OSG/PATh Experiences with DevOps

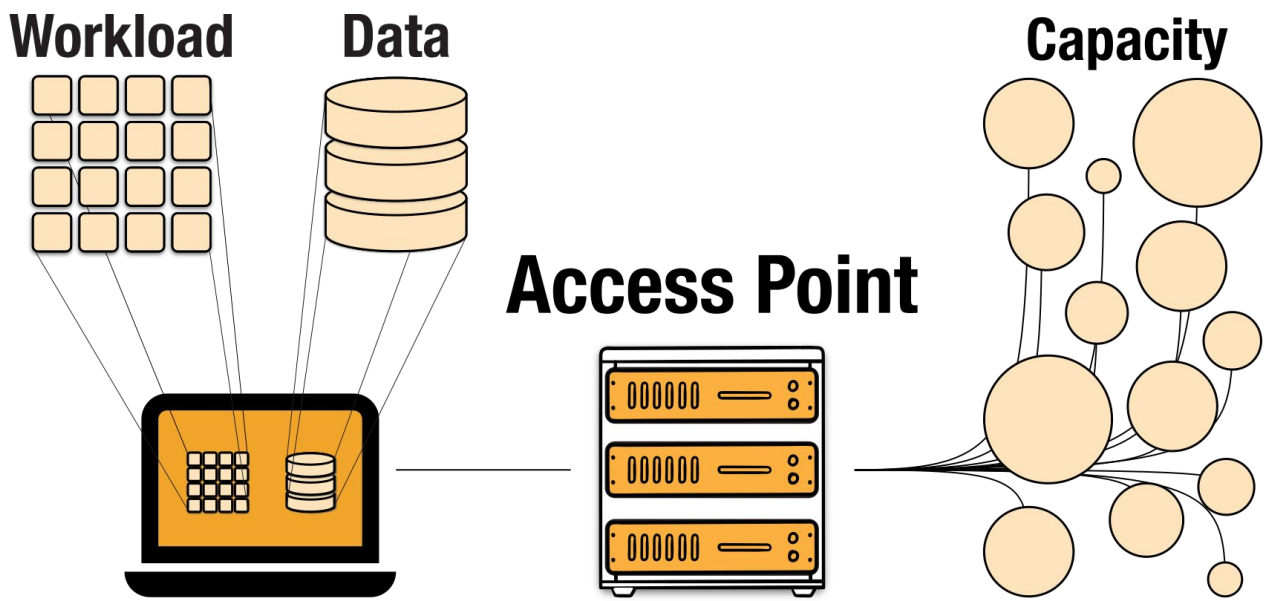
Brian Lin
OSG Software Area Coordinator
Center for High Throughput Computing
University of Wisconsin–Madison

PATh PARTNERSHIP to ADVANCE THROUGHPUT COMPUTING

HT CENTER FOR HIGH THROUGHPUT COMPUTING

The OSG Consortium

- OSG Consortium operates pools of shared capacity for distributed High Throughput Computing (dHTC)
- OSG staff provide:
  - OSG Fabric of Services
  - Software distribution, integration and maintenance
  - Research Facilitation
  - …and much more!



PATh

- PATh: **P**artnership to **A**dvance **Th**roughput Computing
  - OSG Consortium
  - Center for High Throughput Computing (CHTC)
  - Webinar: Harnessing the Power of Distributed HTC ft. Miron Livny
- CHTC is home to the PATh DevTeam, developers of the HTCondor Software Suite (HTCSS)
- HTCSS is a critical part of the OSG Software Stack

PATh PARTNERSHIP to ADVANCE THROUGHPUT COMPUTING

OSG

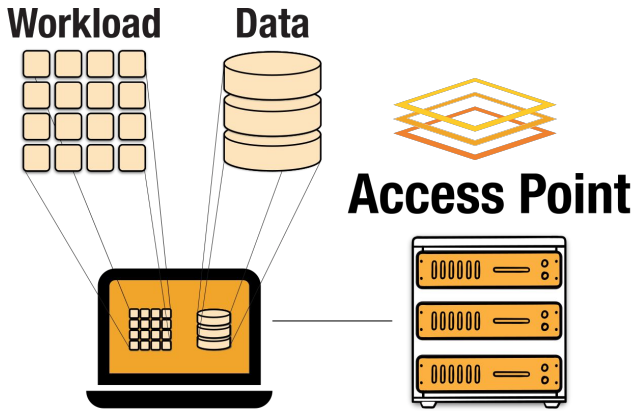
HT CENTER FOR HIGH THROUGHPUT COMPUTING

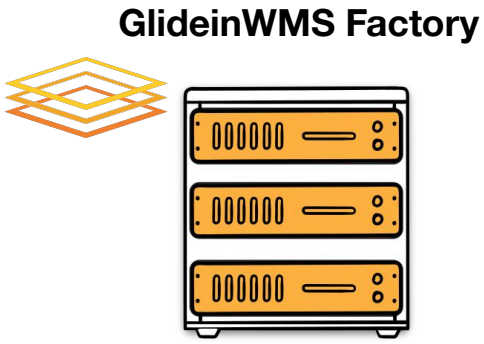
DevOps: combining ***software development*** and ***operations*** to provide ***higher quality*** software, ***more quickly***

OSPool: PATh DevTeam + OSG Ops

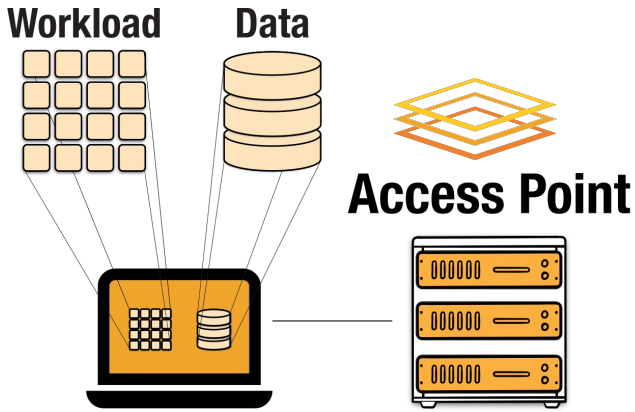
- The OSPool is a dynamic, dHTC service for open science with contributions from hundreds of sites
- Consists of distributed services operated by a distributed OSG Operations Team
- Services all based on HTCSS
- End-user support and documentation provided by OSG Research Facilitators
- No brainer for PATh DevOps!

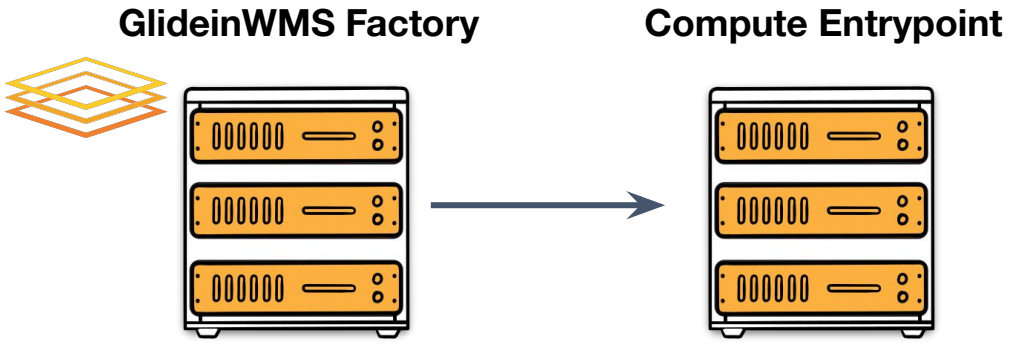
Under the Hood



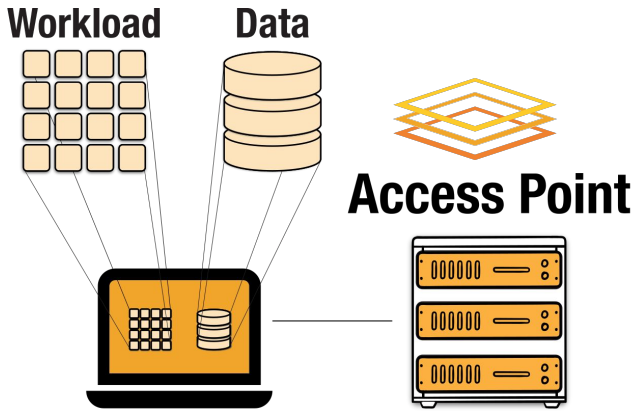


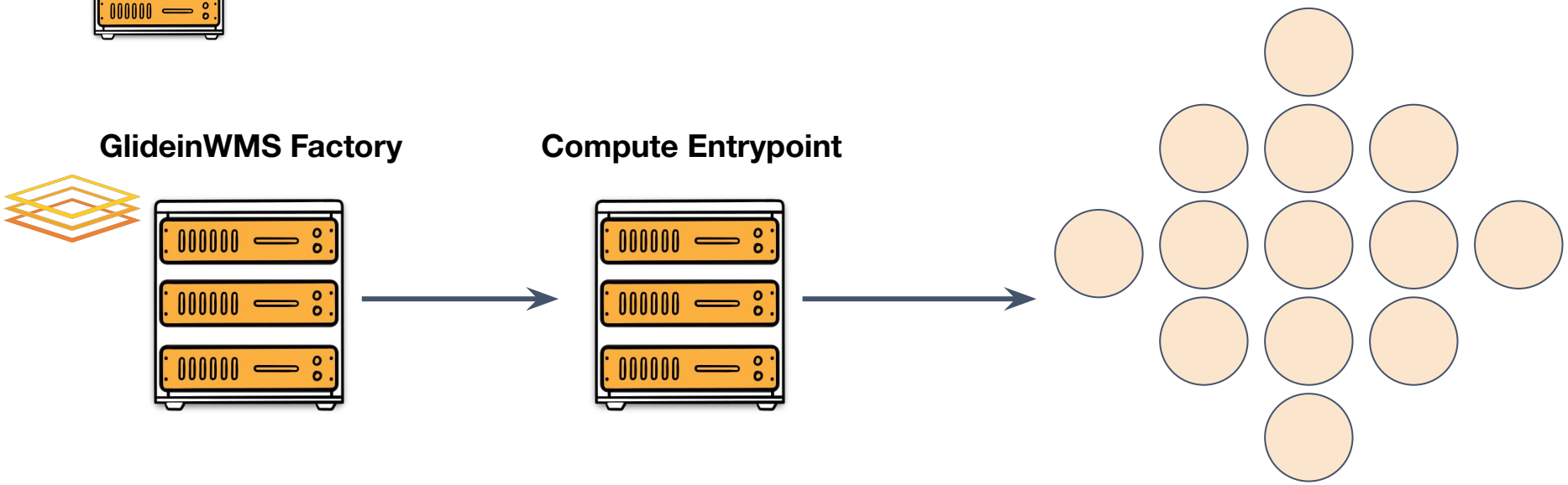
Under the Hood



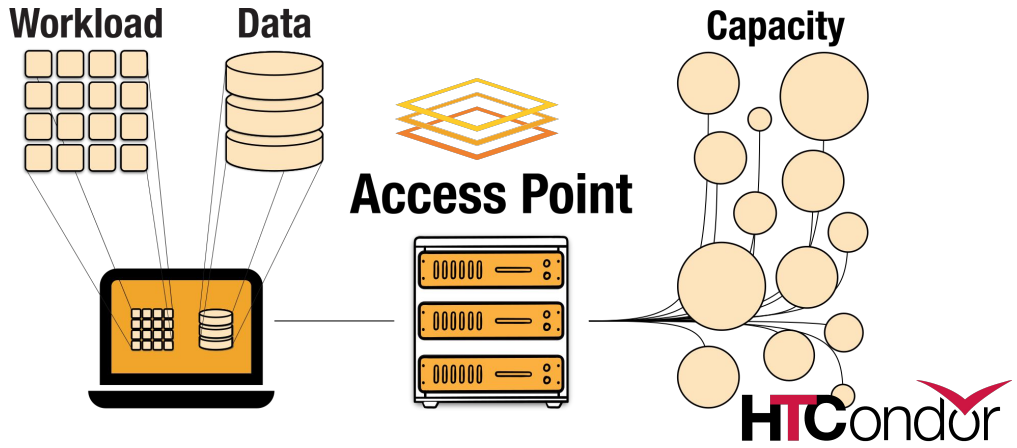


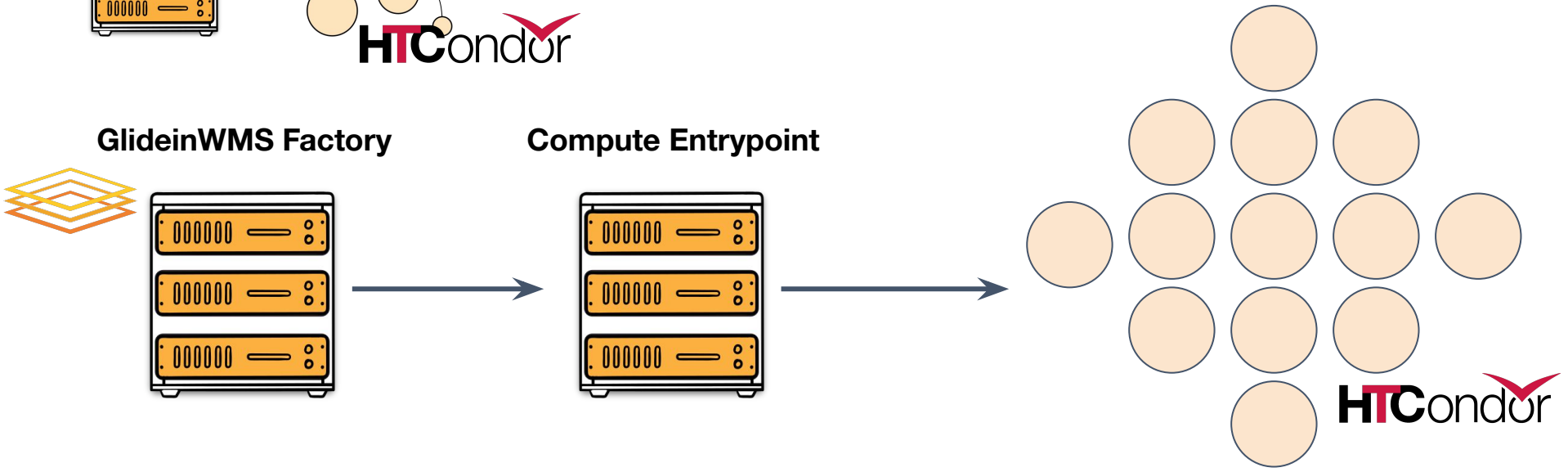
Under the Hood



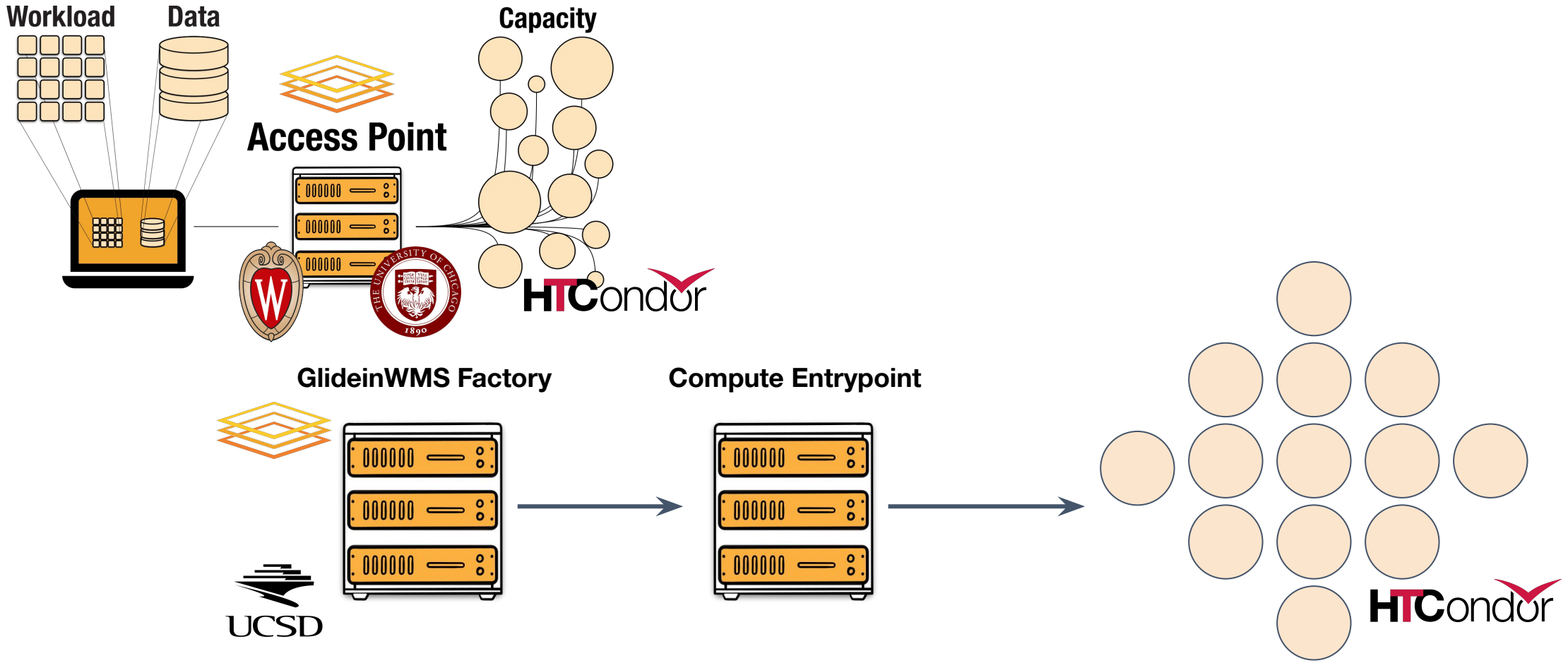


Under the Hood

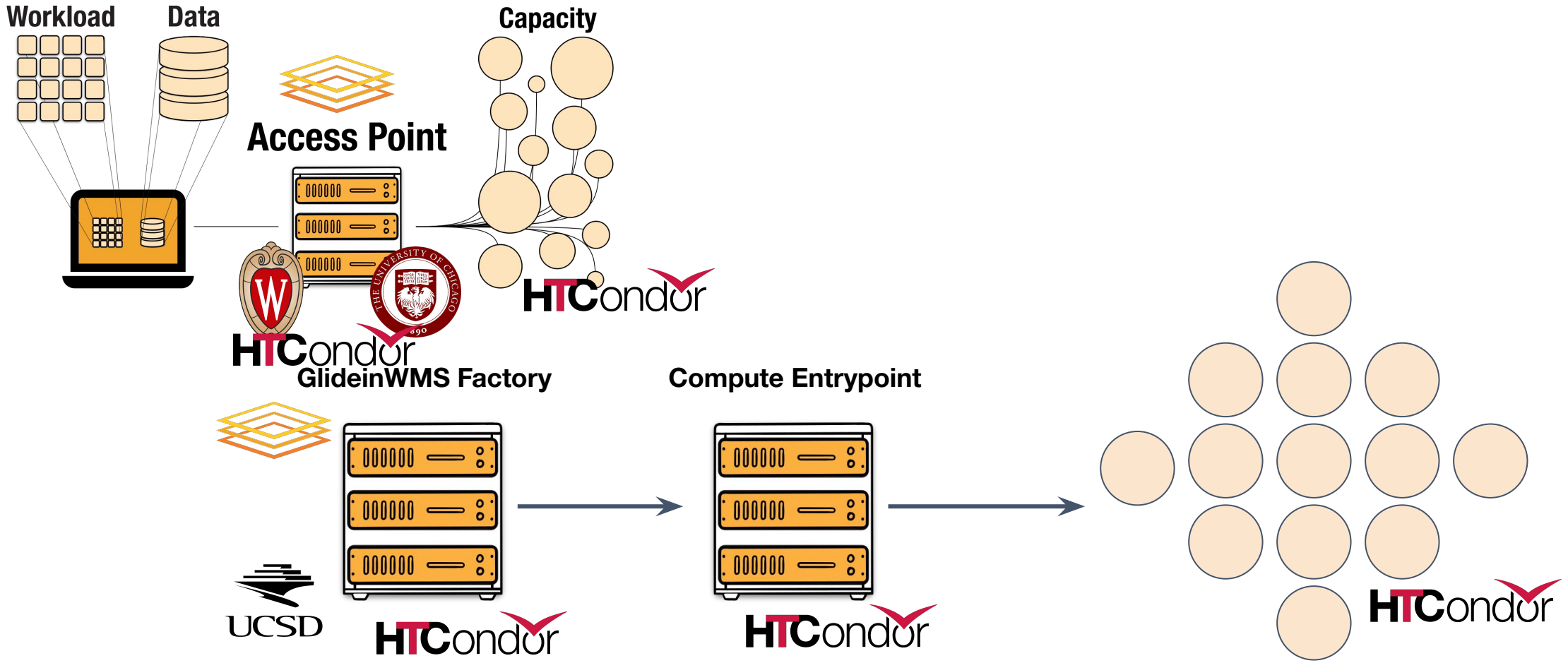




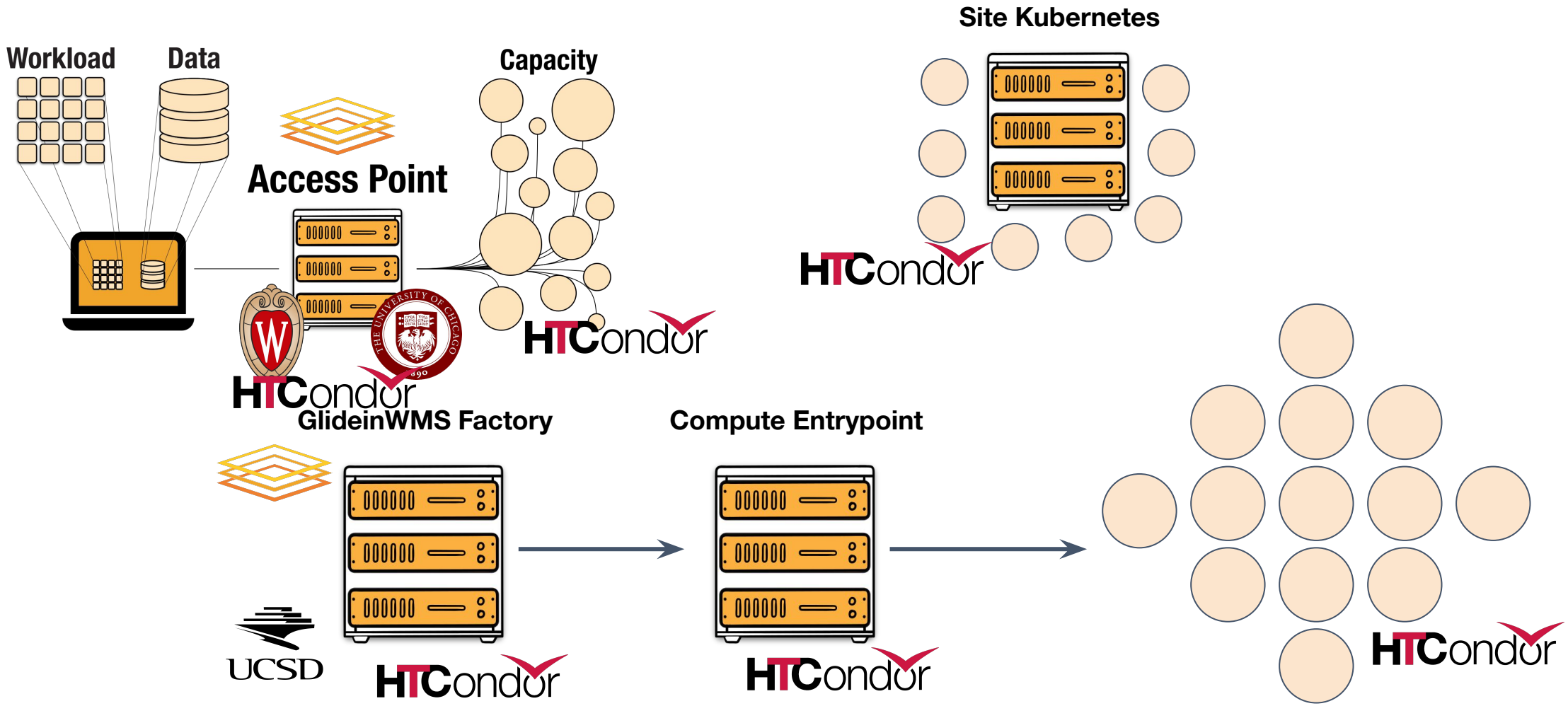
Under the Hood



Under the Hood



Under the Hood



OSPool: PATh DevTeam + OSG Ops

Then…

- Monthly HTCSS releases
- Features and bug-fixes designed for general public
- OSG-Operated Access Points used LTS versions of HTCSS
- OSPool upgraded for specific bug-fixes

Now…

- HTCSS releases and OSPool upgrades every Tuesday
- Features and bug-fixes planned by PATh Dev and OSG teams
- OSG-Operated Access Points use feature versions of HTCSS
- OSPool changes rolled into general public, monthly releases

Building (Virtual) Relationships

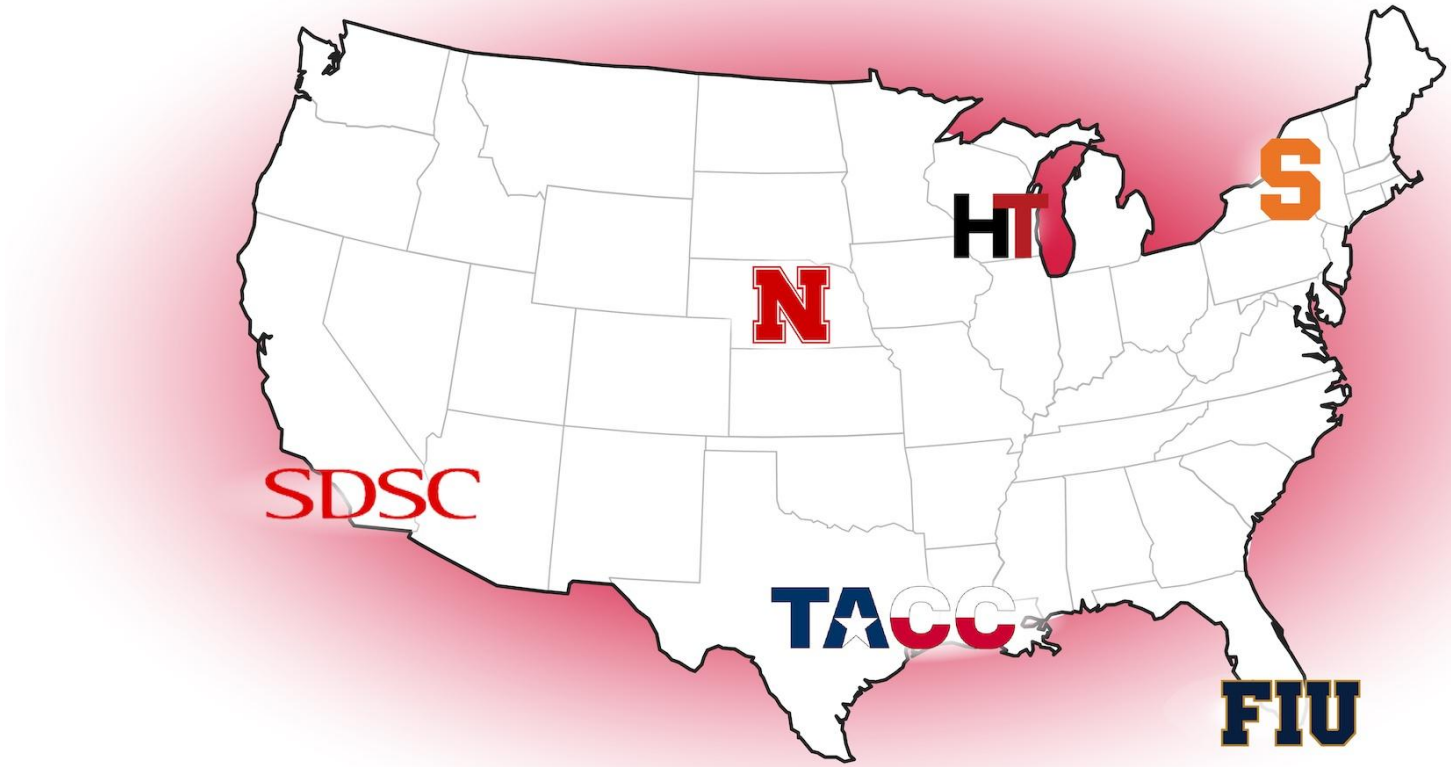
- Minimal technical and policy work to get from Then → Now
- Heaviest lifts around new organizational goals and structure
  - Monthly town hall meetings for leadership to discuss strategic direction and receive input from staff
  - Weekly OSPool operational and all staff meetings
  - Dedicated Slack channels for day-to-day communications
- Our solutions reflect our distributed team in a virtual-only world
  - PATh awarded in September 2020
  - Highlighted weaknesses in our onboarding processes
  - Excited for an in-person All Staff meeting!

OSPool Special Projects

- Historically, parallel developments in different parts of the software stack added support for Apptainer user jobs
- Difficult to untangle due to feature disparity and expertise distributed among teams
- New relationships and processes leveraged to turn an intractable problem into a multi-week project
  - OSPool users will be able to take advantage of currently missing HTCSS features
  - Other HTCSS users will see improvements to Apptainer support

The PATh Facility

Powering distributed high throughput computing



PATh PARTNERSHIP to ADVANCE THROUGHPUT COMPUTING

OSG

HT CENTER FOR HIGH THROUGHPUT COMPUTING

PATh Facility: PATh Production + OSG

- National-scale dHTC service: 30k cores, 36 A100 GPUs
  - Each site maintains hardware and networking
  - PATh Production Services Team provisions Kubernetes at remote sites:
    - Florida International University
    - Syracuse University
    - University of Nebraska
    - University of Wisconsin
  - San Diego Supercomputer Center and Texas Advanced Computing Center provision their own hosts with Kubernetes
- Developers in PATh Production Services deploy dHTC services across distributed Kubernetes infrastructure

PATh PARTNERSHIP to ADVANCE THROUGHPUT COMPUTING

OSG

HT CENTER FOR HIGH THROUGHPUT COMPUTING

PATh Facility GitOps

- GitOps: tracking and making operational changes with Git and Continuous Deployment (CD) systems
  - Service states are enshrined in history, allowing for rebuild in case of catastrophic failure
  - “Who”, “what”, “when”, and hopefully “why” recorded for auditing and knowledge sharing
- Continuous deployment through Flux
- Secrets encrypted with SealedSecret
- Dex for managing operator access

PATh Facility Challenges

- Coordination of troubleshooting hardware and networking issues in a distributed ops environment
- Slightly different Kubernetes environments at sites may require different strategies and/or development efforts
- System complexity can result in difficult-to-troubleshoot issues
  - Potentially mitigated via Operations and architecture manuals
  - New technologies complex at first → organizational education opportunities

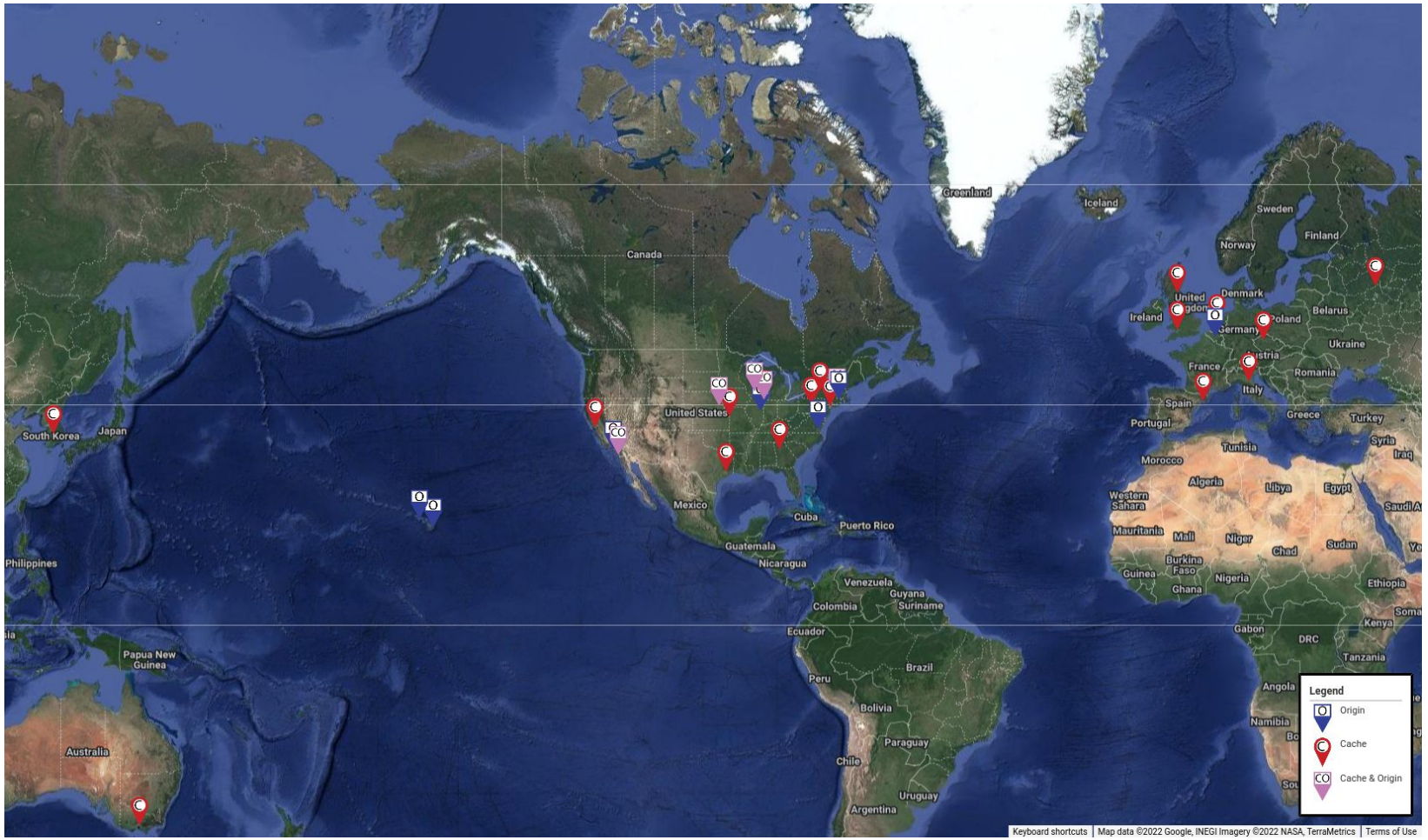
OSG Fabric of Services

- OSG Operations manages a variety of central services for members of the Consortium
- OSG Operations team spread across time zones: coordination (daily standups) and GitOps are critical
- Many of these services have been migrated to Kubernetes clusters at the University of Wisconsin and University of Chicago
- Services fall on the DevOps spectrum, depending on what can be taken off the shelf vs created in-house

OSG Hosted Compute Entrypoints (CEs)

- CEs provide an entrypoint/translation layer from glideins to the local site scheduler
- CEs can be hard for new sites to install and maintain (e.g., CC*)
- OSG Hosted CEs are one of our software-as-a-service offerings:
  - OSG Operations deploys Hosted CE Helm charts, allowing sites to contribute capacity to the Consortium with just an SSH connection
  - Developed by the OSG Software Team and requires close coordination between OSG Software and Operations
  - Where does your product fall on the DevOps spectrum and for whom? We experienced some pain with sites running their own Hosted CEs

Open Science Data Federation



PATh PARTNERSHIP to ADVANCE THROUGHPUT COMPUTING

OSG

HT CENTER FOR HIGH THROUGHPUT COMPUTING

Open Science Data Federation

- Service deployment analogous to CEs:
  - Sites can choose to host their own
  - Or allow OSG Operations to deploy caches/origins within the site's Kubernetes cluster
- Cache and Origin container development by the OSG Software Team, deployment by OSG Operations
- Helm chart development in progress, needs to support different Kubernetes environments

Monitoring

- Monitoring is crucial to DevOps with distributed services
- Various levels of real-time monitoring specific to each service
  - Icinga for monitoring host health
  - Prometheus for health of Kubernetes deployments
  - Check_mk for remote and custom checks for services
  - HTCSS monitoring in Grafana, Ganglia, and custom RRD displays for OSPool health
- Accounting as a final verification tool

Speeding Up Traditional Ops

- Some sites in the Consortium were interested in shorter software development → service deployment lifecycles
- OSG Software Team packaged many services as containers
- Stretch of “DevOps” as developers and operators fall under different administrative domains
- Tight coupling replaced with clear, published policy:
  - Different container tags indicate level of testing performed
  - Immutable container tags to allow sites to revert updates

Building Community Knowledge

- Bi-weekly Kubernetes Hackathons
- Finding opportunities for Kubernetes-based projects involving new teams and team members
- Kubernetes-focused Slack channels
- Cross-pollination via Kubernetes meetups with other groups in the Consortium

Questions?

- Ping us on the IRIS-HEP Slack
- Reach out in #software on the OSG Slack
- Send us email at help@osg-htc.org

This material is based upon work supported by the National Science Foundation under Grant Nos. 1148698, 1836650 and 2030508. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

PATh PARTNERSHIP to ADVANCE THROUGHPUT COMPUTING

OSG

HTC CENTER FOR HIGH THROUGHPUT COMPUTING