Table of Contents

Table of Contents

EMI Review of the EGI QC Documents (Version 2)

See also the review of version 1: EmiEgiQcReviewV1 whether your previous comments need to be reiterated or have been addressed.

Reviewers

A.Aimar, M.Cecchi, L.Field, P.Fuhrmann, J.White

Deadline for the review

- **25 July 2011**
- Write your feedback in the corresponding section below

Material to Review

UMD Quality Criteria: https://documents.egi.eu/document/364 ⧉

Direct links are in each section.

Compute Capabilities Quality Criteria (Marco Cecchi)

Direct link to document to review⧉

JOBEXEC_IFACE_1

- "The test suite must include tests for all the documented functions." maybe EGI itself should define a minimum set of primitives an interface should provide. - " Invalid output should throw an exception as documented" It is not clear whether the interface should be accessed directly through the network or via APIs. In both cases, the term exception could not apply everywhere, so I would speak in generic terms of "error". - EMI-ES is not listed among the possible interfaces.

JOBEXEC_JOB_2:

Typo: Non-empy -> Non empty

JOBEXEC_JOB_3:

Testing job cancellation involves different scenarios depending on where the job was found in the submission chain at the moment the cancellation is triggered. This test should possibly be broken down into subtests such as: cancel a job when the job in state (submitted, waiting, scheduled, running, already cancelled, aborted, etc).

JOBEXEC_EXECMNGR_1:

I would add: check if the middleware requires to have outbound connectivity on the WN and check if the job requires an identity switch. These criteria should not be blocking, but should be well known.

JOBEXEC_EXECMNGR_3:

A list of supported BS would be expected here, as for JOBEXEC_EXECMNGR_2.

INTERACTIVE_JOB_1

Providing interactive login to remote machine also involves specific configuration on the WN. That can only depend on the site administration policies. Of course we are not speaking of interactive root access to a virtual machine here. In any case, I'd relax the original statement into something like "provide interactive, shell-like access to a worker node". That could simply be done, for example, by redirecting stdin&out over a socket.

INTERACTIVE_JOB_3/INTERACTIVE_JOB_4

INTERACTIVE_JOB_4 encompasses what described in INTERACTIVE_JOB_3

JOBSCH_EXEC_1

EMI-ES is missing. In general, most of the above comments (JOBEXEC_IFACE_1, JOBEXEC_JOB_*) apply to JOBSCH_EXEC_* items as well.

JOBSCH_JOB_7

Submission of collection should be limited to a determined maximun number of nodes.

JOBSCH_JOB_8

DAG jobs should work for all the CEs supported by the metascheduler, not only for a subset. Also, the ability to support workflows (jobs with cyclic dependencies whose exit condition is to be evaluated at run-time) should be taken into account.

JOBSCH_WMS_1:

it is not clear in this context the meaning of "Input from Technology Provider: Test and for checking resubmission mechanism"

JOBSCH_WMS_2

"A test to submit a job and check if it is accepted or rejected, specially for big JDLs", repetead in JOBSCH_WMS_3 maybe it should be something about resubmission, i.e. the sentence in JOBSCH_WMS_1. In fact JOBSCH_WMS_3 correctly reports another: "A test to submit a job and check if it is accepted or rejected, specially for big JDLs"

Data Quality Criteria (Patrick Fuhrmann)

Direct link to document to review⧉

(With input from Oliver and Soonwook)

General remarks

- It might be just the phrasing but in some cases the section "Input from Technology Providers" refers to "Test to check ..." in other cases the documents reads "Test suite for ...". The technology providers (which translates to PT's in the EMI world I guess) are not supposed to provide test suites. Testing is done within the framework of the PT and is reported in order to get the packages certified by EMI. PT's are happy to provide information on what is tested but a test suite again would be a product and as such would undergo the same procedures as a 'normal' product and would have to be negotiated between EGI and EMI.
- As an example: 2.1.1 "It must include tests for all the documented functions." That is certainly envisioned but rather naive and an enormous effort. The LFC is just an example. We should make this a medium term goal. We should focus on 'most used' funcationality first. Please keep in mind that we all only have limited efforts which needs to be used in a focused way.

1.1 and 1.2

The sentence "Data Access Appliances must implement (at least one of) the OGSA-DAI and WS-DAI realizations and support all the functionality included in the interfaces" is not correct. Not all Data Access Appliances are supposed to provide those interfaces. The EMI Storage Elements are data access appliances but don't and won't those interfaces. Please rephrase.

2.2.2 Amga functionality : METADATA_AMGA_FUNC_2

- This is a mixing of creating entries and managing attributes in METADATA_AMGA_FUNC_2. We would suggest :
- Test1: Create a new entry. List entry.
- Test2: Create a new set of entries. List entries.
- Test3: Remove existing entry. List entry. (failed)

Generic Quality Criteria (Alberto Aimar)

Direct link to document to review⧉

GENERIC_DOC_4

Online help mandatory: Currently the online help or the --h command options is not mandatory but most command have it.

GENERIC_DOC_5

API documentation mandatory: Currently the API documentation is not a mandatory document of our documentation review. But of course all public APIs are documented.

GENERIC_DOC_1, _2, _3, _7

All OK. Functional Description, Release Notes, User Doc, Admin Doc are all also mandatory in the EMI Documentation Policy.

GENERIC_REL_1

Software Licenses are required and tracked by EMI. But what are the EGI compatible licenses for using them on the EGI Infrastructure? How can the TP know that?

GENERIC_REL_2

The part about clear and readable code is a bit generic. But all EMI code is publicly available.

GENERIC_REL_3, _4,

These are also in the EMI Policy.

GENERIC_REL_5

Similar requirement exists in EMI. The big question here is the granularity of testing. Now V2 contains mention of this comment.

GENERIC_REL_6

This is also an EMI requirement.

GENERIC_REL_7

This is new compared to Version 1. Should be a ticketing submission channel not a "bug tracker" where EGI submits bugs. EGI does not submit to the trackers of EMI but submits GGUS tickets and some are bugs others are request of clarifications, etc.

GENERIC_SERVICE_1

Service control and status commands. This is not a specified requirement for EMI services at the moment.

GENERIC_SERVICE_2

Log files. This is not a specified requirement for EMI services at the moment.

GENERIC_SERVICE_3

No such requirement is explicitly part of EMI policies. Nevertheless, Product Teams are required to perform scalability and performance tests as part of their product certification.

GENERIC_SERVICE_4

No such requirement is explicitly part of EMI policies. Nevertheless, Product Teams are required to perform scalability and performance tests as part of their product certification.

GENERIC_SEC_1

Not part of the current EMI requirements.But should be added to the EMI requirements.

GENERIC_SEC_2

Not part of the current EMI requirements.But should be added to the EMI requirements.

Information Capabilities Quality Criteria (Laurence Field)

Direct link to document to review⬈

In general these requirement are desribed in a very simplistic way. I would recommend that they are revised and written in more detail.

INFOMODEL_SCHEMA_1

The description states that "Information exchanged in the EGI Infrastructure must conform to GlueSchema". What information does this refer to? Any information that is exchanged or a subset?

In "Input from Technology Provider" it states, "Test that the information published by the product conforms to the GlueSchema v1.3 Technology and v2.0 (optionally)" However in "Pass/Fail Criteria" it states "Information published must be available in GlueSchema v1.3 and GlueSchema v2". This is a contradition.

INFODISC_IFACE_1

The description states that "Information published by the appliance must be available through LDAPv3 protocol". What is an applicance and what information does this refer to?

INFODISC_AGG_1

The description states what the software must not do. It should define what the software must do. Filtering out irrelevant information is not really a requirement. Provding only relvant information is a requirement.

In general I think that this requirement does not make sense and needs to be revised.

INFODISC_AGG_2

In general I think that this requirement does not make sense and needs to be revised. - --+++ INFODISC_AVAIL_1 The description states "Information Discovery appliances must be able to handle big amounts of data". How big is big? How many data sources, how much data, how many changes, in what data format etc.

This requirement is too simplistic and needs to be revised.

INFODISC_AVAIL_2

Description The description statsthat the "information discovery service should be able to handle load under realistic conditions." What are realistic loads?

For the Pass/Fail Criteria, where does the value of 20 come from? Is this realistic?

MSG_IFACE_1

In the Pass/Fail Criteria, it statest that either JMS 1.1 or AMQP must be supported. As far as I am aware the recomendation from EMI is to use STOMP.

Operations Capabilities Quality Criteria (Laurence Field)

Direct link to document to review⤴

In gerneral I have a concern about the style of this document. What is being described? Requirements, Tests , Sepcification etc. The current style seems to mix these concepts and as such makes it difficult to understand. The desciptions need to be improved to clear be inline with what is being described. From the phrasing it should be clear that this is a a test or a requirement. Everything should also be made more abstract. For example it should state that dataset A should be moved to point B by time T rather than requiring that cron is used to publish accounting information.

MON_NCG_1

More details are required for the Test Description. What information? Which database? How do we know it is working?

MON_NCG_2

The description does not sound like a requirement. The "NGI has to understand", isn't that and NGI issue?

More details are required for the Test Description. What information? Which database? Test contains no deatils now to test for redundency which seems to be the perpose of the test. How do we know it is working?

MON_PORTAL_1

Ok, but some phrasing could be improved for clarity.

MON_PORTAL_2

Ok, but some phrasing could be improved for clarity.

MON_PORTAL_3

Ok, but some phrasing could be improved for clarity.

MON_PORTAL_4

Ok, but some phrasing could be improved for clarity.

MON_PORTAL_5

How fast is fast, how soon is soon, how much is too much? Please be more specific.

MON_DB_1

Ok, but some phrasing could be improved for clarity.

MON_PROBE_JOBEXEC_1

"Pass/Fail" Check sentence.

MON_PROBE_JOBEXEC_2

"Pass/Fail" Check sentence.

MON_PROBE_JOBEXEC_3

"Pass/Fail" Check sentence.

MON_PROBE_JOBSCH_1

"Pass/Fail" Check sentence.

MON_PROBE_METADATA_1

"Pass/Fail" Check sentence.

ACC_JOBEXEC_1

All the actions??? Are you sure?

ACC_JOBSCH_1

All the actions??? Are you sure?

Security Capabilities Quality Criteria (John White)

Direct link to document to review⧉

Is this a "Quality" document describing the quality of software or a set of requirements? It reads more like a set of requirements. The test requirements themselves look reasonable and pretty much mirror the test suites that are used in EMI certification of security components.

One question that comes to mind is that is EGI expecting a new set of test-suites to run in a particular framework or will they use the results of EMI certification? Statement that the document as "OK" does not imply acceptance of re-writing/reformatting our tests for EGI.

ATTAUTH_ WEB_2

This follows the previous (EGEE/LCG) policy documents? If so, OK.

AUTHZ_ PDP_1

"PDPs must support the XACML interface" This is a bit general. Internally Argus uses XACML but the WHOLE XACML spec is not used.

Storage Capabilities Quality Criteria (Patrick Fuhrmann)

Direct link to document to review⬀

(With input from Oliver K.)

General remarks

- Same remark as wtih 'DATA' : It might be just the phrasing but in some cases the section "Input from Technology Providers" refers to "Test to check ..." in other cases the documents reads "Test suite for ...". The technology providers (which translates to PT's in the EMI world I guess) are not supposed to provide test suites. Testing is done within the framework of the PT and is reported in order to get the packages certified by EMI. PT's are happy to provide information on what is tested but a test suite again would be a product and as such would undergo the same procedures as a 'normal' product and would have to be negotiated between EGI and EMI.
- Just to avoid misunderstandings : Different storage software providers in EMI provide different file access mechanisms at different times within EMI-2 and later. (Especially FILETRANS_API_3). Consequently, testing a capability must not be applied before the software provider adds this capability to the release notes as being available for production usage.

2 File Access (a remark on POSIX)

- Typo : The ID : FILEACC_API_1 is used twice.
- Here as well it is important to only check capabilities which are described as available. So some storage elements may not allow modification of existing data (as it is on tape already). Same is true for 'append' as described in the "Input of the Technology Providers" of FILEACC_API_2 (which is actually the second occurrence of FILEACC_API_1)
- FILETRANS_API_2 : Although we expect FTS to support http(s) at some point, this is not yet agreed and must only be tested if officially declared 'available for production usage'.
- FILETRANS_API_3 : same as FILETRANS_API_2.

5.1 SRM interface STORAGE_API_1

The bit with SRM is tricky. The sentence "Execute a complete test suite for the SRM v2.2 that covers all the specification" has two requirements :

- Test complete Test Suite: There is a group of storage providers covering CASTOR, DPM, BestMAn, StoRM and dCache which make sure the existing SRM infrastructure doesn't break. This goups agreed to use the S2 Test as a basis for this effort. So I would strongly suggest, EGI uses the official S2 Test Suite provided by that group. (Please talk to Andrea Sciaba, CERN for details)
- All the specification: That's a bit naive. Each SRM provider will offer a set of tests which the corresponding storage element can be tested against. I don't want to discuss this further in this wiki. Here again, EGI should contact the SRM Test Group leader (Andrea)
    - ♦ STORAGE_API_2 "Related Information" already refers to this issue.

Incomplete sentence in Pass/Fail criteria : Exceptions to the specification may (What ?).

5.2 STORAGE_DEVICE_1

- Ok if you are talking about BDII and Glue1.3/2.0

This topic: EMI > EmiEgiQcReviewV2
Topic revision: r11 - 2011-07-28 - PatrickFuhrmann