第3章 IPv4的问题

本章主要讨论IPv4中存在的问题。虽然IPv4已经取得了令人难以置信的成功，但是仍有一些值得改进的地方。其中最显眼和最值得注意的可改进之处在于其地址空间的大小。其他议题与性能及IP头字段的设计和使用相关。本章还将讨论安全性、性能和管理控制等议题。

3.1 修改还是替换

考虑到IPv4存在的时间，它确实工作得不错。那为什么还要用其他的东西来替换它呢？毕竟如果把IPv4替换掉的话，网络中的所有系统均需要升级。升级到最新的微软 Windows易如闲庭信步，但IPv4的升级对于大型组织来说，简直就是一场恶梦。我们讨论的网络可能包括十亿甚至更多的遍布全球的系统，上面运行着不知道多少种不同版本的 TCP/IP连网软件、操作系统和硬件平台。要求对其中所有系统同时进行升级是不可想象的。

那么有没有办法可以避免 IP升级可能带来的纷乱和不幸呢？答案是也许有，也许没有。这取决于对新协议的需求程度。换句话说，如果协议的唯一问题仅仅在于地址的匮乏，通过使用诸如后面所讨论的划分子网、网络地址翻译或无类域内选路等现有工具和技术，也许可以使该协议在相当长的时间内仍可继续工作。但是，这种权宜之计不可能长期有效，实际上，这些技术已经使用了很多年，如果不实现对 IP的升级，它们最终将阻碍未来 Internet的发展，因为它们限制了可连接的网络数和主机数。

本章还将讨论IPv4的其他问题，除了地址缺乏的问题外，还包括更普遍的扩展性问题、管理问题、选路困难、服务的改进和服务质量特性的交付以及安全性问题。

最后，拥有多年IPv4工作经验的工程师们作出的决定是替换而不是修补 IPv4。我们知道IPv4中哪些工作良好，哪些只是可以工作，哪些可以工作得更好。现在的情形不是用未知量来取代已知量。IPv6的设计者们将这个新协议建立在 IPv4的基础上，沿用IPv4工作良好的部分，改进可以工作的部分，去掉影响性能和功能的部分，另外还增加了当前特别需要的功能。

本节的其余部分讨论目前用于解决IPv4缺点的一些方法，然后讨论IPv4向IPv6升级的协议过渡的含义。

协议的补丁和扩展

IPv4面临的最紧迫的问题是地址空间的大小问题，主要研究方向也定位在如何减少地址空间的浪费并提高使用效率上。其他议题，包括选路、网络管理、配置及 IPv4扩展选项有时也与地址空间有关。

1. IPv4选路

在互联网或内联网上传输的 IPv4包必须从一个网络选路到另一个网络以到达其目的地。选路协议可以使用动态机制来确定路由，但是所有选路最终依赖于某个路由器查看不同路由的列表并确定正确的路由。选路表包含网络的列表和连接到这些网络的接口的列表。路由器查看包，确定包所在的网络(或该网络可能在的网络)，然后把包发送到适当的网络接口。

现在的关键问题在于路由表的长度将随着网络数量的增加而变长。而路由表越长，路由器在表中查询正确路由的时间就越长。如果只需要了解 10个、100个或1000个网络，这不是问题。但是对诸如现在的Internet，拥有大量的网络，在骨干路由器上通常携带超过 11万个不同网络地址的显式路由，此时选路变成了一场恶梦。

选路议题影响到性能，它对互联网增长的影响远比地址空间的匮乏更紧迫。 IPv4地址可能在5年内使用殆尽，但如果不使用分级寻址来集聚和简化选路， Internet的性能可能在最近甚至现在就变得不可接受。

2. 划分子网

对子网的合理使用将增加地址使用的效率，但它对于效率的改进是有限的。如果想了解原因的话，先考虑原来的网络地址分配方式：一个机构可以申请到一个 A、B或C类地址。如果能够证明自己需要相当数量的主机地址，机构也许能获得一个 B类地址；否则，获得的将会是一个C类地址。无论申请人的网络中的主机是 200台、20台还是2台，他们都将获得一个C类地址，这样就占用了254个主机地址。如果他们能够使权威机构确信他们的确需要一个 B类地址的话，即便他们只有1000台主机，他们仍将获得完整的B类地址，这样一来又占用了65 534个主机地址。

获得这些地址后，从外部发往网络内任一处的业务流都在一个路由器接口处理，该路由器将把这些数据重新选路到本机构内的目的地。这种体系结构意味着用户可以按照自己的愿望来设计网络。图3-1 中显示了两种方案。两个网络都连接到Internet，但C类网在本机构内只提供了一个网络的连接能力，而 B类网络把机构划分成三个子网，通过内部路由器彼此连接，并通过第二个路由器连接到Internet。

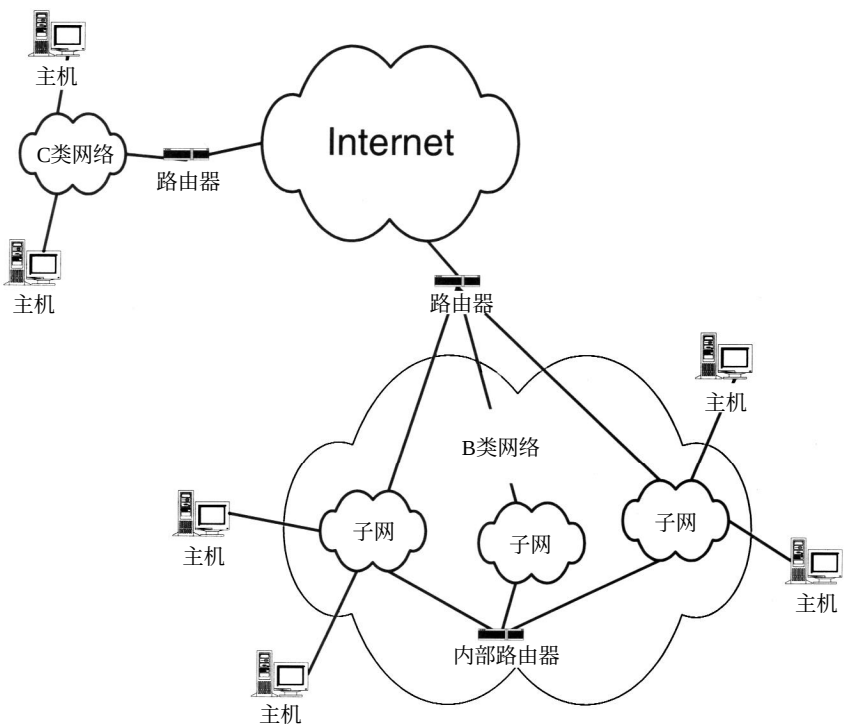

图3-1 子网有助于组织网络业务流，并改善网络地址使用的效率

当本地网络媒体在网络大小或连接的主机数量上到达极限时，就需要划分子网，同时它也可用来反应机构的体系结构。图中没有明确显示出子网不必在同一个建筑物内或同一个城市内。路由器重定向数据可以经过本地连接，也可以经过长途数据通信链路。这意味着一个机构可以与不同的分支、操作单位或子公司一起共享一个网络地址。

划分子网的问题在于它只适用于某种特定规模的机构——或者是C类网络或者是B类网络。例如，一个大型机构使用一个B类地址的网络有以下优点：使用8位子网掩码(换句话说，从B类地址的16位主机地址中借用8位)意味着可以有256个子网，其中每个子网可以最多有254个主机。如果使用9位子网掩码，还可以把子网数量倍增至512，当然每个子网上的主机数量降至了126。通过增加或减少位数量，可以很好的调整子网结构使之适应整个机构的体系结构。

对于需要B类地址的用户，不幸的是，除非是已经有了B类网络，否则目前很难得到B类地址。地址授权机构现在将C类地址成块分配给ISP，并通过它们再重新分配给用户。一个C类网络最多只能处理254台主机(绝对是最大值)，如果划分子网后，其主机数将进一步减少。因此，划分子网的C类网络可以用于包含不超过8(或16)个子网的小公司，同时每个子网上的主机数量少于30个(或14个)。即便如此，这两种配置还是把网络能够容纳的主机数量限制为不能超过210个，降低了地址分配的效率。

子网有助于在一个机构内组织其业务流，同时可以使来自外部源的数据报的选路更简单。外部源无需知道目的子网的任何情况，因为所有的子网是在同一个网络地址下，且所有去往该网络中任何地址的数据报都要首先经过一个路由器，然后由该路由器决定把数据向哪个子网发送。

划分子网的有趣特点在于可以对一个已经划分子网的网络进一步划分子网。在图 3-2中，一个B类网络被分为三层。第一层的路由器连接在Internet上，没有子网操作。但是，在该机构内，Internet路由器意识到有4位用于子网。这意味着最多可能有16个子网；那些子网中的任何一个都可以像图中所示一样进一步划分子网。在这个例子中，它们每一个都为最低层的子网又使用了4位，但机构内部的不同组可以选择不同的方法来分配其地址。例如，一个组具有多个小组但每个小组中的主机数量较少，可以使用6位作为子网，此时使得子网掩码的总长度达到10位；而另一个具有较少分支但每个小组都比较大，因此只使用了3位，从而使总的子网掩码达到7位。

3. 无类域间选路

无类域间选路(CIDR)技术有时也被称为超网，它把划分子网的概念向相反的方向作了扩展：通过借用前三个字节的几位可以把多个连续的C类地址集聚在一起。换句话说，就像所有到达某个B类地址的数据都将发给某个路由器一样，所有到达某一块C类地址的数据都将被选路至某个路由器上。

称做无类选路的原因在于它使得路由器可以忽略网络类别(C类)地址，并可以在决定如何转发数据报时向前再多看几位。另外一个与子网划分不同的特点在于，对于外部网络来说，子网掩码是不可见的；而超网路径的使用主要是为了减少路由器上的路由表项数。例如，一个ISP可以获得一块256个C类地址。这可以认为与B类地址相同，只不过前3位不是10x而是110。有了超网后，路由器可设定为包含地址块的前16位，然后把地址块作为有8位超网的一条路由来处理，而不再是为其中包含的每个C类地址处理最多可能256项路由。由于ISP经常负责为他们的客户的网络提供路由，于是他们获得的通常就是这种地址块，从而所有发往其客

户网络的数据可以由ISP的路由器以任何一种方式选路。

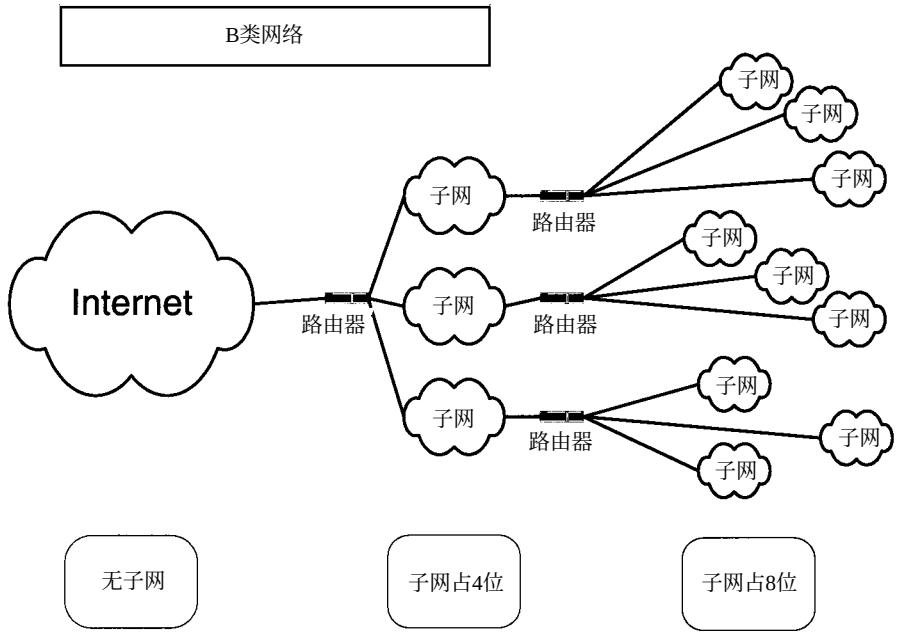

图3-2 子网可以进一步划分，产生更复杂的网络

由于B类网络相对缺乏和C类网络相对富余，这种把C类地址捆在一起的方法对于中等规模的机构来说很有用。此外，CIDR还缩短了路由表，这大大增加了选路的效率。但是，虽然CIDR增强了网络地址分配的效率，可它却并不能增加IPv4下总的主机数量，因此这只是一种短期解决办法而不是对于IPv4问题的长期解决方案。

4. 网络地址翻译

网络向外泄露的信息越少，网络的安全性就越高。对于TCP/IP网络来说，这意味着可能需要在内部网络和外部网络间设立一个防火墙，由它来接收所有请求。既然内部主机与外部主机失去了直接联系，那么IP地址就无所谓全球唯一，换句话说，如果内部主机不需要由Internet上的主机直接寻址，那么就可以为它们任意选择一个IP地址。实际上，许多与Internet没有任何联系的机构采用的就是这种方法。但当他们确实需要把二者连接在一起时，就必需对所有主机重新编号。

曾经有一段时间，许多公司无论是否打算连接Internet，都急于先申请到一段全球唯一的地址，因为这样可以使他们将来不必为主机重新编号。但是，随着专用IP网络的发展，为避免减少可分配的IP地址，有一组IP地址被拿出来专门用于专用IP网络。任何一个专用IP网络均可以使用包括一个A类地址(10.0.0.0)、16个B类地址(从172.16.0.0到172.31.0.0)和256个C类地址(从192.168.0.0到192.168.255.0)在内的任何地址。同时正如RFC 1918中的定义，把这些专用网络连接到公用网络的路由器不转发该网络上的任何数据。

网络地址翻译(NAT)在专用网络和公用网络之间的接口实现，该系统(一般是防火墙或路由器)了解专用网络上所有主机的地址，并将其翻译为可访问的公用网络地址，这样所有的内部主机就可以与外部主机通信。

虽然这种办法对于提高IP地址的分配效率有所帮助，但是网络设计人员在决定一个网络

是否使用NAT之前必须非常小心，要先确定其是否适用。对于那些永远不需要与其他网络合并或直接访问公用网络的网络， NAT很适合。潜水艇上的IP网络就可能非常适合使用专用地址：它不太可能与另一艘潜水艇上的网络合并在一起，也不太可能需要直接连接到其他网络或公用网络。如果两个以上使用专用网络地址的网络需要合并，例如两个使用专用 IP地址的银行要把他们的ATM机合并，那么最终形成的网络很可能需要进行重新编号以避免 IP地址的冲突。

NAT为一些小型机构提供了一种自己管理其地址空间的简单方法，无需依赖于地址授权机构为他们现在及将来的需要来分配足够的地址空间。 NAT还使得一些机构可以非常快速和灵活地定义临时地址或真正的专用网络地址。与 CIDR不同， NAT确实提供了一种可以真正减少IP地址需求的办法，尽管它使用起来有很大随意性，并且在重新对专用 IP网络编址时将花费较长的时间和昂贵的代价。

5. 网络管理与配置

设计IPv4和大多数其他TCP/IP应用协议集的目的都不是易于使用。例如，原始的文件传送协议(FTP)依靠的就是非常神秘的请求和响应代码，并使用了类似天书般的命令。提到这一点的原因在于：实际上设计这些显然很复杂的命令和控制机制的目的是实现跨平台的标准化，并简化对理解这些协议的软件的访问。一个使用 IPv4的系统必须使用一组特别复杂的参数来进行正确的配置，其中一般包括：主机名、 IP地址、子网掩码、默认路由器及其他 (根据应用而有所不同)。这种做法很复杂，意味着进行这些配置的人必须理解所有这些参数，或者至少由真正理解它的人来提供这些参数。这意味着将一个系统连接到 IPv4网络将十分复杂、非常耗时且代价高昂。

启动协议(BOOTP)是将主机连接到网络的简化过程的第一步。这个相对比较简单的协议为只具备极少配置信息的主机 (通常是一个简单的终端)提供了到BOOTP服务器获取其IP配置信息的方法。由于它只提供了将 IP地址及其他配置信息与链路层地址 (例如以太网卡地址)映射的方法，故它并不足以解决所有问题。要使用 BOOTP管理100台主机，则必须为每台主机指定其IP地址。

地址管理和主机配置提出了至少两大问题：首先，如果配置主机很困难，将耗费钱财；其次，如果无论是否已连接，均为每个主机捆绑一个 IP地址，这将浪费地址。如果可以使主机的配置变成即插即用，那将是一种好方法。即，只需把系统插到网络上，它将自动配置。在多个主机间共享IP地址也是一种好方法，如果有100台主机，在任意时刻同时上网的主机数不超过一半，那么只需使用50个IP地址让它们共享即可。

试图解决这些问题的结果是：在 BOOTP框架之上构造了另一个称为动态主机配置协议(DHCP)的协议。它使用的仍然是客户机/服务器模型，与BOOTP一样，客户机可以使用DHCP来向服务器查询配置信息。但是， DHCP更具灵活性，因为它可以随着IP地址的分配办法的不同而提供不同的配置信息。地址的分配有以下三种机制：

- 自动分配。主机申请IP地址，然后获得一个永久地址，可在每次连接网络时使用。
- 手工分配。服务器根据网络管理员提供的表格为每个主机分配一个特定 IP地址。无论主机是否需要，这些地址都将被保留。
- 动态分配。服务器按照先来先服务的方法分配 IP地址，主机在一个特定时间范围内使用该IP地址，然后该地址“借用”期满。

无论是自动分配还是手工分配都可能使得 IP地址分配效率很低。自动分配可能占用 IP地址，如果一个机构的主机数量多于用户数量，使用这种方法将占用与主机数量相同的地址。手工分配意味着网络管理员必须为每个主机配置一个 IP地址，而不管其连接网络的时间是一个小时还是一年。动态分配使得可以在一个大的用户数量的前提下共享少量的 IP地址。

不幸的是，DHCP由于其与状态相关的特性而无法实现真正的即插即用。用户不得不建立一个了解其主机的DHCP服务器，并且要使支持DHCP的主机了解最近的DHCP服务器。真正的即插即用，其实是移动性问题的一部分，而这正是 IPv4不能支持的。下面我们会看到，IPv4不具备支持移动性和网络管理能力，这也增强了升级到 IPv6的呼声。

6. 服务类型

IP使用的是包交换网络体系结构。这意味着包可以使用许多不同的路由到达目的地。这些路由的区别在于：有的代价比较高，有的吞吐量比较大，有的延时比较小，还有的可能会比其他的更可靠。第 2章讨论的IPv4服务类型(TOS)字段，允许应用程序告诉 IP如何处理其业务流。一个需要大吞吐量的应用，如 FTP，可以强制TOS为其选择具有更大吞吐量的路由；一个需要更快响应的应用，如Telnet，可以强制TOS为其选择具有更小延时的路由。

这确实是一个好想法，但却从来没能在实际应用中真正实现。一方面，这需要选路协议彼此协作，除提供基于开销的最佳路由外还要提供可选路由的延时、吞吐量和可靠性的数值。另一方面，还需要应用开发者实现一个功能，使其可以提出可能影响性能的服务请求。 TOS是一种选择，如果用户认为低延时对于其应用最重要，则应用的吞吐量或可靠性将爱到影响。

7. IP选项

第2章中曾提到，IPv4头包含了一个可变长的选项字段。IP选项用于指示一些特殊的功能。在最初的规范中没有定义这些选项，但最终增加了关于安全性和选路功能的选项。选路选项中包括一个记录路由的功能，让每个处理包的路由器都将自己的地址记录到该包中，另一个时间戳功能让每个路由器在包中记录自己的地址和处理包的时间。另外还有源选路选项：“宽松源选路”指明包在发往其目的地的过程中必须经过的一组路由器，而“严格源选路”则指定了该包只能由列出的路由器处理。

IP选项的问题在于它们是特例。大多数 IP数据报不包括选项，并且厂商按不包括选项的数据报来优化路由器。IP头如果不包括选项，则5字节长，易于处理，尤其是在路由器设计优化了对这种头的处理之后。对于路由器的销售而言，性能是关键，且由于大部分数据报不支持IP选项，因此路由器往往把这种包作为特例，搁置起来，只有在不会影响路由器总体性能时才加以处理。

尽管使用IPv4选项有很多好处，但由于它们对于性能的影响已使得它们很少使用。

8. IPv4安全性

很长时间以来，都认为安全性不是网际层的任务。在这种情况下，安全性意味着对净荷数据的加密。其他安全性概念还包括对净荷的数字签名、密钥交换、实体的身份验证和资源的访问控制。这些功能一般由较高层处理，通常是应用层，有时是传输层。例如，广泛应用的安全套接字层(SSL)协议由IP之上的传输层处理，而应用相对较少的安全 HTTP(SHTTP)则由应用层处理。

最近，随着虚拟专用网(VPN)软件和硬件产品的引入，安全隧道协议和机制有所扩展。这些产品通常会对一个IP数据报流加密，即把这些包本身作为另一些 IP数据报的净荷。IP数据报

可想象为一个包装好的盒子，里面还包含着一个小盒子，在小盒子中还包含另一个盒子。最小的盒子中包含的是应用数据，下一个盒子中是传输层数据，而最外面的盒子包含 IP数据。实际上隧道的工作方法就是把一个IP盒子放在不同地址信息的另一个IP盒子中。

隧道协议，如微软的点到点隧道协议 (PPTP)，首先对IP数据报加密，然后打包，再发送到隧道上。

所有这些关于IP安全性的办法都有问题。首先，在应用层进行加密使很多信息被公开。尽管应用层数据本身是加密的，携带它的 IP数据仍会泄露参与处理的进程和系统的信息。在传输层加密要好一些，并且SSL为Web的安全性工作得很好，但它要求客户机和服务器应用程序都要重写以支持SSL。隧道协议也工作得不错，但却被缺乏标准的问题所困扰。

IETF的IP安全性(IPsec)工作组一直致力于设计一种机制和协议来同时保证 IPv4和IPv6业务流的安全性。虽然已有一些基于 IP选项的关于IPv4安全性的机制，但在实际应用中并不成功。IPsec的目标是使这些工具可用，并在IPv6中集成更加完整的安全性。

3.2 过渡还是不过渡

毫无疑问，IPv4需要一些改变以使得它能够在网络协议的发展中得以继续生存。增长中的网络正在消耗有限的IP地址空间资源，这一简单问题意味着地址空间必须扩充。前一节中列举了一些可以帮助IPv4延长生命的著名方法，但是众所周知那不过是临时的办法。现在已经清楚，地址问题并不是IPv4中存在的唯一问题：网络越多意味着路由表越大，同时还导致路由器的性能下降。同样，难以实现IPv4选项意味着这些选项中实现的功能对用户不可用。

考虑一下如果只是简单地倍增 IP地址的长度而不修改协议的其他部分，将会发生什么。所有的TCP/IP协议栈将需要同时被更新。落在后面的人将丧失与 Internet连接的能力。尽管这种改变已经相对简单，但是由于错误配置而导致的系统瘫痪仍将产生巨大影响。对于任何人来说这种改变的代价都是巨大的，因为它意味着使用 IPv4的所有机构都需要定位系统中的每一台主机，对于拥有许多用户和主机的大型机构，这绝不是一件简单的事情。更复杂的是，那些系统中有许多是比较老的或过时的甚至是已经废弃的系统，在这些系统上运行的网络软件可能已经过期并且没有人再提供支持。

任何对于现有系统进行升级的请求都可能导致混乱。对于 IPv4的修补，无论是临时加入一个补丁还是用另一个重新设计的协议来替换，都将导致混乱。既然与其他方法相比，升级不会带来更多的痛苦，那么在可能存在一个更强健的修补方案时，何必再使用一个一个单独的补丁呢？

IPv6协议规范在1995年底提交IETF并获得批准。软件厂商最早在 1996年就已经开始提供IPv6网络协议栈的测试版。 1997年，测试产品和实验性的IPv6骨干网(6BONE)已经就位，但是到了1998年升级的势头缓慢下来。无论如何，最近还无法确定一个明确的“ 交割 ”日期。相反，将逐渐出现向IPv6过渡(后续章节讨论)，IPv4与IPv6将共存并交互。

向IPv6的过渡最有可能从高端而不是从最终用户开始，即一些机构和 ISP可能会先在其骨干网络中首先实现IPv6。即便如此，这些机构中也会有一部分会先去解决两千年问题，从而进一步降低过渡的速度。但是不管怎么说，只要应用开发者开始递交基于 IPv6的新产品，那么向IPv6发展的速度就将大大加快。尽管与两千年问题相比，该问题在最终期限上具有较大的灵活性，IPv6的计划也不容拖延太久。