只要在这一条记录的网关列中填写接入互联网的路由器地址，当匹配不到其他路由时[①]，网络包就会被转发到互联网接入路由器。因此这条记录被称为默认路由，这一行配置的网关地址被称为默认网关。在计算机的TCP/IP设置窗口中也有一个填写默认网关的框，意思是一样的。计算机上也有一张和路由器一样的路由表，其中默认网关的地址就是我们在设置窗口中填写的地址。

路由表中子网掩码为0.0.0.0的记录表示“默认路由”。

这样一来，无论目标地址是表示一个子网还是表示某台设备，都可以用相同的方法查找出转发目标，而且也避免了不知道转发到哪里的问题。

3.3.6 包的有效期

从路由表中查找到转发目标之后，网络包就会被转交给输出端口，并最终发送出去，但在此之前，路由器还有一些工作要完成。

第一个工作是更新IP头部中的TTL（Time to Live，生存时间）字段（参见第2章的表2.2）。TTL字段表示包的有效期，包每经过一个路由器的转发，这个值就会减1，当这个值变成0时，就表示超过了有效期，这个包就会被丢弃。

这个机制是为了防止包在一个地方陷入死循环。如果路由表中的转发目标都配置正确，应该不会出现这样的情况，但如果其中的信息有问题，或者由于设备故障等原因切换到备用路由时导致暂时性的路由混乱，就会出现这样的情况。

发送方在发送包时会将TTL设为64或128，也就是说包经过这么多路由器后就会“寿终正寝”。现在的互联网即便访问一台位于地球另一侧的

① 由于匹配的比特数越长优先级越高（最长匹配原则），因此子网掩码为0.0.0.0的记录优先级是最低的，只有当找不到其他匹配的记录时，才会选择这条记录。

服务器，最多也只需要经过几十个路由器，因此只要包被正确转发，就可以在过期之前到达目的地。

3.3.7　通过分片功能拆分大网络包

路由器的端口并不只有以太网一种，也可以支持其他局域网或专线通信技术。不同的线路和局域网类型各自能传输的最大包长度也不同，因此输出端口的最大包长度可能会小于输入端口[①]。即便两个端口的最大包长度相同，也可能会因为添加了一些头部数据而导致包的实际长度发生变化，ADSL、FTTH 等宽带接入技术中使用的 PPPoE[②] 协议就属于这种情况。无论哪种情况，一旦转发的包长度超过了输出端口能传输的最大长度，就无法直接发送这个包了。

遇到这种情况，可以使用 IP 协议中定义的分片功能对包进行拆分，缩短每个包的长度。需要注意的是，这里说的分片和第 2 章介绍的 TCP 对数据进行拆分的机制是不同的。TCP 拆分数据的操作是在将数据装到包里之前进行的，换句话说，拆分好的一个数据块正好装进一个包里。从 IP 分片的角度来看，这样一个包其实是一个未拆分的整体，也就是说，分片是对一个完整的包再进行拆分的过程。

分片操作的过程如图 3.15 所示。首先，我们需要知道输出端口的 MTU[③]，看看这个包能不能不分片直接发送。最大包长度是由端口类型决定的，用这个最大长度减掉头部的长度就是 MTU，将 MTU 与要转发的包长度进行比较。如果输出端口的 MTU 足够大，那么就可以不分片直接发送；如果输出端口的 MTU 太小，那么就需要将包按照这个 MTU 进行分片，但

① 最大包长度是由各个通信规格定义的，如果包超过了这个最大长度就不符合相应的规格，也就不能传输了，因此输入端口收到的包不会超过最大长度。

② PPPoE：PPP over Ethernet。它是一种控制 ADSL、FTTH 等宽带网络的方式，4.3.2 一节将会对它进行介绍。

③ 一个包能传输的最大数据长度，2.3.1 节介绍过。

在此之前还需要看一下 IP 头部中的标志字段，确认是否可以分片[①]。

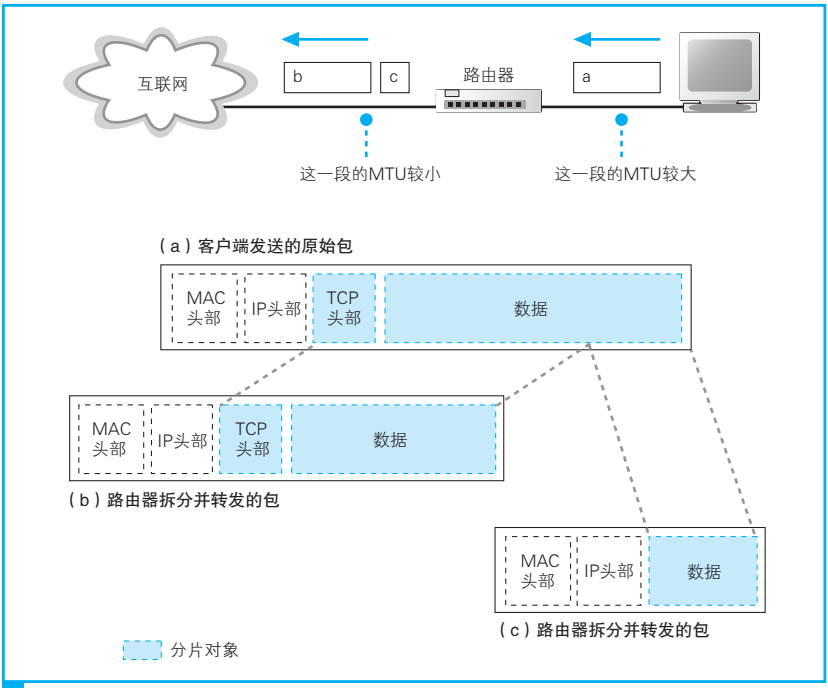

图 3.15　对包进行拆分的分片功能
尽管 TCP 头部不是用户数据，但从 IP 协议的角度来看它也是数据的一部分。

如果查询标志字段发现不能分片，那么就只能丢弃这个包，并通过 ICMP 消息通知发送方。否则，就可以按照输出端口 MTU 对数据进行依次拆分了。在分片中，TCP 头部及其后面的部分都是可分片的数据，尽管 TCP 头部不属于用户数据，但从 IP 来看也是 TCP 请求传输的数据的一部分。数据被拆分后，每一份数据前面会加上 IP 头部，其大部分内容都和原本的 IP 头部一模一样，但其中有部分字段需要更新，这些字段用于记录分片相关的信息。

① 一般来说都是可以分片的，但下面两种情况不能分片：1）发送方应用程序等设置了不允许分片；2）这个包已经是经过分片后的包。

3.3.8　路由器的发送操作和计算机相同

到这里，发送前的准备工作就完成了[①]，接下来就会进入包的发送操作。

这一步操作取决于输出端口的类型。如果是以太网端口，则按照以太网的规则将包转换为电信号发送出去；如果是 ADSL 则按照 ADSL 的规则来转换，以此类推。在家庭网络中，路由器后面一般连接 ADSL 等线路接入互联网，因此路由器会根据接入网的规则来发送包。不过，要理解具体的操作过程，需要先理解相应的通信线路[②]，比较复杂，因此我们留到下一章探索互联网内部时再讲解。这里，我们假设路由器位于公司等局域网的内部，即输出端口也是以太网，看看这种情况是如何操作的。

以太网的包发送操作是根据以太网规则来进行的，即便设备种类不同，规则也是相同的。也就是说，其基本过程和协议栈中的 IP 模块发送包的过程是相同的，即在包前面加上 MAC 头部，设置其中的一些字段，然后将完成的包转换成电信号并发送出去。下面来简单复习一下这个过程。

首先，为了判断 MAC 头部中的 MAC 地址应该填写什么值，我们需要根据路由表的网关列判断对方的地址。如果网关是一个 IP 地址，则这个 IP 地址就是我们要转发到的目标地址；如果网关为空[③]，则 IP 头部中的接收方 IP 地址就是要转发到的目标地址。知道对方的 IP 地址之后，接下来需要通过 ARP[④] 根据 IP 地址查询 MAC 地址，并将查询的结果作为接收方

① 实际上还有一项工作。IP 头部中有一个用于错误检验的字段“校验和”，在路由器更新 TTL 和分片的过程中，IP 头部的内容发生了改变，因此必须重新计算校验和。这里之所以没有详细讲解这个过程，是因为和以太网以及通信线路本身的错误校验机制相比，IP 校验和的可靠性很低，因此大多数路由器都不去校验这个值，就当它不存在一样。

② ADSL 等通信线路会在下一章介绍。

③ 第 2 章我们讲过网关的 IP 地址和接口的 IP 地址相同时，表示 IP 头部中的接收方 IP 地址就是我们要转发的直接目标，但这段内容是针对 Windows 计算机的。路由器和 Windows 不一样，当包可以直接发送到最终接收方时，一般网关列是留空的。

④ ARP 是根据 IP 地址查询 MAC 地址的协议，2.5.5 节介绍过。

MAC 地址。路由器也有 ARP 缓存，因此首先会在 ARP 缓存中查询，如果找不到则发送 ARP 查询请求。

路由器判断下一个转发目标的方法如下。

- 如果路由表的网关列内容为 IP 地址，则该地址就是下一个转发目标。
- 如果路由表的网关列内容为空，则 IP 头部中的接收方 IP 地址就是下一个转发目标。

路由器也会使用 ARP 来查询下一个转发目标的 MAC 地址。

接下来是发送方 MAC 地址字段，这里填写输出端口的 MAC 地址[①]。还有一个以太类型字段，填写 0080（十六进制）。

网络包完成后，接下来会将其转换成电信号并通过端口发送出去。这一步的工作过程和计算机也是相同的。例如，当以太网工作在半双工模式时，需要先确认线路中没有其他信号后才能发送，如果检测到碰撞，则需要等待一段时间后重发。如果以太网工作在全双工模式，则不需要确认线路中的信号，可以直接发送。

如果输出端口为以太网，则发送出去的网络包会通过交换机到达下一个路由器。由于接收方 MAC 地址就是下一个路由器的地址，所以交换机会根据这一地址将包传输到下一个路由器。接下来，下一个路由器会将包转发给再下一个路由器，经过层层转发之后，网络包就到达了最终的目的地。

3.3.9 路由器与交换机的关系

关于路由器的基本工作，也就是包转发，到这里就全部讲完了，下面来整理一下路由器与交换机的关系。

① 端口的 MAC 地址一般也是在硬件生产过程中写入 ROM 中的。

要理解两者之间的关系，关键点在于计算机在发送网络包时，或者是路由器在转发网络包时，都需要在前面加上 MAC 头部。之前的讲解都是说在开头加上 MAC 头部，如果看图 3.16 大家可以发现，准确的说法应该是将 IP 包装进以太网包的数据部分中。也就是说，给包加上 MAC 头部并发送，从本质上说是将 IP 包装进以太网包的数据部分中，委托以太网去传输这些数据。IP 协议本身没有传输包的功能，因此包的实际传输要委托以太网来进行。路由器是基于 IP 设计的，而交换机是基于以太网设计的，因此 IP 与以太网的关系也就是路由器与交换机的关系。换句话说，路由器将包的传输工作委托给交换机来进行[①]。当然，这里讲的内容只适用于原原本本实现 IP 和以太网机制的纯粹的路由器和交换机，实际的路由器有内置交换机功能的，比如用于连接互联网的家用路由器就属于这一种，对于这种路由器，上面内容可能就不适用了。但是，如果把这种“不纯粹”的路由器拆分成“纯粹”的路由器和“纯粹”的交换机，则它们各自都适用上面的内容。

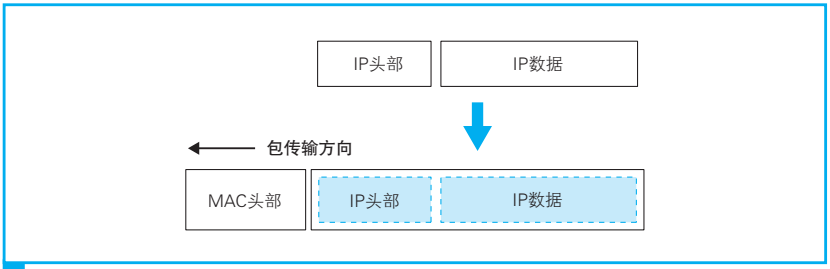

图 3.16　将 IP 包装进以太网包的数据部分

从包的转发目标也可以看出路由器和交换机之间的委托关系。IP 并不是委托以太网将包传输到最终目的地，而是传输到下一个路由器。在创建 MAC 头部时，也是从 IP 的路由表中查找出下一个路由器的 IP 地址，并通

① 除了使用交换机，还可以使用集线器，或者用交叉双绞线直接连接到路由器端口都可以。关键是，在委托传输时，只要能按照以太网规则传输包，不管是什么样的设备都可以。

过 ARP 查询出 MAC 地址，然后将 MAC 地址写入 MAC 头部中的，这表示 IP 对以太网的委托只是将包传输到下一个路由器就行了。当包到达下一个路由器后，下一个路由器又会重新委托以太网将包传输到再下一个路由器。随着这一过程反复执行，包就会最终到达 IP 的目的地，也就是通信的对象。

到这里我们已经梳理了路由器与交换机之间的关系。简单来说，IP（路由器）负责将包发送给通信对象这一整体过程，而其中将包传输到下一个路由器的过程则是由以太网（交换机）来负责的。

当然，网络并非只有以太网一种，还有无线局域网，以及接入互联网的通信线路，它们和 IP 之间的关系又是什么样的呢？其实只要将以太网替换成无线局域网、互联网线路等通信规格就可以了。也就是说，如果和下一个路由器之间是通过无线局域网连接的，那么就委托无线局域网将包传输过去；如果是通过互联网线路连接的，那么就委托它将包传输过去。除了这里列举的例子之外，世界上还有很多其他类型的通信技术，它们之间的关系也是一样的，都是委托所使用的通信技术将包传输过去。

IP 本身不负责包的传输，而是委托各种通信技术将包传输到下一个路由器，这样的设计是有重要意义的，即可以根据需要灵活运用各种通信技术，这也是 IP 的最大特点。正是有了这一特点，我们才能够构建出互联网这一规模巨大的网络。

IP（路由器）负责将包送达通信对象这一整体过程，而其中将包传输到下一个路由器的过程则是由以太网（交换机）来负责的。

3.4 路由器的附加功能

3.4.1 通过地址转换有效利用 IP 地址

刚才我们介绍了路由器的基本工作过程，现在的路由器除了这些基本功能之外，还有一些附加功能。下面来介绍两种最重要的功能——地址转换和包过滤。

首先，我们先了解一下地址转换功能出现的背景。所谓地址，就是用来识别每一台设备的标志，因此每台设备都应该有一个唯一不重复的地址，就好像如果很多人的地址都一样，那么快递员就不知道该把包裹送给谁了。网络也是一样，本来互联网中所有的设备都应该有自己的固定地址，而且最早也确实是这样做的。比如，公司内网需要接入互联网的时候，应该向地址管理机构申请 IP 地址，并将它们分配给公司里的每台设备。换句话说，那个时候没有内网和外网的区别，所有客户端都是直接连接到互联网的。

尽管互联网原本是这样设计的，但进入 20 世纪 90 年代之后，互联网逐步向公众普及，接入互联网的设备数量也快速增长，如此一来，情况就发生了变化。如果还用原来的方法接入，过不了多久，可分配的地址就用光了。如果不能保证每台设备有唯一不重复的地址，就会从根本上影响网络包的传输，这是一个非常严重的问题。如果任由这样发展下去，不久的将来，一旦固定地址用光，新的设备就无法接入了，互联网也就无法继续发展了。

解决这个问题的关键在于固定地址的分配方式。举个例子，假如有 A、B 两家公司，它们的内网是完全独立的。这种情况下，两家公司的内网之间不会有网络包流动，即使 A 公司的某台服务器和 B 公司的某台客户端具有相同的 IP 地址也没关系，因为它们之间不会进行通信。只要在每家公司自己的范围内，能够明确判断网络包的目的地就可以了，是否和其他公司的内网地址重复无关紧要，只要每个公司的网络是相互独立的，就不会出现问题。

解决地址不足的问题，利用的就是这样的性质，即公司内部设备的地址不一定要和其他公司不重复。这样一来，公司内部设备就不需要分配固定地址了，从而大幅节省了 IP 地址。当然，就算是公司内网，也不是可以随便分配地址的，因此需要设置一定的规则，规定某些地址是用于内网的，这些地址叫作私有地址，而原来的固定地址则叫作公有地址[①]。

私有地址的规则其实并不复杂，在内网中可用作私有地址的范围仅限以下这些。

10.0.0.0 ~ 10.255.255.255
172.16.0.0 ~ 172.31.255.255
192.168.0.0 ~ 192.168.255.255

在制定私有地址规则时，这些地址属于公有地址中还没有分配的范围。换句话说，私有地址本身并没有什么特别的结构，只不过是将公有地址中没分配的一部分拿出来规定只能在内网使用它们而已。这个范围中的地址和其他公司重复也没关系，所以对于这些地址不作统一管理，不需要申请，任何人都可以自由使用。当然，如果在公司内部地址有重复就无法传输网络包了，因此必须避免在内网中出现重复的地址。

尽管这样的确能节省一部分地址，但仅凭这一点还无法完全解决问题。公司内网并不是完全独立的，而是需要通过互联网和其他很多公司相连接，所以当内网和互联网之间需要传输包的时候，问题就出现了，因为如果很多地方都出现相同的地址，包就无法正确传输了。

于是，当公司内网和互联网连接的时候，需要采用图 3.17 这样的结构，即将公司内网分成两个部分，一部分是对互联网开放的服务器，另一部分是公司内部设备。其中对互联网开放的部分分配公有地址，可以和互联网直接进行通信，这一部分和之前介绍的内容是一样的。相对地，内网部分则分配私有地址，内网中的设备不能和互联网直接收发网络包，而是通过一种特别的机制进行连接，这个机制就叫地址转换。

① 在互联网规格中写作 Globally Unique Address 或者 Public Address。

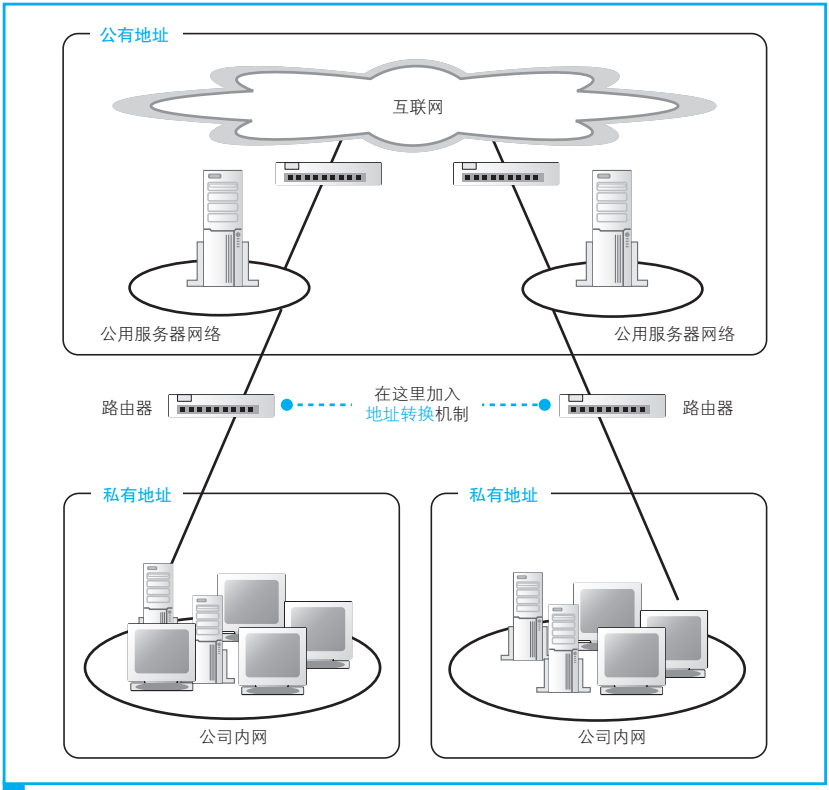

图 3.17　私有地址和公有地址分别管理

3.4.2　地址转换的基本原理

地址转换的基本原理是在转发网络包时对 IP 头部中的 IP 地址和端口号[①]进行改写。具体的过程我们来看一个实际的例子，假设现在要访问 Web 服务器，看看包是如何传输的。

首先，TCP 连接操作的第一个包被转发到互联网时，会像图 3.18 这样，将发送方 IP 地址从私有地址改写成公有地址。这里使用的公有地址是

① 这里的端口号指的是 TCP 和 UDP 的端口号，不是路由器和集线器连接网线的那个端口。

地址转换设备[①]的互联网接入端口的地址。与此同时，端口号也需要进行改写，地址转换设备会随机选择一个空闲的端口。然后，改写前的私有地址和端口号，以及改写后的公有地址和端口号，会作为一组相对应的记录保存在地址转换设备内部的一张表中。

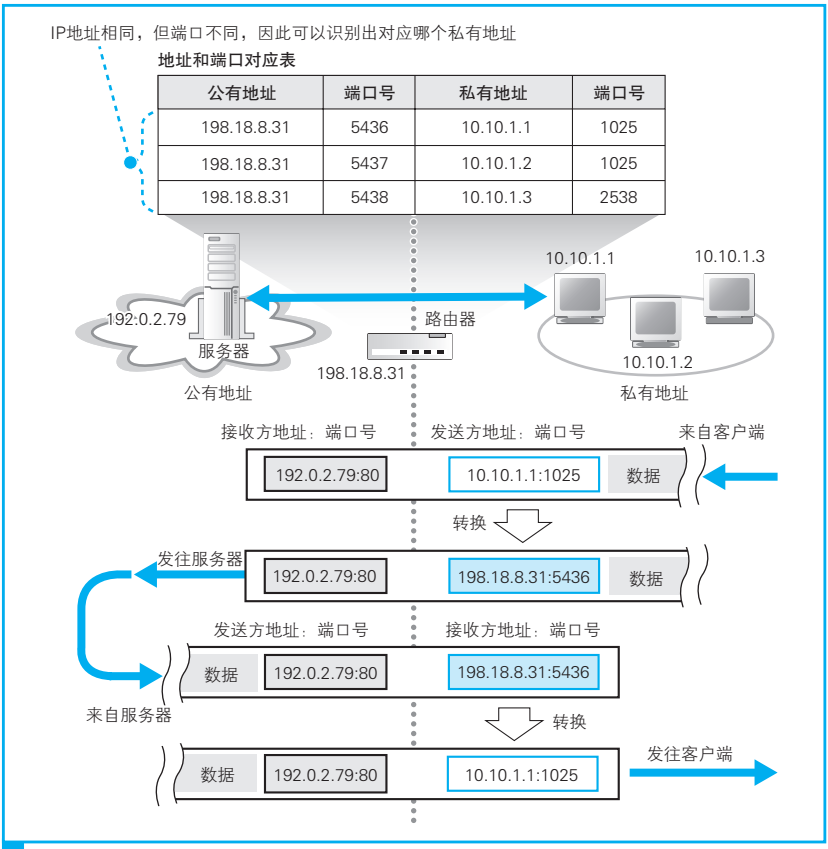

图 3.18　利用端口号改写 IP 地址
在对外只能使用一个公有地址的情况下，可以用不同的端口号来区别内网中的不同终端。

① 具备地址转换功能的设备不仅有路由器，有些防火墙也有地址转换功能，它的工作方式和路由器是相同的，因此这里我们虽然用了地址转换设备这个词，但在这里的上下文中指的就是路由器。

改写发送方 IP 地址和端口号之后，包就被发往互联网，最终到达服务器，然后服务器会返回一个包。服务器返回的包的接收包是原始包的发送方，因此返回的包的接收方就是改写后的公有地址和端口号。这个公有地址其实是地址转换设备的地址，因此这个返回包就会到达地址转换设备。

接下来，地址转换设备会从地址对应表中通过公有地址和端口号找到相对应的私有地址和端口号，并改写接收方信息，然后将包发给公司内网，这样包就能够到达原始的发送方了。

在后面的包收发过程中，地址转换设备需要根据对应表查找私有地址和公有地址的对应关系，再改写地址和端口号之后进行转发。当数据收发结束，进入断开阶段，访问互联网的操作全部完成后，对应表中的记录就会被删除。

通过这样的机制，具有私有地址的设备就也可以访问互联网了。从互联网一端来看，实际的通信对象是地址转换设备（这里指的是路由器）。

上面是以公司内网为例来进行介绍的，家庭网络中的工作过程也是完全相同的，只是规模不同而已。

3.4.3　改写端口号的原因

现在我们使用的地址转换机制是同时改写地址和端口号的，但早期的地址转换机制是只改写地址，不改写端口号的。用这种方法也可以让公司内网和互联网进行通信，而且这种方法更简单。

但是，使用这种方法的前提是私有地址和公有地址必须一一对应，也就是说，有多少台设备要上互联网，就需要多少个公有地址。当然，访问动作结束后可以删除对应表中的记录，这时同一个公有地址可以分配给其他设备使用，因此只要让公有地址的数量等于同时访问互联网的设备数量就可以了。然而公司人数一多，同时访问互联网的人数也会增加。一个几千人的公司里，有几百人同时访问互联网是很正常的，这样就需要几百个公有地址。

改写端口号正是为了解决这个问题。客户端一方的端口号本来就是从

空闲端口中随机选择的，因此改写了也不会有问题。端口号是一个 16 比特的数值，总共可以分配出几万个端口[①]，因此如果用公有地址加上端口的组合对应一个私有地址，一个公有地址就可以对应几万个私有地址，这种方法提高了公有地址的利用率。

3.4.4 从互联网访问公司内网

对于从公司内网访问互联网的包，即便其发送方私有地址和端口号没有保存在对应表中也是可以正常转发的，因为用来改写的公有地址就是地址转换设备自身的地址，而端口号只要随便选一个空闲的端口就可以了，这些都可以由地址转换设备自行判断。然而，对于从互联网访问公司内网的包，如果在对应表中没有记录就无法正常转发。因为如果对应表中没有记录，就意味着地址转换设备无法判断公有地址与私有地址之间的对应关系。

换个角度来看，这意味着对于没有在访问互联网的内网设备，是无法从互联网向其发送网络包的。而且即便是正在访问的设备，也只能向和互联网通信中使用的那个端口发送网络包，无法向其他端口发送包。也就是说，除非公司主动允许，否则是无法从互联网向公司内网发送网络包的。这种机制具有防止非法入侵的效果。

不过，有时候我们希望能够从互联网访问公司内网，这需要进行一些设置才能实现。之所以无法从互联网访问内网，是因为对应表里没有相应的记录，那么我们只要事先手动添加这样的记录就可以了（图 3.19）。一般来说，用于外网访问的服务器可以放在地址转换设备的外面并为它分配一个公有地址，也可以将服务器的私有地址手动添加到地址转换设备中，这样就可以从互联网访问到这台具有私有地址的服务器了[②]。

---

① 16 比特可以表示 65 536 个端口号，但并不是所有这些端口都可以用于地址转换。

② 这种配置中，需要将地址转换设备的公有地址添加到 DNS 服务器中。

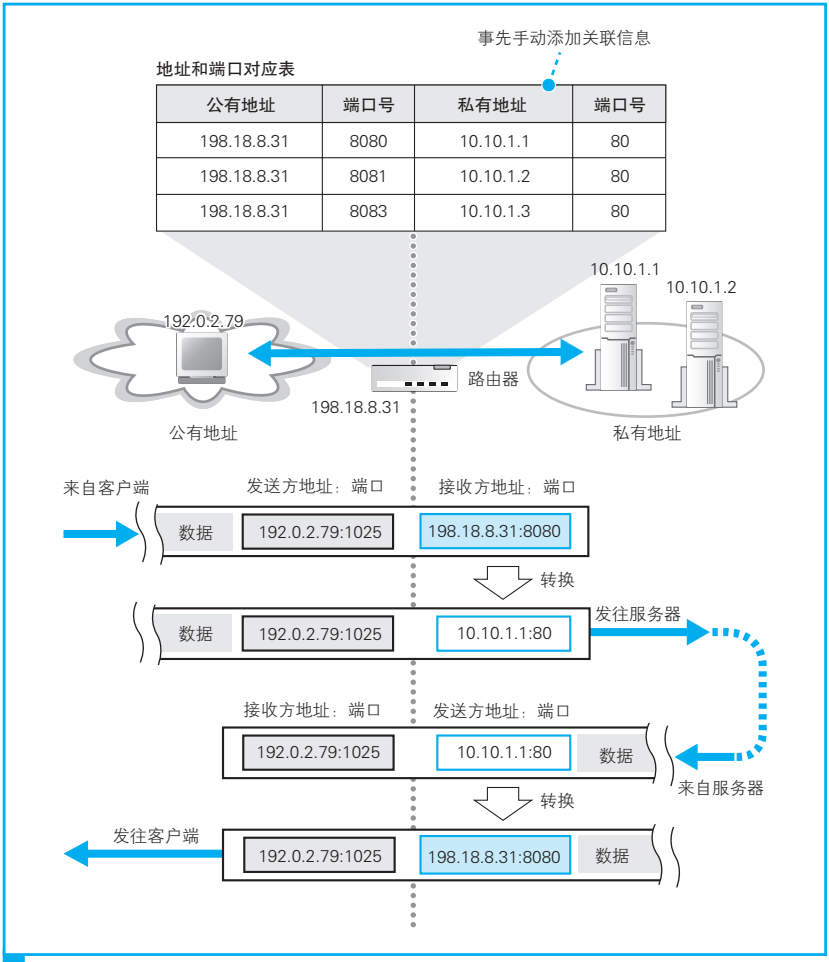

图 3.19　从互联网访问公司内网
只要事先将地址和端口的关联信息添加到地址转换设备的对应表中，就可以从互联网访问内网中的设备了。

3.4.5　路由器的包过滤功能

下面来介绍一下包过滤功能。包过滤也是路由器的一个重要附加功能，刚才的地址转换看起来有点复杂，不过包过滤的机制并不复杂。包过滤就

是在对包进行转发时，根据 MAC 头部、IP 头部、TCP 头部的内容[①]，按照事先设置好的规则决定是转发这个包，还是丢弃这个包。我们通常说的防火墙设备或软件，大多数都是利用这一机制来防止非法入侵的[②]。

包过滤的原理非常简单，但要想设置一套恰当的规则来区分非法访问和正常访问，只阻止非法入侵而不影响正常访问，是非常不容易的。举个例子，为了防止从互联网非法入侵内网，我们可以将来自互联网的所有包都屏蔽掉，但是这会造成什么结果呢？正如我们第 2 章介绍过的 TCP 的工作过程一样，网络包是双向传输的，如果简单地阻止来自互联网的全部包，那么从内网访问互联网的操作也会无法正常进行。

这个话题其实非常有趣，由于包过滤的使用方法和服务器的工作相关，所以我们在探索服务器时再详细介绍吧。

当网络包通过互联网接入路由器之后，它终于要进入互联网内部了，下一章将对这一部分进行探索。

小测验

本章的旅程告一段落，我们为大家准备了一些小测验题目，确认一下自己的成果吧。

问题

1. 局域网中使用的双绞线中为什么要将信号线缠绕在一起？
2. 将输入的信号广播到所有端口上的设备是交换机还是集线器？
3. 用来指定网络号和主机号比特数的值叫什么？
4. 将大网络包进行拆分的功能叫什么？
5. 路由器的路由表中有时可以看到子网掩码为 0.0.0.0 的记录，这代表什么意思？

---

① 也有些设备可以根据 TCP 头部后面的数据内容设置过滤规则，不过一般不太常见。

② 也有一些防火墙是用其他机制来防止非法入侵的。