

存储器访问时间较长。在这种设计中没有全局共享的存储器。多计算机（消息传递系统）比（共享存储器）多处理机系统容易构建，但是编程比较困难。可见，每种类型各有其优点。

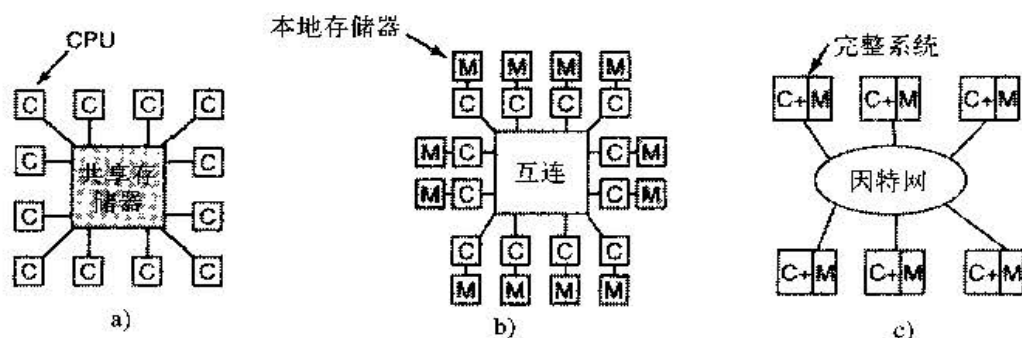


图8-1 a) 共享存储器多处理机；b) 消息传递多计算机；c) 广域分布式系统

第三种模型参见图8-1c，所有的计算机系统都通过一个广域网连接起来，如因特网，构成了一个分布式系统（distributed system）。每台计算机有自己的存储器，当然，通过消息传递进行系统通信。图8-1b和图8-1c之间真正惟一的差别是，后者使用了完整的计算机而且消息传递时间通常需要10~100ms。如此长的延迟造成使用这类松散耦合系统的方式和图8-1b中的紧密耦合系统不同。三种类型的系统在通信延迟上各不相同，分别有三个数量级的差别。类似于一天和三年的差别。

本章有四个主要部分，分别对应于图8-1中的三个模型再加上虚拟化技术（一种通过软件创造出更多虚拟CPU的方法）。在每一部分中，我们先简要地介绍相关的硬件。然后，讨论软件，特别是与这种系统类型有关的操作系统问题。我们会发现，每种情况都面临着不同的问题并且需要不同的解决方法。

8.1 多处理机

共享存储器多处理机（或以后简称为多处理机，multiprocessor）是这样一种计算机系统，其两个或更多的CPU全部共享访问一个公用的RAM。运行在任何一个CPU上的程序都看到一个普通（通常是分页）的虚拟地址空间。这个系统惟一特别的性质是，CPU可对存储器字写入某个值，然后读回该字，并得到一个不同的值（因为另一个CPU改写了它）。在进行恰当组织时，这种性质构成了处理器间通信的基础：一个CPU向存储器写入某些数据而另一个读取这些数据。

至于最重要的部分，多处理机操作系统只是通常的操作系统。它们处理系统调用，进行存储器管理，提供文件系统并管理I/O设备。不过，在某些领域里它们还是有一些独特的性质。这包括进程同步、资源管理以及调度。下面首先概要地介绍多处理机的硬件，然后进入有关操作系统的问题。

8.1.1 多处理机硬件

所有的多处理机都具有每个CPU可访问全部存储器的性质，而有些多处理机仍有一些其他的特性，即读出每个存储器字的速度是一样快的。这些机器称为UMA（Uniform Memory Access，统一存储器访问）多处理机。相反，NUMA（Nonuniform Memory Access，非一致存储器访问）多处理机就没有这种特性。至于为何有这种差别，稍后会加以说明。我们将首先考察UMA多处理机，然后讨论NUMA多处理机。

1. 基于总线的UMA多处理机体系结构

最简单的多处理机是基于单总线的，参见图8-2a。两个或更多的CPU以及一个或多个存储器模块都使用同一个总线进行通信。当一个CPU需要读一个存储器字（memory word）时，它首先检查总线忙否。如果总线空闲，该CPU把所需字的地址放到总线上，发出若干控制信号，然后等待存储器把所需的字放到总线上。

当某个CPU需要读写存储器时，如果总线忙，CPU只是等待，直到总线空闲。这种设计存在问题。在只有两三个CPU时，对总线的争夺还可以管理，若有32个或64个CPU时，就不可忍受了。这种系统完全受到总线带宽的限制，多数CPU在大部分时间里是空闲的。

这一问题的解决方案是为每个CPU添加一个高速缓存（cache），如图8-2b所示。这个高速缓存可以位于CPU芯片的内部、CPU附近、在处理器板上或所有这三种方式的组合。由于许多读操作可以从本地