IPv6

Protocol#*41*

Copyright

If you find this document useful or further it's distribution we would appreciate you letting us know.

History & 'why not to panic'.

This is has all happened before...

- IPv4 became the Internet's protocol on Jan. 1, 1983
  - Was preceded by NCP
  - The Internet ran dual-stack IP & NCP until 1983-01-01
- DoD standardized on IP in 1984
- Concern about the viability of IPv4 as early as 1991.
  - The IETF ROuting & ADdress group.
- DoD will be fully IPv6 by the end of 2008.
- IANA exhausts IPv4 allocations by May of 2010
  - RIRs will exhaust addresses by April 2011
- Backbones and root DNS already support IPv6.

...this will all happen again.

Is it for real?

- Microsoft Vista & Server 2008 are native IPv6.
- LINUX supports IPvy since mid-life 2.4.x kernels.
- AAAA (IPv6 records) enabled in root DNS server.
  - Febuary 6th, 2008
- IPv4 addresses will be exhausted by 2010
  - http://www.ripe.net/news/community-statement.html
- ICANN says *move now*.
  - http://www.searchcio.com.au/topics/article.asp?DocID=6100812
- Sprint & Verizon will be IPv6 by the end of 2008
- All federal agencies are migrated or migrating to IPv6.

IPv4 is very tired.

- IPv4 addresses will be exhausted by 2010
  - More and more (every?) devices are IP enabled.
    - The problem only gets worse and worse.
  - A constrained address space means....
    - ...networks are NAT'd
      - NAT destroys the point-to-point model.
      - NAT hobbles real-time and multimedia applications
    - ... subnetting into smaller and smaller chunks.
      - Routing tables get bigger and bigger
        - Routing gets slower (or at least not faster)
- IPv4 is cumbersome.
  - Every router in the hop checks the CRC
  - Fragment reassembly

Dual Stack

- How can a network run both IPv4 & IPv6?
  - Easy!
    - The first 4 bits of both an IPv4 and IPv6 packet are a version field.
      - An IPv4 only device should ignore IPv6 traffic.
      - An IPv6 only device should ignore IPv4 traffic.
      - A dual-stack device should pass the packet to the appropriate stack.
  - Many networks used to be multi-stack.
    - Rember NetBIOS, IPX, VINES, etc...

This is has all happened before.

Advantages of IPv6

- Funtional auto-configuration.
- Working multicast
  - Have you ever tried multicast in IPv4?
- Much larger address space.
  - Has many subtle advantages.
- Simpler faster routing.
  - No checksums, no fragments.
- Jumbo packets
  - Individual packets can be up to 4Gb.
  - IPv4 packets are limited to 64k.
- Tunelling is native to the protocol.

IPv6 Problems

- Not more secure.
  - If you transition insecure applications to IPv6...
  - Built-in security features are over sold.
    - Most services do not support them.
    - Still require sane configuration.
- Implementation issues persist.
  - Older operating systems will never support it.
  - Older hardware will never support it.
  - Applications need to be patched / upgraded.
    - Configurations need to recognize IPv6 addresses.
- Some anonymity advocates take issue with IPv6.

Unsupported Interface Types

- Never to be supported
  - SLIP
  - PLIP
  - Raw IP
  - ISDN
    - */dev/isdn?*

- Not currently supported
  - Syncronous PPP
    - */dev/ippp?*
  - TUN

It's not IPv4

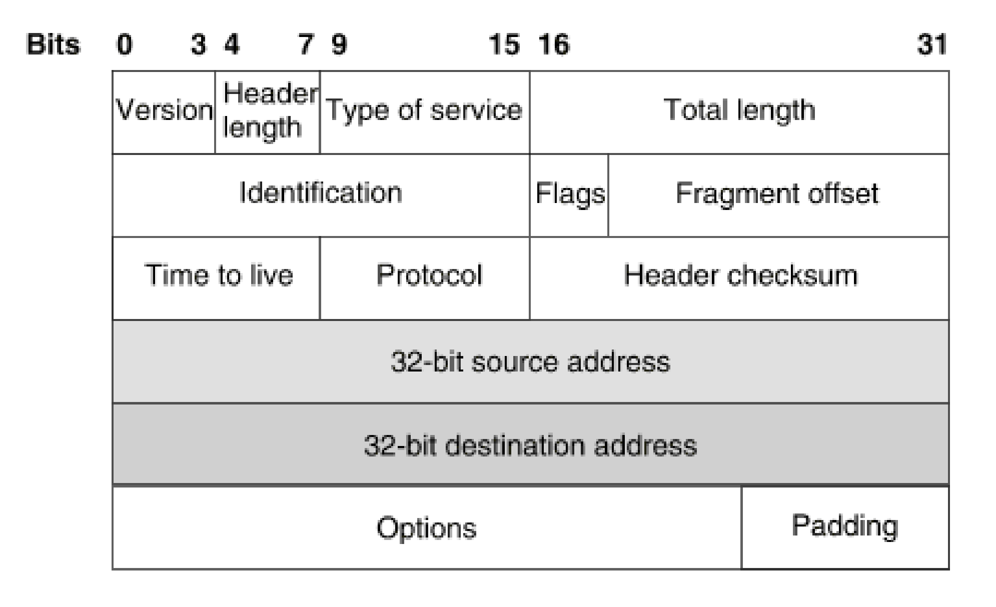
- Multiple addresses per interface.
  - No eth0:1, eth0:2, etc...
  - An interface will almost always have multiple addresses
- Addresses are interface ids, not host ids.
- Multiple types of addresses.
- Netmasks are [usually] fixed.
  - Networks are assigned as /48.
    - Every network has 65,535 subnets.
  - Netmask is usually a /64
    - First 64 bits are the network id.
    - Second 64 bits are the host id.
    - Every subnet has 18,446,744,073,709,551,616 hosts.

It's so not IPv4

- No more broadcast.
- No more ARP
  - Uses link neighbor discovery
- No checksum at the network layer
  - Much more efficient routing.
- No fragmentation.
  - Source host determines packet size.
- Tunneling is inherent.
  - A VPN or tunnel in IPv6 is just an IPv6 connection.

IPv4 Header

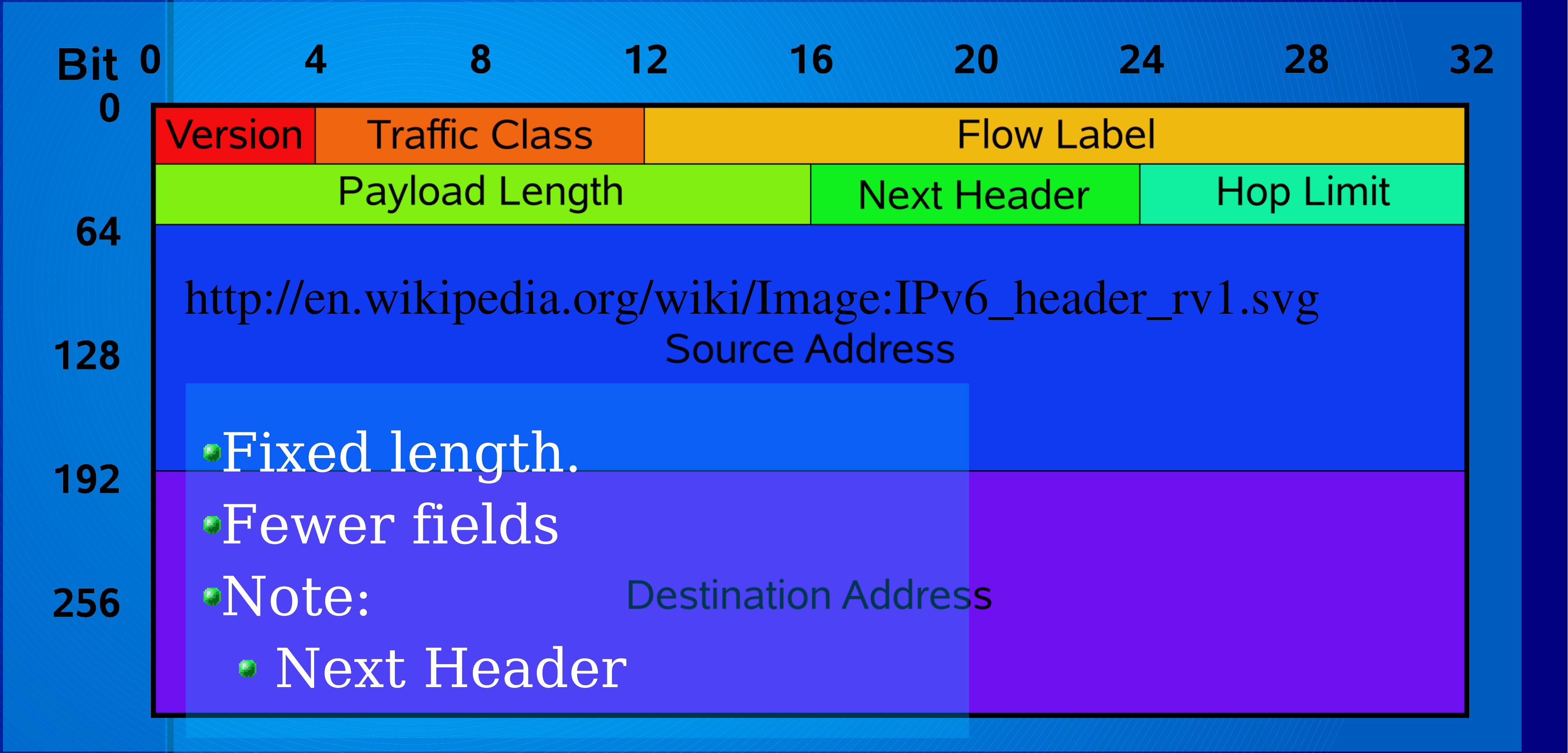
- Variable length.
- Note
  - Checksum field
  - Fragment field
- QoS field (?)



Bits	0 3	4 7	9 15	16	31
	Version	Header length	Type of service	Total length	
	Identification			Flags	Fragment offset
	Time to live		Protocol	Header checksum	
	32-bit source address				
	32-bit destination address				
	Options				Padding

Diagram from:
http://www.synopsys.com/products/designware/dwtb/articles/ethernet_mac/ipv4.gif

IPv6 Header



Bit	0	4	8	12	16	20	24	28	32
0	Version	Traffic Class		Flow Label					
	Payload Length				Next Header		Hop Limit		
64	Source Address								
128									
192	Destination Address								
256									

http://en.wikipedia.org/wiki/Image:IPv6_header_rv1.svg

- Fixed length.
- Fewer fields
- Note:
  - Next Header

http://en.wikipedia.org/wiki/Image:IPv6_header_rv1.svg

IPv6 Addresses

- IP Addresses are 128 bits
  - 0000:0000:0000:0000:0000:0000:0000:0000
    - Each :0000: unit is 16 bytes
    - Eight units of 16 bytes (8 * 16 = 128)
- There are three categories of addresses:
  - Unicast addresses
  - Multicast addresses
    - Identifies a group of nodes
      - Either link local , site local, or global.
      - Packets are delivered to all nodes.
  - Anycast addresses
    - Identifies a group of nodes.
      - Packets are delivered to the nearest node.

Why 128 bits?

$$HD = \frac{\log(allocatedObjects)}{\log(possibleObjects)}$$

0.80 maximum realistic efficiency

0.85 overcrowded

See RFC3194 for additional details.

- IPv4
  - 0.8 corresponds to 50 million hosts.
  - We have passed the point of congestion,
- IPv6
  - 0.8 corresponds to 5.07(10*30)
  - 1,000 hosts for every gram of planet Earth.
  - Allows rather laissez-faire allocation strategies.
  - Expected life-span of 30 years.

IPv6 Unicast Addresses

Normal IP Addreses

- There are five kinds of unicast addresses.
  - Global – A public address.
    - 2001:/16 – Allocated to RIRs
    - 2002:/16 – Allocated to 6to4
  - Link-local – Only exist on the local wire (ethernet).
    - fe80:/10 – Assigned via SLAAC
  - Site-local – Only relevant within a network.
    - Rather like 192.168.x.x and 10.x.x.x
    - fec0:/10
  - Unique local – Lost in bureaucratic purgatory.
    - fc00:/8 – Managed by a yet-to-be-established registrar.
    - fd00:/8 – Uses a random 40 bit string to make a unique /48.
  - Special - Special.

SixXS

https://noc.sixxs.net/tools/grh/ula/

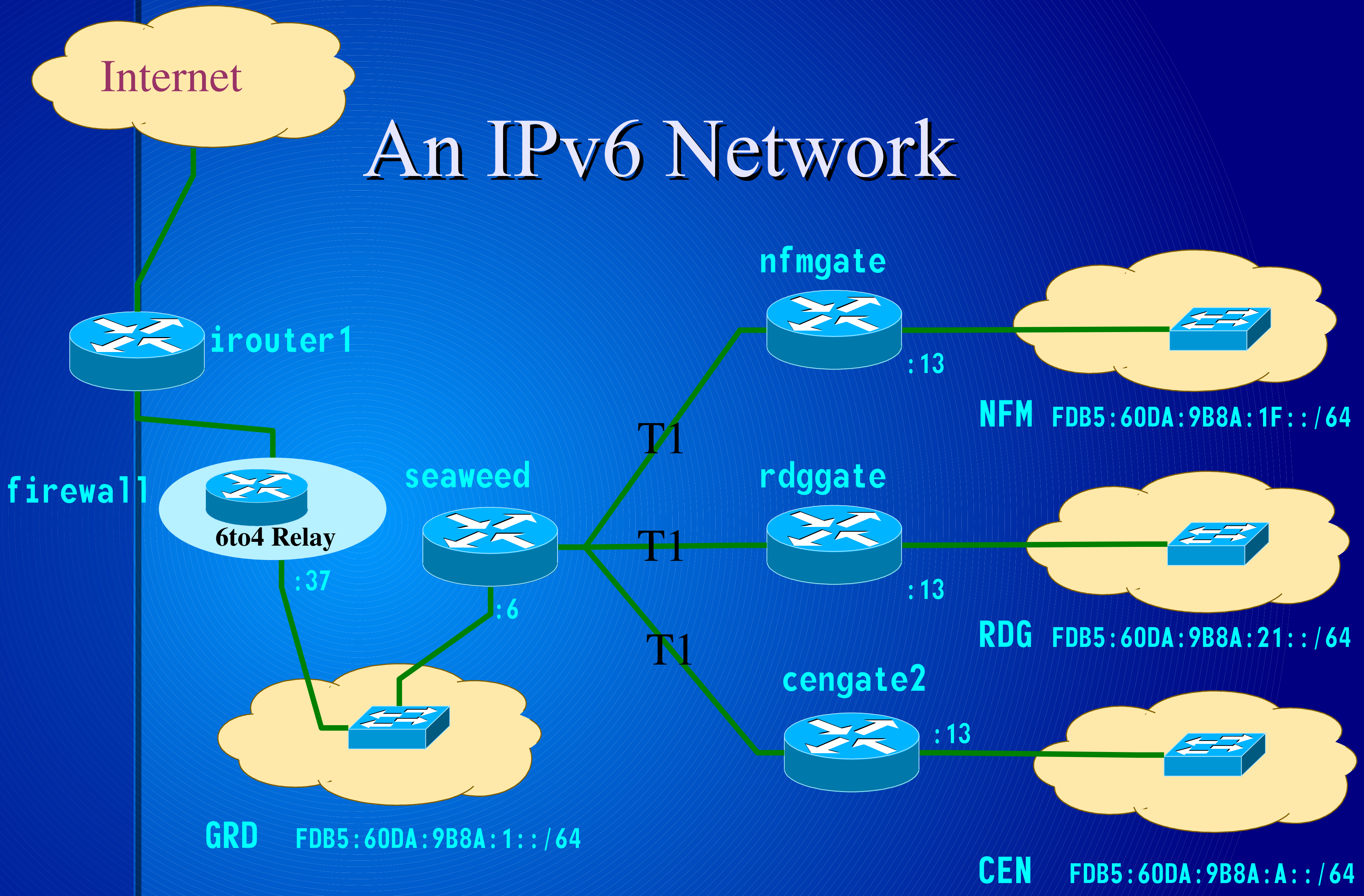
- SixXS operates a rogue registry for generating site local addresses in the fd00:/8 space.
  - Submit a MAC# and SixXS generates and records an allocation.
- Will an official registrar ever be created?
  - ??

fdb5:60da:9b8a::/48	Morrison Industries
fdc9:af33:6bd5::/48	Whitemice Consulting

IPv6 addresses are really long...

- The :: notation shortens represeting IPv6 address.
  - :: replaces a sequence of :0000:
  - 2001:0DB8:0400:965a:: is short for
    - 2001:0DB8:0400:965a:0000:0000:0000:0000
  - 2001:0DB8:0400:965a:0042::1 is short for
    - 2001:0DB8:0400:965a:0042:0000:0000:0001
  - You can't use :: more than once in an address.
    - That would result in ambiguous addresses.
- You can also drop leading zeros.
  - 2001:0DB8:400:965a:0042::1 becomes
    - 2001:DB8:400:965a:42::1

An IPv6 Network


Internet
irouter1
firewall
6to4 Relay
:37
seaweed
:6
T1
T1
T1
nfmgate
:13
NFM FDB5:60DA:9B8A:1F::/64
rdggate
:13
RDG FDB5:60DA:9B8A:21::/64
cengate2
:13
CEN FDB5:60DA:9B8A:A::/64
GRD FDB5:60DA:9B8A:1::/64

ifconfig

```
tyr:~ # /sbin/ifconfig
eth0      Link encap:Ethernet  HWaddr 00:0D:60:1C:9F:26
          inet addr:192.168.1.46  Bcast:192.168.1.255  Mask:255.255.255.0
          inet6 addr: 2002:ce72:2d6b:1::46/64 Scope:Global
          inet6 addr: fdb5:60da:9b8a:1::46/64 Scope:Global
          inet6 addr: fdb5:60da:9b8a:1::42/64 Scope:Global
          inet6 addr: fdb5:60da:9b8a:1::61/64 Scope:Global
          inet6 addr: fe80::20d:60ff:fe1c:9f26/64 Scope:Link
          inet6 addr: fdb5:60da:9b8a:1::48/64 Scope:Global
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1

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:16436  Metric:1
```

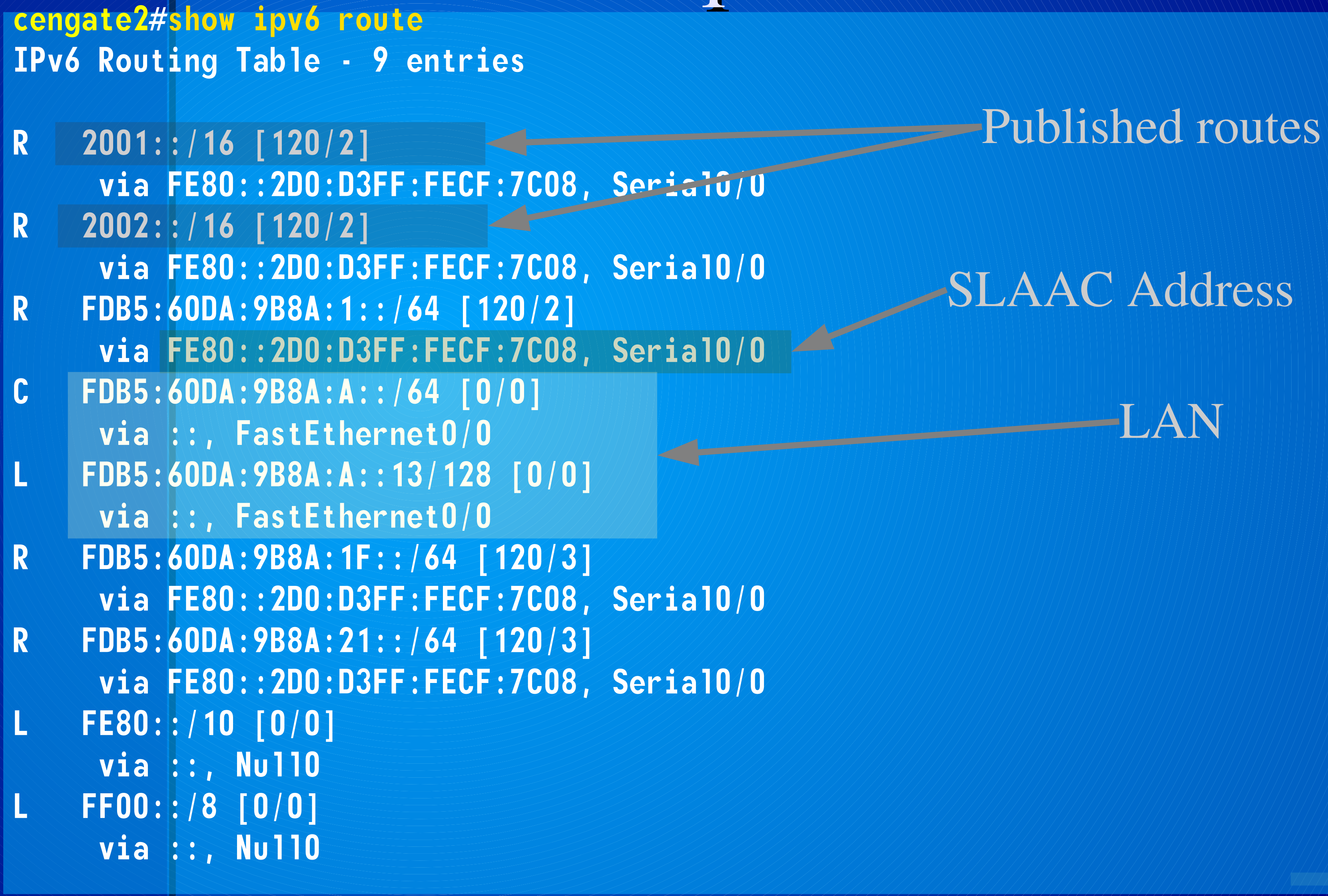
show ipv6 interface

```
Serial2/0/1:0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::21A:2FFF:FE89:9608
  Description: GRD-RDG PtP
  No global unicast address is configured
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::9
    FF02::1:FF89:9608
  MTU is 1500 bytes
  ICMP error messages limited to one every 100 milliseconds
  ICMP redirects are enabled
  Output features: Service-policy
  Service-policy output: voip
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  Hosts use stateless autoconfig for addresses.
```

Routing Table

```
tyr:~ # netstat --inet6 -rn
Kernel IPv6 routing table
Destination                   Next Hop                Flags Metric Ref      Use Iface
::1/128                       ::                      U     0      2163401   3  lo
2002:ce72:2d6b:1::46/128      ::                      U     0      20        1  lo
2002:ce72:2d6b:1::/64         ::                      U     256    0         0  eth0
fdb5:60da:9b8a:1::42/128      ::                      U     0      0         1  lo
fdb5:60da:9b8a:1::46/128      ::                      U     0      13        1  lo
fdb5:60da:9b8a:1::48/128      ::                      U     0      0         1  lo
fdb5:60da:9b8a:1::61/128      ::                      U     0      0         1  lo
fdb5:60da:9b8a:1::/64         ::                      U     256    0         0  eth0
fe80::20d:60ff:fe1c:9f26/128  ::                      U     0      19        1  lo
fe80::/64                     ::                      U     256    0         0  eth0
ff00::/8                      ::                      U     256    0         0  eth0
::/0                          2002:ce72:2d6b:1::55    UG    1024   0         0  eth0
```

show ipv6 route



```
cengate2#show ipv6 route
IPv6 Routing Table - 9 entries

R   2001::/16 [120/2]
     via FE80::2D0:D3FF:FECF:7C08, Serial0/0
R   2002::/16 [120/2]
     via FE80::2D0:D3FF:FECF:7C08, Serial0/0
R   FDB5:60DA:9B8A:1::/64 [120/2]
     via FE80::2D0:D3FF:FECF:7C08, Serial0/0
C   FDB5:60DA:9B8A:A::/64 [0/0]
     via ::, FastEthernet0/0
L   FDB5:60DA:9B8A:A::13/128 [0/0]
     via ::, FastEthernet0/0
R   FDB5:60DA:9B8A:1F::/64 [120/3]
     via FE80::2D0:D3FF:FECF:7C08, Serial0/0
R   FDB5:60DA:9B8A:21::/64 [120/3]
     via FE80::2D0:D3FF:FECF:7C08, Serial0/0
L   FE80::/10 [0/0]
     via ::, Null0
L   FF00::/8 [0/0]
     via ::, Null0
```

Published routes

SLAAC Address

LAN

Route Selection.

- Routing based on longest matching prefix.
  - Packet destination is 2001:abcd::1
    - Router has routes to:
      - 2001:/16
      - 2001:ab::/24
        - This route gets chosen.

Source Address Selection

- An interface almost always has multiple addresses.
- Which address is the source address of a connection?
  - Packet destination is 2001:abcd::1
    - Router has routes to:
      - 2001:/16
      - 2001:ab::/24
        - This route gets chosen.
  - Just like in selecting a route the nearest IP is chosen as the source address of the connection.

Scope

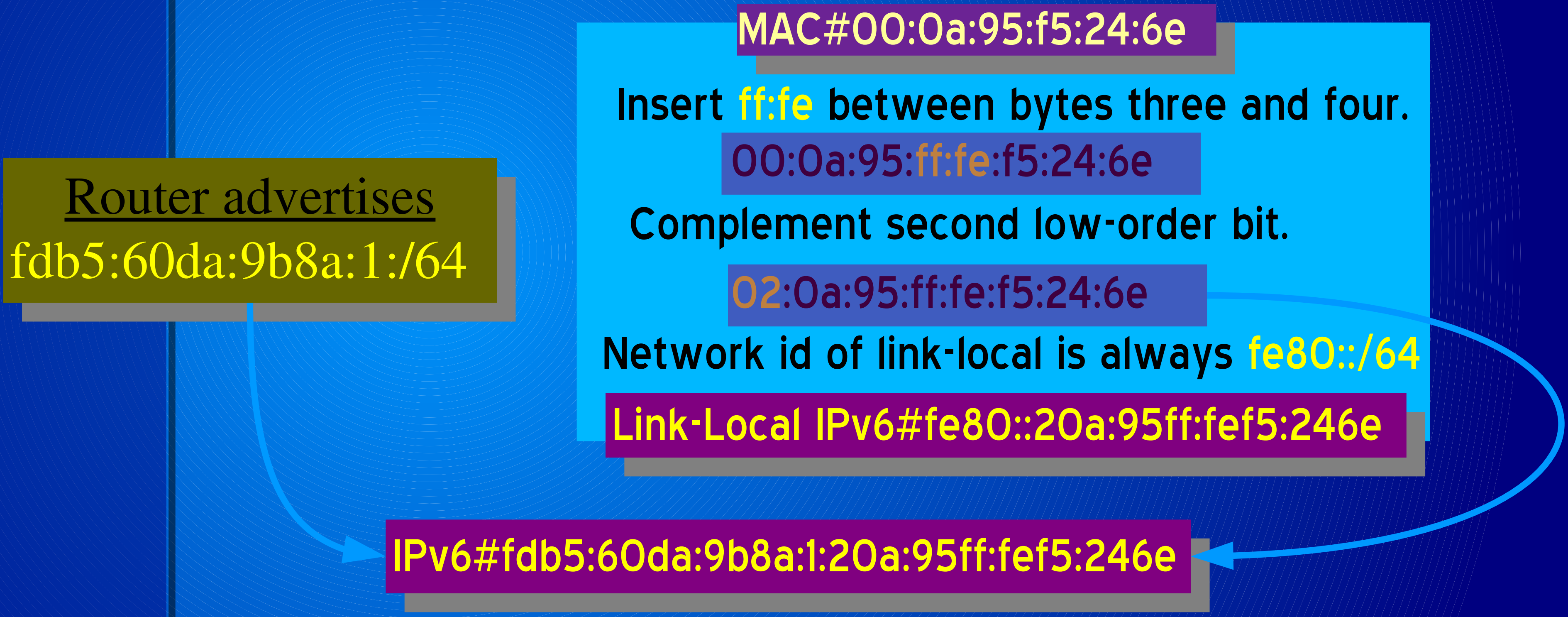
- What if a host has two ethernet interfaces?
  - eth0 has address fe80::216:36ff:fedd:6d36
  - eth1 has address fe80::20d:60ff:fe1c:9f26
  - All link-local addresses appear as the same subnet!
    - Which interface do I use to communicate with another host with a link-local address?

Scope

```
$ ping6 fe80::211:25ff:fe40:9a56%eth0
PING fe80::211:25ff:fe40:9a56(fe80::211:25ff:fe40:9a56) from
fe80::216:36ff:fedd:6d36 eth0: 56 data bytes
64 bytes from fe80::211:25ff:fe40:9a56: icmp_seq=1 ttl=64 time=3.38 ms
...
--- fe80::211:25ff:fe40:9a56 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 3001ms
rtt min/avg/max/mdev = 0.339/1.200/3.382/1.266 ms
```

SLAAC

Stateless Address Auto Configuration



Router advertises
fdb5:60da:9b8a:1:/64

MAC#00:0a:95:f5:24:6e

Insert ff:fe between bytes three and four.

00:0a:95:ff:fe:f5:24:6e

Complement second low-order bit.

02:0a:95:ff:fe:f5:24:6e

Network id of link-local is always fe80::/64

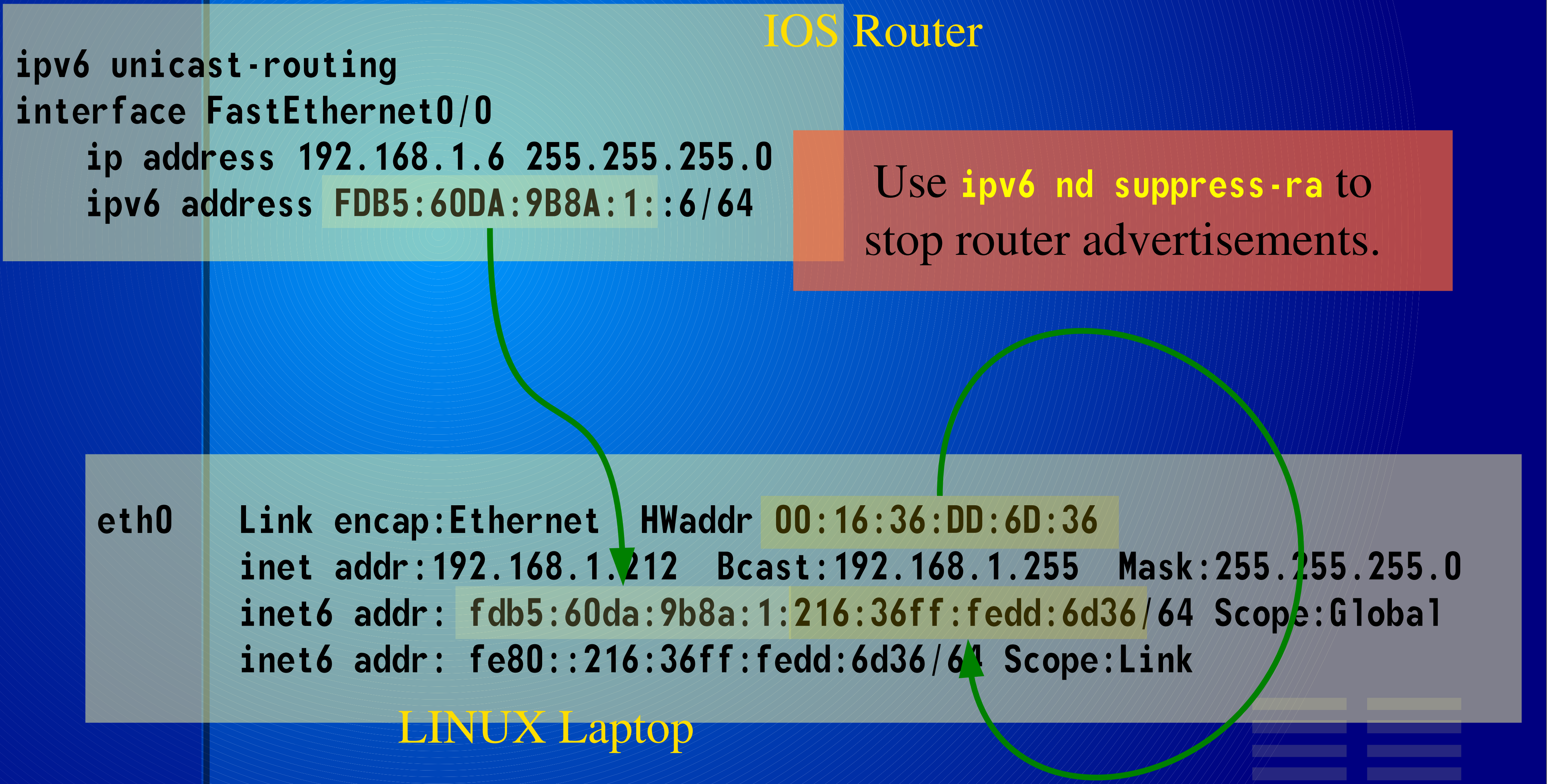
Link-Local IPv6#fe80::20a:95ff:fef5:246e

IPv6#fdb5:60da:9b8a:1:20a:95ff:fef5:246e

If you don't have a traditional router you can run radvd on any LINUX host that supports IPv6.

http://www.litech.org/radvd/

SLAAC in play



IOS Router

```
ipv6 unicast-routing
interface FastEthernet0/0
   ip address 192.168.1.6 255.255.255.0
   ipv6 address FDB5:60DA:9B8A:1::6/64
```

Use **ipv6 nd suppress-ra** to stop router advertisements.

```
eth0      Link encap:Ethernet  HWaddr 00:16:36:DD:6D:36
          inet addr:192.168.1.212  Bcast:192.168.1.255  Mask:255.255.255.0
          inet6 addr: fdb5:60da:9b8a:1:216:36ff:fedd:6d36/64 Scope:Global
          inet6 addr: fe80::216:36ff:fedd:6d36/64 Scope:Link
```

LINUX Laptop

Router Configuration

SEAWEED

```
ipv6 unicast-routing
!
interface FastEthernet0/0
 ip address 192.168.1.6 255.255.255.0
 ipv6 address FDB5:60DA:9B8A:1::6/64
 ipv6 rip mi enable
!
interface Serial2/0/1:0
 ipv6 enable
 ipv6 rip mi enable
!
ipv6 route 2001::/16 FDB5:60DA:9B8A:1::37 100
ipv6 route 2002::/16 FDB5:60DA:9B8A:1::37 100
ipv6 router rip mi
 redistribute static
!
```

RDGGATE

```
ipv6 unicast-routing
!
interface FastEthernet0/0
 ipv6 address FDB5:60DA:9B8A:21::13/64
 ipv6 rip mi enable
!
interface Serial2/0/1:0
 ipv6 enable
 ipv6 rip mi enable
!
ipv6 router rip mi
!
```

ping6

```
awilliam@WM_ADAM1:~> host rdggate
rdggate.morrison.iserv.net has address 192.168.33.19
rdggate.morrison.iserv.net has IPv6 address fdb5:60da:9b8a:21::13
awilliam@WM_ADAM1:~> ping6 fdb5:60da:9b8a:21::13
PING fdb5:60da:9b8a:21::13(fdb5:60da:9b8a:21::13) 56 data bytes
64 bytes from fdb5:60da:9b8a:21::13: icmp_seq=1 ttl=63 time=12.7 ms
64 bytes from fdb5:60da:9b8a:21::13: icmp_seq=2 ttl=63 time=12.2 ms
64 bytes from fdb5:60da:9b8a:21::13: icmp_seq=3 ttl=63 time=10.8 ms
64 bytes from fdb5:60da:9b8a:21::13: icmp_seq=4 ttl=63 time=11.4 ms
64 bytes from fdb5:60da:9b8a:21::13: icmp_seq=5 ttl=63 time=10.8 ms
64 bytes from fdb5:60da:9b8a:21::13: icmp_seq=6 ttl=63 time=11.0 ms
64 bytes from fdb5:60da:9b8a:21::13: icmp_seq=7 ttl=63 time=10.8 ms

--- fdb5:60da:9b8a:21::13 ping statistics ---
7 packets transmitted, 7 received, 0% packet loss, time 6014ms
rtt min/avg/max/mdev = 10.840/11.439/12.727/0.695 ms
```

DNS (bind)

/etc/named.conf

```
options {
     ...
     listen-on-v6 { any; };
     ...
};

zone "a.8.b.9.a.d.0.6.5.b.d.f.ip6.arpa" IN {
  type master;
  database "ldap ldap://192.168.1.9/zoneName=a.8.b.9.a.d.0.6.5.b.d.f.ip6.arpa,...
  allow-update { none; }
}
```

Nibble format.

Reverse lookups occur in the **ip6.arpa** domain.

AAAA

IPv6 addresses are recorded in AAAA records, similar to IPv4's A records.

```
littleboy:~ # host tyr
tyr.backbone.local has address 192.168.1.46
tyr.backbone.local has IPv6 address fdb5:60da:9b8a:1:20d:60ff:fe1c:9f26
tyr.backbone.local mail is handled by 10 kyack.backbone.local.
```

```
dn: relativeDomainName=tyr,zoneName=backbone.local,...
objectClass: top
objectClass: dNSZone
zoneName: backbone.local
relativeDomainName: tyr
dNSTTL: 999
dNSClass: IN
mXRecord: 10 sardine.morrison.iserv.net.
aRecord: 192.168.1.46
aAAARecord: fdb5:60da:9b8a:1:20d:60ff:fe1c:9f26
```

ip6.arpa

(!(ip6.int))

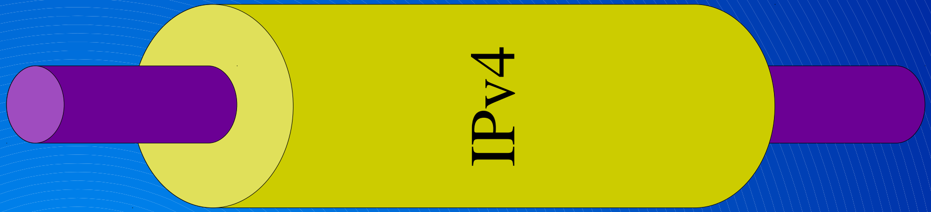
```
littleboy:~ # host fdb5:60da:9b8a:1:20d:60ff:fe1c:9f26
6.2.f.9.c.1.e.f.f.f.0.6.d.0.2.0.1.0.0.0.a.8.b.9.a.d.0.6.5.b.d.f.ip6.arpa domain
 name pointer tyr.backbone.local.
```

```
dn: relativeDomainName=6.2.f.9.c.1.e.f.f.f.0.6.d.0.2.0.1.0.0.0,zoneName=a.8.
 b.9.a.d.0.6.5.b.d.f.ip6.arpa,....
objectClass: top
objectClass: dNSZone
pTRRecord: tyr.backbone.local.
relativeDomainName: 6.2.f.9.c.1.e.f.f.f.0.6.d.0.2.0.1.0.0.0
zoneName: a.8.b.9.a.d.0.6.5.b.d.f.ip6.arpa
```

Nibble format.

6to4

http://www.onlamp.com/pub/a/onlamp/2001/06/01/ipv6_tutorial.html



- 6to4 delivers IPv6 trafic via an IPv4 tunnel.
  - A 6to4 address is created from an IPv4 address
    - The 2002:/16 is allocated to 6to4
    - So 2002 + your 32bit IPv4 address…
      - … and you have a /48 of your own.

```
# ipv4="206.114.45.107"
# printf "2002:%02x%02x:%02x%02x:%04x::1\n" `echo $ipv4 | tr "." " "`
2002:ce72:2d6b:0000::1
```

- 2002:ce72:2d6b:/48 is my IPv6 allocation.

Setting up 6to4

```
ip tunnel add tun6to4 mode sit ttl 32 remote any local 206.114.45.107

ip link set dev tun6to4 up

ip -6 addr add 2002:ce72:2d6b:0000::1/16 dev tun6to4

ip -6 route add 2000::/3 via 2002:c058:6301:: dev tun6to4 metric 1

ip6tables --policy INPUT ACCEPT
ip6tables --policy FORWARD ACCEPT
ip6tables --policy OUTPUT ACCEPT
```

IPv6 address

IPv4 address

Magic anycast address
(RFC3068)

Obviously not what you'd want to do in real life.

curl

On the client:

```
$ curl -6 http://gourd-amber/
```

On the web server:

```
$ tail /var/log/apache2/access_log
fdb5:60da:9b8a:1:216:36ff:fedd:6d36 - - [18/Mar/2008:16:18:18 -0400]
"GET / HTTP/1.1" 200 921 "-" "curl/7.16.4 (i686-suse-linux-gnu)
libcurl/7.16.4 OpenSSL/0.9.8e zlib/1.2.3 libidn/1.0"
```

XPsp2 / Server 2000

- Windows XP has *hacked-in* IPv6 support.
  - To install IPv6
    - ipv6 install
    - This automatically creates a 6to4 interface.
  - To list IPv6 interfaces
    - ipv6 if
  - Does not support…
    - IPv6 DHCP (http://internecine.eu/software/dibbler_dhcpv6.html)
    - DNS over IPv6
    - File & Print (via CIFS) over IPv6
  - Prefers IPv6 over IPv4 if an IPv6 address is provided via DNS.
  - Support for 6to4

Vista / Server 2008

- Windows Vista has complete IPv6 support.
  - Supports IPv6 DHCP & DNS
  - Prefers IPv6 over IPv4 if an IPv6 address is provided via DNS.
    - Except for Internet Explorer, sometimes.
  - 6to4 tunnelling is active by default.

XP

- Windows XP has *hacked-in* IPv6 support.
  - To install IPv6
    - ipv6 install
    - This automatically creates a 6to4 interface.
  - To list IPv6 interfaces
    - ipv6 if
  - Does not support…
    - IPv6 DHCP (http://internecine.eu/software/dibbler_dhcpv6.html)
    - DNS over IPv6
    - File & Print (via CIFS) over IPv6
  - Prefers IPv6 over IPv4 if an IPv6 address is provided via DNS.