802.11n: The Standard Revealed

The Next Generation Wireless

Executive Summary

The IEEE 802.11 Working Group has now completed 802.11n, the multi-year effort to standardize an upgrade to the 802.11 radio. 802.11n provides a new set of capabilities dramatically improving the reliability of communications, the predictability of coverage, and the overall throughput of devices.

There are no additional mandatory features introduced between the widely adopted draft 2.0 version of 802.11n and the final version. Therefore, customers who already have chosen the draft 2.0 devices can continue to operate these devices with confidence and without any hardware or software changes. Also, those who have been holding back waiting for the standards to be finally ratified before moving over to the 802.11n can start migration now.

The 802.11n protocol has several enhancements in the physical layer and the MAC sublayer that provide exceptional benefits to wireless deployments. The four key features are:

- Multiple-input multiple-output (MIMO). MIMO uses the diversity and duplication of signals using the multiple transmit and receive antennas.
- 40-MHz operation bonds adjacent channels combined with some of the reserved channel space between the two to more than double the data rate
- Frame aggregation reduces the overhead of 802.11 by coalescing multiple packets together.
- Backward compatibility, which makes it possible for a/b/g and 802.11n devices to coexist, thereby allowing customers to phase in their access point and/or client migrations over time.

This white paper describes these new capabilities in detail, explains how the benefits provided by 802.11n are achieved, and examines the compatibility of this new standard with existing deployments. This white paper also describes the issues to address when planning migration of an existing 802.11a/g deployment to 802.11n and the results that can be expected from such a migration. But first, it is important to understand just what 802.11n is, and what it is not.

The Role of the IEEE and the Wi-Fi Alliance

802.11n was a 7-year endeavor at the IEEE, with three major phases: Study Group, Task Group and Sponsor Ballot. The High Throughput Study Group first met on September 11, 2002, and this led to the 802.11n Task Group a year later. Considerable work ensued to explore the core set of features that provided the maximum benefit for the broadest set of devices; these features were finalized in draft 2.0. Cisco provided valuable leadership toward robust backward compatibility and the security of the enhanced protocols.

The Wi-Fi Alliance, an industry organization that provides interoperability certification for 802.11 devices, first began certifying the interoperability of draft 2.0 802.11n devices in June 2007. The certification tested the core features of 802.11n. The program was broadly adopted by the Wi-Fi vendors and their equipment enjoyed strong market penetration. To date several hundred draft 2.0 802.11 products have been certified by the Wi-Fi Alliance and several tens of millions of devices have been deployed worldwide.[1]

---

[1] 802.11n: Ready for Business: http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps6973/white_paper_c11-457039_ns767_Networking_Solutions_White_Paper.html

Over time the IEEE progressed on the other functionality that was incompletely addressed by draft 2.0. At each revision of the draft, optional features, providing incremental customer benefits, were refined. This process continued until June 2009 and draft 11.0. The official version of 802.11n is the one produced by the 802.11 Working Group and ratified by the IEEE. This version was ratified by the IEEE Standards Board on September 11, 2009.

Underpinning discussions on changes to 802.11n after draft 2.0 was the strong goal of the Wi-Fi Alliance and its members to ensure that differences between the final ratified 802.11n standard and 802.11n draft 2.0 were such that any device built to the 802.11n draft 2.0 specifications could be upgraded to the final, ratified standard with software-only changes. This strategy proved to be highly effective. In fact the final 802.11n standard has no mandatory features in addition to those defined in draft 2.0.

Pre-802.11n Products

Even before there was a draft of the 802.11n standard that had reached any degree of consensus in the 802.11 Working Group, there were many consumer-grade products available that claimed to be "pre-n." Each product was based on one particular proposal or another out of the many proposals that were made to the 802.11 Working Group. Most of these products are not compatible or interoperable between different vendors.

802.11n Technology

The goal of the work on 802.11n is to dramatically increase the effective throughput of 802.11 devices available to end-user applications, not to simply build a radio capable of higher bit rates.

The difference between these goals is like the difference between replacing commuter cars with buses and redesigning commuter cars to do away with the back seat. Although the car is shorter and so more cars can be packed more densely in the lane than buses, the benefit of the redesign is dwarfed by the continuing need for a safe following distance, the engine bay, and the trunk. In the same way, increasing the effective throughput of an 802.11 device requires more than providing a higher bit rate: every aspect of 802.11 that introduces overhead needs to be minimized as far as possible.

This is the essence of 802.11: not only a faster physical layer, but also a more efficient MAC layer so that more of the speed benefits are available to users. The important improvements in 802.11 are:

- Reliability and bit rate are increased by MIMO.
- The bit rate is further increased by 40-MHz operation.
- Frame aggregation ensures more of the increased bit rate is available as higher throughput to applications.
- Robust backward compatibility means that it is straightforward for networks to be progressively upgraded to 802.11n.

In addition:

- Opportunities for client power saving are extended.

Let’s look at each of these improvements in turn.

MIMO

Multiple-input multiple-output (MIMO) is the heart of 802.11n. This technical discussion of MIMO provides a basis for understanding how 802.11n can raise reliability and reach data rates many times higher than 802.11a/b/g in the same radio spectrum. These benefits improve the experience of all wireless users, and are especially valuable for customers with difficult RF environments, such as thick walls or small rooms, and for customers with high throughput or quality-of-service (QoS) requirements, such as voice or video users.

Radio Operation Basics

To understand the improvement brought by MIMO technology, it is important to understand some of the basics that determine how well a traditional radio operates. In a traditional, single-input single-output radio, the amount of information that can be carried by a received radio signal depends on the amount by which the received signal strength exceeds the noise at the receiver, called the signal-to-noise ratio, or SNR. SNR is typically expressed in decibels (dB). Higher SNR enables more information to be carried on the signal and recovered by the receiver.

To understand the improvements that MIMO technology brings to 802.11, imagine trying to drive quickly down a road. The bigger the car engine, the better its tires and suspension, and the lower the wind resistance, the faster the car can go. At the same time, a road that is full of pot-holes or wind slows the car down. In this analogy, the capabilities of the car are the signal, and the poor road quality is the noise. And better SNR means a faster experience.

Even when maximum speed is not the objective, having extra power in reserve means it is easier to pass slow traffic and to maintain speed under difficult circumstances along the journey such as steep ascents or head-winds. Cruising speed can be maintained with greater reliability. Having a high SNR helps in much the same way: being able to improve the SNR means being able to reduce dead-spots and attain a desired data rate with much less sensitivity to environmental variability.

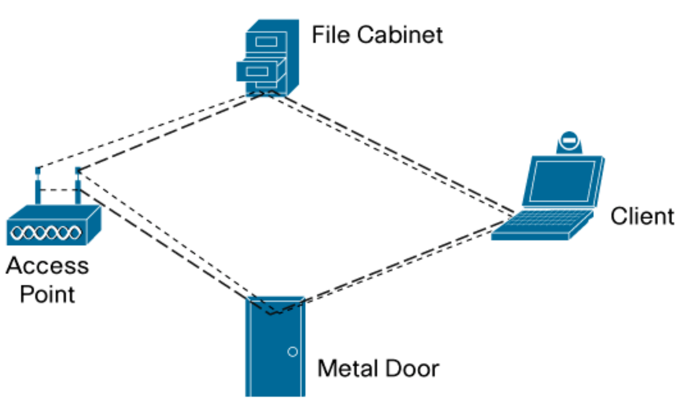
Once the minimum SNR is achieved to allow information to be exchanged at the desired rate, any additional SNR is an additional benefit. That additional SNR can be spent on increasing the information rate, increasing the distance, or a little bit of both. However, you can't spend the same dB more than once, just as you can't spend the same dollar more than once (at least not without encountering some unpleasant consequences).

Spatial Propagation for MIMO

In typical indoor WLAN deployments—for example, enterprise offices, classrooms, hospitals, and warehouses—it is rare for the radio signal to only take the direct, shortest path from the transmitter to the receiver. Often there is no "line of sight" between the transmitter and the receiver: there is a cube wall, door, or other structure that obscures the line of sight. All of these obstructions reduce the strength of the radio signal as it passes through them, perhaps to an unusable level.

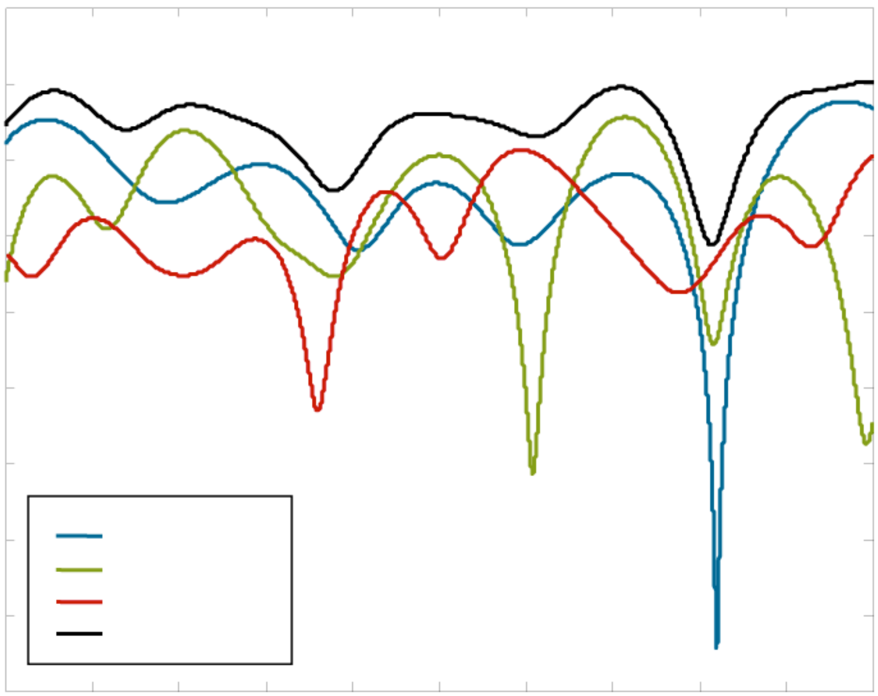
Luckily, there are other ways for the radio signal to travel from transmitter to receiver. Most environments are full of surfaces that reflect a radio signal as well as a mirror reflects light. Imagine that all of the metallic surfaces, large and small, that are in an environment were actually mirrors. Nails and screws, door frames, ceiling suspension grids, pipes and structural beams are all reflectors of radio signals. It would be possible to see the same WLAN access point in many of these mirrors simultaneously. Some of the images of the access point would be a direct reflection through a single mirror. Some images would be a reflection of a reflection. Still others would involve an even greater number of reflections. This phenomenon is called *multipath* (see Figure 1).

Figure 1. Multipath



When a signal travels over different paths to a single receiver, the time that the signal arrives at the receiver depends on the length of the path it traveled. The signal traveling the shortest path will arrive first, followed by copies or echoes of the signal slightly delayed by each of the longer paths that the copies traveled. When traveling at the speed of light, as radio signals do, the delays between the first signal to arrive and its copies is very small, only nanoseconds. (A rule of thumb for the distance covered at the speed of light is roughly one foot per one nanosecond.) This delay is enough to be able to cause a significant transformation of the signal at a single antenna because all the copies interfere with the first signal to arrive. Sometimes the interference is constructive, sometimes neutral, and sometimes destructive. Indeed, the nature of the interference varies randomly across the band of the signal and from antenna position to antenna position (see Figure 2).

Figure 2. Variation of Signal Quality from Constructive and Destructive Interference across the Frequency Band of the Signal and across Antennas



A MIMO radio sends multiple radio signals at the same time and takes advantage of multipath. The signals may be careful modifications of the same information, in order to improve reliability, or totally different information, in order to improve throughput. At its simplest, each signal is sent from its own antenna, using its own RF chain. (An RF chain takes a low-frequency signal and shifts it up to the right RF radio channel ready for transmission out of the antenna, or vice versa). Because there is some space between each of these antennas, each signal follows a slightly different path to a receive antenna. The receiver has multiple antennas as well, each with its own RF chain, experiencing relatively independent multipath. Each of the receive antennas separately harvests the arriving signals (see Figure 3). Then, each antenna's received signal is combined with the signals from the other receive antennas. With a lot of complex math, MIMO provides for three profoundly valuable capabilities as shown in Figure 4:

- The ability to use multiple transmit antennas to improve the SNR of the signal at the receiver
- The ability to use multiple receive antennas to improve the SNR of the signal at the receiver, known as MIMO equalization

- The ability to send two or more signals at the same time over the same spectrum. Each of these transmitted signals is called a **spatial stream** and sending multiple spatial streams in this way is known as **spatial division multiplexing**. 802.11n standard allows for up to 4 spatial streams. Beyond the contemporary two spatial streams, performance enhancements are incremental in nature.

Figure 3. Spatial Division Multiplexing

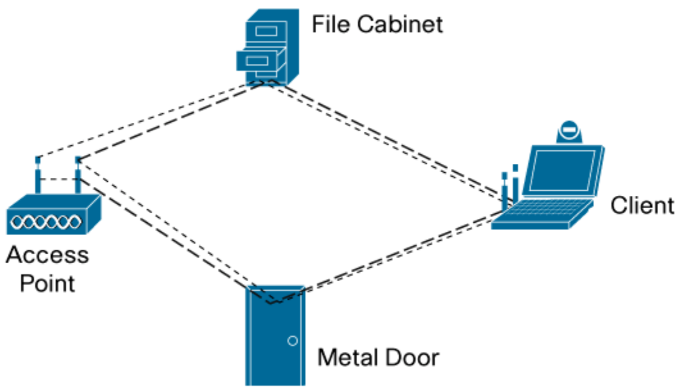
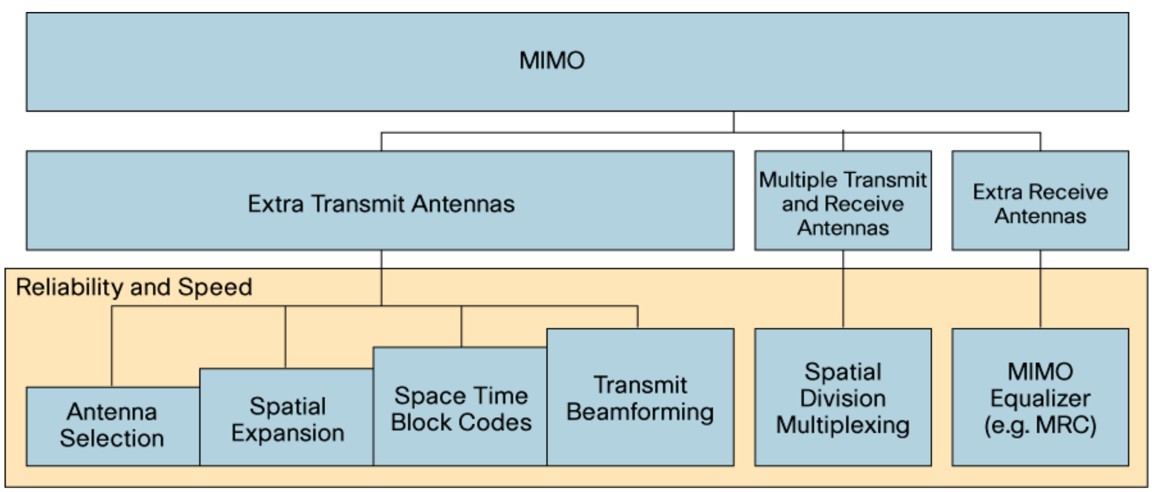


Figure 4. Reliability and Speed Improvements Enabled by MIMO



MIMO links describe how one device (for example, an access point) transmits to a second device (for example, a client). The important parameters are the number of transmit antennas (with RF chains) in the access point and the number of receive antennas (with RF chains) in the client. For example, 2x1 is “two by one,” signifying two access point transmit antennas and one client receive antenna.

Confusingly, **MIMO devices** are also described by similar language: the number of transmit and receive antennas in a device—for example, a 2x3 access point means that the access point has two transmit antennas and three receive antennas. But in this white paper we are concerned with MIMO links, not MIMO devices.

802.11n defines a number of different communication options according to the number of transmit antennas and the number of receive antennas in a MIMO link. Some antennas are used for spatial division multiplexing. And if there are more transmit or receive antennas than needed for spatial division multiplexing, they are used to improve the SNR for the link. Also, MIMO links are **reciprocal**, in the sense that the signal sent one way from device A to device B is affected by the RF environment in the same way as a signal sent from device B to device A. This means, for instance, that a 2x3 link is much like a 3x2 link. Still, making full use of the MIMO link depends on having the right MIMO technologies at one or both ends.

Thus the first significant benefit of MIMO is to improve the SNR. With suitable MIMO technologies available, the gain in SNR is large for each step from 1x1 to 1x2 or 2x1 to 2x2 and to 2x3 or 3x2, but the improvement with 3x3 and beyond is incremental in nature.

Using Transmit Antennas to Improve the SNR

Having (and using) more transmit antennas than spatial streams can help to improve the SNR of a MIMO link through one of the following techniques (in order from most helpful to least helpful): transmit beamforming, Space Time Block Coding (STBC), spatial expansion (SE, and also called cyclic delay diversity), and antenna selection.

To understand the difference between these techniques, consider the following analogies. If antenna selection is like changing lanes on a freeway to finding a better lane for a time, then spatial expansion is like randomly moving cars to random lanes so that congestion never gets very dense. STBC is like driving down two lanes at a time yet being able to combine free space in one lane with free space on another lane, while transmit beamforming is like moving all other cars out of your lane whenever there is space.

Thus, despite this proliferation of techniques, the main fact is that **transmit beamforming** is the most powerful mechanism to improve the SNR of a link.

Transmit Beamforming

When there is more than one transmit antenna, **transmit beamforming** is the method to coordinate the signal sent from each antenna so that the signal at the receiver is dramatically improved. This technique was originally conceived for situations in which the receiver has only a single antenna and when there are few obstructions or radio-reflective surfaces (for example, in open storage yards). However, it turns out that transmit beamforming with **MIMO-orthogonal frequency division multiplexing** (MIMO-OFDM) is a tractable problem, and the solution is just as valuable in the rich indoor multipath environment. 2x1 transmit beamforming can improve SNR by 3 to 6 dB depending on the data rate.

To understand transmit beamforming, consider a radio signal as a wave shape, with a wave length that is specific to the frequency of the signal. When two radio signals are sent from different antennas, these signals are added together at the receiver's antenna (see Figure 5). Depending on the distance that each radio signal travels, they are very likely to arrive at the receiver out of phase with each other. This difference in phase at the receiver affects the overall signal strength of the received signal. By carefully adjusting the phase of the radio signals at the transmitter, the received signal can be maximized at the receiver, increasing SNR. This is what transmit beamforming does for each individual subcarrier in the signal: it effectively focuses the transmit antennas on a single receiver, as shown in Figure 6.

Figure 5. Destructive Interference

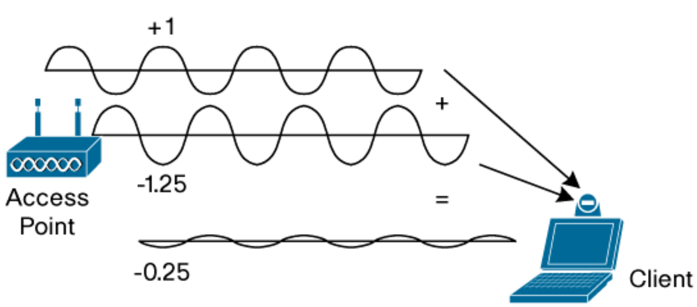
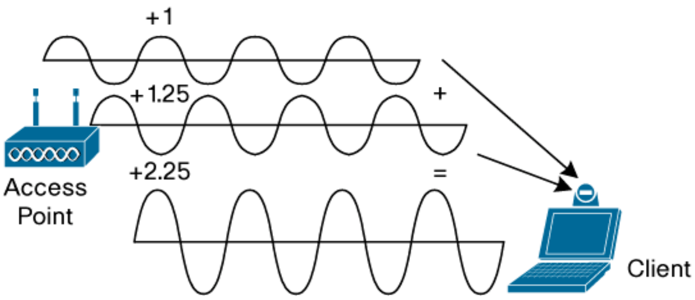


Figure 6. Transmit Beamforming (Synthesizing Constructive Interference)



At first glance, transmit beamforming is not something that can easily be done at the transmitter without information from the receiver about the received signal. Protocols for obtaining this feedback in multiple, optional ways is standardized by the 802.11n amendment. These protocols require that both access point and client support a common subset of the options. Yet because of these fragmented options and the complexity of implementing them, the first generation of mainstream 802.11n access point and client chipsets did not implement 802.11n transmit beamforming. Accordingly, 802.11n transmit beamforming is not yet certified. 802.11n transmit beamforming will come eventually, but not until the next generation of chipsets take hold in the market. Meanwhile, Cisco has already demonstrated leadership by providing the benefits of transmit beamforming without requiring client support[2].

Meanwhile, no protocols exist for 802.11a, b, or g devices. Yet these devices, lacking multiple receive antennas, need transmit beamforming the most! Cisco® ClientLink transmit beamforming helps here since it does not depend upon 802.11n for feedback and instead it gathers everything it needs from the packets it receives directly from each client The client need not even be aware that transmit beamforming is occurring, it just receives packets with more reliability or at a higher rate, or a bit of both.

Transmit beamforming is most valuable for increasing the data rate when transmitting to a single Wi-Fi receiver. This is because it is not possible to optimize the phase of the transmitted signals when sending broadcast or multicast transmissions. Consequently, transmit beamforming does not increase the coverage area of an access point, since that is determined, in large part, by the ability to receive the broadcast beacons from the access point.

Space Time Block Coding

Whereas transmit beamforming raises the quality of a channel by coordinating the signals sent from each antenna, **space time block coding (STBC)** makes best use of the unimproved channel.

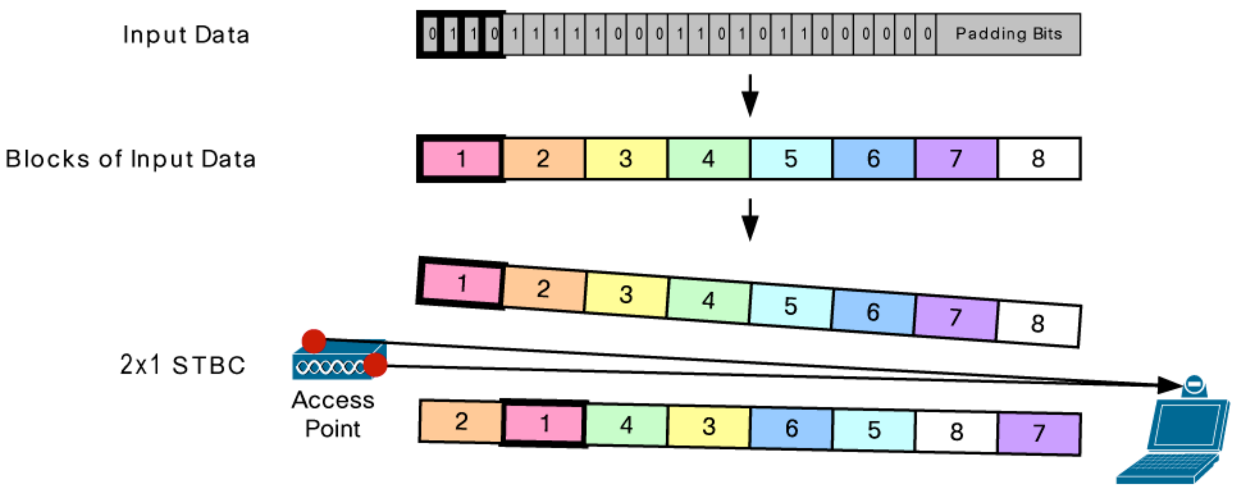
There are 2x1, 3x2, 4x2, and 4x3 flavors of STBC, though only 2x1 STBC is certified by the Wi-Fi Alliance. 2x1 STBC divides the data into blocks (sized according to how many data bits fit into an OFDM symbol). Next 2x1 STBC duplicates the blocks as two streams, one per antenna, where the second stream is a reordering of the first. Consider just the first block, which is transmitted twice: first on one transmit antenna and then later on the opposite antenna. This means if the SNR from either transmit antenna to the receive antenna is good, the data can be recovered reliably; it takes destructive multipath on both antennas for failure to occur.

This trick for just the odd-numbered blocks is well known as **fast time diversity**, and on its own increases reliability yet reduces the data rate by 50 percent. What 2x1 STBC adds is the ability to transmit the even-numbered data blocks too, as shown in Figure 7. All data blocks are sent twice, over different antennas, so the maximum diversity from the unimproved channel is harvested without a reduction in the peak data rate. The signals are cleverly modified to help with this process, and so STBC support at the receiver is necessary to make sense of the received, overlapped signals. There is, however, a minor loss in average data rate noticeable for shorter packets, since the

[2] Cisco ClientLink: Optimized Device Performance with 802.11n:
http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps10092/white_paper_c11-516389.html

transmitted data must be padded to evenly full a multiple of two OFDM symbols (instead of just one). STBC improves SNR by 1 to 4 dB over Spatial Expansion, according to the data rate.

Figure 7. 2x1 STBC



In the absence of transmit beamforming, 2x1 STBC, when supported, is well-suited for access-point-to-handset transmissions, given that most handsets do not have room for a second 802.11n RF chain or antenna. However, this lack of a second antenna also prevents handsets from using STBC when transmitting to their associated access point.

STBCs can work well for both unicast and multicast/broadcast packets, since there is no customization for individual receivers. However, STBC is a physical-layer technology requiring silicon support that is optional in 802.11n. It takes just one legacy client—or one 802.11n client that does not support STBC—and STBC cannot be used for multicast/broadcast data transmissions. Moreover, STBC was disabled during the 802.11n draft 2.0 certification testing. Because STBC is only an optional component of the ratified certification, it is not available or certified in the majority of 802.11n clients. Therefore, the benefits of STBC to multicast/broadcast are unimportant for the current client mix.

Spatial Expansion

Spatial expansion (SE), also called **cyclic delay diversity**, is the rudimentary method to map a smaller number of spatial streams to a larger number of transmit antennas. SE is needed to avoid unintentionally sending too much energy in a random direction, which is what happens if the signals are just duplicated across antennas. It turns out SE provides some modest diversity benefits also.

SE has most benefit in certain environments where there are no long multipath echoes, only a few short echoes. It turns out that these are particularly variable channels: subcarriers tend to be affected by multipath the in the same way, so if one subcarrier is weak, most all subcarriers are weak, with the result that the SNR is poor. SE works by sending additional copies of the signal, like echoes, from different antennas, thereby synthesizing artificial multipath is synthesized at the transmitter. SE introduces greater variability in individual subcarriers so that the overall variability averages out. Irrespective of the physical channel, the receiver tends to see an “average” channel. However, spatial expansion is applied blindly, independently of the actual channel, so ultimately its benefits are rather modest.

We have seen vendors offering unfortunately labeled “3x3” devices, which turn out to be 2x3 devices plus cyclic delay diversity to synthesize a third transmit signal. Although the third transmit signal has its own RF chain and antenna, the input to the third RF chain is not extra information, so there is no speed increase.

Although Cisco Aironet® 802.11n access points offer spatial expansion as a matter of course, Cisco recommends transmit beamforming where possible, since transmit beamforming offers the greatest benefits for customers in difficult propagation environments or where reliability at range is most needed.

Antenna Selection

Antenna selection is most used in 802.11a/b/g access points. These typically have one RF chain so they lack the wherewithal for transmitting beamforming, STBC, or even SE. Using measures of packet error rate or received signal strength, antenna selection determines which antenna would be best to use to transmit to the client. Compared with transmit beamforming or STBC, antenna selection is a blunt instrument because it cannot optimize the signal on a per subcarrier basis, and the selection of the optimal antenna can require the transmission of many packets.

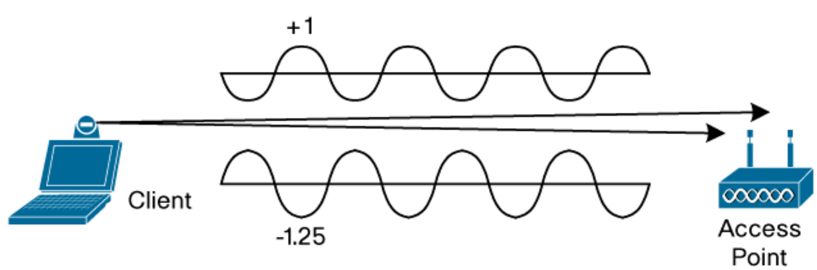
MIMO Equalizer

The MIMO equalizer is complementary to transmit beamforming. Whereas transmit beamforming enables the transmitter to make most use of its transmit antennas, the MIMO equalizer enables a receiver to combine the signals from its receive antennas, as shown in Figure 8. For the case of a single spatial stream, the MIMO equalizer is better known as maximum ratio combiner (MRC).

It comes back to the mirror-like (multipath) nature of the environment: if a spatial stream is sent from a single transmit antenna to a receive antenna, the stream follows a set of paths that may interfere constructively or destructively at the receive antenna. If there is a second receive antenna, the spatial stream follows a slightly different set of paths because there is some space between the two receive antennas. Sometimes one set of paths leads to destructive interference at one receive antenna, yet it is highly unlikely that both sets of paths simultaneously lead to destructive interference at both antennas. This behavior, called **spatial diversity**, provides improvements (captured within the MIMO equalizer) to SNR just as dramatic as with transmit beamforming (see the “All” curve in Figure 2).

MIMO equalization is a solid way to improve the reliability and predictability of wireless communications. 802.11n enterprise vendors including Cisco generally offer three receive antennas and RF chains, which provide significant benefits to uplink one and two spatial stream signals, such as laptop backups, voice services, and video surveillance.

Figure 8. MIMO Receiver Harvesting Multiple Signals



Spatial Division Multiplexing

The use of multiple antennas for spatial division multiplexing provides the second significant benefit of MIMO: namely, the ability to use each spatial stream to carry its own information, which provides dramatically increased data rates.

802.11a/b/g only allows the transmission of a single spatial stream at a time, which makes it like driving down a single lane road. With spatial division multiplexing, the single lane is transformed into a multilane highway.

Using spatial division multiplexing for two streams requires at least two transmit antennas and at least two receive antennas in the link. More generally, S spatial streams always require at least S transmit antennas at the transmitter and S receive antennas at the receiver. For instance, if the transmitter or receiver have only one antenna, the peak

802.11n physical layer rate possible in a 40-MHz channel is 150 Mbps. Yet if they both have two antennas and the required spatial division multiplexing functionality, the peak 802.11n rate is 300 Mbps.

Achieving the maximum data rate becomes progressively harder the more spatial streams that are available. A higher SNR is needed, so the devices need to be in closer proximity and the transmissions are more sensitive to colliding transmissions from non-Wi-Fi interferers and neighboring devices. A direct implementation of the highest rates requires upwards of 35 dB SNR and a nominal range of less than 20 feet. Thus for enterprise use, it is necessary to buttress spatial division multiplexing with a great many auxiliary improvements to improve the SNR. These additional improvements include having more antennas than spatial streams for transmit beamforming and/or MIMO equalization; improved error correction; unequal modulation (described later); the most optimal receiver math; and non-Wi-Fi interference mitigation.

The 802.11 draft 2.0 certification tested two spatial stream devices only. With the advent of the certification of ratified 802.11n, three spatial streams are optionally tested also. However three spatial streams offer limited incremental benefits over two spatial streams unless the stringent SNR requirements are addressed, and the current client mix is dominated by 802.11a/g and one- and two-spatial-stream 802.11n devices, and so three spatial streams are presently unwarranted.

MIMO Example

As an example to summarize these ideas, an access point with two transmit antennas transmitting to a client with two receive antennas has several options. If the access point and client are close together so that the SNR is already good, the access point can send two spatial streams to the client for highest throughput. This is analogous to viewing a hiking trail on a bright sunny day, and gaining greatest appreciation from seeing the scene in full color. Alternatively, if the access point and client are distant, so that the SNR is poor and needs improving, the access point should send only one spatial stream. In order to reliably use a higher data rate over that single spatial stream, the access point can use its otherwise spare antenna for transmit beamforming and/or the client can use its additional antenna to provide another input to its MIMO equalizer, thereby improving the SNR of the link. This is like following a hiking trail in dim light, where most detail is discerned by viewing the scene in black and white.

To summarize the benefits of MIMO:

- Benefits of extra receive antennas can be achieved even if the transmit side radio is 802.11a/b/g.
- Benefits of Cisco ClientLink can be achieved even if the receive side does not support MIMO or beamforming. However, standards-based beamforming requires support both on the transmit and receive side of a common beamforming mode
- The benefits of Spatial Division Multiplexing requires MIMO on both sides

802.11n Radio Enhancements

In addition to MIMO technology, 802.11n makes a number of additional changes to the radio to increase the effective throughput of the WLAN. The most important of these changes are increased channel width, higher modulation rates, and reduced overhead. This section will describe each of these changes and the effect they have on WLAN throughput.

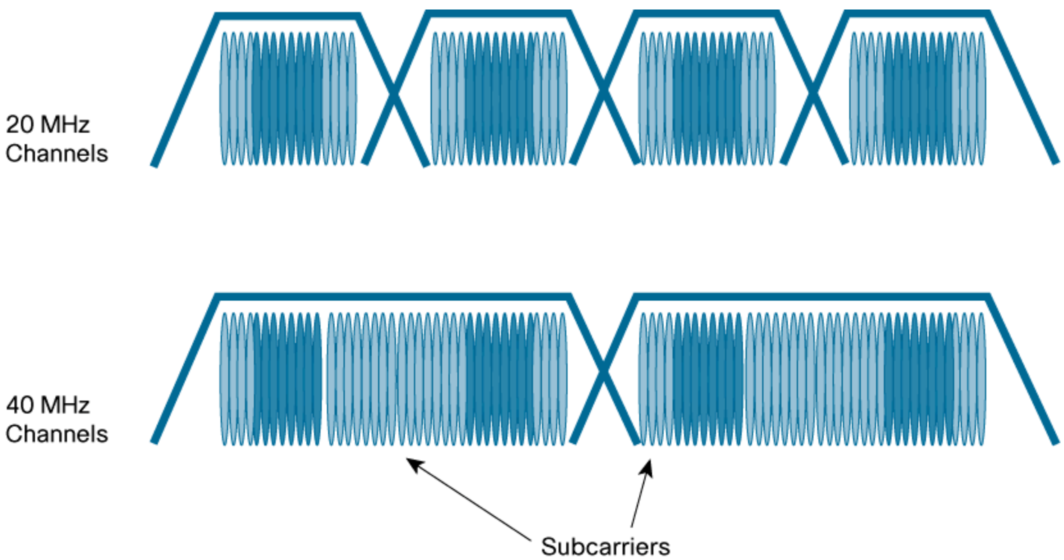
20- and 40-MHz Channels

The original 802.11 direct sequence radio and the 802.11b extension to the base standard use a radio bandwidth that is 22 MHz wide, with a radio channel spacing slightly higher, at 25 MHz. 802.11a and 802.11g both use radio bandwidths that are 20 MHz wide. Because 802.11g is an extension to 802.11b, 802.11g spaces its channels just as 802.11b does, every 25 MHz. The data rate per bandwidth of the radio channel is an important measure of the efficiency of the radio. This is called the **spectral efficiency** and is measured in bits per second per Hertz. The spectral efficiency of 802.11b is half a bit per second per Hertz (for example, 11 Mbps in 22 MHz). 802.11a and 802.11g have higher spectral efficiency: as much as 2.7 bits per second per Hertz at 54 Mbps.

Using exactly the same technology as 802.11a and 802.11g, some proprietary WLAN systems are available that provide up to 108 Mbps. These proprietary systems use a simple technique to double the data rate of 802.11a and 802.11g. They use two channels at the same time. This is called **channel bonding**. With channel bonding, the spectral efficiency is the same as 802.11a and 802.11g, but the channel bandwidth, now 40 MHz, is twice as great. This provides a simple way of doubling the data rate.

802.11n uses both 20-MHz and 40-MHz channels. Like the proprietary products, the 40-MHz channels in 802.11n are two adjacent, 20-MHz channels, bonded together. When using the 40-MHz bonded channel, 802.11n takes advantage of the fact that each 20-MHz channel has a small amount of the channel that is reserved at the top and bottom, to reduce interference in those adjacent channels. When using 40-MHz channels, the top of the lower channel and the bottom of the upper channel don't have to be reserved to avoid interference. These small parts of the channel can now be used to carry information. By using the two 20-MHz channels more efficiently in this way, 802.11n achieves slightly more than doubling the data rate when moving from 20-MHz to 40-MHz channels (see Figure 9), and a commensurate increase in spectral efficiency.

Figure 9. 20-MHz and 40-MHz Channels



Higher Modulation Rates

As 802.11 evolved from its 1 and 2 Mbps origins, though 802.11b (11 Mbps), and 802.11a/g (54 Mbps), the sophistication of the modulation method, and its spectral efficiency, improved exponentially. 802.11a and 802.11g adopted a method called **orthogonal frequency division multiplexing (OFDM)**. OFDM divides a radio channel into a large number of smaller channels, each with its own subcarrier signal, as shown in Figure 9. Each of these subcarrier signals is able to convey information independent of all the other subcarrier signals. It is roughly the same as having a group of independent radios bunched together.

For 802.11a and 802.11g, a symbol lasts 4 microseconds, including an 800 nanosecond **guard interval**. For the highest data rate, 54 Mbps, each symbol carries 216 data bits. These data bits are spread out over 48 subcarriers. In addition, there are 72 error-correction bits sent in each symbol at 54 Mbps, for a total of 288 bits in the symbol (that is, a rate 3/4 encoder). To pack this many bits on each subcarrier, the subcarrier is modulated using 64 quadrature amplitude modulation (QAM). This means that each subcarrier is able to carry 6 bits (a combination of data and error correction bits) per symbol.

802.11n continues the modulation evolution. 802.11n uses OFDM and a default symbol period of 4 microseconds, similar to 802.11a and 802.11g. However, 802.11n increases the number of subcarriers in each 20-MHz channel from 48 to 52. 802.11n provides a selection of eight data rates for a transmitter, including a data rate using 64 QAM with a new rate 5/6 encoder. Together these changes marginally increase the data rate to a maximum of 65 Mbps, for a single-transmit radio. Via spatial division multiplexing, 802.11n also increases the number of transmitters allowable to four. In total, 802.11n provides up to 32 data rates for use in a 20-MHz channel. For two transmitters, the maximum data rate is 130 Mbps. Three transmitters provide a maximum data rate of 195 Mbps. The maximum four transmitters can deliver 260 Mbps. Spectral efficiency is now raised a remarkable five times, and reaches a 13 bits per second per Hertz.

When using 40-MHz channels, 802.11n increases the number of subcarriers available to 108. This provides a maximum data rate of 135 Mbps, 270 Mbps, 405 Mbps, and 540 Mbps for one through four transmitters, respectively. Similarly, there are eight data rates provided for each transmitter, 32 in total, for the 40-MHz channel. Spectral efficiency reaches 13.5 bit/sec/Hz with the extra subcarriers.

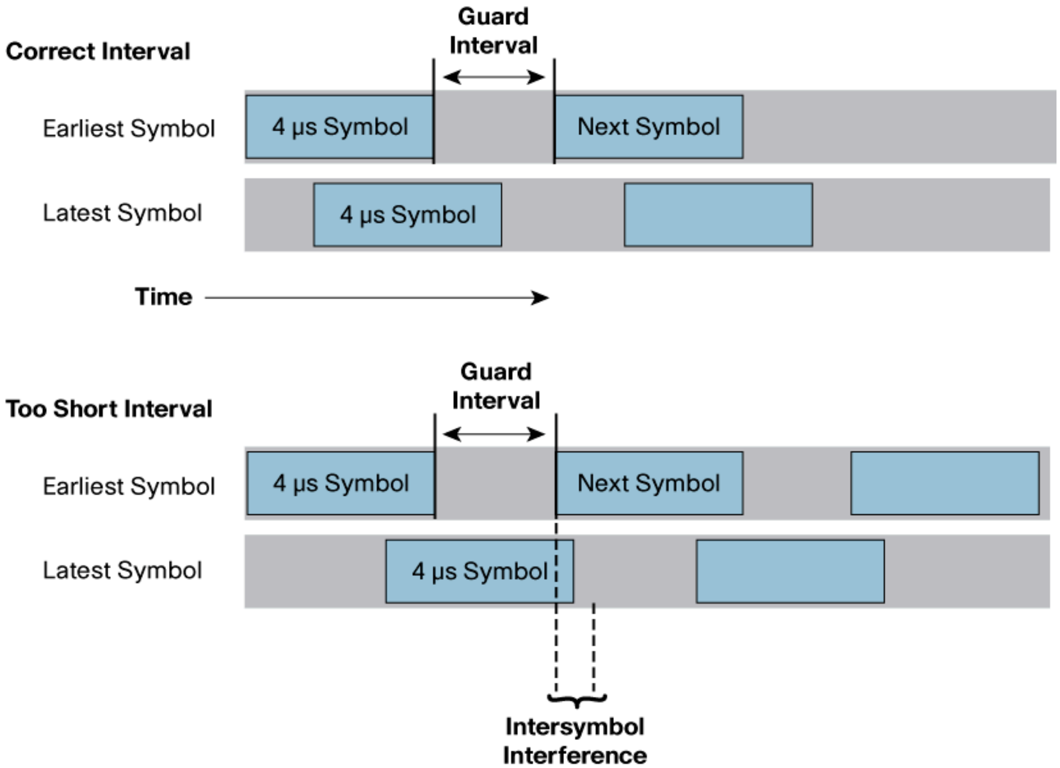
The rates described so far use the same modulation (called **equal modulation**) on all of the subcarriers: for example, all the subcarriers use quadrature phase shift keying (QPSK) or 64 QAM. This is the same as 802.11a and 802.11g. 802.11n adds the ability to modulate different spatial streams using different modulation methods: that is, some spatial streams use QPSK, some other spatial streams use 16 QAM, and yet other spatial streams use 64 QAM. This dramatically increases the number of data rates available to be used. In fact, there are dozens more possible data rates, using this **unequal modulation** method. However, it is unlikely that many practical implementations would be able to take advantage of this method, as it requires a significant amount of feedback from the receiver to the transmitter to identify the individual spatial streams that must use each of the different modulation methods.

As can be seen, the 40 MHz marks one of the most important methods of enhancing throughput in the 5-GHz mode and therefore customers are advised to keep their legacy clients on 2.4-GHz while moving their dual-band clients to 5-GHz either through manual configuration or innovative techniques such as bandsteering.

Lowered Overhead: Guard Interval

The **guard interval** that is part of each OFDM symbol is a period of time that is used to minimize intersymbol interference. This type of interference is caused in multipath environments when the beginning of a new symbol arrives at the receiver before the end of the last symbol is done. These two symbols arrive over two different paths. The "late" symbol has not yet been completely received when the new symbol arrives because it traveled a longer path than the new symbol (see Figure 10). When this situation occurs, the interference it causes reduces the effective SNR of the radio link. The guard interval is a quiet period between symbols that provides for the arrival of late symbols over long paths. The length of the guard interval is selected for the severity of the multipath environment. 802.11a and 802.11g use 800 nanoseconds as the guard interval, allowing for path differences of 800 feet.

Figure 10. Guard Interval



The default mode for 802.11n is also to use 800 nanoseconds as the guard interval. However, if the multipath environment is not as strict about a required allowance for 800 foot differences in paths between transmitter and receiver, 802.11n also provides a reduced guard interval of 400 nanoseconds. This reduces the symbol time from 4 microseconds to 3.6 microseconds. This reduced symbol time has a corresponding effect in increasing data rates. For 20-MHz channels, maximum data rates for one to four transmitters with the reduced guard interval are 72, 144, 216, and 288 Mbps for a 20- MHz channel and 150, 300, 450, and 600 Mbps for a 40-MHz channel.

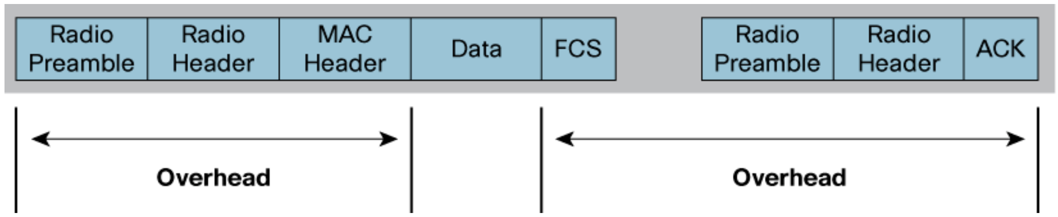
MAC Enhancements

There is only so much improvement that can be obtained in 802.11 by increasing the data rate of the radio. There is a significant amount of fixed overhead in the MAC layer protocol, and especially in the interframe spaces and acknowledgements of each frame transmitted. At the highest data rates, this overhead alone can be longer than the entire data frame. In addition, contention for the air and collisions also reduce the maximum effective throughput of 802.11. 802.11n addresses these issues by making changes in the MAC layer to improve on the inefficiencies imposed by this fixed overhead and by contention losses.

Frame Aggregation

Every frame transmitted by an 802.11 device has fixed overhead associated with the radio preamble and MAC frame fields that limit the effective throughput, even if the actual data rate was infinite (see Figure 11).

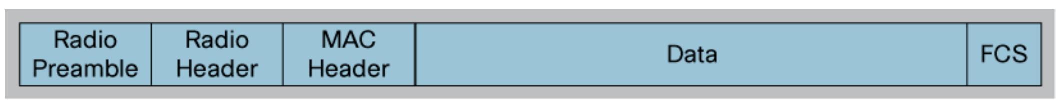
Figure 11. Overhead



To reduce this overhead, 802.11n introduces **frame aggregation**. Frame aggregation is essentially putting two or more frames together into a single transmission.

802.11n introduces two methods for frame aggregation: MAC Service Data Units (MSDU) aggregation and MAC Protocol Data Unit (MPDU) aggregation. Both aggregation methods reduce the overhead to only a single radio preamble for each frame transmission (see Figure 12). This is especially beneficial for customers transmitting lots of smaller packets, such as voice frames, TCP ACKs, and so forth.

Figure 12. Aggregation



Because multiple frames are now sent in a single transmission, the number of potential collisions and the time lost to backoff is significantly reduced. The maximum frame size is also increased in 802.11n, to accommodate these large, aggregated frames. The maximum frame size is increased from 4 KB to 64 KB.

One limitation of frame aggregation is that all the frames that are aggregated into a transmission must be sent to the same destination; that is, all the frames in the aggregated frame must be addressed to the same mobile client or access point.

Another limitation is that all the frames to be aggregated have to be ready to transmit from the client or access point at the same time, potentially delaying some frames to wait for additional frames, in order to attempt to send a single aggregate frame.

A third limitation of aggregation is that the maximum frame size that can be successfully sent is affected by a factor called **channel coherence time**. Channel coherence time depends on how quickly the transmitter, receiver, and other items in the environment are moving. The faster things are moving, the smaller the maximum frame size can be as the data rate is reduced—that is, the time for the transmission must be less than the channel coherence time.

There are slight differences in the two aggregation methods that result in differences in the efficiency gained. These two methods are described next.

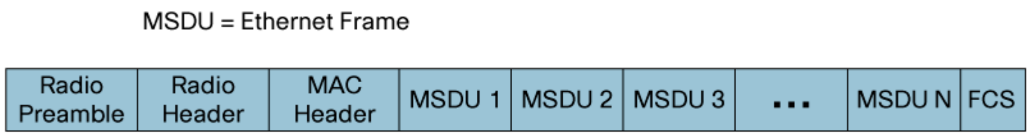
MAC Service Data Unit Aggregation

MSDU aggregation is the more efficient of the two aggregation methods. It relies on the fact that an access point receives frames from its Ethernet interface, to be translated to 802.11 frames and then transmitted to a mobile client. Similarly, most mobile client protocol stacks create an Ethernet frame, which the 802.11 driver must translate to an 802.11 frame before transmission. In both these cases, the "native" format of the frame is Ethernet, and it is then translated to 802.11 format for transmission.

Theoretically, MSDU aggregation allows frames for many destinations to be collected into a single aggregated frame for transmission. Practically, however, MSDU aggregation collects Ethernet frames for a common destination, wraps the collection in a single 802.11 frame, and then transmits that 802.11-wrapped collection of Ethernet frames (see

Figure 13). This method is more efficient than MPDU aggregation up to the maximum allowed MSDU aggregate size of 3839 or 7935 bytes (according to client capability), because the Ethernet header is much shorter than the 802.11 header.

Figure 13. MSDU Aggregation



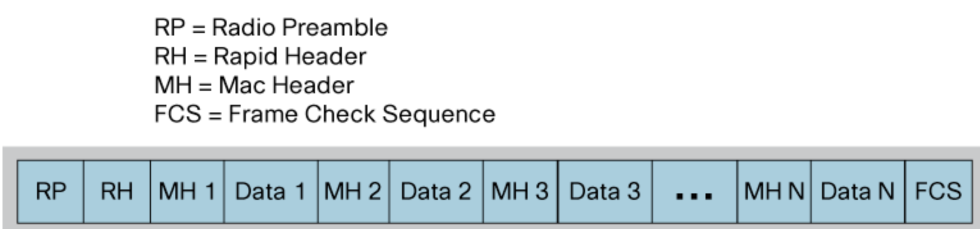
For a mobile device, the aggregated frame is sent to the access point, where the constituent Ethernet frames are forwarded to their ultimate destinations. For an access point, all of the constituent frames in the aggregated frame must be destined to a single mobile client, since there is only a single destination in each mobile client.

With MSDU aggregation, the entire, aggregated frame is encrypted once using the security association of the destination of the outer 802.11 frame wrapper. A restriction of MSDU aggregation is that all of the constituent frames must be of the same QoS level. It is not permitted to mix voice frames with best-effort frames, for example.

MAC Protocol Data Unit Aggregation

MPDU aggregation is slightly different from MSDU aggregation. Instead of collecting Ethernet frames, MPDU aggregation translates each Ethernet frame to 802.11 format and then collects the 802.11 frames for a common destination. The collection doesn't require a wrapping of another 802.11 frame, since the collected frames already begin with an 802.11 MAC header (see Figure 14).

Figure 14. MPDU Aggregation



MPDU aggregation does require that all the 802.11 frames that constitute the aggregated frame have the same destination address. However, this results in the same behavior as MSDU aggregation, since the destination of all frames sent by a mobile client is that client's access point, where the 802.11 frames are translated to Ethernet and forwarded to the ultimate destination. Similarly, the destination of any frame sent by the access point is a single mobile client.

With MPDU aggregation, it is possible to encrypt each constituent frame independently, using the security association for each individual 802.11 destination address. This does not have any effective difference from the encryption done in MSDU aggregation, as all frames sent by a mobile client are encrypted using the security association for the access point, and all frames sent by the access point are encrypted using the security association for the single mobile client that is the intended recipient of the frame.

Similar to MSDU aggregation, MPDU aggregation requires that all of the constituent frames be of the same QoS level.

The efficiency of the MPDU aggregation method is lower than that of the MSDU aggregation method for short and medium-sized aggregates, because of the extra overhead of the individual 802.11 frame headers for each constituent frame. The efficiency is further reduced when the encryption is used. Encryption adds overhead to each of the constituent frame in MPDU aggregation, where MSDU aggregation incurs overhead for a single encryption of the outer 802.11 wrapper. On the other hand, MPDU aggregation can create frames up to 65,535 bytes (according to client capability), so it is the preferred scheme when lots of data is available for aggregation.

MAC aggregation requires both the transmit and receive side to support 802.11n. However, the benefits are significant.

Block Acknowledgement

For the 802.11 MAC protocol to operate reliably, each of the frames that is transmitted an individual address—i.e. not multicast or broadcast frames—is immediately acknowledged by the recipient. MSDU aggregation requires no changes to this operation. The aggregated frame is acknowledged, just as any 802.11 frame is acknowledged. This is not the case for MPDU aggregation. For MPDU aggregation, each of the individual, constituent 802.11 frames must be acknowledged. The mechanism to deal with this requirement that 802.11n introduces is called **block acknowledgement**.

Block acknowledgement compiles all the acknowledgements of the individual, constituent frames produced by MPDU aggregation into a single frame returned by the recipient to the sender. This allows a compact and rapid mechanism to implement selective retransmission of only those constituent frames that are not acknowledged. In environments with high error rates, this selective retransmission mechanism provides improvement in the effective throughput of a WLAN using MPDU aggregation over that of one using MSDU aggregation, because much less is retransmitted when an error affects some of the constituent frames of an MPDU aggregated frame as compared to an MSDU aggregated frame.

Certification of Aggregation

The 802.11n certification verifies that all devices are able to receive frames using both MPDU and MSDU aggregation, but does not test that devices are able to transmit frames using either MSDU aggregation. This need not be a concern for customers of Cisco Aironet 802.11n access points, since these access points are able to both transmit and receive frames using MSDU and MPDU aggregation. Cisco Aironet access points select MSDU aggregation for flows with short packets while also enabling dynamic selection between MSDU and MPDU aggregation for flows with longer packets in order to offer the greatest reliability, as well as low-latency and efficiency.

Certification of ratified 802.11n includes testing devices that transmit MDPU aggregation. That is, despite its superior efficiency, MSDU aggregation is not a required feature of all vendors' access points.

Lower Overhead: Reduced Interframe Space

When aggregation of frames is not possible, 802.11n provides a mechanism to reduce the overhead involved with transmitting a stream of frames to the same destination, or different destinations if using PSMP. This mechanism reduces the interframe space between sequences of transmitted frames. The 802.11e extension for quality of service added the ability for a single transmitter to send a burst of frames during a single, timed **transmit opportunity**. During the transmit opportunity, the sender does not need to perform any random backoff between transmissions, separating its frames by the smallest allowable interframe space, the short interframe space (SIFS).

802.11n improves on this mechanism, reducing the overhead between frames, by specifying an even smaller interframe space, called the **reduced interframe space (RIFS)**. RIFS cuts down further on the dead time between frames, increasing the amount of time in the transmit opportunity that is occupied by sending frames.

The two unfortunate aspects of using RIFS are first, aggregation is more efficient that RIFS when transmitting to the same destination; and second, RIFS is restricted to being used only in greenfield deployments—that is, only in deployments in which there are no legacy 802.11a, b, or g devices in the area.

Power Savings

RF chains in a MIMO radio are power hungry. Operating several RF chains requires even more power. To address this situation, 802.11n has extended the power management capability of the 802.11 MAC. There are two extensions beyond the existing mechanisms established in the original standard and the automatic power save delivery added in 802.11e. The two new mechanisms provided by 802.11n are Spatial Multiplexing Power Save mode and Power Save Multi-Poll mode.

Spatial Multiplexing Power Save

The Spatial Multiplexing (SM) Power Save mode allows an 802.11n client to power down all but one of its RF chains. This power save mode has two submodes of operation: static operation and dynamic operation.

The static SM power save mode has the client turn off all but a single RF chain, becoming essentially equivalent to an 802.11a or 802.11g client. The client's access point is notified that the client is now operating in the static, single-RF chain mode, requiring the access point to send only a single spatial stream to this client until the client notifies the access point that its additional RF chains are again enabled and operating. This notification of the access point is done using a new management frame, defined by 802.11n, telling the access point that the client is in static SM power save mode.

The dynamic SM power save mode also turns off all but one of the client's RF chains. But in this mode of operation, the client can rapidly enable its additional RF chains when it receives a frame that is addressed to it. The client can immediately return to the low power state by disabling its additional RF chains immediately after its frame reception is complete. In this mode of operation, the access point typically sends a request-to-send (RTS) frame to the client, to wake its RF chains, prior to sending the client a data or management frame. On receiving the RTS frame, the client enables its RF chains and responds with a **clear-to-send** (CTS) frame. All of its RF chains are now ready to receive the multiple spatial streams sent by all the RF chains in the access point. To use this power save mode, the 802.11n client sends a new management frame to its access point, informing the access point that it is in dynamic SM power save mode.

Power Save Multi-Poll

The Power Save Multi-Poll (PSMP) mode is a polling mechanism somewhat like HCF controlled channel access (HCCA), where HCF denotes hybrid coordination function. This mechanism is defined in 802.11e. Using HCCA, an access point accesses the channel using its native high priority in order to reserve the channel for an extended duration, called a contention free period (CFP). During the CFP, the access point polls individual associated clients. This scheduling mechanism reduces the contention between clients and between the client and the access point. Reducing contention also reduces the time the client spends in backoff and reduces the number of times a frame must be transmitted before it is delivered successfully. However, since clients can be polled at any point, they must remain awake and consuming power for the duration of the CFP.

With PSMP, the access point begins by transmitting a schedule of times allocated for downlink (access point to client) and uplink (client to access point) broadcast, multicast and unicast transmissions. Clients can immediately determine when they are required to be awake, and consequently can doze for the remaining duration. In this way, PSMP offers the reduced contention of a polling mechanism, while also offering clients a power saving opportunity. Furthermore, RIFS is allowed during PSMP for still greater efficiency.

Nonetheless, PSMP is not as effective as unscheduled automatic power-save delivery (U-APSD) from 802.11e for saving power since PSMP is driven by the access point— PSMP requires that the client remain awake to receive the PSMP schedule.

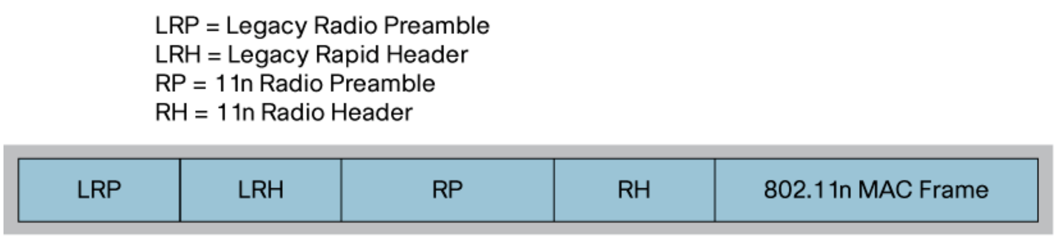
Backward Compatibility

Compatibility with existing 802.11a, b, and g devices is a critical issue addressed in 802.11n. Just as 802.11g provides a protection mode for operation with 802.11b devices, 802.11n has a number of mechanisms to provide backward compatibility with 802.11 a, b, and g devices, allowing these devices to understand the information necessary to allow 802.11n devices to operate in the same area.

For quite a long time, 802.11n will need to operate in the presence of legacy 802.11a, b, and g devices. This mixed-mode operation will continue until all the devices in an area have been upgraded or replaced with 802.11n devices. The mixed-mode protection mechanism for 802.11n is quite similar to the protection mechanism of 802.11g.

Like 802.11g, 802.11n transmits a header that can't be decoded by devices built to an earlier standard. To avoid descending into chaos in the presence of massive interference and collisions, 802.11n operating in mixed mode transmits a mixed-mode packet format, called **HT-mixed**, that begins with a radio preamble and header field that can be decoded by 802.11a and 802.11g radios (see Figure 15). This provides enough information to the legacy radios to allow them to indicate that there is another transmission on the air and how long that transmission will last. Following the legacy preamble and header field, the 802.11n device sends the remaining information using 802.11n rates and its multiple spatial streams, including an 802.11n preamble and header field.

Figure 15. HT-Mixed Frame Format

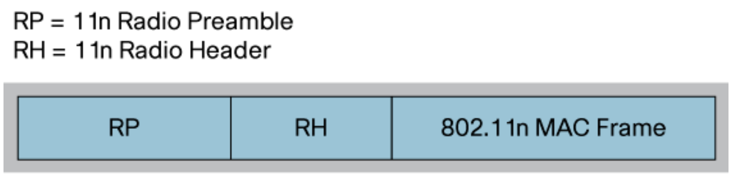


In addition to the legacy preamble and header field, it can also be necessary to use additional protection mechanisms provided by 802.11g to allow the MAC in legacy devices to correctly determine when it is allowed to transmit and when it must perform backoff before transmission. The mechanism provided by 802.11g and utilized by 802.11n when either 802.11g or 802.11a devices are present is the CTS-to-self mechanism. CTS-to-self allows the 802.11n device to transmit a short CTS frame, addressed to itself, that includes the timing information necessary to be communicated to the neighboring legacy MACs that will protect the 802.11n transmission that will follow. The CTS frame must be transmitted using one of the legacy data rates that a legacy device will be able to receive and decode.

The cost of this additional legacy preamble and header field, as well as the CTS-to-self, is more overhead on every 802.11n transmission. This reduces the benefits of all the 802.11n improvements, resulting in significantly lower effective throughput by 802.11n devices in mixed environments. Similar to 802.11g, legacy devices don't need to be associated to the same access point as an 802.11n device to require the use of protection mechanisms. If there are legacy devices on the same channel on any nearby access points, they will cause protection mechanisms to be invoked as well.

Once protection mechanisms are not needed, there is also a non-backward-compatible packet format, called **HT-greenfield**, defined by 802.11n and shown in Figure 16. HT-greenfield dispenses with the legacy preamble and header fields. It reduces the packet overhead by 12 microseconds, so it can improve efficiency relative to HT-mixed format for short and medium length payloads that cannot be aggregated further. On the other hand, HT-greenfield format is still 4 microseconds longer than the 802.11a/g packet format. Thus HT-greenfield format is narrowly beneficial for (possibly aggregated) medium length packets sent using multiple spatial streams.

Figure 16. HT-Greenfield Frame Format



It can be expected that protection mechanisms will be in use in the 2.4-GHz band (802.11b and 802.11g) until nearly every legacy device has disappeared. This is because there are too few channels available in that band to effectively overlay pure 802.11n WLANs in the same areas as legacy 2.4-GHz WLANs. Given the larger number of channels available in the 5-GHz band in many countries, it is possible that two completely separate WLANs could be operating in the same area, with 802.11a operating on one set of channels and 802.11n operating on a different, nonintersecting set of channels. This would result in 802.11n operating in pure high-throughput (greenfield) mode, achieving the highest effective throughput offered by this new standard.

Wireless Intrusion Detection Systems

Attackers are at an advantage in sites lacking 802.11n access points. This is the case because wireless intrusion detection systems (wIDS) have a reduced ability to discover rogues and classify transmissions since 802.11n transmissions are not understood by 802.11a/b/g access points. Although rogue 802.11n access points should transmit beacons at legacy rates (in order to be discovered by legacy clients), this behavior can be disabled by configuration without precluding certain kinds of attacks. For this reason, customers with security concerns are advised to upgrade to the latest 802.11n access points.

Backward Compatibility: Use of 40 MHz in 2.4 GHz

In most parts of the world, the 2.4-GHz band provides at most three or four nonoverlapping channels for 802.11b/g. The use of 40 MHz in this band poses multiple problems:

- Legacy equipment is most commonly deployed on channels 1, 6 and 11—that is, for channels that are 25 MHz apart, yet the mixed-mode preamble provides protection for bonded channels that are 20 MHz apart.
- There is a small percentage of equipment occupying intermediate channels (such as channels 2, 3, or 4) for which the mixed-mode preamble does not help; nor does CTS-to-self sent using higher data rates. Even CTS-to-self sent at 1 or 2 Mbps may help only partially.
- There is a perception that the use of a 40-MHz channel by 802.11n at 2.4 GHz may limit the available spectrum available to other systems in 2.4 GHz (such as Bluetooth, although the experimental evidence is less clearcut.

For these reasons, 802.11n sets a high bar on the use of 40 MHz at 2.4 GHz: there must be no legacy access points on overlapped channels (except for the primary channel) before an 802.11n access point can use 40 MHz. Even while using 40 MHz, the access point and its clients must regularly scan for overlapping legacy access points. It is even possible for a nearby 802.11b/g/n device to order an 802.11n access point to stop using 40 MHz. Finally, an 802.11n access point that can detect non-802.11 communication devices (such as Bluetooth or Zigbee)—and that does detect them—cannot use 40 MHz unless the access point also implements a coexistence mechanism.

We see that 802.11n allows 40-MHz operation at 2.4 GHz, yet in a careful, “good-neighbor” fashion. These requirements do make it very difficult for 40 MHz to be used in practice. This is as it should be, since 40 MHz is best used in isolated locations where overlapped access points and legacy access points are uncommon. In the enterprise, there are two additional concerns:

- The 2.4-GHz band supports only a single 40-MHz channel plus a 20- (or 25-) MHz channel, so it is a significant challenge to find a sensible channel plan with 40 MHz.
- There is a lot of overhead associated with a CTS-to-self packet sent at 1 or 2 Mbps, so that one 40-MHz channel is a less efficient use of the band than two 20 MHz channels. There ends up being twice as much overhead, and the overhead is not negligible

Due to the limited likelihood of being allowed to use 40 MHz, and the detrimental effect 40 MHz has on the available capacity of an enterprise network, Cisco Aironet 802.11n access points are not certified for 40-MHz operation at 2.4 GHz. Rather, the place for 40 MHz is at 5 GHz.

Chipsets

An access point offers features implemented via a combination of hardware and software. Cisco is the only vendor that drives its specific requirements into creating a custom, enterprise-grade chipset for its access points that solves significant customer problems. Other vendors use off-the-shelf, consumer-grade chipsets, usually with poor implementation of many features. Furthermore, the software for a Cisco Aironet access point is custom-written to optimize enterprise requirements, whereas other vendors retrofit enterprise requirements into a consumer-optimized code base. Therefore, performance cannot be judged merely through a checklist of features but by truly deploying the equipment in an enterprise-grade environment.

Summary of 802.11n Technology

To summarize the benefits of 802.11n technology, it is simplest to say that there are two major areas of improvement over previous 802.11 devices. The first area of improvement is in the use of MIMO technology to achieve spatial division multiplexing and greater SNR on the radio link. The second area of improvement is in the greater efficiencies in both radio transmissions and the MAC protocol. These improvements translate into benefits in four areas: reliability, predictable coverage, raw speed, and throughput.

Reliability

Greater SNR on the radio link translates directly to more reliable communication, often at higher data rates. Higher SNR means that more interference is needed to corrupt a transmission. This means greater client densities can be supported.

Predictable Coverage

The use of multiple spatial streams provided by MIMO technology means that there will be fewer dead spots in a coverage area. Areas that previously suffered from destructive multipath interference now make use of that same multipath effect to provide robust communication.

Raw Speed

Spatial division multiplexing, 40-MHz operation and the short guard interval together provide considerable improvement in raw speed over 802.11a or g.

Throughput

The efficiency improvement in the 802.11 MAC provides a greater transfer of the high bit rates of the 802.11n radio to effective throughput seen by actual applications, at least in greenfield deployments. Even in mixed-mode deployments with legacy 802.11 devices, 802.11n will provide greater effective throughput, although significantly less than the greenfield mode.

Migration to 802.11n

The migration to 802.11n is well underway, triggered first by the certification of 802.11n draft 2.0, and second by the certification of ratified 802.11n. 802.11n client devices, beginning with laptops, are already available. New client devices will continue to appear over the next several quarters, until 802.11n is the default WLAN adapter in any portable or mobile device. 802.11n is already the default wireless adapter in laptops. These client devices are completely compatible with existing 802.11a, b, and g access points and will operate just as existing devices do today. Planning to migrate the infrastructure portion of a network to support 802.11n on these new client devices is straightforward.

Planning

There are several areas to consider when planning the migration of a network to support 802.11n. Because of the higher speeds and greater power requirements of 802.11n access points, planning the migration needs to take into account more than just the access point.

Radio Bands

802.11n operates in both the 2.4-GHz (802.11b and g) and 5-GHz (802.11a) radio bands. Planning for each of the radio bands should be done independently, because of the constraints that are sometimes very different for each band.

The 2.4-GHz Band

The 2.4-GHz band is no more than 100 MHz wide, and much less than that in many countries. The same channelization that is used for 802.11b and 802.11g should be used for 802.11n operating in this band. However, the use of the 40-MHz mode of operation of 802.11n is not recommended in this band, because a significant portion of the band will suffer from interference from a single 40-MHz transmitter. In addition, it is required that the second 20-MHz channel, concatenated with the original 20 MHz channel to form the 40-MHz channel, must be free of any legacy transmissions. This drastically reduces the chance that any 40 MHz operations will be feasible in this band.

In much of the world where it is typical to utilize three nonoverlapping channels in this band, a single 40-MHz access point will present a significant challenge to developing a channel plan that will provide adequate capacity in most enterprises. Even when all legacy 802.11b and g devices are removed from the band, it will be difficult to deploy access points utilizing the 40-MHz channels in this band. There is just not enough bandwidth available to even begin to duplicate the three nonoverlapping channels of the legacy layout.

The 5-GHz Band

The 5-GHz band has been opened up significantly in much of the world, due to recent changes by many regulatory agencies. There are significantly more channels available in the 5-GHz band than in the 2.4-GHz band. The larger number of channels in this band makes planning the deployment of an 802.11n network much simpler, even while allowing for 40-MHz operation.

There are at least two possible ways to migrate to 802.11n in the 5-GHz band. The first way is to replace individual legacy access points with 802.11n access points as budget allows and user demand for additional capacity dictates. This gradual migration can be accomplished over a planned period of time or as the need arises. This migration method would have the new 802.11n access point operating on the channel of the legacy access point it replaced. The new 802.11n access point would support 802.11n clients, as well as legacy 802.11a clients. It would operate in mixed mode, providing protection for the legacy 802.11a clients. Eventually, as the last legacy access point is replaced and the last legacy client is retired, the entire set of new 802.11n access points could be switched to operate in greenfield mode.

The second way to migrate to 802.11n would be to reassign the channels on some of the legacy access points to free a set of channels that could be used for 802.11n exclusively. Then as budget allows and demand dictates, new

802.11n access points would be added to the existing WLAN deployment, operating in parallel in overlapping areas with the legacy access points. The new 802.11n access points, however, would support only 802.11n devices and be able to operate in the greenfield mode, providing the greatest benefits of the new standard. Eventually, as 802.11n access points cover the entire area covered by the legacy access points-802.11n clients would have the ability to operate in greenfield mode everywhere, while the legacy access points still provide service to the legacy clients. Once the last legacy client is retired, the legacy access points can also be retired.

Wired Infrastructure Stresses

Today's dual-band access points can theoretically put a load on their Ethernet connections of as much as 108 Mbps. Practically, however, due to the inefficiencies of the 802.11 protocol, they top out at a peak load of 50 to 60 Mbps.

802.11n access points can demand much more of their Ethernet connection. With the higher bit rates on the air and the improved efficiency of the protocol, it is possible that a single, dual-band, 802.11n access point supporting a 20-MHz channel in the 2.4-GHz band and a 40-MHz channel in the 5-GHz band can place a peak demand on its Ethernet connection of as much as 300 to 400 Mbps. Obviously, this is greater than a single or double 100-Mbps Ethernet connection can support.

For this reason, planning for a migration to support 802.11n should also include planning to upgrade the edge Ethernet switching capabilities to support a 1-Gbps connection to each 802.11n access point. This will eliminate any bottlenecks that might occur in areas of high-capacity demand by the 802.11n clients.

Power Requirements

Most current 802.11 access points can be powered using Power over Ethernet (PoE) or 802.3af switches. As chipsets progressively add more functionality and potentially become more power hungry there are two ways of solving this. The first one is the 802.3at standard from the IEEE 802.3 (Ethernet) Working Group that provides 30 watts at the power source. The others are innovating at the chipset or optimizations in technology in order to do more within the 802.3af power budget. The niche alternatives of using power injectors or power supplies also exist.

Access Point Deployment

Planning the positioning of the 802.11n access points should also be considered during migration planning. If the migration plan is to gradually replace the existing legacy access points, there is no further deployment planning necessary. However, if the new 802.11n access points are to be deployed in a new installation or along side an existing deployment, it is possible to use the increased SNR provided by 802.11n to provide greater throughput for clients or to cover greater areas per access point. Remember that SNR is like money in the bank. It can be used for either increased data rate, increased range, or a little of both. But it can't be used for the maximum of both at the same time.

Promises and Expectations

There is great promise of robustness and throughput in the 802.11n standard. Robustness of 802.11n could be measured in application availability and reliability of the link. Only when operating in a mode where all clients and access points support all functionalities can an 802.11n deployment reach its full potential.

Since most of the legacy devices continue to operate in the 2.4-GHz band and congest it, 802.11n can increase robustness and throughput only somewhat in that band. Meanwhile in the 5-GHz band, there will be a major thrust from cost-optimized netbooks and handhelds with 802.11n devices but only one or two spatial streams. For this reason, tests measuring the peak throughput of 802.11n for comparison are much less relevant. To achieve the best performance from their wireless networks, customers need careful planning of their applications and devices coupled with optimization around real-life scenarios such as mixed-mode clients.

Conclusions

With the ratification of 802.11n, customers can start deploying new 802.11n clients, and migrating to or adding to their 802.11n network infrastructure with confidence. Some care, however, should be taken when selecting the 802.11n equipment to install. Throughput tests in a controlled environment rarely indicate the best experience over time. Those same access points should be tested against the widest variety of client devices possible. It's important to look for consistent rate vs. range and to provide the maximum available frequencies, as well as standards-based management protocols and performance optimized in several scenarios like mixed mode, high density and so on.

802.11n has the ability to dramatically increase the capacity of a WLAN, the effective throughput of every client, and the reliability of the client's networking experience. The time to begin moving to this new standard is as soon as it is necessary to add a new access point, in order to address the demand for additional capacity or reliability in the WLAN.

Printed in USA C11-427843-00 09/09

