**Active Flash: Towards Energy-Effificient, In-Situ Data Analytics on**

主机处理器➕固态硬盘嵌入式处理器

应用领域：高性能计算（科学仿真）

主机处理器➕固态硬盘嵌入式处理器

应用领域：高性能计算（科学仿真）

**Extreme-Scale Machines**

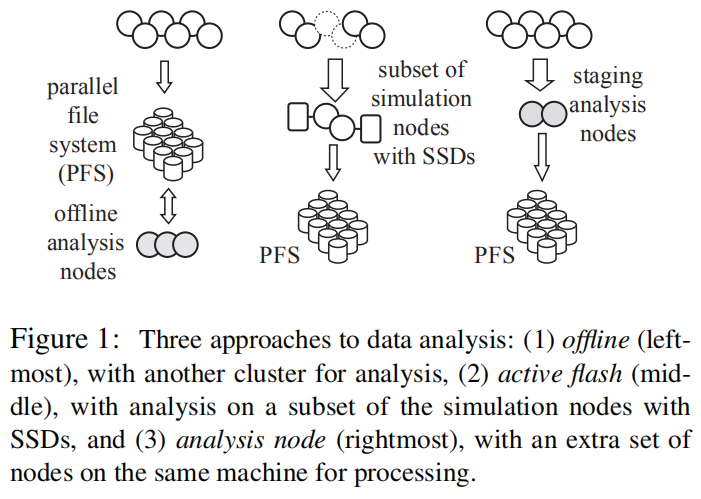
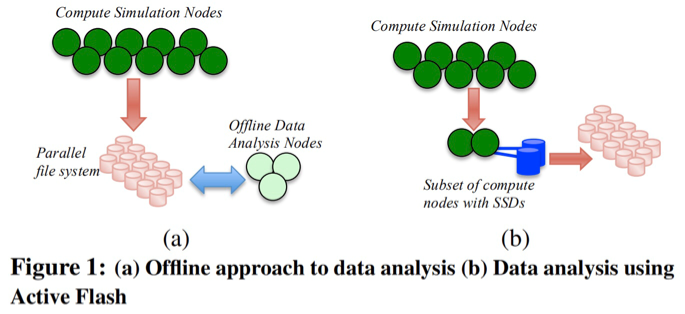
【Abstract】

现今许多科学发现均是在大规模超级计算模拟环境中诞生的，这伴随着大量的数据分析。然而，计算机及存储系统间的数据访问读取使其性能下降且能耗较高。

为解决上述问题，文章提出了Active Flash，它在存储数据的SSD中进行数据分析，是一种于原地进行的科学性数据分析方法。

性能模型及能耗模型表明，Active Flash能解决上述问题且不降低高性能计算的性能。同时，也在商用SSD控制器上创建了Active Flash的原型，以作为此idea的另一支撑凭证。

【Introduction】



3种数据分析方法：

**①传统方法**（offline，图1最左）：

高端计算（HEC, high-end computing）类机器的CPU hours很昂贵 → 同时兼顾simulation(main computation即通过计算产生数据)和post-processing(data analysis)不划算 → 因此采用offline模式，完成simulation(main computation)后，于另外的更小的集群上（并非昂贵的超算机）进行post-processing(data analysis)（两者共享PFS，即结果存入存储子系统并由PFS组织管理，数据分析（计算）时再通过PFS访问数据）

→ 冗余I/O：存储子系统与计算子系统之间过度的数据访问 → 性能与能量效率均不高

**②专用分析节点**（图1最右）：

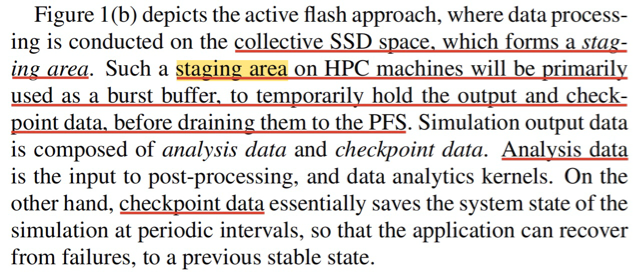
阶段一：结果数据输出；阶段二：写入存储子系统并由PFS组织管理；此方法在阶段一及阶段二之间进行数据分析计算，即在数据传输过程中进行“原地”数据分析

在simulation进程中同时进行computation → 减少了冗余I/O → 但数据分析计算额外占用了昂贵的CPU hours，干扰了simulation进程，致使对昂贵资源的利用率不高，即：

→ 计算节点除处理simulation外还要处理post-processing → 产生额外能耗

→ 仅一部分节点或核用于simulation → 会影响simulation性能

**③Active Flash**（图1中间）：利用SSD的计算能力进行数据分析（模拟完成后，利用SSD控制器的空闲时间完成对存于SSD中的数据进行计算）

这里有staging area的概念；

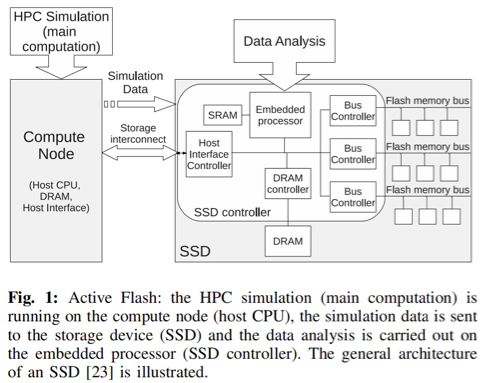
staging area:即SSD，在simulation产生数据被PFS管理前，暂存于SSD中进行数据分析。

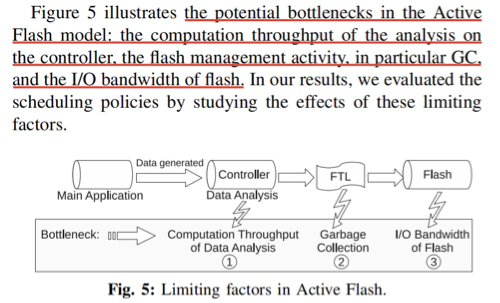
→ 结果数据已存储于SSD，计算不额外占用昂贵的CPU hours，不干扰simulation进程

→ 于数据存储原处（不使用另外的簇）进行计算，减少了冗余I/O

（②和③的区别：②使用部分超算机进行数据分析，③使用SSD嵌入式处理器进行数据分析）

（考虑到计算的超大规模，全部模拟节点的SSD都用做数据分析会使成本激增，故只部分计算节点如此（图1中间图示也如此，注解为subset of simulation nodes with SSDs））



【Background】

Active Flash的可行性分析：

①高I/O吞吐量和内部带宽

②SSD控制器上空闲周期的可利用性：HPC负载具有明显的I/O阶段和计算阶段（间歇性），其I/O活动据周期性且持续时间占比小，则可利用空闲时间进行计算

③多核SSD控制器：多核进一步增加了I/O间空闲时间的利用性

**Summary：**

美国东北大学和橡树岭国家实验室于2012年提出在高性能计算领域中利用主机处理器和固态硬盘内的嵌入式处理器共同进行数据分析（比如边缘探测、寻找局部极值、心跳检测和数据压缩等），并分析了该混合计算模型在性能和能耗方面存在的权衡。当更大比例的计算任务在硬盘内进行时，整体计算性能会下降，但能耗效率更高。仿真结果表明，相比于完全利用主机处理器进行数据分析，该混合计算模型能够以较小的性能代价大幅度提高能耗效率。

美国北卡罗来纳州立大学、东北大学和橡树岭国家实验室于2013年对数据分析任务进行了建模，包括SSD内的分期比率、性能、能耗模型。并应用这些模型将SSD设计成一个专用的数据分析的节点，定义数据分析命令。由Host端下发命令经SSD内调度，实现数据分析的盘内加速。