

即分布式系统 (distributed system)。这些系统与多计算机类似, 每个节点都有自己的私有存储器, 整个系统中没有共享的物理存储器。但是, 分布式系统与多计算机相比, 耦合更加松散。

首先, 一台多计算机的节点通常有CPU、RAM、网卡, 可能还有用于分页的硬盘。与之相反, 分布式系统中的每个节点都是一台完整的计算机, 带有全部的外部设备。其次, 一台多计算机的所有节点一般就在一个房间里, 这样它们可以通过专门的高速网络通信, 而分布式系统中的节点则可能分散在全世界范围内。最后, 一台多计算机的所有节点运行同样的操作系统, 共享一个文件系统, 并处在一个共同的管理之下, 而一个分布式系统的节点可以运行不同的操作系统, 每个节点有自己的文件系统, 并且处在不同的管理之下。一个典型的多计算机的例子如一个公司或一所大学的一个房间中用于诸如药物建模等工作的512个节点, 而一个典型的分布式系统包括了通过Internet松散协作的上千台机器。在图8-29中, 对多处理机、多计算机和分布式系统就上述各点进行了比较。

项目	多处理机	多计算机	分布式系统
节点配置	CPU	CPU、RAM、网络接口	完整的计算机
节点外设	全部共享	共享exc., 可能除了磁盘	每个节点全套外设
位置	同一机箱	同一房间	可能全球
节点间通信	共享RAM	专用互连	传统网络
操作系统	一个, 共享	多个, 相同	可能都不相同
文件系统	一个, 共享	一个, 共享	每个节点自有
管理	一个机构	一个机构	多个机构

图8-29 三类多CPU系统的比较

通过这个表可以清楚地看到, 多计算机处于中间位置。于是一个有趣的问题就是: “多计算机是更像多处理机还是更像分布式系统?” 很奇怪, 答案取决于你的角度。从技术角度来看, 多处理机有共享存储器而其他两类没有。这个差别导致了不同的程序设计模式和不同的思考方式。但是, 从应用角度来看, 多处理机和多计算机都不过是在机房中的大设备机架 (rack) 罢了, 而在全部依靠Internet连接计算机的分布式系统中显然通信要多于计算, 并且以不同的方式使用着。

在某种程度上, 分布式系统中计算机的松散耦合既是优点又是缺点。它之所以是优点, 是因为这些计算机可用在各种类型的应用之中, 但它也是缺点, 因为它由于缺少共同的底层模型而使得这些应用程序很难编程实现。

典型的Internet应用有远程计算机访问 (使用telnet、ssh和rlogin)、远程信息访问 (使用万维网 (World Wide Web) 和FTP, 即文件传输协议)、人际通信 (使用e-mail和聊天程序) 以及正在浮现的许多应用 (例如, 电子商务、远程医疗以及远程教育等)。所有这些应用带来的问题是, 每个应用都得重新开发。例如, e-mail、FTP和万维网基本上都是将文件从A点移动到另一个点B, 但是每一种应用都有自己的方式从事这项工作, 完全按照自己的命名规则、传输协议、复制技术以及其他等。尽管许多Web浏览器对普通用户隐藏了这些差别, 但是底层机制仍然是完全不同的。在用户界面级隐藏这些差别就像有一个人一家提供全面服务的旅行社的Web站点中预订了从纽约到旧金山的旅行, 后来发现她所购买的只不过是—张飞机票、一张火车票或者一张汽车票而已。

分布式系统添加在其底层网络上的是一些通用范型 (模型), 它们提供了一种统一的方法来观察整个系统。分布式系统想要做的是, 将松散连接的大量机器转化为基于—种概念的一致系统。这些范型有的比较简单, 而有的是很复杂的, 但是其思想则总是提供某些东西用来统一整个系统。

在上下文稍有差别的情形下, 统一范例的一个简单例子可以在UNIX中找到。在UNIX中, 所有的I/O设备被构造成像文件一样。对键盘、打印机以及串行通信线等都使用相同的方式和相同的原语进行操作, 这样, 与保持原有概念上的差异相比, 对它们的处理更为容易。

分布式系统面对不同硬件和操作系统实现某种统一性的途径是, 在操作系统的顶部添加一层软件。这层软件称为中间件 (middleware), 如图8-30所示。这层软件提供了一些特定的数据结构和操作, 从而允许散布的机器上的进程和用户用一致的方式互操作。