Networking with TCP/ IP

MODULE 3

MODULE III: Internet Protocol: Purpose And Importance, Datagram Encapsulation, Time To Live (IPv4) And Hop Limit (IPv6),Forwarding In An Internet, Transmission across a single Network, The IP Forwarding Algorithm .The Internet Control Message Protocol, Error Reporting Vs. Error Correction, Testing destination reachability and status (Ping),Checksum computation.

Internet Protocol - *Purpose And Importance*

Internet Architecture And Philosophy

A user thinks of an internet as a single virtual network that interconnects all hosts, and through which communication is possible; its underlying architecture is both hidden and irrelevant. In a sense, an internet is an abstraction of physical networks

Conceptually, a TCPIIP internet provides three sets of services as shown in Figure; their arrangement in the figure suggests dependencies among them. At the lowest level, a connectionless delivery service provides a foundation on which everything rests.

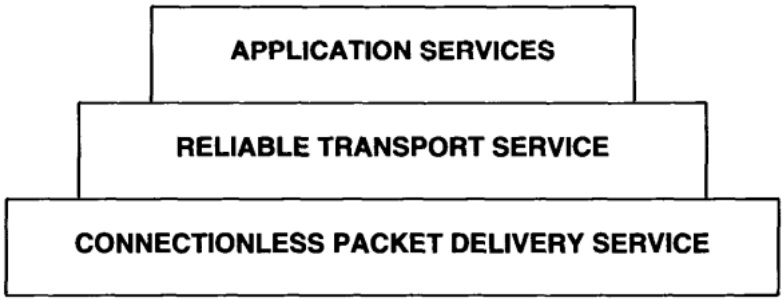
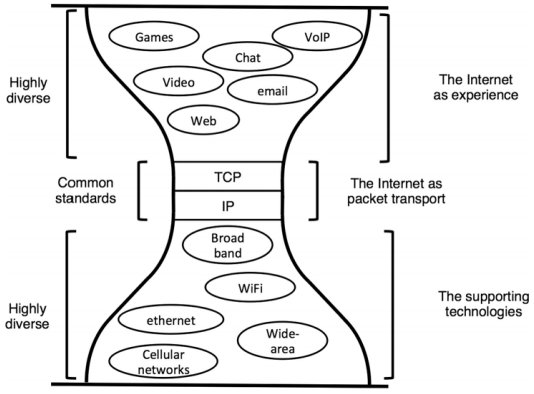


Figure 7.1 The three conceptual layers of internet services.

At the next level, a reliable transport service provides a higher level platform on which applications depend. This can be achieved by "narrow waist" design. One of the most significant advantages of this conceptual separation is that it becomes possible to replace one service without disturbing others.



Connectionless Delivery System

The most fundamental internet service consists of a packet delivery system. Technically, the service is defined as an **unreliable, best-effort, connectionless packet delivery system**, analogous to the service provided by network hardware that operates on a best-effort delivery paradigm. The service is called unreliable because delivery is not guaranteed. The packet may be lost, duplicated, delayed, or delivered out of order, but the service will not detect such conditions, nor will it inform the sender or receiver.

The service is called connectionless because each packet is treated independently from all others. A sequence of packets sent from one computer to another may travel over different paths, or some may be lost while others are delivered. Finally, the service is said to use best-effort delivery because the internet software makes an earnest attempt to deliver packets.

Purpose Of The Internet Protocol

The protocol that defines the unreliable, connectionless delivery mechanism is called the Internet Protocol and is usually referred to by its initials, IP. **IP provides three important definitions**.

First, the IP protocol defines the basic unit of data transfer used throughout a TCP/ IP internet. Thus, it specifies the exact format of all data as it passes across the internet. Second, IP software performs the routing function, choosing a path over which data will be sent.

Third, in addition to the precise, formal specification of data formats and routing, IP includes a set of rules that embody the idea of un reliable packet delivery. The rules characterize how hosts and routers should process packets, how and when error messages should be generated, and the conditions under which packets can be discarded.

The Internet Datagram

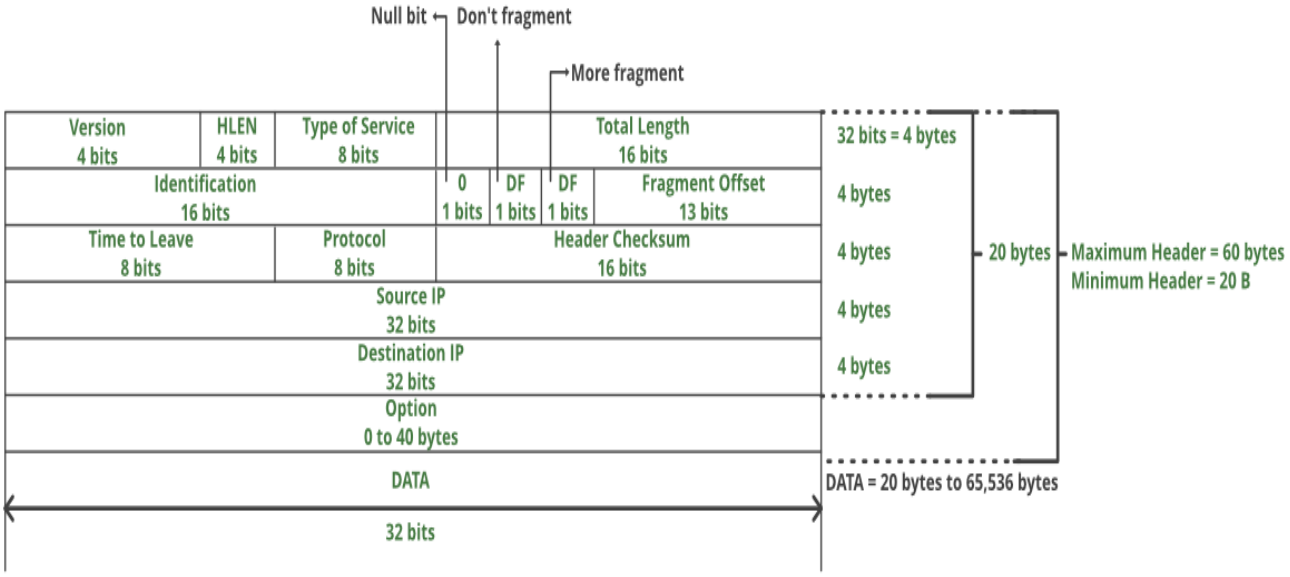
On a physical network, the unit of transfer is a frame that contains a header and data, where the header gives information such as the (physical) source and destination addresses. The

internet , its basic transfer unit internet datagram, sometimes referred to as an IP datagram or merely a datagram.

Figure 7.2 General form of an IP datagram, the TCP/IP analogy to a network frame. IP specifies the header format including the source and destination IP addresses. IP does not specify the format of the data area; it can be used to transport arbitrary data.

Like a typical physical network frame, a datagram is divided into header and data areas. Also like a frame, the datagram header contains the source and destination addresses and a type field that identifies the contents of the datagram. The difference, of course, is that the datagram header contains IP ad dresses whereas the frame header contains physical addresses.

Datagram Format



VERSION: Version of the IP protocol (4 bits), which is 4 for IPv4

HLEN: IP header length (4 bits), which is the number of 32 bit words in the header.

Type of service: Low Delay, High Throughput, Reliability (8 bits)

Total Length: Length of header + Data (16 bits), which has a minimum value 20 bytes and the maximum is 65,535 bytes.

Identification: Unique Packet Id for identifying the group of fragments of a single IP datagram (16 bits)

Flags: 3 flags of 1 bit each, indicate that the datagram has already been fragmented.

Fragment Offset: Represents the number of Data Bytes ahead of the particular fragment in the particular Datagram.

Time to live: Datagram's lifetime (8 bits), It prevents the datagram to loop through the network by restricting the number of Hops taken by a Packet before delivering to the Destination.

Protocol: Name of the protocol to which the data is to be passed (8 bits) Ex TCP, UDP, ICMP etc

Header Checksum: 16 bits header checksum for checking errors in the datagram header

Source IP address: 32 bits IP address of the sender

Destination IP address: 32 bits IP address of the receiver

Option: Optional information such as source route, record route. Used by the Network administrator to check whether a path is working or not.

Datagram Encapsulation

Datagrams are handled by software. They can be of any length the protocol designers choose. We have seen that the Pv4 datagram format allots 16 bits to the total length field, limiting the datagram to at most 65,535 octets. The entire IP datagram fits into one physical frame, making transmission across the physical net

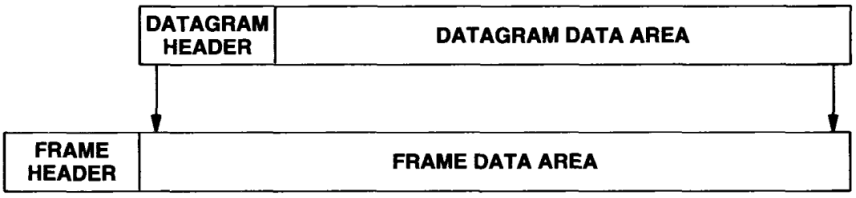


Figure 7.7 The encapsulation of an IP datagram in a frame. The physical network treats the entire datagram, including the header, as data.

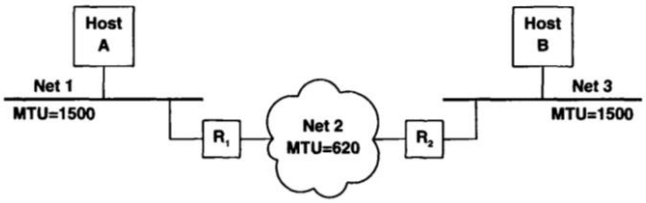
The idea of carrying one datagram in one network frame is called encapsulation. To the underlying network, a datagram is like any other message sent from one machine to another. As in Figure shows, when one machine sends an IP datagram to another, the entire datagram travels in the data portion of the network frame

Datagram Size, Network MTU, and Fragmentation

Each packet switching technology places a fixed upper bound on the amount of data that can be transferred in one physical frame, refer to these limits as the network's **Maximum transfer unit or MTU.** Allowing datagrams to be larger than the MTU in an internet means that a datagram may not always fit into a single network frame.

TCP/IP software chooses a convenient initial datagram size and arranges a way to divide large datagrams into smaller pieces when the datagram needs to traverse a network that has a small MTU.

The small pieces into which a datagram is divided are called fragments, and the process of dividing a datagram is known as fragmentation. The router receives a datagram from a network with a large MTU and must send it over a network for which the MTU is smaller than the datagram size.



In the figure, both hosts attach directly to Ethernets which have an MTU of 1500 octets. Thus, both hosts can generate and send datagrams up to 1500 octets long. The path between them, however, includes a network with an MTU of 620. If host A sends host B a datagram larger than 620 octets, router R, will fragment the datagram. Fragments must be reassembled to produce a complete copy of the original da tagram before it can be processed at the destination.

Fragmentation Control

Three fields in the datagram header, **IDENTIFICATION, FLAGS, and FRAGMENT OFFSET, control fragmentation and reassembly of datagrams**. Field IDENTIFICATION contains a unique integer that identifies the datagram. Its primary purpose is to al low the destination to know which arriving fragments belong to which datagrams. As a fragment arrives, the destination uses the IDENTIFICATION field along with the datagram source address to identify the datagram

For a fragment, field **FRAGMENT OFFSET** specifies the offset in the original datagram of the data being carried in the fragment. To reassemble the datagram, the destination must obtain all fragments starting with the fragment that has offset 0 through the fragment with highest offset.

The low-order two bits of the 3-bit **FLAGS field control fragmentation**. The first control bit aids in such testing by specifying whether the datagram may be fragmented. It is called the do not fragment bit because setting it to 1 specifies that the datagram should not be fragmented.

Time to Live (TTL) or Hope Limit

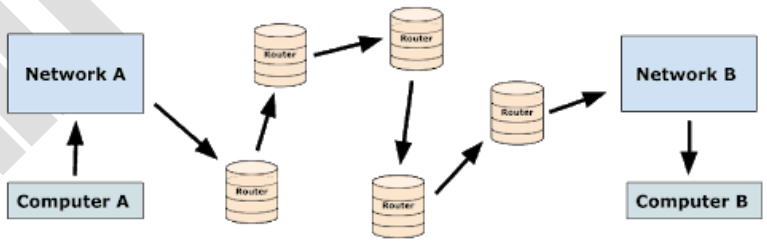
In principle, field TIME TO LIVE specifies how long, in seconds, the datagram is allowed to remain in the internet system. The idea is both simple and important: when ever a computer injects a datagram into the internet, it sets a maximum time that the datagram should survive. Routers and hosts that process datagrams must decrement the TIME TO LIVE field as time passes and remove the datagram from the internet when its time expires.

Each router along the path from source to destination is required to decrement the TIME TOLIVE field by I when it processes the datagram header. Whenever a TIME TO LIVE field reaches zero, the router discards the datagram and sends an error message back to the source.

Internet Protocol: Routing IP Datagrams

In a packet switching system, routing refers to the process of choosing a path over which to send packets, and router refers to a computer making the choice. Routing occurs at several levels. For example, within a wide area network that has multiple physical connections between packet switches, the network itself is responsible for routing packets from the time they enter until they leave.

Remember that the goal of IP is to provide a virtual network that encompasses multiple physical networks and offers a connectionless datagram delivery service. Thus, we will focus on IP forwarding, which is also called internet routing.



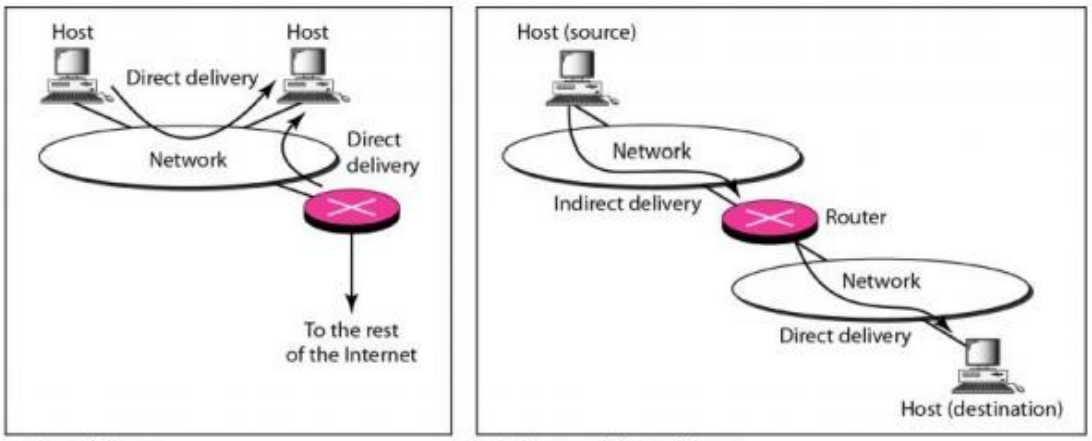
An internet is composed of multiple physical networks interconnected by **computers called router.** Routing in an internet can be difficult, especially among computers that have multiple physical network connections. Ideally, the routing software would examine net work load, datagram length, or the type of service specified in the datagram header when selecting the best path.

Direct And Indirect Delivery

Direct delivery, the transmission of a datagram from one machine across a single physical network directly to another, is the basis on which all internet communication rests. Two machines can engage in direct delivery only if they both attach directly to the same underlying physical transmission system (e.g., a single Ethernet)

IP addresses are divided into a network-specific prefix and a host-specific suffix. To see if a destination lies on one of the directly connected networks, the sender extracts the network portion of the destination IP address and compares it to the network portion of its own IP address. A match means the datagram can be sent directly.

Because the internet addresses of all machines on a single network include a common network pre& and extracting that pre& requires only a few machine instructions, testing whether a machine can be reached directly is extremely efficient.



a. Direct delivery

b. Indirect and direct delivery

In Indirect delivery, when one host wants to send to the other, it encapsulates the datagram and sends it to the nearest router. Once the frame reaches the router, software extracts the encapsulated datagram, and the IP software selects the next router along the path towards the destination. The datagram is again placed in a frame and sent over the next physical network to a second router, and so on, until it can be delivered directly.

Table-Driven IP Routing

The usual IP routing algorithm employs an Internet routing table (sometimes called an IP routing table) on each machine that stores information about possible destinations and how to reach them. Because both hosts and routers route datagrams, both have IP routing tables. Whenever the IP routing software in a host or router needs to transmit a datagram, it consults the routing table to decide where to send the datagram.

Routing table for A

Destination	Route
Host B	R1, R2, B

Routing table for R1

Destination	Route
Host B	R2, B

Routing table for R2

Destination	Route
Host B	B

(a) Routing tables based on route

If every routing table contained information about every possible destination address, it would be impossible to keep the tables current.

Next-Hop Method versus Route Method

One technique to reduce the contents of a routing table is called the next-hop method. In this technique, the routing table holds only the address of the next hop instead of information about the complete route (route method).

a. Routing tables based on route

Destination	Route
Host B	R1, R2, host B

Routing table for host A

Destination	Route
Host B	R2, host B

Routing table for R1

Destination	Route
Host B	Host B

Routing table for R2

b. Routing tables based on next hop

Destination	Next hop
Host B	R1

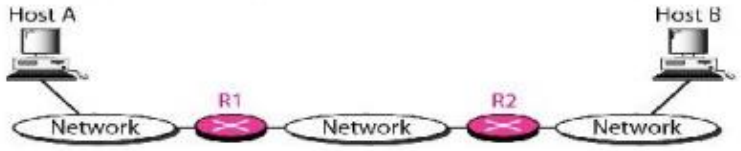
Routing table for host A

Destination	Next hop
Host B	R2

Routing table for R1

Destination	Next hop
Host B	---

Routing table for R2



Network-Specific Method versus Host-Specific Method

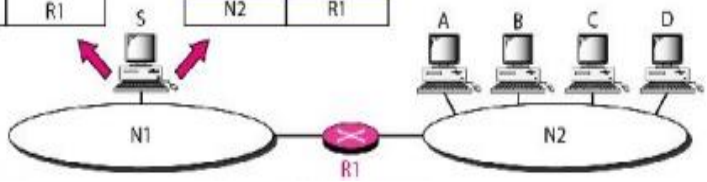
A second technique to reduce the routing table and simplify the searching process is called the network-specific method.

Routing table for host S based on host-specific method

Destination	Next hop
A	R1
B	R1
C	R1
D	R1

Routing table for host S based on network specific method

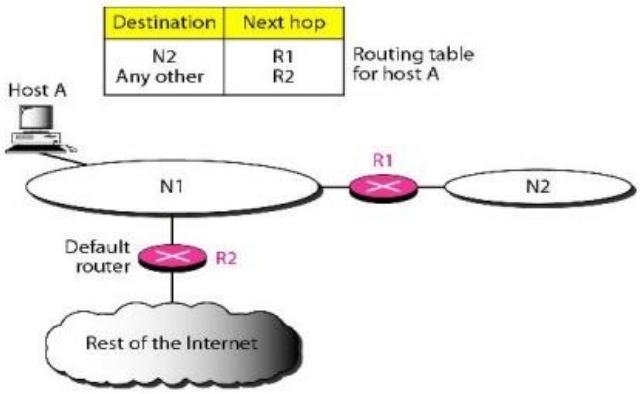
Destination	Next hop
N2	R1



Here, instead of having an entry for every destination host connected to the same physical network (host-specific method), we have only one entry that defines the address of the destination network itself

Default Method

Another technique to simplify routing is called the default method. Host A is connected to a network with two routers. Router R1 routes the packets to hosts connected to network N2. However, for the rest of the Internet, router R2 is used.



So instead of listing all networks in the entire Internet, host A can just have one entry called the default (normally defined as network address 0.0.0.0).

The IP Routing Algorithm

Algorithm:

RouteDatagram (Datagram, RoutingTable)

```
Extract destination IP address, D, from the datagram
    and compute the network prefix, N;
if N matches any directly connected network address
    deliver datagram to destination D over that network
    (This involves resolving D to a physical address,
    encapsulating the datagram, and sending the frame.)
else if the table contains a host-specific route for D
    send datagram to next-hop specified in table
else if the table contains a route for network N
    send datagram to next-hop specified in table
else if the table contains a default route
    send datagram to the default router specified in table
else declare a routing error;
```

Routing With IP Addresses

IP routing does not alter the original datagram. In particular, the datagram source and destination addresses remain unaltered; They always specify the IP address of the original source and the IP address of the ultimate destination

When IP executes the routing algorithm, it selects a new IP address, the IP address of the machine to which the datagram should be sent next. The new address is most likely the address of a router. After executing the routing algorithm, IP passes the datagram and the next hop address to the network interface software , for the physical network over which the datagram must be sent.

The network **interface software binds the next hop address to a physical address**, forms a **frame using that physical address**, **places the datagram in the data portion of the frame, and sends the result.**

After using the next hop address to find a physical address, the network interface software discards the next hop address.

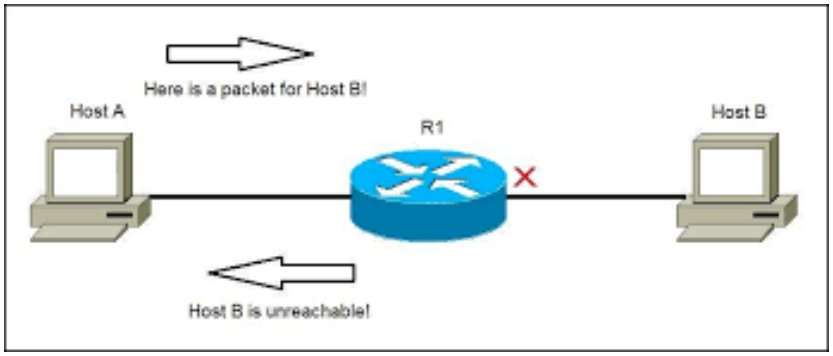
Error And Control Messages (ICMP)

In the connectionless system we have described so far, each router operates autonomously, routing or delivering datagrams that arrive without coordinating with the original sender. The system works well if all machines operate correctly and agree on routes.

Besides failures of communication lines and processors, IP fails to deliver datagrams when the destination machine **is temporarily or permanently disconnected from the network, when the time-to-live counter expires, or when intermediate routers become congested that they cannot process the incoming traffic.**

To allow routers in an internet to report errors or provide information about unexpected circumstances, the designers added a special-purpose message mechanism to the TCP/IP protocols. **The mechanism, known as the Internet Control Message Protocol (ICMP**

Like all other traffic, ICMP messages travel across the internet in the data portion of IP datagrams. The ultimate destination of an ICMP message is not an application program but the **Internet Protocol software on that machine.** That is, when an ICMP error message arrives, the ICMP software module handles it.

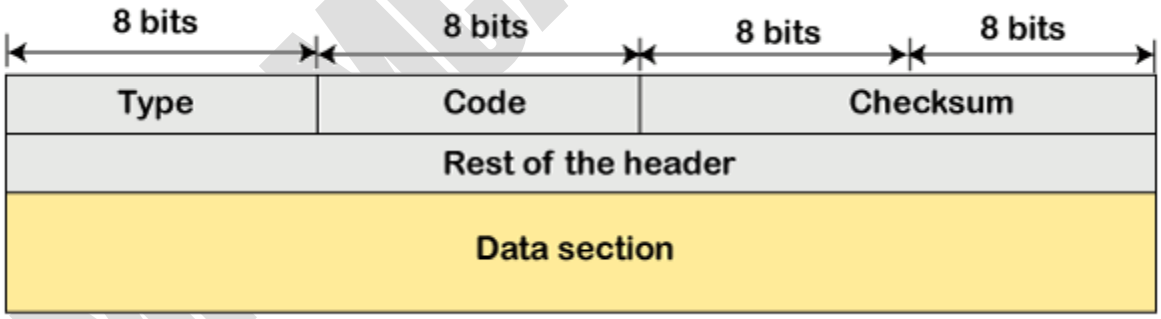


The Internet Control Message Protocol allows routers to send error or control messages to other routers or hosts; ICMP provides communication between the Internet Protocol software on one machine and the Internet Protocol software on another.

ICMP Message Format

The message format has two things; one is a category that tells us which type of message it is. If the message is of error type, the error message contains the type and the code. The type defines the type of message while the code defines the subtype of the message.

The ICMP message contains the following fields:



- **Type:** It is an 8-bit field. It defines the ICMP message type. The values range from 0 to 127 are defined for ICMPv6, and the values from 128 to 255 are the informational messages.
- **Code:** It is an 8-bit field that defines the subtype of the ICMP message
- **Checksum:** It is a 16-bit field to detect whether the error exists in the message or not.

ICMP messages

Category	Type	Message
Error-reporting messages	3	Destination unreachable
	4	Source quench
	11	Time exceeded
	12	Parameter problem
	5	Redirection
Query messages	8 or 0	Echo request or reply
	13 or 14	Timestamp request or reply

Error Reporting vs. Error Correction

ICMP messages are divided into error-reporting messages and query messages. The error-reporting messages report problems that a router or a host (destination) may encounter. The query messages get specific information from a router or another host.

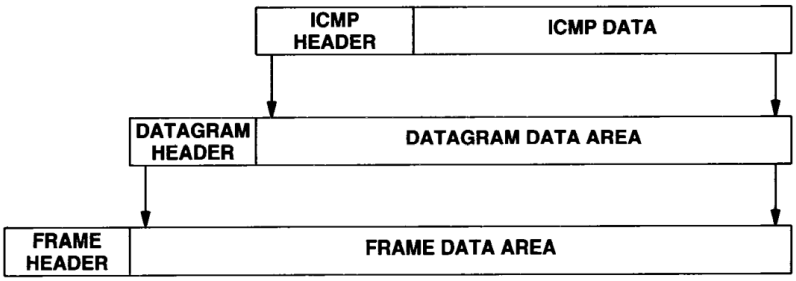
IP, as an unreliable protocol, is not concerned with error checking and error control. ICMP was designed, in part, to compensate for this shortcoming. ICMP does not correct errors, it simply reports them.

ICMP encapsulation and Message delivery

Each ICMP message travels across the internet in the data portion of an IP datagram, which itself travels across each physical network in the data portion of a frame.

Datagrams carry in ICMP messages are routed exactly like datagrams carrying information for users; there is no additional reliability or priority. Thus, error messages themselves may be lost or discarded. Furthermore, in an already congested network, the error message may cause additional congestion.

It is important to keep in mind that even though ICMP messages are encapsulated and sent using IP



Testing Destination Reachability And Status (Ping)

TCP/IP protocols provide facilities to help network managers or users identify network problems. One of the most frequently used debugging tools invokes the ICMP echo request and echo reply messages.

A host or router sends an ICMP echo request message to a specified destination.

Any machine that receives an echo request formulates an echo reply and returns it to the original sender. The request contains an optional data area; the reply contains a copy of the data sent in the request. The echo request and associated reply can be used to test whether a destination is reachable and responding.

PING

On many systems, the command users invoke to send ICMP echo requests is named ping. Sophisticated versions of ping send a series of ICMP echo requests, capture responses, and provide statistics about datagram loss.

PING is an acronym for Packer InterNet Groper

Because both the request and reply travel in IP datagrams, successful receipt of a reply verifies that major pieces of the transport system work. First, IP software on the source computer must route the datagram. Second, intermediate routers between the source and destination must be operating and must route the datagram correctly.

Third, the destination machine must be running (at least it must respond to interrupts), and both ICMP and IP software must be working. Finally, all routers along the return path must have correct routes.