NASA Technical Memorandum 100721

MODIS Information, Data, and Control System (MIDACS) System Specifications and Conceptual Design

(NASA-TM-100721) MODIS INFORMATION, DATA, AND CONTROL SYSTEM (MIDACS) SYSTEM SPECIFICATIONS AND CONCEPTUAL DESIGN (NASA) 122 p CSCL 05B

N89-15448

Unclas

G3/43 0187931

D. Han, V. Salomonson, J. Ormsby, P. Ardanuy, A. McKay, D. Hoyt, S. Jaffin, B. Vallette, B. Sharts, D. Folta, E. Hurley, and D. MacMillan

December 1988

NASA

NASA Technical Memorandum 100721

MODIS Information, Data, and Control System (MIDACS) System Specifications and Conceptual Design

D. Han, V. Salomonson and J. Ormsby
Space Data and Computing Division
Goddard Space Flight Center
Greenbelt, Maryland

P. Ardanuy, A. McKay, D. Hoyt,
S. Jaffin and B. Vallette
Research and Data Systems
Greenbelt, Maryland

B. Sharts and D. Folta
General Sciences Corporation
Laurel, Maryland

E. Hurley and D. MacMillan
Interferometrics, Inc.
Vienna, Virginia

NASA
National Aeronautics and
Space Administration

Information Management
Division

1988

PREFACE

The initial sections in this document discuss system level requirements, the overall operating environment in which requirements must be met, and a breakdown of the MIDACS into component subsystems, which include the Instrument Support Terminal, the Instrument Control Center, the Team Member Computing Facility, the Central Data Handling Facility, and the Data Archive and Distribution System, each of which handles a particular subset of required MIDACS functions. Appendix A contains a data dictionary which defines the data flow items shown in the specifications. The specifications include sizing estimates for the processing and storage capacities of each data system element, as well as traffic analyses of data flows between the elements internally, and also externally across the data system interfaces. This version of the System Specification and Conceptual Design follows an earlier release of a preliminary draft version. The specifications for the data system as a whole, as well as for the individual planning and scheduling, control and monitoring, data acquisition and processing, calibration and validation, and data archive and distribution components, do not yet fully specify the data system in the complete manner needed to achieve the scientific objectives of the MODIS instruments and science team. Indeed, the team members have not yet been selected and the team has not yet been formed. However, it has been possible to develop the specifications and conceptual design based on the present concept of EosDIS, the Level-I and Level-II Functional Requirements Documents, the Operations Concept, and through interviews and meetings with key members of the science community.

The study team is indebted to: Wayne Esaias, Chris Justice, and Joel Susskind for detailed information regarding the science requirements; Bill Barnes, John Barker, and Bruce Guenther for information regarding MODIS instrument concepts; H. Lee Kyle, and Dick Stonesifer for their insight into aspects of data processing, instrument control, and data storage; and to Al Fleig for his assistance in applying the guidelines being set forth by EosDIS.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

1. INTRODUCTION

Since a preliminary version of the Moderate Resolution Imaging Spectrometer (MODIS) specification is needed to support the integration of the MODIS and the High Resolution Imaging Spectrometer (HIRIS) data systems, the preparation of this specification was moved forward in the sequence of planned MODIS events. As a consequence, a number of supporting studies that were originally planned for completion before the specification was written are not yet finished. The information in this specification is the best available at the time when the specification was prepared.

1.1 BACKGROUND: POLAR ORBITING PLATFORM, EOS AND EosDIS

To provide improved sources of remotely sensed Earth data and to support the investigation of environmental and geophysical phenomena on a global scale, NASA is planning to launch a series of remote sensing satellites during the mid- and late-1990's. Four platforms are planned; two will be launched as U.S. satellites under NASA control, one will be launched as a European satellite under European Space Agency direction, and one will be under Japanese control launched under the direction of Japan's Science and Technology Agency. The program is known as the Earth Observing System (EOS) and it envisions using remote sensors in near-polar orbit to obtain complete Earth observation coverage. The first launch will be a NASA satellite designated the NASA Polar Orbiting Platform-1 (NPOP-1); launch of this satellite is anticipated around 1996.

Each platform will be equipped with a number of research instruments to perform the actual sensing functions. Instruments planned for NPOP-1 include the MODIS (the instrument of interest in this specification), and a similar instrument, HIRIS. Both instruments generate images of the Earth as seen in selected bands of the visible-light spectrum. To support Earth temperature studies, the MODIS instrument also senses selected bands in the infrared spectrum. HIRIS provides greater resolution of Earth features, but to avoid generating an unmanageable volume of data, it can only be used for targets of special research interest. MODIS routinely generates sensing data for the entire Earth and as a consequence, it generates a larger volume of output data than the HIRIS instrument.

Data from the MODIS instrument are combined with data from other on-board instruments and data from platform-specific sensors, called the transfer frame, and relayed from space to Earth using the facilities of the Tracking and Data Relay Satellite System (TDRSS). Once data are returned to the ground, header information attached to the data is interpreted and used to route the individual data blocks from the various on-board instruments to appropriate processing facilities for each instrument. Taken as a single entity, the data system that supports the entire complement of EOS instruments is designated as the Earth Observing System Data and Information System (EosDIS).

While it is thought that the various EOS instruments may well share some physical data processing facilities, the precise nature of any sharing that may occur is not yet defined. The initial designs for the MODIS data systems are being formulated as if each were a stand-alone system. The formal designation for the stand-alone MODIS Data System is the MODIS Information, Data, and Control System (MIDACS).

1.2 INTRODUCTORY DESCRIPTION OF MODIS INSTRUMENT

The baseline MODIS instrument design includes a gimballed mirror that allows the instrument to sequentially scan an Earth observation area as the instrument is carried along on the orbiting platform. To achieve complete coverage of the Earth, the instrument scans in the cross-track direction.

However, analysis has shown that a simple instrument that looks downward and scans to the sides cannot provide all required observation data. Specular reflections (sunglint) obscure observations in some instrument-sun geometries. Also, observations of a given region are needed at several observation angles to help determine atmospheric parameters and the effect of the atmosphere on Earth observations. Furthermore, the Bidirectional Reflectance Distribution Function (BRDF), which measures radiation reflected from the Earth as a function of the zenith angle of the illumination source (the sun), the zenith angle of the observation, and the azimuthal separation of the illumination source and observation, is needed for as many observation angles as possible.

These considerations have led to the partitioning of MODIS functions among two instruments, one of which looks directly downward and scans to the sides as described above. This instrument is designated MODIS-N (nadir). The other instrument is rotatable or tiltable about an axis perpendicular to the track direction and is designated MODIS-T (tilt) The MODIS-T instrument can be tilted fore or aft in the along-track direction by as much as plus or minus 50 degrees from nadir. For regular observations, the instrument tilt angle is held fixed, and scanning is to the sides much as for the MODIS-N instrument. The MODIS-T tilt angle can be periodically adjusted in fixed increments by remote command from the MODIS Instrument Control Center.

Spectral characteristics are important for the interpretation of remote sensing data. The spectral bands needed for the MODIS-T instrument are regularly spaced and can be generated using diffraction grating techniques. Sixty-four spectral bands in the wavelength region between 0.4 and 1.04 microns will be observed by MODIS-T. Each band is 10 nanometers in width. The bands for the MODIS-N instrument are not regularly spaced and are not of uniform bandwidth. This instrument observes 31 or, optionally, 40 bands in the region between 0.4 and 14.2 microns. Altogether, the baseline design includes up to 64 + 40 = 104 spectral bands for the MODIS instrument.

At nadir, the resolution of the MODIS-T instrument is 1 km. The MODIS-N design accommodates several resolutions (250 m, 500 m, and 1 km), which are used with a number of spectral bands that are useful in interpreting terrestrial data. Terrestrial features can vary significantly on this distance scale. Atmospheric and oceanic features are normally homogeneous over larger distances; 1 km resolution is adequate for those spectral bands used to examine atmospheric and oceanic features. The majority of the MODIS-N bands operate at 1 km resolution.

1.3 OUTSTANDING ISSUES AND UNCERTAINTIES

With an expected launch of the first NASA Polar Orbiting Platform (NPOP-1) in late 1996, we have some eight years to wait before the first data taken in flight by the MODIS-N and MODIS-T instruments is received and processed. Before this milestone is achieved, the MODIS data system must be fully designed, built, and tested. At this stage, however, the Phase-B design studies for the instruments are not yet completed, the members of the MODIS science team have not yet been selected, and the structure and concept of the EosDIS is still evolving. It is not surprising that there exist some uncertainties, at this time, in specific areas of the MIDACS operations concept. The uncertainties fall into two categories: those driven through a lack of specific definition of the EosDIS environment, and those driven due to a lack of specific knowledge concerning the MODIS instruments' capabilities and the science team members' proposed research objectives.

1.3.1 Real-Time Monitoring of MODIS Data

Engineering and science data taken by the MODIS instruments, as well as selected platform ancillary data, must be monitored in the Instrument Control Center (ICC). Ideally, all data of possible utility would be available in real time with a 100% duty cycle. However, the primary downlink will be through the TDRSS, and it is anticipated that the platform will have access to the TDRSS for only portions of each orbit. Each TDRSS access will be scheduled well in advance of the actual contact. While there will be an alternate, multiple-access low-rate data path for the transmission and verification of emergency commands, the TDRSS link will be the sole path for downlinked data destined for the monitoring function within the ICC. Engineering and science data will be stored on board the platform for playback and downlink at the time of the next TDRSS contact.

Under these circumstances, it will not be possible to monitor the instrument with a 100% duty cycle in real time. With priority-playback processing at the DHC, the data may be monitored with a 100% duty cycle shortly after reception by the DIF. It is anticipated that the platform and instrument ancillary data (engineering/housekeeping) will be packetized separately from the science data, and these packets will be automatically routed to the ICC.

There is a requirement that science team or IOT members located in the ICC monitor four channels each of MODIS-N and MODIS-T data in real time, and that the choice of the channels being monitored be selectable in real time. There are three possible methods for achieving this:

a. The data are buffered on board the platform by the instrument data system during each scan. Data from different channels are packetized separately. All MODIS data is treated as either real-time or priority-playback data by the DHC. A split pipe flow may exist, with Level-0 processing functions performed twice for some MODIS data, allowing the MODIS ICC to receive the data quickly, and the Central Data Handling Facility (CDHF) to receive the data after the gaps in coverage have been filled. Within the ICC, the headers of the Level-0 data packets are examined, and data from selected channels only are ingested into the monitoring data base; non-selected data are lost.

b. As in the previous example, the science data is buffered and reorganized prior to the generation of data packets. However, in this scenario, the ICC uplinks to the on-board data system in real-time the selected subset of channels to be monitored. The on-board data system installs the real-time/priority-processing designation and ICC address in the header of the packets of data required for monitoring. Upon acquisition at the DHC, the designated packets of data are delivered immediately to the ICC, perhaps with only partial Level-0 processing. Within the ICC, the data are received at a relatively low rate, unpacked, and inserted into the monitoring database. Data from unselected channels are not available with this timeliness, but may be analyzed at a later time (24+ hours after observation) upon completion of processing at the CDHF.

c. In this case, we assume that the on-board processor does not buffer the MODIS data during a scan (on the order of ten megabytes for MODIS-T). Each packet of science data then contains data from many spectral channels multiplexed together. All MODIS data is treated as either real-time or priority-playback data by the DHC. As with the first example, there is no interactive, real-time selection of channels to be monitored between the ICC and the on-board processing system. Because of the absense of on-board data sorting by channel, the science data required for monitoring at the ICC must be collated from observations contained on many packets. All MODIS data packets are required, and many aspects of the Level-0 and -1A processing are required to unpack and reorganize the data into a form useful for monitoring. The size of processor required will exceed that normally associated with typical control and monitor functions and may reside on the CDHF.

1.3.2 Implementation of Algorithms for Standard Product Processing

The implementation of algorithms for standard product processing is to be specified in a future version.

1.3.3 Capabilities of the On-Board Processor

The capabilities of the on-board processor are discussed as an interface to the ground data system in Section 3.3.1.

1.3.4 Non-MODIS Instrument Data Availability

Non-MODIS instrument data availability from other EOS and non-EOS sources will be specified in a future version.

1.3.5 Near-Real-Time Data Communication

Communication of near-real-time data from the CDHF/Data Archive and Distribution System (DADS) to field experiments will be specified in a future version.

1.3.6 Data Processing Operations Concept

The level of meaningful detail in the processing operations concept is limited by our knowledge of the processing algorithms. Details of the processing scenarios and concepts such as the logical data record, Earth location, and calibration accuracy will evolve with our understanding of processing algorithms and end-user requirements.

1.3.7 Capabilities and Interfaces of the MIDACS with the DHC

The current capabilities of the DHC are to support the real-time, near-real-time, and routine science data processing timelines that are required by the MIDACS. The DHC will process some data for real-time delivery to the MIDACS (ICC or CDHF) for instrument monitoring. The DHC will also pass the priority playback data to the MIDACS, but it is still uncertain how this will be accomplished. This DHC/MIDACS interface has not been fully clarified.

1.3.8 On-Board Processing

The impact of the ICC workload for planning, scheduling, control, and monitoring functions will be increased if the ICC has the added responsibility to uplink commands for the selection of channels to be monitored. To accomplish this, the ICC personnel would have to work more directly with the EMOC/PSC/NCC for the scheduling of TDRSS contact. Also, the increase of work due to many requests for data or the rapid selection of channels would pressure ICC personnel to make quick decisions without the ability to follow defined procedures and guidelines for the checking and verification of all command loads.

1.3.9 Storage of the MODIS Science, Engineering, and Ancillary Data

The MODIS science, engineering, and ancillary data will be stored at the MIDACS DADS.

1.3.10 Implementation of DARS for Real-Time Field Experiments or Instrument Calibration

The number of requests is unknown at this time. Any request for real-time support would effect the TDRSS scheduling, and therefore need to be coordinated with the appropriate facility to satisfy the request.

1.3.11 Hierarchy of Requests

The hierarchy of priorities for the request for the support of field experiments or real-time monitoring and the availability of workstations at the ICC to provide support is to be defined.

1.3.12 Command Tracking of TOOs and Real-Time Requests

The ICC must verify the command load for all commands uplinked to the MODIS before they are transmitted to the EMOC. The procedure and timeline for doing this must be defined.

1.3.13 MODIS/HIRIS and Joint Scheduling With Other Instruments

The MODIS and HIRIS will have some degree of interaction, not only for the planning and scheduling phase, but also for the impacts of requests for data in the real-time and near-real-time for the support of field experiments for MODIS and coincident observations for HIRIS. A prioritized scheduling process must be developed.

1.3.14 MIDACS Support Personnel

The role of the product support analysts and the system operators with respect to the production of standard data products must be defined.

1.3.15 Error Correction/Grade of Service

The level of MODIS data encoding for error correction is TBD. At a bit error rate of 10^{-12}, on average only one bad MODIS bit will be encountered every day. However, at a bit error rate of only 10^{-8}, 10,000 bad bits will be encountered daily. The packets with uncorrectable errors should be flagged as such by the DHC. The MODIS science team may require a bit error rate of no worse than 10^{-8} or Grade II service in data communication. (A bit error rate of 10^{-10} is preferred and may be required by the science team.)

1.3.16 Data Coverage

Because MODIS data will be used to produce products with global coverage, missing packets will degrade the quality of the final product. Completeness to only the 99% level would result in a loss of 15 minutes of coverage; at 6.5 km per second, this is a 51 or 5600 km swath along the orbit. The MODIS science team may require coverage to no less than 99.9%. Systematic coverage loss may be subject to different requirements than random coverage loss, perhaps by a factor of ten.

1.3.17 CDHF Element Specification

Data processing estimates are highly dependent upon horizontal resolution and coverage assumed for the products. If spatial resolution for a product was relaxed from $(1\ km)^2$ to $(5\ km)^2$, the processing load would change by a factor approaching 25.

More data processing scenarios are needed. In fact, as the science team progresses with its plans for MODIS, there will eventually be a data processing scenario for every product. Between now and that time, the more realistic and well thought out scenarios that can be developed the better. As new scenarios are included in the processing estimation, the categories of algorithms will increase while the number of products in each category will decrease, and the accuracy of the processing estimate will increase. This, of course, assumes that the processing scenarios are not only realistic and accurate, but are truly relevant to the ultimate utilization of the MODIS observations.

1.4 STRUCTURE OF THIS DOCUMENT

The initial sections in this document discuss system-level requirements, the overall operating environment in which requirements must be met, and the break out of the MIDACS into component subsystems, each of which handles a particular subset of required MIDACS functions. Section 2 is entitled "MIDACS System Level Requirements," Section 3 is the "MIDACS Operating Environment," and Section 4 is entitled "MIDACS Architecture".

These sections are followed by a series of "mini-specifications," each devoted to a description of a particular MIDACS subsystem. Section 5 is entitled "The Instrument Support Terminal," Section 6 deals with "The Instrument Control Center," Section 7 addresses "The Team Member Computing Facility," Section 8 specifies requirements for "The Central Data Handling Facility," and Section 9 addresses "The Data Archive and Distribution System."

The Appendices contain relevant information that is not properly part of the formal specification. Appendix A contains a Data Dictionary defining the data flow items shown in the specification.

2. MIDACS SYSTEM LEVEL REQUIREMENTS

This section contains MIDACS system level requirements. The requirements in this section generally describe the functions and performance characteristics that the desired data system must achieve. Hardware performance requirements derived from fundamental system requirements are treated as design parameters and are treated in separate sections dealing with system and subsystem design. This section simply describes the overall functions and performance of the desired instrument data system.

Section 2.1 is entitled "Functional Requirements" and it generally contains qualitative definitions of the functions that the system must perform (i.e., what it is that the system must do). Section 2.2 is entitled "Performance Requirements"; it generally specifies the quantitative performance the system must achieve (i.e., how much data must be processed, how quickly data must be processed, etc.).

2.1 FUNCTIONAL REQUIREMENTS

The MIDACS data system performs three basic functions: it provides the basic data system structures that will be used to direct the overall MODIS project effort to achieve desired scientific goals, it provides for direct on-board control of the MODIS instrument, and it processes, stores, and distributes the data returned by the instrument. Each of these functions is addressed individually in the sections that follow.

2.1.1 Direct Overall MODIS Project Efforts

The direction of MODIS project efforts to attain scientific objectives is the responsibility of the MODIS Science Team Leader. The Science Team Leader is equipped with an Instrument Support Terminal (IST) that provides remote access to instrument monitoring information available from the ICC. The IST provides the Science Team Leader with instrument monitor and control links needed to monitor and control the scientific content of MODIS instrument operations.

The Science Team Leader, working with the Science Team Members, is also responsible for all science-related matters pertaining to the processing of instrument data to generate data products. He monitors the generation of MODIS data products using the communications facilities of the Team Member Computing Facility (TMCF). The science team approves all standard algorithms and quality checks used with the data and is held accountable for the quality of MODIS Data Products produced. Data products are released for general distribution only at the team's direction.

The overall operation of the EOS will be directed by the EOS Project Manager. A MODIS Instrument Operations Team (IOT) Manager will be appointed to direct day-to-day MODIS instrument control operations. Working with the Science Team Leader, the IOT Manager will manage MODIS instrument operations to achieve scientific goals and, working within guidelines provided by the Science Team Leader, the Science Data Processing Manager will direct the generation, storage, and distribution of MODIS data products.

2.1.2 Control Instrument Operation

Instrument control necessarily includes the receipt and processing of instrument monitoring data as well as the actual generation of instrument command sequences. The sections that follow first discuss monitor data requirements, and then the instrument commands that must be supported by the system.

2.1.2.1 *Monitor Instrument Operations and Performance*

2.1.2.1.1 *Required Instrument Monitoring Information.* Sufficient instrument science and engineering data will be monitored to determine the state-of-health of the instrument and to transmit safing commands if required. For example, the MIDACS shall process and display the following information for use in monitoring the MODIS instrument[1,2]:

---

[1]This list is doubtlessly wrong and/or incomplete and will be updated in the next version.

[2]This list includes monitoring functions that are generated in real-time for use in the ICC.

MODIS Engineering Data

Power status (On/Off)[3]
Observing status (Daytime/Nighttime)[4]
Data quantization level (Low/High)[5]
Instrument component temperatures[6]
Instrument safing status (Safe/Loose)
Calibration source status[7]
Instrument self-test results[8]
Star-tracker status[9]

Science data

Tilt angle of MODIS-T instrument
Four spectral channels of MODIS-N data[10]
Four spectral channels of MODIS-T data

MODIS science data may be routed to the ICC as it is received by the DHC (i.e., in the DHC's priority playback mode). The ICC may then select the subset of data of interest for monitoring. The ICC may also periodically request a sample data product from the CDHF. The science data for the MODIS-N and MODIS-T instruments shall be user

---

[3]Does the instrument support a standby status (power on but no observations being generated?

[4]It is assumed that MODIS spectral channels cannot be individually switched on and off and that only two spectral channel activation states are possible:

Daytime - all spectral channels active
Nighttime - only thermal channels active

[5]It is thought that more than one digitization scale may be used to allow the instrument to better adjust to changing intensity conditions such as might occur when switching between land and sea observations. Would such a change of scale be applied uniformly to all spectral bands or would the scale be individually selectable for each spectral band?

[6]The number and locations of temperature monitoring points have not yet been specified.

[7]Not yet specified.

[8]Does the instrument have a self-test capability? What items can be tested? Is the built-in test capability intelligent, e.g. can it declare alarm status and initiate instrument safing sequences? How are test results transmitted to the ground? Is there an instrument status word that conveys the results of the built-in test? How often would the built-in test sequence by repeated?

[9]Whether the MODIS instrument will have it's own unique star-tracker is not determined at this time.

[10]Level-1A, -1B, and -2.

selectable in real time (i.e., the user shall be able to examine any four channels of MODIS-N data and any four channels of MODIS-T data without perceptible delay[11]).

2.1.2.1.2 *Required Platform Monitoring Information.* The basic platform data provided to the MIDACS by the Data Handling Center (DHC) shall be processed for convenient use in MODIS instrument control. Information to be generated might include past, present, or projected future instrument views of the Earth presented as a set of instrument scan lines and their envelope delineated on a geographic map. Each scan line can be fixed with a time designator that indicates the time when that scan was or will be achieved.

The geographic map might include political boundaries, major cities, and other Earth features adequate to allow the ready determination of the satellite position with respect to features that may be of observational interest.

2.1.2.1.3 *Displays.* Instrument engineering information (e.g. Power On/Off) shall be displayed in standard control center monitor format on standard devices (e.g., monitor pages, graphs, strip charts).

Scene data shall be displayed on a high-resolution color monitor, instrument monitoring displays shall reflect a one-to-one mapping of picture elements (pixels) as generated by the MODIS instrument and a rectilinear display matrix used in the ground data system to present the instrument data (i.e., no correction for instrument perspective is required in instrument monitoring displays[12]).

Cross-track scan lines generated by the instrument shall display directly as horizontal lines on the monitoring display and along-track pixels shall display as vertical lines. The horizontal display scale shall be selectable to include the option of displaying the entire cross track scan in the width of the display screen. The vertical display shall be correspondingly adjusted to maintain true perspective as determined at the MODIS instrument. In addition, the system shall support user selection of magnified or "close-up" images with integer magnifications of n = 2,4,8.

The MIDACS shall support the generation and superposition of outline map overlays that reference display scenes to geographically-fixed Earth locations.

2.1.2.1.4 *Real-Time Availability of Monitor Data.* MODIS monitoring will typically use data played back from the tape recorders. The MIDACS architecture, however, will not preclude the reception and processing of science data in real-time directly from the MODIS instruments. The MODIS instrument engineering and housekeeping data will be returned to the ground in real-time mode during TDRSS contact periods. The data to be displayed at the ICC can then be selected on the ground. Real-time instrument data is also recorded during TDRSS contact periods and will subsequently be routed into Routine Processing channels along with other recorded data.

2.1.2.1.5 *Retention of Monitor Data.* The entire stream of MODIS Engineering Data described in 2.1.2.1.1 shall be retained in storage for the life of the MIDACS. The science data described in that paragraph shall be retained until duplicate scientific data

---

[11]This includes Near-Real-Time data available from the DHC.

[12]Such a display would make it very difficult to superimpose an outline map showing geographic features over the display. Perhaps the image must be rectified to geographic coordinates?

is available from normal data processing channels. Science data generated for monitoring purposes may be discarded if redundant standard data products are available.

2.1.2.2 *Generate Instrument Commands*

2.1.2.2.1 *Nature of Commands.* Generally, there are two ways of executing an instrument operation: (1) by commands loaded into the on-board processor and executed autonomously according to a time schedule; and (2) by single event ground command while in real-time contact with the spacecraft. Instrument functions that can be remotely controlled through the facilities of the MIDACS could include the following:

Instrument power (On/Off)
Observation status (Daytime/Nighttime)
Instrument safing (Safe/Loose)
Calibration source activation[13]
Built-In-Test initiate[14]
Star tracker commands[15]
Tilt angle for MODIS-T instrument

2.1.2.2.2 *Procedures.* Stored instrument commands are generated in a two-phase sequence: the first step is to plan and schedule instrument operation and the second is to actually generate instrument command uploads. Functional requirements for planning and scheduling and for generating command uploads will be discussed separately.

Plan and Schedule Instrument Operations

The planning and scheduling function must provide a model of instrument operation to support these activities:

a. Determine the feasibility of a data acquisition objective;

b. Determine instrument resource requirements for a proposed observation sequence;

c. Predict instrument status throughout an observation sequence.

Planning and scheduling is iterative by nature and will include operator interaction to generate new trial schedules.

Instrument Command Modes

The MIDACS shall support a stored command mode and a real-time command mode. For routine operations, command sequences will be uploaded to the on-board satellite data system and stored for execution according to the execution time tag attached to each command. The system shall also support the generation and upload of real-time instru-

---

[13]Not sure what is involved here.

[14]What corrective measures can be taken if a failure is detected? Corrective measures that can be initiated under ground control must be supported by the MIDACS.

[15]See Star-Tracker footnote in 2.1.2.1.1.

ment commands that are executed in real-time as the commands are received at the platform.

During TDRSS contact periods, real-time commands may be uploaded using the normal TDRSS command uplink. For emergency commands, NASA may also provide an alternative low-data-rate command link that will be available for virtually any satellite location. The MIDACS may support emergency instrument control using this alternative link. However, this may not be true for platform emergencies. The NCC is charted to interrupt an active TDRSS schedule to accommodate emergencies through the TDRSS. Real-time instrument control would normally be used only in an emergency to verify instrument operation or to redirect instrument observations to "targets of opportunity" that arise unexpectedly in the course of previously planned operations.

Generate Instrument Command Uploads

Command uploads may support either stored-command or real-time operation. Command upload sequences shall be annotated to provide the uninitiated user with a description of the function performed by each command. These descriptors go to EMOC, but will be stripped out prior to upload to the PSC. Prior to transmission, command uploads shall be validated using the instrument model described in 2.1.2.2.1.

2.1.3 Produce, Store, and Distribute Data Products

2.1.3.1 *Description of Standard Data Products*

2.1.3.1.1 *Major Product Designàtions.* The following designations indicate the general nature of MIDACS data products.

Level-0. Reconstructed, unprocessed instrument data in the same format and time order as produced by the instrument, with duplicate data removed. These are the data supplied to the MODIS data system by the Data Handling Center (DHC), after correction for transmission errors. Level-0 data are not available as a standard MODIS data product.

Level-1A. MODIS instrument data that have been reformatted and appended with ancillary and engineering data needed to complete processing. They are reversible to Level-0.

Level-1B. Level-1B data are irreversibly processed from Level-1A. Earth locations and radiometric calibration has been applied to generate observed Earth radiances. (MODIS Level-1B data may in fact be reversible to Level-0, given some ancillary data off the Level-1A product.)

Level-2. Geophysical or environmental parameters derived from Level-1B data and retaining the resolution of Level-1 data. Candidate Level-2 data products are listed in Table 1. The algorithms required to generate these products are in various stages of development; many are not completely specified at this point. Table 2 provides a summary of candidate Level-2 products.

Level-3. Observed radiances or other geophysical parameters geometrically rectified and resampled onto space-time grids. Images from several satellite passes may be mosaicked to generate Level-3 products.

Table 1
Candidate Level-2 Data Products

LAND SURFACE COMPOSITION
- Soil types
- Rock types
- Available soil moisture (radiance ratios)
- Soil thermal inertia
- Soil particle size

LAND SURFACE BIOLOGICAL ACTIVITY
- Reflected near infrared radiation (Wm^{-2})
- Reflected photosynthetically active radiation (Wm^{-2})
- Leaf area indices
- Plant and crop types
- Ecological zone classifications
- Plant temperature (K)
- Plant productivity
- Plant stress indices
- Photosynthesis rate
- Canopy state

OCEAN CIRCULATION
- Velocity vectors (cm/sec)
- Areal extent of eddies (km^2)
- Sea surface temperature (K)
- Suspended sediment concentration ($mgcm^{-3}$)

OCEAN AND LAKES BIOLOGICAL ACTIVITY
- Primary production rates
- Pigment concentration or groups ($mgcm^{-3}$)
- Suspended sediment concentration ($mgcm^{-3}$)
- Gelbstoffe concentration ($mgcm^{-3}$)
- Chlorophyll concentration
- Phalophatin concentration
- Marine humis concentration
- Fulvic acid concentration
- Species composition
- Phytoplankton biomass
- Chlorophyll fluorescence

AEROSOLS (over water)
- Optical depth (N2 wavelengths)
- Aerosol size distribution
- Aerosol height distribution

CLOUD PROPERTIES
- Cloud top height (mbs)
- Cloud cover (%)
- Cloud albedo(%)
- Cloud top temperatures (K)
- Cloud emissivities ( )
- Cloud radiative forcing (Wm^{-2})
  - Longwave

Shortwave
Net
Precipitation (mm/day)
Identification of cloud free areas for HIRIS, TIMS, etc.

EARTH RADIATIVE BUDGET
Planetary albedo (%)
Surface albedo (%)
Surface emissivity ( )
Outgoing longwave radiation (Wm^{-2})
Surface longwave radiation (Wm^{-2})
Upward
Downward
Net longwave at surface (Wm^{-2})
Net longwave loss from atmosphere (Wm^{-2})
Latent heat flux (Wm^{-2})
Sensible heat flux (Wm^{-2})
Heat flux into Earth (Wm^{-2})
Net radiation at atmospheric top (Wm^{-2})
Bowen ratio
Rate of entropy production (Wm^{-2}/K)

SURFACE TEMPERATURE (K)

ATMOSPHERIC TEMPERATURE AND COMPOSITION
Temperature at N levels (K)
Specific humidity at N levels ($gmcm^{-3}$)
Ozone content (Dobson units)
Carbon dioxide content (ppm by volume)
Total precipitable water (cm)

SNOW AND ICE COVER
Snow and ice extent
albedo
age
emissivity
surface temperature
Polynya area

SEA-ICE COVER
Sea-ice extent
albedo
age
emissivity
surface temperature

Table 2
Summary of Number of Level-2 Products

	Count		Coverage		Net
Global	26 + 2*(N)	= 56	* 1.0	=	56
Oceans	25 + N2	= 33	* 0.6	=	21
Land surface	21	= 21	* 0.4	=	8
Total		=110		=	83

Level-4. Global, regional, and/or long-term models or results from analyses of lower level MODIS products, optionally including in-situ data measured at ground surface, or data from other satellite instruments. Candidate Level-4 products are listed in Table 3.

Table 3
Candidate Level-4 Data Products

Bi-directional models derived from MODIS data
Net carbon dioxide uptakes/release from biological activity
Wind speeds from clouds (m/sec)
- At cloud top height
- Derived at surface

Ocean speeds from surface tracers (cm/sec)
Net meridional energy transport (J/day)
Meridional water vapor transport (gm/sec)
Deforestation rates per unit area
Relationship between ecological zone classification and surface temperature and precipitation
Radiative transfer models vs. MODIS measured radiances
GCM predicted parameters vs. MODIS measured parameters
Mesoscale model predictions vs. observed mesoscale variations
Teleconnection studies

2.1.3.1.2 *Browse Data.* Browse data products assist data users in selecting data that is suitable for their purposes. It is not meant to be used as input to processing algorithms that produce higher-level parameters from lower-level products.

General Requirements

The MIDACS shall support operator designation of those specific MODIS Data Products for which Browse Products will be generated. Options shall include the designation of any, all, or none of the Standard and Specialized MODIS Data Products that the MIDACS produces.

Browse data will be single byte spectral and spatial summaries/subsets of the full resolution MODIS products. This data will require storage equivalent to less than 0.5% of the full resolution MODIS data. Two types of browse data have been defined:

a. Twenty kilometer resolution, single byte latitude/longitude scenes, are specified with four spectral channels each (e.g., thermal infrared, near IR, shortwave IR, and visible) for MODIS-N and -T, and for many (e.g. 50%) of the Level-2 parameters. Eight of these scenes cover the Earth, and an additional eight are defined for specific domains of general interest (e.g., Antarctica, tropical Pacific Ocean). Browse products will be annotated with labels, latitude/longitude grid, and geographical boundaries.

b. Four kilometer (at nadir) resolution single byte, cross-track/along-track scenes are specified, with two spectral channels each for MODIS-N and -T. Twenty of these scenes to cover each orbit (ten during daylight only for MODIS-T).

Specifics of Browse Products

The MIDACS shall support the designation of up to four spectral bands each for the MODIS-N and MODIS-T instruments to be used to generate Level-1B and possibly Level-3 Browse Products for these instruments. Browse products shall be generated only for the designated spectral bands. Levels-2 and -4 and most Level-3 data products do not require the designation of spectral bands for display since they display geophysical parameters. Eight bits of stored information shall be allocated to control the color or gray tone of each displayed pixel.

Depending on whether the original data product that the Browse Product supports has been resampled to an Earth-referenced grid, two types of Browse Products are possible. Level-1 and -2 MODIS product images have not been resampled to an Earth-referenced grid and retain the scanning or Field-Of-View (FOV) characteristics of the MODIS instrument as the observations were originally generated. Browse images that support these products also retain the instrument field-of-view characteristics of the product that they support. Level-3 and most Level-4 products have been resampled to Earth-referenced grids and Browse Products at these levels are also Earth referenced.

Field-of-View Products

To reduce data volume and facilitate data transmission to the user, FOV Browse Products shall be generated at 4 km by 4 km resolution. For FOV products, cross-track scan lines shall be displayed horizontally and the along track dimension shall be displayed vertically. The display shall be Earth-referenced by providing latitude/longitude coordinates for each of the four corners of the image. Coordinates shall be displayed on-screen in the approximate location of the image corner that is being referenced. In addition, the date and time of observation (UT), instrument type designator (MODIS-N or T), product type designation, spectral band (for Level-1 products), and instrument tilt angle (MODIS-T only) for each image shall be displayed on the browse image so as to be readily accessible to user view with minimal obstruction of the underlying image. An overlay map depicting Earth features is not required for MODIS field-of-view products.

Earth-Referenced Products

For Earth-referenced products, two types of browse data are defined:

a. Earth-Octant and Global Views Products are generated using 20 km by 20 km pixels and a latitude/longitude grid system to provide a global perspective on data. The entire Earth can be displayed at this resolution on eight 512 by 512 element screens. This data presentation is useful to a researcher investigating large-scale or global events. The MIDACS shall also support the composite display of the entire Earth on a single 2,048 by 1,024 element display for a researcher having access to a very high-resolution display.

b. Regional View Products support researchers interested in a specific continent or region of the Earth. The MIDACS shall support a 20 km by 20km resolution, 512 by 512 element display of the eight rectangular regions specified in Table 4.

For the Earth-referenced Browse Products described above, the MIDACS shall provide a geographic overlay map that allows the viewer to reference the image to known Earth features. The overlay map shall include latitude and longitude grid lines, policical boundaries, major cities, and other Earth features sufficient to allow the ready Earth location of images. In addition, the date or dates of observations, instrument type designators (MODIS-N and/or T), product type designation, spectral band (for Level-1 products), and instrument tilt angle (MODIS-T only) for each image shall be displayed on the browse image so as to be readily accessible to user view with minimal obstruction of the underlying image.

Table 4
Candidate Regions

Domain	Center
Antarctica	90°S, 0°W
Arctic	90°N, 0°W
East Tropical Pacific	0°N, 135°E
West Tropical Pacific	0°N, 135°W
Southeast Asia	0°N, 90°E
Tropical Atlantic	0°N, 50°W
Continental U.S.	35°N, 110°W
Africa/Sahel	0°N, 110°W

2.1.3.1.3 *Metadata Requirements.* All MIDACS data shall be handled in blocks that are appended with a descriptive header describing the data within the block. Descriptive information shall include the following:

MODIS-N/MODIS-T sensor identification
Product sequence number/version number
Processing date
Calibration algorithm identification number/version number
Product start and stop times
Orbit number(s)
Geographic boundaries of the product
Channel identification
Data quality flags
Calibration quality flags

Land/Ocean flags
Measures of cloudiness
Instrument tilt (MODIS-T)
Scan number(s)
Attitude information
Platform ephemeris
Solar and satellite zenith angle information
Time of observation
Calibration coefficients
- Level-2 products descriptors and quality indicators
- Level-3 products with resolutions and domains

2.1.3.1.4 *Geographic Grids for MODIS Data Products.* It is important that the MODIS Level-3 products, as well as those from other related EOS instruments such as HIRIS and AIRS, be produced on common grids. The use of such common grids would greatly facilitate research applications requiring data sets from more than one EOS instrument. For some applications, such as global climate, a latitude/longitude grid system may be optimal. For other applications, such as Earth resources, equidistant grid systems have proven to be the most useful. The choice of grid type and resolution often depends on the spatial scale chosen.

Latitude/Longitude Grid Systems

The types of grids to be employed could include a global 1° mesh (with 180 x 360 = 64,200 elements), a 2.5° mesh (with 72 x 144 = 10,368 elements), and a 5° mesh (with 36 x 72 = 2,592 elements). Advantages of this type of grid include a compatibility to many previous data sets, including Earth radiation budget and meteorological, as well as intuitive familiarity with the Earth's geography, and direct compatibility with many presently existing plotting packages. A disadvantage of this type of grid include a variable spatial resolution, with a much higher resolution near the poles.

Equidistant Grid Systems

The types of grids to be employed could include regional 1 km, 5 km, 10 km, 20 km, 50 km, and 100 km meshes, as well as a global or near-global mesh with resolutions of 100 km or better. Grid resolutions of less than 1 km would certainly be desireable for HIRIS. Such a system should be coregistered to the common MODIS/HIRIS grids at lower resolutions, but would not directly apply to MODIS unless some of the MODIS-N non-baseline channels were selected.

MODIS/HIRIS Data As Scenes

By defining a "scene" as being composed of five minutes (about 2000 km) of MODIS data at full swath width, a further type of data treatment becomes possible. About 20 of these scenes would be generated during each orbit. These scenes could then be located into a common scene data base. A user would then be able to access this metadata and conveniently identify those days for which specific regions of interest were sampled. These MODIS scenes should relate directly to corresponding HIRIS scenes or subscenes at the higher HIRIS spatial resolution.

2.1.3.2 *Description of Specialized Data Products*

Specialized data products are data products that are considered to be part of a specific research investigation and are produced for a limited region or time period. While new or experimental products may eventually be accepted by the research community as Standard Products to be routinely produced, until they have been formally designated as Standard Products, they are Specialized Data Products.

2.1.3.3 *Data Processing and Distribution Priorities*

2.1.3.3.1 *Real-Time.* Real-Time Data are transmitted and processed essentially as the data become available from the instrument, with only minimal delays for data storage in buffers. The MIDACS architecture shall not preclude such processing.

2.1.3.3.2 *Near-Real-Time.* Near-Real-Time Processing must be completed within 3-8 hours of the original observation.

2.1.3.3.3 *Routine.* All Level-1 MODIS Data Products that are routinely produced shall be available for distribution within 48 hours of the observation that produced the product. All Level-2 products routinely produced[16] shall be available within 72 hours after the observation. Routinely-produced Level-3 products shall be available within 96 hours of the observation. Fixed time limits are not specified for Level-4 products.

2.1.3.4 *Retention of Data Products*

2.1.3.4.1 *Products that are Discarded.* Data products generated to support the monitoring of instrument function are not retained within the MIDACS except as required to support the instrument monitoring function.

2.1.3.4.2 *Products to be Archived.* All Standard Products generated within the CDHF are routinely archived for potential user access. Specialized Data Products are also archived, if they are provided to DADS. The specialized data products to be archived should follow the same requirements as standard products.

2.2 PERFORMANCE REQUIREMENTS

Examples of quantitative performance requirements include overall system response times, communications link and processor throughput capabilities, and data storage capacities. Such performance measures can be specified both for the overall system operating as a whole and for the individual components that make up the system. Performance requirements are determined by the MODIS Science and Instrument Team members and by the various EosDIS organizations and committees that set standards and goals for all data systems supporting EOS instruments (including MODIS).

2.2.1 Nature of Requirements[17]

2.2.2 Instrument Control

2.2.2.1 *Monitor Scenes, Data Volume and Timeliness*

2.2.2.2 *Frequency of Instrument Commands and Timeliness Requirements*

---

[16]The exact role of interactive processing or processing by demand is not well defined here.

[17]For the time being, these requirements are TBD.

2.2.3 Data Products

2.2.3.1 *Daily Processing Volume and Throughput Requirements*

2.2.3.2 *Daily Storage Volume*

2.2.3.3 *Daily Data Distribution Volume and Timeliness Requirements*

Level-1B and Level-2:

Define a "heavy global user" as one who needs: (1) 1 km resolution global coverage; (2) 10 days of data; and (3) 10 MODIS radiances and/or parameters. Define a "heavy regional user" as one who needs 1 km resolution with: (1) coverage over 1/20 of the Earth; (2) 365 days of data; and (3) 10 MODIS radiances and/or parameters. The typical data request will be for on the order of 10^{12} bits of data. We define a "moderate" MODIS data user as one who orders 10% of a heavy user, such as with only one MODIS parameter. We define a "light" MODIS data user as one who orders 1% of a heavy user (e.g., only one month of regional data or a single day of global data. We assume the user community is equally partitioned into light, moderate, and heavy users.

For each transaction of Level-1B and Level-2 data, we have:

Light user: 10^{10} bits
Moderate user: 10^{11} bits
Heavy user: 10^{12} bits

Level-3:

Three general types of Level-3 MODIS data will exist: (1) global; (2) regional; and (3) rectified image. We will assume that the global Level-3 MODIS data, as derived from Level-1B and Level-2 products, will exist on grids with 20 km and 100 km resolutions (with 10^5 and 10^6 elements, respectively). We will assume that the regional Level-3 MODIS data, as derived from Level-1B and Level-2 products, will exist on grids with 5 km and 20 km resolutions (with 10^6 and 10^7 elements, respectively). We will assume that the rectified Level-3 MODIS image data, as derived from Level-1B and Level-2 products, will exist on a 1 km resolution grid of 2,048 x 2,048 lines and pixels (with $4*10^6$ elements). Assuming that 10 parameters are required, a Level-3 data transaction involves:

All users: 10^9 bits

Browse Data:

Browse data will be single byte image data over a 512 x 512 line/pixel array (eight of which cover the entire Earth at about a 20 km equatorial resolution). For the browse data, every transaction involves:

All users: 10^7 bits

3. MIDACS OPERATING ENVIRONMENT

3.1 SOFTWARE STANDARDS[18]

3.1.1 Operating System Portability

3.1.2 Coding Standards

3.1.3 Standard Formatted Data Units

3.1.4 Machine Specific Formats

3.2 HARDWARE

3.2.1 System Expandability and Pre-Planned Improvements

During the hardware design process, potential system improvements that would provide increased hardware capability shall be considered and formulated as pre-planned system improvements. Pre-planned improvements provide a visible and well-defined path to system upgrade as system requirements increase.

3.2.2 System Reliability, Maintainability, and Availability

MIDACS Reliability, Maintainability, and Availability (RMA) requirements should be derived from the fundamental requirement that a failure within the MIDACS shall not lead to loss of instrument control or observation data. Besides those portions of the data system dedicated exclusively to serving MODIS needs, the complete data system includes higher level control facilities and data backup facilities that provide some redundancy with MIDACS functions. Therefore, MIDACS RMA requirements must be derived considering functional alternatives that are available in other portions of the overall data system. This section begins with a discussion of redundancy and backup functions within the overall system and it concludes with a derivation of specific MIDACS RMA requirements.

3.2.2.1 *Redundancy and Backup Functions within the Overall Data System*

3.2.2.2 *MIDACS RMA Requirements*

3.3 DATA TRANSFER AND REQUIRED INTERFACES

A conceptualization of the MODIS data system (MIDACS) in the EosDIS environment is presented in Figure 1. The corresponding MIDACS context diagram, illustrating external interfaces and data flows, is illustrated in Figure 2.

3.3.1 On-Board Processing

The MODIS-N and -T instruments will contain processing capabilities which are yet to be fully defined. In addition to controlling the physical operation of the instrument, the analog-to-digital conversion of the measurements, and data packet generation, the

---

[18]System-wide EosDIS software standards will apply to the MIDACS. Once a common set of software standards has been developed for all the Eos instruments, these standards will be directly incorporated for MIDACS use.

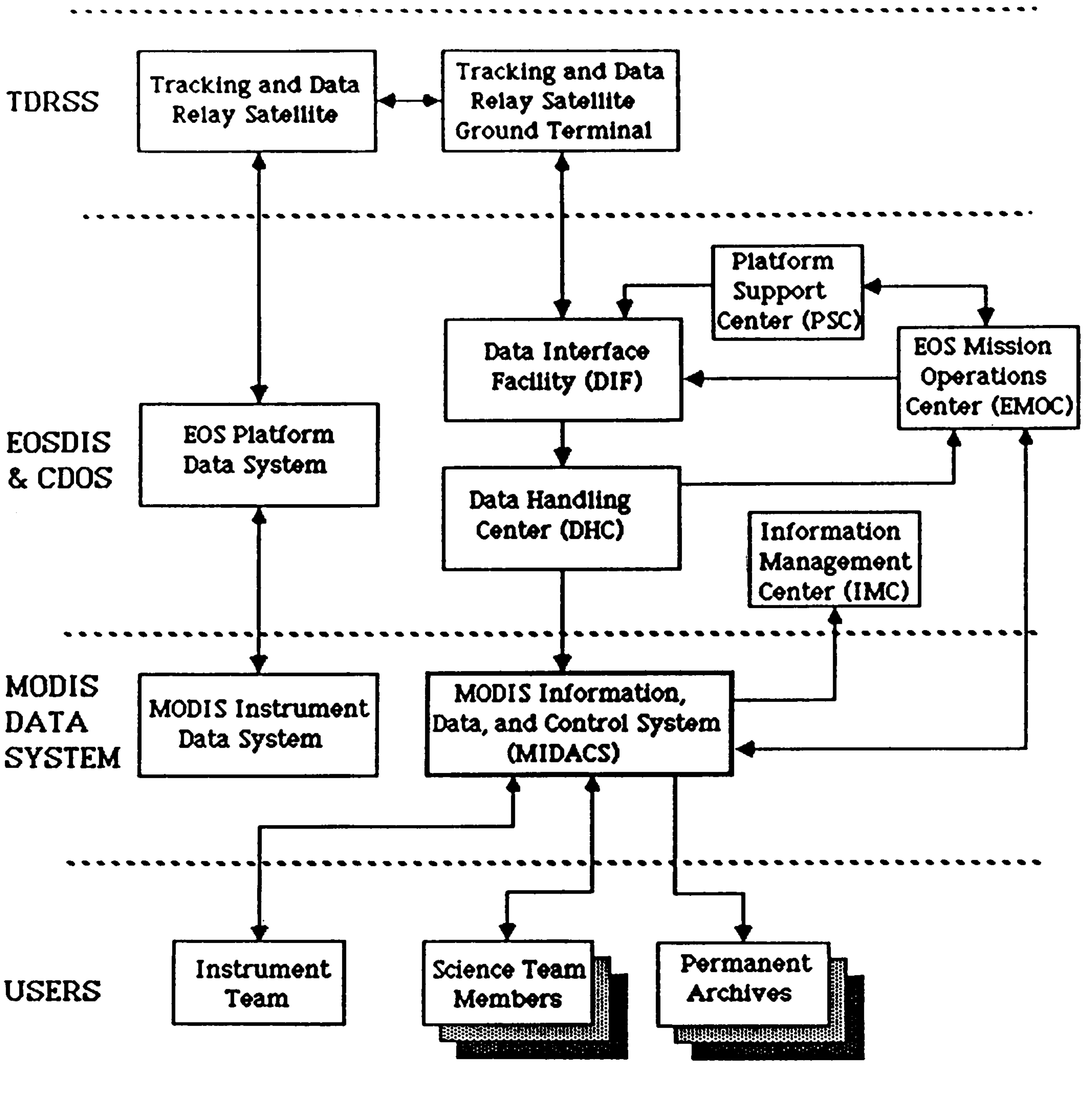


Figure 1. The MODIS Data System in the EosDIS Environment

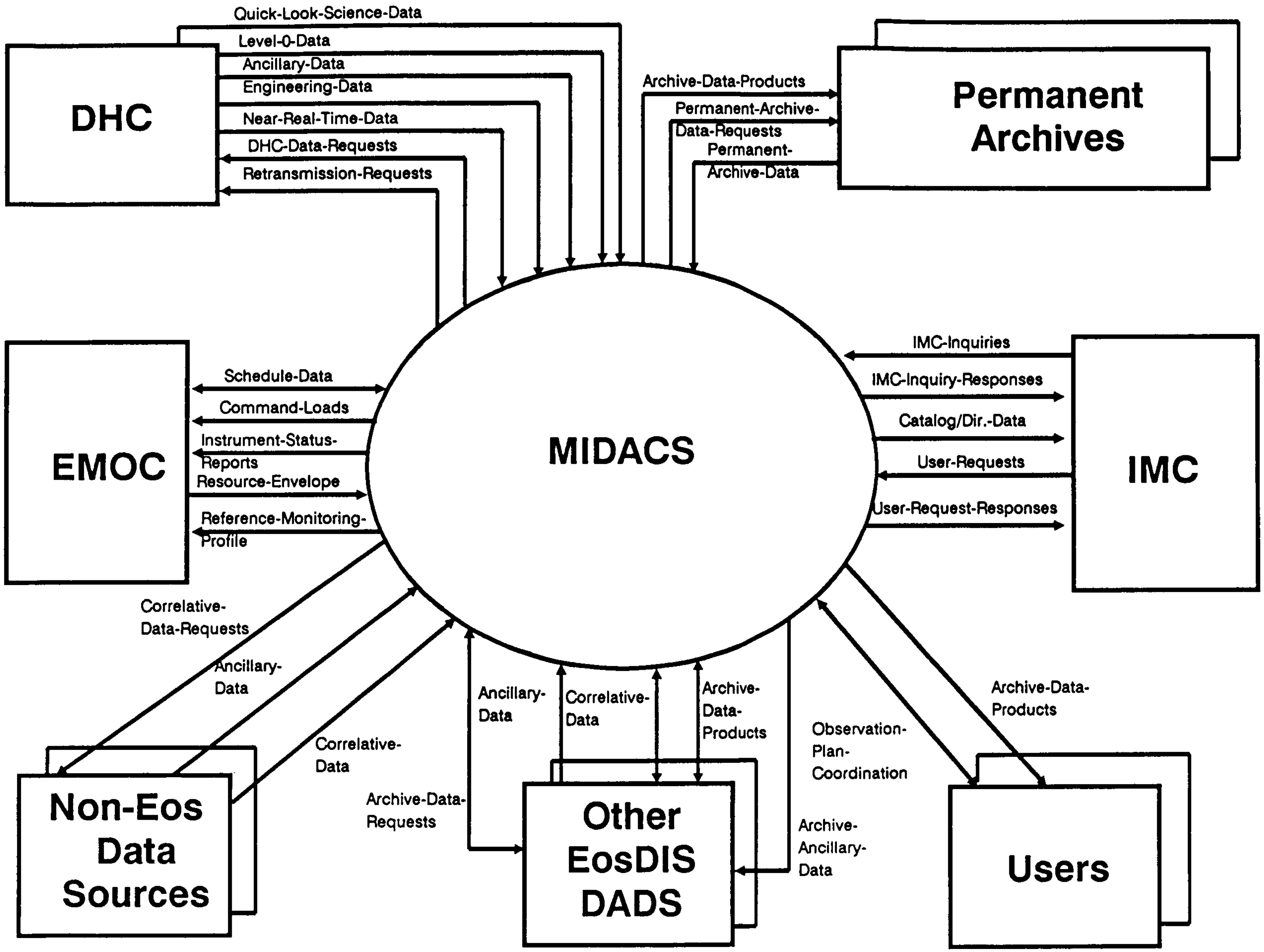


Figure 2. MIDACS Context Diagram

processor will be responsible for many other functions. The on-board processor is not a part of the ground data system supporting the MODIS data reduction, and due to the presence of the TDRSS, DIF, and DHC, no physical interface exists between the two data systems. However, there is a clear functional interface between the on-board and ground data systems, both of which comprise the MODIS data system. The reason for considering the on-board processing in this specifications document is simple: the two systems must function as a single entity, and data must flow transparently between the two. There are functions which must be performed by either the on-board or ground data systems (such as data buffering and sorting), and functions which must be reversed on the ground if performed in flight (such as data compression).

3.3.1.1 *On-Board Data Buffering and Formatting*

It would be desireable if the on-board data system could buffer complete scans of data, and sort the measurements by channel prior to packetization.

3.3.1.2 *On-Board Data Compression*

It would be desireable if the on-board data system could support the performance of either a lossless or lossy data compression.

3.3.1.3 *On-Board Data Processing*

On-board processing of data could generate some products which could be used to control the instrument (e.g., set detector gains for land/sea, alter sensing routine based on cloud cover, etc.).

3.3.1.4 *On-Board Packet Addressing*

On-board addressing (e.g., MODIS CDHF or ICC) and priority designation (e.g., real-time or priority playback) of packets to support real-time monitoring and near-real-time support of field experiments.

3.3.1.5 *On-Board Data Packet Building*

The on-board data system could build data packets for ancillary and engineering/housekeeping data, and other special purpose packets of interest to the ground data system (e.g., from the platform LAN, generate platform ancillary packets, or packets summarizing the duty cycles and selected observations from other instruments on the platform).

3.3.1.6 *Command and Scenario Storage and Execution*

Command and scenario storage and execution functions (e.g., sunglint avoidance, automatic gain changes for high solar zenith angles, internal calibration sequences) are TBD.

3.3.1.7 *Improve Instrument Reliability*

To improve instrument reliability, the data system will provide or support built-in instrument self-test capability, provide fault detection and automatic instrument safing, or support command alternatives to correct instrument failures (e.g., built-in instrument redundancy).

3.3.2 Control Data Exchange with the EMOC

3.3.3 Data Availability at the Data Handling Center

When MODIS instrument data is received at the Data Handling Center (DHC), MODIS data packets containing the MODIS instrument data are embedded in larger blocks of data (transfer frames) that also contain data packets from other instruments and on-board systems. The header and trailer (optional) appended to transfer frames provide the primary data addressing and error correction capability across the satellite-to-ground-station communications link.

At the DHC, the bit order of received data is corrected if the data was initially recorded on on-board tape recorders and reverse order tape playback was used. The headers and trailers of the transfer frames are removed, the contents of the transfer frame are error corrected, if possible, and instrument packets addressed to the individual instrument data systems are separated and grouped according to instrument. Packets for each instrument are arranged in ascending time order, and when a complete ascending sequence of packets has been received, data is transferred from the DHC to the appropriate instrument data system. Except for possible errors introduced during data transmission, the data being transferred to the instrument data system is a reconstruction of the original data stream at the output of the on-board instrument data processing. This reconstructed stream is designated Level-0 data.

3.3.4 Transfer of Data to the Long-Term Archives

During the 15 year system lifetime, the DADS will transfer MODIS datasets to NASA, NOAA, and USGS long-term archives. Transfer activities will be transparent to DADS users. After the data transfers are completed, the affected data will be available only from the archive facility. Users whose queries reference this data will be advised on both the data's location and the need for contacting the respective archive. The IMC/DADS will continue to store and make available the catalog and metadata relevant to the transferred datasets.

3.3.5 Data Distribution Media

3.3.6 Other External MIDACS Interfaces

Information Management Center

The Information Management Center (IMC) is a facility provided by the EosDIS to support user access to stored data products for all Eos instruments. Besides serving as a single point of access for users desiring any of the Eos data products, the Center provides Browse Products for MODIS and other Eos instruments, it stores catalog and directory information for all Eos Data Products, it processes and distributes product orders from users and it serves as a repository of operations information for the MIDACS and other data systems. Available information includes planned instrument operations, processing status or expected delivery of data after observations are completed, and processing status of user data requests after data products have been ordered.

Other Eos DADS

The MIDACS will access information stored in DADS facilities for other instruments. Some of the data products generated by MIDACS require data products from other instruments as input. If the DADS facilities for all Eos instruments reside in the same physical facility, transfer of data to the MIDACS will be simple to implement. Since the volume of data to be transferred is appreciable, separate DADS facilities would require substantial data communication facilities between DADS nodes.

Non-Eos Data Sources

Non-Eos data will be available from a variety of sources and data will be in a variety of formats. Perhaps a typical data exchange will involve the exchange of magnetic tapes. Other possible exchange media include data communication lines, magnetic or optical disks, or perhaps in the mid-1990's, optical tape.

4. MIDACS ARCHITECTURE

4.1 INTRODUCTION TO ELEMENTS

The MIDACS consists of five basic structural elements defined for the EosDIS. Two of these elements support the MODIS instrument command process: the ICC and the IST. The CDHF processes observation data returned by the MODIS instrument, and the DADS stores and distributes MIDACS data. The TMCFs provide resources for nonroutine MIDACS tasks: development of data processing algorithms, determination of instrument calibration constants needed in the algorithms, validation of final MODIS results by comparison with results obtained using other methods, and the production of specialized data products not routinely generated by the CDHF.

Each MIDACS element (IST, ICC, TMCF, CDHF, and DADS) is a logical entity that performs a particular subset of required MIDACS functions (see Figure 3). The division of overall system function among modules dedicated to instrument control, data processing, data storage and retrieval, etc. presents a natural opportunity to select separate, specialized hardware components for each module. Each hardware unit can then be chosen to provide support specifically tailored to a corresponding software function. Distinct hardware components naturally allow the physical separation of MIDACS elements when remote access is needed. Therefore, the MIDACS is being specified to allow the possibility that each MIDACS element will be a physically distinct entity.

If each MIDACS element is physically distinct, then the data interfaces shown in the overall MIDACS context diagram and the MIDACS element context diagrams are not just logical interfaces, but physical interfaces as well. The logical structure of the system and the general nature of interfaces is documented in the data flow diagrams and the corresponding data dictionary descriptions of the data flows.

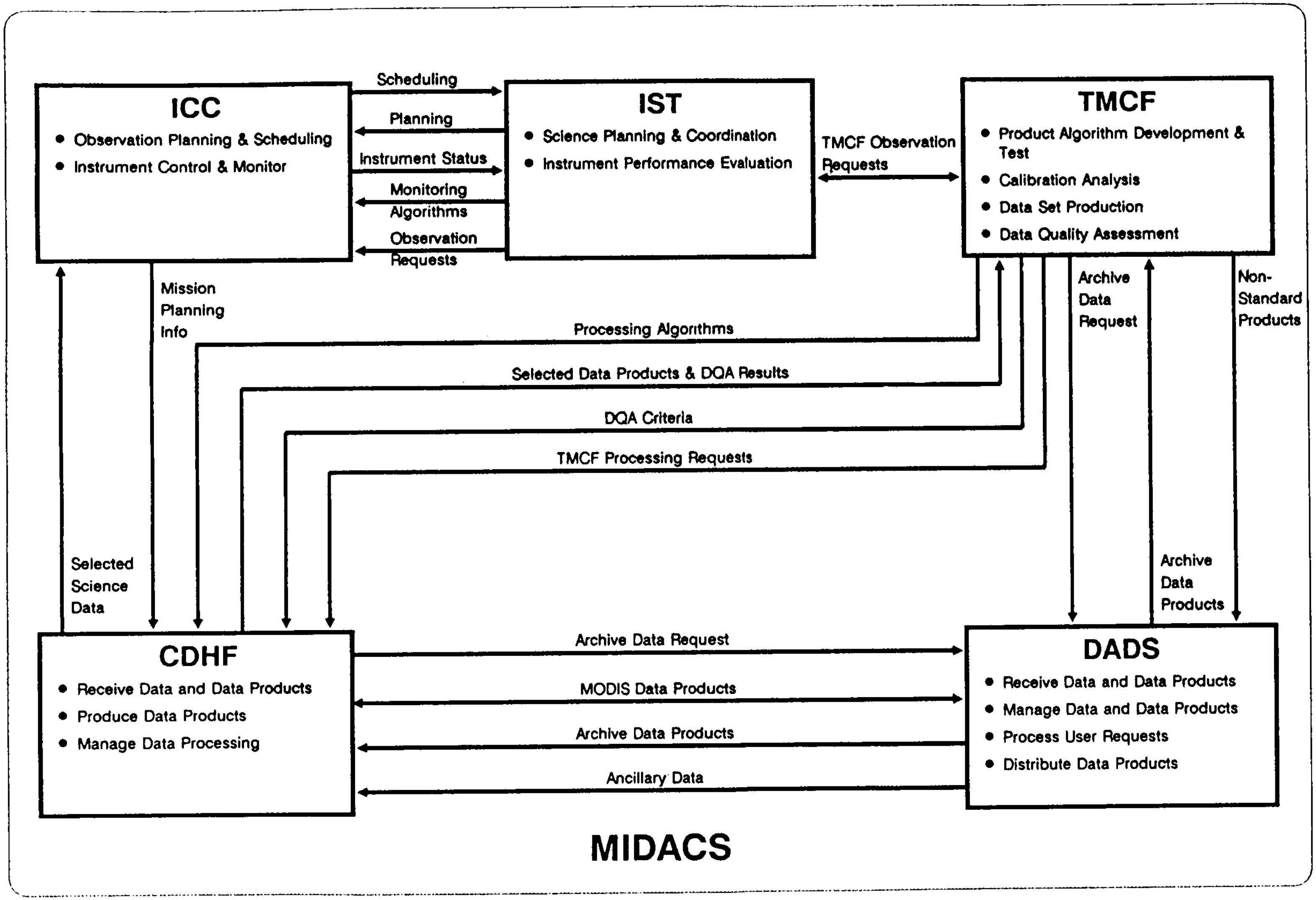
4.2 GENERAL CONSIDERATIONS

5. ELEMENT SPECIFICATION OF THE MODIS INSTRUMENT SUPPORT TERMINAL (IST)

5.1 REQUIREMENTS

5.1.1 Functional Requirements

The MODIS IST is the Science Team Leader's terminal and interface to the MODIS ICC. The functions supported by this terminal are performing observation planning and coordination and monitoring instrument performance.



DQA – Data Quality Assessment

Figure 3. MIDACS Element Functional Allocation Diagram

5.1.2 Performance Requirements

The MODIS Science Team Leader will direct the development and initial operation of MODIS-specific elements of the MIDACS. He will ensure that a basic MODIS science plan, instrument operation models, and instrument operations and monitoring procedures all exist. Following an initial orbital verification period, it is expected that changes to the science plan, instrument models and monitoring algorithms will be made infrequently.

The demand for the Science Team Leader's active research involvement will be lessened as operations become routine. He may monitor the operational aspects of the instrument via his IST and through his involvement with the Science Team Members. Therefore, there are no performance requirements identified for the IST at this time. It is anticipated that the IST will be but one of many workstations networked to the ICC's control and monitor system. The IST will access the same monitoring databases and data displays as other ICC workstations.

The IST may have some restrictions as to access to command transmission facilities and may not enjoy the full capabilities of the ICC's science workstation for planning and commanding. The current envisioned IST capability will allow the Science Team Leader to monitor and to respond appropriately to MODIS anomalies or special situations.

5.1.3 Hours of Operation

The IST interfaces with the ICC will be available at all times.

5.2 IST INTERFACES

Figure 4 shows the context of the IST and illustrates the functional interfaces between the IST and EosDIS, as well as other MIDACS elements.

5.2.1 Interfaces to EosDIS

In this context, IST is used synonymously with Science Team Leader. The MODIS Science Team Leader is a member of the Investigator Working Group (IWG) which meets regularly to determine modifications or changes to the science plan. This interface is shown in the context diagram with USERS and is exercised physically by the Science Team Leader's participation at the IWG meetings and over the telephone (e.g., electronic mail). It is anticipated that there may be a flurry of activity twice per year for these IWG functions. The next paragraph describes a more routine user access to MODIS operations. A user may request a MODIS observation not already covered by the science plan by submitting a request through the IMC. This same interface may serve as conduit for observation plan information.

5.2.2 Interfaces to Other MIDACS Elements

The MODIS Science Team Leader will interface with his Science Team Members at the TMCF(s) via the IST. As shown in Figure 4, this interface is used primarily for requesting MODIS observations not covered by the science plan, and for requesting special data handling, possibly to support field experiments. These requests will be formatted into a Data Acquisition Request (DAR) prior to transmittal to the ICC.

5.2.3 Operator Interfaces - N/A

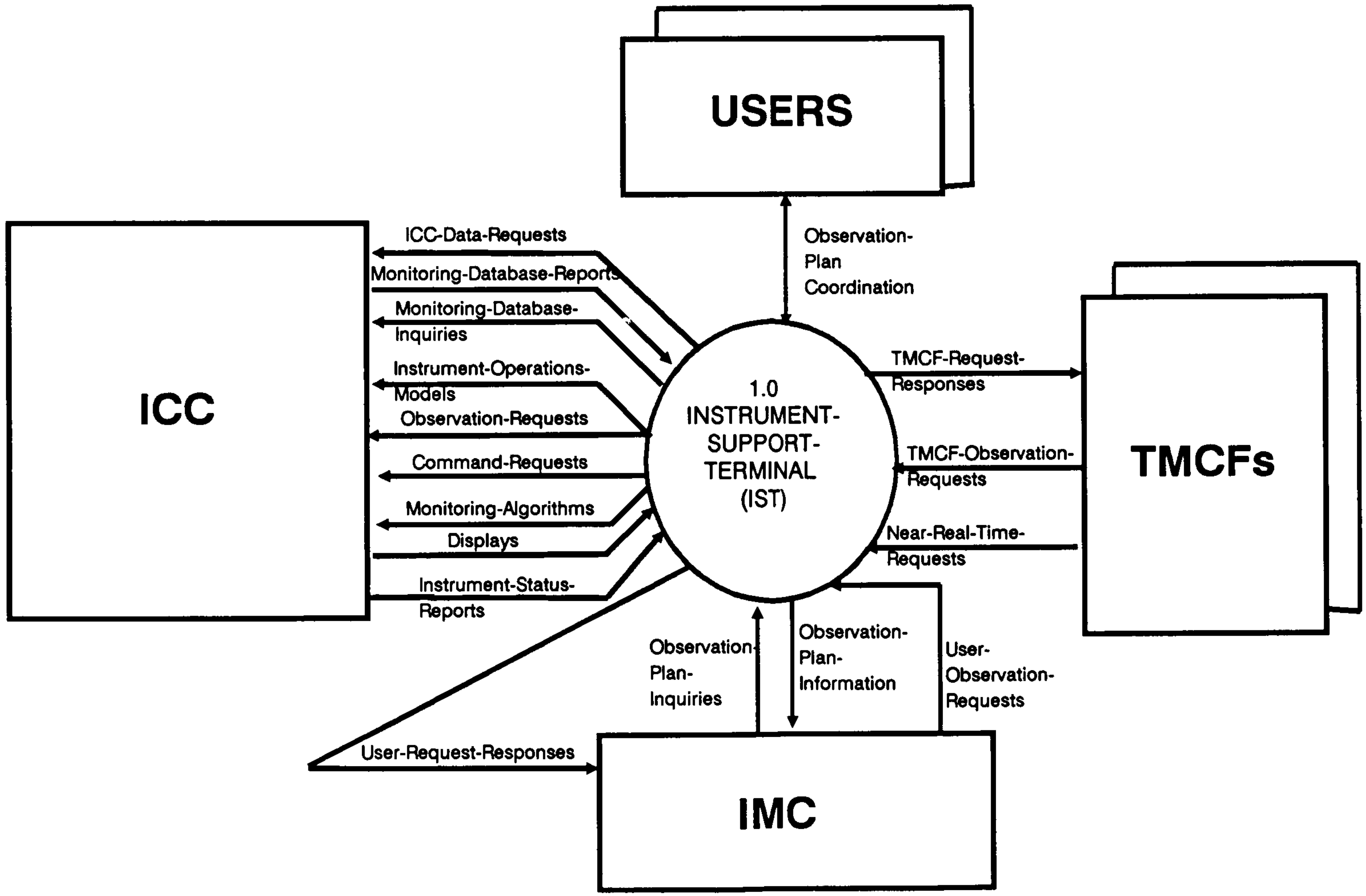


Figure 4. IST Context Diagram

5.3 ARCHITECTURAL DESIGN CONSIDERATIONS

This section presents certain considerations which are intended to be useful inputs toward a conceptual and detailed design of this MIDACS element. A summary of computational assumptions; followed by a traffic analysis across all interfaces shown in the context diagram; then, a discussion of storage requirements; and, finally figures pertaining to a conceptual design are presented.

5.3.1 Assumptions

MODIS is essentially a 100% duty cycle instrument, making its planning and scheduling more generic and possibly relatively simple. Specific data acquisition requests (DARs) are expected to be the exception rather than the rule. Reasons for a DAR may be of the form of a special calibration sequence, or a deviation from the nominal tilt plan for MODIS-T. DARs may be initiated by MODIS Science Team Members or by external research users. The TMCF Observation Requests will be made directly to the IST. A User Observation Request will be submitted to the IMC and relayed from there to the IST. It is anticipated that the sum of these exceptional data requests will number 100 or less per month. A DAR will be about one page in length (letter size), or about 4000 bytes. The contents will be formatted in a near-natural language and submitted electronically to the IST. Upon review and approval, the MODIS Science Team Leader will electronically forward the DAR to the ICC's planning and scheduling processor. It is assumed that the response to an observation request may be contained in an 80 byte field and may go over electronic mail.

The Science Team Leader is also a Science Team Member. As such, we assume that the MODIS Science Team Leader's proposal specified a reasonably powerful workstation on which to perform scientific research. Although not required, this workstation may be attached to the ICC's control and monitor network and function as the IST.

The Science Team Leader may wish to display selected quick-look science data at the IST. Assume that the IST has a 2048 x 2048 pixel image monitor; that there are 2 bytes per pixel in 3 colors; and that the image data to be transferred to the IST will be pulled from the ICC on-line data store onto the narrow-band network. This configuration would require about 20 seconds to build a full MODIS scene at the IST.

The IST will also have access to any of the 80 (or so) engineering display pages (monitor size, i.e. 24 lines by 80 columns) routinely monitored at the ICC workstations. In the worst case, it is assumed that the Science Team Leader may spot check all 80 pages once per day. Further, it is assumed that 60 of these pages are text at about 2000 bytes each, and 20 pages are graphics. If we assume that a typical ICC monitor has a 640 x 350 pixel field at one byte each, then the total graphics field may require about 0.2 MB of information to be transferred to the IST. If we assume, however, that for engineering plots, only about 20% of the graphic pixels are illuminated by non-zero values, then these 20 graphics pages may be only about 0.045 MB each. On a 10 Mbps narrow-band network, each graph would require less than 1 second for transfer to the IST.

The Science Team Leader will have access to all ICC monitoring databases and data stores. The IST will receive periodic instrument status and database from the ICC. It is assumed that an instrument status report may be a daily summary of status and events condensed to one letter size page of about 4000 bytes.

It is also assumed that a more extensive database report will be generated weekly, monthly and annually; and, that the Science Team Leader will request a database report once per month. It is assumed that each report is 20 letter size pages in length and that at least half is in the form of graphics. Therefore, a database report may consist of:

10 pages of text @ 4000 bytes	= 40 kB
10 pages of graphics @ 0.09*MB	= 0.9 MB
total report	= 0.94 MB

*two engineering plots per page

On the 10 Mbps narrow-band network, each report would take less than one second to transfer to the IST.

5.3.2 Traffic Analysis

The types of traffic across the IST fall predominately into the areas of planning and monitoring.

5.3.2.1 *Planning*

For the reasons stated earlier as to the longer term science planning and coordination, the traffic between the USERS and the IST is small enough and of such low duty cycle as to permit use of electronic mail over standard telephone data links.

Observation Plan Coordination (between USERS and IST): <10 kB/month

For planning purposes, specific data acquisition requests will be the exception rather than the rule. DARs may be initiated to the IST by the TMCF(s) or by external users via the IMC. It is assumed the Science Team Leader will review, approve, and forward (or reject) 100 DARs per month. Therefore,

100 DARs/month * 4000 B/DAR = 0.4 MB/month

are anticipated inputs to the IST.

The distribution of the DAR traffic to the IST will be assumed to be 90% from the TMCF(s) and the rest from the IMC. Therefore,

TMCF Observation Requests (from TMCF to IST): 0.36 MB/month

Near Real Time Requests (from TMCF to IST): included in above

TMCF Request Response (from IST to TMCF <10 kB/month

User Observation Request (from IMC to IST): 0.04 MB/month

User Request Response (from IST to IMC): <1 kB/month

Upon review and approval of the 100 DARs per month, the MODIS Science Team Leader will forward the formatted Observation Requests to the ICC for planning and scheduling.

It is also assumed that an IST-initiated Command Request will be forwarded to the ICC in the form of a DAR. Therefore,

Observation Requests (from IST to ICC): 0.4 MB/month

Command Requests (from IST to ICC): included in above.

It is expected that the following two interfaces would be better made directly with the EMOC and will not be discussed in this context.

Observation Plan Information (from IST to IMC): N/A
Observation Plan Inquiries (from IMC to IST): N/A

5.3.2.2 *Scheduling*

The MODIS instruments will be modeled to the extent that an operations (resource) envelope (e.g., thermal, power, LAN) will be dictated by the EMOC as an operating constraint on MODIS operations scheduling. It is assumed that such instrument models will be provided by the Science Team Leader and maintained by the Science Team Members. It is further assumed that once established, the models will remain static for long periods of time. The traffic on the interface is, therefore,

Instrument Operations Models (from IST to ICC): <1 kB/month

5.3.2.3 *Commanding*

It is expected that IST originated command requests will be in the form of a DAR and such traffic is included under planning.

5.3.2.4 *Monitoring*

The IST is a remote workstation connected to the ICC's control and monitor network. Therefore, this traffic takes the form of requests for pre-defined displays of control and monitor data; receipt of trending data and/or status reports; occasional upgrades to monitoring algorithms; and selected science quick-look monitoring data.

a. The ICC will receive the initial MODIS monitoring algorithms from the MODIS Science Team Leader. However, these algorithms are expected to change very infrequently. Therefore, traffic across this interface is negligible.

   Monitoring Algorithms (from IST to ICC): <1 kB/month

b. The MODIS Science Team Leader may request data from the ICC in the form of an instrument status report or a display page of particular engineering or science interest for monitoring the performance of the MODIS. This request may also be for some trending analysis report which may come from the monitoring database. In the following subsections, it is assumed that the Science Team Leader will request one each of these per day. The data request should be very brief as far as traffic across the interface. It is assumed that this request can be accommodated in one 80B line. Therefore,

   ICC Data Request (from IST to ICC):

   2 requests/day * 80 B/request * 30 days/month = 4.8 kB/month

Monitoring Database Inquiry (from IST to ICC):

1 request/day * 80 B/request * 30 days/month = 2.4 kB/month

c. A MODIS instrument status report will be made available on request or once per day to the IST. It is assumed that this report takes the form of a one letter-size page management summary of MODIS T/N status. Therefore, this traffic may be:

Instrument Status Reports (from ICC to IST):

4 kB/day * 30 day/month = 0.12 MB.3.5/month

d. As explained in Section 5.3.1, the IST is a terminal on the ICC's Control & Monitor network and as such, will have access to any of the 80 (or so) engineering display pages routinely monitored at ICC workstations as well as selected image monitoring data. Therefore, the traffic across this interface is:

Displays (from ICC to IST):

60 pages (text) * 2000 B/page * 1/day * 30 days/month = 3.6 MB/month

20 pages (graphics) * 0.045 MB/page * 1/day * 30 days/month = 27 MB/month

25.2 MB/scene * 1 scene/day * 30 days/month = 755 MB/month

e. A monitoring database report may contain 1000 parameters on any given day and similar reports may be based on increasingly condensed data on a weekly, monthly, and annual basis. Assume that a report is 20 pages in length as given in Section 5.3.1, of which about half is in the form of graphics. Further assume that the Science Team Leader will request a daily report only about once per month; that he will receive a weekly report each week, a monthly report each month, and annual report each year; and that each report is about the same length. Therefore, the traffic on this interface may be:

Monitoring Database Reports (from ICC to IST):

1 report (daily) • 0.94 MB/report • 1/month = 0.94 MB/month

1 report (weekly) * 0.94 MB/report * 4.33 times/month = 4.1 MB/month

1 report (monthly) * 0.94 MB/report * 1/month = 0.94 MB/month

1 report (annual) * 0.94 MB/report * 1/12 month = 0.08 MB/month

5.3.2.5 *Training and Testing:* N/A

5.3.2.6 *System Input/Output*

A MODIS traffic summary is as follows:

a. Planning - 0.8 MB/month.

Observation Plan Coordination (between USERS and IST): <10 kB/month

TMCF Observation Request (from TMCF to IST): 0.36 MB/month

Near Real Time Request (from TMCF to IST): included in above

TMCF Request Response (from IST to TMCF): <10 kB/month

User Observation Request (from IMC to IST): 0.04 MB/month

User Request Response (from IST to IMC): <1 kB/month

Observation Requests (from IST to ICC): 0.4 MB/month

Command Requests (from IST to ICC): included in above

b. Scheduling - 1 kB/month

Instrument Operations Models (from IST to ICC): <1 kB/month

c. Monitoring - 792 MB/month

Monitoring Algorithms (from IST to ICC): <1 kB/month

ICC Data Request (from IST to ICC): 4.8 kB/month

Monitoring Database Inquiry (from IST to ICC): 2.4 kB/month

Instrument Status Reports (from ICC to IST): 0.12 MB/month

Displays (from ICC to IST): 786 MB/month

Monitoring Database Reports (from ICC to IST): 6 MB/month

5.3.2.7 *External Interfaces:* TBS

5.3.3 On-Line Storage Requirements

As an active workstation on the ICC's Control and Monitor network, the Science Team Leader will presumably have 10 MB or less of on-line work space allocated on the ICC's on-line storage devices.

Since the IST has access to all ICC databases and data stores, there will be no requirement to duplicate those at the IST.

The IST will need a database allocation to manage the incoming DARs. At 0.4 MB per month, about 5 Mb of on-line storage would be required to store a year's worth of data.

As far as monitoring algorithm maintenance, assume there are 100 algorithms for engineering monitoring each of which has 10 lines of code at 80 bytes per line. Further assume that three versions are retained at any time. Therefore about 240 kB of on-line storage are required for algorithm maintenance.

For instrument model maintenance, assume the model requires 1000 lines of code at 80 bytes each and that three versions are retained at any time. Therefore an additional 240 kB of on-line storage are required for model maintenance.

5.3.4 Off-Line Storage Requirements

For the purpose of gathering a permanent record of DARs input to the IST, approximately 75 MB of off-line storage will accommodate 15 years of instrument operation. Since monitoring algorithms and instrument models are assumed to require only 80 kB each, older versions of these may be retained on off-line devices (e.g., floppies) at the Science Team Leader's discretion.

5.3.5 Conceptual Design

Figure 5 depicts the functional data flow for the IST. The data rates on the interfaces show the peak traffic between the elements. The traffic between the IST and IMC, TMCF(s) and USERS is assumed to be over standard telephone data lines via electronic mail. Therefore, peak data rates across these interfaces is limited by modems at either end (i.e., 300, 1200, 2400, or 9600 baud).

The peak data rate between the ICC and IST will be limited by the capacity of the narrow-band network. The network used here for illustration purposes has a 10 Mbps capacity. The greatest stress to this capacity is assumed to be during times when monitoring image data is transferred to the IST.

In current technology, 2048 x 2048 image display monitors are connected to processors exceeding 1 MIPS capacity. However, there is nothing in the specification context requiring an IST processor greater than 1 MIPS.

6. ELEMENT SPECIFICATION OF THE MODIS INSTRUMENT CONTROL CENTER (ICC)

This section outlines the design parameters used as input to the preliminary MIDACS system specification and conceptual design.

6.1 REQUIREMENTS

6.1.1 Functional Requirements

The MODIS ICC conceptual architecture will be influenced by several significant functional drivers. They are as follows:

a. The instrument control and monitor function is a twenty-four hour a day operation, seven days a week. This is very much a mission operations context in which real-time or near-real-time commanding and/or monitoring are required capabilities;

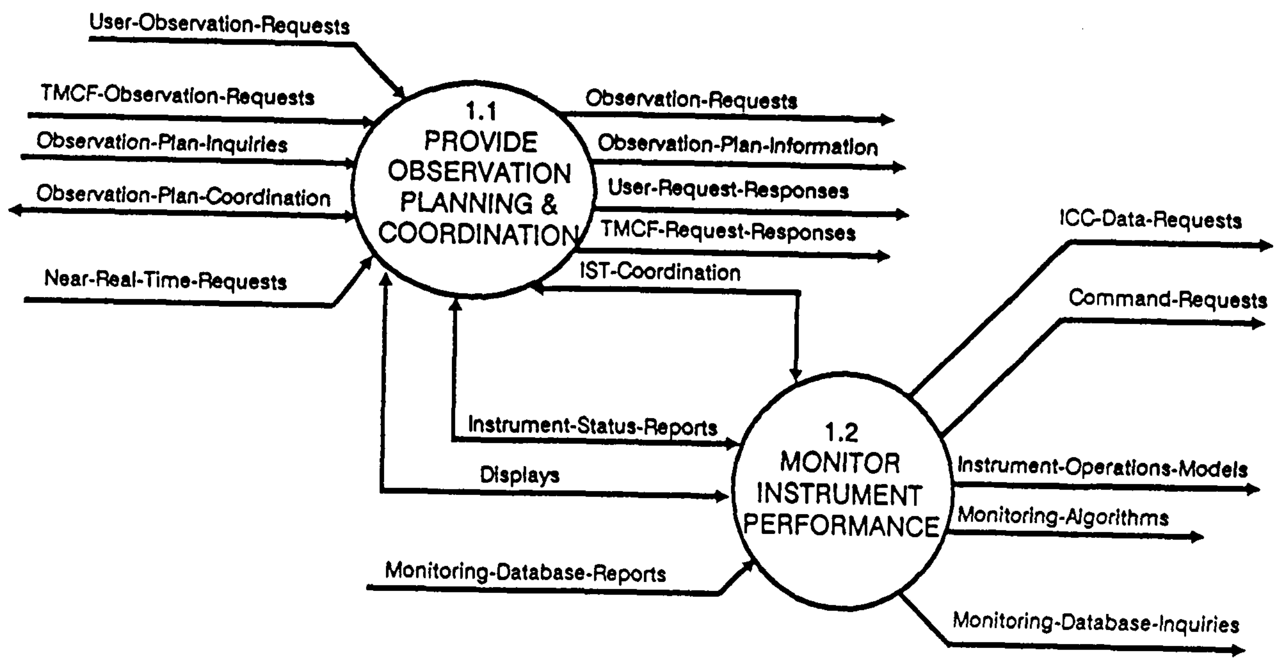


Figure 5. IST Functional Data Flow

b. Real-time or near-real-time displays of instrument status and state-ofhealth telemetry (engineering type data); and,

c. Near-real-time displays of full scenes (approximately 2048 x 2048 pixels) of 1 km resolution spectral bands of processed instrument science data to support instrument monitoring and field experiments.

6.1.2 Performance Requirements

The MODIS ICC's real-time nature and round-the-clock operational requirements are driven by the following performance requirements summarized from Section 4.2 of the MODIS Functional Requirements Document.

a. The MODIS ICC must be capable of responding to a request to support a target-of-opportunity (TOO) in near real-time (within the time period of one orbital revolution). This TOO support may involve a perturbation to an existing schedule, which will necessitate intensive coordination with EMOC, PSC and NCC for any real-time command and TDRSS support and with the DHC for real-time monitoring or near-real-time data processing.

b. Not necessarily associated with the above are separate requirements on the ICC to respond to instrument command decisions and requests for near-real-time data products within the time period of one orbit.

c. There is a requirement on the Eos system, of which the MODIS ICC is a part, that loop delays between command transmission and interpretation of telemetered response shall not exceed 10 seconds.

d. The MODIS ICC is also required to process and display any real-time control and monitor data within 60 seconds from on-board transmission.

e. The MODIS ICC's monitoring function is required to generate alarms within TBD seconds of receipt of data whenever a parameter exceeds a predetermined range.

f. A related requirement is for the ICC to generate a predetermined safing command sequence within TBD seconds for transmission to the EMOC.

g. The MODIS ICC must retain the MODIS engineering history file for trend analysis and anomaly investigation for the operational life of the instrument.

h. The ICC's wide-band ingest processor must be capable of handling a peak data input of about 20 Mbps and be capable of connecting to a wide-band communications network. This processor must also be capable of selecting and transferring any 8 MODIS 1 km channels in real-time. The resultant 8 channels will have a composite data rate of about 1 Mbps. Any or all of the selected data will be processed up to level 2 and displayed in near-real-time on high resolution image monitors capable of 2048 x 2048 pixel resolution.

6.1.3 Hours of Operation

The MODIS ICC will operate twenty-four hours a day, seven days a week. There will be at least a one-for-two redundancy in the critical areas of control and monitor, e.g., workstations (including processors) and communications lines.

6.2 ICC INTERFACES

Figure 6 shows the context of the ICC and illustrates the functional interfaces between the ICC and EosDIS/CDOS as well as other MIDACS elements.

6.2.1 Interfaces to EosDIS/CDOS

The ICC will have interfaces with the DHC, EMOC and the IMC. Essentially, traffic associated with these interfaces will be planning, scheduling and commanding oriented. Various types and quantities of status information are also exchanged among these elements.

6.2.2 Interfaces to Other MIDACS Elements

The ICC will interface with the IST and the CDHF. The IST will have access to any data or display available in the ICC. The MODIS Team Leader (at the IST) is responsible for maintaining instrument operations models and monitoring parameters implemented in the ICC. The Team Leader will use the IST to convey specific observation requests and command requests to the ICC. This interface will also convey instrument status, engineering trending, and performance information to the Team Leader for evaluation.

One output product of the ICC's planning and scheduling process will be mission planning information which will be sent to the CDHF for their use in production job planning. Under architectural review is the possibility of the CDHF processing and sending selected science data to the ICC for quick-look evaluation. Associated with this interface would be a control interface between the ICC and CDHF for the purpose of making the science data selection in real-time.

6.2.3 Operator Interfaces

It is anticipated that there will be at least nine workstations in the MODIS ICC. They are as follows:

- 1 - Supervisor
- 1 - Planning and Scheduling
- 3 - Control and Monitor (1:2 redundancy)
- 2 - Quick-Look Science
- 1 - Data/Communications
- 1 - Instrument Support Terminal (not collocated)

All these workstations will be able to communicate with each other via a network. In addition there will be the following specific external interfaces:

Supervisor: The supervisor will have voice communications capability with the IST, DHC, EMOC, PSC, IMC, and CDHF;

Planning and Scheduling: Planning and scheduling will have voice and data communications with the IST, EMOC, IMC, and CDHF;

Control and Monitor: Control and monitor will have voice and data communications with the IST, EMOC, and DHC;

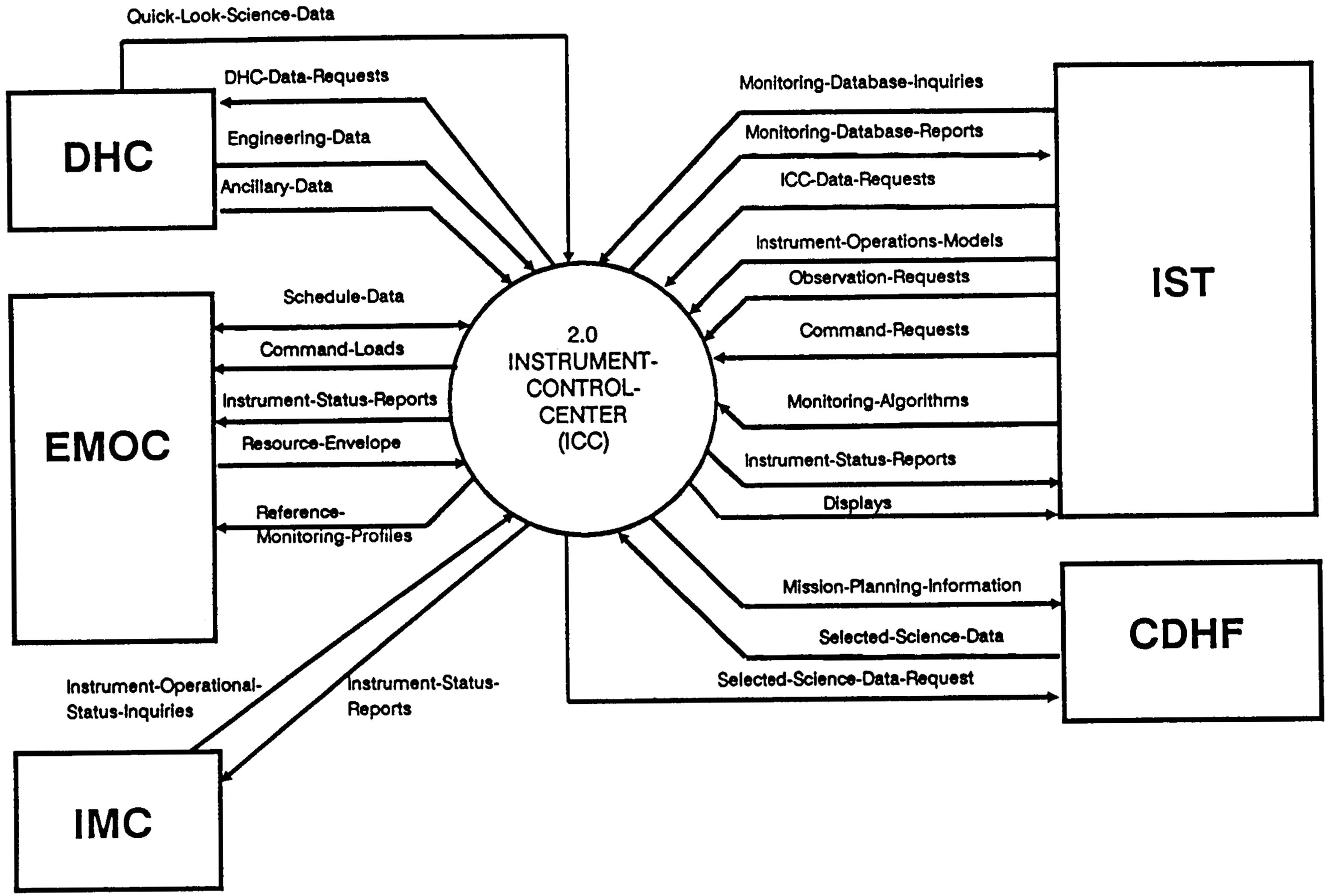


Figure 6. ICC Context Diagram

Science: The science workstations will have voice and data communication with the IST and DHC (CDHF).

6.3 ARCHITECTURAL DESIGN CONSIDERATIONS

This section presents certain considerations which are intended to be useful inputs toward a conceptual and detailed design of this MIDACS element. A summary of computational assumptions; followed by a traffic analysis across all interfaces shown in the context diagram; then, a discussion of storage requirements; and, finally figures pertaining to a conceptual design are presented.

6.3.1 Assumptions

Since MODIS is essentially a 100% duty cycle instrument, its planning and scheduling is generic and therefore relatively simple. Specific data acquisition requests (DARs) are expected to be the exception rather than the rule. Reasons for a DAR may be a special calibration sequence, a deviation from the nominal tilt plan for MODIS-T and the like. It is anticipated that the total of these exceptional data requests will number 100 or less per month. A DAR will be about one page in length (letter size), or about 4000 bytes. The contents will be formatted in a near-natural language and submitted electronically to the IST. Upon review and approval, the MODIS Science Leader will electronically forward the DAR to the ICC's planning and scheduling processor.

The ICC's control and monitor network will support eight local workstations and one remote workstation (the IST). All workstations will be capable of displaying graphics as well as engineering pages (monitor size, i.e., 24 lines by 80 columns). The local workstations may be capable of monitoring up to 80 display pages. It is assumed that 60 of these pages are text at about 2000 bytes each, and 20 pages are graphics. If we assume that the graphics monitors have a 640 x 350 pixel field at one byte each, then the total graphics field may require about 0.224 MB of information. If we assume, however, that for engineering plots, only about 20% of the graphics pixels are illuminated by non-zero values, then these 20 graphics pages may be only about 0.045 MB each. On a 10 Mbps narrow-band network, each of these graphs would require less than 1 second for transfer to the IST. It's assumed that the MODIS instrument will transmit instrument engineering packets separate from science packets for the purpose of ICC monitoring.

The MODIS Science Team Leader may also wish to display selected quick-look science data at the IST. Assume that the IST has a 2048 x 2048 pixel image monitor; that there are two bytes per pixel in three colors; and that the image data to be transferred to the IST will be pulled from the ICC on-line data store onto the narrow-band network. This configuration would require about 20 seconds to build a full MODIS scene at the IST.

The Science Team Leader will have access to all ICC monitoring databases and data stores. The IST will receive periodic instrument status and database reports from the ICC. It is assumed that an instrument status report may be a daily summary of status and events condensed to one letter size page of about 4000 bytes.

It is also assumed that a more extensive database report will be generated weekly, monthly and annually; and, that the Science Team Leader will request a daily database report once per month. It is assumed that each report is 20 letter size pages in length and that a least half is in the form of graphics. Therefore, a database report may consist of:

10 pages of text @ 4000 bytes	=	40 kB
10 pages of graphics @ 0.09*MB	=	0.9 MB
total report	=	0.94 MB

*two engineering plots per gage

On the 10 Mbps narrow-band network, each report would take less than one second to transfer to the IST.

6.3.2 Traffic Analysis

The MODIS traffic analysis for the ICC is somewhat simplified by the anticipated generic nature of the instrument science plan; i.e., all MODIS-N and -T channels on during scene-day, only 15 MODIS-N channels on during scene-night (all else off).

6.3.2.1 *Planning*

For planning purposes, specific data acquisition requests will be the exception rather than the rule. We assume that the Science Team Leader will review, approve and forward 100 DARs per month to the ICC. Therefore,

100 DARs/month * 4000 B/DAR = 0.4 MB/month

On the ICC Context Diagram (Figure 6), the observation requests and command requests coming from the IST to the ICC are assumed to be in the form of a DAR (a recent development).

Observation Requests (from IST to ICC): 0.4 MB/month

Command Requests (from IST to ICC): included in above

6.3.2.2 *Scheduling*

As stated above, the ICC must deal with both generic MODIS plans and DARs. We have assumed under planning that DARs may take the form of a formatted near-natural language. For scheduling traffic passing between IST, ICC and EMOC, it is assumed that each schedule event may be described in one 80-byte line consisting approximately as follows:

12B (time)+20B(mnemonic)+48B(comment) = 80 B/event

a. Generic: There may be 28 mode-change events per day. These relate to turning channels "off" that are "on" and again "on" that are "off" once each orbital revolution at 14 revolutions per day. It is assumed that there will be a standard macro for these events. There have also been identified 644 (reference MIDACS Functional Requirements Document, Appendix D) tilt events per day. Therefore,

   672 events/day * 80 B/event * 30 day/month = 1.6 Mb/month.

b. DARs: Assume there are ten events per DAR at 80 bytes each. Therefore,

100 DARs/month * 10 events/DAR * 80 B/event = 0.08 MB/month,

and the traffic on this interface is:

Schedule Data (between EMOC and ICC): 1.68 MB/month.

Under this section, we will also deal with instrument operations models, resource envelopes, reference monitoring profiles, mission planning information and specialized monitoring data requests.

c. The MODIS instruments will be modeled to the extent that an operations (resource) envelope will be dictated by the EMOC as an operating constraint on MODIS operations scheduling. It is assumed that such instrument models will be provided by the Science Team Leader and maintained by the Science Team Members. It is further assumed that once established, the models will remain static for long periods of time. The traffic on the interface is therefore:

Instrument Operations Models (from IST to ICC): <1 kB/month.

d. The operations envelope (formerly called a resource envelope) for Eos instruments is provided to the EMOC by the PCS. The envelope is expected to change slowly with time. The resource allocation to EMOC is managed by EMOC and partitioned out to instrument scheduling centers like the MODIS ICC. It is expected that the MODIS envelope, once established will remain unchanged for long periods of time. The traffic on this interface is therefore:

Resource Envelope (from EMOC to ICC): <1 kB/month.

e. It is anticipated that an output of the scheduling simulation (modeling) will be an engineering reference monitoring profile which will be made available to IST, ICC and EMOC monitoring functions.

This reference profile may contain such parameters that may be in the operations envelope, like: on/off status, power consumed, temperatures and disturbance torques. As to the number of engineering parameters, it is assumed that there will be at least one voltage and two status indicators for each of the 104 MODIS channels. Further, it is assumed that there will be at least one tilt status indicator for MODIS-T, two power indicators and 10 temperature sensors per instrument. Also included is a factor of two indicating a 100% uncertainty at this point in the GDS development. The net result is a 650 parameter engineering reference monitoring profile. It is assumed that traffic will occur on this interface by exception, i.e., whenever a change occurs. These 650 parameters have been broken down approximately as follows:

status: 320 parms * 1 b/parm * 1/8 b/B = 40B; therefore,

(40B (status) + 12B (time)) * 4 times/rev * 14 rev/day * 30 days/month = 0.09 MB/month.

tilt position: 644 tilt changes per day (MODIS-T); therefore,

(1B(tilt) + 12B(time)) * 644 times/day * 30 days/month = 0.25 MB/month.

Of the remaining 330 parameters, about half change relatively slowly in time, i.e., update every 10 minutes, and the other half may be monitored with simple limit checks which are updated only about once a year. For each parameter update, it is assumed that there will be 12 bytes of time, 20 bytes of mnemonic and 10 bytes of value. Therefore,

((20B (item)+10B (value)) * 330/2 +12B (time)) * 144 chx/day * 30 days/month = 21.4 MB/month;

and for the limit checked parameters:

((20B(item) + 10B (value)) * 330/2 + 12B (time)) * 1/yr * 1/12 yr/month = 414 B/month.

Therefore, the total for this interface is:

Reference Monitoring Profile (from ICC to EMOC): 21.8 MB/month.

f. Mission planning information is another product of the scheduling simulation. It is anticipated that the schedule will be iterated between EMOC, ICC and possibly IST until an approved MODIS schedule is authorized by EMOC. At this time, certain mission planning information may be made available throughout MIDACS. For example, a discipline data timeline (e.g., over land and ocean) could be made available to the CDHF. This timeline may also include special events, e.g., near-real-time science support, which will give notice to the CDHF as to priority resource requirements. It is estimated that each mission planning entry may be an 80 byte field as follows:

12B (time) + 6B(rev.no.) + 1B (land/ocean flag + std scene no.) + 1B (tilt step) + 1B (RT/NRT flags + no. of scenes) + 59B (comment) = 80B.

Assume an entry is made by exception, i.e., only when a change occurs, and that a land/ocean crossing occurs eight times per orbit; that there are 20 scenes per orbit; and, that we accommodate one real-time and one near-real-time event per orbit. Therefore, the traffic on this interface may be,

Mission Planning Information (from ICC to CDHF):

80B * (8 L/O/rev + 20 scenes/rev + 1 NRT/rev) * 14 rev/day * 30 days/month = 1 MB/month.

g. The MODIS ICC must receive a certain amount of wideband (science) data to support instrument monitoring requirements and also to support field experiments. The data itself is discussed under monitoring, below, however the data may have to be requested from the CDHF and/or the DHC. For this, it is assumed that the traffic and coordination will occur on both interfaces. It is estimated that a data request may be an 80-byte field as follows:

12B (start time) + 12B (stop time) + 13B (channel selections) + 2B (packet ID) + 41B (comment) = 80B.

For worst case, assume the MODIS ICC makes one near-real-time data request per orbit. Therefore the traffic on these interfaces may be:

DHC Data Requests (from ICC to DHC):

80B* (1 NRT/rev) * 14 rev/day * 30 days/month = 0.04 MB/month,

Selected Science Data Request (from ICC to CDHF): 0.04 MB/month.

6.3.2.3 *Commanding*

a. Generic commands: It is assumed that there are 28 status change events per day affecting 91 of MODIS' 104 channels, that there are 644 tilt commands per day on MODIS-T and that there are 64 bytes per command. Therefore, assuming that each channel is individually commanded, the monthly traffic of this type is:

   (28 events/day * 91 CMDs/event + 644 CMDs/day) * 64 B/CMD = 0.2 MB/day and 0.2 MB/day * 30 days/month = 6.1 MB/month for generic commanding.

b. DAR event commands - assume, as above, that there are 10 events per DAR and that each event may consist of up to four commands. Therefore, this traffic may be:

   100 DARs/month * 10 events/day * 4 CMDs/event * 64 B/CMD =
   0.3 MB/month.

c. Real-Time Commanding: negligible.

d. Emergency Commanding: negligible.

   Therefore, the total traffic on this interface is:

   Command Loads (from ICC to EMOC): 6.4 MB/month.

6.3.2.4 *Monitoring*

Monitoring data will consist of narrow-band data, i.e., engineering and ancillary data, and wide-band data, i.e., selected MODIS science data. Also discussed below will be monitoring algorithms provided by the Team Leader and maintained by Team Members; ICC displays and instrument status reports; and ICC-data request and database reports.

a. Narrow-band data: This data consists of MODIS engineering and platform ancillary data. It is assumed that by some means to be determined the DHC will directly send the MODIS engineering data packets to the ICC (the alternative may be to have the CDHF strip out engineering data and send it to the ICC). Earlier in this section, we have assumed 650 MODIS engineering parameters; further assume that about half this many will be ancillary data of interest to the ICC.

   As before, the engineering data breaks out as follows, 320 status bits and 330 engineering values of two bytes per parameter; each of the 650 parameters requires three bytes for identification.

Assume there are 325 ancillary parameters of two bytes each and that for each of these, there is two bytes for identification. For every second that all this data is available, assume there are two bytes of time-code information. The traffic, then is as follows:

Engineering Data (from DHC to ICC):

2B (time-code) + 650 IDs * 3 B/ID + 320 status bits * 1/8 B/b + 330 parms * 2 B/parm = 2652 B/sec * 86400 sec/day = 229 MB/day.

Ancillary Data(from DHC to ICC):

325 parms * (2 B/parm + 2(ID) B/parm) * 1/sec * 86400 sec/day = 112 MB/day.

b. Wide-Band Data: For quick-look science data, assume the ICC will receive MODIS instrument data from the DHC at a worst case data rate of about 17 Mbps. There, a wide-band ingest processor will accommodate any eight channels at a composite data rate of about one Mbps for 10^3 seconds, no more than 28 times per day. Therefore,

Quick Look Science Data (from DHC to ICC):

17 Mbps * B/8b * 10^3 sec * 28 day^{-1} = 60 GB/day.

Selected Science Data (from CDHF to ICC):

1 Mbps * 10^3 sec * 28 day^{-1} = 28 = 3.5 GB/day.

This last interface is the alternative wide-band interface for ICC science monitoring. It is felt that the scientist will potentially have more channel selection control via the MIDACS architecture with this interface.

c. The ICC will receive the initial MODIS monitoring algorithms from the Science Team Leader, however, these algorithms are expected to change very infrequently. Therefore, traffic across this interface is negligible.

Monitoring Algorithms (from IST to ICC): <1 kB/month.

d. The MODIS Sciences Team Leader may request data from the ICC in the form of an instrument status report or a display page of particular engineering or science interest for monitoring the performance of the MODIS. This request may also be for some trending analysis report which may come from the monitoring database. In the following subsections, it is assumed that the Science Team Leader will request one each of these per day. The data request should be very brief as far as traffic across the interface. It is assumed that this request can be accommodated in one 80B line. Therefore,

ICC Data Request (from IST to ICC):

2 requests/day * 80 B/request * 30 days/month = 4.8 kB/month.

Monitoring Database Inquiry (from IST to ICC):

1 request/day * 80 B/request * 30 days/month = 2.4 kB/month.

e. A MODIS instrument status report will be made available on request or once per day to the EMOC, IMC, and IST. It is assumed that this report takes the form of a one letter-size page management summary of MODIS T/N status. Therefore, this traffic may be:

Instrument Status Reports (from ICC to EMOC):

4 kB/day • 30 days/month = 0.12 MB/month.

Instrument Status Reports (from ICC to IMC):

4 kB/day • 30 days/month = 0.12 MB/month.

Instrument Status Reports (from ICC to IST):

4 kB/day * 30 days/month = 0.12 MB/month.

It is assumed that MODIS instrument status reports will be distributed routinely each day. The IMC may request a special edition, but that will probably happen very rarely. The traffic on this interface is, therefore:

Instrument Operational Status Inquires (from IMC to ICC): <1 kB/month.

As explained in Section 6.3.1, the IST is a terminal on the ICC. Control & Monitor network and as such, will have access to any of the (or so) engineering display pages routinely monitored at ICC workstations as well as selected image monitoring data. Therefore, the traffic across this interface is:

Displays (from ICC to IST):

60 pages (text) * 2000 B/page * 1/day * 30 days/month = 3.6 MB/month

20 pages (graphics) * 0.045 MB/page * 1/day * 30 days/month = 27 MB/month

25.2 MB/scene * 1 scene/day * 30 days/month = 755 MB/month.

g. A monitoring database report may contain 1000 parameters on any given day and similar reports may be based on increasingly condensed data on a weekly, monthly, and annual basis. Assume that a report is 20 pages in length as given in Section 6.3.1 of which about half is in the form of graphics. Further assume that the Science Team Leader will request a daily report only about once per month; that he will receive a weekly report each week, a monthly report each month, an annual report each year; and that each report is about the same length. Therefore, the traffic on this interface may be:

Monitoring Database Reports (from ICC to IST):

1 report (daily) * 0.94 MB/report * 1/month = 0.94 MB/month,

1 report (weekly) • 0.94 MB/report * 4.33 times/month = 4.1 MB/month,

1 report (monthly) * 0.94 MB/report * 1/month = 0.94 MB/month,

1 report (annual) * 0.94 MB/report * 1/12 month = 0.08 MB/month.

6.3.2.5 *Training and Test*

Between EMOC and the MODIS ICC, assume less than 4 MB/month.

6.3.2.6 *System Input/Output*

A MODIS traffic summary is as follows:

a. Planning - 0.4 MB/month.

   Observation Requests (from IST to ICC): 0.4 MB/month.

   It is assumed that any command requests from the IST will be in the form of a DAR.

   Command Requests (from IST to ICC): included in above.

b. Scheduling: 24.6 MB/month. This traffic will be predominately between EMOC and ICC and has not been adjusted for scheduling iterations.

   Schedule Data (between EMOC and ICC): 1.68 MB/month

   Instrument Operations Models (from IST to ICC): <1 kB/month

   Resource Envelope (from EMOC to ICC): <1 kB/month

   Reference Monitoring Profile (from ICC to EMOC): 21.8 MB/month

   Mission Planning Information (from ICC to CDHF): = 1 MB/month

   DHC Data Requests (from ICC to DHC): 0.07 MB/month

   Selected Science Data Request (from ICC to CDHF): 0.07 MB/month

c. Commanding: 6.4 MB/month.

   Command Loads (from ICC to EMOC): 6.4 MB/month

d. Monitoring - 60.4 GB/day

   Engineering Data (from DHC to ICC): 229 MB/day

   Ancillary Data (from DHC to ICC): 112 MB/day

   Quick Look Science Data (from DHC to ICC): 60 GB/day

   Selected Science Data (from CDHF to ICC): (3.5 GB/day alternative)

   Monitoring Algorithms (from IST to ICC): <1 kB/month

ICC Data Request (from IST to ICC): 4.8 kB/month

Monitoring Database Inquiry (from IST to ICC): 2.4 kB/month

Instrument Status Reports (from ICC to EMOC): 0.12 MB/month

Instrument Status Reports (from ICC to IMC): 0.12 MB/month

Instrument Status Reports (from ICC to IST): 0.12 MB/month

Instrument Operational Status Inquires (from IMC to ICC): <1 kB/month

Displays (from ICC to IST): 786 MB/month

Monitoring Database Reports (from ICC to IST): 6 MB/month

e. Training and Test: 4 MB/month between EMOC and ICC. This data will be partitioned on a non-interference basis on the schedule, command and monitor interfaces.

6.3.2.7 *External Interfaces:* TBS

6.3.3 On-Line Storage Requirements

The sum total on-line storage required by the ICC is estimated to be 1.54 GB. The components contributing to the requirement are described in the following paragraphs.

6.3.3.1 *Planning*

The ICC should accommodate about three months worth of Observation Requests. At 0.4 MB per month, this will require 1.2 MB on-line storage.

6.3.3.2 *Scheduling*

Since an active schedule period is one week in duration, the 25 MB per month from Section 6.3.2.6.b should be adequate for retaining a copy of the active schedule, and three successive layers of candidate schedules. Therefore, 25 MB of on-line storage should suffice.

6.3.3.3 *Commanding*

Again, since command uploads are performed for every active schedule period, 7 MB of on-line storage will accommodate a month's worth of command loads.

6.3.3.4 *Monitoring*

The ICC must have enough on-line storage to accommodate three days worth of ingested Engineering and Ancillary Data. Therefore approximately 1 GB of on-line storage will be required.

The ICC must have enough on-line storage to accommodate ingesting eight MODIS channels at 1 Mbps for 1,000 seconds. Approximately 200 MB of on-line storage will be adequate to accommodate any one quick-look event.

To accommodate all reports and various models and algorithms, the ICC must have an additional 300 MB of on-line storage.

6.3.4 Off-Line Storage Requirements

The ICC's off-line storage requirements are driven by the science monitoring requirements. The sum total off-line storage requirement is about 100 GB of which about 5% is to be retained permanently in the ICC as described in the following paragraphs.

6.3.4.1 *Planning*

The ICC must have off-line storage sufficient to accommodate about one years worth of MODIS instrument plans. This will require about 5 MB of off-line storage.

6.3.4.2 *Scheduling*

Looking at the scheduling components of Section 6.3.2.6.b, one can see that only about 10% of the monthly storage allocation need be saved for archive. Therefore, 2.5 MB per month for 15 years results in 450 MB off-line storage requirement. It is clear that only the approved (or authorized) schedules are archived.

6.3.4.3 *Commanding*

As with authorized schedules, all authorized command loads will be archived by the ICC. Therefore, 7 MB per month for 15 years results in 1.26 GB of off-line storage.

6.3.4.4 *Monitoring*

Approximately once per day, an engineering trending compression will be invoked which should reduce the daily volume of engineering and ancillary data by about 1,000. Therefore, 0.341 MB per day for 15 years results in about 1.9 GB off-line storage requirement.

For quick-look science data, it is assumed that data stored on-line for each event will be off-loaded to off-line storage devices prior to the next event. At 3.5 GB per day, approximately 0.1 TB of off-line storage would be required to store one month's worth of data. It is assumed that one or more scientists would be interested in taking possession of this data and therefore, will not be retained longer than one month by the ICC.

Figure 7 shows the functional data flow for the ICC. The data rates shown on the interfaces, show the peak traffic between the elements.

The peak wideband data rate from the DHC to ICC will be about 17 Mbps. It is expected that these data rates will occur for a duration of about 1,000 seconds, less than 28 times per duty.

The narrow-band data rate shown in probably very near the peak data rate since engineering and ancillary data are assumed to be handled nearly uniformly in time.

The peak data rate between the ICC and IST will be limited by the capacity of the narrow-band network. The network used here for illustration purposes has a 10 Mbps capacity. The greatest stress to this capacity is assumed to be during times when monitoring image data is transferred to the IST.

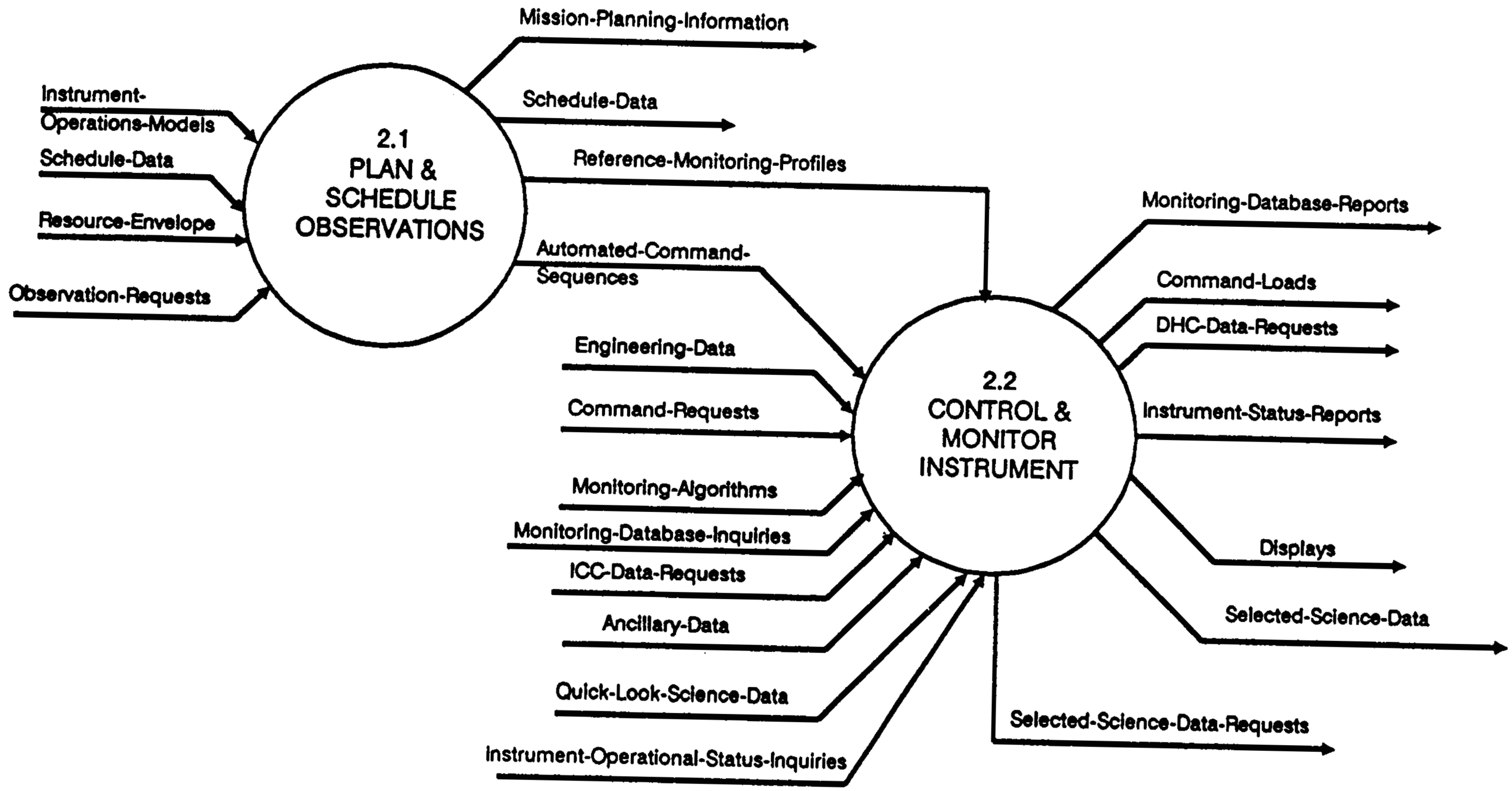


Figure 7. ICC Functional Data Flow

The peak data rate to the CDHF will probably occur once a week when the MODIS ICC publishers the mission planning information. This information will be broadcast in approximately two Mb bursts over the above 10 Mbps network. Therefore the peak data rate is expected to be limited by the net traffic on the ICC control and monitor network and by the communications link which connects to it from the CDHF.

The exception to the above estimate comes if the quick-look science (wide-band) ingest processor is located in the CDHF. In that event, the ICC will receive about 3.5 GB per day across that interface. This quick-look science data would be received at the ICC in real-time at a rate of about one Mbps. These data rates would occur for duration of about 1,000 seconds, less than 28 times per day.

The peak data rate to the EMOC is expected to occur once a week during the ICC's routine planning and scheduling process. All traffic is expected to be limited by the ICC's 10 Mbps narrow-band network. The three heaviest traffic items are schedule data, monitoring profile and command loads. None of these is expected to occur simultaneously with any other. The largest will be the 44 Mb burst of reference monitoring data about once per week.

Figures 8 and 9 illustrate a conceptual architecture for the MODIS ICC. The architectural drivers will be the wide-band ingest processor and the image data processors. The ingest processor and the image data processors. The ingest processor is felt to be a combination of a decommutator and processor. The decommutator extracts and formats (serial to parallel) requested data (8 of 104 channels) into data packets. In the following calculation, this ingest processor is assumed to operate on two byte words; that 70% of its capacity is being utilized; that 40% additional capacity is needed for the hardware implementation of the decommutator; that the ingest processor manages the traffic on the wide-band network at seven instruction per word; and, that a factor of five is applied to reflect equal volume of levels zero, one and two image data going to output devices (i.e., stop, display), and two additional channels being brought out of the store to output devices.

8ch/104 ch 16.83 Mbps * B/8b * W/2B * 1 instructions/W * 5 * 1.4 * 1/0.7 = 6 MIPS

Therefore, the ingest processor would require about a six MIPS capability.
The image processor must be capable of raising level zero data to level two in real-time. The input data rate will consist of eight (1km) channels at a composite 0.85 Mbps; the cumulative path length for level two is assumed to be 40 instructions per list; and, we assume 70% processor utilization.

0.85 Mbps * 40 * 1 = 50 MIPS

Therefore the main processor requirement is for 50 MIPS.

7. ELEMENT SPECIFICATION FOR THE TEAM MEMBER COMPUTING FACILITY (TMCF)

Figure 10 is a context diagram of the TMCF. The TMCF is distributed and is composed of project-provided computing facilities used to develop scientific and calibration algorithms, verify and validate data, and to generate some specialized data sets (see Figure 11). As an organizational unit, the TMCF is where the Science Team Leader provides planning and coordinating for the MODIS Science Team Members and for MIDACS. The TMCF is a distributed network of workstations at Science Team Member

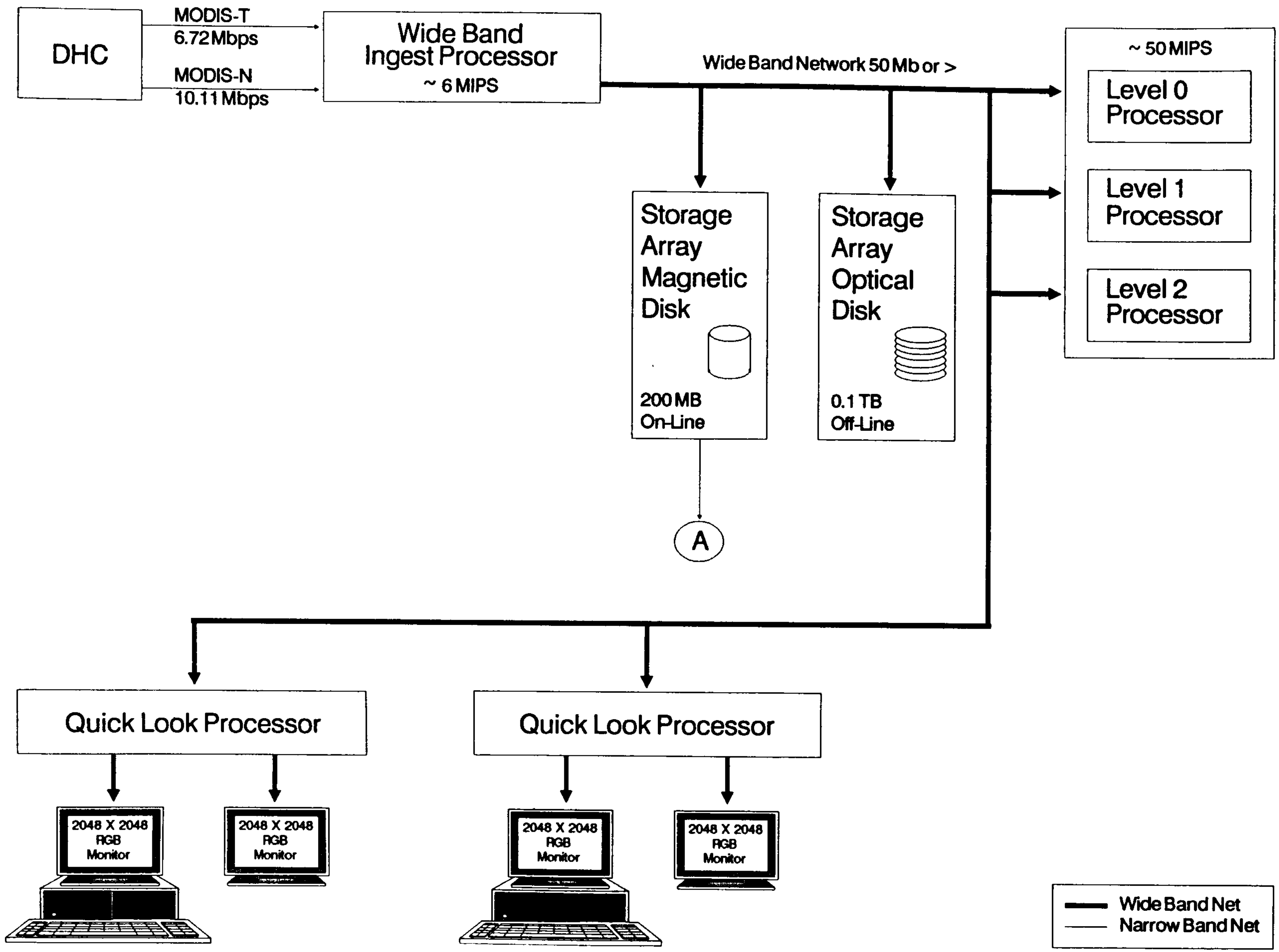


Figure 8. ICC Wide Band Context

ORIGINAL PAGE IS
OF POOR QUALITY

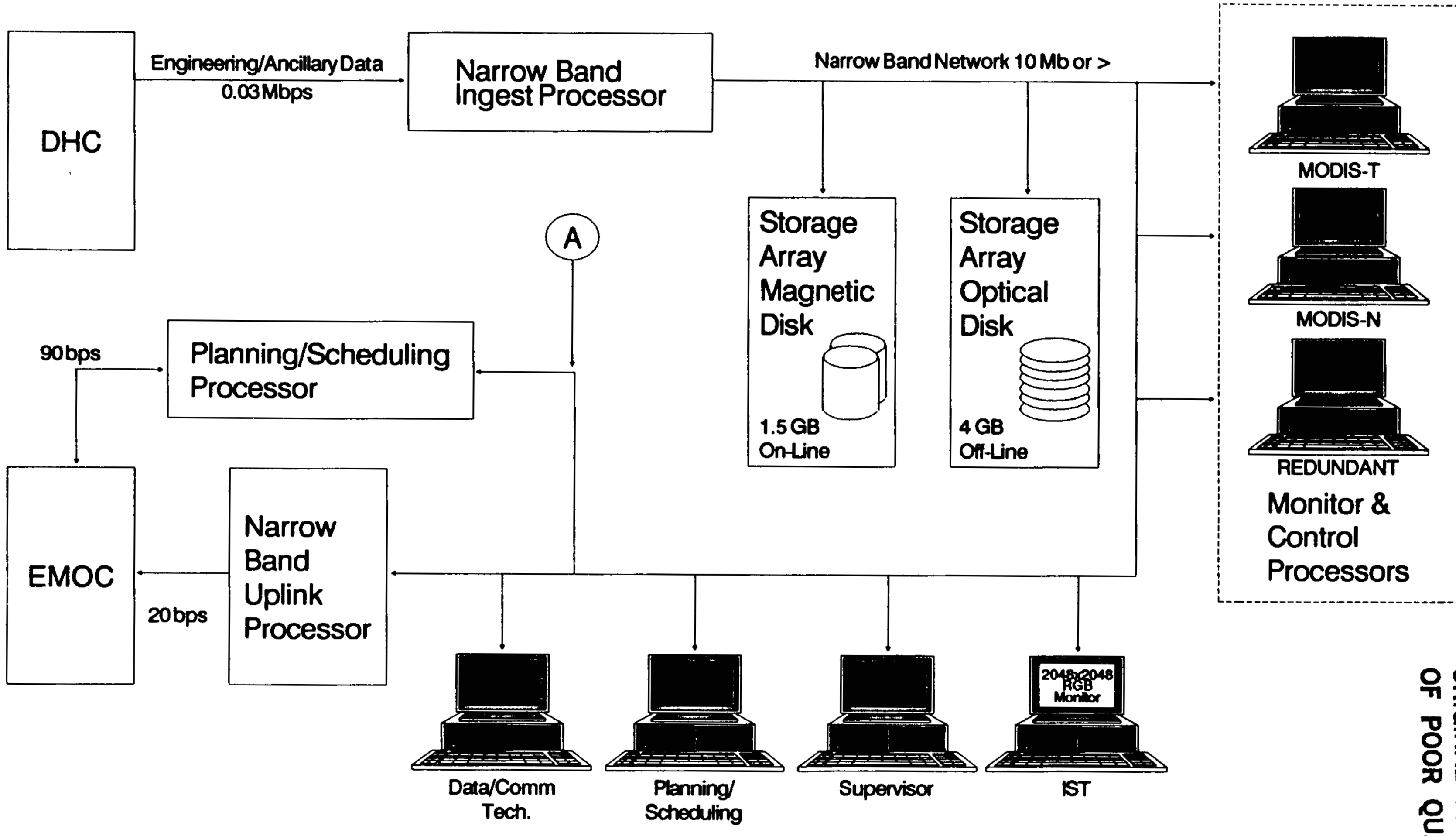


Figure 9. ICC Narrow Band Context

ORIGINAL PAGE IS OF POOR QUALITY

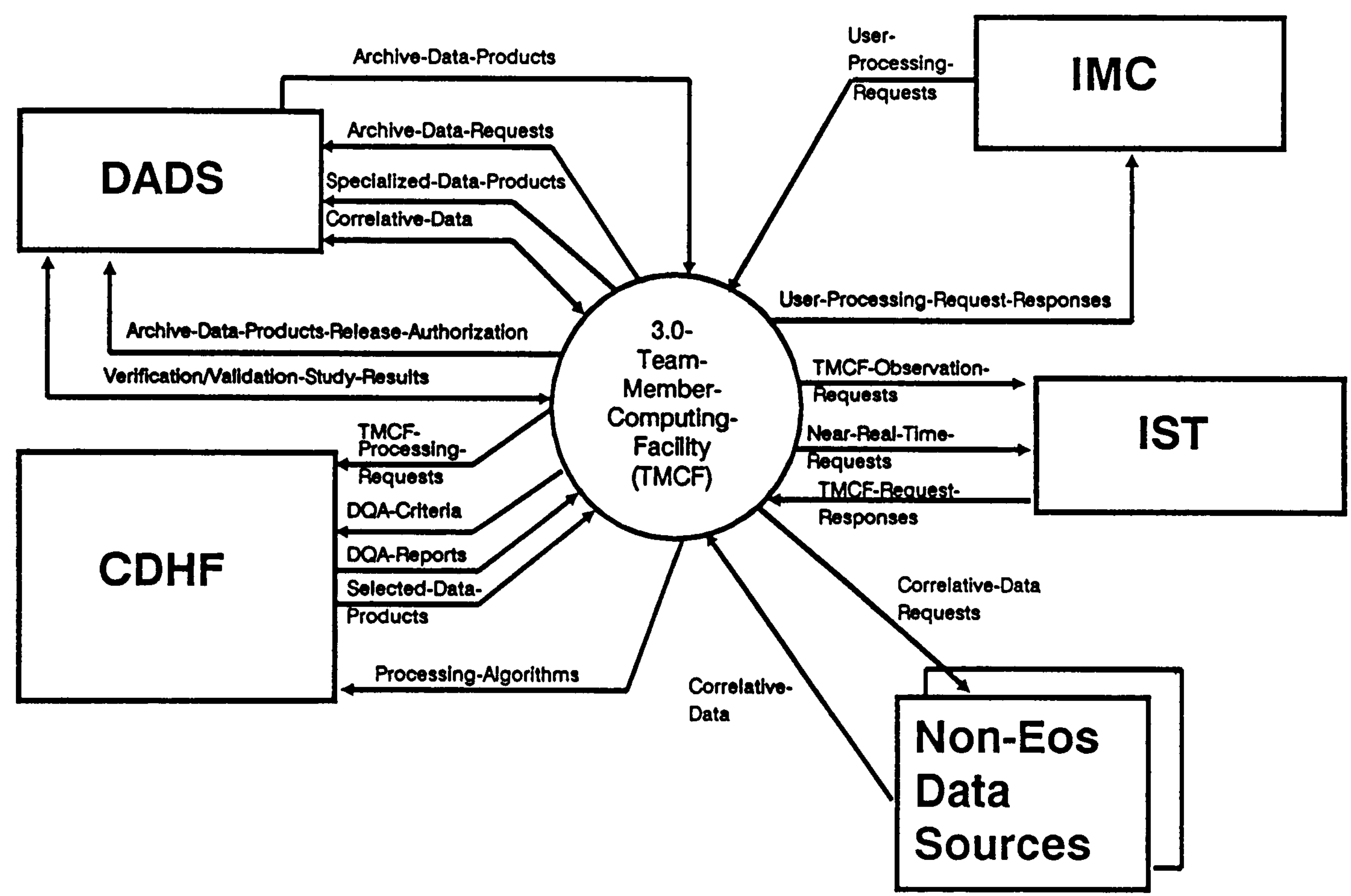


Figure 10. TMCF Context Diagram

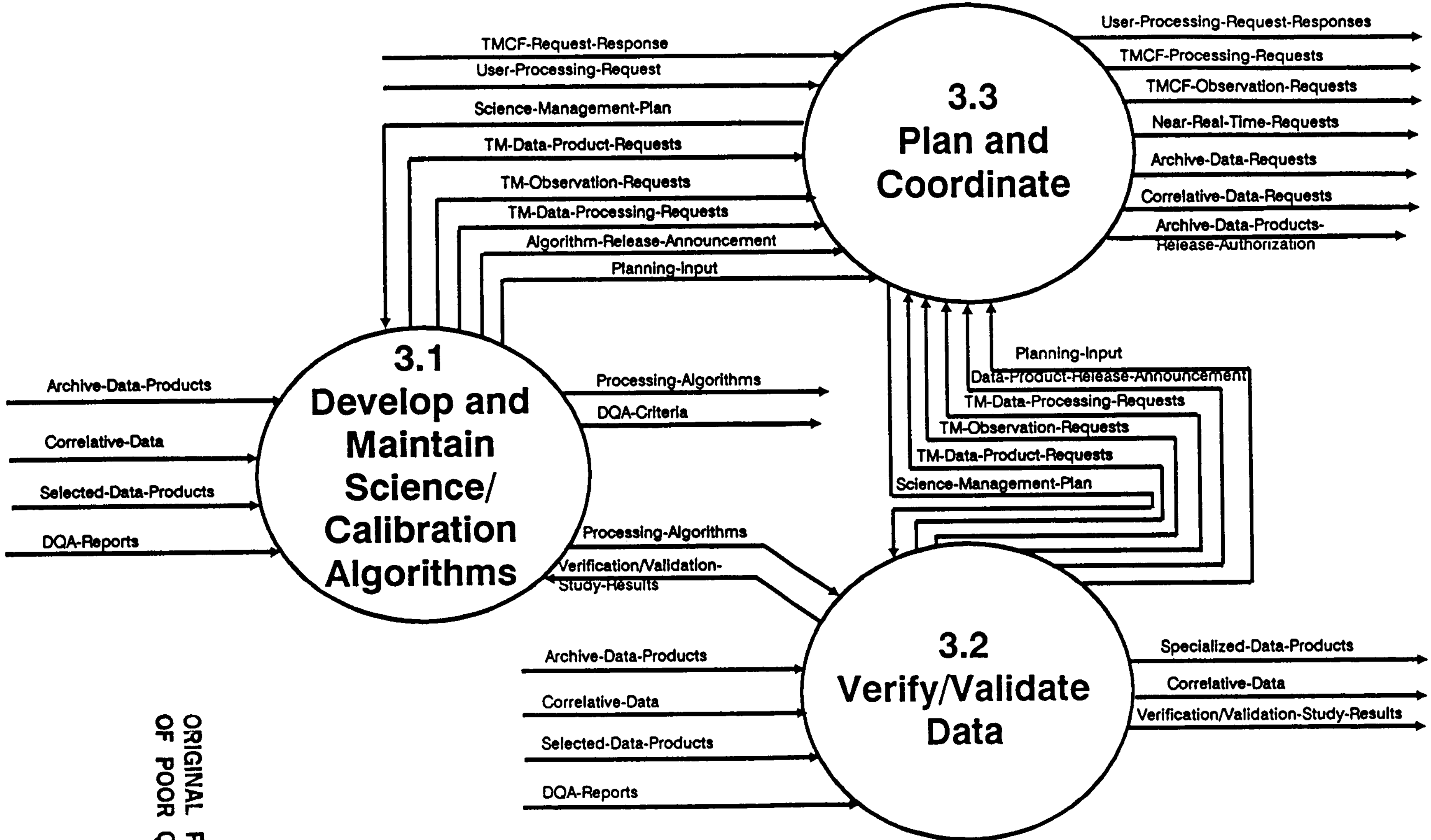


Figure 11. TMCF Functional Data Flow

ORIGINAL PAGE IS
OF POOR QUALITY

locations and perhaps temporarily at the site of a field experiment. The network node at GSFC is where several Science Team Members, including the Science Team Leader, are expected to reside. Also resident at the GSFC will be the Calibration Support Team (CST) and a Science Support Team (SST) which is a group of computer scientists engaged in making the algorithms developed by the Science Team Members more efficient and in developing software which would have general utility to all Team Members. These computer scientists may be part of the CDHF as well. The GSFC TMCF node is central to the TMCF network and will probably have the greatest amount of project-provided computing facilities.

In addition to communications which may be required between the TMCF's computers, each TMCF will require communications with: 1) the CDHF, 2) the DADS, 3) the Information Management Center (IMC), and 4) non-EOS data sources. A communication controller located at the GSFC TMCF node can provide a means of communication for remote TMCF's directly to the CDHF for acquisition of Level 2 data, to the DADS for acquisition of older data products, or to an any other component of MIDACS.

Communications will consist of textual messages (as with the IST), interactive database inquiries (as with the IMC), and the exchange of data products, browse data products, and algorithms (as with the CDHF and the DADS). A low rate (e.g., 9600-baud line) can handle many of these communications, but a high speed bus will be required for the transfer of large data sets.

7.1 FUNCTIONAL REQUIREMENTS

7.1.1 Algorithm Development

Four types of algorithms can be identified as input/output algorithms, plotting and imaging algorithms, calibration algorithms, and science algorithms. The Science Team Members using the TMCF are responsible for the development of calibration and science algorithms. Input/output algorithms and plotting and imaging algorithms which are of use to all Science Team Members will be developed either by the CDHF or by computer scientists at the GSFC TMCF node, or will be purchased from commercial sources. The code will, of course, be in a high level language such as Fortran 77 so that it can be developed on a workstation and transported to mainframe computer. If the CDHF acquires a computer which incorporates vector processing or parallel computing, the ANSI Standard languages require modifications since the input/output is handled differently. These language extensions, which take advantage of the parallel nature of the computer, may have effects on the choice of the Science Team Member workstations and may change the way code will be developed and tested.

7.1.2 Verify/Validate Data

The TMCF will verify spectral radiance measurements, and validate derived geophysical parameters. To a limited extent, as resources allow, it may also generate specialized data products. Sample activities are listed below. The Calibration Support Team (CST) performs studies of instrument properties using standard sources such as lamps, blackbodies, or the solar diffuser plate, or using other stable sources such as Earth targets or the moon. The CST also participates in MODIS/HIRIS intercomparisons, comparisons of MODIS to other EOS platform instruments, and comparisons to other satellite measurements. The CST may also perform other studies such as detector noise analyses, calibration independent studies, spectral calibration studies, and so forth. All these studies are designed to monitor instrument stability and provide traceability to NIST standards.

A typical Science Team Member using the TMCF will be engaged in validation studies, which are designed to certify that a MODIS derived geophysical parameter and its corresponding in-situ measurement are in agreement. They will also want to compare their measurements to other satellite derived measurements of the same or similar parameters. Both of these activities require image analysis such as attribute clustering or image differencing, which in turn imply that an image processing workstation that can import and transform Landsat, Spot, HIRIS, TIMS, GOES, AVHRR, and other images be available to the Science Team Member.

A Science Team Member will also generate special data products such as prototypes of potential standard data products, or Level 4 type data products which are rarely generated and require limited processing capabilities.

7.1.3 Plan and Coordinate

The science team will be responsible for science planning and coordination. The Science Team Leader will develop a Science Management Plan, coordinate the Science Team Member observation requests, processing requests, and data product requests. Ground truth or in-situ field experiments will be organized so that they are coordinated with the MODIS instrument operations.

7.2 PERFORMANCE REQUIREMENTS

The only two performance requirements identified in the MODIS Functional Requirements document are:

a. The TMCF shall verify that the pointing accuracy of the MODIS instrument data sets is within the required accuracy of 0.1 kilometers.

b. The TMCF shall provide the CDHF with calibration coefficients in sufficient time such that Level 1B data processing can occur on schedule.

The first requirement will probably be satisfied by using DQA Criteria, which will involve algorithms on the CDHF computers to automatically and continuously check the pointing accuracy of the MODIS instruments as data are processed at the CDHF. A portion of the Level 1A data (TBD %) may be transmitted from the CDHF to the CST within the TMCF for more detailed analysis.

The second requirement requires that DQA Criteria and/or control charts or their equivalents be incorporated within the CDHF processing so that the calibration can be routinely updated following pre-established procedures. In addition a portion of the Level 1A data stream (about 2.0% of the Level 1A traffic) will be sent to the CST for more detailed analyses. The pre-established criteria may evolve as a result of these analyses and evidence may be found to warrant reprocessing the data.

There may be other as yet unspecified performance requirements relating to validation studies and to algorithm development. These requirements might be that code developed for the generation of Level 2 products on the CDHF computer take advantage of vector or parallel processing capabilities of the CDHF computers so that the most efficient use of the machines be made. A performance requirement for validation studies might be that the accuracy of the data products produced be within some specified tolerance. These performance requirements will be specified either in a later version of this document or may be specified by the Science Team Members.

7.3 INTERFACES

The TMCF Context Diagram summarizes the interfaces of the TMCF within MIDACS. The Calibration Support Team has additional contacts with the instrument contractor and all Science Team Members concerned with calibration. All of these interfaces are low rate (e.g., 1200/2400/9600 baud) except for the transfer of Level 1A data from the CDHF to the TMCF, the transfer of specialized data products from the TMCF to the DADS (low rate on average, but high peak data rate transfer), and possibly the transfer of correlative data from non-EOS sources to the TMCF, which may come in on tape or other media. The interfaces will be described in more detail after a traffic analysis is presented.

7.4 ARCHITECTURE DESIGN

In this section, preliminary traffic estimates are made for a typical TMCF starting with the CST, which is assumed to be one of the most active nodes and hence the driver for the TMCF hardware selection as a whole. This is followed by a summary of the resulting computational assumptions and a description of the hardware architecture and software architecture.

7.4.1 Traffic and Sizing Assumptions

The following analysis of the data flow traffic starts with an analysis of the data flows to and from the CST which is one of the more active nodes on the TMCF network and for which we have a better idea of their activity than the other TMCF nodes. As a result, it is expected to be one of the drivers for the bandwidth of the busses used in the network. After the analysis of the CST traffic flows, a traffic analysis of "light, medium, and heavy global" Science Team Members as defined in 2.2.3.3 will be made and compared to the CST traffic analysis. In anticipation of this later analysis, we will conclude that the "heavy user" acquires an average of 21 GB/month, the "medium user" obtains 2.1 GB/month and the "light user" obtains 0.2 GB/month. The average user thus acquires about 8 GB/month which is more the calculated 6 GB/month which is derived by assuming a typical Science Team Member is one-half as active as the CST. We now proceed to provide the reasoning which leads us to these numbers for the average traffic flows starting with an analysis of the CST traffic flows.

7.4.1.1 *Traffic Assumptions for CST*

As a convenience in analyzing the traffic flows, the concept of an image cube is introduced here. An image cube is defined as an array of observations, with two spatial dimensions and one spectral dimension. A typical image cube might have a 1024 by 1024 spatial dimension and 100 wavelengths, or 105 million pixels. Typically an image cube will be a view of an Earth target at 100 wavelengths which has resulted from about 2.5 minutes of observations. However, any 2.5 minute period of observation will produce an equal amount of data, whether that observation is a scan of an Earth target, a lamp plus the dark side of the Earth, or a spectral calibrator. The traffic analysis will estimate the number of image cubes acquired per unit time period such as one year and then sum them up to deduce the total traffic flow.

For on-board targets such as lamps or blackbodies the following assumptions are made as to their frequency:

a. MODIS-T observes lamps at any one of 3 levels plus off. These observations are assumed to be performed once per week. The lamps are assumed to be on at each level for 2.5 minutes. The dark side of the Earth will also be observed during these sequences.

b. MODIS-T observes the solar diffuser plate for a 2.5 minute time period once per day. The frequency for this measurement is not yet determined. It may be once per orbit or once per week depending upon the instrument design.

c. MODIS-T observes the spectral calibrator for a 2.5 minute time period once per week. The dark side of the Earth is also observed.

d. MODIS-N observes lamps at any one of 3 levels plus off. These observations are assumed to be performed once per week. The lamps are assumed to be on at each level for 2.5 minutes. The dark side of the Earth will also be observed during these sequences.

e. MODIS-N observes the solar diffuser plate for a 2.5 minute period once per day. The frequency of this measurement is not yet determined. It may be once per orbit or once per week depending upon instrument design.

f. MODIS-N observes the spectral calibrator for a 2.5 minute period once per week. The dark side of the Earth is also observed.

g. From the above eight assumptions we estimate 1615 image cubes per year are acquired by both MODIS-T and MODIS-N for calibration work (i.e, 52 x 4 +365 + 52 + 52 x 4 + 365 = 1198 image cubes/year).

For verification that the radiance observations are stable over time, selected Earth targets will be viewed from time to time. For MODIS the following assumptions are made:

a. Ten verification sites exist.

b. Each site is re-visited once per week. We estimate five image cubes per week or 260 image cubes per year are taken by each MODIS instrument.

The moon will be observed on six orbits each lunar month, with the moon observed 8.24 seconds each time. The data generated by this procedure is equivalent to less than one image cube per month and may be neglected in the traffic analysis. Satellite comparisons of MODIS to other satellite instruments are anticipated. We assume:

a. MODIS will be compared sporadically to other satellites as opportunity presents itself. We estimate 12 image cubes per year will be generated from these comparisons.

MODIS will be compared to other EOS platform instruments other than HIRIS as opportunity presents. The following assumptions are made:

a. Each HIRIS verification site is also observed by MODIS.

b. Periodic comparisons (monthly) are made as follows: 1) MODIS-N vs. TIMS, 2) MODIS-N vs. AIRS, 3) MODIS-N vs. HIRIS, 4) MODIS-N vs. MODIS-T, and 5) MODIS-T vs. HIRIS. From these we expect 6 image cubes per month or 72 image cubes per year.

Calibration independent studies may also be performed. We estimate 1 image cube per month or 12 image cubes per year.

From the above paragraphs, about 1542 image cubes per year will be devoted to calibration. Since each image cube has 0.1 GB, there will be 154 GB/year for calibration. Each image cube is 105 Mpixels (1024 x 1024 x 100). MODIS-T and MODIS-N will acquire about 288 such image cubes per day or 105000 image cubes per year. Calibration thus represents about 1.5% of the Level 1A data load handled by the CDHF (44000 GB/year).

These instrument properties, such as spectral or spatial calibrations, or detector noise analysis, will be analyzed using the data acquired above such as night image cubes for the visible channels to perform detector noise studies.

DQA reports will be used by the Science Team Members and the CST. The following assumptions are made:

a. Each report consists of 5 KB of information.

b. 30 reports per month are issued which gives a data transfer of 0.3 MB per month. This is a low data rate which can be handled by a 9600 baud line.

The above paragraphs describe the traffic flows for the CST alone for which we have more information on their operations. Since the CST is assumed to be one of the most active nodes on the network, the above traffic flow analysis should be adequate to specify the bandwidth of the transmission busses.

7.4.1.2 *Traffic Assumptions for Typical TMCF*

We assume at this point that each Science Team Member will have one-half the traffic of the CST and that there are 15 Science Team Members. This last assumption will be examined in terms of "light, medium, and heavy users" at the end of this sub-section. Based upon these assumptions and the following specified assumptions, the table below summarizes the estimated traffic flows for each of the external interfaces for Science Team Member's as given in the TMCF Context Diagram:

Central Data Handling Facility (CDHF)

a. Provide science algorithms and their updates

    1. Assume 100 algorithms and their revisions with accompanying documentation is sent each month from all Science Team Members.

    2. Assume each algorithm is 2000 lines time 80 characters per line and 1 bytes per character or 16 MB.

    3. This implies 16 MB (error checked) per month which can be accommodated on a 9600 baud line.

b. Acquire Level 1A data for special processing

    1. The analysis above gave 154 GB for the CST each year, or 13 GB/month.

    2. With 15 Science Team Members, each about one-half as active as the CST, we get an additional 100 GB/month for all TMCF. Further com-

ments on this data flow are provided below. Each Science Team Member averages about 6 GB/month from this analysis. Using the concept of "light, medium, and heavy users" of section 2.2.3.3, Science Team Members average about 8 GB/month and much of this data will be acquired from the DADS.

3. For local users, a high rate (100 Mbps) fiber optic line could accommodate this data load. Typical workstations may not be able to handle data transfers of more than a few Mbps in any case.

4. For remote TMCF users, mid-1990's technology may still be equivalent to T1 lines (1.5 Mbps). Therefore, "heavy users", as previously defined in section 2.2.3.3, at remote sites may need to acquire a substantial portion of their data on magnetic or optical media.

5. Real-time and near-real-time processing will generate data flows from the CDHF to the TMCF. Section 8 discusses this data flow which does not strongly effect the monthly averages.

c. Request CDHF computing time

1. Assume 2Kb per request and 100 requests per day, then 0.75 MB of data per month is transferred.

2. This can easily be handled by a 9600 baud line.

d. Acquire Data Quality Assessment (DQA) reports

1. Assume each report consists of 5 KB of information.

2. Assume 30 reports per month are issued which gives a data transfer of 0.15 MB per month for each Team Member.

3. This is a low data rate which can be handled by a 9600 baud line.

Instrument Support Terminal/Instrument Control Center (IST/ICC)

a. Provide science scheduling requirements

1. If we assume that the planning traffic from the IST to the ICC is equal to the planning traffic from the TMCF to the IST, we have 0.4 MB/month being transferred.

2. This can easily be handled by a 9600 baud line.

Data Archive and Distribution System (DADS)

a. Provide science algorithms and coefficients, with history of their use, for archiving

1. Assume 100 algorithms and their revisions with accompanying documentation is sent each month from all Science Team Members.

2. Assume each algorithm is 2000 lines time 80 characters per line or 0.16 MB.

3. This 16 MB data rate per month can easily be handled by a 9600 baud line.

b. Receive archive data sets for further studies

1. Assuming that the rate for a "heavy user", as previously defined in Section 2.2.3.3, a scienc team member acquires 21 GB/month.

2. For local users, a high rate (100 Mbps) fiber optic line could accommodate this data load. Typical workstations may not be able to handle data transfers of more than a few Mbps in any case.

3. For remote TMCF users, mid-1990's technology may still be equivalent to T1 lines (1.5 Mbps). Therefore, "heavy users" as defined previously, at remote sites may need to acquire a substantial portion of their data on magnetic or optical media.

c. Provide special science data sets

1. Assume that the data rate is 0.1% of the MODIS data collection rate for each TMCF.

2. For 15 Science Team Members, this gives 1.8 GB/month sent to the DADS.

3. Since these data sets are sent infrequently, but individually may consist of image type data, it is appropriate for remote Science Team Members to send it on magnetic or optical media. Local Science Team Members can use the fiber optic line.

d. Data requests to this and other EOS DADS via IMC

1. Assume that Science Team Members occasionally communicate with their own DADS via IMC and always with the other EOS DADS via IMC.

2. These data requests are sporadic and consist of text type communications, which are easily handled by a 9600 baud line. An estimated 5 to 10 MB/month of queries are generated per month.

Non-EOS Data Sources

a. Acquire correlative data

1. Much of this data will be acquired on media such as tapes although the option that some of it could arrive via a fiber optic line is left open.

2. Electronically acquired data is included in the 100 GB per month rate quoted above.

As an alternative to the above traffic analysis, we now consider that the Science Team Members are "light, medium, and heavy global" users as defined in section 2.2.3.3. If we convert the number of bits requested to image cubes, we fine that a heavy user requests about 600 image cubes, a medium user about 60, and a light user about 6. If they make these requests once every three months, the light, medium, and heavy users are requesting 24, 240, and 2400 image cubes per year. The CST, on the other hand requested 1550 image cubes per year, making it a heavy user, as was earlier surmised. With five of each type of user on a 15 member Science Team, a total of 13320 image cubes would be requested, which corresponds to about 12.7% of the Level 1B data.

7.4.1.3 *Summary*

The analysis above stated that the CST would acquire about 13 GB/month and each Science Team Member would average about 6 GB/month. These data are being acquired predominantly from the CDHF, but a significant portion of the Science Team Member's data may come from the DADS. The "heavy user" acquires an average of 200 image cubes per month or 21 GB/month. The "medium user" obtains 2.1 GB/month and the "light user" obtains 0.2 GB/month. The average user thus acquires about 8 GB/month which is more than the assumed 6 GB/month derived by assuming a typical Science Team Member is one-half as active as the CST. A 100 Mbps bandwidth fiber optic bus in a candidate TMCF network is capable of handling these data flows, although some of the larger data flows might be handled by shipping the data on hard media.

7.4.2 Summary of Computational Assumptions

The following few paragraphs are an attempt to estimate the computational needs of the TMCF computers. The analysis will proceed in two parts:

1. We will examine the CST and its needs.

2. We will examine the SST and its needs. Both of these teams are located at the GSFC TMCF node and are under the science team leader's control. Some team members may receive project-provided computing facilities and they will indicate their needs in these requests. Other team members will be using their own computers for algorithm development, data validation, etc. There may be three large TMCFs, one each for land, ocean, and atmospheres. Other TMCFs may be more limited in their capabilities and needs. Each TMCF will in reality be different from every other TMCF. These sections try to concentrate on facilities that are project-provided and for which information is available about their operation.

Some calibration work will be done at the CST workstation, but a large portion of the calibration work will need to be automated and performed on the CDHF computers (see Section 8.3.2). The traffic analysis above deduced that about four image cubes per day would be sent to the CST for more detailed studies and for interactive studies for which a workstation is a more suitable environment than the CDHF computer. If the number of arithmetic instructions per pixel is 70 (simple linear regression with 10 samples) or 210 computer instructions, then a 1.4 MIPS computer is required assuming 70% utilization. For standard deviations (arithmetic instructions per pixel = 15; computer instructions per pixel = 45), we add 0.5 MIPS to above. For control charts (arithmetic instructions per pixel about 5 and hence computer instructions about 15), we add 0.5 MIPS to above. The peak rate for interactive analysis is three times the average rate since only 8 hours per day out 24 is devoted to this activity. If each image undergoes histogram analysis or equivalent, which is about twice as computationally intensive as converting counts to radiances (see Section 8.3.2), then an additional 2 MIPS is required. If the workstation is used in determining pointing accuracy as well, this requires about 31 computer instructions per pixel (see Section 8.3.2), requiring about 1 MIPS.

The implied total is 10.2 MIPS per workstation from this analysis. Not all possible calibration activities have necessarily been included in this scenario and the calculated MIPS rating is for average conditions. The MIPS rating here may need to be increased to perhaps twice the rate quoted (i.e., 21 MIPS) since the following types of analyses may be performed as well: (1) importing HIRIS and other images and converting them to the same spatial and spectral resolution as MODIS images, (2) detector crosstalk analyses, (3) modulation transfer function studies, (4) dark current analyses, etc. These analyses will probably be performed on a second workstation which is also a backup to the first workstation. Each MODIS instrument will thus have four workstations. The frequency of performing these special studies is not yet determined and will be settled in a later edition of this document. If the CST is assumed to consist of four workstations, two for each MODIS instrument, and each workstation has about 20 MIPS, then the total computing capacity of the CST would be around 82 MIPS.

Alternatively, using Landsat Thematic Mapper calibration as a model, a 5000 by 5000 pixel scene with 100 detectors required 30 minutes of CPU on a VAX 11/780 (0.9 MIPS rating). The extent of the calibration work done in this time period was not specified but it may be only a subset of the procedures in the paragraph above. For MODIS (4848 detectors, 3 times code length, 1024 by 1024 pixel scenes), each scene would require 6.1 hours CPU on a VAX 11/780. At 4 scenes per day, this comes to 4 times 6.1 = 24.4 hours. Multiplying by three again for interactive work, we get 73.2 hours on a 0.9 MIPS machine each day. If this is to be accomplished in 8 hours, then 73.2/8 x 0.9 = 8.2 MIPS machine is required. The 8.2 MIPS rating is close to the 10.2 MIPS rating estimated above.

Next, we consider the amount of random access memory needed. If a software system such as the Land Analysis System (LAS) used by Landsat is used, about 200,000 lines of code are involved. A compiled version would required about 16 MB of core. An additional 16 MB of core would be useful for image storage, which means a workstation with 32 MB of random access memory would suffice.

Finally, we consider the storage needs of the CST. To store one year of image cubes, each workstation needs about 100 GB of storage. However, only a fraction needs to on-line. With 4 GB of storage, approximately 4000 workstation monitor images of 1024 by 1024 resolution could be stored. More input by people actually involved in the calibration work will be needed to provide a better estimate of these storage capacities.

Network and external links are required by the workstation. With a 100 Mbps fiber optic bandwidth bus, each image cube, which is actually 100 images on the workstation monitor or 105 MB or 840 Mb, could be transferred in 8 seconds, which is adequate.

Science Support Team (SST) Node

The SST will be engaged in validation activities and with analyses which increase computing efficiency. The details of their activities are not as yet specified, but it is estimated that the computational needs for these activities will equal or exceed the needs of the CST.

7.4.3 Hardware Architecture

The network node at GSFC which includes the CST activities is envisioned to consist of four workstations, two for each MODIS instrument. One workstation is used for interactive image processing activities such as occur in verification studies. A second workstation also provides a backup to the first workstation when it is down and may be involved in some routine calibration activities depending upon how the CST chooses to use the CDHF computers. It will also be used for calibration algorithm development.

The instrument engineering workstation (one each for MODIS-T and MODIS-N) will have the following uses:

a. Development of calibration algorithms.

b. Interactive analysis of MODIS imagery.

c. Interactive analysis of engineering data.

d. Interactive analysis of DQA reports.

e. Access to the CDHF batch processor.

f. For calibration, capability to derive calibration coefficients on subset of MODIS data.

To serve these purposes, a machine with about 10 to 20 MIPS processor, 32 MB of main memory, 1 GB of disk storage, 1024 x 1024 pixel display with 12 bit planes, would probably suffice, although all these numbers are preliminary.

The image analysis processor/scientific algorithm developer (one each for MODIS-T and MODIS-N) workstation would be used by most Science Team Member's and the CST. Images from MODIS, HIRIS, TIMS, Landsat, Spot, GOES, etc. must be imported to be used in validation and verification studies. It is also a backup to the engineering workstation. The workstation would have the following uses:

a. Development of scientific algorithms.
b. Validation studies.
c. Verification studies.
d. Generation of specialized data products.
e. Access to the CDHF (via high rate line).
f. Access to the IST (via 9600 baud line).
g. Access to Science Team Leader TMCF (via 9600 baud line).
h. Access to other Science Team Members (via 9600 baud line).
i. Access to the DADS (via high rate line).

A high MIPS rating is advantageous for verification studies which involve image processing. A preliminary MODIS TMCF workstation would then be:

a. CPU with speed of 10 MIPS for high speed image analysis.
b. 32 MB of main memory (16 MB for software and operating system).
c. 4 GB of on-line storage (for 4000 images on the workstation screen).
d. Links to the DADS/CDHF (via high rate lines).
e. Links to EOS network (via 9600 baud and possibly a high rate line).
f. Links to the IST (via 9600 baud line).

g. Links to Science Team Leader/Science Team Members (via 9600 baud line).
h. Links to HIRIS (via 9600 baud line and possibly a high rate line).
i. Tape drive/optical disk reader/etc. required for image input/output.
j. A high resolution color monitor for analysis of MODIS and other images.

Currently available workstations with a 10 MIPS rating may be adequate for both the engineering workstation and for the image analysis verification studies, but a 20 MIPS rating is preferred. This implies four such workstations will exist for MODIS as a whole for calibration work. The redundancy of workstations can be utilized for backups when they have down time.

7.4.4 Software Architecture

It is a safe assumption that the computer architecture at the TMCF's will generally be quite different from the computer architecture at the CDHF and, in fact, quite different from each other.

The architecture of computers can be described in many ways. There are two main catagories of architecture: serial and parallel architectures. Within parallel architectures we can distinguish those with switched processors from those with a network of processors. Switched processor computers can be further subdivided into those with shared memory and those with distributed memory. Network computers can be divided into those with mesh, cube, hierarchical, or reconfigurable networks of combined processors and memories.

The MODIS instrument, as with any cross-track scanner, generates image type data. To compute calibrated radiances or geophysical parameters, the same mathematical operations are performed on each pixel within the image. This suggests that use of a parallel architecture computer may be the most efficient approach so that computations for all of the pixel elements can be performed simultaneously. The alternative of serial type computations (where all the computations are performed on one pixel, and then repeated on the next pixel, and so forth) would appear to be an inefficient approach.

Consider then two extremes: 1) The TMCF workstations as serial type computers. 2) The CDHF computer as a massively parallel processor using a cube network. In this case, software developed on the workstations may not efficiently run on the parallel processor. In fact, since parallel processors often require that ANSI standard languages have extensions to them for parallel data input and output, it is possible that the TMCF software might not run on the CDHF computer at all. One solution is to forbid parallel processors from being employed. (This may implicitly occur if ANSI language standards are strictly adhered to). A second solution is to wait for ANSI standards to be modified to account for recent technical developments.

If the CDHF computer is a massively parallel type, it is probably more efficient for algorithm development work to be done on the mainframe computer rather than at workstations whose architecture will undoubtedly be quite different. If the CDHF computer is a more conventional serial design or even a switched processor design, most of the algorithm development can be done on separate workstations. The choice of the CDHF computer architecture will have major effects on the TMCF network and hardware just as it will affect the development of software. Within the MIDACS, there is a need for a central software development group to convert code developed on the TMCF computers to code which efficiently utilizes the architecture of the CDHF computer. This group could be part of GSFC TMCF node or part of the CDHF.

In summary, the CDHF computer architecture will have major implications on the TMCF potentially altering the coding of the software and the TMCF hardware choice.

8. ELEMENT SPECIFICATION FOR THE CENTRAL DATA HANDLING FACILITY (CDHF)

The context diagram for the CDHF is illustrated in Figure 12. External interfaces and all data flows are noted.

8.1 FUNCTIONAL REQUIREMENTS

The MODIS CDHF conceptual architecture will be centered about three major functional drivers (see Figure 13).

8.1.1 Receive DHC Data

a. The CDHF shall receive Level-0 data and ancillary data from the DHC. The Level-0 data shall be in a form that is sequenced by time, by focal plane, by along-track distance, and by band configuration along the scan direction. Ancillary data shall have been checked and validated by comparisons with orbit and attitude reference profiles. The Level-0 data shall have had transmission errors corrected and redundancies removed.

b. The CDHF shall perform TBD acceptance checking of Level-0 data and request any necessary retransmission of data from the DHC.

c. The CDHF shall perform any necessary reformatting of received Level-0 data and will append standard headers.

8.1.2 Produce Data Products Levels 1-4

a. Level-1A data shall be data which are reformatted reversibly, with earth location, calibration data, and other ancillary data.

b. Level-1B data shall be Level-1A data to which the radiometric calibration algorithms have been applied, perhaps irreversibly, to produce values of the instrument measurements (e.g., radiances or irradiances) to which the Earth location and zenith angle algorithms have been applied.

c. The Level-1 processor shall organize the science data into logical data records that are TBD. The natural blocking of the MODIS data is by observation swath (64 km x 1502 km for MODIS-T and 8 km x 1502 km for MODIS-N).

d. The Level-2 processor shall receive Level-1B data and any ancillary data necessary for the Level-2 processing step. The Level-2 product shall contain geophysical parameters derived from Level- 1B data by the application of geophysical parameter algorithms.

e. The Level-3 processor shall receive Levels-1 and - 2 data and any ancillary data necessary for the Level-3 processing step. The Level-3 data product shall contain Earth gridded geophysical parameter data including radiances, etc., from Level-1 averaged or composited in time and in space.

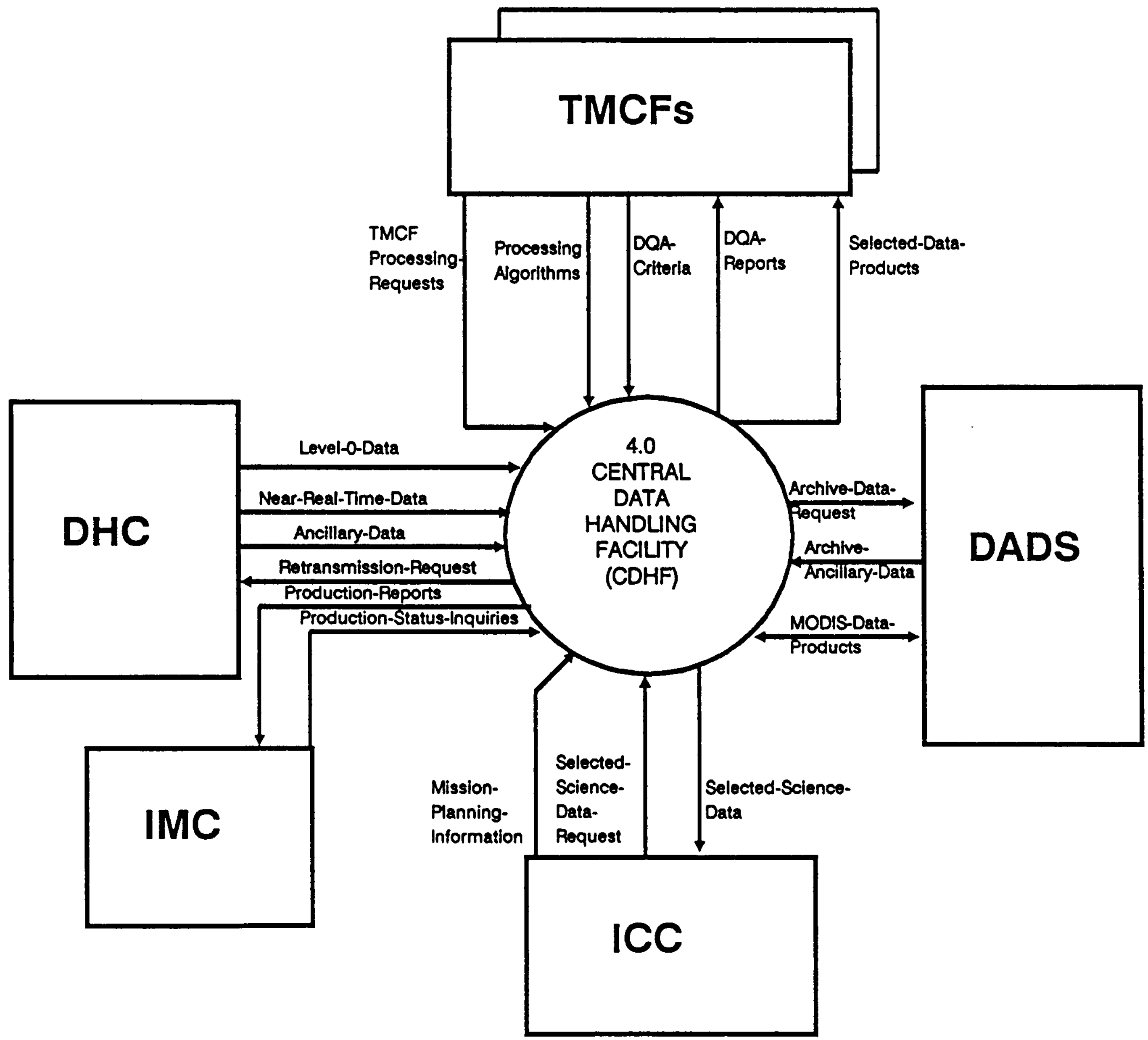


Figure 12. CDHF Context Diagram

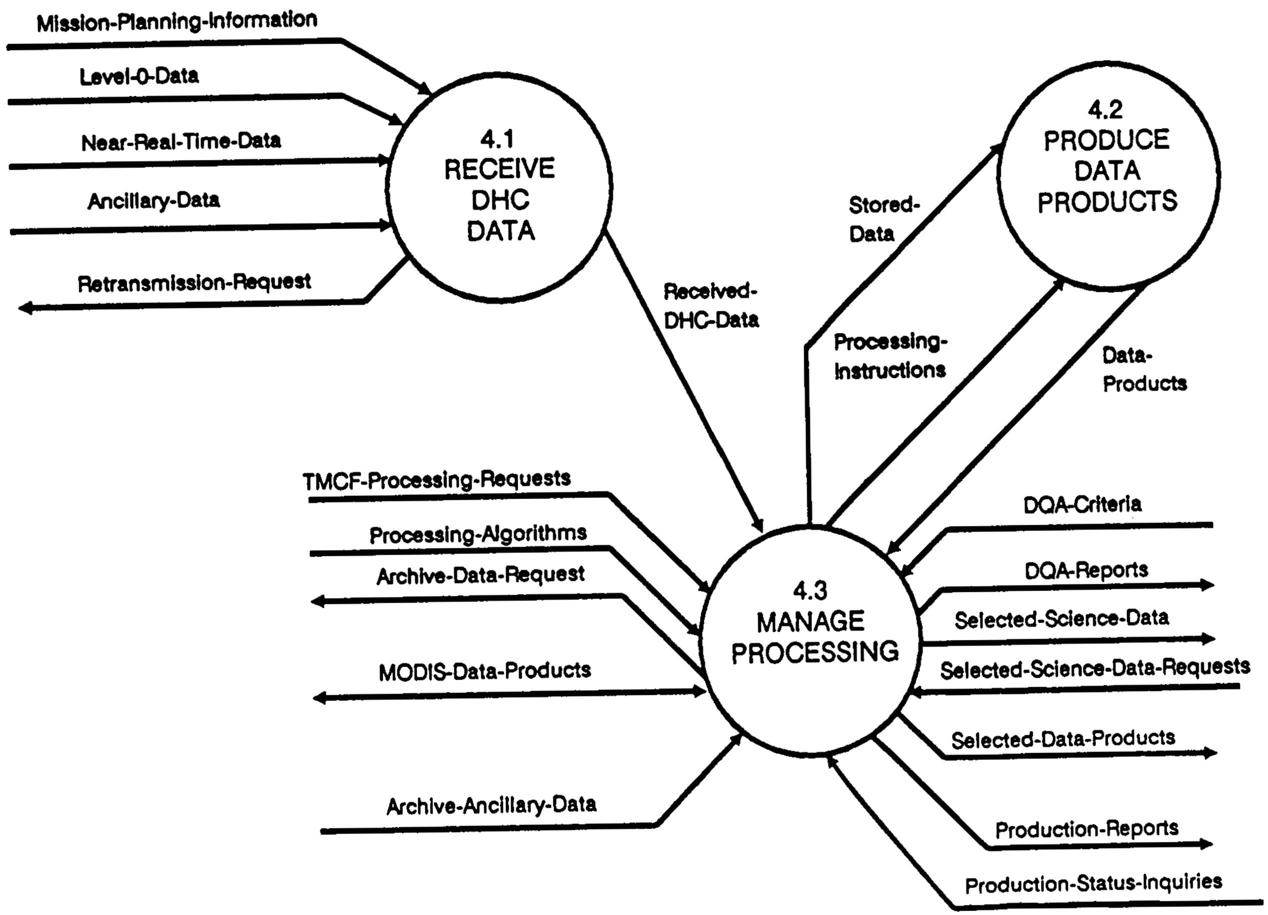


Figure 13. CDHF Functional Data Flow

f. The Level-4 processor shall receive Levels-1, -2, and -3 data and any ancillary data or correlative data necessary for the Level-4 processing step. The Level-4 product shall contain TBD analysis of the lower levels of MODIS data products.

g. The Levels-1, -2, -3, -4 processors shall each be capable of performing reprocessing, special request processing, near real-time processing, and backlog processing in addition to the standard processing of data.

8.1.3 Manage Processing

a. The CDHF shall acquire calibration coefficients and algorithms from the TMCF for use in routine processing of Level-1A data to Level-1B.

b. The CDHF shall temporarily catalog and store data for routine processing, special processing, or reprocessing.

c. The CDHF shall provide selected data products, (subsets of standard and near real-time data) to the TMCF upon request.

d. The CDHF shall transmit selected subsets of MODIS science data to the ICC for monitoring of instrument performance.

e. The CDHF shall perform daily job accounting functions and send reports to the IMC.

8.2 PERFORMANCE REQUIREMENTS

The MODIS CDHF functional drivers will have the following performance requirements:

8.2.1 Receive DHC DATA

a. The CDHF will accept Level-0 data, near real-time data, and ancillary data at transfer rates of up to 300 Mbps.

b. The CDHF will perform acceptance checking of TBD record length and content.

c. The CDHF will reformat and append headers of applicable records to meet TBD file criteria.

8.2.2 Produce Data Products Levels 1-4

The following performance requirements assume a delay of at most 24 hours from observation for delivery of Level-0 data to the CDHF.

a. All Level-0 data sets received within a 24-hour period shall be completely Level-1A processed within 12 hours. The results of Level-1A processing shall be available to authorized investigators within 48 hours of original observation.

b. All Level-1A data shall be completely Level-1B processed within 48 hours of observation.

c. All Level-1B data period shall be completely Level-2 processed within 72 hours of observation.

d. All Level-2 data shall be completely Level-3 processed within 96 hours of observation.

e. The amount of data to be processed to Level-4 is equal to TBD times the raw MODIS instrument data.

f. All near real-time data processing shall be completed within three to eight hours of observation at all Levels (Level-1A to -3). The precise timeliness requirements will depend on the specific data requirements of the field experiments.

g. The data processing function shall routinely process reduced volume (0.1% - 10%) Level-0 to Level-3 data on a daily basis for use as browse data sets.

h. Produce catalog.

8.2.3 Manage Processing

a. The CDHF shall provide a storage system utilizing fast access storage for online working storage and for protection against data loss. This storage system shall be capable of storing data at the volume of TBD Gbytes/day.

b. The CDHF shall distribute standard MODIS data products to the DADS as soon as they are generated.

c. The CDHF shall distribute near real-time data products to the requestor within 3-8 hours of observations.

d. So as to optimize the throughput in the CDHF and DADS, processors should be designed to process physical record n, while simultaneously writing data associated with physical record n-1 and reading physical record n+1. In this manner, the processing will not be I/O bound and the highest throughput will be achieved. While processing data from the physical record, the calculations will be repetitive, thus providing the opportunity for vector or parallel processing.

8.3 LEVEL-0 TO LEVEL-1B CDHF PROCESSING ESTIMATES

The following subsections (8.3.1 through 8.3.3) analyze the data processing involved in transforming Level-0 data to Level-1B data and provide an estimate of the sizing of the computing system, expressed in MIPS, required for this processing.

8.3.1 Level-0 to Level-1A Processing

Level-0 to Level-1A processing consists of (according to the MIDACS Functional Requirement Document and the Phase A study): 1) acceptance checking, 2) reformatting data (unpacking and re-ordering), and 3) appending headers. The data will be unpacked from 12 bits to 16 bits. The Phase A studies estimate that 15 computer instructions per pixel are required to unpack the data and to re-order the data. With 15 instructions per pixel a 28 MIPS rating is implied. We assume that appending headers is neglible computationally (1 MIPS). We assume that acceptance checking involves making sure data received are data requested, etc. and is low level computationally (1 MIPS). Initially, in these calculations we assume 24 hours is available to process 24 hours of acquired data. This assumption is examined later.

The following MIPS rating is obtained: 1) acceptance checking equals 1 MIPS, 2) reformatting data equals 28 MIPS, and 3) appending headers equals 1 MIPS. The total Level 0 to Level 1A MIPS rating then is 30.

8.3.2 Level-1A to Level-1B Processing

This section describes the Level-1A to -1B processing requirements based in part upon some timing runs performed on an IBM 3081 computer. Several simple types of calibration scenarios are considered followed by some comments on possible complications such as striping. The section ends with a summary of the Level-0 to -1B processing requirements and a comparison of different estimates of these requirements.

Three candidate calibration equations were tested by timing runs on an IBM 3081 computer (nominal MIPS rating of 14.5) to derive the MIPS rating in going from Level 1A data to Level 1B data. This processing consists of converting raw voltage counts (represented by X here) to spectral radiances (Y). The three candidate calibration equations, tested by timing runs on the IBM 3081, were:

a. Linear: $Y = A*X + B$,
b. Temperature corrected: $Y = A*(1. + C*T)*X + B$, and
c. Non-linear: $Y = D*(1. + E*T)*X^2 + A*(1. + C*T)*X + B$.

A,B,C,D and E are calibration constants such as the gains and offsets of the detectors.

For linear radiometric calibration, the number of computer instructions per pixel is 13 and the MIPS rating is 24 based upon the notes of the MODIS Data Systems Team (9/23/88). The MODIS-T calibration may be non-linear (J. Barker, 11/15/88). The number of computer instructions per pixel for linear calibration with temperature corrections is 22, implying that a MIPS rating of 41 is required. The number of instructions per pixel for a non-linear calibration including temperature corrections is 40 implying that a MIPS rating of 74 is required. The number of arithmetic instructions per pixel for earth location calculations is 436 (using a Nimbus-7 subroutine to calculate Earth location). If 1% of the pixels have their locations calculated to this detail and the other 99% of the pixels are interpolated using 6 arithmetic instructions per pixel, the average number of arithmetic instructions per pixel is 10.3. Assuming three computer instructions for each arithmetic instruction, 31 computer instructions per pixel are required for each pixel. This implies that a MIPS rating of 58 is required.

The following MIPS rating is implied by the above calculations:

a. A basic linear calibration requires 24 MIPS

b. A calibration with a temperature correction to the gain requires 41 MIPS

c. A non-linear calibration with a temperature correction to the gains requires 71 MIPS

d. Earth location calculations require an additional 58 MIPS

The total rating for Levels-1A to -1B is in the range of 82 to 129 MIPS. Implicit in deriving these numbers is that all aspects of calibration have been included. However, in scanning spectrometers, images may appear with stripes which arise from individual detectors not being fully self-consistent with one another. Using the Land Analysis System (LAS) of Landsat as a guide, the de-striping of an image requires about 88% more time than it does to perform a linear calibration of the same image. This implies that, if striping is a problem, an additional 45 MIPS of computing power is needed, or 135 MIPS if re-processing also requires de-striping. Proper calibration techniques and good instrument design may make this processing requirement unnecessary.

In addition, other computations may be required if the instrument suffers from astigmatism, field curvature, coma, spherical aberration, or other distortions. These effects may require that corrections be applied to an image which are dependent upon the position within the image. The computational requirements for these calculations, if required, are at present unknown. It is anticipated that the MODIS instruments will be built such that these computations will not be required.

There are two other computations that may occur for which we do not yet have reliable estimates of their computational load:

a. Derivation of calibration coefficients will be automatically derived using one orbit's worth of data. This will prevent scene-to-scene discontinuities when mosaics re formed. Presumably all the standard lamp or blackbody measurements and all the space observations will be used in these computations, which requires about 2% of the total data load. Thus, it is not expected that these computations will add significantly to the required processor capabilities.

b. A modulation transfer function (MTF) reversal may be performed. The computational load from this procedure is potentially intense and since it is uncertain if it will occur, we do not yet include this in the MIPS rating. Previous satellite measurements have not done MTF reversals so there is no easy way as yet to determine its impact on the processing needs.

8.3.3 Summary

The total estimated MIPS rating required to go from Level 0 data to Level 1B data is, therefore, in the range of 112 to 159 MIPS with the potential of an additional 45 MIPS for de-striping. The computers are assumed to be 70% utilized in this analysis.

The above MIPS ratings do not consider reprocessing. With two reprocessings going on concurrently with the processing, the MIPS ratings above are multiplied by three to give a range of 336 to 477 MIPS (plus a probable 135 MIPS for de-striping). If near-real-time processing is also considered to equal about 3% of the original processing we must add 3 to 5 MIPS to the above to give a range of 339 to 482 MIPS. This MIPS rating is in a sense a worse case since all detectors are assumed to be non-linear; whereas probably only the thermal infrared detectors will be non-linear.

The above MIPS ratings also do not consider down time for the mainframe computers. Using the computer center at the National Center for Atmospheric Research as a guide, the computers will be down about 10% of the time for maintenance. This increases the needed MIPS rating to between 373 to 530.

The MIDACS Functional Requirements Document (page E-3) derives 89 to 336 MIPS for Level 0 to 1A and 134 to 804 MIPS rating for Level 1A to 1B assuming 24 hours is available for this processing. The total of 223 to 1340 MIPS brackets the range of estimates derived here. The high level MIPS rating of 1340 assumes that 240 computer instructions per pixel are involved. The analysis here, in contrast, derives 70 to 90 computer instructions per pixel.

The Phase A study, on the other hand, deduced that a 22 MIPS machine would be adequate for Level 0 through 1A processing for all the data within 8 hours based upon the equation they use. To process Level 1A to 1B is about 30% more computationally intensive, meaning an additional 27 MIPS is required for a total of 49 MIPS. Their assumed data rate is only about one third of the presently assumed data rate; so the revised MIPS rating will be 165.

It should be noted that two other MODIS data system documents contain estimates of the Level-0 to Level-1 processing requirement. Thus, there are three estimates, each based upon a different set of assumptions: 1) This study gives 373 to 530 MIPS (plus a probable additional 135 MIPS for de-striping). 2) The Functional Requirements Document gives up to 1340 MIPS. 3) The Phase A study deduced about 165 MIPS. Automatically calculating calibration coefficients and MTF reversal calculations may increase the required MIPS ratings derived here.

8.4 LEVELS-2 AND -3 PROCESSING ANALYSIS

Three MODIS data product scenarios have been developed from current Levels-2 and -3 processing algorithms and are described in the following sections. Then in Section 8.4.4, a methodology for estimating the routine processing requirement for all MODIS Level-2 through -3 data products is presented. This overall estimate presented below is based only upon the processing requirements for these three scenarios. This estimate will be improved as more processing scenarios for more individual products are included.

8.4.1 Vegetative Index (C. Justice, GSFC)

a. Input Data Volume: Six visible/near IR channels at 500 m and 2 visible (250 m) channels yield 6(500 m channels) x 4 + 2 (250 m channels) x 16= 56 equivalent channels per 1 km pixel. The MODIS-N data rate without over-sampling is 10.11 Mbps for 94 equivalent channels per 1 km pixel. Assuming that 30 percent of the globe is land and only daytime data (40 percent) is used, the volume of Level-1B input data is

$$1.2 \times 10.11 \text{ Mbps} \times 86400 \text{ sec} \times 56/94 \times 0.30 \times 0.4 = 0.075 \text{ Tb/day.}$$

The factor of 1.2 accounts for a 10 percent increase in the Level-0 data from merging and reformatting and another 10 percent increase from Level-1B processing (EosDIS Baseline Report, July 29, 1988).

b. Processing Requirement: Current processing of AVHRR data on an HP 1000 A-900 require 10 min of CPU time per orbit (50% algorithm and 50% I/O) to process GAC (Global Area Coverage) data, which is 1 km AVHRR data processed on board the satellite to yield one radiance for each block of five pixels in every third scan line. Data are read in from tape and output to disk.

MODIS Processing at 0.5 km of the GAC product (global map) would require

$$10 \text{ min/orbit} \times 15 \text{ orbits/day} \times (5 \text{ km} \times 3 \text{ km}/0.5^2 \text{ km}^2) = 9 \text{ kmin/day.}$$

We assume that processing four local area coverage (LAC) maps (North America, South America, Africa, and Eurasia) requires about the same processing as for the global map. We further assume that the vegetative index is additionally produced at 250 m for 10 percent of the global land areas, and a factor of 0.1/0.25 = 0.4 is obtained. Then the algorithm processing time (50%) becomes

0.5 x (1 + 1 + 0.4) x 9 kmin/day = 180 hours

If the HP 1000 ran at an effective capacity of 3.0 MIPS and the conversion 3 MIPS/MFLOPS is appropriate (EosDIS Baseline Report), then the processing capacity per product would be

180 hours/24 hours x 3.0 MIPS / (3 MIPS/MFLOPS) =
7.5 MFLOPS/product

where one product, the vegetation index, has been derived. This estimate does not include atmospheric correction processing.

c. Output Data Volume: Since the size of each GAC and LAC 10-day map is 2.6 MB, the average daily data volume of the five maps is

5 x 2.6 MB x 8 bits/B /10 days = 10.4 Mb/day

8.4.2 Chlorophyll Product (W. Esaias, GSFC)

a. Input Data Volume: The Level-1B volume of 64 MODIS-T channels with no-oversampling is

1.2 x 0.4 x 6.7 Mbps x 86400 sec/day = 0.3 Tb/day.

The volume increase factor of 1.2 arises from Level-0 and -1B processing (EosDIS Baseline Report). It is also assumed that MODIS-T data is measured for 40 percent of each orbit.

b. Processing Requirement: Current Processing of CZCS data on a MicroVax III requires 1 min of CPU time to process 0.7 MB of radiance data (5 visible channels). At 8-bit digitization, algorithm processing requires

1 min/0.7 MB x 1B/8 bits x 8 bits/chan x 5 chan/pix = 7.2 min/Mpix.

It is estimated that radiance correction takes 60 percent of the current processing time and calculation of 2 derived products (chlorophyll and diffuse attenuation coefficient) takes the remaining 40 percent. Then the processing time for the radiance calculation is equivalent to that of 3 derived products and the processing time of 7.2min is for 5 equivalent derived products.

The MODIS pixel (1 km) rate is

1294 FOV x 6.59 km/sec x 86400 sec = 0.74 Gpix/day

Assuming that 70 percent of the MODIS data is over ocean, one day of daytime ocean data would require an algorithm processing time of

7.2 min/Mpix x 0.74 Gpix/day x 0.7 x 0.4 = 25 hours

If the MicroVax III runs at an effective rate of 2.7 MIPS, then the computing capacity per equivalent derived product would be

25 hours/24 hours/5 products x 2.7 MIPS/ (3 MIPS/MFLOPS) = 0.2 MFLOPS/product

c. Output Data Volume: The daily output Level-2 volume of 20 water leaving radiances, 20 diffuse attenuation coefficients, 9 geophysical parameters, and earth location (4B for latitude and 4B for longitude) per 1 km pixel would be

(49 + 4) words/pix x 0.74 Gpix/day x 0.7 x 0.4 x 2B/word = 22 GB/day

The Level-3 product would consist of arrays of size 2048 x 1024 for each of the 49 parameters above and their standard deviations. These products will be produced daily, weekly, and monthly so that the overall Level-3 volume will be

49 parms x 2 x 2048 x 1024 x 2B/parm x (1 + 1/7 + 1/30) = 484 MB/day

8.4.3 Cloud Properties and OLR (J. Susskind, GSFC)

a. Input Data Volume: Fifteen thermal channels of MODIS-N have a volume of

1.2 x 0.74 Gpix/day x 15 channels x 12 bits/channel = 0.16 Tb/day.

b. Processing Requirement: The current processing of HIRS-2/MSU data on a CYBER 205 takes 20 min of CPU time (80% algorithm processing and 20% I/O) to process one day of temperature profiles and moisture profiles at 60 km resolution and 4 geophysical parameters (effective cloud fraction, cloud top pressure, outgoing longwave radiation (OLR), and longwave cloud radiative forcing) at 30 km resolution. Data is read in from disk and output to disk.

Computation of 2 cloud parameters (effective cloud fraction, and cloud top pressure) requires about 9 percent of current processing. Averaging and gridding data requires about 15 percent. It was estimated that OLR and longwave cloud radiative forcing each would take 8 percent of the current processing time. Producing these products at 1 km MODIS resolution instead of at 30 km yields a factor of 30 x 30.

With these assumptions, algorithm processing would take

0.8 x 20 min x (0.09+2 x 0.08) x 30 x 30 = 60 hours.

Assuming an effective rating of 150 MFLOPS for the CYBER-205, we get a capacity for 4 products of

60 hours/24 hours x 150 MFLOPS = 375 MFLOPS

and an average capacity of 94 MFLOPS/product.

If the four products are derived at 5-km resolution, the processing requirement drops by a factor of 15.7 as shown below. In this case, an increase in processing time of 15 percent is required for averaging and gridding. Therefore, the processing time becomes

$$0.8 \times 20 \text{ min} \times (0.09 + 2 \times 0.08 + 0.15) \times (30 \text{ km}/5\text{km})^2 = 3.84 \text{ hours}$$

Thus, again assuming 150 MFLOPS for the CYBER-205, the four 5-km resolution products require

$$3.84 \text{ hours}/24 \text{ hours} \times 150 \text{ MFLOPS}/4 \text{ products} = 6 \text{ MFLOPS per product}$$

c. Output Data Volume: Level-2 Volume = 5 parameter x 0.74 Gpix/day x 2B/parm = 7.4 GB/day = 0.06 Tb/day

At 5-km resolution, the Level-2 volume would be 0.06 Tb/25 = 2.4 Gb/day.

8.4.4 Overall Processing Estimates

A methodology is described for determining the CDHF processing capacity requirement for routine Level-2 and -3 processing. Using this methodology, a processing capacity estimate is made based on the scenarios described above. This estimate will be improved as more scenarios are developed and the algorithm classification scheme is broadened and refined. The MODIS data products included below have been grouped into three categories and the processing requirements for products within each category have been estimated without regard to specific product coverage and resolution. As these product requirements become known, they will be factored into the processing estimate.

Three generic types of algorithms have been identified based upon the processing in these scenarios. The three algorithm categories are as follows:

Type 1: Land product algorithm: a simple function of radiances (e.g., a ratio of radiances) and a sophisticated mapping of the geophysical parameter to a geographical grid.

Type 2: Atmospheric retrieval algorithm requiring radiative transfer calculations and iterative mathematical operations.

Type 3: Ocean biological activity algorithms: a simple function of the radiances and a relatively simple atmospheric correction.

The processing requirement per product for each type of algorithm was based on the processing requirement and the number of products derived in each of these scenarios. The processing requirements from the first two scenarios depend upon the conversion between MIPS and MFLOPS. Here the factor 3 MIPS/MFLOPS was used. This factor is the same as that used in the EosDIS Baseline Report. However, it should be noted that such factors and also the meaning of MIPS and MFLOPS are highly machine-dependent.

The product classifications have been made according to the algorithm categories and are summarized below.

a. Type 1:

Land surface composition: five products (soil type, rock type, available soil moisture, soil thermal inertia, soil particle size).

Land surface biological activity: eight products (reflected near infrared radiation-3 wavelengths, vegetative index, leaf area indices, plant and crop types, plant stress indices, and canopy state).

Earth radiation budget: five products (surface albedo, surface emissivity, net longwave flux at surface, net longwave loss from atmosphere, Bowen ratio)

Land/snow and ice cover: four products (snow and ice extent, albedo, age, emissivity).

Sea-ice cover: six products (sea-ice extent, albedo, age, emissivity, surface temperature, polynya area).

b. Type 2:

Cloud properties: seven products (cloud top pressure, cloud cover fraction, cloud albedo, cloud emissivity, cloud radiative forcing-longwave and shortwave, and precipitation).

Earth radiation budget: six products (outgoing longwave radiation, surface longwave radiation- upward and downward, latent and sensible surface heat fluxes, heat flux into the earth).

Atmospheric temperature and composition: temperature profile, humidity profile, total ozone content, carbon dioxide content, total precipitable water. In the cloud product scenario above, the fraction of current processing to produce four products was 40 percent. The fractions that correspond to the temperature and humidity profile calculation were estimated to be about 40 percent and 20 percent respectively. Based on these percentages, the processing requirements for the temperature and humidity profile products at 2 km resolution are weighted by factors of 4 and 2, respectively. Therefore, the list of five products shown here is assumed to be equivalent to nine equivalent derived products.

Aerosols: four products (one optical depth in the visible and one in the near-infrared, aerosol size distribution, aerosol height distribution).

Surface temperatures: four products (sea surface temperature, land surface temperature, plant temperature, snow and ice surface temperature).

c. Type 3:

Ocean biological activity: nine products (Gelbstoffe concentration, chlorophyll concentration, phalophatin concentration, phytoplankton pigment, species composition, chlorophyll fluorescence, suspended sediment concentration, marine humis concentration, fulvic acid concentration).

There will be 20 radiance products (water-leaving radiances) or 60 equivalent derived products. Including 20 diffuse attenuation coefficients and the 9 derived ocean products identified above yields a total of 60 + 20 + 9 = 89 equivalent derived products.

In order to compute the overall processing requirement we have summed the processing for each of the above products using the average processing capacity for each type of algorithm. The results are given in Table 5.

Table 5
Estimated Capacity to Process 24 Hours of Data

Algorithm Type	Capacity Per Product (MFLOPS)	Number of Products	Capacity (MFLOPS
1	7.5	28	210
2	94 (6)	30	2820 (180)
3	0.2	89	18
TOTALS		147	3048 (408)

The effective estimate given by this table is the capacity to process 24 hours of MODIS data to Levels-2 and -3 in 24 hours. These numbers are extremely preliminary and not meant to be taken as definitive until substantial refinement is applied. This calculation assumes that at Level-2 all products will be derived at full resolution (1 km) except for temperature and humidity profiles which are at 2 km. If the Type 2 product is derived at a 5-km resolution, then the numbers in parentheses are obtaied. In Section 8.6, these numbers are modified by factors to account for reprocessing, near real-time processing, browse processing, special processing, and processor utilization. Only the science algorithm processing has been included in these estimates. In the algorithm scenarios discussed above input/output requires a large amount of CPU time. However, it is assumed that in the EOS era the input/output CPU requirement will be based on the volume of Level-1B input data to the Level-2 processor.

8.5 CDHF INTERFACES/TRAFFIC ANALYSIS

In this section, the CDHF interfaces with the DHC, IMC, ICC, DADS, and the TMCFs are presented. Each data flow will be discussed. For each flow, the average data rates and the required peak communication rates will be given. Each data flow will be categorized as local, domestic or, foreign.

A local data flow is defined as internal to GSFC. A domestic flow will be defined as a data flow which is both economically and technologically feasible using leased fiber optic cable, i.e., communication via high speed telephone line. Foreign data flows will use satellite communication channels and/or low speed telephone lines.

The instrument data rates used in this discussion are taken directly from Appendix D of the draft MIDACS Functional Requirements Document. All data rates include the anticipated volume of calibration and engineering data. In this analysis, the assumption is made that the communications channels will operate at near 100% efficiency.

The CDHF Context diagram lists eighteen data flows. These correspond to sixteen real data flows since the paths from the DHC to the CDHF of the Level-0 Data and the near real- time-data are identical and the DQA criteria are contained in the processing algorithms. These sixteen data flows can be divided into three data rate categories: high, medium, and low.

The high rate data flows require transfer rates at or above 10 Mbps. The medium rate data flows are those for which a rate of ~ 1 Mbps is sufficient. The low rate data flows can be done with transfer rates of < 10 kbps. The low rate data flows can be handled by 9600 baud modems and hence will have no significant impact on the system design/ specification.

8.5.1 DHC

The DHC will send Level-0, near real-time, and ancillary data to the CDHF and will receive retransmission requests from the CDHF. The Level-0 and near real-time data will be discussed together since they represent the same type of data communicated along the same channel. It is assumed that the DHC to CDHF link will be local, i.e., within GSFC. Depending on the duty cycle of the DHC, the DHC to CDHF links will operate up to 24 hours.

a. Level-0 and Near Real-Time Data: The DHC will transmit Level-0 data to the CDHF. This will be done either in the normal operating mode or in the real time/near real-time mode.

The daily volume of received Level-0 data is determined by the MODIS instrument data rate. The daily average of the raw data rate will be 9.23 Mbps. For this study, it is assumed that there will be no oversampling. It is further assumed that merging and reformatting the raw instrument data will add 10% to the data volume. The daily volume of Level-0 data is given by

$$1.1 \times 9.23 \text{ Mbps} \times 86{,}400 \text{ sec/day} = 0.88 \text{ Tb/day}.$$

Near real-time data will be received and processed in addition to the normal daily data volume. It is assumed that the DHC will select the data which is to undergo near real-time processing and transmit that subset of the Level-0 data as quickly as possible.

It is assumed that the data which is selected for near real-time processing will be duplicated in the data transmitted for normal processing. This will increase the data volume to be transmitted and may increase the data processing required. However, this scenario will greatly simplify the data management and cataloging tasks.

Near real-time support will be required for a total of 15 field experiments (5 each for ocean, atmosphere, and land) each day. It is assumed that each experiment will require one 2000 x 2000 km scene. This will require approximately 5 minutes of MODIS data for each scene. With the further assumption that each scene will require up to half of the MODIS data channels, the daily data volume is

$$15/\text{day} \times 300 \text{ sec} \times (16.83 \text{ Mbps} \times 1.1) \times 0.5 = 0.042 \text{ Tb/day}$$

where 16.83 Mbps is the peak data rate and the factor of 1.1 is the data expansion to Level-0.

The CDHF will be required to receive data at the same rate as it is transmitted by the DHC which is TBD. The estimated transfer rate of priority playback data is 44 Mbps (SAR Level-1 Processing Report). Hence, the communication link will require a capacity of at least 44 Mbps to avoid delay due to transmission in near real-time processing scenarios.

A TBD fraction of the Level-0 data will be retransmitted by the DHC in response to retransmission requests from the CDHF. It is assumed that the retransmission will not be more than 1% of the total data which does not require a significant increase in the capacity of the DHC- CDHF communication channel.

It is assumed that this link will require a minimum capacity of 50 Mbps. If the DHC does not have a 100% duty cycle, the required transfer rate will be increased. Assuming a 20% duty cycle for the DHC to CDHF link, a data transfer rate of 250 Mbps is required. Existing technology can easily achieve data rates of 300 Mbps over distances of a few kilometers using dedicated fiber optic links.

b. Ancillary Data: There will be a small amount of ancillary data transmitted from the DHC to the CDHF and required for routine processing. Typical of this data would be platform attitude and solar zenith angle. It is estimated that ancillary data will have approximately 10^{-4} times the volume of the Level-0 data.

   The daily data volumes and required transfer rates are obtained by multiplying 10^{-4} by the volumes and rates given for Level-0 data. The daily data volume will be approximately 70 Mb and will require a transfer rate of 5-25 kbps.

c. Retransmission Request: CDHF will send retransmission requests to the DHC. The CDHF will test the incoming data for completeness and accuracy. When the received data are unacceptable, a retransmission request will be sent to the DHC. Each retransmission request is expected to be ~200 bytes with the expected number of requests TBD. This is a low-rate, local communication channel.

8.5.2 IMC

The CDHF will receive production status inquiries from the IMC and send production reports to the IMC. It is anticipated that there will be one inquiry of size 4 kb and one report of 40 kb per day. The required storage and transfer rates are negligible and the link will be local.

8.5.3 ICC

The ICC will send mission planning information and selected science data requests to the CDHF and will receive selected science data from the CDHF.

a. Mission Planning Information: The ICC will send the CDHF mission planning information. This information will be used to determine whether all the instrument data has been received from the DHC. The planning information will include start and stop times for each channel and a calibration schedule. While the exact volume and content of the planning information is TBD, it is assumed that the volume will not be larger than 250 kb/day.

b. Selected Science Data Requests: The CDHF will receive requests for selected science data from the ICC. If there are 10 requests per day at 2 kb per request, the daily data volume is only 20 kb.

c. Selected Science Data: The CDHF will send selected science data to the ICC for instrument monitoring. It is estimated that the ICC will require four channels of data at the instrument data rate. This corresponds to roughly 4% of 17 Mbps or 0.7 Mbps for this data path. This is a moderate data flow which are assumed to be local. It is assumed that this interface will operate at 1 Mbps for 367 minutes per day to transfer 3.5 GB of data per day.

8.5.4 DADS

The CDHF will send MODIS data products to the DADS for archiving. The CDHF will issue archive data requests to the DADS for either archived MODIS data products or archived ancillary data.

a. MODIS Data Products to Archive: The CDHF will send Level-1A through Level-4 data to the DADS for archiving. We assume that approximately 80% of the volume will be due to Level-1A, -1B, and -2 data products. It will be necessary to process 24 hours of data in 8 hours which implies that all levels of data products should be transmitted to the DADS in 8 hours. The baseline report estimates that the Level-0 data will be 1.1 times the raw data volume. The Level-1A data will have a volume of 1.1 times the Level-0 data, and the -1B data will be of the same size as the -1A data. The Level-2 will have twice the volume of Level-1A. The Functional Requirements Document states the MODIS data rate as 9.23 Mbps. After Level-0 processing, this rate increases to 10.2 Mbps. The required rates for transferring data products to the DADS are:

   Level-1A: 10.2 Mbps x 24/8 x 1.1 = 33 Mbps,
   Level-1B: 33 Mbps,
   Level-2: 2 x 33 Mbps = 66 Mbps.

   These three products will require a total transfer rate of 132 Mbps with a 100% duty cycle over eight hours.

   There will be much smaller volumes of Level-3 and -4 data also sent to the CDHF. The current best estimate is that Level-3 will be 15% of Level-2 data. This would require a data rate of 10 Mbps. The volume of Level-4 data is TBD, but we assume that it is small enough to have no significant effect on the required data rate.

   The total of Level-1, -2, and -3 data products from the CDHF to the DADS will require a data rate of 142 Mbps with a 100% duty cycle. The assumption of a more reasonable duty cycle of 75- 50% for the CDHF to DADS link will increase the required capacity to 190-285 Mbps. The daily data volume is obtained by multiplying 142 Mbps by 8 hours to obtain 4.1 Tb/day. The data volume to be archived by the DADS is more than 4 times the volume of the Level-0 data.

b. Archive Data Requests: The CDHF will send requests for archived data to the DADS. The number of archive data requests is TBD. However, at 2 kb per request, the required data rate is small, i.e., < 100 kb per day.

c. Archived MODIS Data Products: It is anticipated that all data will be reprocessed twice. In most cases it will be necessary to retrieve this data from the DADS. Depending on whether only Level-1A or all products are retrieved, the volume of data to be returned will be 1.0 to 4. 1b for a 24 hour period.

   It is assumed that 8 hours will be available to retrieve the data, i.e., the data set to be reprocessed is returned during the 8 hours that another data set is being processed. The DADS to CDHF reprocessing connection will require a data rate of 38-142 Mbps with a 100% duty cycle.

   It is possible/likely that standard or special Level-3 or -4 processing will require that some data be recovered from the DADS for routine processing. If this must be done, it will increase the required communication capability. The provision must also be made for a less than 100% duty cycle. With the assumption of recovery of additional data and a 50% duty cycle, it is estimated that the DADS to CDHF link will, at times, be required to handle a data rate of 300 Mbps.

d. Archived Ancillary Data: The CDHF will receive archived ancillary data from the DADS. This will include correlative data and other Eos instrument data and science products. The ways in which these data will be used and the anticipated volume are TBD. For the purpose of this study, this data flow has been estimated to be approximately 10% of the MODIS raw data which will produce a daily data volume of 88 Gb. This will require a 1 Mbps data rate with a duty cycle of 100%.

8.5.5 TMCF

The CDHF will send selected data products and DQA reports to the TMCFs and receive from the TMCFs processing requests and processing algorithms. The processing algorithms will contain the DQA criteria. There will be multiple TMCFs which may be local, domestic, or foreign.

a. Selected Science Products: The CDHF will send selected data products to the TMCFs for algorithm development and various validation studies. It is estimated that the volume will be approximately 0.5% of the MODIS data. This will require a transfer rate of < 100 kbps and produce a daily data volume of 24 Gb if Levels-0 through -2 are transmitted.

   The scenes produced in near real-time processing done on the CDHF are considered selected science products. These scenes must be transmitted to a TMCF that may be located at a remote field experiment. With the assumption that only one product will be transmitted, each processed scene will be approximately 64 Mb of data (2000 x 2000 pixels x 2 bytes/pixel). If a scene is transmitted to the field at a rate of 10 kbps, the transmission can be done 107 minutes. A larger data rate may be needed to achieve the timelines requirement for near real-time processing in support of a field experiment. If 25 products are transmitted, the data volume becomes 1600 Mb which would require 27 minutes to transfer a scene at an effective data rate of 1 Mbps.

b. DQA Reports: The CDHF will send DQA reports to the TMCF. The size and number of DQA reports is TBD. At 40 kb per report and 115 reports/day, the total volume is 0.6 Mbps. This will have negligible impact on the system design.

c. TMCF Processing Requests: The CDHF will receive processing requests from the TMCFs. It is anticipated that these requests will use approximately 2 kb. The number of processing requests to be expected is 100 per day. However, these requests will have no significant effect on the required transfer rates.

d. Processing Algorithms: The CDHF will receive processing algorithms and algorithm updates from the TMCFs. The algorithms will contain the DQA criteria. An algorithm may be relatively large but algorithm uploading will occur infrequently. A 0.16 MB (1288 kb) algorithm transfer could be done 3 times per day with a total transfer of 3.9 Mb per day.

8.5.6 Traffic Summary

There are three high rate data flows to or from the CDHF. The DHC to the CDHF, the DADS to the CDHF, and the CDHF to the DADS. Each of these channels is local and will require data rates as large as 300 Mbps.

There are four data flows which will require data transfer rates of 10 kb to 1 Mbps: ancillary data from the DHC to the CDHF, selected science data from the CDHF to the ICC, archived ancillary data from the DADS to the CDHF, and selected data products from the CDHF to the TMCF. These channels are all local with the exception of the communication to the TMCFs which may be located in remote areas.

The remainder of the data flows illustrated in the CDHF context diagram can all be handled at data rates of < 10 kbps. This will have minimal impact on the communication requirements of the CDHF.

8.6 SYSTEM SPECIFICATION

8.6.1 Processing Capacity

The capacity to process 24 hours of MODIS data to Level-2, -3 in 24 hours is estimated in the last Section 8.4.4 to be 3048 MFLOPS. This is be multiplied by a factor that accounts for 2 reprocessings, near real- time processing, special processing (10%), browse processing (7%), maintenance (10%), and a processor-utilization term (70%).

$$\text{Overall capacity} = 1/0.7 \times (1 + 2 + 0.38 + 0.10 + 0.07 + 0.10) \times 3048 \text{ MFLOPS} = 16 \text{ GFLOPS}$$

The factor of 0.38 for near real-time processing is derived from the ratio of near real-time Level-0 data to the routine Level-0 data (Section 8.5) and a 3 hour processing time.

$$(0.042 \text{ Tb/day})/(0.88 \text{ Tb/day}) \times 24 \text{ hours}/3 \text{ hours} = 0.38$$

A reduced spatial resolution of (5 km) for Type 2 algorithm products (see Section 1.3.17 and Section 8.4.4) yields a factor of 7.5 (3048 MFLOPS/408 MFLOPS) reduction in required overall processor capacity. In this case,

$$\text{Overall Capacity} = 2 \text{ GFLOPS}$$

NOTE: These numbers are extremely preliminary and uncertain at this time and will change as more scenarios are developed and the algorithm classification scheme is refined.

8.6.2 Data Storage Requirements

Based on the traffic analysis in Section 8.5, daily data inflow of Level-0 data is 0.88 Tb/day.

Accounting for 2 reprocessings, the daily data outflow is approximately the total sent to DADS which is 4.1 Tb/day x 3 = 12.3 Tb/day.

Assuming that the CDHF will need to be able to store approximately 2 days of this input plus output data, the Storage volume = (0.88 Tb + 12.3 Tb) x 2 = 26 Tb.

9. ELEMENT SPECIFICATION FOR THE DATA ARCHIVE AND DISTRIBUTION SYSTEM (DADS)

The DADS is the repository for Levels 1-4 MODIS datasets. Other MODIS data stored in the DADS includes catalog and metadata, ancillary and specialized data, and browse data. MODIS data are available from the DADS for retrieval and analysis by the MODIS user community.

9.1 DADS FUNCTIONAL REQUIREMENTS

DADS functional requirements address four areas as follows:

9.1.1 Receipt of MODIS Data

Command histories are received from the ICC. Specialized data products and algorithms, correlative data, processing algorithms, and verification/validation study results are received from the TMCF. Standard and reprocessed MODIS data products are received from the CDHF, including both browse and metadata. In response to user requests, permanent archive data is received from the NSSDC or other long-term archives. Ancillary data is received from other EosDIS and also non-EosDIS sources.

9.1.2 Manage MODIS Data

Header and MODIS data stored in the DADS includes Levels 1-4 data, investigator-generated products, command histories, browse files of reduced resolution data, calibration procedures, summary archives, attribute files, bibliographies of published and unpublished material, calibration sources, and the documentation for the current CDHF and ICC processing software.

In response to user status inquiries, IMC requests, and periodic MODIS reporting requirements, DADS status reports summarizing retrieval and internal DADS activities are produced. Users receive the in-process statuses for queries and requests still being processed. DADS reports for the IMC include job accounting data for catalog use, archival loading, data volumes handled, and any backlogs.

Catalog and directory requirements include information on MODIS data location, ownership, data processing levels, project, platform, geographic location, start/stop times, catalog access, and periodic IMC catalog update. The Eos data catalog contains data pertaining to project, platform, instrument, data processing levels, versions, time, and geographic location.

9.1.3 Process User Requests

As a function of a specific dataset request or one or more dataset or product attributes specified in the user's query to the IMC, the DADS retrieves the necessary MODIS products. These products reside in the DADS, other EosDIS DADS, or one of the three permanent archives.

9.1.4 Distribute Data

Products retrieved for each request are copied to the appropriate media and shipped to the user. Specific quantities of MODIS products can also be electronically sent to the user. Browse and catalog data is distributed in a manner similar to regular MODIS products, being made available for interactive access in the IMC.

9.2 DADS PERFORMANCE REQUIREMENTS

DADS performance requirements address four areas as follows:

9.2.1 Receive Data

Based on the information provided by the dataset's producing organization (usually the CDHF), the DADS catalog will be updated on a daily basis. The catalog entry associated with each archive dataset will be inserted into the catalog after the dataset is inserted into the DADS.

9.2.2 Manage Data

The DADS archive will maintain data on media that provides lifetimes consistent with the MODIS lifetime, rapid access, and economical storage. The DADS will be sized to support a user community possibly ranging from 1,000 to 10,000. From 50 to 200 users are expected to be active at any time, each ordering from 5 to 10 tapes per month, with 1 to 10 users ordering large data volumes. The average number of simultaneous users will be 100. The DADS will be capable of storing data at the volume of 0.6 TBytes/day. The average and maximum total archive data retrieval volume for electronic distribution is TBD GBytes. An expected 10^6 bytes/day of catalog entries are provided to the IMC by the DADS.

9.2.3 Process User Requests

The estimated maximum daily number of retrieval orders for electronic or off-line distribution is 250/day. Up to 100 simultaneous daily interactive catalog/browse/ordering system users will be supported.

The IMC will provide a minimum 9,600 baud dial-up capability. The expected maximum number of queries/day is 1000. The DADS will be accessed from hardwired or remote terminals via menus or natural language interface, supplemented by a free-form command language. The interactive system will provide an average 15 second acknowledgement/response time for search commands.

9.2.4 Distribute Data

The DADS will be capable of satisfying the need for data by multiple users, each user requiring a different time line and path. The retrieval and transfer of catalog information is accomplished with a bit error rate of 10^{-12}. Browse files will be visually searchable via attributes including day, position, time, channel, and parameter. An hardcopy image browse catalog will be available.

Within 40 seconds of an order's receipt, the DADS will locate and/or retrieve each order's first dataset and initiate transmission to the user or shipping media. The DADS will locate and/or retrieve the first dataset of orders for off-line data within 30 minutes for 90 percent of these orders. The DADS will respond to orders requesting permanently (externally) archived data within three working days.

9.3 DADS INTERFACES

DADS interfaces are presented in terms of interfaces with systems or EosDIS entities outside of the DADS, and interfaces among the four DADS functions.

9.3.1 External Interfaces

The context diagram in Figure 14 presents the DADS interfaces with other EosDIS and non-EosDIS processing environments. Each interface's contents are labeled in terms of data flows in one or more directions. Data exchanges will take place electronically. The balance of this subsection presents each external interface's characteristics and data flow directions, with data flow estimates also provided.

9.3.1.1 *TMCF Interface*

This interface is expected to have a moderate transfer rate, an occasional frequency of use, routine timeliness requirements, and local and remote path lengths. Four data flows are from the TMCF to the DADS. However, Correlative-Data flows in both directions, and Archive-Data-Products flow from the DADS to the TMCF. These data flows are estimated at 16 MBytes/month.

9.3.1.2 *CDHF Interface*

This interface is expected to have a high transfer rate, a regular (daily) frequency of use, both near-realtime and routine timeliness requirements, and a local path length. The CDHF and DADS may be co-located. Archive-Data-Requests flow from the CDHF to the DADS, Archive-Ancillary-Data flows from the DADS to the CDHF, and MODIS-Data-Products flow in both directions. The CDHF is expected to send approximately $0.6 * 10^{11}$ Bytes/day in MODIS datasets. As two reprocessings/dataset are anticipated, the full dataflow rate is estimated as $1.8 * 10^{12}$ Bytes/day.

9.3.1.3 *Other EosDIS DADS Interface*

This interface is expected to have a high transfer rate, an occasional frequency of use, routine timeliness requirements, and both local and remote path lengths. Three data flows are two-way, with Ancillary-Data flowing from other EosDIS DADS to the MODIS DADS, and Archive-Ancillary-Data flowing from the MODIS DADS to other EosDIS DADS.

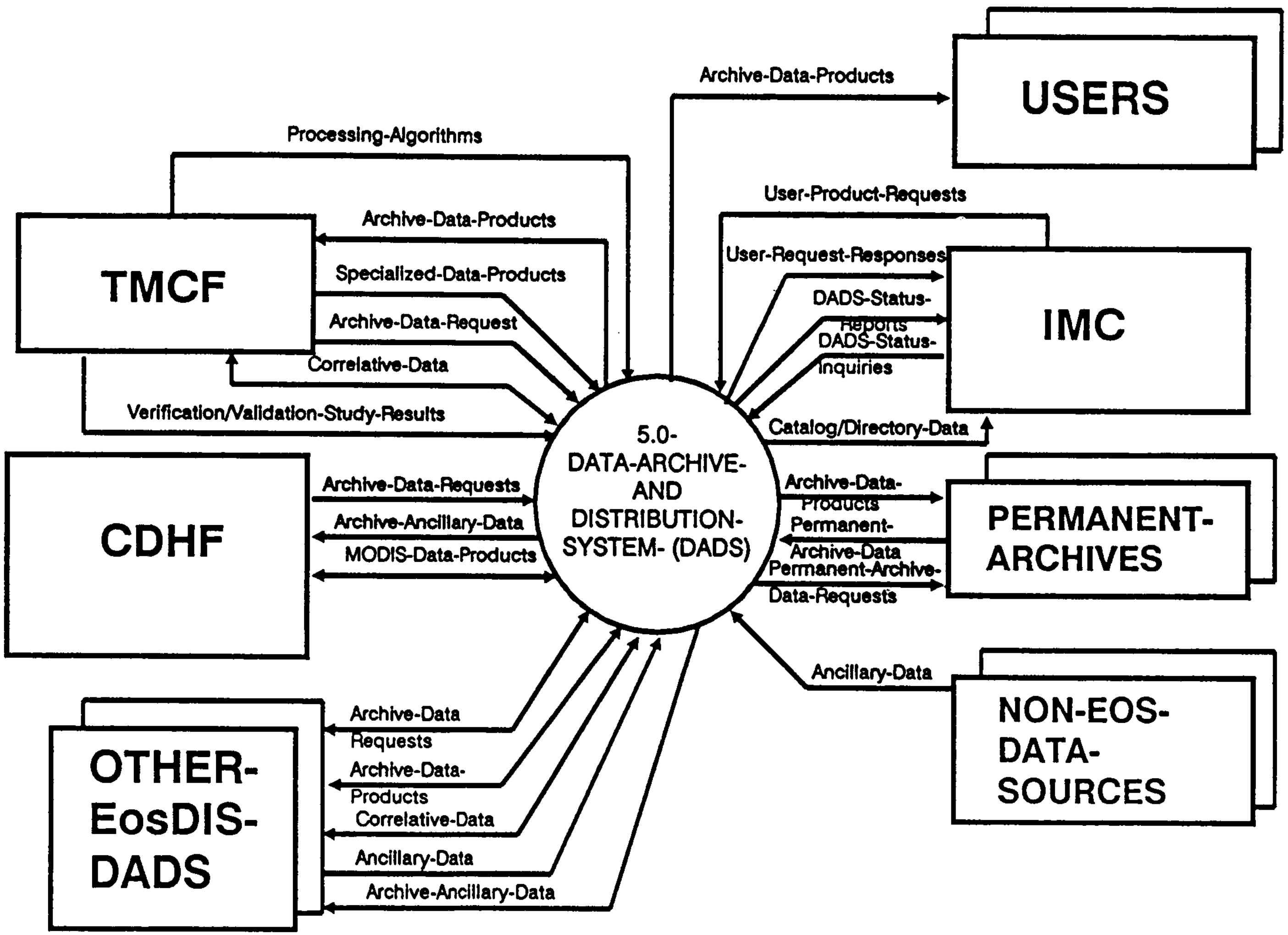


Figure 14. DADS Context Diagram

9.3.1.4 *User Interface*

This interface is expected to have a moderate transfer rate, an occasional frequency of use (per user), routine timeliness requirements, and both local and remote path lengths. Archive-Data-Products flow from the DADS to the user, either in low levels of electronically transmitted data or on media shipped to the user. With the exception of Science Team members using their TMCF facilities, user access of the DADS will be routed through the IMC. With the exceptions of standing orders and TMCF requests for specific datasets, user queries and other data requests will be first processed by the IMC, with DADS processing following as a result of IMC action. The estimated dataflow to the user on off-line media is estimated as 4.7 GBytes/day.

9.3.1.5 *IMC Interface*

This interface is expected to have a moderate transfer rate, with regular frequency of use, routine timeliness requirements, and a local path length. User-Product-Requests and DADS-Status-Inquiries flow from the IMC to the DADS. User-Request-Responses and DADS-Status-Reports flow from the DADS to the IMC. The activity boundary between the DADS and the IMC is as follows:

Activity/Product	IMC	DADS
User Access	Single contact point	Ships data products
Queries	Accepts/parses	Executes
Standing Orders	N/A	Schedules/executes
On-line Catalog	Stores from DADS	Generates
Hardcopy Catalogs	Provides to user	N/A

The IMC will be in effect a manager and filter for user access to the MODIS data in the DADS. Estimated data flows for electronically transmitted datasets, user requests, and status reports is 50 MBytes/day.

9.3.1.6 *Permanent Archive Interface*

This interface is expected to have a high transfer rate, with occasional frequency of use, routine timeliness requirements, and both local and remote path lengths. Archive-Data-Products and Permanent-Archive-Data-Requests flow from the DADS to the Permanent Archive. Permanent-Archive-Data flows from the Permanent Archive to the DADS.

9.3.1.7 *Non-EosDIS Data Source Interface*

This interface is expected to have a high transfer rate, with occasional frequency of use, routine timeliness requirements, and both local and remote path lengths. Ancillary-Data flows from these non-EosDIS sources to the DADS.

9.3.2 Internal Interfaces

The data flow diagram in Figure 15 presents the interfaces among the four DADS functions. The contents of each interface are labeled in terms of data flows in one or more directions. Data exchanges take place electronically with non-DADS elements, and within the DADS processor suite for DADS elements. The balance of this subsection presents each internal interface's characteristics and data flow directions.

C-2

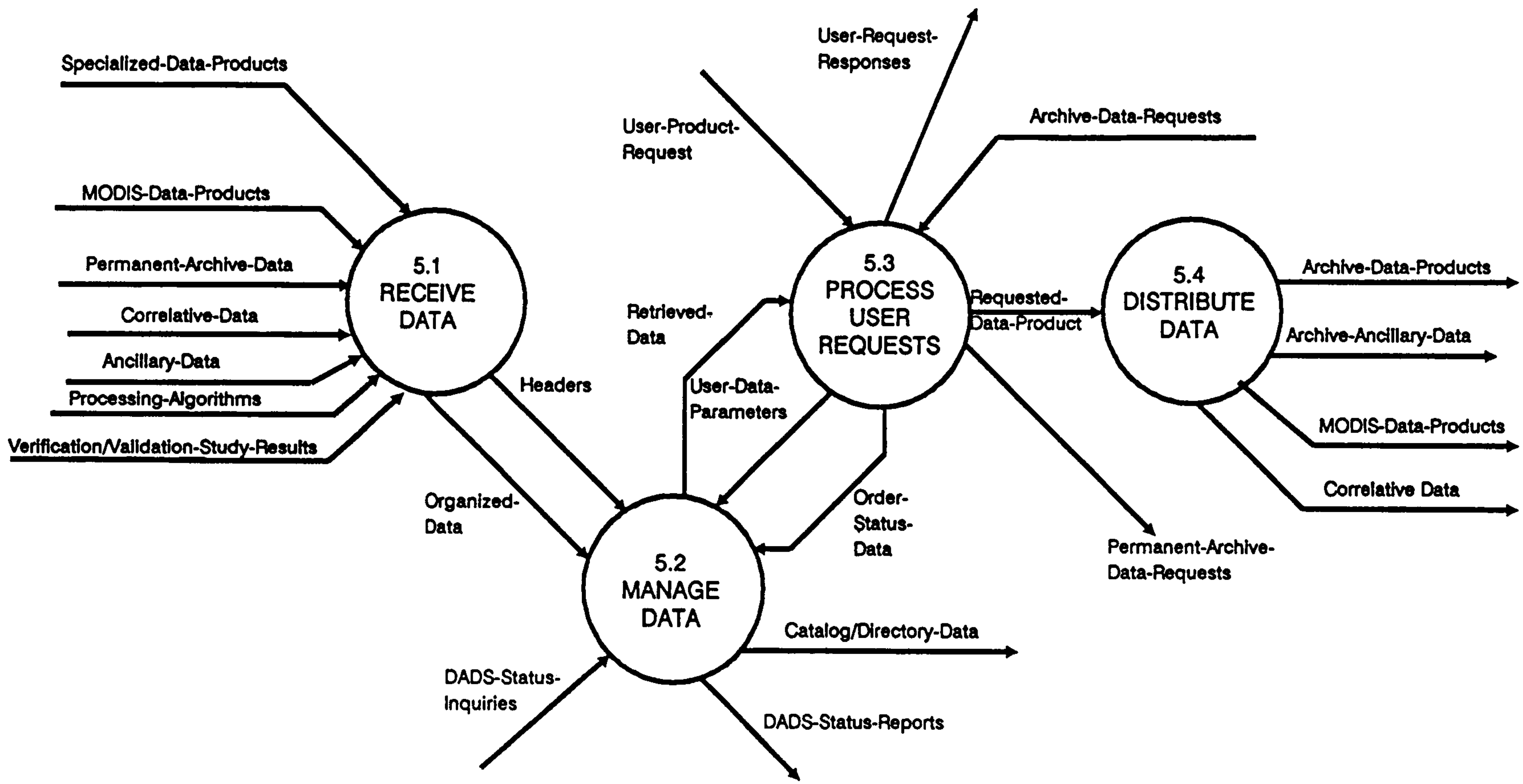


Figure 15. DADS Functional Data Flow

9.3.2.1 *Receive Data*

Data will be ingested and subjected to TBD acceptance checking. Incoming headers will be separated and stored for subsequent generation of catalog and other descriptive data. The remaining data will be organized for subsequent storage within the DADS as datasets or subsequent placement on media for shipment to a user.

9.3.2.2 *Manage Data*

Data is stored as completed data sets. Header data will be processed into catalog or directory data and sent to the IMC. These processes generate activity data that is combined with order status data to generate daily DADS status reports. These reports are sent to the IMC.

9.3.2.3 *Process User Requests*

Each IMC-generated user request results in obtaining data from either the DADS archives or the permanent archives. The status of each query-based or standing order will be kept current. When all the requested datasets are retrieved, they are routed for electronic or off-line media distribution.

9.3.2.4 *Distribute Data*

Retrieved and requested data products will be sent to the requesting users on appropriate media. The media selected is based on the data types and data quantities, with datasets destined for other data centers being transmitted electronically. Off-line media such as optical or magnetic tapes will be shipped to satisfy user queries or standing orders.

9.4 DADS ARCHITECTURAL DESIGN

Figure 16 presents a representative DADS computer system architecture. Foreground processing is reserved for communications support of data sources and system users. Background processing is reserved for file processing. This includes updating the archived MODIS datasets and preparing retrieved datasets for copying to computer-readable media. This is indicated by the dotted line.

The sizing calculations that follow are based on estimates of data quantities and data flow rates. They are the basis for the overall levels of processor power in the estimated DADS configuration.

9.4.1 Receive Data

The DADS processes and makes available to users, data from the CDHF, TMCF, and other sources within 8 hours of receipt. The calculations for the data volumes and levels that drive computer processing requirements are as follows:

9.4.1.1 *MODIS Datasets*

The quantities of data received in the DADS are calculated as follows:

Level 0 = 1.1 * Raw Rate of 10^{12} bits/day = 10^7 bits/sec
(Note: Level 0 is not sent to the DADS; Level-1A and beyond unpacked to 2 bytes/observation.)

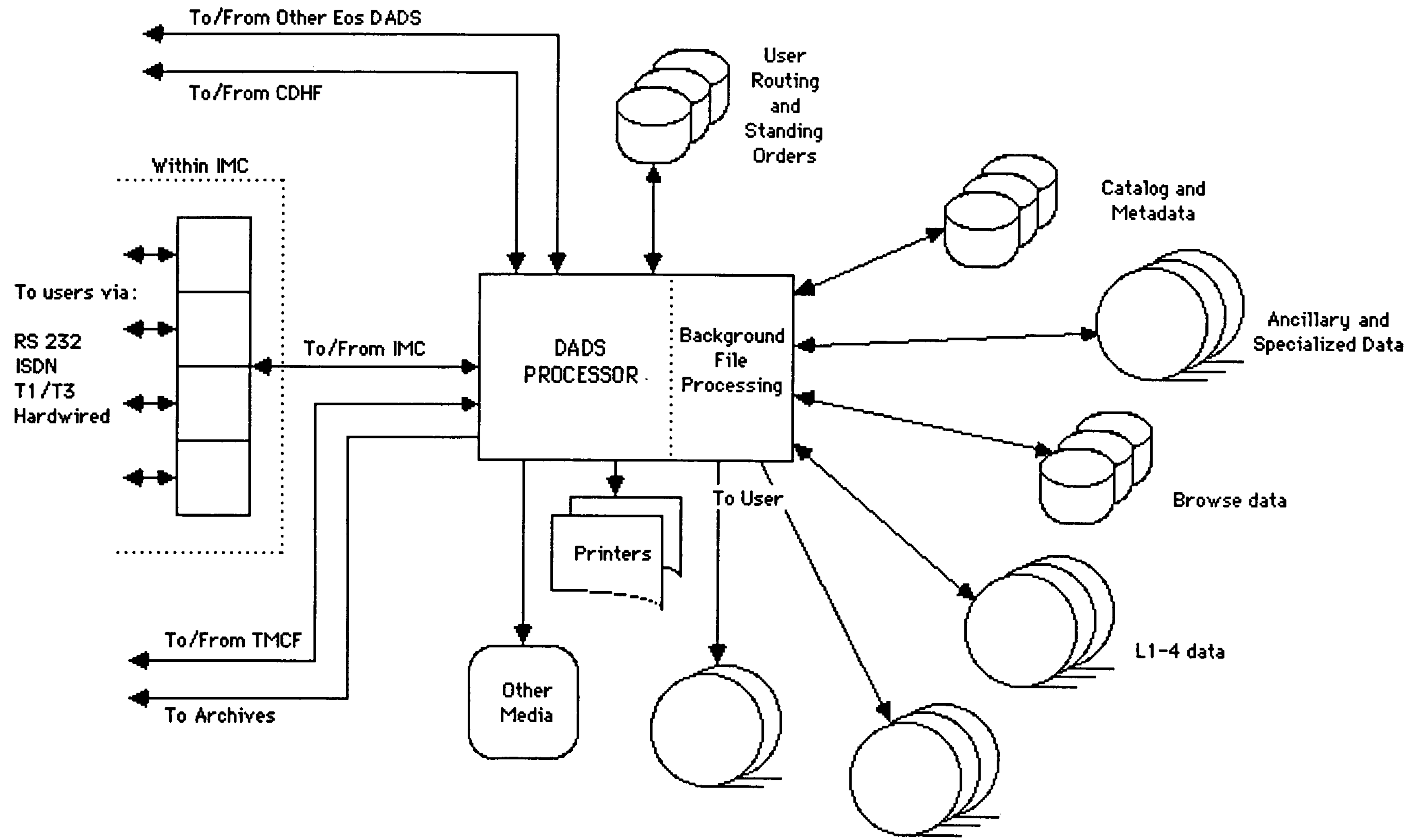


Figure 16. Representative Configuration for the DADS

Level 1A = 1.1 * Level 0 * (16/12) = $1.5 * 10^7$ bits/sec
Level 1B = 1.0 * Level 1A = $1.5 * 10^7$ bits/sec
Level 2 (initially) = Level 1B = $1.5 * 10^7$ bits/sec
Level 3 = 0.15 * Level 2 = $.23 * 10^7$ bits/sec
Level 4 (initially) = 0.5 * Level 3 = $.12 * 10^7$ bits/sec

This is a total of approximately $5 * 10^7$ bits/second. As there are approximately 10^5 seconds/day, this is a total of approximately $5 * 10^{12}$ bits/day. As there are 8 bits/byte, this is a total of approximately $0.6 * 10^{11}$ Bytes/day.

Total for MODIS datasets = $0.6 * 10^{11}$ Bytes/day.

9.4.1.2 *Browse Data*

There are two data sources, MODIS-N and MODIS-T, and two types of browse data: scan coordinates and Earth coordinates. Resolution for along-track/across-track browse data is 4 km. The platform velocity is expected to be 6.5 km/sec. There are 1294 pixels/swath. There are approximately 10^5 seconds/day, There will be two channels displayed, visible and IR. There is 1 byte of data. Therefore,

$$2 * (1/4)^2 * 6.5 * 1294 * 10^5 * 2 * 1 = 2 * 10^8 \text{ Bytes/day}$$

Resolution for latitude/longitude data is 20 km. There are 8 images to cover the world. Each image has 512 * 512 pixels. One byte of data is used. There are also assumed to be eight regions of interest as well. There are 4 channels displayed, and we assume that half of the (100) derived geophysical parameters will be displayed as well. There are 2 instruments. Therefore,

$$(8 + 100 / 2) * 512 * 512 * 1 * 2 * 4 * 2 = 2.4 * 10^8 \text{ Bytes/day}$$

$$(2 * 10^8) + (2.4 * 10^8) = 4.4 * 10^8 \text{ Bytes/day}$$

Total (approximate) for Browse data = $4 * 10^8$ Bytes/day.

9.4.1.3 *Catalog Data*

The number of products for Levels 1A, 1B, 2, 3, and 4 are estimated at 1, 1, 100, 100, and 50, respectively. One page of catalog data/product is also assumed, with a page estimated at 50 lines each of 80 characters. Therefore,

$$252 \text{ products/day} * 4000 \text{ Bytes/product} = 1 * 10^6 \text{ Bytes/day.}$$

Total for Catalog data = 10^6 Bytes/day.

9.4.1.4 *Metadata*

There are an expected 30 metadata parameters to characterize the data in terms of complete scans (0.94 seconds) of MODIS data. For MODIS-N this is 8 km of data, and for MODIS-T this is 1/8 scan or 64/8 km.

$$(86400 \text{ seconds/day}) / (0.94 \text{ seconds/scan}) = 9.2 * 10^4 \text{ scans/day}$$

There are 30 metadata parameters and 252 MODIS products. Therefore,

30 parameters/product scan • 252 products • $9.2 * 10^4$ scans/day = $7 * 10^8$ Bytes.

Total for Metadata is $7 * 10^8$ Bytes/day.

9.4.1.5 *Estimated Sizings for Receive Data*

Every eight hours the DADS will receive quantities of data estimated as follows:

Product Name	Bytes per Day
MODIS Data Products	$6 * 10^{11}$
Browse Data	$4 * 10^8$
Catalog Data	$1 * 10^6$
Metadata	$7 * 10^8$
Total (rounded)	$6 * 10^{11}$

This data will arrive over an eight hour period. This is a rate of approximately $2 * 10^7$ Bytes per second. The amount of DADS processing is minimal as it involves moving data from one buffer to another, from which it is subsequently written to off-line media. The DADS processor provides mainly I/O and media device management. These are not computation-intense activities, and a level of one instruction per 10 bytes of data is projected.

$(2 * 10^7$ Bytes/second$) * 0.1 = 2 * 10^6$ instructions/second

The estimated processing power requirement for the Receive Data function is 2.0 MIPS.

9.4.2 Manage Data

As calculated in the previous section, approximately $6 * 10^{11}$ bytes of data are to be transferred to DADS electronic media for storage. The processing power requirement is estimated to be the same as Receive Data plus 10 percent to manage I/O processing. There are activity reports and analyses the DADS provides to the IMC. They are not expected to require more than a low level of computer power, and are expected to remain in the noise level in terms of processor resources. Part of the reported information would normally be generated by the operation system. The estimated processing power requirement for the Manage Data function is 2.2 MIPS.

9.4.3 Process User Requests

One hundred user requests from three categories of users are expected to be simultaneously satisfied. A "heavy global user" is one who needs: (1) 1 km resolution global coverage, (2) 10 days of data, and (3) 10 MODIS radiances and/or parameters. A "heavy regional user" is one who needs: (1) 1 km resolution coverage over 1/20 of the globe, (2) 365 days of data, and (3) 10 MODIS radiances and/or parameters. The typical data request is expected to be for 10^{12} bits of data. A "moderate" MODIS data user is one who orders 10 percent of the data a heavy user would order. A "light" user is one who orders one percent of the data a heavy user would order, such as only one month of regional data or one day of global data. The user community is assumed to be evenly divided into light, moderate, and heavy users.

Quantities of data being retrieved are categorized and calculated as follows:

User Type	Number of Users	Number of Bits	Number of MBytes	Mbytes Processed
Light	33	10^{10}	1.25	41.25
Medium	33	10^{11}	12.5	412.5
Heavy	34	10^{12}	125	4250
	100			4703.75

A total of 4704 MBytes are being processed and transferred at any given time. The locations of these data on specific DADS media are determined, and the data are being transferred from DADS media to shipment media. With query processing being provided by the IMC, very little additional processing is expected to be necessary. One instruction is assumed to be required for each 100 bytes, or approximately 47 MIPS of computer power.

User queries will be in different stages of completion. As MODIS Level 1-4 data will be stored off-line, query processing will be dependent on tape locating and mounting/dismounting. This reduces the amount of data being transferred at any given time from DADS storage media to the media being shipped to the user. This means that while 100 users will be expected to active at a given time, a large percentage of their interactive jobs will be in wait states, pending media and/or media unit availability. As a result, 25 users' interactive jobs are expected to be actually effecting data transfer or other processing at a given time. As this is 25 percent of the expected number of users,

$$47 \text{ MIPS} * .25 = 11.75 \text{ MIPS}$$

is a reasonable approximation of the required processing power. This processor power level is expected to be more than adequate for also providing user-requested status on outstanding requests, and responding to other non-computational user requests. Total computer power for Process User Requests = 12 MIPS.

9.4.4 Generate Data Product

From the previous section 4704 MBytes of data are moved to the designated media for shipment to the user. This is a straightforward I/O operation, and is expected to require computer power similar to what is needed for Manage Data. This is estimated to be 2.2 MIPS. Total computer power for Generate Data Product = 2.2 MIPS.

9.4.5 Projected DADS Computer Processing Power Requirements

DADS Function	Estimated Processor Requirements (MIPS)
Receive Data	2
Manage Data	2.2
Process User Requests	12
Generate Data Product	2.2
Total	18.4

These four activities are not expected to be occurring during the same periods of time. For example, Receive Data and Manage Data are off-hours activities. This means the additional computer power from other functions is available as needed, and the computer power associated with these two functions is available for other functions, as needed. In addition to the data flows defined above, two reprocessings of MODIS datasets are anticipated. As eight hours are needed for the initial MODIS dataset flows, this same time period would be expected to be needed for the other two dataset flows. This means three eight-hour flows per day, or continuous dataset flows.

For this reason the DADS MIPS requirements will be treated as additive, meaning they must be available fulltime for each individual DADS activity. Given the need to process user requests as expeditiously as possible, these massive data flows must remain transparent to these users. This necessitates sizing for the peak load, and ensuring the peak load does not overload the DADS processor. Therefore,

18.4 MIPS / .7 = 26.38 MIPS.

The DADS computer processing power requirement is 27 MIPS.

10. SUMMARY OF MIDACS DESIGN PARAMETERS

10.1 INTRODUCTION

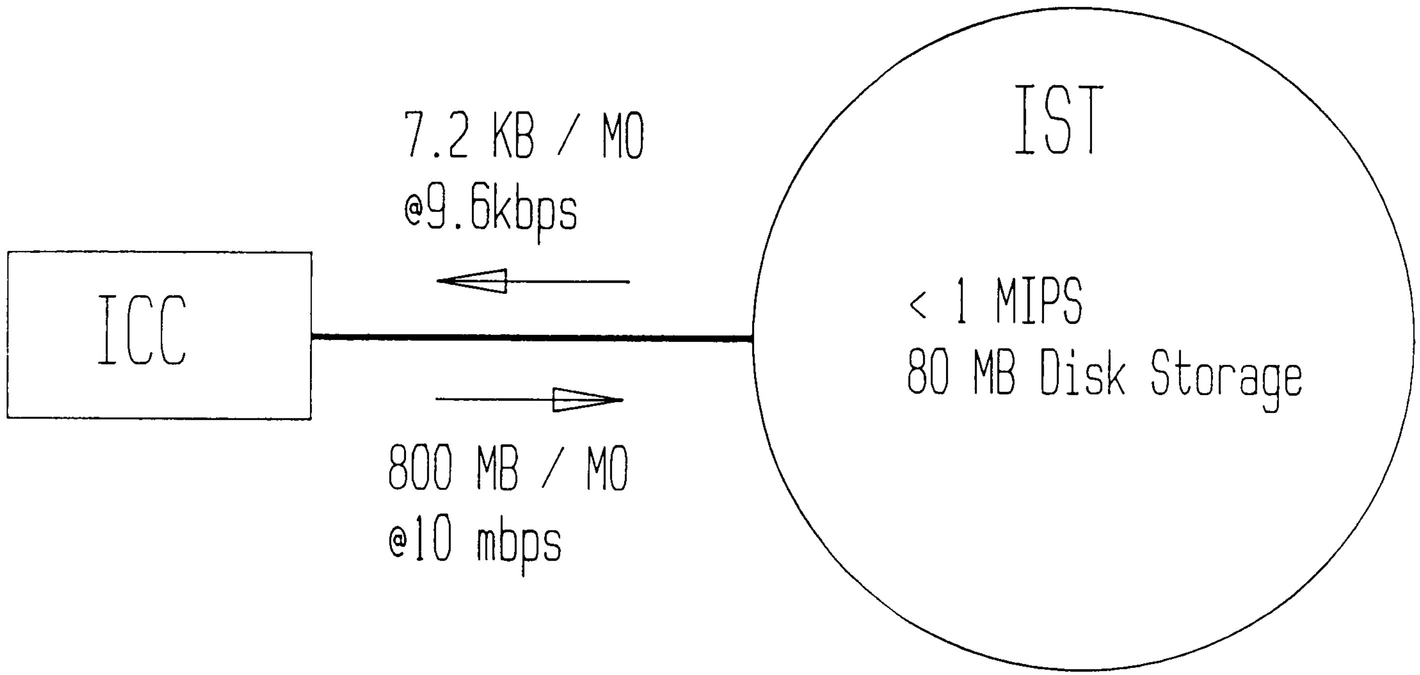
This section of the document presents a quick summary of the MIDACS design parameters that were derived in the preceeding sections. Results are presented individually for each of the MIDACS elements, except that low data rate requirements are summarized in a separate diagram depicting the overall system configuration for low rate communications. This was done since each MIDACS element has at least some data links that can be met by low rate lines, and as a practical matter, all low rate requirements will likely be met by a single grade of low rate service.

10.2 MIDACS ELEMENT FIGURES

The five figures that follow (17 through 21) present summarized information for each of the MIDACS elements, i.e. one figure each is presented for the IST, the ICC, the TMCF, the CDHF, and the DADS. The information in the central circle summarizes processing and data storage requirements for that node. High rate data links with other MIDACS elements or elements external to the MIDACS are shown as channels with arrows attached to indicate directions of data flow. Each arrow is annotated with a number that indicates total volume of data flowing in the indicated direction and an interval of time (day or month) during which that total flow is expected. The peak required flow rate is indicated below that number. Data volumes are indicated in Megabytes (MB), flow rates are indicated in Megabits per second (Mbps).

10.3 LOW-RATE DATA FLOWS

Figure 22 shows the overall architecture for low data rate communications. Elements internal to the MIDACS are shown above the bus; external elements are shown below the bus. Representative types of data exchanged by each element are shown in the boxes denoting the individual MIDACS elements. We consider the term low data rate to be defined as communications over standard telephone lines at up to 9,600 bps. At this rate, up to 3 GB of information could be transferred in one month. Practically, however, the data volume would be much smaller. A more realistic twenty full pages of text (@ 4,000 bytes) per day would yield a monthly volume of 3 MB.



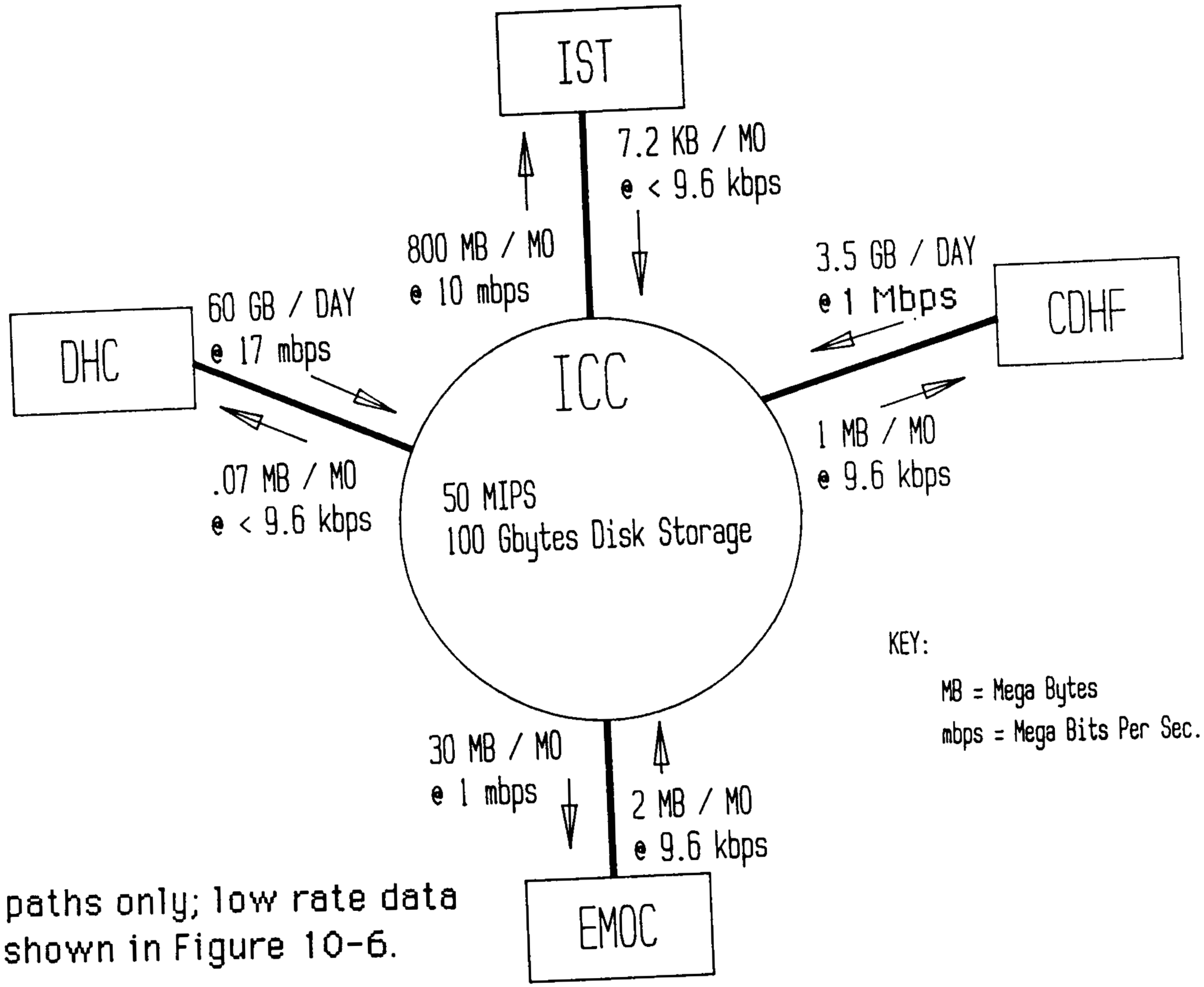
Note: High-rate paths only; low rate data paths are shown in Figure 10-6.

KEY:

MB = Mega Bytes

mbps = Mega Bits Per Sec.

Figure 17. IST Performance Summary



Note: High-rate paths only; low rate data paths are shown in Figure 10-6.

Figure 18. ICC Performance Summary

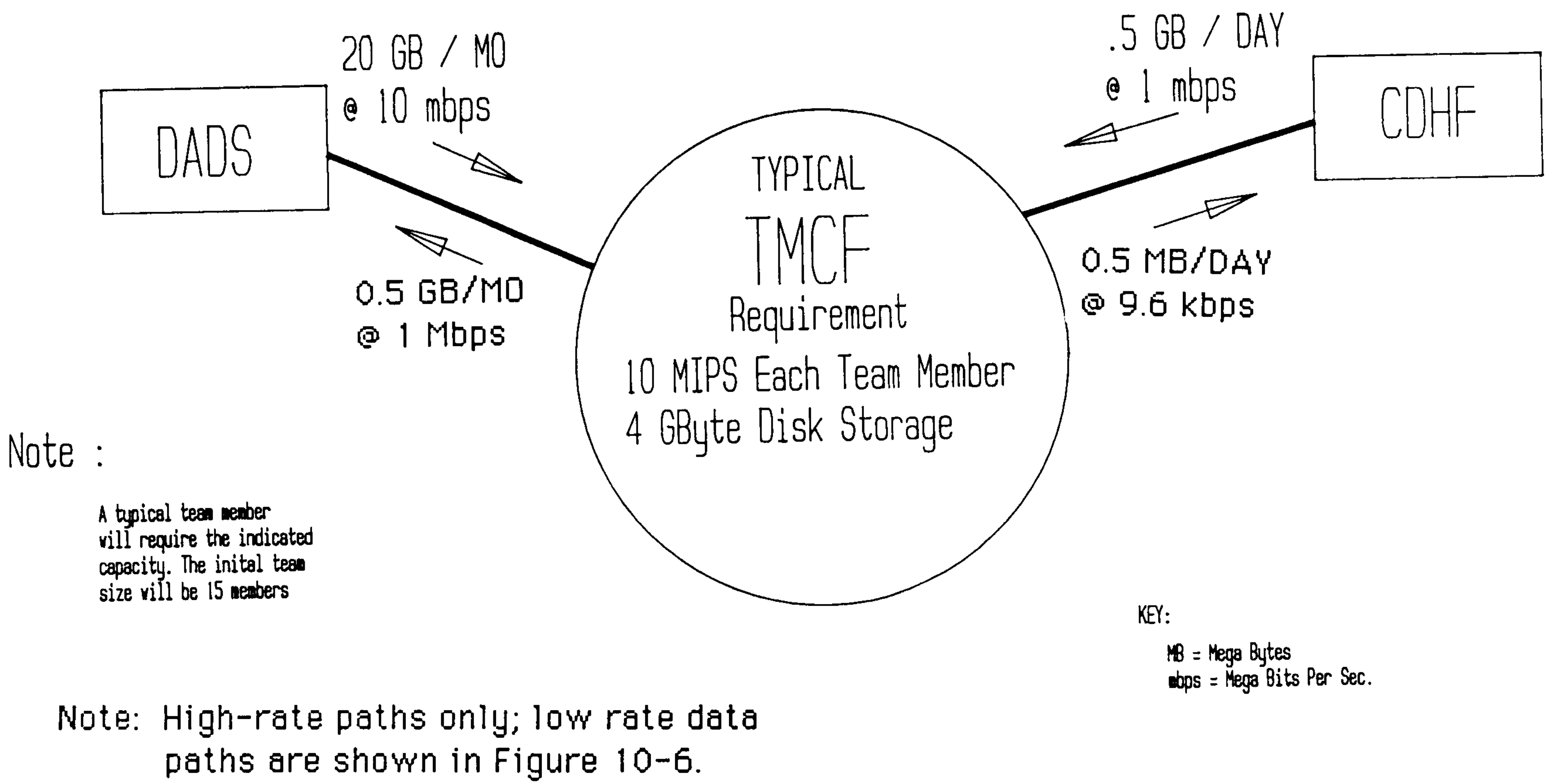
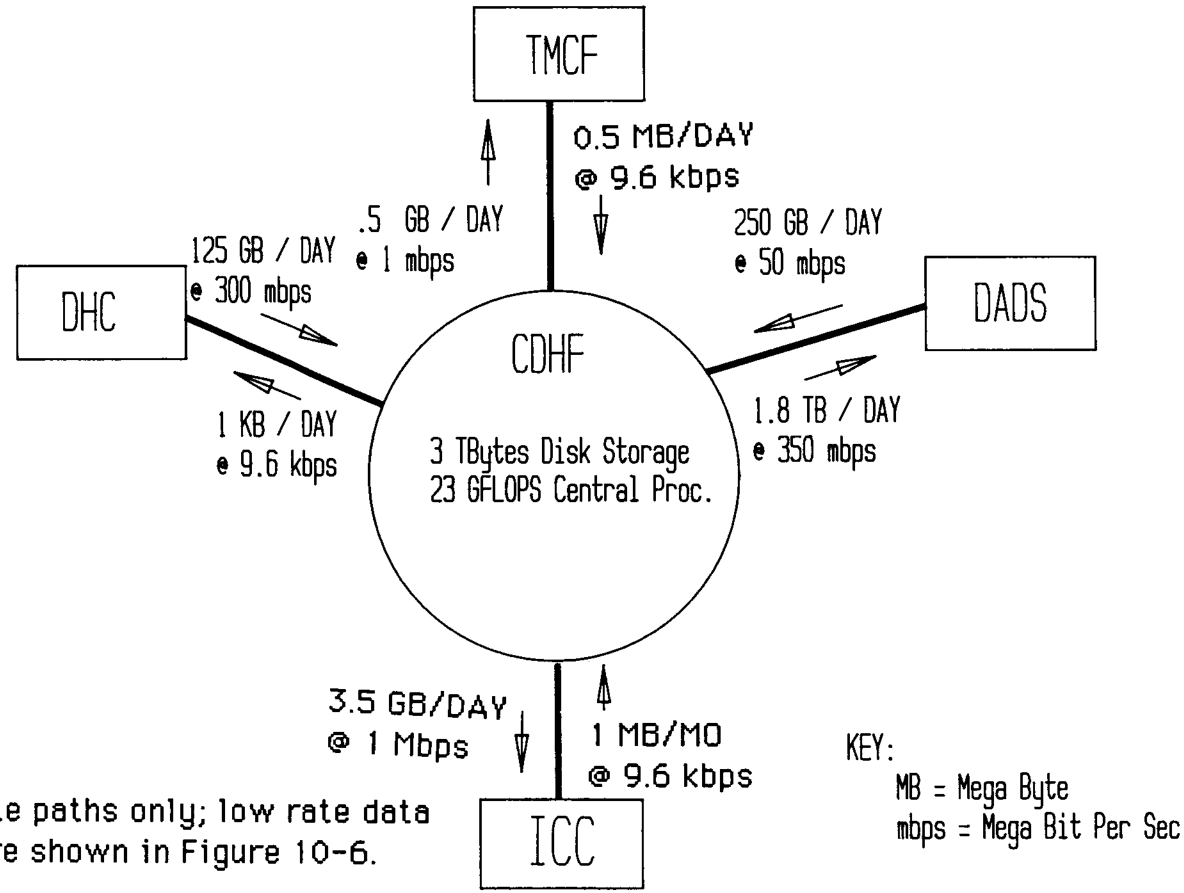
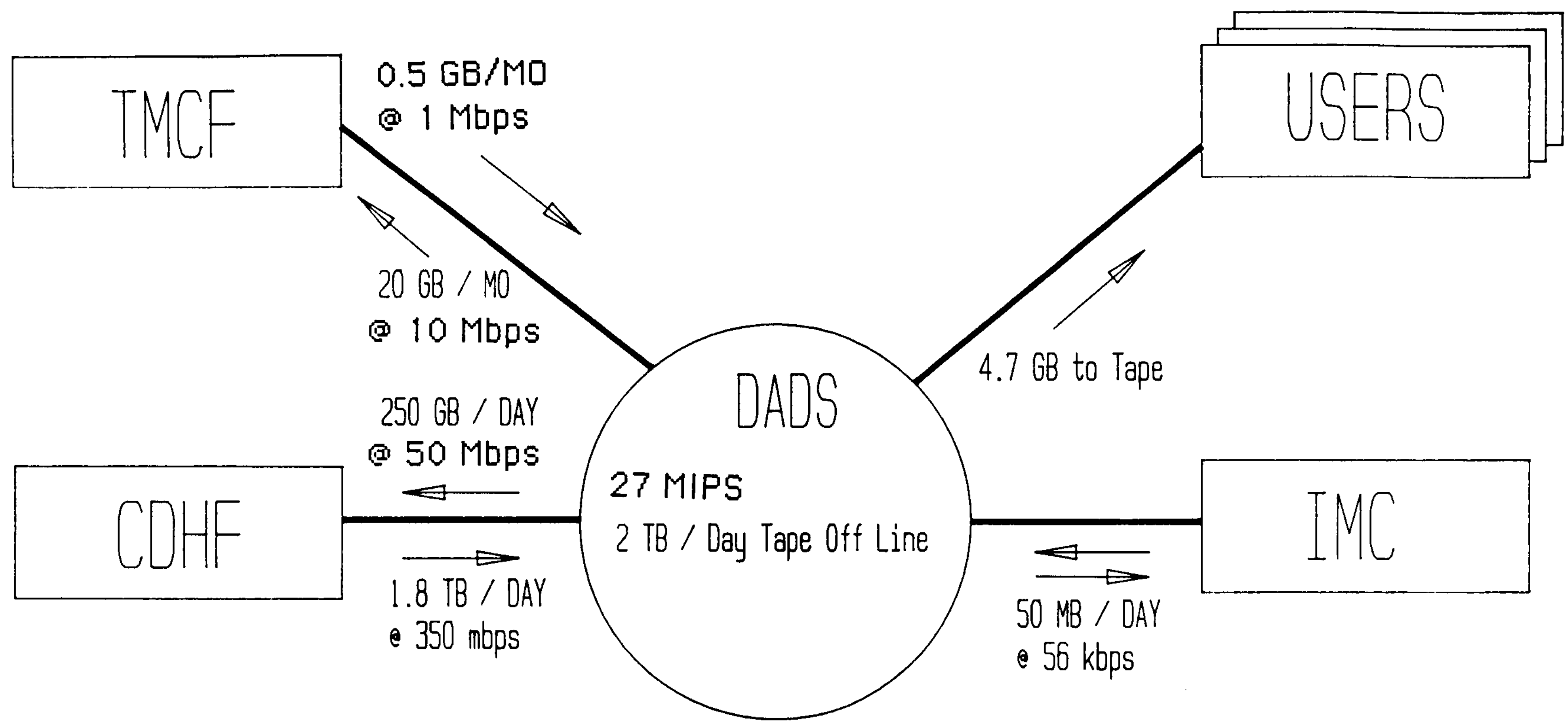


Figure 19. TMCF Performance Summary



Note: High-rate paths only; low rate data paths are shown in Figure 10-6.

Figure 20. CDHF Performance Summary



KEY:

MB = Mega Bytes

mbps = Mega Bits Per Sec.

Note: High-rate paths only; low rate data paths are shown in Figure 10-6.

Figure 21. DADS Performance Summary

ORIGINAL PAGE IS
OF POOR QUALITY

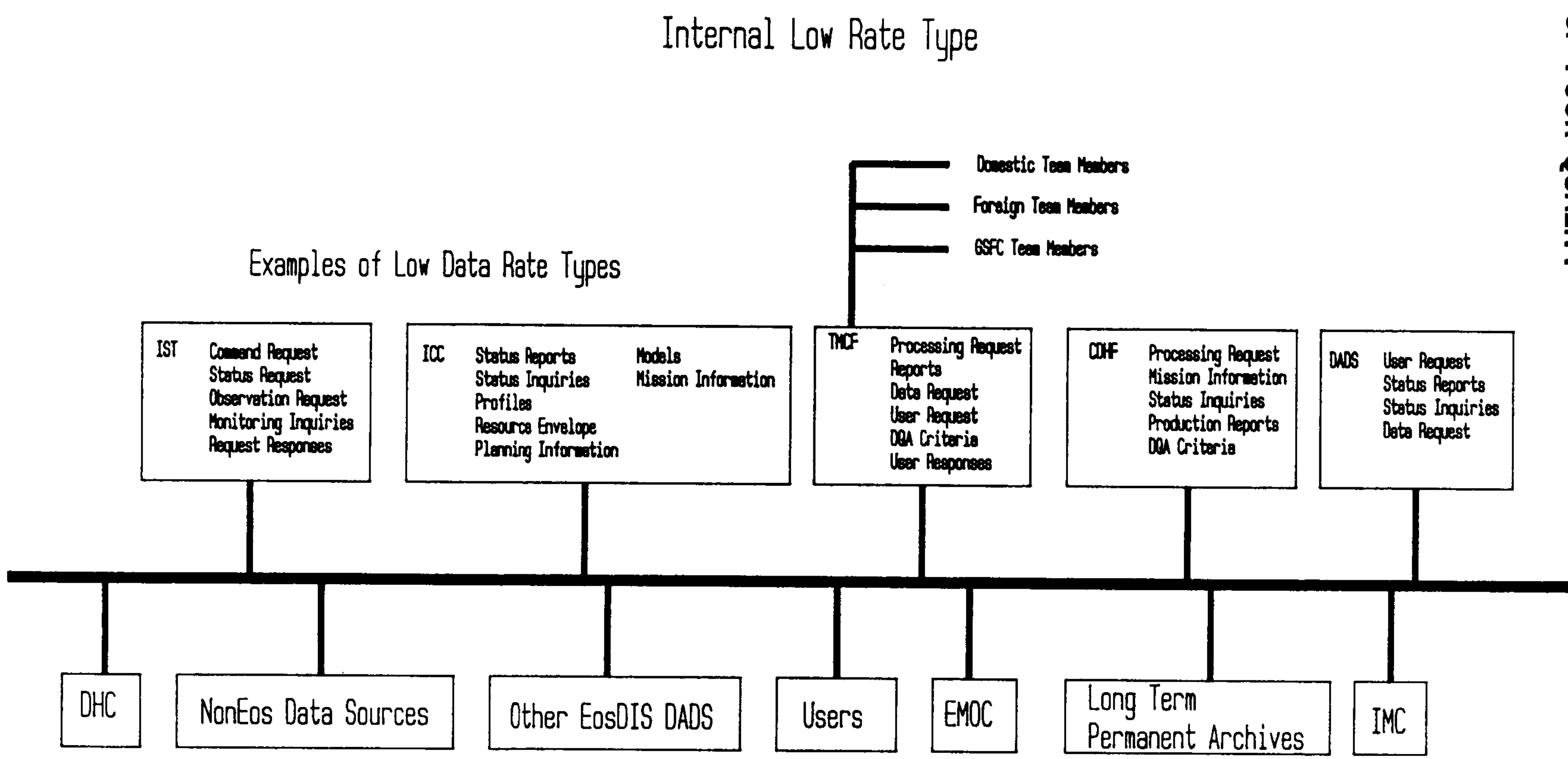


Figure 22. MIDACS External Low Rate Types

APPENDIX A
DATA DICTIONARY

Accepted-Data	=	*DADS ingested data that has been quality checked.*
Algorithm-Release-Announcement	=	*Announcement to the team that a debugged, working processing algorithm is now in use, containing information such as version numbers, availability of user's guide, etc.*
Analysis-Results	=	*Results of analysis of instrument performance over a long period which reflects trends in performance.*
Ancillary-Data	=	*Data other than MODIS-Instrument-Data required to perform MODIS data processing (such as platform and other instrument data).*
Archive-Ancillary-Data	=	*Ancillary data retrieved from the MODIS DADS.*
Archive-Data-Products	=	*MIDACS products routinely archived for potential user access and distributed in response to a product request.*
Archive-Data-Products-Release-DADS Authorization	=	*Authorization to release data products in the to data users.*
Archive-Data-Request	=	*A request for data to be retrieved from any EosDIS DADS.*
Authorized-Schedule-Data	=	*A schedule containing instrument resources and timelines, that have been approved by the EMOC through iteration with the ICC.*
Automated-Command-Sequences	=	*A human readable sequence of commands generated by the planning and scheduling processes and used for the generation of command loads.*
Browse-Data	=	*Subsets of a data set other than the directory and metadata that facilitates user selection of specific data having the required characteristics. For example, image data of a single channel with degraded resolution.*
Candidate-Observation-Sequence	=	*A human readable form of the instrument resources and timelines necessary to perform the observation request. These data are sent to the EMOC for approval.*

Catalog/Directory-Data	=	*A description of data available from the MIDACS DADS listed by platform, instrument, data processing level, algorithm identifier, parameter, time, geographic location, or combination.*
CDHF-Accepted-Data	=	*Data from the DHC that has undergone TBD data validation checks.*
CDHF-Database-Inquiry	=	*Request for CDHF temporary storage catalog information.*
CDHF-Database-Report	=	*Response to CDHF database inquiry.*
CDHF-Data-Products	=	*Levels 1-4 MODIS data products.*
CDHF-Ingested-Data	=	*Data received from the DHC that has been blocked by TBD methods.*
CDHF-Stored-Data	=	*Any subset of data sets cataloged and stored in the CDHF temporary storage.*
Command-Loads	=	*Encoded MODIS instrument command sequences as required by the on-board MODIS instrument control system and constructed so as to affect a specific action; e.g., "HV PWR ON"..*
Command-Request	=	*A command load generated by the IST, verified by the ICC, and immediately transmitted to the EMOC.*
Coordinated-Observation-Plan	=	*Any data received by the IST which has been selected as an observation request which has been coordinated to determine its consistency with the EOS science objectives.*
Correlative-Data	=	*Scientific data not from the MODIS instrument used to verify, interpret, or validate MODIS data products.*
Correlative-Data-Request	=	*Information required to initiate and support the transfer of Correlative-Data to the requestor.*
Critical-Command-Request	=	*A command request issued by the monitoring function when the state-of-health of the instrument or its performance is degraded.*
DADS-Status-Inquiry	=	*Request for a specific type of DADS-Status-Report.*
DADS-Status-Report	=	*Description of the DADS status, resources utilization, and performance.*

Data-Products-Release-Announcement	=	*Announcement to the team that a validated specialized or correlative data product is available to the scientific community.*
DHC-Data-Request	=	*Request to redesignate packet handling and processing priorities.*
Displays	=	* Plots, images, a list of requested data or status of the instrument or ground system communicated visually.*
Distribution-Request	=	*Request to send stored data, for production of MODIS data products, for archiving at the DADS or for use by the TMCF.*
DQA-Criteria	=	*Factors used to assess data quality.*
DQA-Reports	=	*Results of routine data quality assessment associated with data receipt and data product operations.*
Engineering-Data	=	MODIS-Engineering-Data + Platform-Engineering-Data.
Formatted-Observation-Request	=	*Any observation request received by the ICC that has been processed for input into the generation of instrument timelines.*
Generic-Observation-Request	=	*Observation requests that are predetermined and are consistent with the original science plan.*
Geography-Data	=	*Data that can be used to identify land and ocean boundaries or other Earth features necessary for the implementation or generation of the instrument commands. Used in generating instrument timelines.*
Headers	=	*Information about the attributes of standard, non-standard, or data products.*
ICC-Data-Request	=	*A request for information from the ICC.*
IMC-Inquiry	=	*Request for information on the operational status of the MODIS instrument or the MIDACS data system.* Production-Status-Inquiries + DADS-Status-Inquiries + Observation-Plan-Inquiry + Instrument-Operational-Status-Inquiries

IMC-Inquiry-Response	=	Production-Reports + DADS-Status-Reports + Instrument-Status-Reports + Observation-Plan-Information
Ingested-Data	=	*All products and data received by the DADS.* MODIS-Data-Products + Specialized-Products + Permanent-Archived-Data + Correlative-Data + Ancillary-Data + Processing-Algorithms
Instrument-Operational-Status-Inquiry	=	*An inquiry made by the IMC to determine the status of the ICC and/or the MODIS instrument.*
Instrument-Operations-Models	=	*Computer-compatible mathematical analogs of the MODIS instrument, used to estimate resource requirements during a modeled operation.*
Instrument-Status-Reports	=	*Information on the operating configuration of the MODIS instrument.*
IST-Coordination	=	*Planning and performance information relating to the MODIS instrument.*
Level-0-Data	=	*MODIS-Instrument-Data at original resolution, time order restored, with duplicates removed.*
Level-1A-Data	=	*Level 0 data which are reformatted reversibly, with Earth location, solar and instrument zenith angle, calibration data, and other ancillary data appended.*
Level-1B-Data	=	*Level-1A data to which the radiometric calibration algorithms have been applied, perhaps irreversibly, to produce radiances or irradiances, and to which the Earth-location and zenith-angle algorithms have been applied at the grid points.*
Level-2-Data	=	*Geophysical parameter data derived from the Level-1B data by application of geophysical parameter algorithms.*
Level-3-Data	=	*Earth-gridded geophysical parameter data (including Level-1 radiances), which have been averaged or composited in time and space.*
Level-4-Data	=	*TBD analyses of the lower levels of MODIS data.*

Management-Information	=	*Internal information about the DADS data store and catalog.*
Mission-Planning-Information	=	*Instrument operations schedule; information provided by the ICC to the CDHF to verify receipt and specify complete handling of data sets.*
MODIS-Data-Products	=	Levels 1-4 Data Products + Browse Data
MODIS-Engineering-Data	=	*Data other than MODIS-Science-Data generated within the MODIS instrument.*
MODIS-Instrument-Data	=	*Data originating within the MODIS instrument.*
	=	MODIS-Science-Data + MODIS-Engineering-Data
MODIS-Science-Data	=	*Unprocessed observations as generated by the MODIS instrument.*
Monitoring-Algorithms	=	*A procedure for transforming information for processing and displaying MODIS science and engineering data for monitoring instrument performance.*
Monitoring-Database-Inquiry	=	*Inquiry of the monitoring database to determine what instrument monitoring reports, data, and analyses are currently available.*
Monitoring-Database-Report	=	*Report of instrument monitoring functions and availability.*
Near-Real-Time-Data	=	*MODIS-Instrument-Data designated for Priority Processing.*
Near-Real-Time-Processing	=	*Processing accomplished within three to eight hours after input data become available.*
Near-Real-Time-Request	=	*Request to handle data in Priority Mode.*
Non-Standard-Products	=	*Products not routinely produced, standard products produced by an alternate algorithm, or combinations of standard products.*
Observation-Plan-Coordination	=	*Information exchange between a user requesting special MODIS services and the MODIS Instrument Team Leader. The exchange should culminate in a plan for MODIS Instrument Operation.*

Observation-Plan-Information	=	*Information describing observations planned for the MODIS instrument.*
Observation-Plan-Inquiry	=	*A request for information on planned MODIS instrument observations.*
Observation-Planning-Data	=	*Any data received by the IST which has been selected as a possible observation request which will undergo coordination to determine its consistency with the EOS science objectives.*
Observation-Request	=	*MODIS measurement request not covered by the current schedule or data handling procedures. The request is consistent with general science objectives and science mission plans and goals.*
Observation-Resource-Requirements	=	*Predicted instrument resources necessary to fulfill the objectives of the observation request.*
Orbit-&-Attitude-Data	=	*Data that describes the current or predicted orbital position and pointing of the platform or instrument axes.*
Order-Found-Status	=	*Status of the product order when located and retrieved. This information can be sent to the user via the IMC.*
Order-Placed-Status	=	*Status of the Product Order when the user request has been processed. This information can be sent to the user via the IMC.*
Order-Status-Data	=	*Status and billing of the product ordered through IMC.* Order-Found-Status + Order-Placed-Status
Organized-Data	=	*Data products which have been grouped according to the header, e.g., Level 1A data, Land data, or Ocean data.*
Permanent-Archive-Data	=	*Data retrieved from permanent archival storage.*
Permanent-Archive-Data-Request	=	*Request for data from the permanent archive.*
Planning-Input	=	*Information supplied to the Team Leader by the Team Members used to developed the Science Management Plan.*

Platform-Engineering-Data = *Data produced by the platform sensors that are used for operating the platform or as ancillary data.*

Preliminary-Algorithms = *Recently developed algorithms which have not been fully tested.*

Preliminary-Specialized-Data-Products = *Specialized data products which have not yet been verified or validated.*

Priority-Processing = *Immediate processing of designated data items without considering data item position in processing queues (cf. Routine Processing). Includes both Real-Time-Processing and Near-Real-Time-Processing.*

Priority-Ranked-Requests = *Received requests which the Team Leader has given priority ranking in concordance with the Science Management Plan. The Team Leader provides the schedule to the Team Members.*

Processed-Data = *Results of analyzing engineering data in real-time or trend analysis functions.*

Processed-User-Request = *Data request that has been processed by the Receive-User-Request function and is used to locate and retrieve the data.*

Processing-Algorithms = *A mathematical procedure used by the CDHF or the TMCF to process the MODIS data.*

Processing-Instructions = *Procedure and scheduling instructions that command the processing steps and the type of processing (routine, near-real-time, special, or reprocessing).*

Production-Report = Production Schedule +
Production Status

Production-Status-Inquiry = *Request for a production report.*

Quick-Look-Science-Data = *A subset (up to 100%) of MODIS-Science-Data (Real-Time or Priority Playback) used to monitor MODIS instrument performance (may not be completely processed).*

Real-Time-Processing = *Processing accomplished essentially as input data becomes available with only minimal storage delays for data buffering.*

Received-CDHF-Data = *Data received by the CDHF (received DHC data, archive data products, processing algorithms, DQA criteria).*

Received-Data	=	*Cataloged data stored at the TMCF's used for algorithm development and validation/verification of data.*
Received-DHC-Data	=	*Level-0 data and ancillary data from the DHC that has been acceptance checked and reformatted and has had a header appended.*
Received-Requests	=	*A TM data products request, TM observation request, or TM data processing request received by the Team Leader from Team Members in TMCF.*
Reference-Monitoring-Profile	=	*Expected MODIS instrument engineering parameter levels annotated with limits at which alarm status should be declared.*
Requested-Data-Products	=	*Data retrieved by the Process User Request function for distribution on a physical medium.*
Requested-Direct-Data	=	*Products generated on physical medium for distribution.*
Resource-Envelope	=	*Maximum allowable resource consumption levels for the MODIS instrument.*
Retransmission-Request	=	*Request for retransmission of data packets that do not meet quality standards.*
Retrieved-Data	=	*Data retrieved from DADS storage by the Process-User-Request function to fill a product order.*
Revised-Algorithms	=	*Algorithms which are tested and are ready for implementation.*
Routine-Processing	=	*Processing that considers data item position in data processing queues. Cf. Priority Processing.*
Schedule-Data	=	*English language descriptions of planned platform maneuvers or instrument operations.*
Science-Management-Plan	=	*A plan which states the areas of responsibility, technical goals, and time tables of each Team Member, developed by the Team Leader and Team Members.*
Selected-Data-Products	=	*Subsets of standard, near-real-time or specialized data product.*

Selected-Science-Data	=	*A subset of MODIS-Science-Data or MODIS data products used to monitor MODIS instrument performance. These data are transmitted to the ICC by the CDHF to analyze past instrument performance and are not used for real-time monitoring.*
Selected-Science-Data-Request	=	*A request for selected science data for monitoring instrument performance.*
Specialized-Data-Products	=	*Data products which are considered part of a specific research investigation and are produced for a limited region or time period, or data products which are not accepted by the project as standard items.*
Status-&-Trending-Data	=	*Instrument engineering and/or science data that has been analyzed to determine the operating status of the instrument and long-term trends. These data will be used to update any instrument models or algorithms used in analyzing engineering data.*
Stored-Data-Request	=	*Request for stored data from the CDHF temporary storage.*
Stored-Processed-Data	=	*Data which have been processed in real-time and stored for analyzing long-term trends in performance.*
TMCF-Observation-Request	=	*A TM-Observation-Request after approval by the Team Leader.*
TMCF-Processing-Request	=	Standard-Processing-Requests + Reprocessing Requests (approved by Team Leader) + Specialized-Data-Processing-Request
TMCF-Request-Response	=	*Response of the Team Leader to a TMCF-Processing-Request.*
TM-Data-Processing-Request	=	*Team Member data processing request not yet approved by the Team Leader for general release.*
TM-Data-Products-Request	=	*Team Member data products request not yet approved by the Team Leader for general release.*
TM-Observation-Request	=	*Team Member observation request not yet approved by the Team Leader.*

User-Data-Parameters	=	*Parameters used to locate and retrieve data to respond to a data request. Parameters used to identify data type, location, etc. in order to access the data.*
User-Observation-Request	=	*A special observation request not included in the current schedule but consistent with general science objectives and the science mission plan.*
User-Processing-Request	=	*Request by a User to generate MODIS-Data-Products not previously available.*
User-Product-Request	=	*Requests that distributed data products be delivered to a User from the MIDACS DADS.*
User-Request	=	User Product Request + User Observation Request + User Processing Request.
User-Request-Response	=	*Response to a user's request.*
Valid-Observation-Sequence	=	*A time-ordered sequence of instrument operations reflecting observation requests which have been determined to be consistent with the science plan, feasible from an orbit/geometry point-of-view and possibly coordinated with specific scenes.*
Verification/Validation-Study-Results	=	*Results of correlative and modeling studies.*

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No.
TM 100721

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle

MODIS Information, Data, and Control System (MIDACS) System Specifications and Conceptual Design

5. Report Date

December 1988

6. Performing Organization Code

636

7. Author(s)

D. Han, V. Salomonson, J. Ormsby, P. Ardanuy, A. McKay, D. Hoyt, S. Jaffin, B. Vallette, B. Sharts, D. Folta, E. Hurley, D. MacMillan

8. Performing Organization Report No.

89B00065

10. Work Unit No.

9. Performing Organization Name and Address

Space Data and Computing Division
Goddard Space Flight Center
Greenbelt, MD 20771

11. Contract or Grant No.

NAS5-29373
NAS5-28795

13. Type of Report and Period Covered

Technical Memorandum

12. Sponsoring Agency Name and Address

National Aeronautics and Space Administration
Washington, DC 20546-0001

14. Sponsoring Agency Code

15. Supplementary Notes

D. Han, V. Salomonson, and J. Ormsby: Goddard Space Flight Center, Greenbelt, MD 20771; P. Ardanuy, A. McKay, D. Hoyt, S. Jaffin, and B. Vallette: Research and Data Systems, Greenbelt, MD 20770; B. Sharts and D. Folta: General Sciences Corporation, Laurel, MD 20707; E. Hurley and D. MacMillan: Interferometrics, Inc., Vienna, VA 22180

16. Abstract The MODIS Information, Data, and Control System (MIDACS) Specifications and Conceptual Design Document discusses system level requirements, the overall operating environment in which requirements must be met, and a breakdown of MIDACS into component subsystems, which include the Instrument Support Terminal, the Instrument Control Center, the Team Member Computing Facility, the Central Data Handling Facility, and the Data Archive and Distribution System. The specifications include sizing estimates for the processing and storage capacities of each data system element, as well as traffic analyses of data flows between the elements internally, and also externally across the data system interfaces. The specifications for the data system, as well as for the individual planning and scheduling, control and monitoring, data acquisition and processing, calibration and validation, and data archive and distribution components, do not yet fully specify the data system in the complete manner needed to achieve the scientific objectives of the MODIS instruments and science teams. The teams have not yet been formed; however, it was possible to develop the specifications and conceptual design based on the present concept of EosDIS, the Level-I and Level-II Functional Requirements Documents, the Operations Concept, and through interviews and meetings with keys members of the scientific community.

17. Key Words (Suggested by Author(s))

Moderate Resolution Imaging Spectrometer MODIS
High Resolution Imaging Spectrometer HIRIS
Earth Observing System EOS
EOS Data and Information System EosDIS
Data Systems

18. Distribution Statement

Unclassified - Unlimited

Subject Category 43

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	125	

NASA FORM 1626 OCT 86