



RCS-ICT Technical Committee

The RCS-ICT Technical Committee is an inter-departmental and inter-organizational governance body. Establishes and maintains a roadmap of ICT services and projects jointly engaging the communities in the RCS sector.

Notes of the 13th RCS-ICT Technical Committee held on Friday 26th April 2024

Date: 26th April 2024

Main topic(s): Containers registry, CDS future, Update on the Collaborative Tools PSO, Action List Status

Agenda (<https://indico.cern.ch/event/1318715/>)

- Containers registry
- CDS future
- Update on the Collaborative Tools PSO
- Action List Status

Attendance

Present: Alexander Yohei Huss, Andreas Morsch, Andreas Peters, Anika Churilova, Armin Nairz, Ben Couturier, David South, Dirk Duellman, Enrico Gamberini, Eric Grancher, German Cancio, Gianluca Cerminara, Jan Van Eldik, Latchezar Betev, Liz Sexton-Kennedy, Lorenzo Moneta, Marc Dobson, Marco Cattaneo, Micha Moskovic, Nicola Tarocco, Pere Mato, Ricardo Rocha, Shasham Arora, Simone Campana, Stefano Piano, Wainer Vandelli, Xavier Espinal, Zach Marshall, Zacharias Zacharadimos, Zhechka Toteva

Excused: Stefan Roiser

Adoption of Meeting Minutes

Zhechka thanked the contributors to the minutes' review. **Minutes were approved.**



Containers registry (Ricardo Rocha)

Discussion

Harbor is used in many areas, including ATS and RCS, IT, ATLAS, and etc, such as the Docker images for Invenio. Also, ATLAS uses it with external image providers.

As part of the discussion on Harbor's special features, the following points were raised:

- Harbor scans an image and finds all the packages, creates a document, creates a database with the dependencies and can track vulnerabilities later on.
- On the first pull, the image is pre-fetched, but if it's not used for a while, it's garbage collected, and the next time it's needed, it's fetched again.
- There can be automatic replication per project, per a couple of projects, inside a project it can be per a couple of repos or per tags, regular expressions in the replication rules' criteria are quite popular.
- External workflows can easily be hooked into it.

The following conclusions were reached during the discussion of moving images from GitLab to Harbor:

- Harbor can cope with content of more than 20 TB of over 10K images without issues at all.
 - Nevertheless, a scale test should be organised before the HL LHC.
- The native integrations offered by GitLab make it convenient to use.
 - Meanwhile, GitLab features that have never been activated will be difficult to add now.
- When pulling from far distances, Harbor measures slower than Docker.
 - One of the reasons is that Docker hub uses CDN behind.
 - A caching layer can be deployed, similarly to squid and CVMFS.
 - Harbor can be used directly with CVMFS (the backend repo can be CVMFS)

We need a pilot and test how it fits in the WLCG architecture and which use case can fit well to use the Harbour.

Action: Form a working group across the community to tackle the following tasks:

- Run a pilot of container registry in distributed computing (eg. Harbour) caches outside CERN, eg. US/BNL, including profiting from CVMFS.
- Test the scalability of containers' usage in WLCG for HL-LHC: a joint request from the 4 LHC experiments.

CDS future (Nicola Tarocco)

Discussion

In the context of putting in place the right helpers for a global search between the old and new CDS, it was also discussed creating a global search engine at CERN. The IT-PW group is developing the CERN Search, but the CDS search will remain.



In the scope of the Scientific Information Landscape Project, where the CDS team is actively participating in the working group, a decision has been made as to what the GLANCE project and CDS project's responsibilities are.

For a while, the old and the new CDS will run in parallel. The migration to the new CDS will not be slowed down by gathering requirements for new features. The old CDS was developed long before the GLANCE project. The GLANCE project is now more feature-complete. The features that need to be implemented in CDS can be implemented quickly. It is possible to shorten the process with the help of the members of the working group.

Action: Understand better the CDS place in the Scientific Information Landscape Project and gather the information from the communities and present it back at the Technical Committee before the end of the year.

Update on the Collaborative Tools PSO (Eric Grancher)

Presentation

The short grace period for the CERN accounts termination upon departure of the organisation is one of the issues with collaborative tools.

Eligibility discussions are primarily based on financial considerations. The eligibility matrix has been presented to the department heads, while checking if everything makes sense. It might be a good idea to reconsider some restrictions if there is no financial benefit or no legal risk.

As part of the Google workspace discussion, it was strongly recommended that professional and personal accounts be separated.

After the latest CERN council meeting, it was decided that accounts of users from Russian institutes who participate in CERN communities will remain "external" after November 2024 to allow them to complete their analysis. GRID certification authority does not support externals. Passports are not checked for external members. Passports can be considered already checked for users who transition from user and external status. Additionally, CERN can keep a list of ex-users. Finally, the certificates from the Russian certification authorities are trusted, thereby solving the authentication issue. Nevertheless, the entire certification chain should be tested.

Each community should identify users who can assist with testing the transition from a user account to an external account. The transition tests should be conducted first in LHC experiments. The non-LHC experiments will join later. Ideally, the tests should be run before the summer.



Action: Open the PSO document on Collaborative tools to the RCS-ICT TC for comments.

Action: Contact the RCS-ICT members to get representatives and/or to provide feedback about the access to computing services for people with Non-CERN accounts. Gather real needs by users with non-CERN full-accounts from the RCS communities and match it to the Eligibility Criteria.

Action all RCS communities and Simone: Identify users to assist the testing of the transition from user to external accounts.

Critical Services (Xavier Espinal)

Xavi announced that at May's RCS-ICT Technical Committee a presentation will be given on critical services provided by the IT-CD group. It will be the first of three consecutive sessions covering the scope of IT Critical Services.

Action List Review (Zhechka Toteva)

Zhechka presented the current state of the action list. The IT department is studying the options to introduce a workflow in JIRA to follow up on the actions triggered at the Technical Committees.

AOB

The IT engagement team would like to hear from the communities regarding AI Coding Assistants, in particular their current usage in RCS and the needs of RCS.

- The CERN CMS team uses ChatGPT and Copilot.
- SFT is using AI coding assistants only for initial scripting. The production code they write themselves.
- A question was raised about Copyright and licence issues when using proprietary code in the CERN code published in these tools.
- The implications of deliberately teaching these tools with CERN code should be considered.

Action: Collect the RCS communities feedback about AI Coding Assistants (in 2-3 weeks) and share it with Ricardo and Vincent. Understand and discuss the legal implications of using AI Coding Assistants.

Action List

- **2024-13-A1 - Action on Simone and Ricardo:** Form a working group across the community to run a pilot of container registry in distributed computing (eg. Harbour) caches outside CERN, eg. US/BNL, including profiting from CVMFS.
- **2024-13-A2 - Action on Simone and Ricardo:** (the same container registry working group) Test the scalability of containers' usage in WLCG for HL-LHC: a joint request from the 4 LHC experiments.



- **2024-13-A3 - Action on Nicola**: Understand better the CDS place in the Scientific Information Landscape Project and gather the information from the communities and present it back at the Technical Committee before the end of the year.
- **2024-13-A4 - Action on Eric**: Open the PSO document on Collaborative tools to the RCS-ICT TC for comments.
- **2024-13-A5 - Action on Eric**: Contact the RCS-ICT members to get representatives and/or to provide feedback about the access to computing services for people with Non-CERN accounts. Gather real needs by users with non-CERN full-accounts from the RCS communities and match it to the Eligibility Criteria.
- **2024-13-A6 - Action all RCS communities and Simone**: Identify users from Russian institutes to assist the testing of the transition from user to external accounts (in 2-3 weeks).
- **2024-13-A7 - Action on Zhechka**: Collect the RCS communities feedback about AI Coding Assistants (in 2-3 weeks) and share it with Ricardo and Vincent. Understand and discuss the legal implications of using AI Coding Assistants.

Comments/Amendments

- ...