

第2章 透明桥接

在第1章中，我们谈到了交换机、网桥和路由器所能提供的网段类型，但是，没有对网桥或交换机怎样作出转发决定加以讨论。本章将对以太网环境中的网桥 / 交换机的功能加以介绍。实际上，网桥和交换机以相同的方式对帧进行转发，在本章的大部分内容中，只使用网桥这一词。不过本章所提到的各种概念也同样适用于交换机。

因为网桥只基于目标 MAC 地址作出转发决定，所以它不像路由那样会在路由器中形成地址层次。这样，每个节点将不得不经过“获取”后存放在 RAM（内存）或缓存中的表中。本章将描述这一过程，然后介绍过程中将出现的一些问题以及解决这些问题的方法。

在以太网中，作出转发决定的过程称为透明桥接。但是，在令牌环网环境中，这一过程有一点不同，称为源路由桥接，这将在第3章中加以讨论。

2.1 透明网桥的三个功能

透明网桥主要用在以太网环境中，并设计成对于末端节点透明。透明网桥有如下的三个主要功能。

- 1) 获取功能。
- 2) 转发和过滤功能。
- 3) 消除循环功能。

当透明网桥的三个主要功能都能使用时，它们是在网络中是同时起作用的。应该注意，在有些书中列出了网桥的四种功能，在这四种功能中，转发功能和过滤功能成为两个单独的功能。因为 Cisco 认为这两种功能是同一种功能，我也比较赞同这种看法，所以在本书中只列出了三种功能。

2.1.1 获取

网桥基于目标 MAC（介质访问控制）地址作出转发决定。所以它必须“获取” MAC 地址的位置，这样才能准确地作出转发决定。每个以太网物理网段上的工作站都将对所有在网段上传输的帧进行监听。当网桥与物理网段连接时，它会对它监测到的所有帧进行检查。网桥读取帧的源 MAC 地址字段后便作出一个假定。这个假定是，如果它监测到一个来自特殊端口上节点的帧，发送帧工作站的信息就必须驻留在这个端口上。网桥将这个信息放置在一个网桥表中，它在将条目引入网桥表之前，还将执行 FCS，以阻止错误的条目进入网桥表。在 Catalyst 交换机中，这个表称为 CAM（content addressable memory，内容可寻址内存）表。网桥表和 CAM 表基本上是相同的，只有一些小的差别，在后面将会提到。

图2-1给出了四个工作站：A、B、C和D。四个工作站分为两个物理网段，网段中间为网桥，网段通过网桥的两个端口与网桥相连。当工作站 A 向工作站 B 传输信息时，网桥和工作站 B 都将收到这个信息。

网桥收到这个信息后，得知工作站 A 与端口 1 相连，因为从这个端口收到了帧信息。网桥

把条目引入记录工作站 A 的 MAC 地址的网桥表，如图 2-2 所示。

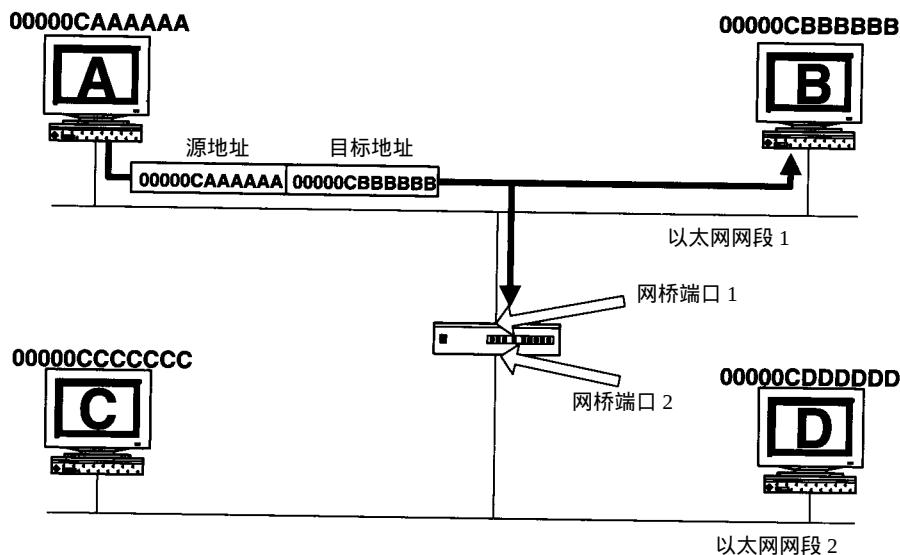


图2-1 工作站A将信息传递至以太网网段1上的工作站B

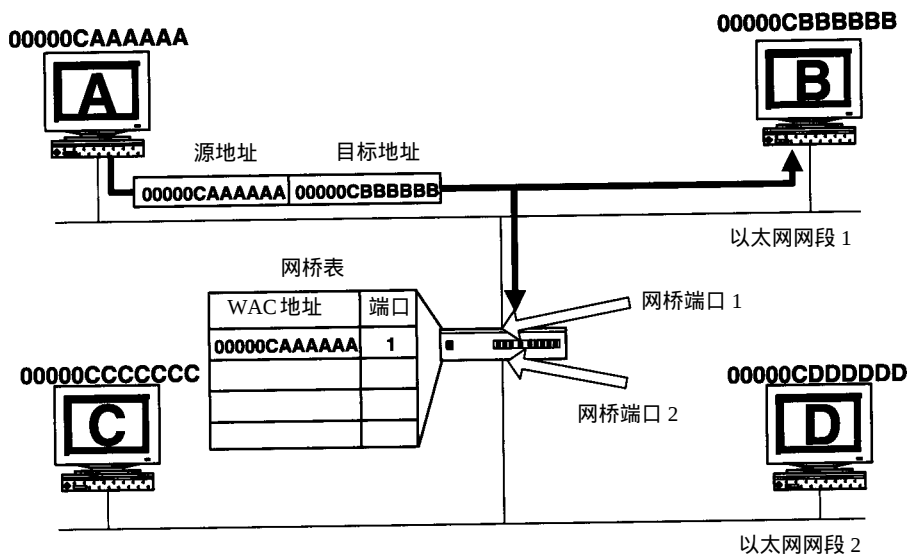


图2-2 网桥得知工作站A与端口1相连

相反，当工作站 B 对工作站 A 的信息作出反应后，网桥监测到工作站 B 反发送出的帧，并将其 MAC 地址作为条目引入网桥表中，如图 2-3 所示。

网桥连续地进行“获取”。这时条目保存在网桥表中。如果这时没有监测到来自 MAC 地址的信息，5 分钟后网桥将停止“获取”。这个时间间隔对于几乎所有的交换机和网桥都是可以配置的，称为老化时间。条目也可以以手工方式输入网桥表。

最后所有的 MAC 地址都将被网桥获取（假定所有的工作站都在使用中）。

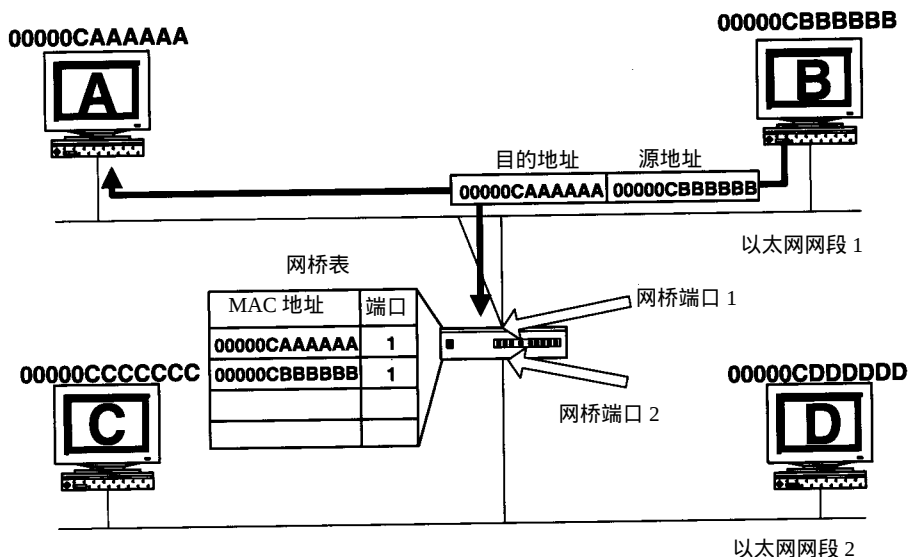


图2-3 网桥得知工作站B与端口1相连

2.1.2 转发和过滤

网桥的第二个功能是转发功能和过滤功能。使用网桥表后，网桥可以作出转发或不转发（即过滤）帧的决定。这个决定取决于帧报头中的目标 MAC 地址。

如果工作站 A 向工作站 C 传送信息，并且工作站 C 在网桥表中有一个条目，网桥把帧发送至以太网网段 2（见图 2-4）。

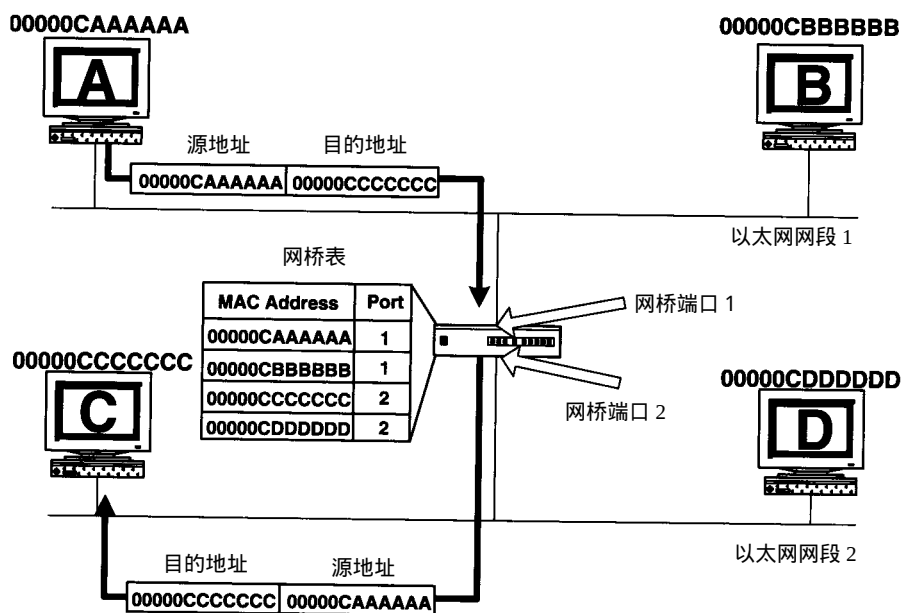


图2-4 转发

如果工作站 A 向工作站 B 传送信息，帧的发送没有必要，因为工作站 B 与工作站 A 在同一个

物理网段上。这时，网桥进行过滤，即不发送帧信息（见图 2-5）

如果工作站A向工作站C发送帧，而在网桥表中没有工作站C，那会出现什么情况呢？网桥将把这个发往网桥不知道的目标MAC地址的帧发至所有端口。在这种情况下，网桥充当的是集线器的角色，以确保它没有使信息停止传送（见图 2-6、见图2-7）。

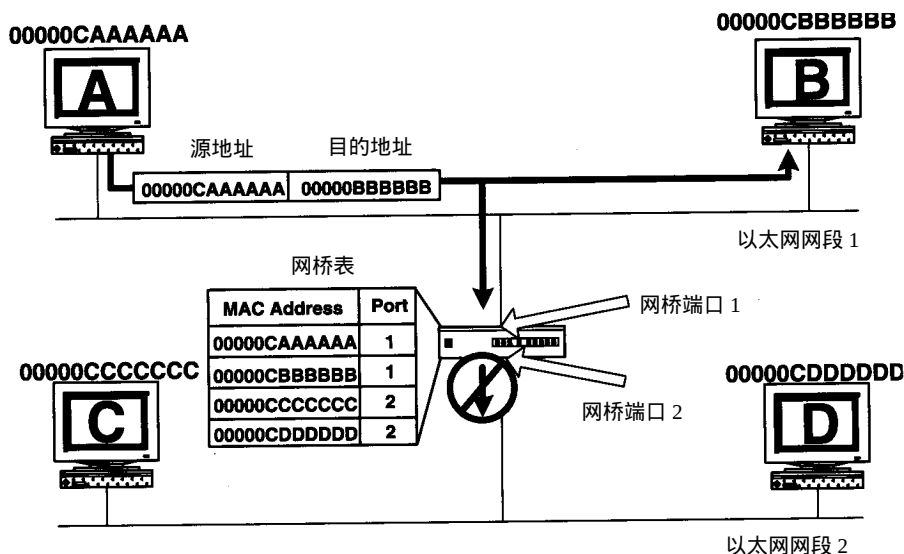


图2-5 过滤（即不转发）

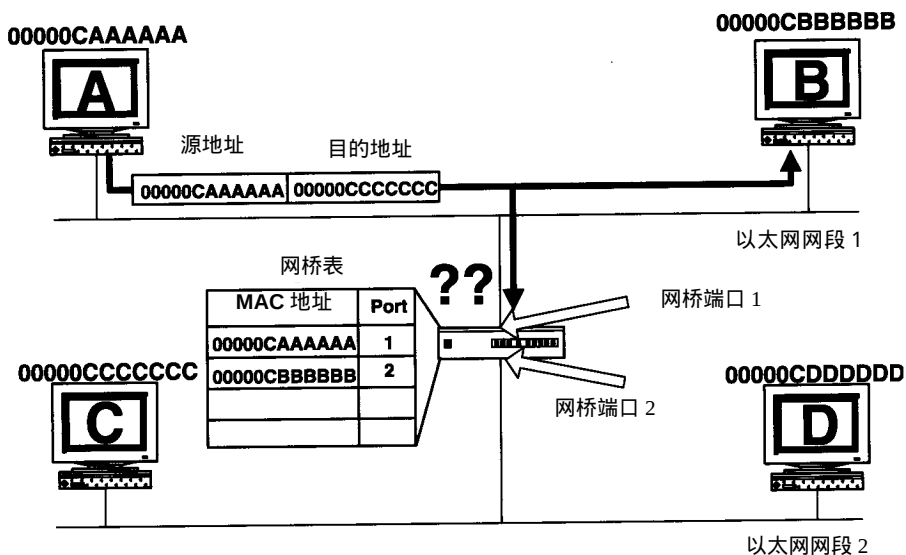


图2-6 网桥表中没有目标MAC地址的情况

如果网桥没有对不知道目标MAC地址的帧进行发送，工作站A不会与工作站C进行连通，直至工作站C传送一个帧。这种情况是不能接受的。

网桥也转发广播而且向端口进行多点传送，其方式和不知道目标端口的帧的发送方式一样。

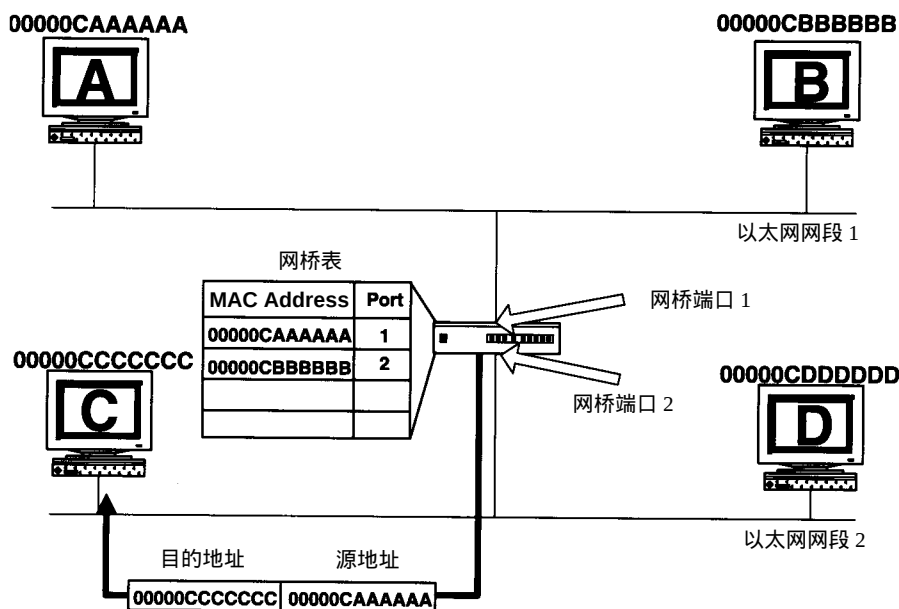


图2-7 网桥向所有的端口发送帧

2.1.3 消除循环

透明网桥的最后一个功能是消除循环，这个功能与其他两个功能比较起来是最难理解的。图2-8为一个有冗余的多重网桥。广播信息从图中右上角的工作站发出。

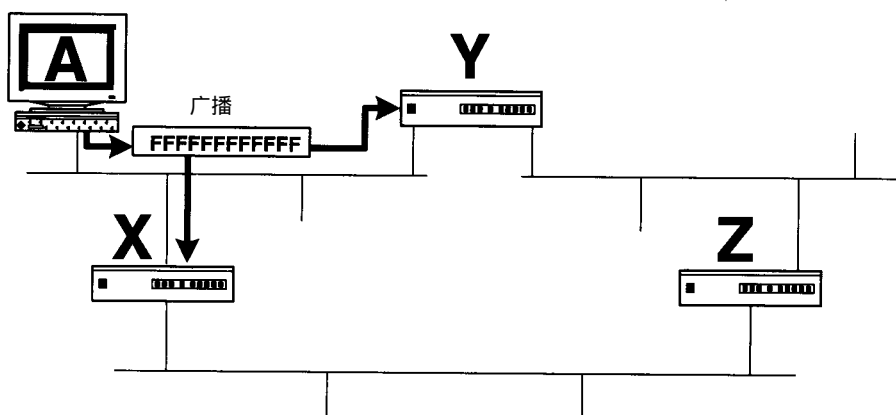


图2-8 网桥循环

在图2-9中，网桥 X、Y 都与第一个物理网段相连，它们先对广播进行监测，监测到广播信息后，便将信息发送至所有的其他端口，这个例子中只有一个其他端口，即端口 Z。

在图2-10中，广播在第二个和第三个物理网段进行传播，而这两个网段都和网桥 Z 相连，网桥 Z 监测到这个广播后，又将广播发送到第三个和第二个物理网段。

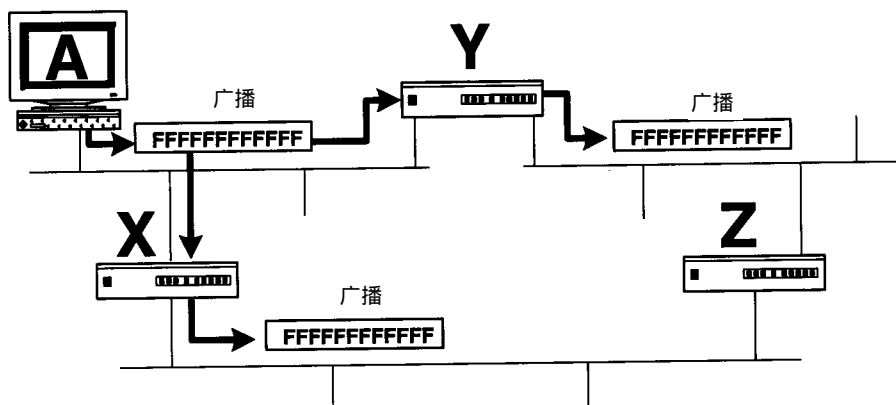


图2-9 网桥循环

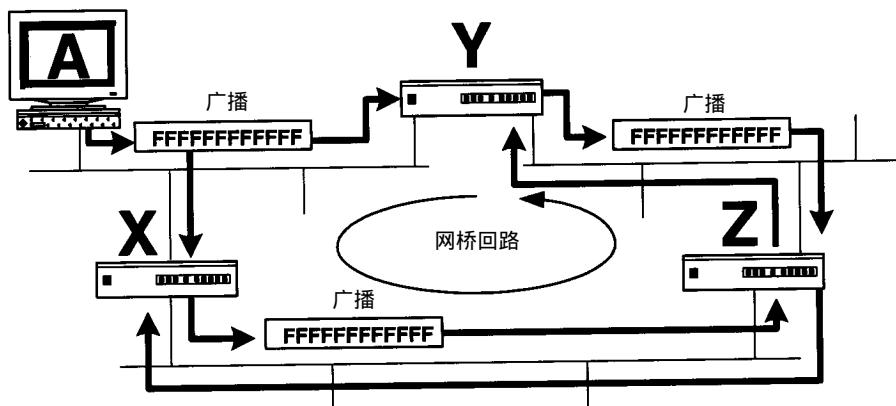


图2-10 网桥循环

这种现象称为网桥循环。在网桥循环中，广播将对第三个物理网段产生损害，如图 2-8到 2-10所示。在实际的运用中，存在数百个物理网段。当工程师们设法对循环进行定位时，这种情况可能会导致数百次死机。一旦循环位置确定后，唯一的解决办法是切断所有连接。

网桥的第三个功能是对循环进行定位并切断多余的连接。为了达到此目的，网桥必须知道其他网桥。当网络上存在一个以上的网桥时，源 MAC地址位于接收帧的端口这一假设将不成立。网桥和另一网桥交换信息的协议称为 STP（生成树协议）

STP有几种不同的版本，而且这些版本相互之间并不相容。对所有的网桥和交换机使用相同的STP版本进行验证是非常重要的，因为不同类的 STP版本将导致网桥循环，这将对以太网的物理网段造成损害。两种最通用的 STP版本为DEC和IEEE。另外还有几种不是很流行的版本。所以，保证不使用非通用版本是非常重要的，因为将来这些版本可能会造成潜在的问题。大多数交换机一般都只支持 DEC和IEEE。Cisco Catalyst交换机在以太网端口上使用的是 IEEE版本，在令牌环端口上使用的是 IBM版本的STP。

2.2 生成树协议

网桥为了让其他网桥知道它的存在，必须向其他端口传送小的信息包，这是可能导致网

桥循环的潜在因素，这些小的信息包称为 BPDU（bridge protocol data unit，网桥协议数据单元）。一个 Catalyst 交换机以每 10 分钟一个的速度从所有的活动端口发出 BPDU。网桥接收到这个 BPDU 后，便利用一个称为 STA（spanning tree algorithm，生成树算法）的数学公式进行计算。通过 STA 计算，网桥就可以知道网络上是否存在循环，当存在循环时，网桥就作出冗余的端口应该被切断的决定。切断端口的过程称为阻塞。受到阻塞的端口仍然是一个活动的端口，也就是说，它仍然可以接收和读取 BPDU。这一过程一直持续到出现失败或者拓扑变化消除循环才能结束。当这一过程结束后，端口便开始发送帧，因为这时循环已经不存在了（图 2-11）。

BPDU 和 STA 的目的是创建一个“非循环”的环境。例如，现实生活中的树是一种进行自然循环的树。所有树干的底部都有树根，树干朝上分成大树枝，大树枝再分成小树枝等等。但是，树枝从来都不会蔓延到其他树枝，即形成一个“非循环”的环境。

STA 看起来是稍微有点复杂，但是只要知道几个小概念，就可以知道哪个端口将处于“阻塞”状态。

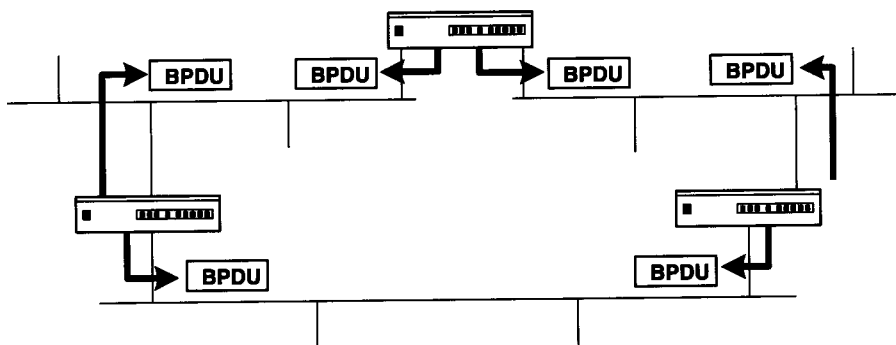


图2-11 所有网桥将发送BPDU

2.2.1 根网桥

所有的树都有根，生成树也有根。在这里介绍一种特殊的网桥，它有一个非常恰当的名字：根网桥。在网络中，所有的网桥都分配了一个数值，这个值称为网桥优先级值，优先级值最小的网桥称为根网桥。默认的情况下，Catalyst 交换机的网桥优先级值为 32768，所以，

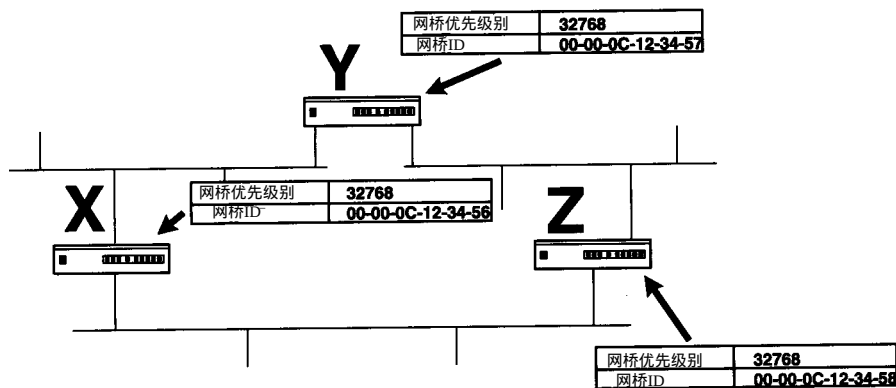


图2-12 所有网桥有相同的优先级

如果只使用 Catalyst 交换机的话，所有网桥的优先级都为最低。在根网桥的判定中，要用到一种决胜法，决胜法与 NFL 延长赛的方式有所不同。在 Catalyst 交换机中，地址池中最低优先级的 MAC 地址分配给监控机。这个 MAC 地址通常称为网桥 ID。网桥 ID 其实就是一个决胜方法，也就是说，网桥 ID 值为最小的网桥将是根网桥。将 MAC 地址作为网桥 ID，是为了确保有且只有一个网桥的 ID 值最小，因为 MAC 地址在全世界都是唯一的。图 2-12 中，所有网桥的优先级相同，所以决胜的方法只能是网桥 ID 值。网桥 X 的网桥 ID 值最小，所以它是根网桥。

2.2.2 应该阻塞的端口

使用 BPDU 后，网桥必须对循环进行定位，并且切断或阻塞导致循环的端口。在网桥循环中，陷入循环中的端口通常为两个或两个以上，哪个端口将处于阻塞状态呢？这是一个必须作出的决定。在桥接环境中，通常给端口分配一个称为端口成本的参数，这个参数的值通常取一个基于介质速度的默认值。Catalyst 的默认值也取决于端口的介质。这个默认值由 1000 除以每秒钟多少兆字节的介质速度而得，如以太网默认的端口成本值为 $1000/10\text{MB/s}$ 或 100MB/s 。通常情况下，网桥一般寻找使根网桥成为最经济的配置端口作为阻塞端口。这好比从 JFK 去纽约旅行。你可以搭乘出租车，也可以乘坐豪华轿车。有时坐豪华轿车更便宜，有时坐出租车更便宜。最佳的方式当然是选择成本最低的方式，除非你的身体很不舒服，而且花多少钱都无所谓。

现在，与根网桥有关的所有问题都非常清楚了。对于网桥，通向根网桥只需要一个端口。所以，与根网桥距离最近的端口或“最便宜”的端口便称为“根端口”。

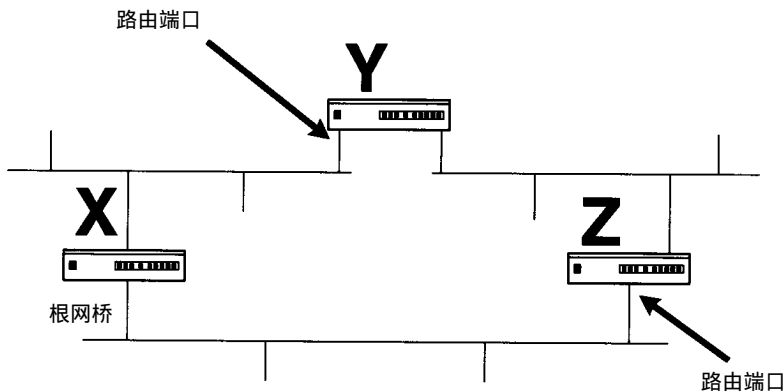


图2-13 指定端口

注意图 2-13 中 3 个网桥的情况，可以看出，与根网桥直接连接的端口将是根端口，因为它是网桥 Y 和网桥 Z 的两个端口中最“便宜”的端口。在这种情况下可能形成循环，所以网桥 Y 或网桥 Z 两个端口中的一个必须处于“阻塞”状态，以阻止循环的产生。为了确定哪一个端口应该关闭，必须知道指定端口（指定网桥）的概念。指定端口是这样一个端口，它关闭通往物理网段上的根网桥的端口。看一下图 2-13 中的第 3 个物理网段，很明显有两个端口可供选择。现在，有人会认为，这两个端口对根网桥来说是对等的。但是，决胜法是称为端口 ID 的一个参数，通常是端口的 MAC 地址。所以，在所有的物理网段上都存在指定端口（网桥），而且这个指定端口是对根而言通过 BPDU 花费最低进行发送的网桥。在端口完全对等的情况下，

MAC地址为最低的端口ID将是指定端口，因为 MAC地址在全世界都是唯一的，所以能防止无法确定指定端口的情况发生。

在图2-14中所有的端口成本都是相同的，所有 MAC地址较大的端口将是进入“阻塞”状态的端口。在这个例子中，网桥 Y的端口ID为00-00-0C-00-00-0A，它比网桥 Z的端口ID要大。所以网桥Y的端口将进入“阻塞”状态（见图 2-15）。

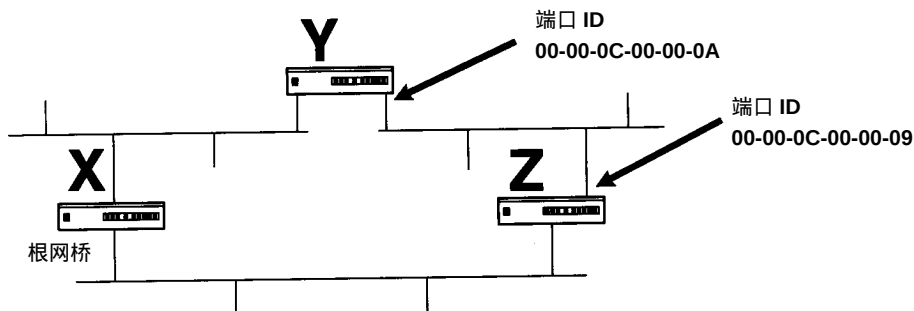


图2-14 端口ID

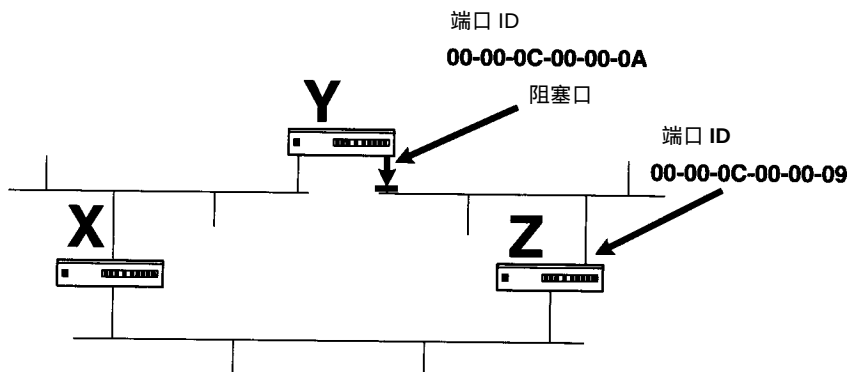


图2-15 工作站Y的网桥将被阻塞

指定端口（网桥）取决于物理网段。与根网桥进行直接连接的两个物理网段使用根网桥的两个网段作为指定端口。根网桥通常是所有物理网段上的指定端口，通过这个端口所有的物理网段都与之进行连接。

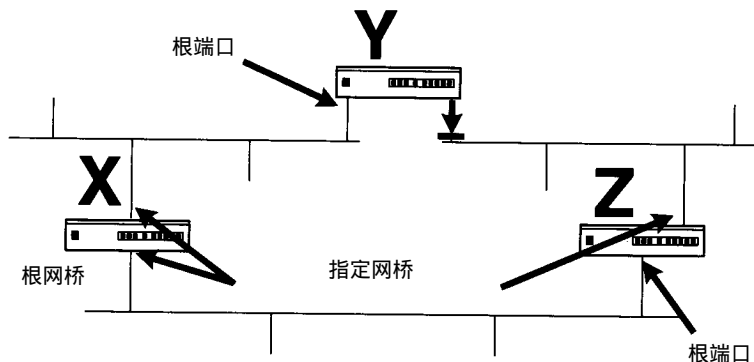


图2-16 根端口和指定网桥端口

网桥对哪个端口是根端口进行计算，然后确定这些端口中的哪一个是物理网段的指定端口，它阻塞所有的端口。在图 2-16 中，确定了所有的根端口和指定网桥，只剩下一个端口用来阻塞。记住，根网桥是它所有端口的指定网桥，所以它不会阻塞它的任何端口。

为了更进一步说明这些概念，下面来看一个更加复杂的网络环境，这有点像实际生活当中的环境。图 2-17 给出了一个网络表，在网络表中，工作站已经省去。在这种配置中产生了

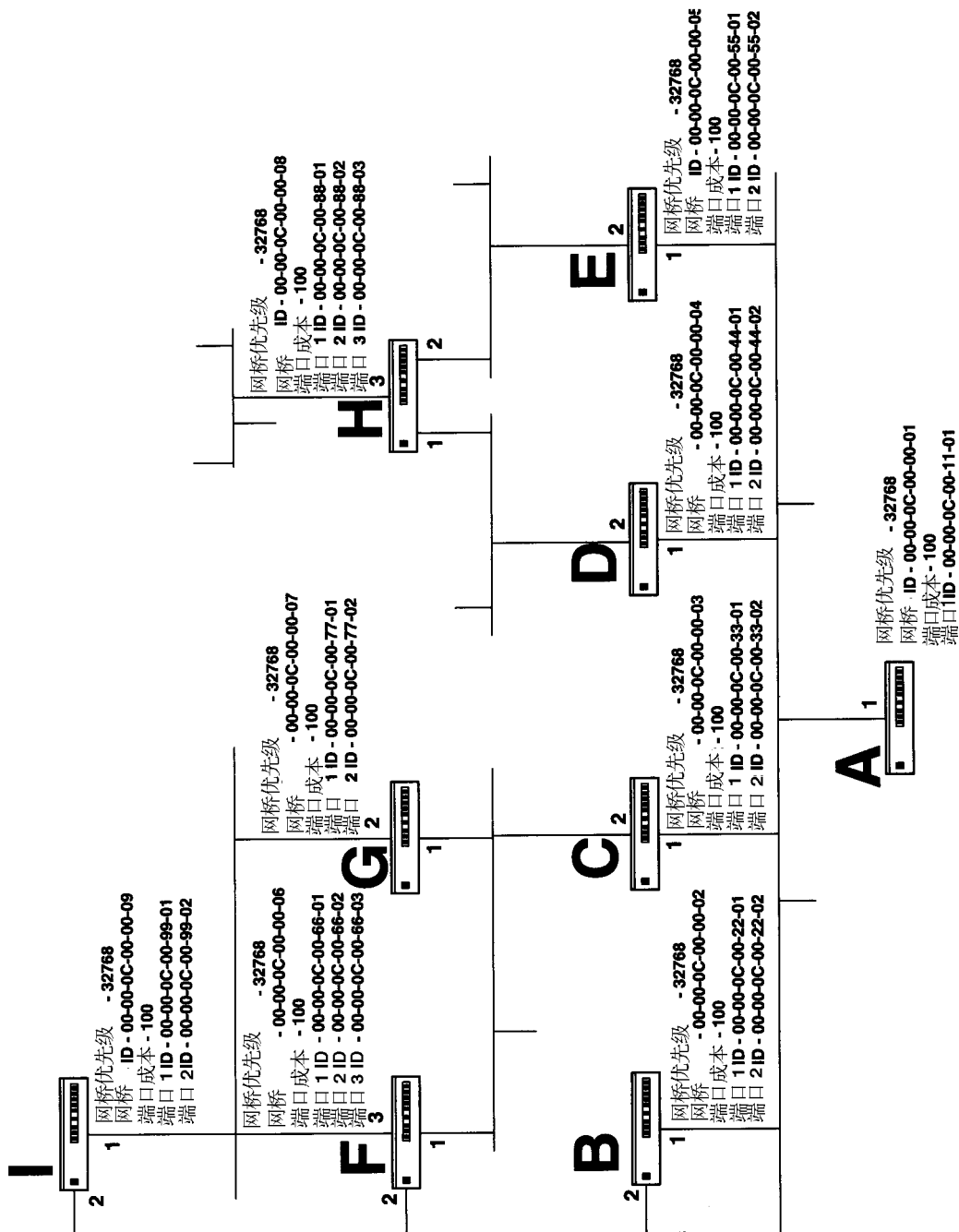


图2-17 复杂的生成树

很多的冗余，因为配置网络的组织可能非常担心会出现某些问题。

在图中列出了所有的端口 ID 和网桥 ID，还列出了网桥的优先级和端口成本。在很多情况下，这些参数作为默认值省略。在分析网络时，对每个物理网段分别进行查看是非常重要的，

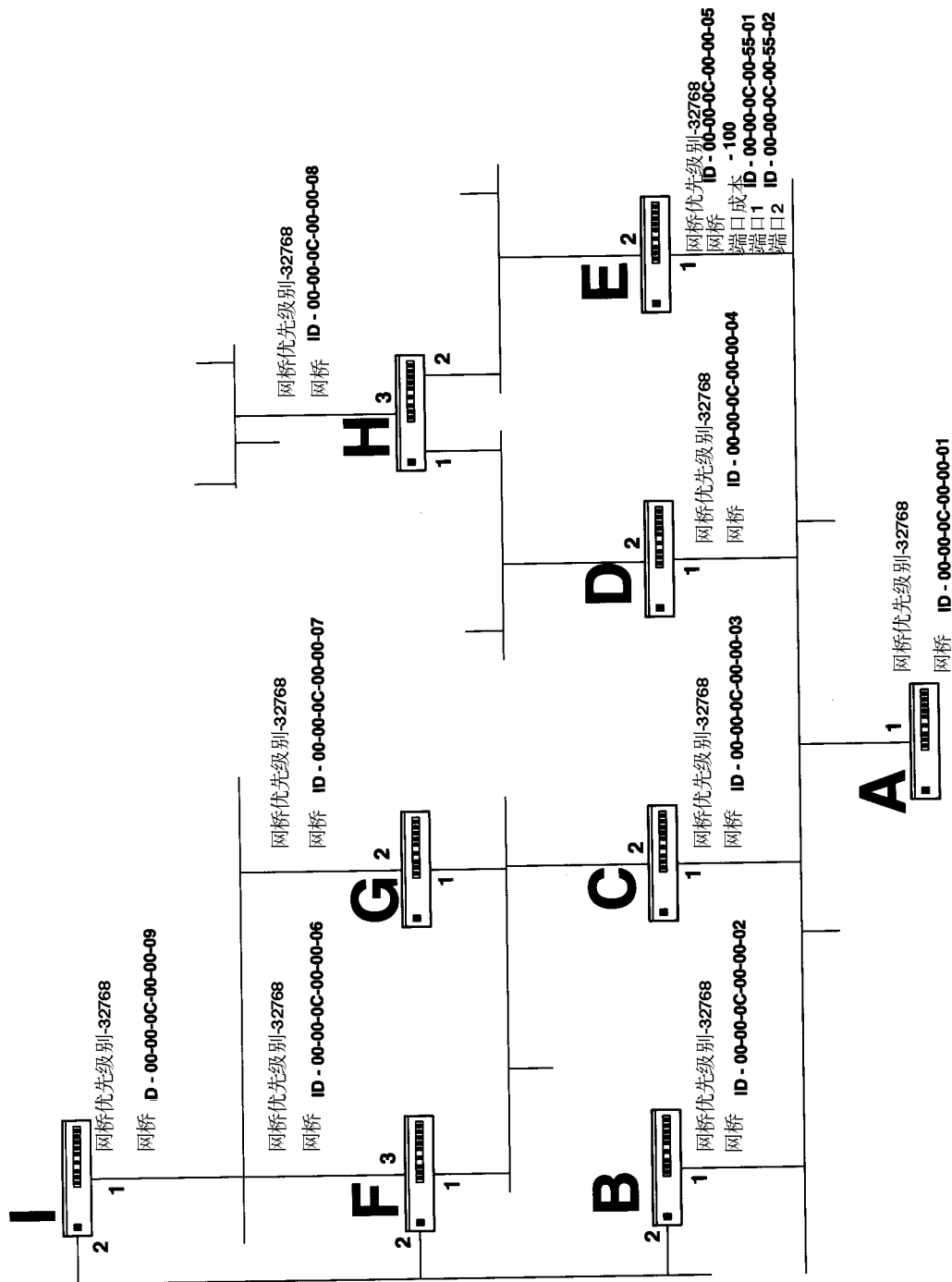


图2-18 确定根网桥

所以先将网络分开，然后对之逐个进行解决。为了确定哪个端口进行发送、哪个端口处于阻塞，建议按照下面的步骤进行。

1) 确定根网桥。网桥优先级最低的网桥将是根网桥。如果不能确定，那么网桥 ID 将作为决胜的手段。

2) 确定指定网桥。花费最低的网桥端口将是指定端口，在端口完全对等的情况下，端口 ID 将是决胜的手段。

3) 确定根端口。在所有的网桥中，哪个端口对根网桥来说花费最低，哪个端口就是根端口。

4) 确定阻塞端口。完成上述三个步骤后，就可以确定除了指定网桥或根端口以外，其他所有的端口都是阻塞端口。

5) 确定发送端口。所有的根端口和指定网桥都是发送端口。

下面对这些步骤进行详细的说明。

步骤 1：确定根网桥。这个例子，所有网桥的优先级都一样，所以应该使用网桥 ID 来确定网桥。因为网桥 A 的网桥 ID 值最小，所以它是根网桥，如图 2-18 所示。

步骤 2：确定所有的物理网段中的指定网桥。第一个物理网段是与根网桥相连的网段，所以在这个网段上确定指定网桥是最简单的。很明显，根网桥对根网桥的成本将是最低的，如图 2-19 所示。

第 2 个物理网段中有三个网桥，图 2-20 表明网桥 B 和根网桥最近。因为所有的端口成本相同，所以，到根网桥进行转发的次数最少的网桥将是指定网桥。这和距离向量路

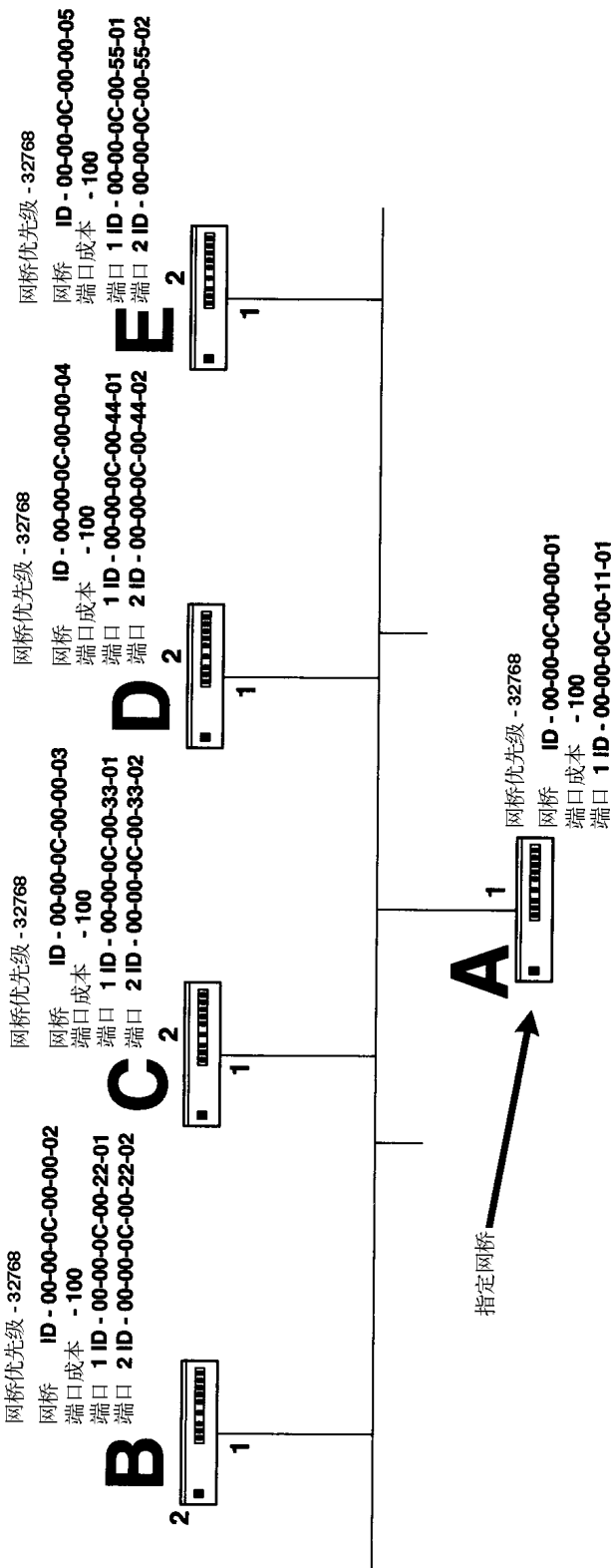


图2-19 第1个物理网段的指定网桥

由协议RIP非常相似。但是，如果端口成本不同，可以将它与链路状态路由进行比较。由于网桥B与根网桥的转发次数最少，所以它是物理网段 2 的指定网桥，如图2-20。

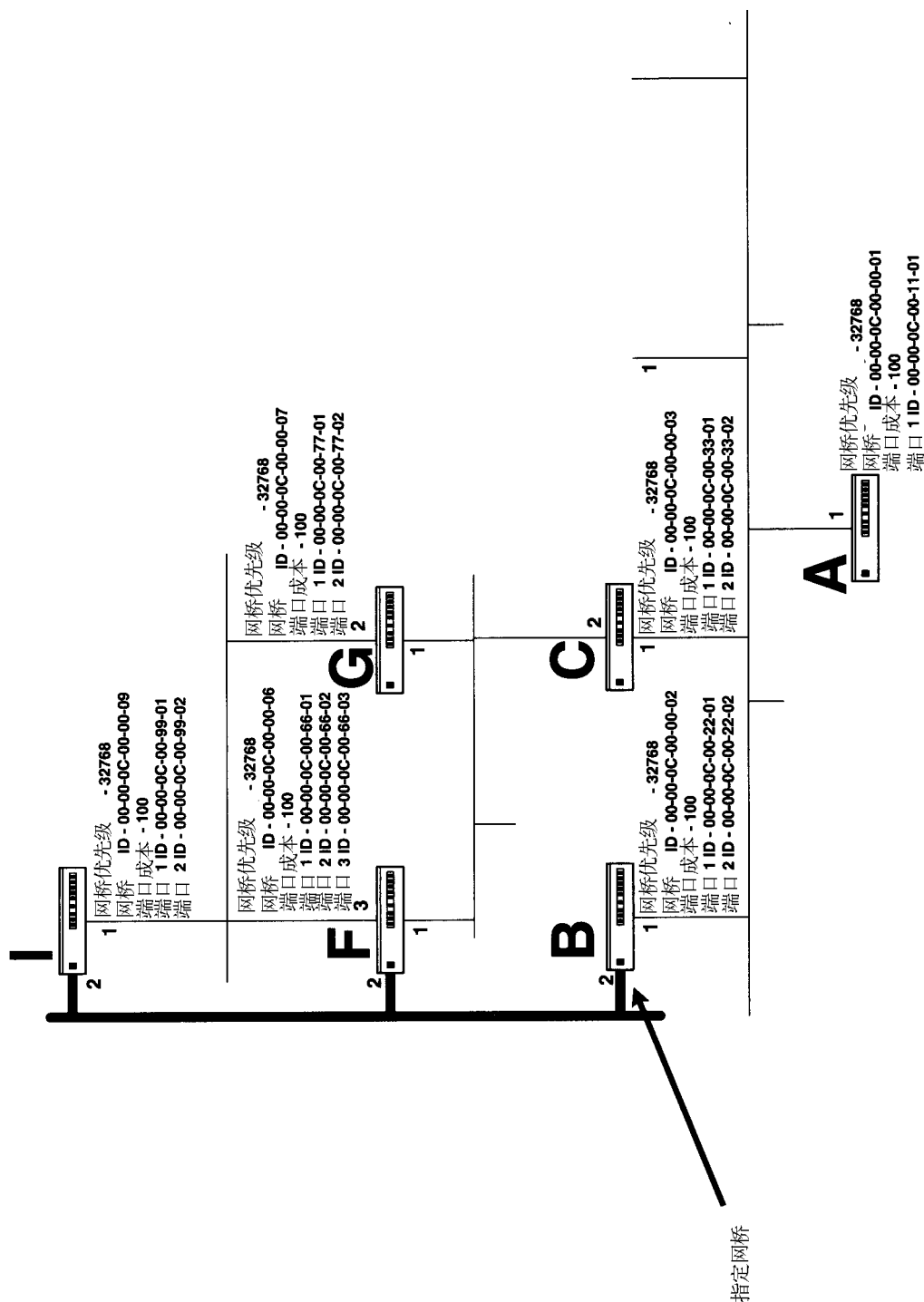


图2-20 第2个物理网段的指定网桥

第3个物理网段使用网桥 C 作为指定网桥，因为对根网桥来说，它的转发次数最少，如图 2-21 所示。

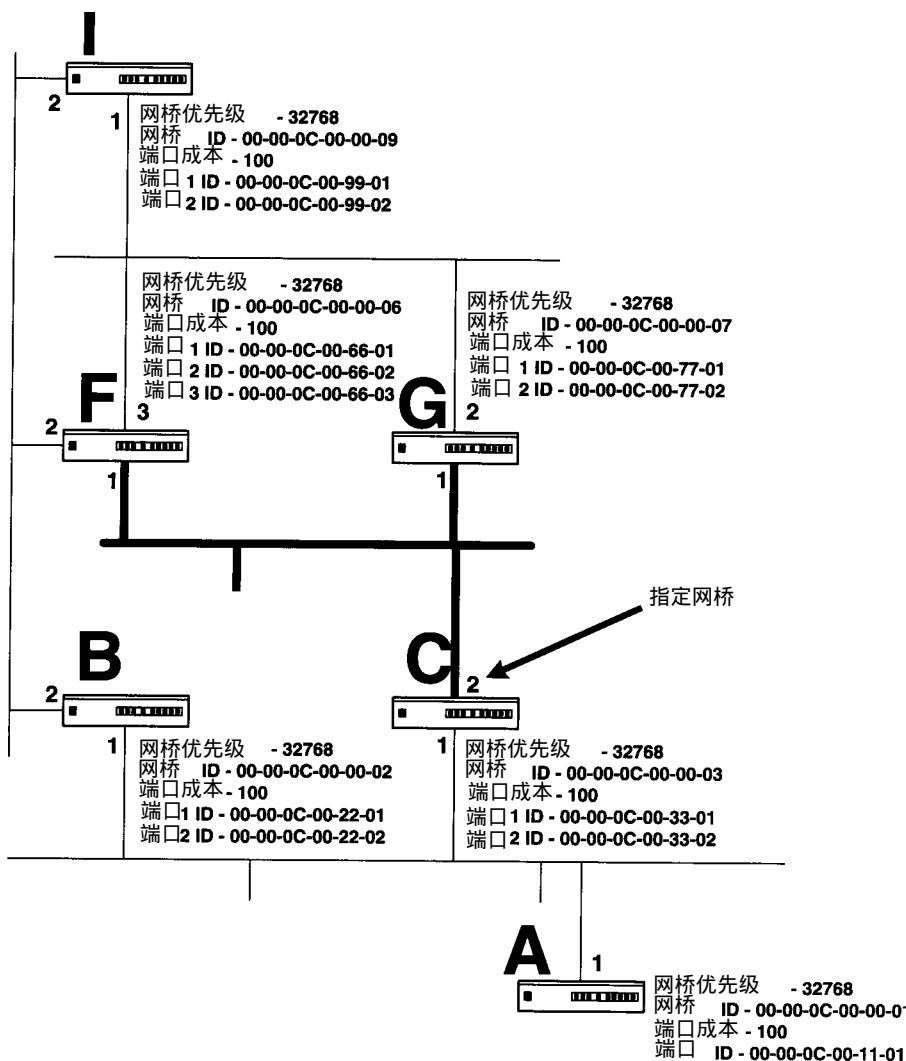


图2-21 第3个物理网段的指定网桥

第4个物理网段稍微有点复杂。网桥 F 和网桥 G 从根网桥都只有一次转发。所以在这里，将端口 ID 值作为决胜法。由于网桥 F 的端口 3 的 ID 值比网桥 G 的端口 2 的 ID 值小，所以网桥 F 将是第四个物理网段的指定网桥（见图 2-22）。这样就可以很容易地看出网桥 D 和网桥 E 分别是第 5 个物理网段和第 6 个物理网段的指定网桥（见图 2-23 和图 2-24）。

网桥 H 是第 7 个物理网段上的唯一网桥，所以它就是指定网桥（图 2-25）。

步骤 3：确定根端口。最靠近根网桥的端口是根端口。对于网桥 B、C、D 和 E 来说，这个端口比较容易确定。但是对于网桥 F，有两条路径通向根网桥，每条路径的成本相同。这时，又要用端口 ID 来做最后的决定。由于网桥 B 的端口 2 的 ID 小于网桥 C 中端口 3 的 ID，网桥 F 的根

端口为F的端口2，对于网桥H，确定根端口的的方法和F一样，由于网桥D的端口2的ID小于网桥E的端口2的ID，所以网桥H的端口1为根端口，参见图2-26。

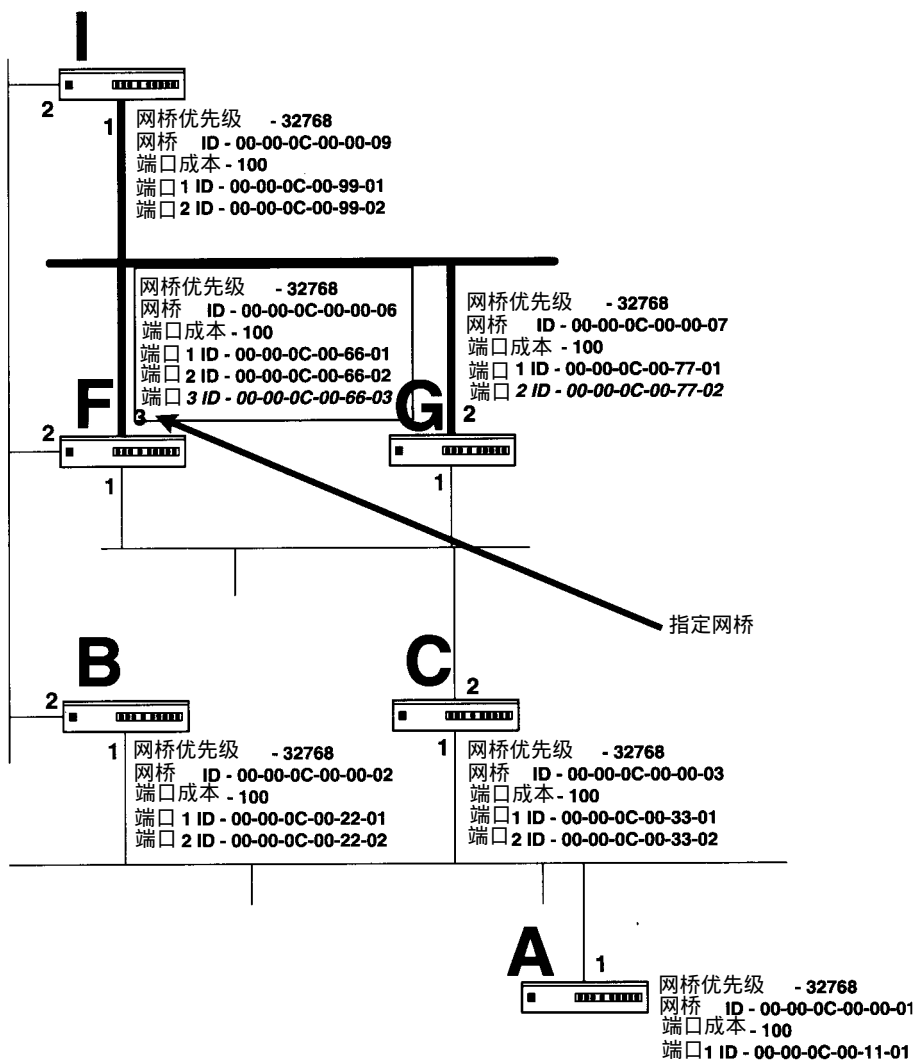


图2-22 第4个物理网段的指定网桥

步骤4：确定阻塞端口。指定网桥端口和根端口确定以后，剩下的所有端口便为阻塞端口（图2-27）。

步骤5：确定转发端口。删除所有的阻塞端口后，我们可以看到一个生成树（图 2-28）。网桥A是根网桥，所以在它的所有端口都可以进行转发。可以看出，图中形成的是“灌木”，并不是真正意义上的树。树一般只有一个树干，而对于这个生成树，它更像是一棵“夹竹桃”或一盆“杜鹃花”。但不管它像什么，它不会影响我们对此问题进行的讨论。

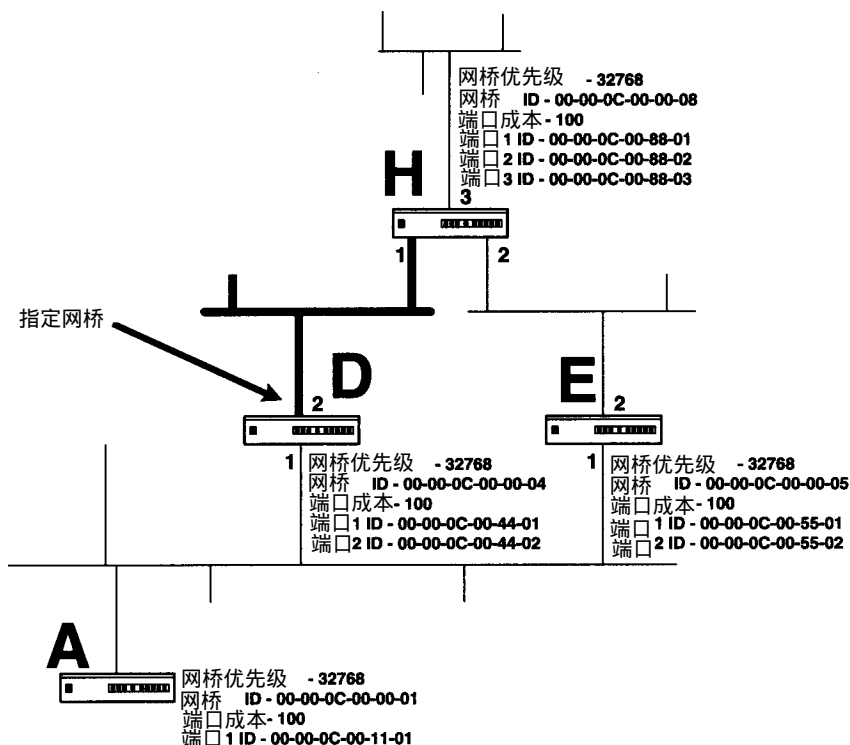


图2-23 第5个物理网段的指定网桥

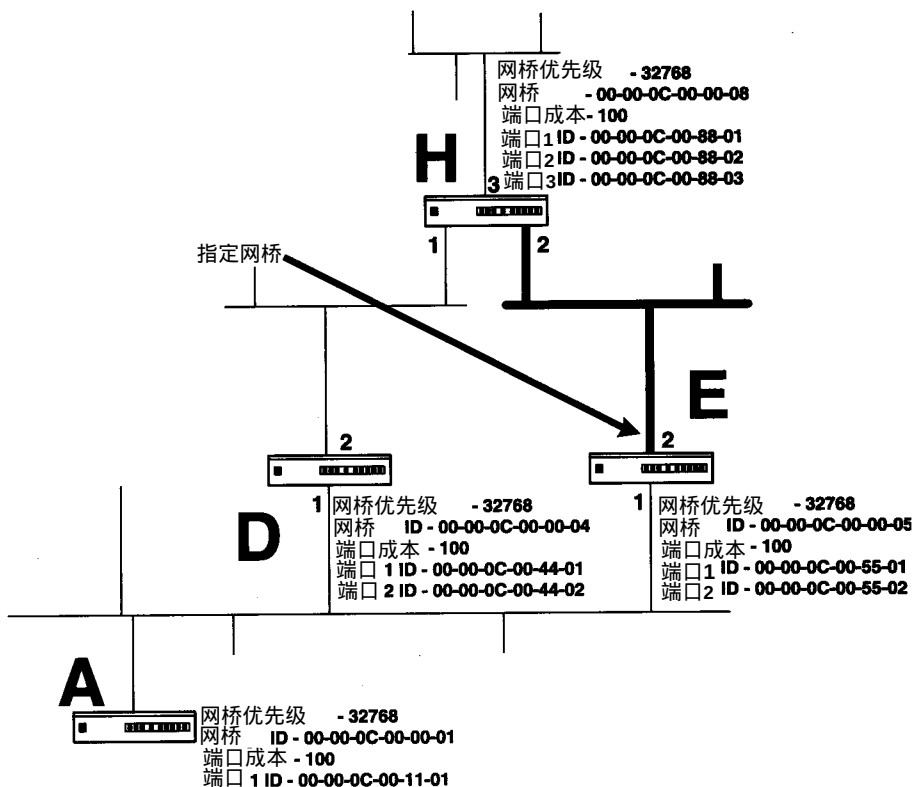


图2-24 第6个物理网段的指定网桥

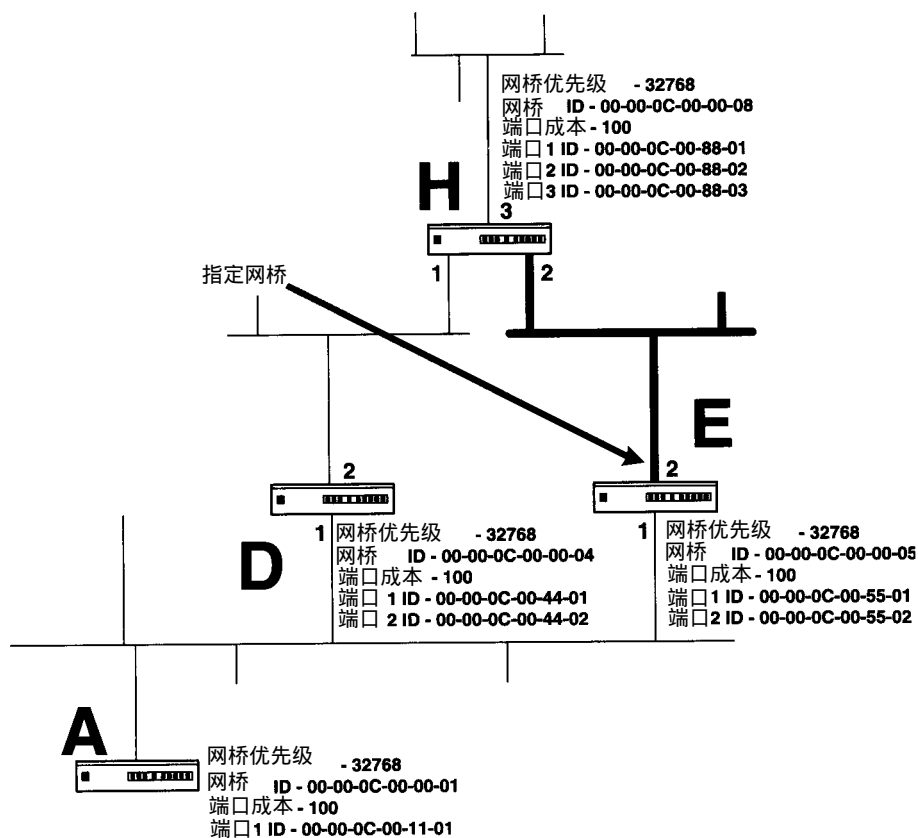


图2-25 第7个物理网段的指定网桥

2.2.3 生成树端口状态

网桥或交换机中有活动端口时（即连接灯是亮的），在它开始进行转发或被阻塞前，将经历几种状态。首先，端口进入监听状态，对 BPDU 进行监听。它将读取这些 BPDU，然后确定它的根端口和与端口相连的网桥。如果确定某个端口需要设置成阻塞状态，那么该端口将进入阻塞状态，同时监听状态结束。如果确定某个端口应该进入发送状态，那么这个端口将进入称为“获取”的中间状态，之所以这么称呼，是因为端口先对源 MAC 地址进行检测，然后再将条目导入网桥表中。“获取”状态结束后，发送状态立即开始（如图 2-29）。如果这时连接状态或生成树的拓扑发生变化，生成树的状态将从头开始。这种现象将导致短暂的网络中断。为了避免产生这种中断，建议不要在任务临界状态时对桥接环境进行实质性的改变。

当网桥或交换机的端口处在监听和获取状态时，它不会对信息进行发送，这正是将交换的网络应用到桌面时会产生许多问题的原因。工作站直接与交换机的端口相连时，不管机器是否是处于重启状态，它的连接状态都将产生变化。例如，早晨用户进入办公室后便打开他的工作站，这时交换机的端口将经历生成树状态这一过程。这个过程需要花费的时间为一分钟或取决于交换机的设置。这时，运行在工作站上的 NOS 客户软件可能会由于连接的延迟而出现一些问题。机器一旦载入这些客户软件，如 Novell DOS 或 Windows Requester 或 Windows NT 后，工作站便会寻找服务器。如果没有找到服务器，工作站将不会与它所需要的任何资源