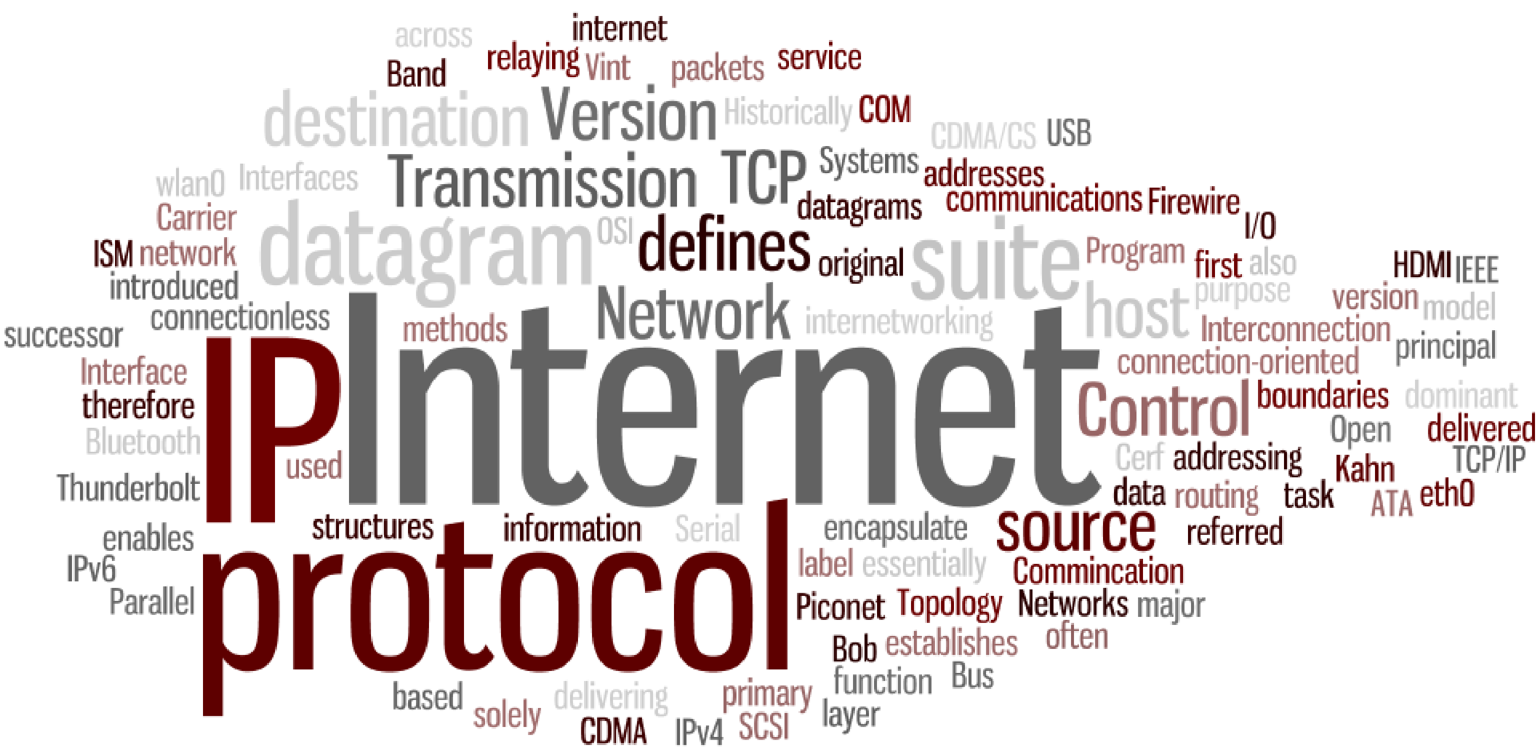
**Duale Hochschule Baden-Württemberg Mannheim**
**IMBIT, WIBI15A, 2016**

Sebastian Welter

# IP Address Scheme



# IP v4 vs. IP v6

- The current IP addressing scheme is called “IP v.4” and uses 32 bit (4 byte) per address.
  - Address space: $2^{32}$ - 4’294’967’296 addresses, approx. 3.7 billion usable
  - e.g. 9.155.23.123

- The next system is called “IP v.6” and will use 128 bit (16 byte) per address.
  - Address space: $2^{128}$ – 340 sextillion addresses
  - e.g. 2001:0db8:85a3:0042:1000:8a2e:0370:7334

  - IP v.6 handles many more addresses e.g. for every mobile phone, car, coffee machine.
  - IP v.6 offers enhancements in other areas

- Adoption of IPv6:
  - Implemented in all major operating systems and many applications
  - Backbones and major routers are becoming IPv6 compatible
  - Very less usage in the general public

# Addressing in Ethernet with TCP/IP

- Physical Address: MAC
  - Ethernet
  - IEEE 802 MAC, 48bit
  - 00:00:00:00:00:00

- Logical Address: IP
  - IPv4
    - 32 bit
    - 192.168.100.100
  - IPv6
    - 128 bit

- Network layer routing

# ARP: Address Resolution Protocol (RFC 826)

1. Sender is checking is Destination and Target are in the same network → subnet mask check
2. If it is outside of the network → deliver to Gateway
3. If it is in inside the network →
   - Check if we have the MAC in the ARP cache
   - → Enter the address into the DLL frame and send
4. If we don't have the address pair in the ARP cache:
   - Send an ARP message as broadcast to all nodes in the network
     - Node with the asked IP shall return its physical address back
   - Every node in the network is receiving the message
     - The right one sends an ARP-REPLY back
5. Both nodes append their ARP cache

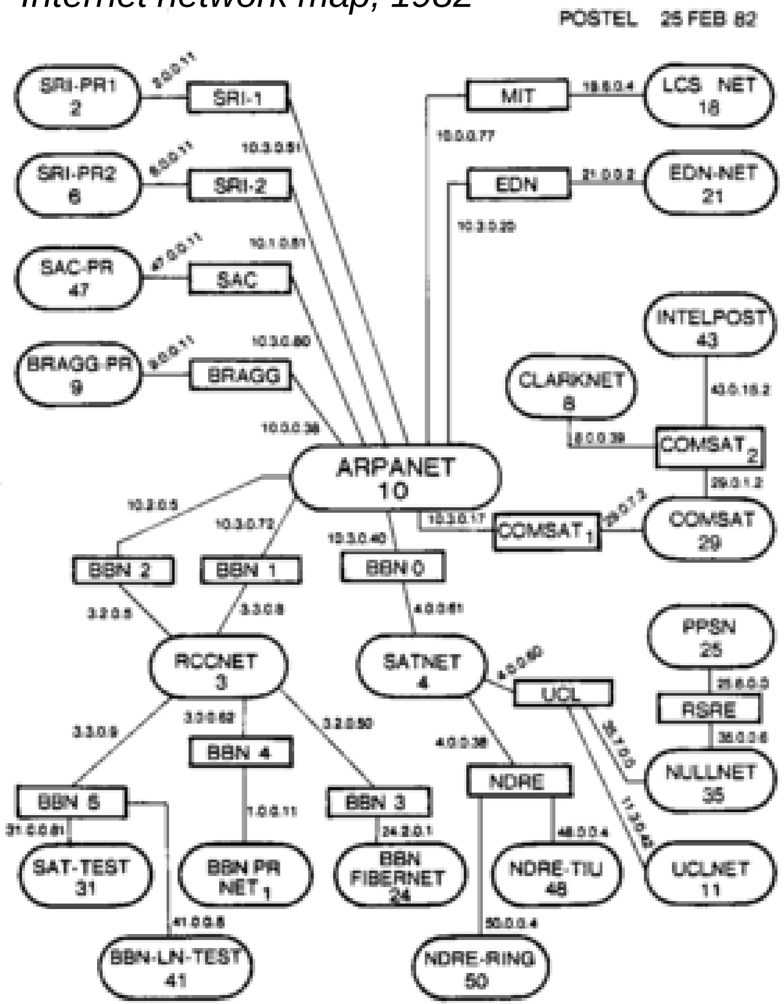
| Internet Protocol (IPv4) over Ethernet ARP packet | | |
|---|---|---|
| bit offset | 0 – 7 | 8 – 15 |
| 0 | Hardware type (HTYPE) | |
| 16 | Protocol type (PTYPE) | |
| 32 | Hardware address length (HLEN) | Protocol address length (PLEN) |
| 48 | Operation (OPER) | |
| 64 | Sender hardware address (SHA) (first 16 bits) | |
| 80 | (next 16 bits) | |
| 96 | (last 16 bits) | |
| 112 | Sender protocol address (SPA) (first 16 bits) | |
| 128 | (last 16 bits) | |
| 144 | Target hardware address (THA) (first 16 bits) | |
| 160 | (next 16 bits) | |
| 176 | (last 16 bits) | |
| 192 | Target protocol address (TPA) (first 16 bits) | |
| 208 | (last 16 bits) | |

# IPv4 subnetworks

- Not all addresses are in one big network
- Different sub-networks
  - Each one handled by its own boundary
  - Each router only knows its neighbors

- IP address was designed to consist of an
  - Network number field
    - Identification of a network
    - Subnet mask
  - Host identifier
    - Identification of a endpoint

*Internet network map, 1982*




- Subnet mask
  - Subnet mask: Definition of how many bytes define the network identifier
  - Instead of having the network number as fixed part of the IP address, specify it as a subnet mask

# IPv4 Classful network addressing scheme

**1981 till 1993**

| Class | Leading bits | Size of network number bit field | Size of rest bit field | Number of networks | Addresses per network | Start address | End address |
|---|---|---|---|---|---|---|---|
| **A** | 0 | 8 | 24 | 128 ($2^7$) | 16,777,216 ($2^{24}$) | 0.0.0.0 | 127.255.255.255 |
| **B** | 10 | 16 | 16 | 16,384 ($2^{14}$) | 65,536 ($2^{16}$) | 128.0.0.0 | 191.255.255.255 |
| **C** | 110 | 24 | 8 | 2,097,152 ($2^{21}$) | 256 ($2^8$) | 192.0.0.0 | 223.255.255.255 |
| **D** | 1110 | (multicast) | | | | 224.0.0.0 | 239.255.255.255 |
| **E** | 1111 | (reserved) | | | | 240.0.0.0 | 255.255.255.255 |

Src: Wikipedia: Classful network, Subnetwork

```
Class A IP Address:       0.0.0.0   = 00000000.00000000.00000000.00000000
Default Subnet Mask:    255.0.0.0   = 11111111.00000000.00000000.00000000
                                      nnnnnnnn.HHHHHHHH.HHHHHHHH.HHHHHHHH

Class C IP Address:       192.0.0.0 = 11000000.00000000.00000000.00000000
Default Subnet Mask: 255.255.255.0 = 11111111.11111111.11111111.00000000
                                      nnnnnnnn.nnnnnnnn.nnnnnnnn.HHHHHHHH
```

# IPv4 Address space registry

| Prefix | Designation |
|---|---|
| 000/8 | IANA - Local Identification |
| 001/8 | APNIC |
| 002/8 | RIPE NCC |
| 003/8 | General Electric Company |
| 004/8 | Level 3 Communications, Inc. |
| 005/8 | RIPE NCC |
| 006/8 | Army Information Systems Center |
| 007/8 | Administered by ARIN |
| 008/8 | Level 3 Communications, Inc. |
| 009/8 | IBM |
| 010/8 | IANA - Private Use |
| 011/8 | DoD Intel Information Systems |
| 012/8 | AT&T Bell Laboratories |
| 013/8 | Xerox Corporation |
| 014/8 | APNIC |
| 015/8 | Hewlett-Packard Company |
| 016/8 | Digital Equipment Corporation |
| 017/8 | Apple Computer Inc. |
| 018/8 | MIT |
| 019/8 | Ford Motor Company |
| 020/8 | Computer Sciences Corporation |
| 021/8 | DDN-RVN |
| 022/8 | Defense Information Systems Agency |
| 023/8 | ARIN |
| 024/8 | ARIN |
| 025/8 | UK Ministry of Defence |
| 026/8 | Defense Information Systems Agency |
| 027/8 | APNIC |
| 028/8 | DSI-North |

| Prefix | Designation |
|---|---|
| 029/8 | Defense Information Systems Agency |
| 030/8 | Defense Information Systems Agency |
| 031/8 | RIPE NCC |
| 032/8 | AT&T Global Network Services |
| 033/8 | DLA Systems Automation Center |
| 034/8 | Halliburton Company |
| 035/8 | Administered by ARIN |
| 036/8 | APNIC |
| 037/8 | RIPE NCC |
| 038/8 | PSINet, Inc. |
| 039/8 | APNIC |
| 040/8 | Administered by ARIN |
| 041/8 | AFRINIC |
| 042/8 | APNIC |
| 043/8 | Administered by APNIC |
| 044/8 | Amateur Radio Digital Communications |
| 045/8 | Administered by ARIN |
| 046/8 | RIPE NCC |
| 047/8 | Administered by ARIN |
| 048/8 | Prudential Securities Inc. |
| 049/8 | APNIC |
| 050/8 | ARIN |
| 051/8 | UK Gov. Dept. for Work and Pensions |
| 052/8 | E.I. duPont de Nemours and Co., Inc. |
| 053/8 | Cap Debis CCS |
| 054/8 | Administered by ARIN |
| 055/8 | DoD Network Information Center |
| 056/8 | US Postal Service |
| 057/8 | SITA |

058/8 till 126/8 & 128/8 till 223/8 are RIR-assigned

224/8 till 239/8 are multicast

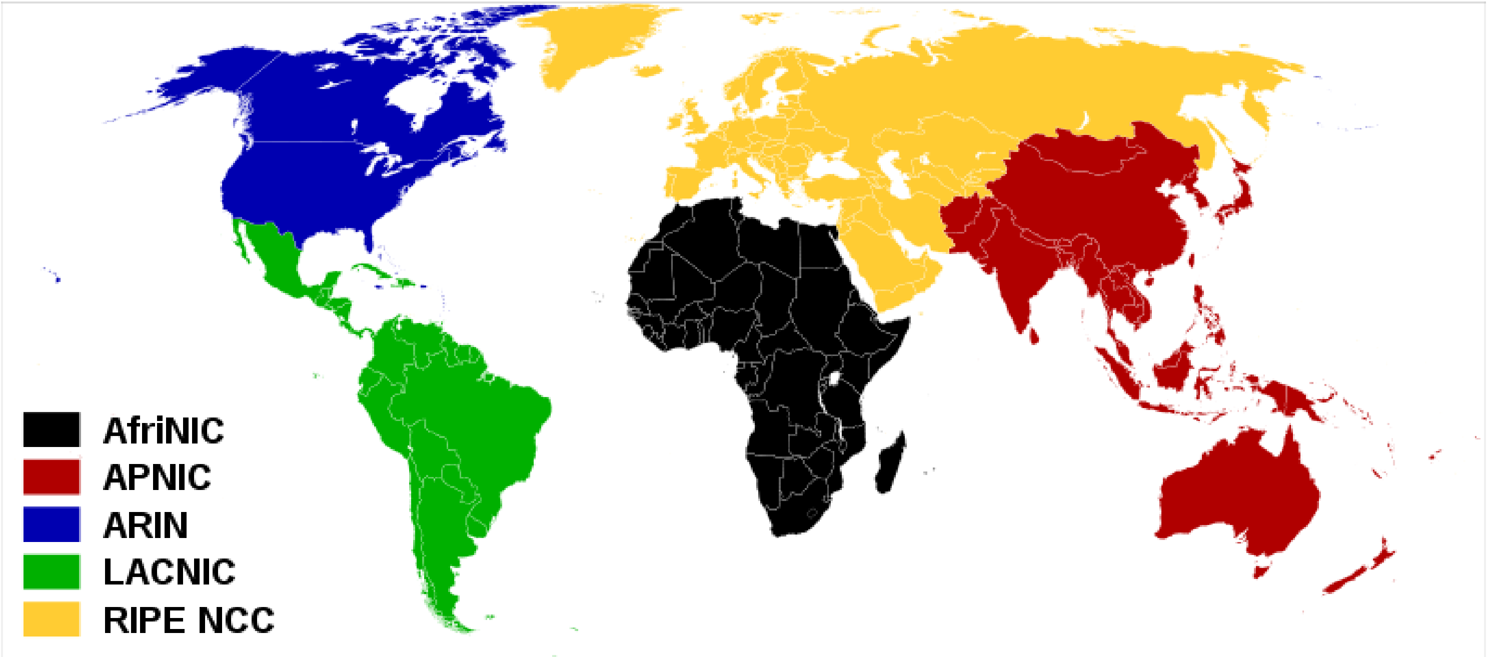
240/8 till 225/8 are reserved

127/8 is loopback

http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml (Last Updated: 2013-03-22)

# CIDR Block assignments

- IANA manages all subnets and their owners
  - Large blocks are issued to Regional Internet Registries (RIR)
- Each RIR is responsible for selling their address space to Internet Providers (ISP) or large organizations



# Classless Inter-Domain Routing (CIDR) … Recap

- **11000000.10101000.00000101.10000010** IP Address 192.168.005.130
- **11111111.11111111.11111111.00000000** Subnet mask 255.255.255.000

- **11000000.10101000.00000101.00000000** Network prefix 192.168.005.000
- **00000000.00000000.00000000.10000010** Host part 000.000.000.130

- **00001010.00001010.00000001.00011110 = 10.10.1.30**
- **11111111.11111111.11111111.00000000 = 255.255.255.0 = /24**
- **00001010.00001010.00000001. = 10.10.1.0**

- **00001010.00001010.00000001.00011110 = 10.10.1.30**
- **11111111.11111111.11111111.11100000 = 255.255.255.224 = /27**
- **00001010.00001010.00000001.000 = 10.10.1.0**

- **00001010.00001010.00000001.00100000 = 10.10.1.32**

# IPv4 address exhaustion I: Classless Inter-Domain Routing

- Classfull network addressing is obsolete
  - Internet grew faster and larger … too less addresses
  - To many large unused networks as companies got B and A classes
    - e.g. IBM owns the A-Class 9.0.0.0, GE owns 3.0.0.0, Ford owns 19.0.0.0

- New more dynamic subnetworking: **Classless Inter-Domain Routing (CIDR)**
  - In place since 1993
  - RFC 1518, 1519 and 4632
  - Variable length subnet masking
  - Much more, smaller, flexible subnets
  - Centrally managed by the IANA
    - Who owns which subnet (currently)

# Classless Inter-Domain Routing (CIDR)

- Variable length subnet masking
  - Requires a indication of how many bits the subnet mask uses for a defined network

- CIDR notation: <IP>/<Subnet>
  - <subnet> being the subnet prefix size in decimal notation (leading 1 bits in the address which compose the network number)

- 192.168.5.130/24

```
11000000.10101000.00000101.10000010   IP Address       192.168.005.130
11111111.11111111.11111111.00000000   Subnet mask      255.255.255.000

11000000.10101000.00000101.00000000   Network prefix   192.168.005.000
00000000.00000000.00000000.10000010   Host part        000.000.000.130
```

# Classless Inter-Domain Routing (CIDR)

Src: Wikipedia: Classless Inter-Domain Routing

| IP/CIDR | Δ to last IP addr | Subnet Mask | Hosts (incl 0/255) | Size |
|---|---|---|---|---|
| a.b.c.d/32 | +0.0.0.0 | 255.255.255.255 | 1 | 1/256 C |
| a.b.c.d/31 | +0.0.0.1 | 255.255.255.254 | 2 | 1/128 C |
| a.b.c.d/30 | +0.0.0.3 | 255.255.255.252 | 4 | 1/64 C |
| a.b.c.d/27 | +0.0.0.31 | 255.255.255.224 | 32 | ⅛ C |
| a.b.c.d/26 | +0.0.0.63 | 255.255.255.192 | 64 | ¼ C |
| a.b.c.d/25 | +0.0.0.127 | 255.255.255.128 | 128 | ½ C |
| a.b.c.0/24 | +0.0.0.255 | 255.255.255.000 | 256 | C |
| a.b.c.0/23 | +0.0.1.255 | 255.255.254.000 | 512 | 2 * C |
| a.b.c.0/22 | +0.0.3.255 | 255.255.252.000 | 1,024 | 4 * C |
| a.b.0.0/16 | +0.0.255.255 | 255.255.000.000 | 65,536 | B |
| a.0.0.0/8 | +0.255.255.255 | 255.000.000.000 | 16,777,216 | A |
| 0.0.0.0/0 | +255.255.255.255 | 000.000.000.000 | 4,294,967,296 | 256 * A |

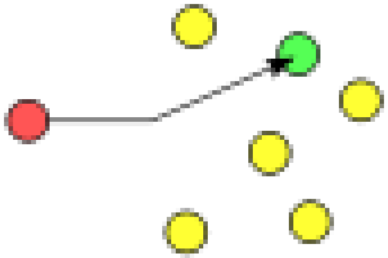
For all networks, except /0 and /32,/31, the first and the last address are reserved:

- The first address is used to identify the network itself
- The last address is the broadcast address

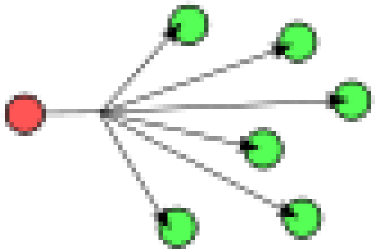
```
192.168.0.0/24
192.168.0.0   = network
192.168.0.255 = broadcast
```

# Unicast, Multicast, Broadcast

- Unicast



- Broadcast

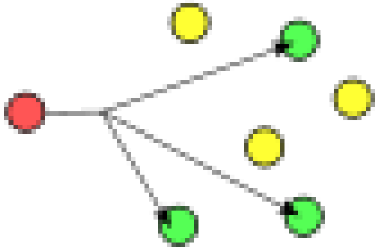


The largest address of a network

192.168.0.255 in 192.168.0.0/24

ff00::0/8

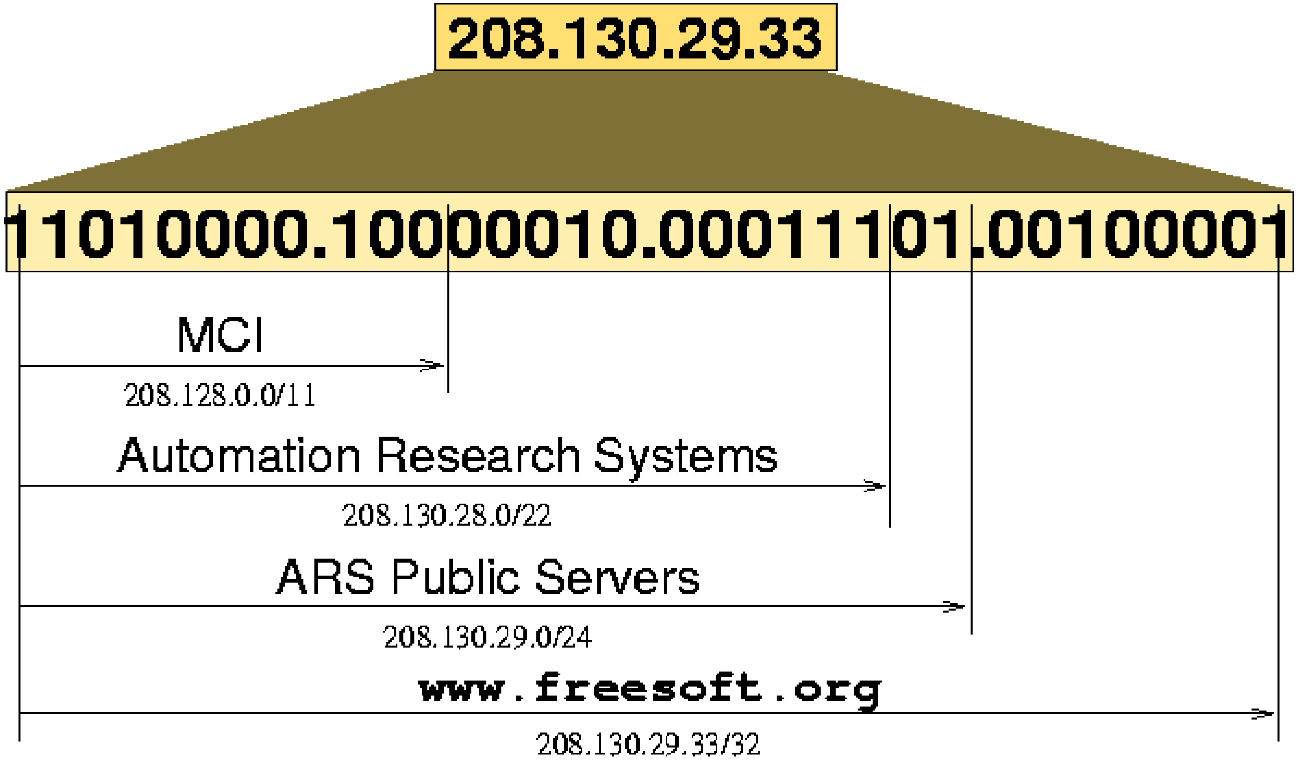
- Multicast



224/8 till 239/8

- Anycast

# CIDR Block assignments




http://en.wikipedia.org/wiki/File:CIDR_Address.png

# ARP: Address Resolution Protocol (RFC 826)

1. Sender is checking is Destination and Target are in the same network → subnet mask check
2. If it is outside of the network → deliver to Gateway
3. If it is in inside the network →
   - Check if we have the MAC in the ARP cache
   - → Enter the address into the DLL frame and send
4. If we don't have the address pair in the ARP cache:
   - Send an ARP message as broadcast to all nodes in the network
     - Node with the asked IP shall return its physical address back
   - Every node in the network is receiving the message
     - The right one sends an ARP-REPLY back
5. Both nodes append their ARP cache

| Internet Protocol (IPv4) over Ethernet ARP packet | | |
|---|---|---|
| bit offset | 0 – 7 | 8 – 15 |
| 0 | Hardware type (HTYPE) | |
| 16 | Protocol type (PTYPE) | |
| 32 | Hardware address length (HLEN) | Protocol address length (PLEN) |
| 48 | Operation (OPER) | |
| 64 | Sender hardware address (SHA) (first 16 bits) | |
| 80 | (next 16 bits) | |
| 96 | (last 16 bits) | |
| 112 | Sender protocol address (SPA) (first 16 bits) | |
| 128 | (last 16 bits) | |
| 144 | Target hardware address (THA) (first 16 bits) | |
| 160 | (next 16 bits) | |
| 176 | (last 16 bits) | |
| 192 | Target protocol address (TPA) (first 16 bits) | |
| 208 | (last 16 bits) | |

# Ethernet Frame

- Ethernet Frame

| PRE | SOF | DA | SA | L/T | Data | CRC |
|---|---|---|---|---|---|---|
| 7b | 1b | 6b | 6b | 4b | 46-1500b | 4b |
| Preamble | Start of Frame | Destination Address | Source Address | Data length and type | Data | |
| 101010… | 10101011 | MAC | MAC | *Ethertype* VLAN Tag | (Data) | 32 bit CRC Checksum |
| Synchronization | fix | | | Prio, Signal, 12 bit VLAN ID | (Padding) | |

Ethertype: 0800 IPv4, 0806 ARP, 0842 WOL, 809B AppleTalk, 86DD IPv6, 8870 Jumbo Frames

# IPv4 header

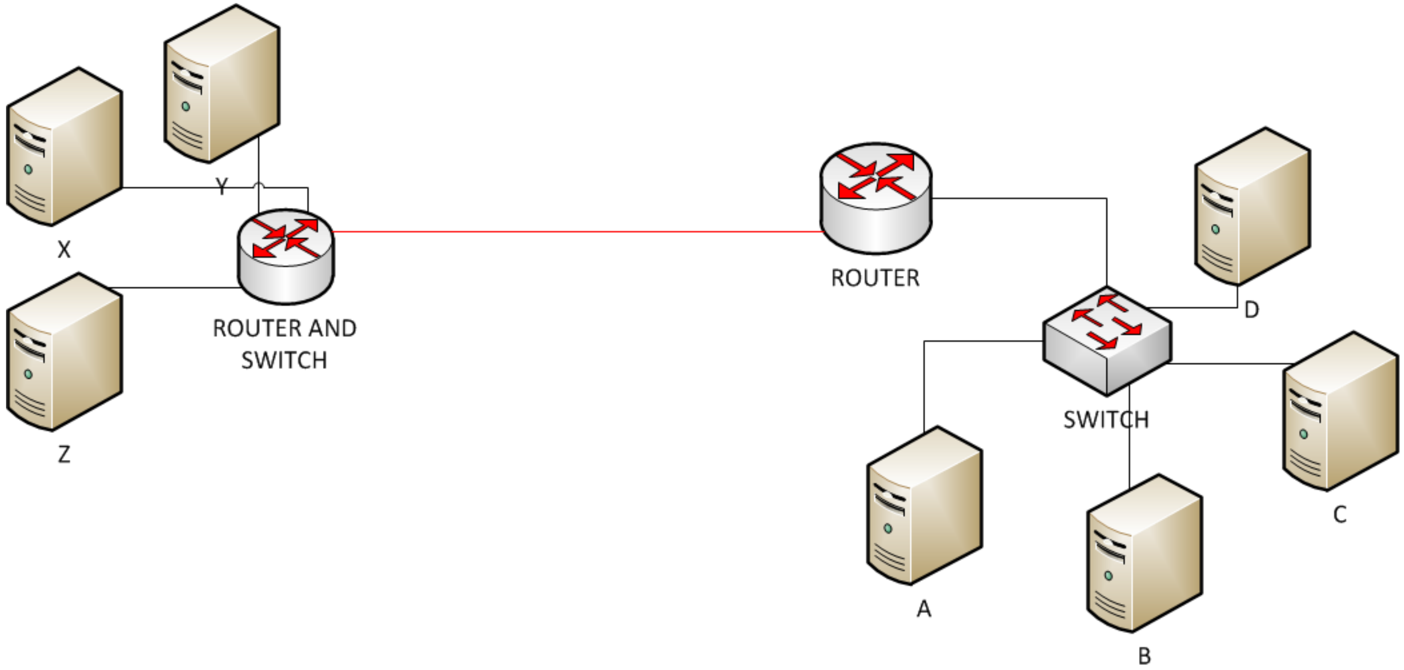
<table>
<tr><th>0</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th><th>21</th><th>22</th><th>23</th><th>24</th><th>25</th><th>26</th><th>27</th><th>28</th><th>29</th><th>30</th><th>31</th></tr>
<tr><td colspan="4">Version</td><td colspan="4">IHL</td><td colspan="6">DSCP</td><td colspan="2">ECN</td><td colspan="16">Total Length</td></tr>
<tr><td colspan="16">Identification</td><td colspan="3">Flags</td><td colspan="13">Fragment Offset</td></tr>
<tr><td colspan="8">Time To Live</td><td colspan="8">Protocol</td><td colspan="16">Header Checksum</td></tr>
<tr><td colspan="32">Source IP Address</td></tr>
<tr><td colspan="32">Destination IP Address</td></tr>
<tr><td colspan="32">Options (if IHL > 5)</td></tr>
</table>

- IHL = Internet Header Length in 32bit words
- DSCP = Differentiated Services Code Point, partially used in Voice-over-IP
- ECN = Explicit Congestion Notification, optional
- Length of the total package
- Identification = multiple fragment identification
- Flags: 0|DF|MF  -- Don't Fragment |  More Fragments
- Fragment Offset, in 8-byte units
- TTL, Header checksum – changed by every router
- Protocol indication of the transported data (TC, UDP, ICMP, …)
- Source and target addresses

# Routing – the base concepts and components

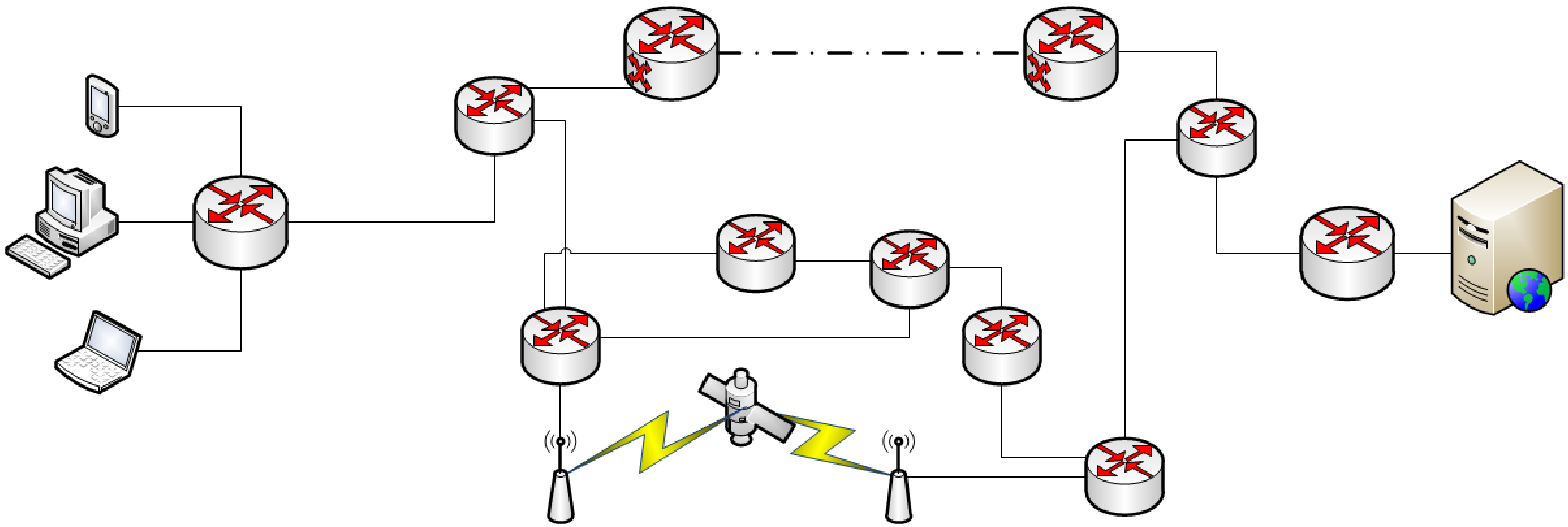
- Gateway ("Standard Gateway") in the common wording of operating systems for IP network interfaces is a ROUTER

- Routers connect different networks and relay packages between
  - Each other
  - The router and its connected endpoint entities (your PC)



# Routing

- Routing: How Do I move a package from source A to destination Z
  - A single router in a path does not know the entire path
  - Each route has a assigned **metric**:
    How much does the route to an address cost through this way



# Routing - Autonomous systems

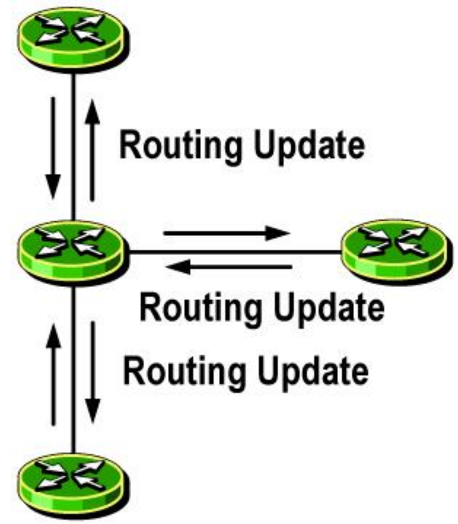
- Autonomous systems (AS)
  - Registered with IANA - *Autonomous System Number*, ASN
    - Currently there are about 37000 assigned (src: Wikipedia)
  - AS are owned by an entity
    - Carrier-Class routers and large enterprise routers
  - One AS: One or many subnets, connected via internal routing protocols

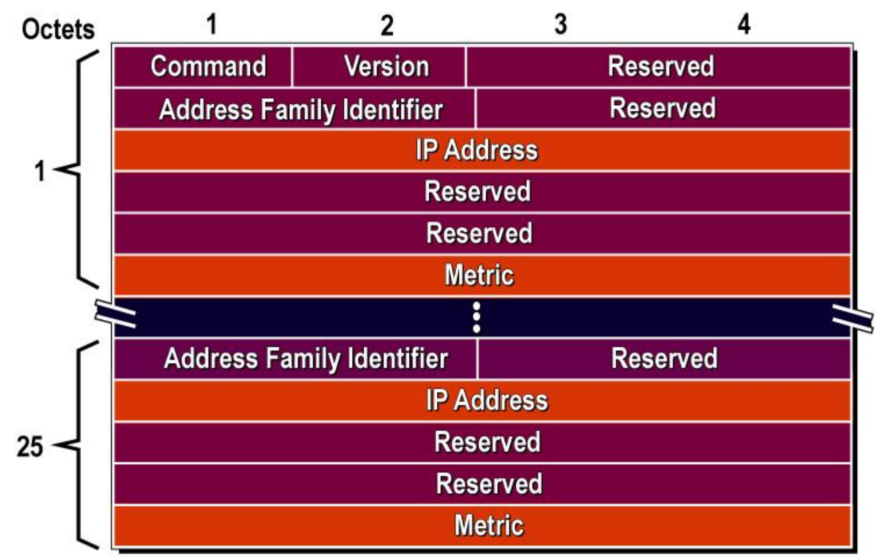
- Communication between AS:
  - Exterior gateway protocol
    - Common: Border-Gateway-Protocol

- Communication within AS:
  - Interior gateway protocol
  - Open Shortest Path First (OSPF)
  - Intermediate System to Intermediate System (IS-IS)
  - Simple: Router Information Protocol (RIP)

# Routing – simple example within an autonomous system: RIP

- RIP – Router Information Protocol
  - Interior gateway protocol
- Metric is the hop-count
  - 16 = infinity
- A router communicates with all of direct neighbors
  - Give me your routing table
  - I give you my routing table
- Newer:
  - RIP v2 with next-Hop and CIDR
    → Subnet support




Images src: http://www.hill2dot0.com/wiki/index.php?title=Routing_Information_Protocol

# Reserved addresses in the IPv4 space

| Range | Description | Reference |
|---|---|---|
| `0.0.0.0/8` | Current network (only valid as source address) | RFC 5735 |
| `10.0.0.0/8` | Private network | RFC 1918 |
| `100.64.0.0/10` | Shared Address Space (ISP and subscribers) | RFC 6598 |
| `127.0.0.0/8` | Loopback | RFC 5735 |
| `169.254.0.0/16` | Link-local | RFC 3927 |
| `172.16.0.0/12` | Private network | RFC 1918 |
| `192.0.0.0/24` | IETF Protocol Assignments | RFC 5735 |
| `192.0.2.0/24` | TEST-NET-1, documentation and examples | RFC 5735 |
| `192.88.99.0/24` | IPv6 to IPv4 relay | RFC 3068 |
| `192.168.0.0/16` | Private network | RFC 1918 |
| `198.18.0.0/15` | Network tests | RFC 2544 |
| `198.51.100.0/24` | TEST-NET-2, documentation and examples | RFC 5737 |
| `203.0.113.0/24` | TEST-NET-3, documentation and examples | RFC 5737 |
| `224.0.0.0/4` | IP multicast (former Class D network) | RFC 5771 |
| `240.0.0.0/4` | Reserved (former Class E network) | RFC 1700 |
| `255.255.255.255/32` | Broadcast | RFC 919 |

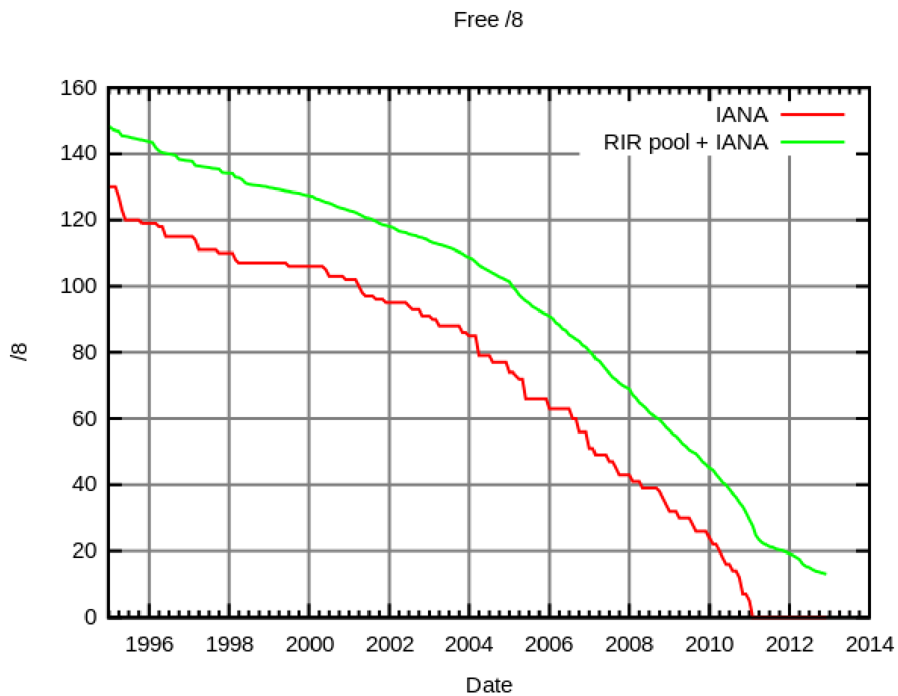
Taken from wikipedia

# How to Get IP Addresses?

- IP addresses for public use are usually assigned by the Internet Service Provider – for single machines or whole networks.
  - In your own network without visibility to the Internet you may choose addresses from special reserved ranges, e.g. 192.168.x.x.
  - A NAT node (network address translation) can translates between internal private addresses and assigned public ones.

# IPv4 address exhaustion II

- We don't have enough IPv4 Addresses
- Even more devices are on the Internet
  - Mobile devices
  - Always on connections
- Still inefficient legacy networks




*Exhaustion of IPv4 addresses since 1995 (wikipedia)*

➢ Reclaiming old, unused legacy block
  - Stanford University released their Class A block in 2000

➢ Name-based virtual hosting of websites
  - One IP address serves multiple domains

➢ Network address translation
  - Transparent local subnetworking with local non-routed addresses

# NAT – Network Address Translation to mitigate the address exhaustion

- Network address translation as one-to-many translation
  - Encapsulate one network in one IP
  - Private networks which are exposed by one IP
    - Homes and offices connected to the Internet using a router
    - Router = IP from ISP
    - Internal devices = IP within a local private network

- Enables millions of users and devices to be on the Internet without consuming an IP

- The device connecting external and internal network needs to manage the NAT:
  - Masquerade packages between external IP and internal IPs
  - Masquerading tables
  - Problem of incoming connections
  - Breaks the end-to-end-connection notion