TONI JANEVSKI

FUTURE FIXED AND MOBILE BROADBAND INTERNET, CLOUDS, AND IoT/AI

IEEE Press

WILEY

Future Fixed and Mobile Broadband Internet, Clouds, and IoT/AI

IEEE Press
445 Hoes Lane
Piscataway, NJ 08854

Future Fixed and Mobile Broadband Internet, Clouds, and IoT/AI

Toni Janevski
Ss. Cyril and Methodius University
Skopje, North Macedonia

IEEE Press

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey. Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Names: Janevski, Toni, author.
Title: Future fixed and mobile broadband Internet, clouds and IoT/AI / Tori Janevski.
Description: Hoboken, New Jersey : Wiley, [2024] | Includes index.
Identifiers: LCCN 2023057788 (print) | LCCN 2023057789 (ebook) | ISBN 9781394187966 (hardback) | ISBN 9781394187973 (adobe pdf) | ISBN 9781394187980 (epub)
Subjects: LCSH: Broadband communication systems. | Cloud computing. | Internet of things. | Artificial intelligence.
Classification: LCC TK5103.45 .J36 2024 (print) | LCC TK5103.45 (ebook) | DDC 621.39/81—dc23/eng/20240128
LC record available at https://lccn.loc.gov/2023057788
LC ebook record available at https://lccn.loc.gov/2023057789

Cover Design: Wiley
Cover Image: © Yuichiro Chino/Getty Images

Set in 9.5/12.5pt STIXTwoText by Straive, Chennai, India

To my great sons, Dario and Antonio, and to the most precious woman in my life, Jasmina.

Contents

About the Author

Toni Janevski, Ph.D., is Full Professor in telecommunications at the Faculty of Electrical Engineering and Information Technologies (FEEIT), Ss. Cyril and Methodius University (UKIM), Skopje, North Macedonia. During 1996–1999, he worked for T-Mobile Macedonia. Since 1999, he has been with FEEIT. During 2005–2008, he was Member of the Commission of the Agency for Electronic Communications of Macedonia. From 2008 to 2016, he was Member of the Senate of UKIM. In 2009, he established the ITU Center of Excellence (CoE) of FEEIT, and served as its head during the period 2009–2022. Since 2009, he has successfully tutored and/or coordinated many international ITU courses in the ITU Academy every year. He received the "Goce Delchev" state award for science in 2012 and "UKIM best scientists" award for 2013, both of which can be received once in a lifetime. Professor Janevski has written multiple books with the well-known publishers Wiley and Artech House, as well as over 200 scientific papers in the field of telecommunications in journals and conferences. He is elected Head of Telecommunications Institute of FEEIT for the term 2023–2026. Prof. Janevski is elected member of the ITU Group on Capacity Building Initiatives (GCBI) for two 4-year terms, 2019–2022 and 2023–2026.

1

Fixed and Mobile Broadband Evolution

The telecommunication world has demonstrated its power to digitalize the life in the past years, providing all services and applications from the physical world to the cyber (or otherwise called digital) world. When we talk about the digital world, digitalization, etc., we mean the open Internet, which exists parallel to our physical world. So the open Internet is the main "platform" for all the services we enjoy today in our work and life and society in general.

When did the growth of the Internet begin? For the world as a whole, the growth of the Internet began in the 1990s and continues at the same pace until today [1–3]. Nowadays, Internet technologies are the main network technologies in telecommunication networks, including fixed and mobile.

What was crucial for Internet/IP to become what it is today, a major networking technology in today's telecommunications world? Well, it is due to the design of Internet Protocol (IP, as a protocol) to be flexible to accommodate different underlying transport technologies (e.g. Ethernet, Wi-Fi, mobile access networks, optical transport networks, and satellite networks) and all the various applications and services running over the top (therefore called Over The Top – OTT services/applications).

1.1 Evolution of Fixed and Mobile Telecommunications

The evolution of telecommunications started with fixed access networks dedicated to telephony at the end of the 19th century, which continued through most of the 20th century. Telecommunications include all technologies that are available at the given time for transfer of different types of information, such as audio (e.g. voice, or in other words, telephony), video (e.g. television at the beginning and many new video services at the present time), and data (everything else that is not included as audio or video as media, such as various types of services, e.g. email and Web). What technologies existed in the 19th century? Well, that was electricity, so different types of information could be transmitted on distance by using electric signals (e.g. DC – direct current electric signals).

1.1.1 Initial Telecommunication Technologies

What is the first telecommunication service? Well, it's telegraphy, which is a type of data service, because telegrams were messages transmitted by electrical signals over long-distance wires. Telegraphy predates telephony, which was invented in 1876 by Alexander Graham Bell. Working

Future Fixed and Mobile Broadband Internet, Clouds, and IoT/AI, First Edition. Toni Janevski.

telegraphy began in the early 19th century, with the first commercial electric telegraph service opened in London in 1839 with a system created by Charles Wheatstone [4]. In the United States, Samuel Morse created the well-known Morse code in 1844 for use in telegraphy. At the same time in 1843, Alexander Bain patented in the United Kingdom the first image transmission system, which is considered a precursor to the fax services that were used later in the 20th century for business communications. Thus, the first telecommunication service was actually telegraphy, which belongs to data services. With the advent of telegraphy came national and international telegraph services, and international traffic raised the issues of politics, language, and economics, which are also issues nowadays that telecommunications must constantly deal with in parallel with all the technological changes over time. For example, a telegram (that is, a message transmitted through a telegraphic system) written in one language (e.g. English) needed to be translated into the recipient's language (e.g. German, Italian, Spanish, and French). It refers to the content of the message. However, the correct transmission of messages from country A to country B required the same approach to encoding and decoding messages adopted by both countries. Also, the financial part, such as who will pay for sending a telegram (sender, receiver, or partly both parties) was also an issue that needed to be resolved at the national and especially the international level.

The telegraph connected major cities in many different countries around the world in a short period of time. The first submarine cable was used in 1850 to connect France and Great Britain, while the first transatlantic cable was laid in 1858 to connect North America to Europe. All that imposed the need for international agreements between governments, i.e. administrations, and the need to process standardization in telecommunications (at that time telecommunications was basically just telegraphy). This led to the founding of the International Telegraph Union (ITU) in 1865. Later the ITU got the name International Telecommunication Union where the word "telegraphy" was changed to "telecommunication," considering that telephony appeared in 1876 and television was demonstrated for the first time in 1925 in London [4].

But before television, wireless telecommunications began with the invention of the wireless telegraph, invented in 1895 by Guglielmo Marconi in Bologna [5]. The wireless, that is radio telegraphy, was analogous to wire telegraphy used earlier in the 19th century. The Marconi Wireless Telegraph Company was later formed to provide World Wide Wireless (the initial WWW about a century before the emergence of the WWW as well-known web services).

Radio broadcasting and television were the next telecommunication services, after telegraphy as the first and telephony as the second. So, after the data service (telegraphy) and audio service (telephony) appeared radio broadcasting (as an audio-based service) and television (TV) as a multimedia service, consisting of video with accompanying audio. However, the spread of radio broadcasting was in the first half of the 20th century and TV broadcasting (initially also based on radio) in the second half.

Telecommunications originally used so-called analog signals obtained by modulating electrical signals in copper cables or radio signals (for radio/wireless communication) with audio, video, or information data and transmitting such a signal from a transmitter at one end to a receiver at the other. However, it required separate networks for different types of signals, such as separate telephone networks, separate radio broadcast networks, separate TV broadcast networks/equipment, and separate data networks. With the transition of legacy telecom world to Internet technologies it became possible for all services and applications to be provided over the same broadband IP-based networks and services (as shown in Figure 1.1). That provided the possibility to have one broadband network (with fixed and mobile access) for all existing and future services.

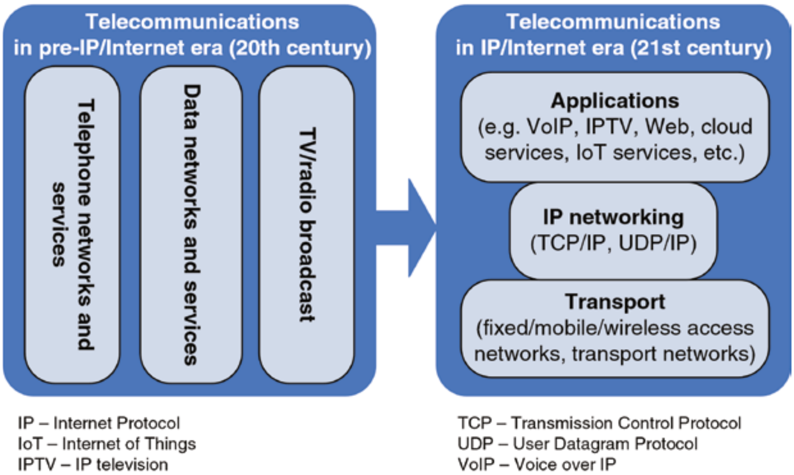

Figure 1.1 Convergence of legacy telecommunication to IP-based networks and services.

1.1.2 Digital Telecommunication World

The driver for the convergence of different types of information into one network was digitalization of signals and systems, driven by the introduction of computer science and informatics into the telecommunications world gradually from the 1960s. By digitizing naturally analog signals, all signals are represented as a series of digits. In the world of telecommunications, it is an unwritten rule that the simplest solution that gets the job done is often the best solution. Thus, although different digital systems can be defined (with a digital base of 2 digits, 3 digits, 4 digits, and so on), the usual approach to encoding information is with the binary system, which consists of two digits, one (1) and zero (0). The simplicity lies in the fact that when a given signal representing a binary 1 or a binary 0 (with noise added to the transmission path) is received, a single threshold is required to decide at the receiver's end whether the digit sent in the given time interval was a binary "1" or "0."

With the transformation of telecom networks from analog to digital since the 1970s and 1980s, it became possible to use the same network for different types of information or media (i.e. audio, video, multimedia, and data). However, different types of media and different services also required different capacity (expressed in bits per second) and different performances by the networks (in terms of end-to-end delay and losses). At that time, it was evident that video requires the most bandwidth (i.e. bitrates), which directly was related to television and its transition from analog to digital. On the other side, voice services (without accompanying video) require much less bitrate than video (e.g. TV).

1.1.2.1 Circuit Switching

Originally, digital telecommunications networks followed the same approach as analog networks before them, that is, they used an allocation of a fixed (i.e. dedicated) amount of bandwidth per flow (for example, a voice call in a given direction) called circuit switching. In fixed digital telephony, the dedicated circuit switching channel was 64 kbit/s in each direction over the telecommunications networks it traversed. The allocation of dedicated bandwidth (e.g. a time slot on a given frequency in a wired or wireless medium) was based on the high synchronization in the network required to multiplex different input streams into larger aggregated streams. Multiplexing is the technique of placing many signals over a single transmission medium (for example, copper, fiber optics, or radio). In general, there are two basic multiplexing schemes used in all digital networks:

- Frequency Division Multiplexing (FDM) allocates different fractions of a given frequency band to different connections (human- or machine-initiated) sharing the same transmission medium. However, they are isolated from each other because they use different frequencies at the same time. When FDM is used in the access network it is called Frequency Division Multiple Access (FDMA). This is the first multiplexing scheme in telecommunications (it was unique in analog telecommunications networks), which will be used at all times now and in the future, considering that telecommunications use frequency bands in every available medium (copper, fiber optics, and radio). Of course, the frequency range used depends on the type of medium and its characteristics in terms of signal attenuation, as well as whether the signals are electrical, optical, or radio signals over copper, optical, and radio links, respectively.
- Time Division Multiplexing (TDM) uses different time intervals, called time slots, which are assigned to different links at the same time using the same frequency in the case of copper cables or radio transmission, or the same wavelength in the case of fiber as the transmission medium. All digital telecommunication transmission systems use TDM. When TDM is used in access networks, it is called Time Division Multiple Access (TDMA). The use of TDM/TDMA requires a certain level of synchronization in the networks.

For multiplexing purposes, circuit-switched transport networks were based on Synchronous Digital Hierarchy (SDH) [3], which uses a centrally positioned primary reference clock with the highest accuracy, and its reference clock is distributed to all nodes in the SDH network via synchronization paths. The American version of SDH was called Synchronous Optical Network (SONET). SDH was based on time slots with a bitrate of 64 kbit/s, originally designed to carry digital voice traffic, although later (from the 1990s onward) the same SDH/SONET transport networks were also used to carry data traffic (e.g. Internet traffic).

What was the main disadvantage of circuit switching? It was the deterministic allocation of resources, such as the mentioned 64 kbit/s in digital telephony. With the aim to provide efficient allocation of network resources, such as frequencies and time slots (as the noted examples of FDM and TDM), circuit switching was inefficient due to several reasons including (but not limited to) slotted resources (e.g. multiples of 64 kbit/s, based primarily on resources needed for digital telephony in the 20th century [3]) as well as use of fixed amount of resources even in cases when there is no traffic to transfer over the given channel (lower flexibility). Therefore, the next stage in the evolution of telecommunications was packet switching, which aimed to provide efficiency and flexibility in resource allocation (frequencies and time slots) in telecom networks.

Figure 1.2 shows mobile and fixed networks based on circuit switching. As for mobile networks, the first digital circuit switching networks belong to the so-called 2G (second generation mobile systems), of which Global System for Mobile Communications (GSM) is the most famous representative that appeared in the 1990s (2G era), standardized by the European Telecommunications Standardization Institute (ETSI). Also, the next generation of mobile networks called 3G had its Circuit-Switching (CS) part, at least the 3G mobile systems that evolved from GSM mobile networks, and were standardized by 3G Partnership Project (3GPP), led by ETSI [6]. The main service in 2G and 3G mobile networks based on circuit switching was voice (i.e. mobile telephony). On the other side, all fixed telephone networks in the 20th century (until the 1990s) were based on circuit

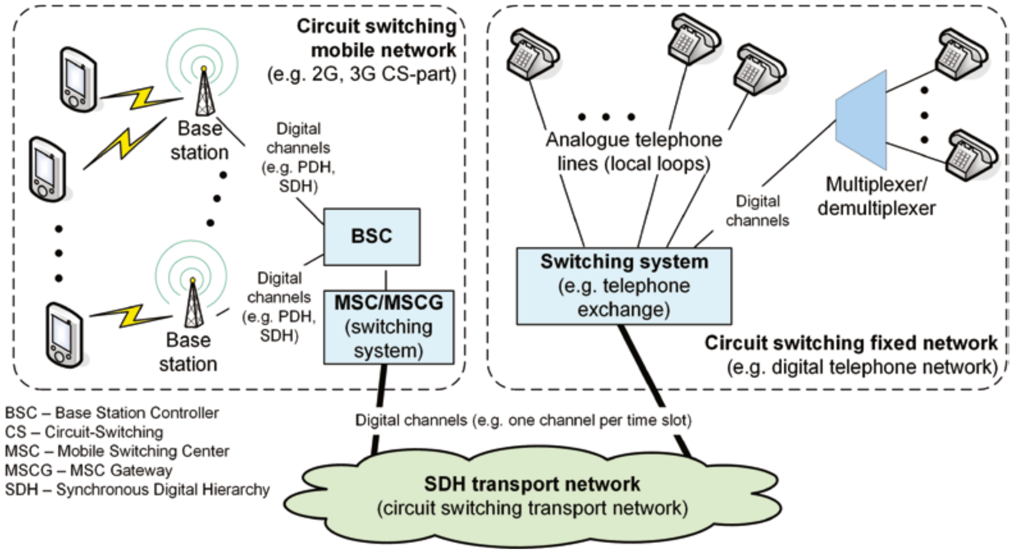

Figure 1.2 Fixed and mobile networks based on circuit switching.

switching (also shown in Figure 1.2). So, the telecom world in the 20th century was based on circuit switching until the 1990s (i.e. before the spread of the Internet/IP technologies).

1.1.2.2 Packet Switching

Packet switching is based on the transmission of information through the so-called packages. What is a packet in telecommunication terms?

Well, a packet refers to a small portion of data that is transmitted based on information contained in so-called packet headers, which carry the necessary control and addressing information for the packet. So, in general, packets are similar to traditional postal packets, with packet "content" (that is, the data, such as digital form of audio, video, and other data, in the packet payload) and an address for the packet's destination (in the header on the packet).

Packet switching is possible only in digital networks, while circuit switching is possible in both analog and digital telecom networks, because the concept of packet header and payload is only possible when the data is in digital form, mainly in the form of ones and zeros (i.e. binary form).

There have been and still are a variety of packet switching technologies. As usual in the world of telecommunications, any new technology is first implemented in the control part of the network, which mainly refers to signaling as a standardized exchange of various control information. For example, in telephony, the important control information is the telephone number for which the dialed digits are transmitted by signaling. Thus, the world's first packet switching system implemented in telecommunication networks was Signaling System 7 (SS7), which was standardized by the ITU for signaling in telecommunication networks in the last two decades of the 20th century.

On the other hand, for packet switching technology to transfer user data there were two main candidates, Internet technologies developed in the United States and Asynchronous Transfer Mode (ATM) developed and supported by European countries. However, ATM, although a packet technology, still largely implemented a circuit switching philosophy, which requires signaling connections to be established before any data transmission. On the other hand, the Internet was created on a best-effort basis, with no guarantees that a given connection will be established or a packet will reach the destination, but the IP network does its best to deliver the IP packet to the destination. The best-effort approach to Internet networking won the packet switching "battle" until the late 1990s (with ATM as the main competition), and then the path was traced to establish Internet technologies as the new telecommunications paradigm in the 21st century for all types of services, including legacy, existing and future ones.

1.2 Internet Evolution

The Internet was the winning technology in the telecommunications world's transition from circuit-switched networks in the 20th century to packet-switched networks in the 21st century. The evolution of the Internet in the world of telecommunications began in the 1990s, although the initial work on Internet protocols began in the 1970s, and especially in the 1980s when the main Internet protocols were standardized, some of which still exist half a century later initial standardization (for example, the first version of the IP, called IP version 4, which was standardized in 1981, was then implemented in all telecommunication networks in all countries around the world in the decades after its introduction [6]).

All the technologies that make the Internet functional and work are standardized by the Internet Engineering Task Force (IETF). For example, the noted IP version 4 (IP) is standardized by RFC 791 [7]. The RFC series was started in 1969 by Steve Crocker and was actually an approach used to take working notes of the ARPANET program (the predecessor of the current Internet). RFC editors' operations were funded by the US government's Defense Advanced Research Projects Agency (DARPA) until 1998, and since then they have been governed by the Internet Society.

What were the main reasons for the Internet's success as a packet switching technology for the telecommunications world? Well, the Internet was created on several principles that made it a global success, of which the following can be considered the most important:

- There is separation of Internet applications and services from the underlying transport technologies (e.g. mobile or fixed access networks, transport networks) via the common networking protocols for all nodes in the network (including end user nodes and network nodes such as switches and routers). For example, Ethernet (the IEEE 802.3 family of standards, started in the 1980s) has been initially created for Internet protocol stack to be implemented over it.
- All network nodes and end user/machine devices have the main Internet protocol stack based on transport layer protocols, which were primarily User Datagram Protocol (UDP) [8] and Transmission Control Protocol (TCP) [9] over the IP, which currently exists in its two versions, IP version 4 (standardized with RFC 791 in September 1981 [7]) and IP version 6 (originally appeared in December 1995 with RFC 1883, was later updated in 1998 with RFC 2460, and finally standardized in 2017 with RFC 8200 [10]).
- Best effort principle of Internet (and generally IP networks) which provided lower costs for Internet network equipment when compared with the costs of the traditional telecom networks in the circuit switching era (i.e. until the 1990s). In fact, in native Internet networking approach there were no performance guarantees end to end, so signaling (which was mandatory in the telephone networks, which were the main circuit-switched telecom networks in the 20th century) became obsolete for many services (e.g. Web services).
- The emergence of Hyper Text Transfer Protocol (HTTP) in the early 1990s and with it the World Wide Web (WWW). In fact, the HTTP as the communication protocol for WWW appeared as an invention by Tim Berners Lee in 1989–1990; however, the crucial moment was not patenting the

HTTP by CERN (where the HTTP inventor has worked in that period), but generally providing to the Internet community (i.e. IETF, to become an open standard) in early 1993. Later appeared standardized HTTP 1.0 (defined by RFC 1945 in May 1996 [11]) and then the long-lasting standard HTTP 1.1 (defined initially by RFC 2068 in 1997 [12], and later updated with RFC 2616 in June 1999 [13]). Until the end of the 1990s, the Web together with email (which was initially standardized in 1970s) was a worldwide success, driving the commercialization of the Internet (moving from being a US federal project to a global network managed by representatives from the players from the Internet ecosystem).

1.2.1 Comparison of Internet and Legacy Telecommunications

A comparison of the traditional telecommunications layering protocol and the IP model is given in Figure 1.3. Originally, the IP model in the early days (in the 1970s) was based on three layers: an interface layer at the bottom, a Network Control Protocol (NCP) in the middle, and an application layer at the top. In 1981, NCP split into TCP (or UDP) over IP, so they became the four-protocol layer model as the native Internet model from the 1980s. However, the network interface layer is typically split into the physical layer and data-link layer by all Standards Development Organizations (SDOs), so with such classification the basic Internet protocol layering model has five layers.

Layering does not forbid multiple protocols to be implemented within a given layer, which in jargon is called layer splitting. Thus, as an example, for audio/video streaming, Real Time Protocol (RTP) is used over UDP, where both protocols (RTP and UDP) belong to Layer 4 — the transport layer. The underlying transport technologies define the network interface, which includes OSI protocol layers 1 and 2. In that manner, these two layers at the bottom of the protocol stack are specific to each transport or access technology, and they are defined in specifications (standards) by the SDO that develops that technology. In that manner, for example, 4G LTE (Long Term Evolution) and 5G NR (New Radio) mobile networks used worldwide are standardized by 3GPP, so OSI layers 1 and 2 for these technologies are standardized in the 3GPP specifications. Another

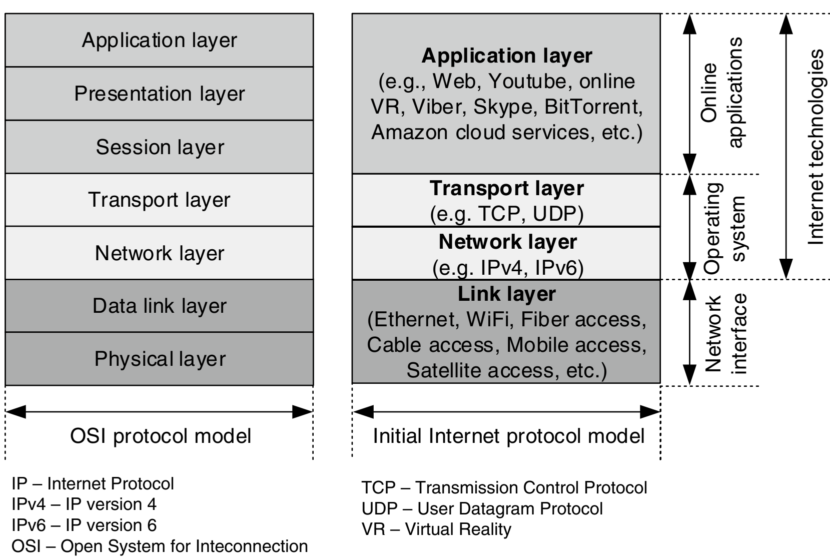

Figure 1.3 Internet protocol layering model.

example, Ethernet (IEEE 802.3 family of standards) and Wi-Fi (IEEE 802.11 family of standards) are standardized by the Institute of Electrical and Electronics Engineers (IEEE) which only standardize the lowest two protocol layers (OSI layers 1 and 2) and assume that Internet protocols will be implemented from Layer 3 (network layer) and above.

Crucial protocols for Internet are IP on OSI layer 3 (network layer) and UDP and TCP on OSI Layer 4 (transport layer). These two protocol layers, which include TCP/IP or UDP/IP, are implemented in the Operating System (OS) of all Internet hosts, including all terminals (user devices such as computers and smartphones, Internet of Things – IoT, devices, and servers) and all network nodes in Internet/IP networks are called routers (because they route IP packets on the path from source to destination). The applications on Internet are implemented on the top protocol layers which include the session, presentation, and application layers under the OSI layering model, and only the application layer under the Internet layering model. Such applications include standardized application protocols such as HTTP and Hyper Text Markup Language (HTML) for Web services, Simple Mail Transfer Protocol (SMTP) and Post Office Protocol 3 (POP3) for email services, as well as numerous proprietary (i.e. not standardized) services, which may be found in application online stores (e.g. Google Play and Apple Store) and other websites on the open Internet (e.g. YouTube, Viber, WhatsApp, Skype, Facebook, Instagram, Twitter, and TikTok).

In fact, the Internet provides access to communications and access and exchange of information between people or machines (i.e. computers), similar to the electrical distribution network for electrical appliances (e.g. washing machines, vacuum cleaners, and mobile phone chargers). For example, power plugs and sockets are the same for different electrical appliances. In the case of the Internet, Internet access is actually a "socket" for all kinds of telecommunication services that allow access to certain information (e.g. certain web content, video content, and various data) or interactive exchange of information (e.g. Voice over IP – VoIP and messaging).

As the winner technology in the packet switching "battle" in the 1980s and 1990s due to noted advantages, from the beginning of the 21st century, Internet technologies gradually penetrated into the legacy telecom networks run by national telecom operators in each country worldwide, supported by convergence of standardization work of different SDOs, including ITU, IETF, 3GPP, IEEE, and others on the national, regional, and global scales. However, IP-based telecom networks differ from the open (or, in other words, public) Internet. Why? Well, because not all telecommunications services are implemented over the open Internet (e.g. carrier-grade telephony and TV services) even though the networks of telecommunications operators have become all-IP.

Figure 1.4 shows the relation between open Internet and IP-based telecom networks. What is the Internet and what are IP networks?

IP networks are all networks that use IP for their interconnection and exchange of information (regardless of its type) between end hosts that are attached to them. All telecom operators have transited to all-IP networks (with few non-IP access networks as exceptions, mainly used for

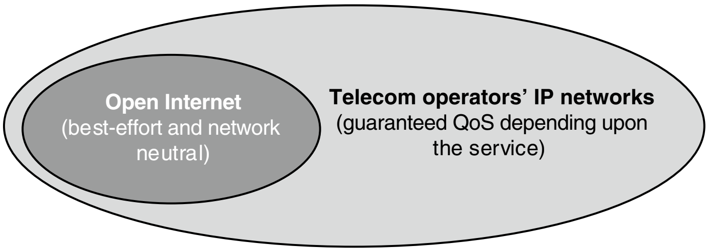

Figure 1.4 Internet vs. IP networks.

connecting low power IoT devices), which are used for open Internet access as well as for provision of legacy telecom services over IP networks with guaranteed Quality of Service (QoS), which are not part of the open Internet network. Such telecommunications services include legacy services such as telephony, TV, and leased lines (for services to businesses), which were transferred from circuit-switched to fully packet-switched telecom networks (here, packet-switched networks are in fact the IP networks), as well as any new services that require guarantees of network performances (e.g. throughput in bits per second, end-to-end delay, and delay variation).

So, the open Internet consists of interconnected IP networks. On the other side, not all IP-based networks are part of the open Internet. As already stated, IP networks carrying carrier-grade telephony, carrier-grade TV, and carrier-grade business services are not part of the open Internet, although such networks are fully based on Internet technologies (e.g. TCP/IP and UDP/IP, as those used in the open Internet). In that way, open Internet can be seen as a subset of telecom operators' IP networks.

But are the open Internet and carrier-grade services (provided by telecom operators with guaranteed QoS and standardized signaling) deployed using physically separated IP networks?

Well, they are not deployed using separate networks; instead, the same fixed and mobile access networks are used for delivery of both open Internet access (or, in other words, Internet Access Service – IAS) and carrier-grade services (i.e. services provided by telecom operators with guaranteed QoS); however, the different telecom services are logically separated from each other (e.g. Internet, telephony, TV, and business services). The same approach of logical separation (i.e. slicing) of deployed network resources and capacity is expected to continue with increasing pace in the future, considering many new emerging services (e.g. IoT services with critical quality and security requirements, such as use cases in industry, transportation, and healthcare).

1.2.2 The Broadband Evolution of Open Internet and IP Networks

Open Internet is characterized by many heterogeneous services that can be provided over it. One of the main features of the open Internet is its network neutrality, which may be considered as a built-in feature from the beginning. In short, network neutrality refers to the principle that all Internet traffic should be treated equally [14]. So, according to the network neutrality rules, the traffic from a niche website should be treated in the same manner as the traffic from global big online services providers such as Google, Amazon, Facebook, and others. However, in 1998, Google was also a niche website, and it has grown on the "wings" of network neutrality of open Internet, which was applicable almost everywhere even without explicit regulation or awareness about it at the beginning of the public Internet (in the 1990s and 2000s).

Network neutrality provided access to the global market for all services and application on the Internet, something that was not possible before (e.g. telephony was the main telecommunication service before the appearance of the open Internet, and it was heavily regulated in all countries). So, the open internet, with its natively built-in network (i.e. net) neutrality, opened up unimaginable opportunities for innovations without asking for permission (e.g. from governments and regulators) that changed the way we lived and worked in the decades since. This has resulted in many new services emerging on the open Internet and shaping the new telecommunications horizon of the 21st century. All such services provided over the open Internet are referred (more like jargon than a strong definition) as OTT services and applications. So, one may say that all services provided through the open Internet access are OTT services, which includes online Web services (e.g. all public websites), online video streaming and video on demand (e.g. YouTube and Netflix), online cloud services (e.g. Google Docs and Amazon cloud), online social networking (e.g.

Facebook, Twitter, Instagram, and TikTok), online gaming (e.g. Steam, Epic games store, and PlayStation gaming platforms), and all other services and applications (including also those with niche market shares) provided through the open Internet access. What is the business logic for telecom operators for provision of open Internet access based on network neutrality and best-effort approach?

Well, with the transition of telecom operators to IP networks from the 2000s onward, Internet access services have emerged as one of the main service offerings from telecom operators. In the last century, until the 1980s, almost all countries had one national telecom operator, managed by the state, aimed at providing telephony and telegraphy as the main services. The advent of the Internet triggered telecom operators to become Internet Service Providers (ISPs), as the Internet itself has used existing telecommunications infrastructure since its inception, first with dial-up modems in digital telecom networks in the 1990s and early 2000s. Later the telecommunications network transitioned to IP-based networking, which was naturally perfect for providing IAS, over either fixed or mobile networks. Telecom operators charge for their IAS services by time usage (that was in the era of dial-up modem access in the early years of the commercial Internet, the 1990s and early 2000s), volume-based, or by a flat fee (in the era of IP-based telecommunications networks). Thus, telecom operators receive revenues by providing access to the Internet as a whole, while online service providers (i.e. OTT providers) usually offer access to certain services for free (e.g. by obtaining revenues from commercial ads on websites) or for a fee (e.g. with appropriate authentication and authorization of the customers). However, without OTT services end users will not need Internet access services from telecom operators. So, it can be said that telecom operators and OTT service providers are related to each other like gas and a car.

The most demanding services in the evolution of the Internet in the 2000s were video services, which required higher bitrates (i.e. higher throughputs) compared to other services consumed by human end users (e.g. VoIP, web, and email). Video bitrates are highly dependent on video resolution and video coders/decoders [15]. Providing video services has multiple dependencies. On one hand, video services require fixed and mobile access networks as well as transport networks to support individual bitrates for video services ranging from hundreds of kbit/s (for lower resolutions, e.g. 480p) up to multiple Mbit/s for High Definition (HD) and tens of Mbit/s for higher video resolutions (e.g. 4K), as shown in Table 1.1. However, one should note that the video compression ratio of the codec (that is, coder at transmitter side – e.g. video server, and decoder at the receiver side – e.g. video player) is dependent upon the processing capabilities of end user devices, such as personal computers, laptops, and others. Higher compression ratio of the codec (which results in lower bitrates needed for a given video stream) on average requires higher processing power of the receiving device to perform rendering in real time with appropriate initial buffering of the video content, needed for smooth reproduction of the video due to delay variations of IP packets carrying

Table 1.1 Required bitrates for video with different resolutions.

Video resolution and frame rate	Approximation of the required bitrate (varies due to codec and frames per second – fps rate)
Below SD (Standard Definition): 180–270p	90 kbit/s to 1 Mbit/s
SD (Standard Definition): 360–540p	150 kbit/s to 4 Mbit/s
HD (High Definition): 720–1080p	500 kbit/s to 5 Mbit/s
Above HD (High Definition), Ultra HD: 1440–2160p	1.5 Mbit/s to 45 Mbit/s

the video data. So, the required bitrates for a given video stream are inversely proportional to the compression ratio of the video, which on the other side is proportional with the processing power of end devices which play the video content.

What is needed for video streaming and video on demand as well as many other different services with various requirements to be deployed on the same physical network infrastructure, either fixed or mobile/wireless one?

Well, the network needs to provide high data rates, or in other words bitrates. Having enough high bitrates (i.e. capacity) to individual end devices, which can provide services with satisfactory quality, is denoted as broadband. Broadband can have various adjectives such as high and ultra when needed to denote speeds that are higher than existing broadband at a given point in time. However, in general, the term "broadband Internet" in telecommunications networks is relative. Why?

Well, because what is broadband today may be considered as narrowband in the future (in a decade or two). For example, in the 1990s and early 2000s individual bitrates of several hundreds of kbit/s were considered as broadband. However, only a decade or two after that time, in the 2010s and 2020s, such bitrates of few hundreds of kbit/s are considered as narrowband. In that manner, bitrates in range of Mbit/s or tens of Mbit/s considered as broadband in the 2010s and early 2020s can be considered as narrowband in the 2030s or 2040s.

The initial development of broadband in the last two decades of the 20th century was intended for transfer of HD video contents, such as HD television. However, the real broadband development was triggered with the spread of Internet, which was suitable for all different types of information, including videos of different resolutions, easily searchable through the WWW, with video content hosted on websites. However, it is impossible to say whether broadband has driven higher resolution video content or whether video content has driven the development of broadband technologies. But it is clear that their development, broadband and video, has been synchronous since the mid-2000s, since the largest video-sharing platform YouTube appeared in 2005.

Figure 1.5 shows fixed Internet access speeds from 1995 onward [16]. The prediction of the average growth of speeds in the future until 2060 is made with the assumption of 20% average annual

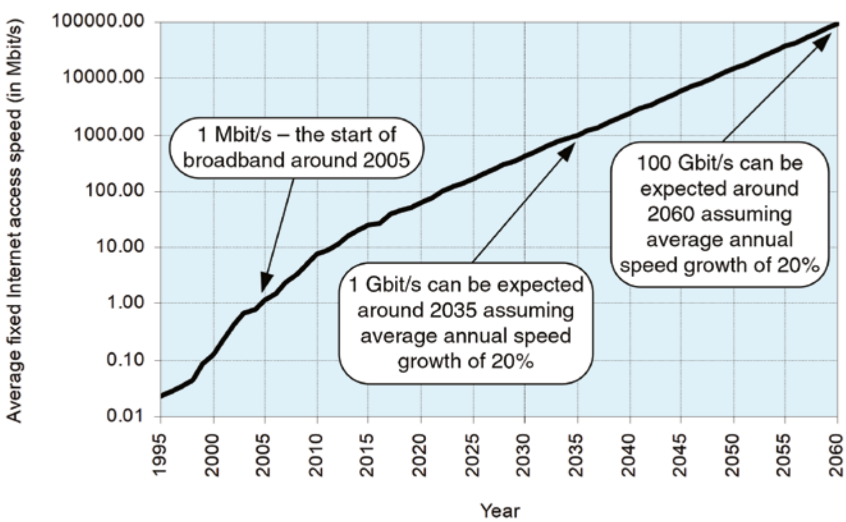

Figure 1.5 Average speeds for fixed Internet access, current and future predictions, 1995–2060.

growth. Thus, with such an assumption, the average speed of fixed Internet access can be expected to reach 1 Gbit/s around 2035 and 100 Gbit/s around 2060 (as shown in Figure 1.5). This is also consistent with Moore's Law [17], which states that processor capacity roughly doubles every two years by doubling the number of transistors on a chip of a given size, with advances in their production over time.

Video sharing platforms such as YouTube were primarily aimed at residential end users, meaning that they emerged at a time when fixed broadband technologies were available for residential users, such as Asymmetric Digital Subscriber Line (ADSL) and cable networks (originally deployed in the late 20th century to broadcast television over fixed access networks). Of course, ADSL and xDSL (digital subscriber line) technologies in general were reusing the existing copper infrastructure consisting of twisted pair lines that were deployed for telephone networks during the 20th century. Twisted pair local loop and coaxial cable access infrastructure (in countries where it already existed) facilitated the rapid deployment of broadband in the 2000s, given that most telecommunications network infrastructure costs come from the deployment of access networks.

However, the highest capacity of all three main media in telecommunication networks has the fiber. It has better performances than copper; therefore, all newly deployed fixed networks are fiber-based (e.g. no one deploys copper, i.e. metallic access, although it continues to be used until fiber is deployed on a given location). Also, fiber has better performances (in terms of capacity and packet losses) than any radio interface, including wireless (e.g. Wi-Fi), mobile (e.g. mobile networks such as 4G and 5G), and satellite broadband access. So, in the long run, all fixed networks will become optical networks. Fiber began to penetrate telecommunications networks at the end of the 20th century, and in the 21st century, especially from the 2010s onward, fiber is the main and only fixed broadband access technology that is deployed globally in all countries. The capacity of the fiber is in the range of tens of Gbit/s in access networks and it is increasing toward the future, while single fiber can carry multiple Tbit/s (Terabits per second, where 1 Tbit/s = 1000 Gbit/s) by using wavelength multiplexing (that is, the data is transferred simultaneously over multiple wavelengths in a single fiber). Of course, fiber used in transport networks needs more capacity to carry the aggregated traffic to/from access networks, and therefore uses more wavelengths than fiber deployed in access networks.

In parallel with fixed broadband, mobile broadband also started with 3G in the 2000s, continued with 4G in the 2010s and 5G in the 2020s, and will continue with 6G in the 2030s, and assuming the pace of innovation already in mobile technologies, there will probably be 7G in the 2040s, and so on. With each new mobile generation, access speeds are approaching those available on fixed networks, and the capabilities of mobile devices are becoming closer to those of desktop computers and laptops, so mobile devices have also become heavy broadband users, especially in the era of 5G (2020s). On the other side, the mobile broadband access contributes to higher affordability and more frequent access to broadband Internet due to two main reasons:

- Every mobile user has their own mobile device, a smartphone, at all times, so the availability of mobile broadband is much greater for end users due to the ubiquitous mobile networks in land-based locations.
- Many countries with developing telecommunications markets had almost no copper (i.e. metallic) access infrastructure due to the absence of higher penetration of telephone networks in the 20th century, resulting in mobile broadband networks in many countries being the main mode of broadband access to the Internet since the 2010s (when mobile broadband became available in most parts of the world) [18].