

GPU 助力普林斯顿海洋模式 (POM) 区域海洋模拟业务

背景

地球系统数值模拟教育部重点实验室于2010年成立，通过联合清华大学地学中心、计算机系、环境系及数学系相关学科，开展地球系统数值模拟研究中的物理模型、数学模型、计算方法和计算支撑技术及应用等方面的研究。

地球系统数值模拟是研究全球变化的重要方法，而高性能计算又是地球系统数值模拟的重要技术支撑。GPU作为一种新兴的异构高性能计算平台，提供了很高的计算能力和显存带宽，但却又使用了较低的功耗。使用GPU作为气候模拟的核心平台，越来越受到模式开发人员的重视。

挑战

全球气候变化及极端天气频发已经成为重大科学问题，通过地球系统模式模拟计算是解决此类问题的重要手段。当前，地球系统模式朝着更高分辨率更多物理过程的方向发展，这意味着模式运行需要巨大的计算资源。以0.1度全球海洋模式POP为例，在传统CPU平台下，15000核上运行1天只能模拟10年左右；而研究一般需要模拟千年尺度的海洋变化。这意味着，程序每次运行至少耗时100天。如果分辨率再提高10倍，计算量还要再提高到1000倍左右，计算量巨大。

地球系统模式程序的特点是程序较长且热点函数不明显。如果把某段或者某几段函数移植到GPU上，数据在显存和主存上的迁移开销会抵消掉GPU加速带来的优势。同时，模式程序主要表现为内存带宽受限，而GPU提供了多种片上缓存，有效地利用片上缓存是有效使用GPU加速地球系统模式的关键。

方案

我们以普林斯顿海洋模式POM为例，对海洋模式在GPU平台下的整体移植项目进行了深入研究。海洋模式的计算过程主要由平流项和水平扩散项、垂直混合项、状态方程、物理参数化方案等部分组成。计算过程核心表现为Stencil计算，但不同计算项又有不同的表现形式。同时，通信和I/O也需要被充分优化。

在单卡平台下，通过使用 read-only data cache 和 local memory (L1 cache) 等片上缓存和 kernel-fusion 等优化方法，单卡 (NVIDIA® Tesla™ K20X) 相当于多核 CPU (Xeon E2670, 8-cores) 有 4 倍的性能提升。

在 1 机 4 卡 (NVIDIA® Tesla™ K20X) 平台下，通过新型的 computation-communication overlapping 方法和 computation-I/O overlapping 方法，运行效率有 10% 以上的性能提升。

综合来看，在 1 机 4 卡 (NVIDIA® Tesla™ K20X) 平台下，运行性能相当于清华大学探索-100 集群 (Xeon X5670) 400 核规模。从而实现了一个 4 卡工作站可以替代一个通用集群进行区域海洋模拟。

影响

海洋模式POM在GPU上的移植和优化相关成果已经发表在地球系统模式开发领域顶级期刊的在线讨论区内 (<http://www.geosci-model-dev-discuss.net/7/8031/2014/gmdd-7-8031-2014-discussion.html>)。

地球系统数值模拟教育部重点实验室作为地球系统模式开发的国内领先研究机构，也是国内首家CUDA卓越中心 (CCOE)，和NVIDIA有着深入的合作。同时，校企合作也在不断加深。通过将GPU应用于地球系统模式，极大地加快了地球系统模式的运行效率，同时降低了运行功耗和运行成本。GPU巨大的计算潜能将助力地球系统数值模拟达到新的高度。