



中國地質大學
China University of Geosciences

艰苦朴素 求真务实

温家宝

艰苦朴素
求真务实
温家宝

1 Catalyst介绍



李健



1 前言

近年来，计算系统在速度和能力方面快速增长。大规模计算在成千上万个处理器的并行系统上执行。因此，并行计算方法和系统是现代计算模拟的核心。随着计算核心数的增加，输入输出（IO）成为限制因子。

计算能力和IO带宽对大规模计算产生深刻影响。



1 前言

传统的模拟过程分为：前处理（划分网格、设置边界条件和计算参数等）、计算和后处理（分析和可视化结果）。3个过程独立，但读写大数据成为大型计算的瓶颈。

当以较低频率保存结果，例如非恒定分析，每10步保存1次结果，将放弃90%的数据。有时，甚至保存单步的完整数据也将超出IO系统的能力。

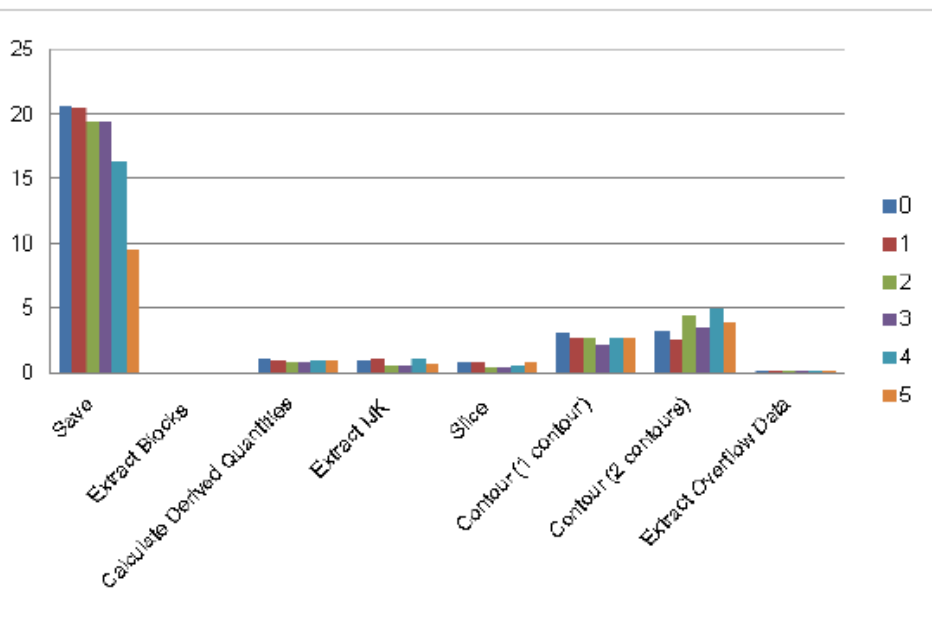


图1.1在桌面上启用6进程在某些分析计算和保存完整数据耗时的比较



1 前言

Catalyst方法将传统的3步整合为1步，即将后处理直接整合到计算过程中：

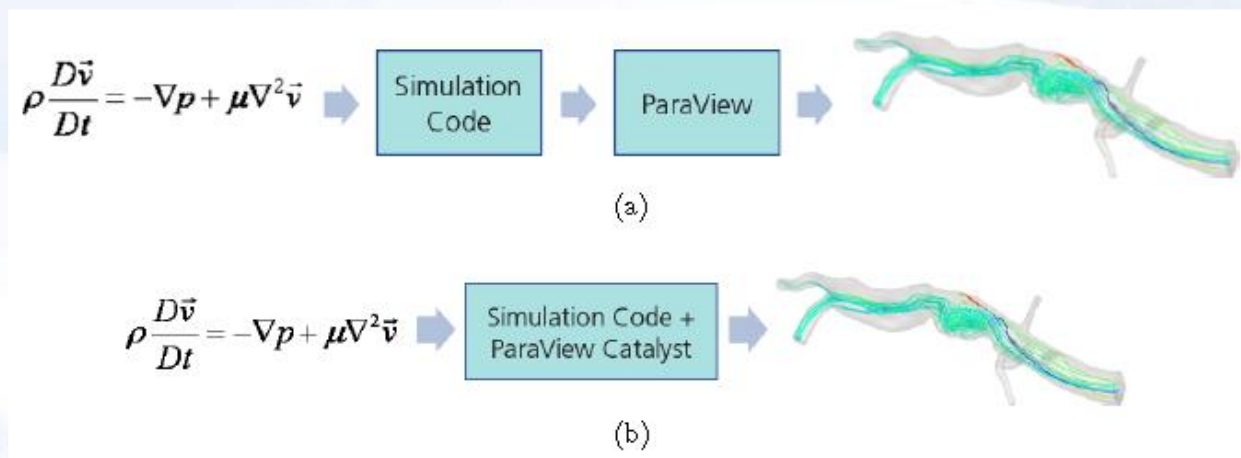


图1.2 传统的(a)和Paraview Catalyst (b)的工作流程



1 前言

不用保存完整的数据集到磁盘，IO只要提取相关的信息变量即可。数据提取，诸如等值线（面）、切面或流线等，比完整数据集要小几个数量级。因此，输出提取数据将显著降低IO成本。

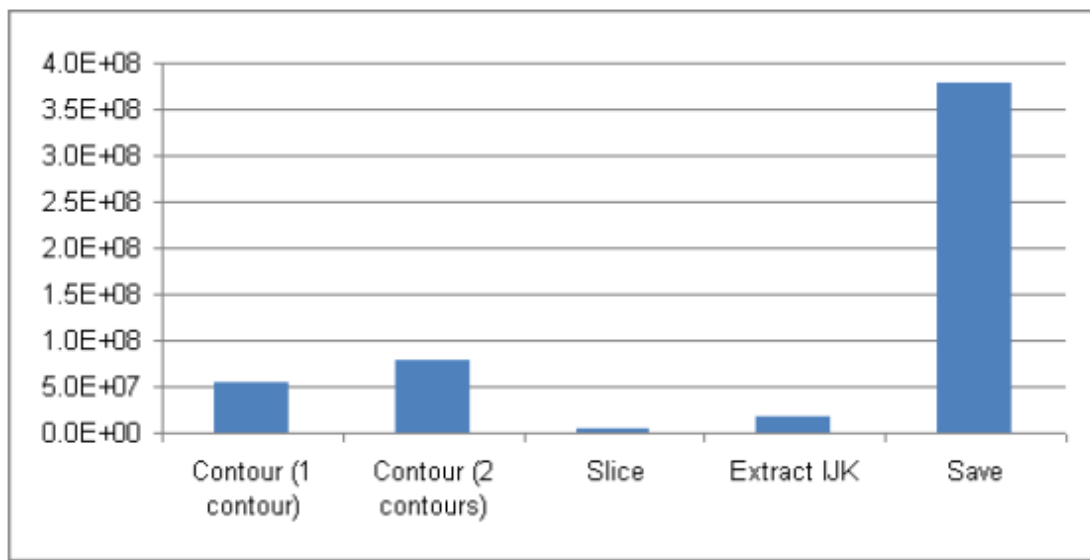


图1.4 比较保存完整数据集与保存特殊分析输出的文件大小（单位：bytes）



1 前言

Co-processing还可以及时发现问题，停止计算。

使用Catalyst做in-situ可视化的典型流程是：使用Paraview GUI定义模拟的输出，创建Python脚本实施初始化；然后，当模拟启动时，加载该脚本；然后，在模拟执行期间，以同步方式（即模拟在运行中）生成分析和可视化输出。

Catalyst可生成图片，计算统计量，生成制图和提取衍生信息，如多边形数据或等值面，来可视化几何形体或数据。



1 前言

具体实施Catalyst时考虑2方面：

(1) 是否降低IO成本是重要的？

(2) 是够这些管线合理地**尺度化**了？否则可视化时会使模拟过程**阻塞**，危害或影响整体分析周期时间。

但是，Paraview和VTK系统都是并行化的，通常是用于大部分的应用。



中國地質大學
China University of Geosciences

艰苦朴素 求真务实

温家宝

艰苦朴素
求真务实
温家宝

中国地质大学