

A horizontal banner image featuring a sequence of celestial bodies: a blue planet (Earth), a brown planet (Mars), and a white planet (Jupiter) against a dark space background with stars. The text "Planetary Data System" is overlaid in white on the right side of the banner.

Planetary Data System

System Architecture

PDS 2010 System Review
March 22-24, 2010

Sean Hardman

Topics

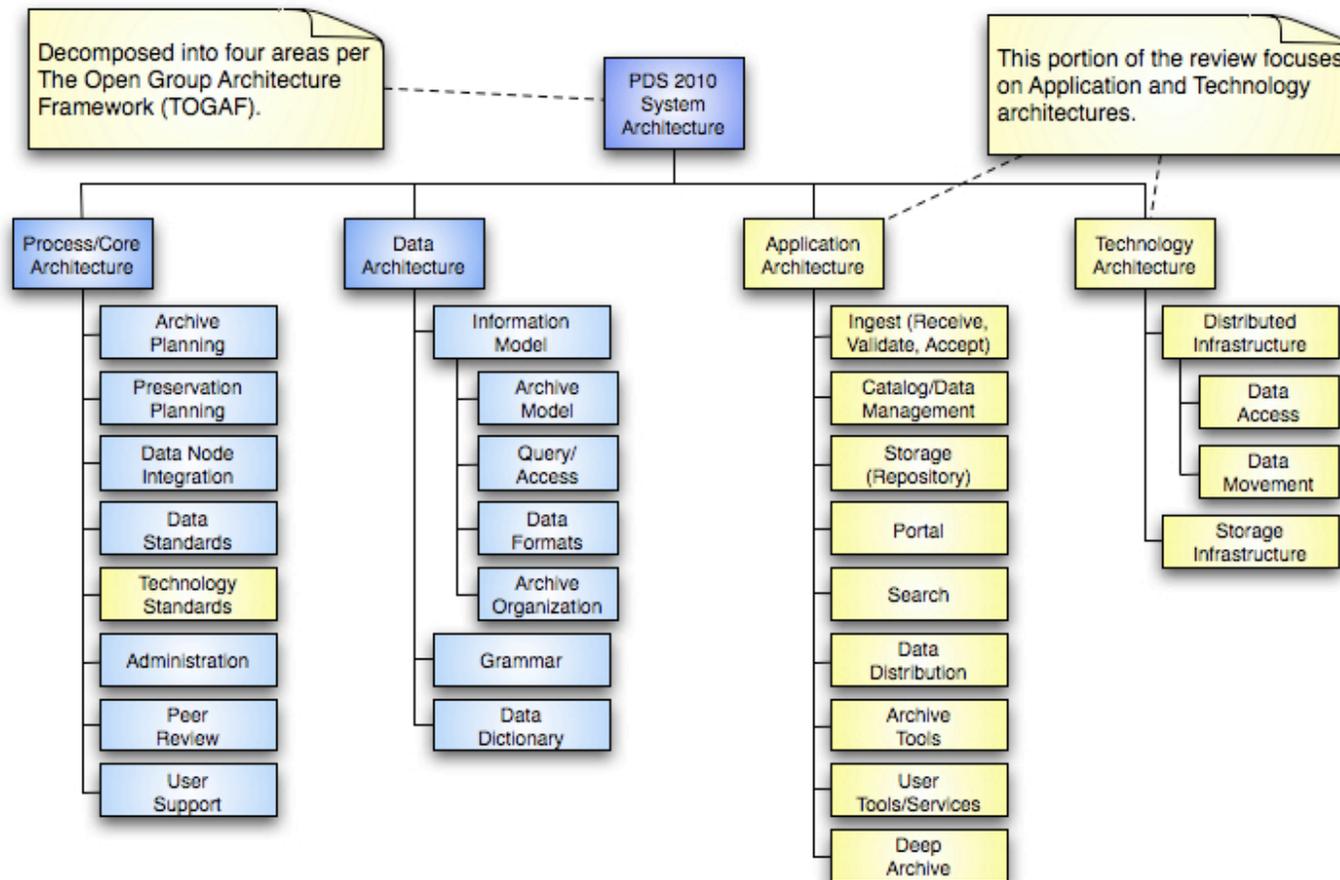
- Overview and Context
- Registry Components
- Ingestion Components and Flows
- Distribution Components and Flows
- Operations Components and Flows
- Wrap Up

Overview

- The System Architecture for PDS, encompasses all PDS-2010 and future projects.
 - This includes projects developed at the Engineering Node as well as the Discipline Nodes.
- The System Architecture facilitates the development of PDS-2010 by providing:
 - Consistent use of common terminology for ease of integration.
 - Commonly defined and implemented interfaces to increase usability and portability of applications.
 - Well defined and loosely coupled services to increase scalability and adaptability for future expansion.

Architectural Elements

(Elements Derived from Level 3 Requirements)



Architecture Decisions

- Service-Based Design
 - Support remote access to data and services to bring the federation together both for ingestion and distribution.
- System of Registries
 - Adopt a system of registries to support improved tracking and access.
- Enhanced Tool Suite
 - A tool-based approach is still appropriate for certain functions.

Service-Based Design

- There are several advantages to adopting a Service-Oriented Architecture (SOA):
 - Captures many of the best practices of previous architectures.
 - Well suited for a distributed system.
 - A service-based architecture provides currency and timeliness for the system.
- Currently developing a SOA solution that suits PDS.
- Service-based functionality will focus on public interfaces for search, retrieval and value-added processing (science services) of data.

System of Registries

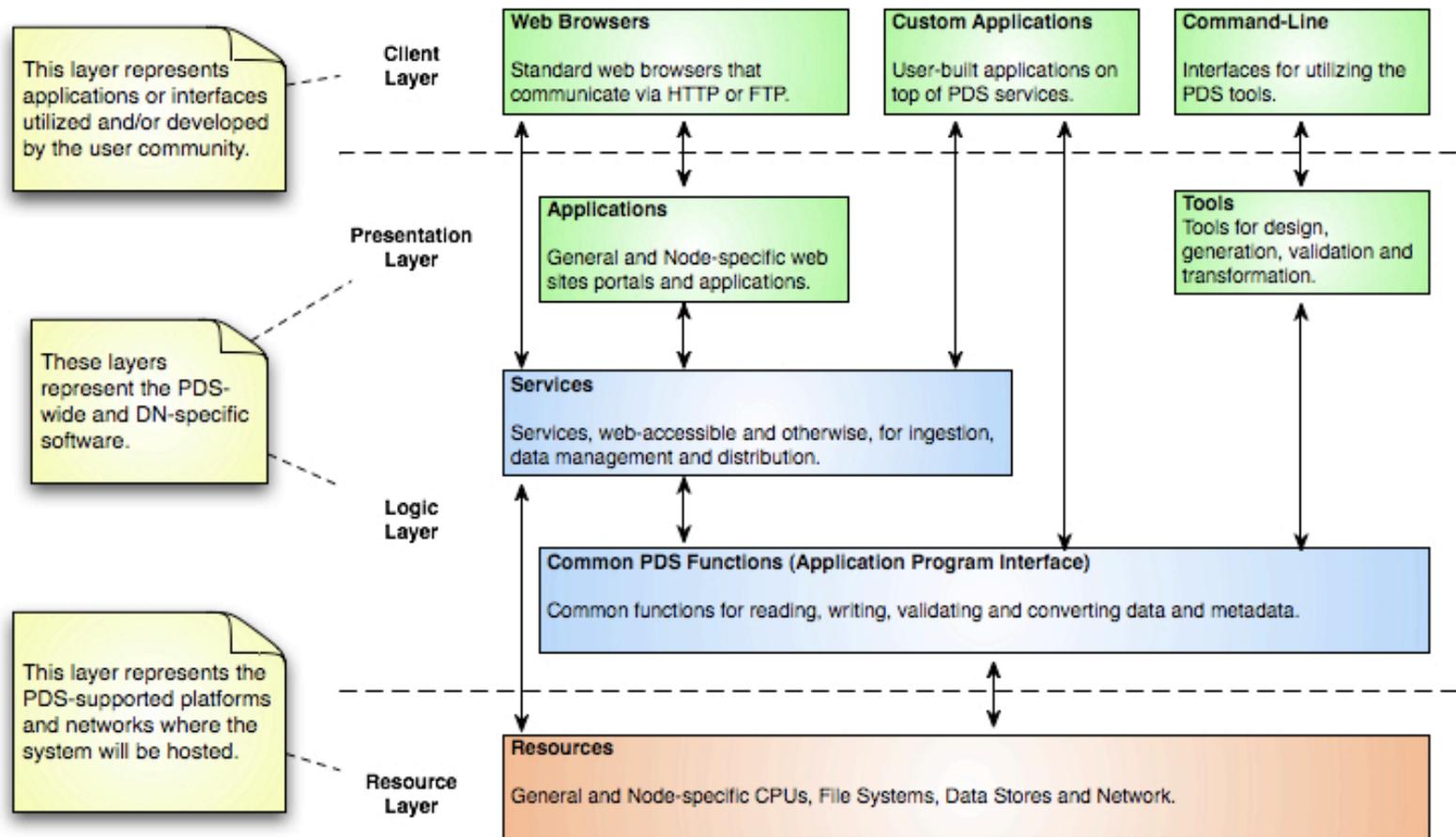
- A registry provides services for sharing content and metadata.
- A federated registry allows cooperating registries to appear and act as a single virtual registry.
 - Provides seamless information integration and sharing
 - Preserves local governance
- A query into the federation returns results from all cooperating registries.

Enhanced Tool Suite

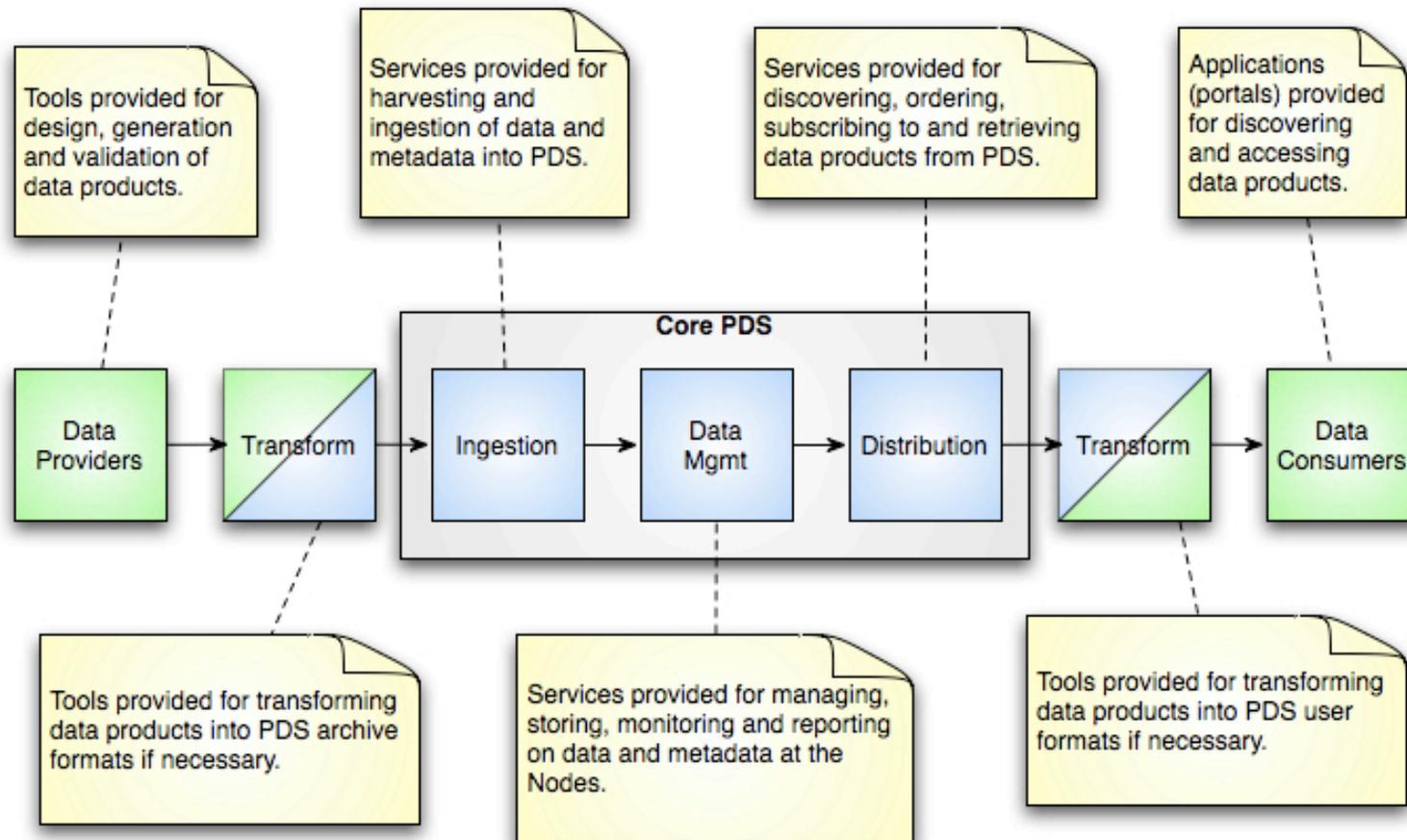
- There are several functions related to product labels and data where a tool is an appropriate interface:
 - **Design** of product label schemas
 - **Validation** of products and collections of products
 - **Generation** of product labels in a pipeline
 - **Transformation** of product formats (labels and data)
 - **Visualization** of image data
- These capabilities will be provided in the form of software libraries with command-line interfaces.
 - Enables incorporation into system services as well as user-developed software.

System Layering

(Layering improves reuse, maintainability and scalability.)

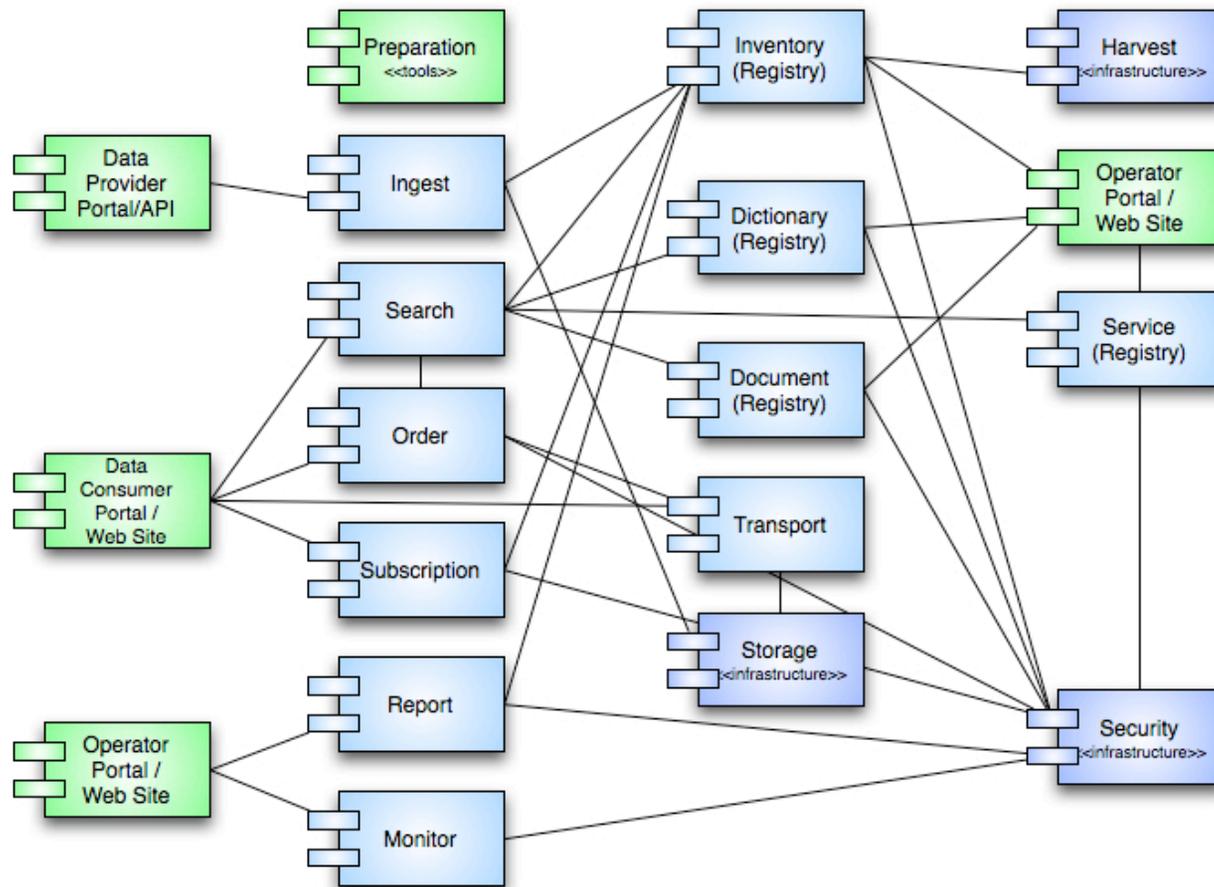


Information Flow



System Context

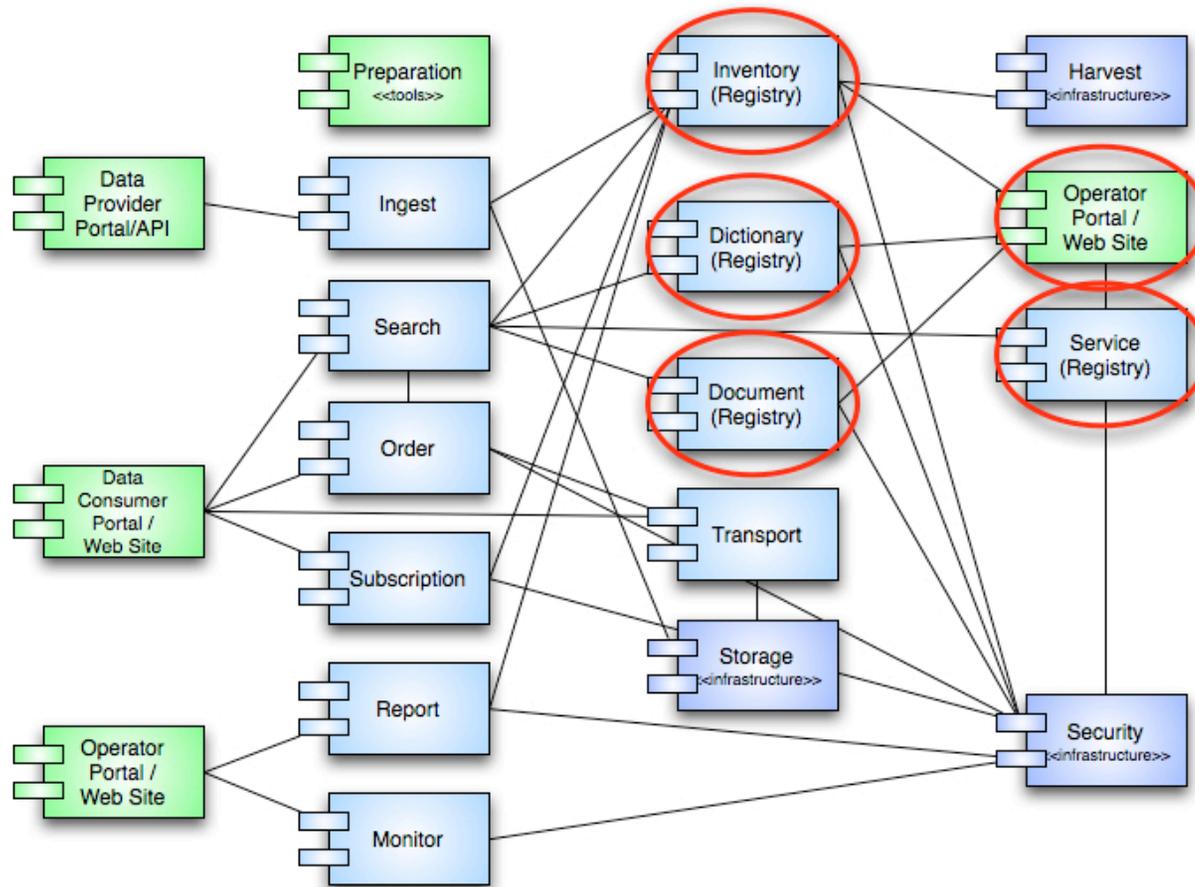
(Components Derived from Architectural Elements)



Topics

- Overview and Context
- Registry Components
- Ingestion Components and Flows
- Distribution Components and Flows
- Operations Components and Flows
- Wrap Up

Registry Related Components



Registry Service

- Provides functionality for tracking, auditing, locating, and maintaining artifacts within the system.
 - Artifacts range from products consisting of data files and label files, schemas, dictionary definitions for objects and elements, service definitions, etc.
- Provides a common implementation for registry service instances.
- Design based on CCSDS effort which builds on the ebXML standard.

Registry Service Instances

- Inventory
 - Captures catalog and product metadata in a number of distributed instances deployed at the Engineering and Discipline Nodes.
 - Tracks catalog and product artifacts from mission delivery to deep archive.
- Dictionary
 - Captures the data dictionary, which consists of object/element definitions and their associations.
- Document
 - Captures project documents, product label schemas, etc.
- Service
 - Captures descriptions of PDS services and their associations with data collections.

Operator Portal

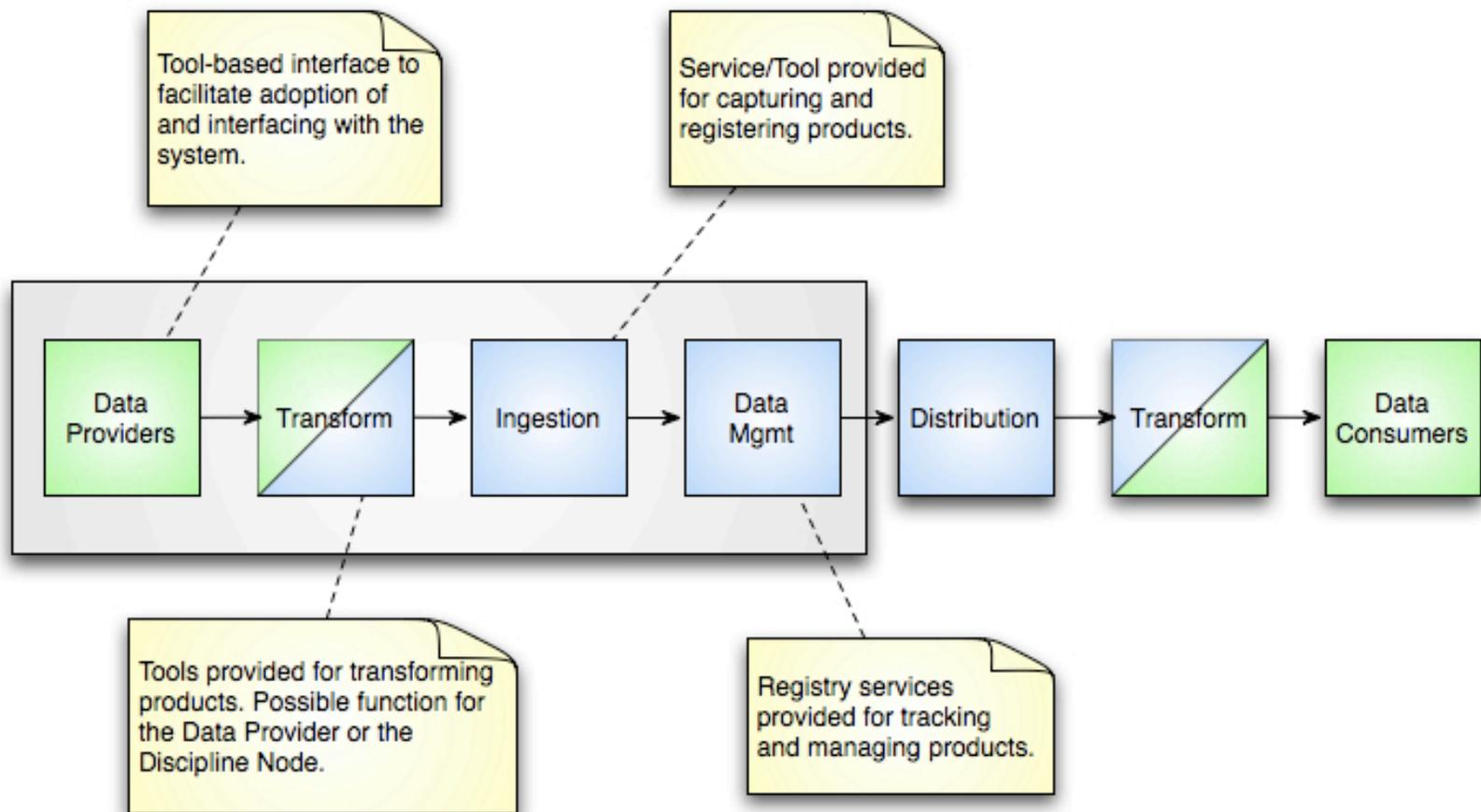
- A general web-based interface for managing registry policy, content and end-to-end tracking.
- The interface will utilize the Registry service API for accessing any of the registry instances.
- The interface is deployable for local instances of the Registry service at the Nodes.
- Access to the interface will have access control requiring authentication and authorization via the Security service.

Topics

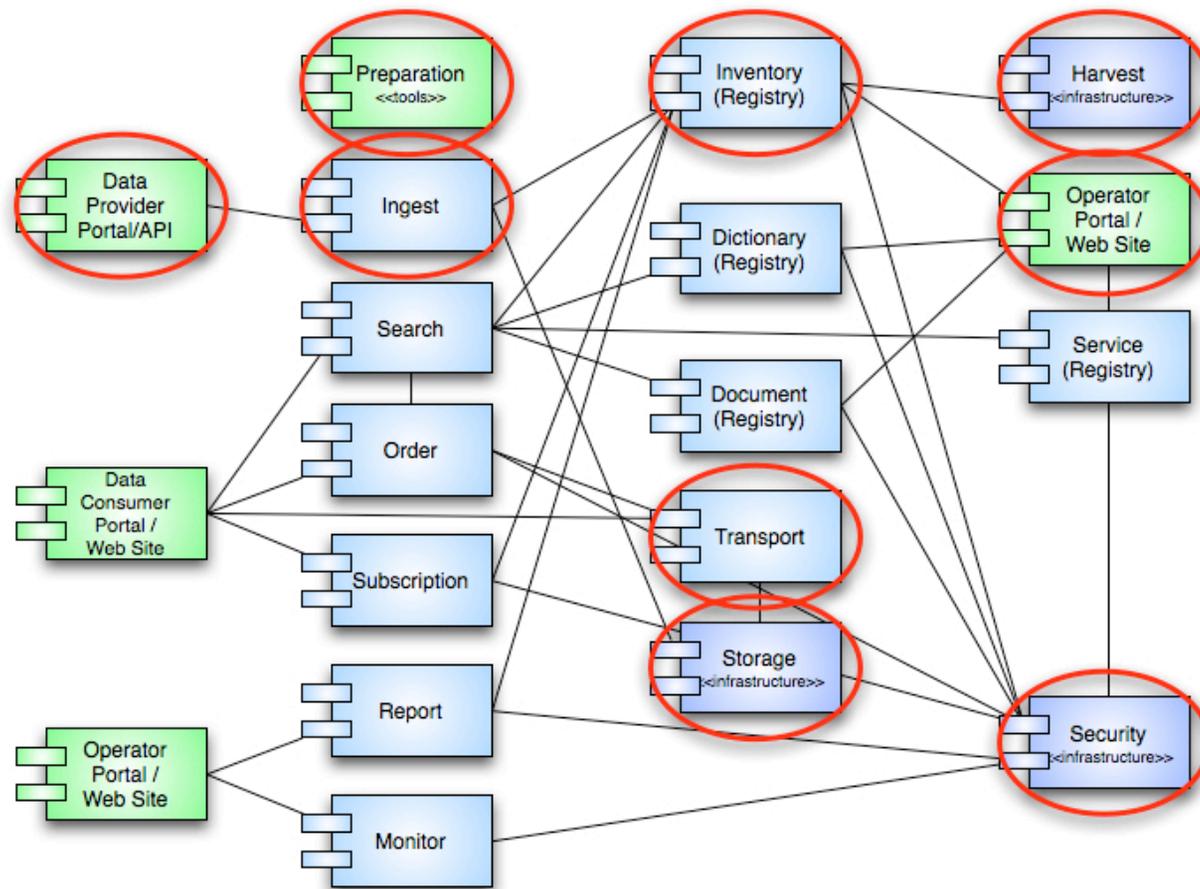
- Overview and Context
- Registry Components
- Ingestion Components and Flows
- Distribution Components and Flows
- Operations Components and Flows
- Wrap Up

Ingestion

(Capture and Registration of Products into the System)



Ingestion Related Components



Phase I Components

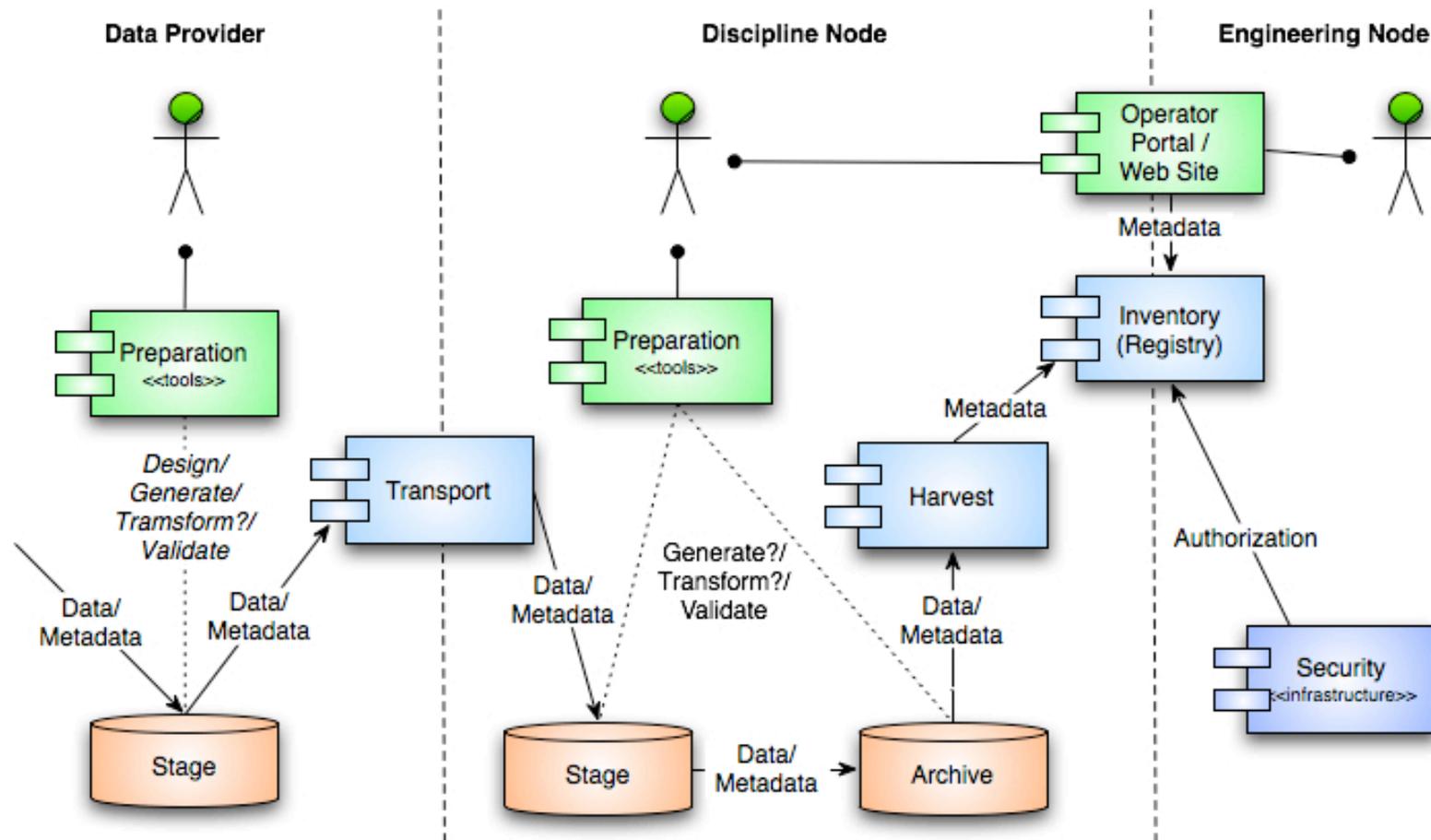
- Preparation Tools
 - Suite of tools for preparing data for ingestion into PDS focusing on design, generation and validation.
 - Allows for existing Node processes and procedures to be utilized for ingestion of data products.
 - Minimizes up-front interface changes for Data Providers.
- Transport Service
 - Represents continued support for FTP and Data Brick delivery mechanisms.
- Harvest Tool
 - Crawler-based tool for capturing and registering product metadata.
 - Allows for periodic or on-demand registration of products.
 - Configurable to support registration of products residing in PDS3 and PDS4 archives.

Phase III Components

- Ingest Service
 - Provides functionality for receiving data and metadata from Data Providers for ingestion into PDS.
 - Leveraged from capabilities developed for the Preparation and Harvest tools.
- Storage Service
 - Provides management of a Node's data repository.
- Data Provider Portal/API
 - An interface to the Ingest service enabling integration with data production pipelines for delivery of products in real time.
 - Also accommodates Principal Investigator (PI) (small volume) submissions.

Ingestion Flow

(Focus on Phase I Capabilities)



Ingestion Flow Details

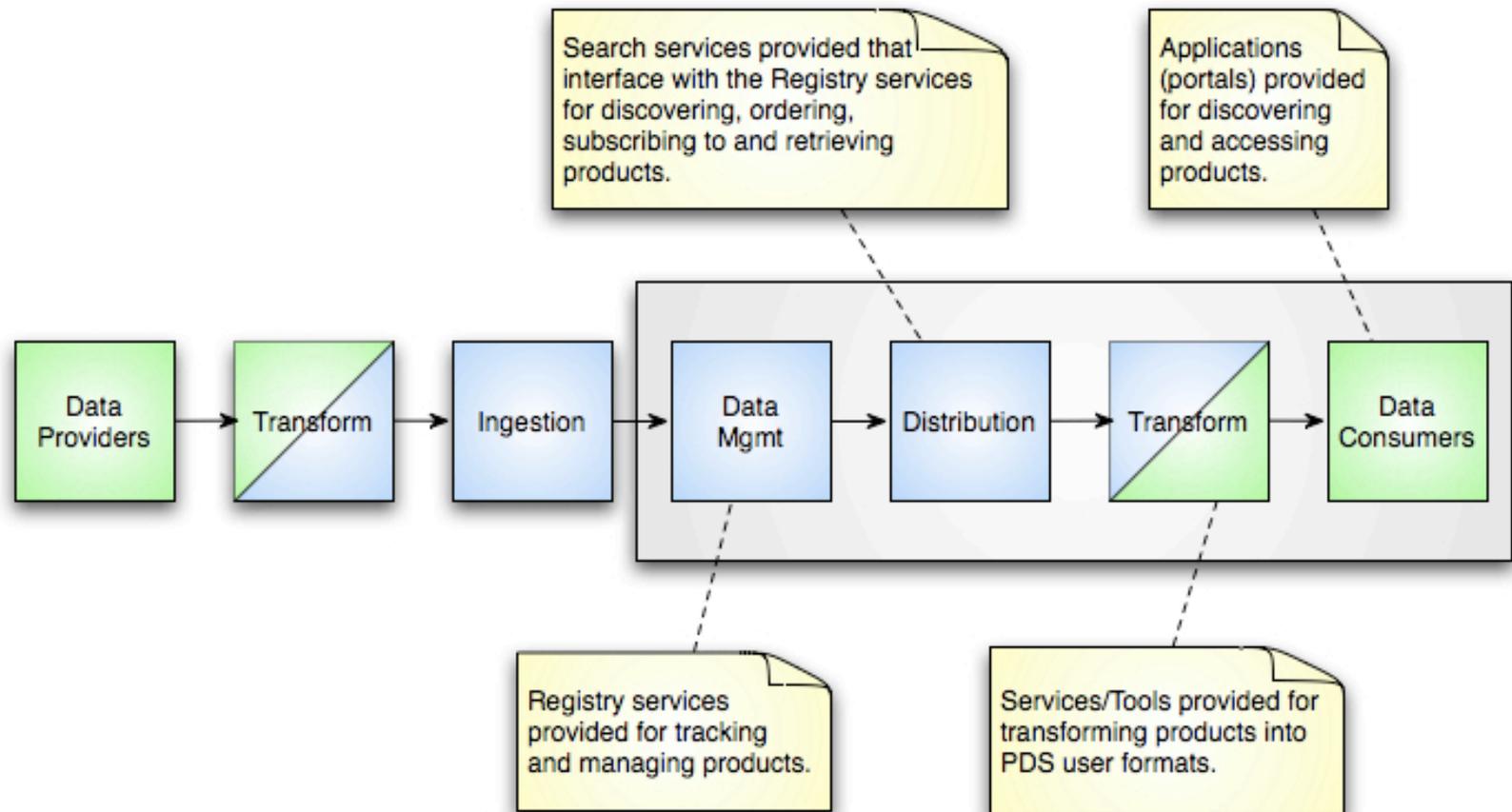
1. Data Provider receives data from the source (e.g., Project, Instrument Team, etc.).
2. Data Provider utilizes PDS provided tools to prepare the data for submission.
3. Data Provider submits transformed/labeled data to the Discipline Node via an agreed interface (e.g., FTP, Data Brick, etc.).
4. Discipline Node receives data/metadata from the Data Provider and stages it in local storage.
5. Discipline Node utilizes PDS tools or tools based on a common library to prepare the data for archive.
6. Discipline Node initiates harvesting of the archive, which registers product metadata in the Inventory service. Metadata registrations are authorized by the Security service.
7. Discipline Node manages housekeeping information and/or augments metadata for search enhancement via the Operator Portal.

Topics

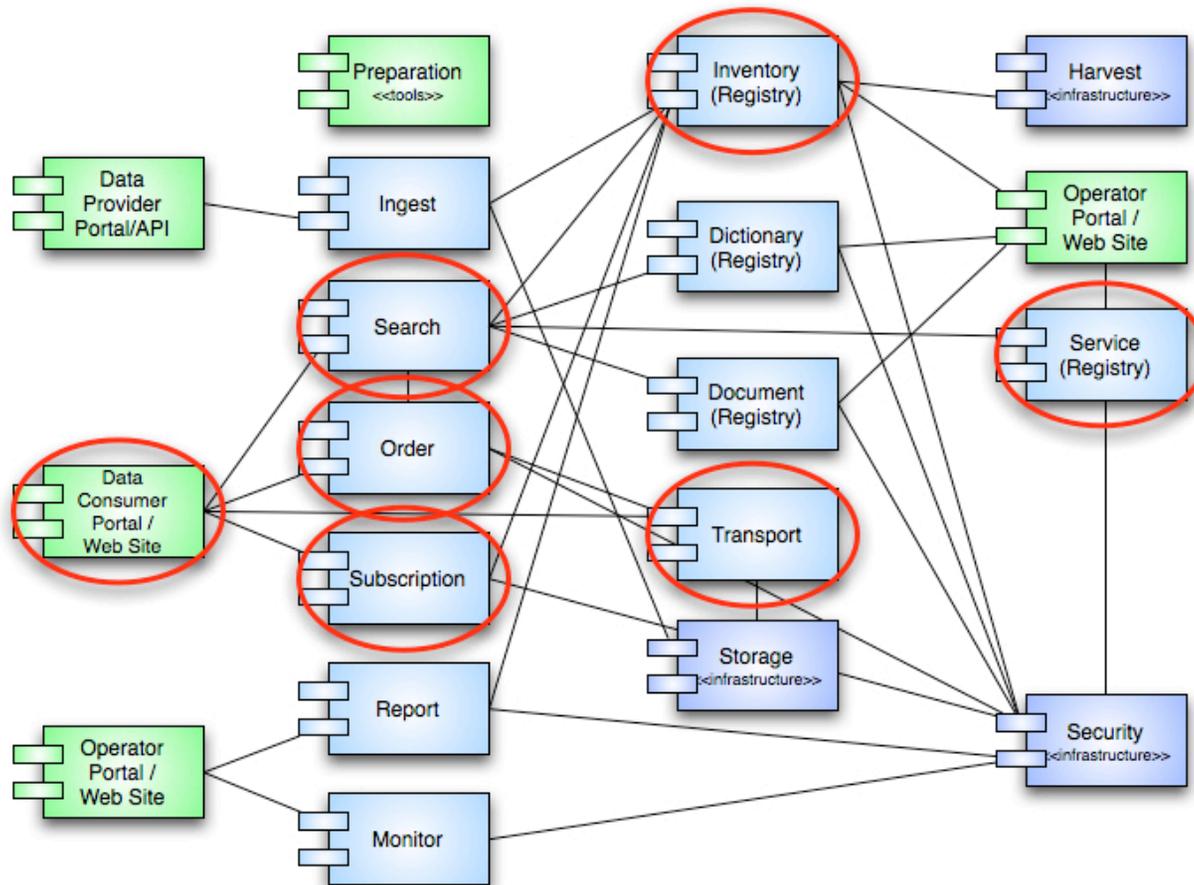
- Overview and Context
- Registry Components
- Ingestion Components and Flows
- Distribution Components and Flows
- Operations Components and Flows
- Wrap Up

Distribution

(Discovery and Distribution of Products)



Distribution Related Components



Phase II Components

- Search
 - Provides indexed metadata utilizing common facets to facilitate search across the Nodes along with Node-specific facets that could be utilized by Node-specific search applications.
 - Provides a REST-based interface for product discovery.
- Transport
 - Integrate existing delivery mechanisms (e.g., FTP, HTTP, etc.).
- Data Consumer Portal
 - Integrate the PDS-wide portal (<http://pds.nasa.gov/>) with the Search service.

Phase III Components

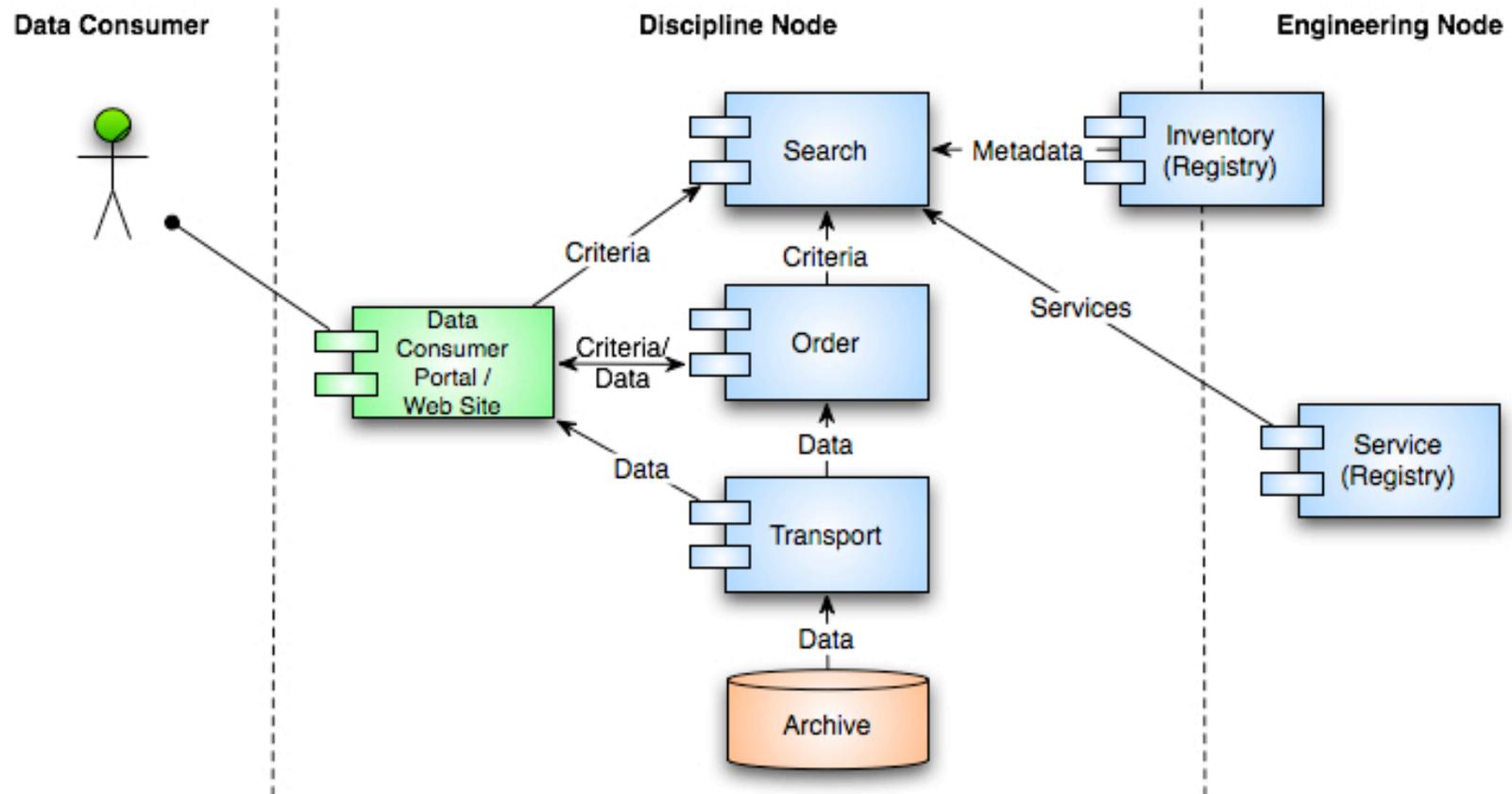
- Preparation Tools
 - Tools for transformation and visualization of products.
- Order
 - A container for PDS-wide and Node-specific science-related services (see next slide).
- Subscription
 - Provides subscription to data, document and software release announcements.
- Transport
 - Incorporate other mechanisms for delivery (e.g., GridFTP) and transformation on-the-fly capabilities.
- Data Consumer Portal
 - Node-specific portals will provide web applications for accessing product-level information.
 - Integration of existing Node-specific applications.

Science-Related Services

- Product-Level Search
 - Discipline-Based, Cross-Mission, Cross-Discipline, Record-Level and Subset
- Data Format Conversion
- Coordinate System Transformation
- Product Visualization
- Calibration on the Fly
- Map Overlays
- All-Purpose Geometry Engine

Distribution Flow

Request Initiated at the Discipline Node



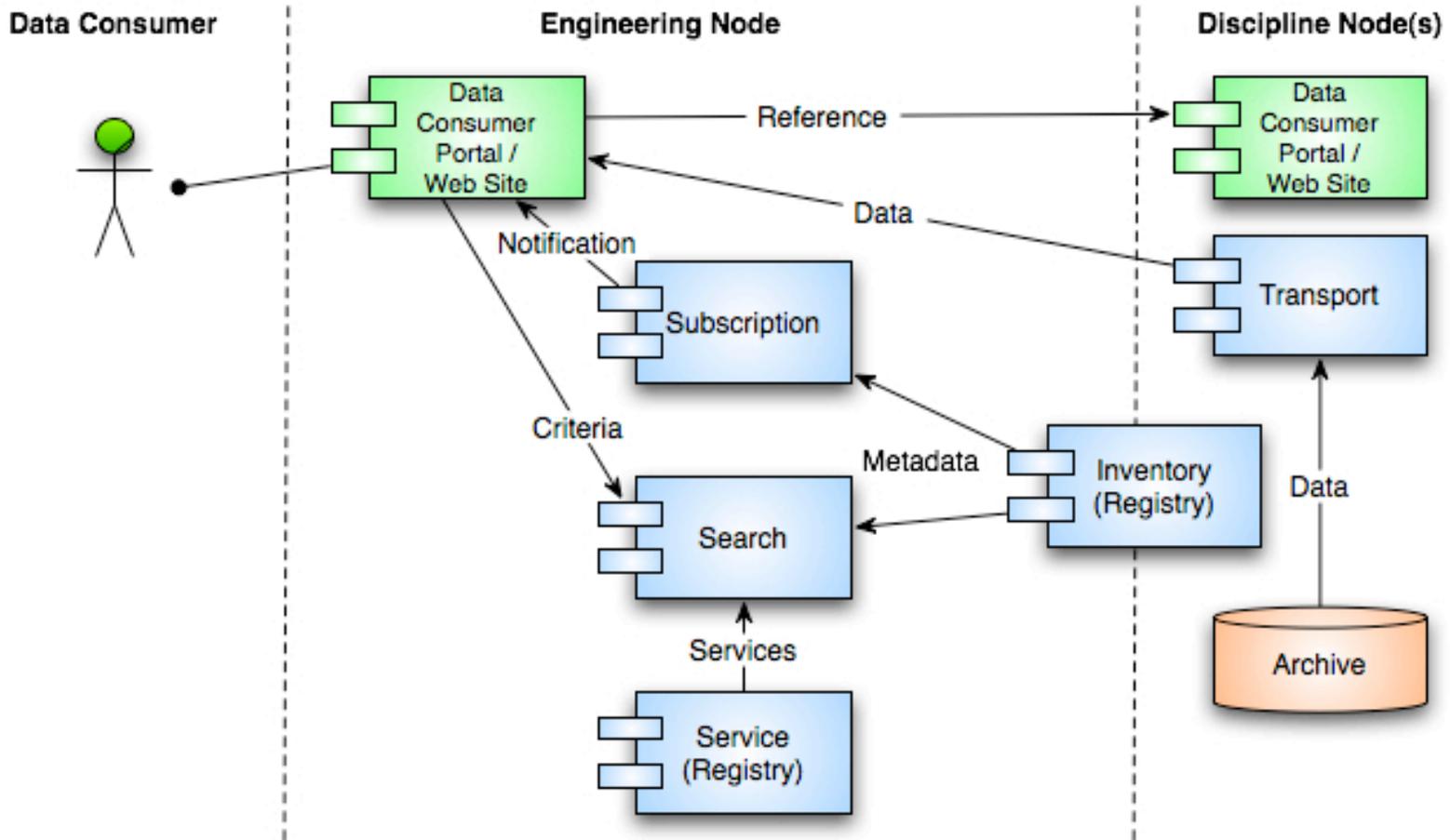
Distribution Flow Details

Request Initiated at the Discipline Node

1. Search service generates a search index utilizing the Service registry to discover the appropriate Inventory service(s) for obtaining the metadata for the index. Tailoring of the search index enables support for the Node-specific search tools.
2. Data Consumer submits a query for data through a portal / web site interface.
3. Portal / web site interface forwards the query to the local Search service.
4. Search service returns results to the portal / web site interface with options for retrieving product(s) that match the query criteria. The Data Consumer may place a special request to the Order service, if one is available at the Node, based on the results returned.
5. Data Consumer makes a request to the Transport service for delivery of the product(s).

Distribution Flow

Request Initiated at the Engineering Node



Distribution Flow Details

Request Initiated at the Engineering Node

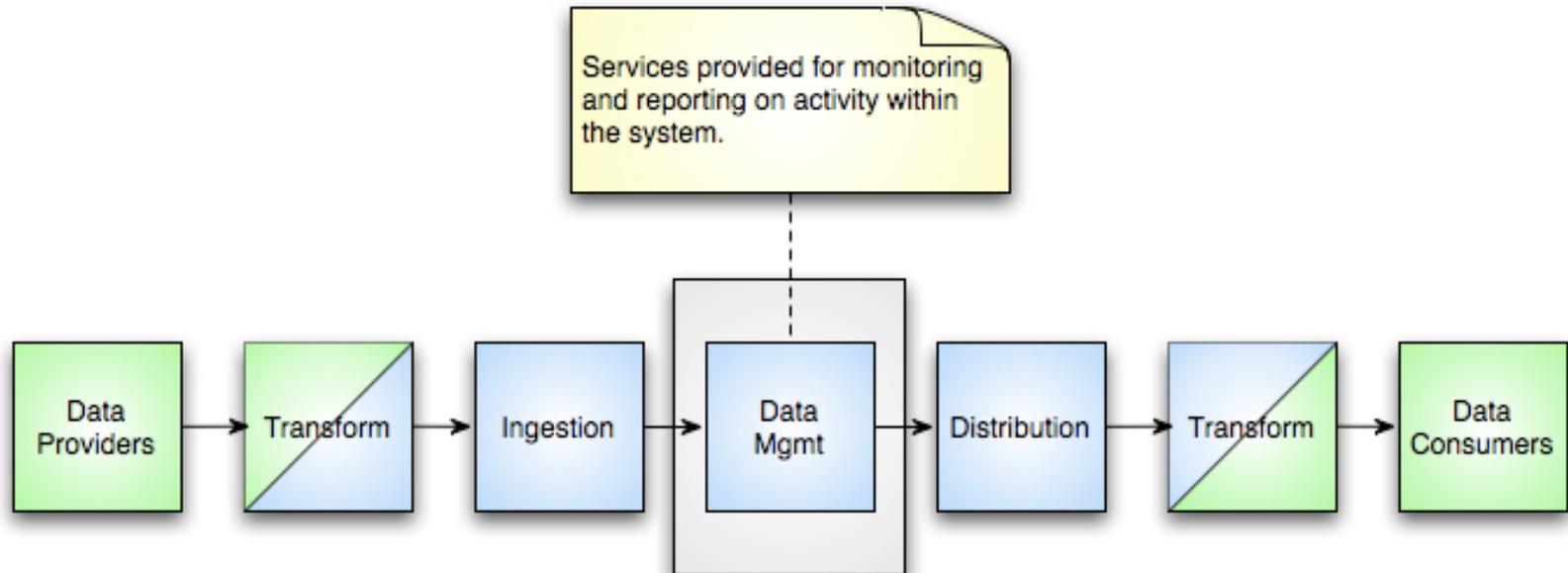
1. Search service generates a search index utilizing the Service registry to discover the appropriate Inventory service(s) for obtaining the metadata for the index.
2. Data Consumer submits a query for data through a portal / web site interface. The Data Consumer may also subscribe to release information via the Subscription service.
3. Portal / web site interface forwards the query to the Search service.
4. Search service returns results to the portal / web site interface with options for retrieving product(s) that match the query criteria.
5. Data Consumer makes a request to the Transport service from the appropriate Node for delivery of the product(s).

Topics

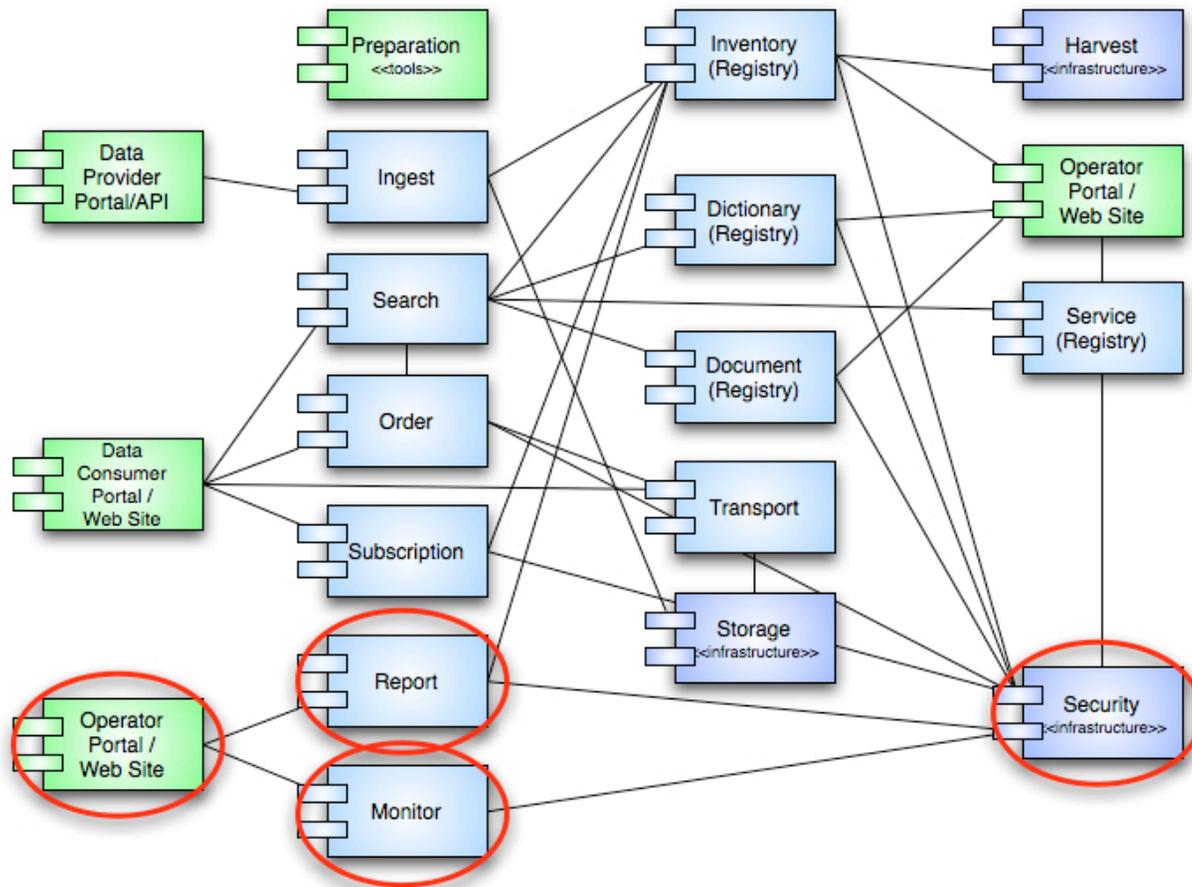
- Overview and Context
- Registry Components
- Ingestion Components and Flows
- Distribution Components and Flows
- Operations Components and Flows
- Wrap Up

Operations

(Monitoring the System and Reporting Metrics)



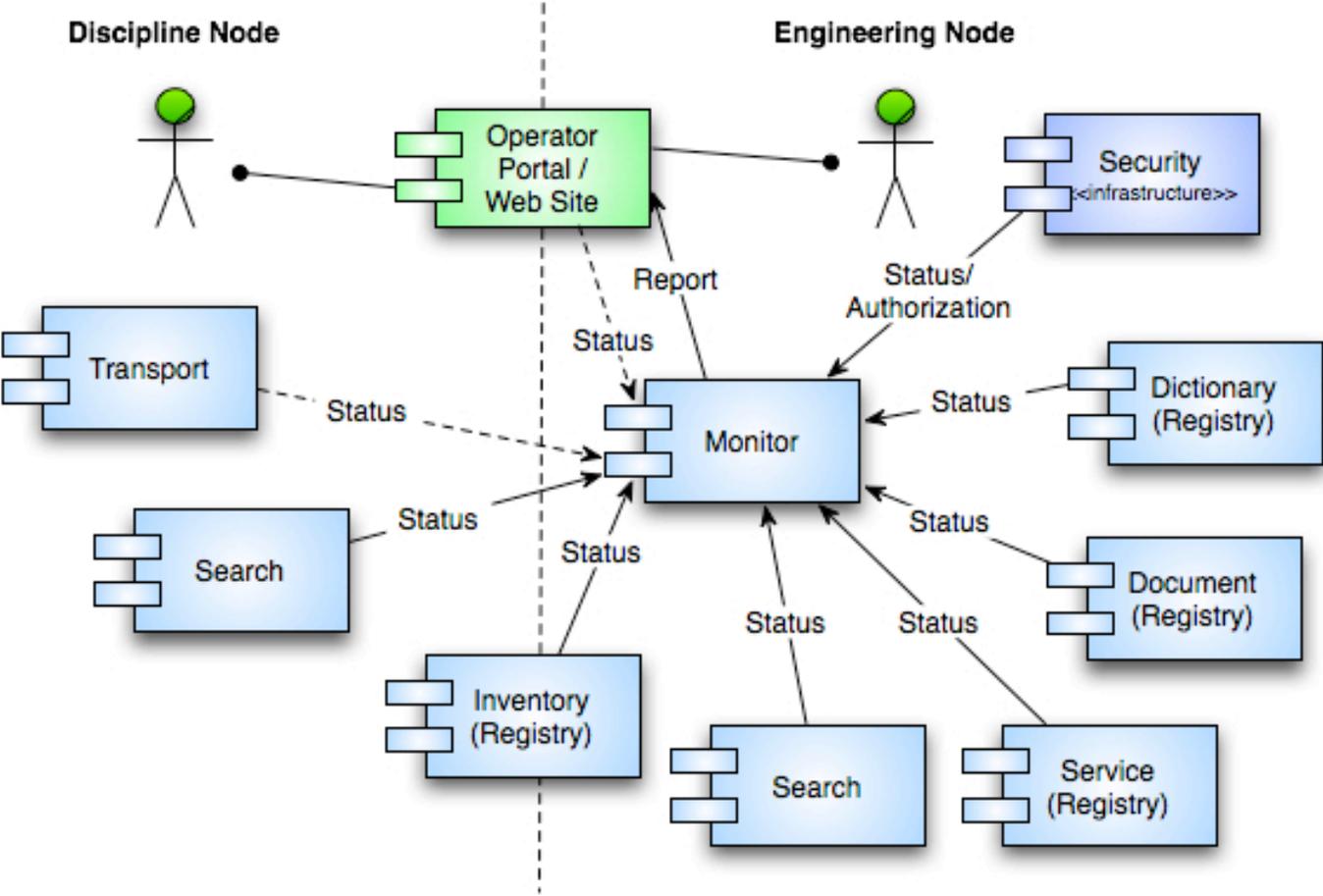
Operations Related Components



Phase I/II Components

- Monitor
 - Monitor service status across the system and facilitate notification of appropriate Node staff if a service were inoperable.
- Report
 - Capture and report metrics collected from across the system including system services and FTP/web logs.
- Security
 - Provides the authentication and authorization functions for the system where necessary (e.g., Monitor, Report, Registry instances).
- Operator Portal
 - Customize/integrate COTS interfaces where feasible.

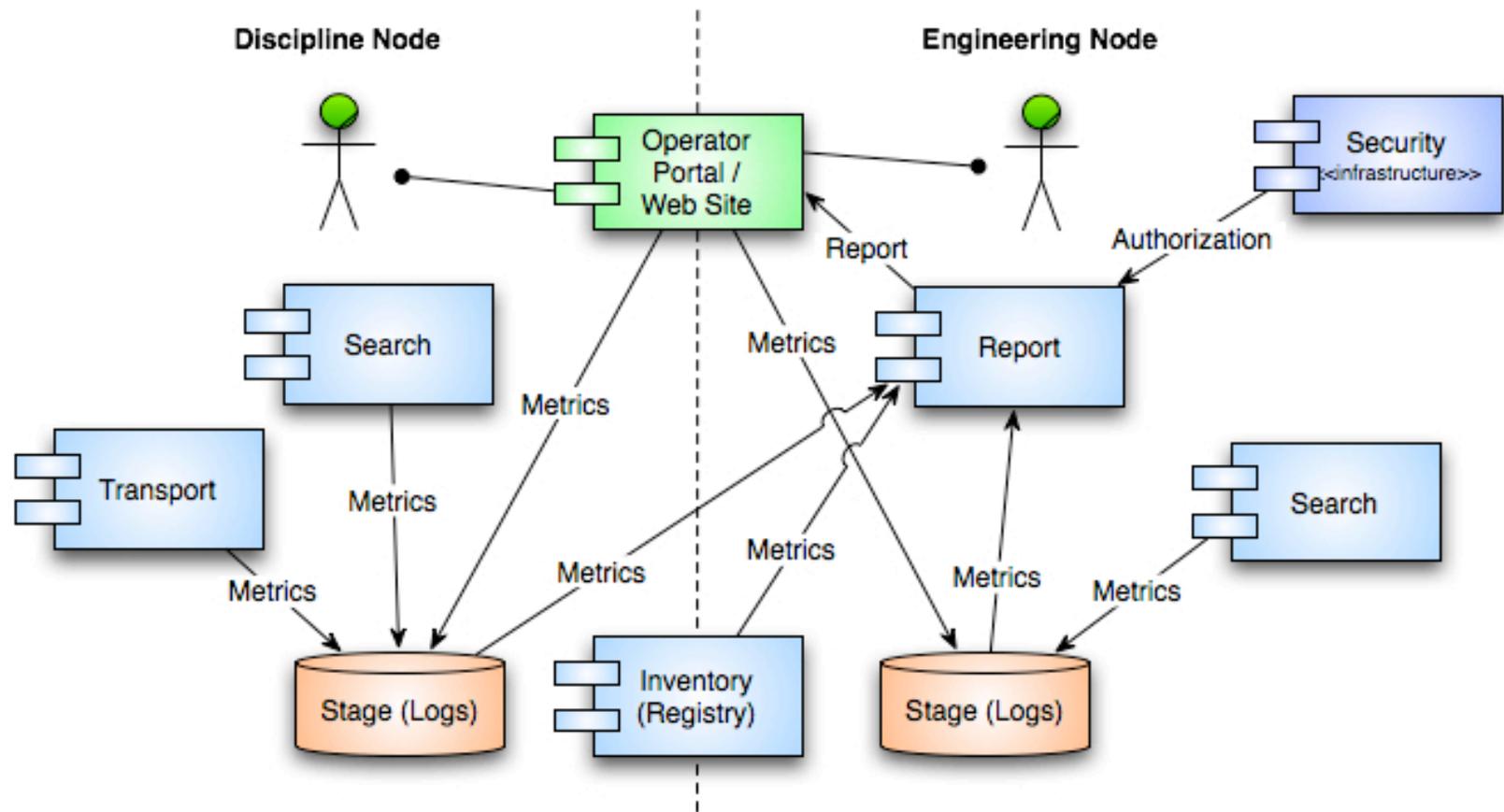
Monitoring Flow



Monitor Flow Details

1. Operator (Discipline Node or Engineering Node) accesses a portal / web site interface, authorized by the Security service, to view the system status.
2. Monitor service receives constant status updates from the system services.

Reporting Flow



Reporting Flow Details

1. Operator (Discipline Node or Engineering Node) accesses a portal / web site interface, authorized by the Security service, to generate a report.
2. The Report service receives periodic metrics submissions from the system services.

Topics

- Overview and Context
- Registry Components
- Ingestion Components and Flows
- Distribution Components and Flows
- Operations Components and Flows
- Wrap Up

Wrap Up

- Presented the scope and decomposition of the System Architecture.
- Introduced the major architectural decisions.
- Identified and defined the components (services, tools and portals) of the system.
- Detailed the provisioning of the components and their interactions within the system.
- Laid the conceptual groundwork for design and development of PDS 2010.

Questions / Comments