

※ 源自 CERN 官方网站。

图 2.9 世界上第一个 Web 网站（复原版）

2.3

TCP 普及期

1995 年—2006 年

从 20 世纪 90 年代中期开始，具备拥塞控制等基本功能的 TCP 随着互联网面向普通用户的广泛普及而迅速推广开来。本节就简单介绍一下这个“TCP 普及期”。

Windows 95 发售（1995 年）与操作系统一起普及的 TCP/IP

1995 年，微软发售了面向大众的桌面操作系统 Windows 95。Windows 95 支持 GUI（Graphical User Interface，图形用户界面），与当时的其他操作系统相比，它极具革新性，发售之初便广受讨论，甚至成了社会现象般的热门话题。后来，它更是广泛流行，成了桌面操作系统的默认标准。

TCP 历史上最重要的一个节点，就是 Windows 95（从 OSR 2 开始）默认支持 TCP/IP 协议。随着互联网大面积推广，万维网也流行开来。在这个背景之下，Windows 95（从 OSR 2 开始）基于“让普通用户更容易地接入互联网”的商业战略，将发售版本设计为“在初始配置的情况下便可以使用 TCP/IP 协议”。受此影响，“使用 TCP/IP 接入互联网”的方式开始在不怎么熟悉互联网的人群中推广开来。

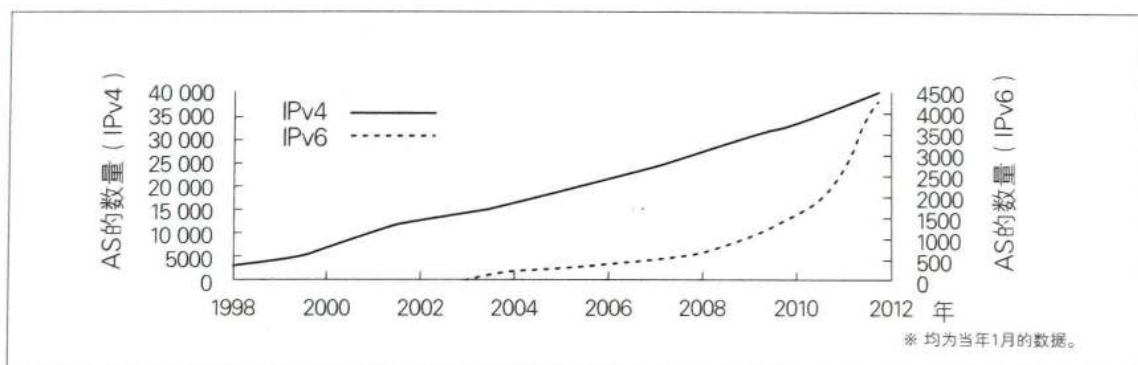
IPv6 投入使用（1999 年）缓慢推进的 IPv6 迁移

当时，IP 地址主要使用 IPv4 的 32 位地址。20 世纪 80 年代以前，IP 地址被批量分配给各个组织。随着 20 世纪 90 年代互联网用户激增，有人指出了 IP 地址可能枯竭的问题。

作为解决方案而开发出来的就是 IPv6，由于它采用 128 位地址，所以实际上可以使用的 IP 地址接近无限个。1999 年，IPv6 地址开始分配。

然而，从不同 IP 版本的 AS^①（Autonomous Systems，自治系统）的数量变化（图 2.10）上可以看出，旧的 IPv4 在 IPv6 出现之后仍被长时间使用，IPv6 的普及其实花费了相当长的一段时间。2010 年左右，IPv6 终于发展起来，IP 地址向 IPv6 的迁移也不断往前推进。后来，随着互联网上 TCP 通信终端设备的增加，IP 地址向 IPv6 的迁移也在不断进行。可以说，IP 地址向 IPv6 的迁移是 TCP 相关技术发展的重要背景。

^① 主要指 ISP 网络。



※ 出处: IPv6 takes one step forward, IPv4 two steps back in 2012

图 2.10 不同 IP 版本的 AS 的数量变化

无线 LAN 出现（1999 年）IEEE 802.11

有线通信曾经因其在速度和可靠性上的明显优势而被广泛使用，但在 2000 年前后，无线通信开始逐渐流行起来。

无线通信接入互联网时最常用的标准就是无线 LAN。图 2.11 展示了无线 LAN 相关的标准。无线 LAN 的总称是 Wi-Fi。1999 年，IEEE 802.11b（2.4 GHz 频段）和 IEEE 802.11a（5 GHz 频段）被先后制定出来。随后，IEEE 802.11g、IEEE 802.11n（Wi-Fi 4）和 IEEE 802.11ac（Wi-Fi 5）等多个标准也被制定出来，相关的产品也开始发售。

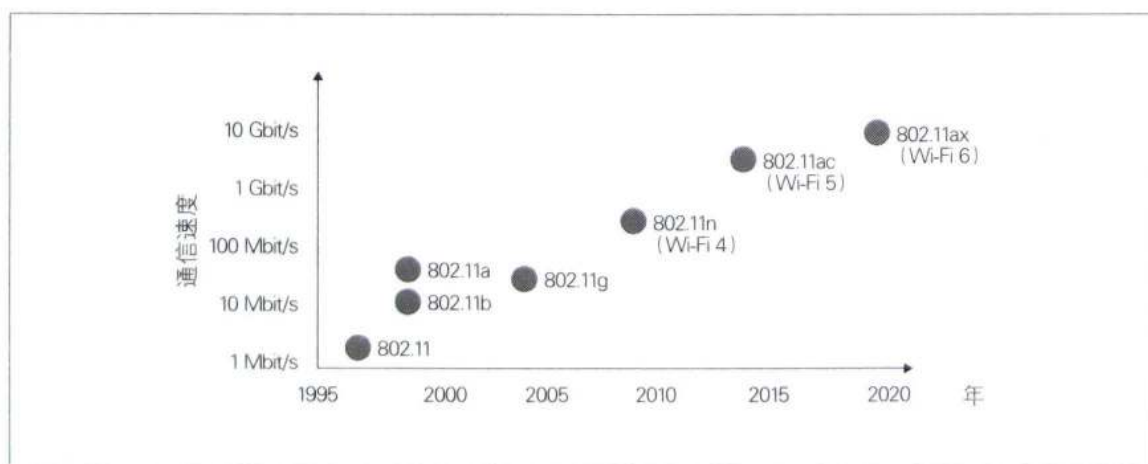


图 2.11 IEEE 802.11 标准化的历史

不仅如此，截至 2019 年，IEEE 802.11ax（Wi-Fi 6）的标准化仍在推

进之中，它在高速化的同时也在逐步地向更大的范围普及。

IEEE 802.11 的特点是，在 MAC 层使用 CSMA/CA 技术。如前文所述，此技术首先确认周围的节点是否在发送数据，然后才开始发送，其目的便是尽可能地规避多个节点的信号冲突。CSMA/CA 技术虽然简单，但避免冲突的效果很好，对于无线 LAN 来说是一个非常重要的技术。

大部分 TCP 拥塞控制算法是“一旦检测到数据包丢失（packet loss，丢包），就减少数据发送量”，因此像 CSMA/CA 这样的底层丢包规避技术，间接对 TCP 通信起了十分重要的作用。

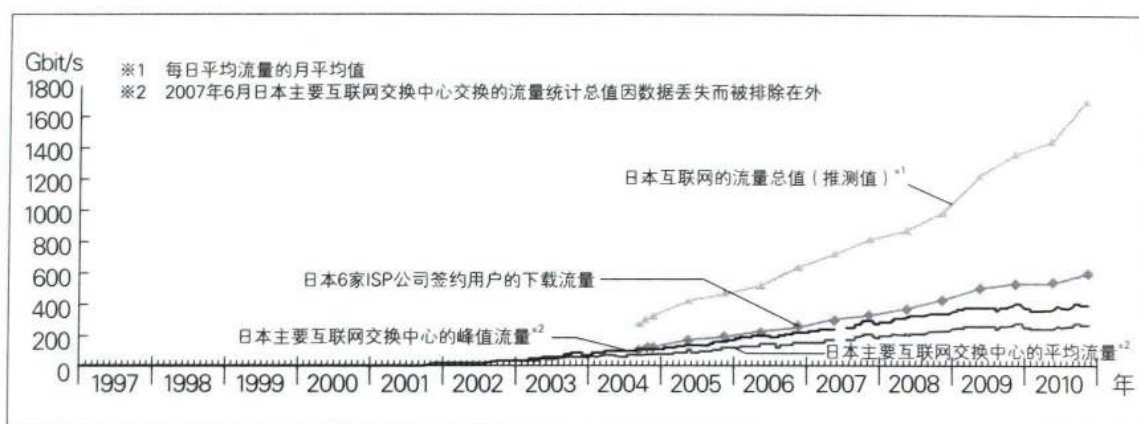
各应用服务 各式各样的互联网应用服务（2004 年—2006 年）特点各异

进入 21 世纪后，随着第 2 代和第 3 代（2nd Generation/3rd Generation，2G/3G）移动通信系统手机的普及，宽带服务和各种互联网应用服务相继登场，并广泛地推广开来。其中，宽带服务采用了由电话线路实现高速通信的 ADSL（Asymmetric Digital Subscriber Line，非对称数字用户线路）、直接将光纤接入普通用户家庭的网络线路 FTTH（Fiber To The Home，光纤入户），以及 CATV（Common Antenna Television，公共天线电视）等线路。

下面列举互联网应用服务中具有代表性的几个例子：任何人都能随意修改的互联网百科全书 Wikipedia（2001 年）、互联网电话服务 Skype（2003 年）、视频分享网站 YouTube（2005 年），以及 SNS 网站 Facebook（2004 年）、mixi（2004 年）和 Twitter（2006 年）等。

从以上例子可以看出，如今仍被经常使用的一些应用服务，在这个时期就已经开始运营了。从图 2.12 可以看出，伴随着各种应用服务的出现，日本的互联网流量不断增长。

从 TCP 通信的视角来看，比较重要的一点是，在互联网上“流量的多少与特性会根据应用服务的不同而变化”。例如，以文本为主的 Web 网站浏览，只需要断断续续地有少量的数据通信即可，而如果是 YouTube 之类的视频分发应用服务，就必须连续下载一定大小的数据。



※ 出处:《2011 年信息通信白皮书》(日本总务省)

图 2.12 日本国内的互联网流量变化

只要分析清楚应用程序生成的流量的特征，我们就可以采取相匹配的流量控制方法。因此，关注网络上的流量有时非常重要。

2.4

TCP 扩展期

21 世纪 00 年代后半期—

从 21 世纪 00 年代后半期开始，随着智能手机和 LTE 的普及，以及云服务的推广，互联网上的流量逐步增加，相应的使用方式也不断变化。本节将这一时期称为 TCP 扩展期。

智能手机普及（2007 年）移动网络连接与 Wi-Fi 连接

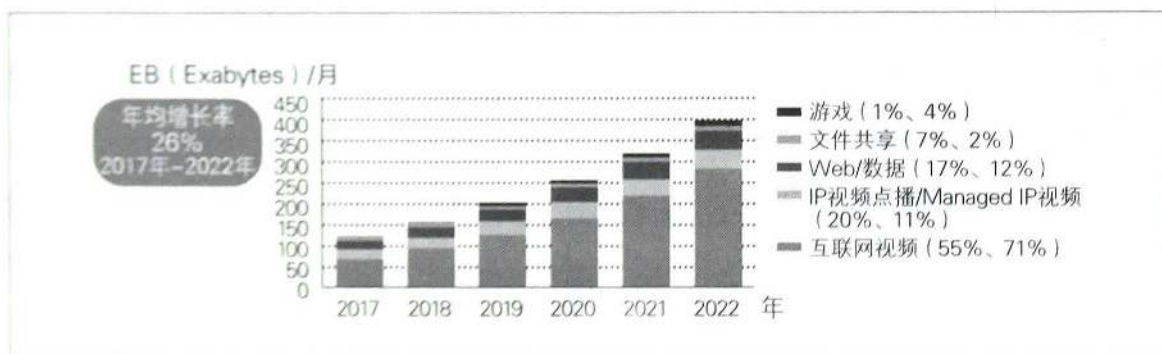
21 世纪 00 年代后半期，智能手机（smart phone）开始发售，取代了一直以来流行的功能机（feature phone），一瞬间便风靡开来。功能机除了提供通话和短信（Short Message Service, SMS）功能以外，只提供了受限的互联网浏览功能，并不支持安装和删除应用程序之类的功能。虽然智能手机的定义比较模糊，但关于它有一点是毋庸置疑的，即黑莓手机（BlackBerry）是初期智能手机的代表。然而，黑莓手机此时仍然以面向商

业为主，普通用户对它并不了解。

智能手机真正的爆发性发展，是从2007年iPhone的发售开始的。究其原因，是触摸屏支持点击即用的直观操作，以及各种应用程序提供的种种方便功能。现在，iPhone手机使用的iOS操作系统和谷歌开发的安卓操作系统在智能手机操作系统中占据的份额最大。

智能手机与以往的手机一样，需要接入移动网络之中（详见后文）。此外，它一般还搭载了Wi-Fi功能。随着各种应用程序的普及，智能手机的功能逐渐多样化，用户使用智能手机的时间也越来越长。从20世纪00年代后半期开始，移动通信流量增长的势头十分迅猛。

在智能手机带来的网络流量中，除了一般的互联网访问流量，观看视频所占的比例大幅增长，具体情况如图2.13所示。另外，向SNS投稿照片和视频导致的上传流量增长，以及因OS和应用程序定期更新导致的下载流量剧增等现象也引起了关注。



※ 出处：Cisco Visual Networking Index: Forecast and Trends, 2017-2022 (Cisco Systems, Inc., 2018)

图 2.13 不同类型的互联网流量变化

云计算出现（2006 年） 远程通信流量增加

通过网络提供计算功能的服务从很早以前就一直存在，但云计算（cloud computing）这个词及其概念直到2006年—2008年左右Google App Engine和Amazon EC2的出现，才一下子普及开来。

以往的（非云类型的）服务或者应用程序通常是安装和搭建在本地计算机系统上的，用户在搭建之后，直接使用本机系统的服务。与之相对，

云计算（图 2.14）则是由服务提供商先把系统搭建在数据中心，之后用户通过互联网访问远程的系统，然后使用其中的功能与服务。

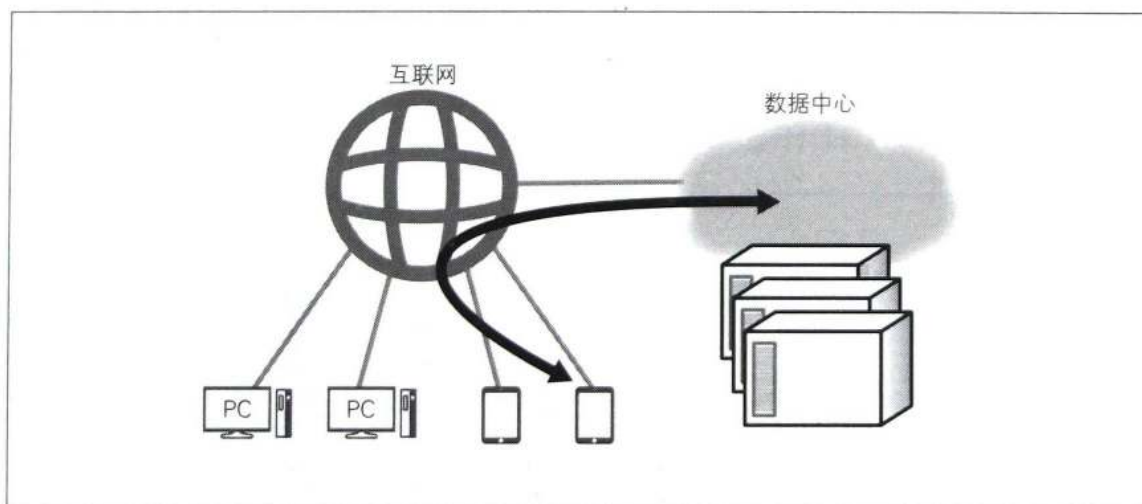


图 2.14 云计算

此类服务提供方式的优点是，用户端只需要一台可以连接到互联网的 PC 或智能手机等普通终端设备即可，因此各种相关的应用服务迅速地普及开来。此外，服务提供商的数据中心集中配置了大量服务器和存储设备，逐步趋向大型化。这就意味着，用户端的设备通过互联网连接到远程的数据中心内的计算机集群时，连接产生的通信流量会变得很大。

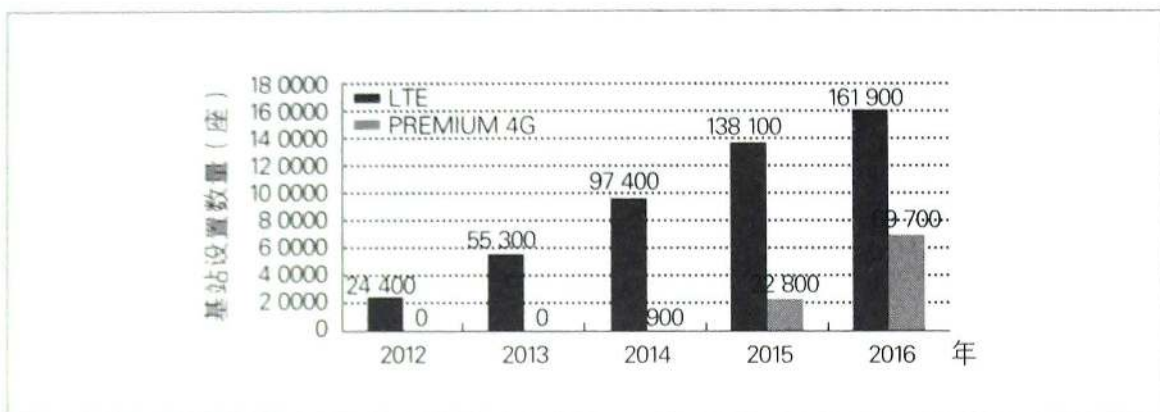
无线网络的特点与对 移动网络的高速化（2010 年、2015 年）TCP 的性能要求

移动通信系统，也就是移动网络（mobile network/cellular network），从诞生之日起便一直在高速化。20 世纪 80 年代，其第 1 代（1st Generation, 1G）的通信速度只有几十 Kbit/s，而且只供汽车电话等少数应用程序使用，从第 2 代（2G）开始发展为分组交换式，能支持电子邮件和互联网访问。

之后，第 3 代（3G）通信速度进一步提高，到了 2010 年左右，3.9G 技术 LTE 开始普及。3.9G 从字面含义来看是一个比 4G 稍微落后一点的技术，但是实际上电信从业者却是以 4G 的名义对外提供 LTE 服务的。LTE 的通信速度达到几十 Mbit/s，应用范围逐渐扩大，人们开始使用移动

网络进行浏览视频等操作。

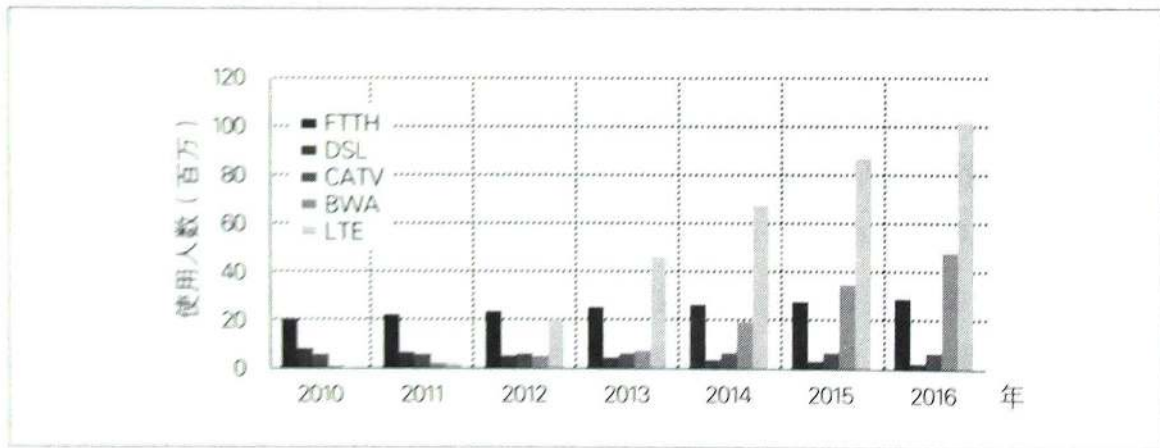
随后，在 2015 年左右，第 4 代（4th Generation，4G）技术 LTE Advanced 开始普及（图 2.15 中的 PREMIUM 4G），其理论上的速度能达到约 1 Gbit/s。此外，2017 年第 5 代（5G）标准完成制定，想必今后会在全世界范围内普及。



※ 出处 NTT DOCOMO Annual Report 2017

图 2.15 NTT docomo 基站设置数量的变化

移动网络的通信速度逐渐提高，与此同时，使用光纤等固定线路的宽带签约用户数量却增长缓慢（图 2.16）。移动设备产生的网络数据在互联网通信流量中占据的比例不断上升。此外，图 2.16 中所示的 DSL 指的是 ADSL 和与之相似的通信方式，BWA（Broadband Wireless Access，宽带无线接入）则指的是无线高速通信标准，其代表之一是 WiMAX。



※ 出处《2017 年信息通信白皮书》（日本总务省）

图 2.16 日本的宽带使用者的人数变化

与有线通信不同的是，当传输媒介是无线时，更容易产生比特差错等，这导致通信对 TCP 的性能要求也有所变化。因此，3G 网络下使用的 W-TCP（详见 1.6 节）等几个面向移动端的 TCP 协议被开发出来。W-TCP 设置用于 TCP 中转的网关，网关为移动设备和通信服务器提供中转服务。换句话说，就是在移动终端与网关之间，针对丢包使用更健壮的 W-TCP 进行通信，而在网关与通信服务器之间使用普通的 TCP 协议，通过两者的结合实现更为高效的通信。

物联网的大众化（2015 年）低功耗、远程的数据通信服务

2015 年左右，“物联网”一词与相应的应用服务火遍各地。“控制各种连接到网络上的物体”的概念本身最早在 20 世纪 80 年代就已经被人们讨论过了，而且也出现了不少实际例子。近年来，随着智能设备（smart device）的普及和无线协议的推广，物联网的概念迅速流行开来。

如图 2.17 所示，想必今后在互联网上，物联网相关的流量一定会继续增加。



※ 出处：Cisco Visual Networking Index: Forecast and Trends, 2017-2022 (Cisco Systems, Inc., 2018)

图 2.17 物联网相关流量的变化

物联网中使用较多的无线通信协议是被称为 LPWA（Low Power Wide Area，低功耗广域技术）的协议群。LPWA 是低功耗、远程通信技术的总称，其通信距离从几百米到几千米不等。LPWA 的代表性技术标准有 LoRaWAN、SIGFOX 和 NB-IoT（Narrow Band IoT，窄带物联网）等。这些技术的主要特点是通过控制传输速度，保持几十 Kbit/s 左右的低速、间断型通信来减

少电力消耗等。

此外，物联网设备通常处理性能不高，而且还有一个特征是会有大量设备同时接入互联网中。另外，还有使用电力受限、设备架设之后更新与更换非常困难等情况。

由于物联网服务的这些特点，其在 TCP 通信方面的新问题愈发突出。例如，为高性能处理设备设计并经过调优的复杂控制算法在物联网设备上很难顺利运行，或者旨在提高网络速度的拥塞控制算法与低速且安装在不稳定环境中的设备之间的兼容性并不好。第 7 章将详细介绍近年来的这些问题和发展情况。

2.5

小结

本章通过梳理不同时期的相关技术动向与出现的各种应用服务，对 TCP/IP 从问世之初的背景到迄今为止的各种变迁进行了介绍。

本章首先介绍了 1968 年到 1980 年左右，TCP 基本成形之前的“TCP 黎明期”。这段时间出现了多项对于 TCP/IP 普及非常重要的技术，例如阿帕网和 ALOHAnet 的搭建、UNIX 和以太网的开发等。总的来说，阿帕网开始使用 TCP/IP 作为通信协议，与此同时 UNIX 系操作系统 BSD 开始默认搭载 TCP/IP 协议，这些都最终推动了 TCP/IP 的普及。

接下来介绍了 1980 年到 1995 年左右，新加入了拥塞控制等新功能的“TCP 发展期”。在这个时期，网络上的流量逐渐增加，拥塞等问题逐渐暴露出来，人们开始担忧拥塞崩溃。针对此问题，Nagle 算法和拥塞控制算法 Tahoe 被开发出来，这些根据情况调整数据发送量的技术被引入 TCP 中。

然后，当时既存的各种网络相互连接形成互联网，彼时诞生的万维网也成为主流的应用程序，而 TCP/IP 作为互联网所用的通信协议一下子快速发展起来。另外，1995 年到 20 世纪 00 年代中期被称为“TCP 普及期”。这段时间，Windows 95 发售，无线 LAN 出现，SNS、YouTube 等各

式各样的互联网应用服务面世。TCP/IP 随着互联网向普通用户的普及而快速渗透到方方面面。

最后，本书介绍了 20 世纪 00 年代后半期的“TCP 扩展期”。在这个时期，智能手机和云计算开始普及，移动网络也通过 LTE 和 LTE-Advanced 技术实现了高速化。其结果是，智能手机通过互联网与设置于远程数据中心的云服务器进行通信，此类通信的流量急速增长。到了 2015 年左右，受智能设备与无线通信协议 LPWA 的普及等情况影响，物联网服务也快速推广开来。

TCP/IP 就像这样随着互联网的普及而推广开来，不断地发展进化。不仅如此，新技术和新应用服务的普及又促使 TCP/IP 进行了各种技术改良，最终形成了目前 TCP/IP 所用的各项技术。

下一章将介绍 TCP 的工作机制及各种拥塞控制算法。届时请结合本章所介绍的历史发展情况，去理解各种技术是在什么样的背景下被开发出来的，又有着什么意义。相信理解了这些，一定有助于你更好地掌握现代 TCP 所使用的各项技术机制与算法。

参考资料

- 维基百科
 - 日文版维基百科 “インターネットの歴史”（互联网的历史）条目
 - 日文版维基百科 “ARPANET” 条目
 - 日文版维基百科 “ALOHAnet” 条目
 - 日文版维基百科 “イーサネット”（以太网）条目
- Joshua Gancher. TCP Congestion Avoidance [EB/OL]. 2016.
- Andrew S.Tanenbaum. 计算机网络 [M]. 熊桂喜，王小虎，译．北京：清华大学出版社，1999.
- Internet Society. Brief History of the Internet [EB/OL]. 1996.
- 日本总务省．2017 年信息通信白皮书 [R/OL]. 2017.
- NTT DOCOMO. Annual Report [R]. 2017.
- Cisco Systems, Inc., . Cisco Visual Networking Index: Forecast and Trends 2017—2022 [R/OL]. 2018.

第 3 章

TCP 与数据传输

实现可靠性与效率的兼顾

TCP 的主要特性是确保了可靠性，即第 1 章所定义的“将发送方待发送的数据无乱序、无丢失地发送给接收方”。与此同时，TCP 还会规避网络拥塞，尽可能高效地传输数据。

可靠性与高效率两者想要兼顾并不容易，TCP 为了这一目标进行了无数次优化。我们可以从无数次优化的背后窥得如何灵巧地控制肉眼无法看见的网络，以及 TCP 艰辛的进化之路。

本章将介绍 20 世纪 90 年代所确定的 TCP 的基本功能。

3.1

TCP 的数据格式

数据包与首部的格式

TCP 的功能究竟是基于什么样的机制实现的呢？

本节将介绍传输层通信所使用的数据结构。

数据包格式 首部与数据部分

不单是 TCP 协议，只要通信设备想要实现协议通信，就必须先定义好数据的格式。

从格式上来说，数据基本上分为两部分：存储传输层控制信息的首部和存储应用程序所需数据的数据部分。

参与通信的设备根据首部格式中指定位置的信息完成通信。一般来说，一个通信协议的内容是基于“首部定义了什么信息”和“基于这些信息要对设备进行什么样的处理”来确定的。例如，一方如果想让对方进行某种操作，就需要把指示信息写入首部的预定义位置上，然后发送出去；收到数据的设备读取首部中的对应信息，按照指示完成操作。人与人之间靠互相沟通可最终完成任务，而收发数据双方通过首部的信息也能实现人类那样的沟通，并最终完成通信。

TCP 报文段 MTU、MSS、链路 MTU 获取和分片

TCP 协议追求的第一目标是“确保通信可靠性”，但是也追求“实现高网络利用率的数据传输”。

从效率的观点来看，数据不应被切割成小份传输，而应该整块传输，这样效率才更高。然而，通信链路都有自己的最大数据传输量。这一限制主要是由下层数据链路层规定的，因此 TCP 必须在这个限制下确定传输的数据量大小。接入网络中的设备，通过以太网、PPPoE（Point-to-Point

Protocol Over Ethernet, 以太网上的点对点协议) 和 ATM (Asynchronous Transfer Mode, 异步传输模式) 等各种不同的数据链路连接到网络上, 不同的数据链路对应的最大帧长度均为固定值。1.6 节简单提到过, 此长度称为 **MTU** (最大传输单元)。不同网络的 **MTU** 值如下: 以太网是 1500 字节, PPPoE 是 1492 字节, 而 ATM 是 9180 字节。近年, 由各台设备自行设置 **MTU** 的方式逐渐流行起来。

在一般情况下, 数据的传输单位称为包 (但如果是以太网这种数据链路层, 其传输单位则称为帧)。数据在发送之前, 必须先被分割为对应数据链路的最大数据传输量 (即 **MTU**) 的大小。在 TCP 中, 分割之后的包称为“TCP 报文段”。

这时, TCP 能分割的最大包长度称为 **MSS** (Maximum Segment Size, 最大段长度)。TCP 首先确定 **MSS** 值, 再开始通信。**MTU** 定义在数据链路层, 因此它的值的大小是包含了 IP 层首部的。也就是说, **MSS** 的值如图 3.1 所示, 是从 **MTU** 中减去 IP 和 TCP 首部的长度之后得到的。包含第 5 层以上的首部的应用程序数据需要被分割或调整为小于等于 **MSS** 的大小。



图 3.1 MSS 的设置

因此, TCP 需要获取所经过链路的 **MTU** 大小, 这主要是通过 IP 层的 ICMP 协议实现的。这个过程称为“链路 **MTU** 获取” (path MTU discovery)。得到 **MTU** 之后, **MSS** 值便可结合最小的 **MTU** 值计算出来。

如果不进行链路 **MTU** 获取, TCP 就必须随后进行 IP 数据包分片 (IP fragmentation, IP 分片), 这需要的操作就更多了。因此, 要想在 **MTU** 的限制下实现更高效的数据传输, TCP 需要提前设置好合适的 **MSS** 值。

TCP 在分割应用程序传过来的一系列数据时, 只将这些数据看作无格式的比特序列。它在彻底忽略应用层的基础上分割数据, 并将其发送出去。接收方应用程序要想完整地复原收到的数据, 完成通信, 就必须通过