CERES Data Management Team Working Group Report

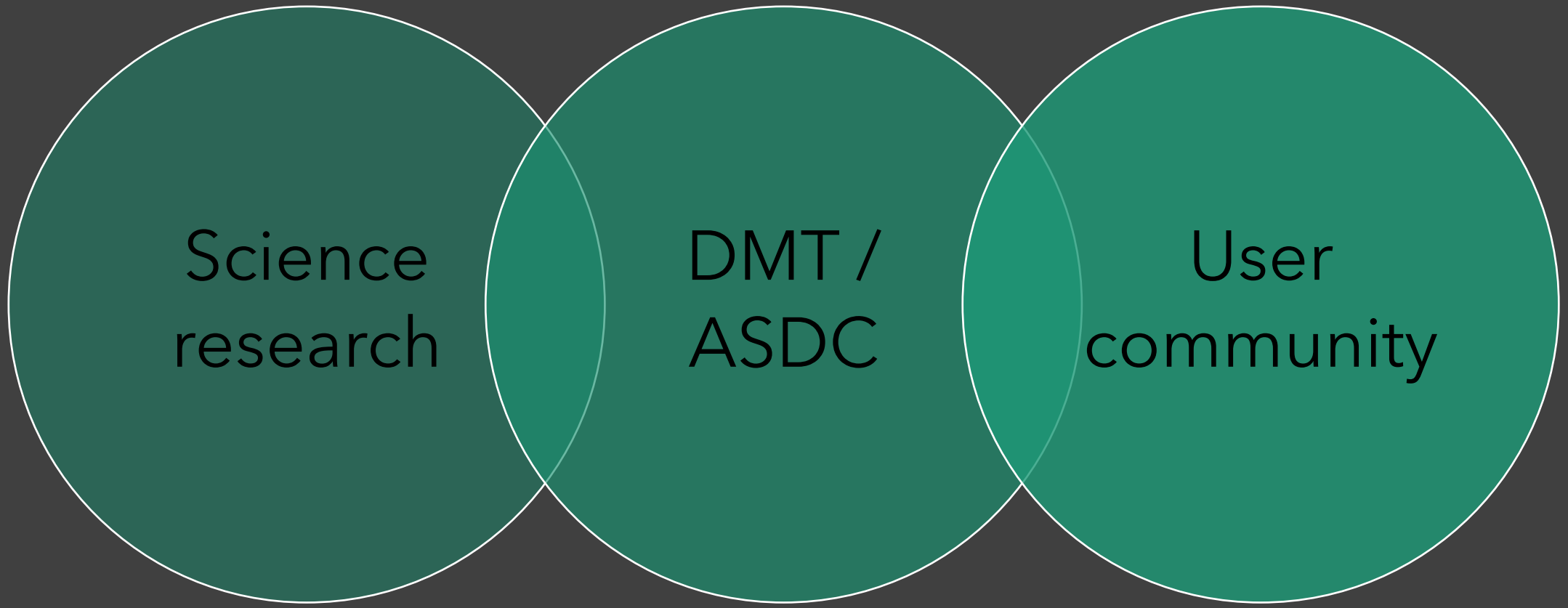
Katie Dejwakh
April 28, 2020

Outline

- DMT Overview
- Product Availability
- DMT Activities
- Code Re-architecture
- Systems & CM

DMT Overview

Data Management Team (DMT)



Data Management Team (DMT)

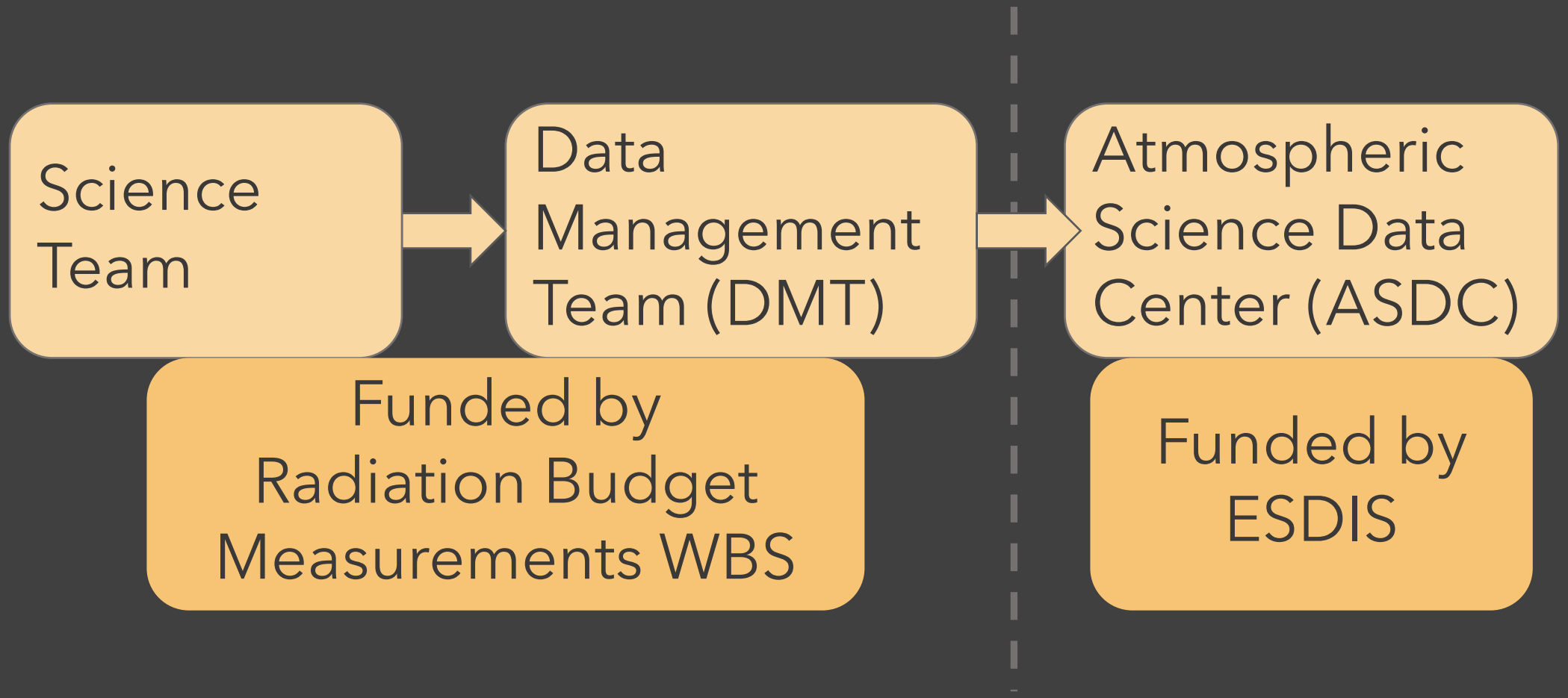
Instrument	ERBE-like	Clouds
Inversion	SARB	TISA
	Systems & CM	

Data Management Team (DMT)

Responsibilities:

- Algorithm implementation
- Software:
  - Maintenance
  - Configuration management
  - Processing
- Output validation

DMT-ASDC Interface



Product Availability

Edition 4 Terra & Aqua

Product	Platform	Processed Thru	Publicly Available?
BDS	Terra, Aqua	Dec. '19	Yes
SSF		Dec. '19	Yes
SSF1deg-Hour		Nov. '19	Yes
SSF1deg-Day/-Month		Nov. '19	Yes
SYN1deg-Hour/3Hour/MHour	Terra+Aqua	Nov. '19	Yes
SYN1deg-Day/-Month		Nov. '19	Yes

Edition 4 Terra & Aqua

Product	Platform	Processed Thru	Publicly Available?
CldTypHist	Terra+Aqua	Nov. '19	Yes
FluxByCldTyp	Terra+Aqua	Dec. '19	May '20
EBAF	Terra+Aqua	Nov. '19	Yes
EBAF ToA	Terra+Aqua	Dec. '19	Yes

Edition 1 S-NPP

Product	Platform	Processed Thru	Publicly Available?
BDS	S-NPP	March '20	Yes
SSF		Feb. '20	Yes
SSF1deg-Hour		Sep. '19*	Yes
SSF1deg-Day/-Month		Sep. '19*	Yes
SYN1deg-Hour/3Hour/MHour	Terra+S-NPP	Nov. '17*	Yes
SYN1deg-Day/-Month		Nov. '17*	Yes

* SSF1deg is paused due to RAPS mode

DMT Activities

New Product – FluxByCldTyp

- 10 years in the making!
- Gridded, L3 product
- 42 cloud types indexed by:
  - Optical depth (6)
  - Cloud pressure (7)
- Available within the next month

Forthcoming Data Product Editions

Edition 2 S-NPP

- Processing thru: Instrument & ERBE-like
- Available early Fall 2020
- Need VIIRS data
- No L3 products from 9/2019 onward, while S-NPP in biaxial mode

Edition 1 NOAA-20

- Processing thru: Instrument, SSF, and Level 3
- Available early Summer 2020

Notable Deliveries

Ed 1 NOAA-20/Ed 2 S-NPP work

- Sampling strategy updates
- Spectral correction coefficients
- SSF, SSF1deg implementation

FluxByCldTyp initial PGE

Libraries

- CERESLib netCDF4/HDF5 support
- CERESLib metadata update support
- PerlLib scoping changes

Migrated CERES Websites

- Multiple platforms, predominantly:
  - NASA LaRC OCIO's WordPress environment
  - ASDC's OpenShift Container Platform (OCP)

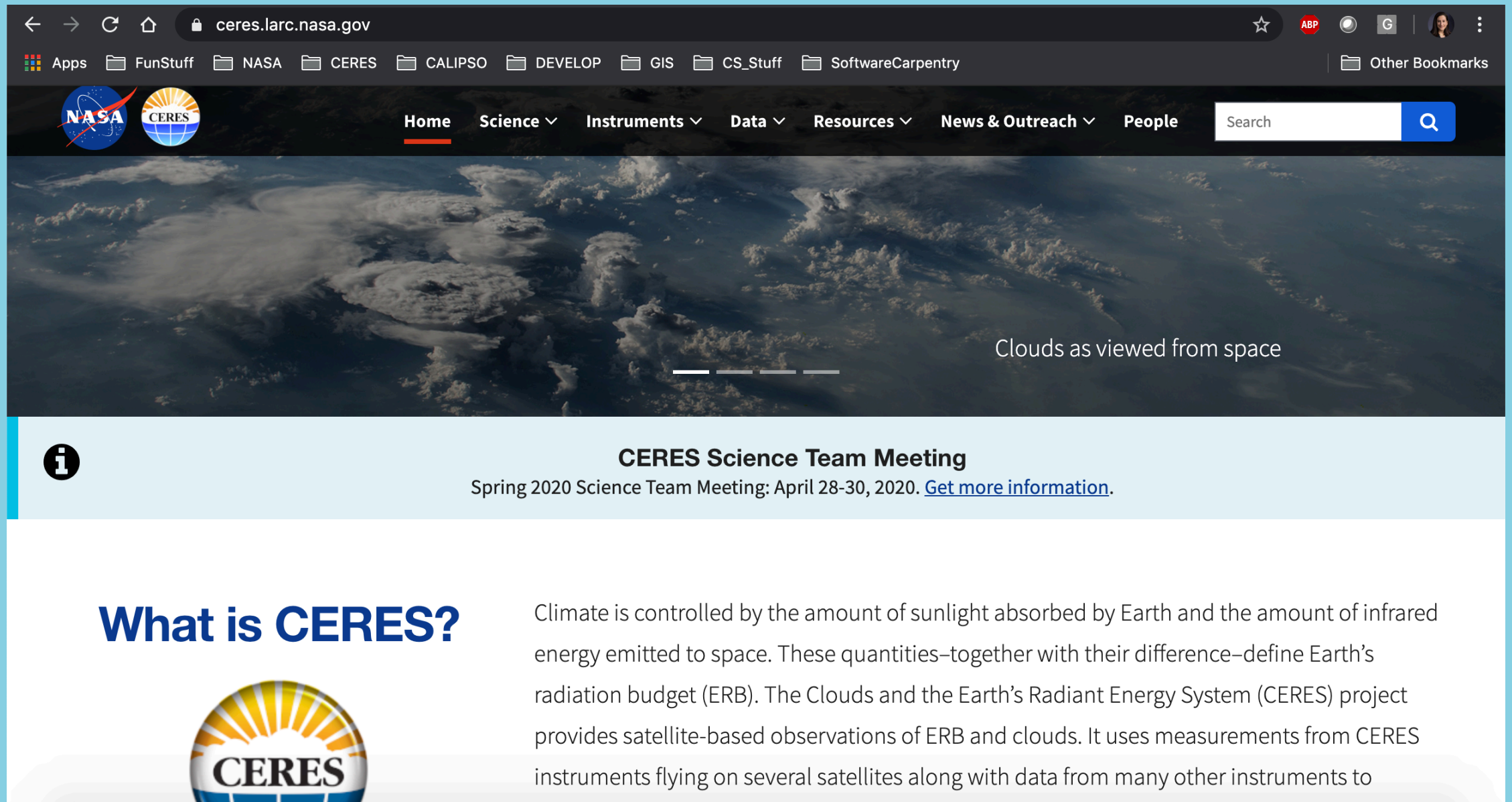
WordPress

- GEWEX SRB/RFA

OCP

- PR Tool
- CM website
- CERES-WG website
- CERES website (both)

New CERES Website



New CERES Website

- Emma Brand – part of Pathways work
- Mobile-friendly
- Runs in OCP
- Easier to promote changes
- ceres.larc.nasa.gov – check it out!

Running CERES PGEs in OpenShift

- Emma Brand & Nelson Hillyer
- 75% CERES PGEs running in OCP thus far
- Next up:
  - "Vertical slicing" runs
  - CATALYST integration
- Flexibility to harness cloud computing power

Code Improvement Activities

CERES Code Re-architecture

Readability

- Eliminate:
  - “Dead” code
  - Unreachable code
- Cap length:
  - Files
  - Functions
- No “magic” numbers

Maintainability

- De-duplicate:
  - Functions
  - Functionality
- Generalize

CERES Code Re-architecture

Modularity

- Single functionality per unit
- Interchangeable units
- Abstraction to library

Two groups: Clouds and TISA

Clouds Re-architecture

- Legacy code
- Varying levels of complexity
- GOTO statements
- Redundancy throughout codebase
- Hard-coded inputs
- “Magic” numbers

Clouds Status

- New team-member for re-architecture, Steve Kohler
- Cloud mask work
- Near-term Goals:
  - Evaluate current code
  - Position for better:
    - Ability to unit test
    - Exception handle
    - Extend
  - Validate, validate, validate

TISA Re-architecture – Parallel Efforts

"In-place" (Ed. 4-5)

- Remove "dead" code
- De-duplicate
- Generalize
- Collect multi-purpose routines for library

System-level (>= Ed. 5)

- Implement new framework
- Build general library
- Add spatiotemporal flexibility

daily_mod_o.f90, daily_mod_new.f90 – Araxis Merge

New Comparison | Start/Stop | Recompare | Layout | Comparison | Preferences | Changes | Current Pane | Find | Undo/Redo | Edits | Indent | Merging | Replace Content | Swap Panes | XML | Sync Links | Bookmarks

/Volumes/Ceres/tisa/msun/git/tmp/fluxbycloudtype_edition4/tisa_grid/CER9.0P5/src/main/daily_mod_o.f90

```
          CALL interm_read(isat, interm_rec, ios)

          IF (ios == 0) THEN
            !--- region =1, MAX_REGIONS_PER_ZONE*NUM_ZONES
            zn = ZONE_NUMBER(interm_rec%region_nest)
            reg =  interm_rec%region_nest - ((zn - 1) * MAX_REGIONS_PER_ZONE)
            ihr = interm_rec%ihour       !--- Cathy Aug 23 2014

            IF (ihr==0) ihr =1
            vza(reg,zn,ihr,isat) = interm_rec%vza
            IF (vza(reg,zn,ihr,isat) > 0.0 .AND. vza(reg,zn,ihr,1) == 0.0) THEN
              CALL store_data(zn,reg,ihr,interm_rec)
            ENDIF

            !--- choose the smaller vza of the 2 satellites
            IF (vza(reg,zn,ihr,isat) > 0.0 .AND. vza(reg,zn,ihr,1) > 0.0 .AND. &
                vza(reg,zn,ihr,isat) < vza(reg,zn,ihr,1) ) THEN
              CALL store_data(zn,reg,ihr,interm_rec)
            ENDIF

            nRecs = nRecs + 1
          ELSE IF (ios == DA_EOF )THEN
            !   print*, "in read_bin_write_daily_with_sfc_nc FINISHED READing ",ios
            EXIT
          ELSE
            PRINT *, "Error in reading record: ", nRecs
            EXIT
          ENDIF
        ENDDO    !read
        CALL interm_close(isat, ios)
      ENDIF   !--- isat 2
    ENDDO     !--- isat loop
  ENDIF   !--- status

  !reorganize the global_solar_incoming onto a nested mapped output format
  !sum up the number of observations
  zone_num = 0

  DO i = sequence_num_beg, sequence_num_end
    region = nested_number(i)

    zone_num = zone_number(region)
    reg = region - ((zone_num - 1) * MAX_REGIONS_PER_ZONE) !convert back to 1-360 f
    nest_factor = NESTING_FACTOR(zone_num)
    regidx = reg
    DO j = 1, nest_factor
      DO hour_index = 1, HOURS_IN_DAY
        hour_of_month = ((iday-1) * HOURS_IN_DAY) + hour_index
        m_toa_sw_insol(regidx,zone_num,hour_index) = global_solar_incoming(region,h
      END DO
      regidx = regidx + 1 !next idex
```

/Volumes/Ceres/tisa/msun/git/tmp/fluxbycloudtype_edition4/tisa_grid/CER9.0P5/src/main/daily_mod_new.f90

```
          nRecs = 2
          DO
            CALL interm_read(isat, interm_rec, ios)

            IF (ios == 0) THEN
              !--- region =1, MAX_REGIONS_PER_ZONE*NUM_ZONES
              zn = ZONE_NUMBER(interm_rec%region_nest)
              reg =  interm_rec%region_nest - ((zn - 1) * MAX_REGIONS_PER_ZONE)
              ihr = interm_rec%ihour       !--- Cathy Aug 23 2014

              IF (ihr==0) ihr =1
              vza(reg,zn,ihr,isat) = interm_rec%vza
              if (vza(reg,zn,ihr,isat) /= real4_dflt) then
                if ( vza(reg,zn,ihr,1) == real4_dflt) then
                  CALL store_data(zn,reg,ihr,interm_rec)
                else if (vza(reg,zn,ihr,isat) < vza(reg,zn,ihr,1)) then
                  CALL store_data(zn,reg,ihr,interm_rec)
                endif
              endif

              nRecs = nRecs + 1
            ELSE IF (ios == DA_EOF )THEN
              !   print*, "in read_bin_write_daily_with_sfc_nc FINISHED READing ",io
              EXIT
            ELSE
              PRINT *, "Error in reading record: ", nRecs
              EXIT
            ENDIF
          ENDDO    !read
          CALL interm_close(isat, ios)
        ENDIF   !--- isat 2
      ENDDO     !--- isat loop
    ENDIF   !--- status

    !reorganize the global_solar_incoming onto a nested mapped output format
    !sum up the number of observations
    zone_num = 0

    DO i = sequence_num_beg, sequence_num_end
      region = nested_number(i)

      zone_num = zone_number(region)
      reg = region - ((zone_num - 1) * MAX_REGIONS_PER_ZONE) !convert back to 1-360
      nest_factor = NESTING_FACTOR(zone_num)
      regidx = reg
      DO j = 1, nest_factor
        DO hour_index = 1, HOURS_IN_DAY
          hour_of_month = ((iday-1) * HOURS_IN_DAY) + hour_index
          m_toa_sw_insol(regidx,zone_num,hour_index) = global_solar_incoming(region,
        END DO
        regidx = regidx + 1 !next idex
```

TISA Status

- ~20% overall code reduction
- Meet once/week
- Using Jira/Bitbucket/Confluence
- Arun – Ed 4-5 code changes:
- Josh – Ed 5+ framework-level construction
  - Fresh look at TISA – as a system
  - “Event Storming” session

Event Storming

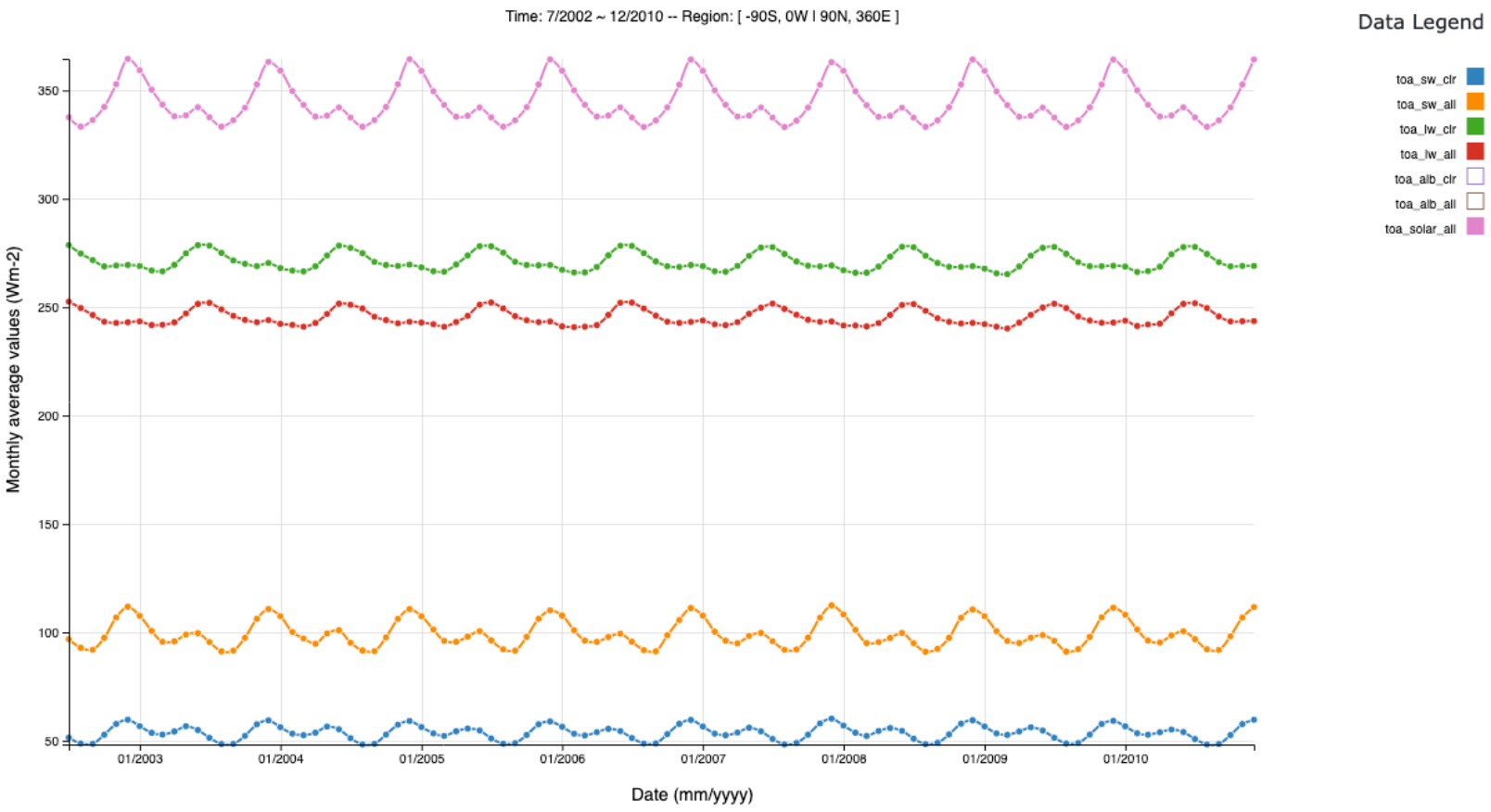
Subsetter/Ordering Tool

- Serve parameter-level subsetting of CERES data

- Two versions:
  - Internal/Science Team-only
  - Public

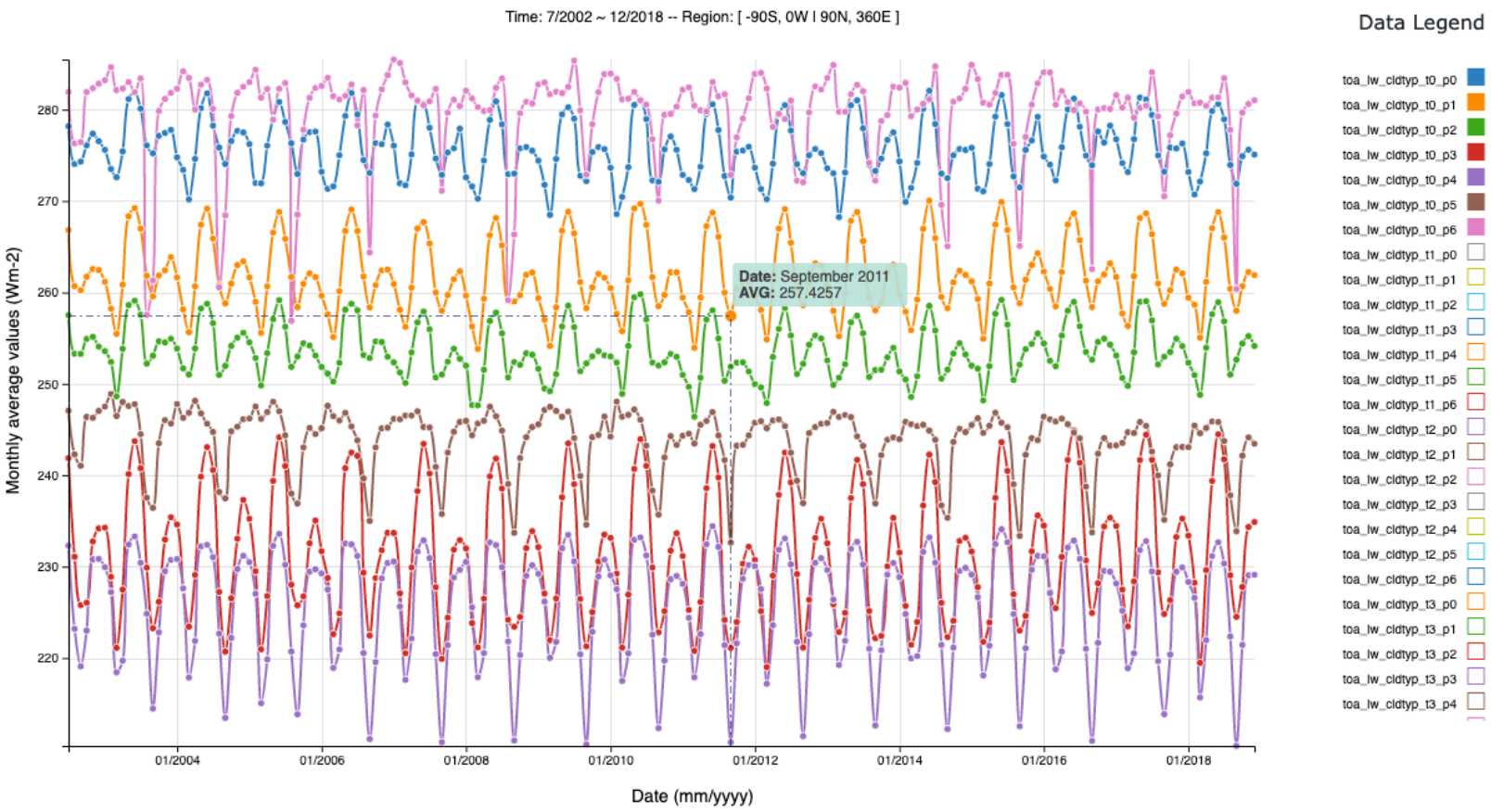
- Prep for FluxByCldTyp:
  - Monthly means map
  - Time series – Babak Samani

CERES_FluxByCldTyp_Ed4.1-MON_TERRA-AQUA



48.2485 | 364.5251 | Redraw | Reset | Save data

CERES_FluxByCldTyp_Ed4.1-MON_TERRA-AQUA



210.38 | 285.4818 | Redraw | Reset | Save data

Subsetter/Ordering Tool

- Migrating:
  1. Virtualization
  2. Containerization (OpenShift)
- Emphases: availability and flexible ordering

Energy Balanced and Filled (EBAF)

Climate Data Record (CDR) of monthly TOA fluxes and consistent computed surface fluxes and clouds suitable for analysis of variability at the intra-seasonal, inter-annual, and longer time scales.

Data Product Information	Description	Parameter	Resolution	Version/ Availability	Order Data
EBAF	Monthly and climatological averages of observed TOA and computed surface all-sky, clear-sky (spatially complete), and cloud radiative effect (CRE) fluxes and associated clouds. The TOA net flux is constrained to the ocean heat storage. EBAF Data Quality Summary				Browse & Subset

Synoptic TOA and surface fluxes and clouds (SYN)

Hourly gridded observed TOA and Fu-Liou RT surface fluxes and clouds

Systems & CM

New Systems

EMC Isilon Storage

- Computational Research Facility (CRF)
- Configuration on hold due to COVID-19

HPE DL360 Servers

- CRF and ASDC
- Virtualized
  - Red Hat virtual machines
  - OpenShift PaaS virtual machines

Updated Data Formats

Upcoming Editions:

HDF5/netCDF4

- Many iterations with ASDC – ingest/metadata check
- First products: Ed 1 NOAA-20 and Ed 2 S-NPP **SSF**
- CM Team – lots of work!

Updated Metadata

Upcoming Editions

- Conventions:
  - Unified Metadata Model (UMM)
  - Climate-Forecasting (CF)
  - Attribute Convention for Data Discovery (ACDD)
  - CERES-specific

- Attributes in HDF5/netCDF4 files

- Separate "*.met" file until no longer necessary

Updated Metadata

Clouds and the Earth's Radiant Energy System (CERES)

Data Management System

Metadata Specification

Version V2

Primary Authors

Nelson Hillyer, Denise Cooper, Joshua Wilkins, Tammy Ayers, Dennis Keyes, Hunter Winecoff

Science Systems and Applications, Inc. (SSAI)
One Enterprise Parkway, Suite 200
Hampton, VA 23666

Updated Metadata

Metadata Checker Tool – Hunter Winecoff

- Checks against CERES metadata standards
- Must run
  - before code delivery
  - as part of integration testing

Summary

- New products and editions on the way!
- Virtualizing
- Engaging in cloud preparation
- Refreshing file formats, metadata
- Improving user experience