University of Amsterdam

InterNetworking and Routing

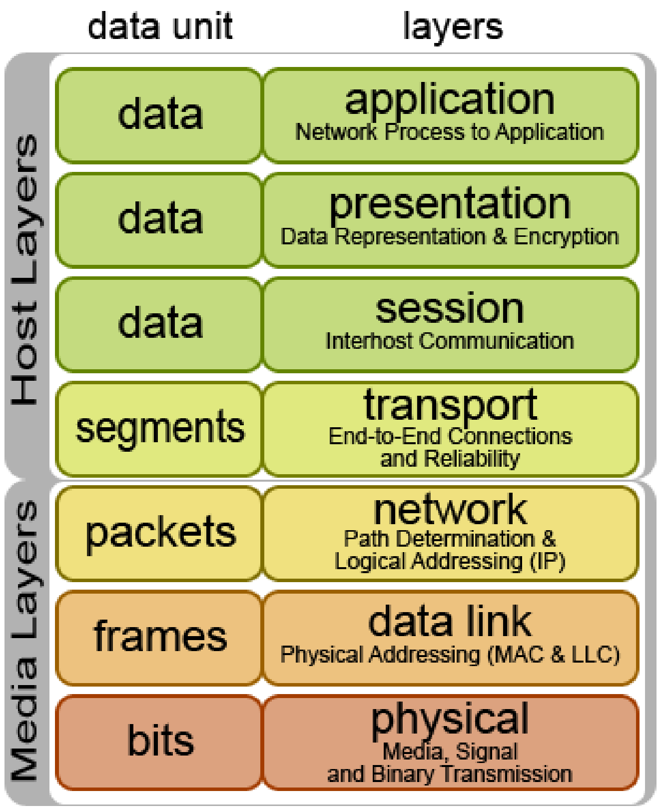
Xavier Torrent Gorjón
Xavier.TorrentGorjon@os3.nl

March 17, 2015

Contents

1 Overview

1.1 OSI Model



1.2 Interfaces and Protocols

Interfaces Interfaces connect different layers on the same computer. Uses Protocol Data Units (PDU).

Protocols Protocols are used to communicate between parties data on a specific layer. Uses Service Data Units (SDU) inside Service Access Points (SAP).

1.3 Encapsulation and Multiplexing

Encapsulation When data units go down one level in the layer model, headers are added to add information regarding the current layer.

Multiplexing Multiple protocols can coexist on the same layer. However, when going down the layer model, these protocols should be treated equally. For example, TCP and UDP are multiplexed down at the IP level, and demultiplexed back when reading the information of IP packets.[1]

1.4 ES Models: Strong vs Weak

Strong ES Model Hosts suppress packets with a destination address that references another of its interfaces.

Weak ES Model Hosts accept packets that match with one of its interfaces addresses, even if it does not receive it on that interface.

1.5 IP Addressing (IPv4)

- 32-bit addresses
- Decimal-dotted notation (a.b.c.d, 0 =< a,b,c,d =< 255).
- Special addresses:

  0.0.0.0 IP address unknown.

  127.0.0.1 Loopback address.

  Host part all 0 Subnet identifier.

  Host part all 1 Directed broadcast.

  255.255.255.255 Local subnet broadcast.

- Private addresses:

  10.0.0.0/8

  172.16.0.0/12

  192.168.0.0/16

  169.254.0.0/16

1.6 Subnetting

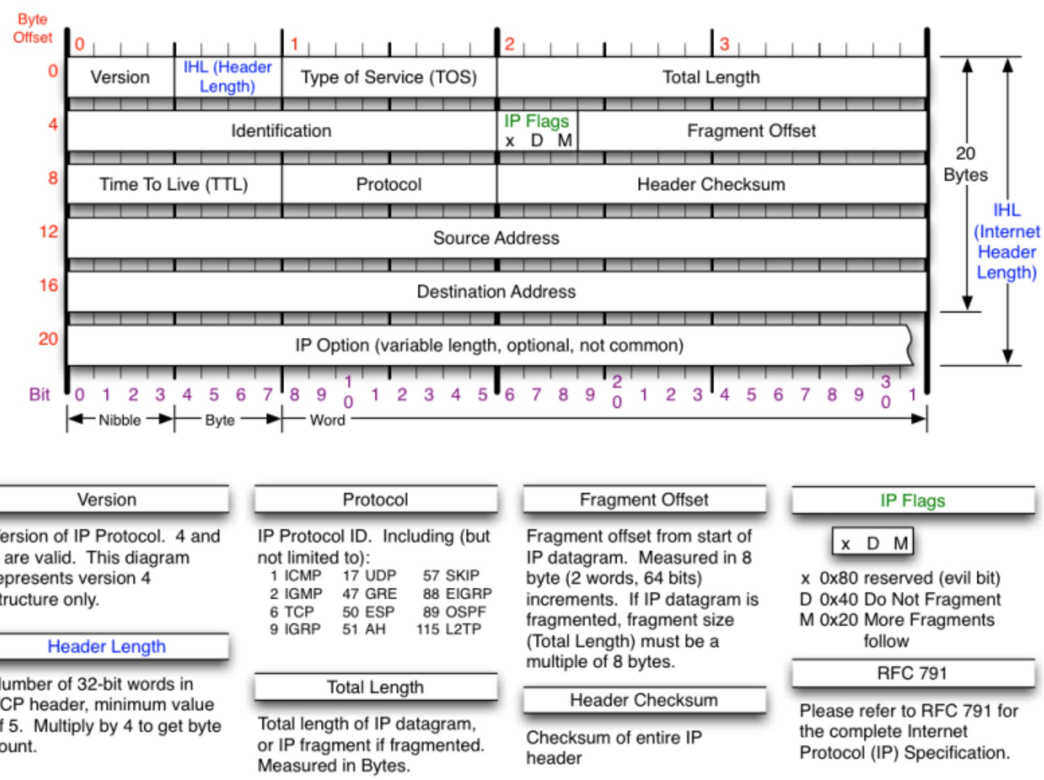
- Originally classful subnetting (subnets in A/B/C ranges, with 24, 16 and 8 bits of network addresses respectively; D range for multicast and an unused E range).[2]
- Classless Inter-Domain Routing (CIDR), with network masks to mark the difference between network address and host address. Routing done by selecting most specific match.

[1] http://www.tcpipguide.com/free/t_TCPIPProcessesMultiplexingandClientServerApplicati-2.htm

[2] http://en.wikipedia.org/wiki/Classful_network#Introduction_of_address_classes

- Variable Length Subnet Masks (VLSM) to use different subnets that do not have the requirement of having the same size. Add the possibility of subnets inside subnets. This was not possible in RIPv1.
- A "link" is defined as the topological area in which a packet with $TTL = 1$ can be delivered (aka. not being forwarded).
- A "subnet" is the topological area in which the interfaces receive the same network prefix.

1.7 IP Packet Format



2 CLEAN: Calculating, Legacy, Endianness, Addressing, Networks

2.1 Calculating: Counting

Counting Process that starts with $n = 0$ as the initial count. Every counted object is labeled with the actual n value, and n is updated to $n = n + 1$. Process ends when all objects have been counted.

2.2 Legacy

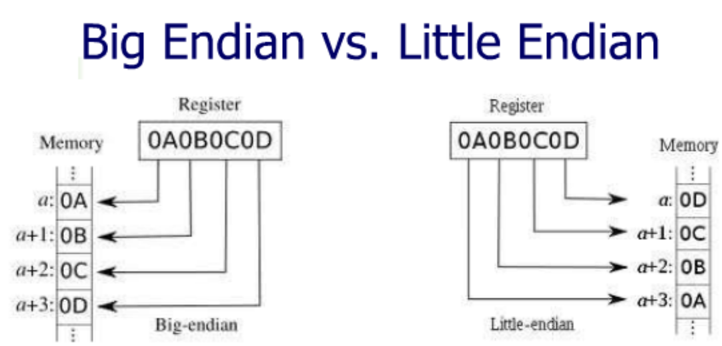
- Everybody knows what Karst thinks of legacy.

2.3 2-adic vs Binary

2-adic	Binary	2-adic to base-10	Binary to base-10
1	0	1	0
2	1	2	1
11	00	3	0
12	01	4	1
21	10	5	2
22	11	6	3
111	000	7	0

- The whole point is that binary resets at every range increase.

2.4 Big-endian and Little-endian



2.5 Addressing

- `http://www.exploringbinary.com/binary-converter/`

3 IPv6

3.1 Rationale

- $4x$ address space size increase = 2^{96} address number increase.
- Headers have a fixed size of 40 bytes. Supports extended headers for additional functionality.
- NATs no longer needed due the vast amount of addresses.

3.2 Addressing

- 128-bit addresses.
- 8 blocks of 4 nibbles ($8x4x4 = 128$ bits)
- Consecutive blocks of all-zeroes can be replaced by :: once.
- No broadcasts, no subnet masks.
- `http://www.iana.nl/assignments/ipv6-address-space/ipv6-address-space.xhtml`

Reserved addresses

::/8	Special-purpose
100::/8	Special-purpose
2000::/3	Global unicast
fc00::/7	Unique local unicast
fe80::/10	Link-local (Link-scoped) unicast
ff00::/8	Multicast

Special-purpose addresses

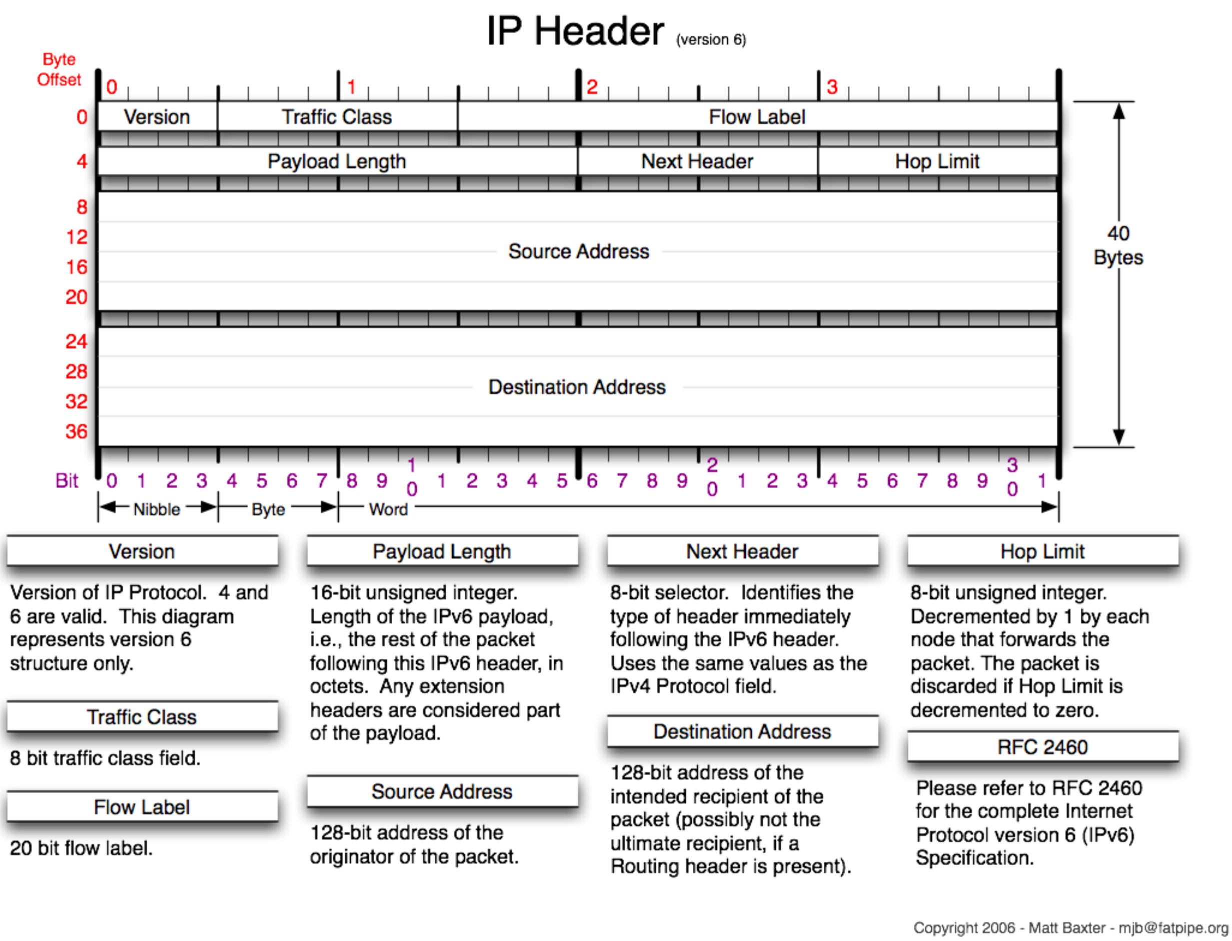
::/128	Unspecified address
::1/128	Localhost address
::a.b.c.d/128(from ::/96)	IPv4-compatible addresses
::ffff:a.b.c.d/128 (from ::ffff:0:0/96)	IPv4-mapped addresses
64:ff9b::/96	Well-known prefix
100::/64	Discard-only address block

3.3 Neighbour Discovery Protocol (NDP)

- IPv6 does not use ARP. Uses ICMPv6 instead.
- ICMPv6 types for NDP:

133	Router Solicitation
134	Router Advertisement
135	Neighbor Solicitation
136	Neighbor Advertisement
137	Redirect Message

3.4 IPv6 Header



4 Layer 2: Bridging and Switching

4.1 Layers 1 and 2

Layer 1 Repeaters, hubs. Same collision domain. Same link segment.

Layer 2 Bridges, switches. Same collision domain. Same link segment.

4.2 Layer 2: MAC and LLC

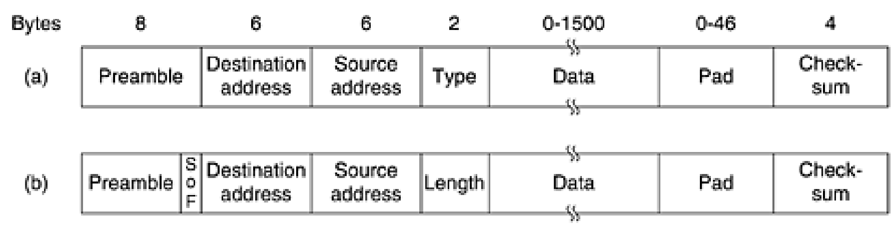
MAC Media Access Control. Work from IEEE 802.3[3].

CSMA/CD Carrier Sense Multiple Access With Collision Detection. Ethernet is the most common.

1. Is my frame ready for transmission? If yes, move to 2.
2. Is medium idle? If not, wait until it becomes ready.
3. Start transmitting.
4. Did a collision occur? If so, go to collision detected procedure.
5. Reset retransmission counters and end frame transmission.

LLC Logical Link Control. Multiplexing mechanisms, flow control and error management. Interface between MAC and Network Layer.

4.3 Frame Formats



DIX Ethernet DEC-Intel-Xerox initial frame structure[4].

8bytes Preamble Each one containing 10101010. Used for synchronization.

6bytes Destination Address

6bytes Source Address

2bytes Type Indicator of the used transport protocol.

0-1500bytes Payload

[3] `http://en.wikipedia.org/wiki/IEEE_802.3`
[4] `http://www.epubbud.com/read.php?g=5HEKFDZU&two=1&tocp=38`

0-46bytes Pad Valid Ethernet frames have, at least, 64 bytes (not counting Preamble!). If a frame is less than that (Payload ¡ 46bytes), a pad is added to it. This is done to ease collision detection.

4bytes Checksum CRC of all the frame fields.

802.3 Ethernet Changes from DIX:

7bytes Preamble + 1byte Start of Frame Same as before but changing last byte to have compatibility with 802.4 and 802.5 (Token Bus and Token Ring).

2bytes type -¿ 2bytes length IEEE tried to change the purpose of the field (and move ”type” information to *inside* the Payload), but some people did not change. Rule of thumb: if its value is over 1500, it is a Length field, otherwise it is a Type field.

Changes on the Payload There are 8 additional bytes of ’metadata’ on the payload that are used mostly for nothing, effectively reducing the MTU from 1500 to 1492.

4.4 MAC Addresses

MAC48 Physical, obsolete.

EUI48 Virtual, including physical.

EUI64 Extended.

OUI First 24bits of the MAC address. Identify who issued the device.

4.5 EUI48 to EUI64

[0:23bits]:FF:FE:[24:–bits]. When used to generate IPv6 address, the 7th most significative bit is swapped.

4.6 Ethernet Types

0x0800 IP

0x0806 ARP

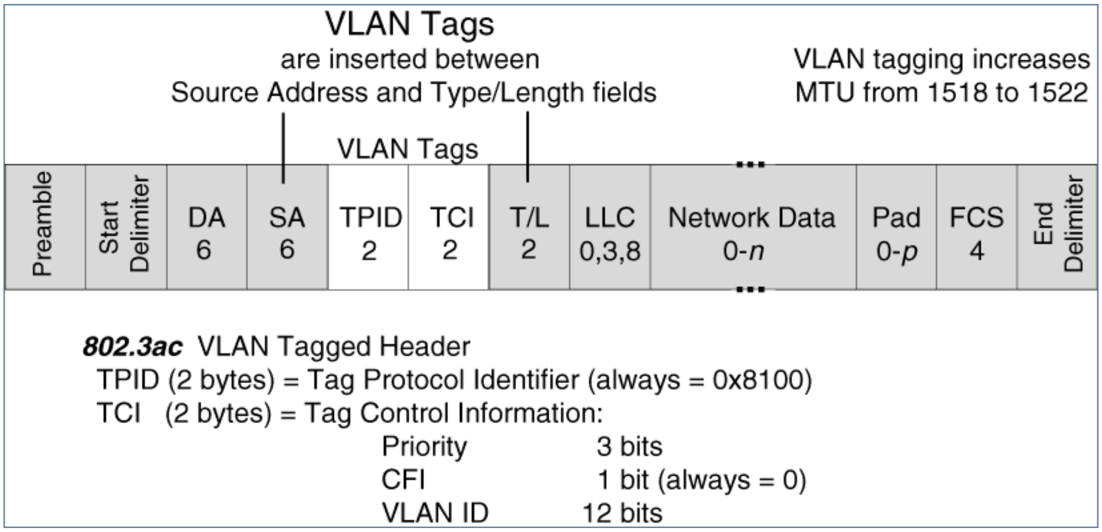
0x86DD IPv6

4.7 Bridges and Switches

Transparent Bridges Use Store-and-Forward. Copy data from one port to another (or multiple ports). They can learn (and remember) where other devices are when they receive messages.

Switches are synonyms of Bridges Usually refer to bridges with multiple interfaces.

4.8 VLANs (802.1Q-2011)



4.9 Layered Extensions

TO-DO

4.10 PBB-TE, TRILL, SPB

TO-DO

5 STP Protocol

5.1 Goals and properties

1. Eliminate edges (connections) until there are no possible loops.
2. After performing the algorithm, graph turns into a tree.
3. Topology changes cause changes on the tree.
4. Protocol works by electing a Root Node.

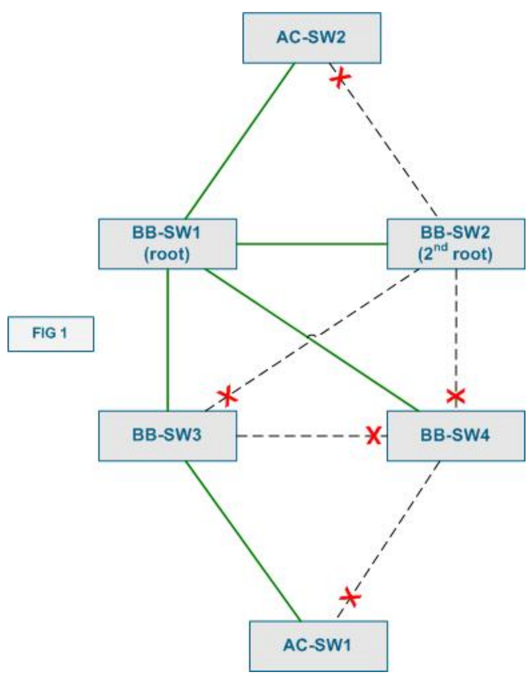
5.2 Configuration Messages

ID based on a variable priority and its MAC address.

Root Node Elected as the node with the lowest ID on the system.

Root ports on each Node Chosen by[5]:

1. Lower advertised Root ID.
2. Lower advertised cost to Root.
3. Lower transmitting bridge ID.
4. Lower port ID.



[5] https://www.youtube.com/watch?v=iB7BxtZVy3c

5.3 Timing Parameters

Hello Time Time between two configuration messages.

Max Age Parameter to discard messages that are too old.

Forward Delay Half of the delay before transitioning from blocking to forwarding.

1. Can be understood as two different waiting times. During that waiting, it does not forward packets.
2. First waiting: Listen for neighbours (other bridges).
3. Second waiting: Learn the location of MAC addresses.

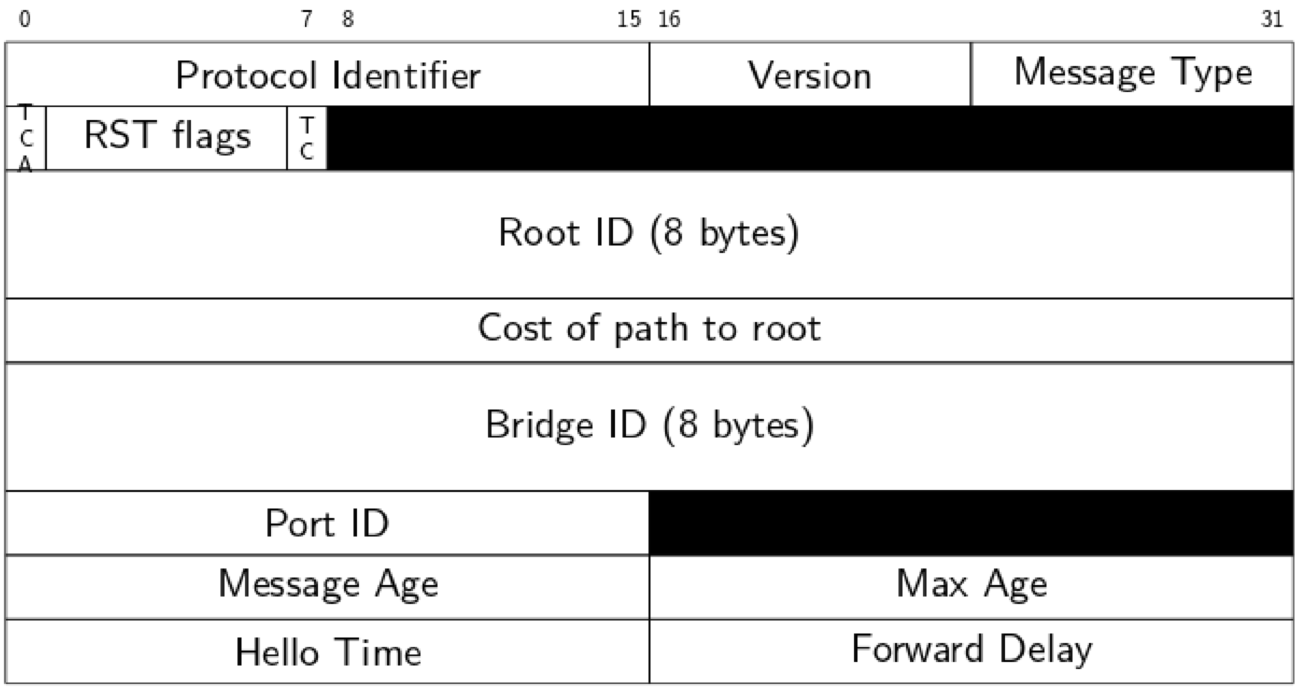
5.4 Topology Change Mechanism

Memory Bridges remember where other bridges are located.

Stable Topology When the topology is stable, bridges have a long caching time.

Topology Changes When a topology change is detected, the bridge detecting the change (and subsequent ones) sends a Topology Change Notification on his Root Port. When the message reaches the Root Bridge, it sets up its Topology Change flag. This causes other bridges to switch to the short caching delay.

5.5 Bridge Protocol Data Unit (BPDU)



Protocol Identifier

Version

Message Type

Flags

Root ID

Cost to Root

Bridge ID

Port ID

Message Age

Max Age

Hello Time

Forward Delay