SSL Architecture

Lincoln Bryant
Enrico Fermi Institute
University of Chicago

IRIS-HEP Retreat @FNAL
September 13, 2019

Reminder of our goals

- **To build an architecture that implements...**
  - a scalable community platform
  - supporting groups and projects
  - bespoke resources & configurations
  - declarative & reproducible deployments
  - services to build & manage artifacts
  - reduce cognitive load for developers and deployers

Since we last met..

- Successfully converted the RIVER cluster from UChicago Computer Science into an SSL resource
- ~3600 logical cores, ~20TB RAM, solid state storage, 10Gb networking
- Kubernetes v1.15
- SLATE-enabled

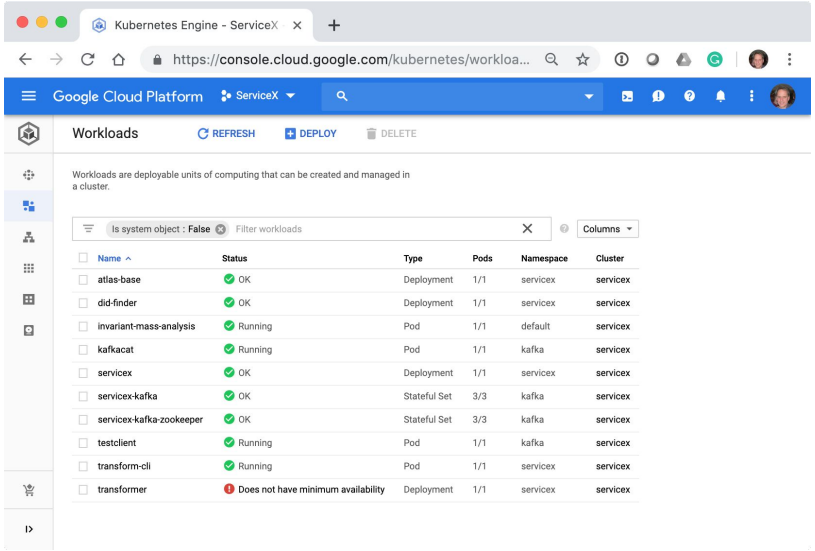
SSL Kubernetes Features

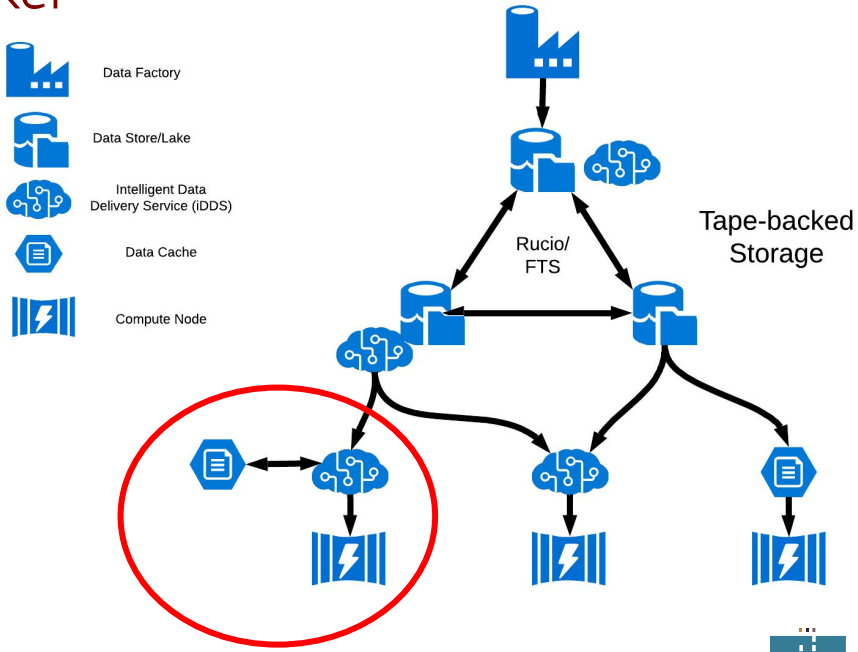
- Each node has 48 cores, 256GB RAM, 800GB SSD, 10Gbps connectivity
- CVMFS available
- MetalLB load balancer available with ~50 public IPs
- NGinx ingress controller
- ”Backfill” and ”Normal” pod priorities; preemption enabled
- External DNS service, monitoring to be added

Successes thus far

ServiceX

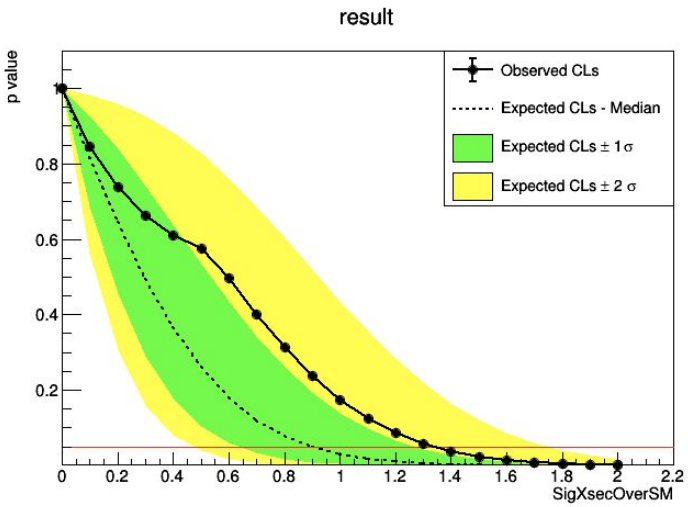
DOMA iDDS development on Google Cloud reproduced on IRIS-HEP SSL with Docker containers and Kubernetes





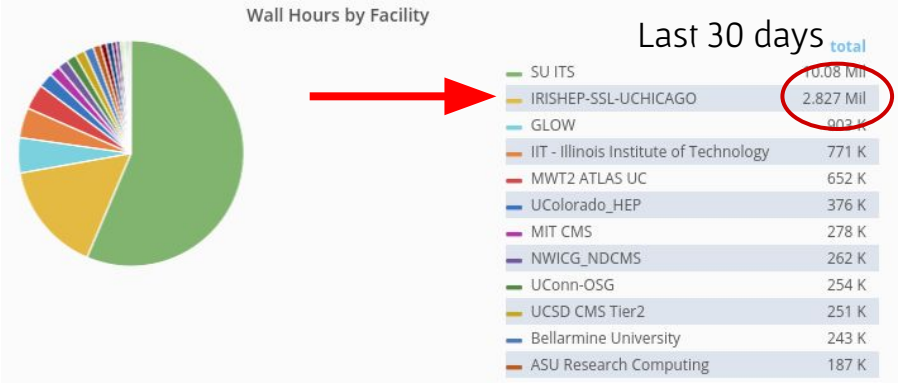
REANA

- UChicago CS Student (Neha Lingareddy) was able to deploy the REANA framework on SSL over the summer
- Started with no knowledge of Kubernetes, REANA, etc – was able to produce ATLAS "Recast demo" plots in a couple of days



Backfill from OSG

- Deployed OSG-like worker containers via SLATE
  - Most of the usual trimmings: CVMFS, glidein startup scripts (no singularity yet)
  - Flocking from OSG Connect
- Large contributor to OSG!
  - 2.83M core-hours over last 30 days



Looking forward

Loose Roadmap

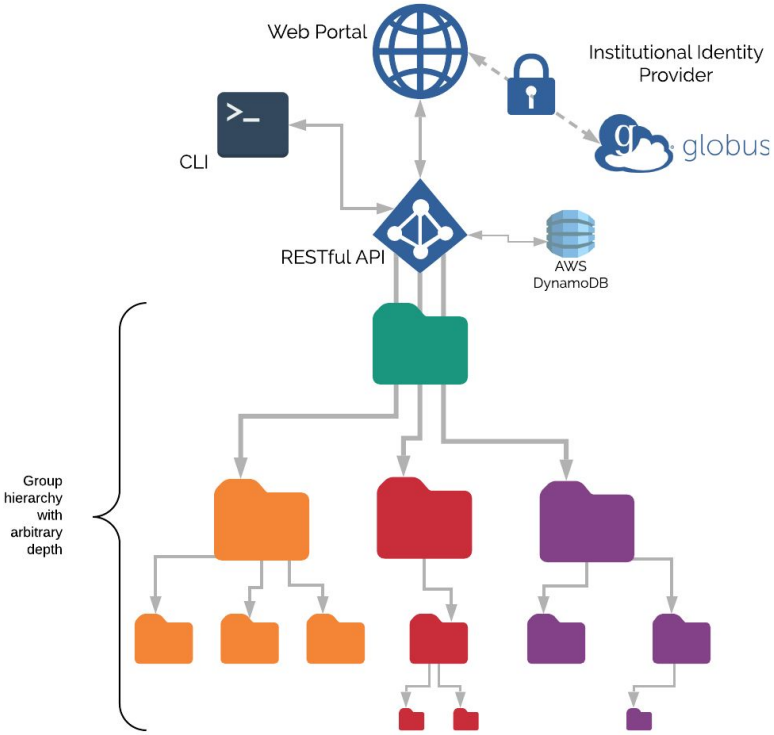
- Immediate goals:
  - Stand up clustered storage solution, e.g. Rook
  - Web UI with federated identity for registering groups and users, mapped onto Kubernetes namespaces.
  - Bring in applications from other parts of IRIS-HEP
- Near-term goals:
  - Stand up SSL resources at other sites
  - Investigate federation technologies (Kubernetes 'native' federation, service meshes, Admiralty, etc)

Storage and Federated Storage

- Currently no clustered storage solution is in place–
  - Investigating Rook/Ceph for storage on the UC cluster
- Will need to understand federated storage across SSL clusters
  - Perhaps we can take some cues from PRP? e.g. EdgeFS?
  - Something custom with Rucio et al?

Groups and projects

- Web and CLI interfaces for user management
- Users can invite others, create sub groups, etc.
- Provisioning clients can map groups onto Kubernetes namespaces
- Need to figure out strategy for sync between SSL resources



Using and contributing to SSL

Accessing SSL

- Today this is done manually, come see me if you want access!
  - Creating Kubernetes user accts / namespaces with some in-house scripts
  - You will need a 'kubectl' client on your laptop to access the cluster.
- Eventually will be replaced with the group system

How to contribute

- SSL Working Meeting
- **Fridays 2–4PM Central Time,** once a month
- Work on standing up Kubernetes resources at participating sites, troubleshoot software deployments, experiment with interoperability technologies, federation/mesh, etc
- Hop on / hop off as desired– no fixed agenda topics

Some Discussion Points

- Are we meeting the needs of the community?
- Have we missed any major required components?
- What do we need to do to get AS, DOMA, IA applications and codes on SSL?
- Are any other institutions willing to contribute resources?

SSL Substrate Planning

- Which other sites are ready to join?
  - Next 3 months (Nov 1)
  - 6 months (Mar 1)
  - 1 year
- "Federation" strategy?
  - Clearinghouse of service deployment targets
    - Lightweight – simple list for the near term

DOMA demonstrator requirements

- 6 months time frame
- 120 TB (60B) needed for multi-PB transfers of nano-AOD over http
- ServiceX to Skyhook to Coffea (in an analysis facility)
  - Want columns; awkward arrays (arrow buffer backend)
  - Skyhook storage
  - Spark cluster
  - Coffea, notebooks

Extra slides

Desired SSL capabilities

- Support a diverse catalog of deployment patterns & models
- Experiment patterns (scalability tests)
- Usability
  - Modality, Reservation
  - Metrics, logging, analytics
- Operation & Support
- Openness
  - to providers to contribute
  - to developers to conduct experiments
- Recording value
  - Analysis platform "blueprints" :
    - Single site/region deployments
    - Multi-region deployment
    - Multi-cloud hybrids – e.g. SSL+GCP+CERN, etc..
  - Demonstrations & archival of demo artifacts

An architecture that implements...

- a community platform
- supports groups and projects
- bespoke resources & configurations
- declarative & reproducible deployments
- services to build & manage artifacts
- scalable up and back down
- reduce cognitive load for developers and deployers

Community platform

- Open to all working on software infrastructure in HEP
- CILogon, Globus to provide single-sign on and federated identity
- Lightweight user and group (project) management system
- Infrastructure itself composable, reusable

Resources

- Mix of bespoke dedicated resources and capability for users to bring allocations on others
- Container-based service orchestration on dedicated resources
- VC3-like technology to connect to HPC/HTC resources for batch
- Facilitate integration of commercial cloud resources when needed

Orchestrating services in the SSL

- Need flexible infrastructure for supporting the workloads we expect from SSL
- Dynamically reconfigure existing hardware to be a HTCondor cluster today, Spark tomorrow, whatever is needed.
- Containerized services are getting a lot of attention in Industry right now– can we take advantage of the momentum?
- Want to "glue" clusters together, but abstract away infrastructure to whatever extent possible – clear a smooth road for the developers
- Potentially mimic cloud native groupings: e.g. create "zones" of resources

CS cluster – SSL base platform services

- Repurposed UChicago CS research cluster
- Vintage but nice: (~50)
  - CPU: 2 x Intel Xeon E2650 v3 12-core processor, 2.3GHz, 30MB cache
  - DRAM: 16 x 16GB TruDDR4 Memory 2133MHz, 256GB
  - Disks: 2 x 800GB SATA MLC SSD, 1.6TB
  - 10G NICs
- 2x40 Gbps to SciDMZ
- Rebuilding as Kubernetes
- Explore federation to aggregate w/ others

Federated ID access (institutional, CERN account), edge services hosting, Unix account provisioning, LHC software env.

Kubernetes (k8s)

- Open source container orchestration platform
- Automate deployment, management, scaling
- Has origins in Google/Borg
- Supported/managed by Cloud Native Foundation
- Declarative model for deployments

Declarative infrastructure

- Want infrastructure built under the SSL to be easily reusable and deployable to other sites
  - No more twikis with install guides!
- Declarative nature of Kubernetes is a good fit and gets us a long way down that road.
- SSL as an incubator for projects which then "graduate" to become full-fledged infrastructures that run on production resources.

Federating platforms

- Expect users to outgrow the dedicated pool of resources we have now.
- Need an interface and mechanism to allow users of the SSL reach into resources at a heterogeneous collection of sites
- Many approaches in the Kubernetes community, waiting to see what survives & what will be most appropriate for us

Extending into HPC/HTC

- Foresee workloads that require some service infrastructure in the SSL, but want to do batch computing elsewhere
- Want to facilitate by using technologies derived from VC3, HEPCloud and others
- Provision compute schedulers, data managers on SSL, schedule workers to HPC resources via overlays

Artifact build and management

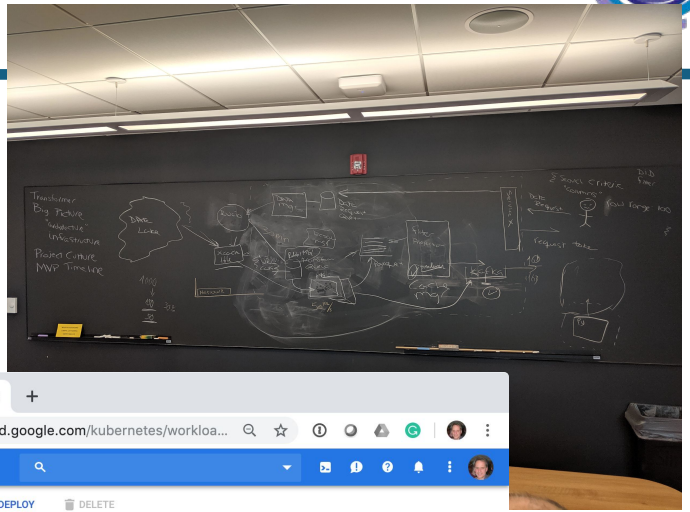
- Provide resources for building and registering containers, compiling software, etc.
- Off-the-shelf tools plugged into SSL resources with a little bit of glue.
  - Why wait 30 minutes for DockerHub to build your container?
- Is this obviated by CERN services? Perhaps more valuable for non-LHC experiments

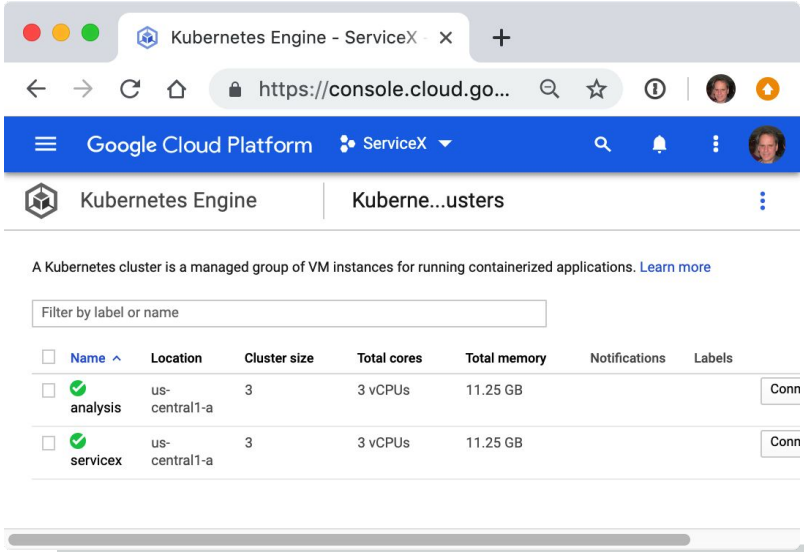
SSL "Glass"

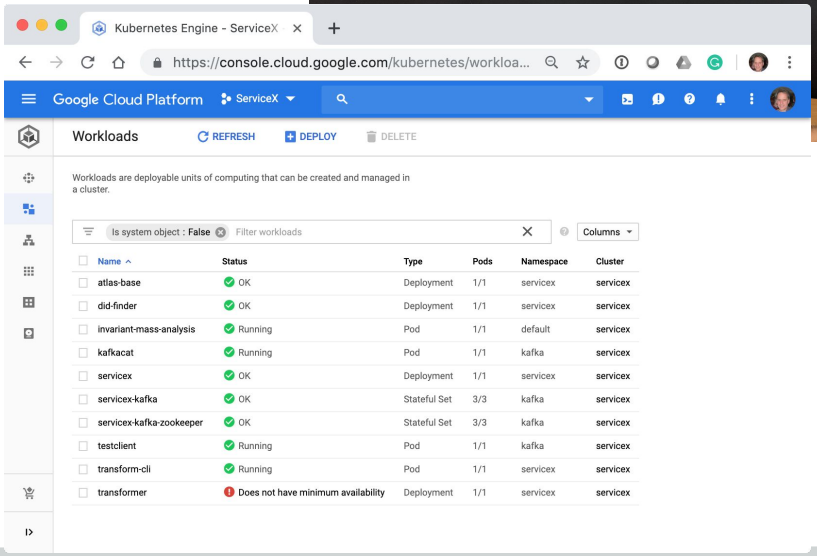
- Portal for visibility and organization
  - groups (projects)
  - resources
  - artifacts
- Metrics, logging, analytics
- Regional, national and international scopes

WBS 6.3 Functional Testing

Until SSL base platform operational we can use GKE for testing. Early deployments for iDDS/ServiceX







Current status

- Group/identity bits are being developed for multiple projects, being repurposed for SSL.
- Kubernetes conversion of River cluster underway.
  - 4 nodes online, backfilling w/ OSG via SLATE
  - Brave early adopters come talk to me afterwards!
- Looking for partners to contribute infrastructure and a bit of effort – experiment with how to federate resources.

Wrap up

- To briefly recap:
  - Institutional Identity and group management
  - Container-based, declarative software deployment and service orchestration
  - Mix of dedicated and non-dedicated resources
  - Exploring options for Federation
  - Building tightly integrated "pane of glass" for it all
- Integrate with industry best practices where practical!

Discussion

- All of this is very nice, but we need to meet the needs of the community.
- We need input from Analysis Systems and others!