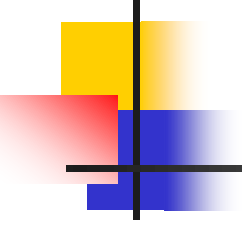


網際網路工程

Network Protocols

NICT01/Argawal ch10 (4.5 hrs)

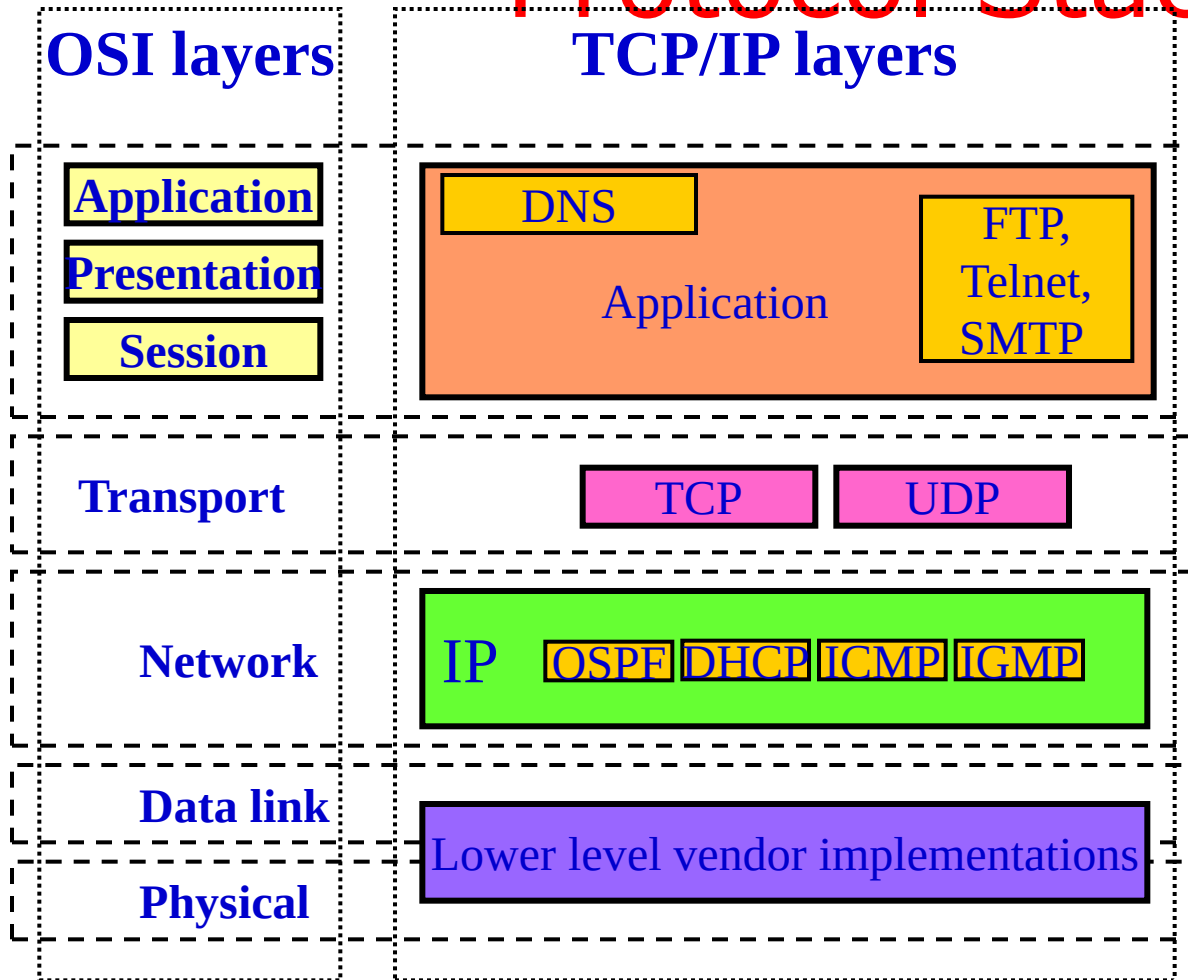
Additional reference: Jim Kurose and Keith Ross ,
Computer Networking: A Top Down Approach
6e, Addison-Wesley, Global edition



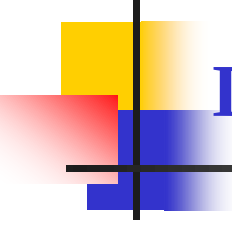
Outline

- Protocol: Set of defined rules to allow communication between entities
- Open Systems Interconnection (OSI) reference model: a need for a common reference for protocol development
- Transmission Control Protocol/Internet Protocol (TCP/IP): extremely popular Internet's backbone

(4e updated) TCP/IP Protocol Stack



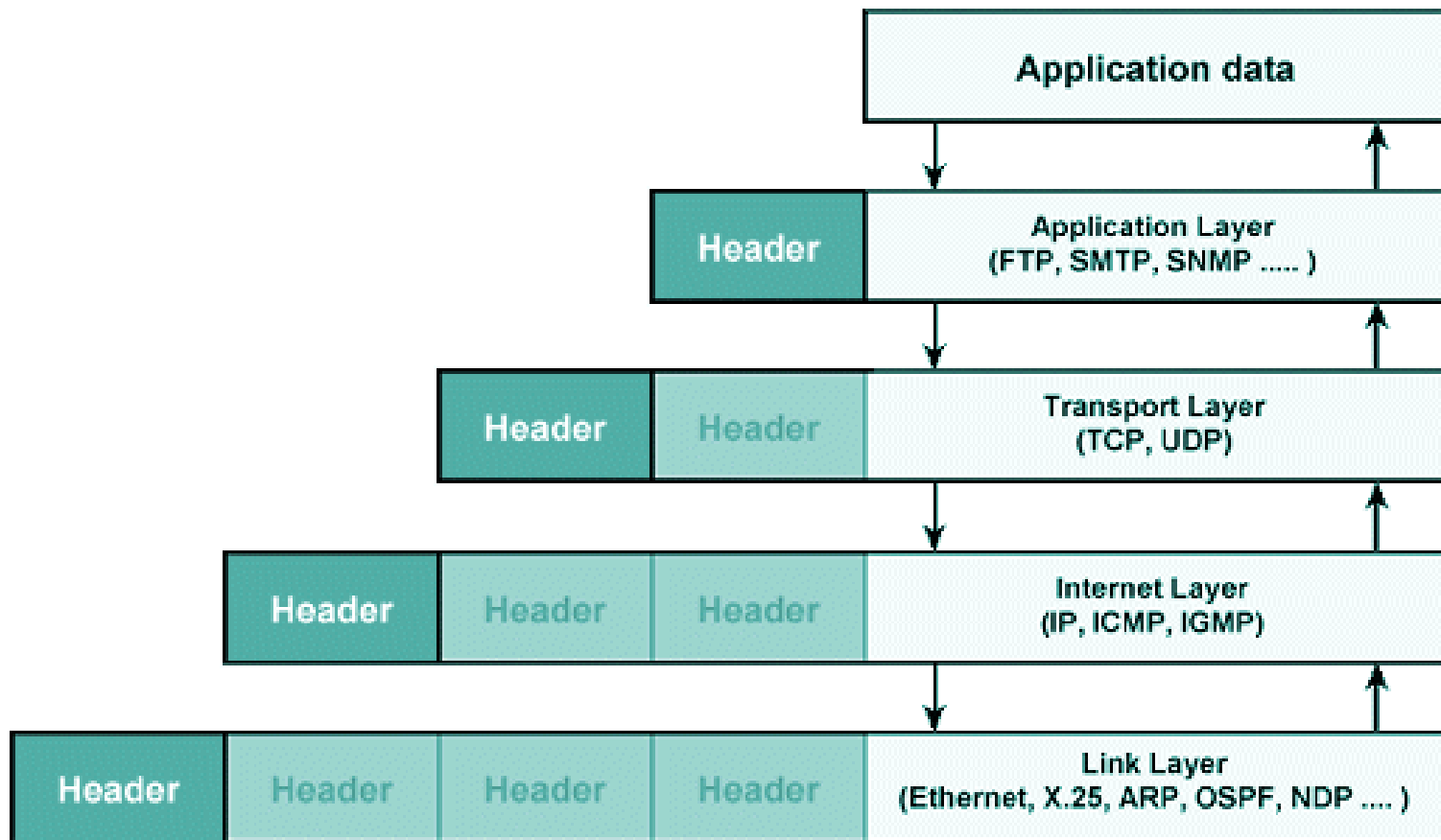
OSPF: Open Shortest Path First (routing)
DHCP: Dynamic Host Configuration Protocol (auto-configuration of IP address, gateway, DNS, mask, etc.)
ICMP: Internet Control Message Protocol
IGMP: Internet Group Management Protocol (for multicast)

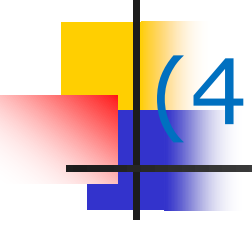


Differences between OSI and TCP/IP Models

- TCP/IP protocol consists of fewer layers.
- The lower four layers correspond to the layer of the OSI model.
- The application layer of the TCP/IP model represents the three topmost layers (session, presentation, application) of the OSI model.

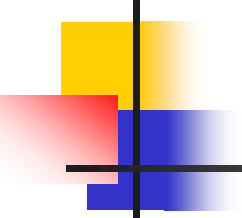
Go to lower layer, get one more envelope/pack





(4e new) Physical Layer Functions

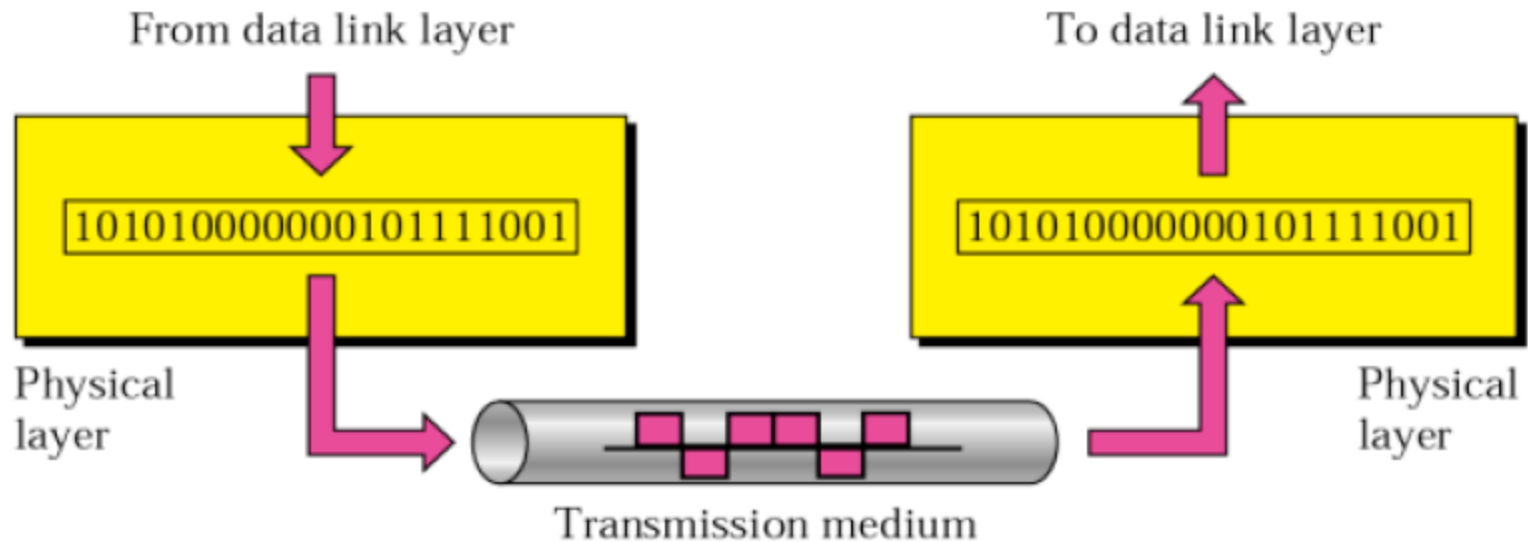
- Process for effective use of **communication resources** (**power, bandwidth etc.**)
- Physical characteristics of interfaces and media
- Representation of **bits** which could be medium access control, MAC, (layer 2 bottom half) address, packet length data etc.
- Link configuration (modulation, coding, multi-antennas, etc.)
- Physical topology (**bus, ring, star, mesh**)

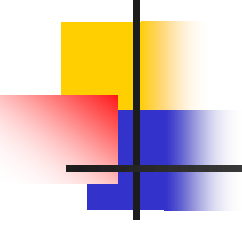


Physical Layer Functions

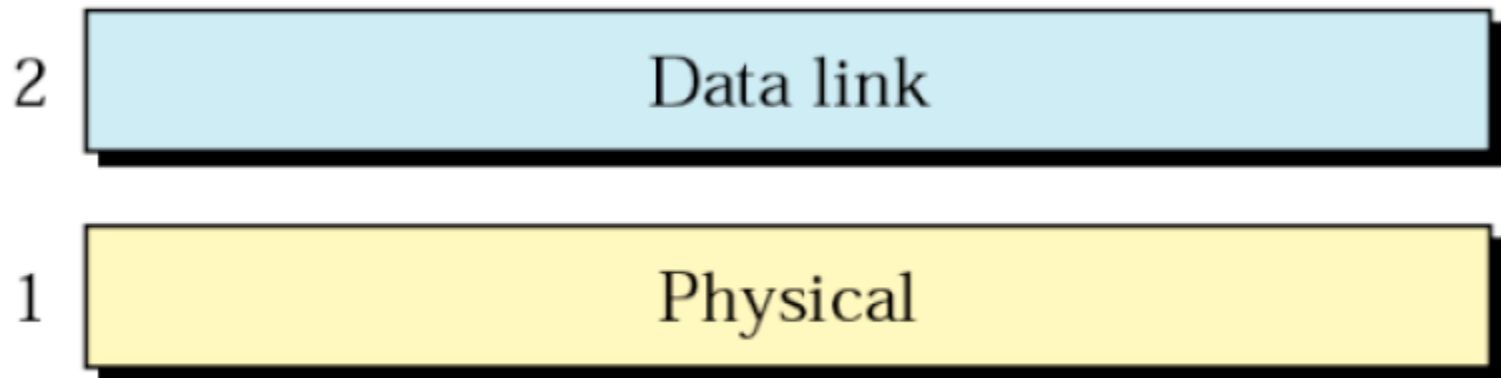
- Performs services requested by the Data Link layer. Concerned with the physical characteristics of interfaces and media.
- Modulation, source/channel coding, channel estimation, interference mitigation, spatial multiplexing/spatial diversity/beamforming, CDMA/TDMA/FDMA/OFDMA, equalization, synchronization, adaptive modulation and coding,, multi-user and cooperative diversity, NOMA(non-orthogonal multiple access), CA (carrier aggregation), mmWave

Physical layer





Data link layer



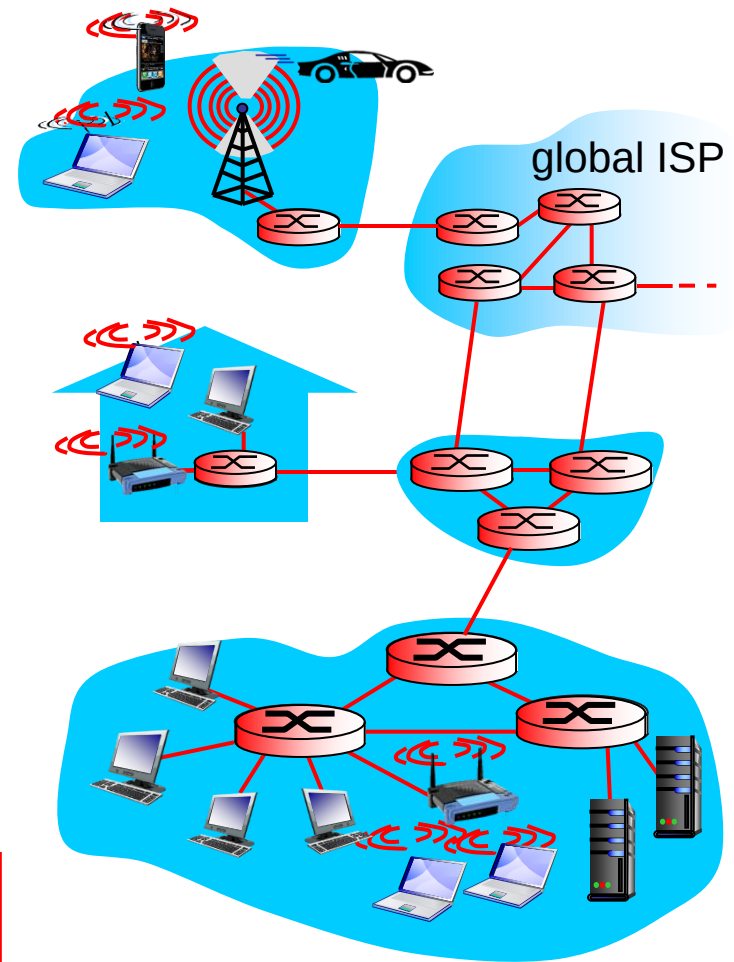
Data Link layer: introduction

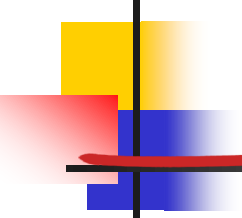
terminology:

- hosts (MS) and routers: **nodes**
- communication channels that connect adjacent nodes along communication path: **links**
 - wired links
 - wireless links
- layer-2 protocol data unit encapsulates (like p.5) datagram

data-link layer has responsibility of transferring datagram from one node to **physically adjacent** node over a link (node-to-node)

[Kur12]





Data Link layer: context

- ❖ datagram transferred by different link protocols over different links:
 - e.g., Ethernet on first link, frame relay on intermediate links, 802.11 on last link
- ❖ each link protocol provides different services
 - e.g., may or may not provide **reliable data transfer like ARQ** (automatic repeat request) over link

transportation analogy:

- trip from Princeton to Lausanne
 - limo: Princeton to JFK
 - plane: JFK to Geneva
 - train: Geneva to Lausanne
- tourist = **datagram**
- transport segment = **communication link**
- transportation mode = **link layer protocol**
- travel agent = **routing algorithm** 旅行社規劃行程



Data Link layer services

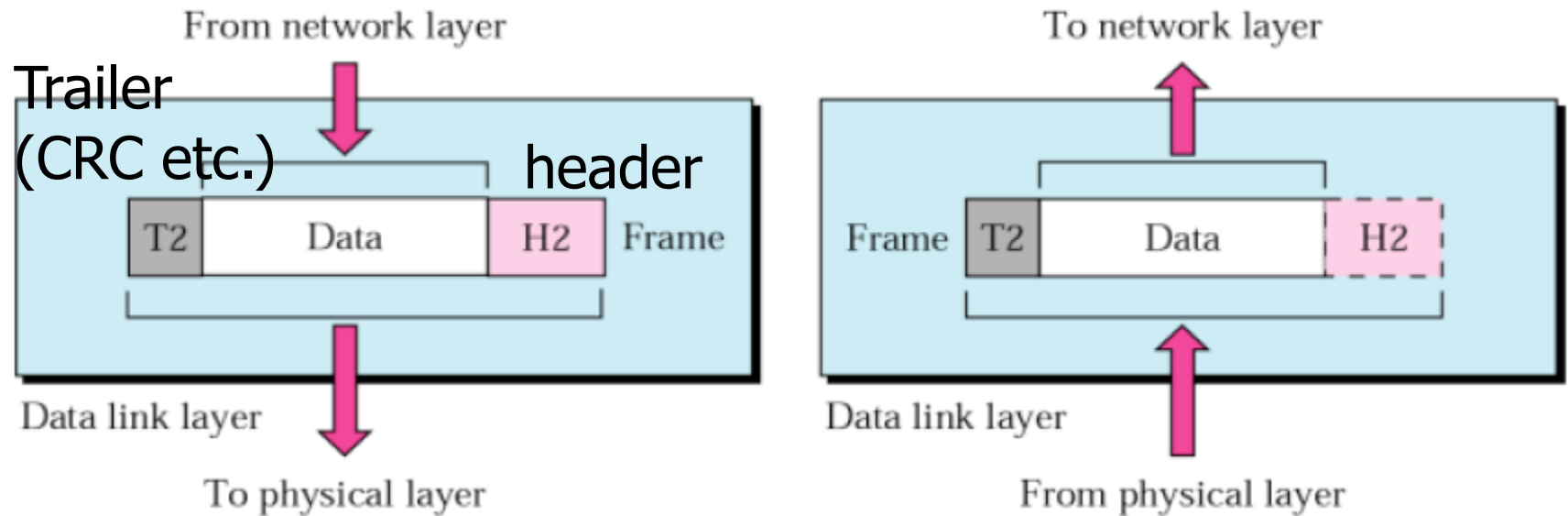
- *framing, link access:*

- encapsulate datagram into frame, adding header, trailer
- channel access if shared medium
- “MAC” addresses used in frame headers to identify source, destination (e.g. 社區內編號 D3-05)
 - different from IP address!

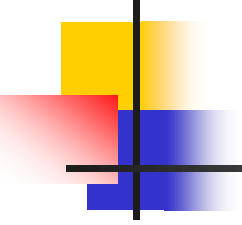
- *reliable delivery between adjacent nodes*

- seldom used on low bit-error link (fiber, some twisted pair)
- wireless links: high error rates
 - *Q:* why both link-level and end-end reliability?
 - *A:* Link ARQ for wireless link saves end-to-end delay

Data link layer



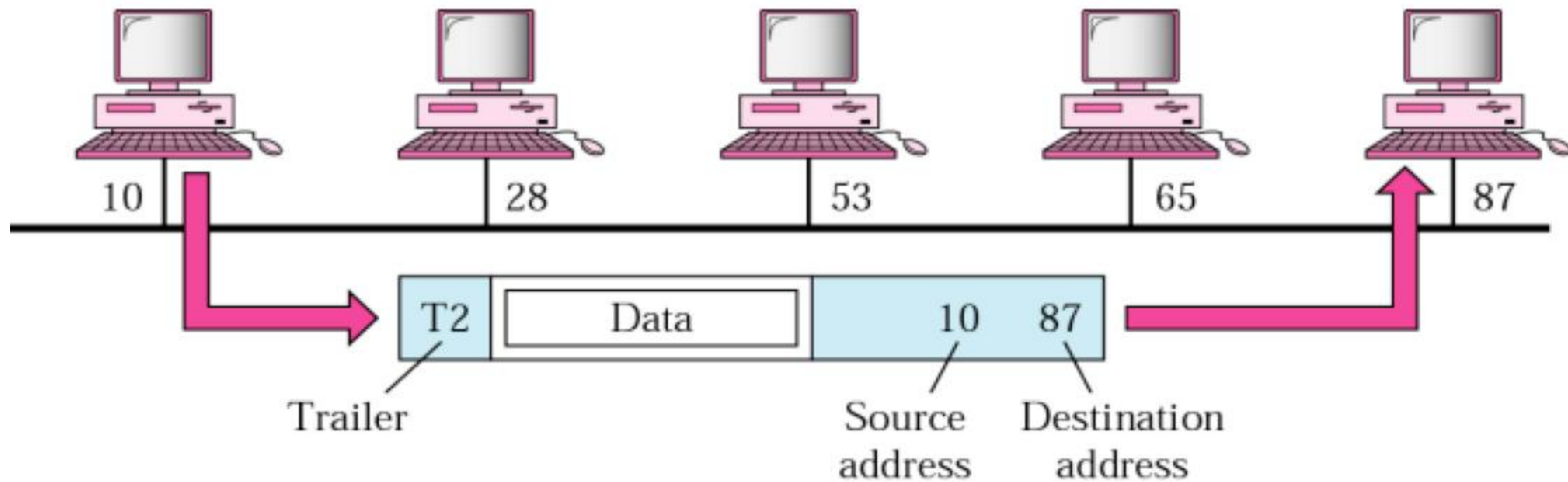
Frame=> layer 2 data unit(較明確)

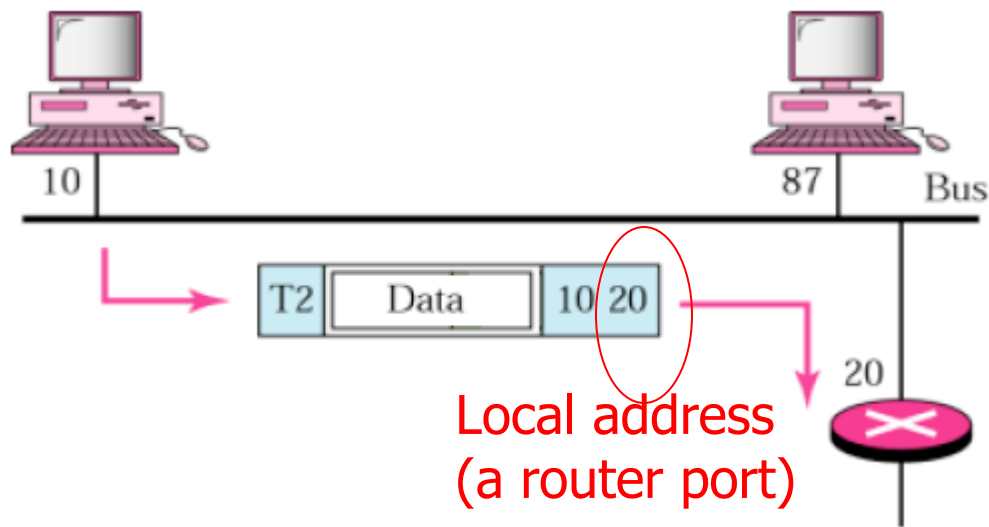


Data Link Layer Functions

- Provides functional and procedural means to transfer data between network entities (nodes).
- Responds to service requests from the network layer and issues requests to the physical layer.
- Concerned with:
 - Framing. (breaking into segments, reassembling the segments)
 - **local** addressing. (e.g. Ethernet address 0010.dc65.f037 is unique)(node-to-node: inside LAN)
 - Flow control (deal with congestion) (ex. Ramp metering 高速公路儀控)
 - Error Control (node-to-node ARQ)
 - Access Control. (e.g. WiFi's CSMA/CA, Ethernet's CSMA/CD in ch6)

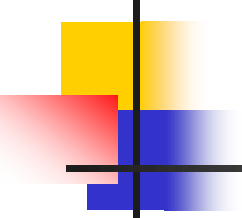
Local addressing



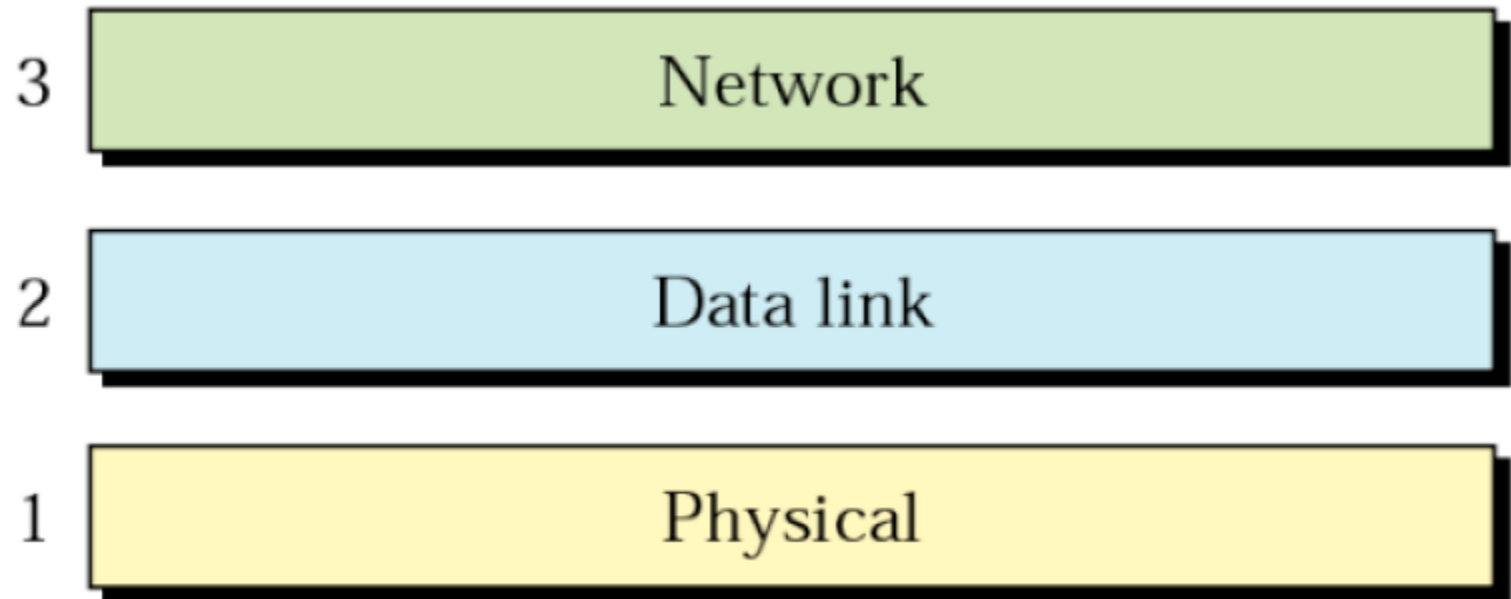


Local address
(a router port)

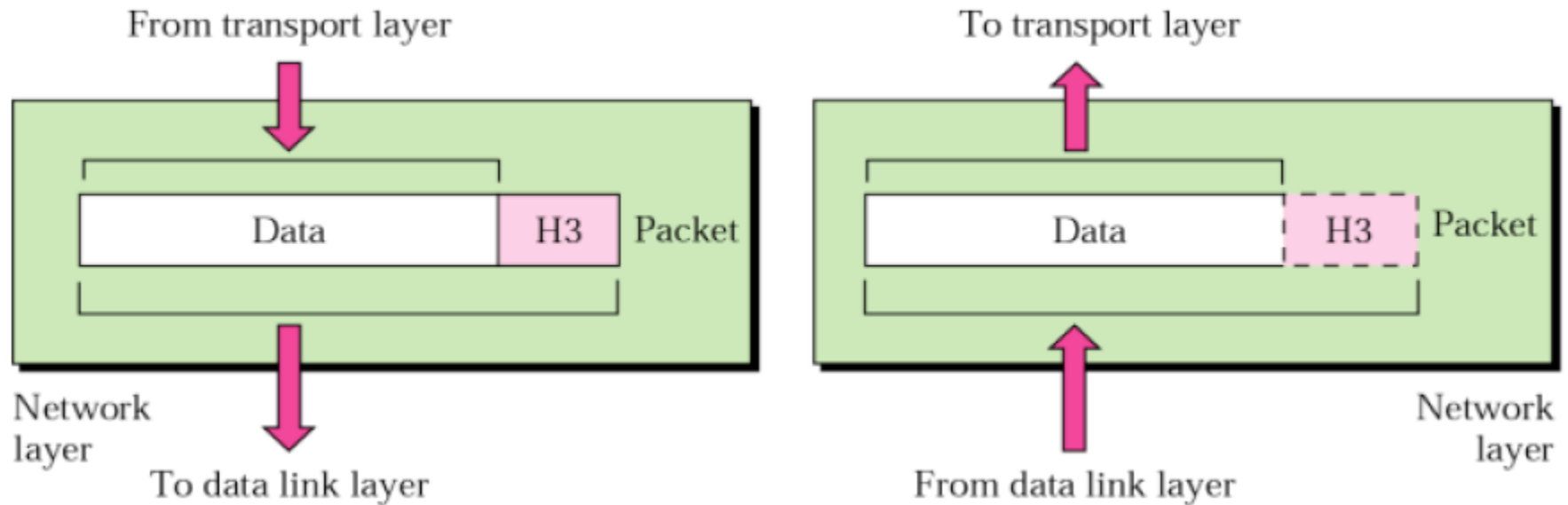
Router (between
LAN and MAN)



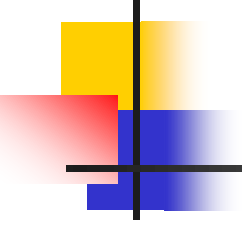
Network layer



Network layer

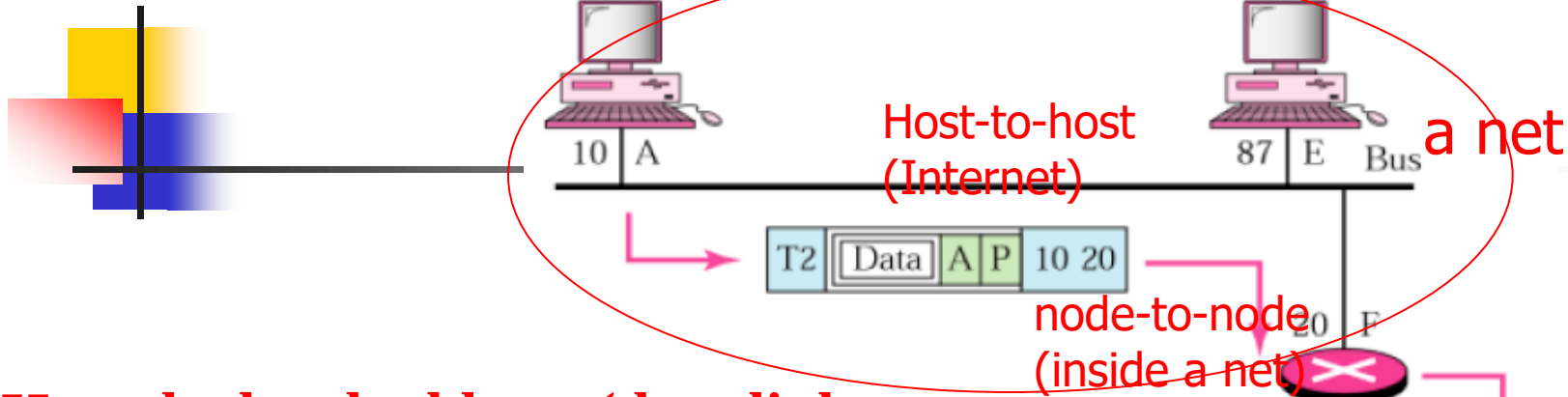


packet=> layer 3 data unit(較明確)

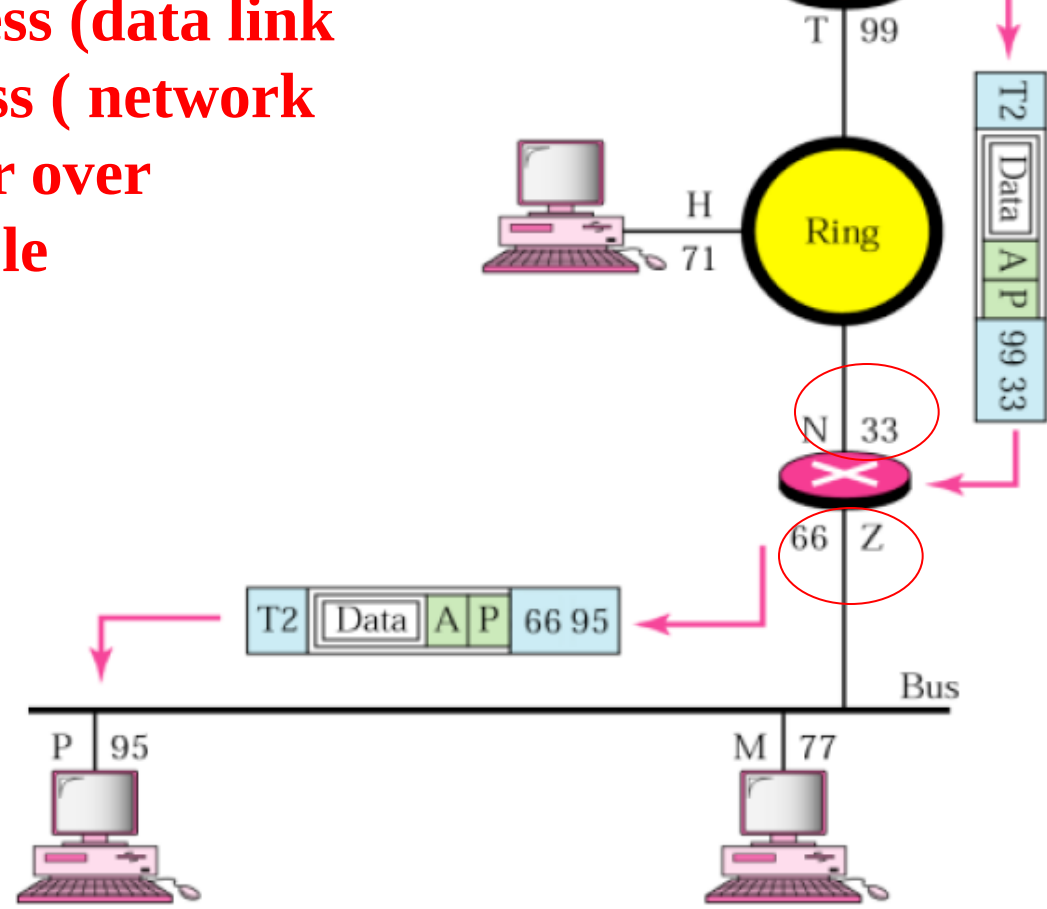


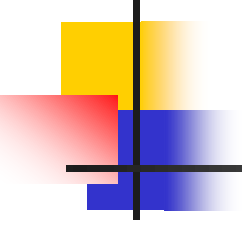
Network Layer Functions

- Provides for transfer of variable length sequences from source to destination via one or more networks (Internet).
- Responds to service requests from the transport layer and issues requests to the data link layer.
- Concerned with:
 - Logical addressing. (IP address 門牌號碼非local addressing 社區內編號: host-to-host across different nets such as Ethernet, ATM, Fiber networks)
 - routing.

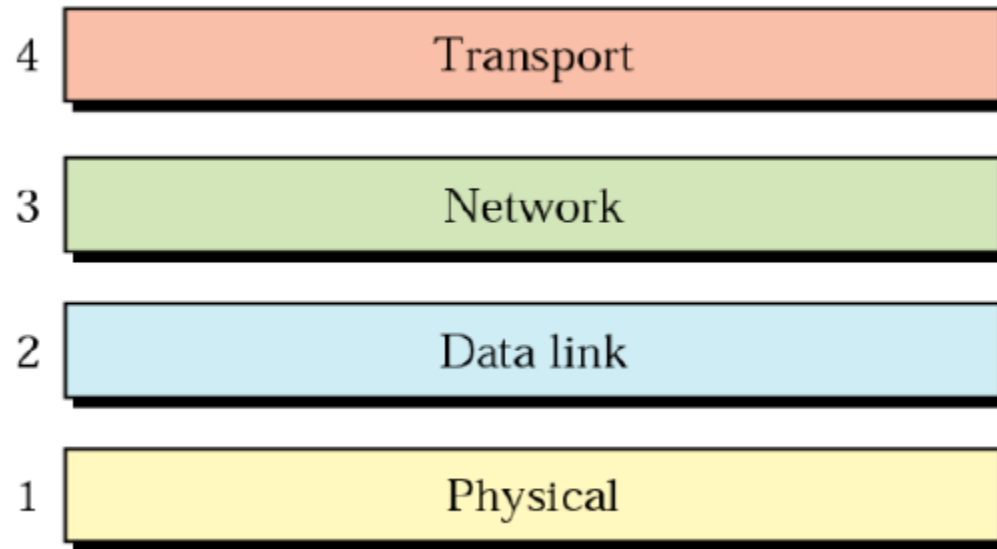


How the local address (data link layer) and IP address (network layer) work together over Internet? An example



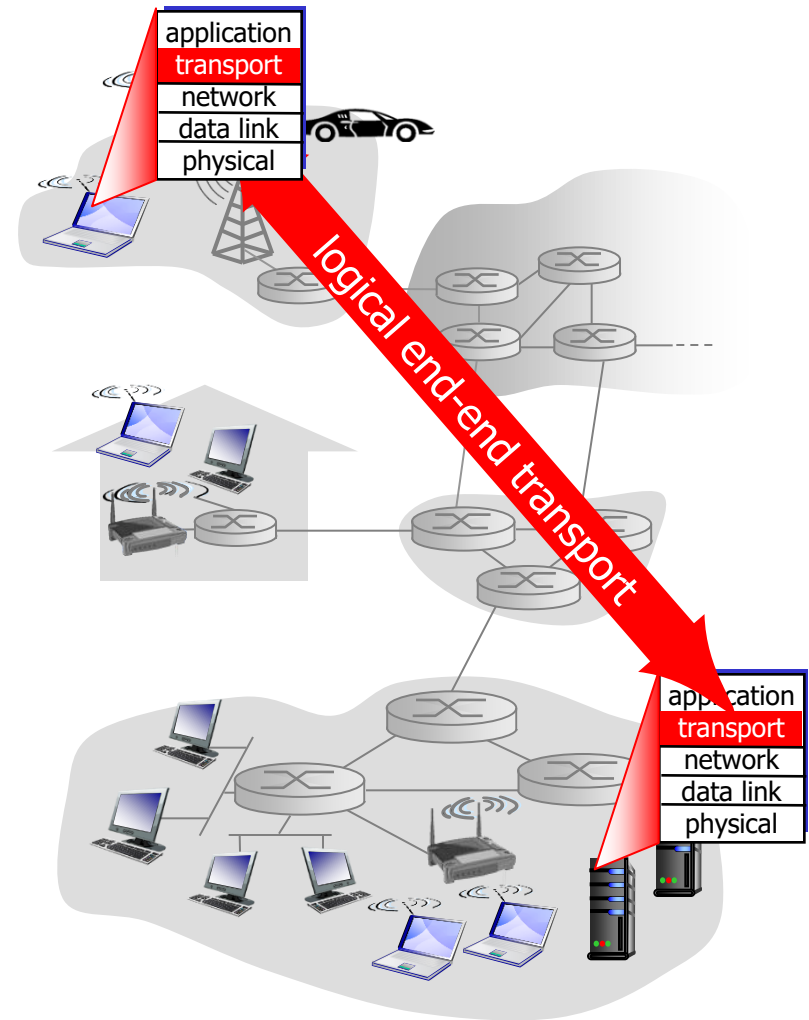


Transport layer

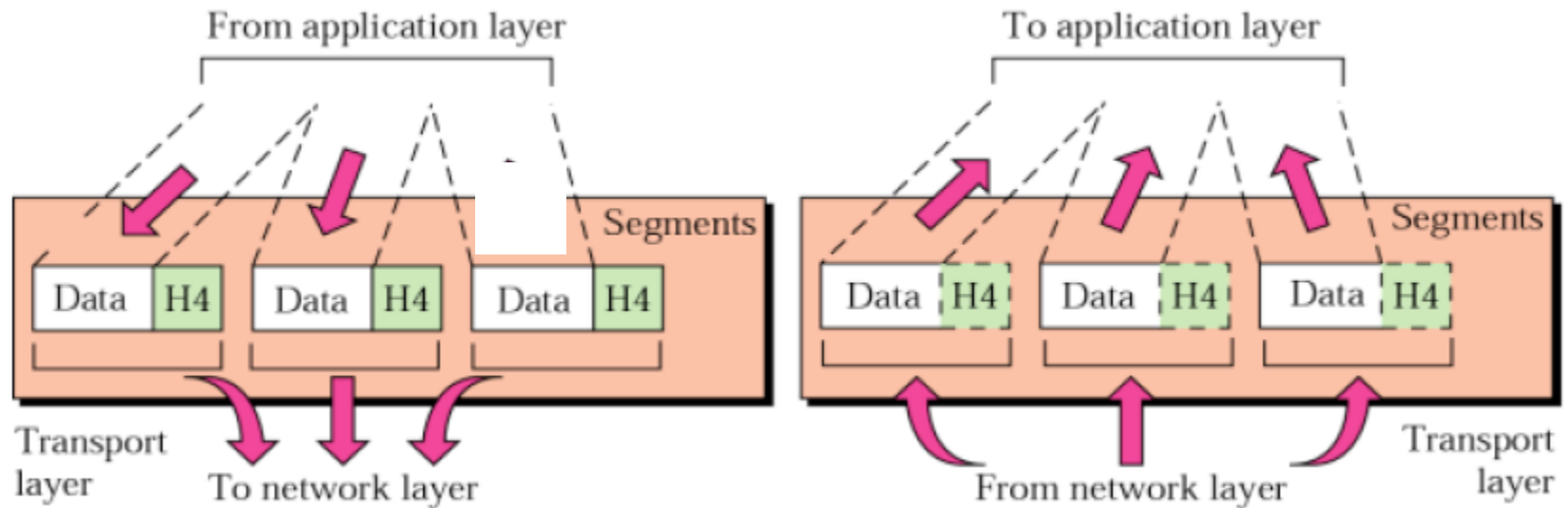


[Kur12] Transport services and protocols

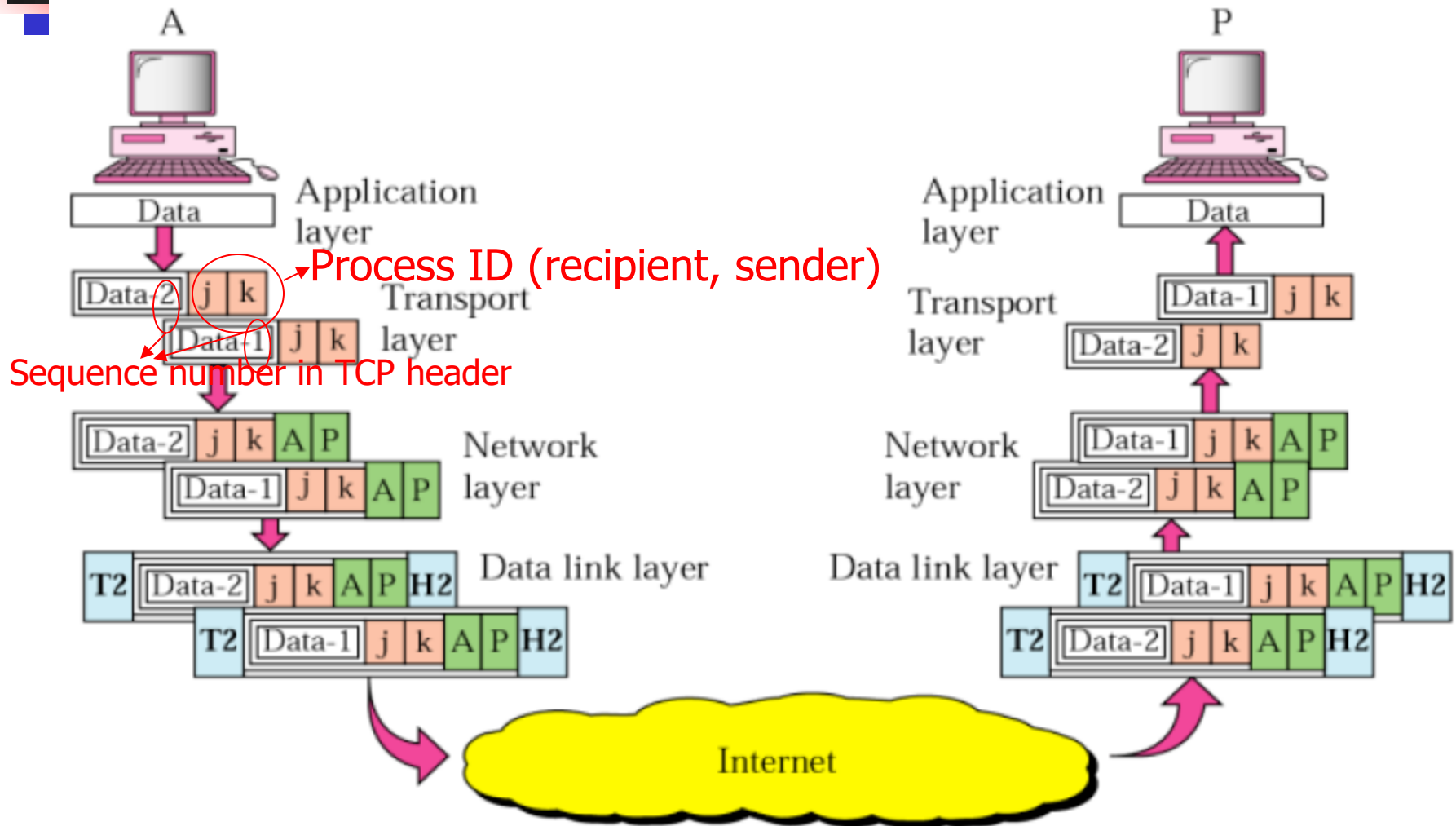
- ❖ provide *logical communication* between app processes running on different hosts
- ❖ transport protocols run in end systems
 - send side: breaks app messages into *segments*, passes to network layer
 - receive side: reassembles segments into messages, passes to app layer
- ❖ more than one transport protocol available to apps
 - Internet: **TCP** and **UDP**

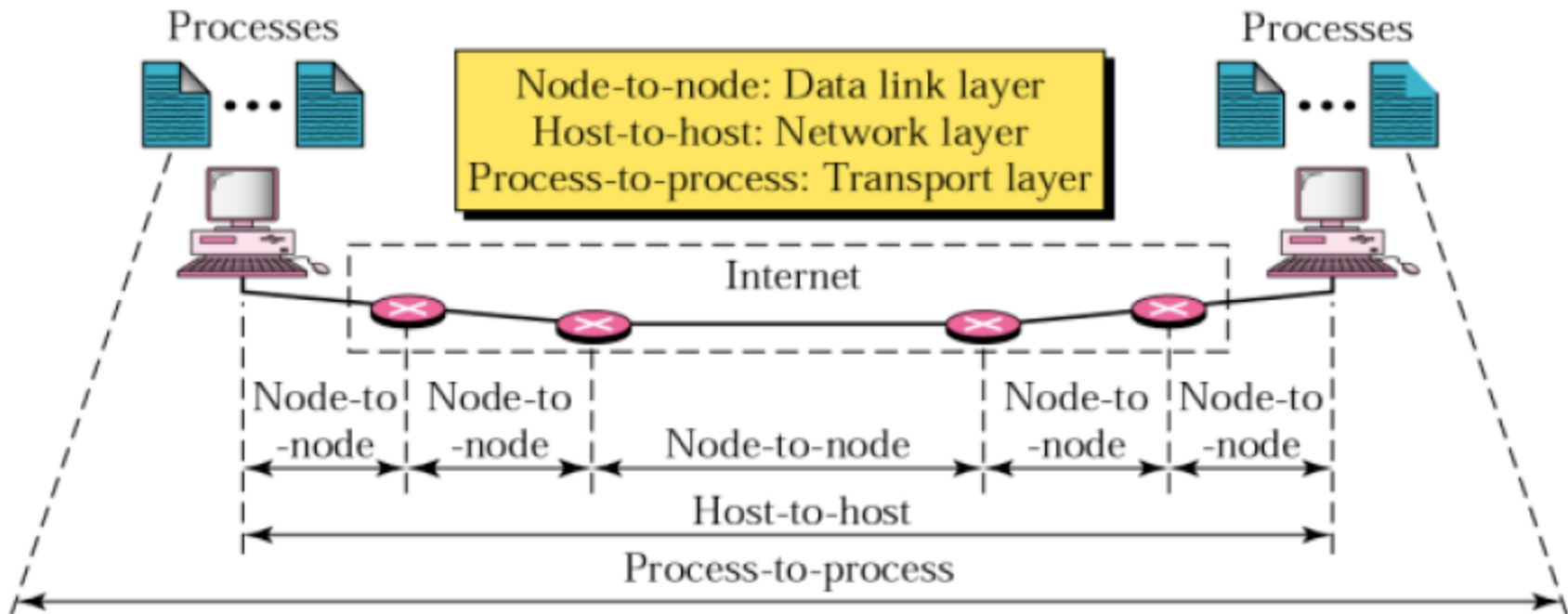
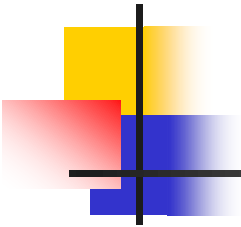


Transport layer



Transport layer example





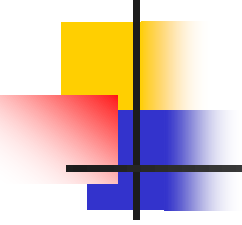
[Kur12]Transport vs. network layer

- ❖ *network layer*:
logical
communication
between **hosts**
- ❖ *transport layer*:
logical
communication
between **processes**
 - relies on, enhances,
network layer
services

household analogy:

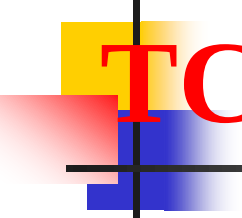
12 kids in **Ann's** house sending
letters to 12 kids in **Bill's**
house:

- hosts = houses
- processes = kids
- app messages = letters in envelopes
- **transport protocol** = Ann and Bill who **demux** to in-house siblings
- **network-layer protocol** = postal service (house-to-house)



12.3

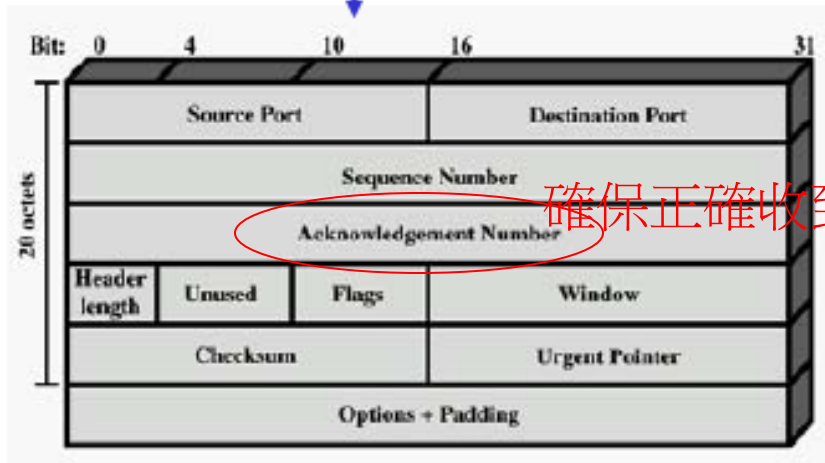
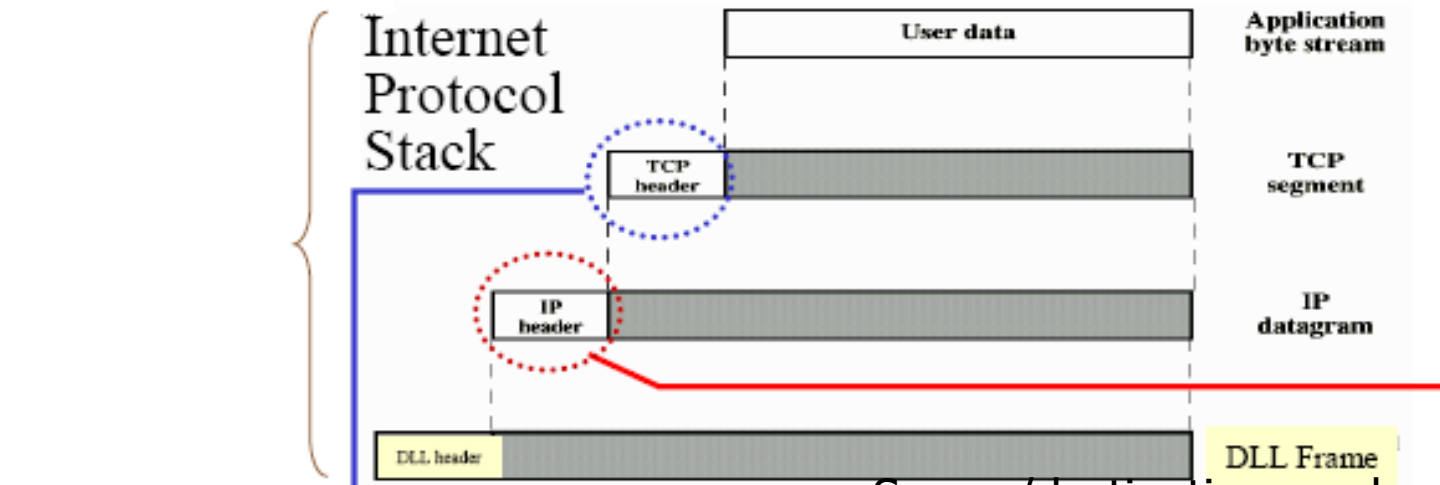
TCP/IP Protocol



TCP/IP transport layer protocols

- Transmission Control Protocol (TCP)
 - (1) error-free transmission, ex. Email, file transfer protocol (FTP), web browsing, etc.
 - (2) automatic repeat request (ARQ)- Argawal ch4 ◦
- User Datagram Protocol (UDP)
 - (1) do not guarantee correctness
 - (2) no ARQ,
 - (3) immediate response transmission, ex. connection with DNS using UDP. If fails, just try again. Real-time video streaming

Protocol Data Unit (PDU) in TCP/IP Architecture



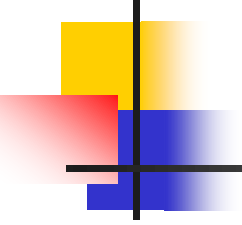
TCP header

Source/destination can have 2^{16} port (process)- too much for one PC. So we use Network Address Translation (NAT) in gateways such that many PCs Can share the same IP.

Ex. lib.ntut.edu.tw:3050 (access IEEE Xplore Database outside the campus)

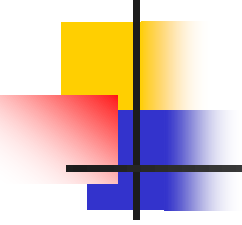
how much overhead?

- ❖ 20 bytes of TCP
- ❖ 20 bytes of IP
- ❖ = 40 bytes + app layer overhead



Internet Protocol (IP)

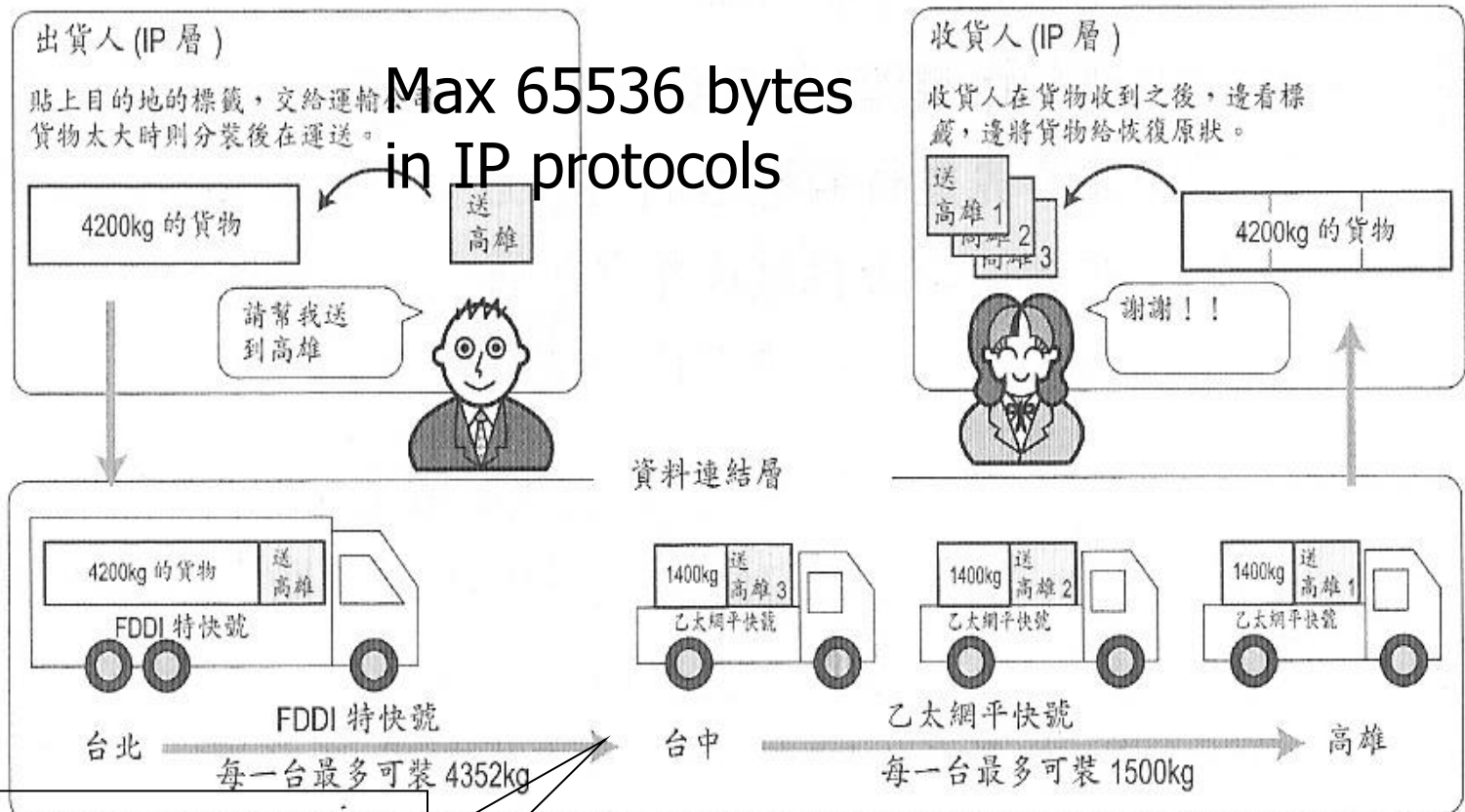
- connectionless, best-effort service for delivery of packets through the inter-network.
- cf. ATM, frame-relay, X.25 provides connection-oriented (**handshaking**) service (virtual circuit, evolved from telephony, dumb end system and complicated network), but not used in today's Internet (smart end system and simple network)[Kur12]
- Best-effort: No error checking or tracking done for the sequence of packets (datagrams) being transmitted.
- Upper layer should take care of sequencing.



Internet Protocol (IP)

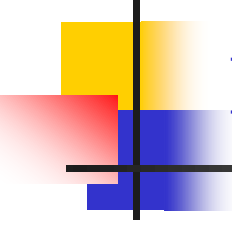
- Datagrams transmitted independently and may take different routes (e.g., a router breakdown) to reach same destination.
- Fragmentation and reassembly supported to handle data links with different maximum – transmission unit (MTU) sizes

Maximum transfer unit (MTU)



不同的資料連結層有不同的MTU
(framing) ex. 1000-1500 bytes in Ethernet

在台中站將貨物下載，分裝成乙太網平快號可以運送的尺寸。接著，在每一個貨物上，將標籤複製之後貼在上面然後進行運送。每一個標籤都被加上編號，以便將來能夠將貨物恢復成原狀。



Internet Control Message Protocol (ICMP)

- Companion protocol to IP.
- Provides mechanisms for error reporting (NOT fixing error) and query to a host or a router.
- Query message used to probe the status of a host or a router.
- Error reporting messages used by the host and the routers to report errors.

When network connection fails, you can (計網中心網路狀態)

- Ping gateway
- Ping DNS

```
C:\Documents and Settings\Shu-Ming Tseng>ping -a  
168.95.192.1
```

```
Pinging http1.hinet.net [168.95.192.1] with 32 bytes of data:
```

```
Reply from 168.95.192.1: bytes=32 time=2ms TTL=246
```

```
Reply from 168.95.192.1: bytes=32 time=2ms TTL=246
```

```
Reply from 168.95.192.1: bytes=32 time=2ms TTL=246
```

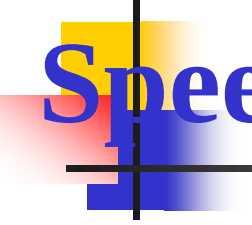
```
Reply from 168.95.192.1: bytes=32 time=2ms TTL=246
```

```
Ping statistics for 168.95.192.1:
```

```
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
```

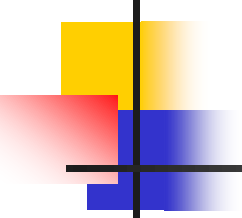
```
Approximate round trip times in milli-seconds:
```

```
    Minimum = 2ms, Maximum = 2ms, Average = 2ms
```



Speedtest ping (ms), DL/UL (Mbps)

- Ping measure round trip delay to/from a server: vital for online video game (response time)
- round trip delay is not important for video streaming (DL speed matters)



tracert www.yahoo.com

tracert: Warning: www.yahoo.com has multiple addresses; using

216.115.102.78

tracert to www.yahoo.akadns.net (216.115.102.78), 30 hops max, 38 byte packets

1 swtn184-1.adsl.seed.net.tw (211.74.184.1) 61.209 ms 63.117 ms 59.422 ms

2 139.175.169.1 (139.175.169.1) 60.172 ms 64.253 ms 60.136 ms

3 R58-38.seed.net.tw (139.175.58.38) 61.669 ms 59.185 ms 60.218 ms

4 R58-201.seed.net.tw (139.175.58.201) 68.353 ms 70.353 ms 71.605 ms

5 R57-99.seed.net.tw (139.175.57.99) 71.593 ms 70.570 ms 71.657 ms

6 R58-166.seed.net.tw (139.175.58.166) 367.829 ms 350.511 ms 355.516 ms

7 64.124.33.168.available.above.net (64.124.33.168) 218.493 ms 218.259 ms

219.996 ms

8 * * *

9 ge-2-3-0.msr1.pao.yahoo.com (216.115.101.34) 370.618 ms ge-3-3-

0.msr2.pao.yahoo.com (216.115.101.38) 337.389 ms 349.591 ms

10 vl21.bas2.snv.yahoo.com (216.115.100.229) 218.355 ms 217.388 ms

vl20.bas1.snv.yahoo.com (216.115.100.225) 221.096 ms

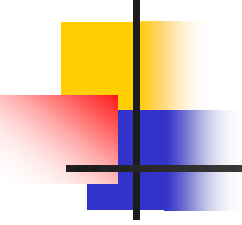
11 w6.snv.yahoo.com (216.115.102.78) 370.140 ms 339.934 ms 340.845 ms

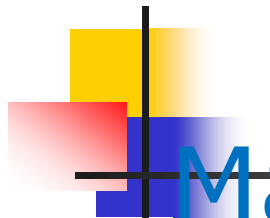
為何去日本玩「用不了AI」

核心網路，接取網路，tracert

- 資安研究機構用tracert發現eSIM 多數封包繞經中國移動香港的**核心網路，雖然接取網路用當地aoft bank etc.(手機顯示)**
- 一是資安：IMSI、地理位置、DNS 查詢等敏感元資料全經中國管轄區，台灣旅客的行蹤等於攤在中國當局眼前
- 二是使用體驗：因為 IP 被判定為香港或中國，使用者出國時 ChatGPT、Claude、Gemini 等 AI 工具會直接被擋。

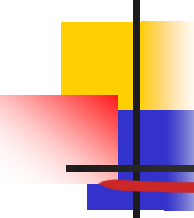
https://www.bnext.com.tw/article/90901/hidden-cost-of-cheap-travel-esim-china-mobile-routing-and-ncc-regulation?utm_source=bn_daily&utm_medium=email&utm_campaign=05-12-2026&bx_heid=5912099802

- 
- traceroute command send UDP packet with TTL=1 to the destination . First routing node finds TTL= 0 and return a ICMP packet to the source.
 - When the source receive a ICMP packet, it checks if the returning node of the ICMP packet is the destination. If not, the source send a UDP packet again with TTL+1.
 - No all routing node will return ICMP packets. For example * * * means the routing node 8 does not send ICMP packet due to firewall (Agrawal ch16).



(4e new) Internet Group Management Protocol (IGMP)

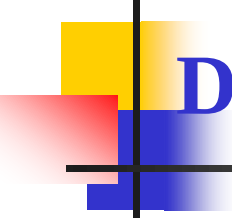
- Used to maintain multicast group membership within a domain
- Similar to ICMP, IGMP query and reply messages are used by routers to maintain multicast group membership
- Periodic IGMP query messages are used to find new multicast members within the domain
- A member sends a IGMP join message to the router, which takes care of joining the multicast tree



IP addresses: how to get one?

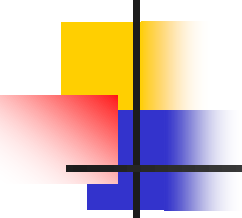
Q: How does a *host* get IP address?

- hard-coded by system admin in a file
 - Windows: control-panel->network->configuration->tcp/ip->properties
 - UNIX: /etc/rc.config
- **DHCP: D**ynamic **H**ost **C**onfiguration **P**rotocol:
dynamically get address from as server
 - “plug-and-play”



Dynamic Host Configuration Protocol (DHCP)

- Used to assign IP addresses dynamically in a domain.
- Node Requests an IP address from DHCP server.
- Helps in saving IP address space by using same IP address to occasionally connecting hosts (no time overlapping)
- Easy to set up TCP/IP parameters (IP address, DNS gateway, mask)
- Manually entering IP address may cause errors (IP address conflict)

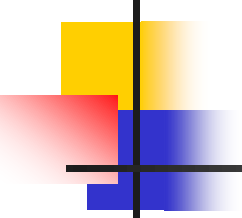


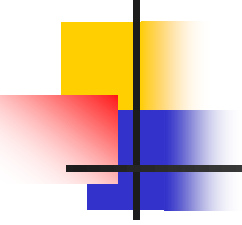
IPv4 Header Format

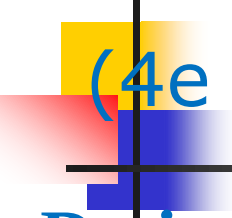
Flag: ex. last fragment in datagram or not

Version(4 bits)	Header length (4 bits)	Type of service (8 bits)	Total length (16 bits)	
Identification (16 bits)			Flags (3 bits)	Fragment offset (13 bits)
Time to live (8 bits)		Protocol (8 bits)	Header checksum (16 bits)	
Source address (32 bits)				
Destination address (32 bits)				
Options and padding (if any)				

See MTU page

- 
-
- 20bytes or more
 - ***Version(4)***
 - ***Header Length (4)***: min 5(20bytes), max 15(60bytes)
 - ***Type of Service(8)*** : reliability / speed requirement. The leftmost 3 bits represent the priority ◦
 - ***Total Length (16)***: max 65535 bytes

- 
- ***Time to Live(8)***: minus 1 if passing a routing node (one hop). Discard the packet if TTL=0 (packet expiration time is up)
 - ***Protocol(8)*** : upper layer protocol is TCP or UDP.
 - ***Header Checksum)(16)*** : check the header for every hop (CRC in ch4)
 - ***Source Address & Destination Address*** : (net-id, e.g. 140.124) (host-id, e.g., 72.71), total length 32bits.



(4e new) Internet Protocol Version 6 (IPv6)

Designed to address the unforeseen growth of the internet and the limited address space provided by IPv4
Features of IPv6:

- **Enhanced Address Space:** *128 bits long, can solve the problem created by limited IPv4 address space (32 bits)*

- **Resource Allocation:** *By using “Flow Label”, a sender can request special packet handling (Non-default QoS)*

Ex. real-time audio/video, delay-constrained, resource reservation

- **Modified Address Format:** *Options and Base Header are separated which speeds up the routing process*

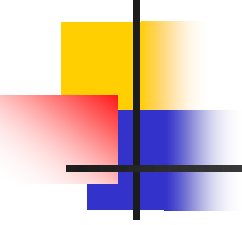
- *Routers don't check options header like IPv4 does*

- *Base header is of fixed length (faster)*

- **Support for Security:** *Encryption and Authentication options are supported in option header*

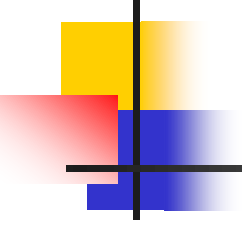
(4e new) Format of IPv6 header

Name	Bits	Function
Version	4	IPv6 version number
Traffic Class	8	Internet traffic priority delivery value
Flow Label (purpose: resource allocation)	20	Used for specifying special router handling from source to destination(s) for a sequence of packets
Payload Length (total length in IPv4)	16, unsigned	Specifies the length of the data in the packet. When set to zero, the option is a hop-by-hop Jumbo payload
Next Header (protocol in IPv4)	8	Specifies the next encapsulated protocol. The values are compatible with those specified for the IPv4 protocol field
Hop Limit (time to live TTL in IPv4)	8, unsigned	For each router that forwards the packet, the hop limit is decremented by 1. When the hop limit field reaches zero, the packet is discarded. This replaces the TTL field in the IPv4 header that was originally intended to be used as a time based hop limit
Source Address	128	The IPv6 address of the sending node
Destination Address	128	The IPv6 address of the destination node



Router/Switch

- A router select the best route for data delivery (mobile switching center in cellular networks has the same function) ,



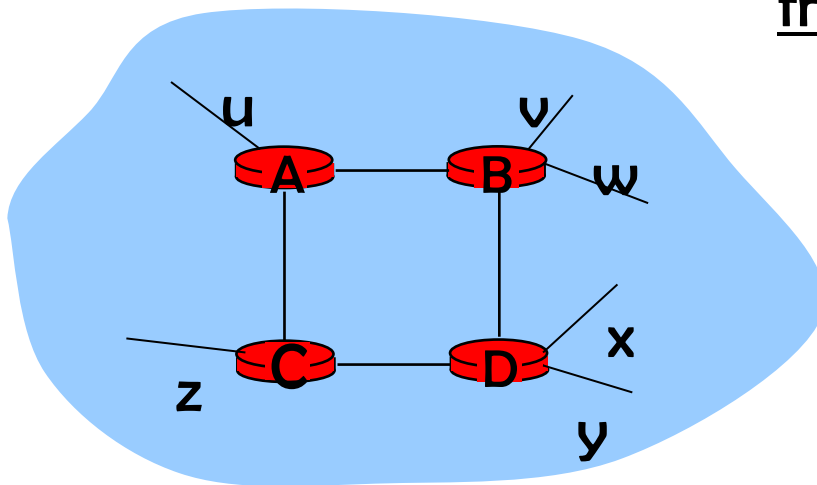
Internet Routing Protocols

most common routing protocols today:

- **Routing Information Protocol (RIP):** distance vector. – small network (DV exchange every 30sec)
- **Open Shortest Path First (OSPF on p.3):** link state- based. Neighbor information is conveyed to all the nodes in the network. – large network (info exchange if router changes or already a long time)

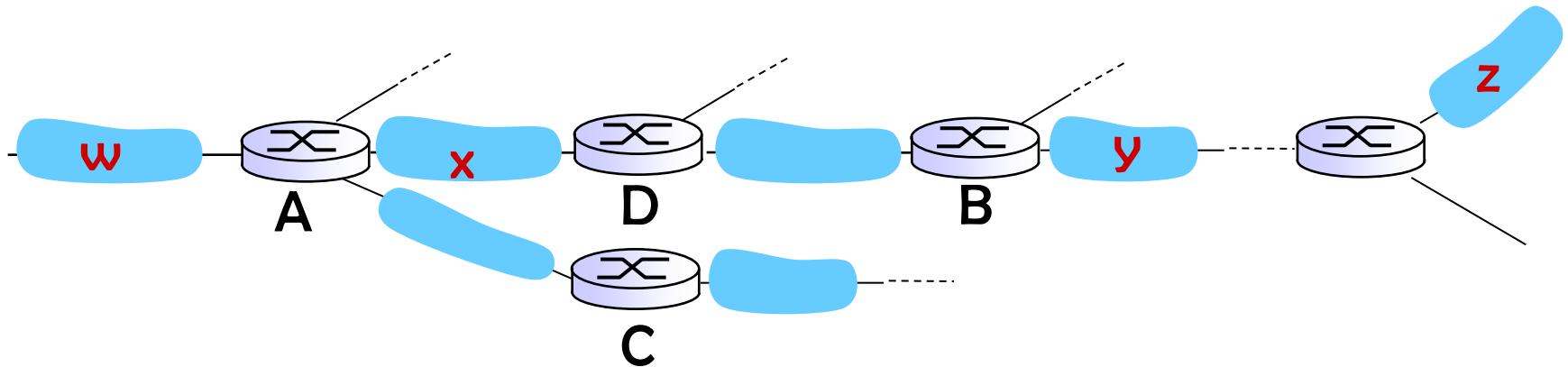
RIP (Routing Information Protocol)

- included in BSD-UNIX distribution in 1982
- distance vector algorithm
 - distance metric: # hops (max = 15 hops), **each link has cost 1**
 - DVs exchanged with neighbors every 30 sec in response message (aka **advertisement**)
 - each advertisement: list of up to 25 destination **subnets** (in IP addressing sense)



from router A to destination **subnets**:

<u>subnet</u>	<u>hops</u>
u	1
v	2
w	2
x	3
y	3
z	2



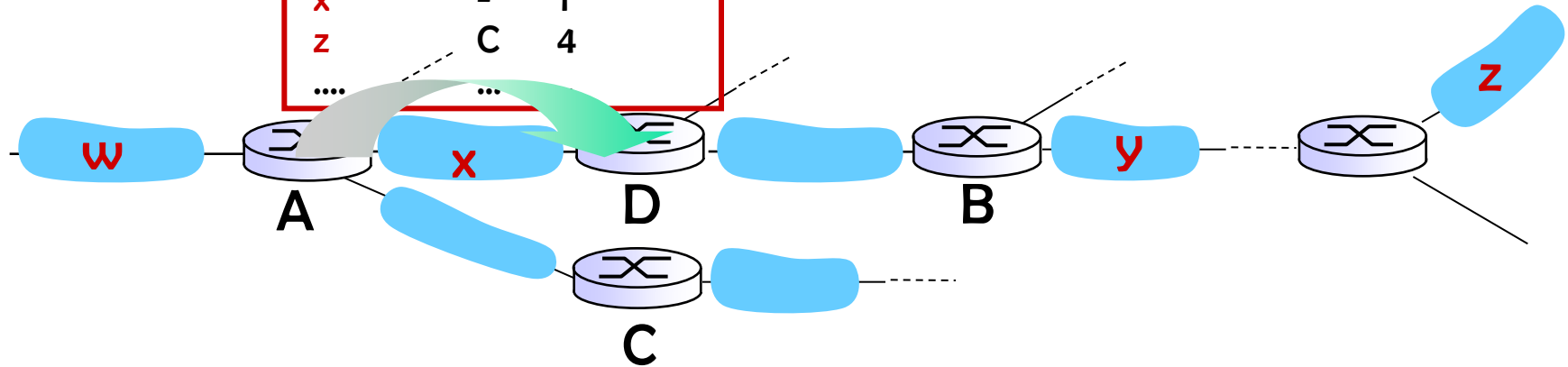
routing table in router D

destination subnet	next router	#hops to dest
w	A	2
y	B	2
z	B	7
x	--	1
....		

RIP: example

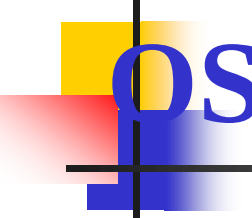
A-to-D advertisement

dest	next	hops
w	-	1
x	-	1
z	C	4
....	...	



routing table in router D

destination subnet	next router	#hops to dest
w	A	2
y	B	2
z	~~B~~ → A	~~7~~ → 5=4+1
x	--	1
....		



OSPF - Open Shortest Path First

- “Open” because it is a open standard (Internet Engineering Task Force, IETF)
- “SPF” from an alternate name for a link state routing protocol
 - View network as a collection of links between nodes
 - topology map at each node (RIP don't need global info)
 - Must know state (up/down) of all links
 - Allows for variable metrics (cost) per link
 - Computes best path for all routes using Dijkstra's shortest path algorithm

A Link-State Routing Algorithm

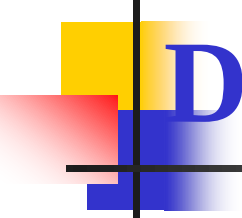
Dijkstra's algorithm (discrete math, computer introduction in freshman/sophomore)

- net topology, link costs known to all nodes
 - accomplished via “link state broadcast”
 - all nodes have same info
- computes least cost paths from one node (‘source’) to all other nodes
 - gives *forwarding table* for that node

iterative: after k iterations, know least cost path to k dest.'s

notation:

- $c(x,y)$: link cost from node x to y ; $= \infty$ if not direct neighbors
- $D(v)$: current value of cost of path from source to dest. v
- $p(v)$: predecessor node along path from source to v
- N' : set of nodes whose least cost path definitively known



Dynamic programming (演算法)

- Break into simpler subproblems in a recursive manner
- Ex. Dijkstra's algorithm for finding shortest path in graph,
- Viterbi algorithm for find minimum distance in convolutional code trellis

Dijsktra' s Algorithm pseudo code

1 *Initialization:*

2 $N' = \{u\}$

3 for all nodes v

4 if v adjacent to u

5 then $D(v) = c(u,v)$

6 else $D(v) = \infty$

7

8 *Loop*

9 find w not in N' such that $D(w)$ is a minimum

10 add w to N' (因為從其他節點來一定更遠 see ex. next page)

11 update $D(v)$ for all v adjacent to w and not in N' : (改經 w 可能會更好)

12 $D(v) = \min(D(v), D(w) + c(w,v))$

13 /* new cost to v is either old cost to v or known

14 shortest path cost to w plus cost from w to v */

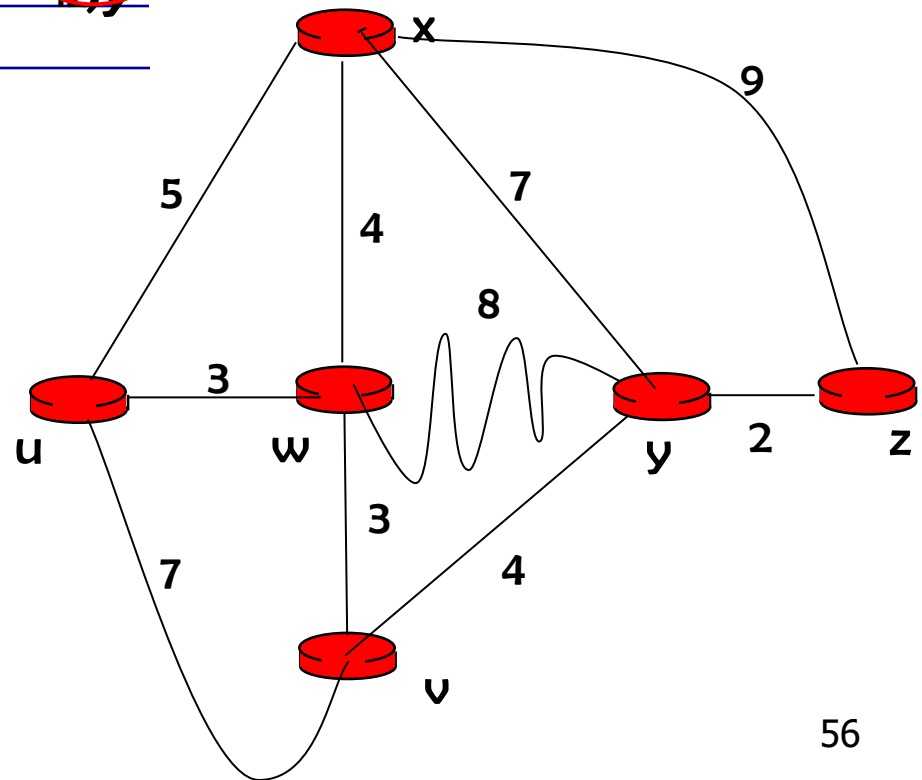
15 *until all nodes in N'*

Dijkstra's algorithm: example

Step	N'	D(v) p(v)	D(w) p(w)	D(x) p(x)	D(y) p(y)	D(z) p(z)
0	u	7,u	3,u	5,u	∞	∞
1	uw	6,w		5,u	11,w	∞
2	uwX	6,w			11,w	14,X
3	uwXv				10,v	14,X
4	uwXvy					12,y
5	uwXvyz					

notes:

- ❖ construct shortest path tree by tracing predecessor nodes
- ❖ ties can exist (can be broken arbitrarily)





OSPF “advanced” features (not in RIP)

- *security*: all OSPF messages authenticated (to prevent malicious intrusion)
- **multiple** same-cost **paths** allowed (only one path in RIP)
- for each link, multiple cost metrics for different Type of Service (**ToS**)(larger bandwidth or shorter response time)
- **hierarchical** OSPF in large domains. (to all nodes => to all domain gateways)

P10.2. What are the different between OSI and TCP/IP protocol models?
Explain clearly.

[Solution]

These are the differences between OSI and TCP/IP protocol models:

The OSI model consists of 7 layers, whereas the TCP/IP model consists only 5 layers.

In the TCP/IP protocol suite, the application layer serves the purpose of the three combined layers of application, presentation and session. The OSI model made a clear distinction between the top three layers. Each application in the TCP/IP suite has to independently implement the session and presentation layer functions.

P10.6. What is DHCP? How does DHCP support dynamic address allocation?

[Hint] by dynamically allocate IP address to host that need access Internet once in a while. propose to reduce the problem of the limited IPv4 address space

P10.23. IPv6 supports resource allocation. Explain how this is achieved.

[Solution]

IPv6 can support resource allocation by the use of the flow label field in the IP header. The flow label can be used to support transmission of real time audio and video, which require resources such as high bandwidth, large buffers, long processing time, etc.

- P10.7. With suitable examples, explain the differences between connection-oriented and connectionless protocols.

[Solution]

- A connection-oriented protocol needs the client and the server to establish a connection with the help of control packets, before the transmission of any data packets. This is termed as a handshake (2G~5G control channels to set up calls)
- A connectionless protocol does not involve handshaking mechanism. The sender can directly send a data packet to the receiver irrespective of whether the receiver is ready to accept it or not.
- The advantage of a connectionless protocol is that it does not require the overhead involved with establishing a connection, thereby improving the time required for the first packet to be received by the receiver. This is useful when the sender needs to intermittently send a few packets, and so the time spent on establishing the connection is not justified.