



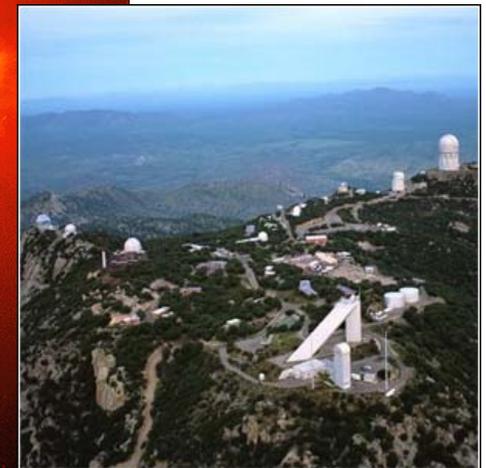
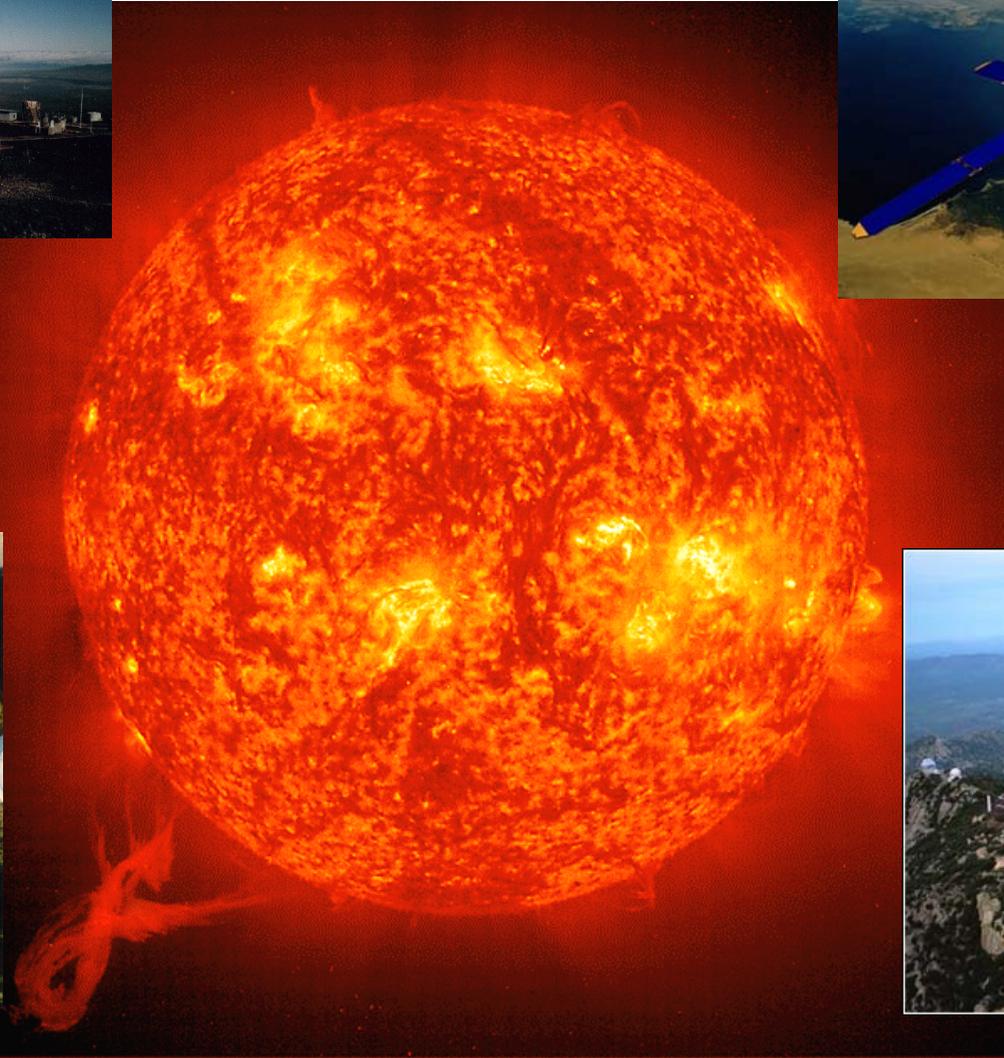
# **Model-Driven Architecture for building Data Grids: the EGSO experience**

Giacomo Piccinelli

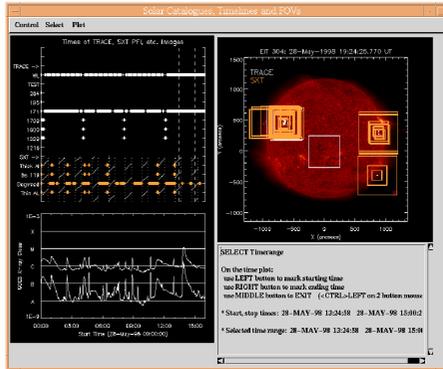
**University College London**



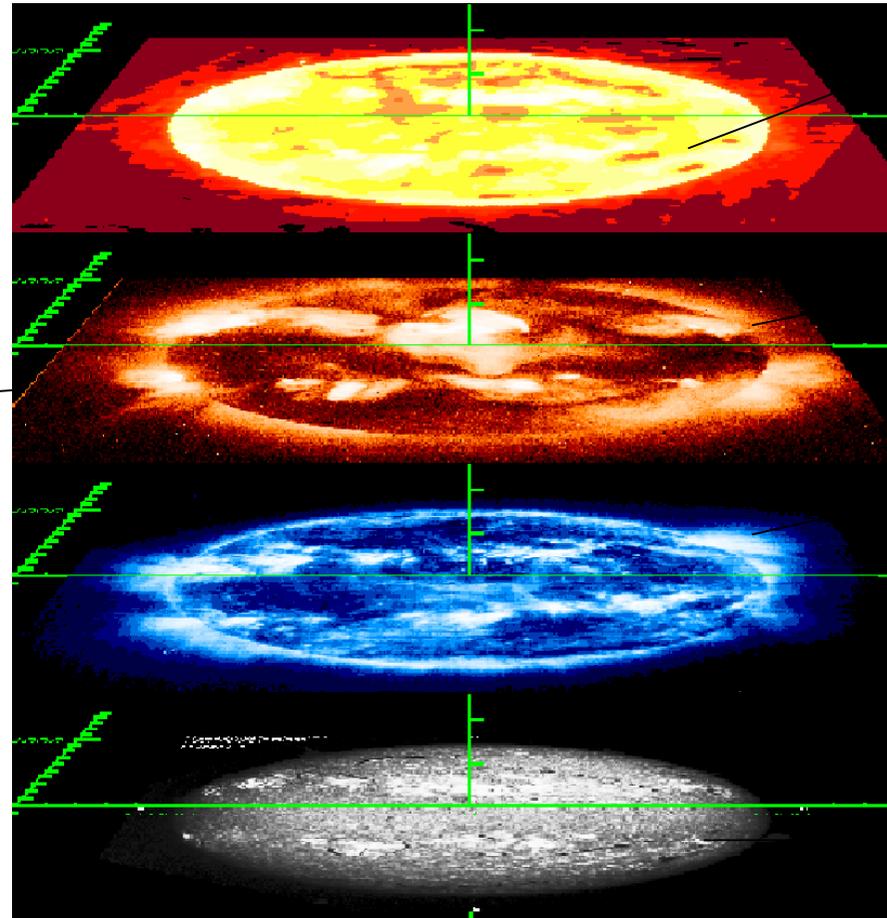
# Solar Observations



# A GRID for Solar Observations



- § Aggregation
- § Calibration
- § Correlation
- § Simulation
- § Forecasting
- § ...



# Contents

- **Motivation**
  - User Requirements
  - Related Works
  - Integration and Reusability
- **Architecture**
  - Consumers
  - Providers
  - Brokers
- **Service Provision**
  - Metamodel
  - UML Profile
- **Validation**
- **Conclusions**



# User Requirements

- Reardon K., Giordano S., and Antonucci E. (Eds.) “*User and science requirements document*” EGSO Report EGSO-WP1-D2-20021031, 2002.

<i>Requirement Classification</i>	<i>Description</i>
Administrative Metadata	<i>Resource description metadata</i>
Structural Metadata	<i>Organizational information</i>
Descriptive Metadata	<i>Information related to description of data contents</i>
User Interface	<i>GUI, API, and other means for user to access system</i>
Provider Interface	<i>Communications between resource providers and system</i>
Data Identification	<i>Data queries and identification of data of interest</i>
Query Handling	<i>Management and storage of queries</i>
Data Processing	<i>Calibration and processing of data</i>
Data Delivery	<i>Transfer of selected data to user</i>
User Environment	<i>User preferences and customization</i>
Resource Utilization	<i>Control and optimization of resource usage</i>
Security	<i>Access Control to resources</i>
Resource Description	<i>Description of data, metadata, and other sources that comprise system</i>
User Collaboration	<i>The means to allow multiple user's coordinated access to system</i>
Implementation	<i>General implementation guidelines</i>

# Related Works

- Piccinelli G. (Ed.) “*EGSO Architecture*” EGSO Report EGSO-WP1-D4-20030217, 2003.

7.1 European DataGrid (EDG)

7.2 Grid Physics Network (GriPhyN)

7.3 Particle Physics Data Grid (PPDG)

7.4 International Virtual Datagrid Laboratory (iVDGL)

7.5 Biomedical Informatics Research Network (BIRN)

7.6 Earth System Grid (ESG)

7.7 Virtual Solar Observatory (VSO)

7.8 Collaborative Sun-Earth Connector (Co-SEC)

7.9 National Virtual Observatory (NVO)

7.10 Astrophysical Virtual Observatory (AVO)

7.11 Astrogrid

7.12 International Virtual Observatory Alliance (IVOA)

7.13 Summary of Observations

## Architectural Model

### Data & Metadata:

- Structural Aspects
- Usage Processes

### Technology

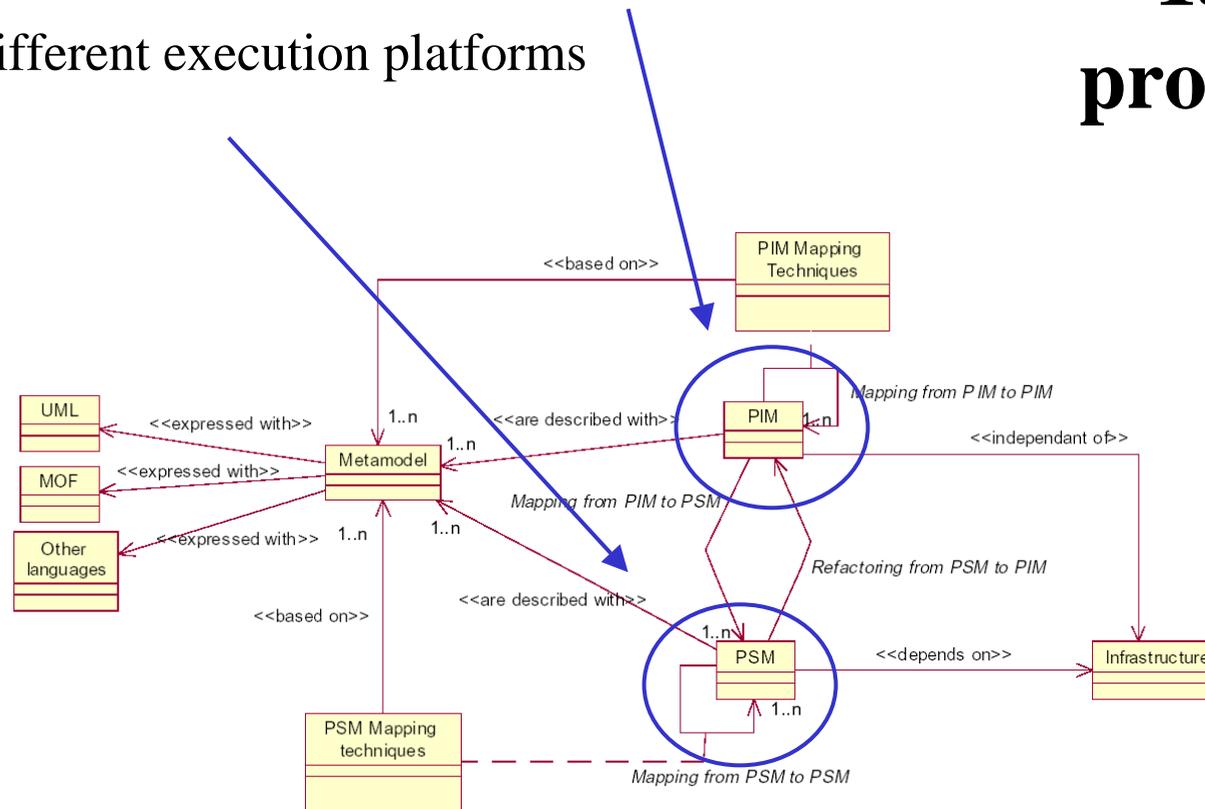


# Integration and Reusability

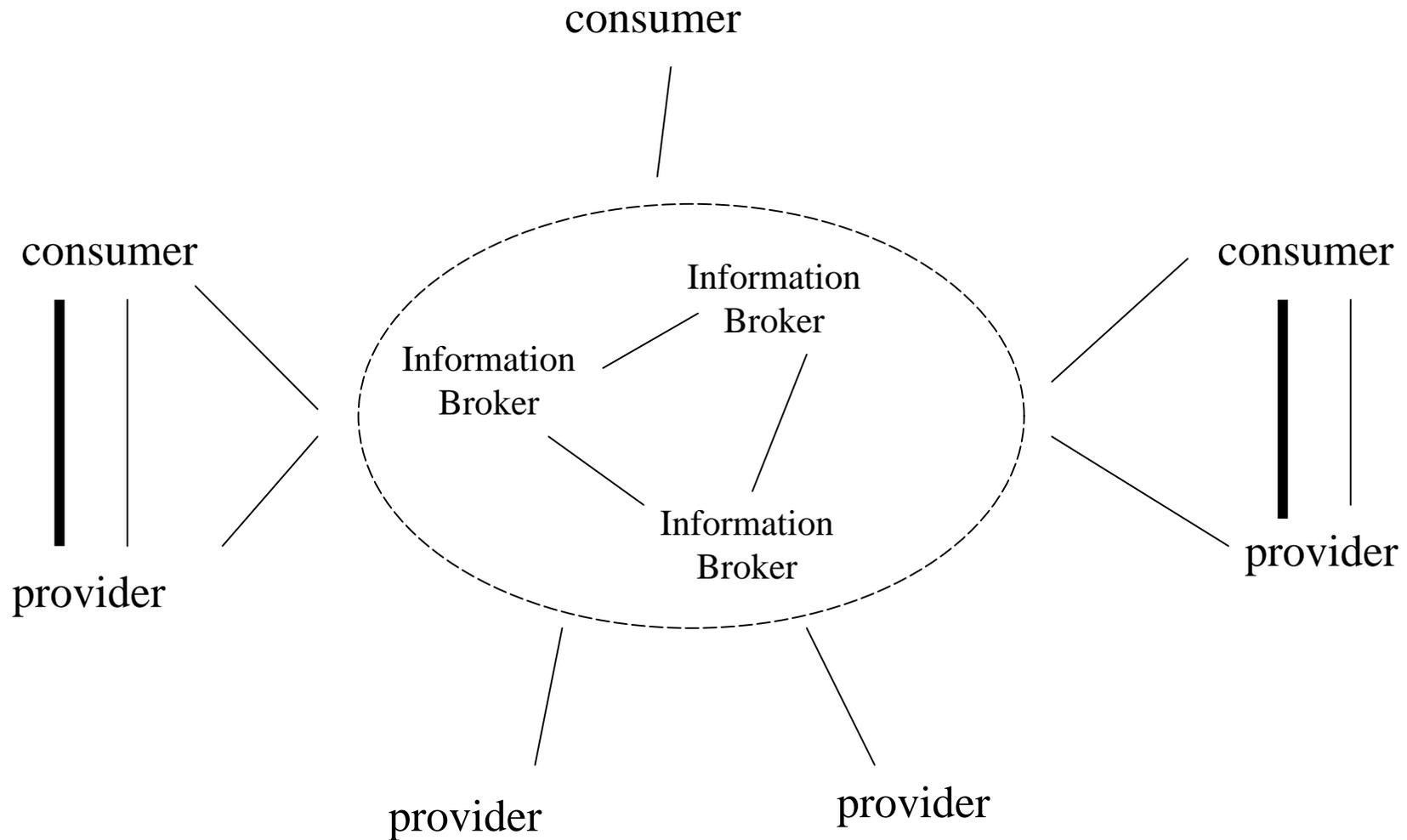
§ Different System and Data models

§ Different execution platforms

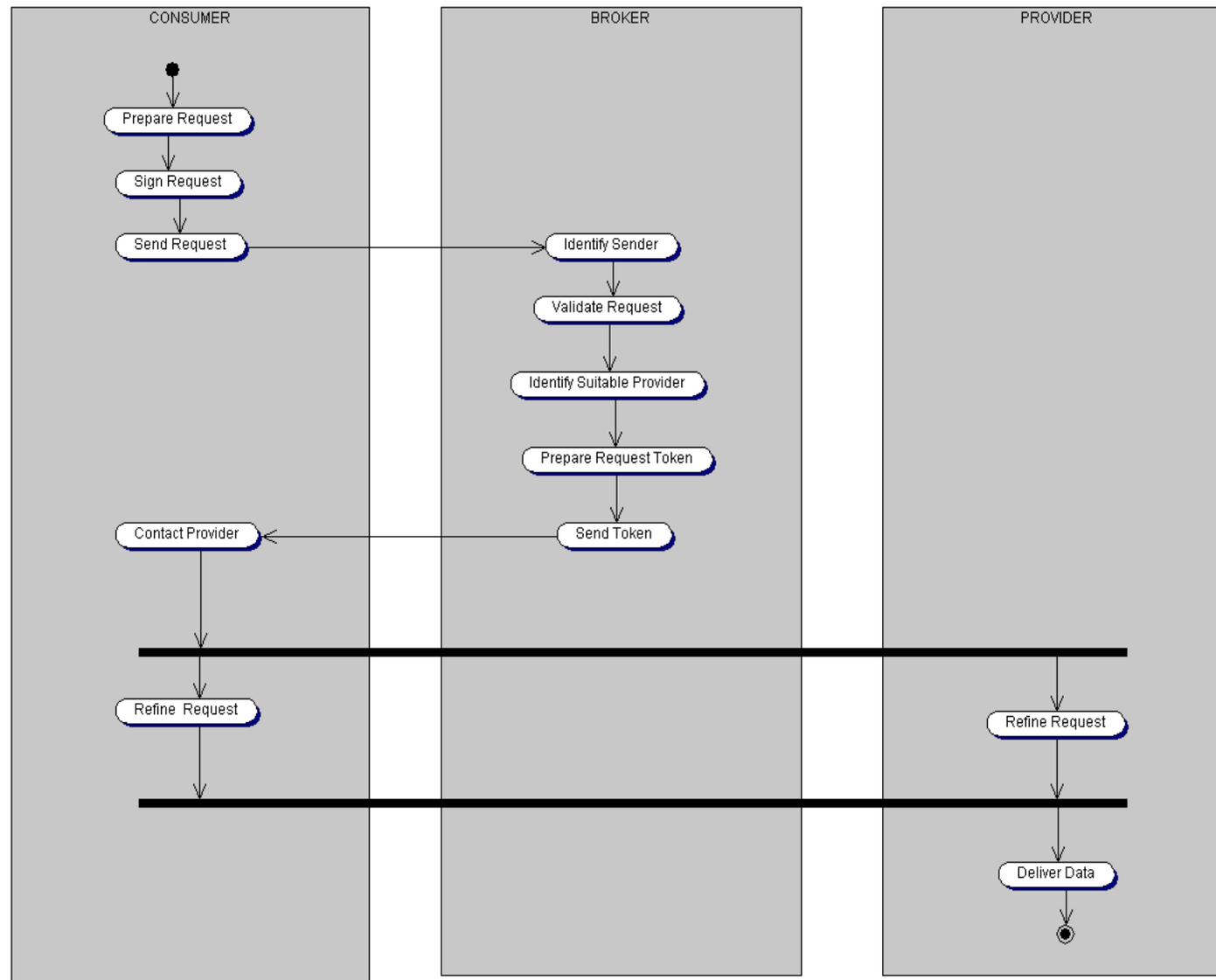
**150+**  
**providers**



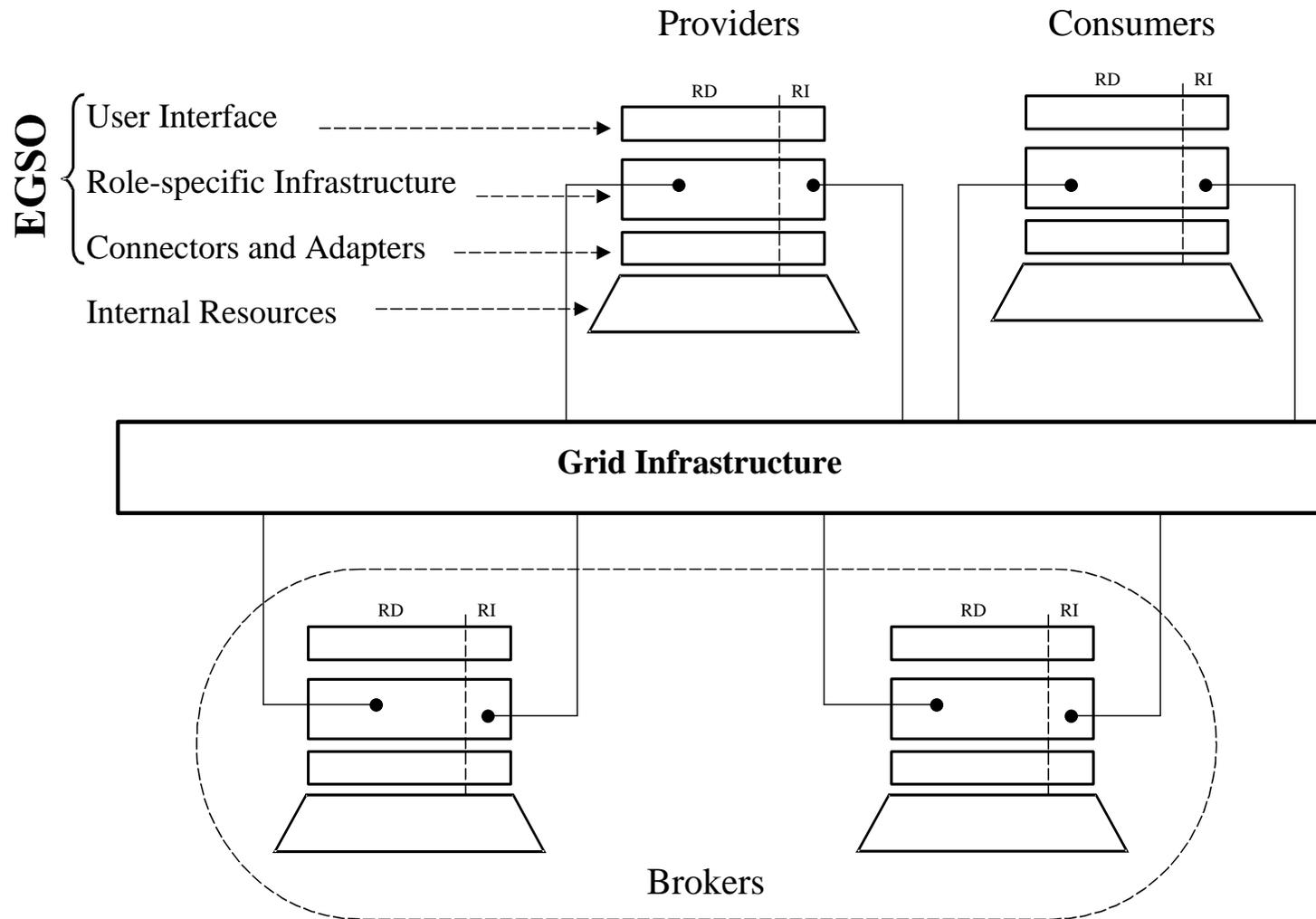
# Roles and Relations



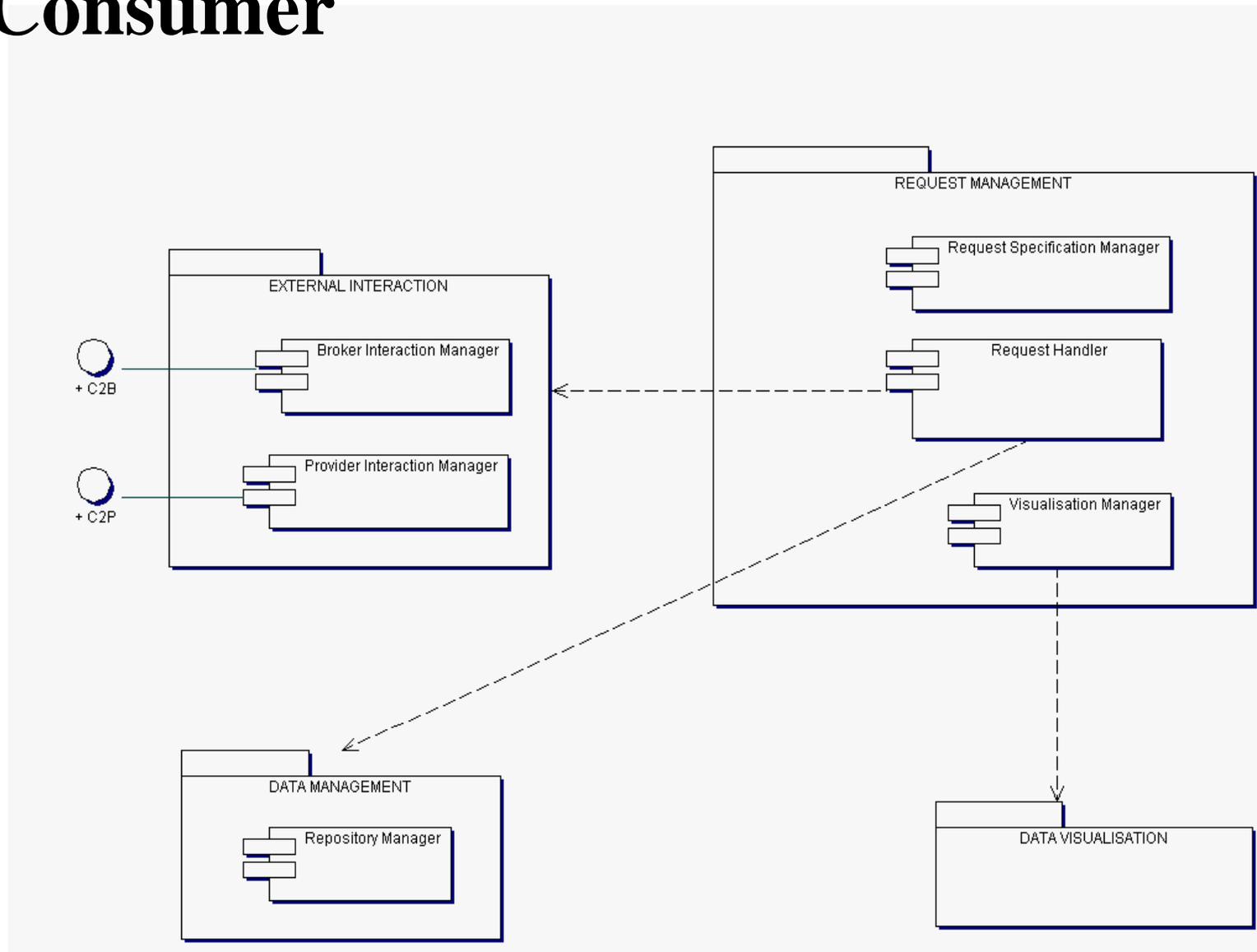
# Dataflow for a Request



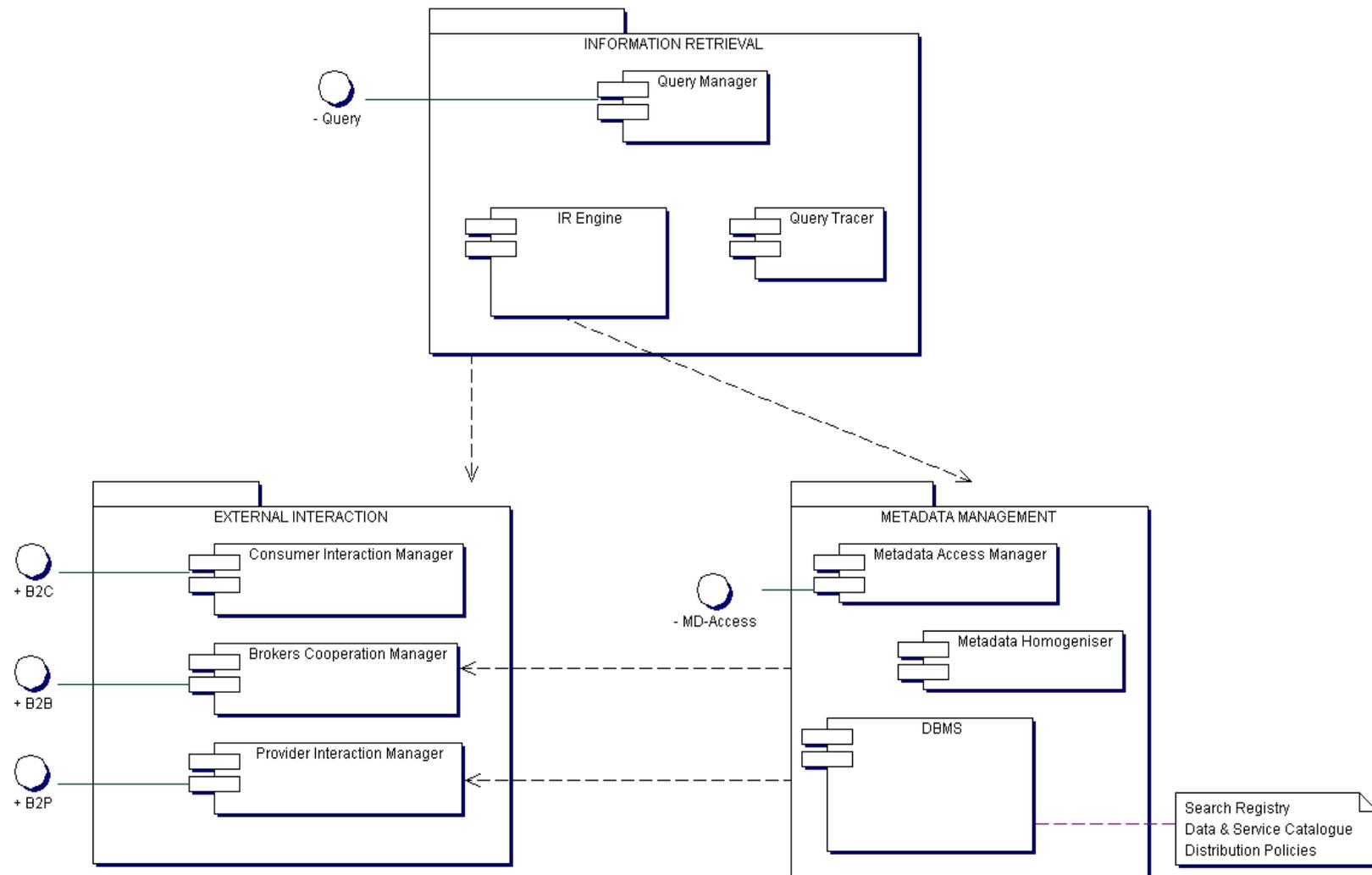
# 3-Tier Architecture



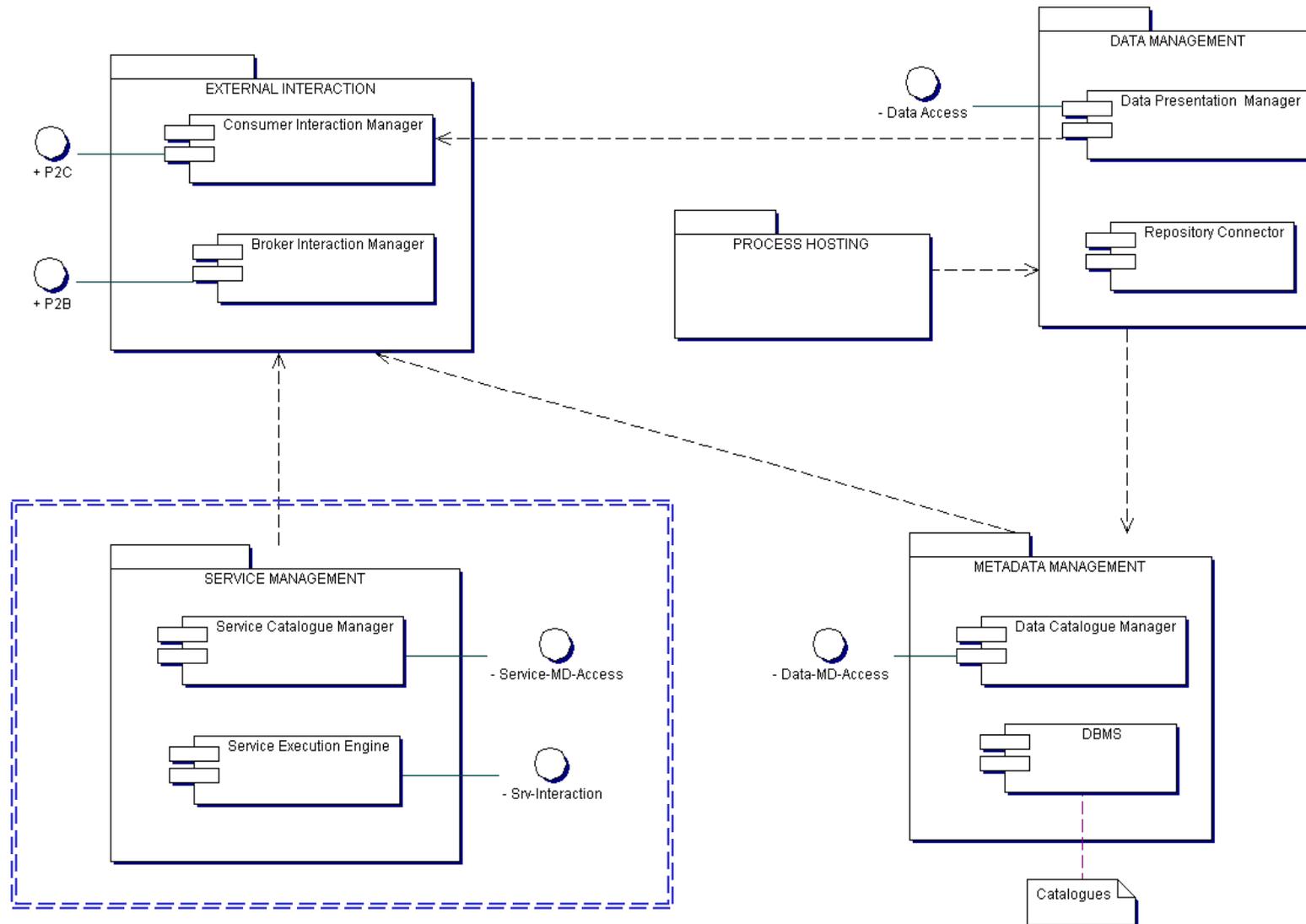
# Consumer



# Broker



# Provider



# Architecture Animation (Consumer)

The screenshot displays the EGSO Consumer application interface, which is divided into several panels:

- Request Execution Panel:** Contains a tree view for 'Request Specifications' (Basic, Features, Events), an XML specification detail, and a workflow summary diagram. The XML is:
 

```

      <?xml version="1.0" encoding="UTF-8"?>
      <root>
      <module name="defineRange"/>
      <module name="defineTolerance"/>
      <module name="dataGathering"/>
      <module name="getInstrumentData"/>
      <module name="calibration"/>
      <module name="getSynopticData"/>
      </module>
      <module name="archiveResults"/>
      </root>
      
```

 The workflow diagram shows a cycle: 'Data Gathering Begin' (yellow diamond) leads to 'Get Instrument Data' (green square), then 'Calibration' (green square), then 'Data Gather End' (yellow diamond), then 'Get Synoptic Data' (green square), and finally back to 'Data Gathering Begin'.
- Input/Output Panel:** Includes text boxes for 'Input' and 'Output' and an 'Execute' button.
- Visualisation Manager Panel:** Shows 'Installed Modules' with three entries:
  - Foo:** Description: Visualisation Module Foo; Dependencies: Bar, Baz.
  - Bar:** Description: Visualisation Module Bar; Dependencies: Foo, Baz.
  - Baz:** Description: Visualisation Module Baz; Dependencies: Bar, Foo.
- Provider Interaction Manager Panel:** Displays a table of provider constraints and usage summary.
 

Hostname	Ping	Availability	Preferred
provider.mssl.ucl...	17ms	62%	<input checked="" type="checkbox"/>
provider.dasop.o...	37ms	1%	<input type="checkbox"/>
provider.rl.ac.uk	146ms	19%	<input type="checkbox"/>
provider.cs.ucl.ac...	19ms	53%	<input type="checkbox"/>
provider.bradford...	137ms	72%	<input type="checkbox"/>
provider.ias.u-ps...	20ms	94%	<input type="checkbox"/>
provider.fh-aarga...	75ms	30%	<input type="checkbox"/>
provider.polito.it	91ms	42%	<input type="checkbox"/>
provider.gsfc.nas...	82ms	8%	<input type="checkbox"/>
provider.inaf.it	82ms	23%	<input type="checkbox"/>
provider.nso.edu	28ms	61%	<input type="checkbox"/>



# Architecture Animation (Broker)

The screenshot displays the EGSO Broker application with several open windows:

- Query Manager:** Shows the status of query instances in a table.
- Query Tracer:** Shows a query transaction audit table.
- Metadata Manager:** Displays a tree view of metadata access and homogeniser settings.
- Consumer Interaction Manager:** Shows a table of consumers serviced with their ping and availability.
- Provider Profile Manager:** Shows a tree view of provider profiles.
- Consumer Profile Manager:** Shows a tree view of consumer profiles.

**Query Manager - Status of Query Instances**

Query Instance	Request Instance	Consumer	Status	Providers
AB1234CD56	AB1234CD56	foo.bar.org	Complete	{foo.bar.org, baz.bif.o...
AB1234CD56	AB1234CD56	foo.bar.org	Blocked	foo.bar.org
AB1234CD56	AB1234CD56	foo.bar.org	Executing	{foo.bar.org, baz.bif.o...
AB1234CD56	AB1234CD56	foo.bar.org	Blocked	
AB1234CD56	AB1234CD56	foo.bar.org	Complete	foo.bar.org
AB1234CD56	AB1234CD56	foo.bar.org	Complete	{foo.bar.org, baz.bif.o...
AB1234CD56	AB1234CD56	foo.bar.org	Complete	{foo.bar.org, baz.bif.o...
AB1234CD56	AB1234CD56	foo.bar.org	Executing	
AB1234CD56	AB1234CD56	foo.bar.org	Executing	foo.bar.org
AB1234CD56	AB1234CD56	foo.bar.org	Complete	{foo.bar.org, baz.bif.o...

**Query Tracer - Query Transaction Audit**

Query Instance	Consumer	Provider	Data Involved
AB1234CD56	consumer.ias.u-p...	provider.fh-aarga...	foo.xml
AB1234CD56	consumer.bradfor...	provider.nso.edu	bar.xml
AB1234CD56	consumer.fh-aarg...	provider.rl.ac.uk	foo.xml
AB1234CD56	consumer.nso.edu	provider.ias.u-psu...	baz.xml
AB1234CD56	consumer.ias.u-p...	provider.mssl.ucl...	bar.xml
AB1234CD56	consumer.gsfc.n...	provider.polito.it	foo.xml
AB1234CD56	consumer.rl.ac.uk	provider.cs.ucl.ac...	bif.xml
AB1234CD56	consumer.inaf.it	provider.mssl.ucl...	foo.xml
AB1234CD56	consumer.mssl.u...	provider.dasop.ob...	bar.xml
AB1234CD56	consumer.mssl.u...	provider.ias.u-psu...	bar.xml
AB1234CD56	consumer.polito.it	provider.inaf.it	baz.xml

**Consumer Interaction Manager - Consumers Serviced**

Hostname	Ping	Availability
consumer.mssl.ucl.ac.uk	133ms	64%
consumer.dasop.obspm.fr	111ms	70%
consumer.rl.ac.uk	64ms	86%
consumer.cs.ucl.ac.uk	9ms	19%
consumer.bradford.ac.uk	49ms	91%
consumer.ias.u-psud.fr	87ms	80%
consumer.fh-aargau.ch	42ms	35%
consumer.polito.it	3ms	97%
consumer.gsfc.nasa.gov	31ms	0%
consumer.inaf.it	12ms	97%
consumer.nso.edu	22ms	50%

**Provider Profile Manager - Provider Profiles**

- Providers
  - provider.mssl.ucl.ac.uk
  - provider.dasop.obspm.fr
  - provider.rl.ac.uk
  - provider.cs.ucl.ac.uk
  - provider.bradford.ac.uk
  - provider.ias.u-psud.fr
  - provider.fh-aargau.ch
  - provider.polito.it
  - provider.gsfc.nasa.gov
  - provider.inaf.it
  - provider.nso.edu



# Architecture Animation (Provider)

The screenshot displays the EGSO Provider application with several open windows:

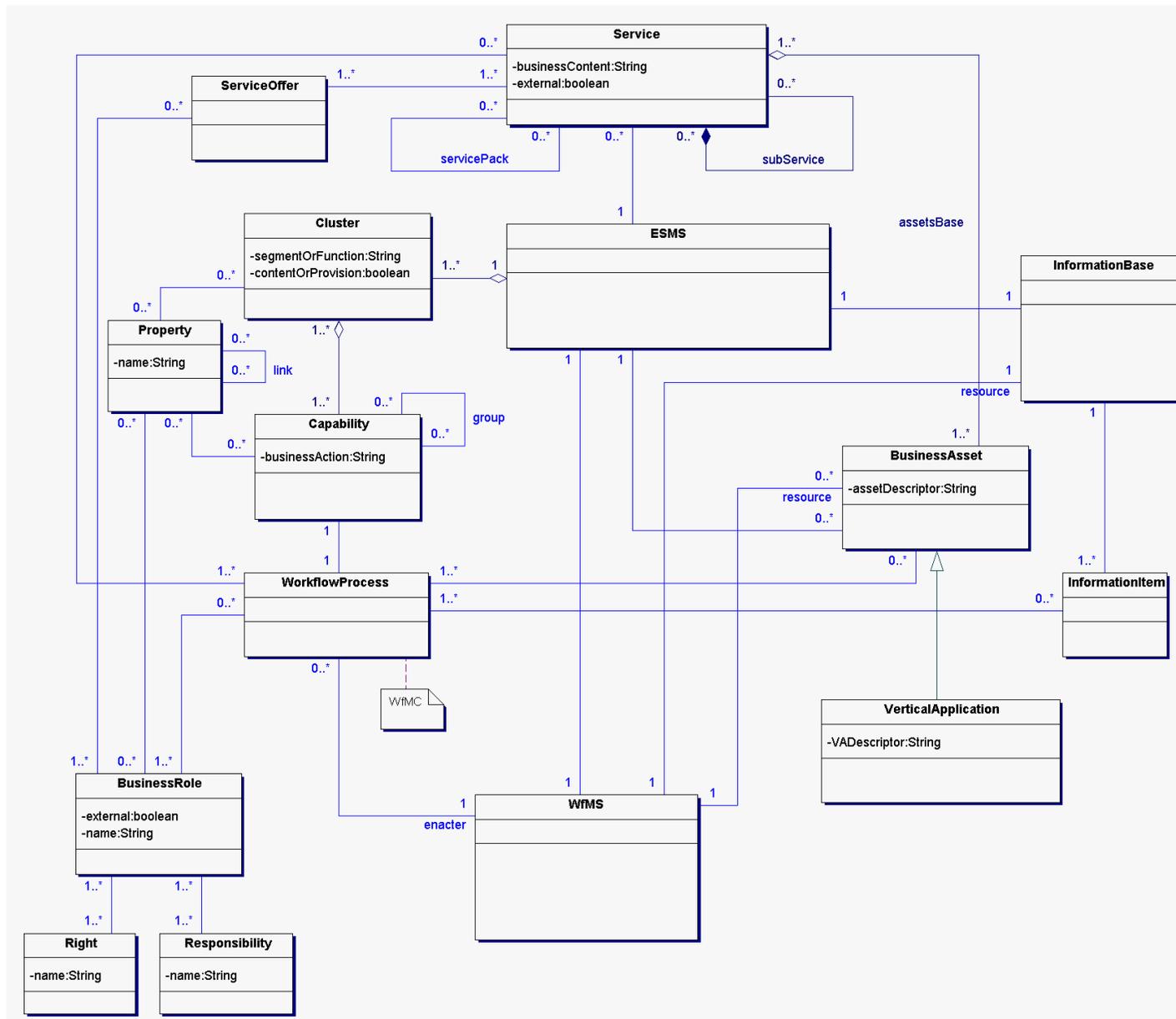
- Data Manager:** Shows 'Connected Repositories' with a tree view containing:
  - file://home/nching/data
  - ftp://ftp.mssl.ucl.ac.uk/pub/solar/nching/data
  - jdbc:mysql://localhost/data?user=nching
- Service Manager:** Displays 'Status of Service Instances' in a table:
 

Request Instance	Type	Status	Input	Output
AB1234CD56	basic.Foo	Complete	file:///foo/Bar	file:///foo/Bar
AB1234CD56	basic.Bar	Executing	file:///foo/Bar	
AB1234CD56	feature.Baz	Complete	file:///foo/Bar	file:///foo/Bar
AB1234CD56	event.Foo	Blocked	file:///foo/Bar	
AB1234CD56	basic.Foo	Executing	file:///foo/Bar	
AB1234CD56	event.Baz	Complete	file:///foo/Bar	file:///foo/Bar
AB1234CD56	feature.Bar	Blocked	file:///foo/Bar	
- Data Catalogue Manager:** Shows 'Registered Catalogues' with a tree view containing:
  - Catalogue Foo
  - Catalogue Bar
  - Catalogue Baz
  - Catalogue Bif
- Service Catalogue Manager:** Shows 'Registered Catalogues' with a tree view containing:
  - Catalogue Foo
  - Catalogue Bar
  - Catalogue Baz
  - Catalogue Bif
- Process Manager:** Displays 'Processes Deployed' in a table:
 

Request Instance	Owner	Signature Verifi...	Code Base	Deployment Scr...	Input Sandbox	Output Sandbox
AB1234CD56	consumer.mssl...	<input checked="" type="checkbox"/>	foo.jar	build.xml	foo/	bar/
AB1234CD56	consumer.mssl...	<input checked="" type="checkbox"/>	bar.jar	build.xml	foo/	bar/
AB1234CD56	consumer.rl.ac...	<input checked="" type="checkbox"/>	baz.jar	build.xml	foo/	bar/
AB1234CD56	consumer.cs.u...	<input checked="" type="checkbox"/>	bif.jar	build.xml	foo/	bar/



# Service Provision - Matamodel



# Service Provision – UML Profile

Metamodel element name	Stereotype name	UML Base Class	Parent	Tags	Constraints
BusinessAsset	BusinessAsset	Class	N/A	assetDescriptor	
BusinessRole	BusinessRole	Class	N/A	name external	
Capability	Capability	Class	N/A	businessAction	
InformationBase	InformationBase	Class	N/A		
InformationItem	InformationItem	Class	N/A		
Property	Property	Class	N/A	name	
Right	Right	Class	N/A	name	
Responsibility	Responsibility	Class	N/A	name	
Service	Service	Class	N/A	external businessContent	
ServiceOffer	ServiceOffer	Class	N/A		
ESMS	ESMS	Class	N/A		
Cluster	Cluster	Class	N/A	businessFunction	
VerticalApplication	VerticalApplication	Class	BusinessAsset	VADescriptor	
WfMS	WfMS	Class	N/A		
WorkflowProcess	WorkflowProcess	Class	N/A		

# Service Provision – UML Profile

Metamodel attribute name	Tag	Stereotype	Type	Multiplicity	Description
assetDescriptor	assetDescriptor	BusinessAsset	String	1	
businessAction	businessAction	Capability	String	1..*	
businssContent	businssContent	Service	String	1..*	
contentOrProvision	contentOrProvision	Cluster	Boolean	1	
external	external	BusinessRole	Boolean	1	default = false
external	external	Service	Boolean	1	default = true
name	name	BusinessRole	String	1	
name	name	Property	String	1	
name	name	Right	String	1	
name	name	Responsibility	String	1	
segmentOrFunction	segmentOrFunction	Cluster	String	1	
VADescriptor	VADescriptor	VerticalApplication	String	1	

# Metamodel Validation

$$\begin{array}{l}
 IB :: CT_i - \alpha \rightarrow IB' :: CT_i' \\
 IB :: CT_i - \alpha^{(*)} \rightarrow IB' :: CT_i' \\
 IB :: CT_i - \alpha^{(*,i)} \rightarrow IB' :: CT_i'
 \end{array}
 \quad
 \begin{array}{l}
 \text{(CMI)} \frac{IB :: CT_i - \alpha \rightarrow IB' :: CT_i'}{IB :: \{ CT_1 \dots CT_i \dots CT_N \} - \alpha^{(i)} \rightarrow IB' :: \{ CT_1 \dots CT_i' \dots CT_N \}} \\
 \text{(CMI)} \frac{IB :: CT_i - \alpha \rightarrow IB' :: CT_i' \quad IB :: CT_{j_1} - \alpha^{(*)} \rightarrow IB' :: CT_{j_1}' \dots IB :: CT_{j_n} - \alpha^{(*)} \rightarrow IB' :: CT_{j_n}'}{IB :: \{ \dots CT_{j_1} \dots CT_i \dots CT_{j_n} \dots \} - \alpha^{(i)(j_1 \dots j_n)} \rightarrow IB' :: \{ \dots CT_{j_1}' \dots CT_i' \dots CT_{j_n}' \dots \}}
 \end{array}$$

$$\begin{array}{l}
 IB :: CL_i - \beta \rightarrow IB' :: CL_i' \\
 IB :: CL_i - \beta^{(*)} \rightarrow IB' :: CL_i' \\
 IB :: CL_i - \beta^{(*,i)} \rightarrow IB' :: CL_i'
 \end{array}
 \quad
 \begin{array}{l}
 \text{(CLMI)} \frac{IB :: \{ CT_1 \dots CT_i \dots CT_N \} - \alpha^{(i)M} \rightarrow IB' :: \{ CT_1 \dots CT_i' \dots CT_N \}' \text{ and } \exists t | (t \in M \cup \{i\} \text{ and } t \in \{k_1, \dots, k_n\})}{IB :: CL \{k_1, \dots, k_n\} - \alpha^{(\beta)} \rightarrow IB' :: CL' \{k_1, \dots, k_n\}}
 \end{array}$$

$$\begin{array}{l}
 IB :: S_i - \gamma \rightarrow IB' :: S_i' \\
 IB :: S_i - \gamma^{(i)} \rightarrow IB' :: S_i'
 \end{array}
 \quad
 \begin{array}{l}
 \text{(S)} \frac{IB :: \{ CL_1 \dots CL_g \} - \beta^{(*)} \rightarrow IB' :: \{ CL_1 \dots CL_g \}'}{IB :: S_i \{i_1 \dots i_o\} - \alpha^{(\beta)} \rightarrow IB' :: S_i' \{i_1' \dots i_o'\}} \\
 \text{(SD)} \frac{IB :: \{ CL_1 \dots CL_g \} - \beta^{(*,j)} \rightarrow IB' :: \{ CL_1 \dots CL_g \}'}{IB :: S_i \{i_1 \dots i_j \dots i_o\} - \alpha^{(\beta)} \rightarrow IB' :: S_i \{i_1' \dots i_j' \dots i_o\}}
 \end{array}$$

# Metamodel Validation

Process	Event	Component
Consumer	consumer[1..2].input_query	consumer.request-management. request-specification-manager
	consumer[1..2].read_stored_query	consumer.data-management.repository-manager
	consumer[1..2].store_query	
	consumer[1..2].visualise	consumer.data-visualisation
Consumer, ConsumerIF	consumer[1..2].availability_query. ref_brok[1..2].ref_dset[1..3]	consumer.external-interaction.broker-interaction-manager, consumer.request-management. request-execution-manager
	consumer[1..2].transparent_query. ref_brok[1..2].ref_dset[1..3]	
	consumer[1..2].broker_result. ref_prov[1..2]	
	consumer[1..2].direct_query. ref_prov[1..2]	
Broker, ConsumerIF	broker[1..2].broker_result. ref_cons[1..2].ref_prov[1..2]	manager
	broker[1..2].fwd_available_req. (+suffix)	broker.external-interaction.broker-coordination-manager , broker.information-retrieval.query-manager, broker.external-interaction.consumer-interaction-manager
broker[1..2].fwd_available_ack. (+suffix)		
broker[1..2].fwd_available_nack. (+suffix)		
(suffix).ref_cons[1..2].ref_dset[1..3].fwd_co[1..2]		
ConsumerIF	broker[1..2].no_store. ref_cons[1..2]	
Broker,	broker[1..2].ref_prov[1..2]. ref_dset[1..3].set_item	broker.information-retrieval.ir-engine,

Events: consumer.1.input\_query, consumer.1.availability\_query.ref\_brok.1.ref\_dset.3, broker.1.availability\_query.ref\_cons.1.ref\_dset.3, broker.1.ref\_prov.1.ref\_dset.3.read\_item.0, broker.1.skip.ref\_prov.1, broker.1.ref\_prov.2.ref\_dset.3.read\_item.0, broker.1.fwd\_available\_req.ref\_cons.1.ref\_dset.3.fwd\_co.1, broker.2.fwd\_available\_ack.ref\_cons.1.ref\_dset.1.fwd\_co.2, broker.2.ref\_prov.1.ref\_dset.3.read\_item.0, broker.2.skip.ref\_prov.1, broker.2.ref\_prov.2.ref\_dset.3.read\_item.1, broker.2.broker\_result.ref\_cons.1.ref\_prov.2, consumer.1.broker\_result.ref\_prov.2, consumer.1.visualise.



# Summary

- ✓ **MDA: architecture Reuse and Integration**
- ✓ **Metamodels: formal validation**
- ∅ **Automation beyond design?**
- ∅ **Platform-specific models beyond components?**

