HUGHES INFORMATION TECHNOLOGY SYSTEMS

ERRATA NOTICE

EOS Core System (ECS) Project Contract No. NAS5-60000

April 19, 1996

Document No.: 305-CD-031-002

Title: Release B Langley DAAC Design Specification for the ECS Project

Enclosed please find change pages for the subject document. Please replace the pages as follows:

Remove	Insert
3-33 through 3-39	3-33 through 3-39
3-61 through 3-62	3-61 through 3-62
3-99 through 3-112	3-99 through 3-112

If you have any questions, please contact our Data Management Office at (301) 925-1037.

305-CD-031-002

EOSDIS Core System Project

Release B Langley DAAC Design Specification for the ECS Project

March 1996

Hughes Information Technology Systems
Upper Marlboro, Maryland

Release B Langley DAAC Design Specification for the ECS Project

March 1996

Prepared Under Contract NAS5-60000
CDRL Item #046

APPROVED BY

Rick Kochhar /s/ 3/26/96

Rick Kochhar, Release B CCB Chairman Date
EOSDIS Core System Project

Hughes Information Technology Systems
Upper Marlboro, Maryland

305-CD-031-002

This page intentionally left blank.

Preface

This document is one of nineteen comprising the preliminary design specifications of the SDPS and CSMS subsystem for Release B of the ECS project. A complete list of the design specification documents is given below. Of particular interest are documents number 305-CD-020, which provides an overview of the subsystems, and 305-CD-039, the Data Dictionary, for those reviewing the object models in detail. A Release B SDPS and CSMS CDR Review Guide (510-TP-004-001) is also available.

The SDPS and CSMS subsystem design specification documents for Release B of the ECS Project include:

305-CD-020-002	Release B Overview of the SDPS and CSMS Segment System Design Specification
305-CD-021-002	Release B SDPS Client Subsystem Design Specification
305-CD-022-002	Release B SDPS Interoperability Subsystem Design Specification
305-CD-023-002	Release B SDPS Data Management Subsystem Design Specification
305-CD-024-002	Release B SDPS Data Server Subsystem Design Specification
305-CD-025-002	Release B SDPS Ingest Subsystem Design Specification
305-CD-026-002	Release B SDPS Planning Subsystem Design Specification
305-CD-027-002	Release B SDPS Data Processing Subsystem Design Specification
305-CD-028-002	Release B CSMS Segment Communications Subsystem Design Specification
305-CD-029-002	Release B CSMS Segment Systems Management Subsystem Design Specification
305-CD-030-002	Release B GSFC Distributed Active Archive Center Design Specification
305-CD-031-002	Release B LaRC Distributed Active Archive Center Design Specification
305-CD-033-002	Release B EROS Data Center Distributed Active Archive Center Design Specification
305-CD-034-002	Release B ASF Distributed Active Archive Center Design Specification
305-CD-035-002	Release B NSIDC Distributed Active Archive Center Design Specification
305-CD-036-002	Release B JPL Distributed Active Archive Center Design Specification

305-CD-037-002	Release B ORNL Distributed Active Archive Center Design Specification
305-CD-038-002	Release B System Monitoring and Coordination Center Design Specification
305-CD-039-002	Release B Data Dictionary for Subsystem Design Specification

This document is a contract deliverable with an approval code 2. As such, it does not require formal Government approval, however, the Government reserves the right to request changes within 45 days of the initial submittal. Once approved, contractor changes to this document are handled in accordance with Class I and Class II change control requirements described in the EOS Configuration Management Plan, and changes to this document shall be made by document change notice (DCN) or by complete revision.

Any questions should be addressed to:

Data Management Office
The ECS Project Office
Hughes Information Technology Systems
1616 McCormick Drive
Upper Marlboro, MD 20774

Abstract

The Release B *Langley Research Center* (LaRC) Distributed Active Archive Center (DAAC) Design Specification describes the ECS subsystems at the ECS portion of the Langley DAAC. ECS Subsystem-Specific Design Specifications provide detailed design descriptions of the subsystems. This document shows the specific implementation of that design at the ECS portion of the Langley DAAC, including the identification of the specific software, hardware and network configuration for the ECS portion of the Langley DAAC.

Keywords: LaRC, Langley, DAAC, configuration, design

This page intentionally left blank.

Change Information Page

List of Effective Pages	
Page Number	**Issue**
Title	Submitted as Final
iii through xii	Submitted as Final
1-1 through 1-4	Submitted as Final
2-1 through 2-4	Submitted as Final
3-1 through 3-120	Submitted as Final
4-1 and 4-2	Submitted as Final
AB-1 through AB-4	Submitted as Final
A-1 through A-6	Submitted as Final

Document History			
Document Number	**Status/Issue**	**Publication Date**	**CCR Number**
305-CD-031-001	Preliminary	October 1995	95-0723
305-CD-031-002	Submitted as Final	March 1996	96-0236

This page intentionally left blank.

Contents

Preface

Abstract

1. Introduction

1.1 Identification 1-1
1.2 Scope 1-1
1.3 Purpose 1-2
1.4 Status and Schedule 1-2
1.5 Document Organization 1-2

2. Related Documentation

2.1 Parent Documents 2-1
2.2 Applicable Documents 2-2
2.3 Information Documents Not Referenced 2-4

3. Langley DAAC Configuration

3.1 Introduction 3-1
3.1.1 Langley DAAC Overview 3-1
3.1.2 DAAC-Specific Mission and Operations Activities 3-2
3.2 Langley External Interfaces 3-7
3.3 Computer Software Component Analysis 3-16
3.3.1 Software Subsystem Overview 3-17
3.3.2 Software Subsystem Analysis Summary 3-19
3.4 DAAC Hardware and Network Design 3-32

3.4.1 Langley DAAC LAN Configuration 3-32
3.4.2 DAAC Hardware Configuration 3-36
3.5 Hardware/Software Mapping 3-111

4. Future Releases

Figures

3.1.1-1 SDPS Subsystems and Configuration Items 3-2
3.1.1-2 CSMS Subsystems and Components 3-3
3.2-1 EDC ECS DAAC External Interfaces 3-7
3.4.1-1 Langley DAAC LAN Topology 3-32
3.4.2-1 Langley ECS DAAC Hardware Configuration Diagram 3-37
3.4.2-2 DMGHW CI RAID Disk Configuration 3-62

Tables

3.1.2-1 Langley Operations Support Functions 3-6
3.2-1 LaRC External Interfaces 3-11
3.3.2-1 Langley Components Analysis 3-21
3.4.1-1 Release B Networking Hardware for Langley DAAC LAN 3-34
3.4.1.1-1 Estimated Release B Data Flows for the Langley DAAC 3-35
3.4.2.2-2 Release B Datasets Held by the Data Server Subsystem 3-48
3.4.2.2-3 LaRC DBMS Repository Sizing (GB) 3-50
3.4.2.3-1 LaRC User System Accesses per Day for Science Users 3-54
3.4.2.3-2 LaRC Science User Search Types for Gateway CI Services 3-54
3.4.2.3-3 LaRC Searches per hour for Science Users 3-55
3.4.2.3-4 LaRC User System Accesses per Day for Non-Science Users 3-55
3.4.2.3-5 LaRC Non-Science User Search Types for Gateway CI Service 3-55
3.4.2.3-6 LaRC Science User Transaction Analysis for Gateway CI Service 3-56
3.4.2.3-7 LaRC Science User Transaction Analysis for Advertising Service CI 3-56

3.4.2.3-8 LaRC Non-Science User Transaction Analysis for Gateway CI Service..... 3-57
3.4.2.3-9 LaRC Non-Science User Transaction Analysis for Advertising Service CI m) ... 3-57
3.4.2.3-10 DBMS Transaction Analysis Summary..... 3-58
3.4.2.3-11 DBMS Transaction Sensitivity Analysis Results 3-58
3.4.2.3-12 Vendor Platform Performance Estimates..... 3-59
3.4.2.3-13 DMGHW CI Disk Capacity Requirements..... 3-60
3.4.2.3-14 LaRC DAAC Hardware Configuration..... 3-63
3.4.2.4-1 LaRC DAAC Specific Ingest Volume Requirements 3-64
3.4.2.4-2 Ingest HWCI Component Descriptions 3-65
3.4.2.4-3 Ingest HWCI Component Sizing for the LaRC DAAC Configuration..... 3-66
3.4.2.6-1 Breakdown of Instrument Production at LaRC 3-72
3.4.2.7.1-1 LaRC Static Model Summary Results..... 3-79
3.4.2.7.1-2 Comparison of LaRC SPRHW Requirements and Specification..... 3-91
3.4.2.8.1-1 LaRC HP Openview Collection Processing Requirement..... 3-100
3.4.2.8.1-2 LaRC Log Entry Storage Volume - Release B..... 3-102
3.4.2.8.1-3 MDA Data Conversion to Sybase Processing Requirement 3-102
3.4.2.8.1-4 CSS/MSS Server Configuration - Requirements Estimate 3-104
3.4.2.8.1-5 DDTS Benchmark Results..... 3-105
3.4.2.8.1-6 MSS Workstation #1 Configuration - Requirements 3-106
3.4.2.8.1-7 MSS Workstation #2 Configuration - Requirements 3-107
3.4.2.8.1-8 COTS Product Disk Requirements 3-109
3.4.2.8.1-9 LaRC CSS Release B Storage Requirement..... 3-109
3.4.2.8.1-10 LaRC MSS Release B Storage Requirements 3-110
3.5-1 Langley Hardware to Software Component Mapping 3-112
A-1 Larc Detailed SPRHW Configurations for Release A and Release B A-2

Appendix A

Abbreviations and Acronyms

This page intentionally left blank.

1. Introduction

1.1 Identification

This Release B Langley DAAC Design Specification for the ECS Project, Contract Data Requirement List (CDRL) Item 046, with requirements specified in Data Item Description (DID) 305/DV2, is a required deliverable under the Earth Observing System Data and Information System (EOSDIS) Core System (ECS), Contract NAS5-60000.

1.2 Scope

Release B of ECS supports functional capabilities and services required to meet driving requirements and milestones including:

- Functionality/services required to support mission operations for the continuation of TRMM, as well as the initiation of LANDSAT 7, COLOR, ADEOS II, and EOS AM-1. This includes planning and scheduling, command and control, production data processing, data distribution and other ECS functions.
- Functionality/services required to support mission operations for the initiation of SAGE III (METEOR) and ACRIM Flight of Opportunity (FOO). This includes production data processing, data distribution and other ECS functions.
- Provide information management, data distribution and a high level archive for the SAR data from the ERS-1/2, JERS-1 and RADARSAT spacecraft.
- Functionality/services required to support EOS ground system interface testing which includes end-to-end mission simulations, communication services for EBnet, network management services and other ECS services.
- Functionality/services required for V0 Interoperability.
- Functionality/services required for Science Software I&T Support for TRMM, LANDSAT 7, COLOR, ADEOS II, EOS AM-1, SAGE III (METEOR), and ACRIM FOO.

Several of the driving requirements and milestones were initially supported by Release A but are expanded upon for Release B. For example, infrastructure Data Flow, End-to-End Testing and Simulation Readiness Testing were supported early-on by Release A, and are fully supported by Release B during the final phases of testing. Similarly, V0 interoperability (one way) is supported by Release A for GSFC, LaRC, and EDC DAACs and is expanded to two way interoperability for all DAACs at Release B.

ECS will provide support to eight Distributed Active Archive Centers (DAACs). The DAACs are tasked with generating EOS standard data products and carrying out NASA's responsibilities for data archive, distribution and information management. The DAACs serve as the primary user

interface to EOSDIS. These DAACs are located at: Goddard Space Flight Center (GSFC) Greenbelt, MD; Langley Research Center (LaRC) Hampton, VA; Oak Ridge National Laboratory (ORNL) Oak Ridge, TN; EROS Data Center (EDC) Sioux Falls, SD; National Snow and Ice Data Center (NSIDC) Boulder, CO; Jet Propulsion Laboratory (JPL) Pasadena, CA; the Alaska SAR Facility (ASF) at the University of Alaska Fairbanks; and the Consortium for International Earth Science Information Network (CIESIN) located in Michigan.

This document is part of a series of documents comprising the Science and Communications Development Office (SCDO) design specification for the Communications and System Management segment (CSMS) and the Science and Data Processing Subsystem (SDPS) for Release B. The series of documents includes an overview, a design specification document for each subsystem, and a design implementation document for each DAAC involved in the release, as well as one for the System Monitoring and Coordination Center (SMC).

This document specifically focuses on the Langley DAAC ECS configuration and capabilities at Release B. It is released, and reviewed at the formal Release B Critical Design Review (CDR).

This document reflects the February 1996 Technical Baseline maintained by the contractor configuration control board in accordance with ECS Technical Direction No. 11, dated December 6, 1994.

1.3 Purpose

The purpose of this document is to show the elements of the Release B ECS science data processing and communications design that will support the ECS portion of the Langley DAAC in meeting its objectives. The Release B SDPS/CSMS Design Specification for the ECS Project (305-CD-020-002) provides an overview of the ECS subsystems and should be used by the reader in order to get a basic understanding of ECS design components. The Release Plan Content Description for the ECS Project (222-TP-003-008) provides a detailed mapping of functional capabilities and services that will be available for each release. While some DAAC configurations vary depending on the mission/capability requirements for ECS at their DAAC, the Langley DAAC at full ECS capability will include all of the ECS subsystems.

1.4 Status and Schedule

This submittal of DID 305/DV2 meets the milestone specified in the Contract Data Requirements List (CDRL) for Critical Design Review (pre-CDR) of NASA Contract NAS5-60000. The submittal will be reviewed during the Release B (CDR), and changes to the design which result from that review will be reflected in subsequent updates.

1.5 Document Organization

This document is organized to describe the design of the ECS portion of the Langley DAAC as follows:

Section 1 provides information regarding the identification, scope, status and schedule, and organization of this document.

Section 2 provides a listing of the related documents which were used as source information for this document.

Section 3 provides a description of the ECS design at the Langley DAAC. It includes a description of the DAAC external interfaces, ECS software implementation, including identification of Off the Shelf (OTS) products, hardware configuration and operational activities.

- Subsection 3.1 establishes the context for the technical discussions with an overview of the specific ECS portion of the Langley DAAC mission and LaRC Release B operations. It identifies the key ECS related mission and operations activities that are supported via the ECS functionality at the DAAC.
- Subsection 3.2 addresses the external interfaces of the ECS subsystems as implemented at the ECS portion of the Langley DAAC. Major interfaces include those with the users, and with SDPF, EDOS, Version 0 DAACs, NOAA ADC, and the CERES, MISR, MOPITT, SAGE III, and ACRIM Scientific Computing Facilities (SCF).
- Subsection 3.3 provides a software component analysis. There are 10 ECS data processing and communications subsystems that contain Hardware Configuration Items (HWCIs) and Computer Software Configuration Items (CSCIs). This section addresses the CSCIs and their corresponding lower level Computer Software Components (CSCs). The CSCs are described in detail in their respective subsystem design specification documents. In this section, the CSCs are captured in a single table, broken down by Subsystem/CSCI. The table lists the CSCIs and the associated CSCs. Notes are provided to expand upon generic explanations from the body of the Subsystem Design Specifications to describe what makes the particular CSC specific to the DAAC. In addition, when a CSC is identified as Off-the-shelf (OTS), the candidate product is identified.
- Subsection 3.4 provides a DAAC-specific discussion of the ECS data processing and communications Hardware Configuration Items (HWCIs). This section identifies the HWCI components and indicates the specific components and quantities that are resident at the DAAC. It includes the Local area network (LAN) configuration and the rationale for the specific hardware configuration.
- Subsection 3.5 provides a software to hardware configuration mapping.

Section 4 gives a description of what can be expected in the next release of ECS.

Appendix A provides detailed configurations for the Data Processing Subsystem's Science Processing hardware suite.

The section, Abbreviations and Acronyms, contains an alphabetized list of the definitions for abbreviations and acronyms used in this document.

This page intentionally left blank.

2. Related Documentation

2.1 Parent Documents

The parent documents are the documents from which the scope and content of this Release B Langley DAAC Design Specification is derived.

194-207-SE1-001	System Design Specification for the ECS Project
209-CD-005-005	Interface Control Document Between the EOSDIS Core System (ECS) and Science Computing Facilities (SCF)
209-CD-006-005	Interface Control Document Between the EOSDIS Core System (ECS) and the National Oceanic and Atmospheric Administration (NOAA) Affiliated Data Center (ADC)
210-CD-001-003	Risk Assessment Report for the ECS Project
220-CD-001-004	Communications Requirements for the ECS Project, Revision 1
305-CD-020-002	Release B SDPS/CSMS Design Specification Overview for the ECS Project
305-CD-021-002	Release B SDPS Client Subsystem Design Specification for the ECS Project
305-CD-022-002	Release B SDPS Interoperability Subsystem Design Specification for the ECS Project
305-CD-023-002	Release B SDPS Data Management Subsystem Design Specification for the ECS Project
305-CD-024-002	Release B SDPS Data Server Subsystem Design Specification for the ECS Project
305-CD-025-002	Release B SDPS Ingest Subsystem Design Specification for the ECS Project
305-CD-026-002	Release B SDPS Planning Subsystem Design Specification for the ECS Project
305-CD-027-002	Release B SDPS Data Processing Subsystem Design for the ECS Project
305-CD-028-002	Release B CSMS Communications Subsystem Design Specification for the ECS Project
305-CD-029-002	Release B CSMS Systems Management Subsystem Design Specification for the ECS Project

305-CD-039-002	Release B Data Dictionary for the ECS Project Subsystem Design Specification
307-CD-004-001/ 329-CD-004-001	Release B Science Data Processing Segment (SDPS) Release and Development Plan for the ECS Project
307-CD-005-001/ 329-CD-005-001	Release B Communications and System Management Segment (CSMS) Release and Development Plan for the ECS Project
311-CD-002-004	Science Data Processing Segment (SDPS) Database Design and Database Schema Specifications for the ECS Project
311-CD-003-004	Communications and System Management Segment (CSMS) Database Design and Database Schema Specifications for the ECS Project
313-CD-006-002	Release B SDPS/CSMS Internal Interface Control for the ECS Project
402-CD-003-001/ 319-CD-006-001	Release B System and Segment Integration and Test Plan for the ECS Project
222-TP-005-001	Release B IDR Engineering Plan for the ECS Project
222-TP-006-002	Release B Document Tree for the ECS Project
222-TP-010-001	Release B IDR Design Development Plan for the ECS Project
510-TP-003-001	Release B (EOS-AM1/Landsat-7) SDPS/CSMS IDR Review Guide for the ECS Project

2.2 Applicable Documents

The following documents are referenced within this Specification, or are directly applicable, or contain policies or nding upon the content of this document.

206-CD-001-002	Version 0 Analysis Report for the ECS Project
209-CD-001-003	Interface Control Document Between EOSDIS Core System (ECS) and the NASA Science Internet
209-CD-005-005	Interface Control Document Between EOSDIS Core System (ECS) and Science Computing Facilities (SCF)
209-CD-011-004	Interface Control Document Between EOSDIS Core System (ECS) and the Version 0 System
209-CD-027-001	Interface Control Document Between EOSDIS Core System (ECS) and the SAGE III Mission Operations Center (MOC)
304-CD-002-002	SDPS Requirements Specification for the ECS Project
305-CD-038-002	Release B System Monitoring and Coordination Center Design Specification for the ECS Project
313-CD-006-002	Release B CSMS/SDPS Internal Interface Control Document for the ECS Project
403-CD-002-001	Release B Verification Specification for the ECS Project

409-CD-002-001	ECS Overall System Acceptance Test Plan for Release B
513-CD-002-001	Release B Hazard Analyses for the ECS Project
514-CD-001-004	Security-Sensitive Items List for the ECS Project
515-CD-002-002	Release B Availability Models/Predictions for the ECS Project
516-CD-002-002	Release B Reliability Predictions for the ECS Project
518-CD-002-001	Release B Maintainability Predictions for the ECS Project
604-CD-001-004	Operations Concept for the ECS Project: Part 1-- ECS Overview
604-CD-002-003	ECS Operations Concept Document, Part 2: Release B
605-CD-002-001	Release B SDPS/CSMS Operations Scenarios
615-CD-002-001	Release B Special Maintenance and Test Equipment for the ECS Project
622-CD-002-001	Release B Training Plan for the ECS Project
627-CD-001-002	Security Risk Management Plan for the ECS Project
705-CD-005-002	Release B IDR Presentation Package for the ECS Project
160-TP-004-001	User Pull Analysis Notebook [for the ECS Project]
210-TP-001-006	Technical Baseline for the ECS Project
160-TP-002-001	Version 1 Data Migration Plan
222-TP-003-008	Release Plan Content Description for the ECS Project
222-TP-005-001	Release B IDR Engineering Plan for the ECS Project
222-TP-006-002	Release B Document Tree for the ECS Project
420-TP-001-005	Proposed ECS Core Metadata Standard Release 2.0, Technical Paper
420-TP-010-001	Transition to Release B, Technical Paper
430-TP-001-001	SDP Toolkit Implementation with Pathfinder SSM/I Precipitation Rate Algorithm
440-TP-006-002	Production Topologies: A Trade-Off Study Analysis for the ECS Project
440-TP-007-001	Platform Families for the ECS Project
440-TP-008-001	Distributed and Parallel Processing for ECS Science Algorithms: A Trade Analysis
510-TP-004-001	Release B (EOS-AM1/Landsat-7) SDPS/CSMS CDR Review Guide for the ECS Project

2.3 Information Documents Not Referenced

The following documents, although not referenced herein and/or not directly applicable, do amplify and clarify the information presented in this document. While not binding on the content of this design specification, these and many additional ECS documents are available via the EDHS. EDHS can be accessed via the World Wide Web (WWW) at the following Universal Reference Location (URL): http://edhs1.gsfc.nasa.gov/. Please note that Internet links cannot be guaranteed for accuracy or currency.

302-CD-003-001	Release B Facilities Plan for the ECS Project
101-303-DV1-001	Individual Facility Requirements for the ECS Project
333-CD-003-002	SDP Toolkit Users Guide for the ECS Project
601-CD-001-004	Maintenance and Operations Management Plan for the ECS Project
101-620-OP2-001	List of Recommended Maintenance Equipment for the ECS Project
828-RD-001-002	Government Furnished Property for the ECS Project
430-TP-001-001	SDP Toolkit Implementation with Pathfinder SSM/I Precipitation Rate Algorithm, Technical Paper
440-TP-001-001	Science Data Server Architecture Study [for the ECS Project]
423-16-01	Goddard Space Flight Center, Data Production Software and Science Computing Facility (SCF) Standards and Guidelines
423-16-02	Goddard Space Flight Center, PGS Toolkit Requirements Specification for the ECS Project
423-41-02	Goddard Space Flight Center, Functional and Performance Requirements Specification for the Earth Observing System Data and Information System (EOSDIS) Core System
540-022	Goddard Space Flight Center, Earth Observing System (EOS) Communications (ECOM) System Design Specification
560-EDOS-0211.0001	Goddard Space Flight Center, Interface Requirements Document Between EDOS and the EOS Ground System (EGS)

3. Langley DAAC Configuration

3.1 Introduction

3.1.1 Langley DAAC Overview

The LaRC Distributed Active Archive Center (Langley DAAC) is one of the eight DAACs that are part of the NASA Earth Observing System Data and Information System (EOSDIS). These DAACs are organized to support specific scientific disciplines. The objective of the Langley DAAC is to ingest, process, archive, and distribute science data and provide support services to its users in the discipline areas of radiation budget, clouds, aerosols, and tropospheric chemistry.

The requirements to support the following activities have driven the design of the ECS portion of the Langley DAAC: the TRMM Mission, EOS AM-1 mission, the NASA Flight-Of-Opportunities missions, the Version 0 to Version 1 transition, Version 0 interoperability, Affiliated Data Center (ADC) interoperability, data flow, end-to-end testing and simulation readiness testing. The Release Plan provides a description of the missions and the driving requirements that must be satisfied to support these activities. This document provides a description of the ECS design components that are specific for the ECS portion of the Langley DAAC. This document also elaborates upon design-generic components that are of special interest to the Langley users, data producers, and DAAC staff.

Langley-specific design benefited greatly from the Langley Version 0 DAAC activities. The lessons learned from the Langley Version 0 DAAC activities are guiding the planning for the Version 0 to Version 1 transition. The goal of the Langley Version 0 DAAC was prototyping in areas with high technical risk for ECS Version 1. One labor- and time-intensive area is the use of the HDF data format and the necessary upgrade of version 0 metadata to access the HDF-related services. Another area is the ingesting, archiving, and distributing data associated with many different types of research projects (e.g. ERBE, SAGE II, field campaigns like FIRE). The Langley Version 0 DAAC established an archive composed of multiple media types (e.g. optical WORM platters, optical rewritable disks, one-off CD-ROM, 3480 tapes). The Langley Version 0 DAAC developed, in close concert with the Langley User Working Group, an information management system. The Langley Version 0 DAAC integrated and continues to enhance the ERBE project's multiple-satellite, multiple instrument data processing system. The data products produced by this system are categorized as live data sets, a DAAC-unique extension. The ECS designers have carefully considered each of these contributions. A description of these contributions is provided in the Version 0 Analysis Report. Issues related to ingest of Version 0 data migration are discussed in the LaRC Version 0 to Version 1 Data Migration Plan (160-TP-002-001).

Figures 3.1.1-1 and 3.1.1-2 illustrate the ECS SDPS and CSMS subsystems and their components for Release B. The bulk of this document focuses on the selected elements of the ECS design that are use to achieve Release B objectives at the DAAC. Section 2.1 of this document identifies CDR-era Design Specifications which provide detailed information on each subsystem.

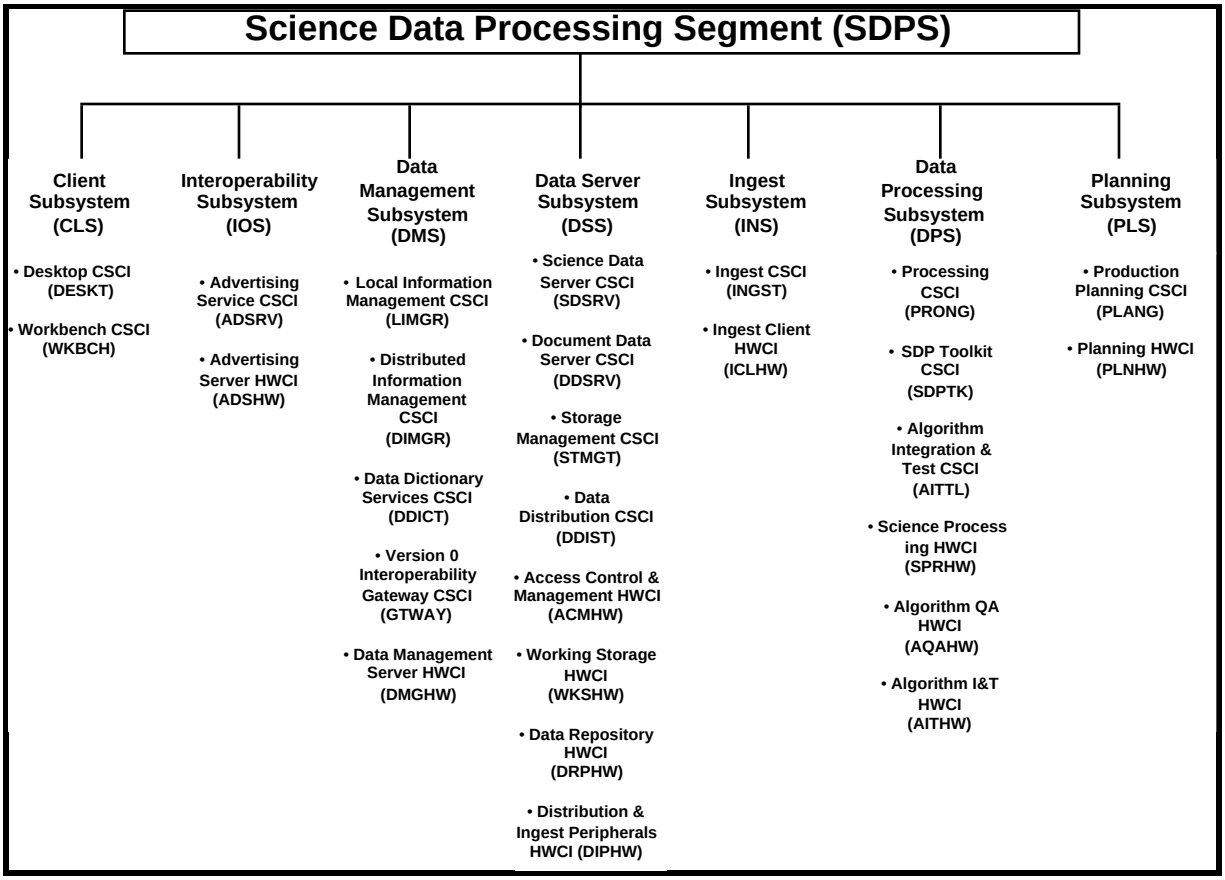

Figure 3.1.1-1. SDPS Subsystems and Configuration Items

3.1.2 DAAC-Specific Mission and Operations Activities

The Langley-specific objectives of Release B are to provide the following:

- TRMM CERES mission support;
- EOS AM-1 CERES, MISR and MOPITT mission support;
- SAGE III (aboard the Russian Space Agency METEOR-3M satellite) mission support;
- ACRIM Flight-Of-Opportunity mission support;
- Independent Verification and Validation (IV&V) Support;

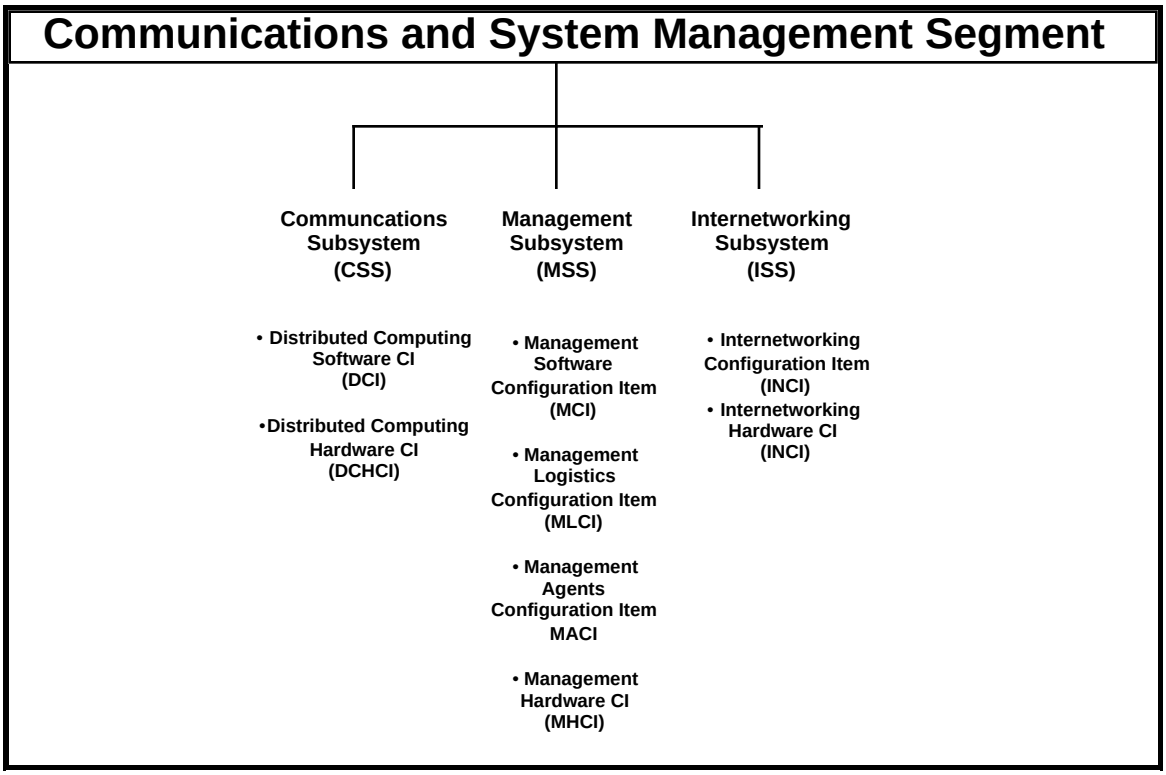

Figure 3.1.1-2. CSMS Subsystems and Components

- Interoperability (NOAA, V0);
- Science Software Support;
- Building from Version 0;
- Version 0 Data Migration;
- DAAC Site Activation

TRMM CERES, EOS AM-1 CERES, MISR, MOPITT, SAGE III (METEOR) and ACRIM Flight-Of-Opportunity Missions Support: ECS provides science software integration and test support, data ingest, processing, archiving, and distribution for these instruments. DAAC-unique components of the Langley DAAC will provide some ancillary and correlative data to support the processing and science QA for these missions.

Independent Verification and Validation (IV&V) Support: Prior to the Release Readiness Review (RRR), the IV&V contractor can witness and/or monitor release acceptance testing and document nonconformances. Upon successful completion of the RRR, the IV&V contractor verifies that the ECS release operates correctly within the EOS Ground System (EGS). The ECS contractor, specifically the Independent Acceptance Test Organization (IATO), supports the IV&V contractor in this effort for a period of one month following RRR at the operational sites. The IATO coordinates personnel, facilities,

and equipment support in the resolution of ECS nonconformances identified during IV&V testing. ECS contractor Maintenance and Operations personnel also support IV&V activities at operational centers, as necessary.

Interoperability: Interoperability involves two different capabilities. First, outgoing interoperability allows users to log into the ECS and access ECS services, including the ability to access non-ECS data products from a site external to ECS directly from the ECS user interface. Second, incoming interoperability allows users, who are logged into a non-ECS site, to access ECS data products directly from the non-ECS user interface, using non-ECS IMS services.

Two-way interoperability between ECS Version 0 and Version 1 is required prior to the V0/V1 transition to ease the transition process.

One-way outgoing interoperability with ADCs (ECS to ADCs) is also required to ease the V0/V1 transition. Two-way interoperability with ADCs is not required for Release B.

Science Software Support: The ECS project recommends that the required SDP hardware strings be made available several months prior to science software delivery (beta, engineering, or mission versions) to allow the SCF users to perform remote testing prior to formal science software delivery, integration, and test (SSI&T).

To support full end-to-end testing of the algorithms, necessary ECS infrastructure software (ancillary/auxiliary data ingest and preparation, DAAC-to-DAAC dat transfers, Level 0 data validation, algorithm delivery, and algorthm product QA services) must be in place by the end of the Mission I&T for each instrument. At the completion of testing, the algorithms will be integrated with other ECS components.

Building from Version 0: Building on Version 0 for a release implies that the release will be capable of matching (in general) the functionality of Version 0 plus adding some features that Version 0 does not have (i.e. building on to, or enhancing, existing Version 0 capabilities). This does not mean the release will match every individual function/capability of Version 0. It will be possible (through interoperability) to access some Version 0 functions, without having to make them part of ECS.

The ECS V1 Client will be implemented in Release B, and will completely replace the V0 Client implemented in Release A. In Release B, ECS will deploy an infrastructure that exceeds that in V0, and which will provide the foundation from which to add future enhancements which exceed V0.

Version 0 (V0) Data Migration: Version 0 (V0) data migration includes the ability to transition V0 data sets from V0 to V1; and provide support, data management, search, and access capabilities for these data sets. A select number of V0 data sets were made available at Release A at the Release A DAACs. Additional data migration takes place during Release B operations. For further details, refer to the Version 1 Data Migration Plan.

DAAC Site Activation: The EOSDIS DAACs have the mission of processing, archiving and distributing earth science data. Langley DAAC site activation was initially performed

for IR1, with phased activations required for Release A and Release B. These phased activations are characterized by increases in hardware and software functionality, including added operational responsibilities. This section discusses activities that will take place to facilitate a smooth transition from Release A to Release B.

The ECS contractor will schedule a series of site coordination trips to all DAACs. The objective of these trips is to ensure that the ECS contractor and the DAAC managers are in agreement with all operational issues. When ECS delivers Release B to the sites, ECS will work with the host organizations to ensure that hardware and software installation and segment and system testing all occur in a pre-planned manner that is sensitive to the mission of the host organization. Coordination topics include facility requirements, locations of Release B equipment and personnel, installation and test periods, etc.

Issues of when training and to whom training is provided on Release B products are critical because of the potential impact on operations and user support. Training on COTS hardware and software, and application software, regardless of the development track, is an absolute necessity. If the site's user services are unable to handle issues about a Release B product, additional demands on developers' time will be made to isolate, remedy, or suggest work-arounds to the issues.

The facility access dates must be at least 2 months prior to the scheduled initial Release B installation date to provide time for site verification inspection, completion of Government facility preparations, and receipt of COTS HW and SW. Installations of HW and SW take between 2 and 6 weeks, depending on whether the site is an initial installation (requiring LAN installation) and the quantity and complexity of the configurations to be installed.

After installation, staffing and training of the maintenance and operations (M&O) staff is accomplished. M&O training occurs in conjunction with the 3-month system integration and acceptance testing.

Another key objective is the transition of the Release A Langley DAAC to Release B. This will involve installing and testing Release B custom software, COTS software and hardware in a currently operational environment. Reference the Transition to Release B Technical Paper (240-TP-010-001) for a more detailed discussion.

ECS subsystems provide mission and operations functionality for Release B. Key ECS-related mission and operations activities supported by the ECS portion of the Langley DAAC include information management, data distribution and a high level data archive.

In addition to automated support, ECS subsystems provide the capability for the ECS operations staff to perform a number of roles in support of these activities. These operational roles for the Langley DAAC are identified in Table 3.1.2-1. The table identifies the corresponding SDPS or CSMS subsystem that enables the DAAC ECS operations staff to perform a particular role/function. Detailed descriptions of these activities are captured in the ECS Operations Concept Document, Part 2: Release A (604-CD-002-003) document. The Release B SDPS/CSMS Operations Scenarios Document (605-CD-002-001) provides additional detailed scenarios for these activities.

Table 3.1.2-1. Langley Operations Support Functions

ECS DAAC Operational Roles	Capability
User Services - Support user with data expertise - Generate and maintain data interface	Data Management Subsystem
Data Ingest - Monitor electronic ingest - Handle media ingest	Ingest Subsystem
Production Planning	Planning Subsystem
Resource Planning	Planning Subsystem Systems Management Subsystem
Production Monitoring and Control	Processing Subsystem
Archive Management	Data Server Subsystem
Data Distribution - Monitor electronic distribution - Handle media distribution	Data Server Subsystem
Resource Management	Processing, Ingest, Distribution & Data Server Subsystems in coordination with Systems Management Subsystem
Science Software Integration Support	Processing Subsystem (1)
Database Maintenance	Data Management Subsystem Data Server Subsystem Application specific (2)
System and Performance Analysis	Systems Management Subsystem
Security	Systems Management Subsystem
Accounting and Billing	Systems Management Subsystem
Sustaining Engineering	Office Support Systems Management Subsystem Communication Subsystem
S/W and H/W Maintenance	Office Support Systems Management Subsystem Communication Subsystem
Configuration Management (change control)	Systems Management Subsystem
Testing, training, property management, integrated logistics support, library administration	Office Support Systems Management Subsystem Communication Subsystem

Notes:

1. Capability used to support science software integration and testing for TRMM, EOS AM-1, SAGE III, and ACRIM FOO readiness and production updates.

2. Included to ensure that the number of small DBMSs throughout the system are not explicitly excluded (e.g., Planning Subsystem has a DBMS)

3.2 Langley External Interfaces

The ECS portion of the Langley DAAC will interface with multiple entities external to the DAAC. The word "external" for the interfaces in this section connotes "external to the ECS portion of the Langley DAAC" and does not connote the less restrictive meaning "external to ECS" . The ECS subsystem-specific DID305 design documents address the interfaces generically in a series of tables supported by textual explanations. For details, the reader is referred to those documents in addition to the various Interface Control Documents (ICDs).

Figure 3.2-1 schematically illustrates the interfaces between the ECS subsystems at the Langley DAAC and the entities external to ECS (sinks and sources of data). The figure enumerates data flows which are elaborated upon in Table 3.2-1.

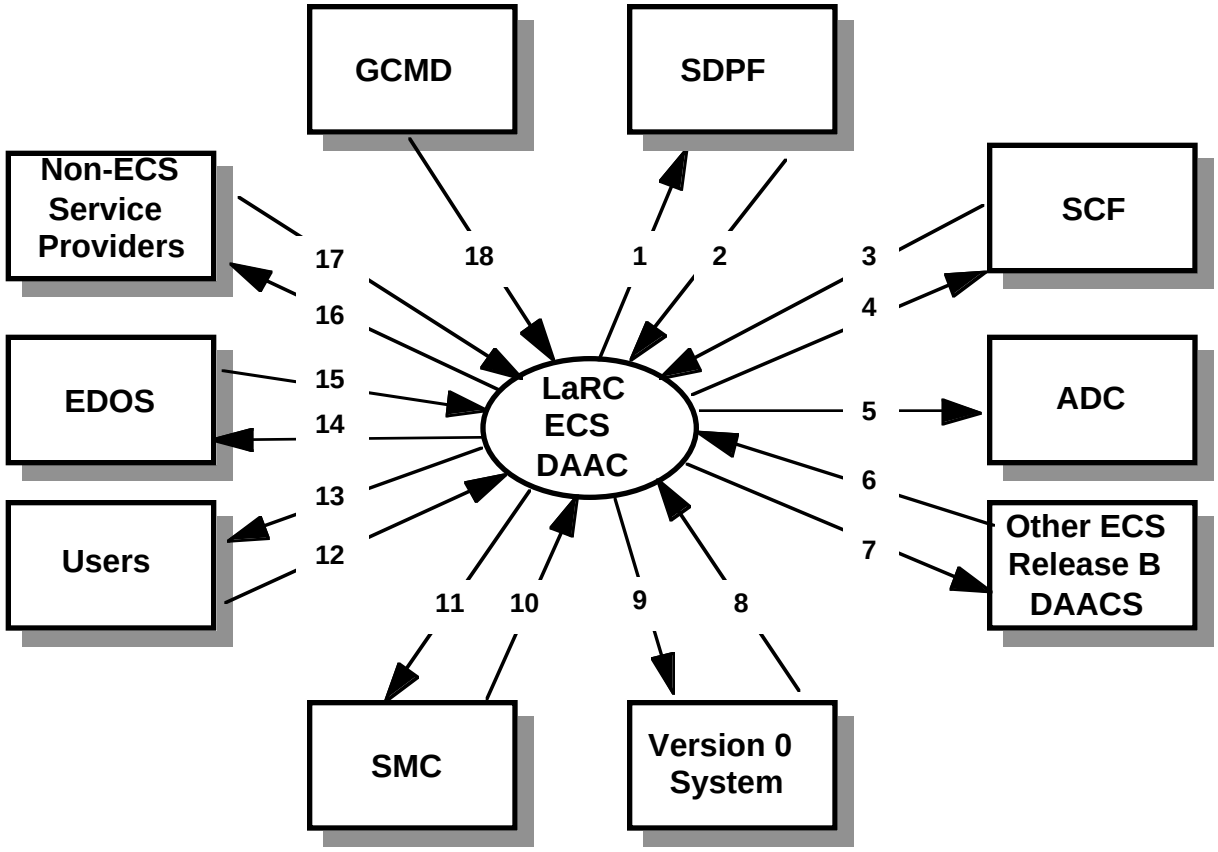

Figure 3.2-1. EDC ECS DAAC External Interfaces

A description of the external entities follows:

- SDPF - This interface supports the transmission of the TRMM CERES Level 0 data, metadata, predictive orbit and definitive orbit data from the TRMM platform and FDF and backup data and expedited data. It is almost a real time interface as there will be little delay

between receipt at the SDPF and re-transmission to the ECS portion of the Langley DAAC. Further details on this interface can be found in the TRMM SDPF Consumers ICD.

- CERES SCF–This interface is required for the CERES science software integration and testing of updates to the operational TRMM and EOS AM-1 science software. It is also required for the scientific quality assurance during TRMM and EOS AM-1 CERES data processing. CERES science software, metadata, status, quality control products, standard products, calibration data, correlative data, science software updates and documentation will traverse this interface. The CERES scientific quality assurance activity resides at both Langley and non-Langley sites.

- MISR SCF–This interface is required for the MISR science software integration and testing of the EOS AM-1 version science software and of updates to the operational EOS AM-1 science software. It is also required for the scientific quality assurance during EOS AM-1 MISR data processing. MISR science software, metadata, status, quality control products, standard products, calibration data, correlative data, science software updates and documentation will traverse this interface. This SCF is located at JPL.

- MOPITT SCF–This interface is required for the MOPITT science software integration and testing of the EOS AM-1 version science software and of updates to the operational EOS AM-1 science software. It is also required for the scientific quality assurance during EOS AM-1 MOPITT data processing. MOPITT science software, metadata, status, quality control products, standard products, calibration data, correlative data, science software updates and documentation will traverse this interface. This SCF is located at NCAR and University of Toronto.

- SAGE III SCF–This interface is required for the SAGE III science software integration and testing of the METEOR-3M Mission release science software and of updates to the operational METEOR-3M Mission science software. It is also required for the scientific quality assurance during METEOR-3M Mission SAGE III data processing. This interface is also required for the Space Station SAGE III science software integration and testing effort. SAGE III science software, metadata, status, quality control products, standard products, calibration data, correlative data, science software updates and documentation will traverse this interface. This SCF is located at LaRC.

- ACRIM SCF - This interface may be required for the ACRIM science software integration and test effort. ACRIM science software and SSI&T products may traverse this interface. This SCF may be located at JPL.

- NOAA ADC[1]–This interface has two purposes. This interface supports the access to non-EOS data sets (stored at the NOAA ADC) to satisfy ECS user queries. This interface is also required to support the ancillary data requirements for the TRMM CERES, EOS AM-1 CERES MISR, MOPITT, and the SAGE III FOO data processing. Further interface details are provided in the ICD Between ECS and the NOAA ADC (209-CD-006-005).

- Other ECS DAACs - This interface has several purposes. It supports primary ancillary data requirements of CERES, MISR, and MOPITT data products produced at the ECS portion of the Langley DAAC. For example, access to the ECS portion of the GSFC DAAC will satisfy the following CERES primary ancillary data sets: TOMS column ozone, gridded products (originally from NOAA NMC), MODIS and VIRS level 1B radiances and column aerosol data, and the TMI (water path and microwave humidity) data. This interface supports primary ancillary data requirements of the EOS AM-1 MISR data processing. For example, some MISR ancillary data requirements will be satisfied by the snow and ice cover data from the ECS portion of the NSIDC DAAC, provisional land cover from the ECS portion of the EDC DAAC, and MODIS radiances, column water vapor, cloud properties, and scene ID from the ECS portion of the GSFC DAAC. Some MOPITT ancillary data requirements will be satisfied by MODIS cloud products and temperature and moisture profile products from the ECS portion of the GSFC DAAC.

  This interface also supports primary ancillary data requirements of non-Langley DAAC instruments. For example, MODIS processing at the ECS portion of the EDC DAAC is dependent on MISR data products produced at the ECS portion of the Langley DAAC.

  Also, all Release A data products that are migrated from their respective Version 0 DAACs will be available to the users. Guide, inventory, standard products and other related information, identified in Table 3-1, will flow across this interface.

- Version 0 - This interface is required to support the primary ancillary data requirements of the TRMM CERES data processing. The TRMM CERES primary ancillary data sets that are archived at the Langley Version 0 DAAC include SAGE II and the ISCCP D1, D2, or DX products. These data sets are currently provided operationally to the Langley Version 0 DAAC from the SAGE II Project and NASA Goddard Institute for Space Studies, respectively.

  This interface is also required to support some of the primary ancillary data requirements of the EOS AM-1 MISR data processing. The EOS AM-1 MISR primary ancillary data sets that are archived at the Langley Version 0 DAAC include SAGE II aerosols and aerosol climatology data products. These data sets are currently provided operationally to the Langley Version 0 DAAC from the SAGE II Project.

  The migration of Version 0 datasets from the Langley Version 0 DAAC to the ECS portion of the Langley DAAC will occur via this interface. Selected Langley Version 0 data sets have already been identified for initial inclusion into Release A: ERBE S-4, S-4G, S-4GN, S-8, S-9, and S-10N; ISCCP DX, D1, and D2; and SAGE II Level 2 & 3 products (aerosol profile, ozone profile, water vapor profile, nitrogen dioxide profile, cloud occurrence, and monthly averaged ozone mixing ratio). Additional products, TBD, will be migrated during the release B lifetime. See the Langley DAAC Data Migration Plan for details. These data sets will be among the first maintained within the ECS Data Server paradigm, as described in Section 3.3.

  This interface to the ECS portion of the Langley DAAC supports access, using the Version 0 System IMS, to the Version 0 holdings that have not currently migrated to ECS.

This interface is also used to support the interoperability interface to provide cross-DAAC access.

- SMC - This interface provides the capability for the ECS portion of the Langley DAAC to receive performance information, policy data and user registration information. Policy data includes that established by the ESDIS project. The ECS portion of the Langley DAAC sends its system performance and status reports to SMC as part of this interface.

- Users – This interface is the mechanism for user community access to ECS data and services. It is the mechanism by which advertisements, user registration, order and product status, desktop object manipulations, and command languages capabilities are utilized.

- EDOS - This interface supports the transmission of the EOS AM-1 DAAC-to-EDOS Data Set (DEDS) from LaRC to EDOS, and Product Data Sets (PDSs) and Expedited Data Sets (EDS)s from EDOS to LARC. The latter include Level 0, metadata, orbit/attitude and expedited data from the EOS-AM-1 CERES, MISR and MOPITT instruments. It is almost a real time interface, as there will be little delay between receipt at EDOS and re-transmission to the ECS portion of the Langley DAAC.

- Non-ECS Service Providers - This interface is required for specialized users who use ECS data to provide and advertise value-added services. These providers include commercial, institutional, or other government agencies, as well as IPs, SCFs, and ADCs.

- GCMD - The Global Change Master Directory (GCMD) is a multi-disciplinary database of information about data holdings of potential interest to the scientific research community. It contains high level descriptions of data set holdings of various agencies and institutions. It also contains supplementary descriptions about these data centers, as well as scientific campaigns and projects, sources (spacecraft, platforms), and sensors. This interface will allow any ECS DAAC to import directory level information from the GCMD via GCMD export files and generate ECS data product advertisements.

The following external entities do not directly interface with the ECS portion of the Langley DAAC; however, these entities are of interest to the DAAC:

- FDF – TRMM satellite refined orbit/attitude data from the FDF is required for the TRMM CERES processing. This data is provided by FDF to SDPF where the FDF refined orbit/attitude data file is attached to the CERES level 0 data and provided to the ECS portion of the Langley DAAC via the SDPF interface.

For the following table, the phrase "as required" refers to the requirements agreed upon by the source and the destination via IRDs, ICDs, etc.

Table 3.2-1. LaRC External Interfaces (1 of 6)

Flow No.	Source	Destination	Data Types	Data Volume	Frequency
1	Ingest	SDPF	L0 data	medium	as requested by Processing for L0 reprocessing
2	SDPF	Ingest	L0 Data	Instrument Specific (CERES, LIS)	daily
2	SDPF	Ingest	Expedited Data	medium	three times a day
2	SDPF	Ingest	Predictive Orbit Data	medium	daily
2	SDPF	Ingest	Definitive Orbit Data	medium	daily
2	SDPF	Ingest	Back-up Data	medium	as required
3	SCF	Ingest	Metadata/updates	low	as required
3	SCF	Ingest	Documents	low	as required
3	SCF	Ingest	Algorithms/Updates	medium	as required
3	SCF	CSS (DAAC Ops via email)	Test Reviews by SCF	low	as required
3	SCF	CSS (DAAC Ops via email)	Request for Resource Usage	low	as required
3	SCF	CSS (DAAC Ops via email)	Reprocessing Request	low	as required
3	SCF	Data Server	QA Data request	low	as required
3	SCF	Data Server	QA Data Subscription	low	as required
4	Data Server	SCF	Status	low	as required
4	Data Server	SCF	Metadata/updates	low	as required
4	Data Server	SCF	Calibration data	medium	as required
4	Data Server	SCF	Correlative data	medium	as required
4	Data Server	SCF	Documents	low	as required

Table 3.2-1. LaRC External Interfaces (2 of 6)

Flow No.	Source	Destination	Data Types	Data Volume	Frequency
4	Data Server	SCF	Algorithms/updates	medium	as required
4	Data Server	SCF	Standard Products	medium	daily as required for QA
4	CSS (DAAC Ops via email, EDHS)	SCF	Toolkit Delivery and Update Package	low	as required
4	CSS (DAAC Ops via email, kftp)	SCF	Test Results, QA, and Production History Data	low	as required
4	CSS (DAAC Ops via email, kftp)	SCF	Resource Usage	low	as required
4	CSS (DAAC Ops via email, kftp)	SCF	Status	low	as required
5	NOAA ADC	data server	Advertising Information	low	as required
5	NOAA ADC	Client/Data Server	Dependent Valids Update	low	as required
5	NOAA ADC	Client	User Authentication Requests	low	as required
5	NOAA ADC	Client	Guide Query Results	low	as required
5	NOAA ADC	Client	Inventory Query Results	low	as required
5	NOAA ADC	Client	Cost Estimates	low	as required
5	NOAA ADC	Client	Product Delivery Status	low	as required
5	NOAA ADC	ECS (DAAC operations)	Schedule Adjudication via telephone	N/A	as required
6	Other ECS DAACs	Ingest	Ancillary Data	high	as required

Table 3.2-1. LaRC External Interfaces (3 of 6)

Flow No.	Source	Destination	Data Types	Data Volume	Frequency
6	Other ECS DAACs	Ingest	Correlative Data	high	as required
6	Other ECS DAACs	Ingest	Calibration Data	medium	as required
6	Other ECS DAACs	Ingest	QA Data	medium	as required
6	Other ECS DAACs	Interoperability	Advertisements	medium	as required
6	Other ECS DAACs	Data Server	Result Sets	medium	as required
6	Other ECS DAACs	Client	Product Results	medium	as required
7	Data Server	Other ECS DAACs	Standard Products	high	as required
7	Data Server	Other ECS DAACs	Metadata	medium-high	as required
7	Data Server	Other ECS DAACs	Ancillary Data	high	as required
7	Data Server	Other ECS DAACs	Correlative Data	high	as required
7	Data Server	Other ECS DAACs	Calibration Data	high	as required
7	Data Server	Other ECS DAACs	Documents	medium	as required
7	Data Server	Other ECS DAACs	Orbit/Attitude Data	medium	as required
7	Data Server	Other ECS DAACs	Data Availability Schedules	medium	as required
7	Data Server	Other ECS DAACs	Algorithms	high	as required
7	Data Server	Other ECS DAACs	L0 Data	high	as required
7	Data Server	Other ECS DAACs	Expedited or Validated Data	medium	as required
7	Data Server	Other ECS DAACs	QA Data	medium	as required
7	Data Mgmt	Other ECS DAACs	Data Dictionary	medium	as required

Table 3.2-1. LaRC External Interfaces (4 of 6)

Flow No.	Source	Destination	Data Types	Data Volume	Frequency
7	Interoperability	Other ECS DAACs	Advertisements	medium	as required
7	Client	Other ECS DAACs	Product Requests	medium	as required
8	Version 0 System	Data Server	Inventory	low	as required
8	Version 0 System	Data Server	Guide	low	as required
8	Version 0 System	Data Server	Browse data	medium	as required
8	Version 0 System	Data Server	Dependent Valids	low	as required
8	Version 0 System	Data Management	V0 Directory search request	low	as requested
8	Version 0 System	Data Management	V0 Inventory search request	low	as requested
8	Version 0 System	Data Management	V0 browse request	low	as requested
8	Version 0 System	Data Management	V0 product order request	low	frequency dependent on user input
8	Version 0 System	Ingest	Migration Data	high	varies depending on migration strategy
9	Data Mgmt	Version 0 System	V0 Browse Result	low-medium	in response to ECS browse result
9	Data Mgmt	Version 0	V0 inventory result set	low-high	in response to ECS inventory result request
9	Data Mgmt	Version 0	V0 directory search result set	low-high	in response to ECS request
9	Data Mgmt	Version 0 System	V0 product order response	low	in response to product request result
9	Data Server	Version 0 System	Result Sets	medium-high	in response to request
9	Data Server	Version 0 System	Session Mgmt responses	low	in response to request
9	Data Server	Version 0 System	Product Request Status	low	as required
10	SMC	MSS	Policies	low	as required
10	SMC	MSS	Conflict Resolution	low	as required

Table 3.2-1. LaRC External Interfaces (5 of 6)

Flow No.	Source	Destination	Data Types	Data Volume	Frequency
10	SMC	MSS	Procedures	low	as required
10	SMC	MSS	Directives	low	as required
11	MSS	SMC	Conflict Resolution Request	low	as required
11	MSS	SMC	Status	low	as required
11	MSS	SMC	Performance	low	as required
12	Users	Client	User registration information	low	as requested
12	Users	Client	User login information	low	as requested
12	Users	Client	Search requests	low	as requested
12	Users	Client	Product requests	low	as requested
12	Users	Client	Acquisition requests	low	as requested
12	Users	Client	Desktop manipulate commands	low	as supplied by user
12	Users	Client	Configuration/Profile information	low	as supplied by user
12	Users	Client	Data manipulate requests	low	as requested
12	Users	Client	Command language request	low	as requested
12	Users	Client	Advertisements, Software, and Documents	low	as supplied by user
12	Users	Ingest	User Methods	medium	as required
12	Users	Ingest	Ingest Status Requests	low	as required
13	Data Server	Users	Metadata	low	as requested
13	Data Server	Users	Documents	low	as requested
13	Data Server	Users	Data Products	medium	as requested
13	Data Server	Users	Browse Products	medium	as required
13	Data Server	Users	Product Request Status	low	as requested
13	Data Server	Users	Schedules	low	as requested

Table 3.2-1. LaRC External Interfaces (6 of 6)

Flow No.	Source	Destination	Data Types	Data Volume	Frequency
13	Client	Users	Results Set	medium	as requested
13	Client	Users	Application user interfaces	low	as requested
13	Client	Users	Formatted data	medium	as requested
13	Client	Users	Desktop Objects	low	as requested
13	Client	Users	Advertisement and Software	low	as requested
13	Client	Users	Error and Status information	low	as available
13	Ingest	Users	Ingest Status	low	as requested
14	Ingest	EDOS	PDS Delivery Verification	low	as requested
14	Ingest	EDOS	EDS Delivery Verification	low	as requested
14	Data Server	EDOS	DEDS	high	rare
15	EDOS	Ingest	PDSs (L0 Data)	high	several times a day
15	EDOS	Ingest	PDS Delivery Record	low	several times a day
15	EDOS	Ingest	EDSs (Expedited Data)	high	several times a day
15	EDOS	Ingest	EDS Delivery Record	low	several times a day
16	Interoperability	Non-ECS Service Providers	Notifications	low	in response to subscriptions
17	Non-ECS Service Providers	Interoperability	Advertisements	low-medium	as required
17	Non-ECS Service Providers	Interoperability	Subscriptions	low	as required
18	GCMD	Interoperability	Advertisements	low-medium	as required

In the table, where an exact number is unavailable, the data volume is estimated as low (less than 1 MB), medium (between 1 MB and 1 GB), or high (greater than 1 GB) per use defined in the frequency column .

3.3 Computer Software Component Analysis

The ECS software subsystems are described in detail in the ECS Subsystem-specific DID 305 documents (refer to Section 2 for a list of DID 305 documents). This section provides a brief overview description of each of the subsystems, then, as part of the analysis, addresses the CSCIs for each subsystem, focusing upon those CSCIs that are specific to the ECS portion of the Langley DAAC. For the most part, the software is the same for all ECS DAACs. However, the content of

the databases and schema constructions may differ. In addition, the purchase of different OTS packages for the DAACs may be required.

3.3.1 Software Subsystem Overview

The ECS software subsystems applicable to the Langley DAAC are described in detail in the ECS Subsystem-specific DID305 documents. This section provides a brief overview description of these subsystems.

Client Subsystem (CLS): This software consists of graphical user interface (GUI) programs, tools for viewing and/or manipulating the various kinds of ECS data (e.g., images, documents, tables) and libraries representing the client application program interface (API) of ECS services. The client subsystem components will available to users for installation on their workstations, and will also be deployed on workstations within the DAAC in support of normal operations, including User Services support.

Interoperability Subsystem (IOS): The interoperability subsystem is an advertising service. It maintains a database of information about the services and data offered by ECS, and allows users to search through this database to locate services and data that may be of interest to them. The advertising service will be implemented as a Web server application with a DBMS back-end.

Data Management Subsystem (DMS): This subsystem includes functions which provide uniform access to descriptions of the data and the data elements offered by the ECS repositories, and provides a bi-directional gateway between ECS and Version 0. This subsystem also includes distributed search and retrieval functions and corresponding site interfaces.

Data Server Subsystem (DSS): This subsystem provides the physical storage access and management functions for the ECS earth science data repositories. Other subsystems can access DSS directly, or via the data management subsystem (if they need assistance with searches across several of these repositories). The subsystem also includes the capabilities needed to distribute bulk data via electronic file transfer or physical media. Other components include, for example, administrative software to manage the subsystem resources and perform data administration functions (e.g., to maintain the database schema); and data distribution software, e.g., for media handling and format conversions. The main components of this subsystem are:

- database management system - uses an off-the-shelf DBMS (Illustra) to manage its earth science data and implement spatial searching, as well as for the more traditional types of data (e.g., system administrative and operational data). It will use a document management system to provide storage and information retrieval for guide documents, scientific articles, and other types of document data.
- file storage management systems - used to provide archival and staging storage for large volumes of data. Provides hierarchical storage support and device/media independence to the remainder of DSS and ECS.

- data type libraries - they will implement functionality of earth science and related data that is unique and not available off the shelf (e.g., spatial search algorithms and translations among coordinate systems). The libraries will interface with the data storage facilities, i.e., the database and file storage management systems.

Ingest Subsystem (INS): This subsystem deals with the initial reception of all data received at an ECS facility, and triggers subsequent archiving and processing of the data. Given the variety of possible data formats and structures, each external interface, and each ad-hoc ingest task, may have unique aspects. Therefore, the ingest subsystem is organized into a collection of software components (e.g., ingest management software, translation tools, media handling software) from which those components required in a specific situation can be readily configured. The resultant configuration is called an ingest client. Ingest clients can operate on a continuous basis to serve a routine external interface; or they may exist only for the duration of a specific ad-hoc ingest task.

Data Processing Subsystem (DPS): The main components of the data processing subsystem, the science software, will be provided by the science teams. The data processing subsystem will provide the necessary hardware resources, as well as software for queueing, dispatching, and managing the execution of the science software in an environment which eventually will be highly distributed and consist of heterogeneous computing platforms. The DPS also interacts with the DSS to cause the staging and de-staging of data resources in synchronization with processing requirements.

Planning Subsystem (PLS): This subsystem provides the functions needed to pre-plan routine data processing, schedule on-demand processing, and dispatch and manage processing requests. PLS provides access to the data production schedules at each site, and provides management functions to the operations and science users for handling deviations from the schedule.

System Management Subsystem (MSS): The System Management Subsystem (MSS) provides enterprise management (network, system and application management) for all ECS resources: commercial hardware (including computers, peripherals, and network routing devices), commercial software, and custom applications. Enterprise management reduces overall development and equipment costs, improves operational robustness, and promotes compatibility with evolving industry and government standards. Consistent with current trends in industry, the MSS thus manages both ECS's network resources per EBnet requirements and ECS's host/application resources per SMC requirements. Additionally, MSS also supports many requirements allocated to SDPS and FOS for management data collection and analysis/distribution. The MSS allocates services to both the system-wide and local levels. With few exceptions, the management services will be fully decentralized. No single point of failure exists which would preclude user access. In principle, every service is distributed unless there is an overriding reason for it to be centralized. MSS has two key specializations: Enterprise Monitoring and Coordination Services and Local System Management Services.

Communications Subsystem (CSS): The CSS services include Object Services, Distributed Object Framework (DOF) and Common Facility Services. Support in this

subsystem area is provided for peer-to-peer, advanced distributed, messaging, management, and event-handling communications facilities. These services typically appear on communicating end-systems across an internetwork and are not layered, but are hierarchical in nature. Additionally, services to support communicating entities are provided, including directory, security, time, and other ancillary services. The services of the Communications Subsystem are functionally dependent on the services of the Internetworking Subsystem. The services of the common facility, object, and DOF are the fundamental set of interfaces for all management and user access (i.e., pull) domain services.

Internetworking Subsystem (ISS): The Internetworking Subsystem provides for the transfer of data transparently within the DAACs, SMC and EOC, and for providing interfaces between these components and external networks. ECS interfaces with external systems and DAAC-to-DAAC communications are provided by the EOSDIS Backbone Network (EBnet). EBnet's primary function is to transfer data between DAACs, including both product data and inter-DAAC queries and metadata responses. Other networks, such as NSI, will provide wide-area services to ECS. In addition, "Campus" networks, which form the existing networking infrastructure at the ECS locations, will provide connectivity to EOSDIS components such as SCFs and ISTs.

3.3.2 Software Subsystem Analysis Summary

The subsystems that comprise SDPS and CSMS have already been described in detail in companion CDR documents, and have been summarized above. This section addresses the CSCIs within each subsystem and identifies the ECS portion of the Langley DAAC. Generally, the software is the same for all ECS DAACs. The content of the databases and schema constructions may differ. In the case of OTS packages the possibility arises for the purchase of different versions for different DAAC hardware, but this will be minimal for Release B. In this section, each of the subsystems will be addressed in a general manner to point out whether or not there are any Langley DAAC-specific portions.

- Client Subsystem - The client software will not have any DAAC-specific portions except for the possibility of different versions of OTS packages due to different types of hardware. Since the services offered by the client are required by operations, user services, and systems administrators, the ECS portion of the Langley DAAC will have clients installed on several different ECS furnished workstations. In addition the Langley V0 DAAC may desire the client on some of their existing workstations to provide additional user access.

- Data Server -The software components of the Data Server Subsystem (DSS) are largely the same for all Data Servers, at all DAACs. The two basic areas in which the Data Server Subsystem software will vary from DAAC to DAAC are configuration and special components.

  Data Server software is designed to be highly configurable in order to allow a wide variety of DAAC-unique policy implementations. These unique configurations will enable the data server software installations to vary behavior, meeting the DAAC-specific needs.

Examples of configuration parameters include number of concurrent connections, number of requests per client, inactivity timeout period, and allocation of software components to hardware.

Another facet of the DSS software that supports the specific DAAC capabilities is in which actual components are installed at the ECS portion of the DAAC. These opportunities for DAAC specificity are driven by the types of distribution available to the DAAC's Data Server clients, and in the types of data (and their data type services) available. There will be portions of the Data Server software specific to the Langley DAAC that are used to add special distribution devices.

However, the primary portion of the DSS software that will be specific to the Langley DAAC will be the specific data types supported at the DAAC. These software components are a portion of the Science Data Server (SDSRV) CSCI. The SDSRV is designed to allow complete flexibility in the data types (specifically, Earth Science Data Types, or ESDTs) that offer their services via the Data Server. Examples of services offered by ESDTs include Insert, Acquire, Browse and SpatialSubset. These data types are organized by separate CSCs, generally one per source instrument (i.e., CERES, MISR, MOPITT, and SAGE III). Additional information about the Langley instruments can be found in the February 1996 ECS Technical Baseline.

- Data Management - None of the data management software will be unique to the ECS portion of the Langley DAAC. The V0 Gateway (GTWAY) will interface with the data servers at each site. Local and cross-DAAC searches on V0 DAACs' data holdings are provided via capabilities resulting from integrating the components of the V0 System IMS into ECS.

- Ingest - The software portions for ingest at the ECS portion of the Langley DAAC may differ from those of other ECS DAACs because of dataset dependencies and differences related to non-homogeneous computer hardware across the ECS portions of the DAACs. Data ingest procedures must match the peculiarities of the ingested data sets. Several types of ingest clients are described in the Ingest companion document. The primary client was based on an approach used by the TRMM SDPF . Since the SDPF is the source of TRMM CERES L0 data, that type of ingest client represents one of the set to be implemented for the ECS portion of the Langley DAAC.

- Interoperability - There are no DAAC-specific portions of the Interoperability Subsystem.

- Production Planning - The actual development code and OTS packages required for the ECS portion of the Langley DAAC will be generic, however there will be a considerable amount of configuration and database information that will be specific. Scripts, while they may use the same language, will be different and will trigger different responses in response to faults and other error conditions .

- Data Processing - Due to dataset characteristics there will be some specific software for the ECS portion of the Langley DAAC in the area of Science Data Processing. In addition, as with the data server and ingest subsystems, differences in hardware types, driven by algorithm requirements, will also result in some differences in software.

- Communications Subsystem - There are no DAAC-specific portions of this subsystem.
- Systems Management - This subsystem is composed of a variety of management applications, providing services such as fault, performance, security and accountability management for ECS networks, hosts, and applications. Two tiers of "view" (domain of management service interface) are provided by the applications in this subsystem. Only the local management view is provided at the ECS portion of the Langley DAAC. The MSS capabilities will be available to the non-ECS portions of the Langley DAAC through an ECS-provided API.
- Internetworking Subsystem - There are no DAAC-specific portions of this subsystem.

Table 3.3.2-1 lists the ECS subsystems and associated CSCIs and CSCs. For each CSC, there is an indication of the type of component. As defined in the DID 305 subsystem-specific documents, type indicates whether the component is custom developed (DEV), off the shelf (OTS), or a wrapper (WRP) that encapsulates OTS, or a combination of these types. In addition, each CSC type can be augmented with a "reuse" tag, which means that the CSC is reused from another subsystem's CSC. The Use column indicates whether a generic-for-all-DAACs (Gnrc) form of the CSC is implemented, or specific (Spfc) tailoring or use is required at a DAAC. The Notes column is included to comment about the characteristics of the system, data, and/or software that makes the CSC specific, as well as to provide any additional information about the generic CSCs. The OTS products are also listed in this column.

Table 3.3.2-1. Langley Components Analysis (1 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Client	DESKT	Desktop Manager	DEV	Gnrc	
Client	WKBCH	Comment/Survey Tool	OTS DEV	Gnrc	WWW Browser
Client	WKBCH	Data Acquisition Request Tool	DEV	Gnrc	
Client	WKBCH	Data Dictionary Tool	DEV	Gnrc	
Client	WKBCH	Document Search Tool	OTS	Gnrc	
Client	WKBCH	Earth Science Search Tool	DEV	Gnrc	
Client	WKBCH	E-mailer Tool	OTS	Reuse	CSS-provided
Client	WKBCH	Hypertext Authoring Tool	OTS	Gnrc	MS Office/TBD Public Domain
Client	WKBCH	Hypertext Viewer	OTS	Gnrc	WWW Browser
Client	WKBCH	Logger/Reviewer Tool	DEV	Gnrc	
Client	WKBCH	News Reader Tool	OTS	Reuse	CSS-provided
Client	WKBCH	Product Request Tool	DEV	Gnrc	
Client	WKBCH	Session Management Tool	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (2 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Client	WKBCH	User Preferences Tool	DEV	Gnrc	
Client	WKBCH	User Registration Tool	DEV	Gnrc	
Client	WKBCH	Visualization Tool	DEV	Gnrc	
Communication	DCCI	Bulletin Board	OTS	Reuse	CSS-provided
Communication	DCCI	Directory/Naming Services	OTS/ DEV	Gnrc	OODCE
Communication	DCCI	Distributed File System (DFS)	OTS	Gnrc	DCE
Communication	DCCI	DOF Services	OTS	Gnrc	OODCE
Communication	DCCI	Electronic Mail Services	OTS/ DEV	Gnrc	native operating system
Communication	DCCI	Event Logger Services	OTS/ DEV	Gnrc	DCE
Communication	DCCI	File Access Services	OTS/ DEV	Gnrc	ftp, kftp, DCE
Communication	DCCI	Life Cycle Services	OTS/ DEV	Gnrc	OODCE
Communication	DCCI	Message Passing Services	OTS/ DEV	Gnrc	Developed with OODCE
Communication	DCCI	Security Services	OTS/ DEV	Gnrc	OODCE
Communication	DCCI	Thread Services	OTS	Gnrc	OODCE
Communication	DCCI	Time Services	OTS/ DEV	Gnrc	OODCE
Communication	DCCI	Virtual Terminal Services	OTS	Gnrc	native operating system
Data Management	DDICT	Client Library	DEV	Gnrc	
Data Management	DDICT	Configuration/Setup	DEV	Gnrc	
Data Management	DDICT	DBMS Server	OTS	Gnrc	Sybase DBMS
Data Management	DDICT	Maintenance Tool	DEV	Gnrc	
Data Management	DDICT	Persistent Data	DEV	Gnrc	
Data Management	DDICT	Request Processing	DEV	Gnrc	
Data Management	DDICT	Server	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (3 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Data Management	DIMGR	Configuration/Setup	DEV	Gnrc	
Data Management	DIMGR	Server	DEV	Gnrc	
Data Management	GTWAY	Configuration/Setup	DEV	Gnrc	
Data Management	GTWAY	Server	DEV	Gnrc	
Data Management	GTWAY	V0 Back End	OTS	Gnrc	From V0
Data Management	GTWAY	V0 Client Interface	DEV	Gnrc	
Data Management	GTWAY	V0 External Interface	DEV	Gnrc	
Data Management	GTWAY	V0 Front End	OTS	Gnrc	From V0
Data Management	LIMGR	Client Library	DEV	Gnrc	
Data Management	LIMGR	Configuration/Setup	DEV	Gnrc	
Data Management	LIMGR	Database Interface	OTS	Gnrc	RogueWave DBTools
Data Management	LIMGR	External Interface	DEV	Gnrc	
Data Management	LIMGR	Mapping Layer	DEV	Gnrc	
Data Management	LIMGR	Request Processing	DEV	Gnrc	
Data Management	LIMGR	Server	DEV	Gnrc	
Data Processing	AITTL	Binary File Comparison Utility	DEV	Gnrc	
Data Processing	AITTL	Code Analysis Tools	OTS	Spfc	CASEVision SPARCWorks Specific for science software language.
Data Processing	AITTL	Data Visualization Tools	OTS DEV	Gnrc	IDL
Data Processing	AITTL	Documentation Viewing Tools	OTS	Gnrc	SoftWindows/MS Office Ghostview
Data Processing	AITTL	ECS HDF Visualization Tools	DEV	Gnrc	Reused from Client subsystem, WKBCH CSCI, Data Visualization (EOSView) CSC

Table 3.3.2-1. Langley Components Analysis (4 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Data Processing	AITTL	HDF File Comparison Utility	DEV OTS	Gnrc	Custom IDL program
Data Processing	AITTL	PGE Processing GUI	DEV	Gnrc	
Data Processing	AITTL	PGE Registration GUI	DEV	Gnrc	
Data Processing	AITTL	Product Metadata Display Tool	Reuse DEV	Gnrc	Reused from Data Processing subsystem, AITTL CSCI, HDF File Comparison Utility CSC
Data Processing	AITTL	Profiling Tools	OTS	Spfc	CASEVision Specific for science software language.
Data Processing	AITTL	Report Generation Tools	OTS/ DEV	Gnrc	OTS: SoftWindows/MS Office, Dev: SSI&T manager
Data Processing	AITTL	SDP Toolkit-related Tools	DEV	Gnrc	
Data Processing	AITTL	SSAP Processing GUI	DEV	Gnrc	
Data Processing	AITTL	Standards Checkers	OTS/ DEV	Spfc	FORCHECK for Fortran 77; otherwise, native compilers and UNIX lint. Specific for science software language.
Data Processing	AITTL	Update Data Server GUI	DEV	Gnrc	
Data Processing	AITTL	Update PGE Database GUI	DEV	Gnrc	
Data Processing	PRONG	COTS	OTS	Gnrc	AutoSys and AutoXpert
Data Processing	PRONG	COTS Management	DEV	Gnrc	
Data Processing	PRONG	Data Management	DEV	Gnrc	
Data Processing	PRONG	Data Pre-Processing	DEV	Spfc	Specific based on uniqueness of ancillary data products
Data Processing	PRONG	PGE Execution Management	DEV	Gnrc	
Data Processing	PRONG	Quality Assurance Monitor Interface	DEV OTS	Spfc	Specific based on uniqueness of Science software and SCF-DAAC relationship
Data Processing	PRONG	Resource Management	DEV	Gnrc	
Data Processing	SDPTK	Ancillary Data Access	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (5 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Data Processing	SDPTK	Celestial Body Position	DEV	Gnrc	
Data Processing	SDPTK	Constant and Unit Conversions	DEV OTS	Gnrc	UDUNITS Freeware
Data Processing	SDPTK	Coordinate System Conversion	DEV	Gnrc	
Data Processing	SDPTK	EOS-HDF	DEV	Gnrc	
Data Processing	SDPTK	Error/Status Handling	DEV	Gnrc	
Data Processing	SDPTK	Geo Coordinate Transformation	DEV OTS	Gnrc	GCTP (USGS)
Data Processing	SDPTK	Graphics Library	OTS	Gnrc	IDL
Data Processing	SDPTK	File Input/Output Tools	DEV OTS	Gnrc	NCSA HDF
Data Processing	SDPTK	Math Package (IMSL)	OTS	Gnrc	IMSL
Data Processing	SDPTK	Memory Management	DEV	Gnrc	
Data Processing	SDPTK	Metadata Access	DEV	Gnrc	
Data Processing	SDPTK	Process Control	DEV	Gnrc	
Data Processing	SDPTK	Spacecraft Ephemeris and Attitude Access	DEV	Gnrc	
Data Processing	SDPTK	Time Date Conversion	DEV	Gnrc	
Data Server	DDIST	Distribution Client Interface	DEV	Gnrc	
Data Server	DDIST	Distribution Products	DEV	Gnrc	
Data Server	DDIST	Distribution Request Management	DEV	Gnrc	
Data Server	DDSRV	DDSRV	DEV/ OTS	Gnrc	Rogue Wave class libraries
Data Server	DDSRV	DDSRV Client	DEV/ OTS	Gnrc	HTTP libraries
Data Server	DDSRV	DDSRV CSDT	DEV/ OTS	Gnrc	Rogue Wave class libraries
Data Server	DDSRV	DDSRV ESDT	DEV/ OTS	Gnrc	Rogue Wave class libraries
Data Server	DDSRV	DDSRV Search Engine	DEV	Gnrc	Text Search Indexor and HTTP server
Data Server	DDSRV	DDSRV Server	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (6 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Data Server	SDSRV	Administration and Operations	DEV	Gnrc	
Data Server	SDSRV	ACRIM	DEV	Spfc	ESDTs - ACRIM
Data Server	SDSRV	CERES	DEV	Spfc	ESDTs - CERES
Data Server	SDSRV	MISR	DEV	Spfc	ESDTs - MISR
Data Server	SDSRC	MODIS	DEV	Spfc	ESDTs - MODIS
Data Server	SDSRV	MOPITT	DEV	Spfc	ESDTs - MOPITT
Data Server	SDSRV	SAGEIII	DEV	Spfc	ESDTs - SAGE III
Data Server	SDSRV	TES	DEV	Spfc	ESDTs - TES
Data Server	SDSRV	Client	DEV/ OTS	Gnrc	Rogue Wave class libraries and OODCE
Data Server	SDSRV	Configuration and Startup	DEV	Gnrc	
Data Server	SDSRV	CSDT	DEV OTS	Gnrc	HDF-EOS
Data Server	SDSRV	DB Wrappers	DEV OTS	Gnrc	Illustra DBMS
Data Server	SDSRV	Descriptors	DEV	Gnrc	Rogue Wave class libraries
Data Server	SDSRV	General ESDT	DEV	Gnrc	
Data Server	SDSRV	Global	DEV/ OTS	Gnrc	Rogue Wave class libraries
Data Server	SDSRV	GUI	DEV OTS	Gnrc	
Data Server	SDSRV	Metadata	DEV/ Wrpr	Gnrc	Illustra DBMS API
Data Server	SDSRV	Non-Product Science ESDT	DEV	Gnrc	
Data Server	SDSRV	Non-Science ESDT	DEV	Gnrc	
Data Server	SDSRV	Server	DEV/ OTS	Gnrc	Rogue Wave class libraries
Data Server	SDSRV	Subscriptions	DEV/ OTS	Gnrc	Rogue Wave class libraries
Data Server	STMGT	Data Storage	DEV/ OTS	Gnrc	AMASS File Storage Management System
Data Server	STMGT	File	DEV	Gnrc	
Data Server	STMGT	Peripherals	DEV	Gnrc	This CSC encapsulates the CSS-supplied API which supports the OTS FTP product.
Data Server	STMGT	Resource Management	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (7 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Data Server	STMGT	Service Clients	DEV/ OTS	Gnrc	CSC encapsulates the AMASS OTS product (Data Server subsystem, STMGT CSCI, Data Storage CSC) . Also Rogue Wave class libraries.
Ingest	INGST	Client	Reuse	Gnrc	
Ingest	INGST	CERES	Reuse	Spfc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	MISR	Reuse	Spfc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	MOPITT	Reuse	Spfc	Reused from Data Server subsystem SDSRV CSCI
Ingest	INGST	SAGEIII	Reuse	Spfc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Configuration/ Startup	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	CSDT	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Data Storage	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI, CSC and STMGT CSCI
Ingest	INGST	DB Wrappers	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Descriptors	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Distribution Client Interface	Reuse	Gnrc	Reused from Data Server subsystem, DDIST CSCI
Ingest	INGST	Distribution Products	Reuse	Gnrc	Reused from Data Server subsystem, DDIST CSCI
Ingest	INGST	Distribution Request Management	Reuse	Gnrc	Reused from Data Server subsystem, DDIST CSCI
Ingest	INGST	File	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	General ESDT	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Global	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	GUI	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI

Table 3.3.2-1. Langley Components Analysis (8 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Ingest	INGST	Ingest Administration Data	DEV	Gnrc	
Ingest	INGST	Ingest Data Preprocessing	DEV	Spfc	Specific based on uniqueness of ingested data and preprocessing requirements.
Ingest	INGST	Ingest Data Transfer	DEV	Gnrc	
Ingest	INGST	Ingest DBMS	OTS	Gnrc	Sybase DBMS
Ingest	INGST	Ingest Request Processing	DEV	Gnrc	
Ingest	INGST	Ingest Session Manager	DEV	Gnrc	
Ingest	INGST	Metadata	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Non-Product Science ESDTs	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Non-Science ESDTs	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Operator Ingest Interfaces	DEV	Gnrc	
Ingest	INGST	Polling Ingest Client Interface	DEV	Gnrc	
Ingest	INGST	Resource Management	Reuse	Gnrc	Reused from Data Server subsystem, STMGT CSCI
Ingest	INGST	Server	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Service Clients	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	Subscriptions	Reuse	Gnrc	Reused from Data Server subsystem, SDSRV CSCI
Ingest	INGST	User Network Ingest Interface	DEV	Gnrc	
Ingest	INGST	Viewing Tools	Reuse	Gnrc	Reused from Client subsystem, WKBCH CSCI, Data Visualization (EOSView) CSC
Interoperability	ADSRV	AdvNavigationServer	OTS	Gnrc	HTTP server
Interoperability	ADSRV	Client Library	DEV	Gnrc	
Interoperability	ADSRV	Core Library	DEV	Gnrc	
Interoperability	ADSRV	HTML Framework	DEV	Gnrc	
Interoperability	ADSRV	HTML Interfaces	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (9 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Interoperability	ADSRV	Installer	DEV	Gnrc	
Interoperability	ADSRV	Persistent Object Framework	DEV	Gnrc	
Internetworking	INCI	Datalink/Physical	OTS	Gnrc	firmware, vendor-supplied with hardware
Management	MACI	Application MIB	DEV	Gnrc	
Management	MACI	ECS Subagent	DEV	Gnrc	
Management	MACI	Encapsulator for non-Peer Agent	OTS/ DEV	Gnrc	OptiMate
Management	MACI	Extensible SNMP Master Agent	OTS/ DEV	Gnrc	Peer Network's agent, along with its toolkit for Dev
Management	MACI	Instrumentation Class Library	DEV	Gnrc	
Management	MACI	Management Agent Services	OTS/ DEV	Gnrc	Peer and Tivoli/Sentry
Management	MACI	Proxy Agent	DEV	Gnrc	
Management	MACI	SNMP Manager's Deputy	DEV	Gnrc	
Management	MCI	Accountability	DEV	Gnrc	
Management	MCI	Application Management	DEV	Gnrc	
Management	MCI	Automatic Actions	DEV	Gnrc	
Management	MCI	Billing and Accounting Management	OTS/ DEV	Gnrc	Selection in progress
Management	MCI	DCE Cell Management	OTS	Gnrc	HP Account Manager Tool
Management	MCI	Diagnostic Tests	OTS	Gnrc	vendor-supplied with hardware
Management	MCI	Fault Management	OTS/ DEV	Gnrc	Tivoli and HP OpenView
Management	MCI	Management Data Access	DEV	Gnrc	
Management	MCI	Management Data Access User Interface	DEV	Gnrc	
Management	MCI	Management Framework	OTS	Gnrc	HP OpenView Network Node Manager
Management	MCI	Management Proxy	DEV	Gnrc	
Management	MCI	Mode Management	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (10 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Management	MCI	Network Manager	OTS	Gnrc	HP OpenView Network Node Manager
Management	MCI	Performance Management	OTS/ DEV	Gnrc	RFP released
Management	MCI	Performance Management Proxy	DEV	Gnrc	
Management	MCI	Performance Test	OTS	Gnrc	vendor-supplied with hardware
Management	MCI	Physical Configuration Management	OTS	Gnrc	Mountain View
Management	MCI	Physical Configuration Proxy Agent	DEV	Gnrc	
Management	MCI	Report Generation	OTS	Gnrc	No decision yet, evaluation in progress
Management	MCI	Report Generation and Distribution	DEV	Gnrc	
Management	MCI	Report Generation Manager	DEV	Gnrc	
Management	MCI	Resource Class Category	DEV	Gnrc	
Management	MCI	Security Databases	OTS	Gnrc	Operating System Password Files, DCE Registry Database, Router Configuration Files, TCP Wrappers configuration files, Operating System Access Control Lists, DCE Access Control Lists
Management	MCI	Security Management	DEV	Gnrc	
Management	MCI	Security Management Proxy	DEV	Gnrc	
Management	MCI	Security Tests	OTS	Gnrc	CRACK, COPS, SATAN, TRIPWIRE
Management	MCI	Trouble Ticketing Management Service	OTS	Gnrc	Remedy Action Request System
Management	MCI	Trouble Ticketing Proxy Agent	DEV	Gnrc	
Management	MCI	Trouble Ticketing Service Requester	DEV	Gnrc	

Table 3.3.2-1. Langley Components Analysis (11 of 11)

Subsystem	CSCI	CSC	Type	Use	Notes
Management	MCI	Trouble Ticketing User Interface	DEV	Gnrc	
Management	MCI	User Contact Tool	OTS/ DEV	Gnrc	Remedy
Management	MCI	User Profile Server	DEV	Gnrc	
Management	MLCI	Baseline Manager	OTS/ DEV	Gnrc	XRP II
Management	MLCI	Configuration Management	OTS	Gnrc	ClearCase
Management	MLCI	Inventory/Logistics/ Maintenance (ILM) Manager	OTS/ DEV	Gnrc	Vendor evaluation in progress
Management	MLCI	Policies and Procedures Management	DEV	Gnrc	
Management	MLCI	Software Change Manager	OTS/ DEV	Gnrc	ClearCase
Management	MLCI	Software Distribution Management Structure	OTS/ DEV	Gnrc	ClearCase and Tivoli
Management	MLCI	Software Request Manager	OTS/ DEV	Gnrc	DDTS
Management	MLCI	Training Management	DEV	Gnrc	
Planning	PLANG	On Demand Manager	DEV	Gnrc	
Planning	PLANG	PDPS DBMS	OTS DEV	Gnrc	Sybase DBMS
Planning	PLANG	Planning Object Library	OTS	Gnrc	Delphi C++ class libraries
Planning	PLANG	Production Planning Workbench	DEV	Gnrc	
Planning	PLANG	Production Request Editor	DEV	Gnrc	
Planning	PLANG	Resource Planning Workbench	DEV	Gnrc	
Planning	PLANG	Subscription Editor	DEV	Gnrc	
Planning	PLANG	Subscription Manager	DEV	Gnrc	

3.4 DAAC Hardware and Network Design

This section describes the ECS hardware and local area network design supporting the Release B ECS mission at the Langley DAAC. Section 3.4.1 contains an overview diagram from the "networks" point of view, and detailed descriptions of the Release B LANs. Section 3.4.2 contains a hardware overview diagram of all of the ECS subsystems at Langley DAAC, followed by detailed descriptions and rationale for each subsystem.

3.4.1 Langley DAAC LAN Configuration

The Release B Langley DAAC LAN design builds upon the design for Release A. The major design drivers and rationale for the Release A design are presented in the Release A Langley DAAC Design Specification (305-CD-031-001), and that document may provide useful background information and context for the current Release B design.

The Langley DAAC LAN topology is illustrated in Figure 3.4.1-1. The topology consists of a User FDDI Network, a Production FDDI Network, and a HiPPI Network (for processing flows). The creation of separate User and Processing networks allows processing flows to be unaffected by user pull demands, and the introduction of the high-speed HiPPI Network provides adequate bandwidth to the Processing and Data Server subsystems to transfer high volumes of data. Each of the networks is discussed in detail below.

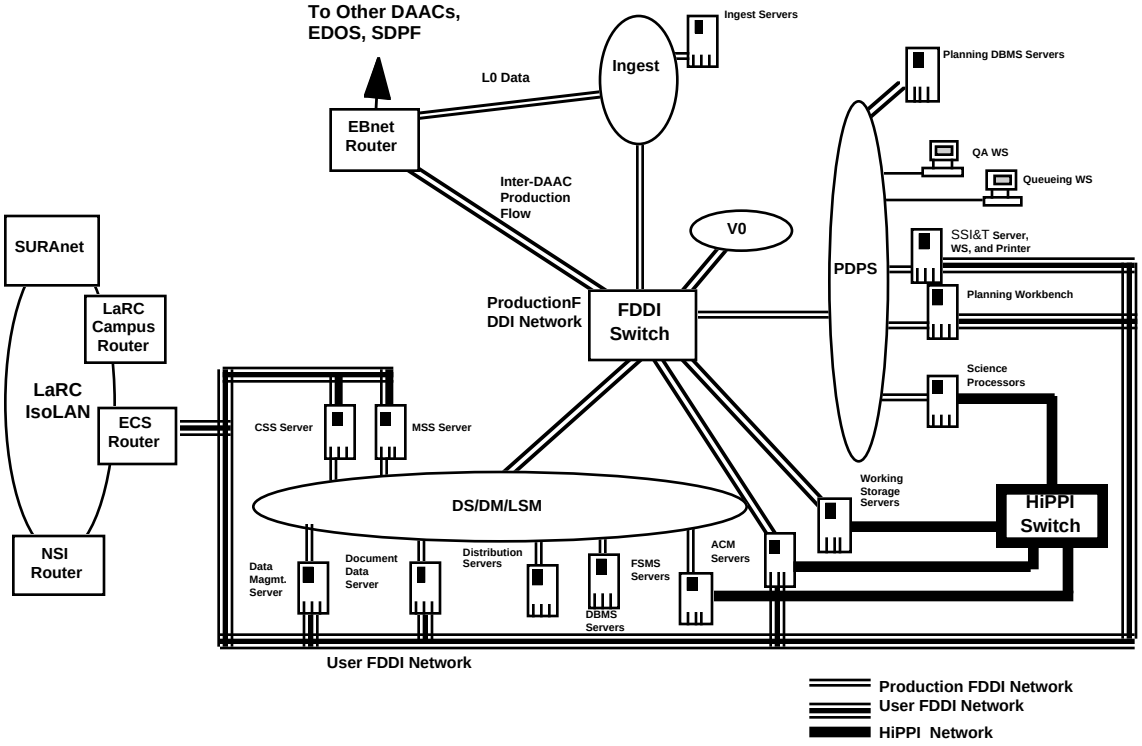

Figure 3.4.1-1. Langley DAAC LAN Topology

The Production Network consists of multiple FDDI rings supporting the DAAC subsystems and connections to external production systems (such as EDOS and other ECS DAACs) via EBnet. The separation and aggregation of hosts and subsystems onto FDDI rings is driven mostly by RMA and data flow requirements. For instance, Ingest is contained on an individual FDDI ring because of the strict RMA requirement for receipt of Level 0 data (0.998 with MDT of 15 minutes). RMA also dictates Ingest's direct connection to the EBnet router. Some Data Server hosts may be contained on dedicated FDDI rings in order to provide adequate bandwidth for the DAAC-to-DAAC processing flow requirements (refer to Table 3.4.1.1-1). The DM, LSM, and some Data Server hosts are contained on a single ring because their flows are expected to be fairly small given that user traffic will be processed on the separate User Network (see discussion below). Another ring provides access to the EBnet router to handle DAAC-DAAC flows. The FDDI Switch is the central device connecting the FDDI rings together, and it provides the necessary routing and filtering control. Note that the device counts in the figure are approximate and (for clarity) workstations and printers are not shown.

The User Network is an FDDI-based LAN connecting users (via NSI, local campuses, general Internet, etc.) to DAAC hosts responsible for providing user access. It has the main advantage of separating user and production flows. This allows DAAC processing data flows to be unaffected by user demand, so that even unanticipated user pull will not hinder the production network. Basically, the User Network provides access to Data Manager hosts and to a subset of Data Server hosts that interact directly with users. Users will not have access to any other hosts, such as Ingest or Processing devices. CSS and MSS servers are connected to the User Network but will not allow direct user access. These connections are required for communications with outside networks for such things as name lookups and receipt of Internet mail, as well as communication with and monitoring of the DAAC's interfaces to the user community (such as NSI and the local campus).

The User Network will connect to the LaRC Isolation LAN through an ECS router which will provide the necessary routing and filtering controls. NSI, the local LaRC Campus, and other Internet providers such as SURAnet will also be connected to the LaRC Isolation LAN. ECS will have an additional direct FDDI connection to the NSI router because of the large user flows expected at LaRC. Connectivity to V0 will be achieved via a direct connection through the DAAC FDDI switch. This connection to V0 will be used for migrating V0 datasets onto ECS. V0 will have a separate connection to EBnet for all other V0 data flows.

Individual FDDI rings for the Production Network will be implemented with FDDI concentrators to provide ease of wiring and central points of management. All Processing Network DAAC hosts will have FDDI interfaces and will be attached directly to the FDDI rings. Workstations will have single-attached FDDI cards, whereas the high-performance servers and processors on the Production Network will have dual-attached FDDI cards to provide redundancy. Dual-attached hosts will be dual-homed to two separate FDDI concentrators to provide an additional level of redundancy in the event of a hub failure. Interfaces of User Network hosts will be single-attached, except for the Data Management Server, which will be dual-attached. Printers and X-terminals will be connected to the DS/DM/LSM FDDI ring via an FDDI-to-Ethernet hub.

The HiPPI Network interconnects Data Server hosts/devices and Science Processors in order to provide a high-speed network to handle the large data transfers between the two subsystems. The HiPPI network will be implemented via a central HiPPI switch with switched 800 Mbps interface ports connected directly to the high-powered processing and storage hosts. The HiPPI Network shifts the numerous transfers of large volumes of data onto a dedicated high-speed topology, freeing the FDDI-based Production Network to handle control flows and DAAC-DAAC processing flows.

ECS will implement IP over the HiPPI Network. The prototyping effort undertaken by ECS to determine the performance of IP over HiPPI has shown that with proper tuning, the level of throughput needed for processing-to-dataserver flows can be achieved. An IP-based lightweight protocol called Bulk Data Service (BDS) will be used so that processing and dataserver applications can make effective use of the HiPPI fabric. Refer to the Release B CSMS Communications Subsystem Design Specification (305-CD-028-002) for more detail on BDS.

Quantities of networking hardware components to upgrade the Langley DAAC network from Release A to Release B are presented in Table 3.4.1-1.

Table 3.4.1-1. Release B Networking Hardware for Langley DAAC LAN

Networking Component	Release A Quantity	Release B Additional Quantity	Comments
FDDI Concentrator	9	2	Bay Networks 2914-04 concentrator with 12 M & 1 A/B port
FDDI Cables	82	33	Multimode fiber cables with MIC connectors
FDDI-to-Ethernet Hub	1	1	Cabletron MicroMMAC-22E; used for X-terminals
Ethernet Cables	8	12	10baseT connection to printers and X-terminals
FDDI Switch/Router	1	0	Interconnects Production Network
FDDI Router	0	1	Interconnects User Network
HiPPI Switch	0	1	High-speed switch connecting Data Server and Processing
HiPPI Cables	0	26	Copper cable

3.4.1.1 Sizing/Performance Rationale

The data flow estimates used as input to the design process for the Langley DAAC LAN topology are contained in Table 3.4.1.1-1. The table, based on dynamic analysis of the February 1996 AHWGP baseline (results for Epoch k (3Q99)) and the February 1996 User Pull Baseline (results for the greater of July 1999 and January 2000), is arranged according to the source and sink of the flow. It provides both raw 24-hour average data flows (including any applicable reprocessing)

which are the output of ECS models, as well as weighted flows containing all overhead and contingency factors. The "Factors Applied" column shows which factors (listed beneath the table) were applied to each data flow.

Table 3.4.1.1-1. Estimated Release B Data Flows for the Langley DAAC

Major Data Flow Description *	Raw Volume (in Mbps)	Factors Applied	Weighted Volume (in Mbps)
Ingest to Working Storage Server	3.4	2,3,4,5,6	9.7
Working Storage Server to/from FSMS Server	117.4	1,4,5,6,7,8	266.5
Working Storage Server to Processing	170.3	1,4,5,6,7,8	386.6
Working Storage Server to Distribution Server	24.7	2,3,4,5	69.5
ACM Server to/from other DAACs	9.9	1,2,3,4,5,6	33.5
User Pull	14.7	2,3,4,5	41.3

*Other flows such as session establishments amongst the subsystem hosts and subsystems to and from MSS are trace amounts and are not included in the table.

Overhead Factors:

(1) SSI&T Factor: 1.2. This factor not applied to Ingest or User flows. Accounts for capacity for integration and test flows.

(2) TCP/IP/FDDI Protocol Overhead: 1.25. This is applied only to the FDDI flows and not to HiPPI flows. Accounts for overhead associated with FDDI, IP, TCP, and other protocols (such as DCE).

(3) FDDI Maximum Circuit Utilization Factor: 1.25. This is applied only to the FDDI flows and not to HiPPI flows. Accounts for amount of 100 Mbps bandwidth that is actually usable for sustained data rates.

(4) Average-to-peak Conversion Factor: 1.5. This provides elasticity in the network by converting the 24 hour averages provided by the model into peaks.

(5) Scheduling Contingency: 1.2. This reflects the ability for the network to recover within 24 hours from a 4 hour down-time (24/20=1.2).

(6) Operational Hours Factor: 1.0 at LaRC. Accounts for percentage of day/week operations are performed. This factor is applied only to production flows, not to user flows.

(7) HiPPI Protocol Efficiency: 1.05. This accounts for the protocol overhead associated with the HiPPI protocol contained in the packet stream. A factor of "1.05" was chosen because overhead comprised of the (LE and FP protocols) is very low when accessing the HiPPI fabric. This factor is applied only to flows over the HiPPI network.

(8) HiPPI Maximum Circuit Utilization Factor: 1.0. This factor is applied only to flows over the HiPPI network. A factor of "1.0" was chosen because HiPPI is a dedicated medium and not a shared medium like FDDI. Thus, it is assumed each host will have access to the entire 800 Mbps channel connecting it to the HiPPI switch and that the HiPPI switch is non-blocking. The limiting factor in the HiPPI flows is the host (and peripheral access).

3.4.2 DAAC Hardware Configuration

The ECS DAAC hardware suite at LaRC hosts the Data Server, Data Management, Ingest, Interoperability, Production Planning, Science Data Processing, Management and Communications subsystems. The hardware and COTS software selected for the LaRC configuration is illustrated in Figure 3.4.2-1, Langley ECS DAAC Hardware Configuration Diagram. These configurations represent the candidate hardware selections which most closely satisfy the processing, storage capacities and communications bandwidth requirements described in the following sections. In some cases the selected configuration appears to significantly exceed the requirements to due to the sizing increments provided by the selected vendor, when in reality, our analysis and selection process has provided cost effective solutions to each problem.

3.4.2.1 Client Subsystem

There is no dedicated hardware support (HWCI) for the Client Subsystem. The Client software configurations are supported by: (1) non-ECS provided hardware platforms, in the case of Client software utilized by the user community, or (2) ECS provided workstations utilizing Client software in support of operations users (network management, DAAC operations, etc.).

3.4.2.2 Data Server Subsystem

During the Release B time frame at the Langley DAAC, a Data Server configuration is supplied to support the following data storage and access services: V0 migration data, TRMM instruments, AM-1 instruments, and SAGE III data as described in the February 1996 Technical Baseline. The Data Server hardware configuration for Release B is discussed in the subsections that follow. Sizing refinements to the IDR document are based on policy decisions by ESDIS project in the areas of reprocessing, backup, user pull sizing, produce versus store policies, and adjustments to the technical baseline driven by processing modifications by the instrument teams. The reader is encouraged to refer to the Data Server specific volume of DID305 for more details on the operational and engineering concepts that form the basis for the DSS Release B design and a more detailed discussion of these design variables.

The Data Server configuration has made some assumptions in the area of reprocessing in regards to disk, robotics, tape drives, and general I/O. The dynamic model used for sizing produced peak and average usage statistics for Data Server resources, monitored and calculated over the full processing cycles of the Release B products (i.e., daily, weekly, monthly, etc.). The peaks are caused by different data staging requirements of specific PGEs. The Processing Subsystem's disk usage patterns are consistent with this pattern. The dynamic model has also shown, that Data Server archive hardware components' sizing to two times the average computed requirements will be sufficient to keep up with the overall level of demand. Reprocessing load, estimated to be equal to one time the processing load, will be worked off in the off-peak processing times, as shown by the analysis done after IDR B. An expansion to the two times reprocessing load can be achieved by horizontal scaling of the components, i.e. adding additional quantities of hardware to the existing configuration.

LaRC at Rel B (1 of 2)

Added for Release B

Note: All systems come with OS, C, C++, DCE, OODCE, Clearcase agent, SNMP agent (Optima), Sybase client, Tivoli client, CD-ROM, FDDI. RAID quantities are USABLE.

SPRHW

SPRHW-LARC-1
Science Processor
SGI PC XL 12CPU
1GB RAM; 4GB disk
14*R10000 CPUs
+HiPPI;+8GB Disk;
+1GB RAM
CASEVision
IDL, IMSL
F77, F90, Ada
DCE Dev Kit
Networker BU SW
Autosys Agent

Note: Rel A RAID in cabinets; change to rack-mounted; server console rather than 20" color monitor

- SPRHW-LARC-1A: 8mm Stkr
- SPRHW-LARC-1B: 68GB RAID; +76GB RAID
- SPRHW-LARC-1C: 4mm Stkr
- SPRHW-LARC-1D: 68GB RAID; +76GB RAID

SPRHW-LARC-2,3,4
X-Terminal
NCD HMX Pro
16MB
X-Terminal
X-Terminal

SPRHW-LARC-12
X-Terminal
NCD HMX Pro
16MB

SPRHW-LARC-14
X-Terminal

SPRHW-LARC-5
Science Processor
SGI PC XL 12CPU
1GB RAM; 4GB; 4mm
+1GB RAM;+HiPPI;
+8GB disk
CASEVision
IDL, IMSL
F77, F90, Ada
DBTools.h++, Tools.h++
Autosys Agent

- SPRHW-LARC-5A: 51GB RAID; +17GB RAID; +Rack Mounts; 20x4.3GB

SPRHW-LARC-6
Science Processor
SGI PC XL 18CPU
1GB RAM; 4GB
+1GB RAM; HiPPI;
+8GB Disk
CASEVision
IDL, IMSL
F77, F90, Ada
Autosys Agent

- SPRHW-LARC-6A: 51GB RAID; +51GB RAID; +Rack Mounts; Each Rack Mounted as 15x4.3GB

SPRHW-LARC-7
Queuing Srvr
SUN Sparc 20/71
384MB, 4GB
+1CPU, +128MB
AutoSys
WABI/MSOffice
Sybase

- SPRHW-LARC-7A: 6GB Disk
- SPRHW-LARC-7B: 6GB Disk

SPRHW-LARC-13
Queuing Srvr
SUN Sparc 20/712
512MB, 4GB
AutoSys
WABI/MSOffice
Sybase

- SPRHW-LARC-12A: 6GB Disk
- SPRHW-LARC-12B: 6GB Disk

SPRHW-LARC-8
SGI PC XL
16CPU; 2GB RAM;
HiPPI; 18GB Disk
CASEVision
IDL, IMSL
F77, F90, Ada
Autosys Agent

- SPRHW-LARC-8A: 144GB RAID
- SPRHW-LARC-8B: 144GB RAID
- SPRHW-LARC-8C: 144GB RAID

SPRHW-LARC-8(A/C): 3*20x9GB ea in 1 rack

SPRHW-LARC-9
SGI PC XL
24CPU;2GB RAM;
HiPPI;18GB Disk
CASEVision
IDL, IMSL
F77, F90, Ada
Autosys Agent

- SPRHW-LARC-9A: 144GB RAID
- SPRHW-LARC-9B: 144GB RAID

SPRHW-LARC-9(A/B): 20x9GB; 20x9GB in 1 rack

SPRHW-LARC-10
SGI PC XL
16CPU; 2GB RAM;
HiPPI; 18GB Disk
CASEVision
IDL, IMSL
F77, F90, Ada
Autosys Agent

- SPRHW-LARC-10A: 144GB RAID
- SPRHW-LARC-10B: 144GB RAID
- SPRHW-LARC-10C: 144GB RAID

SPRHW-LARC-10(A/B/C): 3*20x9GB in same rack

SPRHW-LARC-11
SGI PC XL
24CPU;2GB RAM;
HiPPI;18GB Disk
CASEVision
IDL, IMSL
F77, F90, Ada
Autosys Agent

- SPRHW-LARC-11A: 144GB RAID
- SPRHW-LARC-11B: 144GB RAID

SPRHW-LARC-11(A/B): 20x9GB; 20x9GB in 1 rack

PLNHW

PLNHW-LARC-1
PDPS DBMS Server
SUN Sparc 20/71
384MB, 4GB
+128MB RAM;
+1CPU
AutoSys
SPARCworks
Sybase
SoftPC/MSOffice
DBTools.h++, Tools.h++

- PLNHW-LARC-1A: 6GB Disk
- PLNHW-LARC-1B: 6GB Disk

PLNHW-LARC-2
PDPS DBMS Server
SUN Sparc 20/712
512MB; 4GB
AutoSys
SPARCworks
Sybase
SoftPC/MSOffice
DBTools.h++, Tools.h++

- PLNHW-LARC-2A: 6GB Disk
- PLNHW-LARC-2B: 6GB Disk

Note: 22GB re-distributed into 12GB/PDPS DBMS Server

AQAHW

AQAHW-LARC-1
QA WS
SGI Indigo2 XZ
256MB; 10GB
Upgrade to R10000
IDL, IMSL
CASEVision
ZMail

- AQAHW-LARC-1A: 17GB RAID

AITHW

AITHW-LARC-1
AIT WS
SUN Sparc 20/50
64MB, 6GB
SPARCworks
IDL, IMSL
F77, F90, Ada
SoftPC/MSOffice
Forcheck

AITHW-LARC-2
AIT WS/DBMS server
SUN Sparc 20/50
128MB, 8GB
SPARCworks
IDL, IMSL
F77, F90, Ada
SoftPC/MSOffice
Sybase

AITHW-LARC-3
Printer
HP LaserJet 4M+
12ppm/14MB RAM

AITHW-LARC-6
Printer
HP LaserJet 4M+
12ppm; 14MB
(Ir1 - A Transition)

AITHW-LARC-4
AIT WS
SUN Sparc 20/712
256MB; 6GB
(Ir1 - A Transition)

AITHW-LARC-5
AIT WS
SUN Sparc 20/50
64MB; 6GB
(Ir1 - A Transition)

MSS

MSS-LARC-6
Ir1 MSS Server
SUN 20/712
256MB; 6GB
(Ir1 - A Transition)
WABI/MSOffice
ZMail

AITHW-LARC-7
X-Terminal
NCD HMX Pro
16MB

AITHW-LARC-8
X-Terminal
NCD HMX Pro
16MB

AITHW-LARC-9
X-Terminal
NCD HMX Pro
16MB

AITHW-LARC-10
X-Terminal
NCD HMX Pro
16MB

MSS-LARC-3
MSS WS
SUN Ultra 170
128MB; 4GB
WABI/MSOffice
ZMail, SPARCworks
Billing & Acctg Server & Client
DBMS Rept Generation
Statistical Analysis / Trending SW

MSS-LARC-2
Printer
HP LaserJet 4M+
12ppm; 14MB

MSS-LARC-1
MSS WS
SUN Ultra 170
128MB;8GB
Clearcase Server
DDTS (Puresoft)
XRP (HTG)
ZMail, SPARCworks
ILM SW
WABI/MSOffice
Billing &
Acctg Client
SW License Mgmt

MSS-LARC-4
MSS Server
HP J210/1
256MB; 4GB
+1CPU;+128MB
OpenView (HP)
TME (Tivoli)
ARS (Remedy)
MountainView/PNM
WABI/MSOffice
DBTools.h++, Tools.h++
Sybase Server
Courier (Tivoli)

- MSS-LARC-4A: 20GB RAID; +16GB RAID

MSS-LARC-5
Backup Server
SUN Ultra 4-slot
256MB; 4GB
Legato Networker SW

- MSS-LARC-5A: DLT Stacker

MSS-LARC-6
Impact
Printer

CSS

CSS-LARC-1
CSS Server
HP J210/1
256MB; 4GB
+1CPU;+128MB
DCE Name
DCE Security
Softbench
WABI/MSOffice
Mail Server
Master Sybase Srvr

ISS

FDDI Switch	Quantity: 1
Alantec Power Hub w/4 FDDI cards; 3 power supplies	
FDDI Concentrator:	Quantity: 11
Bay Networks 2914-04	
FDDI Cables:	Quantity: 76
Ethernet Hub:	Quantity: 1
Cabletron MicroMAC-22E w/BRIM-F6	
Ethernet Cables:	Quantity: 4
LAN Analyzer:	Quantity: 1
Comm Cabinets:	Quantity: 2
FDDI Router	Quantity: 1
FDDI Concentrators	Quantity: 2
FDDI Cables	Quantity: 29
HiPPI Switch	Quantity: 1
HiPPI Cables:	Quantity: 22
Ethernet Hub:	Quantity: 1
Ethernet Cables:	Quantity: 12

DMGHW

DMGHW-LARC-1
Data Spec WS
SUN Sparc 20/50
64MB; 6GB
SPARCworks
Zmail

DMGHW-LARC-2
Data Spec WS
SUN Sparc 20/50
64MB; 6GB
ZMail

DMGHW-LARC-3
Data Spec WS
SUN Sparc 20/50
64MB; 6GB
ZMail

DMGHW-LARC-4
DMG Server
HP K200/1
256MB;2GB; 8mm
Upgrade to K400
2 CPU; 512MB RAM
6GB Disk
Sybase, Web Srvr.
Softbench
DBTools.h++, Tools.h++

- DMGHW-LARC-4A: 10GB RAID

DMGHW-LARC-5
DBA Ops WS
HP 715/64
64MB; 6GB
Softbench

LaRC at Rel B (2 of 2)

Added for Release B

Note: All systems come with OS, C, C++, DCE, OODCE, Clearcase agent, SNMP agent (Optima), Sybase client, Tivoli client, CD-ROM, FDDI. RAID quantities are USABLE.

ACMHW

ACMHW-LARC-1
Ops WS
SUN Sparc 20/50
64MB; 4GB
SPARCworks
ZMail

ACMHW-LARC-2
Ops WS
SUN Sparc 20/50
64MB; 4GB
ZMail

ACMHW-LARC-3
APC Server
SGI Challenge L
256MB; 4GB
Upgrade to 10CPU XL
+768MB;+2GB
Illustra
Tools.h++
NFS

ACMHW-LARC-3A
40GB RAID
+28GB RAID
ACMHW-LARC-3B 144GB RAID
ACMHW-LARC-3C 144GB RAID
ACMHW-LARC-3D 144GB RAID

ACMHW-LARC-4
APC Server
SGI Challenge L
256MB; 4GB
Upgrade to 6CPU XL
+256MB;+2GB
Illustra
Tools.h++
NFS

ACMHW-LARC-3A: 20x4.3GB
RAID Rack Mounted: ACMHW-LARC-3B/D: 20x9GB each

WKSHW

WKSHW-LARC-1
WKS Host
SGI Challenge XL
512MB;6GB;6CPU
HiPPI
NFS

WKSHW-LARC-1A
144GB RAID
WKSHW-LARC-1B 144GB RAID
WKSHW-LARC-1C 144GB RAID
WKSHW-LARC-1D 36GB RAID

WKSHW-LARC-2
WKS Host
SGI Challenge XL
256MB;6GB;4CPU
HiPPI
NFS

RAID Rack Mounted:
WKSHW-LARC-1A/C: 20x9GB each
WKSHW-LARC-1D: 5x9GB each

DIPHW

DIPHW-LARC-1
Distribution Srvr
SUN Sparc 20/712
256MB; 6GB
Upgrade to 4CPU
SUN Ultra
256MB; 6GB
SPARCworks
C, C++
Tools.h++

DIPHW-LARC-1A
20GB RAID
+220GB
DIPHW-LARC-1B 6250 Tape
DIPHW-LARC-1C 8mm Stkr
DIPHW-LARC-1D 4mm Stkr
DIPHW-LARC-1E CD-ROM Recordbl
DIPHW-LARC-1F 3480 Drive
DIPHW-LARC-1H 3490 Drive
DIPHW-LARC-1G FAX/ Scanner

DIPHW-LARC-2
Distribution Srvr
SUN Sparc 20/712
256MB; 6GB
Upgrade to 4CPU
SUN Ultra
256MB; 6GB
C, C++

DIPHW-LARC-3
Printer
HP LaserJet
12ppm; 14MB

DIPHW-LARC-4
Printer
HP LaserJet
12ppm; 14MB

DRPHW

DRPHW-LARC-1
FSMS Server
SGI Challenge XL
256MB; 6GB
+2CPU;+768MB
HiPPI
NFS, C++
AMASS

DRPHW-LARC-1A
68GB RAID

DRPHW-LARC-2
FSMS Server
SGI Challenge XL
256MB; 6GB
+2CPU;+768MB
HiPPI
NFS, C++
AMASS

DRPHW-LARC-5
EMASS AML Model 2
1 Tall Quadro Tower; 1 Arm;
+1 Arm, +8 3590 Drives
DRPHW-LARC-5A
DRPHW-LARC-5B
DRPHW-LARC-5C
DRPHW-LARC-5D
DRPHW-LARC-5E
DRPHW-LARC-5F
DRPHW-LARC-5G
DRPHW-LARC-5H
DRPHW-LARC-5I
DRPHW-LARC-5J

DRPHW-LARC-6
EMASS AML Model 2
1 Tall Quadro Tower; 2 Arms;
12 3590 Drives
DRPHW-LARC-6A
DRPHW-LARC-6B
DRPHW-LARC-6C
DRPHW-LARC-6D
DRPHW-LARC-6E
DRPHW-LARC-6F
DRPHW-LARC-6G
DRPHW-LARC-6H
DRPHW-LARC-6I
DRPHW-LARC-6J
DRPHW-LARC-6K
DRPHW-LARC-6L

DRPHW-LARC-7
EMASS AML Model 2
1 Tall Quadro Tower; 2 Arms;
10 3590 Drives
DRPHW-LARC-7A
DRPHW-LARC-7B
DRPHW-LARC-7C
DRPHW-LARC-7D
DRPHW-LARC-7E
DRPHW-LARC-7F
DRPHW-LARC-7G
DRPHW-LARC-7H
DRPHW-LARC-7I
DRPHW-LARC-7J

DRPHW-LARC-3
DBMS Server
SGI Challenge XL
256MB; 6GB
Sybase, NFS, C++
DBTools.h++
Tools.h++

DRPHW-LARC-3A
10GB RAID
+17GB RAID

DRPHW-LARC-4
DBMS Server
SGI Challenge XL
256MB; 6GB
Sybase, NFS, C++
DBTools.h++
Tools.h++

ICLHW

ICLHW-LARC-1A
8mm Stkr

ICLHW-LARC-1
Ingest Server
SGI Challenge L
256MB; 6GB
AMASS
Builders Xcessory
CASEVision
Sybase
ZMail

ICLHW-LARC-1B
51GB RAID
ICLHW-LARC-1C 51GB RAID
+17GB RAID

ICLHW-LARC-2
Ingest Server
SGI Challenge L
256MB; 6GB; 4mm
AMASS
Builders Xcessory
CASEVision
Sybase
DBTools.h++
ZMail

ICLHW-LARC-2A
8mm Stkr

ICLHW-LARC-3
X-Terminal
NCD HMX Pro
16MB

ICLHW-LARC-4
EMASS AML/E
1 Quadro Tower; 1 Arm;
2 3590 Drives
ICLHW-LARC-4A
ICLHW-LARC-4B

ICLHW-LARC-5
Ingest Server
SGI Indigo2
256MB; 6GB

ICLHW-LARC-6
Ingest Server
SGI Challenge DM
256MB; 4GB
(lr1 - A Transition)

DDSHW

DDSHW-LARC-1
Document Server
SUN Sparc 20/712
256MB; 6GB
Netscape, Topic,
Illustra

DDSHW-LARC-1A
6GB RAID

DDSHW-LARC-2
Document Server
SUN Sparc 20/712
256MB; 6GB
SPARCworks
Netscape, Topic,
Illustra

This page intentionally left blank.

The Data Server Subsystem hardware configuration at Langley DAAC consists of five hardware CIs and is sized for the mission support as described above, for a period of one calendar year beyond AM-1 launch date (Epoch k of the Technical Baseline) plus V0 data migration for the same time period.

- ≠ ***Access Control & Management (ACMHW HWCI)*** -- The access hardware allows for client access (both the client subsystem and direct "push/pull" user access) to the Data Server subsystem and provides tools and capabilities for system administration. It is broken down into two components; Administration Stations (AS) which consist typically of operations support workstations, and Access/Process Coordinators (APCs) which consist of server class machines with host attached storage.

- ≠ ***Working Storage (WKSHW HWCI)*** -- Working Storage (WS) hardware of the Data Server supplies storage used for temporary file and buffer storage within the data Server architecture. All data in WS regardless of storage duration is temporary in nature and not part of the permanent archive.

- ≠ ***Distribution and Ingest Peripheral Management (DIPHW HWCI)*** -- The hardware of the Distribution and Ingest Peripheral Management CI supports the hard media distribution methods for data dissemination from the system, as well as hard media ingest of data into the system. The hardware provided by this HWCI includes a variety of media and media drives, jukeboxes/stackers as necessary, and server hosts and disk storage for network distribution.

- ≠ ***Data Repository (DRPHW HWCI)*** -- Data Repositories (DRs) are the hardware components that store and maintain data permanently. This consists of DBMS repositories and archive tape library based repositories. This HWCI provides the disk, server, archive robotics, media and archive tape drives required to support the permanent storage repositories.

- ≠ ***Document Data Server (DDSHW HWCI)*** -- This HWCI provides the disk and server required to support the Document Data Server portion of the DSS.

3.4.2.2.1 Rationale

The following subsystem-wide assumptions were applied in sizing the Data Server hardware components. Data Server Subsystem is sized for : TRMM instruments, AM-1 instruments, and SAGE III data as described in the February 1996 Technical Baseline, as well as V0 migration data support, for a period of one calendar year beyond AM-1 launch date. A combination of static and dynamic modeling was used to size the permanent data repository components, such as the number of robotic arms, tape drives, and production related staging disk. The ECS technical baseline for February 1996 was used both in dynamic modeling as well as in static analysis. User modeling data was used in dynamic models for estimating the user access rates to the system where appropriate as well as the pattern of user interaction with the system. However, DSS sizing is still being driven by the ECS Project direction to size for a distribution volume two times the site's nominal production volume. DAAC to DAAC flow estimates are as published on March 3 1996.

***ACMHW* —** Analysis was undertaken primarily for the sizing of the APC server hosts and their attached storage. The administration workstations are assumed to continue to be minimally configured workstations designed to perform various operations functions (e.g. DBMS administration, repository administration, etc.). Client desktop services and X protocol access to Data Server hosts are the driving sizing factors for the workstations.

APC Server -- The APC server provides the session establishment point for the client, routes service requests to the other CSCIs and HWCIs, and functions as the electronic distribution channel for the DSS user pull and push loads, as well as for the electronic portion of the DAAC to DAAC transfers . The APC server host runs the following software processes and applications: ScienceDataServer Process, ScienceDataServer Administration Process, SubscriptionServer Process, Network ResourceManager Process, PullMonitor Process, CSS DCE client, CSS logging API, and an MSS agent. The internal I/O to be sustained by the processor averaged over 24 hours of operation is not a driving design factor.

The following is the CPU allocation (CPU loading for peripheral support is based on SGI configuration guidelines):

0.5 CPU Monitoring/Managing Tasks (estimated)
+ *4 CPUs Subsetting/Subsampling (Dynamic Modeling)*
+ *1 CPU Operating System and Applications (estimated)*
+ *2.3 CPUs for client support (session threads)*
+ *0.075 CPU for 1 Differential SCSI -II Interfaces (1 SCSI [disk] at 5 MB/sec - 7.5% of a CPU for each 5 MB/sec)*
+ *0.03 CPUs for SCSI RAID controller support (1.5% of a CPU for each 5 MB/sec of controller I/O; 10.48*0.015/5)*
+ *0.079 CPUs for FDDI I/O support (7.5% of a CPU for each 5 MB/sec of SCSI I/O: 5.24 MB/sec/5 MB/sec * 0.075)*
+ *0.05 CPUs in support of subscriptions.*

This requires a total of 9.78 equivalent CPUs. The SGI Challenge XL server with 10 CPUs and 1 GB of RAM will be selected as the APC server. The server will be equipped with 6 GB of local disk, 1 differential SCSI-II controller for the local disk, 15 SCSI RAID Controllers (see *APC Host Storage* below), and 2 DAS FDDI Interfaces. A degraded mode failover will be provided by the second SGI Challenge XL equipped with 6 CPUs, 512 MB, and an identical complement of peripheral interfaces. Here and elsewhere at the LaRC site all SGI CPUs are 200 MHz R10000.

AP*C Host Storage* -- The APC server host disk is sized for electronic ingest (almost exclusively from sources external to the DAAC), as well as electronic distribution (again, almost exclusively to the recipients external to the DAAC) since this pool of hosts is designed to manage the requests to the Data Server as well as the service response. APC storage must also be sized to support functions such as subsetting/subsampling and storage of user session context, which keeps track of user session interactions that may be suspended and resumed.

For Release B the electronic distribution, as defined within the ECS Technical Baseline for CDR, is "one time" the total Data Server ingested volume. This "one times" volume is made up of data

that is "pushed" and "pulled" by the user. The distribution of this electronic data is assumed to occur over a 24-hour period. It is further assumed, that on the average each granule placed on the distribution volumes is pulled/pushed to two separate users, and hence is counted twice in the total daily output. This assumption is based on the workings of subscriptions (assuming that a granule is subscribed to more than once) and the fact that "interesting" data will be requested more than once when inserted while some data may not be requested at all. Distribution of data to users is estimated at 2.66 MB/s (229 GB/day), with 3.11 MB/s (2.8 GB accumulation in 15 minutes disk latency) received from users and other DAACs and going out to other DAACs. It is assumed, that 100% of DAAC to DAAC transfers are electronic. While both the I/O capacity and the disk allocations take the transfers into account, the rate is not a driving factor in the design. Allowing for two days worth of capacity on a fast access medium and adding to it 21 GB of disk required for subsetting/subsampling, as estimated by dynamic modeling, and 1 GB for session context storage this yields approximately 480 GB of RAID.

WKSHW — The rationale for allocation of resources to the WKSHW HWCI has changed from that presented in the IDR version of this document for several reasons. The most significant change is the requirement to store interim products generated by the processing subsystem. Storage of interim products within data server is provided within the WKSHW HWCI rather than within the deep archive (DRPHW HWCI) for two reasons. Principally, it is because these products are required to be stored temporarily rather than permanently. Different data access and archive management requirements apply to data that is not stored permanently, reducing the complexity of the management of these interim products. This requirement for temporary storage, whose volume was undefined at IDR-B, increases the overall WKSHW storage requirement at LaRC by approximately 315 GB.

It has been established by dynamic modeling, that no performance benefit is derived at LaRC with the use of more than 15 minutes of Working Storage. 27 GB of RAID provides sufficient storage to support 15 minutes worth of data received from processing, hard media ingest, electronic V0 migration data ingest, DAAC data ingest, and client support. Working Storage RAID is designed to be a temporary staging area only, not a long-term repository, so 15 minutes worth of space provides adequate time to either write the data to tape or forward it to its destination. The total quantity of the Working Storage RAID required is 342 GB.

Working storage resources had also been assumed at IDR-B to be network attached and largely pooled, allowing sharing of disk resources among servers from the other DSS HWCIs. The technology required to accomplish the pooling of disk resources did not mature as anticipated in some cases, or is only available in a homogeneous vendor solution, and is prohibitively expensive in others. Therefore, the WKSHW HWCI requirement has been expanded to include a host. The following processes and applications will run on this server: ResourceManager Process for the staging disk, StagedDataMonitor Process, CSS DCE client, CSS logging API, and an MSS agent.

The CPU allocation for this server is as follows:

0.5 CPU Monitoring/Managing Tasks (estimated)
\+ *1 CPU Operating System and Applications (estimated)*
\+ *2.3 CPUs for client support (session threads)*
\+ *0.075 CPU for 1 Differential SCSI -II Interfaces (1 SCSI [disk] at 5 MB/sec -*

*7.5% of a CPU for each 5 MB/sec; 1 * 0.075)*

\+ *0.14 CPUs for SCSI RAID controller support (1.5% of a CPU for each 5 MB/sec of controller I/O; 45.13*0.015/5)*

\+ *0.15 CPUs for FDDI I/O support ({total data and control flow over FDDI approximated at 10 MB/sec}/ 5 MB/sec * 0.075)*

\+ *1.75 CPUs for HIPPI I/O support (1.75 CPU * 1 channels).*

This requires a total of 5.9 equivalent CPUs. The SGI Challenge XL server with 6 CPUs will serve as the WKSHW server. 6 GB of local disk and 512 MB of RAM are also required for this server. A 4 CPU SGI Challenge XL with 256 MB RAM will serve a spare for degraded mode of operation if the primary machine fails.

DIPHW — The DIPHW configuration at LaRC includes primarily server host and disk units to serve media based distribution of LaRC data (as well as for some types of Ingest), and includes a number of peripheral form factors. Two servers (Sun Ultra 4-slot) are provided for temporary staging support and peripheral support.

The servers are designed to support the media distribution load and ingest loads. It is designed to handle the I/0 for 1x (one times) distribution of the Data Server ingested volume per twenty-four hours of operation, servicing this load in a nominal 8 hr. total shift period seven days per week. The platforms are sized to handle the network transfer traffic to local disk from the FSMS Server Host source, as well as the media preparation and media ingest I/O. The software processes/applications mapped to the servers are: DistributionServer process, ResourceManager Processes for CD-ROM, various tape, and printers, CSS DCE client, CSS logging API, and an MSS agent. Two server hosts total are provided in order to comply with RMA requirements for the function of archiving and distributing data with the required availability of 0.98 and the mean down time not to exceed 2 hours.

The disk associated with the hard media distribution platform is sized for hard media distribution of 1x (one times) of the total Data Server ingested volume per twenty-four hour operations. The same assumption is used here as was used in the APC disk sizing, namely that on the average each granule placed on the distribution volumes are written twice to two separate users. 24 hours accumulation capacity, plus ten percent, for potential hard media ingest is sized. The ingest activity is assumed primarily for the 75% of the V0 ingest volume via hard media. The activity related to the Ingest Subsystem is assumed to be negligible, since the Ingest Subsystem is sized to handle the high volume electronic ingest loads.

The servers included in the DIPHW must, in aggregate, support 8.3 MB/s of distribution (24 hours of accumulation in an 8 hour shift), a negligible flow from Version 0 hard media ingest, and 0.17 MB/s for backup of GSFC DAAC data. The output load consists of 2.71 MB/s including the LaRC backup data sent to GSFC on hard media. The CPU allocation for each of the two servers is as follows:

0.5 CPU Monitoring/Managing Tasks (estimated)

\+ *1 CPU Operating System and Applications (estimated)*

\+ *1 CPU for approximately 9 MB/sec of I/O to/from the Differential SCSI-II peripheral interfaces*

\+ *0.2 CPUs for Fiber Channel RAID controller support*

\+ *0.1 CPUs for FDDI I/O.*

This requires a total of 2.8 equivalent CPUs. The Sun Ultra 4-slot with 4 CPUs 6 GB of local disk and 256 MB of RAM is specified for these servers. A total of 235 GB of RAID will also be allocated across the two servers in this HWCI.

Peripheral Support -- The peripherals supported at the LaRC site for Release B were selected based on Level Four requirements: S-DSS-30440, S-DSS-30470, S-DSS-30480 (reference SDPS Requirements Specification for the ECS Project, 304-CD-002-002). We have retained the single Release A 6250 tape drive. V0 operational experience has shown that the use of the 6250 drive as a heritage device is low enough to make the provision of a spare drive unnecessary. In the event of the drive failure the low workload will allow waiting for the drive repair or replacement without a noticeable impact on operations. The HWCI complement may be easily scaled for both media types as well as capacity. The peripherals supplied here are included in the configuration to primarily support distribution functions. However, the Ingest Subsystem (Ingest Client) residing at LaRC may utilize peripherals to perform some media based ingest, as necessary, based on media form received for storage. This applies only to peripherals not already configured into the Ingest complement for performance reasons.

The types of media form factors/formats selected for Release B include:

- 8mm Tape,
- 6250 Tape (heritage),
- CD-ROM,
- 3480/3490,
- 4 mm Tape,
- FAX/scanner

The aggregate bandwidth for each of the 8mm and 4mm tape devices will meet the entire required distribution and ingest bandwidth plus allowance for RMA. The rest of the devices will be furnished in quantities dictated by the RMA considerations. CD-ROM drive quantities will be sufficient to allow no less than 25% of the aggregate bandwidth.

This HWCI also provides the network laser printers for the DAAC DS operations.

DRPHW — The DRPHW configuration at LaRC includes both archive based as well as DBMS based physical repositories. They are sized as follows:

Archive Repository -- The archive component was sized through a combination of both static analysis as well as dynamic simulation. The principal model employed was a static model that utilized as input both equipment performance parameters and data flow estimations that were output from the dynamic model. The equipment performance parameters are vendor specifications derated based on reported user experience. The dynamic system model was based on a discrete event simulation with constrained resources assumed. Modeling runs based on the February 1996 Technical Baseline were performed. The dynamic modeling results were also used as a check of

the static model results. Specific data regarding disk, tape drive and robotic resource utilization was obtained from the static model. This data was coupled with the key driving requirements with respect to distribution (e.g. the 2x distribution cap, the User Model service access predictions for epoch k), flow analysis, and data with respect to hardware and software COTS selections.

AMASS was chosen for the Release B archive FSMS. For the FSMS Manager client server, the platform is selected on the basis of FSMS(AMASS)/platform COTS S/W compatibility. Memory and cache estimates are currently based on vendor recommendations, reference "AMASS Archival Management and Storage System, Installation on Silicon Graphics", EMASS Part Number 600149, AMASS Version 4.2.4, March 1995. Aside from the AMASS FSMS the following processes and client applications will run on this server: ResourceManager Process for the staging disk, StagedDataMonitor Process, CSS DCE client, CSS logging API, and an MSS agent. FSMS cache requirement is 42.6 GB of RAID.

The CPU allocation for this server is as follows:

0.5 CPU Monitoring/Managing Tasks (estimated)
\+ *1 CPU Operating System and Applications (estimated)*
\+ *0.2 CPU FSMS Allocation (25.6 MB/sec total flow to and from the ATL *0.8% of a CPU per MB of flow)*
\+ *1.35 CPU for Differential SCSI -II Interfaces (2 SCSI [disk] at 5 MB/sec - 7.5% of a CPU for each 5 MB/sec + 30 SCSI [tape] at 5 MB/sec - 0.8 % CPU/ 1 MB/sec; 2 * 5 * 0.075/5 + 30*5*0.008/1)*
\+ *0.02 CPUs for FDDI I/O support(control flow = 10% of HIPPI Data Flow; 1.47*.075/5)*
\+ *0.04 CPUs for SCSI RAID controller support (1.5% of a CPU for each 5 MB/sec of controller I/O; 14.7 *0.015/5)*
\+ *1.75 CPUs for HIPPI I/O support (1.75 CPU * 1 channel).*

This requires a total of 4.86 equivalent CPUs. An SGI Challenge XL server with 6 CPUs will serve as the DRPHW server. 6 GB of local disk and 512 MB of RAM is also required for the server. A 4 CPU, 256 MB of RAM and an identical peripheral interface complement SGI Challenge XL will be provided for failover.

For the Archive Tape Library Robotics an EMASS Automated Management Library system AML/2 was selected for Release B based on its ability to accommodate multiple media form factors, storage density, growth capacity, floor space utilization, compatibility with COTS software and a number of other factors. (This selection process is documented in the SDPS Storage Technology Insertion Plan white paper (June 1995, #420-WP-003-001) and is not fully discussed here). Given data with respect to the robotic device, and assuming the performance signature of 3590 drives in the robotic unit, the static model analysis resulted in a utilization factor for sizing estimation purposes as given below:

LaRC: Configuration: 3 tall Quadro Tower units

Calculated Robotics Utilization: 2 Robotic Arms at each Tower

Calculated Tape Drive Resource Utilization: 10 units in each

The Release A LaRC configuration assumed the use of 3590 (NTP) technology, yielding a total robotic storage density of approximately 57 TB The Release A LaRC required volume was 7 TB, therefore a minimum amount of media was specified for delivery. Media requirements were calculated based on the one year Release A operations at LaRC, and media to support Release A V0 migration. By June of 1999 the accumulated LaRC data volume is projected to be approximately 52 TB assuming a conservative internal backup volume of the 2% of the entire holdings and additional 2% of GSFC's holdings stored at LaRC. Therefore, from purely a storage technology point of view, the Release A tape solution may be used for the Release B time frame up through epoch K. At that point a decision may be made whether to migrate to a denser technology or to stay with the 3590 solution. It must be noted, that the vendor is promising a 50% density increase for the cartridge (to 15 GB) within less than a year. The three tall Quadro towers will allow storage capacity cumulation to 171 TB with the 10 GB cartridges and to 256 with 15 GB. Additional storage units may then be added as needed if the decision is made to remain within the linear technology.

The volume calculations are based on the Technical Baseline of February 1996 for the Release B product data volumes and for the identified V0 migration. It is assumed that the product data accumulates over the 9/97 through 6/99 time period. No compression is assumed. Backup capacity of 2% (0.96 TB) on-site plus 2% of GSFC capacity (3.59 TB), Browse data capacity of 1MB per granule (12.72 TB), and a spare tape capacity of 10% are added. The total calculated media support requirement for Release B and V0 operations is as follows:

Release B ECS = 49,440 GB + 2% LaRC backup capacity + 2% GSFC backup capacity + browse data + 10% spare media capacity = 69,559 GB --> 6,956 count of 3590 tapes required.

DBMS Repository -- The Data Base Management System (DBMS) Repository component was sized as follows based on static data size analysis as well as transaction based analysis. The transaction analysis is based on both "push" (production metadata update) and "pull" (user access and distribution) loads. Transaction rate was modeled based on the user service request rates as described at PDR time in the User Pull Analysis Notebook, 160-TP-004-001, Question 47 and a cross section of query types derived from the DBMS Benchmark Report, 430-TP-003-001.

The DBMS Server Host was sized based on the transaction analysis mentioned above, as well as platform suitability analysis based on the DBMS COTS software selection for Release B Data Servers (Illustra). Platform suitability is based on the DBMS software manufacturer's compatibility recommendations, benchmark data, and project bench marking activities. Aside from the Data Base engine, the following processes will run on this host: CSS DCE client, CSS logging API, and an MSS agent.

DBMS Server Disk was sized based on the *core metadata* associated with TRMM, AM-1, and SAGE III data as well that associated with the V0 data sets identified for migration within Release B together with metadata associated with other data sets identified for migration with the Release B time frame. This time frame is from 1997 through 1999. Refer to table 3.4.2.2-3 for DBMS Repository Sizing beyond the Release B time frame.. Table 3.4.2.2-2 identifies the Release A and B datasets used for the sizing. The size of the metadata granule for V0 was assumed to be half (0.91 KB) that of the full ECS metadata granule of 1.82 KB. That assumption is based on the 50% mapping of V0 attributes to ECS core attributes across all products throughout the Release B duration.

Table 3.4.2.2-2 Release B Datasets Held by the Data Server Subsystem (1 of 2)

Product	Instrument	Product Description	Platform
CER01	CERES	Bi-Directional Scan Product (BDS)	TRMM
CER02	CERES	ES-8 ERBE-Like Science Product	TRMM
CER03	CERES	ES-9 ERBE-Like Product	TRMM
CER09	CERES	Single Satellite Instrument Earth Scan Product (IES)	TRMM
CER13	CERES	ES-4 ERBE-Like Product	TRMM
CER14	CERES	ES-4G ERBE-Like Product	TRMM
CER16	CERES	Clear Reflectance and Temperature History (CRH)	TRMM
CER04	CERES	Single Satellite Cloud-Radiation Pixel Product (CRS)	TRMM
CER05	CERES	Hourly Single-Satellite Flux Product (FSW)	TRMM
CER06	CERES	Monthly & Regional Flux Data Product (SRBAVG)	TRMM
CER07	CERES	Synoptic Cloud-Radiation Data Product (SYN)	TRMM
CER11	CERES	Single Satellite TOA Fluxes, Surface Fluxes, and Cloud Properties (SSF)	TRMM
CER12	CERES	Surface and TOA Fluxes (SFC)	TRMM
CER15	CERES	Zonal, and Global Monthly Average Data Product (ZAVG)	TRMM
CASTR	CERES		TRMM, AM, PM
CCIDA	CERES		TRMM, AM, PM
CCRHA	CERES		TRMM, AM, PM
CER00	CERES		TRMM, AM, PM
ACR01	ACRIM	Level-1A Product, ACRIM	ACRIMSAT
CER01	CERES	Bi-Directional Scan Product (BDS)	AM-1
CER02	CERES	ES-8 ERBE-Like Science Product	AM-1
CER03	CERES	ES-9 ERBE-Like Product	AM-1
CER09	CERES	Single Satellite Instrument Earth Scan Product (IES)	AM-1
CER13	CERES	ES-4 ERBE-Like Product	AM-1
CER14	CERES	ES-4G ERBE-Like Product	AM-1
CER16	CERES	Clear Reflectance and Temperature History (CRH)	AM-1
CER04	CERES	Single Satellite Cloud-Radiation Pixel Product (CRS)	AM-1
CER05	CERES	Hourly Single-Satellite Flux Product (FSW)	AM-1
CER06	CERES	Monthly & Regional Flux Data Product (SRBAVG)	AM-1
CER07	CERES	Synoptic Cloud-Radiation Data Product (SYN)	AM-1
CER11	CERES	Single Satellite TOA Fluxes, Surface Fluxes, and Cloud Properties (SSF)	AM-1
CER12	CERES	Surface and TOA Fluxes (SFC)	AM-1
CER15	CERES	Zonal, and Global Monthly Average Data Product (ZAVG)	AM-1
MIS01	MISR	Level-1A Product	AM
MIS02	MISR	Level-1B1 Product	AM

Table 3.4.2.2-2 Release B Datasets Held by the Data Server Subsystem (2 of 2)

Product	Instrument	Product Description	Platform
MIS03	MISR	Level-1B2 Product	AM
MIS04	MISR	TOA and Cloud Product	AM
MIS05	MISR	Aerosol Product	AM
MIS06	MISR	Surface Product	AM
MIS07	MISR	Gridded TOA and Cloud Product	AM
MIS08	MISR	Gridded Aerosol Product	AM
MIS09	MISR	Gridded Surface Product	AM
MIS10	MISR	Ancillary Geographic Product	AM
MOP01	MOPITT	Level 1B Radiance (Ancilary Data including CR Pressure), MOPITT	AM
MOP02	MOPITT	CH4 Column (Total Burden)	AM
MOP03	MOPITT	CO Profiles	AM
MOP04	MOPITT	CO Column (Total Burden)	AM
MOP05	MOPITT	CH4 Column (Total Burden) Gridded Product (Fourier Coef Form)	AM
MOP06	MOPITT	CO Column (Total Burden) Gridded Product (Fourier Coef Form)	AM
MOP07	MOPITT	CO Column (Total Burden) Gridded Product (Fourier Coef Form)	AM
SAG01	SAGE-III	Level-1B Transmission Profiles (65 wave lengths) Solar SAGE-III	RSA/CNES SPACESTATION
SAG02	SAGE-III	Aerosol Extinction Profiles (at 7 wavelengths)	RSA/CNES SPACESTATION
SAG03	SAGE-III	Cloud Height, Top	RSA/CNES SPACESTATION
SAG04	SAGE-III	H20 Conc & Mixing Ratio	RSA/CNES SPACESTATION
SAG05	SAGE-III	H20 Conc & Mixing Ratio	RSA/CNES SPACESTATION
SAG06	SAGE-III	N03 Conc & Mixing Ratio, Lunar	RSA/CNES SPACESTATION
SAG07	SAGE-III	03 Conc & Mixing Ratio	RSA/CNES SPACESTATION
SAG08	SAGE-III	OCIO Conc & Mixing Ratio	RSA/CNES SPACESTATION
SAG09	SAGE-III	Pressure	RSA/CNES SPACESTATION
SAG10	SAGE-III	Temperature Profile	RSA/CNES SPACESTATION

The key assumptions associated with the DBMS repository sizing are as follows:

- The products lists have been derived from the DAAC instrument teams representatives, the ECS Technical Baseline and in coordination with Science Metadata sizing for the SDSRV DBMS DRPHW CI.
- The period of data capture for Release A products on the TRMM mission is 8/17/97 to 12/31/98.
- The period of data capture for Release B is 1/1/1997 through 12/31/1999.
- All products are assumed to conform to the Proposed ECS Core Metadata Standard v2.0, 420-TP-001-005, Dec. 1994.
- The metadata sizing has been calculated from the Metadata Expected with each granule included in the table on Page 94 in the Proposed ECS Core Metadata Standard v2.0, 420-TP-001-005, Dec. 1994.
  - The calculated size of 1.823 KB per granule has been obtained from this data source.
- An overhead factor of 2.36 for implementation in Illustra has been estimated based on the benchmarking activities as outlined in the DBMS Benchmark Report, 430-TP-003-001. This assumes a high level of query/insertion activity and a low level of update/deletion activity.
- An estimated overhead of 80 MB will be made for the Illustra product code. Note that this includes sizing for the Illustra database product, the 2-D Spatial Data Blade, and the 3-D Spatial Data Blade.
- All instruments on TRMM are considered to exist in the Data Server as a continuation of Release A.
- All instruments on AM-1 were included in the product lists.
- Keyword metadata per document is derived from Release A, phase2 DDSRV. Average size is 1727 bytes.
- Sizing for document metadata and storage is assumed during initial Release B operations because time phasing of document production is indeterminate.
- Sizing includes requirements for a RAID 5 overhead factor of 1.125.

Table 3.4.2.2-3. LaRC DBMS Repository Sizing (GB)

DAAC	1997	1998	1999	2000	2001	2002	EOC
LaRC	0.2	1.5	5.5	9.9	14.1	18.2	21.8

The calculated disk capacity for the LaRC repository (static analysis) results in a computed requirement of 5.5 GB. Due to the operational experience with the user space requirements, at least

5 GB of disk space must be allocated for a high use data base functioning. Therefore, at least 5 GB beyond the calculated storage requirement will be allocated - 10.5 GB. (The closest available quantity of disk equal or exceeding 10.5 GB will be purchased.) Dual host configuration will allow for failover. Host type will be a SGI XL class machine.

Document Data Server — Document handling is handled via a dedicated Data Server implementation, geared to the predicted document ingest and access volumes and the nature of the COTS S/W requirements imposed on the support hardware. The Document Data Server is provided as a simple server configuration with network access. The following assumptions were made in the preparation of the LaRC Document Data Server configuration:

1. Documents and document metadata together have been considered as a basis for the sizing calculations.
2. Totals for MSFC are included in the GSFC totals.
3. Document related data includes: 5 guide document types, algorithm descriptions, production plans and reference papers. These from the DDSRV Detailed Design for ESDTs (DID 305).
4. Guide document sizes (5 guides) from Release A, phase 2 of DDSRV are sized at 1.5 MegaBytes each.
5. Production plans are assumed to be on per DAAC, sized at 1.0 MegaBytes each.
6. Reference papers are sized at 6.5 MegaBytes each; one per instrument per platform.
7. As a typical example of other documents, the CERES ATBD and LaRC Handbook file sizes were used as the basis for the ALGORITHM DOCUMENTATION sizing for each TRMM instrument. One ATBD is assumed per instrument.
8. The creation of an ATDB is dependent on the establishment of a new dataset. Therefore, the accumulation of individual granules has no effect on the growth in the number of ATDBs.
9. Sizing includes requirements for a RAID 5 overhead factor of 1.125.
10. The figures are approximations, which will be refined over time. The Document Data Server architecture is scaleable.
11. The Document Data Server continues to exist as a separate server. As the design effort for Release B continues relative to the Sybase/Illustra selection, the Document Data Server could be collapsed into the DBMS Server. This would change the document storage sizing and cause it to be added into the DBMS Server sizing.
12. The entire document Data Server DBMS sizing is assumed to be available at the beginning of the Release B time frame. This is to allow sufficient capacity to be available to handle both historical document conversion and new document requirements. This includes the maximum required by the end of Release B.

A 2 CPU SMP server was selected based upon operational experience with the EDF EDHS. A WAIS-like, full text indexer, an http server, and additional custom developed software will reside on this host. The following processes/applications run on this host: Document Data Server Process, WWW Server Process, Document Repository Process, Client Applications Process, CSS DCE client, CSS logging API, and an MSS agent.

The disk complement was sized to hold the document metadata for the data product collections associated with the TRMM, AM-1, and SAGE III mission and for the V0 data sets identified for migration. Sizing for document metadata was based on available V0 guide document sizing, and the 2.0 Core metadata baseline. The calculated required disk capacity for all document collection alone is 4.1 GB.

3.4.2.2.2 Configuration

The specific sizing for the Release B LaRC Data Servers, derived from the rationale described above, is synopsized below. Figures (EDS provided figures) provide the a preliminary design for the site's configuration. Additional details on specific component configurations and sizing are provided within the figures.

For the LaRC Data Server:

ACMHW

- Admin. Workstations: 2 ea. of SUN Ultra
- APC Hosts: one 10 CPU SGI Challenge XL, configured with 6 GB local disk and 1 GB RAM, and one identically configured 6 CPU Challenge XL with 512 MB RAM.
- 480 GB RAID

WKSHW

- WKSHW Hosts: one 6 CPU SGI Challenge XL, configured with 6 GB local disk and 512 MB RAM, and one 4 CPU Challenge XL with 256 MB RAM.
- 342 GB RAID

DIPHW

- Staging Server Host: two 4 CPU SUN Ultra 4-slot, with 6 GB of local disk, 256 MB of Ram each, and shared access to 235 GB disk total.
- Standard Release B Peripheral Set: 8mm Tape drives and stackers, 4mm Tape drives and stackers, 6250 Drive, CD-ROM drive and jukebox, FAX/scanner, 3480/3490 outboard drives.

DRPHW

- FSMS Server Host: 6 CPU SGI Challenge XL, configured with 6 GB local disk and 512 MB RAM and 4 CPU SGI Challenge XL, 43 GB RAID.
- Archive Tape Library Robotics: 3 AMASS AML Model 2 Tall Quadro Tower system, 2 robotic arms each

- Tape Drives: 10 - 3590 drives each (30 total)
- Tape Media: 5,751 - 3590 tape cartridges
- DBMS Server: two of 2 CPU SGI Challenge XL, with 12 GB of shared disk

For the LaRC Document Data Server....

- WAIS/http Data Server 2 of 2 CPU SMP Server
- Data Server Disk: 6 GB mirrored in two machines for Release-B

3.4.2.3 Data Management Subsystem

The Data Management Subsystem (DMS) consists of a single Hardware CI (HWCI) that will also support the Release B processing requirements of the Interoperability Subsystem (IOS) at the LaRC DAAC site.

The DMS is responsible for supporting Advertising Service CI, Data Dictionary CI, Gateway CI, Local Information Manger CI and Distributed Information Manager CI processing activities generated directly from user "pull" search invocations. The DMGHW CI consists of three major components: 1) DBMS/Web Server, 2) Database Management Workstation, 3) Data Specialists and User Support Workstations.

The DBMS/Web server is the primary hardware component in the Data Management Subsystem. The server provides DBMS storage, input/output (I/O), and processing resources in support of the Advertising Service CI, Data Dictionary CI, Gateway CI, Local Information Manager CI and Distributed Information Manager CI in Release B.

The DMGHW CI configuration provided in Section 3.4.2.3.2 depicts the migration of the Release A hardware design to the Release B hardware design. The Release B hardware design accommodates Release B platform design issues concerning scalability, RMA and evolvability. The hardware design is tailored to Release B LaRC DAAC specific processing needs in support of Advertising Service CI, Data Dictionary CI, Gateway CI, Local Information Manger CI and Distributed Information Manager CI processing functions. The Release B LaRC DMGHW CI is designed to support TRMM (CERES) and AM-1 (CERES, MISR, MOPITT) mission datasets. Section 3.4.2.3.1 provides the rationale behind the recommended Release B hardware configuration and is subject to change as Data Management software CI (under investigation in the incremental development track) prototyping results become available.

3.4.2.3.1 Rationale

The performance impact on the DMGHW CI DBMS/Web server in Release A has been determined to be negligible. In order to avoid unnecessary swap-outs of hardware after Release A, the DMGHW CI has been sized according to Release B processing predictions; therefore, the DMGHW CI Release B design builds on the Release A design without the need for large scale changes. The performance drivers for sizing the DMGHW CI server for Release B are:

- User Characterization analysis of science and non-science user search invocations

- DBMS/CI transaction rate (performance) analysis
- DBMS/CI prototype/benchmark analysis
- Hardware Scalability/RMA/Evolvability Analysis

User Characterization Analysis: User Characterization data provides the projected number of science and non-science users, frequency of search invocations per time period, and the percentage of invocations for different types of searches to be supported in the Release B time frame. In Release B, it is expected that science users will primarily access the DM services (Gateway CI, Data Dictionary CI, Local Information Manager CI, Distributed Information Manager CI) while the bulk of the accesses to the Interoperability service (Advertising Service CI) will originate from within the non-science community.

It is assumed that non-science users will access the Advertising Service CI 86% of the time, and DM CIs 14% of the time as documented in "User Characterization and Requirements Analysis" (19400312TPW). The number of searches per hour being processed by the DM CIs in response to queries by science users is based on the ECS science user scenarios in which users are assumed to be accessing the system through the client. Because of the increasing popularity and ease of use of the WWW, it is also expected that science users will make use of the Advertising Service at a rate equal to 25% of DM searches. Data provided by the User Characterization Team apply to epoch m (first quarter of 2000) since the data are meant to represent maximum usage loading (this will occur at the end of Release B).

Tables 3.4.2.3-1 and 3.4.2.3-2 summarize the number of science user system accesses per day and the fraction of invocations per search type for DM services as documented by the User Characterization Team. Table 3.4.2.3-3 summarizes the total number of searches per hour for the busiest hour of the day.

Table 3.4.2.3-1. LaRC User System Accesses per Day for Science Users (Epoch m)

DAAC	User System Accesses per Day
LaRC	352

Table 3.4.2.3-2. LaRC Science User Search Types for Gateway CI Services (Epoch m)

Search Type	Fraction of Total Invocations
Simple Search/1 site	.263
Simple Search/multi-site	.279
Match-up Search/1 site	.272
Match-up Search/multi-site	.185
Coincident Search/1 site	0.0
Coincident Search/multi-site	.000374

Table 3.4.2.3-3. LaRC Searches per hour for Science Users (Epoch m)

DAAC	Searches per hour (busiest time of day)
LaRC	42

The data for searches submitted by science users is categorized into six different types (simple/1, simple/multi, match-up/1, match-up/multi, coincidence/1, and coincidence/multi) for DM services; however, there are only three types of searches for the non-science user data. Each of the three search types that exist in the non-science user data was subdivided into one-site vs. multi-site by applying the proportions of one-site vs. multi-site that exist for science users to the non-science user data. For example, the relative proportions of simple search/1 site and simple search/multi-site for science users is 0.263 and 0.279, respectively. The number of simple searches submitted by non-science users was divided into one-site and multi-site using these same proportions. Only one search type (simple/1) pertains to the Advertising Service CI.

Tables 3.4.2.3-4 and 3.4.2.3-5 summarize the number of non-science user system accesses per day and the fraction of search invocations for DM services as documented by the User Characterization Team.

Table 3.4.2.3-4. LaRC User System Accesses per Day for Non-Science Users (Epoch m)

DAAC	User System Accesses per Day
LaRC	850

Table 3.4.2.3-5. LaRC Non-Science User Search Types for Gateway CI Service (Epoch m)

Search Type	Fraction of Total Invocations
Simple Search/1 site	.31
Simple Search/multi-site	.29
Match-up Search/1 site	.09
Match-up Search/multi-site	.06
Coincident Search/1 site	0.0
Coincident Search/multi-site	.25

DBMS Transaction Rate Analysis: In order to size the DBMS server it is necessary to estimate the size of the Interoperability (Advertising Service CI) and Data Management (Gateway CI, DDICT CI, LIMGR CI, DIMGR CI) services and then determine the transaction rates, or database throughput that must be provided in support of the "pull" search activities that will be invoked by the user community. The transaction rate analysis is based on assumptions regarding the amount of processing associated with the different types of search requests that pertain to the Interoperability and Data Management software CIs. Release B transaction assumptions were made to define a transaction loading value per search request. Number and type of search request are provided by the User Characterization Team. Depicted transaction loading values are assumptions that are based on search complexity. The loading values for search requests will be

refined with actual performance benchmarks as future prototypes are completed. The observed transaction loading from future prototyping/benchmarking activities will be compared to the predicted ones (documented below) and the sizing analysis will be updated as a result (these transaction loading assumptions are defined as "nominal" cases). The transaction data provided is projected for the Release B time-frame.

The processing (transactions per search invocation) assumptions are based on preliminary transaction analysis results for the Advertising Service and DM CIs and will be revised based on future prototyping/benchmarking results as they become available. Tables 3.4.2.3-6 and 2.4.2.3-7 list the predicted transaction load associated with the Gateway CI and Advertising Service CI based on frequency of science user search invocations. Searches/hour are calculated for the busiest time of day at the LaRC DAAC site.

Table 3.4.2.3-6. LaRC Science User Transaction Analysis for Gateway CI Service (Epoch m)

Search Type	Percentage Invoked	Searches/hour	Processing Assumptions (Transactions per Search Type)	Transactions/ hour
Simple Search/1 Site	26.3	11	5	55
Simple Search/Multi-Site	27.9	12	20	240
Match-up Search/1 Site	27.2	11	5	55
Match-up Search/Multi-Site	18.5	8	20	160
Coincident Search/1 Site	0.0	0	5	0
Coincident Search/Multi-Site	.0374	.02	25	.5

Table 3.4.2.3-7. LaRC Science User Transaction Analysis for Advertising Service CI (Epoch m)

Search Type	Percentage Invoked	Searches/hour	Processing Assumptions (Transactions per Search Type)	Transactions/ hour
Simple Search/1 Site	100	11	5	55

Tables 3.4.2.3-8 and 3.4.2.3-9 list the transaction analysis for the Gateway CI and Advertising Service CI based on daily non-science user accesses as depicted in "User Characterization and Requirements Analysis" (19400312TPW). The estimated total non-science user system accesses per day for the Gateway CI and the Advertising Service CI is estimated to be 850. The percentage that each search type, pertaining to the Gateway CI and Advertising Service CI, is invoked is also taken from the same document. Tables 3.4.2.3-8 and 3.4.2.3-9 are completed with the assumption that there will be at least ten search invocations per non-science user access on average.

Table 3.4.2.3-8. LaRC Non-Science User Transaction Analysis for Gateway CI Service (Epoch m)

Search Type	Percentage Invoked	Searches/hour	Processing Assumptions (Transactions per Search Type)	Transactions/ hour
Simple Search/1 Site	31	15	5	75
Simple Search/Multi-Site	29	14	20	280
Match-up Search/1 Site	09	4	5	20
Match-up Search/Multi-Site	06	3	20	60
Coincident Search/1 Site	0.0	0	5	0
Coincident Search/Multi-Site	25	12	25	300

Table 3.4.2.3-9. LaRC Non-Science User Transaction Analysis for Advertising Service CI (Epoch m)

Search Type	Percentage Invoked	Searches/hour	Processing Assumptions (Transactions per Search Type)	Transactions/hour
Simple Search/1 Site	100	305	5	1525

Preliminary transaction analysis results for Interoperability and Data Management CI processes are depicted in Table 3.4.2.3-10. At this time the Data Dictionary CI transaction analysis is equated to the transaction analysis of the Advertising Service CI since the type and frequency of transactions are predicted to be very similar. Due to the fact that DIMGR and LIMGR CI processes are perceived as being the most expensive (in terms of cost to the CPU), and that prototypes will not be developed until after Release B CDR, preliminary sizing estimates have been achieved by doubling the transaction load of the Gateway CI and applying the result to LIMGR CI and DIMGR CI processes. Although equating Advertising Service CI and Data Dictionary CI processes, and doubling the Gateway CI transaction load to produce LIMGR CI and DIMGR CI transaction loading does not pin-point the performance cost that will be levied on the DMGHW CI, it does provide preliminary, expected CPU activity in the absence of real-time prototyping/benchmarking results. The preliminary transaction results for the Advertising Service CI, Gateway CI, Data Dictionary CI, LIMGR CI and DIMGR CI will be revised based on future prototyping/benchmarking analysis results as the Incremental Track software design matures.

Table 3.4.2.3-10 summarizes the science, and non-science user transaction loading per hour for the Advertising Service CI , Gateway CI, Data Dictionary CI, LIMGR CI and DIMGR CI services.

Table 3.4.2.3-10. DBMS Transaction Analysis Summary (Epoch m)

DAAC	User Type	Service	Searches/hour	Transactions/ hour	TPM
LaRC	Science	Gateway	42	511	9
LaRC	Science	Advertising	11	55	1
LaRC	Science	Data Dictionary	11	55	1
LaRC	Science	LIMGR	84	1022	17
LaRC	Science	DIMGR	84	1022	17
LaRC	Non-Science	Gateway	48	735	12
LaRC	Non-Science	Advertising	305	1525	25
LaRC	Non-Science	Data Dictionary	305	1525	25
LaRC	Non-Science	LIMGR	96	1470	25
LaRC	Non-Science	DIMGR	96	1470	25
Totals:			1082	9390	157

A sensitivity analysis has been performed with larger loading allocations; the results are depicted below in Table 3.4.2.3-11.

Table 3.4.2.3-11. DBMS Transaction Sensitivity Analysis Results (Epoch m) (loading has been doubled for both search invocations and transaction rates)

DAAC	User Type	Service	Searches/hour	Transactions/ hour	TPM
LaRC	Science	Gateway	84	2042	34
LaRC	Science	Advertising	22	220	4
LaRC	Science	Data Dictionary	22	220	4
LaRC	Science	LIMGR	168	4084	68
LaRC	Science	DIMGR	168	4084	68
LaRC	Non-Science	Gateway	96	2940	49
LaRC	Non-Science	Advertising	610	6100	102
LaRC	Non-Science	Data Dictionary	610	6100	102
LaRC	Non-Science	LIMGR	192	5880	98
LaRC	Non-Science	DIMGR	192	5880	98
Totals:			2164	37550	627

Future Incremental Track Development prototyping/benchmarking activities will provide a more detailed performance analysis of Advertising Service CI, Gateway CI, Data Dictionary CI, Local Information Manager CI and Distributed Information Manager CI processes; therefore, performance transaction analyses will be revised accordingly.

DBMS Prototyping/Benchmarking Analysis: Currently, preliminary Incremental Track Development performance data is being used to size the processing capacity of the DMGHW CI

DBMS/Web server. Performance analysis results will be revised as planned prototyping/benchmarking activities are completed. Major prototyping activities that will affect performance estimates for the DMGHW CI include, but are not limited to: 1) Prototype workshop 2, and 2) EP7 prototype.

DBMS performance estimates provided in "DBMS Benchmark Report" technical paper (430-TP-003-001), show that for multi-user (32 users) queries (20 similar queries accessing different parts of the test database) running concurrently, the test-bed platform's CPU became saturated (SUN SPARCstation 20/50). A vendor supplied TPM benchmark for the selected platform (HP K400) for Release B operations is shown in Table 3.4.2.3-12. As a rule vendor supplied Transaction Per Second/Minute (TPS/TPM) ratings tend to be a maximum, or high-end value and do not take into account processing overhead associated with other system processes. Processes that will run on the DMGHW CI in Release B include DCE client, MSS agent, HTTP server, Sybase SQL Server, Sybase Replication Server, Sybase Backup Server, Operating System Services, Advertising Service Server, Gateway Server, Data Dictionary Server, LIMGR Server and DIMGR Server.

Table 3.4.2.3-12. Vendor Platform Performance Estimates

Platform	TPM	MIPS
HP K400 (SMP) with 1 processor (PA 7200 CPU)	1000	146

Disk Capacity Sizing: Disk storage for the DMGHW CI has been determined for each DAAC site based on preliminary Interoperability and Data Management CI DBMS application sizing estimates plus vendor inputs for the following COTS software: 1) DBMS software, 2) Development software, 3) HTTP server software, 4) Operating System software, 5) Communications and Utilities software. Capacity sizing for the Interoperability and Data Management databases was achieved by multiplying the expected byte size for core and collection specific attribute definitions by the total number of core and collection specific attributes. Temporary workspace has also been allocated for Interoperability CI and Data Management CI services dependent on frequency and variation of queries. For example, the DIMGR service has a larger capacity requirement than the Data Dictionary service because it requires more temporary workspace since it handles a greater number and more varied types of queries. The expected capacity of Interoperability CI, Data Management CI, COTS, Operating System and Communications and Utilities software to be installed on the DMGHW CI at the Release B LaRC DAAC site is depicted in Table 3.4.2.3-13.

Table 3.4.2.3-13 is filled with preliminary disk sizing results for Release B Interoperability and Data Management software CIs, operational databases and COTS software packages that will be installed on the DMGHW CI at the LaRC DAAC site. Some of the results are estimates (such as database sizes) since the DBMS design will mature and impact disk capacity sizing as the Advertising Service CI, Gateway CI, Data Dictionary CI, Local Information Manager CI, and Distributed Information Manager CI under-go future evaluation and prototyping.

Table 3.4.2.3-13. DMGHW CI Disk Capacity Requirements

S/W Component	Release B Capacity
COTS Software:	
Sybase System	300 MB
HTTP Server	10 MB
	Total: 310 MB
Databases:	
Sybase Master Database	3 MB
Sybase Tempdb Database	100 MB
Sybase Model Database	2 MB
Advertising Database	150 MB (Estimate)
Advertising DB Workspace	150 MB (Estimate)
Advertising DB Log	100 MB (Estimate)
Advertising HTML Files	100 MB (Estimate)
Data Dictionary Database	400 MB (Estimate)
Data Dictionary DB log	100 MB (Estimate)
DMS Working Store Database	500 MB (Estimate)
DMS Working Store DB log	100 MB (Estimate)
	Total: 1705 MB
Operating System & Utilities:	
Operating System Software	700 MB
Utilities	200 MB (Estimate)
DCE Client	46 MB
	Total: 946 MB
	Total: 2961 MB

3.4.2.3.2 Configuration

The selected DMGHW CI DBMS/Web server to be implemented in Release B is a low-end SMP server (HP K400) that is scaleable from one to four processors. A single physical DBMS/Web server will be implemented at the Release B LARC DAAC site. At this time a two CPU configuration has been determined to be appropriate for the DMGHW CI DBMS/Web server due to RMA requirements. In Release A the configuration called for a primary, or "active" server and secondary, or "standby" server, each configured with a single CPU. For Release B, the configuration will be changed to a single host server configuration with redundant CPUs (see Figure 3.4.2-2). Note that for Release A an HP K200 low-end SMP server (primary/active) was selected for the DMGHW CI platform. In order to accommodate redundant network and FWD SCSI-2 interface cards and a redundant power supply unit at Release B the HP K200 will be upgraded to an HP K400. The upgrade will be completed to obtain the additional slots for the network and FWD SCSI-2 cards and power supply only, there is no difference in the scalability, or type of CPUs between the K200 and K400; therefore, the upgrade does not consist of “swapping out” hardware; the same physical servers that were used in Release A DM configurations will be upgraded. Also, since a single server is called for in the Release B design and not redundant servers, the redundant (secondary) K200 servers located at Release A DAACs

(LaRC, GSFC, EDC) will be upgraded to K400 servers and re-located to Release B DAACs as primary servers.

The Release B HP K400 single server configuration will offer redundancy in the form of: 1) dual processors, 2) dual power supply units, 3) dual FWD SCSI-2 cards, and 4) dual FDDI network cards 5) duplicate OS boot/application disk. The components listed above are hot swap-able units which allow them to be replaced without shutting down the server. Also, the HP-UX operating system features memory page de-allocation which automatically blocks out any portion of memory in which an error has been detected; therefore, a failure to memory will not bring operations to a halt. The single server host configuration will allow applications to be run in parallel across both processors, which enhances load balancing and availability/recovery capabilities. The DMGHW CI DBMS/Web server will automatically reconfigure itself in a single CPU configuration in the event of failure to a single CPU. The dual power supply units, FWD SCSI-2 cards and FDDI network cards will also provide continued availability in the case of failure to a single component (per function). The redundant configuration of the DMGHW CI DBMS/Web server has been analyzed by the ECS Reliability Engineering Group using COTS vendor provided data to ensure that all functional availability requirements are met. Preliminary analysis results revealed that the DMGHW CI DBMS/Web server has an MTBF (mean-time-between-failures) of greater than 20,000 hours which meets all pertinent RMA requirements.

A RAID disk unit will be attached via FWD SCSI-2 (dual ported) to the DMGHW CI DBMS/Web server (see Figure 3.4.2-2). The RAID disk until will provide operational and mirrored sets of disk devices to the DMGHW CI server in order to provide uninterrupted data availability in the event of a disk failure. The RAID disk configuration will also be implemented such that a failure to a single disk will be recoverable via a "hot-swap" disk capability. The RAID disk until will be comprised of ten disks: five operational and five mirror disks. Each disk contained in the RAID unit will provide a capacity of 2.1GB.

Since the Release B processing requirements for the Local Information Manager CI and Distributed Information Manager CI are largely unknown at this time, the flexibility of the recommended hardware design assures minimum risk. The design allows for 100% growth in both processing and storage capacity in support of the Advertising Service, Gateway, Data Dictionary, Local Information Manager and Distributed Information Manager CIs in Release B.

The following configuration diagram, Figure 3.4.2-2, depicts the recommended server host/RAID disk unit configuration:

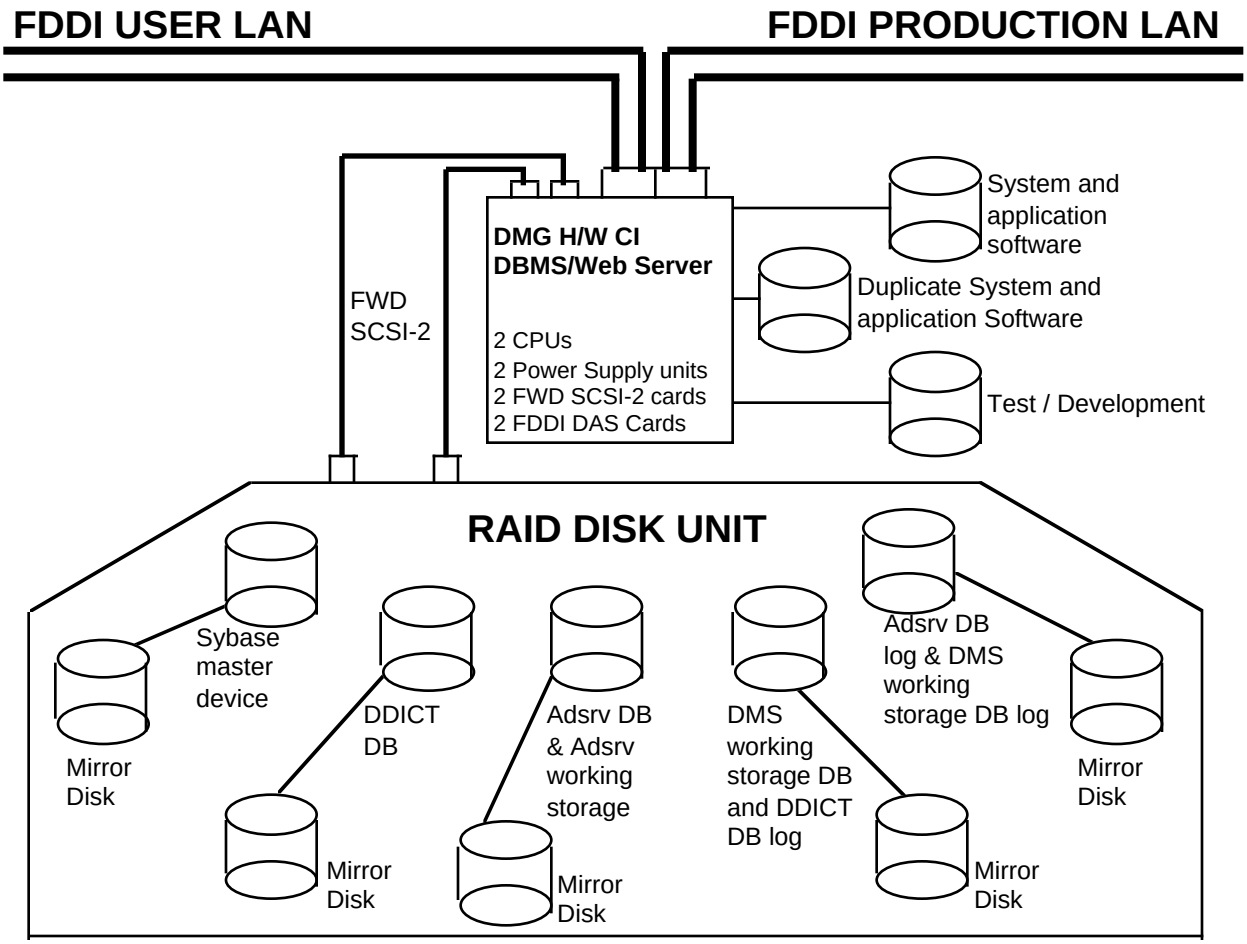

Figure 3.4.2-2. DMGHW CI RAID Disk Configuration

Although the estimates for Advertising Service and Data Management operational databases, COTS software, and operating system and utilities sizing are relatively small, the total disk volume for the DMGHW CI server has been increased in support of Sybase swap (workspace) area, 100% growth capacity for the core operational software, and mirror disk units. The Data Management databases will be replicated (using Sybase replication server) at each DAAC site. The Interoperability and Data Management software CIs will be distributed over multiple disk drives contained in the RAID disk unit in order to enhance performance (vendor and developer recommended); therefore, the total disk volume is well above the actual capacity needed to support the software alone. Additional disk capacity (internal to the server) in the form of a redundant system/application disk, and a test/development disk have also been added to the configuration in support of RMA, and integration and test requirements.

A single 8mm tape drive unit will be configured on the DMGHW CI server in Release B. The 8mm tape drives will be used to backup Advertising Service and Data Management CIs/databases, as well as perform DBA and routine maintenance operations.

A low-end DBMS/DBA uniprocessor workstation will be used for database/system administration activities. A single 8mm tape drive unit will be configured on the DBMS/DBA workstation in Release B. The 8mm tape drive will perform backup/recovery and routine maintenance operations in support of the DMGHW CI DBMS/Web host servers. A small pool of low-end uniprocessor workstations will support Data Specialist/User Support operations. At a minimum the DBMS/DBA and Data Specialist/User Support workstations will be configured with six gigabytes of local disk each. Workstation disk capacities were sized based on IR-1 workstation installation results which included recommendations for additional space per workstation to accommodate the following: 1) User/Temp workspace, 2) Personal development space, 3) Testing, 4) Software upgrades, 5) Working with large files/datasets, 6) future growth/flexibility. Exact capacities for disk drives are dependent on the procurement process as the type, and size of disks being offered for workstation platforms may fluctuate.

Table 3.4.2.3-14 summarizes the recommended DMGHW CI processing configuration for implementation at the Release B LARC DAAC.

NOTE: The HP K400 SMP depicted in table 3.4.2.3-14 is a low-end SMP class server that is scaleable from 1-4 CPUs.

Table 3.4.2.3-14. LaRC DAAC Hardware Configuration

Component	Class/Type	Platform	Qty.	Number of Processors	Memory	Disk Capacity
DBMS/Web Server	SMP	HP K400	1	2	512 MB	27.3 GB
DBA Workstation	Uniprocessor	HP 715/64	1	1	64 MB	6 GB
Data Specialist and User Support Workstations	Uniprocessor	SUN Sparc 20/50	3	1 (each)	64 MB (each)	6 GB (each)

<u>DBMS/Web Server Platform Technical Specifications:</u>

Make: Hewlett Packard

Model: K400 (Low-End SMP class server)

CPU: PA7200 (upgradeable to future PA800 processor)

Clock Frequency: 100 MHz

Number of processors: 1 to 4

MIPS: 146 (1 processor)

TPM: 1000 (1 processor) 3160 (4 processors)

SPECint92: 136 (1 processor)

SPECfp92: 215 (1 processor)

Memory: Expandable to 2GB RAM

Internal Processor-Memory bus bandwidth: 960 MB/sec (peak)

I/O Bandwidth: 128 MB/sec (peak)

3.4.2.4 Ingest Subsystem

Ingest subsystem hardware at LaRC is responsible for the ingest, preprocessing, and storage of CERES, MISR, MOPITT, and SAGE III Level 0 data products. Ingest subsystem hardware will also be involved in the migration of existing Version 0 data into ECS Release B. Subsystem configuration and specific component sizing rationale are provided in the following paragraphs.

3.4.2.4.1 Rationale

The sizing of Ingest Subsystem hardware both from a system level and a component level is based on the 2/7/96 version of the ECS Technical Baseline. Among the information included in the baseline is:

- data by instrument,
- average daily data volume by level,
- and data destination.

Table 3.4.2.4-1 provides a synopsis of the Ingest data volumes required for Release B at LaRC.

Table 3.4.2.4-1. LaRC DAAC Specific Ingest Volume Requirements

Release B DAAC	Daily L0 Volume (GB/day)	Annual L0 Volume (GB/year)	Version 0 Data Migration Volume (GB)
LaRC	47.6	17,374	354

The average expected daily and annual data volumes at each site were calculated from this information and used to determine the required ingest hardware capabilities. Ingest client hosts are sized to accommodate the required ingest volumes as well as I/O and CPU capabilities to support internal data transfers associated with metadata validation and extraction and transfer of data to the Data Server or Processing Subsystems. Working storage disks are sized to accommodate the above functions, as well as provide contingency space for the transfer of more than one days worth of data within a 24-hour period. Since high RMA is a driver for the Ingest Subsystem, all critical components also include some type of sparing or redundancy to ensure that availability requirements are met.

Queuing Analysis Model An Ingest Queuing Model (Imodel) was developed to assist in the sizing of Ingest Subsystem components for the Release A LaRC configuration. This model was also used in a preliminary assessment of the Release B volumes. However, the additional interfaces and increased data volumes to be supported in Release B require the use of a dynamic model to better depict the flow of data into, within, and out of the Ingest Subsystem. A more detailed and comprehensive modeling effort will be completed prior to the Release B CDR. The

configuration presented in the following section provides an estimate of the types and classes of hardware required at the DAAC based on the basic paper analysis and queuing model estimates.

3.4.2.4.2 Configuration

The LaRC configuration for Release B is sized for the TRMM, AM-1, and METEOR missions for one year of operations. Brief descriptions of the generic components, provided within the LaRC configuration at Release B are provided in Table 3.4.2.4-2. The results of the system design analysis results in a recommended configuration consisting of mid-level Symmetric Multi-Processing (SMP) 32-bit machines, capable of supporting multiple network (FDDI and/or HiPPI) and direct-connect (SCSI II) devices.

Working storage devices are RAID 5 units with a minimum of two days worth of space allocated to ingest working storage required to support the functions of acquiring, processing, validation, and archiving L0 data. This volume of working storage allows for one days worth of L0 data to be staged for processing, an additional days worth available for subsequent ingest, and an additional 25% available to service additional Ingest Subsystem needs (e.g., retrieval support, pre-processing, quality checking). Additional magnetic disk resources are supplied within the Ingest Subsystem to support items such as:

- client host operating system
- application software and L0 archive database directory information

Table 3.4.2.4-2 Ingest HWCI Component Descriptions

Component Name	Class/Type	Comments
Client Host	SMP Server W/S	Mid-level SMP servers.
Working Storage	RAID disk	Multiple RAID units, site capacity sized. RAID 5 for random access protocol.
L0 Repository	Archive robotics and tape drives	Single robotics unit with multiple tape drives.

The specific sizing derived from LaRC Release B requirements is synopsized within Table 3.4.2.4-4 and is highlighted in Section 3.4.2 within the site configuration overview Figure 3.4.2-1. The site overview figure provides additional details on specific component configurations and sizing. The Ingest Subsystem resources at LaRC are sized principally for the operational ingest and storage of Level 0 data sets, and to support the migration of data products from Version 0. Additional available system resources are for contingency.

Table 3.4.2.4-3. Ingest HWCI Component Sizing for the LaRC DAAC Configuration

Ingest Component	Component Class	Quantity	Comments
Client Host (ICLHW)	SGI Challenge L SMP	2	L0 Ingest Client hosts. Hosts are adapted to ECOM I/F and ESN. Host attached disk. SCSI I/Fs to RAID working storage.
Working Storage (ICLHW)	RAID Disk (host attached)	2	Host adapted RAID disk arrays. RAID 5. SCSI II adapted & cross strapped to Ingest Client hosts.
L0 Repository (ICLHW)	Archive robotics	1	Robotics unit sized to store one year's worth of Level 0 data plus a small amount for long-term storage of L0 data without L1A counterpart (e.g., MOPITT).
L0 Repository (ICLHW)	Archive tape drives	2	Drives mounted in archive robotics unit to support writing and retrieval of L0 data.
Client Host (ICLHW)	8 mm tape stacker	2	Support for hard media ingest from Version 0.
Client Host (ICLHW)	X-Terminal	1	OPS support for Data Ingest Technician(s).

3.4.2.5 Interoperability Subsystem

For the Release B time frame, the hardware support for the Interoperability Subsystem, particularly the Advertising capabilities are provided by the Data Management HWCI. Please see Section 3.4.2.3 for a complete description of this capability.

3.4.2.6 Production Planning Subsystem

The Release B SDPS Planning Subsystem Design Specification, 305-CD-026-002, Section 5, describes the functionality and architecture of the Planning Subsystem (PLS) Hardware Configuration Item (PLNHW CI). Because of the close coupling of scheduling/queuing functions with planning (i.e., use of the PDPS database), Section 5, also addresses the scheduling/queuing hardware within the Data Processing Subsystem (DPS) Science Processing Hardware Configuration Item (SPRHW CI).

This section describes the sizing of the Planning Subsystem (PLS) Hardware Configuration Item (PLNHW CI) for the LaRC DAAC site. Because of the close coupling of scheduling/queuing functions with planning (i.e., use of the PDPS database), this section also addresses the scheduling/queuing hardware within the Data Processing Subsystem (DPS) Science Processing Hardware Configuration Item (SPRHW CI).

The LaRC PLNHW CI consists of two Production Planning/Management SMP workstation servers which run the Planning Workbench software and the PDPS database. A Production Planning/Management workstation supports the planning operations staff in performing their routine production planning and management functions. A workstation is provided for operations personnel to access production planning GUIs via the Planning Workbench application. These functions include candidate plan creation, plan activation, entry of production request information and report generation.

The Planning Subsystem is responsible for maintaining the Planning and Data Processing (PDPS) database. This database serves both planning and scheduling. Planning uses the database in support of the Planning Workbench application used for production planning and resource planning activities. Scheduling uses the database in support of the AutoSys application which runs the production schedule.

The LaRC scheduling/queuing hardware in the SPRHW CI consists of two Queuing Servers SMP Workstations which run AutoSys and the AutoSys database. Production Monitor Stations run the AutoSys AutoXpert display GUIs.

The scheduling/queuing portion of the DPS is responsible for maintaining the AutoSys database. An active plan is loaded into the AutoSys database from the Planning Subsystem. AutoSys then proceeds to execute this production schedule (the result of activating this plan). AutoSys uses both the PDPS database and its own database during its execution of the production schedule.

Both the PDPS and AutoSys databases are implemented using Sybase. The PDPS database is used for the persistent storage of critical data while the AutoSys database contains the current data for executing a daily production schedule.

3.4.2.6.1 Analysis

Hardware Design Drivers

Drivers for sizing the hardware to support these functions include:

- database transaction rates,
- replication of critical data,
- volume of production status to display,
- job throughput,
- the need for a fail-soft science processing architecture.

The primary drivers for sizing the PDPS database are data criticality, support of science processing's fail-soft environment, and the volume of database transactions. The PDPS data is critical because it consists of invaluable PGE information and run-time statistics used for the creation of future plans. The PDPS database supports AutoSys in execution of the production schedule. If the database fails, the launching of production jobs by AutoSys will stop. The PDPS database also needs to support multiple database accesses from both planning and processing.

The primary drivers for sizing the AutoSys database are support of science processing's fail-soft environment, job throughput, and the display of production status. AutoSys launches jobs which stage data, run PGEs, report PGE status, and destage data. If AutoSys fails, no new production jobs will be executed and production will stop. AutoSys must also be able to execute enough jobs (throughput) to process the daily production schedule within the allocated time frame (8 or 24 hours which varies by DAAC). AutoSys' production status also needs to be displayed without saturating the person monitoring production.

Assumptions

The sizing of the planning and scheduling hardware is based on an analysis of the PDPS database transactions and Planning Workbench functions, AutoSys job throughput, and AutoXpert display requirements. The following paragraphs summarize the assumptions used in this analysis.

1. *Technical Baseline* -- The February 1996 technical baseline information was used to derive key parameters that effect the planning and processing workload at each DAAC site. These key parameters include the number of job activations per day for each PGE at each DAAC and the number of PGEs maintained at each DAAC. For example, at LaRC during the Release B period (Epoch k, 3Q99), the following parameters were derived:

   - No. PGEs at the DAAC 94
   - Number of PGE activations per day 778
   - Total number of jobs required, including reprocessing of 200% 9,336

2. *OPS Concept* -- These are the operational assumptions.

   a. The production scheduler/planner at each DAAC will prepare and publish a 30 day plan every two weeks. This plan is used to provide some assistance in longer range planning. The 30 day plan is only prepared and published every two weeks - if changes occur after a plan is published, the changes will only be incorporated in the next 30 day plan.

   b. The production scheduler/planner at each DAAC will prepare and publish a 10 day plan every week. This plan will provide a finer grain description of planned activities. Like the 30 day plan, this plan will not be replanned and distributed except on the regular weekly boundaries.

   c. The production scheduler/planner at each DAAC will prepare and activate a daily plan or schedule once per day. This schedule will be replanned as required during the course of the day.

   It needs to be noted in this regard, and as described in more detail in the Planning Subsystem design document, that the manner in which each DAAC operations production scheduler/planner decides to conduct planning and scheduling can vary from DAAC to DAAC. The planning and scheduling tools are sufficiently flexible to support a variety of planning and scheduling strategies. The sizing of the storage and the performance of the server is based upon the assumptions given here. It should also be noted that it is not essential that the operations staff conduct replanning for the active schedule if events arise

(e.g., processor failures) that would cause the predictions of the active schedule to depart from reality. Processing will continue regardless since jobs will be released for execution as resources (processors and input data) become available.

3. It was assumed that the reprocessing workload is equal to (1x) the standard processing workload in the Epoch k time frame and that the reprocessing workload is two times (2x) the standard processing workload in the Epoch o time frame (near the end of Release B). The hardware selected for LaRC is sized to support 2x reprocessing.

4. Several other parameters were estimated as part of the process of developing a model for the database sizing and processing activity. Some of these parameters and their values are:

   - No. of working hours per day 24
   - No. of long term (30 day) plans stored 5
   - No. of short term plans stored 5
   - No. of production requests per PGE 4
   - No. of replans (active plans) per day 2

5. *Planning Activities* -- The design for the Planning Subsystem was represented as a collection of several activities or functions (e.g., Subscription Manager, Planning Workbench: Candidate Plan Creation, etc. See Planning Subsystem Design). Functions that were activated only infrequently (e.g., Subscription Submittal) were ignored. The activation frequency of these subsystem activities was then identified, where possible from the February 1996 technical baseline summary (e.g., number of jobs activations per day). Subactivities within these functions were identified and an estimate of the processing load was made based upon the complexity of the activity. These processing load estimates were made in terms of the number of transactions (updates, inserts, and reads) made to the PDPS database. These load estimates were then totaled across the complete set of Planning Subsystem activities for the specified frequency of activation for each DAAC site. The PDPS DBMS server sizing was then estimated by scaling upwards from the base processor size.

PDPS Transaction Rate Analysis

Database transaction rates were modeled for the PDPS database by estimating the number of updates, inserts, and reads made to the database when performing planning and scheduling functions. These functions include:

- Creating Data Processing Requests to be added to AutoSys database (DPRs),
- Release of DPR jobs for execution,
- Updating Data Processing Request Job,
- Getting DPR Job Status,
- Data Initialization,

- Local Data Management,
- Data Staging,
- Data Destaging,
- Initiate PGE execution,
- Monitor Execution,
- Execute Post Processing,
- Plan Creation/Activation,
- Subscription Notifications, and
- On-Demand Production Requests.

For each of the activities listed above, the number of average daily transactions (updates, inserts, and reads) made to the PDPS database was estimated based on the number of PGE activations per day. The average daily transaction load at LaRC based on 778 PGE activations per day is 2,000,000 (including Sybase replication). This would require 23 transactions per second (TPS). The peak transaction load is 1,751,000 worse case, to create a 30-day plan.. The transaction rate of a Sun 75 MHz 2-CPU 20/712 machine is 305 TPS. It would therefore take approximately 96 minutes to create a 30-day plan containing 778 PGE executions at LaRC. A 10 day plan containing 778 PGE executions would take approximately 22 minutes and a 1 day plan would take approximately 7 minutes.

PDPS Database Sizing

The PDPS DBMS server supports the PDPS database that contains all the information central to the functioning of the Planning Subsystem. To size the PDPS DBMS server it was necessary to estimate the number of inserts made into the database in order to determine how quickly and by how much the database would grow each day.

The database classes (tables), associated with the PDPS Transaction Rate Analysis functions above, are:

- DpPrExecutable,
- DpPrDataMap,
- DpPrPGE,
- PlProductionRequest,
- PlDpr,
- PlDataGranule,
- PlODProdRequest.

The number of inserts made to the database are from the PDPS Transaction Rate Analysis. The number of inserts (additional rows added to the Sybase tables) were then input into a Sybase system procedure called sp_estspace. This procedure returns the estimated megabytes of disk space required to store a particular number of rows for a given class (table) above.

The number of PGE activations for LaRC is 778. The disk capacity required to store 90 days worth of PDPS database growth is estimated to be 780 MB.

Planning Hardware

The Production Planning Subsystem is made up of two 20/712 Sun SMP workstations. Each workstation is attached to two 6.3 GB Sun SPARCstorage MultiPacks (external disk) via fast wide SCSI controllers. Each 20/712 machine will run the Planning Workbench software and Sybase (the PDPS DBMS). Both machines can use the Planning Workbench, one for Production Planning, the other for Resource Planning. Sybase will run on one machine while Sybase replication is run on the second machine, thus creating a copy of the PDPS database and allowing automatic switch over to the replicated database upon a failure of the primary PDPS database. Sybase replication ensures the integrity of the critical data that the PDPS database contains.

The Sun operating system and applications will reside on disk internal to each of the 20/712 machines. The Sybase log, index, and data files will be on external disks. Two sets of three 2.1 GB external disks will be accessible via 2 fast wide SCSI controller cards for each machine.

Each 20/712 machine can be upgraded to use emerging hyperSPARC technology CPUs running at 150 MHz. Future experience with the Planning Workbench software and the PDPS Database will determine if this extra capacity will be needed.

Each SPARCstorage MultiPack disk can be increased to a maximum of 12 disks for a total of 25.2 GB if required.

AutoXpert Display Sizing

Two pairs of X-terminals (in addition to the two workstation's displays) will be used to show production status by using the AutoSys AutoCons' display and the AutoXpert JobScape GUI. Additional AutoXpert GUIs: TimeScape and HostScape are also available. If an operator wants to view the progress of CERES (AM) production, he would modify his view of the entire active queue to display only those jobs associated with CERES (AM). To properly do this requires two displays, one for AutoCons and another for JobScape. The other two pairs of displays can be used to display TimeScape and HostScape for the same instrument's production status or another instrument's production could be viewed or maybe the progress of reprocessing jobs could be displayed. Table 3.4.2.6-1 shows the contribution that each instrument's production (PGE activations) makes on total daily production at LaRC. Reprocessing at Epoch k (3Q99) will be approximately 100% of normal production growing to 200% of normal production by the end of Release B.

Table 3.4.2.6-1: Breakdown of Instrument Production at LaRC

Instrument	Number of Instrument PGE Activations per day/total Number of all PGE Activations per day
SAGE	.13%
MOPITT	.55%
CERES (AM)	13.28%
CERES (TRMM)	13.27%
MISR	72.77%

The RAM requirements for each of the AutoXpert and AutoCons GUIs can be calculated using the formula:

20,000 kbytes + (5.5 kbytes x # of jobs loaded) = amount of RAM required.

To open all 4 GUIs when the entire days' processing schedule is loaded would require:

[20,000 kbytes + (5.5 kbytes x 9,336 jobs)] x 4 GUIs = 286 Mbytes of RAM minimum.

The 9,336 jobs is derived by multiplying the number of PGE activations at LaRC for a 24 hour day (778) by the number of jobs it takes AutoSys to activate a PGE (4) plus an additional 200% increase in job activations due to reprocessing (x3),

778 PGE activations x 4 jobs per PGE x 3 = 9,336 job activations per day.

AutoSys Database Sizing

The size of the AutoSys Database is dependent upon the number of jobs loaded into AutoSys at one time. The database, when loaded with a test of 20,000 jobs, consumed 6 MB of disk space. This number would vary as a function of the complexity of job dependencies.

The AutoSys events database will also require disk space. The events database saves information (starting and run times, success or failure of job, alarms, etc.) about the job activations launched by AutoSys. Estimates show that an additional 53 MB of disk space would be needed to archive 14 days worth of event information for LaRC.

AutoSys Job Throughput Analysis

Experience by other AutoSys customers has indicated that AutoSys can sustain 45 jobs per minute while running on a 75 MHz Sun SPARC 20/712 2-CPU machine. Operating in our high availability configuration reduces performance by approximately 35%. Therefore, AutoSys could execute a sustained 29.25 jobs per minute when operating in its Dual Data Server Mode (high availability option). Therefore, a 2-CPU Sun 20/712 machine could launch a maximum of 42,120 jobs for processing within a 24 hour period.

The predicted number of jobs for one 24 hour period at LaRC for Epoch k (3Q99) including an additional 200% increase in job activations due to reprocessing is 9,336 based on the February 1996 technical baseline.

Scheduling Hardware

The scheduling function of the Data Processing Subsystem is done by a pair of Sun 20/712 workstations. Each workstation is attached to two 6.3 GB Sun SPARCstorage MultiPacks (external disk) via fast wide SCSI controllers. AutoSys software runs its Event Processor (using Sybase) on the primary workstation while the AutoSys Shadow Event Processor runs on the secondary workstation. When the Shadow Event Processor detects that the Event Processor has malfunctioned, the Shadow Event Processor takes charge of the execution of the schedule, thereby allowing the continued launching of production jobs. AutoSys will also be run in Dual Data Server Mode which allows AutoSys to replicate its database on separate disks. When a disk or database failure occurs on the primary database, AutoSys continues to operate, using the backup database.

The Sun operating system and applications will reside on disk internal to each of the 20/712 machines. The Sybase log, index, and data files will be on external disk. Two sets of three 2.1 GB external disks will be accessible via 2 fast wide SCSI controller cards for each machine.

Each 20/712 machine can be upgraded to use emerging hyperSPARC technology CPUs running at 150 MHz. Future experience with AutoSys and its Sybase database will determine if this extra capacity will be needed.

Each SPARCstorage MultiPack can be increased to a maximum of 12 disks for a total of 25.2 GB if required.

3.4.2.6.2 Configuration

- Production Planning/Management and PDPS DBMS Server: Quantity: 2

  - Sun SPARC 20/712

  - Clock: 75 MHz

  - CPUs: 2

  - Disk: 4.2 GB

  - RAM: 512 MB

  - Fast wide SCSI controller card: 2

  - SPECrate_int92 = 5,726, SPECrate_fp92 = 5,439

- Planning Disk: Sun SPARCstorage Multipack Quantity: 4

  - 6.3 GB (3, 2.1 GB disks)

  - Hot swappable drives with multiple active spindles

  - Maximum of 12 drives for 25.2 GB

- Queuing Server: Quantity: 2

  - Sun SPARC 20/712

- Clock: 75 MHz

- CPUs: 2

- Disk: 4.2 GB

- RAM: 512 MB

- Fast wide SCSI controller card: 2

- SPECrate_int92 = 5,726, SPECrate_fp92 = 5,439

• Queuing disk storage: Sun SPARCstorage Multipack Quantity: 4

- 6.3 GB (3, 2.1 GB disks)

- Hot swappable drives with multiple active spindles

- Maximum of 12 drives for 25.2 GB

• Production Monitor: NCD HMX Pro Quantity: 4

- X Terminal, color

- RAM: 16 MB

- Size: 21"

3.4.2.7 Data Processing Subsystem

The Data Processing Subsystem (DPS) includes three hardware CIs:

1) *Science Processing (SPRHW)* — The Science Processing HWCI (SPRHW) is the primary HWCI in the Data Processing Subsystem and contains processing resources (processors, memory, disk storage, and input/out subsystems) necessary to perform first-time processing, reprocessing, and Algorithm Integration & Test (AI&T). SPRHW also provides the hardware resources necessary to support management of the science processing, in the form of a Queuing Server and Production Planner Stations. However, the queue management function in science processing is closely coupled to the ECS planning function, and the Queuing Server and Production Planner Stations are therefore discussed in the Planning Subsystem Design Specification (305-CD-026-002). This chapter describes only the hardware resources required to execute the science software.

2) *Quality Assessment and Monitoring (AQAHW)* — The Algorithm Quality Assurance HWCI (AQAHW) provides hardware resources to support DAAC operations personnel in performing planned, routine, non-science QA of product data.

3) *Algorithm Integration and Test (AITHW)* — The Algorithm Integration & Test (AI&T) HWCI (AITHW) provides hardware resources to support the integration and test of science software at the DAAC, and system level validation, integration and test. It is important to note that this HWCI provides workstations and tools for software integration

and test, but does not provide the compute environment or compute capacity required for science software test. This integration and test compute capacity is included in SPRHW.

The specification of these HWCIs at LaRC is described below.

3.4.2.7.1 SPRHW

The design specification for SPRHW is derived by first reviewing the requirements for SPRHW in Section 3.4.2.7.1.1. The requirements are traced from the first principles established in the system requirements, through the detailed processing requirements levied by the instrument teams, and through the models built to simulate system behavior under the load defined by the instrument teams. The technology assessments performed to support SPRHW specification are reviewed in Section 3.4.2.7.1.2. The conclusions of the technology assessments are then applied to the derived, detailed system requirements to produce the SPRHW specification, provided in Section 3.4.2.7.1.3. Details of the specification are discussed in Section 3.4.2.7.1.4 to show how the system requirements are met.

3.4.2.7.1.1 SPRHW Requirements Analysis

The specification of SPRHW must satisfy sizing, expandability, reliability, maintainability, availability, compatibility, and interoperability requirements. These are discussed in the sections below.

SPRHW Sizing and Expandability. The SPRHW sizing and expandability requirements are derived from several sets of fundamental system requirements established in the ECS Functional and Performance Requirements Specification (F&PRS). Specific numerical detail is provided for these requirements by the processing plans provided by the instrument team inputs to the Ad Hoc Working Group on Production (AHWGP). Because the processing plans are quite complex, with many inter-relationships between products, significant analysis is required to understand the processing plans in order to reach a processing hardware design. This analysis has been performed using static and dynamic modeling. The modeling and design efforts have identified some design parameters that are not addressed by the AHWGP inputs; to address these design parameters, additional information concerning the processing plans has been solicited from the instrument teams in the form of a survey. The information derived at each of these steps has been used to determine the design parameters (requirements) for the science processing hardware. The following sections describe that process.

Functional And Performance Requirements. The F&PRS provides several requirements specifying the throughput and timeliness of ECS science processing.

Requirements EOSD1050, EOSD1060, EOSD1070, and EOSD1080 state the timeliness requirements for producing products for first-time processing of data. Level 1, Level 2, and Level 3 products must be produced within 24 hours of the time that ECS receives all of the data required to produce the product. Level 4 products must be produced within seven days of the time that ECS receives all of the data required to produce the product.

Requirement EOSD1040 states that ECS shall have the capacity to perform reprocessing at twice the rate of first-time processing. Requirement PGS-1300 states that ECS shall have the capacity to

perform AI&T, production of prototype products, ad hoc processing for "dynamic browse" or new search and access techniques, and additional loads due to spacecraft overlap at one times the rate of first-time processing.

Requirement PGS-1301 states the requirement that vendor peak processing estimates for processors be derated by a factor of four for design purposes.

Requirement PGS-1270 states the requirement that the processing design and implementation have expandability by a factor of up to three without design change, and by a factor of up to ten without major design change.

The Technical Baseline And The AHWGP Inputs. Appendix C of the F&PRS is included by reference in the F&PRS requirements. This appendix contains estimates of the volume of processing to be performed by ECS. Because these estimates have changed over time and become significantly more detailed as instrument team plans have matured, the material in Appendix C of the F&PRS has been replaced by a group of separately maintained documents referred to as the ECS Technical Baseline.

The Technical Baseline contains material that defines the level of processing required by each instrument for first-time processing. Much of this data has been provided by the instrument teams in the form of inputs to the AHWGP. These inputs are updated approximately every six months; the most recent updates to the AHWGP inputs and the ECS Technical Baseline were baselined in February 1996. The February 1996 Technical Baseline is therefore the basis for the CDR design for SPRHW.

AHWGP processing requirements are expressed in terms of Product Generation Executives (PGEs), which are the smallest scheduled increment of science processing. For each PGE, the AHWGP input identifies the DAAC at which the PGE will be executed, and provides the execution frequency, the list of input data requirements, the list of output data products, and an estimate of the number of millions of floating point instructions (MFPI) performed with each execution. For each input file, an estimate is given of the fraction of the file that is read by the PGE. The baseline also identifies the calendar quarters during which the PGE will be executed; this allows the instrument teams to identify how they will phase in their processing after the instruments are put on orbit. As a result, the processing demand for each instrument may vary by calendar quarter.

The Technical Baseline documents program plans and directives that affect the SPRHW specification. The hours of operation for each DAAC are provided. For DAACs having less than 24 hour per day, seven day per week operations, the assumption is made that all processing must be accomplished during DAAC operating hours.

The Technical Baseline also documents the program directive to phase the implementation of processing capacity for each instrument relative to the date of the instrument's launch. The purpose of the phasing factors is to provide sufficient early processing resources to support AI&T, without purchasing the required full-up capacity before it is needed. With **X** defined as the processing resources required to do first-time processing for an instrument and **L** defined as the launch date of the instrument, the phasing factors are defined below:

- **0.3X for L-2 < t < L-1.** Pre-launch AI&T requires 0.3X during the period from one to two years before launch.
- **1.2X for L-1 < t < L+1.** Pre-launch AI&T and system I&T requires 1.2X during the year before launch. Standard instrument processing requirements (X) begin from the launch date and last for the remainder of the life of the instrument.
- **2.2X for L+1 < t < L+2.** Post-launch AI&T, standard processing, and reprocessing of data require 2.2X starting at launch plus one year.
- **4.2X for t > L+2.** Post-launch AI&T, standard processing, and reprocessing of data require 4.2X starting at launch plus two years.

The launch dates for instruments supported by ECS are documented in the Technical Baseline. Because the PGE processing plans are expressed on a quarterly basis, for the purpose of applying the phasing factors the launch dates are brought forward to the first day of the appropriate calendar quarter. The large majority of processing requirements supported by Release B are for instruments on two platforms, TRMM and AM-1. TRMM is scheduled for launch during the third quarter of (calendar year) 1997 (3Q97), and AM-1 is scheduled for launch during the third quarter of (calendar year) 1998 (3Q98). Therefore the phasing factors for SPRHW are assessed at 3Q96, 3Q97, 3Q98, 3Q99, and 3Q00. However, Release A of ECS has already put resources in place to cover the 3Q96 requirements; hence the focus of Release B design is for the latter four dates.

Specifications for SPRHW have been designed for each of the four Release B dates identified above. However, for the purpose of discussing the design, the 3Q99 period (also referred to within ECS as Epoch k) has been selected as a common frame of reference across the entire set of Release B Design Specifications. At Epoch k, TRMM instruments are supported at the 4.2X level of processing, and AM-1 instruments are supported at the 2.2X level of processing. For the sake of simplicity, and because its processing requirements are relatively small, the SAGE III instrument is phased as if it was launched simultaneously with AM-1.

Static Modeling. The AHWGP inputs define over two hundred PGEs, executed at five DAACS for eleven instrument teams. The PGEs create over three hundred products, and require over one hundred ancillary and Level 0 input data sets. The interdependencies of the products and PGEs creates a network so complex that it cannot be understood by hand. The inputs are therefore analyzed using models, to reduce the volume of data into aggregates for each instrument and DAAC, and to understand the dynamic behavior of the interdependencies. The first step in this analysis is to use a static model to analyze the PGEs.

The static model represents each PGE within ECS on a single line of a spreadsheet. The columns of the spreadsheet contain the attributes for the PGE: its process identifier, instrument, processing site, frequency of initiation, number of executions per initiation, total input file size, total output file size, number of input files, number of output files, volume of files staged per execution, volume of files destaged per execution, and number of MFPI per execution.

Most PGEs are executed once per set of inputs, and those inputs are produced (by the instrument and/or by other PGEs) periodically. However, some PGEs may be executed several times on a given set of inputs that are periodically produced. An example is a MODIS PGE that produces

products defined in terms of geographic tiles; for each set of input data covering the whole earth, the PGE may be executed 355 times, once for each MODIS tile. Hence it is necessary to track both the frequency of initiation of the PGE (the periodicity of the inputs) and the number of times the PGE is executed for a given number of inputs.

The numbers and sizes of input and output files are evaluated by summing the inputs provided for the PGE by the instrument team. The staging and destaging inputs to the static model are derived from the AHWGP inputs. Staging refers to the process of moving data to the science processing resources, and destaging refers to moving data away from the science processing resources. SPRHW does not archive data; it retains only a very limited amount of ancillary data, identified by the instrument teams as permanent, for extended periods (more than 24 hours). All other data required by or produced by processing must be staged from or destaged to the ECS archive. The staging and destaging entries in the static model spreadsheet are calculated by subtracting the permanent ancillary file volumes from the total file volumes.

These inputs are then used to calculate average resource usage levels for processing, network input/output (I/O), archive I/O, and disk I/O. The processing usage is calculated as the number of millions of floating point operations per second (MF) required to meet the PGE's requirements based on the PGE's frequency of initiation, number of executions per initiation, and MFPIs per execution. Similarly, the data flow to the archive (due to destaging), from the archive (due to staging), over the network (due to staging and destaging), and over the disk I/O channels (due to processing, staging, and destaging) are calculated. These values are then aggregated (summed) for each instrument and DAAC.

One method of static model analysis is to calculate the average level of resource usage required to support each PGE, instrument, and DAAC, assuming that the processing is spread smoothly over time. This analysis ignores the timeliness requirements for products, but it produces a baseline estimate of the minimum resources required to keep up with the processing requirements. It is this analysis that defines the value of X used for phasing requirements.

Another approach for the static analysis provides an estimate of the processing capacity required on the worst case day. This analysis takes into account the timeliness requirements for the PGEs (24 hours for Level 1, 2, amd 3 products, and one week for Level 4 products); further, it assumes that each periodic product (e.g., weekly or monthly) is executed at least once on the worst case day. Thus the load for the worst case day assumes the completion of at least one instance of every PGE associated with the DAAC. This produces a (generally) worst case estimate of the resources required to meet the timeliness requirements for first-time processing. (Because the static model cannot take into account the dynamic interactions between PGEs, it is theoretically possible that the worst case for the system is actually exceeded by this analysis, but the dynamic model has not shown this to be the case.)

Table 3.4.2.7.1-1, *LaRC Static Model Summary Results*, shows summary static model results for processing at LaRC for Epoch k (3Q99). The values in the table represent the static model estimates for resource requirements by instrument for first-time processing only. These results represent a time-averaged analysis, as opposed to a worst day analysis; however, LaRC has only a handful of PGEs with an execution periodicity of greater than one day. These PGEs (principally the CERES monthlies) do not significantly drive the requirements for SPRHW resources; their

impact is chiefly upon the LaRC archive. Thus the SPRHW load at LaRC has very little day to day variation, and the average analysis presented below provides a good estimate of the minimum resources required at LaRC for first-time processing. (Staging I/O requirements are better estimated by the dynamic model, as discussed below.)

Table 3.4.2.7.1-1. LaRC Static Model Summary Results

Instrument	Processing (MF)	Processing I/O (MB/s)	Staging I/O (MB/s)	Destaging I/O (MB/s)
CERES TRMM	895	1.5	0.5	1.0
CERES AM-1	1,827	3.8	0.5	3.3
MISR	3,299	18.6	5.3	2.7
MOPITT	9	0.1	0.1	<0.05
SAGE-III	3	<0.05	<0.05	<0.05
DAAC Total	6,033	24.0	6.4	7.0

Dynamic Modeling. The static model has significant limitations. It cannot take into account the dynamic interdependencies of the PGEs or the system's computing resources. As a result, it cannot accurately predict the end to end clock time required to produce a stream of products once Level 0 data has been received. It also provides extremely inflated estimates of the system staging requirements, because it must assume that every (non-permanent) file required for each execution of each PGE is staged for that execution; it cannot take into account that the file may already have been staged to SPRHW in support of some other PGE. On the other hand, it assumes uniform utilization of the SPRHW resources, not accounting for the fact that one type of resource (such as a processor) may be idle while the system uses another type of resource (a network).

To overcome these limitations, a dynamic model of a subset of ECS has been implemented. The dynamic model is implemented using the Block Oriented Network Simulation (BONeS) tool. BONes is a discrete-event simulation tool for analysis and design of communication networks and distributed processing systems. Components of a distributed processing system are represented by nodes, which have resources associated with them that get allocated as events request them. First time production of products by DPS is simulated by the Processing module, in conjunction with the Data Handler module, the Event Driven Scheduler, and the Ingest module. The Data Handler simulates the behavior of the ECS Data Server. The Data Handler stores and retrieves data from the permanent archive, routes data to the requesting subsystems, and manages tiered storage and staging resources. The Event Driven Scheduler monitors the availability of data, requests data to be staged from the Data Handler to Processing, routes newly created data to the appropriate Data Handler or Processing resource, and initiates execution of a process when all required inputs are present. It should be noted that the Event Driven Scheduler is not intended to serve as a simulator of the Planning Subsystem; rather, it simulates the effects of data arrival at ECS and the actions of the Planning Subsystem. The Ingest module simulates the Ingest Subsystem, which accepts data from external systems and users and contains rolling storage of Level 0 instrument data.

The dynamic model is run by creating an input event stream representing the arrival times of Level 0 and ancillary data files. The PGEs are defined to the model in terms of their resource requirements and file dependencies. The system design is also represented as an input to the model, in terms of the number and capacity of system resources for example, the number of SPRHW processors and their throughput in MF. The Event Driven Scheduler then executes the event stream, which is driven by the external events (receipt of Level 0 and ancillary files) and by derived events (the availability of the data and computing resources required to execute a PGE). The simulation is run over an extended simulation period typically 21 days to examine the effects of products produced periodically. The simulation is started with an empty ECS system; there is therefore a starting transient in the simulation's output results. Because of this, the results of the first five days of the simulation are generally discarded.

The results of the dynamic model are analyzed to identify peak and average resource requirements, and average and maximum system response characteristics. Graphs showing resource utilization over time during the period of the simulation can also be generated.

The model explicitly represents each processor within SPRHW, and constrains the simultaneous execution of PGEs to the number of processors in the design, operating at the vendor's peak MF rating derated by a factor of four per the requirements. Average and peak processor utilization and average and peak processor queue lengths are reported, as well as processor utilization curves over the period of the simulation. These results identify whether processing backlogs are occurring, and whether there is excess processing capacity in the design. In this respect the dynamic modeling is an iterative process; an initial, generous configuration of resources is fed to the model, and the results are examined. Design choices are made, and a refined configuration of resources is fed to the model. This process continues until all of the design decisions have been made, and the model indicates that the configuration will satisfy the system's requirements.

For disk storage at the science processors, the model does not constrain the available space; instead, it implements an algorithm to depict the purging of unneeded data from the SPRHW storage array, and monitors the demand for storage. By plotting the storage demand over the period of the simulation, it can be determined when the peak occurs, and for what duration. This information is used to size the SPRHW storage arrays. The dynamic model indicates the following storage requirements for first-time processing:

A. CERES TRMM: 60 GB;

B. CERES AM-1: 432 GB;

C. MISR: 182 GB;

D. MOPITT: 8 GB; and

E. SAGE-III 1 GB.

Dynamic model runs against the February 1996 Technical Baseline that included reprocessing were not available at the time of this writing; however, such runs were performed against the August 1995 Technical Baseline for the CERES and MISR instruments. (Extrapolation from dynamic modeling results using the August 1995 Technical Baseline is possible for these instruments because the AHWGP inputs for CERES were essentially unchanged in the new

baseline, and were only slightly changed -- a few percent in processing, for example -- for MISR.) These results indicate that CERES disk requirements do not increase significantly when a one X reprocessing load is added to a production CERES configuration, but that MISR disk requirements almost double when a one X reprocessing load is added to its processing configuration. These results can be traced back to the characteristics of the PGEs that dominate processing for the instruments.

Network and disk I/O are averaged over time to provide average utilization rates. The model represents these as constrained resources; hence, the model results provide assurance that a given level of network and disk I/O throughput will satisfy the system's requirements, and provide an approximation of the fraction of these resources that is used on average. Based upon dynamic modeling results, the staging requirements per instrument are estimated as follows:

A. CERES TRMM: 0.2 MB/s;

B. CERES AM-1: 0.2 MB/s;

C. MISR: 2.7 MB/s;

D. MOPITT: <0.05 MB/s; and

E. SAGE-III <0.05 MB/s.

The processing I/O and destaging rate estimates shown in Table 3.4.2.7.1-1, based on static modeling, are unchanged by the dynamic model results.

Memory Usage And I/O Survey. The inputs to the AHWGP do not provide any information as to how the PGEs will use processing memory, or about the detailed characteristics of their disk I/O.

Although all of the potential platforms considered for SPRHW make use of virtual memory, the ratio of physical memory (RAM) to virtual memory required during the operation of a system is important for performance reasons. The rate at which memory is accessed by a PGE and the pattern of these accesses can also have important performance impacts. For the SPRHW specification, these requirements affect the amount of RAM, number of memory channels, amount of disk (for virtual memory), and number of disk channels (for virtual memory access). Because the SPRHW processing requirements are very demanding and generally require sophisticated and expensive processors, it is desirable to get as much throughput as possible from the processors by keeping as much of a PGE in RAM as practical.

Similarly, the size and pattern of disk I/O requests strongly affects the performance that can be expected from a disk I/O subsystem. Because of various overheads associated with each disk access, throughput from a disk subsystem increases strongly with the average size of the disk request, up to maximum limits imposed by the disk, interface, and controller technologies. While vendors usually quote I/O subsystem performance for large request sizes (hence peak performance values), it is important to know whether these request sizes are typical for ECS applications; and if not, to derate the vendor performance specifications accordingly.

The type of memory usage and file I/O characteristics that are needed to determine the system requirements for these resources can generally be determined either by code inspection or by measurement during execution. ECS solicited this information from the Release B instrument

teams in January of 1996. Unfortunately, it is still too early for most teams to provide accurate assessments of their requirements, and little data has been received to date in response to the survey. The integration and test of beta science software for Ir1 during the first half of 1996 may provide requirements in time to support the procurement of the first increment of Release B hardware.

For purposes of the generation of the Release B CDR design, assumptions have been made in the absence of requirements data in these areas. It is assumed that each PGE will require approximately 128 megabytes (MB) of RAM. This figure is viewed as a reasonable upper limit for most algorithms, especially if the algorithms are designed with consideration of their memory requirements. It is recognized that some algorithms may require large data structures to iteratively refine their data using models, and will have correspondingly larger memory requirements. These uneven needs between algorithms are mitigated in part by the Symmetric Multiprocessor (SMP) architecture that has been selected for SPRHW.

It is also assumed that ECS SPRHW file I/O will be buffered by the platform operating system, with the average request size characterized by the operating system I/O buffer size. This assumption is made because it is believed that this correctly characterizes the use of HDF, which is expected to be the dominant form of disk access on the science processors. The buffer size varies by operating system, but is typically small (64 kilobytes or less) relative to request sizes used for direct file I/O (256 KB to 4 MB). This means that the vendor disk I/O performance ratings will be derated to reflect small request sizes.

SPRHW Reliability, Maintainability, and Availability Requirements. The product generation function of DPS has an availability requirement of 0.96 (96%) and a mean down time requirement of not to exceed four hours. Implicit in these requirements is also the requirement that SPRHW be able to achieve its timeliness and throughput requirements while allowing for 4% non-availability and component outages of up to four hours.

SPRHW Compatibility and Interoperability Requirements. The Release B ECS software must satisfy ECS requirements for the use of open systems technology, and must be compatible with other Release B subsystems and with hardware provided to the DAACs for Release A. Specifically, the SPRHW hardware suite must be supported by OODCE. SPRHW must also support the network interfaces to be implemented by ECS. For LaRC, these interfaces must support data exchange between DPS and DSS at aggregate rates (in Epoch o) of on the order of 30 MB per second.

3.4.2.7.1.2 Technology Assessments

The requirements identified in Section 3.4.2.7.1.1 can be compared to the capabilities of current and future technologies to identify candidate technical solutions. This type of evaluation can be done on paper, in the form of trade studies or technology evaluations, or it can be done in the laboratory, through prototyping and benchmarking. The paragraphs below discuss the technology assessment activities that have been performed to support the Release B design process, with emphasis on the activities performed since IDR-B.

Technology Evaluations. A number of technology evaluations have been performed under ECS to identify the best technical approaches and products for the SPRHW suite. Three of these evaluations are discussed below.

Production Platform Families. The requirements for first-time processing for instruments at the DAACs range from a few MF for the less compute-intensive instruments, up to 4.8 GF for MODIS at GSFC. Computing requirements in the GF range have historically been the turf for vector supercomputers, but advances in massively parallel processors, symmetric multiprocessors, and workstation farms have pushed these technologies into this range of performance. An evaluation was performed for Release A and Release B of the following candidate platform family architectures for SPRHW:

1) Single processor workstations;
2) Farms of workstations connected via local area networks;
3) Symmetric Multiprocessors (SMPs);
4) Massively Parallel Processors (MPPS); and
5) Vector Supercomputers.

The key criteria used to evaluate the platform families were lifecycle cost per unit of processing power, expandability, ease of science software development, and flexibility. The evaluation found that the requirements for the less compute-intensive instruments could be met by single processor workstations, which are the least expensive and most flexible resources. For the compute-intensive instruments, however, only the SMPs, MPPs, and vector supercomputers offered the expandability required. Among these three, the vector supercomputers were found to have significant cost disadvantages. The SMPs were found to have significant advantages over the MPPs in the ease with which applications can be developed, and the evaluation concluded that the SMP platform class was best suited for ECS's high-end SPRHW requirements.

This evaluation is documented in *Platform Families for the ECS Project*, 440-TP-007-001.

Distributed And Parallel Processing. Parallel processing in both shared memory and distributed memory architectures is viewed as an emerging technique for solving large processing problems. An analysis has been performed to examine the benefits of distributed and parallel computing for ECS. The analysis studies various processing alternatives for ECS science algorithms and provides information on processing technologies. The analysis examines the applicability of using the OSF/Distributed Computing Environment (DCE), symmetric multiprocessing with shared memory, distributed multiprocessing (including workstation clusters), and massively parallel processing for ECS science software.

The SMP architecture is found to provide significant flexibility: it fully accommodates non-parallelized code, supports easy migration of non-parallel code to shared memory parallel processing, and can also support distributed memory parallel processing. This flexibility is viewed as giving SMP systems a significant advantage for ECS processing. Although a need to parallelize particular PGEs has not been identified to date, the ability in the future to parallelize ECS science

software is an important consideration, especially with the large processing requirements of Release B and beyond.

This evaluation is documented in *Distributed and Parallel Processing For ECS Science Algorithms: A Trade Analysis*, 440-TP-008-01. ECS Science and Technology Lab prototyping efforts that have demonstrated parallel program development, based primarily on parallelization tools, are documented in 194-00569TPW and 194-430-TPW-001.

Production Topologies. A technical analysis has been performed examining the advantages and disadvantages of distributing processing tasks from one or more instruments across one or more processing clusters. The purpose of the study was to examine how the interactions between the processing requirements of different instruments might increase or decrease the total hardware resources required to perform the aggregate load. The results of this analysis are reflected in the SPRHW design. The study concluded that the compute-intensive instruments at LaRC (CERES TRMM, CERES AM-1, and MISR) were large enough to require whole platforms to meet their requirements. It will be beneficial to cross over the execution of certain CERES TRMM monthly processes to the CERES AM-1 processing suite, because this allows the two sets of monthly processes to share some data sets, and reduces the load on the LaRC archive. No similar synergy between CERES and MISR was found. The less compute-intensive instruments at LaRC (MOPITT and SAGE III) have requirements so small that they can easily be accommodated within the spare capacity provided by the CERES and MISR platforms; the ECS queuing and planning capabilities will be used to schedule them.

This analysis is documented in *Production Topologies: A Trade-Off Study Analysis for the ECS Project*, 440-TP-006-001.

Prototypes and Benchmarks. Prototypes and benchmarks are used to assess the applicability of a technology or product to a set of requirements. Two sets of recently acquired benchmarks have been used in refining the ECS SPRHW design for Release B. The first set of benchmarks, measuring the performance of IP over HiPPI, was performed in the ECS development facility. The second set of benchmarks, measuring I/O subsystem performance on the SGI platforms, was performed at the University of Minnesota and by SGI.

Network Benchmarking. Benchmarking tests of the performance of IP over HiPPi have been performed in the ECS development facility. The purpose of these tests was to measure the throughput that can be achieved using IP over HiPPI, to determine whether IP introduces a significant overhead in HiPPI communications.

The tests used two SGI Challenge XL systems, equipped with SGI HIO HiPPI adapters. These systems were running version 5.3 of the Irix operating system. Two different benchmarking tools were used to send TCP streams from the memory of one system to the memory of the second system. Throughput rates were measured as the configuration parameters for TCP/IP were varied.

The maximum transfer rates observed in the test were approximately 55 MB per second, just over half of the 100 MB per second theoretically possible with HiPPI. SGI has stated that using version 6.2 of Irix, rates of up to 90 MB per second have been observed in their laboratories. This indicates that IP does not burden HiPPI with a substantial overhead, and that IP-based protocols can be used

over HiPPI in ECS. This eliminates the need to develop custom protocols for ECS using raw HiPPI interfaces.

I/O Subsystem Benchmarking. The Laboratory for Computer Science and Engineering (LCSE) at the University of Minnesota, in affiliation with the Army High Performance Computing Research Center (AHPCRC), has performed a number of benchmarking tests on I/O subsystem performance on SGI platforms. These tests have included performing high speed data transfers over HiPPI using a lightweight TCP/IP protocol, and performing high data rate disk transfers. These tests have demonstrated that the SGI architecture can sustain I/O rates in excess of 150 MB per second between machines, and over 500 MB per second to locally attached file systems.

The HiPPI experiments performed at LCSE demonstrated the ability of the SGI platforms to support data visualization. In these experiments, a single Challenge server, configured with four HiPPI connections, was used to transfer data from its file system to the graphics processors of two SGI Onyx systems. A lightweight TCP/IP protocol referred to as "NFS-Bypass" was used to perform the transfer. This protocol, which provides an NFS-like interface, transfers data between systems using a socket to socket connection, bypassing most of the overhead normally imposed by NFS. Transfer rates for a single HiPPI connection, from disk to graphics processor, of 65 MB per second have been observed. Using multiple HiPPI connections transfer rates from a single machine of over 150 MB per second have been observed.

Research at LCSE has also been performed to evaluate the maximum transfer rates available from various storage configurations on SGI platforms. These experiments have focused on the use of striped, SCSI-2 based RAID file systems using SGI's XFS file system and direct I/O. The LCSE researchers have built file systems capable of sustaining over 500 MB per second transfer rates on a Challenge system. A key aspect of their findings in the building of fast file systems is the need to use large block request sizes. They found that block request sizes associated with buffered I/O — I/O buffered by the operating system's I/O cache — are 64 KB or less. Disk subsystem performance at these request sizes is substantially below peak levels; at this request size, throughput of less than 10 MB per second is observed on SCSI-2 channels having a theoretical bandwidth of about 20 MB per second. Using direct I/O, which is not buffered by the operating system's cache, and with request sizes of one MB or more, throughput rates rapidly approach the per channel limitation imposed by SCSI-2.

SGI has further tuned the NFS Bypass software to make efficient use of the Challenge memory subsystem, and is productizing the software. The product, Bulk Data Service (BDS), will be offered as an extension to Irix 6.2 in the second quarter of 1996. Significantly, since BDS is built on top of TCP/IP, it is not specific to HiPPI; BDS may be used to implement highspeed file transfers over any TCP/IP connection.

SGI has also recently performed benchmarking tests on their newest generation of RAID controllers. These tests indicated sustained transfer rates of eight MB per second per controller are achieved when small request sizes (64 KB) are used. This benchmark is used as a basis for sizing the disk I/O subsystems for SPRHW in Release B.

3.4.2.7.1.3 SPRHW Specification

SPRHW has three top-level component types: Science Processors, Queuing Servers, and Production Planner Stations. As discussed earlier, the Queuing Server and Production Planner Station components are specified in the Planning Subsystem Design Specification.

Appendix A provides detailed configuration information for the LaRC SPRHW suite at Release A and at Release B, including Epochs c, g, k, and o. The intended function and the processing, memory, I/O subsystem, disk channel, storage device, network, enclosure, and console configuration are specified for each system in the SPRHW suite. Upgrades to systems can be identified by comparing the configurations for a system at different epochs.

The Science Processors are characterized by their vendor, enclosure, processors, memory, I/O subsystems, internal disk drives, external disk systems, backup and update devices, network interfaces, and monitors. The sections below provide further explanation for the specification provided in Appendix A. The system designators defined in Appendix A (e.g., SPRHW-LaRC-1, or just -1) are used in the text below.

SPRHW Vendor. The Release B SPRHW components, with the exception of tape libraries provided by Exabyte, will be provided by Silicon Graphics (SGI). The architecture trade-offs show that SMP systems provide clear advantages for ECS science processing, and SGI is a leading vendor of SMP systems. When considerations are taken into account for cost, availability of required ECS software capabilities (such as OODCE), availability of highspeed network components (particularly HiPPI), and ease of re-use of Release A hardware, the logical choice for Release B hardware is SGI.

Final vendor selection has not been made for the external disk systems required by SPRHW. A baseline design is presented here which features RAID-5 devices using SCSI-2 interfaces. This technology is mature and its performance is fairly well understood. However new technologies, particularly fibre channel based RAID arrays, are making their way into the market. If these products are found to be stable and cost effective before the first Release B procurement, they may be selected as the Release B baseline.

Processors. For new equipment purchased for Release B, SPRHW will use the MIPS R10000 processor. This chip has a superscalar 64 bit architecture, capable of performing two floating point operations per clock cycle. SGI has announced that the chip will be offered in a 275 MHz implementation in the second half of 1996. Using the derating factor of 0.25 defined in F&PRS requirement PGS-1301 for vendor specified peak throughput ratings, this processor is estimated to provide 137.5 MF (derated).

LaRC will retain two systems, purchased for Ir1 and Release A and designated SPRHW-LaRC-5 and SPRHW-LaRC-6, that have 90 MHz R8000 processors. These processors are estimated to provide 90 MF (derated) each.

A third system purchased for Ir1 and Release A, SPRHW-LaRC-1, will be upgraded from 75 MHz R8000 processors to 275 MHz R10000 processors.

Decisions about how many processors to configure in a single system were driven by the processing requirements for the instruments and for AI&T. At LaRC, the disk and network I/O requirements are small relative to the processing (MF) requirements; it was not necessary to split

processing loads across systems because of the I/O loads. As a result, the systems at LaRC are configured to provide increments of processing power useful in satisfying the needs of the instruments.

The decision to retain two systems with R8000 processors is considered a cost-effective means of providing for the requirements of CERES TRMM. These machines (-5 and -6) are already in place, and CERES TRMM will begin production operations on them. Together they provide the Epoch k requirement for 3X processing for CERES TRMM.

First time processing requirements for CERES AM-1 can be met using 20 R10000 processors. These can be housed in a single enclosure (-8) with significant room for expansion (see the discussion below in Section 3.4.2.7.1.4). When a one X reprocessing load is added at Epoch k, a second system (-11) is added, and the required processing resources are balanced across the two machines. At Epoch o, when an additional one X reprocessing load is added, processors are added to the -8 and -11 systems.

For MISR, 32 processors are required for first-time processing. While it would be feasible to configure these in a single enclosure, this would allow very little room for growth; hence, two machines (-9 and -12) are specified. At Epoch k, processors are added to these systems to handle the first X of reprocessing. At Epoch o a third system (-12) is added; it is identical to the Epoch k -9 and -12 systems.

The requirement for AI&T processing power for TRMM instruments is 0.2X at Epochs c and g, and 1.2X at Epochs k and o. The requirement for AI&T processing power for AM-1 instruments is 0.2X at Epochs c, g, and k, and 1.2X at Epoch o. The value of X for each instrument is defined above in Table 3.4.2.7.1-1, *LaRC Static Model Summary Results*, as determined by the static model. System SPRHW-LaRC-1 is configured to meet the AI&T requirements at Epochs c, g, and k; at Epoch o, a second system (SPRHW-LaRC-13) is added. Note that for Epochs c and g, SPRHW-LaRC-1 is configured with significantly more processing power than strictly required for AI&T. This is so that this system can be used as a failover system, should any of the production machines go out of service.

Memory. It was noted above that firm data for memory requirements are not available yet, and thus it would be assumed that each processor requires approximately 128 MB of RAM.

For SMP configurations, the SGI memory architecture features leaf controllers, which process requests to a subset of the system memory. Each memory board can have one or two leaf controllers. A system can have up to eight memory boards (subject to backplane limitations), but only up to eight memory controllers. The I/O benchmark results by the group at the University of Minnesota have shown that memory interleaving is important for I/O performance. It is also expected to be important for science algorithm performance on systems with many processors. Therefore, machines with more than one GB of RAM and machines performing high rates of I/O will be configured with eight memory controllers; lesser machines will generally be configured with four memory controllers.

SGI offers RAM at two chip densities. Combined with options for up to eight memory boards, there are a large number of memory configurations offered. Unfortunately, however, there is no configuration offered with eight way interleaving and three GB. As a result, systems at LaRC with between 12 and 28 processors are configured with two GB of RAM and eight way interleaving.

I/O Subsystems. For SMP systems, the SGI architecture provides configurable I/O subsystems which attach to the backplane; each subsystem occupies one backplane slot and provides up to 320 MB per second of bandwidth. The subsystem card is referred to as a PowerChannel 2 or IO4 card. A system may have up to six IO4 cards, subject to backplane limitations.

The first IO4 card in an SMP system provides console and ethernet connections. Each IO4 provides serial and parallel connections, two fast wide differential (FWD) SCSI-2 channels, and space for two HIO controller cards. HIO controller card offerings include a HiPPI card, a FDDI card, and a card supporting three SCSI-2 channels.

The number of IO4 cards specified for each SMP is determined by allocating HIO slots to the FDDI and HiPPI interfaces, and counting the number of SCSI-2 interfaces required. The number of SCSI-2 interfaces required is determined by the number of internal and external SCSI-2 devices supported by the system. In general, it is assumed that the internal slow SCSI-2 devices (CD-ROMs, floppy disk drives, and tape drives) will be aggregated on the first SCSI-2 channel. Internal disk drives will be allocated to the second SCSI-2 channel. External disk arrays will be allocated to subsequent SCSI-2 channels; the number of channels required is based on the desired throughput of the external file systems. At LaRC, the disk I/O requirements are satisfied by using between two and four SCSI-2 channels for the external storage arrays; therefore each LaRC machine is configured with one SCSI-2 HIO card.

The Release A LaRC systems (-1, -5, and -6) have VME FDDI cards which attach directly to the backplane via special VME slots. These controllers will be retained in Release B, and reduce the number of HIO slots needed on these machines.

Appendix A shows how each IO4 HIO slot and each SCSI-2 channel is allocated on the LaRC systems.

Internal Disk Drives. The bulk of storage for the science processors will be provided by external storage arrays. Internal disk drives will only be used to provide swap space for the operating system, and to provide file system space for the operating system and applications. Applications in this context does not include the PGE executables or any temporary file space required by the PGEs; rather it refers to file system space requirements for the Autosys client software and ECS custom code. The file system requirement for the operating system and applications is not expected to exceed two GB.

The allocation for swap space is estimated at four times the size of the physical memory (RAM). This exceeds what is generally configured for servers, and probably represents an upper bound; if it is found that systems require virtual memory larger than four times their physical memory, it is likely that physical memory will also have to be upgraded to reduce paging.

SGI currently offers 2 GB and 4.3 GB internal disk drives and will soon offer 9 GB internal disk drives. A combination of 4.3 GB and 9 GB drives has been selected for each system to satisfy the estimated space requirement. These drives will be allocated to a single SCSI-2 channel.

External Disk Arrays. The disk size and throughput requirements for SPRHW are determined on a DAAC by DAAC and host by host basis, using the static and dynamic modeling results. These requirements are translated into numbers of drives and controllers based upon the

configurations available from the vendor. The current product of choice for external disk arrays is the RAID-5 product from SGI.

The SGI RAID-5 arrays use one redundancy disk for each four data disks. Arrays are built in groups of five disks, with up to four groups (20 disks) in an enclosure. Up to four enclosures can be put in a rack.

An enclosure will support one or two controllers. Each controller can access one or more groups of disks in the enclosure. A group of disks within an enclosure can be accessed by both controllers; however, only one controller may access the group at a time. This form of dual attachment is useful only for implementing failover of the controllers (or hosts) without moving cables. Each controller within an enclosure can also be tied to more than one host; however, only one host may mount the associated file system at a time, and therefore this form of dual attachment is also only useful to implement failover. The controllers in an enclosure can be connected to different hosts, and in fact this configuration is specified for some DAACs.

The throughput to the SGI arrays is limited by the either the interface mechanism (at large request sizes) or by overheads in the host and the controller (at small request sizes). The SCSI-2 interface used in the SGI arrays allows a maximum bandwidth of 20 MB per second to a channel. At small request sizes, however, the overheads are expected to limit performance to not more than eight MB per second per controller. In order to avoid contention on the SCSI-2 channels that would significantly reduce throughput for large block requests, and would also modestly impact throughput for small block requests, each array controller is configured on its own SCSI-2 channel.

The number of SCSI-2 channels and array controllers is determined by dividing the throughput rate required for a system (determined from the static and dynamic modeling results) by the eight MB per second rate per channel assumed for small request sizes. This number is then rounded upward. (Although the disk I/O performed by a system will actually be a mixture of large and small block requests, no effort is made to determine an average throughput rate; rather, the slower rate associated with small block requests is used to represent the entire load. This increases the design margin for these components.) The number of enclosures and racks is then matched to this number of enclosures.

SGI offers 2 GB and 4.3 GB disks in its arrays and soon will also offer a 9 GB disk. The 2 GB disks are not specified for use in SPRHW because the SPRHW disk requirements are too large to make use of these drives practical.

Although the number of groups and size of disks within groups can be varied from one controller to another, the number of such variations assigned to a single host has been deliberately limited. This is because it is expected that file systems will be defined so that they are striped across multiple controllers. Such striped file systems are necessary to support very high speed devices such as the HiPPI network. Striped file systems can only be built using partitions of uniform size, which are most easily built when each controller within the stripe is controlling the same configuration of disks.

The disk enclosures will be configured with three power supplies (one redundant supply) and a minimum amount of write cache. (Vendor benchmarks suggest that write cache will not significantly improve performance for the request mix that SPRHW is expected to produce.)

Backup and Update Devices. Each science processor will be equipped with an internal CD-ROM, as this is the delivery mechanism employed by SGI for updates to operating system software.

The SPRHW suite will include two tape libraries to provide backup capability for SPRHW and for other subsystems, using the FDDI backbone network. For Ir1 and Release A, Exabyte EXB-210 and EXB-218 tape libraries were purchased and put in place at LaRC. The EXB-210 is an 8 mm Digital Audio Tape (DAT) library with eleven cartridge slots and two drives. The EXB-218 is a 4 mm DAT with 19 cartridge slots, two operational drives, and one hot spare drive. These systems will continue to be used for Release B.

Monitors. In general, the science processors are to function as compute servers, and are not configured with sophisticated graphics processing capabilities. It is expected that the machines will be housed in raised floor environments where space is at a premium. The science processing hosts require a console in order to monitor the system at startup, and to perform some system administration tasks; however, this console need not be a bulky graphics monitor. ECS Operations and Maintenance are investigating ways to rack the system consoles, and to allow systems to share the same console.

Network Interfaces. An FDDI subnetwork will be implemented at LaRC to support the Planning and Data Processing Subsystems (PLS/DPS). Each processing component of SPRHW (including the Queuing Server and the Production Planner Stations) will be interfaced to the PLS/DPS FDDI subnetwork; the design of this subnetwork is specified in the Design Specification Overview, CD-305-020-002, and in Section 3.4.1 of this document.

At LaRC the data transfer requirements between DPS and DSS necessitate the implementation of a switched HiPPI network. The HiPPI network will be implemented via a central HiPPI switch with switched 800 Mbps interface ports connected directly to the SPRHW and DSS hosts; this design is also documented in the Design Specification Overview and in Section 3.4.1 of this document.

Enclosures. The actual enclosures for the science processors and their disk arrays are generally of little interest, except that they impose constraints on expansion capabilities, and they occupy floor space at the DAACs. All of the science processors in the LaRC DAAC will use Power Challenge XL enclosures. The Power Challenge XL uses a full height cabinet.

Rack-mounted enclosures are specified for the external disk storage units, in order to save floor space. Units that were cabinet mounted in Ir1 and Release A will be rack-mounted for Release B.

3.4.2.7.1.4 SPRHW Design Discussion

This section provides a discussion of how the specification provided in Section 3.4.2.7.1.3 meets the requirements identified in Section 3.4.2.7.1.1. The sections below provide a general discussion of how the specifications meet the sizing and expandability requirements, and what the general SPRHW failure recovery strategy will be.

Sizing and Expandability. It is important to note that although the science processors are divided into clusters of resources configured to deal with specific processing requirements, and are expected to be tuned and allocated for these purposes, this does not imply that the use of a processing cluster, or a processor within that cluster, is limited to one operating mode or

instrument. The DAAC operations staff, using the planning tools provided by ECS, will determine how the DAAC work load is distributed across the pool of SPRHW resources at the DAAC.

Table 3.4.2.7.1-2, *Comparison Of LaRC SPRHW Requirements And Specification*, compares the resources required at LaRC to those provided by the specification for each epoch of interest. The required values represent 1.2X, 2.2X, and 4.2X levels of processing (at the appropriate epoch for each instrument), calculated from the static model. The row titled "Specified" shows what the planned configuration provides. Each SCSI-2 disk channel is estimated to provide eight MB per second throughput, and each HiPPI interface is assumed to provide 70 MB per second throughput. The value for Specified Interleaving for each epoch is the sum of the number of memory leaf controllers for all SPRHW systems in the specification (typically eight per system). Note that question marks are shown for the memory requirements, as these requirements are not accurately known at this time. Requirements for disk space, disk bandwidth, and network bandwidth for AI&T are set to zero, as AI&T requirements for resources other than processing are not easily predictable; however, the specification provides ample resources in these categories. The table shows that the specification meets or exceeds the requirements for each resource at each epoch.

Table 3.4.2.7.1-2. Comparison of LaRC SPRHW Requirements and Specification

			c 3Q97	g 3Q98	k 3Q99	o 3Q00
CPU	Throughput (MF)	Required	7,240	8,134	15,062	25,339
		Specified	12,600	12,600	15,625	25,525
RAM	Size (MB)	Required	?	?	?	?
		Specified	12,288	12,288	14,336	18,432
	Interleaving (N)	Required	?	?	?	?
		Specified	48	48	56	72
Disk	Space (GB)	Required	60	562	1,264	1,806
		Specified	1,180	1,180	1,900	2,476
	Bandwidth (MB/s)	Required	2.7	36.7	70.7	102.0
		Specified	152.0	152.0	176.0	240.0
Network	Bandwidth (MB/s)	Required	1.2	11.3	21.3	30.2
		Specified	420.0	420.0	490.0	630.0

The ability to expand the SPRHW configurations merits a separate discussion. A DAAC's science processing capabilities can be expanded in two ways: by adding resources to existing hosts and disk arrays, and by adding new hosts and disk arrays. The former approach is referred to as upgrading in place.

For the SMP hosts, the ability to upgrade in place is limited by the number of backplane slots. The Power Challenge XL has 15 backplane slots, although this limit can be expanded to 25 by using a backplane extender. The maximum number of processors supported by Irix 6.2 on the Power Challenge is 36; this would require nine processor boards. Up to eight memory boards can be used, although only eight leaf controllers can be configured on those boards. Up to six IO4 boards

can be put in the system. Thus a maximally configured Power Challenge XL would require 23 of the 25 available slots. Such a system would provide 4.95 derated GF, with 16 GB of RAM. The I/O subsystem would support multiple HiPPI connections and SCSI-2 controllers. (In benchmarking experiments at the University of Minnesota, Challenge systems have been configured with as many as four HiPPI interfaces and as many as 40 SCSI-2 controllers but not simultaneously.)

At Epoch k, the specification calls for the use of seven Power Challenge XLs at LaRC. If each of these systems was upgraded to the maximum configuration described above, these systems would provide an aggregate processing capability of 34.65 GF (derated). This would provide 122 percent growth over the current configuration. The aggregate memory would grow by 700 percent, and the number of disk channels would increase 1172 percent.

If growth beyond these scales is required, it will be necessary to add more hosts to the SPRHW configuration. The HiPPI switch supports 16 ports, but if necessary the switch can be cascaded; adding a second switch would thus provide 30 ports for DPS and DSS hosts. If the ratio of DPS hosts to DSS hosts is assumed to be two to one, this would allow connections for 20 DPS hosts, or a maximum SPRHW compute power of 99 GF (derated). This is roughly 400% of the 25.25 GF specified for LaRC at Epoch o.

If processing power greater than 99 GF (derated) is required, it would be necessary to modify the design slightly by adding a second HiPPI network, and partitioning the HiPPI network traffic. (Cascading the HiPPI network to three switches is not recommended by the vendor, although it may be feasible.)

It should be noted that adding SPRHW hosts potentially has the effect of increasing network traffic, because the SPRHW storage subsystems are locally (host) attached, and can not be shared between hosts. As more hosts are added to process the same load, the potential that data will have to be replicated to multiple host file systems increases, and the aggregate file system size, disk I/O, and network I/O also increase.

Failure Recovery Strategy. The general strategy in the event of the failure of an SPRHW component is to re-distribute the first-time processing load to SPRHW components having sufficient resources to handle the load. Reprocessing and AI&T, although important to ECS over the long haul, can be delayed or reduced until the failed components are replaced. Because SPRHW has redundant capacity to support AI&T and reprocessing, and because scheduling of SPRHW resources is performed dynamically by the queuing and planning resources, most failures will have little impact on the timeliness of production processing.

SPRHW Failure Recovery. The DPS reliability, maintainability, and availability (RMA) requirements of 96% availability and mean down time of less than four hours are met by the SPRHW design. An analysis demonstrating this is provided in *Availability Model/Predictions for the ECS Project*, 515-CD-001-003.

In the event of a failure in the SPRHW suite, the DAAC operators will isolate the fault and determine its severity. Vendor maintenance will be called and repair and/or replacement of the affected part will be initiated.

Depending on the severity and location of the failure, other management actions and decisions may be required. If the failed component only affects resources that were allocated to AI&T, there is no impact to DAAC production operations, and no further action to recover production activity is required.

If the failed component affects resources that were allocated to production, an assessment will be made whether any PGEs need to be re-started or re-run because of possible corruption. If this is the case, ECS production planning capabilities will be used to reschedule the PGEs. If the failed component has affected a limited subset of the production resources, the MSS Subsystem will notify the Queuing Server that these resources are no longer available, and the Queuing server will appropriately stop scheduling these resources. If the remaining resources available for production processing are sufficient to keep up with the first-time processing plan, an immediate re-plan may not be necessary; if the first-time processing PGEs are prioritized above reprocessing PGEs (generally the case), the first-time processing will remain on schedule. If it is necessary to shut down SPRHW resources in order to facilitate a repair, a re-plan can be generated, allocating time in the schedule for the repair.

If a failure reduces the working SPRHW resources allocated to production below the level required to support first-time processing, the DAAC operators may choose to re-allocate AI&T resources to production. This may be done in one of two ways: by using the production and queuing software to logically re-allocate and re-plan the use of the system's resources, or by using the physical hardware in the AI&T hosts to replace failed hardware in the production hosts. The first approach, using the ECS system software to schedule production work on the AI&T platforms, is generally preferred, because it is not intrusive to the hardware. However, there may be circumstances when the second approach is simpler and faster. If the AI&T software environment is significantly different than the production environment (suppose AI&T was testing a new version of an operating system), and if the hardware failure in production is simple (a failed power supply), then changing hardware may be easier than changing software.

At LaRC, for Epochs c and g, a system has been configured for AI&T which has sufficient resources to take over the first-time processing responsibilities of any other system in the complement. At Epochs k and o, because of reprocessing capacity, each group of systems designated for an instrument has sufficient resources to perform first-time processing even if one machine becomes unavailable.

Data backup and recovery are a comparatively minor issue for SPRHW, because SPRHW does not provide long term, secure storage for data. Although SPRHW has large storage arrays, these arrays are used to hold ancillary data files and data granules for short periods of time. If a file system failure occurs within SPRHW, the algorithms, ancillary data files, and data granules can be recovered from the Data Server. Therefore the only data that would be backed up from SPRHW, and restored in the event of a failure, would be operating system and configuration files (i.e., the system disk). In the event that recovery is needed, incremental and full backup tapes would be used to re-build the required file systems.

Network Failure Recovery. The dual-ring implementation of FDDI provides a significant degree of fault tolerance. Most media failures within the FDDI fabric will not result in any loss of service and no reconfiguration would be necessary in these cases. Given the inherent fault tolerance of

FDDI, it is not required to have multiple physical communications paths to each host. Hosts within SPRHW will use dual-attached station cards.

Failure recovery for the HiPPI switch will be supported by stocking spares for the Line Replaceable Units of the switch (power supplies, interface cards, fan). If an individual interface card fails, a host can be re-configured to a hot spare interface card by moving two cables and sending a single software command to the switch; this process can be done in minutes, without disrupting other hosts. If the control module fails, it will be swapped out and the switch will be reconfigured. In the very rare event of an entire switch chassis failure, the switch would either be replaced or repaired. All of these failure recoveries involve activities on the switch; modifications to the attached hosts are generally not required.

3.4.2.7.2 AQAHW

QA of ECS products will include non-science QA, in-line QA, and SCF-based QA. Non-science QA will generally entail data integrity checks on data products and metadata. In-line QA is a form of science QA in which the content of the QA is evaluated using science algorithms. Processing capacity for in-line QA, to the extent that it is specified in the AHWGP inputs and Technical Baseline, is included in the Science Processing HWCI (see Section 7, SPRHW). SCF-based QA is also a form of science QA, and is specified by the product development team. ECS provides support for SCF-based QA to the extent of providing archive and communications capacity for the SCFs to sample the products for QA.

The types of non-science QA to be performed at the DAACs will be specified by the DAAC operations staff. These requirements are as yet largely unspecified. The sections below define the working assumptions being made as to AQAHW requirements, and the specifications which flow from those assumptions.

3.4.2.7.2.1 AQAHW Requirements Analysis

The requirements for AQAHW are based upon the need to have a DAAC-based quality assurance capability to ensure the integrity of the products produced by the DAAC.

DAAC-based non-science QA processing requirements are to be defined through interactions with the DAAC operations personnel. The current design assumption is that DAAC-based non-science QA processing will be performed at the DAACs in parallel with the other forms of QA (SCF-based QA and in-line QA). The design baseline thus includes a local (DAAC) QA workstation to support these DAAC data integrity checks. This local QA workstation is actually similar to a Science User workstation equipped with core Client Subsystem functionality. The QA workstation acts as a client to the Data Processing and Data Server Subsystems. The current operations concept assumes that the QA workstation hosts ECS as well as DAAC supplied processes (as deemed necessary by the DAAC operations personnel), which use a subset of the ECS services to "pull" production data sets using the subscription mechanism. The need for visualization support will be explored as product specific QA processes and requirements are worked jointly with the DAAC operations teams.

3.4.2.7.2.2 AQAHW Technology Assessment

The AQAHW requirements identified to date do not present any significant technical challenges, and therefore no special technology assessment efforts (prototyping, benchmarking, or product evaluations) have been performed to support the specification of AQAHW.

3.4.2.7.2.3 AQAHW Specification

The AQAHW hardware suite at LaRC will consist of a graphics workstation connected to ECS via the PLS/DPS FDDI subnetwork.

The AQAHW graphics workstation was selected to provide a software execution environment equivalent to the AI&T software execution environment in order to facilitate use of the AQA workstation for AI&T when necessary. The AQA workstation is also equipped with a graphics capability sufficient to support sophisticated visualization techniques.

The AQA workstation will be an SGI Indigo2 IMPACT 10000 workstation equipped with 128 MB of RAM. This workstation uses the R10000 chip, which is expected to be available in a 275 MHz implementation in the second half of 1996. The workstation will be equipped with an internal CD-ROM and two 4.3 GB internal disk drives. A SCSI-2 connection will be used to access an external disk array; the external disk array will consist of twenty 4.3 GB disks controlled by a single storage processor and managed in a RAID 5 configuration. This will provide 69 GB of external disk space to the workstation.

The Indigo2 IMPACT 10000 enclosure is a desktop configuration. The workstation will be configured with a 19 inch color graphics monitor.

The AQAHW workstation will reside on the ECS PLS/DPS FDDI network. This network will provide direct access between AQAHW components and the remainder of the PLS/DPS suite. The AQAHW components may also communicate with other ECS components and the external world via the DAAC FDDI switch.

3.4.2.7.2.4 AQAHW Design Discussion

Because the requirements for non-science quality assurance at the DAACs are not well defined, a minimal complement of AQAHW has been specified for each DAAC. This complement could be expanded as necessary in a variety of ways.

If LaRC expresses a need for additional workstations for QA personnel, but does not require visualization support for these staff, X-terminals could be assigned to the AQAHW to provide these seats. These X-terminals would be configured in the same way that the AI&T X-terminals are configured (see Section 3.4.2.7.3.3), and would be attached to the ECS PLS/DPS ethernet subnetwork.

If additional visualization seats were required, additional workstations (configured as above) could be added to the ECS PLS/DPS FDDI subnetwork to support the requirement.

If the data requirements from the Data Server (or from Science Processing) for QA are so large as to burden the ECS PLS/DPS FDDI subnetwork, the AQA workstation may be configured with a

HiPPI interface and connected to the DSS/DPS HiPPI network. This would be subject to the availability of a port on the HiPPI switch, and a slot within the workstation. (The workstation graphics boards and interface boards contend for the same set of slots; a workstation configured with the maximum graphics capabilities offered for this product Maximum IMPACT graphics could not support multiple network interfaces.)

If additional memory is required in the AQA workstation, its memory can be expanded to a maximum of 640 MB.

The Indigo2 IMPACT 10000 workstation is offered with three graphics options; the entry-level option has been selected for use in the AQA configuration. If more sophisticated graphics are required, the graphics in the workstation can be upgraded, subject to the availability of a backplane slot as noted above.

3.4.2.7.3 AITHW

The Algorithm Integration & Test (AI&T) HWCI (AITHW) provides hardware resources to support the integration and test of science software at the DAAC, and system level validation, integration, and test. It is important to note that this HWCI provides workstations and tools for software integration and test, but does not provide the compute environment or compute capacity required for science software test. This integration and test compute capacity is included in the Science Processing HWCI (SPRHW).

3.4.2.7.3.1 AITHW Requirements Analysis

The requirements for AITHW are based upon the need to have a software development, configuration management, and test environment at the DAAC to support the integration and test of science software delivered to the DAAC by the instrument teams.

The AI&T activity is expected to be characterized by intense activity at certain program milestones (the initial delivery of science software, for instance) and a much lower level of activity on an ongoing basis. There are no explicit ECS requirements identifying capacity, throughput, or response time requirements for AITHW; therefore, the sizing of the hardware has been based upon the size of the integration and test effort anticipated for each instrument and upon experience with the ECS Ir1 release.

The number of AI&T seats provided at each DAAC has been specified by allocating two seats for each instrument supported at the DAAC. At LaRC, two seats are provided for each of CERES TRMM, CERES AM-1, MISR, MOPITT, and SAGE-III. The number of AI&T stations must be expandable to provide surge capacity during especially busy periods. The AI&T stations must provide the capability to support the AI&T tools, which provide graphical user interfaces. The AI&T stations do not have stringent reliability, maintainability, availability, and backup requirements, as they do not directly support production processing. The AI&T stations must be compatible with and interoperable with the target SPRHW hardware and the AI&T tools server.

The AI&T tools server must provide sufficient capacity to support the use of the development, configuration management, and test tools by the AI&T stations. Because the number of AI&T

stations at LaRC will be ten, and the tools server is not also acting as the software build or test environment (which is provided by SPRHW), the AI&T server requirements are minimal. The AI&T tools server does not have stringent reliability, maintainability, availability, and backup requirements, as it does not directly support production processing. The AI&T tool server must be compatible with the AI&T stations and the target SPRHW hardware.

AITHW must provide a network printing capability to support the AI&T task.

3.4.2.7.3.2 AITHW Technology Assessment

The AI&T hardware requirements do not present any significant technical challenges, and therefore no special technology assessment efforts (prototyping, benchmarking, or product evaluations) have been performed to support the specification of AITHW.

Because the CPU resources required for AI&T will be provided by target machines in the SPRHW suite, or by the AI&T tools server, the AI&T station requirements can be met with X-terminals. The use of X-terminals for these stations has several advantages, including cost effectiveness, ease of maintenance, expandability, and compatibility.

3.4.2.7.3.3 AITHW Specification

The AI&T hardware suite at LaRC will consist of ten X-terminals, one server, and one network printer. These components will be connected to ECS via the ECS local area network at the DAAC.

AITHW Components. The AITHW components will consist of workstations, a tools server, and a network printer.

AITHW Workstations. The requirements for AITHW workstations will be met at minimum cost by providing NCD HMX-PRO X-terminals. These X-terminals will be configured with 20" color monitors and 16 megabytes of memory to support the use of the tools specified for the AI&T tools server. Users will log on to either the AI&T tools server to use its software, or to a target machine in the SPRHW suite to build and test software in that environment.

The X-terminals will be configured to allow the addition of memory if experience indicates this is necessary; experience with Ir1 resulted in upgrades to the memory of X-terminals purchased for Ir1. Additional X-terminals can be purchased and integrated quickly and easily if expansion of the AI&T seating capacity at a DAAC becomes a critical requirement. Other ECS operations workstations, such as those in the Algorithm Quality Assurance HWCI and the Planning Subsystem HWCI, can also be used to support short term increases in AI&T activity.

Note that LaRC received a Sun workstation (AITHW-LaRC-1) for use as an AI&T workstation in Release A. This workstation will be replaced by an X-Terminal, and will be re-used elsewhere in an ECS DAAC. LaRC also received three X-terminals (SPRHW-LaRC-2, -3, and -4) as part of its Ir1 and Release A science processing complement. Although these X-terminals will not be re-designated (that is, with AITHW-LaRC-N designators), they will be included in the count of AI&T workstations.

AITHW Server. LaRC was provided with a Sun 20/50 SPARCstation for Ir1 and Release A. This system is configured with 128 MB of RAM and four gigabytes of internal disk. The processing requirements for the AI&T server are minimal and this equipment is expected to be satisfactory for Release B.

AITHW_Network_Printer. The AITHW network printing requirement will be satisfied by the printer provided for Ir1, an HP LaserJet 4M+ with 14 megabytes of RAM. This printer has 12 page per minute throughput.

AITHW Interfaces. The AITHW server will reside on the ECS PLS/DPS FDDI subnetwork. The AITHW workstations and printer will reside on the ECS PLS/DPS ethernet subnetwork. These networks will provide direct access between AITHW components and the SPRHW suite. The AITHW components may also communicate with other ECS components and the external world via the DAAC FDDI and ethernet switches.

3.4.2.8 MSS and CSS Subsystems

The MSS and CSS Subsystem hardware have been sized and configured in a redundant configuration in order to provide for high availability of communications infrastructure and management services. The sizing rationale, therefore, applies to both MSS and CSS servers and will be presented in a single subsection.

The MSS Subsystem consists of a single hardware configuration item (MSS-MHWCI), which provides the servers, workstations, and printers needed for all local system management functions. The MSS-MHWCI provides processing and storage for the following MSS software components:

- Management Software Configuration Item (MCI) - provides system monitoring and control (via HP Openview), the database management system (Sybase), trouble ticketing (Remedy), fault and performance management (Tivoli), physical configuration management (Accugraph), security management, accountability management, billing and accounting system, mode management service, performance trending capability, report generation and distribution, and management data access (custom code/scripts used to import log file data to the relational data base management system)

- Management Logistic Configuration Item (MLCI) - Site and SMC maintenance and operations staffs will rely on configuration management to provide software change control (Clearcase), change request management (DDTS), baseline management (XRP), inventory/logistics/maintenance (ILM) management, training management, policy and procedure management, software distribution management (Tivoli), and software license management.

- Management Agent Configuration Item (MACI) - Agents are processes used to monitor and/or control managed objects distributed across heterogeneous platforms. Current COTS technology for network management uses network protocols such as simple network management protocol (SNMP) to provide a way for the manager, the managed objects, and their agents to communicate. SNMP defines specific messages, referred to as commands, responses, and notifications.

The CSS Subsystem consists of a single hardware CI (CSS-DCHWCI), which provides the server for all CSS functionality. CSS contains a single CI, the Distributed Communications CI, which provides the following services:

- Common Facility Services - includes electronic mail, file access, bulletin board, virtual terminal, and event logger services
- Object Services - includes security, naming, message passing, event, thread, time and life cycle services
- Distributed Object Framework - includes OODCE framework functionality.

3.4.2.8.1 Rationale

The MSS/CSS processing complement for LaRC was designed and sized for both the TRMM and AM-1 missions. The sizing of MSS/CSS subsystem hardware is based on the February 1996 version of the technical baseline. Storage requirements have been rounded upward.

Processing Requirements **—** Processing requirements for the MSS and CSS subsystem are driven by the following types of transactions:

- HP Openview data collection from managed objects and ad hoc queries (server)
- Conversion/import of HP Openview and log file data to MSS Sybase DBMS (server)
- DBMS usage for report generation/ad hoc queries (server)
- Fault & performance management notification (server)
- Trouble ticketing (server)
- Order request tracking (server)
- Billing & accounting (workstation)
- Mode management (server)
- Usage for configuration, baseline, training, license, inventory, change request, software distribution, inventory/logistics/maintenance and associated report generation (workstation)
- DCE logical server transactions (directory, security, time).

Server Sizing — ECS already has experience with many of the COTS products to be loaded on the MSS server from previous work in Evaluation Packages (EPs) and EDF installations. Based on this experience, a profile of the MSS/CSS server that is operating under nominal load (e.g., HP Openview map is displayed, but no collections are in process) has been developed. To this, processing requirements have been added for specific types of transactions.

In the EDF, an HP 9000/735/125, rated at 160 MIPS, was loaded with HP Openview, DCE client, Sybase server, X-server, and operating system. Tests were run to examine the impact of various types of HP Openview functions on CPU utilization. HP Openview was configured to discover

approximately 500 nodes within EDF and then displayed them as a node map. Minimal status polling was performed at 15 minute intervals. A variety of HP Openview on-line reports were generated to show such items as packet throughput and CPU utilization. During the testing, processes resident on the server were monitored. CPU utilization remained extremely low (i.e., less than 3%) except during operator queries and initialization. At system start-up, initialization of the various daemons used by HP Openview generated a load of approximately 50%. After start-up, functions that involved initialization of x-windows screens (e.g., generation of the node map or display of a performance graph) generated loads of 25-40% for a brief (less than 15 seconds) period of time. Multiple SNMP queries on a router increased CPU usage to approximately 20 percent, with the primary driver appearing to be the x-windows server. Simultaneous queries of two routers (to two different x-window screens) consumed a total of 50-60% of the CPU. Based on this benchmark, we assume that a basic configuration of a server, including HP Openview, Sybase, DCE client, and the operating system will require approximately 72 MIPS, and will provide adequate resources for routine HP Openview operations. To this must be added processing capacity to handle DCE server functions, HP Openview monitoring, processing of log files, Sybase report generation/ad hoc query capability, Remedy trouble ticketing, Tivoli monitoring, Tivoli software distribution management, mode management, order request tracking, and mail.

HP Openview and log file-to-Sybase data conversion are primary processing drivers that are expected to vary by DAAC. Table 3.4.2.8.1-1 shows estimated numbers of transactions for HP Openview data collection. HP Openview data collection is driven by the number of managed objects to be monitored and the number and frequency of management information base (MIB) objects to be collected for each. Managed objects for each MIB type were counted based on the Release B hardware plan for LaRC. The number and frequency of data collection for each class of managed objects was provided by MSS developers as specified in the CSMS Database Design and Database Schema Specification, (311-CD-003-003, Appendix B). HP Openview provided an estimate of 100,000 instructions per transaction. Using this information, an average number of instructions per second required for HP Openview data collection was developed. These estimates appear to be reasonably in line with HP-provided performance information, which indicates that an HP 9000/735, a machine rated at 125 MIPS, is capable of performing approximately 1300 collections per second.

Table 3.4.2.8.1-1. LaRC HP Openview Collection Processing Requirement

	# MIB Objects	Average Size (Bytes)	LaRC Managed Objects	Collections per hour*	Collections per second	Estimated MIPS
Release B LaRC (Hosts, RDBMS, Router, hubs)	1,953	4	302	286,074	79	8

*Note that the number of collections per hour was derived by multiplying each class of MIB objects (e.g., MIB II objects) by the number of managed objects within that class, and summing the results.

An estimate of 100,000 instructions per transaction was assumed for the conversion of each logged event to Sybase, based on the number of source lines of code for the MSS MDA component involved and an estimate of instructions needed to update the Sybase database. Instructions per transaction was multiplied by the number of logged events, including both HP Openview events and events collected from applications via the logging API. HP Openview events (transactions) are described in the previous paragraph. The number of application-generated entries was developed using the following assumptions:

- One log entry is generated for every system transaction, by every process that is included in the transaction thread.
- The number of "pull" transactions is based on the user model and reflects user service requests by DAAC. Pull transactions (e.g., directory, inventory search requests) are assumed to generate a conservative estimate of 5 log entries each from CIDM and data server processes.
- Order request tracking is dependent on the request for data by a user and the request for status of a data product by a user. For every user request for data, an EcRequest is stored in the management DBMS and updated as required by the DSS. The transaction frequency for EcRequest storage is related to the number of granules requested by user per DAAC. Updates made by the DSS to the EcRequest are considered to be a small percentage of the total granules requested. For every user request for status of a product, the appropriate EcRequest is retrieved from the management DBMS and made available to the client.
- The number of transactions on the "push" side includes external (DAAC-to-DAAC and L0) file transfers (4 log files each), processing-to-archive requests (4 log files each), and PGE execution (2 log files each). Push transactions were based on AHWGP data, which showed that LaRC executes 778 PGEs per day and each PGE requires an average of 39 input/output file requests at Release B.
- In addition, major processes generate log entries of approximately 512 K (based on the MSS application MIB) once every 15 minutes. There are estimated to be 15 processes at each DAAC that will generate log entries every 15 minutes.
- Log files and HP Openview data will be kept for 14 days prior to archiving in long-term Data Server Storage.
- For Billing & Accounting, there are expected to be approximately 17,000 total daily user accounts. Each account will be logged and on demand available for information tracking. An approximate number of user accounts per DAAC was estimated from the February 1996 user pull technical baseline.
- To implement the mode management service, multiple modes are assigned to each logged activity and can be simultaneously executed. The overhead required to provide mode management capability is estimated at 30% of the total logged activity.

Log entry storage volumes are given in Table 3.4.2.8.1-2.

The MIPs required to import the total number of log files per day are given in Table 3.4.2.8.1-3.

Table 3.4.2.8.1-2. LaRC Log Entry Storage Volume - Release B

Log File	Log Events per Transaction	Transaction Frequency per Hour	Total Logged Events per Hour	Bytes per Transaction	Total Size of Bytes/Hr	14-Day Storage Requirements (MB)
User requests	10	324	3,240	420	1,360,800	458
Request tracking	10	162	1,620	420	680,400	229
PGE execution	2	98	196	420	82,320	28
External file transfers	4	9	36	420	15,120	5
Processing-to-Archive Requests	4	3822	15,288	420	6,420,960	2,158
Application MIB poll	15	7	105	512	53,760	18
Billing & accounting activity logging	2	92	184	420	77,280	26
Subtotal Rel B			20,669		8,690,640	2,920
Total Rel B (includes x30% for multiple modes)			26,870		11,297,832	3,796

Table 3.4.2.8.1-3. MDA Data Conversion to Sybase Processing Requirement

	Total HP Openview Events/Day	Total Log File Events/Day	MIPS for 8 hour Sybase import
Release B	6,865,776	644,880	26

At Release B, ad-hoc queries will be performed and statistical analysis collected from the Sybase database. Ad-hoc reports will be generated that include the following type of information; user accesses, trend analysis, fault occurrences, resource utilization, data production jobs and security events. Benchmarks are being run on a prototype Sybase database to evaluate performance. The

prototype database was developed to do real-time benchmarking queries of designated working attributes that are expected to be of reporting interest (i.e., performance).

DCE has been installed in the EDF and used in the Engineering Packages (EPs). Running on an HP 715, rated at 77 MIPS, the DCE server functions used 8% of the CPU, or approximately 6 MIPS. An analysis was performed to determine how much additional load would be placed on the DCE server at Release B.

Load imposed on the DCE server is a function of the number of directory, security and time look-ups from client applications. A client application maintains its own cache containing the most recently accessed directory and security information, and will only access the server when a user is not found in its own cache. Many client applications will only access other clients in the DAAC, and so will never exceed their cache. CIDM and the Data Server APC, however, will be directly accessed by external user clients and so will need to access directory and security information for each user access. At LaRC, the user model reflects a maximum of 324 users accessing per hour. Given that a directory and security lookup typically requires less than 0.5 seconds, it is unlikely that there will be more than 1-2 simultaneous hits on the DCE server. We estimate that 1 additional MIP processor capacity will be sufficient for the level of DCE accesses required.

In the EDF "Mini-DAAC" facility, Tivoli performance was evaluated on an HP 9000 J210/1 rated at 176 MIPS and with 256 MB of RAM. The Tivoli COTS package will be used primarily for performance management, fault management and software distribution (most likely through Tivoli Courier). Performance will be monitored, statistics collected, and faults detected via Tivoli GUI screens. Tests were run to determine Tivoli GUI screen CPU utilization. The benchmark was performed with one user and the following configuration; the Tivoli application, the platform operating system, xwindows, and the performance tool (glance plus). CPU utilization was minimal as expected with no applications running (.5% of the system CPU) and approximately 56 processes active.

Following initialization of the Tivoli application, CPU utilization remained low (<2%), with the Tivoli Management Enterprise (TME) desktop enabled and 61 processes active. CPU loading became more prevalent when an administration GUI was selected from the TME desktop. Peak utilization was recorded at 9% of the system CPU for a period of 10 seconds and 73 processes active. Max peak CPU utilization (approximately 11%) and IO throughput (13.5 MB/s) was recorded when enabling the policy region desktop. In steady state, CPU utilization measured approximately 3% of the system CPU. Opening multiple GUIs did not increase demands on CPU utilization (remained <11%) but linearly required more memory. The Tivoli vendor for this reason recommends 96 MB of RAM dedicated. Total CPU utilization allocation for Tivoli based on benchmark results is estimated at 11% of the system CPU or 20 MIPS. The targeted platform at each DAAC site will be upgraded from the platform used for benchmark calculations. Recognizing the emphasis by Tivoli for memory and moderate processing needs, additional processing and memory capabilities were added to the MSS management server in Release B to provide adequate resources in support of the Tivoli product.

Remedy was evaluated on the HP 9000/735/125 for CPU utilization. The application required very little CPU allocation (<1%). A more significant load was present when performing browse or ticket assignment functions (approximately 6%). Submittal and processing of a trouble ticket required less than 1% of the CPU capacity.

The server requirements, as dictated by the rationale given above, is synopsized in Table 3.4.2.8.1-4.

Table 3.4.2.8.1-4. CSS/MSS Server Configuration - Requirements Estimate

Server Load Sources	Estimated R-B MIPS
Basic configuration (includes HP Openview and DCE client)*	72
Additional HP Openview data collection*	8
Sybase Server and Client*	50
Tivoli *	20
Remedy	11
MDA (log conversion to Sybase)	26
MSS Agent*	3
DCE server (including additional processing for peak directory and security transactions)*	7
Word Processor	1
Spreadsheet	1
Other Common Services (Mail, file transfer, etc.)*	5
Total * These items were considered to be potentially active at the same time. MDA database update is assumed to be run in off-peak hours, and not concurrently with Sybase report generation functions.)	175

Workstation Sizing — There will be two MSS workstations at each DAAC site. Workstation # 1 will primarily contain the MLCI software. This includes software change management (clearcase), change request management (DDTS), baseline control management (XRP), software license manager and inventory/logistics/maintenance (ILM) management. Policy & procedures management and training management will be configured on workstation # 2. Each MSS Workstation will contain the Sybase client, DCE client, Tivoli client, MSS agent, and operator tools.

The DDTS tool was evaluated for performance in the EDF facility on a Sun SPARC 20/50 rated at approximately 130 MIPS and 64 MB of RAM (DAAC targeted platform will be an upgrade version). DDTS is the change request manager and maintains and tracks potential changes (via configuration change requests) to the ECS System. Configuration change requests (CCRs) will be created, logged into the DDTS database and tracked by a CM specialist. Tests were performed to

determine CPU utilization for implementation of these tasks. The benchmark was performed with one user, the DDTS application configured with the platform operating system, xwindows and the performance tool (proctool). Following initialization of the DDTS application, CPU utilization as expected was very low, <1%. For each instance that a CCR was either submitted, modified or logged, the CPU utilization remained below 3% and memory utilization less than 8%. Table 3.4.2.8.1-5 shows that processing utilizations increased significantly when queries were made to the DDTS database.

Table 3.4.2.8.1-5. DDTS Benchmark Results

Benchmark Test	Number of Records	CPU Utilization (% of system CPU)	Memory Utilization (% of system memory)	IO Throughput (KB/s)
CCR submittal/creation	-	1.4 %	6.6 %	4
CCR registration	-	2.3 %	7.1 %	4
EP4 database query	128	9.7 %	7.2 %	11.3
EP6 database query	279	13.5 %	7.3 %	11.7
EP4 + EP6 database query	407	16.5 %	7.5 %	11.8
DDTS (inclusive) database query	1232	30.8 %	7.5 %	12.2

CPU utilizations ranged from 10% of the system CPU to approximately 30%. The number of records in the development environment is expected to be substantially higher than at the DAAC sites. For this reason, a conservative estimate of a maximum of 400 records is used to result in a CPU utilization allocation of approximately 21 MIPS. Memory utilization and IO throughput were moderate and appeared constant for each test performed.

Processing requirements for baseline management COTS (XRP), was estimated from vendor specifications. For a 30 user system, the XRP vendor specifies a processing requirement of 100 MIPS. Each DAAC site is assigned to have 2 XRP users and therefore will require approximately 7 MIPS.

Vendor specifications suggest an allocation of 35 MIPS for the Clearcase Virgin Object Base (VOB) Server. The VOB server is the most compute intensive of the Clearcase server applications due to its required database processing. In the EDF, Clearcase was installed on a SPARCstation 10, equipped with 120 MB RAM, rated at 109 MIPS, and with an ethernet interface. The SPARCstation 10 was initially used for Tool kit development, as well as CM of the Evaluation Packages. With moderate numbers of users, the SPARCstation 10 provided good performance. At peak use (15-20 simultaneous users viewing items, manipulating the contents of the database, and executing directly out of Clearcase), performance was adversely affected. Usage at the DAAC is not anticipated to require more than 5 simultaneous users, frequency of use is anticipated to be much lower, and applications will not be executed from the Clearcase tool. Additional benchmarks will be run as ECS code and science algorithms become available to help determine the precise

Clearcase processing requirements at the DAAC. EDF experience suggests that a workstation configuration in the SPARCstation 20 range should be adequate to support Clearcase, other MLCI COTS, DCE and billing & accounting and Tivoli clients.

On MSS Workstation #2, the COTS products expected to exert the larger processing loads are billing and accounting (B&A) and performance trending. Other primary load contributions come from training management, policy and procedures, and the DBMS report generator.

Major B&A processing loads will occur during nightly batch imports to the sybase server which is not resident to MSS Workstation #2. Processing of B&A transactions such as accounts received, purchase orders placed and products delivered are expected to exert a moderate load on the MSS Workstation due to the expected number of user requests for data as provided by user modeling.

Selection for the performance trending statistical analysis package is in progress. Statistical and historical performance data will be analyzed to assure optimum usage of system resources. A determined number of performance attributes will be analyzed by the statistical tool. Performance trending and other resident COTS packages such as training manager, policy and procedure manager and DBMS report generator are expected to require a small to moderate load on the MSS Workstation.

Tables 3.4.2.8.1-6 and 3.4.2.8.1-7 show configuration requirements for the MSS workstations. They reflect a best estimate of load to be imposed on each MSS workstation. It assumes that most functions run concurrently. Operator functions can be spread across workstations in such a way as to balance processing loads.

Table 3.4.2.8.1-6. MSS Workstation #1 Configuration - Requirements

Workstation Load Sources	Estimated MIPS
Basic configuration (includes Clearcase and Operating System)*	50
Software License Management*	5
DDTS*	21
XRP*	7
Tivoli Client*	5
Sybase Client*	10
Word Processor	1
Spreadsheet	2
Graphics	1
Inventory/Logistics/Maintenance Management	15
MSS Agent*	2
DCE Client*	5
B&A Client*	5
Other Common Services (Mail, file transfer, etc.)	5
Total *These items are considered to be potentially active at the same time	125

Table 3.4.2.8.1-7. MSS Workstation #2 Configuration - Requirements

Workstation Load Sources	Estimated MIPS
Basic configuration (includes Billing & Accounting and Operating System)*	35
Training Management*	5
Performance Trending*	15
DBMS Report Generation*	10
Policy & Procedures*	5
Tivoli Client*	5
Sybase Client*	10
Clearcase Client*	10
Word Processor	1
Spreadsheet	2
Graphics	1
MSS Agent*	2
DCE Client*	5
Other Common Services (Mail, file transfer, etc.)	5
Total *These items are considered to be potentially active at the same time	102

Storage Requirements — Major datastores for the MSS and CSS subsystems include: HP Openview files, application log files (including request order tracking, billing & accounting and mode management), the Management DBMS, and Clearcase-managed data for software change management.

Other datastores include DCE directory, security data, mail, trouble ticketing, Tivoli, DDTS, baseline control management data (XRP), ILM, billing & accounting client, training management, policy & procedures, and DBMS report generation.

The size of the data storage for HP Openview has been estimated from the determination of the frequency of transmission of the necessary information of all the appropriate attributes of the managed objects during one hour period. It was assumed that fourteen days worth of HP Openview data are stored.

A description of how application log file volume was estimated is in the previous section (Processing Requirements). Log file volume is provided in Table 3.4.2.8.1-2 based on an assumption of fourteen days storage prior to archiving in the data server archive.

The storage requirement for the Management DBMS was based on a worst case assumption that all the records from both the log files and HP Openview are stored in the Management DBMS, with an additional 10% for table overhead and summarization records. It is assumed that one months worth of data are maintained in the Management DBMS at a time.

Storage requirements for Clearcase are based on the assumption that Clearcase will store two copies of all source code (including ECS application source and algorithms) and two copies of all

executables. This will enable recovery of the previous version of any application if required. In addition, Clearcase will store test data and configuration files.

Tivoli sizing estimates are based on the number of performance attributes that will be monitored as specified by MSS developers in the Release B CSMS System Management Subsystem Design Specification, (305-CD-029-001, Section 6.6). These include system, application, process, and disk performance metrics. The monitoring frequency is dependent on the performance attribute. A worst case polling frequency of once per minute for all attributes was used in sizing calculations. The size of a typical fault/performance notification was estimated at 256 bytes.

Approximately 400 trouble tickets per day are estimated to be assigned, or approximately 17 per hour. The size of a trouble ticket is approximately 256 characters. Trouble ticket frequency and size are worst case.

To determine DDTS sizing requirements, the frequency non-conformance reports (NCRs) are generated on a daily basis was identified with the report size. A NCR was evaluated due to its similarity to a configuration change request (CCR). The number of CCRs generated at the DAAC sites are not considered to be more necessary in a developmental environment.

There are expected to be approximately 16 periodic reports that will be produced on a daily, weekly, monthly and annual basis. Reporting areas include data production, fault identification, user accesses, resource utilization, user services activity and trouble ticketing. The size of an average report is estimated to be 15 KB. The aggregate number of reports generated is approximately 1 per hour.

The cumulative datastores of XRP, ILM, billing & accounting client, training management, and policy & procedures was estimated based on vendor provided information and experience in the development facility.

Disk space requirements of the MSS management server COTS applications are listed in Table3.4.2.8.1-8. These applications will be stored in RAID and available for download to local disk. The RAID device interface is Fast/Wide SCSI which offers application access times comparable with local disk.

The storage requirement for the Sybase DBMS is estimated to be 9 GB, Clearcase 3.9 GB, Tivoli 114 MB, Remedy 2 MB, DDTS 7 MB, DBMS Report Generator 5 MB, and 145 MB for all other datastores combined. Storage requirements for DCE directory and security stores are based on the number of predicted users as provided by user modeling. The total storage requirement for CSS is estimated to be 496 MB for Release B as specified in Table 3.4.2.8.1-9.

Additional RAID storage is allocated for safeguard of HP Openview functions and storage of billing and accounting transaction logs. Other real time functions (i.e. Tivoli and Remedy) will be replicated to the CSS server. A copy of all management data will be stored in RAID on a daily basis and safestored into a DLT tape drive via the management backup server. As required, the management data will then be stored into ECS data server archive.

The total storage requirement for Release B is estimated to be between 28 and 30 GB as specified in Table 3.4.2.8.1-10 (includes additional storage for Sybase swap space).

Table 3.4.2.8.1-8. COTS Product Disk Requirements

COTS Product	Disk Requirement (MB)
HPOV	2,000
Tivoli	100
Trouble Ticket (Remedy)	50
Sybase Server	1,000
Clearcase	2,000
DCE Server	200
Accugraph	50
TOTAL	5,540

Table 3.4.2.8.1-9. LaRC CSS Release B Storage Requirement

CSS Data Store	Number of Users	Size of Record (# Bytes)	14-Day Storage Requirements (MB)
DCE Directory	17,000	1,000	238
DCE Security	17,000	1,000	238
Mail	348	4,000	20
Total Storage Requirement			496

Table 3.4.2.8.1-10. LaRC MSS Release B Storage Requirements

Datastore	Freq. of Events/Hr	Size in Bytes/ Transaction	Size in Bytes Transmitted/Hr	14-Day Storage Requirements (MB)
HP Openview Datastore	286,074	5	1,430,370	481
Application log files	26,870	420*	11,297,832	3,796
Sybase DBMS				9,409
Clearcase				4,000
Tivoli	1320	256	337,920	114
Remedy	17	256	4,352	2
DDTS	8	2400	19,200	7
DBMS Report Generator	1	15,000	15,000	5
Other Datastores (ILM, XRP, B&A client, training, policy & procedures)				145
Application Disk Space Requirements				5,400
Total Storage Requirement				23,359

*Application polling generates 512 byte logs. These have been included in the per hour total.

Processor Selection — Choice of the MSS/CSS Server platform was based on Release B processing requirements, COTS to be hosted on the platform, and price/performance data provided by EDS. Based on the Release B processing requirements,— a medium-range server class platform was chosen. HP is the preferred vendor, since HP Openview and OODCE will be principal COTS products on these platforms, and HP is one of the principal developers of DCE and OODCE.

3.4.2.8.2 Configuration

The following configuration will be provided for the LaRC LSM for Release B, which includes the MSS MHWCI and the CSS DCHWCI.

- MSS Local Management Server and CSS Communications Server: 2 HP 9000 J210/2 processors, 384 MB of RAM and 4 GB of storage.
- RAID Storage: 36 GB total storage
- Workstations:
  - 1 Sun SPARC Ultra 170, 128 MB of RAM and 8 GB of storage (This workstation will house configuration management software)
  - 1 Sun SPARC Ultra 170, 128 MB of RAM and 4 GB of storage
- Management Data Backup Server
  - 1 Sun Ultra 4-slot with 128 MB of RAM and 4 GB of storage
- Printer
  - 1 HP Laser Jet 4M+ Printer, 12 ppm/14 MB

The Langley DAAC will contain two primary servers for its LSM configuration, cross-strapped to RAID disk to enable warm backup. MSS and CSS applications will run on separate systems but in case of contingency, either system will be capable of running both subsystems.

The HP 9000 J210 is a high performance processor specifically designed for compute intensive and graphic applications.

The configuration at LaRC will include two Sun SPARC Ultra 170 workstations. One of the workstations which will house configuration management software will be configured with higher memory and higher storage (128 MB of RAM and 8 GB of hard drive).

The Langley DAAC will also include a Digital Linear Tape (DLT) Library which implements helical scan technology which averages 20GB storage capacity per cartridge with a data transfer rate of about 3 MB/second.

3.5 Hardware/Software Mapping

With the exception of the Client subsystem, each subsystem has been designed to incorporate hardware CIs that include the components (processors, servers, archive robotics, etc.) on which the software components run. Table 3.5-1 provides a mapping of Langley ECS Release B hardware components to the applicable software components.

Table 3.5-1. Langley Hardware to Software Component Mapping (1 of 9)

HWCI/Units	Subsystem	CSCI	CSC
SPRHW/science processors	Data Processing	AITTL	Binary File Comparison Utility Code Analysis Tools HDF File Comparison Utility Profiling Tools Standards Checkers
		PRONG	Data Pre-Processing PGE Execution Management Resource Management
		SDPTK	All CSCs
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
SPRHW/queuing management server	Data Processing	PRONG	PGE Execution Management Resource Management Product
	Planning	PLANG	All CSCs
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy

Table 3.5-1. Langley Hardware to Software Component Mapping (2 of 9)

HWCI/Units	Subsystem	CSCI	CSC
AITHW/AI&T DBMS server	Data Processing	AITTL	Binary File Comparison Utility Code Analysis Tools Data Visualization Tools Documentation Viewing Tools ECS HDF Visualization Tools HDF File Comparison Utility Profiling Tools Report Generation Tools Standards Checkers
		PRONG	COTS COTS Management Data Management
		SDPTK	All CSCs
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
AITHW/AI&T Operations workstations	Client	DESKT	All CSCs
		WKBCH	All CSCs, except - User Registration Tools
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services

Table 3.5-1. Langley Hardware to Software Component Mapping (3 of 9)

HWCI/Units	Subsystem	CSCI	CSC
	Data Processing	AITTL	Data Visualization Tools Documentation Viewing Tools ECS HDF Visualization Tools PGE Processing GUI PGE Registration GUI Product Metadata Display Tool SDP Toolkit-related Tools SSAP Processing GUI Update Data Server GUI Update PGE Database GUI
		PRONG	COTS COTS Management Data Management
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
AQAHW/QA workstations	Client	DESKT	All CSCs
		WKBCH	All CSCs, except - User Registration Tools
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Data Processing	PRONG	Quality Assurance Monitor Interface
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
PLNHW/PDPS DBMS server	Planning	PLANG	All CSCs
	Data Processing	PRONG	PGE Execution Management Resource Management
	Client	DESKT	All CSCs

Table 3.5-1. Langley Hardware to Software Component Mapping (4 of 9)

HWCI/Units	Subsystem	CSCI	CSC
		WKBCH	All CSCs, except User Registration Tools
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
ICLHW/ingest server	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Ingest	INGST	All CSCs
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
ACMHW/administration and operations workstations	Client	DESKT	All CSCs
		WKBCH	All CSCs, except - User Registration Tools

Table 3.5-1. Langley Hardware to Software Component Mapping (5 of 9)

HWCI/Units	Subsystem	CSCI	CSC
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
ACMHW/APC servers	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Data Server	SDSRV	All CSCs
		STMGT	Service Clients File Peripherals Resource Management
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
DIPHW/distribution servers	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Data Server	DDIST	All CSCs

Table 3.5-1. Langley Hardware to Software Component Mapping (6 of 9)

HWCI/Units	Subsystem	CSCI	CSC
		STMGT	Service Clients File Peripherals Resource Management
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
DRPHW/FSMS servers	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Data Server	STMGT	All CSCs
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
DRPHW/archive robotics - EMASS			******This has no CSCIs allocated to it. It is tied to the FSMS server***
DRPHW/DBMS servers	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
	Data Server	SDSRV	DB Wrappers (Illustra DBMS)

Table 3.5-1. Langley Hardware to Software Component Mapping (7 of 9)

HWCI/Units	Subsystem	CSCI	CSC
DDSHW/document server	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Data Server	DDSRV	All CSCs
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
DMGHW/data specialist workstations	Client	DESKT	All CSCs
		WKBCH	All CSCs
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
DMGHW/administration and operations workstations	Client	DESKT	All CSCs
		WKBCH	All CSCs
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Management	MACI	All CSCs

Table 3.5-1. Langley Hardware to Software Component Mapping (8 of 9)

HWCI/Units	Subsystem	CSCI	CSC
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
DMGHW/DBMS servers	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	Data Management	DDICT	All CSCs
		DIMGR	All CSCs
		GTWAY	All CSCs
		LIMGR	All CSCs
	Interoperability	ADSRV	All CSCs
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
MSS/MSS workstations	Client	DESKT	All CSCs
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
	ISS	INCI	All CSCs
	Management	MACI	All CSCs
		MCI	All CSCs
		MLCI	All CSCs
		MHCI	All CSCs
MSS/MSS LSM Server	ISS	INCI	All CSCs
	Management	MACI	All CSCs
		MCI	All CSCs
		MLCI	All CSCs

Table 3.5-1. Langley Hardware to Software Component Mapping (9 of 9)

HWCI/Units	Subsystem	CSCI	CSC
		MHCI	All CSCs
	Communication	DCCI	Electronic Mail Services Event Logger Services File Access Services Life Cycle Services Message Passing Services Thread Services Time Services
CSS/CSS server	Communication	DCCI	All CSCs
	Management	MACI	All CSCs
		MCI	Automatic Actions Fault Management Management Proxy Performance Management Performance Management Proxy
User workstation	Client	WKBCH	All CSCs
	Ingest	INGST	User Network Ingest Interface

Note: "All CSCs" refers to those CSCs for a specific CSCI that is identified in the Component Analysis Table in Section 3.3.2 of this document.

4. Future Releases

This document has described the design of ECS subsystems for the ECS portion of the Langley DAAC at Release B. Two other releases are currently planned. The next release, Release C, is scheduled for December 1999. The Release Plan Content Description for the ECS Project describes in detail the capabilities being provided in Releases C and D. In summary, Release C will provide evolutionary enhancements to Release B, Flight Operations for EOS PM-1, and the CORBA (Common Object Request Broker Architecture) and Trader designs.

An updated version of this document will precede Release C and will reflect the implementation corresponding to that release. A description of some of the enhanced interface functionality follows.

- The interface between the ECS portion of the Langley DAAC and the CERES SCF will be additionally required to support the CERES science software integration and testing of EOS PM-1 releases, the ingesting of the EOS PM-1 CERES Level 0 data, the ingesting of ancillary data for the EOS PM-1 CERES data processing, the EOS PM-1 CERES data processing, the scientific quality assurance for EOS PM-1 CERES processing performed by the CERES science team and the CERES data management team, and the distribution of the CERES data to the science community.
- The interface between the ECS portion of the Langley DAAC and the SAGE SCF will be additionally required to support the SAGE III science software integration and testing of Space Station releases, the ingesting of the Space Station SAGE III Level 0 data, the ingesting of ancillary data for the Space Station SAGE III data processing, the Space Station SAGE III data processing, the scientific quality assurance for Space Station SAGE III processing performed by the CERES science team and the CERES data management team, and the distribution of the CERES data to the science community.

In addition to the impacts caused by the addition of ECS capabilities, this document will also change to reflect how ECS interacts with DAAC-unique capabilities. The preliminary list of Langley Version 0 DAAC-unique elements follows:

a) Langley Version 0 DAAC Special services (see Langley DAAC Data Migration Plan): IMS, metadata, subsetting, and Read software,

b) Langley DAAC heritage systems: ERBE processing system, Whitlock's Surface Radiation Budget production system (recently funded Pathfinder research project), and Baum's Pathfinder Cloud Retrieval Algorithm Prototypes (recently funded Pathfinder research project),

c) Langley DAAC "live" data sets (see ICD Between ECS and Langley DAAC): ISCCP B3 and DX, SAGE II

d) DAAC-unique tools: databases from Langley Version 0 DAAC that allow User Services to maintain a history (e.g., LaRC DOTS)

e) Biomass Burning and SCAR-B campaign.

This page intentionally left blank.

Appendix A

Table A-1 provides detailed Science Processor Hardware (SPRHW) configurations for Release A and Release B.

Table A-1. Larc Detailed SPRHW Configurations for Release A and Release B

Release A	C 3Q97	G 3Q98	K 3Q99	O 3Q00
SPRHW-LaRC-1 Function: LaRC AI&T Cabinet: Power Challenge XL Console: Character CPU: 12 x 75 MHz R8000 RAM: 1 GB/2-way interleaved	**SPRHW-LaRC-1** Function: LaRC AI&T Cabinet: Power Challenge XL Console: Character CPU: 20 x 275 MHz R10000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-1** Function: LaRC AI&T Cabinet: Power Challenge XL Console: Character CPU: 20 x 275 MHz R10000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-1** Function: LaRC AI&T Cabinet: Power Challenge XL Console: Character CPU: 14 x 275 MHz R10000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-1** Function: LaRC AI&T (1 of 2) Cabinet: Power Challenge XL Console: Character CPU: 26 x 275 MHz R10000 RAM: 2 GB/8-way interleaved
VME-1: FDDI IO4: One HIO-1 (1,1): SCSI HIO-2 (1,2): Unused SCSI-0 (1,0,1): CD-ROM, 8 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks, SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-2 SP1 SCSI-4 (1,2,3): Unused	VME-1: FDDI IO4: Two HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI HIO-3 (2,1): Unused HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM, 8 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks, Two 4.3 GB Internal Disk SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	VME-1: FDDI IO4: Two HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI HIO-3 (2,1): Unused HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM, 8 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks, Two 4.3 GB Internal Disk SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	VME-1: FDDI IO4: Two HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI HIO-3 (2,1): Unused HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM, 8 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks, Two 4.3 GB Internal Disk SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	VME-1: FDDI IO4: Two HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI HIO-3 (2,1): Unused HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM, 8 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks, Two 4.3 GB Internal Disk SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2
RAID-1: 20 x 4.3 GB (Cabinet) RAID-2: 20 x 4.3 GB (Cabinet)	RAID-1: 20 x 9 GB (Rack 1) RAID-2: 20 x 9 GB (Rack 1)	RAID-1: 20 x 9 GB (Rack 1) RAID-2: 20 x 9 GB (Rack 1)	RAID-1: 20 x 9 GB (Rack 1) RAID-2: 20 x 9 GB (Rack 1)	RAID-1: 20 x 9 GB (Rack 1) RAID-2: 20 x 9 GB (Rack 1)
SPRHW-LaRC-2 Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-2** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-2** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-2** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-2** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal
SPRHW-LaRC-3 Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-3** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-3** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-3** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-3** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal
SPRHW-LaRC-4 Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-4** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-4** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-4** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal	**SPRHW-LaRC-4** Function: LaRC AI&T Cabinet: Desktop Console: Graphics CPU: X Terminal
SPRHW-LaRC-5 Function: LaRC AI&T Cabinet: Power Challenge XL Console: Character CPU: 12 x 90 MHz R8000 RAM: 1 GB/2-way interleaved	**SPRHW-LaRC-5** Function: LaRC AI&T Cabinet: Power Challenge XL Console: Character CPU: 12 x 90 MHz R8000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-5** Function: CERES TRMM Cabinet: Power Challenge XL Console: Character CPU: 12 x 90 MHz R8000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-5** Function: CERES TRMM Cabinet: Power Challenge XL Console: Character CPU: 12 x 90 MHz R8000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-5** Function: CERES TRMM Cabinet: Power Challenge XL Console: Character CPU: 12 x 90 MHz R8000 RAM: 2 GB/8-way interleaved
VME-1: FDDI IO4: One HIO-1 (1,1): SCSI HIO-2 (1,2): Unused SCSI-0 (1,0,1): CD-ROM, 4 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks,	VME-1: FDDI IO4: One HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI SCSI-0 (1,0,1): CD-ROM, 4 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks,	VME-1: FDDI IO4: One HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI SCSI-0 (1,0,1): CD-ROM, 4 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks,	VME-1: FDDI IO4: One HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI SCSI-0 (1,0,1): CD-ROM, 4 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks,	VME-1: FDDI IO4: One HIO-1 (1,1): HiPPI HIO-2 (1,2): SCSI SCSI-0 (1,0,1): CD-ROM, 4 mm Tape Stacker SCSI-1 (1,0,2): Two 2 GB Internal Disks,

Table A-1. Larc Detailed SPRHW Configurations for Release A and Release B

Release A	C 3Q97	G 3Q98	K 3Q99	O 3Q00
	Two 4.3 GB Internal Disk	Drives Two 4.3 GB Internal Disk	Drives Two 4.3 GB Internal Disk	Drives Two 4.3 GB Internal Disk Drives
SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1
SCSI-3 (1,2,2): Unused	SCSI-3 (1,2,2): RAID-1 SP2	SCSI-3 (1,2,2): RAID-1 SP2	SCSI-3 (1,2,2): RAID-1 SP2	SCSI-3 (1,2,2): RAID-1 SP2
SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused
RAID-1: 15 x 4.3 GB (Cabinet)	RAID-1: 20 x 4.3 GB (Rack 2)	RAID-1: 20 x 4.3 GB (Rack 2)	RAID-1: 20 x 4.3 GB (Rack 2)	RAID-1: 20 x 4.3 GB (Rack 2)
SPRHW-LaRC-6	**SPRHW-LaRC-6**	**SPRHW-LaRC-6**	**SPRHW-LaRC-6**	**SPRHW-LaRC-6**
Function: CERES TRMM	Function: CERES TRMM	Function: CERES TRMM	Function: CERES TRMM	Function: CERES TRMM
Cabinet: Power Challenge XL	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL
Console: Character	Console: Character	Console: Character	Console: Character	Console: Character
CPU: 18 x 90 MHz R8000	CPU: 18 x 90 MHz R8000	CPU: 18 x 90 MHz R8000	CPU: 18 x 90 MHz R8000	CPU: 18 x 90 MHz R8000
RAM: 1 GB/2-way interleaved	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved
VME-1: FDDI	VME-1: FDDI	VME-1: FDDI	VME-1: FDDI	VME-1: FDDI
IO4: One	IO4: One	IO4: One	IO4: One	IO4: One
HIO-1 (1,1): SCSI	HIO-1 (1,1): HiPPI	HIO-1 (1,1): HiPPI	HIO-1 (1,1): HiPPI	HIO-1 (1,1): HiPPI
HIO-2 (1,2): Unused	HIO-2 (1,2): SCSI	HIO-2 (1,2): SCSI	HIO-2 (1,2): SCSI	HIO-2 (1,2): SCSI
SCSI-0 (1,0,1): CD-ROM	SCSI-0 (1,0,1): CD-ROM	SCSI-0 (1,0,1): CD-ROM	SCSI-0 (1,0,1): CD-ROM	SCSI-0 (1,0,1): CD-ROM
SCSI-1 (1,0,2): Two 2 GB Internal Disks,	SCSI-1 (1,0,2): Two 2 GB Internal Disks,	SCSI-1 (1,0,2): Two 2 GB Internal Disks,	SCSI-1 (1,0,2): Two 2 GB Internal Disks,	SCSI-1 (1,0,2): Two 2 GB Internal Disks,
	Two 4.3 GB Internal Disk	Drives Two 4.3 GB Internal Disk	Drives Two 4.3 GB Internal Disk	Drives Two 4.3 GB Internal Disk Drives
SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1
SCSI-3 (1,2,2): Unused	SCSI-3 (1,2,2): RAID-2 SP1	SCSI-3 (1,2,2): RAID-2 SP1	SCSI-3 (1,2,2): RAID-2 SP1	SCSI-3 (1,2,2): RAID-2 SP1
SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused	SCSI-4 (1,2,3): Unused
RAID-1: 15 X 4.3 GB (Cabinet)	RAID-1: 15 x 4.3 GB (Rack 2)	RAID-1: 15 x 4.3 GB (Rack 2)	RAID-1: 15 x 4.3 GB (Rack 2)	RAID-1: 15 x 4.3 GB (Rack 2)
RAID-1: 15 X 4.3 GB (Cabinet)	RAID-2: 15 x 4.3 GB (Rack 2)	RAID-2: 15 x 4.3 GB (Rack 2)	RAID-2: 15 x 4.3 GB (Rack 2)	RAID-2: 15 x 4.3 GB (Rack 2)
SPRHW-LaRC-7	**SPRHW-LaRC-7**	**SPRHW-LaRC-7**	**SPRHW-LaRC-7**	**SPRHW-LaRC-7**
Function: LaRC Queuing Server	Function: LaRC Queuing Server	Function: LaRC Queuing Server	Function: LaRC Queuing Server	Function: LaRC Queuing Server
	SPRHW-LaRC-8	**SPRHW-LaRC-8**	**SPRHW-LaRC-8**	**SPRHW-LaRC-8**
	Function: LaRC AI&T	Function: CERES AM-1	Function: CERES AM-1 (1 of 2)	Function: CERES AM-1 (1 of 2)
	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL
	Console: Character	Console: Character	Console: Character	Console: Character
	CPU: 20 x 275 MHz R10000	CPU: 20 x 275 MHz R10000	CPU: 16 x 275 MHz R10000	CPU: 20 x 275 MHz R10000
	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved
	IO4: Two	IO4: Two	IO4: Two	IO4: Two
	HIO-1 (1,1): FDDI	HIO-1 (1,1): FDDI	HIO-1 (1,1): FDDI	HIO-1 (1,1): FDDI
	HIO-2 (1,2): SCSI	HIO-2 (1,2): SCSI	HIO-2 (1,2): SCSI	HIO-2 (1,2): SCSI
	HIO-3 (2,1): HiPPI	HIO-3 (2,1): HiPPI	HIO-3 (2,1): HiPPI	HIO-3 (2,1): HiPPI
	HIO-4 (2,2): Unused	HIO-4 (2,2): Unused	HIO-4 (2,2): Unused	HIO-4 (2,2): Unused
	SCSI-0 (1,0,1): CD-ROM	SCSI-0 (1,0,1): CD-ROM	SCSI-0 (1,0,1): CD-ROM	SCSI-0 (1,0,1): CD-ROM
	SCSI-1 (1,0,2): Two 9 GB Internal Disks	SCSI-1 (1,0,2): Two 9 GB Internal Disks	SCSI-1 (1,0,2): Two 9 GB Internal Disks	SCSI-1 (1,0,2): Two 9 GB Internal Disks
	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1	SCSI-2 (1,2,1): RAID-1 SP1
	SCSI-3 (1,2,2): RAID-2 SP1	SCSI-3 (1,2,2): RAID-2 SP1	SCSI-3 (1,2,2): RAID-2 SP1	SCSI-3 (1,2,2): RAID-2 SP1
	SCSI-4 (1,2,3): RAID-3 SP1	SCSI-4 (1,2,3): RAID-3 SP1	SCSI-4 (1,2,3): RAID-3 SP1	SCSI-4 (1,2,3): RAID-3 SP1
	RAID-1: 20 x 9 GB (Rack 3)	RAID-1: 20 x 9 GB (Rack 3)	RAID-1: 20 x 9 GB (Rack 3)	RAID-1: 20 x 9 GB (Rack 3)
	RAID-2: 20 x 9 GB (Rack 3)	RAID-2: 20 x 9 GB (Rack 3)	RAID-2: 20 x 9 GB (Rack 3)	RAID-2: 20 x 9 GB (Rack 3)
	RAID-3: 20 x 9 GB (Rack 3)	RAID-3: 20 x 9 GB (Rack 3)	RAID-3: 20 x 9 GB (Rack 3)	RAID-3: 20 x 9 GB (Rack 3)
	SPRHW-LaRC-9	**SPRHW-LaRC-9**	**SPRHW-LaRC-9**	**SPRHW-LaRC-9**
	Function: LaRC AI&T	Function: MISR (1 of 2)	Function: MISR (1 of 2)	Function: MISR (1 of 3)
	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL	Cabinet: Power Challenge XL
	Console: Character	Console: Character	Console: Character	Console: Character
	CPU: 16 x 275 MHz R10000	CPU: 16 x 275 MHz R10000	CPU: 24 x 275 MHz R10000	CPU: 24 x 275 MHz R10000
	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved	RAM: 2 GB/8-way interleaved

Table A-1. Larc Detailed SPRHW Configurations for Release A and Release B

Release A	C 3Q97	G 3Q98	K 3Q99	O 3Q00
	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2
	RAID-1: 10 x 9 GB (Rack 4) RAID-2: 10 x 9 GB (Rack 4)	RAID-1: 10 x 9 GB (Rack 4) RAID-2: 10 x 9 GB (Rack 4)	RAID-1: 20 x 9 GB (Rack 4) RAID-2: 20 x 9 GB (Rack 4)	RAID-1: 20 x 9 GB (Rack 4) RAID-2: 20 x 9 GB (Rack 4)
	SPRHW-LaRC-10 Function: LaRC AI&T Cabinet: Power Challenge XL Console: Character CPU: 16 x 275 MHz R10000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-10** Function: MISR (2 of 2) Cabinet: Power Challenge XL Console: Character CPU: 16 x 275 MHz R10000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-10** Function: MISR (2 of 2) Cabinet: Power Challenge XL Console: Character CPU: 24 x 275 MHz R10000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-10** Function: MISR (2 of 3) Cabinet: Power Challenge XL Console: Character CPU: 24 x 275 MHz R10000 RAM: 2 GB/8-way interleaved
	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2
	RAID-1: 10 x 9 GB (Rack 4) RAID-2: 10 x 9 GB (Rack 4)	RAID-1: 10 x 9 GB (Rack 4) RAID-2: 10 x 9 GB (Rack 4)	RAID-1: 20 x 9 GB (Rack 4) RAID-2: 20 x 9 GB (Rack 4)	RAID-1: 20 x 9 GB (Rack 4) RAID-2: 20 x 9 GB (Rack 4)
			SPRHW-LaRC-11 Function: CERES AM-1 (2 of 2) Cabinet: Power Challenge XL Console: Character CPU: 16 x 275 MHz R10000 RAM: 2 GB/8-way interleaved	**SPRHW-LaRC-11** Function: CERES AM-1 (2 of 2) Cabinet: Power Challenge XL Console: Character CPU: 20 x 275 MHz R10000 RAM: 2 GB/8-way interleaved
			IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-2 SP1 SCSI-4 (1,2,3): RAID-3 SP1	IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-2 SP1 SCSI-4 (1,2,3): RAID-3 SP1
			RAID-1: 20 x 9 GB (Rack 5) RAID-2: 20 x 9 GB (Rack 5)	RAID-1: 20 x 9 GB (Rack 5) RAID-2: 20 x 9 GB (Rack 5)

Table A-1. Larc Detailed SPRHW Configurations for Release A and Release B

Release A	C 3Q97	G 3Q98	K 3Q99	O 3Q00
			RAID-3: 20 x 9 GB (Rack 5)	RAID-3: 20 x 9 GB (Rack 5)
				SPRHW-LaRC-12 Function: MISR (3 of 3) Cabinet: Power Challenge XL Console: Character CPU: 24 x 275 MHz R10000 RAM: 2 GB/8-way interleaved
				IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2
				RAID-1: 20 x 9 GB (Rack 6) RAID-2: 20 x 9 GB (Rack 6)
				SPRHW-LaRC-13 Function: LaRC AI&T (2 of 2) Cabinet: Power Challenge XL Console: Character CPU: 28 x 275 MHz R10000 RAM: 4 GB/8-way interleaved
				IO4: Two HIO-1 (1,1): FDDI HIO-2 (1,2): SCSI HIO-3 (2,1): HiPPI HIO-4 (2,2): Unused SCSI-0 (1,0,1): CD-ROM SCSI-1 (1,0,2): Two 9 GB Internal Disks SCSI-2 (1,2,1): RAID-1 SP1 SCSI-3 (1,2,2): RAID-1 SP2 SCSI-4 (1,2,3): RAID-2 SP1 SCSI-5 (2,0,1): RAID-2 SP2
				RAID-1: 20 x 9 GB (Rack 6) RAID-2: 20 x 9 GB (Rack 6)

This page intentionally left blank.

Abbreviations and Acronyms

ACMHW	Access Control and Management HWCI
ACRIM	Active Cavity Radiometer Irradiance Monitor
ADC	Affiliated Data Center
ADEOS	Advanced Earth Observing System (Japan)
ADSHW	Advertising Service HWCI
ADSRV	Advertising Service CSCI
AHPCRC	Army High Performance Computing Research Center
AHWGP	Ad Hoc Working Group on Production
AI&T	Algorithm Integration and Test
AITHW	Algorithm Integration & Test HWCI
AITTL	Algorithm Integration and Test Tools (CSCI)
ALT	Altimeter
AM-1	EOS AM Mission Spacecraft 1, morning series
APC	Access/Process Coordinator(s)
API	Application Programming Interface
AQA	Algorithm QA
AQAHW	Algorithm QA HWCI
AS	Administrative Station
ASF	Alaska SAR Facility (DAAC)
ATL	Automated Tape Library
B&A	Billing and Accounting
BDS	Bulk Data Service
BONeS	Block-Oriented Network Simulator
CCR	Configuration Change Request
CDR	Critical Design Review
CDRL	Contract Data Requirements List

CD-ROM	Compact Disk, Read-Only Memory
CERES	Clouds, and the Earth's Radiant Energy System
CI	Configuration Item
CIDM	Client, Interoperability, Data Management
CLS	Client Subsystem
COLOR	Ocean Color (EOS Color Mission)
CORBA	Common Object Request Broker Architecture
COTS	Commercial Off-The-Shelf
CPU	Central Processing Unit
CSC	Computer Software Component
CSCI	Computer Software Configuration Item
CSA	Canadian Space Agency
CSMS	Communications and Systems Management Segment
CSS	Communication Subsystem (CSMS)
D3	Type of StorageTek digital helical scan tape
DAAC	Distributed Active Archive Center
DADS	Data Archive and Distribution System
DAO	Data Assmilation Office
DAT	Digital Audio Tape
DBA	Data Base Administration
DBMS	Data Base Management System
DCCI	Distributed Computing Configuration Item
DCE	Distributed Computing Environment
DCN	Document Change Notice
DDICT	Data Dictionary CSCI
DDIST	Data Distribution CSCI
DDSHW	Document Data Server HWCI
DDSRV	Document Data Server CSCI
DDTS	Distributed Defect Tracking System

DESKT	Desktop CSCI
DEV	Developed code
DFA/MR	Dual Frequency Altimeter /
DID	Data Item Description
DIM	Distributed Information Manager
DIMGR	Distributed Information Management CSCI
DIPHW	Distribution & Ingest Peripheral Management HWCI
DM	Data Management
DMGHW	Data Management HWCI
DMS	Data Management System
DMS	Data Management Subsystem
DOF	Distributed Object Framework
DPR	Data Processing Request
DPS	Data Processing Subsystem
DR	Data Repository
DRPHW	Data Repository HWCI
DSS	Data Server Subsystem
EBnet	EOSDIS Backbone Network
ECS	EOSDIS Core System
EDC	EROS Data Center (DAAC)
EDF	ECS Development Facility
EDHS	ECS Data Handling System
EDOS	EOS Data and Operations System
EGS	EGS Ground System
EMASS	E-Systems Modular Automated Storage Systems
EPn	Evaluation Package n (EP4, EP6, EP7, etc.)
ERS-1/2	European Remote Sensing Satellite
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System

ERS	Earth Resources Satellite
ESA	European Space Agency
ESDIS	Earth Science Data and Information System
EXB	Exabyte
FAX	Facsimile
FDDI	Fiber Distributed Data Interface
Flops	Floating Point Operations Per Second
FOO	Flight Of Opportunity
FOS	Flight Operations Segment
FP	Framing Protocol
F&PRS	Functional and Performance Requirements Specification
FSMS	File and Storage Management System
FWD	Fast Wide Differential
GB	GigaByte
GF	GigaFlops
GSFC	Goddard Space Flight Center
GTWAY	Version 0 Interoperability Gateway CSCI
GUI	Graphical User Interface
HDF	Hierarchical Data Format
HiPPI	High Performance Parallel Interface
HP	Hewlett Packard
HPOV	HP OpenView
HP-UX	HP UNIX
HTTP	HyperText Transfer Protocol
HWCI	Hardware Configuration Item
I&T	Integration and Test
ICD	Interface Control Document
ICLHW	Ingest Client HWCI
IDR	Incremental Design Review

ILM	Integrated Logistics Management
INCI	Internetworking CSCI
INGST	Ingest CSCI
I/O	Input/Output
IOS	Interoperability Subsystem
IO4	SGI I/O subsystem backplane card
IP	International Partner
IP	Internet Protocol
Ir1	Interim Release 1
IRD	Interface Requirements Document
IsoLAN	Isolation LAN
ISS	Internetworking Subsystem
JERS-1	Japanese Earth Resources Satellite
JPL	Jet Propulsion Laboratories
KB	KiloByte
LAN	Local Area Network
LANDSAT	Land Remote Sensing Satellite
LaRC	Langley Research Center
LCSE	Laboratory for Computer Science and Engineering
LE	Link Encapsulation
LIM	Local Information Manager
LIMGR	Local Information Management CSCI
LIS	Lightning Imaging Sensor
LSM	Local System Manager
MACI	Management Agent CSCI
MB	Megabyte
MB/s	Megabytes per Second
MCI	Management CSCI
MD	Maryland

MDT	Mean Down Time
MF	MegaFlops (Millions of Floating Point Operations per Second)
MFPI	Millions of Floating Point Instructions
Mhz	MegaHerz
MIB	Management Information Base
MIC	Media Interface Connector
MicroMMAC	Micro media access control module
MHCI	Management Hardware CSCI
MLCI	Management Logistics CSCI
MISR	Multi-Imaging Spectro-Radiometer
mm	Millimeter
M&O	Maintenance & Operations
MOC	Mission Operations Center
MODIS	Moderate Resolution Imaging Spectrometer
MOPITT	Measurement of Pollution In The Troposhpere
MPP	Massively Parallel Processor
MPS	Millions of Instructions Per Second
MSS	Management Subsystem (CSMS)
MTBF	Mean Time Between Failure
NASA	National Aeronautics and Space Administration
NASDA	National Space Development Agency of Japan
NCD	Network Computing Devices
NCR	Non-Conformance Report
NFS	Network File System
NOAA	National Oceanic and Atmospheric Administration
NSI	NASA Science Internet
NSIDC	National Snow and Ice Data Center (DAAC)
NTP	Network Transfer Protocol
ODC	Other Data Center

OODCE	Object-Oriented DCE
ORNL	Oak Ridge National Laboratory (DAAC)
OS	Operating System
OSF	Open Software Foundation
OTS	Off-The-Shelf
PDPS	Plannning and Data Processing Segment
PDR	Preliminary Design Review
PDS	Product Data Set
PGE	Product Generation Executive
PGS	Product Generation System (archaic; replaced by PDPS and/or SDP)
PLANG	Planning CSCI
PLNHW	Planning HWCI
PLS	Planning Subsystem
PM-1	EOS PM Mission Spacecraft 1, afternoon series
ppm	Pages Per Minute
PRONG	Processing CSCI
nQyy	Quarter n, year yy
QA	Quality Assurance
RAID	Redundant Array of Inexpensive Disks
RAM	Random Access Memory
RDBMS	Relational Data Base Management System
RID	Review Item Discrepancy
RMA	Reliability, Mantainability, Availability
RSA/CNES	Russian Space Agency/Centre National d'Etudes Spatiales
SAGE	Stratospheric Aerosol and Gas Experiment
SAR	Synthetic Aperture Radar
SCDO	Science and Communications Development Office
SCF	Science Computing Facility
SCSI	Small Computer System Interface

SDP	Science Data Processing
SDPS	Science Data Processing Segment
SDPTK	SDP Toolkit CSCI
SDR	System Design Review
SDSRV	Science Data Server CSCI
SGI	Silicon Graphics, Incorporated
SMC	System Montoring and Coordination Center
SMP	Symmetric Multi-Processor
SNMP	Simple Network Management Protocol
SPARC	Single Processor Architecture
SPECfp	(floating point)
SPECint	(integer)
SPRHW	Science Processing HWCI
SSAP	Science Software Algorithm Package
SSI&T	Science Software Integration and Test
SSM/I	Special Sensor for Microwave/Imager
STK	StorageTek
STMGT	Storage Management CSCI
SURAnet	Southeast Universities Research Association network
S/W	Software
TB	TeraByte
TBD	To Be Determined
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TES	Tropospheric Emission Spectrometer
TPM	Transactions Per Minute
TPS	Transactions Per Second
TRMM	Tropical Rainfall Measuring Mission
UR	Universal Reference

URL	Universal Reference Location (or Locator)
VOB	Version Object Base
V0	Version 0
V1	Version 1
WAIS	Wide-Area Information Server
WKBCH	Workbench CI
WKSHW	Working Storage HWCI
WRP	Wrapper
WS	Working Storage
WWW	World Wide Web
XFS	x File System
nX	“n” times (1X, 2X, 3.4X, etc.)
2-D	Two-Dimensional
3-D	Three-Dimensional