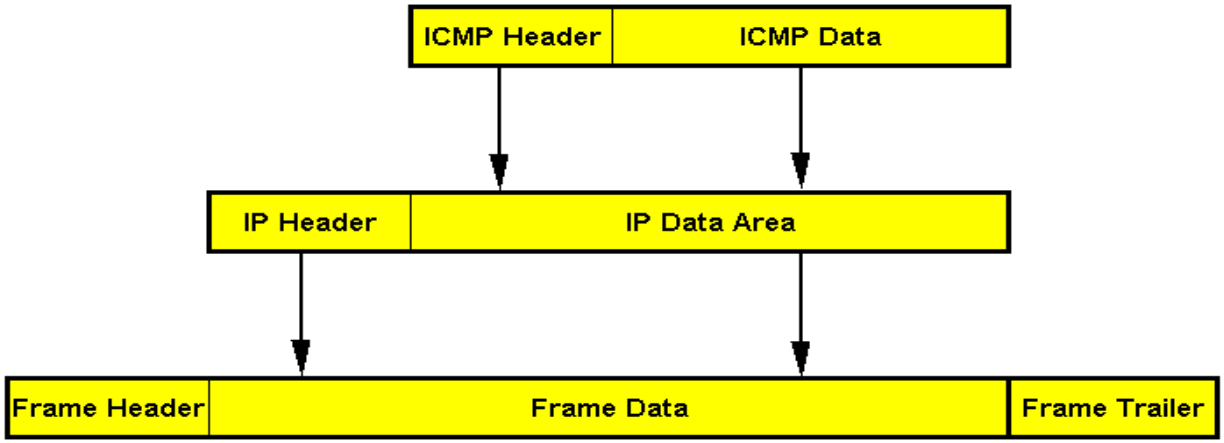
Internet Control Message Protocol ICMP



The Internet Control Message Protocol (ICMP) is a control protocol that is considered to be an integral part of IP, although it is architecturally layered upon IP - it uses IP to carry its data end-to-end. ICMP provides error reporting, congestion reporting, and first-hop router redirection.

Introduction

- IP provides best-effort delivery
- Delivery problems can be ignored; datagrams can be "dropped on the floor"
- Internet Control Message Protocol (ICMP) provides error-reporting mechanism

IP and ICMP

Application Presentation	FTP	Telnet	SMTP	HTTP	Ping	DNS
Session	SSL					
Transport	TCP		UDP		ICMP	
Network	IP					
Datalink	LLC	HDLC	PPP	LAP-B	LAP-F	LAP-D
	Ethernet	Token Ring	FDDI	ATM	DQDB	Frame Relay
Physical	Optical Fiber	UTP	Coaxial Cable	Microwave	Satellite	STP

ICMP Features

- ICMP uses IP as if ICMP were a higher-level protocol (that is, ICMP messages are encapsulated in IP datagrams). However, ICMP is an integral part of IP and must be implemented by every IP module.
- ICMP is used to report some errors, not to make IP reliable. Datagrams may still be undelivered without any report on their loss. Reliability must be implemented by the higher-level protocols that use IP.

ICMP Features

- ICMP can report errors on any IP datagram with the exception of ICMP messages, to avoid infinite repetitions.
- For fragmented IP datagrams, ICMP messages are only sent about errors on fragment zero. That is, ICMP messages never refer to an IP datagram with a non-zero fragment offset field.

ICMP Features

- ICMP has rules regarding error message generation to prevent *broadcast storms*
- ICMP messages are never sent in response to datagrams with a destination IP address that is a broadcast or a multicast address.
- ICMP messages are never sent in response to a datagram which does not have a source IP address which represents a unique host. That is, the source address cannot be zero, a loopback address, a broadcast address or a multicast address.

Error Message Generation Rules

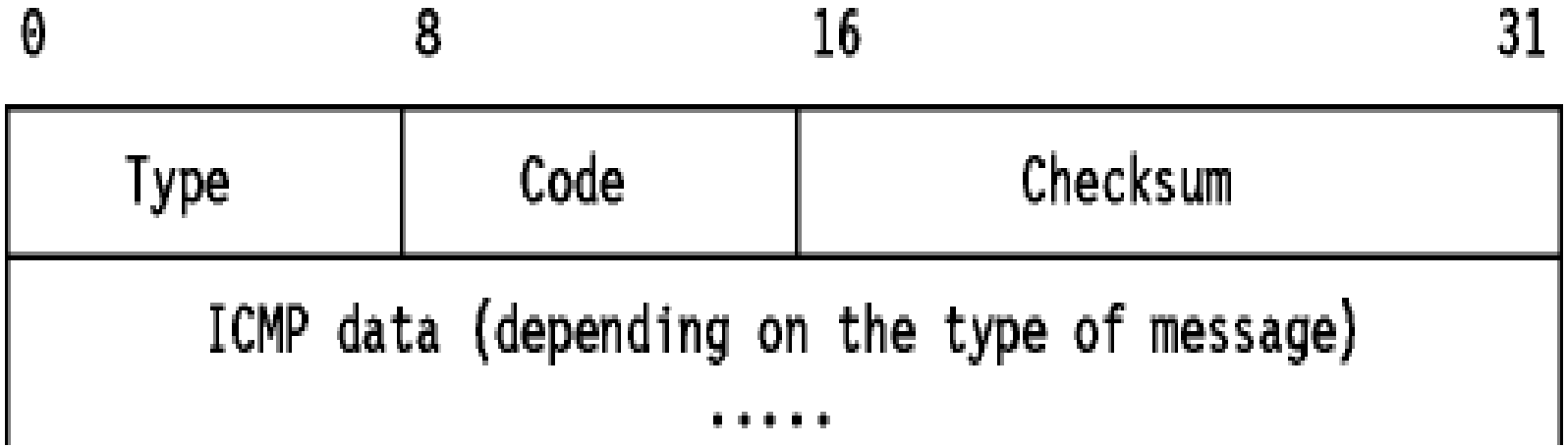
- ICMP errors messages are not generated in response to
  - an ICMP error message
  - datagrams destined to an IP broadcast address
  - datagrams sent as a link-layer broadcast
  - a fragment other than the first
  - a datagram whose source address does not define a single host

ICMP Features

- ICMP messages are never sent in response to ICMP error messages. They may be sent in response to ICMP query messages (ICMP types 0, 8, 9, 10 and 13 through 18).
- RFC 792 states that ICMP messages "may" be generated to report IP datagram processing errors, not "must". In practice, routers will almost always generate ICMP messages for errors, but for destination hosts, the number of ICMP messages generated is implementation dependent.

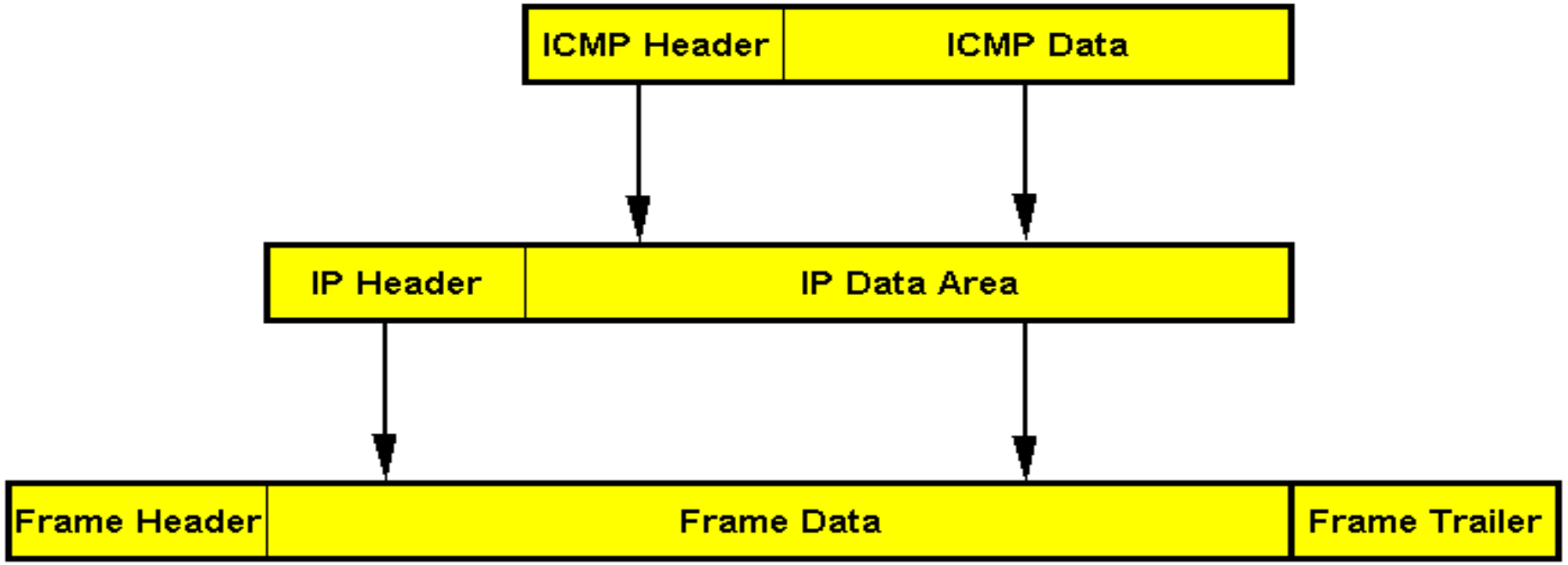
ICMP Message Format

- ICMP messages are described in RFC 792 and RFC 950, belong to STD 5 and are mandatory.
- ICMP messages are sent in IP datagrams. The IP header will always have a Protocol number of 1, indicating ICMP and a type of service of zero (routine). The IP data field will contain the actual ICMP message in the format shown in the figure below:



ICMP Message Transport

- ICMP encapsulated in IP
- But ... how can that work?
- ICMP messages sent in response to incoming datagrams with problems



Error Detection

- Internet layer can detect a variety of errors:
  - Checksum (header only!)
  - TTL expires
  - No route to destination network
  - Can't deliver to destination host (e.g., no ARP reply)
- Internet layer discards datagrams with problems
- Some - e.g., checksum error - can't trigger error messages

Types of Messages

- ICMP defines two types of messages: error and informational messages
- Error messages:
  - Source quench
  - Time exceeded
  - Destination unreachable
  - Redirect
  - Fragmentation required
- Informational messages:
  - Echo request/reply
  - Address mask request/reply
  - Router discovery

ICMP: Message Types

Type	Message
0	Echo reply
3	Destination unreachable
4	Source quench
5	Redirect
8	Echo request
11	Time exceeded
12	Parameter unintelligible
13	Time-stamp request
14	Time-stamp reply
15	Information request
16	Information reply
17	Address mask request
18	Address mask reply

ICMP Message Types

Type	Code	Description	Query	Error
0	0	Echo reply	•	
3		Destination unreachable:		
	0	Network unreachable		•
	1	Host unreachable		•
	2	Protocol unreachable		•
	3	Port unreachable		•
	4	Fragmentation needed		•
	5	Source route failed		•
	6	Destination network unknown		•
	7	Destination host unknown		•
	8	Source host isolated		•
	9	Destination net prohibited		•
	10	Destination host prohibited		•
	11	Network unreachable for TOS		•
	12	Host unreachable for TOS		•
	13	Communication prohibited		•
	14	Host precedence violation		•
	15	Precedence cutoff in effect		•
4	0	Source quench		•

Type	Code	Description	Query	Error
5		Redirect		
	0	Redirect for network		•
	1	Redirect for host		•
	2	Redirect for TOS and Net		•
	3	Redirect for TOS and Host		•
8	0	Echo request	•	
9	0	Router advertisement	•	
10	0	Router solicitation	•	
11		Time exceeded		
	0	TTL equals 0 during transit		•
	1	TTL equals 0 during reassembly		•
12		Parameter problem		
	0	IP header bad		•
	1	Required option missing		•
13	0	Timestamp request	•	
14	0	Timestamp reply	•	
15	0	Information request	•	
16	0	Information reply	•	
17	0	Address mask request	•	
18	0	Address mask reply	•	

ICMP and Reachability

- An internet host, A, is reachable from another host, B, if datagrams can be delivered from A to B
- ping program tests reachability - sends datagram from B to A that A echoes back to B
- Uses ICMP echo request and echo reply messages
- Internet layer includes code to reply to incoming ICMP echo request messages

Destination Unreachable Codes

Code	Meaning
0	Network unreachable
1	Host unreachable
2	Protocol unreachable
3	Port unreachable
4	Fragmentation need and don’t fragment bit set
5	Source route failed
6	Destination network unknown
7	Destination host unknown
8	Source host isolated
9	Communication with dest net administratively prohibited
10	Communication with dest host administratively prohibited
11	Network unreachable for type of service
12	Host unreachable for type of service

ICMP Unreachable Error

- Unreachable errors are generate for a number of reasons
  - network unreachable
  - host unreachable

type (3)	code (0-15)	16-bit checksum
unused (must be 0)		
IP header (including options) + first 8 bytes of IP datagram data		

Handling of ICMP Messages

Type	*Code*	*Description*	*Handled by*
0	0	Echo reply	User process
3		Destination unreachable:	
	0	Network unreachable	"No route to host"
	1	Host unreachable	"No route to host"
	2	Protocol unreachable	"Connection refused"
	3	Port unreachable	"Connection refused"
	4	Fragmentation needed	"Message too long"
	5	Source route failed	"No route to host"
	6	Destination network unknown	"Network is unreachable"
	7	Destination host unknown	"No route to host"
	8	Source host isolated	"No route to host"
	9	Destination net prohibited	"Network is unreachable"
	10	Destination host prohibited	"No route to host"
	11	Network unreachable for TOS	"Network is unreachable"
	12	Host unreachable for TOS	"No route to host"
	13	Communication prohibited	(ignored)
	14	Host precedence violation	(ignored)
	15	Precedence cutoff in effect	(ignored)
4	0	Source quench	Kernel for TCP; ignored by UDP

Handling of ICMP Messages

Type	*Code*	*Description*	*Handled by*
5		Redirect	
	0	Redirect for network	Kernel updates routing table
	1	Redirect for host	Kernel updates routing table
	2	Redirect for TOS and Net	Kernel updates routing table
	3	Redirect for TOS and Host	Kernel updates routing table
8	0	Echo request	Kernel generates reply
9	0	Router advertisement	User process
10	0	Router solicitation	User process
11		Time exceeded	
	0	TTL equals 0 during transit	User process
	1	TTL equals 0 during reassembly	User process
12		Parameter problem	
	0	IP header bad	“Protocol not available”
	1	Required option missing	“Protocol not available”
13	0	Timestamp request	Kernel generates reply
14	0	Timestamp reply	User process
15	0	Information request	Kernel generates reply
16	0	Information reply	User process
17	0	Address mask request	Kernel generates reply
18	0	Address mask reply	User process

Ping Program

- Ping stands for "Packet InterNet Groper"
- The ping program tests whether another host is reachable
- The program works by sending an ICMP echo request to a host, expecting an ICMP echo reply to be returned
- Normally ping is used as a diagnostic tool to test network connectivity

Client/Server Programs

- Ping is is an example of a client/server program
  - the client sends the request to a server
  - the server returns the reply
- Most TCP/IP implementations support the ping server directly in the kernel

Ping Implementation

- Unix implementations set the identifier field to the process ID of the sender
- The sequence number starts at 0 and is incremented every time a new echo request is sent
- Ping operates in one of two modes
  - send a single request, if a response is received the host is alive
  - send one request every second
- Ping in Action
- IP Record Route Option

ICMP and Internet Routes

- List of all routers on path from A to B is called the route from A to B
- traceroute uses UDP to non-existent port and TTL field to find route via expanding ring search
- Sends ICMP echo messages with increasing TTL
  - Router that decrements TTL to 0 sends ICMP time exceeded message, with router's address as source address
  - First, with TTL 1, gets to first router, which discards and sends time exceeded message
  - Next, with TTL 2, gets through first router to second router
  - Continue until message from destination received
- traceroute must accommodate varying network delays
- Must also accommodate dynamically changing routes

ICMP and Path MTU Discovery

- Fragmentation should be avoided
- How can source configure outgoing datagrams to avoid fragmentation?
- Source determines path MTU - smallest network MTU on path from source to destination
- Source probes path using IP datagrams with don't fragment flag
- Router responds with ICMP fragmentation required message
- Source sends smaller probes until destination reached

Information Request/Reply:

- This request is intended for a diskless system to obtain its subnet mask
- Set source and destination addresses to 0 in the request and broadcast
- Server replies back with your IP address
- (Not used. Replaced by RARP and BOOTP)
- Address Mask Request/Reply: What is the subnet mask on this net? Replied by "Address mask agent"

type (17 or 18)	code (0)	16-bit checksum
identifier (can be set to anything)		sequence (can be set to anything)
32-bit subnet mask		

ICMP Messages: Fragmentation Required

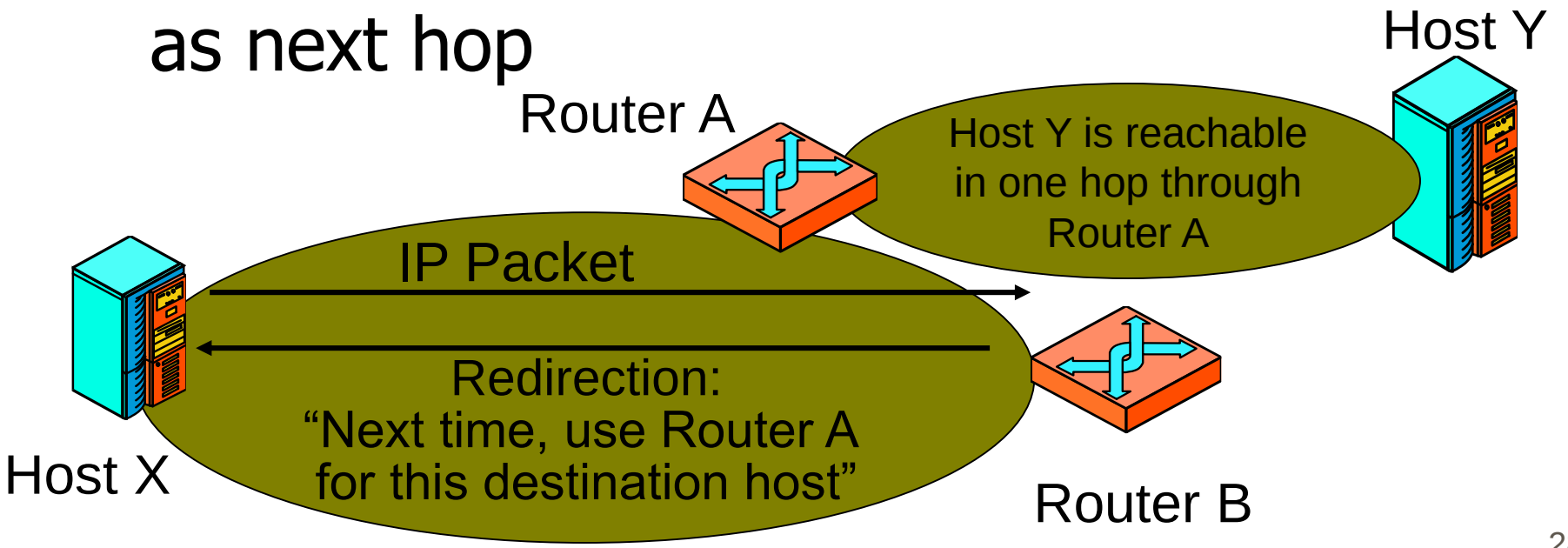
- Fragmentation Required: Datagram was longer than MTU and "No Fragment bit" was set.

ICMP Messages: Time Exceeded

- Time Exceeded: Time to live field in one of your packets became zero.” or “Reassembly timer expired at the destination.

ICMP Messages: Redirect

- Default route may cause extra hop
- Router that forwards datagram on same interface sends ICMP redirect
- Host installs new route with correct router as next hop



ICMP Messages: Redirect

- Redirect: Please send to router X instead of me.
- 0 = Redirect datagrams for the network
- 1 = Redirect datagrams for the host
- 2 = Redirect datagrams for the type of service and net
- 3 = Redirect datagrams for the type of service and host

ICMP Timestamp Request & Reply

- Used to return the current time from another host
- Could be used for (primitive) time synchronization protocol, but NTP and XNTP do a much better job

type (13 or 14)	code (0)	16-bit checksum
identifier (can be set to anything)		sequence (can be set to anything)
32-bit originate timestamp		
32-bit receive timestamp		
32-bit transmit timestamp		

ICMP Timestamp Request & Reply

- The recommended value to be returned is the number of milliseconds since midnight, Coordinated Universal Time (UTC).
- A drawback is that only the time since midnight is returned. The caller must know the date form some other means

ICMP Summary

- Internet layer provides best-effort delivery service
- May choose to report errors for some problems
- ICMP provides error message service
- ICMP is the control sibling of IP
- ICMP is used by IP and uses IP as network layer protocol - Encapsulated in IP datagram - Not reliable
- Feedback about problems
  - e.g. time to live expired
- ICMP is used for ping, traceroute, and path MTU discovery
- Transfer of (control) messages from routers and hosts to hosts