

第2章、 并行计算概念及其所需的网络服务

读者已经从第一章得到了一些关于微机集群的一个初步印象：微机集群就是用网络设备连接起来的一群微机。性急的读者也许会问：自己所在的实验室已有很多安装了 Linux 系统的微机连接在网上，这些微机是不是属于微机集群？

它们可以属于微机集群，也可以不属于微机集群。

是否微机集群并不在于连接在网上微机数量的多少，而在于微机互相连接的网络结构和作此连接的用途，以及为此用途所作的系统上的设置。从网络结构上说，这样的一群微机已经可以作为一个微机集群，但它们还需满足一些其他的条件：

1. 实验室的一群微机一般除了普通的网络连接和设置之外，还没有为并行计算所需的网络功能进行过专门的设置，也没有安装并行计算所必需使用的一些软件或说并行计算环境（平台），所以还需要对它们进行必要的系统设置和软件安装，使它们能被用作并行计算。
2. 实验室的一群微机一般是用集线器（hub）进行网络连接，而微机集群通常用交换机（switch）进行网络连接（集线器与交换机的差别可以参见附录关于微机集群硬件的介绍）。虽然两者的连接在物理结构上是完全相似的，但是集线器很难满足微机集群同时迸发的大量数据通讯要求的。而这样的大数据量通讯又是在采用高速微处理器的微机集群的并行计算中经常出现的。除非专门的并行计算用途，而且这种并行计算应用又是数据通讯量很小的，否则应采用交换机。当然究竟用什么网络设备进行连接还得看微机集群的用途和投资。
3. 实验室的一群微机一般情况下每台都配有显示器、键盘、鼠标等，即每台都可以独立地用作任何用途，而实际上它们也被用作并行计算以外的其他用途。从使用效率上说，微机集群中的每台微机应该只作并行计算的用途。否则，这样的分布式并行计算系统的负载或者太沉重，或者太空闲，使用效率会十分低下。微机集群中的节点机都有固定用途，一般通过服务器（或主控节点机）控制节点机，所以节点机都不需配置显示器、键盘、鼠标等。

在未满足上述条件之前，它们还没有组成真正意义上的微机集群，还不适合被用作并行计算。因此，实验室中连接在网上的一群安装了 Linux 的微机还不能被称作微机集群。但是，通过上面的解释，读者也许可以有一个微机集群的更清晰的图象——微机集群是用快速网络设备连接在一起的、微机数量可扩展的、专门用作并行计算的一群微机。

此话怎讲？这就要从在这种分布式并行计算机上作并行计算的特点说起——组建微机集群最好从了解什么是并行计算以及分布式并行计算机的结构开始。

同时，在动手具体组建微机集群之前，也需要对微机集群这种类型的并行计算机的工作原理有个大致的了解。微机集群从网络结构上来说，是一种分布存储式的并行计算机。因此，本章的主要内容就是介绍并行计算的基本概念和微机集群的网络结构。