

# CERES Data Management

Working Group Report  
CERES Science Team Meeting

Katie Deiwakh  
Walter Miller  
CERES DMT

OCT. 17, 2023

## Agenda

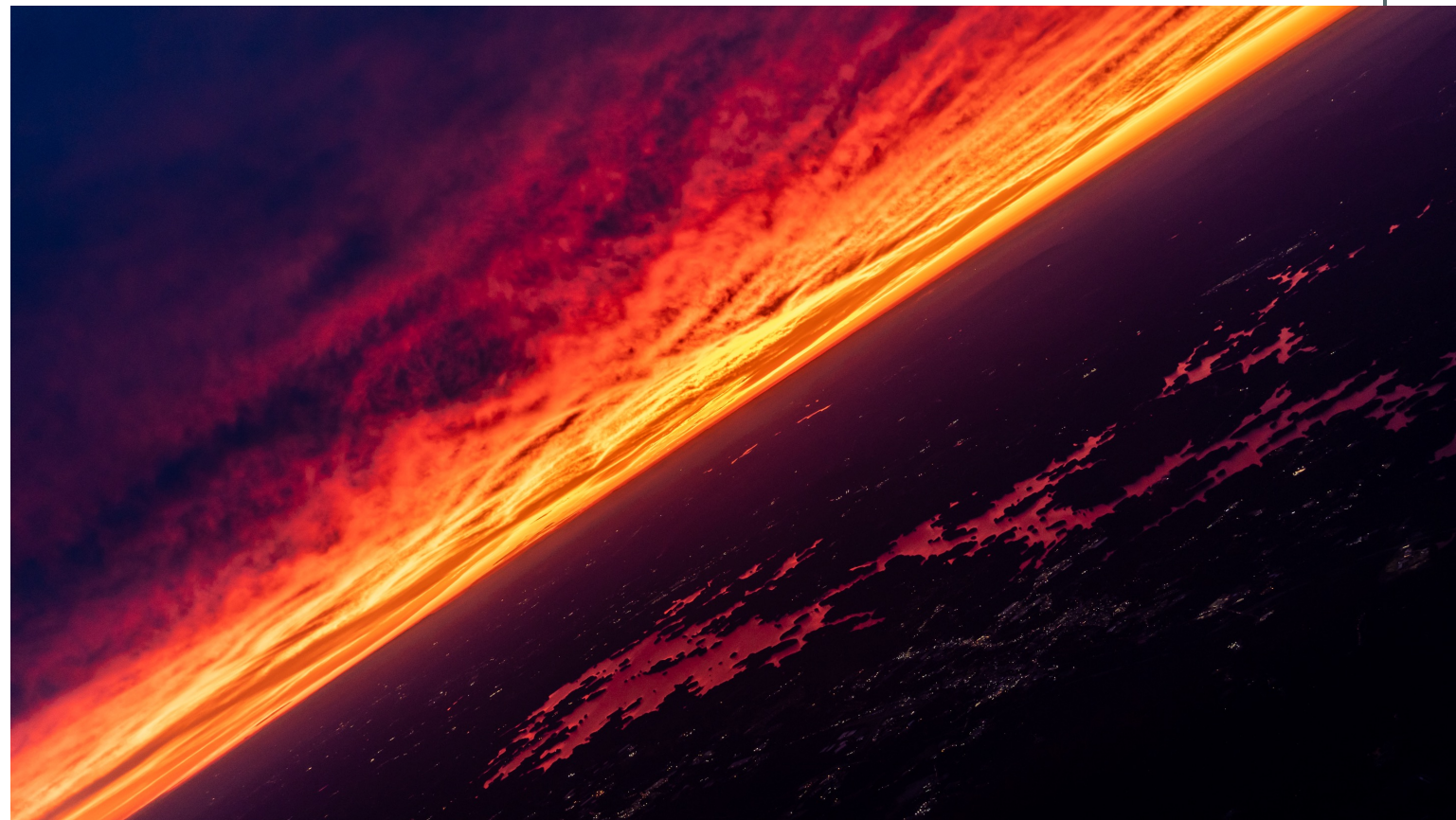
Background

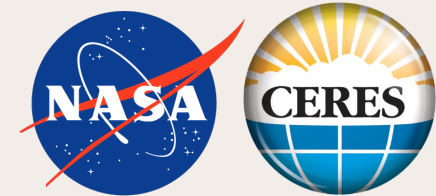
Team Highlights

Improvements:

- Edition 5
- Code Re-architecture
- Systems

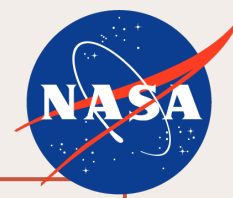
Conclusion





# Background

Instrument

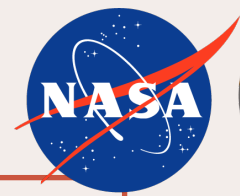


Denise Cooper  
Thomas Grepiotis  
Dianne Snyder  
Hunter Winecoff

ERBE-like

Dale Walikainen  
Jeremie Lande

Clouds



Sunny Sun-Mack

Walter Miller

Ricky Brown

Rita Smith

Elizabeth Heckert

Yan Chen

Steven Kohler

Inversion



Victor Sothcott  
Igor Antropov

Surface and Atmospheric Radiation Budget (SARB)

Thomas Caldwell

Temporal Interpolation and Spatial Averaging



Beau Branch  
Edward Kizer  
Cathy Nguyen  
Joshua Wilkins

FLASHFlux

Jay Garg  
PC Sawaengphokhai  
Hunter Winecoff

Configuration Management (CM) & Infrastructure



Tammy Ayers  
Willinda Evans  
Nelson Hillyer  
Dennis Keyes

Production Request (PR) Tool

Carla Grune  
Elizabeth Heckert



## Visualization and Ordering Tool



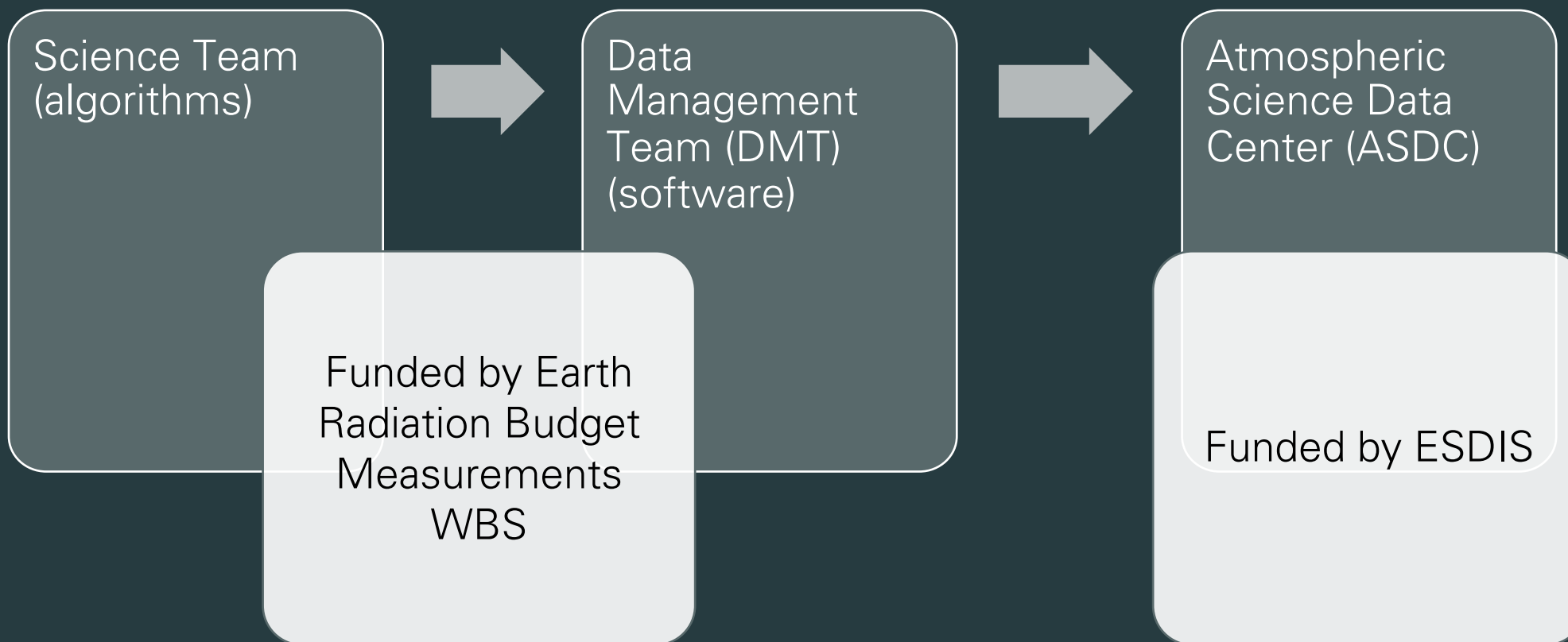
Churngwei Chu  
Pamela Mlynczak  
Babak Samani

## Atmospheric Science Data Center (ASDC) Support

Karen Brown  
Sharon Dukes-Allen  
Christopher Harris

## Background

CERES/RBSP – not a true Science-Investigator Processing System, or “SIPS”





# Team Highlights

add new PR PGE Template for FLASHFlux TISA Grid PGE, FLASH13-9.2P3

to the CATALYST server to adjust job execution priority and limits for the number of high I/O jobs running at any one time.

map 401 to MERRADStream for CER12.1P3

add new PR PGE Template for FLASHFlux TISA Average PGE, FLASH13-5.1P1

of FLASHFlux TISA Average

deactivate FLASH13-3.2P2 PGE, PRs and CATALYST status flags

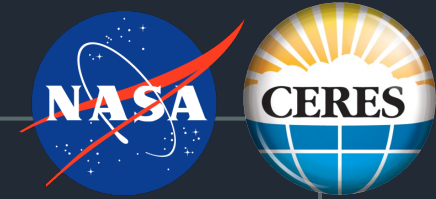
Perf\_Lib module being added in SCCR 1951 to fix the issue with mails not being able to send out emails from the p

to library to handle sending email messages on AMI or AMI-p interactive or grid engine execution hosts.

Labels	Description
1865	Initial delivery of FLASHFlux Inversion subsystem (13-3.0P6, 13-3.1P6)
1865	Create CERES Archiver Profiles for FLASH Inversion PGEs.
1865	Create CATALYST PGE modules for FLASH Inversion PGEs.
1939	Initial delivery of FLASHFlux TISA Grid Subsystem (13-9)

10/17/2023

QC specific CERES Archiver profiles for Baseline1-QC processing for PGEs CER1.1P10, CER2.2P1, and CER2.2P2.



## Software Deliveries (61)

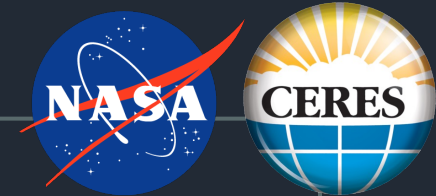
### Meteorological Ozone and Aerosol (MOA):

- First Edition 5 product generation executive (PGE) code delivery
- Associated software deliveries for production and archiving

### SARB:

New PGE to support new “ice age” value input source (Land Information System)

**TISA Gridding:** Incorporating NOAA-20 VIIRS data into CldTypHist & FluxByCldTyp



## Software Deliveries (61)

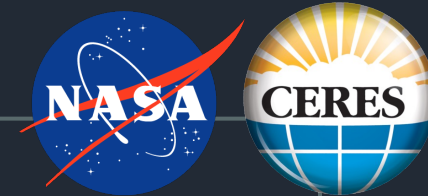
### FLASHFlux:

- Update to prevent unnecessary filtering of good NOAA-20 values in TISA
- New PGE to produce MOA with GMAO GEOS-IT instead of FP-IT
- Quarterly spectral correction coefficient files

### Instrument and ERBE-like:

Several rounds of gains and spectral response function updates for instruments

TISA Averaging: Interpolation adjustment to support surface flux processing



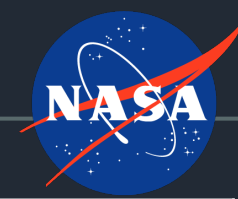
## Software Deliveries (61)

### CATALYST, PR Tool, and CERESlib:

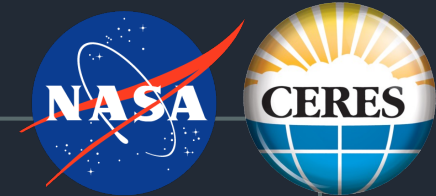
- Many deliveries increasing production resilience to missing or additional data
- Edition 5 support

Various SORCE TSI, MATCH aerosol, and gridded GEO ancillary product updates

# Product Availability



Product	Platform				Availability Thru
	Terra	Aqua	S-NPP	NOAA-20	
BDS	x	x	x	x	July 2023
SSF	x	x	x	x	July 2023
CRS	x	x			2018-2022
SSF1deg-Hour	x	x	x	x	July 2023
SSF1deg-Day/-Month	x	x	x	x	July 2023
SYN1deg-1Hour/-MHour	Terra+Aqua			Terra+NOAA-20	June 2023
SYN1deg-Day/-Month	Terra+Aqua			Terra+NOAA-20	June 2023
CldTypHist	Terra+Aqua			Terra+NOAA-20	July 2023
FluxByCldTyp-Day/-Month	Terra+Aqua			x	July 2023
EBAF	Terra+Aqua			x	June 2023
EBAF ToA	Terra+Aqua			x	July 2023

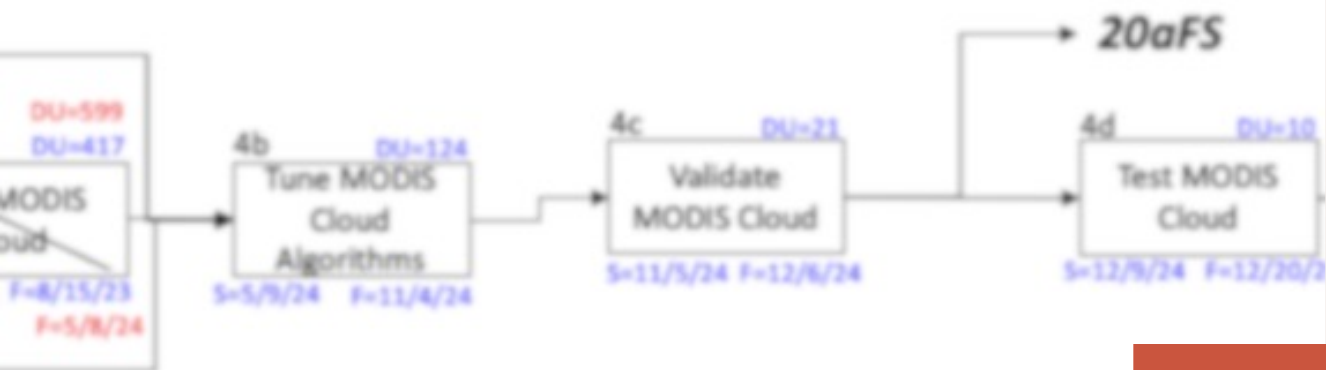


## Upcoming Production

### Edition 4B SYN1deg reprocessing:

- Finishing GEO reprocessing
  - Two-channels satellites: improved neural net; NetCDF4
  - MET8-10 rerun with updated code
- Twilight striping correction
- Updated interpolation scheme
- Reprocessing begun – projected finish: Spring 2024



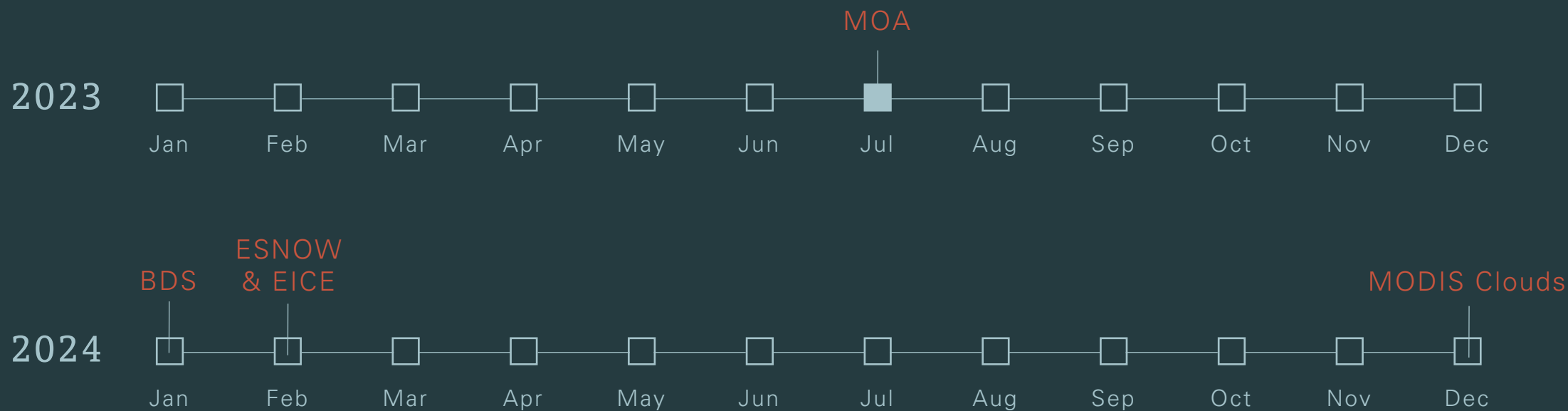


# Improvements





## Edition 5 Progress





## Edition 5 Progress

- Edition 5 MOA delivered
- “AlphaN” tests commencing (Alpha1, first)
  - Internal-only
  - Cloud properties and fluxes through gridding
- Updating schedule as contingencies learned
- Next delivery: BDS or ESNOW & EICE

## TISA Code Re-Architecture

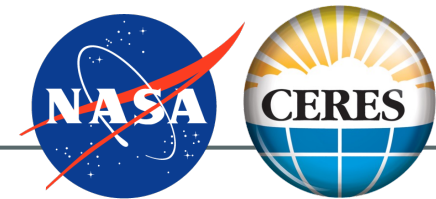


- SW, LW and WN
- Modular code, JSON configuration files
- Ample, generalized subroutine usage
- Sharing of/cross-training on code



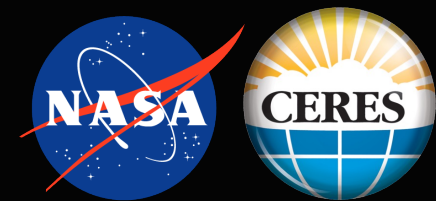
## TISA Code Re-Architecture

- Generic netCDF data product reader
  - “Packing list” JSON file defines expected contents
  - Increases modularity of codebase
  - Easier to debug
- Surface type look-up table finalized
  - Read 10-min. static IGBP and dynamic snow/sea-ice maps:  
Resample starting at 1/6<sup>th</sup> degree (SW & LW)
  - 20 IGBP types → 6 SW and 7 LW
  - IGBP type assignment by weighted partitioning
- SW narrowband to broadband module updated



## TISA Code Re-Architecture

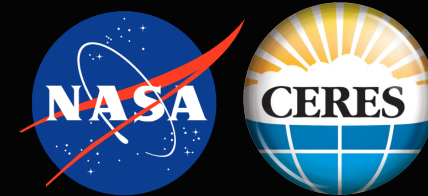
- Interpolation: new interpolation scheme refactoring across editions
- Gridding:
  - C++ Re-architecture:
    - JSON configured metadata support
    - Implemented cosine SZA weighting, non-binned multidimensional variable support, offset capability
  - Prototype CRS1deg product in development
    - Piggybacks on resurrected CRS product
    - First exercise of new, C++ codebase



## Clouds Code Re-Architecture

- Concluded cloud mask “magic number” removal
- Commencing in-depth testing

```
f(!isPolar_){ //legacy var newPolFlg_inMask
if( intlGeosphereBiosphereProgramme_ == (static_cast<int>(InternationalGeosphereVegetationType::WATER_BODIES))
  && snowMap_ <= SNOW_MAP_INCH_MIN
  && iceMap_ == ICE_MAP_PERCENT_MIN
  && sunGlint_ < LUT_CERESThresholds_->getValue("sunglint_prob")
){
ratio213_065_cor_ = nonCommonConfigValues_->ratio213_065_corNonPolarNonSunGlintModifier1 //legacy val 0.2
                  * cos(viewZenithAngle_ * PI / nonCommonConfigValues_->ratio213_065_corNonPolarNonSunGlintModifier2) //legacy val 180.)
                  + nonCommonConfigValues_->ratio213_065_corNonPolarNonSunGlintModifier3; //legacy val 0.58678;
}
else if( intlGeosphereBiosphereProgramme_ == (static_cast<int>(InternationalGeosphereVegetationType::WATER_BODIES))
  && snowMap_ <= SNOW_MAP_INCH_MIN
  && iceMap_ == ICE_MAP_PERCENT_MIN
  && sunGlint_ >= LUT_CERESThresholds_->getValue("sunglint_prob")
){
ratio213_065_cor_ = nonCommonConfigValues_->ratio213_065_corNonPolarSunGlintModifier1 //legacy val 0.18
                  * cos(viewZenithAngle_ * PI / nonCommonConfigValues_->ratio213_065_corNonPolarSunGlintModifier2) //legacy val 180.)
                  + nonCommonConfigValues_->ratio213_065_corNonPolarSunGlintModifier3; //legacy val 0.625;
}
else if( intlGeosphereBiosphereProgramme_ != (static_cast<int>(InternationalGeosphereVegetationType::SNOW_AND_ICE_PERMANENT))
  && snowMap_ <= SNOW_MAP_INCH_MIN
  && iceMap_ == ICE_MAP_PERCENT_MIN
  && (intlGeosphereBiosphereProgramme_ == (static_cast<int>(InternationalGeosphereVegetationType::SHRUBLANDS_OPEN))
    || intlGeosphereBiosphereProgramme_ == (static_cast<int>(InternationalGeosphereVegetationType::BARE_SOIL_AND_ROCKS))
  )
){
ratio213_065_cor_ = nonCommonConfigValues_->ratio213_065_corNonPolarParm7; //legacy val 0.80
}
else if( intlGeosphereBiosphereProgramme_ != (static_cast<int>(InternationalGeosphereVegetationType::SNOW_AND_ICE_PERMANENT))
  && snowMap_ <= SNOW_MAP_INCH_MIN
  && iceMap_ == ICE_MAP_PERCENT_MIN
){
```



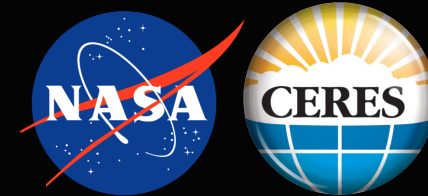
## Clouds Code Re-Architecture

- Concluded cloud mask “magic number” removal
- Commencing in-depth testing

```
ratio213_065_cor_ = nonCommonConfigValues_->ratio213_065_corNonPolarSunGlintModifier1 //legacy val 0.18
                    * cos(viewZenithAngle_ * PI / nonCommonConfigValues_->ratio213_065_corNonPolarSunGlintModifier2) //legacy val 180.)
                    + nonCommonConfigValues_->ratio213_065_corNonPolarSunGlintModifier3; //legacy val 0.625;
```

```
else if( intlGeosphereBiosphereProgramme_ != (static_cast<int>(InternationalGeosphereVegetationType::SNOW_AND_ICE_PERMANENT))
        && snowMap_ <= SNOW_MAP_INCH_MIN
        && iceMap_ == ICE_MAP_PERCENT_MIN
        && (intlGeosphereBiosphereProgramme_ == (static_cast<int>(InternationalGeosphereVegetationType::SHRUBLANDS_OPEN))
            || intlGeosphereBiosphereProgramme_ == (static_cast<int>(InternationalGeosphereVegetationType::BARE_SOIL_AND_ROCKS)))
        )
    ){
    ratio213_065_cor_ = nonCommonConfigValues_->ratio213_065_corNonPolarParm7; //legacy val 0.80
    }
else if( intlGeosphereBiosphereProgramme_ != (static_cast<int>(InternationalGeosphereVegetationType::SNOW_AND_ICE_PERMANENT))
        && snowMap_ <= SNOW_MAP_INCH_MIN
        && iceMap_ == ICE_MAP_PERCENT_MIN
        )
    )
```

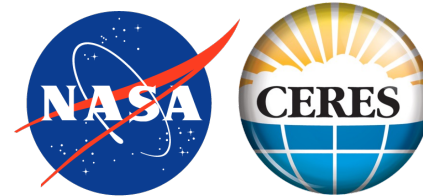




## Clouds Code Re-Architecture

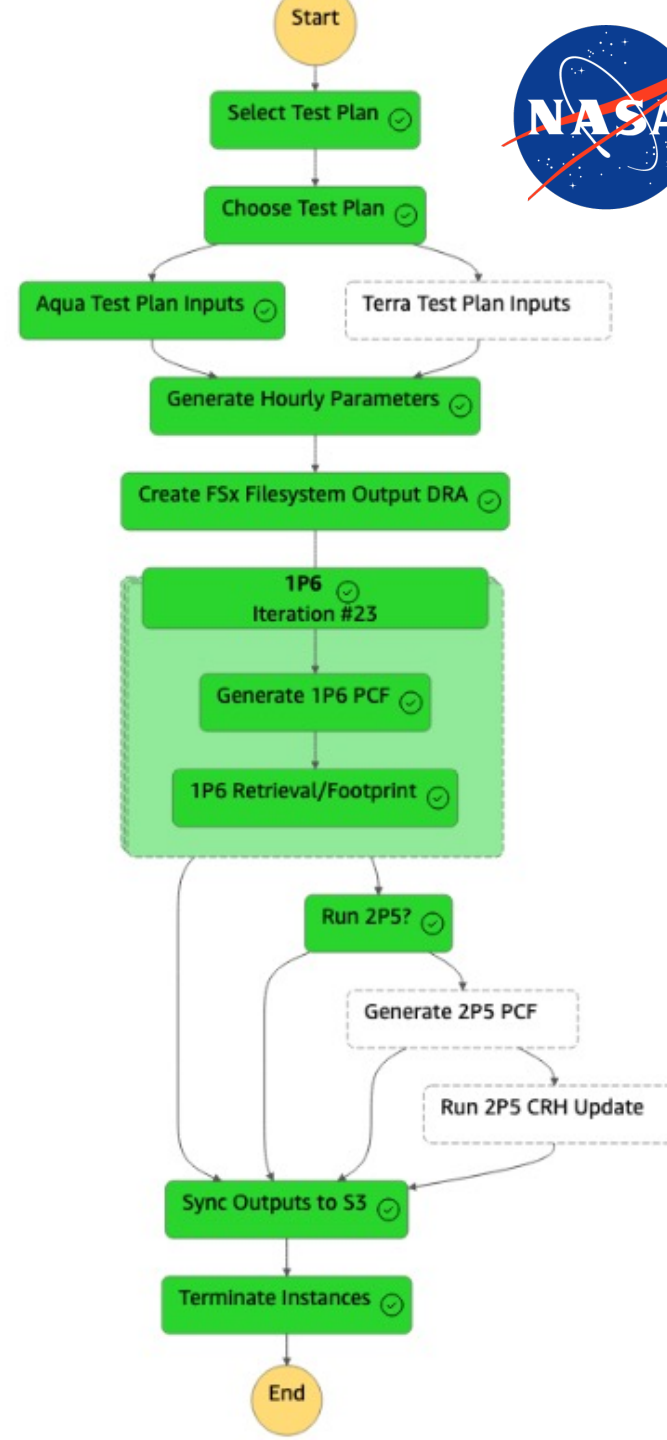
- Concluded cloud mask “magic number” removal
- Commencing in-depth testing

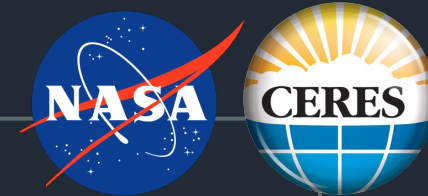
```
>ratio213_065_corNonPolarSunGlintModifier1 //legacy val 0.18  
* PI / nonCommonConfigValues_>ratio213_065_corNonPolarSunGlintMo  
_>ratio213_065_corNonPolarSunGlintModifier3; //legacy val 0.625;
```



# Clouds Code Re-Architecture

- AWS run prototyping:
  - ECS container orchestration
  - Tested Fargate and EC2 for scaling
    - EC2 more flexible
    - Enables FSX (filesystem) usage
  - Largely as performant as on premises runtime
  - Most significant gains potentially algorithm, software design
- Continuing to improve AWS process





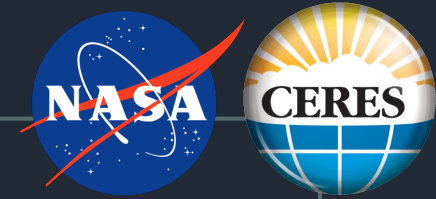
## Libera Data Management Working Group

- LASP-RBSP meeting every other week
- LASP-ASDC finalizing interface control document (ICD)
- Continuing open communication



## Systems

- Summer 2023:  
Replaced production storage hardware
- Currently auditing ASDC data product holdings for:
  - clean-out
  - migration to new hardware



## ASDC & SCF Support

- Continued support for processing and storage infrastructure
- Collection-level metadata issue still resolving
  - Core CERES and FLASHFlux products ~100% complete
  - GEO Clouds
    - 5 active GEOs complete
    - Recently used and two-channel GEOs: ~65%
- OpenShift4 transition complete



## Conclusion

- Continuing to develop Edition 5 algorithms and code
- Spring 2024 transition to Edition 4B for SYN1deg products
- Code re-architecture ongoing
- Working with ASDC to upkeep infrastructure & metadata