Society of Cable
Telecommunications
Engineers

SCTE Operational Practice Series

IPv6 Deployment Best Practices: Fundamentals

STANDARDS

First Edition

Prepared by the SCTE IPv6 Working Group

www.scte.org

Society of Cable Telecommunications Engineers

SCTE Operational Practice Series

IPv6 Deployment Best Practices:

Fundamentals

Prepared by the SCTE IPv6 Working Group

Date of Issue: September 2014

140 Philips Road, Exton, PA 19341

ACKNOWLEDGEMENTS

SCTE would like to acknowledge the work of the SCTE IPv6 Working Group and Comcast for their assistance in the creation of this document.

TABLE OF CONTENTS

INTRODUCTION AND BACKGROUND

Understanding the requirements for IPv6 functionality expressed in the various sets of Data Over Cable Service Interface Specification (DOCSIS) standards from CableLabs helps to create a strong foundation of IPv6 knowledge. Incorporating support for broad use of IPv6 was one of the key motivations for the creation of the DOCSIS 3.0 set of specifications. As such, the DOCSIS 3.0 standards describe both required and optional IPv6-related functionality for cable modems and for CMTS platforms. When CableLabs initially released its DOCSIS 2.0 specifications, IPv6 had already been defined for several years but relatively little pressure in the form of IPv4 address space exhaustion existed. For this reason, the initial DOCSIS 2.0 specifications did not encompass use of IPv6 on CMTS systems or cable modems. When the need to manage DOCSIS 2.0 cable modems with IPv6 became more apparent, CableLabs created the DOCSIS 2.0 + IPv6 Cable Modem Specification beginning in 2009. This document places requirements on cable modems for their management via IPv6 and for use of IPv6 in the CPE space (attached to or integrated within the modem).

MOTIVATION TO ADOPT IPV6

As operators under American Registry of Internet Number, or ARIN, we need to be prepared for growth in North America. The IPv4 pool is dry, operators need additional address space to scale networks for the future. IPv6 provides more address space than IPv4. IPv6 is a 128 bit/16 bytes or 32 hexadecimal character address as compare to IPv4 which is a 32 bit or 4 bytes decimal address. 96 more bits than IPv4! How big is IPv6 you ask? In IPv6, there are 340 trillion, trillion, trillion addresses or 10x10^38 addresses. That's 340,282,366,920,938,463,463,374,607,431,768,211,456 addresses.

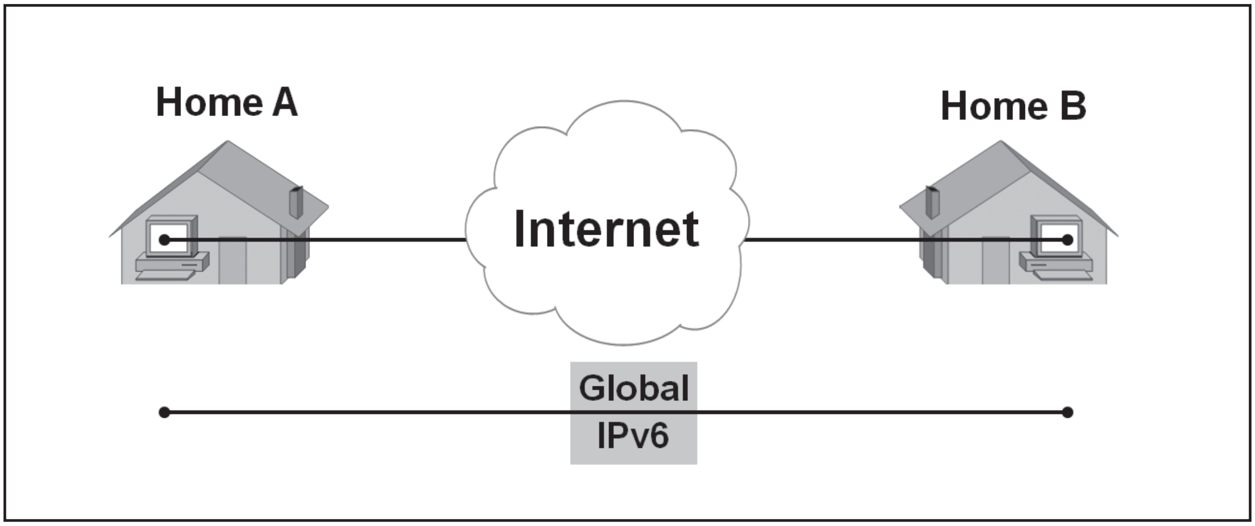


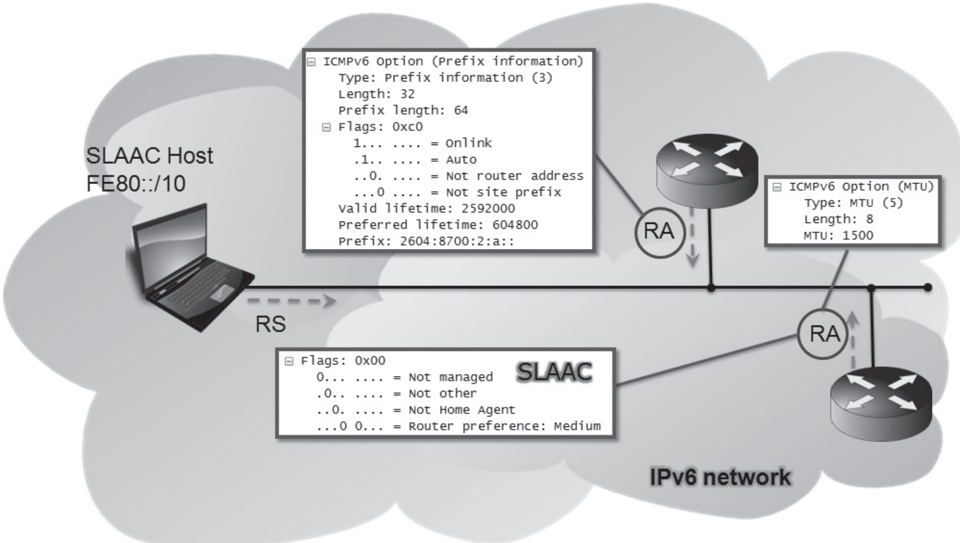
Figure 1: NAT not required for IPv6

A larger address pool means no more Network Address Translation, or NAT, to slow the consumption of IP addresses. Native IPv6 without NAT is the goal of future communications in a global transparent network. IPv6 requires less infrastructure and allows renumbering with the use of Unique Local Addressing (ULA) or Global Unicast Addressing (GUI). IPv6 allows coherent end-to-end packet delivery. IPv6 improves the potential for use of end-to-end security tools for encryption and authentication. IPv6 allows for widespread deployment of peer-to-peer applications like Session Initialization Protocol (SIP) and IP Multimedia Subsystem (IMS).

Features of IPv6

For cable operators to unlock the features of IPv6, nodes must operate in IPv6. When core routers, RANs, CMTSs and OSS run dual stack and communicate using IPv6 interior gateway protocols (IGPs), we can realize the full potential of IPv6. More IPv6 in the infrastructure means more testing of native IPv6.

Hierarchical addressing will improve IPv6 address aggregation. A redesigned eight (8) field header with less fields than IPv4's fourteen (14) field header, will support improved routing. Neighbor Discovery (ND) protocol replaces the broadcast ARP with multicast communication. Auto-configuration features allow clients to configure IPv6 address information without the use of Dynamic Host Configuration Protocol (DHCP), using ND protocol. Features such as Internet Protocol Security (IPSec), added security mechanisms like SEcure Neighbor Discovery (SEND) and encryption features will increase network security. Mobility for IPv6 tackles issues with Mobile IPv4 (MIPv4), like triangular routing. The addition of jumbograms provides a larger datagram size or network layer packet. The packet can be: 2^32 power = 4,294,967,295 bytes = 4.2GB. Since nodes are able to perform Path Maximum Transmission Unit (PMTU), it adds a benefit for moving large IP video datagrams. IPv6 works with scopes instead of using address classes. Scopes limited the area packets travel within the network.



```
⊟ ICMPv6 Option (Prefix information)
    Type: Prefix information (3)
    Length: 32
    Prefix length: 64
  ⊟ Flags: 0xc0
      1... .... = Onlink
      .1.. .... = Auto
      ..0. .... = Not router address
      ...0 .... = Not site prefix
    Valid lifetime: 2592000
    Preferred lifetime: 604800
    Prefix: 2604:8700:2:a::
```

```
⊟ ICMPv6 Option (MTU)
    Type: MTU (5)
    Length: 8
    MTU: 1500
```

```
⊟ Flags: 0x00
    0... .... = Not managed
    .0.. .... = Not other
    ..0. .... = Not Home Agent
    ...0 0... = Router preference: Medium
```

Figure 2: Path MTU

In an all IP quad play the challenge of network convergence become an issue. There are different networks providing different services – for instance, a separate network for data, a separate network for voice, a separate network for video and a separate network for mobility. All the networks, when combined together, have a large need in terms of IP address requirements. Individually all these networks have their own private IP address space but when combined together they create a converged network that creates a huge requirement on the number of IP addresses desired for that converged network. Since the MSOs are running out of IPv4 address space, they will not have the capacity to provide IP addresses for all the devices in this big converged network.

However IPv6 can alleviate such a problem. If an ISP is assigned a /32 address space, that leaves 2 raised to the power of 96 addresses or 79,228,162,514,264,337,593,543,950,336 IPs. Every device in the converged network would be able to have an IPv6 global distinct address.

Another concern is industry consolidation where the merger of two different IP networks. In today's world, we are increasingly seeing scenarios where one company acquires the assets of another. In this scenario, there are two different private IP networks belonging to two different companies. If these two companies decide to merge, there will be an address collision because the private addresses which initially were separate now have overlapping addresses. There is a need to create a new address space and renumber the whole network for the merged company.

Renumbering IP address space creates a huge operational complexity for the MSO. Again, if you use IPv6 it would avoid the renumbering process and save the company money. Renumbering of devices is a method related to auto configuration, implemented using DHCP in IPv4. In IPv6, routers can specify an expiration interval for network prefixes when auto configuration is done. Routers then can send a new prefixes to tell devices to regenerate their IP addresses. Devices can maintain deprecated addresses in case a fall back method is necessary.

The other need for IPv6 is the MSOs' requirement to provide future services such as global transparency, the ability to communicate between two devices sitting in different parts of the world without having to worry about if the other address is behind a NAT server or not. In a native IPv6 network, the need for NAT is removed from the IP network.

In Plug and Play (PnP) home networking, the ability to buy an IP device off the shelf, take it home, plug it into the home network without worrying about configuring an IP address and other system parameters such as Domain Name Service (DNS), Default Gateway (DG) or a subnet mask is important.

We also need access transparency, the ability to seamlessly switch between different access networks such as a cellular network, a cable network or a business services network.

Global Transparency

We just introduced the problems MSOs have and how IPv6 can solve these problems. In one use case we have two home networks, Home A and Home B which are connected together by the Internet.

The user of Home A would like to talk to the user of Home B through a Voice over IP (VoIP) call. The networks of both these homes are using private address space. The user of Home A does not know the IP address of the user of Home B and vice versa. The Voice over IP call is not possible in this scenario.

Now if you go to IPv6, the addresses in these homes become global so then the addresses of each of these users are known to one another. A Voice over IP call can very easily be possible with IPv6 in these two homes without going thru a NAT server. IPv6 eliminates the need for NAT, Session Border Controllers (SBC), transition gateways (TrGW) and restores global transparency to the network. Do keep in mind that eliminating NAT does not mean eliminating the firewalls. Eliminating NAT is saying we do not need NAT that arises from the assignment of private IP address space. However, the firewall can still be used with IPv6 to prevent unauthorized access to homes, businesses, etc.

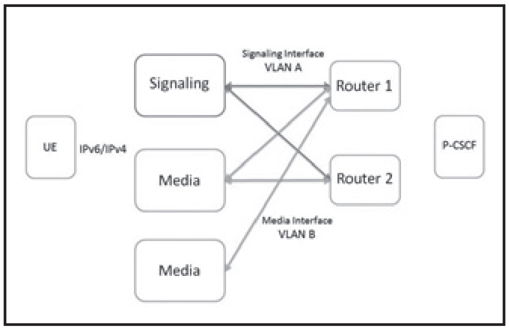


Figure 3: IPv4 to IPv6 example of SBC and TrGW

NAT64, or Network Address Translation IPv6 to IPv4, can be used as an IPv6 migration strategy and should not be confused with standard NAT servers. A NAT64 server has two interfaces, one interface connected to an IPv4 network and one interface connected to an IPv6 network. The NAT64 will take a sender's IPv6 packets and translate them to IPv4 and also make the reverse connection. For IPv4 to IPv6 connections, Interactive Connectivity Establishment (ICE) is needed with NAT64. NAT-PT is depreciated and replaced by NAT64.

PnP Home Networking

Let's talk about how IPv6 enables PnP Home Networking. In future networks we expect to have different types of devices. For instance, networks will have computers, smart phones, home entertainment devices like wireless Blueray, home IP phones, music systems, IP security devices and all these are IP enabled and connected to the broadband network via a home Gateway Router (GWR).

In IPv6, the MSO or cable operator's DHCPv6 server will hand out a prefix such as a /56 (this prefix is not a suggestion) to the customer's home GWR. The GWR will automatically carve out a prefixes (e.g., /64) and assign different addressing to the CPE within the home. For instance a single /64 network supports 2^64 hosts. There are so many addresses that a /64 for a customer is not a big deal in IPv6. There is no need for each of these CPE devices to be manually configured for IP or other IP system parameters. IPv6 has an auto configuration feature beyond DHCP which allows them to just plug into the network, listen to prefixes from the home GWR, and automatically configure themselves for a globally unique IPv6 address. Thus IPv6 offers true PnP in the home network.

The /64 addressing idea is still evolving. For instance, a /120 still provides 2^8 IPs for CPE in the home however interferes with services such a EUI-64. If service provider 1 is assigned a 2001:DB8::/32, then it can create 2^16 or 65,535 - 1 /48 networks. This would work for a small cable operator. Some ISPs have been talking about using a /56 instead of the /48. This would yield 16,777,216 /48s which is the direction large MSOs should take.

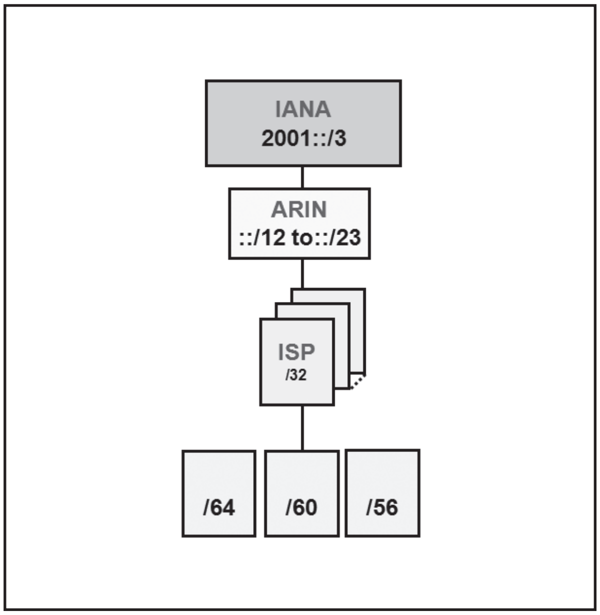


Figure 4: Address Delegation Example

Access Transparency

IPv6 improves mobility across transparent access networks. In the diagram below we see a mobile user using a phone that supports Global System for Mobile (GSM) and Wireless Fidelity (Wi-Fi). This mobile user could be a user driving his or her car and accessing the Internet through a GSM cellular network or making a phone call on the GSM cellular network. When the user enters his or her home, the active session may need to be switched over from the GSM cellular network to the Wi-Fi home base broadband network. Mobile IPv6 enables seamless transition from one access network such as cellular network to another access network such as a broadband network at home without letting the user know that a transition is happening, thus IPv6 enables access transparency.

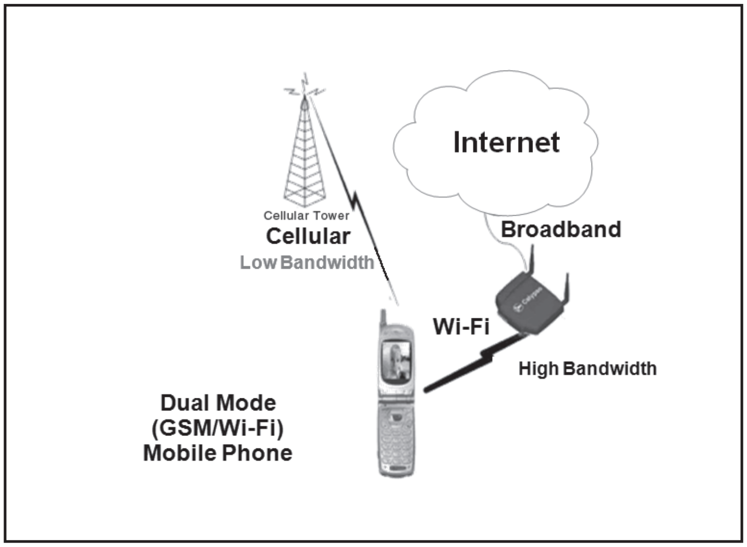


Figure 5: Mobile IPv6

Reinforcing the Need for IPv6

IPv6 allows fairness in IP allocation which was not there in IPv4. There is a strong need for global equitable distribution of IPv6 addresses. IP addresses are provided to those who can demonstrate a need for them in accordance with Regional Internet Registry (RIR) policies. The addresses assigned in IPv4 are not uniform. IPv6 ensures that addressing is assigned in a more uniform manner. Some operating system vendors have made IPv6 the primary networking choice while other operating systems such as major Linux distributions have adopted IPv6.

DOCSIS 3.0 has built in support for IPv6 and the new 'DOCSIS 2.0 + IPv6' standard also supports IPv6, which may only require a firmware upgrade on the cable modem side. Japan and Asia Pacific (APAC) governments are heavily promoting IPv6. Japan has strong motivation for the millions of cell phones and tech-savvy products their population demands. The United States Department of Defense (DoD) mandated that the backbones of all federal agencies support IPv6. This means all routing, path determination and other backbone services must be compliant with IPv6. In addition, the United states, will require federal agencies to run native IPv6 on their HTTP, POP/SMTP, ISP, and domain name servers and services. China, has a country-wide plan to roll out IPv6 since it has exceeded the address requirements of that of the US and will only continue to scale.

ICANN, NRO and IAB announce global transition to IPv6. The Number Resource Organization (NRO) and Internet Corporation for Assigned Names and Numbers (ICANN) and Internet Architecture Board (IAB) held a ceremony to announce the global transition to the next generation of Internet addresses. The Internet Assigned Number Authority (IANA) IPv4 free pool was depleted! RIRs will continue to deplete their pools such as APNIC, and shortly others will follow such as the North American registry ARIN.

As mentioned before, NAT and application layer gateways (ALG) connect IPv4 networks, breaking end-to-end IP communication model. NAT makes fast re-routing difficult and mandates the network keeps connection state. When many devices need reachability from outside, NAT becomes an issue. IPv6 restores global transparency by removing NAT from the network providing end-to-end connectivity.

New provisioning features like SLAAC appear in the new IPv6 protocol providing PnP Home Networking. Access Agnostic and transparency is supported where multiple access networks may be used to deliver IPv6. IPv6 may provide seamless roaming across networks for fixed/mobile convergence and WiMax.

Given all these reasons there is a strong and compelling need to go to IPv6.

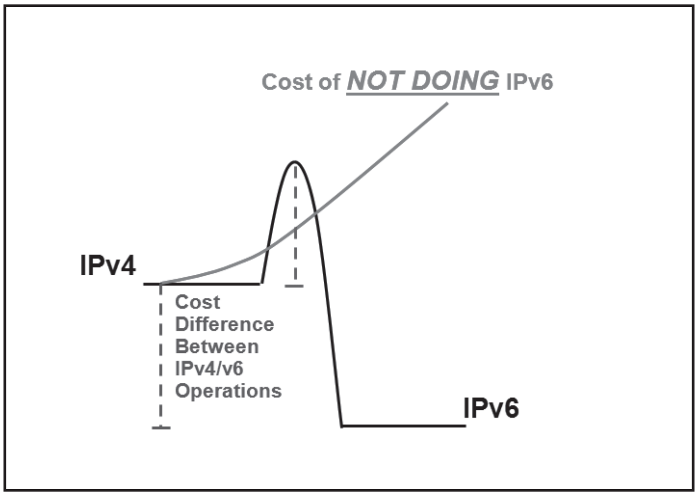


Figure 6: Cost of Not Migrating to IPv6

Can the MSOs afford not working with IPv6? Transition to IPv6 is inevitable! There are some costs associated with transitioning from IPv4 to IPv6, initially. In the long run, the cost to maintain an IPv6 network is considerably less. However if you look at the red line in the figure above, the cost of not doing IPv6 keeps increasing as time goes by. The longer the delay, the higher the cost. The increasing costs of not doing IPv6 can be offset by the long-term reduced cost of maintaining an IPv6 network. It is wise for the MSOs to transition to IPv6, avoiding a loss of potential revenue from new services.

Shortage of IPv4 leads to higher costs due to:

- On-going renumbering

- Loss of global visibility

- Longer deployment cycle

- Impact on Business Continuity

- Degraded Customer experience

BINARY

The IP network mask provides a way for the node to calculate the network ID using bitwise binary ANDing logic. The process uses binary mask and the "AND" logic to determine the network ID.

Bitwise Binary ANDing Logical		
0	0	0
0	1	0
1	0	0
1	1	1

In the example below we see an IPv6 address with a mask of /64 bits. A node will use "AND" logic with the IPv6 and /64 to compute the network ID. In this example, the network ID is 2001:0db8:abcd:1234 before zero compression.

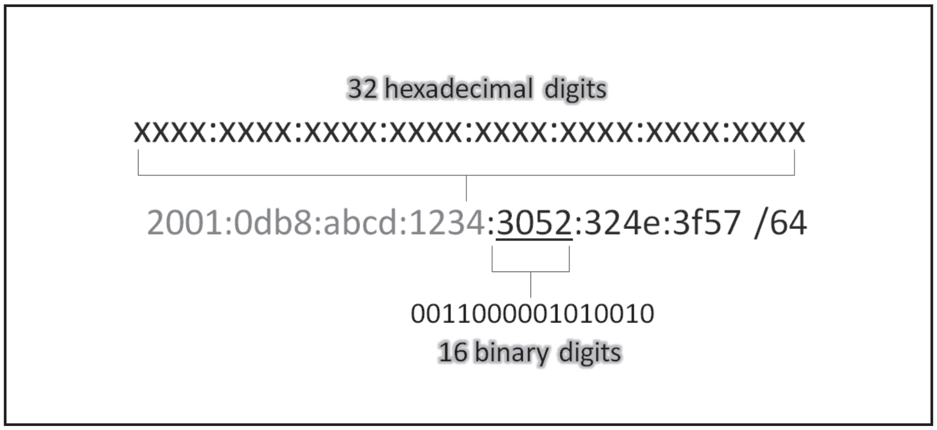

32 hexadecimal digits
xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx
2001:0db8:abcd:1234:3052:324e:3f57 /64
0011000001010010
16 binary digits

Figure 7: IPv6 Addressing with Hexadecimal and Binary

Binary uses positional weighting, the same as with decimal numbers.

For example 205 in decimal would equal:

2 x 100 = 200

0 x 10 = 0

5 x 1 = 5

The same decimal value 205 using binary positional weighting:

11001101 =1x2^7 + 1x2^6 + 0x2^5 + 0x2^4 + 1x2^3 + 1x2^2 + 0x2^1 + 1x2^0

2^7=2 x 2 x 2 x 2 x 2 x 2 x 2 = 128

2^6=2 x 2 x 2 x 2 x 2 x 2 = 64

2^5=2 x 2 x 2 x 2 x 2 = 32

2^4=2 x 2 x 2 x 2 = 16

2^3=2 x 2 x 2 = 8

2^2=2 x 2 = 4

2^1=2

2^0=1

128 + 64 + 8 + 4 + 1 = 205

As shown previously, each section of an IPv6 address is 16 binary bits of 4 hexadecimal digits. A small chart like the one shown below can be used to convert decimal to binary and binary to decimal.

Positional Weighting and Binary			
2^3	2^2	2^1	2^0
8	4	2	1

The chart above will produce 0 to 15 in decimal or 0000 to 1111 in binary.

Example #1:

A decimal 15 is equal to one 8, one 4, one 2, one 1 (1111).

Example #2:

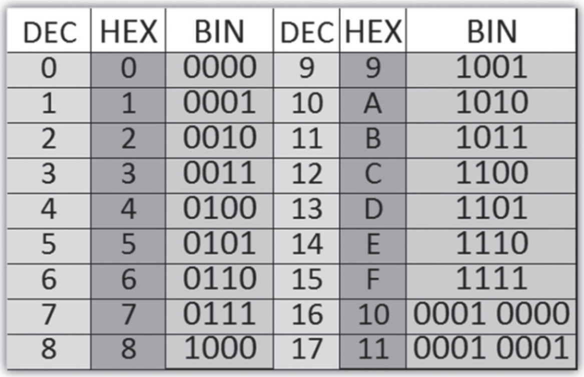
A binary 1100 is equal to one 8, one 4, zero 2s, zero 1s. Also written 8 + 4 = 12.

The chart is the primary tool that makes the process so easy; no math is involved.

A small trick (really the rules of binary) to the chart which adds efficiently during conversion is that the column to the right is always half of the previous column. Another small trick if every column to the right of a column is filled with ones it will be one less than that column. For example: When everything to the right of column 8 is filled with 1s, then the total decimal value is 15.

HEXADECIMAL

Hexadecimal, commonly referred to as “hex”, is base 16 numbering system. IPv6 addresses use hex to shorten the length of the IPv6 addresses. Converting the hexadecimal numeral system to decimal and binary numeral systems is required to fully understand the IPv6 addressing format and how addresses are assigned in the operator’s network. There will be times to assign different lengths of addressing to the various cable system customers. A commercial or business services customer may need a mask of 56 bits and a residential CM may need a mask of 64 bits.



DEC	HEX	BIN	DEC	HEX	BIN
0	0	0000	9	9	1001
1	1	0001	10	A	1010
2	2	0010	11	B	1011
3	3	0011	12	C	1100
4	4	0100	13	D	1101
5	5	0101	14	E	1110
6	6	0110	15	F	1111
7	7	0111	16	10	0001 0000
8	8	1000	17	11	0001 0001

Figure 8: Decimal to Hexadecimal to Binary

Hex uses sixteen distinct symbols to provide more possible combinations taking less memory by using fewer symbols than binary or decimal. Hex is a numeral system that uses symbols (0-9) and (A-F) as shown in the figure above. For example, the symbol D in hex equals a 13 in decimal and 1101 in binary. Where can we find hex numbering? Believe it or not, we have been using it already.

A common use of hex are with Media Access Control (MAC) addresses typical called the Ethernet addresses. Using hex allows a network card manufacture to create 16 million unique network cards under one vendor code or two to the 24th power addresses. They are used to create Wired Equivalent Privacy (WEP) security keys (64 and 128 bit keys) in Wi-Fi communications. Most important, they are also used to create the Internet Protocol Version 6 (IPv6) addressing structure that will be used for networking our cable infrastructures.

The hex numbering system allows us to represent two to the 128th IPv6 possibilities using 32 hex numerals! That is 3.4 times 10 to the 38th power addresses.

Converting hex numbers to binary number system is fairly easy, it's almost trivial. The only requirement is that a person knows the equivalent binary and decimal value of each hexadecimal "digit" (0 to F) shown in the previous chart. A couple of facts about hex:

- Decimal is base 10 which equals 10 symbols / numerals or 0 to 9.
- Hexadecimal is base 16 which equals 16 symbols zero to F.
- Binary is base 2 which equals 2 symbols 0 to 1, referred to as off and on in digital.

To allow for the understanding of more advanced topics in IPv6 addressing, it is important to learn how to convert hex into other numeral systems such as decimal. Converting allows us to create IP subnets of addressing, map IPv6 addresses to layer 2 (L2) MAC addresses and understand the representation of the numbers themselves. Representation of a number could be a route summary or routing prefix.

The first step to converting hex to decimal is to understand base 16 (hex) to base 10 (decimal). The base 16 to base 10 chart will help us convert to decimal. Any valued raised to the power of zero will always equal a 1. Any valued raised to the power of one will always equal itself, 16 in this case. The next value 16 raised to the power of 2 is really 16 x 16, which equals 256. If the chart continued, 16 to the third is really to 16 x 16 x 16 or 4096 in decimal.

Positional Weighting and Hexadecimal			
16^3	16^2	16^1	16^0
4096	256	16	1
0	0	C (12)	5

Let's say we would like to convert 0x0C5. The 0x tells a person this number is a hex value. We drop that from number. The number is plugged into the chart.

Working from right to left. A 5 in the 1s column yields 5 x 1 or 5. A "C" or 12 in the 16 column yields 12 x 16 or 192. A 0 in the 256 column yields 256 x 0 or a zero. A 0 in the 4096 column yields 4096 x 0 or a zero.

Add the 192 and the 5 which yields the decimal value of 197 which is equal to the hexadecimal value 0x0C5.

Most hex conversion only involves 2 digits and should be easier than this example.

Hex to Binary

The first step to converting hex to binary is to understand base 2 (binary) to base 10 (decimal). The base 2 to base 10 chart will help us convert to binary. Remember binary only represents 2 bits a zero and a one.

Just like before…Any valued raised to the power of zero will always equal a 1. Any valued raised to the power of one will always equal itself, 2 in this case. The next value 2 raised to the power of 2 is really 2 x 2, which equals 4. The next value 3 raised to the power of 2 is really 2 x 2 x 2, which equals 8. Since it only takes four bits to represent a hex digit the chart stops. If you raise 2 to the power of 4 bits, it equals 16 possibilities.

A single hex digit is 4 bits!

Let’s say we would like to convert 0xC. Remember the 0x tells us this is a hex value. We drop that from number. We convert C to a 12.

Using the chart from left to right.

Positional Weighting and Binary			
2^3	2^2	2^1	2^0
8	4	2	1
1	1	0	0

How many eights are in 12? One or none? The answer is one.

The twelve now becomes a 4. 12 minus the 8.

How many fours are in 4? One or none? The answer is one.

The four now becomes a 0. 4 minus the 4.

How many twos are in 0? One or none? The answer is zero.

How many ones are in 0? One or none? The answer is zero.

C in hex = 1100 in binary.

INTERNET PROTOCOL VERSION 6

There are only eight fields in the fixed 40 byte IPv6 header. The IP header checksum was removed in IPv6. Now optional layer 4 checksum fields are in the UDP packet. This will allow IPv6 to detect errors using UDP, TCP and data-link protocols. Integrated traffic class and flow labels will provide classification of packets and a more efficient superior Quality of Service (QoS) experience. Header extensions via next header provides more efficient forwarding, future scalability and flexibility and the simplified addition of options.

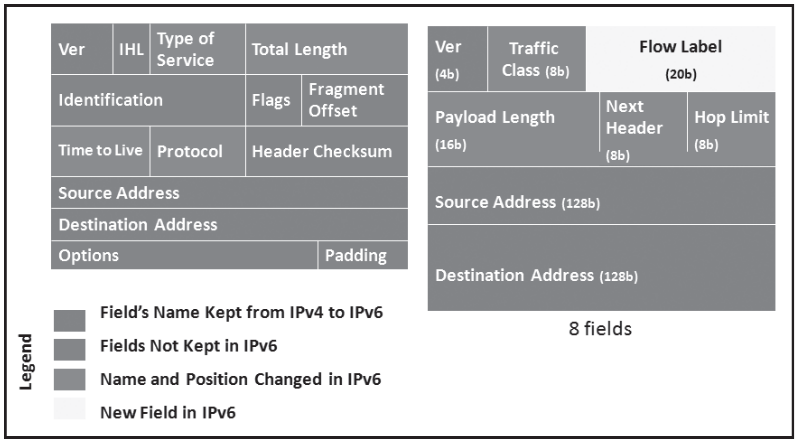


Figure 9: IPv4 to IPv6 Header Comparison

IPv6 Fields	
Version	Identifies the version of IP used to generate the datagram. This field is used the same way as in IPv4, except of course that it carries the value 6 (0110 binary).
Traffic Class	This field replaces the Type of Service (ToS) field in the IPv4 header. It is used not in the original way that the ToS field was defined (with Precedence, D, T and R bits) but using the new Differentiated Services (DS) method defined in RFC 2474. That RFC actually specifies QOS techniques for both IPv4 and IPv6.
Flow Label	This large field was created to provide additional support for real-time datagram delivery and quality of service features. The concept of a flow is defined in RFC 2460 as a sequence of datagrams sent from a source device to one or more destination devices. A unique flow label is used to identify all the datagrams in a particular flow, so that routers between the source and destination all handle them the same way, to help ensure uniformity in how the datagrams in the flow are delivered. For example, if a video stream is being sent across an IP internetwork, the datagrams containing the stream could be identified with a flow label to ensure that they are delivered with minimal latency.
Payload Length	This field replaces the Total Length field from the IPv4 header, but it is used differently. Rather than measuring the length of the whole datagram, it only contains the number of bytes of the payload. However, if extension headers are included, their length is counted here as well.
Next Header	This field replaces the Protocol field and has two uses. When a datagram has extension headers, this field specifies the identity of the first extension header, which is the next header in the datagram. When a datagram has just this “main” header and no extension headers, it serves the same purpose as the old IPv4 Protocol field and has the same values, though new numbers are used for IPv6 versions of common protocols. In this case the “next header” is the header of the upper layer message the IPv6 datagram is carrying.
Hop Limit	This replaces the Time to Live (TTL) field in the IPv4 header; its name better reflects the way that TTL is used in modern networks (since TTL is really used to count hops, not time.) Decremented by 1 by each node that forwards the packet. The packet is discarded if Hop Limit is decremented to zero.
Source Address	The 128-bit IP address of the originator of the datagram. As with IPv4, this is always the device that originally sent the datagram.
Destination Address	The 128-bit IP address of the intended recipient of the datagram; unicast, anycast or multicast. Again, even though devices such as routers may be the intermediate targets of the datagram, this field is always for the ultimate destination.

Encapsulation

Similar to IPv4, delivery of data over IPv6 internetworks is accomplished by encapsulating higher-layer data into IPv6 datagrams. However, encapsulating higher-layer data has been redesigned as part of the overall changes represented by IPv6. Rather than a single base header that contains all fields for the datagram like IPv4, the IPv6 datagram supports a “main” header and then extension headers for additional information as needed. The extension headers via the next header field allow for a great deal of extra information to accompany the IPv6 datagrams as needed.

Optional Internet-layer information is encoded in separate headers that may be placed between the IPv6 header and the upper-layer header in a IPv6 datagram. Since there is no option field in IPv6, next header is now used. The result of using the next header is a fixed length, 40-byte IP header. The IPv6 header will have a base header linked to extension headers via the next header field. The next header field provides the ability for IPv6 packets to link to upper layer protocols, support fragmentation, authentication, encapsulation, hop by hop options, extend routing and support destination options.

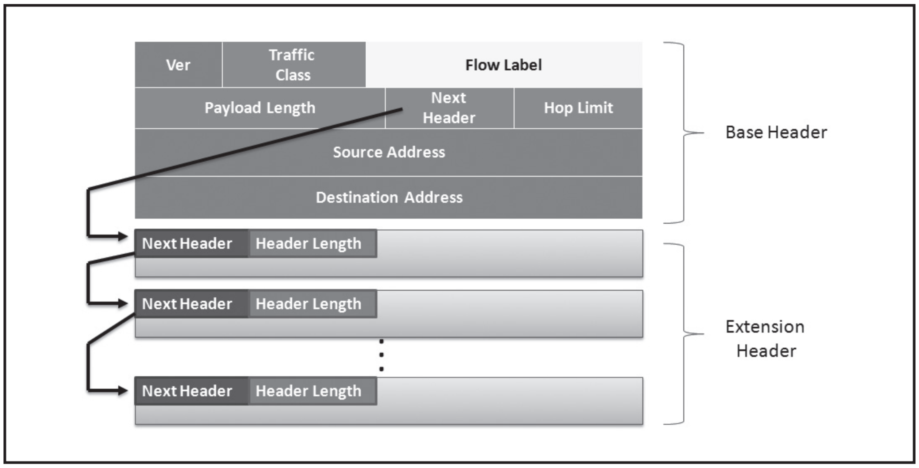


Figure 10: IPv6 Extension Header Diagram

For example, when a IPv6 datagram reaches a destination, the destination checks for a routing header. If there is at least one segment left, that address is copied from the routing header and the packet is forwarded to that address. Otherwise, the routing header is removed and the next routing header is processed.

The Extension Headers have a particular order as recommended by RFC 2460 and that the extension header feature is used no more than once except for the destination options.

IPv6 Extension Header Order	
Header Type	Next Header Value
IPv6 Main Header	-
Hop-by-Hop Options	0
Destination Options	60
Routing	43
Fragment	44
Authentication and Encapsulating Security Payload	51
Upper Layer TCP/UDP	6 / 17
Mobility	135

Options to be processed by the first destination that appears in the IPv6 destination address field plus subsequent destinations listed in the routing header. Extension headers are not examined or processed by any node along a datagram's delivery path. An IPv6 packet can carry as many extension headers as it needs.

Fragmentation

Fragmentation information was moved out of fixed IPv4 fields in base header to an extension header. Similar to v4 fragmentation, the fragment extension header is included in the fragmented IPv6 datagrams via the next header field to provide the information necessary (identification) to allow the fragments to be reassembled by the end host. The extension header is placed between the IPv6 main header and payload when the extension header is used. As with IPv4, IPv6 source which is responsible for fragmentation arranges for destination (end host) to perform re-assembly – changes were made that avoid fragmentation by routers.

The IPv6 source may use a guaranteed minimum maximum transmission unit (MTU) of 1280 bytes or PMTU discovery to identify minimum MTU along path and take advantage of paths with PMTU greater than the IPv6 minimum link MTU.

IPv6 fragmentation is end to end, IPv6 allows fragmentation of any datagram that is too large for the MTU of network over which the datagram must travel. For example a sending host will send a packet using an MTU size of 1500 bytes (octets). If the MTU size was too large the end host will request the minimum MTU size of 1280 bytes (octets). Finally the IPv6 source which is responsible for fragmentation adjusts the Path MTU to 1280 bytes (octets) and continues transmission of the datagrams. A critical component to fragmentation process is the ICMPv6, used for communication when packets exceed MTU size and packets are dropped on the network.

Each fragment is routed independently and must be a multiple of 8 bytes or octets. Here are the fields of the extension header for fragmentation.

Extension Header Fragmentation Fields

Field	Size	Description
Next Header	1 byte	Contains the protocol number of the next header after the fragment header. Used to link headers together as described above.
Reserved	1 byte	Not used, set to zero.
Fragment Offset	13 bits	Similar to the IPv4 fragmentation field, it specifies the offset, or position, in the overall message where the data in this fragment appears.
Reserved	2 bits	Not used, set to zero.
M Flag	1 bit	Similar to the IPv4 fragment field flag, set to 0, it is the last fragment mark; set to 1 indicates more fragments.
Identification	4 bytes / 32 bits	Similar to the IPv4, a larger field with a specific value that is common to each fragment, ensure fragments are not mixed together.

IPv6 Address Structure

IPv6 looks much different than IPv4. The IPv6 128-bit address is 39 characters long including colons with 8 sections. The IPv6 128 bit address is divided along boundaries of 16 bits. Each 16 bit section is equivalent to a 4 digit hex value. The resulting representation is called colon-hexadecimal or hextet. This is in contrast to the 32 bit IPv4 address represented in dotted-decimal format or octet, divided along 8 bit sections, and then converted to its decimal equivalent, separated by periods.

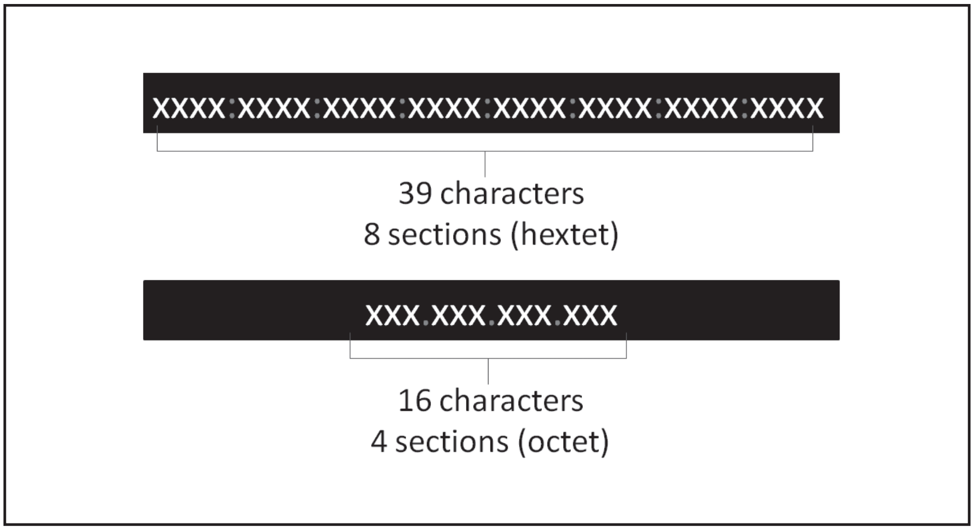


Figure 11: IPv6 address vs. IPv4 address

In the figure below are the sections of an IPv6 unicast address. This is not a standard for cable operators but it is the most widely used approach. In the first place, an IPv6 unicast address breaks down into 4 parts: the Global routing prefix, the cable operator subnet ID, cable operator/customer subnet and the interface ID. The structure below allows a cable operator to have a potential of up to 32 bits for subnetting IPv6 networks.

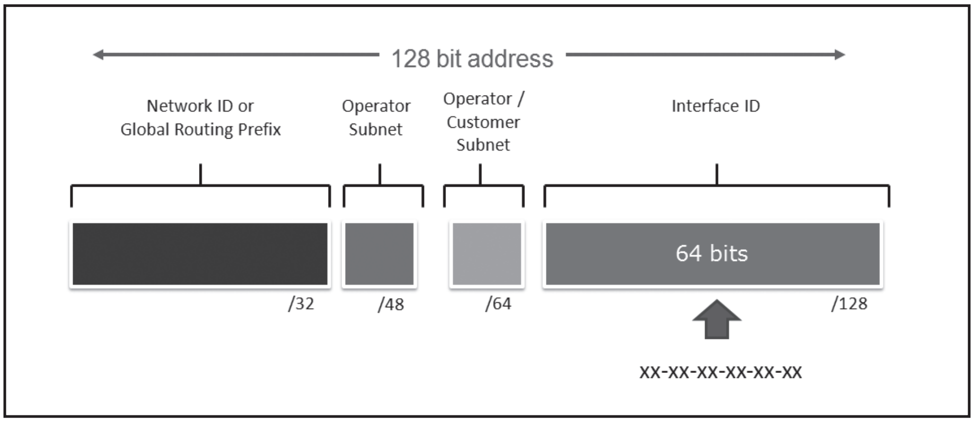
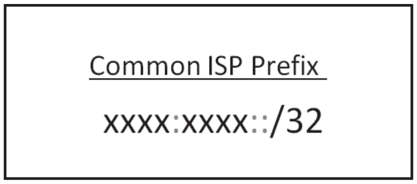


Figure 12: IPv6 Address Format

The trailing interface ID is like the host part of an IPv4 address, used to differentiate multiple hosts on the same IPv6 subnet. It can be arbitrarily long but most often used and recommended are 64 bits because that is when it will be used to effectively carry a MAC address embedded in an IPv6 address.

The other parts of the address, the Global Routing Prefix and the subnet ID, are together equivalent to the first part of an IPv4 address, the network part. This part is used in routing to an address, an actual IPv6 subnet or local area network. The figure shows a breakdown of the addresses typically used and the recommended prefix length for the individual fields as far as the routing prefix goes.

The high-order bits, or the bits that begin an IPv6 address, specify the network ID, the rest of the bits specify particular subnets in that network and interface ID. Thus all the addresses in one network have the same first high order bits. Those first high order bits are called the "prefix." A common subnet prefix will have the following format:



Common ISP Prefix

xxxx:xxxx::/32

Figure 13: IPv6 Prefix

As defined in the Internet Engineering Task Force (IETF) Request for Comment (RFC), the operator has 16 bits at their disposal for creating subnets, or 32 bits when the customer subnet is used. Since these 16 bits are binary values, the range of possible values using the exponentiation of 2^16 will yield 65,536 – 1 subnets under a particular high order set of bits or network prefix. An operator subnet prefix will have the following format if only using 16 bits are used to subnet, xxxx:xxxx:xxxx::/48.

IPv6 was original designed for the customer to have 16 bits at their disposal for creating subnets as defined in the IETF RFC. Operators ultimately have control of how many bits a customer will be able to utilize, for example a residential customer may be provisioned without subnets. In this case the customer will have zero bits to subnet with, or a single subnet using the format xxxx:xxxx:xxxx:xxxx::/64. However, since these 16 bits are binary values, the range of possible values using the exponentiation of 2^16 will yield 65,536 - 1 subnets under a particular high order set of bits or network prefix.

In order for IPv6 features like SLAAC to function correctly the last 64 bits, or /64, are recommended for the interface ID. An interface ID of 64 bits, has the range of possible values using the exponentiation 2^64 =18,446,744,073,709,551,616 host IPs.

Another important aspect of IPv6 addressing is how are the addresses allocated? The allocation process was recently updated by the registries:

- IANA allocates from 2001::/16 to regional registries
- Each regional registry allocation is a ::/23
- ISP allocations from the regional registry is a ::/36 (immediate allocation) or ::/32 (initial allocation) or shorter with justification
- Policy expectation that an ISP allocates a ::/48 prefix to each customer

Zero Compression

For the IPv6 addresses with long sequences of zeros a technique called zero compression can be used to reduce the length of the IPv6 address. The 128-bit IPv6 address can be reduced using two rules:

- Rule one: Leading zeroes within a 16-bit value may be omitted. For example, the address in the figure may reduce each hextets of zeros as shown below.
- Rule two: A single occurrence of consecutive groups of zeroes within an address may be replaced by a double colon as shown below.

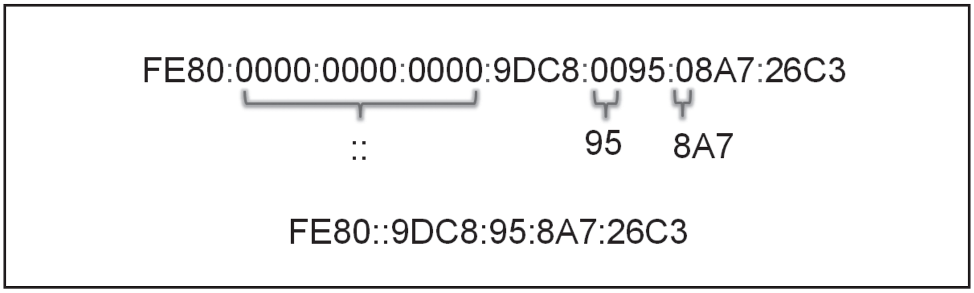


Figure 14: Zero Compression

IPv6 Address Types

IPv6 addresses are assigned to interfaces. Interfaces are expected to have multiple addresses in IPv6, this thinking has changed from IPv4 model. Addresses have a scope, or reach, in the network. IPv6 supports several types of address shown below in the table.

IPv6 Addresses		
Type of IPv6 Address	**Short Description**	**Address Example**
Unicast	One address on a single interface with delivery to single interface.	
Multicast	Address of a set of interfaces with delivery to all interfaces in the set.	FF00::/8 (1111 1111) is for multicast.
Anycast	Address of a set of interfaces with delivery to a single interface in the set.	
Loopbacks	Required and are assigned to single interfaces on a node.	::1
Link-Local Addresses (LLA)	Required on all interfaces on a link, similar to IPv4 loopbacks. Link- Local is a segment of the network separated by routers, similar to the broadcast domain idea in IPv4.	FE80::/10 is the prefix for link-local addresses.

Type of IPv6 Address	Short Description	Address Example
Unique Local Addresses (ULA)	Are used similar to RFC 1918. ULAs are private addresses dedicated to a site.	FC00::/7 is for unicast-local and FF00::/8 is for multicast addresses.
Global Unicast Addressing (GUA)	Public addresses. IPv6 unicast address global space is similar to a public address in IPv4.	2000::/3
Auto configured 6 to 4 tunnels	Only if IPv6 public addresses are available.	
Auto-configured IPv4	Highly discouraged.	
Solicited Node Multicast	Required for ND. Formed by taking the low-order 24 bits of an address (unicast or any cast) and appending those bits to the prefix FF02:0:0:0:0:1:FF00::/104	
All Node Multicast	Globally anonymous and globally published.	
Unspecified Address	Address of all zeroes and is used very similar to the equivalent address 0.0.0.0 in IPv4. Used as a placeholder when no address is available, for example with an initial DHCPv6 request or DAD.	::/8

Other values (*approximately 7/8 of total) are currently unassigned and further details are available at the Internet Assigned Numbers Authority (IANA) site.

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

MIGRATION TECHNIQUES

A wide range of techniques have been identified and may be implemented, basically falling into four categories:

1. Dual-stack techniques, to allow IPv4 and IPv6 to co-exist on the same devices, CMTS and core/aggregate network routers and OSS/BSS. However, dual stack may not be the choice for subscriber control plane since operators may have exhausted current private IPv4 space in the network forcing IPv4 renumbering. In addition, operators are running low on public IPv4 to address the CPE devices in the home.

2. Tunneling techniques using IPv4 as a transient network with a tunnel could connect IPv6 clouds together. Later using the IPv6 transient network with a tunnel could connect IPv4 clouds together.

Translation techniques (translators), to allow IPv6-only devices to communicate with IPv4-only devices.

3. NAT offers options for the operator to transition IPv4 clouds to IPv6. After the network transitions to IPv6 NAT allows subscribers to use IPv4 over the providers IPv6 network.

4. Native IPv6 or IPv6 only, having exhausted the IANA/8 space, some operators must use pure IPv6 for future customers, current customers and new services such as Wi-Fi, IPTV, IP STB, eDVAs and eRouters.

Expect all of these to be used, in combination. To determine which method is the most desirable operators should test the available transition techniques. For the purposes of a scalable solution in the long run this document will only explore native IPv6 and dual stack.

Dual Stack

Dual stack means IPv4 and IPv6 running together in the same network on the same device. For example, a server or workstation will operate with an IPv6 and IPv4 address connected to a network with IPv6 and IPv4 addressing. In dual stack operation, a client is able to query DNS in IPv4 and IPv6. IPv4 uses A records and IPv6 uses AAAA (quad A) records. Based on the preferred protocol of the operating system, one address will be used, for example, the IPv6 address.

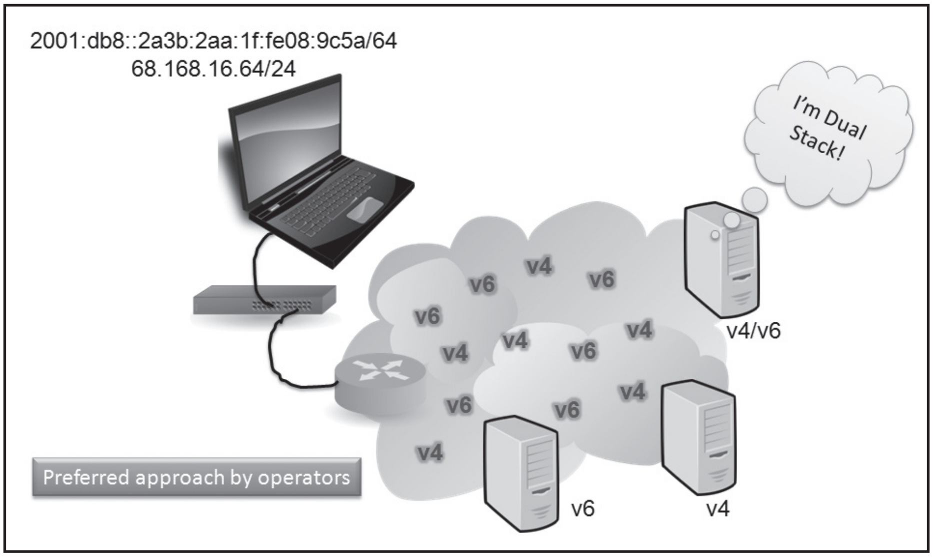

2001:db8::2a3b:2aa:1f:fe08:9c5a/64
68.168.16.64/24
I'm Dual Stack!
v6
v4
v6
v6
v4
v4
v6
v4
v4/v6
v6
v6
v4
v4
Preferred approach by operators
v6
v4

Figure 15: IP Dual Stack

Tunneling

IPv6 offers tunneling techniques which include tunneling IPv4 into IPv6. A tunnel may encapsulate IPv6 traffic within an IPv4 packet or the opposite. An example of a type of tunnel is shown below, called “6to4”, designed for subscribers who are unable to get IPv6 addressing from their provider. The 6to4 tunnel requires an address assignment on each router. The router to router, or p2p, connection will automatically create a unicast tunnel under RFC 3056. The tunnel allows devices to move IPv6 packets over an IPv4 network such as the Internet or aggregate network.

The downside of a 6to4 tunnel is there is no guarantee that devices will be reachable by all native IPv6 hosts using 6to4 tunnels. Where the Internet is used as a tunnel, the relay router will be outside the provider’s control. That means the provider is unable to know the return path to a given subscriber.

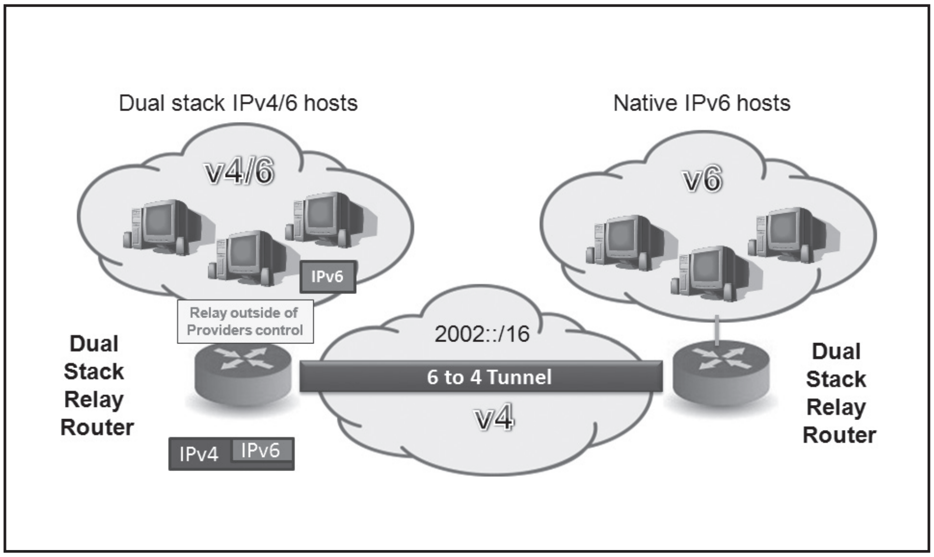


Figure 16: IPv6 to IPv4 Tunnel

Dual stack is needed in tunnels. 6to4 is being used whether MSOs have deployed 6to4 relays or not. A number of operating systems and home network equipment supports 6to4 by default. Deployment of 6to4 replies dramatically reduces latency in the network, 50% or more.

Using 6 Rapid Deployment (6rd) hosts will be reachable from all native IPv6 addresses. With 6rd the ISP provides the tunnel, allowing hosts to be reachable from all native IPv6 addresses. The 6rd solution solves the problem with the unknown return path. No need for DHCPv6 servers and ND here.

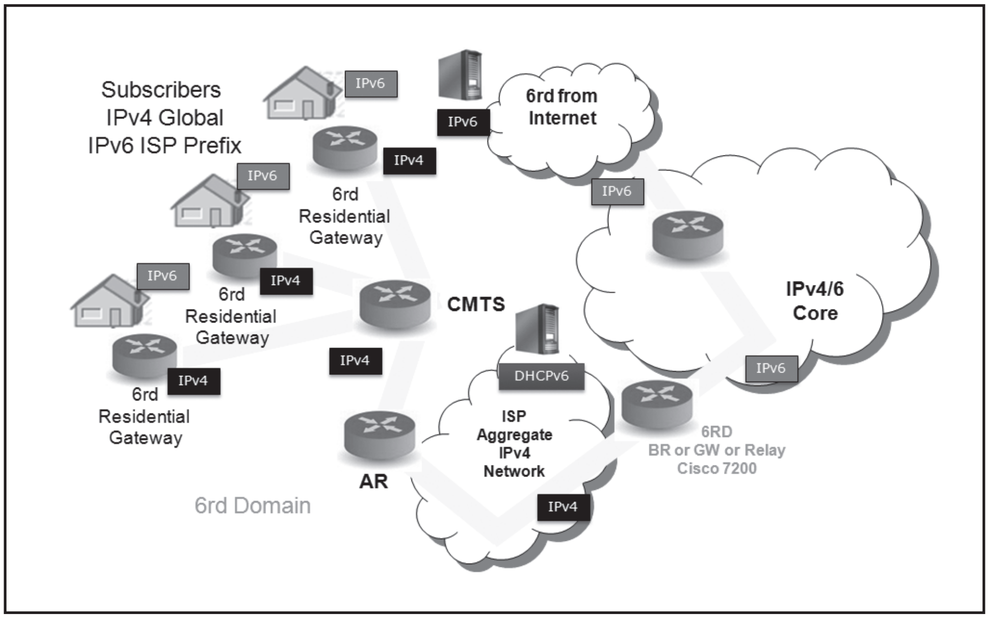


Figure 17: 6 Rapid Deployment

This solution works great when the cable provider is doing an incremental deployment of IPv6 and the aggregation network is IPv4 only. 6rd is still tunneling of IPv6 over IPv4 where subscribers use an ISP IPv6 prefix derived from the IPv4 address of the GWR that supports 6rd. The GWR encapsulates the IPv6 packets inside IPv4 and forwards them across the Internet backbone. A centrally located 6rd Border Relay/Router gateway is at the ISP to handle the 6rd packets headed towards the subscriber. It is important to understand with limited 6RD support in residential CPE, manual configuration may be required. DHCP servers will require enhancements to support 6rd DHCP options.

Another tunnel called Intrasite Automatic Tunnel Addressing Protocol (ISATAP) derives an interface address called the interface ID. The ID is the last 64 bits of an IPv6 address from any IPv4 address assigned to the node. The ISATAP tunnel uses the “5EFE” hex numerals in the address which makes it easy to ID this address.

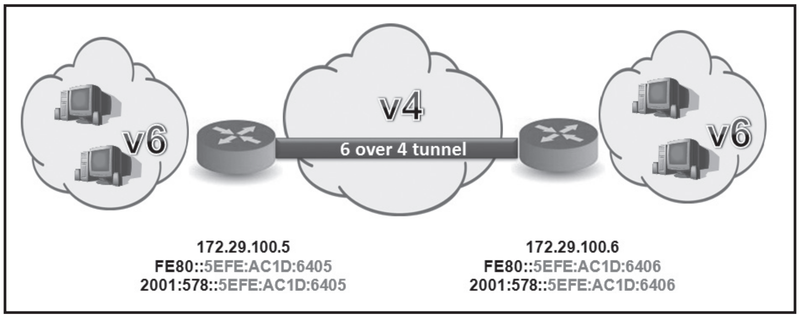


Figure 18: Intrasite Automatic Tunnel Addressing Protocol

Dual Stack Lite (DS-Lite) allows tunneling IPv4 over IPv6 networks to help solve the issue of the continued use of IPv4 addresses, such as in dual stack addressing. In this scenario, the cable customer can share private IPv4 data over an IPv6 cloud using the home GWR. The GWR uses an IPv6 provisioned address to create a 4 to 6 tunnel, encapsulating v4 inside a v6 packet, to the Address Family Transition Router (AFTR) in the provider’s network. The private v4 is recovered to a public v4 address at the AFTR, allowing sharing of IPv4 addresses between DS-Lite enabled networks.

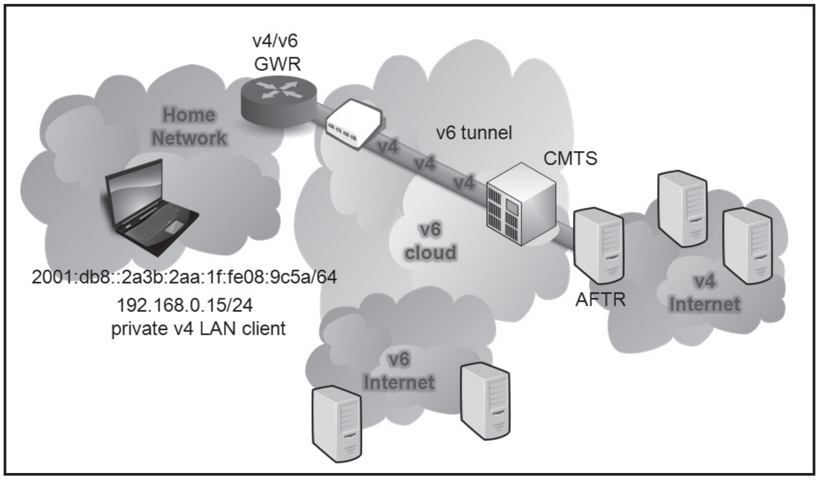


Figure 19: Intrasite Automatic Tunnel Addressing Protocol

Network Address Translation

Another protocol level translation technique is the use of NAT. One example that has been suggested for cable operators is Carrier Grade NAT (CGN) or Large Scale NAT (LSN) which is commonly label NAT444. NAT444 supports many subscribers using private IPv4 address assigned by the provisioning server for deploying this migration technique. The idea with NAT444 is to translate multiple private IPv4 addressing in a single public IPv4 address, supporting millions of translation states. Some of the disadvantages to NAT444 may be a costly solution, another device in the infrastructure and management may be complicated.

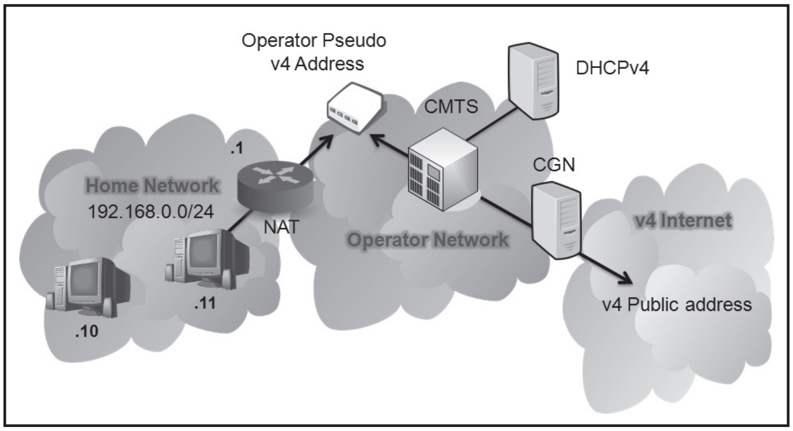


Figure 20: NAT444 / CGN / LSN

Native IPv6

Native IPv6 allows the client to access resources and the Internet using IPv6 only.

IPV6 SERVICES

Internet Control Message Protocol Version 6

The primary purpose of the Internet Control Message Protocol (ICMP) is to relay connectivity information over an IP network. ICMP travels directly within the IP datagram, so like UDP, ICMP is unreliable. ICMPv6 messages are similar to ICMPv4 messages but use different types/codes. The table below has some examples of the messages and type codes.

ICMP version 6	
Message	Type
Destination unreachable	1
Packet too big	2
Time exceeded	3
Parameter problem	4
Echo Request	128
Echo Reply	129

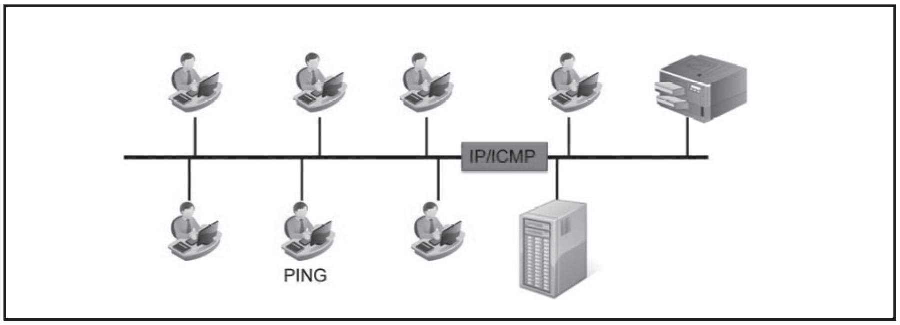


Figure 21: Internet Control Message Protocol

Neighbor Discovery with ICMP

IPv6 uses multicasted ICMPv6 ND to replaces broadcast protocols such as Address Resolution Protocol (ARP) found in IPv4. ND is used to multicast ICMPv6 functions such as redirects, router discovery (RD) and node discovery. The reachability of neighbors is a function of NUD. Hosts use ND to discover routers by using a Router Solicitation (RS) messages and State-Less Address Auto-Configuration (SLAAC) to auto configure IP addressing information for the link. Duplicate Address Detection (DAD) verifies a node's IP address is not in use on the link.

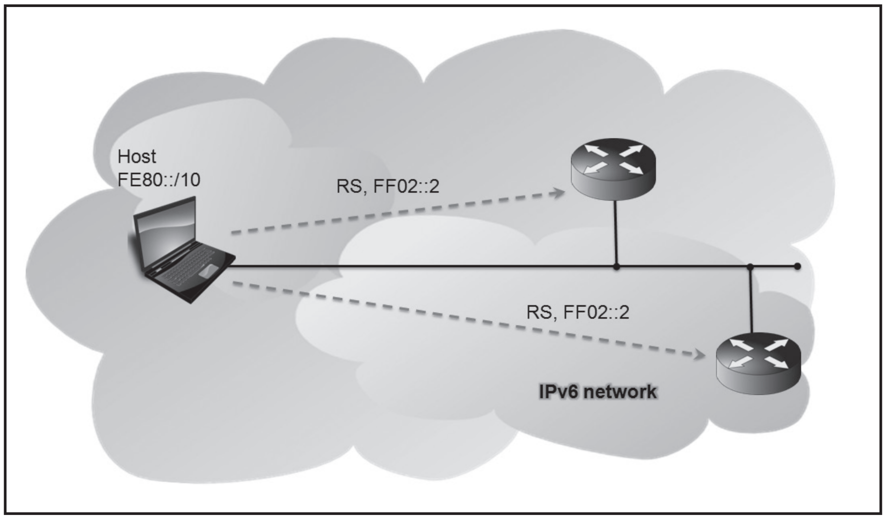


Figure 22: Neighbor Discovery using ICMPv6

Here we have a side-by-side comparison between IPv4 and IPv6 provisioning which indicates that IPv6 leverages existing IPv4 provisioning tools; it redesigned others and it implemented new ones. For example, for address assignment, IPv6 can use DHCP similar to IPv4 but it can also use stateless address auto configuration (SLAAC), similar to IPX and other reconfiguration capabilities. For address resolution, IPv4 has been using broadcast style ARP and reverse ARP, while IPv6 implements the same function using multicasted Neighbor Solicitation (NS) and Neighbor Advertisement (NA) in the context of ND. ND can detect half-link failures (not in 2way state).

RD has been implemented in IPv6 with the help of ICMPv6 ND packets. In IPv6, these functions are performed by RS and Router Advertisement (RA) packets which are also part of the ND. Name Resolution in IPv4 and in IPv6 is implemented with the help of DNS. QoS in IPv6 is the same as in IPv4 with the additional flow label field added in the IPv6 header. IPSec is mandated in the original architecture; however, in practice, IPSec is not used extensively in the data path. Mobile IPv6 removes the triangular path forwarding by introducing path optimization where the correspondent node can talk directly to the Mobile Node.

So why is it that in the context of IPv6 we are so interested in Stateless Address Provisioning mechanisms? The larger address space is going to accommodate thousands of devices and you cannot expect to manually set an IP address for all of them and DHCP may not be the right way to manage thousands of clients because that would generate a lot of state. A simple auto-configuration mechanism which scales would be necessary. Maintaining state for the addresses assigned to all these devices might lead to scalability concerns. Also, some of the new devices that are going to be IP enabled are not going to have very much computational power. Devices such as sensors might not be able to support the 4-way handshake that characterizes a stateful IP assignment.

Nevertheless, the power of stateful DHCP should never be underestimated. Service providers will always want to keep a very tight control over the addresses that are assigned to subscribers.

ND Multicast Messages

ND Message	ICMPv6 Type	Description
NS	135	determine the link-layer of a neighbor, L2 reachability, DAD
RS	133	used by booting hosts to request router RAs immediately (for link prefixes)
NA	136	answer to NS, or change of link-layer address
RA	134	periodic presence advertisement containing prefixes for the link or response to RS
Redirect	137	inform hosts of a better first hop for a destination

On multicast-capable links, ND for IPv6 uses ICMPv6 messaging to determine associations between neighboring nodes. Under ICMPv6, ND consolidates the functions performed by multiple IPv4 protocols.

ND is used by hosts to discover neighboring routers, discover addresses, address prefixes and other configuration parameters.

ND is used by routers to advertise presence, host configuration; link prefixes; or inform hosts of a better next-hop address to forward packets. ND is used by nodes to dispel the link-layer address of a neighboring node, ascertain when the link-layer address of a neighbor changed or to ascertain whether a neighbor is still reachable – called NUD, RFC 2461. Other uses include IPv6 address resolution (replacement for ARP), SLAAC, DAD, renumbering and redirection. The ND consists of the IPv6 header, ICMPv6 header, ND header and ND options.

Important: the ARP functionality, which in IPv4 was implemented in this odd layer in between Layer 2 and 3, has been integrated in the case of IPv6 in ICMP, the natural framework for a control plane mechanism.

A host that becomes active on a link will discover local routing resources (prefixes on a link) by sending out a multicast request called an RS. For example, the host sends an RS packet to inquire about the presence of a router on the link when booting. The RS is sent to all routers listening to this multicast group using multicast address: FF02::2. The source IP address is either a link-local address (FE80::/10) or an unspecified IPv6 address (::).

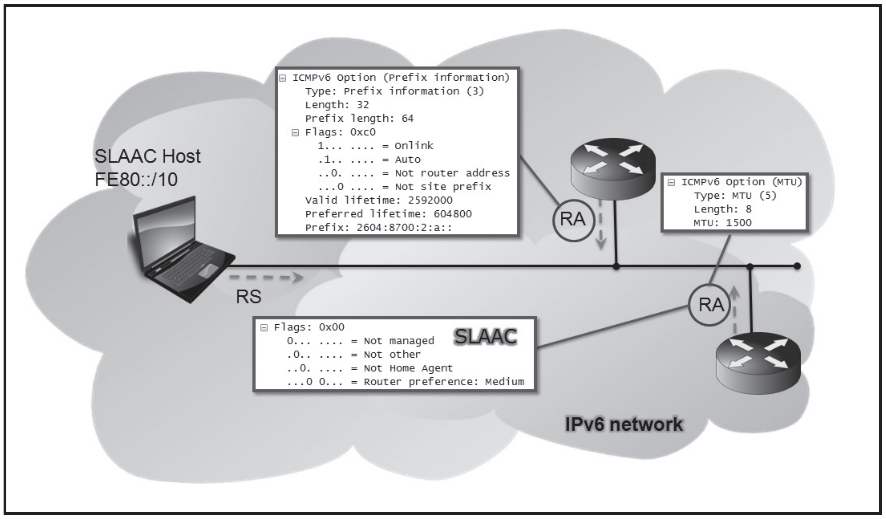
Routers can advertise their services on a link with the help of RA packets. Through an RA, a router will announce its availability; it will announce all the operational prefixes on the link and the preferred address provisioning mechanism, SLAAC or stateful DHCPv6.

RAs are sent in response to a received RS solicitation from hosts or they can be periodically sent. RAs are sent to all nodes using multicast address FF02::1, in which case all hosts on a link will receive that packet OR they are sent directly to the host who requested it. The RAs contain a set of Internet operational parameters that are important to hosts: a default hop limit that hosts should use in ongoing packets. A link's MTU ensures all hosts use the same MTU value.

Router lifetime is the lifetime associated with the default router in units of seconds. The maximum value of 18.2 hrs and a value of 0 means it is not a default router.

DHCPv6 Flags

The M flag or managed flag, when set to "1" indicates that the client uses stateful address provisioning via DHCP. The O flag, or other flag, when set to "1" indicates administered stateless or the client can use other parameters of DHCPv6 such as DNS or default gateway. An RA with an "O" bit set to "1" and an "M" bit set to "0" allows the client to use SLAAC and use DHCPv6 for other information. An RA with an "O" bit set to "0" and an "M" bit set to "0" allows the client to use SLAAC only. The Home Agent (HA) flag is used for Mobile IPv6 (RFC 3775). The default router preference indicates whether to prefer this router over other default routers.



```
⊟ ICMPv6 Option (Prefix information)
    Type: Prefix information (3)
    Length: 32
    Prefix length: 64
  ⊟ Flags: 0xc0
      1... .... = Onlink
      .1.. .... = Auto
      ..0. .... = Not router address
      ...0 .... = Not site prefix
    Valid lifetime: 2592000
    Preferred lifetime: 604800
    Prefix: 2604:8700:2:a::
```

```
⊟ ICMPv6 Option (MTU)
    Type: MTU (5)
    Length: 8
    MTU: 1500
```

```
⊟ Flags: 0x00
    0... .... = Not managed
    .0.. .... = Not other
    ..0. .... = Not Home Agent
    ...0 0... = Router preference: Medium
```

Figure 23: Neighbor Discovery and Router Advertisements

Let's look at a RS message and RA response. This figure shows the packet exchange during the RD process. A host that does not intend to wait for the periodic RAs are going to send the RS to the router's multicast address. The host queries for the local routing resources.

All routers on the link will pick up the packet and will respond with the RA that is sent to the node's multicast address. This packet will advertise the router and all the operational information relevant to that link. The key point is that IPv6 implements dynamic mechanisms to discover and advertise routing resources on a link.

Duplicate Address Detection

Once an IPv6 address has been auto generated, a host must verify its uniqueness with the help of the DAD mechanism. In this example, Host A auto generated its own link-local address. Host A will place this address in a tentative state until it verifies its uniqueness.

To do so, Host A will send a NS message for itself. The destination address of this packet will be the solicited node multicast addresses corresponding to the auto generated link-local address.

The packet will query for the auto generated link-local address of Host A.

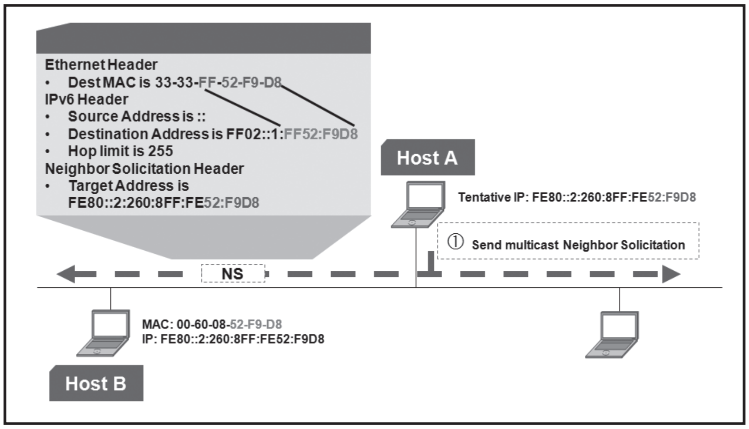

Ethernet Header
• Dest MAC is 33-33-FF-52-F9-D8
IPv6 Header
• Source Address is ::
• Destination Address is FF02::1:FF52:F9D8
• Hop limit is 255
Neighbor Solicitation Header
• Target Address is FE80::2:260:8FF:FE52:F9D8
Host A
Tentative IP: FE80::2:260:8FF:FE52:F9D8
① Send multicast Neighbor Solicitation
NS
MAC: 00-60-08-52-F9-D8
IP: FE80::2:260:8FF:FE52:F9D8
Host B

Figure 24: Duplicate Address Detection

In the example above, the address auto generated by Host A is already in use by Host B on the same link. Hence Host B will take up the NS message and it will reply with the NA destined to all nodes on that link to indicate that it is using that particular address. Host A will also receive the message; it will realize that its address is not unique and it will invalidate the auto-generated address.

Redirects

Another message identified in the ND is the Redirect message. Similar to IPv4, the Redirect message is used by the router to inform a host of a better gateway. In this example, Host A through the RD process identifies R2 as its default gateway so all traffic destined outside of the link will be sent at Layer 2 from Host A to Router R2. If R2 realizes that R1 is a better gateway for Host A, it will send a redirect message to Host A indicating that it should change its default gateway to R1.

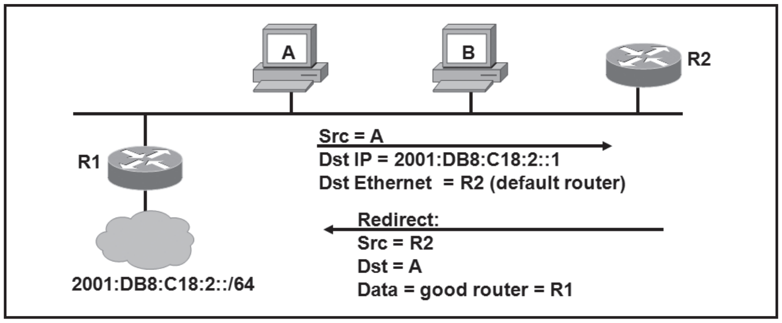


Figure 25: Redirects

Address Resolution

The Layer 2/Layer 3 address mapping in IPv6 is performed with the help of NS (ICMPv6 135) and NA (ICMPv6 136) messages.

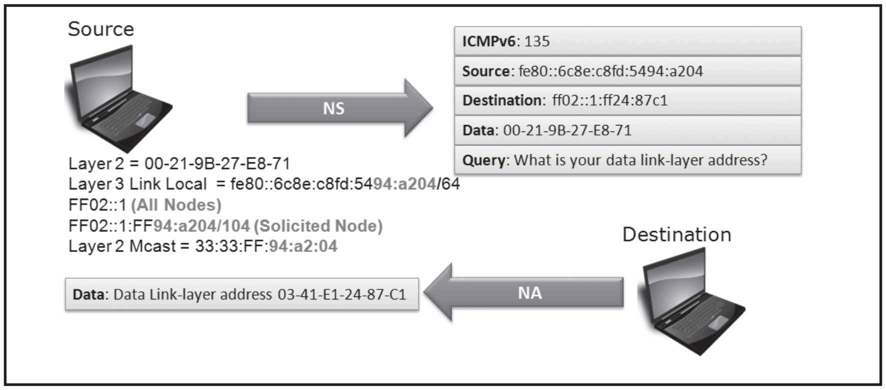

Source
NS
ICMPv6: 135
Source: fe80::6c8e:c8fd:5494:a204
Destination: ff02::1:ff24:87c1
Data: 00-21-9B-27-E8-71
Query: What is your data link-layer address?
Layer 2 = 00-21-9B-27-E8-71
Layer 3 Link Local = fe80::6c8e:c8fd:5494:a204/64
FF02::1 (All Nodes)
FF02::1:FF94:a204/104 (Solicited Node)
Layer 2 Mcast = 33:33:FF:94:a2:04
Destination
Data: Data Link-layer address 03-41-E1-24-87-C1
NA

Figure 26: Address Resolution

At boot time, every IPv6 node has to join two special multicast groups for each network interface. All-nodes multicast group: FF02::1. A link-local solicited-node multicast group: FF02:1:FFxx:xxxx (the "x's" will be derived from the lower 24 bits of the node's address). Solicited node multicast addresses are used in ND for obtaining layer 2 link-layer addresses of other nodes.

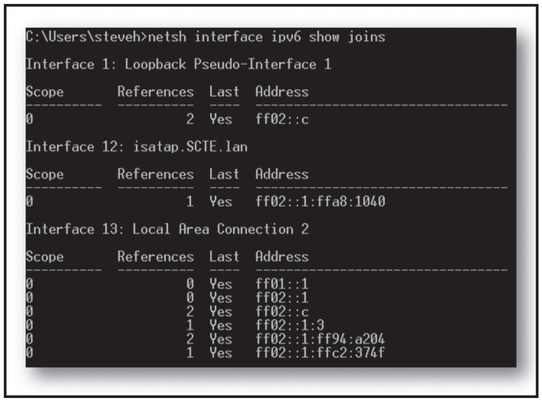
The process is initiated by the source using an NS which has as a layer 2 24 bit destination address and the solicited node multicast Layer 2 address corresponding to the target IPv6 solicited node address – in this case "24:87c1". The Layer 2 source is the address of the interface, "00-21-9B-27-E8-71" of the originating host. The Layer 3 source is the Link-Local address "fe80::6c8e:c8fd:5494:a204/64" of the sending node. This packet contains a query "What is your Link-layer address?" The response to NS is delivered in a NA (ICMPv6 136) message. NAs are also used to inform of a change in the link layer address of a host.

With this information, Host A and Host B can exchange information directly on this link. IPv6 implements dynamic mechanisms to discover and advertise routing resources on links.

Multicast Joins

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

All-nodes multicast group: FF02::1. A link-local solicited-node multicast group: FF02:1:FFxx:xxxx (the "x's" will be derived from the lower 24 bits of the node's address)

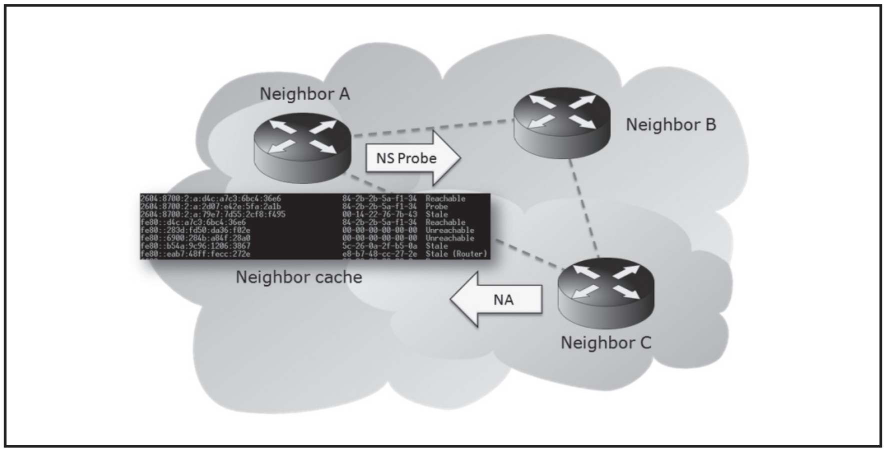


```
C:\Users\steveh>netsh interface ipv6 show joins

Interface 1: Loopback Pseudo-Interface 1

Scope       References  Last  Address
----------  ----------  ----  ---------------------------------
0                    2  Yes   ff02::c

Interface 12: isatap.SCTE.lan

Scope       References  Last  Address
----------  ----------  ----  ---------------------------------
0                    1  Yes   ff02::1:ffa8:1040

Interface 13: Local Area Connection 2

Scope       References  Last  Address
----------  ----------  ----  ---------------------------------
0                    0  Yes   ff01::1
0                    0  Yes   ff02::1
0                    2  Yes   ff02::c
0                    1  Yes   ff02::1:3
0                    2  Yes   ff02::1:ff94:a204
0                    1  Yes   ff02::1:ffc2:374f
```

Figure 27: Multicast Joins

Neighbor Unreachability Detection

Communication failures via neighbors or to a neighbor could occur at any time. Here is an example of how NUD works for all paths between nodes and neighboring nodes.



```
2604:8700:2:a:d4c:a7c3:6bc4:36e6       84-2b-2b-5a-f1-34  Reachable
2604:8700:2:a:2d07:e42e:5fa:2a1b       84-2b-2b-5a-f1-34  Probe
2604:8700:2:a:79e7:7d55:2cf8:f495      00-14-22-76-7b-43  Stale
fe80::d4c:a7c3:6bc4:36e6               84-2b-2b-5a-f1-34  Reachable
fe80::283d:fd50:da36:f02e              00-00-00-00-00-00  Unreachable
fe80::6900:284b:a84f:28a0              00-00-00-00-00-00  Unreachable
fe80::b54a:9c96:1206:3867              5c-26-0a-2f-b5-0a  Stale
fe80::eab7:48ff:fecc:272e              e8-b7-48-cc-27-2e  Stale (Router)
```

Figure 28: Neighbor Discovery and Router Advertisements

An IPv6 node builds a cache with all the neighbors it discovers. NUD maintains reachability state for all the neighbors by periodically probing all the entries in its ND cache with NS. UD can also target very specific neighbors by sending unicast NS to those nodes. When a target responds with a NA it will be flagged as reachable in the ND cache. If a NA is not received, the target is in an unreachability state.

Address Autoconfiguration

The IPv6 protocol implements an Address Auto Configuration mechanism which is unavailable to IPv4. By default, an IPv6 host configures a link-local (FE80::/64) address for each interface using DAD. IPv6 interfaces can automatically acquire addresses even in the absence of a stateful protocol such as DHCPv6. Interfaces will auto generate new local addresses by default.

With the help of a RD mechanism RS, they can also acquire additional unicast addresses, addresses of larger scope such as global addresses or Unique-Local Addresses. They can also learn the IP addresses of default gateways and they can learn the configuration parameters that are relevant on that link.

States of an Autoconfigured Address

Auto configured IPv6 addresses are dynamic in nature. They can be in various operational states as shown in the illustration below.

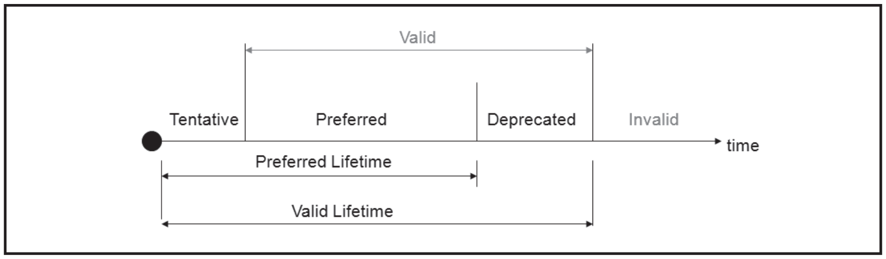


Figure 29: Neighbor Discovery and Router Advertisements

An address that has been verified for uniqueness is in a tentative state and once uniqueness has been confirmed the address becomes valid. A valid address that can be used in an unrestricted way is in a preferred state. An address that can be used for system communication, but is discouraged for future use, is in a deprecated state. And finally, an address that can no longer be used is in an invalid state.

There is one important aspect of the IPv6 address lifetime that we must highlight here. A host which has been using the deprecated address to establish a set of sessions or communications with other hosts will continue to use that address even though the address is now in a deprecated state. Any new sessions that this host will establish must use the new preferred address instead. This mechanism enables the host to migrate from an old address to a new address with minimal impact on users' experience.

Renumbering

Renumbering a network is a complex and demanding event.

IPv6 with ND can simplify the process to a certain extent. A router's interface can be configured with a new prefix alongside the old prefix.

The new prefix can be configured with a longer lifetime and the old prefix with a shorter lifetime. This information will be learned by the hosts on the link from RA. Hosts will deprecate the old prefix in time and they will prefer the new prefix. After awhile none of the hosts will be using the older prefix so the old prefix can be removed from the router interface.

While IPv6 offers this interesting and simple mechanism for changing addresses of hosts within a link, it is important to remember that renumbering is a very complex process where we have to change addresses and routers; we have to change policies, access lists, firewalls; and we have to look for cached entries of old IP addresses. So despite this enhancement, it would be misleading to state that IPv6 resolves the entire renumbering problem.

Path MTU Discovery

IPv6 nodes SHOULD implement PMTU Discovery in order to discover and take advantage of paths with PMTU greater than the IPv6 minimum link MTU.

Domain Name Service

As we know it is impossible to memorize millions of IP addresses, in IPv6 how about an undecillion? The Resource Records (RR) has to be updated to support v6 as shown here. Here we see the A record or host record and the pointer record or reverse host record. The 13 root servers (A-M.root-servers.net) around the world (10 in the United States) are reachable on an IPv6 transport.

Add transport, resource records and dynamic DNS.

ROUTING IPV6

Routing Information Protocol

Routing Information Protocol (RIP) is a distance-vector routing protocol, which prevents routing loops by implementing a limit on the number of hops allowed in a path. RIP has a maximum hops limit of fifteen which also limits the size of networks that RIP can support. A hop count of sixteen is considered an infinite distance in RIP and used to undesirable routes in the path selection process.

RIP implements stability features such as split horizon, route poisoning and hold-down mechanisms to prevent incorrect routing information from being propagated throughout the network. RIPv2 was created to deal with VLSM in modern network infrastructures and to address security concerns.

RIP next generation (RIPng), defined in RFC 2080, is an extension of RIPv2 for support of IPv6. The main differences between RIPv2 and RIPng are:

- Support for IPv6.
- RIPng does not support RIPv1 updates authentication since IPSec is used for authentication.
- RIPng does not support attaching arbitrary tags to routes like RIPv2.
- RIPv2 encodes the next-hop into each route entry, while RIPng requires specific encoding of the next hop for a set of route entries.
- RIPng sends updates on UDP port 521 using the multicast group FF02::9.

Open Shortest Path First

To address the needs of larger cable networks, Open Shortest Path First (OSPF) was developed. OSPF is a link state protocol that is available in two versions; 2 and 3. OSPF version 3 (OSPFv3) was designed for IPv6 by supporting mechanisms for ND and adjacency formation. The size of the OSPFv3 header was reduced from OSPFv2's 24 bytes to 16 bytes. The new header design will continue to support the same five OSPF packet types (hello, database description, link state request, link state update and link state ack). One important note about OSPFv3 is that the protocol will continue to use a Router ID, or RID, and area id of 32 bits. IPv6 was designed with integrated support for IP security (IPSec) allowing the authentication fields of the OSPF header to be suppressed in OSPFv3.

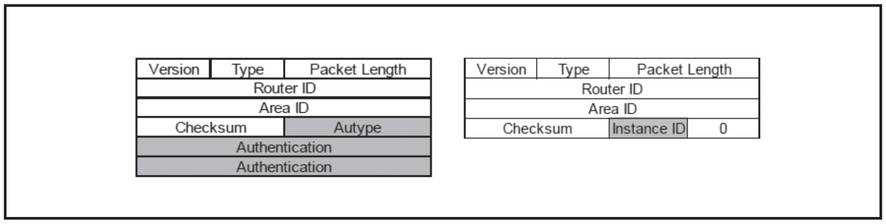


Figure 30: OSPFv2 Header vs. OSPFv3 Header

Intermediate System to Intermediate System

Another routing protocol that addresses the needs of larger cable networks is called Intermediate System to Intermediate System (IS-IS). For IPv6 support, integrated IS-ISv6 allows a single instance of the protocol for routing IPv4 and IPv6 domains in the cable network. Additional type, length, value (TLV) fields were added to the protocol that provide a mechanism for IS-IS to advertise neighbors and IP prefixes on the network.

Border Gateway Protocol

Border Gateway Protocol, or BGP is a path vector exterior gateway routing protocol used by cable networks for inter-autonomous system routing. There is a requirement for cable operators to interconnect networks and provide routing. BGP is used between cable ISPs and their larger private clients to exchange routing information. The updated version of BGP, Multiprotocol Border Gateway Protocol version 4 (MP-BGP4) supports multiprotocol extensions (e.g., IPv6 and MPLS) and the use of header extensions, allowing BGP to carry routing information for protocols other than IPv4.

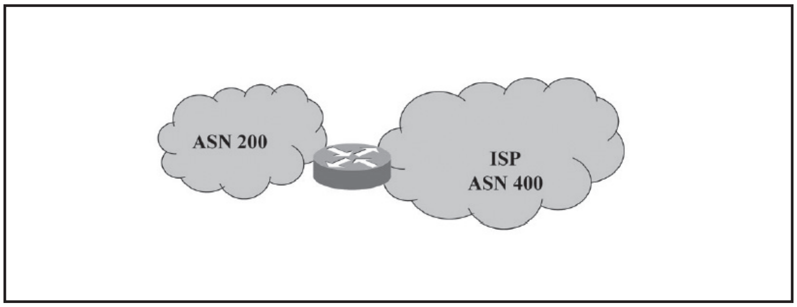


Figure 31: Autonomous Systems connected together using a router running BGP

SECURITY

This section will explore the security considerations for IPv6 deployment in cable networks.

ND Attacks: Packet Looping

The network core which is comprised of high-end routers and high speed connectivity is one of the first areas where IPv6 should be enabled. Enabling IPv6 within the core network first is recommended since the devices in the core are typically in a position to support Dual Stack without major vendor support issues. Further it allows the core network to stabilize before addressing access and edge networks where there is the potential for more specialized network elements, which could have varying degrees of IPv6 support. For these reasons it also makes sense to insure that the core network has a suitable security posture to support an IPv6 deployment.

Neighbor Discovery Protocol Exhaustion Attack

A significant attack vector on the core of networks is exhaustion of the ND table on core routers. A standard IPv6 subnet with a /64 network mask will have 264 or 18,446,744,073,709,551,616 possible host addresses. Common layer 3 switches will have on the order of less than 10,000 possible ND entries. ND entries on routers are needed in order to forward traffic at layer 2. ARP provides for a similar function in IPv4. The characteristic of layer 3 devices having relatively small ND tables compared to the possible number of addresses on a standard /64 subnetted network presents an opportunity for an attacker to execute a denial of service attack on the core routers. Using tools such as flood_advertise6 an attacker can send a large number of false ND announcements to the target router effectively disabling the router's ability to learn new neighbors and possibly flushing valid neighbor entries from the existing table. This particular attack requires the attacker to be on the same layer 2 network as the router under attack.

A similar attack can be executed against point-to-point links within the core of an IPv6 network without the requirement for the attacker to be on the same layer 2 interface as the target. This is achieved by sending an ICMPv6 message to a non-existent address within the scope of the IPv6 subnet configured on a point-to-point link. This in turn causes the router under attack to create a new cache entry in the ND table marked as INCOMPLETE and send a NS message out the point to point interface. Since there is a potential of a significantly large number of addresses on a link with a /64 subnet mask an attacker can generate ICMPv6 message to this large range of addresses and flood the link with NS message while simultaneously consuming the ND table on the router.

Due to the potential of this type of DOS attack on IPv6 Point to Point links it is generally recommended to use a /127 mask on point to point links with the core of any IPv6 network. While this goes against some current recommendation which dictate that a /64 subnet mask should be configured everywhere, at the time of this writing there is an effort within the IETF community to make this a standard. In the meantime, it is generally a good idea to use /127 on point-to-point links as there is a significant risk of DOS attacks with very little advantage of keeping a /64 on these point-to-point links.

Packet Looping DOS Attack on Point to Point Links

The unique traffic forwarding rules for point-to-point (p2p) links open the door to a special case of DOS attacks. On a p2p interface routers do not perform layer 2 address resolution, a process which would enable them to determine if an IP address within the subnet of that p2p link is active. Routers operate under the assumption that there is only one physical host that can be reached out of a p2p interface even if the IP subnet configured for that link supports more than two (one for each end of the link) IP host addresses. A packet sent by mistake, or maliciously, to an IP address that belongs to the subnet of a p2p link – an address that is however not assigned to the two ends of the link – will bounce between the two routers until its TTL expires, opening the door for an amplification mechanism which can be used to attack the routers. Upon receiving the packet, neither router checks whether the destinations really exists on the link. Both routers will simply put the packet back on the link with a decreased TTL.

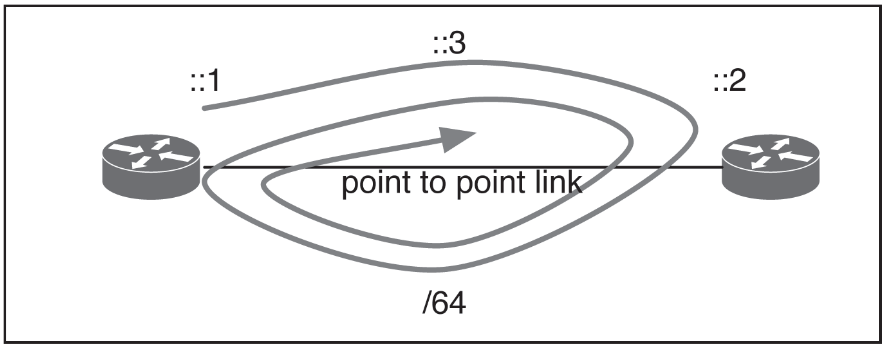


Figure 32: Packet Looping Denial of Service Attack

One solution to this problem is to assign to p2p links prefixes that support only two hosts. In the case of IPv4, this is not only easy but also desired since addressing resources are scarce. It is thus practical and also common to assign a /30 or /31 on a p2p link. In this scenario, the threat mentioned above is not applicable. However, the use of /30s or /31s on p2p links is not a rule or recommendation that is always observed. The problem is exacerbated in the case of IPv6 by specific address architecture constraints. According to RFC4291, the Interface ID of an IPv6 address MUST be 64 bits long, thus requiring that the prefix length on any link not be longer than 64 bits. If RFC 4291 is strictly observed, p2p links offer [(2^64)-2] host addresses to which attack packets can be directed to leverage the amplification of the threat mentioned above. Even if RFC 4291 is not strictly observed, /127 prefix lengths cannot be used on p2p links leaving us with /126 as the longest prefix that can be assigned. This means that we will always have 2 host addresses that can be used for a DOS attack.

This potential threat was directly addressed in the ICMPv6 specification (RFC 4443):

One specific case in which a Destination Unreachable message is sent with a code 3 is in response to a packet received by a router from a point-to-point link, destined to an address within a subnet assigned to that same link (other than one of the receiving router's own addresses). In such a case, the packet MUST NOT be forwarded back onto the arrival link.

The amplification threat would be eliminated by the implementation of the recommendations of RFC 4443. Because of inconsistent stack implementations or because of very specific feature characteristics, these recommendations are not always observed, thus leaving the door open for attacks.

As such it is very important for network providers to be aware of this attack and remediate it by subnetting the point-to-point links with a /127. If that is not an option then applying appropriate access-lists and filters to the edge of the network would be desired so as to filter traffic destined to point-to-point links in the core of the network. Normally there would be no reason why traffic originating from peer networks or subscribers should be communicating with the point-to-point subnets in the core.

This can be extended further by using Link-Local IPv6 addressing on all core point-to-point links. Since there is no need for outside communication to these links, Link-Local communication will allow the core neighboring routers to communication buy eliminate direct communication to point-to-point subnets from outside the core.

Use of Access Control Lists to Secure the Core

Access lists are one of the fundamental tools, which can be used to secure the core IPv6 network. Much of the traffic which you are filtering at an IPv4 layer will be similar in IPv6. However, there are notable exceptions which need to be addressed. First is you might need to allow ICMP traffic through your IPv6 access lists in order for fundamental IPv6 features to work such as the MTU path discovery. For IPv6, if an operator routinely blocks ICMP in IPv4 this policy will have to be adjusted.

A difference in IPv6 compared to IPv4 is the presence of extension headers. Extension headers handle many of the function which were previously handled IP header options in IPv4. They have been removed from the main header and now handled with extension headers in IPv6. Various functions are handled by IPv6 extension headers such as Routing Headers, headers for IPv6 (AH/ESP) and fragmentation headers. Routers can be configured to look and act on traffic which contains specific extension headers. An example of such is an access list provided below.

```
IPv6 access-list EH-type
deny IPv6 any 2001:2B8:1:1::/64 routing
```

This access list denies traffic destined to the 2001:2B8:1:1::/64 with the Routing extension header will be dropped. The routing header, sometimes referred to as RH0, provides very similar functionality to source routing in IPv4. Because of its ability to be used in DOS attacks, the official support of RH0 headers has been deprecated by the IETF. Nonetheless, it is important at the edge of networks to insure that packets are not being delivered from peering networks with the extension header of RH0. Using an access list with extension header support is an effective way of controlling RH0 packets entry into your network.

Further, any network device which is required to parse upper layer protocols such as TCP and UDP are required to be able to parse IPv6 Extension Headers. Any higher layer filters could fail without the ability to parse extension headers. It is also important to evaluate the performance of network devices which parsing extension headers at varying offsets within the packet since the path of the packet can vary widely between different network devices and vendors.

Finally, the Hop by Hop extension header requires that the intermediate router explicitly process this packet. This in turn requires that the packet to be punted to the main CPU of the router for processing which can impact forwarding performance. Unless there is an explicit need for Hop by Hop extension headers, then they also should be filtered out at the edge of the network.

Securing Access to Backend Systems

The security of backend provision and billing systems typically located in the data center would take a similar approach as IPv4 again with notable exceptions. All security policies and procedure which are currently applied to these systems need to be mapped to IPv6. Some of the technical tools which might be used to secure these systems are Intrusion Detection and Prevention (IDS/IPS) systems and network based firewalls.

Considering IDS/IPS systems there are design factors which change when they are monitoring an IPv6 network. By nature IDS/IPS systems are memory intensive functions in that the packets being examined have to be copied to working memory in real time. One well known way for attackers to bypass or completely overload IDS/IPS systems is to heavily fragment the attack traffic. This in turn causes the IDS/IPS system to copy all of the fragments to memory first and examine the payload once all fragments have been received. By flooding the IDS/IPS system with a great deal of fragmented packets, the system might become so low on resources that it can no longer alert on subsequent attacks.

The above is exacerbated when consider IPv6. Since we are dealing with a 128bit address vs. 32 bits, the amount of memory required for packet copies increases significantly. Further the attacker could expand the packets even more by adding unnecessary extension headers. When deploying IDS/IPS systems to monitor and react to attacks on IPv6, it is important to benchmark how the vendor in question will handle resource constraints in IPv6.

Making Reconnaissance of IPv6 Network Harder for Attackers

From a network attacker perspective reconnaissance of IPv4 network was traditionally performed by scanning the IPv4 subnet for hosts. This is either done with simply ping sweeps and/or Nmap scans of the subnet. Due to the extremely large number of hosts possible on a /64 IPv6 subnet, this approach is no longer viable for attackers. Scanning an IPv6 network using traditional methods would take upwards of thousands of years so attackers are moving their reconnaissance to other methods for IPv6.

DNS is being relied on far more heavily as organizations move to IPv6 as the practicality of remembering and typing addresses becomes even more onerous. Since one can assume that more hosts will be populated in DNS in the context of IPv6 network this becomes a natural point for an attacker to begin their reconnaissance. Brute forcing DNS databases is a typical method of retrieving a list of hosts to attack. Many tools exist to take a dictionary and perform and large amount of DNS lookups against a target DNS server. Some of these dictionaries will try common host names first such as firewall, firewall1, server, mail, etc. It can be extended further with dictionaries based on movies, pop culture, etc.

It is very common for DHCPv6 servers to start in at a common point such as :1000 and increment sequentially. Therefore it can be found that hosts in an IPv6 network are not equally distributed across the address space but instead highly clustered. Once this cluster has been identified it is easy to enumerate the rest of the host. Therefore is there is an option in your DHCPv6 to provide a degree of randomness across the /64 it is recommended that this feature be used.

Finally it is common to find DNSSec being deployed alongside IPv6. DNSSec can also be used to assist in performing reconnaissance NSEC records in DNSSec indicate that nothing exists between two separate records in the DNS database. So for example:

```
cisco.com.     10800  IN    NSEC   content.cisco.com. A NS SOA MX TXT RRSIG NSEC DNSKEY
```

indicates that no DNS records exist between cisco.com and content.cisco.com. So if someone says that beta.cisco.com exists you can prove that it does not. The problem with this approach is that these records are in the database and can be queried by anyone. So by asking if beta.cisco.com exists the attacker is automatically finding out that content.cisco.com exists. NSEC3 records remediate this vulnerability and should be used whenever possible in DNSSec deployments to make this method of IPv6 reconnaissance harder for an attacker.

RA Hijacking and Mitigation

In IPv4 man-in-the-middle attacks on a layer 2 VLANs are somewhat difficult as ARP cache poisoning needs to be performed. This basically corrupts the ARP cache of target hosts so that the mac address of the attackers machine is placed as the next hop MAC address for the default gateway. With this, even though the target still has the proper IPv4 gateway address configured, the traffic now is being directed to the attacker's machine at layer 2. From there the attacker can record and change information while simultaneously passing to the destination network.

This type of attack becomes fundamentally easier with IPv6. All an attacker needs to do to become a man-in-the-middle is to advertise their device as a router with an IPv6 RA message. The attacker can even set the priority of their announcement to take precedence over other announcements from valid routers serving as default gateways.

There are functions such as RA Guard in service provider grade switches. RA Guard attempts to lock down the port where RAs are expected. So for example, one would configure a switch port directly connected to a router to allow RAs to pass while switch ports connected to servers or other devices would block RAs. This works well unless the attacker is sophisticated as this type of control can be bypassed by fragmenting the RA packet and inserting a Destination Options Header before the ICMPv6 RA message. If the packet is fragmented so that the first fragment contains the Destinations Options header and the second fragment also contains a destinations options header followed by the RA then the switch will never see the RA and will not drop the advertisement. Unfortunately, since Destination Options headers are by design read by end stations; the end station will reassemble the fragment, read past the destinations options header and receive the malicious RA.

While Secure Neighbor Discovery (SEND) is often proposed as a solution to this problem is it fraught with problems. First is that it isn't universally supported by server and host operating systems. Further the deployment of SEND adds additional complexity to the network which can be solved in an MSO environment by simply disabling RAs on networks where one is sure that stateless address auto configuration will not be used. This is often the case on network links in the core and access networks of an MSO which are not connected to hosts and gain no advantage to having SLAAC disabled. In short, turn of RAs unless you need them.

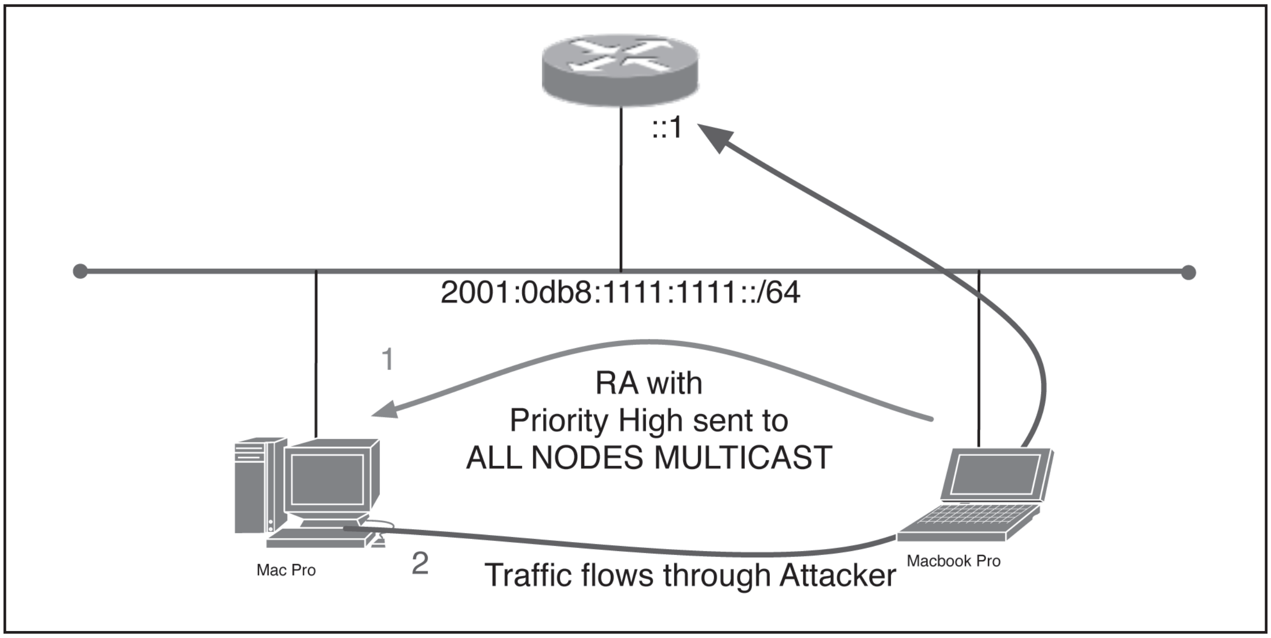


Figure 33: Router Advertisement (RA) Man in the Middle Attack

ND Security Threats

The ND faces similar security threats as its IPv4 counterparts.

Some examples of Redirect Threats are malicious last hop routers, NS or NA spoofing, and spoofed redirect messages that lead to back hauled or captured traffic. In ND, the sending host multicasts a NS, and the destination host, if reachable, responds with an NA containing its layer two address. These exchanges are completely unsecure; an attacker is able to generate an NA with his own layer two address as belonging to other hosts on the link.

The RA Hijacking attack becomes fundamentally easier with IPv6. All an attacker needs to do to become a man-in-the-middle is to advertise themselves as a router with an IPv6 RA message. The attacker can even set the priority of their announcement to take precedence over other announcements from valid routers serving as default gateways.

Service provider grade layer 3 switches support RA guard which attempts to lock down the port where RAs are expected.

The ND also faces Denial of Service (DoS) threats where bogus prefixes are advertised on links or malicious devices can claim to have the IPv6 address of real users, thus tampering with specific address detection mechanisms. Links can be overwhelmed with ND messages and ND messages with the wrong parameters can be used to tamper with stateless address auto configuration and renumbering mechanisms.

A typical router or host may support a small amount of ND entries, around 10,000 for a layer 3 device. Exhaustion of the ND table on a router will disable the router's ability to learn new neighbors and possibly flush valid neighbor entries from the existing table.

thc-ipv6 Toolkit – Attacking the IPv6 Protocol
http://www.darknet.org.uk/2010/07/thc-ipv6-toolkit-attacking-the-ipv6-protocol

Unlike IPv4, however, IPv6 has unique mechanisms that protect it against such threats. IPSec Authentication Headers (AH) can be used in the communication between hosts and between hosts and routers using transport (host to host) or tunnel mode (router to router). Authentication is applied to the entire IPv6 packet (with the mutable fields in the IP header zeroed out). IPSec can only be used with a manual configuration of security associations and the number of manually configured security associations needed for protecting ND can be very large, making this approach impractical for most purposes.

Secure Neighbor Discovery

A more secure version of ND called SEcure Neighbor Discovery (SEND) has been defined in RFC 3971. In IPv6, SEND offers mechanisms for protecting against address spoofing on a link and for certifying routers on a link. Nevertheless, as in the case of IPv4, we must remember that the link layer must be secure prior to any discussion of Layer 3 security. As in the case of ARP, basic ND faces similar security threats. It is thus imperative to secure Layer 2 before looking into the security considerations for Layer 3 and ND.

SEND is not universally supported by server and host operating systems.

Further, the deployment of SEND adds additional complexity to the network which can be solved in an MSO environment by simply disabling RAs on networks where one is sure that stateless address auto configuration will not be used. This is often the case on network links in the core and access networks of an MSO which are not connected to hosts and gain no advantage to having SLAAC disabled. In short, turn off RAs unless you need them.

DOCSIS FUNDAMENTALS

DOCSIS 3.0 requires support from network devices for the ND process. A cable modem must be capable of sending a NS message with its link-local address to the CMTS, and the modem must also perform DAD. Specific to a DOCSIS network, DAD ensures that no two devices attached to a given CMTS have conflicting link-local addresses. In addition, the modem must be capable of sending an RS message to the CMTS, and the CMTS must respond properly with a RA. This functionality is also required of DOCSIS 2.0+ IPv6-compliant cable modems. The ND process, along with that of DHCPv6 which is addressed below, creates potential for flooding. As such, a throttling capability on the CMTS against ND and DHCPv6 floods is very useful, protecting both the CMTS itself and the connected devices of the service provider's IP backbone.

MAC Domain Descriptor

DOCSIS 3.0 creates the framework needed for general use of IPv6. One important aspect of this is the MAC Domain Descriptor (MDD) defined by the DOCSIS 3.0 standards. A DOCSIS 3.0 CMTS is required to create the MDD and transmit it to all downstream channels in the MAC domain. The MDD serves a variety of purposes, including describing channel bonding configurations, but for IPv6, it controls the provisioning mode in use on CMs on a given MAC domain. All modems attempting to range and register on a MAC domain are required to interpret and comply with the parameters in the MDD message.

The IP Initialization Parameters TLV in the MDD message communicates the information necessary for the modem to perform IP Initialization, the IP Provisioning Mode and the pre-registration Downstream Service Identifier (DSID). The IP Provisioning Mode TLV tells the CM to provision in IPv4 mode, IPv6 mode, Alternative Provisioning Mode (APM), or Dual Stack Provisioning Mode (DPM). The pre-registration DSID is used by the modem to forward the multicast traffic required for IPv6 provisioning to the CM's IP stack.

CM Provisioning Modes

The DOCSIS 3.0 MAC and Upper-Layer Protocols Interface (MULPI) specification establishes four (4) possible modes of IP provisioning for cable modems:

1. IPv4 only
2. IPv6 only
3. Alternate Provisioning Mode (APM)
4. Dual-Stack Provisioning Mode (DPM)

These modes affect the address or addresses to be used for the management interface of the cable modem. IPv4-only and IPv6-only modes are self-explanatory, so only APM and DPM will be described herein. APM directs the modem to first attempt to configure itself using IPv6. Should this effort fail, either due to failure to acquire an IPv6 address or to a failure of the TFTP download process, the modem must automatically repeat the configuration process but using IPv4 instead of IPv6. The goal of DPM is to establish the capability of managing the modem via either IPv4 or IPv6. Successful registration of a modem with DPM requires it to acquire both an IPv4 and an IPv6 management address and also download its TFTP configuration file. The modem is required to attempt to download the TFTP file first using IPv6, and then (if that attempt fails) using IPv4. DOCSIS specifications require the use of DHCPv4 (for IPv4 provisioning) and DHCPv6 (for IPv6 provisioning). DHCPv6 functions in a similar manner as DHCPv4, but the specific messages exchanged between the host and the server are unique for IPv6.

The provisioning mode configured in the CMTS and communicated via the MDD is only reflected in the operational states of connected modems if said modems are not commanded to provision in a different mode via the IP Provisioning Mode Override. State information for a particular cable modem can be obtained from the CMTS. Most CMTS vendors support some form of the CLI command "show cable modem" for this purpose, and this command is normally capable of providing state information that differentiates between IPv4 provisioning and IPv6 provisioning.

MDD IP Provisioning Override

Because the contents of the MDD message apply to a MAC domain, the IP Provisioning Mode cannot be customized for an individual modem or a group of modems. Thus, it became necessary to provide a means of overriding the IP provisioning mode in the MDD message. Cable modems must also support IP provisioning mode override, which is also referred to as MDD IP Mode Override (MIMO).

The IP provisioning mode override is a group of MIB attributes which tell the CM how to perform IP Provisioning and when to apply the IP Provisioning. The CM is told to honor the IP Provisioning Mode in the MDD message or to override the IP Provisioning Mode in the MDD with either IPv4 provisioning or IPv6 Provisioning.

IP provisioning mode override support atop core IPv6 functionality is essential to the seamless migration of devices leveraging IPv6. Thoughtful planning and execution in conjunction with the functionality outlined here are enablers for MSOs adopting IPv6 to control the IP version used by devices and operators for management. These are all key variables to consider when planning a migration from IPv4 to IPv6. These techniques are enablers to specifically validate device IPv6 capabilities per device. The approach described, at a high level, affords MSOs a multitude of alternatives when planning the testing and deployment of IPv6 for device management including post deployment activities like troubleshooting. Less seamless and impactful alternatives include the migration of all DOCSIS devices to IPv6 in smaller increments or groups. Migration groups that include entire CMTS platforms for example offer much less control and increase the opportunity for service disruption. The presence of devices on CMTS platforms where DOCSIS devices do not properly support IPv6 could result or require an operator to back out or migrate back to IPv4 to minimize impact to affected customers. Cable modem behavior on a production network is a major variable and consideration for operators planning for the use of IPv6 for device management, in particular devices that may be running older firmware.

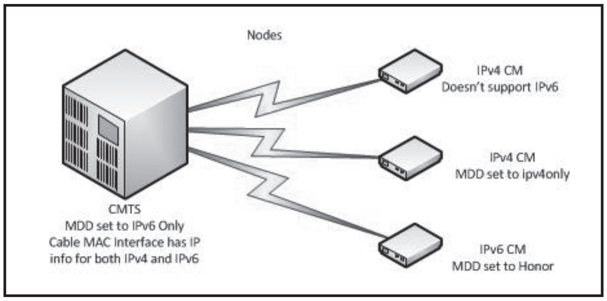


Figure 34: MAC Domain Descriptor (MDD) – Deterministic Approach

CMTS Support for IPv6

From a CMTS perspective, DOCSIS 3.0 creates the framework needed for general use of IPv6. One important aspect of this is the MDD defined by the DOCSIS 3.0 standards. A DOCSIS 3.0 CMTS is required to create the MDD and transmit it to all connected cable modems. The MDD serves a variety of purposes, including describing channel bonding configurations, but for IPv6 it controls the provisioning mode in use on a given MAC domain. All modems attempting to range and register on a MAC domain are required to interpret and comply with the provisioning mode established by the CMTS via the MDD.

As previously mentioned, cable modems must support IP provisioning mode override or MDD IP Mode Override (MIMO). IP provisioning mode override support atop core IPv6 functionality is essential to the seamless migration of devices leveraging IPv6. Thoughtful planning and execution in conjunction with the functionality outlined here are enablers for MSOs adopting IPv6 to control the IP version used by devices and operators for management. These are all key variables to consider when planning a migration from IPv4 to IPv6. These techniques are enablers to specifically validate device IPv6 capabilities per device. The approach described, at a high level, affords MSOs a multitude of alternatives when planning the testing and deployment of IPv6 for device management including post deployment activities like troubleshooting. Less seamless and impactful alternatives include the migration of all DOCSIS devices to IPv6 in smaller increments or groups. Migration groups that include entire CMTS platforms for example offer much less control and increase the opportunity for service disruption. The presence of devices on CMTS platforms where DOCSIS devices do not properly support IPv6 could result or require an operator to back out or migrate back to IPv4 to minimize impact to affected customers. Cable modem behavior on a production network is a major variable and consideration for operators planning for the use of IPv6 for device management, in particular devices that may be running older firmware.

The provisioning mode configured on the CMTS and communicated via the MDD is often reflected in the operational states of connected modems. DOCSIS specifications require the use of DHCPv4 (for IPv4 provisioning) and DHCPv6 (for IPv6 provisioning). DHCPv6 functions in a similar manner as DHCPv4, but the specific messages exchanged between the host and the server are unique for IPv6. State information for a particular cable modem can be obtained from the CMTS. Most vendors support some form of the CLI command “show cable modem” for this purpose, and this command is normally capable of providing state information that differentiates between IPv4 provisioning and IPv6 provisioning. For example, the CMTS defines the following (among other) cable modem states:

1. CMTS has received a DHCPv4 DISCOVER message from CM

dhcpv4done — IP address has been assigned

dhcpv6start — CMTS has received a DHCPv6 Solicit message from the CM

dhcpv6done — CMTS has sent a DHCPv6 Reply message to the CM

The DOCSIS 3.0 MDD is sent by the CMTS to all CMs and carries information about the HFC topology and other control information for the CM. The HFC topology is used in an efficient process that determines to which downstream service group the CM is connected. The MAC Domain Descriptor (MDD) is a MAC management message that defines plant topology and other parameters that are shared in a MAC domain and which the CM needs to know about. It is mandatory that the plant topology (fiber nodes [FN] and how they are split/combined) is configured in the CMTS. Ambiguity Resolution is an efficient process for determining which FN a CM is physically connected to (or more precisely, which downstream service group it is connected to).

MDD is defined in section 6.4.28 of the MULPI specification (CM-SP-MULPIv3.0-I01-060804).

Common DHCP Options for IPv6

DOCSIS 3.0 identifies a number of DHCP options that facilitate use of IPv6 in a cable network, more details of these options are provided in the provisioning section of the document. Some of these DHCP Options are defined as mandatory by DOCSIS specifications, including the following option: Identity Association for Non-temporary Addresses (IA_NA). The IA_NA option is used to assign an IPv6 management address to a given device. IA_NA is commonly employed for IPv6 Prefix Delegation, which is a means of assigning a prefix to a router or gateway device. The router or gateway device then automatically assigns addresses to connected devices using the delegated prefix.

Another option is the Vendor Class Option, the CableLabs specifications require a modem to use this option to identify itself as a DOCSIS 3.0 device. The Type Length Value (TLV) Operational Identifier (OID) 5 option is used by the CM to advertise its capabilities for a wide range of possible functions. Other DHCP Options are used for critical functions such as identifying the ToD and TFTP servers, and showing the configuration file name.

Not required but commonly used options include the following two examples:

- Rapid Commit Option – this enables the use of a streamlined process for a host to obtain its IP address
- Vendor-Identifying Vendor Specific Information Option (TLV 125). This option is actually used with DHCPv4 rather than DHCPv6. It includes various sub-options, including Sub-option 2 which allows for the specification of multiple IP addresses for the TFTP server.

TFTP-related Functions

Once the cable modem has successfully completed the DHCP process, it can communicate with a TFTP server and acquire its configuration or boot file. In a typical DOCSIS network, the CMTS serves as a proxy for the modem in its communication with the TFTP server. The CMTS forwards the modem's request for a config file to the TFTP server, and in some cases the CMTS buffers or stores the configuration file. This allows the CMTS to perform integrity checks on the config file itself and on the process by which the modem receives and installs the file. If the contents of the config file that is loaded on the modem have been modified, or if the config file loaded on the modem does not match that specified by the DHCP server, the CMTS is able to detect these anomalies and, if so configured, take action in response to a possible theft of service.

PROVISIONING FUNDAMENTALS

Provisioning Flow

There are five main steps for cable modem initialization on a DOCSIS network after the PHY layer has been established. These steps are:

1. Initialize the Layer 2 connection to the CMTS
2. Initialize IP Layer 3, including obtaining an IP address and other configuration information
3. Get the current time of day using the RFC 868 time protocol
4. Download the DOCSIS configuration file using TFTP
5. Complete the registration with the CM

The first step of the CM initialization is to establish Layer 2 connectivity. This includes the L2 DOCSIS 3.0 MDD message sent by the CMTS that controls the IP provisioning modes. The MDD message specifies the use of IPv4, IPv6 or APM as the preferred mode for the CM provisioning and management. The MDD message could also specify dual stack management.

IPv6 provisioning is new to DOCSIS 3.0 and includes several steps that are specific to IPv6 and because of this; DOCSIS 3.0 IPv4 provisioning is very similar to DOCSIS 2.0 provisioning.

One significant difference between IPv4 and IPv6 is that IPv6 includes SLAAC, which allows the device to use an IPv6 address without interacting with the DHCP server. DOCSIS 3.0 continues to use DHCP for IPv6 address assignment rather than using SLAAC. DHCPv6 minimizes the changes in the cable operator or MSO operational models and allows the cable operators or MSOs to have explicit control over the IP addresses assigned to CMs and CPEs. By continuing this explicit control over the IP address assignment, the cable operator or MSO can reliably configure control mechanisms to limit access to the cable operator or MSO network to only those devices that have been approved by the cable operators or MSOs for access.

In addition, the cable operator or MSO can update its DNS service from the DHCP server (Dynamic DNS) rather than relying on the CMs and CPEs to perform reliable and authorized updates to its DNS service. For backward compatibility with DOCSIS 2.0 devices, the IPv4 provision flow in DOCSIS 3.0 is the same as in DOCSIS 2.0. In the IPv6 protocol suite, RAs include bits that control the use of SLAAC and DHCP by hosts using IPv6. The DOCSIS 3.0 specification does not rule out the use of SLAAC in the future.

Initialization

There are an array of messages and actions used by the CM when it initializes its IPv6 stack. Here are a list of the messages and actions:

- The CM receives the DOCSIS 3.0 MDD message which directs the CM to use the IPv6 for provisioning
- Before assigning a Link-Local address to the CM management interface, the CM performs DAD on the Link-Local address (LLA) to confirm that the address is not already in use on the HFC
- The CM sends a NS message to the candidate Link-Local address
- If the CM receives no response from other devices on the link, the candidate Link-Local address, it is okay to use this address and the CM assigns it to its management interface. DAD will fail only in the case when some other CM or CPE has used the same EUI-64 interface identifier so that it responds to the NS message from the CM checking on its Link-Local address. This situation clearly represents a problem in the network because it means that two devices are using the same EUI-64 interface identifier and therefore are using the same MAC address. The CMTS may detect and report DAD failure.
- The CM assigns itself a Link-Local address. This step is part of a normal IPv6 initialization process. The CM forms its Link-Local address by combining the EUI-64 interface identifier with the reserve link-local prefix FE80::/10.
- Once the CM has assigned a Link-Local address to its management interface it can receive an RA message from the local IPv6 router/CMTS. To trigger the RA message, the CM sends an RS message to the IPv6 router/CMTS. The RA message is multicast to all CMs and CPEs so a single RA can initialize multiple CMs and CPEs, reducing the traffic on the network during CM and CPE initialization. The RA contains information about the IPv6 prefixes assigned to the network and control bits (M/O) directing the CM to use DHCPv6 for address assignment.
- The CM next initiates the DHCP message exchange by sending a DHCPv6 discover message called a “Solicit.”
- The CMTS acting as a DHCP relay receives the discover message or “Solicit” and forwards it in a relay forward message to the DHCPv6 server.
- The server chooses the IPv6 address and other configuration information for the CM and returns that information as an advertise message. The message to the CM includes a list of TFTP servers the CM should use, the name of the configuration file the CM should download, the RFC 868 time server the CM should use, and other configuration information. The complete list of parameters returned by the DHCPv6 server can be found in the DOCSIS 3.0 specification.
- The server sends that advertise in a relay reply message sent to the CMTS relay agent, which in turn forwards the advertise message to the CM.
- The CM continues the DHCPv6 message exchange by sending a request message to the relay agent which in turn forwards the request message in a relay forward message to the server.
- The DHCPv6 server responds with a reply message forwarded to the relay agent in a reply message. The DHCPv6 includes a mode of operation called Rapid Commit which emits the advertise and request messages from the DHCPv6 message exchange. This optimization streamlines the DHCPv6 message exchange into just 2 messages between the CM and DHCPv6 server forwarded through the relay agent. The use of rapid commit is negotiated between the DHCPv6 client (in this case the CM) and the server. The CM includes an option in its solicit message indicating it is willing to use the rapid commit mode. If the server is also willing to use rapid commit it immediately responds with a reply message rather than an advertise message. DOCSIS 3.0 requires the CM to include the rapid commit option in its initial solicit message. The cable operator or MSO network administrator then configures the DHCPv6 server to use rapid commit or to ignore the rapid commit request from the CM.

- Once the CM has received the reply message with its assigned IPv6 address, it uses DAD to confirm the address is not already in use. As with the Link-Local address, it is an error condition if the assigned address is already in use on the network and if the CM receives a response to its DAD process, it will restart the provisioning process.
- After completing IP Provisioning, the CM obtains the Time of Day (ToD) through the RFC 868 protocol with the request to the time server sent over the version of IP in use for provisioning. In this example, using IPv6 provisioning, the CM will send the ToD request message over IPv6. The CM is given the IPv6 addresses of available network time protocol (NTP) servers as an option in the DHCPv6 messaging from the DHCPv6 server. There are two DHCPv6 options for time configuration.
- One option provides a list of NTP servers to the host and is described in RFC 4075 [RFC4075].
- The other option provides time zone information to the host and is currently specified by the Internet Draft (http://tools.ietf.org/pdf/draft-ietf-dhc-dhcpv6-opt-tz-00.pdf) Timezone Specifier Option for DHCPv6.
- The next step in CM Provisioning is to obtain the configuration file, CL_OPTION_CONFIG_FILE_NAME. As in DOCSIS 2.0 the CM uses TFTP to download the configuration file from a configuration server, CL_OPTION_TFTP_SERVERS. The CM uses whichever version of IP it was directed to use for provisioning by the initial DOCSIS 3.0 MDD message. In this example, the CM will use TFTP over IPv6 to obtain its configuration file. The CM gets the name of its configuration file and a list of IPv6 addresses for the configuration servers as options in the DHCPv6 reply message from the DHCP V6 servers. The CM tries to download the configuration file from the first server in its list of servers. If the download fails, it will re-try from that same server. If it is still unable to download the configuration file, the CM will try the second TFTP server in the list. Therefore the list of TFTP servers provides additional reliability for configuration file download.
- The final step in CM provisioning is to register with the CMTS. If the CM is configured for dual stack management it will initialize the IPv4 protocol-stack after registration.

Dynamic Host Configuration Protocol

The version of DHCP for IPv6 is typically referred to as DHCPv6. For the most part, the DHCPv4 and DHCPv6 protocols are very similar in many aspects. Nevertheless, there are certain differences between the two.

DHCP Message	IPv4	IPv6
Initial Message Exchange	4 way handshake	
Client -> Server	Discover	Solicit
Server -> Client	Offer	Advertise
Client -> Server	Request	
Server -> Client	Acknowledge	Replay

Table Dynamic Host Configuration Protocol Message Comparison

As we see from the table above the initial message exchange is still a four-way handshake in both IPv4 and IPv6. The message types in IPv4 use broadcast and unicast packets, while in IPv6 they are multicast and unicast packets respectively. DHCPv6 is solely a layer three protocol in the OSI model, not simply an upgrade to DHCPv4; it is a separate and distinct protocol.

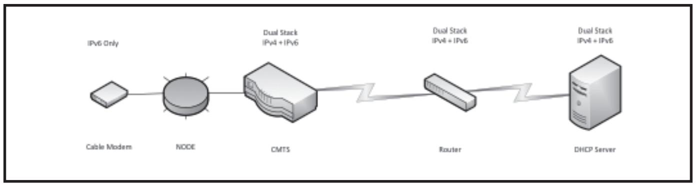


Figure 35: DHCP Flow

IPv6 can leverage a set of new tools for provisioning which are available only through IPv6 and not through IPv4. IPv6 can use the following DHCP-based mechanisms to provide addressing:

1. Stateful DHCPv6

2. Stateless DHCPv6 or DHCP Lite

3. DHCPv6 Prefix Delegation w/SLAAC - where

The latter is a mechanism available only through IPv6. IPv6 may also use the IPv6-specific SLAAC option where hosts learn the information necessary to acquire an address and additional operational information from the RA. IPv6 can also leverage its specific renumbering mechanism in order to facilitate the changing of prefixes on links. Hosts learn the preferred provisioning mechanism on the link from an RA.

Remember the Rapid-Commit is also available as part of the new DHCP protocol, this is a two-message exchange where the first message is a “SOLICIT” and the second is the “REPLY.”

Router Advertisement Flags

There are flags that may be used to manage the provisioning of devices on an IPv6 network. The RA message and enabled flags can be used to control the type of IP configuration stateless, stateful or both.

RA Flag	Meaning	Purpose
M	Managed Address Configuration	Used for Stateful DHCPv6, the client will begin to find a DHCPv6 server by generating a DHCP solicit message
O	Other Stateful Configuration	
A	Autonomous address-configuration	When the flag is set to “1” both configuration
L	On-link	

Table Router Advertisement Flags

The RA message flags setting will determine the type of autoconfiguration.

Meaning	M	O	A	L
Stateful DHCP, address and other operational configuration information is obtained	1	0	0	0
Stateless is where a prefix option is provided in the RA message	0	0	1	0
Stateless is where a prefix option is provided in the RA message and other options are enabled	0	1	1	0
Both Stateful and Stateless enabled. When the "M", "A" and "O" are set to "1" this indicates the preferred DHCP Prefix Delegation mechanism within the link.	1	1	1	0

Table Router Advertisement Configuration Options

Stateless DHCPv6 is used in conjunction with SLAAC in order
to provide additional operational information to hosts. A DHCPv6 server provides information and not address resources, so it does not need to maintain state for each client. A client which intends to acquire information by Stateless DHCPv6 mechanisms sends an Information-Request (IR) message to a DHCP server. The server responds with a Reply message which includes the requested information. The protocol server implementation for Stateless DHCPv6 is very simple and requires minimal resources so it can be easily deployed in routers rather than as a centralized service.

The SLAAC process is an IPv6-specific provisioning mechanism.

Hosts that follow the SLAAC process will go through the following steps:

1. Upon boot-up they will auto-configure a link-local address (FE80::/64)

2. They will join the solicited node multicast address corresponding to this auto-generated link-local address

3. And they will perform DAD to verify the uniqueness of the link-local address

Once the new local address is verified, the host can perform RD and the host can learn from the RAs. RAs allow the host to learn the configuration parameters used on the link, the operational prefixes valid on the link, and the preferred provisioning mechanisms for each prefix. For prefixes that prefer the use of SLAAC for provisioning, the host can auto-generate another set of addresses of a larger scope such as global or unique local. DAD will be used again to verify the uniqueness of these new auto-generated addresses. For prefixes that do not recommend the use of SLAAC for provisioning, the host will use Stateful DHCPv6. On the other hand, SLAAC hosts can also use Stateless DHCPv6 to obtain additional information in addition to addressing.

The DHCPv6 Prefix Delegation (DHCPv6 PD) protocol is a natural extension of DHCPv6. In the IPv6 world where subscribers are provided not just single addresses but entire prefixes, the DHCPv6 PD offers leases for prefixes to the subscriber's requesting router, independent of the media access type and network structure on the subscriber's premises. DHCPv6 operates in flexible deployment models such as client server or client relay server.

The operation of DHCPv6 PD is very similar to DHCPv6. The difference is that the client is a requesting router and the request is for a prefix instead of an address. The server is the delegating router as the ISP. In conjunction with other IPv6 features, such as general prefixes, DHCPv6 PD can prove to be a very powerful provisioning mechanism by providing DNS and NTP server information to CPE.

Host renumbering is done by modifying the RA to announce the old prefix with a short lifetime and the new prefix with a long lifetime.

The DHCPv6 message consists of a header and a set of options. The header contains a message type field and a transaction ID field. The first option in a DHCPv6 packet is a DHCP Unique Identifier (DUID); the 10 byte DUID (00:03:00:01:01:02:03:04:05:06) uniquely identifies a client. It is permanent and it does not change even if the network interface hardware is changed. The DUID will be used in the DHCPv6 4-way handshake messaging. RFC 3315 (Stateful DHCP) identifies three types of DUIDs. The first is generated from a link-layer address and a time stamp; it is generated and stored at first power-up. The second type is vendor-assigned and is based on an Enterprise-unique Number; it must be stored as a read-only value. The third type is based on a link-layer address and it is only used if the network interface hardware cannot be changed.

The second option in the DHCPv6 message is the Identity Association (IA) which represents a bundle of addresses and prefixes, managed under a given interface, that are assigned to a client. The IA is unique for each interface; however an interface can use multiple IAs at the same time. The addresses and prefixes under a given bundle have their own characteristics and have their own lifetimes but they are managed together and, for example, they are renewed at the same time. The IA includes various information such as an identifier, lease extension times, address options and various other options.

From an operational perspective, DHCP uses UDP packets for message exchange. Clients listen to UDP port 546 while Servers and Relay Agents listen to UDP port 547. Clients multicast their messages to the DHCP relay agent and server's address – FF02::1:2 – a lean scope multicast address to which all servers and relay agents subscribe. All DHCP Servers use multicast "FF05::1:3". This could be used by a relay agent to communicate with servers, either because the relay agent wants to send messages to all servers or because it does not know the unicast addresses of the servers. Note that in order for a relay agent to use this address, it must have an address of sufficient scope to be reachable by the servers. All servers within the site are members of this multicast group.

A client can target a specific server in its messages as long as it includes that server's specific DHCPv6 DUID. A client never needs a global re-routable address in order to be able to communicate with a DHCP server – just a link-local address (FE80::/64). That's due in part to the use of multicast addresses in its messages. Nevertheless a server may allow a client to also use unicast for its communication with the server. Servers and relay agents will always respond using unicast packets.

The DHCPv6 message header contains a message type field. The following message types have been identified:

1. SOLICIT: A client sends a Solicit message to locate servers.
2. ADVERTISE: A server sends an Advertise message to indicate that it is available for DHCP service, in response to a Solicit message received from a client.
3. REQUEST: A client sends a Request message to request configuration parameters, including IP addresses, from a specific server.
4. A client sends a Confirm message to any available server to determine whether the addresses it was assigned are still appropriate to the link to which the client is connected.
5. A client sends a Renew message to the server that originally provided the client's addresses and configuration parameters to extend the lifetimes on the addresses assigned to the client and to update other configuration parameters.
6. REBIND: A client sends a Rebind message to any available server to extend the lifetimes on the addresses assigned to the client and to update other configuration parameters; this message is sent after a client receives no response to a Renew message.
7. A client sends a Reply message containing assigned addresses and configuration parameters in response to a Solicit, Request, Renew, or Rebind message received from a client. A server sends a Reply message containing configuration parameters in response to an Information-request message. A server sends a Reply message in response to a Confirm message confirming or denying that the addresses assigned to the client are appropriate to the link to which the client is connected. A server sends a Reply message to acknowledge receipt of a Release or Decline message.

8. A client sends a Release message to the server that assigned addresses to the client to indicate that the client no longer uses one or more of the assigned addresses.
9. DECLINE: A client sends a Decline message to a server to indicate that the client has determined that one or more addresses assigned by the server are already in use on the link to which the client is connected.
10. RECONFIGURE: A server sends a Reconfigure message to a client to inform the client that the server has new or updated configuration parameters, and that the client is to initiate a Renew/Reply or Information-request/ Information Reply transaction with the server in order to receive the updated information.
11. INFORMATION-REQUEST: A client sends an Information-request message to a server to request configuration parameters without the assignment of an IP address to the client.
12. RELAY-FORWard: A relay agent sends a Relay-forward message to relay messages to servers, either directly or through another relay agent. The received message, either a client message or a Relay-forward message from another relay agent, is encapsulated in an option in the Relay-forward message.
13. RELAY-REPL: A server sends a Relay-reply message to a relay agent containing a message that the relay agent delivers to a client. The Relay-reply message may be relayed by other relay agents for delivery to the destination relay agent.

The server encapsulates the client message as an option in the Relay-reply message, which the relay agent extracts and relays to the client.

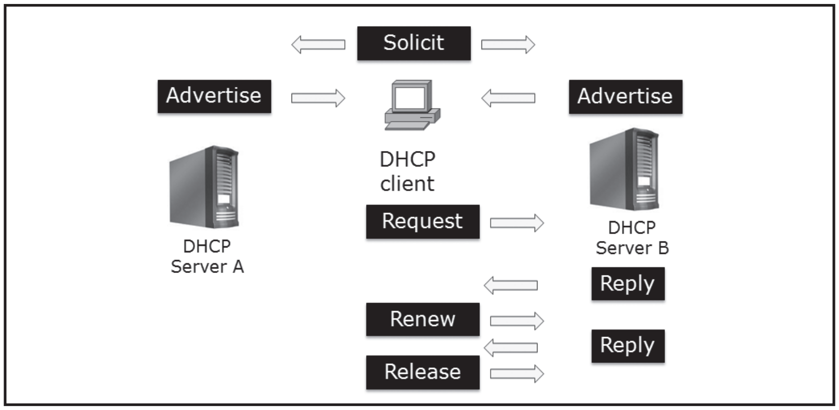
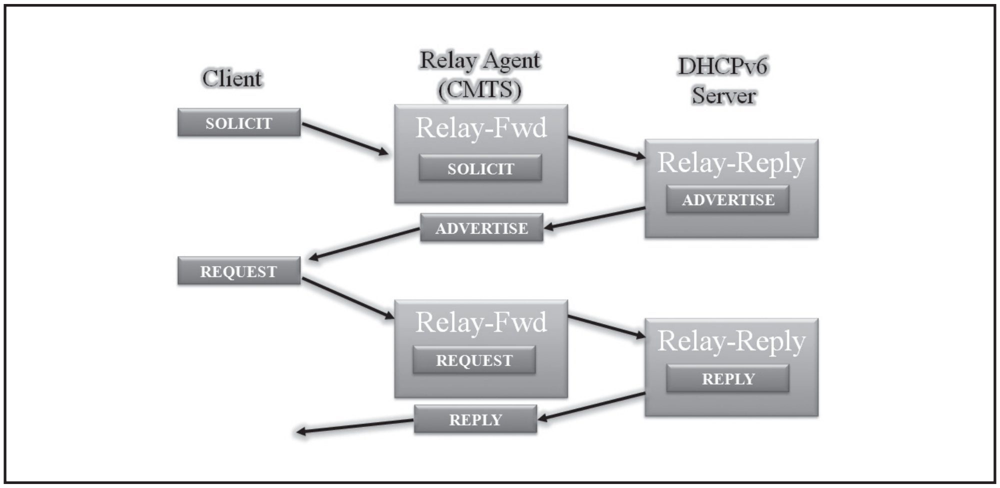


Figure 36: DHCPv6 message types and exchange process

This figure shows the message exchange during the basic operation of DHCPv6. A DHCP client sends a “Solicit” message in order to discover the DHCP servers. The solicit message includes the client DUID and a request option. It also includes an empty Identity Association for Non Temporary Addresses (IA_NA) option to hold addresses.

Both servers will receive the solicit message, and both will reply with an advertise message. In an advertise message, a server includes a client DUID, a server DUID, and other configuration information. It also adds assigned addresses to the IA_NA option. The client receives both advertise messages and selects one server for further communication. In this example, the client selects server B.

The client will send a “Request” message for the assigned addresses. The assignment is acknowledged by the server through a “reply” message. From this point on, the client can use the assigned addresses for the duration of the lease. The client can also renew the lifetime of the addresses by sending a renew message to the server. The renew message is acknowledged through a reply message. When the client decides to shut down, it sends a release message to the server to indicate that it will no longer use the addresses.



Additional messages are used in the case where a relay agent such as a Cable Modem Termination System (CMTS) is brokering the communication between the client and the server.

- The client solicit message is included by the relay in a Relay Forward message and sent to the server
- The server replies with a Relay Reply that includes an Advertise message
- The Advertise message is in turn forwarded to the client. The client will respond with a Request which the relay is going to include in a Relay Forward message.
- The server responds with a Relay reply which includes the Reply and the reply finally makes it to the client

A DOCSIS 2.0+ IPv6 cable modem is required to use DHCPv6 in order to acquire its management address just like a DOCSIS 3.0 modem must. In fact, the CableLabs DOCSIS 2.0+ IPv6 cable modem specification refers in many cases to DOCSIS 3.0 standard documents such as MAC and Upper Layer Protocols Interface (MULPI).

It is important to note that multiple levels of relays are supported by the protocol. Also, the relay server communication can be secured by the use of IPSec.

Trivial File Transfer Protocol

The TFTP is a protocol used by operators to deliver configuration files and firmware to devices such as cable modems and other types of DOCSIS capable devices. Once the cable modem has successfully completed the DHCP process, it has the information needed to communicate with a TFTP server to acquire its configuration or bootfile. In a DOCSIS network, the CMTS may serve as a proxy for the modem in its communication with the TFTP server. The CMTS forwards the modem's request for a configuration file to the TFTP server, and in some cases the CMTS buffers or stores the configuration file on behalf of the requesting device. This allows the CMTS to perform integrity checks on the configuration file itself and on the process by which the modem receives and installs the file. If the contents of the configuration file that is loaded on the modem have been modified, or if the configuration file loaded on the modem does not match that specified by the DHCP server, the CMTS is able to detect these anomalies and, if so configured, take action in response to a possible theft of service.

Time of Day

The ToD protocol is used by operators in the next step of the provisioning process for the cable modem is to establish Time of Day. The cable modem will request the ToD from a DOCSIS provisioning server called the ToD server. The ToD server is used in the cable network to provide the time of day to cable modems during the registration process. After DOCSIS 1.0 ToD became an optional parameter however, it is important to understand operators find this critical to the overall timing accuracy of the cable network.

Domain Name Service

DNS facilitate the use of IP communications. Humans are accustomed to assign names to resources while IP is using IP addresses in order to identify resources. DNS is a distributed database covering multiple, locally managed computers which store the mapping between names and IP addresses and between IP addresses and names.

In a nutshell, Domain Name System (DNS), is a distributed database which manages Resource Records (RRs). The RRs are of various types such as IPV4 and IPV6. They contain information such as Start of Authority (SOA), Name Server (NS), Addresses – which in the case of IPv4 are stored in A records and in the case of IPv6 are stored in AAAA records – and Pointers.

DNS runs on top of IP. It uses both UDP and TCP packets via Port 53. Additional information on DNS can be found in the list of references.

A user intending to acquire IPv6 information from SCTE's web page will type in the browser "www.scte.org/ipv6". DNS will translate this name into an IPv6 address that the computer can use in order to reach this information resource via an IP.

It is important to note that in the case of IPv6, DNS is expected to be more heavily used. IPv6 addresses are extremely long and almost impossible to remember. So, it is expected that users will be more likely to use names in order to talk to other users or to discover information resources. The scalability implications of using DNS with such a large address space will be very interesting.

In order to perform reverse mapping and to support normal recursive and iterative queries, DNS reserves a special domain name. In the case of IPv4, this domain name is called IN-ADDR.arpa. In IPv6, this domain name is IP6.arpa. With the exception of the difference in the domain names, the operation of reverse DNS for both IPv4 and IPv6 is the same. With the integration of IPv6 in today's networks, the challenges do not come from the introduction of new concepts with respect to DNS. They come from the coexistence of DNS for the two protocols, IPv4 and IPv6.

In a world where IPv4 and IPv6 coexist, a resource may be reachable by either IPv4 or IPv6 so its name will bind to both an IPv4 and an IPv6 address.

In the DNS database, this resource will be represented by an A record and a quad-A record.

Resolving DNS with a dual stacked host which intends to reach this resource will send a query to the DNS server corresponding to the resource name. A DNS server will reply with both IPv4 and IPv6 addresses that map to this name. The host will then have to decide which address it will use in order to reach the targeted resource. Understanding the address selection mechanisms and overall operation of DNS in a dual-stacked environment is essential for successful IPv6 deployments.

Cable Modem Termination System Requirements

CMTS support for IPv6 is wide ranging and essential to a successful deployment. CMTS platforms must offer parity between IPv4 and IPv6 from an operational and performance point of view. Further, cable specific functionality to facilitate a successful deployment of IPv6 for device management requires additional functionality specifically CMTS platforms must also minimally support full native dual stack operations along with different IP provisioning modes.

A CMTS must include several new capabilities to support IPv6 in a DOCSIS 3.0 deployment. The CMTS can act as a bridge or a router. If it acts as a bridge, there must be a router on the HFC to provide the required IPv6 protocols. The CMTS is required to act as a relay agent for DHCPv6 messages sent between the CM and the DHCPv6 server. This relay agent function is very similar to the DHCPv4 relay agent function in DOCSIS 2.0. The relay agent inserts some relay agent options forwarded to the server to provide additional information about the CM, such as the CM's MAC address. The full set of relay agent options is given in the DOCSIS 3.0 specification. The DHCPv6 server also sends some relay agent options to the relay agent.

CMTS platforms will need to support four (4) possible provisioning modes also referred to as IP provisioning mode Type Length Values (TLVs) including IPv4 only, IPv6 only, APM, or DPM. The outlined DOCSIS functionality specific to IPv6 and provisioning along with foundational IPv6 support will allow for predictable migration and management of DOCSIS devices leveraging IPv6. In conjunction with the mandatory CMTS platform support for IPv6, cable modems must also support IP provisioning mode override, which is also referred to as MDD IP Mode Override (MIMO).

The CMTS router implements the IPv6 ND for DAD and address resolution and sends RA messages to control the IP address configuration used by CMs and CPEs.

DOCSIS 3.0 includes the use of IPv6 multicast and the CMTS must be capable of using Any Source Multicast (ASM) and Source-Specific Multicast (SSM) as well as meeting the requirements for multicast from IPv6 control tasks such as ASM, SSM, and forwarding IPv6 control traffic such as MDD, DAD and RA.

Finally, the CMTS includes backward compatibility with DOCSIS 2.0 so that legacy modems can interoperate with DOCSIS 3.0 CMTSs.

Bridge Requirements

In DOCSIS 3.0, the CM can operate either as a bridge as in DOCSIS 2.0 or as a router. Here we'll look at the requirements for IPv6 functions when the CM functions acts as a bridge.

The CM uses DHCPv6 for address assignment and other configuration information during provisioning. It parses the DOCSIS 3.0 MDD messages to determine whether the CM should use APM and whether the CM should configure itself for dual stack management. To accommodate dual stack management, the CM could be managed either over IPv4 or over IPv6 simultaneously, and the CM can operate both its IPv4 and IPv6 protocol stacks at the same time.

The version of IP used by the CM for provisioning is independent of data forwarding so the CM forwards both IPv4 and IPv6 traffic for the CPEs. A DOCSIS 3.0 CM is required to be backwards compatible with DOCSIS 2.0 CMTSs so the DOCSIS 3.0 CM can use DOCSIS 2.0 IPv4 provisioning when deployed with a DOCSIS 2.0 CMTS. This allows IPv4 and IPv6 data forwarding from CPEs, regardless of how the CM is provisioned.

A DOCSIS 2.0+ IPv6 cable modem is required to use DHCPv6 in order to acquire its management address just like a DOCSIS 3.0 modem must. In fact, the CableLabs DOCSIS 2.0+ IPv6 cable modem specification refers in many cases to DOCSIS 3.0 standard documents such as MULPI.

The bridge will be managed via SNMP over IPv4 or IPv6 or dual stack IPv4 and IPv6.

eRouter Requirements

Based on experience with IPv6 service deployed elsewhere, the architecture for IPv6 service in DOCSIS 3.0 provides for a network in the subscriber premises that uses global IPv6 addresses for complete connectivity between the CPEs and the Internet. The architecture includes a component called the eRouter which is a full IPv6 router imbedded in the CM. The eRouter provides IPv6 routing and other protocol functions in support of the operation of the subscriber network. The eRouter includes an implication of DHCPv6 based Prefix Delegation or DHCPv6PD through which the ISP assigns the subscriber prefix to the eRouter. The eRouter then subdivides that assigned pre-fix to assign prefixes to the links in the subscriber network.

Typically the eRouter will direct the CPEs on the subscriber network to use IPv6 SLAAC. The eRouter may include a DHCPv6 server function that can provide address assignment if the subscriber wants to use DHCPv6 in the home network. For example, if the home network includes multiple IPv6 routers, the DHCPv6 server in the eRouter can use DHCPv6PD to configure the other home network routers. The eRouter implements ND and RA queries for home CPE as part of the IPv6 protocol suite. In some deployments, CPE may need specific information from the MSO such as IPTV servers for STBs. If the eRouter function is enabled, the eRouter will respond to any DHCPv6 messages from the STBs and the cable operator's DHCPv6 server will not be able to pass that configuration information to the STB.

DOCSIS 3.0 includes the DHCPv6 configuration options (DNS, ToD, TFTP servers) through the cable operator's DHCPv6 server which can deliver additional information to the eRouter's DHCPv6 server. The eRouter's DHCPv6 server then delivers this information to the STBs. The eRouter participates in both versions of multicast listener discovery (MLDv1/v2) protocol. Finally, similar to the CM acting as a bridge, the CM acting as a router is backwardly compatible with DOCSIS 2.0 CMTSs.

Simple Network Management Protocol

Once the network is addressed the cable operator will need to manage the network. The most common management system used by cable operators today is called a Network Management System, or NMS. IP based NMS utilize the Simple Network Management Protocol, or SNMP packets and is by far the most widely accepted application-layer protocol tool for monitoring, securing or managing cable operator TCP/IP network devices. SNMP provides a standardized framework and a common protocol used for the monitoring and management of devices in a cable network. An example is configuring the CM SNMP remote query settings allowing a software package, operated from a management client, to query the CM agent's performance statistics, reporting conditions that may warrant administrative intervention. By polling the CM periodically using SNMP and caching these parameters and configuration (e.g. IP address, MAC address, S/N ratio and RF metrics), operators will know at a glance the state of a single modem and overall status of the access network. SNMP can also be used with routers, switches, CMTS, VoD servers, fiber transport and encoders.

Multicast DSID Forwarding

Another fundamental construct of DOCSIS 3.0 which is required for IPv6 is Multicast DSID Forwarding (MDF). MDF is a means by which the modem filters and forwards downstream multicast traffic. Multicast DSID Forwarding is required to support IPv6 configuration of CPE devices and for some user-joined multicast (using IGMPv3 or MLD protocols).

There are three modes of Multicast DSID Forwarding: GMAC Promiscuous Multicast DSID Forwarding, GMAC Explicit Multicast Forwarding, and MDF Disabled (where GMAC refers to the Group MAC address of the multicast packets). Both GMAC Promiscuous Multicast DSID Forwarding and GMAC Explicit Multicast DSID Forwarding Modes are considered modes in which MDF is "enabled." When MDF is "enabled," the CM filters and forwards all multicast traffic based on the DSID value. In GMAC Explicit MDF Mode, the modem additionally filters multicast traffic on the GMAC.

A DOCSIS 3.0 cable modem must support the GMAC Promiscuous MDF. A DOCSIS 2.0+IPv6 cable modem must support either GMAC Explicit MDF or a proprietary means of forwarding the multicast traffic required for IPv6 provisioning.

The CMTS enables Multicast DSID Forwarding during the registration process. When Multicast DSID forwarding is enabled, the CMTS communicates the Multicast DSIDs to the CM in the Registration Response message or Dynamic Bonding Change message. The CMTS then labels all multicast traffic intended for a CM with a DSID value known by that CM.

Multicast DSID forwarding must be enabled in order to provide IPv6 provisioning of CPE devices. This is because the CM cannot forward the multicast traffic required for IPv6 provisioning unless multicast DSID Forwarding is enabled and the CMTS labels the IPv6 provisioning multicast with a known DSID. If MDF is disabled, the multicast traffic required for IPv6 provisioning will not be passed to CPE devices behind the CM.

CONCLUSION

From the early days of DOCSIS 3.0 where development to introduce support for IPv6 was in its infancy, IPv6 has clearly evolved. As outlined in this best practice, the preparation, deployment, and enablement of IPv6 requires a great deal of planning and coordinated execution, especially when it is imperative to ensure that the introduction of IPv6 support is seamless and minimally disruptive. The goal of this best practice was to provide an understanding to the key areas that adopters of IPv6 must consider when organizing their IPv6 deployment efforts.

APPENDIX A: ABBREVIATIONS AND ACRONYMS

6RD	6 Rapid Deployment
AFTR	Address Family Transition Router
ALG	Application Layer Gateways
APM	Alternate Provisioning Mode
CGN	Carrier Grade NAT
DAD	Duplicate Address Detection
DG	Default Gateway
DHCPv6 PD	DHCPv6 Prefix Delegation
DNS	Domain Name Service
DPM	Dual-Stack Provisioning Mode
DS	Differentiated Services
GSM	Global System for Mobile
GUA	Global Unicast Addressing
GWR	Gateway Router
IAB	Internet Architecture Board
IANA	Internet Assigned Number Authority
ICANN	Internet Corporation for Assigned Names and Numbers
ICE	Interactive Connectivity Establishment
IGP	Interior Gateway Protocols
IMS	IP Multimedia Subsystem
IPSec	Internet Protocol Security
ISATAP	Intrasite Automatic Tunnel Addressing Protocol
LLA	Link-Local Addresses
LSN	Large Scale NAT
MAC	Media Access Control
MDD	MAC Domain Descriptor
MDF	Multicast DSID Forwarding
MIMO	MDD IP Mode Override
MP-BGP	Multiprotocol Border Gateway Protocol version 4
MTU	Maximum Transmission Unit
MULPI	MAC and Upper-Layer Protocols Interface

NA	Neighbor Advertisement
ND	Neighbor Discovery
NRO	Number Resource Organization
NS	Neighbor Solicitation
NUD	Neighbor Unreachability Detection
OSPF	Open Shortest Path First
PMTU	Path Maximum Transmission Unit
RA	Router Advertisement
RR	Resource Records
RS	Router Solicitation
SBC	Session Border Controllers
SEND	SEcure Neighbor Discovery
SLAAC	State-Less Address Auto-Configuration
SNMP	Simple Network Management Protocol
SoA	Start of Authority
ToS	Type of Service
TrGW	Transition Gateways
ULA	Unique Local Addressing
VoIP	Voice over IP

APPENDIX B: IPV6 RELEVANT DOCUMENTS

RFC 2460 Internet Protocol, Version 6 (IPv6) Specification

RFC 3306 Unicast-Prefix-based IPv6 Multicast Addresses

RFC 3315 Dynamic Host Configuration Protocol for IPv6

RFC 3587 IPv6 Global Unicast Address Format

RFC 3633 IPv6 Prefix Options for Dynamic Host Configuration Protocol

RFC 4193 Unique Local IPv6 Unicast Addresses RFC 4291 IP Version 6 Addressing Architecture

RFC 4861 Neighbor Discovery for IP version 6

RFC 4862 IPv6 Stateless Address Autoconfiguration

RFC 5156 Special Use IPv6 Addresses

RFC 6106 IPv6 Router Advertisement Options for DNS Configuration

LIST OF FIGURES

SCTE Operational Practice Series

IPv6 Deployment Best Practices: Fundamentals

With today's rapid growth of the Internet it is evident that far more IP addresses are necessary to connect new devices in the future than the IPv4 address space has available. Internet Protocol version 6 (IPv6) is the latest version of the communications protocol that provides identification and location for computers on networks and routes traffic across the Internet. IPv6 was developed to deal with the long-anticipated problem of IPv4 address exhaustion.

Cable system operators need additional address space to scale networks for the future and IPv6 solves that by providing more address space than IPv4. IPv6 is a 128 bit/16 bytes or 32 hexadecimal character address as compare to IPv4 which is a 32 bit or 4 bytes decimal address. 96 more bits than IPv4! How big is IPv6 you ask? In IPv6, there are 340 trillion, trillion, trillion addresses or 10x10^38 addresses. That's 340,282,366,920,938,463,463,374,607,431,768,211,456 addresses.

The Society of Cable Telecommunications Engineers, Inc. (SCTE) has produced *IPv6 Deployment Best Practices: Fundamentals, First Edition*, to prepare cable professionals for the next version of Internet Protocol. Unlock the potential of your cable network by learning the fundamentals and features of IPv6, including, migration techniques, routing, security, DOCSIS® fundamentals, and provisioning.

SCTE®
Society of Cable
Telecommunications
Engineers

140 Philips Road
Exton, PA 19341-1318
www.scte.org

SCTE Product Code: TS53
ISBN #: 1-930634-33-1

SCTE Mission—Providing technical leadership for the telecommunications industry and serving its members through excellence in professional development, standards, certification, and information.

SCTE is proud to serve as the technical and applied science leader for the cable telecommunications industry. Since its beginning in 1969, the Society has been dedicated to providing meaningful resources and programs for its members and the industry.