ICMP (Internet Control Message Protocol)

Internet Protocols

CSC / ECE 573

Fall, 2005

N. C. State University

Today’s Lecture

I. ICMP Overview

II. ICMP Error Reporting

III. ICMP Query / Response Messages

IV. ICMP Message Processing

ICMP OVERVIEW

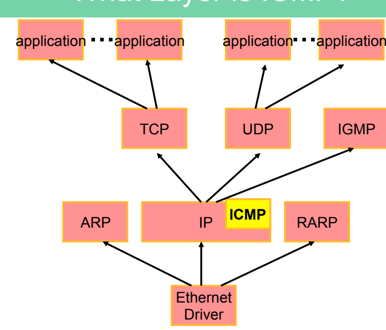
ICMP (RFC 792)

- Communicates…
  - network-level errors
  - information about unexpected circumstances
  - information about the network, in response to queries

ICMP (RFC 792)

- ICMP messages are sent only to the source of the packet causing the message, not to routers
  - Why?
- Error reporting only
  - does not specify corrective action to take
  - kernel, other protocol, or application must decide what to do

What Layer is ICMP?



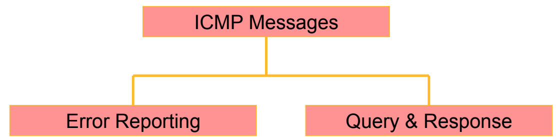
RFC-792 – “ICMP, uses the basic support of IP as if it were a higher level protocol, however, ICMP is actually an integral part of IP, and must be implemented by every IP module.”

ICMP Message Format

0 1 2 3 4 8	16	24 31
Type	Code	ICMP Message Checksum
--- Rest of ICMP Header ---		
ICMP Data (Original IP Header + 8 bytes datagram)		

- Checksum over entire ICMP message
- ICMP Data usually contains…
  - IP header (including Options, but normally = 20 bytes) of datagram that caused error
  - at least 8 bytes of data from this datagram (usually includes fields needed to identify the cause of the error)

ICMP Message Types



ERROR REPORTING WITH ICMP

Why ICMP for Reporting Errors?

- Protocol-specific messages?
- For what protocols or functions are ICMP error messages appropriate?

When Not to Send ICMP Error Messages

- An ICMP error message is never generated in response to:
  1. an ICMP error message
  2. a datagram whose source address does not define a single host (address cannot be zero, loopback, broadcast, multicast)
  3. A datagram whose destination address is an IP broadcast address
  4. a datagram sent as a link-layer broadcast
  5. a fragment other than the first one of a datagram
- For each of the above, why?

#1 Destination Unreachable Msgs

- Upon failure to forward/deliver, router sends ICMP message to source before "dropping" datagram
  - IP is best-effort delivery, but discarding datagrams should not be taken lightly
- Several reasons for failure (next slide), but…
  - not all errors can be diagnosed properly (e.g., host IP address changes)

Type	Code	ICMP Message Checksum
(unused)		
ICMP Data (Original IP Header + 8 bytes datagram)		

Reasons for Destination Unreachable Messages

- Network unreachable (reason?)
- Host unreachable (reason?)
- Protocol (TCP, UDP) not enabled
- Port not bound to a service
- Fragmentation needed, but DF Flag set
- Source route failed

Path MTU Discovery (RFC 1191)

- Host sets DF Flag and transmits a large datagram
- If datagram size exceeds MTU on some link, the router discards datagram and sends back ICMP Destination Unreachable message
  - message includes size of Next Hop MTU

0 1 2 3 4	8	16	24 31
Type	Code	ICMP Message Checksum	
(unused)		Next Hop MTU	
ICMP Data (Original IP Header + 8 bytes datagram)			

Path MTU Discovery (cont'd)

- Host receiving this error message knows to reduce maximum packet size to the Next Hop MTU
- Periodically host will increase the packet size and try again
  - Why?

#2 Time Exceeded ICMP Message

- Sent if router has detected that the hop count (TTL) has reached zero (code 0)
  - usually means a routing error (loop) occurred
  - why would loops occur if routing protocols work right?
- Or, sent if destination host timeout occurred while waiting for fragments to arrive (code 1)
  - normal timeout interval on the order of 60-120s

#3 Router Redirect Messages

- Hosts normally initialize their forwarding table from a (static) configuration file at startup
  - contains minimal info (e.g., address of single default gateway) for simplicity
  - if network topology changes, this info is obsolete
  - how learn of such changes (host don't run routing protocols)?
- Redirect messages do not solve the problem of propagating routes in a general way
  - dynamic routing protocols are used for this

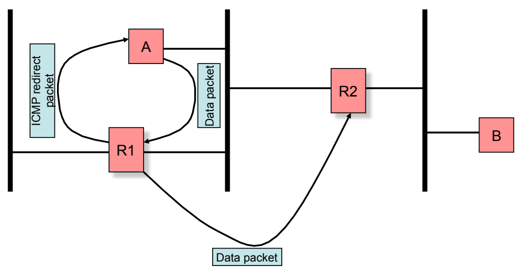
#3 Router Redirect Messages (cont'd)

1. When router detects a host using a "suboptimal" route…
   2. send ICMP Redirect message to the host requesting that it change its forwarding table
   3. forwards original datagram towards its destination

- How to detect a suboptimal route?
  - if router forwards packet out the same interface it came in on
- Which routes should be updated; only for this specific destination?

Example

- Packet from A to B should go through R2, but A sends to R1 first (i.e., A is misconfigured)



Redirect Message Format

- 3 addresses needed
  - IP address that caused redirect (in "Original IP Datagram" header)
  - IP address of router that sent redirect (in IP header of ICMP Message datagram)
  - correct router IP address (in Redirect message)

Restrictions on Redirection

- Redirect messages sent only by "first hop" router
- No Redirect message if Source Routing Option present
- + a few more restrictions (not covered here)

#4 Congestion and Datagram Flow Control

- IP is connectionless
  - does not reserve buffer space or bandwidth
  - potential for congestion, resulting in packet dropping by routers
- ICMP Source Quench message was used to report congestion to original source
  - a host receiving this message is expected to slow down
  - no ICMP message exists to reverse the effect of a source quench

#4 Congestion and Datagram Flow Control (cont'd)

- Not used any more
  - tends to create rather than solve congestion (why?)
  - congestion control in the Internet is now done mostly in the transport layer

#5 "Parameter Problem" Message

- Some error was detected in the IP header
- Pointer indicates byte offset from start of IP header to the "offending" parameter

0 1 2 3 4	8	16 24 31
Type=12	Code=0 or 1	ICMP Message Checksum
Pointer	*(unused)*	
ICMP Data (Original IP Header + 8 bytes datagram)		

A Clever Program: traceroute

- Allows us to see the path taken by the packet
  - why not just use IP record route option?

1. Send UDP datagram with TTL=1
   - first router decrements TTL, notices it is 0, sends ICMP Time Exceeded error message back to sender
   - this error message has IP address of the incoming interface of the router generating the error – now we know the first hop!

A Clever Program: traceroute (cont’d)

2. Now send UDP datagram with TTL=2
   - second router sends back “time exceeded” message, with its IP address
3. Etc…
4. How tell when the destination is reached?
   - the UDP datagram is addressed to an “unlikely” port (>30,000)
   - error message sent back by destination is Destination Unreachable (“port not bound to a service”) ICMP error message

Example Traceroute Output

```
C:\Documents and Settings\Douglas>tracert www.ietf.org

Tracing route to www.ietf.org [132.151.6.21]
over a maximum of 30 hops:

  1    <1 ms    <1 ms    <1 ms  cmdfhub-6509msfc-1.ncstate.net [152.14.19.2]
  2    <1 ms    <1 ms    <1 ms  ncsugw-gew3-1.ncstate.net [152.1.7.1]
  3    <1 ms    <1 ms    <1 ms  ncsugsr-gw-to-ncsu-lan.ncni.net [128.109.23.65]
  4     1 ms     1 ms     1 ms  rtp1-gw-to-core-oc48.ncren.net [128.109.52.6]
  5     1 ms     1 ms     1 ms  ge-6-1-101.hsa2.Raleigh1.Level3.net [64.158.228.1]
  6     1 ms     1 ms     1 ms  ge-7-0-0.mpls2.Raleigh1.Level3.net [209.244.22.37]
  7    10 ms    10 ms    10 ms  so-6-1-0.bbr1.Washington1.Level3.net [64.159.0.106]
  8    10 ms    10 ms    10 ms  so-6-0-0.edge1.Washington1.Level3.net [209.244.11.10]
  9    11 ms    11 ms    11 ms  uunet-level3-oc48.Washington1.Level3.net [209.244.219.158]

 10    11 ms    11 ms    11 ms  0.so-0-3-0.XL1.DCA5.ALTER.NET [152.63.43.170]
 11    13 ms    13 ms    13 ms  0.so-7-0-0.XL1.DCA6.ALTER.NET [152.63.42.190]
 12    13 ms    13 ms    13 ms  0.so-0-0-0.XR1.DCA6.ALTER.NET [152.63.35.113]
 13    14 ms    14 ms    14 ms  285.at-6-1-0.XR1.TCO1.ALTER.NET [152.63.33.58]
 14    15 ms    14 ms    14 ms  193.ATM6-0.GW5.TCO1.ALTER.NET [152.63.39.93]
 15    24 ms    18 ms    17 ms  cnrl-gw.customer.alter.net [157.130.44.142]
 16    19 ms    23 ms    19 ms  www.ietf.org [132.151.6.21]

Trace complete.
```

- “tracert” on Windows machines

QUERYING THE NETWORK WITH ICMP

#1 Echo Request and Reply Messages

- Used to see if destination interface is reachable and functioning

0 1 2 3 4 – 8	8 – 16	16 – 24 – 31
Type=8 or 0	Code=0	ICMP Message Checksum
Identifier		Sequence Number
Data		

- Echo Request
  - Contains Identifier and Sequence Numbers to help match Replies with Requests
- **Echo Reply**
  - is not mandated! reasons for not sending Echo Reply?
  - data sent by Request must be returned in Reply

Program Using Echo Request: ping

- Even if you can’t ping a host, it might be reachable (i.e., ping is disabled on that host but other services are not)
- Identifier = process number of application sending the ping
- Sequence Number starts at 0 and is incremented by each successive Request
  - can tell if replies are missing, duplicated, or reordered
- Round-trip time can be calculated
  - client puts sending time into Request, subtracts from receiving time when Reply comes back

ping Example

```
ping -s kronos.csc.ncsu.edu
PING kronos.csc.ncsu.edu: 56 data bytes
64 bytes from kronos.csc.ncsu.edu (130.207.8.17): icmp_seq=0 time=47 ms
64 bytes from kronos.csc.ncsu.edu (130.207.8.17): icmp_seq=2 time=47 ms
64 bytes from kronos.csc.ncsu.edu (130.207.8.17): icmp_seq=3 time=48 ms
64 bytes from kronos.csc.ncsu.edu (130.207.8.17): icmp_seq=4 time=38 ms
^C
----kronos.csc.ncsu.edu PING Statistics----
5 packets transmitted, 4 packets received, 20% packet loss
round-trip (ms)  min/avg/max = 38/45/48
```

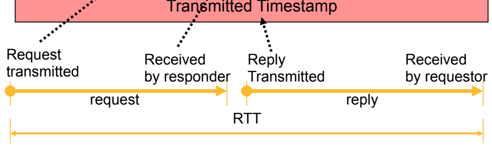

Clock Synchronization

- Each machine maintains its own notion of the current time
  - clocks that differ widely can confuse users of distributed system software
- To synchronize clocks, you need an estimate of round-trip delay
  - simplest technique: ICMP Timestamp Request & Reply messages

#2 Timestamp Request/Reply Messages

0 1 2 3 4 8 16 24 31

Type=13 or 14	Code=0	ICMP Message Checksum
Identifier		Sequence Number
Originating Timestamp		
Received Timestamp		
Transmitted Timestamp		



#2 Timestamp Request/Reply Messages (cont'd)

- Reported in milliseconds since midnight, coordinated universal time (UTC)
- Sending time = request received– request xmitted
- Receiving time = response received– reply xmitted
- RTT = Sending time + Receiving time
  - not affected by synchronization problems (why not?)

RTT Estimation Problems

- Accurate estimation of round-trip delay can be difficult
  - round-trip delays over Internet may have high variance
  - datagrams can be dropped, delivered out of order → taking many measurements may not guarantee consistency
- Alternative 1: Network Time Protocol (RFC 1305)
  - much more sophisticated (and complicated)
  - ms accuracy in LAN/WAN
- Alternative 2: GPS receivers at every node
  - μs accuracy, but cost and other limitations?

#3 Router Discovery (RFC 1256)

- Routers advertise their presence to hosts
  - using either limited broadcast, or a special *multicast* address
- Preference level indicates the "desirability" as a default gateway
- Router Advertisement message

0 1 2 3 4 8 16 24 31

Type=9	Code=0	ICMP Message Checksum
# of router addresses	Address Size	Advertisement lifetime (in seconds)
Router 1 Address		
Router 1 Preference Level		
(…additional router addresses/preference levels)		

#3 Router Discovery (RFC 1256) (cont'd)

- Does not indicate what route a host should use to reach a specific destination!
- Routers periodically broadcast (or multicast) this information
  - time between advertisements roughly every 10 minutes
  - default lifetime is 30 minutes
  - disabling a router interface: advertise with a lifetime of 0
- Hosts can request this information
  - on bootup, host broadcasts a Router Solicitation message

#4 Address Mask Request / Reply (RFC 950)

- Subnet masks needed for classless addressing / routing (we will discuss this later)
- Host sends Subnet Mask Request to its gateway
- ICMP Subnet Mask Reply message contains the 32 bit mask for the subnet from which the request was received

0 1 2 3 4 8 16 24 31

Type=17 or 18	Code=0	ICMP Message Checksum
Identifier		Sequence Number
Address Mask		

ICMP MESSAGE PROCESSING

Processing of ICMP Messages

- ICMP covers a wide range of conditions
- Each message handled differently, e.g…
  - ignored (source quench to UDP)
  - handled by kernel (redirect, source quench to TCP)
  - passed to user process (time exceeded, echo/ timestamp reply)
  - discarded (if no user processes have registered with the kernel to receive ICMP messages)
  - …

Processing of ICMP Messages (cont'd)

Type	Code	Description	Query / Reply / Error	Result / Message
0	0	Echo reply	R	(used by ping)
3	1	Network unreachable	E	application request fails
3	2	Host unreachable	E	application request fails
3	3	Protocol unreachable	E	application request fails
3	5	Fragmentation needed but DF Flag set	E	reduce packet size
3	6	Source route failed	E	respecify route
3	-	Other reasons	E	-

Processing of ICMP Messages (cont'd)

Type	Code	Description	Q / R / E	Result / Message
4	0	Source quench	E	reduction in TCP send rate
5	1	Redirect for host	E	updates routing table
8	0	Echo request	Q	send a reply
9	0	Router advertisement	R	updates routing table
10	0	Router solicitation	Q	send a advertisement

Processing of ICMP Messages (cont'd)

Type	Code	Description	Q / R / E	Result / Message
11	0	Time exceeded (TTL=0)	E	application request fails
12	0	IP header bad	E	?
13	0	Timestamp request	Q	send a reply
14	0	Timestamp reply	R	application calculates RTT
17	0	Address mask request	Q	send a reply
18	0	Address mask reply	R	update mask for interface

Summary

1. ICMP is a "swiss army knife" for lots of problems and small functions
   – common protocol for reporting error conditions
   – also used to query network conditions
2. Some older ICMP functions have been superseded by more powerful, and specialized, protocols
3. ICMP continues to be extended for new purposes

Next Lecture

- User Datagram Protocol (UDP)