



图8-30 在分布式系统中中间件的地位

在某种意义上，中间件像是分布式系统的操作系统。这就是为什么在一本关于操作系统的书中讨论中间件的原因。不过另一方面，中间件又不是真正的操作系统，所以我们对中间件有关的讨论不会过于详细。较为全面的关于分布式系统的讨论可参见《分布式系统》(Distributed Systems, Tanenbaum 和 van Steen, 2006)。在本章余下的部分，首先我们将快速考察在分布式系统（下层的计算机网络）中使用的硬件，然后是其通信软件（网络协议）。接着我们将考虑在这些系统中的各种范型。

8.4.1 网络硬件

分布式系统构建在计算机网络的上层，所以有必要对计算机网络这个主题做个简要的介绍。网络主要有两种，覆盖一座建筑物或一个校园的LAN（局域网，Local Area Networks）和可用于城市、乡村甚至世界范围的WAN（广域网，Wide Area Network）。最重要的LAN类型是以太网（Ethernet），所以我们把它作为LAN的范例来考察。至于WAN的例子，我们将考察Internet，尽管在技术上Internet不是一个网络，而是上千个分离网络的联邦。但是，就我们的目标而言，把Internet视为一个WAN就足够了。

1. 以太网（Ethernet）

经典的以太网，在IEEE802.3标准中有具体描述，由用来连接若干计算机的同轴电缆组成。这些电缆之所以称为以太网（Ethernet），是源于发光以太，人们曾经认为电磁辐射是通过以太传播的。（19世纪英国物理学家James Clerk Maxwell发现了电磁辐射可用一个波动方程描述，那时科学家们假设空中必须充满了某些以太介质，而电磁辐射则在以太介质中传播。不过在1887年著名的Michelson-Morley实验中，科学家们并未能探测到以太的存在，在这之后物理学家们才意识到电磁辐射可以在真空中传播）。

在以太网的非常早的第一个版本中，计算机与钻了半截孔的电缆通过一端固定在这些孔中而另一端与计算机连接的电线相连接。它们被称为插入式分接头（vampire tap），如图8-31a中所示。可是这种接头很难接正确，所以没过多久，就换用更合适的接头了。无论怎样，从电气上来看，所有的计算机都被连接起来，在网络接口卡上的电缆仿佛是被焊上一样。

要在以太网上发送包，计算机首先要侦听电缆，看看是否有其他的计算机正在进行传输。如果没有，这台计算机便开始传送一个包，其中有一个短包头，随后是0到1500字节的有效信息载荷（payload）。如果电缆正在使用中，计算机只是等待直到当前的传输结束，接着该台计算机开始发送。

如果两台计算机同时开始发送，就会导致冲突发生，两台机器都做检测。两机都用中断其传输来响应检测到的碰撞，然后在等待一个从0到 T 微秒的随机时间段之后，再重新开始。如果再一次冲突发生，所有碰撞的计算机进入0到 $2T$ 微秒的随机等待。然后再尝试。在每个后续的冲突中，最大等待间隔加倍，用以减少更多碰撞的机会。这个算法称为二进制指数补偿算法（binary exponential backoff）。在前面有关减少锁的轮询开销中，我们曾介绍过这种算法。

以太网有其最大电缆长度限制，以及可连接的最多的计算机台数限制。要想超过其中一个的限制，就要在一座大建筑物或校园中连接多个以太网，然后用一种称为桥接器（bridge）的设备把这些以太网