
Supporting the Snowmass 2021 Process

Pascal Paschos

University of Chicago & OSG Collaboration Support

John Stupak

University of Oklahoma

The logo for SnowMass2021, featuring the text "SnowMass2021" in a stylized font. "Snow" is in light blue, "Mass" is in white, and "2021" is in white. The text is set against a dark grey rectangular background.

SnowMass2021

<http://snowmass21.org>

OSG Council Meeting

September 3, 2020



Open Science Grid

Outline



- Snowmass Process
- Energy Frontier MC taskforce goals
- Dedicated OSG submit host - "Snowmass Connect"
- Software support
- Local storage & access from OSG
- Distributed storage (plans)
- Supporting Tutorials
- Outlook

Snowmass Process



- Particle physics community planning exercise
 - Organized by the Division of Particles and Fields (DPF) of the American Physical Society (APS)
 - Every 8-10 years, since 1982
- The goal is for the entire particle physics community to come together to identify and **document a scientific vision for the future of particle physics** in the U.S. and its international partners.
 - The Particle Physics Project Prioritization Panel (P5) will take the scientific input from Snowmass and develop a strategic plan for U.S. particle physics that can be **executed over a 10 year timescale, in the context of a 20-year global vision for the field.**
- [More details and historical context](#)

Snowmass Frontiers

ENERGY FRONTIER

NEUTRINO PHYSICS FRONTIER

RARE PROCESSES AND PRECISION

COSMIC FRONTIER

THEORY FRONTIER

ACCELERATOR FRONTIER

INSTRUMENTATION FRONTIER

COMPUTATIONAL FRONTIER

UNDERGROUND FACILITIES

COMMUNITY ENGAGEMENT FRONTIER

Energy Frontier



- Energy frontier reliant on high-energy particle colliders
- Within Snowmass, goal is to assess physics potential/complementarity of various proposed colliders
 - 374 letters of interest submitted so far!

Collider	Energy							
	250	350	380	500	1000	1500	3000	
linear ee	250	350	380	500	1000	1500	3000	GeV
circular ee	m_z	$2m_w$	240	$2m_t$				
hh	14	75	100	150				TeV
eh	1.3	3.5						
$\mu\mu$	3	10	14	30				

Energy Frontier MC Task Force



- To support these studies, large statistics of Standard Model (SM) background Monte Carlo (MC) samples are required
- Backgrounds are:
 - Common among many studies
 - CPU intensive to produce
 - Large
- Goals of the TF:
 - Produce comprehensive SM background samples for each proposed collider
 - Provide access to anyone in the community
 - Provide resources/documentation for individuals to produce their own signal MC samples

Name	Institution
John Stupak (chair)	University of Oklahoma
Robert Gardner	University of Chicago
Simone Pagan Griso	LBNL
Stefan Hoeche	FNAL
Fabio Maltoni	CP3, Catholic University of Louvain
Meenakshi Narain	Brown University
Isabel Ojalvo	Princeton University
Laura Reina	Florida State University
Michael Schmitt	Northwestern University
Alessandro Tricoli	Brookhaven National Laboratory

Snowmass2013

OSG rep.

BSM rep.

MC expert

MC expert

EF convener

EWK rep.

EF convener

QCD rep.

EF convener

Estimating Compute Requirements

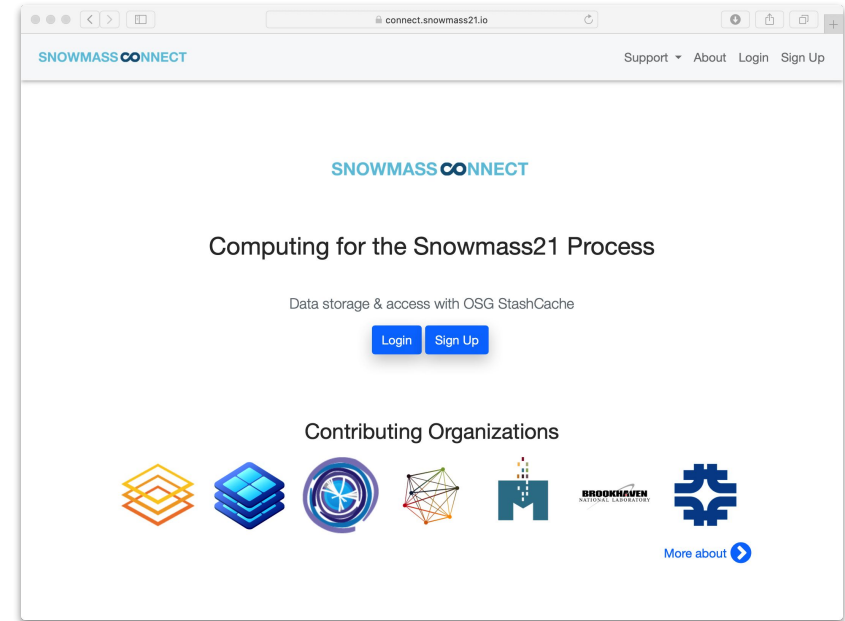


- During Snowmass 2013
 - ~14k jobs per day
 - Peak usage: 100k CPU/day
- Expect to scale up in compute time:
 - Support studies for a wider array of proposed colliders
 - May need larger statistics in samples collected
 - The sample content might be enhanced wrt 2013
- Effort to consolidate dependencies on multiple CVMFS repos into a single one
- Will be testing for computational scale on OSG of Delphes; the “Snowmass Detector” implementation

Snowmass21 Connect (to the OSG)



- Deployed dedicated OSG Connect instance
- 40 users at present from 20 different institutions
- Subgroups created for cosmic and energy frontier working groups
- Accounting, access to login/submit host and JupyterLab portal



Credit: SLATE, IRIS-HEP, OSG, UC-PSD

Submit node

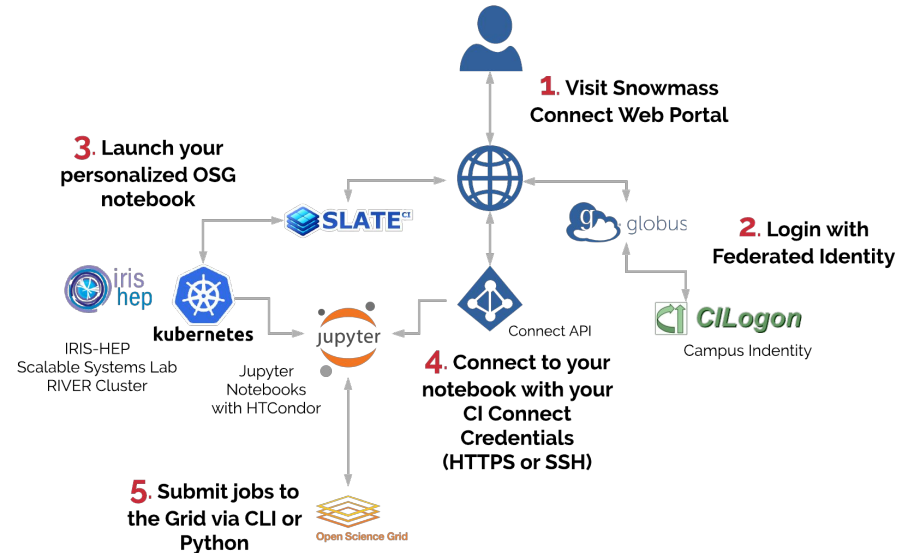


- Login host: login.snowmass21.io
 - OpenStack VM 64 GB 16 VCPU
 - NFS mounted home
 - Access to OSG Stash and CVMFS
 - Monitoring dashboards in Grafana, ganglia
- Physical machine will be deployed as they scale up:
 - 48 logical cores, 200TB of local disk, 10Gbps, 192GB RAM
 - Will provision scratch space on disk for users and local software builds
- Purpose:
 - HTCondor submit point
 - An interactive compute resource for software builds
 - Pre/post-processing results and local analysis jobs

JupyterLab to OSG Capability



- **Under development**
Capability to deploy a personal JupyterLab + HTCondor application from the SLATE catalog using Kubernetes
- IRIS-HEP SSL cluster at UC used on the backend to host the notebook
- Users can submit to OSG *from the notebook*



Credit: SLATE and IRIS-HEP

Data Management



- Current estimate is 300 TB storage. Scale up as needed
- Discussed in meetings with stakeholders and OSG technical teams
- Planning on three origin servers (UChicago, BNL & FNAL)
- Will use the OSG data federation (StashCache)
 - Write once, read many
 - stashcp for data access
- Single OSG redirector. Single namespace in CVMFS yet to be organized, e.g. by role, group, user

Software



- Created OSG OASIS external software repo:
`/cvmfs/snowmass21.opensciencegrid.org`
- Origin: `osg-cvmfs.grid.uchicago.edu`
- Add a collection of compiled code, consolidate tools and make it available over CVMFS jobs on OSG
- Software so far: Delphes, LCIO/miniDST, CEPC simulation framework and dependencies
- Containers as a longer term solution

Training and Tutorials



The Snowmass Connect login server as a component in the delivery of tutorials currently organized for taskforce members

- CEPCSoft training event Sept 8 (Dan Yu IHEP) [link](#)
- MC ILC training event Aug 28 (Jenny List DESY, Jan Strube Oregon)
 - 70 people signed up for the tutorial on LCIO and Jupyter notebooks. About half of them used the Snowmass Connect node.

Physical Resource Contributions



- UChicago Physical Sciences Division providing login server and 200 TB local storage
- Brookhaven lab - pledged storage for all Snowmass 2021 Computing
- Fermilab will provide some storage
- Depending on demands OSG pool, may want to organize a *cause-campaign* to allow interested institutions to contribute with priority

Outlook

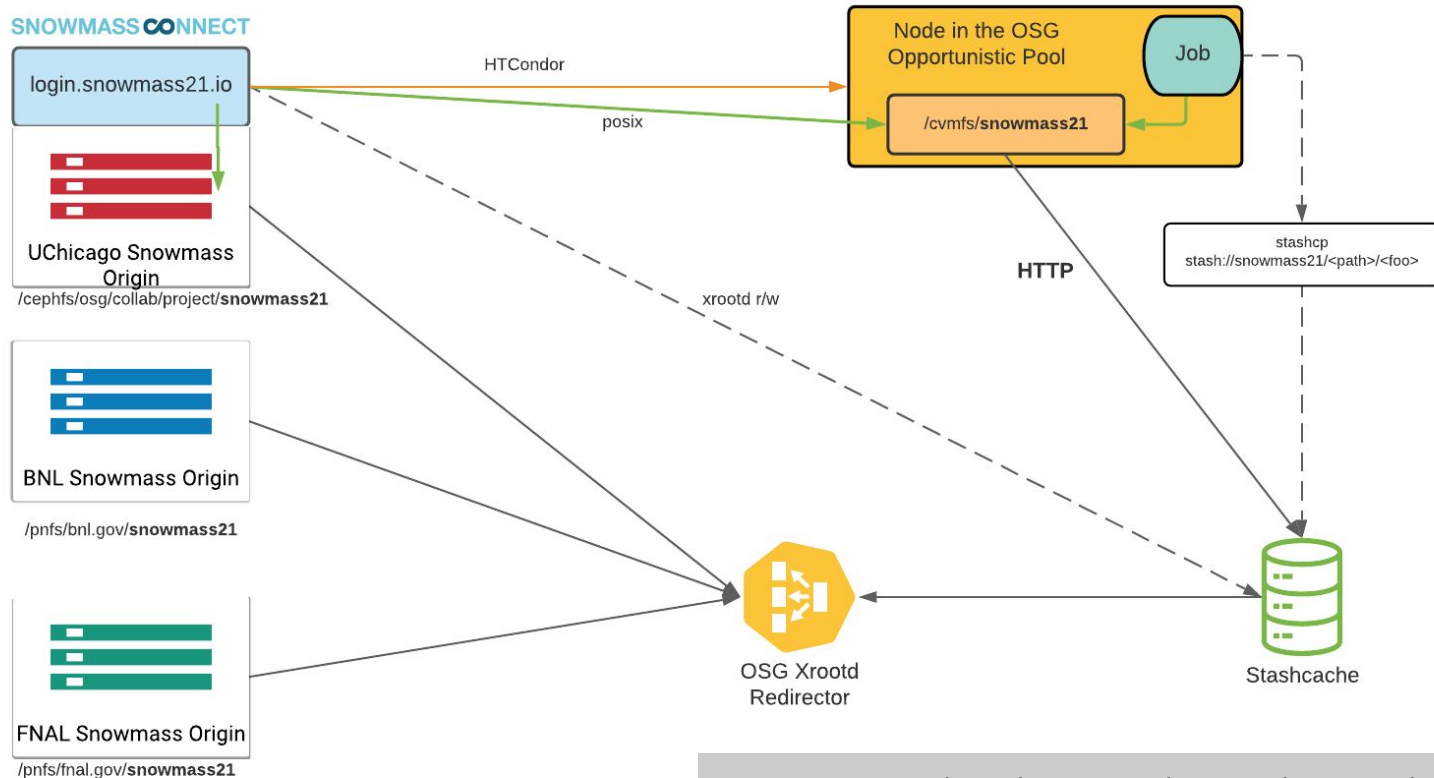


- Infrastructure coming together
 - Setting up physical server
 - Testing and validating data and software origins
 - Work on shared CVMFS namespace among the origins & redirector -
Derek Weitzel, Edgar Fajardo
- Moving to production for Jupyter user notebooks from the portal
- Refining user support model, expanding documentation
- CEPC Software with updated version bundles of common software versions CLHEP Geant4 HepMC ROOT
- Expect production jobs by October 2020

Extras follow



StashCache for Snowmass



Data Access



- 50 TB of CephFS in stash storage in user and project space
 - Private user namespace in `/collab/user/<user_id>`
 - Public project namespace for data in `/collab/project/snowmass21/data`
 - Installed software in `/collab/project/snowmass21/software`
- Project namespace is stashcp accessible from the grid
- Project namespace is also HTTP accessible from <http://stash.osgconnect.net/collab/project/snowmass21>
- Project accessible from a Globus door deployed via SLATE for Snowmass21 and other collaborations
 - Globus collection: [#OSG Connect CI Logon](#)

Documentation & Support



- Put together some getting started user [documentation](#)
 - <https://maniaclab.uchicago.edu/snowmass-connect-docs/>
- Help in onboarding new users to the OSG and DHTC training
- Provide specific instructions for data management
- Provide specific information for installed HEP simulation as requested by the MC taskforce
- Connect users with help@opensciencegrid.org
- Setup **#snowmass21-connect** channel in the **snomwass2020.slack.com** workspace