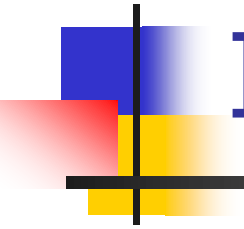
Nov 2010



IPv4 (Part III)

รศ.ดร. อนันต์ ผลเพิ่ม

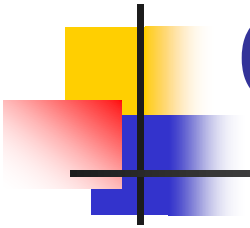
Asso. Prof. Anan Phonphoem, Ph.D.

anan.p@ku.ac.th

http://www.cpe.ku.ac.th/~anan

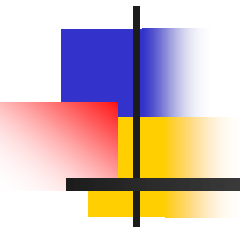
Computer Engineering Department

Kasetsart University, Bangkok, Thailand

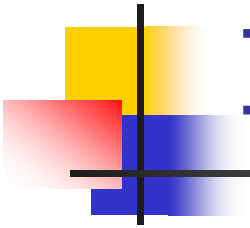


Outline

- IP Fundamental Operation
- Internet Protocol
- Addressing
- Supporting Protocol
  - ARP
  - ICMP: ping + traceroute
  - NAT
  - DHCP

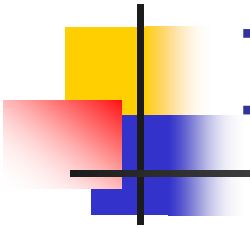


Internet Control Message Protocol (ICMP)



ICMP

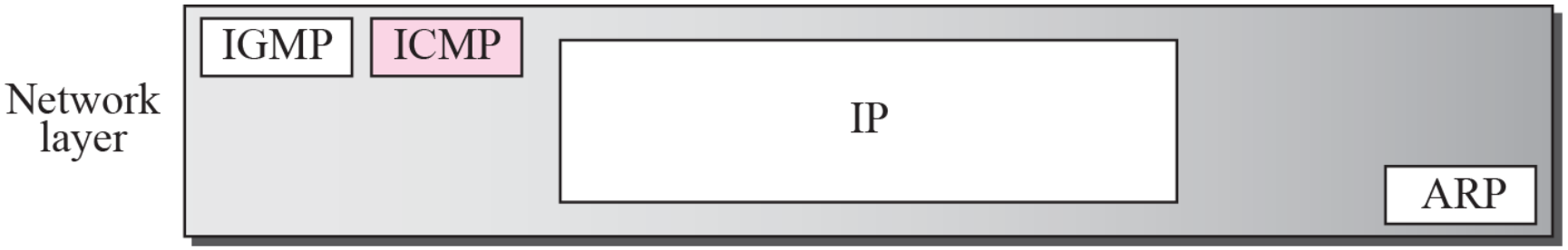
- IP supporter
- For error generating
  - Transmission problem
  - TTL exceed
  - Destination unreachable
  - etc.
- Serve as useful diagnostics
  - ping, traceroute



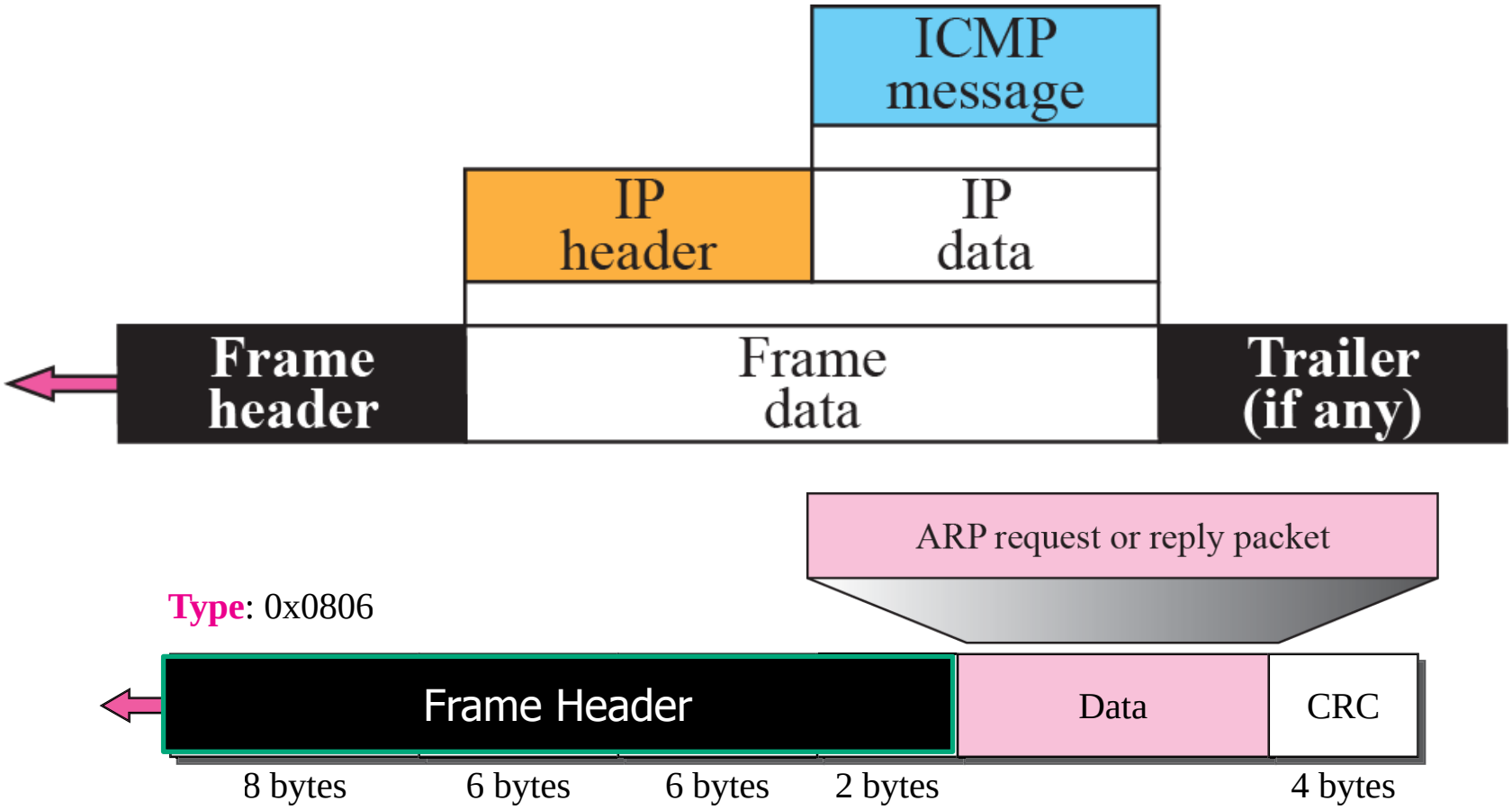
ICMP

- ICMP error messages never generates due to:
  - ICMP error messages selves
  - Broadcast/Multicast (prevent broadcast Storms)
- What are Broadcast Storms ?
  - A large number of broadcast datalink frames transmitted nearly simultaneous
  - LAN may be freeze !

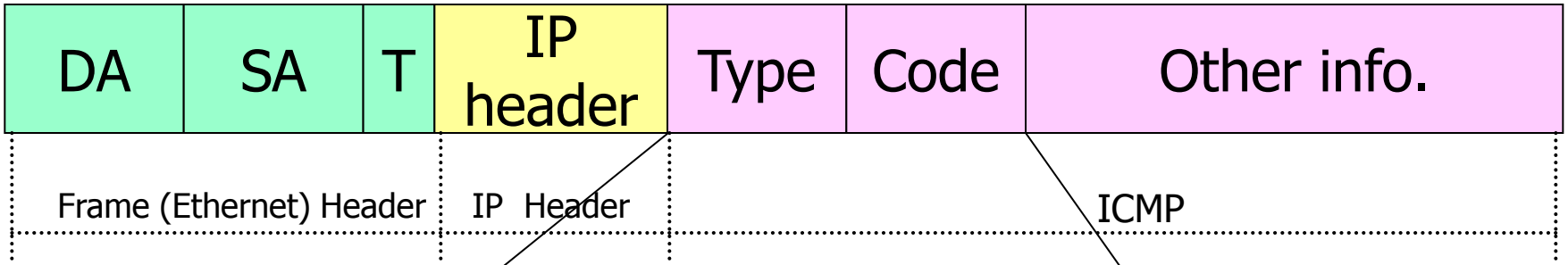
Positioning of ICMP



ICMP encapsulation



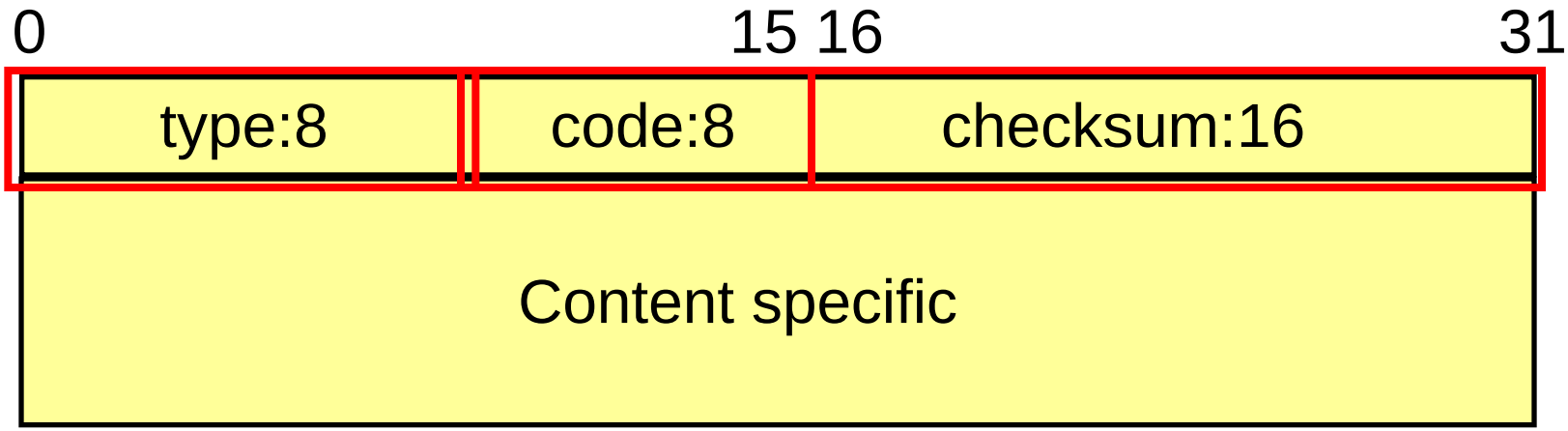
ICMP packet



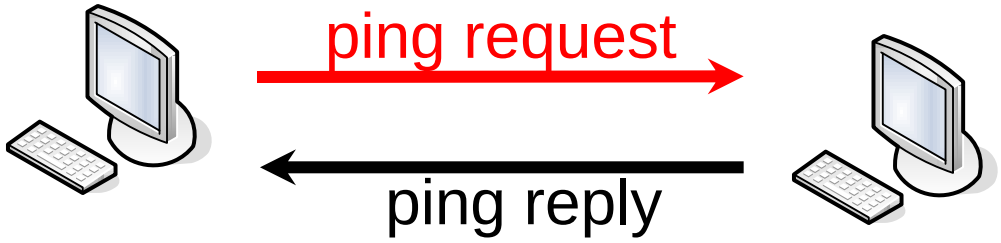
Type	Code	Description
8	0	Echo request
0	0	Echo reply
11	0	Time exceed
3	3	Port unreachable

ICMP header

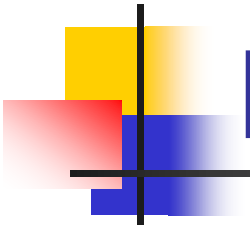
- type - relevant ICMP message
- code - more detail information
- checksum - covers ICMP header/data (not IP hdr)



ping



- Generate an ICMP echo request
- Receive the ICMP echo reply
- All TCP/IP node is supposed to implement ICMP and respond to ICMP echo



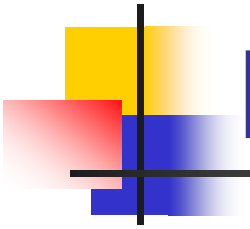
ping command

- Send an echo request message every seconds
- Records the time it takes for each reply
- Every echo request contains a unique sequence number to match replies and request
  - Record round-trip timing
  - Perform packet lost statistics

ping example

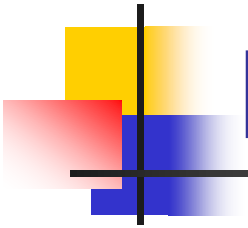
```
[anan@alpha anan]$ ping iwing.cpe.ku.ac.th
PING iwing.cpe.ku.ac.th (158.108.32.199) from 158.108.32.31 : 56(84) bytes of data.
Warning: time of day goes back, taking countermeasures.
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=0 ttl=252 time=1.187 msec
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=1 ttl=252 time=601 usec
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=2 ttl=252 time=594 usec
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=3 ttl=252 time=594 usec
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=4 ttl=252 time=585 usec
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=5 ttl=252 time=590 usec
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=6 ttl=252 time=584 usec
64 bytes from iwing.cpe.ku.ac.th (158.108.32.199): icmp_seq=7 ttl=252 time=587 usec

--- iwing.cpe.ku.ac.th ping statistics ---
8 packets transmitted, 8 packets received, 0% packet loss
round-trip min/avg/max/mdev = 0.584/0.665/1.187/0.198 ms
```



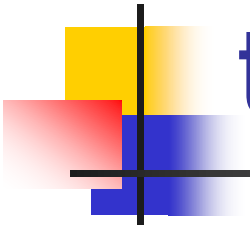
ping as debugging tools

- What do we get from ping?
  - Timing information
  - Connection reliability
  - Destination is reachable (routable)
  - Layer is functional, but not guaranteed telnet!



ping results

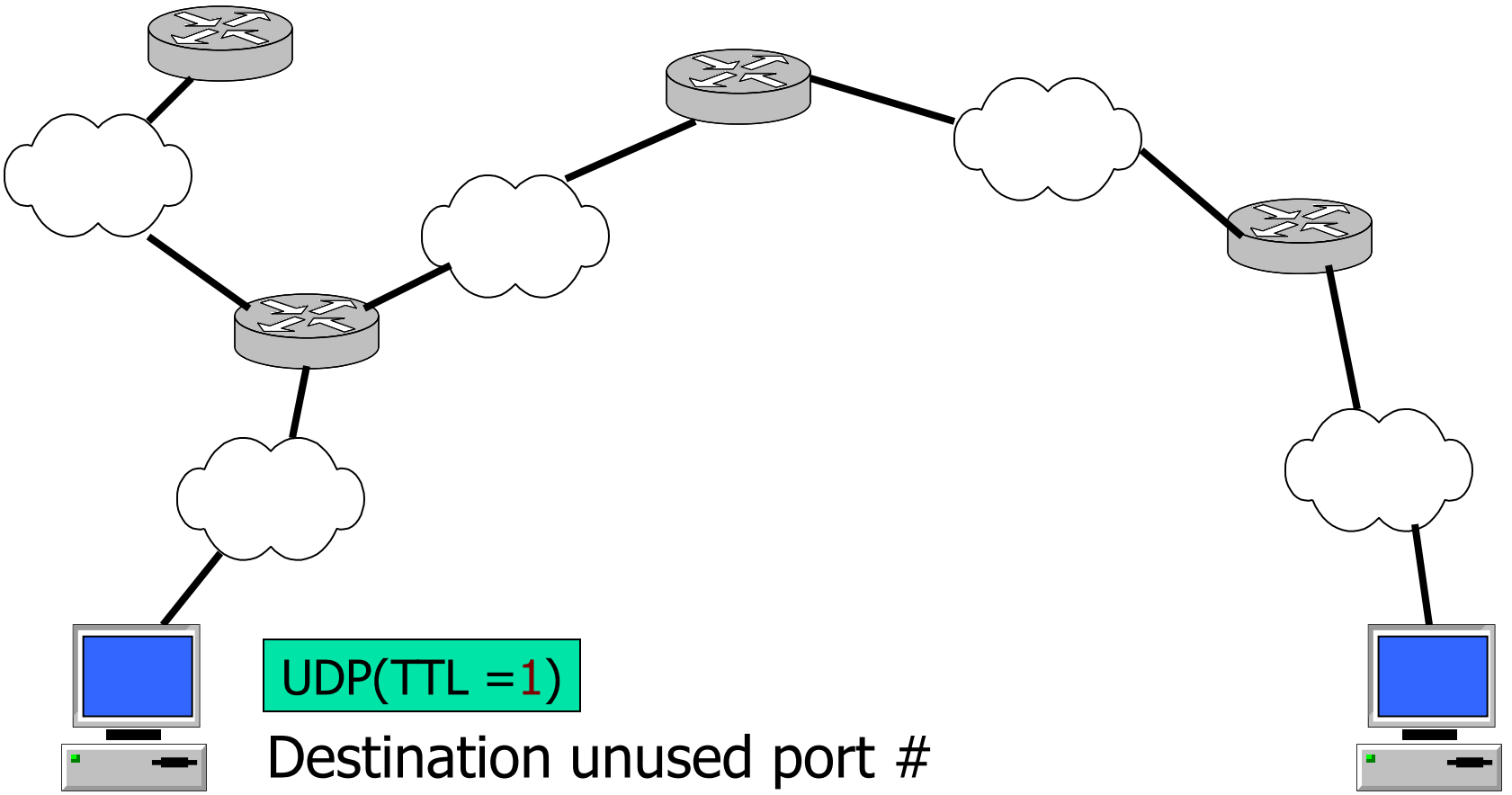
- no response
  - no end node, no connection
- lost packet (significant when >2-3%)
  - transmission error on WAN/LAN, overloading bridges/routers
- time acknowledge vary
  - host/network overloading, >100 ms make telnet less acceptable)
- no lost and echo time is reasonably constant
  - Congratulation! That's all we want.



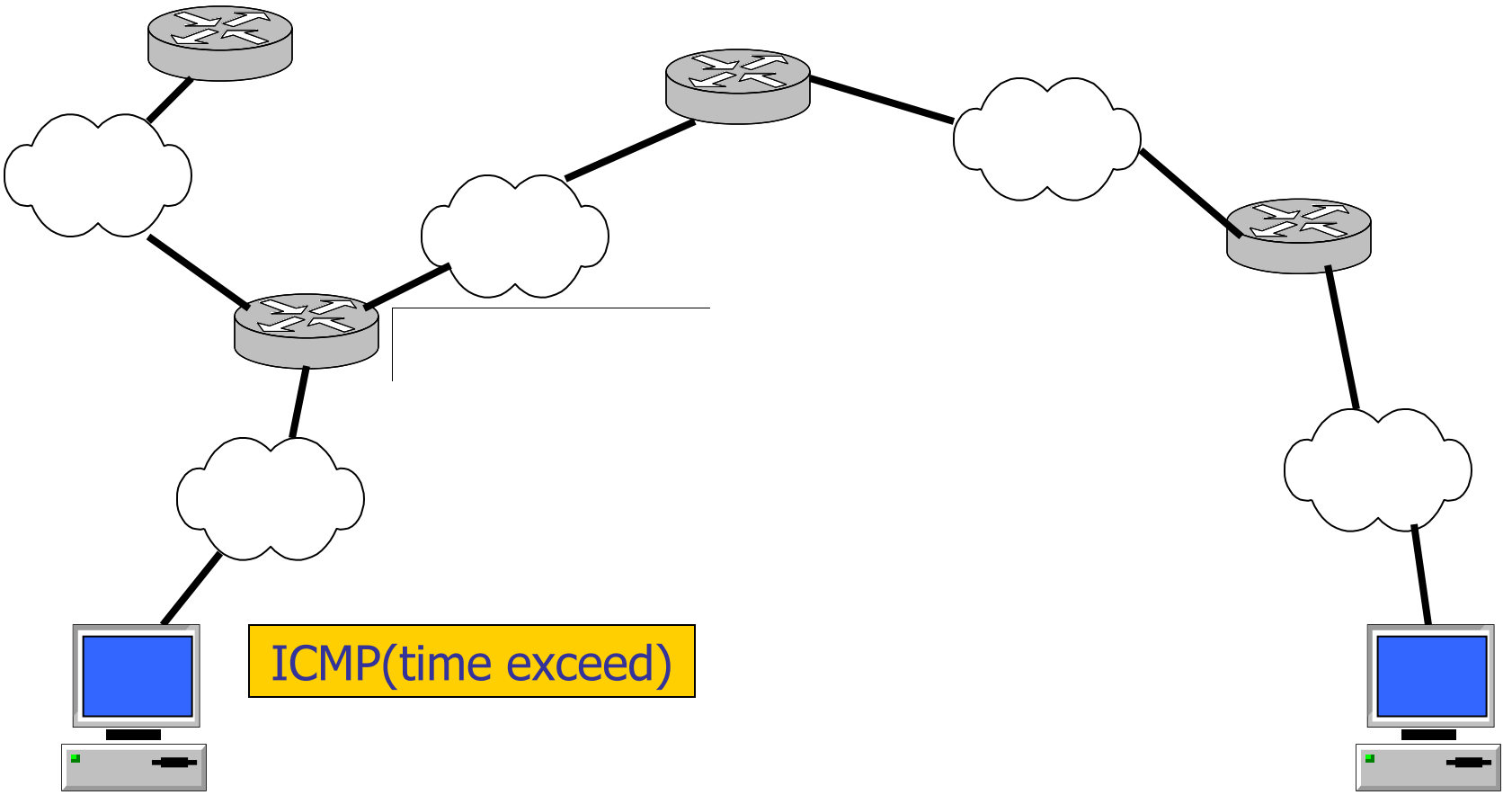
traceroute

- Command to determine the active route to a destination address
- How to?
  - send a UDP messages to an unused port on the target host with ttl=1
  - router decrease ttl to 0, it has to return an ICMP time exceed message
  - traceroute sets ttl =2 and retransmits, this time go one more hop
  - ttl++ until UDP messages reach the destination.
  - the target returns an ICMP service unavailable because there is no UDP port service.

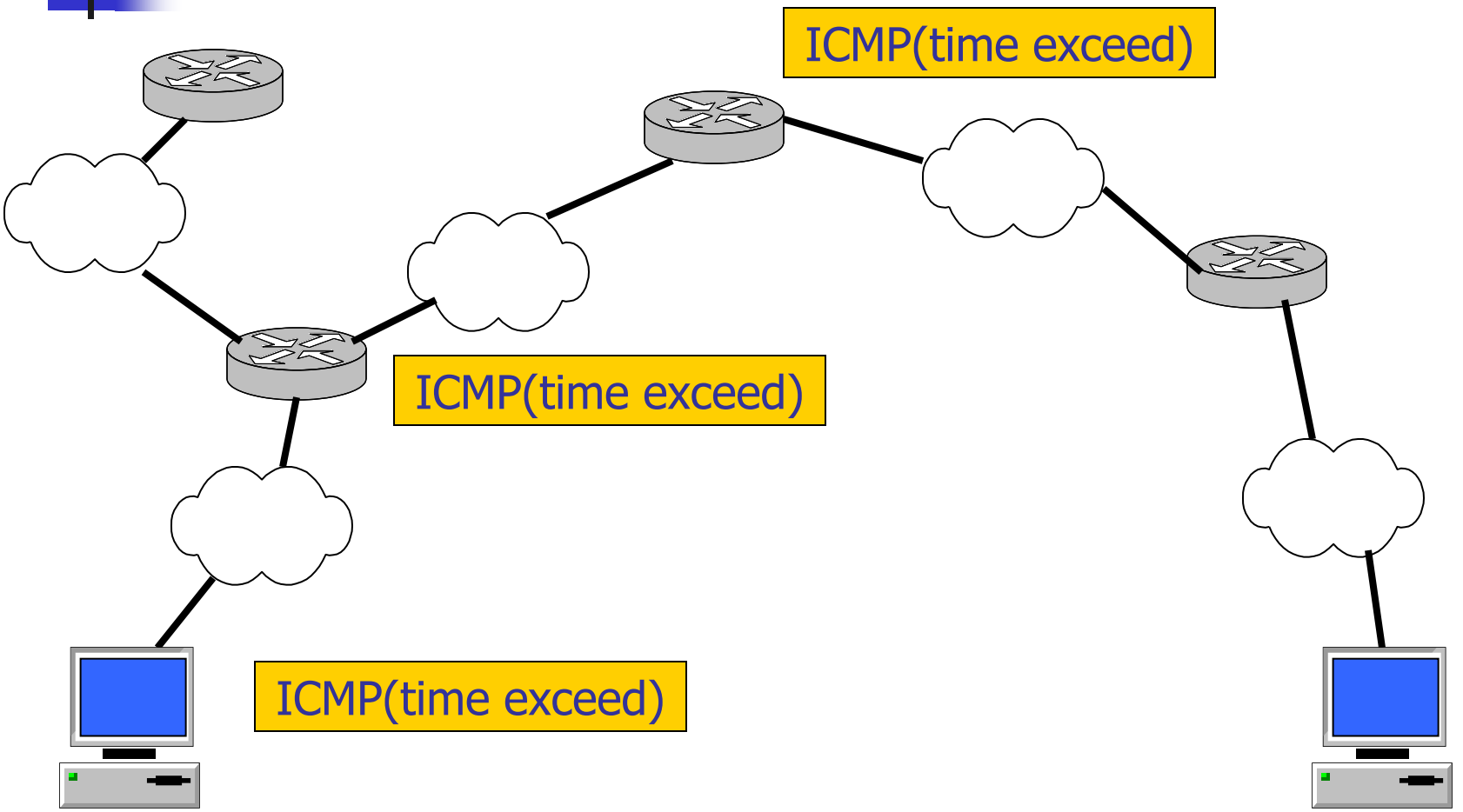
How traceroute works?



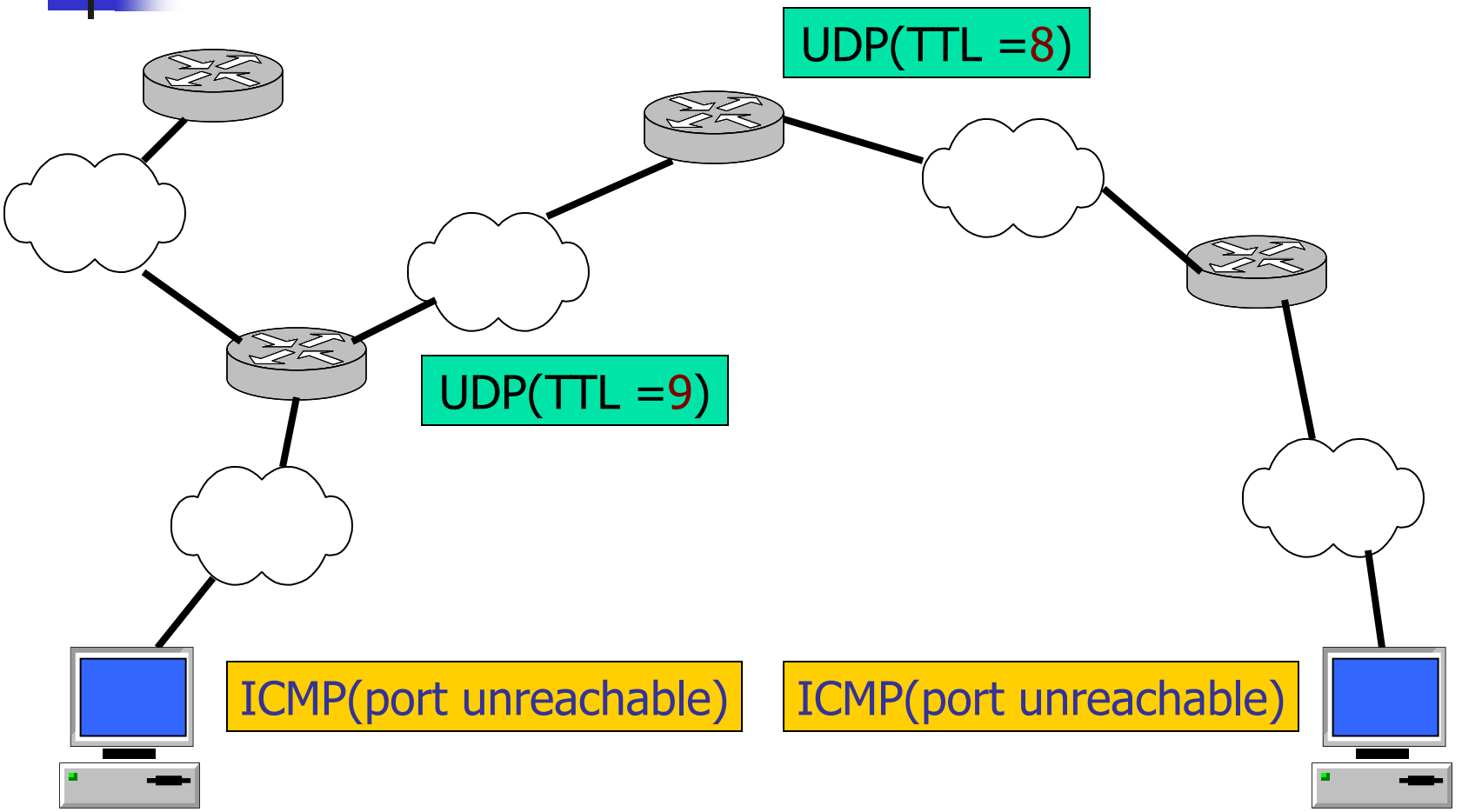
How traceroute works?



How traceroute works?



How traceroute works?



Various of traceroute: TCP sync (not common)

Traceroute example

```
[anan@alpha anan]$ /usr/sbin/traceroute iwing.cpe.ku.ac.th
traceroute to iwing.cpe.ku.ac.th (158.108.32.199), 30 hops max, 38 byte packets
 1  fe-cpegw2-server (158.108.32.1) 0.851 ms  0.782 ms  0.683 ms
 2  gb-cpegwbb-cpegw (158.108.35.10) 0.387 ms  0.368 ms  0.337 ms
 3  gb-cpec4k6-cpec6k (158.108.35.114) 0.685 ms  0.654 ms  0.613 ms
 4  iwing (158.108.32.199) 0.506 ms  0.439 ms  0.418 ms
```

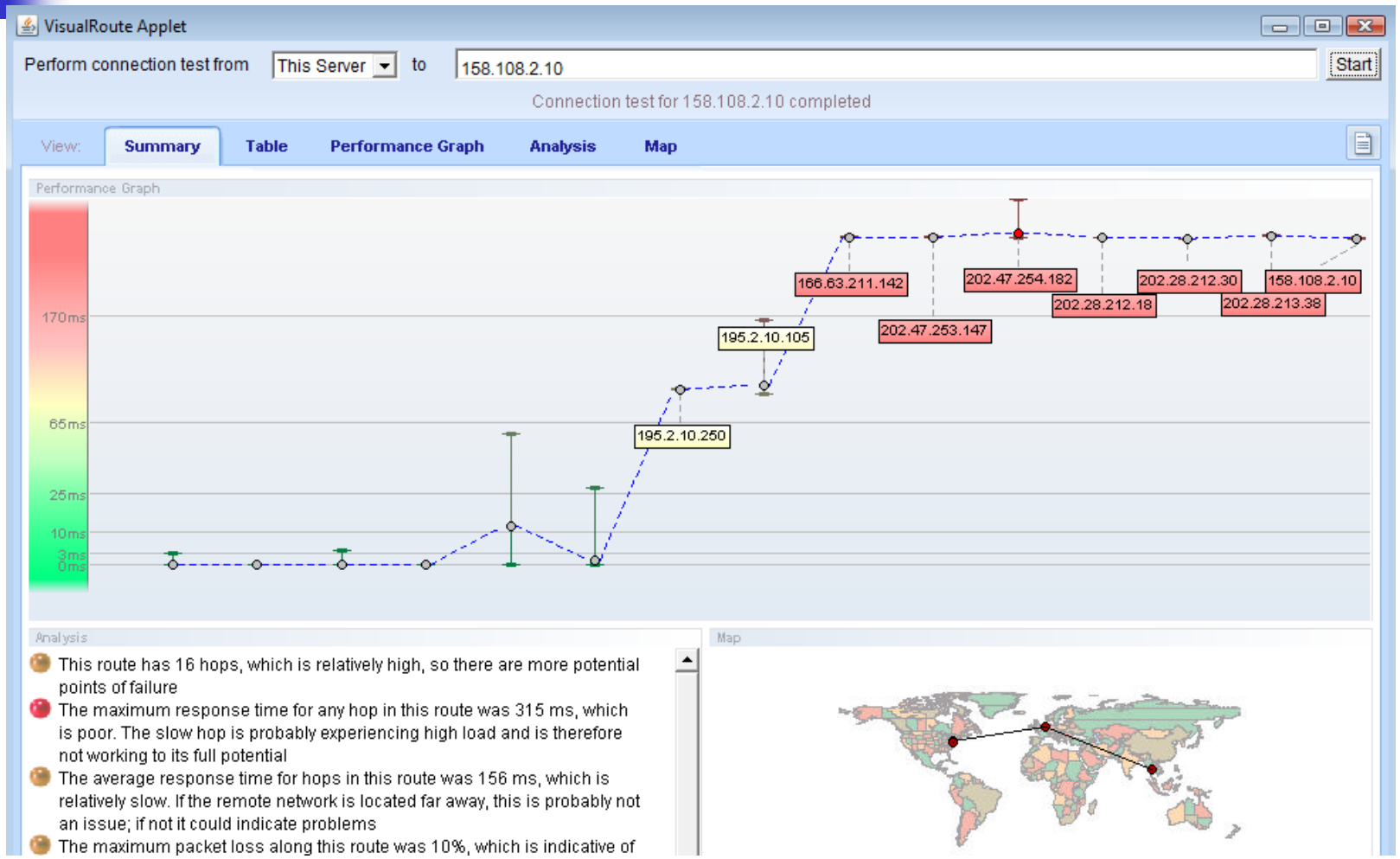
Traceroute example

```
[anan@alpha anan]$ /usr/sbin/traceroute www.umass.edu
traceroute to www.umass.edu (128.119.101.5), 30 hops max, 38 byte packets
 1  fe-cpegw2-server (158.108.32.1) 0.855 ms  0.737 ms  0.700 ms
 2  gb-cpegwbb-cpegw (158.108.35.10) 0.430 ms  0.409 ms  0.359 ms
 3 158.108.254.37 (158.108.254.37) 0.488 ms  0.469 ms  0.401 ms
 4 158.108.251.54 (158.108.251.54) 0.558 ms  0.617 ms  0.733 ms
 5 158.108.251.57 (158.108.251.57) 1.121 ms  0.919 ms  1.046 ms
 6 202.28.213.1 (202.28.213.1) 1.311 ms  1.758 ms  1.154 ms
 7 202.28.212.29 (202.28.212.29) 1.531 ms  1.445 ms  1.189 ms
 8 202.28.212.2 (202.28.212.2) 1.456 ms  1.532 ms  1.151 ms
 9  S1-1.R00.LA-POP.uni.net.th (202.28.28.162) 226.026 ms  226.043 ms  225.962 ms
10 63.216.18.53 (63.216.18.53) 253.741 ms  239.317 ms  249.022 ms
11  snvang-losang.abilene.ucaid.edu (198.32.8.95) 233.765 ms  239.165 ms  240.522 ms
12  dnvrng-snvang.abilene.ucaid.edu (198.32.8.2) 258.216 ms  258.599 ms *
13  kscyng-dnvrng.abilene.ucaid.edu (198.32.8.14) 269.012 ms  268.717 ms  318.331 ms
...
19  nox300gw1-PEER-NoX-UMASS-192-5-89-102.nox.org (192.5.89.102) 310.155 ms  310.240 ms
344.973 ms
20  lgrc-rt-106-8.gw.umass.edu (128.119.2.193) 323.127 ms  325.108 ms  313.802 ms
21  lgrc-rt-106-6.gw.umass.edu (128.119.2.185) 310.291 ms  321.111 ms  309.874 ms
22 * * *
23 * * *
```

Example GUI Traceroute program:

Visual Route

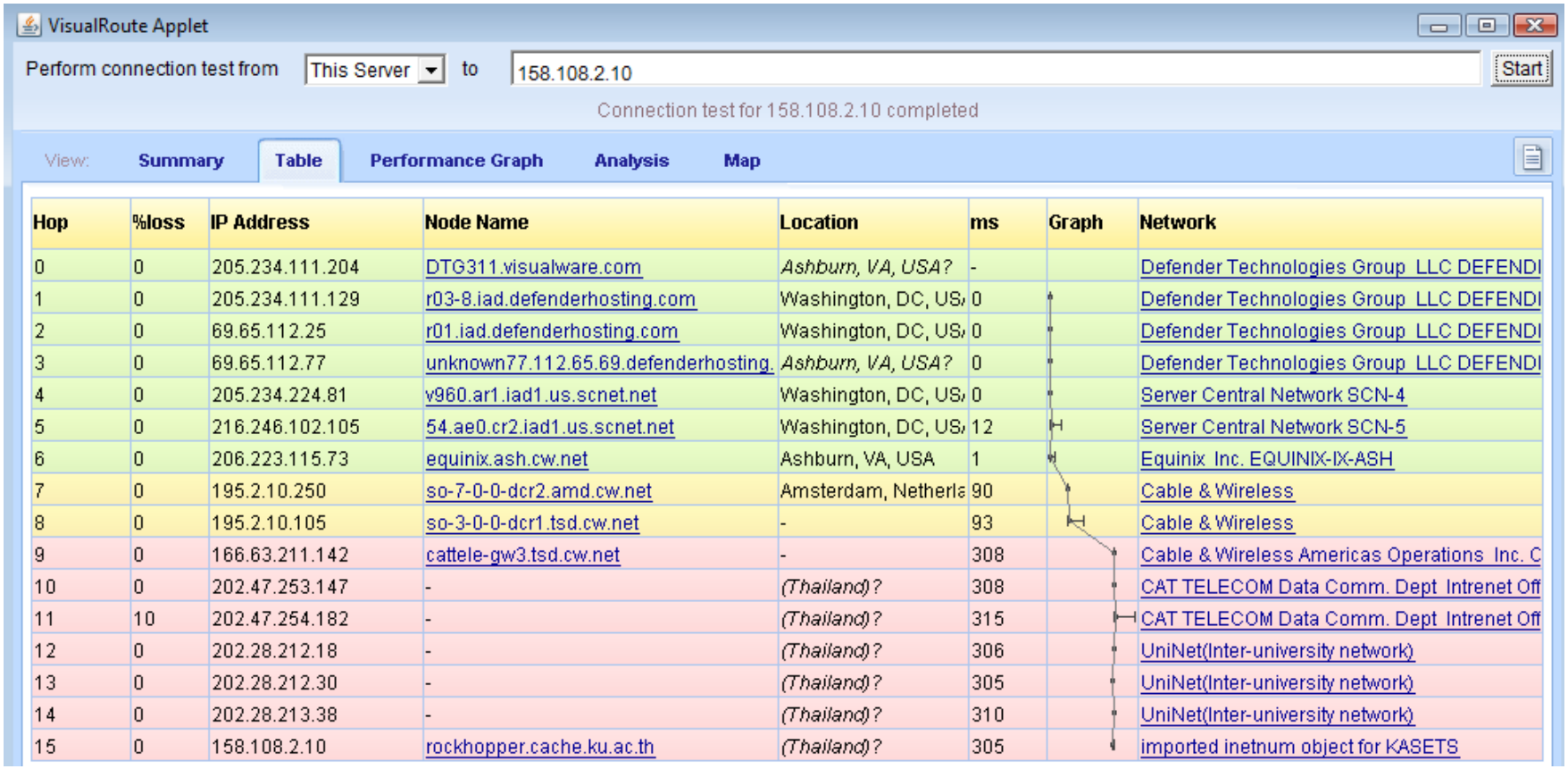
http://visualroute.visualware.com/



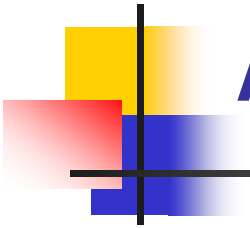
Example GUI Traceroute program:

Visual Route

http://visualroute.visualware.com/



Hop	%loss	IP Address	Node Name	Location	ms	Graph	Network
0	0	205.234.111.204	DTG311.visualware.com	*Ashburn, VA, USA?*	-		Defender Technologies Group LLC DEFENDI
1	0	205.234.111.129	r03-8.iad.defenderhosting.com	Washington, DC, US/	0		Defender Technologies Group LLC DEFENDI
2	0	69.65.112.25	r01.iad.defenderhosting.com	Washington, DC, US/	0		Defender Technologies Group LLC DEFENDI
3	0	69.65.112.77	unknown77.112.65.69.defenderhosting.	*Ashburn, VA, USA?*	0		Defender Technologies Group LLC DEFENDI
4	0	205.234.224.81	v960.ar1.iad1.us.scnet.net	Washington, DC, US/	0		Server Central Network SCN-4
5	0	216.246.102.105	54.ae0.cr2.iad1.us.scnet.net	Washington, DC, US/	12		Server Central Network SCN-5
6	0	206.223.115.73	equinix.ash.cw.net	Ashburn, VA, USA	1		Equinix Inc. EQUINIX-IX-ASH
7	0	195.2.10.250	so-7-0-0-dcr2.amd.cw.net	Amsterdam, Netherla	90		Cable & Wireless
8	0	195.2.10.105	so-3-0-0-dcr1.tsd.cw.net	-	93		Cable & Wireless
9	0	166.63.211.142	cattele-gw3.tsd.cw.net	-	308		Cable & Wireless Americas Operations Inc. C
10	0	202.47.253.147	-	*(Thailand)?*	308		CAT TELECOM Data Comm. Dept Intrenet Off
11	10	202.47.254.182	-	*(Thailand)?*	315		CAT TELECOM Data Comm. Dept Intrenet Off
12	0	202.28.212.18	-	*(Thailand)?*	306		UniNet(Inter-university network)
13	0	202.28.212.30	-	*(Thailand)?*	305		UniNet(Inter-university network)
14	0	202.28.213.38	-	*(Thailand)?*	310		UniNet(Inter-university network)
15	0	158.108.2.10	rockhopper.cache.ku.ac.th	*(Thailand)?*	305		imported inetnum object for KASETS

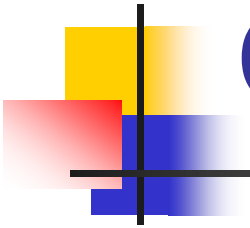


Assignment

- Select 2 URLs from any site in the world
  - Not the same continent
- On different times (e.g. 9AM, 3PM, 11PM) of the day
  - ping
  - traceroute (from your machine)
- **Create**
  - an example graphical route for each URL
  - a comparison table for different times / sites
- **Summarize** and **criticize** the results

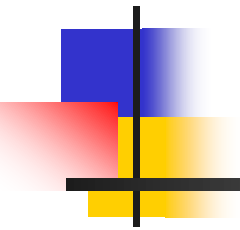
Ping Example

	Ping www.umass.edu		
Time	**results**	**Average (ms)**	**% loss**
9AM	[128.119.101.5] with 32 bytes of data: Reply from 128.119.101.5: bytes=32 time=322ms TTL=127 Reply from 128.119.101.5: bytes=32 time=506ms TTL=127 Reply from 128.119.101.5: bytes=32 time=502ms TTL=127 Reply from 128.119.101.5: bytes=32 time=325ms TTL=127	413	0
3PM			
11PM			

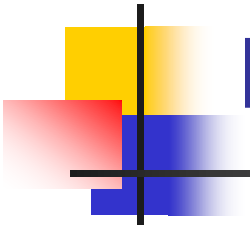


Outline

- IP Fundamental Operation
- Internet Protocol
- Addressing
- Supporting Protocol
  - ARP
  - ICMP: ping + traceroute
  - NAT
  - DHCP



Network Address Translation (NAT)



Network Address Translation (NAT)

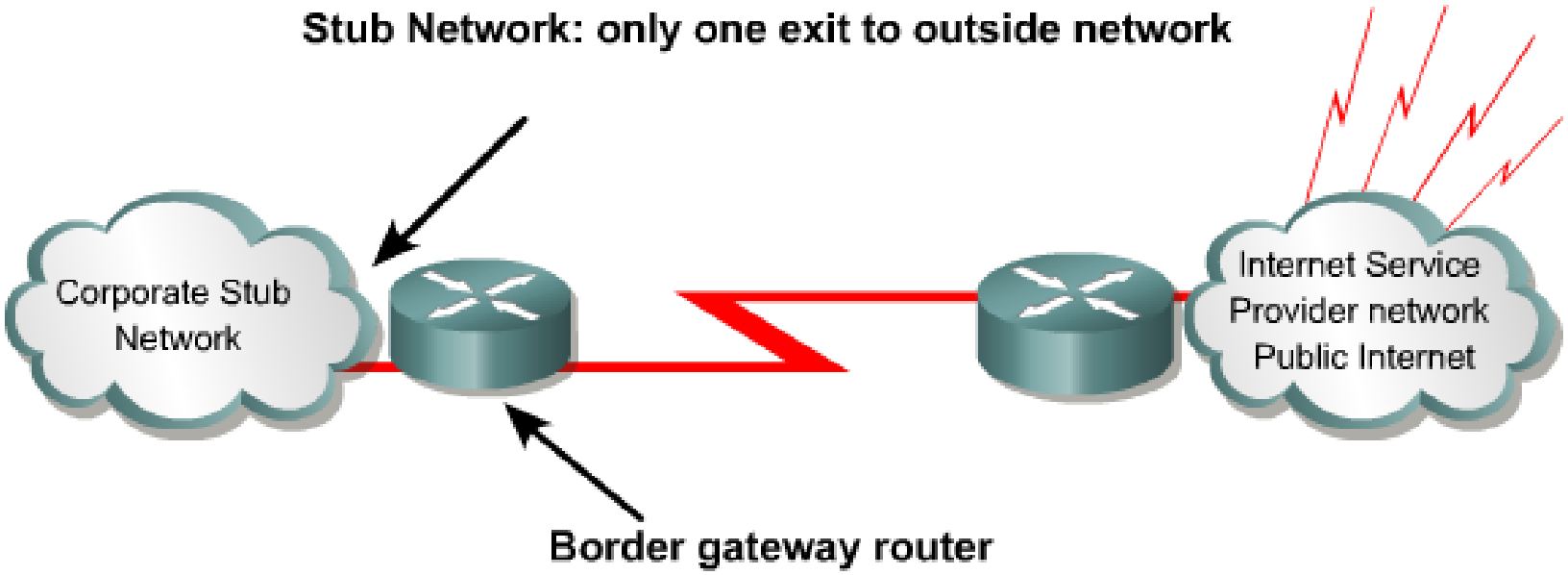
- Private Network
  - Good practice to use private address
- Map local addresses to (real) public IP address(es)
- Security (not expose internal details)
- Alleviate IP depletion

Private IP address

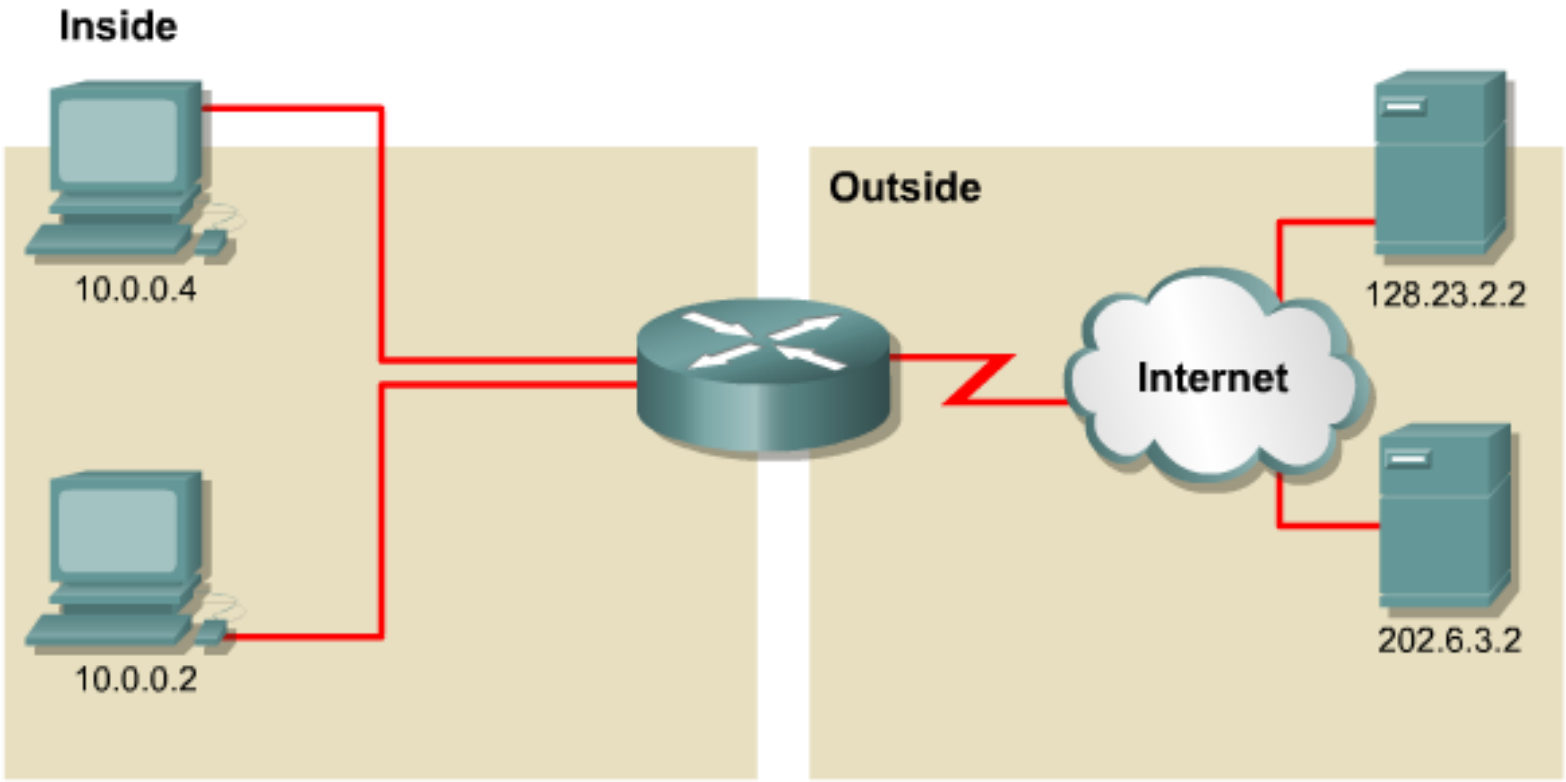
Class	RFC 1918	CIDR prefix
A	10.0.0.0 – 10.255.255.255	10.0.0.0/8
B	172.16.0.0 – 172.16.255.255	172.16.0.0/12
C	192.168.0.0 – 192.168.255.255	192.168.0.0/16

NAT

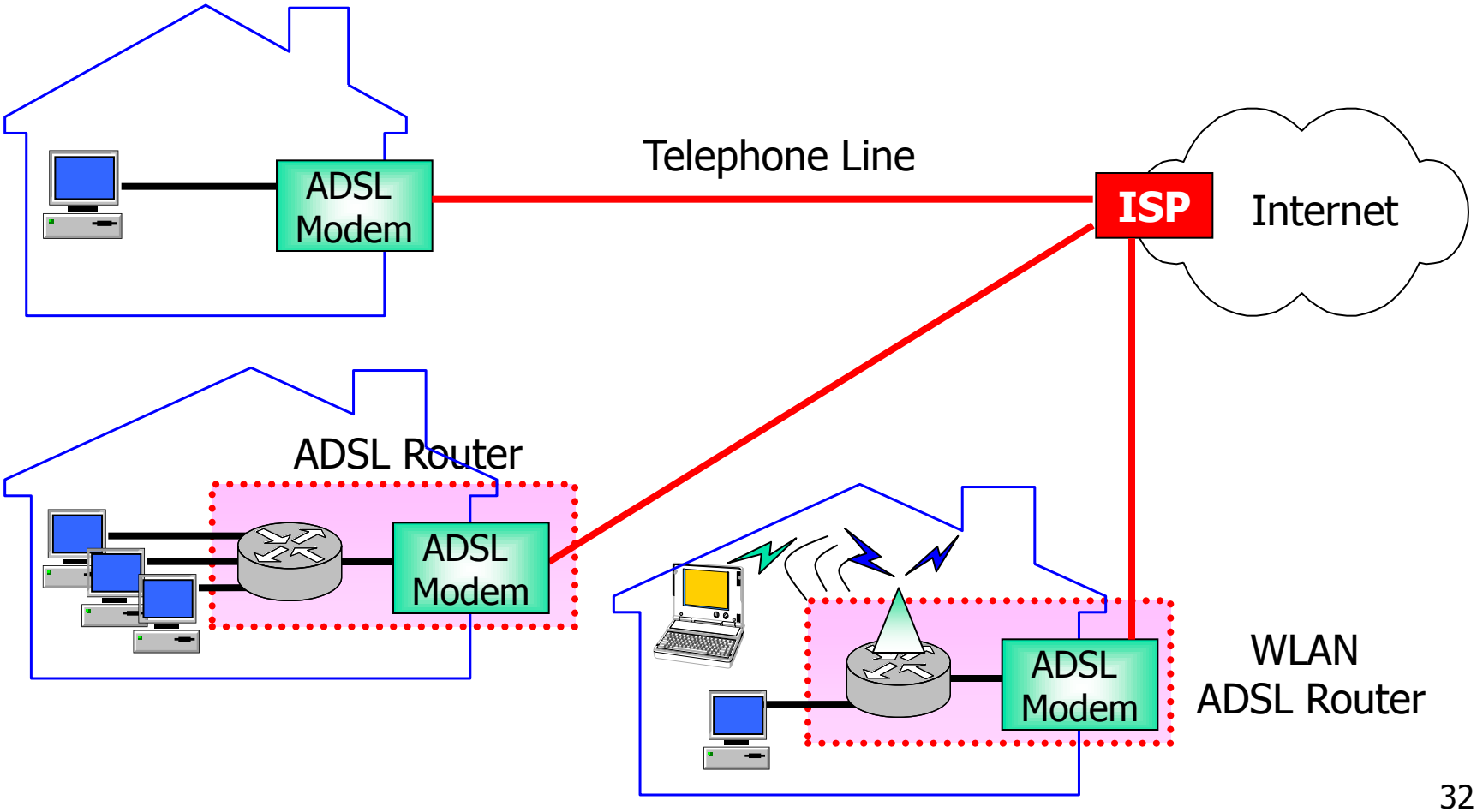
- Operates at the border of a stub network



NAT

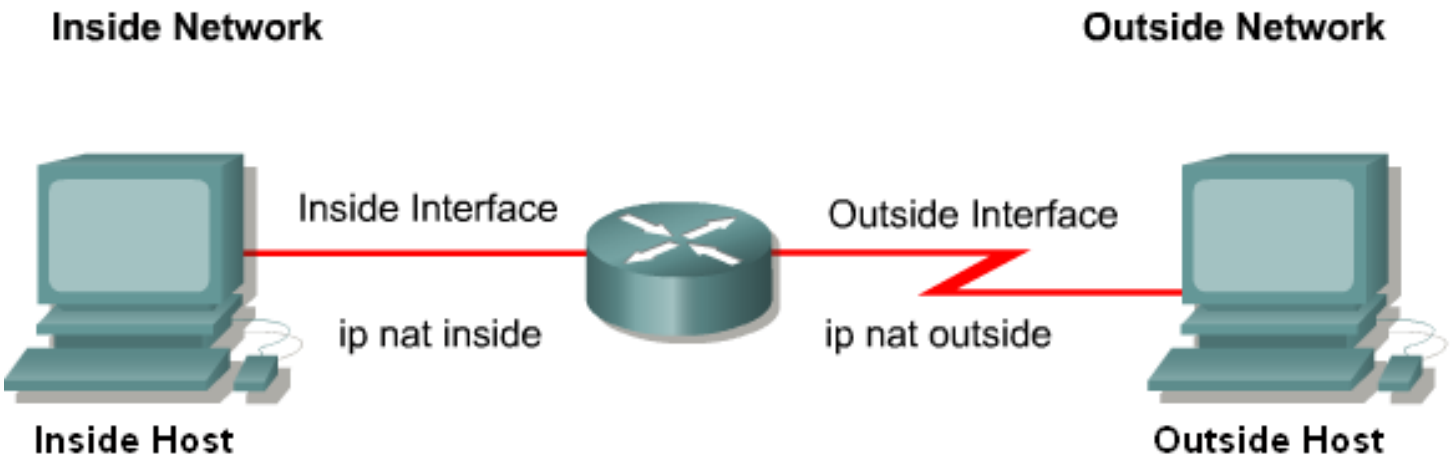


ADSL Connection

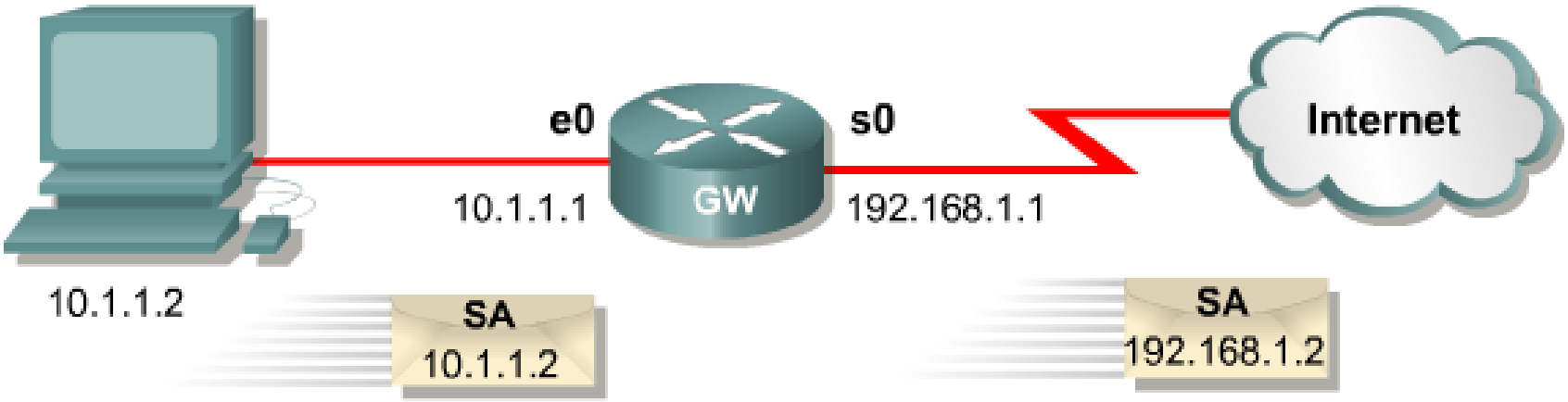


Configure NAT

- Static Translation
- Dynamic Translation

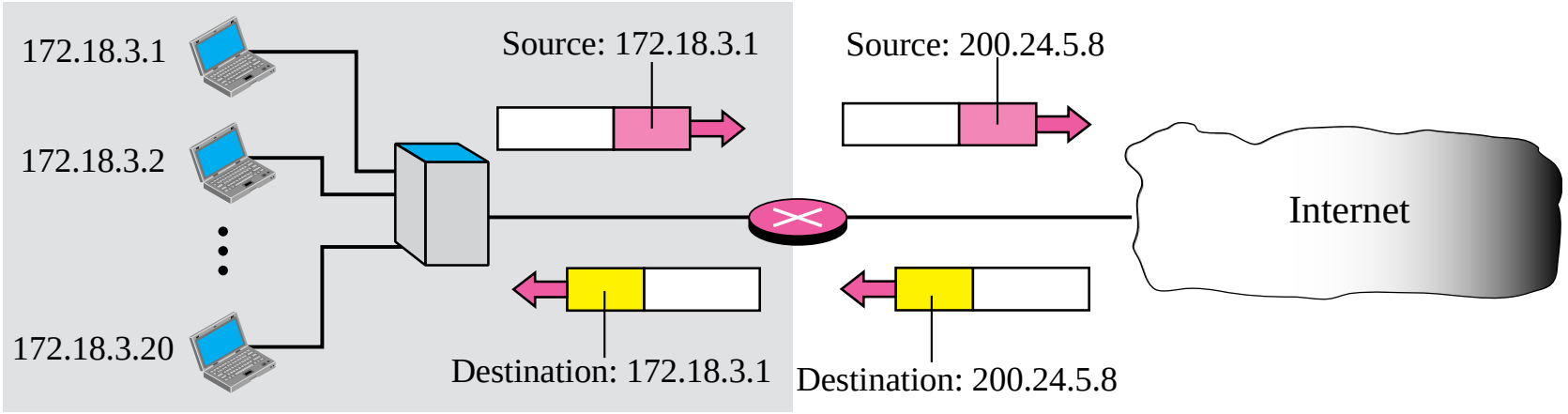


Static Translation



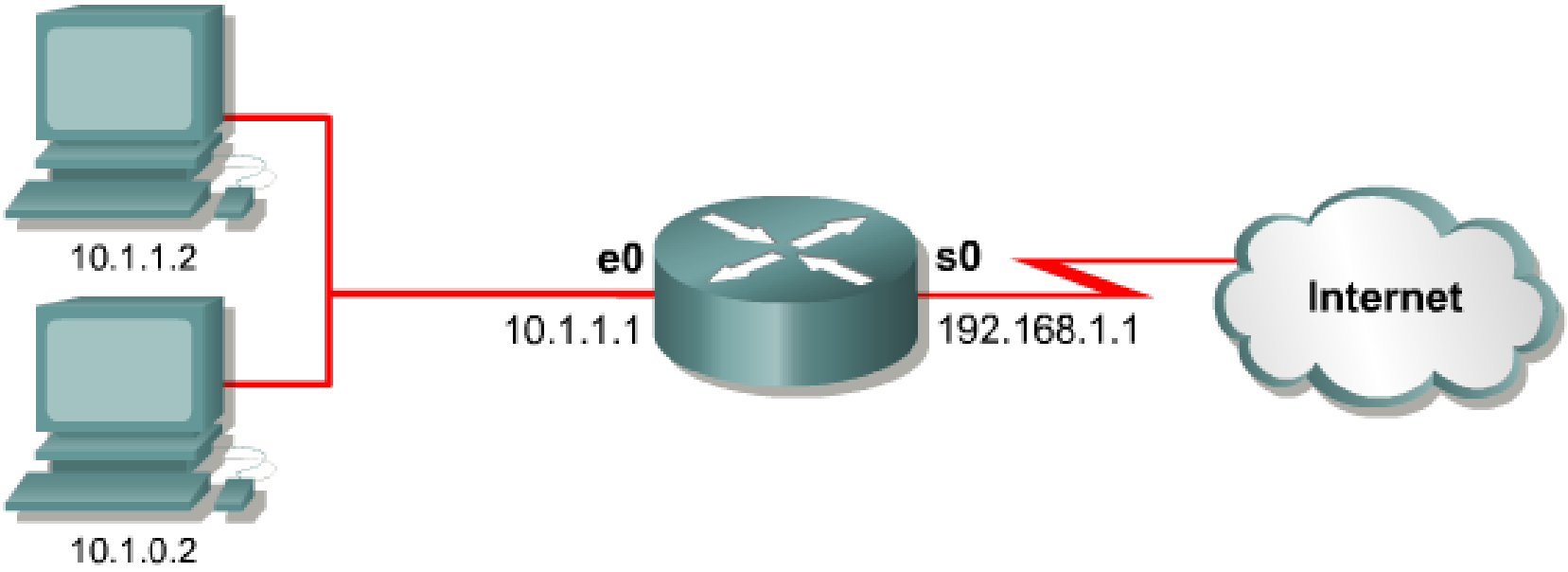

```
hostname GW
!
ip nat inside source static 10.1.1.2 192.168.1.2
!
interface ethernet 0
  ip address 10.1.1.1 255.255.255.0
  ip nat inside
!
interface serial 0
  ip address 192.168.1.1 255.255.255.0
  ip nat outside
!
```

Example

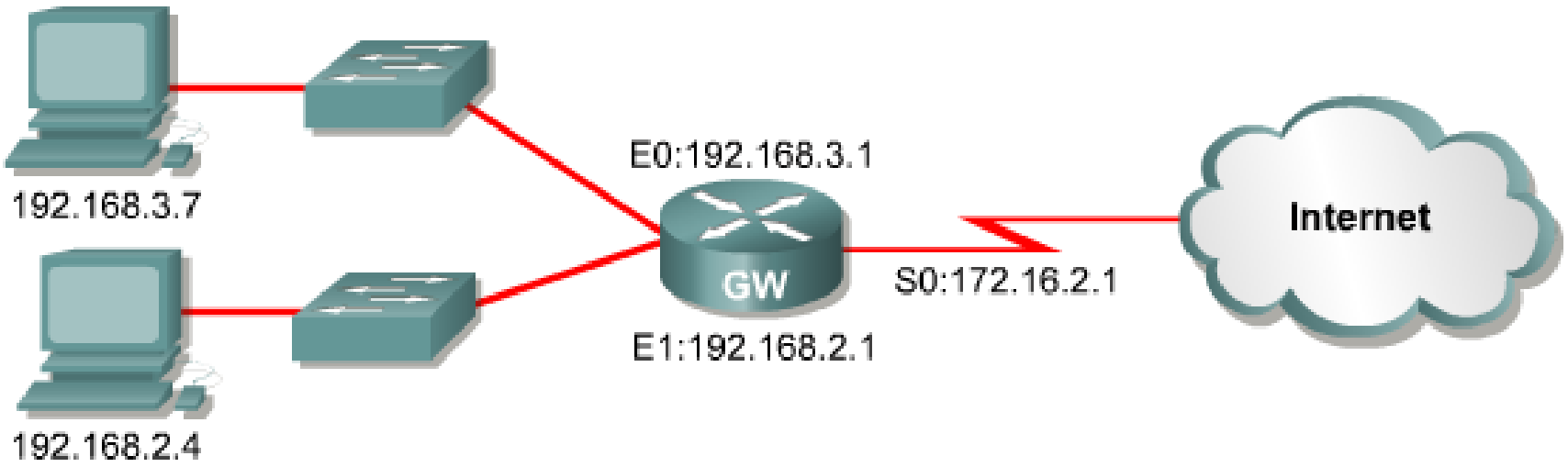


Site using private addresses

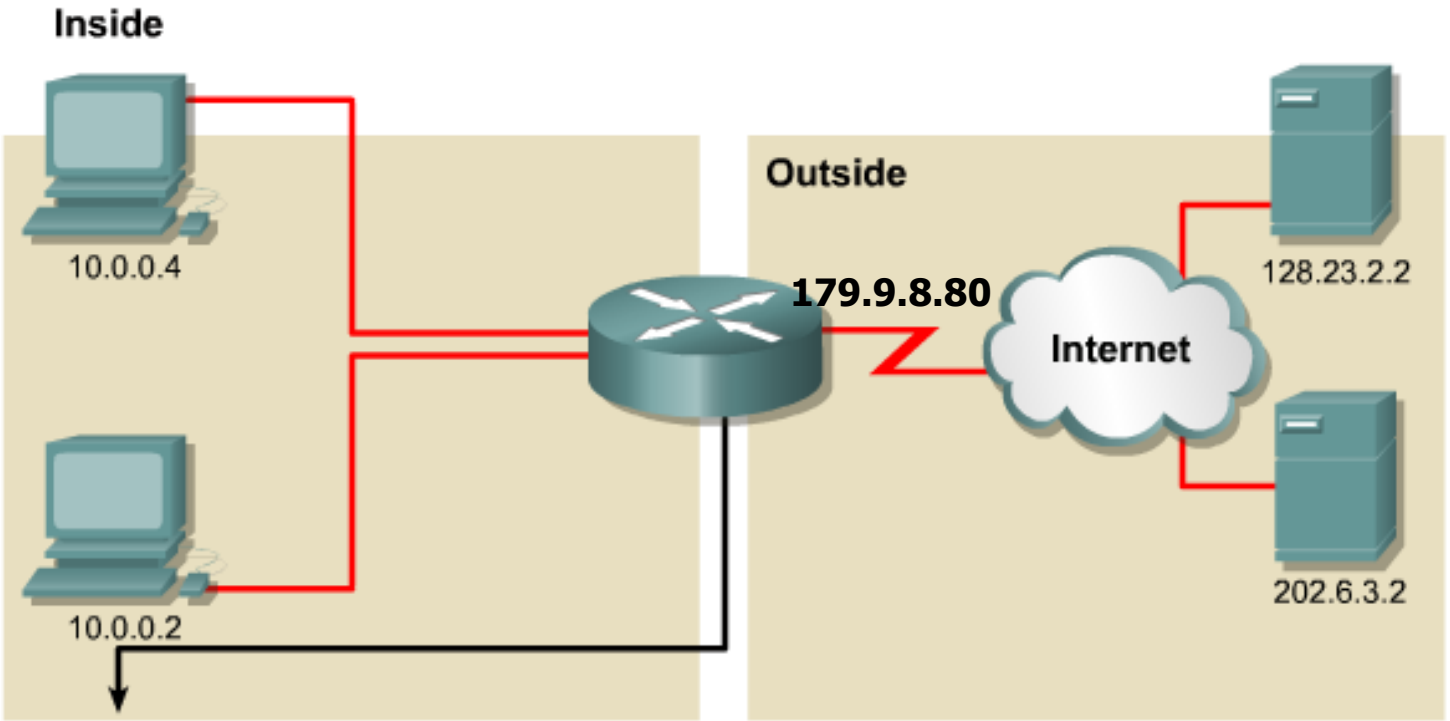
Dynamic Translation



PAT: Port Address Translation (Overloaded NAT)

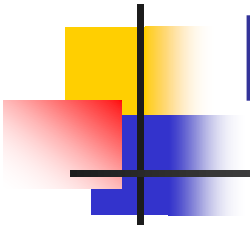


PAT



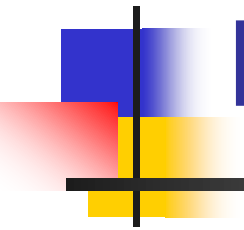
NAT Table with Overload

Inside Local IP Address	Inside Global IP Address	Outside Local IP Address	Outside Global Address
10.0.0.2:1331	179.9.8.80:1331	202.6.3.2:80	202.6.3.2:80
10.0.0.4:1555	179.9.8.80:1555	128.23.2.2:80	128.23.2.2:80

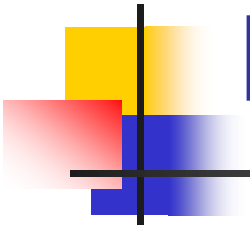


Disadvantages of NAT

- Delay
- Loss of end-to-end ability
- Might not work with some applications



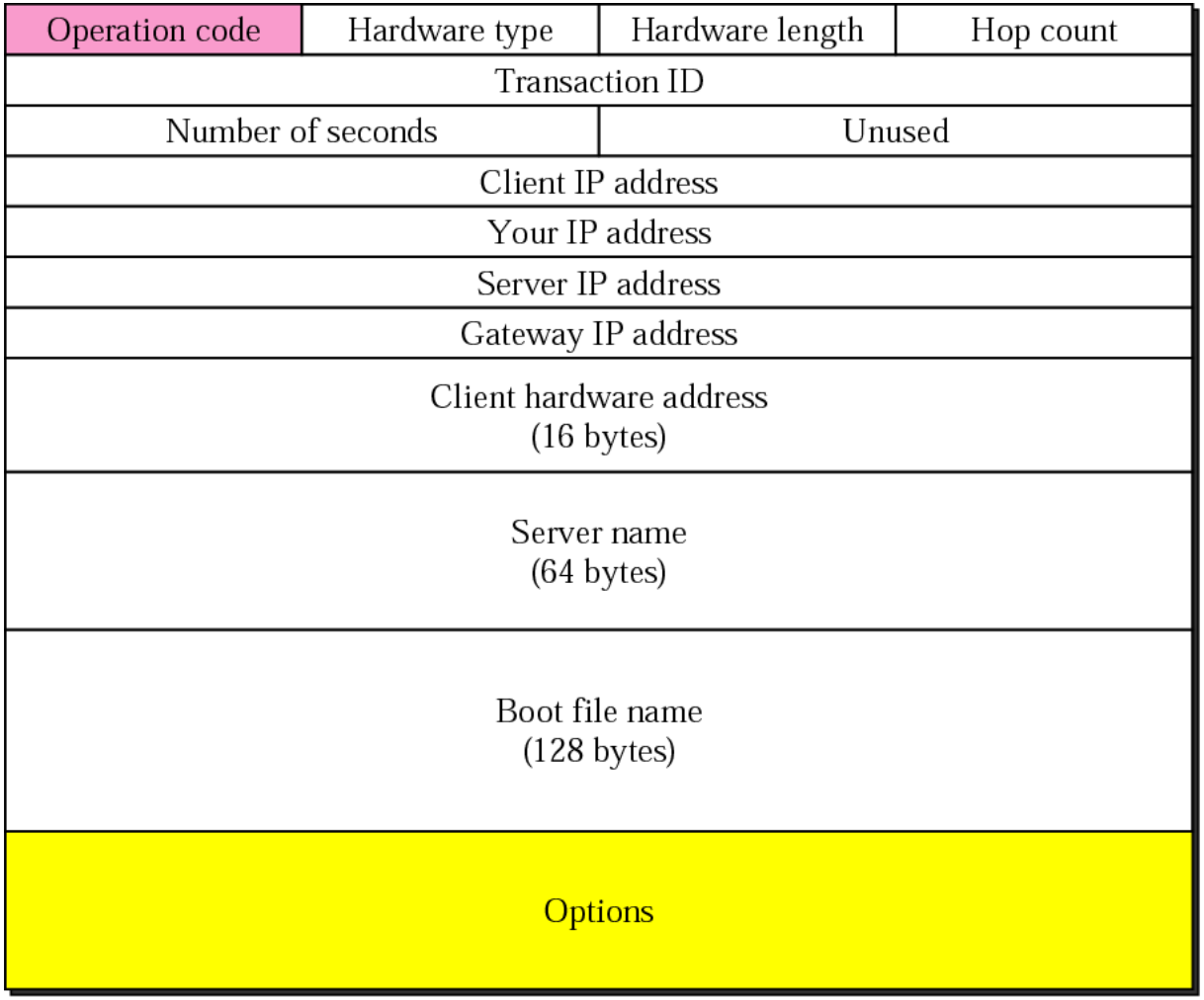
Bootstrap Protocol (BOOTP)



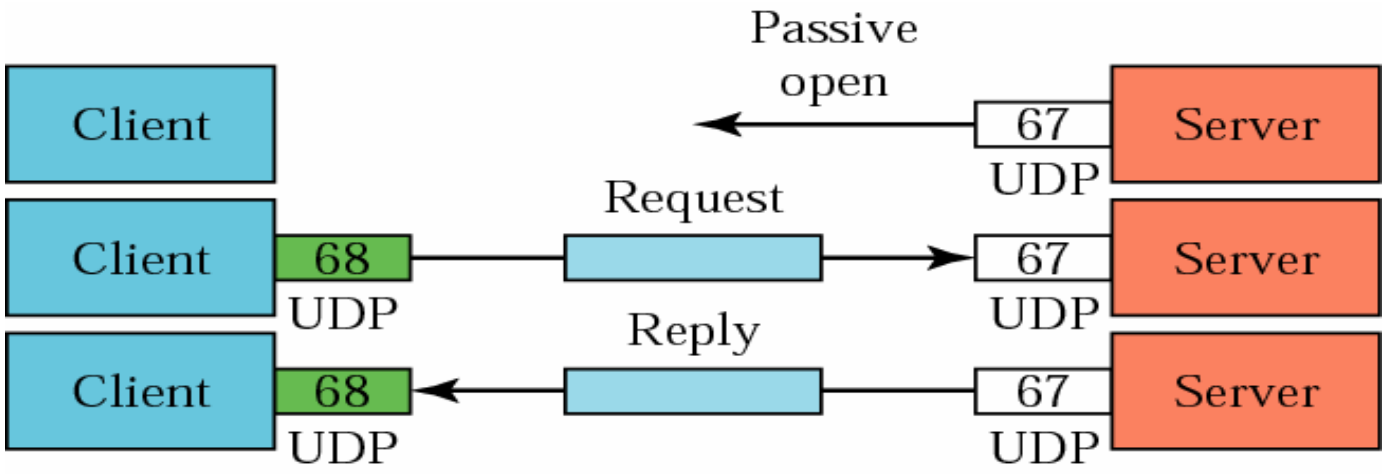
BOOTP

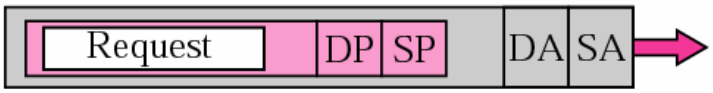
- Diskless terminal
- Discover its own IP address
- Download executable image file
- Small program built in chip
  - BOOTH and TFTP
- Application Protocol
  - Encapsulated in IP and UDP

BOOTP packet format



Operation



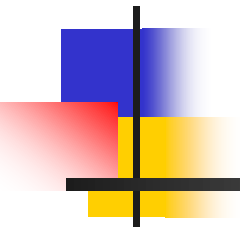


SP: Source port (68)
DP: Destination port (67)

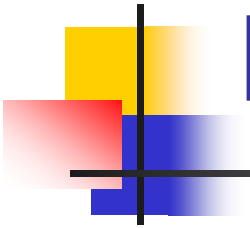
SA: Source address (All 0s)
DA: Destination address (All 1s)

SP: Source port (67)
DP: Destination port (68)

SA: Source address (Server unicast address)
DA: Destination address (All 1s or client unicast address)



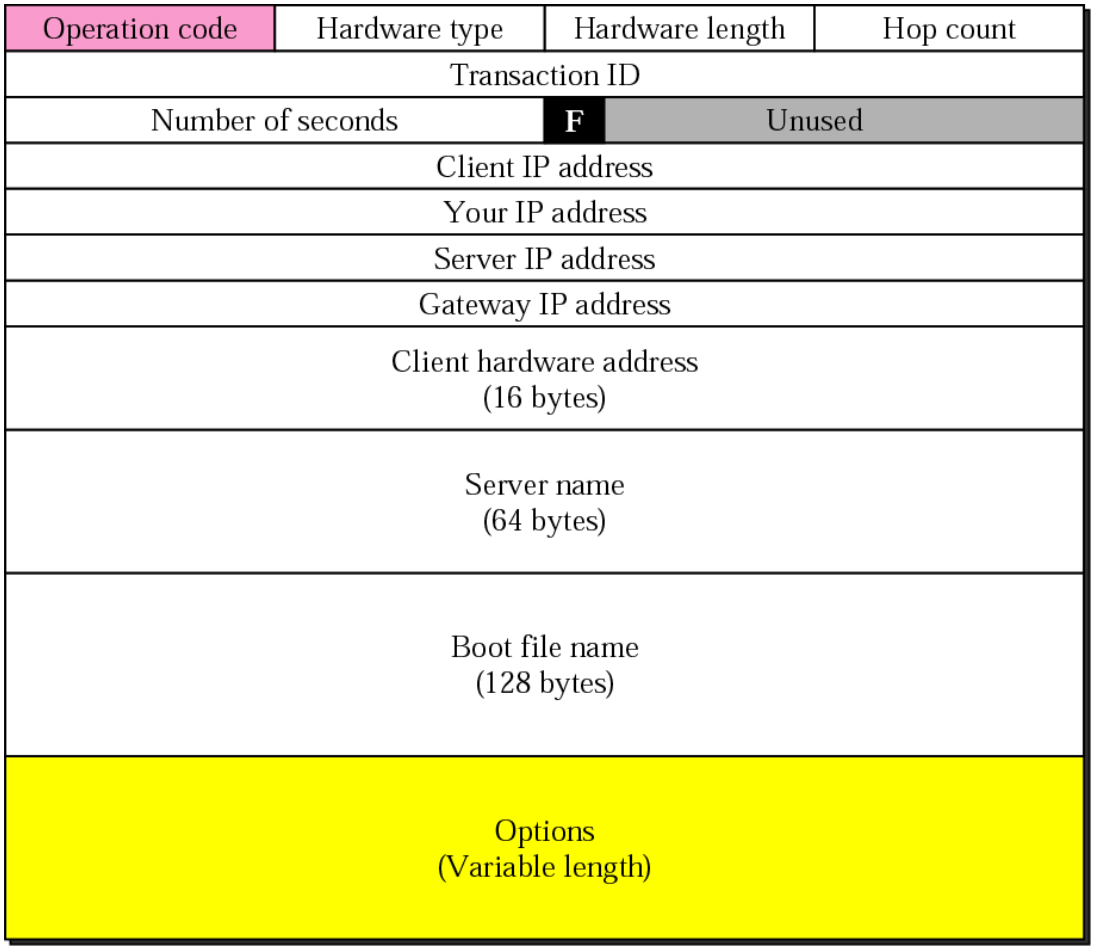
Dynamic Host Configuration Protocol (DHCP)



DHCP

- BOOTP Enhancement
- Same message structure as BOOTP
- Can choose among many DHCP servers

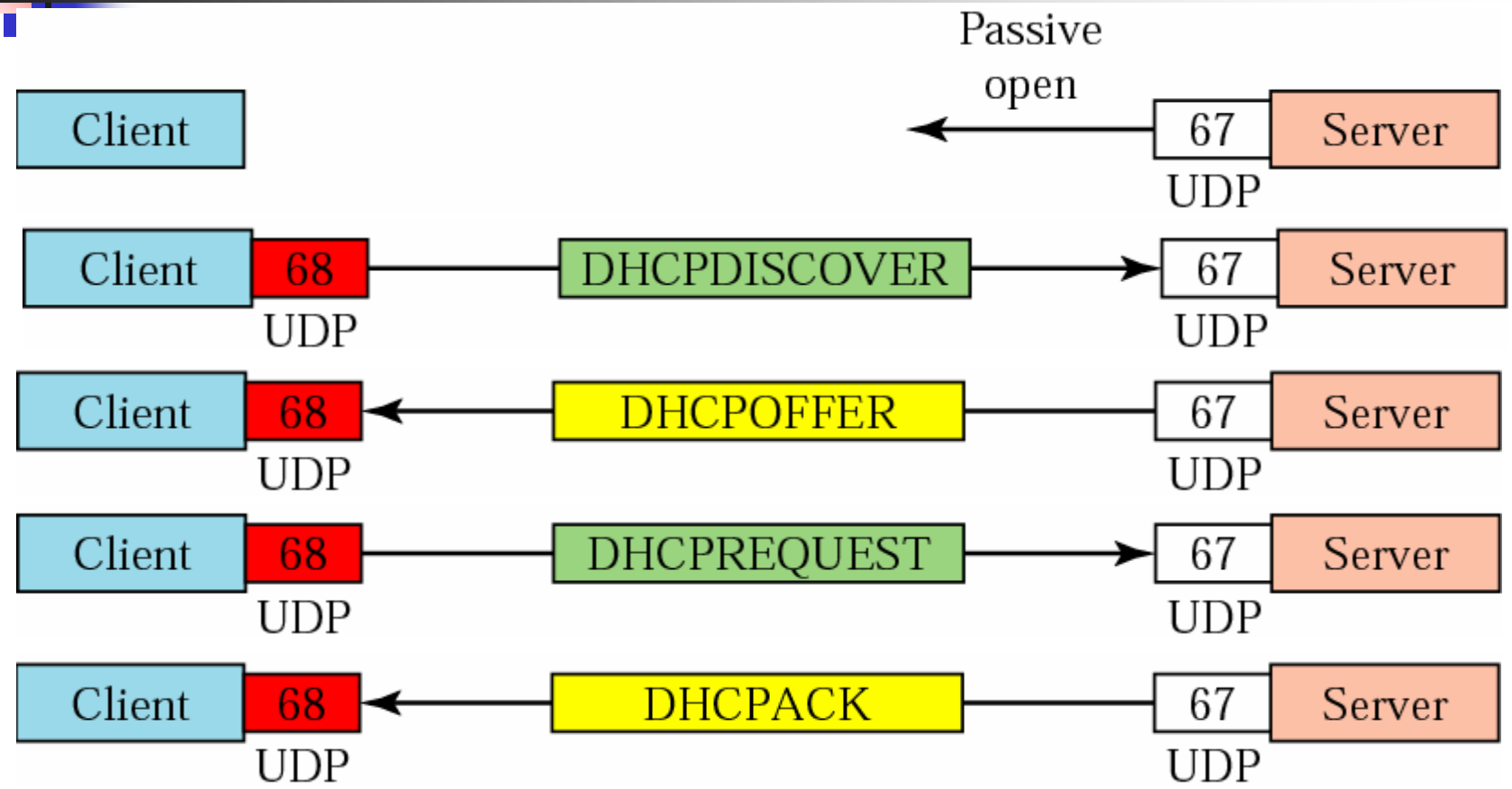
DHCP packet


Operation code
Hardware type
Hardware length
Hop count
Transaction ID
Number of seconds
F
Unused
Client IP address
Your IP address
Server IP address
Gateway IP address
Client hardware address
(16 bytes)
Server name
(64 bytes)
Boot file name
(128 bytes)
Options
(Variable length)

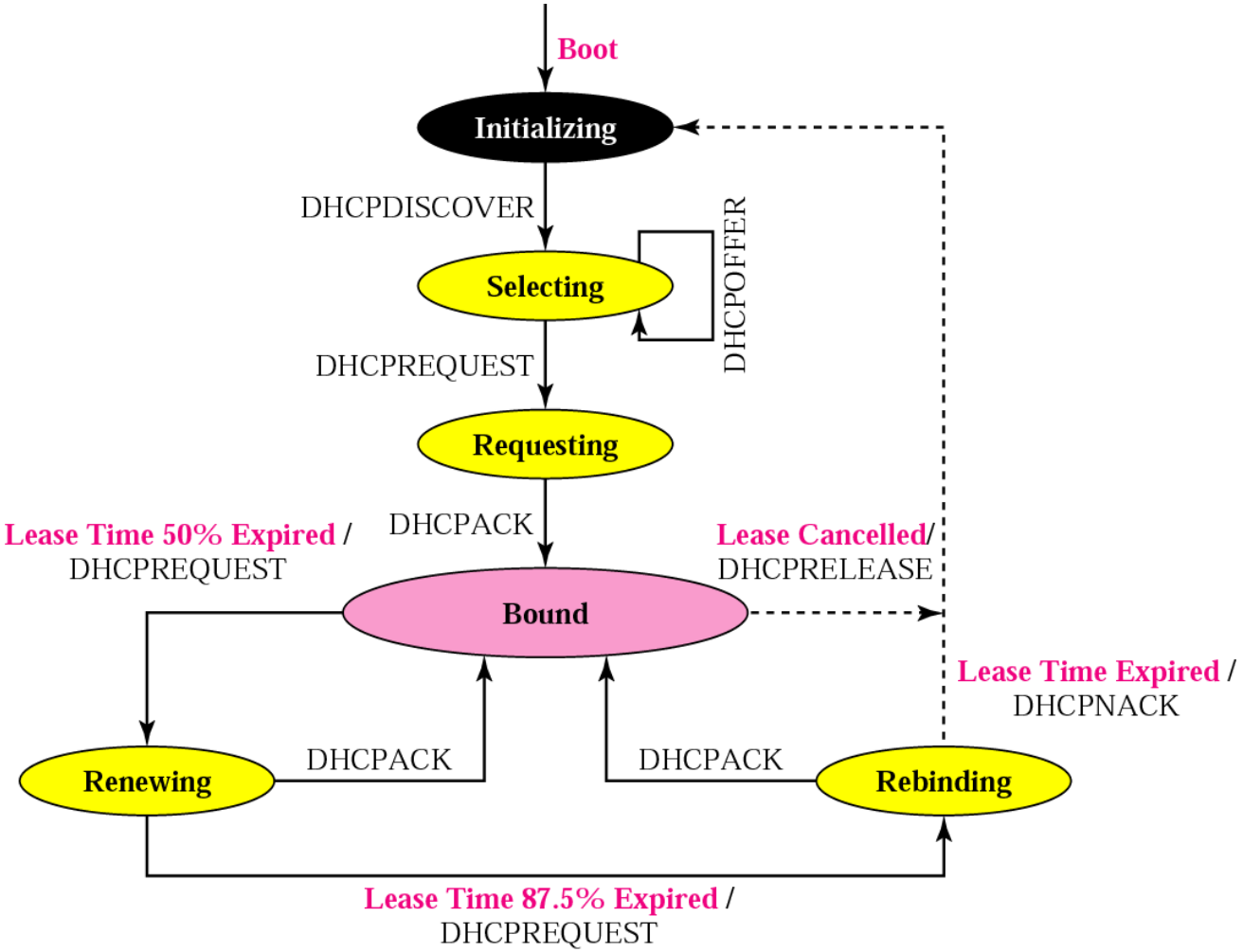
Operation

Ethernet Frame		IP		UDP	DHCP Request	
SRC MAC: MAC A		IP SRC: ?		UDP	CIADDR: ?	GIADDR: ?
DST MAC: FF:FF:FF:FF:FF:FF		IP DST: 255.255.255.255		67	Mask:?	CHADDR: MAC A

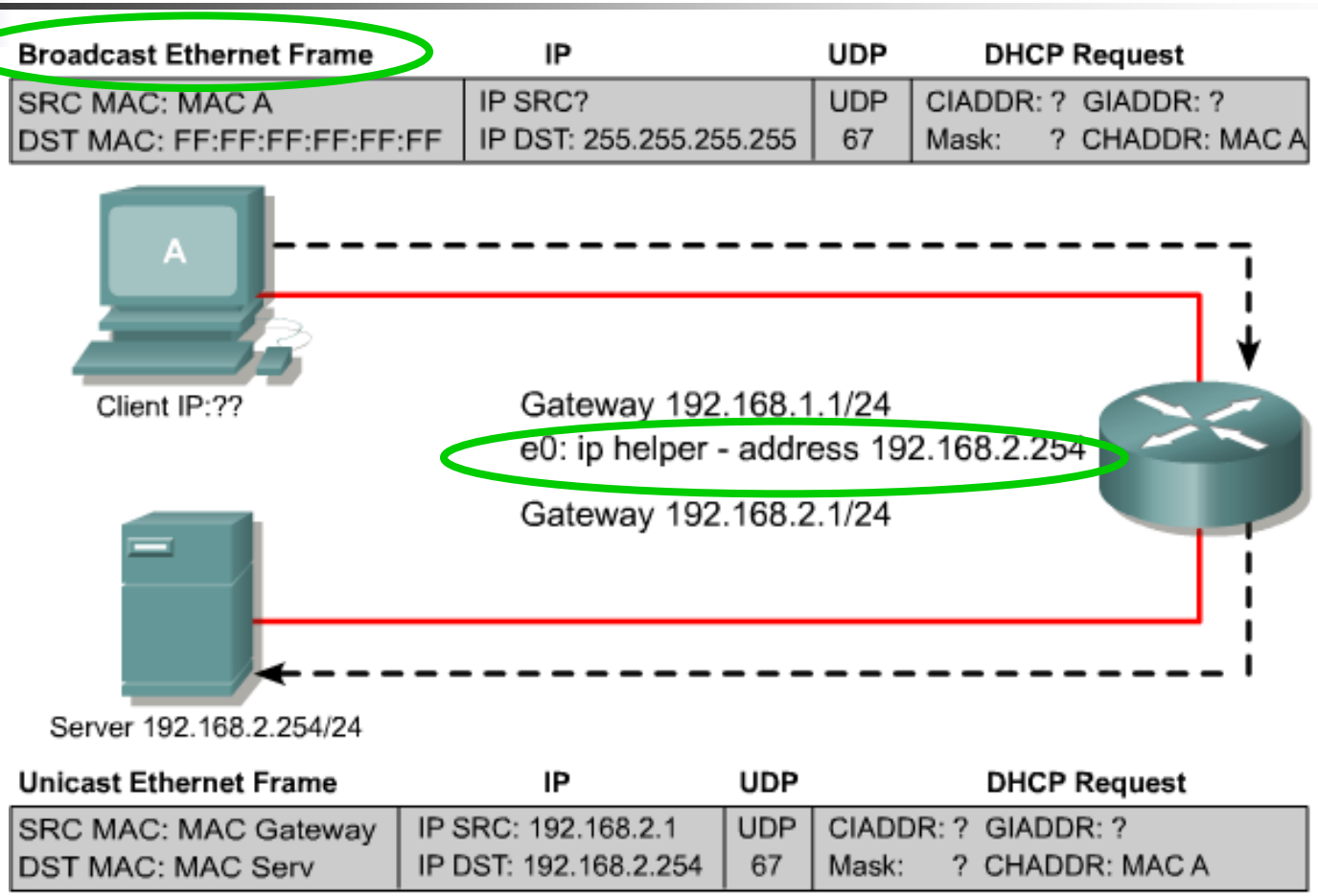
DHCP Message



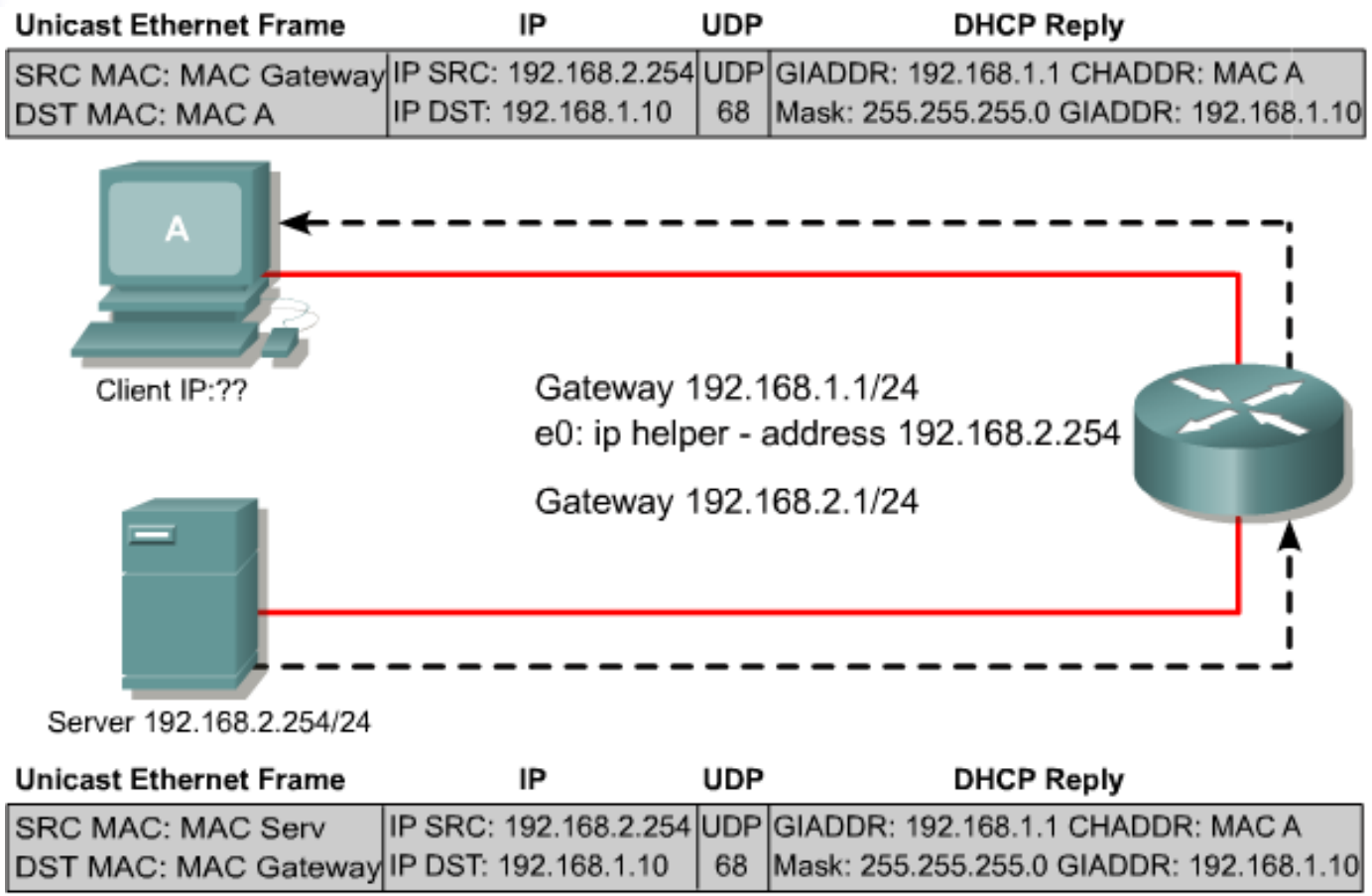
DHCP transition diagram

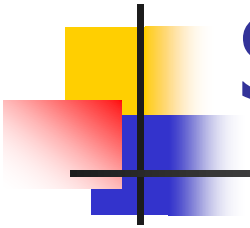


DHCP Relay


Broadcast Ethernet Frame
IP
UDP
DHCP Request
SRC MAC: MAC A
DST MAC: FF:FF:FF:FF:FF:FF
IP SRC?
IP DST: 255.255.255.255
UDP
67
CIADDR: ? GIADDR: ?
Mask: ? CHADDR: MAC A
A
Client IP:??
Gateway 192.168.1.1/24
e0: ip helper - address 192.168.2.254
Gateway 192.168.2.1/24
Server 192.168.2.254/24
Unicast Ethernet Frame
IP
UDP
DHCP Request
SRC MAC: MAC Gateway
DST MAC: MAC Serv
IP SRC: 192.168.2.1
IP DST: 192.168.2.254
UDP
67
CIADDR: ? GIADDR: ?
Mask: ? CHADDR: MAC A

DHCP Relay


Unicast Ethernet Frame
IP
UDP
DHCP Reply
SRC MAC: MAC Gateway
DST MAC: MAC A
IP SRC: 192.168.2.254
IP DST: 192.168.1.10
UDP 68
GIADDR: 192.168.1.1 CHADDR: MAC A
Mask: 255.255.255.0 GIADDR: 192.168.1.10
A
Client IP:??
Gateway 192.168.1.1/24
e0: ip helper - address 192.168.2.254
Gateway 192.168.2.1/24
Server 192.168.2.254/24
Unicast Ethernet Frame
IP
UDP
DHCP Reply
SRC MAC: MAC Serv
DST MAC: MAC Gateway
IP SRC: 192.168.2.254
IP DST: 192.168.1.10
UDP 68
GIADDR: 192.168.1.1 CHADDR: MAC A
Mask: 255.255.255.0 GIADDR: 192.168.1.10



Summary

- IP Fundamental Operation
- Internet Protocol
- Addressing
- Supporting Protocol
  - ARP
  - ICMP: ping + traceroute
  - NAT
  - DHCP