CS2105 Live Class Lec6: Network Layer

$_/|/|U_Ch@NgRu!$

May 5, 2021

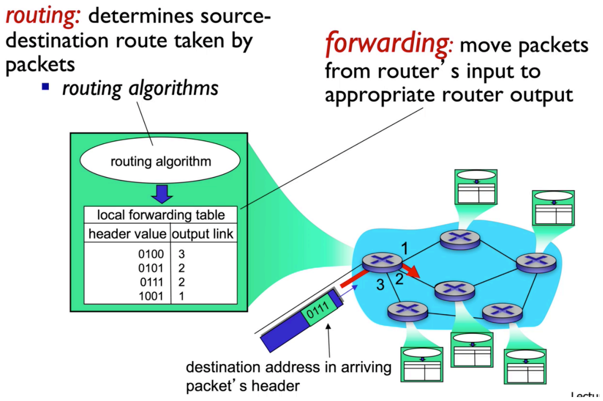
Network layer is responsible for delivering packets to receiving hosts. Routers will examine header fields of IP datagrams passing it and direct it to the right destination

1 Network layer

transport segment from sending to receiving host on sending side: encapsulates segments into datagrams
on receiving side, delivers segments to transport layer
network layer protocols in every host, router
router examines header fields in all IP datagrams passing through it]

2 Two key network-layer functions

1. forwarding: move packets from routers input to appropriate router output
2. routing: determine route taken by packets from source to destination


routing: determines source-destination route taken by packets
▪ routing algorithms
forwarding: move packets from router's input to appropriate router output
routing algorithm
local forwarding table
header value | output link
0100 | 3
0101 | 2
0111 | 2
1001 | 1
1
2
3
0111
destination address in arriving packet's header

The routing algorithm decides which router the file should be outputed each time

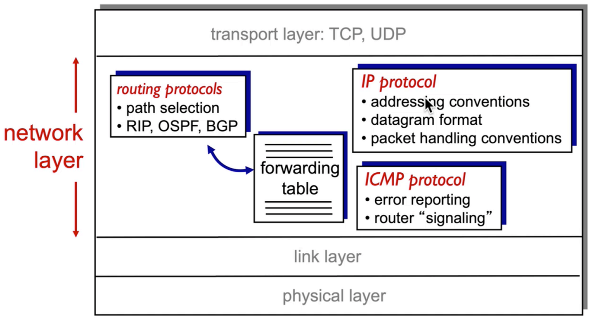
2.1 The network layer can be divided into two layers

1. Data plane
   (a) local, per-routwer functgion
   (b) determine how datagram arriving on router input port is forwarded to router output port
   (c) forwarding function
2. Control plane
   (a) network-wide logic
   (b) determines how datagram is routed among router along end-end path from source host to destination host
   (c) two control-plane approaches:
      i. traditional routing algorithm: implemented in routers, interact with other router and run algprithm in router
      ii. software-defined networking(SDN): there is a centralised remote server that gives intellectual advice(global plan); The router in SDN becomes less functional(switch)

2.2 analogy:taking a trip

forwarding: process of getting through single interchange
routing: process of planning trip from source to destination

3 Types of network protocol



1. The IP protocol is shared by every nodes in the network

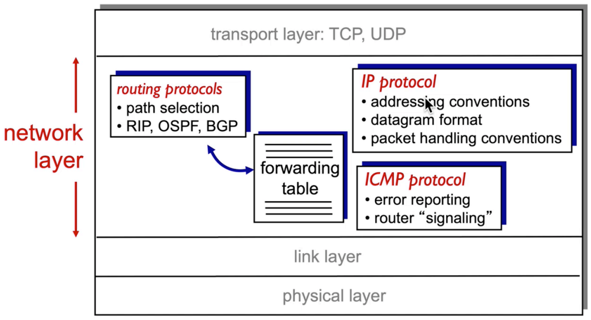
2. routing protocol: path selection such as RIP(intra-domain), OSPF(intra-domain), and BGP(only inter-domain protocol)
3. ICMP protocol: error reporting; router signaling

In reality, some servers may avoid certain protocol for some reasons, except IP protocol, which we can be sure that every nodes will run

4 IP Address

IP v4 address is a 32-bit(4 bytes) binary integer used to identify a host
A host can get an IP address either

1. manually configured by system administrator
2. automatically assigned by a DHCP server



4.1 interface

connection between host/router and physical link

1. router's typically have multiple interface
2. host typically has **one or two or more** interfaces(e.g., wired Ethernet, wireless 802.11(wi-fi)); Typically today's host has at least two interface

IP address associated with each interface, so the IP address is used to identify the interface

5 IP address and Network Interface

An IP address is associated with a network interface

1. A host usually has one or two network interfaces(e.g wired Ethernet and WiFi) A router typically has multiple interfaces(e.g. subnet)

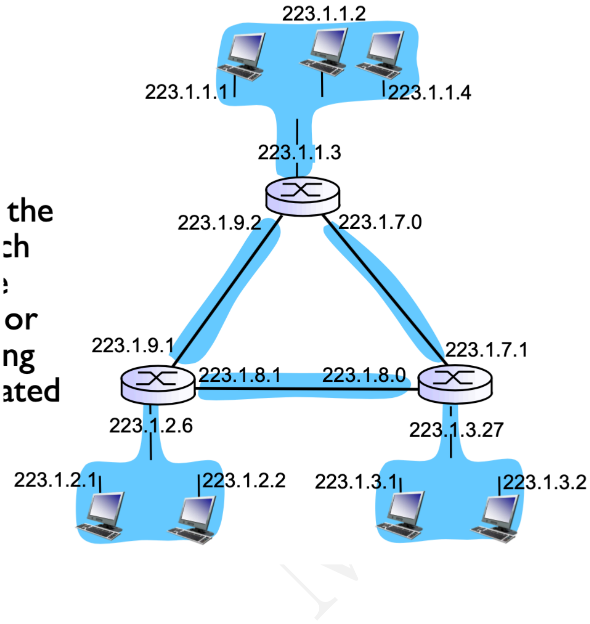
5.1 IP address and Subnet

An IP address logically comproises two parts: $\overbrace{\underbrace{Network(subnet)prefix}_{n\,bits}\,|\,\underbrace{host\,ID}_{32-n\,bits}}^{32\,bits}$

5.1.1 Subnet

Subnet is a network formed by a group of **"directly"** interconnected hosts
Properties of subsets

1. Hosts in the same subnet have the same network prefix in their IP address
2. Hosts in the same subnet can physically reach each other without intervening router
3. They connect to the ouside world through a router

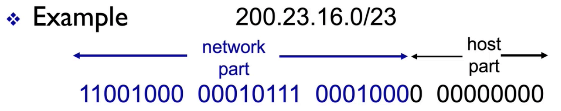


The IP protocol implemented by router that determines the routing path and transport information from one subnet to another subnet
For the network above, there are six subnet
The way to decide the number of subnets: detach each interface from its host or router, creating islands of isolateds network(recipe)
IMPORTANT: the subnet does not depend on whether there are hosts connected with it or not, my simple way to decide how many subnets there are is to remove all router and see how many connected structure there are, in the example above, there are 6 independent connected structure(containing three links but still independent)

6 IP address assignemnt in Internet

Internet's IP address assignemnt strategy is known as **Classless Inter-domain Routing**(CIDR). For ciroirations, they will be offered with a continnum of IP address with fixed subnet prefix by an ISP

1. Subnet prefix of IP address is of arbitrary length
2. Address format: a.b.c.d/x, where x refers to the number of bits in subnet prefix of IP address(the x here has the same function with subnet mask)



6.1 Subnet mask

Subnet mask is used to determine which subnet an IP address belong to
Subnet mask is calculated by setting all subnet prefix bits to 1 and hosts ID bits to 0

6.2 ICANN

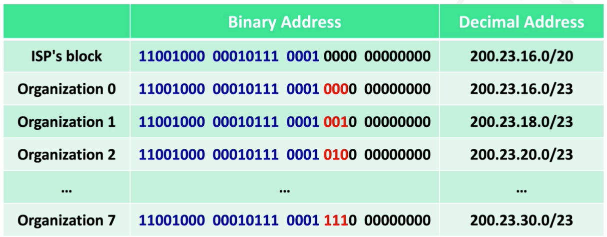
Internet corporation for ASSIGNED nAMES AND NUMBERS
http://www.icann.org./

1. allocate addresses
2. manages DNS
3. assigns domain names, resolves disputes

Special Address	Present Use
0.0.0.0/8	Non-routable meta-address for special use
127.0.0.0/8	Loopback address A datagram sent to an address within this block loops back inside the host This is ordinarily implemented using only 127.0 0.1/32; when u call the address as the identification, the message will look back to the same host the purpose of the address is for debugging purpose
10.0.0.0/8 172.16.0.0/12 192.168.0.0/16	Private addresses They can be used without any coordination IANA or an Internet registry can assume nobody else use it, so can safely use it in private net Within the organization, do the internal test
255.255.255.255/32	Broadcast addresses All hosts on the same subnet receive a datagram with such a destination address try to broadcast certain information to everyone

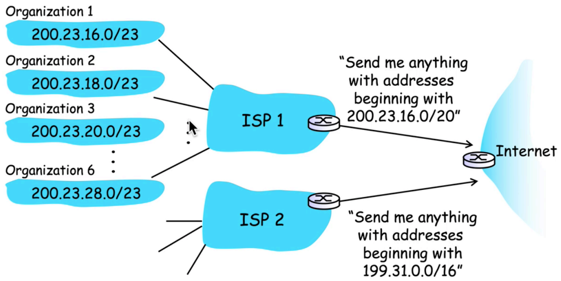
6.3 Hierarchical Address

1. This hierarchical addressing via restricting subnet prefix at all levels allows efficient advertisement of routing information(which is known as route aggregation)



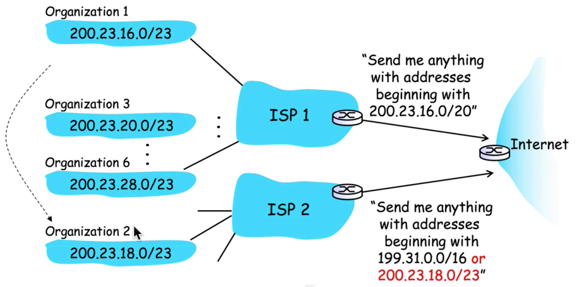
	Binary Address	Decimal Address
ISP's block	11001000 00010111 0001 0000 00000000	200.23.16.0/20
Organization 0	11001000 00010111 0001 0000 00000000	200.23.16.0/23
Organization 1	11001000 00010111 0001 0010 00000000	200.23.18.0/23
Organization 2	11001000 00010111 0001 0100 00000000	200.23.20.0/23
...	...	...
Organization 7	11001000 00010111 0001 1110 00000000	200.23.30.0/23

The higher organization can divide its s IP space and sell
Each level can divert/branch the IP address to a more specific lower level(with a more specific subnet mask)



This mechanism is called route aggregation, every organization router under ISP1 send pkt to ISP1, and it's ISP1's job to further differentiate; The ISP1 is gonna send the IP address with the subnet mask, which denotes the part of network it belongs to. With the subnet mask, all the pkt from or to organization under the ISP1 would be the same source/destination IP(i.e. ISP1) for other outer routers(So for the internet, it should not differentiate whether a pkt is for organization 1 or 2, all are belongs to ISP1, and it's ISP1's job to further differentiate and forward to the inividual customers
This is a hierachical addressing structure

2. Router uses longest prefix match in forwarding table when determining to which next hop IP datagram is sent


Organization 1
200.23.16.0/23
Organization 3
200.23.20.0/23
Organization 6
200.23.28.0/23
Organization 2
200.23.18.0/23
ISP 1
ISP 2
"Send me anything with addresses beginning with 200.23.16.0/20"
Internet
"Send me anything with addresses beginning with 199.31.0.0/16 or 200.23.18.0/23"

In the example above, if one host original belongs to ISP1, but later, it moves to ISP2, it would be inefficient for the host(200.23.18.0/23) to inform every one about the change; The way to solve the problem is to attach a "or 200.23.18.0/23" behind the filter condition of ISP2, and **longest prefix matching** is applied and the ISP2's condition givens longer prefix for 200.23.18.0/23(i.e. more specific information)

6.3.1 In the example above, what if the organization 2 later moves to a third ISP say ISP3(or just move back ISP1), in this case, both ISP 3 and ISP 2 will have the ”or 200.23.18.0/23” appendence, in this case, the two ISPs both provide the same length prefix, how to differentiate this?

If 200.23.18.0/23 moves from ISP 2 to a new ISP 3, ISP 2 will need to stop advertising that block of address since it does not own it anymore. Instead, ISP 3 will need to announce 200.23.18.0/23.
In reality, it is possible that two ISPs announce alternative routes to the same destination, since there might be multiple routing paths (via different ISPs) towards the destination. The receiving ISP will use its own routing policy (using inter domain routing protocol BGP, which is not covered in this introductory module) to choose which route path to use.

7 IP and routing

Two ways:

1. hard-coded by system admin in a file
   (a) Windows: control-panel ->network->configuration->tcp/ip->properties
   (b) UNIX:/etc/tc.config

8 Dynamic Host Configuration Protocol(DHCP)

DHCP allows a host to dynamically obtain its IP address from DHCP server when it joins network. It has the following benefits:

1. IP address is renewable,it can renew its lease on address in use
2. allow reuse of address(only hold addrress while connected)
3. support mobile user who want to join network
4. It has a ”plug and play” or zero configuration property, everything is automati
5. within the same subnet as the host

8.1 DHCP IP assignment

Assignment of IP address by DHCP server involves a 4-step process:

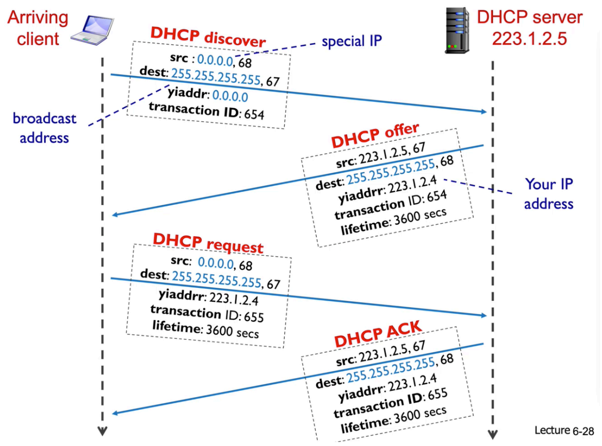
1. **Host broadcasts ”DHCP discover” message**
   Format:

A discover pkt is sent because the hosts knows nothing
src:0.0.0.0:68, //the host does not has IP yet
dest:255.255.255.255.255:67 //the host also does not know the IP of DHCP server, so just broadcast
yiaddr: 0.0.0.0 // the IP which will be allocated, so before allocated this field is 0
transaction ID:654 //Indicate the request

2. **DHCP server responds with ”DHCP offer” message**
   Format:
   src: 223.1.2.5.67(IP address of DHCP server) //The IP address of the DHCP
   dest: 255.255.255.255:68 //The DHCP server also does not know the IP address of the host, the broadcast
   yiaddr: 223.1.2.4 // The allocated IP
   transaction ID: 654
   lifetime: 3600 secs

3. **Host requests IP address: ”DHCP request” message**
   The host may get multiple offers, after choosing the offer it want Format:
   src: 0.0.0.0:68
   dest: 255.255.255.255:67 //Even if the host know the IP address of the server, it still broadcast it, perhaps to also let other DHCP server know
   yiaddr: 233.1.2.4
   transaction ID: 655
   lifetime: 3600 secs

4. **DHCP server sends address: ”DHCP ACK” message**
   Format: src: 223.1.2.5:67
   dest: 255.255.255.255:68
   yiaddr: 223.1.2.4 transaction
   ID: 655
   lifetime: 3600 secs



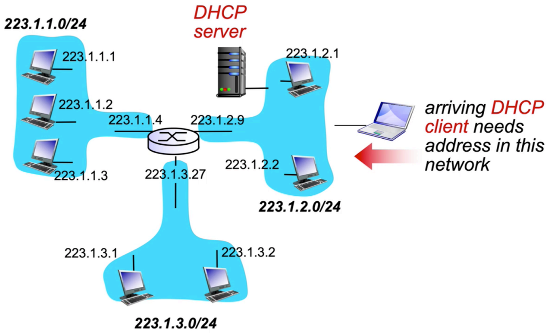
In addiction to host IP address asssignment, DHCP may alsxo provide a host additional network information, such as:

1. IP address of the first-hop router(the default router or default gateway)
2. IP address of local DNS server
3. Network Mask(network prefix)
4. DHCP server port number: 67
5. DHCP client port number: 68

IMPORTANT: DHCP runs over UDP

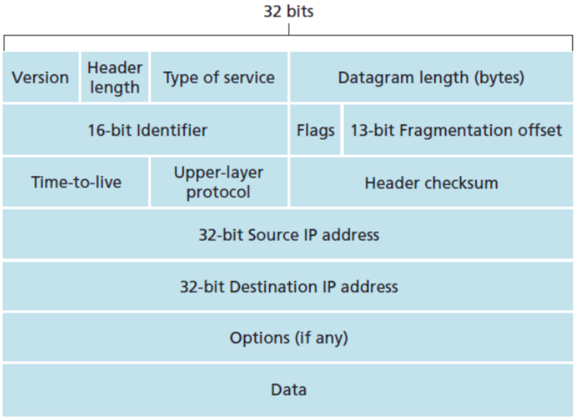
Special Address	Present Use
0.0.0.0/8	Non-routable meta-address for special use
127.0.0.0/8	Loopback address A datagram sent to an address within this block loops back inside the host This is ordinarily implemented using only 127.0 0.1/32; when u call the address as the identification, the message will look back to the same host the purpose of the address is for debugging purpose
10.0.0.0/8 172.16.0.0/12 192.168.0.0/16	Private addresses They can be used without any coordination IANA or an Internet registry can assume nobody else use it, so can safely use it in private net Within the organization, do the internal test
255.255.255.255/32	Broadcast addresses All hosts on the same subnet receive a datagram with such a destination address try to broadcast certain information to everyone

8.1.1 Excercise



If there is one DHCP server serves for the assignment for the network consists of multiple subnet; when a new host comes(with no IP address), how can the DHCP applocates the IP address to the new host through the router, which depends on the IP address(the new host has no IP)
In each subnet, there is certain software agents(with IP address), that will communicate with the DHCP through the router and get the IP address for the new hosts in the subnet

8.2 IP v4 Datagram format


32 bits

Version	Header length	Type of service	Datagram length (bytes)	
16-bit Identifier			Flags	13-bit Fragmentation offset
Time-to-live		Upper-layer protocol	Header checksum	
32-bit Source IP address				
32-bit Destination IP address				
Options (if any)				
Data				

The key fields in the IP v4 datagram are the following

1. Version Number: These 4 bits specify the IP protocol
2. Datagram length: Total length of IP datagram(header plus data), measured in **bytes**
3. Identifier, flag, fragmentation offset are used in IP fragmentation
    - identifier: a 16 bit value that is unique for every datagram for a given source address,

    destination address & protocol, such that they don't repeat within tha maximum datagram lifetime(MDL)

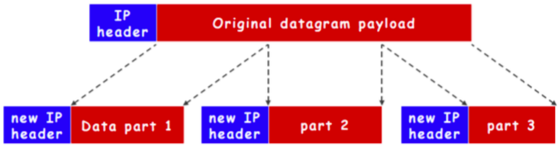
  - flag: a three-bit field that helps you to control and identify fragments. Bit 2: means more fragments
  - The Fragment Offset field (13 bits) is used to indicate the starting position of the data in the fragment in relation to the start of the data in the original packet; in unit of 8 bytes

4. Time-to-live(TLL) refers to number of remaining hops. decremented at each router
5. Header checksum is used to detect error for the IP header
6. source and destination IP address refers to the initial sender and final receiver's IP address

A typical IP v4 header is 20 bytes long

9 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.



9.1 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

9.2 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

9.3 ID

Same datagram will use the same ID

9.4 Length

Length field will change accordingly in smaller packets

10 Intra-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.

10.1 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

10.2 Distance Vector Algorithm

Routers know physically-connected neighbours and link costrs to neighbors. Router exchange local views with neighbours anb 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

Let c(x,y) be thew cost of link between router x and y. c(x,y) =∞ 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

10.2.1 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 ddown 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

11 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

12 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 ont he Internet

12.1 Implementation of NAT

NAT enabled routers must:

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

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 wihtout nitifying 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

13 Internet Control Message Protocol (ICMP)

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.

13.1 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 hostp