Enterprise Information Integration and the OMG’s MDA and MOF

Randall M. Hauch
VP Development, Chief Architect – Metadata Management

OMG's Third Workshop on UML™ for Enterprise Applications:
Model Driven Solutions for the Enterprise
October 21-24, 2002
Overview

• The Need for Enterprise Information Integration
• Model-Driven Information Integration
• Case Studies
• Summary
Enterprises and Integration?

- Each application has specific view of data
- Multiple access mechanisms
- Duplicate information
  - Which is the master?
  - Are all sources up-to-date?
  - Compatible information models?
- Unused information
- External sources
  - Are they even used?
- Impact of new or changing sources?
  - Extra costs!
  - Longer time to market!
What Information Consumers Desire

• Single source
  – One access mechanism
  – All information integrated and relatable

• Abstract where information really comes from
  – Master information always correct
  – New information “shows up” and is already integrated

• But what about the other applications?
  – Their expected views are different!
Model-Driven Information Integration

- Consumers provided with their integrated view of information
  - Their access mechanism
- Actual sources still used
  - No copies needed!
- Model the integration
  - Relationships and transformations
- Integrate in real time per the models
Modeling the Information: MOF

• Various things have to be modeled
  – Schemas, documents, mappings, transformations, etc.
  – One modeling language insufficient and impracticable

• OMG’s Meta-Object Facility (MOF)
  – Defines modeling languages (metamodels)
    • Examples: UML, CWM, EDOC, etc.
  – MOF-based modelers allow creation of models in languages natural to what’s being modeled
OMG’s Meta-Object Facility (MOF)

- meta-metamodel: boxes are meta-meta-metadata
- metamodel: boxes are meta-metadata (meta-entities)
- model: boxes are metadata (entities, instances)
- data:
  - OODBMS
  - RDBMS
  - Data Dictionary
  - XML Document
OMG’s Model-Driven Architecture™

• MDA
  – “…addresses the complete lifecycle of designing, deploying, integrating and managing applications [and] data using open standards.”*

• For information integration
  – Model the consumers and sources and the mappings/transformations between them
  – Deploy models to an execution engine
    • May require PIM->PSM transformation
  – Integrate information in real time per request
  – Manage the integration needs of the enterprise

* Taken from OMG’s Executive Overview of MDA (http://www.omg.org/mda/executive_overview.htm)
Model-Driven Information Integration

![Diagram of MetaMatrix Server and MetaBase](image)

- Online Applications
- Reporting Applications
- Web Services
- Other Applications

Tools:
- MetaBase Modeler
- MetaViewer
- QueryBuilder
- Connector Dev. Kit
- Console

- Enterprise Information Systems
- Connectors

Model and relate
Model-Driven Information Integration

Enterprise Information Consumers (EICs)
- Business Intelligence
- Call Center
- CRM
- Risk Systems

Design-Time Metadata

Virtual Database

Run-Time Metadata

Enterprise Information Systems (EISs)

1. Model Physical
2. Model Virtual
3. Relate
4. Deploy
5. Access
MetaMatrix MetaBase Modeler™

• Uses different metamodels
  – Specific terminology and constructs
  – User doesn’t need to map system to a modeling language

• Uses UML view of different metamodels
  – Shows structure with class and package diagrams
  – Displays in a consistent way
  – Valuable for discovery and management
MetaMatrix MetaBase Modeler™

Modeling Disparate Systems

- Metamodels dictate the constructs and instantiation rules
- Platform-specific terminology maintained and presented to user
MetaMatrix MetaBase Modeler™

- Model the sources
  - Relational, web, services, files, etc.
- Model the consumers
  - Relational, XML
  - Defines ‘virtual’ sources
- Define transformation and mappings
  - Selects, joins, criteria, functions, unions, etc.
  - From physical or virtual sources
- Dependency and impact analyses
MetaMatrix MetaBase Modeler™

- Model XML Schema
  - Import/export
  - Manual
- Build model of XML documents
  - Recursion, choices, etc.
  - Can be built using XML Schemas
- Map XML Document models to physical or virtual sources
- Enables accessing data as XML
Model-Driven Information Integration

• Integration layer
  – Uses OMG’s MDA and MOF technologies
  – Facilitates real-time access to integrated information
  – Centralizes integration
  – Decouples consumers from sources and builds in flexibility

• Models contain valuable information
  – Models are all relatable
    • Keys, relationships, datatypes, etc.
  – Expose to allow customers to discover the day they want to use
Case Study 1

Merrill Lynch

• Goal: Decouple consumers from sources for core data
  – Lower costs: increase data currency, eliminate copies
  – Simplify and unify access
  – Publish firm-wide catalog of available data

• With model-driven information integration
  – Real-time integration of hundreds of sources
    • Realize goal quickly with modeling effort
  – Publish/distribute data from non-XML sources using standardized XML Schemas
    • Browse/search models to find information they want
  – Improve data quality and management
    • Datatype models to help understand details and semantics of information relationships
Case Study 2

**Homeland Security**

- Goal: Share and integrate information
  - Many agencies, 3000+ data sources, real-time integration
  - Noninvasive: change little but integrate everything
  - Security/logistics/legal barrier to combining data

- With model-driven information integration
  - Can easily model and control how agency’s information exposed to others
  - Fits into architecture
    - Agencies continue to use their own sources
    - Real-time access with JDBC, SOAP, XML
Case Study 3

FBI

• Goal: Inventory all information assets to facilitate integration with Homeland Security and migration
  – Create single inventory of thousands of sources
  – Facilitate migration of sources
• With model-driven information integration
  – Improve data quality and facilitate source management
    • Discovery and searching of available information
    • Reduce duplication of information storage as well as usage
  – Facilitates migration
    • Bring "new" data online instantly to other consumers (modeling task, not programming task)
    • Some apps can use new sources, some can still use old
Case Study 4

Department of Justice

- Goal: Integrate state criminal record systems for real-time background checks
  - Single-point to query criminal systems
  - Union information from multiple sources
    - Add many more sources as time goes on
  - Simplify usability of system
- With model-driven information integration
  - Fit existing architecture
  - Modeled in one day
  - Decouples application from sources
Summary

• MDA is fundamental change
• MDA isn’t just for application development
  – Useful to define systems and their behavior; includes development systems and operational systems
  – Effective information integration requires MDA
• MOF is key to extensibility and ease-of-use for model-driven information integration
  – Information sources and consumers modeled naturally
  – Models of information sources and consumers are reusable components that can be related and exposed
• Information integration requires abstraction
  – Consumers exposed to what they need in form they want without knowing the complex integration logic
  – Changes are isolated by model-driven integration layer
  – Great increases in flexibility and return on investment (ROI)