

# CRIC Validation Workflow

Boris Vasilev

Julia Andreeva

Panos Papadopoulos

Dimitrios Christidis

Efstathia Chioteli

# Current state of the accounting workflow (1)

Accounting data is auto-generated on WLCG sites.

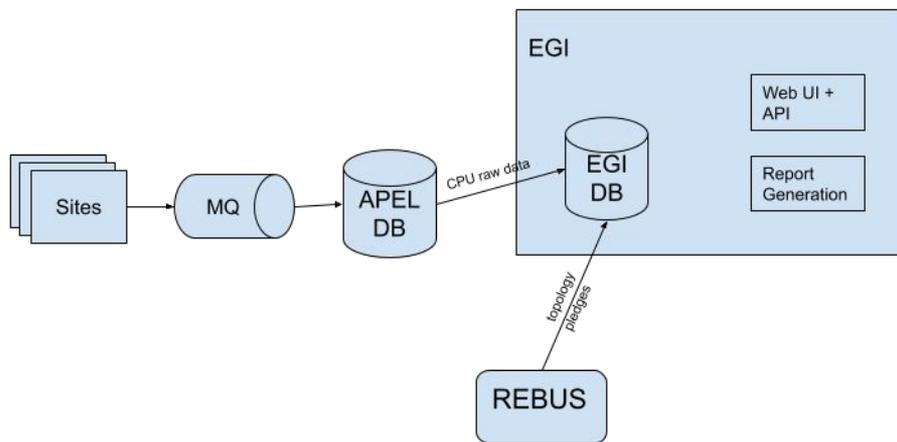
This data is sometimes faulty.

However, faulty or not, the data is still pushed from sites to APEL DB and further to the EGI Accounting portal (CPU accounting data) and collected by WLCG Storage Space Accounting (WSSA) (storage accounting data).

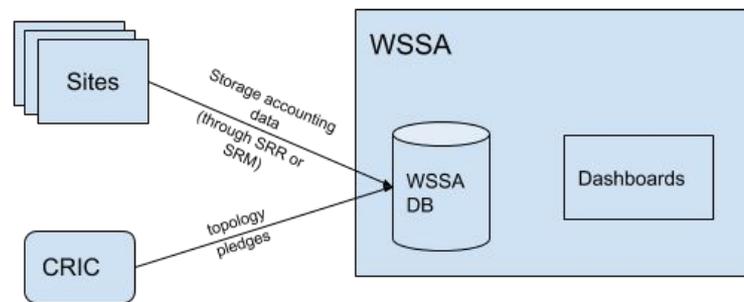
Data collected in EGI accounting DB and WSSA DB is used for generation of the WLCG accounting reports

# Current state of the accounting workflow (2)

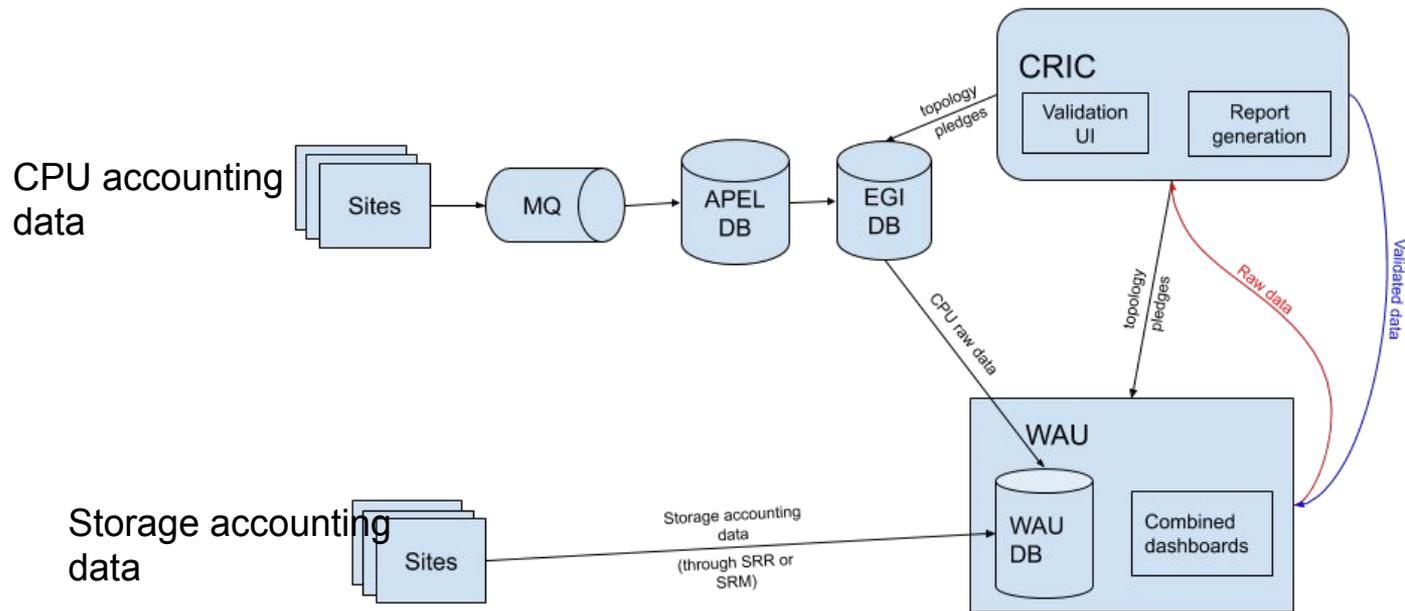
CPU accounting



Storage accounting



# New accounting workflow with validation



# Improvements of the accounting workflow

Introducing a validation mechanism to enable correction of the auto-generated accounting data coming from sites.

Taking back accounting report generation from EGI to use the validated data.

Produce a combined Grafana dashboard with Influx database backend combining data for both CPU and Storage.

# CRIC validation mechanism

Home Core Core API WLCG WLCG API Docs Admin Logs Help echotel Exit

## Accounting Data for 'AGLT2' Resource Centre Site

Form Submission Error!

### Validate CPU data for ATLAS

Generated Wallclock Time WSSA: 0 hours	Please provide your input (in hours) if you disagree with the autogenerated value: 22555 <small>This value has not been saved, please provide a valid one!</small>
Generated Wallclock Work WSSA: 0 HS06hours	Please provide your input (in HS06hours) if you disagree with the autogenerated value: 202

### Validate Disk data for ATLAS

Generated Total WSSA: 5.7 PB	Please provide your input (in PB) if you disagree with the autogenerated value:  <small>This value has not been saved, please provide a valid one!</small>
Generated Used WSSA: 5.38 PB	Please provide your input (in TB) if you disagree with the autogenerated value: 12.3

Save

CRIC WLCG Web Portal server: stathoula-dell | release: core-unknown wlcg-unknown

# WAU - WLCG Accounting Utility

WAU = WSSA + CPU accounting

WAU is a set of 3 combined Grafana dashboards

- Raw dashboard (Implemented)- contains raw auto-generated accounting data
- Validated dashboard (To do) - contains data with updates from validation
- Comparison dashboard (To do) - contains comparison of raw and validated.

Used to detect problems with auto-generated data and follow up on sites producing faulty data

# WAU - WLCG Accounting Utility

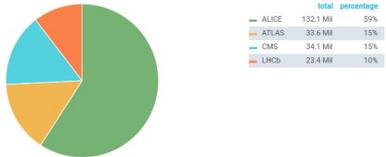
\_user bvasilev cpu\_accounting\_test1 - Jun 2, 2015 22:22:39 to May 31, 2023 22:22:39

## Summary

federation	WCT (hours)	WCW (HEPSPEC06 hours)	Disk Used	Disk Free	Disk Total	Tape Used	Tape Free	Tape Total
CH-CHIPP-CSCS	27 Mil	282 Mil	2.46 PB	902.42 TB	3.37 PB	-	-	-
CH-CERN	321 Mil	3 Bil	48.94 PB	10.10 PB	59.04 PB	217.99 PB	63.41 PB	280.99 PB
ND-Tier3s	-	-	7.99 TB	19.50 TB	27.49 TB	-	-	-

## Pie Charts

Wall-Clock Time (in hours)



Wall-Clock Work (HEPSPEC06 hours)



\_user bvasilev cpu\_accounting\_test1 - Dec 1, 2018 22:22:39 to Dec 1, 2019 22:22:39



# CRIC report generation

Similar to EGI report generation but instead of raw data will use validated data.

Fixing some bugs of the EGI accounting portal for the reports.

Will enable T2 federation reports similar to T1 reports

Improved data quality

# Validation workflow

1. Auto generated raw data is received from sites.
2. In the beginning of every month CRIC sends a request to WLCG site admins to validate accounting numbers for the previous month. Agree or provide correct number.
3. After two weeks, the validation phase closes. Validated data is sent to WAU.
4. At the end of the month reports are generated from CRIC using the validated data.
5. Follow up is required for the sites where measurements generated by the sites are much different compared to ones provided by site admins (threshold to be decided).

# Current status

In the past few months we have been working on this new workflow which will improve data quality.

Implemented:

- CRIC validation mechanism
- Raw data view in WAU (Storage + CPU)
- CRIC report generation (Tier 0+1)

# Next steps

Notification to site admins to validate the accounting data.

Improve WAU database performance by 'caching' fine-grain data (hourly) into coarse-grain stores (daily, monthly).

Report generation for individual Tier-2 and Tier-2 federations

In October we plan to send first notification requests to validate accounting data for September and generate new accounting reports (not official yet) by CRIC. These accounting reports will be sent along with the official ones.

Full transition is planned by the end of the year

Thank you for your attention!