



Advancing Open Science through distributed High Throughput Computing

**Frank Würthwein
OSG Executive Director
Professor of Physics
UCSD/SDSC**

October 22nd 2020





What is OSG ?



- **OSG is a consortium** dedicated to the **advancement** of all of open science via the practice of Distributed High Throughput Computing, and the advancement of its state of the art.
- It is a collaboration between IT, software, and science organizations.
 - **The consortium aggregates human and computing resources from many funded projects.**
- It is governed by the OSG Council, maintaining its by-laws, and electing an executive director for 2 year renewable terms to coordinate a program of work.

- All of open science irrespective of discipline
- Advance the maximum possible diversity of science, groups, and institutions
 - From **individual undergraduates** to international collaborations with thousands of members.
 - From **small colleges, museums, zoos**, to national scale centers of open science.
- Advancing this entire spectrum requires us to have a **diversified portfolio of services**

**OSG appears differently to
different people ...**

**Trainer, service provider,
compute and data infrastructure,
scientific user community, ...**

OSG serves 4 distinct groups

- The **individual researchers** and small groups
- The **campus Research Support Organizations**
 - Teach IT organizations & support services so they can integrate with OSG
 - Train the Trainers (to support their researchers)
- **Multi-institutional Science Teams**
 - XENON, GlueX, SPT, Simons, and many many more
 - Collaborations between multiple campuses
- The 4 **“big science”** projects:
 - US-ATLAS, US-CMS, LIGO, IceCube

- Researchers have **command line access on a submit host**.
- We operate a **compute federation**
 - An HTCondor batch system across many clusters
- We provide a **uniform runtime environment** across the compute federation
- We operate a data federation to provide **uniform data access** across the compute federation.



Concept of a “Resource Pool”



- We enable resource owners to dynamically share their hardware resources with many different organizations.
 - Contribute to other people’s “Resource Pools”
 - Fine grained control of priorities stays with the “other people”
- We operate one such resource pool for the benefit of all of open science, focused on individuals and small groups.
 - The “Open Pool”
- We allow spare capacity to “slosh around” between resource pools if the owners so desire.
 - Communities share resources between each others pools.

- OSG's objective is to “Advance Open Science through distributed High Throughput Computing”
- OSG thinks of its science stakeholders in terms of 4 categories:
 - Individual Researchers
 - Campus Research Computing Organizations
 - Multi-campus Science Teams
 - “Big Science” Collaborations
- OSG offers a diversified portfolio of services to support these different science stakeholders.

Contact us at: support@opensciencegrid.org