

Data Management Overview

Lynne Case

lcas@eos.hitc.com

18 April 1996

Agenda



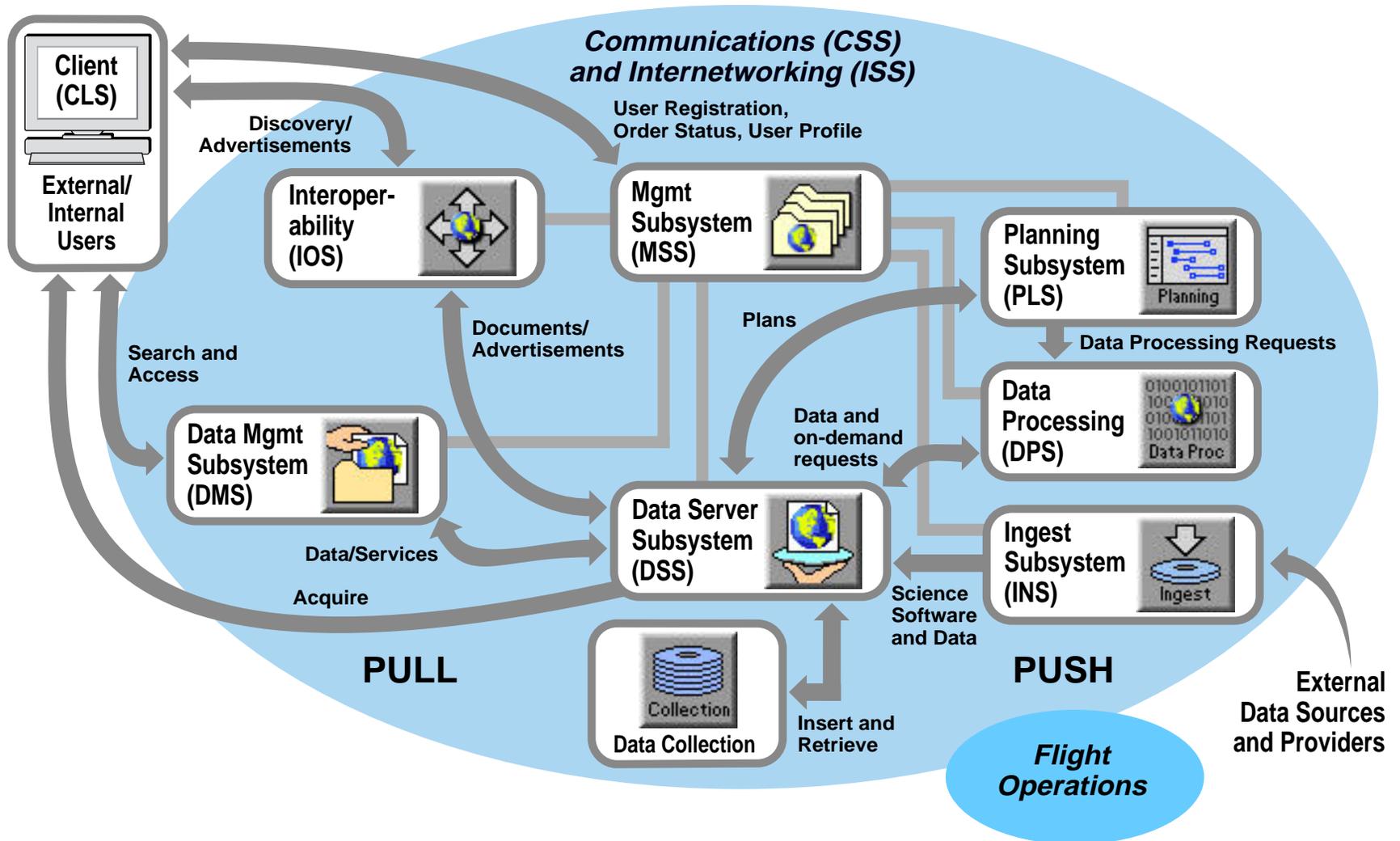
Overview - Lynne Case	1:00 - 1:45
Data Dictionary Design - Hassan Rifky	1:45 - 2:10
LIM/DIM/GTWAY Configuration - Hassan Rifky	2:10 - 2:35
BREAK	2:35 - 2:50
DIM/LIM/GTWAY Object Models and Scenarios	
• CSC Overview - Lynne Case	2:50 - 3:00
• Client-Server Interface - Alain Petit	3:00 - 3:30
• Request Processing and Mapping - Geetha Srikantan	3:30 - 4:10
• External Interfaces - George Chang	4:10 - 4:40
Two-way Interoperability Model - Lynne Case	4:40 - 5:00

Subsystem Overview



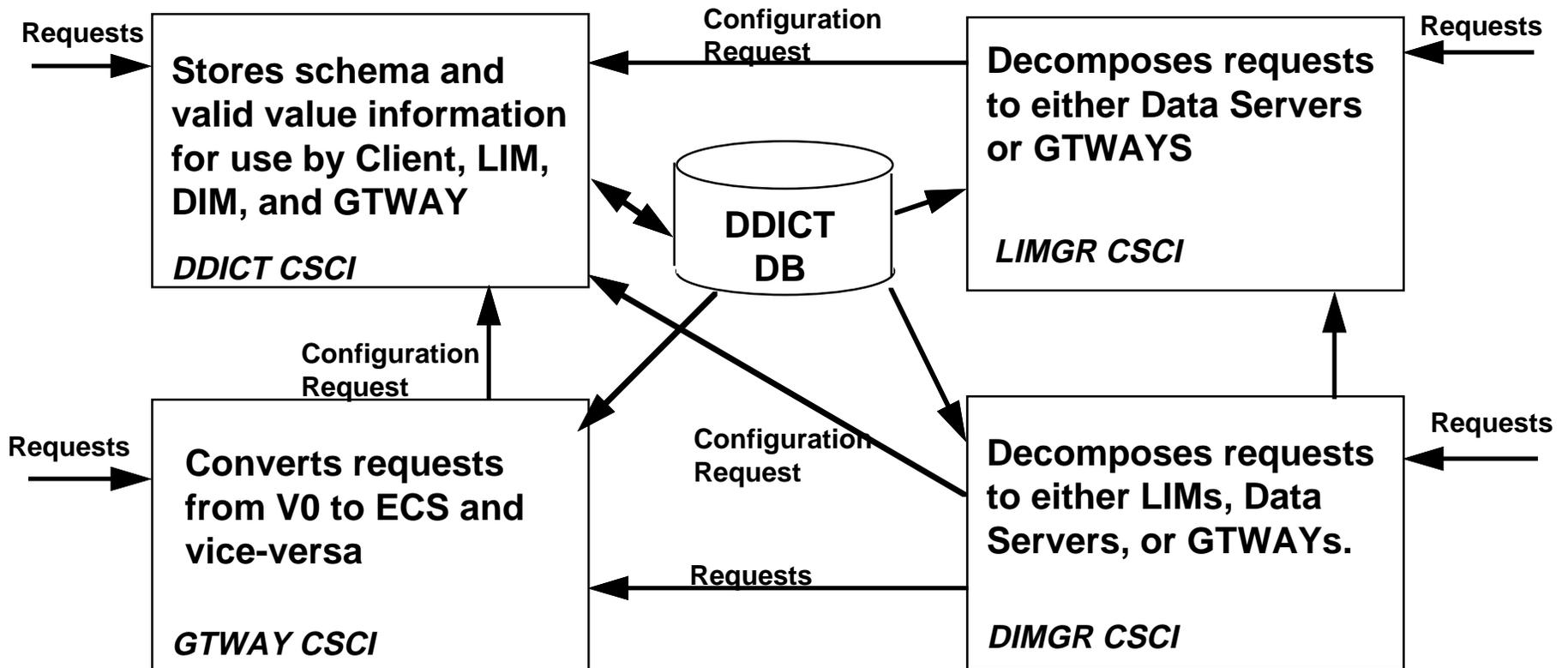
- **ECS context**
- **Driving Requirements**
- **HW and SW Architecture**
- **COTS selections**
- **IDR & workshop issues workoff**
- **Prototypes and Trades**
- **Design changes from IDR**
- **305 Errata**

ECS Context





Software Architecture Overview



Data Management Subsystem



Driving Requirements - DDICT

Data Dictionary (DDICT) CSCI - 305-CD-023-002 Section 4

- **Provides interface to repository containing the following information**
 - **Collection-level metadata describing data collections and their location in the system.**
 - **Attribute descriptions, data types, and domains.**
 - **Dependencies between attribute domains (for example, instrument to platform)**
 - **Phenomenology and climatology data and how (or if) the terms relate to collections in the system.**
 - **Aliasing of terms**
- **DDICT is all new in Release B**
 - **This replaces and significantly enhances the local client database used in the Release A client for valid value information.**

Driving Requirements - LIM & DIM



Local Information Manager (LIMGR) CSCI - 305-CD-023-002 Section 6

- Provides access to data and services at other components (Science Data Server, Document Data Server, V0 Gateway) within a site.
- Provides for site enhancements and autonomy of physical schema representation.
- The LIMGR is all new in Release B

Distributed Information Manager (DIMGR) CSCI - 305-CD-023-002 Section 5

- Provides access to data and services at multiple sites.
- Accesses the components for Science Data Server, Document Data Server, LIMGR, and V0 Gateway depending on the configuration of each DAAC.
- Operates only on a common set of schema attributes known as the ECS metadata standard (311-CD-004-002). For site specific attributes, searches must go to the lower level components.
- DIMGR is all new in Release B

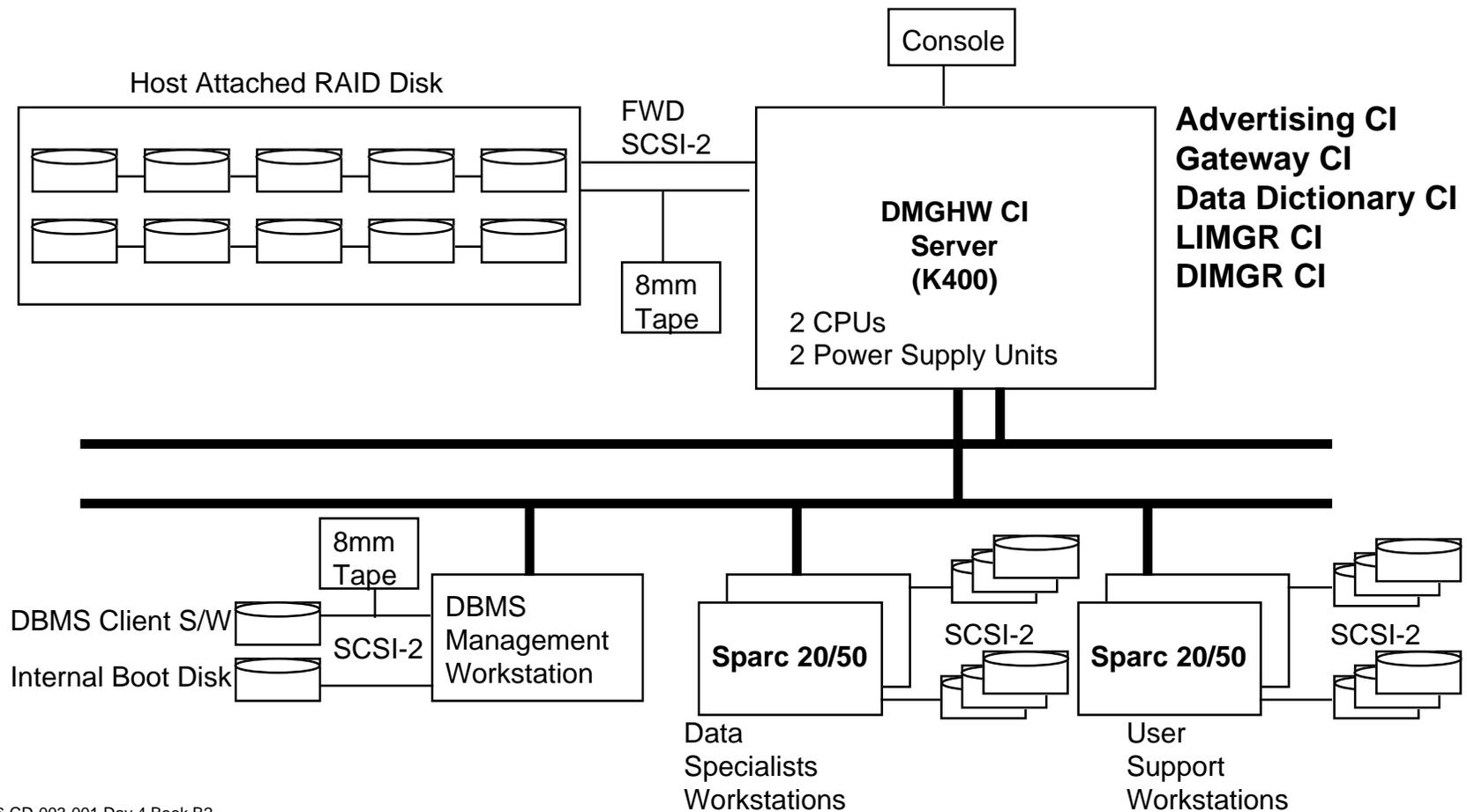


Driving Requirements - GTWAY

V0 Interoperability Gateway (GTWAY) CSCI - 305-CD-023-002 Section 7

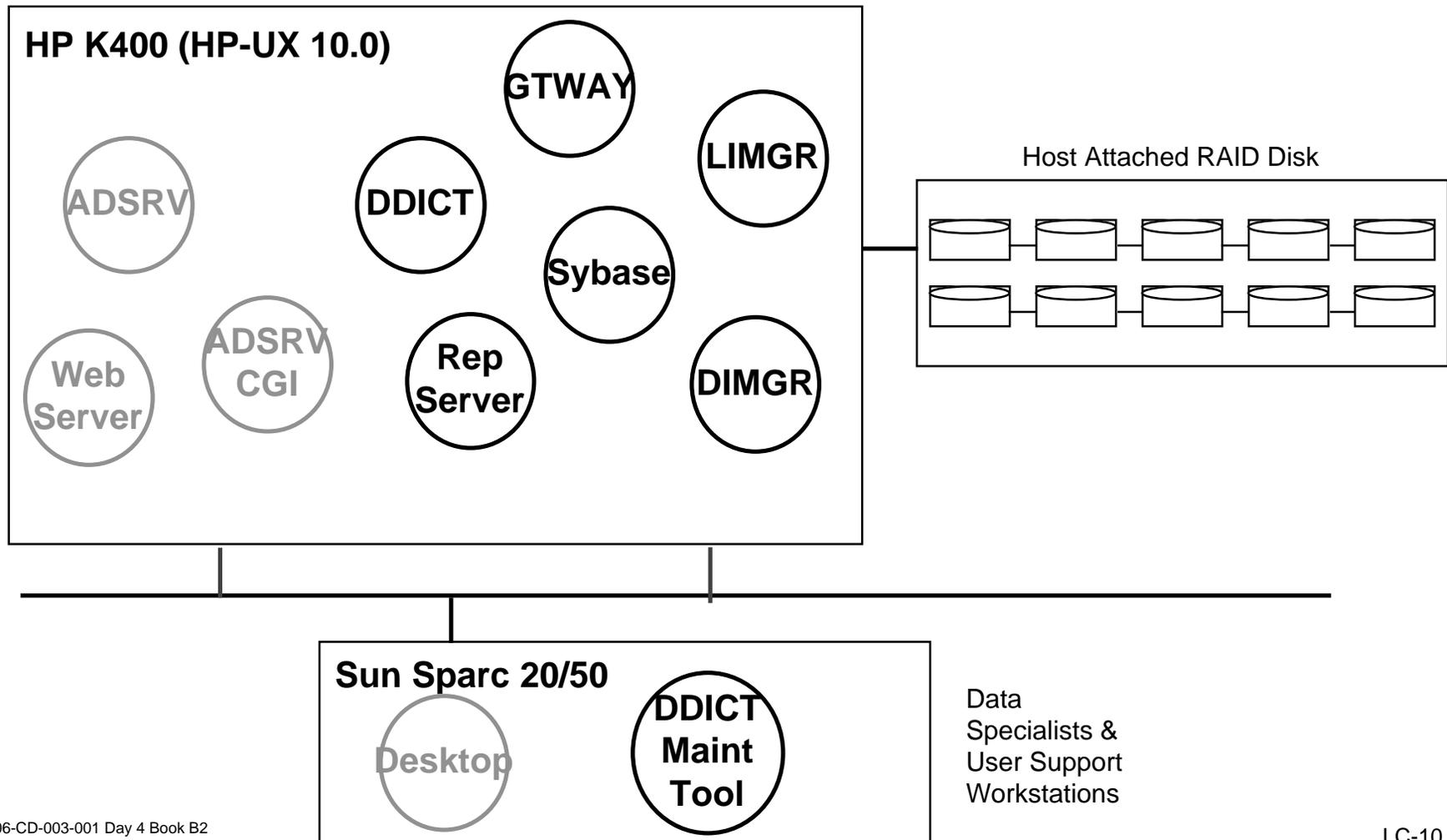
- Provides access to V0 data and services from the ECS client and to ECS data and services to the V0 client.
- Only provides the subset of services and data that the V0 protocol allows, but can easily be extended for others.
- Maps between ECS terminology and V0 terminology.
- New Features in Release B
 - Provides the ECS to V0 direction for interoperability
 - Modifies the V0 to ECS direction to provide the DIM-view for some international partners. (See this more in the Two-way Interoperability discussion).

Data Management Hardware Architecture





Hardware Architecture Overview



COTS Selections



COTS Used in Release A

- Sybase 10, including Replication Server

COTS Used in Release B

- Sybase 11, Including Replication Server

Status

- Reviewed other COTS solutions, see Prototypes and Trades slide.



IDR Issues and RIDs

Status

- IDR

- RID #11 - Granule Packages - Data Server responsibility. Closed.
- RID #21 - Not enough Design shown - CDR and Final Design Review to fix this. Closed.
- RID #24 - Coincident searches not conforming to all ECS data dictionary requirements. Supported by the client to inform the user of choices made that eliminate data sources based on inadequate data schemas. Prototype in EP7. Data Dictionary provides the client with this information. Closed.
- RID #29 - Reference to and/or access to user supplied data sets. Can be supported with LDAS, which comes from Data Server. External Data Provider options discussed next week. Closed.
- Architecture Concept and Baseline Configuration- presented at Ops Workshop in Jan. and again today.
- Data Dictionary Caching Scheme - timestamps being applied to Data Dictionary Contents such that Client can determine which pieces need to be downloaded.

Ops Workshop Issues



Ops Workshop

- **Purpose of the LIM? - (1) Encapsulate site enhancements, (2) Provide better join performance by doing local joins, and (3) Use by External Data Providers to map non-ECS schemas to the federated schema.**
- **Investigate methods to tune Web Server - Web server belongs to IOS, but lives on DMS server. IOS will investigate the methods.**
- **Data Dictionary Log entry should include user submitting the update. - YES!**



Prototypes and Trades

Status

- **CDR Baseline**
 - **Coincident Search Study**
 - >> Reviewed possible coincident search methods.
 - >> Determined that all possible coincident operations should occur at the Data Server level since they have the knowledge of the spatial structures and methods. LIMs and DIMs will simply coordinate the searches.
 - **COTS DIM/LIM Study**
 - >> Submitted RFI to Data Warehousing and DBMS vendors.
 - >> Responses did not provide enough justification for proceeding with a procurement.
 - >> Most systems still lacking in support for spatial data.
 - >> Also breaks the encapsulation model established for evolutionary enhancements.
 - >> Prototyping and study will continue for Release C, but will not be included in the Release B baseline.



Key Mechanisms

Mechanisms through which processes acquire or provide software services:

- **Process Framework (PF)**
 - **Mechanism for incorporating infrastructures required to support distributed computing in the ECS environment**
 - **Provides process initialization and life cycle support**
 - **Sets parameters for naming/directory/security services**
 - **Interfaces to mode management, event logging, synchronous message passing**



Key Mechanisms (Cont.)

- **Universal Reference (UR)**
 - Mechanism for referencing system wide data and service objects
 - Provides externalize and internalize services for objects
- **Server Request Framework (SRF)**
 - Mechanism for constructing servers and client/server APIs
 - Provides asynchronous request processing service
- **Subscription Server (SS)**
 - Mechanism for implementing event-action model based on a producer/consumer paradigm
 - Provides subscription and event/action processing services
- **Request Tracking**
 - Provides near real-time end-to-end status of selected requests (e.g., user orders, ingest, and system backups)
 - Provides access to resource cost information over the life of a request

Design Changes from IDR



Coincident Search Resolution

- **IDR design had the DIM/LIM performing coincident search operations.**
- **Study showed the cost of this design was too high based on the expected frequency of execution of the service.**
- **The coincident searches will be resolved using the Data Server to perform the coincidence operation.**

305 Errata



Only the Public Interface of the DDICT, LIMGR, DIMGR, and GTWAY is described in the 305.

Some changes have occurred in the client-server interface of these components.

OMT diagrams and documentation updates are available.