SAP BusinessObjects USER CONFERENCE

October 5-7, 2010 | Orlando, Florida
Walt Disney World Swan and Dolphin

Metadata Madness

Alan Mayer
Solid Ground Technologies

Session 303

ASUG Real Experience. Real Advantage.

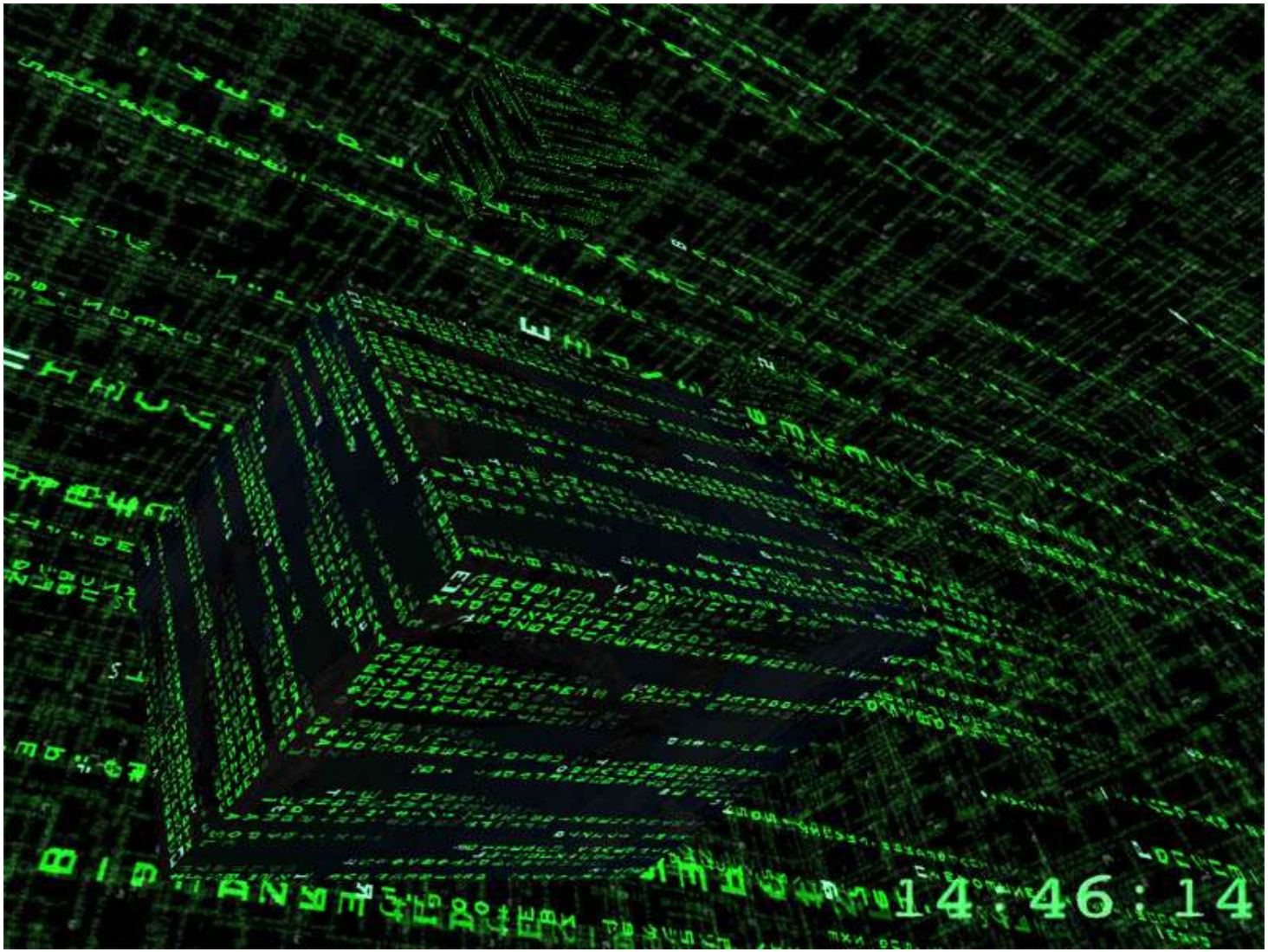
Agenda

- Introduction
- Metadata Basics
- Architecture
- Deployment
- Harvesting Metadata
- Consuming Metadata
- Conclusion

Introduction

Introduction

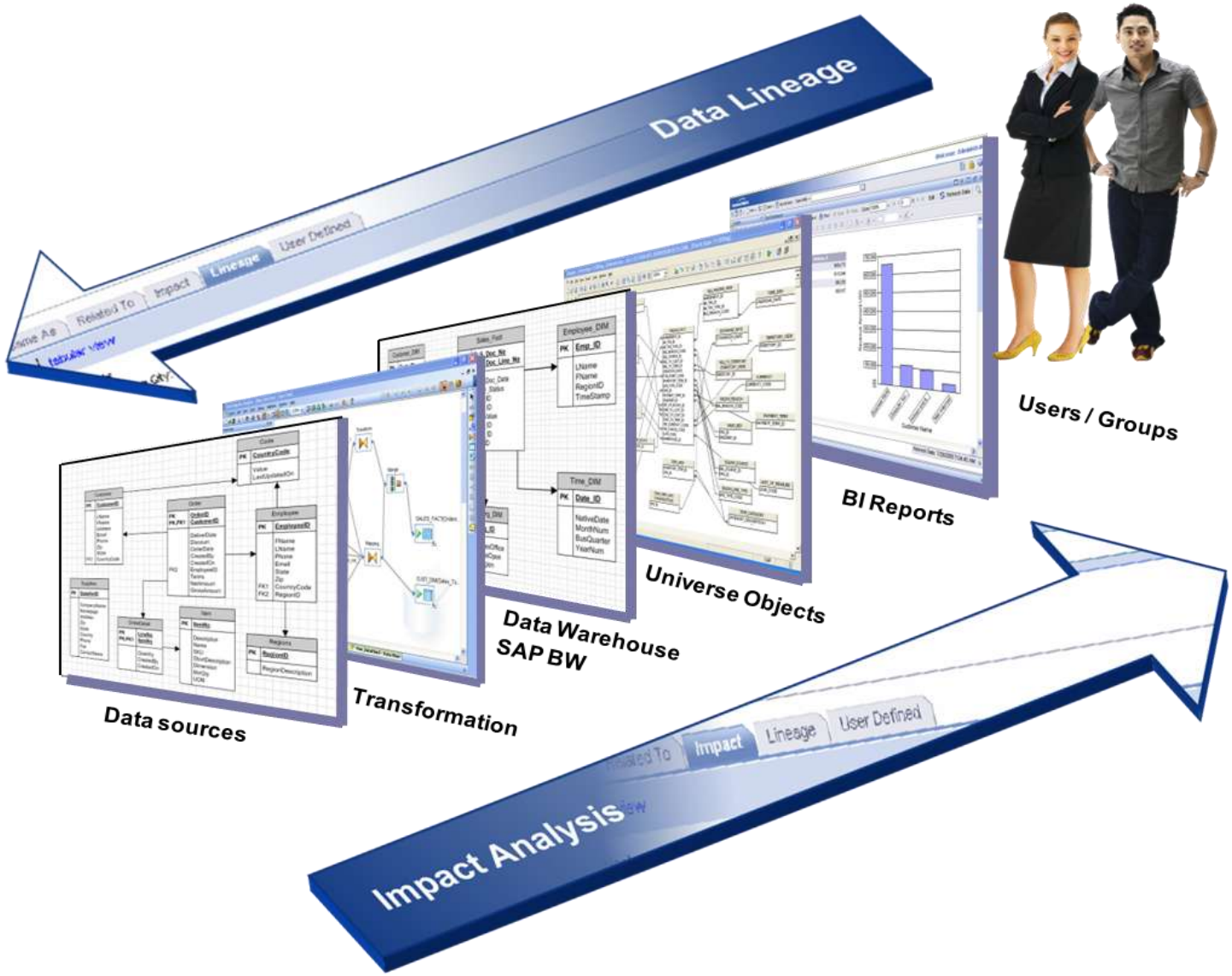
Introduction



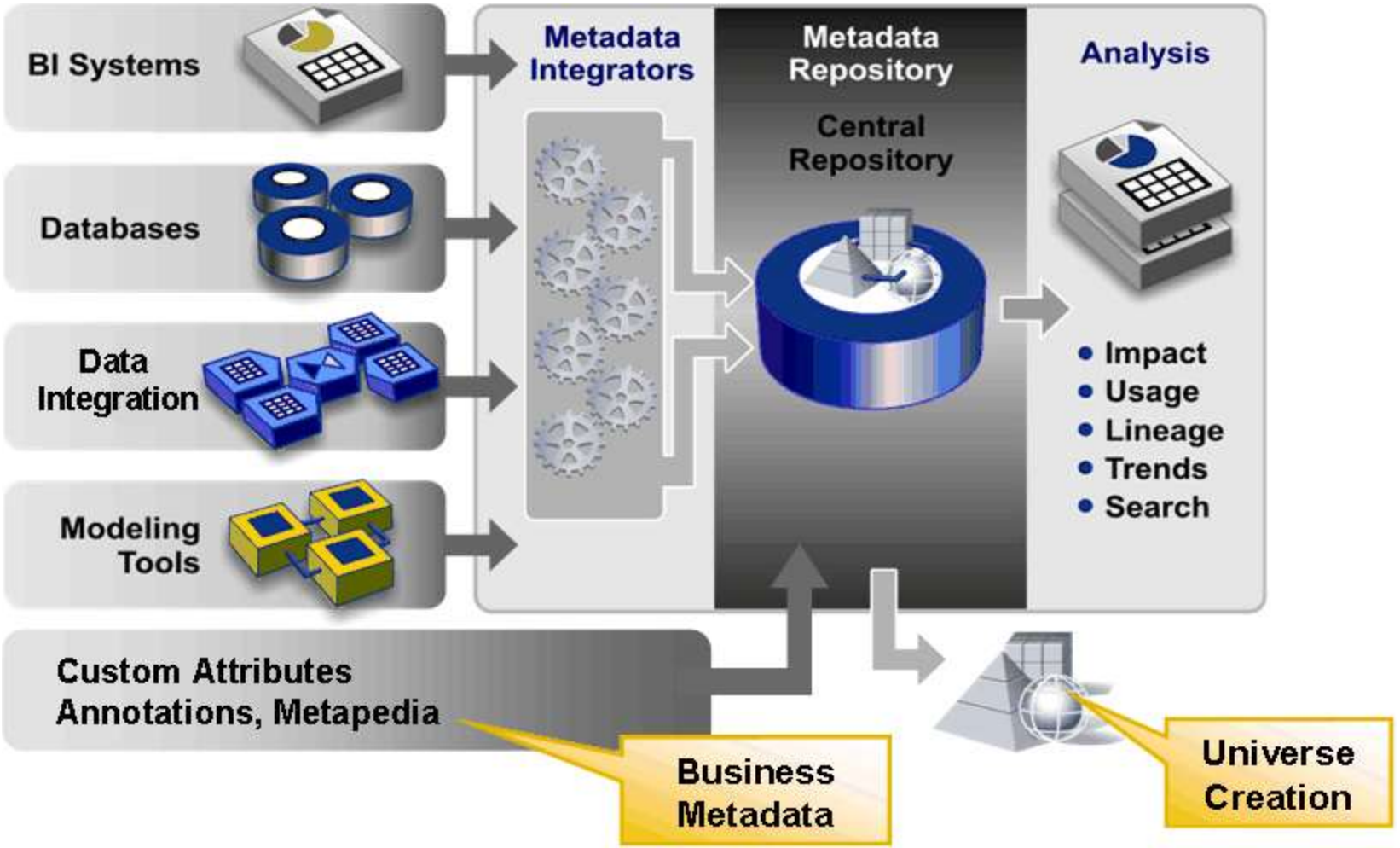
Introduction, cont'd

Metadata Basics

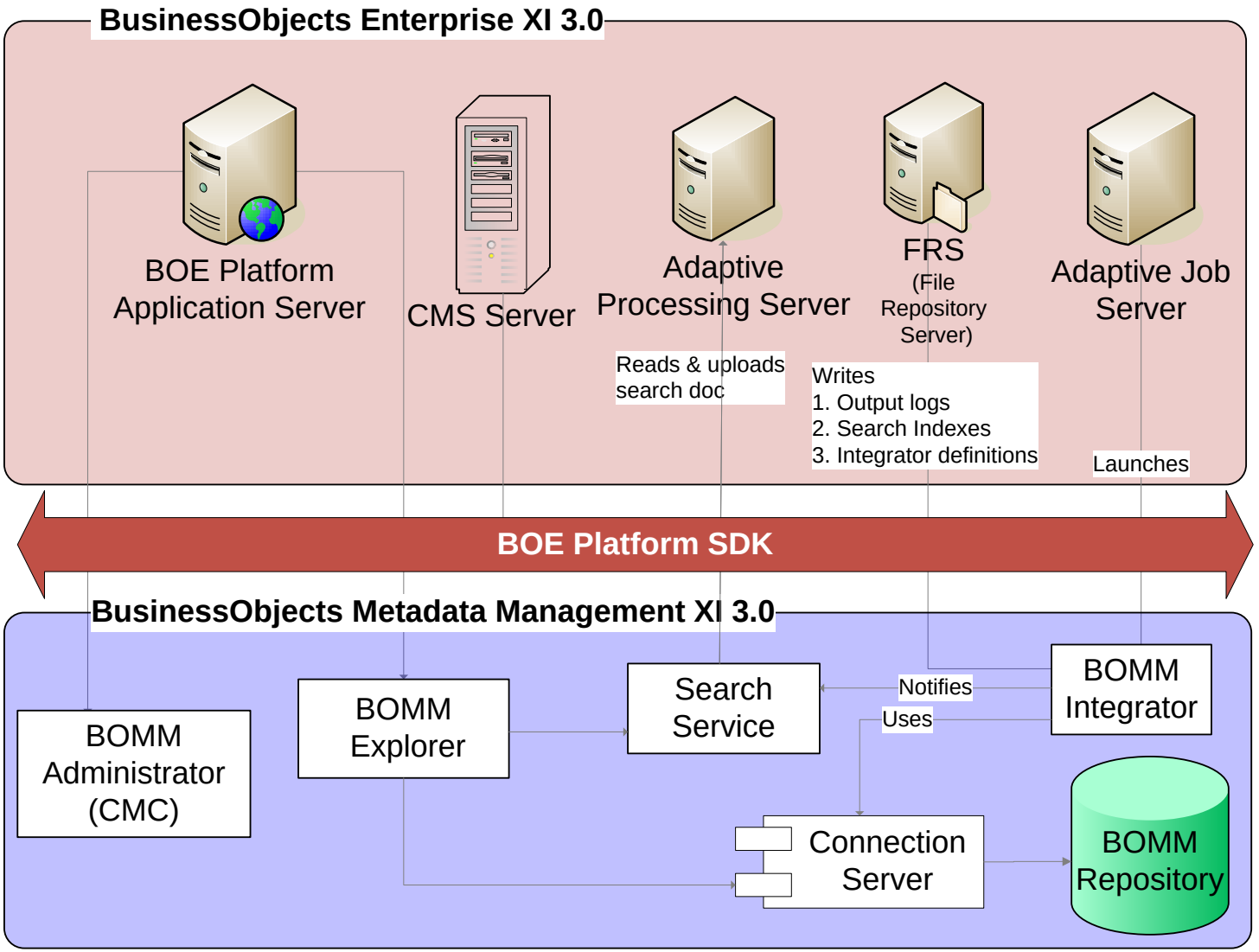
Metadata Basics



Architecture

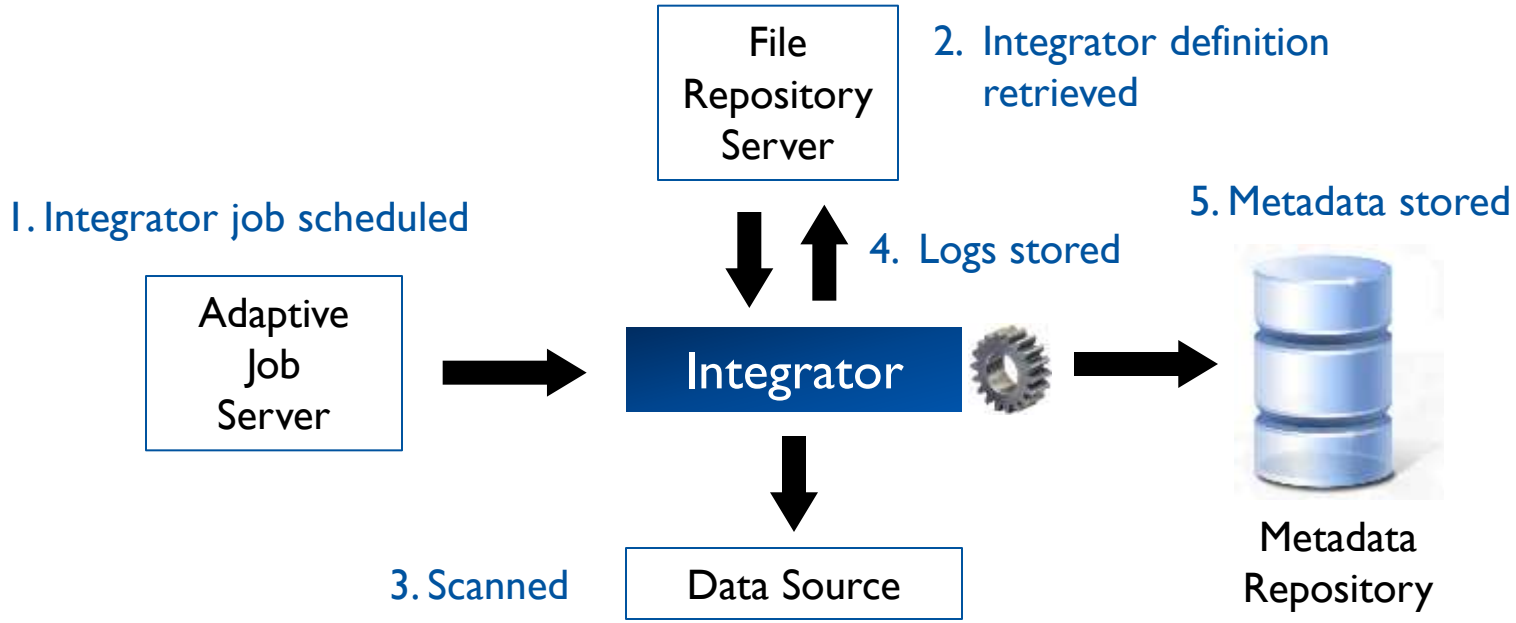


Architecture: System View



Architecture: Components

- **Metadata Integrator**
  - The hardest working component
  - Collects information from a variety of sources
  - Stores that metadata in a database repository



Architecture: Components

- Metadata Integrator, cont'd
  - Many types of sources can be scanned

SAP BusinessObjects
BO Enterprise
Data Services
Data Federator
Business Warehouse (BW)

Databases
Oracle
MS SQL Server
DB2
Teradata
MySQL

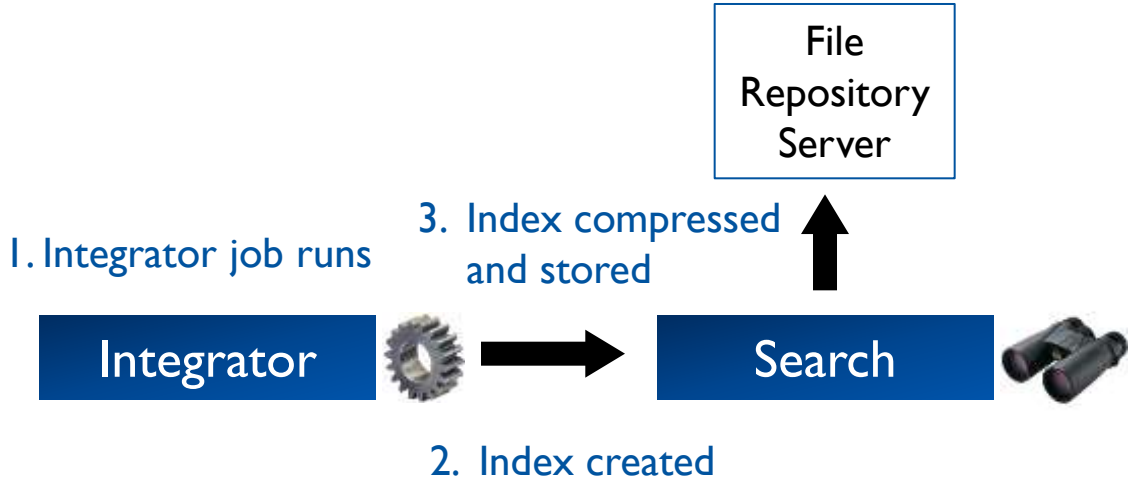
MITI Bridge (Included)
COBOL Copybook
COOL
Cognos
Datastage
Embarcadero
Erwin
Hyperion Application Builder / Essbase Integration Services
Informatica Metadata Manager, PowerCenter
Infosphere
Lotus Notes

CA Model Manager
Oracle BI
Oracle Data Integrator / Designer / Warehouse Builder
Silverrun-RDM
Sybase PowerDesigner
… and many more

Specifications
OMG UML
W3C XML DTD, Schema
Common Warehouse Model (CWM)

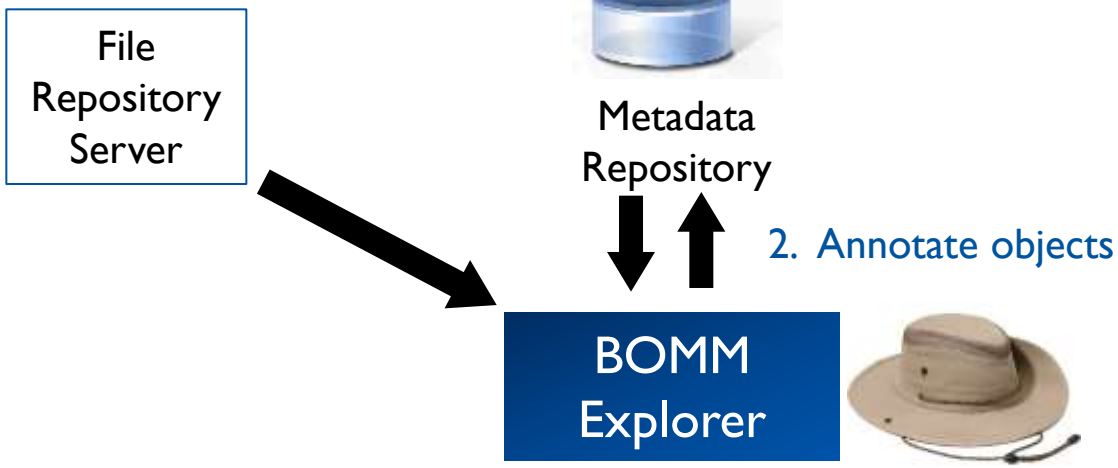
Architecture: Components

- **Search Server**
  - Locates an object regardless of source
  - Creates a search index while sources are scanned
  - Updates index with Metapedia changes



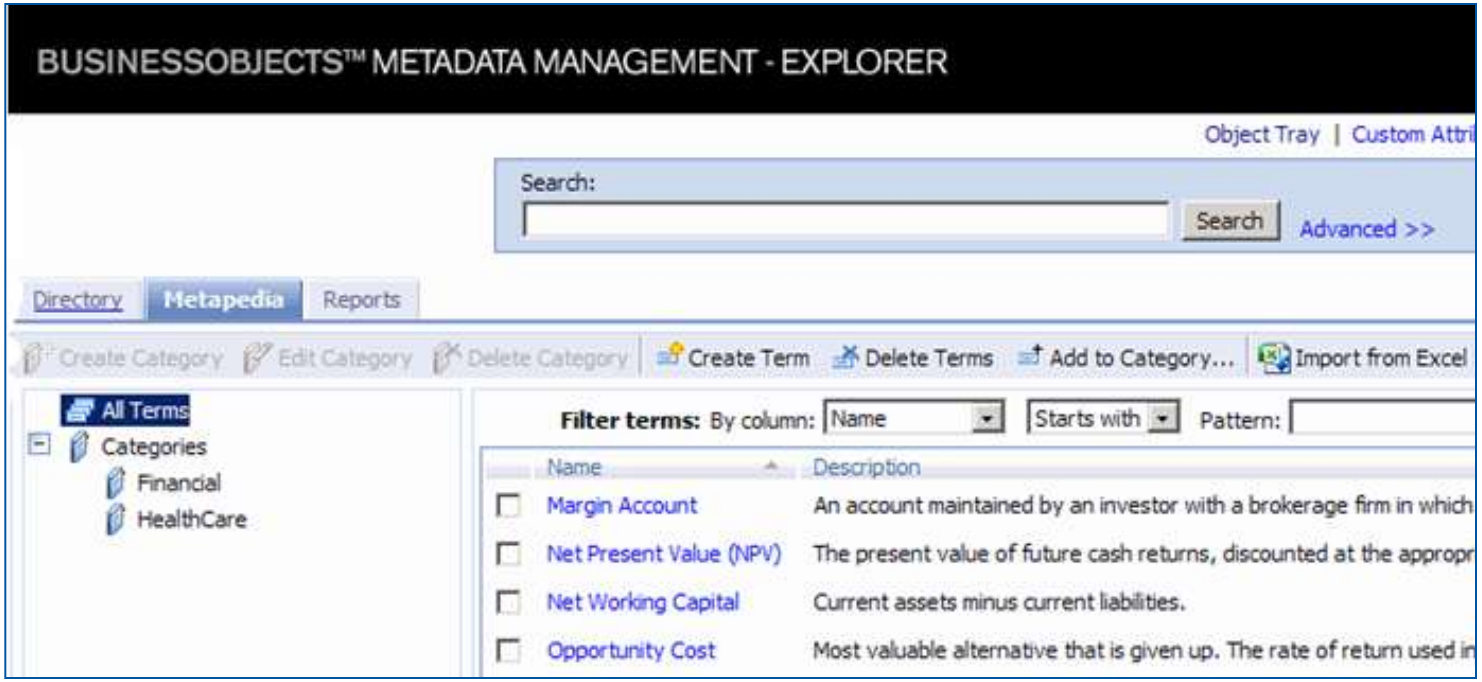
Architecture: Components

- **BOMM Explorer**
  - BO Metadata Management Explorer
  - Web application used to search and view metadata
  - Annotate scanned objects with attributes, relationships



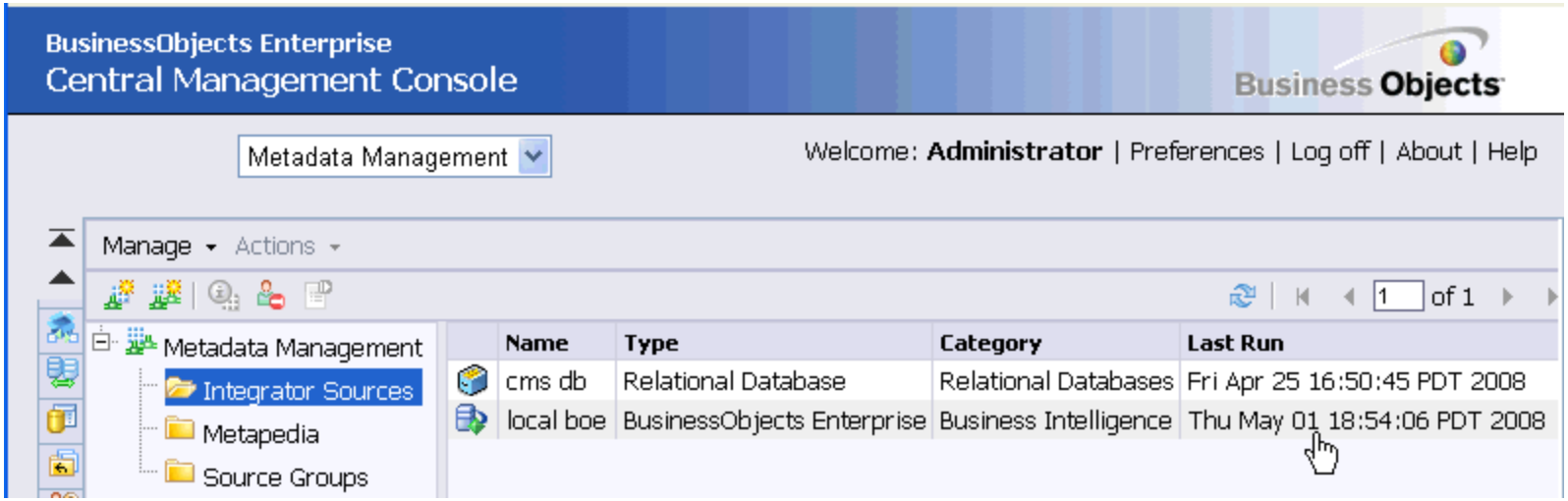
Architecture: Components

- **Metapedia**
  - Glossary of business terms
  - Terms are entered manually
  - Part of BOMM Explorer



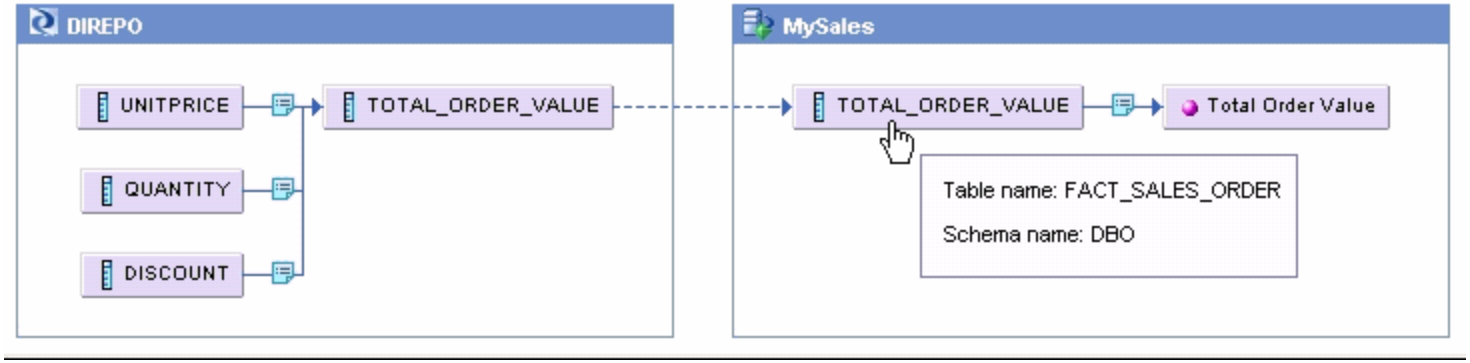
Architecture: Components

- **MM CMS Repository Objects**
  - Collection of custom JSP pages
  - Allows for administration of Metadata Management
  - Customizes the CMC by adding MM functionality



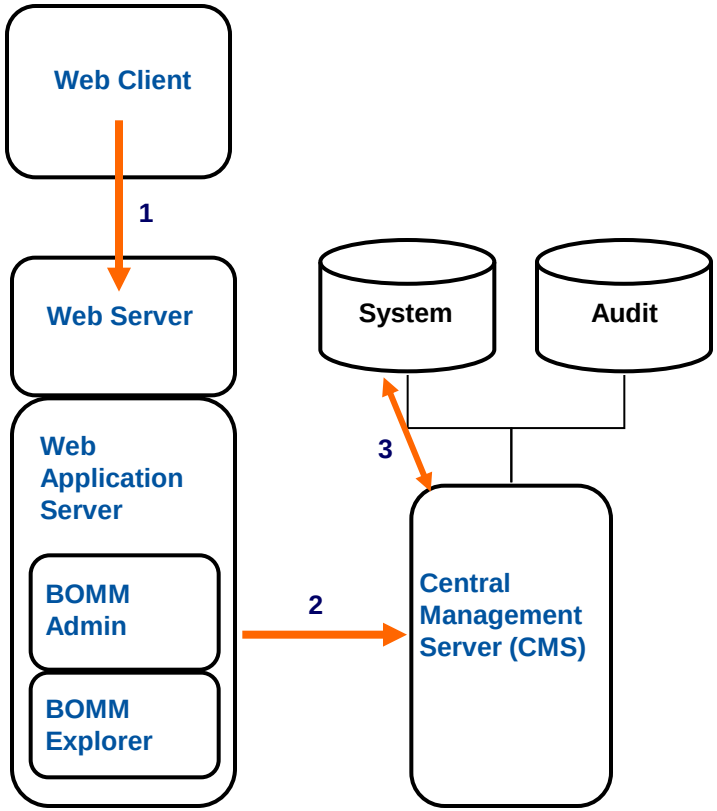
Architecture: Components

- **MM InfoView Pages**
  - Adds hyperlinks for lineage, impact analysis
  - All or nothing – cannot hide by user or via permission



Architecture: Process Flows

Scheduling an Integrator Job

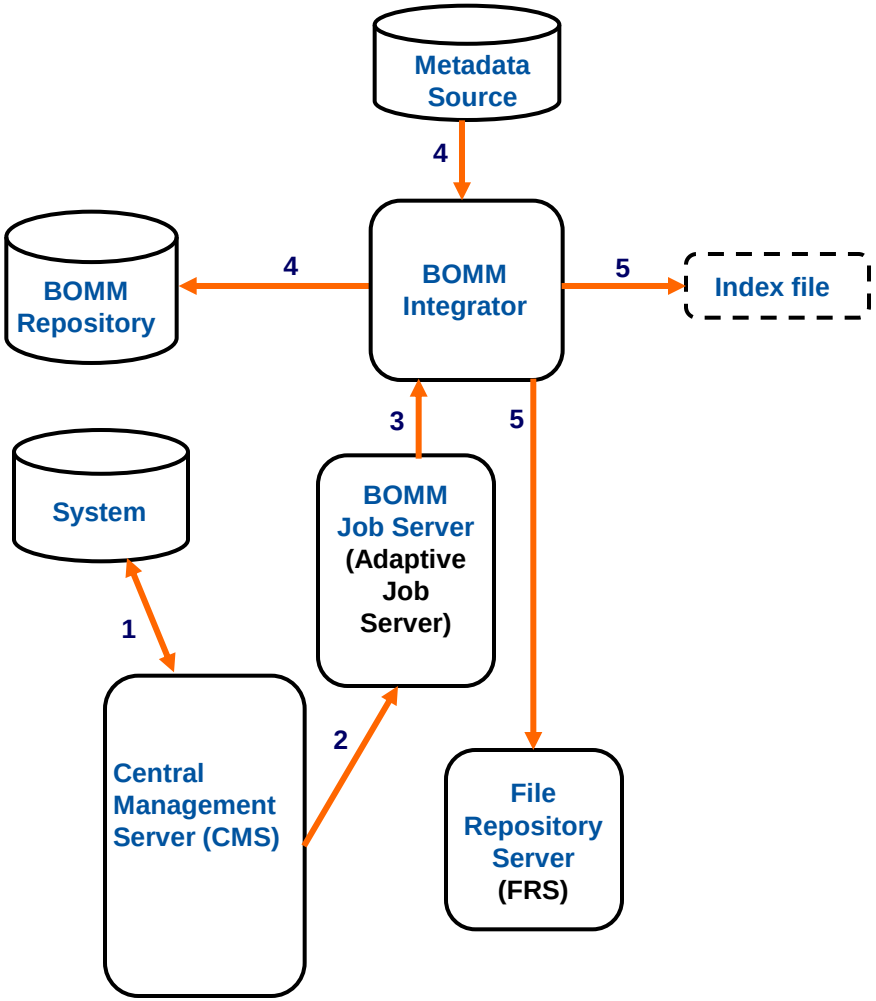


Steps

1. Client schedules an integrator job through the Central Management Console (CMC)
2. The BOMM Administrator forwards the request to the Central Management System (CMS)
3. The CMS saves the schedule in the system database and places the instance on a queue in memory

Architecture: Process Flows

Running an Integrator Job - Part I

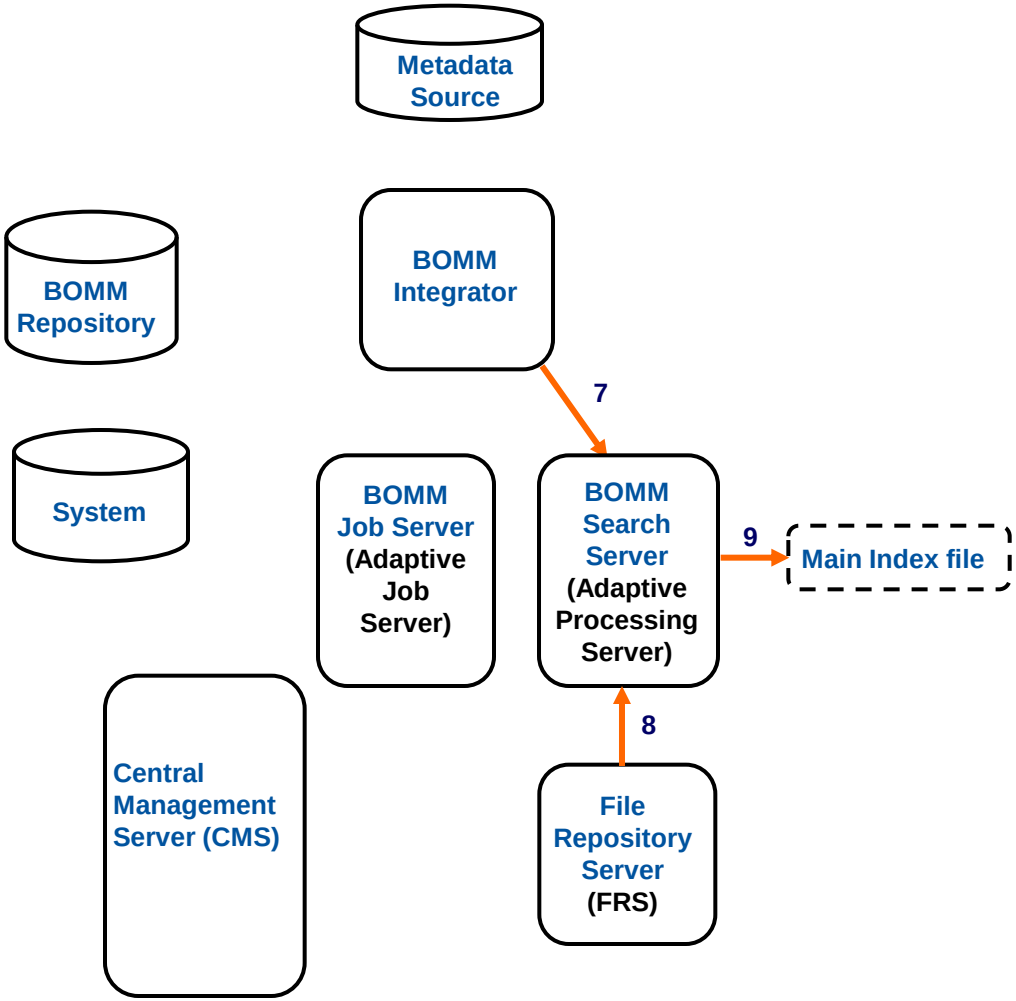


Steps

1. CMS checks for available integrator jobs. Send request to BOMM Job Server
2. CMS sends request to BOMM Job Server
3. The Job Server launches the approriate Integrator
4. The Integrator collects metadata and stores that data in the BOMM Repository
5. The Integrator also creates search indexes and stores them

Architecture: Process Flows

Running an Integrator Job - Part 2

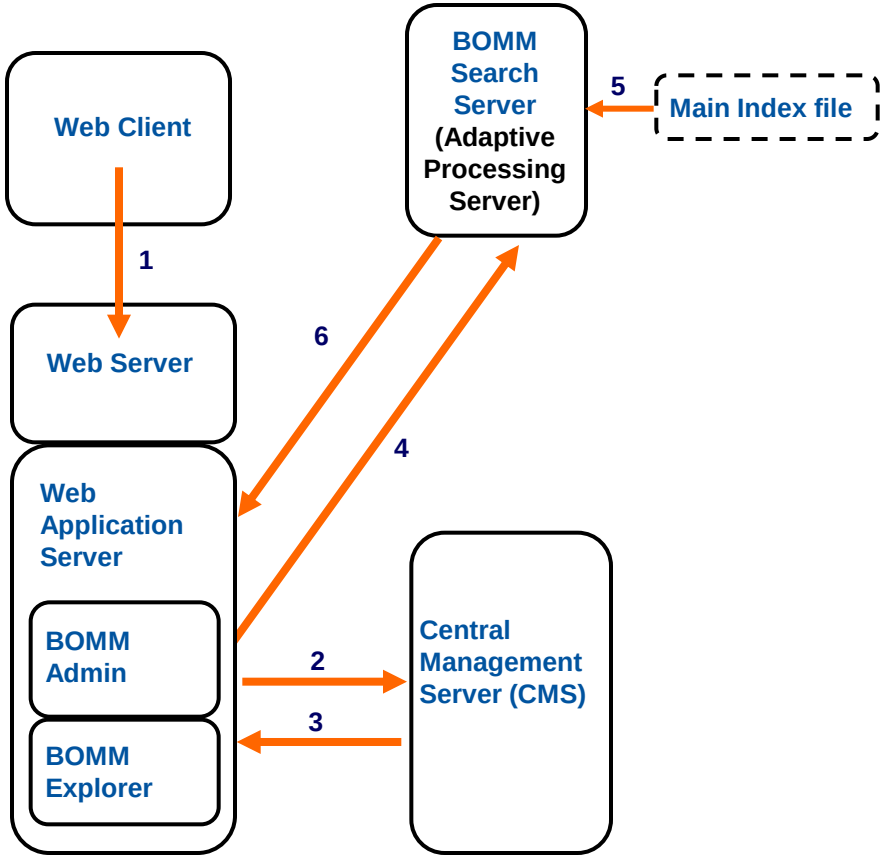


Steps

7. After storing the search index, the Integrator notifies the Search Server
8. The Search Server uploads the search index, deleting it from the File Repository Store
9. The Search Server then merges that index with the main index file

Architecture: Process Flows

Searching for Metadata

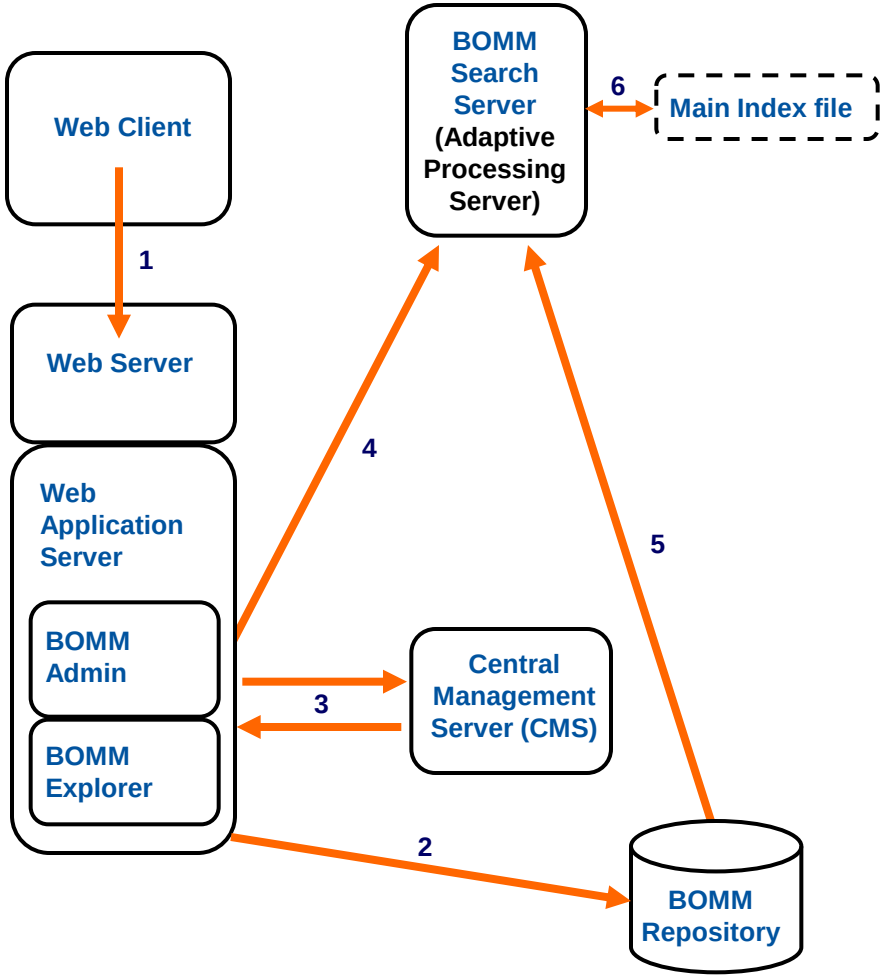


Steps

1. A customer searches for metadata through the web
2. The BOMM Explorer (web application) requests an available Search Server from the CMS
3. The CMS returns the credentials of an available server
4. The BOMM Explorer send the query to the Search Server
5. The Search Server uses its main index to retrieve the answer
6. The query result is sent back to the Explorer

Architecture: Process Flows

Adding User-Defined Metadata



Steps

1. A customer adds Metapedia data through BOMM Explorer
2. The BOMM Explorer (web application) adds that data to the Metadata Repository
3. The BOMM Explorer (web application) requests and receives an available Search Server from the CMS
4. The Explorer batches additional user requests then contacts the Search Server
5. The Search Server reads the additional user metadata
6. … and merges it with the main index

Deployment

- Most BOMM Components are services within BO Enterprise Servers
- Servers can be created with just MM-related services

Adaptive Job Server (AJS)

MM Scheduling Service

Adaptive Processing Server (APS)

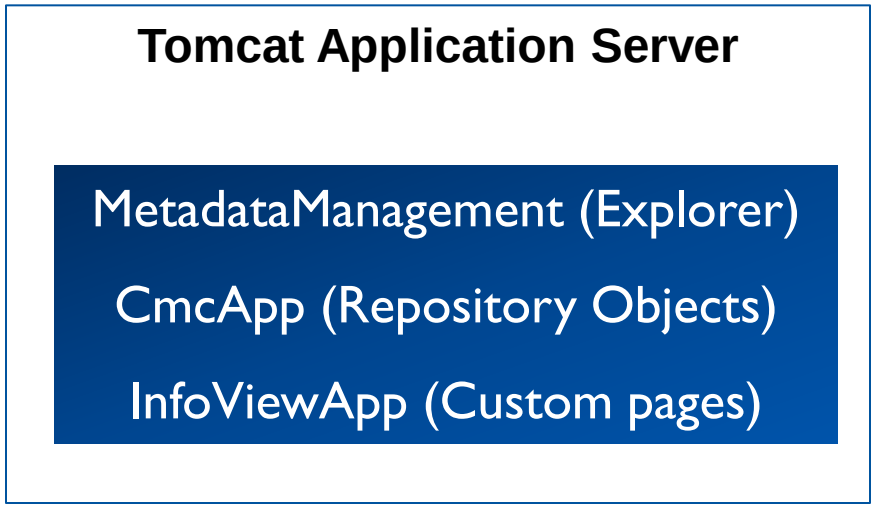
MM Integration Service

MM Search Service

MM Relationship Service

Deployment

- BOMM Explorer and Repository Objects are the exception
- Explorer is an additional web application
- CMS Objects links into the CMC
- Infoview is also customized with lineage / impact links



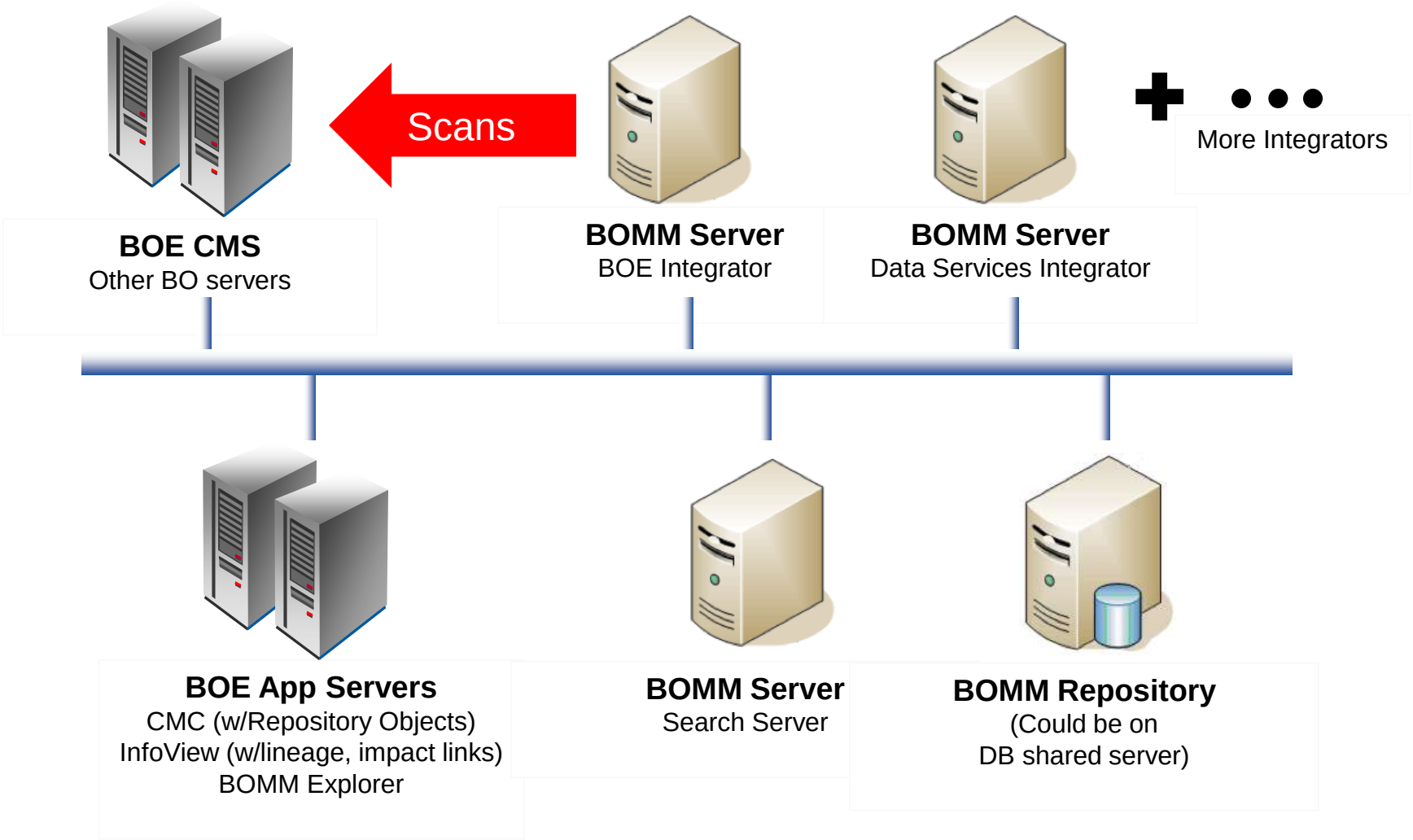
Deployment

- As you've seen, Metadata Management **REQUIRES** Windows
- This forces an all-Windows deployment
- … or a heterogeneous deployment (Windows and UNIX)
- … or something else
- That last point was a teaser. Stay tuned!

Deployment: General Rules

- Metadata Management is process-intensive
- It may require several servers to scan, store, and serve up metadata
- Do not be surprised by the next few slides

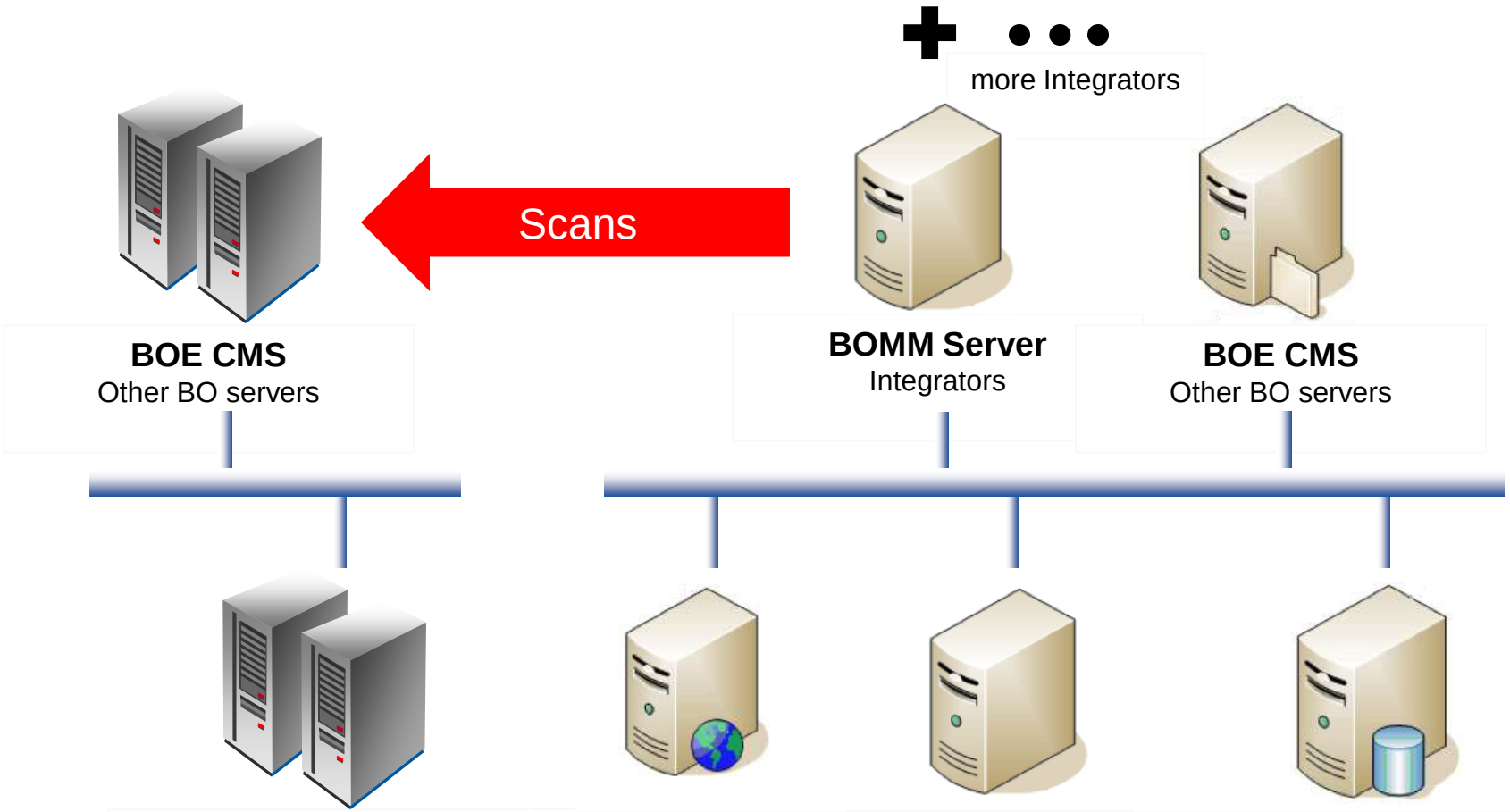
Deployment: Integrated



Deployment: Integrated Notes

- **Infoview Impact / Lineage Links**
  - Requires integrator on same cluster
  - Uses relationship service to create graphical diagrams
  - Integrators and relationship service must run on Windows
- **Heterogeneous clusters must be treated carefully**
  - Heterogeneous = Adding UNIX and Windows servers in the same cluster
  - CMC clustering is not supported
    - UNIX servers must host the CMS
    - Windows services reference UNIX-based CMS
  - Windows services should be trimmed to Metadata Management
    - Most UNIX customers want processing performed on UNIX
    - Windows services are made available as last option

Deployment: Standalone



Deployment: Standalone Notes

- **Advantages**
  - Completely Windows based (No Unix/Windows clustering)
  - Many different environments can be scanned
  - Much more scalable as a result – one MM installation suffices
- **Disadvantages**
  - Reporting accomplished through standalone MM Repository
  - Infoview impact/lineage links will not work
    - Those links depend on Windows-based MM services
    - Those services MUST be part of the cluster

Deployment: Working with Fewer Servers

- What if you can't afford a server per MM process?
- Depending on load, several servers can be combined:
  1. Search server + RDBS Integrator + Data Services Integrator
  2. BOE Integrator + Common Warehouse Model (CWM) Integrator
  3. MM Repository on same machine that hosts busiest Integrator
- Combinations are based on processing requirements
  - RDBMS and DS Integrator can be scheduled at different times (Combo 1)
  - BOE integration will use the most resources; CWM the least (Combo 2)
  - Shortest path to populate the repository (Combo 3)

Harvesting Metadata

- Collecting data is simple
  - Create a MM data source within the CMC
  - Create an Integrator job against that data source
  - Schedule the job using a recurring pattern (every day, week, …)
  - Monitor the results
- Making sense of the data collected is harder
  - The metadata integrators are very particular when scanning data
  - Pay close attention to the error logs generated
  - Some errors can be ignored, other cannot

Consuming Metadata

- Metadata can be retrieved many ways …
  - BOMM Explorer categorizes types of data as hyperlinks
  - These links can be selected to retrieve the scanned data
  - Explorer also offers several parameterized Crystal Reports
  - Integrated only – Infoview hyperlinks while viewing a report or query
  - Universe built on MM repository tables
  - That last option may take some work but is the most flexible

Demonstration

Conclusion

- Metadata may be MORE important than the data it describes
- Knowing what you've got is extremely important!
  - Create new data flows and semantic layers is costly
  - Answer administrative questions with ease
  - Find the lineage of any object used within your BI solution
  - Know the impact that each piece has (objects, reports, queries, universes, …)
- You may/will have to create additional reports
  - A relational repository database helps
  - A universe can be created to report on BOMM tables and relationships

Questions ?

Alan Mayer

alan.mayer@solidgrounded.com
214-295-6250 (office)
214-755-5771 (mobile)

SolidGround
Technologies

Thank you for participating.

Please remember to complete and return your evaluation form following this session.

For ongoing education on this area of focus, visit the Year-Round Community page at www.asug.com/yrc

SESSION 303