OpenCGA v1.2.0 Released

We are pleased to announce new version 1.2.0. This is probably one of the biggest and most significant OpenCGA release in terms of new features and improvements, during this release a total of 23,528 additions and 16,107 deletions were applied. To highlight some of the most significant improvements: now you can define internal *releases*, more secure session tokens with JWT, users can now belong to more than one group, huge performance improvement in ACL queries, new variant *basic* merge mode, better Solr integration, many new acceptance tests, ... Here you can find an overview of the most notable improvements.

General

- **QA and testing**: We have now more than **400 acceptance tests** with Fitnesse! This release has added about 50 new tests. The test coverage is now about 60%.
- **Command line**: Some bugs as *count* functionality are fixed. Some cleanups, new parameters added and an improved help are other improvements.
- **REST web services**: Authentication tokens are not passed in the URL any more, they are added to the HTTP Header. Also, some new web services and parameter have been added and Swagger has been improved.

Catalog

- **Allow to create Project *releases***: A new *release* field has been added to data models. This allows users to define and create internal *releases* automatically for any project without having to replicate the data. Therefore, Catalog and Variants data are now be associated with a particular *release*, you can query Catalog data in any previous release.
- **JSON Web Tokens (JWT)**: Session tokens have been migrated to **JWT** (https://jwt.io/) standard. JWT allows to implement federated systems by authenticating users across different applications. JWT session tokens are not stored in Catalog any more, this has reduced the number queries needed during **login**. Tokens are now sent in the HTTP header instead of old *sid* query parameter, note that *sid* query parameter will still work during next releases.
- **Rewrite ACLs database implementation**: ACL permission storage in the database have been redesigned and re-implemented to improve performance and reduce the number of queries needed, this also improves robustness and reduce source code. For instance thanks to this permission checks are now part of the query what reduce the number of queries needed. This change is completely transparent for any developer, no API or data model has been changed.
- **Multi-group support**: Users can now belong to more than one group. Permissions defined in different groups of one user are resolved as the union of them.
- **Improve synchronisation from LDAP**: Users will be registered and imported automatically the first time they *login*. Also, the groups they belong to will be synced in each *login*, this will provide and automatic validation and synchronisation with LDAP groups.
- **Private Variable Sets and Annotations**: Any Variable Set can now be defined as *private* (called *confidential*). In such a case, only users with the a new permission will be able to see or edit *confidential* annotations from those variable sets..
- **Propagate permission between Samples and Individuals**: A new *propagate* parameter has been added to both *Sample* and *Individual* ACL web service to also propagate permissions to the other related entries.
- **Improve data model consistency**: Some data models such as *File* contained an array of *samplelds* instead of an array of *Sample* objects, this is inconsistent with other data models. This type of change has been applied to *File, Job* and *Cohort* data models.
- **New *members* group**: A new default special group *members* has been added to the Study. This group will keep track of all the users from any *group* or with any permission assigned in the study. This members group allows to keep track of all users in one study making easier to remove users. Also, permissions can be set to *members* groups.
- **New REST web services**: New search web service in project has been added.

As a result of some of these changes **Catalog performance** has been improved significantly and the number of queries executed have dropped to **50%**.

Variant Storage

- **New Variant *merging* mode**: We have developed a new merge (or aggregate) algorithm when loading variants, this is called *basic* mode, you can refer to the old one as *advanced* mode. The *basic* mode is much faster and more appropriate for clinical projects while the *advanced* mode is designed for research or population studies.
- **Apache Solr integration**: In this release Solr has been integrated with both backends (MongoDB and HBase). As a result of this integration the performance of many queries, specially the very complex ones, have been significantly improved. This is a transparent change in how the query engine executes the queries.
- **Implement remove of File and Study**: It is possible now to remove (from the database and index) files. You can also remove a whole Study.

- Many small bug fixes and performance improvements

Breaking changes

In order to implement some of the improvements and new features in **Catalog** such as *release* we had to do some changes in the database. We have implemented several migration scripts that can be run to upgrade to new schema. You can find the scripts at https://gist.github.com/pfurio /2ca0cb2da46eac9e309101066f8758f5.

Another migration script has been developed to fix a small bug in **Variant Storage** database, you can find it at https://gist.github.com/j-coll/3dec01abc70644943d33de78105c633e.

Sorry for any inconvenience caused. We do not expect many more changes in coming releases, as always we will try to minimise them.

Issues and Release Notes

You can find more detailed information about all issues at **https://github.com/opencb/opencga/issues? q=is%3Aissue+milestone%3Av1.2.0+is%3Aclosed**

Release notes and links to the issues can be found at: **http://docs.opencb.org/display/opencga /Release+Notes**