

# CERES Data Management Activity

---

Presented to  
CERES Science Team

NASA Langley Research Center  
April 27, 2010

Jonathan Gleason  
[jonathan.l.gleason@nasa.gov](mailto:jonathan.l.gleason@nasa.gov)

# Topics to be Covered

---



- CERES Background/Overview
- CERES Processing Approach
- Edition 2 Status
- Edition 3 Overview
- FM5 Update
- Process Optimization Effort
- Development and Production Platforms

# CERES Statistics

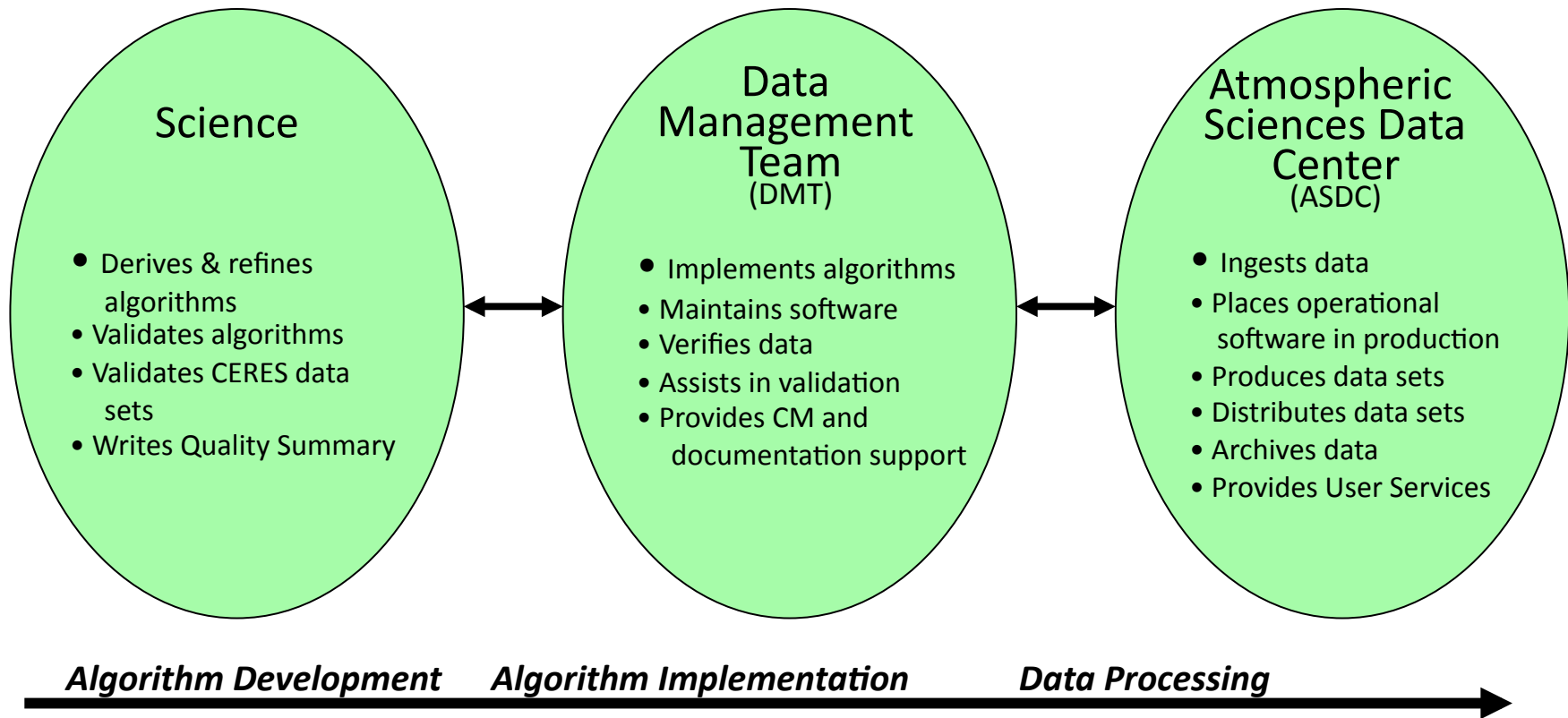
(from the Terra & Aqua Senior Review)

---



- High level of data fusion
  - 11 Instruments on 7 satellites
- 25 unique input data sources
- 18 CERES data products
- Over 90% of CERES data product data volume involves 2+ instruments
- Individual data products include up to 260 unique parameters
- Approximately 1.7 million lines of QC and validation codes
- Approximately 1.4 million lines of production codes

# CERES Organization



# CERES Subsystems

---



- CERES is made up of 7 Working Groups
  - Instrument                      - SOFA
  - ERBElite                        - SARB
  - Clouds                            - TISA
  - Inversion or ADM
- Code organized into 12 Subsystems
  - Each subsystem tied to 1 or more working groups
- Each Subsystem made up of 1 or more Product Generation Executives (PGEs)
  - Currently there are about 70 active PGEs

# Data from other Instruments used by CERES

---



- CERES Instrument/ERBELike only subsystems that can process when only CERES data available.
- CERES directly uses the following MODIS data sets:
  - MYD02SS1/MOD02SS1\* (19 channel radiance subset of every other pixel every other scanline)
  - MYD03/MOD03\* (geolocation)
  - MYD04\_L2/MOD04\_L2 (5 min 10 km aerosol swath)
  - MYD08\_D3/ MOD08\_D3 (daily 1 deg aerosol)
    - Critical data sets; must have matched pairs to process.
- Additionally CERES uses, Geostationary satellite data:
  - MET-5, MET-6, MET-7, MET-8, MET-9
  - GOES-8, GOES-9, GOES-10, GOES-11, GOES-12
  - GMS-5, MTSAT-1R

# CERES Processing Software



Subsystem Number	Subsystem Name	LOC (to nearest 1K)	Publicly Available Date Products	Product Frequency	Comments
	CERESlib	132K			All Satellites
1	Instrument/Pre-Processor	4K			NPP only
1	Instrument	110K	BDS	1/day	All Satellites
2	ERBElite/ Inversion	33K	ES-8	1/day	All Satellites
3	ERBElite/ TSA	16K	ES-9, ES-4	1/month	All Satellites
12	MOA	10K			Run monthly
4.1 – 4.4	Clouds	359K			All Satellites
4.5 – 4.6	Inversion	153K	SSF	1/hour	All Satellites
5	SARB	178K	CRS	1/hour	All Satellites
6 & 9	TISA-Gridding	43K	FSW, SFC, ISCCP-D2like-Day/Nit	60/month, 36/month, 1/month	All Satellites
11	GGEO	172K	ISCCP-D2like-GEO	1/month	Geostationary
7.2	Synoptic SARB	47K			All Satellites
7.1 & 8 10	TISA-Averaging	211K	SYN, AVG, ZAVG SRBAVG	1/day, 1/month, 1/month 5/month	All Satellites
	<b>TOTAL LOC</b>	<b>1,468K</b>			

# Current CERES Processing Approach



<b>Baseline1-QC</b>	<b>Edition1-CV</b>	<b>Edition2</b>
<ul style="list-style-type: none"><li>• Processed daily</li><li>• Run Instrument &amp; ERBElke Inversion subsystems</li><li>• Use Composite Snow Map</li><li>• Not publicly available</li></ul>	<ul style="list-style-type: none"><li>• Processed monthly</li><li>• Run Instrument and ERBElke subsystems</li><li>• Use actual Snow map and wait for all expected instrument inputs</li><li>• CV stands for “Calibration/Validation”</li><li>• Primary Instrument &amp; ERBElke products made publicly available</li></ul>	<ul style="list-style-type: none"><li>• Processed in blocks of 4+ months at a time</li><li>• Run all CERES subsystems as inputs become available</li><li>• All primary archival products made publicly available</li></ul>



# Edition 2 Processing



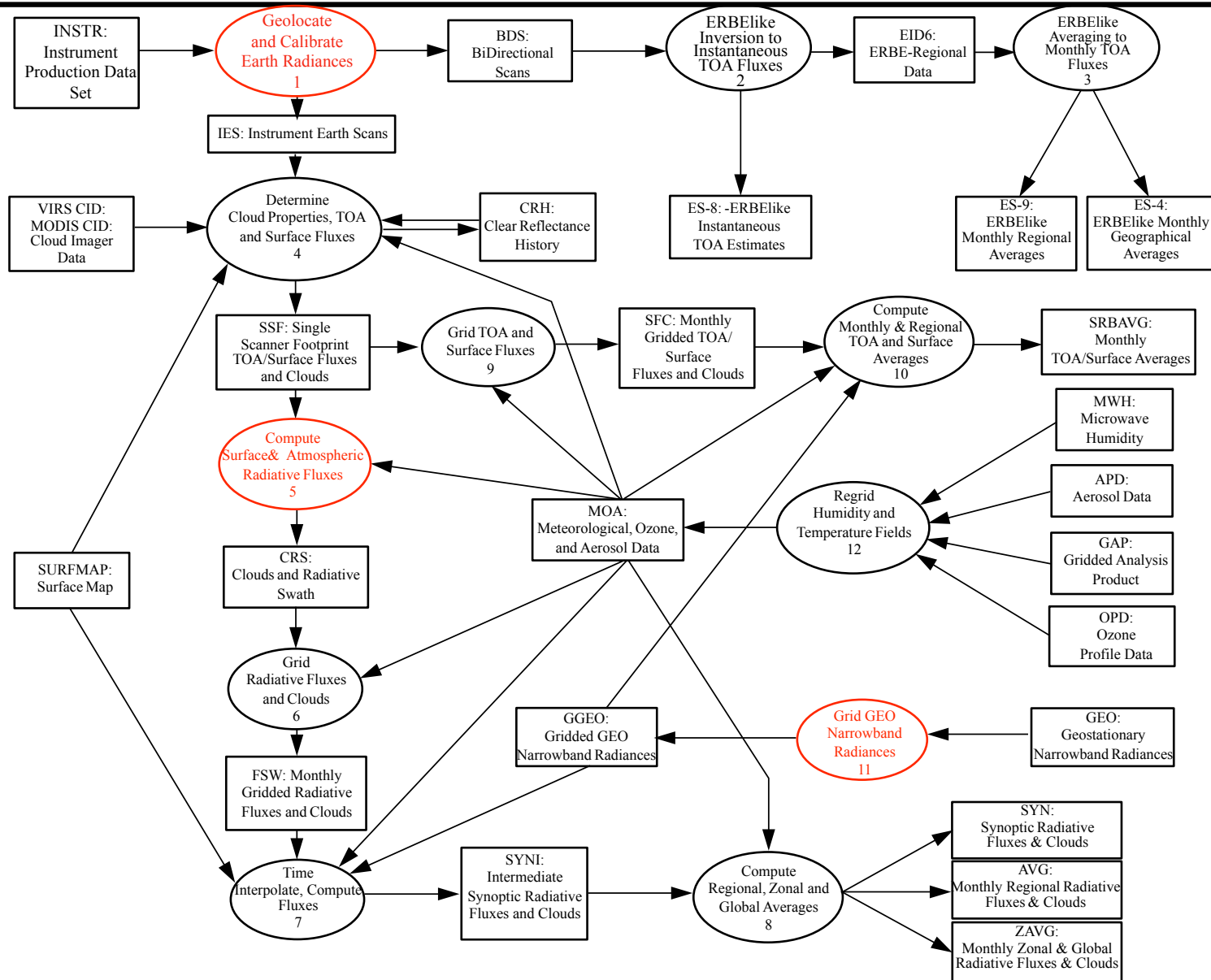
	Stage 1	Stage 2	Stage 3	Stage 4
<b>Latency</b>	6 – 24 months (Wait for Gains and SRF based on Edition1-CV)	~6 months after SFC (Wait for Cal Coef for geo-sat)	Wait for aerosol inputs (MATCH-like data)	Wait for GGEO and FSW availability
<b>Processing</b>	Instrument (BDS, IES)* ERBEl like Inv (ES8)* ERBEl like TSA (ES4, ES9) MOA (MOA) Clouds (Temp)* Inversion (SSF)* TISA-gridding (SFC)	GGEO (GGEO) TISA-Averaging (SRBAVG)	SARB (CRS)* TISA-Gridding (FSW)	TISA-averaging (TSIB) Synoptic SARB (SYNI) TISA-Averaging (SYN, AVG, & ZAVG)  * Instantaneous

# Main Terra & Aqua Edition2 Data Sets

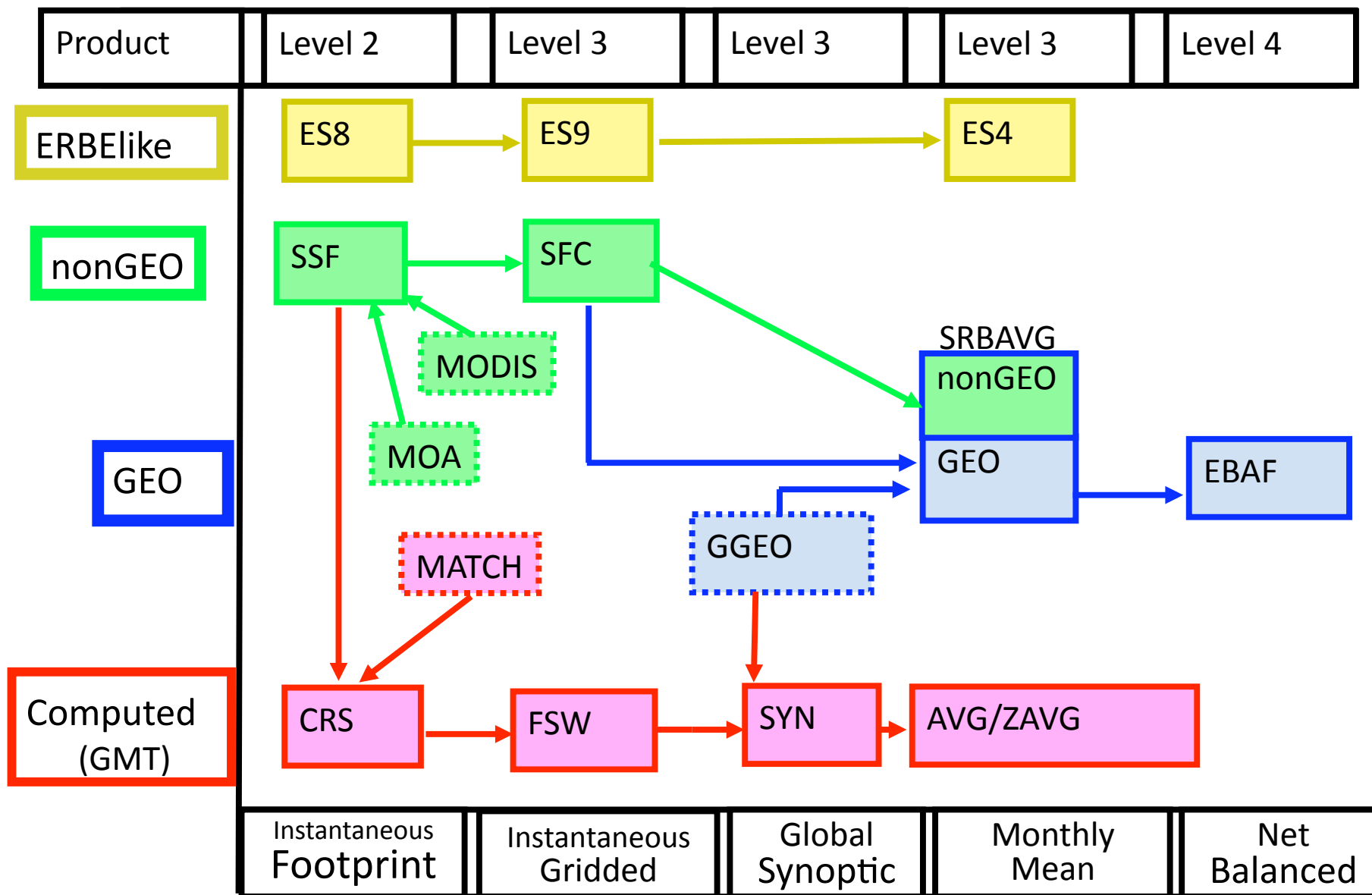


<b>Product</b>	<b>Latest Edition</b>	<b>Data available through</b>	<b>Comments</b>
BDS, ES8, ES4, ES9	Edition2 (T, A)	Sep'09	Process thru 3/1/10 by end of May
SSF	Edition2G (T) Edition2C (A)	Sep'09 Dec'07	Processing planned thru 2/2010 Aqua will be Edition2D
SFC	Edition2G (T) Edition2C (A)	Sep'09 Dec'07	“
SRBAVG	Edition2D (T) Edition2A (A)	Oct'05	Waiting on MTSAT coefficients
CRS	Edition2G (T) Edition2C (A)	Mar'09 Dec'07	Processing planned thru 2/2010, Need 2010 MATCH data Aqua will be Edition2D
FSW	Edition2G (T) Edition2C (A)	Mar'09 Dec'07	Processing planned thru 2/2010 Aqua will be Edition2D
SYN, AVG, ZAVG	Edition2C (T) Edition2B (A)	Oct'05	Waiting on MTSAT coefficients
ISCCP-D2like-Day, ISCCP-D2like-Nit	Beta1	Aug'07	Edition 2 in pre-op testing
ISCCP-D2like-Geo	Beta1	Oct'05	Edition 2 in pre-op testing

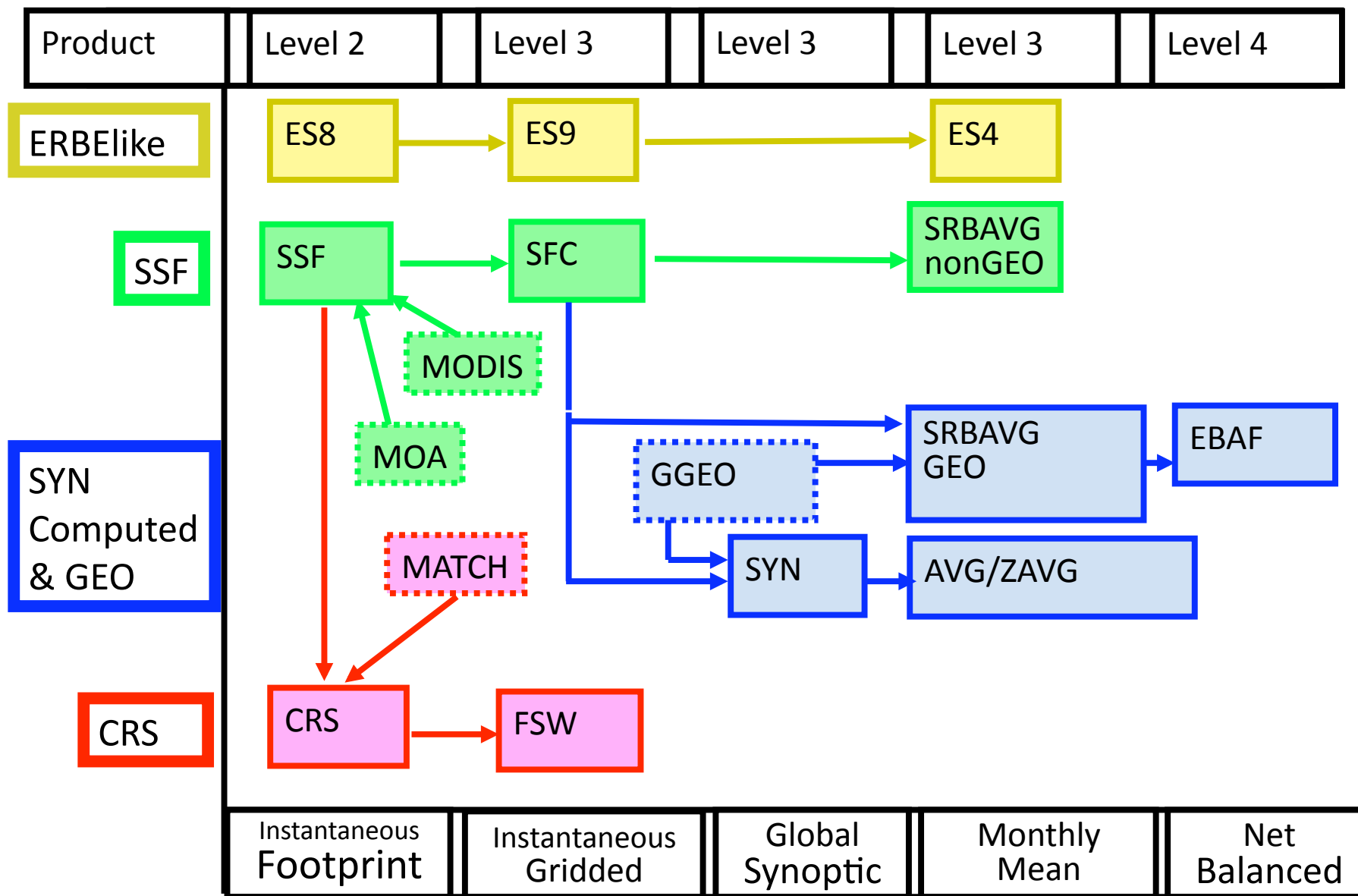
# CERES Edition 2 Processing Diagram



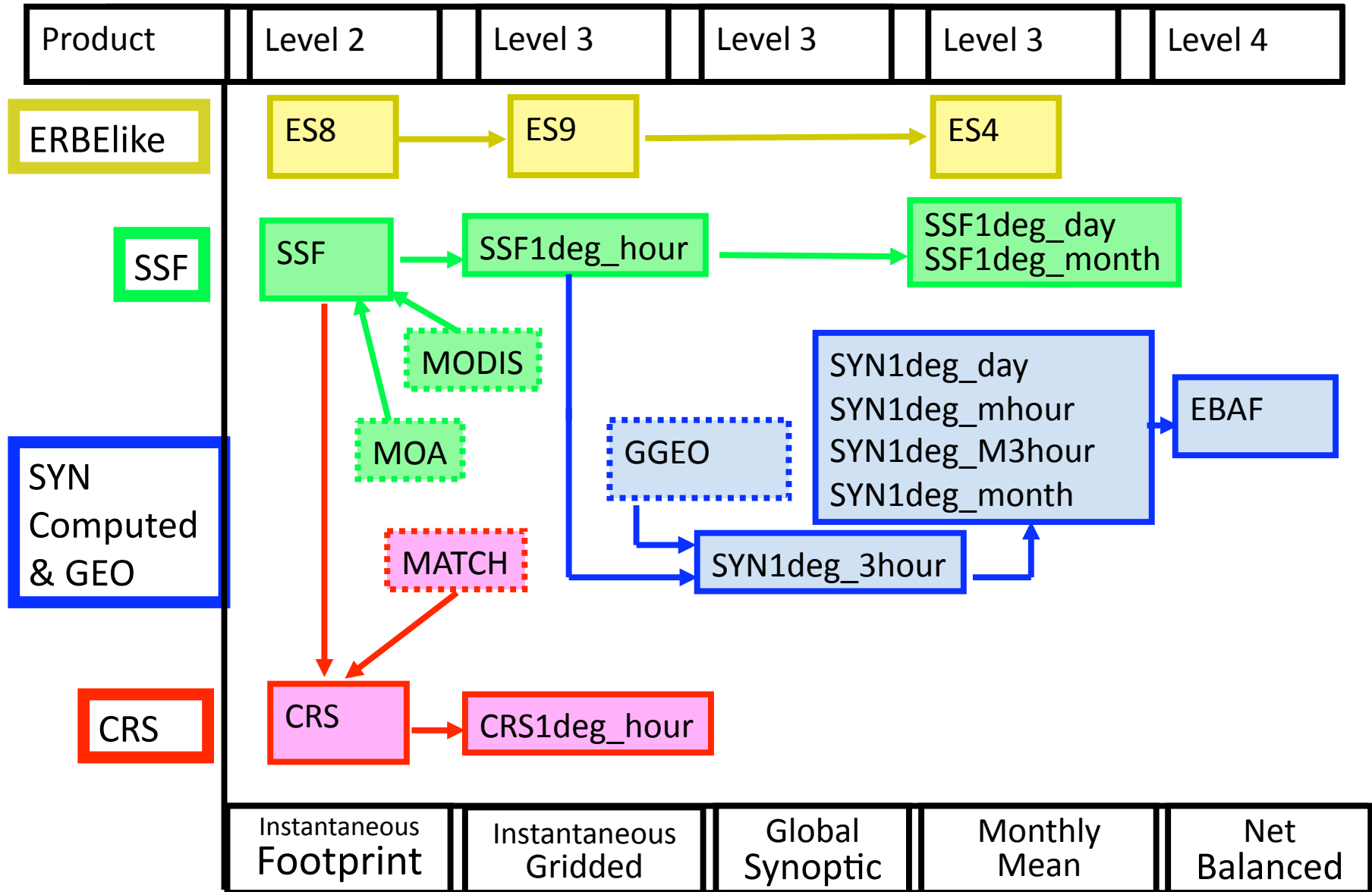
# CERES Edition 2 Processing Flow



# CERES Edition 3 Processing Flow



# CERES Edition 3 Product file name convection

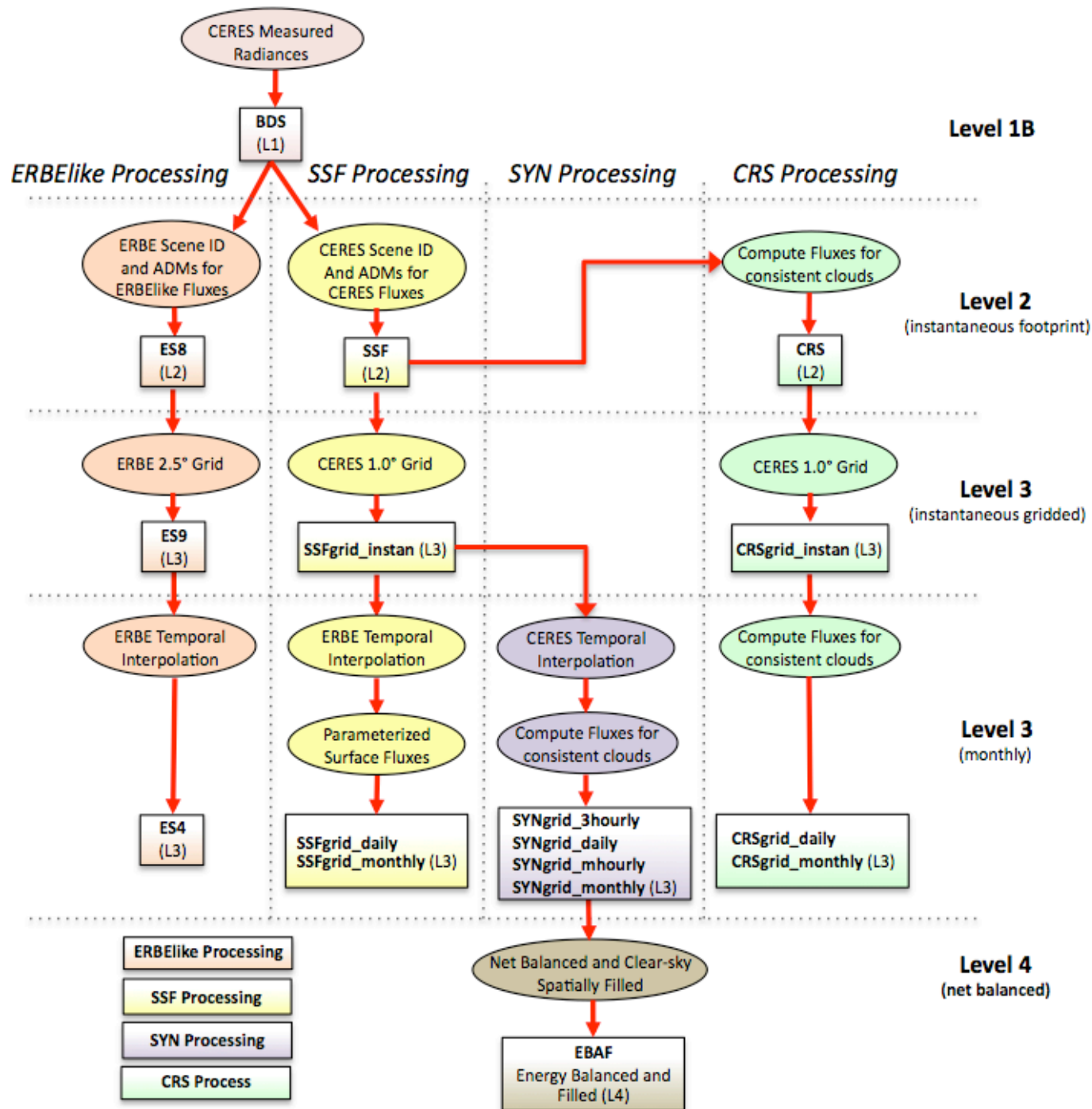


# Edition 3 Naming Convention



Edition 2 Product Name	Edition 3 Product Name
SFC	SSF1deg-Hour
SRBAVG (nonGEO)	SSF1deg-Day
	SSF1deg-Month
FSW	CRS1deg-Hour
SYN	SYN1deg-3Hour
AVG	SYN1deg-M3Hour
SRBAVG (GEO)	SYN1deg-Day
	SYN1deg-MHour
	SYN1deg-Month
AVG/ZAVG	SYN1deg-Month

\*No Change for Instrument or ERBElike





# Edition 3 Deliveries and Processing

---

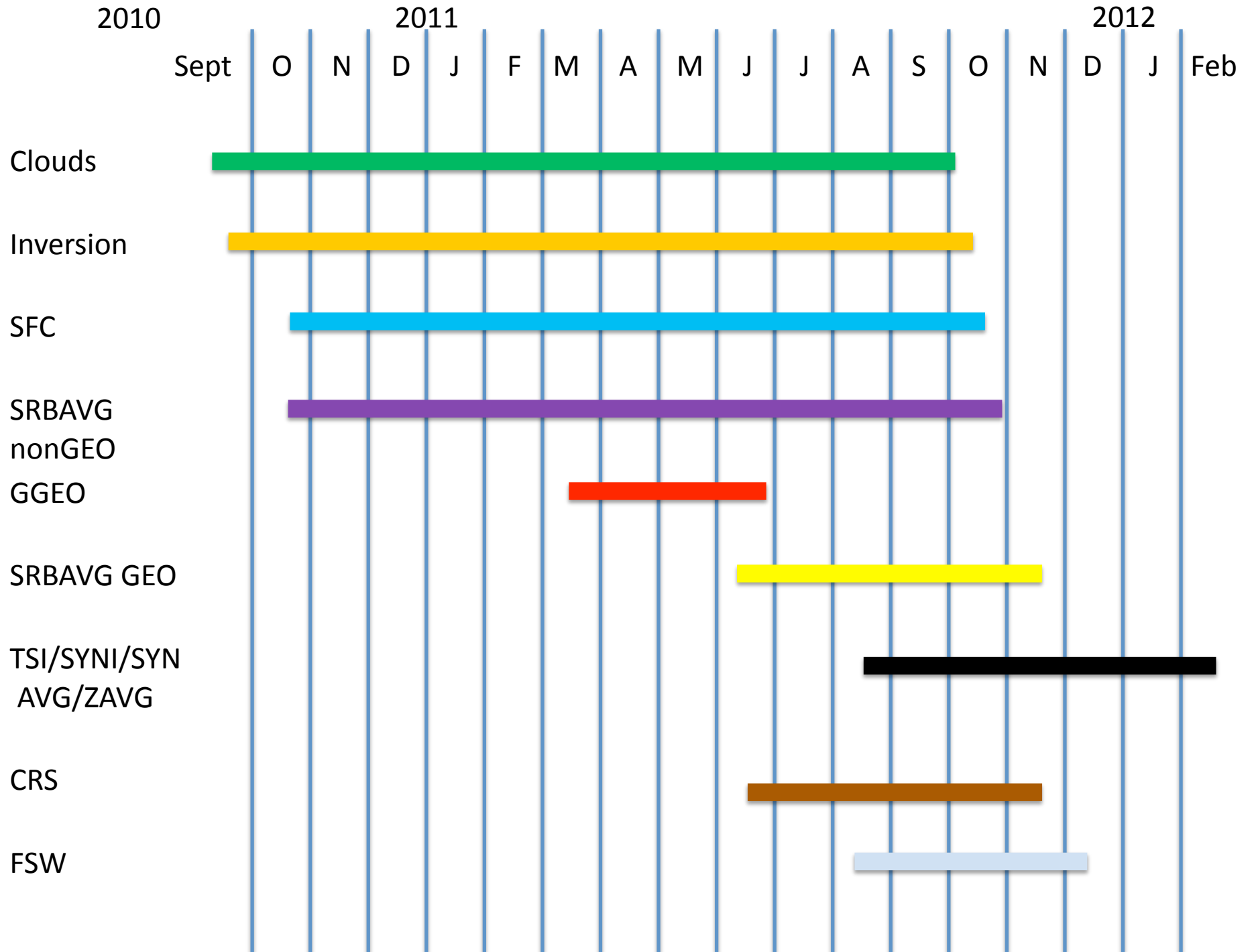
Where are we?

# Edition 3 Roadmap



Delivery	Beta	Edition 3	Production Ready	10 Years Data
Clouds	done	June 25, 2010	September 2010	October 2011
Inversion	done	July 23, 2010	September 2010	October 2011
SFC	May 21, 2010	August 2010	October 2010	October 2011
SRBAVG nonGEO	June 11, 2010	September 2010	October 2010	October 2011
GGEO	Oct 8, 2010	February 2011	March 2011	June 2011
SRBAVG GEO	Dec 3, 2010	April 2011	June 2011	November 2011
TSI/SYNI/SYN AVG/ ZAVG	Feb 2011	June 2011	August 2011	February 2012
CRS	Nov 12, 2010	May 2011	June 2011	November 2011
FSW	Jan 2011	July 2011	August 2011	December 2011

# Processing 10 years of Terra Edition 3



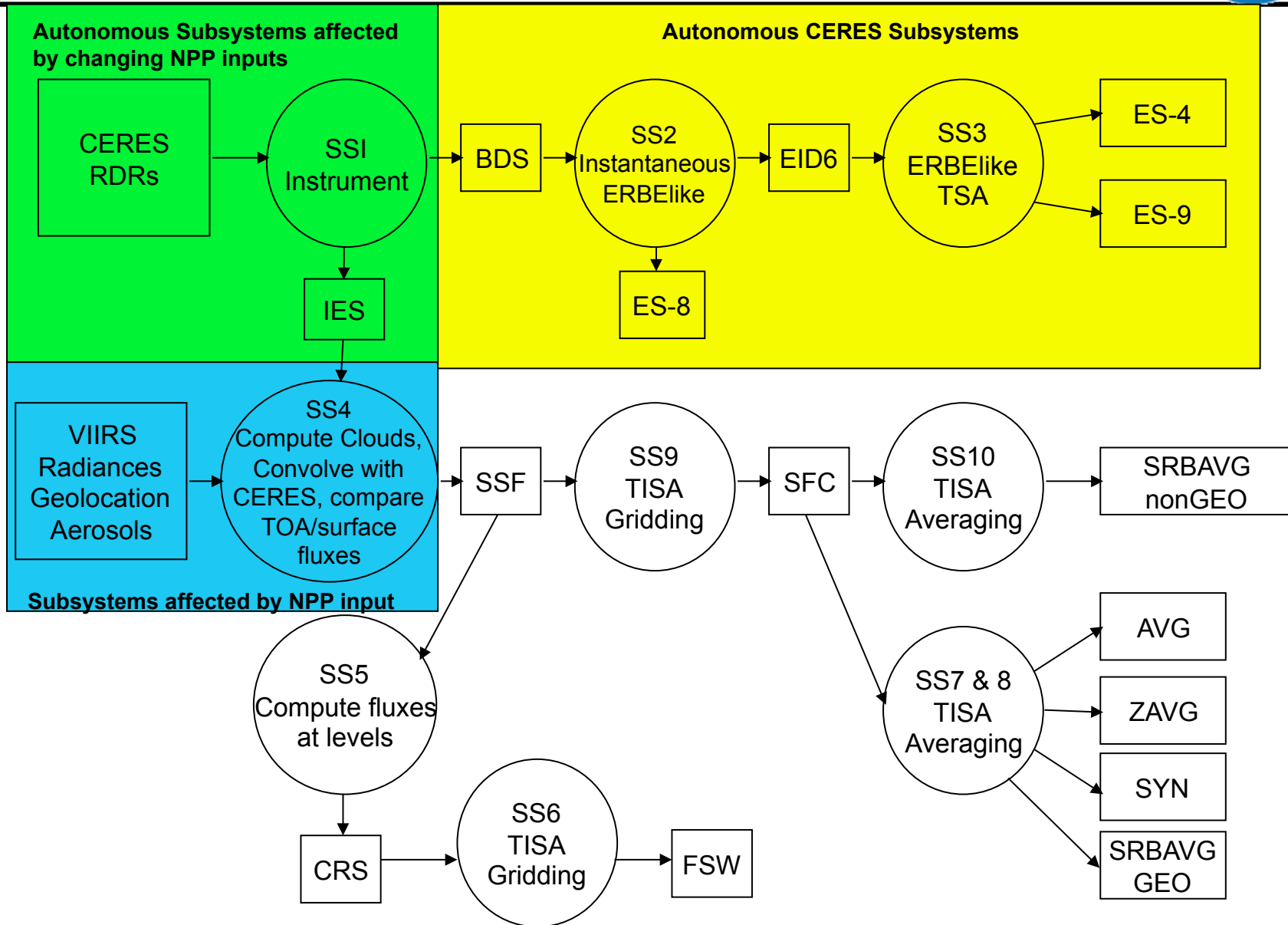
# Relevant NPP Issues

---



- Imager input data required for Climate Data Record (CDR) must be of climate quality and consistently calibrated over entire period.
  - In NPP era, Land PEATE provides CERES aggregated radiance and geolocation files and sub-sampled data files using CERES provided code. Land PEATE also provides AOT files that correspond to sub-sampled radiance/geolocation.
  - For Terra/Aqua, MODAPS provides radiance, geolocation, and aerosol files from a collection that begins at covers open.
- **NPP CERES made use of already existing interfaces.**
  - Cost savings by using existing infrastructure.
  - Land PEATE already getting VIIRS data. Agreed to also obtain CERES RDRs.
  - Network between Land PEATE and ASDC exists for Terra/Aqua.

# Simplified CERES Processing Flow



# FM5 Code Development

---



- Instrument only subsystem modified
  - Preprocessor will convert data to format consistent with Terra and Aqua
  - Instrument subsystem currently implemented with Ada
  - Convert Ada to C++ and deliver preprocessor for Ada as schedule risk mitigation
- Seven total code deliveries
  - Deliver C++ FM5 (AMI P6)
  - Deliver Ada FM5 (AMI x86) (Risk Mitigation)

# Process Optimization Effort

---



- Proposed in 2008 CERES Senior Review
  - Migrate off SGI to AMI (DPO)
  - Optimize CERES production software via parallelization
  - Communicate Production Requests
- Production Request Database
  - Automate PRs (Currently via .pdf)
  - Run scripts can read required input
  - Online for real time review by subsystem teams

# AMI Script Development

---



- DMT developing production scripts for each CERES PGE
  - Input file checking (currently manual)
  - Sun Grid Engine job submittal
  - Environment setup
- Complete set required for each PGE delivered to AMI – deliveries on hold
- First draft in review by ASDC SSIT



# Production Systems

---



- Unexpected challenges with AMI system
  - Originally experienced poor performance – degraded state / instability
  - Unexpected scripting effort required by DMT
- No PGEs currently in production on AMI
  - Generally speaking all Edition 3 code still targeted for AMI
- Most Edition 2 code expected to remain on Magneto for near term future



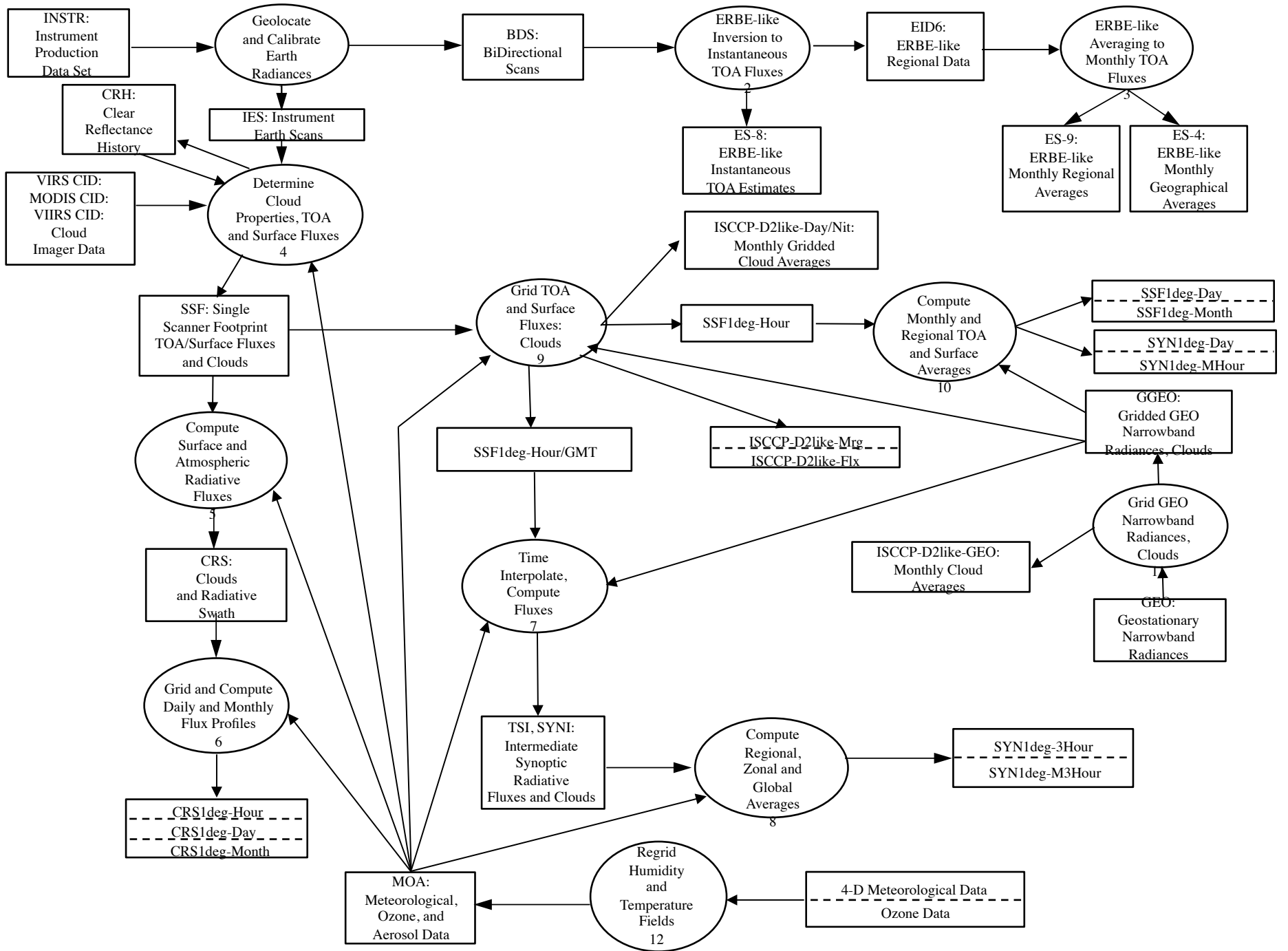
# Science Data Product URLs and Contacts

---

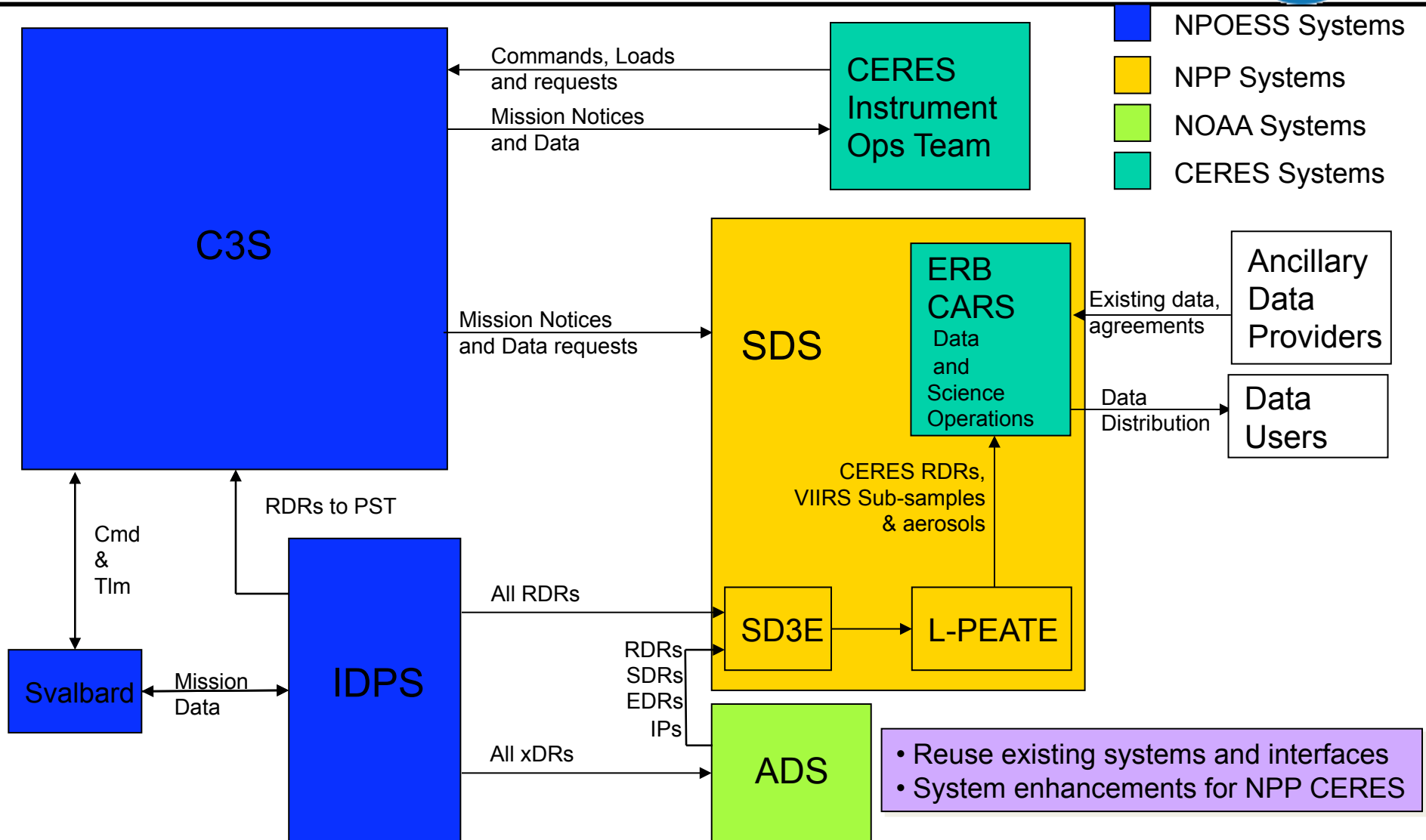
- Ordering Data
  - <http://www-pm.larc.nasa.gov/SATGIF1/ceres-ordering-tool/CERESExample/index.php>
  - [http://eosweb.larc.nasa.gov/HBDOCS/langley\\_web\\_tool.html](http://eosweb.larc.nasa.gov/HBDOCS/langley_web_tool.html)
- Subsets of SSF, CRS, and ES8 are available
  - Order data using Java version 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](mailto:larc@eos.nasa.gov)
    - Langley ASDC Customer Service
- CERES News (e-mail)
  - Subscribe from CERES Data Products webpage
  - All new public datasets are announced soon after public release
  - Mechanism for distributing CERES information



# Questions and Comments



# NPP CERES Operational Data Flow



# Delivery Process

---



- Data Management personnel create expected output on target production machine
  - Verify that expected output looks as expected on target machine
- CM untars delivery and compiles source code on target machine, runs to reproduce expected output
  - Verify that delivery tar file includes all necessary components prior to turning delivery over to ASDC for testing
- Once delivery is in production do not immediately begin running an Edition data set
  - Run ValRx for all instrument/input combinations
    - Ensure production environment not altering output
    - Ensure correct files were delivered
    - Ensure scripts set up correctly

# Edition 3 Roadmap



Delivery	Beta	Edition 3	Production Ready	10 Years Data
Clouds	done	June 25, 2010	September 2010	February 2011
Inversion	done	July 23, 2010	September 2010	February 2011
SFC	May 21, 2010	August 2010	October 2010	March 2011
SRBAVG nonGEO	June 11, 2010	September 2010	October 2010	April 2011
GGEO	Oct 8, 2010	March 2011	April 2011	July 2011
SRBAVG GEO	Dec 3, 2010	April 2011	June 2011	August 2011
TSI/SYNI/SYN AVG/ ZAVG	Feb 2011	June 2011	August 2011	February 2012
CRS	Nov 12, 2010	May 2011	June 2011	October 2011
FSW	Jan 2011	July 2011	August 2011	November 2011

# Edition 3 Roadmap



Delivery	Beta	Edition 3	Production Ready	10 Years Data
Clouds	done	June 25, 2010	September 2010	September 2011
Inversion	done	July 23, 2010	September 2010	September 2011
SFC	May 21, 2010	August 2010	October 2010	September 2011
SRBAVG nonGEO	June 11, 2010	September 2010	October 2010	October 2011
GGEO	Oct 8, 2010	February 2011	March 2011	June 2011
SRBAVG GEO	Dec 3, 2010	April 2011	June 2011	October 2011
TSI/SYNI/SYN AVG/ ZAVG	Feb 2011	June 2011	August 2011	February 2012
CRS	Nov 12, 2010	May 2011	June 2011	October 2011
FSW	Jan 2011	July 2011	August 2011	November 2011



# Processing 10 years of Terra Edition 3

