Skyhook: Managing Columnar Data Within Storage

Tuesday 13 September 2022 15:00 (10 minutes)

The advent of high-speed network and storage devices like RDMA-enabled networks and NVMe SSDs, the fundamental bottleneck in any data management system has shifted from the I/O layer to the CPU layer resulting in reduced scalability and performance. This issue is quite prominent in systems reading popular data formats like Parquet and ORC which involve CPU intensive tasks like decoding and decompression of data on the client. One solution to this problem is adopting computational storage, where CPU intensive tasks like decoding, decompression, and filtering are offloaded/distributed to often underutilized storage server CPUs, getting back scalability and accelerating performance. We build Skyhook, a programmable data management system based on Apache Arrow and Ceph that enables offloading of query processing tasks to storage servers. Skyhook does not require any modifications to Ceph nor assumes computational storage devices, rather its unique design embeds query processing in Ceph objects. This approach makes adding query offloading capabilities to the storage systems a breeze for practitioners. We use Skyhook to manage HEP datasets with storage, i.e., minimizing the creation of additional copies. The current release is deployed in University of Nebraska and University of Chicago and supports offloading of Nano Event filtering and projection queries. Our roadmap includes supporting joins with a distributed query execution framework that partitions Substrait query plans and distributes them for execution on clients, worker nodes, and storage objects. For the execution we plan to use the Acero (Arrow Compute Engine). For generating Substrait query plans we are planning to use Ibis. We are also collaborating with Argonne National Lab to extend Skyhook to other storage systems such as the Mochi software-defined storage system using RDMA for data transport to accelerate Skyhook query performance.

Author: CHAKRABORTY, Jayjeet

Co-author: MALTZAHN, Carlos Presenter: CHAKRABORTY, Jayjeet Session Classification: Plenary Session Tueday