

BioMOBY Asynchronous Service Call Proposal

Call Proposal GNV5-06/01

(v 2.2)

Contributions:

Enrique de Andrés José-María Fernández Johan Karlsson Sergio Ramírez José Manuel Rodríguez Carrasco Roman Rosset

Coordinators: Oswaldo Trelles, David González-Pisano

Node GNV-5: Integrated Bioinformatics University of Málaga

Málaga, 07 August 2006





1. - Preliminaries

Throughout this document, we will assume that the reader is already familiar with the basic concepts of BioMOBY. For more information about BioMOBY, please study the official documentation¹.

This document contains the proposal from INB² of how to deal with asynchronous services in BioMOBY. The proposal is a result of discussions during the INB Meeting in Málaga (July, 2005) with the participation of Martin Senger and Edward Kawas, and in the INB mailing lists. During further discussion in the "MOBY-dev" mailing list, a suggestion was put forward to investigate if it was possible to base the messaging on an OASIS standard called WSRF.

The aim of the proposal is to contribute to the *standardisation* of asynchronous service calls in BioMOBY. A new extended set of service calling operations and defined XML messages are detailed.

The main motivation for the proposal is to facilitate the implementation of "long-running" services, i.e. services that demand enough computational resources to need more than a few minutes to compute the result.

INB[®]

¹ http://www.biomoby.org

² Instituto Nacional de Bioinformática (INB), Spain, http://www.inab.org

2- Current BioMOBY specification

In the **MOBY-S 0.86.3** version of the BioMOBY API, services present their interfaces as Simple Object Access Protocol (SOAP³) RPC.

A MOBY compliant service (registered as having the service protocol "moby") is one that uses only object/service classes defined in the MOBY Central registry, presents its service interface via SOAP, and registers this service interface in MOBY Central. In the coming months it will be expanded to allow the registration of non-MOBY SOAP services, as well as CGI services, but this will not affect the API described below for MOBY-SOAP services.

Table 1 - Current BioMOBY definition of "MOBY compliant service"

A service can be called through a single procedure (operation), by using the name the service was registered with in the MOBY Central catalog.

After retrieving a service description (currently in the form of a simplistic, but legitimate, WSDL document) from MOBY Central, client programs will subsequently communicate directly with the service provider, first by sending a request, in the form of an input message and then, if all goes well, by receiving an output message. The communication takes the form of a very simplistic SOAP RPC call: the name of the remote procedure call is the same as that when it was registered in MOBY Central.

The URI (uniform resource identifier) looks like a URL (uniform resource locator), but is subtly different. Where a URL is the address of a document on the Internet, a URI is an abstract identifier which allows the service to be uniquely identified.

At this time, the URI for this procedure call is always http://biomoby.org, as in:

http://biomoby.org/#your_procedure_call_name

This is regardless of the URI for the service provider! This is useful because the same service might be available from several providers. If they all use the same URI, then a computer (or human) can infer that they are equivalent, and swap one for the other, based on availability, or other criteria.

Table 1 - Current BioMOBY specification for SOAP RPC calls

HTTP is the usual transport protocol. Although not forced by the BioMOBY specification, most of BioMOBY service providers install a web server that handles the SOAP requests.

The most straightforward paradigm is to have a single SOAP server running as a CGI script, and this listener hands-off requests to the appropriate code module as requests arrive.

Table 2 - from "Constructing MOBY-S Compliant Services"

Several problems related to long-running services have been identified in the current BioMOBY specification. The same transaction is used to request the execution of a service and to wait for the results. This causes the system to remain occupied and non-responsive in the meanwhile. Also, it is a de facto standard to set a connection timeout both in server and clients, closing the socket being used for client-service communication. Often the connection timeout is no more than a few minutes, making it impossible to call long-running services.

³ <u>http://www.w3.org/TR/soap/</u>



Figure 1 - Illustration of the problem with long-running services in the current BioMOBY specification



3- Specification of proposal

To enable calls to long-running services, it is better to dedicate a separate connection for each step (thereby using an asynchronous communication model). Unfortunately, the current BioMOBY specification does not support this.

In this chapter, we suggest how to add such support to the BioMOBY specification. We have selected a polling approach. This requires that the service maintains a state. Therefore, we will use an OASIS standard called WSRF to help implementation of state-aware services.

Before we continue, let us describe some concepts that we will use throughout this document:

Batch-call (or session):

This is a service invocation and corresponds to one MOBY/mobyContent message.

Job:

A job is one of the multiple service execution requests that belongs to a service invocation, and corresponds to one mobyData message.

3.1. WSRF - Web Services Resource Framework

In this section we aim to provide a brief, general overview of WSRF. Those interested in details of WSRF and the operations defined there are invited to study the documentation⁴.

WSRF is an OASIS standard that can be used to implement interoperable state-aware webservices (WS). WSRF uses the W3C standard WS-Addressing to achieve transport-neutral addressing for web-services. The concept of End-Point-References (EPR) is central in WS-Addressing. An EPR is used to refer to a specific service instance.



Table 1: General example of EPR for service MyService.

An EPR is returned by a WS-Addressing compliant service as a response to an invocation of the service. An EPR can hold several fields but important for this proposal are the Address and ReferenceParameters fields. The Address tells the client what URL should be used in future communication (it may or may not be the same as the HTTP connection URL). The optional ReferenceParameters field contains an *opaque* reference to the specific service instance at the Address URL (similar to a ticket or identifier). The EPR is used by the client and the service in subsequent message exchanges.

A central WSRF concept is "WS Resources". It can be used to represent resources such as a shopping cart, hardware such as a printer or a print job created in a printer. WSRF defines the messages sent between client and service and API functions that the client uses to interact with these resources. A WS Resource is represented by a property document. This document represents the state of the resource and is communicated as XML in the SOAP body.

⁴ http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsrf

In WSRF there are several APIs that can be used:

- WS-Addressing
- WS-Notification
- WS-Policy
- WS-Security.

However, only WS-Addressing is mandatory. The mandatory operations defined there by WSRF must be implemented by all WSRF web-services. For a full list and information which operations are mandatory and which are optional, please see the official WSRF documentation.

A WS Resource consists of a set of properties (called a document), each with a particular name that is unique within the document. Interaction with WS Resource properties from the client (for example to get or set the value of the properties) is performed by using operations from the WS-Resource specification. If a call with such an API function fails, there are several pre-defined error messages (called Faults). These Faults can also be extended to provide service-specific error messages.

Since WSRF is WS-Addressing compliant, any WSRF compliant service must use an EPR to refer to a particular WS Resource.

3.2. Operations, faults and properties

We propose that asynchronous BioMOBY services represent a batch-call as one WS Resource and therefore that the way to refer to such a batch-call is by using the EPR.

A BioMOBY asynchronous service must implement the following WSRF operations⁵:

- *GetResourceProperty*: This mandatory WSRF operation (from WS-Resource) allows a requestor to retrieve the value of a *single* resource property of a WS-Resource.
- GetMultipleResourceProperties: This optional WSRF operation (from WS-Resource) allows a requestor to retrieve the values of *multiple* resource properties of a WS-Resource.
- Destroy: This optional WSRF operation (from WS-ResourceLifetime) allows a service requestor to request the immediate destruction of a WS-Resource.

Furthermore, both BioMOBY asynchronous services and BioMOBY asynchronous clients must be able to process the following standard WSRF faults.

- *ResourceUnknownFault*: Used to indicate that the resource identified in the message is not known to the Web service.
- *ResourceUnavailableFault*: Used to indicate that the Web service is active, but unable to provide access to the resource.
- InvalidResourcePropertyQNameFault: Used to indicate that resource property specified in the request message did not correspond to a resource property element of the WS-Resource referred to in the request message.
- *ResourceNotDestroyedFault*. Used to indicate that a WS-Resource was not destroyed for some reason.

⁵ For a detailed documentation about each WSRF specification, operation and fault, please see http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsrf

The following groups of properties are mandatory for a BioMOBY asynchronous service:

- Status properties:
 - This group represents the status of the jobs resulting from a batch-call. The batch-call itself is identified through the EPR, so the name of the property will be used to identify the job. The service must have a property called "status_queryID" for each job (input mobyData). For example, if a batch-call contains two jobs with queryIDs q1 and q2, these properties must exist: "status_q1" and "status_q2". The status will be given using the OMG's LSAE standard⁶ schema for Notification Events.
- Result properties:
 - This group represents the result from the particular jobs belonging to a batchcall. If a client asks for a result property before a particular job has finished, an WSRF Fault "InvalidResourcePropertyQNameFault" will be returned. The batch-call itself is identified through the EPR, so the name of the property will be used to identify the job. The properties are named in a similar way as the status properties but following the pattern "result_queryID".

Please note, that these properties must be read-only. WSRF specifies some optional operations that can be used to delete or modify such properties. The easiest way to not allow access to these operations is to simply not implement them (resulting in a SOAP client fault if a client tries to call them). However, if a service author wishes to implement these operations for modification of other properties of a batch-job not specified here, any attempt to change the properties discussed in this proposal must be stopped on the implementation level and an additional WSRF Fault should be returned by the service ("UnableToModifyResourcePropertyFault") signaling that these properties are read-only.

3.3. Registering asynchronous services



Services must be able to indicate to clients if their services can work in an asynchronous mode or not. Such information must be made part of the service registration procedure. We propose to add a Boolean parameter to the service registration API call (asynchronous).

For reasons of compatibility with synchronous clients, it is mandatory to provide a synchronous version of the service. As before, only the synchronous service will be registered in MobyCentral.

- If a service is provided in synchronous mode only (asynchronous=false), the service provider will implement just one SOAP operation that is named exactly like the service (e.g. doBlastAnalysis).
- If a service is provided in asynchronous mode (asynchronous=true), the service provider must implement the following SOAP operations (in addition to the synchronous operation used for synchronous mode):
 - For submitting (starting) a job:
 - The service name registered in MobyCentral with _submit appended (e.g. doBlastAnalysis_submit)
 - For polling the status of a job:
 - GetResourceProperty/GetMultipleResourceProperties for the property "status_qname", where qname is the queryID of the corresponding mobyData input.

⁶ Life Sciences Analysis Engine (LSAE) final adopted specification - <u>http://www.omg.org/cgi-bin/doc?dtc/2005-04-01</u> BioMOBY clients must be able to understand the pre-defined events in this specification: Heartbeat event, Percent progress event, Job State changed event, Step progress event, Time progress event. For more information, please check the LSAE specification.

- For retrieving results of a job/jobs:
 - GetResourceProperty/GetMultipleResourceProperties for the property/properties "result_qname", where qname is the query name of the corresponding mobyData input.
- For destroying a job:
 - *Destroy* for the batch-call identifier created during submission.

Please note that although the service provider must implement and publish these additional operations, *only* the synchronous version will be registered in MobyCentral. If asynchronous equals true during service registration, MobyCentral (the registry) will produce WSDL for the additional SOAP methods (and, of course, the service must implement these SOAP methods). Therefore, the knowledge if a service can be called in an asynchronous way must be stored in the registry.

Even if a service is asynchronous, it must always be possible⁷ to call it in a synchronous mode. Naturally, in this case it is possible (as before) that the connection between client and service is closed because of a timeout.

3.3. Determining if a service is asynchronous capable

A client discovers services by using the findService⁸ API call of BioMOBY. We suggest allowing a new value in the input field protocol and the output field category 'moby-async', adding to the normal values 'moby', 'cgi' and 'soap'.

In BioMOBY, clients receive a WSDL from MobyCentral by using the retreiveService operation from the BioMOBY API. Clearly, the operations relating to WSRF must also be described in the WSDL⁹ that MobyCentral generates.

The ability to call a service asynchronously can also be deduced from LSID resolution¹⁰ (returning RDF describing the service). Here, we suggest that this information should be stored in the form of a "hasCallingDetail" parameter, where one of the properties of a callingDetail includes whether it is asynchronous or not (remember, that according to the proposal, asynchronous also implies that there is a synchronous way to call the service).

An example as N3:

a :operation; :hasCallingDetail [a :callingDetail;

:hasSynchType moby:asynchronous];

An example as RDF:

<hasOperation rdf:parseType="Resource">

<rdf:type rdf:resource="http://www.mygrid.org.uk/ontology#operation"/>

- <hasCallingDetail rdf:parseType="Resource">
- <rdf:type rdf:resource="http://www.mygrid.org.uk/ontology#callingDetail"/>
- <hasSynchType rdf:resource="http://biomoby.org/RESOURCES/MOBY-S/ServiceDescription#asynchronous"/> </hasCallingDetail>

-

</hasOperation>

⁷ Note that the SOAP method should be available. If, however, the author of the service knows that the service will *never* finish before a time-out, it is reasonable to return an empty result directly together with a MobyException with error code 701 (Specific errors from the BioMOBY service) and the message "Service must be invoked asynchronously."

⁸ First, we note that the fields *protocol* and *category* fields are named differently but represent the same thing in the input and output. It would be better to keep a consistent naming. Second, the fact that a service is capable of asynchronous communication is *by itself* not a reason to choose this service over another synchronous-only service.

⁹ http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsrf

¹⁰ http://biomoby.open-bio.org/index.php/for-developers/moby_extensions/moby_metadata

3.3. Invoking asynchronous services – Communication sequence

The sequence for an asynchronous MOBY service execution is the following:

- A. The client starts the session
 - If the client is only capable of synchronous communication, the services (both sync and async) will work in synchronous mode (current BioMOBY behavior)
 - If the client and service are capable of asynchronous communication, the client sends the request message to the *servi cename_submit* SOAP operation to inform the service that they wish to run the service in asynchronous mode, and that it is able to cope with asynchronous calls. The message contains the standard MOBY content and data needed to execute the service.

If the service is synchronous, then it will construct the result and return it. However, asynchronous services will accept the task and return back a WSRF message (containing a batch-call identifier or "ticket" inside of an EPR), instead of the actual result. From this point both client and server will work in asynchronous modes.

B. Polling for service execution status

The client uses the **GetResourceProperty/GetMultipleResourceProperties** WSRF operation to retrieve status information using the batch-call identifier (in the EPR) and the queryID of the corresponding mobyData input. This step can be repeated as many times as needed until the service execution is done and the result is ready, using the SOAP operation to periodically request for execution status information. Additional notification information (i.e. "step 2: sequence alignment done") can be included to report progress status from the requested service to the client.

C. Request service execution result

When the client receives a status from the service that indicates that the execution is completed, the client can ask the service for the result by sending a message containing the asynchronous batch-call identifier and queryID to the **GetResourceProperty/GetMultipleResourceProperties** WSRF operation for the properties in the "result" group.

The response from this operation call will contain WS Resource properties whose formats are MOBY compatible and will include, any MobyExceptions that might have occurred¹¹. These properties are all available until the resource is no longer available. A resource becomes unavailable when the service either by its own initiative cleans the result or after a client explicitly asks that the resource is cleaned.

D. The client destroys the session

Once the client has retrieved the results of all jobs and do not wish to retrieve them again, it should destroy the session created during the step A. This is done by sending a message containing the EPR to the **Destroy** WSRF operation¹².

¹¹ Note that this means that if an error relating to Moby happens (for example, not properly formatted BioMOBY input or invalid object), then the proper way to handle this situation is the following: report a *status* using LSAE that signals that the job is finished (several options in LSAE for this) and then report a *result* as a normal Moby message with whatever MobyException that is suitable.

¹² It is possible that the service provider by his own initiative cleans the results of old batch-calls. This depends however on the policy of the service provider and this proposal does *not* specify any "minimum" time that the results should be available but clearly most providers might want to clean "old" results at some point. If the resource has been cleaned and a client request the resource then the service should respond with the WSRF Fault "ResourceUnknownFault" (since there no longer exists any trace of this resource/batch-call)



Sequence diagram for an async-async communication using BioMOBY

3.4. Asynchronous messages

Service execution request is the only message that will follow exactly the normal BioMOBY standard, the other messages will be WSRF messages.

Regarding error-handling: naturally, in all these SOAP calls it is possible that we get SOAP related faults but we do not specify these here.

When using the standard WSRF operations it is possible to get the standard WSRF faults. We give a general example in the appendix to show readers how such a fault might look like. In this section we list the documented WSRF faults in each situation. For more information about WSRF operations and faults, we again refer the reader to the official WSRF documentation.

The messages are as follows:

1. Requesting asynchronous service execution: This message is identical to the BioMOBY XML to request synchronous service execution. The only difference is that the client sends the request to the *servi cename_submit* SOAP method.

SOAP XML request for asynchronous service execution	(Unmodified)
<soap: envel="" ope=""></soap:>	
<suap. header=""></suap.>	
<pre><soap: body=""></soap:></pre>	
< <i>servi cename</i> submi t>	
<pre><moby: moby=" http://www. bi omoby. org/moby-s" ns:="" xml=""></moby:></pre>	
<moby:mobycontent></moby:mobycontent>	
<pre><moby: mobyData queryI D=' queryI d00 ></pre>	
<pre><!-- Standard BIOMOBY XML for Input--></pre>	
milbby moby Data	
<pre><iiioby and="" biomory="" for="" input="" xml=""></iiioby></pre>	
<td></td>	

- 2. Accepted asynchronous request: The service recognizes the asynchronous request, and communicates to the client that its request was accepted and that the service will work in asynchronous mode. For this accepted asynchronous request (and for the polling, and polling response operations), a standard WSRF message is sent, containing two important parts in the EPR:
 - ► Address: This is the address where the EPR is a valid reference to the resulting batch-call.
 - ReferenceParameters: The ticket representing the service provider identifier for the service execution job. The value of this ticket has no intrinsic meaning. The service provider can choose it be any legal XML fragment. Clients should not attempt to interpret the value of the ticket; it is simply an identifier and should remain opaque from the point of view of the client.

<pre><soap: envel="" ope=""> <soap: header=""> </soap:> </soap:></pre>	SOAP XML response for accepted asynchronous service execution (New)	
<pre> <soap: body=""> <servicename_submitresponse> <submitreference> <submitreferenceparameters> </submitreferenceparameters> <td><soap: envel="" ope=""> <soap: header=""></soap:></soap:></td><td></td></submitreference></servicename_submitresponse></soap:></pre>	<soap: envel="" ope=""> <soap: header=""></soap:></soap:>	
	<pre> </pre>	

This EndPointReference is included in the SOAP header in the subsequent messages and is used to refer to the batch-call.

3. Polling for service status: We assume a polling model where the client queries the service asking for the status of its request. The client makes the polling requests by sending a message to the GetResourceProperty or GetMultipleResourceProperties WSRF operations of the service. The structure of the polling message is a WSRF block with the asynchronous job "ticket" (in the soap header) and the original request query identifiers (properties status_queryID).



Valid fault messages are ResourceUnknownFault, ResourceUnavailableFault and InvalidResourcePropertyQNameFault.

4. **Response for polling request:** The response sent from the service includes the current process status for the requested query identifiers.





Valid fault messages are ResourceUnknownFault, ResourceUnavailableFault and InvalidResourcePropertyQNameFault.

5. Requesting the result: Once the client knows that the service execution is finished and that the result is ready, it should retrieve results by sending a message to the GetResourceProperty or GetMultipleResourceProperties WSRF operations in the server. The message structure is the same as in the step before, but in this case asking for resul t_queryI D properties.





Response for requesting the result: The content of the requested resul t_queryI D
properties is a standard BioMOBY response message containing the result of the
service execution. Note that if a BioMOBY related error occured during the execution,
the BioMOBY response will contain empty mobyData with corresponding
mobyException.







It will be a client task (at API level) to compose a standard BioMOBY message from the WSRF response.

Valid fault messages are ResourceUnknownFault, ResourceUnavailableFault and InvalidResourcePropertyQNameFault.

7. Destroying the resource for asynchronous service execution: After client has retrieved the results of all query identifiers, it should destroy the resource it was created during asynchronous service execution request by sending a message to the Destroy WSRF operation in the server. The structure of this message is a WSRF block with the asynchronous job ticket (in the soap header).



8. Response for destroying the resource:



Accepted fault messages are ResourceUnknownFault, ResourceUnavailableFault and ResourceNotDestroyedFault.

APPENDIX A – WSRF Faults

In this appendix we give brief example of a WSRF Fault:

WSRF fault (Ne	w)
<soap: envel="" ope=""></soap:>	
<soap: header=""></soap:>	
<pre> <wraketype:section content="" of="" se<="" section="" td="" the=""><td></td></wraketype:section></pre>	
 <soap: body=""> <soap: faul="" t=""> <faul tcode="">soap: Server</faul> <faul tstring="">Application error</faul> <detail> <wsbf: <i="">Faul tMessage></wsbf:></detail></soap:></soap:>	
<pre><wsbf: mestamp="" ti="">faul t_ti mestamp</wsbf:></pre>	

These errors come as standard SOAP Faults but with WSRF specific faults inside the details section. All WSRF faults are extended from the standard WSRF BaseFault using the XML Schema extension mechanism. For details, see WS-BaseFaults¹³.

 $^{^{13} \} http://docs.oasis-open.org/wsrf/wsrf-ws_base_faults-1.2-spec-os.pdf$