Introduction to Routing in Internet

Internet basics

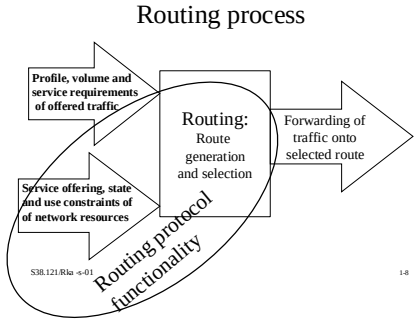
IPv4 and ICMP

Internet Addressing

ARP - Address Resolution Protocol

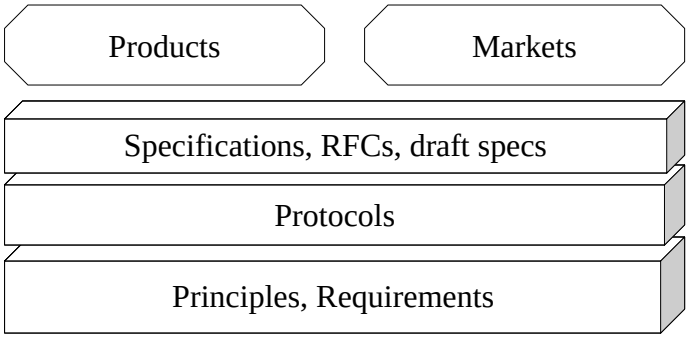
Routing Information (Distance Vector) Protocol Principles

Internet routing is based on routing protocols that collect the input data



- No off-line route planning, off-line only dimensioning.
- Routing is fully automated.
- Routing is divided to interior and exterior.
  - This course will concentrate only on Interior routing.
  - S38.191 will talk about exterior routing

Levels of analysis - we deal with principles, protocols and specifications



Internet Architecture Principles

- Does it exist?
- End-to-end principle *by Dave Clark*
  - all e.g. error and flow control in end stations
  - trusting networks is redundant
  - more reliable transport works for IP
  - no state information/connection in the network
  - same as distributed systems

Principles - IP over everything

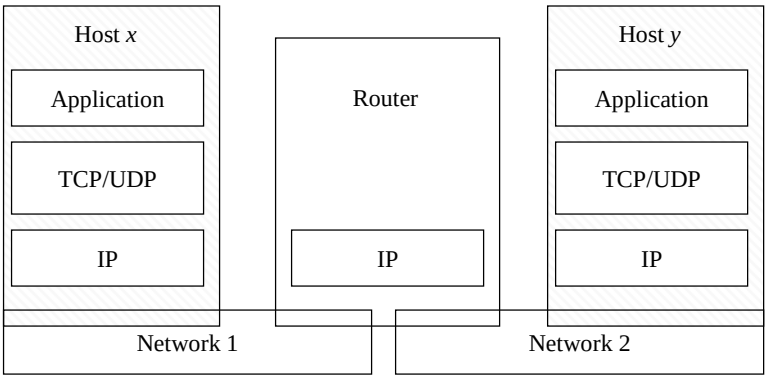
by Vinston Cerf

- Interconnection based on IP overlay over all kinds of networks
  - framing or encapsulation
  - address mapping
    for each transport technology
- interconnection based on translation:
  - e.g. signaling inter-working - imperfect mapping
  - IPv4 to IPv6 mapping!

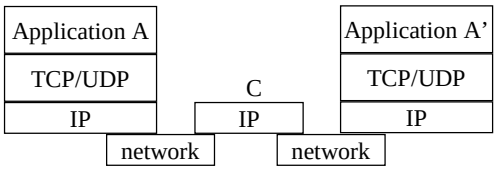
Principles - Connectivity is its own reward

- Be liberal with what you receive, conservative with what you send
  - try to make your best to understand what you receive
  - maximum adherence to standard when sending
- snowballing effect keeps all interested in connectivity thus keeps adhering to standards

Internet layer model - hosts and routers



Message forwarding in Internet layers



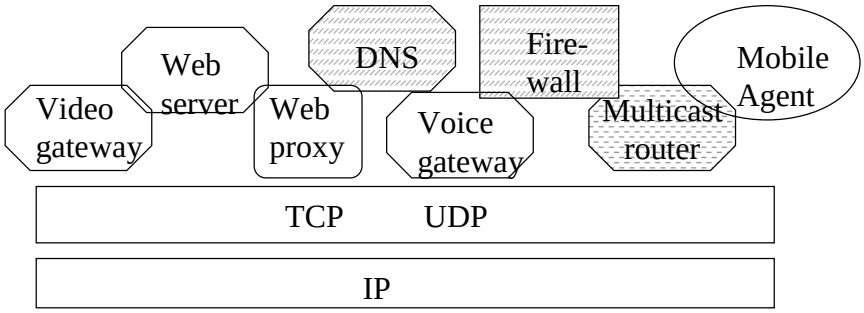
Encapsulation:

a1 → c1, IP	A → B, TCP	TCP header + data
Ethernet header	IP header	

Encapsulation:

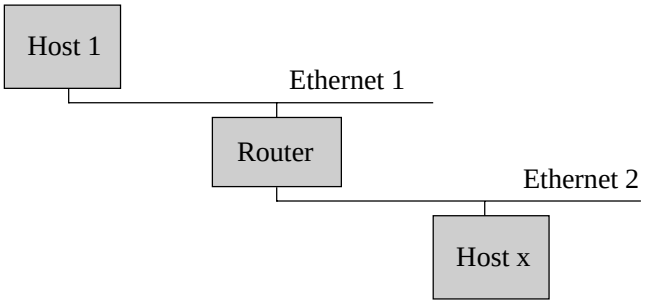
c2 → b1, IP	A → B, TCP	TCP header + data
Ethernet header	IP header	

Internet architecture includes a set of Service level components on top of TCP/IP

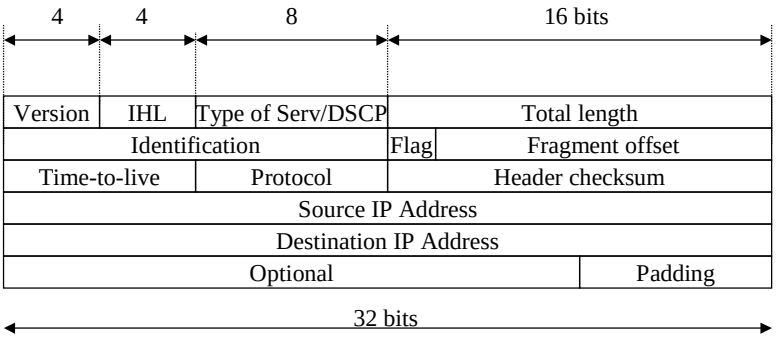


In this course we may touch some of these but only in their relation to routing.

By connecting Ethernet segments with a router, the traffics can be kept separate



IPv4 packet header



We assume that the sender knows its own IP address, if not self configuration protocols such as *RARP, BOOTP, DHCP - dynamic host conf. protocol* are used

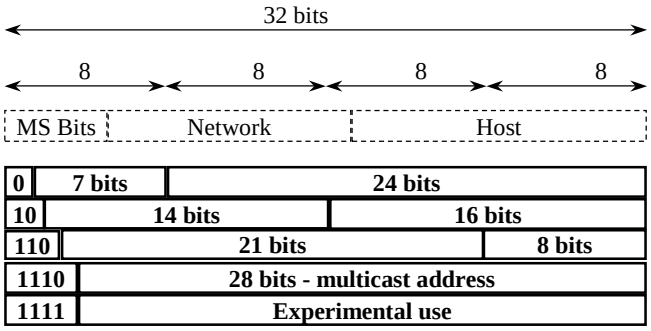
DSCP - DiffServ Code Point, IHL - IP header length

IP version	IP version number. Current version is 4
IHL	Internet header length. Expressed as number of 32 –bit words
Type of Service/ DSCP	TOS contains 3MSBits: packet priority + service type. DSCP – is proposed use for Differentiated Services
Total length of the packet	Expressed as nrof octets in the payload and in the header
Identifikation, Flags and Offset	Are used when large packets are fragmented when underlaying network has maximum packet length.
TTL	Time-to-live. The value is decremented with an integer representing the quality of the network on each router a path of the packet. Packet is deleted when TTL reaches

Protocol	Protocol, that the receiving host should use to process the datapacket, e.g. TCP
Checksum	Header checksum. Calculated as 16 bit one's complement sum
Source Address	IP address of the sender of the packet.
Destination Address	IP address of the destination host
Options	Used for special types of information or "tricks". One packet can carry many option fields

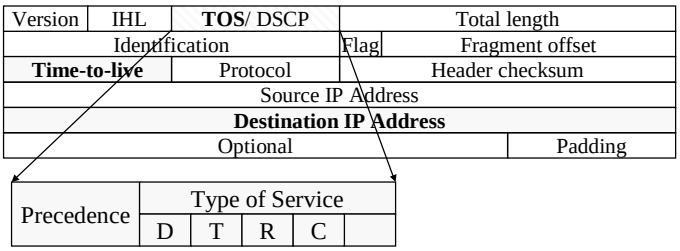
IPv4 address formats

This is the model prior to CIDR



IP -address points to an interface (rather than the host).
(In addition Mask is used to separate the sub-network and the host.)

Destination Address and the TTL are used for Routing



TOS = ***route selection criteria: D - minimization of delay or***
T - maximization of bandwidth or
R - maximization of reliability or
C - minimization of cost

This Schema was never widely adopted!

priority - highest value --> must be served first in the queue.

Options: for example: source routing. Used very seldom because routers tend to serve packets with options last.

Internet address structure

network number	subnet	host

Mask	Address	Net	Subnet	Host
0xFFFF0000	10.27.32.100	A: 10	27	32.100
0xFFFFFE00	136.27.33.100	B: 136.27	16 (32)	1.100
	136.27.34.141	136.27	17(34)	0.141
0xFFFFFFC0	193.27.32.197	C: 193.27.32	3(192)	5

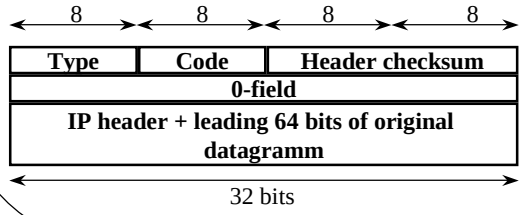
High order bits:
0 ..... 0 - 127. --> A-class
10.... 128. - 191. --> B-class
110...192. - 223. --> C-class

without right zeroes (and with right zeroes)

ICMP - Internet Control Message Protocol gives feedback to the sender about the network state

- All hosts and routers must support ICMP. (To battle Denial of Service Attacks not always a good idea).
- ICMP packet is sent backwards if e.g.
  - the receiver is unreachable
  - router deletes a packet
  - TTL = 0
- If ICMP message is deleted, a new one is not generated to avoid the snowballing effect.

ICMP messages



Type=0 - Echo reply
3 - destination unreachable
4 - (source quench)
5 - Redirect
8 - Echo
9 - Router advertisement
10 - Router solicitation
11 - Time exceeded
12 - Parameter problem
13 - Timestamp
14 - Timestamp reply
15 - Information request
16 - Information reply

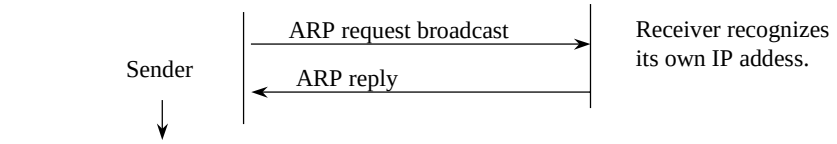
(4 - source quench="slow down" has been dropped from recommendations)

Code =
0 - net unreachable
1 - host unreachable
2 - protocol unreachable
3 - port unreachable
4 - fragmentation needed and DF set
5 - source route failed

ARP - Address resolution protocol (RFC-826) maps IP to the underlaying protocol.

Sender works like this:

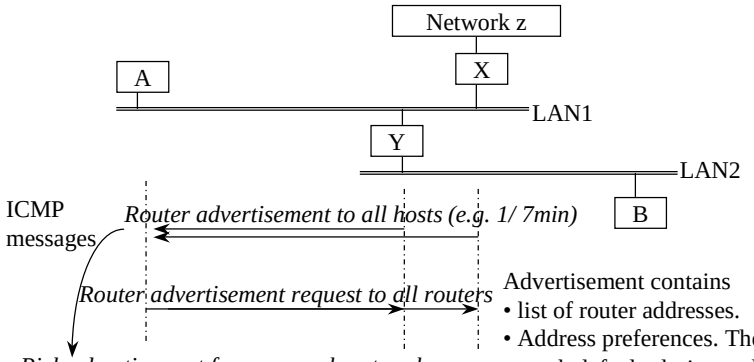
1. Compare masked values of own and destination IP addresses to find out whether the destination is in the same sub-network. If =, destination is in the same sub-network, if not the packet must be sent to a router.
2. Find media address (MAC address) of the next hop.



Store in local address cache

Each network technology requires its own ARP adaptation. Adaptation is easy if the network supports broadcast or multicast.
E.g. ATM requires a special ARP - server.

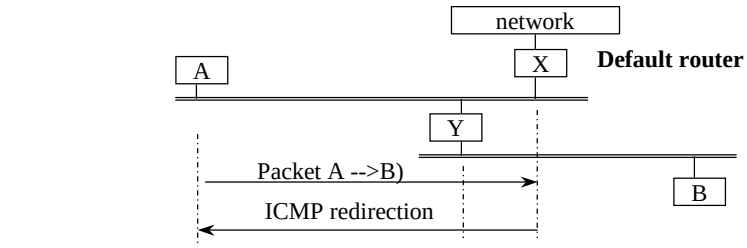
A network may have many routers, closest to destination must be found

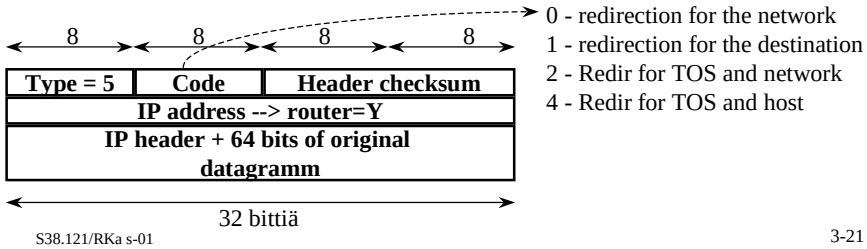


- *Pick advertisement from own sub-network*
- *Default router preference is highest.*

Advertisement contains
- list of router addresses.
- Address preferences. These mark default, designated,etc. router or router address
- validity period (e.g. 30 min)

Router can send redirection packet to hint to a better route towards a destination





0 - redirection for the network
1 - redirection for the destination
2 - Redir for TOS and network
4 - Redir for TOS and host

Redirect is a slow mechanism. Hot-standby addressing is an improvement

- Virtual router redundancy protocol (RFC 2338 - 4/98)
  - a router may have a virtual IP address
  - a router can take the IP and MAC addresses of a failed router (in the same segment)
  - After recovery routers negotiate about address assignments
  - Clients are configured with a static (virtual) router address
  - Cisco and DEC have equivalent proprietary protocols
- Host can listen to RIP or OSPF
  - not recommended but used sometimes anyway

Host must have feedback from the first router to avoid sending to a “black hole”

Feedback may be

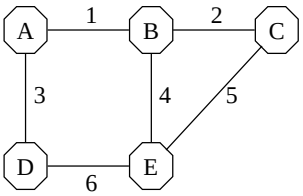
- TCP acknowledgements
- Router advertisements
- ARP-replies
- ICMP echo reply

Between routers, routing protocols provide similar feedback and help in detecting failed router neighbors.

Routing can be static or dynamic

- Static routing is based on manually configured routing tables.
  - Static routing is used when e.g. two peer providers do not trust each other or
  - To connect an organization to a Service Provider with a single connection
  - Static routing is difficult to maintain
- Dynamic routing is based on routing protocols which create and maintain the routing tables automatically
  - examples of routing protocols are RIP, OSPF, BGP...
  - E.g. to connect an organization with multiple links to the Internet

RIP - Routing Information Protocol is a basic protocol for interior routing



Example network with nodes A, B, C …
Link numbers: 1, 2, 3 …
In a real network no global link numbering exists.

RIP is a distance vector protocol.

Let us study DV protocol principles:

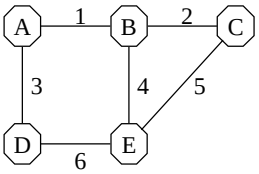
Initial state: Nodes know their own addresses and interfaces, no more.

Node A creates its routing table:

From A to Node	Link	Cost
A	local	0

The corresponding DV is: A=0.

Generation of routing tables starts when all routers send their DVs on all interfaces



A=0 →

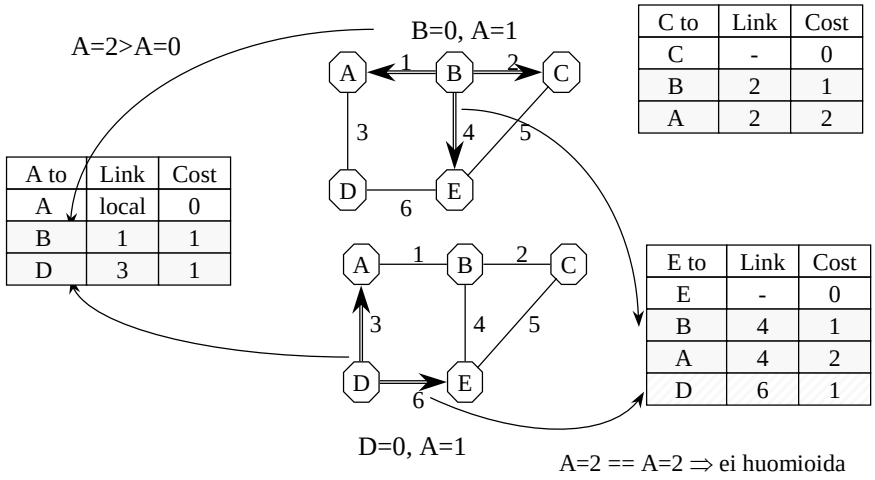
Let's look at reception in Node B

From B to Node	Link	Cost
B	local	0

1. B increments the DV + 1 ⇒ A=1 and
2. B looks for the result in its routing table, no match
3. B adds the result to its RT, result is

	From B to Node	Link	Cost
	B	local	0
New entry →	A	1	1

B creates its own DV and sends it to all neighbors



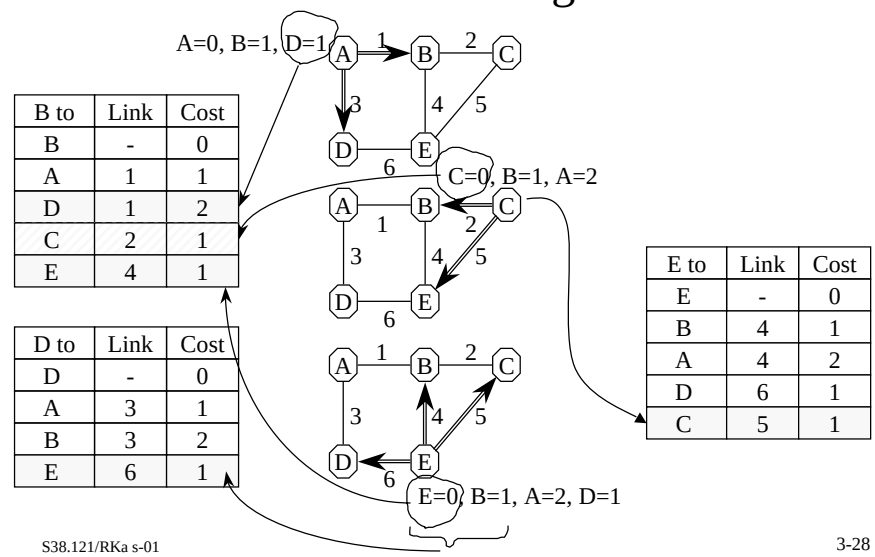
A to	Link	Cost
A	local	0
B	1	1
D	3	1

C to	Link	Cost
C	-	0
B	2	1
A	2	2

E to	Link	Cost
E	-	0
B	4	1
A	4	2
D	6	1

A=2 == A=2 ⇒ ei huomioida

Nodes with changed RTs create DVs and send them to neighbors

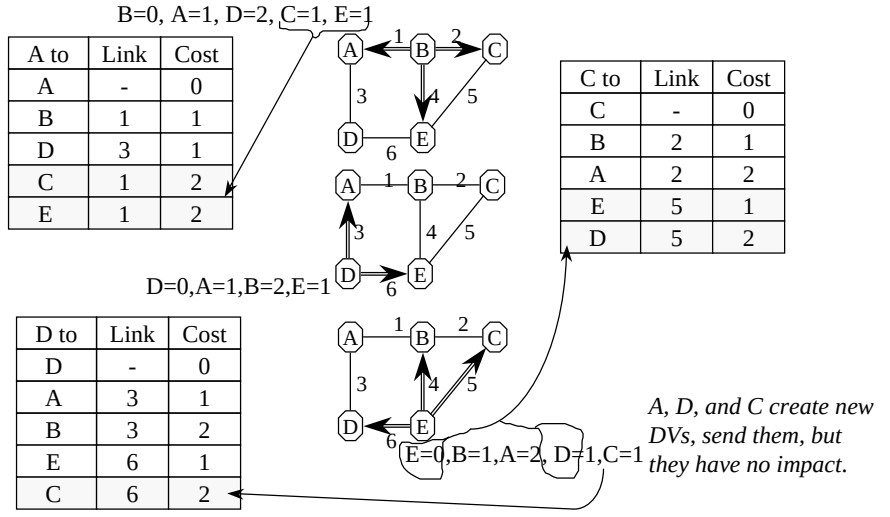


B to	Link	Cost
B	-	0
A	1	1
D	1	2
C	2	1
E	4	1

D to	Link	Cost
D	-	0
A	3	1
B	3	2
E	6	1

E to	Link	Cost
E	-	0
B	4	1
A	4	2
D	6	1
C	5	1

Again the changes are sent ...



A to	Link	Cost
A	-	0
B	1	1
D	3	1
C	1	2
E	1	2

C to	Link	Cost
C	-	0
B	2	1
A	2	2
E	5	1
D	5	2

D to	Link	Cost
D	-	0
A	3	1
B	3	2
E	6	1
C	6	2