**ICMP (INTERNET CONTROL MESSAGE PROTOCOL)**

Ken Wong  
Washington University in St. Louis  
kenw@arl.wustl.edu  
www.arl.wustl.edu/~kenw

---

**INTERNET CONTROL MESSAGE PROTOCOL**

- Allows IP systems to send error and administrative messages
- Required part of any IP implementation
- Usage
  - Errors: Routers report problems (e.g., can’t route datagram; congestion)
  - Queries
    - Defined in request/reply pairs
    - e.g., Hosts test reachability (ping)
- **ICMP is an error reporting (not correction) mechanism**
  - Error message is sent to the datagram source
  - Can not be used to directly inform intermediate routers of a problem; e.g.,
    - Router Rk in path "R1, R2, ... , Rj; Rk" detects a routing problem
    - Rj has a bad routing table ... Rk can only tell R1 there was an error

---

**ICMP MESSAGE DELIVERY**

- An ICMP message is encapsulated in an IP datagram
- Datagram protocol field = 1 => Message is carried in an IP datagram
- Applications send/receive ICMP messages through raw IP interface
- ICMP messages that cause an error are silently dropped

---

**ICMP MESSAGE STRUCTURE**

<table>
<thead>
<tr>
<th>Unreachable, Time Exceeded, Source Quench:</th>
<th>Parameter Problem:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image_url" alt="Diagram" /></td>
<td><img src="image_url" alt="Diagram" /></td>
</tr>
</tbody>
</table>

- **Unreachable, Time Exceeded, Source Quench:**

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Msg Checksum</th>
<th>IP Hdr and Bad Packet</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>At least 28 bytes</td>
</tr>
</tbody>
</table>

- **Need Fragmentation:**

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Msg Checksum</th>
<th>nxtmtu</th>
<th>IP Hdr and Bad Packet</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At least 28 bytes</td>
</tr>
</tbody>
</table>

- **Parameter Problem:**

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Msg Checksum</th>
<th>ptr</th>
<th>IP Hdr and Bad Packet</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>At least 28 bytes</td>
</tr>
</tbody>
</table>
**ICMP QUERY AND ERROR MESSAGE TYPES**

- **Query Message Types**
  - ICMP_ECHO/ECHOREPLY: Echo request/reply: 8/0
  - ICMP_TIMESTAMP/TIMESTAMPREPLY: Timestamp request/reply: 13/14
  - ICMP_MASKREQ/MASKREPLY: Address mask request/reply: 17/18
  - ICMP_INFO/INFORMREPLY: Information request/reply (Obsolete)
  - ICMP_ROUTERADVERT/_ROUTERSOLICIT: Router advertisement/solicitation: 9/10

- **Error Message Types**
  - ICMP_REDIRECT(4 codes): Better route available: 5
  - ICMP_UNREACH(16): Destination unreachable: 3
  - ICMP_TIMEXCEED(2): Time (TTL or reassembly TTL) exceeded: 11
  - ICMP_PARAMPROB(2): Problem with IP header: 12
  - ICMP_SOURCEQUENCH: Request to slow down transmission: 4

---

**ICMP_REDIRECT EXAMPLE**

- Router detects a better route available
- Allows host to have small routing table

---

**ICMP_TIMEXCEED**

- **Time (TTL or reassembly TTL) exceeded**
  - Time-To-Live (TTL) counter (Hop count) in IP header (Code 0)
    - Prevent infinite datagram cycling when there are routing cycles
    - Counter is decremented after IP processing
    - Datagram is discarded when TTL reaches 0
  - **Reassembly TTL** (Code 1)
    - Timer started when first fragment of a multiple-fragment datagram arrives
  - **Message Content**
    - Unused word; IP header and first 8 bytes of datagram data
ICMP_ECHO AND ICMP_ECHOREPLY

- Echo request/reply (ping)
  - Test if destination is reachable/responding
- Request contains an optional data area, identifier (process id), and sequence number
- Reply contains a copy of the request data area, identifier, and sequence number

PING EXAMPLE

```
> ping -s -R -v yahoo.com
64 bytes from yahoo.com (204.71.177.35): icmp_seq=0. time=200. ms
IP options: <record route> ncrc-eng.wustl.edu (128.252.5.133),
  fe0-0.starnet2.starnet.net (199.217.254.193),
  st11.vcp.verio.net (129.250.16.98),
  st10.st11.verio.net (129.250.2.214),
  dfw2.st10.verio.net (129.250.2.218),
  iad3.dfw2.verio.net (129.250.2.209),
  iad3.verio.net (129.250.0.23),
  verio.s0-0-1.br1.DCA.globalcenter.net (206.165.113.146),
  pos3-0-0-155M.br1.DCA.globalcenter.net (204.152.166.5)
64 bytes from yahoo.com (204.71.177.35): icmp_seq=1. time=127. ms
IP options: <record route> ncrc-eng.wustl.edu (128.252.5.133),
... Repeated ...
```

TRACEROUTE

- Uses UDP, ICMP and TTL field in IP header
  - Recommended TTL = 64, but some set as high as 255
- Each router along path decrements TTL by 1 or number of seconds it holds datagram
- TTL prevents infinite loops
- When TTL = 0, router returns ICMP "time exceeded" error and router IP address to source
- Traceroute Operation
  - Send UDP datagram to unlikely port at destination with TTL = 1, 2, 3, ...
  - Discover routers along path as ICMP message returns

TRACEROUTE EXAMPLE

```
traceroute to yahoo.com (204.71.177.35), 30 hops max, 40 byte packets
1 gateway.cs.wustl.edu (128.252.165.249) 1.573 ms 0.985 ms 0.986
2 wustl-fddi-starnet.wustl.edu (128.252.5.254) 2.459 ms 2.045 ms
3 fe0-0.starnet1.starnet.net (199.217.254.194) 2.747 ms 2.223 ms
4 vcp.st11.verio.net (129.250.16.97) 2.906 ms 2.243 ms 3.179 ms
5 st11.st10.verio.net (129.250.2.213) 3.080 ms 2.736 ms 2.990 ms
6 st10.dfw2.verio.net (129.250.2.217) 20.986 ms 19.754 ms 20.199
7 dfw2.iad3.verio.net (129.250.2.210) 65.729 ms 63.791 ms 64.099
8 iad3.iad0.verio.net (129.250.2.177) 64.419 ms 64.609 ms 63.755
... 26 pos1-0-622M/cr1.NUQ.globalcenter.net (206.251.0.73)
     141.579 ms 148.512 ms 137.012
27 pos5-0-0-155M/hr1.NUQ.globalcenter.net (206.251.0.121)
     137.256 ms 137.226 ms 124.934
28 yahoo.com (204.71.177.35) 129.703 ms 126.576 ms 137.004 ms
```