Steven VanDevender
Hervey Allen
Network Startup Resource Center

PacNOG 6: Nadi, Fiji

IP Basics

Layers

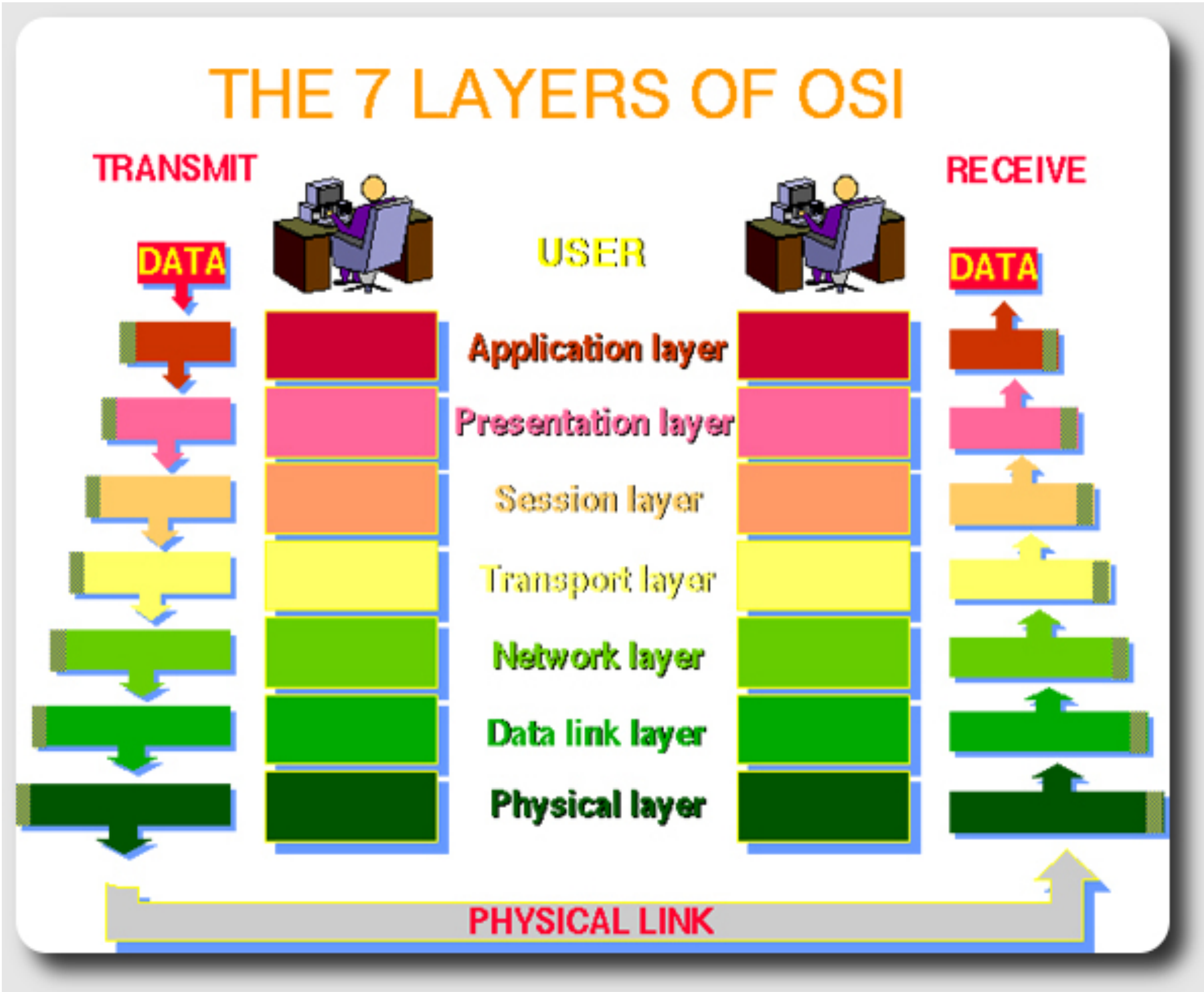
Complex problems can be solved using the common divide and conquer principle. In this case the internals of the Internet are divided into separate layers.

- Makes it easier to understand
- Developments in one layer need not require changes in another layer
- Easy formation (and quick testing of conformation to) standards

Two main models of layers are used:

- OSI (Open Systems Interconnection)
- TCP/IP

OSI Model


THE 7 LAYERS OF OSI
TRANSMIT
RECEIVE
DATA
USER
DATA
Application layer
Presentation layer
Session layer
Transport layer
Network layer
Data link layer
Physical layer
PHYSICAL LINK

OSI

Conceptual model composed of seven layers, developed by the International Organization for Standardization (ISO) in 1984.

Layer 7 – Application (servers and clients etc web browsers, httpd)

Layer 6 – Presentation (file formats e.g pdf, ASCII, jpeg etc)

Layer 5 – Session (conversation initialisation, termination,)

Layer 4 – Transport (inter host comm – error correction, QOS)

Layer 3 – Network (routing – path determination, IP[x] addresses etc)

Layer 2 – Data link (switching – media acces, MAC addresses etc)

Layer 1 – Physical (signalling – representation of binary digits)

Acronym: All People Seem To Need Data Processing

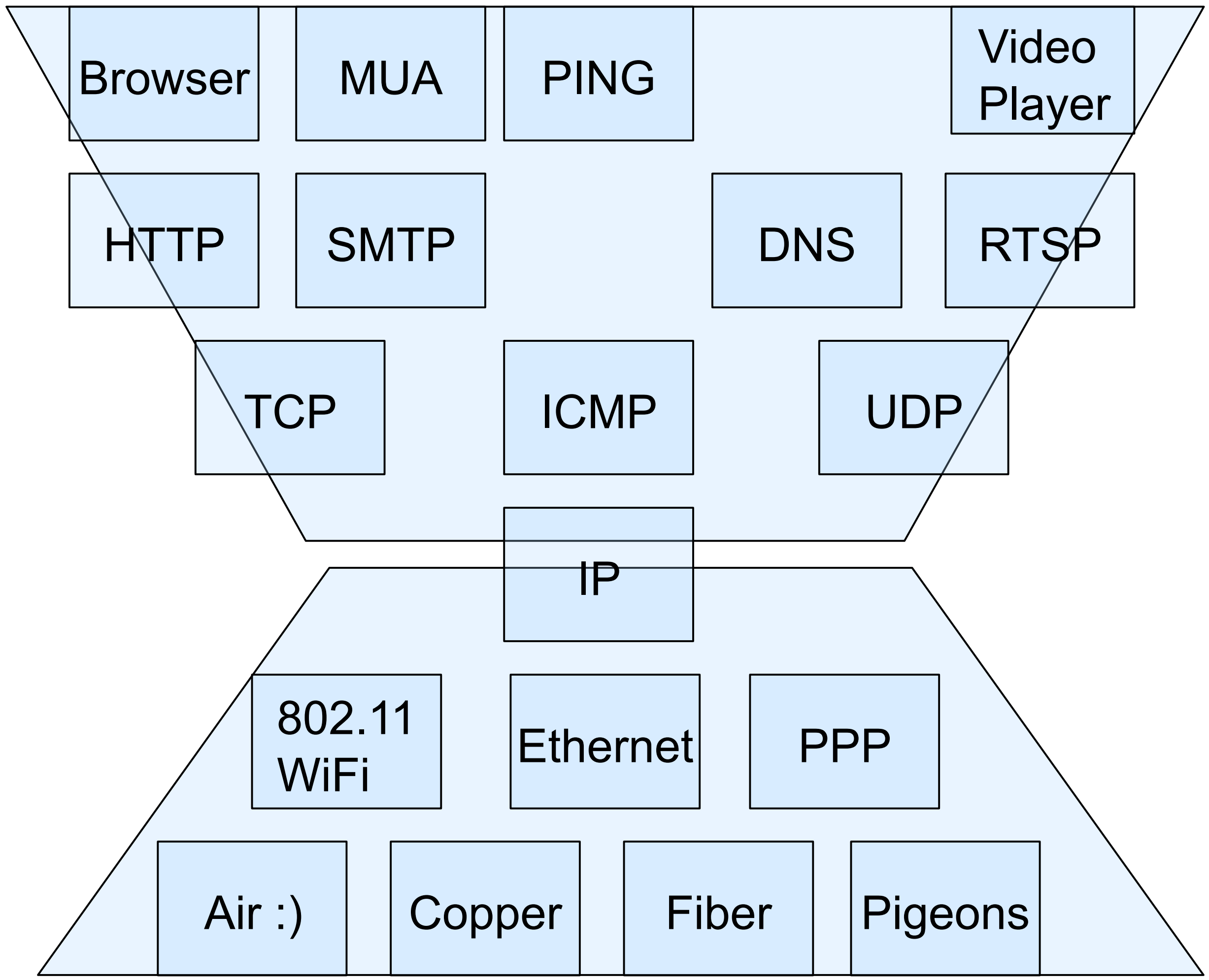
TCP/IP

Generally, TCP/IP (Transmission Control Protocol/Internet Protocol) is described using three to five functional layers. We have chosen the common DoD reference model, which is also known as the *Internet Reference Model*.

- Process/Application Layer consists of applications and processes that use the network.
- Host-to-host transport layer provides end-to-end data delivery services.
- Internetwork layer defines the datagram and handles routing of data.
- Network access layer consists of routines for accessing physical networks.

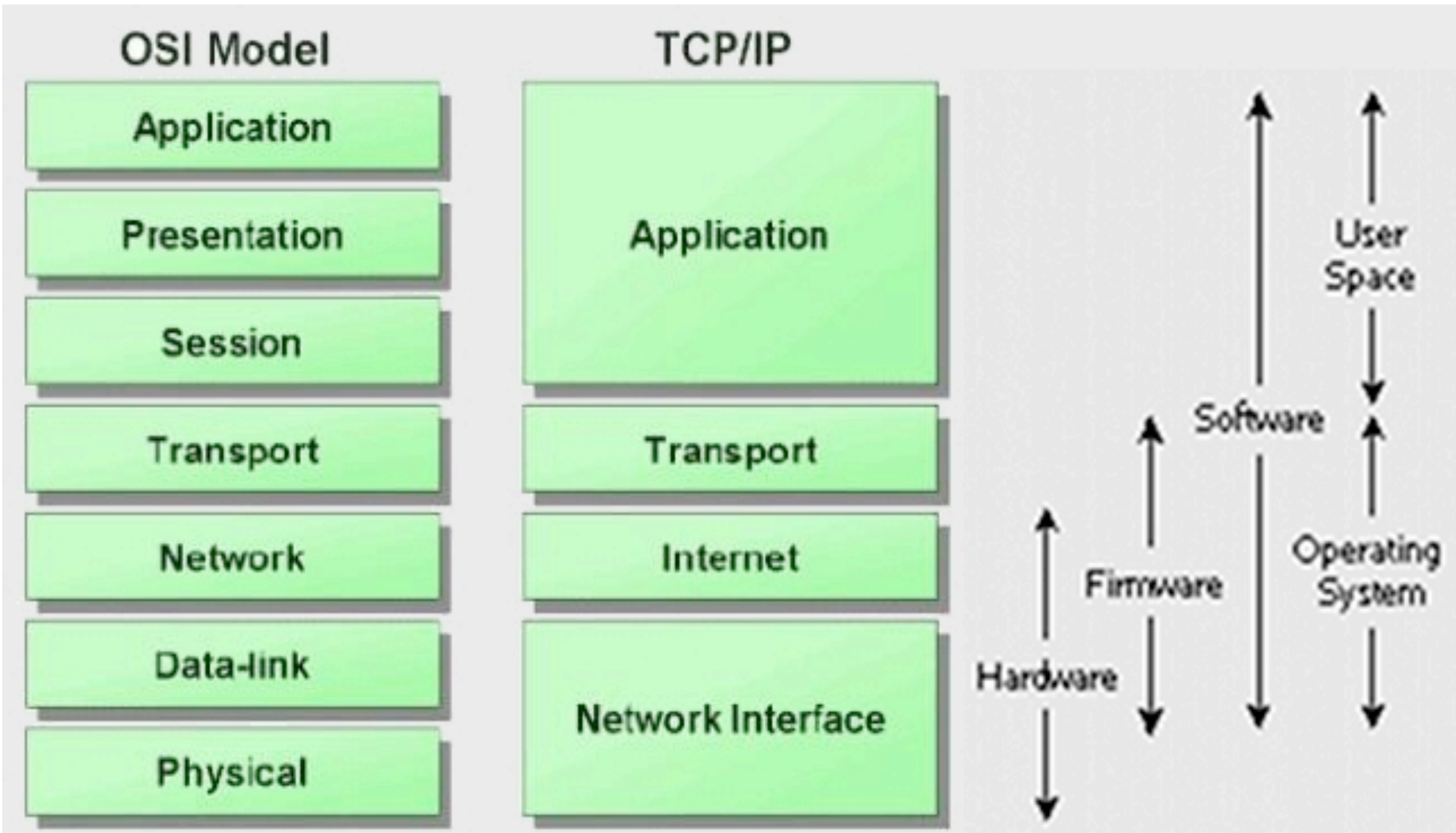
(See
http://en.wikipedia.org/wiki/

TCP/IP model – the “hourglass”



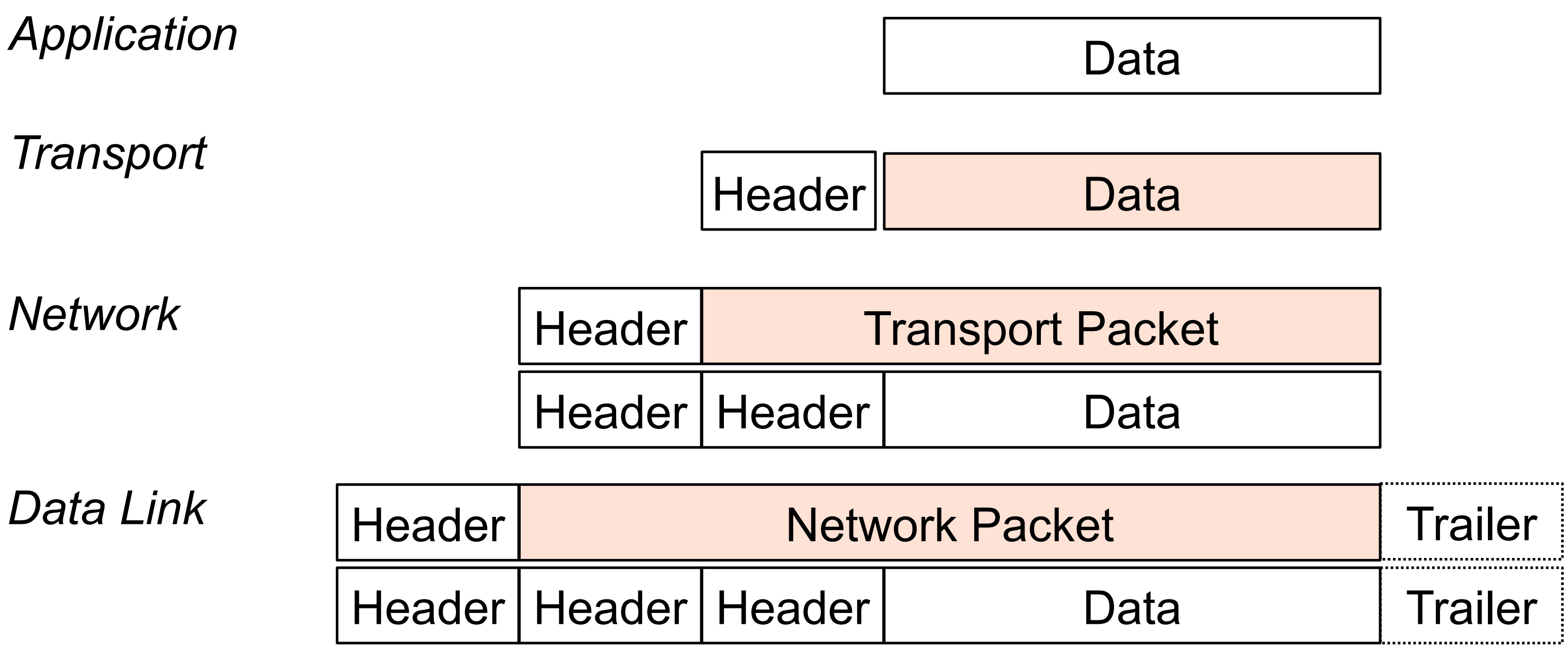
Notice that we do not really have clear 3, 4 or 5 layers here … :)

OSI and TCP/IP



Encapsulation & Decapsulation

Lower layers add headers (and sometimes trailers) to upper layers packets



Frame, Datagram, Segment, Packet

Different names for packets at different layers:

- Ethernet (link layer) frame
- IP (network layer) datagram
- TCP (transport layer) segment

Terminology is not strictly followed

we often just use the term “packet” at any layer

Summary

Networking is a problem approached in layers.

- OSI Layers
- TCP/IP Layers

Each layer adds headers to the packet of the previous layer as the data leaves the machine (encapsulation) and the reverse occurs on the receiving host (decapsulation)

So What is an IPv4 Address Anyway?

32 bit number (4 octet number) can be represented in lots of ways:

133	27	162	125	Decimal
10000101	00011011	10100010	01111101	Binary
85	1B	A2	7D	Hexadecimal

Network Masks

Network Masks help define which bits are used to describe the Network Part and which for hosts

Different Representations:

- decimal dot notation: 255.255.224.0 (128+64+32 in byte 3)
- binary: 11111111 11111111 111 00000 00000000
- hexadecimal: 0xFFFFE000
- number of network bits: /19 (8 + 8 + 3)

Binary AND of 32 bit IP address with 32 bit netmask yields network part of address

More to the Structure

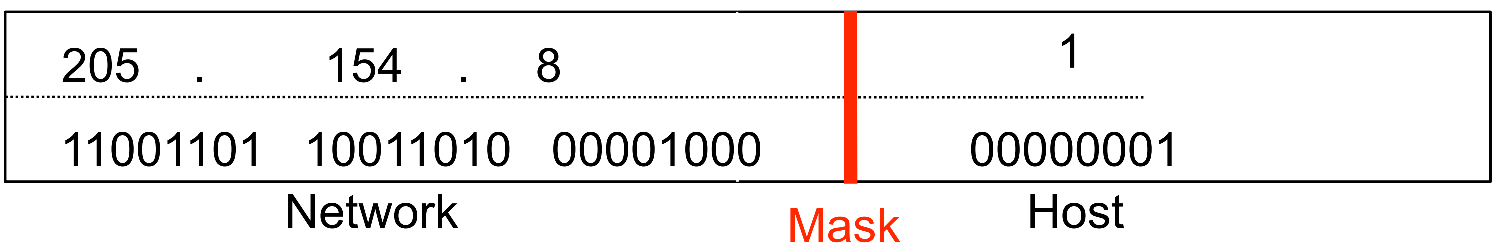
Hierarchical Division in IP Address:

<u>Network Part (Prefix)</u>

Describes which network

<u>Host Part (Host Address)</u>

Describes which host on that network



<u>Boundary can be anywhere</u>

Used to be a multiple of 8 (/8, /16/, /24), but not usual today

Sample Netmasks

137.158.128.0/**17** (netmask **255.255.128.0**)

1111 1111	**1111 1111**	**1 | 000 0000**	**0000 0000**
1000 1001	1001 1110	1 | 000 0000	0000 0000

198.134.0.0/**16** (netmask **255.255.0.0**)

1111 1111	**1111 1111** |	**0000 0000**	**0000 0000**
1100 0110	1000 0110 |	0000 0000	0000 0000

205.37.193.128/**26** (netmask **255.255.255.192**)

1111 1111	**1111 1111**	**1111 1111**	**11 | 00 0000**
1100 1101	0010 0101	1100 0001	10 | 00 0000

Allocating IP Addresses

The subnet mask is used to define size of a network

E.g a subnet mask of 255.255.255.0 or /24 implies 32-24=8 host bits

- 2^8 minus 2 = 254 possible hosts

Similarly a subnet mask of 255.255.255.224 or /27 implies 32-27=5 host bits

- 2^5 minus 2 = 30 possible hosts

Special IP Addresses

All 0's in host part: Represents Network

- e.g. 193.0.0.0/24
- e.g. 138.37.128.0/17
- e.g. 192.168.2.128/25 (Why?)

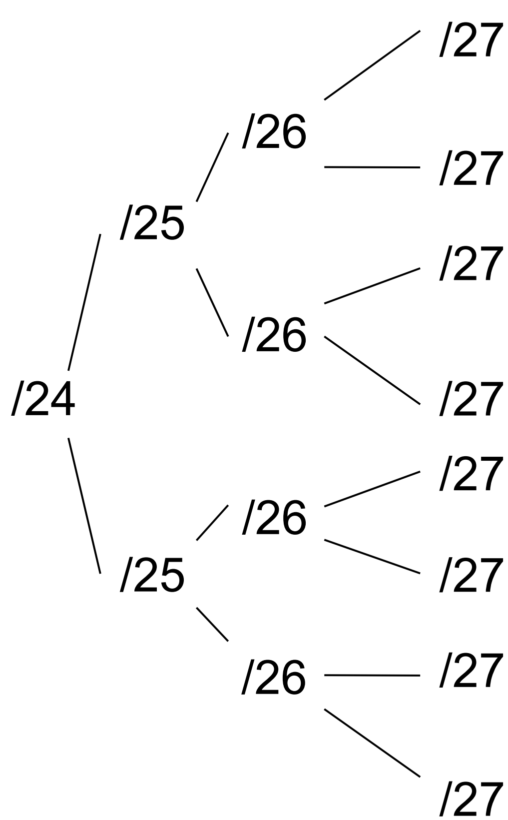
All 1's in host part: Broadcast (all hosts on net)

- e.g. 137.156.255.255 (137.156.0.0/16)
- e.g. 134.132.100.255 (134.132.100.0/24)
- e.g. 192.168.2.127/25 (192.168.2.0/25) (Why?)

127.0.0.0/8: Loopback address (127.0.0.1)

0.0.0.0: Various special purposes (DHCP, etc.)

Networks – Super- and Subnetting

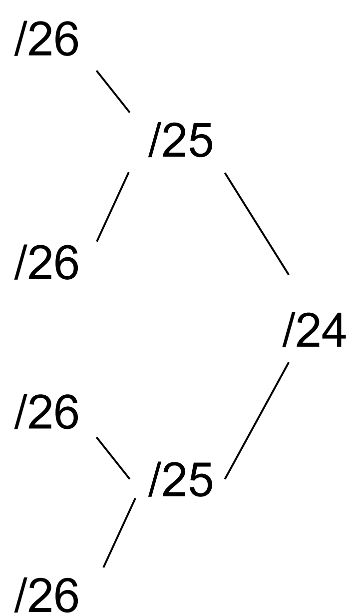


By adding one bit to the netmask, we subdivide the network into two smaller networks. This is *subnetting.*

i.e.: If one has a /26 network (32 – 26 = 6 => 2^6 => 64 addresses), that network can be subdivided into two subnets, using a /27 netmask, where the state of the last bit will determine which network we are addressing (32 – 27 = 5 => 2^5 => 32 addresses). This can be done recursively (/27 => 2 x /28 or 4 x /29, etc...).

Example: 192.168.10.0/25 (.0 - .127) can be subnetted into 192.168.10.0 / 26 and 192.168.10.64 / 26

Networks – Super- and Subnetting



Inversely, if two networks can be "joined" together under the same netmask, which encompasses both networks, then we are *supernetting*.

Example:

Networks 10.254.4.0/24 and 10.254.5.0/24 can be "joined" together into one network expressed: 10.254.4.0/23.

Note: for this to be possible, the networks must be *contiguous*, i.e. it is not possible to supernet 10.254.5.0/24 and 10.254.6.0/24

Numbering Rules

Private IP address ranges (RFC 1918)

- 10/8 (10.0.0.0 – 10.255.255.255)
- 192.168/16 (192.168.0.0 – 192.168.255.255)
- 172.16/12 (172.16.0.0 – 172.31.255.255)

Public Address space available from your Regional Internet Registry

Choose a small block from whatever range you have, and subnet your networks (to avoid problems with broadcasts, and implement segmentation policies – DMZ, internal, etc...)

Regional Internet Registries

RIRs

- AfriNIC
  Africa
- APNIC
  Asia and Pacific
- ARIN
  United States, Canada and parts of Caribbean
- LACNIC
  Latin America and Caribbean
- RIPE NCC
  Europe, Middle East and Central Asia

There is overlap and much history behind these registries.

Some Linux IP-related Settings

Files (Debian/Ubuntu)

`/etc/network/interfaces`

`/etc/hosts`

`/etc/hostname` *(optional)*

Commands

- `ifconfig eth0 192.168.100.x/24`
- `route add default gw 192.168.100.1`
- `hostname pc1.workshop.domain`

Routing

Every host on the internet needs a way to get packets to other hosts outside its local network.

This requires special hosts called **routers** that can move packets between networks.

Packets may pass through many routers before they reach their destinations.

The Route Table

All hosts (including routers) have a **route table** that specifies which networks it is connected to, and how to forward packets to a gateway router that can talk to other networks.

A simple route table (as might be shown by `route -n`):

Destination	Gateway	Genmask	Flags	Interface
128.223.142.0	0.0.0.0	255.255.254.0	U	eth0
0.0.0.0	128.223.142.1	0.0.0.0	UG	eth0

What do Route Table Entries Mean?

Destination	Gateway	Genmask	Flags	Interface
128.223.142.0	0.0.0.0	255.255.254.0	U	eth0
0.0.0.0	128.223.142.1	0.0.0.0	UG	eth0

The **Destination** is a network address.

The **Gateway** is an IP address of a router that can forward packets (or 0.0.0.0, if the packet doesn't need to be forwarded).

The **Genmask** is a netmask.

Flags indicate the type of route.

The **Interface** is the network interface that is connected to that network.

How the Route Table is Used

A packet that needs to be sent has a destination IP address.

For each entry in the route table (starting with the first):

1. Compute the logical AND of the destination IP and the **Genmask** entry.
2. Compare that with the **Gestination** entry.
3. If those match, send the packet out the **Interface**, and we're done.
4. If not, move on to the next entry in the table.

Reaching the Local Network

Suppose we want to send a packet to using this route table.

Destination	Gateway	Genmask	Flags	Interface
128.223.142.0	0.0.0.0	255.255.254.0	U	eth0
0.0128.223.143.42.0.0	128.223.142.1	0.0.0.0	UG	eth0

In the first entry:

128.223.143.42 AND 255.255.254.0 = 128.223.142.0

This matches the **Destination** of the first routing table entry, so send the packet out **Interface** eth0.

The first entry is called a **network route**.

Reaching Other Networks

Suppose we want to send a packet to 72.14.213.99 using this route table?

Destination	Gateway	Genmask	Flags	Interface
128.223.142.0	0.0.0.0	255.255.254.0	U	eth0
0.0.0.0	128.223.142.1	0.0.0.0	UG	eth0

1. 72.14.213.99 AND 255.255.254.0 = 72.14.212.0
2. This does not match the first entry, so move on to the next entry.
3. 72.14.213.99 AND 0.0.0.0 = 0.0.0.0
4. This does match the second entry, so forward the packet to 128.223.142.1 via eth0.

The Default Route

Note that this route table entry:

Destination	Gateway	Genmask	Flags	Interface
0.0.0.0	128.223.142.1	0.0.0.0	UG	eth0

…matches every possible destination IP address. This is called the **default route**.

The gateway device (128.223.142.1) has to be a router capable of forwarding traffic.

More Complex Routing

Consider this route table:

Destination	Gateway	Genmask	Flags	Interface
192.168.0.0	0.0.0.0	255.255.255.0	U	eth0
192.168.1.0	0.0.0.0	255.255.255.0	U	eth1
192.168.2.0	0.0.0.0	255.255.254.0	U	eth2
192.168.4.0	0.0.0.0	255.255.252.0	U	eth3
0.0.0.0	192.168.1.1	0.0.0.0	UG	eth0

This is what a router's routing table might look like. Note that there are multiple interfaces for multiple local networks, and a gateway that can reach other networks.

Forwarding Packets

Any UNIX-like (and other) operating system can function as gateway:

In Linux in /etc/sysctl.conf set:

```
net.ipv4.ip_forward=1
```

Without forwarding enabled, the box will not forward packets from one interface to another: it is simply a host with multiple interfaces.