User Batch Questions

Kian-Tat Lim

Original Concept

- We will have a production batch system; open it up to users with resource limits
- Allow any executable
- Provide same framework as LSST production, but do not require it

Issues

- User accounts
  - RSP-only or infrastructure
- Access to data
  - VO services
  - Butler
  - Staging scripts/wrappers
  - CernVM-FS (???)
  - Direct filesystem
  - Direct object store
- Compute resources
  - RSP-only or shared with production

VMs, Containers, or Processes

- Execution and storage overheads
- Ease of environment creation/maintenance
- Security

- Containers: Docker vs. Singularity

User-Directed Parallel Processing

- Programmatic: Dask, Spark, Parsl
- Command Line (low-level): HTCondor, SLURM
- Command Line (high-level): Pegasus, PanDA
- Kubernetes: K8s Jobs, Argo Workflows
- BPS, PipelineTask wrapper
- UWS

Follow CADC?

- skaha is generalized JupyterHub
- Plans to extend to non-interactive batch containers
- VOSpace data access

- But also OpenStack VMs