



EGI.eu Software Provisioning

FINDINGS OF DRY RUNS USING EMI COMPONENTS

Document identifier	EGI-SA2-Dry Run results-V3.docx
Document Link	https://documents.egi.eu/document/479
Last Modified	28/04/2011
Version	3



COPYRIGHT NOTICE

Copyright © EGI.eu. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

The work must be attributed by attaching the following reference to the copied elements: "Copyright © EGI.eu (www.egi.eu). Using this document in a way and/or for purposes not foreseen in the license, requires the prior written permission of the copyright holders. The information contained in this document represents the views of the copyright holders as of the date such views are published.

I. DOCUMENT LOG

Issue	Date	Comment	Author/Organization
1.0	27/04/2011	First draft	Michel Drescher, EGI.eu
2.0	27/04/2011	Second draft	Michel Drescher, EGI.eu
3.0	28/04/2011	Third draft	Michel Drescher, EGI.eu
4.0			



II. ORGANISATION SUMMARY

To support science and innovation, a lasting operational model for e-Infrastructure is needed – both for coordinating the infrastructure and for delivering integrated services that cross national borders. The objective of EGI.eu (a foundation established under Dutch law) is to create and maintain a pan-European Grid Infrastructure in collaboration with National Grid Initiatives (NGIs) in order to guarantee the long-term availability of a generic e-infrastructure for all European research communities and their international collaborators.

In its role of coordinating grid activities between European NGIs, EGI.eu will:

- Operate a secure integrated production grid infrastructure that seamlessly federates resources from providers around Europe
- Coordinate the support of the research communities using the European infrastructure coordinated by EGI.eu
- Work with software providers within Europe and worldwide to provide high-quality innovative software solutions that deliver the capability required by our user communities
- Ensure the development of EGI.eu through the coordination and participation in collaborative research projects that bring innovation to European Distributed Computing Infrastructures (DCIs)

The EGI.eu is supporting ‘grids’ of high-performance computing (HPC) and high-throughput computing (HTC) resources. EGI.eu will also be ideally placed to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit the user communities within the European Research Area.

EGI will collect user requirements and provide support for the current and emerging user communities. Support will also be given to the current heavy users of the infrastructure, such as high energy physics, computational chemistry and life sciences, as they move their critical services and tools from a centralised support model to one driven by their own individual communities.

The EGI community is a federation of independent national and community resource providers, whose resources support specific research communities and international collaborators both within Europe and worldwide. EGI.eu, coordinator of EGI, brings together partner institutions established within the community to provide a set of essential human and technical services that enable secure integrated access to distributed resources on behalf of the community.

The production infrastructure supports Virtual Research Communities – structured international user communities – that are grouped into specific research domains. VRCs are formally represented within EGI at both a technical and strategic level.



III. TABLE OF CONTENTS

1	Introduction.....	5
2	First dry run.....	6
2.1	Preparations.....	6
2.2	Dry run execution.....	6
2.3	Results.....	7
2.3.1	Issue 1: Child tickets for StagedRollout not automatically created	8
2.3.2	Issue 2: Encoding of large amounts of free text as XML CDATA	8
2.3.3	Issue 3: Fixed short names for products, per Technology Provider.....	8
2.3.4	Issue 4: Verification document(s) not publicly accessible	8
3	Second Dry Run	9
3.1	Preparations.....	9
3.2	Results.....	10
3.2.1	Issue 5: EMI distinction between “products” and “internal component”	10
3.2.2	Issue 6: Implications of definition of internal components on update policy	10
3.2.3	Issue 7: Products, internal components and meta-packages	10
3.2.4	Issue 8: Violation of EGI software provisioning policy	10
3.2.5	Issue 9: Two products covered by the same meta-package.....	11
3.2.6	Issue 10: Two meta-packages provided for one product.....	12
3.2.7	Issue 11: Incomplete UMD Release Plan	12
4	Conclusions	14
5	References.....	15



1 INTRODUCTION

This document summarises the activities that have taken place in two dry runs of the EGI Software Provisioning Workflow, using both times components from the upcoming EMI-1 major release.

The document is divided into three parts:

Section 2 covers the first dry run using three components from EMI (VOMS, VOMS-Admin and ARC-CE) to validate the applicability of the EGI Software Provisioning dry run.

Section 3 covers the second dry run that was planned to reassure the efforts that are necessary to sufficiently verify EMI software components. The collected effort figures then were planned to feed into a realistic draft of the UMD Release Plan before it is published.

Section 4 draws conclusions from the previous sections, and calls for further action from all Technology Providers to at least provide constructive input on the presented issues.



2 FIRST DRY RUN

The goals of the first dry run of EMI components through the EGI Software Provisioning workflow were:

1. Applicability of the Software Provisioning process at large
2. Detect bottlenecks within the EGI Software Provisioning process
3. Detect bugs, ambiguities and other issues in the EGI Software Provisioning process
4. Assure that EMI as a key Technology Provider for existing infrastructure understands the process and is enabled to provide all the necessary information.

EMI were asked to identify three EMI products that should be used for executing the dry run. The following products were identified:

- VOMS 2.0 RC3
- VOMS-Admin 2.6.1 RC3
- ARC CE 1.0.0 RC3

Once the products were identified the first dry run was executed. It started on 8 April 2011, and is still on going with ARC-CE in the dry run's StagedRollout phase.

2.1 Preparations

At the start of the dry run the GGUS submission infrastructure was not available.

Also the module that analyses the submitted release.xml artefact and creates product-specific repositories for Verification and StagedRollout was delayed.

Therefore, these steps were manually conducted, and the resulting RT tickets were created by hand with all the necessary information:

- VOMS 2.0 RC3: <https://rt.egi.eu/rt/Ticket/Display.html?id=1728>
- VOMS-Admin 2.6.1 RC3: <https://rt.egi.eu/rt/Ticket/Display.html?id=1727>
- ARC-CE 1.0.0 RC3: <https://rt.egi.eu/rt/Ticket/Display.html?id=1726>

2.2 Dry run execution

As indicated in the previous section, the dry run effectively started with the verification of the products against the Quality Criteria that apply to the respective products.

As much as possible, all products were verified according to the process, as if it were to be included in a UMD release. The process is defined in: https://wiki.egi.eu/wiki/EGI_Quality_Criteria_Verification

VOMS, and VOMS-Admin were formally rejected due to packages missing from the EMI-1 RC3 repository. The relevant verification reports are permanently available at:

- VOMS 2.0 RC3: <https://documents.egi.eu/document/467>
- VOMS-Admin 2.6.1 RC3: <https://documents.egi.eu/document/469>



The initial verification of ARC-CE 1.0.0 RC3 confirmed a bug in the meta-package file provided for this product: A package was referenced in this Meta-Package which

- a) Was not present in the RC3 repository, and
- b) Was mistakenly present in the meta-package for ARC-CE 1.0.0

After resolving this issue, ARC-CE 1.0.0 RC3 was accepted and passed on to Staged Rollout. The Verification Reports for ARC-CE 1.0.0 RC3 are permanently available at <https://documents.egi.eu/document/468>. StagedRollout for ARC-CE 1.0.0 RC3 did not conduct a near-real life test – it just ensured that the release notes for ARC-CE were present, according to the initial objectives of the dry run.

2.3 Results

This dry run revealed issues that were not known beforehand. However, it validated the Software Provisioning workflow as such, and highlighted the usefulness of particular part of the workflow, which was until then questioned. The following gives a brief analysis of the initial objectives compared to the findings outlined above.

1. Applicability of the Software Provisioning process at large

With the exception of the infrastructure parts of the process that were not yet available at that time, the process has been successfully validated. Those parts of the process infrastructure that are available and functional worked as expected: The repository provided product specific areas for easy verification and Staged Rollout. The RT tickets handling the progress and coordination of responsibility and ownership in the overall process worked as expected, and detailed information is available in the RT ticket, and Document database, for review by any interested reader such as Quality Assurance teams, DMSU members, Site Admins, etc.

2. Detect bottlenecks within the EGI Software Provisioning process

No bottlenecks were discovered. Verification of products may be conducted in parallel, reducing the time to publication dramatically. Likewise, the process is designed to support parallel StageRollout, given that sufficient sites take on the role of early adopter for software components that are in StageRollout.

3. Detect bugs, ambiguities and other issues in the EGI Software Provisioning process

Several issues were detected during the dry run. They are collected below (sections)

4. Assure that EMI as a key Technology Provider for existing infrastructure understands the process and is enabled to provide all the necessary information.

This objective was not met, as a release.xml artefact was not provided by EMI. The EMI project was – and still is at the time of writing – occupied with the preparation of the final EMI-1 release. Nonetheless the understanding of the release delivery concept was not demonstrated and may pose a significant risk for delay for future releases.

The following sections discuss individually all issues that fall under objective 3, for better identification and readability.

2.3.1 Issue 1: Child tickets for StagedRollout not automatically created

One issue was discovered during the dry run: The provisioning process is used for both external Technology Providers such as EMI, IGE and SAG, and internal Technology Providers, currently the EGI Trust Anchor “provider” and EGI-InSPIRE JRA1 for SAM. The provisioning process had to be changed from being release based to product based, i.e. being able to accept a set of product updates while rejecting another set of products within the same release, for external technology providers while it had to stay the same for internal Technology Providers. During this transition StagedRollout requested more information to be passed on from the parent ticket into the StagedRollout child ticket. This last minute request was accommodated for at last minute, but unfortunately introduced a bug so that RT tickets modelling and coordinating StageRollout activities for a Product undergoing the provisioning process in a separate StageRollout queue are not automatically created anymore. This issue is currently under investigation and a fix is expected soon.

2.3.2 Issue 2: Encoding of large amounts of free text as XML CDATA

Another issue relates to handling larger amounts of texts, such as release notes and change logs. For the purpose of a permanent record, the release.xml artefact is designed to contain verbatim copies of release notes and change logs of each product declared in that artefact. That type of text most likely contains characters such as “<”, “&” etc. that break any XML document validation. Therefore the relevant sections need to be redefined as CDATA embedded in the release.xml document.

2.3.3 Issue 3: Fixed short names for products, per Technology Provider

For each product declared in the release.xml file, two fields are provided to describe the product: One field for a long product name (“productName”), and one short name field, providing a technical code name for a product (“productShortName”). While the product name is used for display purposes for the end user, the product’s short name is a technical identifier used verbatim as an identifier in the provisioning process infrastructure (i.e. managing the product-specific repositories provided for verification and StagedRollout). Technology Providers must pay special attention to the use of this field, in that:

- The product short name for the same product (e.g. VOMS) does not change for the lifetime of this product
- That the product short name allows for numbers, but that product short names “egi-voms1” and “egi-voms2” would result in the provisioning of two separate products, instead of two major versions of one and the same product.
- The product short names must be unique per Technology Provider (i.e. two different technology provider may use the same product short name) – but this is strongly discouraged.
- For technical reasons the allowed characters for product short names are (in regular expression notation): [aA-zZ][0-9][-,_]

2.3.4 Issue 4: Verification document(s) not publicly accessible

Despite being documented in the process descriptions for Verification, and StagedRollout, the Verification documents for the ARC-CE 1.0.0 RC3 component were not configured in EGI’s Document Database for public access and consumption. This is a manual step, and needs education and training into routine behaviour of the verification officers. However, this issue is considered non-critical.



3 SECOND DRY RUN

During the EGI Userforum 2011 in Vilnius, Lithuania, EGI SA2 together with EGI TSA1.3 decided to conduct a second dry run with EMI components, based on the results of the first dry run and the need to collect reassurance in the necessary efforts to conduct the verification of components of EMI-1.

Without this reassurance it is difficult to provide a realistic estimate, or plan, of how many EMI-1 components may be provided as part of UMD 1.0. Additionally, the OMB met face to face in Vilnius [R 1] produced a prioritised list of EMI-v1 components that the OMB felt were required to be present in UMD 1.0 to ensure continuous support and provisioning of the EGI production infrastructure with product updates and security patches.

Also, the OMB placed an additional constraint on the publication date of UMD 1.0 to the end of June since critical components of gLite 3.1 reach their respective end of support cycle at that date.

Therefore, the second dry run's objectives were different from the first dry run:

- Based on basic effort figures, and the OMB's prioritised component list, provide a first draft of the UMD Release Plan.
- Based on this UMD Release Plan, build a list of components comprising this second dry run.
- Use as much information as possible from the EMI Release Tracker [R 2] to build the required metadata
- Reverse-engineer an appropriate release.xml artefact into the form that EGI would expect had EMI provided this artefact by themselves, to provide a practical and near-real life example for EMI to use as a reference
- Collect effort information for the Verification, and StagedRollout of components of the EMI-1 major release.
- Extrapolate, based on the extracted effort figures, a realistic second draft of the UMD Release Plan.

3.1 Preparations

The preparations for this second dry run were interrupted by informal reviews of the first (non-public) UMD Release Plan, which was considered unacceptable.

A second iteration of the UMD Release Plan now contains 21 components (of which some are EMI products, some are internal components, see below) planned to be included in UMD 1.0, to be published on 4 July 2011. This list of components was informally reviewed by Cristina Aiftimei and considered as good to go for a second dry run. EGI and EMI agreed that EMI-1 RC4 was a good basis for this second dry run.

The required artefacts were generated as far as possible without having access to the RC4 repository, waiting for the final "green light" to proceed with the dry run.

However, as of the time of writing the official "green light" for starting this dry run has not been given by Cristina Aiftimei, the EMI Release Manager.

As available time is running out, EGI SA2 decided to cancel the dry run, and publish the collected results based on the preparations and information available from the EMI Release Tracker, and RC3 repository.

3.2 Results

From the discussions occurring during the preparation of the second internal draft of the UMD Release Plan, and the actual preparations of artefacts for the second dry run, the following issues were detected. Naturally, as the dry run was eventually cancelled, this list of issues focuses on the assembly of the release.xml and the available information from the EMI Release tracker.

The numbering of issues will continue numbering for easy identification and discussion across dry runs.

3.2.1 Issue 5: EMI distinction between “products” and “internal component”

It is unclear from the EMI release tracker which entry represents a “product”, and which entry represents an “internal component”. For example, VOMS [R 3] and CREAM [R 4] are reasonably safe to be considered products, while LCMAPS [R 5] is considered an “internal product” [R 8]. Nothing in the EMI Release Tracker, at least nothing obvious to EGI, indicates which of the tracker entries must be considered an internal component, and which entries in the tracker must be considered internal components.

3.2.2 Issue 6: Implications of definition of internal components on update policy

If a software component is considered an internal component, then this raises the question on the visibility of such components in an official public tracker for consumption of EMI customers. Also, this influences the update plans for such internal components – if they are internal, then to a customer, an update of an internal component should only manifest as an update to the affected product(s) that actually use the updated internal component (see below).

3.2.3 Issue 7: Products, internal components and meta-packages

It is the impression of EGI SA2 and TSA1.3 (i.e. StagedRollout) that EMI is providing meta-packages only for EMI products, but not for internal components, even though internal components are actively tracked in the EMI release tracker. For example, CREAM, VOMS and LCMAPS are tracked as active entries in the EMI Release Tracker, whereas only the first two (i.e. VOMS and CREAM) are EMI products and therefore provide a meta-package for installation. Consequently, those meta-packages reference the necessary packages for LCMAPS as both products use LCMAPS.

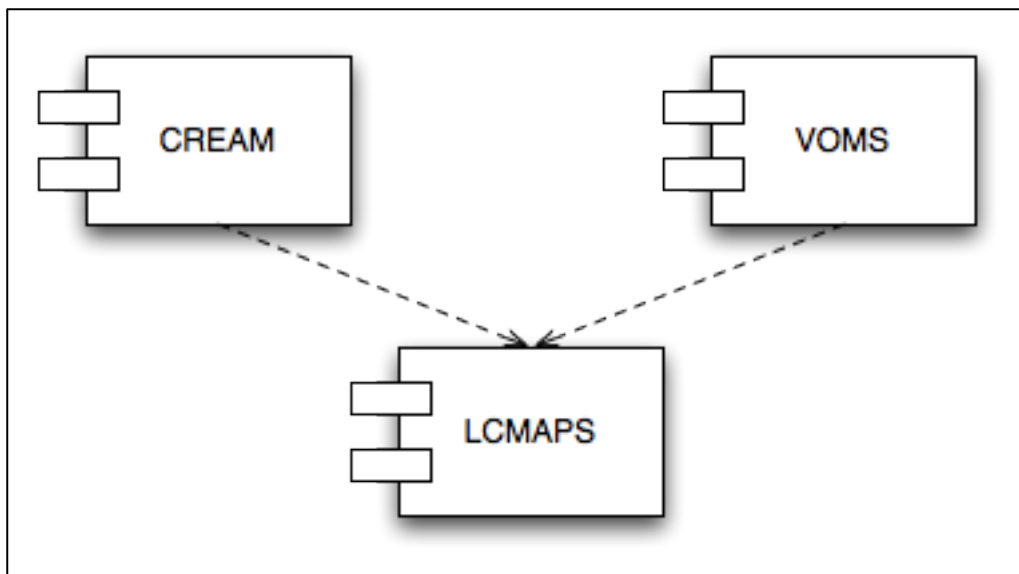
Again, this raises the question why LCMAPS is tracked as, and therefore exposed as, a component on eye-level with VOMS and CREAM.

3.2.4 Issue 8: Violation of EGI software provisioning policy

EGI’s policy on software provisioning clearly requires that any product considered for inclusion into the UMD **must** undergo Verification and StagedRollout. However, EMI’s current model of providing meta-packages only for products (e.g. VOMS, CREAM) but not for internal components (LCMAPS) creates a loophole to circumvent EGI’s provisioning policy. Figure 1 illustrates a typical component dependency within EMI-1. Currently EMI provides meta-packages only for products such as VOMS and CREAM, but not for LCMAPS as an internal component. Therefore updates of products such as CREAM and VOMS would undergo proper EGI software provisioning procedures – but not LCMAPS as there is no meta-package provided for LCMAPS.

The consequence would be that

- a) The EGI Software Provisioning process (which is products based) would ignore those updates of internal components such as LCMAPS until an update of either of the depending products (e.g. VOMS) is provided, or
- b) These components would be downloaded by sites offline, effectively bypassing the EGI Software Provisioning process, or
- c) Preventing either a) or b) EGI would have to manually create parts of, or a complete, release.xml artefact to be able to properly provision the updated internal component for



inclusion in a UMD update – but at significant effort and costs.

Figure 1: Exemplary component dependencies in EMI-1

Therefore EGI requires EMI – and any other Technology Provider in that respect – to provide a proper meta-package for each component that EMI plans to update either as part of a major release, or independently in between major EMI updates.

For the concrete example of CREAM, VOMS and LCMAPS, this would require the creation of a meta-package for LCMAPS, and changing the meta-packages for CREAM and VOMS to reference the LCMAPS meta-package, instead of the LCMAPS packages directly.

This approach provides the added value that component dependencies, as exemplified for VOMS, CREAM and LCMAPS in Figure 1, are far better manageable for and within EMI. Likewise, the whole software provisioning process and effort for EGI will be much more manageable once all “updatable” components are covered by meta-packages, and component dependencies are provided along the lines of meta-packages.

3.2.5 Issue 9: Two products covered by the same meta-package

While analysing the contents of the EMI Release Tracker [R 2] in order to prepare for the second dry-run, EGI discovered that both VOMS 2.0.0 [R 3] and VOMS-Admin 2.6.1 [R 7], though maintained as separate products, in fact are referenced by the same meta-packages (*emi-voms-mysql-1.0.0-*



1.sl5.x86_64.rpm and *emi-voms-oracle-1.0.0-1.sl5.x86_64.rpm*). A manual verification of those meta-packages gave evidence that indeed all packages of VOMS, and VOMS-Admin, were referenced by those meta-packages.

As a consequence, VOMS and VOMS-Admin can only be updated at the same time, since both products are (from a package maintaining perspective) indistinguishable from each other.

EGI therefore requires either the separation of VOMS and VOMS-Admin into two separate products and meta-packages (with the VOMS-Admin meta-package referencing the VOMS meta-package as a dependency), or the fusion of both into one maintained, updated product with only one product name (e.g. VOMS).

3.2.6 Issue 10: Two meta-packages provided for one product

During the analysis of the EMI Release Tracker two products were noticed in that they are maintained as one product, but with two different meta-packages depending on which persistence product (MySQL or Oracle) is selected at installation time. The products in question are VOMS 2.0.0 [R 3] (which also has other issues, see above), and LFC 1.8.1 [R 6].

From a deployment and software architecture perspective, the selection of a persistence solution should be either a configuration detail, or there are profound reasons to tightly integrate with a concrete persistence product (such as MySQL or Oracle) that in turn justifies the maintenance of two different products, for example “VOMS for MySQL” and “VOMS for Oracle”. The first option requires more careful Verification and StagedRollout on the EGI side as the various options are implied in the configuration options – unless it is unambiguously clear that the selection of the RDBMS truly is a configuration detail and behaviour of the application is guaranteed to be identical no matter which supported solution is configured. However, experience shows that this is in fact holds true only for few cases where the relational structure of the persisted data is reasonably straight forward, and manifests identical in all supported RDBMS. The second option provides an explicit statement that the choice of the backing RDBMS may in fact influence application behaviour – or a significant difference in the code base is evident to prevent such difference in behaviour.

EGI prefers the second alternative, making an explicit statement of alternative flavours of the same application, providing tight integration with the respective RDBMS of choice.

3.2.7 Issue 11: Incomplete UMD Release Plan

Based on the issues described above, and the shift from product specific repositories used in gLite, a surprisingly large amount of insecurity and inaccuracy in developing the UMD Release Plan is evident. It is not clear which EMI Release Tracker entries are internal components and which are products. Basing this decision on the provision of a meta-package proved unreliable, as not all what is considered a product actually provides a meta-package (e.g. Hydra), or more than one meta-package (e.g. VOMS).

Therefore the current draft of the UMD Release Plan lists for UMD 1.0 21 entries of EMI Release Tracker entries, of which EGI believes are EMI products. But in fact, the contents of UMD 1.0 will cover many more components of EMI-1, as indicated in Issue 3 (section 3.2.3) and Issue 4 (section 3.2.4). To provide a much more complete picture to the primary customers of the UMD Release Plan,



the Site Administrators deciding on when to upgrade to which component, EGI requires much more explicit information along component dependency trees than is currently available.



4 CONCLUSIONS

EGI conducted one dry run of running three EMI components that are planned to be part of the upcoming EMI-1 release (see section 2). A second dry run planned with significantly more components from EMI-1 was planned and prepared, but cancelled due to time constraints (see section 3).

The first dry-run revealed mainly issues surrounding the EGI Software Provisioning process, for example child tickets in RT not being automatically created (section 2.3.1). The most important issue coming out of this first dry run is issue 3 (section 2.3.3), requiring effort and diligence across all Technology Providers, and therefore some more discussion until a satisfactory solution is assured.

The second dry run, only planned but never conducted, involved a much deeper analysis of the EMI Release Tracker and its contents, and the implications on the EGI Software Provisioning process and policy. It provides a deep insight in the current status of the EMI Release Tracker, how it is used to manage and track the progress of all components that are developed and maintained by EMI Product Teams.

These issues (issue 5 to issue 11) pose more fundamental questions on release management and deployment management with relevance not only to EMI, but in fact all Technology Providers that engage with EGI.eu. Documenting the current misalignment on the technical level between EGI and EMI as manifested in the issues described above, this document intends to use the findings as a concrete example to drive EGI.eu's roadmap towards sustainable large-scale product based software provisioning sourcing technology from more than one Technology Provider.

While issues 1 to 4 are partially EGI internal issues, and partially issues that require only limited technical efforts, issues 5 to 11 require a detailed discussion across all Technology Providers. Having said that, the target audience for this document is perhaps best sought within the members of the TCB, rather the two task forces on Integration and Automation, and to a lesser extend the task force on Software Quality.

5 REFERENCES

R 1	OMB Face to Face meeting at the EGI User Forum 2011 in Vilnius, Lithuania, https://www.egi.eu/indico/conferenceDisplay.py?confId=267
R 2	Official EMI Release Tracker: https://savannah.cern.ch/task/?group=emi-releases
R 3	EMI Release Tracker entry for VOMS 2.0.0 https://savannah.cern.ch/task/?18587
R 4	EMI Release Tracker entry for CREAM 1.13 https://savannah.cern.ch/task/?18534
R 5	EMI Release Tracker entry for LCMAPS 1.4.29 https://savannah.cern.ch/task/?18706
R 6	EMI Release Tracker entry for LFC 1.8.1 https://savannah.cern.ch/task/?18687
R 7	EMI Release Tracker entry for VOMS-Admin 2.6.1 https://savannah.cern.ch/task/?18702
R 8	Skype chat between Balazs Konya and Michel Drescher on Tuesday 26 April 2011