OpenCosmo Soft Launch and Feedback Session

DESC Sprint Week, Oct. 30 2024

The problem

Data are…

- Massive (Petabyte scale)
- Distributed (On many systems)
- Heterogeneous (In different formats)

Scientists are…

- Distributed (Differently from the data)
- Heterogeneous (Interested in different things)
- Time-constrained (no explanation needed)

OpenCosmo

Goal: Create tools that give researchers access to **data** and **compute** without requiring accounts on or direct interaction with the associated HPC systems.

Data from MiraTitanU Snapshots

The Mira-Titan Universe simulation suite was carried out on Mira, a supercomputer at the Argonne Leadership Computing Facility, and Titan, at the Oak Ridge National Laboratory. The simulations cover a range of cosmological models including models with a dynamical dark energy equation of state parameterized via w_0 and w_a. Each simulation covers a $(2.1\text{Gpc})^3$ volume and evolves 3200^3 particles. We provide outputs for 27 redshifts, between z=4 and z=0, including halo information and down-sampled particle information.

Please select one or more models from the list below, then select all the relevant redshifts and data products. The SubmitTransfer button will indicate the number and overall size of the selected files that you aim to transfer. This button will lead you to the Globus interface. The Search box at the top allows you to narrow the model selection by specifying a model number or a numerical value for any cosmological parameter.

Search:

	Model	Ω_{cdm}	ω_b	Ω_ν	h	σ_8	n_s	w_0	w_a
☐	M000	0.22000	0.02258	0.0	0.7100	0.8000	0.9630	-1.0000	0.0000
☐	M001	0.32760	0.02261	0.0	0.6167	0.8778	0.9611	-0.7000	0.6722
☐	M002	0.19970	0.02328	0.0	0.7500	0.8556	1.0500	-1.0330	0.9111
☐	M003	0.25900	0.02194	0.0	0.7167	0.9000	0.8944	-1.1000	-0.2833
☐	M004	0.29710	0.02283	0.0	0.5833	0.7889	0.8722	-1.1670	1.1500
☐	M005	0.16580	0.02350	0.0	0.8500	0.7667	0.9833	-1.2330	-0.0445
☐	M006	0.36430	0.02150	0.0	0.5500	0.8333	0.9167	-0.7667	0.1944

Hydro Simulations

- Galaxy Query
- Halo-Particles Query (Hydro)
- Halo Query (Hydro)

Gravity-Only Simulations

- Halo-Particles Query (Gravity)
- Halo Query (Gravity-Only)
- Map Query (Gravity-Only)

Contact Us

Welcome to the OpenCosmo Portal!

Select a task to get started

Galaxy Query

This flow retrieves and filters galaxies from HACC simulations

Halo-Particles Query (Gravity)

This flow queries the particles of a given halo (run the "Halo Query" first to get a halo ID!). Particles and surface density map are returned in an HDF5 file.

Halo-Particles Query (Hydro)

This flow queries the particles of a given halo (run the "Halo Query" first to get a halo ID!). Particles and surface density maps, diviced by particle species, are returned in an HDF5 file.

Halo Query (Gravity-Only)

This flow retrieves and filters halos from HACC gravity-only simulations

Halo Query (Hydro)

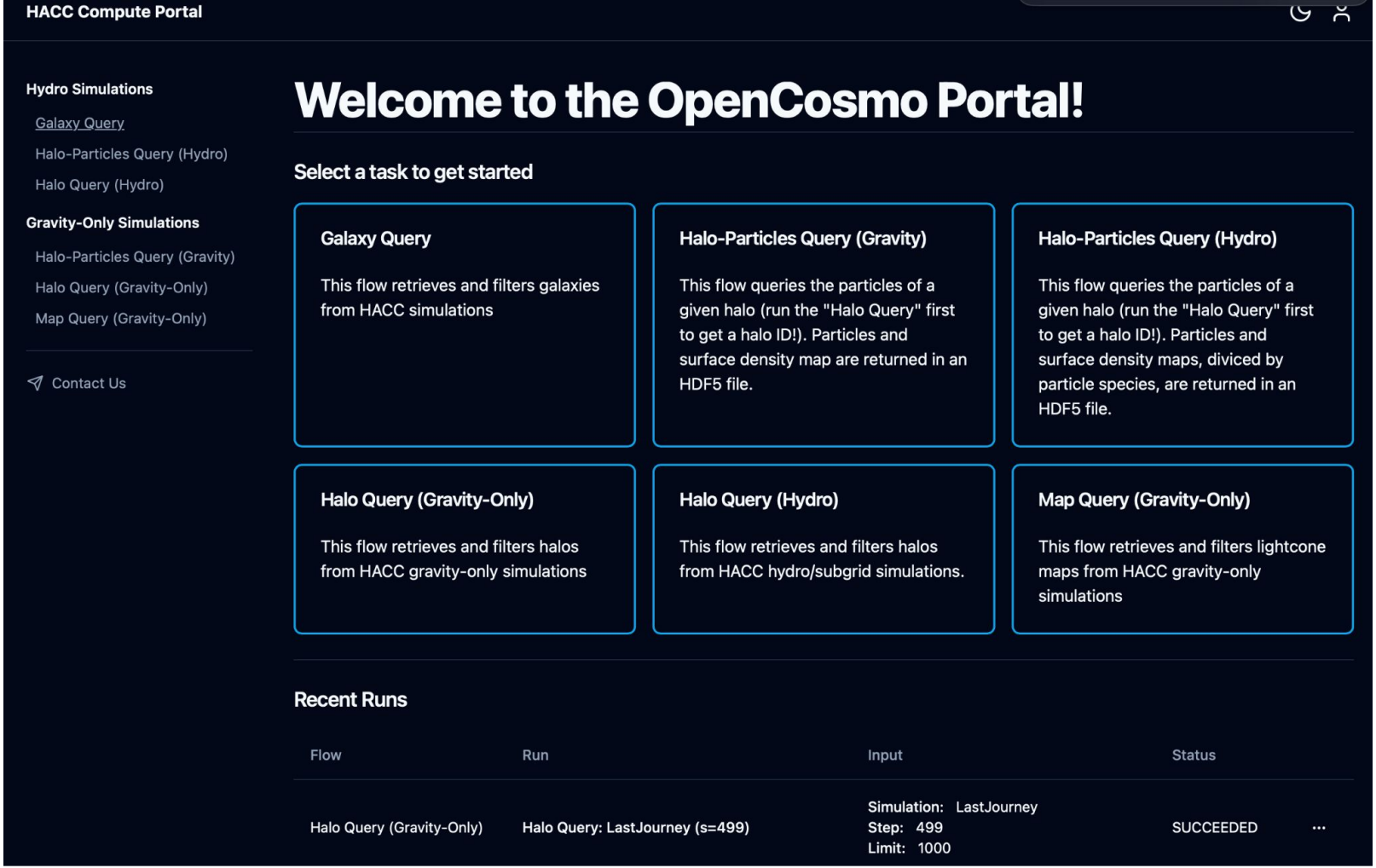
This flow retrieves and filters halos from HACC hydro/subgrid simulations.

Map Query (Gravity-Only)

This flow retrieves and filters lightcone maps from HACC gravity-only simulations

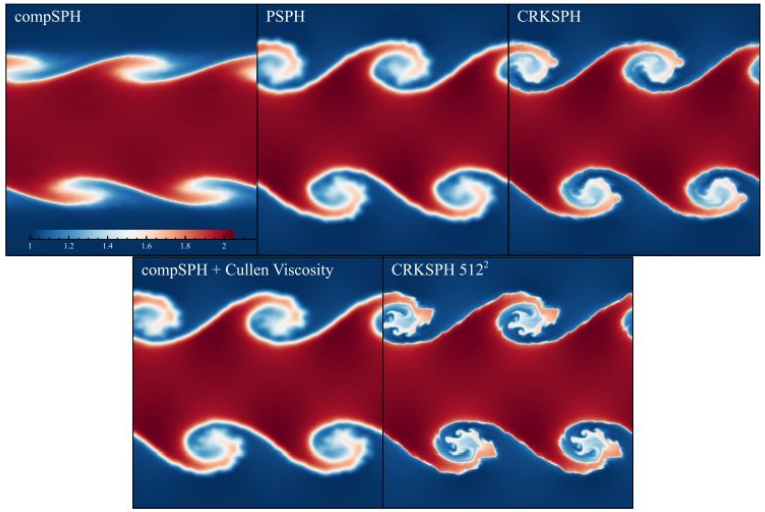
Recent Runs

Flow	Run	Input	Status
Halo Query (Gravity-Only)	**Halo Query: LastJourney (s=499)**	**Simulation:** LastJourney **Step:** 499 **Limit:** 1000	SUCCEEDED ···



Hardware/Hybrid Accelerated Cosmology code (HACC)

- Highly scalable and optimised to run on various HPC architectures
- Gravity only n-body simulations solve the Vlasov-Poisson equations, a 6 dimensional PDE, using tracer particles evolved by gravitational forces. Hybrid solver makes this scalable.
- Hydro solver evolves fluid equations using conservative reproducing kernel SPH formalism (CRK-SPH https://arxiv.org/abs/1605.00725)
- Now with extra subgrid models!

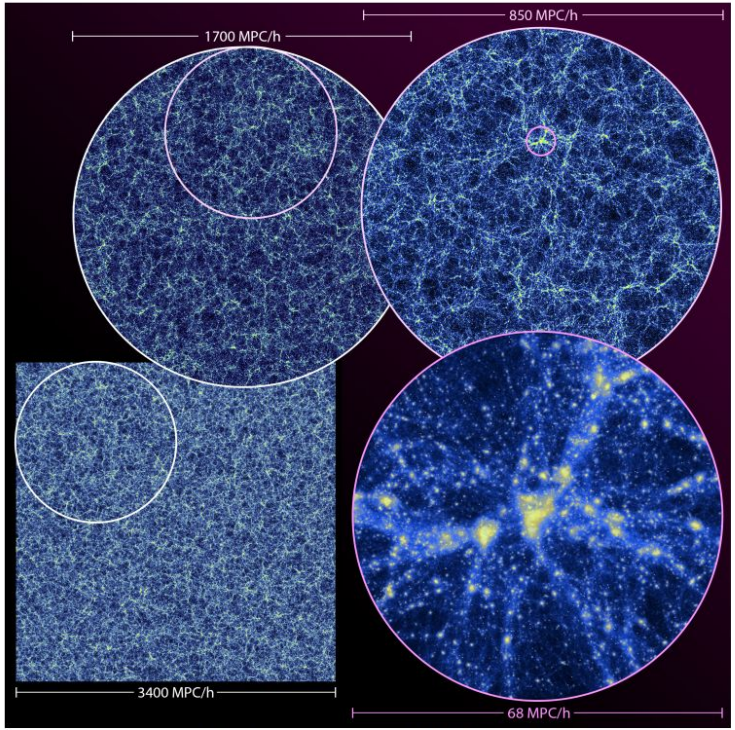


Growth of Kelvin-Helmholtz instability, Frontier et. al 2016

The simulations!

The new Frontier - E

- Large-scale gravity only simulations like Outer Rim, Farpoint, Last Journey, Mira-Titan and more!
- Outer Rim is used as the base for the DESC DC2 simulation which provides the data previews!
- **New sim alert:** Frontier-E GO is a 3.15 Gpc/h gravity-only simulation with a ~10^8 Msun/h mass resolution. Halos are measured on lightcone particles out to z=5, with substructure measurements from halo cores and full particle outputs for massive clusters.
- Full hydro capabilities recently added to HACC. A paired hydro simulation with the same parameters as Frontier-E GO is currently in progress.



Last Journey, Heitmann et. al 2020

Discussion of simulations

Possibly mention cosmodc2 is based on OuterRim?

- HACC description
- Massive gravity-only runs - e.g. Outer Rim, Farpoint, Last Journey, Mira-Titan, New Worlds, Q-continuum. Used for creating mock skies in DESC such as CosmoDC2->DC2->DP0.2
- Full hydro capabilities added, large-scale runs are ongoing
- This simulation is a __Gpc gravity-only box with a ___ mass resolution. Halos are measured on lightcone particles out to z=_, with substructure measurements in the halos and full particle outputs for the halos out to 2r200. A paired hydro simulation is currently in progress.

Goals for the Sprint

Goals for this sprint

- Test-run of the portal
- Feedback on positive/negative aspects of the portal
- Suggestions for useful analysis tools, data access methods

Data highlights

Previous Sprint (Winter 2024)

ALCF resources → highlighting example snapshot hydro data

Video recording of previous sprint
Github with example data and notebook

Resulted in many useful portal updates

This sprint (Autumn 2024)

OLCF resources → highlighting lightcone products from new gravity-only simulations

The machine - Polaris at the ALCF

- A huge thank you to ALCF for hosting us on Polaris!
- 560 node analysis machine
- Each node has: CPU with 512GB memory, and four A100 GPUs

Image credit: Argonne National Laboratory

DEMO? Possibly record something in case we run into issues?

Challenges!

Here are some tasks which should highlight tasks you might like to do in with the simulation, try them out or get data for your own science

- Cluster inspection
- Understanding hydro data
- Make your own galaxy catalog

Try it out now: cosmoexplorer.alcf.anl.gov

Understand the hydro data

- Use the halo query to download halos from the SCIDAC_T001 simulation at step 624 (z=0) with M200c>1.e13, including profiles and galaxies (cap at 100,000 halos)
- Get the haloparticles for FOF_tag=265534890
- The grav-only queries are at the top of the notebook if you're interested in looking at those too

- Download the notebook here showcasing data exploration and follow along for the hydro examples ()
  - https://github.com/ArgonneCPAC/CosmoExplorer-DESCEvent/

Portal offline? Download the data and notebook here

Last Journey cluster inspection

- Locate the most massive cluster in Last Journey step 499 (try asking for a catalog of halos with mass >1.e15, note down the fof_halo_tag)
  - Plot the density profile of this cluster (request profiles)
  - Get the particles and remake this density profile (try a haloparticles query for the tag)
  - Visualize these particles if you like (e.g. using matplotlib)

Frontier-E map query

- From the angular coordinates RA=60, dec=-40 extract the density maps out to 10 degree distance (run the mapquery)
- Inspect your healsparse map!
- Watch this space

Frontier-E cluster inspection

- Locate the most massive cluster on the lightcone (try asking for a catalog of halos with mass >1.e15, note down the fof_halo_tag)
  - Plot the density profile of this cluster (request profiles)
  - Get the particles and remake this density profile (try a haloparticles query for the tag)
  - Visualize these particles if you like (e.g. using matplotlib)
  - From the angular coordinates of the cluster (healpy's vec2ang can help here) extract the density maps around this cluster out to 1 degree distance (run the mapquery)

DIY galaxy catalog

Need to know the halo lightcone queries for this,

- Get all halos with M200c above 10^13, within RA/dec and redshift bounds from the Frontier-E GO simulation
- Use an HOD model to populate these halos with galaxies based on their mass and SOD profile parameters
- Get the density maps in this area to create lensing maps for these galaxies

Vis example

Tasks:

- Visualize the largest cluster
  - Find the largest cluster in the simulation

    (the following won't work surely, I think we haven't set up the vis stuff, but we can show people how to grab the particles)
  - Grab the particles
  - Put them into paraview (tram - hopefully Silvio can set this up)
  - Inspect the halo
  - Look at the density field from the maps about that

Vis example

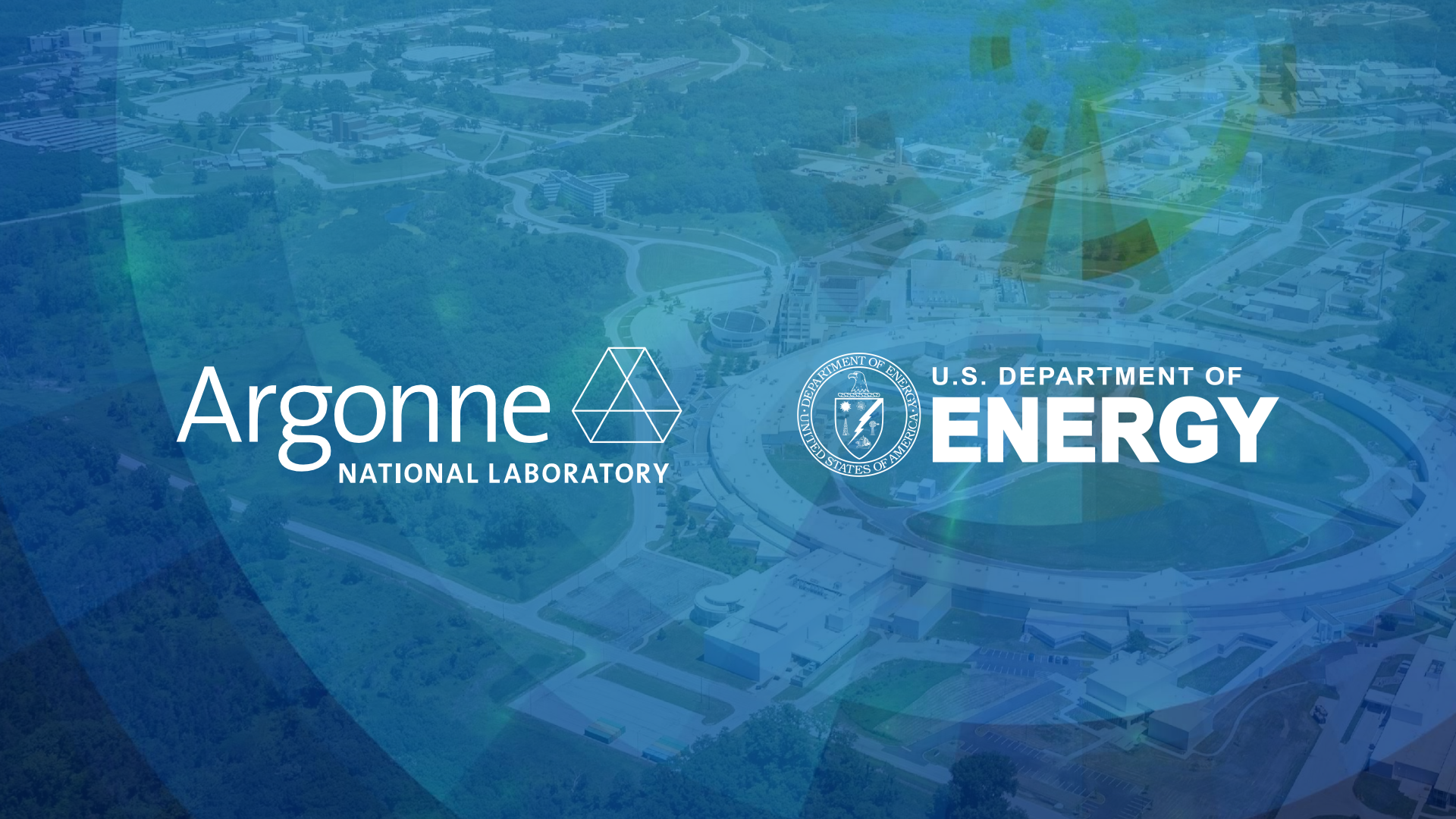
Tasks:

Visualize the largest cluster though properties, particles, and maps:

- Perform a halo properties query, with mass cut > 5.e14, look at outputs and find the tag of the largest halo!
- Halo particles query on that halo tag (do we have this?)
- Grab the particles and vizualize however you like e.g. using matplotlib)
- Find the RA/dec position (healpy vec2ang will do that for you), and extract the map around this step and look at what the density field looks like!

SECTION BREAK


Argonne
NATIONAL LABORATORY


Argonne
NATIONAL LABORATORY
U.S. DEPARTMENT OF
ENERGY