### 用于数据分析的异构云存储配置

摘要

数据分析是公有云服务上一类重要的数据密集型工作负载。然而，为这些应用选择正确的计算与存储配置是很困难的，因为可用选项空间很大和选项之间的交互非常复杂。此外，分析型工作负载访问的不同数据流具有不同的特征，这些特征可以被不同的存储设备更好地服务。

我们介绍selecta，这是一个为数据分析工作负载推荐近最优的云计算和存储资源配置的工具。Selecta使用潜在因素协同过滤来预测应用在不同配置下的性能，基于分析训练工作负载收集的稀疏数据。我们使用上百个SQL和ML应用对selecta进行评估，结果显示selecta选择近最优的性能配置（最优10%内）的概率为94%，近最优的成本配置的概率为80%。我们还使用selecta得出关于云存储系统的重要见解，包括NVME闪存设备的性能-成本效率，云存储关于支持细粒度容量和带宽分配的需求，端到端存储优化的动机。

#### 介绍

随着亚马逊AWS、谷歌云平台和微软AZure把它们的工作负载转移到平台上，公共云市场正经历着前所未有的增长。除了高弹性之外，公有云还承诺减少所有租户的总成本，因为资源是被租户共享的。然而，实现性能和成本效率要求为每一个给定的应用选择合适的配置。不幸的是，大量可用的实力类型和配置选项使为应用选择合适的资源非常困难。

存储选择通常是必要的，特别是对数据密集型分析的云部署。云厂商提供了各种各样的存储选项，包括对象、文件和块存储。块存储可以由硬盘、固态驱动器和高带宽、低延迟的NVMe组成。上述设备可能在运行应用程序的云实例本地或者远程。仅这些选项就会导致存储配置选项在吞吐量、延迟和位成本方面出现数量级的差异。随着基于3D-X poind 的新兴技术的出现，云存储领域变得越来越多样化。

选择正确的云存储配置对性能和成本来说都很重要。考虑SPARK SQL在两张128GB的表上等价查询的结果，我们发现8个EC2节点集群中的实例访问r-HDD比l-NVME查询花费时间要长8.7倍。这与最近的研究形成对比。与之前的spark相比，发现更快的存储只能将任务执行时间中位数最多提高19%。尽管NVMe单位时间成本更高，但NVMe的性能优势使该查询的执行成本降低8倍。如果我们还考虑核数和内存实例的一些选项，性能最好和性能最差VM存储配置之间的性能差距超过30倍。

为分析型应用确定云配置具有挑战性。即使我们将自己限制为单个实例类型并专注于优化性能，为特定应用程序选择存储配置仍然很重要。表一比较了三个spark应用程序的性能，这三个程序分别运行在存储设备分别为I-NVMe、r-SSD和混合(r-SSD用于输入/输出数据，l-NVMe用于中间数据)的i3.8xlarge EC2实例。第一个应用程序是I / O绑定的，并且受益于NVMe Flash的高吞吐量。第二个应用程序有CPU瓶颈，因此对三个存储选项配置表现性能一样。第三个应用程序是I/O绑定的，并且在混合存储选项表现性能最好，因为它最大限度的减少了读写I/O之间的干扰，读写I/Os在Flash上的性能是不对称的。这个结果不应该令人惊讶。分析工作负载访问多个数据流，包括输入和输出文件、日志和中间数据（例如混洗和广播）。每个数据流在访问频率、访问方式和数据生存时间方面都有不同的特征，这些特征使得不同的数据流更适合不同的存储设备。例如表1中的TPC-DS query 80应用程序，输入数据存储在r-SDD和中间数据存储在NVMe闪存比所有数据存储在NVMe表现得更好，因为它隔离了流并且消除了干扰。

我们介绍selecta，一种学习近最优VM和存储配置的工具，用于用户指定的性能成本目标的分析型应用程序。Selecta的目标是对新到达的数据进行频繁或定期重新运行的分析工作。配置由云实例类型（核心计数和内存容量）以及用户输入/输出数据和中间数据的存储类型和容量定义。Selecta使用稀疏性能数据训练在各种云配置上配置的应用程序，以及仅在两个配置上配置的目标应用程序的性能度量。