

第8章 多处理机系统

从计算机诞生之日起,人们对更强计算能力的无休止的追求就一直驱使着计算机工业的发展。ENIAC可以完成每秒300次的运算,它一下子就比以往任何计算器都快1000多倍,但是人们并不满足。我们现在有了比ENIAC快数百万倍的机器,但是还有对更强大机器的需求。天文学家们正在了解宇宙,生物学家正在试图理解人类基因的含义,航空工程师们致力于建造更安全和速度更快的飞机,而所有这一切都需要更多的CPU周期。然而,即使有更多运算能力,仍然不能满足需求。

过去的解决方案是使时钟走得更快。但是,现在开始遇到对时钟速度的限制了。按照爱因斯坦的相对论,电子信号的速度不可能超过光速,这个速度在真空中大约是 30cm/ns ,而在铜线或光纤中约是 20cm/ns 。这在计算机中意味着 10GHz 的时钟,信号的传送距离总共不会超过 2cm 。对于 100GHz 的计算机,整个传送路径长度最多为 2mm 。而在一台 1THz (1000GHz)的计算机中,传送距离就不足 $100\mu\text{m}$ 了,这在一个时钟周期内正好让信号从一端到另一端并返回。

让计算机变得如此之小是可能的,但是这会遇到另一个基本问题:散热。计算机运行得越快,产生的热量就越多,而计算机越小就越难散热。在高端Pentium系统中,CPU的散热器已经比CPU自身还要大了。总而言之,从 1MHz 到 1GHz 需要的是更好的芯片制造工艺,而从 1GHz 到 1THz 则需要完全不同的方法。

获得更高速度的一种处理方式是大规模使用并行计算机。这些机器有许多CPU,每一个都以“通常”的速度(在一个给定年份中的速度)运行,但是总体上会有比单个CPU强大得多的计算能力。具有1000个CPU的系统已经商业化了。在未来十年中,可能会建造出具有100万个CPU的系统。当然为了获得更高的速度,还有其他潜在的处理方式,如生物计算机,但在本章中,我们将专注于有多个普通CPU的系统。

在高强度的数据处理中经常采用高度并行计算机。如天气预测、围绕机翼的气流建模、世界经济模拟或理解大脑中药物-受体的相互作用等问题都是计算密集型的。解决这些问题需要多个CPU同时长时间运行。在本章中讨论的多处理机系统被广泛地用于解决这些问题以及其他科学、工程领域中的类似问题。

另一个相关的进展是因特网不可思议地快速增长。因特网最初被设计为一个军用的容错控制系统的原型,然后在从事学术研究的计算机科学家中流行开来,并且在过去它已经获得了许多新用途。其中一种用途是,把全世界的数千台计算机连接起来,共同处理大型的科学问题。在某种意义上,一个包含有分布在全世界的1000台计算机的系统与在一个房间中有1000台计算机的系统之间没有差别,尽管这两个系统在延时和其他技术特征方面会有所不同。在本章中我们也将讨论这些系统。

假如有足够多的资金和足够大的房间,把一百万台无关的计算机放到一个房间中很容易做到。把一百万台无关的计算机放到全世界就更容易了,因为不存在第二个问题了。当要在一个房间中使这些计算机相互通信,以便共同处理一个问题时,问题就出现了。结果,人们在互连技术方面做了大量工作,而且不同的互连技术已经导致了不同性质的系统以及不同的软件组织。

在电子(或光学)部件之间的所有通信,归根结底是在它们之间发送消息——具有良好定义的位串(bit string)。其差别在于所涉及的时间范围、距离范围和逻辑组织。一个极端的例子是共享存储器多处理机,系统中有从2个到1000个的CPU通过一个共享存储器通信。在这个模型中,每个CPU可同样访问整个物理存储器,可使用指令LOAD和STORE读写单个的字。访问一个存储器字通常需要 $2\sim 10\text{ns}$ 。尽管这个模型,如图8-1a所示,看来很简单,但是实际上要实现它并不那么简单,而且通常涉及底层大量的消息传递,这一点我们会简要地加以说明。不过,该消息传递对于程序员来说是不可见的。

其次是图8-1b中的系统,许多CPU-存储器通过某种高速互连网络连接在一起。这种系统称为消息传递型多计算机。每个存储器局部对应一个CPU,且只能被该CPU访问。这些CPU通过互连网络发送多字消息通信。存在良好的连接时,一条短消息可在 $10\sim 50\mu\text{s}$ 之内发出,但是这仍然比图8-1a中系统的存