NASA Technical Memorandum XXXX, Volume 2

DRAFT

MODIS Technical Report Series

Volume 2, MODIS PGS Data Processing Operations Concept

Ed Masuoka, Al Fleig, J.J. Pan, Tom Goff, John Crocker
Lloyd Carpenter, Philip Ardanuy, Carl Solomon, Jim Storey

October 1993

NASA Technical Memorandum XXXX, Volume 2

DRAFT

MODIS PGS Data Processing Operations Concept

Ed Masuoka
Goddard Space Flight Center
Greenbelt, Maryland

Al Fleig
University of Maryland
College Park, Maryland

J.J. Pan
Tom Goff
John Crocker
Lloyd Carpenter
Philip Ardanuy
Research and Data Systems Corporation
Greenbelt Maryland

Carl Solomon
Jim Storey
Hughes STX
Lanham, Maryland

MODIS Data Processing Operations Concept

Table Of Contents

1.0 INTRODUCTION 4
- 1.1 Purpose 4
- 1.2 Scope 4
- 1.3 MODIS Overview 4

2.0 RELATED DOCUMENTATION 6

3.0 MODIS DATA PROCESSING SOFTWARE (MDPS) CONCEPT 7
- 3.1 Terminology 7
- 3.2 Interface Environments 7
  - 3.2.1 Team Leader Computing Facility (TLCF) Environment 8
  - 3.2.2 Science Computing Facility (SCF) Environment 8
  - 3.2.3 Product Generation System (PGS) Environment 8
- 3.3 Goals and Objectives 8
- 3.4 Description 9
- 3.5 Policies 9
- 3.6 Assumptions 10

4.0 EXTERNAL INTERFACES 11
- 4.1 PGS Scheduler Interactions 11
- 4.2 IMS Work Orders 11
- 4.3 TLCF Interactions 11
- 4.4 SCF QA 12

5.0 CAPABILITIES AND CHARACTERISTICS 12
- 5.1 Software Development Environment 12
  - 5.1.1 Science Software Task 13
    - 5.1.1.1 Algorithm Dependency Diagram Generation 13
    - 5.1.1.2 Science Software Delivery and Test 13
    - 5.1.1.3 Science Software Change 14
  - 5.1.2 MODIS Script Development and Test 14
    - 5.1.2.1 MODIS Script Development 14
    - 5.1.2.2 MODIS Script Test 15
- 5.2 Operational Environment 15
  - 5.2.1 Operational Requirements 15
  - 5.2.2 Operational Requirements for Level 1A 15
  - 5.2.3 Operational Requirements for Level 1B 15
  - 5.2.4 Minimization of Data I/O Redundancy in the MDPS 15
- 5.3 Processing Capabilities 16
  - 5.3.1 MODIS Level 1A Processing Capabilities 16
  - 5.3.2 MODIS Level 1B Processing Capabilities 17
  - 5.3.3 MODIS Level 2 Processing Capabilities 17
- 5.4 Data Processing 18
  - 5.4.1 Input Data Control 19
  - 5.4.2 PGE Execution 19
  - 5.4.3 Processing Mode 19
    - 5.4.3.1 Standard (Routine) Processing 20
    - 5.4.3.2 Quicklook (Near-Real-Time) Processing 20
    - 5.4.3.3 Reprocessing 21

Table Of Contents
(continued)

5.5 Error Handling/Logs 21

6.0 SAMPLE OPERATIONAL SCENARIOS 22
6.1 Standard MODIS QA Scenario 22
6.2 Standard MODIS Processing Scenario 23
6.3 Quicklook (Near-Real-Time Processing) Scenario 26
6.4 Reprocessing Scenario 27
6.5 Special Request Processing Scenario 28
6.6 System Test Scenario 29

7.0 ACRONYMS AND ABBREVIATIONS 30

8.0 GLOSSARY 32

MODIS Data Processing Operations Concept

1.0 INTRODUCTION

Moderate Resolution Imaging Spectroradiometer (MODIS) Data Processing includes MODIS Level 1A, Level 1B, Level 2, and Level 3 data processing. The MODIS Data Processing System (MDPS) is designed, developed, tested, and maintained by the MODIS Science Data Support Team (SDST) at the Team Leader Computing Facility (TLCF). The MDPS is delivered to the Product Generation System (PGS), part of the Earth Observing System (EOS) Data and Information System (EOSDIS) Core System (ECS). The PGS is responsible for the generation, and management of standard MODIS data products. MODIS Level 2 data products are generated by science software provided by the MODIS Science Team (a team of selected Earth scientists from several universities and NASA). These scientists have formed atmospheric, ocean, and land discipline groups. The SDST is responsible for integrating the Level 2 science code developed by the scientists into the MDPS.

1.1 Purpose

The purpose of the MODIS Data Processing Operations Concept is to address the tasks, operations concepts, and processing scenarios involved in the MODIS Data Processing. The operations concept of the MODIS Data Processing provides information on MODIS requirements to the ECS.

1.2 Scope

The MODIS Data Processing operations concept addresses the development and testing of Level 1 software, the integration and testing of Level 2 and Level 3 software received from the MODIS science team, the execution of software for the generation of MODIS data products, the operation and control of various processing conditions, and the interfaces with the PGS, the TLCF, and Science Computing Facilities (SCFs). In addition, this document addresses the assumptions required for the MODIS data processing and concerns related to processing efficiency.

1.3 MODIS Overview

The MODIS objective is to make long-term observations to improve our understanding of the global dynamics and processes occurring on the land surface, ocean surface layer, and lower atmosphere (including surface-atmosphere interactions) by exploiting the visible, near-infrared (IR), and thermal-IR spectrum with observation resolutions of $\frac{1}{2}$ to 2 days and 250 m to 1 km at nadir.

The MODIS measurement objectives include surface temperature (land and ocean), ocean color (sediment, phytoplankton), global vegetation maps, global change (deforestation and desertification), cloud characteristics, aerosol concentrations and properties, atmospheric temperature and moisture structure, snow and ice cover characteristics, and ocean currents. Additional measurement objectives include chlorophyll concentration, primary productivity, sediment transport, standing water, wetland extent, vegetation properties, hemispherical albedo, bi-directional reflectance, cloud properties, and aerosol radiances.

The land objective is to collect data for studies of the spatial and temporal variability in land surface properties (e.g., surface temperature, primary production, evapotranspiration and photosynthesis, vegetation cover and phenology, snow and ice, radiative properties including radiation balance) with emphasis on problems such as desertification, regional vegetation stress due to acid rain or drought, and succession or change in vegetation species due to deforestation and anthropogenic effects.

The ocean objective is to collect data for studies of the spatial and temporal variability of ocean surface thermal and bio-optical properties (e.g., water-leaving radiances, photosynthetic pigments, sea surface temperature, flow visualization, attenuation coefficients and sea ice), with special emphasis on ocean primary productivity.

The atmosphere objective is to collect data for studies of tropospheric dynamics, climatology, and chemistry, as obtained through observations of cloud characteristics (height, type, albedo, optical depth, effective droplet radius, and thermodynamic phase), aerosol properties, water vapor, and temperature.

MODIS will observe nearly the whole Earth twice a day for at least 15 years in two series with three spacecraft in each series. These spacecraft will be flown in 705 km circular, sun-synchronous orbits with 10:30 a.m. descending node equator crossings for one series, and 1:30 p.m. ascending node equator crossings for the other series. The first launch will take place in June 1998. MODIS data products, beginning at Level 1A, are required by the members of the MODIS science team, members of the other EOS instrument teams, the EOS interdisciplinary investigators, and the scientific community at large. The primary data systems include the MODIS instrument and its processor (from sensor signals to data packets), the EOS platform and the Tracking and Data Relay Satellite System (TDRSS) (data transmission from the instrument through the platform and TDRSS, to the ground), and the ground processing system including the EDOS and EOSDIS.

2.0 RELATED DOCUMENTATION

The following is a list of references pertinent to the contents of this document:

1. Earth Observing System (EOS) Data and Information System (EOSDIS) Product Generation System (PGS) Operations Concept (preliminary), March 1993.
2. The Product Generation System Toolkit Study Report, Ver. 1.9a, April 1993.
3. MODIS Software and Data Management Plan in press.
4. MODIS Science Computing Facilities (SCF) Plan, in press, E. Masuoka and A. Fleig.
5. MODIS Science Data Support team (SDST) Coding Recommendations for the MODIS Science team (draft), dated April 10, 1992.
6. MODIS Level 1A Software Baseline Requirements (Draft), July 1993, E. Masuoka, A. Fleig, T. Goff, P. Ardanuy, L. Carpenter, C. Solomon and J. Storey.
7. MODIS Information, Data, and Control System (MIDACS) Operations Concepts, December 1988.
8. Preliminary Draft, EOSDIS Core System (ECS), PGS Toolkit Requirement Specification, 29 January 1993.
9. Science Data Plan for the EOS Data and Information System covering EOSDIS Version 0 and Beyond, Version 2, June 1993.
10. Interface Requirements Document Between EOSDIS Core System (ECS) and Science Computing Facilities, August 1993.
11. EOSDIS Core System (ECS) Requirements Specification, Functional and Performance Requirements for the ECS Project, 12 August 1993.
12. Interface Requirements Document between EOSDIS Core System (ECS) and Science Computing Facilities, August 1993.
13. ECS Operations Concept Document for the ECS Project, August 1993.
14. MODIS Level 1B Requirements Report, in preparation.

3.0 MODIS DATA PROCESSING SOFTWARE (MDPS) CONCEPT

The MODIS Data Processing Software (MDPS) concept introduces the terminology, the goals and objectives of the system, the description of the system functions and interfaces, and the policies for using the system.

3.1 Terminology

The following terminology is used to describe the MODIS Data Processing operations concept:

MODIS science team refers to the selected group of scientists who are members of the MODIS atmosphere, land, and oceans discipline groups and their co-investigators. They are responsible for development of the science software which generates MODIS data products.

Science software refers to the complete package that includes source code, build instruction, documentation, test data, and related information.

Product Generation Executable (PGE) refers to an executable entity that is re-built on the TLCF and the PGS using source code and build instructions from the science software. A PGE produces one or more standard products.

Production script refers to a script file which manipulates the execution of PGE's. The production script can be developed for the generation of a single product, a stream of products or a group of products generated in parallel. Some additional commands (e.g., staging request, resource checking, memory refreshing, etc.) are included in the production script.

MODIS script refers to a production script which logically integrates the scripts for MODIS Level 1A, 1B, 2, and 3 data processing. **MODIS Level 1A script** is the script for Level 1A data processing, **MODIS Level 1B script** is the script for Level 1B data processing, and **MODIS Level 2 script** is the script for all Level 2 data processing. The MODIS 2 script includes the **MODIS Level 2 Ocean script** for Ocean products and the **MODIS Level 2 Atmosphere/Land script** for both atmospheric and land products. **MODIS Level 3 script** is for Level 3 data processing. All of these scripts can also be referred as MODIS scripts.

Dependency tables refers to tables or diagrams which indicate the interdependency of MODIS Levels 1 to 3 algorithms. The dependency tables provide information on all input data required to execute a specified PGE. Using the dependency tables, the PGS scheduler can check all required data sources before starting the execution of MODIS scripts.

3.2 Interface Environments

The MDPS staff is responsible for the execution of science software used to generate MODIS data products. The MDPS is decomposed into Level 1A , Level 1B, and Level 2 processing subsystems. The design, development, and test of the MDPS are performed in the Team Leader Computing Facility (TLCF) environment. Then it is delivered to the PGS for final testing and production processing. The science software is developed by the science team members in their own Science Computing Facilities (SCFs). Figure 1 shows the operational relationship between the PGS, MODIS SCFs, and MODIS TLCF.

3.2.1 Team Leader Computing Facility (TLCF) Environment

The purpose of the TLCF is for the development of Level 1A software and the integration of Level 2 and Level 3 science code into a PGS environment.

The MODIS TLCF is designed to provide the required computer support for the team leader. The TLCF supports hardware, CASE tools, code checkers and optimization tools for MODIS data processing. The TLCF also supports generation of simulated data, prototyping, development of test cases, testing at all levels, software optimization, and configuration management. In the post-launch era, the TLCF will continue the previous support in addition to supporting algorithm updating, refinement or replacement, the generation of special products, and the quality assurance of products.

3.2.2 Science Computing Facility (SCF) Environment

The MODIS science team members are responsible for the software to produce MODIS standard products, the production of special data products (research products generated for a subset of the available data and not scheduled for routine production on the PGS), and the validation of standard and special data products on the SCF. The SDST supports the Science Team by integrating the software into a package that runs efficiently on the PGS. In general, the SCFs are located at the team members' research sites.

3.2.3 Product Generation System (PGS) Environment

The PGS is responsible for processing the MODIS standard data products. The maintenance of operational software requires frequent communication and cooperation between the staffs of the PGS and the SDST. The MODIS Science Data Support Team (SDST) aids MODIS science team members in making their science software comply with EOSDIS software standards.

The PGS is a production-oriented computing facility which houses and executes the MDPS, and software for other instruments, to generate standard data products from Level 1 through Level 3. In addition, the PGS is responsible for event- and data-driven scheduling of jobs, providing control information and generating associated reports and MODIS logs, reprocessing data, providing a toolkit that meets requirements for data production, and verifying ancillary data. The PGS also delivers plots and reports to the SCFs after the execution of PGEs. While the development, modification and testing of MODIS science software does not take place at the PGS, cooperation between the PGS staff, the MODIS SDST, and the MODIS Science Team members are required to ensure that production proceeds smoothly.

3.3 Goals and Objectives

The major goal of the MDPS is to generate valid MODIS science data products. The major objectives of the MDPS include:

- generating the MODIS valid data products;
- handling special and reprocessing requests;
- performing limited product quality assurance.

3.4 Description

The MODIS SDST provides a MODIS script that includes five production scripts:

1. The MODIS Level 1A script: manages the generation of the MODIS Level 1A product.
2. The MODIS Level 1B script: manages the generation of the MODIS Level 1B product.
3. The MODIS Level 2 Atmosphere and Land script: manages the generation of the MODIS Level 2 Atmosphere and Land products.
4. The MODIS Level 2 Ocean script: manages the generation of the MODIS Level 2 Ocean products.
5. The MODIS Level 3 script: manages the generation of the MODIS Level 3 products.

The PGS starts the MODIS data processing by executing the MODIS script when all required resources are available. MODIS science team members analyze the data products, plots, reports, and MODIS logs using interactive quality assurance tools. Figure 2 shows the overview of the MDPS.

The Information Management System (IMS) sends the product orders to the PGS Scheduling Subsystem through the PGS monitoring/accounting subsystem. The product order provides the time window, priority processing request, reprocessing request, standardizing order with timeline, or changes to standing order. The PGS Scheduling Subsystem provides an event or data-driven scheduler to control the script execution. The PGS Scheduling Subsystem is also responsible for the data staging and resources monitoring (such as the disk space, CPU, wall clock time of scripts, script and/or file dependencies, etc.).

The MODIS Levels 1A, 1B, 2 and 3 data processing are activated via production scripts. The staging of data is performed through the PGS utilities during the execution of the MODIS scripts. According to the MODIS dependency tables, some algorithms can be executed either in sequence or in parallel. In order to minimize the input and output redundancy, the data which are required by most of the algorithms (e.g., Level 1B data) are stored in memory or in high-speed RAM disk to reduce the I/O access time.

3.5 Policies

The MODIS SDST is responsible for software integration and testing and for developing MODIS scripts for data product generation. The MODIS SDST at the TLCF ensures that:

- the software is tested using test data sets provided by the science team and the results are compared with those obtained at the SCFs;
- the results are completely analyzed and reported to the science team members;
- the MODIS scripts and software are developed and maintained by the SDST;
- the MODIS PGEs can be executed in sequence or parallel, depending upon the algorithm dependency, data flow, and availability of resources; and
- modifications provided by science team members to the SDST are incorporated into the system quickly.

At this time, MODIS has identified the following key operations policies that are directly related to the SDST/PGS operational interfaces:

1. During the post-launch phase of operations (after Launch & Early Orbit Operations have completed) the PGS is able to configure, test, and move to production small code modifications delivered by the MODIS SDST within five working days of delivery.

2. The SDST is solely responsible for updating MODIS scripts and MODIS software running on the PGS. The SDST shall maintain the latest configured version of the software and deliver it to the PGS I&T staff for rebuilding and final test. The PGS I&T staff submits Software Problem Reports (SPRs) to the SDST should the software require modification.

3. The SDST staff performs software testing, data analysis operations, and QA on large volumes of data products that have been copied to the TLCF from the DAAC. This copy process uses either a high speed network link or high density media.

4. Operationally, the PGS staff communicates with the SDST as a single point of contact regarding status of MODIS processing. For this purpose, the PGS staff does not contact MODIS science team members directly.

5. Operationally, when MODIS processing is initiated using an input data set with gaps and that data set is later updated with the gaps filled, then the PGS runs the appropriate MODIS processing to produce a complete data set for the archive.

6. Operationally, when a primary external input data set is not available, the PGS staff does not initiate a MODIS processing task using a secondary input data set unless so directed by an external organization and the primary input will not be available within a reasonable (TBD) time period.

7. Large volumes of MODIS reprocessing requests will be performed in blocks at given kickoff points periodically. For example, a MODIS algorithm modification (V7.1) is moved into the production stream on July 3, 1999. Processing performed after this date uses the V7.1 software. Another modification (V7.2) is moved in on July 20. Processing performed after this date uses the V7.2 software. The MODIS team determines that sufficient algorithm modifications have occurred to warrant a reprocessing sequence to be performed. The MODIS team and PGS staff agree that a new reprocessing sequence will be started on September 1, 1999. On September 1, 1999, the MODIS reprocessing sequence starts reprocessing blocks of data using the latest available software version at that time.

3.6 Assumptions

The assumptions for the operations and test of the MDPS at the TLCF and PGS are listed below.

1. One or more computing systems in the TLCF run software identical to that of the EOSDIS PGS at the DAACs which do MODIS processing.

2. The ECS routinely obtains and prepares ancillary data for the MDPS. For input data, the ECS needs to perform the following functions:
   - Receive Data
   - Ingest Data
   - Concatenate Data Segments
   - Check for completeness
   - Obtain alternate data as prescribed by MODIS SDST
   - Nominal values are used as default
   - Temporary files can be saved

3. MODIS data are optionally stored in memory or high-speed RAM disk until other products are completed to reduce I/O redundancy.
4. Runtime parameters can be used to switch off specified products.
5. The SDST develops and controls MODIS scripts.
6. The PGS provides the scripting language.

4.0 EXTERNAL INTERFACES

Objects, external to the actual MODIS process, that interact directly or indirectly with the MODIS product generation executables (PGEs) consist of the PGS scheduler functions, IMS derived work orders, TLCF and TM SCFs.

4.1 PGS Scheduler Interactions

The PGS scheduler determines the data availability and sets up staging of all input data sets required for a given initiation of a MODIS executable (a PGE with linked code and an accompanying shell script) before dispatching this executable to the operating system execution queue. This includes all ancillary data sets. After the creation of a MODIS PGE by the scheduling task initiation section, the PGS maintains a ready state for the acceptance of any messages generated by the MODIS executable and routes those messages to the proper receivers, such as the MODIS TLCF, other instrument PIs, PGS error handlers, or TM SCFs. On completion of a PGE, the PGS receives the MODIS termination message with an indication of the quality of the Data Product as determined by the MODIS PGE. If MODIS PGE indicates a successful completion at this point, the output Data Products are then archived by the DADS, and the Metadata is transmitted to the IMS for further validation. Else, the evidence is saved and sent to the TLCF for analysis.

4.2 IMS Work Orders

PGS processing is a work order enabled data driven system. The production of MODIS Data Products is enabled as either standing work orders or special work orders. Standing work orders are generated automatically and cover relatively long periods of time. These standing work orders can be canceled due to abnormal conditions. Special work orders are generated for test integration, quick look, and reprocessing reasons. Special work orders occur frequently just after launch in the algorithm adjustment phase (test and integration), and proceed into periods of intense reprocessing for algorithm historical comparisons. Each work order component initiates a MODIS PGE consisting of a processing script and a Data Product executable. Execution of work orders and the initiation of a PGE are performed by the PGS and are not a MODIS function.

4.3 TLCF Interactions

The various MODIS processors generate entries into the MODIS processing status log and also perform automated data quality test checks as part of routine processing. The MODIS processing log entries are transferred to MODIS TLCF periodically for incorporation into a database that allows MODIS team members and the MODIS SDST to assess processing performance. Systematic processing anomalies are also detected via this database. Automated data quality with processing statistics are transferred to the MODIS TLCF to facilitate human QA. Synopsis of this QA-derived data also becomes a part of the MODIS TLCF hosted database.

4.4 SCF QA

QA data products are sent to the appropriate team member SCFs for each of the algorithms executed in the PGS environment. The QA products are derived from the MODIS data created primarily by the Level 2 and Level 3 PGEs. Calibration and general QA of the lower levels of MODIS products are performed at the TLCF. The SCFs return data QA reports to the DADS to be incorporated into the respective Data Product metadata.

Validation of MODIS data products and science data quality assurance are the responsibility of the science team members.

The PGS provides a software toolkit to assist in the development and porting of science software by providing a consistent set of tools and a means by which the TLCF and each SCF will simulate the PGS environment.

5.0 CAPABILITIES AND CHARACTERISTICS

Major functional capabilities provided by the MDPS and characteristics of the physical and organizational environment for the MDPS are described in this section. The concept of execution control, passage of data across external interfaces, major design constraints, and the PGS toolkit are also addressed.

Basically, the MDPS has the following capabilities:

- The MDPS functions without human interaction except for ancillary data preparation and QC.
- MODIS processing is capable of running with nominal values for missing data sets.
- MODIS processing provides for checkpointing at reasonable intermediate stages, typically at PGE boundaries.
- MODIS processing is initiated using a set of starting rules implemented by the PGS scheduler.

5.1 Software Development Environment

Two major activities are performed in the Software Development Environment:

1) Science Software Task - This activity includes development of the MODIS software, generation of the algorithm dependency diagram, testing individual software before the integration process, and configuration management of algorithm changes.
2) MODIS script development and testing - This activity includes integrating PGEs into MODIS scripts and testing the execution of these scripts.

The following assumptions are made for the software development:

- PGS Scheduling software is made available to the TLCF as soon as it is developed for MDPS test purposes.

- Test data are staged at the TLCF.
- Simulated operational input data are available in time for testing the version 1 software before delivery to the EOSDIS.
- There is a high speed data link between the TLCF and the ECS.
- The PGS is responsible for configuration control of the MDPS software delivered by SDST to the PGS.
- MODIS processing software writes software version identifiers into metadata for identification and heritage purposes.

5.1.1 Science Software Task

The science software task includes the development of MODIS software, the generation of an algorithm dependency diagram, the testing and integration of team member's code, and the management of algorithm change.

5.1.1.1 Algorithm Dependency Diagram Generation

In order to analyze the algorithm dependency, the MODIS SDST staff interacts with scientists, extracts information from: 1) the ATBDs, and 2) Science Processing Database (SPDB) on input data, data products, and algorithms. Using this information, the SDST explores algorithm dependency, and generates algorithm dependency diagrams. See Figures 3 and 4. These diagrams show data flow in a hierarchical structure. These diagrams have been updated many times to accommodate product descoping and updating, and will go through more iterations as algorithms require a change in their input or output data.

Figure 3 shows algorithm dependency for ocean data products. Figure 4 shows algorithm dependency for atmosphere and land data products.

Currently, the SDST is extracting dependency information from the Algorithm Theoretical Basis Documents (ATBDs) and interacting with team members to update the algorithm dependency diagram.

The algorithm dependency diagram and data flow are developed at an early stage in the design, providing information on the process stream for product generation. The data flow indicates the required input data to each algorithm and the products generated by each algorithm.

Some dependencies exist between the MODIS data products and non-MODIS data products. The science team members identify the required and alternate inputs including standard products from MODIS and other instruments, calibration data, and ancillary data. The team member must decide what to do when required data sources are not available.

5.1.1.2 Science Software Delivery and Test

The MODIS team members are responsible for the development of their own algorithms at the SCF. Once their programs work correctly, team members deliver their code and related information, to SDST for algorithm integration. The delivery policies have been specified in the documents "MODIS Software and Data Management Plan" and "MODIS SDST Coding Recommendations for the MODIS Science Team". The SDST runs the delivered code before and

after SDST modification on the TLCF. When results are not consistent, the SDST cooperates with the team member to resolve differences. The iterative interactions and corrections continue until problems are resolved.

Testing of the team members' code is described in an algorithm-specific test plan. The test plan specifies the type and location of test data, the operator/analyst actions necessary to conduct the test, procedures for analyzing test results, and success/failure criteria for the test. A test activity log, which is maintained for each test, serves as a test history file. The log includes items regarding installation, testing, test results, and test status, as well as questions, concerns, and action items.

5.1.1.3 Science Software Change

The details for science product development are given in the Software and Data Management Plan. The details for change control processing are given in the Configuration Management Plan. The SDST supports the activities of science software integration, updating, refinement, or replacement. In addition, the SDST updates the algorithm dependency diagram when algorithms are added or deleted. Any changes to the algorithm are delivered to the PGS after careful examination and testing by the SDST.

5.1.2 MODIS Script Development and Test

In order to match the PGS operational environment, the generation and testing of the production script is accomplished in a simulated PGS environment at the TLCF. The components of the production script include:

- Initialization program - to notify the Monitoring and Accounting (M&A) subsystem and initiate activities in support of the toolkit functions.
- Requests for data staging.
- Code execution (which includes memory allocation).
- Quality assurance execution and associated data set generation.
- Requests for data file cleanup.
- Termination program - to notify the Management and Accounting subsystem and update accounting information including all inputs/outputs, success/failure status, etc.

The production script can contain any number of PGS-allowable UNIX commands in addition to the commands used to initiate the execution of code when all requirements are satisfied. The commands in the production script can be structured in any way provided in UNIX. This means that code can be executed in parallel, in sequence, or as a result of a decision made in the production script. The production script will send processing messages about any abnormal or abandoned processing messages to the PGS control center.

5.1.2.1 MODIS Script Development

Based on the algorithm dependency diagram, the higher level script integrates PGEs into a production stream, i.e., assigns the processing order to a PGE. The script can be modified easily when the scheduler is changed or special processing is requested.

The assumptions on the MODIS script development are listed below:

- MODIS processing executes within the context of the PGS Production scripts.
- MODIS scripts will be sensitive to the most efficient use of processing resources.
- MODIS software runs on dedicated computing systems.
- MODIS scripts are capable of multiprocessing (i.e., several PGEs can be executed in parallel in a multiprocessing computing environment).
- The PGS supports the execution of multiple instances of the same PGE, i.e., multiple PGEs can access the same data. (PGS must ensure file conflicts are avoided.)
- MODIS processes have the capability to allocate memory and create temporary files in the PGS environment.
- MODIS processing will operate on an integer number of full scans.
- When called with the remainder of a data set, the process will replace the previous results with the more complete ones, and, optionally chain this through higher levels of processing.
- MODIS assumes a work order process, allowing calls of special scripts which are different from the standard production operations scripts.

5.1.2.2 MODIS Script Test

In order to pass the acceptance test at the PGS, the MODIS script will be fully tested before porting it to the PGS for routine processing. Because the TLCF and PGS environments are not identical, (for example, the PGS has to consider the whole EOS production generation, not only the MODIS products), there could be some modifications needed to the script after it gets to the PGS. SDST and PGS personnel will work together to resolve any inconsistencies. A test activity log, which will be maintained for each test, will serve as a test history file. The log will include items regarding installation, testing, test results, and test status, as well as questions, concerns, and action items

The majority of testing is done on the TLCF. The PGS staff has the option to observe the testing on the TLCF. Minimal additional testing is required on the PGS.

5.2 Operational Environment

This section discusses the operational requirements and the management of data storage during the MDPS operations.

5.2.1 Operational Requirements

It is assumed that all required input data are available at the PGS before executing MODIS scripts. The MODIS scripts are designed to minimize the need for intervention by the operators.

5.2.2 Operational Requirements for Level 1A

MODIS Level 1 data processing takes place in the PGS and involves the transformation of Level-0 data received from EDOS into Level 1A. Level 1A is reversible to Level-0. Because of the volume of data, this processing is necessarily fully automatic. The primary uses of Level 1A are to expedite

the processing of higher level MODIS data and to archive the data for possible future reprocessing. The 1A product is also used by the MCST for maintenance and calibration.

The Level 1A product is derived from two input data streams: instrument data packets and the Level-0 position and attitude packets. Production of each Level 1A data record requires the assemblage of a 'scan cube' which consists of many packets of raw instrument Level-0 data. The requirements for the input bit error rate, chronology, completeness, coverage and ancillary data are specified in Reference 6.

Level 1A processing appends spacecraft position and attitude to each scan cube. Appended also are header information about data quality, completeness, calibration, instrument temperatures, and algorithm version numbers. In addition, all the information needed for cataloging and data extraction is inserted into the headers. See Reference 6.

5.2.3 Operational Requirements for Level 1B

Level 1B is derived from Level 1A. This transformation is not reversible. Level 1A data structure is preserved to the extent possible. The basic Level 1B data processing consists of: earth location, satellite/solar zenith angle/relative azimuth angle determination, and time calculations; radiometric calibration; data quality assessment; and header record processing. See Reference 14.

5.2.4 Minimization of Data I/O Redundancy in the MDPS

From the algorithm dependency diagram, it is clear that some products are used by more than one PGE. The I/O redundancy can be minimized through production stream I/O file tools. Production Stream I/O files exist for the duration of a complete flow of programs and program control. A production stream file may exist until some criterion such as a processing deadline or production stream completion is met. File deletion criteria may be designed through PGS tools and may be met by the setting of flags by various production processes, completion of specific data products, or termination of specific production streams or programs.

5.3 Processing Capabilities

MODIS processing estimates for storage volume and processing power are an ongoing task. An overview of the memory in megabytes and time required to make each MODIS standard product for an orbit's worth of data is provided by Figure 5.

Estimates of the time required to make each MODIS product is based on the following assumptions: a computer system with sustained performance of 6,000 Gflops and a 20 Mbytes/sec data transfer rate; all ancillary data sets are available and pre-processed so that no delays occur waiting for the arrival of data sets; and estimates of system resources required to pre-process the ancillary data sets will be supplied later and are not included in the current estimate of processing power required to make the standard products.

5.3.1 MODIS Level 1A Processing Capabilities

MODIS Level 1A process ingests Level 0 instrument data packets at an equivalent real time rate of 1.3 Mbytes per second. This corresponds to an orbital rate of 5 Gbytes, for the orbital period of 98 minutes and should be multiplied by a factor of 5 to 10 to represent the actual processing load. The MODIS Level 1A process creates Data Products at a rate of 8 Gbytes per orbit, and will require .1 Gflops of processing power, .2 Gbytes of storage and TBD Gbytes of main memory.

For each execution of the MODIS Level 1A PGE, the software will generate a data set based on Level 0 data recorded during a prescribed time interval. The end points of each such time interval will be provided to the ECS prior to processing time. The time intervals will be selected on the basis of commands for changes in the instrument mode, or for lunar calibration, or on orbital boundaries, etc. There will be short overlaps between consecutive time intervals to assure complete data sets. Provision of the processing time intervals will require the availability of command sequences and orbital predictions in advance.

5.3.2 MODIS Level 1B Processing Capabilities

The MODIS Level 1B process utilizes the MODIS Level 1A Data Products on an orbital basis. This Level 1B process will consume 3000 Mflops of equivalent processing power and 40 Gbytes of storage and 52 Gbytes of main memory.

The MODIS Level-1B processing requires a number of inputs in the course of its normal processing cycle. In addition to the geo-located Level-1A data in memory at the start of Level-1B processing, data will be required from several other sources. The major additional sources of data (in terms of the data rates required) are : 5% of a previous orbit of 1A data acquired in day mode, prior classification masks and ancillary calibration data (data obtained either from other instruments or from ground based calibration). Additional minor inputs are : historical calibration coefficients, location and characteristics of calibration sites and spectral signature inputs to the masking algorithm.

The major outputs from the 1B processing are: orbits of calibrated 1B data which will presumably remain in memory for level 2 processing, 250m, 500m and 1000m resolution classification masks and metadata for both the Level-1B processing and the masking processing. Minor outputs include: updates to the historical database of calibration coefficients, entries to the instrument performance and Quality Assurance databases, updates to the Spectral Signature Database and periodic updates to the Calibration Constant File.

The summarized Level-1B memory and storage requirements are: *52 GByte of memory and 32 GByte of storage per orbit.* In addition some of the outputs generated from previous orbits (most notably the mask product) are required as inputs to the processing of the current orbit. This will require that Level-1B products be stored in a database format which will allow them to be staged rapidly enough to be used in the Level-1B on-line processing.

The data volumes for inputs and outputs are summarized in a following table as well as in Figure 6, illustrating which pieces of data are required in memory during which processing steps. Note that the equivalent real time rate does not take into account any scheduling or scripting overhead, reprocessing, testing, or quick look requirements

I/O requirements for 1 orbit of Level 1B

Major Inputs to Level 1B Processing per Orbit

	Volume (GBytes)	Source
Level 1A Data	7	Resident in Memory
Geo-Location Data	2	Resident in Memory
Polar Data	1	Disk
Prior Masks	10	Disk
Ancillary Data	2	Disk
Calibration Site Data	0.2	Disk
Spectral Signature Data	0.2	Disk
Total Inputs from Disk	13.4	Disk

Major Outputs from Level 1B Processing per Orbit

	Volume (GBytes)	Destination
Level 1B Data	14	Disk/Remain in Memory
Geo-Location Data	2	Remain in Memory
Classification Masks	10	Disk
Level 1B Metadata	4	Disk
Mask Metadata	2	Disk
instrument Performance & QA Data	2	Disk
Total Storage Requirement	32	Disk

5.3.3 MODIS Level 2 Processing Capabilities

The MODIS Level 2 land/atmosphere processes read the Level 1B orbital Data Product and require 1000 Mflops of processing power with 3 Gbytes of storage and TBD Gbytes of main memory. The ocean, land, and atmosphere products require 16 Gbytes of information per orbit.

The summaries of resources required for processing MODIS instrument data up to the orbital Level 2 Data product are 5.7 Gflops to produce 600 Gbytes of data per day. The processing requirements to produce the MODIS Level 3 and 4 Data Products have not been fully explored at the present time.

5.4 Data Processing

The MDPS communicates with the PGS before, during, and after the MODIS script execution. Also, the MDPS reports the status of script execution to the PGS and records the status on a log

for future evaluation. Production script execution consists of two components: (1) Input data control, and (2) PGE execution.

5.4.1 Input Data Control

This component manages and controls all required data, such as the Level 1A and 1B data, ancillary data, and correlative data. Data and resource availability can also be checked by this component in the script. Data validation should be performed by the PGE to avoid any abnormal computation. This component also provides options when input data are not available.

The PGS ensures that input data are "on-line" (i.e., on magnetic disk at the PGS and ready for access) prior to the code execution. The PGS toolkit provides the capability to query for availability of data required by science production runs. Toolkit interfaces to the PGS at the start of a production run provide information about the dimensions of input data sets and memory/disk space required.

As mentioned earlier, in order to avoid any input/output format problems, the team member should use PGS toolkit's to read and write standard science data products from standard structures. The data structure and data format of MODIS products are determined by the team members with support from the SDST.

5.4.2 PGE Execution

This component executes the science software. In case of processing errors, the code uses PGS tools to flag these errors and send a message to the appropriate person and take an appropriate action (e.g., terminate execution, ignore the errors, or abandon job after receiving a command). On termination of a program or production stream, either the production stream must relinquish system resources or the PGS will automatically recover resources used by the stream. MODIS shells and/or PGEs perform automated data product quality assurance and generate browse data. Quick Look data are generated on demand. Metadata are updated after each process.

Since the execution of the code can interrupted or aborted, the software periodically posts completed results to temporary storage. To avoid wasting previous processing, the software is designed with checkpoints at reasonable breakpoints.

System resources requested by one program may need to be shared through a production stream. In this case, tools need to be designed which can differentiate between resource requests for single executions and requests allowing the passing of resources between programs within a production stream.

5.4.3 Processing Mode

The various processing modes are: standard (routine) processing, Quicklook (near-real-time) processing, reprocessing, special request processing, and system test runs. The MDPS accepts an initiation message from the scheduler indicating the mode of processing and input data set names. This message contains the processing mode (e.g., standard, reprocessing, etc.), input file names, and data descriptors. The Scheduler checks the time, the availability of input data, computer memory, and other system resources. The MDPS sends an error message to the PGS if the process fails due to invalid input data or insufficient memory allocation, and the execution is abandoned or halted for PGS checking.

The PGS may send a request to suspend the current execution and resume the execution of a previous process at the last completed location. As mentioned in the "PGE Execution", it is expected that the program can periodically save or post the completed data to a housekeeping storage and set a checkpoint for reference. This allows the process to be interrupted with minimum penalty and maintains output product integrity. If the PGS sends a cancel command, the scheduler performs an orderly abort in which the data sets are properly posted and closed and processing log entries are posted.

The MDPS informs the PGS Scheduler of the completion and validation of data products and associated metadata and browse data. This allows the Scheduler to start other production runs dependent on these data products and pass metadata and browse products to the appropriate systems. In general, the MODIS script process automatically posts and closes generated data sets and passes ownership of those data sets to the PGS Product Manager.

5.4.3.1 Standard (Routine) Processing

Standard processing refers to the MODIS script execution for the generation of standard product. Some Level 2 algorithms require ancillary data. These ancillary data sets could include, for example, Level 1B (or Level 2) data from other instruments, either space- or ground-based. Level 2 processing derives geophysical parameters from these inputs through the application of science algorithms. The Level 2 products will be similar to the Level 1B product in the data structure, that is, it will consist of orbital swaths of geophysical parameter data along with appended information. Each Level 2 product may contain one or more related parameters or attributes.

In the standard processing, data staging is accomplished before initiating the processing script. The scheduler keeps track of the files staged for each PGE and the files needed, which are output from a previous PGE. If there are any difficulties in staging data, the Scheduler issues a warning to the System Management Center. If staging does not occur and no direction is given to the PGS, the production script is aborted and a message is sent to the Monitoring and Accounting subsystem.

5.4.3.2 Quicklook (Near-Real-Time) Processing

Quicklook Products are ordered in two ways: 1) Quicklook Flag in Instrument Telemetry is on. Only fractions of orbits contain this flag. The Quicklook Flag is scheduled to be turned on time so that Quicklook Products can be produced in the IMS. The partial orbit is processed automatically by EDOS and in all applicable EOSDIS components (i.e., a PGS) in a Quicklook fashion with maximum priority. 2) Special real-time event creates need for Quicklook data. In the case of a real-time event such as a volcanic eruption, forest fire, or other natural or man-made disaster, MODIS may request previously acquired or soon-to-be acquired data to be processed quickly and with maximum priority. The MODIS team notifies the appropriate personnel in the EOSDIS to initiate Quicklook processing when these events occur.

The Quicklook processing mode is used to provide immediate access to PGS processing for data items that are needed in less than the usual processing turnaround times. For example, the science team may require some Level 2 data in time to support a field experiment. The appropriate portions of Levels 0, 1A, 1B, and 2 data would be mandated and processed as priority data at the PGS, and similarly processed on a priority basis at the PGS without disruption of routine processing. The resulting Quicklook products would be delivered to the requester within three to eight hours of observation.

The MODIS Data Processing Software will create Quicklook data products, for use during field campaigns and other high priority processing events. MODIS Quicklook data is flagged by the

instrument and telemetered in the MODIS instrument data stream. Quicklook data is expected to consist of partial orbits, covering the areas required for field campaign investigations.

For Level-1A, the Quicklook product will have the same format as the standard Data Product, but will not be padded to full orbit boundaries. It will contain an integer number of full scans and a Quicklook identification in the Data Product header. The extent of Quicklook processing at this time is TBD. If Quicklook Level-1B, Level-2, and Level-3 products are required, the SDST will make every effort to use the same software used for the standard products for the Quicklook processing.

See the Quicklook Processing Scenario in Section 6.3 of this document for further details about Quicklook operations.

5.4.3.3 Reprocessing

The science team may request that Levels 2 through 3 data be reprocessed and would supply updated calibration algorithms and/or science algorithms. Reprocessing occurs without interruption of routine production activities, and at least twice the routine processing rate. The reprocessing could be updating and replacing products with data produced with improved algorithms or calibration, or retrospective processing, once a new product is introduced.

Significant reprocessing of data is being planned in sizing the PGS. As science understanding changes or bugs are detected, improvements and fixes to production software will be made. The PGS will have to manage the reprocessing decisions as well as handle on-going processing requirements.

5.5 Error Handling/Logs

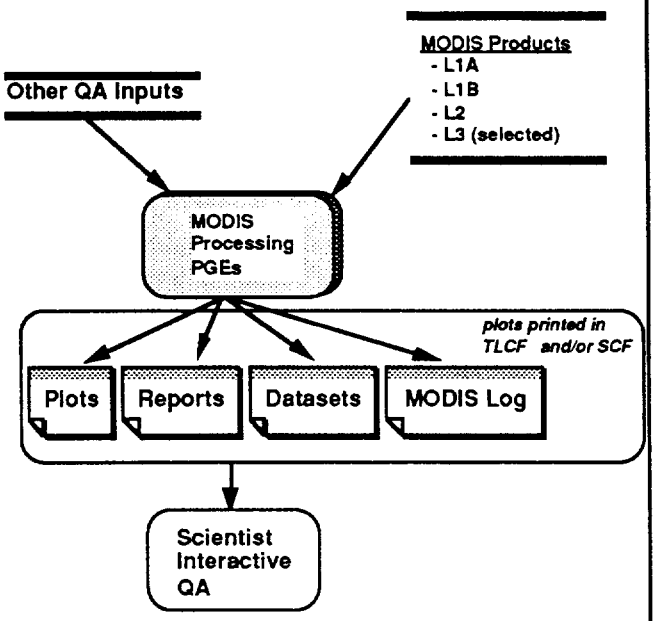
The MDPS generates unsolicited alarm messages to the PGS in the event of data or instrument problems. This class of messages indicates to the PGS that a serious problem has occurred within the MDPS process that could lead to the generation of invalid data. The contents of the message indicates the nature and severity of the problem. The message is expected to contain both indicator flags with predetermined error types and ASCII contents that can be displayed to an operator or transmitted to other EOS computing facilities. These messages are also entered into the processing log.

Error output tools provide the ability for software to report errors or problems to the PGS system managers and the code developer. The MDPS generates messages of a non-critical nature for informational purposes to the PGS for routing to other EOS computing facilities. This is a class of messages that are informational in nature and do not affect the quality of the output data products. They are meant to be passed on to off-line processes. in the EOS environment for resolution. These messages are also posted to the processing log. This facility might, for example, inform the Advanced Spaceborne Thermal Emission and Reflection (ASTER) team of MODIS-detected volcanic activity.

The MDPS posts all relevant processing information to the Processing Log. This log, either common to all the MODIS processes or unique to each process, contains an audit trail of time ordered processing events. The contents of this log file are written by the processes and can be examined, but not altered, by any other process. The location of and access method to this processing file is TBD and may be influenced by performance concerns.

6.0 SAMPLE OPERATIONAL SCENARIOS

6.1 Standard MODIS QA Scenario



MODIS Routine QA Scenario

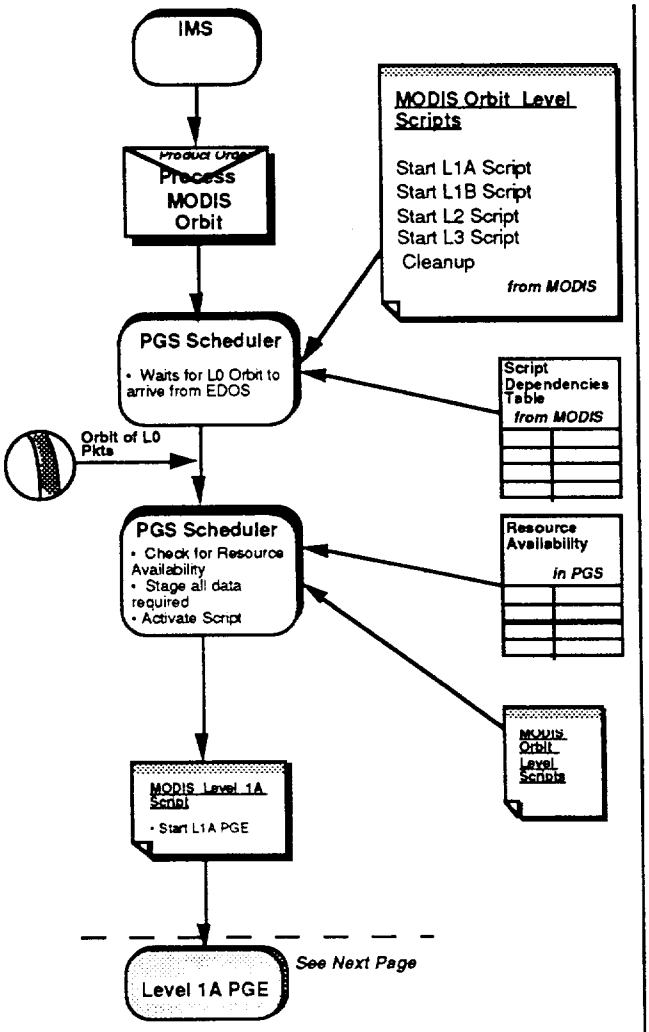
- MODIS PGS Software writes QA messages to MODIS Log and QA information to the product headers and metadata during execution.

- MODIS PGEs executed on daily, weekly, monthly and TBD other intervals.

- MODIS PGEs produce plots and reports in the TLCF and / or SCF using the products in the archive (or temporary PGS storage).

- MODIS Science Team Members analyze the datasets, plots, reports, and MODIS log using interactive QA tools in the TLCF or SCFs.

6.2 Standard MODIS Processing Scenario



MODIS Routine Processing Scenario

Page 1 of 3

- Product Order Arrives From IMS

- PGS Scheduler identifies script associated with Product Order

- PGS Scheduler uses a script dependencies table (previously delivered by MODIS SDST) to determine the processing resources and data sets required prior to script execution.

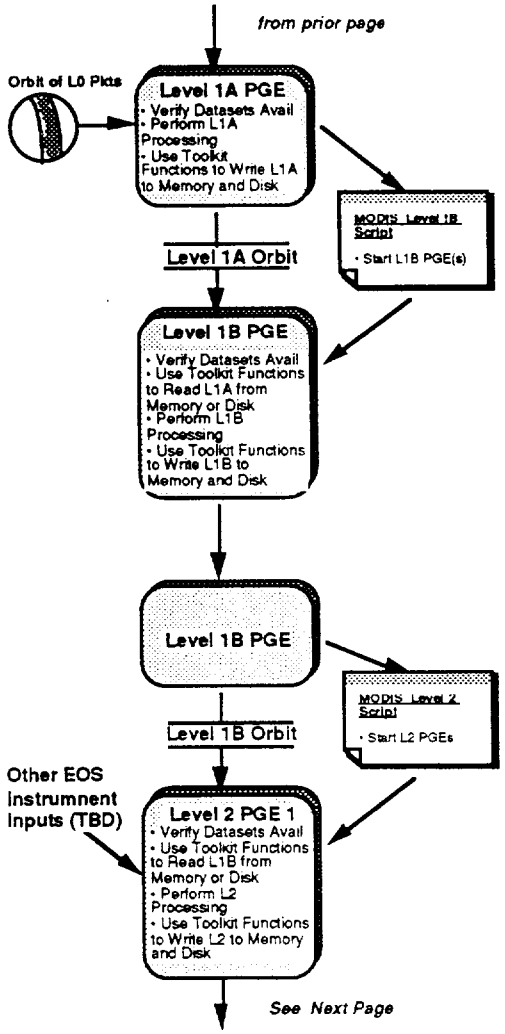
- PGS Scheduler waits for arrival of the MODIS orbit from EDOS

- Orbit of L0 packets arrives.

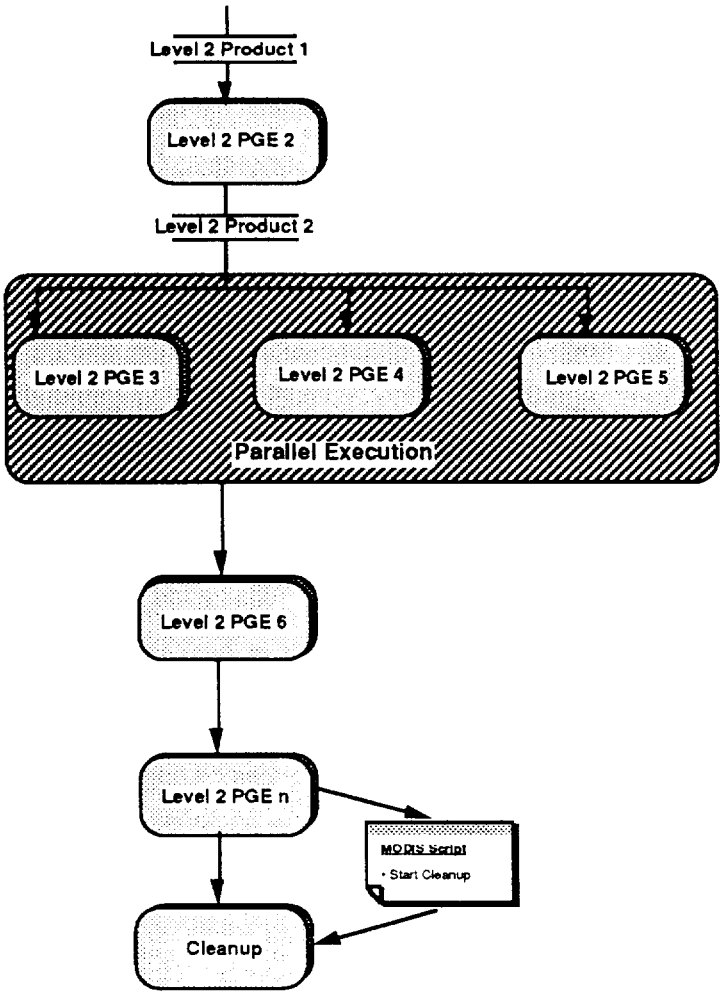
- PGS Scheduler verifies sufficient resources are available to execute the Level 1A script.

- When all required resources are available and allocated and all datasets are staged, PGS Scheduler initiates execution of the *MODIS Level 1A Script.*

- MODIS Level 1A Script starts Level 1A PGE.



- Level 1A PGE verifies required datasets are staged.
- Level 1A PGE performs L1A processing and uses PGS Toolkit functions to write L1A to memory and to disk.
- Level 1B Script is initiated.
- Level 1B PGE verifies required datasets are staged.
- Level 1B PGE performs L1B processing and uses PGS Toolkit functions to read L1A from memory or from disk. PGE writes L1B to memory and to disk.
- Level 2 Script is initiated.
- Level 2 PGE 1 verifies required datasets are staged.
- Level 2 PGE 1 performs processing and uses PGS Toolkit functions to read L1B (also L1A in some cases) from memory or from disk. PGE writes L2 products to memory (in some cases) and to disk.



• Control returned to MODIS Level 2 Script where L2 PGE 2 is executed.

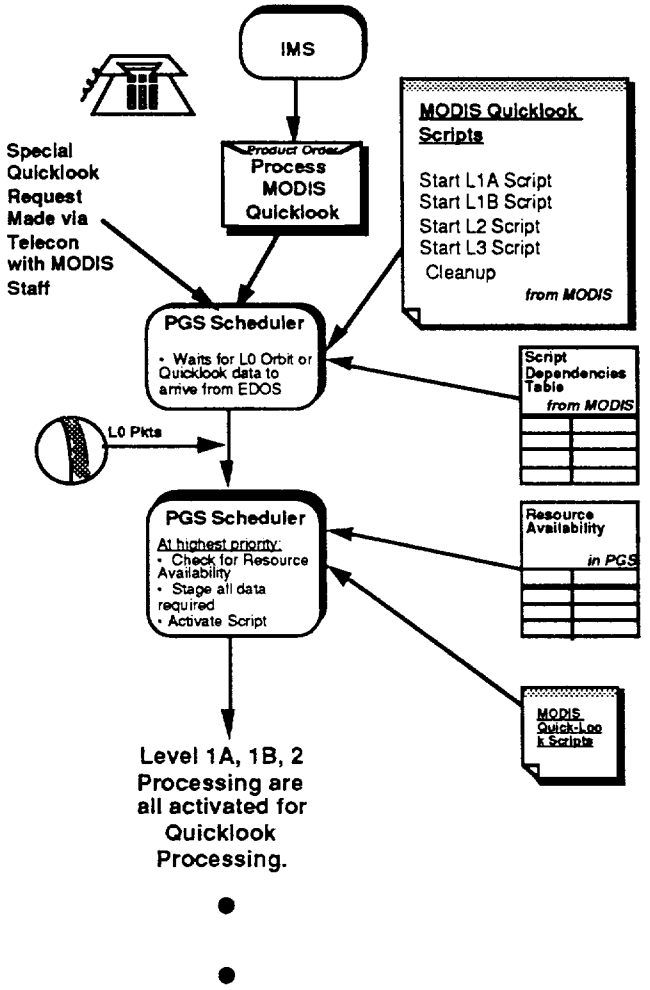
NOTE: The parallel nature of the Level 2 PGEs shown here is an example of the processing approach MODIS is using. The operational system will have many more algorithms operating in parrallel asynchronously.

• Level 2 PGEs 3,4, and 5 all execute in parallel, if resources allow.

• Remaining Level 2 PGEs execute

• Cleanup script executed.

6.3 Quicklook (Near-Real-Time Processing) Scenario



MODIS Quicklook Processing Scenario

• Quicklook Products will be ordered in two ways:

1. Quicklook Flag in Instrument Telemetry is set to on. Only fractions of orbits will contain this flag. The Quicklook Flag will be scheduled to be set to on, resulting in sufficient time to produce Quicklook product orders in the IMS. The partial orbit shall be automatically processed by EDOS and all applicable EOSDIS components (i.e., PGS) in a Quicklook fashion with maximum priority.

2. Special real-time event creates need for Quicklook data. In the case of a real-time event such as a volcanic eruption, forest fire, or other natural or man-made disaster, MODIS may request previously acquired or soon to be acquired data to be processed quickly, at maximum priority. The MODIS Team will notify the appropriate personnel in the EOSDIS to initiate Quicklook processing when these events occur.

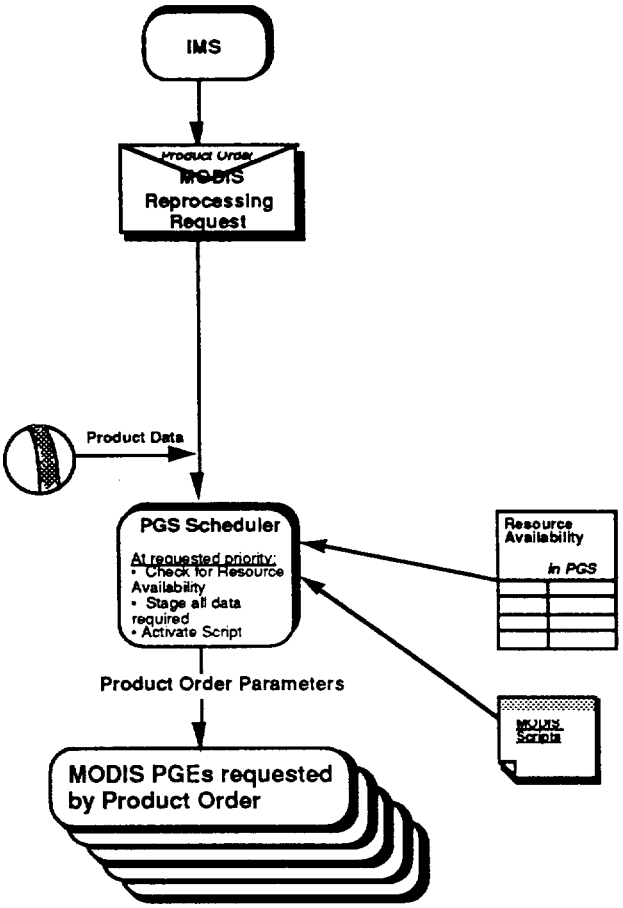
• PGS Scheduler uses a script dependencies table (previously delivered by MODIS SDST) to determine the processing resources and data sets required prior to script execution.

• PGS Scheduler waits for arrival of the MODIS Quicklook data from EDOS

• L0 packets arrive.

• All MODIS processing on this Quicklook data is performed at maximum priority.

6.4 Reprocessing Scenario



MODIS Reprocessing Scenario

• MODIS will make Reprocessing Requests on the EOSDIS for groups of products over a range of time. Typically, MODIS will reprocess long time ranges (ie. weeks or months in length).

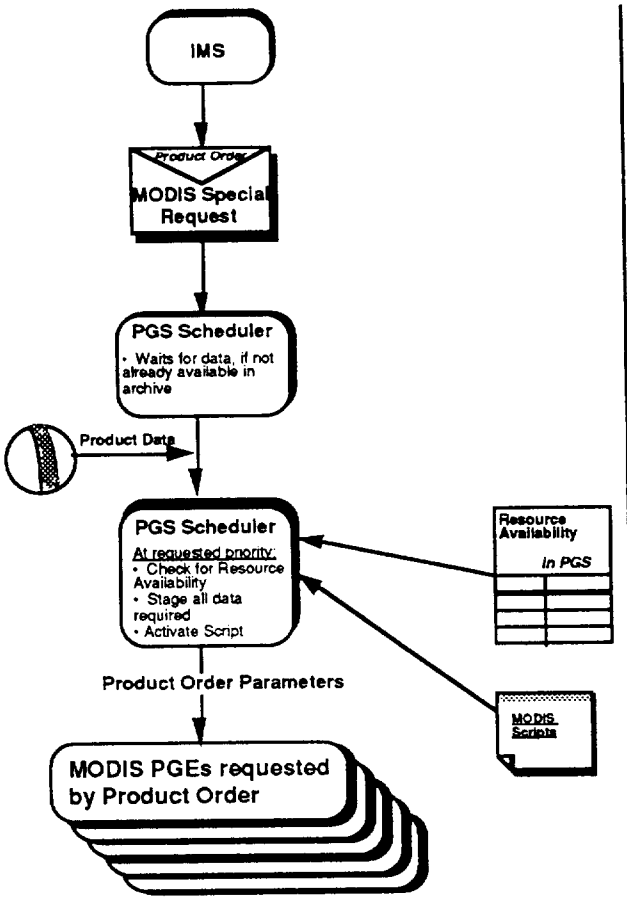
• MODIS will identify the time range and subset of products to be produced in the Product Order. Product Orders contain Product Order Parameters that specify product identifiers and "run-time" options in the processing.

• PGS Scheduler uses the product order parameters to determine the script to activate.

• PGS Scheduler activates appropriate script(s) with product order parameters.

• MODIS processing is performed at reprocessing priority.

6.5 Special Request Processing Scenario



MODIS Special Request Processing Scenario

- MODIS will make Special Requests on the EOSDIS for certain products. MODIS will identify the time range and subset of products to be produced in a Product Order. Product Orders contain Product Order Parameters that specify product identifiers and "run-time" options in the processing.

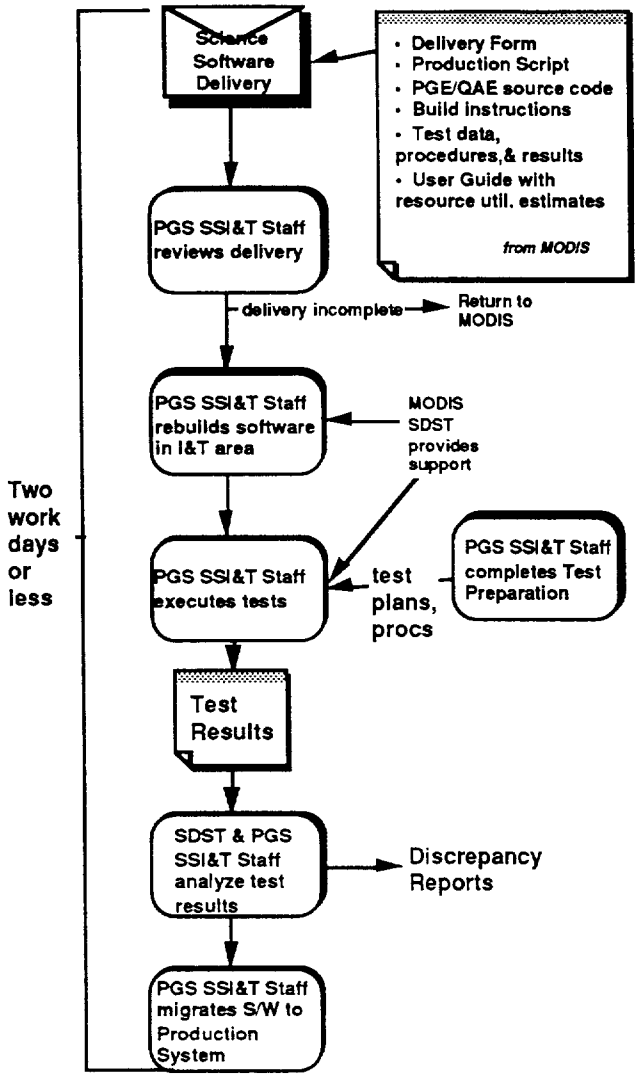
- The product order will also specify the priority of the Special Request Processing to be performed.

- PGS Scheduler uses the product order parameters to determine the script to activate.

- PGS Scheduler activates appropriate script(s) with product order parameters.

- MODIS processing is performed at requested priority.

6.6 System Test Scenario



MODIS System Integration & Test Scenario
Postlaunch Phase

• MODIS SDST delivers software package to the PGS electronically. *Note: The SDST will perform all PGS software deliveries on behalf of the MODIS Science Team.*

• The PGS staff will review the delivery and assess completeness. If the delivery is incomplete then the SDST will work with the PGS staff to solve the problem.

• A substantial amount of the integration and testing performed on the science software will be performed by the SDST in the TLCF.

• The PGS staff will rebuild the software in the I&T area and will address any remaining source code portability issues in cooperation with the SDST.

• PGS Staff executes tests and conducts performance baselining for the PGE's. Variations in the PGE's operational parameters/modes on resource utilization are investigated.

• SDST and PGS Staff analyze the test results and write Discrepancy Reports (DRs), when problems occur.

• SDST performs ALL modifications to the MODIS Software.

• PGS Staff migrates software to Production area for use operationally.

7.0 ACRONYMS AND ABBREVIATIONS

ASTER	Advanced Spaceborne Thermal Emission and Reflection
ATBD	Algorithm Theoretical Basis Definition
CASE	Computer Aided Software Engineering
CCB	Configuration Control Board
CCR	Configuration Change Request
CCSDS	Consultative Committee on Space Data Systems
CRC	Cyclic Redundancy Check
DAAC	Distributed Active Archive Center
DADS	Data Archive and Distribution System
DEM	Digital Elevation Model
DFD	Data Flow Diagram
ECMWF	European Centre of Medium Range Weather
ECS	EOSDIS Core System
EDOS	EOS Data and Operations System
EOS	Earth Observing System
EOSDIS	EOS Data and Information System
GSFC	Goddard Space Flight Center
HDF	Hierarchical Data Format
IMS	Information Management System
IPAR	Intercepted PAR
IR	Infrared
ISCCP	International Satellite Cloud Climatology Program
I&T	Integration and Test
JGOFS	Joint Global Ocean Flux Study
M&A	Monitoring and Accounting
MIDACS	MODIS Information Data and Control System
MODIS	Moderate-Resolution Imaging Spectroradiometer
MDPS	MODIS Data Processing System
PAR	Photosynthetical Active Radiation
PDR	Preliminary Design Review
PGE	Product Generation Executable
PGS	Product Generation System
PM	Product Management

ACRONYMS AND ABBREVIATIONS (continued)

PMS	Product Management System
QA	Quality Assurance
QAE	Quality Assurance Executable
QC	Quality Control
RDC	Research and Data Systems Corporation
SBRC	Santa Barbara Research Center
SCF	Science Computing Facility
SCMA	Scheduling, Control, Monitoring, and Accounting
SDST	Science Data Support Team
SPDB	Science Processing Data Base
SPR	Software Problem Report
SRR	Systems Requirements Review
SST	Sea Surface Temperature
TDRSS	Tracking and Data Relay Satellite System
TLCF	Team Leader Computing Facility
VIS	Visible
WOCE	World Ocean Circulation Experiment

8.0 GLOSSARY

ALGORITHMS consist of the executable programs for science product generation, source code of these executable programs, job control scripts, and algorithm documentation. Algorithms are the result of new or updated science code passing through the integration and testing process, involving the scientist and the MODIS SDST. After formal approval, algorithms are delivered by the PGS to the DADS for storage, and are retrieved as needed to support product generation. The DADS shall also archive algorithms contributed as EOSDIS resources by other data centers. Some frequently used algorithms may also be kept on line in the PGS. Algorithms shall be orderable and distributed to authorized users. Algorithm support is not provided for outside users either by the SDST or the Science team Members.

ALGORITHM UPDATES are delivered to the PGS' integration and testing environment by scientists at an SCF. They represent changes to existing production algorithms, or are new algorithms to produce new Standard Products. Algorithm updates include the source code for the candidate algorithm, its associated documentation, a job step control skeleton, test data sets, and expected test results.

ANCILLARY DATA refers to any data, other than Standard Products, that are required as input in the generation of a Standard Product. This may include selected engineering data from the EOS platform or ephemeris data, as well as non-EOS ancillary data. All ancillary data are received by the PGS from the DADS.

AT-LAUNCH PRODUCTS are data products for which algorithms are scheduled to be completed and operational at launch time. Generation of the at-launch products will begin as soon as data are available.

BAND is a center frequency and filter function width combination that is referenced by an arbitrary number. One channel usually measures data from one band, but MODIS has two channels of data for each of bands 13 and 14.

ß VERSION SOFTWARE is the preliminary version of software used to test migration from the SCF or TLCF to the EOSDIS, exercise interfaces, and test execution in the operational environment.

BROWSE DATA are subsets of a data set other than the directory and metadata that facilitates user selection of specific data having the required characteristics. For example, for image data, browse data could be a single channel of multichannel data with degraded resolution. The form of browse data is generally unique for each type of data set and depends on the nature of the data and the criteria used for data selection within the related science discipline.

CALIBRATION is the process of removing biases from the measurements. The information required to perform calibration of the instrument science data will include instrument engineering data, spacecraft or platform engineering data, pre-flight calibration measurements, in-flight calibrator measurements, ground truth data, and calibration equation coefficients derived using calibration software.

CHANNEL is the data derived from an instrument data gathering electronic channel. Channels of data may be obtained in parallel. Contrast this with a band of data.

CORRELATIVE DATA are data, generally from outside sources, needed to evaluate and validate EOS data products.

DATA QUALITY REQUEST is a request issued by the PGS to a scientist at an SCF to perform quality assurance (QA) of a particular product before future processing or distribution. A time window is applied to the request in keeping with the production schedule.

DOCUMENTS are the hard copy or digitized references or records about an algorithm, or the data products generated by the algorithm. These shall be archived at the DADS.

FRAME is a set of data representing one instance of: all bands of science data and 10 along track spatial elements (day or night mode), or all bands & 10 spatial elements of solar diffuser, black body, or space view data, or specialized data from the SRCA or Engineering/Memory dumps.

INSTRUMENT DATA are data specifically associated with the instrument, either because it was generated by the instrument or included in data packets identified with that instrument. These data consist of instrument science and engineering data, and possibly ancillary data. These data may be assembled for transmission by the instrument, or by an on-board processor of the instrument data.

INTERACTIVE SESSION DIALOG consists of messages that flow between a scientist at an SCF and the PGS that support general communication with the Integration and Test Service. This includes log-ins, mail messages, etc.

LEVEL 0 DATA are raw instrument data at original resolution, time ordered, with duplicates removed.

LEVEL 2 DATA are Level 0 data, which may have been reformatted or transformed reversibly and packaged with needed ancillary, engineering, and auxiliary data.

METADATA are data which describe the content, format, and utility of a Standard Product. It includes such data as algorithm and calibration numbers, size of product, date created, QA information from the PI's, summary statistics, and an audit trail. Metadata are received by each DADS with the corresponding data sets. DADS validates it physically, updates it with inventory information, enters it into a distributed database (to which the Information Management System (IMS) has access), and archives it. Metadata about special products produced at SCFs shall be sent to the DADS along with their associated data products.

METADATA UPDATES are additional or changed metadata items relating to a previously delivered product.

ON TIME QA is a response to a data quality request that is received within the established production time window. It is received from a scientist at an SCF. It consists of data which will be used to complete the QA fields of the metadata. Overdue QA responses are sent directly to the DADS.

ORBIT is the three dimensions of the science data: along track, across track, and band number with ground locations, engineering data (per scan), and quality indications attached.

PIXEL is the smallest decomposed unique science data value. One pixel represents the data from one spatial position and one channel of data.

POST-LAUNCH PRODUCTS are data products for which algorithms are scheduled to be completed at some time after launch.

QUICK-LOOK DATA are data that carry a quick-look identifier which is initially applied at the MODIS instrument in response to a request for quick processing in support of a field experiment or other special purpose. Processing is done in near-real-time as the data sets are received using the same algorithms as for normal processing. Calibration and quality assurance requirements are relaxed for fast turn-around. Data identified as quick-look are also processed in the normal mode with all other data for standard products.

SCIENCE COMPUTING FACILITY is a MODIS science team member's computing facility to support algorithm development and testing and generation of MODIS special products.

SCIENCE DATA SUPPORT TEAM is the software support group designated by the MODIS Science team Leader to develop MODIS science data processing requirements, develop MODIS Level 1 processing software, develop the higher-level processing shell, support the integration of team members' algorithms into the production environment, and provide support for the development of required plans and documentation.

SPATIAL ELEMENT is all channels of data corresponding to a 1-kilometer ground equivalent field of view. This means that 16- to 250-meter detectors, 4- to 500-meter detectors, and 1- to 1-km detectors reside within one spatial element.

SPECIAL PRODUCTS are special science data products consisting of L1A, L1B, L2, L3, and L4 which are produced at the SCF or the TLCF. (See also Standard Products.) These shall be archived at the DADS and distributed to authorized requesters.

TEAM LEADER COMPUTING FACILITY is the SCF designed to provide the required computer support for the MODIS team leader. The TLCF will support MODIS Level 2 and 1B algorithm development, development of the higher-level processing shell, integration of team members' algorithms, generation of special products, algorithm testing, etc.

VERSION 1 SOFTWARE is used to correct any problems in the ß Version, to complete the operator interface, and to generate all messages.

VERSION 2 SOFTWARE is ready for launch. It will be used for final integration, testing of operations procedures, training of operations staff, etc.

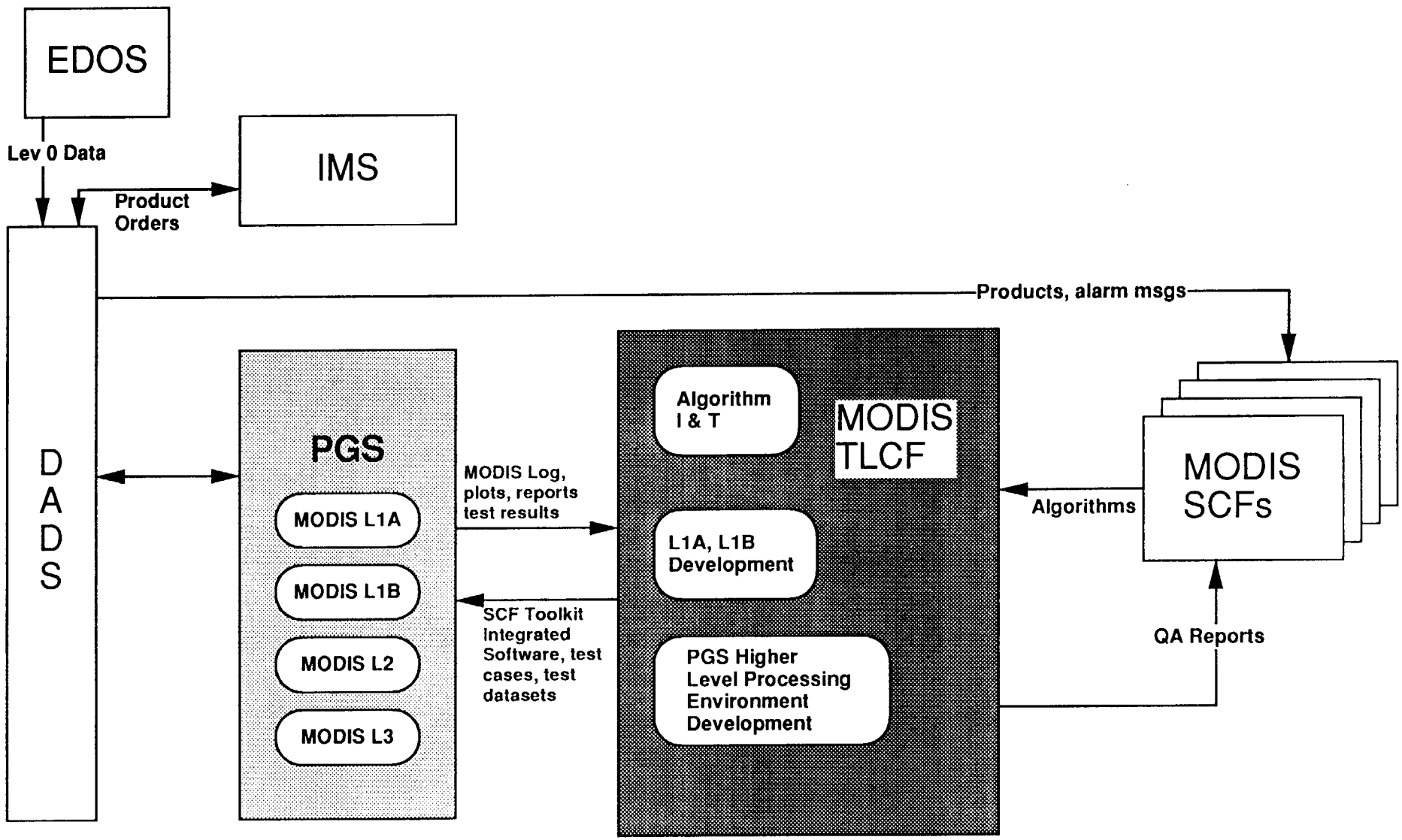


Figure 1. MODIS PGS Operational Processing Context

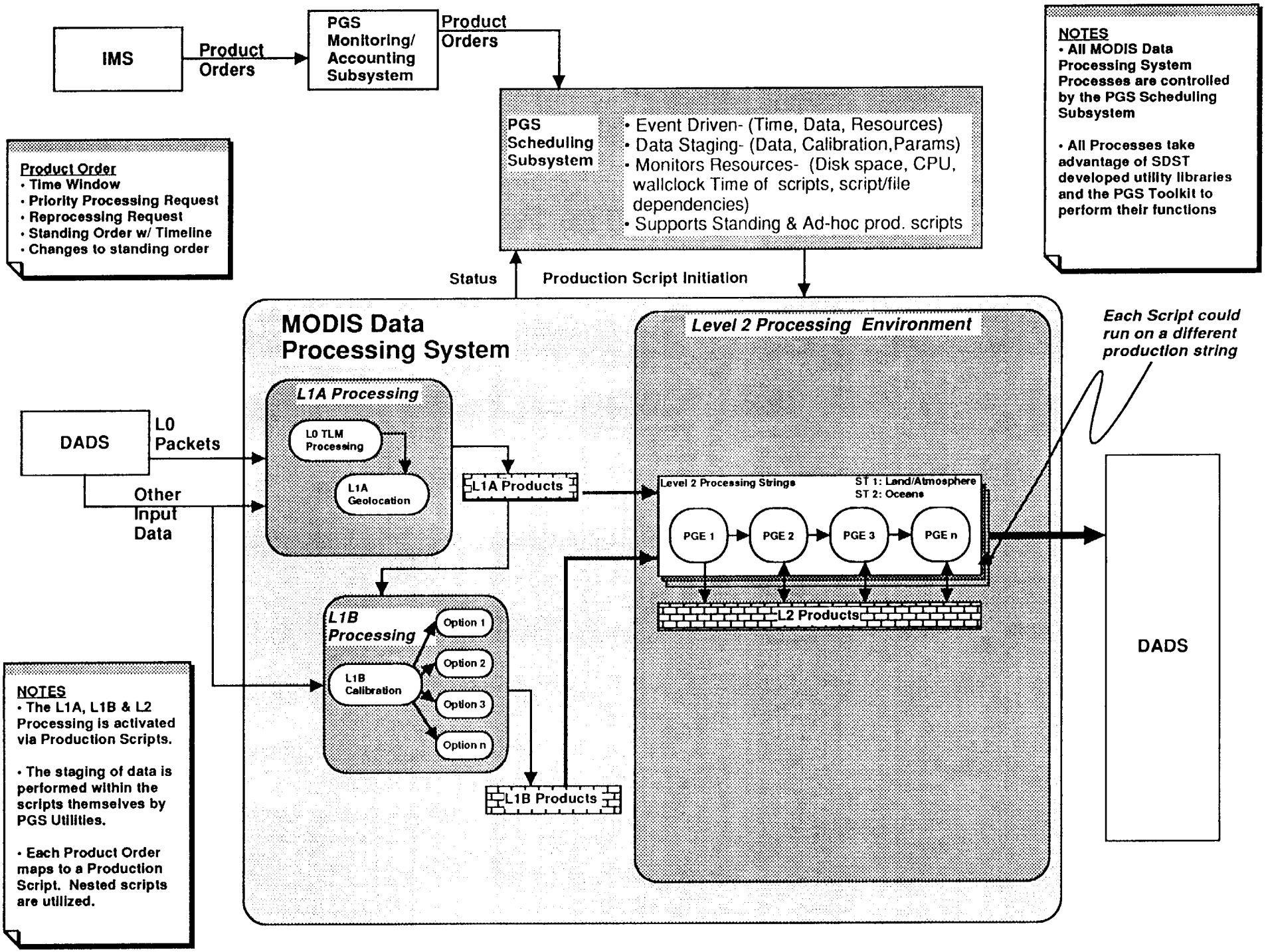


Figure 2. MODIS Data Processing System -Software Architecture-V0

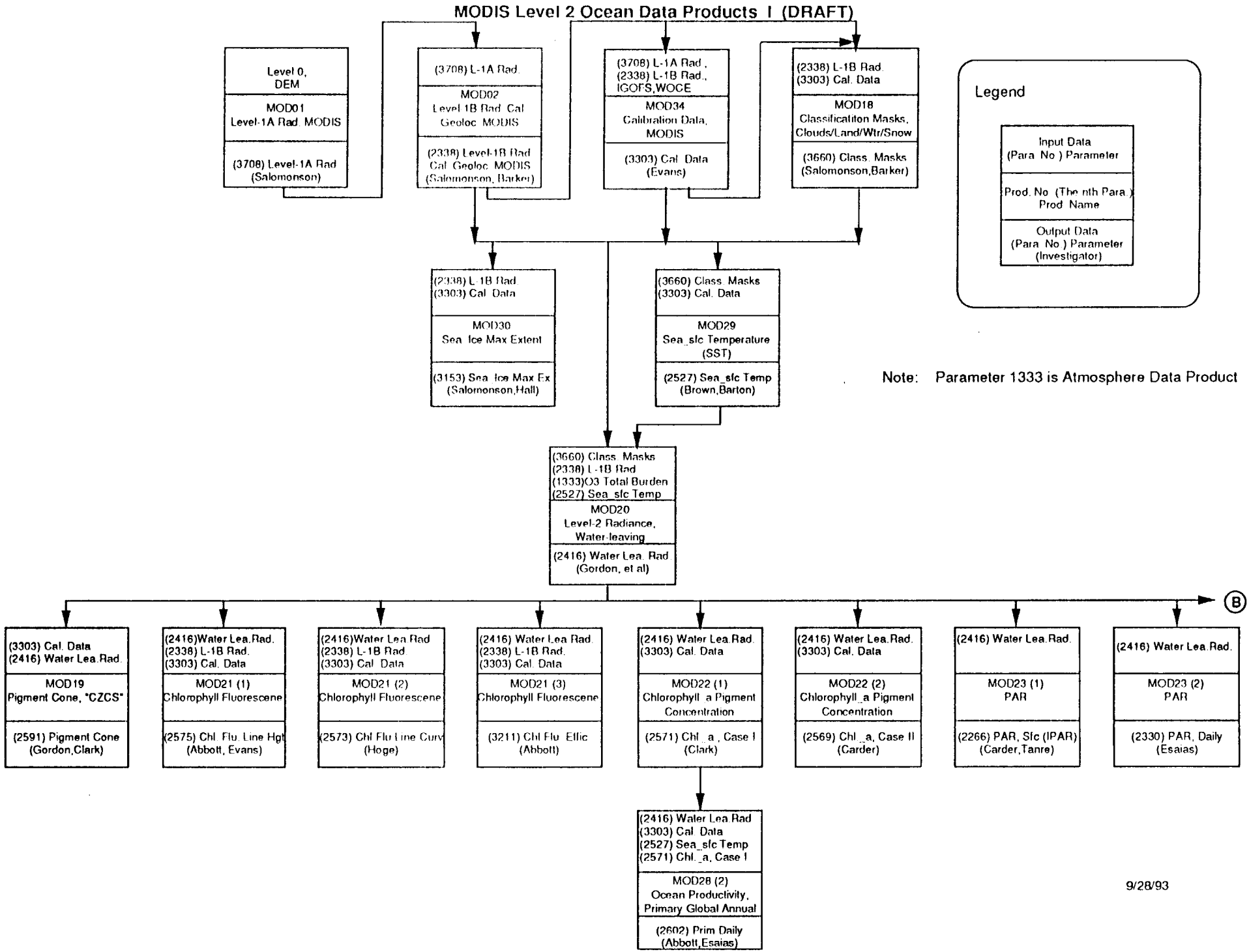
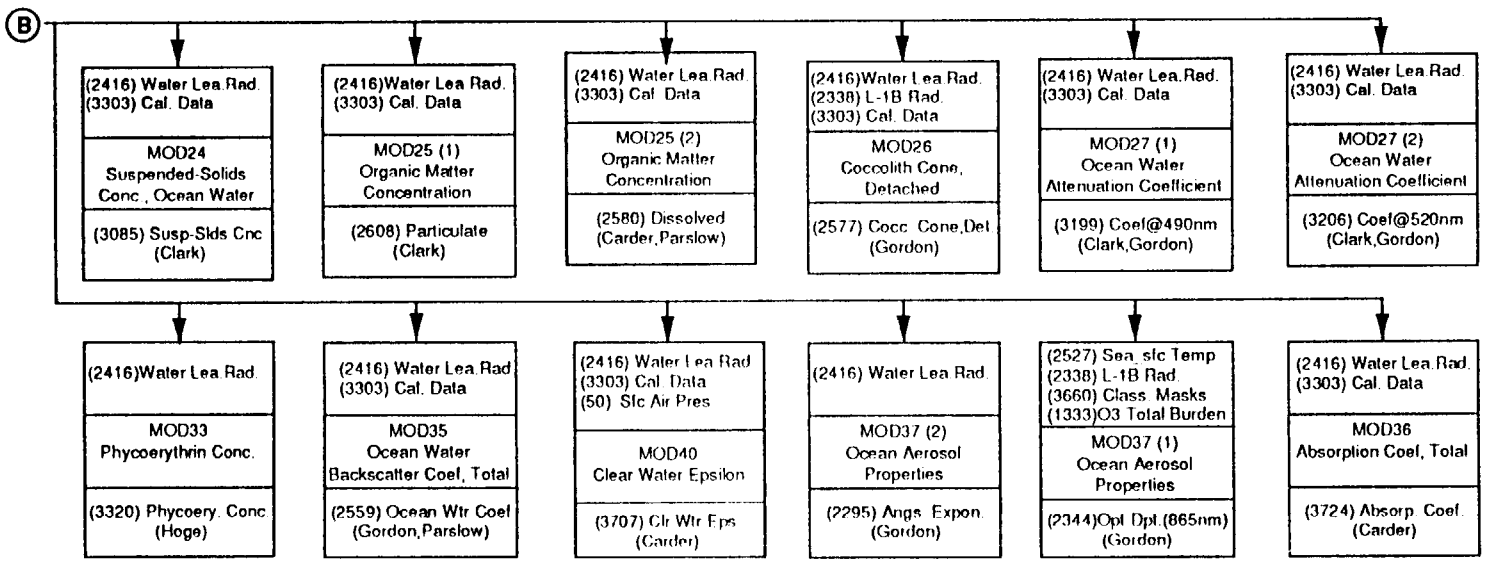


Figure 3. MODIS Level 2 Ocean Products (1 of 3)

Note: Parameter 1333 is Atmosphere Data Product



9/28/93

Figure 3. MODIS Level 2 Ocean Products (2 of 3)

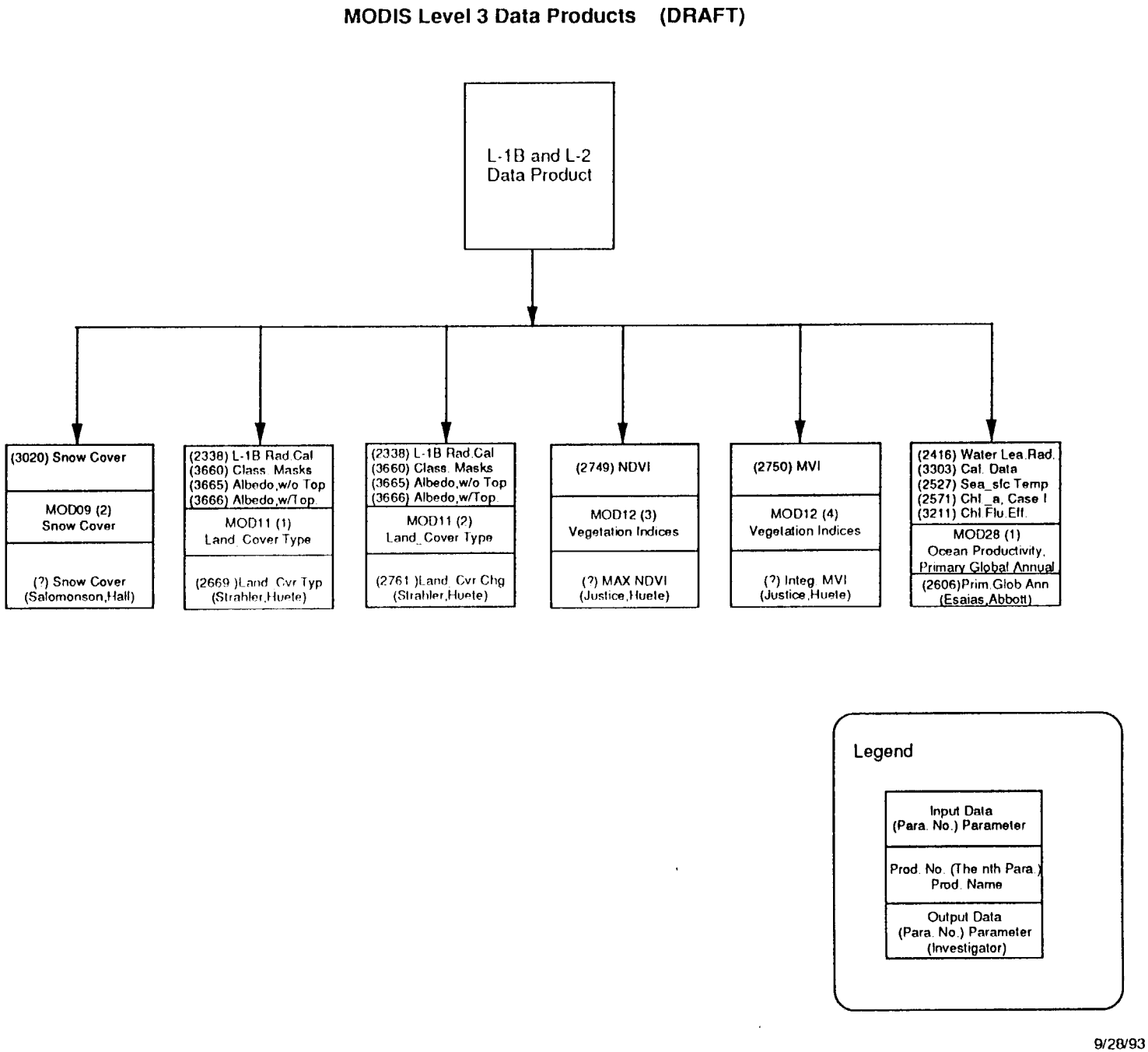


Figure 3. MODIS Level 2 Ocean Products (3 of 3)

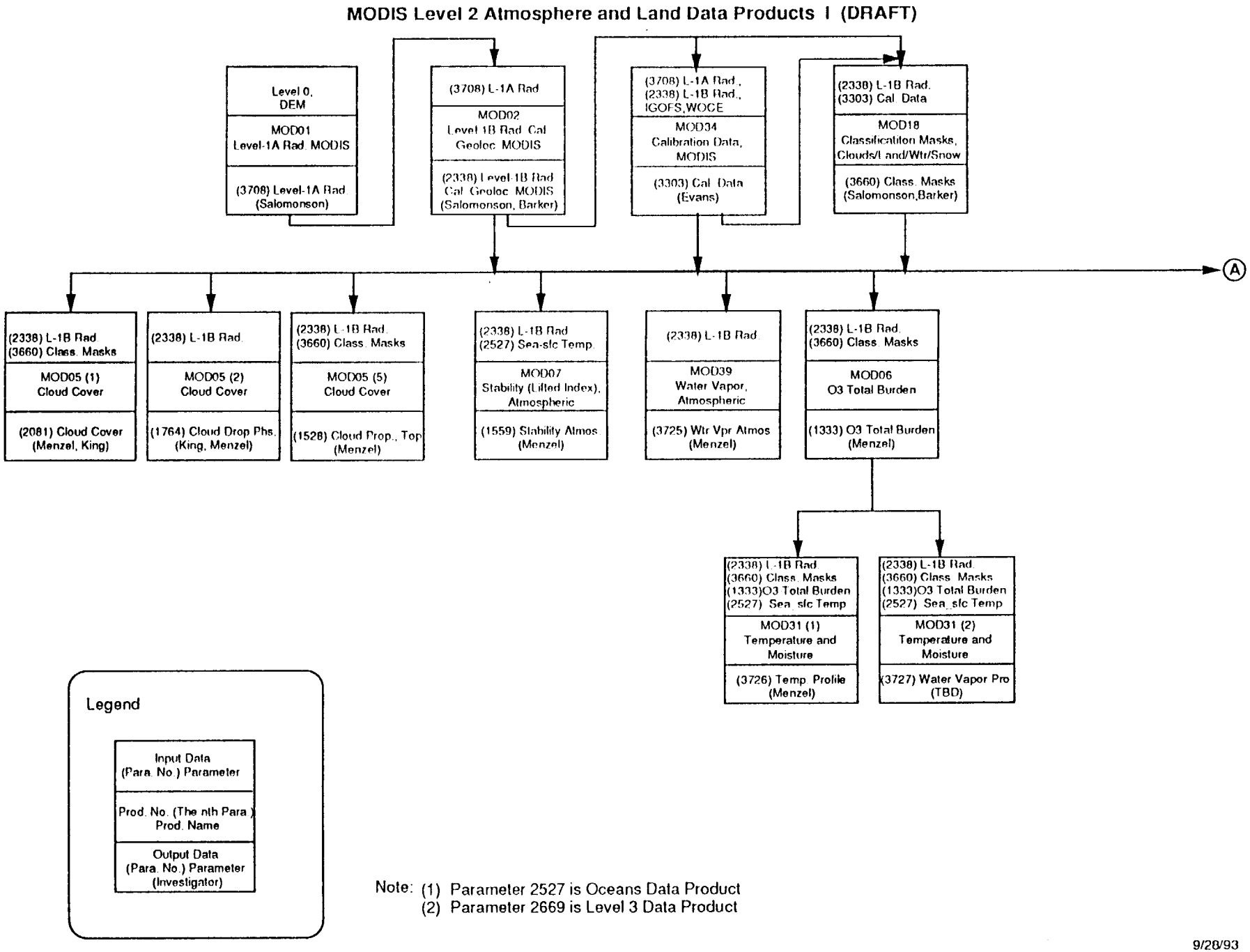


Figure 4. MODIS Level 2 Atmospheric and Land Products (1 of 2)

MODIS Level 2 Atmosphere and Land Data Products (II) (DRAFT)

Note: (1) Parameters 2527 and 2416 are Oceans Data Products
(2) Parameter 2669 is Level 3 Data Product
(3) Parameter 4101 is from ?

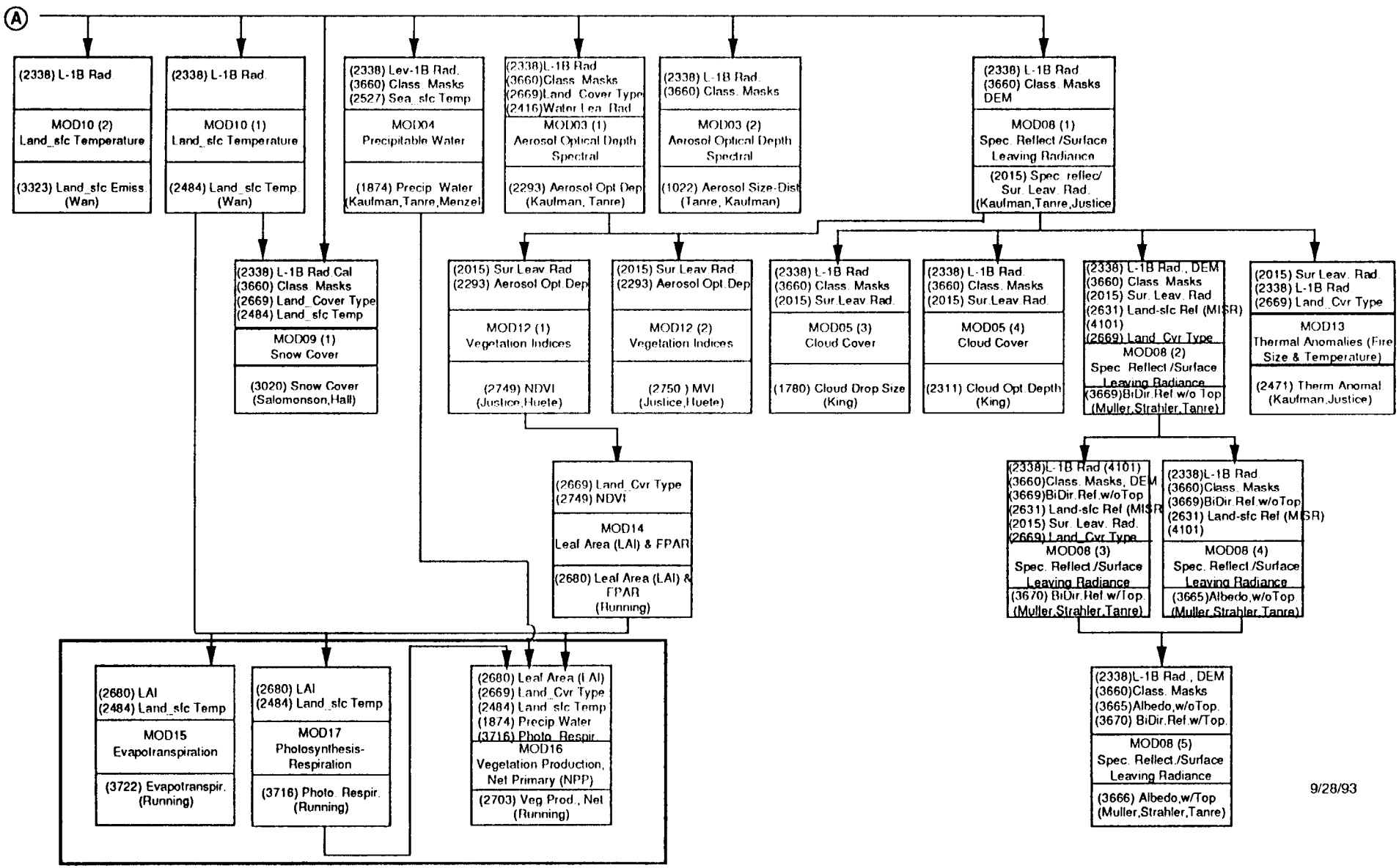


Figure 4. MODIS Level 2 Atmospheric and Land Products (2 of 2)

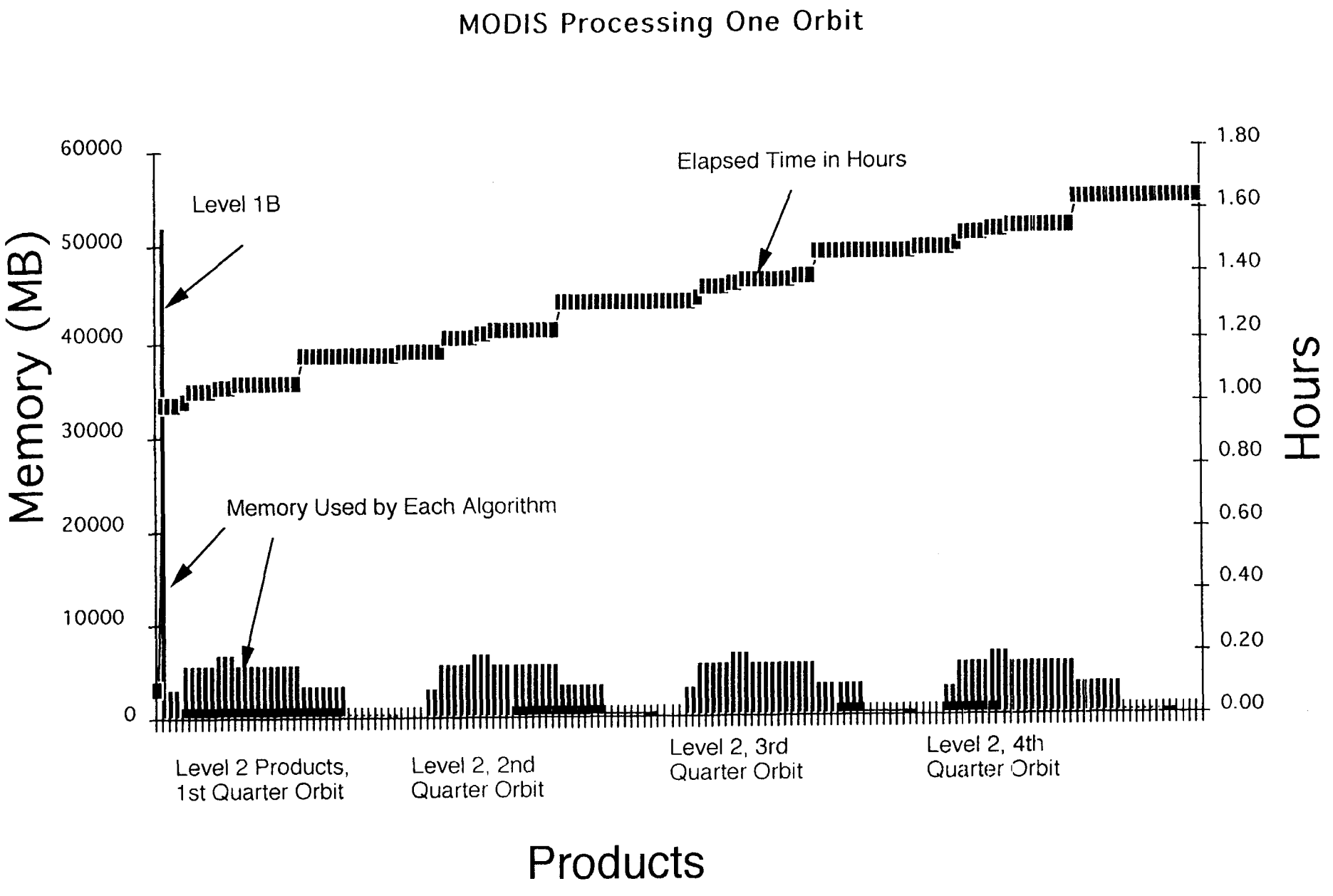
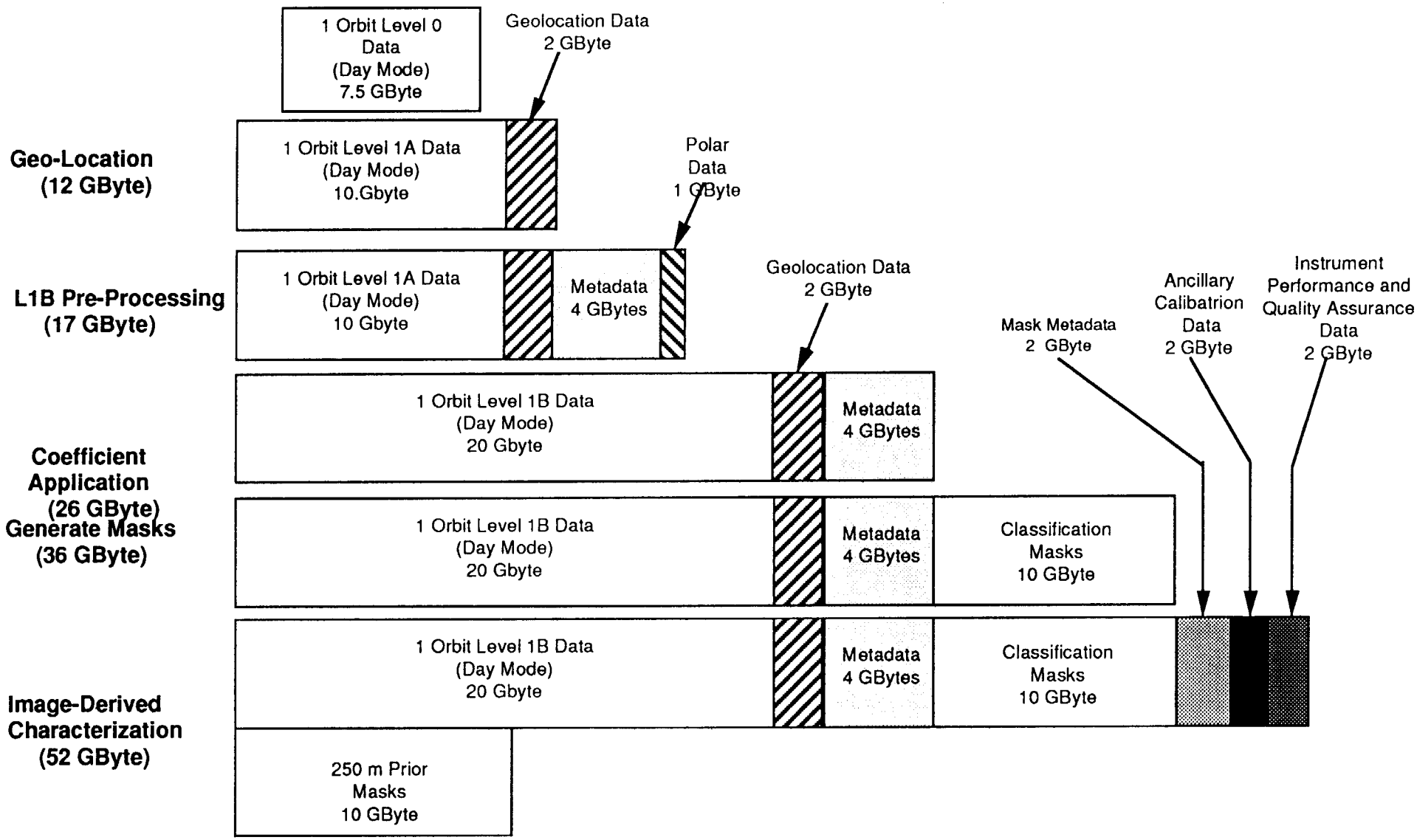


Figure 5. Time and memory required to produce MODIS products for one orbit of data on a computer system with 6,000GFLOP of sustained processing power.



A schematic representation of the Level 1B memory requirements. The major processing steps are shown in bold at the left ,with an estimate of the total memory required to complete that step in parenthesis. The major blocks of memory and their functions are represented by the rows of rectangles. Assuming that memory can be de-allocated and re-used throughout the processing of one orbit's worth of data, the amount of memory required is the memory required to perform the most memory intensive operation, Image-Derived Characterization.

Figure 6. Memory Requirements to Perform 1B Processing on 1 Orbit of Day Mode Data

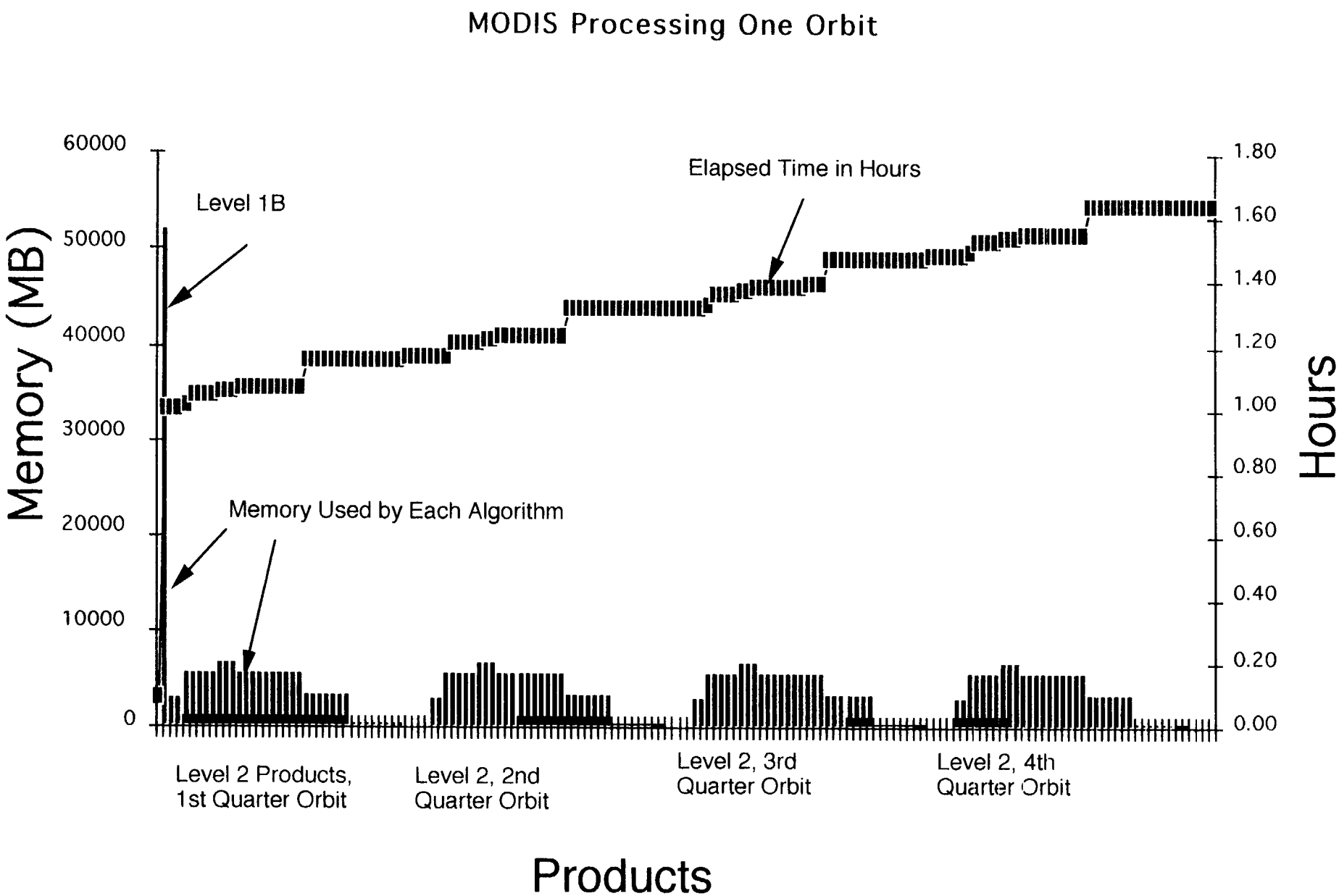


Figure 5. Time and memory required to produce MODIS products for one orbit of data on a computer system with 6,000GFLOP of sustained processing power.