1

Welcome to the Cloud

OBJECTIVES

By the end of this chapter, you should be able to:

1. Describe basic networking, including why networks are drawn as clouds, hosts, addresses, the Internet, and Internet service providers.
2. Explain how the Internet works, how Netflix uses Amazon Web Services IaaS (Infrastructure as a Service) with virtual machines, and a Google SaaS (Software as a Service).
3. Describe messages, fragmentation, multiplexing, and frames versus packets.
4. Describe how single networks operate—especially how switches forward incoming frames.
5. Describe how internets and routers make it possible for hosts on different networks to work together.
6. List the five layers commonly encountered in network standards, describe what each layer does, describe concepts and terms in each layer, identify at which layer a given process is operating, and identify which standards agencies and standards architecture are relevant to that process.

# BOX: By the Numbers

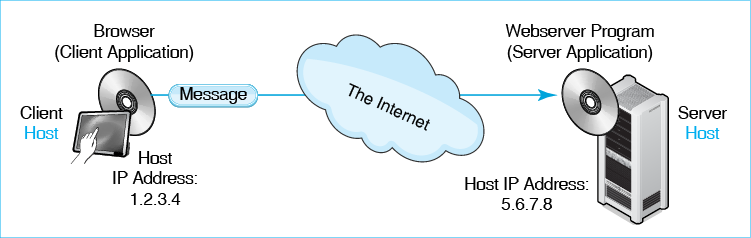
The Internet is enormous, growing, and changing.

* By 2003, there were already more devices connected to the Internet (computers, phones, etc.) than there were human users.[[1]](#footnote-1)
* In 2010, 21% of the world’s population used the Internet. In 2013, it was 39%.[[2]](#footnote-2)
* In 2012, online video viewing overtook DVD and Blu-Ray viewing.[[3]](#footnote-3)
* From 2011 to 2016, global IP traffic will triple, and the number of connected devices will nearly double.[[4]](#footnote-4)
* In 2016, Cisco expects the Internet to carry one zettabyte of data.[[5]](#footnote-5) A zettabyte is 1,000,000,000,000,000,000 (one sextillion) bytes.
* By 2020, there will be 50 billion devices connected to the Internet—ten times the number of human users. The great majority of these will be devices talking to other devices, without human involvement.[[6]](#footnote-6)

# Netflix Dives Into the Amazon[[7]](#footnote-7)

Figure 1-1 shows two computers communicating over the Internet. In this figure, the Internet is a cloud. Just as you do not see inside a real cloud, users should not have to understand the details of the Internet. The Internet should simply work, like electricity, water, and telephones.

Figure 1-1: Internet Communication



Test Your Understanding

1. What is the significance of the Internet being depicted as a cloud?

## Hosts, Messages, and Addresses

By not showing details of the Internet, Figure 1-1 allows us to focus on some concepts and terms we will see throughout this course.

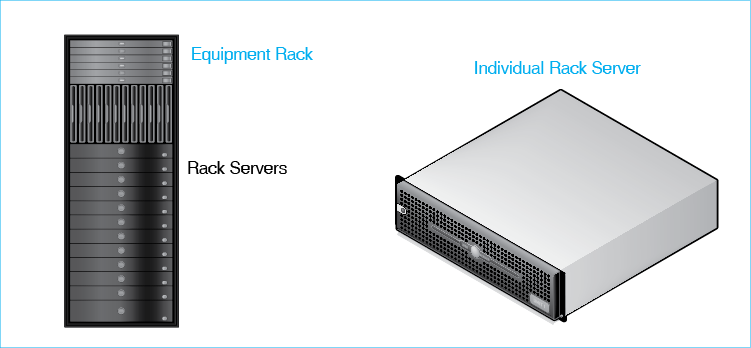
Hosts  The figure shows two devices attached to the Internet. One is a tablet. The other is a larger computer. Both are called hosts. This is an odd but central piece of terminology. We call any device that communicates over a network a host. Hosts also include desktop PCs, laptops, tablets, smartphones, smart glasses, and smart watches. In the future, hosts will include interactive walls, tables, and appliances that will turn your entire home into an immersive interactive environment. In a trend called the Internet of things, even coffee makers, toasters, medical implants, and many other small and large devices around us will be hosts that communicate through networks to work better. In fact, machine-to-machine communication will soon dominate traffic on the Internet. The term host is not an obvious name for computers that attach to networks, but it is the common name for them in networking.

Any device that communicates over a network is a host.

Clients and Servers  Traditionally, hosts have come in two types. Server hosts provide services to other hosts. The hosts that receive these services are called client hosts. Client hosts include things you see every day, such as desktop PCs, laptops, tablets, mobile phones, and many other devices you see each day. You may never have seen a server, and unless you work in an IT department, there is no reason to see one. In practice, most servers resemble pizza boxes. They are 19 inches (48 cm) wide, about two (60 cm) feet deep, and about two to six inches (5 to 15 cm) tall. Multiple servers are stacked in equipment racks. Data centers are filled with these racks. There are also far larger servers, and we will usually show a single large device in illustrations rather than a rack server.

Server hosts provide services to other hosts. The hosts that receive these services are called client hosts.

Figure 1-: Rack Servers and Equipment Rack



Messages  Figure 1-1 shows that communication between application programs is done by exchanging messages. There are many types of messages on the Internet and in other networks. Understanding individual message types (and telling them apart) will be an important challenge for you in this course.

Addresses  Messages require addresses. If you want to send the first author a message, you would need to know his e-mail address, Ray@Panko.com. Hosts also need addresses. On the Internet, these are Internet Protocol addresses or IP addresses. In Figure 1-1, the IP addresses are 1.2.3.4 for the source host and 5.6.7.8 for the destination host.

Test Your Understanding

2. a) What is the term we use in networking for any device that communicates over a network? b) Is your smartphone a host when you use it to surf the ‘Web? c) Are you as a person a host when you use a network?

3. a) How do application programs on different hosts communicate? b) If I call you on your cellular phone, do I have to know your name? (Think about this a second.) c) What do I have to know? d) What kind of addresses do hosts have on the Internet? e) What kind of address is 128.171.17.13?

## Dotted Decimal Notation (DDN)

In Figure 1-1, IP addresses are expressed as four numbers separated by dots (periods), this is called dotted decimal notation (DDN). IP addresses are actually strings of 32-bits (1s and 0s). Computers work directly with these raw bit strings. Human memory and writing, however, need a crutch to deal with long bit strings. Dotted decimal notation is precisely that—a crutch for inferior biological entities like ourselves.

Figure 1-3 shows how to convert a 32-bit IP address into dotted decimal notation.

* First, divide the 32 bits into four 8-bit segments.
* Second, treat each segment as a binary number and convert this binary number into a decimal number. For example, the first segment, 00000001 in binary, is 1 in decimal. How can you do this? Go on the Internet and search for an online binary-to-decimal converter. Just type this need in a search engine. In fact, most search engines let you simply type “convert binary 00000001 to decimal.”
* Third, combine the four decimal field values, separating them by dots. This gives 1.2.3.4.

Figure 1-3: Dotted Decimal Notation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Step | Action | Example | | | |
| 1 | Divide 32 bit IP address bits into four 8-bit segments. | 00000001 | 00000010 | 00000011 | 00000100 |
| 2 | Convert each segment from binary to decimal. | 1 | 2 | 3 | 4 |
| 3 | Assemble the segments in decimal, separated by dots. | 1.2.3.4 | | | |

The 32-bit IP address is not the only type of IP address. A 32-bit IP address is an IP Version 4 (IPv4) address. IPv4 is the dominant IP protocol on the Internet today. However, we are beginning to see significant use of IP Version 6 (IPv6) on the Internet. As we will see in Chapter 8, IPv6 addresses are 128 bits long and are represented for human consumption in a different way.

Test Your Understanding

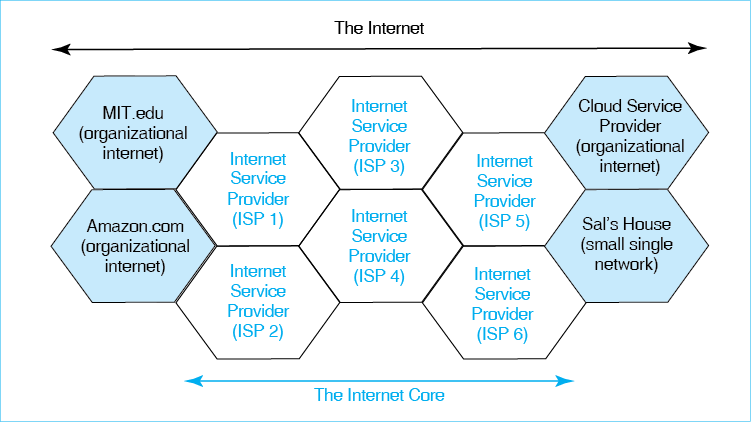
4. a) What do we call the format 128.171.17.13? b) Convert the following binary IP address into dotted decimal notation (spaces are added for easier reading): 10000000 10101011 00010001 00001101. (Check Figure: 10000000 = 128.) c) Convert 5.6.7.138 into a 32-bit IP address. (Check Figure: 5 = 00000101) Show a space between each 8-bit segment. d) What type of IP addresses is 32 bits long? e) What other type of IP address exists, and how long are its addresses?

## The Internet

Having looked at some basic terminology we will use throughout the course, it is time to begin looking inside the Internet “cloud.” Figure 1-4 illustrates that the global Internet is not a single network. Instead, the Internet is a collection of thousands of single networks and smaller internets. All of these single networks and smaller internets link together to form a single transmission system that allows any Internet host reach any other host.[[8]](#footnote-8) Some single networks and smaller internets are owned by user organizations such as Amazon.com or Syracuse University. Even smaller networks are owned by families and even individuals. If you have a network at home, it is a single network on the Internet! Specialized internets link these smaller networks and smaller internets together. We call these linking internets Internet service providers (ISPs). ISPs collectively form the core of the Internet, which is also called the Internet’s backbone.[[9]](#footnote-9) To use the Internet, a customer must connect to an ISP.

The Internet is a collection of single networks and smaller internets. All of these networks and smaller internets link together to form a single transmission system that allows any Internet host reach any other host.

Figure 1-4: The Internet’s Networks and Smaller Internets



Capitalization  When do you capitalize the term internet? In lowercase (except at the beginning of a sentence and in titles), internet refers to any internet—any collection of single networks and smaller internets. We capitalize the term when we refer to the global Internet. For example, the University of Hawaii has an internet on campus. It connects single networks in individual buildings. The University of Hawaii internet is also one of the smaller internets that make up the Internet.

In lowercase (except at the beginning of a sentence and in titles), internet refers to any internet—any collection of single networks and smaller internets. We capitalize the term when we refer to the global Internet.

Ownership, Control, and Funding  Given the importance of the Internet in the world today, you might wonder about its ownership, control, and funding. The actual picture might surprise you.

* Who owns the Internet? The surprising answer is, “Nobody.” The ISPs and other organizations own their pieces of the Internet.
* Who controls the Internet? Again, the answer is nobody. Although the Internet Engineering Task Force (IETF) creates standards, network owners decide which standards to adopt. There is no overall authority to enforce standards or to govern interconnection business practices. Everything is negotiated between the network and internet owners.
* Who pays for the Internet? You do. You need an ISP to use the Internet. You and other users pay ISPs, who work out arrangements with other ISPs to deliver your messages. You probably pay around $30 per month to your ISP. Businesses pay thousands or millions of dollars annually. With rare exceptions, no government money sustains the Internet.

This may strike you as a rather strange situation. However, the worldwide telephone network has always worked this way. Users pay individual local telephone companies, but their calls typically pass through multiple telephone companies to reach to other users. Users pay for the telephone company with their monthly charges. The United Nations creates technical standards and working agreements, and these are usually enforced by individual countries, but only to some extent. In addition, customer payments fund the telephone network.

Test Your Understanding

5. a) Is the Internet a single network? Explain. b) What is the role of ISPs? c) When do you capitalize internet? d) Who controls the Internet? e) Who funds the Internet?

## Netflix Dives into the Amazon

You know from personal experience how individuals use the Internet. The corporate experience is very different. We will illustrate this by looking at how Netflix uses the Internet. Netflix is a commercial streaming video service with tens of millions of customers around the world. Streaming video places a heavy load on network capacity. For a two-hour high-definition movie, Netflix must deliver five million bits every second. This is a total of nine gigabytes for that one movie. On a typical night, Netflix generates roughly a third of the Internet traffic going into U.S. homes.

Requirements  Users expect high video quality, and they will not tolerate delay or unreliability. The Internet was not designed for these requirements. The Internet is a “best effort” delivery system that makes no promises. As you know, it often has insufficient speed and reliability, and it often has too much delay for Netflix users. Netflix had to overcome these limitations.

The Internet is a “best effort” delivery system.

Video streaming also requires vast amounts of server processing capacity. Each movie must be transcoded into many different streaming formats appropriate for different user equipment and network conditions. When a customer requests a movie, streaming servers select the best transcoded format for that particular customer. Transcoding every movie into a hundred or more formats takes enormous server capacity.

In addition, at the heart of Netflix’s business plan is an application that creates personalized viewing suggestions for individual customers. This requires the analysis of massive data about the customer’s viewing habits and the choices of other customers with similar viewing profiles. This also demands massive server capacity.

Outsourcing  In 2008, when Netflix was only delivering movies by mailing DVDs, the company suffered a crippling server outage that halted shipments for several days. That was a wake-up call for Netflix. Management realized that reliability would be critical for the online delivery it would soon introduce. It also realized that while Internet delivery, content delivery, transcoding, recommendation systems, and other applications would be keys to its core business, it had no competitive advantage in managing servers. Rather than developing the expertise it would need for the complex server technologies it use, Netflix decided to outsource server operation to a company that could meet Netflix’s high requirements for capacity, reliability, and agility in responding to sudden demand changes.

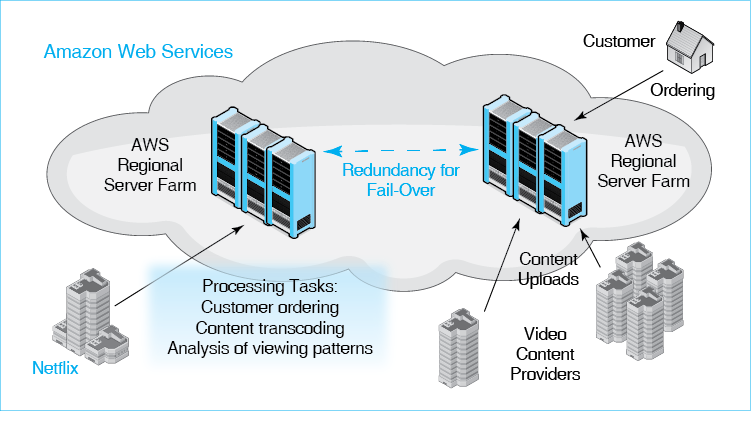
Netflix chose Amazon Web Services (AWS). Amazon’s life depended on its e-commerce servers. The company by necessity became arguably the most advanced manager of servers in the world. Its data centers became so massive that they began to be called server farms. This term was a testament to their size and to the fact that they were usually located in rural areas with low land and utility costs. These server farms today have thousands of rack servers.

Amazon eventually made a business decision to offer managed servers as a service. Today, this is called infrastructure as a service (IaaS). Customers merely connect to servers that Amazon manages. Customers manage their own applications running on these servers. Figure 1-5 shows that this service is generally viewed as a cloud. Just as users do not have to know how the Internet operates, Amazon customers do not have to know how Amazon is managing its servers.

In infrastructure as a service (IaaS) customers do not manage servers, but they manage the applications that run on these servers.

Amazon sells this cloud IaaS service under the name Amazon Web Services (AWS). AWS had the capacity that Netflix needed for customer ordering, transcoding, the analysis of viewing patterns, and other purposes. In addition, Amazon had multiple regional server farms with high fail-over capabilities. Even the loss of an entire server farm would not disrupt service for more than minutes. This brought the reliability that Netflix customers demanded. Netflix customers today log into an AWS server to order videos and to take care of other business transactions with Netflix. There are many login servers, and AWS automatically routes the user to one of them.

Figure 1-: Netflix and Amazon Web Services (AWS)

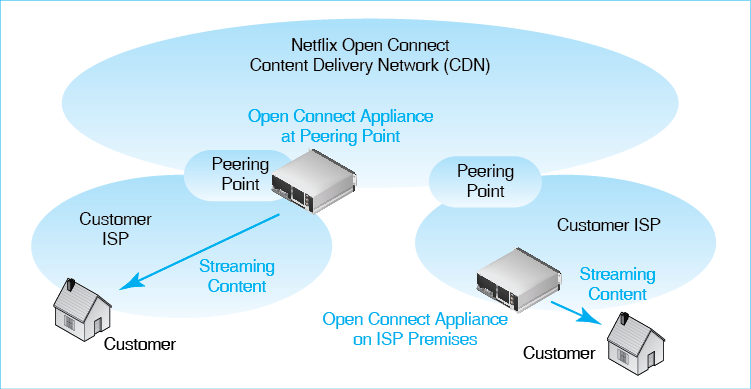


This does not mean that Netflix ignores the operation of AWS server farms. Netflix constantly reviews the quality of AWS service. In addition, it constantly tests reliability by disabling some of its AWS servers briefly to see what the effects are. It does this with a family of programs called the Simian Army. This name stems from the fact that two of the most important disabling programs are Chaos Monkey and Chaos Gorilla. Constant testing leads to the constant refinement of how Netflix uses AWS. Amazon has outages occasionally, but Netflix customers rarely notice them.

Content Delivery  Netflix uses AWS to store more than one petabyte of video content in multiple locations. Of course, having massive storage is worthless unless the transcoded video content can be delivered effectively to users. In particular, content delivery must have very little delay, which in networking is called latency. When people select a video, they want it to start immediately. When they hit the back button, they want the video to start instantly after it jumps back to an earlier scene.

Figure 1-6 shows how Netflix delivers video content to individual customers. It uses an approach called a content delivery network (CDN). Many commercial websites such as news sites have long used this approach. Content delivery networks focus on reducing latency to users by caching (keeping a copy of) content as near as possible to the user. Netflix, which calls its CDN Open Connect, has taken this concept and scaled it enormously.

Figure 1-: The Netflix Open Connect Content Delivery Network (CDN)



For Open Connect, Netflix created its own webserver appliances. They are called appliances because they are simply installed and then operate with little or no local management. (Think of a toaster.) Each is a relatively small box that can fit into a standard equipment rack. The current Open Connect appliance is seven inches (18 cm) high and two feet (61 cm) deep. Although small in size, it holds about 100 terabytes of data on 36 hard disk drives. Its processor is fast enough to stream movies from its built-in webserver software to between 10,000 and 20,000 customers simultaneously. Netflix upgrades its CDN servers about once a year with newer hardware to increase their capabilities.

Figure 1-6 shows that Open Connect is a network on the Internet. It can peer with (connect to) the ISP of a customer. The CDN boxes can be placed at the peering point, so that traffic only travels the relatively small span of the customer’s ISP network. This slashes latency in starting streams of content. Beyond this, many ISPs that serve customers place the appliance on their own premises, at a location near the final customer. Netflix customers like this because it slashes latency even more. ISPs like it because traffic only flows across a small portion of their network.

With only 100 terabytes of data storage, individual appliances cannot store all of Netflix’s one petabyte of content. Consequently, Netflix uses sophisticated analysis to identify the 100 terabytes of content most likely to be demanded by customers served by each specific appliance. It installs this content on the individual CDN server appliances. Customer interests change rapidly, so the content mix for each appliance has to be refreshed endlessly. During quiet periods in demand each day, Netflix deletes content declining in popularity and sends content growing in demand.

Test Your Understanding

6. a) List Netflix’s content delivery requirements. b) How does Netflix use AWS? c) How does the Open Connect CDN reduce streaming delays to customers?

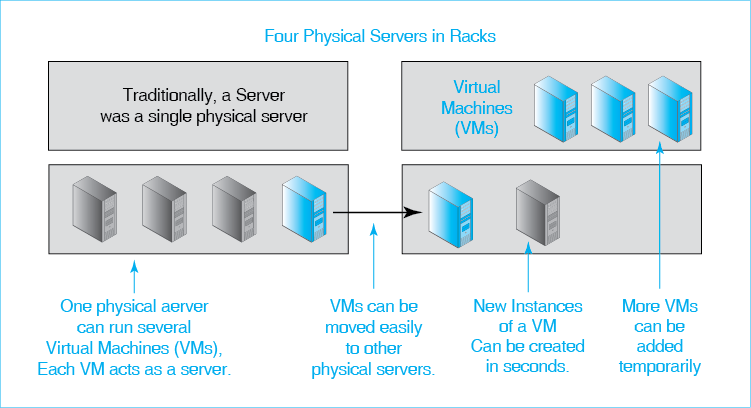
## Virtualization and Agility

Virtual machines are software processes running on the physical server. However, they act like real servers to the outside world.

Figure 1-7 shows that AWS uses virtualization to turn each physical server into several virtual machines (VMs). VMs are not real machines. They are only software processes running on the physical server. However, they act like real servers to the outside world. Each has its own IP address as well as its own data. A VM is even managed like a real server.

Virtual machines are software processes running on the physical server. However, they act like real servers to the outside world.

Figure 1-7: Server Virtualization through Virtual Machines



Virtual machines give a data center agility, which is the ability to make changes quickly—even very large changes. For example, Amazon can move a VM immediately from one physical server to another simply by transferring or moving it as a single large file. It can even move VMs to servers in different regions of the world. New VM instances (individual virtual machines) can be added in seconds. In fact, a company can spawn many additional copies of the same virtual machine if demand surges. For example, Netflix spins up many additional VMs for content delivery during peak viewing time. Later, it spins them down again to save money. Physical servers offer nothing like this degree of agility. To make virtualization even more attractive to customers like Netflix, AWS provides a simple self-service application for customers to use to add new VMs and do many other things themselves. This saves AWS money, and it gives customers direct control of their resources.

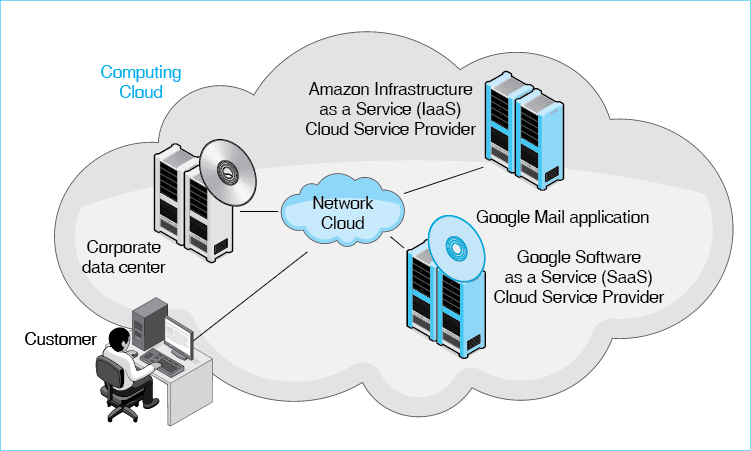
Test Your Understanding

7. a) Distinguish between physical servers and virtual machines. b) What can be done with virtual machines that would be difficult to do with physical servers? c) What are the two advantages of self-service applications that allow Netflix and other customers to add new instances of VMs, spin down VMs when they are not needed, and do many other server management tasks?

## Infrastructure as a Service (IaaS) and Software as a Service (SaaS)

Amazon is a cloud service provider (CSP). Figure 1-8 illustrates this concept. We saw earlier that the Internet and other networks are often depicted as clouds. We also saw that CSPs also operate their services opaquely, forming a second layer of cloud.

Figure 1-8: Cloud Service Providers, IaaS, and SaaS



Infrastructure as a Service  The AWS service that Netflix uses is referred to, generically, as Infrastructure as a Service (IaaS). This ungainly name refers to the fact that AWS provides the computing infrastructure, which consists of server operation, database management systems, and so forth.

However, customers like Netflix must create and manages its own applications for user ordering, transcoding, personalized viewing suggestions, and other matters. By outsourcing server infrastructure operation to AWS, Netflix can focus its efforts more fully in developing and extending its applications, which are the critical parts of its business.

In Infrastructure as a Service (IaaS), the CSP manages servers but customers manage applications.

Software as a Service  Amazon is not the only CSP that Netflix uses. Another is Google. Netflix uses Google Mail for its internal communication. In contrast to just offering IaaS, Google offers application software as well. This is called Software as a Service (SaaS). Here, “software” refers to application software. SaaS actually has been popular for some time. For example, many companies use salesforce.com application software for salesforce management and customer relationship management.

In Software as a Service (IaaS), the CSP manages both servers and applications.

Product versus Service  As a Service refers to pricing. Traditionally, a company bought servers as products, much like you might purchase a new television. However, cloud services are sold like electrical service. You buy nothing. Instead, you pay for the amount of service capacity you use. This eliminates the capital expense (CapEx) of purchasing servers. IaaS makes servers an operating expense (OpEx), which can be managed so that money is spent only when it must be. SaaS, in turn, changes application programs from products to services.

The term “as a service” means that the customer pays on an ongoing basis instead of making an up-front purchase.

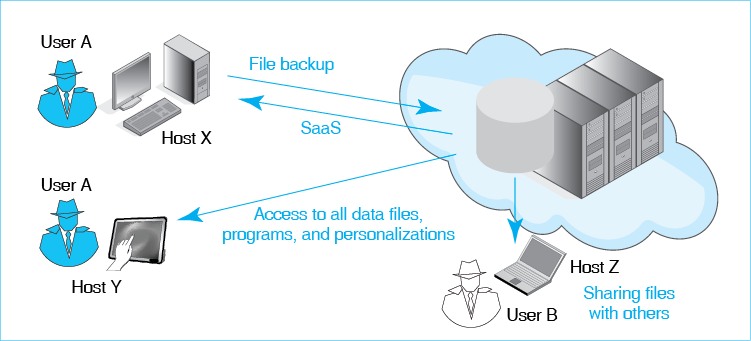
Test Your Understanding

8. a) What is a CSP? (Do not just spell out the acronym). b) Distinguish between IaaS CSPs and SaaS CSPs. c) Is AWS an IaaS or a SaaS for Netflix? d) Is Google an IaaS or a SaaS for Netflix? e) Who owns and manages the servers in IaaSs and SaaSs? f) Who owns and manages the applications in IaaSs and SaaSs? g) In AWS, what does Netflix manage and not manage?

## Cloud Service for Clients

Figure 1-9 shows that individual users also use cloud services. In the figure, User A connects to a cloud service provider for word processing, spreadsheets, and other office productivity applications. These are provided as services, for a monthly fee. In some of these services, software is stored in the cloud until it is needed; it is then downloaded to the user’s current device. In other services of this type, the software is downloaded to reside on each of the user’s devices permanently.

Figure 1-: Cloud Client Services



User A is on Host X, which is a desktop PC. Although User A can store files locally, users also have the option of storing data files in the cloud. The cloud service provider constantly backs up user files, so data losses are rare. (In contrast, users rarely back up files they store on their own devices.)

If User A later uses a different device, say Host Y, the user has full access to the services of the office productivity provider. The user can access his or her data files from any other PC, laptop, tablet, or mobile phone. The user also has full access to his or her office software on any of these devices. Furthermore, any customizations the user has configured on the office productivity software are effective on each device. Screen size differences and other device characteristics will modify the experience somewhat, but as far as possible, it will seem to User A that he or she is still working on Host X.

The service also allows the user to make his or her files available to others selectively. In the figure, User B has been given permission to download a particular file. Services for personal use have a limited number of access rules to apply to control sharing, but these are usually enough for personal situations. In office performance services created for large enterprises, ownership and access can be implemented with complex rules tailored to each situation and based on the organizational role of the person desiring access.

Test Your Understanding

9. a) List the benefits of using a cloud service provider for personal office services. b) How do cloud-based office software services go beyond serving individuals?

## Rain Clouds: Security

A central issue for every cloud customer is security. In cloud computing, companies are putting critical corporate data on computers owned by other organizations. In the case of Netflix, Amazon is actually a competitor in the streaming media market. If cloud service providers fail to protect data from hackers, the potential consequences are enormous. Today, cloud customers wonder if government agencies are demanding access to their data under gag orders that prevent cloud providers from letting them know this is happening. Edward Snowden’s exfiltration of information from NSA servers in 2013 illustrated how even single employees can get access to masses of critical sensitive data.

Figure 1-: Cloud Concerns (Study Figure)

Security

Protection from hackers

The possibility that the government is reading files in the cloud

Stress on the Network

Cloud computing drives network demand

This requires large increases in capacity

Reliability must increase

To deal with security, companies must do due diligence, looking in depth at how cloud service providers handle security. However, there is no way to understand everything about a cloud service provider’s security. For the time being, many organizations are crossing their fingers, whistling in the dark, knocking on wood, and yielding to the attraction of cloud computing’s low cost and agility.

Test Your Understanding

10. What concerns do customers have about cloud security?

## Networks and the Cloud

Networks today must work extremely well, almost perfectly. They must do this while growing at unbelievable rates. They must also do all this using standards that are older than most of today’s network engineers. We have been talking about the Internet. However, as we saw earlier, the Internet is not a single network. It is a jumble of single networks and smaller internets ranging in size from a couple of devices in a dorm room to corporate internets for globe-spanning corporations.

The demands of cloud computing are creating enormous stresses on networks. Cloud service providers themselves generate massive and fast-changing network transmission loads. Customers of cloud services also find themselves with massive increases in Internet and local network traffic. In addition to growing rapidly, networks are also facing increasing demands for reliability because a company that loses contact with its cloud service providers for even brief periods of time will suffer heavy losses.

Test Your Understanding

11. Describe the relationship between cloud computing and networking.

# Speed

The first question people ask about a new baby is, “Is it is a boy or a girl?” For a network, the question is, “How fast is it?” Years ago, when e-mail was the major networked application, speed was not an issue. Today, we have increasingly speed-hungry applications that can only be served by ever-faster networks.

When we talk about buying milk, we quantify the amount we wish to buy in gallons or liters. In networking, we measure speed as the number of bits (1s and 0s) transmitted per second. We use the abbreviation bps for bits per second. Note that network speed is usually not measured in bytes per second. Today, we routinely encounter speeds of millions or billions of bits per second. Consequently, as Figure 1-11 shows, we turn to metric notation. In increasing factors of a thousand, we have kilobits per second (kbps), megabits per second (Mbps), gigabits per second (Gbps) and even terabits per second (Tbps). If you are a little rusty on metric notation, you should read the box “Writing Speeds in Metric Notation.”

Figure 1-: Metric Speed Designations

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Prefix | Abbreviation | U.S. | Meaning | Zeroes | Example |
| kilo | l | Thousand | 1,000 | 3 | 33 kbps is 33,000 bps 43,700 bps is 43.7 kbps |
| mega | M | Million | 1,000,000 | 6 | 3.4 Mbps is 3,400,000 bps or 3,400 kbps 523,750,000 bps is 523.75 Mbps |
| giga | G | U.S. Billion | 1,000,000,000 | 9 | 62 Gbps is 62,000,000,000 bps or 62,000 Mbps or 62,000,000 kbps |
| tera | T | U.S. Trillion | 1,000,000,000,000 | 12 | 1.5 Tbps is 1,500,000,000,000 bps |

How much speed do applications need? Figure 1-12 looks at things from the individual user’s point of view, showing how long it will take to download various types of information at various transmission speeds. Note that e-mail is instantaneous at all but the lowest historical speeds. Streaming video requires a very fast connection, and for disk backup, even gigabit speed may not be enough.

Figure 1-: Application Download Times

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Application / Speed | 1 Mbps | 10 Mbps | 100 Mbs | 1 Gbps |
| E-mail Message (250 words) | — | — | — | — |
| Photograph (2 MB) | 16 sec | 2 sec | — | — |
| 1-Hour HDTV Program (7 Mbps) | 7 hr | 42 min | 4 min | 25 sec |
| Backup 500 GB hard drive | 1.9 mo | 5.8 da | 14 hr | 2 hr |

Corporate networks, of course, must carry the combined transmissions of all users and all machine-to-machine background processes. This creates an enormous aggregate need for speed.

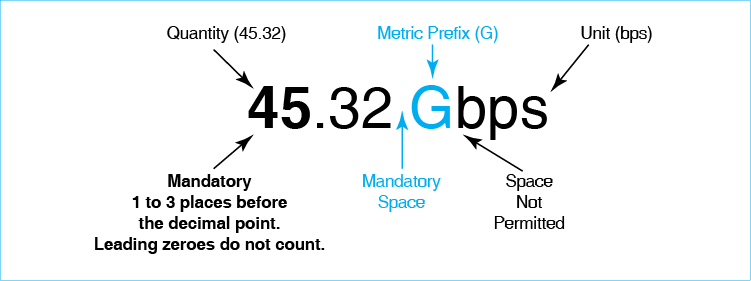
Test Your Understanding

12. a) Is network speed usually measured in bytes per second (Bps) or bits per second (bps)? b) How many bits per second is 56 kbps without a metric prefix? (Check Figure: 56,000 bps) c) How many bits per second is 376.2 Mbps without a metric prefix? d) Express 47,300,000 bps with a metric prefix. (Check Figure: 47.3 Mbps.) e) Express 87,011 bps with a metric prefix. f) Why do you need to know what application you are using to know what connection speed you need? g) Distinguish between speed to individuals and corporate network speeds.

# Box: More on the Metric System and Speeds

Americans tend to be relatively unfamiliar the metric system (officially the International System of Units, SI), which is used heavily in networking. This section looks into metric notation in a bit more detail. It also notes that there are rules for writing metric numbers properly; these rules go beyond metric basics. Figure 1-13 presents basic terminology and writing conventions

Figure 1-: Writing Numbers in the Metric System



## Quantity, Units, and a Mandatory Space

Quantity and Units  Figure 1-13 shows that there are two core parts to a number written in metric. The first is the quantity, which is the number of meters, second, watts, or things being measured. In the figure, the quantity is 14.6. The unit, in turn, is what is being measured. In the figure, the unit is ms, which is milliseconds.

There is a mandatory space between the quantity and the unit. Yes, this is picky, but it is important for writing metric numbers properly. Think of the space as a multiplication symbol.

Prefix and SI Unit  The quantity, mandatory space, and unit are the three core parts of a metric number. However, the unit typically has two parts. The second is the SI unit, such as second, meter, or watt. Before the SI unit comes a prefix, which is a multiplier. As shown earlier, prefixes beyond a hundred are given in multiples of 1,000. The unit ms means one millisecond. Here, m stands for milli, which is 1/1000. Note that while there is a mandatory space between the quantity and the unit, it is mandatory *not* to have a space between the prefix and the SI unit.

There is a mandatory space between the quantity and the unit.

It is mandatory not to have a space between the prefix and the SI unit.

Test Your Understanding

13. a) What are the three parts of a metric number? b) What are the two parts of the unit? c) Where is a space required in metric notation? d) Where is it mandatory not to have a space? e) Write 37kbps properly. f) Write 89k bps properly.

## Capitalization

SI unit capitalization rules follow normal conventions for capitalization in each country. We will look at rules in the United States, which are the rules we will use in this book.

* When SI unit names are written out, they are written in lowercase. We have seconds, meters, hertz (a unit of frequency), and kelvins (a unit of temperature).
* SI abbreviations, in contrast, are written lowercase unless the unit is named after a person, in which case it is capitalized. We have s for second, m for meters, Hz for Hertz, and K for Kelvins.

Test Your Understanding

14. a) When are unit names capitalized? b) When are unit abbreviations capitalized?

## Working with Quantities and Metric Prefixes

Removing Metric Prefixes  Sometimes, you need to change the way a number is expressed. For instance, suppose that you want to express 33 kbps without a metric prefix. The prefix k stands for a thousand, so 33 kbps is 33 times 1,000 bps. This is 33,000 bps. In the second row of the figure, 3.4 Mbps is 3.4 times 1,000,000 bps—3,400,000 bps.

Adding Metric Prefixes  What if you need to go in the other direction—adding a metric prefix? In the first row, we have 43,700 bps. This is 43.7 kbps. How did we get this? We divided the original quantity by 1,000, changing it from 43,700 to 43.7. To compensate, we added the prefix k, which multiplies the SI unit by 1,000. So we have 43,700 / 1000 \* 1000 \* bps. If you divide the quantity by a thousand and also multiply the metric prefix by 1000, you leave the value the same.

For 3,400,000 bps, we divide the quantity by a million and add the metric prefix mega (M). So 3,400,000 bps gives 3.4 Mbps. Note that you always multiply or divide by a factor of a thousand (1,000, 1,000,000, 1,000,000,000, etc.).

How many times should you multiply or divide the quantity by a thousand? The answer is that you do it until you get one to three digits before the decimal point. So 7.89 Mbps (1 place) is good. So are 23.426 kbps (2 places) and 178 Gbps (3 places). However, 4,300 kbps is not good (4 places). Neither is 0.45 bps (no places before the decimal point because leading zeroes do not count).

What if you add a metric prefix to a quantity less than one? Then instead of dividing by a thousand, and increasing the metric prefix, you multiply by a thousand and decrease the metric prefix accordingly. If you have 0.045 Gbps, you multiply 0.045 by a thousand. This gives you 45. To compensate, you divide the metric prefix by 1,000, giving M. So you have 45 Mbps instead of 0.045 Gbps.

Test Your Understanding

15. a) Write 50 bps properly. b) Write 56.780,000 m properly. c) Write 0.25 s properly. d) Write 45.6355 kbps properly. e) Write 37,400 m properly. f) Write 0.032 Mbps properly.

# Messages

In networking, entities communicate by exchanging messages. In this section, we will begin looking at network messages by focusing on the application messages that application programs exchange to coordinate how they work. The key point is that long application messages are fragmented into smaller segments, and these segments are placed inside other messages called frames or packets. The network deals only with these frames or packets.

## Application Messages

The World Wide Web uses the Hypertext Transfer Protocol (HTTP) standard to standardize message exchange between browsers and webserver programs. Figure 1-14 shows that an HTTP request message asks for a file. The subsequent HTTP response message delivers the file or an error message. This exchange is called an HTTP request/response cycle. Browsers and webserver programs are application programs, so the messages they exchange are application messages.

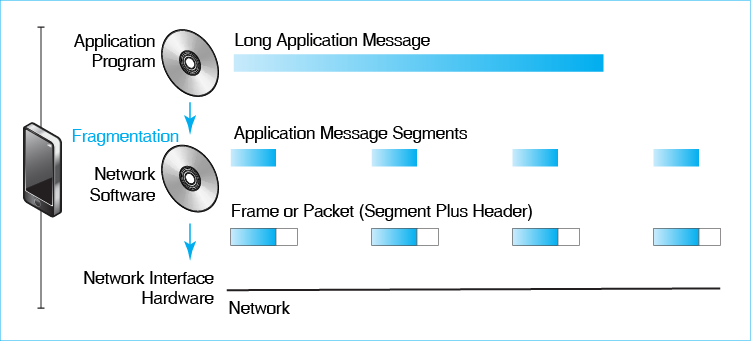
Figure 1-14: Application Request-Response Cycle in HTTP



## Message Fragmentation, Frames, and Packets

As Figure 1-15 shows, application message may be very large. As we saw earlier, a movie transfer requires the transmission of about nine gigabytes of data. Sending a single long application message through a network would be like sending an 18-wheeler trucks through a narrow English country village.

Figure 1-15: Fragmentation, Frames, and Packets



Fragmentation  In addition to application software, hosts also have network software, which handles network transmission for the operating system. As Figure 1-15 illustrates, network software does three things with application messages.

* First, if the application message is large, the network software fragments the long application message into many smaller message segments.
* Second, the network software places each segment in an electronic envelope. This “envelope” is merely a set of added bits in front of the segment (the header) and perhaps another set of added bits after the segment (the trailer). These bits contain delivery information for the segment in the envelope. Importantly, the envelope contains the address of the receiving host. The network uses this information to deliver the envelope to the destination host.
* Third, the network interface hardware on the source host transmits each segment plus envelope over the network.

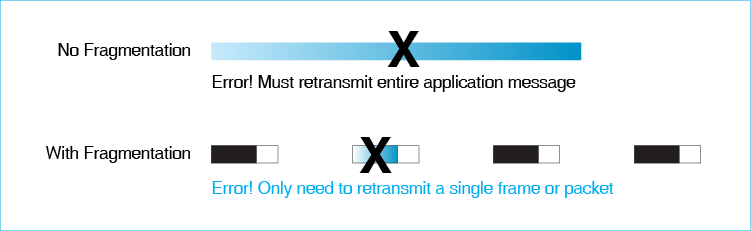
An envelope and its contents are called either a frame or a packet. Later in this chapter, we will see when to use the term frame and when to use the term packet.

Application message segments are carried inside frames or packets that have delivery information.

On the destination host, the network software receives all the frames or packets delivering the application message. It removes the segments, puts them in order, and reassembles the original application message. The network software passes this application message up to the application program. Note that neither of the two application programs knows that fragmentation has been done. The source application simply passes the whole application message down to the network software. The destination application receives the whole message from its network software. We say that fragmentation and reassembly are transparent to the application programs.

Error Correction  Why do fragmentation? One reason is error correction. Suppose there is a transmission error somewhere within the network and a few bits are lost, damaging the message. The incorrect message must be retransmitted. Figure 1-16 shows retransmission with and without fragmentation.

Figure 1-16: Fragmentation and Error Correction

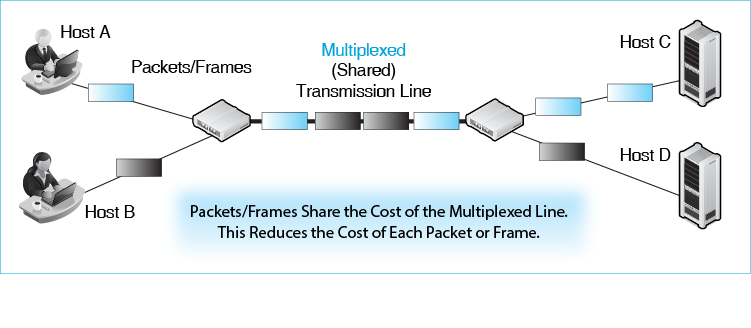


Without fragmentation, the sender must retransmit the entire message. This adds quite a bit to the network’s traffic load. In addition, if a message is long, a transmission error has a high probability of occurring each time it is sent. The message may have to be resent multiple times before it arrives correctly. Transferring a moderate-size file from home to school over a cranky telephone line and without fragmentation once took the first author three hours.[[10]](#footnote-10)

With fragmentation, the sender only needs to retransmit the single damaged frame or packet. An error will therefore increase the transfer time only slightly. The bottom line is that fragmentation reduces delays due to errors, speeding application message delivery.

Multiplexing  Figure 1-17 shows a second reason for fragmentation. Consider this analogy. Would it be nice to drive into school or work in your own personal highway lane? There would be no traffic congestion at all. That would certainly be nice, but having thousands of traffic lanes on a freeway would entail astronomical construction costs. In the real world, we have to drive on shared highways that carry many cars in each lane. The individual cars and trucks “share” the cost of the highway lane.

Figure 1-17: Fragmentation and Multiplexing



In networking, frames or packets share the cost of multiplexed (shared) transmission lines. In the figure, two client PCs are sending messages to different servers. Their frames/packets share the capacity of the multiplexed line. Multiplexing reduces transmission costs.

It is technically possible to multiplex large application messages. However, filling a line efficiently with large messages is impossible. To see why, consider another analogy. Hourglasses work far better with sand than with rocks. Small frames and packets can fill a transmission line very efficiently.

Test Your Understanding

16. a) What is fragmentation? b) What are frames or packets? c) How does fragmentation improve transmission speed through its error handling abilities? d) What is multiplexing? e) How does fragmentation reduce transmission cost through multiplexing?

# Single Networks

An internet is a collection of single networks and smaller internets. We now look at the term “single network.” Single networks are networks that have three defining characteristics. If you do not understand these characteristics immediately, things should be clear after some examples.

* A single network uses a single technology for transmission. All devices must comply with that single technology’s standards.
* Each host has a unique address on the network, like a telephone number.
* Messages in single networks are called frames, not packets.

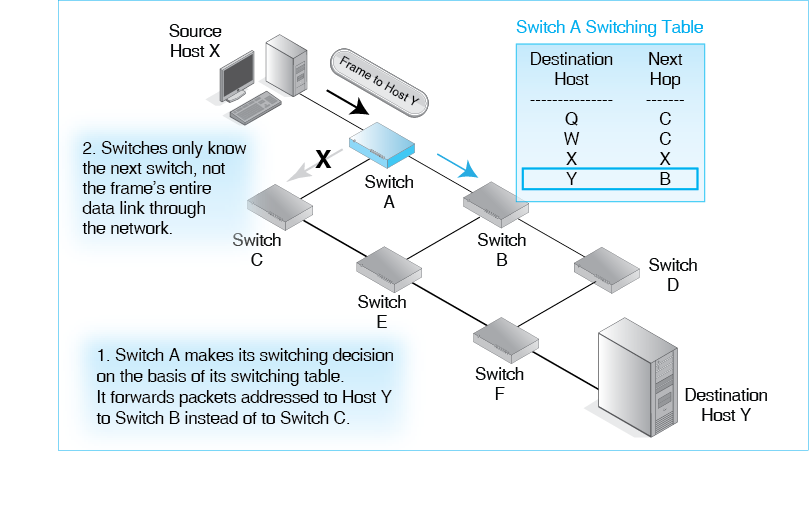
Test Your Understanding

17. What are the three defining characteristics of single networks?

## Switched Single Networks

Most single networks are switched networks, in which the frame must travel through a number of switches to reach the destination host. Figure 1-18 illustrates a switched network.

Figure 1-: Switched Network



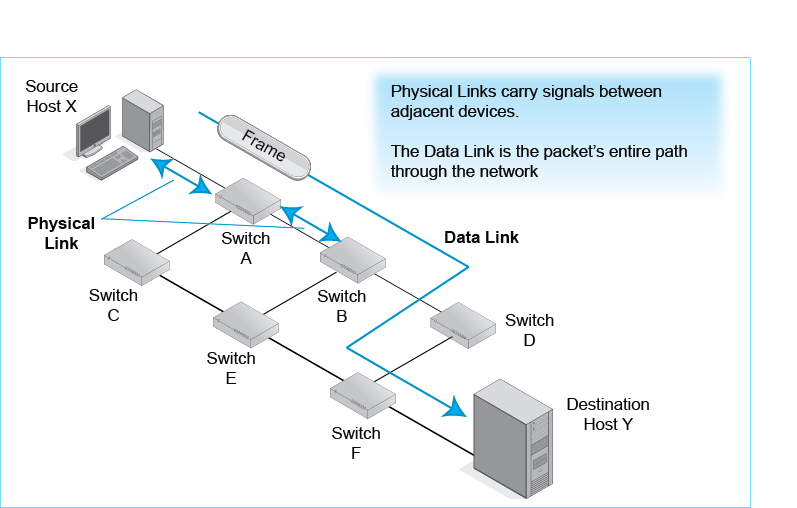
Switching Decisions. In this figure, the source host has transmitted a frame to the switch it connects to, Switch A. Upon receiving the frame, Switch A must make a switching decision. It must decide whether to send the frame on to Switch B or Switch C. To do this, it uses its switching table. The table requires the switch to look at the destination address in the frame. In the figure, the destination host is Host Y, and the switch is Switch B. Based on this information, Switch A passes the frame to Switch B.

Notice that Switch A does not know the frame’s entire path through the network. It only has local knowledge of what to do next. Switch B will also have only local knowledge. It will know whether to send the frame to Switch D or Switch E. To make this decision, it will use its own switching table, which will tell it what switch to send the frame to next if the destination host is Host Y.

Data Links. Figure 1-19 introduces two pieces of terminology that you will use throughout this course. The first is data link. The data link is the path a frame takes through a switched single network from the source host to the destination host. It is the entire path the frame travels. Consequently, there is only a single data link for each frame.

The data link is the path a frame takes through a switched single network from the source host to the destination host.

Figure 1-: Physical Links and Data Links in a Switched Network



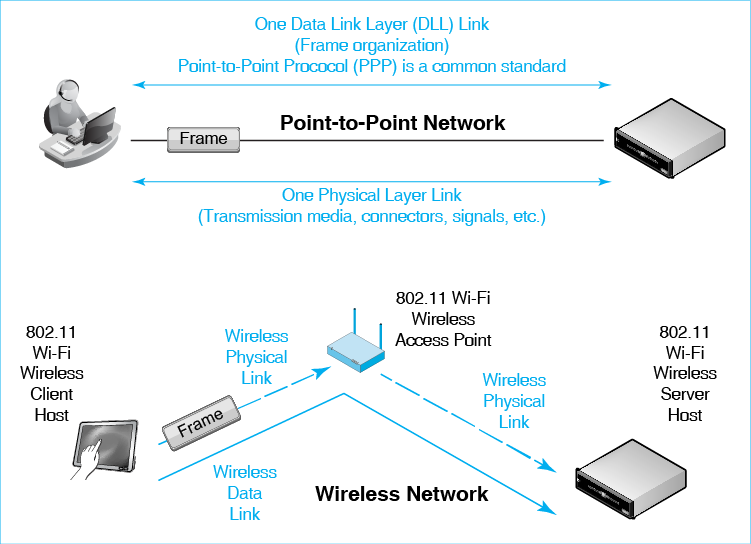
Physical Links. Along its data link, the frame must jump from one switch to the next. We need a name for a single hop between adjacent devices in the overall data link. Networking professionals call it the physical link. In Figure 1-19, two physical links are marked. There will be three more along the data link shown in the figure.

A physical link is a connection between adjacent devices in the data link.

## Nonswitched Single Networks

Not all single networks are switched. Figure 1-20 illustrates two nonswitched single networks.

Figure 1-: Nonswitched Single Network Technologies



Point-to-Point Network. The network on top is a point-to-point network. The source and destination hosts are connected directly. There is one data link for the frame. (There is always one data link). In this case, there is only a single physical link as well. A point-to-point network may not seem like a very useful network, but we will see later in this book that point-to-point networks play important roles in the Internet.

Wi-Fi Wireless Network. The network at the bottom is a wireless network, in fact, a Wi-Fi network. Here, there are two hosts in a house. Using the home’s wireless access point, they can share files, use each other’s printer, and do other things. When the user on the left transmits a frame, it goes to the access point. The access point then passes the frame on to the destination host, Host Y. There is one data link, as there always is in a single network. There are two physical links—one from the source host to the access point, the other from the access point to the destination host.

Test Your Understanding

18. a) When a frame arrives at a switch, what must the switch decide? b) How will it decide this? c) Does a switch know a frame’s entire path through the network? d) What do we call the path the frame takes from the source host to the destination host in a single network? e) What do we call the path between two adjacent devices in a single network? f) How many frames are transmitted when the source wireless host sends a frame to the destination host via an access point in Wi-Fi? g) How many physical links?

## Single-Network Host Addresses

Earlier, we looked at Ethernet addresses. Single network devices also have addresses. To send a frame to another host on the same single network, the source host must know the destination host’s address on that single network. The source host places that destination address in the frame. The destination host accepts the frame because it is addressed to it. If there are switches along the way, they make forwarding decisions based on the destination address in the frame.

We saw earlier that the Internet uses IP addresses, which are 32 or 128 bits long. In single networks, there are many types of address formats. The most widely used type of single-network address is the EUI-48 address. As its name suggests, this address is 48 bits long. The EUI part stands for extended unique identifier. All addresses are unique identifiers, so this part makes sense. The extended part? Uh, just ignore it.

If you have some networking background, you might be asking, “Hey, isn’t that a Media Access Control (MAC) address?” Actually, it is the Institute of Electrical and Electronics Engineers’ (IEEE’s) new name for MAC addresses. You need to know that they mean the same thing because the term “MAC address” is still widely used. However, the IEEE has deprecated the MAC name in favor of EUI-48.

EUI-48 address is the new name for Media Access Control (MAC) address.

Figure 1-: EUI-48 Addresses (Study Figure)

Different single-network standards use different types of addresses

A common type of single-network address is the EUI-48 address

EUI stands for extended unique identifier

It is 48 bits long

Used in Ethernet 802.3, Wi-Fi 802.11, and Bluetooth

However, not all single-network addresses are EUI-48 addresses

Formerly called a media access control (MAC) address

Still commonly referred to as MAC addresses

Written as six pairs of number or letter symbols separated by dashes

Example a1-b2-xx-92-ff-00

An EUI-48 address is written as six pairs of number or letter symbols separated by dashes. For example, an EUI-48 address might be a1-b2-cc-92-ff-00. In Chapter 5, we will look at EUI-48 addresses in detail.

EUI-48 addresses are very common because they are used in 802.3 Ethernet wired networks, 802.11 Wi-Fi wireless networks, and Bluetooth. However, they are not the only kind of single-network address. Do not confuse their common use with universality.

EUI-48 (MAC) addresses are very common because they are used in 802.3 Ethernet wired networks, 802.11 wireless networks, and Bluetooth. However, they are not the only kind of single-network address. Do not confuse their common use with universality.

Test Your Understanding

19. a) What is the most widely used single-network address? b) What else is it called? c) Why is the following NOT an EUI-48 address: a9-00-ff-93-bd? d) Are all single-network addresses EUI-48 addresses?

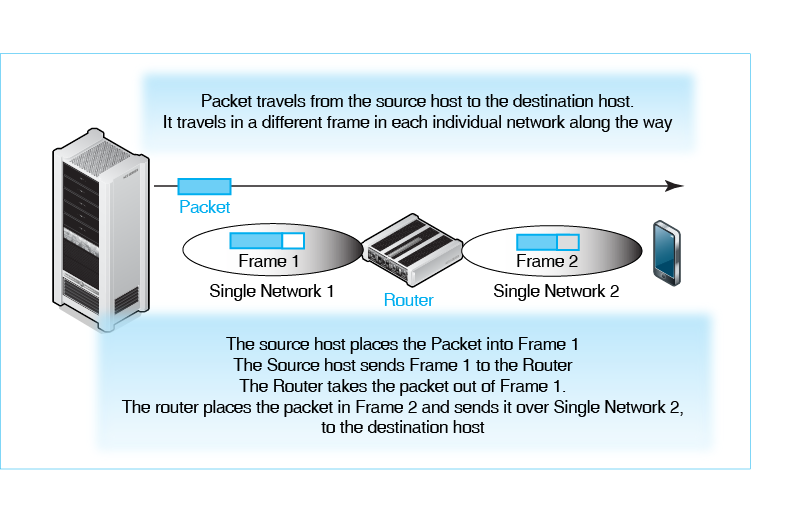
# Internets

## Routes and Packets

Route  Internets contain many single networks. Figure 1-22 shows that internets introduce a new type of path. We have seen that the path a frame travels through a single network is called a data link. In turn, the path that a message must travel through an internet, across many single networks, is called a route.

The path that a message must travel through an internet, across many single networks, is called a route.

Figure 1-: An Internet



Packets  In addition, frames only travel within a single network. We need a new type of message that will travel all the way from the source and destination host across an internet. This new type of message is called a packet.

Why just not let frames travel all the way between the source and destination hosts instead of creating these new packet things? That would certainly simplify things. However, there are many different standards for single networks. These standards govern how frames are forwarded and even the syntax of frames (which we will see in the next chapter). If we sent a frame from one network to the next, there is no guarantee that the next network would even understand how the frame is organized, much less be able to forward it well.[[11]](#footnote-11)

IP Addresses  To go with our new message type, we need a new type of address that is assigned in a way that is globally unique. We have already seen this address. It is the IP address. IP version 4 (IPv4) addresses are the most common. They are 32 bits long and are written for human consumption in dotted decimal notation. IPv6 addresses are beginning to grow in use. They are 128 bits long and are written for humans in a different way that we will see in chapter 8.

Routers  We also need a new type of device to connect different single networks together and pass packets from the source host to the destination host. These devices, which act like switches for packets, are called routers. Initially, they were called gateways, and they are still referred to as gateways sometimes.

Test Your Understanding

20. a) How are frames and packets related in terms of capsulation (when one message is placed inside another message)? b) Create a table with three columns. The first is Description, the second is Single Network, and the third is Internet. Rows should have the descriptions Message Type, Address Type, and Forwarding Device. Fill in the table.

## Frames and Packets

Frames only travel through single networks, while packets travel across internets. You might expect that the two are related, and they are. As we just saw, Figure 1-22 shows a very simple internet with two single networks.

Host A creates a packet for Host B. This packet must travel across both single networks to reach the destination host. In turn, there are two frames, one in each single network. The figure shows that when a packet is transmitted, it travels inside a frame in each network along the way. It is carried by Frame 1 in Single Network 1 and by Frame 2 in Single Network 2.

* Host A first creates the packet. It places the packet inside Frame 1 and transmits the frame across Network 1 to the router.
* The router receives Frame 1 and removes the packet from it. It then decides what interface to send the packet back out. It selects Network 2. It then places the packet in a frame suitable for Network 2. We will call this frame Frame 2. The router sends this frame to the destination host.
* The destination host takes the packet out of the frame. It reads the destination IP address of the packet to ensure that it was intended for the destination host. It is, and the transmission process is ended.

To use terminology we saw earlier, the packet travels over a single route from the source host to the destination host. Along the way, it travels in two frames through two single networks and so two data links.

Test Your Understanding

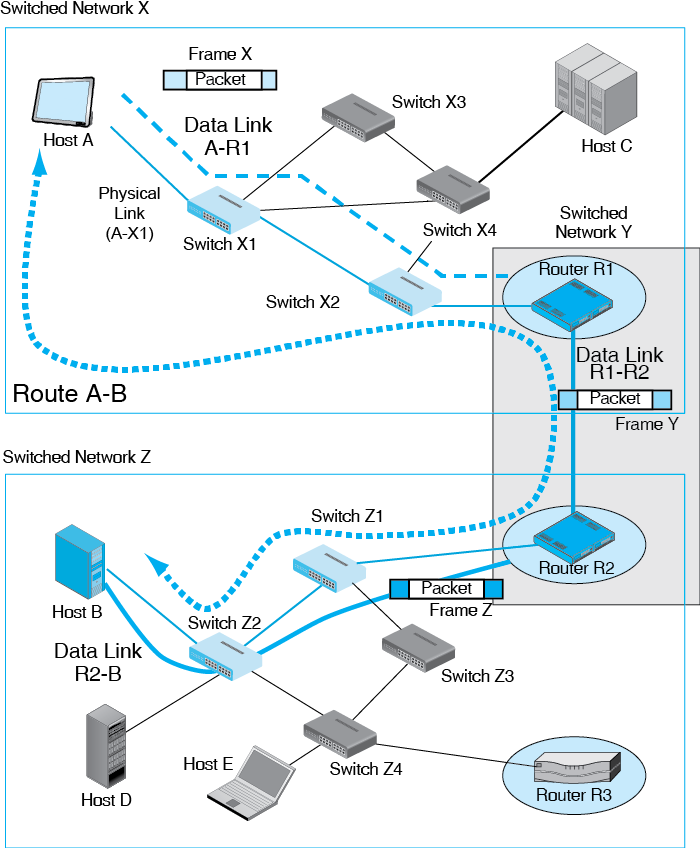
21. a) How are frames and packets related in terms of capsulation (when one message is placed inside another message)? b) What does a router do when a frame arrives?

## An Integrative Exercise

Figure 1-23 shows a more complex internet. It shows Host A on Switched Network X, which sends a packet to Host B on Switched Network Z. Trace the route of the packet between the two hosts. You will see that:

* There are one packet, one route, and two routers.
* There are three frames and three data links.
* There are seven physical links.

Figure 1-: A More Complex Internet



Test Your Understanding

22. a) When Host D in Figure 1-23 transmits a packet to Host C, how many physical links, data links, and routes will there be? b) How many packets will there be? c) How many frames? (Check answers are 7, 3, 1, 1, and 3). c) When Host A in Figure 1-23 transmits a packet to Host C, how many physical links, data links, and routes will there be? e)  How many packets will there be? f) How many frames?

# “Box: Packet Switching”

The generic idea of fragmenting large application messages and sending the segments in small individual messages is known generically as packet switching. However, from the discussion we have just seen, packets are never switched. Rather, frames are switched in single networks and packets are routed in internets. Networking terminology needs a consistency police.

The generic idea of fragmenting large application messages and sending the segments in small individual messages is known generically as packet switching.

Figure 1-: “Packet Switching”

“Packet switching” is a generic name for fragmentation and envelopes

However, packets are never switched

In switched networks, frames are switched

In internets, packets are routed

Test Your Understanding

23. Both single switched networks and routed networks are said to use packet switching. Why is this term confusing?

# Standards Layers

Companies do not want to limit themselves by buying all of their hardware and software from a single vendor. They want to pit vendors against each other so that vendors have to compete for their dollars. Competition brings lower prices. It encourages vendors to keep adding new features to prevent their products from degenerating into commodities. From a risk viewpoint, standards mean that if your vendor falls behind in technology or even goes out of business, your company is not SOL (strictly out of luck).

Competition requires hardware and software from different vendors to be able to interoperate (work together effectively). This requires standards. Standards are detailed rules of operation that specify how two hardware or software processes work together. Standards govern such things as connectors, wiring, wireless operation, switching, routing, frames, packets, and application program messages. We will not look at standards much in this chapter. This is not because they are unimportant. It is because standards are so important that they require an entire chapter (Chapter 2). However, it is important to understand one general concept about standards at this point, namely layers.

Standards are detailed rules of operation that specify how two hardware or software processes work together.

## Five Layers

As discussed earlier, network standards are created in layers. Each layer supports the layer above it. Figure 1-25 shows that there are five general standards layers.

Figure 1-: General Standards Layers

|  |  |  |
| --- | --- | --- |
| Number | Name | Role |
| 5 | Application | To standardize communication between two application programs of a certain type. |
| 4 | Transport | To do application message fragmentation and other functions. |
| 3 | Internet | To transmit a packet across an internet.  Packet organization, router operation, other things needed to transmit a packet across a route in an internet. |
| 2 | Data Link | To transmit a frame across a single network.  Frame organization, switch and access point operation, and other things needed to transmit a frame across a data link in a single network. |
| 1 | Physical | To transmit a signal across a transmission link.  Transmission media, plugs and connectors, signaling. |

## Layers 1 through 3 (Physical, Data Link, and Internet Layers)

We saw that physical links connect adjacent devices, data links are paths that frames take through single networks, and routes are paths that packets take through an internet. They are defined at standard layers 1 through 3. These are the physical, data link, and internet layers.

* Physical link standards describe transmission media and signaling.
* Data link layer standards describe the structure of frames, DLL addresses, and how to deliver frames across a single network.
* Internet layer standards describe the structure of IP packets, IP addresses, and how to deliver packets across a series of routers in an internet.

## Layers 4 and 5 (Transport and Application Layers)

We have also seen the top two layers, although we did not name them earlier. At the top, the application layer standardizes communication between application programs. This includes application messages. HTTP operates at the application layer, standardizing communication between the browser and the webserver application programs. Other application layer standards govern e-mail, FTP, and standards for the Domain Name System, which we will see later in this chapter. Thanks to layering, application designers do not have to worry about transmission through single networks and internets. They concern themselves strictly with application layer matters.

The application layer standardizes communication between application programs.

The transport layer does fragmentation on the source host, which we saw in Figure 1-15. The transport layer also does reassembly on the destination host. This frees application programs from the need to worry about fragmentation. In Chapter 2, we will see that the transport layer has several other functions, including error correction.

The transport layer does fragmentation of the application message on the source host and reassembly on the destination host.

Test Your Understanding

24. a) What is required for competition among vendors? b) What benefits do standards bring? c) Name the five layers from bottom to top. d) Give the number for each layer name. e) What does Layer 1 standardize? f) Layer 2? g) Layer 3? h) Layer 4? i) Layer 5?

## Standards Agencies and Architectures

Standards are created by organizations called standards agencies. Figure 1-26 shows the most important standards agencies. It also shows their standards architectures, which are the broad frameworks within which they create individual standards.

Figure 1-: Networking Standards Agencies and Architectures

|  |  |  |
| --- | --- | --- |
| Architecture | OSI | TCP/IP |
| Standard agency/agencies | ISO and ITU-T | Internet Engineering Task Force (IETF) |
| Architecture name | OSI | TCP/IP |
| Examples of standards | 802.3 Ethernet, 802.11 Wi-Fi, optical fiber | TCP, IP, DHCP, DNS |
| Layers at which dominant | Physical (1) and Data Link (2) | Internet (3) and Transport (4) |

OSI: ISO and ITU-T  Standards at the physical and data link layers are governed by two organizations working together. ITU-T creates standards for telecommunications networks that provide voice, video, and data transmission services. ISO creates many computer standards. Networking involves both computers and transmission services, so this pair of organizations is a natural partnership for creating network standards that combine computers and transmission. ITU-T and ISO develop their standards within a standards architecture called OSI. This is an acronym for Reference Model for Open Systems Interconnection. Mercifully, this is never spelled out. This choice of acronyms is unfortunate because it is easy to confuse ISO (a standards agency) with OSI (a standards framework).

ISO is the standards agency, while OSI is a standards framework.

TCP/IP: The IETF and RFCs   Neither ISO nor ITU-T considered internetworking in detail until internets were well established. Consequently, another organization took the lead in internet standards. This was the Internet Engineering Task Force (IETF). It calls its architecture TCP/IP, after two of its standards, TCP and IP. TCP is the standard for fragmenting application messages. IP is the standard for moving packets across routers through the Internet. Again, terminology is a bit confusing. TCP/IP is a standards architecture; TCP and IP are individual standards, and they are not the only standards in TCP/IP.

The Internet Engineering Task Force (IETF) is the standards agency for the Internet.

TCP/IP is a standards architecture; while TCP and IP are individual standards, and they are not the only TCP/IP standards.

Relative Dominance  Overall, ISO and ITU-T dominate standards development at the physical and data link layers (Layers 1 and 2). In turn, the IETF dominates standards development at the internet and transport layers (Layers 3 and 4). Although there are exceptions, these patterns of domination are very strong.

ISO and ITU-T dominate standards development at the physical and data link layers (Layers 1 and 2).

The IETF dominates standards development at the internet and transport layers (Layers 3 and 4).

The Application Layer  At the application layer, there is no dominance. There are many standards created by many standards agencies. The IETF is even working collaboratively with ISO to develop standards for voice over IP and other applications.

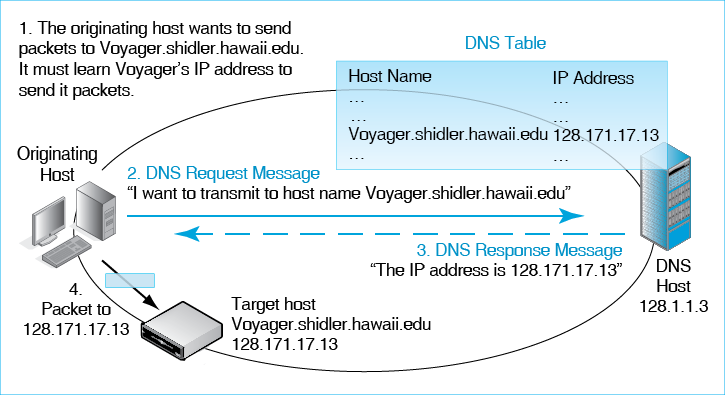
Test Your Understanding

25. a) Distinguish among ISO, OSI, and ITU-T. b) In what two layers do OSI standards dominate? c) What is the standards agency for the Internet? d) In what two layers do its standards dominate? e) What standards agency is dominant at the application layer?

## TCP/IP Supervisory Applications: The Domain Name System (DNS)

TCP and IP (along with the UDP standard we will see in the next chapter) are sufficient for delivering packets over an internet. However, the TCP/IP family of standards also has a large family of supervisory standards that govern matters beyond packet transmission. We will look at only one of them in this chapter, the Domain Name System. You will see many more throughout this book.

Figure 1-: The Domain Name System (DNS)



If you want to call someone from your mobile phone, you need to know their telephone number. If you only know their name, you can call directory service and get their number. Similarly, each Internet host has an IP address, which is its official address on the Internet. Some hosts also have host names, such as Amazon.com and boisestate.edu. IP addresses, not host names, are the official addresses of Internet hosts. You cannot address packets to a host name—only to an IP address. Suppose that you want to shop at Amazon.com. You type its host name in your browser. Your browser, knowing that it needs Amazon.com’s IP address, contacts a domain name system (DNS) server, which is like a directory service for host names and IP addresses. The DNS server looks up the IP address for Amazon.com and sends it back to your browser. Your browser will then send packets to Amazon.com’s IP address.

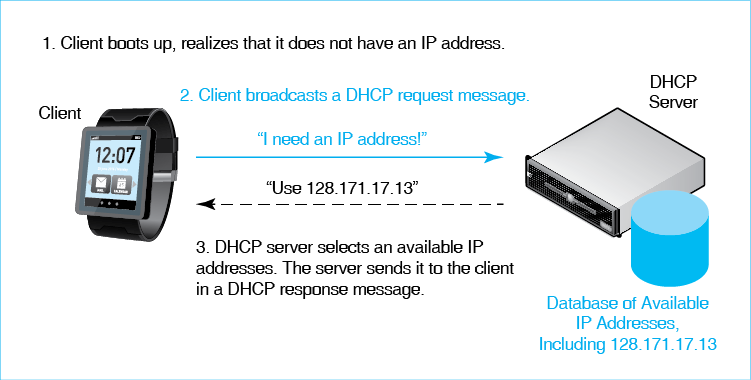
Test Your Understanding

26. a) To send packets to a target host, what must the source host know? b) If the source host knows the host name of the target host, how can it get the target host’s IP address?

## The Dynamic Host Configuration Protocol (DHCP)

The host you use when you surf the ‘Web or do other tasks needs an IP address. How does it get its address? The answer is that client hosts get their IP addresses using the Dynamic Host Configuration Protocol (DHCP). As Figure 1-28 illustrates, when a client device boots up, it realizes that it does not have an IP address. It broadcasts a DHCP request message to its local DHCP server.[[12]](#footnote-12) This message asks for an IP address to use. The server finds an available IP address in its database and responds by sending the client a DHCP response message that includes the IP address. For subsequent packets sent by the client, this is the packet’s source IP address.

Figure 1-: Dynamic Host Configuration Protocol (DHCP)



When the client shuts down, it forgets the IP address. The next time the client boots up, it contacts the DHCP server for a new IP address to use. It typically receives a different IP address each time it does this.

Test Your Understanding

27. a) When a client boots up, why does it need a DHCP server? b) Will the client get the same IP address each time it boots up?

# Conclusion

## Synopsis

The Internet is not the only network. It is, however, the 500-pound gorilla in the room. The Internet is enormous, and it is still growing and evolving at a fierce pace. The Internet is not a single network. Rather, it is a mix of single networks and smaller internets. They work together to deliver messages between any two computers attached to the Internet. We saw that a computer attached to the Internet is a host. The core of the Internet consists of Internet service providers (ISPs), who deliver messages, which we later saw are called IP packets, between customer networks. No organization owns or controls the Internet, and customer fees pay for its upkeep.

“The cloud” is the biggest buzzword in networking today. Unlike many marketing-driven terms, cloud computing is actually important. In fact, it is the major driving force in networking today. We saw how Netflix uses infrastructure as a service (IaaS) from Amazon to host user ordering, transcoding, and customer recommendation analysis. We also saw how it uses software as a service (SaaS) from Google to handle its internal needs. However, Netflix has its own content delivery network for streaming content to individual users. We saw how the use of virtual machines brings agility to IaaS.

Application messages can be extremely long. Network software fragments them into smaller pieces called segments. The network software places these application message segments into envelope messages called frames (in single networks) or packets (in internets). This process of framing/packetization has two major benefits. One is to improve throughput when there are transmission errors. The other is to permit the messages between many pairs of applications to share transmission lines. This sharing, called multiplexing, reduces cost because each conversation only pays for a fraction of the line.

Initially, there were only single networks. A single network uses a single network technology and has a coordinated address space. Messages in single networks are called frames.

In single networks, there are two layers of standards. Physical layer standards govern the physical connection between adjacent devices. Data link layer (DLL) standards specify how frames are organized and delivered over the single network. DLL standards also govern addresses. There is a separate physical link between each pair of devices that are connected. There is only a single data link, which is the path the frame takes in a single network all the way from the source host to the destination host.

In switched single networks, frames travel through multiple switches on their data links from the source host to the destination host. Each switch decides how to forward the frame to the next switch, based on the destination address in the frame. Note that each switch only worries about the next switch along the data link. Individual switches do not know the frame’s entire path through the data link.

What if you were using one single network and the server you wanted to reach was on a different single network? There was no way to do this for many years. Even if the two single networks used the same technology—which was far from certain—some addresses on the different networks may be the same.

Internets solve this conundrum. They do this by adding a third layer of networking, the internet layer. The internet layer defines a new type of message, which is called a packet. Internet layer addresses are assigned globally, so that no two hosts in the world will have the same internet layer addresses. The internet layer also defines a new type of network forwarding device, a router, to connect different single networks. The global Internet uses the Internet Protocol (IP).

Internets do not work independently of single networks. As the packet travels from the source host to the destination host across the internet, it travels through multiple single networks. In each single network, the packet is encapsulated inside a frame specific to that network’s technology. Except in the last single network, the receiving router removes the packet from the arriving frame, puts it in a new frame, and sends the frame over a single network to a next-hop router. This continues until the final single network, which contains the destination host. The final router places the packet in a frame and transmits the frame directly to the destination host. The packet’s passage through the internet is complete. The path that a packet takes all the way across an internet is its route.

Standards are critical to networking because they allow vendors to compete for the customer’s business. To simplify standards (and product) development, the standardization task is divided into five layers. The physical layer, the data link layer, and the internet layer all deal with transmission. The transport layer is responsible for dividing application messages into segments and for other functions we will see in the next chapter. The highest layer, the application layer, governs how application programs understand each other. We looked briefly at an important application layer standard, HTTP.

Standards are set by organizations called standards agencies. At the physical and data link layers, nearly all standards are set jointly by two standards agencies working together, the ITU-T and ISO. Confusingly, their standards architecture is called OSI. At the internet and transport layers, the Internet Engineering Task Force (IETF) dominates standards. Its standards are named TCP/IP standards, after two of its many individual standards, TCP and IP.

Some Internet hosts have host names as well as IP addresses. Only IP addresses are official addresses on Internet. We saw that the domain name system (DNS) standard allows a source host that only knows the host name of a target destination host it wants to communicate with to discover the IP address of the target host. You will see many more supervisory standards as you go through this course.

## End-of-Chapter Questions

Thought Questions

1-1. a) When Host C in Figure 1-23 transmits a packet to Host B, how many physical links, data links, and routes will there be? b) How many packets will there be? c) How many frames? (Check answers are 7, 3, 1, 1, and 3).

1-2. a) When Host E in Figure 1-23 transmits a packet to Host D, how many physical links, data links, and routes will there be? b)  How many packets will there be? c) How many frames?

1-3. a) What type of single network is Network X in Figure 1-23? b) What type of single network is Network Y? c) What protocol does Network Y follow at the data link layer?

1-4. What is the difference between the Internet and the World Wide Web? The answer is not in the text.

1-5. Use Excel’s dec2bin() function to convert from dotted decimal notation to binary. Convert each decimal number segment separately. The IP address in dotted decimal notation is 128.171.17.13. Put a space between each 8-bit segment to make your answer easier to read. Hint: 128 in decimal is 10000000 in binary. Hint: Be sure you have 32 bits in your answer.

1-6. Use the Excel bin2dec() function to convert the following 32-bit IP addresses into dotted decimal notation: 10101010 10101011 00001111 111110000. (Spaces added for ease of reading.) Hint: 10101010 is 170.

1-7. Do you get the same IP address each time your client boots up? This creates a serious problem for peer-to-peer applications in which two clients communicate directly instead of using a server. Can you tell what that problem is?

Troubleshooting Questions

Troubleshooting is an important skill to have when networks go wrong. The job is to find the root cause of the problem from observed symptoms through logical and empirical tests.

* First, understand the symptoms in detail. Often, a small point is the key to identifying the problem.
* Second, know all of the system’s components and decide analytically which ones might be the cause of the problem. This almost always requires you to draw a picture of the network just to identify the elements that need to be considered.
* Third, list all of the possible causes of the problem. You do not start testing them one at a time. To think of one thing and consider it, then do this again and again, is chaotic, unprofessional, and usually futile. Use this approach to answer troubleshooting questions.
* Fourth, exclude as many possibilities as possible logically because they do not fit the details of the situation.
* Fifth, prioritize the alternatives you cannot eliminate logically. Begin with the most likely ones and perhaps the easiest to test.
* Sixth, describe how you would test each alternative.

1-7. A server that you use daily is unusually slow. So are all of the other servers you try. Troubleshoot the problem using the six-step method described above. List the steps in order. Draw the picture.

1-8. You type the URL of a server you use every day. Your browser tells you that the host you are trying to reach does not exist. This message probably comes from your company’s DNS server. Troubleshoot the problem using the six-step method described above. List the steps in order. Be sure to draw a picture of the situation.

Perspective Questions

1-9. What was the most surprising thing for you in this chapter?

1-10. What was the most difficult material for you in this chapter?

1. Suzanne Choney, “US Has More Internet-Connected Gadgets Than People,” *nbcnews.com*, January 2, 2003. http://www.nbcnews.com/technology/us-has-more-internet-connected-gadgets-people-1C7782791. [↑](#footnote-ref-1)
2. Geneva, “[Key ICT Indicators for Developed and Developing Countries and the World (Totals and Penetration Rates),](http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2012/ITU_Key_2006-2013_ICT_data.xls)” *International Telecommunications Unions (ITU)*, February 27, 2013. [↑](#footnote-ref-2)
3. Jared Newman, “Online Video Expected to Overtake DVD, Blu-ray Viewing this Year,” *Techhive*, May 27, 2012. http://www.techhive.com/article/252650/online\_video\_expected\_to\_overtake\_dvd\_blu\_ray\_viewing\_this\_year.html. [↑](#footnote-ref-3)
4. Larry Hettick, “Cisco: Networked Devices Will Outnumber People 3 to 1 in 2016,” *Network World*, June 1, 2012. http://www.networkworld.com/newsletters/converg/2012/060412convergence1.html [↑](#footnote-ref-4)
5. Grant Gross,” Cisco: Global ‘Net Traffic to Surpass 1 Zettabyte by 2016, Cisco Says,” *Network World*, May 31, 2012. http://www.pcworld.com/article/256522/cisco\_global\_net\_traffic\_to\_surpass\_1\_zettabyte\_in\_2016.html [↑](#footnote-ref-5)
6. Ericsson, “CEO to Shareholders: 50 Billion Connections 2020,” press release, April 2010. [↑](#footnote-ref-6)
7. Sources for this section include the following. Brandon Butler, “Three Lessons from Netflix on How to Live in a Cloud,” *NetworkWorld*, October 9, 2013. http://www.networkworld.com/news/2013/100913-netflix-cloud-274647.html. Matt Petronzio, “Meet the Man Who Keeps Netflix Afloat in the Cloud,” *mashable.com*, May 13, 2013. http://mashable.com/2013/05/13/netflix-dream-job/. Kevin Purdy, “How Netflix is Revolutionizing Cloud Computing Just So You Can Watch ‘Teen Mom’ on Your Phone,” *www.itworld.com*, May 10, 2013. http://www.itworld.com/cloud-computing/355844/netflix-revolutionizing-computer-just-serve-you-movies. Ashlee Vance, “Netflix, Reed Hastings Survive Missteps to Join Silicon Valley's Elite,” *Business Week*, May 9, 2013. http://www.businessweek.com/articles/2013-05-09/netflix-reed-hastings-survive-missteps-to-join-silicon-valleys-elite. [↑](#footnote-ref-7)
8. The original term for *internet* was *catanet*. When things are connected together in computer science, they are said to be concatenated. Fortunately, “catanet” never caught on, saving the Internet from a flood of bad feline jokes. [↑](#footnote-ref-8)
9. For simplicity, the figure shows ISPs as if they served nonoverlapping geographic regions. Actually, ISPs often overlap geographically. National and international ISPs may connect at several geographical locations to exchange messages. [↑](#footnote-ref-9)
10. And yes, the bitterness remains to this day. [↑](#footnote-ref-10)
11. Even if the two single networks used the same standards, they might both assign a particular address to a host. If a frame was sent to this address, to which host would the two connected networks deliver it? [↑](#footnote-ref-11)
12. Broadcasting is necessary because the client does not know anything about the network, including the IP address of the local DHCP server. To broadcast the DHCP request message, the client makes the destination IP address thirty-two 1s. When a router receives a packet with an all-1s destination IP address, it broadcasts the IP address to all nearby hosts. All hosts read all broadcast packets. Only the DHCP server responds. If more than one DHCP server responds, the client selects one of them. [↑](#footnote-ref-12)