

# TCP/IP Networking

wangth (2017-2020, CC BY-SA)  
? (1996-2016)

交大資工系資訊中心

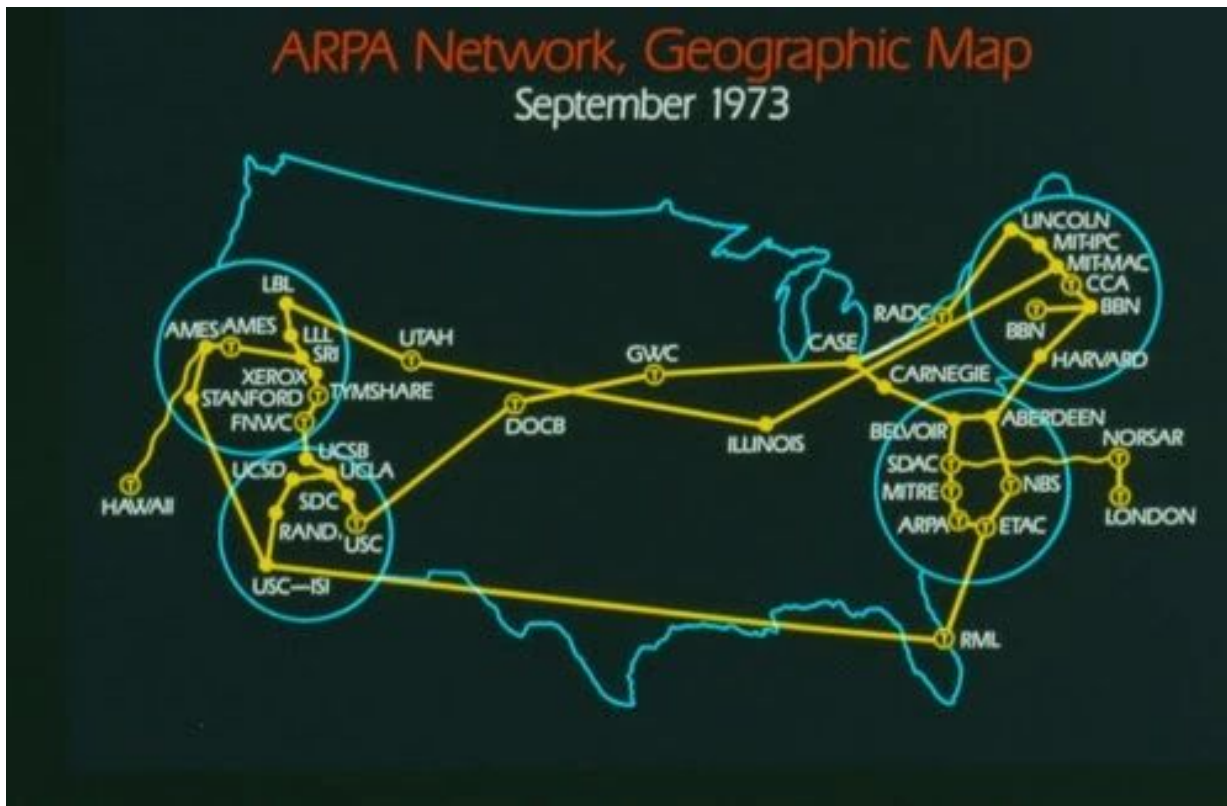
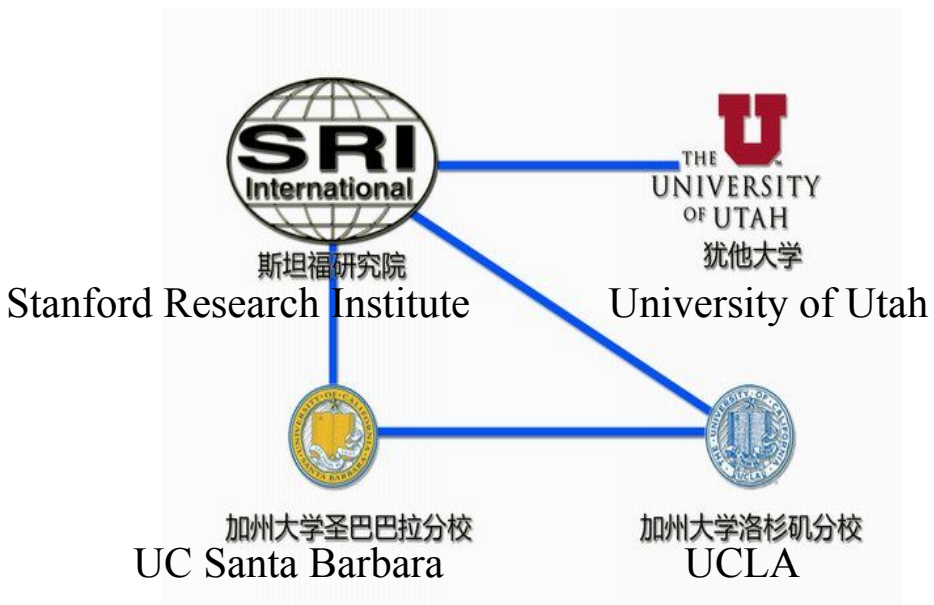
Computer Center of Department of Computer Science, NCTU

# Overview

- Introduction
- Layers of TCP/IP
  - Link Layer
  - Network Layer
  - Transport Layer
  - Application Layer
- Network Interface and Hardware
- Networking
- ARP
- Setting up Network

# Introduction – ARPANET

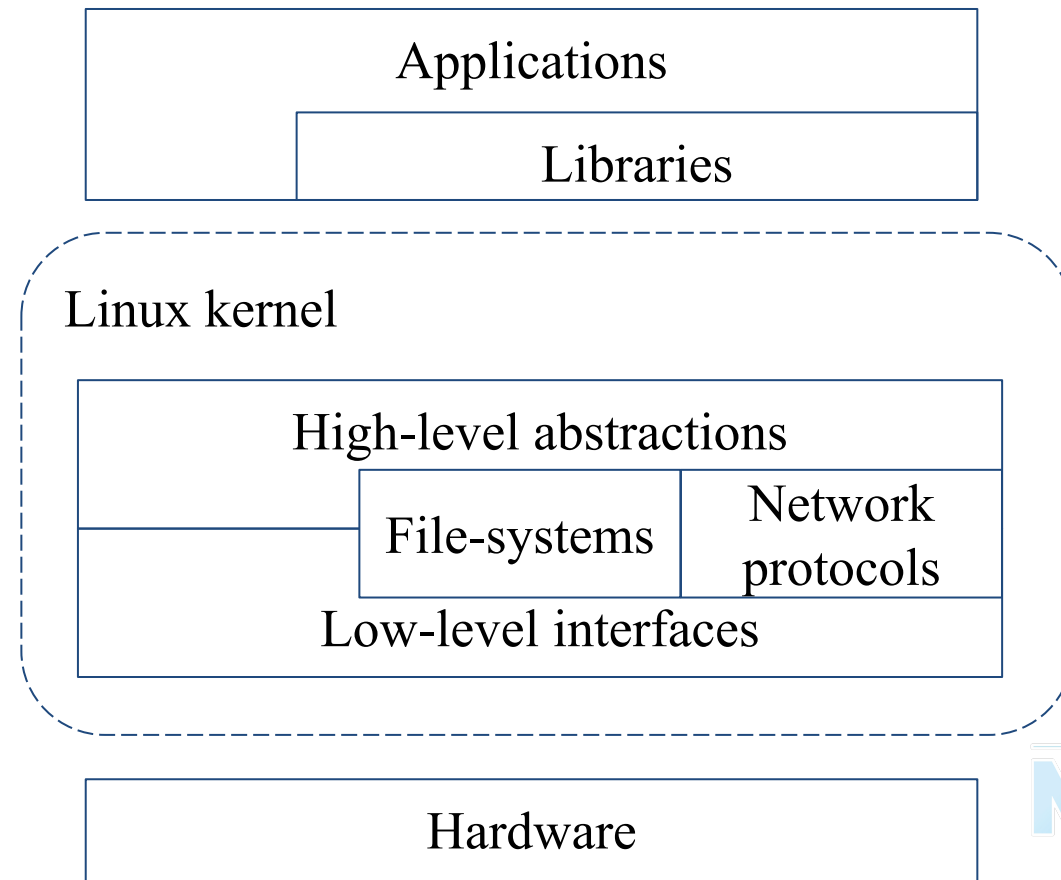
- Advanced Research Projects Agency Network
  - NCP (Network Control Protocol)
    - Allow an exchange of information between separated computers



<https://inventiontourblog.wordpress.com/2015/03/31/internet-advanced-research-project-agency-arpa-develops-the-first-computer-network/>

# Introduction – Why TCP/IP ?

- Transmission Control Protocol / Internet Protocol
- The gap between applications and Network
  - Network
    - 802.3 Ethernet
    - 802.4 Token bus
    - 802.5 Token Ring
    - 802.11 Wireless
    - 802.16 WiMAX
  - Application
    - Reliable
    - Performance



# Introduction – Layers of TCP/IP (1)

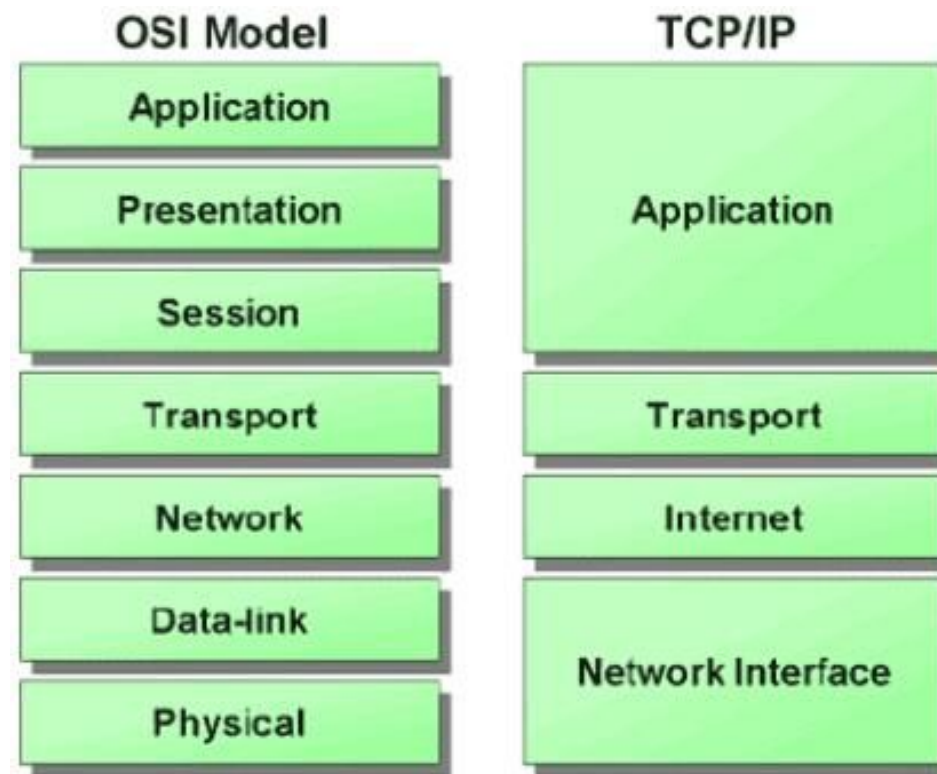
- TCP/IP is a suite of networking protocols
  - 4-layer architecture
    - Link layer (data-link layer)
      - Include device drivers to handle hardware details
    - Network layer (IP)
      - Handle the movement of packets around the network
    - Transport layer (Port)
      - Handle flow of data between hosts
    - Application

Application	Telnet
Transport	TCP, UDP
Network	IP, ICMP, IGMP
Link	Device driver and interface card



# Introduction – Layers of TCP/IP (2)

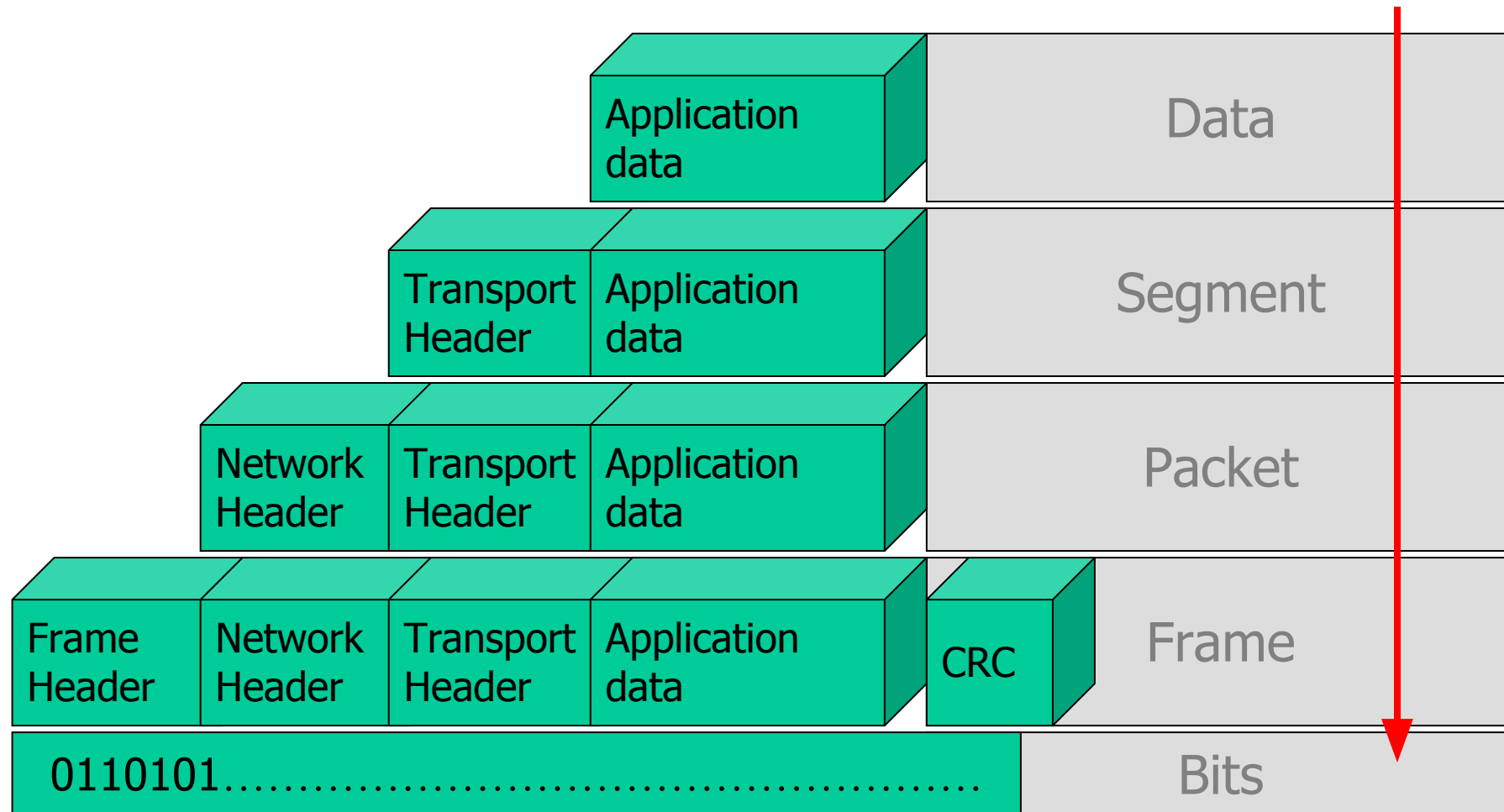
- ISO/OSI Model (International Organization for Standardization / Open System Interconnection Reference Model)
- TCP/IP Model



TCP/IP and the OSI model

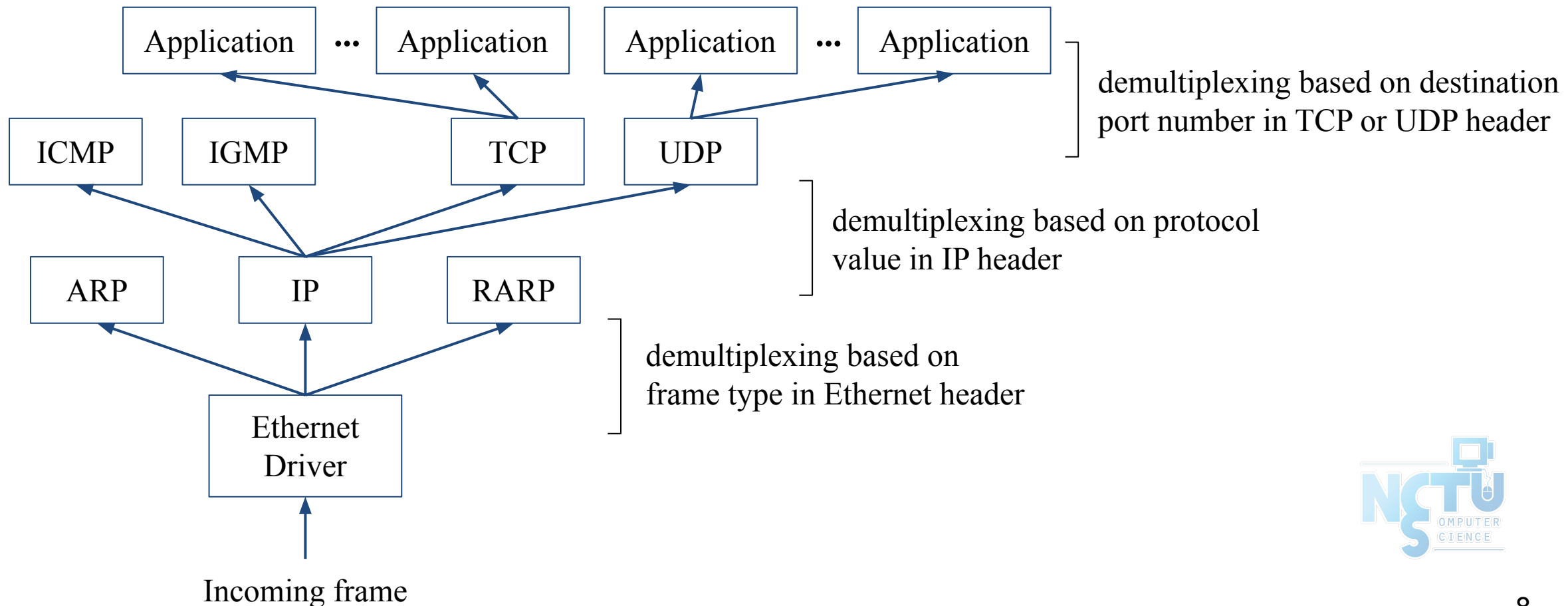
# Introduction – Layers of TCP/IP (3)

- Encapsulation (Multiplexing)
  - Gathering data from multiple sockets, enveloping data with header



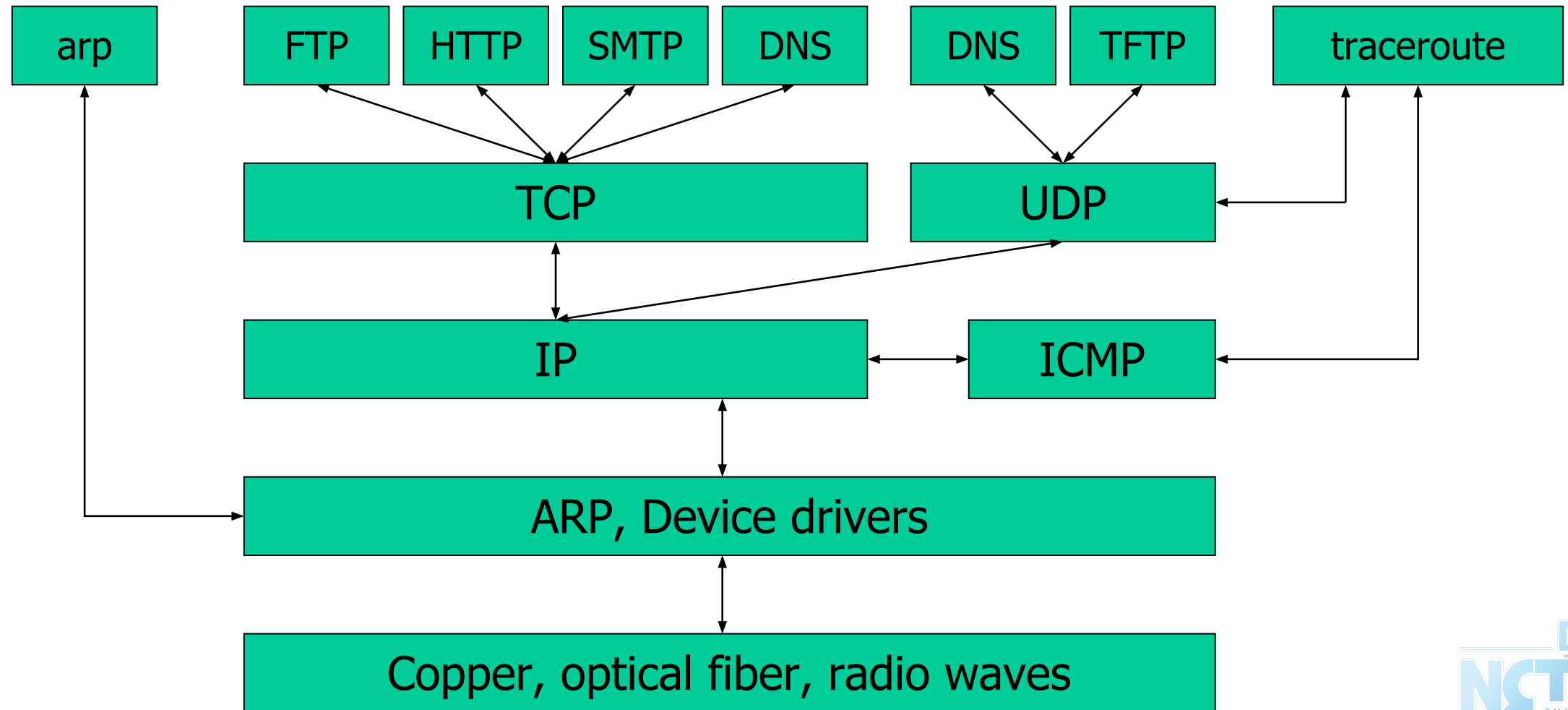
# Introduction – Layers of TCP/IP (4)

- Decapsulation (Demultiplexing)
  - Delivering received segments to correct socket



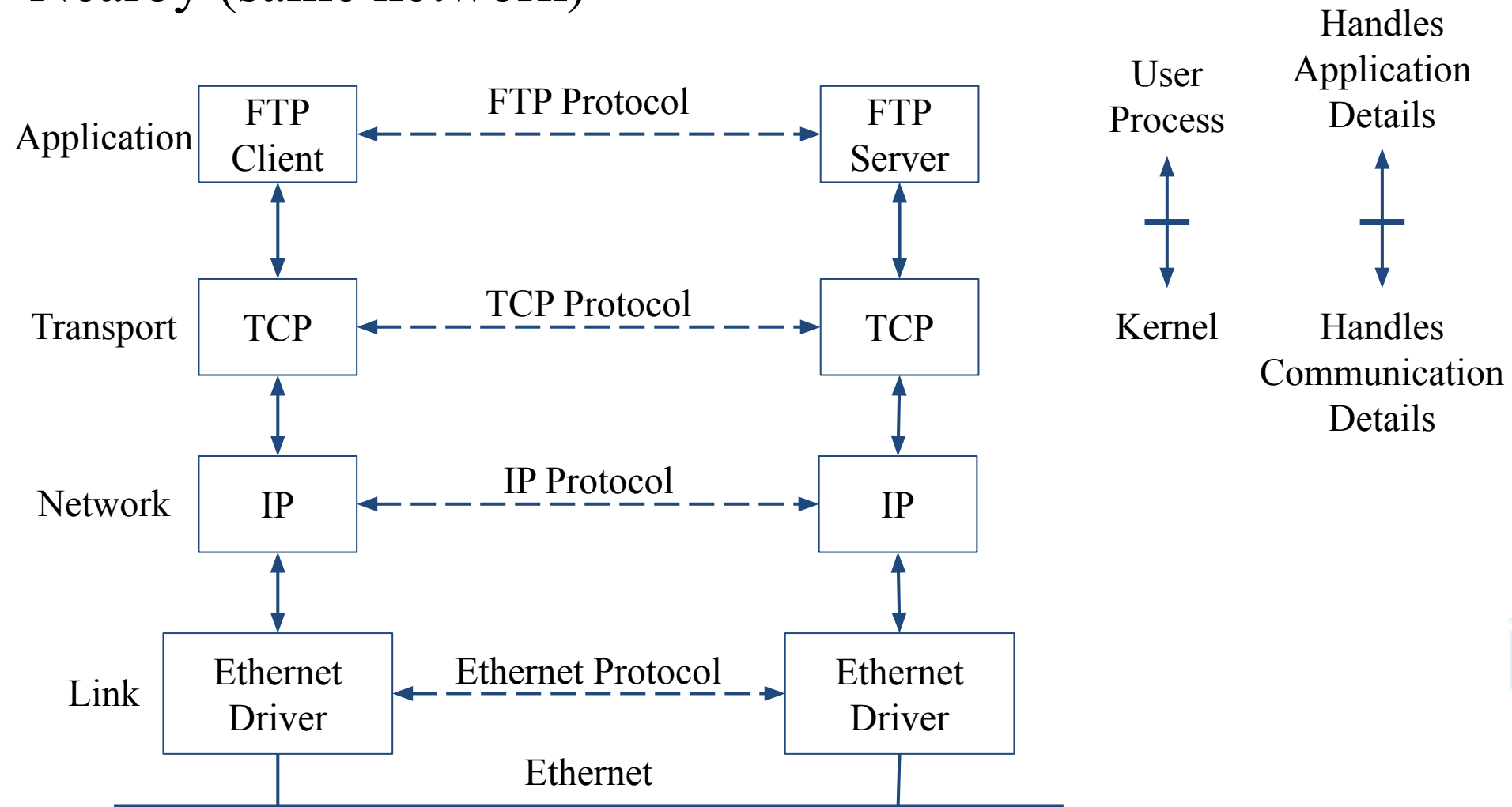


# Introduction – TCP/IP Family



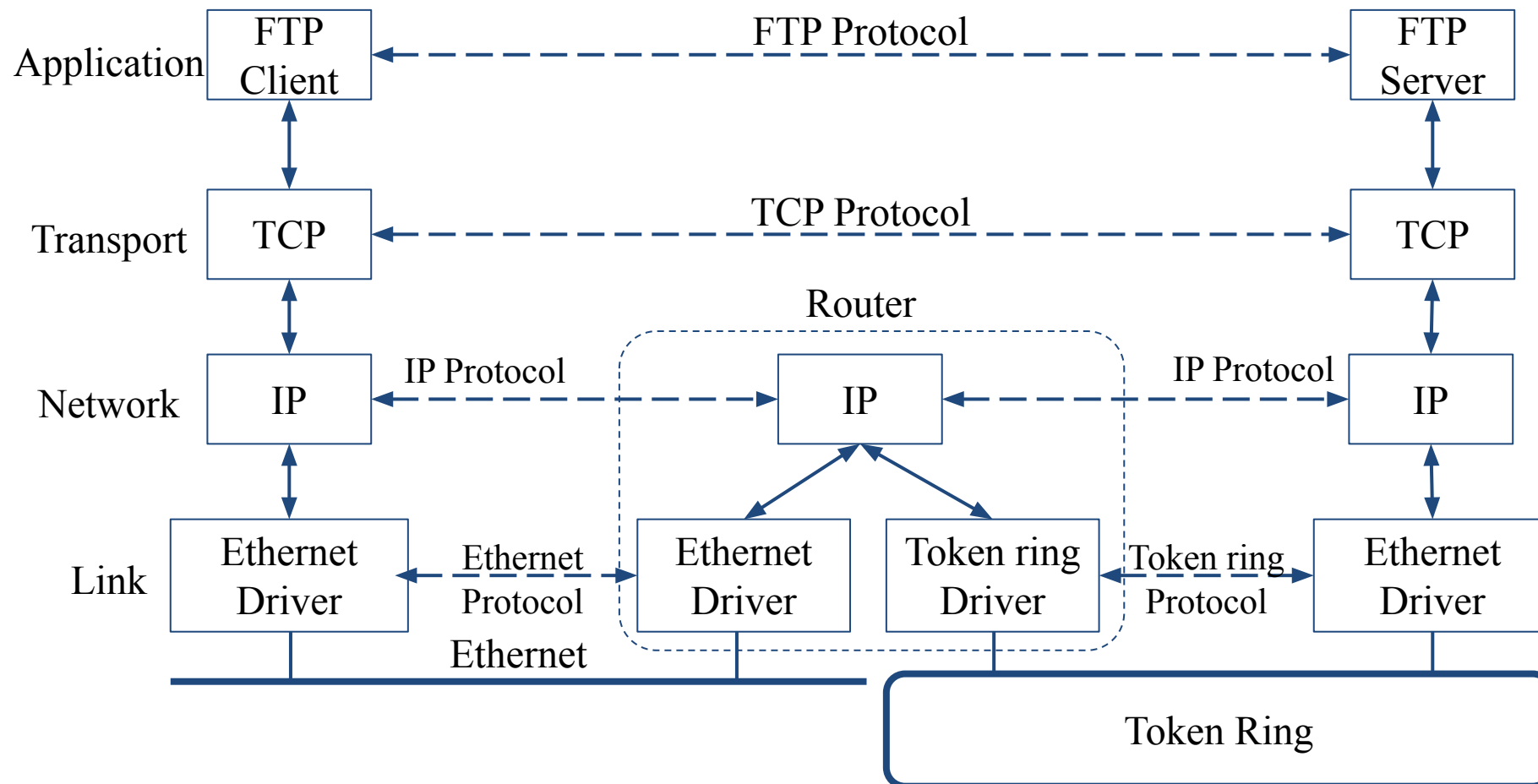
# Introduction – Addressing

- Addressing
  - Nearby (same network)



# Introduction – Addressing

- Addressing
  - Faraway (across network)



# Introduction – Addressing

- IP
  - 32-bits, Unique Internet Address of a host
- Port
  - 16-bits, Uniquely identify application
- MAC Address
  - Media Access Control Address

```
$ ifconfig
sk0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> mtu 1500
    options=b<RXCSUM, TXCSUM, VLAN_MTU>
    inet 140.113.17.215 netmask 0xfffff00 broadcast 140.113.17.255
    inet 140.113.17.221 netmask 0xffffffff broadcast 140.113.17.221
    ether 00:11:d8:06:1e:81
    media: Ethernet autoselect (100baseTX <full-duplex,flag0,flag1>)
    status: active
lo0: flags=8049<UP,LOOPBACK,RUNNING,MULTICAST> mtu 16384
    inet 127.0.0.1 netmask 0xff000000
```

# Link Layer

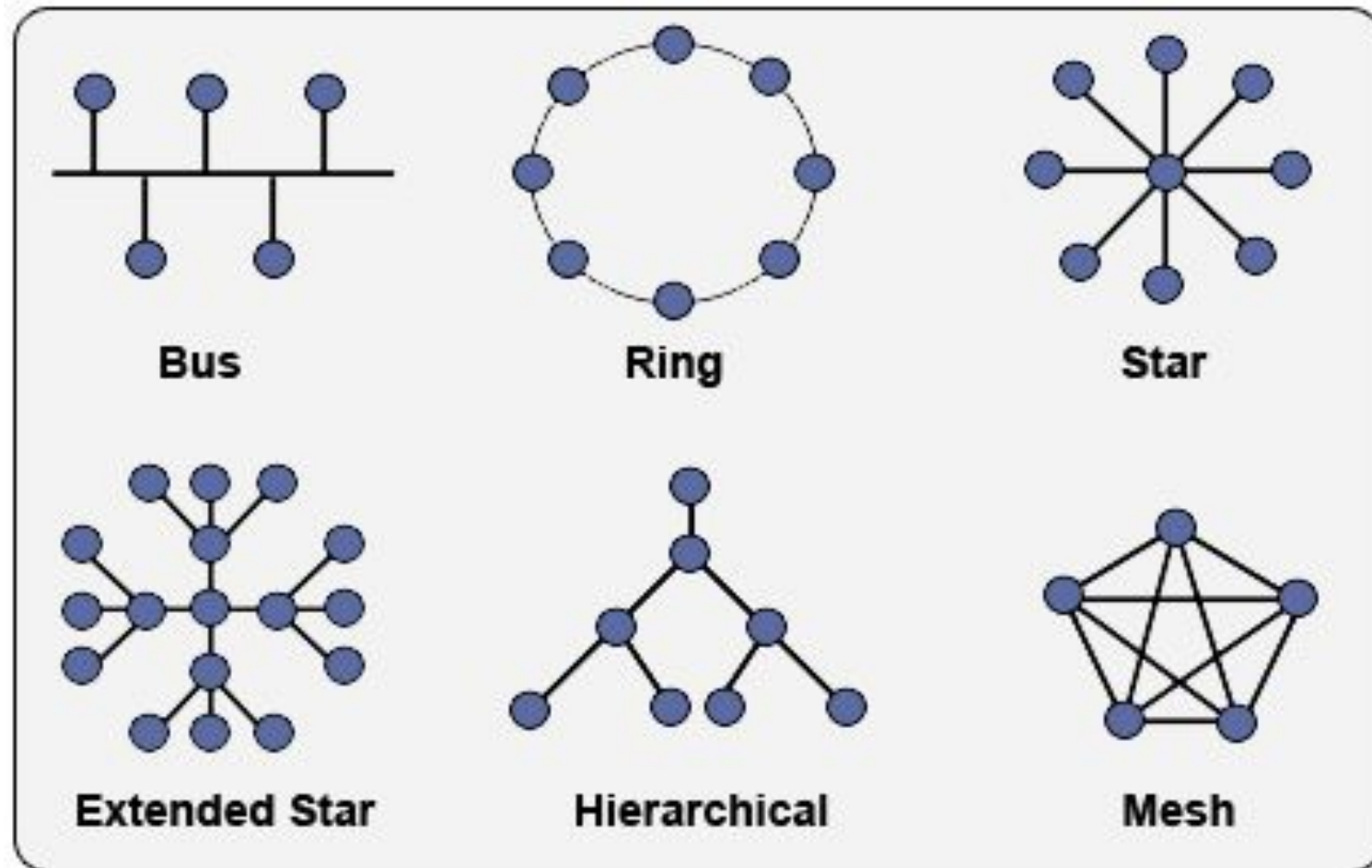
# Network Interface and Hardware

- LAN (Local), WAN (Wide), MAN (Metropolitan)
  - Ethernet, Token-Ring, FDDI
  - PPP, xDSL, ISDN
- Physical Topologies (see next slide)
- Logical Topologies
  - Broadcast, Token-passing
- Common LAN Devices
  - NIC, Repeater, Hub, Bridge, Switch, Router
- Common LAN Media
  - UTP, STP, Coaxial Cable, Fiber Optic Cable



# Network Interface and Hardware

## – Physical Topologies



<https://www.itprc.com/a-guide-to-network-topology/>

# Network Interface and Hardware

## – Media

- Media

- Coaxial Cable

- Thicknet v.s. thinnet
- BNC connector

- Twisted Pair Standards

Pin#	1	2	3	4	5	6	7	8
T568-A	W/G	Green	W/O	Blue	W/Blue	Orange	W/Br	Brown
T568-B	W/O	Orange	W/G	Blue	W/Blue	Green	W/Br	Brown

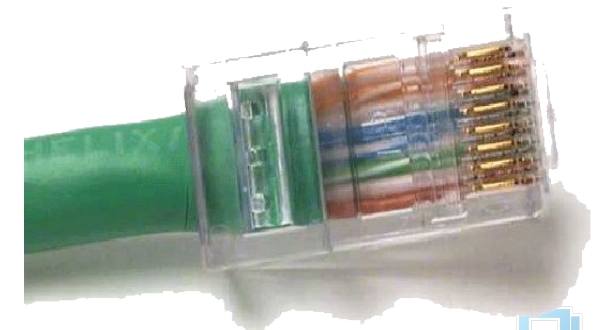
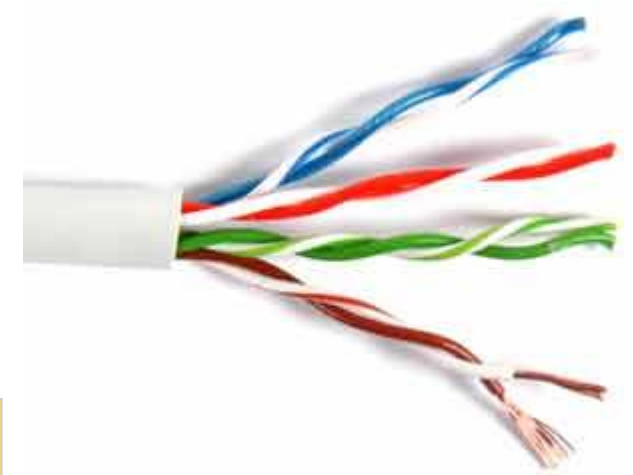
- Straight-through v.s. Crossover
- RJ-45 connector

- Fiber Optic Cable

- Multimode v.s. single mode

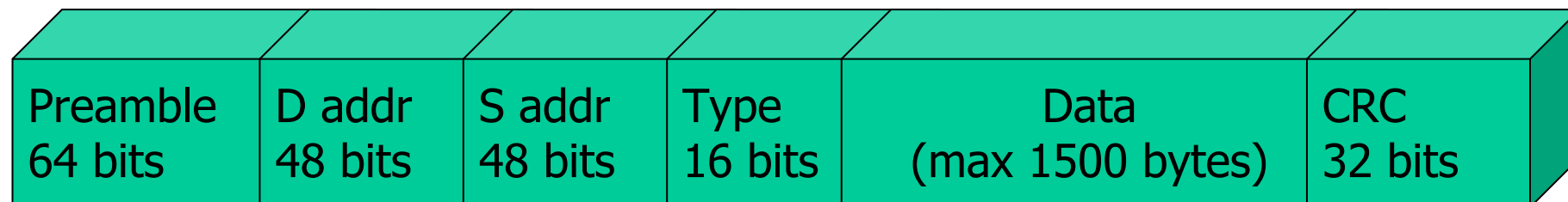
- Wireless

- IrDA, Radio (2.4GHz, 5GHz)



# The Link Layer

- Ethernet: the common LAN
  - 10 Mb/s □ 100 Mb/s □ 1Gb/s
  - 802.3 □ 802.3u □ 802.3z
  - CSMA/CD (Carrier Sense Multiple Access/Collision Detect)
- Ethernet Address (48bits)
  - 00:80:c8:92:0e:e1
- Ethernet Frame
  - Ethernet MTU (Maximum Transmission Unit) is 1500 bytes
  - IP fragmentation
  - Path MTU
    - MTU of various physical device



# Network Layer

# The Network Layer

- Path Determination
  - The Internet Protocol (IP)
    - IP address (32 bits)
- Topics
  - IP Address
  - Subnetting and netmask
  - Address types
  - Routing

# The Network Layer – IP Address

- 32-bit long
  - Network part
    - Identify a logical network
  - Host part
    - Identify a machine on certain network
- IP address category
- E.g. :
  - NCTU
    - Class B address: 140.113.0.0
    - Network ID: 140.113
    - Number of hosts:  $256 * 256 = 65536$

Class	1st byte	Format	Comments
A	1-126	N.H.H.H	Very early networks, or reserved for DOD
B	128-191	N.N.H.H	Large sites, usually subnetted, were to get
C	192-223	N.N.N.H	Easy to get, often obtained in sets
D	224-239	-	Multicast addresses, not permanently assigned
E	240-254	-	Experimental addresses



# The Network Layer

## – Subnetting and Netmask (1)

- Subnetting
  - Borrow some bits from network ID to extends hosts ID
  - E.g.,
    - Class B address : 140.113.0.0  
= 256 Class C-like IP addresses  
in N.N.N.H subnetting method
    - 140.113.209.0 subnet
- Netmask
  - Specify how many bits of network-ID are used for network-ID
  - Continuous 1 bits form the network part
  - E.g.:
    - 255.255.255.0 in NCTU-CS example
      - 256 hosts available
    - 255.255.255.248 in ADSL example
      - Only 8 hosts available

# The Network Layer

## – Subnetting and Netmask (2)

- How to determine your network ID?
  - Bitwise-AND IP and netmask
  - E.g.,
    - $140.113.214.37 \ \& \ 255.255.255.0 \ \square \ 140.113.214.0$
    - $140.113.209.37 \ \& \ 255.255.255.0 \ \square \ 140.113.209.0$
  
    - $140.113.214.37 \ \& \ 255.255.0.0 \ \square \ 140.113.0.0$
    - $140.113.209.37 \ \& \ 255.255.0.0 \ \square \ 140.113.0.0$
  
    - $211.23.188.78 \ \& \ 255.255.255.248 \ \square \ 211.23.188.72$ 
      - $78 = 01001110$
      - $78 \ \& \ 248 = 01001110 \ \& \ 11111000 = 72$

# The Network Layer

## – Subnetting and Netmask (3)

- In a subnet, not all IP are available
- The first one IP □ network ID
- The last one IP □ broadcast address
- E.g.,

```
Netmask 255.255.255.0  
140.113.209.32/24
```

```
140.113.209.0    => network ID  
140.113.209.255 => broadcast address  
1 ~ 254, total 254 IPs are usable
```

```
Netmask 255.255.255.252  
211.23.188.78/29
```

```
211.23.188.72 => network ID  
211.23.188.79 => broadcast address  
73 ~ 78, total 6 IPs are usable
```

# The Network Layer

## – Subnetting and Netmask (4)

- The smallest subnetting
  - Network portion : 30 bits
  - Host portion : 2 bits

=> 4 hosts, but only 2 IPs are available
- ipcalc
  - \$ pkg install ipcalc
  - /usr/ports/net-mgmt/ipcalc

```
$ ipcalc 140.113.235.100/28

Address:    140.113.235.100      10001100.01110001.11101011.0110 0100
Netmask:    255.255.255.240 = 28 11111111.11111111.11111111.1111 0000
Wildcard:   0.0.0.15             00000000.00000000.00000000.0000 1111
=>
Network:    140.113.235.96/28    10001100.01110001.11101011.0110 0000
HostMin:    140.113.235.97       10001100.01110001.11101011.0110 0001
HostMax:    140.113.235.110      10001100.01110001.11101011.0110 1110
Broadcast:  140.113.235.111      10001100.01110001.11101011.0110 1111
Hosts/Net:  14                    Class B
```

# The Network Layer

## – Subnetting and Netmask (5)

- Network configuration for various lengths of netmask

Length	Host bits	Hosts/net	Dec. netmask	Hex netmask
/20	12	4094	255.255.240.0	0xFFFFF000
/21	11	2046	255.255.248.0	0xFFFFF800
/22	10	1022	255.255.252.0	0xFFFFFC00
/23	9	510	255.255.254.0	0xFFFFFE00
/24	8	254	255.255.255.0	0xFFFFF00
/25	7	126	255.255.255.128	0xFFFFF80
/26	6	62	255.255.255.192	0xFFFFFC0
/27	5	30	255.255.255.224	0xFFFFFE0
/28	4	14	255.255.255.240	0xFFFFF0
/29	3	6	255.255.255.248	0xFFFFF8
/30	2	2	255.255.255.252	0xFFFFFC

# The Network Layer

## – IP address crisis

- IP address crisis
  - Run out of class B address
    - The most desirable ones for moderately large organizations
  - IP address were being allocated on a FCFS
    - With no locality of reference
- Solutions
  - Short term
    - Subnetting and CIDR (classless inter-domain routing)
    - NAT (network address translation)
  - Long term
    - IPv6



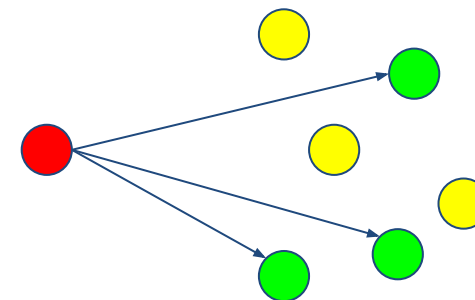
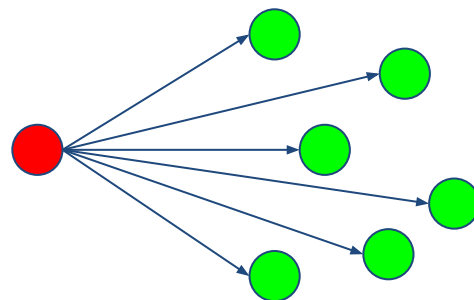
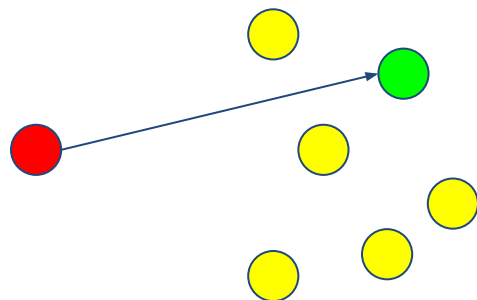
# The Network Layer

## – NAT (1)

- Private Address

- Packets with private address will not go out to the Internet
- 3 private address ranges
  - Depend on the size of your organization

IP class	From	To	CIDR range
Class A	10.0.0.0	10.255.255.255	10.0.0.0/8
Class B	172.16.0.0	172.31.255.255	172.16.0.0/12
Class C	192.168.0.0	192.168.255.255	192.168.0.0/16



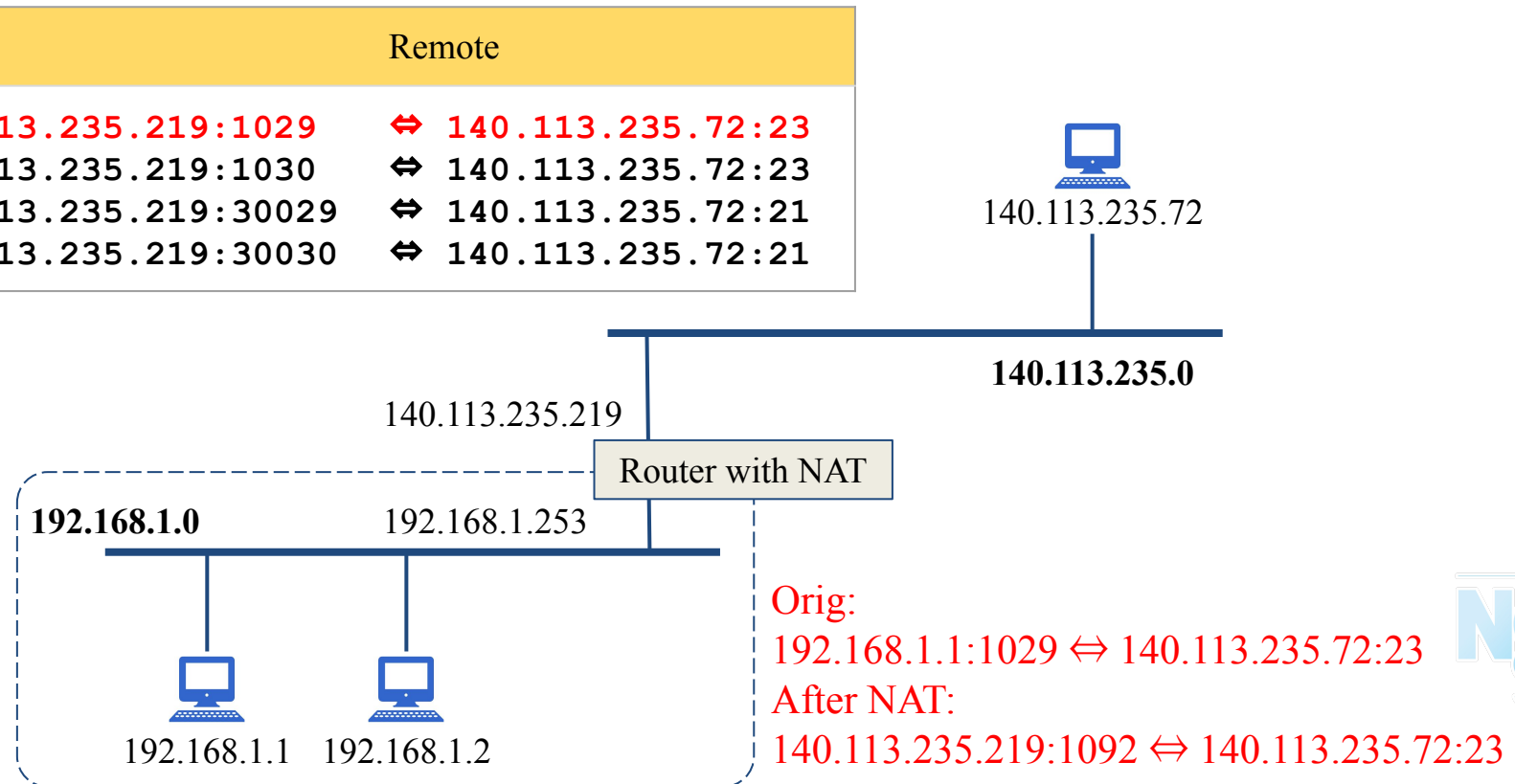
# The Network Layer

## - NAT (2)

- NAT
  - Network Address Translation
  - Allow hosts using private address to talk with outside

### NAT mapping table

Orig	Alias	Remote
192.168.1.1:1029	140.113.235.219:1029	↔ 140.113.235.72:23
192.168.1.1:1030	140.113.235.219:1030	↔ 140.113.235.72:23
192.168.1.2:1029	140.113.235.219:30029	↔ 140.113.235.72:21
192.168.1.2:1030	140.113.235.219:30030	↔ 140.113.235.72:21



# The Network Layer

## – Routing (1)

- Goal: Direct a packet closer to the destination
- Flat v.s. Hierarchical
- Routing table
  - Routing information (which kind of packets to which way)
  - Rule-based information
  - Kernel will pick the most suitable way to route the packets

```
$ netstat -rn
Routing tables

Internet:
Destination          Gateway              Flags      Refs      Use  Netif  Expire
default              140.113.17.254      UGS        0    4439610  dc0
127.0.0.1            127.0.0.1          UH         0       3887   lo0
140.113.17/24        link#1             UC         0         0     dc0
140.113.17.209       00:0d:61:21:02:54  UHLW       1         38     dc0    477
140.113.17.212       00:90:96:23:8f:7d  UHLW       1    22558   lo0
140.113.17.215       00:11:d8:06:1e:81  UHLW       1         17     dc0   1188
```

# The Network Layer

## – Routing (2)

- Static route
  - Statically configured by "route" command
  - E.g.:
    - \$ route add default 140.113.235.254
    - \$ route add 192.168.1.0/24 192.168.1.254
- Dynamic route
  - gated

# The Network Layer

## - Routing (3)

- "ping -R" and "traceroute"

```
$ ping -c 1 -R www.nctu.edu.tw
PING www.nctu.edu.tw (140.113.250.5): 56 data bytes
64 bytes from 140.113.250.5: icmp_seq=0 ttl=61 time=2.249 ms
RR:   ProjE27-253.NCTU.edu.tw (140.113.27.253)
      140.113.0.57
      CC250-gw.NCTU.edu.tw (140.113.250.253)
      www.NCTU.edu.tw (140.113.250.5)
      www.NCTU.edu.tw (140.113.250.5)
      140.113.0.58
      ProjE27-254.NCTU.edu.tw (140.113.27.254)
      e3rtn.csie.nctu.edu.tw (140.113.17.254)
      chbsd.csie.nctu.edu.tw (140.113.17.212)

--- www.nctu.edu.tw ping statistics ---
1 packets transmitted, 1 packets received, 0% packet loss
round-trip min/avg/max/stddev = 2.249/2.249/2.249/0.000 ms

$ traceroute www.nctu.edu.tw
traceroute to www.nctu.edu.tw (140.113.250.5), 64 hops max, 40 byte packets
 1  e3rtn (140.113.17.254)  0.524 ms  0.406 ms  0.512 ms
 2  ProjE27-254.NCTU.edu.tw (140.113.27.254)  0.574 ms  0.501 ms  0.422 ms
 3  140.113.0.58 (140.113.0.58)  0.487 ms  0.583 ms  0.541 ms
 4  www.NCTU.edu.tw (140.113.250.5)  0.673 ms  0.611 ms  0.621 ms
```

# Transport Layer



# The Transport Layer – ports

- 16-bits number
- Preserve ports
  - 1 ~ 1024 (root access only)
- Well-known port
  - /etc/services

chargen	19/tcp	ttytst source	#Character Generator
chargen	19/udp	ttytst source	#Character Generator
ftp-data	20/tcp	#File Transfer [Default Data]	
ftp-data	20/udp	#File Transfer [Default Data]	
ftp	21/tcp	#File Transfer [Control]	
ftp	21/udp	#File Transfer [Control]	
ssh	22/tcp	#Secure Shell Login	
ssh	22/udp	#Secure Shell Login	
telnet	23/tcp		
telnet	23/udp		

IANA Service Name and Transport Protocol Port Number Registry

<https://www.iana.org/assignments/service-names-port-numbers>

# The Transport Layer

- UDP v.s. TCP

Function	UDP	TCP
Connection-oriented	No	Yes
Message boundaries	Yes	No
Data checksum	Optional	Yes
Positive acknowledgement	No	Yes
Time-out and retransmit	No	Yes
Duplicate detection	No	Yes
Sequencing	No	Yes
Flow control	No	Yes

# The Transport Layer – useful commands

- tcpdump, sniffit, trafshow, netstat -s

```
$ tcpdump -n host 140.113.235.131
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on dc0, link-type EN10MB (Ethernet), capture size 96 bytes
11:25:50.996542 IP 140.113.17.212.61233 > 140.113.235.131.22: P
266166194:266166226(32) ack 938637316 win 33304 <nop,nop,timestamp 3368918203
130908112>
11:25:50.998247 IP 140.113.235.131.22 > 140.113.17.212.61233: P 1:33(32) ack 32
win 33304 <nop,nop,timestamp 134993614 3368918203>
11:25:50.998396 IP 140.113.235.131.22 > 140.113.17.212.61233: P 33:65(32) ack 32
win 33304 <nop,nop,timestamp 134993614 3368918203>
11:25:50.998438 IP 140.113.17.212.61233 > 140.113.235.131.22: . ack 65 win 33288
<nop,nop,timestamp 3368918205 134993614>
11:26:36.935422 IP 140.113.17.212 > 140.113.235.131: ICMP echo request, id 28124,
seq 0, length 64
11:26:36.935761 IP 140.113.235.131 > 140.113.17.212: ICMP echo reply, id 28124,
seq 0, length 64
^C
6 packets captured
697 packets received by filter
0 packets dropped by kernel
```

# The Application Layer

- The Client-Server Model
  - Port Numbers:
    - /etc/services
    - The first 1024 ports are reserved ports
  - Internet Services
    - inetd and /etc/inetd.conf
  - DNS

# The Application Layer – inted

- inetd - internet "super-server"
  - add `inetd_enable="YES"` into `/etc/rc.conf`

`/etc/inetd.conf`

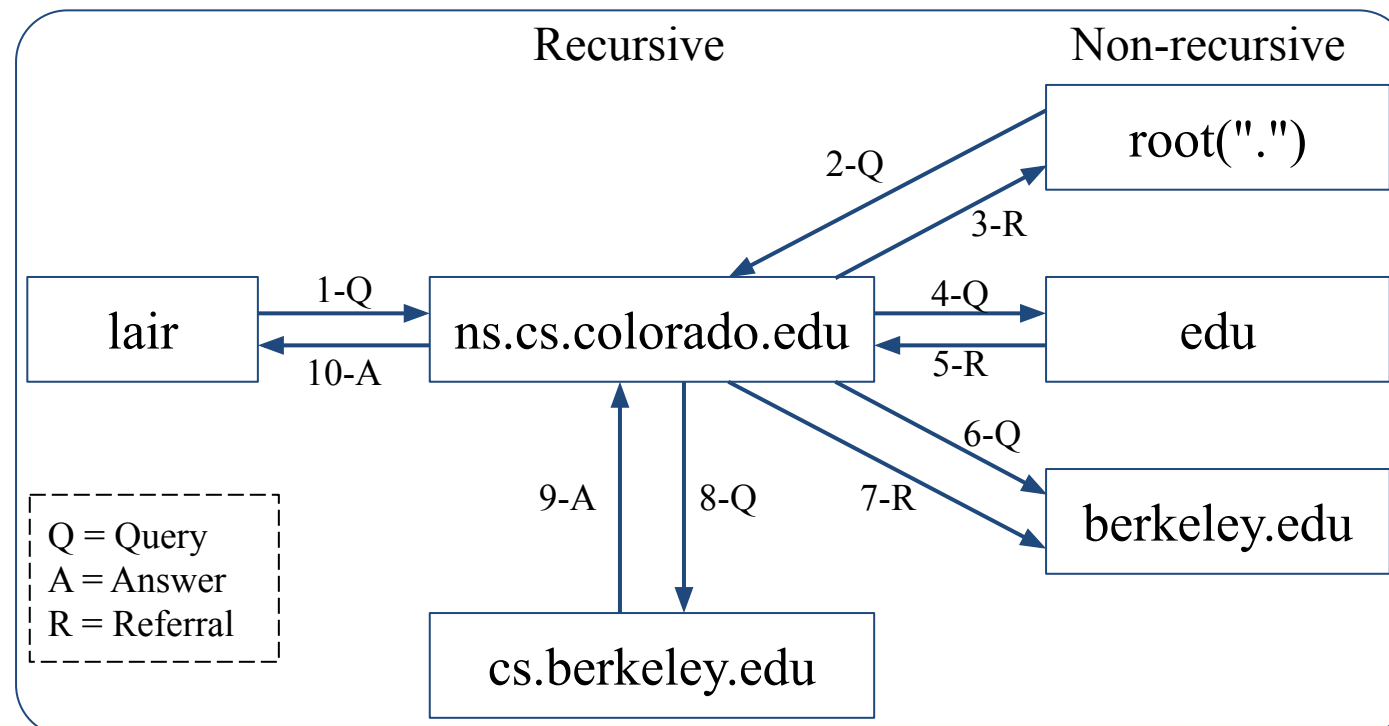
```
daytime  stream  tcp  nowait  root  internal
ftp       stream  tcp  nowait  root  /usr/libexec/ftpd      ftpd -l
ssh      stream  tcp  nowait  root  /usr/sbin/sshd        sshd -i -4
telnet   stream  tcp  nowait  root  /usr/libexec/telnetd   telnetd
pop3     stream  tcp  nowait  root  /usr/local/libexec/popper popper
```

`/etc/services`

```
daytime  13/tcp
ftp-data 20/tcp  #File Transfer [Default Data]
ftp      21/tcp  #File Transfer [Control]
ssh      22/tcp  #Secure Shell Login
telnet   23/tcp
pop3     110/tcp #Post Office Protocol - Version 3
```

# The Application Layer – DNS

- Domain Name System
  - Record IP-hostname mapping
  - DNS query
    - "what is the IP of vangogh.cs.berkeley.edu" from lair.cs.colorado.edu
  - Hierarchical architecture

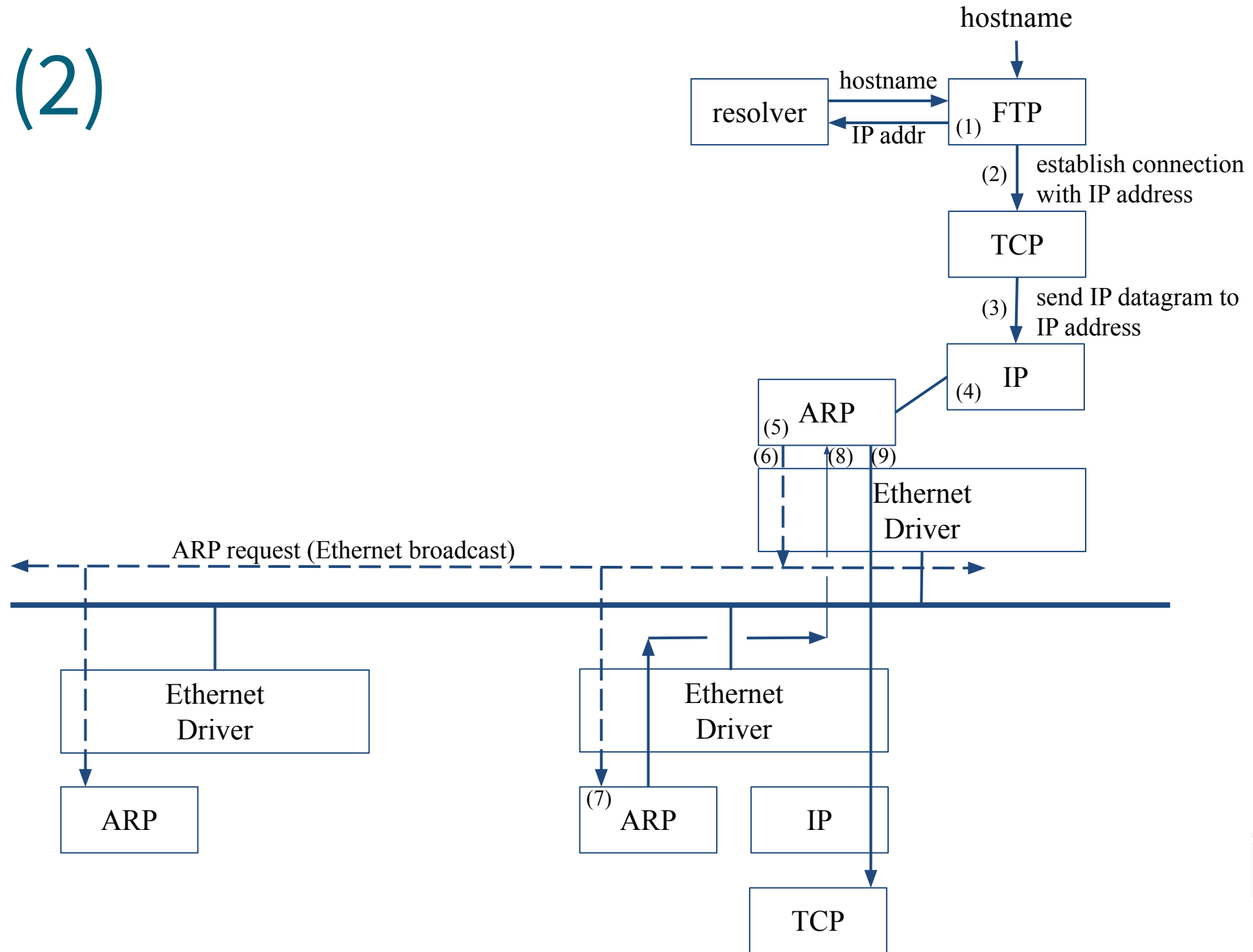


# ARP (1)

- Address Resolution Protocol
  - Ask MAC address of certain IP
  - Broadcast
  - Any one receiving ARP packet and having this IP will reply to the sender
  - When the host owing this IP is not on the same network, sender will use the MAC address of next-hop router to send the packet



# ARP (2)



# ARP (3)

- Maintain recent ARP results
  - Come from both ARP request and reply
  - Expiration time
    - Complete entry = 20 minutes
    - Incomplete entry = 3 minutes
  - Use arp command to see the cache
  - E.g.:
    - \$ arp -a
    - \$ arp -da

```
$ arp -a
crypto23.csie.nctu.edu.tw (140.113.208.143) at 00:16:e6:5b:fa:e9 on fxp1 [ethernet]
e3rtn-208.csie.nctu.edu.tw (140.113.208.254) at 00:0e:38:a4:c2:00 on fxp1 [ethernet]
e3rtn-210.csie.nctu.edu.tw (140.113.210.254) at 00:0e:38:a4:c2:00 on fxp2 [ethernet]
winpc7.csie.nctu.edu.tw (140.113.215.187) at 00:17:31:84:6c:0f on fxp3 [ethernet]
e3rtn-215.csie.nctu.edu.tw (140.113.215.254) at 00:0e:38:a4:c2:00 on fxp3 [ethernet]
e3rtn-216.csie.nctu.edu.tw (140.113.216.254) at 00:0e:38:a4:c2:00 on fxp0 [ethernet]
```

# Setup network connection

- Steps
  - Assign an IP address and hostname
  - Default route
  - DNS
  - Utility to test whether you connect to the Internet

# Setup network connection

## – assign IP, hostname and default route (1)

- FreeBSD

- In /etc/rc.conf

```
defaultrouter="140.113.17.254"  
hostname="chbsd.csie.nctu.edu.tw"  
ifconfig_fxp0="inet 140.113.17.212 netmask 255.255.255.0"  
ifconfig_fxp0_alias0="inet 140.113.17.214 netmask 255.255.255.255"  
ifconfig_fxp1="inet 192.168.1.254 netmask 255.255.255.0"
```

- Linux

- /etc/sysconfig/network
- /etc/sysconfig/network-scripts/ifcfg-eth0

```
NETWORKING=yes  
HOSTNAME=linux3  
GATEWAY=140.113.209.254
```

```
DEVICE=eth0  
BOOTPROTO=static  
BROADCAST=140.113.209.255  
IPADDR=140.113.209.143  
NETMASK=255.255.255.0  
NETWORK=140.113.209.0  
ONBOOT=yes
```

# Setup network connection

## – assign IP, hostname and default route (2)

- /etc/hosts
  - Host name database
  - Each line is a host
    - Internet address
    - Official host name
    - aliases

```
$ less /etc/hosts
127.0.0.1          localhost
140.113.209.72    ccbsd12.csie.nctu.edu.tw
140.113.209.2     ccserv
140.113.209.6     cc duty
140.113.209.7     mailgate
140.113.209.32    qkmj
```

# Setup network connection

## – assign IP, hostname and default route (3)

- Solaris

- `/etc/inet/netmasks` (network and netmask)
- `/etc/inet/hosts` (hosts)
- `/etc/defaultrouter` (default router)
- `/etc/nodename` (host name)
- `/etc/resolv.conf` (domain, nameserver, search)
- `/etc/hostname.interface` (IP, either hostname in hosts or IP)

```
$ cat hostname.rtl1s0 nodename defaultrouter resolv.conf
sun1
sun1.cs.nctu.edu.tw
140.113.235.254
domain cs.nctu.edu.tw
nameserver 140.113.235.107
nameserver 140.113.6.2
```

```
$ /etc/inet/netmasks /etc/inet/hosts
140.113.235.0 255.255.255.0
127.0.0.1 localhost
140.113.235.102 csduty
140.113.235.171 sun1
140.113.235.101 cshome
```

# Setup network connection

## – assign IP, hostname and default route (4)

- Change IP manually

- E.g.,

- `$ ifconfig fxp0 inet 140.113.235.4 netmask 255.255.255.0`

- `$ ifconfig fxp0 up`

- `$ ifconfig fxp0 down`

- Specify default route manually

- E.g.,

- `$ route add default 140.113.235.254`



# Setup network connection

## – configuring DNS

- FreeBSD, Linux

/etc/resolv.conf

```
search cc.cs.nctu.edu.tw cs.nctu.edu.tw
nameserver 10.1.1.1
nameserver 140.113.235.1
```

- Host lookup order

/etc/nsswitch.conf

```
group: compat
group_compat: ldap nis
hosts: files dns
netgroup: ldap
networks: files
passwd: compat
passwd_compat: ldap nis
shells: files
services: compat
services_compat: nis
protocols: files
rpc: files
automount: ldap files
```

# Utilities for network connection

- ping

- Send ICMP ECHO\_REQUEST to a host

```
$ ping -c 1 www.nctu.edu.tw
PING www.nctu.edu.tw (140.113.250.5): 56 data bytes
64 bytes from 140.113.250.5: icmp_seq=0 ttl=60 time=3.022 ms

--- www.nctu.edu.tw ping statistics ---
1 packets transmitted, 1 packets received, 0% packet loss
round-trip min/avg/max/stddev = 3.022/3.022/3.022/0.000 ms
```

- traceroute

- Print the route packets take to network host

```
$ traceroute www.nctu.edu.tw
traceroute to www.nctu.edu.tw (140.113.250.5), 64 hops max, 40 byte packets
 1  e3rtn-235 (140.113.235.254)  0.640 ms  0.449 ms  0.474 ms
 2  140.113.0.210 (140.113.0.210)  0.465 ms  0.310 ms  0.361 ms
 3  140.113.0.166 (140.113.0.166)  0.415 ms  0.379 ms  0.403 ms
 4  140.113.0.149 (140.113.0.149)  0.678 ms  0.536 ms  0.574 ms
 5  www.NCTU.edu.tw (140.113.250.5)  0.533 ms  0.415 ms  0.438 ms
```

# Useful Utilities in ports

- net/mtr
  - Traceroute and ping in a single graphical network diagnostic tool
- net/nload
  - Console application which monitors network traffic in real time
- net/wireshark
- net/tshark
  - A powerful network analyzer/capture tool

# Other issues

- The following issues will be given in NA (Network Administration)
  - DHCP
  - NAT
  - DNS
  - Mail
  - ...