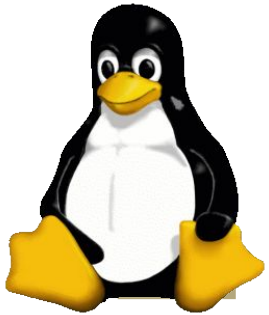


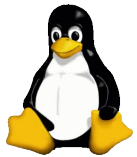
Linux Network Administration

Maruthi Inukonda

16th Feb 2019

Agenda

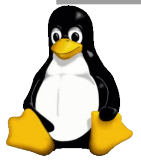
- TCP/IP Stack Concepts
- Physical Layer
- Data Link Layer in Linux
- Network layer in Linux
- Transport layer in Linux
- Application layer in Linux



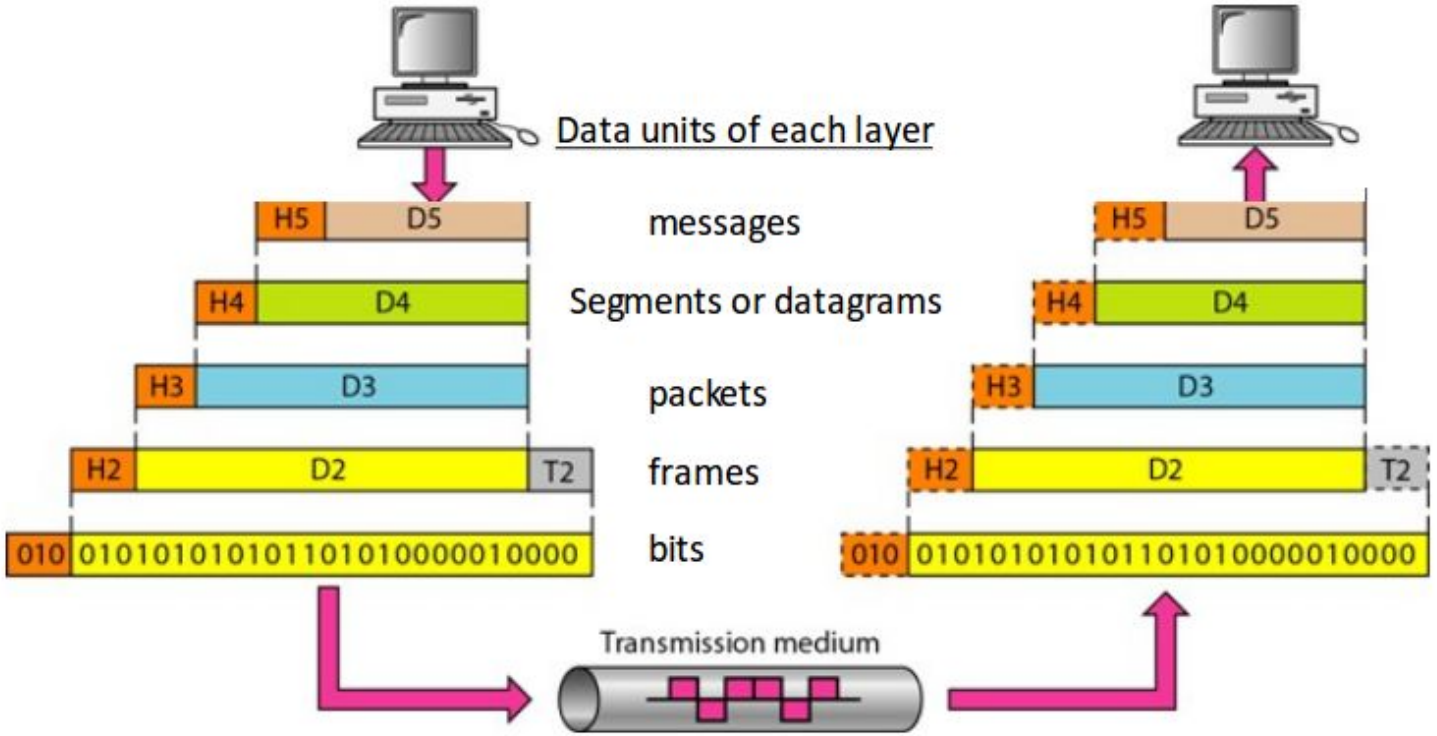
TCP/IP Stack Concepts

TCP/IP Stack - Layers - Addresses

Layer #	Layer Name	Protocols	Protocol Data Unit	Addressing	Implemented in
L5	Application	DHCP, DNS, HTTP, SMTP, SSH, FTP, LDAP, NFS, NTP	Message	Socket Handle/File Descriptor	User space
L4	Transport	TCP, UDP	Datagram	Port number	Kernel space
L3	Network	IP, ICMP, ARP	Packet	IP address	Kernel space
L2	Data Link	Ethernet, WiFI	Frame	MAC address	Hardware
L1	Physical	1000baseT, 802.11	bits	None	Hardware

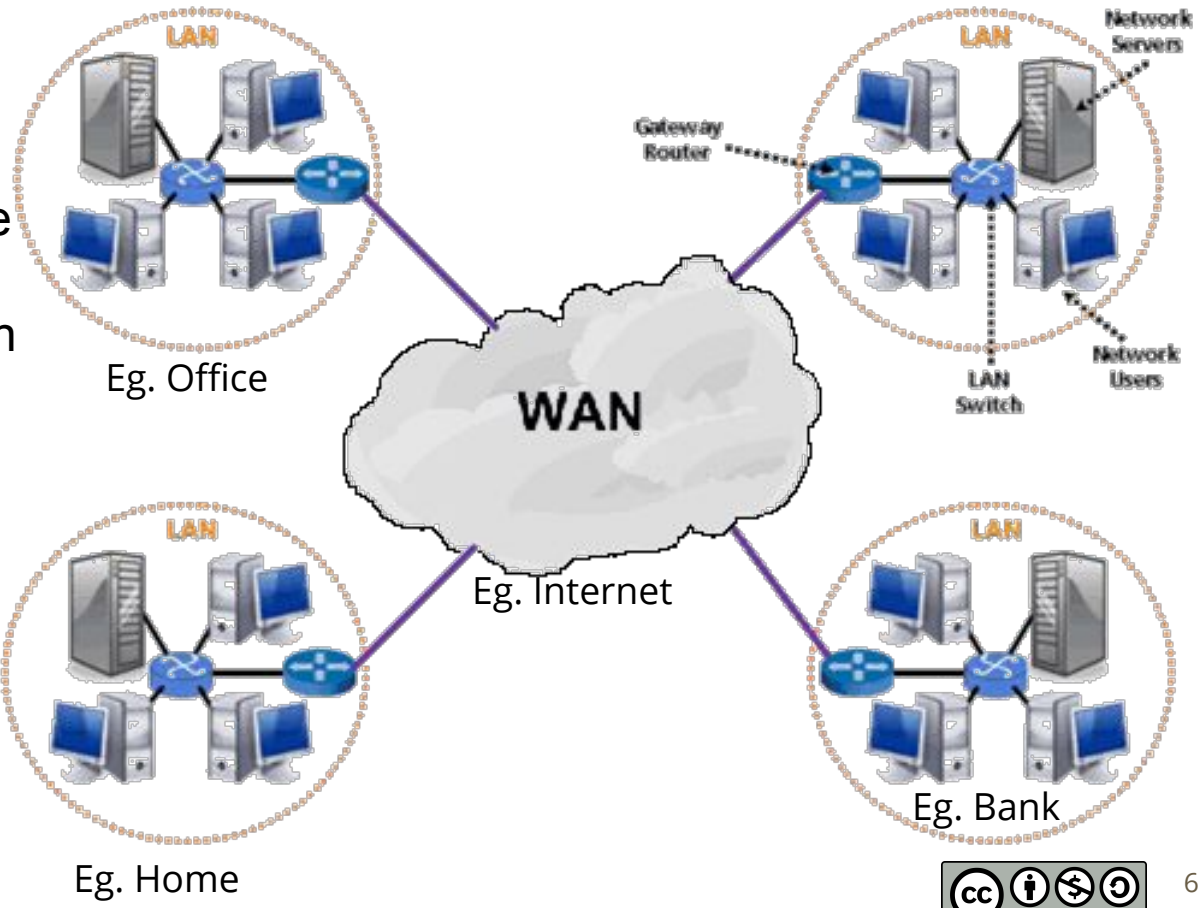


TCP/IP Stack - Layers - Encapsulation - Decapsulation



Local Area Network (LAN), Wide Area Network (WAN)

- LAN : network in small geographical area.
  Eg. college/house/office
  Eg. Intranet.
- WAN : network between networks.
  Eg. Internet

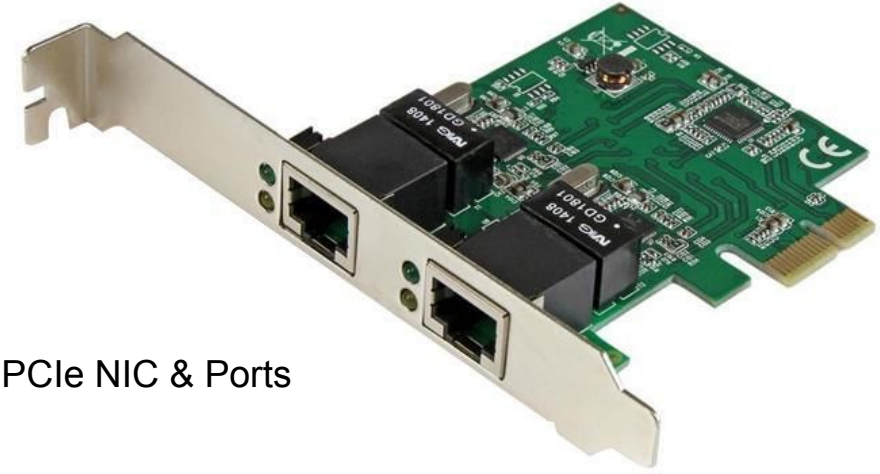


Physical Layer

Network Interface Controller (NIC)

- Used for networking between computers.
- Available in two models
  - Onboard controller
  - Add-on card in PCIe slot.

Onboard NIC Ports

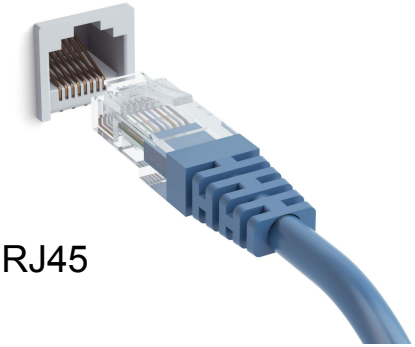


PCIe NIC & Ports

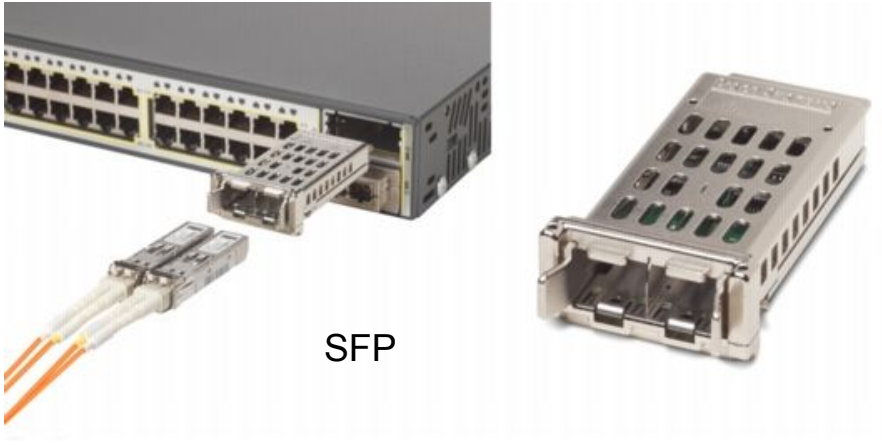
Figures not to the scale

Network Connectors

- Registered Jack 45 (RJ45)
  - Available in 1, 10Gbps
- Small Form-factor Pluggable (SFP)
  - Max : 100 Gbps
  - Also available in 10, 40Gbps.



RJ45



SFP

Figures not to the scale

Hub

- Physical layer (L 1) device
- Does not learn MAC addresses of devices in its ports.
- Broadcasts every bit to every port.
- Contains only one broadcast domain.
- Half-duplex communication
- Max: 10 Mbps.
- Have become extinct now a days.
- Sometimes called unmanaged switch.

Hub

Figures not to the scale

Network cards

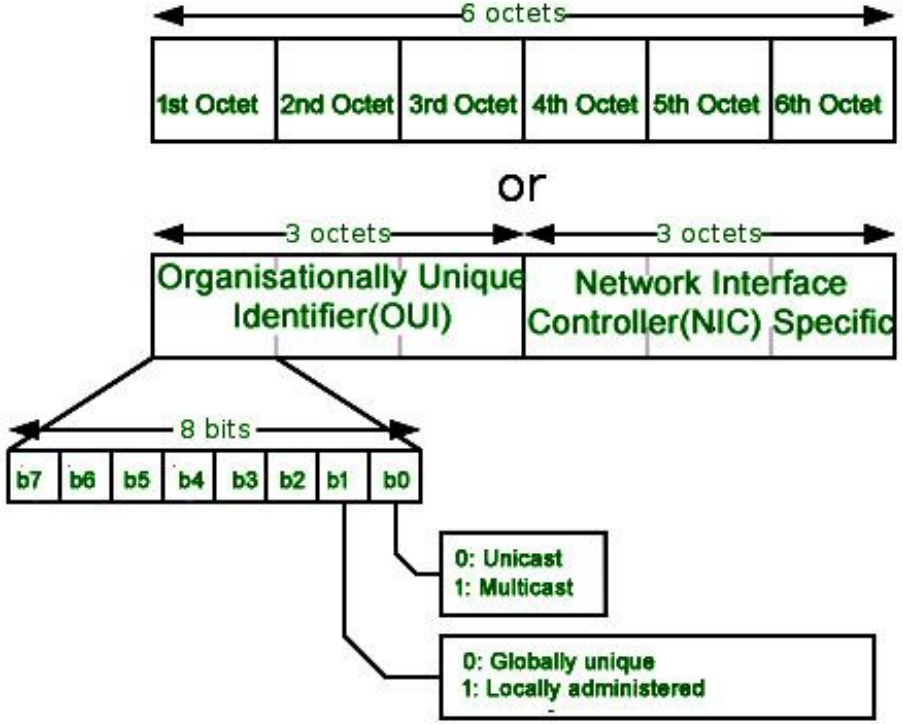
- List network interface cards using `lspci`

```
$ lspci
...
00:1f.6 Ethernet controller: Intel Corporation Ethernet Connection (4) I219-V (rev 21)
...
05:00.0 Network controller: Intel Corporation Device 24fd (rev 78)
...
```

Data Link Layer

Hardware addresses

- MAC address
- 6 bytes (48 bit)
- Written as 6 colon separated octets
  Eg.
  74:86:0b:28:fb:4d
  ac:ed:5c:11:bc:8c
- First 3 octets represent Organizationally Unique Identifier (OUI)
  Eg. 74-86-0b - Cisco
  ac:ed:5c - Intel



http://standards-oui.ieee.org/oui.txt

Switch

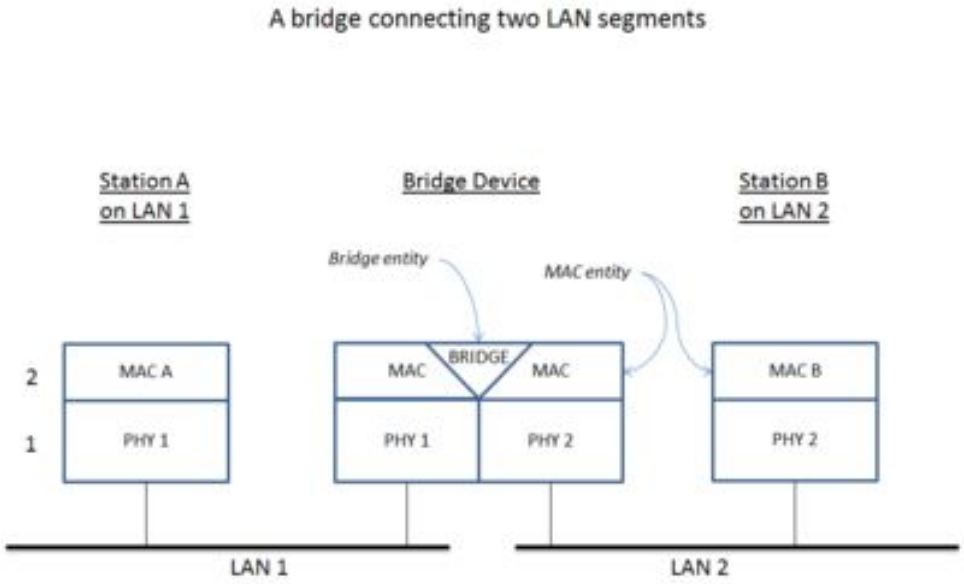
- Data Link layer (L2) device.
- Learns MAC addresses of devices connected to its ports.
- Unicasts frames to right port based on MAC address..
- Full/Half-duplex communication.
- Can contain more than one broadcast domain using VLANs.

Switch

Figure not to the scale

Bridge

- Data Link layer (L2) device.
- Used to connect multiple network segments of same network.
- Learns MAC addresses of devices connected to its ports.
- Unicasts frames to right port.
- Full/Half-duplex communication.



Bridge

Figure not to the scale

WiFi Access Point

- Data link layer (L2) bridge device.
- Used to connect different segments (Ethernet and WiFi) of same network.
- Learns MAC addresses of devices connected to its Ethernet port and Wireless link.
- Unicasts frames to right port based on MAC address.
- Full/Half-duplex communication.
- Used in large enterprises.

Wifi Access Point

Figures not to the scale

Network interfaces

- List network logical interfaces using `ifconfig`

Ethernet

- maximum frame size is 1518 bytes
- 18 bytes are overhead (header and FCS)
- Maximum Transfer Unit (MTU) is 1500 bytes.

```
$ ifconfig -a
enp0s31f6 Link encap:Ethernet  HWaddr 54:e1:ad:28:fb:4d
          BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0
carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)
          Interrupt:16 Memory:f1200000-f1220000

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          inet6 addr: ::1/128 Scope:Host
          UP LOOPBACK RUNNING  MTU:65536  Metric:1
          ...

wlp5s0    Link encap:Ethernet  HWaddr ac:ed:5c:11:bc:8c
          BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0
carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)
```

Wired interface details

- Wired (Ethernet) interface details using `ethtool <interface_name>`

```
# ethtool enp0s10
Settings for enp0s10:
    Supported ports: [ TP ]
    Supported link modes:   10baseT/Half 10baseT/Full
                            100baseT/Half 100baseT/Full
                            1000baseT/Full
    ...
    Advertised pause frame use: No
    Advertised auto-negotiation: Yes
    Speed: 1000Mb/s
    Duplex: Full
    Port: Twisted Pair
    PHYAD: 0
    Transceiver: internal
    Auto-negotiation: on
    MDI-X: off (auto)
Cannot get wake-on-lan settings: Operation not
permitted
    Current message level: 0x00000007 (7)
                drv probe link
    Link detected: yes
```

Wireless interface details

- Wireless (Wifi) interface details using `iwconfig`

```
$ iwconfig
wlp0s2f1u7  IEEE 802.11bgn  ESSID:"meghaduta"  Nickname:"<WIFI@REALTEK>"
          Mode:Managed  Frequency:2.457 GHz  Access Point: 48:EE:0C:46:B5:5E
          Bit Rate:72.2 Mb/s   Sensitivity:0/0
          Retry:off   RTS thr:off   Fragment thr:off
          Encryption key:****-****-****-****-****-****-****-****   Security mode:open
          Power Management:off
          Link Quality=15/100  Signal level=-68 dBm  Noise level=0 dBm
          Rx invalid nwid:0  Rx invalid crypt:0  Rx invalid frag:0
          Tx excessive retries:0  Invalid misc:0   Missed beacon:0
```

Note: `Link Quality` is helpful in finding out a location with better wifi coverage.

Network Layer

IP addresses

- Two versions V4, V6. IP v4 is most commonly used.
- V4 address is 32-bit, divided into 4 octets. (dotted decimal notation).
  Eg. 192.168.1.30
- Maximum 4 billion (2^{32}) addresses.
- 32-bits are divided into network bits and host bits.
- Network mask is used to separate network and host bits from IP address.
- IP address exhaustion is addressed using private and public addresses.
- Historically addresses were split into classes (A, B, C, D, E).
- Currently there are no classless (Classless InterDomain Routing) for flexible network and host bits.
- Subnetting and Supernetting for simplicity.

Private, Reserved, Public IP v4 addresses

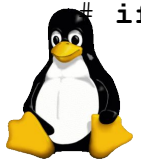
Class	Name #	Private address range	No of IP addresses
A	24-bit block	10.0.0.0 - 10.255.255.255	2^{24} - 2
A	Loopback / Diagnostic	127.0.0.0 - 127.255.255.255	2^{24}- 2
B	Link Local	169.254.0.0 - 169.254.255.255	2^{16} - 2
B	20-bit block	172.16.0.0 - 172.31.255.255	2^{20} - 2
C	16-bit block	192.168.0.0 - 192.168.255.255	2^{16}- 2
D	Multicast	224.0.0.0 - 239.255.255.255	16 x (2^{24}- 2)
E	Reserved	240.0.0.0 - 254.255.255.255	15 x (2^{24}- 2)
E	Broadcast	255.0.0.0.0 - 255.255.255.255	2^{24}- 2

- Private addresses can be used only inside a LAN.
- Addresses other than the Private & Reserved are public addresses routable on the Internet

Static IP (ubuntu)

- Set `iface, address, gateway, netmask, network, dns-nameservers` in `/etc/network/interfaces`

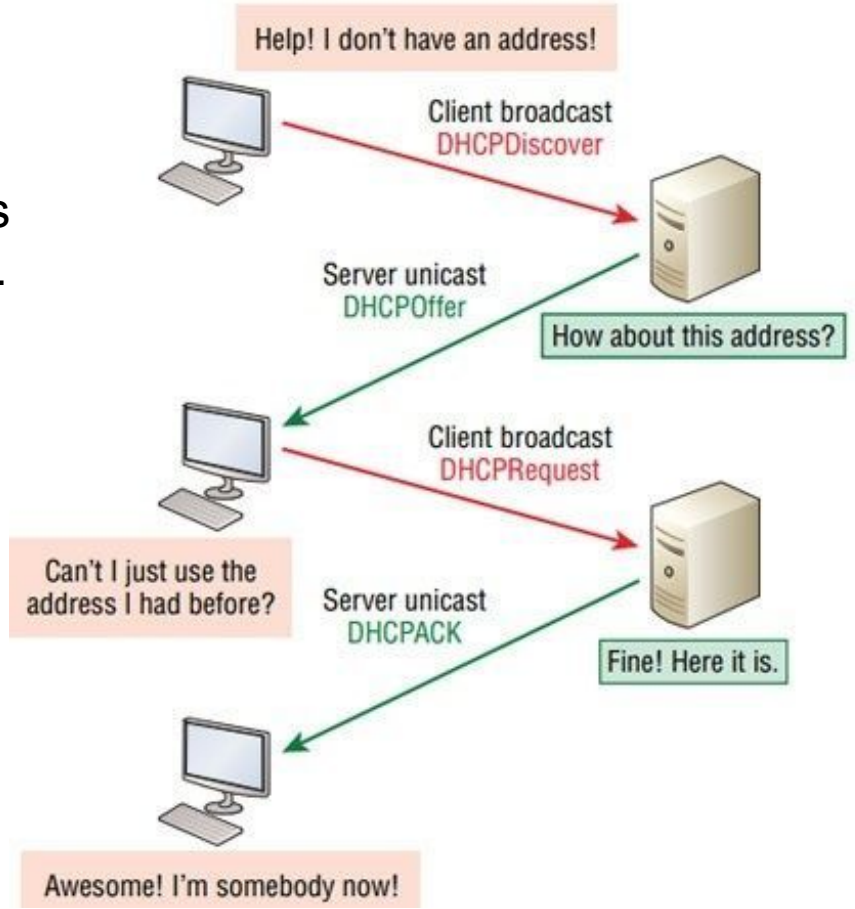
```
$ sudo systemctl disable NetworkManager.service
$ sudo systemctl stop NetworkManager.service
$ sudo cat /etc/network/interfaces
...
##Static IP Configuration enp0s10
auto enp0s10
iface enp0s10 inet static
address 172.16.0.200
gateway 172.16.0.1
netmask 255.255.255.0
network 172.16.0.0
dns-nameservers 172.16.0.1

# ifdown enp0s10
# ifup enp0s10
```



Note: For laptops/desktops it is more convenient to use NetworkManager (GUI)

Dynamic IP

- IPs are dynamically assigned to network devices other than servers (laptops, mobiles, some desktops).
- Dynamic Host Configuration Protocol (DHCP) server does this.
- DORA Process
  - Discover
  - Offer
  - Request
  - Acknowledge



Dynamic IP (ubuntu server)

- For DHCP, edit `auto` and `iface` in `/etc/network/interfaces`

```
# systemctl disable NetworkManager.service
# systemctl stop NetworkManager.service

# cat /etc/network/interfaces
...
##To configure DHCP
auto enp0s7
iface enp0s7 inet dhcp
...

# ifdown enp0s7
# ifup enp0s7
```

Dynamic IP (ubuntu desktop)

- For DHCP, edit network settings using `nmcli` command

```
# systemctl enable NetworkManager.service
# systemctl restart NetworkManager.service
# systemctl status NetworkManager.service

# nmcli dev status
DEVICE     TYPE      STATE          CONNECTION
enp0s31f6  ethernet  connected      DHCP-Ethernet
wlp5s0     wifi      disconnected   --
lo         loopback  unmanaged       --

# nmcli dev connect wlp5s0
# nmcli dev status
DEVICE     TYPE      STATE     CONNECTION
...
wlp5s0     wifi      connected  meghaduta
...

# nmcli dev disconnect wlp5s0
```

Network interfaces

- List network logical interfaces using `ifconfig`

```
$ ifconfig
enp0s31f6 Link encap:Ethernet  HWaddr 54:e1:ad:28:fb:4d
          inet addr:172.16.0.104  Bcast:172.16.0.255
Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:1812542 errors:0 dropped:0 overruns:0 frame:0
          TX packets:979079 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:1354135937 (1.3 GB)  TX bytes:184552688 (184.5 MB)
          Interrupt:16 Memory:f1200000-f1220000

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          inet6 addr: ::1/128 Scope:Host
          UP LOOPBACK RUNNING  MTU:65536  Metric:1
          ...

wlp5s0    Link encap:Ethernet  HWaddr ac:ed:5c:11:bc:8c
          inet addr:172.16.0.105  Bcast:172.16.0.255
Mask:255.255.255.0
          UP BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:52450 errors:0 dropped:0 overruns:0 frame:0
          TX packets:30418 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:31103440 (31.1 MB)  TX bytes:7327547 (
```

Checking Connectivity

- Internet Control Message Protocol (ICMP) echo request and echo reply packets are used to check connectivity.
- Check Network layer connectivity, latency using `ping`

```
$ ping -c 3 172.16.0.1
PING 172.16.0.1 (172.16.0.1) 56(84) bytes of data.
64 bytes from 172.16.0.1: icmp_seq=1 ttl=255 time=2.86 ms
64 bytes from 172.16.0.1: icmp_seq=2 ttl=255 time=1.94 ms
64 bytes from 172.16.0.1: icmp_seq=3 ttl=255 time=1.47 ms

--- 172.16.0.1 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2003ms
rtt min/avg/max/mdev = 1.471/2.093/2.865/0.580 ms
```

Finding MAC address

- Find MAC address of other networked device using `arping`

```
# arping -c 3 -I wlp0s2f1u7 172.16.0.1
ARPING 172.16.0.1 from 172.16.0.105 wlp0s2f1u7
Unicast reply from 172.16.0.1 [48:EE:0C:46:B5:5E]  2.807ms
Unicast reply from 172.16.0.1 [48:EE:0C:46:B5:5E]  2.561ms
Unicast reply from 172.16.0.1 [48:EE:0C:46:B5:5E]  2.566ms
Sent 3 probes (1 broadcast(s))
Received 3 response(s)
```

Note: This command is helpful in detecting redundant IP addresses.

Router

- Network layer (L3) device.
- Used to interconnect different networks.
- Learns MAC addresses and IP addresses of devices connected to its ports.
- Unicasts frames to right port based on IP address.
- Full/Half-duplex communication.
- Contains more than one broadcast domain using network address.

Ethernet Router

Figures not to the scale

WiFi Router

- Network layer (L3) device.
- Used in homes and small offices.
- It also contains

  Hardware:
  - NICs
  - Modem (optional)
  - Bridge
  - Router
  - Wifi access point

  Software :
  - DHCP server
  - NAT server
  - DNS relay server

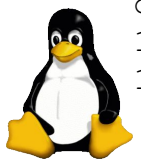
Home Wifi + Ethernet Router

Figures not to the scale

Routing Table

- Find routing table using `route -n` or `netstat -r` or `ip route show`

```
$ route -n
Kernel IP routing table
Destination      Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0          172.16.0.1      0.0.0.0         UG    600    0        0 wlp0s2f1u7
172.16.0.0       0.0.0.0         255.255.255.0   U     600    0        0 wlp0s2f1u7
...

$ netstat -r
Kernel IP routing table
Destination     Gateway         Genmask         Flags   MSS Window  irtt Iface
default         172.16.0.1      0.0.0.0         UG        0 0          0 enp0s31f6
link-local      *               255.255.0.0     U         0 0          0 enp0s31f6
172.16.0.0      *               255.255.255.0   U         0 0          0 enp0s31f6
...

$ ip route show
default via 172.16.0.1 dev enp0s31f6  proto static  metric 100
169.254.0.0/16 dev enp0s31f6  scope link  metric 1000
172.16.0.0/24 dev enp0s31f6  proto kernel  scope link  src 172.16.0.104  metric 100
```



Transport Layer

Port Numbers

- Port number (16-bit) is used to uniquely identify server and client endpoint in transport layer within a computer.
- There are three types
  - Well-known (0 - 1023) : used by standard server programs.
    Eg. 22 for ssh, 53 for dns, 443 for https
  - Registered (1024 - 49151) : used by non-standard server programs.
    Eg 7070 for rtsp, 5353 mdns, 8080 for http
  - Ephemeral (49152 - 65535) : used by client programs

Port scanning

- Find ports open by a computer use `nmap`
- Not all computers allow port scan, as it could be exploited.

```
$ nmap 172.16.0.1

Starting Nmap 7.01 ( https://nmap.org ) at 2019-02-15 21:02 IST
Nmap scan report for 172.16.0.1
Host is up (0.030s latency).
Not shown: 997 closed ports
PORT       STATE SERVICE
80/tcp     open  http
7777/tcp   open  cbt
52869/tcp open  unknown

Nmap done: 1 IP address (1 host up) scanned in 0.62 seconds
```

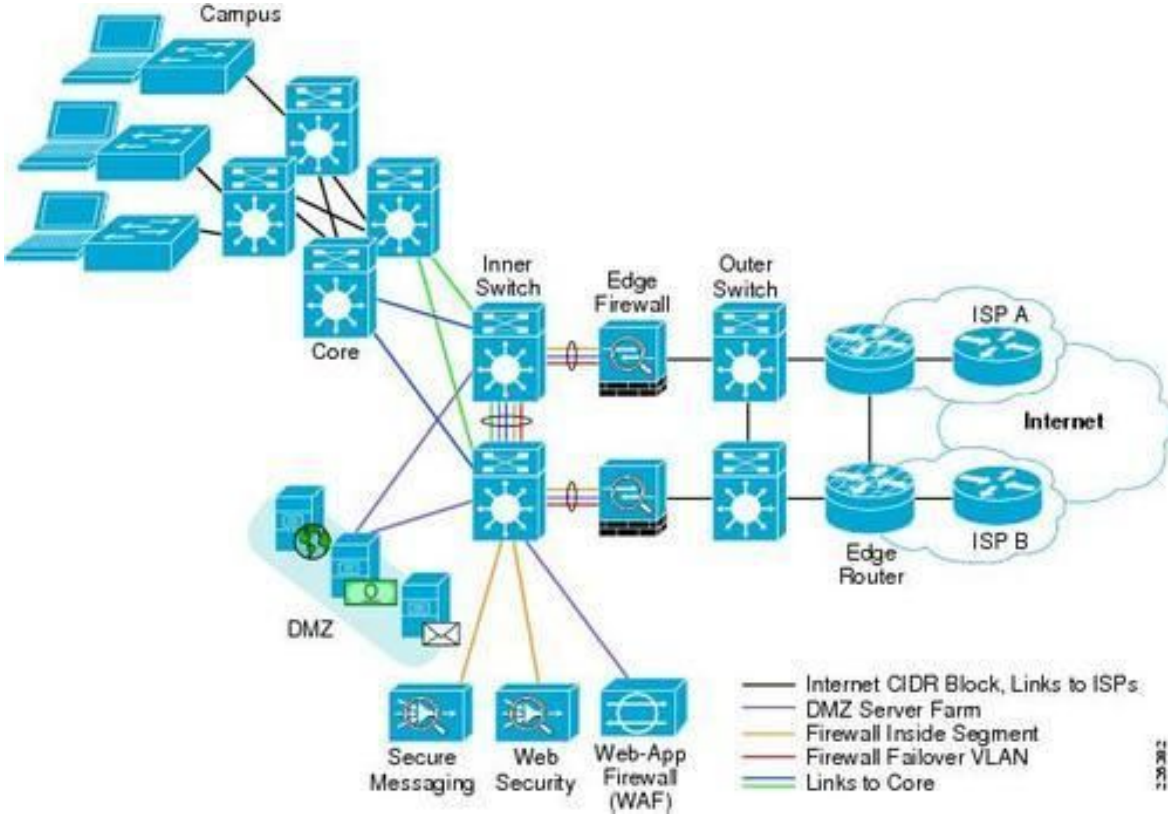
Finding open ports from inside

- Find ports open by a server from inside using `netstat`
- Both listening ports and ephemeral ports are listed

```
$ netstat -np
(Not all processes could be identified, non-owned process info
 will not be shown, you would have to be root to see it all.)
Active Internet connections (w/o servers)
Proto Recv-Q Send-Q Local Address           Foreign Address         State       PID/Program name
tcp        0      0 172.16.0.104:43592      162.125.18.133:https    ESTABLISHED 9057/dropbox
tcp        0      0 172.16.0.104:51704      maa05s09-in-f10.1:https ESTABLISHED 10836/firefox
tcp        32     0 172.16.0.104:52288      162.125.81.3:https      CLOSE_WAIT  9057/dropbox
tcp        0      0 172.16.0.104:58210      bom12s01-in-f14.1:https ESTABLISHED 10836/firefox
tcp        0      0 172.16.0.104:54166      maa05s09-in-f10.1:https ESTABLISHED 10836/firefox
...
Active UNIX domain sockets (w/o servers)
Proto RefCnt Flags       Type       State         I-Node   PID/Program name     Path
unix  2      [ ]         DGRAM                    37427 3199/systemd
/run/user/1001/systemd/notify
...
unix  3      [ ]         STREAM     CONNECTED     39523 3614/compiz
unix  3      [ ]         SEQPACKET  CONNECTED     42292 4201/firefox
...
unix  3      [ ]         STREAM     CONNECTED     42266 3710/dropbox
```

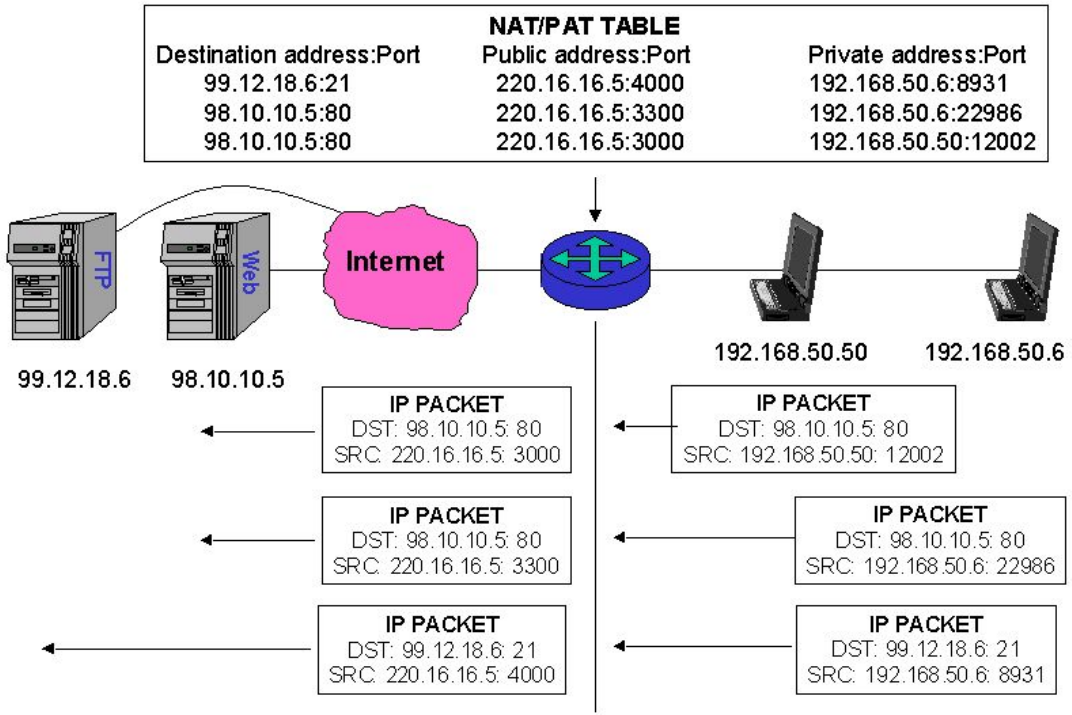
ISP and LANs

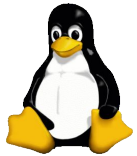
- Internet Service Provider (ISP)
- Multiple ISPs for failover and load-balancing



NAT

- Network Address Translation (NAT)
- Internet facing routers in Home, Office, ISP do private to public address translation and vice-versa.



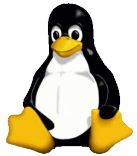


Application Layer

Sockets

- Socket file descriptor (handle) is used to uniquely identify a connection(tcp or udp) in application layer process.
- Find open sockets using `ss` command

```
$ ss -ta
State       Recv-Q Send-Q Local Address:Port               Peer Address:Port
LISTEN      0      128        *:ssh                         *:*
LISTEN      0      5      127.0.0.1:ipp                     *:*
LISTEN      0      128    127.0.0.1:17600                   *:*
...
ESTAB       0      0      172.16.0.104:49786                162.125.34.129:https
CLOSE-WAIT  32     0      172.16.0.104:57242                162.125.34.6:https
TIME-WAIT   0      0      172.16.0.104:54814                172.217.26.161:https
CLOSE-WAIT  1      0      172.16.0.104:59084                34.192.34.151:https
...
```



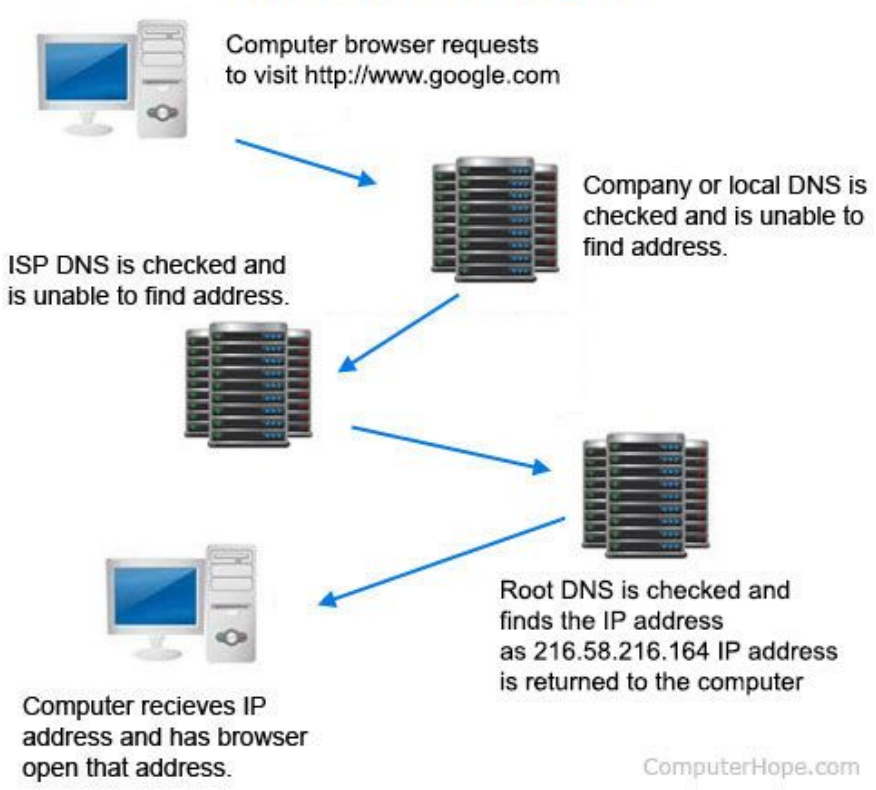
Socket file descriptor

- Find file descriptor associated with open sockets using `lsof` command.
  - `-p` : find file descriptors within a process.
  - `-n` : prevents name-lookup

```
$ lsof -n -p `pidof dropbox`
COMMAND  PID      USER    FD   TYPE DEVICE SIZE/OFF NODE NAME
dropbox 3710 maruthisi   68u  IPv4 208135        0t0  TCP
172.16.0.104:35790->162.125.18.133:https (ESTABLISHED)
dropbox 3710 maruthisi   90u  IPv4  49295        0t0  TCP *:db-lsp (LISTEN)
dropbox 3710 maruthisi   91u  IPv6  49296        0t0  TCP *:db-lsp (LISTEN)
dropbox 3710 maruthisi   92u  IPv4  45242        0t0  UDP *:17500
dropbox 3710 maruthisi   94u  IPv4 145537        0t0  TCP 172.16.0.104:41552->162.125.81.7:https
(CLOSE_WAIT)
dropbox 3710 maruthisi  111u  IPv4  47162        0t0  TCP localhost:17600 (LISTEN)
dropbox 3710 maruthisi  112u  IPv4  45692        0t0  TCP 172.16.0.104:41212->162.125.81.3:https
(CLOSE_WAIT)
...
```

Domain Name Server (DNS)

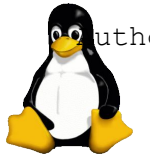
- Name lookup : Convert Fully Qualified Domain Name (FQDN) to Public IP address Eg www.google.com
- Reverse name lookup: Convert Public IP address to FQDN.



DNS lookup

- Hostname → IP address using `nslookup`
- IP address → hostname using `nslookup`

```
$ nslookup www.wikipedia.org
Server:     172.16.0.1
Address:    172.16.0.1#53

Non-authoritative answer:
Name:    www.wikipedia.org
Address: 91.198.174.192

$ nslookup 91.198.174.192
Server:     172.16.0.1
Address:    172.16.0.1#53

Non-authoritative answer:
192.174.198.91.in-addr.arpa    name = text-lb.esams.wikimedia.org.

Authoritative answers can be found from:
```



References

References

- Linux manual pages
- www.wikipedia.org
- Courtesy Google images

Q & A