

Lecture 7: Internetworking

CSE 123: Computer Networks
Chris Kanich



Extra project discussion:
Friday 3:30pm WLH 2204



Lecture 8 Overview



- Internet Protocol
 - Service model
 - Packet format
- Fragmentation
- Addressing
 - Subnetting
 - CIDR

Combining Networks



- Main challenge is heterogeneity of link layers:
 - Addressing
 - » Each network media has a different addressing scheme
 - Bandwidth
 - » Modems to terabits
 - Latency
 - » Seconds to nanoseconds
 - Frame size
 - » Dozens to thousands of bytes
 - Loss rates
 - » Differ by many orders of magnitude
 - Service guarantees
 - » Send and pray vs reserved bandwidth

How to connect different networks?



- **Monopoly**
 - Re-engineer network to use a single set of protocols everywhere
 - Economic cost
- **Translation Gateways**
 - Translates directly between different network formats
 - $O(n^2)$ complexity (n is # of protocols)
 - May not be able to translate perfectly (QoS)
- **Indirection Gateways**
 - Translates between local network format and universal "intermediate" format
 - $O(n)$ complexity
 - May not take advantage of features in underlying network
- Note impact of economics on decision. Engineering not science.

SSI, 2006

CSE 123A -- Lecture 7 -- Internetworking and IP

internetworking



- Cerf & Kahn74,
"A Protocol for Packet Network Intercommunication"
 - Foundation for the modern Internet
- **Routers** forward **packets** from source to destination
 - May cross many separate networks along the way
- All packets use a common **Internet Protocol**
 - Any underlying data link protocol
 - Any higher layer transport protocol

CSE 123 -- Lecture 7: Internetworking

5

Some background...



- 1968: DARPA/ARPANet (precursor to Internet)
 - Advanced Research Projects Agency Network
 - Bob Taylor, Larry Roberts create program to build first wide-area packet-switched network
- 1978: new networks emerge
 - SATNet, Packet Radio, Ethernet
 - All "islands" to themselves -- didn't work together
- Big question: how to connect these networks?

Note: If you want to learn more about Internet history, read "Where Wizards Stay Up Late" by Hafner and Lyon

SSI, 2006

CSE 123A -- Lecture 7 -- Internetworking and IP

DARPAnet Primary Goal: Connect Stuff

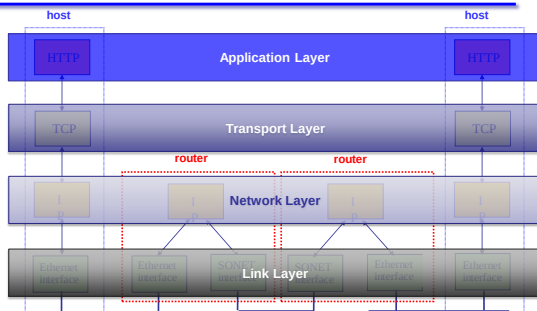


- “Effective technique for multiplexed utilization of existing interconnected networks” – David Clark
 - **Minimal** assumptions about underlying networks
 - » No support for broadcast, multicast, real-time, reliability
 - » Extra support could actually get in the way
 - Packet switched, store and forward
 - » Matched application needs, nets already packet switched
 - » Enables **efficient resource sharing**/high utilization
 - “Gateways” interconnect networks
 - » Routers in today’s nomenclature

SSI, 2006

CSE 123A -- Lecture 7 -- Internetworking and IP

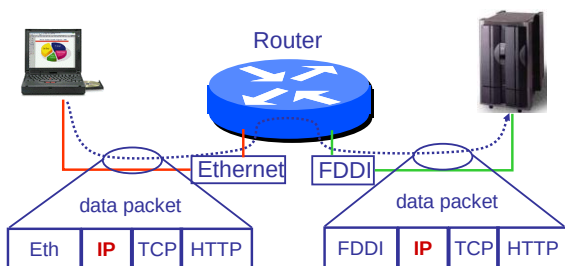
TCP/IP Protocol Stack



CSE 123 -- Lecture 7: Internetworking

8

IP Networking



CSE 123 -- Lecture 7: Internetworking

9

Routers



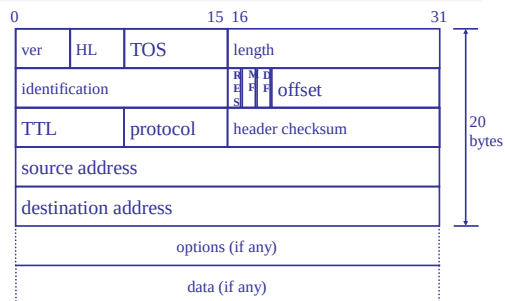
- A router is a store-and-forward device
 - Routers are connected to multiple networks
 - On each network, looks just like another host
 - A lot like a switch, except at the network layer
- Must be explicitly addressed by incoming frames
 - Not at all like a switch, which is transparent
 - Removes link-layer header, parses IP header
- Looks up next hop, forwards on appropriate network
 - Each router need only get one step closer to destination

IP Philosophy



- Impose few demands on network
 - Make few assumptions about what network can do
 - No QoS, no reliability, no ordering, no large packets
 - No persistent state about communications; no connections
- Manage heterogeneity at hosts (not in network)
 - Adapt to underlying network heterogeneity
 - Re-order packets, detect errors, retransmit lost messages...
 - Persistent network state only kept in hosts (fate-sharing)
- Service model: **best effort**, a.k.a. *send and pray*

IP Packet Header



Version field



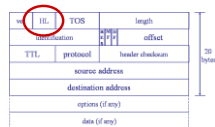
- Which version of IP is this?
 - ♦ Plan for change
 - ♦ Very important!
- Current versions
 - ♦ 4: most of Internet today
 - ♦ 6: new protocol with larger addresses
 - ♦ What happened to 5? Standards body politics.



Header length



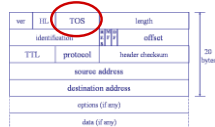
- How big is IP header?
 - ♦ In bytes/octets
 - ♦ Variable length
 - » Options
 - ♦ Engineering consequences of variable length...
- Most IP packets are 20 bytes long



Type-of-Service



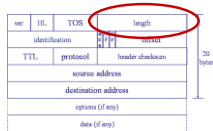
- How should this packet be treated?
 - ♦ Care/don't care for delay, throughput, reliability, cost
 - ♦ How to interpret, how to apply on underlying net?
 - ♦ Largely unused until 2000 (hijacked for new purposes, ECN & Diffserv)



Length



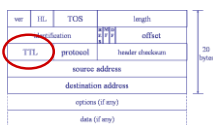
- How long is whole packet in bytes/octets?
 - ◆ Includes header
 - ◆ Limits total packet to 64K
 - ◆ Redundant?



TTL (Time-to-Live)



- How many more routers can this packet pass through?
 - ◆ Designed to limit packet from looping forever
- Each router decrements TTL field
- If TTL is 0 then router discards packet



Protocol



- Which transport protocol is the data using?
 - ◆ i.e. how should a host interpret the data
- TCP = 6
- UDP = 17



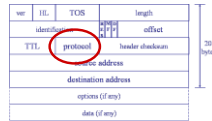
IP Checksum



- Header contains simple checksum
 - Validates content of header *only*

- Recalculated at each hop
 - Routers need to update TTL
 - Hence straightforward to modify

- Ensures *correct* destination receives packet

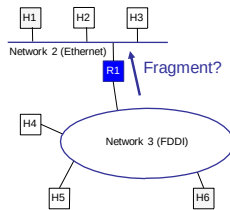


Fragmentation



- Different networks may have different frame limits (MTUs)
 - Ethernet 1.5K, FDDI 4.5K

- Router breaks up single IP packet into two or more smaller IP packets
 - Each **fragment** is labeled so it can be correctly **reassembled**
 - End host* reassembles them into original packet



IP ID and Bitflags



- Source inserts unique value in identification field
 - Also known as the IPID
 - Value is copied into any fragments

- Offset field indicates position of current fragment (in bytes)
 - Zero for non-fragmented packet

- Bitflags provide additional information
 - More Fragments bit helps identify last fragment
 - Don't Fragment bit prohibits (further) fragmentation
 - Note recursive fragmentation easily supported—just requires care with More Fragments bit



Fragmentation Example



length	ID	MF	offset
=4000	=x	=0	=0

One large datagram becomes
several smaller datagrams

length	ID	MF	offset
=1500	=x	=1	=0

length	ID	MF	offset
=1500	=x	=1	=1480

length	ID	MF	offset
=1040	=x	=0	=2960

Costs of Fragmentation



- Interplay between fragmentation and retransmission
 - A single lost fragment may trigger retransmission
 - Any retransmission will be of entire packet (why?)
- Packet must be completely reassembled before it can be consumed on the receiving host
 - Takes up buffer space in the mean time
 - When can it be garbage collected?
- Why not reassemble at each router?

Path MTU Discovery



- Path MTU is the smallest MTU along path
 - Packets less than this size don't get fragmented
- Fragmentation is a burden for routers
 - We already avoid reassembling at routers
 - Avoid fragmentation too by having hosts learn path MTUs
- Hosts send packets, routers return error if too large
 - Hosts can set "don't fragment" flag
 - Hosts discover limits, can size packets at source
 - Reassembly at destination as before
