Contribution ID: **436**   Type: **Talk**

Tuning the CMS Coffea-casa facility for 200 Gbps Challenge

Thursday 24 October 2024 16:51 (18 minutes)

As a part of the IRIS-HEP "Analysis Grand Challenge" activities, the Coffea-casa AF team executed a "200 Gbps Challenge". One of the goals of this challenge was to provide a setup for execution of a test notebook-style analysis on the facility that could process a 200 TB CMS NanoAOD dataset in 20 minutes.

We describe the solutions we deployed at the facility to execute the challenge tasks. The facility was configured to provide 2000+ cores for quick turn-around, low-latency analysis. To reach the highest event processing rates we tested different scaling backends, both scaling over HTCondor and Kubernetes resources and using Dask and Taskvine schedulers. This configuration also allowed us to compare two different services for managing Dask clusters, Dask labextention, and Dask Gateway server, under extreme conditions.

A robust set of XCache servers with a redirector were deployed in Kubernetes to cache the dataset to minimize wide-area network traffic. The XCache servers were backed with solid-state NVME drives deployed within the Kubernetes cluster nodes. All data access was authenticated using scitokens and was transparent to the user.

To ensure we could track and measure data throughput precisely, we used our existing Prometheus monitoring stack to monitor the XCache pod throughput on the Kubernetes network layer. Using the rate query across all of the 8 XCache pods we were able to view a stacked cumulative graph of the total throughput for each XCache. This monitoring setup allowed us to ensure uniform data rates across all nodes while verifying we had reached the 200 Gbps benchmark.

Authors: WIGHTMAN, Andrew (University of Nebraska Lincoln (US)); TOVAR LOPEZ, Benjamin (University of Notre Dame); BOCKELMAN, Brian Paul (University of Wisconsin Madison (US)); LUNDSTEDT, Carl (University of Nebraska Lincoln (US)); WEITZEL, Derek (University of Nebraska Lincoln (US)); ATTEBURY, Garhan (University of Nebraska Lincoln (US)); THILTGES, John (University of Nebraska Lincoln (US)); BLOOM, Kenneth (University of Nebraska Lincoln (US)); SHADURA, Oksana (University of Nebraska Lincoln (US)); ALBIN, Sam (UNL)

Presenter: SHADURA, Oksana (University of Nebraska Lincoln (US))

Session Classification: Parallel (Track 9)

Track Classification: Track 9 - Analysis facilities and interactive computing