

## 路由器原理及路由协议

来源：中国电信网站

本文通过阐述TCP / IP网络中路由器的基本工作原理，介绍了IP路由器的几大功能，给出了静态路由协议和动态路由协议，以及内部网关协议和外部网关协议的概念，同时简要介绍了目前最常见的RIP、OSPF、BGP和BGP-4这几种路由协议，然后描述了路由算法的设计目标和种类，着重介绍了链路状态法和距离向量法。在文章的最后，扼要讲述了新一代路由器的特征。

——近十年来，随着计算机网络规模的不断扩大，大型互联网络（如Internet）的迅猛发展，路由技术在网络技术中已逐渐成为关键部分，路由器也随之成为最重要的网络设备。用户的需求推动着路由技术的发展和路由器的普及，人们已经不满足于仅在本地网络上共享信息，而希望最大限度地利用全球各个地区、各种类型的网络资源。而在目前的情况下，任何一个有一定规模的计算机网络（如企业网、校园网、智能大厦等），无论采用的是快速以太网技术、FDDI技术，还是ATM技术，都离不开路由器，否则就无法正常运作和管理。

### 1 网络互连

——把自己的网络同其它的网络互连起来，从网络中获取更多的信息和向网络发布自己的消息，是网络互连的最主要的动力。网络的互连有多种方式，其中使用最多的是网桥互连和路由器互连。

#### 1.1 网桥互连的网络

——网桥工作在OSI模型中的第二层，即链路层。完成数据帧（frame）的转发，主要目的是在连接的网络间提供透明的通信。网桥的转发是依据数据帧中的源地址和目的地址来判断一个帧是否应转发和转发到哪个端口。帧中的地址称为“MAC”地址或“硬件”地址，一般就是网卡所带的地址。

——网桥的作用是把两个或多个网络互连起来，提供透明的通信。网络上的设备看不到网桥的存在，设备之间的通信就如同在一个网上一样方便。由于网桥是在数据帧上进行转发的，因此只能连接相同或相似的网络（相同或相似结构的数据帧），如以太网之间、以太网与令牌环（token ring）之间的互连，对于不同类型的网络（数据帧结构不同），如以太网与X.25之间，网桥就无能为力了。

——网桥扩大了网络的规模，提高了网络的性能，给网络应用带来了方便，在以前的网络中，网桥的应用较为广泛。但网桥互连也带来了不少问题：一个是广播风暴，网桥不阻挡网络中广播消息，当网络的规模较大时（几个网桥，多个以太网段），有可能引起广播风暴

（broadcasting storm），导致整个网络全被广播信息充满，直至完全瘫痪。第二个问题是，当与外部网络互连时，网桥会把内部和外部网络合二为一，成为一个网，双方都自动向对方完全开放自己的网络资源。这种互连方式在与外部网络互连时显然是难以接受的。问题的主要根源是网桥只是最大限度地吧网络沟通，而不管传送的信息是什么。

#### 1.2 路由器互连网络

——路由器互连与网络的协议有关，我们讨论限于TCP / IP网络的情况。

——路由器工作在OSI模型中的第三层，即网络层。路由器利用网络层定义的“逻辑”上的网络地址（即IP地址）来区别不同的网络，实现网络的互连和隔离，保持各个网络的独立性。路由器不转发广播消息，而把广播消息限制在各自的网络内部。发送到其他网络的数据包先被送到路由器，再由路由器转发出去。

——IP路由器只转发IP分组，把其余的部分挡在网内（包括广播），从而保持各个网络具有相对的独立性，这样可以组成具有许多网络（子网）互连的大型的网络。由于是在网络层的互连，路由器可方便地连接不同类型的网络，只要网络层运行的是IP协议，通过路由器就可互连起来。

——网络中的设备用它们的网络地址（TCP / IP网络中为IP地址）互相通信。IP地址是与硬件地址无关的“逻辑”地址。路由器只根据IP地址来转发数据。IP地址的结构有两部分，一部分定义网络号，另一部分定义网络内的主机号。目前，在Internet网络中采用子网掩码来确定IP地址中网络地址和主机地址。子网掩码与IP地址一样也是32bit，并且两者是一一对应的，并规定，子网掩码中数字为“1”所对应的IP地址中的部分为网络号，为“0”所对应的则为主机号。网络号和主机号合起来，才构成一个完整的IP地址。同一个网络中的主机IP地址，其网络号必须是相同的，这个网络称为IP子网。

——通信只能在具有相同网络号的IP地址之间进行，要与其它IP子网的主机进行通信，则必须经过同一网络上的某个路由器或网关（gateway）出去。不同网络号的IP地址不能直接通信，即使它们接在一起，也不能通信。

——路由器有多个端口，用于连接多个IP子网。每个端口的IP地址的网络号要求与所连接的IP子网的网络号相同。不同的端口为不同的网络号，对应不同的IP子网，这样才能使各子网中的主机通过自己子网的IP地址把要求出去的IP分组送到路由器上。

## 2 路由原理

——当IP子网中的一台主机发送IP分组给同一IP子网的另一台主机时，它将直接把IP分组送到网络上，对方就能收到。而要送给不同IP子网上的主机时，它要选择一个能到达目的子网上的路由器，把IP分组送给该路由器，由路由器负责把IP分组送到目的地。如果没有找到这样的路由器，主机就把IP分组送给一个称为“缺省网关（default gateway）”的路由器上。“缺省网关”是每台主机上的一个配置参数，它是接在同一个网络上的某个路由器端口的IP地址。

——路由器转发IP分组时，只根据IP分组目的IP地址的网络号部分，选择合适的端口，把IP分组送出去。同主机一样，路由器也要判定端口所接的是否是目的子网，如果是，就直接把分组通过端口送到网络上，否则，也要选择下一个路由器来传送分组。路由器也有它的缺省网关，用来传送不知道往哪儿送的IP分组。这样，通过路由器把知道如何传送的IP分组正确转发出去，不知道的IP分组送给“缺省网关”路由器，这样一级级地传送，IP分组最终将送到目的地，送不到目的地的IP分组则被网络丢弃了。

——目前TCP / IP网络，全部是通过路由器互连起来的，Internet就是成千上万个IP子网通过路由器互连起来的国际性网络。这种网络称为以路由器为基础的网络（router based network），形成了以路由器为节点的“网间网”。在“网间网”中，路由器不仅负责对IP分组的转发，还要负责与别的路由器进行联络，共同确定“网间网”的路由选择和维护路由表。

——路由动作包括两项基本内容：寻径和转发。寻径即判定到达目的地的最佳路径，由路由选择算法来实现。由于涉及到不同的路由选择协议和路由选择算法，要相对复杂一些。为了判定最佳路径，路由选择算法必须启动并维护包含路由信息的路由表，其中路由信息依赖于所用的路由选择算法而不尽相同。路由选择算法将收集到的不同信息填入路由表中，根据路由表可将目的网络与下一站（next hop）的关系告诉路由器。路由器间互通信息进行路由更新，更新维护路由表使之正确反映网络的拓扑变化，并由路由器根据量度来决定最佳路径。这就是路由选择协议

（routing protocol），例如路由信息协议（RIP）、开放式最短路径优先协议（OSPF）和边界网关协议（BGP）等。

——转发即沿寻径好的最佳路径传送信息分组。路由器首先在路由表中查找，判明是否知道如何将分组发送到下一个站点（路由器或主机），如果路由器不知道如何发送分组，通常将该分组丢弃；否则就根据路由表的相应表项将分组发送到下一个站点，如果目的网络直接与路由器相连，路由器就把分组直接送到相应的端口上。这就是路由转发协议（routed protocol）。

——路由转发协议和路由选择协议是相互配合又相互独立的概念，前者使用后者维护的路由表，同时后者要利用前者提供的功能来发布路由协议数据分组。下文中提到的路由协议，除非特别说明，都是指路由选择协议，这也是普遍的习惯。

### 3 路由协议

——典型的路由选择方式有两种：静态路由和动态路由。

——静态路由是在路由器中设置的固定的路由表。除非网络管理员干预，否则静态路由不会发生变化。由于静态路由不能对网络的改变作出反映，一般用于网络规模不大、拓扑结构固定的网络中。静态路由的优点是简单、高效、可靠。在所有的路由中，静态路由优先级最高。当动态路由与静态路由发生冲突时，以静态路由为准。

——动态路由是网络中的路由器之间相互通信，传递路由信息，利用收到的路由信息更新路由器表的过程。它能实时地适应网络结构的变化。如果路由更新信息表明发生了网络变化，路由选择软件就会重新计算路由，并发出新的路由更新信息。这些信息通过各个网络，引起各路由器重新启动其路由算法，并更新各自的路由表以动态地反映网络拓扑变化。动态路由适用于网络规模大、网络拓扑复杂的网络。当然，各种动态路由协议会不同程度地占用网络带宽和CPU资源。

——静态路由和动态路由有各自的特点和适用范围，因此在网络中动态路由通常作为静态路由的补充。当一个分组在路由器中进行寻径时，路由器首先查找静态路由，如果查到则根据相应的静态路由转发分组；否则再查找动态路由。

——根据是否在一个自治域内部使用，动态路由协议分为内部网关协议（IGP）和外部网关协议（EGP）。这里的自治域指一个具有统一管理机构、统一路由策略的网络。自治域内部采用的路由选择协议称为内部网关协议，常用的有RIP、OSPF；外部网关协议主要用于多个自治域之间的路由选择，常用的是BGP和BGP-4。下面分别进行简要介绍。

#### 3.1 RIP路由协议

——RIP协议最初是为Xerox网络系统的Xerox parc通用协议而设计的，是Internet中常用的路由协议。**RIP采用距离向量算法**，即路由器根据距离选择路由，所以也称为距离向量协议。路由器收集所有可到达目的地的不同路径，并且保存有关到达每个目的地的最少站点数的路径信息，除到达目的地的最佳路径外，任何其它信息均予以丢弃。同时路由器也把所收集的路由信息用RIP协议通知相邻的其它路由器。这样，正确的路由信息逐渐扩散到了全网。

——RIP使用非常广泛，它简单、可靠，便于配置。但是RIP只适用于小型的同构网络，因为它允许的最大站点数为15，任何超过15个站点的目的地均被标记为不可达。而且RIP每隔30s一次的路由信息广播也是造成网络的广播风暴的重要原因之一。

#### 3.2 OSPF路由协议

——80年代中期，RIP已不能适应大规模异构网络的互连，OSPF随之产生。它是网间工程任务组织（IETF）的内部网关协议工作组为IP网络而开发的一种路由协议。

——**OSPF是一种基于链路状态的路由协议**，需要每个路由器向其同一管理域的所有其它路由器发送链路状态广播信息。在OSPF的链路状态广播中包括所有接口信息、所有的量度和其它一些变量。利用OSPF的路由器首先必须收集有关的链路状态信息，并根据一定的算法计算出到每个节点的最短路径。而基于距离向量的路由协议仅向其邻接路由器发送有关路由更新信息。

——与RIP不同，OSPF将一个自治域再划分为区，相应地即有两种类型的路由选择方式：当



源和目的地在同一区时，采用区内路由选择；当源和目的地在不同区时，则采用区间路由选择。这就大大减少了网络开销，并增加了网络的稳定性。当一个区内的路由器出了故障时并不影响自治域内其它区路由器的正常工作，这也给网络的管理、维护带来方便。

### 3.3 BGP和BGP-4路由协议

——BGP是为TCP / IP互联网设计的外部网关协议，用于多个自治域之间。它既不是基于纯粹的链路状态算法，也不是基于纯粹的距离向量算法。它的主要功能是与其它自治域的BGP交换网络可达信息。各个自治域可以运行不同的内部网关协议。BGP更新信息包括网络号 / 自治域路径的成对信息。自治域路径包括到达某个特定网络须经过的自治域串，这些更新信息通过TCP传送出去，以保证传输的可靠性。

——为了满足Internet日益扩大的需要，BGP还在不断地发展。在最新的BGp4中，还可以将相似路由合并为一条路由。

### 3.4 路由表项的优先问题

——在一个路由器中，可同时配置静态路由和一种或多种动态路由。它们各自维护的路由表都提供给转发程序，但这些路由表的表项间可能会发生冲突。这种冲突可通过配置各路由表的优先级来解决。通常静态路由具有默认的最高优先级，当其它路由表表项与它矛盾时，均按静态路由转发。

## 4 路由算法

——路由算法在路由协议中起着至关重要的作用，采用何种算法往往决定了最终的寻径结果，因此选择路由算法一定要仔细。通常需要综合考虑以下几个设计目标：

——（1）最优化：指路由算法选择最佳路径的能力。

——（2）简洁性：算法设计简洁，利用最少的软件和开销，提供最有效的功能。

——（3）坚固性：路由算法处于非正常或不可预料的环境时，如硬件故障、负载过高或操作失误时，都能正确运行。由于路由器分布在网络联接点上，所以在它们出故障时会产生严重后果。最好的路由器算法通常能经受时间的考验，并在各种网络环境下被证实是可靠的。

——（4）快速收敛：收敛是在最佳路径的判断上所有路由器达到一致的过程。当某个网络事件引起路由可用或不可用时，路由器就发出更新信息。路由更新信息遍及整个网络，引发重新计算最佳路径，最终达到所有路由器一致公认的最佳路径。收敛慢的路由算法会造成路径循环或网络中断。

——（5）灵活性：路由算法可以快速、准确地适应各种网络环境。例如，某个网段发生故障，路由算法要能很快发现故障，并为使用该网段的所有路由选择另一条最佳路径。

——路由算法按照种类可分为以下几种：静态和动态、单路和多路、平等和分级、源路由和透明路由、域内和域间、链路状态和距离向量。前面几种的特点与字面意思基本一致，下面着重介绍链路状态和距离向量算法。

——链路状态算法（也称最短路径算法）发送路由信息到互联网上所有的结点，然而对于每个路由器，仅发送它的路由表中描述了其自身链路状态的那一部分。距离向量算法（也称为Bellman-Ford算法）则要求每个路由器发送其路由表全部或部分信息，但仅发送到邻近结点上。从本质上来说，链路状态算法将少量更新信息发送至网络各处，而距离向量算法发送大量更新信息至邻接路由器。

——由于链路状态算法收敛更快，因此它在一定程度上比距离向量算法更不易产生路由循环。但另一方面，链路状态算法要求比距离向量算法有更强的CPU能力和更多的内存空间，因此链路状态算法将会在实现时显得更昂贵一些。除了这些区别，两种算法在大多数环境下都能很好

地运行。

——最后需要指出的是，路由算法使用了许多种不同的度量标准去决定最佳路径。复杂的路由算法可能采用多种度量来选择路由，通过一定的加权运算，将它们合并为单个的复合度量、再填入路由表中，作为寻径的标准。通常所使用的度量有：路径长度、可靠性、时延、带宽、负载、通信成本等。

## 5 新一代路由器

——由于多媒体等应用在网络中的发展，以及ATM、快速以太网等新技术的不断采用，网络的带宽与速率飞速提高，传统的路由器已不能满足人们对路由器的性能要求。因为传统路由器的分组转发的设计与实现均基于软件，在转发过程中对分组的处理要经过许多环节，转发过程复杂，使得分组转发的速率较慢。另外，由于路由器是网络互连的关键设备，是网络与其它网络进行通信的一个“关口”，对其安全性有很高的要求，因此路由器中各种附加的安全措施增加了CPU的负担，这样就使得路由器成为整个互联网上的“瓶颈”。

——传统的路由器在转发每一个分组时，都要进行一系列的复杂操作，包括路由查找、访问控制表匹配、地址解析、优先级管理以及其它的附加操作。这一系列的操作大大影响了路由器的性能与效率，降低了分组转发速率和转发的吞吐量，增加了CPU的负担。而经过路由器的前后分组间的相关性很大，具有相同目的地址和源地址的分组往往连续到达，这为分组的快速转发提供了实现的可能与依据。新一代路由器，如IP Switch、Tag Switch等，就是采用这一设计思想用硬件来实现快速转发，大大提高了路由器的性能与效率。

——新一代路由器使用转发缓存来简化分组的转发操作。在快速转发过程中，只需对一组具有相同目的地址和源地址的分组的前几个分组进行传统的路由转发处理，并把成功转发的分组的目的地址、源地址和下一网关地址（下一路由器地址）放入转发缓存中。当其后的分组要进行转发时，首先查看转发缓存，如果该分组的目的地址和源地址与转发缓存中的匹配，则直接根据转发缓存中的下一网关地址进行转发，而无须经过传统的复杂操作，大大减轻了路由器的负担，达到了提高路由器吞吐量的目标。

[关闭窗口]

©2000-2010 世界网络 版权所有！

太阳网数据中心提供网络带宽