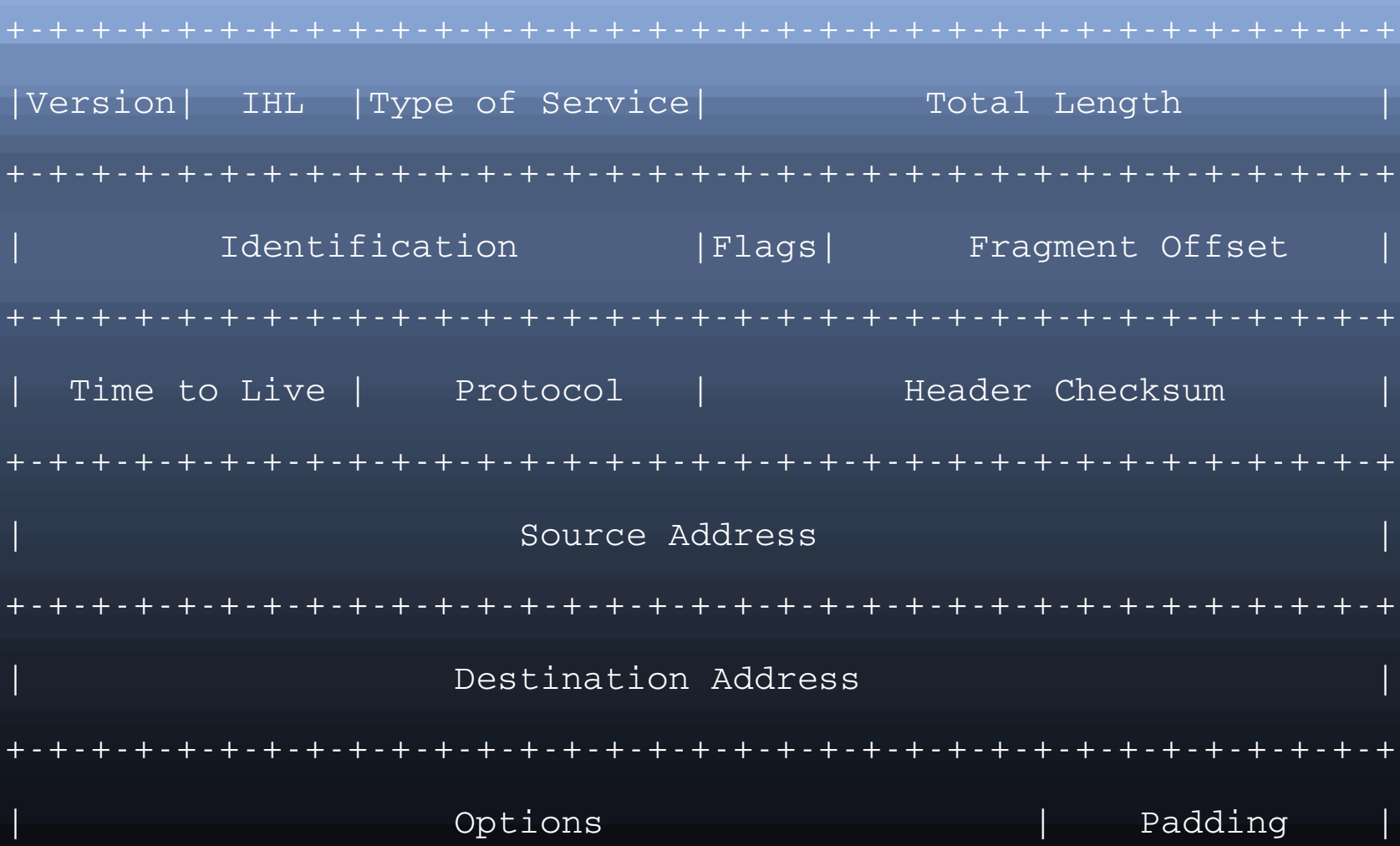
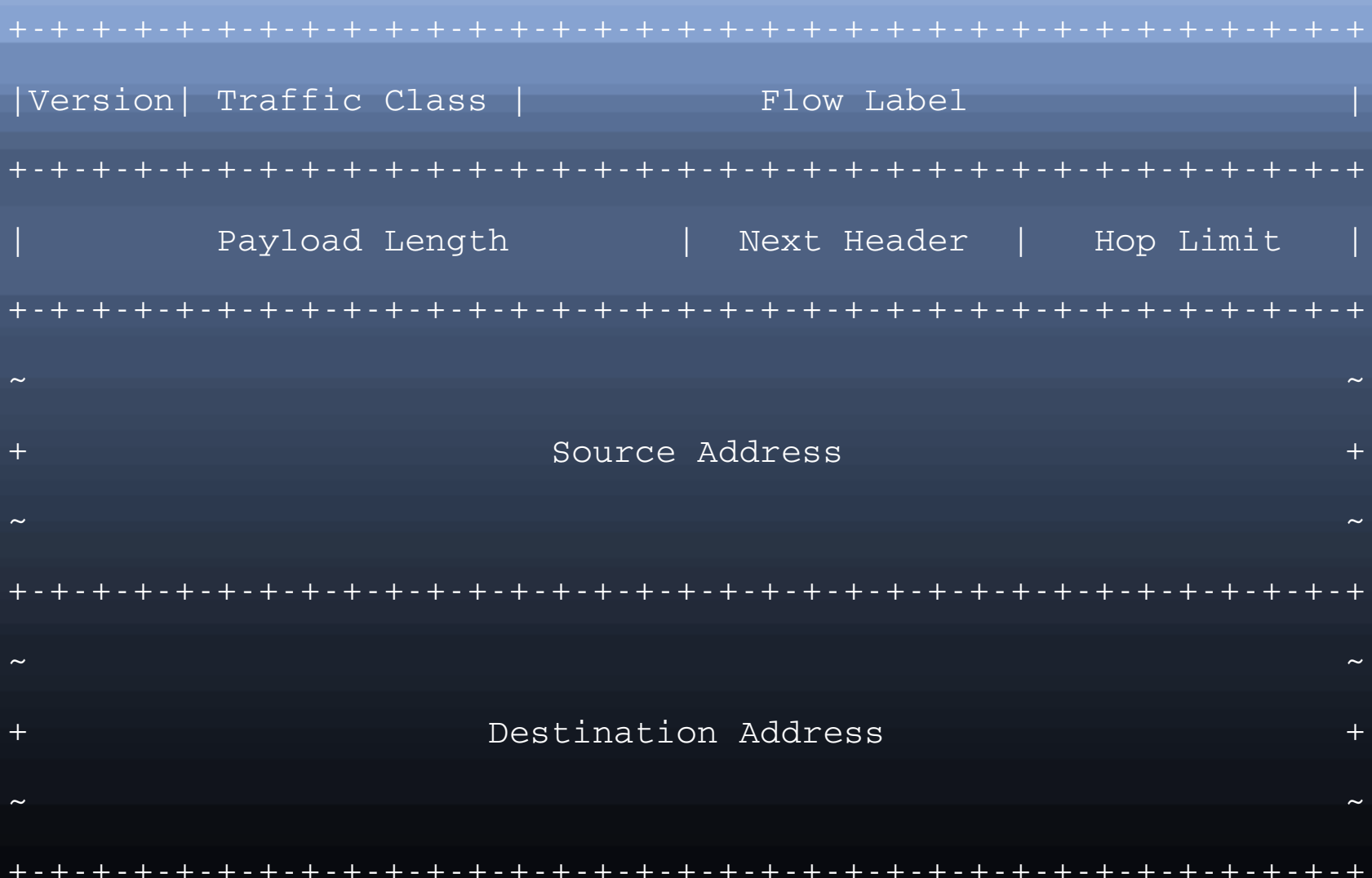
ICS 451: Today's plan

- IP headers
- fragmentation
- IPv6 socket programming
- ICMP
  - ping
  - traceroute

IPv4 header


Version
IHL
Type of Service
Total Length
Identification
Flags
Fragment Offset
Time to Live
Protocol
Header Checksum
Source Address
Destination Address
Options
Padding

IPv6 header



- or see wikipedia, IPv6, under packet format

IP fragmentation

- when a datagram is larger than the MTU of the outgoing interface, IP fragments it
  - in IPv6, only senders fragment, routers don't
  - routers can only fragment if the Don't Fragment (DF) flag is zero
- fragments are IP packets with IP headers
- each fragment carries the same ID
  - hopefully different from IDs of other datagrams that have the same source and destination

IP fragment offset

- the destination needs to reassemble incoming fragments into a whole datagram
- each fragment carries a fragment offset
  - the place where the fragment's payload fits in the reassembled payload
- e.g. fragment offset 144: this payload is to be placed beginning at byte offset 144 in the reassembled payload
- offsets must be a multiple of 8 bytes
  - and the 3 least significant bits are not sent
  - so the offset field is 13 bits long

IP reassembly

- see if already reassembling for this (source,destination,packetID)
  - if not, create an empty reassembled payload
- save the payload at the appropriate offset in the reassembled payload
  - need to keep track of where we have received data for the payload
- if the packet is complete, give to TCP/UDP
  - know size after receiving last fragment
    - identified by MoreFragments (MF) flag = 0

Ping of Death

- late 1980s
- someone sent an artificial fragment with
  - fragment offset + fragment size > 2^{16}
- some implementations copied packet into memory at given offset without checking
  - others overflowed 16-bit variables
- something important was after the buffer
- could crash systems from across the Internet!

  (didn't have to be a ping)

IPv6 fragmentation

- only done by senders, never by routers
- done in an extension header
- fragment extension header carries 13-bit fragment offset, 32-bit packet ID, and 1-bit More Fragments flag

IPv6 socket programming

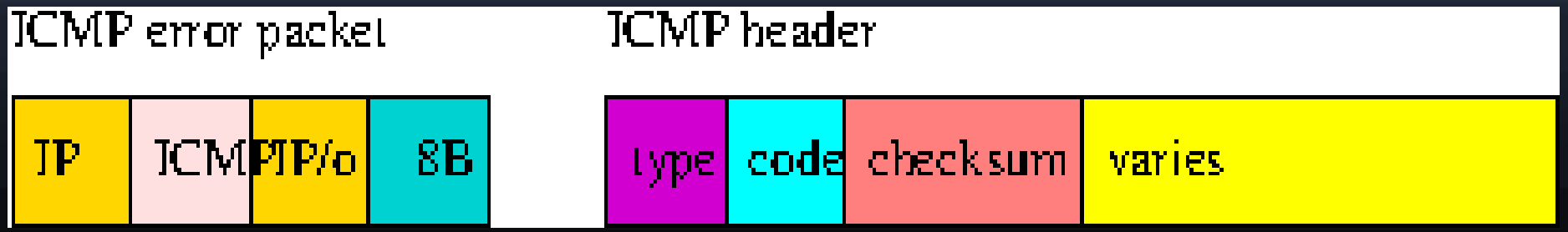
- sockets are generic
  - not much change from IPv4
- `AF_INET` should be `AF_INET6`
- `struct sockaddr_in` should be `struct sockaddr_in6`
  - `sin6_family`, `sin6_port`, `sin6_addr`
  - `sin6_flowinfo` (new in IPv6)
  - `sin6_scope_id` (not used much)

ICMP

- Internet Control Message Protocol, 2 functions:
  - error reporting (never sent in response to ICMP error packets)
  - network debugging (ping)
- common errors packets include
  - destination unreachable (network, host, protocol, or port unreachable)
  - fragmentation needed
  - TTL expired (Time Exceeded)
  - redirect
  - parameter problem

ICMP error message format

- IP header, ICMP header, IP header of error packet, 64 bits (8 bytes) of IP payload of error packet
- ICMP header (8 bytes) has type (e.g. destination unreachable), code (e.g. host unreachable) and 4 or more bytes that vary depending on the type



Ping

- ICMP type ECHO, code 0
- ICMP type ECHO REPLY, code 0
- 4 extra bytes hold two bytes of ID (usually process ID) and two bytes of sequence number
- any additional bytes carry sender's time, in binary (usually 8 bytes)
  - when the packet returns, can compare time received (from system) with time sent (in packet)

Traceroute

- When a router drops a packet, it sends back a Time Exceeded in Transit ICMP error
- so if I send a packet with TTL 1, my router should send me an error
- TTL 2, the next router should send me an error
- …
- the final host should send me port unreachable
  - sometimes blocked by firewalls

Traceroute example, IPv4

```
edo@uhx01 1 % traceroute -n www.ietf.org
traceroute to www.ietf.org (104.20.1.85), 30 hops max, 40 byte packets
 1  128.171.24.193  0.345 ms  0.297 ms  0.186 ms
 2  128.171.1.201  1.478 ms  1.127 ms  1.172 ms
 3  128.171.64.190  1.085 ms  0.945 ms  1.008 ms
 4  205.166.205.48  1.021 ms  0.863 ms  1.083 ms
 5  74.202.119.9  1.541 ms  1.598 ms  14.653 ms
 6  64.129.238.190  70.884 ms  52.823 ms  73.906 ms
 7  * 4.68.71.137  54.391 ms  54.440 ms
 8  4.69.144.138 54.595 ms 4.69.144.202 53.114 ms 4.69.144.74 52.828ms
 9  4.68.70.130  53.027 ms  58.414 ms  55.843 ms
10  62.115.32.214  84.881 ms  81.319 ms  81.322 ms
11  104.20.1.85  62.808 ms  62.954 ms  55.745 ms
```

Traceroute example, IPv6

```
$ traceroute6 -n www.ietf.org

traceroute to www.ietf.org
(2400:cb00:2048:1::6814:55), 30 hops max, 80
byte packets

 1  2001:470:a:446::1  71.683 ms  73.466 ms
75.001 ms

 2  2001:470:0:9b::1  85.305 ms  85.277 ms
85.256 ms

 3  2001:504:16::3417  75.380 ms  75.373 ms
75.324 ms

 4  2400:cb00:28:1024::6ca2:f46a  75.116 ms
2400:cb00:28:1024::6ca2:f415  75.430 ms
2400:cb00:28:1024::6ca2:f46a  75.243 ms
```

Path MTU discovery

- sender sends as big packets as its network MTU can support
  - if TCP MSS option is present in the SYN packets, sender will use the smaller of its own and its peer's MSS
  - TCP also sets the Don't Fragment (DF) bit
    - routers cannot fragment IPv6 packets
- router that cannot forward, sends

  destination unreachable, fragmentation needed
  - payload indicates MTU of interface
- next router may do the same again!

Deprecated ICMP messages

- source quench: send to a sender that is filling up a router's buffers
  - increases traffic during congestion
  - apparently not very effective
- redirect: on this same network, for this destination, use this other router
  - can be used for man-in-the-middle attacks
  - might still work on many systems (there are other ways to do MITM)