2

INTERNET PROTOCOL VERSION 6 (IPv6)

2.1 INTRODUCTION

During the early 1990s, the fanatical adoption of the Internet as the hottest worldwide communications vehicle led organizations throughout the world to inundate Internet Registries with IP address space requests. This surge in demand for IP address space stimulated the IETF, the engineering and standards body of the Internet, to define a new version of the Internet Protocol that would provide more addressing capacity to meet then and anticipated future address requirements. As discussed in Chapter 1, the adoption of techniques such as CIDR and private address space helped stem the flood of public address space requests; however, these strategies were expected only to prolong the availability of IPv4 address space, albeit for another 10 years or so.

The availability of IPv4 address space continues to diminish and every Regional Internet Registry (RIR) has issued notifications to the Internet community at large that IPv4 space availability is limited and will be exhausted within "a few years." RIRs are responsible for IP address allocation to Internet Service Providers, who in turn allocate space to enterprises, service providers, and any organization requiring IP address space. Ultimately, this exhaustion will impact organizations requiring public IP address space. And Microsoft's Vista™, 7, and Server 2008 products enable IPv6 by default. IPv6 may arrive sooner than you think and with Vista or 7, perhaps whether you'd like it or not!

IP Address Management: Principles and Practice, by Timothy Rooney

IP Header
IP Packet Contents

Figure 2.1. IP commonality in header and packet concept.

Version 6 of the Internet Protocol[*] is an evolution from version 4 but is not inherently compatible with version 4. Chapter 15 describes several migration and coexistence techniques. The primary objective for version 6 was essentially to redesign version 4 based on the prior 20 years of experience with IPv4. Real-world application support added to the IPv4 protocol suite over the years was designed into IPv6 from the outset. This included support for security, multicast, mobility, and autoconfiguration.

The most striking difference in the evolution from IPv4 to IPv6 is the tremendous expansion of the size of the IP address field. Whereas IPv4 uses a 32-bit IP address field, IPv6 uses 128 bits. A 32-bit address field provides a maximum of 2^{32} addresses or 4.2 billion addresses. A 128-bit address field provides 2^{128} addresses or 340 trillion trillion trillion addresses or 340 undecillion[†] (3.4×10^{38}) addresses. To put some context around this tremendously large number, consider that this quantity of IP addresses

- averages to 5×10^{28} IP addresses per person on Earth based on a 6.5 billion population;
- averages to 4.3×10^{20} IP addresses per square inch of the Earth's surface;
- amounts to about 14 million IP addresses per nanometer to the nearest galaxy, Andromeda, at 2.5 million light years.

Like IPv4, not every single address will necessarily be usable due to subnetting inefficiencies, but a few undecillion of wasted addresses won't have much impact! Beyond this seemingly incomprehensible number of IP addresses, there are a number of similarities between IPv6 and IPv4. For example, at a basic level, the "IP packet" concept applies equally well for IPv6 as IPv4 in terms of the concept of the packet header and contents (Figure 2.1), as does the basic concept of protocol layering, packet routing, and CIDR allocations. We'll focus on the variety of defined IPv6 addresses in this chapter and discuss IPv6 subnetting and allocation techniques in the next chapter.

2.1.1 IPv6 Key Features

The IETF has attempted to develop IPv6 as an evolution of IPv4. The evolutionary strategy in migrating from IPv4 to IPv6 is intended to enable IPv6 to provide many new

[*] IP version 5 was never implemented as an official version of IP. The version number of "5" in the IP header was assigned to denote packets carrying an experimental real-time stream protocol called ST, the Internet Stream Protocol. If you'd like to learn more about ST, please refer to RFC 1819 (169).

[†] We're using the American definition of undecillion of 10^{36}, not the British definition that is 10^{66}.

features while building on the foundational concepts that made IPv4 so successful. Key IPv6 features include

- *Expanded Addressing*. 128 bits hierarchically assigned with address scoping (e.g., local link versus global) to improve scalability.
- *Routing*. Strongly hierarchical routing, supporting route aggregation.
- *Performance*. Simple (unreliable, connectionless) datagram service.
- *Extensibility*. New flexible extension headers provide built-in extensibility for new header types and more efficient routing.
- *Multimedia*. Flow label header field facilitates quality of service (QoS) support.
- *Multicast*. Replaces broadcast and is compulsory.
- *Security*. Authentication and encryption are built-in.
- *Autoconfiguration*. Stateless and stateful address self-configuration by IP devices.
- *Mobility*. Mobile IPv6 support.

2.1.2 The IPv6 Header

The IPv6 header layout is shown in Figure 2.2. While the size of both the source and destination IP address fields quadrupled, the overall IP header size only doubled. The fields in the IPv6 header are as follows:

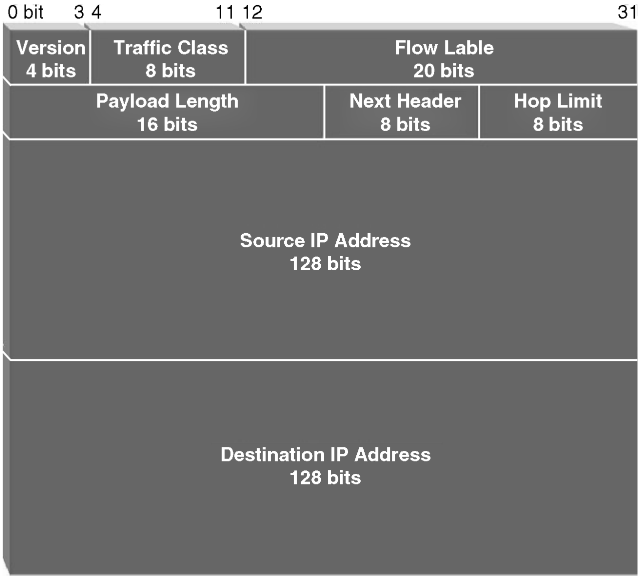

Figure 2.2. IPv6 header (10).

Version. The Internet Protocol version, 6 in this case.

Traffic Class. This field replaces the IPv4 type of service/DS header field and indicates the type or priority of traffic in order to request routing treatment.

Flow Label. Identifies the "flow" of traffic between a source and destination to which this packet belongs as set by the source. This is intended to enable efficient and consistent routing treatment for packets within a given communications session, such as those within a real-time transmission versus a best-effort data transmission.

Payload Length. Indicates the length of the IPv6 payload, that is, the portion of the packet after the base IPv6 header, in octets. Extension headers, if included, are considered part of the payload and are counted within this length parameter.

Next Header. This field indicates the type of header that follows this IP header. This may be an upper layer protocol header (e.g., TCP, ICMPv6, etc.) or an extension header. The extension header concept enables specification of source routing, fragmentation, options, and other parameters associated with the packet only when they are necessary, not as overhead on all packets as in IPv4.

Hop Limit. Analogous to the IPv4 TTL field, this field specifies the number of hops over which this packet may traverse before being discarded. Each router decrements the value of this header field upon forwarding of the packet.

Source IP Address. The IPv6 address of the sender of this packet.

Destination IP Address. The IPv6 address of the intended recipient(s) of this packet.

2.1.3 IPv6 Addressing[*]

Three types of IPv6 addresses have been defined. Like IPv4, these addresses apply to interfaces, not nodes. Thus, a printer with two interfaces would be addressed by either of its interfaces. The printer can be reached on either interface, but the printer node does not have an IP address per se.[†] Of course, for end users attempting to access a node, DNS can hide this subtlety by enabling a hostname to map to one or more interface addresses.

Unicast. The IP address of a single interface. This is analogous to the common interpretation of an IPv4 host address (nonmulticast/nonbroadcast /32 IPv4 address).

Anycast. An IP address for a set of interfaces usually belonging to different nodes, any one of which is the intended recipient. An IP packet destined for an anycast address is routed to the nearest interface (according to routing table metrics) configured with the anycast address. The concept is that the sender doesn't necessarily care which particular host or interface receives the packet, but that

[*] Introductory sections of this chapter are based on material from Chapter 2 of Ref. 11.

[†] Many router and server products support the concept of a "box address" via a software loopback address. This loopback address, not to be confused with the 127.0.0.1 or ::1 loopback addresses, enables reachability to any one of the device's interfaces.

one of those sharing the anycast address receives it. Anycast addresses are assigned from the same address space from which unicast addresses have been allocated. Thus, one cannot differentiate a unicast address from an anycast address by sight. Anycast in IPv4 networks has recently created a buzz in providing similar *closest routing to the intended service*, such as for DNS servers by using a shared unicast IPv4 address. This provides benefits in simplifying client configuration, in having it always use the same [anycast] IP address to query a DNS server, regardless of where on your network the client is connected. We'll discuss DNS deployment using anycast addresses in Chapter 11.

Multicast. An IP address for a set of interfaces typically belonging to different nodes, all of which are intended recipients. This of course is similar to IPv4 multicast. Unlike IPv4, IPv6 does not support broadcasts. Instead, applications that utilized broadcasts in IPv4, such as DHCP, use multicast to a well-known (i.e., predefined) DHCP multicast group address in IPv6.

A device interface may have multiple IP addresses of any or all address types. IPv6 also defines a link local scope of IP addresses to uniquely identify interfaces attached to a particular link, such as a LAN. Additional scoping can be administratively defined per site or per organization, for example, as we'll discuss later in this chapter.

2.1.4 Address Notation

Recall that IPv4 addresses are represented in dotted decimal format where the 32-bit address is divided into four 8-bit segments, each of which are converted to decimal, and then separated with "dots." If you thought remembering a string of four decimals was difficult, IPv6 will make life a little tougher. IPv6 addresses are not expressed in dotted decimal notation; they are represented using a colon-separated hexadecimal format. Jumping down to the bit level, the 128-bit IPv6 address is divided into eight 16-bit segments, each of which is converted to hexadecimal, and then separated by colons. Each hexadecimal "digit" represents four bits as per the mapping of each hex digit (0–F) to its 4-bit binary values below. Each hex digit corresponds to 4 bits with the following possible values.

0 = 0000	4 = 0100	8 = 1000	C = 1100
1 = 0001	5 = 0101	9 = 1001	D = 1101
2 = 0010	6 = 0110	A = 1010	E = 1110
3 = 0011	7 = 0111	B = 1011	F = 1111

After converting a 128-bit IPv6 address from binary into hex, we group sets of four hex digits and separate them with colons. We'll use the term *nibble* to represent a grouping of four hex digits or 16 bits; thus, we have eight nibble values separated by colons, rendering an IPv6 address appearing as shown in Figure 2.3.

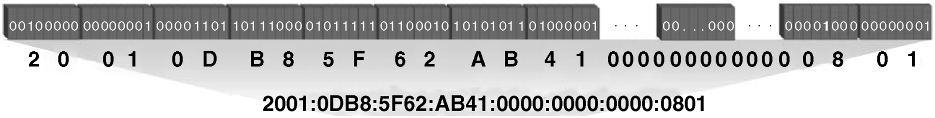

Figure 2.3. IPv6 address: binary to hexadecimal (11).

Instead of dealing with four decimal values, each between 0 and 255, separated by dots in IPv4, IPv6 addresses consist of up to eight hexadecimal values, each between 0 and FFFF, separated by colons. There are two acceptable abbreviations when writing IPv6 addresses. First, leading zeroes within a nibble section, that is, between colons, may be dropped. Thus, the above address could be abbreviated as

2001 : DB8 : 5F62 : AB41 : 0 : 0 : 0 : 801

The second form of abbreviation is the use of a double colon to represent one or more consecutive sets of zero nibbles. Using this form of abbreviation, the above address can be further abbreviated as

2001 : DB8 : 5F62 : AB41 :: 801

Isn't that much better? Note that only one double colon may be used within an address representation. Since there are always eight nibble segments in the address, one can easily calculate how many of them are zero with one double-colon notation; however, it would be ambiguous with more than one.

Consider the address 2001:DB8:0:56FA:0:0:0:B5. We can abbreviate this address as either

2001 : DB8 :: 56FA : 0 : 0 : 0 : B5 or 2001 : DB8 : 0 : 56FA :: B5

We can easily calculate that the double colon denotes one nibble (8 total minus 7 nibbles shown) in the first case and three (8 minus 5 shown) in the second notation. If we attempted to abbreviate this address as 2001:DB8::56FA::B5, we could not unambiguously decode this, as it could represent any of the following possible addresses:

2001 : DB8 : 0 : 56FA : 0 : 0 : 0 : B5
2001 : DB8 : 0 : 0 : 56FA : 0 : 0 : B5
2001 : DB8 : 0 : 0 : 0 : 56FA : 0 : B5

Thus, the requirement holds that only one double colon may appear in an IPv6 address.

2.1.5 Address Structure

The IPv6 address is divided into three fields, as shown in Figure 2.4.

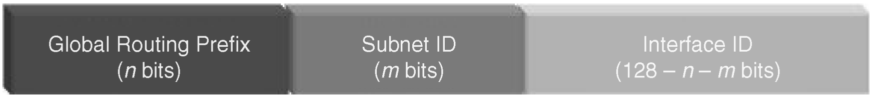

Figure 2.4. IPv6 address structure (12).

The global routing prefix is akin to an IPv4 network number and is used by routers to forward packets to router(s) locally serving the network corresponding to the prefix. For example, a customer of an ISP may be assigned a /48-sized global routing prefix and all packets destined to this customer would contain the corresponding global routing prefix value. In this case, $n = 48$ as per Figure 2.4. When denoting a network, the global routing prefix is written, followed by slash, and then the network size, called the prefix length. Assuming that our example IPv6 address, 2001:DB8:5F62:AB41::801, resides within a /48 global routing prefix, this prefix address would be denoted as 2001:DB8:5F62::/48. As with IPv4, the network address is denoted with zero-valued bits beyond the prefix length (bits 49–128 in this case) as denoted by the terminating double colon.

The subnet ID provides a means to denote particular subnets within the organization. Our ISP customer with a /48 may choose to use 16 bits for the subnet ID, providing 2^{16} or 65,534 subnets. In this case, $m = 16$ as per Figure 2.4. This leaves $128 - 48 - 16 = 64$ bits for the interface ID. The interface ID denotes the interface address of the source or intended recipient for the packet. As we'll discuss a bit later, the global unicast address space that has been allocated for use so far requires a 64-bit interface ID field.

One of the unique aspects of this IPv6 address structure in splitting a network ID consisting of the global routing prefix and subnet ID, from an interface ID, is that a device can retain the same interface ID independent of the network to which it is connected, effectively separating "who you are," your interface ID, from "where you are," your network prefix. As we'll see, this convention facilitates address autoconfiguration, though not without privacy concerns. But we're getting a little ahead of ourselves, so let's jump back up to the macro level and consider the IPv6 address space allocated so far by the Internet addressing authority, the Internet Assigned Numbers Authority (IANA).

2.2 IPv6 ADDRESS ALLOCATIONS

The address space that has been allocated so far by IANA is highlighted in dark gray in Table 2.1 and is discussed in the ensuing text. These allocations represent less than 14% of the total available IPv6 address space.

2.2.1 ::/3—Reserved Space

Address space prefixed with $[000]_2$ is currently reserved by the IETF. Addresses within this space that have unique meaning include the unspecified (::) address and the loopback (::1) address. The IPv6 addressing architecture specification, RFC 4291 (12), requires that all unicast IPv6 addresses, except those within this address space (that is beginning with ::/3 ($[000]_2$)), must utilize a 64-bit interface ID field, and this interface

TABLE 2.1. IPv6 Address Allocations (13)

IPv6 Prefix	Binary Form	Relative Size of IPv6 Space	Allocation
0000::/3	000	1/8	Reserved by IETF: the "unspecified address" (::) and the loopback address (::1) are assigned from this block
2000::/3	**001**	**1/8**	**Global unicast address space**
4000::/3	010	1/8	Reserved by IETF
6000::/3	011	1/8	Reserved by IETF
8000::/3	100	1/8	Reserved by IETF
A000::/3	101	1/8	Reserved by IETF
C000::/3	110	1/8	Reserved by IETF
E000::/4	1110	1/16	Reserved by IETF
F000::/5	1111 0	1/32	Reserved by IETF
F800::/6	1111 10	1/64	Reserved by IETF
FC00::/7	**1111 110**	**1/128**	**Unique local unicast**
FE00::/9	1111 1110 0	1/512	Reserved by IETF
FE80::/10	**1111 1110 10**	**1/1024**	**Link local unicast**
FEC0::/10	1111 1110 11	1/1024	Reserved by IETF
FF00::/8	**1111 1111**	**1/256**	**Multicast**

ID field must utilize the modified EUI-64* algorithm to map the interface's layer 2 or hardware address to an interface ID. Thus, addresses within the ::/3 address space can have any length interface ID field, unlike the remainder of the IPv6 unicast address space, which must utilize a 64-bit interface ID field.

2.2.2 2000::/3—Global Unicast Address Space

The global unicast address space allocated so far, 2000::/3, represents 2^{125} or 4.25×10^{37} IP addresses. Given the 64-bit interface ID requirement defined in the IPv6 addressing architecture [RFC 4291 (12)], the global unicast address format as formally defined in RFC 3587 (14) is shown in Figure 2.5.

The first three bits are $[001]_2$ to indicate global unicast address space. The following 45 bits comprise the global routing prefix, followed by the 16-bit subnet ID and 64-bit interface ID, respectively. Current guidelines call for ISPs allocating /48 networks to their customers, thereby assigning global routing prefixes to customers. Each customer may then define up to 65,534 subnets by uniquely assigning values within the remaining 16-bit subnet ID field for each subnet.

* EUI-64 refers to the 64-bit Extended Unique Identifier defined by the IEEE. We'll cover the modified EUI-64 algorithm later in this chapter.

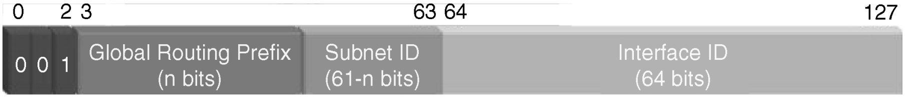

Figure 2.5. Global unicast address format (14).

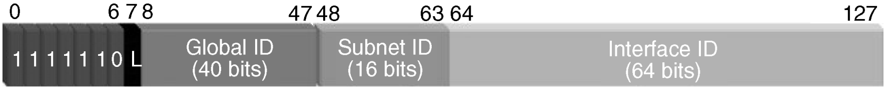

Figure 2.6. Unique local address format (15).

2.2.3 FC00::/7—Unique Local Address Space

The unique local address (ULA) space, defined in RFC 4193 (15), is intended to provide locally assignable and routable IP addresses, usually within a site. RFC 4193 states that “these addresses are not expected to be routable on the global Internet.” Thus, while not as stringent as RFC 1918 in defining private IPv4 address space, the unique local address space is essentially private address space, providing “local” addressing with a high probability of still being globally unique. The format of unique local address space is shown in Figure 2.6.

The first seven bits, bits 0–6, are $[1111\ 110]_2$ = FC00::/7, which identifies a unique local address. The eighth bit, the “L” bit, is set to “1” if the global ID is locally assigned; setting the “L” bit to “0” is currently undefined, though the Internet community (IETF) has discussed enabling this setting for globally unique local addresses, assignable through Internet Registries. The 40-bit global ID field is intended to represent a globally unique prefix and must be allocated using a pseudorandom algorithm, not sequentially. In either case, the resulting /48 prefix comprises the organization’s ULA address space, from which subnets can be allocated for internal use. The subnet ID is a 16-bit field to identify each subnet, while the interface ID is a 64-bit field.

An example pseudorandom approach to derive a unique global ID as described in RFC 4193 recommends computing a hash[*] of

- the current time as reported by a Network Time Protocol (NTP) server in 64-bit NTP format,
- concatenated with an EUI-64 interface ID of an interface on the host performing this algorithm.

The least significant (rightmost) 40 bits of the result of the hash operation are then populated as the global ID.

[*] A hash is created by performing a mathematical operation on the data to be hashed and a random value. A particular mathematical algorithm, the Secure Hash Algorithm 1 or SHA-1, is required in this case.

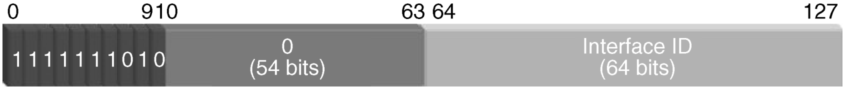

Figure 2.7. Link local address format (12).

2.2.4 FE80::/10—Link Local Address Space

Link local addresses are used only on a particular link, such as an Ethernet link; packets with link local destination addresses are not routed. That is, packets having link local addresses will not reach beyond the corresponding link. These addresses are used for address autoconfiguration and neighbor discovery, which will be discussed later. The format of link local addresses is shown in Figure 2.7.

The FE80::/10 link local prefix is followed by 54 zero bits and the 64-bit interface ID.

2.2.5 FF00::/8—Multicast Address Space

Multicast addresses identify a group of interfaces typically on different nodes. Think of multicast addresses as a scoped broadcast. All multicast group members share the same group ID and hence all members will accept packets destined for the multicast group. An interface may have multiple multicast addresses; that is, it may belong to multiple multicast groups. The basic format of IPv6 multicast addresses is shown in Figure 2.8.

The prefix FF00::/8 identifies a multicast address. The next field is a 4-bit field called "flags." The format of the multicast address depends on the value of the flags field. The scope (also affectionately referred to as "scop") field indicates the breadth of the multicast scope, whether per node, link, global, or other scope values defined below. The value of the flags and scope fields can fortunately be easily discerned by looking at the third and fourth hex digits within the address, respectively, as we'll summarize a bit later.

Flags. The flags field is comprised of 4 bits, which we'll discuss starting from right to left (12):

0	R	P	T

- The T bit indicates whether the multicast address is of transient nature or is a well-known address assigned by IANA. The T bit is defined as follows.

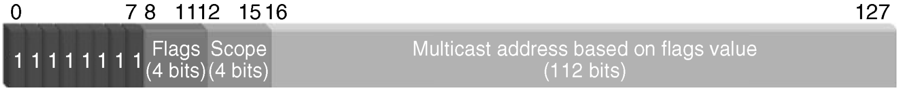

Figure 2.8. Multicast address format (12).

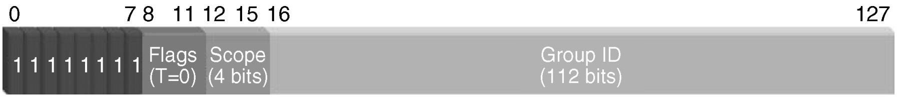

Figure 2.9. Multicast address with flag T = 0.

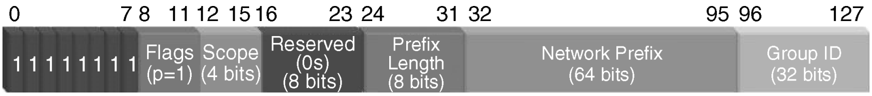

Figure 2.10. Multicast address with flag P = 1 (16).

- T = 0. This is an IANA-assigned well-known multicast address (Figure 2.9). In this case, the 112-bit multicast address is a 112-bit group ID field.

  IANA has assigned numerous group IDs thus far.[*] For example, group ID = 1 refers to all nodes within the associated scope (defined by the scope field), group ID = 2 refers to all routers within the scope, and so on. The scope field is defined below, but example well-known multicast addresses are

  - F01::1 = all nodes on this link.
  - FF02::2 = all routers on this link.
  - FF05::1 = all nodes on this site.
  - FF05::2 = all routers on this site.

- T = 1. This is a temporarily assigned or transient multicast address. This can be an address assigned for a specific multicast session or application. An example might be FF12::3:F:10.

- The P bit indicates whether the multicast address is comprised partly of a corresponding unicast network prefix or not. The P bit is defined[†] as follows:

  - P = 0. This multicast address *is not* assigned based on the network prefix. The format of a multicast packet with P = 0 is as described above (i.e., when T = 0), with the 112-bit group ID field.

  - P = 1. This multicast address *is* assigned based on the network prefix of the unicast subnet address "owning" the multicast address allocation. This enables allocation of multicast space associated with allocated unicast space for simpler administration. If P = 1, the T bit must also be set to 1. The corresponding format of a multicast packet is shown in Figure 2.10.

    When P = 1, the scope field is followed by 8 zero bits (reserved), an 8-bit prefix length field, and a 64-bit network prefix field and a 32-bit group ID field. The prefix length field represents the prefix length of the associated unicast network

[*] Please refer to http://www.iana.org/assignments/ipv6-multicast-addresses for the latest assignments.

[†] The definition of the P bit is documented in RFC 3306 (16).

address. The network prefix field contains the corresponding unicast network prefix, while the group ID field contains the associated multicast group ID.

For example, if a unicast address of 2001:DB8:B7::/48 is allocated to a subnet, a corresponding unicast-based multicast address would be of the form FF3*s*:0030:2001:DB8:B7::*g*, where

- FF = multicast prefix.
- 3 = $[0011]_2$, that is, P = 1 and T = 1.
- *s* = a valid scope as we'll define in the next section.
- 00 = reserved bits.
- 30 = prefix length in hex = $[0011\ 0000]_2$ = 48 in decimal, the prefix length in our example.
- 2001:DB8:B7:0 = 2001:0DB8:00B7:0000 = 48-bit network prefix in the 64-bit network prefix field.
- *g* = a 32-bit group ID.

A special case of this format occurs with P = T = 1 when the prefix length field = FF and $s \leq 2$. In this case, instead of the network prefix field consisting of the unicast network address, this field will be comprised of the interface ID of the respective interface. The interface ID used must have passed the duplicate address detection (DAD) process, which is discussed later in this chapter, to assure its uniqueness. In this special case, the scope field must be 0, 1, or 2, meaning of interface local or of link local scope. This *link-scoped multicast address* format is defined as an extension of the IPv6 addressing architecture via RFC 4489 (17).

- The R bit within the flags field enables specification of a multicast rendezvous point (RP) that enables multicast group would-be subscribers to link in temporarily prior to joining the group permanently. If the R bit is set to 1, the P and T bits must also be set to 1. When R = 1, the multicast address is based on a unicast prefix, but the RP interface ID is also specified (Figure 2.11). The format of the multicast address when R = 1 is identical to the case when R = 0 and P = 1 with the exception that the reserved field is split into a 4-bit reserved field and a 4-bit rendezvous point interface ID (RIID) field.
    - The IP address of the RP is identified by concatenating the network prefix of corresponding prefix length with the value of the RIID field. For example, if an RP on the [unicast] network is 2001:DB8:B7::6, the associated multicast address would be FF7*s*:0630:2001:DB8:B7:*g*, where *s* = a valid scope defined below and *g* = a 32-bit group ID.

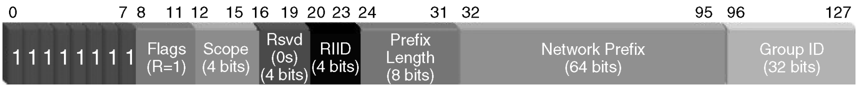

Figure 2.11. Multicast address with flag R = 1.

  - The explicit breakdown of this address is as follows:
    - FF = multicast prefix.
    - 7 = $[0111]_2$, that is, R = 1, P = 1, and T = 1.
    - s = a valid scope defined below.
    - 0 = reserved bits.
    - 6 = RIID field, to be appended to the network prefix field.
    - 30 = prefix length in hex = $[0011\ 0000]_2$ = 48 in decimal, the prefix length in our example.
    - 2001:DB8:B7:0 = 2001:0DB8:00B7:0000 = 48-bit network prefix in the 64-bit network prefix field.
    - g = a 32-bit group ID.
- The first flag bit is reserved and is set to 0.

Multicast Flags Summary. Who thought multicast addressing could be so complicated? But as is typically the case, with complexity comes flexibility! To summarize, the net result of the above bit stipulations yields the following valid values of the flags field as currently defined. Since the flags field immediately follows the first eight "1" bits, we denote the "effective prefix" of these first eight bits followed by the valid 4-bit flags field (Table 2.2).

TABLE 2.2. Multicast Flags Summary

Flags (Binary)	Effective Prefix	Interpretation
0000	FF00::/12	Permanently assigned 112-bit group ID scoped by 4-bit scope field
0001	FF10::/12	Temporarily assigned 112-bit group ID scoped by 4-bit scope field
0011	FF30::/12	Temporarily assigned unicast prefix-based multicast address
0111	FF70::/12	Temporarily assigned unicast prefix-based multicast address with rendezvous point interface ID
All other flags values	–	Undefined

Scope. The scope field identifies, naturally enough, the scope or "reach" of the multicast address. This is used by routers along the multicast path to constrain the reach of the multicast communications with the corresponding scope. Note that scopes other than interface local, link local, and global must be administratively defined within the routers serving the given scope in order to enforce the corresponding reach constraint. Table 2.3 summarizes valid scope values.

TABLE 2.3. Multicast Scope Field Interpretation

Scope Field			
Binary	Hex	Meaning (Scope)	Description
0000	0	Reserved	Reserved
0001	1	Interface local	Scope consists of a single interface on a node and is useful only for loopback transmission
0010	2	Link local	Scope is only the link on which the multicast packet is transmitted
0011	3	Reserved	Reserved
0100	4	Admin local	Scope is limited to the smallest scope administratively configured. This is not based on physical connectivity or other multicast-related configuration
0101	5	Site local	Scope is limited to the site as administratively defined
0110–0111	6–7	Unassigned	N/A
1000	8	Organization local	Scope consists of multiple sites within one organizational entity as administratively defined
1001–1101	9–D	Unassigned	N/A
1110	E	Global scope	Scope is unlimited
1111	F	Reserved	Reserved

2.2.6 Special Case Multicast Addresses

Solicited Node Multicast Address. One form of multicast address that each node must support is the solicited node multicast address. This address is used during the duplicate address detection phase of address autoconfiguration and for the neighbor discovery protocol, which enables identification of IPv6 nodes on a link. The solicited node multicast address is formed by appending the low-order (rightmost) 24 bits of the solicited node's interface ID to the well-known FF02::1:FF00/104 prefix.

For example, let's say a node wishes to resolve the link layer address of the device (interface) with IP address 2001:DB8:4E:2A:3001:FA81:95D0:2CD1. Using the low-order 24 bits, D02CD1 in hex, the device would address its request to FF02::1:FFD0:2CD1 (Figure 2.12).

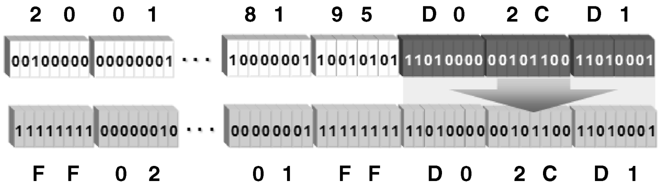

Figure 2.12. Solicited node multicast address derivation (12).

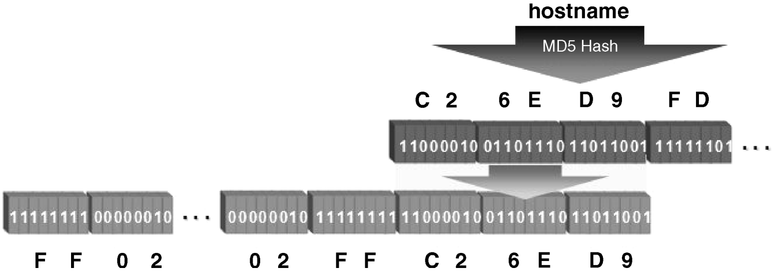

Figure 2.13. Solicited node information query address.

Node Information Query Address. The node information query address is a multicast address enabling solicitation of hostname and IPv6 and IPv4 address information from an IPv6 host (Figure 2.13). If you think this sounds like an overlap with what DNS already provides, you're correct. However, according to RFC 4620 (18), this mode of resolution "is currently limited to diagnostic and debugging tools and network management." And instead of querying a DNS server for this information, a query is issued to the node information query address.

Use of this multicast address format enables an IPv6 address to be formed based only on the hostname of the intended recipient; if the IPv6 address is already known and hostname information is requested, the IPv6 address itself may be used as the destination address. When IP address information is being requested for a known hostname, the canonical hostname[†] is hashed using the 128-bit MD-5 algorithm, and the first 24 bits resulting from the hash are appended to the FF02::2:FF00:0/104 prefix. Each node receiving a message addressed to this node information query address compares the last 24 bits in the address with the first 24 bits of a hash of its own hostname; if it matches, the recipient will reply with the requested information.

2.2.7 IPv6 Addresses with Embedded IPv4 Addresses

We will discuss IPv4 to IPv6 migration and coexistence strategies in Chapter 15, but we'll introduce the IPv4-mapped IPv6 address here (Figure 2.14). This type of address is not routable on the Internet, and is used solely by some translation schemes, and should not generally be used within an IPv6 packet on a communications link. This address

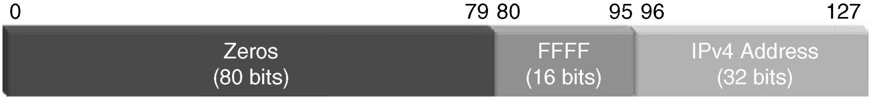

Figure 2.14. IPv4-mapped IPv6 address (12).

[†] The "canonical hostname" is technically the first "label" in the fully qualified domain name in lowercase characters. This terminology is described in detail in Chapter 9 but suffice it to say that this generally is the intended destination hostname.

format consists of 80 zero bits, followed by 16 one bits, followed by the 32-bit IPv4 address.

This address notation combines the familiar IPv4 dotted decimal format appended to the specified IPv6 prefix. Thus, an IPv4-mapped IPv6 address for 172.16.20.5 would be represented as : :FFFF:172.16.20.5.

2.3 IPv6 ADDRESS AUTOCONFIGURATION

One of the advertised benefits of IPv6 is the ability for devices to automatically configure their own IPv6 address that will be unique and relevant to the subnet to which it is presently connecting.[‡] There are three basic forms of IPv6 address autoconfiguration:

- *Stateless.* This process is "stateless" in that it is not dependent on the state or availability of external assignment mechanisms, for example, Dynamic Host Configuration Protocol for IPv6 (DHCPv6). The device attempts to configure its own IPv6 address(es) without external or user intervention.
- *Stateful.* The stateful process relies solely on external address assignment mechanism such as DHCPv6. The DHCPv6 server would assign the 128-bit IPv6 address to the device in a manner similar to DHCP for IPv4 operation. This process will be described in detail in Chapter 5.
- *Combination of Stateless and Stateful.* This process involves a form of stateless address autoconfiguration used in conjunction with stateful configuration of additional IP parameters. This commonly entails a device autoconfiguring an IPv6 address using the stateless method, and then utilizing DHCPv6 to obtain additional parameters or options such as which NTP servers to query for time resolution on the given network.

At the most basic level, the autoconfiguration of an IPv6 unicast address involves concatenating the address of the network to which the device is connected (where you are) and the device's interface ID (who you are). Let's first consider how the device determines the address of the network to which it is connected.

2.4 NEIGHBOR DISCOVERY

The process of *neighbor discovery* in IPv6 enables a node to discover the IPv6 subnet address on which it is connected. Neighbor discovery in general also enables identification of other IPv6 nodes on the subnet, to identify their link layer addresses, to discover routers serving the subnet, and to perform duplicate address detection. Discovery of routers enables IPv6 nodes to automatically identify routers on the subnet, negating the need to configure a default gateway manually within the device's IP

[‡] Note that some IPv4 protocol stacks, such as those provided with Microsoft Windows 2000 and XP, among others, perform address autoconfiguration utilizing the IPv4 "link local" address space, 169.254.0.0/16.

configuration. This discovery enables a device to identify the network prefix(es) and corresponding prefix length(s) assigned to the link.

The discovery process entails each router periodically sending advertisements on each of its configured subnets indicating its IP address, its ability to provide default gateway functionality, its link layer address, the network prefix(es) served on the link including corresponding prefix length and valid address lifetime, as well as other configuration parameters.

The router advertisement also indicates whether a DHCPv6 server is available for address assignment or other configuration. The M bit (managed address configuration flag) in the router advertisement indicates that DHCPv6 services are available for address and configuration settings. The O bit (other configuration flag) indicates that configuration parameters other than the IP address are available via DHCPv6; such information may include which DNS servers to query for devices on this link. Nodes can also solicit router advertisements using router solicitation messages, addressed to the link local routers multicast address (FF02::2).

2.4.1 Modified EUI-64 Interface Identifiers

Once a node identifies the subnet to which it is attached, it may complete the address autoconfiguration process by formulating its interface ID. The IPv6 addressing architecture stipulates that all unicast IPv6 addresses, other than those beginning with binary $[000]_2$, must use a 64-bit interface ID derived using the modified EUI-64 algorithm. The "unmodified" EUI-64 algorithm entails concatenating the 24-bit company identifier issued by the IEEE to each network interface hardware manufacturer (e.g., the initial 24 bits of an Ethernet address) with a 40-bit extension identifier. For 48-bit Ethernet addresses, the company identifier portion of the Ethernet address (first 24 bits) is followed by a 16-bit EUI label, defined as hexadecimal FFFE, followed by the 24-bit extension identifier, that is, the remaining 24 bits of the Ethernet address.

The modification required to convert an unmodified into a modified EUI-64 identifier calls for inverting the "u" bit (universal/local bit) of the company identifier field. The "u" bit is the seventh most significant bit in the company identifier field. Thus, the algorithm for a 48-bit MAC address is to invert the "u" bit and insert the hexadecimal value FFFE between the company identifier and the interface identifier. This is illustrated in Figure 2.15 using a MAC address of AC-62-E8-49-5F-62. The resulting interface ID is AE62:E8FF:FE49:5F62.

For non-Ethernet MAC addresses, the algorithm calls for use of the link layer address as the interface ID, with zero padding (from the "left"). For cases where no link layer address is available, for example, on a dial-up link, a unique identifier utilizing another interface address, a serial number, or other device-specific identifier is recommended.

The interface ID may not be unique, especially if not derived from a unique 48-bit MAC address. Thus, the device must perform duplicate address detection prior to committing the new address. Prior to completing the DAD process, the address is considered tentative.

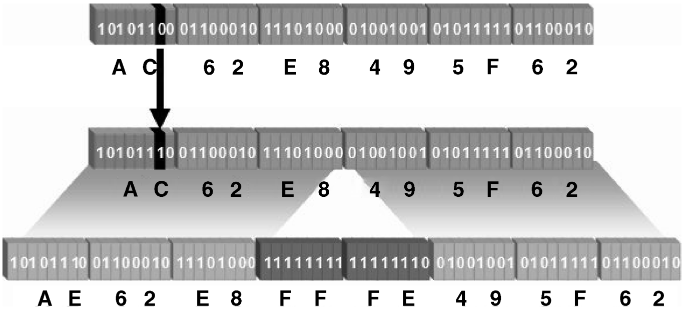

Figure 2.15. Modified EUI-64 interface ID example (11).

2.4.2 Duplication Address Detection

DAD is performed using the neighbor discovery process, which entails the device sending an IPv6 Neighbor Solicitation packet to the IPv6 address it just derived (or obtained from DHCPv6) in order to identify a preexisting occupant of the IP address. After a slight delay, the device also sends a Neighbor Solicitation packet to the solicited node multicast address associated with this address.

If another device is already using the IP address, it will respond with a Neighbor Advertisement packet, and the autoconfiguration process will stop; that is, manual intervention or configuration of the device to use an alternate interface ID is required. If a Neighbor Advertisement packet is not received, the device can assume uniqueness of the address and assign it to the corresponding interface. Participation in this process of Neighbor Solicitation and Advertisement is required not only for autoconfigured addresses but also for those statically defined or obtained through DHCPv6.

IPv6 addresses have a lifetime during which they are valid (Figure 2.16). In some cases the lifetime is infinite, but the concept of address lifetime applies to both DHCPv6 leased addresses and autoconfigured addresses. This is useful in easing the process of network renumbering. Routers are configured with and advertise a preferred lifetime and a valid lifetime value for each network prefix in their Router Advertisement messages. IP addresses that have successfully proven unique through the duplicate address detection process described above can be considered either preferred or deprecated. In either state, the address is valid, but this differentiation provides a means for upper layer protocols (e.g., TCP, UDP) to select an IP address that will likely not change during the ensuing session.

A device refreshes the preferred and valid times with each Router Advertisement message in accordance with the values advertised. When time expires on a preferred prefix, the associated address(es) will become deprecated, though still valid. Thus, the deprecated state provides a transition period during which the address is still functional but should not be used to initiate new communications. Once the valid lifetime of the address expires, the address is no longer valid for use. Should a subnet be reassigned a different network prefix, the router can be configured to advertise the new prefix, and

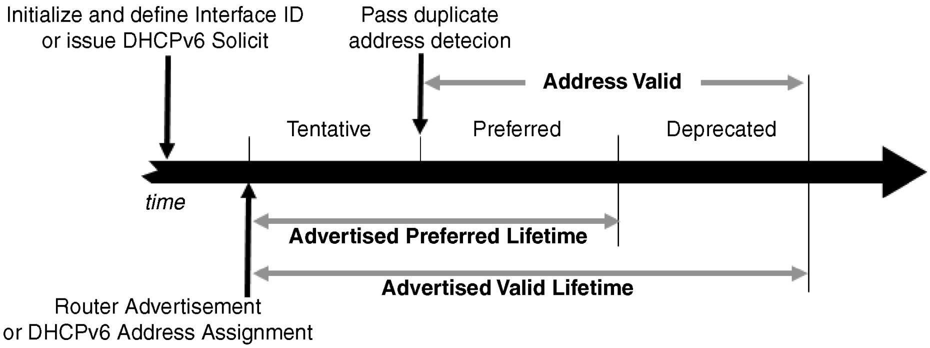

Figure 2.16. IPv6 address lifetimes (figure based on Ref. 19).

devices on the network would undergo the autoconfiguration process using the new prefix as the lifetime of the old prefix expires.

2.5 RESERVED SUBNET ANYCAST ADDRESSES

RFC 2526 (20) defines the format for reserved subnet anycast addresses. These addresses are used by IPv6 devices to route packets to the nearest device of a particular type on a specified subnet. For example, a reserved subnet anycast address can be used to send packets to the nearest mobile IPv6 home agent on a specified subnet. Since the global routing prefix and subnet ID are specified within this address type, it enables a node to locate the nearest node of the desired type on that subnet.

The format of the address takes on one of two forms based on whether the subnet prefix requires formulation of the interface ID field in modified EUI-64 format. Recall that all global unicast addresses other than those beginning with $[000]_2$ must utilize 64-bit interface IDs formulated based on the interface's link layer address and the modified EUI-64 algorithm described previously.

1. If the EUI-64 algorithm is required, the reserved subnet anycast address is formulated by concatenating the following fields (Figure 2.17):
   - 64-bit global routing prefix and subnet ID.
   - 57 bits of all 1s, except the seventh bit in this sequence (the 71^{st} bit from the beginning, counting left to right), which is 0. This seventh bit corresponds to the "u" bit (universal/local bit) of the company identifier field in the hardware

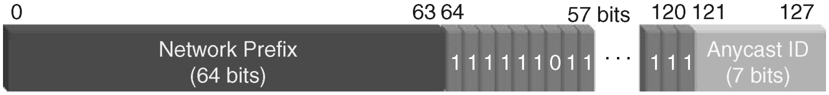

Figure 2.17. Reserved subnet anycast address format when EUI-64 is required (20).

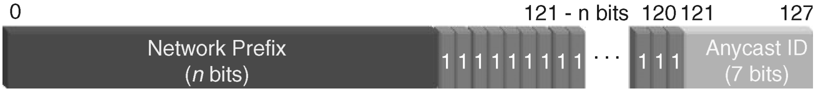

Figure 2.18. Reserved subnet anycast address format when EUI-64 is not required (20).

address when applying the EUI-64 algorithm. This bit is always zero in this particular scenario to represent the "local" setting of the bit.

- 7-bit anycast ID. RFC 2526 defines a single anycast ID of hex 7E for mobile IPv6 home agent anycast. Other anycast ID values are reserved, though IANA may assign additional anycast IDs based on future IETF RFC publications.

2. If EUI-64 is *not* required based on the global routing prefix and subnet ID, then the network prefix length is arbitrary at n bits, followed by $121 - n$ 1 bits, followed by the 7-bit anycast ID (Figure 2.18).

2.6 REQUIRED HOST IPv6 ADDRESSES

RFC 4294 (21) summarizes the requirements for IPv6 nodes, a device that implements IPv6, and for IPv6 routers. In terms of required addresses, all IPv6 nodes must be capable of recognizing the following IPv6 addresses for itself:

- The loopback address (::1).
- Its link local unicast address (FE80::<interface ID> as configured via autoconfiguration).
- The all-nodes multicast address (FF0*s*::1, where $s =$ scope).
- Unicast and anycast addresses configured automatically or manually on each interface.
- The solicited node multicast address for each of its unicast and anycast addresses.
- Multicast addresses for each multicast group to which the node belongs.

A router node is required to support the above addresses plus the following addresses:

- The subnet router anycast address (<subnetwork prefix>::/128, that is, interface ID = 0s).
- The all-routers multicast address (FF0*s*::2, where $s =$ scope).
- Anycast addresses configured on the router.

Other device types such as DHCP and DNS servers must recognize scoped multicast addresses corresponding to group IDs assigned by IANA (i.e., when flags = 0).