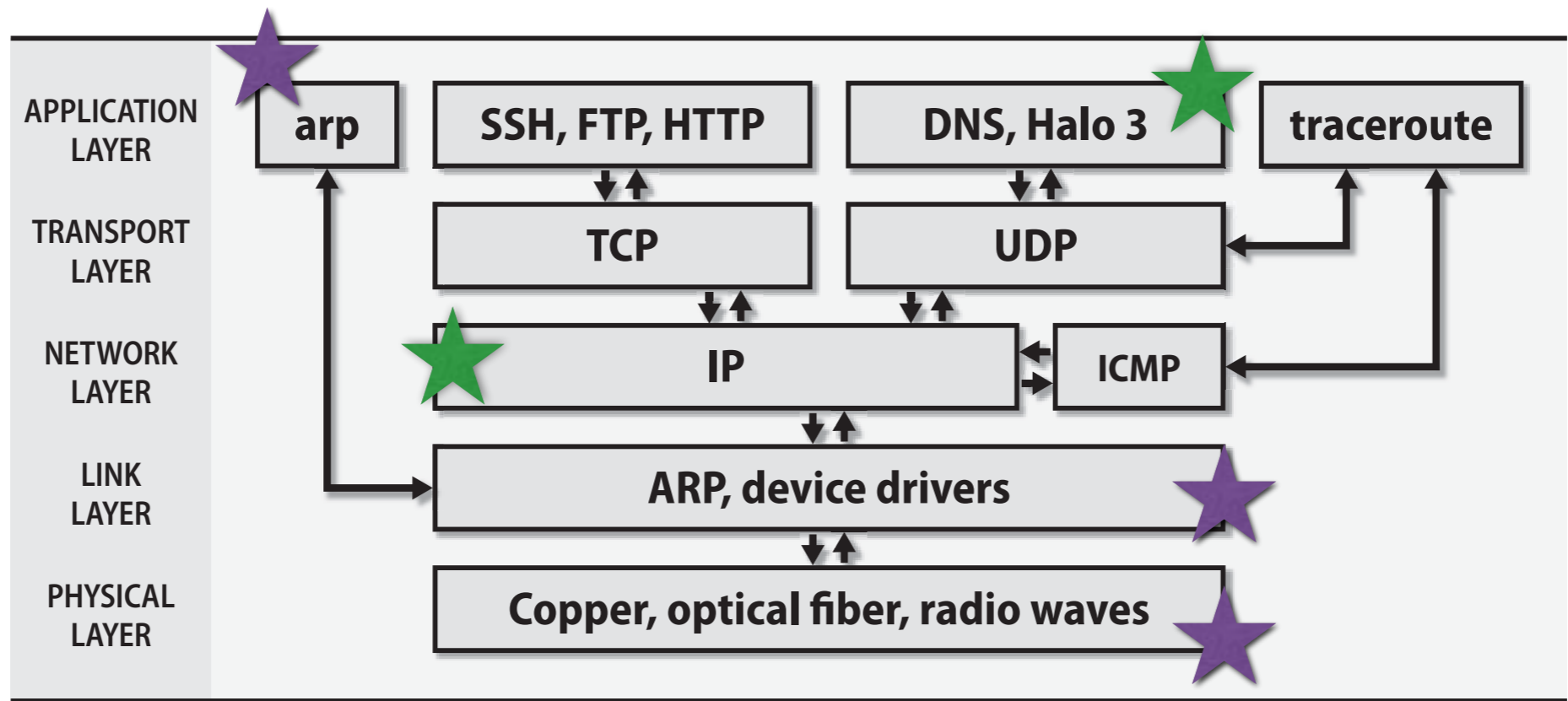


IP Layer

Michael Tsai
2017/04/10



IP (Network layer) 的主要功能

1. Forwarding: Router通常有多個interface (網卡)。把 packet從來源的interface移到目的地方向的interface並發送出去叫做forwarding。
 - ▶ 一般client並不會開啟此一功能!
2. Routing: 找出往目的地方向的一條路徑。通常由 routing algorithms/protocol決定。
 - ▶ 因為系上通常到特定的目的地都只有一條路徑，我們網管的工作通常只會接觸到第一部分。



IP封包的格式(v4)

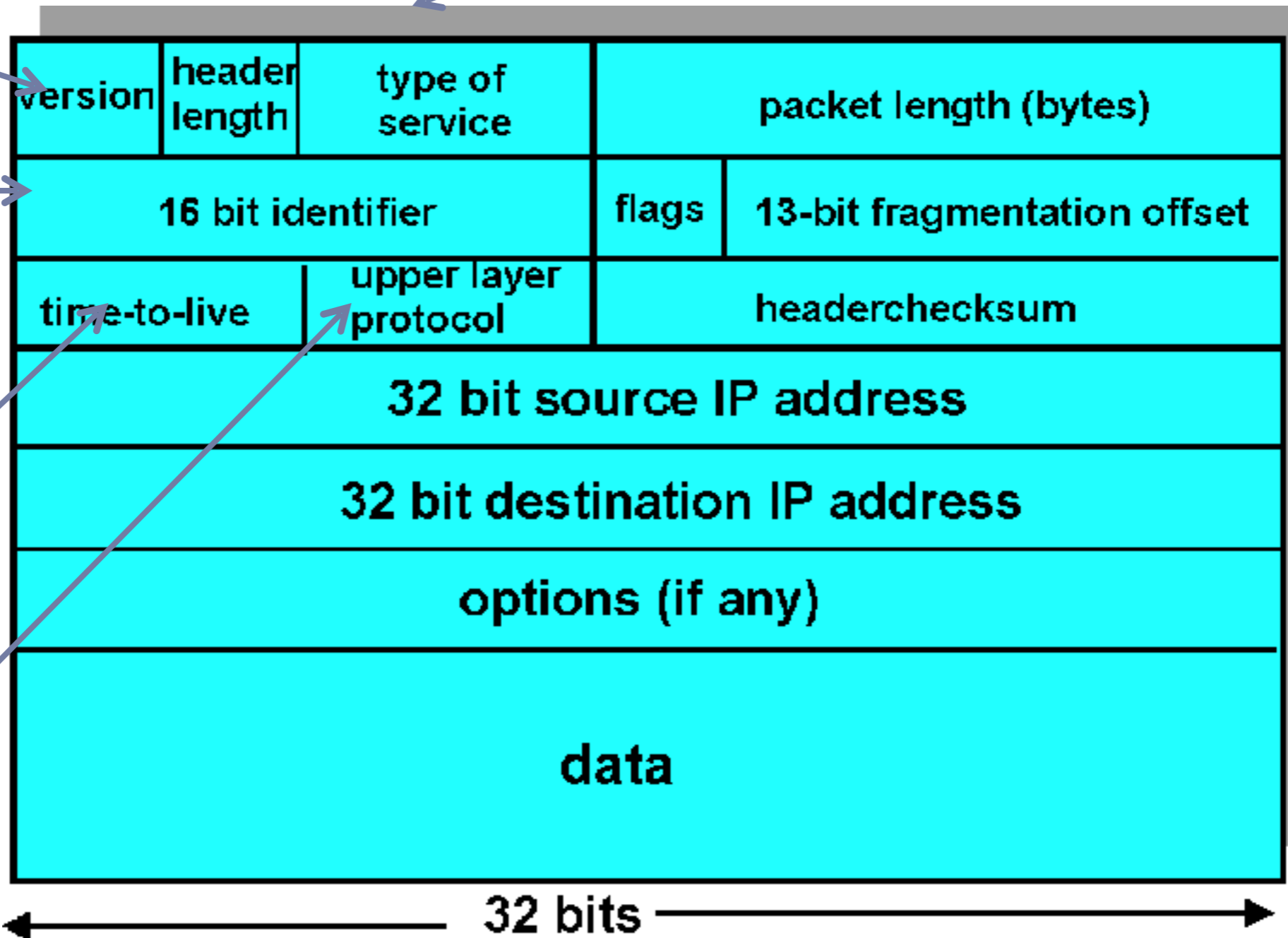
表示是否需要特殊處理(如即時的影像或聲音)

v4 or v6

用來處理
fragmentation
(想想MTU)

最多可以經過
幾台機器(router)

Transport layer使
用的協定
(通常為TCP or
UDP)

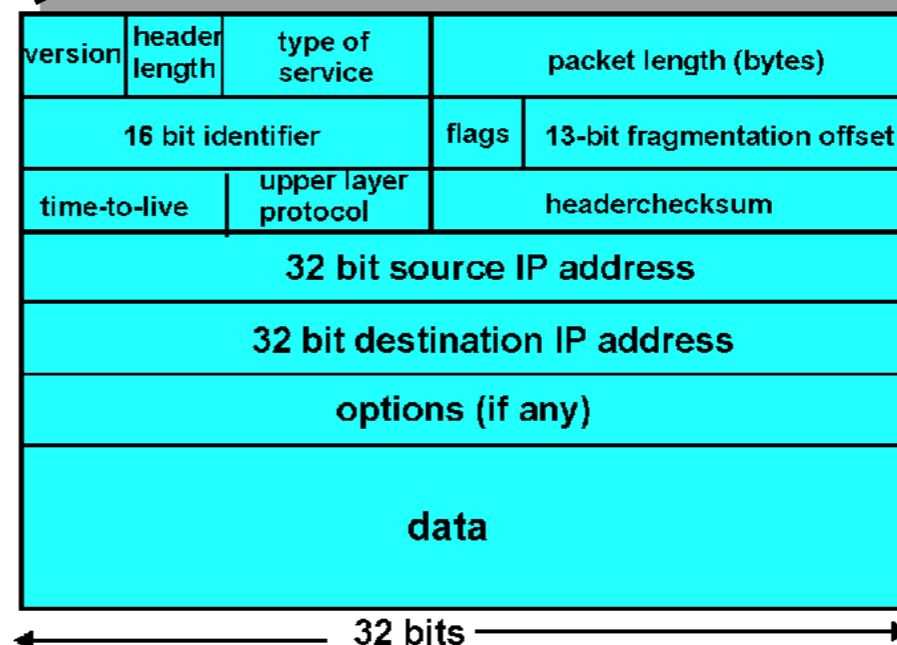


Where is IP packet?

Ethernet **Frame**

Preamble	Start of frame delimiter	MAC destination	MAC source	Length (IEEE 802.3)	802.1Q tag (optional)	Payload	Frame check sequence(3)
7 octets	1 octet	6 octets	6 octets	2 octets	(4 octets)	42–1500 octets	4 octets

IP **Packet** is in Ethernet's payload!



Typical Internet Packet

Ethernet header	IPv4 header	UDP header	Application data	Ethernet CRC
14 bytes	20 bytes	8 bytes	100 bytes	4 bytes

UDP packet (108 bytes)

IPv4 packet (128 bytes)

Ethernet frame (146 bytes)

IP Address (v4)

- AAA.BBB.CCC.DDD (4 bytes) = ? # total hosts
- Network + host address \rightarrow same network address == same network (subnet)

Historical Internet Classes (no mask)

Class	1st byte	Format	Comments
A	1-127	N.H.H.H	Very early networks
B	128-191	N.N.H.H	Large sites (hard to get)
C	192-223	N.N.N.H	Easy to get (often obtained in sets)
D	224-239	-	Multicast addresses
E	240-255	-	Experimental addresses

But this is inefficient

- Most networks only have ~100 hosts
- Class A & B addresses are wasted
- Thus we need to find a way to further split the networks! (subnetting)

Netmask

- Netmask ==
32-bit number with leading 1's + trailing 0's
- Digits mapped to 1's → network address
Digits mapped to 0's → host address
- Expressed as (a) 0xfffffc0 or (b) 255.255.255.192

IP address	128	.	138	.	243	.	0
Decimal netmask	255	.	255	.	255	.	192
Hex netmask	f	.	f	.	f	.	c 0
Binary netmask	1111	.	1111	.	1111	.	1100 0000

Two Special Addresses

- Network address
= “network address” + “host address = 0”
- Broadcast address
= “network address” + “host address = all 1’s”

Setting Interface Address

- `ifconfig -a` —> display all interfaces
- `ifconfig eth0 192.168.25.1 netmask 255.255.255.0`
—> set the IP and netmask of an interface
- `ifconfig eth0 up`
—> enable the interface
- `ifconfig eth0 media auto`
—> set the media type to auto-sense

Why do we need to know the “network address”?

- Answer: we need to know if the destination host can be reached directly (in the same network).
- How? Q: is the network address the same?
- Question: what if it is not on the same network?
- Answer: we ask a host to relay for us.
- Question: but, which host?
(it has to be on the same network)

系上防火牆的Routing table (部分)

192.168.48.0/
255.255.248.0

192.168.55.254

192.168.219.0/
255.255.255.0

192.168.219.254



140.112.30.254

140.112.28.0/
255.255.252.0

Routing Table:

```
192.168.48.0 255.255.248.0 192.168.55.254
192.168.219.0 255.255.255.0 192.168.219.254
140.112.28.0 255.255.252.0 140.112.30.254
0.0.0.0 0.0.0.0 140.112.x.x
```



How to represent a group of destination hosts?

- CIDR == Classless Inter-Domain Routing
- Borrowing the netmask idea:
IPs from 192.144.0.0 to 192.144.7.0,
we can say 192.144.0.0/21 (21 == 255.255.248.0)
- Any IP address falls in that “network”
(though might not be a real network), can be
represented by that CIDR

Private IP

- Private IP
==IPs that are not globally allocated to anyone

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

In-class exercise

- Install “ipcalc” on your VM or the platform of your choice. Learn how to use it.
- Suppose you would like to allocate some private IPs for a sub-network with around 20 hosts. Give the network address and the net mask such that it is not “oversized”.
- Paste the output of ipcalc for that network address.

NAT (Network Address Translation)

只有一塊門牌發給我們，怎麼辦呢？

對照表：

- 菜瓜布有連到8.8.8.8
- 要找助教請轉到192.168.0.4

內部用: 192.168.0.2

菜瓜布

Src: 192.168.0.2
Dest: 8.8.8.8

門牌: 140.112.91.208

馬撒起

Src: 8.8.8.8
Dest: 192.168.0.2

Src: 140.112.91.208
Dest: 8.8.8.8

內部用: 192.168.0.2

凱莉

Src: 8.8.8.8
Dest: 140.112.91.208

內部用: 192.168.0.4

小小郭

內部用門
牌: 192.168.0.254

內部用: 192.168.0.5

Routing Table

- netstat -nr (不看hostname) or netstat -r (看hostname)

```
Kernel IP routing table
Destination      Gateway         Genmask        Flags   MSS Window  irtt Iface
0.0.0.0          140.112.30.254 0.0.0.0        UG      0 0       0 eth0
140.112.30.0     0.0.0.0        255.255.255.0 U        0 0       0 eth0
```

- route add default gw 140.112.30.254
—> all traffic not to local subnets goes to the gw
- route add -net 132.236.220.64 netmask 255.255.255.192
—> all traffic that has destination address with the described network address goes to 132.236.220.64

ICMP (Internet Control Message Protocol)

- ▶ 一些管理用的訊息，用來通知client關於網路的狀況。
- ▶ 常用的用途：
 1. 通知client此路不通。(Destination network/host/protocol/port unreachable or unknown)
 2. Ping使用的echo request & reply

```
C:\Users\Administrator>ping 8.8.8.8

Ping 8.8.8.8 <使用 32 位元組的資料>:
回覆自 8.8.8.8: 位元組=32 時間=20ms TTL=128
回覆自 8.8.8.8: 位元組=32 時間=20ms TTL=128
回覆自 8.8.8.8: 位元組=32 時間=20ms TTL=128
回覆自 8.8.8.8: 位元組=32 時間=20ms TTL=128

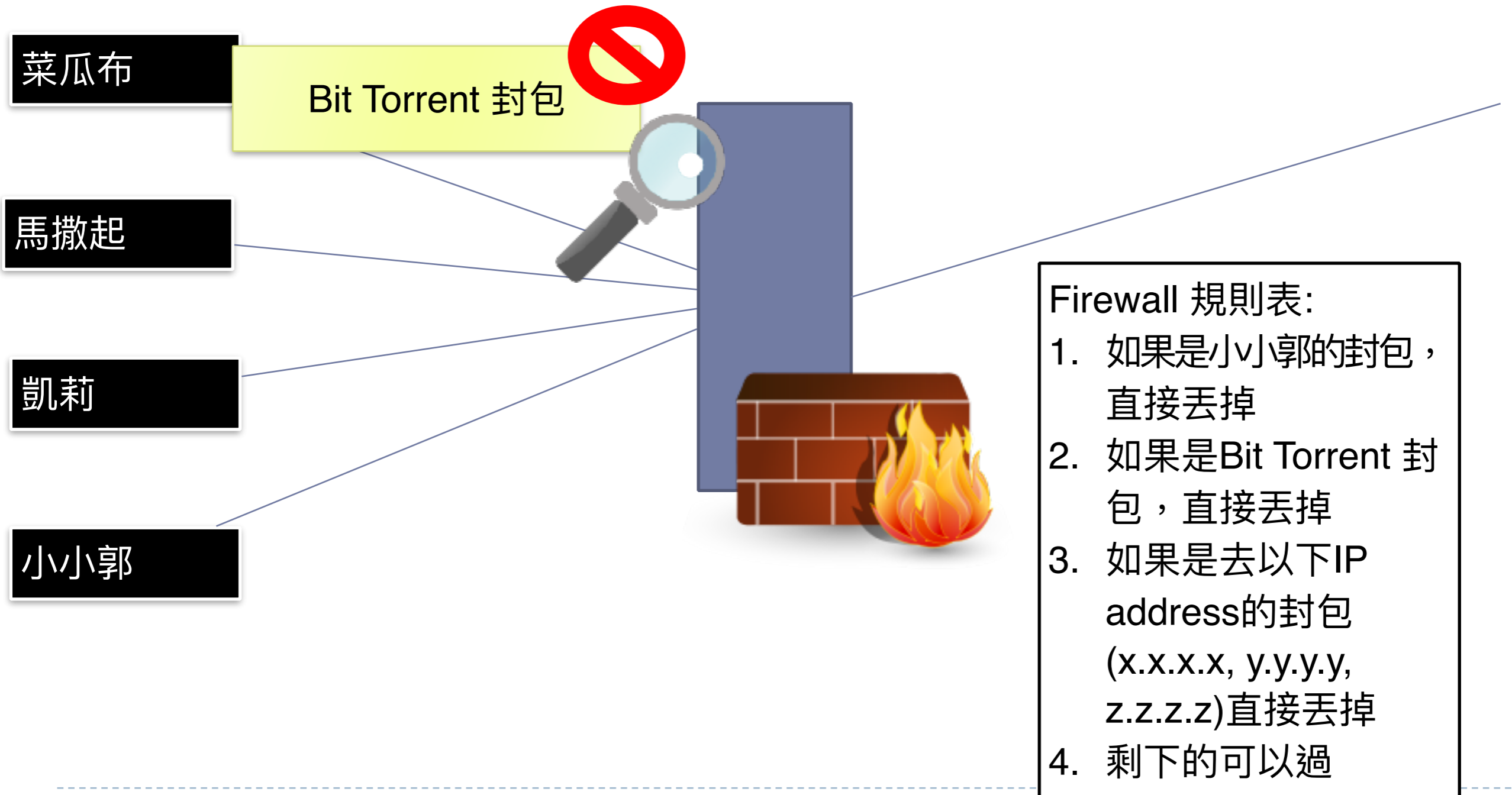
8.8.8.8 的 Ping 統計資料:
    封包: 已傳送 = 4, 已收到 = 4, 已遺失 = 0 (0% 遺失),
    大約的來回時間 (毫秒):
        最小值 = 20ms, 最大值 = 20ms, 平均 = 20ms

C:\Users\Administrator>
```

3. TTL expire (用來偵測或預防路徑中的loop或是traceroute使用)



Firewall

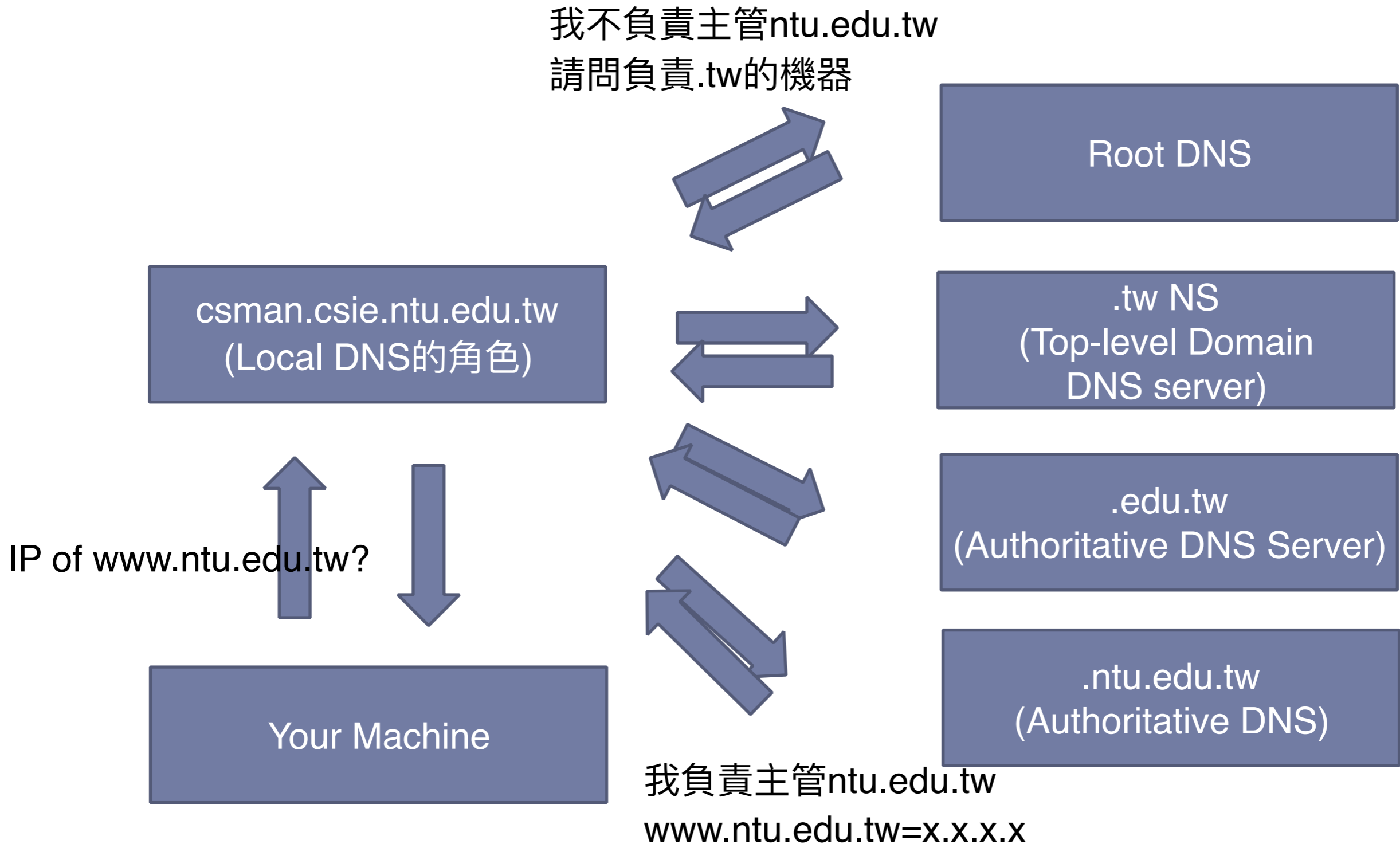


DNS (Domain Name Service)

- ▶ 一言以蔽之: 將名稱轉為IP的服務
- ▶ 常見的轉換種類:
 - ▶ Domain name -> IP (type A):
ntucsv.csie.ntu.edu.tw -> 140.112.30.28
 - ▶ @domainname的mail server (type MX):
csie.ntu.edu.tw -> ms.csie.ntu.edu.tw
 - ▶ Domain name -> domain name (type CNAME):
www.csie.ntu.edu.tw -> ntucsv.csie.ntu.edu.tw
 - ▶ IP -> domain name (type PTR)
140.112.30.21 -> csman.csie.ntu.edu.tw
- ▶ 可以多重宣告: 增加可靠度或分散性.
 - ▶ 例如www.google.com的A指到了6個IP!



分散式的架構: 分層負責 (recursive query)



DNS的細節

- ▶ 如果local DNS本身主管被查詢的domain，則可以直接回覆。
 - ▶ 例如140.112.30.21如果被查詢www.csie.ntu.edu.tw
- ▶ Local DNS可以暫存之前查詢過的結果。
 - ▶ 主要用來減輕主管DNS server及網路的負擔。
 - ▶ 每筆在主管DNS server上的紀錄都有對應的TTL值，規範可以被占存多久。



/etc/resolv.conf

nameserver 140.112.30.21

nameserver 140.112.254.4

nameserver 140.112.2.2

search csie.ntu.edu.tw

- search
—> resolve incomplete names (linux1 —> linux1.csie.ntu.edu.tw)
- nameserver —> specify the address of the DNS server

DNS延伸閱讀

- Top 10 DNS attacks:
<http://www.networkworld.com/article/2886283/security0/top-10-dns-attacks-likely-to-infiltrate-your-network.html#slide1>
- Distributed Reflection DoS attack
- Cache poisoning / DNS hijacking (sol: DNSSEC)
- TCP SYN floods

常用DNS指令

- Examples:
 - `dig @8.8.8.8 -t MX csie.ntu.edu.tw`
 - `dig @140.112.30.21 www.csie.ntu.edu.tw`

```
;; ANSWER SECTION:
www.csie.ntu.edu.tw.      600      IN       A       140.112.30.28

;; AUTHORITY SECTION:
csie.ntu.edu.tw.        86400    IN       NS      csman2.csie.ntu.edu.tw.
csie.ntu.edu.tw.        86400    IN       NS      ntuns.ntu.edu.tw.
csie.ntu.edu.tw.        86400    IN       NS      csman.csie.ntu.edu.tw.

;; ADDITIONAL SECTION:
csman.csie.ntu.edu.tw.  600      IN       A       140.112.30.21
ntuns.ntu.edu.tw.       85489    IN       A       140.112.254.6
csman2.csie.ntu.edu.tw. 600      IN       A       140.112.30.12
```

In-Class Exercise

- 找出linux1到www.nasa.gov經過了哪些機器(domain name可) keyword: mtr, traceroute
- 找出csie.ntu.edu.tw和ntu.edu.tw的mail server們(SMTP)的IP是什麼

延伸閱讀

- 前講師(小小郭)的線上投影片:
<http://xdlab.org/~math120908/slides/nettool.html#/introduction-to-network-tools>