

网络的宏观视角

前两章介绍了很多概念和思想,本章再给读者增加一些对网络的感性认识。

3.1 网络的通信范围

网络的分类有很多种,其中一种是按照网络的作用范围进行分类,这种分类只以距离作为分类的依据,和具体的通信技术无关。

1. 广域网

广域网(Wide Area Network, WAN)一般是进行大距离通信的公用网络,目的是把分布于各地的网络进行延伸,通常由大的电信公司经营管理。

广域网所覆盖的范围可达几千千米,能连接多个地区、城市和国家。

最早的长距离通信需要借助于公用电话网,通信容量小。目前广域网的通信介质大量采用了光纤,容量很大,支持很多用户的同时通信,能够满足大批量与突发性通信,一般提供开放的接口与规范化的协议来提供公共通信服务。

当前有线的广域网中,同步数字体系(Synchronous Digital Hierarchy, SDH)占有重要的地位,而4G、5G等蜂窝通信网是无线广域网。

2. 城域网

城域网(Metropolitan Area Network, MAN)是在一个城市范围内所建立的公用网络,作用距离为5~50km,也可以用于覆盖一个大学、园区等。

传统的城域网有着专门的标准,如分布式队列双总线(Distributed Queue Dual Bus, DQDB)、光纤分布式数据接口(Fiber Distributed Data Interface, FDDI)等,但随着SDH的快速发展,不少城域范围的网络也采用了SDH技术,已经模糊了城域网的概念。

3. 局域网

局域网(Local Area Network, LAN)是在较小地理范围(1km以内,甚至几十米)内将结点(如主机)连接起来而形成的网络,运用广泛。局域网是结构复杂程度最低的网络,主要目的是连接计算机,为最终的用户服务。

局域网产生了很多的产品,但经过几十年的发展,多数已经被淘汰了,目前的有线局域网主要是以太网,无线局域网则是WiFi。

4. 接入网

接入网(Access Network, AN)是局域网和互联网的中介,包括用户终端到互

联网之间的所有设备,接入网长度一般为几百米到几千米,被形象地称为最后一千米问题。接入网也可以分为有线接入网和无线接入网。有线接入网又可以分为铜线接入网、光纤接入网和光纤同轴电缆混合接入网等。

有线接入方式在多数情况下信号传输质量好,相关通信协议可以较为简单,目前一个重要的技术是以太接入网,是以太网为了适应接入而进行的改造。

作为接入技术,都应该有用户认证的功能,也就是需要用户进行相关的登录过程,方便对用户上网进行计费和管理。

5. 需要注意的地方

首先,本书对范围的分类更倾向于针对物理网络,而不是经过互联的网络。

其次,网络用户连上互联网实际上往往经历了很多种网络,如首先经过局域网(以太网、WiFi等),然后经过接入网进行认证,经过授权后才能访问互联网,这时可能会通过若干广域网/城域网,最后,进入网络资源所在的局域网。就如同豪横公司为我们安排的、经过若干交通工具的旅游线路一样。

另外,还有一个常见的名词——校园网(Campus Networks,也称为园区网、企业网),通常是指一个大学或一个企业的内部网,主要特征是网络完全由一个机构来管理。校园网更强调业务性,并且根据园区大小的不同而覆盖范围差别很大,所以一般不把它单独归为一类。

3.2 拓扑结构

为什么常用一朵云来表示网络呢?其中一个因素是物理网络的多样性,结构上难以以一种统一的方式来描述。学术界用拓扑(topology)结构这个术语来描述网络的实现结构。

这种概念是一种基于抽象的方法,只是告诉我们一个网络的几何形状是什么样的,具体是如何连接和如何工作的,不同的网络有不同的实现方法,即便它们可能具有相同的拓扑结构。本书主要介绍基本的网络拓扑,实际的网络拓扑可以由不同的基本拓扑相互组合形成复杂的拓扑。

常见的拓扑有总线型、星状、树状、环状以及网状等。

拓扑中往往会采用线段表示结点之间的链路,这个链路可能是有线的,也可能是无线的。

1. 点到点拓扑

点到点的拓扑是最简单的结构,采用这种拓扑的网络在生活中比较少见,但可以作为两个网络的连接手段,例如,两个路由器之间使用一根串口线实现互连,这根串口线即形成了一个特殊的网络,如图3-1中虚线所圈部分。



图3-1 点到点拓扑

这种拓扑下的网络往往具有双向的通信链路(如具有两根数据线缆,一根负责发送,另一根负责接收),可以保证通信是双向的,这就是通信行业常说的**全双工方式**(通信双方可以

同时收发数据,类似于平时的电话)。有些网络中参与者是不能同时收发数据的,即所谓的**半双工方式**,类似于对讲机。还有一种通信方式是纯单向的,例如,用收音机收听广播,这种方式称为**单工方式**。

还可以把这种拓扑扩展为链型/线型拓扑,如图 3-2 所示。这种拓扑在一些广域网和无线网络中会用到。



图 3-2 链型拓扑

2. 总线型拓扑

总线型拓扑(见图 3-3)是一种重要的拓扑结构,曾经是一种非常常见的拓扑结构,现在在某些领域(如工控网络)中还承担着重要的角色。

总线型拓扑有一根线缆作为总线,是网络中主要的通信媒体,总线两端往往需要有终端电阻用来防止电磁波的反射,而各个通信结点通过分支线连接到主电缆上。总线型拓扑结构简单,需要较少的电缆。但是网络只要有一处线缆接触不良/断开,就会导致全网无法工作。



图 3-3 总线型拓扑

这种拓扑结构的通信链路通常只有一根,不采用特殊技术(如后面所讲的信道复用技术)的情况下,通信方式多数是半双工工作方式,任意时刻网络上只能有一对结点进行通信,且结点间无法做到同时发送和接收数据。

3. 星状

星状拓扑是当前最常见的一种网络拓扑,它将每个通信结点连接到一个中心设备,形成了以该设备为核心的星状结构。结点通过中心设备进行通信,任意两个结点间的通信只需两步。这种网络拓扑既便宜又易于工程安装。人们平时用 WiFi 上网,其基本的拓扑就是星状拓扑。

这种拓扑结构扩展非常方便,对于有线网络,中心设备有几个接口就可以连接几个结点,可以做到热插拔。对于无线网络则更加随意,任一个结点离开(包括宕机)和加入网络都不影响其他结点的通信。但是这种拓扑结构对中心设备的依赖很大,如果中心设备故障,整个网络将瘫痪。

这种拓扑的网络可以是全双工方式,也可以是半双工方式,依据中心设备的不同而不同。如果中心设备只支持半双工方式,则任意时刻网络上只能有一对结点进行通信,且结点无法做到同时发送和接收数据。如果设备支持全双工方式,则多对结点可以同时通信,且结点可以同时发送和接收数据,图 3-4 展示了两对结点同时通信的示意图。

4. 树状

目前常说的树状拓扑是星状拓扑结构的扩展,是把星状网络中的一个或多个结点替换为星状网络,形成多层网络,如图 3-5 所示,具有天然的分级/分层结构,可以把连接互联网的那一个中心设备设为根结点。

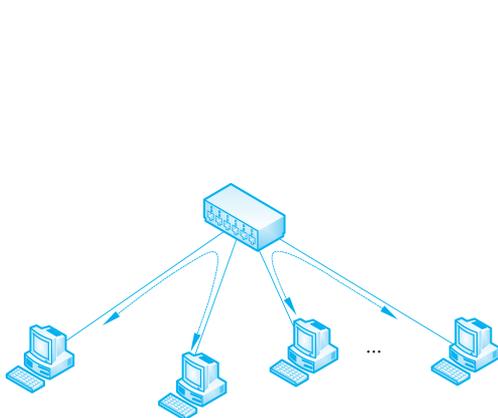


图 3-4 星状拓扑结构

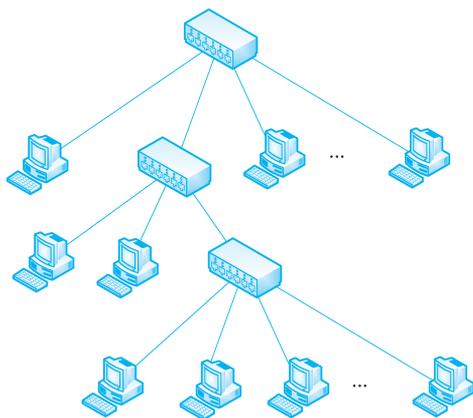


图 3-5 树状拓扑结构

这种拓扑极具扩展性,只要不超过通信距离的限制,可以级联多层,其他特性同星状拓扑。

这种拓扑的通信方式就更加复杂了,有可能是全部全双工工作方式,也有可能是全部半双工工作方式,还有可能一部分是全双工方式,其他的是半双工方式,依赖于中心设备的布置。

5. 环状

环状拓扑结构在局域网中基本上已经见不到了,但是在广域网/骨干网中,是一个常用的拓扑结构。

环状拓扑中,各结点通过相关设备与相邻的结点进行手拉手式的连接,从而形成一条首尾相连的闭合环状链路,如图 3-6 所示。数据一般沿着一个固定的方向(如逆时针)进行流动,依次经过环中各个结点,每个结点都核对信息的目的地址,从而判断是否是发送给自己的,只有接收结点才进行数据的接收。

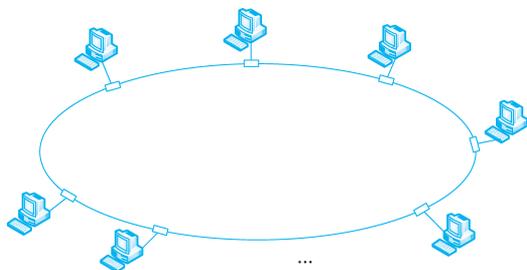


图 3-6 环状拓扑结构

一般情况下,网络中只传输一个发送者的信息,实现半双工通信方式。但是也可以通过特殊的技术实现多个结点同时收发信息。

这种拓扑结构的网络如果不采用故障恢复技术,某个结点的故障将导致整个网络的瘫痪。同样,也需要有特殊处理才能提高网络的可扩展性。

6. 网状

这种拓扑结构如图 3-7 所示,网络中任意两个结点可能存在多条通信的路径。

网状拓扑结构具有较高的可靠性,即使一条链路出现故障,数据也可以通过另一条路径

到达目的结点。但这种拓扑结构较为复杂,实现起来费用较高,不易管理和维护。

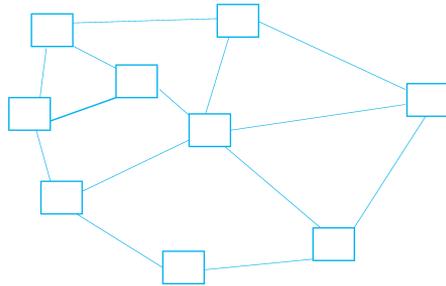


图 3-7 网状拓扑结构

3.3 多种网络互联的示例

图 3-8 给出了一个网络互联的假想示意图。

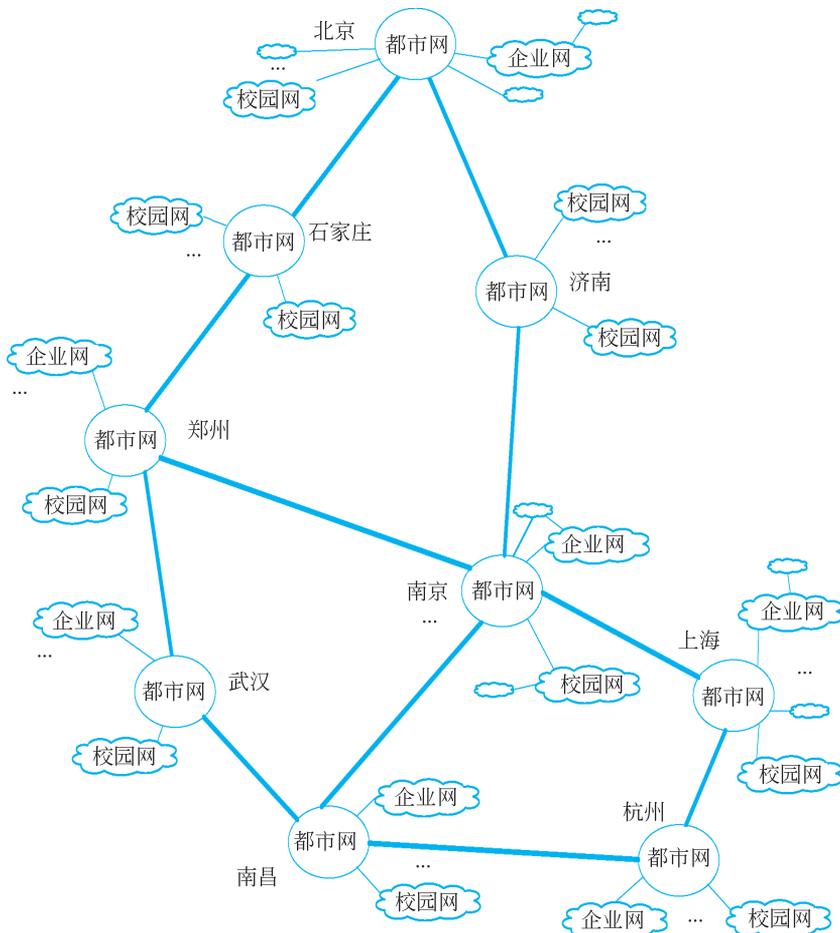


图 3-8 网络互联的假想示意图

1. 非面向用户的网络

城市之间必然以广域网联接,城市范围内以都市网联接,这两个网络都可以采用前面所提的SDH(同步数字体系)。这里假设城市之间采用了点到点拓扑,而城市内部采用了环状拓扑。

需要注意的是,读者不要认为全国就一个这样的结构,实际上每一个大的电信公司(如中国电信、中国移动、中国联通)都可以存在这样的网络结构。

上面两个网络一般不直接面对最终的用户,这种网络强调的是局与局之间的通信,带宽要大。

2. 面向用户的网络

各个单位都可以组建自己的网络为最终的用户服务,小的网络可以是一个局域网,而对于大的单位,单个局域网已经无法满足需求了,为此可以采用园区网的概念形成校园网、企业网。园区网可以采用广域网的技术(如SDH)联接众多的局域网。不管大小,这里不再细分,而是以云进行统一表示。

不管是什么网络,如果希望联入互联网,都需要接入网(可以简单地认为是园区网与都市网之间的一段线路和设备)的支持,从而联入本地电信部门所提供的都市网。

3. 互联网的现状

整个互联网,就是由各种网络为主体,存在多条路径相联的网状网拓扑结构。读者可以再次体会到,人们每天接触的互联网的本质是:网络的网络。

互联网已经成为世界上规模最大、增长速率最快的计算机网络,没有人能够准确说出互联网究竟有多大。

3.4 分组交换

网络联接完毕,就可以进行数据的传输了。通信领域在历史上存在着三类通信机制:电路交换、报文交换和分组交换。互联网采用了分组交换的技术来进行数据的传输。

3.4.1 什么是交换

不管是前面提到的路由器,还是电话局老式的程控交换机,都有一部分类似于图3-9的交换结构(switching fabric),它有很多的接口(interface),负责数据的接收和发送。在需要时,交换结构的核心部分可以通过内部通路把一对接口进行临时的连接,让数据从输入接口转移到输出接口,这就是所谓的交换。

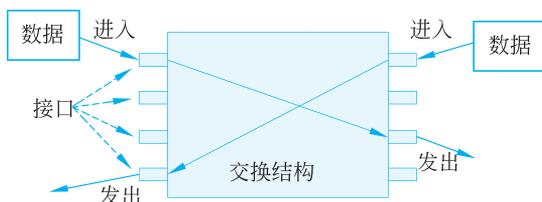


图 3-9 交换结构

交换结构的数据处理过程如下。

- (1) 数据从一个接口(输入接口)进入交换结构(就如同高铁进站)。

(2) 交换结构查表可知数据从哪一个接口(输出接口)发出可以更加接近目标,于是在两个接口间进行临时的连接,把数据由输入接口转发给输出接口(就如同火车站进行的轨道调度)。

(3) 数据从输出接口发送出去(就如同火车出站)。

现在大部分交换结构中可以实现并发性的交换和转发,如图 3-9 所示,有两对接口同时在进行转发。

3.4.2 为什么计算机网络不采用电路交换技术

1. 电路交换技术概述

电路交换很早就出现了,典型的产品就是电话网,曾经是世界上最大的通信网络。电路交换需要经历下面 3 个阶段。

(1) 建立连接:通过拨号建立一条专用的信道,以保证双方通话时所需的通信资源不被其他用户占用。

(2) 通信:双方互相通话。

(3) 释放连接:挂断电话,释放刚才使用的这条专用信道及相关资源。

这个过程在时间轴上的表现如图 3-10 所示。

3.4.1 节中的交换结构在电路交换中表现得有些不同,交换结构在建立连接时进行连接申请的转发,但是一旦建立了连接,交换结构中的连接是不断开的,直到通话双方挂断电话,才释放连接,这个“临时”有点长。

2. 独占性

电路交换是面向连接的服务的一个特殊实例,之所以特殊,是因为电路交换建立的是一条专用的物理通路,这条物理通路可以是一根电线,也可以是采用特殊技术而提供的一部分资源,它们具有的共同点是独占性,即只能通话双方使用。而计算机网络中使用的面向连接的服务都是共享通信资源的,没有独占性。

也正是因为独占性,为了支持更多人同时通话,电路交换不得不采用大量的物理资源。图 3-11 展示了瑞典斯德哥尔摩在 1890 年建设的巨型塔楼,连接了大约 5000 条线路,遍布整个城市的各个方向。

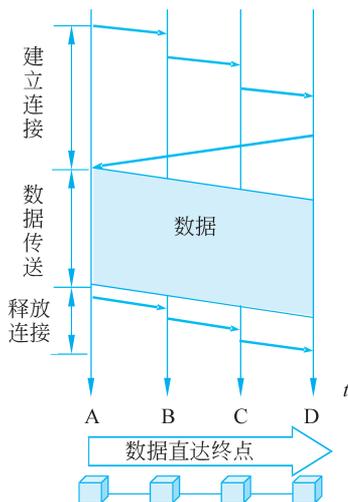


图 3-10 电路交换在时间轴上的表现

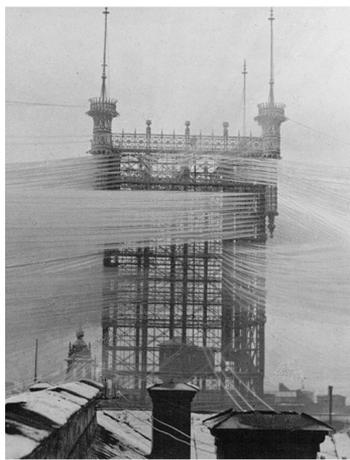


图 3-11 斯德哥尔摩的电信塔楼

这种通信有专门的资源保证,能够很好地支持通信质量需求、实时性好。但是这种技术,连接建立的时间较长,工作过程复杂,并且双方通信效率很低,大多数情况下非常浪费,除非双方正在吵架。

3. 不适合于计算机网络

计算机的通信和电路交换技术出发角度显然不同,计算机的数据大部分是不需要实时的,有一定的延迟问题不大。更重要的是,计算机的数据具有突发性,如果建设庞大的物理资源来应对庞大的突发性,那么在数据量少的时候,资源就太过浪费了(不建设又无法满足庞大人群同时上网)。因此必须提出一种基于共享物理资源的技术来进行计算机网络的建设和。

3.4.3 为什么计算机网络不采用报文交换技术

1. 报文交换技术概述

报文交换最初使用在电报业务上,事先不需要建立连接,如同发送信件一样,把信件送到邮箱/邮局即可。报文交换在时间轴上的表现如图 3-12 所示。

发送方 A 将用户的数据打包成报文 m(包含目的结点的地址),直接通过网络发给下一跳结点 B。B 做三件事情:

- (1) 将 m 在自己的缓存中存储下来,进行排队等候。
- (2) 等到 B 处理完毕 m 之前的报文后,从缓存中提取 m,根据 m 的目的地址,查表得出下一跳为 C。
- (3) 把数据转到连接 C 的接口上,并从该接口把 m 发出给 C。

这个过程就是重要的**存储转发**过程。

C 做同样的存储转发过程,把 m 转发给 D,D 获得 m 中的数据,完成 m 的传输过程。

2. 具有共享性

报文交换技术在发送数据前,不需要源和目的结点之间建立物理的连接,在发送过程中也不需要“霸占”整条路径的物理资源,只是逐段占用线路,并且使用完毕就放弃占用,其他结点可以立即使用,因此具备了共享性。

举例来讲(见图 3-13),S1 发送报文 m1 给 D1,S2 发送报文 m2 给 D2,假设 m1 先被处理并转发,在使用完 R1-R2 之间的链路后,开始使用 R2-R3 之间的链路,此时 m2 可以使用 R1-R2 之间的链路,两者可以同时发送。

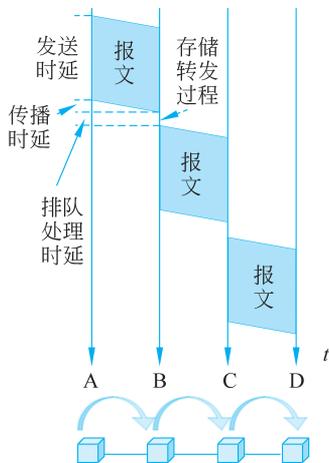


图 3-12 报文交换在时间轴上的表现

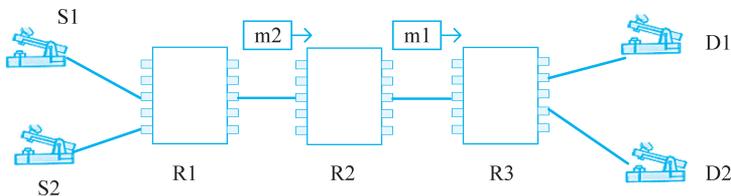


图 3-13 存储转发过程实现共享

3. 仍然不适合于计算机网络

既然报文交换已经可以利用存储转发实现线路共享了,为什么计算机网络不采用报文交换呢?

- 计算机网络传输的数据有大有小,如果数据较大,则数据从源结点发出到目的结点收到,整个过程时延太长。
- 如果数据较大,而交换结点缓存资源有限,两者是一种冲突。
- 如果数据较大,会长时间占用某一段链路,导致共享性下降,不利于公平性。
- 大的报文,如果其中出错 1b,整个报文就传输失败了。

为此,计算机网络在存储转发的机制上,增加了分组的思想。

3.4.4 分组交换技术

1. 概述

分组交换技术如图 3-14 所示。

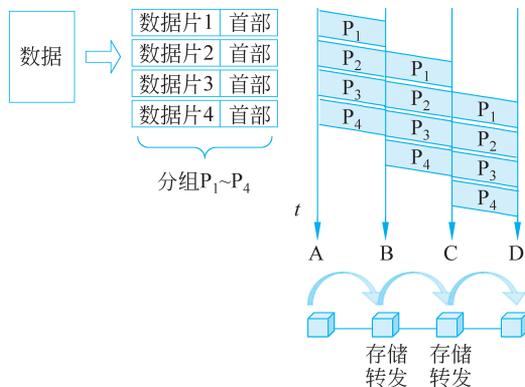


图 3-14 分组交换技术

首先,分组交换技术事先不需要建立连接,并沿用了存储转发的重要思想。

其次,分组交换技术事先将数据切割成数据片,加上首部(包括最重要的地址信息)形成分组,通过共享链路进行传输,每个分组都独立进行传输。

形成分组后,每一个分组每到达一个中间结点,都被执行存储转发的过程,从而一步一步靠近目标,最终到达接收方。

接收方在收到数据分组后,需要做的主要事情就是把这些分组拆掉首部,按照顺序进行组合,合并出原来的数据,交给用户。

2. 分组交换技术的特性

1) 分组可能经由不同的路径

由于分组之间是独立传输的,而计算机网络中源和目的结点之间的路径是可能变化的,甚至可能有多条路径并存的情况,所以分组可能会在不同的传输路径上进行传输,进而导致在到达目的结点的时候,可能是乱序的。

例如,本来从路径 a 走的分组,可能因为路径 a 上某个设备宕机,网络会感知这样的情况并重新计算路径,这时,分组就会在重新计算的路径 b 上进行传输。其实这也是设计网络时考虑的可靠性之一。

2) 尽最大努力

为了保证网络核心部分的简单性和高速性,分组在计算机网络的传输过程中是不保证可靠性的,分组可能丢失,也可能出错、乱序。计算机网络的一个工作原则是尽力就好。

只有当分组到了目的结点,目的结点才知道是否出了问题,这需要设计人员根据实际需要,自己决定在上层是否采用可靠的通信技术予以保证。

3) 提高了并行性

从图 3-14 可以看到,结点 B 可以在收到 P_1 后就立即转发 P_1 给 C,不必等到同一个数据的所有分组都被收齐后再进行转发(报文交换需要收齐某个报文的所有数据后才能进行转发),这样就大大增加了数据转发过程的并行性。

4) 进一步提高共享性和公平性

除了逐段占用链路(不始终“霸占”链路)外,分组交换因为分组的机制而具有了更高的共享性,如图 3-15 所示。A 和 C、B 和 D 同时进行通信,A 和 B 都把大的报文 m_A 和 m_B 分成了若干分组: $P_{A1}, P_{A2}, \dots, P_{An}$ 和 $P_{B1}, P_{B2}, \dots, P_{Bm}$ 。那么在 R1-R2 链路上可以采用公平策略来实现 $P_{A1}, P_{B1}, P_{A2}, P_{B2}, P_{A3}, P_{B3}, \dots$ 的顺序交叉发送分组,相比于报文交换技术(效果相当于 $P_{A1}, P_{A2}, \dots, P_{An}, P_{B1}, P_{B2}, \dots, P_{Bm}$)来说,共享性更高了。

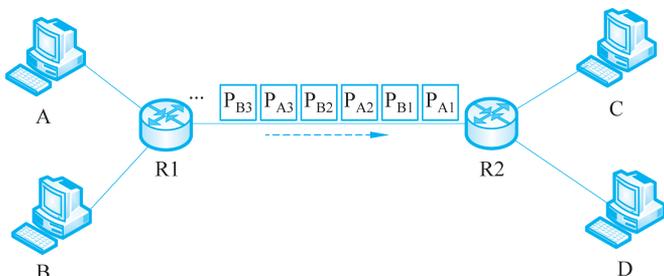


图 3-15 提高共享性示意图

5) 缺点

分组在各结点存储转发时需要排队和接受处理,这会造成一定的时延。这个时延也是计算机网络中占主要部分的时延。另外,每一个分组都必须携带的首部(里面有必不可少的控制信息)也造成了一定的额外开销。

3.5 连接网络的神器——路由器

3.5.1 概述



路由器概述

互联网上存在诸多类型的网络,这些网络的连接绝不是线缆连上即可(即便是电网,也需要有设备进行电压等的转变),负责连接不同网络的设备是路由器,它也是分组转发过程的主要场所。

1. 路由器的作用

路由器有以下三个主要的作用。

- 作为桥梁和翻译,连接不同的网络,形成庞大的互联网。
- 在连接多个网络的基础上计算路由,找出所有的、较好的通信路径,使得网络上任一

主机发送的信息可以被另一台任意主机所接收。这个过程就如同豪横公司为旅客规划旅游路线一样。

- 根据路由结果进行分组的转发,就是从源网络方向接收 IP 分组,通过路由器发往目的网络方向,每个路由器都这样处理分组,使分组能够向目的网络不断接近,最终到达目的网络。这个过程就如同不断转乘一样。

2. 路由器如何实现网络的互联

为了实现上面的作用,路由器需要满足以下前提。

- 所有的网络结点(包括主机和路由器)都使用同一种标准,这个标准就是 IP 协议。
- 路由器要“通晓”不同物理网络的“方言”,目前来说主要是第 1 层和第 2 层的协议。

可以这样类比,IP 协议是豪横公司定义的语言规则(普通话),不同的网络相当于不同的交通设施,其管理部门(A 和 B)有自己的方言。在旅客转乘时,假设的中转服务人员(相当于路由器)必须能听懂 A 的方言,将其转换成普通话,进行相关操作后,以 B 能听懂的方言叙述旅客的下一站旅程事宜。

再来回顾一下 TCP/IP 的体系结构,如图 3-16 所示。

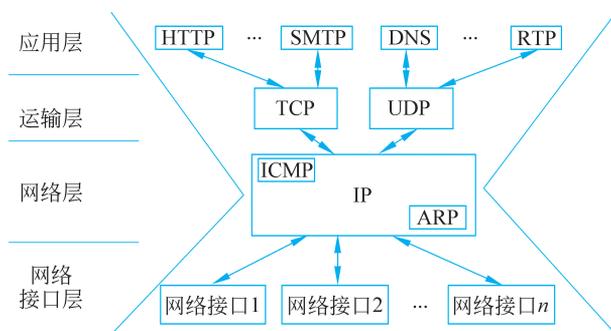


图 3-16 TCP/IP 的体系结构

可以发现,TCP/IP 的体系结构的特点是上下大而中间小,上层的各种协议都向下汇聚到 IP 协议上(everything over IP),而网络接口层的各种协议都需要支持 IP 协议(IP over everything,实际上是各种物理网络的网络接口可以被 IP 所使用)。

IP over everything 正是路由器实现对不同物理网络进行互联的基础。因此,路由器必须实现 TCP/IP 体系的中(IP 协议)、下(多种网络接口)层的内容。

3. 路由器的分类

依据路由器工作的侧重点不同,可以将路由器分为以下几类。

- 接入路由器:使得家庭和小型企业连接到互联网服务提供者(ISP),是一般用户最常见的设备。
- 企业级路由器:可以连接校园或企业内部很多的网络,组成较为庞大的园区网。企业级路由器越来越强大,很多自带防火墙、网络管理等功能。
- 骨干路由器:通常不是为用户直接服务的,主要用来连接长距离骨干网,要求路由器能对少数链路进行高速的数据转发。

早期的企业级、骨干路由器由国外公司(如思科 Cisco)把持,非常昂贵,目前国内公司(如华为等)已占有了很多市场,并把价格拉下来(几千元,甚至更便宜)。

4. 网络和路由器与人类交通的类比

高铁作为城市之间的交通,要求距离远、速度快,可以类比为广域网,满足把若干城市联接起来的任务。

公交、地铁作为城市内的交通,可以把高铁旅客运送到城市内各个方向,可以类比为城域网。高铁和高铁之间、高铁和公交/地铁之间的转换需要一个特殊的地方(如高铁站)进行人流的中转,相当于高性能的骨干路由器。

各个区、大的企业、学校、公司内的共享单车,可以让人员抵达最近的目的地,方便灵活,可以类比为局域网。公交车、地铁的站台这些中转的地方相当于一般的路由器(企业路由器、接入路由器)。

3.5.2 路由器的基本结构

路由器是一种具有多个输入、输出接口的专用计算机。图 3-17 展示了路由器的基本构造。可以看到,路由器从大的方面看可以分为上下两个层次。

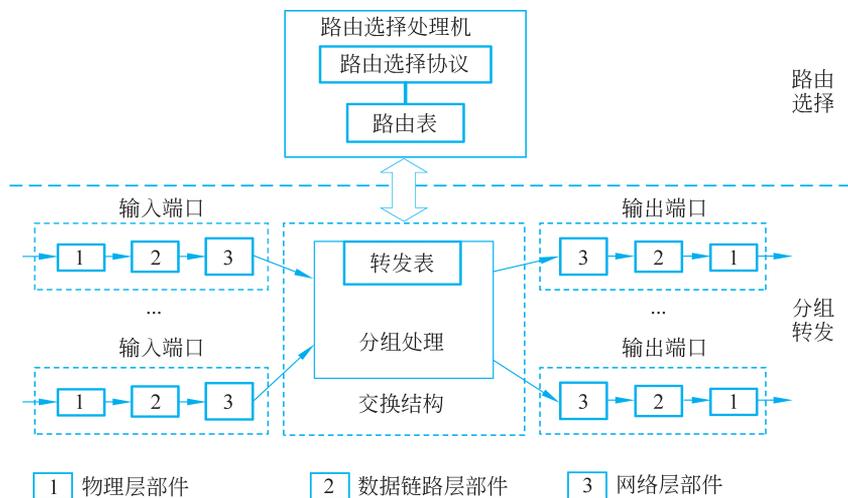


图 3-17 路由器的基本构造

- 路由选择(routing)部分：搜集网络拓扑信息,动态地计算路由(构造路径)。这一部分相当于豪横公司各中转站中的旅途规划办公室,他们之间相互交流道路的情况,计算出旅客实际的出游路线。
- 分组转发部分：将收到的 IP 分组依据路由表给出的方向,交换到指定的输出接口。这一部分相当于豪横公司各中转站中的中转服务人员。

1. 路由选择部分

1) 功能

这一层的核心是路由选择处理机,按照分布式算法经常或定期地和其他路由器交换路由信息、使用指定的路由选择算法计算和构造出路由表(routing table)。路由器需要不断地重复这个过程,防止路由信息过时。

分布式是网络中常见的一个名词,简单地说,就是很多实体分布在网络各处,共同运行,一起协作以完成同一件事情。

很显然,选择通畅快捷的路径(如近路),能够大幅地提高通信的速度,节约网络系统资源,从而让网络系统发挥出更大的效益。因此路由器的路由选择算法是十分重要的,这一部分内容将在后续章节介绍。

2) 路由表

路由表包含很多路由表项,每个表项最少包括<目的网络地址、下一跳>这样的信息,其中下一跳可以理解为后继的传送方向。路由器收到分组时,可以获得目的地址(地址中包括网络地址和主机地址两个信息),从中求出分组的网络地址,只根据该地址进行查表和转发。

只保存目的网络地址,而忽略主机的信息,使得路由表中的表项数目大幅减少,一方面减少了缓存的压力,另一方面可以提高查找的速度。就如同高铁系统只记录火车站的地址,而不必记录乘客的详细地址一样。

2. 分组转发部分

1) 抽取转发表

分组转发部分首先需要从路由表抽取出转发表(forwarding table,相当于电话号码簿),最少包括<目的网络地址、下一跳>信息,分组处理时可以据此得到后续向哪一个方向进行转发。多数人并不区分转发表和路由表。

2) 接口

路由器一般都要连接多个网络,而连接一个网络只需要一个接口,这些接口都是双向的,既可作为输入接口,又可作为输出接口。图 3-17 中标注的输入、输出接口只是为了方便介绍而已。

路由器的接口都具有物理层、数据链路层和网络层这三层处理模块。接口类型可以不同,例如,串口、以太网接口、广域网接口等,其不同主要体现在物理层和数据链路层。

(1) 输入接口。

作为输入接口时,物理层收到信号后,形成二进制数据流交给数据链路层。数据链路层的处理模块将二进制流形成数据帧,剥去帧首和尾部后可以得到分组(相当于豪横公司的中转服务人员听懂 A 的方言,整理成普通话),将其送到网络层。网络层的处理模块设有一个缓冲区(队列),来不及处理的 IP 分组暂时存放在这个队列中,等待后续送到交换结构进行转发。

输入接口的查表和转发功能是非常重要的,将待处理的 IP 分组进行分析,得到目的网络地址,查找转发表,依据转发表给出的方向把数据交换到指定的输出接口。

这一部分最重要的要求是要快,这样才能及时处理缓存中的分组,不至于让分组在缓存中堆积,影响后续的分组。为此采用了很多技术,例如,不断改进的交换结构,使用影子副本(每个输入接口复制一份转发表),使用二叉线索来提高查询性能等。

(2) 输出接口。

在输出接口中,网络层的处理模块也设有缓冲区,缓存那些需要发送出去的分组。数据链路层处理模块得到分组后,将分组加上数据链路层的帧首部和尾部(相当于豪横公司的中转服务人员把普通话改成 B 的方言),交给物理层后发送到外部线路。

很显然,接口工作涉及排队时延和处理时延,在发送过程中涉及发送时延。

路由器中的输入或输出队列可能因为某时刻分组的突发性到达,超出了队列的容量,会

产生溢出(部分分组因为无法进入相关队列而被抛弃),这种情况的出现是造成互联网上分组丢失的重要原因。

3. 交换结构

交换结构是分组转发部分中关键的部件,决定了路由器转发的性能。常用的交换方法有三种:通过存储器、通过总线、通过纵横交换结构。

1) 通过存储器

如图 3-18 所示,当路由器的某个输入接口收到一个分组时,将分组复制到存储器中。路由器查表后将分组复制到合适的输出接口缓存中。

这种方式显然效率较低,严重依赖于存储器的读写速度,每个分组要经历先写后读两次访问存储器的过程,不利于提高并行程度。

这种方式相当于一个小的旧式中转站,只有不多的中转服务人员为所有旅客服务。

2) 通过总线

如图 3-19 所示,分组从输入接口进入后,通过共享的总线传送到合适的输出接口。如果一个输入接口的分组需要通过总线,但是总线又正在被其他接口所占用的话,该接口只能等待。

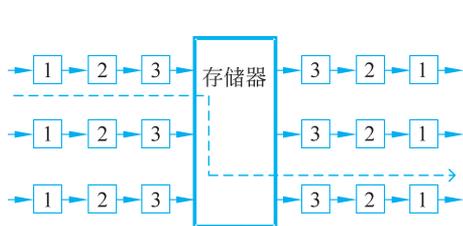


图 3-18 通过存储器进行交换

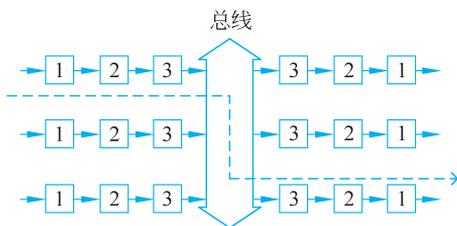


图 3-19 通过总线进行交换

这种方式传送一个分组的过程只需要经历一个步骤,效率比存储器方式有所提高,但因为总线是共享的,同一时间只能有一个分组在总线上进行传输,并行性还有待提升。另外,路由器的转发带宽受总线速率的限制,还可能受总线访问控制策略的影响。

在这种方式下,中转站比上面的方式多了更多的中转服务人员,但是很可惜,中转站硬件设施不太能体现豪横,只有一个通道。

3) 通过纵横交换结构

纵横交换结构又称为互连网络(interconnection network)。

路由器的输入接口与一个方向(如横向)的总线相连,输出接口与另一个方向(如纵向)的总线相连。如果没有分组转发时,横、纵向的总线之间是断开的,如果需要发送分组,则将相关的横、纵向总线相连(如图 3-20 中的黑点),实现输入、输出接口的直通。

这种方式可以极大地提高交换结构的并行性,实现多个输入接口与输出接口之间的并行传输。

在这种方式下,中转站比前面的方式豪横了很多,有更多的中转服务人员,通过很多的通道同时为旅客服务。

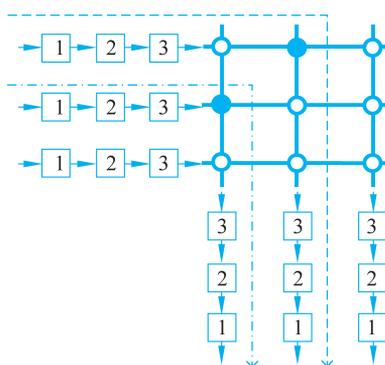


图 3-20 通过纵横交换结构

习 题

1. 按通信范围分,计算机网络分为哪几类?
2. 计算机网络有哪些基本的拓扑结构?
3. 路由器有哪些主要的作用?
4. 路由器在收到一个数据帧后是如何工作的?
5. 为什么计算机网络不采用电路交换技术?
6. 为什么计算机网络不采用报文交换技术?
7. 分组交换技术有什么优势?
8. 分析:为什么分组交换可以打破电路交换的独占信道的情况?

9. 源和目的结点间有 k 段链路(传播时延为 ds),带宽为 b (b/s)。电路交换时电路的建立需 ss 。分组交换时分组长度为 pb ,忽略排队处理等时延。在怎样的条件下,分组交换的时延比电路交换(包括建立连接的时间)的要小?

10. 结点 A 向 B 发送长度为 10^7 b 的报文,经过了两个路由器(3 段链路)。链路带宽为 2Mb/s。忽略传播、处理和排队时延。

(1) 如采用报文交换,从 A 把报文传送到 B 需要多少时间?

(2) 如采用分组交换,将报文分为 1000 个等长分组(忽略分组首部),则把 1000 个分组从 A 传到 B 需要多少时间?

11. [2013 研]主机甲通过 1 台路由器(存储转发方式)与主机乙互连,两段链路的数据传输速率均为 10Mb/s,主机甲分别采用报文交换和分组大小为 10Kb 的分组交换向主机乙发送 1 个大小为 8Mb($1\text{M}=10^6$)的报文。若忽略链路传播延迟、分组头部开销和分组拆装时间,则两种交换方式完成该报文传输所需的总时间分别为多少?

12. [2010 研]在如图 3-21 所示的采用“存储-转发”方式分组的交换网络中,所有链路的数据传输速率为 100Mb/s,分组大小为 1000B,其中,分组头部大小为 20B,若主机 H1 向主机 H2 发送一个大小为 980 000B 的文件,则在不考虑分组拆装时间和传播延迟的情况下,从 H1 发送到 H2 接收完为止,需要的时间至少是多少?

13. [2023 研]两段链路的数据传输速率为 100Mb/s,时延带宽积(即单向传播时延 \times 带宽)均为 1000b。如图 3-22 所示,若 H1 向 H2 发送 1 个大小为 1MB 的文件,分组长度为 1000B,则从 H1 开始发送时刻起到 H2 收到文件全部数据时刻止,所需的时间至少是(注:MB= 10^6 B)? ()

A. 80.02ms

B. 80.08ms

C. 80.09ms

D. 80.10ms

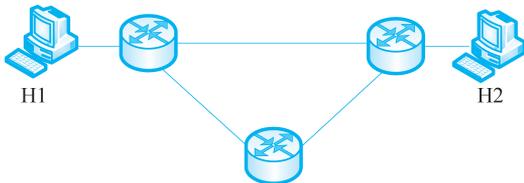


图 3-21 习题 12 用图

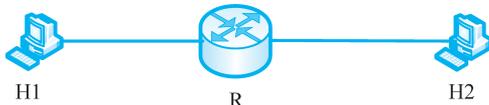


图 3-22 习题 13 用图

14. 试简述分组交换的要点。试从多个方面分析优缺点。