

➤ 分布式数据库的本地性或近地性

分布式数据库系统中最重要目标（至少在使用广域网的情况下，以及某些使用局域网的情况下）是尽量减少对网络的利用，即尽可能减少站点之间的通信次数和通信量。所以，分布式数据库设计中的一个主要原则是使数据和应用实现最大程度的本地性。虽然分布式数据库管理系统能使站点间的通信更加完善，但是开发一个分布式数据库的主要目的是：通过尽可能地使数据靠近使用该数据的应用进行分配，从而提高处理的本地性或近地性，减少通信。因此，在一个精心设计的分布式数据库中，“90%的数据应当在本地站点找到，而只有10%的数据需要在远程站点进行访问。”本地性的优点不仅仅减少了远程访问的次数，而且简化了对该应用执行的控制。

➤ 控制数据的适当冗余

控制数据的适当冗余是分布式数据库系统设计的又一个目标。在分布式数据库系统中，为了提高系统的本地性、并发度和可靠性，需要增加数据的副本。这不仅使应用具有高度的可用性和本地性，而且当数据的任何一个副本不能使用时，可方便地使用在另一站点中的该数据的副本进行恢复，从而提高系统的可靠性。然而另一方面，为维护数据的一致性，减少同步更新的开销，需要控制数据副本的增加。因此，设计者会面临种种折中考验：数据或数据片段要不要复制？复制多少个副本最为合适？等等。

➤ 工作负荷分布

分布式计算机系统的一个重要特征是把工作负荷分布在网络中的各个站点上。分布工作负荷的目的是充分利用每个站点计算机的能力和资源，以提高应用执行的平行程度，从而提高系统的性能。因为工作负荷分布对处理本地性具有相当的影响，所以需要在数据分布设计时考虑到两者之间的矛盾。

➤ 存储的能力和费用

数据库的分布会受到各站点的存储能力的影响。在网络中可以有专用于存储数据的站点，也可以有完全不支持大容量存储的站点。一般来说，数据存储的费用与 CPU，I/O 及传输的费用相比是不重要的，但是必须考虑各站点可用存储空间的限制。