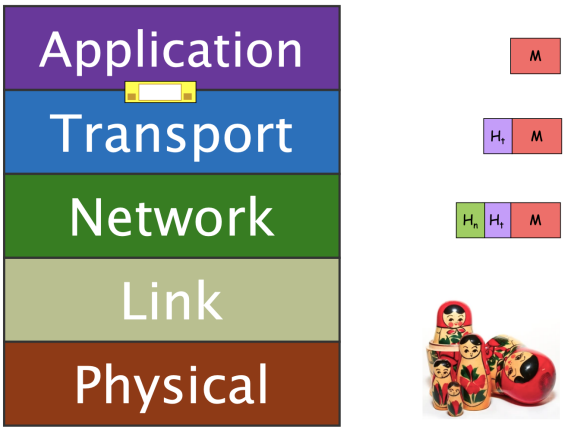
CS2105 Live Class Lec7: Network Layer

$_{/}|/|U_{C}h@NgRu!$

May 5, 2021

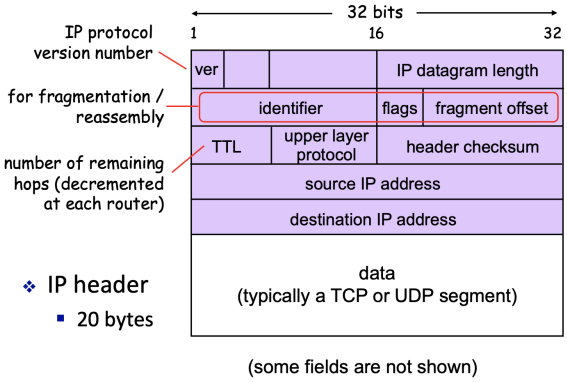
1 The header of network



As the data goes down, more layers are added(like add envelope);

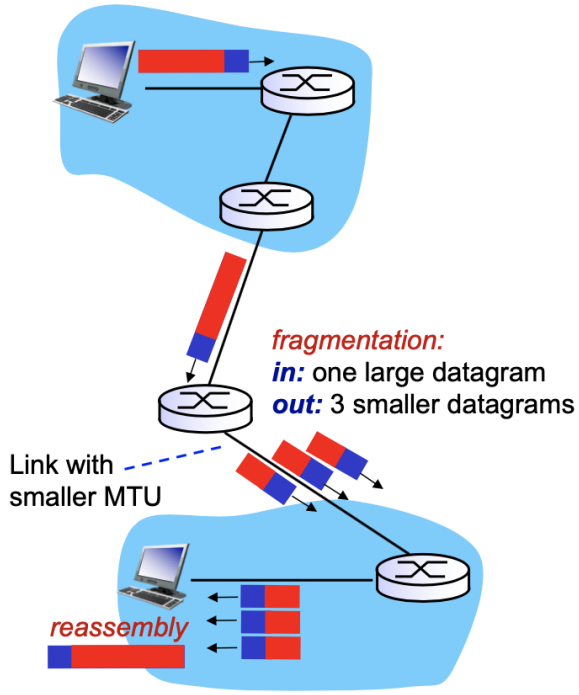
1. In transport layer, TCP/UDP header is added
2. In network layer, new hearder of network layer is addded

2 IP v4 Datagram format



32 bits(4 bytes) long and 5 rows and so 20 bytes in total
The key fields in the IP v4 datagram are the following

1. Version Number: These 4 bits specify the IP protocol
2. IP Datagram length: Total length of IP datagram(**header plus data**), measured in **bytes**
3. Identifier, flag, fragmentation offset are used in IP fragmentation
4. Time-to-live(TLL) refers to number of remaining hops. decremented at each router, if the TTL is 0, no longer forward; The routing algorith make sure there is only one route from the source to destination, so the TTL is to make sure the data is not transported in a inifinity loop(SSPP)
5. Header checksum is used to detect error for the IP header,Protects the **IP header only**
6. source and destination IP address refers to the initial sender and final receiver's IP address
7. upper layer protocol: indicate what upper layer is on above: usually TCP or UDP



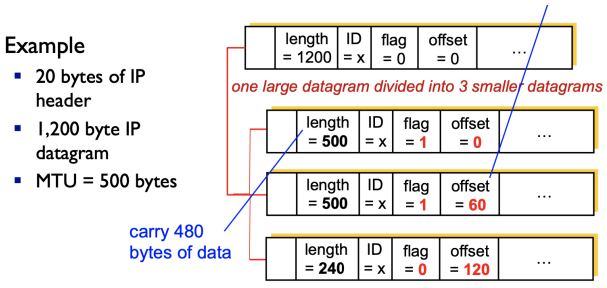
A typical IP v4 header is 20 bytes long

3 IP Fragmentation and reassembly

IP fragmentation is used because different links may have different maximum transfer unit(MTU), specified by underlying link layer, and IP datagrams could be so large that need to be fragmented to fit in to MTU. When destination hosts reassembles the packet, IP header fields are used to identify fragments and their relative order.

Destination host will reassemble the packet

IP header fields are used to identify fragments and their relative order



3.1 Length

IMPORTANT: the "length in the header" contains the header length. For example, the "length of header" of the original packet is 1200, which consist of 20 bytes of the header length and 1180 of data. The original packet was divided into three packet data, first two carries 480 bytes of data each, and the last one carries 220 bytes of data.

3.2 offset

Offset is in expressed in units of 8-bytes(The (MTU-20) is usually divisible by 8
Because each fragmentation

3.2.1 The different between the datagram here and the packet of UDP

It's totally different though the UDP does not add much feature above

3.3 Flag of fragmentation

flag is set to :

1. 1 if there is next fragment from the same segment
2. 0 if this is the last fragment

3.3.1 Potential problem

There is possible that frame with flag 0 comes before the frame with flag 1
Offset and length should be checked so that even if the packet datagram arrives out of order, it still works

3.4 Offset

offset is expressed in unit of 8-bytes, indicating the first byte of the data of this datagram in relative to the original fragmented datagram
For example, the first smaller fragment has 0 bytes of frag ahead, hence the offset is 0/8=0
The second smaller fragment has 480 bytes of frag ahead, hence the offset is 480/8=60
The third smaller fragment has 960 bytes of frag ahead, hence the offset is 960/8=120

3.5 ID

Same datagram will use the same ID

3.6 Length

Length field will change accordingly in smaller packets

4 Intra-AS Routing and Inter-AS Routing

The Internet is a network-of-networks, organised in a hierarchy of autonomous system(AS). Due to the size of the Internet and the decentralised administration of the Internet, routing on the Internet is done hierarchically

* **Intra**-AS routing
    - Finds a good paths between two router within an AS
    - Commonly used protocols: RIP,OSPF
* **Inter**-AS routing
    - Handles the interface between ASes
    - The standard protocol: BGP

The aim of routing may differ between intra-AS routing and inter-AS routing

* **Intra**-AS routing
    - Single administrator, so no policy decisions in needed
    - Routing mostly focus on performance
* **Inter**-AS routing
    - Administrator often wants to control over how its traffic is routed, who routes through its net, etc
    - Policy may dominate over performance

We can abstractly view a network of routers as a graph, where vertices are routers and edges are physical links between routers
Costs can be associated to each link, represented as the weight of respective edge
In this sense, routing is the problem of finding a least cost path between two vertices in a graph.

4.1 Link State Algorithm

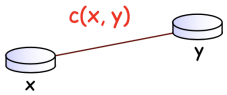
For link state algorithm, all routers have the complete knowledge of network topology and link cost. Routers will periodically broadcast link costs to each other.
The least cost path can be found using Dijkstra algorithm

4.2 Distance Vector Algorithm

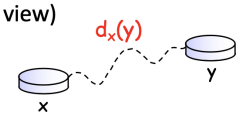
Routers know physically-connected neighbours and link costs to neighbors. Router exchange local views with neighbours and update own local view, based on neighbours' view. It admits an interactive process of computation converging slowly:

- Swap local view with dirct neighbours
- Updata own's local view
- Repeat 1-2 till no more change to local view

- c(x, y): the cost of link between routers x and y
  - = ∞ if x and y are not direct neighbours



- $d_x(y)$: the cost of the least-cost path from x to y (from x's view)



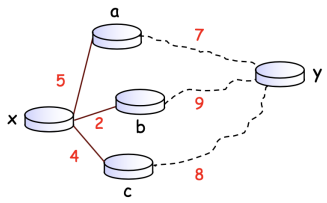
Let c(x,y) be thew cost of link between router x and y. c(x,y) $=\infty$ if x and y are not direct neighbours.

Let $d_x(y)$ be the cost of the least-cost path from x to y, from x point of view. This d is called the distance vector.

Distance vector algorithm then utilises **Bellman-Ford Equation**:

$$d_x(y) = \underbrace{min}_{v\ neighbor\ of\ x} \{c(x,v)+d_v(y)\}$$

The first term is easily obtained since it is mostly manually configured. The distance vector in second term is obtained from exchange of local views between neighbours



$$d_x(y) = min_v\{c(x,a)+d_a(y),\ c(x,b)+d_b(y),\ c(x,c)+d_c(y)\}$$
$$= min\{12, 11, 12\} = 11$$

4.2.1 Bellman-Ford behavior

Router will maintain a NxN table where N is number of routers in the network, storing $d_{from}(to)$

- In the beginning, router x has only c(x,v) knowm. c(x,v) is ∞ if v is not a neighbour. Therefore, $d_x(v)$ is initialised as c(x,v) and others to infinity
- Router will exchange d(v) with each otehr. After each exchange, x will update $d_x(v)$ according to Bellman-Ford Equation
- In addition, x will note down the next hop router to every destination, based on the least cost path. This information will be used to create forwarding table of x
- Eventually this table till converge to provide optimal routing service

5 Routing information protocol(RIP)

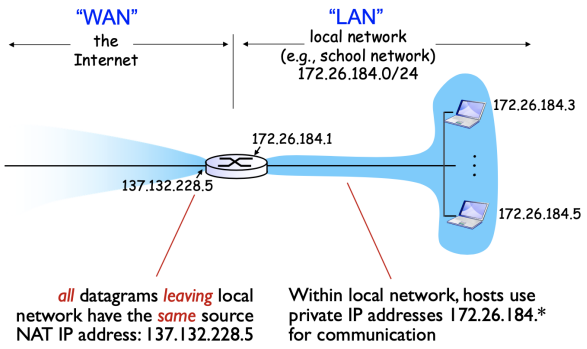
Routing information protocol implements the Distance Vector Algorithm. IT uses hop counts as the cost metric
Entries in the routing table every 30 seconds over UDP port 520
Self-repair mechanism if there is no update from a neighbour router for 3 minutes, the neighbour is assumed to be failed

6 Network Address Translation(NAT)

Network Address Translation is used by routers connecting to both local network and the internet, for translating IP addresses with the aid of port number.
It is necessary since private IP addresses cannot be used for routing on the Internet(the IP v4 address is used up)
Usually switch goes up to layer two, router goes to layer three. The routers in home usually contains switch, router and firewall; The router has both DHCP client and DHCP server inside

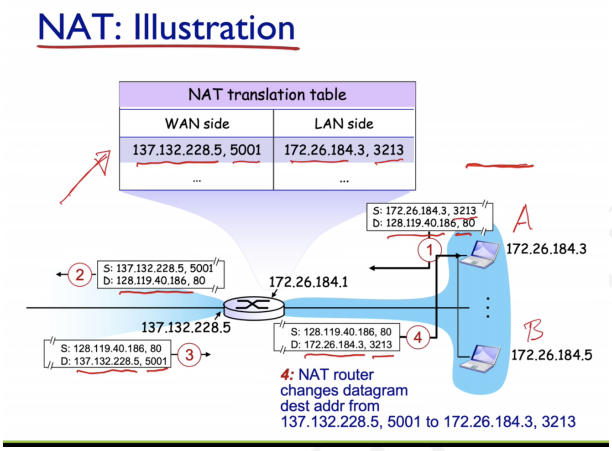
- client: It will get a dynamic assigned address from ISP
- server: allocate private address for the devices from home

IMPORTANT: The private IP address is not routable in the broader internet, because many computers can have the save IP address



For the broader internet, all devices from the save subnet is with the same IP address

6.1 Implementation of NAT



NAT enabled routers must:

1. Maintain a NAT translation table, in which mapping between(source IP address, **new port#**) of the private network and (NAT IP address, **port #**) used on the Internet are stored.

   IMPORTANT: The NAT should give each host a independent port number, otherwise, hosts cannot be differentiated later when some public server response

2. For outgoing datagrams, replace(source IP address,port#) to (NAT IP address,**new port #**)

3. For incoming datagram, replace(NAT IP address, new port #) in destination fields with coresponding (source IP address,port #) by NAT translation table

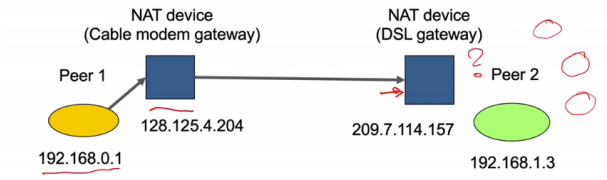
IMPORTANT: the port number in NAT should be differentiated from the port number in transportation layer. The port number in the transportation layer refers to process while the port number in network layer refers to host

6.2 The benefits of NAT include:

1. There is no need to rent a range of public IP addresses form ISP. just one public IP for the NAT router
2. All hosts use private IP addresses. Addresses of hosts in local network without notifying the outside world
3. ISP can be changed without changing addresses of hosts in local network
4. Hosts inside local network are not explicitly addressable and visible by outside world

6.3 Challenges of NAT

Peer-to-peer applications does not work



1. There is one application hosted in a machine with 192.168.1.3 , which is a private IP address of 209.7.114.157
2. If there is one client want to communicate with the application, it will connect with 209.7.114.157 with the help of 128.125.4.204
3. However, there is many machines allocated with 192.168.1.3's private IP and 209.7.114.157 does not know which to connect

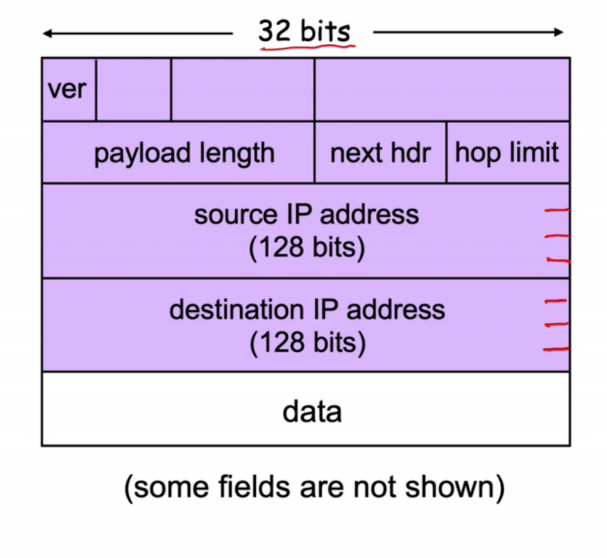
There is protocol specific for peer-to-peer

7 IPv6

Non-examinable in CS2105

1. IPv6 is designed to replace IPv4
2. Primary motivation: 32-bit IPv4 address space is soon to be completely allocated

3. IPv6 datagram: 40 byte header



It would be much harder to remember:
2001:0db8:85a3:0042:1000:8a2e:0370:7334

7.0.1 Interesting features

1. One interesting aspect of IP address is geolocation(it can be teld where is certain IP geographically located)
2. For some applications, it is interesting to know where a host with a specific IP address is located
3. In IPv4 IP address, there is no "hint" of where an address is located
4. There exists databases that map IP addresses to locations
5. IPv6 addresses, could in theory provide more accurate geolocations

8 Internet Control Message Protocol (ICMP)

- ICMP header: Type + Code + Checksum + others.

Type	Code	Description
8	0	echo request (ping)
0	0	echo reply (ping)
3	1	dest host unreachable
3	3	dest port unreachable
11	0	TTL expired
12	0	bad IP header

Selected ICMP Type and subtype (Code)

Internet Control Message Protocol(ICMP) is used by hosts and routers to communicate network-level information.

1. Error reporting: unreachable host/ network / port / protocol
2. Echo request / reply (used by ping)

ICMP messages are carried in IP data gram.ICMP header starts after IP header.

8.1 Ping and traceroute

Command ping sees if a remote host will respond to us- check for connection.
Command traceroute sends a series of small packets across a network and attempts to display the route that the messages would take to get to a remote host

9 Network Security

Network security concerns:

1. Confidentiality
2. Integrity
3. Availability
4. Authenticity

9.1 Subsection of Cryptography

Cryptography is aimed to make it difficult for an unauthorised third party to understand private communications(more for confidentiality)

Two most important part of cryptography is algorithm an the key
There are two main types:

1. Symmetric key
2. Public key session

9.2 Message Integrity and Digital Signatures

Two ways to ensure message integrity and message authenticity are message authentication code(MAC) and digital signature
The basis of both methods is cryptographic hash function

9.3 Cryptographic

Cryptographic hash function is a hash function H which takes an random-length input m, and generate a fixed size string H(M), known as message digest(hash or finger print)
Cryptographic hash function must adhere the property that it is computationally infeasible to find two different messages m and m' such that H(m)=H(m')
The following property makes it impossible for Trudy to forge anotehr message m' with the same message digest
Popular cryptographic hash functions: MD-5 and SHA-1(both not secure) The description of MAC and signature are waived here. Details can be found on the CS2107

9.4 SSL:Secure Socket Layer

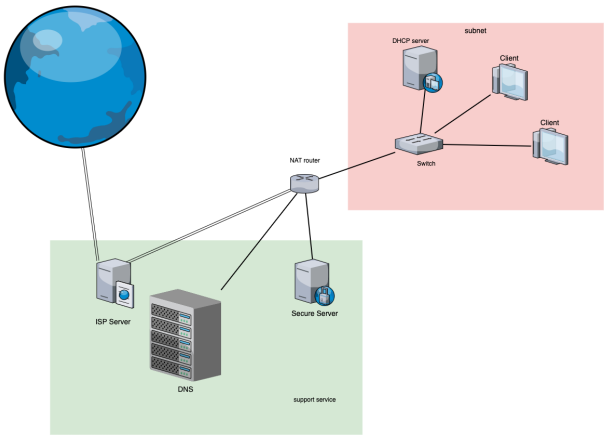
Secure socket layer is an layer interfaced between application and transport layer. It is applicable to TCP connections

9.5 IPSec: Internet Protocol Security

IPSec is a suite of protocols that secure communications by authenticating and encrypting each IP packet of a communication session. It is interfaced between transport layer and IP layer.
Both SSL and IPsec can be used to build VPN

10 Routing conclusion

11 get IP



1. DHCP will give every host in the subnet a dynamic IP(private) and subnet mask

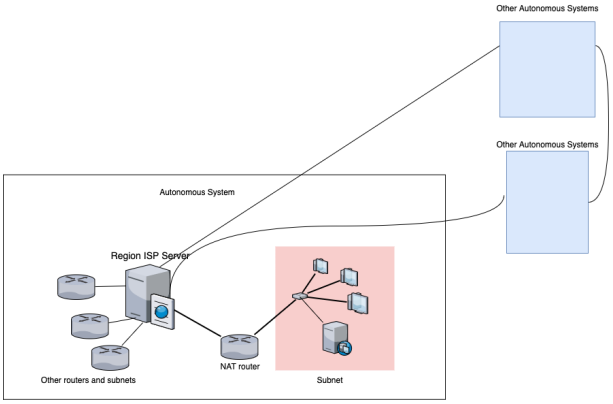
2. Every host also get
   - DNS server IP
   - Default gate way(which is usually the NAT router); Anything that does not belong to subnet will go through default gateway

The Network address translation(NAT) is a method of mapping an IP address space into another by modifying network information in IP headers of pkts while they are transit across a traffic routing device, it may transfer IP:

1. private to public(multiplex)
2. public to private(de-multiplex)

IMPORTANT: switch or hub are both transparent to hosts(the switch and hubs don't have IP address)

11.1 Review of network layer routing



The routing within the autonomous system is based on RIP(Bellman ford and hop based) or OSPF etc, which is performance-oriented

The routing between autonomous system is based on BGP(border gateway protocol; there are other protocol also)