CESM & New Technologies:

Clouds, Containers and Accelerators (GPUs)

Brian Dobbins
NCAR/CGD

June 15th, 2022

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

CESM in the Cloud

CESM2-ARISE Experiments

Geoengineering ensembles:

- 1-degree WACCM case
- >15M CPU hours
- ~1PB of raw output

Full workflow on AWS (build, run, postprocess, archive)

CTSM Workshop

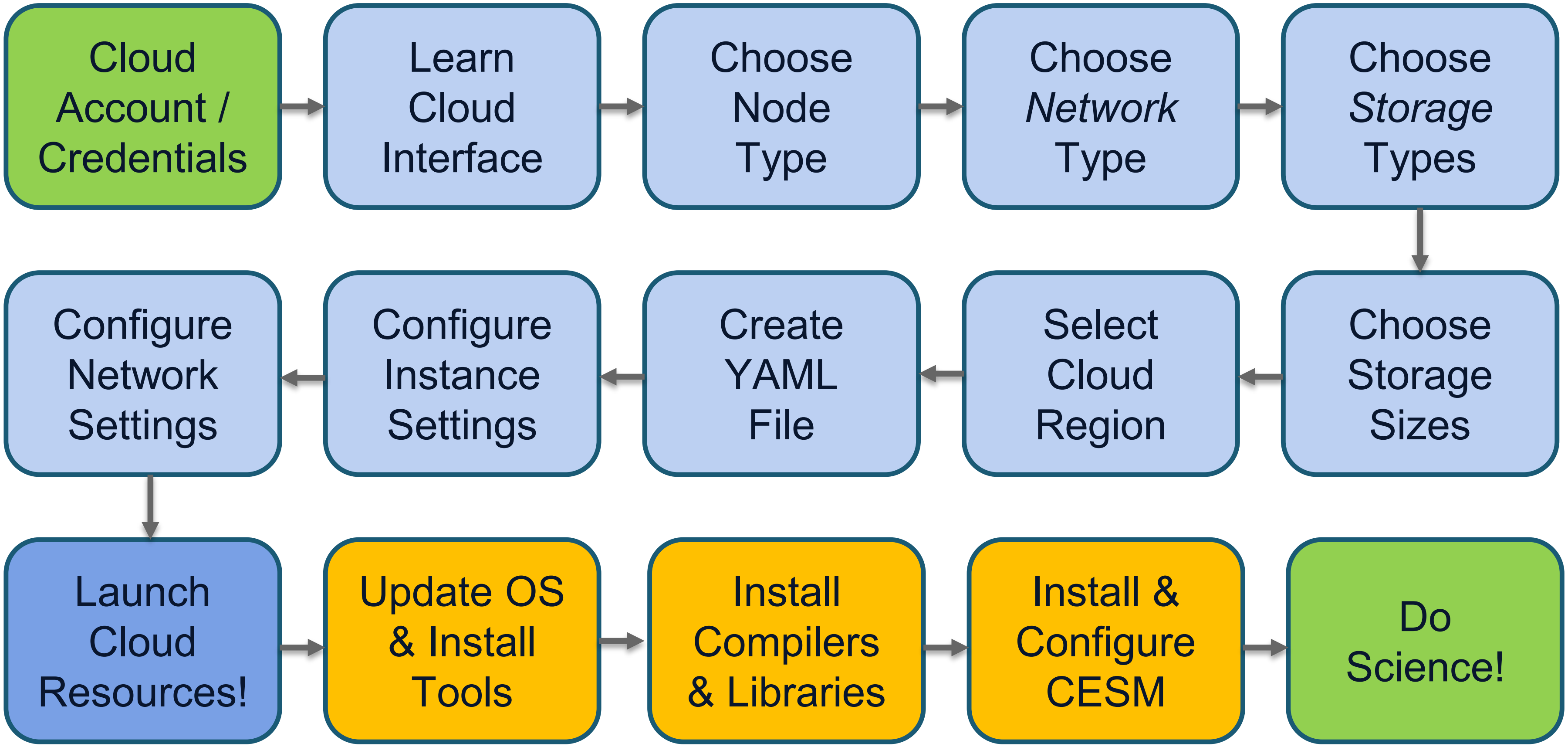
JupyterHub-based training:

- 60+ users
- Auto account creation
- Multiple queues / nodes

Simple, fast deployment for many-user training.

Current projects range from large-scale runs to workshops

How do we use the Cloud?



How do we use the Cloud?

There are lots of decisions that go into running on the cloud; we are simplifying that via our *CESM Cloud API:*

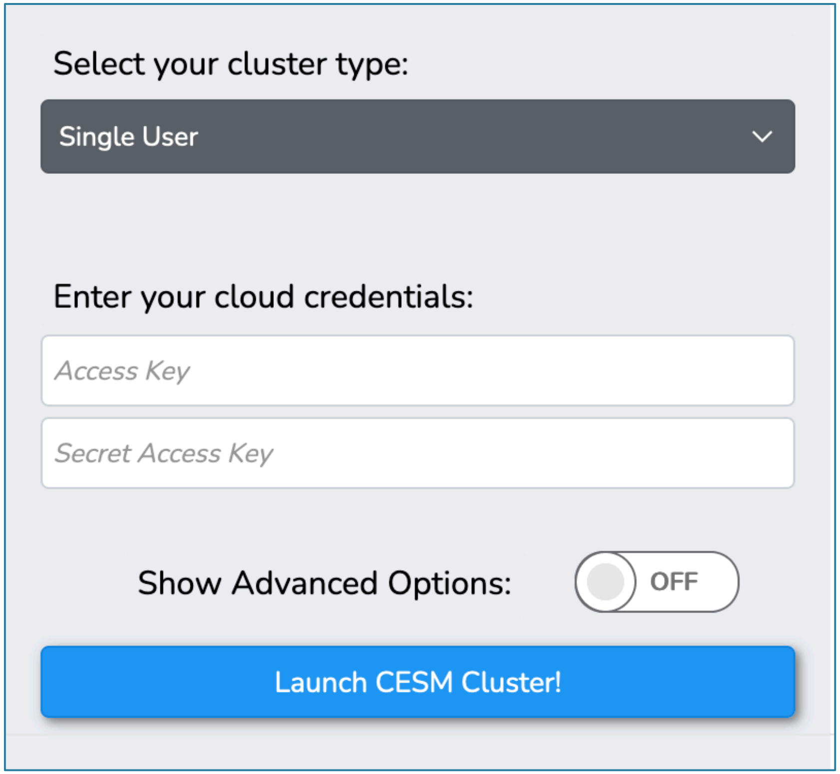
User w/ Cloud Account → **CESM Cloud API** → Fully-Configured System

This offers several key advantages:

- Ease-of-use for scientists
- Multi-cloud support via a single interface
- Extensible features – JupyterHub, ‘Phone Home’, tools, etc.

Cloud API Interface Examples

Website	Linux CLI




```
[bdobbins@cheyenne1] $ cesm create cluster –v
Info: Found AWS credentials: **xT3g..
Info: Mode: Cluster[coupled]
Info: Nodes: C5N.18xlarge
Info: Querying regions for pricing
Info: Selected region US-East-1
Info: Requiring EFA
Info: EBS for /home (40G)
Info: Lustre for /scratch (2.4T)
Status: Creating cluster configuration
Status: Launching cluster
Status: Installing compilers on head node
…
  Success!
   Connect via: ssh bdobbins@52.15.115.58

[bdobbins@cheyenne1] $
```

Cloud Economics (Compute)

CPU Costs (*On-Demand Pricing*)			
Node	AWS C5N	AWS HPC6a	Azure HBv3
Cost (per hour)	$3.88	$2.88	$3.60
Cores (per node)	36	96	120
Cost (per core-hour)	0.11	0.03	0.03

Derecho 'Cost' Equivalent*: **~0.003** / core-hour

(*) Estimate based on ~$32.5M for CPU portion, $1M/year for power, 5-year lifetime, *no personnel costs,* no NWSC facility costs

Cloud Economics (Data)

Cost Comparison		
20 TB	AWS S3	Azure Blob
Storage Cost (Per Month)	$471.04	$425.98
Egress Cost	$1792.00	$1443.50

AWS / ASDI has been *very* supportive and hosts certain *public datasets* for free, like CESM2-LENS and CESM2-ARISE.

For other data, though, the cloud offers two main data hosting options:
- Host pays (potentially unbounded cost / financial denial-of-service)
- Downloader pays (more limited access)

Containers

Clouds offer preconfigured, standardized environments …
on demand from a variety of providers.

Containers offer preconfigured, standardized environments *...*
on your own hardware.

This makes it ‘easy’ to provide *ready-to-run* tools to the community:

- Greatly simplifies ease of porting
- Can ensure cross-system compatibility

CESM/CESM -Lab Containers

Environment (base)	Common OS tools & libraries: compilers, NetCDF, Vi, Emacs, Perl, Python, etc.
↓	
Jupyter Lab (optional)	Jupyter/Pangeo Notebook suite, plus any additions to support AI/ML/Visualizations
↓	
CESM Version (optional)	Specific CESM versions – eg, 2.2, 2.3, CTSM dev tag, SCAM-PBL tag, etc

Goal: A consistent CESM & analysis environment for everyone.

Containers aren't limited to CESM itself..

We are also containerizing a variety of our tools:

- CAM Topology tools
- CESM time-series generation tool
- Compression tools

On Desktops/Laptops, we typically use Docker images.
On HPC systems, we test with Singularity/Apptainer.

Unified Cloud / Container Platforms

The key point about both clouds and containers is that you have full control over the *environment* used by an application.

- Enables ensuring cloud/container compatibility.
  (Learn / test on the container, move to the cloud easily!)

Also working on standardizing *Jupyter* environments:

- EASE – "Earth Analysis Science Environment"
  Preinstalled Conda environment with Pangeo + CESM tools

Goal: Same EASE kernel on NCAR/Cloud/Container environments

Accelerators & CESM

Four key challenges for Accelerator/GPU use in CESM:

- Performance vs Efficiency
- Science Capability vs Capacity
- Code Portability
- Resources

A *complex* topic that's often misrepresented as as simple one!

Performance vs Efficiency

Distinction between *efficiency* and *performance*:

MPAS-Atmosphere (v6), 120km resolution, 58L, DP

Hardware	*Efficiency (SYPD)* (1 GPU or CPU node)
GPU (NVIDIA A100)	**17.90**
CPU (2 x 64c Zen3)	6.53

Performance vs Efficiency

Distinction between *efficiency* and *performance*:

MPAS-Atmosphere (v6), 120km resolution, 58L, DP		
Hardware	***Efficiency (SYPD)*** **(1 GPU or CPU node)**	***Performance (SYPD)*** **(Scaled out)**
GPU (NVIDIA A100)	**17.90**	26.45
CPU (2 x 64c Zen3)	6.53	**103.76**

At our workhorse resolution, MPAS as a proxy shows:

- GPUs are more *efficient* (by ~3x!)
- CPUs scale out and are more *performant* (by ~4x!)

Science Capability vs Capacity

Should NCAR focus on enabling a few uniquely large-scale runs?
Or on supporting a high volume of science at workhorse scales?

Cheyenne: 145K cores

1.27B core-hours/year

~700M CPU hours/year for CESM (~55% of cycles)

B1850 @ 120km	B1850 @ 3.75km
2198 core-hr/yr	~32^3 more work
~320,000 SY / yr (@ 120km)	**9.7 SY / yr** (@ 3.75km)

Code Portability

	NVIDIA	AMD	Intel
Accelerator	*A100*	MI200	PV (Xe)
CUDA	✓	✗	✗
OpenACC	✓	✓	✗
OpenMP	✓	✓	✓

Not every method works on every platform!
- For a *community model*, we <u>need</u> portability.

CISL has been working with an OpenACC –> OpenMP converter which would be a good path forward.

Resources (People & Systems)

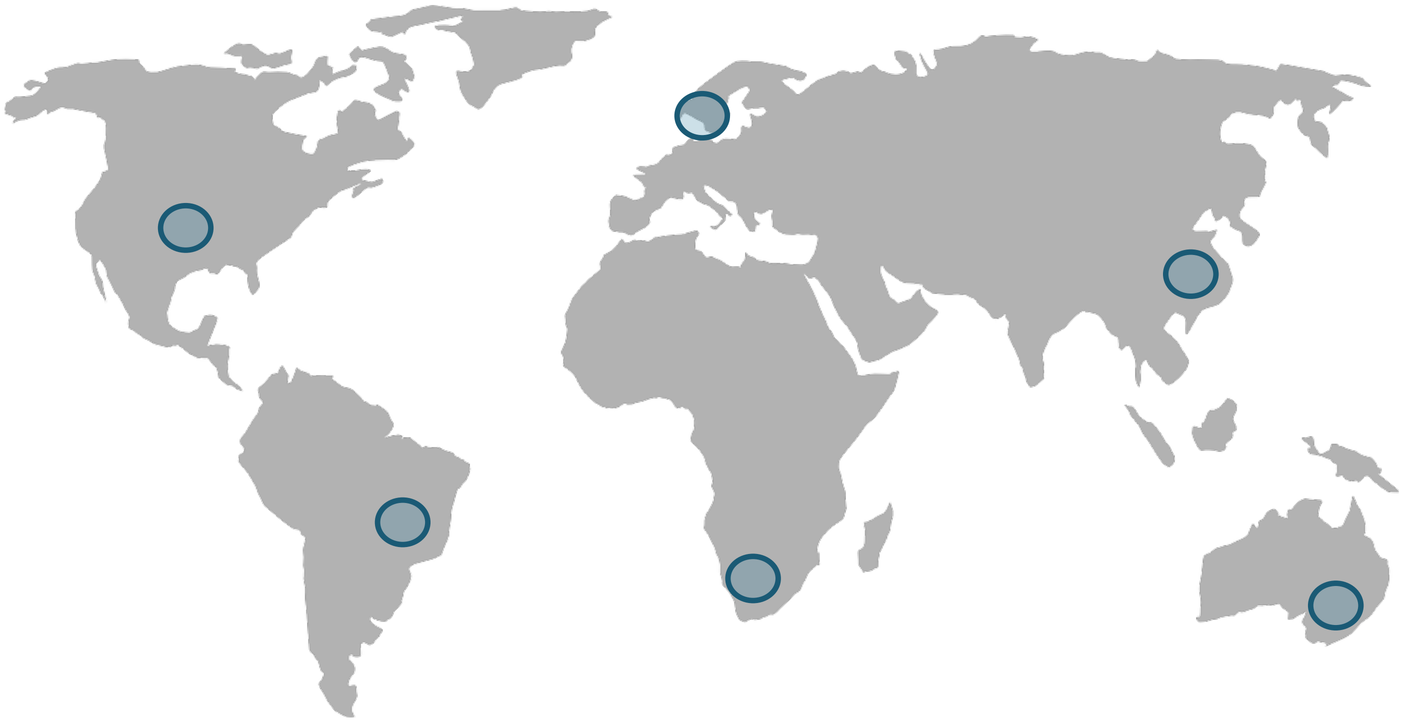
Models of similar complexity have taken >10 FTE yrs to GPUize:

- Do we focus our efforts on this, or on science features?
- *Can we get additional funding or community help to do both?*

Still *relatively* low adoption of GPUs in academic systems:

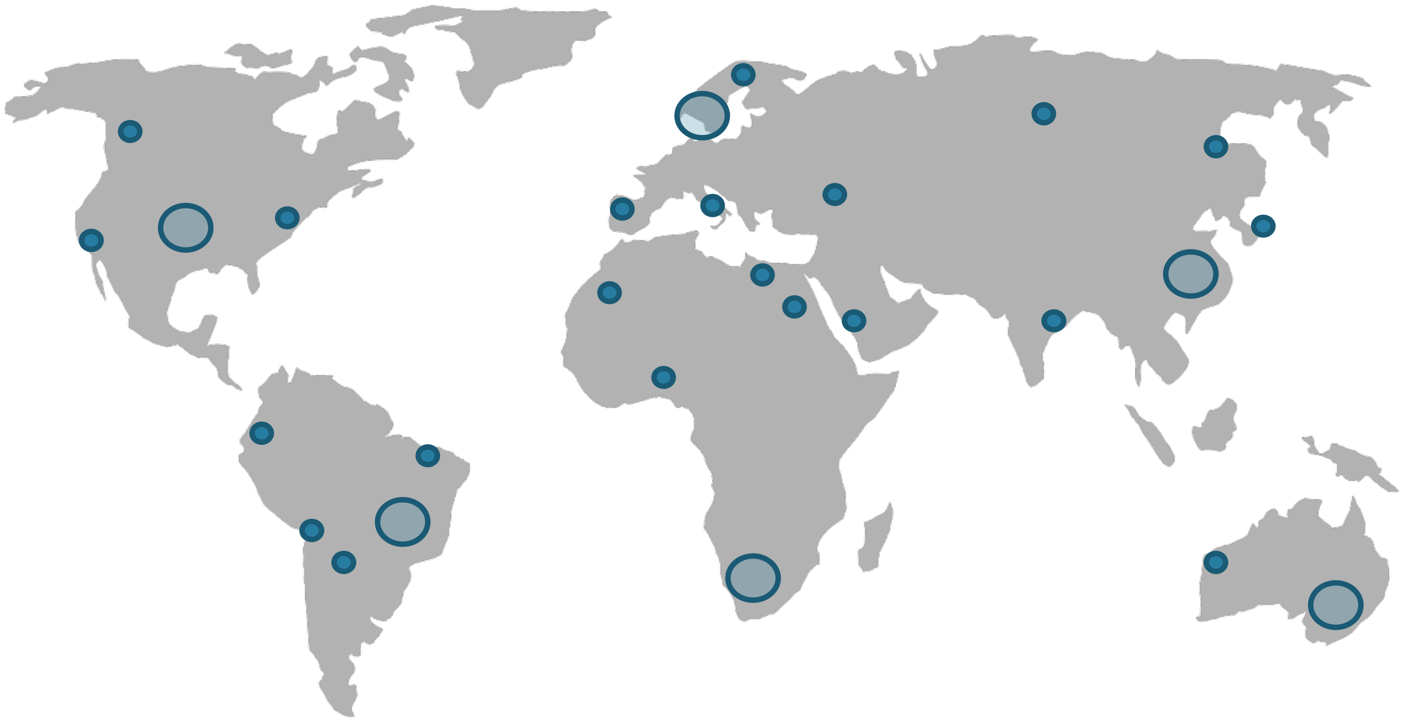
- Top500 US Academic sites
- Informal XSEDE CC Survey

Extra: Climate Data Network



Near-term challenge: Localized input data for model runs

Extra: Climate Data Network



Near-term challenge: Localized input data for model runs
Long-term challenge: Data-sharing network for online analysis

Summary

- CESM & JupyterHub usable via AWS for science & training
  - Easy to use, but not as cost-effective as on-prem systems

- Containers enable ready-to-run applications on your own hardware
  - New tools will likely really benefit the community (eg, time-series)

- Our GPU approach is a careful balance of technology, science, systems and people to do what's *best* for the CESM community.
  - *Efficiency* & *performance* are very different at ~1-degree runs!

- *Data access is a growing issue, and we're looking at some ideas here.*

Thank you!

Questions / Comments?

bdobbins@ucar.edu