Investigating Data Access Models for ATLAS A Case Study with FABRIC Across Borders and ServiceX

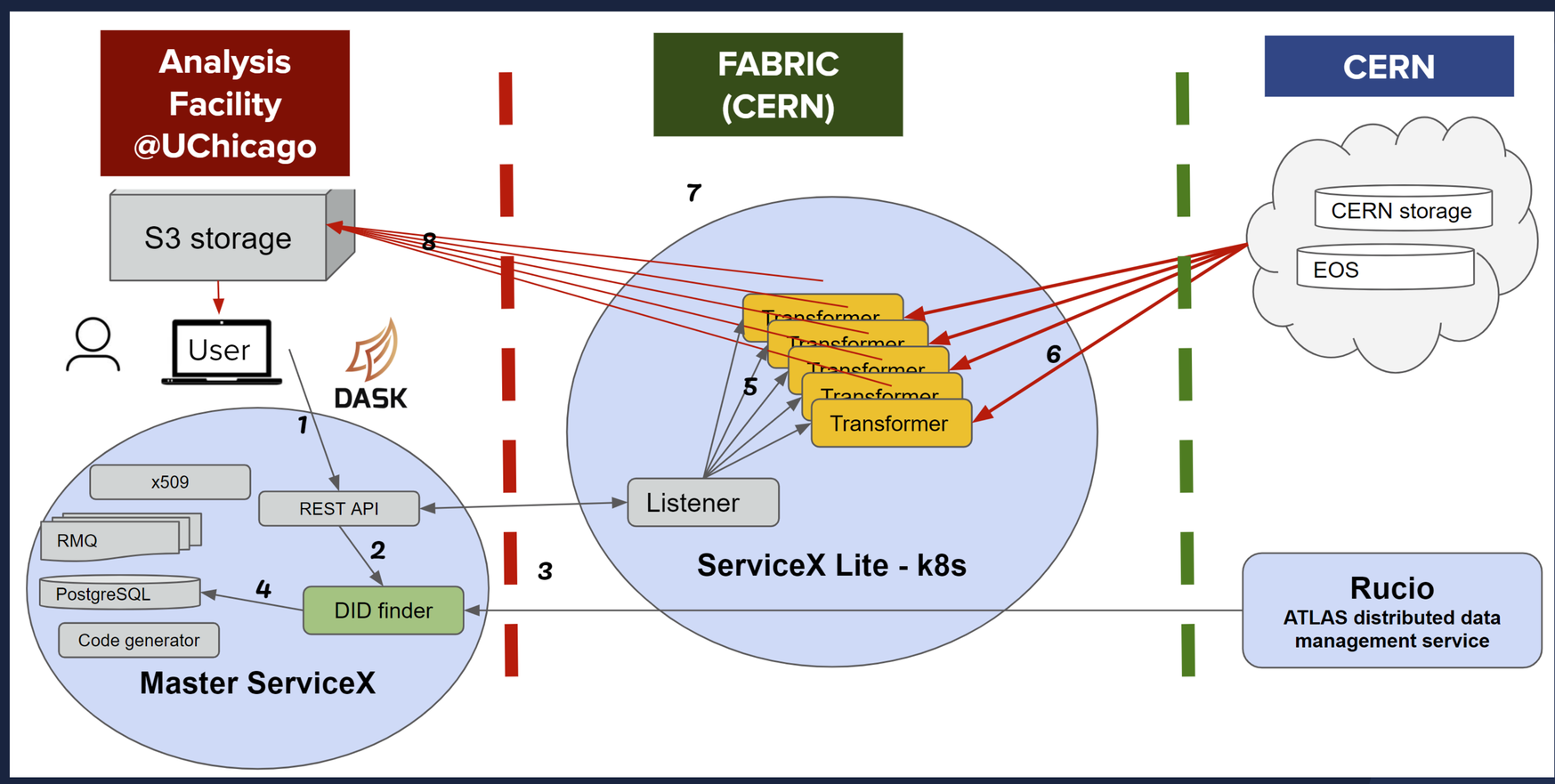
David Jordan[1]; Fengping Hu[1]; Ilija Vukotic[1]; Judith Lorraine Stephen[1]; Lincoln Bryant[1]; Robert William Gardner Jr[1]; Shawn McKee[2]

[1] University of Chicago (US)

[2] University of Michigan (US)

A lightweight, distributed filtering service running near data storage systems and delivering to Analysis Facilities could enable:

- interactive analysis
- efficient WAN use
- simpler data management



1. Introduction

- Current analysis models rely on grid resources for individual data pre-processing. This is a slow process and produces new datasets that are still too large for interactive analysis.
- We set up an analysis system that filters data at the edge resource close to the data, then sends results to an analysis facility where we run pythonic analysis on a Dask cluster (Fig1)

2. Setup

- Users connect to a full-fledged ServiceX instance at the University of Chicago ATLAS Analysis Facility (UChicago AF). This service looks up input data, generates filtering code, manages requests.
- ServiceX Lite is deployed on a Kubernetes cluster on a FABRIC slice in FABRIC Across Borders(FAB). This lightweight deployment only runs ServiceX filtering at up to one thousand cores.
- Output data (ROOT files or Awkward arrays in Parquet files) is uploaded to an S3 instance at UChicago AF over a FABRIC experimental network (Fig 2)
- Data analysis was done from a Jupyter notebook and parallelized on a 200 core K8s deployed Dask cluster.

3. Results

- We processed 9 samples totalling 3TB of data in 21k files. Analysis accessed 5% of the data (all events but only a part of variables).
- Standard analysis (download data + HTCondor processing) takes almost a day and moves all the data over WAN.
- ServiceX local processing of WAN accessed, and pre-cached data takes 15 and 3 minutes, respectively.
- The proposed model finished in ~6 minutes. This time is strongly dependent on percent of data accessed and the output data compression settings.
- **This represents 55% improvement in time-to-result and ~50% reduction in data transported over WAN.**

4. Future

- Further testing with larger data and new ServiceX versions
- Multisite deployment
- Intelligent scaling
- Fair-share support
- Location awareness

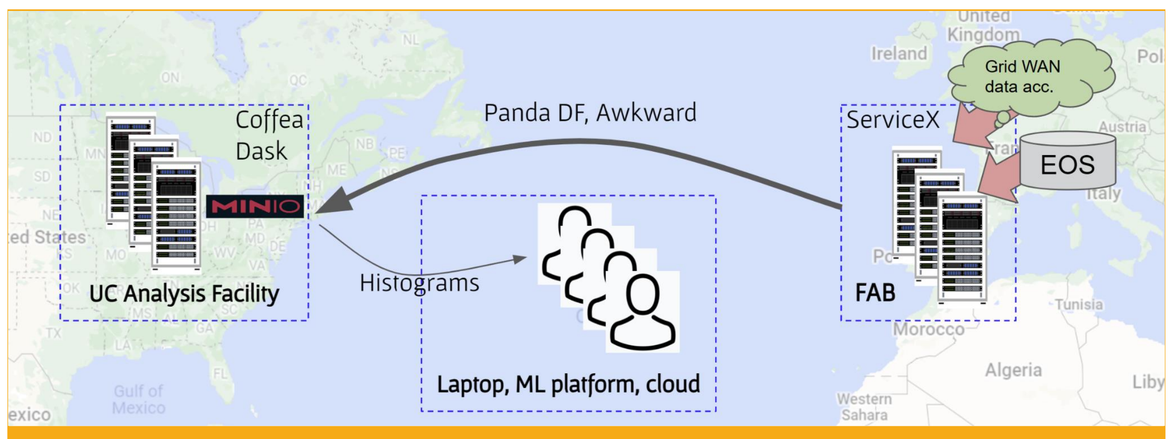


Figure 1. FAB K8s cluster located @CERN has a very fast access to CERN EOS storage and up to 1k cores for filtering. Actual analysis is done at UChicago AF where up to 1k Dask workers are available.

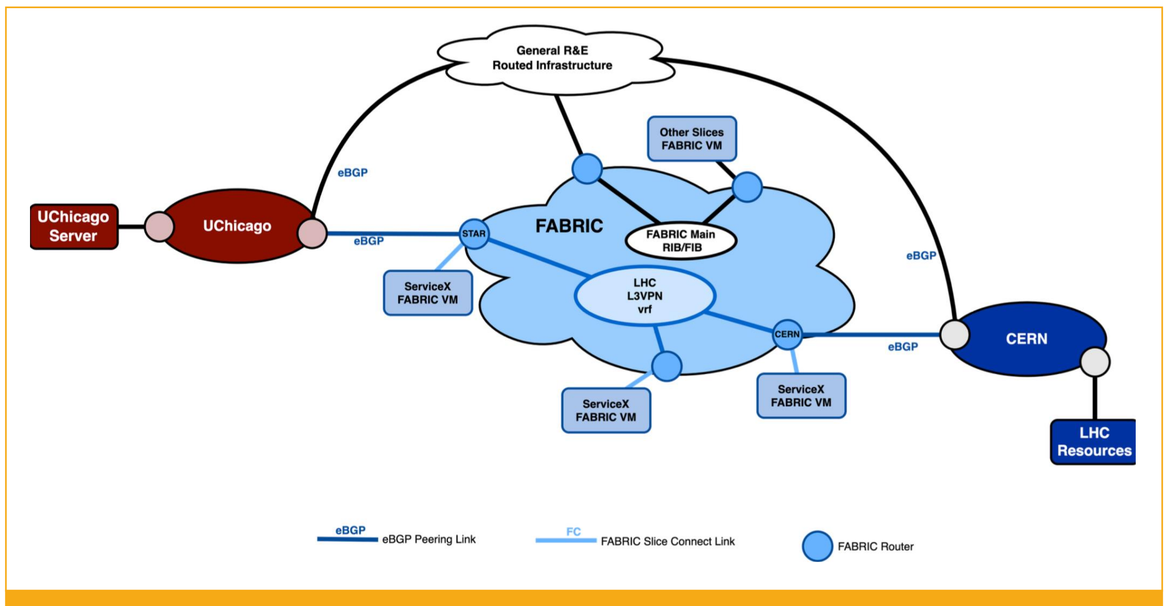


Figure 2. FAB computing cluster features dedicated network peering arrangements that link CERN Tier-0, the FABRIC experimental network, and an analysis center at the University of Chicago

References

1. *FABRIC* *http://portal.fabric-tested.net*
2. *ServiceX* *http://servicex.readthedocs.io*
3. *UChicago Analysis Facility* *http://af.uchicago.edu*

October 19 - 25, 2024
CHEP 2024

National Science Foundation

This work supported in part by NSF awards:
- OAC-2029176 Collaborative Research: IRNC: Testbed: FAB: FABRIC Across Borders
- OAC-1836650 Institute for Research and Innovation in Software for High Energy Physics IRIS-HEP
- PHY-2120747, U.S. ATLAS Operations: Discovery and Measurement at the Energy Frontier

download the **presentation**

ivukotic@cern.ch