IP Basics

Unix/IP Preparation Course
May 29, 2011
Dar es Salaam, Tanzania

Layers

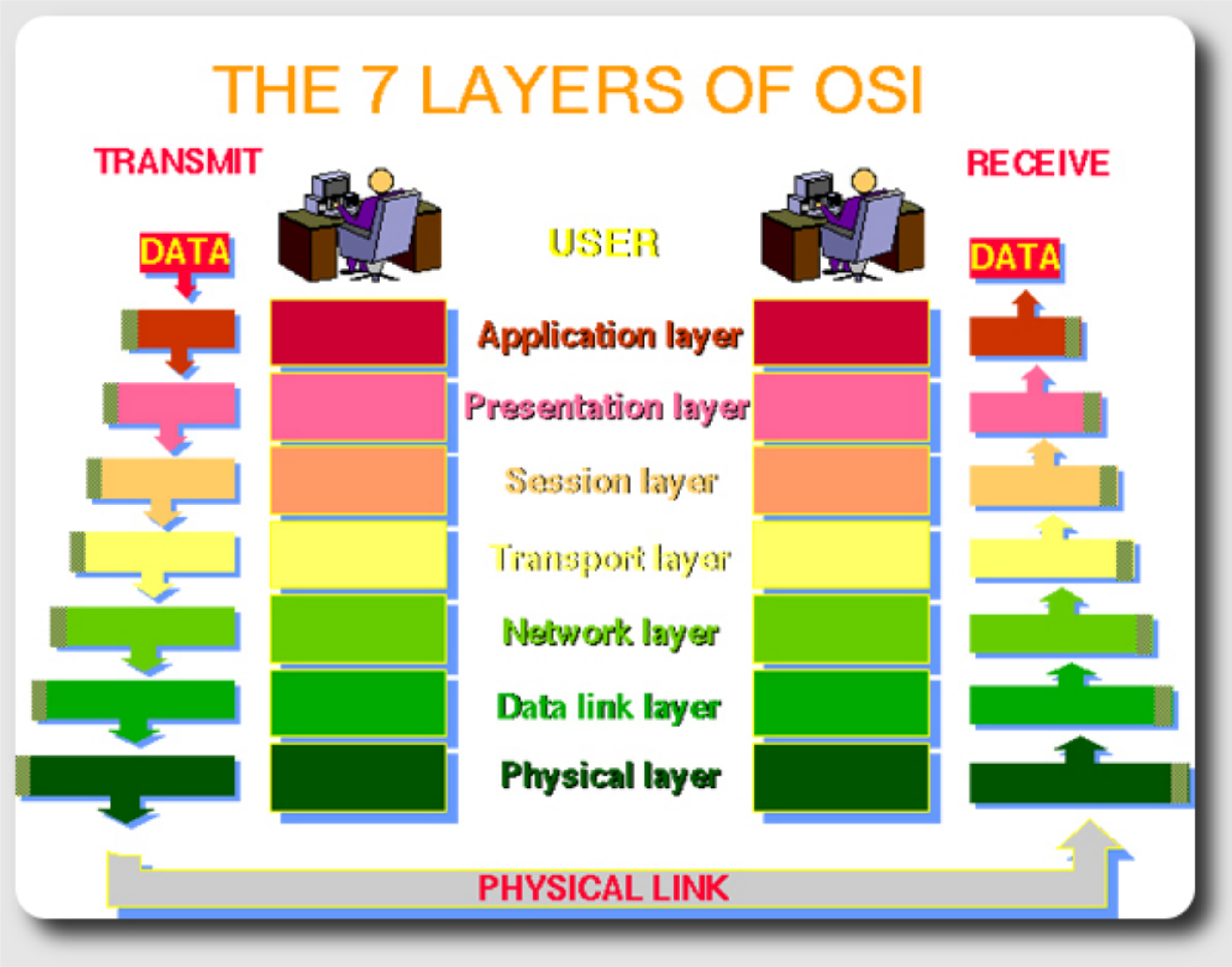
Complex problems can be solved using the common divide and conquer principle. In this case the internals of the Internet are divided into separate layers.

- 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
USER
DATA
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)

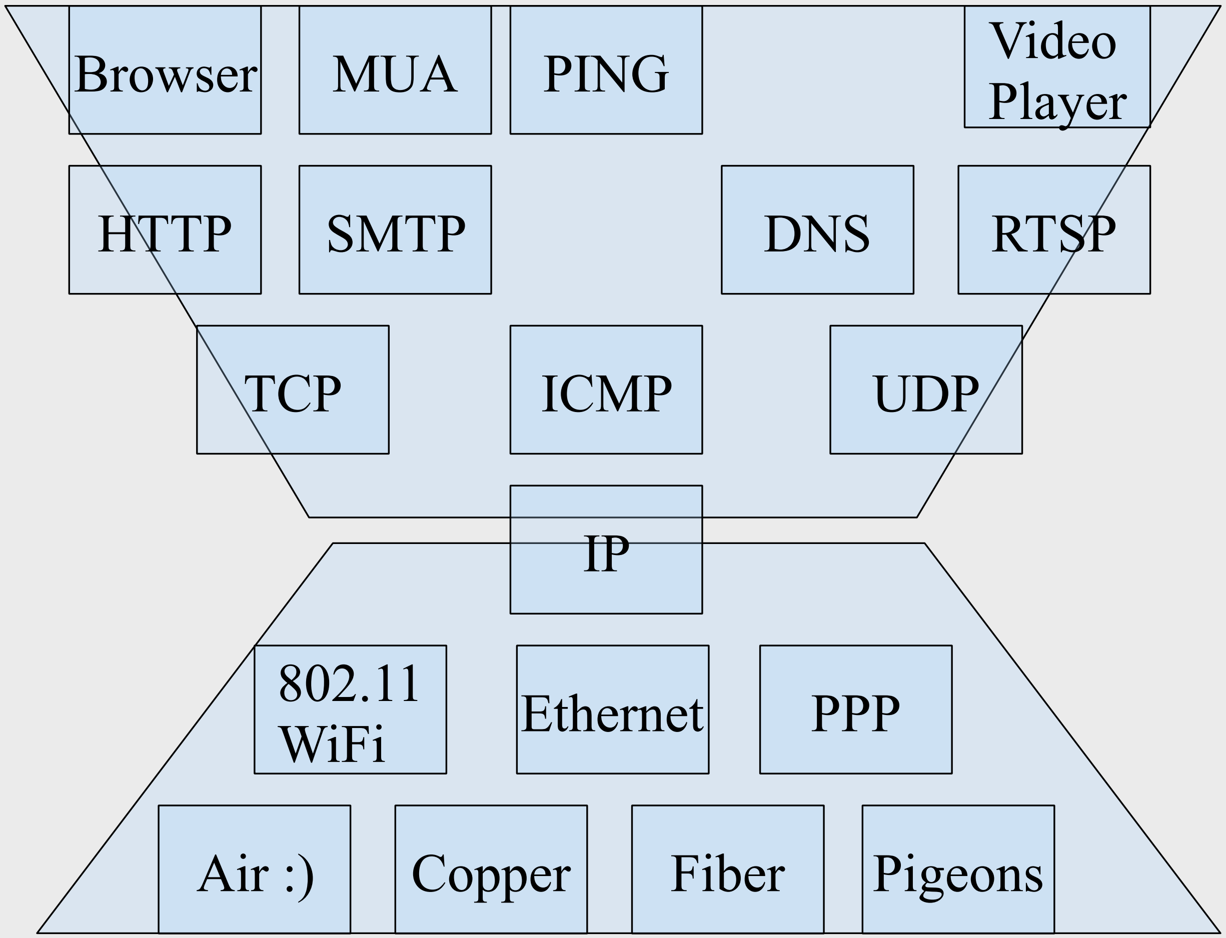
Acronym: All People Seem To Need Data Processing

TCP/IP

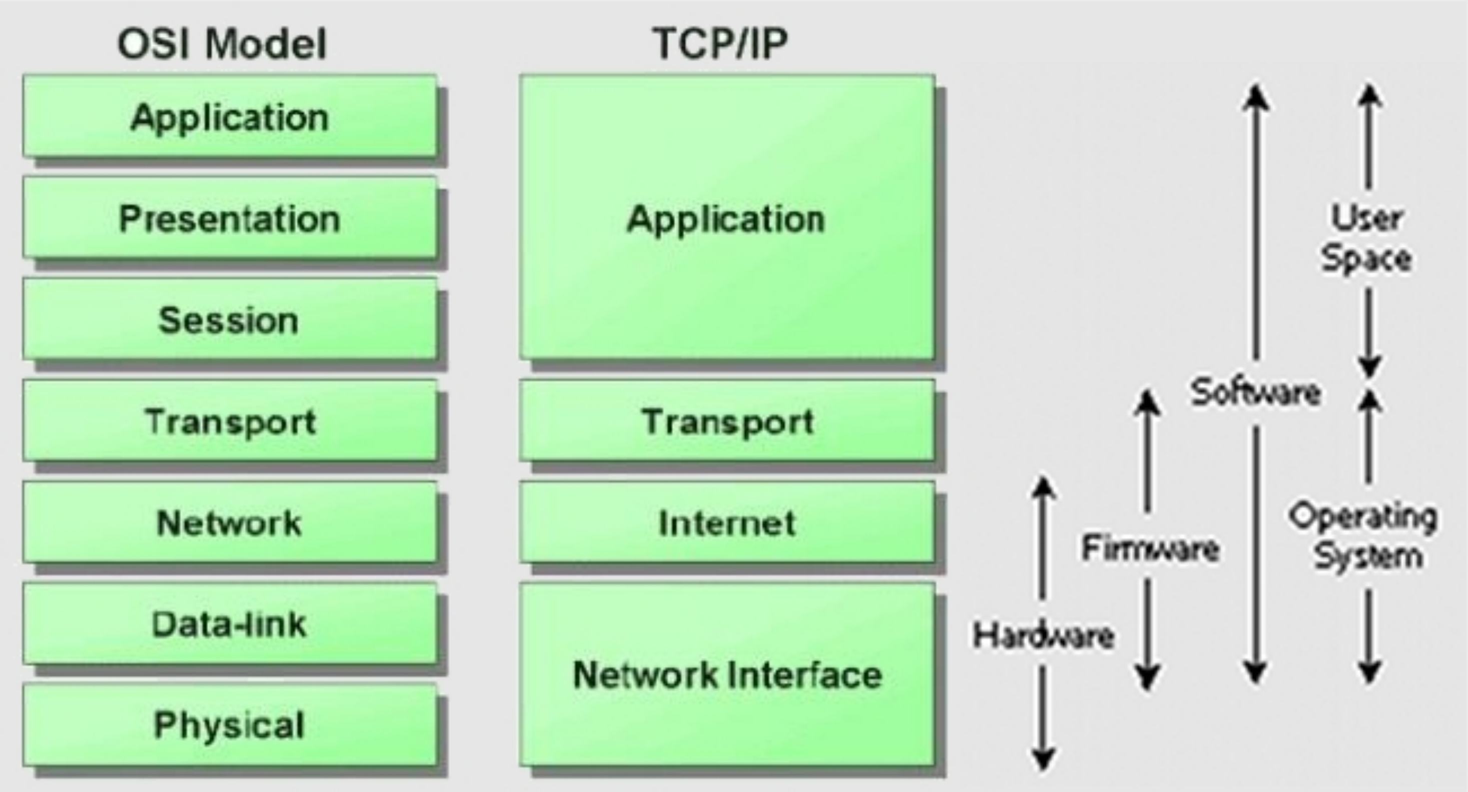
Generally, TCP/IP (Transmission Control Protocol/Internet Protocol) is described using three to five functional layers. We have chosen the common DoD reference model, which is also known as the Internet reference model.

- 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 the routing of data.
- Network access layer consists of routines for accessing physical networks.

TCP/IP model – the "hourglass"

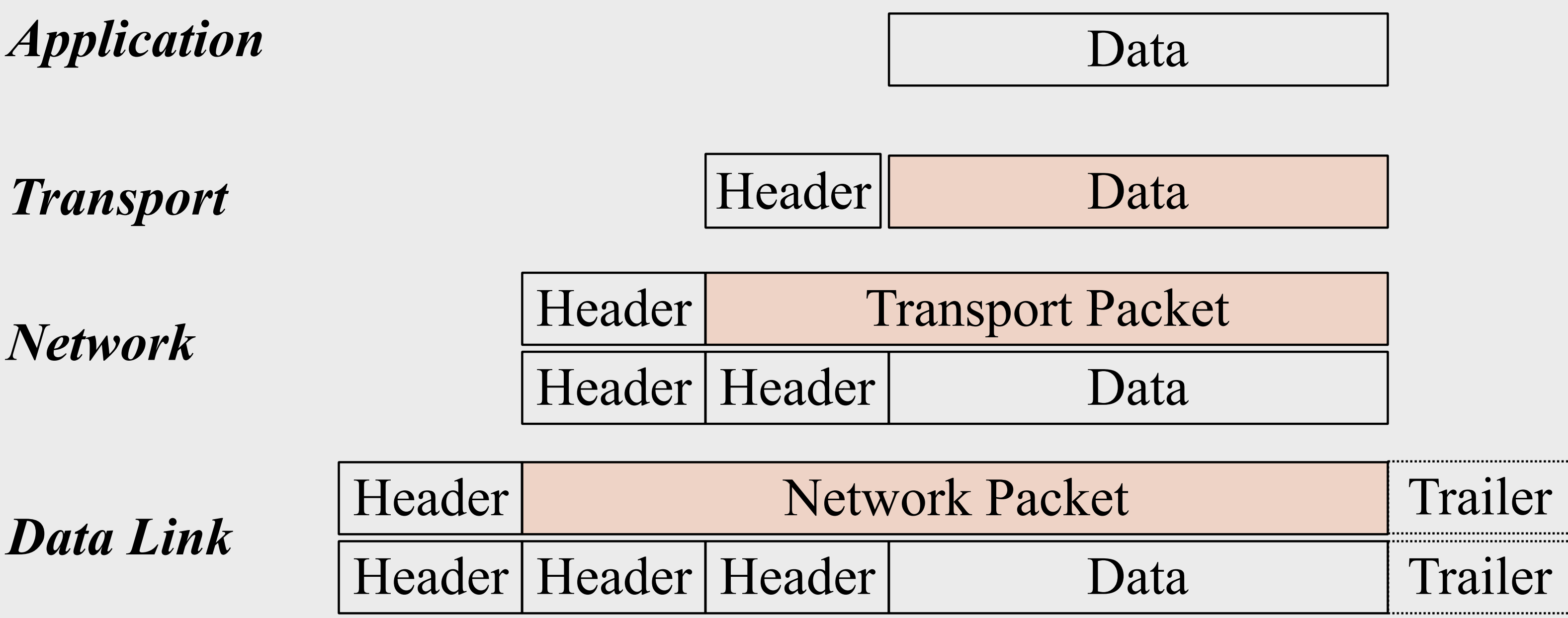


OSI and TCP/IP



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

10000101	00011011	10100010	01111101

85	1B	A2	7D

More to the structure

Hierarchical Division in IP Address:

Network Part (Prefix)

describes which network

Host Part (Host Address)

describes which host on that network

205 . 154 . 8		1
11001101 10011010 00001000		00000001
Network	Mask	Host

Boundary can be anywhere

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

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

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)

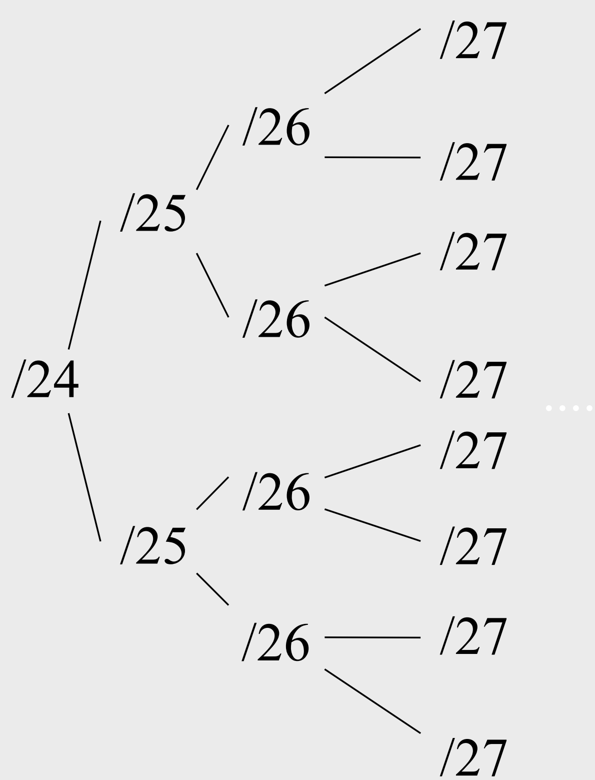
e.g. 134.132.100.255 (134.132.100.0/24)

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

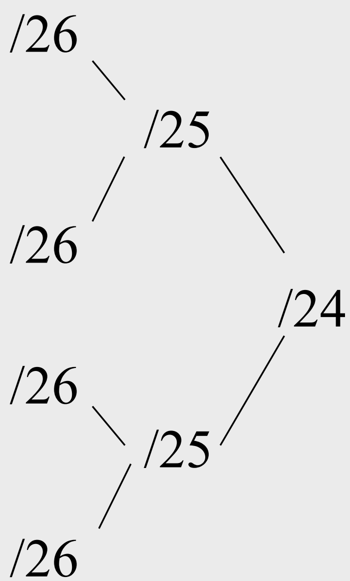


By adding one bit to the netmask, we subdivide the network into two smaller networks. This is *subnetting.*

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 AfriNIC
- 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...)

Network related settings

Files

```
/etc/rc.conf
/etc/netstart
/etc/hosts
/etc/resolv.conf
```

Commands

```
# ifconfig eth0 196.200.218.x/24

# route add default 192.200.218.254        (FreeBSD)
# route add default gw 192.168.218.254     (Linux)

# hostname pcN.ws.afnog.org
```

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.

FreeBSD routing table from "`netstat -anr`"

```
Routing tables

Internet:
Destination        Gateway            Flags    Refs      Use  Netif Expire
default            196.200.218.254    UGS         4     1068   bge0
127.0.0.1          link#3             UH          0       12    lo0
196.200.218.0/24   link#1             U           0        0   bge0
196.200.218.253    link#1             UHS         0        0    lo0

Internet6:
Destination                       Gateway                       Flags      Netif Expire
::1                               ::1                           UH          lo0
fe80::%lo0/64                     link#3                        U           lo0
fe80::1%lo0                       link#3                        UHS         lo0
ff01:3::/32                       fe80::1%lo0                   U           lo0
ff02::%lo0/32                     fe80::1%lo0                   U           lo0
```

What do route table entries mean?

```
Destination        Gateway            Flags     Refs      Use  Netif Expire
default            196.200.218.254    UGS          4     1068   bge0
127.0.0.1          link#3             UH           0       12    lo0
196.200.218.0/24   link#1             U            0        0   bge0
196.200.218.253    link#1             UHS          0        0    lo0
```

- 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).
- **Flags** indicate various attributes for each route:
  - **U Up**: The route is active.
  - **H Host**: The route destination is a single host.
  - **G Gateway**: Send anything for this destination on to this remote system, which will figure out from there where to send it.
  - **S Static**: This route was configured manually, not automatically generated by the system.
  - **C Clone**: Generates a new route based on this route for hosts we connect to. This type of route normally used for local networks.
  - **W WasCloned**: Indicated a route that was auto-configured based upon a local area network (Clone) route.
  - **L Link**: Route involves references to Ethernet hardware.
- **Refs** is the number of active references to this route.
- **Use** is the count of number of packets sent using this route interface
- The **Netif** is the network interface that is connected to that network
- **Expire** is the seconds the ARP entry is valid

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 **destination** 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 128.223.143.42 using this route table.

Destination	Gateway	Genmask	Flags	Interface
128.223.142.0	0.0.0.0	255.255.254.0	U	bge0
0.0.0.0	128.223.142.1	0.0.0.0	UG	bge0

- 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** bge0.
- That first entry is called a **network route**.

Do you notice anything different about this routing table?

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 bge0.

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 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 (for things like NAT):

- In FreeBSD in /etc/rc.conf set:

```
gateway_enable="YES”
```

- In Linux (temporary)*:

```
# echo “ip_forward=1” > /proc/sys/net/ipv4
```

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

*In sysctl.conf add “net.ipv4.ip_forward=1