



CERES Data Management Status

Presented to
CERES Science Team

Mike Little

Erika Geier

(input from Michelle T. Ferebee, ASDC)

NASA Langley Research Center

October 25, 2006

michael.m.little@nasa.gov

erika.b.geier@nasa.gov



Agenda



-
- CERES Systems Issues
 - Technical
 - Code Conversion
 - Computing and Storage Capabilities
 - Work with Atmospheric Sciences Data Center (ASDC) to automate CERES Production
 - Management
 - CERES Data Products Status
 - Edition 3
 - Specific data products



SCF Code Conversion



- Purpose: Permit multi-platform processing with same scientific results
- Migrate from SGI3800 (IRIX) to G5 (OSX, LINUX) & Intel LINUX
 - Scientifically Identical Output (platform invariance)
 - Define criteria for each subsystem
 - Plan for validation and certification
 - Convert code to multi-platform (not just a different one)
 - Scripts and science codes
 - Compiler, processor/memory, software engineering
 - Run code on cluster (SGE, S4PM, environment)
 - Big-endian to Little-endian (flags, compilers, binary data, etc.)
- Production Codes for Subsystems
 - Done: Toolkit, cereslib, Instrument, SARB, MOA, TISA-grid
 - WIP: Inversion, Clouds, TISA Averaging, GGEO
- Analysis Codes - limited so far
 - Initial orientation for users: system architecture and SGE
 - Need to debug code conversion process & hardware environment
 - Develop skills among Data Management Team to provide help



SCF Technical Developments



- Migrate Processing from SGI to G5 Cluster to Intel Cluster
 - Reduce cost, increase capacity and flexibility
 - Ensure maintainability and reliability of facility
 - Use SunGridEngine (SGE) to manage resources (8 queues)
- Migrate Data into Storage Area Network (SAN)
 - Increase reliability, access to data, reduce staging work
- Convert Authentication to Open Directory
 - Use by cluster and SAN
- User Impact
 - No direct interaction with cluster nodes
 - Must understand how to use SGE and architecture
 - New diagnostic tools for cluster jobs
 - Desktop computer to do some processing and interaction



SCF Capacity Time-Phasing



System	FY03	FY04	FY05	FY06	FY07	FY08	FY09
SGI Processors	32	72	72	72	72/40	40	Retire
G5 Processors			10	16/54	54/80	80	80
Intel Processors					8	16	32
SGI Disk (TB)	10	11	11	11	1	1	Retire
Odd Disk (TB)	1	5	10	15	5	5	5
SAN Disk(TB)-R5	20	20	20	20/40	80/100	120	140
-R0					7	7	7
Legato Backup	20	20	20	20	20	20	Retire
Exabyte Backup					200	300	400



CERES Automation at ASDC



- Purpose: Process same quantity with very few operator resources
- Scheduled to be operational in FY08
 - Dependent upon ANGe transformation
 - ANGe Operational by beginning January, 2008
 - Plan to begin CERES Automation integration after LATIS/ECS rolloff
 - WIP to accomplish all long-lead items
 - Design Document
 - Interface Spec for CERES (PRs, PCF)
 - Implementation of PR functionality
 - Modify S4PM to run both SGI and LINUX processing systems
 - Expect to do some PR & limited SS testing before May, 2007
- Requirements for CERES Data Management Team
 - Replace ASCII/PCF Generator scripts (WIP)
 - Revised PR submission model (Planning in Progress)
 - Not dependent upon Code Conversion
- Near term activities
 - Design Document (ASDC)



Management Developments



-
- Support Services Contract moves from SAIC to SSAI
 - Results from 10-month competition
 - Hand-off on December 01, 2006
 - High degree of incumbent employee retention required
 - Replace all (~95%) Government Furnished Equipment (GFE) with contractor supplied equipment
 - Infrastructure Imposes New Constraints
 - Expect to expend considerable resources to comply with Security requirements
 - US National Institutes of Standards and Technology (NIST) Federal Information Processing Standards (FIPS)
 - Assess each individual system in terms of threat, vulnerability, risk, type of data
 - Increased restrictions on sharing of and access to data and processing capacity
 - Increased requirements to report on use of Information Technology and to prove compliance with Security requirements



Instrument Cal/Val Runs in SCF



- Status - Enough data for 1st pass Ed3 cal
 - Need at least 1 more
- Performance Considerations
 - Improved turn-around 2 Satellite-months in an hour
 - All day on the SGI
 - Must be analyzed before running next slug of data
 - Storage System
 - Performance/Reliability still not good enough
 - Migrating proper data to Storage System from diverse hosts,
 - Recent Lessons Learned
 - Processing environment variable setting created processing errors
 - Stack size, memory usage, file size
 - Trapping processing signals and errors different than SGI
- Open Issues
 - C-routine naming conventions (`_name`) compared to SGI, LINUX
 - Storage System



Edition3 Getting Closer



- Edition2 to end with December 2005 data
 - BDS, ERBElike, SSF, SFC, CRS, and FSW complete
 - SRBAVG creeping closer to completion
 - SYN, AVG, ZAVG will be a while...
- Edition3 expected to start processing in 2007
 - Reprocess all data and do forward processing for Terra and Aqua
 - BDS and ERBElike could be available by the next meeting!
 - All other products 2008 or later (Beta-Ed3 runs start in 2007)
 - Product parameters may be added or changed
 - SSF size increase over 25%!
 - Terra and Aqua data sets will be of very similar quality
 - Inputs changing
 - MODIS collection5 will be used
 - MERRA may replace GEOS-4



CERES Science Data Sets

New Since May 2006



Data Set	Comments
CER_SSF_Aqua-FM4-MODIS_Ed2A-NoSW CER_SFC_Aqua-FM4-MODIS_Ed2A-NoSW	<ul style="list-style-type: none"> • 3/30/05 hour 18 through 1/1/06 • Same algorithms and processing as Edition2 • Due to SW channel failure, no daytime SW or LW • SSF publicly available
CER_CRS_Aqua-FMx-MODIS_Edition2A CER_FSW_Aqua-FMx-MODIS_Edition2A	<ul style="list-style-type: none"> • Aqua Edition2A SSF from crosstrack instrument are input • 7/02 - 12/05 • Public release expected
CER_BDS_TRMM-PFM_Edition1-CV CER_ES8_TRMM-PFM_Edition1-CV CER_ES4_TRMM-PFM_Edition1-CV CER_ES9_TRMM-PFM_Edition1-CV CER_BDS_TRMM-PFM_Ed1-CV-TransOps CER_ES8_TRMM-PFM_Ed1-CV-TransOps CER_ES4_TRMM-PFM_Ed1-CV-TransOps CER_ES9_TRMM-PFM_Ed1-CV-TransOps	<ul style="list-style-type: none"> • 12/27/1997 through 8/31/1998; March 2000 (Edition1-CV) • Select transient operation periods during 1999 (Ed1-CV-TransOps) • Calibration/validation data sets that replace Edition1 • Algorithms to be held constant across entire data set • Algorithms consistent between Terra and Aqua • Public release expected



CERES Science Data Sets



Expected Before Next Science Team Meeting (Edition2)

Data Set	Comments	When?
CER_SSF_Aqua-FMx-MODIS_Edition2B	<ul style="list-style-type: none"> • Corrects Model B SW surface fluxes, all other parameters remain the same as Edition2A • Process 7/02 -12/05 • Public release expected 	Late 2006
CER_SFC_Aqua-FMx-MODIS_Edition2B	<ul style="list-style-type: none"> • Edition2B SFC are input • Process 7/02 – 12/05 • Public release expected 	Immediately following Aqua Edition2B SSF
CER_SRBAVG_Aqua-FMx-MODIS_Beta6	<ul style="list-style-type: none"> • Edition2B SFC are input • Process 7/02 – 5/04 and possibly 6/04 -12/05 • Available to Science Team only (no public release) 	Early 2007
CER_CRS_Aqua-FMx-MODIS_Edition2B	<ul style="list-style-type: none"> • Corrects Model B SW Surface fluxes, all other parameters remain the same as Edition2A • Process 7/02 -12/05 • Public release expected 	Spring 2007
CER_FSW_Aqua-FMx-MODIS_Edition2B	<ul style="list-style-type: none"> • Edition2B CRS are input • Process 7/02 – 12/05 • Public release expected 	Immediately following Aqua Edition2B CRS
CER_SYN_Terra-FMx-MODIS_Beta3 CER_AVG_Terra-FMx-MODIS_Beta3 CER_ZAVG_Terra-FMx-MODIS_Beta3	<ul style="list-style-type: none"> • Process 2002 (one year) • Available to Science Team only (no public release) 	Late 2006 or early 2007



CERES Science Data Sets



Expected Before Next Science Team Meeting (Edition3)

Data Set	Comments	When?
CER_BDS_Aqua-FMx-MODIS_Edition3 CER_BDS_Terra-FMx-MODIS_Edition3 CER_ES8_Aqua-FMx_Edition3 CER_ES8_Terra-FMx_Edition3 CER_ES4_Aqua-FMx_Edition3 CER_ES4_Terra-FMx_Edition3 CER_ES9_Aqua-FMx_Edition3 CER_ES9_Terra-FMx_Edition3	<ul style="list-style-type: none"> • The first Edition3 data sets! • Terra to process 2/25/00 forward • Aqua to process 6/18/02 forward 	Spring 2007
CER_SSF_Terra-FMx-MODIS_Beta1-Ed3 CER_SFC_Terra-FMx-MODIS_Beta1-Ed3 CER_SSF_Aqua-FMx-MODIS_Beta1-Ed3 CER_SFC_Aqua-FMx-MODIS_Beta1-Ed3	<ul style="list-style-type: none"> • First Beta run of proposed Edition3 data sets (There will be a second Beta run several months after the first completes.) • Plan to process 8 months of data per instrument • Many new parameters (see SSF - Edition3 in CERES Data Products Catalog) • Some parameters changed or deleted 	Spring 2007



Older CERES Science Data Sets

Still of Interest (Calibration/Validation)



Data Set	Comments
CER_BDS_Terra-FMx_Edition1-CV CER_ES8_Terra-FMx_Edition1-CV CER_ES4_Terra-FMx_Edition1-CV CER_ES9_Terra-FMx_Edition1-CV CER_BDS_Aqua-FMx_Edition1-CV CER_ES8_Aqua-FMx_Edition1-CV CER_ES4_Aqua-FMx_Edition1-CV CER_ES9_Aqua-FMx_Edition1-CV	<ul style="list-style-type: none"> • All Terra and Aqua-FM3 data from time the covers opened through present • Aqua FM4 data from 6/19/02 - 3/29/05 • Replaces Edition1 • Algorithms to be held constant across entire data set • Publicly available
CER_BDS_Aqua-FM4_Ed1-CV-NoSW CER_ES8_Aqua-FM4_Ed1-CV-NoSW CER_ES4_Aqua-FM4_Ed1-CV-NoSW CER_ES9_Aqua-FM4_Ed1-CV-NoSW	<ul style="list-style-type: none"> • 3/30/05 through present • Same algorithms and processing as Edition1-CV • Due to SW channel failure, no daytime SW or LW radiances and fluxes • Available to Science Team
CER_BDS_Aqua-FM4_AnomOps-Ed1-CV CER_ES8_Aqua-FM4_AnomOps-Ed1-CV CER_ES4_Aqua-FM4_AnomOps-Ed1-CV CER_ES9_Aqua-FM4_AnomOps-Ed1-CV	<ul style="list-style-type: none"> • 7/19/05 through 8/30/05 • Same algorithms and processing as Edition2 • Due to instrument temperature changes made in hopes of reversing SW channel problem, data quality varies from Ed1-CV-NoSW • Available to Science Team (no public release)



Older CERES Science Data Sets

Still of Interest (TRMM)



Data Set	Comments
CER_BDS_TRMM-PFM_Edition1 CER_ES8_TRMM-PFM_Edition2 CER_ES4_TRMM-PFM_Edition2 CER_ES9_TRMM-PFM_Edition2 CER_SSF_TRMM-PFM-VIRS_Edition2B CER_SFC_TRMM-PFM-VIRS_Edition2B CER_SRBAVG_TRMM-PFM-VIRS_Edition2B CER_CRS_TRMM-PFM-VIRS_Edition2C CER_FSW_TRMM-PFM-VIRS_Edition2C	<ul style="list-style-type: none"> • TRMM period Jan'98 – Aug'98; Mar'00 • Publicly available
CER_BDS_TRMM-PFM_Transient-Ops2 CER_ES8_TRMM-PFM_Transient-Ops2 CER_SSF_TRMM-PFM-VIRS_Edition2B-TransOps	<ul style="list-style-type: none"> • Transient Operations period Jan'99 – Jul'99 • BDS not publicly released; others publicly available
CER_ES4_PFM+FM1+FM2_Edition2 CER_ES9_PFM+FM1+FM2_Edition2 CER_ES4_PFM+FM1_Edition2 CER_ES9_PFM+FM1_Edition2 CER_ES4_PFM+FM2_Edition2 CER_ES9_PFM+FM2_Edition2	<ul style="list-style-type: none"> • Mar'00 only month of TRMM, Terra overlap • Publicly available
CER_SSF_TRMM-SIM-VIRS_Edition2-VIROnly	<ul style="list-style-type: none"> • Processed only for months when CERES was off • Processed through Jul'01; no additional months expected • Publicly available



Older CERES Science Data Sets

Still of Interest (Terra)



Data Set	Comments
CER_BDS_Terra-FMx_Edition2 CER_ES8_Terra-FMx_Edition2 CER_ES4_Terra-FMx_Edition2 CER_ES9_Terra-FMx_Edition2	<ul style="list-style-type: none">• Rev1 SW corrections available• 2/25/00 – 12/31/05 processed• Publicly available
CER_ES4_FM1+FM2_Edition2 CER_ES9_FM1+FM2_Edition2	<ul style="list-style-type: none">• Processed Mar'00 through Dec'02, production then halted• Replaced by Terra/Aqua combined Crosstrack data sets (FM1+FM3, FM1+FM4 Edition2)• Publicly available
CER_SSF_Terra-FMx_Edition2B CER_SFC_Terra-FMx-MODIS_Edition2C	<ul style="list-style-type: none">• Rev1 SW corrections available• 2/25/00 – 12/31/05 processed• Uses Edition2 instrument data as input• Edition2 Clouds; Final Terra ADMs• Publicly available
CER_SRBAVG_Terra-FMx-MODIS_Edition2D	<ul style="list-style-type: none">• Edition2C SFC are input• 3/00 - 5/04 processed• To extend through 12/05 in early 2007• Publicly available
CER_CRS_Terra-FMx_Edition2B CER_FSW_Terra-FMx-MODIS_Edition2C	<ul style="list-style-type: none">• 3/1/00 – 12/31/05 processed• Edition2B SSF input; latest version of SARB algorithms• Publicly available



Older CERES Science Data Sets

Still of Interest (Aqua)



Data Set	Comments
CER_BDS_Aqua-FMx_Edition2 CER_ES8_Aqua-FMx_Edition2 CER_ES4_Aqua-FMx_Edition2 CER_ES9_Aqua-FMx_Edition2 CER_BDS_Aqua-FM4_Ed2-NoSW CER_ES8_Aqua-FM4_Ed2-NoSW CER_ES4_Aqua-FM4_Ed2-NoSW CER_ES9_Aqua-FM4_Ed2-NoSW	<ul style="list-style-type: none"> • Rev1 SW corrections available • Aqua-FM4 Edition2 processed 6/18/02 – 3/29/05 • Aqua-FM4 Ed2-NoSW processed 3/30/05 - 12/31/05 • Aqua-FM3 Edition2 processed 6/18/02 – 12/31/05 • Edition2 publicly available • Ed2-NoSW available to Science Team
CER_ES4_FM1+FM3_Edition2 CER_ES4_FM1+FM4_Edition2 CER_ES9_FM1+FM3_Edition2 CER_ES9_FM1+FM4_Edition2	<ul style="list-style-type: none"> • Combine only Crosstrack instruments • Jul'02 – Dec'05 available • Uses Terra Edition2 and Aqua Edition2 inputs • Publicly available
CER_SSF_Aqua-FMx-MODIS_Edition2A CER_SFC_Aqua-FMx-MODIS_Edition2A	<ul style="list-style-type: none"> • Aqua ADMs • 7/02 - 12/31/05 processed • Publicly available



Documentation Overview

- Data Quality Summaries
 - Detailed information about a particular data set
 - Always consult Quality Summary prior to using data or publishing research
- Data Products Catalog
 - Parameter lists for each data product
 - Version of pages that apply to data set included with order
- Collection Guides
 - User Guide for data product
- Description/Abstract
 - Record of differences between data sets and configuration codes



Questions about Data Sets



- Look over Data Products Catalog pages
- Reread Data Quality Summary
- Consult Collection Guide, if available
- Specific science questions may be sent to Contact Scientist listed in Section 2.2 of Collection Guide or in Description/Abstract
- All other questions should be sent to User Services

larc@eos.nasa.gov

- For data products for which no Collection Guide or Description/Abstract is available, send all questions to User Services



Science Data Product URLs and Contacts

- Ordering Data
 - http://eosweb.larc.nasa.gov/HBDOCS/langley_web_tool.html
 - Can also order data from EOS Data Gateway
- Subsets of SSF, CRS, and ES8 are available
 - Order data using Java version of Langley Ordering Tool
 - Can subset by parameters or latitude/longitude box
- Contact Points
 - All questions regarding production data products and their use
 - E-mail: larc@eos.nasa.gov
 - Langley ASDC Customer Service
- CERES News (e-mail)
 - Subscribe from CERES datasets webpage
 - All new public datasets are announced as soon as available
 - Mechanism for distributing CERES information