

# SCREAM – Designing a next generation global atmosphere model for exascale

Aaron S. Donahue\* CESM Workshop 2024 June 11, 2024

\*and many more



This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 IM Release Number LLNL-PRES-865311







## What is SCREAM? EAMxx?

- The Simple Cloud Resolving E3SM Atmosphere Model (SCREAM) is the 3.25km configuration of <u>EAMxx</u>.
- EAMxx is the atmosphere model *code* which can be run at any resolution (eventually):
  - a complete redesign of the E3SM Atm. Model in C++/Kokkos



Resolved-scale **fluid dynamics** treated by a non-hydrostatic Spectral Element (**SE**) approach



Microphysical processes handled by Predicted Particle Properties (P3) scheme Rad ext GP



**Radiation** handled by externally-developed, GPU-ready **RRTMGP** package



\* Using coarser grid for physics parameterizations (PG2) \*



### How we got here

- Parallel tracks:
  - validation of SCREAM configuration,
  - and development of EAMxx infrastructure.
- ~4.5yrs + ~6 FTE's





\* Caldwell, PM, et al., "Convection-permitting simulations with the E3SM global atmosphere model" (2019), JAMES, doi: 0.1029/2019MS001870 \*\* Donahue, AS, et al., "To exascale and beyond -- The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales" (2024), JAMES, *under review* 



### **Portability and Performance**

- Portable Performance:
  - Compiler Directives, e.g. OpenMP, OpenACC
  - General Purpose Libraries, e.g. Kokkos, YAKL and Raja
  - Domain Specific Languages (DSL), e.g. GridTools, PSyclone and CLAW
- Transition to C++
  - Necessary to leverage tools like Kokkos
  - Attracts the next generation of software engineers, who may have never worked with Fortran
  - Better compiler support from vendors
  - A number of added benefits, such as object oriented programming.

700 Frontier Summit (SDPD) Perlmutter-GPU 400 350 Perlmutter-CPU 300 day 250 200 wallclock 175 .0 150 125 per 10090 days 80 70 Simulated 60 50 40 Model 30 Atmosphere Dycore 20 384 512 1024 2048 4096 8192 (a) Number of nodes

figure: scaling performance of EAMxx

Added Bonus: A complete rewrite allows us to cleanup legacy code that is either unused or rigid!







### **Memory Management**

- 3.25km → 25.17M columns... x 128 levels = 3.22B dofs
- Output:
  - 2D field (e.g. surface pressure) = 12.9 GB
  - 3D physics field (e.g. temperature) = 1.65 TB
  - 4x greater on dynamics grid (e.g. for restarts)

#### **Rigid Structure**

- Hard-coded persistent variables (phys\_state, phys\_tend, dyn\_state, PBUF)
- Hard-coded set of restart variables



Storage & on-node memory limits



### **Memory Management**

- 3.25km → 25.17M columns... x 128 levels = 3.22B dofs
- Output:
  - 2D field (e.g. surface pressure) = 12.9 GB
  - 3D physics field (e.g. temperature) = 1.65 TB
  - 4x greater on dynamics grid (e.g. for restarts)

#### Rigid Structure

- Hard-coded persistent variables (phys\_state, phys\_tend, dyn\_state, PBUF)
- Hard-coded set of restart variables

#### **Intelligent Memory Management**

- Field Manager:
  - <u>only allocate</u> persistent variables actually used.
  - Restart variables list <u>determined at runtime</u>.
- Atm. Memory buffer <u>shared memory space</u> for local variables.







### **The EAMxx Field Manager**

phys\_state phys\_tend dyn\_state PBUF cam\_in/out Field Manager Class and Field Objects

- All persistent variables are instances of a field object
- All fields managed by a single Field Manager
- Simplifies operations on fields, e.g.
  - Remapping to different grids
  - Adding new fields
  - Storing field metadata





# Directed Acyclic Graph (DAG)

- Reference for **how** state variables are used.
- Quickly **audit** all global variables in the simulation.
- Runtime determination of what is needed for restarts.







# Directed Acyclic Graph (DAG)

- Reference for **how** state variables are used.
- Quickly **audit** all global variables in the simulation.
- Runtime determination of what is needed for restarts.

Incredibly useful for **debugging!** 







### **Atmosphere memory buffer**

- Allocate the **minimum** memory footprint to accommodate local memory needs of all processes.
- Take advantage of pointers to memory, rather than allocated arrays.





Process 3

Memory Buffer





- 2D field (e.g. surface pressure) = 12.9 GB
- 3D physics field (e.g. temperature) = 1.65 TB
- 4x greater on dynamics grid (e.g. for restarts)







- 2D field (e.g. surface pressure) = 12.9 GB
- 3D physics field (e.g. temperature) = 1.65 TB
- 4x greater on dynamics grid (e.g. for restarts)



30°E 60°E 90°E 120°E150°E 180° 150°W120°W90°W 60°W 30°W



Produce output at a set of distinct points.





- 2D field (e.g. surface pressure) = 12.9 GB
- 3D physics field (e.g. temperature) = 1.65 TB
- 4x greater on dynamics grid (e.g. for restarts)



Produce output at a set of distinct points.

Top of Atmosphere Shortwave Cloud Forcing in SCREAMv1 January



Produce output over a sub-region of the globe.





- 2D field (e.g. surface pressure) = 12.9 GB
- 3D physics field (e.g. temperature) = 1.65 TB
- 4x greater on dynamics grid (e.g. for restarts)



Produce output at a set of distinct points.

Top of Atmosphere Shortwave Cloud Forcing in SCREAMv1 January



Produce output over a sub-region of the globe.

#### Online coarsening of output.











E35M Energy Exascale Earth System Model



### Testing

- EAMxx inherits all the standard E3SM testing coverage, + robust unit tests
- This is possible because,
  - Model elements are independent.
  - Take advantage of arbitrary class structure in C++.
- Incredibly useful for **debugging** and **verification**
- SCREAM git-repo CI integrates unit testing through the Autotester.

E3SM-Autotester commented 6 hours ago	(Member) ···
Status Flag 'Pull Request AutoTester' - Jenkins Testing: 1 or	nore Jobs FAILED
Note: Testing will normally be attempted again in approx. 2 I testing will be attempted again on next available autotester in	Irs. If a change to the PR source branch occurs, the un.
Pull Request Auto Testing has FAILED (click to expand)	
SCREAM_PullRequest_Autotester_Mappy # 5172 FAILE	D (click to see last 100 lines of console output)
SCREAM_PullRequest_Autotester_Weaver # 5520 FAIL	ED (click to see last 100 lines of console output)

477 tests passed.				
Name 🔨	Status ٨	Time	History	Summary
ad_ut	Passed	2s 140ms	Stable	Stable
array_io	Passed	3s 230ms	Stable	Stable
atm_proc	Passed	2s 680ms	Stable	Stable
atmosphere_density	Passed	1s 460ms	Stable	Stable
bfbhash	Passed	3s 730ms	Stable	Stable
catch_main_invalid_flags	Passed	3s 990ms	Stable	Stable
catch_main_tests	Passed	3s 640ms	Stable	Stable
check_subcycling	Passed	50ms	Stable	Stable
check_subcycling_tend_check	Passed	320ms	Stable	Stable
check_surf_mom_flux_slices_np1	Passed	370ms	Stable	Stable
check_surf_mom_flux_slices_np2	Passed	380ms	Stable	Stable
check_surf_mom_flux_slices_np3	Passed	340ms	Stable	Stable
check_surf_mom_flux_slices_np4	Passed	370ms	Stable	Stable
check_U_V_slices_fail_diff	Passed	310ms	Stable	Stable
check_U_V_slices_fail_layout	Passed	390ms	Stable	Stable
check_U_V_slices_fail_missing	Passed	350ms	Stable	Stable
check_U_V_slices_np1	Passed	470ms	Stable	Stable
check_U_V_slices_np2	Passed	360ms	Stable	Stable
check_U_V_slices_np3	Passed	370ms	Stable	Stable
check_U_V_slices_np4	Passed	340ms	Stable	Stable
cld_fraction_standalone	Passed	1s 570ms	Stable	Stable
coarsening_remapper_np1	Passed	3s 400ms	Stable	Stable
coarsening_remapper_np2	Passed	2s 620ms	Stable	Stable
coarsening_remapper_np3	Passed	1s 180ms	Stable	Stable
coarsening_remapper_np4	Passed	820ms	Stable	Stable







### **Runtime Options:**

effaw oro = 0.375



- Difficult to read/parse
- Not organized
- Cumbersome to add new options
- Difficult to audit, vulnerable to user error
  - Conflicts only detected if developer adds their own checks.
  - Multiple entries can lead to unexpected behavior.



## **Runtime Options: YAML**

- EAMxx use YAML to handle all runtime options.
- Human readable options organized into relevant sections.
- Strict enforcement of each entry.
- Support for "info" metadata for user support.
- Setting and adding new options is simple. Each process has their own parameter list.
- Has similar support to CIME XML functions
  - ./atmchange
  - ./atmquery



physics: Type: Group atm\_procs\_list: !strings - mac\_aero\_mic rrtmgp mlcorrection compute\_tendencies: !strings – T\_mid - qv horiz\_winds enable\_postcondition\_checks: true enable\_precondition\_checks: true internal\_diagnostics\_level: 0 mac\_aero\_mic: Type: Group atm\_procs\_list: !strings - tms shoc cldFraction spa - p3 cldFraction: compute\_tendencies: !strings [] enable\_postcondition\_checks: true enable\_precondition\_checks: true internal\_diagnostics\_level: 0 number\_of\_subcycles: 1 repair\_log\_level: trace compute\_tendencies: !strings [] enable\_postcondition\_checks: true enable\_precondition\_checks: true internal\_diagnostics\_level: 0 number\_of\_subcycles: 12 p3: compute\_tendencies: !strings [] do\_predict\_nc: true do\_prescribed\_ccn: true enable\_column\_conservation\_checks: false enable\_postcondition\_checks: true enable\_precondition\_checks: true internal\_diagnostics\_level: 0 max\_total\_ni: 740000.0 number\_of\_subcycles: 1 p3\_a\_imm: 0.65 n3 autoconversion prefector: 1350 0

23

	Runtime	<b>Options: YAML</b>
--	---------	----------------------

> ./atmquery mac\_aero\_mic::number\_of\_subcycles namelist\_defaults::atmosphere\_processes::physics::mac\_aero\_mic::number\_of\_subcycles: 6

ngs

n3 autoconversion prefactor: 1350

- > ./atmchange mac\_aero\_mic::number\_of\_subcycles=12 Regenerating .../namelist\_scream.xml. Manual edits will be lost.
- > ./atmquery mac\_aero\_mic::number\_of\_subcycles
- namelist\_defaults::atmosphere\_processes::physics::mac\_aero\_mic::number\_of\_subcycles: 12 **EAMx**



			- horiz_winds
• Huma	<pre>mac_aero_mic: Type: Group atm_procs_list: !strings</pre>	ections.	enable_postcondition_checks: true enable_precondition_checks: true internal_diagnostics_level: 0 mac_aero_mic: Type: Group
<ul> <li>Strict e</li> <li>Suppo</li> <li>Setting their c</li> </ul>	<pre>- tms - shoc - cldFraction - spa - p3 cldFraction:    compute_tendencies: !strings []    enable_postcondition_checks: true    enable_precondition_checks: true    internal_diagnostics_level: 0</pre>	cess has	<pre>atm_procs_list: !strings - tms - shoc - cldFraction - spa - p3 cldFraction: compute_tendencies: !strings [] enable_postcondition_checks: true enable_precondition_checks: true internal_diagnostics_level: 0 number_of_subcycles: 1 repair_log_level: trace compute_tendencies: !strings [] enable_postcondition_checks: true</pre>
• Has sin • ./ • ./	<pre>number_of_subcycles: 1     repair_log_level: trace     compute_tendencies: !strings []     enable_postcondition_checks: true     enable_precondition_checks: true     internal_diagnostics_level: 0     number_of_subcycles: 12 ergy Exascale </pre>		<pre>enable_precondition_checks: true internal_diagnostics_level: 0 number_of_subcycles: 12 p3: compute_tendencies: Istrings [] do_predict_nc:-true do_prescribed_ccn: true enable_column_conservation_checks: false enable_postcondition_checks: true enable_precondition_checks: true internal_diagnostics_level: 0 max_total_ni: 740000.0</pre>
Ear	rth System Model	24	number_of_subcycles: 1 p3_a_imm: 0.65

### **Conclusions:**



 EAMxx is a complete rewrite of the E3SM global atmosphere model infrastructure in C++/Kokkos, eploying modern software best practices.



• SCREAM is the 3.25 km configuration of EAMxx with processes targeting high resolution.



- EAMxx enforces strict memory management as a **necessity** given the shear size of the problem being solved.
- EAMxx infrastructure was designed with testing in mind. Unit testing:
  - Makes debugging substantially easier,
  - Gives more confidence that new features won't impact or break existing features.



- The adoption of YAML makes runtime options:
  - More protected against user error,
  - Human readable,
  - More organized





### Questions? <u>donahue5@llnl.gov</u>

#### LW-Flux Up @ TOM



### Output Control: user\_nl\_eam

```
cld_macmic_num_steps = 6
mfilt = 10000,1,12
cosp_lite = .true.
use_gw_front = .true.
nhtfrq = 1,0,-2
fincl1 = 'Q','CLDLIQ','CLDICE'
fincl2 = 'PS','TMQ','T500'
scm_iop_srf_prop = .true.
avgflag_pertape = 'A','I','X'
iop_pudge_tg_math{tau}
```





### **Output Control: YAML**

```
cld_macmic_num_steps = 6
mfilt = 10000,1,12
cosp_lite = .true.
use_gw_front = .true.
nhtfrq = 1,0,-2
fincl1 = 'Q','CLDLIQ','CLDICE'
fincl2 = 'PS','TMQ','T500'
scm_iop_srf_prop = .true.
avgflag_pertape = 'A','I','X'
iop_pudge_tg = _folce
```







### **Output Control: YAML**

```
cld_macmic_num_steps = 6
mfilt = 10000,1,12
cosp_lite = .true.
use_gw_front = .true.
nhtfrq = 1,0,-2
fincl1 = 'Q','CLDLIQ','CLDICE'
fincl2 = 'PS','TMQ','T500'
scm_iop_srf_prop = .true.
avgflag_pertape = 'A','I','X'
iop_pudge_tg = _felse
```

Averaging Type: Instant Fields: Physics PG2: Field Names: - qv - qc - qi Max Snapshots Per File: 1 filename\_prefix: output.scream.monthly.NAME\_1 output\_control: Frequency: 1 frequency\_units: nmonths

```
Averaging Type: Max

Fields:

Physics PG2:

Field Names:

- ps

- VapWaterPath

- LiqWaterPath

Max Snapshots Per File: 12

filename_prefix: output.scream.monthly.NAME_2

output_control:

Frequency: 2

Frequency:
```

### What is a Field?

- Stores everything the atmosphere model needs to know about a global variable.
- As an object allows SCREAM to define universal operations for all fields,
  - E.g. remapping, arithmetic operations, cloning, property checking, timestep management ...
    - > Name: T\_mid
    - > Units: K
    - > Layout: (columns, levels)
    - > Grid: Physics PG2
    - > Datatype: Real
    - > Timestamp: TS{2024,11,4,8,30,0}
    - > Data Pointer: 6efgkh38sahdlgis0372
    - > Providers: HOMME, SHOC, P3, RRTMGP, Surface-Coupling
    - > Customers: HOMME, SHOC, P3, RRTMGP, Surface-Coupling
    - > Extra Metadata: [Mask, Foo, Bar, ...]

\* Artist depiction





### What is an Atmosphere Process?

- An interface to the dynamics or an atmosphere parameterization.
- EAMxx supports universal operations on atmosphere processes:
  - Unit testing, subcycling, backing out tendencies, log msgs, performance timing, ...
- Has a defined initialization, run and finalization.

#### > Name: SHOC

- > Type: Physics
- > Grid: Physics PG2
- > Timestamp: TS{2024,11,4,8,30,0}
- > Inputs: omega, surf\_sens\_flux, surf\_mom\_flux, ..., T\_mid, qv
- > Outputs: surf\_evap, T\_mid, qv, ... , tke, pbl\_height
- > init\_impl: Do 'xyz' to initialize SHOC
- > run\_impl: Do 'ijk,' then call shoc\_main, do 'qrs'
- > finalize\_impl: Nothing to do

\* Artist depiction





### **Performance Portability (C++/Kokkos)**



climate.atlassian.net/wiki/spaces/ECM/pages/4129325057/2 024-02-15+All-Hands+Presentation+Meeting+Notes



Packs



