Vijay Dialani, IBM Corporation
Abdeslem Djaoui, Rutherford Appleton Laboratory
Steve Fisher, Rutherford Appleton Laboratory
Dieter Gawlick, Oracle Corporation
Christopher Kantarjiev, Oracle Corporation
Cecile Madsen, IBM Corporation
Susan Malaika, IBM Corporation
Shailendra Mishra, Oracle Corporation
March 14, 2005

Motivating Scenarios for INFOD (Information Dissemination)

Status of This Memo

This memo provides information to the Grid community motivating scenarios for the Information Dissemination working group. It does not define any standards or technical recommendations. Distribution is unlimited.

Copyright Notice

Abstract

Grids applications must be able to access, manage and share data in a timely fashion at a very large scale and potentially across organizational boundaries. Challenges arise from the number and variety of involved parties, the size of the data, the number of recipients (consumers) of each data element and the geographical distance. Critical data must be protected while providing efficient access to and distribution of these data with the required Quality Of Service (QOS).

"Information Dissemination" (INFOD) addresses these requirements. INFOD supports timely and efficient data dissemination of customized information mainly based on consumer's needs. It is assumed that the data to be disseminated are created in response to an event. Therefore, a key objective is to disseminate data as soon as it becomes available

INFOD (Information Dissemination) is a working group in GGF focusing on the publication, dissemination and consumption of information. This document explains what INFOD is and presents a selection of both commercial and scientific scenarios. Requirements are inferred from these scenarios and are related to the specification.

Contents

Vijay Dialani, IBM Corporation
Abdeslem Djaoui, Rutherford Appleton Laboratory
Steve Fisher, Rutherford Appleton Laboratory
Dieter Gawlick, Oracle Corporation
Christopher Kantarjiev, Oracle Corporation
Cecile Madsen, IBM Corporation
Susan Malaika, IBM Corporation
Shailendra Mishra, Oracle Corporation
March 14, 2005

1. Introduction

This document first explains the purpose of INFOD and then provides a set of Use Cases to demonstrate the utility of this approach.

Messaging systems naturally start with the message. This approach requires that the publisher of that message predefines all useful messages. The INFOD approach considers the events that can cause messages to be published, and even goes further to include state changes that gave rise to the event. The state change may only be the passage of time. Those interested in the information, which will be delivered as messages, may subscribe to that information in terms of states, events or messages.

The INFOD approach requires ways to define states, events and messages. Now various disciplines have defined, within their own universe of discourse, a way of describing information that they find meaningful. We refer to these styles of information descriptions as vocabularies. Often, vocabularies are defined through XML schemas (XSD, XDR etc), or DTDs or RELAX NG. It is helpful to consider vocabularies defined in other terms, particularly through relational schemas, because much state information, particularly state that changes, is stored in relational databases. Information described through XML schemas and through relational schemas has an extensive set of standardized manipulation languages, e.g., XPath, XSLT, XQuery (standardization underway), SQL, SQL/XML. These general purpose languages are helpful for subscribers and consumers to recognize states, to identify state changes, to define events, to filter and to transform messages in a unified way.

If the publisher of the information provides a vocabulary or vocabularies defining the state then a subscriber to that information is able to define events and the messages produced for consumers.

In this document we refer to the *Components*, meaning a typical implementation of a web service with one or more interfaces. The INFOD specification is written in terms of interfaces - but in this document we suggest practical groupings of these interfaces. The INFOD components mentioned so far are publishers and consumers. Clients of INFOD are known as actors. Some components and actors require their own identity management for security reasons. These are known as principals and they may assume multiple roles. The INFOD principals include publishers, consumers, subscriptions, disseminators, and registration managers.

Information originates from a publisher; this may simply be an actor in which case it is only able to publish messages without being able to be directly influenced by subscribers to those messages. Publishers may implement the consumer interface allowing them to receive control messages. Note that the same consumer interface used by INFOD consumers to receive INFOD messages, is being exposed by publishers to receive control messages.

A publisher may push messages directly to a consumer but more often information is published to the consumer interface of an INFOD disseminator. From there it may be passed through a network of disseminators to the consumer interface of one or more consumers. Depending on the subscription, messages may be delivered directly to a consumer or a single notification message may be sent saying that there is information in the disseminator that the consumer may pull. In the latter case the disseminator will not send another notification message to that consumer, until the consumer has pulled some information from the disseminator.

Communication between disseminators may be via the consumer interface or it may go through a propagator interface with more features. Note that it may be possible to eliminate the propagator interface in favour of a consumer interface with appropriate QoS properties.

Another important component is the INFOD registration manager. The INFOD registration manager provides a service to manage registration of all principals. It has multiple interfaces one for each type of registration, however we think that any practical implementation would co-locate these interfaces, as it is often necessary to combine information registered via the different interfaces. The registrations stored are for publishers, disseminators and consumers and also subscriptions for messages are stored there. The registration manager is not passive but must also send notifications. For example, if a publisher is registered, a notification must be sent to a consumer that is potentially interested in that information.

With these ingredients we are able to deal with:

- *Explicit publishing:* The producer publishes in all cases irrespective of the existence of a subscriber. The producer identifies the consumers and pushes to a set of consumers who have not necessarily subscribed
- *On Demand publishing:* The producer publishes and then the disseminator checks if there is a subscriber before dissemination. If there is no subscriber then the disseminator discards the notification
- *Implicit publishing:* The producer publishes only if there is a subscriber

1.1 Acknowledgements
We would like to thank the following people:

- Members of the retail industry services team in IBM for the retail use case
- Members of the eDiamond project especially Alan Knox of IBM and the eDiamond team at Oxford University, for sharing their ideas and materials
- Thomas Soddemann from Deisa for a replication use case

2. Retail Industry Use Case

2.1 Summary
In this use case, we consider a typical configuration where some data is held at headquarters and other data is held at individual stores. Reasons for keeping information in multiple locations are usually performance or availability (e.g., being able to continue running the stores in the face of network problems). Quite often information is held in operational systems used for running the business, e.g., buying and selling, or in data warehouses, used for analyzing how the business is performing. Industry agreed data models can be used for storing the data such as ARTS [1]

2.2 Customers
The goal in this use case is to automate the information dissemination across a distributed retail system, so the movement of data occurs in response to events or due to time passing, and not as a result of specific human activity.
Actors include:

- Software execution schedulers
- Information providers, e.g., an updated retail price list
- Human requesters such as customers, headquarters staff, or store workers (in ad hoc case)

2.3 Scenarios
In this section the scenarios are described in terms of patterns for information dissemination

- **Bulk:** This activity usually runs according to a schedule or policy moving information from headquarters to one or more stores, often in large quantities. Headquarters or the store can initiate the activity. On receipt the information often requires some kind of processing. Information transmitted in this way could be new releases of software products that may require installation on receipt, or retail price lists which on receipt may be transformed in various ways.
- **Trickle:** This activity is typically event driven, e.g., every 50 sales push the sales information from a store to headquarters, or to other stores. The information volumes may not be significant and the transmissions may not be performance sensitive. Information transmitted in this way could include sales transactions or promotion redemptions.

- **Ad Hoc:** This activity typically occurs when people run applications that issue queries. Replies are expected in real time, so the activity is performance sensitive. The information volume transmitted may be very small and can occur across stores or between stores and headquarters, e.g., to check if an item is in stock at a nearby store, to check eligibility for promotion, or to review customer shopping-lists.
- **Data Replication:** This activity is performed where multiple copies of data are required, e.g., at all the stores, or at headquarters and at the stores. It occurs between databases and can be one-way or bi-directional (headquarters to stores, stores to headquarters), and is scheduled or policy driven. The kind of data that may be replicated includes the product catalog, and promotion information.
- There are various models for data replication in the retail industry. Examples include:
    - *Synchronize*: Synchronization is the process which ensures that structured data in either real-time or near real-time data stores is consistent and as near to current as possible. It can be one way or bi-directional, depending if updates are permitted at all locations
    - *Reconcile*: Reconciliation occurs between operational data stores and data warehouses. It is the process by which data originating in different data sources is cleansed, made semantically and temporally consistent, and often extended with temporal data to enable storing of history, before being stored in warehouse. By definition, reconciliation is a one-way process from operational datasets a warehouse
        - *Derive:* Derivation occurs when the results of queries on a warehouse are stored, e.g., the joining of data from a number of warehouse relational tables, and its summarization and sub-setting. The information dissemination is usually one way and does not include cleansing

2.4 Security

2.5 Performance

3. Medical Screening Use Case

3.1 Summary

The eDiaMoND project [2] is prototyping a Grid to support digital breast imaging in the UK. Mirada Solutions will develop a state-of-the-art breast imaging workstation as a client to the eDiaMoND Grid.

In addition to developing a mammography Grid, the eDiaMoND project will demonstrate its value by addressing breast screening, radiologist training and epidemiology applications. The Grid is also intended to form the basis for research in the areas of medical imaging, computer-aided detection and diagnosis, and in the development of Grid infrastructures.

3.2 Customers

3.2.1 Healthcare

Specifically X-Ray Mammography, but the middleware involved deals with DICOM format medical images and can be generalized to support other medical imaging applications.

Users are Radiographers (data capture and quality control), Clinic Administrators (controlling workflow) and Radiologists (clinical analysis of images).

3.2.2 Research

A collection of digital X-Ray mammography images provides a significant resource for research. It should allow properly authorized researchers access to images to support research and development of data mining and image processing techniques for detection and diagnosis of breast cancer.

Additionally the repository is of value in epidemiological research. It should be capable of being joined to some other study specific data resource for the duration of a particular epidemiological study.

3.3 Scenarios

3.3.1 Screening

A Radiographer takes X-Ray mammograms and scans the resulting films to produce DICOM files that are uploaded to the Grid. At some later time, an Administrator creates work lists of images that need reading by Radiologists. These are managed by a workflow component. Later still, a Radiologist attends a screening clinic to "read" a batch of images. The Radiologist workstation retrieves work lists and fetches the indicated images from the Grid. The list of images may include prior images as well as those recently captured. The Radiologist generates a diagnostic report for each case which is uploaded to the Grid.

In general, images are captured and read at the same clinic. However the system must allow (subject to policy) images captured at one clinic to be read at any other. Details of workflow may vary by geography: e.g. how many times an image is read, whether subsequent readers may see reports from previous readers etc. Image analysis tools to aid in detection and diagnosis should be available in the Grid and the Grid should support research and development to extend this toolset.

3.3.2 Radiologist Training

References to interesting images may be gathered together to form "training rolls" (akin to screening work lists) which may then be associated with other relevant teaching material that is distinct from eDiaMoND. A Training workstation pulls the images from the Grid in the same way as the Radiologist workstation.

3.3.3 Epidemiology

An Epidemiologist collects study data (e.g. dietary information) and builds a grid data resource. This data resource may be joined with the eDiaMoND mammography resource, subject to policy, for the duration of a particular study.

3.4 Involved resources

Each clinic has a server hosting eDiaMoND grid services and data storage systems, and one or more client workstations. The servers are connected by a secure, high bandwidth VPN.

3.5 Functional requirements for OGSA platform

- *Data Management*. Efficient storage and retrieval of medical images. Dynamic adaptation of the eDiaMoND Grid to actual data flows is desirable.
- *Configuration and control*. It should be possible to easily add new clinics (or remove them) from the eDiaMoND Grid. It should be possible to alter the topology of the eDiaMoND Grid transparently to the users without interrupting service.
- *Metering and accounting*. To provide and demonstrate compliance with regulations.
- *Policy*. See security considerations below.
- *Security*. See security considerations below.

3.6 OGSA platform services utilization

- *Name resolution and discovery*.
- *Policy*.
- *Data Access and Integration*. The eDiaMoND Grid builds on OGSA-DAI [3].
- *Provisioning and resource management*.
- *Metering and Accounting*.
- *Monitoring and Administration*.

3.7 Security considerations

The system must be capable of implementing policies that ensure that any data access complies with regulations, ethical considerations, legal requirements, patient consent etc. These policies may vary by clinic or by region.

Confidentiality, authorization, auditing and access control and data integrity are important. Authorization and access control should not require a single overarching central authority. Anonymization is a requirement for images passed to researchers.

3.8 Performance considerations

Four views per breast are taken each time a patient visits the screening clinic. This produces 4 DICOM files of 32Mb each. The new case and one or two prior cases (if available) are viewed by Radiologists. Relatively few queries are required to create work lists for reading and the service call overheads are therefore negligible. The screening application tends to be bandwidth limited.

3.9 Use Case situation analysis

Integration of data into a virtual resource may be done by proprietary mechanisms and the "merged" data resource exposed as a grid service; or by using service-based distributed query processing. The client deals with the same service interface irrespective of the underlying data virtualization mechanism.

Manual reconfiguration of the eDiaMoND Grid, without disruption of service and without reconfiguration of clients, is possible through the manipulation of service registries. Dynamic adaptation of data storage resources is a point of research.

The eDiaMoND Grid is a research prototype is investigating the characteristics of a distributed database of X-Ray images and its value in breast cancer screening and research. A production system would require significant extra integration with PACS and HIS systems and their associated workflows. These issues are being addressed by the Integrating the Healthcare (IHE) initiative [4]; a project designed to advance the state of data integration in healthcare.

3.10 Information Dissemination Patterns

Information dissemination falls into 4 categories:

- *Bulk*. By far the largest category. This includes overnight bulk retrieval of images for a reading session the following day and bulk loading of data to the Grid where the radiographer workflow at a clinic dictates data is loaded in large batches.
- *Trickle*. Loading of data to the Grid where the radiographer workflow at a clinic dictates data is loaded on a case-b-case basis.
- *Ad Hoc*. Images returned as the result of an ad hoc query by a Radiologist during a reading session.
- *Replication*. Patient data and image metadata is held as a separate data resource from the image. Data volumes are negligible in comparison to image data. This data may be federated or replicated as particular circumstances dictate (current prototype model is

DB2 II federation of relational databases at each clinic). The images are large binary objects that are not replicated in the Grid.

Images are always streamed in the eDiaMoND Grid using a pull model. Grid services identify a source and destination for the image among the set of grid nodes and clients. A process at the destination then streams the image directly from the source "out of band".

4. Scientific Scheduling Use Case

4.1 Summary

Efficient scheduling requires up to date knowledge of all the available resources that a user is entitled to exploit.

4.2 Customers

Actors include:

- The user or scheduler/resource broker acting on behalf of the user
- Information providers for each resource

4.3 Scenarios

- A scientist wants to run an interactive simulation/rendering job that must be finished within the next half hour. He must find a computing resource that is fast enough, has enough memory and is linked to a fast network connection. He submits a complex query to an information system and identifies the needed resources. He then submits his job.
  Security:
  The user must know that the resource information published is trustworthy. This implies that those who published the information should be reliably identified. If the information cannot be trusted the user would be obliged to confirm the information by connecting directly to the resource.

  Performance:
  The results should be returned in human response time , i.e., ~1 second. Information about some resources may be protected.

- A long running job specified in a job description file is submitted for batch processing. The scheduler tries to find the best match between the job requirements and the resources available on a Grid, whose characteristics are retrieved from an information service. In particular information such as policies to access resources, characteristics and status of resources and availability of the required application environment is needed. The scheduler does not need to find the best solution, but merely one near the optimum. The scheduler makes some predictions based on historical information.

  Security:
  Here the human of the precious scenario has been replaced by another service. The service must be even more careful about attempting to make use of a bogus resource as the service is trusted by its user. Information about some resources may be protected, however it may be possible to offer the service more information than would be offered to a user, as information about resources is of potential interest to hackers.

  Performance:
  One would hope that the scheduler service might make
  better use of the information system and make less queries to get its job done. This suggests that lower performance would be acceptable, however as the user sees the scheduler as a single service it will be judged by its ability to make multiple queries so probably requires better performance. This could be avoided if the scheduler returns immediately after accepting a job and then starts processing it in the background.

- A scheduler (or related component) might wish to ensure that any large files that are needed for a job are available at a storage resource close to where the job will run. There is no point in copying to a local cache too early in a busy system. Rather an estimate of when the job will run is needed along with an estimate of how long it will take to copy the file so that the data can be obtained just in time.

  Security: Same as previous scenario

  Performance: This will be a background process and so a response of one minute is fine.

5. Network Monitoring Use Case

5.1 Summary

Network monitoring is carried out by network service providers to ensure that the infrastructure is sound, by customers of those services to ensure that they are getting the service defined in some SLA and by schedulers (as described in the scheduling use case) which need the information to optimise data movements.

5.2 Customers

Actors include:

- Network service providers
- Customers of network service providers
- Schedulers/resource brokers

5.3 Scenarios

- A site network manager uses monitoring tools to get performance information such as round trip time (RTT) between pairs of hosts, percentage loss of packets; inter-packet delay variation and throughput achieved for file transfers. He needs to perform these measurements over periods of days or weeks to ensure that the promised QoS is delivered. He also makes the data available to other potential *consumers* including end-users and applications (e.g. schedulers and file transfer tools). The availability of the data for further analysis and evaluation is crucial.
- An end-user or application are more likely to be interested in the most recent measurements to adjust, in “real-time”, their usage of network resources and make sure that the agreed QoS is honored. The availability of the data is necessary if further action is to be taken when the agreed QoS is not delivered

5.4 Security:

Networks are resources so the comments about security above in the previous use case apply here

5.5 Performance:

For interactive use the ~ One second reponse time is expected.

6. Logging and Bookkeeping Use Case

6.1 Summary

Information needs to be available for users who have submitted jobs concerning the detailed status of their current jobs and an abbreviated history of completed jobs. It is assumed that as a job is processed it goes through a certain number of states until it is finally completed and all its output dealt with. The precise set of states is probably dependent upon the design of the scheduler. Computing resources are involved in the Use Case as they must publish some of the state changes of a job.

6.2 Customers

The actors include:

- Scheduler/resource broker
- Computing resources
- Job submitter

6.3 Scenarios

- A user submits a job to a scheduler and a jobId is assigned. It goes through various changes of state until it is dispatched to a computing resource to be queued again and

then executed. The user has expressed an interest in this job and will be informed of each state change.
- A user wants to know the status of all jobs submitted that week
- A user wants to know the status of all incomplete jobs.

6.4 Security:
Again the user wants to be reasonably sure that the information received has a known provenance. It is also very important that the information is protected. Typically users should *only* be able to find out about their own jobs.

6.5 Performance:
For the first scenario, a delay ~ 1 minute before being informed should be fine, however for the interactive scenarios ~ 1 second reponse is expected.

7. Notification of newly processed particle physics data Use Case

7.1 Summary

Particle physicists typically study a particular physics process. Data of interest are extracted from a very large (TBytes) body of data.

7.2 Customers

The actors involved are
- The particle physicist wishing to be notified of the existence of new data.
- The production manager who submits production batch jobs

7.3 Scenarios

The physicist is able to locate all existing data sets containing the physics data (events). He is also able to request notification when new data is obtained. This may be simulated data, newly processed experimental data or data that has been reprocessed. The physicist expresses his interest in a particular physics process. Someone responsible for running large-scale production jobs submits a set of jobs and each time a job runs and produces new data of interest for the physics process of interest the physicist is informed. The physicist might then submit a new job to analyze those new data.

7.4 Security:
Particle physicists don't care about security. Consequently itshould be possible to operate without the overhead implied by security.

7.5 Performance:
A delay of fifteen minutes would probably not be noticed.

8. Deisa Data Replication
In DEISA, a user account gets created at one site. Not the information (user data, certificate, ...) need to be known by all other partners as well. Currently, an LDAP service has been set up which has the information available and each site runs a replica of part of the tree that is updated every night.
Soon, a messaging service will be applied with a publish subscribe schema in the way, that once a user account has been created at the home site, this home sites notifies the other site of the availability of new user information and triggers a workflow that includes copying the necessary user data, creating UNIX account, home directories.

All sites will have subscribed to a message service. At the moment the plan is to run a service at each site for reliability

9. Data Movement for DAIS

9.1 Summary

This section is based on the materials in DAIS Usage Scenarios [5]

9.2 Customers

vdialani@us.ibm.com, A.Djaoui@rl.ac.uk, S.M.Fisher@rl.ac.uk, Dieter.Gawlick@oracle.com, Chris.Kantarjiev@oracle.com, madsen@us.ibm.com, malaika@us.ibm.com, Shailendra.Mishra@oracle.com

9.3 Scenarios

10. ERP Use Case to be provided by Dieter to cover R4

11. RFID Sensors to be provided by Dieter to cover R5 and R12

12. Homeland security – to be provided by Abdeslem to cover R6

13. Bank balance notification – to be provided by Chris to cover R7

14. e-mail filter – to be written by Susan to cover R8

15. Appendix: INFOD Interfaces

The INFOD interfaces are as follows
I1 Consumer Interface
I2 Disseminator Interface (optional)
I3 Propagator Interface (optional)
I4 General Registration Interface (optional)
I5 PublisherRegistration Interface (optional)
I6 SubscriptionRegistration Interface (optional)
I7 ConsumerRegistration Interface (optional)
I8 DisseminatorRegistration Interface (optional)

The only interface that is mandatory for INFOD is the Consumer Interface

16. The INFOD requirements:

These requirements are implied by the Use Cases

16.1 Publishers and Subscribers
R0: Subscribers must be able to select publishers of interest
R1. Subscribers must be able to select messages of interest
R2. Publishers must be able to select consumers of interest
R3. Publishers and subscribers must be able to use vocabulary management to select messages
R4. Subscribers may be able to specify the messages to be published in response to an event
R5. Subscribers may be able to define an event by querying one data source
R6. Subscribers may be able correlate events to form composite events by querying multiple data sources

16.2 Consumers
R7. Consumers must be able to specify dynamically how and when messages are received
R8. Consumers may be able to limit the messages received
R9. Consumers may be able to retrieve past messages on request: Publishers, Subscribers and Consumers
R10. Publishers, Subscribers and Consumers must be able to define QoS
R11: Publishers, Subscribers, and Consumers may have Identity Management

16.3 Filters
R12. Filters must be able to be composed

Author Information

Susan Malaika
IBM Data Management
Hawthorne, NY, US
malaika@us.ibm.com

Steve Fisher
Rutherford Appleton Laboratory (CCLRC)
Chilton
Didcot
Oxon OX11 0QX, UK
s.m.fisher@rl.ac.uk

Intellectual Property Statement

Full Copyright Notice

References

[1] ARTS Association for Retail Industry Standards http://www.nrf-arts.org/
[2] eDiaMoND http://www.ediamond.ox.ac.uk

[3] Open Grid Services Architecture Data Analysis and Integration (OGSA-DAI) http://www.ogsa-dai.org.uk

[4] Integrating the Healthcare Enterprise (IHE) http://www.rsna.org/IHE/index.shtml

[5] DAIS Usage Scenarios

https://forge.gridforum.org/projects/infod-wg/document/DAIS_Usage_Scenarios/en/1

17. Glossary

Consumer
Data Movement
Event
Notification
Propagator
Producer
Publisher
Subscriber