

尽管存储转发包交换灵活且有效,但是它存在通过互连网络时增加时延(延迟)的问题。假设在图8-17中把一个包传送一跳所花费的时间为 $T$ 纳秒。为了从CPU 1到CPU 2,该包必须被复制四次(至A、至C、至D以及到目标CPU),而且在前一个包完成之前,不能开始有关的复制,所以通过该互连网络的时延是 $4T$ 。一条出路是设计一个网络,其中的包可以逻辑地划分为更小的单元。只要第一个单元到达一个交换机,它就被转发到下一个交换机,甚至可以在包的结尾到达之前进行。可以想象,这个传送单元可以小到1比特。

另一种交换机制是电路交换(circuit switching),它包括由第一个交换机建立的,通过所有交换机而到达目标交换机的一条路径。一旦该路径建立起来,比特流就从源到目的地通过整个路径不断地尽快输送。在所涉及的交换机中,没有中间缓冲。电路交换需要有一个建立阶段,它需要一点时间,但是一旦建立完成,速度就很快。在包发送完毕之后,该路径必须被拆除。电路交换的一种变种称为虫孔路由(wormhole routing),它把每个包拆成子包,并允许第一个子包在整个路径还没有完全建立之前就开始流动。

## 2. 网络接口

在多计算机中,所有节点里都有一块插卡板,它包含节点与互连网络的连接,这使得多计算机连成一体。这些板的构造方式以及它们如何同主CPU和RAM连接对操作系统有重要影响。这里简要地介绍一些有关的内容。部分内容来源于(Bhoedjang, 2000)。

事实上在所有的多计算机中,接口板上都有一些用来存储进出包的RAM。通常,在包被传送到第一个交换机之前,这个要送出的包必须被复制到接口板的RAM中。这样设计的原因是许多互连网络是同步的,所以一旦一个包的传送开始,比特流必须以恒定的速率连续进行。如果包在主RAM中,由于内存总线上有其他的信息流,所以这个送到网络上的连续流是不能保证的。在接口板上使用专门的RAM,就消除了这个问题。这种设计如图8-18所示。

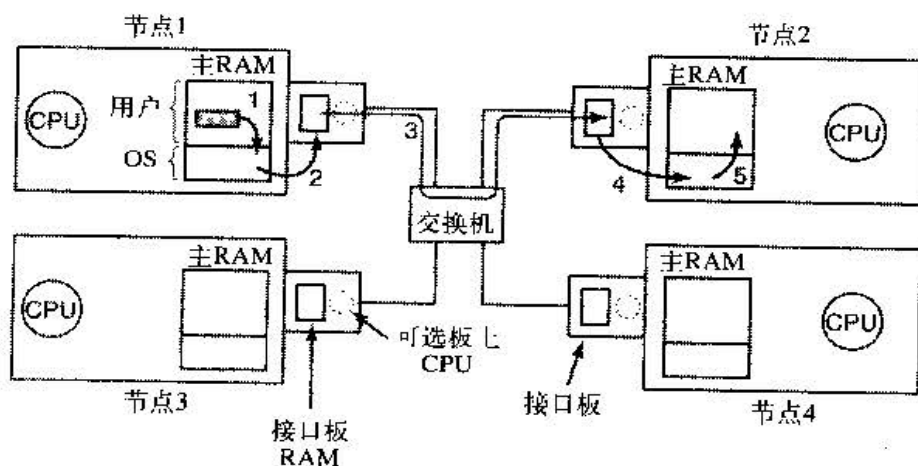


图8-18 网络接口卡在多计算机中的位置

同样的问题还出现在接收进来的包上。从网络上到达的比特流速率是恒定的,并且经常有非常高的速率。如果网络接口卡不能在它们到达的时候实时存储它们,数据将会丢失。同样,在这里试图通过系统总线(例如PCI总线)到达主RAM是非常危险的。由于网卡通常插在PCI总线上,这是一个惟一的通向主RAM的连接,所以不可避免地要同磁盘以及每个其他的I/O设备竞争总线。而把进来的包首先保存在接口板的私有RAM中,然后再把它们复制到主RAM中,则更安全些。

接口板上可以有一个或多个DMA通道,甚至在板上有一个完整的CPU(乃至多个CPU)。通过请求在系统总线上的块传送(block transfer),DMA通道可以在接口板和主RAM之间以非常高的速率复制包,因而可以一次性传送若干字而不需要为每个字分别请求总线。不过,准确地说,正是这种块传送(它占用了系统总线的多个总线周期)使接口板上的RAM的需要是第一位的。

很多接口板上有一个完整的CPU,可能另外还有一个或多个DMA通道。它们被称为网络处理器(network processor),并且其功能日趋强大。这种设计意味着主CPU将一些工作分给了网卡,诸如处理