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

Mid-Year Review Meeting, Sept. 6, 2007

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

* Oracle, IBM 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@ece.utk.edu, 421 Ferris Hall, UT
  – Seddik Djouadi, 865-974-5447, djouadi@ece.utk.edu, 307 Ferris Hall, UT
  – Raghul Gunasekaran, 865-385-5857, raghul@utk.edu, 536 SERF, UT
  – Ming Chen, Ying Sun, Samir Sahyoun, 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, 408-927-1406, vdialani@us.ibm.com, IBM Almaden Research Center
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.
Landscape Assessment

- Traditional information dissemination
  - Using a centralized server where data is stored and queried - single point of failure.
  - Having data logged and matched at a data center adds delay.
  - With increasing number of publishers and consumers, the efficiency of the matching algorithm worsen - do not scale.

- The proposed information dissemination (INFOD) model
  - Data are kept distributively at the local storage. The centralized server is only responsible of matching communities of interests and helping identify what data should be sent to whom and when to send.
  - Publishers determine consumers dynamically based on data constraints.
  - Information is delivered directly to the consumers. Information flow changes dynamically as the condition or state of the publishers and consumers of data changes.

An Information Dissemination Scenario

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

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

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

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

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

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

Alert Msg:
Category: Fire
Severity: Expected
Action: Evacuate

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

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

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

Sends CAP messages

Matches Entities based on entries created

Registers as a Publisher

Registers as a Subscriber, defines subscription identifying dynamic consumers

Registers as a Subscriber, defines subscription identifying consumers and describes message format.
Project Content - INFOD Resources

Subscriber and Consumer are modeled as independent entities. Subscriptions are created by a subscriber, targeting a subset of consumer.

Subscription defines what constitutes an event at the publisher and the message that needs to be generated in response to these events.

Entities identify a community of interest, by defining property constraints.

Structured Information Model, user communities are identified by property and data vocabularies.

Data sources, indicate the variety of information that publisher’s can provide.
Project Content - INFOD Implementation

• Targeting at application scenarios such as first responders and consequence responses
• Vocabulary
  – NIEM (National Information Exchange Model)
  – CAP (Common Alerting Protocol)
• Implementation on Oracle Database 10g

INFOD Registry Service

Subscription Reasoning Component

Registry Manager Component
  - Vocabulary Manager
  - INFOD Entry’s Manager
  - Meta Data Manager

ORACLE PL/SQL

Test Environment

https://forge.gridforum.org/sf/projects/infod-wg
Project Content - INFOD Implementation

Step 1: Vocabulary Registration

**Property Vocabulary (NIEM)**

```xml
<infod:CreatePropertyVocabulary>
  <infod:VocabularyName>SensornetPropertyVocabulary</infod:VocabularyName>
  <infod:VocabularyBody>
    <xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.w3.org/INFOD/Entity">
      <xsd:element name="Organization" type="u:organizationType" nillable="true"/>
      <xsd:complexType name="organizationType">
        <xsd:sequence>
          <xsd:element name="OrganizationName" type="u:TextType"/>
          <xsd:element name="OrganizationID" type="u:IDType"/>
          <xsd:element name="OrganizationDescriptionText"/>
          <xsd:element name="OrganizationLocation"/>
          <xsd:element name="PrimaryContactInformation"/>
          <xsd:element name="OrganizationStatusText" type="u:Text"/>
          <xsd:element name="Classification" type="u:ClassificationType" nillable="true"/>
        </xsd:sequence>
      </xsd:complexType>
      <xsd:element name="Classification" type="u:ClassificationType">
        <xsd:complexType>
          <xsd:sequence>
            <xsd:element name="ClassificationAgencyName"/>
            <xsd:element name="ClassificationDerivationText"/>
          </xsd:sequence>
        </xsd:complexType>
      </xsd:element>
    </xsd:schema>
  </infod:VocabularyBody>
</infod:CreatePropertyVocabulary>
```

**Data Vocabulary (NIEM & CAP)**

```xml
<infod:CreateDataVocabulary>
  <infod:VocabularyName>SensorDataVocabulary</infod:VocabularyName>
  <infod:VocabularyBody>
    <xsd:element name="Activity" type="u:ActivityType" nillable="true"/>
    <xsd:element name="Event" type="u:eventType" nillable="true"/>
    <xsd:element name="capAlertID" type="xsd:string" minOccurs="0" maxOccurs="unbounded"/>
    <xsd:element name="capAlertUrgency" minOccurs="0" maxOccurs="unbounded">
      <xsd:simpleType>
        <xsd:restriction base="xsd:string">
          <xsd:enumeration value="Future"/>
          <xsd:enumeration value="Past"/>
        </xsd:restriction>
      </xsd:simpleType>
    </xsd:element>
    <xsd:element name="capAlertSeverity" minOccurs="0" maxOccurs="unbounded">
      <xsd:simpleType>
        <xsd:restriction base="xsd:string">
          <xsd:enumeration value="Extreme"/>
          <xsd:enumeration value="Severe"/>
          <xsd:enumeration value="Moderate"/>
        </xsd:restriction>
      </xsd:simpleType>
    </xsd:element>
    <xsd:element name="capAlertCategory" minOccurs="0" maxOccurs="unbounded">
      <xsd:simpleType>
        <xsd:restriction base="xsd:string">
          <xsd:enumeration value="Security"/>
          <xsd:enumeration value="Rescue"/>
          <xsd:enumeration value="Fire"/>
        </xsd:restriction>
      </xsd:simpleType>
    </xsd:element>
  </infod:VocabularyBody>
</infod:CreateDataVocabulary>
```

* Only a subset of the adopted schema is shown, to get an understanding of the vocabularies and their content.

Managed by UT-Battelle for the U.S. Department of Energy – Supporting the Department of Homeland Security
Step 2: INFOD Entry's

**Publisher**

- **Registers**
  - `<infod:CreatePublisherEntry>`
  - `<infod:WSReference>http://../Node1</infod:WSReference>`
  - `<infod:PublisherName>Node1</infod:PublisherName>`
  - `<infod:PublisherDescription>Chemical Sensor Node</infod:PublisherDescription>`

- **Property Vocabulary Instance**
  - `<infod:CreatePropertyVocabularyInstance>`
  - `<infod:EntryReference>Node1EPR</infod:EntryReference>`
  - `<infod:PropertyVocabularyReference>….</infod:PropertyVocabularyReference>`

- **Data Source Entry**
  - `<infod:CreateDataSourceEntry>`
  - `<infod:DataSourceEntryName>Node1DataSourceEntry</infod:DataSourceEntryName>`
  - `<infod:DataSourceEntryReference>Node1</infod:DataSourceEntryReference>`

**Consumer**

- **Registers**
  - `<infod:CreateConsumerEntry>`
  - `<infod:WSReference>http://...AlertListener1</infod:WSReference>`
  - `<infod:ConsumerName>AlertListener1</infod:ConsumerName>`
  - `<infod:ConsumerDescription>Monitoring Chemical Hazards</infod:ConsumerDescription>`
  - `<infod:ConsumerPropertyConstraint>`
    - `<capAlertCategory>Env,Health</capAlertCategory>`
  - `<infod:CreateConsumerPropertyConstraint>`

- **Property Vocabulary Instance**
  - `<infod:CreatePropertyVocabularyInstance>`
  - `<infod:EntryReference>AlertListener1EPR</infod:EntryReference>`
  - `<infod:PropertyVocabularyReference>EPR</infod:PropertyVocabularyReference>`

- **Data Source Entry**
  - `<infod:CreateDataSourceEntry>`
  - `<infod:DataSourceEntryName>Node1DataSourceEntry</infod:DataSourceEntryName>`
  - `<infod:DataSourceEntryReference>Node1</infod:DataSourceEntryReference>`

**Subscriber**

- **Registers**
  - `<infod:CreateSubscriberEntry>`
  - `<infod:SubscriberName>AlertListener1</infod:SubscriberName>`
  - `<infod:SubscriberDescription>AlertListener1</infod:SubscriberDescription>`

- **Create Subscription**
  - `<infod:CreateSubscription>`
  - `<infod:SubscriptionName>GetNodeChemicalSensorAlerts</infod:SubscriptionName>`
  - `<infod:SubscriptionDescription>Listening for Node and Chemical Sensor Alerts</infod:SubscriptionDescription>`
  - `<infod:SubscriberReference>`
  - `<wса:Address>http://...</wса:Address>`
  - `<infod:DynamicConsumerConstraint>`
  - `<infod:DataConstraint>`
    - `<infod:DynamicConsumerConstraint>`
  - `<infod:DataConstraint>`
  - `<infod:DynamicConsumerConstraint>`
  - `<infod:DataConstraint>`
Project Content - INFOD Implementation

Step 3: Subscription Evaluation (Matching)
- Triggered whenever a new or a change in Publisher/Consumer/Subscriber/Subscription’s entry and their propertyVocabularyInstances is recorded in the registry.
- Evaluates the subscription, to find a set of matching publishers and consumers from a community of registered INFOD entities, with the knowledge of the data source.
- Further for every publisher from the subscription evaluation, matching consumers are found satisfying both its constraints and the consumer’s constraint.
- Similarly, matching publishers are found for every consumer.

Step 4: Notification
- On having matched publishers and consumer constraints, notification message is sent.

Notification to Publisher
<infod:PublisherNotification>
<infod:SubscriptionReference>
<wsa:Address>http://…AlertListener1EPR</wsa:Address>
</infod:SubscriptionReference>
<infod:ConsumerEntryReference>
<wsa:Address>http://…AlertListener1EPR</wsa:Address>
</infod:ConsumerEntryReference>
<infod:DataConstraint>
http://…vocabEPR.Severity=‘Moderate’ AND http://…vocabEPR.capAlertCategory=“Env”
</infod:DataConstraint>
<infod:DynamicConsumerConstraint>………..</infod:DynamicConsumerConstraint>
</infod:PublisherNotification>

Notification to Consumer
<infod:ConsumerNotification>
<infod:SubscriptionReference>
<wsa:Address>http://…AlertListener1EPR</wsa:Address>
</infod:SubscriptionReference>
<infod:PublisherEntryReference>
<wsa:Address>http://…Node1EPR</wsa:Address>
</infod:PublisherEntryReference>
<infod:ConsumerNotification>
<infod:DataConstraint>
</infod:DataConstraint>
<infod:DynamicConsumerConstraint>………..</infod:DynamicConsumerConstraint>
</infod:ConsumerNotification>

Notification to Subscriber
<infod:SubscriberNotification>
<infod:SubscriptionReference>
<wsa:Address>http://…EPR</wsa:Address>
</infod:SubscriptionReference>
<infod:PublisherEntryReference>
<wsa:Address>http://…EPR</wsa:Address>
</infod:PublisherEntryReference>
<infod:SubscriberNotification>
<infod:DataConstraint>
</infod:DataConstraint>
<infod:DynamicConsumerConstraint>………..</infod:DynamicConsumerConstraint>
</infod:SubscriberNotification>
Project Content - INFOD Implementation

Step 5: Notification Message from Publisher to Consumer

<infod:Notify>
  <infod:NotificationMessage>
    <infod:SubscriptionReference>
      <wsa:Address>
        http://infod.sensornetwork.com/subscriptionEPR/AlertListener1
      </wsa:Address>
    </infod:SubscriptionReference>
    <infod:PublisherReference>
      <wsa:Address>
        http://infod.sensornet.com/PublisherInfodEPR.node1
      </wsa:Address>
    </infod:PublisherReference>
    <infod:Message>
      Description of the event
    </infod:Message>
  </infod:NotificationMessage>
</infod:Notify>
Project Content - Phased Implementation

- Phased implementation approach
  - Start with concrete example vocabulary
  - Simple services prototype to more elaborate capabilities
  - Fixed candidate vocabulary (e.g., NIEM) to general vocabularies
  - Simple matching to extended matching

- Phase 1 - simplest scenario
  - 1 known data vocabulary
  - 1 Publisher and 1 Consumer service added to the Registry
  - A simple subscription with just data constraints and no property constraints.
  - Very simple data is published by the publisher.

- Phase 2
  - 2 Publisher and 2 Consumer services
  - The addition of Property vocabularies and instances of these properties for the publishers and consumers.
  - A subscription which now contains some simple property constraints as well as data constraints

- Phase 3
  - Multiple data vocabularies.
  - Multiple Publisher, Consumer, and Subscriber services added to the Registry.
  - More complex subscriptions with data constraints and property constraints.
  - More complex data published by the publishers.

- Phase 4
  - A standard notification interface such as WSN Notify.
Collaborative Opportunities

• ORNL’s Shelby County Sensor Information Fusion Center project provides a comprehensive target application scenario for INFOD.
  – Help establish common understanding (finding matches) between homeland security response agencies
  – Facilitate the sharing of federated data for first responders and consequence responses

• All the research findings and software developments will be accessible through public domains, maintained at UT
### 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>
<tr>
<td>Task 1: Design, Development, and Evaluation of the DSMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 1.1: Literature survey and document study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 1.2: Prototype design and development</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 1.3: Middleware performance evaluation and refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 2: Demonstration of the DSMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 2.1: Application scenario 1 - collaborative event analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 2.2: Application scenario 2 - establishing data correlation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 2.3: Application scenario 3 - first responders and consequence response at Shelby County/Memphis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The development of the DSMP (Task 1.2) has been divided into a 4-phase implementation plan. Because of the close collaboration with Oracle, IBM, ORNL, and RAL, we are able to finish Phase I ahead of time. With the phase 1 prototype accomplished, the application development tasks can be carried out in parallel to the DSMP development.
- We have also set up a testbed at ORNL for related projects. With that experience, the university can set up a similar testbed for independent evaluation.
Budget Information

• Project budget (June 5, 2007 - May 31, 2009): $400,000
• Spending as of Aug. 1, 2007: $21,015
Commercialization Progress

- Since the project has just started on June 5, 2007, there has not been any progress made toward commercialization.
IP STATUS

• No disclosures have been filed so far.
Summary & Conclusions

- Completed tasks by plan
  - Literature survey on existing information dissemination developments
  - Identified unique features in the proposed INFOD model
  - Identified the ultimate target application scenario (first responders and consequence response at the urban area of Memphis)

- Accomplishments ahead of schedule
  - With the in-kind support from our private sectors and international collaborations, we are able to finish the Phase I implementation of the INFOD prototype
  - A testbed has been set up at ORNL to facilitate application deployment and software evaluation

- Plan of work for the next six months
  - Finish Phase II of the software development
  - Work on the first application scenario of the INFOD prototype (collaborative event analysis)