1 A Brief History of Ethernet (from a Car Manufacturer's Perspective)

1.1 From the Beginning

In 1969, employees at AT&T/Bell Labs developed the first version of Unix. The original intention was to aid the company's internal development of software on and for multiple platforms. Unix had an unpretentious beginning. However, it evolved to be a very widespread and powerful operating system that facilitated distributed computing. Distributed computing requires communication, which is why the success of Unix is closely linked to the history of Ethernet. One important reason for the success of Unix was that, for antitrust reasons, AT&T was neither allowed to sell Unix nor to keep the intellectual property to itself [1]. As a consequence, Unix – in source code form – was shared with everybody interested.

It was especially, but not only, embraced by universities, and the community that evolved provided the basis for the computing environment we are used to today and in which Ethernet has its place. At a time when computing was dominated by large, proprietary, and very expensive mainframe computers few people had access to, Unix created a demand for Local Area Networking (LAN) while at the same time providing an affordable, common platform for developing Unix [2]. As one example, a group at the University of California, Berkeley, created a Unix derivative known as the Berkeley Software Distribution (BSD). It was first released in 1978 and its evolutions became as established as the "BSD-style license" attached to it [3]. Another example is the Transmission Control Protocol (TCP). The first version of this, published in 1974, was implemented for Unix by Stanford University by 1979 [4]. Later, in 1989, the then up-to-date TCP/IP (Industrial Protocol) code for Unix from AT&T was placed in the public domain and thus significantly helped to distribute the TCP/IP Internet Protocol Suite [5].

The advent of Unix represents an important milestone in the early days of computing. It coincides with a time when a significant number of public as well as proprietary research projects were initiated to investigate methods to interchange data locally and at higher speeds than could be provided for by the telephone system [6]. One of the most momentous projects was the one at Xerox PARC. Xerox needed a solution for data transmission between its first personal computer workstations (called "Xerox Alto"), its laser printers, and the early Internet. Thus, Ethernet was invented (1973), patented (1975) [7], and published (1976) [8].

The general opinion (see e.g. [9]) is that the foundation of Ethernet's later success was laid almost as early in time as this, because of the following two choices:

1. **Opening the technology to others:** At the time, it was common for computer companies to try to bind customers to their products by using proprietary technologies or at least restricting competition with the licensing policy of their patents. Xerox held the patents on Ethernet, but there seems to have been an early understanding that they would profit more from the network effects of a widely deployed Ethernet than from selling the technology itself.[1] Seven years after the invention, on September 30, 1980, Xerox published the "DIX Standard" for Ethernet [10] jointly with the Digital Equipment Corporation (DEC) and Intel. They also offered the technology for adoption to the Institute of Electrical and Electronics Engineers (IEEE) 802 group, very shortly after the group had been founded.[2] With several competing technologies being proposed and followed up, it was by no means evident that Ethernet would prevail. But it did, and one of the reasons attributed to this is that Xerox followed a relaxed licensing policy while not trying to dominate the standardization effort [6]. In the authors' view, such an approach substantially supports the proliferation and success of a standard. Unfortunately, it was an attitude as uncommon then as it is today.
2. **Limiting the technical solution to the task at hand:** Ethernet addressed, and still does address, the communication mechanisms needed on the lower 1½ layers of the ISO/OSI layering model only (see also Figure 1.4 in Section 1.2.1), at a time when the ISO/OSI layering model had yet to be completed. It provided a container that gets a packet through a network with multiple participants, but is as independent from the application layer as possible [11]. When new communication systems are being designed today, there is still a tendency to define all layers, often even without adhering to their strict separation. What supposedly allows for "simplifications" and provides the advantages of complete control over the whole communication stack generally makes the system less flexible and less adaptable to future, and hence unknown, requirements. Indeed, Ethernet's adaptability has proven itself to the extent that it has now been successfully introduced in a completely different physical and application environment: Automotive.

In the years that followed, the IEEE became the host for the development of Ethernet. In 1983, IEEE 802.3 published the first of many Ethernet Standards, 10BASE-5 for 10 Mbps over thick coax cable [12]. In the same year, at least 21 companies were mentioned in the trade press to be developing and/or manufacturing Ethernet products [6]. When on January 1, 1984 the AT&T monopoly ended, the existing installed telephone wiring became usable for competing services and applications [13], and a whole new range of possibilities opened to the networking world. Thus, in 1987, SynOptics, a Xerox spin-off, was the first company to prove the feasibility of transmitting Ethernet at 10 Mbps over telephone wire–type cables with a proprietary Ethernet product [6]. The IEEE ratified the respective 10BASE-T standard in September 1990. Because of the many other proprietary versions of Ethernet that had evolved in the meantime, standardization of 10BASE-T was not obvious and required some effort. Nevertheless, when successful, it sealed the victory over other networking technologies in the market [14]. Shortly after, an optical Ethernet version was developed and published as 10BASE-F in 1993.

Meanwhile, the world around Ethernet did not stand still, but continued to provide means and create demands for networking. Various evolutions of TCP and IP were developed. In October 1989, the Internet Engineering Task Force (IETF) published the complete set of protocols in the TCP/IP Internet protocol suite [15, 16]. As mentioned, the success of TCP/IP was fueled by AT&T's public domain implementation of TCP/IP on Unix [5]. In 1991, the Telecommunications Industry Association or TransImpedance Amplifier (TIA) published a standard for inexpensive Unshielded Twisted Pair (UTP) wiring: TIA/EIA-568. Even today, it is hard to imagine an Ethernet network without the 8P8C/RJ-45 connector described in that standard. The World Wide Web was launched in 1994 [17], and the IETF released a specification for IPv4 routers in June 1995 [18]; the Windows 95 Service Pack-1, released on February 14, 1996, automatically included the Microsoft Internet Explorer 2.0 (i.e., built-in TCP/IP networking), bringing the Internet to the masses [19]. Internet Explorer had been available before but had to be purchased separately.

Subsequently, the IEEE amended and enhanced Ethernet, proving Ethernet's adaptability. First, IEEE 802.3 added, and continues to add, new speed grades. Figure 1.1 gives an overview of the increase in data rates for copper and fiber optical channels, in comparison to when similar speed grades were developed for the automotive industry. The largest data rates envisioned today are 40 Gbps for transmission for twisted pair cables and 400 Gbps for optical communication. Figure 1.2 gives an overview of all Ethernet Physical Layer (PHY) variants developed or under development. It is notable that many of the new developments no longer simply increase the previous data rate by a factor of 10, but that the market diversified with many in-between speed grades.

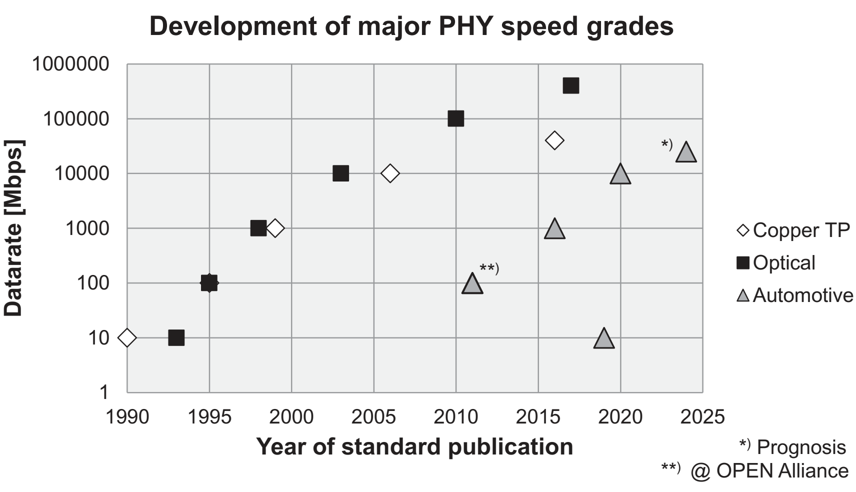

Figure 1.1 Timeline of major PHY speed increases in comparison with automotive speed grades

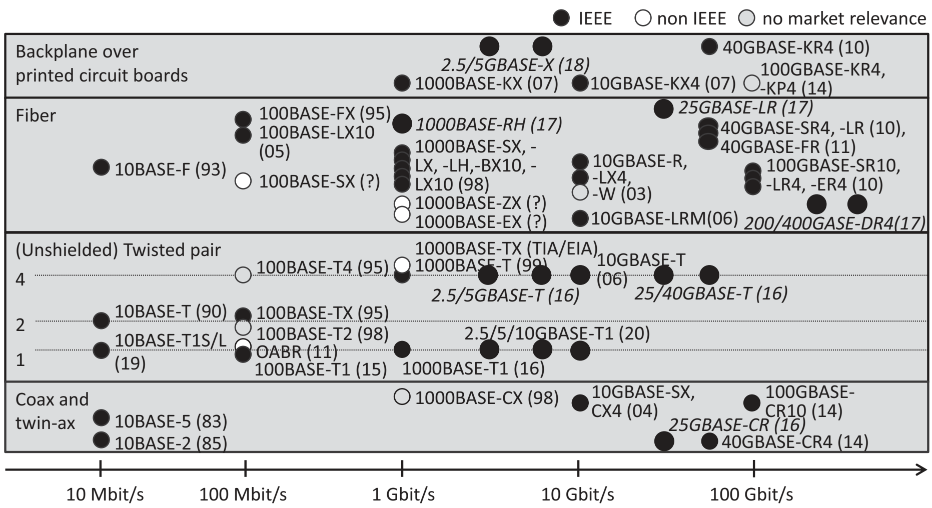

Figure 1.2 2016 overview on Ethernet PHY variants. The (expected) year of release is in brackets, when known [L. Völker, Modified from non-public document, 2011]

Over time, IEEE 802.3 added new functionality and use cases. In 1997, IEEE 802.3 enabled full-duplex communication and flow control to replace the shared media approach prevailing until then. In 1995, IEEE 802.3 added autonegotiation. In 2003, it added Power over Ethernet (PoE), and, in 2010, Energy Efficient Ethernet (EEE). The new use cases include Ethernet in the First Mile (EFM, 2004, see also Section 1.2.2), Ethernet over copper backplane (2007), and finally, in 2013, automotive. The start of work in the automotive space is marked by the establishment of an IEEE 802.3 task force to develop a Reduced Twisted Pair Gigabit Ethernet (RTPGE) (see also Section 5.3.1).

In addition to its PHY-related activities, the IEEE has worked, and is still working on, Quality-of-Service (QoS) schemes for Ethernet and other management functions. Ethernet originally provided quality control only in the form of a CRC check at the receiver, which has no other consequences than offering the possibility to discard packets with detected errors. A pure IEEE 802.3 measure was taken in 1998, when IEEE 802.3 agreed on a four-byte packet length extension in order to accommodate an IEEE 802.1Q header consisting of 802.1 Virtual LAN (VLAN) and priority information. Another important concept was established in 2011, when the IEEE (mainly in 802.1) finalized the first set of standards summarized under Audio Video Bridging (AVB). AVB aimed at improving the quality of audio and video transmissions over an Ethernet network (for more details see Section 7.1). In 2012, the initiative was renamed Time Sensitive Networking (TSN) in order to reflect that use cases with challenging timing requirements beyond audio and video transmission were being addressed in the next set of standards. At the time of writing this in 2020, further enhancements on the AVB/QoS functionalities were still being standardized under TSN at IEEE 802.1.

1.2 The Meaning of "Ethernet"

The term "Ethernet" was first used in 1973, the name referring to the "luminiferous ether" believed by nineteenth-century physicists to be a passive medium between the Sun and Earth, which allows electromagnetic waves to propagate everywhere. The coax used for the inventors' communication system was equally passive and they also intended their data packets to go everywhere [14].

Nevertheless, at first the IEEE did not officially adopt the name (although, unofficially, it did). As an open standards body, the IEEE did not want to give the impression of favoring any company in particular. Despite the fact that Xerox had relinquished their trademark on the name, IEEE 802.3 was instead called "Carrier Sense Multiple Access with Collision Detection (CSMA/CD)" [11]. The official renaming of the IEEE 802.3 efforts into "Ethernet" did not happen until 2007 [20].

As a result, in the various application fields and industries, the name "Ethernet" is used with different meanings, some of which have little in common with what is specified in IEEE 802.3. The following sections thus provide an outline on how different industries use (the term) "Ethernet."

1.2.1 Ethernet in IEEE

Ethernet is standardized in IEEE 802.3 (see Figure 1.3). This comprises the complete PHY and those parts of the Data Link Layer (DLL) that are technology-specific, like the packet format and the medium access method chosen (see also Figure 1.4). Various other aspects also in the IEEE standards (e.g., in IEEE 802.1) affect the implementation of an Ethernet-based communication system. While being relevant, these standards are applicable to all technologies addressed in 802 and are therefore

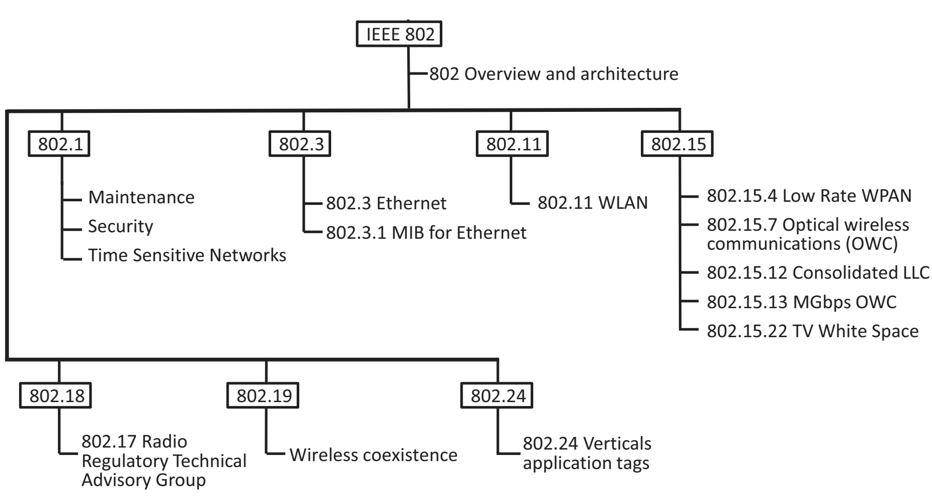

Figure 1.3 IEEE 802 standardization groups active in 2019 [22]

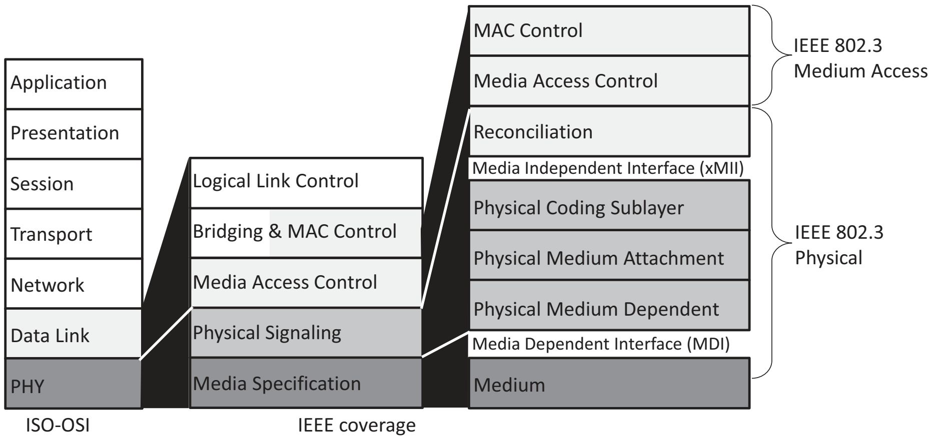

Figure 1.4 Ethernet in IEEE (e.g., [21])

not "IEEE Ethernet"-specific. This is the same for the Logical Link Control (LLC), whose standardization has been concluded in IEEE 802.2 and whose task is to harmonize various methods of medium access toward the network layer [11, 21].

One of the main inventions of the original Ethernet was sharing the media with the help of a CSMA/CD mechanism. CSMA/CD was based on the ALOHA method, which had been developed at the University of Hawaii a few years earlier as a multiuser access method and which more or less simply proposed retransmissions in case collisions were detected [14]. In the case of CSMA/CD, this was enhanced by additionally establishing whether the channel was occupied prior to the start of a transmission. Only if the channel is sensed available is the transmitter allowed to send its packet. Nevertheless, even in this case, collisions can occur, e.g., when another unit had also sensed the channel was available and started transmitting simultaneously. Both transmitters would detect the collision and, in consequence, go into a random back-off period that would increase its potential length with the number of collisions having occurred for one packet [11].

Today, it is hard to find Ethernet installations that still use the CSMA/CD method.[3] The vast majority of Ethernet networks are installed as switched networks with a type of Point-to-Point (P2P) connection in between.[4] In these switched networks, only two PHYs are connected directly, and the switch behind the PHY in the receiving unit forwards the received packets according to their addressing via other PHYs connected to the same switch. The so-called full-duplex[5] operation provides significant advantages in terms of timing and supported link segment lengths [11]. Like in the CSMA/CD mode, the Media Access Control (MAC) is responsible for receiving and transmitting packets in the full-duplex mode. However, the switches necessitated a new sublayer: The MAC control. The general purpose of the MAC control layer is to allow for the interception of Ethernet packets in the case of specific requirements. In the case of a full-duplex switched network, it prevents packet losses in cases where more data is received than can be forwarded, i.e., the MAC control layer enables flow

	Preamble	SFD	Destination MAC address	Source MAC address	Optional 802.1Q tag	Length or Ethertype	Payload/LLC	CRC/ FCS	Inter packet gap
Bytes:	7	1	6	6	4	2	42/46-1500	4	min.12

Figure 1.5 Elements of an Ethernet frame/packet

control to prevent packet losses in the case of network congestion, while allowing for limited resources in terms of the buffering and switching bandwidth. The MAC control layer provides the mechanisms that determine when packets may be sent [21].

The most pronounced and stable element of Ethernet is the Ethernet frame/Ethernet packet (see Figure 1.5). The packet starts with a preamble and the Start Frame Delimiter (SFD). These were originally introduced to help synchronize incoming data in the case of CSMA/CD operation. Starting with 100BASE-TX, more complex signal encoding was introduced, which allows for the deployment of special symbols to detect the beginning and end of a packet. So, when CSMA/CD is not used, preamble and SFD are not needed. They are, nevertheless, kept for backward compatibility reasons.

Each Ethernet interface is assigned a unique 48-bit serial number, often referred to as the “MAC address” or the “hardware address.”[6] Following the preamble, every packet contains information on where the packet is to be sent and where the packet comes from, using the respective MAC addresses. End node MACs initially only examine the destination address to evaluate whether a packet is intended for this end node (as direct/uni-, multi-, or broadcast). If the packet’s destination address matches, the packet is read completely; if the addresses do not match, the packet is ignored. Switch nodes evaluate both the destination address, deciding which port to send the packet to, and the source address, remembering for future incoming packets on which port to find the addressee with that address. This means that there normally is a learning period after start-up in a switched Ethernet network.

The next four bytes represent an optional IEEE 802.1Q tag. The first two bytes identify that this indeed is an 802.1Q header. The remaining two provide the Tag Control Information (TCI) and are divided into three bits for the priority information according to the 802.1p standard, one bit representing the Drop Eligible Identifier (DEI), and 12 bits for the VLAN identifier, which specifies to which VLAN the packet belongs [23]. VLANs represent an important concept for partitioning a physical LAN into various logical domains at layer two (see also Section 7.2).

The next field indicates either the length of the packet or the Ethertype. The Ethertype states what type of data to expect in the payload in respect to the higher layers. It covers content like IP (v4 or v6) or certain AVB packets, but also various proprietary types that have accumulated over time. Ethernet was designed to be a container for whatever data needs to be transmitted; for example, several of the Industrial Ethernet variants – e.g., Profinet, EtherCat, Sercos, Powerlink, High-Speed Ethernet (HSE) – have their own Ethertype (see also Section 1.2.3). The IEEE 802.1Q identifier has the Ethertype 0x8100. A list of Ethertypes is maintained by the IEEE [24]. When the field represents the packet’s length, its value is a number

Table 1.1 Comparison of main Ethernet attributes as originally defined in 10BASE-5 with the "IEEE Ethernet" of today

	Ethernet in 10BASE-5	IEEE Ethernet today
Packet	26+12 bytes overhead, 46–1,500 bytes payload	Optional 4 bytes for 802.1Q tag added
Medium access	CSMA/CD	Full-duplex switches with flow control
QoS	Best-effort traffic without acknowledgments	Additionally, 802.1Q, TSN possible
Signaling	Manchester encoding	Various, e.g., PAM-2/3/4/5, DSQ128, NRZ
Media	Coax	TP, fiber, backplane, Twinax

equal to or less than 1500 (see next paragraph). In this case, the IEEE 802.3 LLC protocol can be used to identify the type of data that is being transmitted.

The payload has a minimum size of 42 bytes when the 802.1Q tag is present, and 46 bytes when it is not.[7] If the data being sent is shorter than the minimum payload, then the remaining bytes of the payload are filled with padding. The maximum payload length is 1,500 bytes. Note that the payload represents user data only from a layer two perspective. Various headers from other layers, like the IP or User Datagram Protocol (UDP) headers, will further reduce the bytes available for the actual application.

Finally, the packet is terminated with a Cyclic Redundancy Check (CRC) called the Frame Check Sequence (FCS). The FCS is 32 bits long and checks the integrity of the various bits of the packet (other than preamble and SFD). Following the packet there must be an InterPacket Gap (IPG) of a minimum of 12 bytes. With a fully loaded payload this means that the header/payload efficiency is over 97%.

Table 1.1 provides an overview of the main components of Ethernet and how they have changed over time. As has been visualized in Figure 1.2, Ethernet has been developed for various media, and almost all but the original one are being addressed today. As a result of higher data rates and advancements in signal processing, the physical signaling has changed with the media and has also been standardized in various forms. The original media access mechanism CSMA/CD is seldom used today. Nevertheless, the principle that Ethernet performs no quality control in form of acknowledgments or retransmits, as well as its "container" function, has been kept. If needed, retransmits have to be initiated on higher layers. Likewise, the Ethernet packet has remained almost unchanged, with only the addition of the optional 802.1Q header.

1.2.2 Ethernet in Telecommunications

The telecommunication providers laid many of the foundations for today's Information Age. Among other things, they were involved in the development of computers, in how to use them (see Section 1.1), and in physically connecting companies and households to the (telecommunication) network, which is still fundamentally important in our networked society today. It thus is surprising that in respect to the adoption of Ethernet and IP the telecommunication industry seems belated and struggling to keep up [25]. However, for a very long time, enabling voice communication between two parties at different physical locations was the main service provided by telecommunications companies. From the invention of telephony in the middle of the nineteenth century [26] to acknowledging the need for changes, the twenty-first century had to arrive [27]. By now, 2020, these changes have become more than just adaptations. Telecommunication providers are in active transition to completely abandoning the original, voice-oriented principles of communication – i.e. circuit-switched/Time Division Multiplex (TDM) communication – for packet-switched IP traffic [28–31]. To explain the developments that led to this and the relationship to Ethernet, a (simplified) distinction is made between the following communication areas: The user domain, the access technologies, and the core telecommunication networks (see also Figure 1.6).

For years, the focus of telecom providers was on improving voice communication. They invested in better voice quality and better coverage. They enabled more simultaneous calls and more connections between the continents and became more cost efficient with improved automation in call switching. The carrier switches that manage the calls between two different subscribers represented a key element in the network and thus were first to go digital. AT&T introduced the first digital carrier switch in 1965 [13]. Ericsson and Siemens developed their commercial and very successful digital switching products AXE and “Elektronisches Wählsystem Digital” (EWSD) in

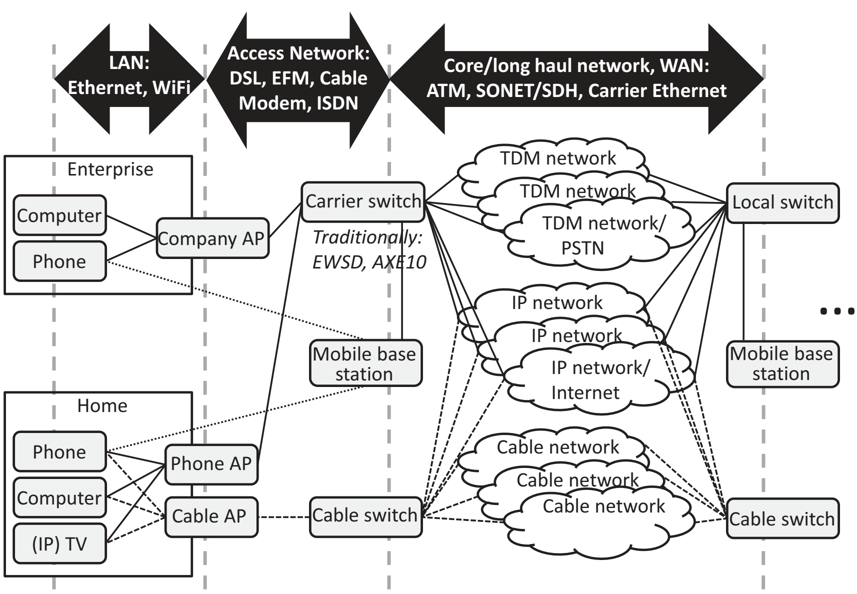

Figure 1.6 Simplified diagram of the relation between the different elements in telecommunications

the late 1970s [32, 33]. All the while, on the subscriber side the communication stayed analog.

The Information Age was thus not driven by the digitization of the Public Switched Telephone Network (PSTN) but by the US military and the computer industry and their desire to share data, which eventually led to the establishment of the Internet. These are the important milestones in this process:

1. The development and publication of IPv4 and TCP as RFCs in 1981 [34, 35].
2. The already mentioned break-up of the AT&T monopoly on January 1, 1984 [13].
3. The creation of the IETF in 1986.
4. The invention of the World Wide Web in 1989/90 [17].
5. The decommissioning of the military run Advanced Research Projects Agency Network (ARPANET).
6. The deregulation of the Internet for commercial use in 1990.

The Internet changed everything. Not only the way people work, live, and communicate but also the paradigm that telephony/voice communication needs to be circuit switched. The application that proved this was Voice over IP (VoIP). There had been very early experiments that showed that voice could be transmitted with data packets. However, the real start of VoIP was in 1995, when the first commercial product allowed real-time, full-duplex voice communication [36]. VoIP first took root in international calls, where the cost savings for users were most significant. The technology improved with the rapidly increasing number of calls and their management in the network became more powerful. With the proliferation of broadband Internet access into the household, VoIP became successful. In 2009, 25% of all voice minutes were using VoIP [37]. By 2013, depending on the actual provider(s), it was possible that a VoIP call was converted to circuit switched for some part of the way or that a regular circuit-switched call was changed into VoIP. The year 2017 saw 1 Billion VoIP users, which is expected to triple to 3 Billion by 2021 [38].

VoIP is one example of the recent changes in telecommunications. The changes apply to all services though. In the past, individual services such as voice, video, data, and mobile were each offered by different service providers. Today, this is no longer the case. Data, web, storage, VPN, voice, and video applications have converged on a technological foundation that is available on the networks of telephone, Internet, mobile phone, and cable TV providers [39]. The underlying paradigm is the provision of IP-based services, and any new telecom network created on a green field would be designed all-IP end-to-end [40]. IP is not a goal in and of itself but enables the use of new methodologies such as Software Defined Networking (SDN) and Network Function Virtualization (NFV) that, in turn, help to simplify network management, reduce operating and equipment costs, create new services, and provide complete and seamless integration [28, 41, 42]. A logical and cost effective way to realize IP is to use Ethernet as well [43, 44]. As a result, the end user can consume hundredfold bandwidth for virtually no added costs.

To deploy Ethernet at the user level (see also Figure 1.6) is simple. Ethernet was primarily designed for enterprise LANs and from there it spread into homes.