

Date: Tue, 26 Jun 2012 19:59:21 +0200 (CEST)
From: secretariat-egicf-emitc2-proceedings@eu-emi.eu
To: santiago.gonzalez@ific.uv.es
Cc: secretariat-egicf-emitc2-proceedings@eu-emi.eu
Subject: [PoS] PoS(EGICF12-EMITC2)028

Dear author(s)

we have now completed the refereeing process of your paper.

Your paper is accepted, although some minor changes are requested. You find below the reviewers' feedback.

NOTE : Please send us an email to acknowledge the reception of this message.

We require the implementation of these recommendations, within one month from the reception of this message. Should you need more time, please contact us at secretariat-egicf-emitc2-proceedings@eu-emi.eu

=====

Instructions:

1: If there are requests for change, to help reviewers on how these have been addressed, please send to secretariat-egicf-emitc2-proceedings@eu-emi.eu a copy of the below feedback, with the suggestions answered point by point.

2: Upload the revised version to PoS

Don't hesitate to contact for any doubt you may have.

=====

Reviewer A

Overall Evaluation:

Rate on a scale of 1(poor) to 5 (excellent)

Relevance	:	4
Presentation	:	4
Originality	:	3
Overall	:	4

Strength points

(summarize two or three reasons for publishing this paper)

- The study is a good and successful example for ATLAS new computing model on the T2D center.
- The study shows the advantage of the adopted model for the point of users.

Weak points

(summarize two or three reasons for rejecting this paper)

- N/A

Detailed comments to Author(s)

Is there any other similar case studies for the implementation of this new computing model for ATLAS?

IF YES; Please mention about the similarities or differences between the result of the studies in the conclusion of the paper.

IF NO; Please indicate the originality of your implementation again in the conclusion of the paper.

Santi (Done)

We have added more information in section 2.2, 2.3, 2.4 and a new reference. Where there are more details..

We have indicated the similarities between the result fo the Studies in the conclusions with other sites/Tier2s.

Reviewer B

Overall Evaluation:

Rate on a scale of 1(poor) to 5 (excellent)

Relevance	:	5
Presentation	:	3
Originality	:	4
Overall	:	4

Strength points

(summarize two or three reasons for publishing this paper)

- Relevant topic: the consequences of the restructured ATLAS data management architecture to the users and to a Tier-2 site.

Weak points

(summarize two or three reasons for rejecting this paper)

- The most important statements are not backed up with description of tests scenarios and measurements
- Some sentences need improvements in terms of English

Detailed comments to Author(s)

- Overall:
 - o GRID -> Grid
 DONE!!!!
 - o Atlas -> ATLAS
 DONE!!!
 - Abstract:
 - o EGI/gLite middleware -> EMI/gLite middleware OR EGI-UMD-gLite middleware
 DONE!!!!
 - o "It is also used..." - what is also used? Rephrase
 DONE!!!, We have rephrased it
 - Section 1:
 - o "High Energy Physics..." - the word institute is Messing
 DONE
 - o "important size as it is the IFIC" - what size?
 DONE

- o LHC runs proton beams

 DONE

- o Higgs particle - if exists! It's not known

 DONE

- o ...affects to the IFIC users

DONE

- o How the Spanish Tier-2 is working - so is IFIC a Tier-2? It's not stated anywhere in the paper.

It is said in section 1 at the beginning:

"IFIC participates in this experiment as one of the members of the Spanish Tier-2 together with *Universidad Autónoma de Madrid (UAM)* and *Institut de Física d'Altes Energies (IFAE)* in"

DONE

- Section 1.1:

- o There are 11, not 10 Tier-1 sites.

Actually taking into account CERN + the new one in Russia there is 12...

But Ok, Taiwán was not working very well, also there is around 70 Ter2

- o What is the meaning of the "cloud" here? Explain

DONE, introduced in section 1.1

- Section 2:

- o Typos: tranfered -> transferred

Done

- o What's "multi-cloud site"?

done, introduce the definition in section 2

- o If the T1s store data, then why does a T2 want to request data from another T2? Why it does not request data only from T1s?

The Tier1 stored Raw data, some time for the analysis people use specific data dedicated for the analysis like D3PD which are not stored at the Tier1 and can be stored at the Tier2, which are dedicated to analysis. So people from another Tier2 can required that data store only a some Tier2s.

We have try to explain that on the paper in section 2 with a new paragrah...
DONE

- Section 2.1

- o "During the first year of operation..." - when?

DONE

- o "During the first year of operatio the hierarchical model has been found insufficient." - this statement needs reference to evidences.

OK, we have rephrase, and this is referenced by the ATLAS collaboration

"The original model was a working starting point but Tier-2 activity was strongly linked to the associated Tier-1 reliability. Some Tier-1s did not have associated Tier-2s and had few tasks to process while they had the storage to host a significant fraction of task outputs. In addition, Tier-2 had the computing resources to do reprocessing but was limited due to required direct access to the Tier-1 database. As a consequence, some of the sites were not used at full capacity, especially Tier-2s."

DONE

- Section 2.2:

- o "As it has been shown..." - needs a referente

DONE

- o "During the latest weeks..." - when?

Ok, as has been during the last year as well, we has changed the paragraph

DONE

- o "During the latest weeks IFIC has fulfilled transfer performance..." - statement needs evidences. Reference to test documentations?

Done, we have added a Plot..

- Section 2.3:

- o What are the meanings of Alpha, Bravo, Charlie, Delta ranks?

Done, we have added more information

- o 90% availability for a site? It seems very low, compared to Web servers, Web services, FTP servers etc.

It is ATLAS specifications...

In any case I have added the definitions for Availability

Site availability for analysis jobs is defined based on the "analysis panda queue status":
(the amount of time queue status = online in the period) / (the whole period).

DONE

- o VOS -> Vos

Done

- o The following part is confusing. Rephrase: "Other VOs are getting more maturity ... can help for his work."

We have rephrased in the next way:

... collaborative

Grid by the physicists, that re getting more and more used to the tools and the environment to perform the work. On the other hand other VOs are getting more and more experienced in the use of the collaborative grid, also benefiting from this framework and the involved processes, which eventually improves the experience for their users.

DONE

- Section 2.4:

- o What is "EC-Cloud"?

DONE

- o Has the described changes really resulted shorter waiting times? Describe the tests and the results. (Or referente them from other papers)

we have added more information in the section a referente to a paper.

- o Figure 3: What"s "ES-Cloud"?

Solved

DONE

- o Figure 3: Hard to read, text too small.

Ok, figures are now bigger to read it better

DONE

- Section 3:

- o "IFIC is well qualified" - in what? What"s the proof?

Done, we added figure 1 and referente to that figure

- o Incorrect sentence: "It is also very important the configuration, ..."

On the other hand the management and support for these resources is important as well.

DONE

- Section 3.1:
 - o Expand the "EGI" acronym when it's first used

DONE

- o Table 1 and Table 2 show different numbers for the IFIC storage space. Harmonise and correct these. If they are

In one case we use cores table 2 and in the other the HEP-SPEC, but both are similar..

There in the number of core a 1, to say the equivalent between cores and HEP-SPEC, and the value is similar in both tables.

For the space isa round 1 PB

correct, explain what's the difference between the tables.

We have added in table 1 that is the resources for the Tier2.in table 2 are resources for Tier2 and GRID-CSIC

DONE

- Section 4:
 - o Couldn't the data transfer route of the old ATLAS model be hidden from the users with high level tools?

In principie Yes, both models has to be transparent for the user. The new one is better because the important data are copied in several Tier2s, so this is good for the anlysis, andi t is the Tier1 is not working, the Tier2s associated with them can work and users can analyse data on them.

- o Besides providing a simpler way for the users for data transfer, is the new data management model more effective than the old one? Test results?

We have provided some results on the test:

For a heavy exotic particles analysis at IFIC, in two weeks 6 users sent 35728 jobs with success to 64 sites (Tier-2s and Tier-1s). 1032 jobs were sent to the Spanish Tier-2 (2.89%) and 815 datasets were used as input files producing 1270 files as output

- Conclusion

o "This example is the best proof of the good throughput..." - it was not a proof. Where are the test cases and testresults?

Related with the text in section 4

Done

=====

Thanks for your prompt attention.

Kind regards,
Secretariat for Publication of EGI CF/EMI 2nd TC proceedings