ServiceX and other deployments on the IRIS-HEP Scalable Systems Laboratory

Lincoln Bryant
Enrico Fermi Institute
University of Chicago

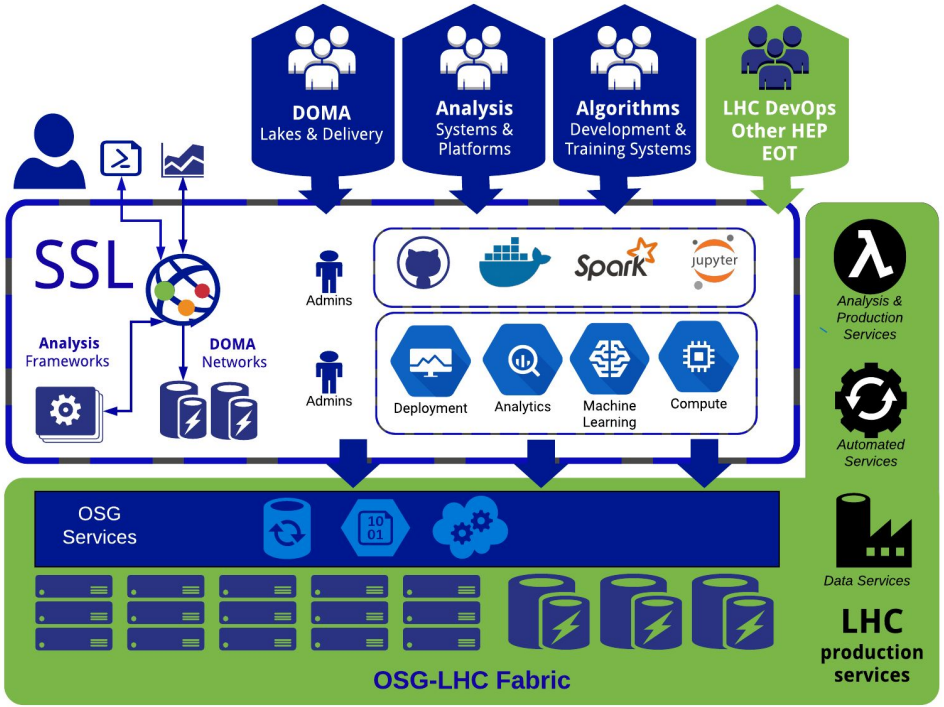
WLCG pre-GDB on Kubernetes
December 10, 2019

SSL as a Path to Production

Provisioning of software **environments** and development tools.

Distributed platforms materialized with tools like containerized edge services.

Integration point with the OSG, WLCG and LHC experiment services (data, analysis).



Orchestrating services in the SSL

- Need **flexible infrastructure** for supporting the workloads we expect
- **Dynamically reconfigure services** on existing hardware
  - e.g. HTC batch vs interactive frameworks
- Can we **leverage** cloud-native momentum?
- Want to **"glue" clusters together to achieve scale**, but abstract away infrastructure differences
  - Clear a smooth road for the developers
- Discussion to **multi-site K8s "substrate"**

SSL and Kubernetes

- Focus not so much on developing new K8S-based technology, but proliferation to institutions
  - We are building all of this great software infrastructure but we may find ourselves struggling to find places to deploy it!
- SSL monthly call – topical presentations and helpdesk of sorts
  - **Gather expertise and spread K8S know-how**

First piece of the substrate

Repurposed R&D cluster at UChicago

- Vintage but nice (3k cores)
- 48 cores, 256 GB RAM, 2x800GB SSD, 10Gbps networking per node
- 2x40g to campus 100g SciDMZ
- Kubernetes for flexibility for services and compute

Federated ID access (institutional, CERN account), edge services hosting, Unix account provisioning, OSG-LHC software environment

Examples of applications on SSL

ServiceX: an IRIS-HEP DOMA service

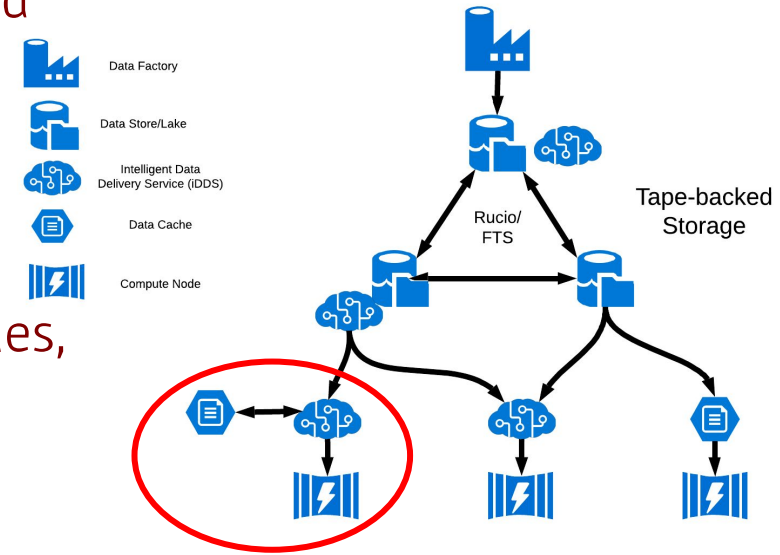
ServiceX focuses on integration with Rucio and reformatting for pythonic tools & **endstage analysis systems**

R&D for data delivery of columnar data

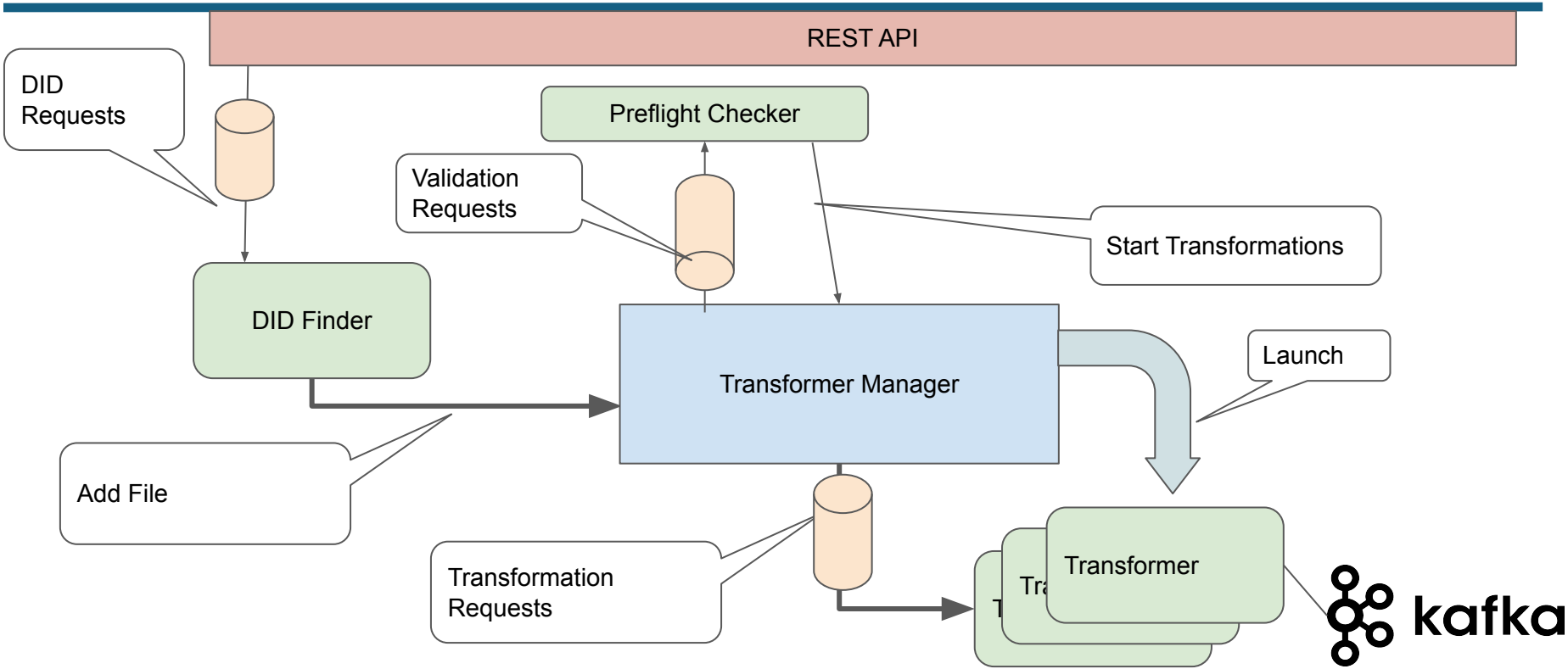
Supports multiple input types (xAOD, flat ntuples, ...) and common data mgt (Rucio, XCache)

Utilize industry standard tools (GKE, on-prem Kubernetes, Helm, Kafka, Redis, Spark, ...)

Reproducible, portable deployments on k8s



ServiceX declarative service components



ServiceX as deployed on the SSL

Services

NAME	TYPE	CLUSTER-IP	EXTERNAL-IP	PORT(S)
release-name-0-external	NodePort	10.104.248.66	<none>	19092:31090/TCP
release-name-1-external	NodePort	10.102.224.207	<none>	19092:31091/TCP
release-name-kafka	ClusterIP	10.110.20.190	<none>	9092/TCP
release-name-kafka-exporter	ClusterIP	10.97.197.166	<none>	9308/TCP
release-name-kafka-headless	ClusterIP	None	<none>	9092/TCP
release-name-minio	ClusterIP	10.97.66.31	<none>	9000/TCP
release-name-postgresql	ClusterIP	10.96.46.224	<none>	5432/TCP
release-name-postgresql-headless	ClusterIP	None	<none>	5432/TCP
release-name-rabbitmq	ClusterIP	10.106.225.45	<none>	4369/TCP,5672/TCP,25672/TCP,15672/TCP,9090/TCP
release-name-rabbitmq-headless	ClusterIP	None	<none>	4369/TCP,5672/TCP,25672/TCP,15672/TCP
release-name-servicex-app	NodePort	10.102.57.95	<none>	8000:31973/TCP
release-name-zookeeper	ClusterIP	10.108.18.68	<none>	2181/TCP
release-name-zookeeper-headless	ClusterIP	None	<none>	2181/TCP,3888/TCP,2888/TCP

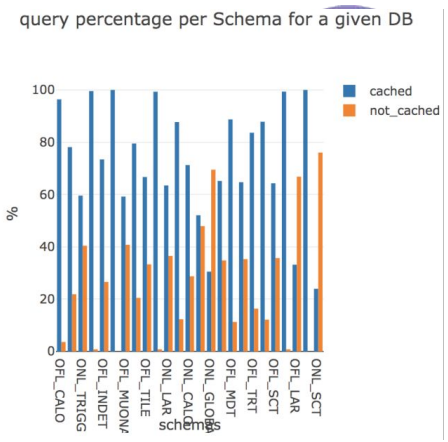
Workloads

```
C:\Users\ivukotic>kubectl get pods -n servicex
NAME                                                READY   STATUS      RESTARTS   AGE
release-name-did-finder-b5b6bfb48-d2bnq             1/1     Running     2          14m
release-name-kafka-0                                1/1     Running     3          17d
release-name-kafka-1                                1/1     Running     1          17d
release-name-kafka-exporter-76cf9c7c69-nbzcq        1/1     Running     2          17d
release-name-minio-5c54b8648-b92n2                  1/1     Running     0          14m
release-name-postgresql-0                           1/1     Running     0          14m
release-name-preflight-cb454d56-f4tmv               0/1     Error       7          14m
release-name-rabbitmq-0                             2/2     Running     0          14m
release-name-servicex-app-9b58465fc-q8xkm           1/1     Running     2          14m
release-name-test-topic-consumer                    0/1     Completed   0          2d23h
release-name-test-topic-consumer-v2                 0/1     Completed   0          2d23h
release-name-test-topic-consumer-v3                 0/1     Completed   0          2d23h
release-name-test-topic-create-consume-produce      0/1     Completed   0          17d
release-name-testclient                             0/1     Error       0          26h
release-name-zookeeper-0                            1/1     Running     1          17d
release-name-zookeeper-1                            1/1     Running     2          17d
release-name-zookeeper-2                            1/1     Running     1          17d
transformer-210b2510-0e16-4006-9055-d9addb51fcaa-xprdk   1/1     Running     0          13m
```

- Currently deployed on Kubernetes cluster on SSL
- Performance testing: Running with hundreds of transformers on cluster
- Has Prometheus monitoring dashboard
- All services and deployments created by Helm chart

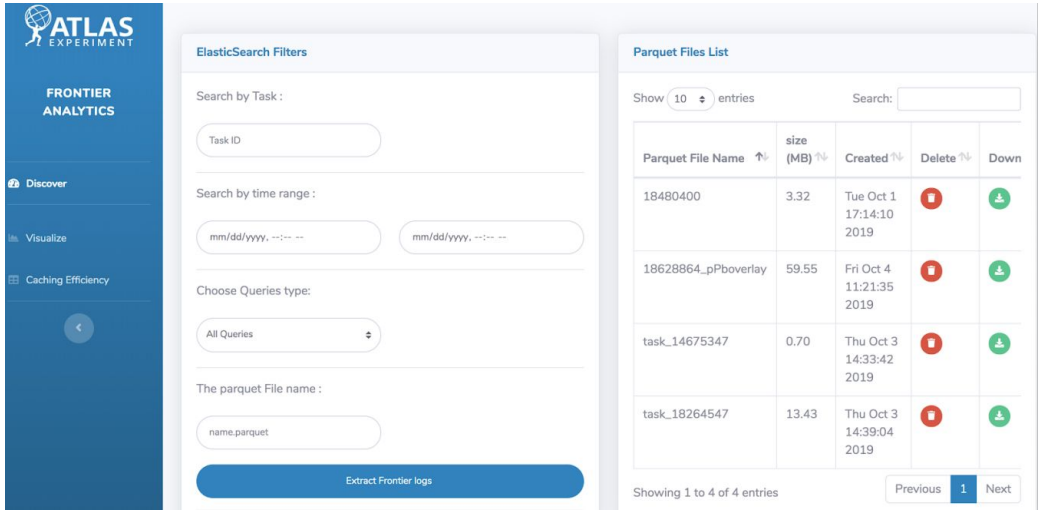
Frontier Analytics on the SSL

Millissa Si Amer
Nurcan Ozturk
Elizabeth Galas
Julio Lozano
Andrea Formica



- Analyze access patterns for ATLAS Conditions Data in a variety of workloads
- e.g., find problematic workloads that have poor cache hit ratios in Frontier
- Help design a more cache-friendly system for Run 3
- Accesses a lot of data, calculation is quite expensive in CPU time and RAM, so everything done asynchronously, temporary results cached, results stored.

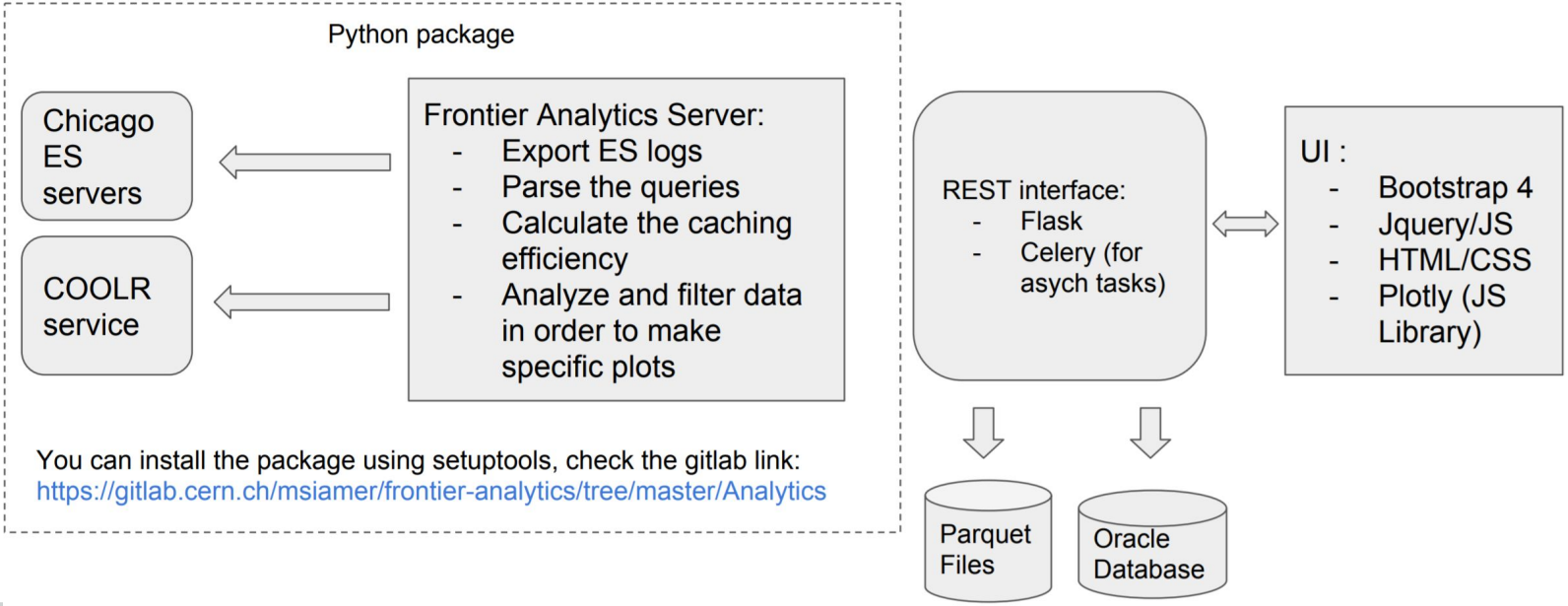
Presented at CHEP 2019.



https://frontier.uc.ssl-hep.org/

Frontier Analytics on SSL, cont'd

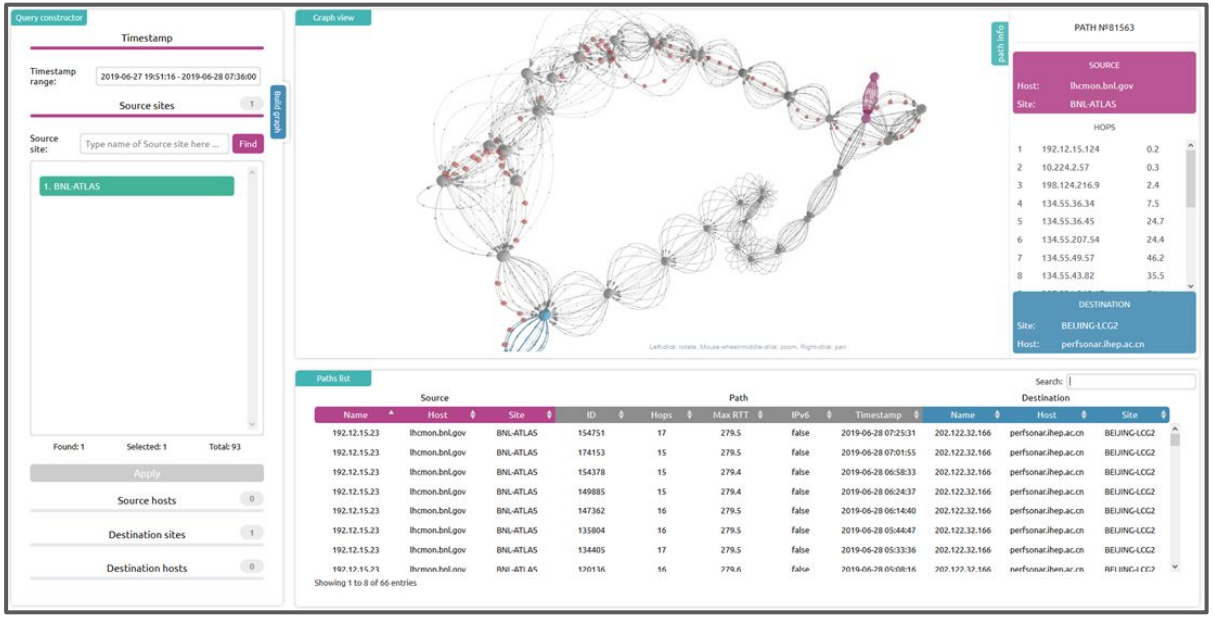
- Using CERN gitlab for source, secrets and image repo.
- One line deployment just doing kubectl create of a directory with yaml files.
- Deployment significantly simplified using cert manager (lets encrypt) for ingress.



PerfSONAR Analytics Platform on the SSL

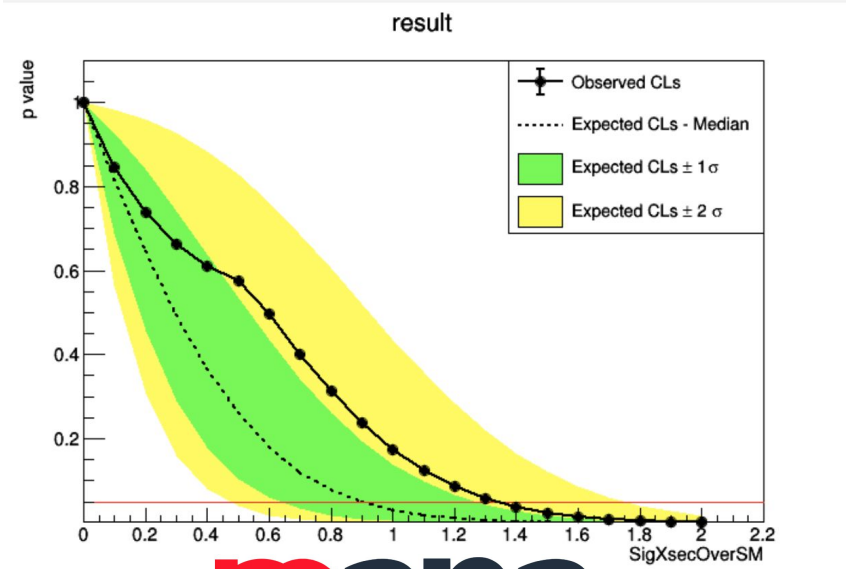
- Analytics service / front-end to data stored in Elasticsearch at UChicago
- 1,2, or 3D network route visualization based on PerfSONAR traces
- Single yaml file deployment – one service, one ingress, one pod, one secret.

https://perfsonar.uc.ssl-hep.org



REANA on the SSL

- Reproducible Data Analysis Platform
  - Create "runnable recipes" for making reproducible data analyses
- Summer student at UChicago was able to successfully deploy REANA on SSL
- She was able to run the ATLAS RECAST demo in just a few days with no prior Kubernetes or physics analysis experience



https://github.com/reanahub/reana-demo-atlas-recast

ATLAS ElasticSearch Snapshot on SSL

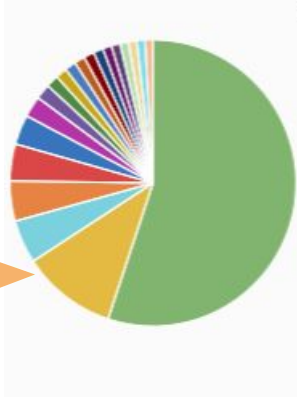
- Elastic stack at UChicago was upgraded to v7.3 over the summer, fraught with problems :(
- Needed to restore access for heavy users!
- Created and restored an entire Elastic 6.8 backup on SSL, via K8S + Helm
  - https://github.com/elastic/helm-charts worked perfectly!
  - 35 nodes, each with 800GB SSD
  - Happily running alongside Condor, ServiceX, other applications

OSG Backfill on SSL

	total
SU ITS	33.6 Mil
IRISHEP-SSL-UCHICAGO	6.54 Mil
GLOW	2.992 Mil
MIT CMS	2.763 Mil
IIT - Illinois Institute of Technology	2.546 Mil
MWT2 ATLAS UC	2.056 Mil
UCSD CMS Tier2	1.452 Mil
UColorado_HEP	1.053 Mil
Vanderbilt ACCRE	858 K
SLATE-K8S-UCHICAGO	787 K

6.5M hours in past 90 days

Consistently a top opportunistic resource provider in OSG



- HTCondor worker SLATE application with CVMFS plugin from CERN+PRP
- Container borrows OSG Glidein startup scripts, job wrapper, etc

```
$ slate app install htcondor-worker --group ssl ...
```

What new capabilities are needed?

- RBAC rules borrowed from SLATE / NRP Controller, but: **accounts need more rigorous authentication than $KUBECONFIG, need traceability**
  - https://github.com/dexidp/dex ?
  - Users would occasionally be presented with a request for authentication token, which would take them through an OAUTH2 workflow.
- Rook/Ceph still on our wishlist at Chicago, need to sit down and understand why it doesn't work currently

Call for presentations, getting sites involved

- Would like to start gathering topical presentations for the community
- Some ideas
  - Deploying K8S on Bare Metal
    - How-to's for: Volume Providers, Ingress, Load Balancers, Monitoring
      - Taken for granted on Cloud, how to reproduce that experience for your users
  - K8S Multi-tenant
    - What is possible in the multi-tenancy space today?
    - Authentication? Dex? Other OIDC providers?
  - K8S Federation
    - Reports on KubeFed, Admiralty, etc.
    - What works? What's broken?
  - Supporting applications on clusters
    - Jupyter notebooks
    - GPU/FPGA/Other architectures
    - Big data movers

SSL Team Google Group

https://groups.google.com/a/iris-hep.org/forum/#!forum/ssl-team

Welcome to join us - setting up next call

https://www.when2meet.com/?8478744-PB1zs

Thank you!

http://iris-hep.org/ssl