INTERNET CONTROL MESSAGE PROTOCOL (ICMP)

Best-Effort Service

- Definition: *Best-effort* is a characteristic of a network system that makes a best attempt to deliver data, but does not guarantee delivery
- IP provides a best-effort service, i.e., it makes a best attempt to deliver data
- However, IP does not guarantee delivery, i.e., datagrams can be
  - Lost
  - Duplicated
  - Delayed
  - Delivered out of order
  - Corrupted
- A best-effort service does not mean a careless service; In fact, IP attempts to
  - Avoid errors
  - Report problems when they occur, if it is possible

An Error Detection Without Reporting

- When a host creates an IP datagram, the host includes a checksum that covers the entire header
- Whenever a datagram is received, the checksum is verified to ensure that the header arrived intact
- Note that, after changing fields in the header, e.g., after decrementing the *TIME-TO-LIVE* field, a router must recompute the checksum before forwarding the datagram to its next hop
- If the checksum verification indicates that some changes in the header bits occurred
  - The receiver can not send an error message back to the sender computer, because the receiver can not trust the source address
  - The receiver can not forward the damaged datagram, because the receiver can not trust the destination address
- Therefore, this datagram must be discarded without further processing, this is because the receiver can not trust any field in the datagram

Error/Information Reporting Mechanism

- Problems that are less severe than transmission errors result in error conditions that can be reported, e.g.,
  - Destination unreachable
  - Routing loop
  - Fragment loss
- TCP/IP suite includes a protocol called *Internet Control Message Protocol* (ICMP) to send error/information messages when conditions such as the one described above arise
- ICMP is a separate protocol inside the internet layer
- IP and ICMP protocols are co-dependent
  - IP uses ICMP when it sends an error/information message
  - ICMP uses IP to transport messages

Source Quench Message

- When a router receives so many datagrams which consume all its available buffer space, i.e., datagram overrun, the router
  - Must discard incoming datagrams, simply because it run out of buffer space
  - Sends a request to source hosts asking them to slow down, i.e., sending them a *source quench* message
- The *source quench* message is a request to the host to cut back the rate at which it is sending traffic to the internet destination
- A router sends a *source quench* message for each discarded datagram
- On receipt of a *source quench* message, the source host should cut back the rate at which it is sending traffic to the specified destination until it no longer receives *source quench* messages from that router
- A router may send the *source quench* message when it approaches its capacity limit, rather than waiting until the capacity is exceeded, i.e., the datagram which triggered the *source quench* message may be delivered.

Time Exceeded Message

- When a router reduces the *TIME-TO-LIVE* field in a datagram to zero, the router
  - Must discard this datagram
  - Notifies the source host with this incident, by sending a *time exceeded* message to the source host
- When a host reaches the reassembly expiration time before the arrival of all fragments from a given datagram, the host
  - Must discard this datagram
  - Notifies the source host with this incident, by sending a *time exceeded* message to the source host

Destination Unreachable Message

- When a router determines that a datagram can not be delivered to its final destination, the router
  - Notifies the source host with this incident, by sending a *destination unreachable* message to the source host
- The *destination unreachable* message specifies whether
  - The network to which the destination attaches is unreachable, e.g., temporarily disconnected from the internet, according to the information in the router, i.e., the routing table
  - The destination host is unreachable, e.g., temporarily off-line
  - The datagram must be fragmented, yet the *DO-NOT-FRAGMENT* flag is on

Redirect Message

- A router, R_1, receives a datagram from a host on a network to which the router is attached
- R_1 extracts the datagram destination network address, X
- R_1 checks its routing table for X and obtains the next hop, R_2
- The router forwards the datagram to R_2
- If R_2 and the source host are on the same network, the router
  - Notifies the source host with this incident, by sending a *redirect* message to the source host
- The *redirect* message advises the source host to send its traffic for network X directly to router R_2, as this is a shorter path to the destination, at least from the *TIME-TO-LIVE* point of view

Parameter Problem Message

- If a router, or a host, processing a datagram finds a problem with the header parameters, e.g., it cannot complete processing the datagram, it
  - Must discard this datagram
  - Notifies the source host with this incident, by sending a *parameter problem* message to the source host
- The *parameter problem* message is only sent if the error caused the datagram to be discarded.
- One potential source of such a problem is with incorrect arguments in an option field
- The *parameter problem* message also identifies the byte of the original datagram's header where the error was detected

Echo Request/Reply Messages

- An *echo request* message can be sent to the ICMP software on any computer
- In response to an incoming *echo request* message, ICMP software is required to send an ICMP *echo reply* message
- The data received in the echo message must be returned in the echo reply message; this data aid the sender in matching the replies with the echo requests
- The *echo request* and *echo reply* are not error messages
- This pair of messages usually tests whether a destination is reachable or not

Timestamp Request/Reply Messages

- Similar to echo request/reply messages
- A *timestamp reply* message returns a time stamp as well
- The time stamp is a 32-bit number representing the passed milliseconds since midnight Universal Time (UT)

Address Mask Request/Reply Messages

- When a host boots, it broadcast an *address mask request* message
- Routers that receive this request send an *address mask reply* message that contains the correct 32-bit subnet mask being used on the network

Information Request/Reply Messages

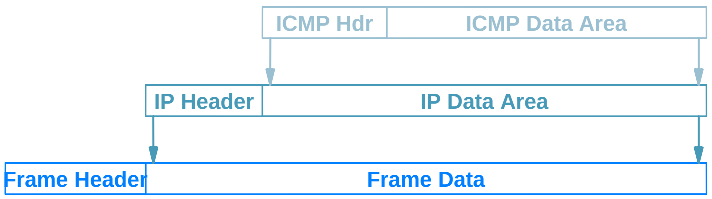
- Similar to address mask request/reply messages
- However, an *information reply* message returns the network number
- This *information request* and *information reply* pair of messages is a way for a host to find out its network number

ICMP Messages Summary

- ICMP messages include error reporting messages, e.g.,
  - Source quench message
  - Time exceeded message
  - Destination unreachable message
  - Redirect message
  - Parameter problem message
- ICMP messages also include information messages, e.g.,
  - Echo request/reply messages
  - Timestamp request/reply messages
  - Address mask request/reply messages
  - Information request/reply messages

ICMP Message Transport

- IP uses ICMP to report problems
- ICMP encapsulates messages in an IP datagram for transmission
  - The IP *TYPE* field specifies that an ICMP message is encapsulated in the IP data area
  - The ICMP message is placed in the data area of the IP datagram
- The datagram is then forwarded as usual, i.e., the complete datagram is encapsulated in a frame for transmission



Two level of encapsulation

ICMP Message Destination

- An ICMP reporting message is always created in response to a certain datagram, including
  - Datagram encounter a problem
  - Datagram carry an ICMP request message to which a router creates a reply
- In either case, a router sends the ICMP message back to the source of the datagram, by
  - Extracting the source address from the header of the incoming datagram
  - Placing the address in the *DESTINATION* field of the header of the datagram carrying this ICMP message

Avoiding an Infinite Loop

- What happens if
  - A datagram, D, causes an ICMP error message, I_1
  - I_1 causes an error, which generates another ICMP error message I_2
  - I_2 causes an error, which generates another ICMP error message I_3
  - Error messages keep cascading
- To avoid an internet from becoming congested due to carrying error messages about error messages, no error messages about ICMP error messages are sent

Using ICMP Messages to Test Reachability

- *Ping* is a Unix command which utilizes the ICMP protocol to see if a given destination can be reached or not
- When invoked, ping sends an IP datagram that contain an ICMP *echo request* message to the specified destination
- After sending the request, it waits a short time for the reply
- If a reply arrives,
  - Ping declares that the remote host is alive
- Else, i.e., if no reply arrives,
  - Ping retransmits the request
- If no reply arrives for the retransmissions, or if an ICMP *destination unreachable* message arrives,
  - Ping declares that the remote machine is not reachable

Using ICMP Messages to Trace a Route

- *Traceroute* is a Unix command which prints all the routers along a path to a given destination host
- To find the i^{th} router along the path to the destination,
  - Traceroute sends a datagram to the destination with the *TIME-TO-LIVE* is set to i
  - When the i^{th} router receives this datagram, it
    - Decrements the *TIME-TO-LIVE* counter in the header
    - Discards the datagram, since the counter reaches zero
    - Sends an ICMP *time exceeded* error message back to the source
  - Traceroute extracts the IP address from the ICMP reply and announces the address of the i^{th} router along the path
- *When should the traceroute procedure be stopped? How?*
- Note that, if routes change between two probes, the sequence of routers found by traceroute may not correspond to a valid path through the internet

Path MTU

- In a router, IP software fragments any datagram that is larger than the MTU of the network over which the datagram is being transmitted
- Although fragmentation solves the problem of heterogeneous networks, it often impacts performance; i.e., memory and CPU time are used to construct/reassemble fragments
- If the application chooses a datagram size less that or equal to the smallest network MTU along the path to the destination (known as the *path MTU*), no router will need to fragment the datagram
- If routes change, i.e., the path changes, the *path MTU* can be changed as well
- However, in many parts of the internet, routes tend to remain stable for days or weeks
- Hence, it makes sense for a computer to find the *path MTU* and create datagrams that are small enough to fit in the network without fragmentation

Using ICMP Messages for Path MTU Discovery

- An IP datagram header contains a bit to specify no fragmentation is allowed
- ICMP sends a *destination unreachable* error message when fragmentation required, but not permitted
- To find *path MTU*,
  - A host sends a sequence of datagrams that have the header bit set to prevent fragmentation
  - If a datagram is larger than the MTU of a network along the path, the router connected to that network will
    * Discard the datagram
    * Send an ICMP *destination unreachable* error message to the source host
  - The host can then send a smaller datagram, until one succeeds