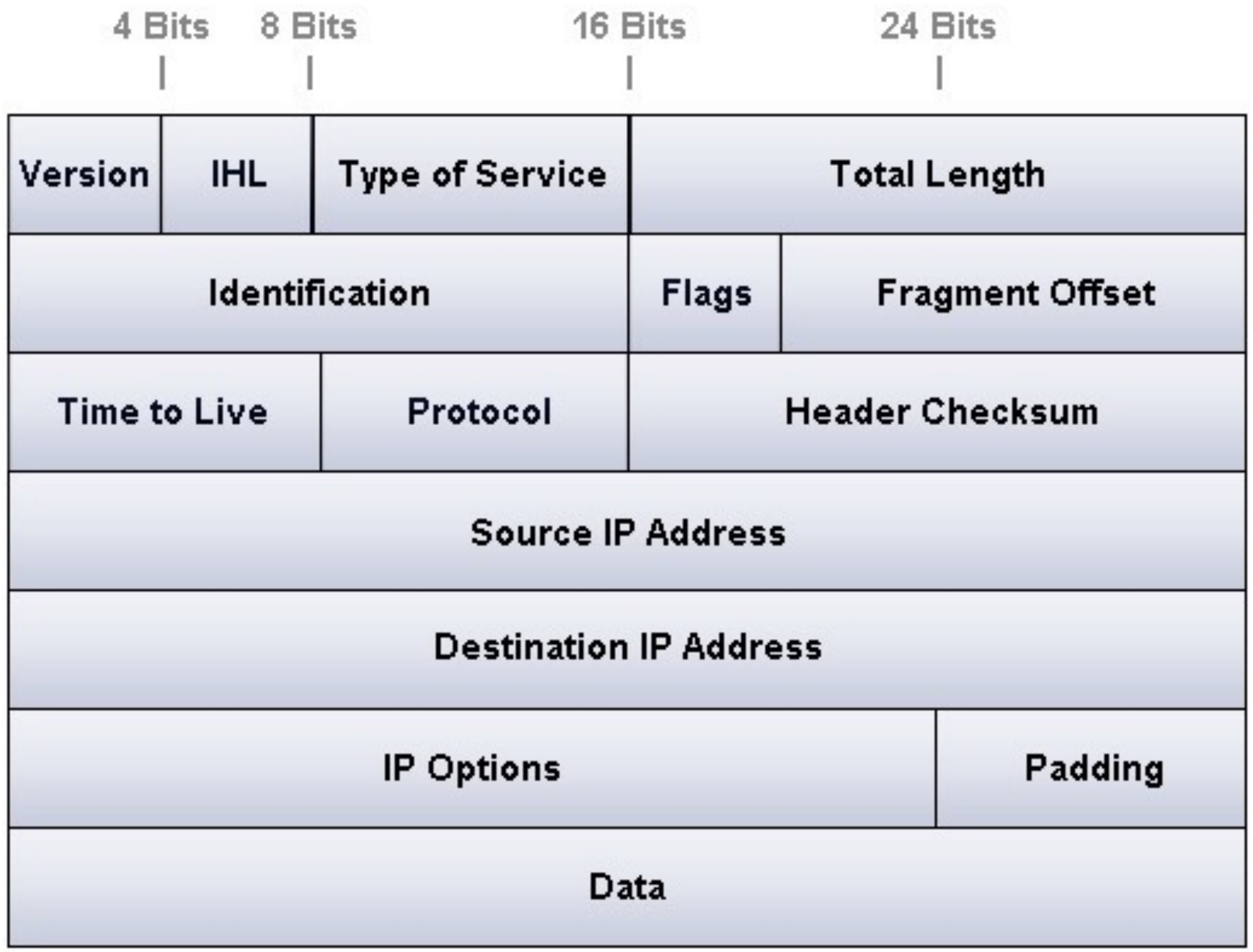
Example

- 4000 byte IP datagram
- MTU = 1500 bytes

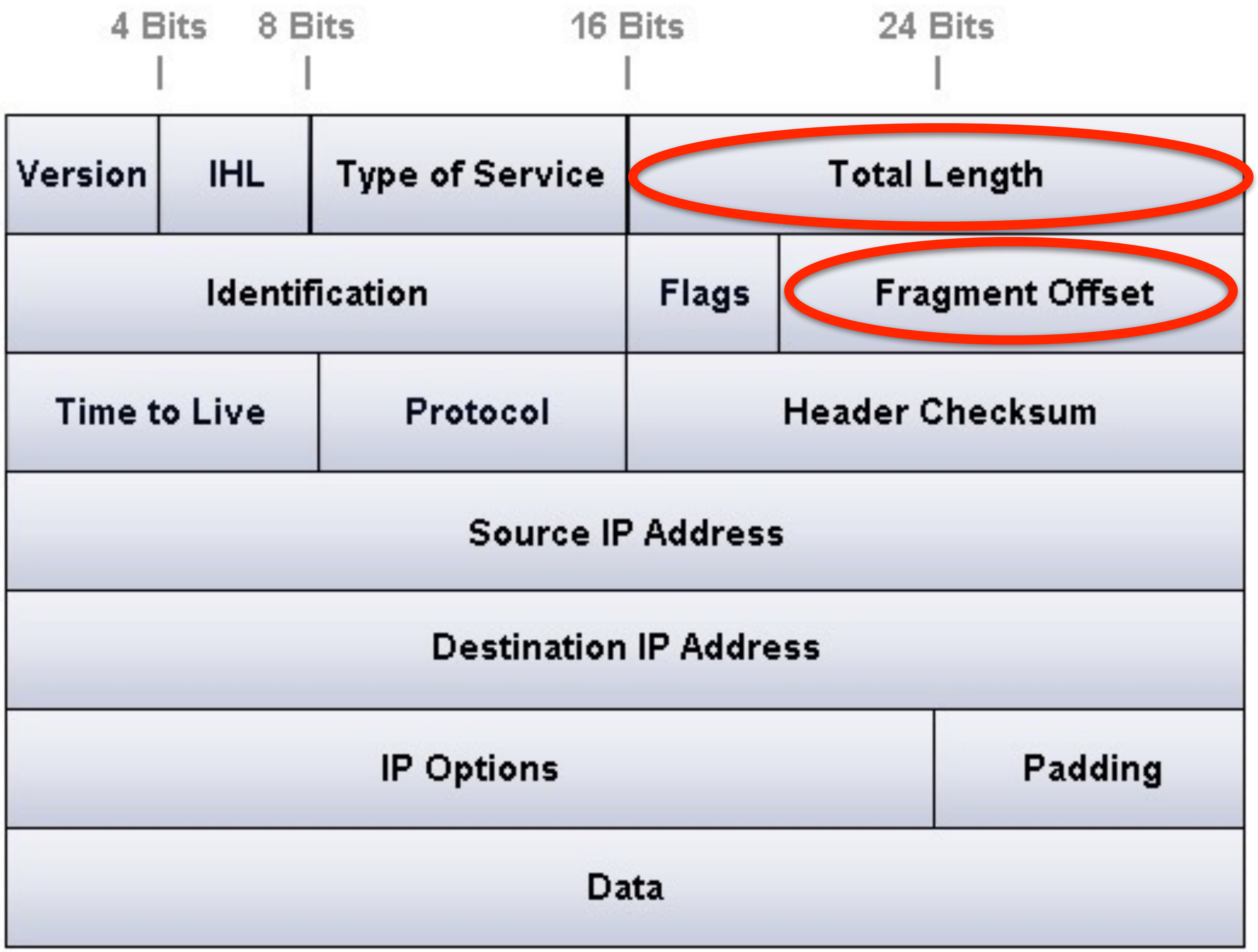


IP Packet Header Revisited

IP Packet Header Revisited



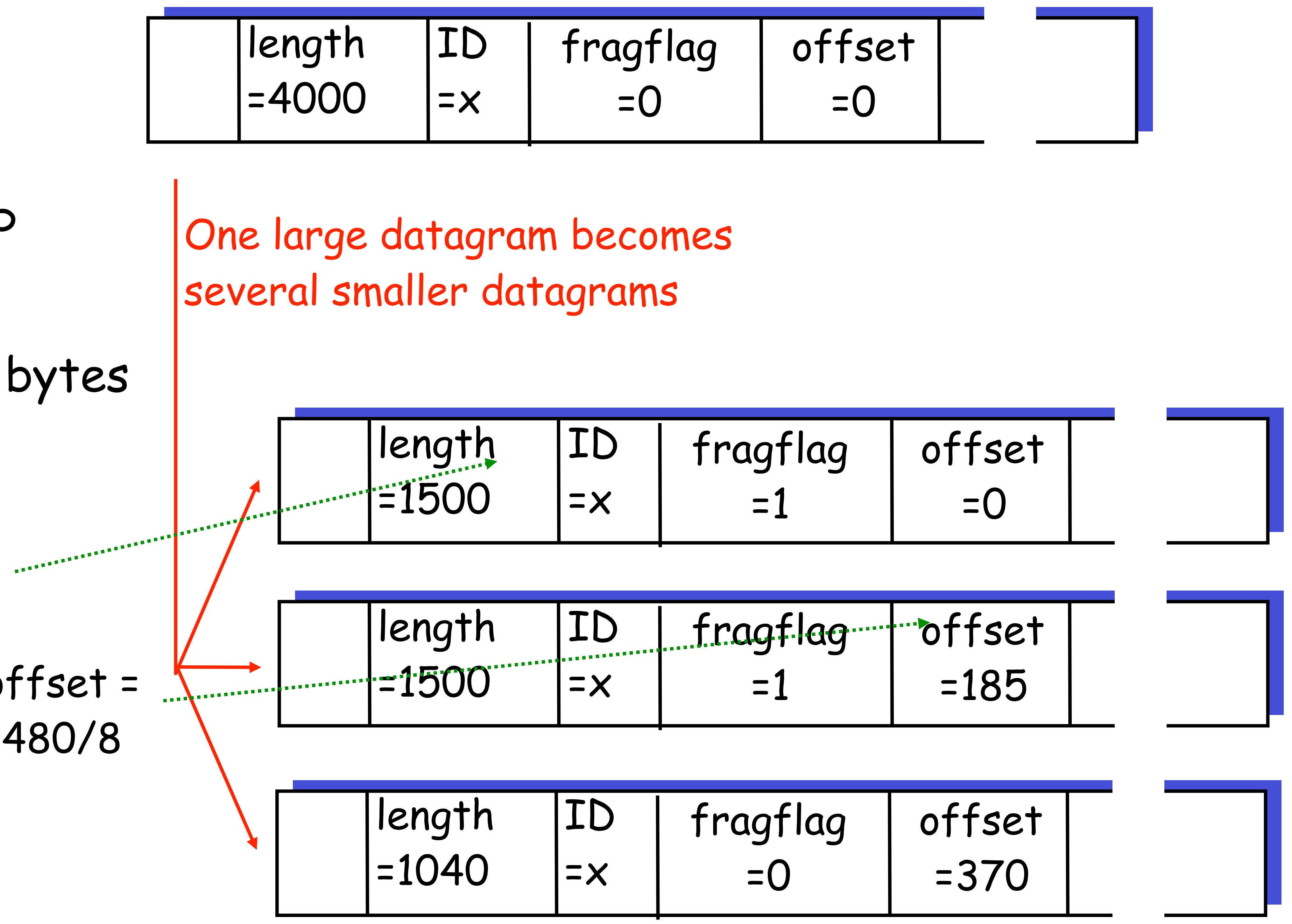
IP Packet Header Revisited



IP Fragmentation and Reassembly

Example

- 4000 byte IP datagram
- MTU = 1500 bytes

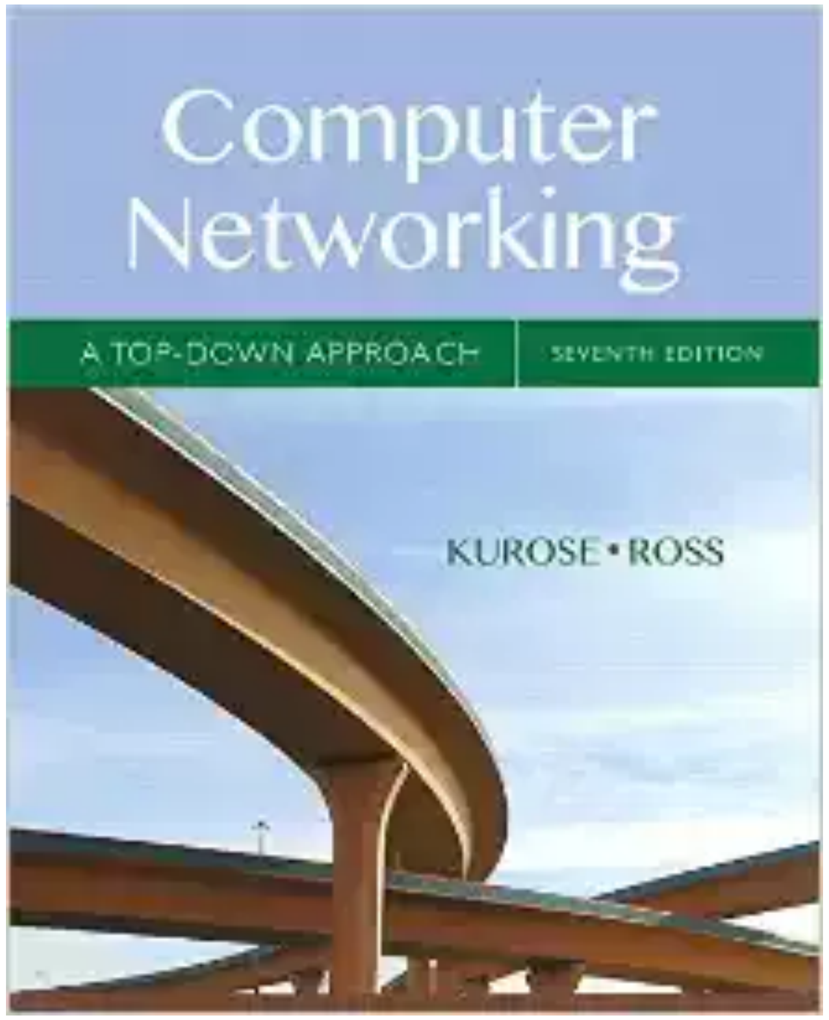


How Do We Know the MTU Size?

How Do We Know the MTU Size?

- Data link layer must "support" a packet of at least 576 bytes

Reading Along ...



- Network layer is chapters 4 & 5
  - 5.6: ICMP

ICMP: Internet Control Message Protocol

ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information

ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol

ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)

ICMP: Internet Control Message Protocol

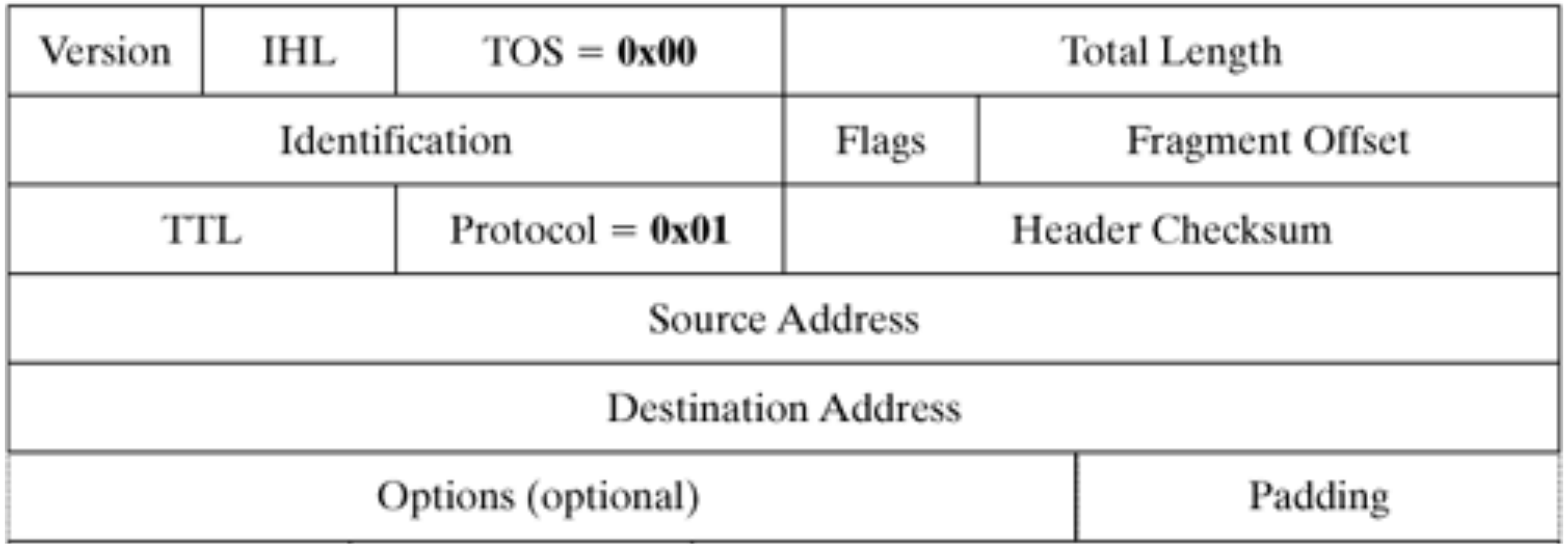
- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP msgs carried in IP datagrams

ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP msgs carried in IP datagrams

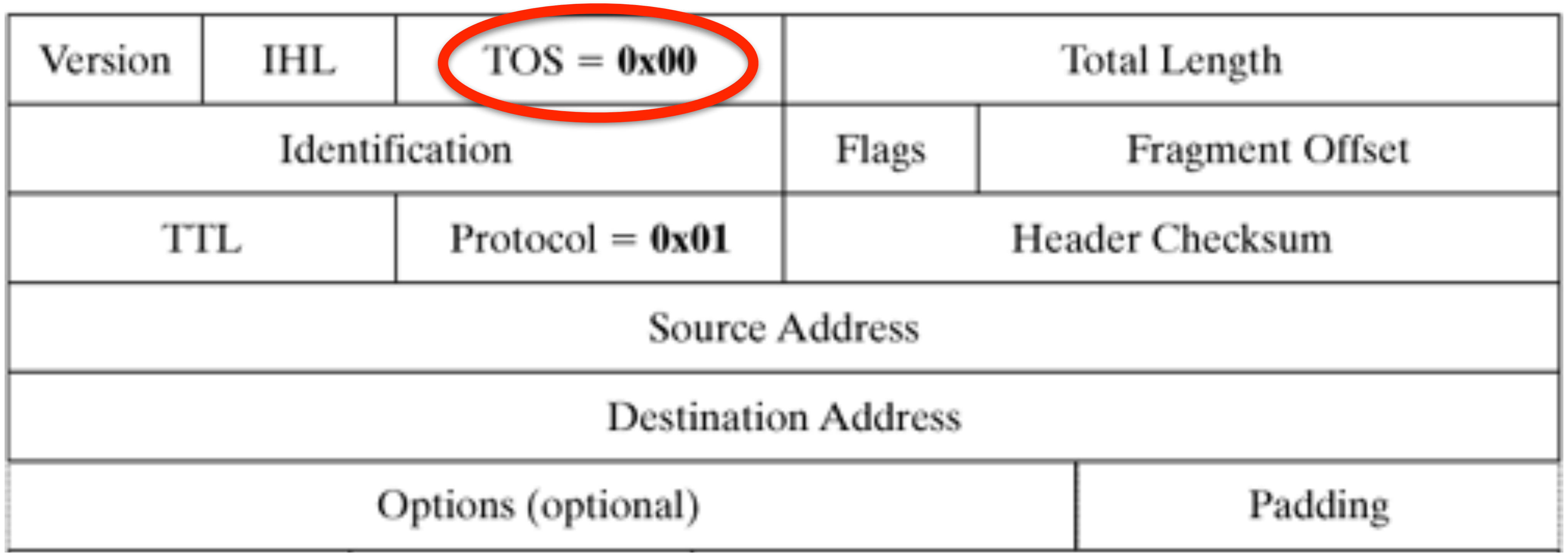
- ICMP message: type, code plus first 8 bytes of IP datagram causing error
  - (sometimes > 8 bytes)

ICMP: Internet Control Message Protocol



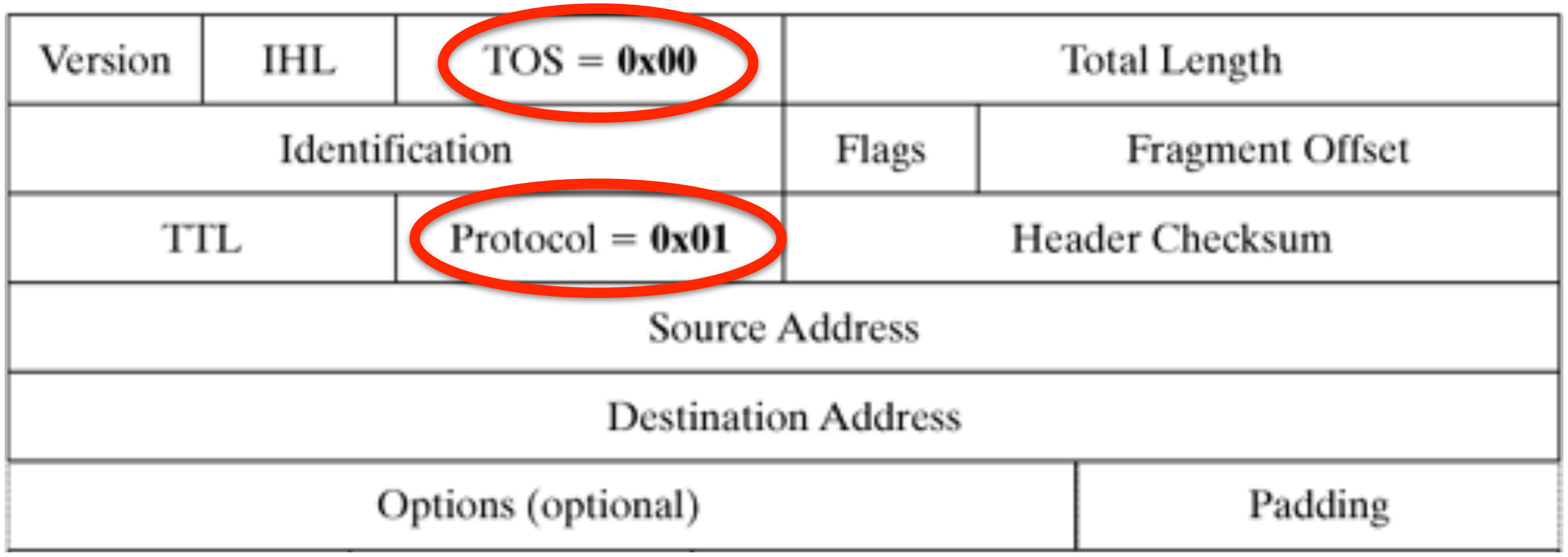
Version	IHL	TOS = **0x00**	Total Length	
Identification			Flags	Fragment Offset
TTL		Protocol = **0x01**	Header Checksum	
Source Address				
Destination Address				
Options (optional)				Padding

ICMP: Internet Control Message Protocol



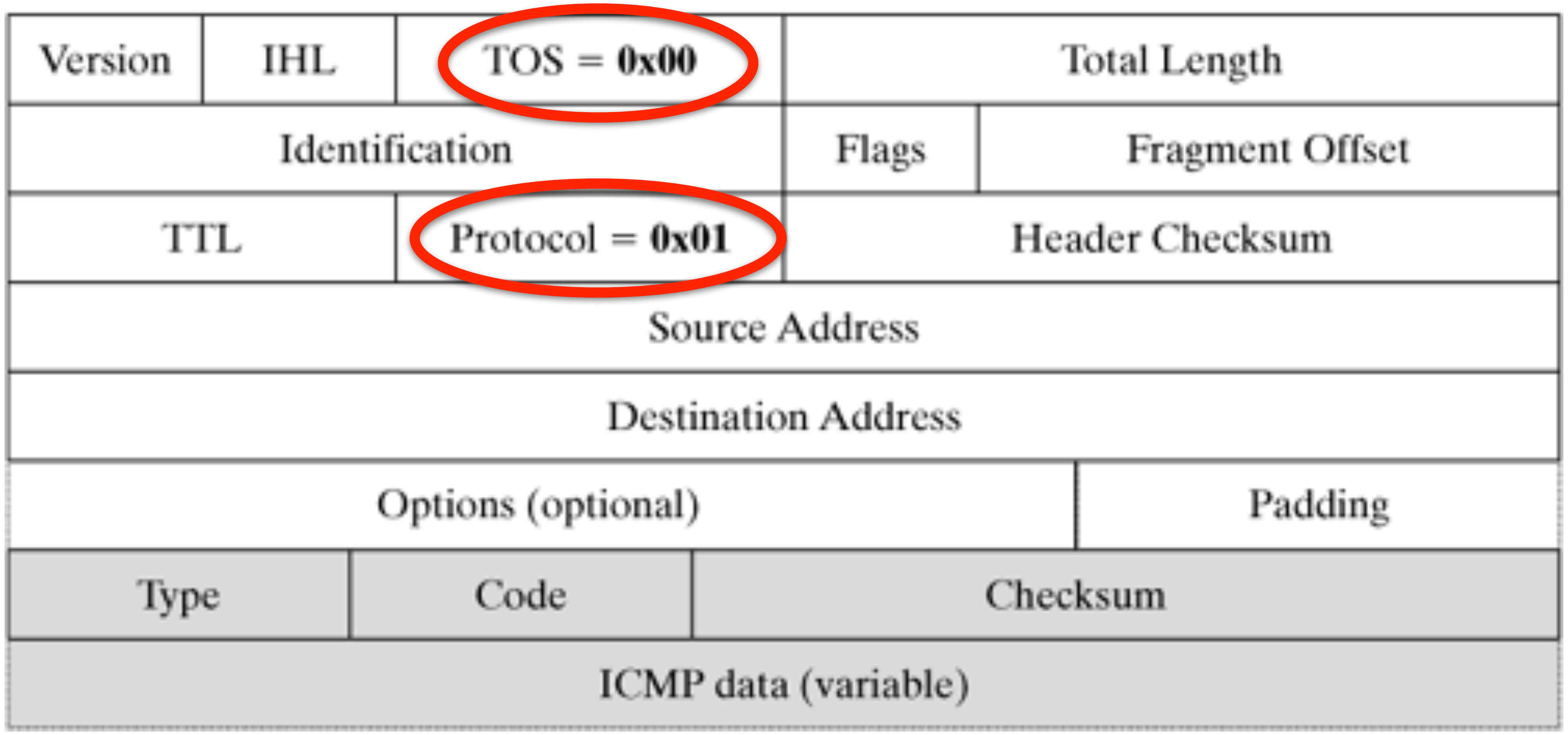
Version	IHL	TOS = **0x00**	Total Length	
Identification			Flags	Fragment Offset
TTL		Protocol = **0x01**	Header Checksum	
Source Address				
Destination Address				
Options (optional)				Padding

ICMP: Internet Control Message Protocol



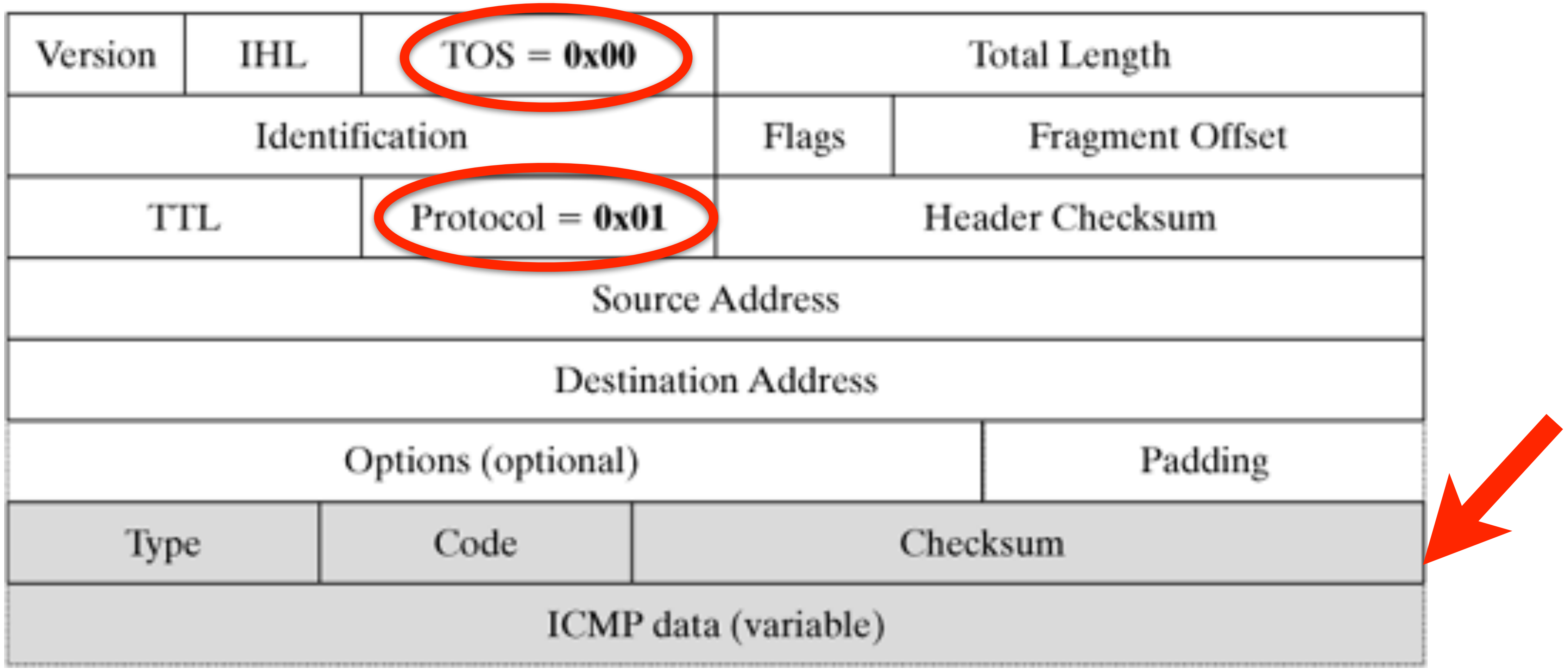
Version	IHL	TOS = **0x00**	Total Length	
Identification			Flags	Fragment Offset
TTL		Protocol = **0x01**	Header Checksum	
Source Address				
Destination Address				
Options (optional)				Padding

ICMP: Internet Control Message Protocol



Version	IHL	TOS = **0x00**	Total Length
Identification		Flags	Fragment Offset
TTL		Protocol = **0x01**	Header Checksum
Source Address			
Destination Address			
Options (optional)			Padding
Type	Code	Checksum	
ICMP data (variable)			

ICMP: Internet Control Message Protocol



ICMP: Internet Control Message Protocol

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	4	packet too big
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ICMP: Internet Control Message Protocol

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	4	packet too big
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ICMP: Internet Control Message Protocol

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	4	packet too big
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ICMP: Internet Control Message Protocol

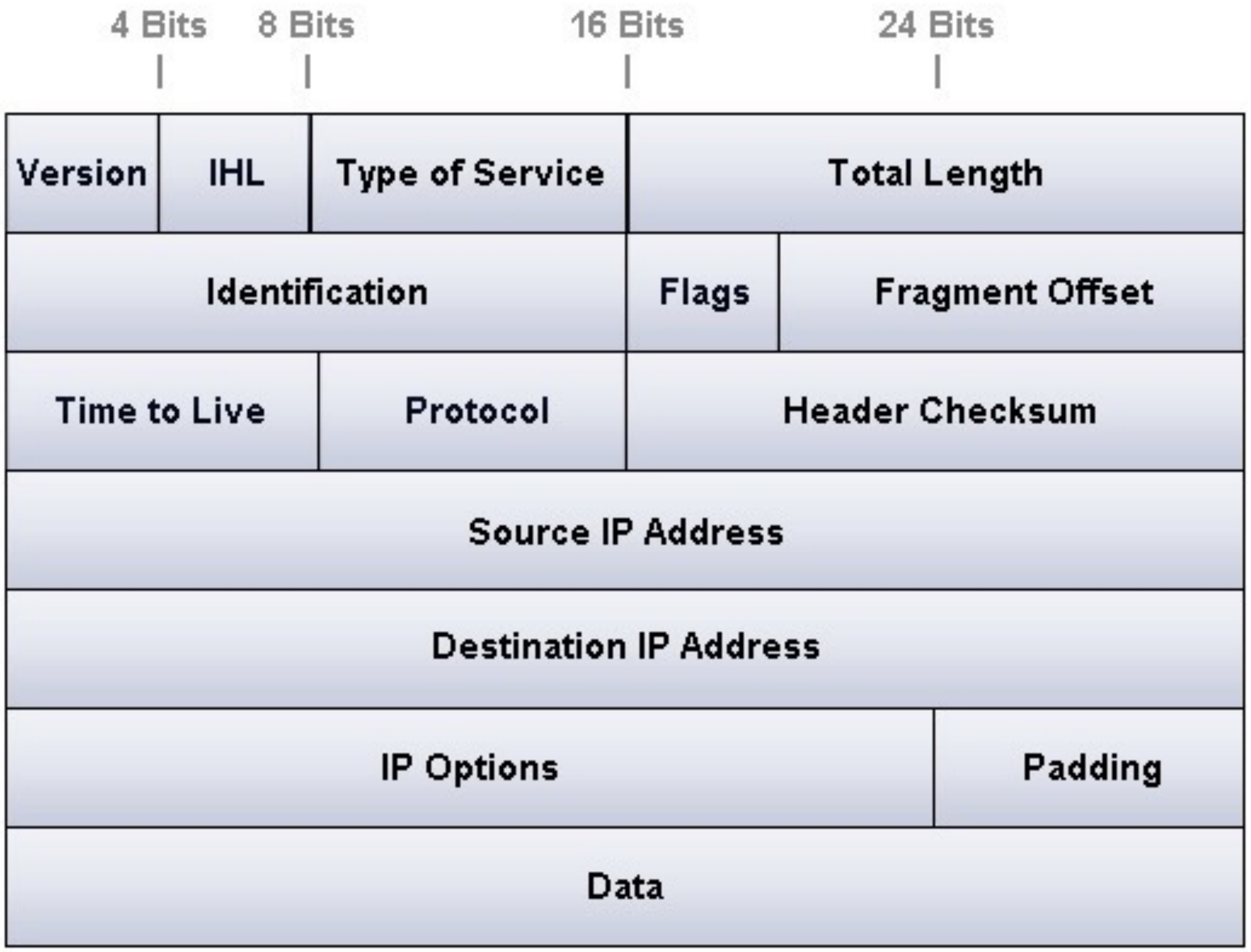
Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	4	packet too big
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ICMP: Internet Control Message Protocol

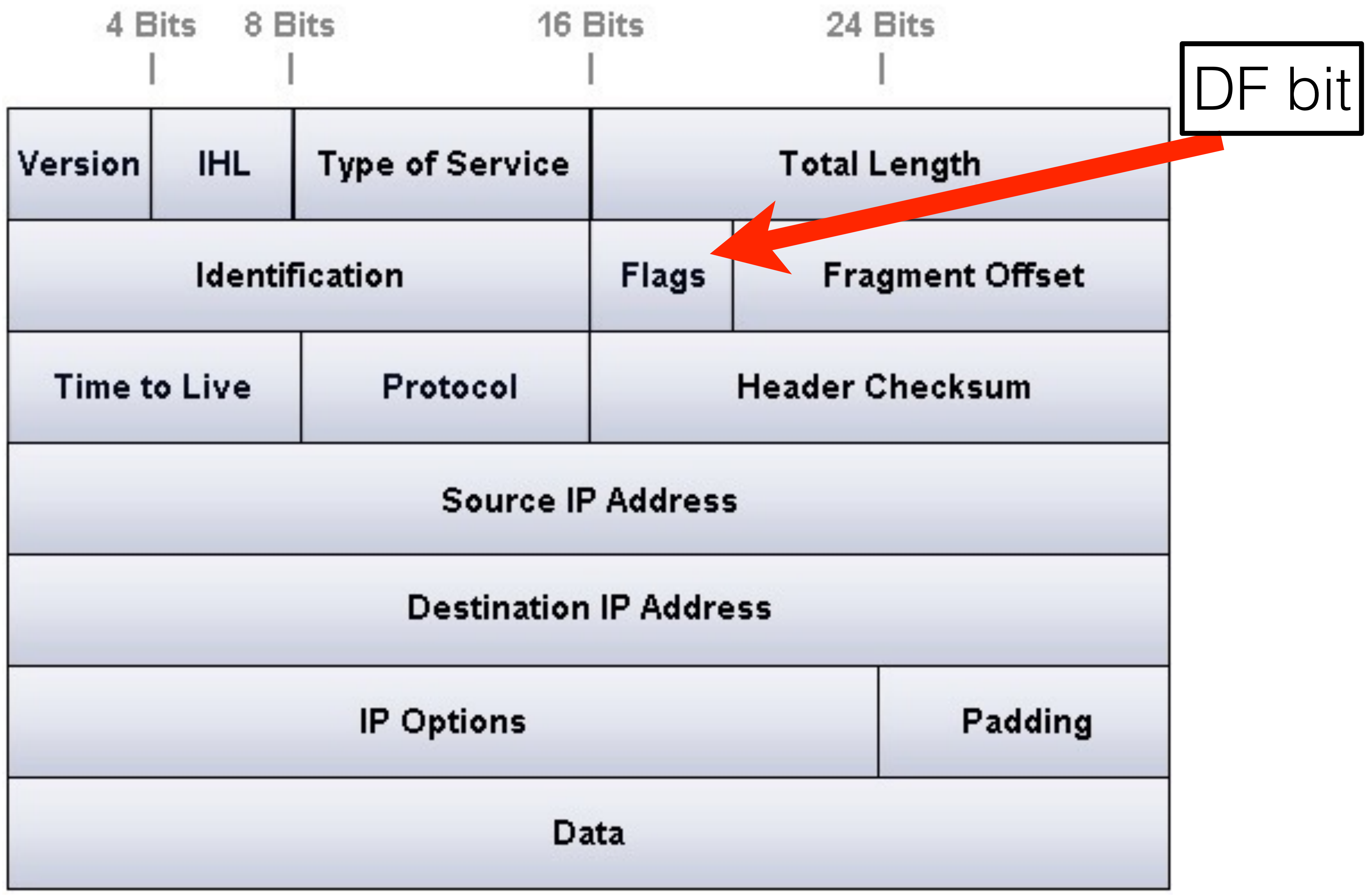
Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	4	packet too big
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Fragmentation … revisited

Fragmentation … revisited



Fragmentation … revisited



ICMP: Internet Control Message Protocol

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	4	packet too big
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Traceroute

Traceroute

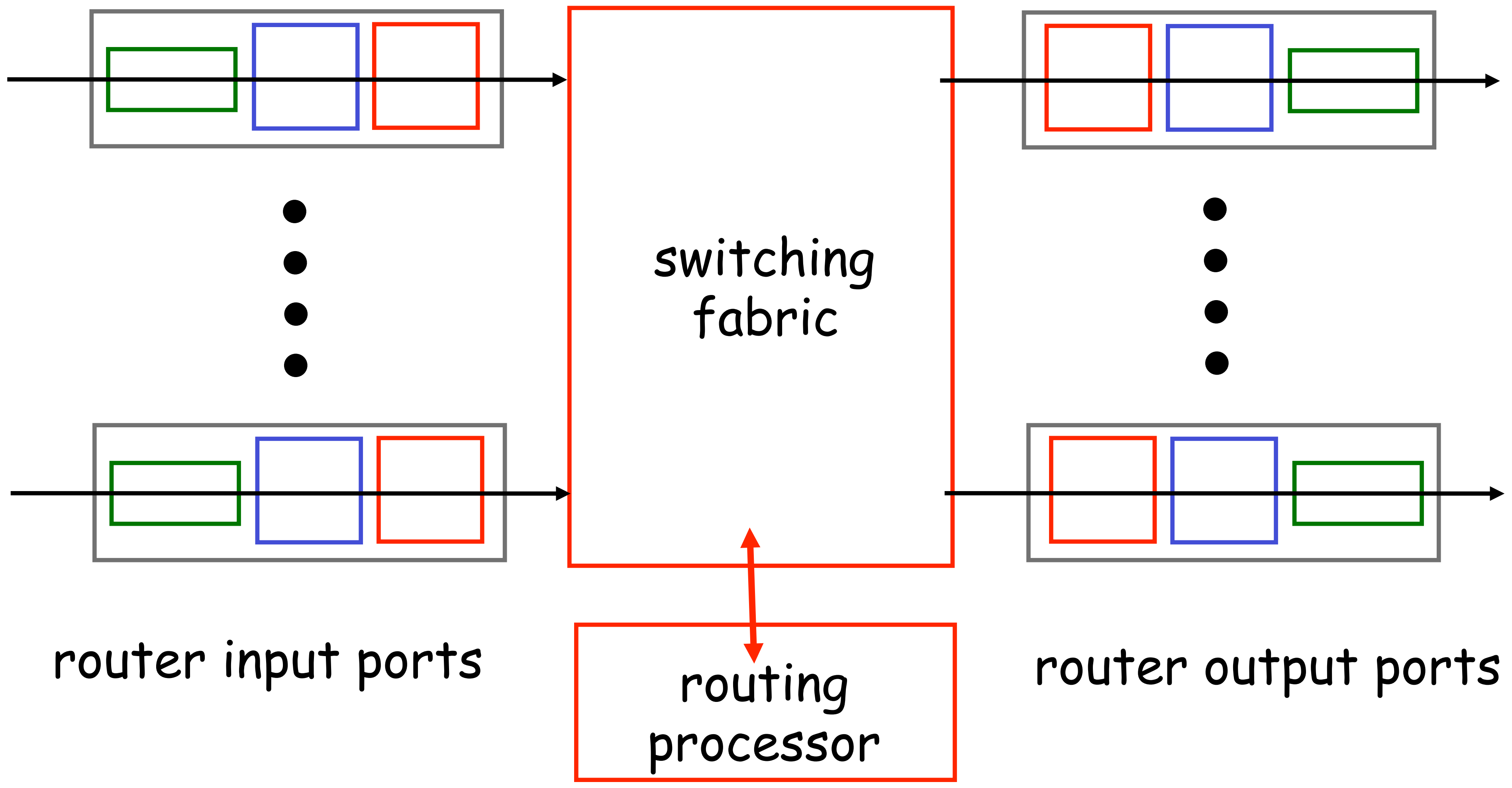
```
% traceroute -n www.icir.org
traceroute to www.icir.org (192.150.187.12), 64 hops max, 52 byte packets
 1  192.168.1.1  4.342 ms
 2  69.222.35.61  4.828 ms
 3  69.222.35.254  10.028 ms
 4  12.81.252.194  19.145 ms
 5  12.122.132.197  23.057 ms
 6  * * *
 7  * * *
 8  * * *
 9  4.15.122.46  164.480 ms
10  137.164.46.144  177.389 ms
11  137.164.50.31  164.832 ms
12  128.32.0.37  172.196 ms
13  128.32.0.83  170.924 ms
14  169.229.0.141  170.854 ms
15  192.150.187.249  170.783 ms
16  192.150.187.12  72.084 ms
```

Traceroute

```
% traceroute -n www.icir.org
traceroute to www.icir.org (192.150.187.12), 64 hops max, 52 byte packets
 1  192.168.1.1  4.342 ms
 2  69.222.35.61  4.828 ms
 3  69.222.35.254  10.028 ms
 4  12.81.252.194  19.145 ms
 5  12.122.132.197  23.057 ms
 6  * * *
 7  * * *
 8  * * *
 9  4.15.122.46  164.480 ms
10  137.164.46.144  177.389 ms
11  137.164.50.31  164.832 ms
12  128.32.0.37  172.196 ms
13  128.32.0.83  170.924 ms
14  169.229.0.141  170.854 ms
15  192.150.187.249  170.783 ms
16  192.150.187.12  72.084 ms
```

- How does this work?

Router … Revisited



Traceroute and ICMP

Traceroute and ICMP

- Source sends series of packets to the destination
  - first has TTL =1
  - second has TTL=2, etc.

Traceroute and ICMP

- Source sends series of packets to the destination
  - first has TTL =1
  - second has TTL=2, etc.

- When n-th packet arrives to n-th router:
  - router discards packet (why?)
  - and sends to source an ICMP message (type 11, code 0)
  - ICMP message includes the IP address of router

Traceroute and ICMP

- Source sends series of packets to the destination
  - first has TTL =1
  - second has TTL=2, etc.

- When n-th packet arrives to n-th router:
  - router discards packet (why?)
  - and sends to source an ICMP message (type 11, code 0)
  - ICMP message includes the IP address of router

- Probe segment eventually arrives at destination host
  - destination returns ICMP "port unreachable" packet (type 3, code 3)
  - when source gets this ICMP, stops