Cisco Self-Study:

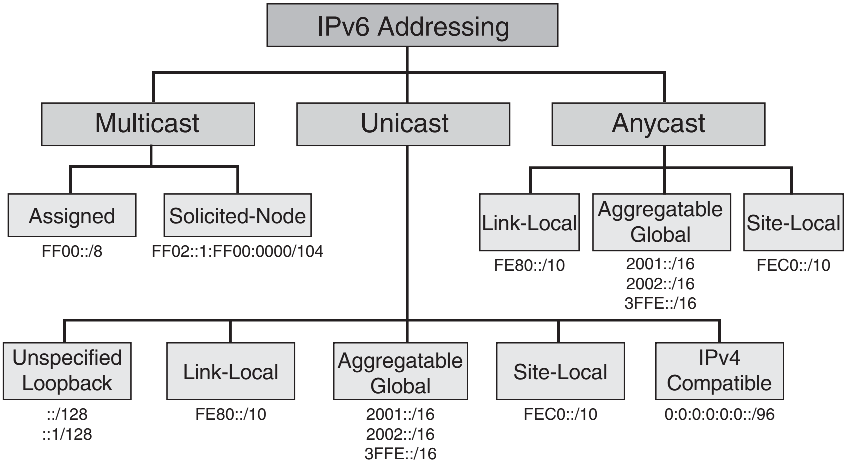
Implementing Cisco IPv6 Networks (IPV6)

Design, build, configure, and support networks based on Version 6 of the Internet Protocol

ciscopress.com

Edited by: **Régis Desmeules**

Types of Addresses of the IPv6 Addressing Architecture



Cisco Self-Study: Implementing IPv6 Networks (IPV6)

Régis Desmeules

Cisco Press

Cisco Press
201 West 103rd Street
Indianapolis, IN 46290 USA

Cisco Self-Study: Implementing IPv6 Networks (IPV6)

Régis Desmeules

Published by:
Cisco Press
201 West 103rd Street
Indianapolis, IN 46290 USA

Printed in the United States of America 2 3 4 5 6 7 8 9 0

Second Printing August 2005

Library of Congress Cataloging-in-Publication Number: 2002100664

ISBN: 1-58705-086-2

Warning and Disclaimer

This book is designed to provide information about implementing Cisco IPv6 networks. Every effort has been made to make this book as complete and accurate as possible, but no warranty or fitness is implied.

The information is provided on an “as is” basis. The author, Cisco Press, and Cisco Systems, Inc. shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc. cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members of the professional technical community.

Reader feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through e-mail at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Publisher	John Wait
Editor-In-Chief	John Kane
Cisco Systems Program Manager	Anthony Wolfenden
Manager, Marketing Communications, Cisco Systems	Scott Miller
Executive Editor	Brett Bartow
Acquisitions Editor	Michelle Grandin
Managing Editor	Patrick Kanouse
Development Editor	Ginny Bess
Project Editor	Marc Fowler
Copy Editor	Gayle Johnson
Technical Editors	Bruno Ciscato Patrick Grossetete Jun-ichiro Itojun Hagino Casimir Samanasu Saeed Bin Sarder
Team Coordinator	Tammi Ross
Book Designer	Gina Rexrode
Cover Designer	Louisa Adair
Production Team	Octal Publishing
Indexer	Tim Wright

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the
Cisco.com Web site at www.cisco.com/go/offices.

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Czech Republic
Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy
Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal
Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Printed in the USA

About the Author

Régis Desmeules is an independent consultant. His specialties include IPv4, IPv6, network architecture and design, security, DNS, multimedia, Cisco routers, LAN switches, UNIX, and Microsoft implementations. He has developed and taught courses related to IPv4, IPv6, multimedia over IP, security, DNS, and MobileIP. He has taught courses in Canada and at different events such as INET, IPv6 Forum, Internet2, and Networld+ Interop. Desmeules was also a consultant for Viagénie, Inc. There he participated in IPv6 projects such as the deployment of IPv6 backbones on CA*net2 and CA*net3; the development and operation of Freenet6.net, one of the first tunnel servers for IPv6; the deployment of a stealth IPv6 DNS root server on the 6bone; the IPv6 Internet exchange called 6TAP and the network game Quake over IPv6. He has served as a collaborator on the IPv6 course and training that were designed for the IOS Learning Services Group at Cisco Systems, Inc. Before working for Viagénie, Desmeules worked for the largest distance education university in Canada, where he built large data, voice, and videoconference networks. He lives in the peaceful town of Quebec City in Canada.

About the Technical Reviewers

Bruno Ciscato is a networking consultant focusing on IPv6. Prior to that, he worked with Cisco Systems, Inc., for six years designing service provider networks and leading the 6net Project, a large-scale IPv6 test bed for the European scientific community funded by the European Commission. He lives in Italy and loves wine and sailing.

Patrick Grossetete is a senior product manager at Cisco Systems. In the Internet Technology Division (ITD), he is responsible for the IPv6 strategy on Cisco IOS. Prior to that, he was Field Distinguished Engineer on the EMEA consulting team, doing network design for customers. He regularly presents on behalf of Cisco at various IPv6 events and represents Cisco at the IPv6 Forum. Before joining Cisco Systems in 1994, Grossetete worked in remedial support and as a network consultant for Digital Equipment, where he was involved with LAN products, ATM, and DECnet/OSI architecture. He lives in France with his wife and two children.

Jun-ichiro Itojun Hagino is a researcher at IIJ, one of the biggest ISPs in Japan. He serves as a co-chair of the IETF v6ops working group. He received a PhD from Keio University for research on object-oriented operating systems.

Casimir Samanasu is a Program Manager with Cisco Systems, Inc. He holds an M.S. Computer Science degree from DePaul University, Chicago, and an MBA from the University of Dallas. Casimir has developed LAN switching courses in the past and was responsible for Cisco IOS curriculum that included advanced technologies such as QoS, Multicast, Security, and VPN. As a curriculum manager, Casimir was recently responsible for the development of training for IPv6 and Mobile IP technologies, and created the "ABCs of IP Version 6" document.

Saeed Bin Sarder has worked in Cisco Systems' High-Speed Switching Group as a development resting engineer for more than two years. His group is responsible for testing control and data plane issues in IOS on Catalyst 6000 products, including IPv4, IPv6, MPLS, QoS, and IP multicast routing and hardware forwarding on a variety of LAN and WAN modules.

Dedication

I would like to dedicate this book to my wife, Caroline, son, Olivier, and daughter, Sarah, for their unconditional support during the nights and weekends dedicated to this book.

A special thank-you goes to my friends Marc Blanchet, Florent Parent, and Hélène Richard at Viagénie, who provided suggestions, help, and encouragement during the writing. They all contributed to the development of tutorials and courses from which this book is inspired.

Acknowledgments

I would like to say thank you to Bill St-Arneau and his team at Canarie in Canada, who have provided support and belief since the early development stages of the IPv6 protocol.

Thanks to the Cisco Press staff, especially Michelle, Chris, Ginny, and all the technical reviewers who contributed to this book—Patrick, Itojun, Cas, Bruno, and Saeed.

Contents at a Glance

Table of Contents

Introduction

Introduced by the IETF in 1992, IPv6 appears today to be a fundamental and well-engineered solution to the IPv4 addressing space shortage. IPv6 is significantly more efficient than IPv4 because its design is based on the past 20 years experience of the IPv4 protocol.

With IPv6, we have to change our thinking, because this protocol was not only designed for computers on networks such as the current IPv4 Internet. The applicability of IPv6 is global to communication devices such as cellular, wireless, phones, PDAs, television, radio, and so on rather than being limited to computers.

One main goal of IPv6 is to make the router the key element of any network by simplifying the deployment, operation, and management of any IP-based network. Moreover, IPv6 is more advanced and scalable than IPv4 for global networks that will be made with billions of nodes, such as the 3G infrastructure. Some advantages of IPv6 are a larger address space, a simpler header, autoconfiguration, renumbering, aggregation, multihoming, transition, and coexistence with the existing IPv4 infrastructure.

In the long term, Internet gurus and high-level analysts agree that the Internet must be upgraded to IPv6. In fact, the ultimate goal of IPv6 is to completely replace IPv4. Therefore, the long-term market for IPv6 is huge, representing billions of nodes and networks all around the world.

Cisco Systems, Inc., is the world's leading supplier of internetworking hardware and software. Cisco has been involved in IETF IPv6 standardization since 1995, at the early stages of IPv6's design. Because Cisco technology carries about 80% of all Internet traffic, Cisco is obviously a key player in the worldwide deployment of IPv6.

NOTE Because it is difficult in this book to keep an updated list of IPv6 features that are or will by supported on the Cisco IOS Software technology for the different platforms, then, you are invited to look at the latest list of available features at www.cisco.com. You can find the latest list in the "Start Here: Cisco IOS Software Release Specifics for IPv6 Features" manual as well in the CCO Feature Navigator.

This Book's Objectives

Understanding the technical mechanisms of IPv6, the new IPv6 features available on Cisco IOS Software technology, and the interoperability between the Cisco routers and IPv6 implementations are fundamental to deploying scalable and reliable IPv6 networks.

Therefore, this book gives you a strong view of the Cisco IPv6 implementations, as well as an in-depth technical reference regarding designing, configuring, deploying, and debugging IPv6 on Cisco routers. You will gain expertise in IPv6 on Cisco technology through practical examples of all the IPv6 features presented in this book.

Who Should Read This Book

This book is intended for professionals in the enterprise and provider markets such as architects, network designers, systems engineers, network managers, administrators, and any technical staff. Professionals who are planning to use Cisco technology to deploy IPv6 networks, provide IPv6 connectivity, and use IPv6 within their network backbones need this book. You will find this book valuable because it provides many examples, figures, IOS commands, and tips for using IPv6 with Cisco IOS Software technology.

You will find everything you need to describe, design, configure, support, and operate IPv6 network backbones based on Cisco routers. To get the most out of this book, you need a minimal background in IPv4 and should be able to operate Cisco routers.

How This Book Is Organized

Although you could read this book cover-to-cover, it is designed to be flexible. You can easily move between chapters and sections to cover just the material you need more work with.

The book is divided into five parts. The first part explains the history of IPv6, the rationale behind it, and its benefits. The second part presents in detail the basic and advanced features of IPv6. Then it explains designing, enabling, configuring, and routing IPv6 networks using Cisco IOS Software technology. The third part describes the main integration and coexistence mechanisms and demonstrates integrating IPv6 on the current IPv4 infrastructure using different strategies. This part shows you examples of internetworking using Cisco IOS Software technology with different host implementations supporting IPv6. The fourth part describes the 6bone design and how this worldwide IPv6 backbone is operated. It also provides information that helps ISPs undertand the steps and rules of becoming an IPv6 provider on the IPv6 Internet. The fifth part contains the appendixes and a glossary.

The following list highlights the topics covered and the book's organization:

- **Part I: Overview of and Justification for IPv6**

  Chapter 1, "Introduction to IPv6"—This chapter introduces and provides an overview of the new IPv6 protocol. More specifically, it discusses the rationale of IPv6 by presenting the issues of IPv4 such as IPv4 address space exhaustion, the fast-growing global Internet routing table, and the many implications of using network address translation (NAT) mechanisms. This chapter also presents the history of IPv6 and provides an overview of IPv6 features such as larger address space, addressing hierarchy, aggregation, autoconfiguration, network renumbering, efficient headers, mobility, security, and the transition from IPv4 networks to IPv6.

- **Part II: IPv6 Design**

  Chapter 2, "IPv6 Addressing"—This chapter discusses the fundamentals of IPv6 and explains the application of basic IPv6 configurations on Cisco routers. More specifically, this chapter describes in detail the new IPv6 header, the IPv6 addressing architecture, the upper-layer protocols UDP and TCP, the representation of IPv6 addresses, and all types of addresses scoped in IPv6 such as link-local, site-local, and many others. This chapter also explains and provides examples of enabling IPv6 on a router, enabling and assigning IPv6 addresses to network interfaces, using the EUI-64 format to configure addresses, and verifying IPv6 configurations on interfaces.

  Chapter 3, "IPv6 in Depth" —This chapter is the key chapter of the book, because it describes IPv6's advanced features and mechanisms such as Neighbor Discovery Protocol (NDP), stateless autoconfiguration, prefix advertisement, duplicate address detection (DAD), the replacement of ARP, Internet Control Message Protocol for IPv6 (ICMPv6), path MTU discovery (PMTUD), the new AAAA record for the domain name system (DNS), DHCPv6, IPSec, and

Mobile IPv6. Then, to help you acquire strong practical knowledge of these advanced IPv6 features, Chapter 3 covers enabling and managing prefix advertisement on Cisco, renumbering a network, and defining IPv6 standard and extended access control lists (ACLs). It also provides examples of verifying, managing, and debugging IPv6 configurations on Cisco routers using IPv6-enabled tools and commands such as **show**, **debug**, **ping**, **traceroute**, **Telnet**, **ssh**, and **TFTP** which are EXEC commands of the IOS.

Chapter 4, "Routing on IPv6"—This chapter explains the differences between the EGP and IGP routing protocols for IPv6 by comparing them to their IPv4 equivalents. As in IPv4, routing protocols are fundamental for the IPv6 routing domains. Chapter 4 starts by presenting an overview of the updates and changes applied on these routing protocols to support IPv6. This chapter covers the interdomain routing protocol BGP4+ and the intradomain routing protocols RIPng, IS-IS for IPv6, and OSPFv3. This chapter also discusses and provides examples of enabling, configuring, and managing these IPv6 routing protocols on Cisco routers. More pratically, it covers configuring static and default IPv6 routes, enabling and configuring BGP4+ with IPv6, establishing multihop BGP4+ configuration, configuring BGP4+ to exchange IPv4 routes between BGP IPv6 peers, configuring prefix filtering and route maps for IPv6 with BGP4+, using link-local addresses with BGP4+, configuring RIPng, enabling and configuring IS-IS and OSPFv3 for IPv6, and redistributing IPv6 routes into BGP4+, RIPng, and IS-IS for IPv6 and OSPFv3. The last section of the chapter presents the commands used in Cisco Express Forwarding for IPv6 (CEFv6). It also describes managing some of these routing protocols using the **show** and **debug** commands.

- **Part III: IPv4 and IPv6: Coexistence and Integration**

  Chapter 5, "IPv6 Integration and Coexistence Strategies"—This chapter covers the main integration and coexistence strategies provided in IPv6 to maintain complete backward compatibility with IPv4 and to allow a smooth transition from IPv4 to IPv6. The integration and coexistence strategies presented in this chapter include the dual-stack approach; the multiple protocols and techniques of tunneling IPv6 packets over IPv4 networks, such as configured tunnel, tunnel broker, tunnel server, 6to4, GRE tunnel, ISATAP, and automatic IPv4-compatible tunnel; and IPv6-only-to-IPv4-only transition mechanisms such as the application-level gateway and NAT-PT. In addition, this chapter covers enabling the dual stack, enabling a configured tunnel, enabling 6to4, using a 6to4 relay, deploying IPv6 over GRE, enabling ISATAP tunnels, enabling NAT-PT, and applying static and dynamic NAT-PT configurations. This chapter also provides examples of verifying and debugging some of these transition techniques.

  Chapter 6, "IPv6 Hosts Internetworking with Cisco"—This chapter covers enabling and configuring IPv6 support on Microsoft Windows NT, 2000, and XP; Solaris 8; FreeBSD 4.x; Linux; and Tru64 UNIX to internetwork with Cisco IOS Software technology. You see examples of internetworking using stateless autoconfiguration, the dual-stack approach, configured tunnel, and 6to4 between the IPv6 host implementations and Cisco routers.

- **Part IV: The IPv6 Backbone**

  Chapter 7, "Connecting to the IPv6 Internet"—This chapter discusses how the IPv6 Internet is built and how to be connected to it. More specifically, this chapter describes the architecture, design, addressing, and routing policy of the 6Bone and how to become a pseudo-TLA on that IPv6 backbone. It also covers policy allocation and how addresses are allocated on the production IPv6 Internet by regional Internet registries (RIRs). It lists the criteria to become an IPv6 provider and describes address allocation, the reassignment of addresses to customers, and how providers may deploy IPv6 connectivity to their customers.

- **Part V: Appendixes**

  Appendix A, "Cisco IOS Software IPv6 Commands"—This appendix lists commands of the Cisco IOS Software technology that are available for IPv6 and that are presented in this book.

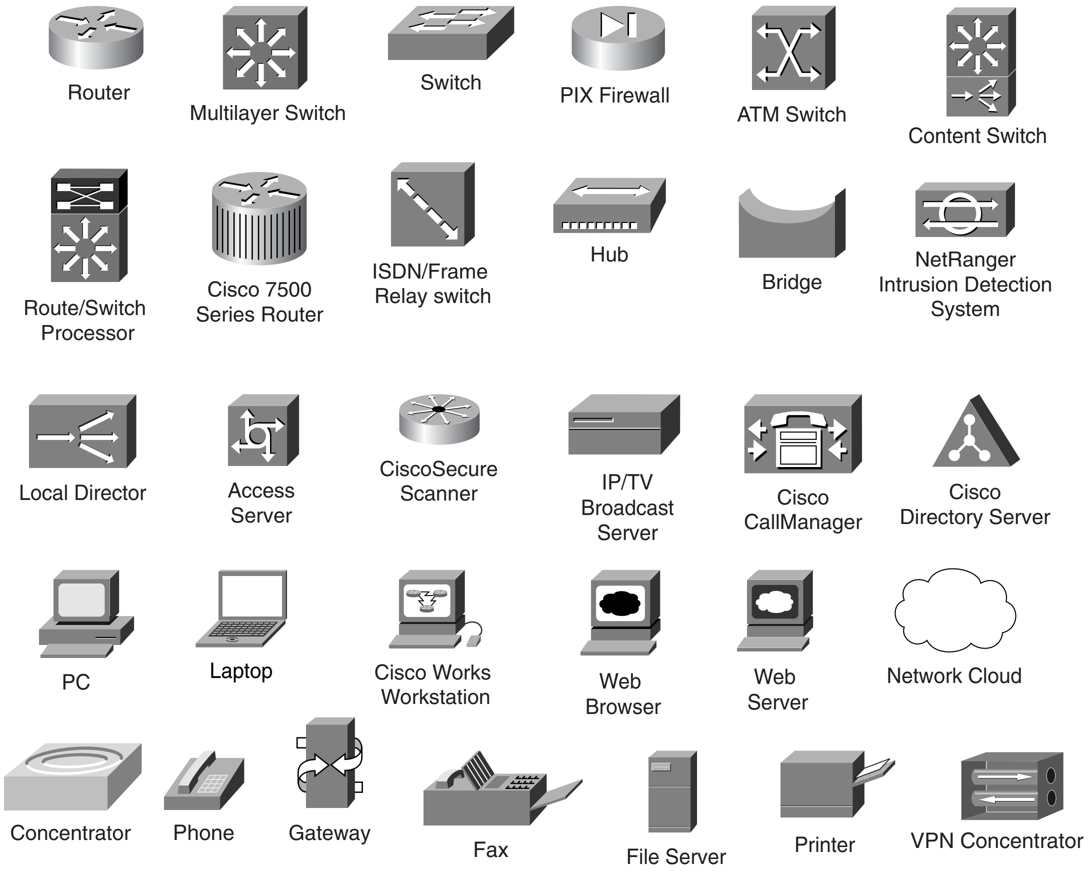
  Appendix B, "Answers to Review Questions"—This appendix provides the answers to each chapter's review questions. The answers to the case study questions can be found at the end of each chapter.

  Appendix C, "RFCs Related to IPv6"—This appendix lists IETF RFCs that explore the technical specifications of IPv6.

  Glossary—This element provides definitions of new technical terms introduced by IPv6.

Icons Used in This Book

Cisco uses the following standard icons to represent different networking devices.
You will encounter several of these icons within this book.



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the *IOS Command Reference*. It describes these conventions as follows:

- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets ([ ]) indicate an optional element.
- Braces ({ }) indicate a required choice.

- Braces within brackets ([{ }]) indicate a required choice within an optional element.
- **Bold** indicates commands and keywords that are entered literally as shown. In configuration examples and output (not general command syntax), bold indicates commands that are manually input by the user (such as a **show** command).
- *Italic* indicates arguments for which you supply actual values.

PART I

Overview of and Justification for IPv6

The Internet has been growing at a very fast rate during the last ten years. The Internet runs over IP version 4 (IPv4), but this protocol was designed 20 years ago for a few hundred computers. This part of the book explains why an upgrade of the IPv4 protocol is needed. It also presents the justifications for and the main benefits of the IP version 6 (IPv6) protocol.

The following chapter comprises this part of the book:

"Everything that can be invented has been invented."

Charles Duell, Commissioner, U.S. Patent Office, 1899

CHAPTER 1

Introduction to IPv6

Before diving into a new technology, it is critical that you understand what problems that technology is designed to solve and what new advantages it provides. When you finish this chapter, you should be able to explain the rationale for using Internet Protocol version 6 (IPv6). This chapter also presents the main features and benefits of the IPv6 protocol.

Rationale for IPv6

You should understand that IPv6 was designed and engineered for many reasons. First and foremost, the Internet Protocol version 4 (IPv4) address scheme is limited by its 32 bits, which causes problems for the long-term growth of the Internet. Moreover, parts of the IPv4 address scheme, such as Class D and E, are reserved for special uses. This also decreases the number of globally unique unicast IPv4 addresses available. Then, very large blocks of globally unique unicast addresses were assigned to organizations in the 1980s, even though the Internet has been growing quickly, especially in Asia and Europe. But, some countries in Asia and Africa received just one Class C address for the entire country because they arrived late to the Internet.

The number of globally unique unicast IPv4 addresses still available is not enough to assign a different IP address to every new device to come. IP is considered by the market as the common denominator to converge different application layers such as data, voice, and audio. However, these new devices require many more IP addresses to interconnect all kinds of IP appliances besides just the computers currently interconnected on the Internet.

The global Internet routing table is huge and continues to grow despite mechanisms such as *classless interdomain routing* (CIDR) and *Network Address Translation* (NAT). Therefore, some studies predict the exhaustion of the current IPv4 address space between 2005 and 2011.

This exhaustion prediction prompted the Internet Engineering Task Force (IETF) to come to the general consensus that there was enough time to engineer a new IP protocol to replace IPv4 before the depletion of the address space. The history behind the development of IPv6 shows that this process has been structured and coordinated between different contributors to solve the problems of the IPv4 protocol.

The version of NAT that was developed during the early days of the web and the commercial Internet to solve critical issues was also seen by the community as a potential solution to the exhaustion of the IPv4 address space. However, a good understanding of the address translation mechanism shows how the NAT mechanism breaks the end-to-end model of the Internet, which causes more limitations to IPv4 than benefits.

IPv4 Address Space

IPv4 is based on a 32-bit address scheme that could in theory enable a total of 4 billion hosts (exactly 4,294,967,296) on the whole Internet. However, this 32-bit scheme was originally divided into five hierarchical classes managed by the *Internet Assigned Numbers Authority* (IANA). The first three classes (A, B, and C) are available as globally unique unicast IP addresses. These classes were assigned to the requesters with a fixed prefix length using different netmask values. A *netmask* is consecutive series of bits preset to 1 designed to "mask" the network part of an IP address.

Table 1-1 shows the five classes of IPv4 addresses, along with their associated ranges and network masks.

Table 1-1 *Hierarchical Classes of IPv4 Addresses*

Classes	Range	Netmask
A	0.0.0.0 to 127.255.255.255	255.0.0.0
B	128.0.0.0 to 191.255.255.255	255.255.0.0
C	192.0.0.0 to 223.255.255.255	255.255.255.0
D	224.0.0.0 to 239.255.255.255	—
E	240.0.0.0 to 255.255.255.255	—

The IANA is an organization dedicated to the central coordination of the Internet. The IANA is responsible for assigning numbers to protocols and blocks of IP addresses to regional Internet registries and large providers. You can find more information about IANA at www.iana.org.

In North America, where early adopters of the Internet were significant in the 1980s, almost all universities and large corporations received Class A or B addresses, even if they had a small number of computers. Today, these same organizations still have unused IPv4 addresses in their assigned blocks of IPv4 addresses, but they have not redistributed them to other organizations. Moreover, many organizations and companies that received an IPv4 address in the 1980s don't exist as such anymore. For example, Digital got bought by Compaq, which might get bought by Hewlett-Packard. Digital and Hewlett-Packard each have a Class A block of addresses.

The redistribution of unused address space is a very important issue of the Internet. In theory, it should be possible to have a global Internet routing table with 4.2 billion entries, but in real life, this represents issues of scalability, performance, and management for large network operators. How is it possible to converge a 4.2 billion-entry database in just a few milliseconds? Just the addition of hundreds of thousands of Class C addresses originating from Class B addresses into the global Internet routing table means doubling the current size of the routing table.

The number of unused IPv4 addresses within these assigned blocks of IPv4 addresses is very large.

Moreover, other large parts of the addressing scheme are not used to assign unique addresses to devices, which decreases the percentage of IPv4 addresses really available as globally unique unicast IP addresses. For example, Class D and E addresses are reserved for multicast and experimental purposes. Networks 0.0.0.0/8, 127.0.0.0/8, and 255.0.0.0/8 are reserved for protocol operations, and 10.0.0.0/8, 169.254.0.0/16, 172.16.0.0/12, 192.168.0.0/16, and 192.0.2.0/24 are special allocations for private networks (defined by RFC 1918). In fact, the sum of all already-assigned Class A and B addresses, unused address spaces, and reserved IP addresses has forced regional Internet registries and ISPs to put a hold on address assignments and distribution. Only small blocks of IPv4 addresses are assigned to organizations, which often means fewer addresses than hosts.

NOTE Three regional Internet registries in the world are responsible for assigning blocks of IP addresses to providers and organizations. ARIN (American Registry for Internet Numbers) serves North America, Central America, and South America; RIPE NCC (Réseaux IP Européens Network Coordination Center) covers Europe and Africa; and APNIC (Asia Pacific Network Information Center) covers IP address assignment in Asia. All three registries have guidelines for the request of IP address spaces. You can find additional information about these registries at www.arin.net, www.ripe.net, and www.apnic.net.

The 32-bit address space of IPv4, like any other addressing scheme such the telephone numbering system, is not optimal. Christian Huitema proposes a logarithmic ratio that is applied to other address spaces, such as the one used in telephone numbers, to compare the efficiency of use.

Each addressing plan has several levels of hierarchy where some margin is provided. However, over time, the hierarchy might change due to growth and the need for mobility. Then, when an allocation is exceeded, a renumbering is needed that involves a very painful and costly operation. Renumbering of telephone area codes in North America has been done since the 1990s mainly due to the growth of new phone service.

At each level of a hierarchy, there is a loss of efficiency. When several hierarchies are present in an addressing plan, the loss of efficiency is much greater. This has a multiplicative effect on the overall efficiency.

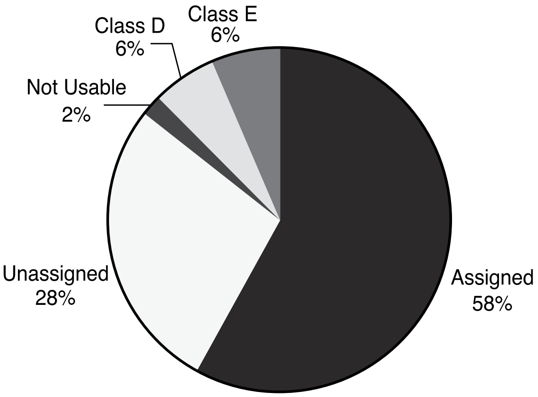
IPv4 is not worse or better than other addressing schemes, but with the class hierarchy (A, B, C, D, and E), in which the most-significant bits of address hierarchy levels are assigned to providers and low-order bits are used for sites and subnets, the address space is less efficient. RFC 3194, *The Host-Density Ratio for Address Assignment Efficiency: An update on the H ratio,* presents detailed information about the HD ratio and the IPv4 address scheme.

The HD ratio is a percentage used to identify the pain level caused by a specific efficiency. A ratio lower than 80% is manageable, but a ratio higher than 87% is hard to sustain. RFC 3194 states that IPv4's 32-bit address space will reach the maximum pain level when 240 million globally unique unicast IP addresses are used on the Internet.

Current IANA IP Address Space Allocation

Figure 1-1 shows the IANA allocation of IP address space in September 2002. Classes D and E, which are unavailable as globally unique unicast addresses, represent a total of 12% of the whole IPv4 address space. The 2% of unusable addresses includes 0.0.0.0/8, 127.0.0.0/8, 255.0.0.0/8, and private address spaces. The biggest slice of the graph (58%) represents the address space already assigned to organizations and regional Internet registries such ARIN, APNIC, and RIPE, meaning that 28% of the remaining IPv4 space is still unallocated.

Figure 1-1 *Percentage of the IPv4 Address Space Assigned in September 2002*



Source: Computed from information published by IANA about IPv4 address space allocation

Future Growth of the Internet

The current situation shows that it will be much harder to get addresses in the future as they become a scarce resource and the Internet continues to grow globally. The problem is obvious in some places, but not in North America, where 75% of the IPv4 address space is allocated for less than 10% of the world population.

Moreover, temporary and semipermanent connections such as dialup are being replaced by connections such as cable-modem/xDSL, which require one permanent IP address per node instead of one temporary IP address for a pool of PPP subscribers. The ratio of subscribers per IP address changes from many:1 to 1:1. Wireless networks are emerging markets, and 802.11b devices and mobile networks are deployed everywhere. However, wireless devices frequently change physical locations, access points, and logical subnets during movement, which means that extra pools of IP addresses are requested for these devices.

Some ISPs are running out of IP addresses; therefore, they must assign private addresses to their customers through NAT. New large networks cannot get IPv4 addresses from regional Internet registries or ISPs. New technologies such PDAs, wireless devices, cellular phones, VoIP, and videoconferencing over IP applications, require globally unique unicast IP addresses. Moreover, the current generation of PCs and operating systems allows people to have their own web servers for their personal data. This also requires permanent IP addresses to be assigned on home networks.

IPv4 Address Space Exhaustion

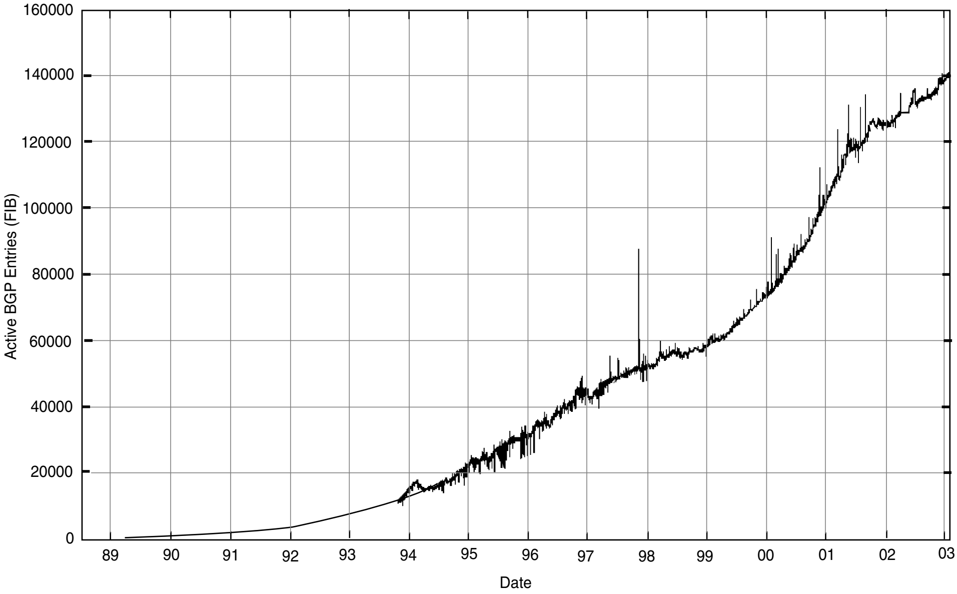
The work on IPv6 at the IETF started when a preliminary study in 1990 concluded that the IPv4 address space would be exhausted. More specifically, the IETF predicted that Class B would be exhausted within four years (1994). This study also identified the necessity to assign several adjacent Class C addresses instead of Class B addresses to organizations. Class C addresses are small, but there are plenty of them (2,097,152).

NOTE Class C is a block that represents 255 IPv4 addresses, whereas one Class B means 65,536 IPv4 addresses. However, in reality, 254 hosts can be addressed on a Class C.

The main technical constraint of that orientation was preserving the global Internet routing table size while keeping it from exploding. With several thousand routes in the global Internet routing table, the addition of hundreds of thousands of new small routes (Class C) was an important issue to avoid. Therefore, the CIDR mechanism adopted in 1992 was put into place to summarize adjacent blocks of IPv4 addresses in one block. CIDR has helped control the growth of the Internet routing table since 1993.

Figure 1-2 shows the global Internet routing table growth since 1989 (active BGP entries). In 2001, the total number of routes was more than 100,000, then later in 2003 it was 140,000 entries (a 40% growth within 24 months) If you want real-time information about this routing table, look for a route server on the Internet. Some of these route servers are freely available for public information and debugging purposes.

Figure 1-2 *Global Internet Routing Table Growth Since 1989*



Source: BGP Table Statistics, Telstra web site, www.telstra.net/ops/bgptable.html

Example 1-1 shows the global Internet routing table on a route server.

Example 1-1 *Looking at the Global Internet Routing Table on a Route Server*

```
#telnet route-server.ip.att.net

route-server>show ip route
show all routes of the routing table of the Internet



route-server>sh ip bgp summary

BGP router identifier 12.0.1.28, local AS number 65000
BGP table version is 665451, main routing table version 665451
117228 network entries and 2373589 paths using 116277944 bytes of memory
37354 BGP path attribute entries using 2091992 bytes of memory
24197 BGP AS-PATH entries using 630776 bytes of memory
402 BGP community entries using 15192 bytes of memory
```

Example 1-1 *Looking at the Global Internet Routing Table on a Route Server (Continued)*

```
24674 BGP route-map cache entries using 493480 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
Dampening enabled. 945 history paths, 751 dampened paths
BGP activity 125101/1203692325 prefixes, 2562479/188890 paths, scan interval 60 secs
```

In Example 1-1, the highlighted line shows 117,228 network entries, which is the total number of routes, 2,373,589 paths, which is the number of BGP AS-PATH entries, and 116,277,944 bytes of memory, which is memory used on that router to handle the routing table.

Another study done by the IETF tried to predict how long it will be before the IPv4 address space is exhausted. This study projected the unavailability of new IPv4 address space between 2005 and 2011. The perception of these results was mixed and contested. Some people argued that the projection was pessimistic, and others thought it was optimistic.

History of IPv6

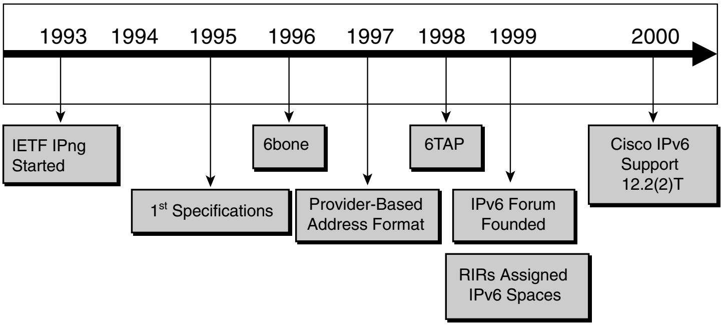
Demonstration of the IP address space exhaustion led to the consensus that there was enough time to design, engineer, and test a new protocol with enhanced functionalities instead of deploying a new protocol that just adds larger addresses. This represented a unique opportunity to fix the limitations related to the IPv4 addressing scheme and to develop a protocol to ensure reliable growth of the Internet over the next decades. The process took care of requirements from various industries, including the cable and wireless industries, electric power utilities, the military, corporate networks, Internet service providers (ISPs), and many others.

In 1993, a call for proposals (RFC 1550) was issued. Three were studied in detail:

- Common Architecture for the Internet (CATNIP) proposed converging the CLNP, IP, and IPX protocols with the use of Network Service Access Point (NSAP) addresses. (Defined in RFC 1707.)
- Simple Internet Protocol Plus (SIPP) proposed increasing the IP address size to 64 bits and improving the IP header. (Defined in RFC 1752.)
- TCP/UDP Over CLNP-Addressed Networks (TUBA) suggested replacing IP (Layer 3) with Connectionless Network Protocol (CLNP), where TCP/UDP and other upper protocols could run on top of CLNP. (Defined in RFC 1347.)

The recommended proposal was SIPP, with an address size of 128 bits. The main author of SIPP was Steve Deering. IANA assigned the version number 6 to the protocol. A working group at IETF called *IP Next Generation (IPng)* was started in 1993 just before the web really led to the explosion of Internet traffic. However, the issue of IPv4 existed before the web. Then, the first specifications came in late 1995 (RFC 1883). The IPng working group was renamed *IPv6* in 2001. Figure 1-3 shows the origin and evolution of IPv6.

Figure 1-3 *History of IPv6*



In 1996, an IPv6 test bed called the IPv6 backbone (6bone) was created over the Internet. The 6bone has used mainly a mix of Cisco IOS Software routers with IPv6 beta implementations and other router software under UNIX platforms. IPv6 prefixes within the IPv6 space 3ffe::/16 were assigned to the 6bone participants. In 1997, a first attempt was made to structure the IPv6 space as a provider-based IPv6 address format. One year later, the first IPv6 exchange, called 6TAP, was deployed at STARTAP in Chicago. In 1999, regional Internet registries (RIRs) started assigning production IPv6 prefixes using the IPv6 space 2001::/16. In the same year, the IPv6Forum, a worldwide consortium of leading Internet vendors and research and education networks, was founded to promote IPv6 in the market and to enable collaboration between vendors. In 2000, many vendors began bundling IPv6 into their mainstream products. Cisco introduced a three-phase road map for the development of IPv6 and made IPv6 support available in Cisco IOS Software Release 12.2(2)T. In 2001, Microsoft announced the availability of IPv6 in the mainstream code of its latest operating system, Windows XP.

NOTE Chapter 7, "Connecting to the IPv6 Internet," provides detailed information about the 6bone and the IPv6 spaces assigned by IANA. Chapter 6, "IPv6 Hosts Internetworking with Cisco," describes the IPv6 support on Microsoft Windows XP.

IPv5

The Internet community uses IPv4 and has used IPv6 for a couple of years. IANA is the organization that has the worldwide responsibility of assigning numbers to everything related to the Internet, which includes versions of the IP protocol. IANA assigned version 6 to the IPng protocol in 1995 following a request by the IPng working group.

What about "IP version 5"? IPv5 is an experimental resource reservation protocol intended to provide quality of service (QoS), defined as the Internet Stream Protocol (ST). It can provide real-time transport of multimedia such as voice, video, and real-time data traffic across the Internet. This protocol is based on previous work of Jim Forgie in 1979, as documented in IETF Internet Experiment Note 199. It consists of two protocols—ST for the data transport and Stream Control Message Protocol (SCMP). IPv5, also called ST2, is documented in RFC 1819 and RFC 1190.

Internet Streaming Protocol version 2 (ST2) is not a replacement for IPv4. It is designed to run and coexist with IPv4. The number 5 was assigned by IANA because this protocol works at the same link-layer framing as IPv4. A typical distributed multimedia application can use both protocols: IP for the transfer of traditional data and control information such as TCP/UDP packets, and ST2 for real-time data carriers. ST2 uses the same addressing schemes as IPv4 to identify hosts. Resource reservation over IP is now done using other protocols such as *Resource Reservation Protocol* (RSVP).

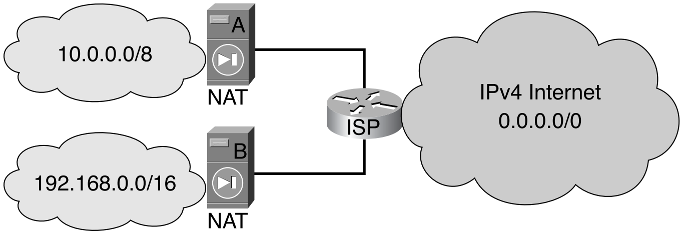
Network Address Translation

Since 1992, CIDR has not been the only mechanism directly involved in slowing down the IPv4 address shortage. Over the years, the NAT mechanism (defined in RFC 1631), seen as a short-term solution, played a key role by allowing organizations to use few Internet globally unique unicast IP addresses for their large networks. NAT typically translates packets from a network, which uses globally unique unicast IP addresses or a private address space as defined by RFC 1918, to the Internet.

NOTE IANA has reserved three blocks of IP addresses for private addressing. Address spaces 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16 are used for address translation with the Internet.

Figure 1-4 shows networks using private addressing. 10.0.0.0/8 and 192.168.0.0/16 are connected to the Internet through the same ISP using NAT. Because private addresses are not routed across the Internet, nodes on these private networks cannot be reached from the Internet.

Figure 1-4 *Networks Connected to the Internet Using NAT with Private Addressing*



Since 1990, the combination of CIDR, NAT, and private addressing has provided benefits to the Internet by slowing the depletion of IPv4 addresses.

Moreover, one of the arguments against deploying IPv6 is the use of NAT. This is seen by some as the *permanent* solution to the shortage of IPv4 address space. However, using NAT has many implications; these were taken into consideration during the engineering of IPv6. Some of these limitations are documented in RFC 2775 and RFC 2993:

- **NAT breaks IP's end-to-end model**—IP was originally designed so that only endpoints (hosts and servers) have to handle the connection. The network itself, the underlying layers, and NAT do not have to handle connections.
- **The need to keep the state of connections**—NAT implies that the network (NAT translator) needs to keep the state of the connections and that NAT has to remember the translation of addresses and ports.
    - — The need to keep the state of the connections in NAT makes fast rerouting difficult in case of a failure of the NAT device or the links near the NAT device. Networks using links and route redundancy can suffer problems.
    - — Organizations deploy high-speed links (Gigabit Ethernet, 10 Gigabit Ethernet) to increase the performance of their backbones. However, address translation requires additional processing, because the state of each connection must be kept with NAT. Therefore, NAT hinders network performance.
    - — For providers and organizations that must keep records of all connections made by their end users for security reasons, the recording of NAT state tables becomes mandatory to trace back to the source of problems.
- **Inhibition of end-to-end network security**—To protect the integrity of the IP header by some cryptographic functions, this header cannot be changed between the origin of the packet, which protects the header's integrity, and the final destination, which checks the integrity of the received packet.

  Any translation on the path of header parts breaks the integrity check. Although many adaptations can partly solve this issue in some cases, the fundamental problem is not easy to solve. The IPSec authentication header (AH) is an example of this problem.

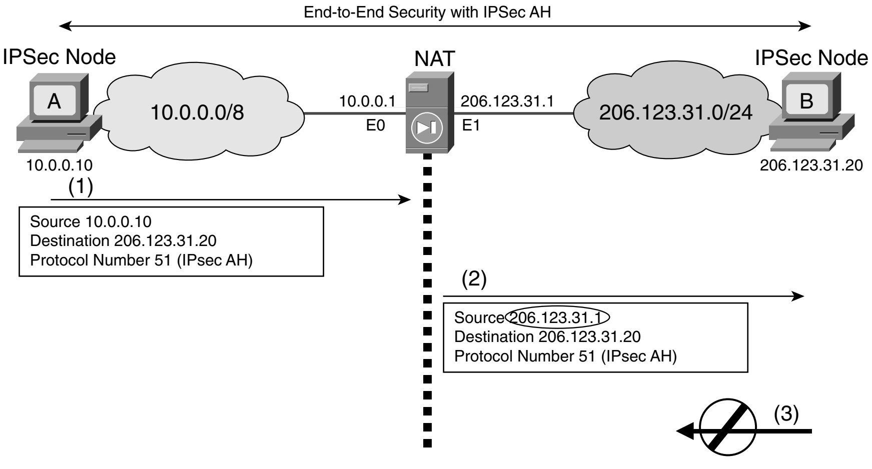
  In Figure 1-5, Computer A (1), which has an IPSec implementation, sends IP packets with protocol number 51 (IPSec AH) to Computer B. NAT, before forwarding the packet (2) to network 206.123.31.0/24, changes the IP source address within the header from 10.0.0.10 to 206.123.31.1. However, the IPSec implementation in Computer B fails the integrity check because something was modified within the packet header during transport.
- **Applications that are not NAT-friendly**—More than just port and address mapping is necessary to forward the packet through the NAT device. NAT has to embed the full knowledge of all applications to do the right tricks. This is especially important in cases with dynamically allocated ports with rendezvous ports, embedded IP addresses in

application protocols, security associations, and so on. The outcome is that the NAT device needs to be upgraded each time a new non-NAT-friendly application is deployed.

- **Address space collision**—When different networks and organizations use the same private address space and have to merge or connect, an address space collision results: Different hosts/servers can have the same address, and routing disables reaching the other network. However, this can be resolved by a few techniques such as renumbering or twice-NAT. But these techniques are very painful and costly and later increase NAT's complications.
- **Ratio of internal and reachable IP addresses**—NAT can be efficient when there is a large number of hosts/servers inside and very few reachable addresses outside. The ratio of internal/reachable addresses must be large to make NAT effective.

  However, many servers behind NAT that must be reached from the Internet is a problem. The same protocol cannot be multiplexed on the same port using the NAT external address, such as in Network Address Port Translation (NAPT) mode. NAPT allows the sharing of one IP address using TCP and UDP ports as tokens for the translation mechanism. For example, two web servers located behind NAT that both use port TCP 80 cannot use the same external IP address without changing the port number. Because many protocols make nodes as servers, it consumes many external addresses. Consequently, NAT is not as useful.

Figure 1-5 *Translation Breaks the Integrity Check of IPSec AH in the End-to-End Model*


End-to-End Security with IPSec AH
IPSec Node
NAT
IPSec Node
A
10.0.0.0/8
10.0.0.1
E0
206.123.31.1
E1
206.123.31.0/24
B
10.0.0.10
206.123.31.20
(1)
Source 10.0.0.10
Destination 206.123.31.20
Protocol Number 51 (IPsec AH)
(2)
Source 206.123.31.1
Destination 206.123.31.20
Protocol Number 51 (IPsec AH)
(3)

IP's original design was based on an end-to-end model. This model led to the design of thousands of Internet standards with predictable behavior for the benefit of the Internet. However, NAT, introduced as a temporary solution, breaks this end-to-end model. NAT was a patch applied to extend IPv4's lifetime for a short time. IPv6 is the long-term solution to retain the end-to-end model and the IP protocol's transparency.

IPv6 Features

After the overview of the main problems related to the IPv4 protocol, you should see that IPv6 solves all these problems and provides new benefits. Here are the main improvements:

- The 128-bit address scheme, which provides plenty of IP addresses for the next decades.
- The larger address space provides globally unique addresses to billions of new devices such as PDAs, cellular devices, and 802.11 systems, that will be manufactured in the future.
- Multiple hierarchy levels help aggregate routes, which promotes efficient and scalable routing to the Internet.
- Multihoming with the preservation of strict route aggregation is possible.
- The autoconfiguration process allows nodes of the IPv6 network to configure their own IPv6 addresses.
- The transition between IPv6 providers is transparent to end users with the renumbering mechanism.
- ARP broadcast is replaced by multicast use on the local link.
- The IPv6 header is more efficient than IPv4. Fewer fields are present, and the header checksum is removed.
- A flow label field can provide traffic differentiation.
- New extension headers replace IPv4's Options field and provide more flexibility.
- IPv6 was designed to handle mobility and security mechanisms much more efficiently than the IPv4 protocol.
- Many transition mechanisms are designed with IPv6 to allow a smooth transition from IPv4 networks to IPv6.

The following sections examine some of these IPv6 features and discuss how they offer improvements over the IP protocol.

Larger Address Space

IPv6 increases by a factor of 4 the number of address bits, from 32 to 128 bits. During the IPv6 design specification, there was a debate about using fixed-length 64-bit addresses versus variable-length addresses up to 160-bit. Table 1-2 compares the arguments for each.

Table 1-2 *IPv6: 64-Bit Versus 160-Bit Proposals*

64-Bit Proposal	160-Bit Proposal
Enough addressing for 10 trillion sites and 10^{15} nodes	Addresses compatible with NSAP addressing
Minimizes the increase of the header size compared to IPv4	Autoconfiguration possible using IEEE 802.x link-layer addresses
—	Variable length of the addresses allows for the use of 64-bit addresses instead of fixed-length addresses. Over time, addresses can be longer.

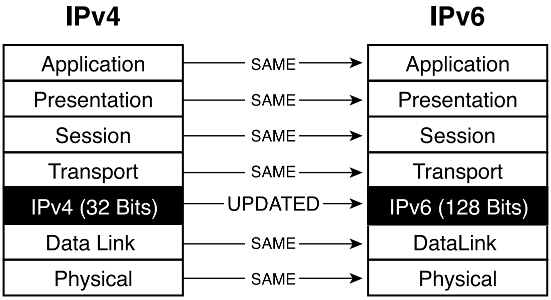
Finally, using fixed-length addresses of 128 bits for IPv6 was found to be the most appropriate choice.

With IPv4, the number of addressable nodes is 4,294,967,296 (2^{32}), which represents about two IPv4 addresses for every three people (based on a world population of 6 billion people in 2001).

By comparison, the 128-bit length of IPv6 represents $3.4 * 10^{38}$ addresses, which allows approximately $5.7 * 10^{28}$ IPv6 addresses for every person in the world. However, as in any addressing scheme, such as IPv4 and telephone systems, not all the addresses can be used, but enough are available for any kind of use. Increasing the number of bits for the address also means an increase in the IP header size. Because each IP header contains a source address and a destination address, the size of the header fields containing the addresses is 64 bits for IPv4 and 256 bits for IPv6.

Comparing the OSI reference model of IPv4 to that of IPv6 (see Figure 1-6), IPv6 represents only a change at Layer 3 (the network layer). Other layers are slightly modified. This was an important consideration during the engineering of IPv6. The other layers of the two OSI reference models are the same, which means that protocols such as TCP and UDP used with IPv4 continue to run on top of IPv6.

Figure 1-6 *Scope of IPv6 with the OSI Reference Model*



Global Reachability

The important issue that initiated the IPv6 effort was the address space exhaustion study to give one globally unique unicast address to each device connected to the Internet. By using a much larger address space than IPv4 (4,294,967,296 addresses), IPv6 enables the use of a global and reachable address for almost every kind of device: computers, IP phones, IP faxes, TV setup boxes, cameras, pagers, wireless PDAs, 802.11b devices, cell phones, home networking, and vehicles. From now until 2006, cellular manufacturers plan to produce billions of new wireless devices that include an IP stack. These next-generation wireless devices will provide subscribers with Internet interactivity and services with their phones.

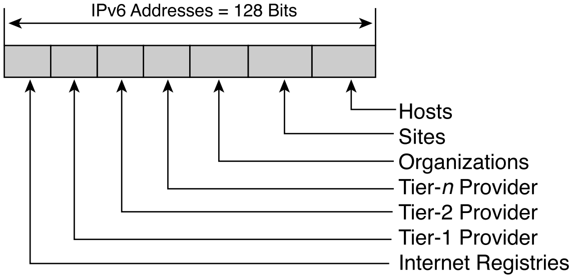
Trying to fit all these devices into the current IPv4 address space is almost impossible. Having a unique IP address for each device enables end-to-end reachability, which was lost over past years with NAT devices and private addressing. The end-to-end model is especially important for telephone call and end-to-end security. IPv6 enables the full support of application protocols without needing special processing by the network itself.

NOTE In IPv6, NAT is undesirable between IPv6-only networks. Plenty of IPv6 addresses are available precisely to preserve the end-to-end model of the IP protocol.

Levels of Addressing Hierarchy

A much larger address space enables the use of multiple levels of hierarchy inside the address space, as shown in Figure 1-7. Each level helps aggregate its IP space and enhance the allocation function. Providers and organizations may have tiered hierarchy and manage the assignment of the space below.

Figure 1-7 *128-bit Address Space Enables Multiple Levels of Hierarchy*



Using multiple levels in the hierarchy provides flexibility and new functionalities to the protocol.

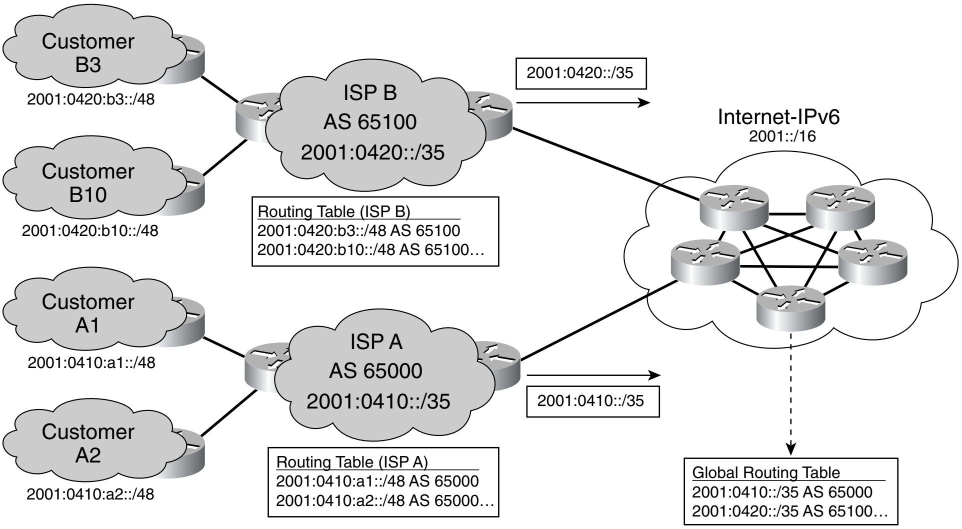
A flexible addressing architecture is key to a network protocol. In the IPv4 world, the small 32-bit address space is an important limitation that has not led to the use of several hierarchy levels, so this affects route summarization (aggregation).

Aggregation

A larger IPv6 address space makes room for large address allocations to ISPs and organizations. Having a large-enough prefix for an organization's entire network enables the use of only one prefix. Moreover, the ISP can summarize routes (aggregation) of all its customers' prefixes into a single prefix and announce it to the IPv6 Internet.

In Figure 1-8, ISP B advertises to the IPv6 Internet that it can route the network 2001:0420::/35, which includes IPv6 spaces assigned to customer B3 (network 2001:0420:b3::/48) and customer B10 (network 2001:0420:b10::/48). ISP A advertises to the IPv6 Internet that it can route 2001:0410::/35, including networks 2001:0410:a1::/48 and 2001:0410:a2::/48.

Figure 1-8 *Providers Aggregate Customer Prefixes and Then Advertise Their Prefixes to the IPv6 Internet*



NOTE When a customer changes its IPv6 provider, it must change its IPv6 prefix to preserve this global aggregation. Changing providers implies network renumbering. However, autoconfiguration (discussed in a moment and presented later in detail in Chapter 3, "IPv6 in Depth") eases the renumbering of hosts within an organization.

This aggregation of routes promotes an efficient and scalable routing. To connect all kinds of devices and networks on the Internet in the future, which can represent several billions of nodes, scalable routing is a requirement. However, there should be many fewer routes in the global IPv6 Internet routing table than in the current IPv4 Internet. Route aggregation in IPv6 is possible because multihomed sites can configure addresses from several upstream providers.

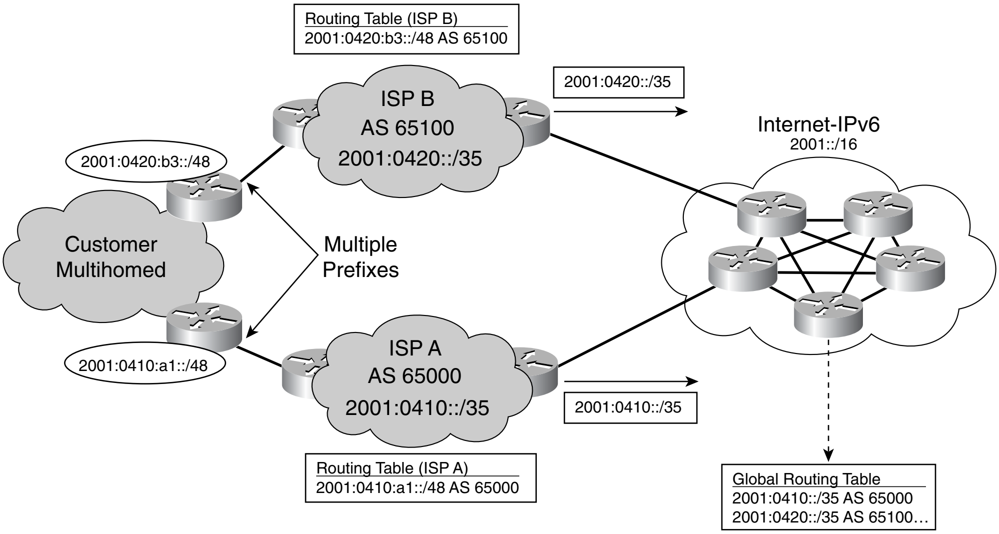
Multiple Addresses

In the IPv4 world, it is not simple to connect a network to multiple providers. One way an organization can do multihoming is to get provider-independent IPv4 space from regional Internet registries. Then, an organization can conclude peering agreements with multiple providers to announce its prefix to the Internet. In the context of provider-aggregatable IPv4 space, the prefix used is part of a provider's address space. Then multihoming is possible if other ISPs used advertise the same prefix to the Internet. At the very least, it breaks any kind of aggregation in the global Internet routing table. However, multihoming is desirable for high network reliability.

Having a much larger address space with IPv6 enables the use of multiple simultaneous prefixes for an organization. An organization connected to several ISPs gets multiple prefixes that are part of these ISPs' IPv6 address spaces. This allows multihoming without breaking the global routing table, which currently is not possible in IPv4.

In Figure 1-9, the multihomed customer is connected to both ISP A and ISP B, which have assigned networks 2001:0420:b3::/48 and 2001:0410:a1::/48a to it. ISP A and ISP B advertise their /35 prefixes to the IPv6 Internet.

Figure 1-9 *IPv6 Enables the Use of Multiple Prefixes for Multihoming*



Multihoming is obviously possible with IPv4, but it has consequences for the size of the global Internet routing table, because the same network prefix may be advertised by different autonomous systems (ASs). One goal of IPv6 is to preserve the global routing table as small as possible.

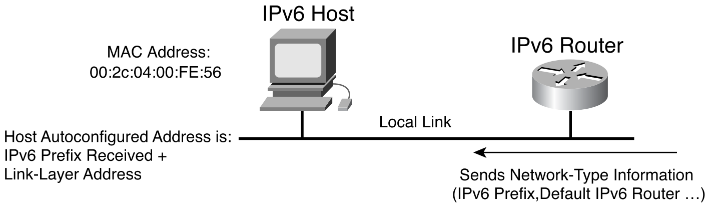
The concept of multiple addresses implies that each network interface of a node might have multiple globally unique unicast IP addresses at the same time. Having multiple addresses on a network and nodes requires source address selection to choice addresses used to initiate connections. Source selection is a mechanism by which each node can select or be forced to prefer an IPv6 prefix when many are available. Moreover, if a link goes down, all routers within the multihomed network should be able to replace the current IPv6 prefix advertised by another one. The source address selection and router renumbering mechanisms are currently being discussed at the IETF. However, a mechanism called autoconfiguration is already available to allow the renumbering of all nodes on IPv6 networks.

Chapter 2, "IPv6 Addressing," presents IPv6's addressing architecture in detail.

Autoconfiguration

Autoconfiguration is a new function enabled by IPv6. By having a much larger address space, IPv6 is designed to enable autoconfiguration of the addresses on a device while keeping the global uniqueness. As Figure 1-10 illustrates, an IPv6 router on the same local link sends network-type information such as the IPv6 prefix of the local link and the default IPv6 route. All IPv6 hosts on the local link listen to this information and then can configure by themselves their IPv6 addresses and the default router. Autoconfiguration is a mechanism by which each IPv6 host and server appends its link-layer address (for example, an Ethernet MAC address) in the EUI-64 format to the globally unique unicast IPv6 prefix advertised on the subnet.

Figure 1-10 *IPv6 Host Autoconfiguring Its IPv6 Address*



NOTE Autoconfiguration (defined in RFC 2462) is also called *IPv6 stateless address autoconfiguration*.

An interface's link-layer address is based on the MAC (Media Access Control) address of the network interface converted to the EUI-64 (Extended Unique Identifier 64) format, which has a 64-bit length. The transformation of the 48-bit MAC address into EUI-64 is covered in detail in Chapter 2. An interface's link address is the lowest 64-bit part of the IPv6 address, and the IPv6 prefix is the high-order 64-bit part of the 128-bit address.

NOTE An IPv6 prefix assigned to a local link has a 64-bit length (/64). The low-order 64-bit part is the interface's link-layer address. Using this concept, IPv6 simplifies subnet addressing within networks by having the same prefix length instead of different netmask values, as in IPv4.

The 128-bit address provided by autoconfiguration is guaranteed to be globally unique, because the 48-bit MAC address is a combination of a 24-bit Organizational Unique Identifier (OUI) assigned to a vendor by the IEEE with a unique 24-bit value generated for each interface built. Because it is possible under special circumstances to modify a network interface's 48-bit MAC address using software that could cause address collision, each IPv6 stack has a process enabled to detect duplicate addresses on the local link. Duplicate address detection (DAD) mechanism is explained in detail in Chapter 3.

NOTE Autoconfiguration is not the only way to assign an IPv6 address to a node's interface. Manually configuring a network interface can still be done in IPv6 and is mandatory for routers. IPv6 hosts can also obtain interface addresses and parameters from a DHCPv6 server. This mode (DHCPv6) is called *IPv6 stateful address configuration* (as opposed to IPv6 stateless address configuration, or autoconfiguration). Finally, another method allows a node to generate a random interface identifier that can be used as the low-order 64-bit part of the address. Random address generation was added to preserve privacy.

Autoconfiguration enables *plug and play,* which connects devices to the network without any configuration or servers such as DHCP servers. This is a key feature to enable deployment of new devices on a very large scale on the Internet such as cell phones, wireless devices, home appliances, and home networks.

Chapter 3 covers the stateless autoconfiguration mechanism in detail.

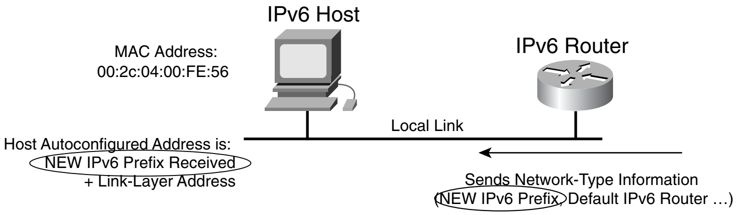
Renumbering

The larger address space provided by IPv6 lets organizations get IPv6 prefixes that provide IPv6 addresses for their production needs. One of the main goals of IPv6 is to keep the smallest global IPv6 routing table possible on the Internet by forcing strict aggregation. However, when an organization changes its IPv6 upstream provider, it must renumber its network.

With IPv4, renumbering is a time-consuming, error-prone task. The organization gets a new IPv4 space first, and then it must change the IPv4 addresses of all its routers, servers, hosts, and other devices on the network. Routing protocols and DNS servers must also be updated with the new IPv4 addresses at the same time. Therefore, renumbering in IPv4 incurs downtime and halts network services.

In IPv6, the renumbering process was designed to be smooth because the transition between unicast IPv6 providers can be completely transparent to end users. The combination of having multiple providers during the transition and the stateless autoconfiguration mechanism enables easy renumbering for hosts by sending the new unicast IPv6 prefix to the network. However, the renumbering of routers represents a burden for network operators, as in IPv4. A lifetime value can be assigned to advertised prefixes, allowing nodes to use the newest prefix after the expiration of the current prefix. Therefore, hosts and servers automatically pick the new global unicast IPv6 prefix and then use the new address. Figure 1-11 shows the IPv6 router on the same local link that sends network-type information such as a new IPv6 prefix and a new default IPv6 route. Hosts on this local link use these new values to autoconfigure their new IPv6 addresses.

Figure 1-11 *A New Unicast IPv6 Prefix Is Advertised on the Local Link During Renumbering*


IPv6 Host
MAC Address:
00:2c:04:00:FE:56
IPv6 Router
Local Link
Host Autoconfigured Address is:
NEW IPv6 Prefix Received
+ Link-Layer Address
Sends Network-Type Information
(NEW IPv6 Prefix, Default IPv6 Router ...)

NOTE In IPv6, a router cannot configure its network interfaces using the autoconfiguration mechanism. IPv6 addresses on the router's interfaces must be configured manually. Moreover, router interfaces are seen by hosts and servers on the local link with another kind of IPv6 address called the *link-local address*. This guarantees that the router can be reached even while a network is renumbering. Obviously, unicast IPv6 addresses assigned to each of a router's network interfaces can change during the renumbering. Link-local address is presented in chapter 2.

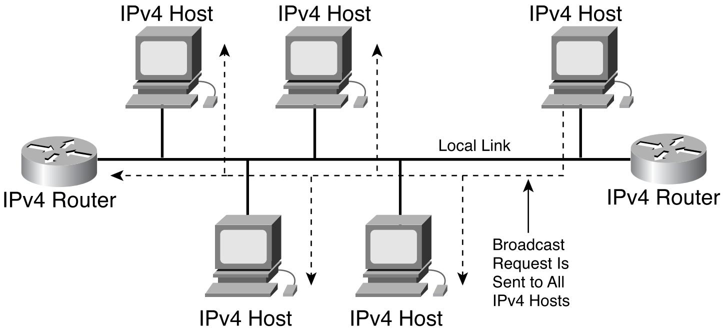
The renumbering process does not prevent hosts and servers from losing their current TCP and UDP sessions at the exact moment the transition occurs, which is only possible with a protocol such as MobileIP.

Chapter 3 presents the mechanisms of IPv6 behind network renumbering.

Multicast Use

The ARP (Address Resolution Protocol) broadcast in IPv4, well-known by the use of the Layer 2 MAC address *ff:ff:ff:ff:ff:ff,* is inefficient for the network. Each time a broadcast request is sent to a local link, it causes at least one interrupt in every computer on the link, even though only one or two nodes are involved. The computer's network interface listens to the broadcast packet. Then it is sent to the operating system, and finally it arrives at the IP stack, where it can be used or simply ignored. In some cases, broadcasts can completely hang up a whole network; this is called a *broadcast storm*. Figure 1-12 shows a broadcast packet in IPv4 that is sent on the local link to every host from one host. This broadcast packet goes up to the IPv4 stack of all nodes on this local link.

Figure 1-12 *ARP Broadcast Request Sent to a Local Link by Any IPv4 Host*

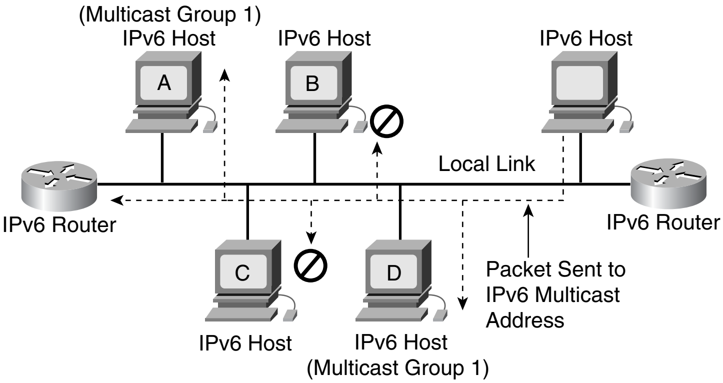


ARP broadcast is not used in IPv6. Multicasting is used instead. As illustrated in Figure 1-13, multicast group 1 defines a group of network interfaces. Network interfaces of computer A and computer D are members of multicast group 1. When a packet is sent to multicast group 1 using this group's multicast address, the packet is processed only by computers A and D, which are members of this group. Every other computer and router on this local link does not process the packet sent to multicast group 1, because they are not members.

Therefore, multicasting enables the efficient use of the network by spreading broad requests to a smaller number of possible computers by using different and specific multicast groups for the different functions. This is less costly in CPU cycles for all computers on a local link and prevents the majority of problems, such as the broadcast storms in IPv4.

Multicasting in IPv6 is used on local links to replace the ARP broadcast traffic, which means that the use of multicast routing is not required on the router infrastructure between local link subnets for that use. However, as in IPv4, this is possible to enable the multicast routing in IPv6 on routers for global use.

Figure 1-13 *Multicast Packet Sent to All Members of a Multicast Group on a Local Link in IPv6*



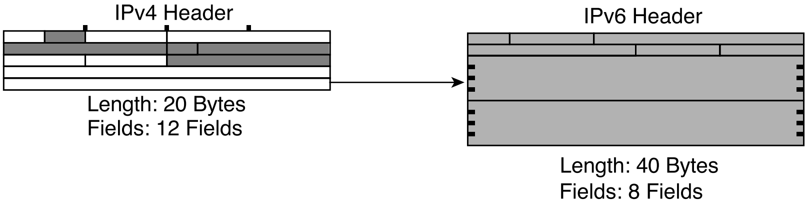
Because the range of multicast addresses in IPv6 is much larger than in IPv4, the allocation of multicast groups should not be limited. For example, a scope has been defined in the whole IPv6 addressing space for any type of multicast use.

Chapter 3 explains in detail the use of multicasting to replace ARP.

Efficient Header

As shown in Figure 1-14, the new IPv6 header is simpler than the IPv4 packet header. Six of the IPv4 header fields are removed in the IPv6 header. The IPv4 header with the Option and Padding fields has 14 fields, and the IPv6 header has eight fields. The basic IPv6 header size is 40 octets, and the IPv4 header without the Option and Padding fields is 20 octets. The basic IPv6 header has a fixed length, and the IPv4 header when used with the Options field may have a variable length.

Figure 1-14 *IPv6 Header Is Simpler and Larger Than the IPv4 Header*



The fewer fields in the IPv6 header and its fixed length mean that it is less costly in CPU cycles for routers to forward IPv6 packets. This has direct benefits for network performance.

All fields in the IPv6 header are aligned to 64-bit, which enables direct storage and access to memory. These enhancements enable hardware-based processing, which provides scalability of the forwarding rate for the next-generation high-speed pipes. However, this remains to be seen because of the following:

- 128-bit addresses are larger than the atomic word size of the current processors, so there is more lookup to do to get the full 128-bit address.
- Performing the longest match prefix to look at 128-bit versus 32-bit before forwarding packets also has a clear impact on performance.
- Packet filtering performed at Layer 4 (TCP/UDP) results in the parsing of optional IPv6 headers (when present), which represents additional CPU cycles for routers.

Moreover, the hardware to process packets is not currently optimized to meet the performance expectations of IPv6. However, in the long term, the 64-bit alignment of IPv6 header fields should improve routing efficiency.

In IPv4, a 16-bit field is used to verify the header's integrity. The packet's sender generates the checksum and then forwards the packet to the network. Because some other fields change within the IP header, such as the TTL (Time-to-Live) value decremented at each hop, a new checksum is generated and then is filled into the IP header each time a router forwards the packet.

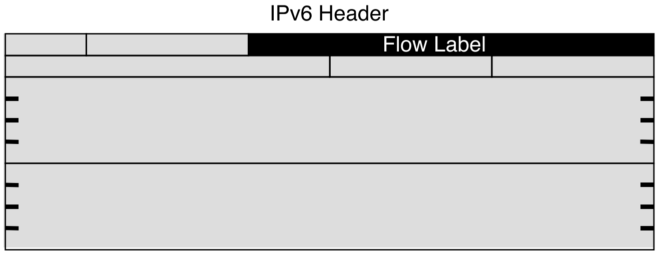
Another improvement to the IPv6 header is related to the *Checksum* field. This header field is simply removed to increase routing efficiency. In fact, all routers in the path do not have to make the checksum recalculation during the forwarding process. Error detection is now handled by data-link layer technologies (Layer 2) and by checksums of the end-to-end connection at the transport layer (Layer 4). The checksum done at Layers 2 and 4 is strong enough to bypass the need for Layer 3 checksum. With IPv6, checksums are required for both transport protocols TCP and UDP. UDP checksum was optional with IPv4.

Fragmentation is handled differently in IPv6. Fragmentation fields of IPv4 are either completely gone or removed and then replaced by extension headers. Chapter 2 presents details about the new way to handle fragmentation and explains the impact on the header.

Flow Label

IPv6 includes a new *Flow Label* field in the IPv6 header, as illustrated in Figure 1-15. A source node can use this special field to request special handling for a specific sequence of packets. The Flow Label field is mainly for end-station processing, not for routers. This can be useful for streaming applications such as videoconferencing and Voice over IP that require real-time data transmission. The flow label enables per-flow processing for applications requiring QoS in routers in the path. This is better than best-effort forwarding.

Figure 1-15 *Flow Label Is a New Field in the IPv6 Header*



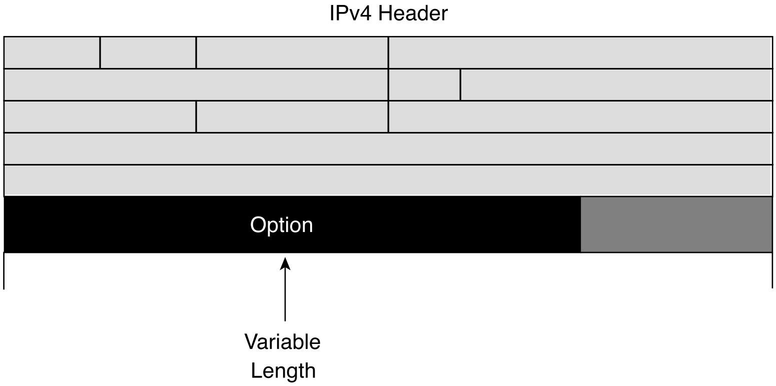
This field differentiates the traffic at the IP layer without doing other tricks to identify the flows. With this label, a router does not have to open the transport inner packet to identify the flow; it finds the information in the IP packet header. The current IETF standard does not specify the details of how to manage and process the label. Interactions with DiffServ, IntServ, RSVP, and MPLS are possible development methods.

Chapter 2 presents the IPv6 header in detail.

Extension Header

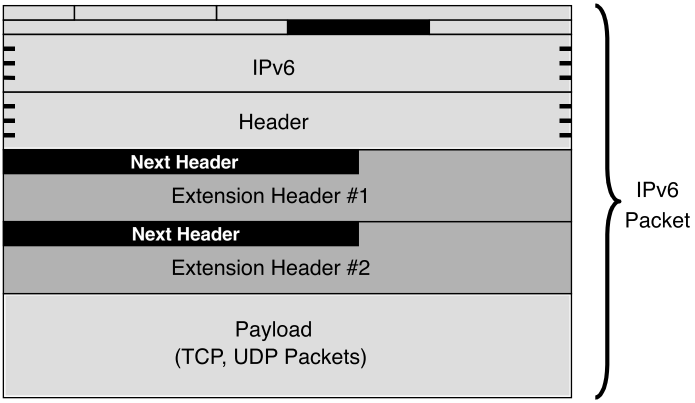
Within an IPv4 packet, an *Options* field (RFC 791) might be present at the end of the header. This *Options* field, when present, has a variable length, depending on the optional feature used between end hosts. Routers all along the path must compute this variable field length within each packet, even though the *Options* field is only used by both end nodes. Figure 1-16 illustrates the Options field within the IPv4 header.

Figure 1-16 *Options Field Length Within the IPv4 Header Has a Variable Length*



IPv6 uses a new approach to manage optional information in the header. Instead of using an Options field at the end of the header, IPv6 uses extension headers. Extension headers form a daisy chain of headers linked by a header field called Next Header, as shown in Figure 1-17.

Figure 1-17 *Extension Headers Are Daisy-Chained at the End of the IPv6 Header*



One *Next Header* field is present within every IPv6 extension header used. Many types of extension headers are defined for the different needs of IPv6 applications. This approach provides better efficiency in the option processing, because it ensures that routers and nodes compute only the headers targeted for them along a path.

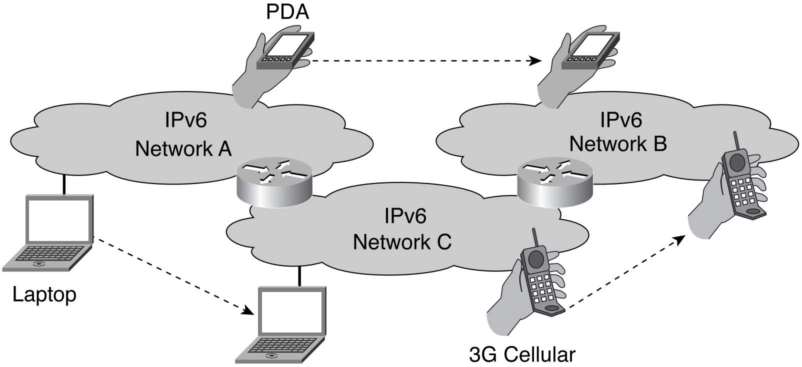
Mobile IPv6 is an example of a protocol using different extension headers for its operation when a mobile node is away from its home network. Extension headers provide important improvements to the Mobile IPv6 protocol as compared to Mobile IP, used in IPv4 networks.

Chapter 2 presents extension headers in detail.

Mobility

Mobility is a highly desirable and important feature for companies, organizations, and employees who want to access the web, e-mail, their bank accounts, and home from outside these networks, even from the car. New Layer 2 wireless technologies such as 802.11b and 3G (Third Generation) can help them fulfill these needs. 802.11b devices are cheap and can provide network connectivity with interesting bandwidth in several business locations such as offices, airports, and hotels. Billions of 3G cellular devices have an IP stack. In addition, cellular operators build IP core backbones based over IPv6. Thus, mobility with IPv6 is mandatory. Figure 1-18 illustrates the mobility provided by IPv6 networks.

Figure 1-18 *Many Devices Moving from IPv6 Networks to Others*



At the IP layer, the MobileIP protocol assumes that a node's IP address uniquely identifies the node's point of attachment to a network. A mobile node must be able to communicate with other nodes after changing its data link layer point of attachment without changing its IP address and breaking current connections. The MobileIP protocol lets nodes move from one IP network to another. The wireless/cellular industry uses the MobileIP protocol to grant IP mobility to wireless data.

MobileIP is available for both IPv4 and IPv6. However, with IPv6 the mobility is built into the protocol instead of being a new function added to IPv4. This means that any IPv6 node can use MobileIP as needed. MobileIPv6 uses the following IPv6 extension headers:

- A routing header for the registration
- A destination header for datagram delivery between mobile nodes and correspondent nodes

Both provide better performance for communications and enhancements to the IP protocol.

Security

The IPSec protocol, an IETF standard for IP network security, can provide several security functions:

- Access control limits access to people who have authorization.
- Authentication certifies that the person who sends data is who the person claims to be.
- Confidentiality ensures that any data carried over a public network, including passwords, is encrypted to make it very hard for anyone to see the exchanged data.