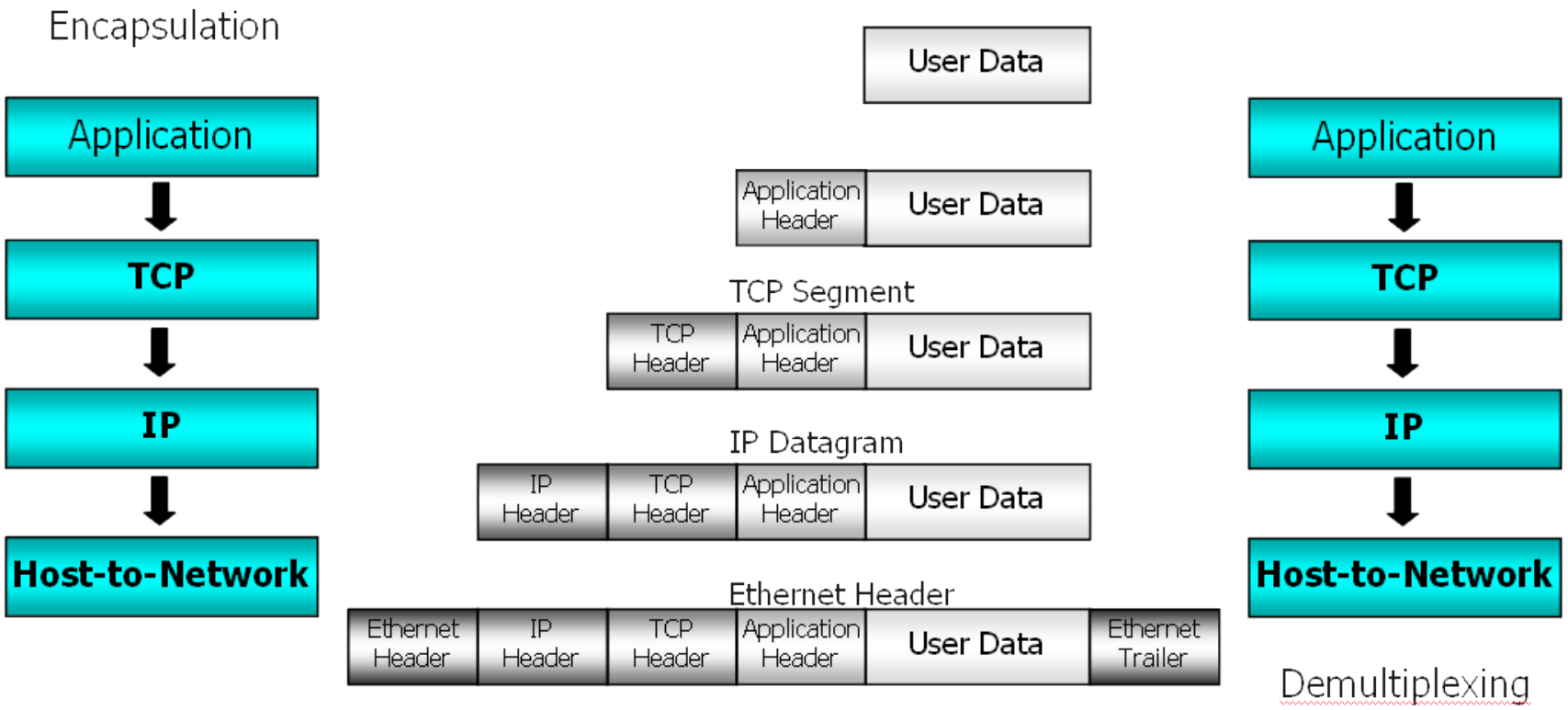
Review IP

Dr. Pipat Sookavatana

Encapsulation Layer



- http://www.us-cert.gov/cas/techalerts/TA04-111A.html
- **ww.cisco.com/warp/public/105/pmtud_ipfrag.html**
- http://erg.abdn.ac.uk/users/gorry/course/inet-pages/ip-cksum.html
- http://www.webclasses.net/Courses/Protocols/7.0/DemoBuild/units/unit01/sec06a.html

TCP/IP Structure

Host to Network Layer (Physical + Data Link)

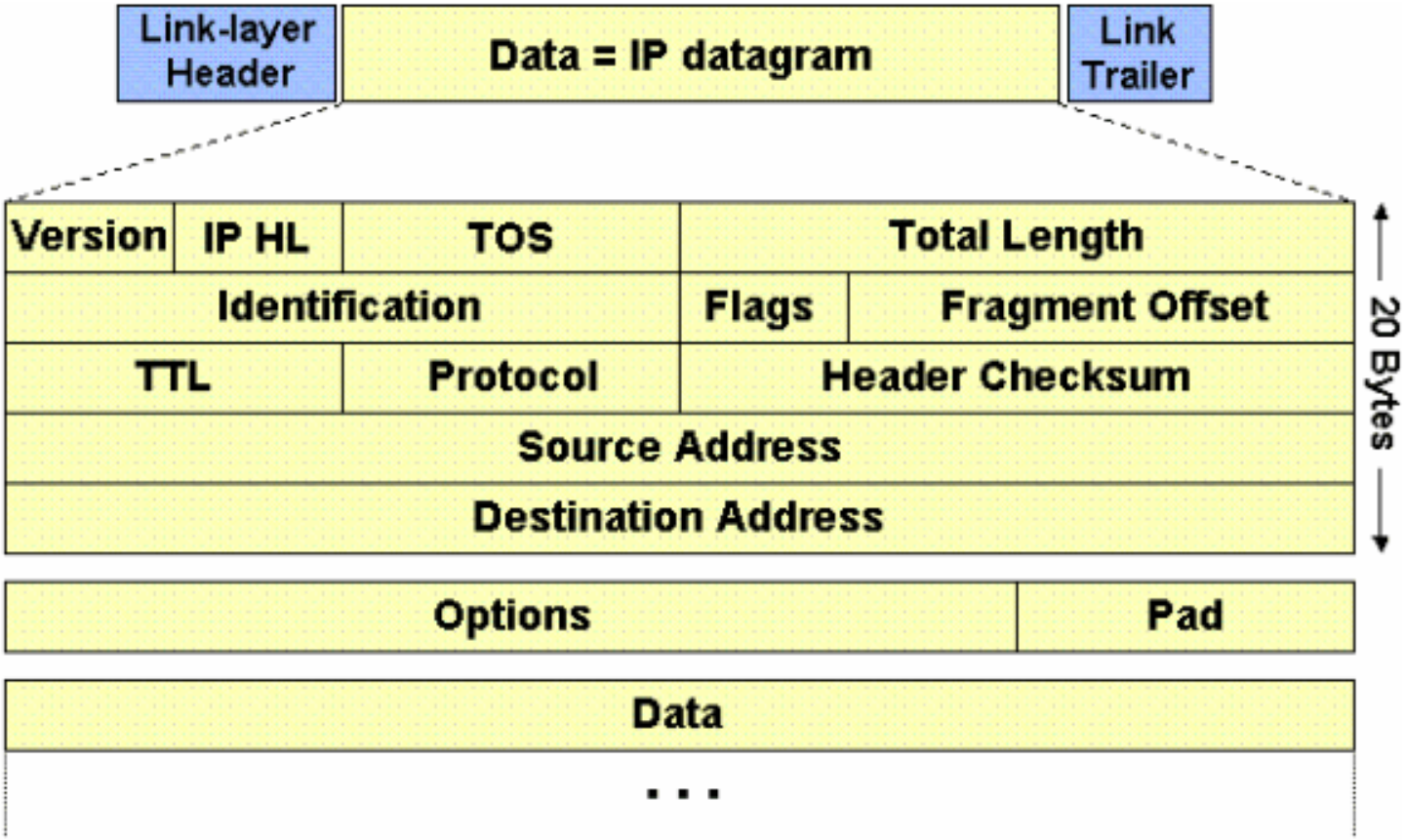
- โพรโตคอลสำหรับการควบคุมการสื่อสารในชั้นนี้เป็นสิ่งที่ไม่มีการกำหนดรายละเอียดอย่างเป็นทางการ หน้าที่หลักคือการรับข้อมูลจากชั้นสื่อสาร IP มาแล้วส่งไปยังโหนดที่ระบุไว้ในเส้นทางเดินข้อมูลทางด้านผู้รับก็จะทำงานในทางกลับกัน คือรับข้อมูลจากสายสื่อสารแล้วนำส่งให้กับโปรแกรมในชั้นสื่อสาร

Internet Layer (Network Layer)

- Packet Switching Network
  - เป็นการติดต่อแบบ Connectionless หลักการทำงานคือการปล่อยให้ข้อมูลขนาดเล็กที่เรียกว่า Packet สามารถไหลจากโหนดผู้ส่งไปตามโหนดต่างๆ ในระบบจนถึงจุดหมายปลายทางได้โดยอิสระ หากว่ามีการส่งแพ็กเก็ตออกมาเป็นชุดโดยมีจุดหมายปลายทางเดียวกันในระหว่างการเดินทางในเครือข่าย แพ็กเก็ตแต่ละตัวในชุดนี้ก็จะเป็นอิสระแก่กันและกัน ดังนั้น แพ็กเก็ตที่ส่งไปถึงปลายทางอาจจะไม่เป็นไปตามลำดับก็ได้
  - IP Protocol (RFC: 791)

- IP Protocol (RFC 791)

- Internet Protocol
  - Addressing
  - Routing (Source to Destination)
  - Fragmentation for difference (MTU)


Link-layer Header
Data = IP datagram
Link Trailer
Version
IP HL
TOS
Total Length
Identification
Flags
Fragment Offset
TTL
Protocol
Header Checksum
Source Address
Destination Address
20 Bytes
Options
Pad
Data
. . .

Controller

- Version: 4 bits
  - The Version field indicates the format of the internet header.
- IHL: 4 bits
  - Internet Header Length is the length of the internet header in 32 bit words, and thus points to the beginning of the data. Note that the minimum value for a correct header is 5 (5 Row at 32 bits each)
  - **TOS: 8 bits**
    - Type of Service ใช้เป็นข้อมูลสำหรับเราเตอร์ในการตัดสินใจเลือกการเราต์ข้อมูลในแต่ละดาต้าแกรม แต่ในปัจจุบันไม่ได้มีการนำไปใช้งานแล้ว (Usually for QoS used)

- Total Length: 16 bits
  - ความยาวทั้งหมดของ Datagram อันนี้ มีขนาดเป็น Octets
  - Max 65535 octets (16 bits)
  - ในการส่งข้อมูลจริง ข้อมูลจะถูกแยก (Fragment) เป็นส่วนๆตามขนาดของ MTU (Maximum Transfer Unit) ที่กำหนดในลิงค์เลเยอร์ และนำมารวมกันอีกครั้งเมื่อส่งถึงปลายทาง
  - Host โดยส่วนใหญ่จะรองรับ MTU ที่ 576 Bytes โดยแบ่งเป็น *Data Block 512 Bytes* และ Header Block 64 Bytes

- Identification: 16 bits
  - เป็นหมายเลขของดาต้าแกรมในกรณีที่มีการแยกดาต้าแกรมเมื่อข้อมูลส่งถึงปลายทาง จะนำข้อมูลที่มี identification เดียวกันมารวมกัน

- Flags: 3 bits
  - Various Control Flags.
    - Bit 0: reserved, must be zero
    - Bit 1: (DF) 0 = May Fragment, 1 = Don't Fragment.
    - Bit 2: (MF) 0 = Last Fragment, 1 = More Fragments.

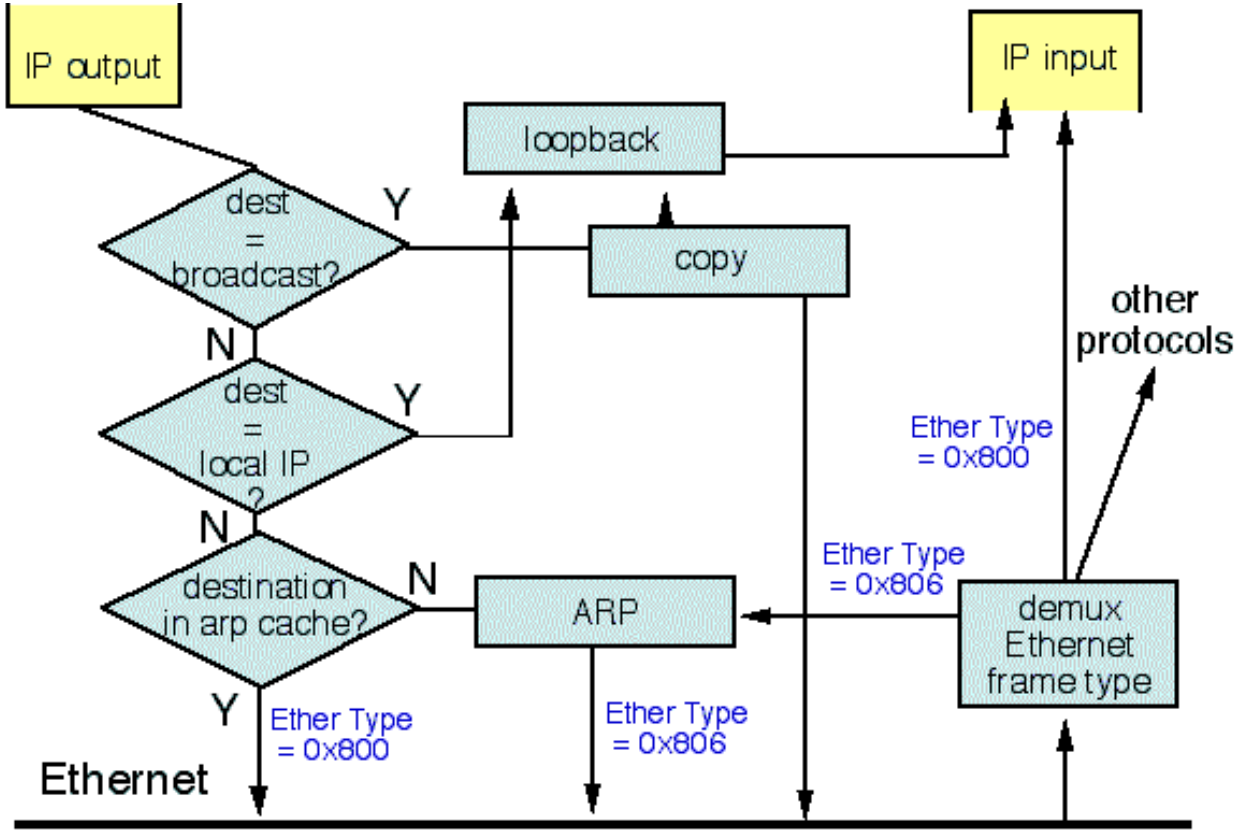
0	1	2
0	DF	MF

- Fragment Offset: 13 bits
  - ใช้ในการกำหนดตำแหน่งข้อมูลใน Datagram ที่มีการแยกส่วน เพื่อให้สามารถนำกลับมาเรียงต่อกันได้อย่างถูกต้อง
- Time to Live: 8 bits
  - จำนวน Hop ที่ Datagram สามารถท่องไปใน ระบบ Network
  - ดูรายระเอียดจาก CPEN1330 Computer Network

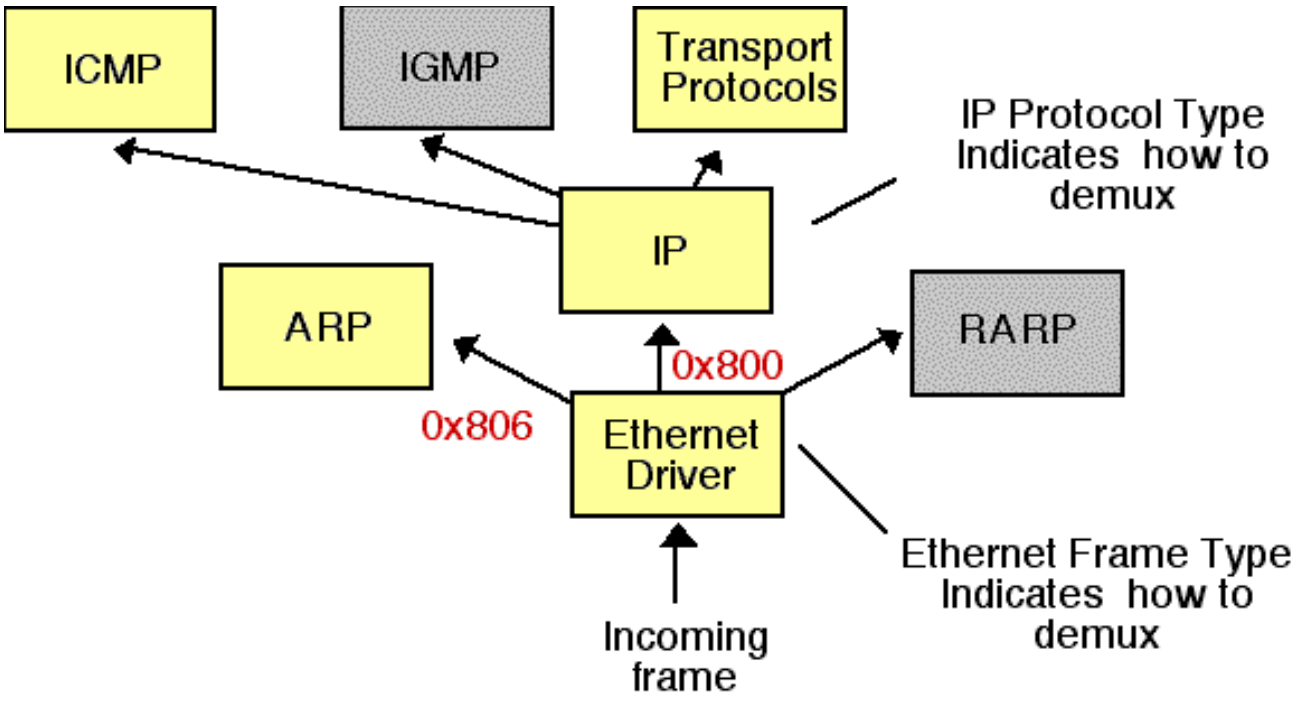
- Protocol: 8 bits
  - Protocol ที่ใช้ Datagram นี้ ถูกอ้างอิงใน
    - *Postel, J., "Assigned Numbers," **RFC 790**, USC/Information Sciences Institute, September 1981.*
  - Octate 1 → ICMP [53,JBP]
  - Octate 6 →TCP [34,JBP]
  - Octate 21 (17 Dec)→ User Datagram [42,JBP]
- Header Checksum: 16 bits
  - อ่านเพิ่มเติมที่ CheckSumIP (Download at berry)

- **Source IP address** : หมายเลข IP ของผู้ส่งข้อมูล
- **Destination IP address** : หมายเลข IP ของผู้รับข้อมูล
- **Data** : ข้อมูลจากโปรโตคอลระดับบน

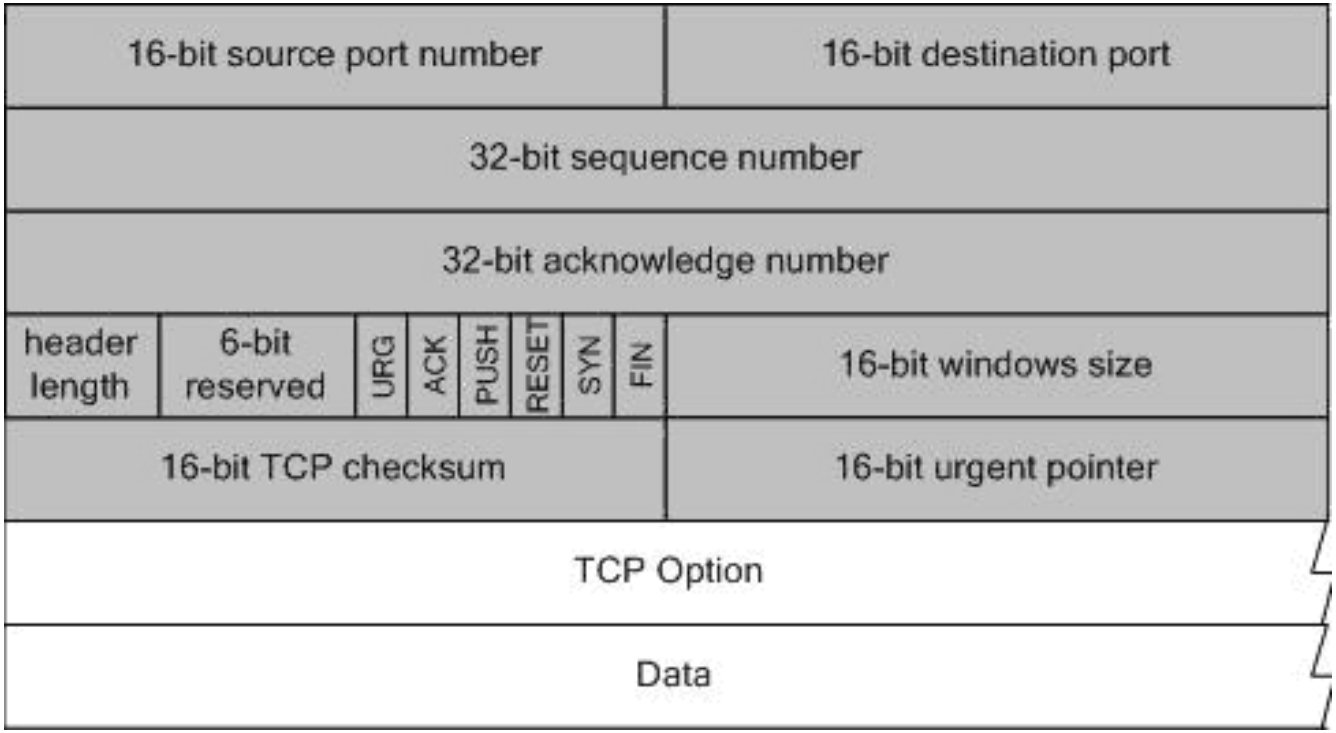
IP Packet Processing
Transmission of a frame over Ethernet



Reception of a frame from Ethernet



TCP : (Transmission Control Protocol) RFC 793



Port	Protocol	Use
21	FTP	File transfer
23	Telnet	Remote login
25	SMTP	E-mail
69	TFTP	Trivial File Transfer Protocol
79	Finger	Lookup info about a user
80	HTTP	World Wide Web
110	POP-3	Remote e-mail access
119	NNTP	USENET news

- **Source Port Number** :
  - หมายเลขพอร์ตต้นทางที่ส่งดาต้าแกรมนี้ โดยทั่วไปพอร์ตนี้จะเรียกว่า "ไคลเอนต์พอร์ต" คือพอร์ตที่ไคลเอนต์เปิดขึ้น มาเพื่อรอการตอบรับจากเซิร์ฟเวอร์ (Ephenumeral port))
- **Destination Port Number** :
  - หมายเลขพอร์ตปลายทางที่จะเป็นผู้รับดาต้าแกรม (Service port)
- **Sequence Number** :
  - ฟิลด์ที่ระบุหมายเลขลำดับอ้างอิงในการสื่อสารข้อมูลแต่ละครั้ง เพื่อใช้ในการแยกแยะว่าเป็นข้อมูลของชุดใด และนำมาจัดลำดับได้ถูกต้อง
- **Acknowledgment Number** :
  - ทำหน้าที่เช่นเดียวกับ Sequence Number แต่จะใช้ในการตอบรับ
- **Header Length** :
  - โดยปกติความยาวของเฮดเดอร์ TCP จะมีความยาว 20 ไบต์ แต่อาจจะมากกว่านั้น ถ้ามีข้อมูลในฟิลด์ option แต่ต้องไม่เกิน 60 ไบต์
- **Flag** :
  - เป็นข้อมูลระดับบิตที่อยู่ในเฮดเดอร์ TCP โดยใช้เป็นตัวบอกคุณสมบัติของแพ็กเก็ต TCP ขณะนั้นๆ และใช้เป็นตัวควบคุมจังหวะการรับส่งข้อมูลด้วย ซึ่ง Flag มีอยู่ทั้งหมด 6 บิต แบ่งได้ดังนี้

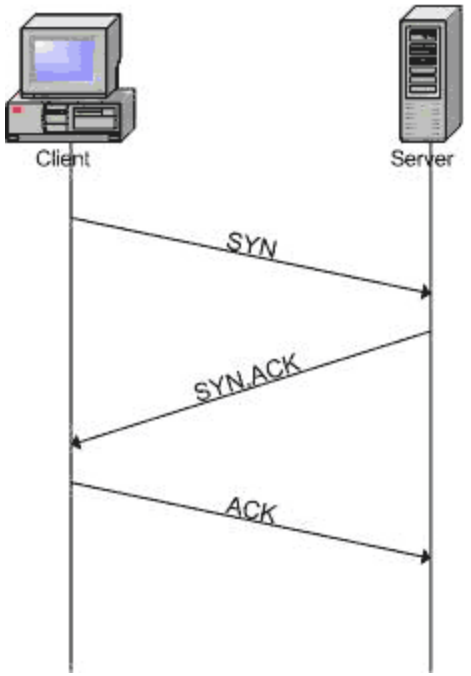
Flag	Details
URG	ใช้บอกความหมายว่าเป็นข้อมูลด่วน และมีข้อมูลพิเศษมาด้วย (อยู่ใน Urgent pointer)
ACK	แสดงว่าข้อมูลในฟิลด์ Acknowledge Number นำมาใช้งานได้
PSH	เพื่อแจ้งให้ผู้รับข้อมูลทราบว่า ควรจะส่งข้อมูล Segment นี้ไปยังโพรเซสที่กำลังรออยู่ทันที
RST	Reset the connection
SYN	Synchronize sequence numbers
FIN	ใช้ส่งเพื่อแจ้งให้ปลายทางทราบว่ายุติการติดต่อ

- Window Size
  - ขนาดของการรับ - ส่งข้อมูลในแต่ละครั้งที่ทางฝ่ายผู้รับจะสามารถรับได้ เนื่องจากในการรับข้อมูลนั้น ทางผู้รับจะต้องจัดเตรียมหน่วยความจำในการพักข้อมูลที่มากจาก TCP และทำการ Demultiplex ออกมา หากไม่มีการตกลง ถึงขนาดที่ทางฝ่ายรับสามารถรับได้ ก็จะทำให้การสื่อสารข้อมูลไม่สมดุล และฝ่ายรับอาจจะประมวลผลทัน ซึ่งจะส่งผลให้ต้องส่ง ข้อมูลซ้ำหลายครั้ง

- Checksum
  - ฟิลด์ที่ใช้ในการตรวจสอบความถูกต้องของข้อมูลใน TCP เซกเมนต์ใช้ระบุหมายเลข Sequence Number ของ TCP เซกเมนต์ล่าสุดที่อยู่ในโหมด Urgent
- Urgent Pointer
  - ข้อมูลเพิ่มเติมซึ่งจะอยู่ใน TCP Header เมื่อมีการตั้งค่า option บางอย่างที่ต้องการข้อมูลเพิ่มเติมซึ่งไม่มีใน TCP Header เช่น MSS, Strict Route

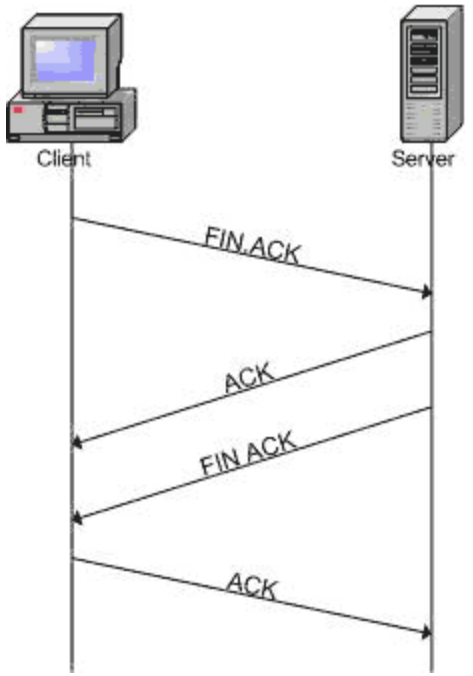
TCP Connection Establishment (3 Ways Handshake)

1. ไคลเอนต์จะทำการส่ง SYN Flag ระบุหมายเลขพอร์ตที่ต้องการติดต่อบนเซิร์ฟเวอร์และระบุหมายเลข ลำดับของข้อมูล (ISN - Initial Sequence Number)
2. เครื่องเซิร์ฟเวอร์เมื่อได้รับข้อมูลเซกเมนต์จากข้อ 1 ก็จะตอบกลับด้วยการเพิ่มค่า ISN ที่ได้รับขึ้นอีก 1 พร้อมทั้งระบุหมายเลขลำดับ (ISN) ของตนเอง และเปิด SYN กับ ACK Flag
3. ไคล์เอนต์เมื่อได้รับการตอบกลับจากเซิร์ฟเวอร์ตามข้อ 2 ก็จะทำการตอบรับกลับไป โดยการเพิ่มค่า ISN ของเซิร์ฟเวอร์ขึ้นอีก 1 และเปิด ACK Flag เมื่อผ่านการสร้าง connection ทั้ง 3 ขั้นตอนแล้ว ตอนนี้ทั้งไคล์เอนต์ และเซิร์ฟเวอร์เปรียบเสมือนมีการเชื่อมต่อถึงกันแล้ว สถานะของการเชื่อมต่อในขณะนี้เรียกว่า **Established**



Connection Termination

1. ไคลเอนต์ทำการส่ง FIN ACK Flag ไปยังเซิร์ฟเวอร์
2. เซิร์ฟเวอร์ทำการ ACK Flag พร้อมกับ FIN ACK Flag ไปยังไคลเอนต์
3. ไคลเอนต์ทำการตอบรับ พร้อมส่ง ACK Flag ไปยังเซิร์ฟเวอร์



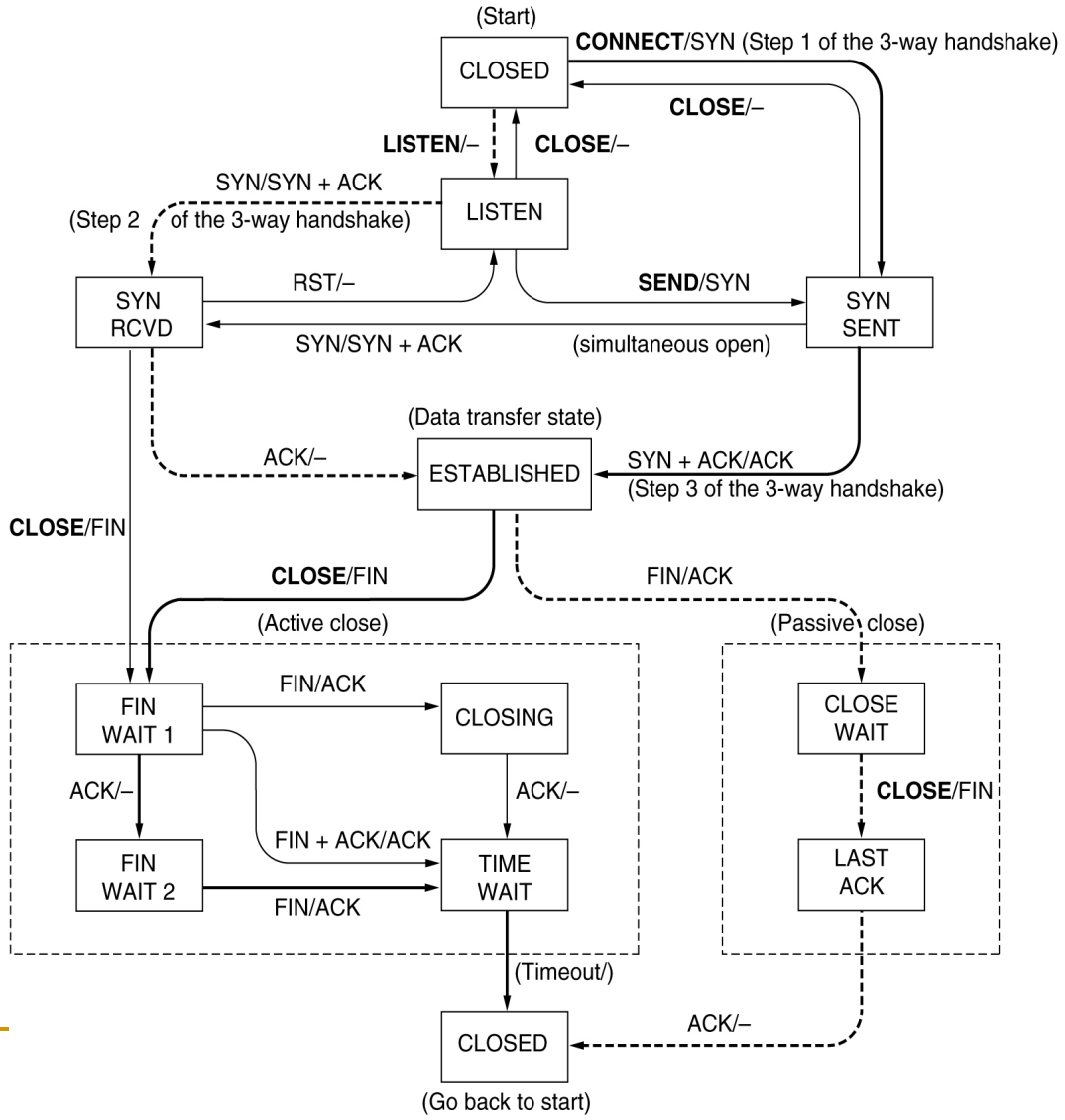
- ในการใช้งานจริง อาจมีการยุติการสื่อสารเพียงด้านเดียว คือหยุดส่งข้อมูล แต่ยังคงเปิดพอร์ตไว้รอรับข้อมูลจากอีกด้านหนึ่ง ทั้งนี้ขึ้นอยู่กับลักษณะการใช้งาน การปิดพอร์ตสื่อสารเพียงด้านเดียวเช่นนี้ เรียกว่า **Half-Close**

TCP Connection Management Modeling

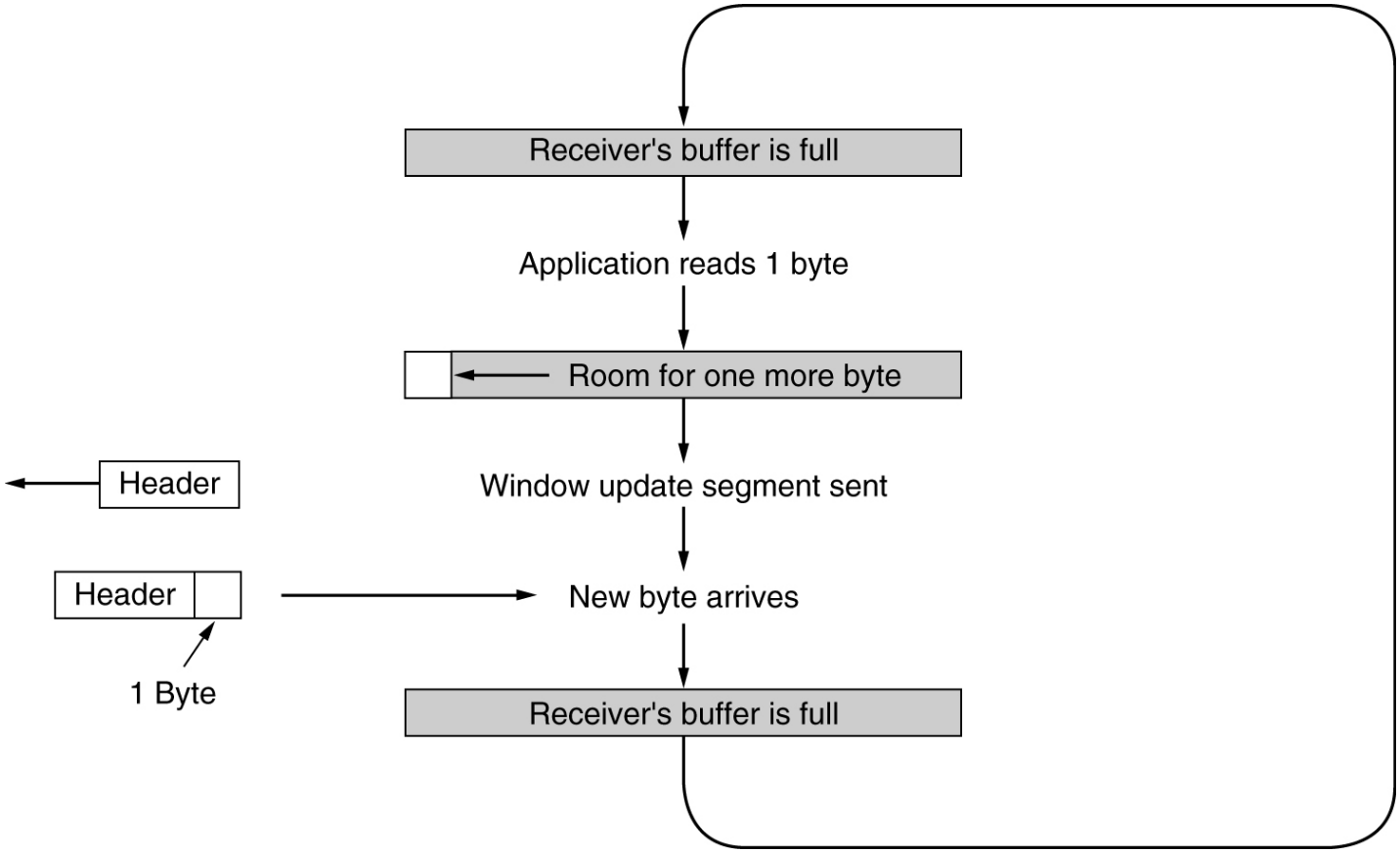
State	Description
CLOSED	No connection is active or pending
LISTEN	The server is waiting for an incoming call
SYN RCVD	A connection request has arrived; wait for ACK
SYN SENT	The application has started to open a connection
ESTABLISHED	The normal data transfer state
FIN WAIT 1	The application has said it is finished
FIN WAIT 2	The other side has agreed to release
TIMED WAIT	Wait for all packets to die off
CLOSING	Both sides have tried to close simultaneously
CLOSE WAIT	The other side has initiated a release
LAST ACK	Wait for all packets to die off

TCP Connection Management Modeling (2)

TCP connection management finite state machine. The heavy solid line is the normal path for a client. The heavy dashed line is the normal path for a server. The light lines are unusual events. Each transition is labeled by the event causing it and the action resulting from it, separated by a slash.



TCP Transmission Policy (2)

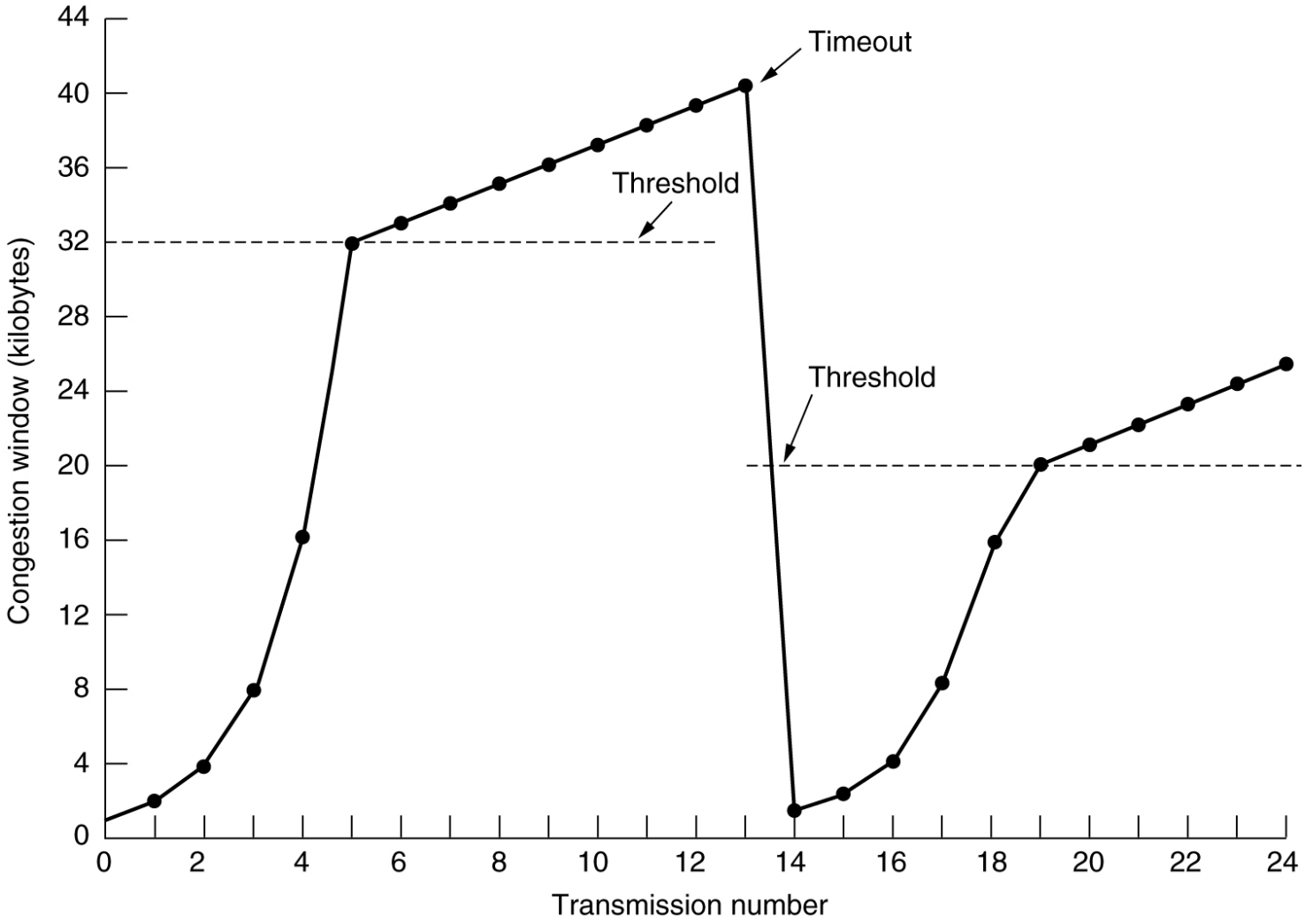


Silly window syndrome.

TCP Transmission Policy (2)

- Clark's solution and Nagle's algorithm
  - แก้ปัญหา Sender ส่งข้อมูล TCP ครั้งละ 1 byte (เนื่องจาก Silly window syndrome)

TCP Congestion Control



An example of the Internet congestion algorithm.

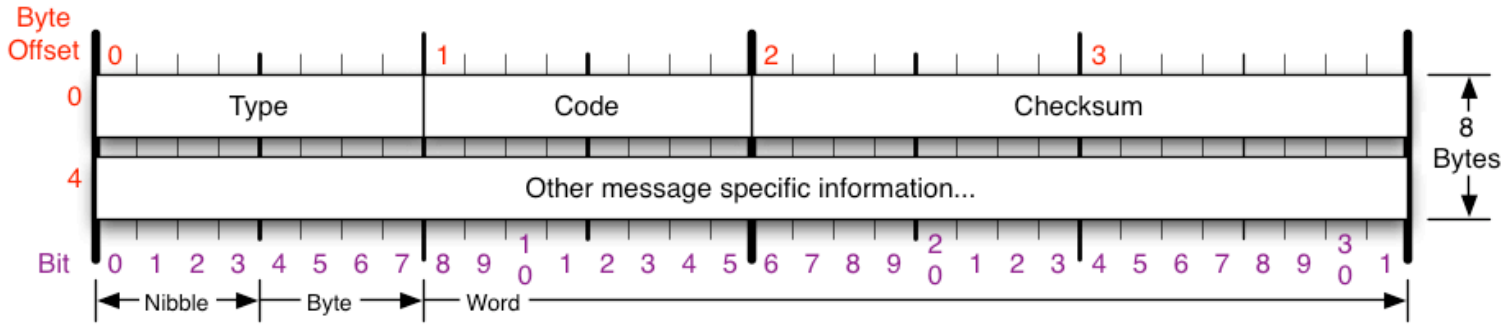
- Max size = 1024
- Initial congestion window = 64 KB

1. เกิดการ Timeout ขึ้น Threshold ถูก Set เป็น ½ (32 KB) และ congestion window = 1 KB
2. congestion window โตขึ้นแบบ Exponentially (Slow Start Algorithm (Jacobson, 1998)
3. มื่อ congestion window โตขึ้นจนถึง Threshold (32 KB) ขนาดของ congestion window จะเพิ่มแบบ Linear (ทีละ 1 KB)

ICMP (Internet Control Message Protocol)

- ICMP เป็นโปรโตคอลที่ใช้ในการตรวจสอบและรายงานสถานภาพของดาต้าแกรม (Datagram) ในกรณีที่เกิดปัญหากับดาต้าแกรม เช่น เราเตอร์ไม่สามารถส่งดาต้าแกรมไปถึงปลายทางได้ ICMP จะถูกส่งออกไปยังโฮสต้นทางเพื่อรายงานข้อผิดพลาด ที่เกิดขึ้น อย่างไรก็ดี ไม่มีอะไรรับประกันได้ว่า ICMP Message ที่ส่งไปจะถึงผู้รับจริงหรือไม่ หากมีการส่งดาต้าแกรมออกไปแล้วไม่มี ICMP Message ฟ้อง Error กลับมา ก็แปลความหมายได้สองกรณีคือ ข้อมูลถูกส่งไปถึงปลายทางอย่างเรียบร้อย หรืออาจจะมีปัญหา ในการสื่อสารทั้งการส่งดาต้าแกรม และ ICMP Message ที่ส่งกลับมาก็มีปัญหาระว่างทางก็ได้ ICMP จึงเป็นโปรโตคอลที่ไม่มีความน่าเชื่อถือ (unreliable) ซึ่งจะเป็นหน้าที่ของ โปรโตคอลในระดับสูงกว่า Network Layer ในการจัดการให้การสื่อสารนั้นๆ มีความน่าเชื่อถือ

ICMP Header



ICMP Message Types

Type	Code/Name
0	Echo Reply
3	Destination Unreachable
	0 Net Unreachable
	1 Host Unreachable
	2 Protocol Unreachable
	3 Port Unreachable
	4 Fragmentation required, and DF set
	5 Source Route Failed
	6 Destination Network Unknown
	7 Destination Host Unknown
	8 Source Host Isolated
	9 Network Administratively Prohibited
	10 Host Administratively Prohibited
	11 Network Unreachable for TOS
3	Destination Unreachable (continued)
	12 Host Unreachable for TOS
	13 Communication Administratively Prohibited
4	Source Quench
5	Redirect
	0 Redirect Datagram for the Network
	1 Redirect Datagram for the Host
	2 Redirect Datagram for the TOS & Network
	3 Redirect Datagram for the TOS & Host
8	Echo
9	Router Advertisement
10	Router Selection
11	Time Exceeded
	0 TTL Exceeded
	1 Fragment Reassembly Time Exceeded
12	Parameter Problem
	0 Pointer Problem
	1 Missing a Required Operand
	2 Bad Length
13	Timestamp
14	Timestamp Reply
15	Information Request
16	Information Reply
17	Address Mask Request
18	Address Mask Reply
30	Traceroute

Checksum

Checksum of ICMP header

RFC 792

Please refer to RFC 792 for the Internet Control Message protocol (ICMP) specification.

ICMP ที่สำคัญ

ชนิดของข้อมูล ICMP	ความหมาย
0 – echo replay	ตอบกลับ Echo (message echo)
8 - echo	ส่ง echo message ไปเพื่อถามว่า ปลายทางยังทำงานหรือไม่
3- Destination Unreachable	Packet ไม่สามารถถูกส่งไปถึง ปลายทางได้
4- Source quench	Route เริ่มที่จะ congest ให้ทำการลดความเร็วในการส่งข้อมูล
11- Time exceeded	ส่งคืน source เมื่อ route ทำการ drop packet