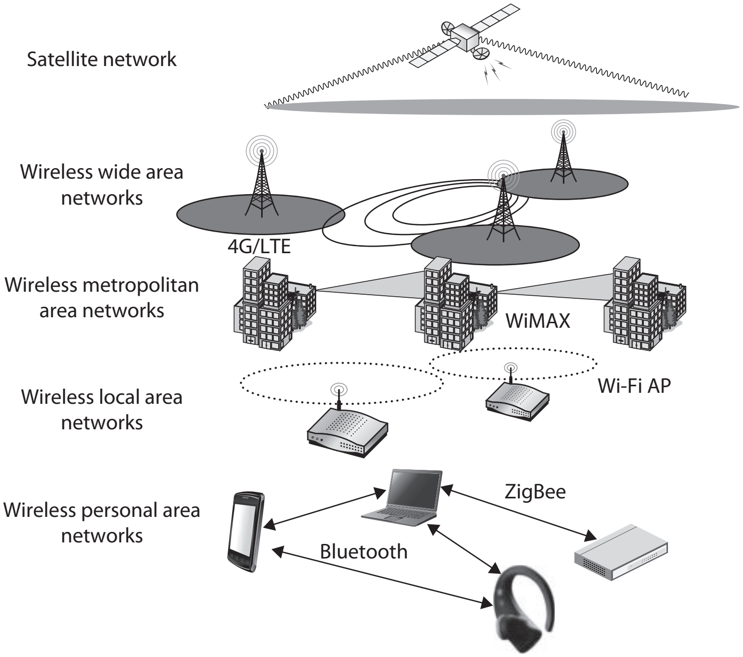
Figure 12-1 Various wireless networks


Satellite network
Wireless wide area networks
4G/LTE
Wireless metropolitan area networks
WiMAX
Wireless local area networks
Wi-Fi AP
Wireless personal area networks
ZigBee
Bluetooth

The *amplitude* of a radio signal indicates its power, which in turn dictates how far it can go. Amplitude usually is measured in watts or milliwatts (one-thousandth of a watt), but you may also see it expressed in decibels per milliwatt (dBm or dBmW), which is a measure of comparison to one milliwatt. For example, a wireless access point may allow you to configure the transmit power in increments from 0 dBm (1 mW) up to 23 dBm (200 mW).

In a wired network, each computer and device has its own cable connecting it to the network in some fashion. In wireless technologies, each device must instead share the allotted radio frequency spectrum with all other wireless devices that need to communicate. This spectrum of frequencies is finite in nature, which means it cannot grow if more and more devices need to use it. The same thing happens with Ethernet—all the computers on a segment share the same medium, and only one computer can send data at any given time. Otherwise, a collision can take place. Wired networks using Ethernet employ the CSMA technology (described in Chapter 11). Wireless LAN (WLAN) technology is actually very similar to Ethernet, but it uses CSMA/CA (collision avoidance). The wireless device sends out a broadcast indicating it is going to transmit data. This is received by other devices on the shared medium, which causes them to hold off on transmitting information. It is all about trying to eliminate or reduce collisions.

A number of techniques have been developed to allow wireless devices to access and share this limited amount of medium for communication purposes. The goal of each of these wireless technologies is to split the available frequency into usable portions, since it is a limited resource, and to allow the devices to share those portions efficiently. The most popular approach is called spread spectrum, though orthogonal frequency division multiplexing (OFDM) is widely used also. Because spread-spectrum technology is so prevalent, we'll get into it in a fair amount of detail in the next section.

Spread Spectrum

Radio frequencies cover a wide range, or *spectrum*, of frequencies. Some parts of this spectrum are allocated by national governments or international agreements for specific purposes. A radio *frequency band* is a subset of the radio spectrum designated for a specific use. For example, the range of radio frequencies between 1.8 and 29.7 megahertz (MHz) is almost universally considered the amateur radio band. A well-known frequency band is frequently labeled using just a single frequency, such as when we refer to the 2.4-GHz band used in many Wi-Fi systems. This band actually corresponds to the range of frequencies between 2.4 and 2.5 GHz. The challenge is how to dynamically allocate individual frequencies to specific sets of transmitters and receivers without them stepping all over each other. This is where spread-spectrum techniques come in handy.

Spread spectrum means that something is distributing individual signals across the allocated frequencies in some fashion. So, when a spread-spectrum technique is used, the sender spreads its data across the frequencies over which it has permission to communicate. This allows for more effective use of the available spectrum, because the sending system can use more than one frequency at a time.

Think of spread spectrum in terms of investments. In conventional radio transmissions, all the data is transmitted on a specific frequency (as in amplitude modulated [AM] radio systems) or on a narrow band of frequencies (as in frequency modulated [FM] radio). This is like investing only in one stock; it is simple and efficient, but may not be ideal in risky environments. The alternative is to diversify your portfolio, which is normally done by investing a bit of your money in each of many stocks across a wide set of industries. This is more complex and inefficient, but can save your bottom line when the stock in one of your selected companies takes a nose-dive. This example is akin to direct sequence spread spectrum (DSSS), which we discuss in an upcoming section.

There is in theory another way to minimize your exposure to volatile markets. Suppose the cost of buying and selling was negligible. You could then invest all your money in a single stock, but only for a brief period of time, sell it as soon as you turn a profit, and then reinvest all your proceeds in another stock. By jumping around the market, your exposure to the problems of any one company are minimized. This approach would be comparable to frequency hopping spread spectrum (FHSS), discussed next. The point is that spread-spectrum communications are used primarily to reduce the effects of adverse conditions such as crowded radio bands, interference, and eavesdropping.

Frequency Hopping Spread Spectrum

Frequency hopping spread spectrum (FHSS) takes the total amount of spectrum and splits it into smaller subchannels. The sender and receiver work at one of these subchannels for a specific amount of time and then move to another subchannel. The sender puts the first piece of data on one frequency, the second on a different frequency, and so on. The FHSS algorithm determines the individual frequencies that will be used and in what order, and this is referred to as the sender and receiver's *hop sequence*.

Interference is a large issue in wireless transmissions because it can corrupt signals as they travel. Interference can be caused by other devices working in the same frequency space. The devices' signals step on each other's toes and distort the data being sent. The FHSS approach to this is to hop between different frequencies so that if another device is operating at the same frequency, it will not be drastically affected. Consider another analogy: Suppose George and Marge work in the same room. They could get into each other's way and affect each other's work. But if they periodically change rooms, the probability of them interfering with each other is reduced.

A hopping approach also makes it much more difficult for eavesdroppers to listen in on and reconstruct the data being transmitted when used in technologies other than WLAN. FHSS has been used extensively in military wireless communications devices because the only way the enemy could intercept and capture the transmission is by knowing the hopping sequence. The receiver has to know the sequence to be able to obtain the data. But in today's WLAN devices, the hopping sequence is known and does not provide any security.

So how does this FHSS stuff work? The sender and receiver hop from one frequency to another based on a predefined hop sequence. Several pairs of senders and receivers can move their data over the same set of frequencies because they are all using different hop sequences. Let's say you and Marge share a hop sequence of 1, 5, 3, 2, 4, and Nicole and Ed have a sequence of 4, 2, 5, 1, 3. Marge sends her first message on frequency 1, and Nicole sends her first message on frequency 4 at the same time. Marge's next piece of data is sent on frequency 5, the next on 3, and so on until each reaches its destination, which is your wireless device. So your device listens on frequency 1 for a half-second, and then listens on frequency 5, and so on, until it receives all of the pieces of data that are on the line on those frequencies at that time. Ed's device is listening to the same frequencies but at different times and in a different order, so his device never receives Marge's message because it is out of sync with his predefined sequence. Without knowing the right code, Ed treats Marge's messages as background noise and does not process them.

Direct Sequence Spread Spectrum

Direct sequence spread spectrum (DSSS) takes a different approach by applying sub-bits to a message. The sub-bits are used by the sending system to generate a different format of the data before the data is transmitted. The receiving end uses these sub-bits to reassemble the signal into the original data format. The sub-bits are called *chips*, and the sequence of how the sub-bits are applied is referred to as the *chipping code*.

When the sender's data is combined with the chip, the signal appears as random noise to anyone who does not know the chipping sequence. This is why the sequence is sometimes called a pseudo-noise sequence. Once the sender combines the data with the chipping sequence, the new form of the information is modulated with a radio carrier signal, and it is shifted to the necessary frequency and transmitted. What the heck does that mean? When using wireless transmissions, the data is actually moving over radio signals that work in specific frequencies. Any data to be moved in this fashion must have a carrier signal, and this carrier signal works in its own specific range, which is a frequency. So you can think of it this way: once the data is combined with the chipping code, it is put into a car (carrier signal), and the car travels down its specific road (frequency) to get to its destination.

Spread Spectrum Types

This technology transmits data by "spreading" it over a broad range of frequencies:

- FHSS moves data by changing frequencies.
- DSSS takes a different approach by applying sub-bits to a message and uses all of the available frequencies at the same time.

The receiver basically reverses the process, first by demodulating the data from the carrier signal (removing it from the car). The receiver must know the correct chipping sequence to change the received data into its original format. This means the sender and receiver must be properly synchronized.

The sub-bits provide error-recovery instructions, just as parity does in RAID technologies. If a signal is corrupted using FHSS, it must be re-sent; but by using DSSS, even if the message is somewhat distorted, the signal can still be regenerated because it can be rebuilt from the chipping code bits. The use of this code allows for prevention of interference, allows for tracking of multiple transmissions, and provides a level of error correction.

FHSS vs. DSSS

FHSS uses only a portion of the total spectrum available at any one time, while the DSSS technology uses all of the available spectrum continuously. DSSS spreads the signals over a wider frequency band, whereas FHSS uses a narrowband carrier that changes frequently across a wide band.

Since DSSS sends data across all frequencies at once, it has higher data rates than FHSS. The first wireless WAN standard, 802.11, used FHSS, but as data requirements increased, DSSS was implemented. By using FHSS, the 802.11 standard can provide a data throughput of only 1 to 2 Mbps. By using DSSS instead, 802.11b provides a data throughput of up to 11 Mbps.

Orthogonal Frequency Division Multiplexing

Besides spread-spectrum techniques, another common approach to trying to move even more data over wireless frequency signals is called *orthogonal frequency division multiplexing (OFDM)*. OFDM is a digital multicarrier modulation scheme that compacts multiple modulated carriers tightly together, reducing the required spectrum. The modulated signals are orthogonal (perpendicular) and do not interfere with each other. OFDM uses a composite of narrow channel bands to enhance its performance in high-frequency bands. OFDM is officially a multiplexing technology and not a spread-spectrum technology, but is used in a similar manner.

A large number of closely spaced orthogonal subcarrier signals are used, and the data is divided into several parallel data streams or channels, one for each subcarrier. Channel equalization is simplified because OFDM uses many slowly modulated narrowband signals rather than one rapidly modulated wideband signal.

OFDM is used for several wideband digital communication types such as digital television, audio broadcasting, DSL broadband Internet access, wireless networks, and 4G/5G mobile communications.

Wireless Networking Fundamentals

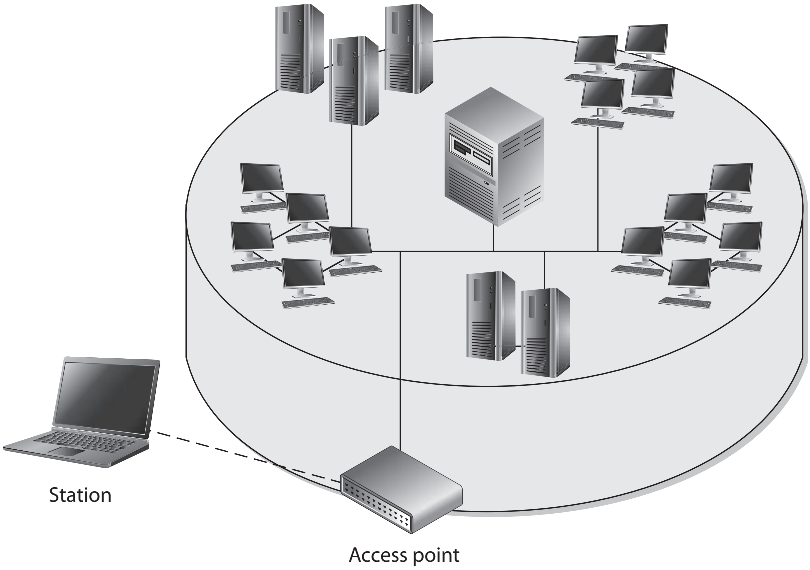
The techniques we've covered so far deal with how we create radio links between devices, but how do we build on those links to create networks? Fundamentally, there are three topologies used to build wireless networks: star, mesh, and point to point. The star topology is by far the most prevalent because it is used in both WLANs and cellular networks, both of which have endpoints connecting to a specialized network device that handles layer 2 forwarding and, in some cases, layer 3 routing. The mesh topology is common for low-power devices in close proximity to each other, such as those used in smart homes, as well as in devices that span a large area, such as environmental sensors in wildlife refuges. Finally, point-to-point wireless topologies are common when connecting buildings as part of a metropolitan area network (MAN).

Before we get into the myriad of wireless protocols that enable the various types of wireless networks, let's take a closer look at what makes a typical WLAN work.

WLAN Components

A WLAN uses a transceiver, called an *access point (AP)*, also known as a wireless access point (WAP), which connects to an Ethernet cable that is the link wireless devices use to access resources on the wired network, as shown in Figure 12-2. When the AP is connected to the LAN Ethernet by a wired cable, it is the component that connects the wired

Figure 12-2 Access points allow wireless devices to participate in wired LANs.



and the wireless worlds. The APs are in fixed locations throughout a network and work as communication beacons. Let's say a wireless user has a device with a wireless network interface card (NIC), which modulates the user's data onto radio frequency signals that are accepted and processed by the AP. The signals transmitted from the AP are received by the wireless NIC and converted into a digital format, which the device can understand.

When APs are used to connect wireless and wired networks, this is referred to as an *infrastructure WLAN*, which is used to extend an existing wired network. When there is just one AP and it is not connected to a wired network, it is considered to be in *standalone* mode and just acts as a wireless hub. An *ad hoc WLAN* has no APs; the wireless devices communicate with each other through their wireless NICs instead of going through a centralized device.

EXAM TIP Ad hoc WLANs are inherently less secure than infrastructure WLANs.

For a wireless device and AP to communicate, they must be configured to communicate over the same channel. A *channel* is a certain frequency within a given frequency band. The AP is configured to transmit over a specific channel, and the wireless device "tunes" itself to be able to communicate over this same frequency.

Any hosts that wish to participate within a particular WLAN must be configured with the proper *Service Set ID (SSID)*. Various hosts can be segmented into different WLANs by using different SSIDs. The reasons to segment a WLAN into portions are the same reasons wired systems are segmented on a network: the users require access to different resources, have different business functions, or have different levels of trust.

NOTE When wireless devices work in infrastructure mode, the AP and wireless clients form a group referred to as a Basic Service Set (BSS). This group is assigned a name, which is the SSID value.

When WLAN technologies first came out, authentication was simple and largely ineffective against many attackers. As wireless communication increased in use and many deficiencies were identified in these networks, a steady stream of improved approaches were developed and standardized. These covered both performance and security issues.

WLAN Standards

Standards are developed so that many different vendors can create various products that will work together seamlessly. Standards are usually developed on a consensus basis among the different vendors in a specific industry. The IEEE develops standards for a wide range of technologies—wireless being one of them.

The first WLAN standard, 802.11, was developed in 1997 and provided a 1- to 2-Mbps transfer rate. It worked in the 2.4-GHz frequency band, which is one of the free industrial, scientific, and medical (ISM) bands established by the International

Table 12-1 Generational Wi-Fi

Technology Supported	Wi-Fi Generation
802.11b	Wi-Fi 1
802.11a	Wi-Fi 2
802.11g	Wi-Fi 3
802.11n	Wi-Fi 4
802.11ac	Wi-Fi 5
802.11ax	Wi-Fi 6

Telecommunication Union (ITU). This means that organizations and users in most countries do not need a license to use this range. The 802.11 standard outlines how wireless clients and APs communicate; lays out the specifications of their interfaces; dictates how signal transmission should take place; and describes how authentication, association, and security should be implemented.

Now just because life is unfair, a long list of standards actually fall under the 802.11 main standard. You have probably seen the alphabet soup of 802.11a, 802.11b, 802.11g, 802.11n, 802.11ac, and 802.11ax (and a bunch of others). While the original 802.11 standard created the world of WLANs, the unrelenting pace of progress required changes and improvements over time. To try and make sense of things, the Wi-Fi Alliance created a scheme for numbering the generations of 802.11 protocols in 2018. This was done to help consumers differentiate products based on the most advanced 802.11-based technology supported by a given device. Table 12-1 lists the six generations of Wi-Fi, which we describe in the following sections.

NOTE Wi-Fi generations 1–3 are not formally defined by the Wi-Fi Alliance but are commonly understood to map to the technologies shown in Table 12-1.

802.11b

This standard was the first extension to the 802.11 WLAN standard. (Although 802.11a was conceived and approved first, it was not released first because of the technical complexity involved with this proposal.) 802.11b provides a transfer rate of up to 11 Mbps and works in the 2.4-GHz frequency range. It uses DSSS and is backward-compatible with 802.11 implementations.

802.11a

This standard uses a different method of modulating data onto the necessary radio carrier signals. Whereas 802.11b uses DSSS, 802.11a uses OFDM and works in the 5-GHz frequency band. Because of these differences, 802.11a is not backward-compatible with 802.11b or 802.11. Several vendors have developed products that can work with both 802.11a and 802.11b implementations; the devices must be properly configured or be able to sense the technology already being used and configure themselves appropriately.

As previously discussed, OFDM is a modulation scheme that splits a signal over several narrowband channels. The channels are then modulated and sent over specific frequencies. Because the data is divided across these different channels, any interference from the environment will degrade only a small portion of the signal. This allows for greater throughput. Like FHSS and DSSS, OFDM is a physical layer specification. It can be used to transmit high-definition digital audio and video broadcasting as well as WLAN traffic.

This technology offers advantages in two areas: speed and frequency. 802.11a provides up to 54 Mbps, and it does not work in the already very crowded 2.4-GHz spectrum. The 2.4-GHz frequency band is referred to as a "dirty" frequency because several devices already work there—microwaves, cordless phones, baby monitors, and so on. In many situations, this means that contention for access and use of this frequency can cause loss of data or inadequate service. But because 802.11a works at a higher frequency, it does not provide the same range as the 802.11b and 802.11g standards. The maximum speed for 802.11a is attained at short distances from the AP, up to 25 feet.

802.11g

The 802.11g standard provides for higher data transfer rates—up to 54 Mbps. This is basically a speed extension for 802.11b products. If a product meets the specifications of 802.11b, its data transfer rates are up to 11 Mbps, and if a product is based on 802.11g, that new product can be backward-compatible with older equipment but work at a much higher transfer rate.

802.11n (Wi-Fi 4)

This standard is designed to be much faster than 802.11g, with throughput at 100 Mbps, and it works at the same frequency range as 802.11a (5 GHz). The intent is to maintain some backward-compatibility with current Wi-Fi standards, while combining a mix of the current technologies. This standard uses a concept called multiple input, multiple output (MIMO) to increase the throughput. This requires the use of two receive and two transmit antennas to broadcast in parallel using a 20-MHz channel.

802.11ac (Wi-Fi 5)

The IEEE 802.11ac WLAN standard is an extension of 802.11n. It also operates on the 5-GHz band, but increases throughput to 1.3 Gbps. 802.11ac is backward compatible with 802.11a, 802.11b, 802.11g, and 802.11n, but if in compatibility mode it slows down to the speed of the slower standard. A major improvement is the use of multiuser MIMO (MU-MIMO) technology, which supports up to four data streams, allowing that many endpoints to simultaneously use a channel. Another benefit of this newer standard is its support for *beamforming*, which is the shaping of radio signals to improve their performance in specific directions. In simple terms, this means that 802.11ac is better able to maintain high data rates at longer ranges than its predecessors.

802.11ax (Wi-Fi 6)

Higher data rates are not always the best way to solve problems, and in the race for faster standards, we took a bunch of shortcuts that made them inefficient. The 802.11ax standard aims to address efficiency rather than faster speeds. A significant improvement it has is a new multiuser OFDM technology that replaces the single-user focused technology used in the 802.11a/g/n/ac standards. This means that multiple stations get to use available channels much more efficiently. In addition, the new standard doubles the number of streams supported by MU-MIMO, which means more stations can use it at the same time. These and many other improvements make 802.11ax much faster and better able to handle very crowded environments.

Other Wireless Network Standards

So far, we've focused pretty heavily on radio-based WLANs. There are other wireless network standards that you should know, at least at a superficial level. These include light-based WLANs and radio-based MANs and PANs. We describe the most important of these standards in the following sections.

Li-Fi

Li-Fi is a wireless networking technology that uses light rather than radio waves to transmit and receive data. It is essentially Wi-Fi using lights instead of radios. You can also think of it as fiber-optic communications without the fiber (i.e., over free space). It turns out that light, like radio, is an electromagnetic wave. The difference is that light is on a much higher frequency range and, thus, can carry significantly more information, at least in theory. Imagine if every light fixture in your home or workplace was able to modulate data onto the light it generates, while your computing device (laptop, smartphone, or whatever) could sense it and use its own light source (maybe the flash on your smartphone) to send data back to the light bulb. Because of the frequencies involved, our eyes are not able to perceive the tiny fluctuations in frequency. Besides, Li-Fi can work over infrared light too, which we can't see anyway.

One of the key benefits of Li-Fi (besides speed and ubiquity) is that it is very constrained to a particular space. Each light bulb has a cone of illumination within which it communicates with specific devices. You don't have to worry about an attacker with a sophisticated antenna picking up your signals a mile away. You can also be pretty confident of who you are communicating with because they have to be right there under the light source. These relatively small areas of service (by a given light source) are called *attocells*. The prefix atto- means quintillionth (which is a pretty small number), but, importantly, it's the next prefix down after femto-, as in *femtocells*, which are tiny cells used in some cellular networks.

At the time of this writing, Li-Fi technology is in its infancy but holds great promise. There are still many challenges to overcome, including co-channel interference (where multiple light sources overlap each other), roaming (seamlessly transferring a communications channel to an adjacent attocell or to an RF-based system if the user wanders out of the supported area), and endpoint interface devices (the sensors and light sources that would have to be built into each laptop, smartphone, etc.). Still, the benefits are many. Apart from the ones mentioned in the previous paragraph, Li-Fi promises to support much higher densities of endpoints, with much lower latencies, and in places where RF can be problematic (e.g., healthcare facilities, aircraft cabins, and power plants).

802.16

IEEE standard 802.16 is a MAN wireless standard that allows for wireless traffic to cover a much wider geographical area, where stations can be as far as 70 km apart. It uses some of the same bands as WLAN standards, specifically 2.4 GHz and 5 GHz, but uses up to 256 subcarriers with variable data rates to efficiently handle lots of traffic across large distances. This technology is also referred to as *broadband* wireless access.

A commercial technology that is based on 802.16 is WiMAX, which was widely touted as a replacement for second-generation (2G) digital cellular networks, particularly in rural areas. While this did not happen across the board (it largely lost out to Long Term Evolution or LTE), 802.16, and WiMAX in particular, remains in widespread use, especially outside the United States. A common implementation of 802.16 technology is shown in Figure 12-3.

NOTE IEEE 802.16 is a standard for vendors to follow to allow for interoperable broadband wireless connections. IEEE does not test for compliance to this standard. The WiMAX Forum runs a certification program that is intended to guarantee compliance with the standard and interoperability with equipment between vendors.

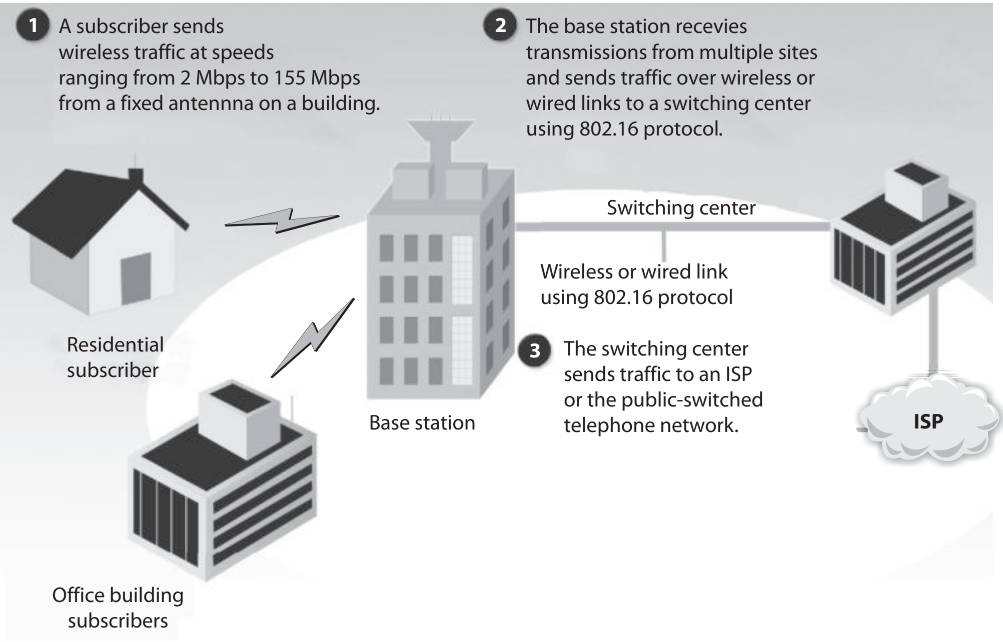

1 A subscriber sends wireless traffic at speeds ranging from 2 Mbps to 155 Mbps from a fixed antennna on a building.
2 The base station recevies transmissions from multiple sites and sends traffic over wireless or wired links to a switching center using 802.16 protocol.
Switching center
Wireless or wired link using 802.16 protocol
Residential subscriber
3 The switching center sends traffic to an ISP or the public-switched telephone network.
Base station
ISP
Office building subscribers

Figure 12-3 Broadband wireless in MAN

802.15.4

The IEEE 802.15.4 standard deals with a much smaller geographical network, which is referred to as a *wireless personal area network (WPAN)*. This technology allows for connectivity to take place among "disadvantaged" devices, which are the ubiquitous low-cost, low-data-rate, low-power, extended-life ones such as the embedded devices introduced in Chapter 7. For example, if you are using active radio frequency identification (RFID) or Industrial Internet of Things (IIoT) devices, odds are that you are using 802.15.4. This standard is optimized for situations in which machines communicate directly with other machines over relatively short distances (typically no more than 100 meters). For this reason, this standard is a key enabler of the Internet of Things (IoT), in which everything from your thermostat to your door lock is (relatively) smart and connected.

The 802.15.4 standard defines the physical (PHY) layer and Media Access Control (MAC) sublayer of the data link layer in the OSI model. At the physical layer, it uses DSSS. For MAC, it uses CSMA-CA. In terms of topology, this standard supports star, tree, and mesh networks. The catch is that, regardless of the topology, 802.15.4 requires a full-function device (FFD) that acts as a central node for the network (even if it is not logically or physically placed at its center). This central device is called the *coordinator* for one or more connected reduced-function devices (RFDs). This makes a lot of sense in a star or tree topology, where you have a regular computer as the hub or root node. It might be a bit less intuitive when you think of mesh networks such as you would find in a smart home network, but we'll get into that when we discuss ZigBee in the next section.

There are multiple extensions to the base 802.15.4 standard that optimize it for specific geographic regions or applications. You may come across the following:

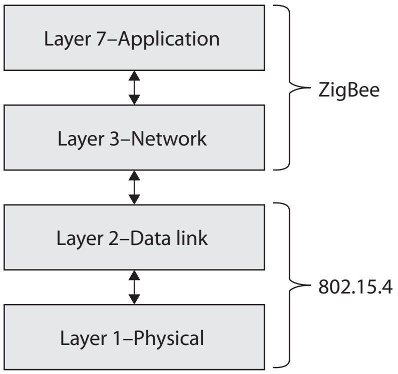
802.15.4c	For use in China
802.15.4d	For use in Japan
802.15.4e	For industrial applications
802.15.4f	For active (i.e., battery powered) radio frequency identification (RFID)
802.15.4g	For smart utility networks (SUNs)

Because this standard was intended to support embedded devices in close proximity to each other, the typical range is only about 10 meters (though it could reach 1 km in optimal conditions) and the data rates are quite low. While nodes frequently communicate at the maximum rate of 250 Kbps, there are also lower rates of 100, 20, and even 10 Kbps for smaller devices that have to last a long time on small batteries. Despite the low data rates, devices that implement this standard are able to support real-time applications (i.e., those that require extremely low latencies) through the use of Guaranteed Time Slot (GTS) reservations. Note that when a GTS is used, the channel access technique used has to be time division multiple access (TDMA) instead of the more typical CSMA/CA. TDMA is a technique that divides each communications channel into multiple time slots to increase the data rates by taking advantage of the fact that not every station will be transmitting all the time.

Security-wise, 802.15.4 implements access control lists (ACLs) by default, so nodes can decide whether to communicate with other nodes based on their claimed physical address. Keep in mind, however, that spoofing a physical address is trivial. The standard also offers (but does not require) two other security mechanisms that you should know. The first is support for symmetric key encryption using the Advanced Encryption Standard (AES) with 128-bit keys, used to protect message confidentiality and integrity. The second is a frame counter feature that protects against replay attacks by tracking the last message received from another node and ensuring a new message is more recent than it.

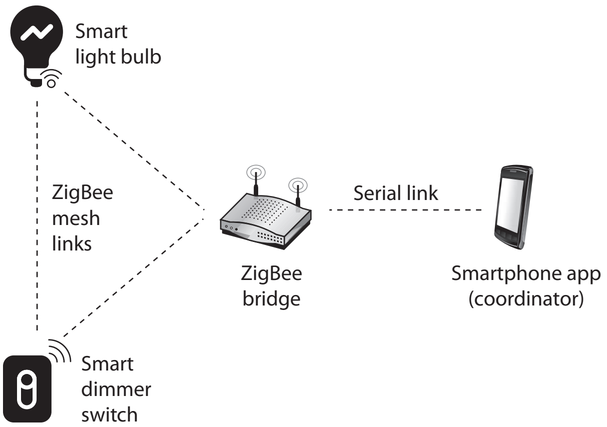
ZigBee

ZigBee is one of the most popular standards based on IEEE 802.15.4. It sits right on top of the layer 1 and layer 2 services provided by 802.15.4 and adds networking and application layer support.



ZigBee is intended to be simpler and cheaper than most WPAN protocols and is very popular in the embedded device market. You can find ZigBee in a variety of home automation, industrial control, medical, and sensor network applications. Figure 12-4

Figure 12-4 ZigBee in a smart home



shows a typical implementation of this standard for controlling lights in a smart home. All light bulbs and switches are able to talk to each other either directly, through the bridge, or by relaying traffic for destination nodes that are too far from the source node. Note that the connection between the bridge and the controller is over a serial link, but this could be implemented as a connection over Wi-Fi, Bluetooth, or any other means.

Because ZigBee is meant to be used in embedded devices that don't have (and can't afford) a bunch of operating system overhead, it assumes what is called an *open trust model*. This means that all applications within a device trust each other, which indirectly extends to all devices in a network as well. It also means that perimeter protection is absolutely critical. This protection should be implemented both physically and logically. At the physical level, ZigBee devices should be tamper-resistant to prevent attackers from simply reading encryption keys or otherwise gaining physical control of a node and using it as a beachhead for future attacks. At the logical level, it means controlling access to the network, which is done primarily through key management.

The ZigBee standard defines three different 128-bit symmetric keys:

- **Network key** Shared by all nodes to support broadcasts
- **Link key** Unique for each pair of connected devices and used for unicasts
- **Master key** Unique for each pair of connected devices and used for the Symmetric-Key Key Establishment (SKKE) protocol from which the other keys are derived

Since embedded devices oftentimes lack a user interface, the ZigBee standard allows multiple ways to distribute and subsequently manage these keys. The most secure way is based on a centralized security model in which the coordinator node acts as a Trust Center. This node is responsible for authenticating new devices that attempt to join the network and then securely sending them the keys they need. To facilitate this, manufacturers of ZigBee devices can install unique certificates at the factory, which are then used by the Trust Center to authenticate them and distribute keys using the Certificate-Based Key Establishment (CBKE) protocol. This is not all that common of an approach apart from high-security commercial systems. More commonly, manufacturers install a unique key in each device, which is then used by the SKKE protocol to derive keys, much like we would do using the Diffie-Hellman algorithm we covered in Chapter 8. This second approach is less secure, doesn't require a Trust Center, and is typical in consumer systems.

EXAM TIP ZigBee is most secure when a coordinator node acts as a Trust Center.

Bluetooth

The *Bluetooth wireless* technology has a 1- to 3-Mbps transfer rate and works in a range of approximately 1, 10, or 100 meters. It was originally invented as an alternative to connecting devices using cables. Unsurprisingly, its most common application today is in cordless headsets for smartphones. However, the technology has plenty of other uses.

If you have a cell phone and a tablet that are both Bluetooth-enabled and both have calendar functionality, you could have them update each other without any need to connect them physically. If you added some information to your cell phone contacts list and task list, for example, you could just place the phone close to your tablet. The tablet would sense that the other device is nearby, and it would then attempt to set up a network connection with it. Once the connection was made, synchronization between the two devices would take place, and the tablet would add the new contacts list and task list data. Bluetooth works in a portion of the frequency band used by 802.11 devices (2.4 GHz).

In early versions of Bluetooth, real security risks existed due to protocol vulnerabilities, but they have been largely mitigated. Still, as with any other technology, it is possible for attackers to compromise the confidentiality, integrity, or availability of Bluetooth devices. One attack type to which these devices are vulnerable is called *Bluesnarfing*, which is the unauthorized access from a wireless device through a Bluetooth connection. This allows attackers to read, modify, or delete calendar events, contacts, e-mails, text messages, and so on. While recent versions of the Bluetooth standard make this much harder, it is still possible to trick unwary users into accepting an attacker's connection attempts.

Another attack type that Bluetooth is vulnerable to is referred to as *Bluejacking*. In this attack, someone sends an unsolicited message to a device that is Bluetooth-enabled. Bluejackers look for a receiving device (phone, tablet, laptop) and then send a message to it. The countermeasure is to put the Bluetooth-enabled device into nondiscoverable mode so others cannot identify this device in the first place. If you receive some type of message this way, just look around you. Bluetooth only works within a 10-meter distance, so it is coming from someone close by.

Other Important Standards

The wireless standards we've covered so far cover the ways in which devices connect to each other and send data over the radio links they create. Over the years, we've discovered that there are a bunch of other features we want in our wireless networks, regardless of the communications standards being used by the radios themselves. These include Quality of Service (QoS), roaming, and spectrum management issues. Let's take a look at another set of standards you should know.

802.11e

This standard provides QoS and support of multimedia traffic in wireless transmissions. Voice, streaming video, and other types of time-sensitive applications have a lower tolerance for delays in data transmission. The problem is that the original 802.11 protocol treated all traffic equally. In other words, an e-mail message that could safely take minutes to get through had exactly the same priority as a video packet whose tolerable latency is measured in fractions of a second. To address this, the 802.11e standard defines four access categories (ACs) in increasing priority: background, best effort, video, and voice. This QoS provides the capability to prioritize traffic and affords guaranteed delivery. This standard and its capabilities have opened the door to allow many different types of data to be transmitted over wireless connections.

802.11f

When a user moves around in a WLAN, her wireless device often needs to communicate with different APs. An AP can cover only a certain distance, and as the user moves out of the range of the first AP, another AP needs to pick up and maintain her signal to ensure she does not lose network connectivity. This is referred to as *roaming*, and for this to happen seamlessly, the APs need to communicate with each other. If the second AP must take over this user's communication, it needs to be assured that this user has been properly authenticated and must know the necessary settings for this user's connection. This means the first AP needs to be able to convey this information to the second AP. The conveying of this information between the different APs during roaming is what 802.11f deals with. The process of transferring between one AP and another is sometimes called *handoff*. It outlines how this information can be properly shared.

802.11h

Because the ISM bands are unlicensed, devices that operate in them are expected to deal well with interference from other devices. This was all good and well before the explosion of WLANs and Bluetooth devices, but quickly became an issue as crowding increased. To make things worse, the 5-GHz band is used not only for Wi-Fi but also for certain radar and satellite communications systems. In this increasingly busy portion of the spectrum, something had to be done to deal with interference.

The 802.11h standard was originally developed to address these issues in Europe, where interference in the 5-GHz band was particularly problematic. However, the techniques it implements are applicable in many countries around the world. Two specific technologies included in the standard are *Dynamic Frequency Selection (DFS)* and *Transmit Power Control (TPC)*. DFS is typically implemented in the WLAN AP and causes it to automatically select channels that have less interference, particularly from radars. TPC causes any device to automatically reduce its power output when it detects interference from other networks.

802.11j

Japan regulates its radio spectrum differently than many other countries, particularly in the 4.9- and 5-GHz bands. Specifically, Japan uses different frequencies, radio channel widths, and wireless operating settings. In order to allow international devices to be interoperable in Japan, the IEEE developed the 802.11j standard. The need for this standard underscores the fact that each country has the sovereign right to regulate its radio spectrum as it sees fit.

Evolution of WLAN Security

To say that security was an afterthought in the first WLANs would be a remarkable understatement. As with many new technologies, wireless networks were often rushed to market with a focus on functionality, even if that sometimes came at the expense of security. Over time, vendors and standards bodies caught on and tried to correct these omissions. While we have made significant headway in securing our wireless networks, as security professionals we must acknowledge that whenever we transmit anything over the electromagnetic spectrum, we are essentially putting our data in the hands (or at least within the grasp) of our adversaries.

802.11

When WLANs were being introduced, there was industry-wide consensus that some measures would have to be taken to assure users that their data (now in the air) would be protected from eavesdropping to the same degree that data on a wired LAN was already protected. This was the genesis of *Wired Equivalent Privacy (WEP)*. This first WLAN standard, codified as part of the original IEEE 802.11, had a tremendous number of security flaws. These were found within the core standard itself, as well as in different implementations of this standard. Before we delve into these deficiencies, it will be useful to spend a bit of time with some of the basics of 802.11.

EXAM TIP If you ever come across WEP in the context of wireless security, you know it's the wrong answer (unless the question is asking for the least secure standard).

The wireless devices using this protocol can authenticate to the AP in two main ways: *open system authentication (OSA)* and *shared key authentication (SKA)*. OSA does not require the wireless device to prove to the AP it has a specific cryptographic key to allow for authentication purposes. In many cases, the wireless device needs to provide only the correct SSID value. In OSA implementations, all transactions are in cleartext because no encryption is involved. So an intruder can sniff the traffic, capture the necessary steps of authentication, and walk through the same steps to be authenticated and associated to an AP.

When an AP is configured to use SKA, the AP sends a random value to the wireless device. The device encrypts this value with a preshared key (PSK) and returns it. The AP decrypts and extracts the response, and if it is the same as the original value, the device is authenticated. In this approach, the wireless device is authenticated to the network by proving it has the necessary encryption key. The PSK, commonly known as the Wi-Fi password, is a 64- or 128-bit key.

The three core deficiencies with WEP are the use of static encryption keys, the ineffective use of initialization vectors, and the lack of packet integrity assurance. The WEP protocol uses the RC4 algorithm, which is a stream-symmetric cipher. *Symmetric* means the sender and receiver must use the exact same key for encryption and decryption purposes. The 802.11 standard does not stipulate how to update these keys through an automated process, so in most environments, the RC4 symmetric keys are never changed out. And usually all of the wireless devices and the AP share the exact same key. This is like having everyone in your company use the exact same password. Not a good idea. So that is the first issue—static WEP encryption keys on all devices.

The next flaw is how initialization vectors (IVs) are used. An IV is a numeric seeding value that is used with the symmetric key and RC4 algorithm to provide more randomness to the encryption process. Randomness is extremely important in encryption because any patterns can give the bad guys insight into how the process works, which may allow them to uncover the encryption key that was used. The key and 24-bit IV value are inserted into the RC4 algorithm to generate a key stream. The values (1's and 0's) of the key stream are XORed with the binary values of the individual packets. The result is ciphertext, or encrypted packets.

In most WEP implementations, the same IV values are used over and over again in this process, and since the same symmetric key (or shared secret) is generally used, there is no way to provide effective randomness in the key stream that is generated by the algorithm. The appearance of patterns allows attackers to reverse-engineer the process to uncover the original encryption key, which can then be used to decrypt future encrypted traffic.

So now we are onto the third mentioned weakness, which is the integrity assurance issue. WLAN products that use only the 802.11 standard introduce a vulnerability that is not always clearly understood. An attacker can actually change data within the wireless packets by flipping specific bits and altering the Integrity Check Value (ICV) so the receiving end is oblivious to these changes. The ICV works like a cyclic redundancy check (CRC) function; the sender calculates an ICV and inserts it into a frame's header. The receiver calculates his own ICV and compares it with the ICV sent with the frame. If the ICVs are the same, the receiver can be assured that the frame was not modified during transmission. If the ICVs are different, it indicates a modification did indeed take place and thus the receiver discards the frame. In WEP, there are certain circumstances in which the receiver cannot detect whether an alteration to the frame has taken place; thus, there is no true integrity assurance.

So the problems identified with the 802.11 standard include poor authentication, static WEP keys that can be easily obtained by attackers, IV values that are repetitive and do not provide the necessary degree of randomness, and a lack of data integrity. The next section describes the measures taken to remedy these problems.

NOTE 802.11 and WEP were deprecated years ago, are inherently insecure, and should not be used.

802.11i

IEEE came out with a standard in 2004 that deals with the security issues of the original 802.11 standard, which is called IEEE 802.11i or *Wi-Fi Protected Access 2 (WPA2)*. Why the number 2? Because while the formal standard was being ratified by the IEEE, the Wi-Fi Alliance pushed out WPA (the first one) based on the draft of the standard. For this reason, WPA is sometimes referred to as the *draft* IEEE 802.11i. This rush to push out WPA required the reuse of elements of WEP, which ultimately made WPA vulnerable to some of the same attacks that doomed its predecessor. Let's start off by looking at WPA in depth, since this protocol is still in use despite its weaknesses.

WPA employs different approaches that provide much more security and protection than the methods used in the original 802.11 standard. For starters, the PSK size was increased to 256 bits and is salted with the SSID of the WLAN to make it harder to crack. This is good, but the greatest enhancement of security is accomplished through specific protocols, technologies, and algorithms. The first protocol is *Temporal Key Integrity Protocol (TKIP)*, which is backward-compatible with the WLAN devices based upon the original 802.11 standard. TKIP actually works with WEP by feeding it keying material, which is data to be used for generating new dynamic keys. TKIP generates a new key for every frame that is transmitted. These changes constitute the variety of this standard known as WPA Personal, which is geared at consumers.

NOTE TKIP was developed by the IEEE 802.11i task group and the Wi-Fi Alliance. The goal of this protocol was to increase the strength of WEP or replace it fully without the need for hardware replacement. TKIP provides a key mixing function, which allows the RC4 algorithm to provide a higher degree of protection. It also provides a sequence counter to protect against replay attacks and implements a message integrity check mechanism.

There is also a more robust version called WPA Enterprise. The main difference is that it also integrates 802.1X port authentication and Extensible Authentication Protocol (EAP) authentication methods. The use of the 802.1X technology (which we'll discuss in its own section shortly) provides access control by restricting network access until full authentication and authorization have been completed, and provides a robust authentication framework that allows for different EAP modules to be plugged in. These two technologies (802.1X and EAP) work together to enforce mutual authentication between the wireless device and authentication server. So what about the static keys, IV value, and integrity issues?

TKIP addresses the deficiencies of WEP pertaining to static WEP keys and inadequate use of IV values. Two hacking tools, AirSnort and WEPCrack, can be used to easily crack WEP's encryption by taking advantage of these weaknesses and the ineffective use of the key scheduling algorithm within the WEP protocol. If a company is using products that implement only WEP encryption and is not using a third-party encryption solution (such as a VPN), these programs can break its encrypted traffic within minutes. There is no "maybe" pertaining to breaking WEP's encryption. Using these tools means it will be broken whether a 40-bit or 128-bit key is being used—it doesn't matter. This is one of the most serious and dangerous vulnerabilities pertaining to the original 802.11 standard.

The use of TKIP provides the ability to rotate encryption keys to help fight against these types of attacks. The protocol increases the length of the IV value and ensures that every frame has a different IV value. This IV value is combined with the transmitter's MAC address and the original WEP key, so even if the WEP key is static, the resulting encryption key will be different for every frame. (WEP key + IV value + MAC address = new encryption key.) So what does that do for us? This brings more randomness to the encryption process, and it is randomness that is necessary to properly thwart cryptanalysis and attacks on cryptosystems. The changing IV values and resulting keys make the resulting key stream less predictable, which makes it much harder for the attacker to reverse-engineer the process and uncover the original key.

TKIP also deals with the integrity issues by using a message integrity check (MIC) instead of an ICV function. If you are familiar with a message authentication code

(MAC) function, this is the same thing. A symmetric key is used with a hashing function, which is similar to a CRC function but stronger. The use of a MIC instead of an ICV function ensures the receiver will be properly alerted if changes to the frame take place during transmission. The sender and receiver calculate their own separate MIC values. If the receiver generates a MIC value different from the one sent with the frame, the frame is seen as compromised and it is discarded.

The types of attacks that have been carried out on WEP devices and networks that just depend upon WEP are numerous and unnerving. Wireless traffic can be easily sniffed, data can be modified during transmission without the receiver being notified, rogue APs can be erected (which users can authenticate to and communicate with, not knowing it is a malicious entity), and encrypted wireless traffic can be decrypted quickly and easily. Unfortunately, these vulnerabilities usually provide doorways to the actual wired network where the more destructive attacks can begin.

The full 802.11i (WPA2) has a major advantage over WPA by providing encryption protection with the use of the AES algorithm in counter mode with CBC-MAC (CCM), which is referred to as the Counter Mode Cipher Block Chaining Message Authentication Code Protocol (CCM Protocol or CCMP). AES is a more appropriate algorithm for wireless than RC4 and provides a higher level of protection. WPA2 defaults to CCMP, but can switch down to TKIP and RC4 to provide backward compatibility with WPA devices and networks.

802.11w

WPA2 was a huge step forward for WLAN security because it provided effective encryption for most wireless traffic. However, there are certain frames that cannot be encrypted because every station (even those that have not yet joined the network) must be able to receive. These are called *management frames*, and they take care of things like beaconing, association, and authentication. While we can't encrypt them, we can take measures to ensure their integrity. The IEEE 802.11w standard provides Management Frame Protection (MFP) that prevents certain types of attacks, such as replay and denial-of-service (DoS) attacks.

A particularly problematic type of DoS attack on WLANs is called a deauthentication (or deauth) attack and it exploits a feature of Wi-Fi that allows WAPs to disconnect rogue devices by sending a deauthentication management frame. You can see how, in an environment without MFP, it would be trivial for an attacker to spoof such messages, claiming to be the real WAP. 802.11w solves this problem for WLANs that are not yet on WPA3.

WPA3

Like any other security mechanism, WPA2 began to crack under intensifying attacks. By 2018, the Wi-Fi Alliance decided that a new approach was needed. The result is WPA3, which is not directly equivalent to any IEEE standard, though it does require 802.11w to protect management frames. Like its predecessor WPA2, WPA3 comes in two flavors: Personal and Enterprise.

WPA3 Personal is aimed at the consumer market and tries to make security transparent to the average user. One of the most important innovations of this standard is that it allows users to choose passwords that, though they might be easily guessable, still provide adequate security. This is done through Simultaneous Authentication of Equals (SAE), which is defined in IEEE 802.11s, instead of relying on WPA2's preshared keys. SAE uses the Diffie-Hellman key exchange method but adds an authentication element based on the (potentially weak) password. The result is a secret session key that is remarkably resistant to password-cracking attempts.

WPA3 Enterprise is similar to its predecessor (WPA2 Enterprise) but makes use of stronger cryptography. It does this by restricting the allowed algorithms to a handful of strong ones that use 192-bit keys. It also requires certificates on both the AP and the wireless device for mutual authentication. The challenge with deploying WPA3 is that many older wireless interfaces, particularly those on most embedded devices, cannot support it, which means you may have to upgrade many (or all) of your endpoints.

802.1X

The 802.11i standard can be understood as three main components in two specific layers. The lower layer contains the improved encryption algorithms and techniques (TKIP and CCMP), while the layer that resides on top of it contains 802.1X. They work together to provide more layers of protection than the original 802.11 standard.

The 802.1X standard is a port-based network access control protocol that ensures a user cannot make a full network connection until he is properly authenticated. This means a user cannot access network resources and no traffic is allowed to pass, other than authentication traffic, from the wireless device to the network until the user is properly authenticated. An analogy is having a chain on your front door that enables you to open the door slightly to identify a person who knocks before you allow him to enter your house.

NOTE 802.1X is not a wireless protocol. It is an access control protocol that can be implemented on both wired and wireless networks.

By incorporating 802.1X, the new standard allows for the user to be authenticated, whereas using only WPA provides *system* authentication. User authentication provides a higher degree of confidence and protection than system authentication. The 802.1X technology actually provides an authentication framework and a method of dynamically distributing encryption keys. The three main entities in this framework are the supplicant (wireless device), the authenticator (AP), and the authentication server (usually a RADIUS server).

The AP controls all communication and allows the wireless device to communicate with the authentication server and wired network only when all authentication steps are completed successfully. This means the wireless device cannot send or receive HTTP, DHCP, SMTP, or any other type of traffic until the user is properly authorized. WEP does not provide this type of strict access control.

Another disadvantage of the original 802.11 standard is that mutual authentication is not possible. When using WEP alone, the wireless device can authenticate to the AP, but the authentication server is not required to authenticate to the wireless device. This means a rogue AP can be set up to capture users' credentials and traffic without the users being aware of this type of attack. 802.11i deals with this issue by using EAP. EAP allows for mutual authentication to take place between the authentication server and wireless device and provides flexibility in that users can be authenticated by using passwords, tokens, one-time passwords, certificates, smart cards, or Kerberos. This allows wireless users to be authenticated using the current infrastructure's existing authentication technology. The wireless device and authentication server that are 802.11i-compliant have different authentication modules that plug into 802.1X to allow for these different options. So, 802.1X provides the framework that allows for the different EAP modules to be added by a network administrator. The two entities (supplicant and authenticator) agree upon one of these authentication methods (EAP modules) during their initial handshaking process.

The 802.11i standard does not deal with the full protocol stack, but addresses only what is taking place at the data link layer of the OSI model. Authentication protocols reside at a higher layer than this, so 802.11i does not specify particular authentication protocols. The use of EAP, however, allows different protocols to be used by different vendors. For example, Cisco uses a purely password-based authentication framework called Lightweight Extensible Authentication Protocol (LEAP). Other vendors, including Microsoft, use EAP and Transport Layer Security (EAP-TLS), which carries out authentication through digital certificates. And yet another choice is Protected EAP (PEAP), where only the server uses a digital certificate.

EAP-Tunneled Transport Layer Security (EAP-TTLS) is an EAP protocol that extends TLS. EAP-TTLS is designed to provide authentication that is as strong as EAP-TLS, but it does not require that each user be issued a certificate. Instead, only the authentication servers are issued certificates. User authentication is performed by password, but the password credentials are transported in a securely encrypted tunnel established based upon the server certificates.

If EAP-TLS is being used, the authentication server and wireless device exchange digital certificates for authentication purposes. If PEAP is being used instead, the user of the wireless device sends the server a password and the server authenticates to the wireless device with its digital certificate. In both cases, some type of public key infrastructure (PKI) needs to be in place. If a company does not have a PKI currently implemented, it can be an overwhelming and costly task to deploy a PKI just to secure wireless transmissions.

When EAP-TLS is being used, the steps the server takes to authenticate to the wireless device are basically the same as when a TLS connection is being set up between a web server and web browser. Once the wireless device receives and validates the server's digital certificate, it creates a master key, encrypts it with the server's public key, and sends it over to the authentication server. Now the wireless device and authentication server have a master key, which they use to generate individual symmetric session keys. Both entities use these session keys for encryption and decryption purposes, and it is the use of these keys that sets up a secure channel between the two devices.

Organizations may choose to use PEAP instead of EAP-TLS because they don't want the hassle of installing and maintaining digital certificates on every wireless device.

Before you purchase a WLAN product, you should understand the requirements and complications of each method to ensure you know what you are getting yourself into and if it is the right fit for your environment.

A large concern with any WLANs using just WEP is that if individual wireless devices are stolen, they can easily be authenticated to the wired network. 802.11i has added steps to require the user to authenticate to the network instead of just requiring the wireless device to authenticate. By using EAP, the user must send some type of credential set that is tied to his identity. When using only WEP, the wireless device authenticates itself by proving it has a symmetric key that was manually programmed into it. Since the user does not need to authenticate using WEP, a stolen wireless device can allow an attacker easy access to your precious network resources.

The Answer to All Our Prayers?

So, does the use of EAP, 802.1X, AES, and TKIP result in secure and highly trusted WLAN implementations? Maybe, but we need to understand what we are dealing with here. TKIP was created as a quick fix to WEP's overwhelming problems. It does not provide an overhaul for the wireless standard itself because WEP and TKIP are still based on the RC4 algorithm, which is not the best fit for this type of technology. The use of AES is closer to an actual overhaul, but it is not backward-compatible with the original 802.11 implementations. In addition, we should understand that using all of these new components and mixing them with the current 802.11 components will add more complexity and steps to the process. Security and complexity do not usually get along. The highest security is usually accomplished with simplistic and elegant solutions to ensure all of the entry points are clearly understood and protected. These newer technologies add more flexibility to how vendors can choose to authenticate users and authentication servers, but can also bring us interoperability issues because the vendors will not all choose the same methods. This means that if an organization buys an AP from company A, then the wireless cards the organization buys from companies B and C may not work seamlessly.

So, does that mean all of this work has been done for naught? No. 802.11i provides much more protection and security than WEP ever did. The working group has had very knowledgeable people involved and some very large and powerful companies aiding in the development of these new solutions. But the customers who purchase these new products need to understand what will be required of them *after* their purchase. For example, with the use of EAP-TLS, each wireless device needs its own digital certificate. Are your current wireless devices programmed to handle certificates? How will the certificates be properly deployed to all the wireless devices? How will the certificates be maintained? Will the devices and authentication server verify that certificates have not been revoked by periodically checking a certificate revocation list (CRL)? What if a rogue authentication server or AP was erected with a valid digital certificate? The wireless device would just verify this certificate and trust that this server is the entity it is supposed to be communicating with.

Today, WLAN products are being developed following the stipulations of this 802.11i wireless standard. Many products will straddle the fence by providing TKIP for backward-compatibility with current WLAN implementations and AES for organizations that are just now thinking about extending their current wired environments with a wireless component. Before buying wireless products, customers should review the Wi-Fi Alliance's certification findings, which assess systems against the 802.11i proposed standard.

Best Practices for Securing WLANs

There is no silver bullet to protect any of our devices or networks. That being said, there are a number of things we can do that will increase the cost of the attack for the adversary. Some of the best practices pertaining to WLAN implementations are as follows:

- Change the default SSID. Each AP comes with a preconfigured default SSID value that may reveal the manufacturer and even model number, which may advertise systems with known vulnerabilities.
- Implement WPA3 Enterprise to provide centralized user authentication (e.g., RADIUS, Kerberos). Before users can access the network, require them to authenticate.
- Use separate VLANs for each class of users, just as you would on a wired LAN.
- If you must support unauthenticated users (e.g., visitors), ensure they are connected to an untrusted VLAN that remains outside your network's perimeter.
- Deploy a wireless intrusion detection system (WIDS).
- Physically put the AP at the center of the building to limit how far outside the facility the signal will reach (and be reachable). The AP has a specific zone of coverage it can provide.
- Logically put the AP in a DMZ with a firewall between the DMZ and internal network. Allow the firewall to investigate the traffic before it gets to the wired network.
- Implement VPN for wireless devices to use. This adds another layer of protection for data being transmitted.
- Configure the AP to allow only known MAC addresses into the network. Allow only known devices to authenticate. But remember that these MAC addresses are sent in cleartext, so an attacker could capture them and masquerade himself as an authenticated device.
- Carry out penetration tests on the WLAN. Use the tools described in this section to identify APs and attempt to break the current encryption scheme being used.

Mobile Wireless Communication

Mobile wireless has now exploded into a trillion-dollar industry, with over 14 billion devices worldwide, fueled by a succession of new technologies and by industry and international standard agreements. So what is a mobile phone anyway? It is a device that can send voice and data over wireless radio links. It connects to a cellular network, which is connected to the public switched telephone network (PSTN). So instead of needing a physical cord and connection that connects your phone and the PSTN, you have a device that allows you to indirectly connect to the PSTN as you move around a wide geographic area.

A cellular network distributes radio signals over delineated areas, called *cells*. Each cell has at least one fixed-location transceiver (base station) and is joined to other cells to provide connections over large geographic areas. So as you are talking on your mobile phone and you move out of one cell, the base station in the original cell sends your connection information to the base station in the next cell so that your call is not dropped and you can continue your conversation.

We do not have an infinite number of frequencies to work with when it comes to mobile communication. Millions of people around the world are using their cell phones as you read this. How can all of these calls take place if we only have one set of frequencies to use for such activity? Individual cells can use the same frequency range, as long as they are not right next to each other. So the same frequency range can be used in every other cell, which drastically decreases the amount of ranges required to support simultaneous connections. A rudimentary depiction of a cellular network, in which nonadjacent cells reuse the frequency sets F0, F1, F2, F3, and F4, is shown in Figure 12-5.

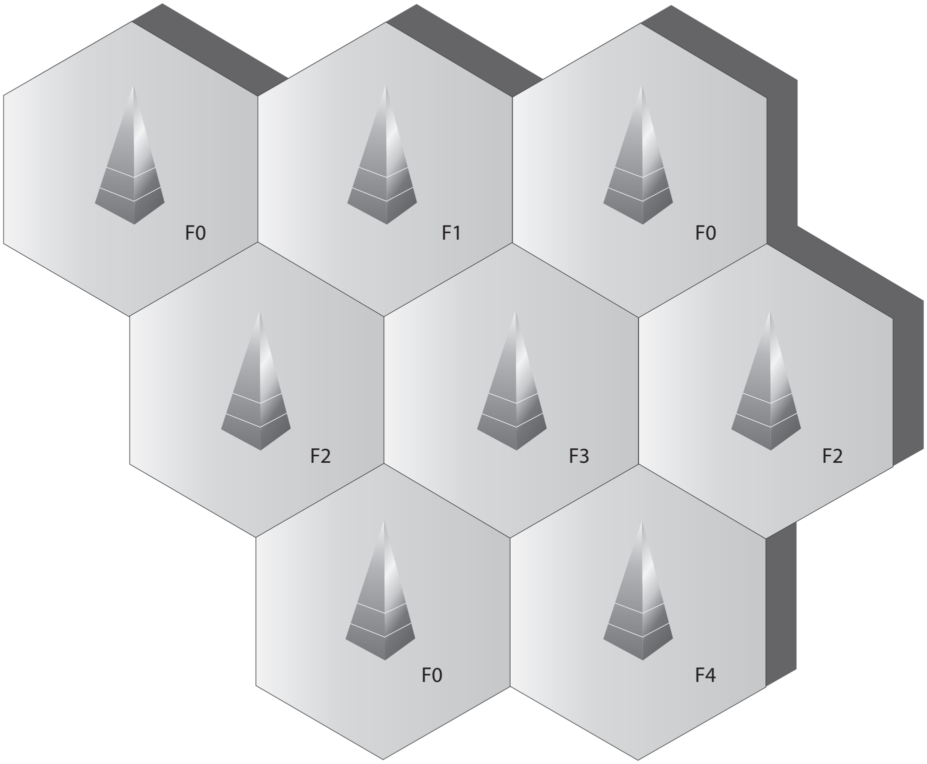


Figure 12-5 Nonadjacent cells can use the same frequency ranges.

Multiple Access Technologies

The industry has had to come up with other ways to allow millions of users to be able to use this finite resource (frequency range) in a flexible manner. Over time, mobile wireless has been made up of progressively more complex and more powerful "multiple access" technologies, listed here:

- Frequency division multiple access (FDMA)
- Time division multiple access (TDMA)
- Code division multiple access (CDMA)
- Orthogonal frequency division multiple access (OFDMA)

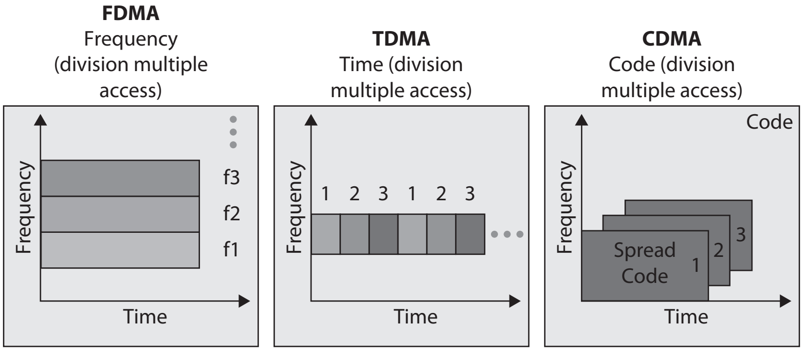
We'll quickly go over the characteristics of each of these technologies because they are the foundational constructs of the various cellular network generations.

Frequency division multiple access (FDMA) was the earliest multiple access technology put into practice. The available frequency range is divided into sub-bands (channels), and one channel is assigned to each subscriber (cell phone). The subscriber has exclusive use of that channel while the call is made, or until the call is terminated or handed off; no other calls or conversations can be made on that channel during that call. Using FDMA in this way, multiple users can share the frequency range without the risk of interference between the simultaneous calls. FDMA was used in the first generation (1G) of cellular networks. Various 1G mobile implementations, such as Advanced Mobile Phone System (AMPS), Total Access Communication System (TACS), and Nordic Mobile Telephone (NMT), used FDMA.

Time division multiple access (TDMA) increases the speed and efficiency of the cellular network by taking the radio frequency spectrum channels and dividing them into time slots. At various time periods, multiple users can share the same channel; the systems within the cell swap from one user to another user, in effect, reusing the available frequencies. TDMA increased speeds and service quality. A common example of TDMA in action is a conversation. One person talks for a time and then quits, and then a different person talks. In TDMA systems, time is divided into frames. Each frame is divided into slots. TDMA requires that each slot's start and end time are known to both the source and the destination. Mobile communication systems such as Global System for Mobile Communication (GSM), Digital AMPS (D-AMPS), and Personal Digital Cellular (PDC) use TDMA.

Code division multiple access (CDMA) was developed after FDMA, and as the term "code" implies, CDMA assigns a unique code to each voice call or data transmission to uniquely identify it from all other transmissions sent over the cellular network. In a CDMA "spread spectrum" network, calls are spread throughout the entire radio frequency band. CDMA permits every user of the network to simultaneously use every channel in the network. At the same time, a particular cell can simultaneously interact with multiple other cells. These features make CDMA a very powerful technology.

It is the main technology for the mobile cellular networks that presently dominate the wireless space.


FDMA
Frequency (division multiple access)
TDMA
Time (division multiple access)
CDMA
Code (division multiple access)
Frequency
Time
f3
f2
f1
1 2 3 1 2 3
Code
Spread Code
1
2
3

Orthogonal frequency division multiple access (OFDMA) is derived from a combination of FDMA and TDMA. In earlier implementations of FDMA, the different frequencies for each channel were widely spaced to allow analog hardware to separate the different channels. In OFDMA, each of the channels is subdivided into a set of closely spaced orthogonal frequencies with narrow subchannels. Each of the different subchannels can be transmitted and received simultaneously in a multiple input, multiple output (MIMO) manner. The use of orthogonal frequencies and MIMO allows signal processing techniques to reduce the impacts of any interference between different subchannels and to correct for channel impairments, such as noise and selective frequency fading. 4G and 5G require that OFDMA be used.

Generations of Mobile Wireless

Multiple access technology development was driven by the dramatic growth in mobile subscribers worldwide. Mobile wireless technologies have gone through a whirlwind of confusing generations. The first generation (1G) dealt with analog transmissions of voice-only data over circuit-switched networks. This generation provided a throughput of around 19.2 Kbps. The second generation (2G) allows for digitally encoded voice and data to be transmitted between wireless devices, such as cell phones, and content providers. TDMA, CDMA, GSM, and PCS all fall under the umbrella of 2G mobile telephony. This technology can transmit data over circuit-switched networks and supports data encryption, fax transmissions, and short message services (SMSs).

The third-generation (3G) networks became available around the turn of the century. Incorporating FDMA, TDMA, and CDMA, 3G has the flexibility to support a great variety of applications and services. Further, 3G replaced circuit switching with packet switching. Modular in design to allow ready expandability, backward compatibility with 2G networks, and stressing interoperability among mobile systems, 3G services greatly expanded the applications available to users, such as global roaming (without changing one's cell phone or cell phone number), as well as Internet services and multimedia.

In addition, reflecting the ever-growing demand from users for greater speed, latency in 3G networks was much reduced as transmission speeds were enhanced. More enhancements

Mobile Technology Generations

Like many technologies, the mobile communication technology has gone through several different generations.

First generation (1G):

- Analog services
- Voice service only

Second generation (2G):

- Primarily voice, some low-speed data (circuit switched)
- Phones were smaller in size
- Added functionality of e-mail, paging, and caller ID

Generation 2½ (2.5G):

- Higher data rates than 2G
- "Always on" technology for e-mail and pages

Third generation (3G):

- Integration of voice and data
- Packet-switched technology, instead of circuit-switched

Generation 3.5 G (3GPP)

- Higher data rates
- Use of OFDMA technology

Fourth generation (4G)

- Based on an all-IP packet-switched network
- Data exchange at 100 Mbps to 1 Gbps

Fifth generation (5G)

- Higher frequency ranges, which cut down range and make interference a bigger deal
- Data rates of 20 Gbps possible
- Supports dense deployment of high-speed, low-latency services

to 3G networks, often referred to as 3.5G or as mobile broadband, took place under the rubric of the Third Generation Partnership Project (3GPP). 3GPP resulted in a number of new or enhanced technologies. These include Enhanced Data Rates for GSM Evolution (EDGE), High-Speed Downlink Packet Access (HSDPA), CDMA2000, and Worldwide Interoperability for Microwave Access (WiMAX).

At the time of writing, 4th generation (4G) mobile networks are dominant (though, as we're about to see, that's going to change soon). Initially, there were two competing technologies that fell under the umbrella of 4G: Mobile WiMAX and Long-Term Evolution (LTE). Eventually, however, LTE won out and WiMAX is no longer used in mobile wireless networks. (Though, as we've already discussed, WiMAX is still used as an alternative to traditional ISP services in WANs.) A 4G system does not support traditional circuit-switched telephony service as 3G does, but works over a purely packet-based network. 4G devices are IP-based and are based upon OFDMA instead of the previously used multiple carrier access technologies. In theory, 4G devices should be able to reach 2-Gbps data rates, though that is seldom the case in practice.

Fifth generation (5G) is the technology that is all the rage right now. Its biggest advantage, at least from users' perspectives, over 4G is speed. 5G is capable of reaching a whopping 20 Gbps, which puts it in the neighborhood of the latest Wi-Fi 6 standard. What are the drawbacks of 5G? In order to achieve those jaw-dropping speeds, 5G uses higher frequencies that, as we already discussed, have shorter ranges and are more susceptible to interference. This means that carriers will have to put up more cellular towers.

Each of the different mobile communication generations has taken advantage of the improvement of hardware technology and processing power. The increase in hardware has allowed for more complicated data transmission between users and hence the desire for more users to use mobile communications.

Table 12-2 illustrates some of the main features of the 2G through 5G networks. It is important to note that this table does not and cannot easily cover all the aspects of each generation. Earlier generations of mobile communication have considerable variability between countries. The variability was due to country-sponsored efforts before agreed-upon international standards were established. Various efforts between the ITU and countries have attempted to minimize the differences.

NOTE While it would be great if the mobile wireless technology generations broke down into clear-cut definitions, they do not. This is because various parts of the world use different foundational technologies, and there are several competing vendors in the space with their own proprietary approaches.

	2G	3G	4G	5G
Spectrum	1,800 MHz	2 GHz	Various	Various 3–86 GHz
Bandwidth	25 MHz	25 MHz	100 MHz	30–300 MHz
Multiplexing Type	TDMA	CDMA	OFDMA	OFDMA
New Features Introduced	Digital voice, SMS, MMS	Mobile Internet access, video	Mobile broadband, HD video	Ultra-HD and 3D video
Data Rate	115–128 Kbps	384 kbps	100 Mbps (moving) 1 Gbps (stationary)	Up to 10 Gbps
Introduction	1993	2001	2009	2018

Table 12-2 The Different Characteristics of Mobile Technology

Hacking Mobile Phones

2G networks (which are still around, believe it or not) lack the ability to authenticate towers to phones. In other words, an attacker can easily set up a rogue tower with more power than the nearby legitimate ones and cause the target's mobile phone to connect to it. This type of attack allows attackers to intercept all mobile phone traffic. Though 3G and 4G networks corrected this serious vulnerability, it is sometimes still possible to force most phones to switch down to 2G mode by jamming 3G, 4G, and 5G towers. In an effort to maintain some form of connectivity, handsets may then switch down to the vulnerable 2G mode, making the attack possible again.

Devices designed to perform this type of attack are called International Mobile Subscriber Identity (IMSI) catchers. Initially intended for law enforcement and intelligence agency use, IMSI catchers are increasingly available to criminals in the black markets. Moreover, it is possible for anyone to build one of these attack platforms for less than $1,500, as Chris Paget demonstrated at DefCon in 2010. This is yet another example of how backward compatibility can perpetuate vulnerabilities in older protocols.

Satellites

Today, satellites are used to provide wireless connectivity between distant stations. For two different locations to communicate via satellite links, they must be within the satellite's line of sight and *footprint* (area covered by the satellite), which tends to be large even for low Earth orbit satellites. The sender of information (ground station) modulates the data onto a radio signal that is transmitted to the satellite. A transponder on the satellite receives this signal, amplifies it, and relays it to the receiver. The receiver must have a type of antenna—one of those circular, dish-like things we see on top of buildings. The antenna contains one or more microwave receivers, depending upon how many satellites it is accepting data from.

Satellites provide broadband transmission that is commonly used for television channels and Internet access. If a user is receiving TV data, then the transmission is set up as a one-way (broadcast) network. If a user is using this connection for Internet connectivity, then the transmission is set up as a two-way network. The available bandwidth depends upon the antenna and terminal type and the service provided by the service provider. Time-sensitive applications, such as voice and video conferencing, can suffer from the delays experienced as the data goes to and from the satellite.

There are two types of orbits that are commonly used in satellite communications networks: geosynchronous and low Earth. Traditional networks, like the ones that broadcast TV and carry transoceanic data links for the major carriers, orbit at an altitude of 22,236 miles, which means they rotate at the same rate as the Earth does. This is called a *geosynchronous orbit*, and it makes the satellites appear to be stationary over the same spot on the ground. The key benefit is that the ground station antenna doesn't have to move. The main drawbacks are that, with that kind of range, you need a pretty big antenna and have to wait about a second for a radio wave to go up to the satellite and come back to Earth. This latency can create challenges for real-time communications like video conferencing.

Other satellites use a low Earth orbit (LEO), which is typically between 99 and 1,243 miles above the surface of the Earth. This means there is not as much distance between the ground stations and the satellites as in other types of satellites. In turn, this means smaller receivers can be used, which makes LEO satellites ideal for international cellular communication and Internet use. The catch is that the data rates tend to be much smaller than geosynchronous satellites and the service plans are pretty expensive.

In most cases, organizations use a system known as a very small aperture terminal (VSAT), which links a station (such as a remote office) to the Internet through a satellite gateway facility run by a service provider, as shown in Figure 12-6. Alternatively, VSATs can be deployed in stand-alone networks in which the organization also places a VSAT at a central location and has all the remote ones reach into it with no need for a gateway facility. The data rates available can range from a few Kbps to several Mbps. Dropping prices have rendered this technology affordable to many midsized organizations, though it is still far from being inexpensive.

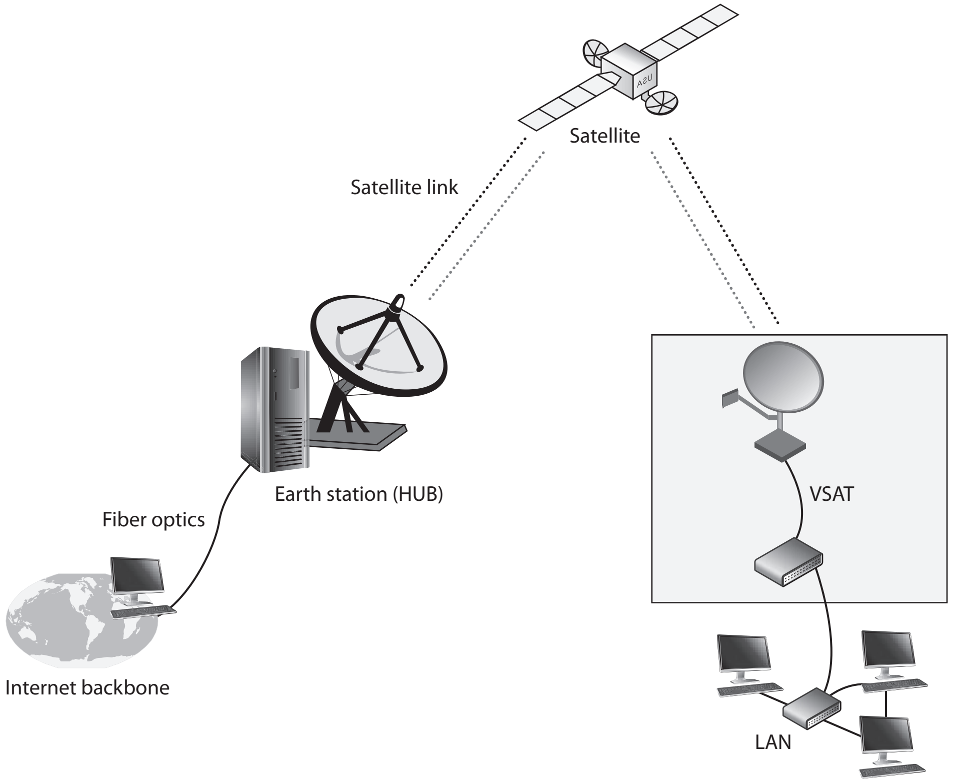


Figure 12-6 Satellite broadband

Chapter Review

Wireless networking is ubiquitous and, over the years, the security community has made great strides to ensure the confidentiality, integrity, and availability of our systems using these technologies. Still, risk can never be driven to zero, and this is particularly true when you transmit into free space, whether you do so using radio or light waves. Best practices for securing wireless networks include using strong cryptography, controlling access, and periodically testing the effectiveness of our controls.

As security professionals, we must always be aware of the myriad of new wireless technologies being developed and sold. For each, we have to compare the benefits (which are always touted by the vendors) to the risks (which may be less obvious and more difficult to identify). The market will constantly push products that promise new features and functionality, even if they come at the cost of security. To be clear, most new technologies incorporate at least some basic security features (and in many cases, advanced security features too), but these are not always implemented in a systematic manner by their adopters. That's where security professionals need to weigh in.

Quick Review

- Wireless communication systems modulate data onto electromagnetic signals like radio and light waves.
- Normally, a higher frequency can carry more data, but over a shorter distance and with more susceptibility to interference.
- Wireless communication systems typically use carrier sense multiple access with collision avoidance (CSMA/CA) as a medium access control (MAC) protocol.
- A radio frequency band is a subset of the radio spectrum designated for a specific use.
- Wi-Fi systems operate in the 2.4-GHz and 5-GHz bands.
- Most wireless communication systems use one of two modulation techniques: spread spectrum or orthogonal frequency division multiplexing (OFDM).
- Spread spectrum modulation techniques include frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS).
- DSSS spreads the data being transmitted over a wider spectrum than would otherwise be needed in order to overcome interference and relies on a chip sequence to let receiving stations know how to reconstruct the transmitted data.
- FHSS uses a single subchannel at a time, but rapidly changes channels in a particular hop sequence.
- Wireless local area networks (WLANs) come in two forms: infrastructure and ad hoc.
- Environments can be segmented into different WLANs by using different SSIDs.
- 802.11a provides up to 54 Mbps and operates in the 5-GHz band.

- 802.11b provides a transfer rate of up to 11 Mbps and works in the 2.4-GHz frequency range.
- 802.11g operates in the 2.4-GHz band and supports data rates of up to 54 Mbps.
- 802.11n, also known as Wi-Fi 4, supports throughputs of up to 100 Mbps and works in the 5-GHz band.
- IEEE 802.11ac (Wi-Fi 5) is an extension of 802.11n that increases throughput to 1.3 Gbps and is backward compatible with 802.11a, 802.11b, 802.11g, and 802.11n.
- The 802.11ax standard aims to address efficiency rather than faster speeds.
- Li-Fi is a wireless networking technology that uses light rather than radio waves to transmit and receive data.
- 802.16 is a metropolitan area network (MAN) wireless standard that allows wireless traffic to cover large geographical areas where stations can be as far as 70 km apart, using the 2.4-GHz and 5-GHz bands.
- The 802.15.4 standard defines the physical layer and Media Access Control sublayer of wireless personal area networks (WPANs).
- ZigBee is a standard for layers 3 (network) and 7 (application) that is built on top of 802.15.4 and is most commonly used in Internet of Things (IoT) and Industrial IoT systems.
- Bluetooth is another standard for WPANs, which is most commonly used to replace the cables connecting peripherals to computers and mobile devices.
- The 802.11e standard provides Quality of Service (QoS) and support of multimedia traffic in wireless transmissions.
- 802.11f standardizes the processes by which access points transfer active connections among themselves, enabling users to roam across APs.
- The 802.11h standard was developed to address interference issues in the 5-GHz band, particularly with regard to radar and satellite systems, through Dynamic Frequency Selection (DFS) and Transmit Power Control (TPC) techniques.
- 802.11j is an example of standards that allow common technologies such as WLANs to be employed in countries (in this case Japan) where local regulations conflict with portions of a broader standard (in this case 802.11).
- 802.11 was the original WLAN standard, which included Wired Equivalent Privacy (WEP); it is now obsolete.
- 802.11i defines Wi-Fi Protected Access 2 (WPA2) and is the most common standard in use in WLANs today.
- The IEEE 802.11w standard provides Management Frame Protection (MFP) that prevents certain types of attacks, such as replay and DoS attacks.
- WPA3 was developed by the Wi-Fi alliance (not IEEE) and is quickly replacing WPA2 for both personal and enterprise use.

- 802.1X is an access control protocol that can be implemented on both wired and wireless networks for user authentication and key distribution.
- Mobile telephony has gone through different generations and multiple access technologies: 1G (FDMA), 2G (TDMA), 3G (CDMA), 4G (OFDM), and 5G (OFDM).
- Satellite communications links provide connectivity across very long distances and in places that would otherwise not be reachable, but may introduce latency challenges.

Questions

Please remember that these questions are formatted and asked in a certain way for a reason. Keep in mind that the CISSP exam is asking questions at a conceptual level. Questions may not always have the perfect answer, and the candidate is advised against always looking for the perfect answer. Instead, the candidate should look for the best answer in the list.

1. Which of the following is not a characteristic of the IEEE 802.11a standard?
   - **A.** It works in the 5-GHz range.
   - **B.** It uses the OFDM spread-spectrum technology.
   - **C.** It provides 52 Mbps in bandwidth.
   - **D.** It covers a smaller distance than 802.11b.
2. Wireless LAN technologies have gone through different versions over the years to address some of the inherent security issues within the original IEEE 802.11 standard. Which of the following provides the correct characteristics of WPA2 in Enterprise mode?
   - **A.** IEEE 802.1X, WEP, MAC
   - **B.** IEEE 802.1X, EAP, TKIP
   - **C.** IEEE 802.1X, EAP, WEP
   - **D.** IEEE 802.1X, EAP, CCMP
3. Which of the following is *not* a characteristic of Li-Fi networks?
   - **A.** Support for high client densities
   - **B.** High latency
   - **C.** Constrained coverage area
   - **D.** Can work on the infrared spectrum
4. How would you best ensure the security of a ZigBee system?
   - **A.** Ensure a coordinator acts as a Trust Center
   - **B.** Use 256-bit encryption keys
   - **C.** Deploy in a ring topology with preassigned slots for each device
   - **D.** Use the Symmetric-Key Key Establishment (SKKE) protocol to derive keys

5. Which of the following is a Bluetooth-specific attack that allows unauthorized read/write access from a wireless device?

   A. Bluejacking

   B. Replay attack

   C. Smurf attack

   D. Bluesnarfing

6. What does the IEEE 802.1X standard cover?

   A. A Management Frame Protection (MFP) that prevents replay and denial-of-service (DoS) attacks

   B. Wi-Fi Protected Access 2 (WPA2)

   C. Security extensions to the physical layer (PHY) and Media Access Control (MAC) sublayer of the data link layer in the OSI model

   D. An access control protocol for user authentication and key distribution

7. Which of the following is not a disadvantage of satellite networks compared to terrestrial ones?

   A. Latency

   B. Cost

   C. Bandwidth

   D. Video conferencing

Use the following scenario to answer Questions 8–10. You are planning an upgrade for the wireless network at one of your manufacturing sites and want to use this as an opportunity to improve network security. The current system is based on 10-year-old wireless access points (WAPs) that implement 802.11g. You're using WPA2 in Personal mode because you have multiple Industrial Internet of Things (IIoT) devices. You can update the firmware on the WAPs, but you really think it's time for an upgrade.

8. What could make it harder for you to switch from WPA2 Personal mode to Enterprise mode?

   A. Enterprise mode requires licenses that can be costly.

   B. The WAPs may not support Enterprise mode.

   C. IIoT devices may not support Enterprise mode.

   D. The return on investment is insufficient.

9. What is the best technology to which you should consider upgrading?

   A. IEEE 802.16

   B. IEEE 802.11w

   C. IEEE 802.11f

   D. IEEE 802.11ax

10. The existing wireless network has recently become unusable, and you suspect you may be the target of a persistent Wi-Fi deauthentication attack. How can you best mitigate this threat?

    A. Deploy WPA3 access points across the facility

    B. Perform MAC address filtering to keep the rogue stations off the network

    C. Immediately update the firmware on the access points to support 802.11w

    D. Change the channel used by the WAPs

Answers

1. **C.** The IEEE standard 802.11a uses the OFDM spread-spectrum technology, works in the 5-GHz frequency band, and provides bandwidth of up to 54 Mbps. The operating range is smaller because it works at a higher frequency.
2. **D.** Wi-Fi Protected Access 2 requires IEEE 802.1X or preshared keys for access control, Extensible Authentication Protocol (EAP) or preshared keys for authentication, and the Advanced Encryption Standard (AES) algorithm in counter mode with CBC-MAC Protocol (CCMP) for encryption.
3. **B.** Latency is the delay in data transfers, which is extremely low in Li-Fi networks.
4. **A.** Using a Trust Center provides a way to centrally authenticate devices and securely manage encryption keys, which are 128 bits (not 256). Without a Trust Center, the SKKE protocol can be used to derive keys, but this approach is not as secure. ZigBee does not support ring topologies.
5. **D.** Bluesnarfing could allow an attacker to read, modify, or delete calendar events, contacts, e-mails, text messages, and so on. Bluejacking is the only other Bluetooth attack option, but this refers to someone sending an unsolicited message to a device.
6. **D.** 802.1X is an access control protocol that can be implemented on both wired and wireless networks for user authentication and key distribution. MFP is covered in 802.11w, WPA2 is covered in 802.11i, and the other option (security extensions) was a distracter.
7. **C.** If you have the budget for it, data rates on satellite networks are comparable with other modes of communication. These systems, however, are typically more expensive and have high latencies, which means they are not well suited for time-sensitive applications, such as voice and video conferencing.
8. **D.** If a WAP supports WPA2, it would do so in either Personal or Enterprise mode as long as it can be connected to the needed backend services (e.g., a RADIUS server), with no need for additional licensing. Thus, the change would not typically be expected to have ROI issues. However, many embedded devices, including IIoT, do not support this mode and would have to be replaced.

9. **D.** 802.11ax is the only standard describing a WLAN among the list of options. 802.16 is used in metropolitan area networks (MANs). 802.11w covers Management Frame Protection (MFP) in wireless networks. 802.11f deals with users roaming among access points.

10. **C.** 802.11w provides Management Frame Protection (MFP) capabilities that would mitigate this type of attack. This is included in WPA3, so either answer would generally work. However, it is probably faster, cheaper, and safer to roll out 802.11w upgrades first, which would likely have no negative effects on the networks, while research and planning continue on how to best implement a WPA3 solution across the enterprise. This is a good example of the types of ambiguous questions you'll see on the CISSP exam.

This page intentionally left blank

CHAPTER 13

Securing the Network

This chapter presents the following:

- Secure networking
- Secure protocols
- Multilayer protocols
- Converged protocols
- Micro-segmentation

More connections to more devices means more vulnerabilities.

—Marc Goodman

Having developed a foundational understanding of networking technologies, we now turn our attention to building secure networks upon this foundation. In this chapter, we circle back to the core networking and service protocols introduced in Chapter 11 and discuss the threats against them and how to mitigate those threats. This discussion is grounded in the secure design principles covered in Chapter 9. We'll take the same approach as we expand our scope of interest from those core protocols and services to include other services, such as e-mail, that are critical to modern networks.

These networks are not as neatly divided as the OSI model could lead us to believe. Increasingly, we are relying on multilayer and converged protocols where concepts from different layers and even network components overlap in ways that have important security implications. The goal of this chapter is to show how, through a thoughtful application of secure protocols and best practices, we can secure our networks and the services they provide.

Applying Secure Design Principles to Network Architectures

A network architecture is just a model of a network. Like any model, it is not 100 percent representative of reality and uses abstractions to simplify some details so that we can focus on the others. By ignoring the little details (for now), we make it easier on ourselves to focus on the more important elements. For example, before we decide how many web servers we need and which operating systems and software we need to run on them, we should first identify the classes of servers and where we would put them. We might have a set of externally accessible servers for our web presence, but we may also need some servers that are for internal use only by all employees, and yet another set that is only for web developers. Where do we put each set and how might we need different controls for them? Maybe we need a demilitarized zone (DMZ), an internal sharing cluster, and a development virtual local area network (VLAN), each with specific sets of controls meant to mitigate their differing risk profiles. A network architecture allows us to answer these high-level questions before we start configuring any boxes.

Now, once we go through all the trouble of coming up with an architecture that works, we shouldn't have to reinvent the wheel. Network architectures also serve as templates for future systems. What's more, they can be codified and shared among similar organizations to reduce work and ensure we all follow best practices. Even if a lot of the details are different, a sound architecture can be reused time and again.

Many of these best practices relate to security. Since we intend our architectures to be reusable, it is imperative that we apply secure design principles when we implement them. In the sections that follow, we will discuss a (wide) variety of networking concepts and technologies that you will need to understand to implement secure design principles in network architectures. Periodically, we circle back and discuss some of these important secure design principles. It is important to note that there is no one-size-fits-all solution in this effort, so you will have to be selective about which of these principles you apply in any specific situation. Still, as a CISSP, you are expected to be conversant with all of them.

Let's start by reviewing the 11 secure design principles we covered in Chapter 9 and look at how they apply to network architectures.

- **Threat modeling** Everything we do in cybersecurity should be grounded in a good understanding of the threats we face. In this chapter, we focus our attention on network security, so we'll illustrate the threats we face as we discuss the various technologies and protocols involved in operating and securing our networks.
- **Least privilege** Traffic should be allowed to flow between any two points that are required to communicate in order to satisfy a valid organizational requirement, and nowhere else. We cover this in depth when we address network segmentation later in this chapter.
- **Defense in depth** While some IT and security professionals equate this principle with having a DMZ for public-facing servers, the principle applies throughout the network and requires that we build concentric defenses around our most valuable assets.
- **Secure defaults** Perhaps the simplest illustration of this principle as it applies to our networks is ensuring firewalls' default configurations are to deny all traffic from any source to any destination (deny all all). However, the principle should apply throughout our network and be consistent with least privilege.
- **Fail securely** The key to applying this principle is asking two questions: What happens when this network system fails? What happens when a packet doesn't match an "allow" rule on the firewall? (Hint: it should not be allowed through.)

- **Separation of duties** Speaking of firewall (and other security appliance) rules, who is in charge of those in your organization? Any sensitive duties should be split up among vetted staff members. At a minimum, if you don't have enough staff, everybody's sensitive work should be regularly checked by someone else.
- **Keep it simple** Unless you are architecting a global network for a multinational corporation, you should try to develop an architecture that can be depicted in a single PowerPoint slide and still describe all the important components.
- **Zero trust** Services and traffic on your network should all be authenticated and encrypted. When two servers are part of a system (e.g., the web server and its backend database), they should authenticate each other and have rules around what requests each is allowed to make of the other.
- **Privacy by design** Encrypting your network traffic is a good start toward protecting privacy, but where is the data being collected and for what purpose? For example, as we prepare for auditability (see the next principle), we need to ensure that we are not casting too wide of a net in terms of the data we log.
- **Trust but verify** Everything that happens on the network should be auditable, meaning that there should be a record of who is talking with whom, when, and why. This is normally done by ensuring logs are properly configured and protected against tampering or accidental loss.
- **Shared responsibility** Odds are that your network architecture will include at least a handful of service providers. Whether these are Internet service providers, cloud service providers, or managed services providers, it is critical to agree on who has responsibility over which aspects of your network.

EXAM TIP You should be prepared to map the various secure design principles to specific scenarios.

With these principles in mind, let's look at specific ways in which we can assess and implement network architectures securely.

Secure Networking

The most prevalent networking standards and protocols we use today (Ethernet, TCP/IP, and so on) were born decades ago (before many of us). Back then, the world was kinder and gentler (at least in the digital realm) and security just wasn't the sort of thing folks thought about when it came to computers and networks. With the explosion of the Internet came immense opportunities for both the law abiding and the criminals. The need for secure networking became apparent, but it was too late. We've been trying to catch up ever since by bolting security onto insecure technologies. One of the most common ways of securing our networks is through the use of encryption, particularly in trusted tunnels through untrusted networks.

Link Encryption vs. End-to-End Encryption

In each of the networking technologies discussed in this chapter, encryption can be performed at different levels, each with different types of protection and implications. Two general modes of encryption implementation are link encryption and end-to-end encryption. *Link encryption* encrypts all the data along a specific communication path, as in a satellite link, a terrestrial T3 leased line, or even between hosts on the same LAN. Because link encryption happens at layers 1 and 2, not only is the user information encrypted, but the (layer 3 and higher) headers, trailers, addresses, and routing data that are part of the packets are also encrypted. The only traffic not encrypted in this technology is the data link control messaging information, which includes instructions and parameters that the different link devices use to synchronize communication methods. Reading this information won't give an attacker any insights into what is being transmitted or where it is ultimately going.

End-to-end encryption (E2EE) occurs at the session layer (or higher), which means the headers, addresses, routing information, and trailer information are not encrypted, enabling attackers to learn more about a captured packet and where it is headed. Transport Layer Security (TLS), which we will discuss shortly, is the most common example of E2EE. Because the routing information is sent in plaintext, attackers can perform traffic analysis to learn details about the network, such as which hosts play which roles in it.

Link encryption, which is sometimes called *online encryption*, is usually provided by service providers and is incorporated into network protocols. All of the information is encrypted, and the packets must be decrypted at each hop so the router, or other intermediate device, knows where to send the packet next. The router must decrypt the header portion of the packet, read the routing and address information within the header, and then re-encrypt it and send it on its way.

With end-to-end encryption, the packets do not need to be decrypted and then encrypted again at each hop because the headers and trailers are not encrypted. The devices in between the origin and destination just read the necessary routing information and pass the packets on their way.

End-to-end encryption is usually initiated by the user of the originating computer. It provides more flexibility for the user to be able to determine whether or not certain

Encryption at Different Layers

Encryption can (and typically does) happen at different layers of an operating system and network stack. The following are just a few examples:

- End-to-end encryption happens within the applications.
- TLS encryption takes place at the session layer.
- Point-to-Point Tunneling Protocol (PPTP) encryption takes place at the data link layer.
- Link encryption takes place at the data link and physical layers.

messages will get encrypted. It is called "end-to-end encryption" because the message stays encrypted from one end of its journey to the other. Link encryption has to decrypt the packets at every device between the two ends.

Link encryption occurs at the data link and physical layers, as depicted in Figure 13-1. Hardware encryption devices interface with the physical layer and encrypt all data that passes through them. Because no part of the data is available to an attacker, the attacker cannot learn basic information about how data flows through the environment. This is referred to as *traffic-flow security*.

NOTE A *hop* is a device that helps a packet reach its destination. It is usually a router that looks at the packet address to determine where the packet needs to go next. Packets usually go through many hops between the sending and receiving computers.

Advantages of end-to-end encryption include the following:

- It provides more flexibility to the user in choosing what gets encrypted and how.
- Higher granularity of functionality is available because each application or user can choose specific configurations.
- Each hop device on the network does not need to have a key to decrypt each packet.

The disadvantage of end-to-end encryption is the following:

- Headers, addresses, and routing information are not encrypted, and therefore not protected.

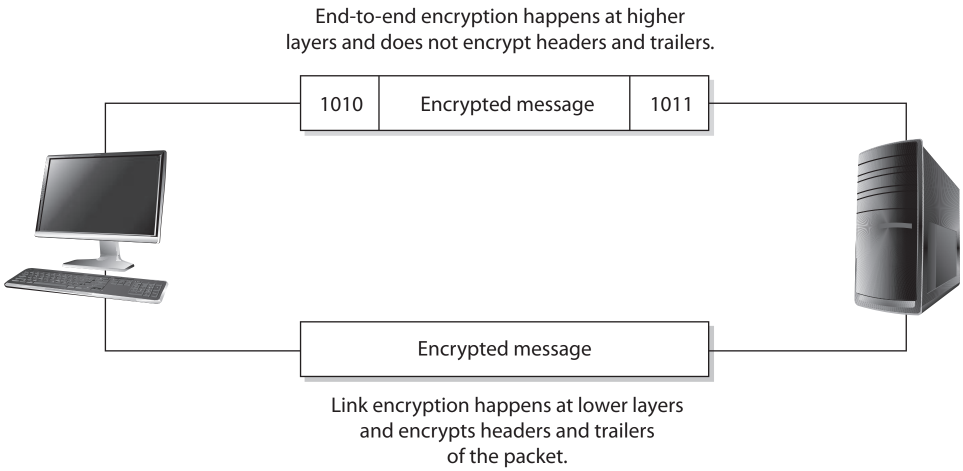

End-to-end encryption happens at higher layers and does not encrypt headers and trailers.
1010
Encrypted message
1011
Encrypted message
Link encryption happens at lower layers and encrypts headers and trailers of the packet.

Figure 13-1 Link and end-to-end encryption happen at different OSI layers.

Hardware vs. Software Cryptography Systems

Encryption can be done through software or hardware, and there are trade-offs with each. Generally, software is less expensive and provides a slower throughput than hardware mechanisms. Software cryptography methods can be more easily modified and disabled compared to hardware systems, but it depends on the application and the hardware product.

If an organization needs to perform high-end encryption functions at a higher speed, it will most likely implement a hardware solution.

Advantages of link encryption include the following:

- All data is encrypted, including headers, addresses, and routing information.
- Users do not need to do anything to initiate it. It works at a lower layer in the OSI model.

Disadvantages of link encryption include the following:

- Key distribution and management are more complex because each hop device must receive a key, and when the keys change, each must be updated.
- Packets are decrypted at each hop; thus, more points of vulnerability exist.

TLS

The most prevalent form of end-to-end encryption is *Transport Layer Security (TLS)*. TLS is a security protocol that provides confidentiality and data integrity for network communications. It replaced the (now insecure) Secure Sockets Layer (SSL) standard. These two protocols coexisted for many years, and most people thought that there were very few differences between SSL and TLS (TLS is currently in version 1.3). However, the Padding Oracle On Downgraded Legacy Encryption (POODLE) attack in 2014 was the death knell of SSL and demonstrated that TLS was superior security-wise. The key to the attack was to force SSL to downgrade its security, which was allowed for the sake of interoperability.

EXAM TIP Because SSL and TLS were (for a time) very closely related, the terms are sometimes still used interchangeably to describe network encryption in general. However, the SSL protocol has been insecure for many years and should not be the correct answer to an encryption question (unless it is asking for an insecure protocol).

Backward compatibility has long been a thorn in the side of those of us trying to improve cybersecurity. TLS 1.3 represents a switch to a focus on security, which shows in the limited number of cipher suites that it supports (just five). This means attackers can no longer trick a server into using an insecure cryptosystem during the connection establishment negotiation. One of the key features of TLS 1.3 is that the handshake used to establish a new connection requires only one client message to the server and one response from the server. There's a lot that happens in there, though, so let's take a look at a summarized version of this handshake.

1. Client "Hello" message, which includes
   - A list of cipher suites and protocols supported by the client
   - Client inputs for the key exchange
2. Server "Hello" message, which includes
   - The server's selection of cipher suite and protocol version
   - Server inputs for the key exchange
3. Server authentication, which includes
   - The server's digital certificate
   - Proof that the server owns the certificate's private key
4. (Optionally) Client authentication, which includes
   - The client's digital certificate
   - Proof that the client owns the certificate's private key

NOTE While TLS 1.3 minimizes the plaintext information transferred between hosts, TLS 1.2 (and earlier) passes a lot more information in the clear, potentially including the server name (e.g., www.goodsite.com).

As mentioned, TLS 1.3 has dramatically reduced the number of recommended cipher suites from 37 (in previous versions) to just five. This is an important improvement because some of those 37 suites were known (or suspected) to be vulnerable to cryptanalysis. By reducing the suites to five and ensuring these provide strong protection, TLS 1.3 makes it harder for attackers to downgrade the security of a system by forcing a server to use a weaker suite. The allowed suites in the latest version of TLS are as follows:

- **TLS_AES_256_GCM_SHA384** The encryption algorithm here is AES with a 256-bit key in Galois/Counter Mode (GCM). GCM is a mode of operation that provides message authentication. The hashing algorithm is SHA-384. This suite provides the best protection but requires the most computing resources.
- **TLS_AES_128_GCM_SHA256** This suite is almost identical to the preceding one, but saves on resources by using a smaller 128-bit key for encryption and a slightly faster SHA-256 for hashing. It is ideally suited for systems with hardware support for encryption.
- **TLS_AES_128_CCM_SHA256** In this suite, AES (again, with a 128-bit key) runs in Counter mode with CBC-MAC (CCM), which uses 16-byte tags to provide message authentication (much like GCM does).

- **TLS_AES_128_CCM_8_SHA256** This suite is almost identical to the preceding one, but Counter mode with CBC-MAC uses 8-byte tags (instead of 16-byte ones), which makes it better suited for embedded devices.
- **TLS_CHACHA20_POLY1305_SHA256** The ChaCha stream cipher (doing 20 rounds), combined with the Poly1305 message authentication code (MAC), is a cipher suite that is a good choice for software-based encryption systems. Many modern systems rely on hardware-based encryption, so the authors of TLS 1.3 wanted to ensure the recommended suites supported multiple devices. Besides, it just makes sense to have at least one encryption algorithm that is not AES.

We already discussed AES (and briefly mentioned ChaCha20) in Chapter 8, and CCM in Chapter 12, but this is the first time we bring up GCM and Poly1305. These are approaches to provide authenticated symmetric key encryption. *Authenticated encryption (AE)* provides assurances that a message was not modified in transit and could only come from a sender who knows the secret key. This is similar to the MAC approach discussed in Chapter 8 but is applied to stream ciphers. TLS 1.3 takes the AE concept to the next level in what is known as *authenticated encryption with additional data (AEAD)*. AEAD essentially computes the MAC over both ciphertext and plaintext when these are sent together. For example, when sending network traffic, there are certain fields (e.g., source and destination addresses) that cannot be encrypted. An attacker could replay an encrypted message later using a different packet, but if we're using AEAD (as TLS 1.3 requires), this bogus packet would automatically be discarded.

Another key feature of TLS 1.3 (which was optional in TLS 1.2 and prior) is its use of *ephemeral keys*, which are only used for one communication session and then discarded, using the Diffie-Hellman Ephemeral (DHE) algorithm. This provides *forward secrecy* (sometimes called *perfect forward secrecy*), which means that if attackers were somehow able to crack or otherwise obtain the secret key, it would only give them the ability to decrypt a small portion of the ciphertext. They wouldn't be able to decrypt everything going forward.

Attackers Use TLS Too!

While TLS is often our first line of defense in protecting our network traffic from prying eyes, attackers use it too, precisely for the same reason. There are many known examples of malware using TLS. Banking Trojans, such as TrickBot, Emotet, and Dyre, make use of TLS to communicate data back to their master server. Ransomware families, such as Jigsaw, Locky, and Petya, have also used TLS to infect machines and transfer information. They way in which attackers use TLS, however, is usually quite different from how it is used in legitimate connections. Analyzing network traffic can often point out some of these differences, such as:

- Offering weak or obsolete cipher suites
- Rarely offering more than one extension (enterprise clients use up to nine)
- Using self-signed certificates

While we focused on TLS 1.3 in this section, it is worth noting that, as of this writing, the Internet Society reports that only 58 percent of the world's top 1,000 websites support this latest version. What does this mean to you? You should balance the enhanced security of this protocol with the needs of your stakeholders. If you are not on TLS 1.3 yet, you may want to ask what percentage of your user base would not be able to communicate securely if you switched. All major browsers support it, so odds are that you'd be in good shape. But even if you're still on TLS 1.2, keep in mind that most of the features described in this section that make 1.3 so much better are optional in the previous version. This should give you a path to gradually improve your security while taking care of your stakeholders. Whatever your situation, TLS is probably the most important encryption technology for securing our networks, particularly our virtual private ones.

NOTE TLS 1.0 and TLS 1.1 were never formally deprecated but are widely considered insecure.

VPN

A *virtual private network (VPN)* is a secure, private connection through an untrusted network, as shown in Figure 13-2. It is a private connection because the encryption and tunneling protocols are used to ensure the confidentiality and integrity of the data in transit. It is important to remember that VPN technology requires a tunnel to work and it assumes encryption.

We need VPNs because we send so much confidential information from system to system and network to network. The information can be credentials, bank account data, Social Security numbers, medical information, or any other type of data we do not want to share with the world. The demand for securing data transfers has increased over the years, and as our networks have increased in complexity, so have our VPN solutions.

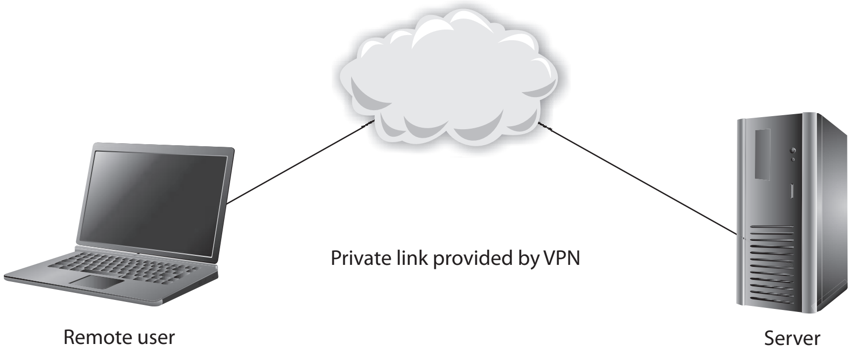


Figure 13-2 A VPN provides a virtual dedicated link between two entities across a public network.

Point-to-Point Tunneling Protocol

One of the early approaches to building VPNs was Microsoft's *Point-to-Point Tunneling Protocol (PPTP)*, which uses Generic Routing Encapsulation (GRE) and TCP to encapsulate Point-to-Point Protocol (PPP) connections and extend them through an IP network (running on TCP port 1723, by default). Since most Internet-based communication first started over telecommunication links, the industry needed a way to secure PPP connections, which were prevalent back then. The original goal of PPTP was to provide a way to tunnel PPP connections through an IP network, but most implementations included security features also since protection was becoming an important requirement for network transmissions at that time. PPTP, like many security protocols, did not age well and is now considered insecure and obsolete.

Layer 2 Tunneling Protocol

The *Layer 2 Tunneling Protocol (L2TP)*, currently in version 3, is a combination of Cisco's *Layer 2 Forwarding (L2F)* protocol and Microsoft's PPTP. L2TP tunnels PPP traffic over various network types (IP, ATM, X.25, etc.); thus, it is not just restricted to IP networks as PPTP was. PPTP and L2TP have very similar focuses, which is to get PPP traffic to an end point that is connected to some type of network that does not understand PPP. Unlike PPTP, L2TP runs on UDP (default port 1701), which makes it a bit more efficient. However, just like PPTP, L2TP does not actually provide much protection for the PPP traffic it is moving around, but it integrates with protocols that *do* provide security features. L2TP inherits PPP authentication and integrates with IPSec to provide confidentiality, integrity, and potentially another layer of authentication.

It can get confusing when several protocols are involved with various levels of encapsulation, but if you do not understand how they work together, you cannot identify if certain traffic links lack security. To figure out if you understand how these protocols work together and why, ask yourself these questions:

1. If the Internet is an IP-based network, why do we even need PPP?
2. If L2TP does not actually secure data, then why does it even exist?
3. If a connection is using IP, PPP, and L2TP, where does IPSec come into play?

Let's go through the answers together. Let's say that you are a remote user and work from your home office. You do not have a dedicated link from your house to your company's network; instead, your traffic needs to go through the Internet to be able to communicate with the corporate network. The line between your house and your ISP is a point-to-point telecommunications link, one point being your home router and the other point being the ISP's switch, as shown in Figure 13-3. Point-to-point telecommunication devices do not understand IP, so your router has to encapsulate your traffic in a protocol the ISP's device will understand—PPP. Now your traffic is not headed toward some website on the Internet; instead, it has a target of your company's corporate network. This means that your traffic has to be "carried through" the Internet to its ultimate destination through a tunnel. The Internet does not understand PPP, so your PPP traffic has to be encapsulated with a protocol that can work on the Internet and create the needed tunnel.

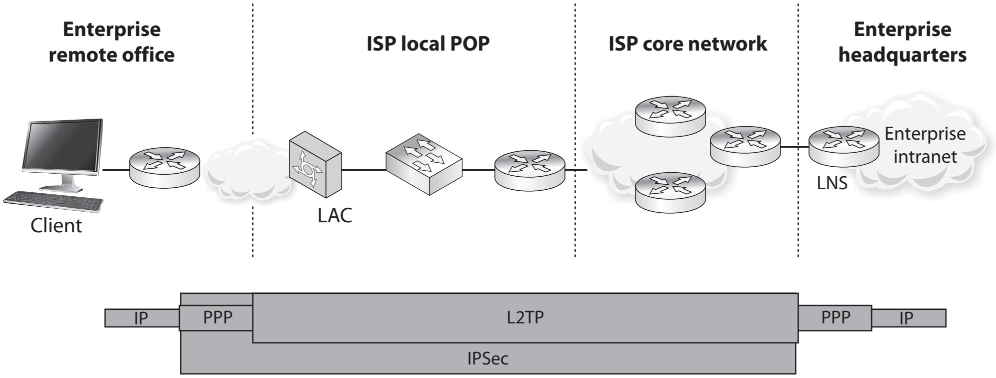


Figure 13-3 IP, PPP, L2TP, and IPSec can work together.

So your IP packets are wrapped up in PPP, and are then wrapped up in L2TP. But you still have no encryption involved, so your data is actually not protected. This is where IPSec comes in. IPSec is used to encrypt the data that will pass through the L2TP tunnel. Once your traffic gets to the corporate network's perimeter device, it will decrypt the packets, take off the L2TP and PPP headers, add the necessary Ethernet headers, and send these packets to their ultimate destination.

Here are the answers to our questions:

1. If the Internet is an IP-based network, why do we even need PPP?
   Answer: The point-to-point line devices that connect individual systems to the Internet do not understand IP, so the traffic that travels over these links has to be encapsulated in PPP.
2. If L2TP does not actually secure data, then why does it even exist?
   Answer: It extends PPP connections by providing a tunnel through networks that do not understand PPP.
3. If a connection is using IP, PPP, and L2TP, where does IPSec come into play?
   Answer: IPSec provides the encryption, data integrity, and system-based authentication.

Here is another question: Does all of this PPP, L2TP, and IPSec encapsulation have to happen for every single VPN used on the Internet? No, only when connections over point-to-point connections are involved. When two gateway routers are connected over the Internet and provide VPN functionality, they only have to use IPSec.

Internet Protocol Security

Internet Protocol Security (IPSec) is a suite of protocols that was developed to specifically protect IP traffic. IPv4 does not have any integrated security, so IPSec was developed to "bolt onto" IP and secure the data the protocol transmits. Where L2TP works at the data link layer, IPSec works at the network layer of the OSI model.

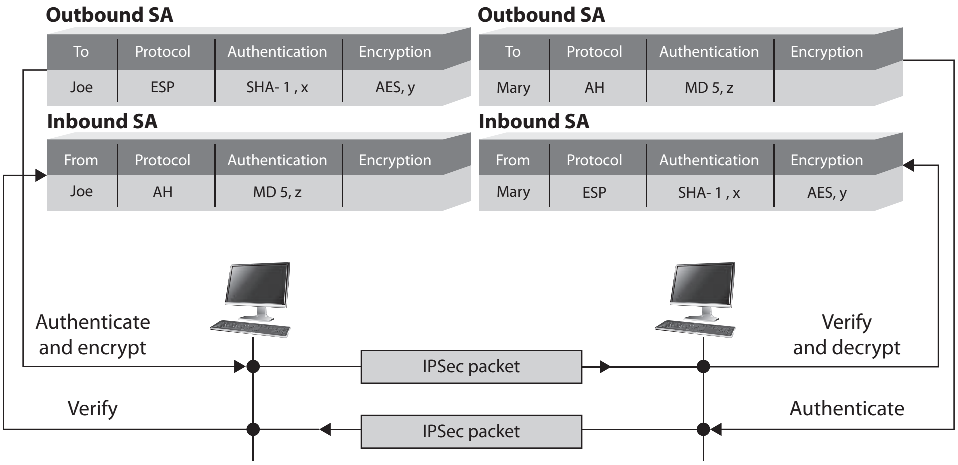
The main protocols that make up the IPSec suite and their basic functionality are as follows:

- **Authentication Header (AH)** Provides data integrity, data-origin authentication, and protection from replay attacks
- **Encapsulating Security Payload (ESP)** Provides confidentiality, data-origin authentication, and data integrity
- **Internet Security Association and Key Management Protocol (ISAKMP)** Provides a framework for security association creation and key exchange
- **Internet Key Exchange (IKE)** Provides authenticated keying material for use with ISAKMP

AH and ESP can be used separately or together in an IPSec VPN configuration. The AH protocols can provide data-origin authentication (system authentication) and protection from unauthorized modification, but do not provide encryption capabilities. If the VPN needs to provide confidentiality, then ESP has to be enabled and configured properly.

When two routers need to set up an IPSec VPN connection, they have a list of security attributes that need to be agreed upon through handshaking processes. The two routers have to agree upon algorithms, keying material, protocol types, and modes of use, which will all be used to protect the data that is transmitted between them.

Let's say that you and Juan are routers that need to protect the data you will pass back and forth to each other. Juan sends you a list of items that you will use to process the packets he sends to you. His list contains AES-128, SHA-1, and ESP tunnel mode. You take these parameters and store them in a security association (SA). When Juan sends you packets one hour later, you will go to this SA and follow these parameters so that you know how to process this traffic. You know what algorithm to use to verify the integrity of the packets, the algorithm to use to decrypt the packets, and which protocol to activate and in what mode. Figure 13-4 illustrates how SAs are used for inbound and outbound traffic.



Outbound SA

To	Protocol	Authentication	Encryption
Joe	ESP	SHA- 1 , x	AES, y

Inbound SA

From	Protocol	Authentication	Encryption
Joe	AH	MD 5, z	

Outbound SA

To	Protocol	Authentication	Encryption
Mary	AH	MD 5, z	

Inbound SA

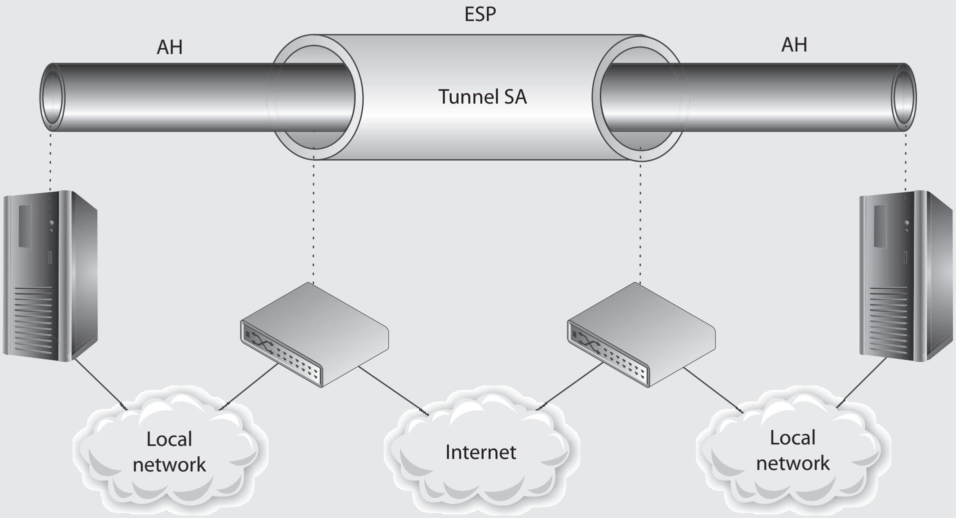
From	Protocol	Authentication	Encryption
Mary	ESP	SHA- 1 , x	AES, y

Figure 13-4 IPSec uses security associations to store VPN parameters.

NOTE The U.S. National Security Agency (NSA) uses a protocol encryptor that is based upon IPSec. A *HAIPE (High Assurance Internet Protocol Encryptor)* is a Type 1 encryption device that is based on IPSec with additional restrictions, enhancements, and capabilities. A HAIPE is typically a secure gateway that allows two enclaves to exchange data over an untrusted or lower-classification network. Since this technology works at the network layer, secure end-to-end connectivity can take place in heterogeneous environments. This technology has largely replaced link layer encryption technology implementations.

IPSec

IPSec can be configured to provide *transport adjacency*, which just means that more than one security protocol (ESP and AH) is used in a VPN tunnel. IPSec can also be configured to provide *iterated tunneling*, in which an IPSec tunnel is tunneled through another IPSec tunnel, as shown in the following diagram. Iterated tunneling would be used if the traffic needed different levels of protection at different junctions of its path. For example, if the IPSec tunnel started from an internal host and terminated at an internal border router, this may not require encryption, so only the AH protocol would be used. But when that data travels from that border router throughout the Internet to another network, then the data requires more protection. So the first packets travel through a semisecure tunnel until they get ready to hit the Internet and then they go through a very secure second tunnel.


ESP
AH
AH
Tunnel SA
Local network
Internet
Local network