Wi-Fi Above 100 Mbps

802.11n

A Survival Guide

O'REILLY®

Matthew S. Gast

802.11n: A Survival Guide

Wireless has finally come of age. With a significant jump in throughput over previous standards, 802.11n is the first wireless technology that doesn't trade speed for mobility, and users have stormed onto wireless networks with a passion. In this concise guide, Matthew Gast—chair of the IEEE group that produced revision 802.11-2012—shows you why wireless has become the default method of connecting to a network, and provides technical details you need to plan, design, and deploy 802.11n today.

Building a network for the multitude of new devices is now a strategic decision for network engineers everywhere. This book gives you an in-depth look at key parts of 802.11n, and shows you how to achieve an Ethernet-free wireless office.

- Learn how MIMO's multiple data streams greatly increase wireless speed
- Discover how 802.11n modifications improve MAC efficiency
- Examine advanced PHY features such as beanforming and space-time code block
- Use advanced MAC features to maintain interoperability with older devices
- Plan an 802.11n network by determining traffic demand, key applications, power requirements, and security
- Choose the architecture, select hardware, and plan coverage to design and build your network

Purchase the ebook edition of this O'Reilly title at oreilly.com and get free updates for the life of the edition. Our ebooks are optimized for several electronic formats, including PDF, EPUB, Mobi, APK, and DAISY—all DRM-free.

Twitter: @oreillymedia
facebook.com/oreilly

US $19.99 CAN $20.99
ISBN: 978-1-449-31204-6

O'REILLY®
oreilly.com

802.11n: A Survival Guide

Matthew S. Gast

O'REILLY®
Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo

802.11n: A Survival Guide
by Matthew S. Gast

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*http://my.safaribooksonline.com*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

Editors: Mike Loukides and Meghan Blanchette
Production Editor: Rachel Steely
Proofreader: Kristen Borg
Cover Designer: Karen Montgomery
Interior Designer: David Futato
Illustrators: Robert Romano and Rebecca Demarest

Revision History for the First Edition:
2012-03-30 First release
See *http://oreilly.com/catalog/errata.csp?isbn=9781449312046* for release details.

ISBN: 978-1-449-31204-6

[LSI]

1333128185

In memory of Nicholas A. Jeremica
(December 5, 1977 - December 12, 2006)

And for Maria, Max, and Nicole,
who no doubt miss him even more than I do

Table of Contents

Foreword

Communications is changing our life in many ways. Our society is increasingly dependent on wireless communications technologies and the applications they support. In the same way that we can't understand how the folks who lived in 1800 managed without anesthesia, television, ... (insert favorite must-have technology here), our children won't understand how we managed to grow up without on-the-go email, social networking, ... (insert technology x[1] here).

IEEE 802.11 is a fundamental component of today's wireless communications technology. We have seen a rapid penetration of this technology into the enterprise (replacing or augmenting wired networks), home (wireless routers) and public spaces (airports, hotels). It is now no longer an optional feature, but an end-user expectation. IEEE 802.11n is the current "must-have" standard for wireless LAN that provides increased throughput and supports new applications and usage models. This book describes the development and operation of the 802.11n communications protocol. It provides background for understanding the impact of various features and settings on network operation and deployment. It is recommended reading for networking professionals and those who are curious as to what goes on under the hood. I have known Matthew Gast through our work together in the IEEE 802.11 working group—he as chair of 802.11REVmb and I as its technical editor. He writes clearly with a minimum of jargon and a maximum of relevant example usage. His personal experience of wireless LAN standards development, product development, and product deployment shines through. He has turned what might be a dry subject into a surprisingly easy read.[2] So, throw another log on the fire, pour your favorite libation, put your feet up, and settle down to...

—Adrian Stephens, January 2012

1. If I knew what "x" was now, I'd make a fortune and retire. Of course, "x" will be obvious to our children and they will wonder how we never thought of it now.
2. Not an easy task, but given that he has managed to turn IEEE 802.11REVmb comment resolution into occasional high comedy, absolutely nothing is impossible.

Preface

People still move. Networks still don't.

A decade ago, I first wrote that people moved, and networks needed to adapt to the reality that people worked on the go. Of course, in those days, wireless LANs came with a trade-off. Yes, you could use them while moving, but you had to trade a great deal of throughput to get the mobility. Although it was possible to get bits anywhere, even while in motion, those bits came slower. As one of the network engineers I worked with put it, "We've installed switched gigabit Ethernet everywhere on campus, so I don't understand why you'd want to go back to what is a 25-megabit hub." He underestimated the allure of working on the go.

In the rubble of the early 2000s Internet bust, wireless LANs were one of the first new technologies to cut through the gloom. As they gained in popularity, they morphed from an expensive toy to show off and became a must-have technology. My apartment had a wireless LAN very early, and I remember when I showed it off with pride. In the days of 2 Mbps 802.11 networks (or, if you were rich, an 11 Mbps 802.11b network!), all you had to do was beat the speed of your Internet link, which was not particularly hard.

What made 802.11n the next big step in the wireless revolution is that it was the first time that a wireless LAN delivered reasonable performance. Wireless LANs first took root in industries with highly mobile employees: hospital medicine, logistics, and education. Trading speed for mobility is a good exchange when you are on the move. It looks less attractive when you are relatively stable. 802.11n was the technology that buried the trade-off of speed for mobility. For the first time, it was possible to build a wireless network without compromise.

To build your network without wires, you don't have a choice: you're about to get familiar with 802.11n.

Audience

This book is about 802.11n, which was itself a major revision to the previous specification. To get the most out of it, you'll need to be familiar with the basics of the 802.11 MAC and how it orchestrates access to the medium. It will help to be somewhat familiar with how 802.11 networks were designed before 802.11n came along. In a sense, this book is the 802.11n-specific companion to the earlier *802.11 Wireless Networks: The Definitive Guide* (2nd edition), which was last revised in 2005.

The intended reader is a network professional who needs to delve into the technical aspects of 802.11n network operations, deployment, and monitoring, such as:

- Network architects responsible for the design of the wireless network at their place of business, whether the 802.11n network is the first wireless LAN, or an upgrade from a previous 802.11 standard
- Network administrators responsible for building or maintaining an 802.11n network, especially those who want to make the transition from earlier 802.11a/b/g technologies

One class of people for which I've specifically not written is the security officer. Ten years ago, "802.11" and "security" were not usually used together in the same sentence, unless it was derisive. Of the many changes that have occurred in the world of 802.11 since I last put virtual pen to virtual paper, none pleases me more than the acceptance of 802.11 as a readily secured network access layer. As an industry, we fought the battle to secure wireless LANs, and we won. My position as chair of the Wi-Fi Alliance's security efforts is generally boring—and I hope it stays that way!

Conventions Used in This Book

The following typographical conventions are used in this book:

Italic
: Indicates new terms, URLs, email addresses, filenames, and file extensions.

`Constant width`
: Used for program listings, as well as within paragraphs to refer to program elements such as variable or function names, databases, data types, environment variables, statements, and keywords.

Safari® Books Online

Safari Books Online (*www.safaribooksonline.com*) is an on-demand digital library that delivers expert content in both book and video form from the world's leading authors in technology and business.

Technology professionals, software developers, web designers, and business and creative professionals use Safari Books Online as their primary resource for research, problem-solving, learning, and certification training.

Safari Books Online offers a range of product mixes and pricing programs for organizations, government agencies, and individuals. Subscribers have access to thousands of books, training videos, and prepublication manuscripts in one fully searchable database from publishers like O'Reilly Media, Prentice Hall Professional, Addison-Wesley Professional, Microsoft Press, Sams, Que, Peachpit Press, Focal Press, Cisco Press, John Wiley & Sons, Syngress, Morgan Kaufmann, IBM Redbooks, Packt, Adobe Press, FT Press, Apress, Manning, New Riders, McGraw-Hill, Jones & Bartlett, Course Technology, and dozens more. For more information about Safari Books Online, please visit us online.

How to Contact Us

Please address comments and questions concerning this book to the publisher:

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472
800-998-9938 (in the United States or Canada)
707-829-0515 (international or local)
707-829-0104 (fax)

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at:

http://oreil.ly/802_11n

To comment or ask technical questions about this book, send email to:

bookquestions@oreilly.com

For more information about our books, courses, conferences, and news, see our website at *http://www.oreilly.com*.

Find us on Facebook: *http://facebook.com/oreilly*

Follow us on Twitter: *http://twitter.com/oreillymedia*

Watch us on YouTube: *http://www.youtube.com/oreillymedia*

Acknowledgments

It has been just long enough since I last wrote a book for me to forget the long slog that is writing. Although some writers will tell you that there is a tyranny in staring at a blank page, I have found that the hardest part is turning a completed rough draft into an actual book. As always, I benefited from a tremendous supporting cast at O'Reilly,

starting with Mike Loukides, who got the project off the ground, and Meghan Blanchette, who took on the much harder task of making me finish.

I could not have asked for a better review team, which includes several 802.11 luminaries who are famous in their own right. My all-star review team consisted of (in alphabetical order):

David Coleman
: A well-known wireless LAN author in his own right, and one of the longest serving technical instructors on Wi-Fi there is, David devoted his own precious time to his review of the manuscript. I can't figure out where he found the time, since David is one of the few people who travels more than I do.

Joe Fraher
: Joe is simply the best tech writer I know in the networking industry, and granted my request for a review without hesitation. When I write, I have a large supporting cast at O'Reilly who take my draft work and turn it into something professional. When Joe writes, he does everything himself, including the illustrations. For a long time, he even had the unthinkable task of producing all of Aerohive's documentation himself.

Craig Mathias
: Craig is one of the best-known wireless LAN analysts in the industry. He has been a champion of the first edition of my book since it came out a decade ago, and I was happy that he took the chance to correct this book before publication.

Bob O'Hara
: Bob was the original technical editor of the 802.11 standard when it came out, and he later co-founded the wireless LAN company Airespace. He graciously gave his time and shared a great deal of historical knowledge about why 802.11 is designed the way it is.

Changming Liu
: When he started Aerohive, Changming went all-in on 802.11n as the future of both wireless LANs and his newborn company. In addition to being dead right about 802.11n, he is a person that I learn something from every time I talk with him. I only wish that I'd had more opportunities to learn from him in the past two years.

Andrew von Nagy
: Andrew is the wireless architect at a major retailer, and creates incredibly deep and detail-oriented posts on the technology at Revolution Wi-Fi. In building a network, you have to blend the theoretical knowledge of the protocol with practical expertise in how equipment works. Andrew contributed commentary based on living, breathing network deployment every day.

Adrian Stephens
: Adrian knows 802.11 as well as anybody I know, with the added benefit that he can explain almost any detail better than anybody else alive. Adrian gets around Cambridge by bicycle, and it is a long-running item of gallows humor in the 802.11

working group that a bus driver in Cambridge having a bad day could easily set back the 802.11 standards effort years by hitting the wrong person. I am indebted to Adrian for his guidance when I served as chair of the 802.11 revision task group, as well as for the foreword.

Tim Zimmerman
: Tim was present at the creation of 802.11. He was a voting member of the working group that approved the first published 802.11 specification, and he was involved in an organization called the Wireless LAN Association, which became the Wi-Fi Alliance. Few people can combine deep historical knowledge with a broad sweep of knowledge across an industry from the vantage point of an analyst.

In addition to the formal review team, I benefited from the assistance of several others.

Many companies treat authors with suspicion, if not outright hostility, and I am lucky that Aerohive is an exception to this practice. Adam Conway, my boss, immediately supported my book proposal, even though having one of his employees write a book was a novel experience for him. Aerohive's marketing team was a source of repeated encouragement throughout the process; I would specifically like to recognize Stephen Philip, Zena Baloca, and Jenni Adair for their many motivational conversations during the book's gestation.

Terry Simons, a talented test engineer for a major Wi-Fi player, was a never-ending source of commentary and suggestions—and I mean that in a good way! Chris Hessing never fails to inspire me to learn more about technology, and offered to read this book before I asked.

I owe special thanks to the Wi-Fi Alliance's marketing staff. Kelly Davis-Felner was incredibly responsive to my permission requests. My day-to-day work as a task group leader at the Wi-Fi Alliance has been ably supported by both Sarah Morris and Kevin Robinson, both of whom contribute to the industry in ways too innumerable to list here.

CHAPTER 1
Introduction to 802.11n-2009

I feel the need...

...the need for speed!

—Maverick and Goose, *Top Gun*

802.11n's appeal can be summed up in one word—speed. In the stone ages of 1 Mbps and 2 Mbps wireless LANs, the most charitable thing that one could say about the technology was that it was faster than DSL. With the benefit of a decade of development, wireless LAN technology is mature and widely viewed as "the way you connect to a network," just as Ethernet was in the mid-1990s. When I started working, visitors would often ask where the nearest Ethernet port could be found. Today, they ask for a password for the wireless LAN. The speed of 802.11n is a large part of the reason why so many devices don't even have the option for Ethernet ports any more. Once 802.11n put wireless LANs on the same performance footing as wired LANs, wireless became the obvious choice.

History

802.11n began life in late 2002, and took exactly seven years to produce from start to finish.[1] As with all IEEE 802 standards, it began as a study group. The High Throughput Study Group was founded to investigate building an 802.11 physical layer that could provide speeds of 100 Mbps, as measured at the Medium Access Control (MAC) layer. An early rule of thumb for wireless LANs is that the throughput at the MAC layer would be half of the transmission speed. In 802.11b, the 11 Mbps transmission speed topped out at 5-6 Mbps of throughput; in 802.11a/g, the 54 Mbps transmission speed resulted in 25-30 Mbps of throughput. 802.11n's goal was to increase the efficiency of the

1. As with all 802.11 standards, the best place to check on status is the official 802.11 timeline. In addition to estimating the milestone dates for drafts that are in progress, it offers a look back at the dates of milestones for existing standards.

protocol so that increases in transmission speed would not be held back by protocol overhead.

The first major milestone in the development of 802.11n was the second draft. Although a few products were made based on 802.11n draft 1.0, interoperability was spotty. After resolving some 12,000 comments on the first draft, a much more robust second draft was produced in early 2007. With a great deal of the ambiguity removed, draft 2.0 products had the expected high degree of interoperability. Around the same time, the Wi-Fi Alliance launched a certification program to drive interoperability of the then-emerging 802.11n standard. That certification program was a wild success. At the member meeting in the summer of 2008, the cake in Figure 1-1 was tangible (and tasty) evidence of the success of the program.

Figure 1-1. 802.11n draft 2.0 celebration cake

I remember when 802.11n was finally approved in September 2009. I was visiting Australia, and woke up to the anticipated news, which was that 802.11n had been approved by the standards board. As a 500-page specification, there was a great deal of work that went into the publication process. One of my favorite bits of memorabilia from my participation in the 802.11 working group is the mouse pad in Figure 1-2. I keep it around as a reminder that it's more important to get something done right than to get it done as quickly as possible.

The Technology of 802.11n

In the "traditional" 802.11a/b/g world, the link layer and physical layer were separate entities. Improving speed was largely viewed as a task for the physical layer. The way to improve network speed was to make bits go faster.

Figure 1-2. 802.11n meeting schedule mouse pad

Physical Layer (PHY) Enhancements

The basis of the raw speed in 802.11n is Multiple Input/Multiple Output (MIMO) technology, which allows a single radio channel to support multiple data streams. Before 802.11n, the transmitter and receiver were Single Input/Single Output (SISO) devices. From the transmitter's antenna, the same data stream flew out in every direction, bouncing off walls and other obstacles, and then arrived at the receiver. If two paths between the transmitter antenna and the receiver antenna were out of sync, the resulting signal could be quite weak due to the interference between paths. This phenomenon, known as multipath interference, was the bane of network designers because moving an access point slightly could dramatically improve coverage. As client devices moved around, they could move from "hot spots" to "cold spots" due to multipath interference. In a MIMO system, the transmitter and receiver can take advantage of multiple paths. Each path gets a different set of data, and therefore, the resulting transmission

is not subject to the same destructive effects of multipath interference. In fact, without multiple paths, the benefits of MIMO are significantly limited.

802.11n came along at an interesting point in the evolution of 802.11. Prior PHYs had been targeted at particular radio bands. The original 802.11 frequency-hopping and direct-sequence PHYs were specified only in the 2.4 GHz ISM band. 802.11a was developed when the 5 GHz band was opened for license-free use in the United States, and subsequently extended when regulations were liberalized in many other countries. 802.11g was an effort to take the technology behind 802.11a and make it available in the 2.4 GHz band. 802.11n, however, was designed when both bands were available, and therefore, it is "frequency agnostic."

802.11n can use both radio bands (2.4 GHz and 5 GHz). Manufacturers use a variety of terms to tell the user what band the radio operates in. One of the more common appellations is "802.11ng" for 802.11n in the 2.4 GHz band, and "802.11na" for 802.11n devices in the 5 GHz band. Dual-band devices may be labeled in a variety of ways, such as "dual-band 802.11n" or "802.11agn."

To increase speed beyond the capabilities offered by MIMO, 802.11n offers the option for wider channels. By doubling channel width, it is possible to double data rates. Network administrators must carefully consider a set of trade-offs in using wider channels. In return for higher speed, radio planning becomes more complex due to a higher demand for spectrum, and coexisting with previously installed networks based on 20 MHz channels becomes a concern. Wider 40 MHz channels also have a higher potential to interfere with non-802.11 technologies such as Bluetooth, which is one of the major reasons why 40 MHz channels must be disabled by default in the 2.4 GHz band.

MIMO technology can also be enhanced for beamforming. With an antenna array, it is possible to arrange transmissions such that the energy is "focused" or "directed" towards a particular physical location. By concentrating energy in one direction, it is possible to improve the signal-to-noise ratio and the transmit speed, though a complex set of trade-offs also limits the raw capability of beamforming.

Link Layer (MAC) Enhancements

A decade ago, it was routine to estimate the flat-out top speed of a wireless LAN by taking the "headline" rate promised by the PHY and cutting it in half. Various forms of protocol overhead, such as medium contention and positive acknowledgment, meant that the best-case scenario for 802.11 was that the 802.11 MAC had approximately 50% efficiency. The network may advertise a 54 Mbps data rate, but the best you can hope for is generally 25-27 Mbps. In many cases, the estimate of 50% efficiency was overly generous. Just as with many other markets, the wireless LAN market has vendors that produce implementations with many optional performance-enhancing

features (both standards-based options and vendor-specific options), in addition to vendors that produce only the bare minimum required to ship a product.

802.11 manages contention by allocating access to the network medium for transmission, but not all transmitters are created equal. As speeds increase, the required interframe spaces have a higher cost. In an 802.11b network, the opportunity cost of a 10 μs short interframe space (SIFS) is about 14 bytes. When the data rate increases to a moderate 802.11n rate, the cost of the same SIFS jumps to well over a kilobyte.

A significant portion of 802.11n is devoted to improving the efficiency of the MAC. While many users of 802.11 equipment focus on the high data rates, the improved efficiency is a major source of the speed improvements. The main technique in 802.11n for enhancing efficiency is frame aggregation. Each transmitter needs to pay a cost to access the medium, but frame aggregation spreads that cost over several smaller frames. Depending on the type of data being transmitted, aggregation can improve efficiency from 50% to about 75%.

Although not directly related to transmission efficiency, the 802.11n MAC extends the power-saving capabilities of 802.11. 802.11n radio cards have high power consumption. While this is not a huge drawback in a device with extensive reserves of battery power such as a laptop, reducing power consumption in smaller battery-operated devices such as phones and tablets is necessary.

802.11n: First We Take the LAN, Then We Take the World

802.11n offers great power, but has significant complexity. That complexity may be expressed in a variety of ways. When the 2012 revision of 802.11 first incorporated the 802.11n amendment, the text grew in size by about 50%. Another favorite factoid of mine is that there is a clause that lays out the possible data rates for each potential value of the array of parameters selected for transmission. The resulting tables are a notable fraction of the overall specification.

802.11n has over 300 different data rates! If you feel confused, you're not alone.

Bob O'Hara, the original technical editor of the 802.11 standard, once admonished me to "never confuse standardization with building a product."[2] The complexity of 802.11n is an excellent illustration of his point. With such a complex standard, it is

2. Or, as one of the reviewers (who you will note I have carefully not named) put it, "there're also parts of the standard that are there for 'political reasons' to get support from particular groups, but which will never be implemented. They are put there to ensure perfection of the standard, because a standard is perfect when everybody is equally unhappy with the outcome."

simply not possible to overcome the hardware engineering challenges and have the maximum capabilities available when the standard is published. In the case of 802.11n, product support of capabilities comes in waves. The first wave of 802.11n that hit in 2007 was composed of products that supported sustained 100 Mbps throughput, and, with the configuration of protocol options for maximum speed, up to a 300 Mbps link rate. The second wave hit in 2011, with products that featured data rates up to 450 Mbps. Meanwhile, many low-power devices upgraded to 802.11n single-stream technology with small improvements in speed.

The interplay between standardization and building product is, in effect, managed by the Wi-Fi Alliance, an industry trade organization. All major wireless LAN vendors are members of the Wi-Fi Alliance. With a complex standard such as 802.11, trade organizations serve a valuable purpose in focusing development effort on the most valuable features for end users.[3]

The Wi-Fi Alliance works to create certification programs for emerging standards. As a general rule, the Wi-Fi Alliance will have an initial certification program based on a solid subset of the standard that is available as the technology is developed, and then they launch a second program based on the approved standard. While the technology is being developed, the Wi-Fi Alliance certification programs often serve as a key target for product manufacturers, who will build products capable of passing the certification test.

3. In the past, various industry groups such as the Frame Relay Forum, the ATM Forum, the MPLS Forum, and the DSL Forum have served similar roles in their respective technologies.

PART I

The PHY

In IEEE 802 networks, the physical layer, or PHY, is responsible for taking bits and turning them into a physical representation for transmission on the network medium. Every 802.11 physical layer has come with its own "big idea" that is the core driver for development. In the 1990s, the driver was just providing connectivity. Once that was established, 802.11b used Complementary Code Keying (CCK) to push the original 2 Mbps speed up to 11 Mbps. Simultaneous development on 802.11a used Orthogonal Frequency Division Multiplexing (OFDM) to establish the blazing speed of 54 Mbps. 802.11a required the use of a new spectrum, and the big idea in 802.11g was to bring the OFDM technology into the 2.4 GHz band to make higher speeds widely available.

In 802.11n, the big idea is MIMO. To transmit data, 802.11n can send multiple simultaneous data streams. Understand MIMO, and you understand the key to 802.11n. 802.11n broke new ground in other ways as well, most notably by standardizing wider channel bandwidths. In addition to the increase in the raw data rate from the PHY, it also included several efficiency enhancements in the MAC. Understanding 802.11n begins with understanding how the physical layer can move data so quickly.

CHAPTER 2

MIMO and the 802.11n PHY

Row, row, row your boat,

Gently down the stream…

—Row, Row, Row Your Boat (traditional)

Increasing the speed of a network can be done in two ways. First, protocol designers can try to increase the raw speed as measured by the transmission rate, which might be termed the "go faster" approach. Increasing transmission rates has been the primary tool for the dramatic increase in wireless LAN speeds from 1 Mbps in the original 802.11 standard to 54 Mbps in 802.11a and 802.11g. Second, protocol designers can increase the speed perceived by users by increasing the efficiency of the protocol, to transmit more bits within a given period of time, an approach which might be called the "efficiency" approach. Although 802.11n uses both techniques, most of the gains come from dramatic increases in data rate. At the core of the "go faster" approach, 802.11n uses MIMO. Early 802.11n products transmitted at a data rate of 150 Mbps, and the standard laid out a clear path to data rates of up to 600 Mbps. These speeds are achievable only through the application of MIMO.

The Big Idea: MIMO and Data Streams

Before MIMO, 802.11 used a single data stream. A transmitter used one antenna, and a receiver used one antenna.[1] The transmission link in pre-802.11n devices can be described in terms of its two components. It was called Single-Input because the receiver used one antenna, and Single-Output because the transmitter used only one antenna. Taken together, the communication system was called Single-Input/Single-Output (SISO).

1. Most 802.11 devices used antenna diversity, which is a way for a device to pick the "best" antenna out of a set. For any given transmission, only one antenna was active.

Between the two endpoints in a SISO system flows one set of data, called a *stream*. If a data link is like a highway, then SISO is a two-lane country highway. (I'm stretching the analogy slightly because highways are bidirectional but 802.11 is half-duplex, so bear with me.) The highway certainly works—traffic can move from one point to the next. 802.11 was a successful technology when 802.11a/g systems were commonplace. However, there was a speed limit. The two-lane country highway had become too popular, and the 802.11 working group needed a solution. One method would be to increase the speed limit along our road. Instead of having cars travel at highway speeds of 60 miles (100 kilometers) per hour, we can impose a new rule that cars on our two-lane highway need to go faster. Instead of what we think of as normal highway speeds, cars now need to travel at 250 miles (400 kilometers) per hour. Extending the highway analogy, the dramatic improvements required for driver alertness and reaction time at higher speeds are like Shannon's Law.[2] Yes, it's possible to increase the speed to increase throughput, but it requires that the endpoints in a communication system distinguish between finer and finer details.

Instead, to increase the throughput of our highway, we decide to widen it. Instead of a single lane in each direction, we have two lanes. Although there are certainly costs involved in widening the highway in terms of land use, perhaps a center divider, and traffic signals, we can keep the speed limit where we know most motorists can handle the driving task. Widening the highway is similar to what the 802.11 working group did with 802.11n. Instead of just a single transmitter and a single receiver in the system, both sides now have multiple antennas. That is, the receiver has multiple inputs, and the transmitter uses multiple outputs.[3]

Figure 2-1 is a simplified high-level comparison of early 802.11 SISO systems to MIMO systems. In the SISO system, a single active antenna transmits to a single active antenna. Although SISO systems may have multiple antennas, only one is active for any given data frame. In the MIMO system, all of the antennas are active simultaneously. Each antenna in the MIMO transmitter sends its own data stream as input into the radio channel (hence, multiple input), and each antenna in the MIMO receiver collects its own data stream as output from the radio channel (hence, multiple output). The figure simplifies the picture by showing the simple case of a data stream going between pairs of antennas; in real-world systems, complex matrix math is used to create multiple data streams through the radio channel. In the ideal case, each pair of antennas in the MIMO system is capable of transmitting its own independent data stream, just as each lane on our highway carries its own set of cars.

2. Shannon's Law is a mathematical relationship between the bit rate of a channel, as described by the bandwidth of the channel and its signal-to-noise ratio. Effectively, it states that the speed of a transmission channel can be made as large as desired as long as the signal-to-noise ratio increases in tandem.

3. One reviewer asked how to work wider channels into my analogy, and the best I could come up with is that wider channels are like wider lanes. If we widen our two-lane highway so the lanes are wider, articulated tractor-trailer trucks are able to use the road and increase the "throughput" of the highway as measured in cargo, instead of the number of motorists per unit of time.

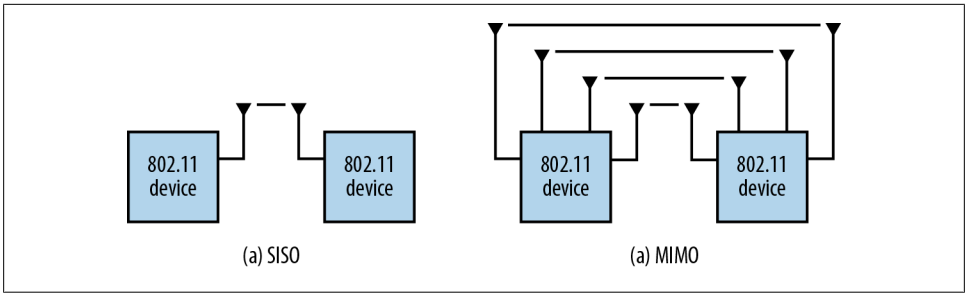


Figure 2-1. Comparison of SISO transmission to MIMO transmission

Antenna Diversity and MIMO Compared

802.11a/b/g devices often had multiple antennas to take advantage of *antenna diversity*. If an 802.11a/b/g SISO device has multiple physical antennas, and an 802.11n MIMO device has multiple antennas, what makes those two technologies different? The answer lies in how the radio chip is laid out, and the complexity of the transceiver.

In an antenna diversity system, there is really only one transceiver. Multiple antennas may feed into that one transceiver, but when the receiver is active, it must choose which of the antennas is connected to the electronics in the receive chain. Only one antenna can be connected to the receiver electronics, no matter how many antennas may sprout from the device. Many high-end 802.11a/b/g devices implemented antenna diversity, and would typically have two antennas for each radio band. When a signal was received, the antenna diversity system would choose the antenna to be connected to the electronics. Typically, 802.11a/b/g devices with diversity would choose the antenna that received the strongest signal.

Upon transmission, an antenna system must select only one of the antennas to use. Again, only one of the antennas in a diversity system may be active. Many 802.11a/b/g devices always used one "primary" antenna for transmission. A few 802.11a/b/g devices used a slightly more sophisticated algorithm. When receiving a frame from another 802.11 device, these APs would save the antenna that was active for the reception and then use the same antenna for future transmissions to that peer.

The distinguishing characteristic of antenna diversity is that there is really only one transmitter, but it does get to choose the "best" single antenna. Antenna diversity had value because the antennas could be located at a precise distance apart so that when a transmission was interfering with itself due to multipath interference, one of the two antennas would be in a "cold spot" and one would be in a "hot spot." Typically, the difference is around half a wavelength, and at 2.4 GHz, the wavelength is only a few centimeters. Because the transceiver is switching between antennas, sometimes antenna diversity may be called *switched diversity*.

In MIMO, each antenna can be driven independently. Antenna One can transmit (or receive) a completely different set of bits from Antenna Two. If the two antennas can transmit streams that remain independent at the client, then you can double your

throughput by using both. Simply put, antenna diversity is about picking the best single path between two devices. MIMO is about using all of the available paths.[4]

Processing of independent paths also opens up two additional uses, both of which will be explored in the PHY part of the book. Transmit beamforming uses the multiple streams to increase signal strength at the receiver. On the receiving side, if the number of antennas exceeds the number of spatial streams, *receiver combining* techniques can also provide a boost to signal strength. At a hardware level, circuitry for MIMO and receiver combining is identical, so improved reception is "free" once you have built a MIMO antenna system.

Spatial Streams

Figure 2-1 in the previous section demonstrated the core idea of MIMO, which is the transmission of multiple data streams across the same radio channel. Each data stream is sometimes called a *spatial stream* because it is a separate path through the space between communication peers. When a MIMO link can only transmit using a single stream, it isn't really MIMO and there isn't any throughput improvement. (Imagine, I guess, our newly expanded four-lane country road being under construction and being back to a single lane in each direction.)

Spatial streams are created by having multiple independent paths through space between two devices; in a typical 802.11 network, those two devices are an access point and a client device. Figure 2-2 depicts a situation that was the bane of every 802.11a/b/g network administrator. In the figure, the transmitter and receiver have two paths. One is a direct line-of-sight path, and a second bounces off a wall and is exactly out of phase. If the two paths are being used to transmit the same set of bits, they will interfere destructively and there will be no signal to read at the receiver. However, MIMO systems perform a radio-channel alchemy of sorts. Instead of the worthless lead of a cold spot, MIMO exploits multiple paths through space and transmutes the lead into the gold of a hot spot with twice as much throughput. Because the two paths do not interfere with each other, independent transmissions can be sent through each path, doubling throughput. In the language of the field, the degree of similarity between the paths is called their *correlation*; in the figure, the two paths are said to be *uncorrelated*. When designing products, the RF designers typically spend quite some time thinking about antenna placement in order to minimize the correlation so that overall system throughput can be improved.[5]

4. It's even possible to combine both diversity and MIMO in the same system, but the benefits are beyond the scope of this book, and certainly beyond the scope of a sidebar.

5. Multipath is a key ingredient to the success of a MIMO system. As one reviewer put it, "In space, nobody can hear you scream, and MIMO doesn't work because there are no uncorrelated paths." That is, if there is only one path between two points, what are you going to bounce off of?

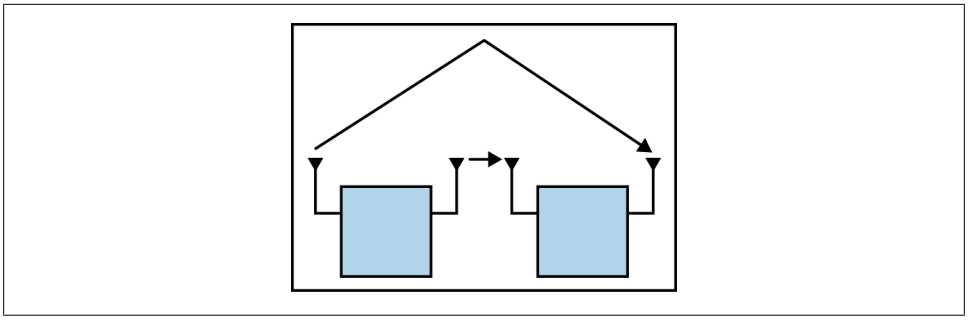


Figure 2-2. Creating spatial streams with multipath

802.11n reuses the channel structure and modulation techniques of 802.11a/g, called Coded OFDM. Each path in an 802.11n MIMO system is roughly equivalent to a single 802.11a/g transmission, but MIMO lets you double, triple, or quadruple up. 802.11n supports the use of up to four spatial streams. Another way to look at how spatial streams benefit 802.11 users is to look at a measurement called the spectral efficiency. It's very easy to increase the speed of a network technology by increasing resource utilization. In 802.11, the key resource that needs to be optimized is the radio spectrum. Table 2-1 compares the 802.11 physical layers in terms of spectral efficiency, which is defined as the number of bits that can be sent per unit of spectrum; obviously, higher spectral efficiency is better. Because 802.11n has two channel widths, there is an entry for each channel width representing first-generation 802.11n hardware.

Table 2-1. Spectral Efficiency Comparison

802.11 PHY	Spectral Efficiency (Mbps/MHz)
802.11 direct sequence / 802.11 frequency hopping	0.09
802.11b	0.5
802.11a/g	2.7
802.11n (20 MHz channels, MCS 15)	6.5
802.11n (40 MHz channels, MCS 15)	6.75

Radio Chains

Between the operating system and antenna, an 802.11 radio interface has to perform several tasks. When transmitting a frame, the main tasks are the inverse Fourier transform to turn the frequency-domain encoded signal into a time-domain signal, and amplification right before the signal hits the antenna so it has reasonable range. On the receive side, the process must be reversed. Immediately after entering the antenna, an amplifier boosts the faint signal received into something substantial enough to work with, and performs a Fourier transform to extract the subcarriers. In an 802.11 interface, these components are linked together and called a *radio chain*. Selecting the components to make up the radio chain is an important task for system designers, especially

for those who make infrastructure devices. Generally, an access point will have much higher-quality components in its radio front-end, so an AP's transmitted signal can go much farther than a client's signal. To ensure that transmission and reception performance is closely matched, a condition referred to as a *symmetric link*, system designers must match the transmit and receive amplifiers; if especially powerful transmit amplifiers and high-gain antennas are used, designers must also ensure that the receive side has equal capabilities.

In the single-stream SISO world of 802.11a/g, only one Fourier transform and one amplifier are needed in the transmit chain. In order to transmit independent bits along each path, however, 802.11n requires multiple radio chains. If each spatial stream is to carry different bits, those bits must be processed by their own individual components. Figure 2-3 shows a simplified block diagram for an 802.11n interface with four radio chains. Building a radio interface that has multiple radio chains is a significantly more complex undertaking than making it work with just one radio chain. (For example, compare it to Figure 13-17 in *802.11 Wireless Networks: The Definitive Guide*, which shows an OFDM PHY block diagram.) One of the major reasons that 802.11n has rolled out in phases is that it was necessary to tackle the relatively easy problem of building two-chain devices before building more complex three- and four-chain devices.[6]

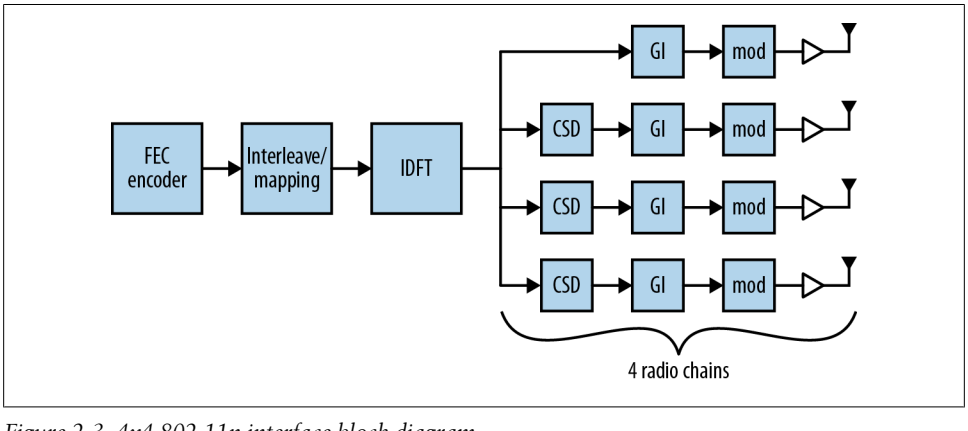


Figure 2-3. 4x4 802.11n interface block diagram

Multiple radio chains can dramatically increase speed, but they have a negative effect as well. Fourier transforms use complex mathematics, and have significant power requirements. Amplifiers, too, are power-hungry. When receivers sample the channel, the power required depends on both the channel width and the number of spatial streams, both of which increase in 802.11n. Multiple radio chains can be used to

6. We'll return to the full block diagram; I've made some simplifications in this picture to show how data flows. The purpose of this figure is to illustrate the difference between a SISO system and a MIMO system, not to offer a full description of an 802.11n interface at this point.

increase speed, but they also consume significant power. Managing power consumption was an important task for the 802.11 working group, and it is an important task for 802.11n system designers building hardware based on power-hungry components.

MIMO Notation

A recurring theme of this book is how 802.11n uses shorthand notation to simplify complex concepts into something compact enough to write or speak easily.

T × R : S

The maximum data rate is determined by the number of streams S. Each stream must have its own radio chain on both the sending and receiving side, so both T and R must be greater than or equal to S. Essentially, each stream must have its own transmitting antenna to leave and its own receiving antenna to arrive at. A two-stream transmission requires that both the transmitter and receiver have at least two radio chains.

As a network administrator, it's important to get all three numbers. Early enterprise-grade 11n devices were advertised as "3×3" devices, which they certainly were because they had three radio chains. In spite of having three radio chains, they were only capable of two spatial streams. That is, they were 3×3:2 devices, and have the maximum speed of a two-stream device. The extra radio chain certainly helps performance, but the difference between a 2×2:2 and a 3×3:2 system is marginal.

Relationship Between Spatial Streams and Radio Chains

Each spatial stream needs to have an independent transmitter and receiver, so as you might have guessed, the number of radio chains must be greater than or equal to the number of spatial streams. As part of the MIMO extensions, 802.11n devices can negotiate the number of streams used, and devices will find the highest common speed they can use. In fact, there is a feature in 802.11n, *Space-Time Block Coding* (STBC), that requires two radio chains to transmit a single spatial stream. By spreading the spatial stream across two radio chains and two paths, it's possible to increase the redundancy in transmission to offset path loss, though at the cost of overall transmission speed.

Even without STBC, "extra" radio chains offer a benefit. Many devices implement *Maximal Ratio Combining* (MRC), which uses information from all radio chains to process a frame. MRC works by cleverly combining the information received at each antenna by taking the strongest components of the received signal from each antenna. As a rough example, say that the transmitted signal consists of the sentence "Wireless networks are cool." At the receiver side, it may be that one antenna receives "Wireless," a second antenna receives "networks," and the third antenna receives "are cool." MRC enables the receiver to put all three pieces together into a coherent whole. At a more technical level, what MRC enables a receiver to do is pick the best antenna for each carrier in the 802.11n channel to compensate for individual fades in sub-channels.

Additional information from the supplemental radio chains increases the accuracy of the MRC process. MRC is widely implemented in enterprise-grade APs, and it offers a benefit for all devices, including reception of transmissions from older non-802.11n devices. The antennas in many battery-powered devices are designed around the device case and aesthetics, not the optimal layout for maximum radio performance. By using a 3×3 AP, network administrators can gain extra sensitivity to decode the weak signals from even single-stream clients. Figure 2-4 shows the conceptual range increase when MRC is used. The solid line on the graph is the rate-over-range plot for reception of a two-stream device by a 2×2 MIMO system; the dashed line is the rate-over-range plot for reception of the same two-stream transmission by a 3×3 MIMO system. There is no increase in peak speed because both systems are operating as two-stream receivers. However, the additional range from MRC means that for a given distance, MRC yields higher data rates. Conversely, for a given data rate, the MRC-enabled receiver can hear weaker signals and enable greater range.

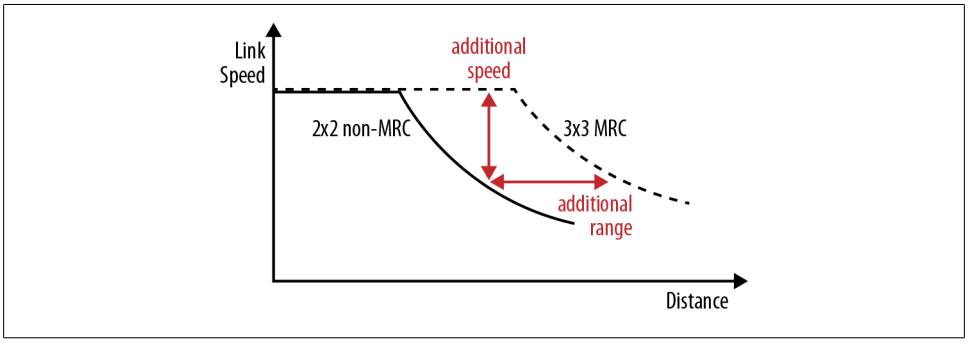


Figure 2-4. MRC range increase

CHAPTER 3
Channels, Framing, and Coding

I like big channels, I cannot lie…

—"Baby Got Bandwidth" (with apologies to Sir Mix-a-Lot)

This chapter moves from the rarefied theoretical discussion of how MIMO works into the details of how the 802.11n PHY interfaces with the physical medium and delves into the details of the techniques that are used to increase speeds.

Channel Structure and Layout

The structure of a channel in 802.11n is basically the same as 802.11a/g. Both are based on OFDM, and therefore, both divide up the radio channel into a number of subcarriers that are packed closely together and precisely enough that they are orthogonal to each other. 802.11n provides several minor improvements to the structure of the channel. Like 802.11a/g, it uses OFDM and re-uses the same modulations and numbering scheme.[1]

Channel Structure

802.11n offers two features to increase the utilization of the radio spectrum. 802.11n retains the common 20 MHz channel width used by prior 802.11 standards. Within the 20 MHz channel, however, 802.11n improves spectral efficiency by adding subcarriers that were unused in 802.11a/g, as shown in Figure 3-1. Even though 802.11n adds four data subcarriers, increasing throughput by about 8%, it does not need to add any pilot subcarriers. Pilot subcarriers are used to provide channel measurement and calibration, and are a form of overhead. Just as MIMO increases the efficiency of a data transmission, it increases the efficiency of the pilot carrier operation. In a MIMO

1. For an introduction to the use of OFDM in 802.11a/g, see Chapter 13 in *802.11 Wireless Networks: The Definitive Guide*.

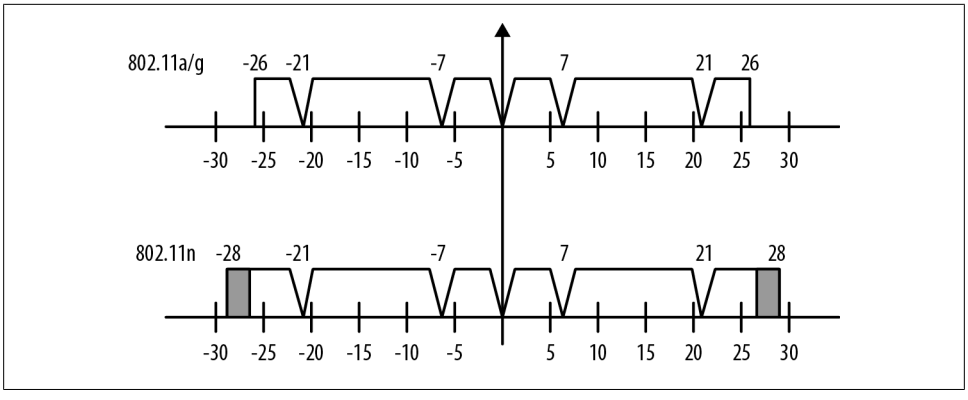


Figure 3-1. Channel structure comparison between 802.11a/g (52 carriers) and 802.11n (56 carriers)

system, each subcarrier will be received through each of the received radio chains, and thus, will provide more information on the state of the channel as in a SISO system.

The second change made by 802.11n is that it supports operation in wider 40 MHz channels. Although the standard describes several methods of operating a 40 MHz channel, by far the most common one is that two adjacent 20 MHz channels are treated as one channel and treated as a single 40 MHz contiguous block.

Even though the center frequency of a 40 MHz channel moves upward, the "name" of the channel does not change. For example, a 20 MHz channel operating on channel 60 can be used next to a 20 MHz channel operating on channel 64. However, an AP advertising a 40 MHz bandwidth at channel 60 occupies the spectrum for both channel number 60 and channel number 64. Figure 3-2 illustrates the differing channel widths.

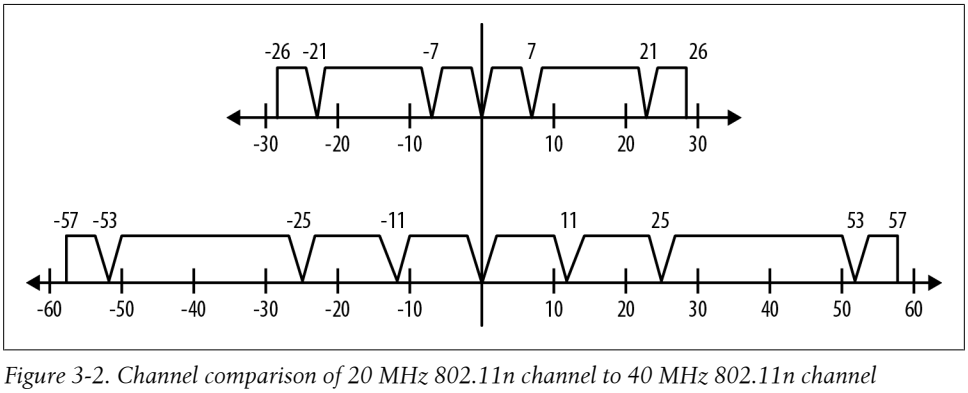


Figure 3-2. Channel comparison of 20 MHz 802.11n channel to 40 MHz 802.11n channel

802.11n's 40 MHz channels do more than double the throughput when compared to the traditional narrow channels because the 40 MHz channel format decreases overhead as a fraction of the channel. Pilot carriers are an "overhead" expense required in OFDM, and do not transmit any data from higher protocol layers. 802.11n doubled

the channel width from 20 MHz to 40 MHz, but only increased the number of pilot carriers by half, as described in Table 3-1. By using the increased effectiveness of pilot carriers in a MIMO system, the spectral efficiency increases by 4%.

Table 3-1. Channel description attributes

PHY standard	Subcarrier range	Pilot subcarriers	Subcarriers (total/data)
802.11a/g	-26 to +26	-21, -7, +7, +21	52 total, 48 usable
802.11n, 20 MHz	-28 to +28	-21, -7, +7, +21	56 total, 52 usable
802.11n, 40 MHz	-57 to +57	-53, -25, -11, +11, +25, +53	114 total, 108 usable

Regulatory Rules and Operating Channels

As with anything related to radio transmission in 802.11, regulatory considerations are an important part of how much data a network can support. Although the 802.11 specifications define a large number of channels, especially in the various 5 GHz bands, they are only available for use if allowed by the regulatory authority in the country where an AP is installed. 802.11n operates in both the 2.4 GHz band used by 802.11b/g as well as the 5 GHz band used by 802.11a.[2] In both cases, 802.11n reuses the channel numbering that was established by previous standards. Figure 3-3 illustrates the latest available information on regulations in the 5 GHz band. Channels are identified by the IEEE channel number across the top of the diagram, and blocks indicate whether a 20 MHz or 40 MHz channel is available for use. The broad band in the middle of the figure, 5.470–5.725 GHz, has been the subject of intense interest. It is generally available throughout the world for use, and represents a substantial increase in available spectrum for building wireless LANs.

One of the reasons why the 5.470 GHz spectrum became available is that its previous use generally allowed for a secondary user of the band. It is widely used for weather radar. Wireless LAN use of the band is secondary, meaning that if radar signals are detected the wireless LAN must be automatically shut down. However, such a high frequency generally does not impinge on operation of indoor wireless LANs. This band is sometimes referred to as the *Dynamic Frequency Selection* (or DFS) band after the required procedures to detect radar and avoid interfering with the allocated use of the band. Certification to operate within the DFS band is generally granted by the radio regulator in a country after testing to ensure compliance with the relevant rules.[3] Use of 40 MHz channels substantially increases the demand on spectrum resources, and the ability to use 40 MHz channels will typically depend on whether the network components can support channels on which DFS operations are required.

2. Potentially confusingly, the operation of 802.11n in the 2.4 GHz band is generally called *802.11ng*, while operation of 802.11n in the 5 GHz band is generally called *802.11na*.

3. 802.11 defined several protocol operations to support the use of DFS, most notably in the 802.11d and 802.11h amendments.

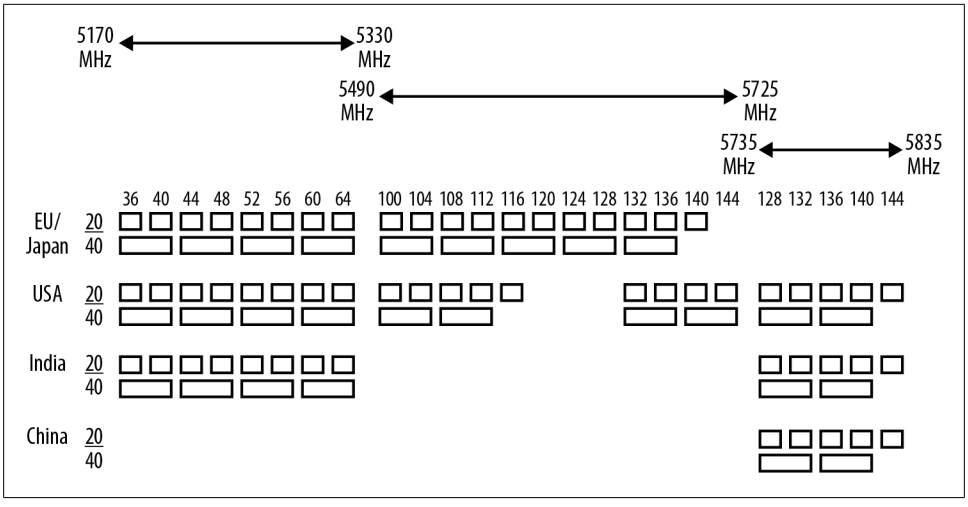


Figure 3-3. Available channel map (5 GHz)

Transmission: Modulation and Guard Interval

When a frame is transmitted, there are two separate parameters that 802.11n stations must agree on. First, the transmitter must select the modulation rate and coding that describes how data will be turned into radio waves. Second, the transmitter must select a guard interval; 802.11n optionally allows a shorter guard interval than previous standards.

Modulation and Coding Set (MCS)

In 802.11n, the Modulation and Coding Set (MCS) number is a value that describes the number of spatial streams, modulation (BPSK, QPSK, 16-QAM, or 64-QAM), and error-correcting code used for a transmission. 802.11n supports both *equal modulation*, in which all spatial streams are transmitted in the same manner, and *unequal modulation*, in which the spatial streams may be modulated differently. 802.11n defines 77 different combinations of modulation and coding. To date, most products shipped support only equal modulation modes, which are the first 32 MCS values. Table 3-3 and Table 3-4 at the end of this chapter have a list of the common MCS values, along with the link rates.

Unequal modulation is useful when one spatial stream is significantly more impaired than others. In transmit beamforming, the mathematical operations that are used to separate out the spatial streams may result in the streams having significantly different signal-to-noise ratios, which requires that some spatial streams be transmitted using a more conservative modulation. Because transmit beamforming has not been widely implemented, unequal modulation has also not seen wide implementation.

Forward Error Correcting Codes in 802.11n

802.11n specifies the use of two forward-error correction (FEC) codes to provide protection of frames as they are transmitted. Forward-error correction codes take the data to be transmitted and encode it with redundant bits to enable correction of errors at the receiver. Such a code decreases the efficiency of the channel by requiring transmission of extra bits, but it enables many errors to be transparently recovered at the receiver. If the efficiency loss from the redundant code bits is less than the efficiency loss from retransmissions, the error-correcting code improves the efficiency of the channel. One key parameter for forward-error correcting codes is the *code rate*, which describes the number of payload bits as a fraction of the total number of bits. For example, a code with rate R=1/2 transmits one payload bit for every two bits on the channel; in other words, out of every two bits, one is a redundant code bit added to detect and correct errors.

802.11n continues with the use of a convolutional code as first used in the OFDM PHY, and adds optional support for a new Low-Density Parity Check (LDPC). In almost all respects, the convolutional code used in 802.11n is identical to the convolutional code used in 802.11a/g. To increase the effective speed of the PHY, 802.11n adds one additional code rate: R=5/6. As with the R=2/3 coder specified in previous 802.11 amendments, the R=5/6 convolutional code is achieved by puncturing the output of the R=1/2 encoder to reduce hardware implementation complexity.[4] When using the convolutional code, a single encoder is used until the PHY rate exceeds 300 Mbps.

LDPC coding is a higher-performance error correction code than the convolutional codes previously used by 802.11. LDPC works by splitting up the data into blocks and adding redundant bits to each block in order to recover errors in transmissions. In 802.11n, LDPC uses the same code rates as the convolutional codes; if the LDPC code rate is R=2/3, then one out of every three bits is a redundant bit added by the encoding process. Compared to convolutional codes, LDPC implementation requires more complex (and therefore more power-hungry) circuitry. Support for LDPC is optional, but beginning to be more widely supported. LDPC operation is transparent to end users because the MAC transparently exchanges information about LDPC capabilities so it is used only when supported by both parties to a transmission.

4. Puncturing a convolutional code to achieve a higher rate code is described in Figure 13-10 of my earlier book, 802.11 Wireless Networks: The Definitive Guide (O'Reilly). That figure illustrates how to puncture the output of a R=1/2 code to achieve R=2/3. To achieve R=5/6 from the output of an R=1/2 encoder, more of the data is discarded during the puncturing process.

Guard Interval

Although it is not a rule in the legal sense, a good rule of thumb used by OFDM system designers is that the guard interval (GI) should be four times the highest multipath delay spread.[5] When 802.11a was being designed, designers used a conservative value of 200 ns for the delay spread, and chose to make the guard interval 800 ns. Implementation experience has shown that most indoor environments rarely have delay spreads of even 100 ns, and often it is closer to 50-75 ns. To wring an extra bit of performance out of the radio link, 802.11n includes an option for the short guard interval, which cuts the guard interval down to 400 ns. Whether you can successfully use the short guard interval depends on the multipath spread, which is reported by advanced RF analysis tools. In general, multipath interference is worse when there are significant reflections due to metal. Most 802.11n networks are able to use the short guard interval without issue.

All other OFDM parameters remain the same. Overall speed is increased because the part of the symbol time devoted to data transmission is still 3.2 μs, but each symbol is shorter. The total symbol length shrinks from 4.0 μs with the long guard interval to 3.6 μs (3.2 μs + 0.4 μs), shrinking OFDM overhead by 10% (as illustrated in Figure 3-4).

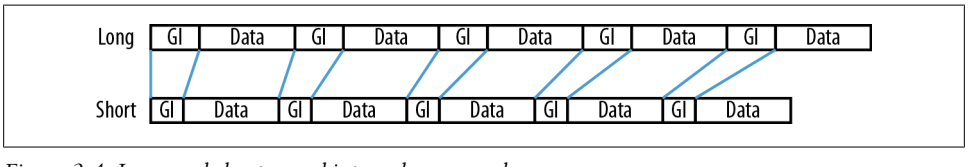


Figure 3-4. Long and short guard interval compared

PLCP Framing

As with previous 802.11 PHYs, the 802.11n specification defines a physical-layer frame using the Physical Layer Convergence Protocol (PLCP). The 802.11n PLCP supports three modes:

- *Non-HT mode*. All 802.11n devices are required to support a "non-11n" mode, which requires that they interoperate with 802.11a/b/g devices. No 802.11n features are available in this mode. The format of the non-HT mode PLCP frames is exactly the same as 802.11a or 802.11g.[6] Some documentation refers to this as "legacy mode" because it is operating according to the exact same rules as older standards.

5. The guard interval is a common component of OFDM system design. For more background on the rationale behind the guard interval, see Chapter 13 of *802.11 Wireless Networks: The Definitive Guide*.

6. See, for example, Figure 13-14 in *802.11 Wireless Networks: The Definitive Guide*.

- *HT mixed mode (HT-MM)*. All 802.11n devices are also required to support a mixed mode, in which the PLCP header is compatible with 802.11a/g PLCP headers, though of course the high-speed 802.11n body cannot be decoded by 802.11a/g devices.
- *HT-Greenfield (HT-GF) mode*. The greenfield PLCP header is slightly shorter than the mixed mode header, and can be used in an area where only 802.11n devices are deployed. In one of the proposals that led to 802.11n, a similar function was originally called *pure mode*.

All devices must support both the non-HT mode and the mixed mode. If an AP is configured to also support greenfield mode, it will be advertised in the network's Beacon frame. An 802.11n device can choose to use any PLCP frame format supported by the receiver. Commonly, non-HT mode is used for short frames without a payload, such as CTS and ACK frames, because it has lower overhead than either of the HT modes. Many 802.11n devices, especially those designed for enterprise use, support only mixed mode and not greenfield mode.

Each of the 802.11n PLCP operating modes is illustrated in Figure 3-5. Non-HT mode is equivalent to the 802.11a/g mode, and HT Mixed Mode is by far the most common method. HT-Greenfield mode is shown for reference. Greenfield mode is not commonly implemented, and is best used only in cases when there are no overlapping networks.

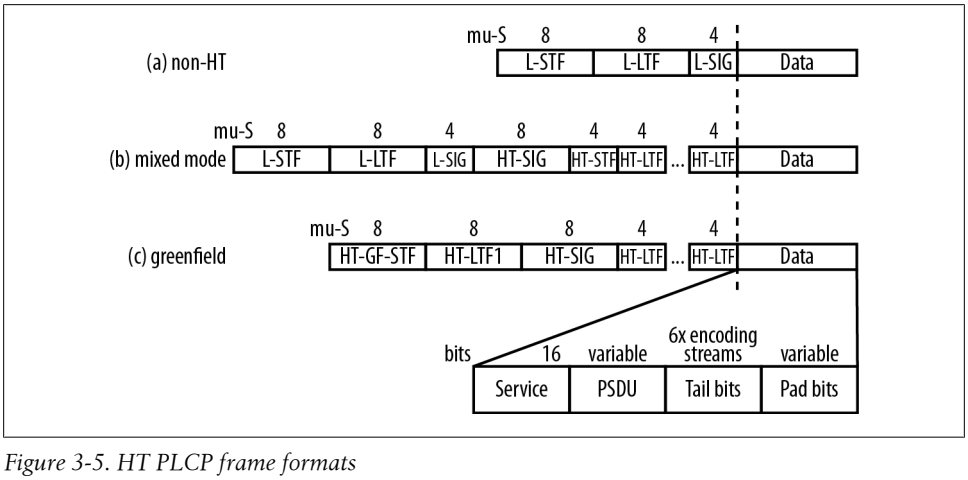


Figure 3-5. HT PLCP frame formats

HT Mixed Mode PLCP Format

The fields in the HT Mixed Mode PLCP frame are:

Non-HT Short Training Field (L-STF) and Non-HT Long Training Field (L-LTF)
: These fields are identical to the fields used in 802.11a/g, and are a sequence of 12 OFDM symbols that are used to assist the receiver in identifying that an 802.11n frame is about to start, synchronizing timers, and antenna selection. These fields can be decoded by any 802.11 device that is capable of OFDM operation.

Non-HT Signal (L-SIG)
: The Signal field is used in 802.11a/g to describe the data rate and length in bytes of the frame. 802.11n devices will set the data rate to 6 Mbps and derive a spoofed length in bytes so that when 802.11a/g stations compute a transmission duration from the "length" at the 6 Mbps rate, it matches the transmission duration required for the HT frame. For complex frame exchanges, the duration derived from the L-SIG field may be used as a form of protection to enable 802.11a/g devices to share the medium with 802.11n devices, as described in Chapter 5.

HT Signal (HT-SIG)
: The HT Signal field is the 802.11n analog of the Signal field used in previous OFDM PHYs. Like the Legacy Signal field, the HT Signal field describes the data rate and length. However, this field can only be understood by 802.11n devices. To get the "true" rate and length information, for example, to display in an analysis tool, it is necessary to decode and interpret the HT Signal field. It also carries additional information about the transmission, including the type of error-correction code, the guard interval, and aggregation.

HT Short Training Field (HT-STF)
: The HT STF serves the same purpose as the non-HT STF. Just as the non-HT training fields help a receiver tune in the signal, the HT-STF assists the receiver to detect a repeating pattern and setting receiver gain.

HT Long Training Field (HT-LTF)
: There are two types of HT-LTF: a Data LTF and an optional Extension LTF used in channel sounding frames as part of transmit beamforming. As with the non-HT LTF, the HT LTF helps tune the MIMO system so that spatial streams can be decoded. In most cases, the extension LTFs are not present. One HT LTF is present for one spatial stream, two HT LTFs are present for two spatial streams, and four HT LTFs are present for three or four spatial streams. Extension LTFs are used to excite spatial streams for transmit beamforming, which is why they are not commonly used.

Greenfield mode removes backward compatibility support by replacing the legacy fields with HT-specific versions of the same fields.

Greenfield Mode

This book doesn't go into the details of greenfield mode because it is not widely used. Greenfield mode yields a small performance increase because the PLCP headers are 8 μs shorter. As with any other protocol option, choosing to implement it is subject to cost/benefit analysis. Significant components of the PLCP are implemented in hardware (or low-level software supplied by the chip vendor), making greenfield mode hard to develop for most wireless LAN product developers.

In addition to the development costs, greenfield mode may cause significant problems when a network is built using it. By design, it is not compatible with existing 802.11a/b/g networks, which makes it potentially dangerous to deploy a greenfield-mode network where there are preexisting wireless pre-11n LAN devices. Wireless LANs of any size have typically existed for a long enough time to accumulate a number of "legacy" devices, for some value of "legacy" (whether it is only 802.11a/g or includes older 802.11b devices). For backward compatibility alone, greenfield mode doesn't seem to make much sense. I have yet to come across a network that requires the marginally higher speed that greenfield mode provides but can nevertheless be restricted only to 802.11n devices. Furthermore, a greenfield-mode network has the potential to destabilize overlapping networks that provide service to 802.11a/b/g devices.

One minor concern for greenfield networks is that because the low-level support is so rare, an AP operating in greenfield mode might be "invisible" to network analysis tools such as sniffers or other wireless security products. This risk seems minimal at this point, however, given the lack of greenfield mode support in wireless LAN devices currently on the market, and the requirement in the standard that Beacons be transmitted using 802.11a/g methods.

HT Signal Field

The HT Signal field, shown in Figure 3-6, serves the same purpose as the Signal field does in 802.11a/g. It is transmitted at a standard rate (MCS 0) and can be readily decoded by a receiver. In an HT-Mixed Mode format transmission, the presence of the HT Signal field is used to determine whether the payload is transmitted using the older 802.11a/g rates or the MIMO rates in 802.11. To assist receivers in determining whether the HT Signal field or an 802.11a/g Data field follows the training symbols, the constellation for the HT Signal field is rotated. 802.11a/g receivers are unable to decode the payload, but 802.11n receivers decode and use the data in the HT signal field aids in decoding the PSDU portion of the packet by describing how it is encoded and modulated.

Modulation and Coding Scheme (7 bits)
: This field describes the modulation and coding scheme, and implicitly gives the receiver the number of spatial streams. Although the MCS value can take on any one of 76 values, the 32 described in "802.11n Speed" on page 29 are by far the most commonly used.

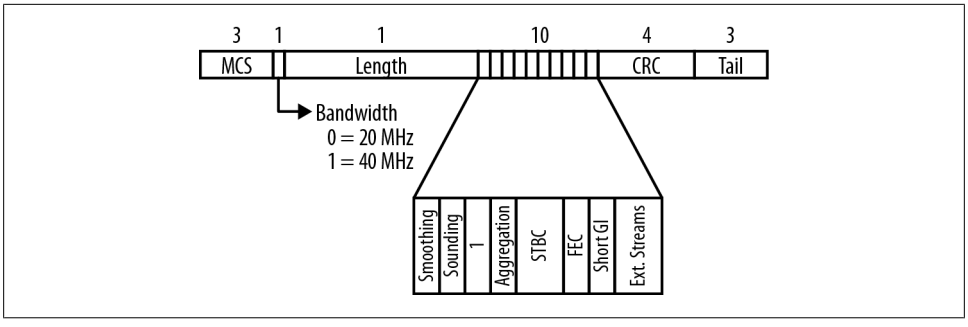


Figure 3-6. HT Signal field

Channel bandwidth (1 bit)
: This bit is set to 0 for 20 MHz channels and 1 for 40 MHz channels.

HT Length (16 bits)
: This field describes the length of the PSDU in bytes, and can range up to 65,536.

Not Sounding (1 bit)
: Sounding frames are used in beamforming, a process that is described in more detail in Chapter 4. When this bit is set to 0, the frame is not a sounding frame, which is the typical operation. When set to 1, the frame is used for sounding operations.

Aggregation (1 bit)
: This bit is set when the payload is an aggregate frame that contains several sub-frames. Aggregation is further described in Chapter 5.

STBC (2 bits)
: Space-Time Block Coding allows multiple radio chains to be used to increase the sensitivity to a single data stream; it is further described in Chapter 4.

Forward Error Correction coding (1 bit)
: When using a convolutional code, this bit is set to 0; to use LDPC, the bit is set to 1.

Short guard interval (1 bit)
: When the short guard interval is in use, this bit is set to 1. When set to 0, it indicates the long guard interval is in use.

CRC (8 bits)
: A CRC allows the receiver to detect corruption of the HT Signal field.

Tail bits (6 bits)
: The HT Signal field is protected by a convolutional code, and requires a trailing six zeroes to "ramp down" the convolutional code.

Data Field

As shown in Figure 3-5, the Data field of the PLCP consists of four components:

SERVICE field (16 bits)
: This field has 16 zeroes to initialize the data scrambler. To avoid long sequences of the same bit, 802.11n uses a scrambler to improve the distribution of ones and zeroes in the payload.

PSDU (variable)
: This field contains the PLCP Service Data Unit, which is a frame from higher protocol layers being prepared for transmission out the radio interface. To improve efficiency, the 802.11n MAC layer may provide an aggregate frame, which is described in Chapter 5.

Tail bits (6 bits for each encoding stream)
: Convolutional coders require tail bits to cleanly terminate the convolutional code. Each encoder requires six bits. Generally speaking, this field will be six bits long, though at speeds greater than 300 Mbps, it is twelve bits.

Pad bits (variable)
: Pad bits are used to ensure that the Data field of the PLCP is an even number of symbols.

Transmission and Reception Process

The block diagram for an 802.11n interface is shown in Figure 3-7. When a frame is ready for transmission, an 802.11n interface runs the following procedure:[7]

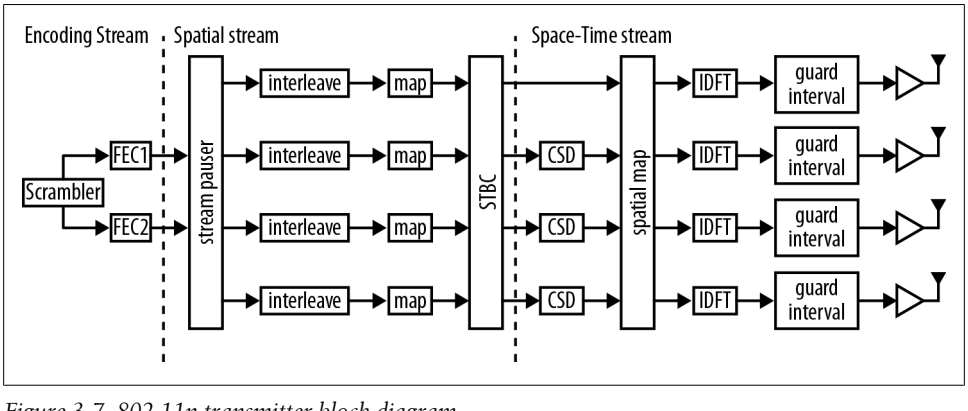


Figure 3-7. 802.11n transmitter block diagram

7. In many ways, the transmission procedure for 802.11n resembles that for the OFDM physical layers used in 802.11a and 802.11g. For more information, see Chapter 13 in *802.11 Wireless Networks: The Definitive Guide*.

1. **Scramble and Forward Error Correction (FEC) encode**: The scrambler is a process that reduces the probability of long strings of identical bits before feeding the resulting bits to the FEC encoder. The FEC is either a convolutional coder or an LDPC coder; Figure 3-7 shows two convolutional coders to illustrate that two encoders must be used when the data rate exceeds 300 Mbps.
2. **Stream parsing:** The stream parser takes the output of the FEC encoder and divides up the encoded bits between each spatial stream. For example, if there are two spatial streams, the stream parser will take the encoded bits and assign each encoded bit to one of the spatial streams. At this point, the bits flowing from the stream parser to the interleaver are a *spatial stream*. Output from the stream parser is sent to the interleaver, which is the first component in the radio chain.
3. **Interleave and mapping:** As in 802.11a/g, the interleaver changes the order of the bits. Convolutional codes correct errors best when errors are confined to short bursts. The interleaver spreads adjacent bits across OFDM subcarriers to assist the convolutional code in isolating and correcting errors. When bits leave the interleaver, they are mapped on to constellation points for transmission.[8] 802.11n does not alter the constellations used by 802.11a/g.
4. **Space-Time Block Coding (STBC):** This optional step, described in more detail in Chapter 4, can be used to transmit one spatial stream across two antennas for extra redundancy. The space-time block coder takes the output of the interleaver/constellation mapper and spreads it across multiple radio chains, transforming the spatial streams into *space-time streams*.
5. **Spatial mapping:** Space-time streams are mapped onto the transmit chains by the spatial mapper. This will often be a *direct mapping*, in which a spatial stream becomes a space-time stream and then is mapped on to a single transmit chain. Direct mapping is a simple, straightforward process.

   As an alternative to direct mapping, the spatial mapper may instead perform a *spatial expansion*, in which all of the space-time streams from the STBC are spread across all the transmit chains. Spatial expansion is similar to *beamforming*, in which the space-time streams are "shaped" to direct energy in a particular direction. Beamforming is discussed in more detail in Chapter 4.
6. **Inverse Fourier transform and cyclic shift:** Taken together, these two components convert the frequency-domain signal into a time-domain signal suitable for transmission. The inverse Fourier transform takes the frequency-domain data from OFDM and converts it to time-domain data. To preserve the neutral nature of the MIMO transmission, a *cyclic shift* adds a small phase delay to each transmit chain. The cyclic shift may be added per transmit chain (as shown in Figure 2-3), or it may be added after the space-time streams are created (as shown in Figure 3-7).

8. Constellation mapping and modulation is a complex topic. For a bit more background with pictures, see my video about QAM and error vector magnitude.

7. **Guard insertion and windowing** improve the signal quality at the receiver, just as in 802.11a/g.
8. Finally, the **radio section amplifies the signal** for transmission out an antenna. At this stage, the final data signal is available and can be placed on to the carrier. A high power amplifier (HPA) increases the power so the signal can travel as far as needed, within regulatory limits.

To receive a frame, the transceiver reverses the steps. A weak signal from the antenna is boosted by a Low-Noise Amplifier (LNA) so that the symbols can be recovered with a Fourier transform. After separating the spatial streams and de-interleaving, errors in the bit stream are corrected by the FEC and the resulting frame is passed to the MAC.

802.11n Speed

Answering the question of "How fast does 802.11n go?" is something that defies a simple explanation. Unlike previous PHYs that had only a handful of options (none of which affected the link rate used) 802.11n has a number of options that together determine the data rate. Almost 2% of the page count of 802.11n is devoted to tables that describe the speed for various options.

This section discusses the speed in 802.11n from the standpoint of a network administrator buying product that is readily available. Two questions will often come to mind. First, how does it compare to pre-802.11n equipment, such as the laptops with 802.11a/g? Second, what types of link rates can I expect to see when I use 802.11n devices?

For simplicity, this section will quote link rates using the long guard interval. It is the default setting on most equipment, and is generally safer to use. The performance boost for switching to the short guard interval about 11%, which is substantial only in environments where the radio channel is saturated. When faced with a bewildering array of choices, most people tend to zero in on a couple of numbers that stick out, and I've found that the most commonly cited numbers are the following:

- 150 Mbps (two spatial streams using 20 MHz channels with a short guard interval at MCS15), because it is the first data rate in the 802.11 world to top 100 Mbps using the traditional 20 MHz channel size
- 300 Mbps (two spatial streams using 40 MHz channels with a short guard interval at MCS15), because it is the top data rate supported by the first products that saw wide release into the market
- 450 Mbps (three spatial streams using 40 MHz channels with a short guard interval at MCS15), because it is the top data supported by three-stream 802.11n products. This data rate was widely available in products that began shipping in 2011.

- 600 Mbps (four spatial streams using 40 MHz channels with a short guard interval at MCS 31), because it is the top data rate described by the 802.11n standard, even if no products based on that data rate have yet been sold.

Comparison 1: 802.11a/g versus 1x1 11n

Many client devices have implemented 802.11n for cost reasons. Once the technology basis for the industry moves, it is hard to get older chips. Even though many battery-powered devices have low overall throughput requirements and therefore, manufacturers would sometimes prefer to keep using 802.11b, chip vendors continuously refresh their product lines and pull their customers towards the latest technology. Even devices made with full attention to reducing battery draw are now made with 802.11n chips, though of course it is a 1x1 design to minimize the number of power-hungry radio chain components. Virtually every Wi-Fi enabled phone uses single-stream 802.11n, as do many tablet computing devices.

In the case of single-stream 802.11n, there is only a small performance gain in link rate. 802.11n's channel has 8% more data subcarriers in a channel, so it can support link rates that are 8% higher, as shown in Table 3-2. Even though the link rate is not higher on paper, it often makes sense to use 802.11n for these devices because the antenna array at the AP can often boost performance through techniques like maximal ratio combining.

Table 3-2. 802.11a/g speed versus 802.11n equivalent

Modulation/coding	802.11a/g speed	802.11n (1x1, 20 MHz, Long GI) speed
BPSK, R=1/2	6	6.5
QPSK, R=1/2	12	13.0
QPSK, R=3/4	18	19.5
16-QAM, R=1/2	24	26.0
16-QAM, R=3/4	36	39.0
64-QAM, R=1/2	48	52.0
64-QAM, R=3/4	54	58.5
64-QAM, R=5/6	not used in 802.11a/g	65.0

Comparison 2: 20 MHz versus 40 MHz channels

The second comparison that is often made is the difference in link rate between the familiar 20 MHz channel width and the new wider 40 MHz channels in 802.11n. Table 3-3 and Table 3-4 show the link rate that can be achieved for various combinations of modulation, coding, and spatial streams. The table shows data rates using the long guard interval; for speeds using the optional short guard interval, add 11%. In 802.11n, each combination is given an MCS number. These two tables only show the

equal modulation MCS numbers; there are many more MCS numbers that make use of unequal modulation, but they are not widely supported.

The first wave of access points and laptops to hit the market supported only two spatial streams, and the common numbers used to describe the performance of these systems was "100 Mbps" (which is the net throughput of a set of frames aggregated together at 135 Mbps, plus a block acknowledgment) or "300 Mbps" (the link rate achieved with 40 MHz channels and a short guard interval).

In 2011, the second major wave of products was brought to market, supporting three spatial streams and speeds of up to 450 Mbps. These products now power all mainstream high-performance access points, and many new client devices now use three-stream 802.11n chipsets. Although four spatial streams have been standardized, it seems unlikely that four-stream products will ever be mainstream because most chip vendors have focused on building four-stream products using 802.11n's even higher-speed successor.

Table 3-3. 20 MHz speeds (long guard interval)

Modulation and coding	1 SS	2 SS	3 SS	4 SS
BPSK, R=1/2	6.5 (MCS 0)	13.0 (MCS 8)	19.5 (MCS 16)	26.0 (MCS 24)
QPSK, R=1/2	13.0 (MCS 1)	26.0 (MCS 9)	39.0 (MCS 17)	52.0 (MCS 25)
QPSK, R=3/4	19.5 (MCS 2)	39.0 (MCS 10)	58.5 (MCS 18)	78.0 (MCS 26)
16-QAM, R=1/2	26.0 (MCS 3)	52.0 (MCS 11)	78.0 (MCS 19)	104.0 (MCS 27)
16-QAM, R=3/4	39.0 (MCS 4)	78.0 (MCS 12)	117.0 (MCS 20)	156.0 (MCS 28)
64-QAM, R=1/2	52.0 (MCS 5)	104.0 (MCS 13)	156.0 (MCS 21)	208.0 (MCS 29)
64-QAM, R=3/4	58.5 (MCS 6)	117.0 (MCS 14)	175.5 (MCS 22)	234.0 (MCS 30)
64-QAM, R=5/6	65.0 (MCS 7)	135.0 (MCS 15)	195.0 (MCS 23)	260.0 (MCS 31)

Table 3-4. 40 MHz speeds (long guard interval)

Modulation and coding	1 SS	2 SS	3 SS	4 SS
BPSK, R=1/2	13.5 (MCS 0)	27.0 (MCS 8)	40.5 (MCS 16)	54.0 (MCS 24)
QPSK, R=1/2	27.0 (MCS 1)	54.0 (MCS 9)	81.0 (MCS 17)	108.0 (MCS 25)
QPSK, R=3/4	40.5 (MCS 2)	81.0 (MCS 10)	121.5 (MCS 18)	162.0 (MCS 26)
16-QAM, R=1/2	54.0 (MCS 3)	108.0 (MCS 11)	162.0 (MCS 19)	216.0 (MCS 27)
16-QAM, R=3/4	81.0 (MCS 4)	162.0 (MCS 12)	243.0 (MCS 20)	324.0 (MCS 28)
64-QAM, R=1/2	108.0 (MCS 5)	216.0 (MCS 13)	324.0 (MCS 21)	432.0 (MCS 29)
64-QAM, R=3/4	121.5 (MCS 6)	243.0 (MCS 14)	364.5 (MCS 22)	486.0 (MCS 30)
64-QAM, R=5/6	135.0 (MCS 7)	270.0 (MCS 15)	405.0 (MCS 23)	540.0 (MCS 31)

Mandatory PHY Features

802.11n is a complex specification, with many protocol features. To help readers keep track of features that are mandatory and optional, Table 3-5 classifies the protocol features into two categories. Generally speaking, the Wi-Fi Alliance test plan validates functionality which is mandatory in the specification and has optional testing only for the most widely supported features.

Table 3-5. Feature classification of PHY features

Feature	Mandatory/optional?	Comments
HT-Mixed Mode	Mandatory	
HT-Greenfield mode	Optional	Validated in Wi-Fi Alliance test plan; supported by approximately 1/4 of certified devices.
800 ns guard interval	Mandatory	
400 ns (short) guard interval	Optional	Validated in Wi-Fi Alliance test plan and displayed on interoperability certificate; supported by approximately 3/4 of devices
20 MHz channel operation	Mandatory	
40 MHz channel operation	Optional	Validated by Wi-Fi Alliance test plan and displayed on interoperability certificate. Only about 1/4 of devices support 2.4 GHz operation with coexistence mechanisms; about 3/4 of enterprise APs support 40 MHz channels in 5 GHz
Single-stream operation (MCS 0 through MCS 7)	Mandatory	Wi-Fi Alliance test plan only requires 1-stream operation by client devices.
2-stream operation (MCS 8 through MCS 15)	Mandatory for APs	2-stream operation is required for all APs. The number of spatial streams tested is displayed on the interoperability certificate.
3- and 4-stream operation (MCS 16 through MCS 31)	Optional	3-stream enterprise APs are mainstream and are tested by the Wi-Fi Alliance; very few 4-stream APs exist, and 4-stream operation is untested
HT Duplicate Mode (MCS 32)	Optional	MCS 32, also called HT Duplicate mode, uses a single 40 MHz channel with one spatial stream at a data rate of 6 Mbps. It is very conservatively coded for high reliability. Because the speed is quite slow, however, it is not widely used in large-scale networks.
MCS 33 through MCS 76 (unequal modulation)	Optional	
Transmit Beamforming	Optional	Not part of Wi-Fi Alliance test plan. See Chapter 4
Low-density Parity Check	Optional	Not part of Wi-Fi Alliance test plan. See Chapter 4
Space-Time Block Coding	Optional	Validated by Wi-Fi Alliance test plan and displayed on the interoperability certificate, but supported by less than 1/5 of certified 11n devices. See Chapter 4.

CHAPTER 4

Advanced PHY Features for Performance

Half the transmit power I pump into my antenna is wasted. The trouble is that I don't know which half.

—John Wanamaker (had he been a wireless network administrator instead of an early advertiser)

In its most basic conception, MIMO is technology that provides a multi-lane highway for 802.11n transmissions. When a transmitter can send multiple streams across a link, throughput improves. This description of MIMO is fairly passive, and does not provide a full picture of everything that a MIMO system can do. With an antenna array, it is possible to create far more sophisticated transmissions than just sending multiple streams. With advanced radio chips, it is possible to send transmissions from the MIMO antenna array in a particular direction in a process called *beamforming*. Individual receiving antennas within MIMO arrays can also cross-check each other to improve signal processing at the receiver. Both applications can increase the signal to noise ratio, which improves speed. In addition to developing the underlying technology for antenna arrays, 802.11n includes protocol features that can improve the signal processing gain. MIMO can also be used to spread a single spatial stream across multiple transmitters for extra signal-processing gain at the receiver, and hence, longer range in a process called *Space-Time Block Coding* (STBC). 802.11n also provides another option to increase signal-processing gain, the *Low-Density Parity Check* (LDPC) code. As discussed in "Relationship Between Spatial Streams and Radio Chains" on page 15, many products also implement some form of ratio combining to increase signal gain, though ratio combining is not a protocol feature.

Beamforming

Prior to 802.11n, most access points were equipped by default with omnidirectional antennas. "Omnis" radiate equally in all directions, and their coverage is often depicted

as a circle with the AP at the center. One of the costs of of an omni is that the signal destined to a receiver is sprayed out evenly in all directions, even though the receiver is only in one of those directions. The best way I have found to explain it is to recall John Wanamaker's famous statement that half of all advertising spending is wasted, but the problem is that nobody knows which half goes to waste. *Beamforming* is 802.11's answer to this conundrum. It enables an antenna array to focus the energy in the direction of a client device. By using the same power but focusing it towards the receiver, it is possible to increase the signal to noise ratio at the receiver. Higher signal to noise ratio enables use of more aggressive coding, and therefore, higher speed. Figure 4-1 shows the basic idea of beamforming. A transmitter, also called the *beamformer*, "focuses" or "steers" the energy so that the transmission tends to go more in one direction, towards the receiver, also called a *beamformee*. (The terms beamformer and beamformee are useful in the context of explicit channel measurement protocols that require frame exchanges.) The figure illustrates one of the classic beamforming trade-offs. To form a beam in one direction, the result is often a figure-8 coverage pattern, as shown in the figure.

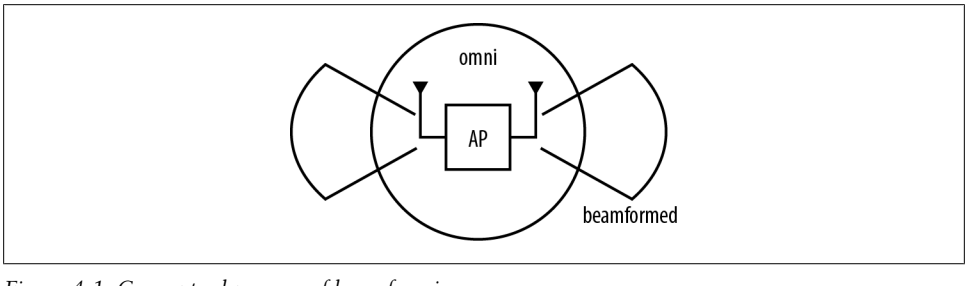


Figure 4-1. Conceptual process of beamforming

In essence, beamforming is an electrically steerable antenna. When transmitting to a given receiver, the antenna is "pointed" in the direction of the receiver, but the direction that the beam is focused can be controlled by subtly altering the phase shifts into the antenna system rather than mechanically moving the antenna. Phase shifts in the transmission sent by each antenna are used to focus energy along lines of constructive interference between the antenna elements.[1]

Naturally, phase shifting the transmissions to individual radio chains requires fairly advanced signal processing within the 802.11 chip. When the transmission is biased towards transmission in one direction, it tends to have a relatively low sensitivity in the opposite direction. For example, in Figure 4-1, the beam is being focused at the receiver

1. In most beamforming systems, the antenna array is still composed of omnidirectional antennas. Complex phased-array antenna systems with many elements can be used as well, but such antennas have significantly higher complexity, and hence, cost. Expensive antenna arrays make a great deal of sense in mobile telephony; when used in Wi-Fi APs, the antenna array tends to lead to compromises in other areas of the system design.

on the right. Interference sources from other directions will tend to be rejected by the antenna because the antenna system is pushing the gain towards the right.

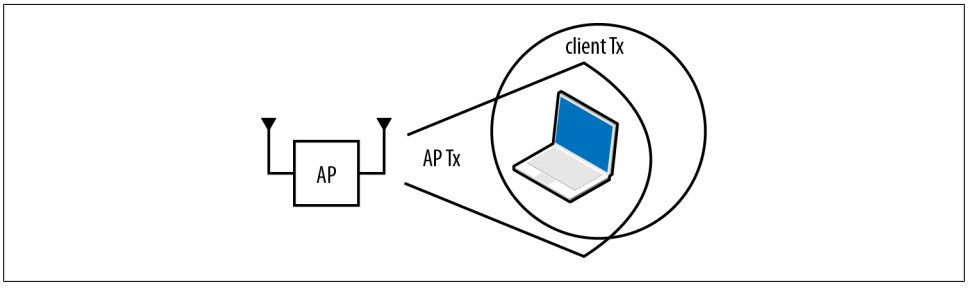
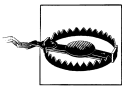


Figure 4-2. Beamforming asymmetry

By concentrating radio energy in a particular direction, the "downstream" link from AP to client gets additional power. With more power, a stronger signal is better able to cut through interference. A gain of just 1-2 dB can be the difference between data rates in an 802.11 device.[2] The disadvantage of beamforming is that its benefits are concentrated at the receiver. Figure 4-2 illustrates the point by showing a typical network with an AP and a laptop attached to it. When the AP transmits to the laptop, it is able to transmit a concentrated beam, potentially with a gain over an omindirectional antenna of 3-5 dB. However, when the laptop sends the 802.11 ACK (or a TCP segment containing a TCP-layer acknowledgment), the AP cannot set up the antenna with sensitivity "pointing at" the client because it may be serving multiple attached client devices. 802.11 does not exercise tight control over transmission timing, so beamforming access points generally receive without any steering and are restricted to using antenna arrays as omnidirectional receivers. As a result, beamforming generally leads to asymmetric links, with the potential for hidden node interference. To eliminate link asymmetry or reduce its negative effects, focus on the receive sensitivity of the AP to offset the lower receive antenna gain.



Asymmetric links are not compatible with all applications. If the transmissions mainly flow downstream from an AP to receivers, it can increase speed. Balanced links that expect upstream data (such as voice or videoconferencing) may not see much gain from beamforming, especially if RTS/CTS exchanges are required to avoid hidden nodes.

Types of Beamforming

Beamforming refers to one of several related technologies that can be used to increase SNR at the receiver. This book uses the term *beamforming* to refer only to technologies

2. Although it doesn't really sound like it, 1 dB is a huge amount in the radio world. Don't believe me? Go ask your boss for a "1 dB raise."

that work by actively steering transmissions from an antenna array by using multiple antenna array elements. If your transmitter isn't using multiple components in an array to steer the beam, it's not beamforming. For example, an array of static directional antennas doesn't do beamforming. It may increase range by using high-gain directional antennas, but it is using the directional antennas in a static way. To be a true beamforming device, either the radio chip or the antenna must be doing something to change the transmissions out of an antenna array, and do so on a frame-by-frame basis.

Whether implemented on the radio chip or in the antenna array, beamforming works by taking the radio signals to be transmitted and applying a mathematical transformation to them that changes the way they are transmitted. Mathematically, the transformation is typically represented as a matrix that takes an incoming signal pre-beamforming and maps it to a number of outputs.[3] With on-chip beamforming, the matrix has dimensions of the number of space-time streams that are inputs and the number of transmit chains. In a 3-stream system with 3 outputs, the matrix will be a square that maps some combination of the inputs to each of the outputs. When beamforming is implemented in the antenna array, the size of the matrix is also the number of inputs and outputs. An antenna array may still take the 3-stream transmission as input, but instead of mapping on to three transmit chains, the antenna may map those three inputs on to ten or more output elements.

The matrix that describes the transformation from space-time streams to transmitted energy is called the *steering matrix*. In the block diagram of an 802.11n interface in Figure 3-7, the spatial mapper uses the steering matrix to alter the transmitted data to have longer reach. A similar process would be carried out within software running on the antenna system in an antenna-based beamforming system.

Broadly speaking, beamforming comes in two major flavors, and both are compared in Table 4-1. *Explicit beamforming* is easy to understand. Before transmitting, a device actively measures the channel, and uses the measurement to compute the steering matrix directly. Active channel measurement is accomplished by transmitting a *sounding* frame to the beamformee, which replies with a frame that indicates how the sounding frame was received. By comparing the known contents of the sounding frame to a representation of its contents at the receiver, the beamformer can compute the steering matrix. The downside to explicit beamforming is that it requires active support on both ends of the radio link. To receive beamformed transmissions, a device must be able to send channel measurements back to the beamformer. Only 802.11n has defined the channel measurement sequences, which means that older 802.11a/g devices will be unable to receive explicit beamformed frames.

Almost all 802.11n access points that support beamforming do so based on *implicit beamforming*. As the name implies, implicit beamforming devices do not use any frame

3. Matrix operations are used because the effects of multipath interference may be different for each spatial stream. Beamforming is used independently for each stream because each stream may have different frequency-specific fading effects.

exchanges that are dedicated to beamforming. Instead, devices estimate the beamforming matrix from received frames, or by inference from frames that are lost. Well-known frames such as ACKs or the data transmitted on pilot channels can be used to estimate the steering matrix. Implicit beamforming is based on less comprehensive measurements, and therefore, does not provide quite as high a level of performance. On the other hand, the implicit variety may be implemented by only one side of a link and may be used when that link supports a pre-11n PHY. For these reasons, implicit beamforming is significantly more common, and serves as a bridge between traditional omnidirectional transmission and the future of standards-based beamforming in 802.11 radio chips.

Table 4-1. Beamforming technology comparison

Attribute	Implicit	Explicit
802.11 PHY support	802.11a, b, g, and n	802.11n only
Client requirement	None	Send channel measurements to AP
Link adaptation	Open loop	Closed loop
Feedback source	Client uplink frames	Client channel measurements
Performance gain	Moderate gain, generally increasing with the number of antenna elements	Higher gain
Implementation location	Software/firmware built into antenna system or layered on top of radio chip	Special features within 802.11n radio interface
Number of implementations	Common	Rare

Explicit Beamforming Exchanges in 802.11n

Beamforming is not yet a common feature on 802.11n APs. As this book was written, very few wireless LAN chips supported beamforming, though nearly all of the 802.11 chips in development tout support for this feature. 802.11n defined several methods for communicating explicit channel feedback. One method uses frames that are specifically identified as "sounding" frames, signaled as such by the sounding bit in the PLCP header. Sounding can also be carried out by using *Null Data Packets* (NDPs), which are frames that have no data but are designed to enable detailed channel measurements. Explicit beamforming is not yet widely supported, in part because there are two methods and no product vendor wants to implement both types.

Space-Time Block Code (STBC)

Space-Time Block Coding (STBC) can be used when the number of radio chains exceeds the number of spatial streams. In effect, STBC transmits the same data stream twice across two spatial streams so that a receiver that misses a data block from one spatial stream has a second shot at decoding the data on the second spatial stream. In effect,

STBC takes the gains from MIMO and uses them nearly exclusively for increased range. A single encoding stream must take two radio chains for transmission, which means that a 2×2 MIMO device transmitting with STBC is effectively operating as a single-stream device. 802.11 interfaces include a step-down algorithm that selects slower and more reliable transmission rates; STBC can be used when the data channel is too poor to support one full stream per radio. In such environments, STBC is worth the 50% penalty on transmitted speed. When an STBC-enabled AP is serving mainly single-stream devices at long ranges, STBC is definitely worth considering.

Low-Density Parity Check (LDPC)

Until 802.11n, all OFDM-based PHYs used a convolutional code as the forward-error correcting (FEC) code. Conceptually, a convolutional code works by "smearing out" errors over time so that a single bit error can be corrected if enough neighboring bits are good. The low-density parity check (LDPC) code that is defined as an option in 802.11n works in a similar manner, but it also offers a coding gain when compared to convolutional codes. Depending on the channel model, simulations show that LDPC increases signal-to-noise ratio between 1.5 and 3 dB.[4]

The LDPC operates on code words that are 648 bits, 1296 bits, or 1944 bits long. When a frame is sent to the 802.11n interface for transmission, it is first divided into blocks based on the length of the code word (Table 4-2). As with the convolutional code, the coding rate defines the number of added bits used to detect and correct errors; no change is required to the tables in "Comparison 2: 20 MHz versus 40 MHz channels" on page 30.

Table 4-2. LDPC block size and rates

Code rate	Data bits per code word
R=1/2	324, 648, or 972
R=2/3	432, 864, or 1296
R=3/4	486, 972, or 1458
R=5/6	540, 1080, or 1620

Use of LDPC must be negotiated as part of the 802.11 association process. An 802.11n device only transmits LDPC-encoded frames to peers that have indicated support for LDPC in either Beacon frames or Association Request frames.

4. One of the early presentations on LDPC's coding gain is IEEE document 11-04/0071.

PART II
The MAC

Before 802.11n, the 802.11 MAC was generally viewed as being about 50% efficient—take the operating data rate of the system during transmission, and cut it in half to account for the various forms of protocol overhead needed to coordinate transmissions between stations. When data rates were relatively low, the cost of inefficiency was manageable. Losing half of a one-megabit network is only 500 kbps. As transmission speeds continued to increase, the loss became more and more significant. 802.11n developed several protocol features to reclaim some of the inefficiently used airtime. Focus on efficiency has paid off: under many conditions, 802.11n can have an efficiency of 70%.

CHAPTER 5
MAC Basics

We must, indeed, all hang together, or most assuredly
we shall all hang separately.

—Benjamin Franklin

802.11n made several modifications to the MAC layer to improve efficiency. Just as jumbo frames on Ethernet improve efficiency by increasing the "timeslice" of medium usage devoted to transferring data, larger frames in 802.11n do the same. However, unlike Ethernet, frames in 802.11 must be acknowledged. When several frames are clustered together in an aggregate container, sending single positive acknowledgments for each would give back many of the efficiency gains. Therefore, 802.11n devices make extensive use of block acknowledgments to improve the efficiency of the required positive acknowledgment process; 802.11n receivers can selectively acknowledge the individual constituent frames inside an aggregate.

Frame Changes

The 802.11 data frame is only slightly changed by 802.11n. Figure 5-1 shows the format of an 802.11 Data frame as modified by 802.11n. The major changes from the traditional 802.11n Data frame are the increase in size, the addition of the optional HT Control subfield, and the fact that the QoS Control field is utilized extensively in block acknowledgment. The payload of the MAC is increased about fourfold, which can be used to aggregate higher-layer frames together for efficiency.

Management frames signal that they are part of an 802.11n network by including the HT Capabilities information element, shown in Figure 5-2. When a station includes the HT IE in transmissions, it is declaring that it is an 802.11n device. The HT IE is included in Beacon frames so that stations can determine that a network supports 802.11n. An 802.11 station will insert the HT IE into Probe Request frames to seek out 802.11n networks and declare to an AP that it is capable of 802.11n operation. The HT IE is also included in Association and Reassociation exchanges so that an 802.11n device associating to an AP can exchange information on capabilities.

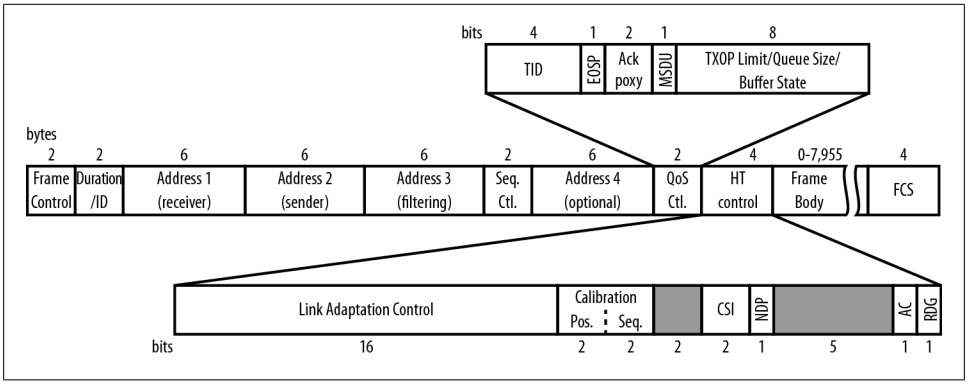


Figure 5-1. Revised format of MAC frame

HT Capabilities Info (2 bytes)
: The basic capability information is used to communicate the types of channels, coding, and coexistence support to reduce potential interference between neighboring networks using 20 and 40 MHz channels. One important bit in this field is the 40 MHz Intolerant bit, which is described in "40 MHz Intolerance for Channel Width Interoperability" on page 64.

A-MPDU Parameters (1 byte)
: This field declares parameters used with A-MPDU aggregation, described in "A-MPDU" on page 45.

Supported MCS Set (16 bytes)
: This large field includes a great deal of information on the data rates supported by the station including the IE. It allows a station to report the data rates it can support, plus any difference between the data rates it can receive and the data rates it can transmit (though this protocol feature is not commonly used).

HT Extended Capabilities (2 bytes)
: This field supports the description of extended capabilities such as Phased Coexistence Operation (PCO) and the Reverse Direction (RD) protocol. Neither of these features are widely implemented, and are not described in this book.

Transmit Beamforming Capabilities (4 bytes)
: These fields support beamforming, which is a topic complex enough to receive its own section in Chapter 4.

Antenna Selection Capabilities (1 byte)
: Antenna selection capabilities are used in systems which have more antenna elements than radio chains. This is not a common configuration for devices, and thus, these bits are not widely used in shipping products.

The HT Operation IE, shown in Figure 5-3, is included in transmissions from the AP to inform client devices of the current state of the network. It is included in Beacon, Probe Response, and (Re)Association Response frames.

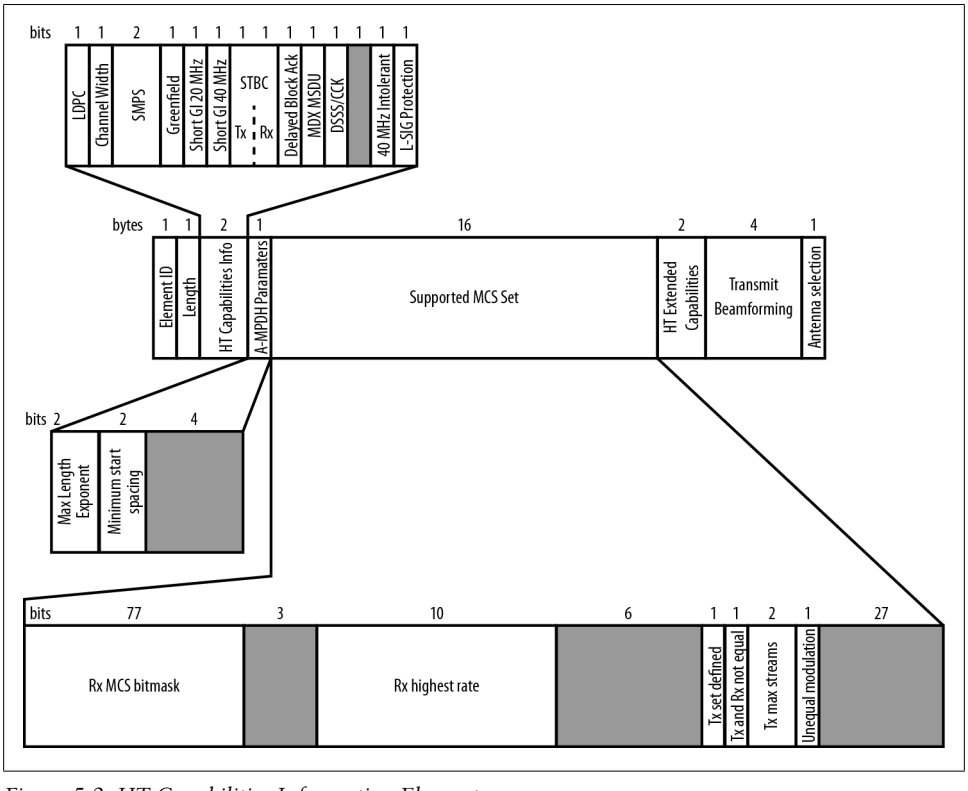


Figure 5-2. HT Capabilities Information Element

Primary Channel (1 byte)
: This field indicates the primary operating channel of the network. This is the channel number that will be used in management interfaces.

Secondary Channel Offset (2 bits)
: This field is set to 1 if the secondary channel has a higher frequency than the primary channel, 3 if the secondary channel is below the primary channel, and 0 when no secondary channel is present. Generally, most of the time, the secondary channel will be 20 MHz higher than the first channel; the secondary channel position is needed for coexistence mechanisms, as described in "Channel Width Selection (20/40 MHz BSS)" on page 61.

Channel width (1 bit)
: This field is set to 0 when the channel is a "narrowband" 20 MHz channel, and 1 when it is any other value.

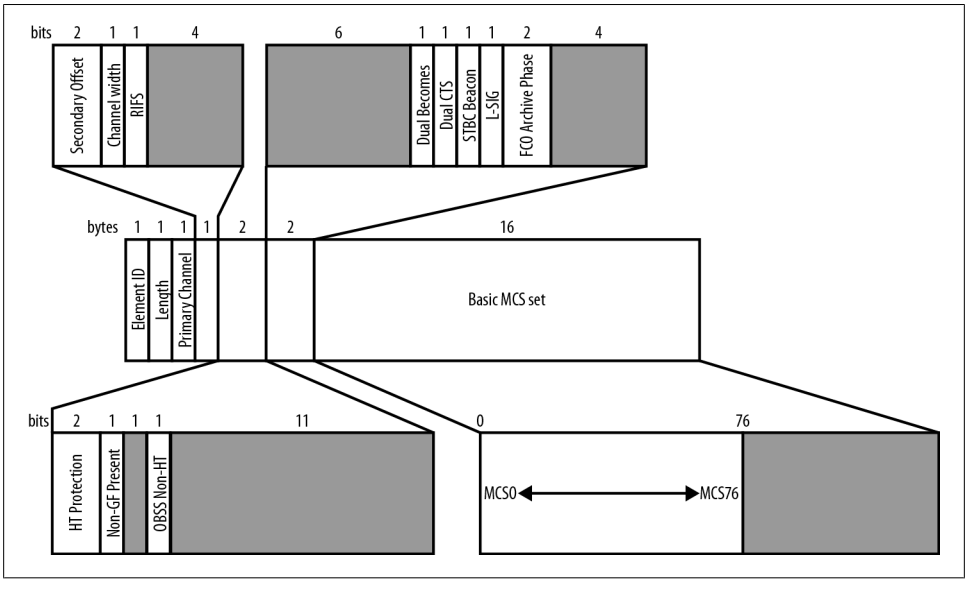


Figure 5-3. HT Operation IE

RIFS (1 bit)
: This field is set to 1 if RIFS operation is allowed, and 0 if it is forbidden. RIFS operation is discussed in more detail in the section "Reduced Interfame Space (RIFS)" on page 52.

HT Protection (2 bits)
: This field indicates the type of protection being used to avoid interference with pre-11n devices. Protection is discussed in more detail in "Protection of Non-HT Transmissions" on page 53.

Non-greenfield stations present (1 bit)
: When an associated station is not capable of greenfield operation, this bit is set to 1.

OBSS Non-HT STAs present (1 bit)
: Indicates that an *overlapping BSS* (OBSS) has non-11n devices associated that require protection.

Dual Beacon, Dual CTS, and STBC Beacon (1 bit each)
: These modes are used when the Beacon is transmitted using Space-Time Block Coding (STBC), which is relatively uncommon because it may render the beacon unintelligible to non-STBC stations.

L-SIG Protection Full Support (1 bit)
: Set to 1 to indicate full support of the L-SIG protection mechanism described more fully in "Protection of Non-HT Transmissions" on page 53.

PCO Active and PCO Phase (1 bit each)
: These bits indicate the use of the Phased Coexistence (PCO) method, which switches a channel between 20 MHz and 40 MHz. These bits are used to indicate that PCO is in operation and whether the channel is currently 20 MHz or 40 MHz. PCO is not widely used, so this book does not describe it in detail.

Basic MCS set (16 bytes)
: This string of 127 bits is used to signal support of the various data rates, defined by a modulation and code set (MCS). Each of the 76 MCS rates is assigned a bit, starting from zero, and each bit is set to 1 if the data rate is supported. Unused bits are reserved and set to 0.

Airtime Efficiency Improvements

The core idea behind aggregation is that gaining access to the wireless medium in 802.11 is time consuming. By forming an aggregate frame, an 802.11 device can spend more time transmitting, and effectively spread the cost of gaining access to the medium over frames that carry several higher-layer packets. 802.11n defines two types of frames: the *Aggregate MAC Protocol Data Unit* (A-MPDU) and the *Aggregate MAC Service Data Unit* (A-MSDU). The two types of aggregation are distinguished by where in the protocol stack they apply aggregation.

Frame aggregation nicely illustrates one of my favorite truths about standards: they tell you how to do something, but not when or why. If you go looking in the 802.11 standard for an explanation of when to build an aggregate or how to manage a transmit queue, you won't find it. It is up to individual product designers to build a queuing algorithm that makes appropriate choices on when to coalesce individual frames into an aggregate, and how to manage the size of aggregates so that large data frames don't block small high-priority frames from voice traffic.[1]

A-MPDU

The more common form of aggregate frame, the A-MPDU, is shown in Figure 5-4. The A-MPDU is a relatively simple form of aggregation in that the higher-layer (IP) packet that would normally be transmitted is given a MAC header and sent back-to-back. In Figure 5-4(a), the IP packet is given a MAC header and trailer, and the (unaggregated) frame is put into a PHY-layer frame for transmission. In contrast, Figure 5-4(b) shows the aggregation process. Several IP packets are each given their own header. Because these individual packets in the aggregation process will be put together, they are referred to as *A-MPDU subframes* at this stage. In the aggregation process, the *MAC Delimiter*

1. I am not aware of any detailed studies that compare various aggregation implementations. The effectiveness of an aggregation implementation will depend a great deal on the traffic mix on your network. If you expect heavy loads, it is worth looking at how aggregation will improve performance with your particular traffic mix.

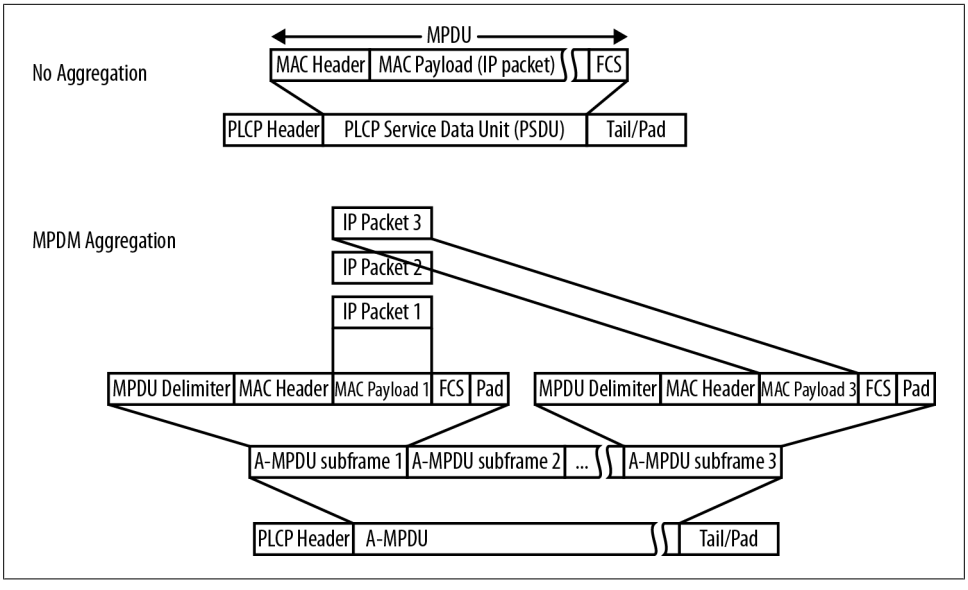


Figure 5-4. A-MPDU aggregation

is inserted to assist a receiver in extracting individual subframes. Each subframe gets a complete 802.11 header, frame check sequence, and is padded so that it aligns on symbol boundaries in the physical transmission. Because each subframe in an A-MPDU gets its own MAC header, encryption is applied to each subframe in isolation. Because each subframe in an A-MPDU has its own frame check sequence, an error in one subframe will only affect that subframe, and other subframes in the aggregate can be recovered. All frames within an A-MPDU must be destined to the same receiver address on the wireless link, but may have multiple destination addresses.[2]

A-MPDU is widely supported, and typically uses hardware-assisted processing. The Wi-Fi Alliance 802.11n certification program requires support for reception of A-MPDU frames. Although A-MPDU is widely supported, it is not typically easy to tell from an analyzer that an A-MPDU has been transmitted. Many analyzers plug into the host protocol stack at a level at which the aggregate frame has already been separated out into its constituent subframes, and each of those subframes is made available to the analyzer separately. An A-MPDU is limited in size only by the 802.11n PLCP, and thus can be up to 65,535 bytes.

A-MPDU Density

Although many receivers can process aggregate frames, the performance of hardware designs can vary. Gaining access to the wireless medium is an expensive process, and

2. See Figures 4-4 and 4-5 in *802.11 Wireless Networks: The Definitive Guide* for a discussion of the difference between the destination address and the receiver address.

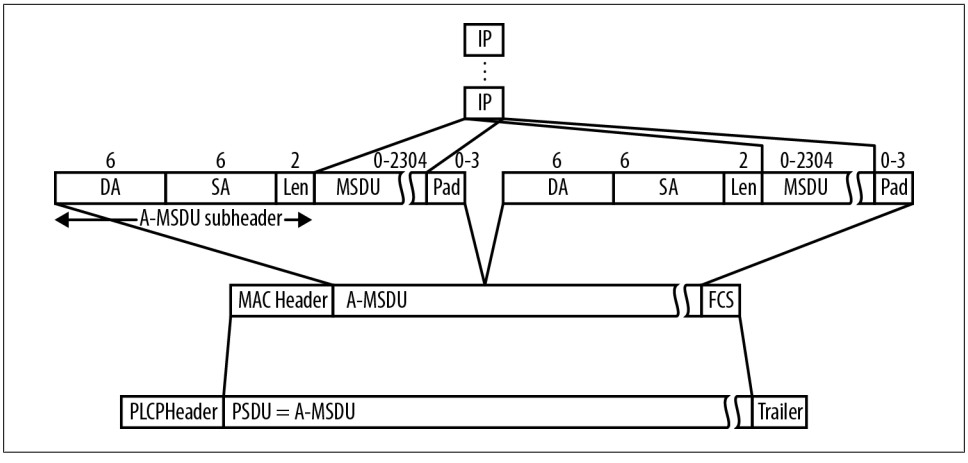


Figure 5-5. A-MSDU aggregation

even if it is necessary to insert padding in between A-MPDU subframes, the performance of a transmission can be improved dramatically over individual transmissions. A-MPDU size is limited both by the total size of an A-MPDU, as well as the time required between A-MPDU subframes. The acceptable values of A-MPDU timing parameters is shown in Table 5-1. In my unscientific experience, many devices are capable of 64 kB A-MPDU processing, even if they require 4 µs or 8 µs spacing between subframes.

Table 5-1. A-MPDU parameters

Parameter type	Allowed values
A-MPDU size (kB)	8 kB, 16 kB, 32 kB, or 64 kB
A-MPDU minimum spacing (µs)	unrestricted, 1/4, 1/2, 1, 2, 4, or 8 µs

A-MSDU

In addition to aggregation just before handing bits to the PHY for transmission, it is possible to pack multiple higher-layer packets into a single MAC frame. Building an Aggregate MSDU (A-MSDU) requires more software support because the network driver must take several higher-layer packets and build them into a single MAC frame payload.[3] The format of the A-MSDU, shown in Figure 5-5, shows how a higher-layer packet is put into an *A-MSDU subframe*, and those subframes are put together into a single MAC frame. In contrast to A-MPDU, this type of aggregation has only one MAC frame. As a result, when security is applied to an A-MSDU, one encryption operation secures the whole aggregate frame.

3. Although the aggregation is happening within the logical entity called "the MAC" in the standard, in implementation terms A-MSDU is typically carried out by the network driver. Aggregation therefore uses the host CPU and memory rather than the dedicated processing power on the radio card.

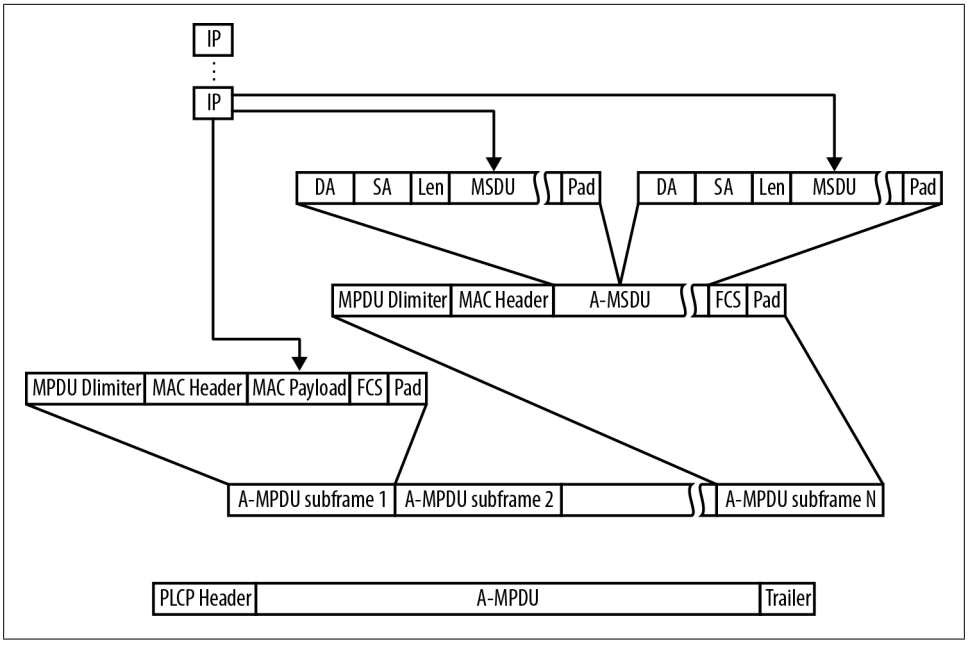


Figure 5-6. A-MPDU of A-MSDU

Destination address (DA)
: This field holds the destination MAC address of the subframe.

Source address (SA)
: This field holds the source MAC address of the subframe.

Length
: This field is the length of the MSDU in bytes.

MSDU
: Each individual constituent MSDU from higher layers appears after the destination/source/length triplet.

Padding (0-3 bytes)
: This field is computed so that the length of the A-MSDU subframe is an even multiple of 4 bytes.

An A-MSDU can have a maximum size of 7,955 bytes because all the aggregated frames must fit in the frame format shown in Figure 5-1.

The two forms of aggregation can be combined. Each of the A-MPDU subframes can itself be an A-MSDU, as shown in Figure 5-6. In the figure, the last subframe in the A-MPDU is itself an aggregate frame containing two MSDUs.

Aggregation Compared

The two methods of aggregation described in the previous sections are subtly different, and therefore, it is worth showing a comparison between the two. Table 5-2 compares A-MPDU and A-MSDU aggregation. Reception of both types of aggregate frames is widely supported, and about half of products support transmission of A-MPDUs.

Table 5-2. 802.11n aggregation types compared

Attribute	A-MPDU	A-MSDU
Payload size	Very large - about 64 kB	Larger than standard 802.11 frames (about 8 kB)
Software support	Lower. Aggregation is performed by hardware interface and presented to the operating system as independent frames.	Higher. Software (generally driver software executing on host platform) must perform A-MSDU assembly and unpacking.
Transmission overhead	Higher. Each subframe has a complete MAC header and FCS.	Lower. Only one MAC header and FCS is required for the aggregate frame.
Transmission reliability and interference resistance	Higher. Each subframe has its own FCS and can be checked independently. Loss of a subframe only loses a single component of the aggregate.	Lower. If the FCS check fails, all subframes are discarded.
Frame encryption	Applied to each individual subframe. Each subframe must be decrypted individually.	Applied to the aggregate as a whole. Only one decryption is necessary to obtain all subframes.
QoS traffic classes	Multiple. One traffic class per subframe.	One traffic class for all subframes.
Reception support required?	Yes, and tested by Wi-Fi Alliance certification.	Yes, and tested by Wi-Fi Alliance certification.
Transmission support required?	No. Optional in both 802.11n and the Wi-Fi Alliance certification test.	No. Not tested by the Wi-Fi Alliance certification test.

Block Acknowledgment

In the 802.11 MAC as it was originally specified in 1997, every frame required a positive acknowledgment. Each transmission was not complete until an acknowledgment was received. Network traffic is generally bursty. For example, a user reading web pages will send out a request, receive a flurry of packets carrying the requested web page, and then the network will generally be idle while the user reads the page. The original conception of the 802.11 MAC required that each frame sent to the receiver be acknowledged separately, as in Figure 5-7(a). The quality-of-service extensions in 802.11e brought in *block acknowledgments* (usually abbreviated as BlockACK), which allowed the sender to transmit a stream of frames and have them all acknowledged at once. Conceptually, the BlockACK extensions work similar to the selective ACK option in TCP. Two forms were defined in 802.11e, and subsequently carried over to 802.11n,

and both are illustrated in Figure 5-7(b). With the Immediate Block ACK form, the sender transmits a series of frames and expects an acknowledgment right away, while in the Delayed Block ACK form, the receiver can send its acknowledgment later.

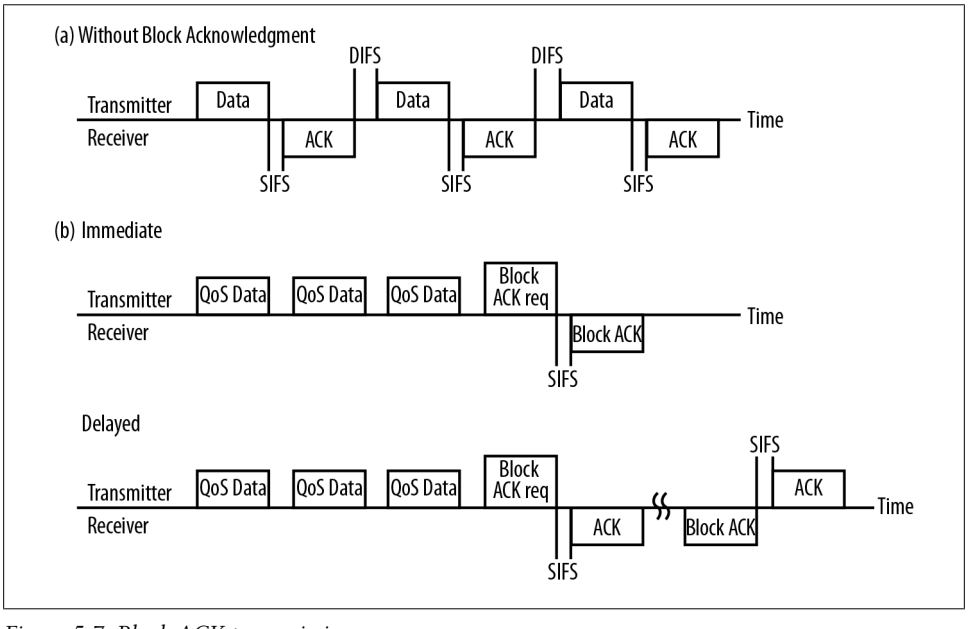


Figure 5-7. Block ACK transmissions

In Figure 5-7(b), the QoS Data frames are sent without any immediate positive acknowledgment. Instead of requiring that each frame be individually acknowledged, a single Block ACK frame exchange handles all acknowledgment. Because the Block ACK exchange is faster than individual acknowledgments, medium efficiency is improved. Block ACKs work by setting up a "window" for acknowledgment, and enable a receiver to selectively acknowledge any or all of the frames within a window. If any one of the three frames in Figure 5-7(b) are lost, the block acknowledgment can request retransmission of just the lost frame. For example, if the second data frame is lost, the Block ACK exchange will acknowledge the first and third frames.

Aggregate frames are well-suited for use with block ACKs because an aggregate frame holds several individual frames. By transmitting all the frames in the aggregate together, much less time is expended in the protocol overhead operations used to gain control of the channel. Block ACK processing was originally optional, but the efficiency gains possible when coupled with aggregate frame transmission are so compelling that Block ACK support is mandatory for 802.11n devices.

In addition to requiring support for Block ACKs, 802.11n defined new extensions to the protocol to reduce resource requirements for receivers of aggregate frames. The most notable of these extensions is the *compressed block ACK*, which takes its name

from the acknowledgment method. In 802.11e, Block ACKs worked on frame fragments, and the Block ACK could acknowledge individual fragments. Maintaining data structures that supported both the sequence numbers for assembled frames and the fragment identifier numbers that compose them required building significant state at the receiver. In 802.11n, the acknowledgment is compressed because it works on whole, unfragmented frames. As part of the standardization process, the task group considered whether to include fragmentation support with A-MPDU. Simulations showed that benefits existed only at higher bit-error rates than are widely used in the real world. Therefore, fragmentation support in block acknowledgment is not required.

To request a block acknowledgment, a station sends the Block ACK Request frame, shown in Figure 5-8. For the purposes of discussing a block acknowledgment in 802.11n, the important fields are the Compressed field, which requests use of the compressed form. The fragment number in the Block ACK Request Information field is set to zero.

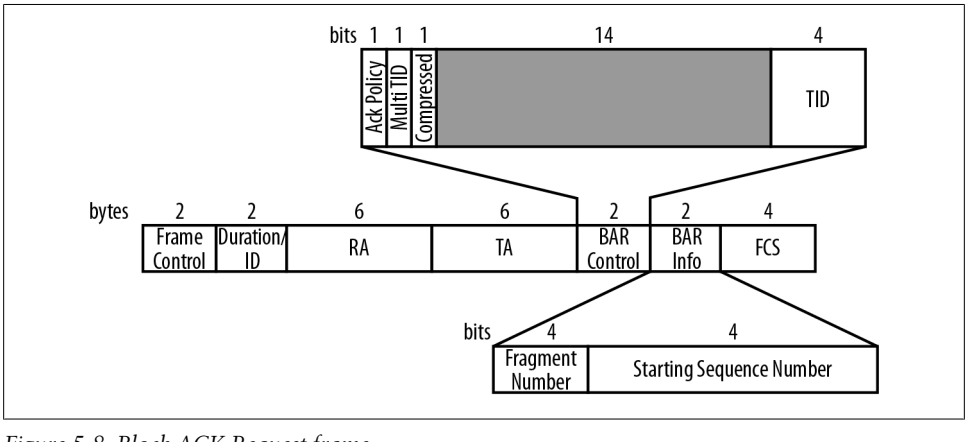


Figure 5-8. Block ACK Request frame

To send a block acknowledgment, the receiver uses the compressed Block ACK frame, shown in Figure 5-9. A single Block ACK frame can be used to acknowledge 64 MSDUs. In its "basic" uncompressed form, the Block ACK frame must acknowledge individual fragments, and is 128 bytes long. By restricting the block ACK to unfragmented frames, the length of the bitmap can be significantly reduced. Each bit in the bitmap acknowledges the frame that has that offset from the starting sequence number. If the starting sequence number is 100, then bit 0 acknowledges sequence number 100, bit 1 acknowledges sequence number 101, and so on.

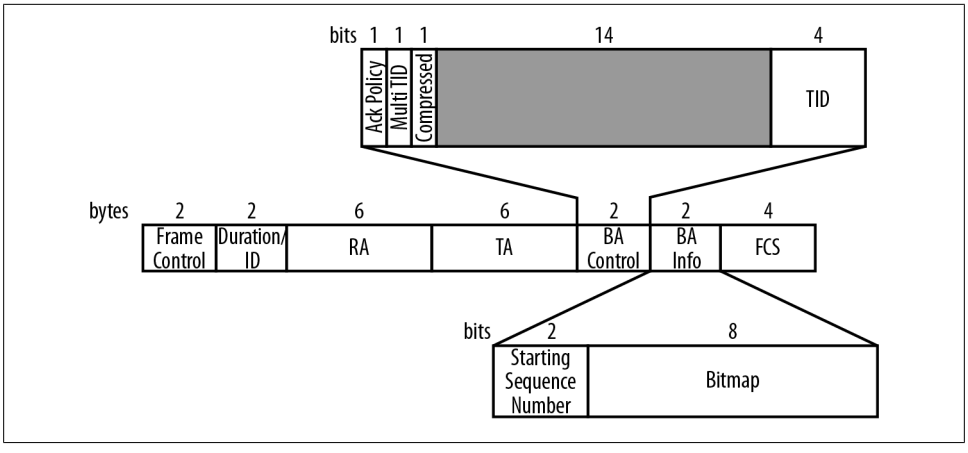


Figure 5-9. Compressed Block ACK frame

Reduced Interframe Space (RIFS)

802.11 basic channel access protocol uses spaces in between frame transmissions as a method of allocating access to the medium.[4] At the core of the interframe space mediation mechanism is that high-priority transmissions, such as acknowledgments, have a shorter waiting period and can jump on the medium in front of the start of new frame exchanges. Prior to 802.11n, the shortest interframe space was the short interframe space (SIFS), and it was used to complete frame exchanges by allowing response frames to be transmitted immediately following their triggers. For example, an acknowledgment can be transmitted after waiting only one SIFS. Likewise, a clear-to-send (CTS) frame transmitted immediately following a request-to-send (RTS) frame need only wait for one SIFS before getting on the medium.

802.11n defines a new interframe space, the *Reduced Interframe Space* (RIFS). It is functionally equivalent to the SIFS, and used whenever the SIFS might be used. Although it is shorter, as shown in Table 5-3, it does not define a new priority level. Its only purpose is to be used in place of the SIFS to increase efficiency. Obviously, it is not available on 802.11a/b/g devices, and, in fact, should not be used when 802.11a/b/g devices are present because it may prevent them from reading the Duration field in the frame header and updating medium access information.

Most devices do not transmit using the RIFS because the efficiency gain is relatively small. However, all Wi-Fi CERTIFIED n devices are tested for the ability to receive frames transmitted after the RIFS.

4. Medium access is discussed in Chapters 3 and 4 of *802.11 Wireless Networks: The Definitive Guide*.

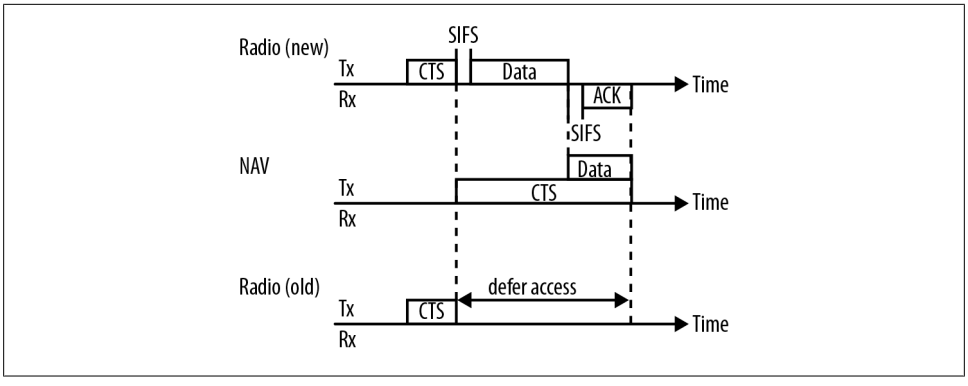


Figure 5-10. Protection basics

Table 5-3. Reduced Interframe Space length

Frequency band	SIFS value (µs)	RIFS length (µs)
2.4 GHz	10	2
5 GHz	16	2

Protection of Non-HT Transmissions

When new wireless LAN technology is introduced, it often transmits in a way that cannot be decoded by older devices. When 802.11g was introduced, the OFDM transmission style could not be decoded by older 802.11b devices. Therefore, 802.11g's standardized the first form of *protection*, named because it protects newer devices from interference by older devices.[5] Protection is not just a good idea: it's the law of the protocol. When migrating from an older technology such as 802.11g to a newer technology, it will be necessary to plan for the reduction in airtime and network capacity during the transition period where protection mechanisms will be most required.

In 802.11g, there was only one protection mechanism defined. Before transmitting in the newer style, a device had the responsibility to transmit in a backward compatible way to make sure older stations correctly deferred access to the medium. The most common way of achieving this is for a device to transmit a CTS frame to itself, using an older modulation that can correctly be processed. This method is shown in Figure 5-10, as a MAC-layer protection. Before transmitting a frame using the "new" style, a station sends a frame in the "old" style that tells receiving MACs to defer access to the medium. Although an older station is unable to detect the busy medium, it will still defer access based on the medium reservation in the CTS.

5. For more information on the protection mechanisms added in 802.11g, see Chapter 14 of *802.11 Wireless Networks: The Definitive Guide*.

Protection Mechanisms

In addition to the MAC-layer protections, 802.11n adds a new PHY-layer protection mechanism. The PLCP contains information on the length of a transmission, and 802.11n sets up the physical-layer header so that it includes information on the length of transmissions.

PHY-layer protection: L-SIG duration in HT-Mixed Mode

A simple method of protection is to alter the expected medium usage time that is calculated from values in the L-SIG frame header on the physical-level frame. In 802.11a/g frames, the SIGNAL header in the PLCP frame contains both a data rate and a length in bytes of the included MAC frame. Stations that receive the PLCP frame calculate a "duration" of the transmission from the data rate and length, and then defer transmitting until the calculated duration has passed. 802.11n devices operating in HT-Mixed Mode will typically set the data rate to 6 Mbps, and then reverse engineer a length field based on the amount of time required for transmission. (Wireless LAN analyzers must be programmed to discard the L-SIG data rate and look instead at the MCS in the HT-SIG field.)

If your wireless LAN analyzer is reporting that nearly all traffic on your network is sent at 6 Mbps, chances are that it is reading the L-SIG header and not the HT-SIG header. Upgrade to the latest version, be sure you are using an 802.11n capture interface, and look at a new capture.

When using L-SIG duration, 802.11n devices will set the L-SIG duration to be the amount of time required to transmit the included HT frame, plus any amount of time required for the next frame in the sequence. During a multi-frame exchange, the MAC duration field may be set for the duration of the entire exchange, but the L-SIG duration will cover the next frame in the sequence. Figure 5-11 illustrates the difference between the two forms of duration setting, and it shows how the L-SIG duration is distinct from the MAC duration. On the radio link, the transmission is simple: a CTS-to-self, followed by a data frame that is acknowledged. At the MAC layer, the NAV is set according to the basic protection rules. In L-SIG protection, there is a second "duration" value calculated from the data rate and length transmitted in a frame's PLCP header. Along the PHY line at the bottom of the figure, the shaded area along the PLCP line shows the extra length added to the legacy Signal field as computed by the PLCP. 802.11n's mixed mode frame header is identical to that used by 802.11a/g, and therefore, it is possible to use the PLCP header to establish the time required for the transmission. Using the PLCP header requires somewhat less time on the medium because it does not require a whole separate frame and interframe space.

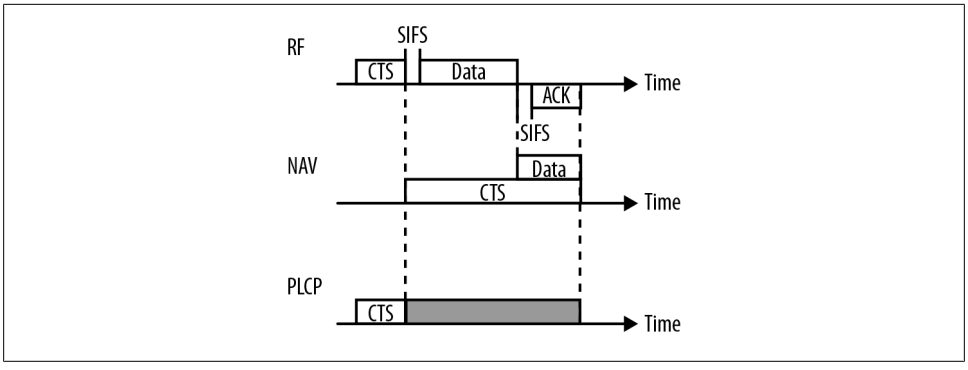


Figure 5-11. L-SIG duration

MAC-layer protection: CTS-to-self

Just as in 802.11g, an 802.11n device may transmit a control frame to lock up the medium for the duration of the entire HT exchange. Typically, the control frame will be a CTS-to-self to cause other devices in the area to defer access to the medium, though it is also possible to use a full RTS/CTS exchange. Control frames used to manage access to the medium are transmitted in a non-HT format so they can be understood by older stations. CTS-to-self protection may also be used to improve reliability.

MAC layer protection: transmission mode switch

A variation on the control frame-to-self mechanism described in the previous section is that the first frames in an exchange can be transmitted at non-HT rates, and set the MAC Duration field to lock up access to the medium. After the first two frames set at non-HT rates, it is then possible for the two devices in the exchange to switch to an HT mode and use whatever protocol options are desirable, including greenfield mode and the RIFS for the protected duration.

Protection Rules

The previous section described how protection works, but not when it is activated. Networks advertise the type of protection used through the HT Operation information element, described in Figure 5-3. 802.11n defines four protection modes, and each network selects one:

No protection
: Just as it sounds, in no protection mode there are no special transmission rules applied. This mode may only be set when all devices are 802.11n devices, and there are no non-802.11n devices detected. It is rare for an 802.11n network to be operating in no protection mode, especially in the 2.4 GHz band, due to the large number of existing 802.11 devices that were installed before 802.11n was standardized.

Non-member protection
: Non-member protection mode is used when a network must protect other devices from 802.11n transmissions. An example of this mode would be when I use an Apple AirPort Express 11n AP in a hotel room. On my network, I am using both an 802.11n laptop and an 802.11n AP, but the hotel's network is not 802.11n. Therefore, non-member protection is used to prevent my 802.11n network from interfering with the hotel's network. Non-member protection is only used when all devices attached to the network in protection mode are 802.11n devices.

20 MHz protection mode
: This mode is not commonly used because it is activated only for a 20/40 MHz network. Most networks are set up for either 20 MHz channels or 40 MHz channels, and do not switch between the two dynamically.

Non-HT mixed mode
: This method, described in "PHY-layer protection: L-SIG duration in HT-Mixed Mode" on page 54, is the most common form of protection used because most networks require protection and many networks have a non-802.11n device attached. In this mode, transmissions are protected implicitly by the use of the mixed mode header, and no special procedures are required for 802.11a/g devices.

802.11n protection also interacts with the existing Use_Protection bit defined in 802.11g. When the 802.11g Use_Protection flag is set, any frame transmissions used for protection are sent at 802.11 direct sequence or 802.11b rates (1, 2, 5.5, or 11 Mbps).

Security

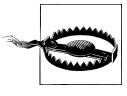
The basic security architecture first standardized in 802.11i did not require significant changes in 802.11n. There are three small changes to the security architecture. First, CCMP, the AES-based cryptographic protocol, was extended to protect the longer frames that appear in Figure 5-1. Second, the standard states that replay detection sets regular MAC payloads and aggregate MAC payloads on equal footing. Third, and most importantly, 802.11n specifies that TKIP is no longer allowed for use with 802.11n.

Whether to eliminate TKIP for use with 802.11n was a subject of debate within the 802.11 working group. TKIP was originally designed as a stopgap measure, and traded some amount of robustness in terms of security so that it could be retrofitted on to the then-existing 802.11b devices. TKIP's design was not easily extended to new protocol features; for example, when QoS capabilities were added, TKIP's design did not protect the contents of the QoS control field, which led to a small attack vector.[6] Further extending TKIP to protect fields added by 802.11n was too great a task.[7] As a result, there was never a codified technical standard for use of TKIP with 802.11n. In fact, the 802.11n standard went so far as to clearly and unambiguously state that the use of TKIP

6. The attack that was published was called Tews-Beck, after the two security researchers who discovered it. One of the better write-ups on the attack is Glenn Fleishman's article for Ars Technica.

with 802.11n was strictly forbidden. Given that many 802.11n networks will be built after a hardware upgrade, it also made sense to the working group to use the 802.11n transition as a way to improve the overall security of wireless LANs.

Nevertheless, some early 802.11n products implemented an 802.11n-compatible version of TKIP. As a result, the Wi-Fi Alliance 802.11n certification program began performing negative tests that prevented the use of TKIP with 802.11n data rates. Many 802.11n products retain TKIP in a "legacy" mode, in which data rates are limited to 802.11a/b/g rates in order to continue the use of TKIP.



TKIP and WEP are not allowed with 802.11n, and are forbidden by Wi-Fi Alliance testing. 802.11a/b/g devices are still allowed to do TKIP. Many certified 11n devices implement TKIP and WEP, but forbid them from using 11n rates.

Mandatory MAC Features

To help readers keep track of features that are mandatory and optional, Table 5-4 classifies the protocol features into two categories. Generally speaking, the Wi-Fi Alliance test plan validates functionality that is mandatory in the specification, and has optional testing only for the most widely supported features.

Table 5-4. Feature classification of MAC features

Feature	Mandatory/optional?	Comments
A-MPDU reception	Mandatory	Required by Wi-Fi Alliance test plan
A-MPDU transmission	Optional	Optionally tested by Wi-Fi Alliance test plan and displayed on interoperability certificate
A-MSDU reception	Mandatory	Required by Wi-Fi Alliance test plan
A-MSDU transmission	Optional	Not tested by Wi-Fi Alliance test plan
Block ACK	Mandatory	In addition to the basic requirements, 802.11n added support for the compressed block ACK
Protection	Mandatory	Basic protection mechanisms such as the CTS-to-self are required, but the L-SIG spoofing option is optional
RIFS receive	Mandatory	
Spatial Multiplexing Power Save	Mandatory	
No use of TKIP with HT rates	Mandatory	Validated by Wi-Fi Alliance test plan
Phased Coexistence	Optional	Not widely implemented, and not discussed in this book
Power Save Multi-Poll	Optional	Not widely implemented; see Chapter 6

7. TKIP is based on WEP, and requires the initialization of the RC4 cipher for every frame it protects. In A-MPDU aggregation, each subframe must be encrypted separately, and there is not enough time to re-initialize an RC4 cipher engine in the quick turnaround time between the processing of subframes.

CHAPTER 6

Advanced MAC Features for Interoperability

In the context of protocol design, *coexistence* can refer to many ways that devices are expected to show the equivalent of good citizenship and not cause undue harm to surrounding devices. 802.11n includes several capabilities that were completely new to 802.11 wireless LANs, and preserving the strong record of interoperability required that these new capabilities be designed in such a way that they could be used in the presence of older devices.

Radio Medium Coordination

802.11 derives a good portion of its total network speed from *frequency (or spatial) reuse*; that is, a single radio channel can be used multiple times in the same network. Each radio channel is subject to the rules of 802.11's CSMA/CA. Neighboring APs that are operating on the same channel must share access to the radio medium and therefore must not transmit at the same time. Neighboring APs operating on different channels can transmit simultaneously, and therefore will typically have much higher total throughput.

802.11 defines the term *overlapping BSS* (OBSS) to refer to another network that uses available airtime. By definition, an OBSS must be both on the same channel and in the same space; if either of those conditions are not met, the two networks do not interfere with each other. Generally speaking, networks are designed to minimize OBSSes, typically through careful frequency assignment. Neighboring APs are assigned to different channels; although their coverage areas overlap, they operate on different channels for increased capacity.

Prior to 802.11n, OBSSes were handled through the CSMA/CA protocol. All networks had similar channel widths, so detecting an overlapping network was not any different from detecting any other 802.11 transmission. After the introduction of wider 40 MHz

channels in 802.11n, however, additional work was required to ensure that overlapping networks using wider channels would be correctly detected.

Clear-Channel Assessment (CCA)

Because 802.11 is a listen-before-talk protocol, determining that the medium is idle is a vital feature for proper protocol operation. Before attempting transmission, an 802.11n device will perform a *clear-channel assessment* (CCA), and it may only proceed with the transmission if the CCA shows that the channel is idle. If the channel is busy, the device must defer transmitting according to the 802.11 protocol rules. CCA rules are built by choosing a level at which the CCA mechanism must detect a transmission, typically in dBm, and stating that any signal stronger than the threshold results in a CCA reading busy. As with previous PHYs, the CCA is based on both *signal detection*, where a device receives and decodes an 802.11 transmission, and *energy detection*, where a device receives a transmission whose energy is far enough above the *noise floor* that it will interfere with an 802.11 transmission. The noise floor is the ambient wireless signal in the area that transmissions must rise above to be heard. As a rough analogy, the noise floor is like the burble of conversation at a party, and a receiver can only detect a transmission when it is louder than the background noise. In a wireless networking context, the noise floor is increased by usage of the medium; in the 2.4 GHz band, the noise floor is increased by Bluetooth, cordless phones, microwave ovens, and even other 802.11 APs.

When an 802.11n network is set up for 20 MHz channels, the rules for detecting the medium as busy are fairly simple:

Signal detection
: When an 802.11n transmission at -82 dBm is received, the CCA is busy. -82 dBm was chosen as the cut-off because it is the minimum sensitivity specified by 802.11n. (Many products are able to substantially exceed that sensitivity level.)

Energy detection
: Energy detection is based on a simple rule that a strong signal will block 802.11 transmissions. To address the need to defer transmission in the face of strong interference, 802.11n specifies that a signal that is 20 dB above the minimum sensitivity will also cause the CCA to set the channel busy. That is, energy of -62 dBm is strong enough to cause the medium to be busy, even if no signal can be decoded.

Greenfield detection
: Even though not all devices can support greenfield mode, they must indicate the medium is busy for any greenfield transmissions received at a strength of -72 dBm or higher.

When the network uses 40 MHz channels, a few extra rules come into play. 40 MHz channels are divided into a primary and secondary channel. Beacons are sent on the primary channel. In essence, the clear channel assessment rules for 40 MHz channels

retain the use of signal detection on the primary channel, and use energy detection on the secondary channel.

Signal detection, primary channel
: When the secondary channel is idle, the primary channel must read as busy when an 802.11 transmission is received at -82 dBm. (This is identical to the 20 MHz signal detection rule.)

Signal detection, both channels
: When a 40 MHz transmission is received on both the primary and secondary channel at -79 dBm or better, the 40 MHz channel must be indicated as busy.

Energy detection, primary channel
: When the secondary channel is idle, the primary channel reads as busy when transmissions of -62 dBm or higher are received. (This is identical to the 20 MHz energy detection rule, which specifies a 20 dB difference from the primary channel rule.)

Energy detection, secondary channel
: If the primary channel is idle, the secondary channel will be read as busy when transmissions of -62 dBm or higher are received.

Energy detection, 40 MHz channels
: Both channels are set to busy when energy greater than -59 dBm is received. (Although this rule is specific to 40 MHz channels, it retains the 20 dB margin used in energy detection rules.)

Greenfield detection, primary channel
: Just as with the 20 MHz rule, greenfield transmissions at -72 dBm or higher cause the channel to be set to busy.

Greenfield detection, 40 MHz channels
: When a valid 40 MHz greenfield transmission is received, even if it cannot be decoded by a non-greenfield receiver, it must set the channel to busy if the strength exceeds -69 dBm.

Channel Width Selection (20/40 MHz BSS)

One of the ways that 802.11n dramatically boosts speed is the new 40 MHz channels. While using the wider channel might seem like a slam-dunk for the network administrator, there is a subtle trade-off between total network throughput and peak speeds. Peak speeds are easy to measure because it is the transmission rate measured by a receiver. Depending on how it is measured, the peak speed might refer to the data rate used on as little as a single frame ("2-stream 802.11n has a peak rate of 300 Mbps") or it might refer to the TCP or application-level throughput as measured by a user ("When I use the new 802.11n network, my download transfers at 10.5 megabytes per second"). The total network throughput is the data moved by all the access points in a network. Instead of being a measurement easily verified by a user, the total network throughput

is best measured by a network administrator, perhaps by adding up the throughput on all the switch ports used by the backbone connecting the APs in a wireless LANs, or using a network management application that performs the same function by querying all the access points in a network.

To decide between 20 MHz and 40 MHz channels, network administrators need to trade off peak throughput for total sustained throughput. 40 MHz channels require that twice as much spectrum be idle before beginning transmission. Even though the instantaneous speed of a 40 MHz channel will be higher at any given moment, there are fewer channels available for use with 40 MHz. As a result, the frequency reuse will not be quite as good in a densely packed network, and it may not be possible to avoid OBSSes. Wider channels are also more susceptible to interference because a 40 MHz channel will avoid transmitting when either of its 20 MHz halves is busy. The effect in the 2.4 GHz band is that a network built out of APs using 20 MHz will often show higher total network throughput, even though the speed measured to any client at a given instant is lower. Even within the 5 GHz band, wider channels can decrease total throughput, but the size of the effect depends on how much the wider channels cause overlap in the radio space. With products that support the full channel set for 802.11n, the effect in 5 GHz is often negligible.

In the 5 GHz band, spectrum is relatively abundant,[1] and channels do not overlap, so 40 MHz operation is straightforward (see Figure 3-3). To set up for 40 MHz channels, a network administrator simply tells the equipment to run at 40 MHz, takes two existing channel numbers, and enjoys the higher speeds.

In the 2.4 GHz band, however, there are no good choices for where to put a 40 MHz channel. Figure 6-1 shows the difficulties of enabling a 40 MHz channel in only 83 MHz of spectrum. Channel layout is made especially acute by the size of the band, which is only slightly larger than 80 MHz, plus the restriction that 802.11n must use existing channel numbers. 40 MHz channels are defined in terms of existing channel numbers, with one legacy channel number designated as the *primary* channel, and the other the *secondary* channel. The secondary channel is always at a 20 MHz offset from the primary. In Figure 6-1, the 40 MHz channel is channel 6, and it is using the 20 MHz below as an extension channel. The secondary channel is channel 2, but any transmission energy on channel 2 is well-placed to interfere with an existing network deployed on channel 1. It is also important to note that it is not possible to build two overlapping networks with 40 MHz channels in the 2.4 GHz band because it is only three legacy channels wide and a 40 MHz channel requires the width of two 802.11g channels.

During the development of 802.11n, there were extensive discussions about whether to allow 40 MHz channels in the 2.4 GHz band. The argument for wider channels is

1. With 40 MHz channels in 802.11n, spectrum is abundant. Future 802.11 standards will use even wider channels for higher speed and will make the 5 GHz band feel as crowded as 2.4 GHz does today. Or, put far more amusingly, Adrian Stephens has his own corollary to Lansford's Law ("Moore's Law does not apply to spectrum"): "Traffic increases to exceed the available spectrum."

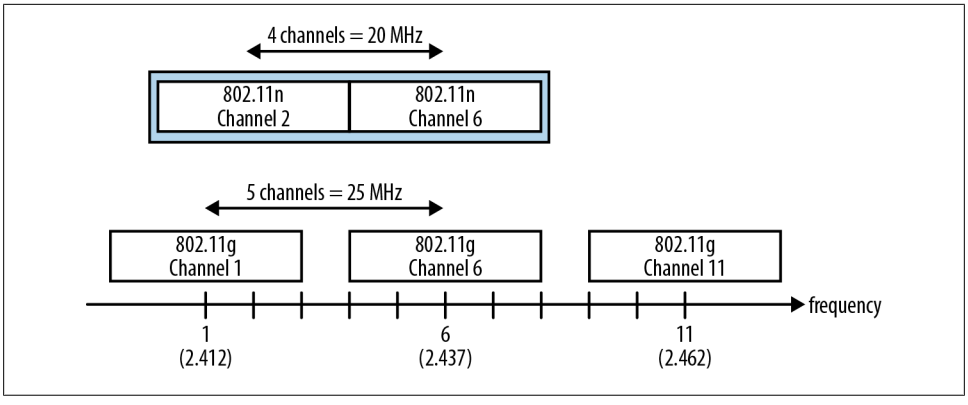


Figure 6-1. 40 MHz channels and the 2.4 GHz band

that it increases peak throughput, especially when there are no overlapping networks. In many residential environments, APs are separated far enough that each AP can operate independently without care for interference. The argument against using wide channels in the 2.4 GHz band is that wide-channel transmissions can cause significant interference with existing networks. As Figure 6-1 shows, the best layout for 40 MHz channels in the 2.4 GHz band means that each channel interferes with two 802.11a/b/g channels. While it is possible to design a network such that the 40 MHz channel is placed so that it overlaps with two older channels, that prevents the use of more than one 40 MHz channel. As a compromise, 802.11n includes the ability to use 40 MHz channels in the 2.4 GHz band, but only with coexistence mechanisms to reduce the potential interference between networks using 20 MHz and 40 MHz channels.[2] As a general rule, 40 MHz operation in 2.4 GHz is acceptable if there are no other networks in range. Although no network is an island, many residential networks are installed in environments where a network is not close enough to its neighbors to cause interference. (High-density apartments and condominiums are an obvious exception.) In a typical large-scale network, capacity requirements dictate that only 20 MHz channels be used because all the APs overlap.

Beacon frames are transmitted in a broadly supported form so that receivers can understand and react to them. In the 5 GHz range where the spectrum for extensive use of 40 MHz channels is readily available, Beacon frames are transmitted in a non-HT (non-11n) format so that any overlapping 802.11a networks will be able to react to them. Likewise, in the 2.4 GHz range, Beacon frames are transmitted using pre-11n modulations such as OFDM (802.11g) or even CCK (802.11b).

2. The Wi-Fi Alliance 802.11n certification program allows 40 MHz channels to be supported in the 2.4 GHz band only if coexistence mechanisms are supported. Most enterprise devices have not implemented 40 MHz support in the 2.4 GHz range, but many consumer-focused devices have.

Channel access rules

Before transmitting a 40 MHz frame, a station is responsible for ensuring that the entire 40 MHz channel is clear. Clear-channel assessment is performed on the primary channel according to the well-understood rules for transmission on an 802.11 channel. Even if a device intends to transmit a 40 MHz frame, the slot boundaries and timing are based on access to the primary channel only. The secondary channel must be idle for the priority interframe space before it may be used as part of a 40 MHz transmission.[3]

The virtual carrier sense mechanism carried in the network allocation vector (NAV) is updated only for the primary channel. Any 20 MHz frames sent on the primary channel update the NAV, and any 40 MHz frames update the NAV for the primary channel only. Devices are not required to update the NAV based on any frames confined to the secondary 20 MHz channel.

40 MHz Intolerance for Channel Width Interoperability

Because 40 MHz transmissions may cause significant impairment to overlapping 20 MHz networks, 802.11n includes a mechanism for a 20 MHz network to request of overlapping networks, "Please don't use 40 MHz transmissions near me." This mechanism, *Forty MHz Intolerance*, shares its name with the bit in the HT Capabilities IE used for signaling (see Figure 5-2), and applies only to networks operating in the 2.4 GHz spectrum. Within the 5 GHz band, there are a sufficient number of nonoverlapping channels that 40 MHz intolerance is not required.

Any device can use the 40 MHz Intolerant bit to disable 40 MHz channels in its immediate vicinity. When a client device sets the 40 MHz Intolerant bit, it instructs the AP to which it is associated that the AP should use only 20 MHz transmissions provided the AP is operating a 20/40 MHz BSS as described in the previous section. Access points that have been restricted to 20 MHz operation will then transmit a 20/40 Coexistence Management frame to indicate to neighboring APs that they should also switch to 20 MHz operation, if possible. Through the use of the 40 MHz intolerance bit, it is possible for an 802.11 device to shut down 40 MHz operation in its immediate vicinity, even on networks that merely overlap.

Power-Saving

802.11n devices can consume a great deal of power compared to their SISO predecessors. Much of the difference is due to the larger number of amplifiers used in 802.11n. Each transmitting radio chain requires a power amplifier to boost the signal before sending it out the antenna, while each receive chain uses a low-noise amplifier to bring the signal directly off the antenna up to a level that it can be used by the remainder of

3. For a more detailed discussion of channel access rules, see Chapter 3 of *802.11 Wireless Networks: The Definitive Guide*.

the components in the chain. Additional components in the transceiver chain also consume power. By turning off whole receive chains, substantial power saving is possible.

Spatial Multiplexing (SM) Power Save

The LNAs at the input end of a receive chain consume significant amounts of power, but are not always required. When a frame is destined to another station, or does not require the use of all receive chains, it is possible to save a significant amount of power by reducing the number of active receive chains. Spatial Multiplexing (SM) Power Save is an 802.11n protocol feature reduces power consumption by cutting down the number of active receive chains, and hence, the number of active amplifiers. Reduced power consumption is especially valuable in battery-powered devices because it can extend the time a device operates between charges.

SM Power Save static mode maintains power to only a single active receive chain; SM Power Save dynamic mode switches the number of receive chains to match the number required for an incoming transmission. In dynamic mode, a device normally operates with just a single receive chain active. To wake up the receiver for multiple receive-chain (and hence, spatial stream) operation, the sender must start the exchange with a frame transmitted using only a single spatial stream, such as an RTS frame. Figure 6-2 shows an example of an SM power save exchange. The sender of a multi-chain (MIMO) frame sends an RTS frame to the SM Power Save receiver, which is normally in a single-stream operating mode. For the receiver, reception of an RTS frame directed to it indicates that all receive chains should be made active. (Action frames may also be used to trigger activation of all receive chains.) The data frame can then be sent using any number of mutually supported radio chains and spatial streams. At the conclusion of the sequence, the SMPS receiver returns to having only a single chain active.

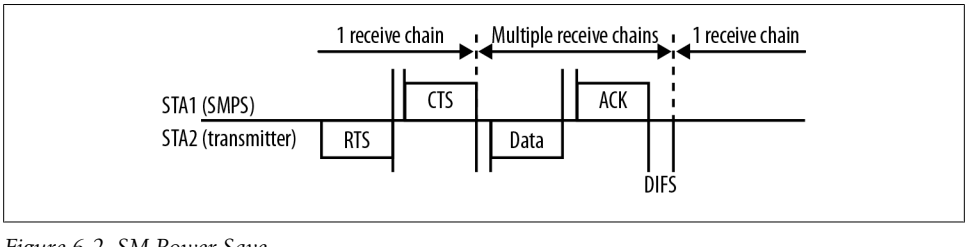


Figure 6-2. SM Power Save

Power-Save Multi-Poll (PSMP)

Some devices transmit small frames periodically. Examples include handheld data-collection terminals or voice telephones. Power-save multi-poll (PSMP) was designed with this type of device in mind. The AP begins a PSMP sequence, shown in Figure 6-3, with a frame that starts the PSMP period. To make more efficient use of airtime,

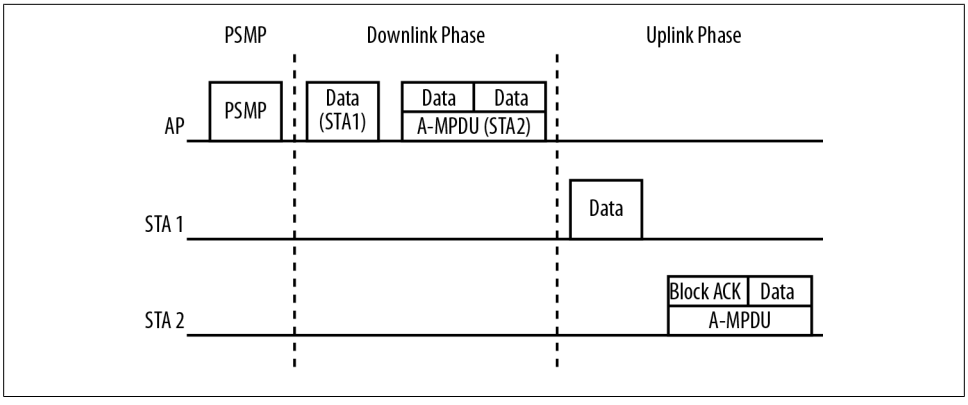


Figure 6-3. PSMP sequence

an AP divides up the PSMP period so that transmissions can occur without waiting for the lengthy medium-access procedures to complete. Before starting the PSMP sequence, the AP software lays out its transmit queue and selects frames to be transmitted during the downlink phase.

In the downlink phase, frames are transmitted as a burst. To improve efficiency, the burst consists of frames separated only by short (or reduced) interframe spaces. More importantly, the shortened interframe spacing allows receivers to power on only for the downlink phase and disable their receivers at other times. A receiving station capable of PSMP operation can examine the PSMP frame and learn the length of the PSMP downlink phase and power on only for the downlink portion.

Following the downlink phase, the uplink phase commences. Just as in the downlink phase, frames in the uplink phase can be transmitted without needing to wait for medium access. Each station is assigned an uplink transmission time for the uplink phase, which further assists in saving power because a device can sleep for the entire uplink phase except its assigned slot.

Although some restrictions exist on the types of frames that can be used within PSMP, it is possible to use both aggregate frames and Block ACK frames to further increase efficiency. Figure 6-3 shows how aggregate frames can be used, and shows the Block ACK being used with an A-MPDU.

PART III

Using 802.11n to Build a Network

One of my first laptops came with built-in Ethernet connectivity. I remember the shift in the vendor's literature from "you can use a parallel port adapter" to "this model now features a built-in Ethernet port." It wasn't a "real" Ethernet port, and the computer required an external adapter. Toting that laptop made me stand out because in the pre-PC Card days, hooking a portable computer up to an Ethernet port was just weird.

It is with some level of nostalgia that I look back on the heyday of Ethernet, since Ethernet ports seem only slightly more useful than an RS-232 port these days. Back in January 2008, the original MacBook Air launched. At the time, I remember thinking about how 802.11n didn't seem quite mature enough to leave off an Ethernet port — and I even worked at a wireless infrastructure vendor at the time!

That's not a question now, and the 8-pin RJ-45 connector is fairly chunky. With laptop case space at a premium, the preference of users for wireless LANs means that the Ethernet port can be eliminated without too much fuss. I've now used my latest computer for almost two years, and I can't remember if I've ever used the Ethernet port. On the occasions when I speak at universities, I often ask students about Ethernet, and it is not uncommon to be speaking to an audience where a majority of people haven't ever had to use a wired Ethernet port. There are a few specialized cases where the throughput of gigabit Ethernet is needed, but they are becoming increasingly rare (and are likely endangered by the gigabit successors to 802.11n).

More importantly, wireless LANs have grown up. After the "convenience" networks of the 802.11a/b/g days, we're now up to 3-spatial stream APs that can push data at up to 450 Mbps. With such great speed, network usage models have changed dramatically. Experience in building large-scale wireless LANs is increasingly widespread, with a variety of choices that give you top-notch radio management. It's now possible to build what is often called the "all wireless office," a network in which devices are connected using wireless LANs by default, and the only reason to use wire is for servers. Ethernet is dying as a desktop connection technology, but there is no crime. We've just found a network technology we like better.

CHAPTER 7

Planning an 802.11n Network

Ethernet is dead.

—Friedrich Nietzsche, had he been a network administrator instead of philosopher

On a very long airplane flight when writing this book, I decided that you have a special affinity for things that you recall when you "came of age." I can't think of a good reason why I find the electro-pop music of the 1980s a guilty pleasure, or why I feel sad that Ethernet is slowly dying. Both eighties music and Ethernet correspond to my coming of age in both fields. I can trace my career in networking to my first experiences with Ethernet, and I was building LANs well before every home had one.

Regardless of my feelings on the matter, Ethernet is dying as an access technology. Given the choice, most users (especially younger users!) will connect to a wireless network instead of digging for a cable and finding an Ethernet port. That is, if they even know what an Ethernet port looks like. In a striking illustration of the point, a university network administrator once told me that out of the thousands of sessions in the university library in the previous semester, less than 100 unique MAC addresses had been detected on the library's wired Ethernet network. It's a good bet that many students now entering universities have never had to deal with Ethernet in the same way that I did.

With 802.11n, performance has reached the stage where it is acceptable as a wire replacement. Home networks can happily run on Fast Ethernet, with 802.11n easily exceeding 100 Mbps. On business networks, gigabit Ethernet never supplanted Fast Ethernet to the desktop. Most laptops now come with 3-stream 802.11n interfaces capable of hundreds of megabits of real, application-level throughput. Ethernet is dying, due in large part to 802.11n. Doing your own bit to be part of the conspiracy starts with planning.

What's On Your Network?

Many of the readers of this book probably like network technology simply for the sake of network technology, and speed is its own reward. For most computer users, though, networks are a way to get work done. As a result, the types of devices the network must support drive the design. Throughout the design section of the book, I will be assuming you've already made the decision to jump to 802.11n. You may have made that decision because your network is running flat out with 802.11a/g and users still complained that it was too slow. Perhaps instead, you have devices that need the increased range of 802.11n. Or it might be as simple as the flattening of the price curve. When 802.11n APs were first introduced, they sold at a significant premium to venerable 802.11a/g designs. Today, you are hard-pressed to find enterprise-class devices that support 802.11a/g only. This book assumes you've already made the decision to go with 802.11n as the core technology, likely for one of the following reasons:

Peak speed/throughput
: Some applications require the highest speed that can be provided. Hospitals were early adopters of 802.11 wireless LANs because healthcare technology revolves around patient needs, and wireless connectivity can enable electronic routing of images. One of the earliest hospitals I worked with adopted 802.11a for radiology carts because the transfer of detailed X-ray images between the emergency room and radiology technicians benefits from high data rates. Many applications that focus on bulk data transfer can benefit from moving to 802.11n. Various imaging applications are an obvious fit, but the data transfer does not have to be file-based to benefit from 802.11n. With the high capacity enabled by 802.11n, it is possible to support many forms of streaming video.

Capacity
: The raw capacity of 802.11n allows a network to provide increased levels of service. Much of the reason why capacity is higher is the improved efficiency of individual transmissions. Many high-end wireless products wind up exploiting the improved efficiency to skew transmissions towards fast 802.11n devices. If you have a network that must support a large number of users, in practice there isn't a choice—just use 802.11n.

Latency
: Some applications benefit primarily from lower latency, especially real-time streaming applications such as voice, videoconferencing, or even video chat. 802.11n doesn't directly improve latency compared to older 802.11 standards, but it does offer an indirect improvement. Latency suffers when a network is operating close to capacity. By dramatically increasing network capacity, 802.11n reduces the amount of time a network operates with heavy congestion.

Range
: Some applications are most sensitive to range, and can benefit from the extended range that MIMO brings to wireless LANs. A common example is that quick service

restaurants will dispatch an employee with a portable computer to take orders from the drive-through line to help the queue move faster. Transactions are quite small, and could be served adequately with any wireless LAN technology, but the extended range of 802.11n enables employees to roam throughout the drive-through line.

Mobile End-User Devices

One strong driver for the use of 802.11n as the network infrastructure is the increased prevalence of 802.11n in devices. For the past few years, when you purchase a laptop with a wireless LAN interface, the underlying technology is 802.11n. The greater proportion of 802.11n devices that are going to be using the network, obviously, the greater driver there is to build an 802.11n infrastructure. A network that is designed to support primarily older 802.11a/b/g devices doesn't need the same level of infrastructure quality as a network designed to replace switched Ethernet connections with 3-stream 802.11n laptops.

Among different devices, the quality of the wireless infrastructure can vary widely. No standard dictates how to embed an 802.11n interface into a device, and the transmit power, receive sensitivity, and design of the antennas are all choices left to product designers. When two devices, both claiming "802.11n compliance," connect to an AP, the quality of the service they receive may vary widely depending on how much effort the product designer devoted to the wireless LAN interface. One device may have a top-quality chipset combined with well-designed and placed antennas, along with high transmit power. If it is placed next to a device that uses the cheapest possible chipset with off-the-shelf antennas that are poorly placed, it will experience poor connectivity, even though the AP is providing exactly the same service.

Almost as important as the type of devices is the number and density of devices. When I started working with 802.11 a decade ago, it was common to joke that APs had become the modern-day equivalent of oases, with users clustered around not a water source but a connectivity source. As wireless networks have grown in popularity, it has become common to see new devices that cannot use anything but a wireless network for connectivity. When phones with Wi-Fi first became prevalent, many people predicted that networks would need to support two devices per user (portable computer and phone). With the emergence of tablets, that prediction, if anything, seems conservative. As far as planning network capacity, most tablets are pure wireless devices without the option to connect to Ethernet, but they are also closer to computers than phones, with large screens that can more effectively present large amounts of data.

802.11n and Single-Stream Clients

The word "legacy" in the phrase *legacy clients* generally means "anything older than what I'm talking about," and as such, tends to shift over time. In this book, the term "legacy" is used to refer to 802.11a/b/g-only devices, and especially anything that is

only using SISO technology. Even 802.11n has single-stream speeds, which are used both by devices that have poor link quality as well as devices that have only a single transmit chain. Portable devices are often built using single-stream 802.11n because each radio chain consumes power, and the amplifiers are especially power-hungry. Using only a single-chain transceiver can dramatically increase battery life. (Apple's iPad is a good example; although it's an 802.11n device, it is a single-stream device with a top speed of 65 Mbps.) Single-stream 802.11n clients work very much like 802.11a/b/g devices, with only minor enhancements to transmission speed based on the improved coding unless 40 MHz channels are supported.

802.11n is not magic, and will not enable your older 54 Mbps client devices to suddenly break through that 54 Mbps barrier. What it does, however, is it improves what is called the *rate-over-range*. At a given distance, the AP with better rate over range will help clients connect at a faster rate. In Figure 7-1, there are two access points shown with the range of each data rate. As with any wireless technology, as you move farther from the AP, the data rate decreases. The older 802.11a/b/g AP at the top is a state-of-the-art design from 2008, and it uses the final generation of 802.11a/b/g chipset. The newer 802.11n AP at the bottom can't use any data rates faster than 54 Mbps because it still has to transmit at a rate the client understands. What it can do, however, is use the MIMO reception to increase the range at which the client can transmit at its highest data rates.[1]

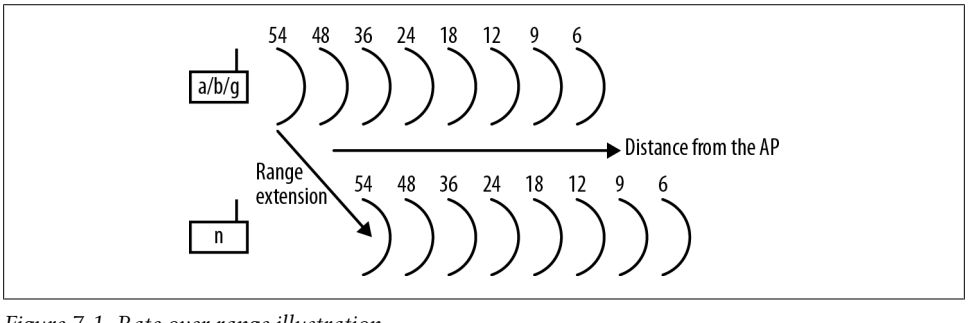


Figure 7-1. Rate over range illustration

No matter what you do, single-stream client devices will never exceed 65 Mbps. The job of the network with respect to single-stream clients is to offer the best service possible, at the longest range possible. To support large numbers of single-stream devices, focus on the rate-over-range performance to ensure that you are extending the best possible speeds out to the client. It is in networks that support a large number of single-stream devices that the backwards compatibility of a device comes into play, and if you have a large number of older devices that only support single-stream operation, your selection process should focus on how well 802.11n can improve the client experience.

1. Higher data rates at a longer range is attractive. It's also means that 802.11n APs providing service for older 802.11g clients provides "better g than g."

Overall speed is determined by the number of streams supported by both ends of the link. If you have a large number of single-stream client devices, consider weighting your AP spending towards a higher number of APs for better frequency reuse, not necessarily the fastest APs you can buy.

On the receive side, technology like Maximum Ratio Combining (MRC) improves the SNR by a few dB, but that can make the difference between one or two data rates. MRC works best when there are "spare" radio chains, that is, when the number of radio chains in use is greater than the number of streams. Every radio chain improves performance a little bit, so if you have a number of clients working at medium range, opting for a 3×3 AP may improve performance over a 2×2 AP by increasing the achievable data rate.

A 3×3 AP working with single-stream clients will have somewhat better performance than a 2×2 AP working with single-stream clients.

Although many devices have converted to 802.11n by this point, low-data-rate devices designed around a specific application may still be using older technology. 802.11a/b/g devices can have significantly lower power consumption, making it preferable for legacy devices where battery life is a key consideration. Older barcode scanners are just moving to single-stream 802.11n.

Bring Your Own Device (BYOD)

The wave of enthusiasm for bring-your-own-device (BYOD) programs, which allow employees to use personally-owned devices on a corporate network, has been driven by mobile devices that increase productivity by putting information quite literally in the hands of users. There is a great deal that can be written about BYOD programs and what they mean for network administrators and IT departments. I'm not going to attempt to write much about BYOD because the field is evolving so quickly.

Generally speaking, what it means for network administrators is that there are, broadly speaking, a few classes of devices that get used on a network: corporate-owned devices, such as a company-supplied laptop, phone, or tablet. *Corporate-owned devices* are often managed and have policy enforcement that requires that you run anti-virus software, perhaps use a web proxy, and possibly even lack administrative rights to the machine. *Guest devices* are brought in by guests, such as visiting sales people or business partners. Guest devices typically are set up with Internet access and nothing else. In the past, networks have often had an ad hoc procedure in place for contractors to enable them to use their own laptops and have some level of privilege. (I once worked at an organization that created an Active Directory account for an attorney that visited our office once a week, but he only used that account for access to the wireless LAN.) Contractors

typically receive some access, but not full access. And that's before employee-owned devices started showing up.

Why am I writing about BYOD in a book about wireless networks? All the productivity-enhancing devices that people want to use (and in many cases just started using without permission) have wireless LAN interfaces. Many of them don't even have the ability to connect to Ethernet networks, and therefore are forced into using the wireless LAN. As that first point of contact, it falls to the wireless network administrator to enable access for employee-owned devices...but not too much access.

If you are thinking of supporting a BYOD program on your wireless LAN, look for flexible security mechanisms that can be used to identify and register employee devices, and a wireless LAN that can effectively segregate traffic based on the type and ownership of devices, and can be used to easily identify and register devices. It's also important to "overbuild" your network for a worst-case scenario. With BYOD, what little control you may once have had over devices on the network is gone. All it takes is one popular device that has poor radio characteristics to ruin your day. As part of starting a BYOD program, pick a "worst case" device performance and design around that for two reasons. First, clearly documenting the minimum supported device will help in explaining performance problems for devices that fall below that baseline. Second, having a clearly documented worst-case scenario provides guidance on how to redesign the network to handle a problematic device. If, for example, you have designed a network to work with devices at an average signal strength of -70 dBm and must cope with a device that requires service of -65 dBm, use your favorite planning tool and see what would be required to redesign the network for a 5 dB increase.

Traffic and Application Mix

As important as the type and number of devices is the activity that those devices support. 802.11n can dramatically increase the throughput of a network that supports data-transfer applications, especially if paired with 802.11n client devices. Data transfer in bulk can be sped up through the application of higher data rates, aggregate frames and block acknowledgements. Fortunately, most applications fall into this type. File transfers obviously benefit—with 802.11n, you can give a wireless device the full benefit of Fast Ethernet-type data rates and transfer at just over 10 megabytes per second. Any application that works like a file transfer benefits similarly; medical imaging applications must often fetch large, detailed images from a server on to a device for display. Even web browsing, widely viewed as a "light" application, benefits from being able to pack multiple frames from the web server's response into downstream data.

Real-time streaming applications may or may not benefit, depending on the characteristics of the application.[2] Block ACK procedures can only be used between a sender and a single receiver, and therefore cannot be used with a multicast application. Mul-

2. Non-real-time streaming applications such as buffered video transmission are much more like bulk data transfers as long as the network can keep the buffer reasonably full.

ticast applications may benefit from the higher data rates in 802.11n, but it depends on the application and your network equipment. When a frame is destined for a group of receivers, it must be transmitted at a data rate supported by all receivers. Most 802.11n networks are designed to support any type of client, and will therefore send certain management and control frames at older 802.11b rates. The easiest way to ensure that multicast frames are transmitted at a rate that can be understood by all receivers is to pick a very low data rate, but this prevents an 802.11n network from being any better than the 802.11a/b/g network it is replacing. To take full advantage of 802.11n for multicast streaming applications, an AP must monitor the data rates that each receiver is capable of. Typically, that is accomplished by monitoring Internet Group Management Protocol (IGMP) messages to determine when a client has joined or left a multicast group. By maintaining a dynamic list of receivers in the multicast group, the AP can select the highest data rate used by all group receivers instead of all receivers. In a network that supports mixed traffic of 802.11b/g devices and 802.11n devices, such an AP can use the much faster 802.11n data rates if the group members are all 802.11n devices.[3]

One streaming application that is not directly helped by 802.11n is voice traffic. Voice is a special case because it requires streaming data, but the nature of voice prevents it from being buffered. It is not possible to speed up voice by using 802.11n because there is very little benefit to transmitting the regularly scheduled voice frames at 802.11n speeds. An 802.11n access point may improve voice traffic indirectly by offering superior reception with a more advanced radio design, as well as by increasing the overall capacity of the network and freeing airtime for voice devices.

IPv6 on Wireless LANs

IPv6 generally works easily on wireless LANs because the MAC layer is distinct from the IP layer. One common shortcut used in early enterprise-grade 802.11n hardware was to use a single broadcast/multicast encryption key per SSID. When IPv6's stateless address auto-configuration (SLAAC) is used, a wireless client device will listen for IPv6 router advertisements, take the advertised IPv6 prefix, and assign the last 64 bits of the IPv6 address from the client MAC address.

If you have multiple VLANs attached to a single SSID, all the IPv6 router advertisements on every VLAN will be received by all clients. Even though a client may be assigned to one particular SSID, it will get IPv6 addresses from every VLAN. At the time this book went to press, some vendors were working to lift this limitation.

3. Some APs can also convert a single multicast frame into a series of unicast frames, which can also boost speeds.

Network Integration

In the parlance of network designers, the *access layer* consists of both wired and wireless networks, and provides the attachment point for user devices. A wireless network can deliver service only to the extent that it is connected to a reliable backbone that supports the desired level of service. After many unsuccessful efforts to deliver gigabit networking to the desktop, the need to provide a high-quality support foundation for 802.11n has driven the need to push gigabit network links out to the wiring closet.

Network Services

The starting point for planning out how to connect a wireless access layer to your network is to create a complete inventory of services the network will need to support. Integration of services into the network framework is shown by Figure 7-2. The figure is meant to be conceptual in that it shows where services might attach to an existing network core, and what services the wireless access layer may depend on. This section classifies services based on the part of the network where they require the highest level of support.

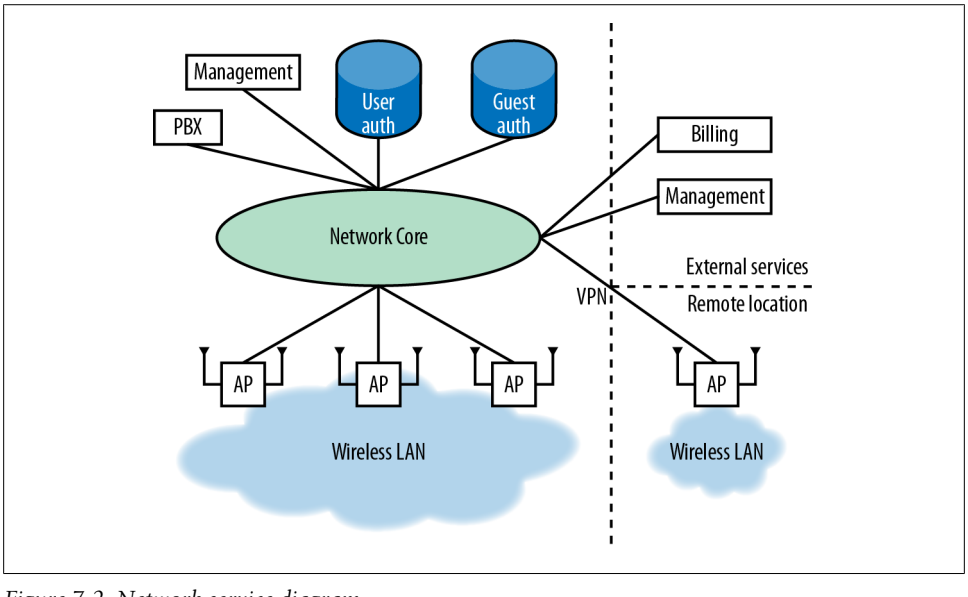


Figure 7-2. Network service diagram

Edge services

Services at the network edge are largely implemented within the wireless LAN equipment itself, though the precise implementation details vary from vendor to vendor.

Quality of Service (QoS)
: Quality of service on wireless LANs is dependent on both the wireless link as well as a path through the core taken by traffic. In general, network core capacity exceeds that of the wireless LAN, and delivering robust service quality depends more on the wireless link than the rest of the network. QoS mechanisms for 802.11 were first standardized in 802.11e-2005, and are tested by the Wi-Fi Alliance as part of the Wi-Fi Multi-media (WMM) certification. WMM works by mapping QoS information from the network core into preferential treatment on the wireless link, and hence it acts as an extension of any QoS policies in the core network.

Fast roaming
: Moving connections rapidly between APs is critical for real-time applications such as voice and videoconferencing. When security must be included as part of the handoff between APs, there are two major implementation paths. Opportunistic Key Caching (OKC) moves the master key between APs and is widely available in network equipment. The emerging 802.11r specification provides a guaranteed fast transition capability, but it is not yet widely implemented due to complexity and the acceptable performance of OKC in many scenarios.

Device/operating system identification
: Providing differentiated services to devices based on type requires that the network identify the class of device. Wireless LAN access points in an ideal position to observe information about devices as they attach, and use information such as DHCP requests to identify the operating system for use in policy enforcement.

Security (encryption)
: Wireless LAN security services are provided at the termination point of the wireless link. Typically, security services will only be enabled once user authentication has completed, and an authorization process has verified a user's right to access the network.

Security (filtering)
: To reduce the load on the network core, the wireless LAN may provide an integrated firewall to restrict access to certain parts of the network.

Spectrum analysis
: With the right hardware support, access points can monitor the physical layer directly and identify non-802.11 devices that can interfere with the network. It is a valuable troubleshooting tool, especially when intermittent interference must be detected and located.

Network-wide services

In contrast to services at the edge, a network-wide service is generally provided at a point beyond the wireless link. Again, depending on a vendor's implementation, parts of these services may be implemented in code that runs on an access point at the edge of the network.

User authentication (captive web portal)
: When the highest levels of security are not a concern, captive web portals are used for authentication. When an unauthenticated device attaches to the network, a filter prevents access beyond the access point until authentication has completed. Captive web portals are often used in public access deployments, such as airports and hotels, because every device has a web browser and most users will easily be redirected to the web portal. Captive web portals can use a variety of databases for authentication, but are unable to provide strong link-layer encryption.

User authentication (RADIUS or LDAP)
: For standard internal users such as employees or students, it is more common to tie access to a database of users. Database access can either be done through a RADIUS server, or by a connection straight into a directory server. When a user store provides authentication, it is possible to use strong link-layer encryption so that each user has his or her own unique per-session key.

Network configuration management
: Network administrators need to be able to change the configuration on network elements. Typing a single change into hundreds of access points is time-consuming, error-prone, and a waste of time. Centralized change control is a practical requirement for a network of any size. Traditionally, network management was provided by management applications that are installed on a dedicated computer. However, the software-as-a-service revolution has come to the wireless LAN, and it may be possible to "rent" access to a full-featured management application for substantially less money than a traditional license.

Network monitoring and reporting
: Once a network is installed, an administrator needs to be able to see the overall health of the network, monitor usage, and inspect system logs. These features are often coupled with configuration tools in the same application.

Guest access (registration)
: With most mobile devices now using wireless LAN interfaces as their primary (or sometimes only) method of network connectivity, it is common to extend network access to guests at an organization. Registering guest accounts can take many forms depending on security requirements, but will often be built using either a self-registration page accessible through a captive web portal or a system that allows network administrators to delegate guest account creation to other employees.

Guest access (billing)
: When a network is built to provide service to the public at large, it typically charges for access. Collecting access charges is a vital component of such a network, and may consist of a credit card gateway, integration with a hotel's billing system, or even a connection to a centralized clearinghouse.

Network admission control
: Admission control denies access to devices which do not meet a baseline security policy. Common examples of an admission control policy are "your anti-virus

definitions are not up to date." Admission control is frequently quite expensive because robust policy enforcement requires client-side software to collect security state, and installing software on individual devices is a complex endeavor.

Location
: Location services assist the network in finding devices, typically on a graphical map, by correlating received signals from multiple listening points. To build a network that supports location, a high density of APs will be required so that the target location zones have multiple APs available to correlate signals.

Wireless intrusion detection
: Wireless intrusion detection systems (WIDS) work by observing radio traffic and searching for patterns that indicate an attack. WIDS can be built as a completely separate system, or, more often, it can be built into a wireless network as a feature of the same access points that provide service.

Remote services

As a final reflection of the way that wireless LANs have taken over the "last hop," it is now possible to build an access point that provides the internal wireless network from any point on the Internet. These combined VPN/access point devices connect from any remote location to the network core using strong VPN technology, and then act exactly like access points connected directly to network core. Remote locations are fully managed, and provide the exact same wireless LAN as used within headquarters. VPN access points make it easy to set up off-site meetings or remote branch offices because users see the wireless LAN exactly the same as at headquarters.

Backbone Connectivity

Generally speaking, connecting 802.11n to the backbone network is fairly simple. Since even a basic 2-stream AP can readily push 100 Mbps per radio at peak, you want to have gigabit switching infrastructure to support an 802.11n access layer. It is possible to run 802.11n on Fast Ethernet, but any 802.11n AP can easily bottleneck at Fast Ethernet speeds. To support a fully functional 802.11n network, start the project by ensuring that backbone connections are fast enough, whether through an upgrade to gigabit Ethernet switches or using Fast Ethernet link aggregation across multiple uplink ports on the AP.

How you connect the network depends on the services that the network needs to support. Generally speaking, management protocols require very little overhead and are not a key consideration. If possible, map out the major flows of traffic on the network and use that research to determine the major sources and destinations. In a network that supports extensive use of virtual desktop infrastructure (VDI), it is likely that most traffic will be to a handful of VDI servers in a data center. On the other hand, a university campus with highly mobile students and departmental infrastructure likely has more of a mesh-style traffic flow that does not have natural choke points. Network access

for guest users over the wireless network typically pulls all guest traffic back to a central point, and naturally fits a centralized forwarding model.

Power Requirements

In a network of any significant size, most APs will be powered from edge switches. As a practical matter, AP product designers have recognized that the 802.3af power-over-Ethernet (PoE) standard is widely installed. APs must either operate within the power limits impose by 802.3af, or give network administrators a reason to install special power equipment. Not surprisingly, most APs work within the 803.3af power limit.

In 802.3af, an Ethernet port supplies 15.4 watts of power at the switch, though the standard only guarantees a little bit less than 13 watts over a maximum-length cable due to losses from cable resistance. For a highly functional 802.11n AP, the power limit is quite strict. The major components in an AP are the radio modules that provide the wireless connectivity, the Ethernet controllers that connect the AP to the network, and the CPU that runs all the software to make it happen. AP designers carefully select components with one eye on power consumption to ensure that an AP will run with only 13 watts.

To run the AP, power can be supplied using one of the following methods:

DC power adapter
: Most APs have the ability to take DC power directly from a wall outlet. Of course, the best place to install APs is typically on the ceiling, and very few buildings are built with AC outlets in the ceiling. Early wireless LAN deployments sometimes put power lines into the ceiling, but hiring licensed electricians and complying with building codes is quite expensive.

Power injector
: Power injectors are network devices that plug into an existing Ethernet port and AC power in the wiring closet, and combine the data signals and power into a single Ethernet cable to the AP. Power injectors can be purchased separately from switches, and can be used for installations where AC outlets are not available at the AP and the edge switch is not capable of supplying power.

802.3af Power over Ethernet (PoE)
: PoE is the most common way to power access points because most access points also require a connection to the network. The 802.3af standard supplies 12.95 watts to the end device. Once the network cable is put in place, the power rides along with it. If you decide to relocate an access point, moving the network connection moves power with the AP so there's no need to worry about where power is physically present. Low-voltage wiring is significantly easier to install than AC power because the wiring is more flexible, easier to handle, and is not quite the same safety hazard.

802.3at power ("PoE plus")
: When it became clear that some devices were going to exceed the 13-watt guarantee of 802.3af, the industry went to work on an improved standard that is often called "PoE plus." It supplies over 30 watts, which is more than enough to power any 802.11n AP you plug in. The downside to using 802.3at is that you may need to purchase mid-span power injectors to use it or upgrade the edge switches to models that support 802.3at.

Multiple Ethernet cables
: The clunkiest solution to the power problem is to require that an AP have two Ethernet cables, each with 802.3af. The AP then draws power over both cables. In a way, this is the worst of all worlds. Cabling costs are very high because two wires are required for each AP location, and two power injectors or PoE ports are required for each AP.

Better than any of these four is to purchase APs that work within the 802.3af power limits, as any high-end 802.11n AP released since 2008 will have been designed to work within the 802.3af power ceiling. Prior-generation APs that consumed more than 13 watts were often designed with a prioritized partial-shutdown list that would power down components to allow the AP to start up within the limit imposed by 802.3af. APs with dual Ethernet ports would often start by shutting down unused Ethernet ports. Typically, the next step would be to limit the clock speed of the CPU, which might limit software functionality. Reducing the functionality of the wireless interfaces was a last resort; 802.11n's power saving specifications allow for the shutdown of individual radio chains when used for communication with less capable clients. Many vendors exploited that by throttling 3×3 APs back to 2×2 operation to save the power of running the third radio chain. Although shutting down a radio chain limits performance, it will still be faster than 802.11a/g.

Power Headroom

Good engineers leave headroom in a specification, and power over Ethernet is no exception. Nominally, 802.3af delivers up to 350 mA at 48 volts of electrical power to the end of a maximum-length (100 meter) Ethernet cable. PoE supplies 15.4 watts into the Ethernet cable, though resistive loss means that the only power guarantee is 12.95 watts at the end of the cable. Fortunately, high-quality components throughout the power system can make a big difference. Higher wiring categories (5e and 6) have lower resistance in the cable and lose less energy to electrical resistance. High-quality power injectors are typically built with headroom, and supply somewhere in the range of 50 to 52 volts to the powered device. Although it doesn't sound like much, that's an extra 4%.

Most Ethernet cable runs are shorter than the maximum 100-meter length. Ethernet cable is quite thin, which leads to fairly high resistance; 802.3af is specified for both the older category 3 and newer category 5 cabling. If the typical cable run is only 40 meters and is high-quality category 5, substantially more power is available at the end

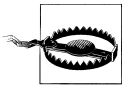
of the cable than the standard specifies. Additionally, most cables are built above what the specification requires, and have lower resistance.

All these factors combine to allow many off-the-shelf power injectors and 802.3af switches to safely supply about 20% more power than nominal, or about 15 watts.

Security

It may seem odd to put security in the overall network planning chapter of the book, but there is a reason for it. Security is both simple and complex with 802.11n. Simple because 802.11n has specified that the only method of security with 802.11n is the Counter Mode with CBC-MAC Protocol (CCMP), sometimes better known as "the AES protocol" in informal conversation.[4] Although this restriction makes choosing security quite simple, it can make upgrading to 802.11n complex because older security options are available. Back when the Robust Security Network (RSN) architecture was being developed, there were two parallel tracks for security. One was a backward compatible method that would work with existing devices that had been shipped since 1999; that approach led to the Temporal Key Integrity Protocol (TKIP). TKIP was widely used in networks in early wireless LANs because it was developed and tested earlier. The second component of the RSN architecture was a forward-looking security system based on AES that was intended to fix wireless security once and for all; it later became CCMP. The RSN architecture was finalized in 2004, and all wireless LAN chips produced since then support both components.

During the development of 802.11n, the question arose as to whether TKIP should be supported with 802.11n. Standards groups do not always speak with a single voice when discussing proposals, but the debate and discussions around whether to allow 802.11n devices to support TKIP was about as one-sided a discussion as I've ever seen. TKIP was showing its age cryptographically, and it was difficult to extend TKIP to any technology developed after it. CCMP was readily extensible, and easily adapted to work with 802.11n. The rule is pretty simple: if you want security with 802.11n, you must use CCMP.



If you plan to encrypt data on an 802.11n network, you must use CCMP. TKIP is not supported.

In fact, the Wi-Fi Alliance certification program for 802.11n specifically tests to ensure that certified devices do not implement TKIP with 802.11n. If you are using 802.11n, it will by definition be using the strongest possible security. The downside is that if you

4. CCMP is sometimes used interchangeably with the name of the Wi-Fi Alliance certification program that tests for CCMP interoperability: Wi-Fi Protected Access, version 2 (WPA2).

have an existing network that is based on TKIP security, you'll need to either continue at 802.11a/b/g rates or make plans to move to CCMP.

Additional Security Features

Some wireless networks use additional security components, such as firewalls to filter traffic between wired and wireless networks or wireless-specific intrusion detection/prevention systems to stop attacks in progress. From the perspective of such security systems, 802.11n is just a faster-speed link. Any firewalls in place will continue to work as they did with 802.11a/b/g networks, and will perform the same purpose. An existing 802.11a/b/g wireless intrusion system will not be able to detect 802.11n transmissions if it does not use 802.11n interfaces. For the most part, any security systems you have built on top of an existing 802.11 network will work after the upgrade to 802.11n unless they need to tap into the wireless link directly.

TKIP Transition Planning and Support

If you are upgrading to 802.11n and have an extensive TKIP deployment, the security transition requires its own planning. 802.11n APs will not support TKIP with the fast data rates of 802.11n. It is possible to continue to use TKIP by using the 802.11n APs in a "legacy" mode where they use 802.11a/b/g data rates with the improved radio technology of 802.11n. Such a network will have improved range as well as rate-over-range, but the performance gain will not be as substantial as if the client devices were unleashed to use 802.11n rates.

Many 802.11n devices will support TKIP, but will only do so with older 802.11a/b/g rates.

There are two common methods of handling the transition away from TKIP. Both methods use TKIP and CCMP simultaneously with the same set of network infrastructure, but they do so in slightly different ways.

Separate SSIDs

Most APs now support the ability to have multiple SSIDs on a single AP. One approach is to take the existing SSID that uses TKIP and duplicate it on the new 802.11n APs. The TKIP SSID works exactly as it did on the 802.11a/b/g network because the 802.11n APs are being used in a backward compatible mode. Users probably won't notice the transition unless they obsessively check the BSSIDs and note you have switched to a new vendor. In this method, the new 802.11n physical infrastructure runs two networks in parallel. Network administrators can monitor utilization of the TKIP SSID and encourage users with CCMP-enabled devices to switch, and assist users in upgrading to

CCMP-capable devices. At some point, the utilization of the TKIP SSID will be so low that it can be deactivated.

Simultaneous support of TKIP and CCMP

As an alternative, both encryption protocols can be run simultaneously on the same SSID, which is sometimes called *mixed mode* because it enables the network to support client devices with a mix of encryption technologies. When both TKIP and CCMP are supported, the network allows clients to choose the method they use to protect unicast frames. Broadcast and multicast frames must be received by all devices connected to the network, and in the case of mixed encryption must always be transmitted using a method supported by all clients—in this case, the lowest common security denominator of TKIP. It is comparatively simpler to use because there is only one SSID, but it has one serious drawback and one potential problem. The serious drawback is that because all multicast frames are transmitted using TKIP, it is not possible to use fast 802.11n data rates for any broadcast or multicast traffic. Due to the need for backward compatibility, all broadcast and multicast frames are transmitted using older 802.11a/g data rates. The potential problem is that simultaneous support of multiple cryptographic modes may not be compatible with some of the original TKIP implementations and may cause problems.

User Authentication

802.11n made no changes to the user authentication framework. Any user authentication system that works with 802.11a/b/g networks will also work with an 802.11n network.[5] EAP-based authentication is designed to work on top of many different physical layers, and therefore it does not require any change when moving to 802.11n. Connections between the wireless network and the user account system should not need to be redesigned.

Design Checklist

When planning a network, use the following checklist:

Client count, density, and mix
: To plan the network, figure out how many clients and estimate the traffic demand. A good rule of thumb is that an AP can serve 20-30 clients with acceptable service. Battery-operated portable devices such as tablets or phones will impose less of a traffic demand, but may require power-saving support. Decide between 3×3 MIMO and 2×2 MIMO systems based on what hardware clients have.

5. See Chapter 22 in 802.11 Wireless Networks: The Definitive Guide for a detailed discussion of building a user authentication system for your wireless LAN.

Applications
: Identify the key applications that must be supported on the network so that you can test them during installation and build a network that provides enough capacity along with appropriate quality of service support.

Backbone switching
: Upgrade to gigabit Ethernet, and check whether jumbo frame support is required.

Power requirements
: Ensure that you have supplied power to the AP mounting locations, either with power over Ethernet switches or power injectors, and ensure that they supply enough power to run your chosen AP hardware.

Security planning
: With 802.11n, TKIP has reached the end of its useful life. Before upgrading to 802.11n, it is worthwhile to consider moving an existing network to CCMP (WPA2) to avoid reconfiguring client devices. No changes will be needed to the user authentication implemented at the link layer.

CHAPTER 8

Designing and Installing an 802.11n Network

With planning out of the way, it's time to get down to the details of network design. Once you are secure in the knowledge of what your network must support, you can run through straightforward project planning to build it. Begin with understanding what it means to extend the access layer of the network out using a wireless link, and then select an appropriate product. Physical installation comes relatively late in the process, followed by basic network monitoring to make sure that the network is meeting your requirements.

Network Architecture for 802.11n

Throughout the evolution of wireless LAN technology, there have been a number of approaches to add the wireless LAN access layer on to an existing wired backbone network. Most approaches share two fundamental attributes:

MAC-layer mobility (also called "layer 2" mobility after the OSI model). 802.11 works by "hiding" the motion of a device from the network backbone. Even though a wireless LAN-equipped device is moving in space, from the perspective of routers and switches, it is remaining on the same broadcast segment out at the edge of the network. Within a single broadcast segment, mobility is easy. A decade ago, providing continuous connectivity across a routed (layer 3) boundary was difficult, but every commercially-available product suitable for use in a large-scale environment has addressed the "subnet roaming" question. Wireless products provide this capability within the MAC layer—as an 802.11 device moves around, it continues to send frames to its default gateway, regardless of where it has attached to the network.[1]

1. The alternative to MAC-layer mobility is to embed the mobility function within the network layer (layer 3 of the OSI model) with a protocol such as Mobile IP. Mobile IP was not widely adopted, in large part because of the extensive modification to the TCP/IP stack required on either the end-user devices or within the elements at the network edge.

802.1X security (also called "WPA2-Enterprise" after the Wi-Fi Alliance certification, or "layer 2 security" to contrast with IPsec or SSL VPNs). "Wireless security" has gone from being a laugh line in 2002 to something that is now taken for granted. Arguably, the typical wireless LAN with strong user authentication and encryption is now more secure than the wired backbone it connects to. In 2006, the Wi-Fi Alliance began requiring that all certified products implement version 2 of the Wi-Fi Protected Access specification (WPA2). Although there are many products that have not been certified, they often make use of components designed for use in certified products. WPA2 is stable, proven security, and is now widely used. WPA2-Enterprise offers network administrators the capability of designing network authentication around existing user databases, and extends the role information stored in those user databases out to the access layer. Typically, the "role" assigned to a user is enforced based on a combination of VLAN assignment at the edge, perhaps with additional IP filtering or QoS information.

Control plane location. In addition to protocol layering such as the familiar 7-layer ISO model, network technologies can be divided into *planes*. Each plane has its own protocol layers, of course, but the plane has a specialized purpose (Figure 8-1). Common planes are the *data plane*, *management plane*, and *control plane*:

Data plane
: Protocols in the data plane move bits from one location to another, and are concerned with moving frames from input interfaces to output interfaces. In an IP network, the main data plane protocols are TCP and IP, with applications such as HTTP riding on top of the network and transport layers.

Management plane
: The management plane provides protocols that allow network administrators to configure and monitor network elements. In an IP network, SNMP is a protocol in the management plane. A vendor's configuration application would also reside in the management plane; wireless LANs may use CAPWAP as a transport protocol in the management plane. Without exception, large scale IP networks use centralized management, and thus, have a centralized management plane.

Control plane
: The control plane helps make the network operate smoothly by changing the behavior of the data plane. An IP network uses routing protocols for control, while switched networks use the spanning tree protocol.[2] The control plane of a wireless LAN is responsible for ensuring mobility between access points, coordinating radio channel selection, and authenticating users, among other tasks.

2. In IP networks, many protocols used for network control use the same transport protocols as the data plane. For example, routing protocols communicate using IP, but also those routing protocols also influence how IP datagrams are routed. In the telephone network, the control messages to set up and tear down telephone calls use a protocol named Signaling System 7 (SS7) and travel over a completely separate network from the telephone calls themselves.

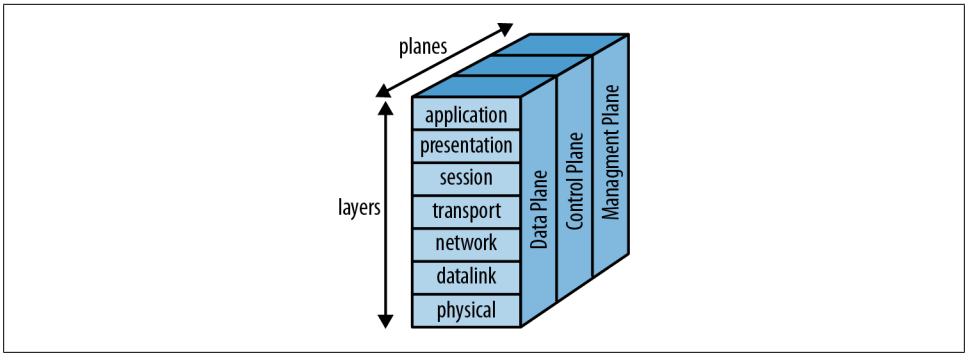


Figure 8-1. Network protocol planes

Wireless networks can be classified based on the location of the control plane, and much of the development across the history of wireless LANs is about refinements to the control plane. Consider the three basic architectures shown in Figure 8-2. When wireless LANs were first developed and network administrators were forced to deal with "stone age" products (roughly 1997-2003), building a wireless LAN was a significant undertaking. The simple explanation for the complexity of building a network is that adding the wireless access layer to the network meant that the number of network devices went through the roof. One network architect I worked with in the first half of the 2000s estimated that adding wireless access to the network had doubled the number of network elements he was responsible for. More formally, the network in Figure 8-2(a) consists of individual "autonomous" APs, and the management plane is distributed among all the APs. A few vendors selling autonomous access points did build centralized management tools, but most did not. More importantly, the control plane in a network of autonomous APs may not exist as a coherent whole, meaning that moving between access points was not guaranteed to work smoothly and there was little coordination between APs in making use of the radio environment.

The "iron age" of wireless LANs came with the emergence of wireless LAN controllers around 2003-4, and led to a design like Figure 8-2(b). The management plane was moved from the APs into controllers, and a centralized management application provided an overview of the entire wireless network. More importantly, the aptly named controllers brought the first coherent control plane functionality to the wireless LAN. Part of the reason for the development of controllers was that the processing horsepower needed for control plane functions was more affordable when it was provided in one location. In a typical controller-based deployment, the access points have limited functionality without the intelligence provided by the controller. Authenticating and authorizing users is handled by the controller, as are algorithms that provide RF management functions such as channel selection. Centralized management and control made it possible to build much larger networks, and most large scale networks that were built prior to the emergence of 802.11n were built using controllers. Early controllers also centralized the data plane as well by requiring that all traffic from APs be sent through the controller; this is often referred to as a *network overlay* because the

wireless network was separate from the existing core network and layered on top of it. In effect, the controller took on the role of a distribution switch for users attached to APs and provided mobility by serving as an anchor for the logical point of attachment.

With the emergence of 802.11n, traffic volumes increased dramatically. 802.11n increased the speed of each individual device, of course, but the high capacity of 802.11n also made it possible to use more mobile devices. The increased load could no longer be put through a single choke point on the network, and led to network designs like Figure 8-2(c), in which APs are directly connected to the network and traffic flows out from the AP on to the network. This approach is sometimes referred to as *distributed forwarding* because the data plane function has moved from the controller out to the AP. Although this architecture looks superficially similar to autonomous APs, it is typically paired with centralized management. More importantly, the availability of increased processing power has made the typical AP significantly more powerful than it was in the iron age. With increased processing power, it is possible to move the control plane out of a centralized location and distribute it through the network. Distributed AP deployments have varying control-plane functionality based on the vendor; typical control-plane functions that may be available in a distributed access point include radio tuning, tunneling to provide mobility support to end-user devices, security functions, and evenly distributing the user load across APs.

Temporary Networks

Temporary networks are set up for many reasons. Typically, they are built for an event, such as a conference or trade show that doesn't justify the cost of permanently installing cable. Temporary networks can range in size dramatically depending on the reason for the setup. At the micro-network extreme, many business travelers pack an AP when they travel, and set up a temporary network in the hotel rooms they use. At the larger end of the temporary network, a wireless LAN set up for a conference or trade show may cover very large buildings with hundreds of APs for just a few days. Some temporary networks may be set up as an offshot of a larger network. For example, auditors carrying a "remote access" AP that connects a wireless LAN back to a corporate network through a VPN.

Architecture Comparison

Building a "micro-network" of a AP or two is easy. With a small number of APs, it is acceptable to manage APs individually. Upgrading to 802.11n is also straightforward: take out your existing 802.11a/b/g APs and replace them with 802.11n APs. At such a small scale, almost anything will work. That is, unless your micro-network is actually one small piece of a larger network, in which case you should skip ahead a couple of sections and start reading. At some point, the overhead of managing individual devices will be too great, at which point you are building a small- or medium-size network. These networks have just as much to gain from 802.11n. Building a small network of

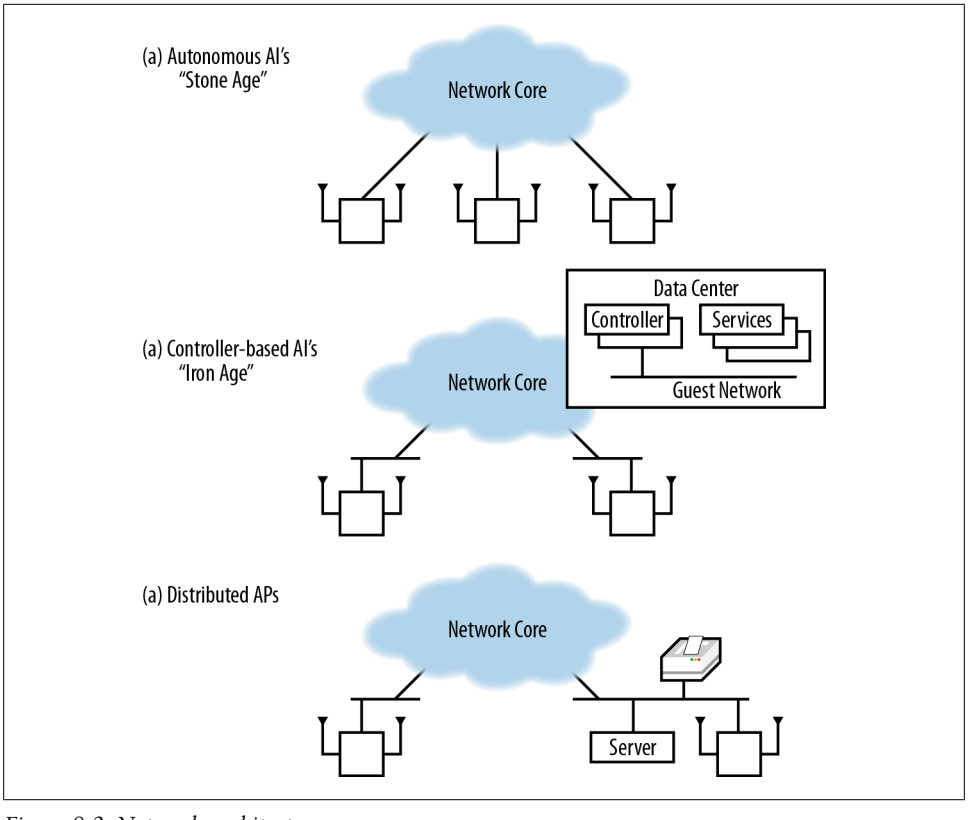


Figure 8-2. Network architectures

five to ten APs is too small to use the "big iron" of controllers, but it is large enough to require centralized management.

Prior to the widespread use of 802.11n, large networks needed a centralized control plane to handle the loads imposed by large numbers of users, and the choice between autonomous APs and controller-based APs was a straightforward one that was almost always resolved in favor of the more advanced centralized control plane. 802.11n enabled a change in user attitudes towards connectivity that has driven the development of distributed control planes. As wireless LANs became viewed as stable, many new devices are being built without Ethernet—the MacBook Air was one of the first examples, and the number and variety of tablet devices is another. With users switching towards mobile battery-operated devices that they can use on the go, a significant portion of the network access layer has moved from traditional wired Ethernet to wireless networks. With the explosion of 802.11 devices now available, network architects have designed higher- and higher-capacity networks, stressing the centralized control plane. Early controller-based networks were able to use a single controller as the focal point for both the control and the data plane, but that assumption no longer holds.

Table 8-1 compares the three basic architectures. In reality, there is some overlap between these architectures when they are implemented in products. It is likely that a large-scale network at any speed, especially one supporting critical applications, will require some degree of decentralization, either by moving some of the data plane functions to the edge of the network, some of the control plane functions to the edge of the network, or both. All three architectures are capable of supporting any set of network requirements, but the cost and availability of the resulting network may vary.

Table 8-1. Architecture comparison

Attribute	Autonomous APs	Controller-based APs	Distributed APs
Location of data plane	Distributed, enabling high network performance	Centralized, potentially limiting performance to the forwarding capacity of a controller. Good mobility support because devices attach through the controller.	Distributed, enabling high network performance. Many products have features to assist with mobility.
Location of management plane	Depends on product; often distributed, imposing very high staff costs	Centralized, lowering operational expenses	Depends on product; often centralized, enabling lower operational expenses
Location of control plane	Distributed, if it exists. Nonexistent control plane limits flexibility of security and radio management.	Centralized, with high functionality for radio management and user management.	Distributed. Functionality of control plane depends on vendor implementation.

Management plane

If you are building a large-scale network, there is nothing to consider regarding the management plane. You need centralized management—otherwise, maintaining configuration across multiple devices will quickly break down. With early wireless LANs, there were some products that didn't have centralized management, but those days are thankfully long behind us. Any 802.11n product designed for use in a network beyond home-size scales has central management, though there are of course large differences in functionality and cost. If you have a relatively small network, the software-as-a-service model may offer you the ability to use a full-featured management system.

Centralized management is non-negotiable beyond just a few access points.

Data plane

This section has presented each of the three architectures with sharp lines between them. In actual products, the location of data forwarding is one of the places where products offer flexibility. Many controller-based APs can be configured to send data

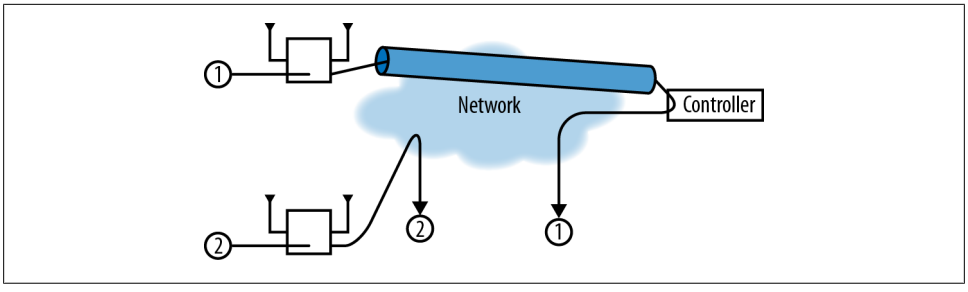


Figure 8-3. Jumbo frame support

either through a centralized controller or directly from the AP to the network. Some common hybrid approaches are to offer a choice on a per-SSID basis so that guest traffic can be centrally forwarded, or that traffic from devices attached to the AP's VLAN can be forwarded by the AP while traffic bound for VLANs attached to other APs must be send through the controller's forwarding engine. Likewise, some distributed data planes offer roaming capabilities that can make any VLAN accessible throughout the network by tunneling between APs.

When a centralized data plane is used, 802.11n APs may also benefit from jumbo frame support and path MTU discovery. In the centralized forwarding model, illustrated by the top AP in Figure 8-3, data frames from client devices are transmitted from the AP to the controller through a tunnel that traverses the network backbone. For a transmitted frame to be sent to its network destination, it is received by the AP, passed to the controller through the tunnel, and then placed on the network by the controller.

For efficiency, client devices generally send maximum-length Ethernet frames of 1,500 bytes. Unless the path between the AP and controller can handle jumbo frames, it will have to fragment client traffic, resulting in a 1,500-byte frame that contains most of the client's frame, followed by a second frame that has the remainder of the client's data. There are four main ways to cope with the potential for fragmentation in a controller-based architecture:

Jumbo frame support
: If the path between the AP and controller supports jumbo frames, nothing happens. The AP takes the client's data frame, puts a tunnel header on it, and sends the slightly larger tunneled frame across the network. There is an theoretical (and, in practice, imperceptible) penalty in speed because the tunneled frame is slightly larger, but the overhead is so minimal that I doubt any real-life network has measured the performance impact.

IP stack fragmentation
: IP itself can fragment frames. When the client data frame comes to the AP, it is given a tunnel header, and the resulting packet for transmission to the controller exceeds 1,500 bytes. As a result, the IP stack will fragment the packet into a maximum-size packet plus a fragment. When both the first packet and its trailing

fragment arrive at the controller, they are reassembled into a single IP packet, the tunnel header is removed, and the client's data frame is processed. Using IP-level fragmentation takes advantage of existing fragmentation code that is well-understood, but it does impose a cost on the controller in terms of packet buffer memory and processor load; the backbone must also forward two packets instead of just one. Depending on the network architecture, the tunnel may also break; most firewalls drop all IP fragments without further analysis.

Controller tunnel (CAPWAP) fragmentation
: As an alternative to using fragmentation and reassembly in the IP stack, the tunneling protocol between the AP and controller can implement its own fragmentation protocol. In fact, the Control and Provisioning of Wireless Access Points (CAPWAP) protocol, specified in RFC 5415, specifies its own fragmentation layer (see section 3.4). Fragmentation at the tunneling protocol improves firewall traversal because most firewalls do not perform detailed analysis of traffic beyond IP headers.

TCP Maximum Segment Size control
: The initial setup of a TCP connection includes a *Maximum Segment Size* (MSS). Some APs can rewrite the MSS value sent by clients so that the clients send frames that are small enough that the client frame plus the tunnel header is exactly the maximum size packet for the AP-to-backbone connection.

Additionally, there is one additional way to connect APs to the backbone network to avoid fragmentation concerns. Instead of using a tunneling protocol to reach a central point of attachment, some APs can connect directly to the network edge and forward traffic through the edge switch, as illustrated by the bottom AP. APs that support direct connection to the network do not require jumbo frame support.

Many commercial products on the market offer a combination of the approaches in Figure 8-3. Some controller-based APs can be configured to send data either through a centralized controller or directly from the AP to the network. Some common hybrid approaches are to offer a choice on a per-SSID basis so that guest traffic can be centrally forwarded, or that traffic from devices attached to the AP's VLAN can be forwarded by the AP while traffic bound for VLANs attached to other APs must be send through the controller's forwarding engine. Likewise, some distributed data planes offer roaming capabilities that can make any VLAN accessible throughout the network by tunneling between APs.

With an 802.11n network of any size, it is likely you need to use both distributed data forwarding as well as the ability to forward wireless traffic across broadcast domain boundaries.

Control plane

The location of the control plane is the source of most architectural differentiation in available products. Centralized control plane technologies have been around for a long period of time, and are supported by mature, tested software, and can be combined with a distributed data plane for efficient use of the core network. Many centralized control planes are in the process of moving towards either a split control plane (where functions are shared between the controller and APs) or a more fully distributed control plane, but it is common that some features are not available when the controller is removed from the network. Distributed control planes can be cheaper, especially when designing for distributed networks with many remote sites.

The location of the control plane may have an effect on the overall reliability and resiliency of the network, which may be important depending on the applications in use and the structure of the network. For example, hospitals are increasingly turning to wireless LANs to support critical applications that distribute medical records, stream real-time data from patient care devices, fill prescriptions, and even make remote doctors available.[3]

To guard against downtime, it is necessary to build a redundant control plane in addition to a redundant data plane. Even with overlapping AP coverage, the network must enable client devices to quickly switch between APs. Simply handing off the wireless connection is easy; the difficulty in moving devices between APs is ensuring that QoS state, security keys, and firewall policy state also move between APs. Neither the distributed or centralized type of control plane is inherently more resilient; a distributed control plane protocol can be resilient by design, while a centralized control plane may require spare controllers.

Carefully evaluate the trade-offs involved in a centralized versus a distributed control plane, and ensure that the control plane you select meets your needs for functionality, resiliency, and cost.

802.11n Hardware

After reviewing network requirements from the previous chapter and deciding on what constraints drive the logical architecture, it's time to pick out access point hardware. Access points all perform the same basic function in that they shuttle frames between radio networks and Ethernet, but there can be tremendous differences in cost and functionality. Comparing access points on the basis of price alone is like saying that a

3. Broadly speaking, the ability to involve remote medical professionals is called "telemedicine," and is enabled by videoconferencing plus a variety of healthcare information technology, supported by a robust network infrastructure. As you might expect, it is far too broad a topic to be discussed at any length in this book—but the supporting technology for that "last hop" to the handheld device is almost certainly going to be 802.11.

Chevy Aveo and a Mazda Miata are the same because they are both cars that get you between two points.[4] To build a network of more than just a handful of access points, you probably want to look beyond the hardware available at electronics stores and at highly functional APs. Even a small network can benefit from corporate-grade APs—if you are running the network for a hedge fund, network performance is critical.

Wi-Fi Alliance certification
: A basic requirement is demonstrated interoperability with the Wi-Fi Alliance's test suite. Certification is not an absolute guarantee of interoperability, but it is an obvious first step. To check on the certification status of a product, visit the Wi-Fi Alliance web site and click on the "Wi-Fi CERTIFIED Products" button on the left-hand side of the page. How to decipher the full certification information is discussed in the next section, Technology "Waves."

High performance
: At the relatively low speeds of 802.11a/b/g, the demand for CPU power to run at "air rate" was relatively low. As speeds increase to the current state-of-the-art 450 Mbps, significantly more processing power is required to forward data, and the effects of tuning the software for speed will be more pronounced. A 5% increase in data rate at 802.11b rates is likely only half a megabit, but a 5% increase at 3-stream 802.11n speeds is in excess of 20 Mbps. Vendors of corporate-grade hardware invest much more heavily in software tuning because their products are sold where every bit per second matters.

Hardware quality and robustness
: Corporate-grade devices are designed to be used for many years before replacement, and therefore, are often designed with future expandability in mind. Components are selected based with a view towards quality and long life instead of primarily based on cost. Sophisticated antennas or other radio front-end components may be used to improve the quality of the network, either in terms of throughput or coverage. Radios will be enabled on all available channels even though the cost of regulatory compliance with DFS channels can be substantial, and software supports automatic configuration of radio channel selection.

Software functionality, upgradability, and quality
: Generally speaking, more expensive devices have significantly more functionality, with advanced features in several areas. Vendors regularly plan for the release of new features, and it is common for new features to be provided midway through a product's life cycle. Additionally, extensive QA testing is used to ensure that corporate-grade devices can be run for months at a time under heavy loads.

Antenna options
: Internal antennas allow an AP to be self-contained, and to blend smoothly into the aesthetic environment. External antennas typically have higher gain, which

4. This is funniest take on the Chevrolet Aveo that I have ever read: *http://www.ginandtacos.com/2009/12/18/npf-ed-drives-tiny-car-hilarity-ensues/*.

improves range. In a deployment based on area coverage instead of density, or a deployment in a challenging radio environment, selecting the right external antenna can be the difference between a poor-quality network and a successful one. External antennas are also frequently used for outdoor deployments. Picking the right external antenna is still something of an art, and the antenna must be matched to the performance characteristics of the AP. A high-gain antenna will dramatically increase the transmit range of an AP, but if the AP has low receive sensitivity, the high-gain antenna will cause more problems than it solves.

Power options

Consumer-grade devices are typically powered with a "wall wart" transformer and must be installed close to existing electrical outlets, while corporate-grade devices can draw power from the device at the other end of the Ethernet cable. Power over Ethernet enables placement of devices in out-of-the way locations, and can be used to provide power even on very high ceilings.

Security

Security is not just about providing solid encryption, though that is the obvious starting point. Corporate-grade products offer flexible authentication through RADIUS and directory interfaces, per-user VLAN mapping, traffic filtering and queuing, and built-in captive web portals for web-based authentication.

Quality of service

At the most basic level, quality of service support is compliance with the Wi-Fi Multi-media (WMM) certification requirements, which divides traffic on the air into four classes of differing priority. More complex queuing systems can be used to improve service quality for voice devices, or to ensure that airtime is balanced fairly between network users.

Manageability

If you are reading this book, you need centralized management. Evaluate management tools for a wireless network in the way that you evaluate management tools for a wired network. Ensure that the management software provides something beyond simple configuration management and can report on the overall state of the network.

Modular Access Point Design

Unlike the PHYs that came before it, 802.11n came with a "roadmap." Once the basic MIMO standard was agreed on, it was straightforward to lay out how the two-stream design would be extended to three and four streams. Product vendors had two choices in designing early products. The approach taken by most was to design a monolithic unit. A single self-contained AP has first-generation 802.11n technology, typically 2-stream MIMO. When future generations of 802.11n technology are introduced to the market, the self-contained early 2-stream AP is replaced by a 3-stream AP.

An alternative approach is to make a modular unit in which "radio cards" can be replaced. While these are often touted as future-proof devices, they are far from being

future proof. An AP "chassis" that can accept radio cards is expensive to produce, and the AP base unit needs to hold the CPU and memory. When designing the AP base, the vendor needs to deliberately oversize the CPU and memory specifications, and the increased cost translates directly into an increased price. The CPU must be just fast enough for the future software. Too little CPU, and the whole purpose of the extra cost of the modular unit is undermined; too much CPU, and the cost is higher than it needs to be. Furthermore, upgrading a network is not just a matter of the cost of parts. The real cost is that somebody needs to get on a ladder to replace APs, and that cost is the same whether you are swapping out radio cards in a modular unit or swapping out fixed-configuration APs.

Typically, purchasers of modular APs find that by the time they are ready to change modules, newer fixed-configuration APs often cost less for more capabilities.

Technology "Waves"

802.11n is a complex standard with interdependent pieces. Previous PHYs were able to come to market all in once piece, so that 802.11g did not exist, and then it did, fully formed. The complexity of 802.11n means that it has come to market in distinct "waves" or "phases." Part of the reason for phases is that the difficulty in standardization is determining what the frame format should be, and what is eventually possible to build into hardware. However, the incremental work in adding 4-stream transmission to 802.11n is very small compared to the hardware engineering needed to put four radio chains in close proximity.

One way of thinking about the market is to think of it in technology waves, where a wave is a set of products that hit the market at about the same time. When building wireless LAN products, the most important component is the radio chipset, a set of silicon chips that implement the components shown in Figure 3-7. Typically, the amplifiers are put on a *radio front-end* card, which allows them to be customized to a particular application. Long-distance transmission or high rate-over-range can be achieved by using higher-quality radio front-ends. The radio chip itself will have a PHY section that demodulates data, possibly from several data streams, and corrects any errors by using the error-correcting code. The resulting bit stream is passed to the MAC where security is applied, any frame deaggregation is done, and so on. APs are built around radio chips, so the technology waves in APs are driven by new radio chipsets.

In 2006, the 802.11n market was tiny, and the draft 802.11n standard was only in its first draft.[5] Several areas of the standard had different interpretations, and radio chipsets for 802.11n were maturing along with the draft standard itself. Many of the earliest

5. The 802.11 timeline site is a fabulous resource for the progression of standards through the drafting process. 802.11n was being feverishly assembled in early 2006. The final votes to create draft 1.0 occurred at the March 2006 meeting in Denver, and the draft was balloted in April. The significantly improved draft 2.0 was not balloted until March 2007.

chipsets only supported operation in 2.4 GHz, and the only way to guarantee interoperability was to buy products that used the same chipset vendor. In retrospect, I labeled the 2006 time frame as the *zeroth wave* because there wasn't really a large market.

The first wave came a year later, with the second draft of 802.11n. After the 12,000 comments received against the first draft, the second draft was much expanded and refined. Due to the perception of non-interoperability, the Wi-Fi Alliance began certifying products against the developing standard, an action which helped focus the standards developers on working out the major contributors to interoperability. The draft 2.0 certification program cemented the major components of the developing standard in place, and created a market where interoperability was tested and became a major focus of product vendors. The first wave was when 5 GHz products came to market, as well as when 802.11n enterprise devices emerged, most of which used a 3x3 MIMO design that supported two spatial streams. Power consumption of these APs was quite high, and often exceeded the 802.3af power budget.

The second wave of products was driven by the ratification of the 802.11n standard in 2009. Customer feedback on the first wave of enterprise devices was that full functionality was required on the 802.3af power budget, and the industry responded. Second-generation 802.11n hardware from chipmakers were 2×2 MIMO designs supporting two spatial streams. Removing the third radio chain and using better chip manufacturing techniques meant that the second-generation of 802.11n enterprise devices operated at similar speeds for the end users, but with much lower power consumption.

The third wave of products in the enterprise came in 2011, as 3×3 MIMO designs supporting three spatial streams came to market. At the time this book was written, high-end APs were all three-stream designs. It is likely, though not definite, that most chip designers will focus their four-stream efforts on 802.11ac, the forthcoming gigabit standard.

Wi-Fi Alliance Certification

The Wi-Fi Alliance is an industry association of companies that are driving wireless LAN technology. The Alliance is best known for the Wi-Fi CERTIFIED interoperability testing program that began in 2000. Prior to 802.11n, the underlying PHYs were simple enough that there was a single set of core capabilities and very few optional features. The statement that "this device supports 802.11a" meant the same thing for any 802.11a device. With 802.11n, however, the underlying technology is much more complex. 802.11n devices certified for interoperability by the Wi-Fi Alliance are given the designation *Wi-Fi CERTIFIED n*, and are required to meet a minimum set of 802.11n technology components.

Once certification is complete and a product is awarded certification, it can be looked up at the Wi-Fi Alliance certified product listing. Each product is also given an *interoperability certificate* that details the individual product features that have been certified.

Figure 8-4 shows a sample interoperability certificate, in this case, for an Aerohive HiveAP 330. In the upper-left hand corner of the interoperability certificate, the Wi-Fi CERTIFIED logo shows the basic interoperability certification (in this case, 802.11a/b/g/n); these logos often appear on product literature and packaging. Below the certification logo, there is a box noting the number of spatial streams supported by the device (in this case, three). As a point of contrast, Figure 8-5 shows a certificate for a client device (in this case, for a Dell card used as a reference design). This card has support for the greenfield preamble and wide channels in the 2.4 GHz band, both of which are displayed in the left-hand column.

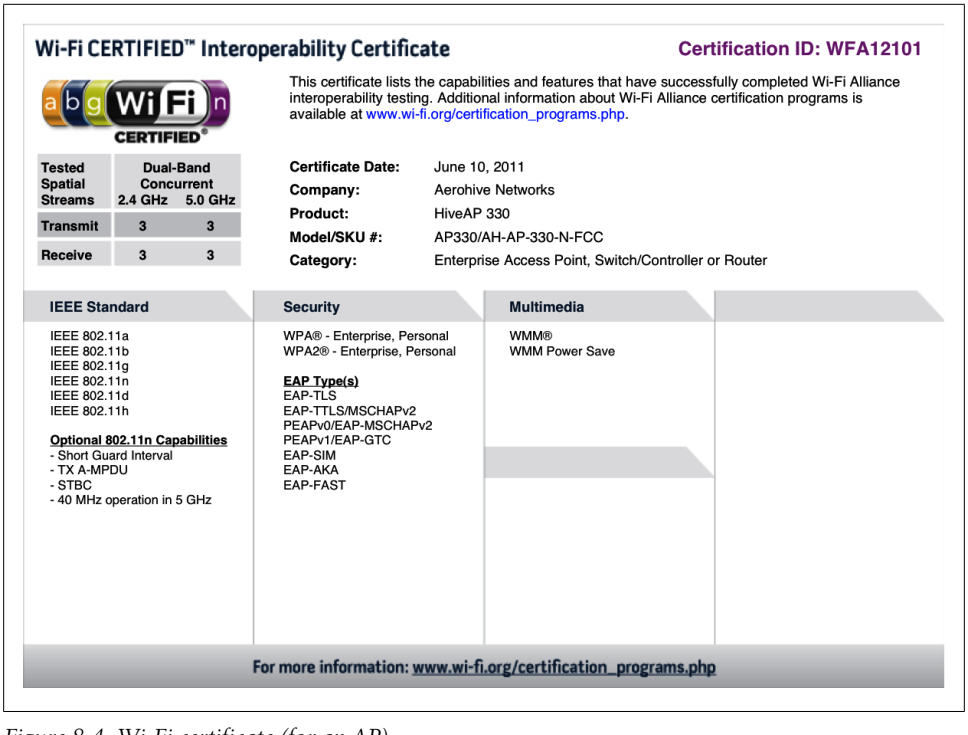


Figure 8-4. Wi-Fi certificate (for an AP)

Mandatory tests

Every device submitted for 802.11n certification must pass a series of basic tests that are expected to be supported by every 802.11n device. These features include:

Minimum number of spatial streams

APs must support at least two streams before being allowed to use the 802.11n logo. No such rule applies to client devices. The number of declared spatial streams, up to 3, is tested and placed on the interoperability certificate. The HiveAP 330

Wi-Fi CERTIFIED™ Interoperability Certificate — **Certification ID: WFA8884**

a b g Wi Fi n CERTIFIED®

This certificate lists the capabilities and features that have successfully completed Wi-Fi Alliance interoperability testing. Additional information about Wi-Fi Alliance certification programs is available at www.wi-fi.org/certification_programs.php.

Tested Spatial Streams	Dual-Band Selectable 2.4 GHz	5.0 GHz
Transmit	2	2
Receive	2	2

Certificate Date: March 18, 2010
Company: Dell
Product: DW1520 Wireless-N WLAN Half-Mini Card
Model/SKU #: DW1520/
Category: Reference Design

IEEE Standard

IEEE 802.11a
IEEE 802.11b
IEEE 802.11g
IEEE 802.11n
IEEE 802.11h

Optional 802.11n Capabilities
- Short Guard Interval
- Greenfield Preamble
- TX A-MPDU
- STBC
- 40 MHz operation in 2.4 GHz with coexistence mechanisms
- 40 MHz operation in 5 GHz

Security

WPA® - Enterprise, Personal
WPA2® - Enterprise, Personal

EAP Type(s)
EAP-TLS
EAP-TTLS/MSCHAPv2
PEAPv0/EAP-MSCHAPv2
PEAPv1/EAP-GTC
EAP-SIM
EAP-AKA
EAP-FAST

Additional Vendor EAP Type(s)
EAP-TLS
EAP-TTLS/MSCHAPv2
PEAPv0/EAP-MSCHAPv2
PEAPv1/EAP-GTC
EAP-SIM
EAP-AKA
EAP-FAST

Multimedia

WMM®
WMM Power Save

Special Features

Wi-Fi Protected Setup™
- PIN
- PBC

For more information: www.wi-fi.org/certification_programs.php

Figure 8-5. Wi-Fi certificate (for a client)

certificate indicates that it can support 3-stream operation for both transmission and reception in both radio bands.

A-MPDU reception
: Any Wi-Fi CERTIFIED n device must be able to receive A-MPDU frames. A-MPDU support is typically provided within the 802.11n chip itself, so this support for option is widespread throughout the history of 802.11n. Devices under test are allowed to self-describe the A-MPDU size supported, so it is impossible to determine whether the A-MPDU implementation supports a high density of back-to-back MPDUs or something less capable.

A-MSDU reception
: In addition to the A-MPDU aggregation, devices must support A-MSDU reception as well before being certified.

Overlapping BSS (both bands)
: The overlapping BSS test ensures that an 802.11n network whose coverage are overlaps with an 802.11a/b/g network does not cause interference to the legacy network.

Reduced Interframe Space (RIFS) support
: Tests validate the use of the RIFS for efficiency. At the start of the Wi-Fi CERTIFIED n program, this capability was not required, but it has been required since the middle of 2008.

Security: TKIP & WEP negative tests
: 802.11n devices may not use TKIP or WEP to protect frames sent at 802.11n data rates. The certification program includes "negative tests," which are tests to ensure that WEP and TKIP cannot be used with 802.11n data rates. Many products implement data rate limits when WEP or TKIP are configured, so that if an 802.11n network is configured for TKIP, its components will avoid using data rates higher than 54 Mbps.

Optional tests

In addition to the mandatory tests described in the previous section, the certification program includes a number of optional capabilities, each of which is called out on the interoperability certificate.

40 MHz operation in 5 GHz
: The use of wide channels in 5 GHz enables greater throughput. This option is commonly supported by enterprise devices intended for large-scale high-throughput environments. At the time this book was written, slightly over three-quarters of enterprise networking devices support this option.

40 MHz operation in 2.4 GHz (including coexistence)
: As with the previous option, this item indicates support for 40 MHz channels in the 2.4 GHz band, with a catch. It is not possible to implement 40 MHz channels in the 2.4 GHz band and pass the test unless the 20/40 MHz coexistence features described in Chapter 6 are also implemented. Because spectrum is scarce in the 2.4 GHz band and 40 MHz channels upset the existing 3-channel plan, this option is not widely certified. Only about a quarter of Wi-Fi CERTIFIED n devices implement this feature.

Short Guard Interval
: Support for the short guard interval is widespread. It increases throughput by about 10%, and is widely supported by the chipsets that power wireless LAN devices. Almost three-quarters of all Wi-Fi CERTIFIED n devices support the short guard interval, with an even higher fraction of enterprise networking devices including support.

Space-Time Block Coding (STBC)
: STBC improves the ability of a signal to travel farther because it uses all of the MIMO signal processing gains to increase range. STBC is not widely implemented, and is certified in less than a fifth of Wi-Fi CERTIFIED n devices.

Transmission of A-MPDUs
: Tests that the device supports sending A-MPDUs. This is the only aggregation test; the certification testing does not validate A-MSDU behavior. About half of certified devices support this feature.

Greenfield mode
: Greenfield mode improves efficiency at the cost of easy interoperability with 802.11a and 802.11g networks. Greenfield mode is supported in about a quarter of all certified devices, but it is extremely rare in enterprise-class APs.

HT Duplicate Mode (MCS 32)
: HT Duplicate mode improves error rate performance. It consists of one mode of operation at 6 Mbps that sends the identical data stream on both halves of a 40 MHz channel. It is implemented by less than a fifth of certified devices, and not discussed in detail in this book.

Access points that implement 802.11n with a single spatial stream are not allowed to claim that they are certified for 802.11n. As a compromise, however, APs that pass any test in this section can be certified and use the logo that indicates "with some n features." For example, a device that implemented single-stream 802.11n and supported 40 MHz channels could be labeled as "Wi-Fi CERTIFIED 802.11a/b/g with some n features."

Untested Features in 802.11n Certification

Some protocol features are not included by the Wi-Fi Alliance interoperability testing for 802.11n devices. Most notably, four-stream operation and beamforming are not part of the Wi-Fi CERTIFIED n test plan. Wi-Fi Alliance tests are built by writing a test plan of the most common features and assembling a "testbed" of devices that implement those features. For maximum assurance of interoperability, the testbed is composed of products from several different vendors so that both the hardware and software are independent implementations of the same specification. As this book was being written at the end of 2011, support for both four-stream operation and beamforming did not have wide support across the industry.

Coverage and Capacity Planning

When 802.11 first emerged, radio planning was a complex undertaking because there was almost a master craftsman's touch to walking a building, estimating coverage areas, and assigning non-overlapping channels to maximize throughput. Those days are long gone. Any wireless LAN system that works at the scale of more than a few APs has an automatic radio tuning system built in. APs continuously assess channel quality on a basket of factors, and choose the "best" channel according to an algorithm that typically uses the number of overlapping APs on each channel, the signal strength of those overlapping APs, and a channel loading figure.

With channel selection handed over to the wireless LAN itself, network administrators can focus on the higher-level questions of ensuring that the network is built to meet its desired goals. Rather than delving into the minutiae of wringing marginal performance gains out of a slightly improved channel map, administrators can spend time on the bigger questions to build the maximum network capacity within a given budget. Generally speaking, the service side of wireless LAN installation is about where to put APs and the assignment of channels to those APs; the cost side is not just about the equipment, but also includes the cost of installation.

Wireless LAN network capacity comes from huge tracts of clean spectrum. 802.11n does not define usage in a new band of spectrum. Rather, it coexists in the 2.4 GHz band with 802.11b/g and in the 5 GHz band with 802.11a. As wireless networks have become accepted (or even loved), the constraints on the 2.4 GHz band and its three channels have become more pronounced. Except for all but the most basic networks, a wireless LAN needs the extra capacity available in the 5 GHz band. The 5 GHz band provides more channels, plus the capability to build smaller coverage cells for better frequency reuse.

Unlike previous physical layers, 802.11n is capable of using both the 2.4 GHz ("802.11ng") and 5 GHz ("802.11na") radio bands. For high capacity, you will need to use both of them at the same time.

The number of APs needed to build a network is determined by the desired performance of the network as a system. Pick a device as the worst-case performance target; the target device will often be a single-stream device such as a phone or tablet. Estimate the number of devices on the network and the required system-wide performance to get a "capacity" target AP count. Next, estimate the coverage area of an AP based on the required signal strength to provide adequate coverage. Many single-stream devices will require that the network be designed around a signal strength of -67 dBm. Using a planning tool, get a "coverage" target AP count. Obviously, pick the larger of the two numbers. In many environments, the AP layout will be a blend between coverage and capacity, with high-density areas getting a closely packed AP deployment while areas with lower user densities will use the more sparse coverage plan. As you set up capacity plans for a network, allow for sufficient capacity so that the network will meet user demands into the future. A variety of tools exist that can assist with network planning, and a few vendors offer free estimates of either the AP count or mounting locations.

AP Mounting Locations

If your 802.11n network is an upgrade from an existing 802.11a/b/g network, a good first step in getting maximum performance for a given cost is to reuse existing mounting locations. Cabling is often one of the major costs of building a wireless LAN, and

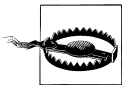
reusing existing cabling is a good first step, especially if the cabling is in place to many of the locations where you would want to put wireless LAN coverage.

802.11n APs might show a small increase in speed by moving to a new mounting location, but the cost involved in resurveying and recabling new locations is almost always prohibitive. From a cost efficiency standpoint, it is cheapest to install 802.11n APs in the same locations as the existing 802.11a/b/g APs, and then add extra capacity where it is required. Typically, 802.11n offers the highest benefits in high-density/high-capacity locations such as auditoriums, conference rooms, and other areas where large numbers of users tend to gather. As demands on the network grow, additional capacity can be added as network usage increases.

When building a new 802.11n network, especially for coverage, it is best to design the network around the needs of 5 GHz coverage. Because radio range at 5 GHz is somewhat shorter, and the 5 GHz band bears the brunt of building a high-capacity network, it is best to lay out a new network with mounting locations selected to obtain desired 5 GHz coverage. Only once that is complete, enable radios on the 2.4 GHz band. Due to the longer reach at the lower-frequency band, only a subset of APs will need to have their 2.4 GHz radios active.

Channel Types and Layout

The easiest component of channel layout is the channel assignments. Just let your network pick channels itself. With rare exceptions, automatic channel selection algorithms converge on a good enough channel plan that the gain in speed from picking channels manually is quite low and not worth the time.[6] The important inputs to channel selection algorithms are the available channels and the capacity of each channel. 802.11n offers two channel widths: a backwards-compatible 20 MHz channel that is identical to 802.11a/g, and a new 40 MHz wide channel that offers slightly more than double the speed.



Do not use 40 MHz channels in the 2.4 GHz band. With only 83 MHz of spectrum, you have space for just two, and that's only if you ignore the overlap problems. Just don't do it.

Practically speaking, an extensive deployment of 40 MHz channels will need support for the worldwide harmonized radio band (channels 100 to 144 in Figure 3-3). Using these channels requires that the AP support Dynamic Frequency Selection (DFS). DFS capabilities are required by radio regulators in each individual country, and support is tested as part of the government certification process required to sell radio devices.

6. Most of the exceptions involve the presence of persistent non-802.11 interference that dramatically reduces the capacity of a channel. However, this can be detected by its effects on throughput of attached devices, a combination of high signal strength with low data rates, or by spectrum analysis capabilities that are built into the AP itself.

Without DFS, wireless networks in Europe are restricted to four indoor 40 MHz channels and wireless networks in the U.S. are limited to six. A key enabling technology for DFS operation is the IEEE 802.11h amendment, which is tested by the Wi-Fi Alliance and appears on the interoperability certificate as shown in Figure 8-4.[7]

DFS support enables outdoor operations in Europe, as well as adding five additional 40 MHz channels.

40 MHz channels reduce the number of allowed channels by half. In order to have a reasonable channel layout, you need 4 to 6 available channels, which means that DFS certification and operation are a practical requirement.

Although DFS operation improves the number of channels, it is not perfect. Spectrum is the lifeblood of wireless LANs, and many channels are available on the condition that wireless LAN operation is secondary to primary uses of the band. Operation on channels where another use takes priority, such as weather radar on channels 100 to 144, requires that APs monitor for higher-priority transmissions and switch channels to avoid causing interference. Getting DFS right is tricky because an AP must cede the channel to a higher-priority use, but if it is too sensitive the resulting channel changes will result in unwanted network disruption. The only remedy for a wireless network administrator is to carefully assess whether channels that require DFS are needed to meet the goals of the network and carefully monitor the network in operation to ensure that any channel changes are the result of actual interference and not merely false positives.

AP Transmit Power Setting

In addition to dynamic channel selection, many APs have the ability to automatically adjust transmit power through a function called Transmit Power Control (TPC). Transmit power adjustments are often described as a way to fill in coverage holes or otherwise help a network to "self-heal" in response to changes in the radio environment. TPC, however, can easily create asymmetric links, as shown in Figure 8-6. In Figure 8-6(a), the client device power is set much higher than the AP transmit power. Although the AP network has optimized itself for high coverage density with relatively low transmission power, a client device set for much higher transmission power may cause the carrier sense in a large surrounding area to read busy. Figure 8-6(b) shows the opposite problem, in which the AP transmit power is set very high and a low-power client is no longer able to reach the AP.

7. If you need lots of wide channels, check on the channels supported by your AP. At one point, the FCC suspended DFS certification testing for quite some time, leaving many products technically capable of supporting the DFS band unable to do so for legal reasons.

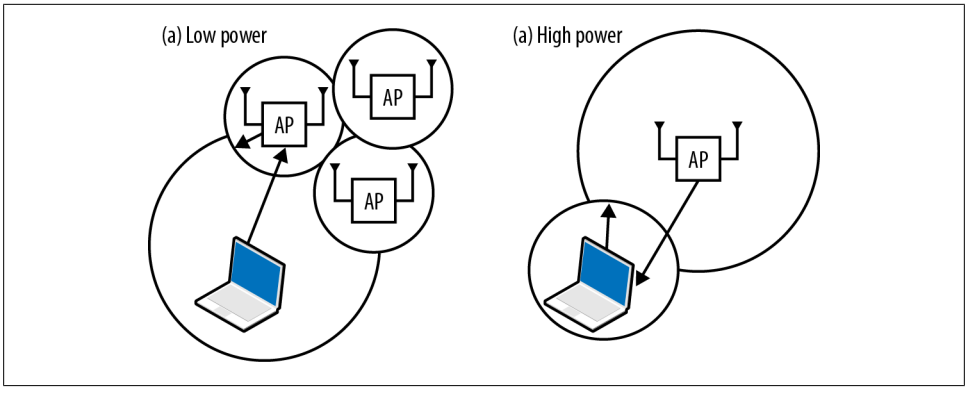


Figure 8-6. Transmit Power Control and link asymmetry

The key for network administrators in configuring TPC is to guard against both types of link asymmetry. Begin by determining the maximum transmit power of mobile devices, and capping the transmission power of each AP at the level of the typical client. Laptops typically transmit at a maximum power of 100 mW (20 dBm), but may be set lower to conserve battery power. Phones are often capable of the same power level, but higher transmit power does decrease battery life. A good compromise is to set the network for a maximum power of 30-50 mW (roughly 15-19 dBm).

Network Management

When problems arise after installing the network, administrators turn to analysis tools to diagnose the problem, find the root cause, and determine what configuration steps can be taken to eliminating the problem. Wireless networks are by no means trouble-free, but the underlying technology is now fairly mature. Network management applications play a critical role in network operations by automating the analysis and troubleshooting of problems. Good network management tools can help you check assumptions on the client device and traffic mix, as well as react to change in the usage of the network over time.

Network Analysis

In the early days of wireless LANs, just getting a packet capture could be a challenge. Thankfully, those days are behind us, and many software products exist to easily obtain a packet trace. For quick analysis, I use Wireshark because it is a cross-platform tool that is equally available on MacOS, Linux, and Windows, and my long history with tcpdump has resulted in the ability to write filter rules using the familiar Berkeley Packet Filter (BPF) syntax. Installation is straightforward, with ready-to-run code available easily.

Wireshark requires system-level privileges to read the packet capture. On MacOS and Linux, make sure that the BPF devices are readable by the user running wireshark.

When doing frame analysis, your analyzer will read frames from the wireless interface and display their contents. Critically, the frame is only available if it is received by your analyzer. With Ethernet, the analyzer can be plugged in, and barring extremely unusual problems or bugs in port mirroring code, will receive frames without issue. On a wireless network, the analyzer only receives what it can decode. To troubleshoot client problems, it is often necessary to physically take the analyzer to the problematic location and set it down next to the device. Ideally, your analyzer will use the same chipset as the client device so that it has similar performance characteristics, though that will not be necessary in most cases.

Some APs include a remote capture tool, which has the advantage of providing the stream of frames as received by the AP. Remote capture from the AP can be used to diagnose problems by looking at the data stream from the AP's vantage point, perhaps while cross-referencing with an analyzer on the client side. For example, if a connectivity problem occurs, comparing captures from the AP and an analyzer near the client device may help you determine if frames are being lost in the air.

Wireshark has several capture modes. For wireless LAN usage, you will want to capture in *monitor mode*, which reports all received frames up the software stack. Monitor mode is required to capture 802.11 header information as well as radio information such as the data rate and *Received Signal Strength Indication* (RSSI).

Network Tuning

The 802.11 MAC manages airtime, which means that performance tuning in 802.11n uses similar techniques to previous PHYs: reduce airtime contention when you can, and pack as many bits into each microsecond when that fails. Just as in 802.11a/b/g, reducing the coverage area of each AP works to increase the frequency reuse and provides independent channels. Just as in 802.11a/b/g, the spectrum in the 5 GHz channel is "cleaner," with less interference, as well as the raw capacity from all the additional channels (see "Channel Types and Layout" on page 105). The past few years have seen greater appreciation of the role of 5 GHz channels in improving throughput. Physics plays a role here: 5 GHz radio waves travel a smaller distance, enabling better frequency re-use.

Although 802.11n changed many rules of the game, it didn't change them all. If you are building a high-capacity network, the 5 GHz band is your friend. Use it.

Many manufacturers select default settings that are generally good for data networking, and will deliver acceptable performance for web-based applications and email. In fact, for 802.11n APs, many APs include a feature that gives priority to high speed 802.11n frames because they move data much more quickly than the older 802.11a/b/g frames. When transmitting a 1500-byte Ethernet frame, 802.11n needs less than a third of the time to move the frame at 300 Mbps than 802.11a/g does at 54 Mbps. When you factor in aggregate frames, the time savings are even more pronounced. Preferential treatment for fast 802.11n frames has the apparent effect of speeding up the network for 802.11n users with only minimal impact to users of older devices.

Tuning for voice

Unlike a data-oriented network, some special configuration may be helpful for networks that support extensive amounts of voice traffic. Voice traffic is demanding because it cannot be buffered, so many of the efficiency enhancements in 802.11n are not used by voice handsets. The core of voice tuning is reducing latency for as much traffic as possible.

QoS configuration: enable Wi-Fi Multi-Media (WMM) and priority queuing
: WMM is a quality-of-service specification that can dramatically improve the quality of voice at the receiver.[8] Not all vendors turn on WMM by default. The single most important configuration change you can make to support higher-quality voice calls is to ensure that WMM is enabled. Some vendors also have an option for strict priority scheduling, which delivers frames in order to the receiver.

Increase data rate used for Beacon frame transmission
: Voice handsets are often very aggressive in roaming between APs, so tuning efforts will focus on decreasing the effective coverage area of APs and reducing large areas of coverage overlap. One of the most effective ways of limiting the effective range of an AP is to make its Beacon transmissions travel a shorter distance. While it is not possible to design a radio wave that stops at a certain distance, increasing the data rate of Beacon frames can be used to limit the effective range of the network. Typically, the Beacon rate will be set at a minimum of 24 Mbps, and sometimes even higher. (802.11a/g rates should be used because many voice handsets do not use 802.11n.)

Limit use of lower data rates
: In combination with using high-data rate Beacon transmission, turn off usage of the lower data rates. Many APs have the ability to disable low data rates, which assists clients in making decisions to move to new APs, and the use of high data rates decreases the overall airtime usage and average latency. Limiting low-speed data rates also helps voice networks function better by reducing the amount of

8. In 2006, I experimented with an early WMM implementation on the Interop show floor to show the dramatic quality improvements with WMM.

airtime required by non-voice clients, leaving a larger share of airtime available for voice devices.

Shorten DTIM interval
: Many voice products use multicast frames for control features or for push-to-talk (PTT) features. Multicast frames are held for transmission until the Delivery TIM (DTIM).[9] Many APs will ship with a DTIM of 3, so multicast transmissions are delivered after every third Beacon. Setting the DTIM to 1 makes multicast delivery more frequent, at the cost of some battery life on handsets that need to power on after every Beacon to receive multicasts.

Reduce retry counters
: Voice is highly sensitive to latency. 802.11 will automatically retry failed transmissions, but retransmissions take additional time. In voice transmission, frames should arrive on time or not at all. Using network capacity to retransmit frames after the target delivery time does not improve call quality, but it can delay other voice frames in the transmit queue. Somewhat counterintuitively, reducing the frame retry count can improve overall latency, and therefore voice quality.

Tuning for multicast

Multicast applications are often similar to voice applications in terms of the demands placed on the network. Multicast traffic streams are often video, and may not be easily buffered. Furthermore, multicast traffic has a lower effective quality of service than unicast traffic on a wireless LAN because multicast frames are not positively acknowledged. In a stream of unicast frames, each frame will be acknowledged and retransmitted if necessary. Multicast transmission has no such reliability mechanism within 802.11, so a stream of multicast frames may not be received.

Shorten DTIM interval
: Just as with voice, many multicast applications depend on promptly receiving data. Setting the DTIM interval as low as possible improves the latency of multicast delivery.

Increase the data rate for multicast frames
: By default, many products will select a low data rate, often 2 Mbps, for multicast transmissions in an effort to be backward compatible. While this is a laudable goal, and the choice of 2 Mbps was reasonable during the 802.11b-to-802.11g transition in 2004, low data rates for multicast no longer serve that goal. Unless there are critical applications running on 2 Mbps devices, or there is a large number of such old devices on the network without any upgrade path, increase the multicast data rate to reduce airtime contention. Many APs can automatically set the multicast data rate to the minimum data rate used for unicast frames to associated clients, or even the minimum unicast rate for clients in the multicast group.

9. For more information on the operation of the DTIM, see Chapter 8 in *802.11 Wireless Networks: The Definitive Guide*.

Enable multicast-to-unicast conversion
: Some APs implement a feature that converts a single multicast frame into a series of unicast frames. Multicast frames must be transmitted at a rate that can be decoded by all receivers, and is therefore often relatively slow. Unicast frames can be transmitted much faster if the receivers are close to the AP. A series of positively acknowledged unicast frames may take approximately the same amount of airtime, but have significantly greater reliability.

Implementation Checklist

When designing and building a network, use the following checklist:

Choose an architecture
: The easy choice in architecture is that the management plane must be centralized. In most cases, a hybrid data plane that blends aspects of both a distributed data plane and centralized forwarding will be the right choice. Carefully evaluate the trade-offs for the location of the management plane based on application requirements and cost.

Hardware selection
: Select hardware that meets requirements for performance and functionality, and is certified by the Wi-Fi Alliance.

Coverage planning
: Lay out a network for 5 GHz coverage, and turn on 2.4 GHz radios as needed to provide coverage. Consider turning on 40 MHz channels if high performance is needed.

Tune for applications
: Adjust the Beacon interval, DTIM, and data rates to suit the applications in use on the network.

Afterword

I have often described my work in the 802.11 working group as the best job I ever had, even though it was a part-time volunteer position that came without pay. As a regular attendee, I had a ringside seat for much of the 802.11n standardization effort, culminating in a vote in the summer of 2009 to approve the task group's final draft. Procedurally, it was a vote like many others I attended, but there was an electricity in the room. After years of exertions to meet the yearning of users for more speed, we were delivering a long-awaited standard. Even though the outcome of that final vote was not in doubt, I went to that meeting in part so that I could say "I was there." As it turns out, I do have an interesting story to tell because the final vote was 53 in favor and one against proceeding, and many people want to understand why there was one no vote.

When I started with wireless LANs, it would have been unthinkable to use them as the primary method of connecting to a network. By delivering 802.11n, some of the smartest people I know have made it unthinkable not to do so. For most practical purposes, wireless networks are now on par with Ethernet. While wireless networks may seem like the obvious choice, few have a firsthand appreciation for technical and intellectual firepower trained within the 802.11 working group that makes it possible.

Even with speeds now in the hundreds of megabits per second with 802.11n, the 802.11 working group continues to drive speeds higher. Without missing a beat, this group began two efforts to achieve gigabit-speed networks; both the future 802.11ac and the 802.11ad draft standards promise speeds of up to 7 Gbps. Products based on both standards are currently in development and should start to hit the market this year. Future efficiency gains based on extending the MIMO technology pioneered in 802.11 with multi-user extensions promise even greater capacity for networks.

With 802.11n, we made users choose not to use wire. With the next generation, we will make users forget wire ever existed.

Matthew Gast

San Francisco, California

January 22, 2011

Glossary

ACK
Abbreviation for "Acknowledgement." ACKs are used extensively in 802.11 to provide reliable data transfers over an unreliable medium. For more details, see "Contention-Based Data Service" in Chapter 3 of 802.11 Wireless Networks: The Definitive Guide.

See also **Block ACK**, **Implicit Feedback**.

AES
Advanced Encryption Standard. A cipher selected by NIST to replace the older Data Encryption Standard (DES) in 2001 after a five-year evaluation. AES is a 128-bit block cipher which uses either 128-, 192-or 256-bit keys. It has been widely adopted by many protocols requiring the use of a block cipher, including CCMP in 802.11, though CCMP uses only 128-bit keys. AES is specified in FIPS Publication 197.

AP
Access Point. Bridge-like device that attaches wireless 802.11 stations to a wired backbone network. For more information on the general structure of an access point, see Chapter 20 of *802.11 Wireless Networks: The Definitive Guide*.

AS
Authentication Server. The network service that validates user credentials. Usually RADIUS in 802.11 networks.

Basic Block ACK
The original block acknowledgement specification in the 802.11e amendment allowed a receiver of a group of frames to selectively acknowledge individual 802.11 fragments. Extensions in 802.11n make the protocol more efficient for use with 802.11n networks.

See also **Compressed Block ACK**.

basic service set
See **BSS**.

beamforming
A method of using precise phase shifts on an antenna array that focuses the resulting transmission in a particular direction. Sending beamformed transmissions may require an exchange of control information to set up the antenna array.

beamformee
The receiver of a beamformed transmission. The beamformee may need to transmit some packets in a beamforming setup exchange, but the main purpose of the beamforming exchange is to receive a directional transmission.

beamformer
The sender of a beamformed transmission. The beamformer may need to receive some packets in a beamforming setup exchange, but the main purpose of such an exchange is to send a directional transmission.

Block ACK
: A mechanism that allows the recipient of a series of frames to transmit one acknowledgement for the entire series. It enables selective acknowledgement of each frame in the series. By transmitting just one umbrella ACK frame, it makes substantially more efficient use of airtime than the traditional positive ACK transmitted in response to a single frame.

Block ACK Request
: The Block ACK Request (BAR) frame is sent prior to a series of frames that the transmitter would like to be acknowledged. Without a Block ACK Request, the receiver cannot send a block ACK.

BPSK
: Binary Phase Shift Keying. A modulation method that encodes bits as phase shifts. One of two phase shifts can be selected to encode a single bit.

BSS
: Basic Service Set. The building block of 802.11 networks. A BSS is a set of stations that are logically associated with each other.

BSSID
: Basic Service Set Identifier. A 48-bit identifier used by all stations in a BSS in frame headers.

code rate
: In the context of a forward-error correcting code, the code rate describes the fraction of bits devoted to error correction, and is typically symbolized by R. For example, an R=1/2 code takes the input data stream and encodes every payload bit as two bits. Codes can be described as conservative, or able to correct large errors. Conversely, a code rate may be aggressive, meaning that error correction capacity is being sacrificed for efficiency. The lower the code rate, the more conservative a code is; coding at R=1/2 enables more error recovery than coding at R=5/6.

Compressed Block ACK
: A new block ACK extension defined by 802.11n. The "compression" referred to in the name refers to the fact that the compressed Block ACK mechanism can only acknowledge non-fragmented frames. 802.11n uses such large aggregate frames that fragmentation is not commonly used, and the block ACK window can be made substantially more efficient by acknowledging at the frame level instead of the fragment level.

  See also **Block ACK**, **Basic Block ACK**.

CCM
: Counter Mode with CBC-MAC. An authenticated block cipher mode defined in RFC 3610. It can be used with any 128-bit block cipher, but is commonly used with AES in wireless LANs for security.

CCMP
: Counter Mode with CBC-MAC Protocol. 802.11i-2004 defined the use of AES with the CCM mode of operation as CCMP. It is the strongest encryption protocol available for use with wireless LANs, and the only security protocol allowed for use with 802.11n.

CRC
: Cyclic Redundancy Check. A mathematical checksum that can be used to detect data corruption in transmitted frames. The CRC is a linear hash function, and should not be used for data security assurance.

CSMA
: Carrier Sense Multiple Access. A "listen before talk" scheme used to mediate the access to a transmission resource. All stations are allowed to access the resource (multiple access) but are required to make sure the resource is not in use before transmitting (carrier sense).

CSMA/CA
: Carrier Sense Multiple Access with Collision Avoidance. A CSMA method that tries to avoid simultaneous access (*collisions*) by deferring access to the medium. 802.11 and

AppleTalk's LocalTalk are two protocols that use CSMA/CA.

CTS

Clear to Send. The frame type used to acknowledge receipt of a Request to Send and the second component used in the RTS-CTS clearing exchange used to prevent interference from hidden nodes.

DA

Destination Address. The MAC address of the station the frame should be processed by. Frequently, the destination address is the receiver address. In infrastructure networks, however, frames bridged from the wireless side to the wired side will have a destination address on the wired network and a receiver address of the wireless interface in the access point.

DBPSK

Differential Binary Phase Shift Keying. A modulation method in which bits are encoded as phase shift differences between successive symbol periods. Two phase shifts are possible for an encoding rate of one data bit per symbol.

DCF

Distributed Coordination Function. The rules for contention-based access to the wireless medium in 802.11. The DCF is based on exponentially increasing backoffs in the presence of contention as well as rules for deferring access, frame acknowledgment, and when certain types of frame exchanges or fragmentation may be required.

Delayed Block ACK

A method of transmitting a block ACK some time after the last data frame in the burst to be acknowledged was successfully received.

DFS

Dynamic Frequency Selection. A spectrum management service required by European radio regulations (European Commission decisions 2005/513/EC and 2007/90/EC, along with ETSI EN 301 893) to avoid interfering with 5 GHz radar systems, as well as spread power across all available channels. DFS was also key to the FCC decision to open up the harmonized frequency band in the U.S.

DIFS

Distributed Inter-Frame Space. The interframe space used to separate atomic exchanges in contention-based services.

See also **DCF**.

DQPSK

Differential Quadrature Phase Shift Keying. A modulation method in which bits are encoded as phase shift differences between successive symbol periods. Four phase shifts are possible for an encoding rate of two data bits per symbol.

DS

Distribution System. The set of services that connects access points together. Logically composed of the wired backbone network plus the bridging functions in most commercial access points.

DSSS

Direct-Sequence Spread Spectrum. A transmission technique that spreads a signal over a wide frequency band for transmission. At the receiver, the widespread signal is correlated into a stronger signal; meanwhile, any narrowband noise is spread widely. Most of the 802.11-installed base at 2 Mbps and 11 Mbps is composed of direct-sequence interfaces.

DTIM

Delivery Traffic Indication Map. Beacon frames may contain the DTIM element, which is used to indicate that broadcast and multicast frames buffered by the access point will be delivered shortly.

ESS

Extended Service Set. A logical collection of access points all tied together. Link-layer roaming is possible throughout an ESS, provided all the stations are configured to recognize each other.

ETSI
: European Telecommunications Standards Institute. ETSI is a multinational standardization body with regulatory and standardization authority over much of Europe. GSM standardization took place under the auspices of ETSI. ETSI has taken the lead role in standardizing a wireless LAN technology competing with 802.11 called the High Performance Radio LAN (HIPERLAN).

Explicit Feedback
: When used with beamforming, this refers to a beamforming method that requires frames to be sent between the two parties to a beamformed transmission. The beamformee must send frames that help the beamformer calibrate future transmissions.

FEC
: Forward Error Correction. A type of code in which the transmitter takes the payload for transmission and encodes it with redundant bits to enable the receiver to correct errors. There are two main types: convolutional codes that work on arbitrary-length streams of data, and block codes that work on fixed-length blocks.

FCC
: Federal Communications Commission. The regulatory agency for the United States. The FCC Rules in Title 47 of the Code of Federal Regulations govern telecommunications in the United States. Wireless LANs must comply with Part 15 of the FCC rules, which are written specifically for RF devices.

FCS
: Frame Check Sequence. A checksum appended to frames on IEEE 802 networks to detect corruption. If the receiver calculates a different FCS than the FCS in the frame, it is assumed to have been corrupted in transit and is discarded.

FIPS
: Federal Information Processing Standard. Public standards used by nonmilitary agencies of the United States federal government and its contractors.

four-way handshake
: The key exchange defined in 802.11i that expands a pairwise master key into the full key hierarchy. The 4-Way Handshake allows a supplicant and an authenticator to agree on dynamically derived encryption keys.

GMK
: Group Master Key. The key used by an authenticator to derive the group transient key.

GTK
: Group Transient Key. Derived from the group master key by combining with the group random number, the GTK is used to derive the group key hierarchy, which includes keys used to protect broadcast and multicast data.

HR/DSSS
: High-Rate Direct-Sequence Spread Spectrum. The abbreviation for signals transmitted by 802.11b equipment. Although similar to the earlier 2-Mbps transmissions in many respects, advanced encoding enables a higher data rate.

HT
: High Throughput. The official name of the 802.11n PHY, and a common abbreviation that is used colloquially to mean "802.11n."

IEEE
: Institute of Electrical and Electronics Engineers. The professional body that has standardized the ubiquitous IEEE 802 networks.

Immediate Block ACK
: A style of Block ACK in which the Block ACK frame is sent immediately following the frames that it is acknowledging.

Implicit Feedback
: A method of beamforming where no explicit communication takes place between the beamformer and beamformee. Implicit feedback often uses the received frames

themselves to estimate the required channel calibration. It does not produce as effective a steering matrix, but it does not require software support at both ends of the link.

ISM

Industrial, Scientific, and Medical. Part 15 of the FCC Rules sets aside certain frequency bands in the United States for use by unlicensed Industrial, Scientific, and Medical equipment. The 2.4-GHz ISM band was initially set aside for microwave ovens so that home users of microwave ovens would not be required to go through the burdensome FCC licensing process simply to reheat leftover food quickly. Because it is unlicensed, though, many devices operate in the band, including 802.11 wireless LANs.

ITU

International Telecommunications Union. The successor to the CCITT. Technically speaking, the ITU issues recommendations, not regulations or standards. However, many countries give ITU recommendations the force of law.

LDPC

Low-density Parity Check. A block error-correction code that can optionally be used in 802.11.

LLC

Logical Link Control. An IEEE specification that allows further protocol multiplexing over Ethernet. 802.11 frames carry LLC-encapsulated data units.

MAC

Medium Access Control. The function in IEEE networks that arbitrates use of the network capacity and determines which stations are allowed to use the medium for transmission.

MCS

Modulation and Coding Set. A number from 0-76 that describes both the modulation and the forward error-correcting code used.

MIMO

Multiple-Input/Multiple-Output. An antenna configuration that uses more than one transmission antenna and more than one receiver antenna to transmit multiple data streams. MIMO antenna configurations are often described with the shorthand "Y×Z," where Y and Z are integers, used to refer to the number of transmitter antennas and the number of receiver antennas, respectively.

MPDU

MAC Protocol Data Unit. A fancy name for frame. The MPDU does not, however, include PLCP headers.

MRC

Maximal Ratio Combining. A method of combining the signals from multiple antennas in an antenna array to boost the signal-to-noise ratio of a received frame. MRC uses the "extra" radio chains in an antenna array to provide additional information.

MSDU

MAC Service Data Unit. The data accepted by the MAC for delivery to another MAC on the network. MSDUs are composed of higher-level data only. For example, an 802.11 management frame does not contain an MSDU.

NAV

Network Allocation Vector. The NAV is used to implement the virtual carrier sensing function. Stations will defer access to the medium if it is busy. For robustness, 802.11 includes two carrier-sensing functions. One is a *physical* function, which is based on energy thresholds, whether a station is decoding a legal 802.11 signal, and similar things that require a physical measurement. The second function is a *virtual* carrier sense, which is based on the NAV. Most frames include a nonzero number in the NAV field, which is used to ask all stations to politely defer from accessing the medium for a certain number of microseconds after the current frame is transmitted. Any receiving stations will process the NAV and defer access, which prevents collisions. For more

detail on how the NAV is used, see "Contention-Based Data Service" in Chapter 3 of 802.11 Wireless Networks: The Definitive Guide.

noise floor

The noise floor is the level of ambient background "static" in an area. Transmissions must rise above the noise floor in order to be received. A good analogy for the noise floor is the burble of conversations within a room where a party is being held. In order to hear and understand a single voice, you have to be able to concentrate on it so you can hear it over the background level.

OBSS

Overlapping BSS. Refers to another network installed in the same physical space on the same channel, whether it is part of the same ESS or not. If two access points were installed next to each other on channel 6, each would be an OBSS of the other.

OFDM

Orthogonal Frequency Division Multiplexing. A technique that splits a wide frequency band into a number of narrow frequency bands and inverse multiplexes data across the subchannels. 802.11a and 802.11g are based on OFDM. 802.11n uses MIMO to transmit multiple OFDM data streams.

PCO

Phased Coexistence. A method by which an 802.11n AP switches between operating a 20 MHz channel and a 40 MHz channel to give the benefits of higher-speed wide channels to those devices that support it while retaining backward compatibility with 20 MHz devices. Operation of the channel is divided into alternating 20 MHz and 40 MHz phases.

PDU

See **protocol data unit**.

PHY

Common IEEE abbreviation for the physical layer.

PLCP

Physical Layer Convergence Procedure. The upper component of the PHY in 802.11 networks. Each PHY has its own PLCP, which provides auxiliary framing to the MAC.

PMD

Physical Medium Dependent. The lower component of the PHY, responsible for transmitting RF signals to other 802.11 stations.

PMK

Pairwise Master Key. The root of all keying data between a supplicant and an authenticator. It may be derived from an EAP method during authentication, or supplied directly as a preshared key.

PPDU

PLCP Protocol Data Unit. The complete PLCP frame, including PLCP headers, MAC headers, the MAC data field, and the MAC and PLCP trailers.

protocol data unit

Layers communicate with each other using protocol data units. For example, the IP protocol data unit is the familiar IP packet. IP implementations communicate with each other using IP packets.

See also **Service Data Unit**.

PS

Power Save. Used as a generic prefix for power-saving operations in 802.11.

PSDU

PLCP Service Data Unit. The data the PLCP is responsible for delivering. Typically it will be one frame from the MAC, with headers. In 802.11, however, the PSDU may consist of an aggregate of several MAC service data units.

PSK

Pre-shared Key. In 802.11i, refers to the authentication method that depends on a statically configured authentication key that must be distributed manually. Also called WPA-PSK.

PSMP
: Power-Save Multi-Poll. A power-saving system specific to 802.11n that improves both power efficiency and airtime efficiency by scheduling transmissions to associated clients.

QAM
: Quadrature Amplitude Modulation. A modulation method that varies both the amplitude and phase simultaneously to represent a symbol of several bits. 802.11n uses both 16-QAM and 64-QAM at higher transmission rates.

QPSK
: Quadrature Phase Shift Keying. A modulation method that encodes bits as phase shifts. One of four phase shifts can be selected to encode two bits.

RA
: Receiver Address. MAC address of the station that will receive the frame. The RA may also be the destination address of a frame, but not always. In infrastructure networks, for example, a frame destined for the distribution system is received by an access point.

RADIUS
: Remote Authenticated Dial-In User Service. A protocol used to authenticate dial-in users that has become more widely used because of 802.1X authentication. The most common type of authentication server used in 802.1X systems.

RLAN
: Radio LAN. A term used by European radio regulations to refer to any wireless network built on radio technology. Although 802.11 is the most popular, others do exist. One of the better known alternative radio network technologies is ETSI'S HIPERLAN.

RF
: Radio Frequency. Used as an adjective to indicate that something pertains to the radio interface ("RF modulator," "RF energy," and so on).

RIFS
: Reduced Interframe Space. A shortened frame separator that allows better use of available airtime when two HT devices are communicating with each other.

RSN
: Robust Security Network. A network that uses the security methods originally defined 802.11i-2004, and does not provide any support for the use of WEP.

RSSI
: Received Signal Strength Indication. This is a value reported for the strength of a frame that has been received, and acts much like a "volume" indicator for the transmission. RSSI may be reported in many different ways, but a common method is in dBm.

RTS
: Request to Send. The frame type used to begin the RTS-CTS clearing exchange. RTS frames are used when the frame that will be transmitted is larger than the RTS threshold.

SA
: Source Address; as disinct from TA. Station that generated the frame. Different when frame originates on the distrbution system and goes to the wireless segment.

SDU
: See Service Data Unit.

Service Data Unit
: When a protocol layer receives data from the next highest layer, it is sending a service data unit. For example, an IP service data unit can be composed of the data in the TCP segment plus the TCP header. Protocol layers access service data units, add the appropriate header, and push them down to the next layer.

  See also protocol data unit.

SIFS
: Short Interframe Space. The shortest of the four interframe spaces. The SIFS is used between frames in an atomic frame exchange.

Spatial stream
: MIMO techniques are sometimes called spatial reuse because a MIMO system will send multiple independent data streams between the transmitter and the receiver. Each data stream is called a spatial stream because it takes a different path through space between the transmitter and receiver. An 802.11n device may have up to four spatial streams. For any given transmission, the maximum number of spatial streams is defined by the lower number

SSID
: Service Set Identity. A string used to identify an extended service set. Typically, the SSID is a recognizable character string for the benefit of users.

STBC
: Space-Time Block Coding. A method of transmitting a single data stream across multiple antennas for additional transmission redundancy.

TA
: Transmitter Address. Station that actually put the frame in the air. Often the access point in infrastructure networks.

TIM
: Traffic Indication Map. A field transmitted in Beacon frames used to inform associated stations that the access point has buffered. Bits are used to indicate both buffered unicast frames for each associated station as well as the presence of buffered multicast frames.

TK
: Temporal Key. 802.11i key hierarchies derive a temporal key to be used for authentication protocols. The temporal key is the main input to link-layer encryption protocols such as TKIP or CCMP.

TKIP
: Temporal Key Integrity Protocol. One of the improved encryption protocols in 802.11i, TKIP uses the fundamental operations of WEP with new keying and integrity check mechanisms to offer additional security. 802.11n clearly forbids the use of TKIP with 802.11n frames.

WEP
: Wired Equivalent Privacy. Derided as Wiretap Equivalence Protocol by its critics. A standard for ciphering individual data frames. It was intended to provide minimal privacy and has succeeded in this respect. In August 2001, WEP was soundly defeated, and public code was released. WEP is not supported with 802.11n devices.

Wi-Fi
: An umbrella term used to refer to wireless LANs in general, and a testament to the strength of the Wi-Fi Alliance's branding activities. "Wi-Fi" is often used interchangeably with "wireless LAN" or "802.11."

Wi-Fi Alliance
: The Wi-Fi Alliance (formerly the Wireless Ethernet Compatibility Alliance) started the Wi-Fi certification program to test interoperability of 802.11 implementation. Originally, the term was applied to devices that complied with 802.11b (11-Mbps HR/DSSS), but further programs have extended PHY interoperability testing to include 802.11a, 802.11g, and 802.11n, as well as security.

Wi-Fi CERTIFIED
: Trademark of the Wi-Fi Alliance used to indicate that a particular device has passed an interoperability test. Once certified, a product's capabilities are published in the Wi-Fi Alliance certification database, and an interoperability certificate lists certified capabilities.

Wi-Fi CERTIFIED n
: A sub-set of Wi-Fi CERTIFIED indicating that a product has passed the certification test specifically for 802.11n. Security is required, and older less secure security protocols (WEP and TKIP) are forbidden. Additionally, an AP must support at least two spatial streams.

with some n features
: A tagline used with Wi-Fi certification of 802.11a/b/g devices to indicate it has passed interoperability testing with some features in 802.11n. It indicates that a single-stream AP supports some 802.11n-specific features.

WPA and WPA2
: Wi-Fi Protected Access. A security standard based on 802.11i draft 3. The Wi-Fi Alliance took 802.11i draft 3 and began certifying compliance with early TKIP implementations to accelerate adoption of 802.11 security protocols. WPA2 is based on the full ratified version of 802.11i-2004. Products certified with 802.11n are only allowed to use CCMP to encrypt high-speed 802.11n frames.

About the Author

Matthew Gast is Director of Product Management at Aerohive Networks, where he leads development of the core software technologies that power Aerohive network devices. He is the past chair of the 802.11-2012 revision task group, and was an officer and contributor to 802.11u-2011. At the Wi-Fi Alliance, he leads efforts to improve wireless network security, most recently through the launch of the Protected Management Frame certification program based on the 802.11w-2009 specification. His first book on wireless LANs, 802.11 Wireless Networks: The Definitive Guide, is now in its second edition and has been translated into six languages. Matthew is an avid photographer, and is currently trying to waste even more money on his hobbies by taking pilot lessons.