Internet Protocols

• Internet Protocols
  — Small set of functions that form basis of all protocols
  — Not all protocols have all functions
    • Reduce duplication of effort
    • May have same type of function in protocols at different levels
  — Encapsulation
  — Fragmentation and reassembly
  — Connection control
  — Ordered delivery
  — Flow control
  — Error control
  — Addressing
  — Multiplexing
  — Transmission services
Encapsulation

- Data usually transferred in blocks
  - Protocol data units (PDUs)
  - Each PDU contains data and control information
  - Some PDUs only control

- Three categories of control
  - Address
    - Of sender and/or receiver
  - Error-detecting code
    - E.g. frame check sequence
  - Protocol control
    - Additional information to implement protocol functions

- Addition of control information to data is encapsulation

- Data accepted or generated by entity and encapsulated into PDU
  - Containing data plus control information
  - E.g. TFTP, HDLC, frame relay, ATM, AAL5, LLC, IEEE 802.3, IEEE 802.11
Fragmentation and Reassembly (Segmentation – OSI)

- Exchange data between two entities
- Characterized as sequence of PDUs of some bounded size
  - Application level message
- Lower-level protocols may need to break data up into smaller blocks. This is called **fragmentation**
- Many reasons for fragmentation
  - Communications network may only accept blocks of up to a certain size
    - ATM 53 octets
    - Ethernet 1526 octets
  - More efficient error control
    - Smaller retransmission
  - Fairer
    - Prevent station monopolizing medium
  - Smaller buffers
  - Provision of checkpoint and restart/recovery operations
Disadvantages of Fragmentation

- Make PDUs as large as possible because
  - PDU contains some control information
  - Smaller block, larger overhead
- PDU arrival generates interrupt
  - Smaller blocks, more interrupts
- More time required to process many smaller PDUs
Reassembly

• Segmented data must be reassembled into messages
• More complex if PDUs have arrived out of order
PDUS and Fragmentation
(Copied from chapter 2 fig 2.4)
Connection Control

- Connectionless data transfer
  - Each PDU treated independently
  - E.g. datagram
- Connection-oriented data transfer
  - E.g. virtual circuit
- Connection-oriented preferred (even required) for lengthy exchange of data
- Or if protocol details must be worked out dynamically
- Logical association, or connection, established between entities
- Three phases occur
  - Connection establishment
  - Data transfer
  - Connection termination
  - May be interrupt and recovery phases to handle errors
Phases of Connection Oriented Transfer

- Connection request
- Connection accept
- Data
- Acknowledgement
- Multiple exchanges
- Terminate-connection request
- Terminate-connection accept
TCP/IP Concepts

Host A

App Y

App X

TCP

IP

Network Access Protocol #1

Physical

Network 1

Router J

IP

NAP 1

NAP 2

Physical

Network 2

Host B

App X

App Y

TCP

IP

Network Access Protocol #2

Physical

Port or service access point (SAP)

Logical connection (TCP connection)

Global network address

Subnetwork attachment point address

Logical connection (e.g., virtual circuit)
Internetworking Terms

• Communications Network
  — Facility that provides data transfer service

• An internet
  — Collection of communications networks interconnected by bridges and/or routers

• The Internet - note upper case I
  — The global collection of thousands of individual machines and networks

• Intranet
  — Corporate internet operating within the organization
  — Uses Internet (TCP/IP and http)technology to deliver documents and resources
Internetworking Terms (2)

- **End System (ES)**
  - Device attached to one of the networks of an internet
  - Supports end-user applications or services

- **Intermediate System (IS)**
  - Device used to connect two networks
  - Permits communication between end systems attached to different networks
Network Architecture Features

- Addressing
- Packet size
- Access mechanism
- Timeouts
- Error recovery
- Status reporting
- Routing
- User access control
- Connection based or connectionless
Architectural Approaches

- Connection oriented
- Connectionless
Connection Oriented

• Assume that each network is connection oriented

• IS connect two or more networks
  — IS appear as ES to each network
  — Logical connection set up between ESs
    • Concatenation of logical connections across networks
    — Individual network virtual circuits joined by IS

• May require enhancement of local network services
  — 802, FDDI are datagram services
Connection Oriented IS Functions

- Relaying
- Routing

- e.g. X.75 used to interconnect X.25 packet switched networks

- Connection oriented not often used
  — (IP dominant)
Connectionless Operation

- Corresponds to datagram mechanism in packet switched network
- Each NPDU treated separately
- Network layer protocol common to all DTEs and routers
  — Known generically as the internet protocol
- Internet Protocol
  — One such internet protocol developed for ARPANET
  — RFC 791 (Get it and study it)
- Lower layer protocol needed to access particular network
Connectionless Internetworking

• Advantages
  — Flexibility
  — Robust
  — No unnecessary overhead

• Unreliable
  — Not guaranteed delivery
  — Not guaranteed order of delivery
    • Packets can take different routes
  — Reliability is responsibility of next layer up (e.g. TCP)
IP Operation

LAN 1

End system (A)

Router (X)

Frame relay WAN

Router (Y)

End system (B)

LAN 2

TCP
IP
LLC
MAC
Physical

TCP
IP
LLC
MAC
Physical

TCP
IP
LLC
MAC
Physical

TCP
IP
LLC
MAC
Physical

$\begin{align*}
& t_1, t_6, t_7, t_{10}, t_{11}, t_{16} \\
& t_2, t_5 \\
& t_3, t_4 \\
& t_8, t_9 \\
& t_{12}, t_{15} \\
& t_{13}, t_{14}
\end{align*}$

$\begin{align*}
& \text{IP-H} \quad \text{TCP-H} \quad \text{Data} \\
& \text{LLC1-H} \quad \text{IP-H} \quad \text{TCP-H} \quad \text{Data} \\
& \text{MAC1-H} \quad \text{LLC1-H} \quad \text{IP-H} \quad \text{TCP-H} \quad \text{Data} \quad \text{MAC1-T} \\
& \text{FR-H} \quad \text{IP-H} \quad \text{TCP-H} \quad \text{Data} \quad \text{FR-T} \\
& \text{LLC2-H} \quad \text{IP-H} \quad \text{TCP-H} \quad \text{Data} \\
& \text{MAC2-H} \quad \text{LLC2-H} \quad \text{IP-H} \quad \text{TCP-H} \quad \text{Data} \quad \text{MAC2-T}
\end{align*}$

TCP-H = TCP header
IP-H = IP header
LLC1-H = LLC header
MAC1-H = MAC header
MAC1-T = MAC trailer
FR-H = Frame relay header
FR-T = Frame relay trailer
Design Issues

- Routing
- Datagram lifetime
- Fragmentation and re-assembly
- Error control
- Flow control
The Internet as a Network

(a) Packet-switching network architecture

(b) Internetwork architecture
Routing

- End systems and routers maintain routing tables
  - Indicate next router to which datagram should be sent
  - Static
    - May contain alternative routes
  - Dynamic
    - Flexible response to congestion and errors

- Source routing
  - Source specifies route as sequential list of routers to be followed
  - Security
  - Priority

- Route recording
Datagram Lifetime

- Datagrams could loop indefinitely
  - Consumes resources
  - Transport protocol may need upper bound on datagram life
- Datagram marked with lifetime
  - Time To Live field in IP
  - Once lifetime expires, datagram discarded (not forwarded)
  - Hop count
    - Decrement time to live on passing through each router
  - Time count
    - Need to know how long since last router
Fragmentation and Re-assembly

- Different packet sizes
- When to re-assemble
  - At destination
    - Results in packets getting smaller as data traverses internet
  - Intermediate re-assembly
    - Need large buffers at routers
    - Buffers may fill with fragments
    - All fragments must go through same router
      - Inhibits dynamic routing
IP Fragmentation (1)

• IP re-assembles at destination only

• Uses fields in header
  — Data Unit Identifier (ID)
    • Identifies end system originated datagram
      – Source and destination address
      – Protocol layer generating data (e.g. TCP)
      – Identification supplied by that layer
  — Data length
    • Length of user data in octets
IP Fragmentation (2)

— Offset
  • Position of fragment of user data in original datagram
  • In multiples of 64 bits (8 octets)

— More flag
  • Indicates that this is not the last fragment
Fragmentation Example

Original datagram
Data length = 404 octets
Segment offset = 0
More = 0

First fragment
Data length = 208 octets
Segment offset = 0
More = 1

Second fragment
Data length = 196 octets
Segment offset = 26 64-bit units (208 octets)
More = 0

Header

Data
Dealing with Failure

- Re-assembly may fail if some fragments get lost
- Need to detect failure
- Re-assembly time out
  - Assigned to first fragment to arrive
  - If timeout expires before all fragments arrive, discard partial data
- Use packet lifetime (time to live in IP)
  - If time to live runs out, kill partial data
Error Control

- Not guaranteed delivery
- Router should attempt to inform source if packet discarded
  — e.g. for time to live expiring
- Source may modify transmission strategy
- May inform high layer protocol
- Datagram identification needed
- (Look up ICMP)
Flow Control

- Allows routers and/or stations to limit rate of incoming data
- Limited in connectionless systems
- Send flow control packets
  - Requesting reduced flow
- e.g. ICMP
Internet Protocol (IP) Version 4

- Part of TCP/IP
  - Used by the Internet
- Specifies interface with higher layer
  - e.g. TCP
- Specifies protocol format and mechanisms
- RFC 791
  - Get it and study it!
  - [www.rfc-editor.org](http://www.rfc-editor.org)
- Will (eventually) be replaced by IPv6 (see later)
IP Services

- Primitives
  - Functions to be performed
  - Form of primitive implementation dependent
    - e.g. subroutine call
  - Send
    - Request transmission of data unit
  - Deliver
    - Notify user of arrival of data unit

- Parameters
  - Used to pass data and control info
Parameters (1)

- Source address
- Destination address
- Protocol
  - Recipient e.g. TCP
- Type of Service
  - Specify treatment of data unit during transmission through networks
- Identification
  - Source, destination address and user protocol
  - Uniquely identifies PDU
  - Needed for re-assembly and error reporting
  - Send only
Parameters (2)

- Don’t fragment indicator
  - Can IP fragment data
  - If not, may not be possible to deliver
  - Send only
- Time to live
  - Send only
- Data length
- Option data
- User data
Options

- Security
- Source routing
- Route recording
- Stream identification
- Timestamping
IPv4 Header

- **Version**: 3 bits
- **IHL**: 4 bits
- **Type of Service**: 8 bits
- **Total Length**: 16 bits
- **Identification**: 16 bits
- **Flags**: 3 bits
- **Fragment Offset**: 13 bits
- **Time to Live**: 8 bits
- **Protocol**: 8 bits
- **Header Checksum**: 16 bits
- **Source Address**: 32 bits
- **Destination Address**: 32 bits
- **Options + Padding**: Variable length

20 octets
Header Fields (1)

• Version
  — Currently 4
  — IP v6 - see later

• Internet header length
  — In 32 bit words
  — Including options

• Type of service

• Total length
  — Of datagram, in octets
Header Fields (2)

- Identification
  - Sequence number
  - Used with addresses and user protocol to identify datagram uniquely
- Flags
  - More bit
  - Don’t fragment
- Fragmentation offset
- Time to live
- Protocol
  - Next higher layer to receive data field at destination
Header Fields (3)

- Header checksum
  - Reverified and recomputed at each router
  - 16 bit ones complement sum of all 16 bit words in header
  - Set to zero during calculation
- Source address
- Destination address
- Options
- Padding
  - To fill to multiple of 32 bits long
Data Field

- Carries user data from next layer up
- Integer multiple of 8 bits long (octet)
- Max length of datagram (header plus data) 65,535 octets
Inter-domain Routing

- Classful network design
- Classless Inter-Domain Routing - CIDR
  - Introduced in 1993 by the Internet Engineering Task Force
  - Goal was to slow the growth of routing tables on routers across the Internet, and to help slow the rapid exhaustion of IPv4 addresses
  - CIDR appends a “/” character to the address and the decimal number of leading bits of the routing prefix
  - Example:
    - 192.168.1.0/24 for IPv4,
    - 2001:db8::/32 for IPv6
IPv4 Address Formats

Class A

0 Network (7 bits) Host (24 bits)

Class B

1 0 Network (14 bits) Host (16 bits)

Class C

1 1 0 Network (21 bits) Host (8 bits)

Class D

1 1 1 0 Multicast

Class E

1 1 1 1 0 Future Use
IP Addresses - Class A

• 32 bit global internet address

• Two parts
  — Network part
  — Host part

• Class A
  — Start with binary 0
  — All 0 reserved
  — 01111111 (127) reserved for loopback
  — Range 1.x.x.x to 126.x.x.x
  — All allocated
IP Addresses - Class B

- Start with binary 10
- Range 128.x.x.x to 191.x.x.x
- Second Octet also included in network address
- $2^{14} = 16,384$ class B addresses
- All allocated
IP Addresses - Class C

- Start with 110
- Range 192.x.x.x to 223.x.x.x
- Second and third octet also part of network address
- \(2^{21} = 2,097,152\) addresses
- Nearly all allocated
  - See IPv6
Subnets and Subnet Masks

- Allow arbitrary complexity of internetworked LANs within organization
- Insulate overall internet from growth of network numbers and routing complexity
- Site looks to rest of internet like single network
- Each LAN assigned subnet number
- Host portion of address partitioned into subnet number and host number
- Local routers route within subnetted network
- Subnet mask indicates which bits are subnet number and which are host number
Routing Using Subnets

LAN X
- Net ID/Subnet ID: 192.228.17.32
- Subnet number: 1

R1
- IP Address: 192.228.17.33
- Host number: 1

LAN Y
- Net ID/Subnet ID: 192.228.17.64
- Subnet number: 2

LAN Z
- Net ID/Subnet ID: 192.228.17.96
- Subnet number: 3

Rest of Internet

A
- IP Address: 192.228.17.57
- Host number: 25

B
- IP Address: 192.228.17.65
- Host number: 1

C
- IP Address: 192.228.17.97
- Host number: 1

D
ICMP

• Internet Control Message Protocol
• RFC 792 (get it and study it)
• Transfer of (control) messages from routers and hosts to hosts
• Feedback about problems
  — e.g. time to live expired
• Encapsulated in IP datagram
  — Not reliable
**ICMP Message Formats**

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Destination Unreachable; Time Exceeded; Source Quench

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Parameter Problem

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(c) Redirect

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(d) Echo, Echo Reply

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(e) Timestamp

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(f) Timestamp Reply

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(g) Address Mask Request

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(h) Address Mask Reply
IP v6 - Version Number

- IP v 1-3 defined and replaced
- IP v4 - current version
- IP v5 - streams protocol
- IP v6 - replacement for IP v4
  —Next Generation
Why Change IP?

• Address space exhaustion
  — Two level addressing (network and host) wastes space
  — Network addresses used even if not connected to Internet
  — Growth of networks and the Internet
  — Extended use of TCP/IP
  — Single address per host

• Requirements for new types of service
IPv6 RFCs

- 1752 - Recommendations for the IP Next Generation Protocol
- 2460 - Overall specification
- 2373 - addressing structure
- others (find them)
- [www.rfc-editor.org](http://www.rfc-editor.org)
IPv6 Enhancements (1)

- Expanded address space
  - 128 bit

- Improved option mechanism
  - Separate optional headers between IPv6 header and transport layer header
  - Most are not examined by intermediate routes
    - Improved speed and simplified router processing
    - Easier to extend options

- Address autoconfiguration
  - Dynamic assignment of addresses
IPv6 Enhancements (2)

- Increased addressing flexibility
  - Anycast - delivered to one of a set of nodes
  - Improved scalability of multicast addresses

- Support for resource allocation
  - Replaces type of service
  - Labeling of packets to particular traffic flow
  - Allows special handling
  - e.g. real time video
IPv6 Structure

- Mandatory IPv6 header
- Hop-by-hop options header
- Routing header
- Fragment header
- Destination options header
- TCP header
- Application data

Octets:
- 40
- Variable
- 8
- Variable
- 20 (optional variable part)
- Variable

IPv6 packet body

Optional extension headers

Mandatory IPv6 header

Next Header field
Extension Headers

- Hop-by-Hop Options
  - Require processing at each router
- Routing
  - Similar to v4 source routing
- Fragment
- Authentication
- Encapsulating security payload
- Destination options
  - For destination node
IP v6 Header

Bit: 0  4  12  16  24  31

Version  Traffic Class  Flow Label

Payload Length  Next Header  Hop Limit

Source Address

Destination Address

10 * 32 bits = 40 octets
IP v6 Header Fields (1)

• Version
  — 6

• Traffic Class
  — Classes or priorities of packet
  — Still under development
  — See RFC 2460

• Flow Label
  — Used by hosts requesting special handling

• Payload length
  — Includes all extension headers plus user data
IP v6 Header Fields (2)

- Next Header
  - Identifies type of header
    - Extension or next layer up
- Source Address
- Destination address
IPv6 Addresses

- 128 bits long
- Assigned to interface
- Single interface may have multiple unicast addresses
- Three types of address
Types of address

- **Unicast**
  - Single interface

- **Anycast**
  - Set of interfaces (typically different nodes)
  - Delivered to any one interface
  - the “nearest”

- **Multicast**
  - Set of interfaces
  - Delivered to all interfaces identified
IPv6 Extension Headers

(a) Hop-by-hop options header; destination options header

(b) Fragment header

(c) Generic routing header

(d) Type 0 routing header
Hop-by-Hop Options

- Next header
- Header extension length
- Options
  - Pad1
    - Insert one byte of padding into Options area of header
  - PadN
    - Insert $N \geq 2$ bytes of padding into Options area of header
    - Ensure header is multiple of 8 bytes
  - Jumbo payload
    - Over $2^{16} = 65,535$ octets
  - Router alert
    - Tells router that contents of packet is of interest to router
    - Provides support for RSPV (chapter 16)
Fragmentation Header

- Fragmentation only allowed at source
- No fragmentation at intermediate routers
- Node must perform path discovery to find smallest MTU of intermediate networks
- Source fragments to match MTU
- Otherwise limit to 1280 octets
Fragmentation Header Fields

- Next Header
- Reserved
- Fragmentation offset
- Reserved
- More flag
- Identification
Routing Header

• List of one or more intermediate nodes to be visited
• Next Header
• Header extension length
• Routing type
• Segments left
  — i.e. number of nodes still to be visited
Destination Options Header

- carries optional info for destination node
- format same as hop-by-hop header
Virtual Private Networks

- set of computers interconnected using an insecure network
  - e.g. linking corporate LANs over Internet
- using encryption & special protocols to provide security
  - to stop eavesdropping & unauthorized users
- proprietary solutions are problematical
- hence development of IPSec standard
IPSec

- RFC 1636 (1994) identified security need
- encryption & authentication to be IPv6
- but designed also for use with current IPv4
- applications needing security include:
  — branch office connectivity
  — remote access over Internet
  — extranet & intranet connectivity for partners
  — electronic commerce security
IPSec Scenario
IPSec Benefits

- provides strong security for external traffic
- resistant to bypass
- below transport layer hence transparent to applications
- can be transparent to end users
- can provide security for individual users if needed
IPSec Functions

- Authentication Header
  - for authentication only
- Encapsulating Security Payload (ESP)
  - for combined authentication/encryption
- a key exchange function
  - manual or automated
- VPNs usually need combined function
- see chapter 21
Summary

- basic protocol functions
- internetworking principles
- connectionless internetworking
- IP
- IPv6
- IPSec