



iDDS status updates

Wen Guan, Tadashi Maeno

Brian Bockelman, Torre Wenaus

Fernando Barreiro, Fahui Lin, Rui Zhang, Misha Borodin, Paul Nilsson

Main structure updates

- **ORM layer (database) updates**
 - CLOB item consumed a lot of database space.
 - Update the iDDS Work/Workflow structure to automatically only store the changed items to db. It fixed the db space consumption issue.
 - CLOB item in contents table slows down big bulk insertion.
 - Replace the CLOB items with varchar2, improved the bulk insertion efficiency.
 - Optimized index on contents and transforms table.
 - Improved the exclusive locks for multiple agents/threads
- **Work/Workflow plugin updates**
 - Update the plugins to adapt the Work/Workflow structure updates.
- **Request management with automatic expiration:**
 - Lifetime: time from creation time
 - Pending time: time from no new jobs/outputs
- **Agent updates:**
 - Reduced the processing steps: Merge several steps together in one step

Use cases updates

- **DOMA LSST exercise**

- Scaling tests with 30K(finished)
- Scaling tests with 230K(subfinished)
 - LSST software kit returns 0 while some jobs failed. As a result, the failed jobs were marked as successful. Need a fix on the software kit.
- Scaling tests simulation on the dev instance
 - Without submission to Google Cloud.
 - Optimize iDDS structure and db structure during the simulation to speed it up.
 - For 30K jobs(24 tasks), iDDS can finish handling request, transformation and processing creation/updating within seconds(10~30 seconds).
 - For 230K jobs(24 tasks), current iDDS still needs about 16 minutes(improved from 38 minutes).
 - Similar results when increasing the VMs from 1 to 3.
 - Maybe it needs a big buffer until the session commission.
 - Thinking how to improve it without breaking session consistency.
 - Simulating job releasing: No big delay found.

- **HPO**

- FastCaloGAN is using it.
- New use case ToyMC is adapting to use the HPO framework.

- **Others**

- Update the plugins to adapt the Work/Workflow updates.

User Interface updates

- **Clients:**
 - **New enriched clients:**
 - Request status displaying with/without details
 - Request management with abort/retry
 - Others
 - Improved the client return values and logging.
 - Integrated iDDS client with panda IAM.
- **Monitoring**
 - ATLAS iDDS monitor
 - DOMA iDDS wfprogress monitor(new)
 - Grafana and Elasticsearch
 - Lite independent monitor(dev)

Summary: iDDS Current Status

❖ Main architecture

- iDDS database, core, REST API
- Plugins
- Agents
- Watchdogs

❖ Documents & monitors

- Home page: <https://idders.cern.ch>
- Codes: <https://github.com/HSF/iDDS>
- Documents: <https://idders.readthedocs.io>
- ATLAS monitor: <https://bigpanda.cern.ch/idders/>
- Different monitors are being enriched.

❖ Instances in production

- ATLAS, DOMA

❖ Use cases

- Data carousel
- HPO(toymc)
- Job level DAG(LSST exercise)
- Task level DAG(ActiveLearning)

❖ Future developments

- Main Structure improvements
- Monitor improvements
- New use cases
 - Dynamic transformation and placement on demand, for example Derivation on Demand
 - Fine-grained data transformation and delivery, such as Event Streaming Service