

CERES Systems Engineering Committee

Members: Tammy Ayers, SAIC
Vertley Hopson, DAAC
Sandy Nolan, SAIC
Sue Sorlie, DAAC

Charter: Serve as a forum for resolving issues which affect more than one working group. Report to CERES Data Management Team

February 08, 2001 10:00 am

Erika Geier, Walt Miller, John Robbins, and Beth Flug joined the SE committee for this meeting.

The designation, logging, reporting and disposition of successful and failed CERES jobs at the DAAC was discussed. Sue explained the processing procedures followed by the DAAC for non-APGS jobs, which include the following:

1. A PGE is promoted to production and a production request is received
2. DAAC operators receive log sheets for requested jobs and as jobs are run, they enter job numbers and other information in the log sheets
3. If a PGE fails with a non-zero exit code, the job is marked failed on log sheet and the exit code is also recorded. No other action is taken. The created data and log files are deleted.
4. If the PGE completes with a zero exit code, the epilogue is run, the required data (as described in the "CERES File Management Policy at the DAAC") is sent to the DAAC archive and /QA, and email reporting these actions is sent.
5. The /QA sweeper runs to remove oldest data on QA disk. Sue also will delete oldest data to make room on the disk as it is required.
6. The "Product Availability Status" chart is no longer available. It has been replaced by the 'Data Reports Page' at URL: http://charm.larc.nasa.gov/data_reports. (This page is still under development. Information found in these data reports is obtained by querying the DAAC archive.)

John said that he would like to see a Web site that would have a processing chart for all CERES subsystems (PGEs). The chart would show the status of all jobs (data dates) that had been requested for that PGE. A user could click on the job/data date and receive information as soon as it is available about the job, including the scheduled processing date, the actual processing date, and information on the success or failure of the job. If the job was successful, then the data file name and archival date should be provided. If a job was unsuccessful, then a reason would be provided for the job failure. Under ERBE, a processing data base was maintained and processing results were available each morning in a customized report. Similar reports were needed for CERES ERBE-like processing. Beth has built a processing data base that is populated with information from the metadata files identified by the production email that is sent out when CERES ERBE-like and SSF data are archived. (There was no DAAC maintained database from which this same information could be obtained). Using this CERES processing database, a "CERES Production Process Tracking" Web site is being developed. Beth gave a demonstration of the information and reports that are currently available at a test Web site..

Walt asked about the previous requirement that the log files, data, and metadata from all failed jobs be tarred up and made available to the CERES IT. Vertley explained that there was not enough disk space to meet that requirement and continue with production runs. Walt added that often by the time a failed Cloud job is identified by the CERES IT, all data, metadata and log files for the failed run had been deleted and were unrecoverable. The Cloud Subsystem scripts are written to send email to the Cloud Working Group when a production job fails, but only a small percentage of these emails are received.

CERES Systems Engineering Committee

Vertley and Walt will both look into why this email is only sent sporadically.

Erika pointed out that the procedures for reporting and examining individual failed jobs, may need to be different from the procedures for a mass of failed jobs. Scientist often want to know why data is available in one version of a data set and not in another. There are several reason that an hourly SSF file could be missing:

1. No CERES data - will not exist in any version of the data set
2. Ancillary data missing at the time of processing
3. Input data set changed (example- different version of VIRS data used)

Erika suggested that one person be assigned to be the Point-of-Contact (POC) /Information Consolidator for each product. The POC would be responsible for collecting and disseminating the reason that data does not exist. Erika said that how this was done would be up to the POC and responsible working group. Vertley pointed out that a common format for all products would be easier to use and maintain. A suggestion was made that in addition to the data product POCs, that there also be one Information Consolidator for all of CERES.

There was a discussion on exit codes and Cloud Subsystem processing. The DAAC only archives products that are produced from runs with an exit code of 0. If the Convolution Subsystem produces an interim SSF with a QA flag equal to "Failed", then the exit code is 0 and the required Cloud files are archived. In this case, the Inversion PGE is run, but fails because of bad input. The only way for the DAAC to know not to start the Inversion job, would be to look at the meta-data in the .met file. If the exit code on the Cloud job was changed to non-zero, then the "good" Cloud data would not be archived. When there is no IES input, then the Cloud PGE is not run. When the IES Quality Flag is set to "Failed", then the Cloud PGE exits with a 202 and the "good" Cloud data are not archived. Erika said that at this time, there is no reason to go back and run the Cloud PGE for the missing hours because the jobs which would use the corrected CRH have already been run.

As a result of this discussion, it was determined that all CERES subsystems need to be informed that if they want information about failed production jobs, then their PGE scripts need to send email containing any necessary information. The subsystems may also be able to negotiate with the DAAC to have small log files and metadata files moved to the /QA disk when jobs fail.

Sue asked about the ERBE-like overlap file names for PGEs CER2.3P1 and CER2.3P2 in the CERES File Management Policy. For the month of January 2000, the overlap days would be December 31, 1999 and February 01, 2000. Erika asked about the ERBE-like monthly combined product for December 2000 FM1 and FM2 and emphasized that it needs to be run along with any other missing data before the promotion of the next Instrument Subsystem delivery. After some discussion, it was decide that the Instrument Subsystem delivery, as long as it was not recom-piled, could be promoted independent of the when the Toolkit upgrade on samantha occurred. Sue or Vertley will email Tammy as soon as they have a definite date for the Toolkit upgrade on samantha.

The /ENG/CERES partition on samantha is currently over 90% full. Tammy will send out email to all subsystems asking them to clean up their /ENG/CERES directories.

CERES Systems Engineering Committee

Tammy distributed copies of the updated CERES Delivery Memo template, the original template and an example of a completed Delivery Memo using the new format. The SE committee is to review the template and send comments to Tammy.

The SE committee will review the CERES Operations Agreement at the next SE meeting, on Thursday, February 15, 2001 at 9:00 am in the SAIC 3rd floor conference room.

Meeting adjourned at 11:50 am. skn