RED NOTES

UNIT 6.- NETWORK LAYER

1.-INTRODUCTION

The **network layer** oversees transporting segments from sending to receiving hosts. On sending side encapsulates segments into datagrams. On receiving side, delivers segments to transport layer. The network layer protocols are in every host or router.

In order to carry packages through the network, the **IP protocol** is in charge of these problems:

- Identifying with addresses of the devices that intervene in the communication (IP addresses).
- Choosing a route in the network that allows you to reach the destination (routing).

The two key network layer functions are:

- **Forwarding**: moving packets from router's input to appropriate router output (router local action).
- **Routing**: determining the route taken by packets from source to destination (network wide process).

As usual, Internet's network layer provides **best-effort service**, that is:

- Timing between packets is not guaranteed to be preserved.
- Packets are not guaranteed to be received in the order in which they were sent.
- Eventual delivery of transmitted packets is not guaranteed.

2.-VIRTUAL CIRCUIT AND DATAGRAM NETWORKS

Datagram network provides network-layer **connectionless** service (like UDP).

Virtual circuit network provides network-layer **connection** service (like TCP).

Both are similar to the transport-layer services, although:

- **Service**: host-to-host.
- **No choice**: network provides one or the other.
- **Implementation**: in network core.

The virtual circuit model has three phases:

1. Connection establishment, circuit set up.
2. Data transfer, circuit is used.
3. Connection teardown, circuit is deleted.

Each packet carries VC identifier (not destination host address). Every router on source-destination path maintains the "state" for each passing connection.

A virtual circuit consists of:

- **Path** from source to destination.

- **VC numbers**, one number for each link along path, it can be changed on each link.
- **Entries in forwarding tables** in routers along path

Advantages:

- Less header overhead.

Disadvantages:

- At least 1 RTT delay before sending data.
- Routers must store status information about VCs
- If a link fails, the connection falls and it must be re-established.
- Buffer space reservation in routers to store packages if necessary.

Datagram VS Virtual circuits summary:

Issue	Datagrams	Virtual circuits
Setup phase	Not needed	Required
Router state	Per destination	Per connection
Addresses	Packet carries full address	Packet carries short labels
Routing	Per packet	Per circuit
Quality of service	Difficult to add	Easier to add

Longest prefix matching: when looking for forwarding table entry for given destination address, use longest address prefix that matches destination address.

3.-IP: INTERNET PROTOCOL

IPv4 datagram fields:

- Version number (4 bits).
- Header length (4 bits), expressed in words of 32 bits (min = 5).
- Datagram length (16 bits), measured in bytes.
- Type of service (8 bits), 3 bits for the priority (ignored), 4 bits for the type of service (only one bit to 1) and 1 bit to zero.
- Time-to-live, this field is decremented by one each time the datagram is processed by a router. It is dropped once 0 is reached.
- Protocol, it indicates the protocol of the datagram encapsulated: TCP (6), ICMP (1) and UDP (17).
- Header checksum, it must be recomputed and stored again at each router.
- Options.

IP address is 32-bit identifier for host, router interface.

An **interface** is a connection between host/router and physical link.

There is one IP address associated with each interface.

IP addresses v4 are represented as four decimal numbers obtained from the four bytes that make up the IP address. Each IP address has two fields: **Network identifier** and **Host identifier**.

Special IP Addresses:

- **Network IP Address**: Network identifier + All 0s
- **Loopback Address**: 127 + any value (localhost).
- **Host Address**: All 0s + All 0s (it is used when the host obtains its IP address automatically).
- **Directed Broadcast Addresses**: Network identifier + All 1s (to send to all host connected to this network).
- **Limited Broadcast Addresses**: All 1s + All 1s (to send to all computers on the network to which is connected the sending host).

There are two types of Internet addressing, depending on how it determines the length of the network prefix:

- **Classful IP addressing**: can take only the predefined number of bits 8, 16 or 24.
- **Classless IP addressing**: any number of bits can be assigned, but it requires a network mask.

Addresses within a private address space will only be unique within the private network. Routers don't route private addresses out into the Internet. The private address ranges are:

- 192.168.0.0/16
- 172.16.0.0/12
- 10.0.0.0/8

Subnet: give bits from the host to the network in order to divide this one.

Subnets can physically reach each other without intervening the router.

Route aggregation: use the common bits as netmask.

Hierarchical addressing allows efficient advertisement of routing information.

The **forwarding tables** contain information about possible destination networks and how to reach them. They should be compact and have only information about destination networks and the next router to reach them. The information in the forwarding tables is:

Destination network | Netmask | Route (next hop) | Output interface

4.-ROUTING ALGORITHMS

Classification of routing algorithms:

- **Global**: all routers have complete topology and link cost information.
- **Decentralized**: routers know physically-connected neighbours and link costs to neighbours. It is an iterative process of computation and exchange of information with neighbours.
- **Static**: routes change slowly over time.
- **Dynamic**: routes change more quickly.

Bellman-Ford equation:

Let:

$d_x(y)$ = cost of least-cost path from x to y.

Then:

$d_x(y) = \min(c(x,v) + d_v(y))$

From time-to-time, each node sends its own distance vector estimate to neighbours. When x receives a new distance vector estimate from neighbour, it updates its own distance vector using B-F equation.

Each node waits for a change in local link cost or a message from a neighbour, then recomputes the estimates and finally, if the distance vector to any destination has changed, it notifies its neighbours.

Considering cost changes:

- Good news travels fast.
- Bad news travels slow.

Solutions to a routing loop caused by a “count to infinity” problem:

- Limiting the diameter of the network.
- Poisoned reverse with Split horizon.

Split horizon: prohibiting a node from advertising a node back onto the interface from which it was learned.

Poisoned reverse: sets the number of cost to the unconnected node to a number that indicates “infinite”.

A Link-State routing algorithm: **Dijkstra’s Algorithm**.

1. **Initialization**:
2. N' = {u}
3. for all nodes v
4. if v adjacent to u
5. then D(v) = c(u,v)
6. else D(v) = ∞
7. **Loop**
8. find w not in N' such that D(w) is a minimum
9. add w to N'
10. update D(v) for all v adjacent to w and not in N’:
11. D(v) = min(D(v), D(w) + c(w,v))
12. until all nodes in N'

Notation:

- c(x, y): link cost from node x to y.
- D(v): current value of cost of path from source to destination v.

- p(v): predecessor node along path from source to v.
- N': set of nodes whose least cost path definitively known.

In order to have a hierarchical routing, we have to collect routers into regions, "autonomous systems" (AS). Each AS is within an ISP. The routers in the same AS run the same routing protocol, the "intra-AS" routing protocol.

Gateway router: it is a router at "edge" of its own AS that has a link to another router in another AS.

The forwarding table is configured by both intra and inter-AS routing algorithm:

- Intra-AS sets entries for internal destinations.
- Inter-AS and intra-AS sets entries for external destinations.

Hot potato routing: send packet towards closest of two routers.

5.-ROUTING IN THE INTERNET

The most common intra-AS routing protocols or **interior gateway protocols** (IGP) are:

- **RIP**: Routing Information Protocol.
- **OSPF**: Open Shortest Path First.
- **IGRP**: Interior Gateway Routing Protocol.

RIP uses UDP (port 520) and the implementation of the distance vector algorithm is:

- The distance is measured in hops (max = 15) and each link has cost 1.
- Distance vector is exchanged with neighbours every 30 seconds in response message (**advertisement**).
- Each advertisement lists up to 25 destination **subnets**.

If no advertisement is heard after 180 seconds, then the neighbour/link is declared dead, that starts these steps:

1. Routes via neighbour are invalidated.
2. New advertisements are sent to neighbours.
3. Neighbours in turn send out new advertisements if tables have changed.
4. Link failure information quickly propagates to entire net.
5. Poison reverse is used to prevent ping-pong loops (infinite distance = 16 hops).

OSPF uses link state algorithm. Its advertisements carry one entry per neighbour and are flooded to the entire AS (directly over IP). The network administrator decides the criteria to define the cost.

OSPF "advanced" features:

- **Security**: all OSPF messages are authenticated.
- **Multiple** same-cost **paths** allowed.
- For each link, there are multiple cost metrics for different ToS.
- Integrated **uni** and **multicast** support.
- **Hierarchical** OSPF in large domains.

The hierarchical OSPF is divided in two levels: **local area** and **backbone**. The link-state advertisements are only in the area. Each node has the detailed area topology, that is, only knows the shortest path to nets in other areas.

The **area border routers** summarize distances to nets in own area and advertise to other are border routers

The **backbone routers** run OSPF routing limited to the backbone.

The **boundary routers** connect to another AS's.

The Internet **inter-AS** routing **BGP** has different goals for different types of routing:

- Intra-AS: performance.
- Inter-AS: reachability.

BGP (Border Gateway Protocol) is the *de facto* inter-domain routing protocol. BGP provides each AS a mean to:

- Obtain subnet reachability information from neighbouring AS's.
- Propagate reachability information to all AS-internal routers.
- Determine "good" routes to other networks based on reachability information and policy.

It allows to subnets to advertise its existence to rest of Internet.

Basics of BGP:

- Uses TCP port 179.
- It is a path vector algorithm, similar to distance vector but with full paths. As it works with full paths, it prevents routing loops.
- Routers do not need to share a direct physical link.

6.-IPv6

IPv6 datagram has a fixed-length of 40 bytes header and the fragmentation is not allowed. This version was created as 32-bit addresses are not enough and the headers of this versions helps to speed the processing and forwarding.

The address is of 128 bits and is written in hexadecimal notation separated by ":". It includes a zero compression technique, then a sequence of zeros is replaced by a pair of ":" (only once). It uses CIDR notation (IPv6 Address / x).

Type of addresses:

- **Unicast**: address of a computer.
- **Multicast**: address of a group of computers (all).
- **Anycast**: address of a group of computers (anyone of the group).

The IPv6 datagram format adds this fields:

- **Priority**: identify priority among datagrams in flow.
- **Flow label**: identify datagrams in same flow.

- **Next header**: identify upper layer protocol for data.

And changes from IPv4:

- **Checksum**: removes entirely to reduce processing time at each loop.
- **Options**: allowed, but outside of header, indicated by “Next header” field.
- **ICMPv6**: new version of ICMP.

Not all routers can be upgraded to IPv6, thus the **tunnelling** technique is used. This technique is based on carry the IPv6 datagram as **payload** in IPv4 datagram among IPv4 routers.