Department Name: Computer Science
Course Name: M. Tech
Semester: 2nd
Paper Name: Advanced Computer Network (MTCS-201)
Topic: ICMP, Routing Protocol-RIP, OSPF, BGP, ATM

A. ICMP (INTERNET CONTROL MESSAGE PROTOCOL)

This protocol discusses a mechanism that gateways and hosts use to communicate control or error information. The Internet Control Message Protocol allows gateways to send error or control messages to other gateways or hosts; ICMP provides communication between the Internet Protocol software on one machine and the Internet Protocol software on another. This is a special purpose message mechanism added by the designers to the TCP/IP protocols. This is to allow gateways in an internet to report errors or provide information about unexpecter circumstances. The IP protocol itself contains nothing to help the sender test connectivity or learn about failures.

Motivation

- IP may fail to deliver datagrams because
  - the destination is not available
  - the time-to-live counter expires
  - routers become congested
- We need to let the sender know what has happened
- ICMP is a required part of IP

Purpose

- ICMP allows routers (and hosts) to send error or control messages to other routers or hosts.
- ICMP provides communication between the Internet Protocol software on one machine and the Internet Protocol software on another.
- Network-layer protocol to allow hosts & routers to communicate network- related information.
- ICMP information is carried as IP payload.

Restrictions

- ICMP messages are not generated for errors that result from datagrams carrying ICMP error messages. Why?
- ICMP is only sent to the original source. Why?

ICMP Segment Structure

- Variable-size segment; 8-byte minimum
- Type: command or status report ID
- Code: status code for the type
- Checksum: Checksum from ICMP header & data
- Rest of header: depends on type
    - Error reports contain the IP header & first 8 bytes of original datagram's data.

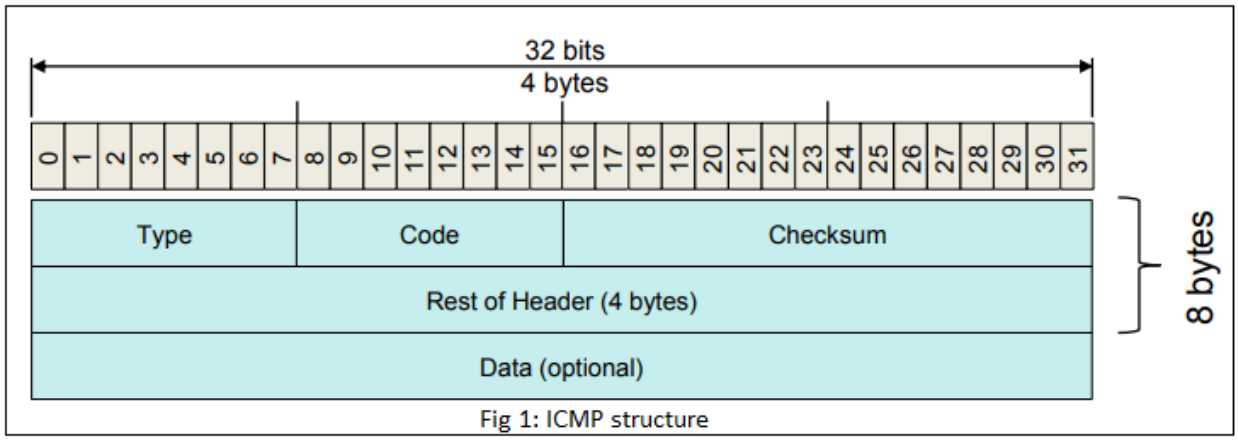


Fig 1: ICMP structure

ICMP Encapsulation

- ICMP is encapsulated in an IP packet, but is considered part of the IP or Internet layer.

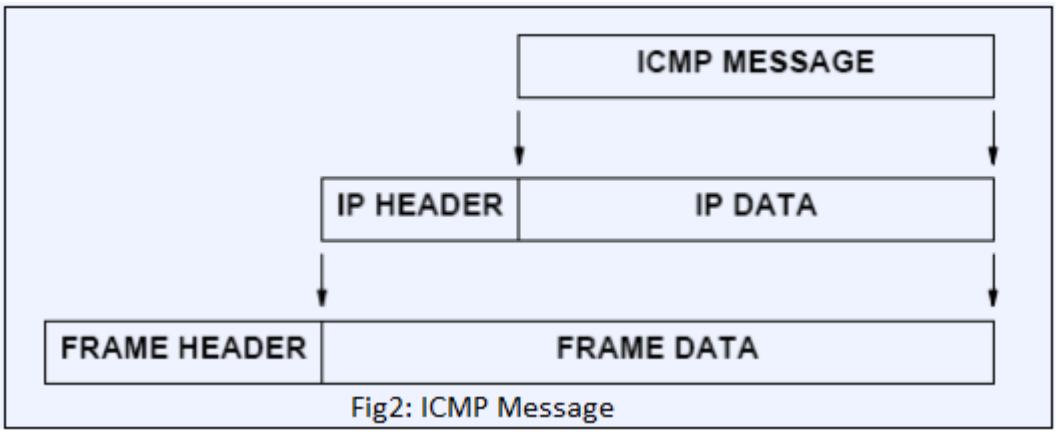


Fig2: ICMP Message

ICMP Messages

- The Common ICMP header

  - Each ICMP message has its own format, they all begin with the same three fields
  - TYPE (8-bit): identifies the message
  - CODE (8-bit): provides further information about the message type
  - CHECKSUM (16-bit):
  - In addition, ICMP messages that report errors always include the header and the first 64 data bits of the datagram causing the problem.

Type Field	ICMP Message Type
0	Echo Reply
3	Destination Unreachable
4	Source Quench
5	Redirect (change a route)
8	Echo Request
9	Router Advertisement
10	Router Solicitation
11	Time Exceeded for a Datagram
12	Parameter Problem on a Datagram
13	Timestamp Request
14	Timestamp Reply
15	Information Request (obsolete)
16	Information Reply (obsolete)
17	Address Mask Request
18	Address Mask Reply

Fig3: ICMP messages fields and types

Echo request and reply message

- Used to test reachability

- An echo request can also contain optional data (the content does not matter)
- An echo reply always returns exactly the same data as was received in the request
- Sent by ping program ™

Host Unreachable

- When a router cannot forward or deliver an IP datagram, it sends a destination unreachable message back to the original source.
- The CODE field specifies details ƒ
  - 0: network unreachable
  - 1: host unreachable ƒ
  - 2: protocol unreachable ƒ
  - 3: port unreachable ƒ
  - 4: fragmentation needed and DF (don't fragment) set ƒ
  - 5: source route failed

Source Quench

- To deal with congestion and datagram flow control
- When routers are overrun with traffic, it is called congestion.
- A machine uses ICMP source quench messages to report congestion to the original source.
- There is no ICMP message to reverse the effect of a source quench. Usually the host gradually increases the rate when no further source quench requests are received.

Route Redirect

- Routers exchange routing information periodically to accommodate network changes and keep their routes up-to-date. However, hosts do not do this.
- **A general rule:** Routers are assumed to know correct routes; hosts begin with minimal routing information and learn new routes from routers.
- When a router detects a host using a nonoptimal route, it sends the host an ICMP redirect message, requesting that the host change its rout.
- Limited to interactions between a router and a host on a directly connected network.

Attacks Using ICMP Messages

- Mapping Network Topology
- Mapping a target network is a very strategic part of most intelligently planned attacks. This

initial step in reconnaissance attempts to discover the live hosts in a target network. An attacker then can direct a more focused scan or exploit toward live hosts only.

- Sending individual ICMP echo: this is what the ping command does.
- Sending ICMP echo requests to the broadcast addresses of a network.
- Sending ICMP echo requests to network and broadcast address of subdivided networks
- Sending an ICMP address mask request to a host on the network to determine the subnet mask to better understand how to map efficiently.

Smurf Attack

- Ping a broadcast address, with the (spoofed) IP of a victim as source address
- All hosts on the network respond to the victim
- The victim is overwhelmed
- Keys: Amplification and IP spoofing
- Protocol vulnerability; implementation can be "patched" by violating the protocol
- Specification, to ignore pings to broadcast addresses
- ICMP echo just used for convenience
- All ICMP messages can be abused this way
- "Fraggle" is the equivalent with UDP

Ping of Death

- ICMP echo with fragmented packets
- Maximum legal size of an ICMP echo packet:
  65535 - 20 - 8 = 65507
- Fragmentation allows bypassing the maximum size:
  (offset + size) > 65535
- Reassembled packet would be larger than 65535 bytes
- OS crashes
- Same attack with different IP protocols

ICMP Redirect Attack

- Ask a host to send their packet to the target "router".
- Useful for man-in-the-middle attacks

- Winfreez(e)
    - Windows
    - ICMP Redirect: YOU are the quickest link to host Z
    - Host changes its routing table for Z to itself
    - Host sends packets to itself in an infinite loop

B. ROUTING PROTOCOL-ROUTING INFORMATION PROTOCOL (RIP)

- The Routing Information Protocol (RIP) is one of the oldest distance-vector routing protocols which employ the hop count as a routing metric.
- RIP prevents routing loops by implementing a limit on the number of hops allowed in a path from source to destination.
- The largest number of hops allowed for RIP is 15, which limits the size of networks that RIP can support.

HOP COUNT :

- Hop count is the number of routers occurring in between the source and destination network.
- The path with the lowest hop count is considered as the best route to reach a network and therefore placed in the routing table.
- RIP prevents routing loops by limiting the number of hopes allowed in a path from source and destination.
- The maximum hop count allowed for RIP is 15 and hop count of 16 is considered as network unreachable.

FEATURES OF RIP :

1. Updates of the network are exchanged periodically.
2. Updates (routing information) are always broadcast.
3. Full routing tables are sent in updates.

4. Routers always trust on routing information received from neighbor routers. This is also known as Routing on rumours.

RIP VERSIONS :

There are three vesions of routing information protocol – RIP Version1, RIP Version2 and RIPng.

a) RIP version 1

- The original specification of RIP, defined in RFC 1058, was published in 1988.
- **RIP v1** is known as *Classful* Routing Protocol because it doesn't send information of subnet mask in its routing update.
- When starting up, and every 30 seconds thereafter, a router with RIPv1 implementation broadcasts to 255.255.255.255 a request message through every RIPv1 enabled interface. Neighbouring routers receiving the request message respond with a RIPv1 segment, containing their routing table. The requesting router updates its own routing table, with the reachable IP network address, hop count and next hop, that is the router interface IP address from which the RIPv1 response was sent.
- RIPv1 enabled routers not only request the routing tables of other routers every 30 seconds, they also listen to incoming requests from neighbouring routers and send their own routing table in turn. RIPv1 routing tables are therefore updated every 25 to 35 seconds.
- The RIPv1 protocol adds a small random time variable to the update time, to avoid routing tables synchronising across a LAN. It was thought, as a result of random initialization, the routing updates would spread out in time, but this was not true in practice. Sally Floyd and Van Jacobson showed in 1994 that, without slight randomization of the update timer, the timers synchronized over time.

- RIPv1 can be configured into silent mode, so that a router requests and processes neighbouring routing tables, and keeps its routing table and hop count for reachable networks up to date, but does not needlessly sends its own routing table into the network. Silent mode is commonly implemented to hosts.

- RIPv1 uses classful routing. The periodic routing updates do not carry subnet information, lacking support for variable length subnet masks (VLSM). This limitation makes it impossible to have different-sized subnets inside of the same network class. In other words, all subnets in a network class must have the same size. There is also no support for router authentication, making RIP vulnerable to various attacks.

b) RIP version 2

- Due to the deficiencies of the original RIP specification, RIP version 2 (RIPv2) was developed in 1993, published as RFC 1723 in 1994, and declared Internet Standard 56 in 1998.
- It included the ability to carry subnet information, thus supporting Classless Inter-Domain Routing (CIDR).
- To maintain backward compatibility, the hop count limit of 15 remained.
- RIPv2 has facilities to fully interoperate with the earlier specification if all *Must Be Zero* protocol fields in the RIPv1 messages are properly specified. In addition, a *compatibility switch* feature allows fine-grained interoperability adjustments.
- In an effort to avoid unnecessary load on hosts that do not participate in routing, RIPv2 *multicasts* the entire routing table to all adjacent routers at the address *224.0.0.9*, as opposed to RIPv1 which uses broadcast.
- Unicast addressing is still allowed for special applications.
- (MD5) authentication for RIP was introduced in 1997.
- Route tags were also added in RIP version 2. This functionality allows a distinction between routes learned from the RIP protocol and routes learned from other protocols.

c) RIPng

- RIPng (RIP next generation), defined in RFC 2080, is an extension of RIPv2 for support of IPv6, the next generation Internet Protocol. The main differences between RIPv2 and RIPng are:

    - Support of IPv6 networking.
    - While RIPv2 supports RIPv1 updates authentication, RIPng does not. IPv6 routers were, at the time, supposed to use IPsec for authentication.
    - RIPv2 encodes the next-hop into each route entry, RIPng requires specific encoding of the next hop for a set of route entries.

- RIPng sends updates on UDP port 521 using the multicast group ff02::9.

RIP V1	RIP V2	RIPNG
Sends update as broadcast	Sends update as multicast	Sends update as multicast
Broadcast at 255.255.255.255	Multicast at 224.0.0.9	Multicast at FF02::9 (RIPng can only run on IPv6 networks)
Doesn't support authentication of update messages	Supports authentication of RIPv2 update messages	–
Classful routing protocol	Classless protocol, supports classful	Classless updates are sent

Fig4: RIP Versions

TIMERS

The routing information protocol uses the following timers as part of its operation:

- **Update Timer:** controls the interval between two gratuitous Response Messages. By default the value is 30 seconds. The response message is broadcast to all its RIP enabled interface.

- **Invalid Timer:** The invalid timer specifies how long a routing entry can be in the routing table without being updated. This is also called as expiration Timer. By default, the value is 180 seconds. After the timer expires the hop count of the routing entry will be set to 16, marking the destination as unreachable.

- **Flush Timer:** The flush timer controls the time between the route is invalidated or marked as unreachable and removal of entry from the routing table. By default the value is 240 seconds. This is 60 seconds longer than Invalid timer. So for 60 seconds the router will be advertising about this unreachable route to all its neighbours. This timer must be set to a higher value than the *invalid timer.*

- **Holddown Timer:** The hold-down timer is started per route entry, when the hop count is changing from lower value to higher value. This allows the route to get stabilized. During this time no update can be done to that routing entry. This is not part of the RFC 1058. This is Cisco's implementation. The default value of this timer is 180 seconds.

LIMITATIONS

- The hop count cannot exceed 15, or routes will be dropped.
- Variable Length Subnet Masks are not supported by RIP version 1 (which is obsolete).
- RIP has slow convergence and count to infinity problems.

C. Open Shortest Path First (OSPF) Protocol

- The Internet is made up of a large number of Autonomous Systems (AS). Each AS is operated by a different organization and can use its own routing algorithm inside.
- Network in an internetwork is independent of all the others, it is often referred to as an **Autonomous System** (**AS**).
- For example, the internal networks of companies *X*, *Y*, and *Z* are usually seen as three ASes if all three are on the Internet. All three may use different routing algorithms internally. Nevertheless, having standards, even for internal routing, simplifies the implementation at the boundaries between ASes and allows reuse of code.
- A routing algorithm within an AS is called an **interior gateway protocol**; an algorithm for routing between ASes is called an **exterior gateway protocol**.
- The original Internet interior gateway protocol was a distance vector protocol (RIP) based on the Bellman-Ford algorithm inherited from the ARPANET.
- In 1988, the Internet Engineering Task Force began work on a successor. That successor, called **OSPF** (**Open Shortest Path First**), became a standard in 1990.
- The OSPF protocol is a link-state routing protocol, which means that the routers exchange topology information with their nearest neighbors. The topology information is flooded throughout the AS, so that every router within the AS has a complete picture of the topology of the AS.
- OSPF supports three kinds of connections and networks:
    1. Point-to-point lines between exactly two routers.
    2. Multiaccess networks with broadcasting (e.g., most LANs).
    3. Multiaccess networks without broadcasting (e.g., most packet-switched WANs).
- A **multiaccess** network is one that can have multiple routers on it, each of which can directly communicate with all the others. All LANs and WANs have this property. Figure 5(a) shows an AS containing all three kinds of networks. Note that hosts do not generally play a role in OSPF.

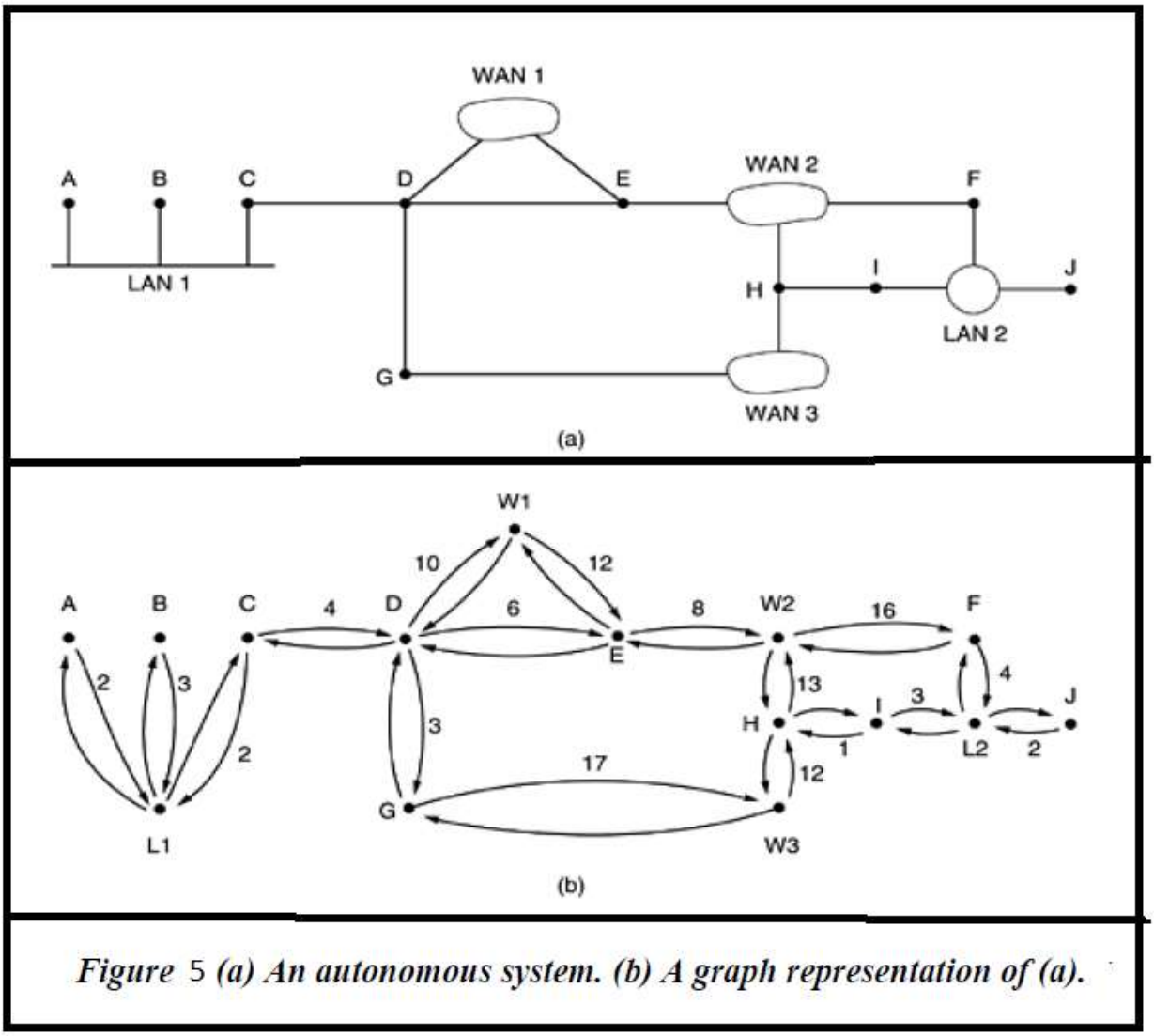


Figure 5 (a) An autonomous system. (b) A graph representation of (a).

- OSPF operates by abstracting the collection of actual networks, routers, and lines into a directed graph in which each arc is assigned a cost (distance, delay, etc.). It then computes the shortest path based on the weights on the arcs. A serial connection between two routers is represented by a pair of arcs, one in each direction. Their weights may be different. A multiaccess network is represented by a node for the network itself plus a node for each router. The arcs from the network node to the routers have weight 0 and are omitted from the graph.

- Figure 5(b) shows the graph representation of the network of Fig. 5(a). Weights are symmetric, unless marked otherwise. What OSPF fundamentally does is represent the actual network as a graph like this and then compute the shortest path from every router to every other router.
- Many of the ASes in the Internet are themselves large and nontrivial to manage. OSPF allows them to be divided into numbered areas, where an area is a network or a set of contiguous networks. Areas do not overlap but need not be exhaustive, that is, some routers may belong to no area. An area is a generalization of a subnet. Outside an area, its topology and details are not visible.
- Every AS has a backbone area, called area 0. All areas are connected to the backbone, possibly by tunnels, so it is possible to go from any area in the AS to any other area in the AS via the backbone. A tunnel is represented in the graph as an arc and has a cost. Each router that is connected to two or more areas is part of the backbone. As with other areas, the topology of the backbone is not visible outside the backbone.
- Within an area, each router has the same link state database and runs the same shortest path algorithm. Its main job is to calculate the shortest path from itself to every other router in the area, including the router that is connected to the backbone, of which there must be at least one. A router that connects to two areas needs the databases for both areas and must run the shortest path algorithm for each one separately.
- During normal operation, three kinds of routes may be needed: intra-area, interarea, and inter-AS. Intra-area routes are the easiest, since the source router already knows the shortest path to the destination router. Interarea routing always proceeds in three steps: go from the source to the backbone; go across the backbone to the destination area; go to the destination. This algorithm forces a star configuration on OSPF with the backbone being the hub and the other areas being spokes. Packets are routed from source to destination "as is." They are not encapsulated or tunneled, unless going to an area whose only connection to the backbone is a tunnel. Figure below shows part of the Internet with ASes and areas.

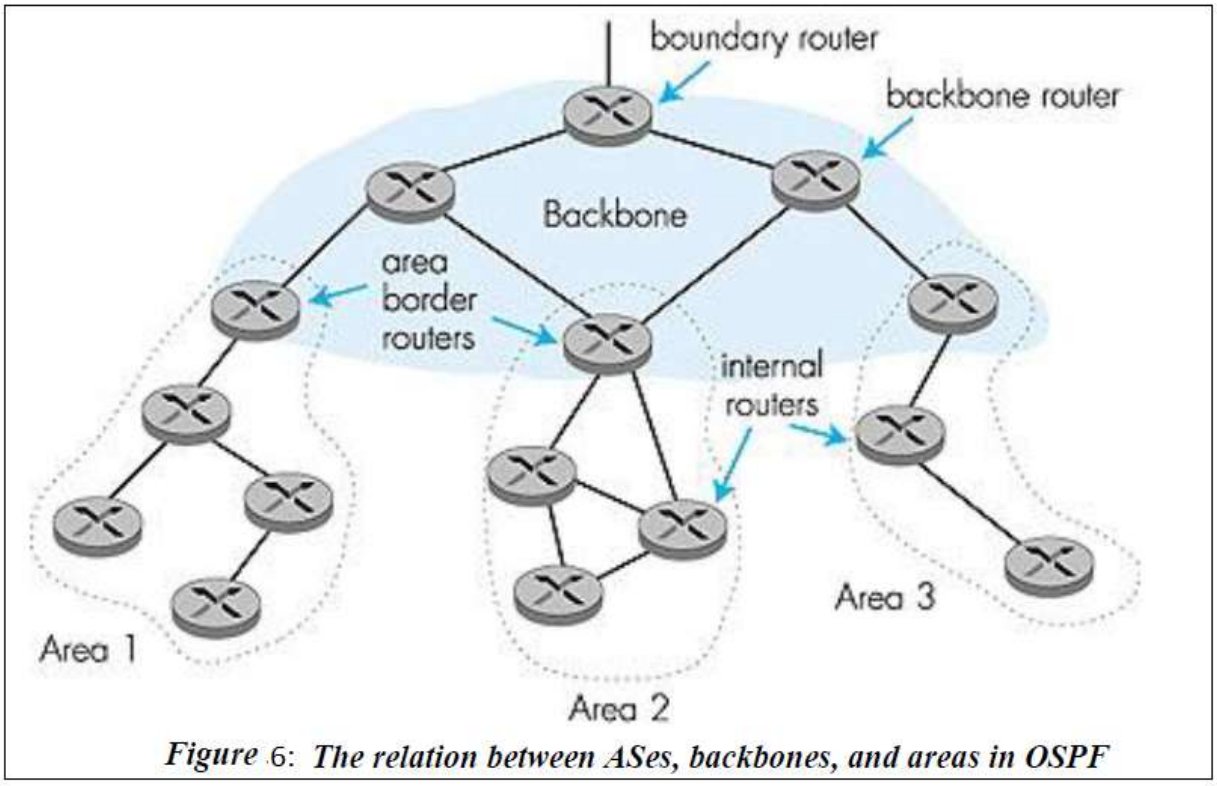


Figure** 6: **The relation between ASes, backbones, and areas in OSPF

- OSPF distinguishes four classes of routers:
    1. Internal routers are wholly within one area.
    2. Area border routers connect two or more areas.
    3. Backbone routers are on the backbone.
    4. AS boundary routers talk to routers in other ASes.

- These classes are allowed to overlap. For example, all the border routers are automatically part of the backbone. In addition, a router that is in the backbone but not part of any other area is also an internal router. Examples of all four classes of routers are illustrated in above Fig 6.
- There are five types of messages are used in OSPF.

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

Advantages and Disadvantages

- The main advantage of a link state routing protocol like OSPF is that the complete knowledge of topology allows routers to calculate routes that satisfy particular criteria. This can be useful for traffic engineering purposes, where routes can be constrained to meet particular quality of service requirements.

- The main disadvantage of a link state routing protocol is that it does not scale well as more routers are added to the routing domain. Increasing the number of routers increases the size and frequency of the topology updates, and also the length of time it takes to calculate end-to-end routes.

- This lack of scalability means that a link state routing protocol is unsuitable for routing across the Internet at large, which is the reason why IGPs only route traffic within a single AS.

OSPF Version 3 (OSPFv3)

OSPF version 2 (OSPFv2) is used with IPv4. OSPFv3 has been updated for compatibility with IPv6's 128-bit address space. However, this is not the only difference between OSPFv2 and OSPFv3. Other changes in OSPFv3, as defined in RFC 2740, include

- protocol processing per-link not per-subnet
- addition of flooding scope, which may be link-local, area or AS-wide
- removal of opaque LSAs support for multiple instances of OSPF per link
- various packet and LSA format changes (including removal of addressing semantics).

D. Border Gateway Protocol (BGP)

- BGP is a standardized exterior gateway protocol designed to exchange routing and reachability information among autonomous systems (AS) on the Internet.
- BGP is classified as a path vector protocol, and it makes routing decisions based on paths, network policies, or rule-sets configured by a network administrator and is involved in making core routing decisions.
- BGP may be used for routing within an autonomous system. In this application it is referred to as **Interior Border Gateway Protocol**, **Internal BGP**, or **iBGP**. In contrast, the Internet application of the protocol may be referred to as **Exterior Border Gateway Protocol**, **External BGP**, or **eBGP**.
- The Border Gateway Protocol has been in use on the Internet since 1994. The current version of BGP is version 4 (BGP4), which was published as RFC 4271 in 2006
- Border Gateway Protocol (BGP) is used to Exchange routing information for the internet and is the protocol used between ISP which are different ASes.
- The protocol can connect together any internetwork of Autonomous System (AS) using an arbitrary topology. The only requirement is that each AS have at least one router that is able to run BGP and that is router connect to at least one other AS's BGP router. BGP's main function is to exchange network reach-ability information with other BGP systems. Border Gateway Protocol constructs an autonomous systems' graph based on the information exchanged between BGP routers.

Characteristics

- **Inter-Autonomous System Configuration:** The main role of BGP is to provide communication between two autonomous systems.
- BGP supports Next-Hop Paradigm.
- Coordination among multiple BGP speakers within the AS (Autonomous System).

- **Path Information:** BGP advertisement also include path information, along with the reachable destination and next destination pair.
- **Policy Support:** BGP can implement policies that can be configured by the administrator. For ex:- a router running BGP can be configured to distinguish between the routes that are known within the AS and that which are known from outside the AS.
- Runs Over TCP.
- BGP conserve network Bandwidth.
- BGP supports CIDR.
- BGP also supports Security.

Functionalities

BGP peers performs 3 functions, which are given below.

- The first function consist of initial peer acquisition and authentication. both the peers established a TCP connection and perform message exchange that guarantees both sides have agreed to communicate.
- The second function mainly focus on sending of negative or positive reach-ability information.
- The third function verifies that the peers and the network connection between them are functioning correctly.

BGP Route Information Management Functions:

- **Route Storage:**

  Each BGP stores information about how to reach other networks.
- **Route Update:**

  In this task, Special techniques are used to determine when and how to use the information received from peers to properly update the routes.
- **Route Selection:**

  Each BGP uses the information in its route databases to select good routes to each network on the internet network.

- **Route advertisement:**

  Each BGP speaker regularly tells its peer what is knows about various networks and methods to reach them.

E. ASYNCHRONOUS TRANSFER MODE (ATM)

- 1980's effort by the phone companies to develop an integrated network standard (BISDN) that can support voice, data, video, etc. •
- ATM uses small (53 Bytes) fixed size packets called "cells".
- Cell switching has properties of both packet and circuit switching Easier to implement high speed switches
- 53 bytes – Small cells are good for voice traffic (limit sampling delays) For 64Kbps voice it takes 6 ms to fill a cell with data
- ATM networks are connection oriented
  – Virtual circuits.

ATM Reference Architecture

- Upper layers
  – Applications
  – TCP/IP
- ATM adaptation layer
  – Similar to transport layer
  – Provides interface between upper layers and ATM
    Break messages into cells and reassemble
- ATM layer
  – Cell switching
  – Congestion control
- Physical layer
  – ATM designed for SONET
    Synchronous optical network TDMA transmission scheme with 125 μs frames.

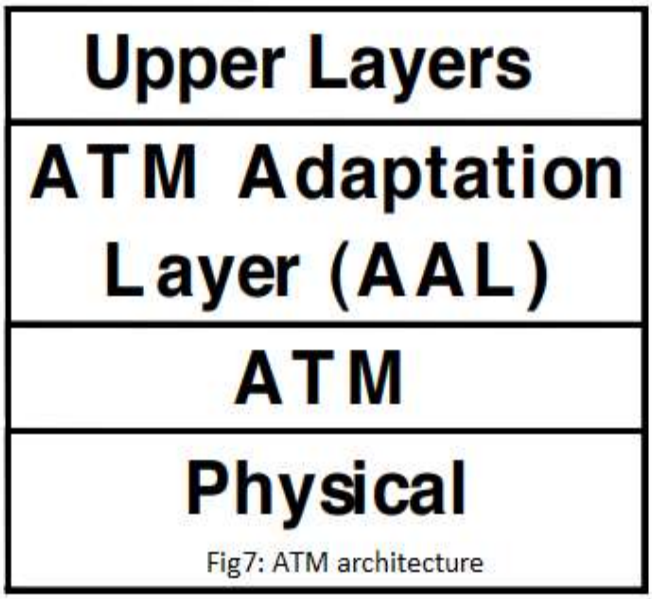
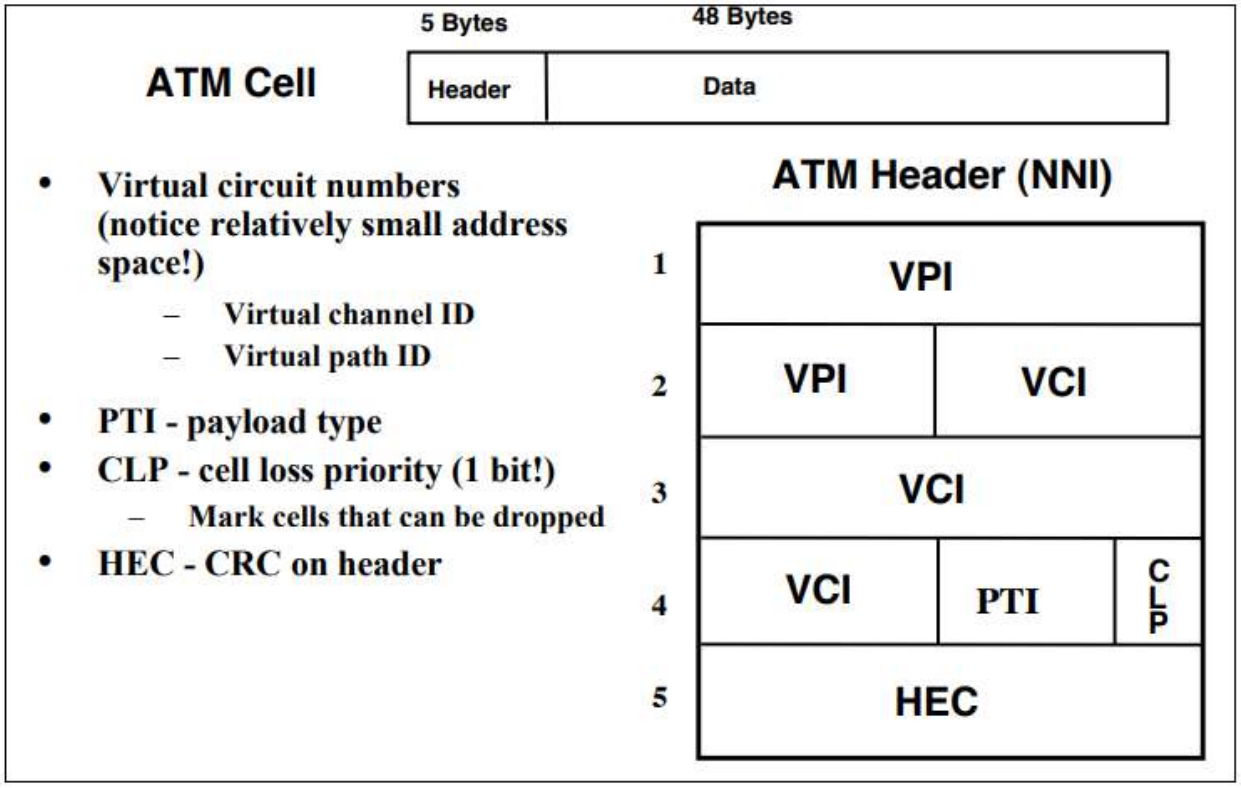


Fig7: ATM architecture

ATM Cell format



- Virtual circuit numbers (notice relatively small address space!)
  - Virtual channel ID
  - Virtual path ID
- PTI - payload type
- CLP - cell loss priority (1 bit!)
  - Mark cells that can be dropped
- HEC - CRC on header