ServiceX and Kubernetes

Interactive Columnar Analysis at Scale
What Does ServiceX Do?

- Converts experiment-specific datasets to columns
- Extracts data from flat ROOT files
- Simple cuts
- Simple derived columns
- Specified fields
- REST Interface
- Autoscales
- Object store results
- Transactional

Example Use Case

ServiceX delivers columns with condition:

\[ N_{jet} > 0 \quad \&\& \quad pT_{jet} > 10 \text{ GeV} \]
Delivery of columnar data to analysis system

- xAOD
- Flat N-tuple
- NanoAOD
- ...

ServiceX

- Arrow Buffers
- Awkward Array
- Slimmed / Skimmed Root
- ...

[Diagram showing data flow from various formats to ServiceX]
Architecture

REST Interface

DID Finder

Pre Flight Check

Code Generator

Transformer Manager

Transformer

RabbitMQ

Minio

Kafka

Kubernetes

Open source

Experiment Specific

Cross Experiment
Kubernetes Implementation

- Deployments:
  - ServiceX App
  - DID Finder
  - Code Generator
  - Preflight Check
- Secrets
  - X509 Cert and Key
- ConfigMaps
  - Flask Config
  - Generated Code
  - Rucio Config

- Ingress
  - REST Interface
  - Object Store Browser

- Roles:
  - Job Manager
  - Secret Manager
  - ConfigMap Manager
HELM Chart

- Easy to deploy on most Kubernetes Clusters
  - Laptops
  - University of Washington Physics Cluster (Tev)
  - University of Chicago River Cluster
- Extensive customization via values.yaml
  - Whether to use persistent volumes throughout
  - Rucio preference for local replicas
  - Use autoscaling
Dependent Charts

- RabbitMQ
- MinIO
- Postgres (optional)
Access to Service

• All interactions are via REST
• Gordon Watts’ ServiceX_Frontend Library
• Internal to cluster via Service
• Optional JWT Authentication
• Optional external ingress for
  • REST Service
  • Minio Object Store
Next Steps

- CI Build and Test of Helm Chart
- Prometheus Monitoring of Application
- LogStash log mining
- Migrate to GlobusAuth or CERN JWT
- Obtain Service Accounts for CMS and ATLAS
## Acknowledgements

<table>
<thead>
<tr>
<th>Institution</th>
<th>Contributors</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Chicago</td>
<td>Marc Weinberg, Rob Gardner, Ilija Vukotic, Lincoln Bryant, Chris Weaver, Neha Lingareddy</td>
</tr>
<tr>
<td>University of Washington</td>
<td>Gordon Watts, Mason Proffitt</td>
</tr>
<tr>
<td>FNAL</td>
<td>Lyndsey Gray, Jim Pivarski</td>
</tr>
<tr>
<td>University of Illinois</td>
<td>Mark Neubauer</td>
</tr>
<tr>
<td>University of Texas</td>
<td>KyungEon Choi</td>
</tr>
</tbody>
</table>

This project is supported by National Science Foundation under Cooperative Agreement OAC-1836650. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.
bengal1@illinois.edu