

FAIR Commissioning & Control Working Group

Notes from the meeting held on June 1st, 2016

e-mail distribution: [FAIR-C2WG-ALL at GSI.de](mailto:FAIR-C2WG-ALL@GSI.de), [participants list](#)

Agenda:

- FAIR Archiving Specification – Presentation in view of Approval, V. Rapp
- AOB

1. Discussion on the specification of the FAIR Archiving System

Discussion:

Data validation

J. Stadlmann: How is valid data defined?

V. Rapp: Invalid data means it's not sent in the expected frequency or out of bounds

J. Stadlmann: This might lead to filtering out what you want to see. Don't filter too much!

V. Rapp: It's about the definition of rules

R. Steinhagen: Someone's noise is someone else's signal. Prevent overloading of the system

R. Bär: Data is already preprocessed e.g. out of bounds is already checked

J. Stadlmann: What if a device has a weird value for some reason.

R. Bär: What is sent to the archive system should be stored in the archive system basically – unless it's really completely corrupted

J. Stadlmann: Access to that system must be protected

Backpropagation of archived data

R. Steinhagen: What happens to less-than-perfect calibrated data (e.g. wrong gain factor)? At LHC this was particularly important for accurate intensity/luminosity figures. These historic records had an important impact on the physics analysis and precision of the physics cross-section analysis. Should we consider back-propagating known intensity calibration changes similarly for FAIR?

F. Herfurth: Archive system shouldn't be smart, just store, don't validate

J. Stadlmann: Back-propagation would lead to data loss

R. Bär: Could lead to storing raw data, using already calibrated data

R. Steinhagen: Apply post correction, used for most important data at CERN

F. Herfurth: Should post correction be part of archive system or another system?

R. Steinhagen: Every tool has to know of the post-processing system when it's an extra system

R. Bär/R. Steinhagen: Needs to be decided (not part of spec)

Broker

R. Steinhagen: Time scale for cache?

V. Rapp: This is configurable (hours → days → weeks, if necessary, only related to RAM/disk space)

UNICOS / Data reduction

J. Fitzek: Data from UNICOS only going to long-term storage?

V. Rapp: It's not ready at the moment, not clear if it can be part of the short-term-storage

U. Weinrich: It must be part of short-term-storage

R. Bär: UNICOS has its own short-term-storage

U. Weinrich: This means it is reduced?

R. Bär: Not necessarily, it depends on when you access the data

R. Steinhagen: Specification needs to be more specific on data reduction. Not all features necessarily have to be implemented initially, but the interfaces for future upgrades should be already provided.

U. Weinrich: The scales for data reduction must be defined. May not reduce data newer than one week

R. Steinhagen: The present planning is to keep the raw data at least for 3-4 weeks (tbd. if more is needed).

R. Bär: If someone has made a certain observation, this data can be prioritized and won't be reduced. In the end it's a matter of disk space.

U. Weinrich: Would be great to see the average pressure first and then dig down into the details.

Technical concept

R. Steinhagen: Are both databases filled at the same time?

V. Rapp: Transferring data from short-term-storage is faster

Spec process

R. Bear: Spec must be finished. What are the user requirements? Did we forget something?

A. Reiter: What is in the spec? meta-data / predefined

R. Steinhagen: Specific use-cases can also be added later, but knowing some now helps during implementation

U. Weinrich: Who will check the spec? Have users being asked?

R. Steinhagen: Everyone has been invited to this meeting, and the specs circulated for the initial engineering check to the FC²WG list. Since this includes all MPLs/MCs and related further distribution lists, this should cover most (if not all) important stake holders.

Handle device changes

R. Steinhagen: Specification must be more specific on data extraction and reduction. How to handle change of devices (New devices, device retired, BPM model changes, BLMs at different location)?

V. Rapp: It's mentioned in specification, changed are stored in FESA database, data model of devices will be stored in meta data

R. Steinhagen: Two paradigms by data objects or by variables?

V. Rapp: FESA property as 1:1 (same as CERN is planning to go), same way of tracking changes, much faster

R. Steinhagen: When using the by objects approach for long terms trending we need to unwrap this structure, expected as part of extraction interface

R. Steinhagen: Interface must be self-explanatory, real-world variable names

J. Fitzek: Already a requirement for FESA devices to have such names

V. Rapp: Changing data storage concept might be expensive

Spec process / finished specification handling

A. Reiter: How to ensure the right details will be added later (next steps)

R. Steinhagen: Another round of circulating the spec (for engineering check), then meet eet again in one month. Deadline is June 30th, circulate current draft with changed applied (based on this this meeting)

R. Bär: Next step is a conceptual design document, will contain more detailed information

R. Steinhagen: Two steps - functional requirement (use cases, scope) and detailed requirement (technical)

Reported by Ch. Hillbricht