

面向高能物理数据处理的可计算存储系统设计与实现

Thursday, 11 August 2022 11:40 (15 minutes)

随着高能物理实验装置复杂度和规模的不断扩大，实验数据的存量和增量均逐年增加。因此，高能物理实验对数据的处理能力和存储能力提出了新的挑战。高能物理数据处理普遍采用“存算分离”的计算模式，数据处理时需要在计算节点和存储节点间频繁搬运，占用大量的访存带宽和网络带宽。当系统中的计算作业达到一定规模时，便会遇到“内存墙”问题，以及各种网络问题、文件系统故障等，使整个计算系统处于不稳定状态。

可计算存储技术尝试将计算单元的部分任务卸载到存储单元，以减少数据移动，提升系统整体的处理效率。还可通过异构计算部件来补充存储单元的计算能力，达到计算加速的目的。

本文提出将可计算存储技术应用到高能物理数据处理中，并设计实现了相应的可计算存储服务器。该服务器以 CPU 为核心，通过 PCIe 总线挂载若干硬盘用于数据存储，以及 FPGA 芯片用于计算加速。利用 FPGA 可定制性强、并行度高的特点，对高能物理领域中部分常见的计算密集型任务和 I/O 密集型任务进行加速，只将存储单元的计算结果反馈给计算单元，从而减少或不进行节点间的数据移动。通过 FPGA 与硬盘之间的“数据零拷贝”，减少了计算加速时服务器内部的数据搬运。利用 FPGA 动态配置功能，可实现应用的快速切换。根据应用的特点，本系统在算法、应用、服务三个层次提供可计算存储服务，可明显缓解高能物理数据处理过程中由大量数据搬运造成的 I/O 瓶颈问题。在提升计算效率的同时，也增强了系统运行的稳定性。

Primary author: 高, 宇 (高能所)

Co-author: Dr CHENG, Yaodong (IHEP)

Presenter: 高, 宇 (高能所)

Session Classification: Parallel Session VIII (5): Particle Detector Technology

Track Classification: 粒子物理实验技术