

CERES Data Management Status

Presented to
CERES Science Team

Mike Little
Erika Geier
NASA Langley Research Center
May 22, 2002

m.m.little@larc.nasa.gov

e.b.geier@larc.nasa.gov

CERES Data Management Status

Agenda

Science Data Product Status

Processing Environment Status

CERES Science Data Customers

Technology Advances

CERES Science Data Source Status

Platform	Satellite	Source	Status	Comments
a	TRMM	Instrument Data Stream	Complete	
a	Terra-CERES	Instrument Data Stream	Incoming	
a	Aqua-CERES	Instrument Data Stream	CoversOn	Started 5/11/02 data date; Beta1 processing.
e	TRMM	Spacecraft	Incoming	Current for Edition 2- VIRS only SSF
e	Terra	EDOS-EMOS-FDS-LDAAC-LaTIS	Incoming	Sporadic data missing (2000-6.99%, 2001-0.1%, 2002-1.5%). Work in progress to obtain back data and to correct l
e	Aqua	EDOS-LaTIS	Incoming	Data started 5/11/02 data date.
eris	TRMM	Spacecraft	Incoming	Current for Edition2- VIRS only SSF
eris	Terra	EDOS-LDAAC-LaTIS	Incoming	Data missing (2000-3.32%, 2001-0.2%, 2002-1.5%). in progress to obtain back data.
eris	Aqua	EDOS-LaTIS	Testing	Started 5/11/02 data date.
	TRMM-VIRS	Instrument	Incoming	Edition 2-VIRS only SSF
	Terra - MODIS	GDAAC	Incoming	Code written at LaRC, data subsetted at GDAAC for delivery to LaTIS
	Aqua-MODIS	GDAAC	In test	Subsetting code from Terra being modified for Aqua GDAAC, next week.
		DOA	Incoming	Quality improving, but not yet adequate for CERES u
		ECMWF	Sample	Through 7/31/01. Award 06/01/02 for 8/1/01-8/30/02
		GHRC μ wave humidity	Incoming	Precipitable water

CERES Science Data Products Delivered Since May, 2001-TRMM

Data Set	Comments
SF_TRMM-PFM-VIRS_Edition2A	9 months of TRMM data processed using TRMM Edition2 Clouds and Edition2A ADMs; Edition2B to be processed shortly with same clouds, improved Aerosol A optical depths, Edition2B ADMs, and improved surface fluxes
SF_TRMM-PFM-VIRS_Edition2A-TransOps	Same code as Edition2A, but run for data collected during periods when CERES was briefly powered on in 1999.
SF_TRMM-SIM-VIRS_Edition2-VIROnly	Starts Sept 1998 and is available whenever CERES instrument is powered on (no Edition2A or Edition2A-TransOps available) Only cloud parameters and Aerosol A optical depths available - NO CERES data. CERES FOV geolocation is simulated based on attitude and ephemeris. Still processing.
CRS_TRMM-PFM-VIRS_Beta3 CRS_TRMM-PFM-VIRS_Subset-Beta3	Most recent CRS produced. Used TRMM Edition2A SSF as input. Processed all TRMM months. Subset-Beta3 contains only FOVs which fall over valid regions. Beta3 contains all FOVs.
CRS_TRMM-PFM-VIRS_Beta3-TransOps	Beta3-TransOps used TRMM Edition2-TransOps SSF as input. (Select hours during year 1999 when CERES was operational)
FC_TRMM-PFM-VIRS_Beta4	All TRMM months expected to be run. Last Beta run before Edition2B.
RBAVG_TRMM-PFM-VIRS_Beta3	1 month processed - Feb'98
SF_TRMM-PFM-VIRS_Edition2B SF_TRMM-PFM-VIRS_Edition2B-TransOps	Updated ADMs; improved aerosols; updated surface algorithms. Apr'98 & Feb'98 processed for review. Remaining 7 months to process next week.
SW_TRMM-PFM_VIRS_Beta2	May 98 processed using CRS Beta3 as input.

CERES Science Data Products Delivered Since May, 2001- TERRA

Data Set	Comments
SF_Terra-FM1-MODIS_Beta2/ SF_Terra-FM2-MODIS_Beta2	4 months processed: Nov'00, Dec'00, Apr'01, May'01. Edition1 IES used as input (no FM2 drift correction)
SF_Terra-FM1-MODIS_Beta2-overARM/ SF_Terra-FM2-MODIS_Beta2-overARM	Nov'00 - Jul'01 based on MODIS V003 availability. 4 ARM sites: Mannus, Nauru, Barrow, Oklahoma. Edition1 IES used as input (no FM2 drift correction)
SF_Terra-FM1-MODIS_Beta2-overCLAMS/ SF_Terra-FM2-MODIS_Beta2-overCLAMS	Chesapeake Lighthouse site. July'01 only during period of CLAMS experiment. Edition1 IES used as input (no FM2 drift correction)
SF_Terra-FM1-MODIS_Beta3/ SF_Terra-FM2-MODIS_Beta3	March 2001, December 15-21, 2000, June 1-7, 2001. first SSF containing MODIS aerosols. Edition1 IES used as input (no FM2 drift correction)
SF_Terra-FM1-MODIS_Beta3-overARM/ SF_Terra-FM2-MODIS_Beta3-overARM	Nov'00 - Jul'01 based on MODIS V003 availability. 4 ARM sites: Mannus, Nauru, Barrow, Oklahoma. first SSFs containing MODIS aerosols. Edition1 IES used as input (no FM2 drift correction)
CRS_Terra-FM1-MODIS_Beta2/ CRS_Terra-FM2-MODIS_Beta2	Beta2 SSFs used as input. 4 months processed: Nov'00, Dec'00, Apr'01, May'01
CRS_Terra-FM1-MODIS_Beta2-overARM/ CRS_Terra-FM2-MODIS_Beta2-overARM	Beta2-overARM SSFs used as input. Nov'00 - Jul'01 based on MODIS V003 availability. 4 ARM sites: Mannus, Nauru, Barrow, Oklahoma
CRS_Terra-FM1-MODIS_Beta2-overCLAMS/ CRS_Terra-FM2-MODIS_Beta2-overCLAMS	Beta2-overCLAMS SSFs used as input. Chesapeake Lighthouse site. July'01 only during period of CLAMS experiment
CSW_Terra-FM1-MODIS_Beta2/ CSW_Terra-FM2-MODIS_Beta2	Processed 4 months of Terra Beta2 CRS (Nov'00, Dec'00, Apr'01, May'01) using Beta2 CRS as input. Expect no public release.

CERES Science Data Products

TRMM Soon-to-be-Processed

Data Set	Comments	When?
CRS_TRMM-PFM-VIRS_Edition2B CRS_TRMM-PFM-VIRS_Edition2B-TransOps	First validated CERES SARB product to be made public	Expected May'02; public release Summer'02.
SFC_TRMM-PFM-VIRS_Edition2B	First validated CERES TISA gridded product to be made public	Expected June'02; public release July'02.
RBAVG_TRMM-PFM-VIRS_Beta4	9 mo of TRMM Beta4 SFC are input; all the latest algorithm changes	Expected May'02; public release May/June'02.
RBAVG_TRMM-PFM-VIRS_Edition2B	First validated CERES TISA averaged product to be made public	Expected June/July'02; sch public release Jul'02.
CSW_TRMM-PFM-VIRS_Beta3	9 mo of TRMM Edition2B CRS are input	Expected July'02; public release Aug'02.

CERES Science Data Products

Terra Soon-to-be-Processed

Data Set	Comments	When?
BDS_Terra-FM1_Edition2 BDS_Terra-FM2_Edition2	FM2 drift corrected. Mar'00 - present	Expected May'02; public release 06/20/02.
CS8_Terra-FM1_Edition2 CS8_Terra-FM2_Edition2 CS9_Terra-FM1_Edition2 CS9_Terra-FM2_Edition2 CS9_FM1+FM2_Edition1 CS4_Terra-FM1_Edition2 CS4_Terra-FM2_Edition2 CS4_FM1+FM2_Edition2	FM2 drift corrected; Snow/ice bug corrected. Mar'00 - present	Expected May'02; public release scheduled 06/20/02.
SSF_Terra-FM1-MODIS_Beta4 SSF_Terra-FM2-MODIS_Beta4 SSF_Terra-FM1-MODIS_Beta4-overARM SSF_Terra-FM2-MODIS_Beta4-overARM	Improved cloud properties; final TRMM Edition2B ADMs; includes MODIS aerosols Edition1 IES are input. 2 global weeks; Nov'00-Jul'01 over 4 ARM sites	Expected late May'02; public release
SSF_Terra-FM1-MODIS_Edition1A SSF_Terra-FM2-MODIS_Edition1A	1st validated SSF for Terra; uses TRMM Ed2B ADMs; & MODIS aerosols. Ed2 IES are input (FM2 drift corrected) Nov'00 - Jul'02	Expected Jun '02; public release scheduled 07/10/02.
FC_Terra-FM1-MODIS_Beta1 FC_Terra-FM2-MODIS_Beta1	1st Beta runs of this product.	Expected Aug'02; public release scheduled 09/10/02.
RBAVG_Terra-FM1-MODIS_Beta1 RBAVG_Terra-FM2-MODIS_Beta1	1st Beta runs of this product. Uses MODIS aerosols.	Expected Oct'02; public release scheduled 11/10/02.

CERES Science Data Products

Aqua Soon-to-be-Processed

Data Set	Comments	When?
DS_Aqua-FM3_Beta1 DS_Aqua-FM4_Beta1	First Aqua products!	Expected after cover
S8_Aqua-FM3_Beta1 S8_Aqua-FM4_Beta1 S9_Aqua-FM3_Beta1 S9_Aqua-FM4_Beta1 S9_FM3+FM4_Beta1 S4_Aqua-FM3_Beta1 S4_Aqua-FM4_Beta1 S4_FM3+FM4_Beta1	First Aqua products!	Expected after cover

Older CERES Science Data Products

Still of Interest

Data Set	Comments
BDS_TRMM-PFM_Edition1	
BDS_TRMM-PFM_Transient-Ops2	Current, validated BDS data sets for TRMM
ES8_TRMM-PFM_Edition2	
ES8_TRMM-PFM_Transient-Ops2	
ES9_TRMM-PFM_Edition2	
ES4_TRMM-PFM_Edition2	Current, validated ERBElke data sets for TRMM
BDS_Terra-FM1_Edition1/ BDS_Terra-FM2_Edition1	Soon to be replaced by Edition2 data set (FM2 drift corrected)
ES8_Terra-FM1_Edition1/ ES8_Terra-FM2_Edition1	
ES9_PFM+FM1_Edition1/CER_ES9_PFM+FM2_Edition1/ ES9_PFM+FM1+FM2_Edition1	
ES9_FM1+FM2_Edition1/ ES9_Terra-FM1_Edition1/ ES9_Terra-FM2_Edition1	
ES4_PFM+FM1_Edition1/ ES4_PFM+FM2_Edition1/ ES4_PFM+FM1+FM2_Edition1	
ES4_FM1+FM2_Edition1/ ES4_Terra-FM1_Edition1/CER_ES4_Terra-FM2_Edition1	Soon to be replaced by Edition2 ERBElke data sets (FM2 drift corrected)

Science Data Product URLs and Contacts

Ordering Data

- http://eosweb.larc.nasa.gov/HBDOCS/langley_web_tool.html

Contact Points

- All questions regarding production data products and their use
 - E-mail: larc@eos.nasa.gov
 - Langley ASDC Customer Service
- Contents of this presentation: m.m.little@larc.nasa.gov

SCF Configuration and Planning

Objectives and capabilities

Provide environment for development and testing of science codes

Provide computing capacity for unscheduled requirements

Improvements to Communications

ASDC to SCF - Gigabit backend isolated from other traffic

SCF to local Users - Upgrade to 100Mbps via switch in progress

SCF to Offsite Gateway- continues to be 10Mbps until upgrades complete

- ETA end of summer 2002

SAIC development environment to 100Mbps

ITS Issues

Access control improvements - VPN in testing by LaRC ITS

- Expect to begin CERES testing by July

Threat Conditions - over 60 actual break-in attempts every two weeks

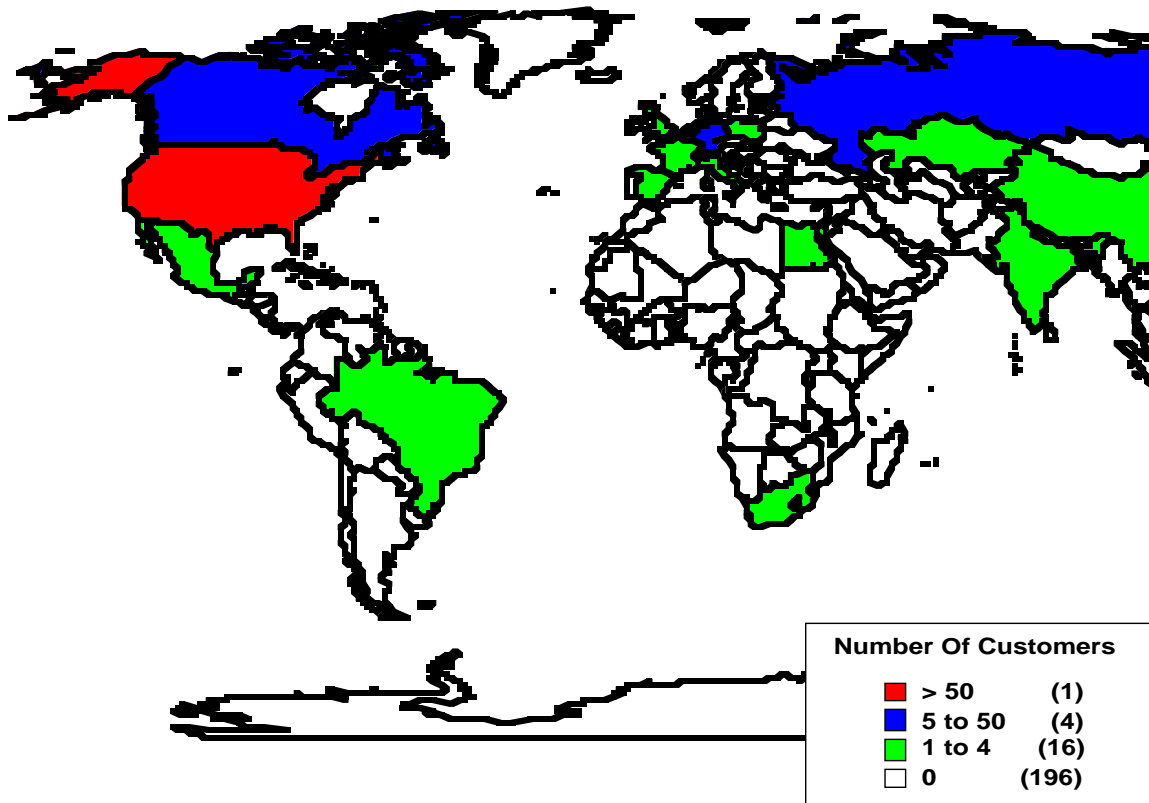
Consolidation of archives to ASDC

Tape Library

DMSS

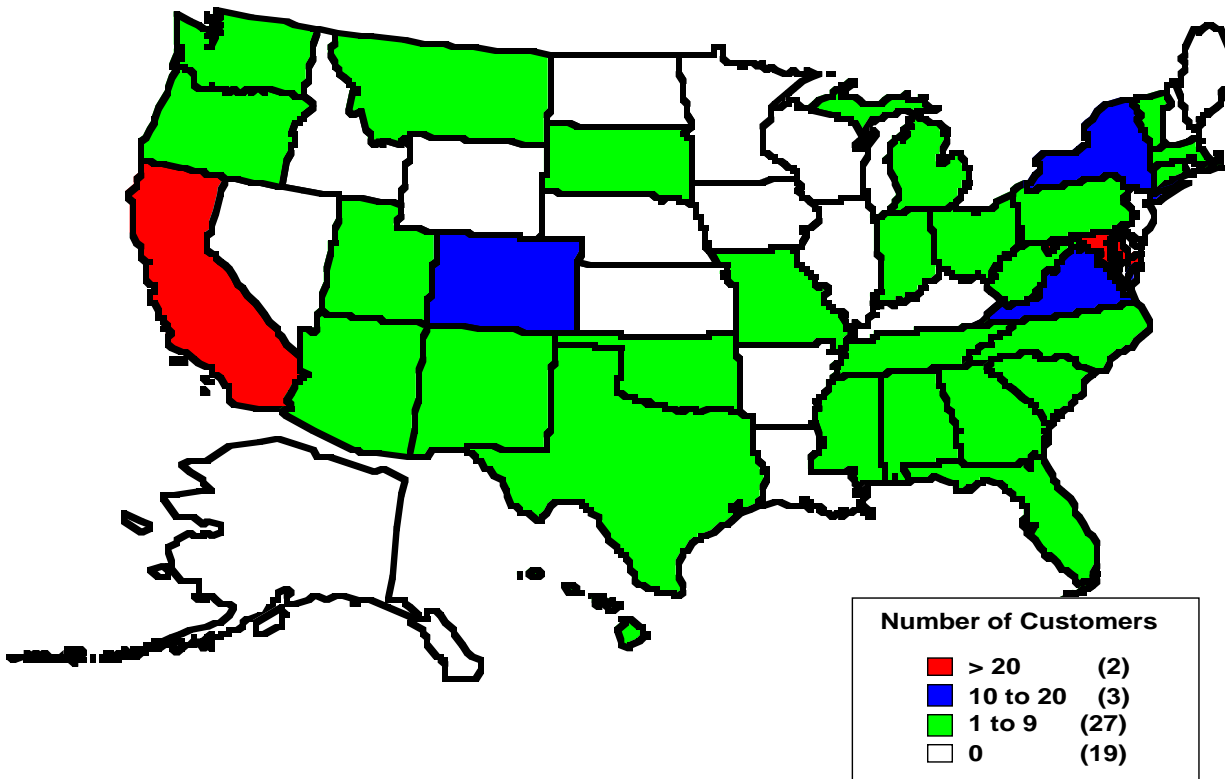
CERES Customers FY2001

CERES Customer Distribution - FY 2001



CERES Customers FY2001

CERES US Customer Distribution - FY 2001



CERES Data Distribution Volumes

		FY01	1Q02
BDS	Terra	85,787	Not Avail
ES4	Terra	3,858	741
ES8	TRMM	891,741	5,439
	Terra	867,102	174,672
ES9	TRMM	5,621	793
	Terra	31,006	244
SFC	TRMM	15,023	Not Avail
SSF	TRMM	658,376	218,261
	Terra	47,269	107,073

Earth Science Technology Advances

Computational Technology Requirements Workshop

Identify needs for NASA funded work during FY04-FY06

- NASA unique issues

Problems faced by Related Communities

- Lack of sufficient, affordable processing capacity
 - Throughput of initial products
 - Reprocessing of products with improved algorithms
- Data Management
 - Data Mining or subsetting
- Visualization of data
- Code Development
 - Delay time to re-implement a PGE with a new algorithm
 - Shortage of qualified developers
 - Cost

Earth Sciences Modeling Framework

Joint development between NASA, NOAA, DoD, DoE, NSF

CERES Data Management Status Back-up Charts

1. ESTO Computational Technology Workshop Summary

Earth Science Technology Office
Computational Technology
Requirements Workshop
Trip Report

Mike Little, LaRC
Troy Anselmo, SAIC
04/30/02 to 05/01/02

Outline

Workshop Objectives

STO Background

Challenges and Requirements

- Climate modeling
- Weather prediction
- Solid Earth modeling

Workshop Objectives

Earth Science Enterprise research and applications strategy defined for next decade

WE working to

- Identify computational technologies (CT) necessary to support research objectives

- Identify appropriate areas for NASA investment

- Develop budget guidelines and justification for FY04 – FY06

Objective of workshop is to:

- Evaluate computing needs and evaluate against current capabilities

- Identify gaps between now and 2010

- Solicit inputs from research community for technologies required

The workshop placed substantial emphasis on technologies required to meet prediction goals in the areas of climate, weather, and solid earth

Background

Product Lines in ESTO

Instrument Incubator Program

Information Science Program

Computational Technology Program

Focus on Prediction and Modeling function

Regional impact is important to this Administration

Need revolutionary computing environment to achieve breakthroughs required

- Funded for evolution, not revolution
- 5 fold increase in data acquisition storage & retrieval
- Need migrate SW to new architectures as they emerge & to exploit their features

Weather quality must be NWS performance, not just research world.

New ways of funding model development

Most funding from Research Announcements

- How to cooperate
- Present leveraging resources with cooperative partners
 - Resources being duplicated

Summary - Climate

Goals for 2010

- 6-12 month routine seasonal prediction, 12-24 month experimental prediction
- Coupled atmosphere-land-ocean-sea-ice models
- 1/4° atmosphere, 1/10° ocean; 100-member ensembles
- Multiple scale ensembles (1/4 degree short term (1st season), 1/2 degree longer term (2nd season))
- Resolving the mesoscale with non-hydrostatic models (downscaling for regional forecasts)
- Expanded physics (High latitudes, Stratosphere, Clouds, Dynamic vegetation)
- Prediction of extreme events

Current capabilities

- Distributed Oceanographic Data System (DODS) used to subset data
- FORTRAN environment
- Auto differentiation tool to assess model sensitivities

Capabilities required for 2010

- 500TB for multi year ensembles or extreme events
- Short Term: 10^3 X current performance for single image (~80 Tflop/s sustained aggregate requirement)
- Medium-term: 6.6×10^4 X current performance for single image (~ 660 Tflop/s sustained aggregate requirement)
- Problem solving environment (ESMF)
 - Enhanced model development, facilitate model inter-changeability, provide buffer to underlying hardware architecture
 - advanced data assimilation, integrated tools for data management, analysis, and visualization

Distributed access to archive

- 500TB per image, 5TB per “experiment”, 25 PB per multi ensemble experiment
- Ability to share > 30TB for analysis (extreme event for 1 season)

Summary - Weather

Goals

Accuracy and latency requirements driven by NOAA and US Navy

Challenges

Data Management/Data Assimilation is highest priority

- Massive increases in number of sources and data volumes
- Improved algorithms for data reduction and extraction required (need reduction in assimilation of 10^3)

Systems engineering for HPC

- Insure commodity IT will be adequate

Requirements

Technology Development

- SW to integrate systems (Schedulers, performance tools, file storage systems)
- Interface technology, especially high speed communications

Software tools

- System performance tools, improved performance compilers, auto parallelization

Development/Analysis environments

- Facilitate parallelization, hardware independence
- Improved visualization and analysis tools

Summary - Solid Earth

Challenges

Data and Metadata

- Difficult to deal with due to large data volumes, heterogeneous and distributed
- Provide access within 5 minutes and/or automatically

Modeling complexity

- Multi-scale, multi-component, multi-source
- Real time while event is occurring

Understanding the data

- Multi-dimensional
- Bottlenecks due to differences in registrations, map projections, and layers

HPC performance

Requirements

Create virtual observatory for solid earth research

- > 1000 sites, 1TB – 1PB of data
- Provide access to 100GB files from 40TB sets on demand within 5 minutes

Single, integrated solid earth problem solving environment

- Parallel algorithms, model assessment methodology and tools, data assimilation tools
- Performance at 100TFLOPS sustained by 2010

Advanced Visualization tools

- 1TB data sets
- Data mining and understanding, pattern/wavelet recognition capabilities

Science Data Product Status

Product	ID	Function	Maturity	ETA Public	Comments

Science Data Production Status

Product	Source	Period	Ver	Comments

Processing Environment Status

CF

SDC

- Covered Separately by Bruce Barkstrom

Subsystem Codes

- Simulator

- Instrument

- Subsystems 1-11

CERES Subsystem Status

SS	Subsystem	TRMM	TERRA	Comments
1	Instrument	Edition 2	Edition2	PGE CER1.3p3-
2	ERBE-like Inversion	Edition 1	Edition1 Edition2	PGE CER2.4p1-spectral response
3	ERBE-like Averaging	Edition 1	Edition1	
4.1-4.4	Clouds (SSF)	Edition 2B	Beta2 Beta3 over ARM	
4.5-4.6				
5	SARB	Edition 1	Beta3	
6	Gridding		Beta4	
7.1	TISA		Beta1	
7.2	Synoptic SARB			
8	TISA Averaging		Beta1	
9	TISA Gridded		Beta4	
10	TISA Ave (Mo TOA/Surface)		Beta1	
11	GGEO		Beta3	
12	MOA		ECMWF-GEOS3	
	CERESlib			