Introduction to IPv6 (Part B)

Athanassios Liakopoulos (aliako@grnet.gr)

Greek IPv6 Training, Athens, May 2010

Copy … Rights

Contributions

- B. Tuy
- L. Toutain
- A. Vives
- Athanassios Liakopoulos

Outline

Introduction to IPv6

Outline

Introduction

- IPv6 Addressing Format
- Address Types
- Interface IDs
- IPv6 Addressing Schema

Associated protocols

- Neighbour Discovery
- Path MTU
- Stateless / Stateful Autoconfiguration

IPv6 Addressing

Introduction to IPv6

IPv6 Addressing Format

- 128 bit long addresses
  - Allow hierarchy
  - Flexibility for network evolutions
- Use CIDR principles:
  - Prefix / prefix length
    - 2001:660:3003::/48
    - 2001:660:3003:2:a00:20ff:fe18:964c/64
  - Aggregation reduces routing table size
- Hexadecimal representation
- Interfaces have several IPv6 addresses
- Further reading
  - RFC4291 defines IPv6 addressing scheme
  - RFC3587 defines IPv6 global unicast address format

IPv6 - Addressing Model

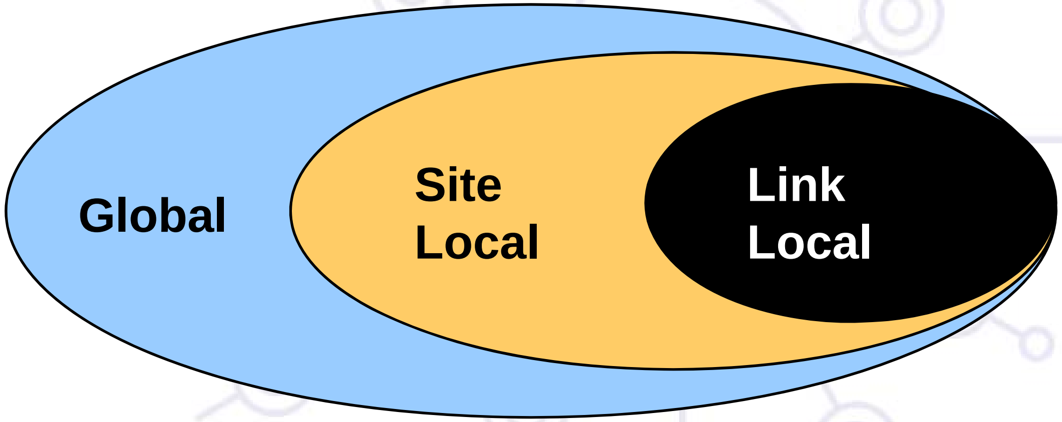
Addresses are assigned to interfaces

- Change from IPv4 model

Interface 'expected' to have multiple addresses

Addresses have scope:

- Link Local
- Site Local
- Global



Addresses have lifetime

- Valid and Preferred lifetime

Site-Local Address deprecated in RFC3879. Replaced by **Unique Local Address (ULA)** in RFC 4193

Address Types

Introduction to IPv6

IPv6 Address Types

- Unicast (one-to-one)
  - Global
  - Link-local
  - Unique Local (ULA)
  - site-local (deprecated in RFC3897)
  - IPv6-mapped
  - IPv4-compatible (deprecated in RFC4291)
- Multicast (one-to-many)
- Anycast (one-to-nearest)
- Reserved

Textual Address Format

- Preferred Form (a 16-byte Global IPv6 Address):

2001:0DB8:3003:0001:0000:0000:6543:210F

- Compact Format:

2001:DB8:3003:1::6543:210F

- IPv4-mapped: ::FFFF:134.1.68.3
- Literal representation
  - [2001:DB8:3003:2:a00:20ff:fe18:964c]
  - http://[2001:DB8::43]:80/index.html

IPv6 Address Type Prefixes

Address Type	Binary Prefix	**IPv6 Notation**
Unspecified	00...0 (128 bits)	**::/128**
Loopback	00...1 (128 bits)	**::1/128**
Multicast	1111 1111	**FF00::/8**
Link-Local Unicast	1111 1110 10	**FE80::/10**
ULA	1111 110	**FC00::/7**
Global Unicast	(everything else)	
IPv4-mapped	00...0:1111 1111:IPv4	**::FFFF:IPv4/128**
Site-Local Unicast (deprecated)	1111 1110 11	**FEC0::/10**
IPv4-compatible (deprecated)	00...0 (96 bits)	**::IPv4/128**

- Global unicast assignments actually use 2000::/3 (001 prefix)
- Anycast addresses allocated from unicast prefixes

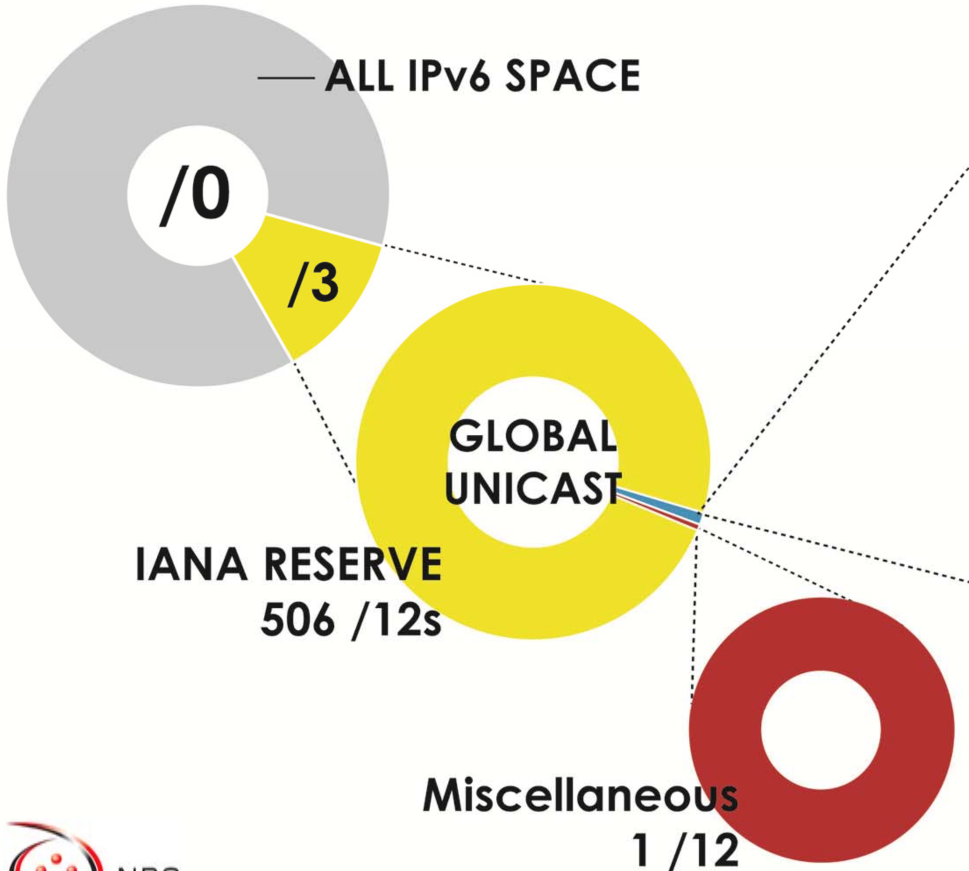
IPv6 Address Space

- Aggregatable Global Unicast Addresses (001): 1/8
- Unique Local Unicast addresses (1111 1110 00): 1/128
- Link-Local Unicast Addresses (1111 1110 10): 1/1024
- Multicast Addresses (1111 1111): 1/256

For	Future	Use	In Use
1/2	1/4	1/8	1/8

- More info: http://www.iana.org/assignments/ipv6-address-space

IPv6 Address Space



RIRs 5 /12s (October 2006)

RIR	IPv6 ADDRESS
AfriNIC	2C00:0000::/12
APNIC	2400:0000::/12
ARIN	2600:0000::/12
LACNIC	2800:0000::/12
RIPE NCC	2A00:0000::/12

NRO

Special-Purpose Unicast Addresses

- **Unspecified address** 0:0:0:0:0:0:0:0 <u>or</u> ::/128
  - Used as a placeholder when no address is available
- **Loopback address** 0:0:0:0:0:0:0:1 <u>or</u> ::1/128
  - for sending packets to itself:
- **Documentation prefix**: 2001:db8::/32

(Further info in RFC5156)

Link-Local & Site-Local Unicast Addresses

- Link-local addresses for use during auto-configuration and when no routers are present (FE80::/10)

10 bits	*54 bits*	*64 bits*
1111 1110 10	**0 . . . 0**	**Interface ID**

FE80

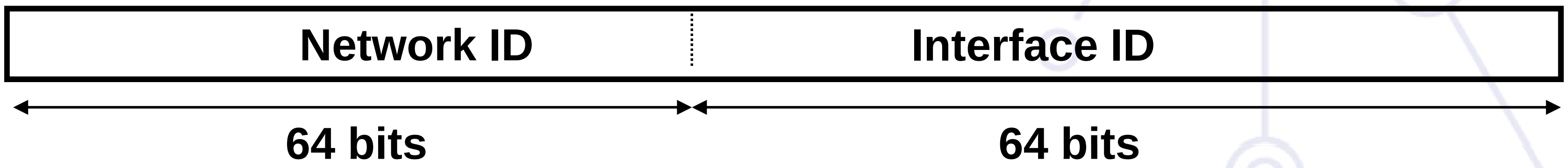
- Site-local addresses for independence from changes of TLA / NLA* (FEC0::/10) (deprecated by RFC3879)

10 bits	*54 bits*	*64 bits*
1111111011	**Subnet ID**	**Interface ID**

FEC0

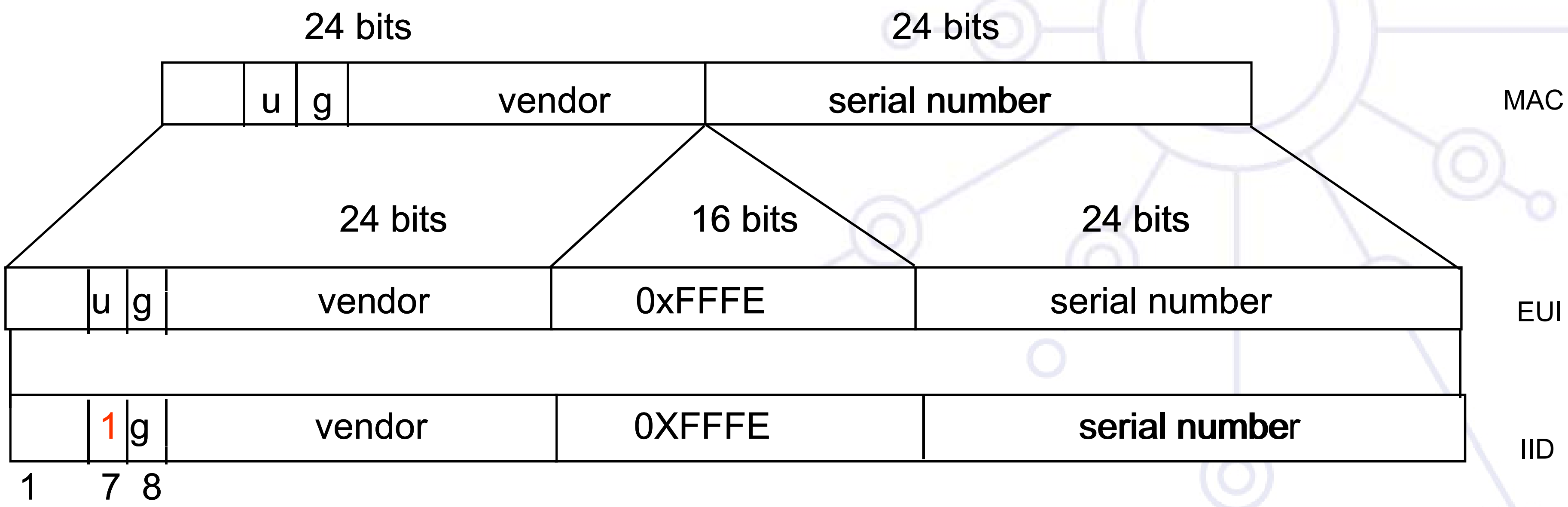
Interface IDs

- The lowest-order 64-bit field of unicast addresses may be assigned in several different ways:
  - auto-configured from a 64-bit MAC address
  - auto-configured from a 48-bit MAC address (e.g., Ethernet) expanded into a 64-bit EUI-64 format
  - assigned via DHCP
  - manually configured
  - auto-generated pseudo-random number (to counter some privacy concerns)
  - CGA (Cryptographically Generated Address)
  - possibly other methods in the future



Autoconfigured Interface IDs (1)

- 64 bits to be compatible with IEEE 1394 (FireWire)
- Eases auto-configuration
- IEEE defines the mechanism to create an EUI-64 from IEEE 802 MAC addresses (e.g. Ethernet, FDDI, etc)



Autoconfigured Interface IDs (2)

- Links with non global identifier → fill first left bits with 0
  - e.g., the Localtalk 8 bit node identifier

- For links without identifiers, there are different ways to proceed (e.g., tunnels, PPP) to have a subnet-prefix-unique identifier:
  - Choose the universal identifier of another interface
  - Manual configuration
  - Node Serial Number
  - Other Node-Specific Token

Multicast Addresses

11111111	flags	scope	group ID
8	4	4	112 bits

- Flags: **0RPT**
  - **0**: The high-order flag is reserved, and must be initialized to 0
  - **T**: Transient, or not, assignment (RFC4291)
  - **P**: Assigned, or not, based on network prefix
  - **R**: Rendezvous Point Address embedded, or not
- Scope field: 1 - Interface-Local
  2 - link-local
  4 - admin-local
  5 - site-local
  8 - organization-local
  E – global
  (3,F - Reserved), (6,7,9,A,B,C,D - Unassigned)
-

Unique Local IPv6 Unicast Addresses (1)

- ULAs are defined in RFC4193:
  - Globally unique prefix with high probability of uniqueness
  - Intended for local communications, usually inside a site
  - They are not expected to be routable on the global Internet
  - They are routable inside of a more limited area such as a site
  - They may also be routed between a limited set of sites
  - Locally-Assigned Local addresses vs. Centrally-Assigned Local addresses

Unique Local IPv6 Unicast Addresses (2)

- ULA characteristics:
  - Well-known prefix to allow for easy filtering at site boundaries
  - ISP independent and can be used for communications inside of a site without having any permanent or intermittent Internet connectivity
  - If accidentally leaked outside of a site via routing or DNS, there is no conflict with any other addresses
  - In practice, applications may treat these addresses like global scoped addresses
  - Sites can be merged without any renumbering of the ULAs
  - Sites can change their provider-based IPv6 unicast address without disrupting any communication that uses Local IPv6 addresses
  - Can be used for inter-site VPNs

Unique Local IPv6 Unicast Addresses (3)

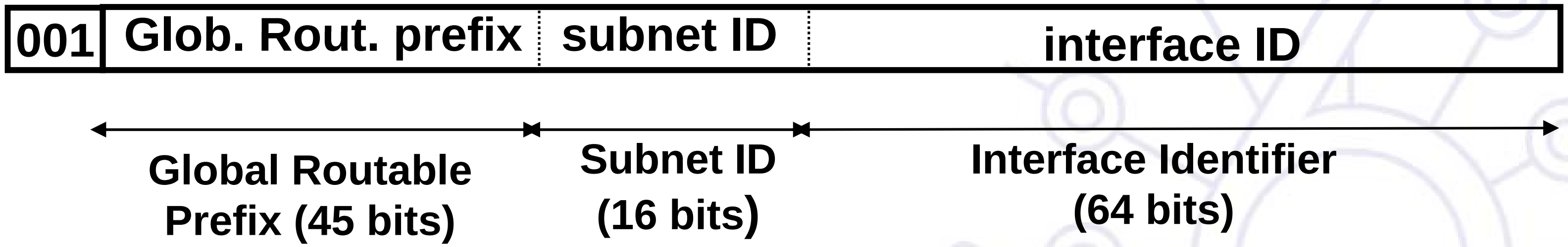
- Format:

Prefix	L	global ID	subnet ID	interface ID
7 bits	1	40 bits	16 bits	64 bits

- FC00::/7 prefix identifies the Local IPv6 unicast addresses
  - L=1, if the prefix is locally assigned
  - L= 0, may be defined in the future (in practice used for centrally assigned prefixes)
- ULA are created using a pseudo-randomly allocated global ID
  - This ensures that there is not any relationship between allocations and clarifies that these prefixes are not intended to be routed globally

Global Unicast Addresses

- Defined in RFC3587



- The global routing prefix is a value assigned to a zone (site, a set of subnetworks/links)
  - It has been designed as an hierarchical structure from the Global Routing perspective
- The subnetwork ID, identifies a subnetwork within a site
  - Has been designed to be an hierarchical structure from the site administrator perspective

Anycast Addresses

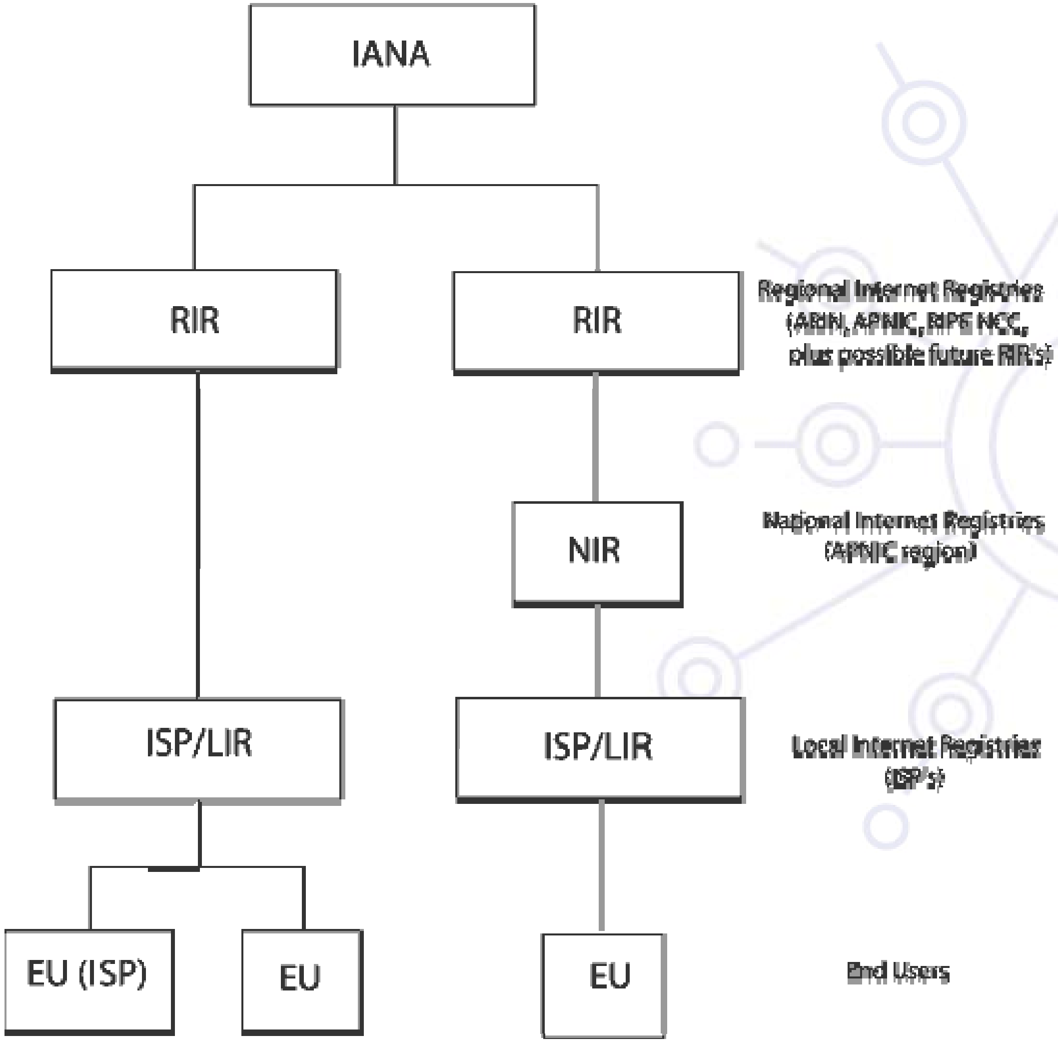
- Identifier for a set of interfaces (typically in different nodes).
  - A packet sent to an anycast address is delivered to the "nearest" interface (routing protocols' distance)
- Taken from the unicast address space (of any scope).
  - Not syntactically distinguishable from unicast addresses
- The nodes assigned a anycast address must be explicitly configured to know that it is an anycast address
  - Assigned only to routers
- Reserved anycast addresses are defined in RFC2526
- The Subnet-Router anycast address mandatory on all routers
  - Used for communication with routers in one or multiple remote networks

n bits	*128 – n bits*
Subnet Prefix	00..00

Addressing Schema

Introduction to IPv6

Production Addressing Scheme (1)



Production Addressing Scheme (2)

001	Glob. Rout. prefix	subnet ID	interface ID
	Global Routable Prefix (45 bits)	Subnet ID (16 bits)	Interface Identifier (64 bits)

- LIRs receive by default /32
  - Production addresses today are from prefixes 2001, 2003, 2400, etc.
  - Can request for more if justified
- /48 used only within the LIR network, with some exceptions for critical infrastructures
- /48 to /128 is delegated to end users
  - Recommendations following RFC3177 and current policies
  - /48 general case, /47 if justified for bigger networks
  - /64 if one and only one network is required
  - /128 if it is sure that one and only one device is going to be connected

Production Addressing Scheme (3)

- Source: www.iana.org/assignments/ipv6-unicast-address-assignments

```
IPv6 Global Unicast Address Assignments [0]
[last updated 2008-05-13]
Global Unicast Prefix Assignment     Date         Note
--------------------- ----------     ------       ----
2001:0000::/23        IANA           01 Jul 99    [1]
2001:0200::/23        APNIC          01 Jul 99
2001:0400::/23        ARIN           01 Jul 99
2001:0600::/23        RIPE NCC       01 Jul 99
2001:0800::/23        RIPE NCC       01 May 02
2001:0A00::/23        RIPE NCC       02 Nov 02
2001:0C00::/23        APNIC          01 May 02    [2]
2001:0E00::/23        APNIC          01 Jan 03
2001:1200::/23        LACNIC         01 Nov 02
. . .
```

RIR Allocation Policies

- AfriNIC:
  - http://www.afrinic.net/IPv6/index.htm
  - http://www.afrinic.net/docs/policies/afpol-v6200407-000.htm *
- APNIC:
  - http://www.apnic.org/docs/index.html
  - http://www.apnic.org/policy/ipv6-address-policy.html *
- ARIN:
  - http://www.arin.net/policy/index.html
  - http://www.arin.net/policy/nrpm.html#ipv6 *
- LACNIC:
  - http://lacnic.net/sp/politicas/
  - http://lacnic.net/sp/politicas/ipv6.html *
- RIPE-NCC:
  - http://www.ripe.net/ripe/docs/ipv6.html
  - http://www.ripe.net/ripe/docs/ipv6policy.html *

  - *describes policies for the allocation and assignment of globally unique IPv6 address space

RIR Allocation Statistics

- AfriNIC:
  - http://www.afrinic.net/statistics/index.htm
- APNIC:
  - http://www.apnic.org/info/reports/index.html
- ARIN:
  - http://www.arin.net/statistics/index.html
- LACNIC:
  - http://lacnic.org/sp/est.html
- RIPE-NCC:
  - http://www.ripe.net/info/stats/index.html

- See http://www.ripe.net/rs/ipv6/stats/

Associated Protocols

Associated Protocols

New Protocols (1)

- New features are specified in IPv6 Protocol -*RFC 2460 DS*
- Neighbor Discovery (NDP) - *RFC 4861 DS*
- Auto-configuration :
  - Stateless Address Auto-configuration -*RFC 4862 DS*
  - DHCPv6: Dynamic Host Configuration Protocol for IPv6 -*RFC 4361 PS*
  - Path MTU discovery (pMTU) -*RFC1981 DS*

New Protocols (2)

- MLD (Multicast Listener Discovery) –*RFC 2710 PS*
  - Multicast group management over an IPv6 link
  - Based on IGMPv2
  - MLDv2 (equivalent to IGMPv3 in IPv4)
- ICMPv6 (RFC 4443 DS) "Super" Protocol that :
  - Covers ICMP (v4) features (Error control, Administration, ...)
  - Transports ND messages
  - Transports MLD messages (Queries, Reports, ...)

Neighbour Discovery

Associated Protocols

Neighbor Discovery for IPv6 (1)

- IPv6 nodes (hosts and routers) on the same physical medium (link) use Neighbor Discovery (NDP) to:
  - discover their mutual presence
  - determine link-layer addresses of their neighbors
  - find neighboring routers that are willing to forward packets on their behalf
  - maintain neighbors' reachability information (NUD)
  - not directly applicable to NBMA (Non Broadcast Multi Access) networks)
    - ➔ NDP uses link-layer multicast for some of its services.

Neighbor Discovery for IPv6(2)

- Protocol features:
  - Router Discovery
  - Prefix(es) Discovery
  - Parameters Discovery, e.g. link MTU, Max Hop Limit, etc
  - Address Autoconfiguration
  - Address Resolution
  - Next Hop Determination
  - Neighbor Unreachability Detection
  - Duplicate Address Detection
  - Redirect

NDP: Comparison with IPv4

- The IPv6 Neighbor Discovery protocol corresponds to a combination of the IPv4 protocols:
  - Address Resolution Protocol (ARP)
  - ICMP Router Discovery (RDISC)
  - ICMP Redirect (ICMPv4)
- Improvements over the IPv4 set of protocols:
  - Router Discovery is part of the base protocol set
  - Router Advertisements carry link-layer addresses and prefixes for a link, and enable Address Autoconfiguration
  - Multiple prefixes can be associated with the same link.
  - Neighbor Unreachability Detection is part of the base protocol set
  - Detects half-link failures and avoids sending traffic to neighbors with which two-way connectivity is absent
  - By setting the Hop Limit to 255, Neighbor Discovery is immune to off-link senders that accidentally or intentionally send ND messages.

NDP Messages (1)

- NDP specifies 5 types of ICMP packets :
  - **Router Advertisement** (RA) :
    - periodic advertisement or response to RS message (of the availability of a router) which contains:
      - list of prefixes used on the link (autoconf)
      - address configuration
      - a possible value for Max Hop Limit (TTL of IPv4)
      - value of MTU
  - **Router Solicitation** (RS) :
    - the host needs RA immediately (at boot time)

NDP Messages (2)

- **Neighbor Solicitation** (NS):
  - to determine the link-layer @ of a neighbor
  - or to check a neighbor is still reachable via a cached L2 @
  - also used to detect duplicate addresses (DAD)

- **Neighbor Advertisement** (NA):
  - answer to a NS message
  - to advertise the change of physical address

- **Redirect** :
  - Used by routers to inform hosts of a better first hop for a destination

Address resolution

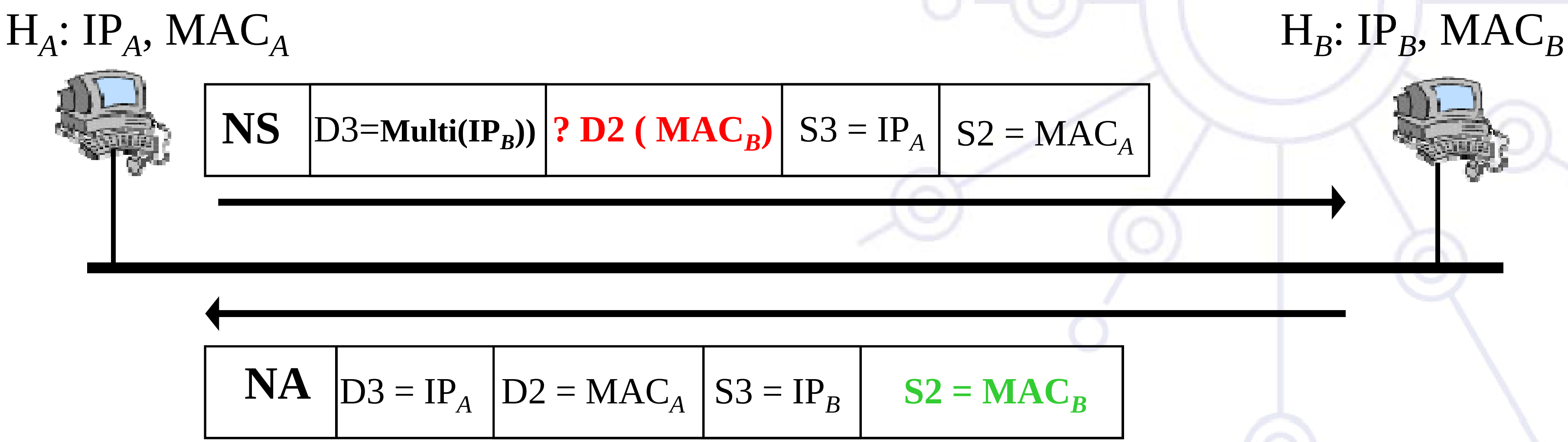
- Address resolution is the process through which a node determines the link-layer address of a neighbor given only its IP address.
- Find the mapping:
- Dst IP @ ➔ Link-Layer (MAC) @
- Recalling IPv4 & ARP
  - ARP Request is broadcasted
    - Request is sent to `ethernet address:`

      `FF-FF-FF-FF-FF-FF`
    - Request contains the src's link local address
  - ARP Reply is sent in unicast to the source
    - Reply contains the destination's link local address

Address resolution with NDP

At boot time, every IPv6 node has to join 2 special multicast groups for each network interface:

- All-nodes multicast group: **`ff02::1`**
- Solicited-node multicast group: **`ff02::1:ffxx:xxxx`**
  - derived from the lower 24 bits of the node's address



Address resolution (3) : multicast solicited address

- Concatenation of the prefix FF02::1:FF00:0/104 with the last 24 bi of the IPv6 address

Example:

- Dst IPv6 @: 2001:0660:010a:4002:4421:21FF:FE24:87c1
- Sol. Mcast @: FF02:0000:0000:0000:0000:0001:FF24:87c1
- Ethernet: 33-33-FF-24-87-c1

Path MTU

Associated Protocols

Path MTU discovery (RFC 1981)

- Derived from RFC1191 (IPv4 version of the protocol)
  - Path = set of links followed by an IPv6 packet between source and destination
- Link MTU = maximum packet length (bytes) that can be transmitted on a given link without fragmentation
- Path MTU (or pMTU) = min { link MTUs } for a given path
- Path MTU Discovery = automatic pMTU discovery for a given path

Path MTU discovery (2)

- Protocol operation
  - makes assumption that pMTU = link MTU to reach a neighbor (first hop)
  - if there is an intermediate router such that
    - link MTU < pMTU
    - ➔ it sends an ICMPv6 message: "Packet size Too Large"
  - source reduces pMTU by using information found in the ICMPv6 message
  - ...

=> Intermediate network element aren't allowed to perform packet fragmentation

Stateless/Stateful Autocofiguration

Associated Protocols

Stateless Autoconfiguration

- Host should be plug & play
- Uses some of the Neighbor Discovery ICMPv6 messages
- When booting, the host asks for network parameters:
  - IPv6 prefix(es)
  - default router address(es)
  - hop limit
  - (link local) MTU

Stateless Autoconfiguration

- Only routers have to be manually configured
  - And/or can use the *Prefix Delegation* option
  - RFC 3633
- Hosts can get automatically an IPv6 address
  - BUT it isn't automatically registered in the DNS

➢ Servers should be manually configured

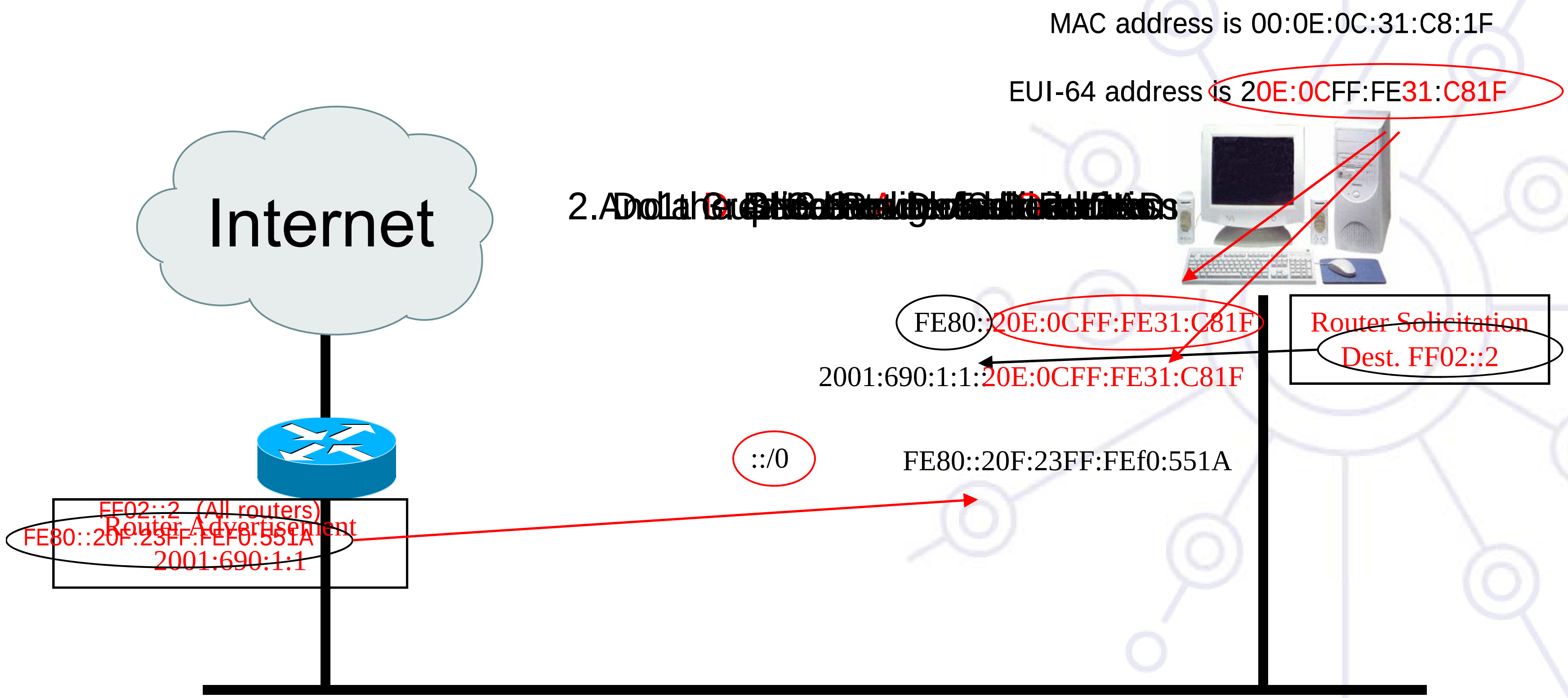
Stateless Autoconfiguration

- IPv6 Stateless Address Autoconfiguration is described in RFC 2462
- Hosts are listening for Router Advertisements (RA) messages, periodically sent out by routers on the local link
- RA messages coming from the router(s) on the link identify the subnet
- Allows a host to create a global unicast IPv6 address from:
  - Its interface identifier (EUI-64 address)
  - Link Prefix (obtained via Router Advertisement)
- Global Address = *Link Prefix* + *EUI-64 address*

Stateless Autoconfiguration

- Usually, the router sending the RA messages is used, by hosts, as the default router
- If the RA doesn't carry any prefix
  - The hosts don't configure (automatically) any global IPv6 address (but may configure the default gateway address)
- RA messages contain two flags indicating what type of stateful autoconfiguration (if any) should be performed
- *It's impossible to automatically send DNS server addresses*
- IPv6 addresses depends on NIC card

Stateless Autoconfiguration



Statefull Autoconfiguration DHCPv6

- Dynamic Host Configuration Protocol for IPv6
  - RFC 3315
  - stateful counterpart to IPv6 Stateless Address Autoconfiguration.
- According to RFC 3315 DHCPv6 is used when:
  - no router is found
  - Or if Router Advertisement message enables use of DHCP

Statefull Autoconfiguration DHCPv6

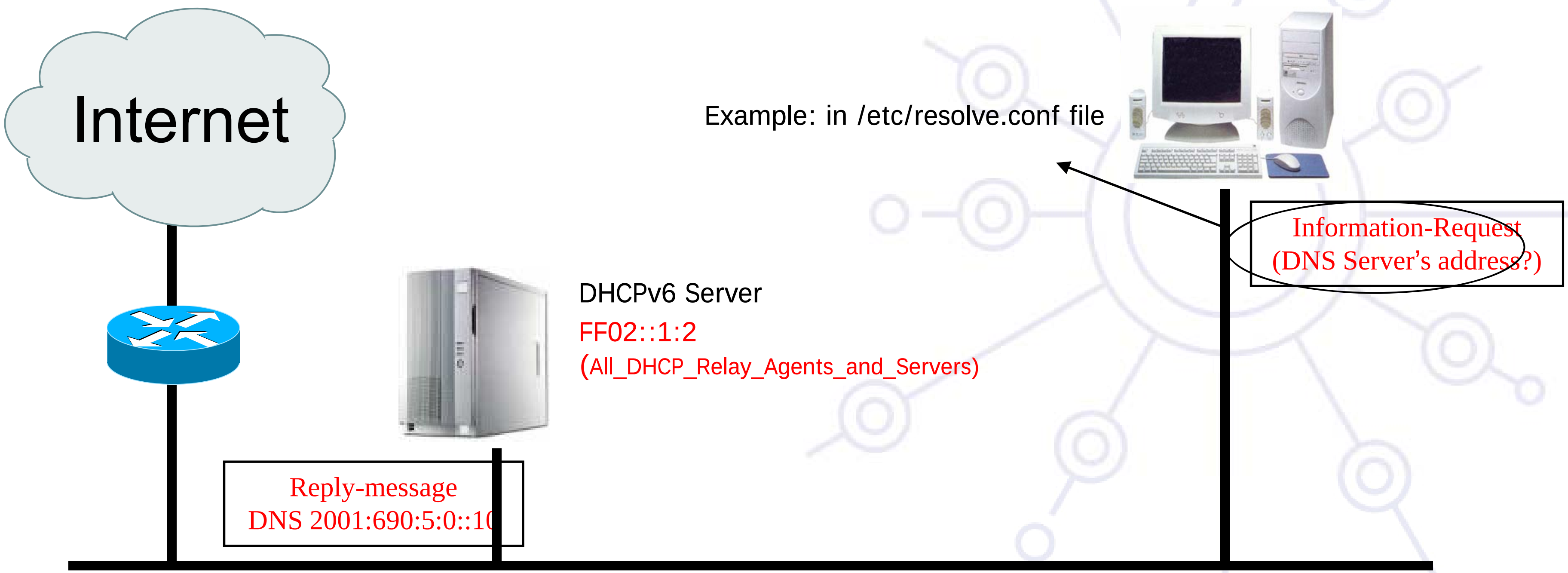
- DHCPv6 works in a client / server model
  - Server
    - Responds to requests from clients
    - Optionally provides the client with:
      - IPv6 addresses
      - Other configuration parameters (DNS servers...)
    - Is listening on multicast addresses:
      - All_DHCP_Relay_Agents_and_Servers (FF02::1:2)
      - All_DHCP_Servers (FF05::1:3)
    - Memorizes client's state
    - Provides means for securing access control to network resources

Statefull Autoconfiguration DHCPv6 /3

- Client
  - initiates requests on a link to obtain configuration parameters
  - uses its link local address to connect the server
  - Sends requests to FF02::1:2 multicast address (All_DHCP_Relay_Agents_and_Servers)

- Relay agent
  - node that acts as an intermediary to deliver DHCP messages between clients and servers
  - is on the same link as the client
  - Is listening on multicast addresses:
    - All_DHCP_Relay_Agents_and_Servers (FF02::1:2)

Statefull Autoconfiguration DHCPv6



Conclusion

- The two types of configuration complement each other
  - Example: we can obtain the address from stateless autoconfiguration and the DNS server address from DHCPv6
- In dual-stack networks we can obtain DNS server addresses from DHCPv4
- DHCPv6 clients aren't still available natively in all Operating Systems.
  - So, we still need to install manually a client
  - Not transparent to users

Questions?