PDS 2010 System Design

Technical Session

June 10-11, 2009

Distributed Infrastructure Design Team

Overview of System Design

Provide an overview of the system design showing the end-to-end flow and the supporting services, tools and applications that will be necessary.

Topics

- Design Team
- Design Principles/Goals
- Service-Based Design
- Ingestion Scenario
- Distribution Scenario

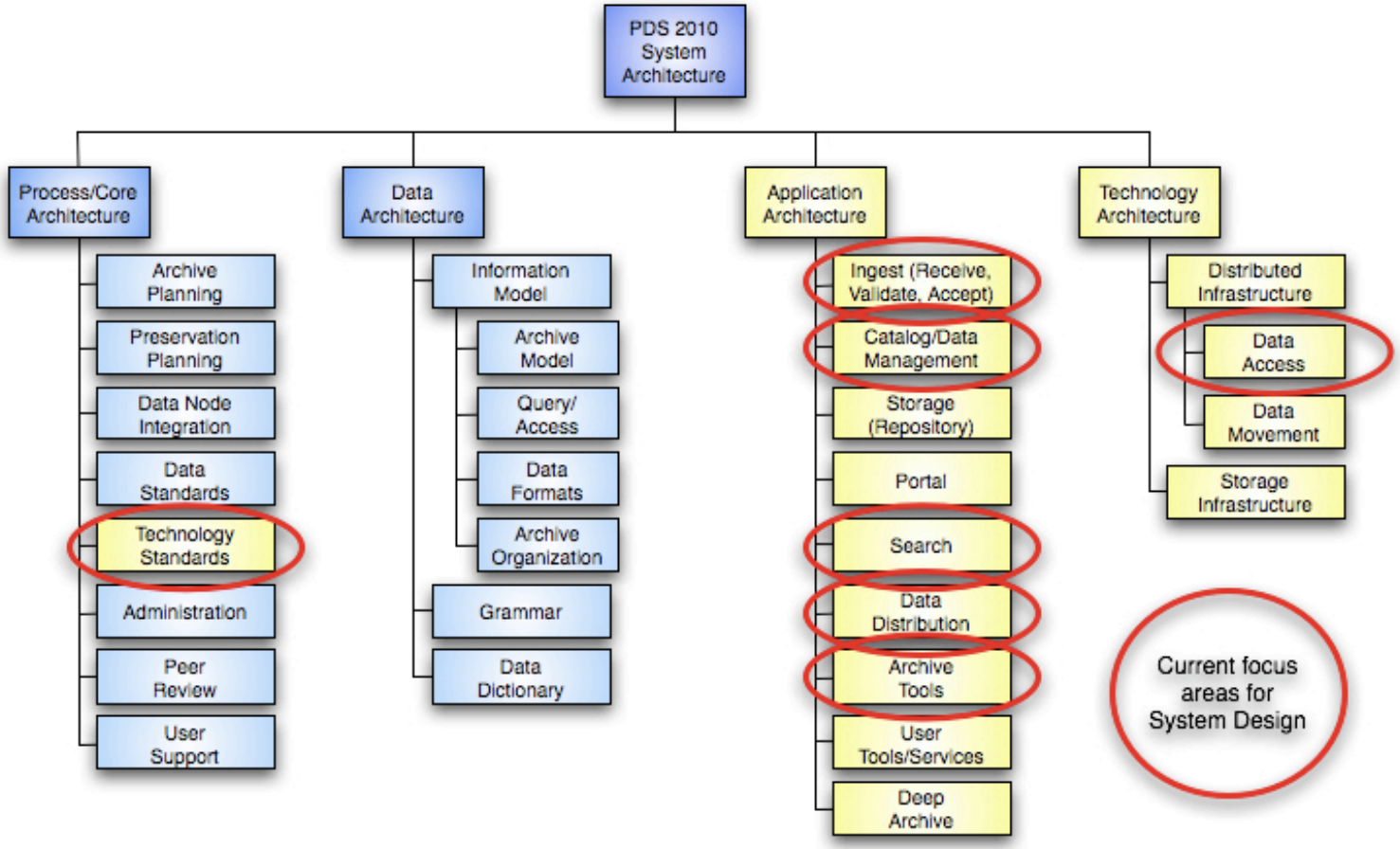
Design Team

- Formed the design team back in January, which consists of the following personnel:
  - Sean Hardman (Engineering)
  - Todd King (PPI)
  - Mike Martin (Management)
  - Paul Ramirez (Engineering)
  - Alice Stanboli (Imaging)
  - Tom Stein (Geosciences)
- Weekly teleconferences (more or less) are held on Tuesday mornings, formerly Thursday afternoons.
- Current artifacts are captured on the PDS Wiki and Engineering Node web sites:
  - http://oodt.jpl.nasa.gov/wiki/pages/viewpage.action?pageId=2600
  - http://pds-engineering.jpl.nasa.gov/index.cfm?pid=100&cid=134

Design Team
Sub-Project/Team Objectives

- Investigate and select the core technologies to be utilized in the development and operation of PDS 2010.
- Initiate development of some of the core services that will serve as building blocks for development of the system.
  - Core services include: Registry, Security, Report, Dictionary and Distributed Access Infrastructure.
- Recent focus has shifted towards defining ingestion and distribution functionality.
- Capture technology standards and service development guidelines for the PDS.

Design Team
Sub-Project/Team Focus



Design Team
Engineering Approach

- Prepare a brief white paper identifying the state-of-the-practice for each service and whether there are COTS or open source solutions available.
- Identify use cases and/or requirements for the service.
- Prepare a design for implementing the service from scratch or for integrating a COTS or open source solution.
- Implement/integrate the service per the design.
- Test the service against the requirements.
- Deploy the system to the target environment (e.g., DN, EN).

Design Principles

- Introduce common software, where appropriate, that is extensible to accommodate discipline-specific needs.
- Isolate technology choices from functionality to facilitate future upgrades.
- Minimize tight-coupling between components to facilitate phased deployment and component replacement.
- Simplify component and user interfaces to facilitate adoption and use of software.
- Utilize open source and COTS solutions when available.

Design Goals

- Improve ingestion efficiency (catalog and data products).
- Facilitate tracking and improve integrity of the archive.
- Facilitate data product search across nodes.
- Improve usability of data products.
- Improve delivery of data to users and deep archive.
- Increase common software development and deployment.

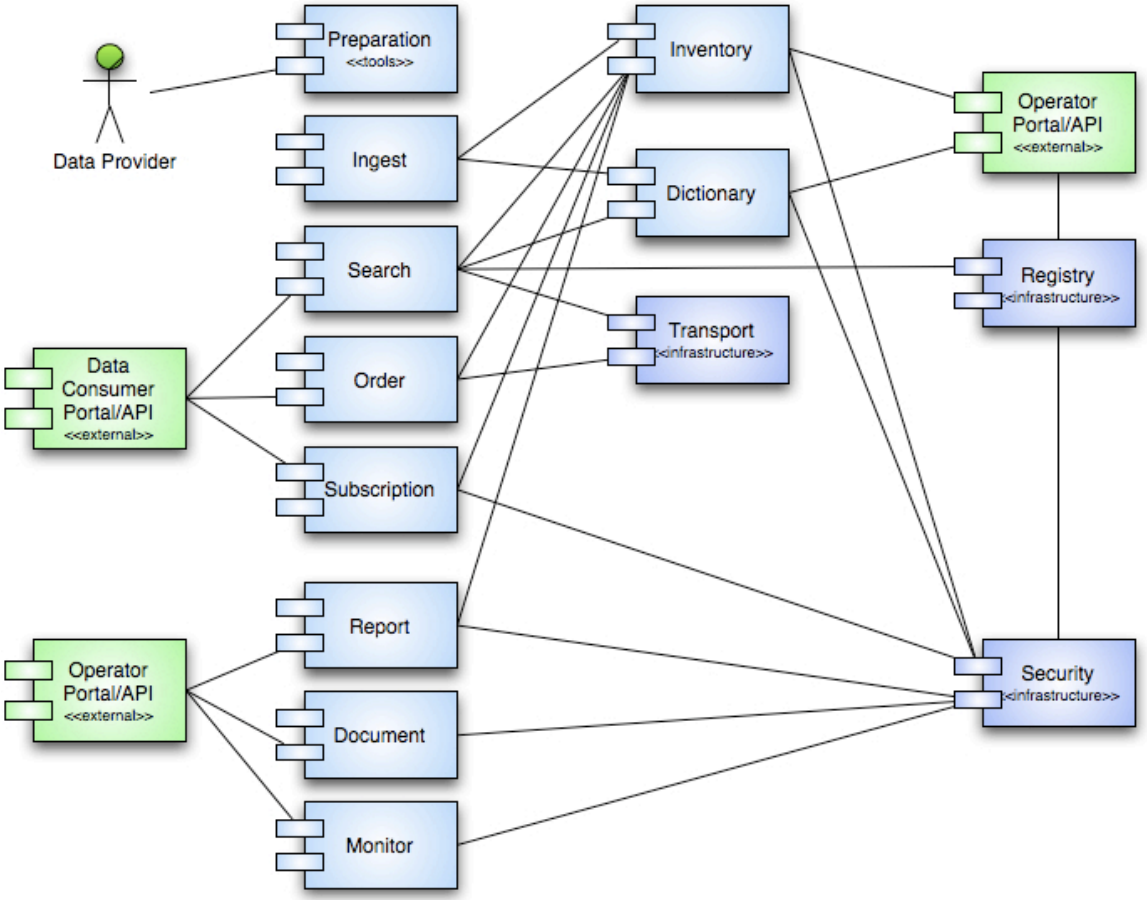
Service-Based Design

- 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.
  - Promotes "loose coupling", "software reuse", "encapsulation" along with other hot buzz phrases in software development today.
  - A service-based architecture provides currency and timeliness for the system.
- Currently working towards a SOA solution that suits PDS.
- Not every function needs to be wrapped by an Internet-accessible service.
  - There is still a place for tools when it comes to preparing data to be archived.
- Service-based functionality will focus on search and retrieval of data.

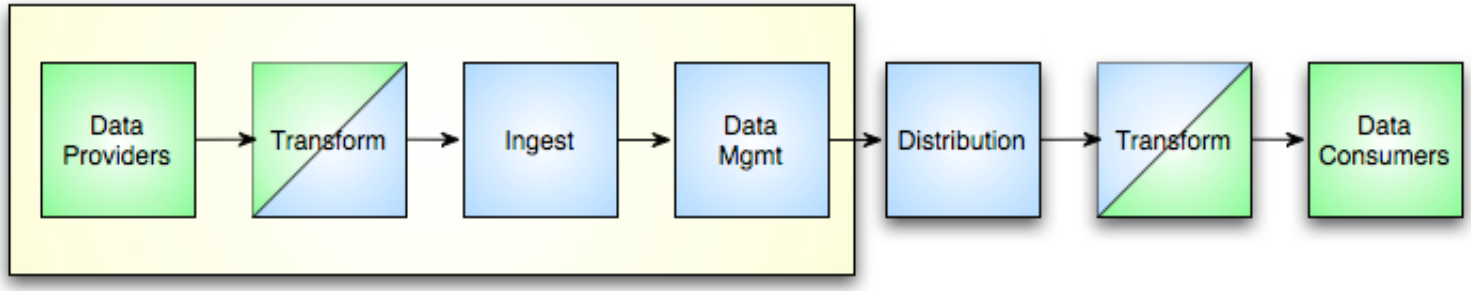
Service-Based Design
Service Vision

- Plans include developing a Service Specification to guide future service and component development for PDS personnel.
  - Will provide details on such things as interface and message content requirements.
  - Will facilitate development of node-specific services/ components (e.g., transformation) that can be integrated with PDS 2010 services.
- The goal is to design and build an extensible system that can grow and have functionality added to over time.

Service-Based Design
Component Identification

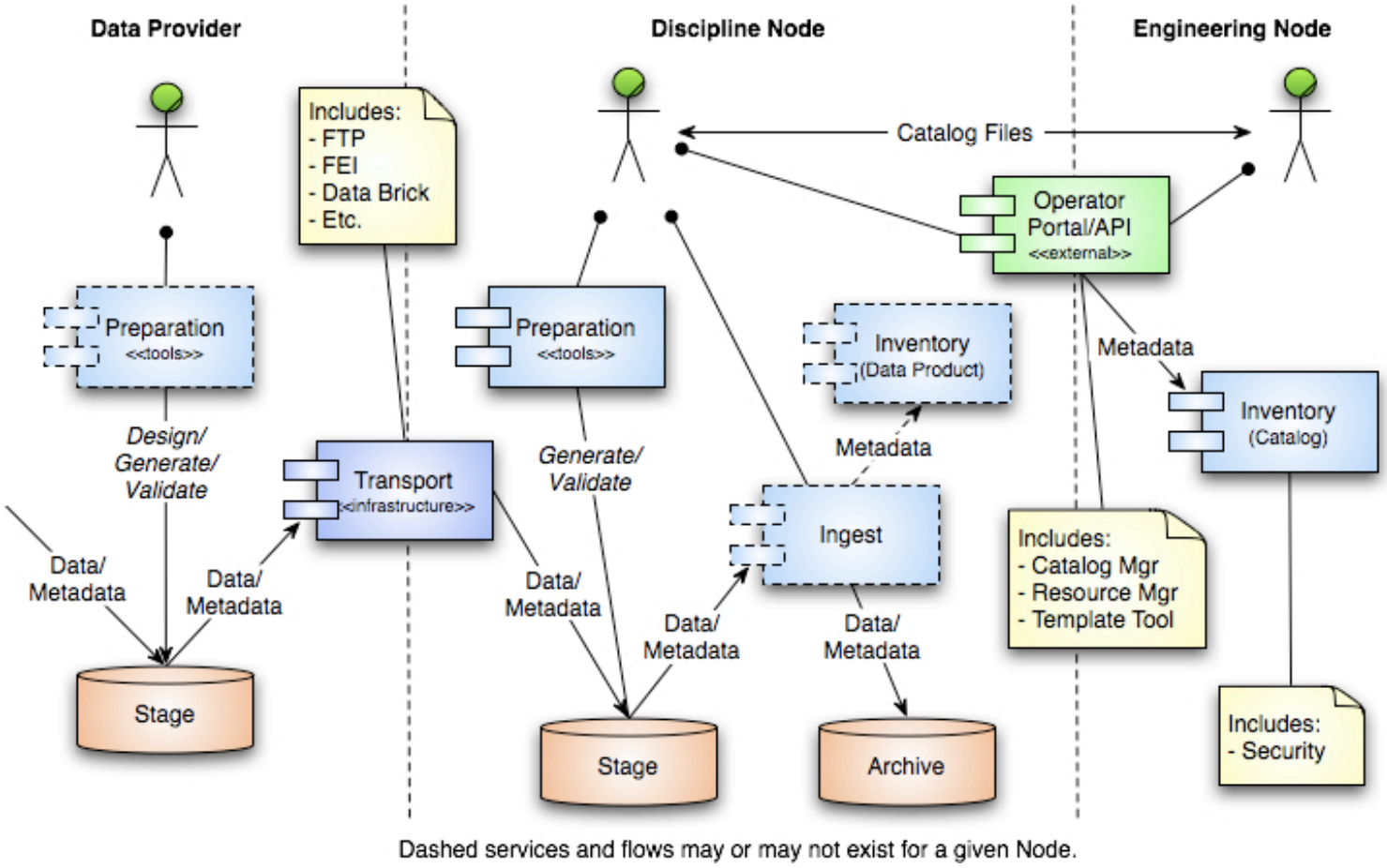


Ingestion Scenario

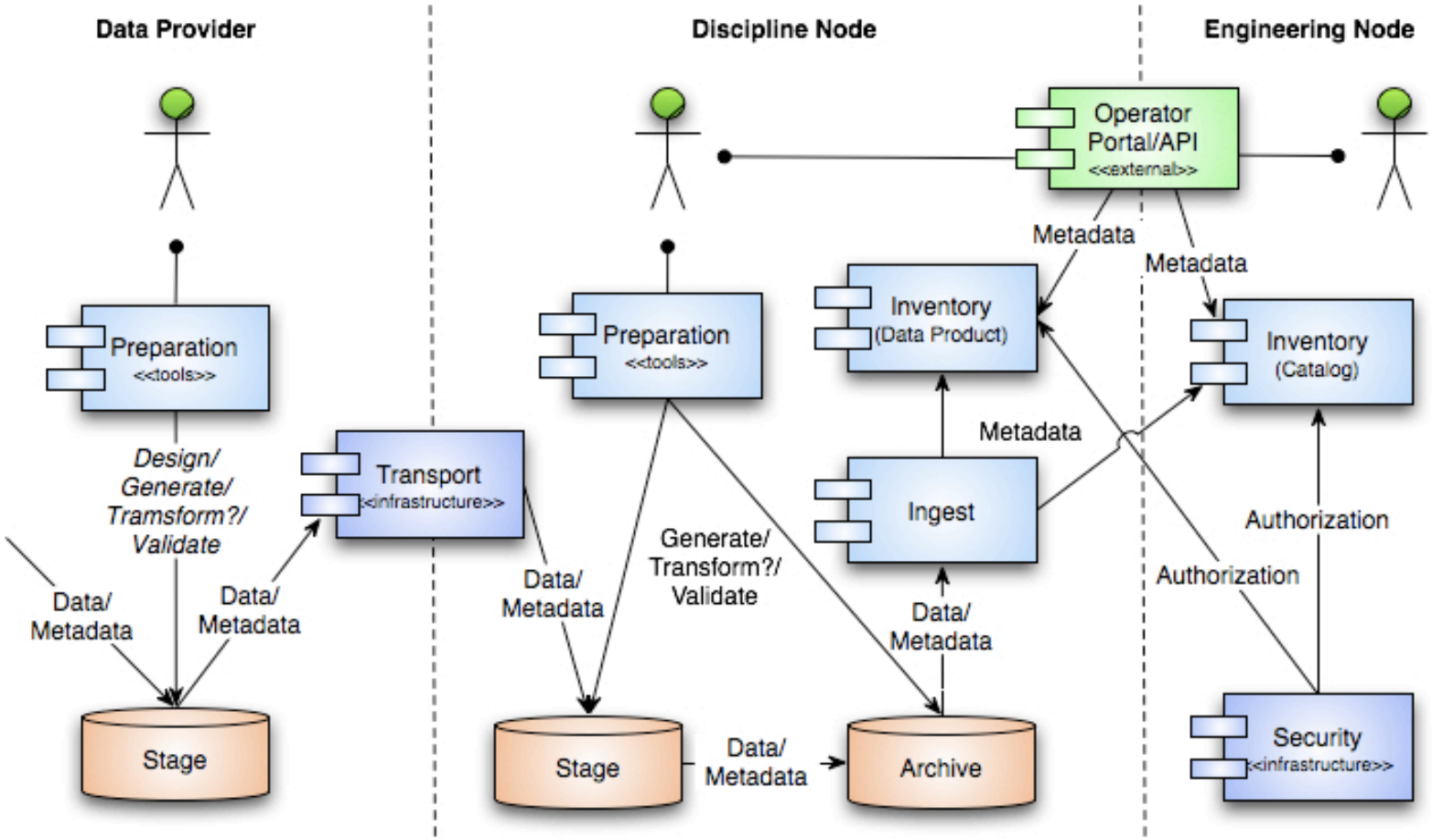


- The ingestion scenario covers ingestion of catalog and data product metadata into their respective Inventory services.
- The proposed ingestion design will be contrasted with the current design and focuses on the area in the end-to-end diagram above highlighted by the yellow box.

Ingestion Scenario
Current Design



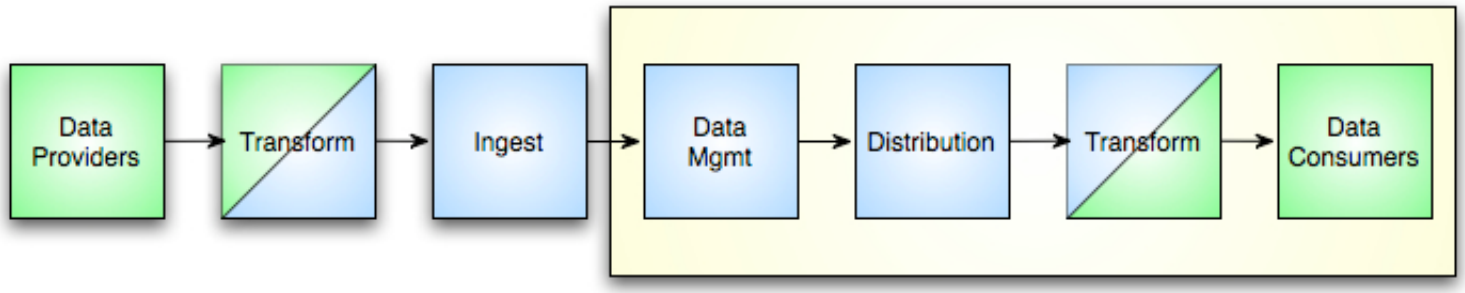
Ingestion Scenario
Proposed Design



Ingestion Scenario
Design Differences

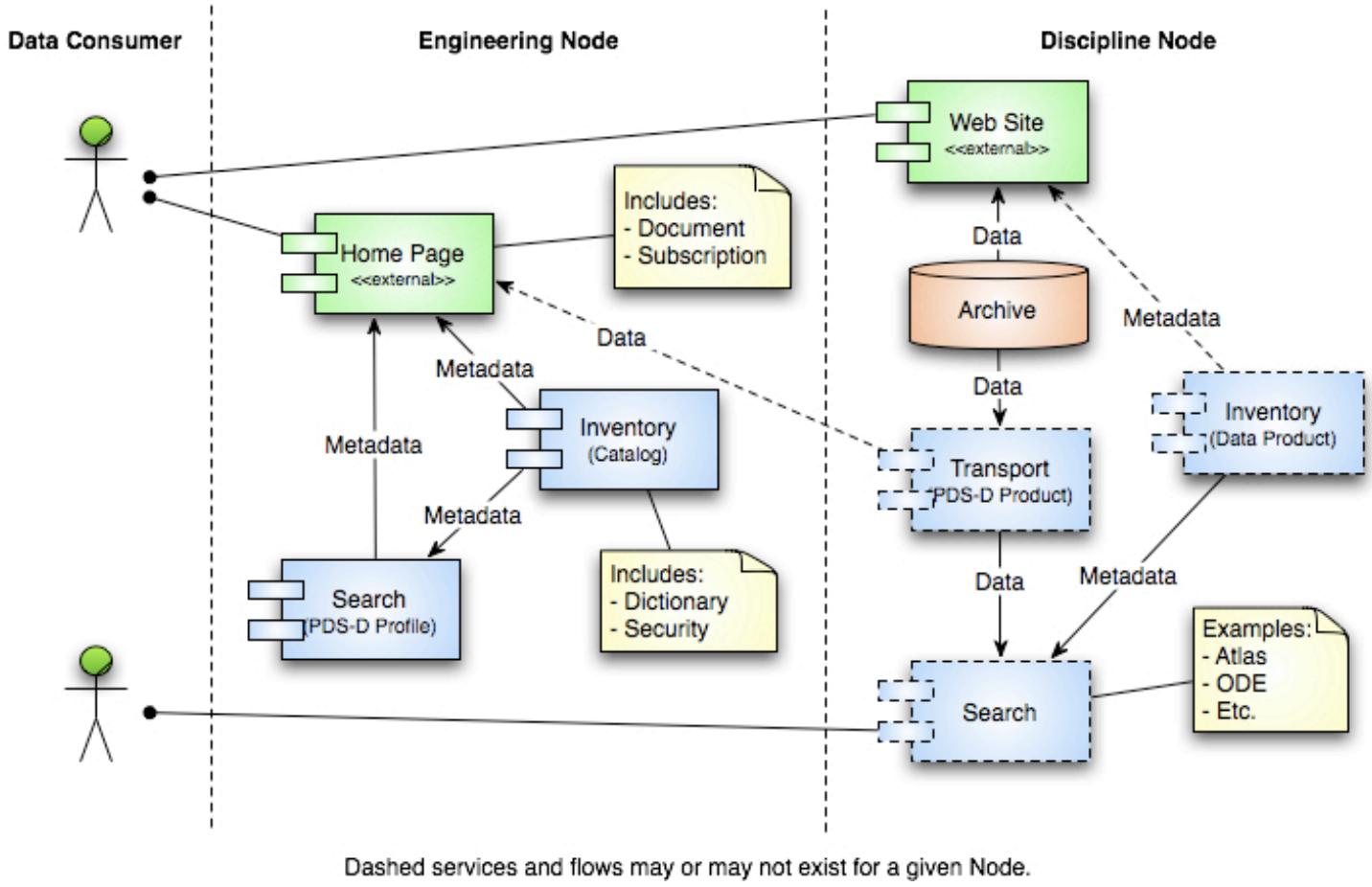
- Transformation of incoming data/metadata is shown as a possible function for the Data Provider or the Discipline Node via a tool.
- A portal/API will be utilized for submission of catalog-level metadata.
  - Alleviating the need for pushing catalog files around the system.
- An Ingest Service is introduced for capturing and registering data product and catalog-level metadata.
- An extensible Node-specific Inventory Service is introduced for tracking data product submissions at a minimum but could support data product search at the Node.

Distribution Scenario

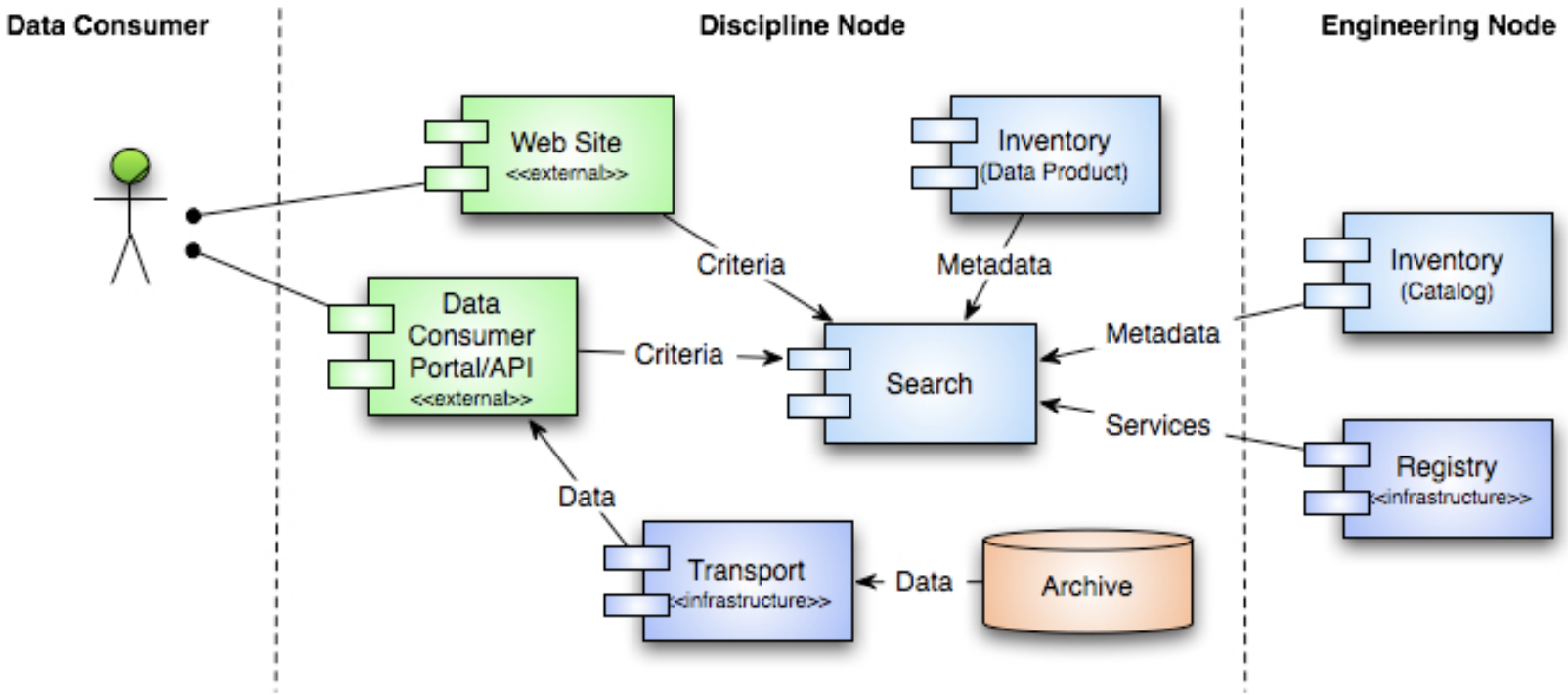


- The distribution scenario covers search of the catalog and data product metadata and distribution of associated data.
- The proposed distribution design will be contrasted with the current design and focuses on the area in the end-to-end diagram above highlighted by the yellow box.
- The proposed design includes scenarios for DN and EN initiated searches.

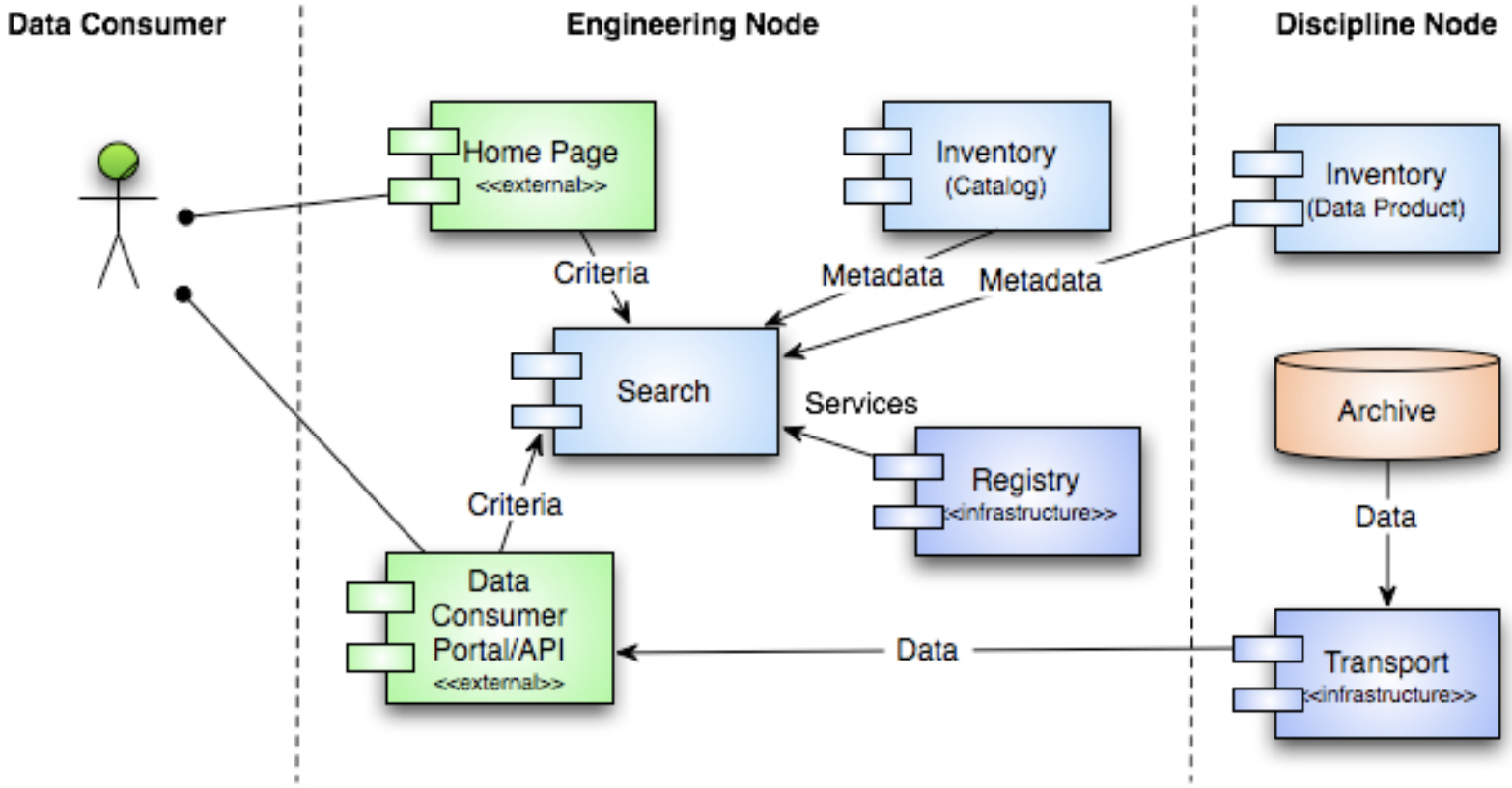
Distribution Scenario
Current Design



Distribution Scenario
Proposed Design (DN Search)



Distribution Scenario
Proposed Design (EN Search)



Distribution Scenario
Design Differences

- Utilization of a common Search Service for interfacing with the services hosting catalog-based and Node-specific data product metadata.
- Extensible but common, Transport Service is introduced to facilitate access and usability (i.e., transformation) of data products.
- Introduction of a Data Consumer portal/API for discovering and retrieving data/metadata.

Service and Tool details are forthcoming.