**EGI-InSPIRE**

**VO monthly A/R statistics**

|  |  |
| --- | --- |
| Document Link | https://documents.egi.eu/document/<DOCID> |
| Last Modified | 19/11/2012 |
| Version | 0.1 |
| Contact Person | Diego Scardaci |
| Document Status | Draft |

|  |
| --- |
| Abstract |

1. Copyright notice

This work by members of the EGI-InSPIRE collaboration is licensed under a Creative Commons Attribution 3.0 Unported License (see a copy of the license at <http://creativecommons.org/licenses/by/3.0>). This license let you remix, tweak, and build upon this work, and although your new works must acknowledge EGI.eu, you do not have to license your derivative works on the same terms. Reproductions or derivative works must be attributed by attaching the following reference to the copied elements: “Based on work by members of the EGI-InSPIRE collaboration used with permission under a CC-BY 3.0 license (source work URL: specify if known)”.

1. Authors list

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Name** | **Partner/Activity/Organisation/Function** | **Date** |
| **From** |  |  |  |

1. Document Log

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Comment** | **Author/Organization** |
| **1.0** |  |  |  |
| **2.0** |  |  |  |

1. ACKNOWLEDGMENT

This work is partially funded by <project, organisation>

TABLE OF CONTENTS

introduction 4

1 Solution 5

1.1 Which VOs? 5

1.2 VOs aggregation and resources A/R 5

1.3 Algorithms 5

2 References 7

Appendix A Numbered appendices 8

A.1 Appendix level 2 8

A.1.1 Appendix level 3 8

# introduction

This document described the design of the solution to provide the VO monthly A/R statistics as described in the RT4348: <https://rt.egi.eu/rt/Ticket/Display.html?id=4348>.

The text of the original requirement is below reported:

*At the PY2 review we got the comment that availability statistics should reflect more the experience of a VO.*

*I would like you together with SAM team and JRA1 to do a feasibility study of an A/R reporting system in the operations portal according the following guidelines:*

* *VO specific A/R reports are generated monthly, using the OPS VO (yes, this is a limitation, but we cannot rely on the existing of a VO SAM instance)*
* *Only reports for "high activity" VOs are generated. With high activity we can follow the definition of the accounting portal, i.e. high activity week if*
* *CPU time consumed > 1 year/week*
* *CPU time consumed/month > 4 year/week*
* *the list of service end-points supporting VO x needs to be extracted from top-BDII (e.g. WMS, LFC, DPM, CE....)*
* *availability results are summarized on hourly, daily, weekly and monthly basis as we do for top-BDII*
* *the A/R calculation algorithm needs to be exactly defined, but it could be something like that: (OR CEs) AND (OR DPMs) AND (OR LFCs) .... AND (OR WMS)*

*Can you study the feasibility of this?*

# Solution

In this section we describe the technical details of the proposed solution.

## Which VOs?

We agreed to generate only reports for "high activity" VOs. With high activity VO we decided to follow the definition of the accounting portal, i.e. high activity week if

* CPU time consumed > 1 year/week
* CPU time consumed/month > 4 year/week

Although the list of the “high activity” VOs as above defined is a quite dynamic information (we observed big fluctuation in the past), we decide to query the accounting portal each quarter to identify the VOs list. We have to consider this solution as an approximation that could be improved in the future.

* SA1 + Accounting Portal: please, complete this section putting the technical details to get the list of “high activity” VOs as above defined

## VOs aggregation and resources A/R

We chose the ops portal internal DB (the one used in the Resources Browser module) as data source for the VOs aggregation with the caveat of cross-checking the monitored services in GOCDB. Indeed, you can have services in BDII with no OPS monitor tests configured.

Now, in the operation portal, we can get the list of VOs supported for each registered resources. We have to reverse this view to get the list of resource associated to each VO. The best frequency to perform the query has to be identified.

Then, we will query SAM to get the A/R for each resource associated to the VOs.

* Ops Portal please provide input on technical details about how to do the VOs aggregation
* Ops Portal + SAM, frequency SAM tests query to be agreed, hourly or daily?

Ops Portal contribution:

* Establish the high activity VO list . This list should be provided by the Accouting Portal in a automatic way (e.g xml feed)
* Foreach VO of this list query the BDII and extract a list of services for it
* Filter this list with the list of the type of services to be consider (CE , SRM , LFC , WMS , ??)
* Filter the list of services with the ones computed by SAM ( http://grid-monitoring/atp/api/search/vofeeds/json?json&vo=ops&is\_monitored=on )
* Foreach service in the list - Connect to SAM and retrieve A/R data.
* Compute the data with the algorithm based on a daily basis .

Limitations :

Bdii publication is not always accurate . The list of services will change in the time and sometimes some services will not be published properly.This part is very delicate and we will probably cause some distortions in the results of the computation. It seems that the WMS publications are not very reliable .

The computation of A/R based on the ops vo will be completely false for some services like LFC , WMS (which could be ok for one vo and down for the other ones)

We need to identify the list of services that could be reliable . I mean a list of services which have a similar behaviour with the VO ops and the other VO.

## Algorithms

Once collected the data about A/R for the resources associated to the “high activity” VOs, we have to compute the VO monthly A/R statistics.

We have to choose one or more algorithms to do it.

Our idea is to start with the simplest one that could be expressed as follow:

(OR CEs) AND (OR DPMs) AND (OR LFCs) .... AND (OR WMS)

Using this algorithm we don’t obtain information about the perceived status from an user point of view, but just to a re-aggregation of the data on a VO basis.

In the future we will adopt more complex algorithm that taking in consideration averages and weigh in a proper way the different kind of resources.

* Everibody: Describe here the details of the algorithms and its limitations, i.e.

# References

|  |  |
| --- | --- |
| R 1 |  |
| R 2 |  |
| R 3 |  |
| R 4 |  |
| R 5 |  |

1. Numbered appendices
	1. Appendix level 2
		1. Appendix level 3
			1. Appendix level 4
				1. Appendix level 5