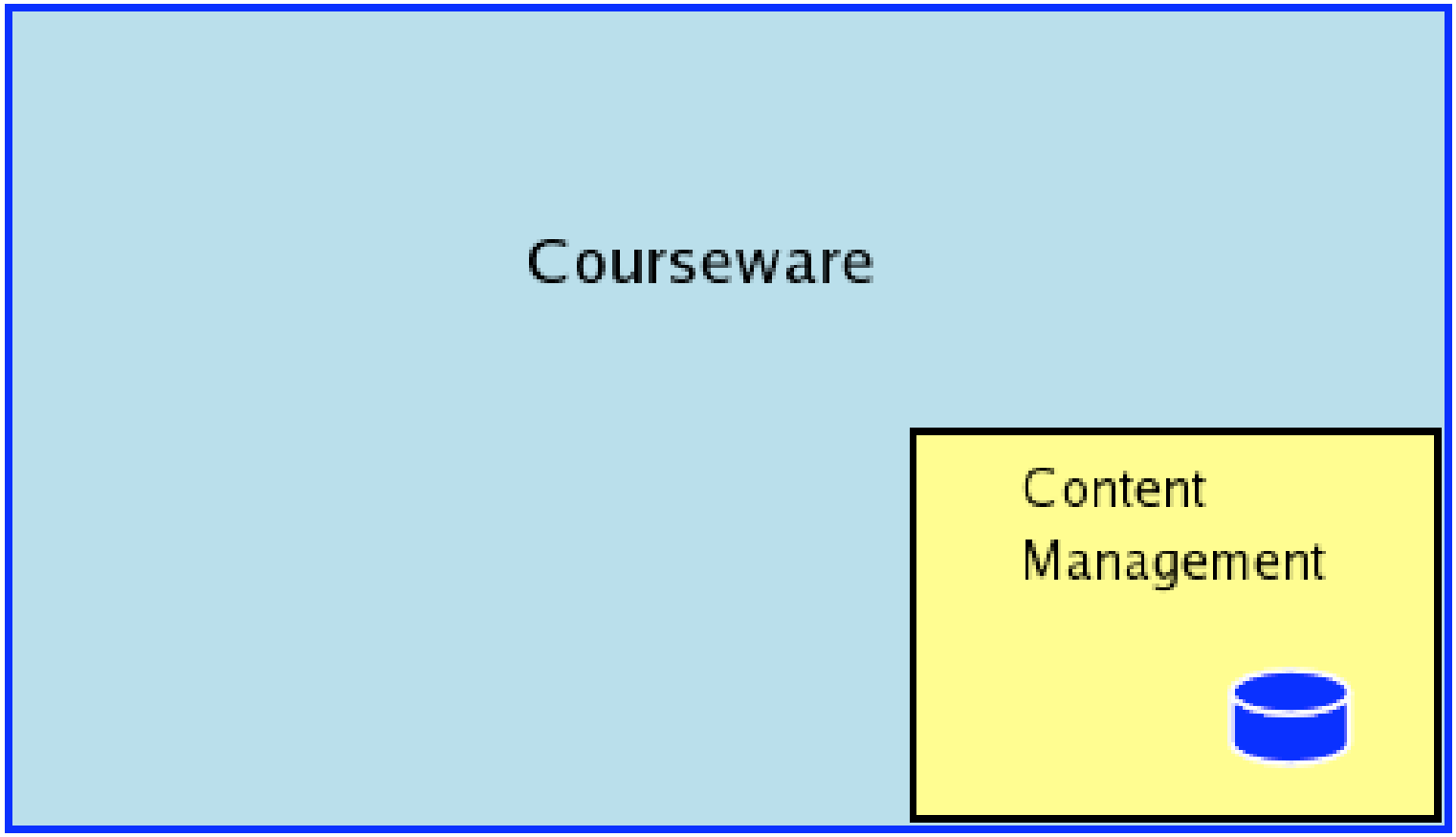
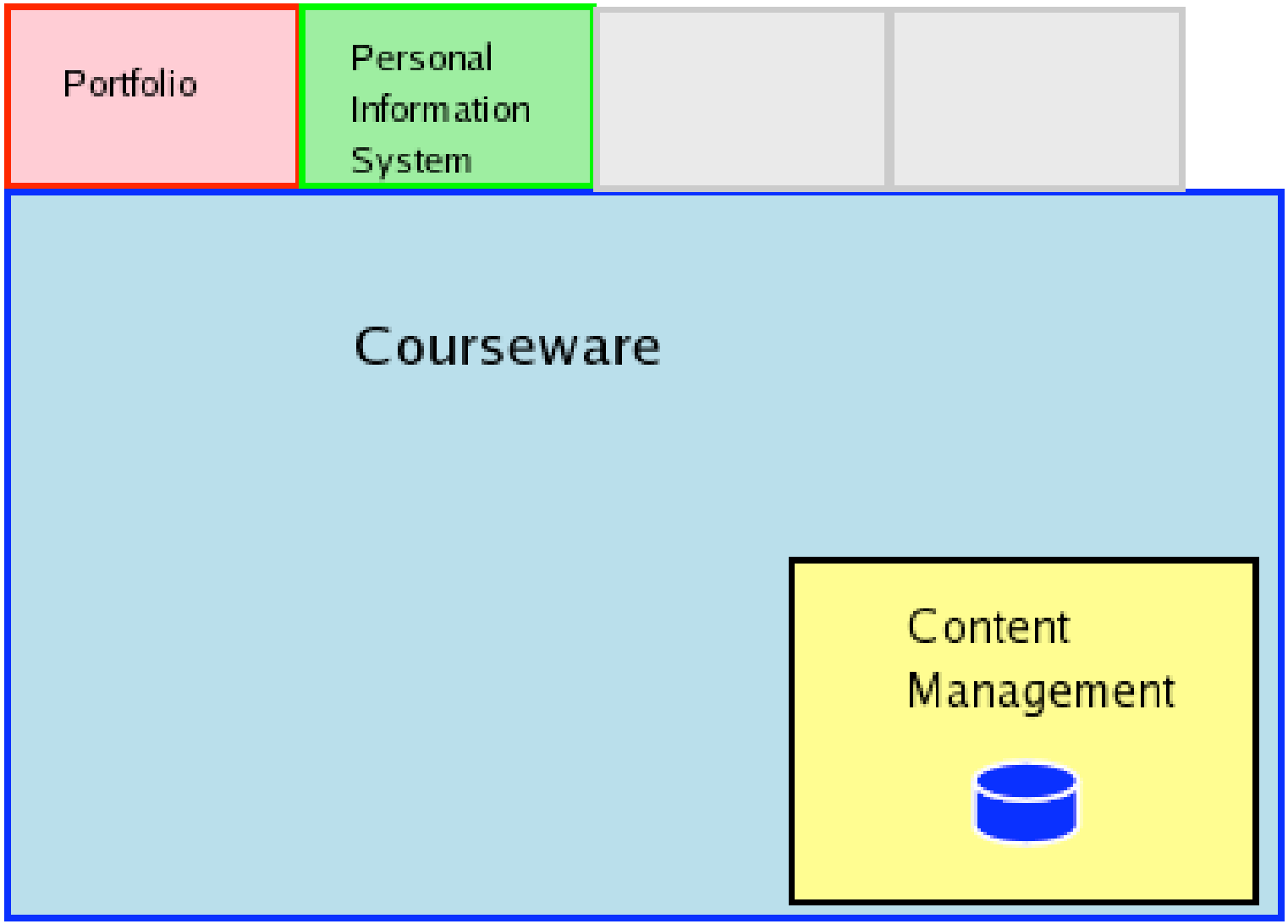
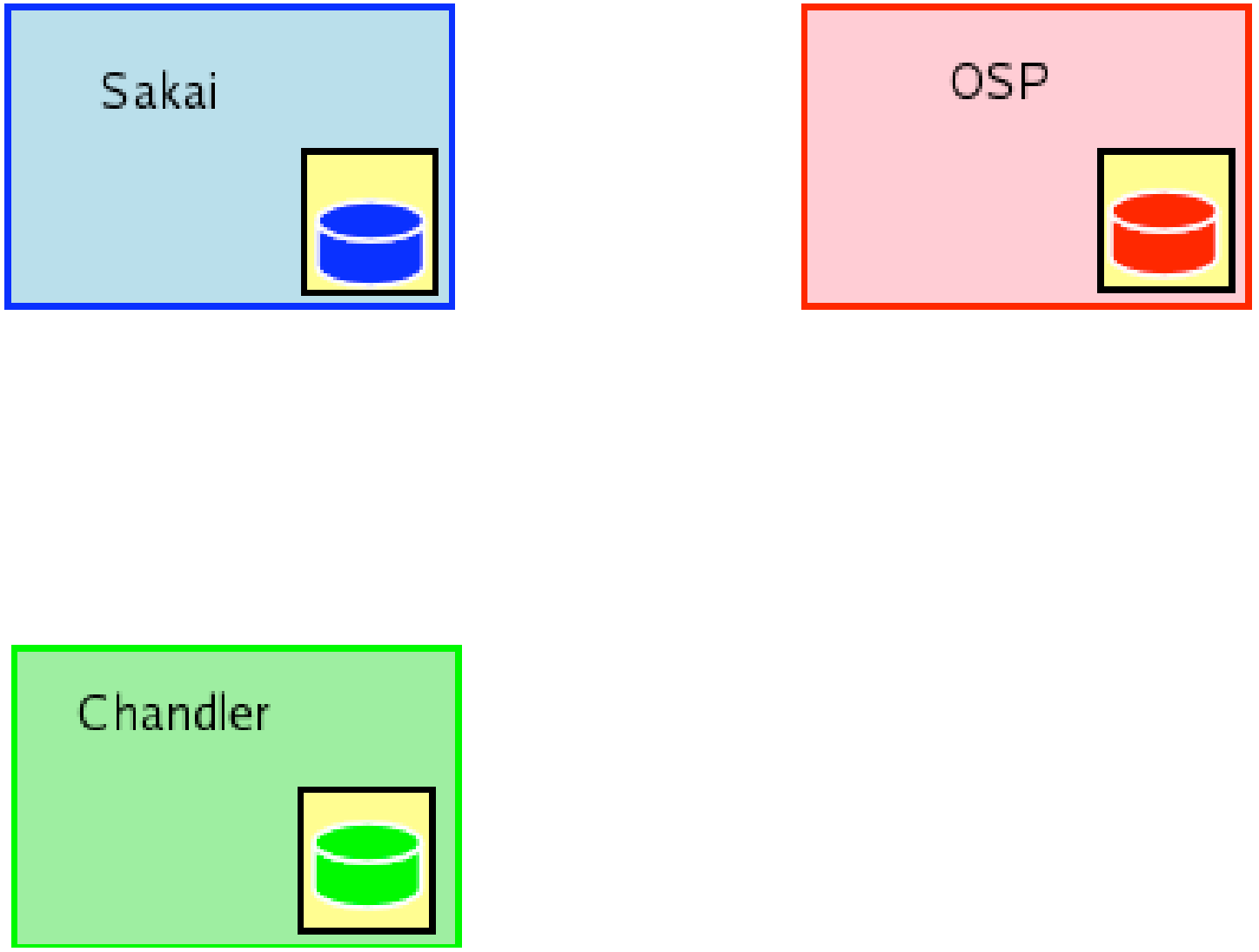
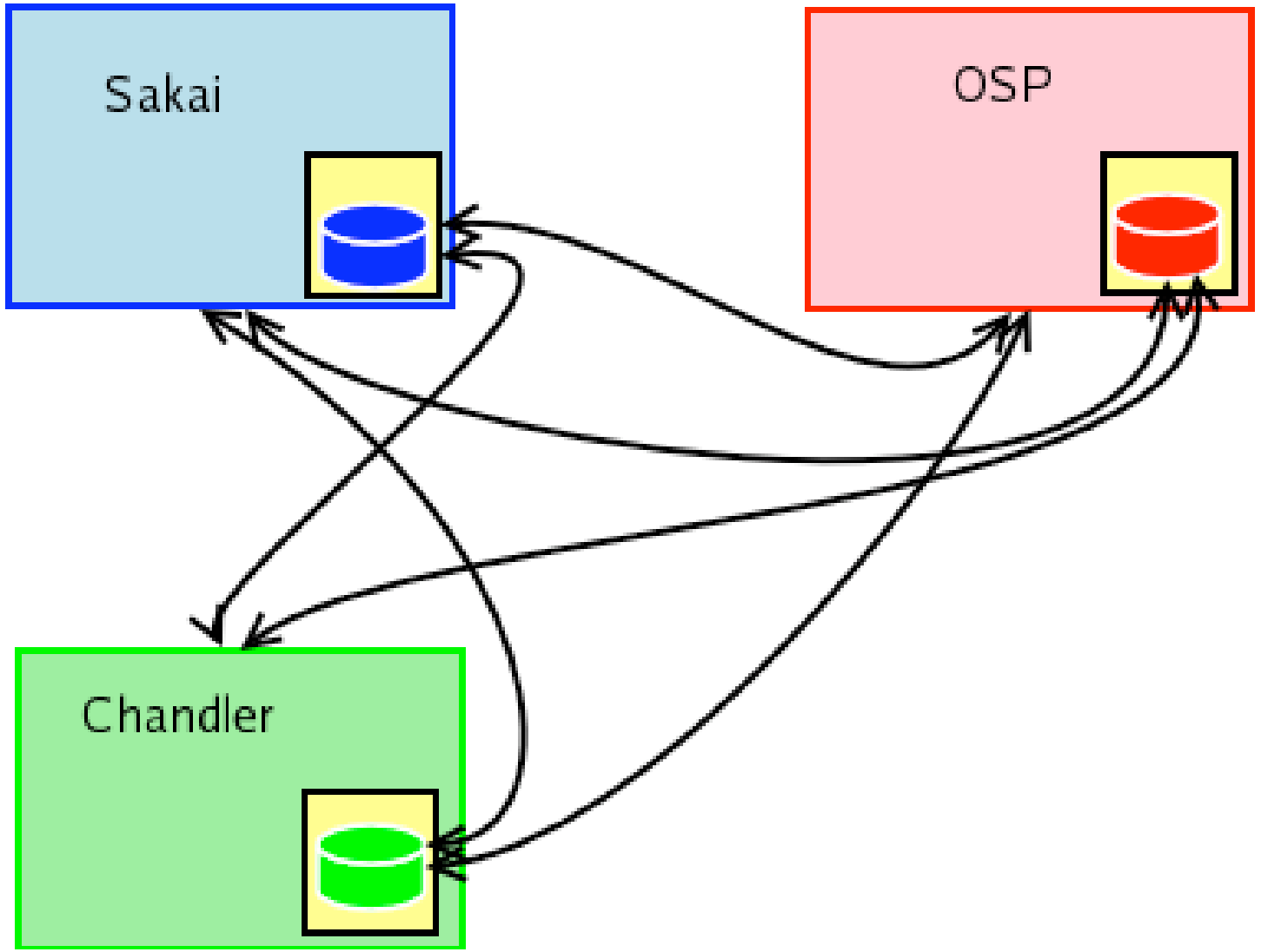
A Technology Analysis of Repositories and Services

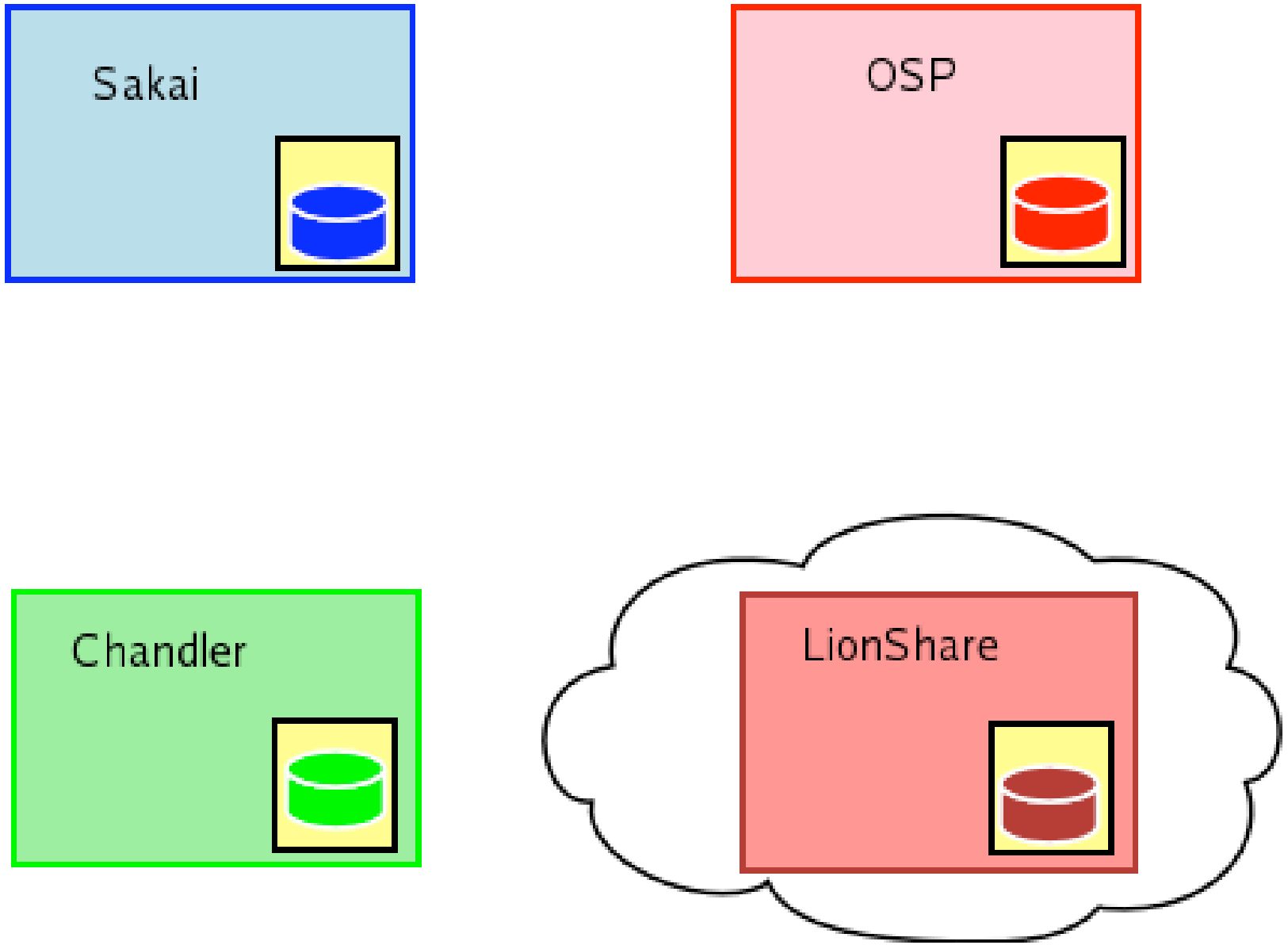
Sayeed Choudhury
Tim DiLauro

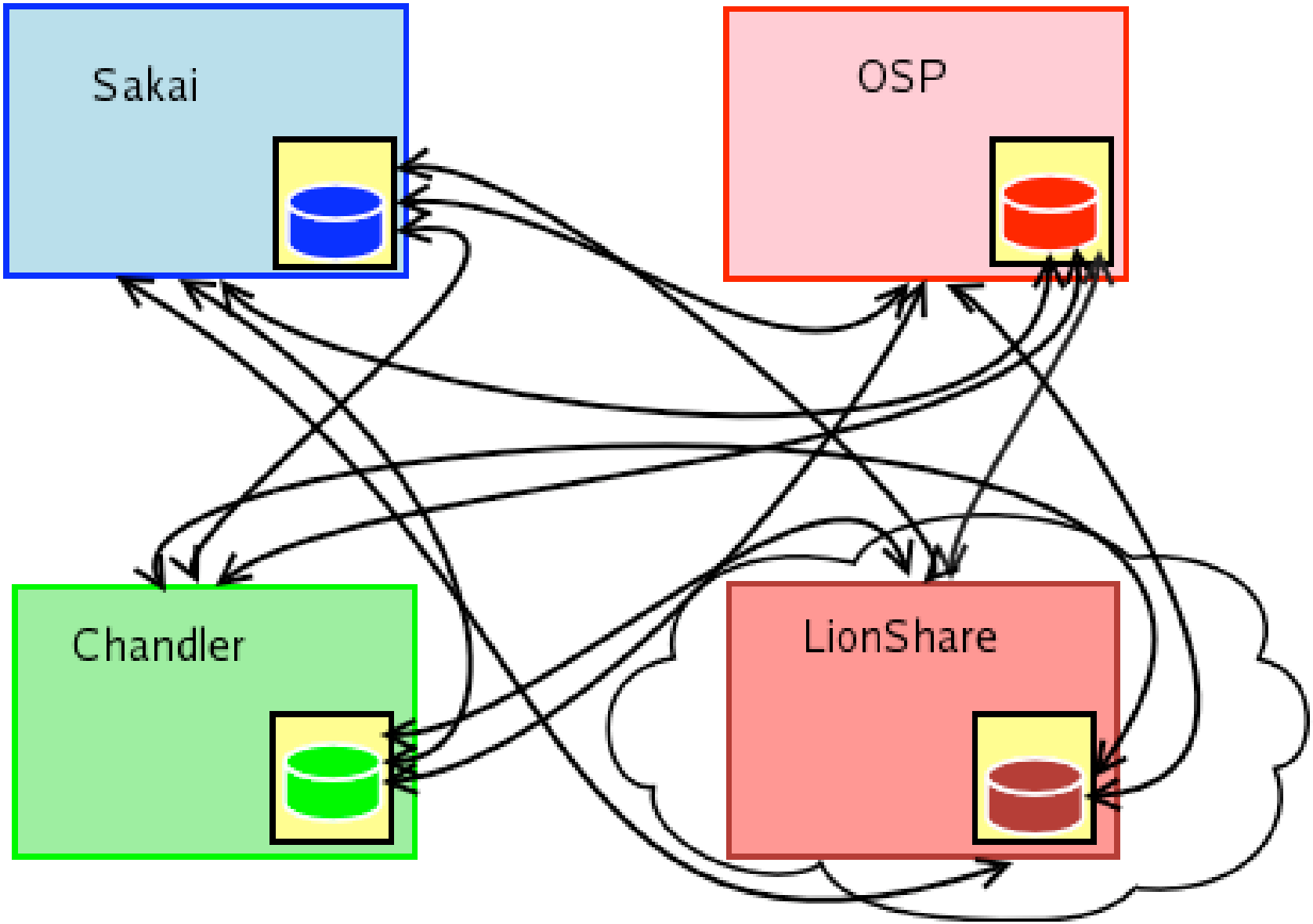

Courseware
Content
Management

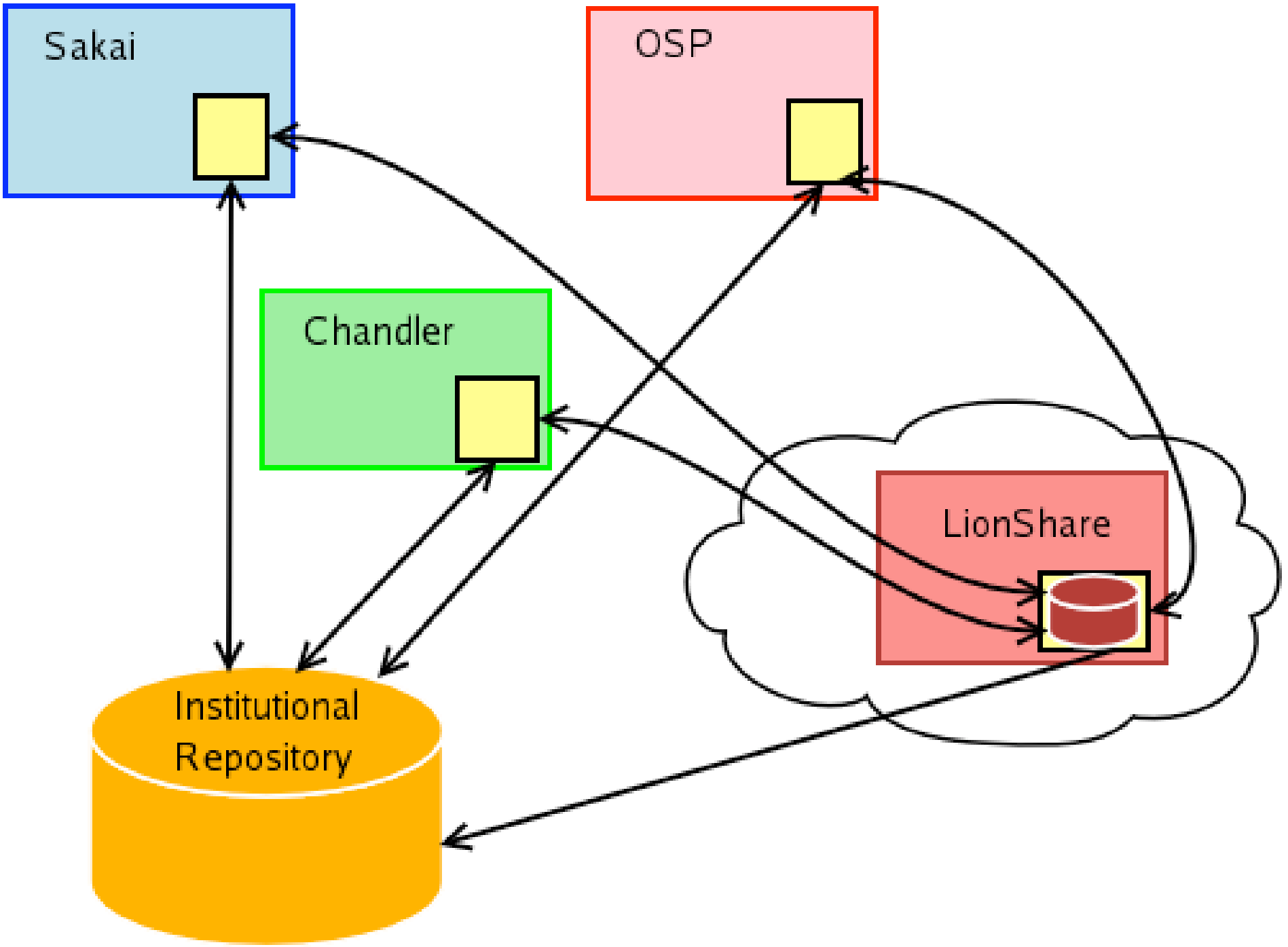

Portfolio
Personal
Information
System
Courseware
Content
Management

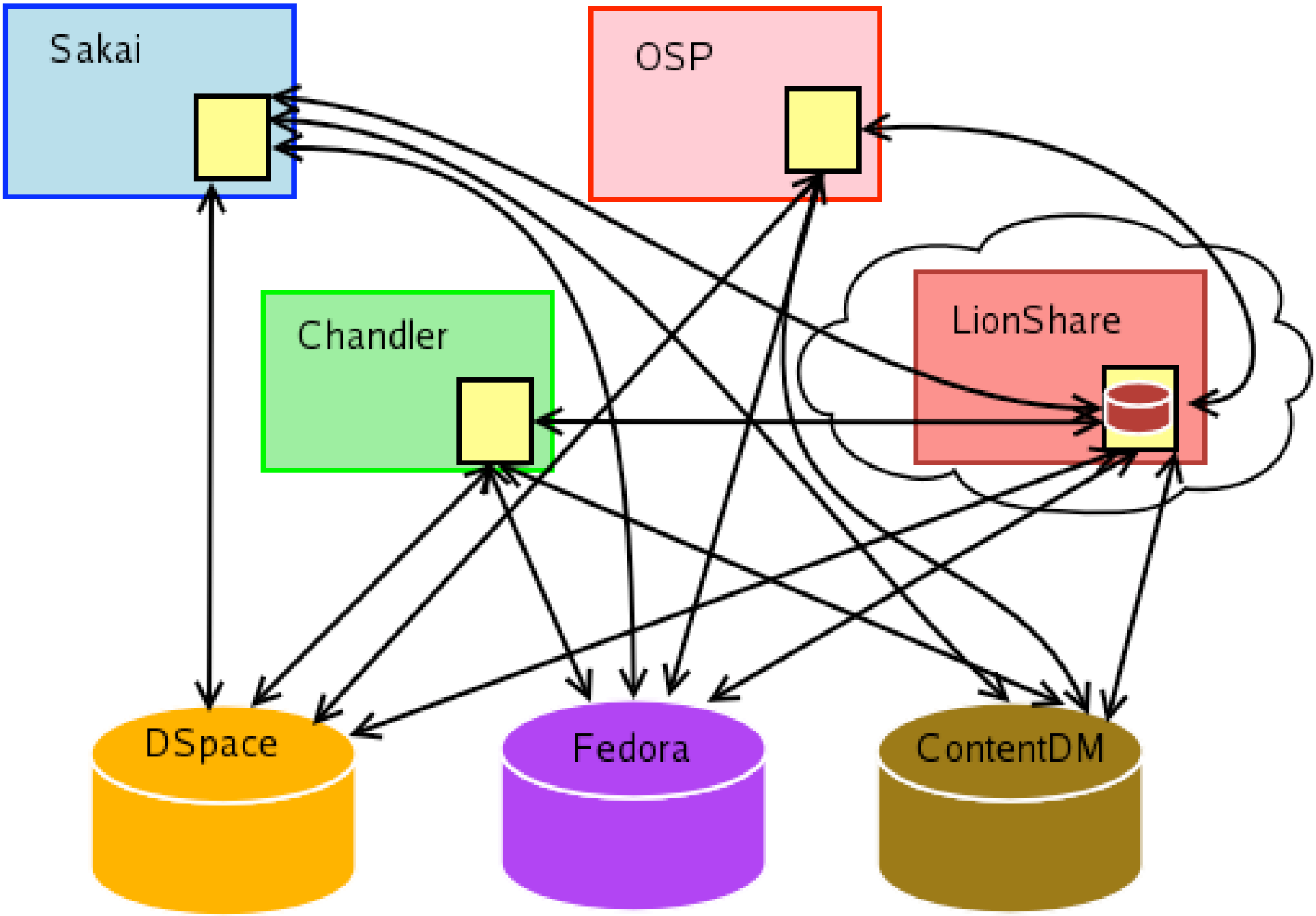

Sakai
OSP
Chandler

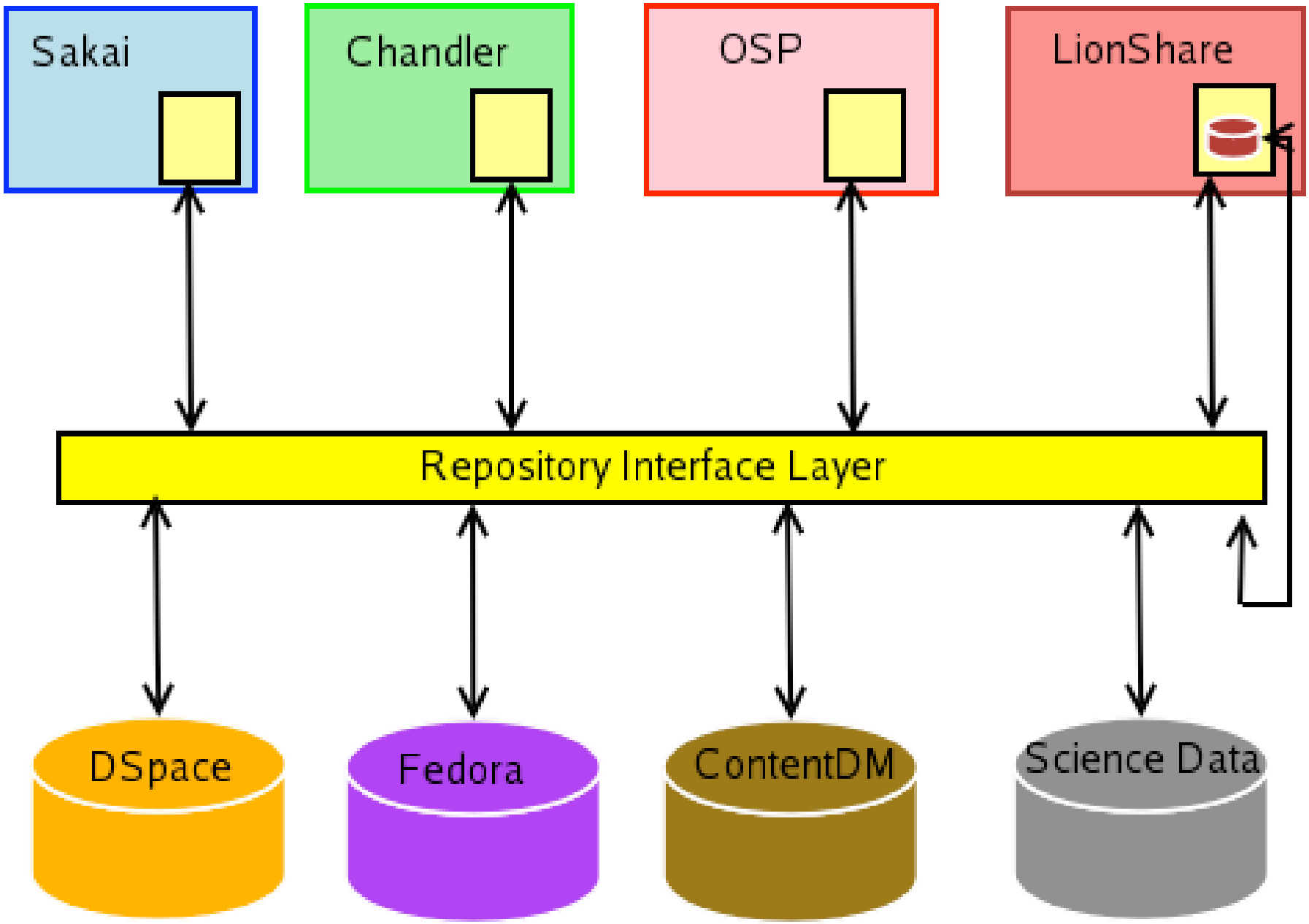

Sakai
OSP
Chandler


Sakai
OSP
Chandler
LionShare


Sakai
OSP
Chandler
LionShare


Sakai
OSP
Chandler
LionShare
Institutional
Repository


Sakai
OSP
Chandler
LionShare
DSpace
Fedora
ContentDM


Sakai
Chandler
OSP
LionShare
Repository Interface Layer
DSpace
Fedora
ContentDM
Science Data

Our Use Case Approach

- Start with scenarios, or “stories”
- Record sequence of key events
- Cluster events across scenarios
- Develop use cases from event clusters
  (e.g., fetch a digital object, given an identifier)

Motivation for Approach

- Foster shared vocabulary, based on user requirements
- Ease evaluation of repository-related technologies
- Reduce complexity of application integration
- Support gap analysis
- Create an extensible, reusable mechanism

AIHT Format Migration Scenario

Johns Hopkins has ingested into a repository the contents of the George Mason University The September 11 Digital Archive as provided by LC for Archive Ingest and Handling Test (AIHT). Recently, it has been determined that the JPEG format can no longer be supported in our environment.

Miles Parker, the repository manager, has decided to transform each JPEG datastream to a corresponding TIFF, storing some information about the original JPEG in one of the TIFF tags. Some code has already been written to perform this JPEG to TIFF transformation. Miles needs to run this code on every datastream for which the current version is a JPEG and then store the results as the current version of the datastream, keeping the JPEG as a previous version.

Because the repository contains more than 57,000 digital objects, Miles creates a program to do this. This program iterates over each object in the repository, checking the current version of each datastream for a MIMEType of "image/jpeg". For each datastream that matches the criteria, the transformation is performed and the new version is stored into the repository. In addition, metadata describing the transformation is also stored in the affected digital object.

Format Migration Use Case

Use Case Name:	**Migrate Content to New Format**
Summary:	Content datastreams within a repository must be systematically migrated from one (input) format to another (output) format when, for example, the former goes out of scope.
Actors:	Data Curator or automated data management system
Basic Course of Events:	1) The curator or automated process initiates a function to perform the format migration, providing criteria for selection of datastreams that should be migrated. 2) The function fetches the next matching datastream from the repository, performs the transformation, and stores the output datastream as a new version of the input datastream. 3) Metadata for the new datastream and metadata tracking the transformation is added to the associated digital object. 4) Steps (2) and (3) are repeated until no more datastreams match the selection criteria.
Alternative Paths:	• In step (2), an agent is injected into the repository instead of a function pulling content out of the repository. • In step (2), the output of the transformation replaces the original input datastream.
Exception Paths:	
Extension Points:	
Triggers:	The curator or automated system becomes aware that a format is going out of scope or that a format migration is required for some other reason.
Assumptions:	Software to convert from the input to the output format exists or can be created.
Preconditions:	
Postconditions:	
Author:	Tim DiLauro
Date:	Created 2005-04-11. Last modified 2005-04-11.
Scenarios:	AIHT Format Migration

Applying Use Cases

- Develop functional requirements from use cases
- Investigate support for these requirements in repositories
- Map interface specifications to functional requirements

Functional Requirements Derived from Use Cases

	Use Case 1	Use Case 2	Use Case 3	Use Case 4
Functionality 1	*	*	*	*
Functionality 2		*	*	
Functionality 3	*			
Functionality 4		*	*	*
Functionality 5	*			
Functionality 6		*		
Functionality 7			*	
Functionality 8				*
Functionality 9	*			*
Functionality 10		*	*	*

Determine Repository Support for each Functionality

	Fedora 2.0	Dspace 1.2.1	Other Repo	…
Functionality 1	*	*		*
Functionality 2	*		*	
Functionality 3	*			
Functionality 4		*	*	*
Functionality 5	*	*		
Functionality 6	*	*	*	
Functionality 7		*	*	
Functionality 8				
Functionality 9	*	*		
Functionality 10	*	*	*	

Map Interface Specifications to Functional Requirements

	JSR-170 0.16.3	DR OSID 2.0	IMS DRI 1.0	Other API
Functionality 1	*	*		
Functionality 2	*	*		
Functionality 3	*	*		
Functionality 4		*	*	*
Functionality 5	*	*		
Functionality 6	*		*	*
Functionality 7	*		*	
Functionality 8	*	*		
Functionality 9	*	*	*	
Functionality 10	*	*	*	

Gap Analysis

	Use Cases				Repositories			API Specifications			
	UC 1	UC 2	UC 3	UC 4	Fedora	Dspace	Other	JSR-170	DR OSID	IMS DRI	Other
Functionality 1	*	*	*	*	*	*	-	*	*	-	-
Functionality 2		*	*		*	-	*	*	*	-	-
Functionality 3	*				*	-	-	*	*	-	-
Functionality 4		*	*	*	-	*	*	-	*	*	*
Functionality 5	*				*	*	-	*	*	-	-
Functionality 6		*			*	*	*	*	-	*	*
Functionality 7			*		-	*	*	*	-	*	-
Functionality 8				*	-	-	-	*	*	-	-
Functionality 9	*			*	*	*	-	*	*	*	-
Functionality 10		*	*	*	*	*	*	*	*	*	-

Gap Analysis

	Use Cases			Repositories			API Specifications			
	UC 2	UC 3	UC 4	Fedora	Dspace	Other	JSR-170	DR OSID	IMS DRI	Other
Functionality 1	*	*	*	*	*	-	*	*	-	-
Functionality 2	*	*		*	-	*	*	*	-	-
Functionality 3				*			*	*		
Functionality 4	*	*	*	-	*	*	-	*	*	*
Functionality 5				*	*		*	*		
Functionality 6	*			*	*	*	*	-	*	*
Functionality 7		*		-	*	*	*	-	*	-
Functionality 8			*	-	-	-	*	*	-	-
Functionality 9			*	*	*	-	*	*	*	-
Functionality 10	*	*	*	*	*	*	*	*	*	-

Acknowledgements

The Mellon Foundation
Library of Congress

University of Virginia
Massachusetts Institute of Technology

Write scenarios or use cases for repository connectivity for your own projects!

Send to us at <scenarios@dkc.jhu.edu>

Questions?