Data Sharing Middleware Prototype (DSMP) for Information Dissemination Among Heterogeneous Sources

Mid-Year Review Meeting, Sept. 17, 2008

Hairong Qi (PI), Raghul Gunasekaran, University of Tennessee
Xiaorui Wang (co-PI), Seddik Djouadi (co-PI), UT
Oak Ridge National Laboratory*
Oracle Corporation*
Microsoft Research
Rutherford Appleton Laboratory, UK*

* Oracle, Microsoft Research, and ORNL verbal commitments for in-kind support (consulting and research software)
Contact Information

• Academia
  – Hairong Qi, 865-974-8527, hqi@utk.edu, 1508 Middle Dr., 319 Ferris Hall, EECS Department, University of Tennessee, Knoxville, TN 37996
  – Xiaorui Wang, 865-974-0627, xwang@eecs.utk.edu, 421 Ferris Hall, UT
  – Seddik Djouadi, 865-974-5447, djouadi@eecs.utk.edu, 307 Ferris Hall, UT
  – Raghul Gunasekaran, 865-385-5857, raghul@utk.edu, 536 SERF, UT
  – Ming Chen, Ying Sun, Samir Sahyoun, Ben Taylor, UT Graduate Students

• Research Laboratories
  – Frank DeNap, 865-576-8786, denapfa@ornl.gov, Oak Ridge National Laboratory, PO Box 2008, MS6085, Oak Ridge, TN 37831
  – Mallikarjun Shankar, 865-574-2704, shankarm@ornl.gov, Oak Ridge National Laboratory, PO Box 2008, MS6085, Oak Ridge, TN 37831
  – Steve Fisher, RAL, s.m.fisher@rl.ac.uk, Rutherford Appleton Laboratory (RAL), UK

• Industry, Private sectors
  – Dieter Gawlick, Ronny Fehling, Aravind Yalamanchi, 650-560-8706, {dieter.gawlick, ronny.fehling, aravind.yalamanchi}@oracle.com, Oracle Corporation
  – Vijay Dialani, Microsoft Research Center
Collaborative Team

• Academia
  – University of Tennessee
  – Vanderbilt University

• Research Laboratory
  – ORNL (Oak Ridge National Laboratory, US)
  – RAL (Rutherford Appleton Laboratory, UK)

• Industry
  – Microsoft Research
  – Oracle
Project Description

• The objective of this project is to develop a data sharing middleware that is able to handle multiple distributed data sources and dynamically changing items, and to assist in real-time information dissemination across multiple agencies for homeland security purposes.

• The ultimate target scenarios are first responders and consequence response at the urban area of Memphis (e.g., Shelby County) with stakeholders including the Fire Department, Weather Services, the E911 Operations Center, Law Enforcement Agencies, etc.
Comments and New Issues

- Clarification on the distinctive advantages of the INFOD (INFOrmation Dissemination) model
- Real-time guarantee
- New: Security issues
Landscape Assessment

**Traditional Model**

- Publisher ➔ Consumer
- Publisher ➔ Pull
- Publisher ➔ Subscriber/Consumer

**INFOD Model**

- INFOD Registry
- Subscriber (Subscription)

- Publisher ➔ Consumer

**Right Info → Right Person @ Right Time**

<table>
<thead>
<tr>
<th>Traditional Model</th>
<th>INFOD Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Service-oriented system (binding AFTER event)</td>
<td>• Establishes a framework for info flow</td>
</tr>
<tr>
<td>• Repository – data center, processing center</td>
<td>• Matches publishers and consumers based on information needs expressed through subscriptions and limited by properties</td>
</tr>
<tr>
<td>• Static system. Extending the system is difficult.</td>
<td>• Event-based system (binding BEFORE event)</td>
</tr>
<tr>
<td></td>
<td>• Registry - NOT a data (event) repository</td>
</tr>
<tr>
<td></td>
<td>• Better handling of dynamics. Extensibility is good. (Vocabulary, e.g., NIEM)</td>
</tr>
</tbody>
</table>
Existing Alerting Systems

DMIS (Disaster Management Interoperability Service)

Right Info → Right Person @ Right Time
An Information Dissemination Scenario

Notification Message differs based on the entries created and the subscriptions.

Registers as a Subscriber, defines subscription based on client necessities.

Service Providers

ABC Chemicals

Hospital

INFOD Registry Service

Weather Station

County Office

Fire Station

Police/E911

INFOD Registry Service

Notification Message

Notification Message

Notification Message

Notification Message

Notification Message

Notification Message

Notification Message

Notification Message

Notification Message

Registers as a Publisher

Registers as a Subscriber and defines subscription identifying dynamic consumers.

Matches Entities based on entries created.

Registers as a Publisher

Registers as a Publisher

Registers as a Subscriber, defines subscription identifying consumers and describes message format.

Registers as a Consumer

Registers as a Consumer

Registers as a Consumer
An Information Dissemination Scenario

Alert Msg:
Category: Fire
Severity: Extreme
Chemical Types: XXX
Action: Stop

Alert Msg:
Category: Fire
Urgency: Immediate
Action: Stop

Alert Msg:
Category: Fire
Urgency: Immediate
Action: Detour

Alert Msg:
Category: Fire
Urgency: Immediate
Action: Evacuate

Alert Msg:
Category: Fire, Chemical
Urgency: Immediate
Action: Evacuate

Alert Msg:
Category: Fire, Chemical
Severity: Extreme
Chemical Types: XXX

Alert Msg:
Category: Fire, Chemical
Urgency: Immediate
Action: Detour

Alert Msg:
Category: Fire, Chemical
Severity: Extreme
Chemical Types: XXX
An Information Dissemination Scenario

Alerting Consumers based on dynamically changing events

Alert Msg:
Category: Fire
Urgency: Immediate
Action: Stop

Alert Msg:
Category: Fire
Urgency: Immediate
Action: Detour

Alert MSG:
Category: Chemical
Urgency: Immediate
Action: Evacuate

Alert MSG:
Category: Chemical
Urgency: Immediate
Action: Stay Indoors

Alert MSG:
Category: Fire, Chemical
Urgency: Immediate
Action: Evacuate

Alert MSG:
Category: Fire, Chemical
Urgency: Immediate
Action: Detour East

Alert MSG:
Category: Fire, Chemical
Severity: Extreme, (Minor, Moderate)
Chemical Types: XXX

Alert MSG:
Category: Fire, Chemical
Severity: Extreme
Chemical Types: XXX
INFOD Resources

Subscription defines the event of interest at the publisher, the message to be generated in response to an event and helps identify consumers dynamically.

Entries characterize real world entities and define constraints identifying other entities of interest.

Also, data source entry details on the publishers information - associated to a data vocabulary.

Vocabulary instances characterize entities and constraints are evaluated on instances created.

---

Publisher

Subscriber

Consumer

INFOD Registry

Publisher Entry

Subscriber Entry

Subscription

Consumer Entry

Data Source Entry

Property Vocabulary Instance

Data Vocabulary

Property Vocabulary

Entry

Resource – not an entry

Creation of resource

Reference (EPR)

Notification (by INFOD registry)

Notification (by Publishers)
Publisher, Subscriber and Consumers are modeled as web services
Matching Entities

- In generic pub/sub models – subscription only binds publishers to consumers of information.
- In INFOD, apart from the subscription, every entity (publisher, data source, consumer, and subscriber) can define constraints that grant or limit information flow.
Control-based Real-time Metadata Matching

• Some facts
  – Hundreds even thousands of subscriptions are registered;
  – Different subscriptions have different priorities;
  – Updates arrive with unpredictable intervals;
  – Valuable Information at the Right Time (VIRT).

• Challenges
  – The execution time of reevaluating a subscription may vary significantly;
  – Reevaluating all subscriptions may cause severe system overload and unacceptable long delays;

• Goals
  – Primary Goal: Average response time of subscription reevaluation meets with the soft deadline.
  – Secondary Goal: Maximize the number of low-priority subscription reevaluation upon each update.
Controller Design and Evaluation

- A feedback controller is designed and analyzed;
- The average response time converges to the set point in spite of workload variations;
- Better control accuracy and system quality of service than two baselines;
- Theoretical analysis with extensive experiments on a physical test-bed verifies our conclusions.

“Control-based real-time metadata matching for information dissemination,” 14th IEEE Int. Conf. on Embedded and RT Comp Sys and App, Taiwan, August 2008.
INFOD - Security Considerations

- **Authentication**
  - Users establish identity with the registry
  - Publishers authenticate to the consumer
  - Web Service Security Specification

- **Authorization**
  - RBAC: Role-based access control policies
    - INFOD users are associated with Roles
    - Limit operational and resource access of users with the registry
  - Add security constraints to limit or grant mapping between entities matched through subscription

Collaborative Opportunities

• ORNL’s Shelby County Sensor Information Fusion Center
  – Provides a comprehensive target application scenario for INFOD
• ORNL’s SensorNet Group
  – Testbed setup
• Oracle
  – In-kind support
• Vanderbilt University
  – GME (Generic Modeling Environment)
• All the research findings and software developments are accessible through public domains, maintained at UT
Shelby County Fusion Center
SNAPS+POM+NOAA+INFOD

1. Shelby County Sheriff
   SNAPS II Mobile System
   - 8 chem/5 rad/5 video /1 weather sensors

2. Port of Memphis Sensors
   - 5 chem/ 1 weather sensors

3. NOAA Live Regional Weather
   - HTTPS: XML-RPC, SOAP
   - WFS, OLS,...
   - Filter, Agents
   - Access control
   - Replicated storage, image, video server

4. Distributed Wide-Area Middleware
   - Prototype and Analysis
   - Distributed querying and top-down programming
   - Policy-based data-sharing
   - Asynchronous messaging

INFO-D

Fusion Center Portal and Viewer (Web Server; Database; GIS (Google); HPAC plume modeling)

Consumer

Publisher

UT
ORNL
Industry
Oracle, MS

Application info
Emergency updates
Responder data

Contact: ORNL UT-Battelle; HT Hunter; Hunterht@ornl.gov; 865-574-6297

Managed by UT-Battelle for the U.S. Department of Energy – Supporting the Department of Homeland Security
### Project Timeline

<table>
<thead>
<tr>
<th>Tasks</th>
<th>06/07 - 05/08</th>
<th>06/08 - 05/09</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1</td>
<td>Q2</td>
</tr>
</tbody>
</table>

**Task 1: Design, Development, and Evaluation of the DSMP**
- Task 1.1: Literature survey and document study
- Task 1.2: Prototype design and development
- Task 1.3: Middleware performance evaluation and refinement

**Task 2: Demonstration of the DSMP**
- Task 2.1: Application scenario 1 - collaborative event analysis
- Task 2.2: Application scenario 2 - establishing data correlation
- Task 2.3: Application scenario 3 - first responders and consequence response at Shelby County/Memphis

- The development of the DSMP (Task 1.2) has been divided into a 4-phase implementation plan. Because of the close collaboration with Oracle, Microsoft Research, ORNL, and RAL, we are able to finish all four phases of prototype development **ahead of schedule**.
- Phase 1 - simplest scenario with a **known data vocabulary and a trivial subscription**
- Phase 2 - 2 publishers services, 2 consumer services with the addition of property vocabularies
- Phase 3 - multiple data vocabularies, publish, consumer, and subscriber services
- Phase 4 - a standard notification interface
Budget Information

• Project budget (June 5, 2007 - May 31, 2009): $400,000
• Spending as of August 31, 2008: $197,708
Commercialization Progress

- The INFOD working group is planning on merging with OGC (Open Geospatial Consortium) to get more publicity of the product on geospatial and location based services
- Potential feature to Oracle product line
IP STATUS

• Open source development
  – Will be available through sourceforge to stimulate broader participation
Achievements

- Identified first responder as the ultimate use case scenario for DHS interest
- Finished all four phases of the prototype development (ahead of schedule)
- Finished simulation of collaborative event analysis (MS thesis)
- Proposed a new plume model with promising simulation results
- Visited Shelby County Fusion Center and discussed potential integration scenario
- Papers accepted
- Students graduated
- Presentations
  - “INFOD Use Case Scenario & Demo,” Open Grid Forum (OGF), Feb 2008, Boston
DEMO
INFOD Web Application
INFOD Example Scenario: Tornado Relief Campaign

• Need
  – Monitor activities in the affected region for a short time period.
  – Track first responders in the region – ambulances, police, fire service, doctors, etc …
  – First responder needs to be aware of the resources available and contact information

• Requirements
  – A functional system instantly available that would cater to the current application
  – Instant setup by domain experts
Example Scenario: Tornado Relief Campaign

Register vocabularies

Plug and Play

Notification message on resource availability

Entities update on current status

Create Subscriptions

INFOD Registry

Publishers and consumers communicate specific needs directly

New Consumer

Affected Region

New notification messages sent on current resource availability

- Existing subscriptions applicable.
- Notification messages sent on current resource availability.

Entities register; create entries and property instances
Example Scenario: Tornado Relief Campaign

Information dissemination in 4 steps

1. Register community property and data vocabularies in the INFOD registry

2. Define subscriptions binding entities, defining events and which entity needs to be alerted on which other entities presence.

3. Entities register to the INFOD registry
   - Create entries
   - Create and update property instances

4. Notification message sent to matched entities.
Create or edit an INFODweb user

First Name
Last Name
Username
Password
Valid

Click Me Reset

Login as an INFODweb user

Username
Password

Log In Reset
INFOD Web Services

Logged in as user4

INFOD consumer
INFOD publisher
INFOD subscriber
INFOD subscription
INFOD propertyvocabulary
INFOD propertyvocabulary
INFOD datavocabulary
INFOD datasource
INFOD getmetadata
INFOD getnotificationmessage

Message Record

Ownership record

Refresh
## Property Vocabulary

### Community User Property Vocabulary

<table>
<thead>
<tr>
<th>Property Vocabulary Predicates</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification</td>
<td>Identifier for the Unit/ Person/ Organization</td>
</tr>
<tr>
<td>Name</td>
<td>Name of the Unit/ Person/ Organization</td>
</tr>
<tr>
<td>Description</td>
<td>Text Description</td>
</tr>
<tr>
<td>Location</td>
<td>Physical location information</td>
</tr>
<tr>
<td>Contact Information</td>
<td>Person to me contacted</td>
</tr>
<tr>
<td>Resource Type</td>
<td>Description of the resource</td>
</tr>
<tr>
<td>Resource Location</td>
<td>Resource Physical location information</td>
</tr>
</tbody>
</table>

Vocabulary predicates are an abstraction from NIEM (National Information Exchange Model)
# Data Vocabulary

## Alert Message Data Vocabulary (NIEM and CAP - Common Alerting Protocol)

<table>
<thead>
<tr>
<th>Data Vocabulary Predicates</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>Description between time period</td>
</tr>
<tr>
<td>Event</td>
<td>Description at a specific time</td>
</tr>
<tr>
<td>Substance</td>
<td>Description of a chemical material</td>
</tr>
<tr>
<td>Status</td>
<td>Actual, Exercise, System, Test</td>
</tr>
<tr>
<td>Message Type</td>
<td>Alert, Update, Cancel, Ack, Error</td>
</tr>
<tr>
<td>Scope</td>
<td>Public, Restricted, Private</td>
</tr>
<tr>
<td>Urgency</td>
<td>Immediate, Expected, Future, Past</td>
</tr>
<tr>
<td>Response Type</td>
<td>Shelter, Evacuate, Prepare, Execute, Monitor, Assess, None</td>
</tr>
<tr>
<td>Severity</td>
<td>Extreme, Severe, Moderate, Minor</td>
</tr>
<tr>
<td>Certainty</td>
<td>Very likely, Likely, Possible, Unlikely</td>
</tr>
<tr>
<td>Category</td>
<td>Geo, Met, Safety, Security, Rescue, Fire, Health, Env, Transport, Infra, CBRNE, Other</td>
</tr>
</tbody>
</table>
INFOD Web Services

Step 1: Register Property and Data vocabularies for a community

Logged in as user 7

<table>
<thead>
<tr>
<th>propertyvocabularyname</th>
<th>FirstResponderVocabu</th>
</tr>
</thead>
<tbody>
<tr>
<td>propertyvocabularydescription</td>
<td>Station on First responders</td>
</tr>
<tr>
<td>propertyvocabularybody</td>
<td>ads\firstResponder.xsd</td>
</tr>
<tr>
<td>propertyvocabularyentryreference [for drop and replace on]</td>
<td>Browse...</td>
</tr>
<tr>
<td>executionmode [for drop only]</td>
<td></td>
</tr>
</tbody>
</table>

Managed by UT-Battelle for the U.S. Department of Energy – Supporting the
INFOD Web Services

Step 2: Create Consumers and Publishers

**Property Constraints**

for $publishers in fn:collection("$$INFODpublisher")
where $publishers//@OrganizationSubUnitName="RedCross"
INFOD Web Services

Message Record

```xml
<?xml version="1.0" encoding="UTF-8"?>
<env:Envelope
  xmlns:env="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:ns0="http://infodregistry/WebService.wsdl/types/"
>
  <env:Body>
    <ns0:createconsumerentryResponseElement/>
    <ns0:result>5715EAFDC11D5D28E0405BA0952C158F</ns0:result>
  </env:Body>
</env:Envelope>
```
### INFOD Web Services

#### Step 3: Subscribers and Subscriptions

**Property Constraints**

```xml
for $publisher in fn:collection("$$INFODpublishers")
  where $publisher//OrganizationSubUnitName="E911Center"
for $firstresponders in fn:collection("$$INFODconsumers")
  where $firstresponders//OrganizationSubUnitName="RedCross"
```

**Data Constraints**

```xml
declare namespace $data =http://infod.firstrespondernet.com/AlertDataVocabulary;
let $msg1 := for $firstresponders
  where $data:AlertStatus = 'Actual' and
  $data:EventCategory = 'CBRNE' and
  $data:EventSeverity > 'Moderate'
  return {$data, $data:Instruction = 'Evacuate people in the region' }
let $msg2 := for $firstresponders
  where $data:capAlertStatus = 'Actual' and
  $data:EventCategory = 'Fire' and
  $data:EventSeverity > 'Moderate'
  return { $data:Substance, $data:Volume, $data:EventCategory }
```

**Dynamic Consumer Constraint**

```xml
for $firstresponders
  where $firstresponders//OrganizationSubUnitName="Police"
  return msg1
for $firstresponders
  where $firstresponders//OrganizationSubUnitName="FireService"
  return msg2
```