

"油藏数值模拟"在 973 项目中的进展

油藏模拟是在高性能计算机上数值模拟油气资源在开采过程中的复杂流动,以了解和控制油气田的生产动态,为控制和选择优化的开发方案提供科学依据。精细油藏数值模拟新技术对于研究剩余油分布,促进我国相当部分已进入高含水后期老油田的高产、稳产具有重要意义。

精细油藏模拟实际需要百万~千万量级网格点,历史拟合 20 年以上。作为生产需要,希望八小时之内完成一次模拟。1998 年立项时,利用某串行油藏模拟商业软件,在 Power Challenge 机上试算大庆百万网格点问题,经过两个月的运算,模拟结果尚未完成。2001 年大庆油田在引进的 SGI2000 上用 8 个进程的商业并行软件 PVIP2000 模拟同样问题,计算时间超过 70 个小时。PVIP2000 并行软件为美国 Harriburton-Landmark 石油服务公司的商业油藏模拟软件,它的核心软件和求解技术由 Argonne 实验室和 the University of Texas-Austin 完成。

在获得国家科技进步二等奖的“高性能分布式并行数值代数软件研究与开发”项目中,我们于 1999 年在多个国产并行机环境下用自主研发的并行软件 PRISV1.1 为大庆油田实现了三维三相百万网格点生产数据的长时间模拟,在 16 处理器机群系统上的模拟速度为 63.9 小时,达到与进口并行模拟系统 PVIP+Origin2000 (10 个处理器)相当的水平;2000 年在曙光 3000 上用 16 处理器运行 PRISV1.1,求解同一问题,历时 13.97 小时(其通讯时间 3.86 小时),计算速度为进口商业模拟系统的 5 倍以上,这成为 2000 年度获得的最佳绝对计算性能。

在国家 973 项目启动时,在高性能计算机如曙光 3000 上,即使使用 32~1024 个处理器求解大庆百万节点问题,计算时间也不会小于 4 个小时,主要制约因素在于通讯瓶颈。当时的情况是,利用现成并行软件,任何高性能并行计算机对大庆百万节点问题的加速比上限为 $60 \times 24 / 4 = 360$ 倍,曙光 3000 的理论加速比上限估计为 $60 \times 24 / 10 = 144$ 倍,曙光 2000 的理论加速比上限估计为 $60 \times 24 / 24 = 60$ 倍,以太网微机机群上的理论加速比上限估计为 $60 \times 24 / 40 = 36$ 倍。因此,如何大幅度减少全局通讯频率,如何进一步提高非线性求解器的计算效率和效果,如何研制出在 32~1024 处理器的微机机群上都具有良好可扩展性的并行软件,成为大规模精细油藏数值模拟的当务之急。在 973 项目立项时,我们制定的目标是,在微机机群上,达到两个数量级的加速比,其中,算法的改进和并行实现技术贡献一个数量级。

基于在 973 项目中的并行预条件子算法、非线性问题求解技术和并行实现等研究成果,我们于 2002 年提出了由八种不同性质预条件组成的并行预条件子系统,并应用于油藏数值模拟软件,从而改进了并行油藏模拟软件的算法与性能,再加上基于全局通讯极小化的并行实现,形成了 PRISV2.1。

以模拟大庆油田 116 万网格点的历史的实际数据为例,在 LSSC-II 的 Cluster 机群上 (2.0GHZ 主频, Myrinet 网络),利用 64 个节点运行 64 个进程,历时 0.94 小时。与 1998 年立项时相比,提高绝对计算性能约 1600 倍,其中机器主频加速 5 倍,64 个节点的并行软件实现技术加速不超过 64 倍,显然,非线性求解器在算法上贡献的加速至少 5 倍。

与国家科技进步二等奖的软件相比，现有并行软件的绝对性能提高了 10 倍以上，可扩展性已经达到 256 处理器。并且，我们将该软件在高性能并行计算机上的理论加速比上限估计提升为 $60 \times 24 / 0.4 = 3600$ 倍，LSSC-II 上的理论加速比上限估计提升为 $60 \times 24 / 0.6 = 2400$ 倍。更为重要的是，根据这几年在大规模科学计算研究与实现方面的经验，针对求解耦合偏微分方程组（如 CFD 等）的 Newton-Krylov-Schwarz 方法，我们已经在算法应用和并行实现方面获得了系列成果，经验和教训给我们提供了在超大规模并行计算（512~2048 处理器上的并行算法和并行程序设计思路）中的研究思路和实现方案。这些知识的积累，为我们在高性能计算机上的进一步性能提高打下了坚实的基础。