In today’s connected business world, ubiquitous access to digitized information and data is critical. Networking has not only reshaped the information landscape, but it has also vastly increased speeds at which information moves and is consumed. In fact, networking is the key to accessing the complex services and applications within which information plays a starring role. Although access to the Internet is essential, all networking begins at the local level. For modern networks, Ethernet is the standard infrastructure upon which local networks rest.

Ethernet comprises a family of frame-based protocols and technologies designed to connect local area networks (LANs) — computers and devices situated in close proximity and able to communicate with one another. Ethernet draws its name from the fanciful Latin terminology luminiferous aether, which translates to a light-bearing medium once thought to fill space and propagate magnetic waves. As such, Ethernet is a metaphorical reference for the physical transmission medium in the workplace (and at home) that propagates digitized information in electronic form. The term “Ethernet” was originally coined by
Sensing a Carrier and Handling Collisions

Two people on opposite ends of a phone conversation can sense carrier presence (either a dial-tone or a connected call) and handle collisions (overlapping conversations). Similarly, a partnership paired over Ethernet requires an ability to sense activity, detect collisions, and handle contention to keep communications running smoothly. Ethernet follows a set of standards to establish connectivity and avoid contention so that communicating parties can share a common medium. Without conversational politeness and common courtesy, everybody would be shouting at each other all the time, and nobody could hear anything.

Detecting wire transmission

Even in a room filled with speakers, the human ear is sensitive enough to distinguish individual conversations. Conversational details and voices, however, become increasingly unclear over distance or when competing sounds intrude. Likewise, Ethernet technology has fixed ranges and limitations. For the most part, you and I can socialize freely in a bar or at a convention despite other ongoing conversations. Instead, consider a conference call with multiple speakers and listeners — it’s hard to make out details when several people speak at once.

Similarly, network communications don’t happen randomly: there’s a clearly defined set of rules called protocols that govern Ethernet transmissions. For two devices to communicate, both must understand and use a shared protocol. Detecting wire transmission requires listening in on the wire for an active conversation, much like participants in a conference call. Devices wait their turns before transmitting information just like conference callers wait for breaks and pauses to speak. No two parties should ever talk over each other unless neither wishes to be heard. It’s just that simple.
Detecting collisions

Although it's considered rude for two people to talk at once, it can happen accidentally. Eventually a gap in the conversation creates an opening for someone else to speak, and two people jump in at once. The same concept applies to digital communications. For Ethernet, collisions are detected by comparing transmitted and received data. Any sign of difference indicates two stations transmitting at once and a jam signal results, causing a random delay in both stations to prevent more collisions. Compare this situation to two people on a phone call who start talking simultaneously, after which one person pauses and asks the other person to “go ahead.”

Ethernet’s solution to problems with simultaneous communications is a process called carrier sense multiple access with collision detection (CSMA/CD):

✔ Carrier sensing: Each station continuously listens for traffic on the wire to identify gaps in transmission, like awaiting pauses in a two-way conversation.

✔ Multiple access: Numerous stations may transmit at any time whenever they sense the carrier is not in use, which means no current conversations are active.

✔ Collision detection: If two or more stations in a shared CSMA/CD network or collision domain transmit close to the same time, their bit streams will collide. Special circuitry detects the ensuing garbled signals.

Even with nothing to transmit, a listening station on an Ethernet network constantly listens for activity. CSMA/CD operates at the Media Access Control (MAC) layer of the Open Systems Interconnection (OSI) Reference Model Layer 2 and continually monitors the physical medium for traffic. Whenever the medium is busy, CSMA/CD waits its turn by delaying any pending transmissions of its own.

Carrier sensing goes something like this:

I have something to say. (A frame is ready to send.)

Is anyone already speaking? (Is there any contention on the line?)

If not, I'll take my turn to speak. (Transmit the frame.)
Did I or someone else speak out of turn? (Did a collision occur?)

If so, I’ll wait to speak again. (Reschedule the transmission with a timeout interval.)

The “frame” being transmitted in the CSMA/CD example is an Ethernet frame, shown in Figure 1-1.

![Figure 1-1: Format of a basic Ethernet frame.]

**Backing off and retransmission**

What happens when two transmissions collide and two conversations overlap? Both parties wait a certain time interval before attempting to retransmit. Of course, it would be of no use if both parties waited the same time to try again — the collision would repeat endlessly! Ethernet uses exponential backoff to reschedule transmission following a collision, much like two parties in a phone conversation pause when speaking over one another. An exponential backoff algorithm (also called truncated binary exponential backoff, where truncation simply means the exponential increase stops at a certain point) determines an acceptable rate of delay to avoid further collisions. Although the back-off delay is arbitrary among humans, it’s a matter of precise calculation in terms of Ethernet devices. After encountering \( n \) collisions, a random slot time between 0 and \( 2^{n-1} \) is chosen. During the first encounter, each station may wait 0 or 1 slot times; following the second collision, a station may wait between 0 to 3 slot times until the delay is sufficient to permit retransmission without conflict.
One-Way Broadcast Only, Or Bidirectional Traffic?

You can think of the directional flow of communications between two people in terms of sending and receiving. You’ve probably heard the phrase, “All talk and no listen.” That’s precisely what happens in a one-way conversation where only one person speaks and the other remains silent. For digital systems, this is a form of broadcast communication called unidirectional, in which one station transmits and all other stations receive. However, digital systems enable more complex, two-way conversations in which sending and receiving information occurs at the same time.

Duplex operation

When two devices or parties communicate in both directions (in effect, sending and receiving or speaking and listening) they are operating in full-duplex mode. Full-duplex transmission provides two-way or bidirectional communication between parties to carry a conversation.

Half-duplex transmission

Half-duplex transmission supports bidirectional signals but permits only sending or receiving at any given time — not both at once. Typically, the other party must wait for the transmitting device to stop transmission before sending a reply. Broadcast radios are half-duplex devices. When two people interact over citizen band radio, each party ends with “Over” to mark the end of a statement and to signal the other party to reply.

Half-duplex transmissions allow collisions to occur. Worst case, collision detection between the two most-distant stations on a network takes twice as long (the collision window or slot time) as it takes to send a signal from one end of a network to the other — something called signal propagation time. The compromise was to choose a maximum network diameter and establish a minimum frame length to permit detection of worst-case scenario collisions — this works well for 10 megabits per second (Mbps) connections. However, half-duplex hampers faster Ethernet versions, such as Fast Ethernet at 100 Mbps or Gigabit Ethernet at 1,024 Mbps. Half-duplex
transmissions inconveniently halve available network bandwidth because sending stations can only transmit while recipients await their turns to send.

**Full-duplex transmission**

*Full-duplex transmission* means data can be sent and received simultaneously. Cellular phones, for example, use different radio frequencies to carry bidirectional communications simultaneously. Similarly, Ethernet traffic can move in both directions at the same time without interruption or interference. From an Ethernet perspective full-duplex operation is simpler than half-duplex transmission.

How? There’s no media contention, no collisions, and no delayed retransmissions to manage. The end result: more available time for transmission and effective doubling of the existing bandwidth because each pathway supports full-rate, two-way transmissions.

**Ethernet flow control**

In terms of networking, *flow control* manages the rate of transmission between nodes to prevent fast senders from overwhelming slow receivers. Flow control is a rate limiter that eases heavily loaded receivers from getting flooded with requests from lightly loaded senders, as sometimes happens on large-scale networks with many endpoints.

Flow-control devices are classified by whether or not the recipient issues feedback to the sender. Transmit flow control can occur independently of travel direction, meaning that send rate can differ from receive rate. Flow control can also be *open loop* (no feedback between parties) or *closed loop* (receiver reports pending congestion to transmitter).

**Ethernet Speed Limits**

There are fast talkers and then there are slow talkers. Each of us has our own pace. Ethernet conversations have a similar “wire speed” that dictates how quickly conversation happens.
**Auto-negotiation concepts**

Ethernet devices seek to set mutually acceptable transmission parameters through a process called *auto-negotiation*. Duplex operation and transmission speed are decided during initial setup to establish a mutually workable transmission mode. It's because of auto-negotiation that higher-speed and lower-speed devices can exchange information, much like a slow-talker asking a fast-talker to slow down and speak more clearly.

Every device declares its available parameters and reaches a satisfactory compromise to find the best possible mode of operation. Parallel detection helps devices capable of auto-negotiation identify operational parameters for devices incapable of negotiating automatically. In that case, an auto-negotiation device cannot choose full-duplex, so it always defaults to half-duplex mode.

**Auto-negotiation signaling**

Ethernet auto-negotiation uses electrical pulses sent when no device is sending or receiving data to detect device connectivity. A network device can detect link failures when auto-negotiation packets or pulses do not return within an acceptable time period. Failures can result from a broken transmission medium or the recipient device itself.

**Auto-negotiation operation**

Auto-negotiation seeks to automatically configure two devices on a shared network segment to their best capabilities. Auto-negotiation parameter exchanges involve these steps:

1. Link partners transmit initial negotiation parameters with pulses containing Acknowledgment bits.
2. Stations identify one another as auto-negotiation–capable within the first pulses of an initial exchange.
3. Stations await a few more pulses to enter an Acknowledge Detect state, then exchange code words with Acknowledgment bits set.
4. After reception of three complete, consecutive, and consistent Acknowledgment pulses, both stations enter an Acknowledge Complete state.

5. Several more Acknowledgment bursts follow, with an optional Next Page exchange if necessary.

6. Upon completing a Next Page exchange, stations pick link parameters unless they lack common values.

Link failures, network delays, performance problems, and network troubleshooting difficulties may be eased through optimal link configuration. Auto-negotiation helps compliant Ethernet devices automatically achieve correct and optimal implementation of shared features and functions.

**Standard Ethernet (10 Mb)**

Standard Ethernet or 10BASE-T operates over twisted-pair Category 3 (Cat 3) or Category 5 (Cat 5) cable. The twists prevent crosstalk, which is electromagnetic interference (EMI) produced from neighboring wire pairs and external sources.

A category rating such as Cat 3 or Cat 5 indicates various grades of cable-signal integrity with higher numbers corresponding to higher grades. Standard 10 Mb Ethernet provides the basic transmission rate or “wire speed” for half-duplex and full-duplex network devices.

**Fast Ethernet (100 Mb)**

Fast Ethernet or 100BASE-TX uses Cat 5 and Cat 5e (for enhanced) cable for the greater signal integrity required at higher transmission speeds. Fast Ethernet is capable of carrying basic network traffic along with voice services. Cat 5 and Cat 5e are both capable of 100 megahertz (MHz) signaling as used in 100 Mb networking.

Although there are several Fast Ethernet standards, 100BASE-TX or twisted-pair is by far the most commonly available form in use today. Fast Ethernet’s success is due to its ability to co-exist with established network installations, and many adopters today support standard and Fast Ethernet designs all because of auto-sensing and auto-negotiation strategies.
Chapter 1: Understanding Ethernet

Gigabit Ethernet (GbE)

If you thought Fast Ethernet was fast, Gigabit Ethernet (GbE) or 1000BASE-T is a dizzying blur. Just as Fast Ethernet eventually augmented and replaced standard Ethernet designs, GbE is taking over as the LAN standard. GbE describes technologies that deliver data rates at a speedy one billion (1,073,741,824, actually $2^{30} - 1$) bits per second. GbE is carried primarily across optical fiber and short-haul copper-backbones, into which 10/100 Mb cards can also feed directly.

10 Gigabit Ethernet (10 GbE)

At the upper end of Ethernet speeds you’ll find 10 Gigabit Ethernet (10 GbE, or 10 GBASE-T). It operates at a wire speed 10 binary orders of magnitude faster than GbE (or $2^{40} - 1$ bits per second). Where GbE supports half-duplex operation (but typically operates full-duplex), 10 GbE supports only full-duplex links. Not only is half-duplex linkage unsupported, but CSMA/CD is off limits to 10 GbE. By current Ethernet LAN standards, 10 GbE is a disruptive technology (technology that unexpectedly displaces established technology) that offers faster, more efficient, and less expensive data shuttling.

10 GbE technology uses optical fiber and can replace complex Asynchronous Transfer Mode (ATM) switches and Synchronous Optical Networking (SONET) multiplexers. It targets LANs, wide area networks (WANs), and metropolitan area networks (MANs) using familiar IEEE 802.3 MAC protocols and frames. On multi-mode fiber, 10 GbE goes up to 300 meters, but single-mode fiber spans up to 40 kilometers!

Coming down the road: IEEE developmental Ethernet

40 GbE and 100 GbE networking sit at the edge of two high-speed horizons. The following sections describe each one.

40 GbE

Following after 10 GbE comes 40 GbE, an optical high-speed standard based on the Ethernet protocol format intended
primarily for server and high-end computing, such as linking LAN data centers to a WAN. 40 GbE is ideal for ultra-fast end-to-end switching and routing across enterprise backbones, converged campus networks, supercomputing clusters, grid-computing designs, and high-speed storage networks. 40 GbE service is entering a growth phase that, when finalized, should meet networking demands for moderate capacity until 100 GbE is commercially available.

100 GbE
A much faster forthcoming 100 GbE standard seeks to address aggregate and core networking applications. Single-mode 100 GbE optical transmissions can reach distances up to nearly 40 kilometers (or just over 10 meters over copper), a remarkable feat for Ethernet even by today’s metrics. Aggregate 10 GbE shuttling data across IP routers coupled with 100 GbE ports for switch-to-switch interconnection can operate in an existing 10 GbE infrastructure. 100 GbE is seen as a logical step in the convergence of Ethernet and optical network transport that begins with 10 GbE. As such, industry standards bodies and network operators favor 100 GbE as a preferred mode of transport for future network design and development.

The IEEE Higher Speed Ethernet Study Group (HSSG) has ruled that both 40 GbE and 100 GbE will comprise the next forthcoming Ethernet standard, because both areas exhibit different growth rates that can’t be serviced properly by a single rate. HSSG has been working on this new standard, dubbed IEEE 802.3ab, since 2006. Efforts are split between telephone companies that favor 100 GbE and data center vendors who prefer 40 GbE.

Local and Backbone Ethernet
The transition from consumer-grade to carrier-grade Ethernet begins at the backbone or “core” network. Initially, GbE was deployed in high-capacity backbone network links and enterprise network circles but quickly found its way into the consumer market. Backbone networks interconnect separate LANs and subnetworks and can unite diverse networks within the same buildings or across separate areas.
Basic Ethernet LANs are ill-equipped for network core deployment and ill-suited for Ethernet implementations that a communications carrier might deliver to a customer (usually called Carrier Ethernet). A need for more robust and reliable network capability becomes more apparent when data rates and endpoint distances increase. Current LAN standards, particularly those related to subnet size and subnet bridging, do not scale to carrier levels. The IEEE and Metro Ethernet Forum are working to create a series of standards that elevate Ethernet to carrier-grade specifications.

Three key issues with deploying the Ethernet backbone are maintaining scalability, providing Quality of Service (QoS), and improving resiliency against large-scale failures. Ethernet backbone applications can be broadly divided, as follows:

- **Interface**: Ethernet offers a point-to-point link layer protocol between devices (such as IP routers).
- **Network**: Ethernet serves as a sub-layer over which IP and other higher-level protocols travel.

What does it take to endow Ethernet with carrier-grade capability? Well, for starters, there are five basic attributes: standardized services, high scalability, high reliability, QoS, and service management. We dig into each of these aspects in later chapters where appropriate.

---

**Metro Ethernet Forum (MEF)**

The MEF comprises more than 150 organizations of various kinds: cable operators, service providers, product testers, product vendors, equipment manufacturers, and research laboratories. The MEF leverages these players to define and develop standard performance measures for network equipment for use in metro Ethernet networks (MENs). In addition, MEF standards seek to establish minimum expectations for behavior and performance levels that promote interoperability among carrier-capable network elements.

Product vendors and service providers that adhere to the defined MEF standards receive MEF certificates to establish credibility when selecting equipment and services to implement a Carrier Ethernet solution.
Better Spanning Tree Protocols (STPs) and bridging between subnets are also required for carrier-grade networking. The Rapid Spanning Tree Protocol (RSTP) provides better convergence in large Ethernets, and the Multiple Spanning Tree Protocol (MSTP) supports virtual LAN (VLAN) bridging. However, RSTP and MSTP both present unique challenges, and there is current work underway to improve Provider Backbone Bridging-Traffic Engineering (PBB-TE), itself based on layered VLAN tags and MAC-in-MAC encapsulation defined in Provider Backbone Bridges (PBBs). PBB-TE differs in that it eliminates flooding, dynamic forwarding tables, and STP altogether.

**Ethernet Interfaces**

The basic Ethernet interface has undergone numerous transformations since its inception with D-sub connectors to bayonet coaxial hook-ups and eventually RJ-45 connectors. Although RJ-45 remains the current interface standard, Ethernet speeds and package designs continue to evolve.

**Attachment unit interface**

An Ethernet attachment unit interface (AUI) is a physical port on a computer that combines with a media attachment unit (MAU) that may be internal or external. Together, these components form the basic interconnect between computer and an Ethernet medium. The MAU provides services that correspond to the Physical layer of the OSI Reference Model.

Older AUIs resemble a 15-pin D-sub connector that links VGA monitors to video cards. However, modern Ethernet AUIs for 10BASE-T connections look like telephone plugs but with a much larger RJ-45 connector and interface, as shown in Figure 1-2. (A typical fiber cable connector is also shown for purposes of comparison. Fiber cabling is discussed in Chapter 2.) A modern AUI also uses several more wire pairs than a telephone connector and is entirely incompatible.

**Media-independent interface**

The media-independent interface (MII) shown in Figure 1-3 is a set of electronic components that links Ethernet MAC in the network device to a physical layer device (PHY), which sends
signals onto the network medium. MII may bind to an external
transceiver or connect to chips on the same device. Media-
independence means that any of several PHY devices may be
bridged to support both 10 Mb and 100 Mb operation for con-
nectivity among 10BASE-T and 100BASE-T devices.

Figure 1-2: RJ-45 (Ethernet cable), RJ-11 (telephone cable),
and common fiber connectors.

**GbE media-independent interface**

The GbE MII interfaces between the MAC and PHY on a device
at speeds up to 1000 Mb. The interface is designed for back-
ward compatibility with the MII specification and can also
operate at fall-back speeds of 10/100 Mb in accordance with
the MII specification. Data on the interface is framed using the
Ethernet standard and therefore includes all the same prop-er-
ties: a preamble, frame delimiters, headers, protocol-specific
data, and a cyclic redundancy check (CRC).

**Ethernet network interface**

An Ethernet network interface card (NIC) is a complete pack-
age that incorporates PHY electronics and AUI ports within a
computer. A NIC may be a separate add-in PCI or PCI-Express
(PCle) card, or it can be integrated on the computer’s moth-
erboard. Most NICs are built for a particular combination of
network, protocol, and media, but others can serve multiple
network types.
Each NIC has a unique 48-bit serial number called a MAC address, stored in its read-only memory. The MAC address corresponds to a particular manufacturer, product type, and interface and is used in IP networks to translate between hardware addresses and protocol addresses.

**Plug It In, and It Works!**

Perhaps the biggest benefit of Ethernet design is its simplicity. At the basic level you can just plug a wire between switch and computer and establish connectivity. Because of this, Ethernet lends itself well to scalable internal networks but can also adapt to larger-scale network designs.

Ethernet network devices are designed to auto-sense operational speeds and parameters, and fall-back to mutually compatible settings when interfaces are mismatched. In Chapter 2, we dig deeper into the nature of Ethernet media and SONET and introduce the topic of wireless Ethernet.