Week3 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22

RUFENG DING

Contents

1 IP backgrounds

A internetwork is :

- A collection of individual networks.
- Connected by intermediate networking devices.
- That functions as a single large network.

1.1 Internet Protocol suite

ISO/OSI and Internet Protocol suites

Table 1: Internet Protocol suites

7	Aplication	FTP/SMTP/Telnet/HTTP	Host-to-Host Protocols
6			
5			
4	Transport	TCP UDP	
3	Internet	IP (ICMP/ARP/RARP...)	
2	Data link	802.x/PPP/FR/ATM	Network specified ”outside” Internet specification
1	Physical	LANs/PSTN/LL/ADSL	

1.2 Two providers

A service describtion of Internet
The Internet allows Distributed Applications running on it end systems to exchange data with each other. It provides two service two its distributed applications:**a connection-oriented reliable service** and**a connectionless unreliable service**. It does not yet provide a service that makes promises about how long will it take to deliver the data from sender to receiver.

1.3 router

It is the main component of the Internet. It is responsibile to, step by step, forward packets to the destination. It can be simple and cheap but also very expensive.

Path determination:
route taken by packets from source to destination. Routing algorithms.

Forwarding:
move packets from router's input to appropriate router output.

Inside the router:
Two key router functions:

- run routing algorithms/protocol(RIP/OSPF/BGP).
- switching datagrams from incoming to outgoing link.

2 Addressing

2.1 IP internet Protocol RFC791

- Connectionless service
- Network addressing
- Best effort delivery
  - IP datagrams may arrive at destination host damaged duplicated of order, or not at all
  - no end-to-end delivery guarantees
- Handles data forwarding using routing tables prepared by other protocols such as:
  - Open shortest path first (OSPF)
  - Routing information protocol (RIP)
- Fragmentation and reassembly

2.2 IPv4 Datagram Structure

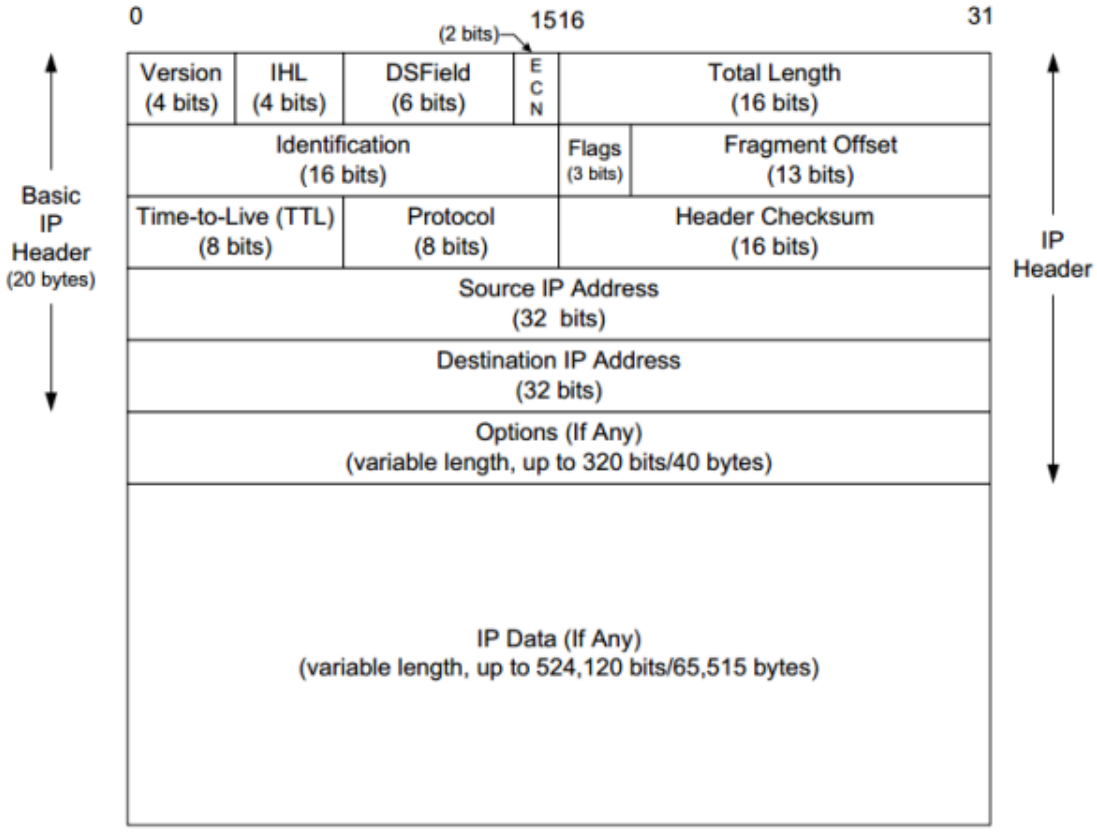

Figure 1: IPv4 Datagram Structure

2.3 IPv4 Addressing

Two main functions:

- uniquely identify a computer in a given internet.
- provide information so that routers deliver the packet to the correct destination.

They have 32 bits (4Bytes) and are representd by dot notation. eg.: *138.77.45.3*

10001010 01001101 00101101 00000011

addressdes will have two parts: The left part will identify a particular network/LAN. The right part will identify the host inside that network.

Important : The amount of bits for each part will vary.

2.4 Special IPv4 addresses

- All 0 host suffix - netwrok address 128.10.0.0 is a network with possible hosts like : 128.10.3.4, 128.10.0.3
- All 0s network - this network eg. 0.0.0.2 host 2 on this network
- all 1s host suffix - broadcast to all host on the same subnet
- public IP address are controlled by the internet
- private IP addresses RFC1918
  - any organization can use these inside their network. However thses addresses can't go on the internet
  - 10.0.0.0 - 10.255.255.255
  - 172.16.0.0 - 172.31.255.255
  - 192.168.0.0 - 192.168.255.255
- Loopback address : 127.0.0.1 All computers "have" this IP address

2.5 Classful addressing

host ID of all 0's indicate Network ID

host ID of all 1's indicate broadcast to Network ID

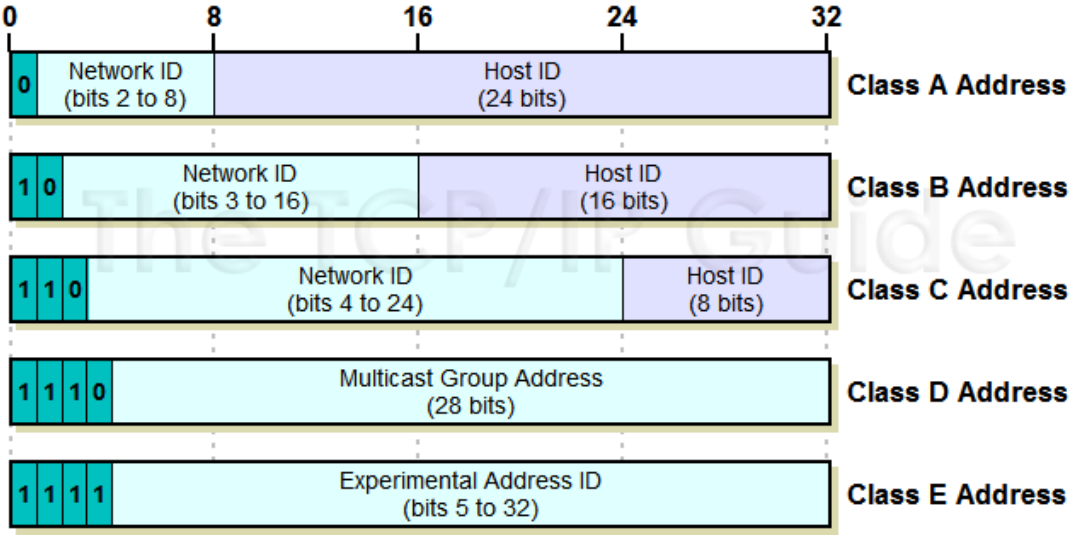

Figure 2: addresses classes

- 0.x.x.x -127.x.x.x : 128 big networks 16 million hosts.

- 128.x.x.x - 191.x.x.x : 16384 Medium Networks 65 thousand hosts.
- 192.x.x.x - 223.x.x.x : 2 million small networks 254 hosts.
- 224.x.x.x -239.x.x.x used for multicast group.
- 240.x.x.x - 255.x.x.x reserved for future use.

In this scheme each address is said to be self-identifying because the boundry between prefix and suffux can be computed from the address alone.

2.6 subnetting

sometiems organization may wish to partition their network, This can be done with subnetting.
Classless and subnet address extensions (CIDR)
class scheme is too rigid and many address can be wasted, subnettting permits to split class A,B,or C in smaller network. Host ID field is split in subnetwork field and host field. A subnet mask (Net ID + subnet field) identify all hosts that belong to a specific subnetwork address space.

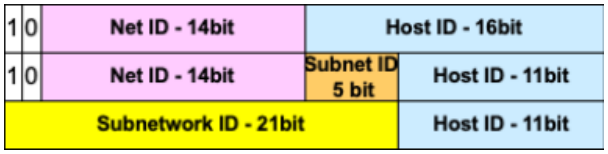

Figure 3: eg CIDR

In this example subnet mask is 255.255.248.0 or FF.FF.F8.00 Hex. This network permits to allocate up to 2046 Hosts.

IP address netmasks
Bit mask for the network part of the address: eg. 255.0.0.0/8. 255.255.240.0/20, etc.

eg: 128.16.20.1/16

/16 - 255.255.0.0

128.16.20.1 1000 0000 0001 0000 0001 0100 0000 0001

255.255.0.0 1111 1111 1111 1111 0000 0000 0000 0000

128.16.0.0 1000 0000 0001 0000 0000 0000 0000 0000

2.7 Problems with IPv4

- Sortage of IP addresses The 32bits address system in IPv4 can theoretically recongnize 4.3 billion hosts. This is not enough for widespread adoption of IP in multiple devices

- Insufficient security functions Scalability requires that robust security measures on the IP datagram are availble. In IPv4 security is typically a funciton of the upper layers
- Fragmentation introduces complexity
- No Quality of Service support
- Complex header

2.8 IPv6 Datagram Structure

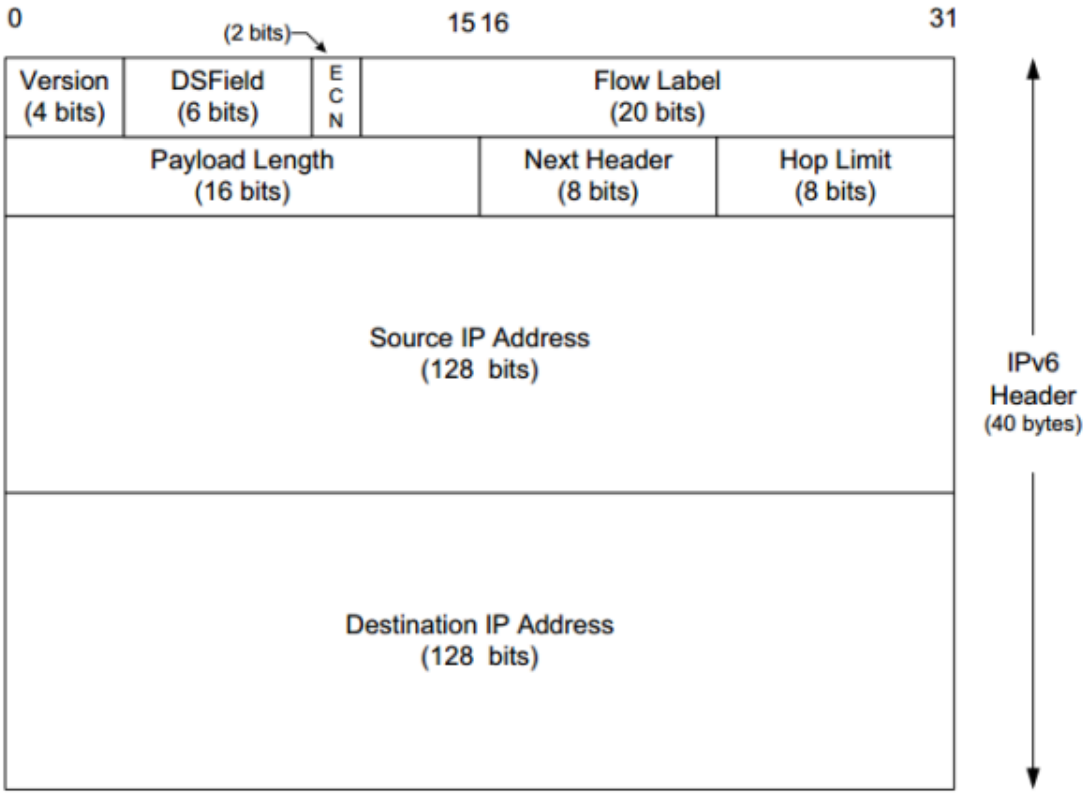

Figure 4: IPv4 Datagram Structure

2.9 IPv6 provides

- Expansion of IP address : An astronomical number is catered for.
- Hierarchical IP addresses
- IPsec function is installed as standard
- QOS control function
- No Fragmentation

- Automatical location of IP addresses
- Simplified header
- Allows Jumbograms (very big packets)

2.10 IPv6 address

2.10.1 notation

IPv6 uses 8 groups of 4 hexadecimal digits: 2001:0630:0013:0200:0000:0000:0000:ace0. This can be "tided up" by removing leading 0's and eliding runs of "0"s: 2001:630:13:200::ace0 . The boundary between network and host part is indicated using /. eg. 2001:630:13:200::ace0/64 indicates a network address of 2001:630:13:200::and a host address of ::ace0.

2.10.2 address blocks

IPv6 addresses are allocated to interfaces. An interface can and will have many addresses. IPv6 address space has been split into blocks. RFC4291 describles IPv6 addressing.

There are :

special purpose blocks:

- 0000::/8 is reserved by IANA and includes
- The unspecified address (all 0's) and the loopback address (::1) are assigned from this block
- IPv6 addresses mapped from IPv4 (E.g. ::128.40.42.82 is a valid IPv6 address,but cannot be globally routed)

multicast: Equivalent to addresses from the IPv4 block: 224.0.0.0/3 and IPv6 multicast range is FF00::/8

Note: There is no concept of broadcast in IPv6 : multicast must be used instead.

link-local unicast: Allocated from the block FE80::/10 and 64bit Interface ID appended to make an address. Interface ID constructed from interface MAC address and this address cannot be globally routed.

local IPv6: Designed to replace RFC1918 private IP addresses.

Block prefix: FC00::/7 8th bit indicates global or local management. Only a value of 1 (local management) has currently been standardised: Prefix becomes FDOO::/8 inpractice.

The remaining 56bits of the network address consist of a random 40bit ID and 16bit subnet number. The ID is randomly generated such that there is a good chance it is globally unique. Allows two organisations to merge Locally addressed networks without renumbering.

Global unicast: Equivalent to a standard IPv4 address such as 128.40.42.82 and allocated from the block: 2000::/3

Special case unicast address: *Anycast*: An anycast address may appear on several interfaces on different hosts,but the network layer only delivers packets to one of them.

A note on IPv6 addresses with embedded IPv4 addresses: Two forms of these were specified:

- IPv4 compatible lPv6 addresses: ::0000:128.40.42.82
  - Designed to allow automatic tunnelling of IPv6 over IPv4
  - This range has been deprecated and will no longer be used
- IPv4 mapped IPv6 addresses: ::FFFF:128.40.42.82
  - Designed to allow IPv6 hosts to exchange packets with IPv4 hosts

2.10.3 address allocation policy

The IPv6address assignment policies were designed to result in efficient routing tables. Assignments are hierarchical with the Regional Internet Registries getting large /12 allocations. The assignment policy recommended in RFC3177 is to allocate /48 prefixes to organisations and private individuals. Very large organisations may be assigned a /47 or a set of /48 prefixes. IPv6 was designed to a;low for 245 networks (/48 prefix, 35x1012). Compare IPv4 with 2.2x106 networks.

2.10.4 address in practice

recommended is to use 64/64 scheme:

- 64 bits for network part
- 64 bits for host part

Host part derived using EUI64. uses the 48 bit layer 2 MAC address (padded to 64bits). Means a subnet can have 224 hosts = 16M. Practivally, subnets should never run out of addresses.

3 Multicasting

3.1 Multicast in IPv4

Multicast is the transmission to a given set of members of a group.This group may contain members anywhere in the Internet. This process is quite complex.

Addresses of tthese groups must be in the range of 224.X.X.X to 239.X.X.X

Nodes need to subscribe to a multicast group (using multicast addresses). More on this later.

3.2 Multicast addresses in IPv6

Multicast is an integral part of IPv6 and is used for:

- Router discovery
- Address resolution
- Well-known service discovery

A multicast address is indicated by FF in first byte and the next byte is formed of 4 flag bits and 4 scope bits. The other 112 bits are the multicast group ID.

In practice,to make mapping to ethernet adresses simple,the group ID is usually 32bits.

3.3 Special Multicast addresses in IPv6

There are many permanently assigned multicast addresses:

- FF02::1 = All nodes on a link (LAN)
- FF02::2 = All routers on a link
- FF05::2 = All routers in a site
- FF05::3 = AII DHCP servers in a site

FF02::1:FFXX:XXXX is the solicited node multicast address. The last 24bits are copied from the last 24bits of the node's unicast address. This is likely to be site unique.

3.4 Mapping multicast addresses to Ethernet

The Multicast IP address has to be mapped to an Ethernet MAC address so that hosts can receive the datagrams.

In IPv4 the MAC prefix of 01:00:5e is used along with the last 24bits of the multicast IP address. (eg. 224.15.31.23 maps to 01:00:5e:0f:1f:17) ln IPv6 the MAC prefix of 33:33 is used along with the last 32bits of the IP address.

4 Resolution and Autoconfiguration

4.1 ARP address resolution protocol

Host broadcasts a request to ask for the MAC address of IP address. And the host whose IP is the one reply back the MAC address.

Sending a package:

Is the package for our network?

- Y (We need the Ethernet address) - Is the Ethernet address for this IP address in the cache?
  - Y Prepare an Ethernet Frame with the address - send the frame
  - N Send an ARP broadcast asking for the Ethernet address - Get the response update the Cache - go to Y
- N Looking in our routing table to check what is the router to send the packet - now the packet ”is” for our network. We are going to send the packet to the router.(Then go to Y)

4.2 ARP notes

ARP Cache timeout: typically 20 minutes.

Proxy ARP: sometimes router will reply ”instead of the host” to ”trick” the host into sending them the packets.

Gratuitous ARP: a ”reply” sent by a host without a reuest. Typically sent by a host waking up.

Remember: ARP is not just for Ehternet. It works for any layer 2 technology.

4.3 Replacement of ARP in IPv6

In IPv4 on a LAN, ARP is used to discover the link layer address so datagrams can be exchanged.

IPv6 uses a different method. **Remember there is no broadcast in IPv6**. A Neighbour Solicitation message is sent to the **neighbour solicitation multicast address**.

The node in the multicast group (remember, there should only be one) replies with a Neighbour Advertisement which contains the link-layer address.

The Address to link-layer mapping is stored in a cache in the host. Hosts can send periodic unsolicited neighbour advertisement message to the all nodes multicast address (FF02::1).

5 obtain own IP address: DHCP/SLAAC

5.1 DHCP

Sometimes we want an IP address allocated dynamically. For this we usually use DHCP.

- It allow dynamic configurations: automatically assigned address leasing.
- Uses LAN broadcast
- Require servers: central store of configuration information
- Useful for
  - mobile hosts
  - large numbers of hosts (can use static/manual address assignment)
- usually also returns:
  - default router
  - netmasks
  - DNS servers
- usually sent using UDP port 64 (destination 255.255.255.255 source 0.0.0.0)

5.2 SLAAC

IPv6 Host Autoconfiguration SLAAC = StateLess Address Autoconfiguration

How a host get its addresses:
as a host boots, it will bind eadch network interface to the following addresses:

- FF02::1 all nodes multicast address
- ::1 Loopback address

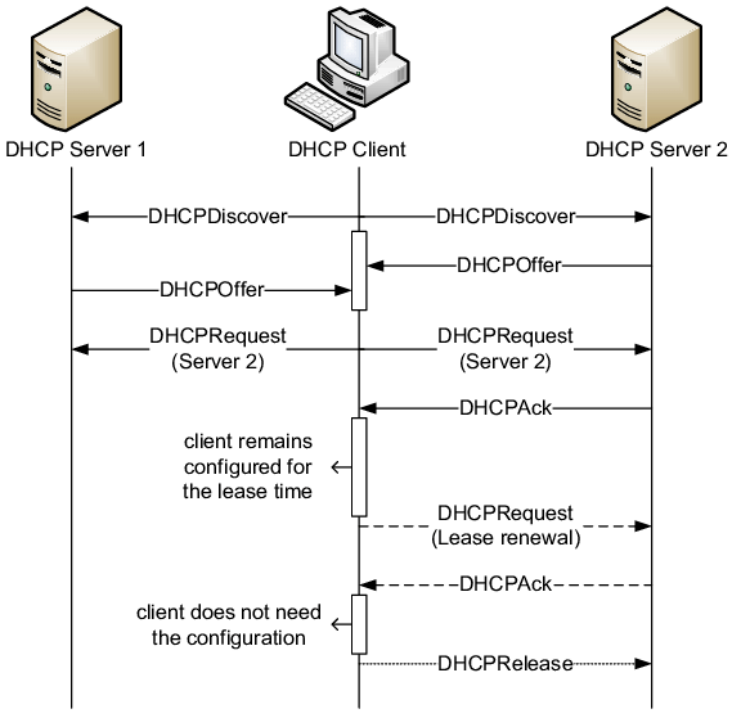

Figure 5: DHCP process

- FF02::1:ffxx:xxxx node solicitation multicast address(xx:xxxx = last 24 bits of MAC address(usually))
- FE90::interface ID (This is marked as a "Tentative address" and cannot be used as the source address for any datagrams yet.)

5.3 Host autoconfiguration

At this point, a host has a usable IPv6 address. However, this address is link-local.(FE80 prefix).Which means it can not be used to talk to discover any router on the network.

The next stage is to discover any router on the network. A router solicitation message is snet to the all-routers multicast address FF02::2. All routers on the link reply with a router advertisement. Each advertisment that has the autoconfiguration flag set will cause the host to construct an address from the advertised router prefix and the host's interface ID. If there is more than one router(and network prefix), an address for each network prefix will be assigned.

5.4 DAD

DAD = duplicate address detection: use a multicast mechanism

The host constructes a ICMP Neighbour Discovery packet which will be sent to the node solicitation

multicast address corresponding to the last 24 bits of the interface ID.

This will discover if there is another node using the same address if there is, autoconfiguration will stop and the node will have to be assigned an address by another means. The other node will use a Neighbour Advertisement message to signal its presence.

Otherwise, the IPv6 address is unique and there usable.

5.5 Privacy Extension RFC4941

computer picks a series of bits randomly, and fills in the last 48 bits with the random bits. Reduces privacy concerns. A new IP address can be generated with varying frequency. In theory even one per connection. And it is widely available in operation systems.

6 fragmentation

6.1 Fragmentation in IPv4

in the beginning the Internet slogan was "IP over everything".

Because IP has to use several underlying technologies, IP packets may have to be fragmented. Fragmentation is done by routers. Defragmentation is done by end systems.

Related parts: Total length / Identificaiton / Flags /Fragment Offset.

6.2 Fragmentation in IPv6

IPv6 router do not do fragmentation. Instead senders implement a process called **Path MTU discovery**

Senders send a first packet. If a link in the middle cannot cope with that size, it drop the packet and sends back a ICMP message saying "packet too big". The process continues until the packet reached the destination.

7 ICMP

Internet Control Message Protocol (ICMP) ICMP is used by IP to send error and control messages.

It is sent by end hosts or routers.

Table 2: ICMP in IPv4

IPv4 Header	ICMP header	ICMP data

Table 3: ICMP in IPv6

IPv6 Header	IPv6 Extension Headers(if present)	ICMP header	ICMP data

7.1 ICMP messages

ICMP messages are carried in IP packets. But conceptually we see ICMP at the same level of IP.

7.2 ICMPv6

in IPv6 ICMP isresponsible not only for error and informational messages but also for IPv6 router and host configuration. Most messages achieve similar goals as IPv4.

Some messages:

0-127 errors.

128-255 informational.

7.3 ICMP tools

7.3.1 PING

used to see whether a specified IP address is reachable. Tool is available in Microsoft Windows operationg system and UNIX paltforms.

7.3.2 TraceRoute

Sned a packet with time to live = 1

the first router discards the packet and send a ICMP 'time to live exceeded message'

send a packet with time to live = 2

The second router discards the packet and send an ICMP 'time to live exceeded message'. This is repeated until a response is received from the destination.

Table 4: ICMP

Type 8 bits	Code 8 bits	Checksum 16 bits

Table 5: ICMP messages

type	ode	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	6	dest. network unknown
3	7	dest. host unknown
4	0	source quench(congestion control not used)
8	0	echo request(ping)
9	0	route advertisement
10	0	routerr discovery
11	0	TTL expired
12	0	bad IP header

8 NAT

once a network is created, LANs can be point-to-point and routers have more than one LAN.

Network address tranlation NAT

sometimes the organization/ISP/home has many more computers than IP addresses. NAT is used to covert private IP address to public IP addresses and vice versa. Normally present as a router function. Used to connect an intranet using private IP addresses to the Internet. It may enhance security.

eg:

- inner host 10.0.0.1 sends datagram to 128.119.40.80
- NAT router changes datagram source address from 10.0.0.1 to 138.76.29.7,5001 updates table (NAT translation table)
- reply arrives destination address: 138.76.29.7,5001
- NAT router changes datagram destination address form 138.76.29.7 to 10.0.0.1,3345

8.1 NAT for UDP and ICMP

in UDP there is no connection termination. NAT routers need to use timers to check if they can reuse a port.

And it is similar for ICMP.

8.2 NAT traversal

originating connections TO a node behind a NAT is hard. There is no port to put in the TCP SYN packet.

There are many techniques: STUN, TURN, uPnP. Generally involve registering with a node with a public IP address which will then relay the connections.

8.3 Carrier Grade NAT

NAT can be done inside the network. Many networks behind the NAT box. Sometimes thousands of users.

Problems: port exhaustion. Some applicaitons use lots of TCP connections.

9 Mobility

9.1 Mobile IP

Nodes move to a different network. But they want to still receive the packets desined to their 'old' address. Sending applications do not need to be aware of the movement. Just one of the techniques for IP Mobility...

3 Step:

1. Discovery
2. Registration
3. tunnelling

9.2 Three Musketeers of Mobile IP

How does the Mobile Node find out where it is?

- Agent Discovery - ICMP Router Discovery

How does the Mobile Node Inform its current location?

- Registration - Authentication, location update and deregistration

How are packets delivered?

- Tunnelling - IP in IP or GRE

9.3 Mobile IP terminology

1. Agent Discovery To discovery a foreign agent a Mobile IP node uses ICMP Router Discovry (RFC 1256)

   Router periodically broadcasts or multicasts ICMP RD messages on all its links. Mobile IP advertisements contain defined extensions to ICMP RD.

2. Agent Advertisement routers propagate advertise packet in an ICMP packet.

3. registering

   - The mobile sends a registration request to the foreign agent
   - The foreign agent relays this request to the home agent
   - The foreign agent relays this reply to the mobile node

4. Tunnelling the home agent arrived at a foreign network and it discovered a router willing to act as a foreign agent and registered with its home agent.

9.4 IP over IP

senders sends the packet to the destination unaware of any tunnelling/mobility : original packets - encapsulated packet

9.5 The Triangle problem

Route not optimal. Server and Mobile node could communicate directly but...

Authenticaiton would have to be made with potentially every node in the internet.

9.6 Route Optimization

In IPv6 things are different when both nodes have IPv6 enabled.

The mobile node can send a IPv6 Destination options header message to the corresponding node. Packets can then be sent directly.