WLCG Operational Costs
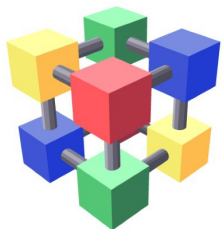
A. Forti, J. Flix, A. Sciaba',
M. Dimou, M. Alandes-Pradillo
On behalf of the WLCG Operations Coordination Team
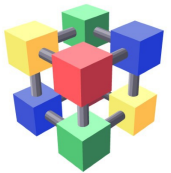
WLCG Workshop
11th April 2015

Outline

- Ops Cost and Survey

- Service Maintenance

  - Support

  - Batch System

- A "Simple Tier2"

- New Technologies (A new type of Tier2)

- Monitoring

- Communication

- Participation

Ops cost and survey

- WLCG Operations Coordination team was tasked to evaluate experiments and sites operational costs at the end of 2014
  - Some feedback from the experiments
  - Site survey
- Initial presentation about FTEs and communication given by J. Flix and March GDB.
- This presentation concentrates on the suggestions from sites in the site survey and some random thoughts
  - Most suggestions even those more practical open further questions

Service maintenance
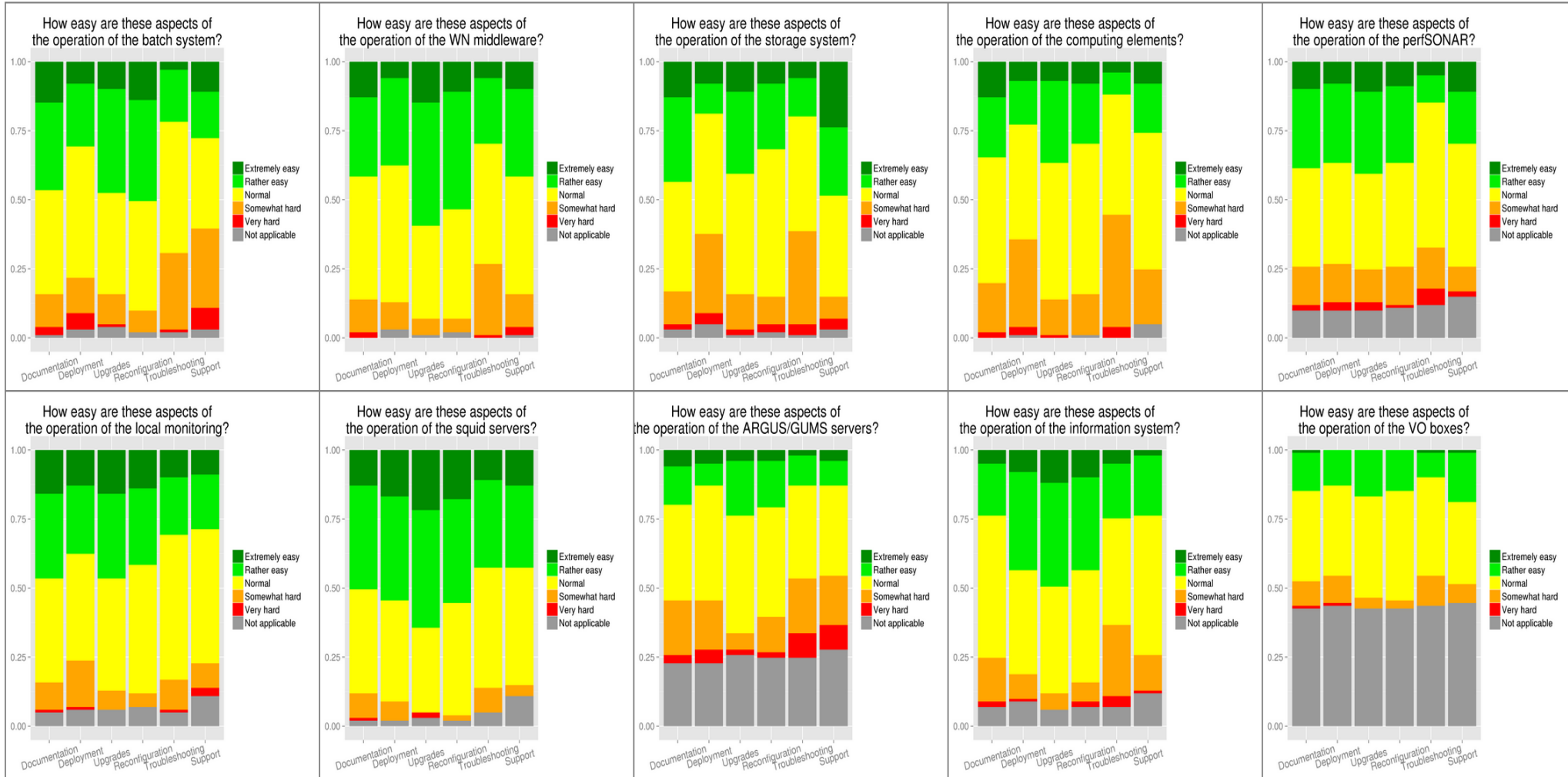
Support

- Peaks of negativity for deployment and troubleshooting services in particular SEs and CEs

  - Poor error reporting, poor documentation, lack of log parsing tools.

- YAIM/EGI/EMI future is unclear.

  - YAIM created problems at few more sophisticated sites but many others are still using it underneath puppet when available.

- The number of services not really required for operations should be reduced.

  - WLCG services should be simpler and fewer

  - Particular grudges with glexec and perfsonar

- The lack of proper support for ARGUS and other tools (Torque/Maui) is a concern

- Time spent managing hardware, producing user documentation should also be considered a WLCG activity.

Batch system: torque/maui

- Major worry is the lack of support.
  - Unofficially supported by Nikhef on a best effort basis and if it suits also Nikhef needs
    - Provided multicore procedure with easy to install script
    - Looking into cgroups plugins
  - Sites cannot open a bug or request a feature and expect to be answered as they would do with official support
  - Community support also practically not existent anymore
- Several sites asked for direction or recommendation from experiments and WLCG.
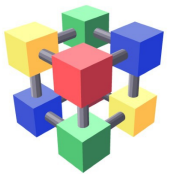 - Cannot tell sites what to do but can express a preference.

A "simple Tier2"

- There is this new concept of a "simple Tier2" but unless we revolutionise our sites there is very little that can be removed. Some suggestions

  - APEL box has been mentioned by several sites as a burden but it is not strictly necessary

    - ARC-CEs (and OSG) publish directly into APEL not clear why all the other CREAM sites couldn't do the same

  - Push new sites towards ARC-CE which is simpler and more robust

    - Push new sites towards Htcondor as community is building up

    - Alternatively SLURM or if they can afford it UGE or latest LSF

  - Re-evaluate the need of an heavy weight BDII, mostly used for service discovery and getting few unreliable numbers in Rebus.

    - Are all the lines it publishes really needed?

    - Service in itself is light weight **if something fills the values for you. And YAIM is fading.**

  - Keep up the work to reduce the number of storage protocols
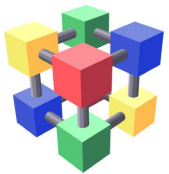
New Technologies
A new type of site

- Cloud

  - WLCG should provide tested and optimised disk images compatible with any cloud infrastructure. Ideally, running a WLCG site should mean running few (1-3) virtual machines with WLCG-provided images configured by a small and well documented configuration file

    - Some of these ideas more expanded in the Resource Provisioning presentation

- Containers

  - WLCG should move to a model whereby site admins no longer are required to install local services beyond core cluster functionality;

  - WLCG should invest in containerisation of middleware services to reduce the workload on sites. Keep the barrier of entry for potential opportunistic resource providers at an extreme minimum: provide a grid model but leveraging container technologies such as Docker or Rocket.

    - Some of these ideas have started to be discussed also in the HSF context although not concentrating on grid services.

New Technologies Storage

- Have HTTP-based storage federations

  - Couples with other requests to "stick to industry standards" when evaluating new technologies.

- A big number of sites are looking into CEPH as a storage technology. Among many pros, it enables SSD caching out of the box. This could be a game changer for more efficient WAN transfers. So supporting CEPH as an SE technology could be very beneficial.
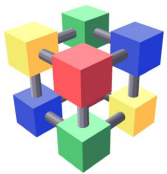
Monitoring

- Still conflicting messages between the necessity to monitor the whole experiment chain and having the capability to isolate the "site functionality" problems properly for availability calculations.

  - Sites are upset by false positives (test fails but jobs run), but they are also upset by false negatives (tests fails, but not site fault and availability drops)

    - Repainting is a chore both for the site and for the experiment
    - Should publicize SAM integration in local nagios more
      - But old ops tests sent nagios alarms without sites having to integrate
    - Local monitoring essential to catch problems before jobs do.
      - Wild variety of sites setup.
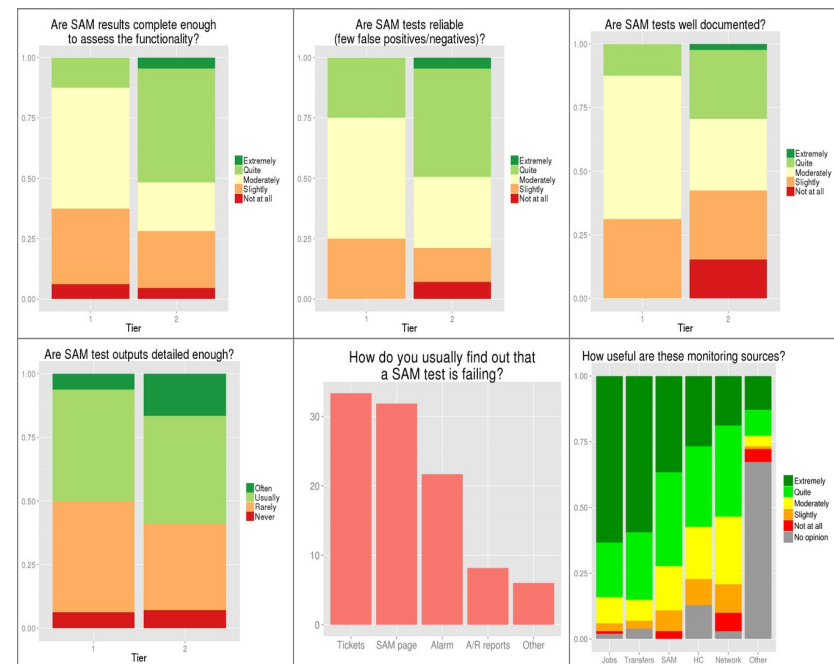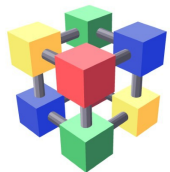
Monitoring: SAM tests

- Troubleshooting difficult due to poor error reporting and lack of documentation links in the test. These are old favourites

    - Time wasted always to google for references to the errors (which are slowly disappearing because pages are progressively less searchable)

    - There is no description of what the tests do and it is not possible to replicate manually

        - Experiments in charge of the metrics so what is there is experiment dependent

    - It is not possible to rerun the tests once the problem has been fixed

- Several requests for a site unified page showing

    - All the experiments SAM tests for a site

        - A top level view like this existed and has been eliminated during consolidation for lack of usage. Either a communication problem or new admins just repeat old requests.

    - All the experiments tests on top of SAM for a site

        - Could be done using SSB but requires some manpower.

The operational cost of protecting pages

- One of the recurrent requests talking about troubleshooting, documentation and monitoring is that quite a lot of information is nowadays protected and not searchable.

  - The experiments twikies and JIRA tickets are protected too and let's be honest the twiki/JIRA search functionality ain't no google.... and "I found this savannah ticket that says..." ain't no more.

  - On the monitoring side it's practically impossible to access any meaningful monitoring unless one belongs to the experiment in question

    - ATLAS – the only experiment keeping the monitoring open is now also closing one of the most important monitoring services.

- We should be aware that all this has a dramatic operational cost

  - If it was bad before with distributed documentation now it is worst because google has lost its effectiveness and site administrators are forced to open CERN accounts to access information.
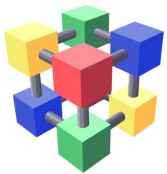
WOC ↔ sites
Formalisation

- Formalisation of the procedure to make requests to sites

    - Political: important service requests should come from WLCG as formal requests, endorsed by the WLCB MB. WLCG Ops Coordination role should be further enhanced to make the experiments converge in their demands.

        - This is important to discuss and take a decision because we know 58% of sites support at least 2 LHC VOs. And would definitely go towards a simplification of communication

    - Practical: all reasonable requests but may require some work

        - Make clearer the difference between requests and requirements

        - Have clearer deadlines

        - Use more formal tracking of requests

            - We know GGUS is a favourite but also WLCG broadcasting

        - Add a WOC GGUS SU for communication from sites

            - Probably excessive but the egroup should be publicized more
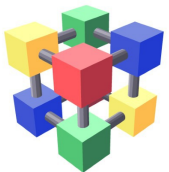
WOC ↔ sites
Planning

- WOC directions are not clear. It's all hidden in the minutes but not clearly expressed.

    - Some keywords of this set of suggestions: planning, timeline, short and medium term, checklist

        - Or as a site put it "the outcome of the WOC meeting should be planning not minutes"

- This was sort of discussed at the last GDB and we agreed (at least I think it is a good idea) that adding a "Sites action" section to the minutes (with the deadlines?) would already cover most of these problems.

    - Further broadcast of that minutes section should act as reminder for all sites to read the minutes.
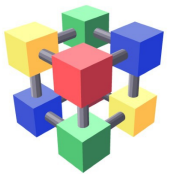
WOC ↔ sites Consolidation

- Consolidate the planning function to improve the consistency of messages to sites

  - Avoid having many ways of managing, tracking and tracing operational developments

- Consolidate information in a single entry point to updated information about the middleware, known issues, recipes, and the experiment software, **public and indexed by search engines** (x4)

  - Avoid having important information hidden in tickets – if a problem/fix is identified, everyone should be notified

  - Have fewer sources of information and more things coming from official channels
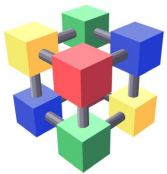
Sites ↔ Sites

- Suggestions here are abundant here are three that could be implemented
  - Creation of an open wlcg-discussion list for sites to discuss on new technologies, new services, deployments, validations
    - Sounds suspiciously like lcg-rollout for some things but it may help create a sense of community, new technologies output may be fed back to HSF (?)
    - Mailing lists and their archives must be searchable.
  - Re-introduce a yearly WLCG jamboree for sites to meet and discuss WLCG related matters
    - They used to be very well attended, not sure why they were dropped.
    - Hepix and GDB well liked as fora for sites
  - Open and searchable wikis, provided they are carefully kept up to date for official documentation

WOC, TF, WG Participation

- How to improve WOC participation
  - Start a bit later to allow US sites to attend i.e. 4 pm, limit the duration to one hour and focus the meeting on a subset of issues making the purpose of the meeting clearer
    - All sensible suggestions
    - More active involvement of T2 representatives is another
- Participation to the TF (lack of manpower aside)
  - Timezone difference mostly for Asian sites
  - Effort from site admins should rather go into maintaining a central documentation and training repository
    - How many other sites would subscribe this?
  - The work of the TFs and WGs should be better advertised in a public forum
    - WOC reports and TF/WG presentations at the GDB not enough?
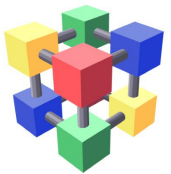
Conclusion

- To improve operations suggestions pointed at reinforcing the community ties and enhancing the role of WOC by formalising communication channels and planning.

  - Some of the feedback exposed are practical and sensible suggestion that can be easily implemented.

- WLCG monitoring could be further improved

  - Expose more experiment monitoring to sites

- Some suggestions on how to simplify a grid Tier2 infrastructure

  - Eliminating some of the services and replacing others with other solutions, moving to using standard tools rather than HEP grown.

- Further simplification especially in view of loss of staff requires diversification of technology.

  - Really simple sites may not run every application.

  - Some work is already ongoing
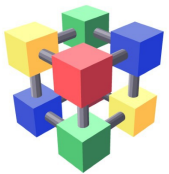
Acknowledgements

- Many thanks to all the 100 sites that replied the survey and offered their advice and suggestions.

Questions Discussion

WLCG Operations Coordination: el WOC