

Configuration

Overview

A full OpenCGA installation requires three configuration files, all of them are written in **YAML** format:

1. **configuration.yml**: general configuration for different OpenCGA components such as *Catalog* or *Servers*
2. **storage-configuration.yml**: storage-related configuration settings to index and query genomic data
3. **client-configuration.yml**: client configuration required for the command line (*opencga.sh*) and RESTful client libraries (*Java*, *Python* and *JavaScript*)

The first two files (*configuration.yml* and *storage-configuration.yml*) are needed for running the server, while the third one (*client-configuration.yml*) is required only by CLI and client libraries to interact with the REST or gRPC services. You can find all the configuration files in the **conf** folder of the OpenCGA installation.

Configuring at run time

After installing OpenCGA you will have to edit the different configuration files. You can find more information about how to do this in the pages in the section:

1. [configuration.yml](#)
2. [storage-configuration.yml](#)
3. [client-configuration.yml](#)

Configuring at build time with Maven

In order to make easier the deployment we have defined several [Maven Properties](#) to automatically set the values when building OpenCGA from source code. These are defined in a default [Maven Profile](#) at main OpenCGA **pom.xml**. Note that you can define several maven profiles at *~/m2/settings.xml*, each of these profiles can set different values for different environments, e.g. test or production cluster can have different hosts or passwords. You can easily activate one profile when building using maven **-P** option.:

```
mvn clean install -P PROFILE_NAME
```

By doing this maven will automatically substitute the maven property value in the corresponding configuration file. For example, in the client configuration file we can find the following section:

```
## REST client configuration options
rest:
  host: "${OPENCGA.CLIENT.REST.HOST}"
  batchQuerySize: 200
  timeout: ${OPENCGA.CLIENT.REST.TIMEOUT}
  defaultLimit: 2000
```

Both `OPENCGA.CLIENT.REST.HOST` and `OPENCGA.CLIENT.REST.TIMEOUT` will be replaced for the corresponding value in the maven properties in your profile:

```
<OPENCGA.CLIENT.REST.HOST>localhost:8080/opencga</OPENCGA.CLIENT.REST.HOST>
<OPENCGA.CLIENT.REST.TIMEOUT>30000</OPENCGA.CLIENT.REST.TIMEOUT>
```

The names of the Maven properties have been designed to indicate in which file the property will be injected. Properties starting with `OPENCGA.STORAGE` will go to *storage-configuration.yml* while properties starting with `OPENCGA.CLIENT` will go to *client-configuration.yml*. The rest of the properties go to *configuration.yml*.

At this moment this is the default profile at main *pom.xml* file:

Table of Contents:

- [Overview](#)
 - [Configuring at run time](#)
 - [Configuring at build time with Maven](#)

Maven Default Profile

```
<profile>
  <id>default-config</id>
  <activation>
    <activeByDefault>true</activeByDefault>
  </activation>
  <properties>
    <opencga.war.name>opencga-${opencga.version}</opencga.war.name>

    <!-- General -->
    <OPENCGA.INSTALLATION.DIR>/opt/opencga</OPENCGA.INSTALLATION.DIR>
    <OPENCGA.USER.WORKSPACE>/opt/opencga/sessions</OPENCGA.USER.
WORKSPACE>
    <OPENCGA.JOBS.DIR>${OPENCGA.USER.WORKSPACE}/jobs</OPENCGA.JOBS.DIR>
    <OPENCGA.TOOLS.DIR>${OPENCGA.INSTALLATION.DIR}/tools</OPENCGA.
TOOLS.DIR>
    <OPENCGA.DB.PREFIX>opencga</OPENCGA.DB.PREFIX>
    <OPENCGA.EXECUTION.MODE>LOCAL</OPENCGA.EXECUTION.MODE>

    <!-- Client -->
    <OPENCGA.CLIENT.REST.HOST>http://localhost:8080/${opencga.war.name}
</OPENCGA.CLIENT.REST.HOST>
    <OPENCGA.CLIENT.REST.TIMEOUT>30000</OPENCGA.CLIENT.REST.TIMEOUT>
    <OPENCGA.CLIENT.GRPC.HOST>localhost:9091</OPENCGA.CLIENT.GRPC.HOST>
    <OPENCGA.CLIENT.ORGANISM.SCIENTIFIC_NAME></OPENCGA.CLIENT.ORGANISM.
SCIENTIFIC_NAME>
    <OPENCGA.CLIENT.ORGANISM.COMMON_NAME></OPENCGA.CLIENT.ORGANISM.
COMMON_NAME>
    <OPENCGA.CLIENT.ORGANISM.TAXONOMY_CODE></OPENCGA.CLIENT.ORGANISM.
TAXONOMY_CODE>
    <OPENCGA.CLIENT.ORGANISM.ASSEMBLY></OPENCGA.CLIENT.ORGANISM.
ASSEMBLY>

    <OPENCGA.SERVER.REST.PORT>9090</OPENCGA.SERVER.REST.PORT>
    <OPENCGA.SERVER.GRPC.PORT>9091</OPENCGA.SERVER.GRPC.PORT>
    <OPENCGA.MONITOR.PORT>9092</OPENCGA.MONITOR.PORT>

    <!-- Catalog -->
    <OPENCGA.CATALOG.DB.HOSTS>localhost:27017</OPENCGA.CATALOG.DB.
HOSTS>
    <OPENCGA.CATALOG.DB.USER></OPENCGA.CATALOG.DB.USER>
    <OPENCGA.CATALOG.DB.PASSWORD></OPENCGA.CATALOG.DB.PASSWORD>
    <OPENCGA.CATALOG.DB.AUTHENTICATION_DATABASE></OPENCGA.CATALOG.DB.
AUTHENTICATION_DATABASE>
    <OPENCGA.CATALOG.DB.CONNECTIONS_PER_HOST>20</OPENCGA.CATALOG.DB.
CONNECTIONS_PER_HOST>

    <!-- Storage -->
    <OPENCGA.STORAGE.DEFAULT_ENGINE>mongodb</OPENCGA.STORAGE.
DEFAULT_ENGINE>
    <OPENCGA.STORAGE.CACHE.HOST>localhost:6379</OPENCGA.STORAGE.CACHE.
HOST>
    <OPENCGA.STORAGE.SEARCH.HOST>http://localhost:8983/solr/</OPENCGA.
STORAGE.SEARCH.HOST>
    <OPENCGA.STORAGE.SEARCH.TIMEOUT>30000</OPENCGA.STORAGE.SEARCH.
TIMEOUT>
    <OPENCGA.STORAGE.SEARCH.HOST>http://bioinfodev.hpc.cam.ac.uk/solr/<
/OPENCGA.STORAGE.SEARCH.HOST>

    <!-- Storage Variants general -->
    <OPENCGA.STORAGE.VARIANT.DB.HOSTS>localhost:27017</OPENCGA.STORAGE.
VARIANT.DB.HOSTS>
    <OPENCGA.STORAGE.VARIANT.DB.USER></OPENCGA.STORAGE.VARIANT.DB.USER>
    <OPENCGA.STORAGE.VARIANT.DB.PASSWORD></OPENCGA.STORAGE.VARIANT.DB.
PASSWORD>

    <!-- Storage Alignments general -->
    <OPENCGA.STORAGE.ALIGNMENT.DB.HOSTS>localhost:27017</OPENCGA.
STORAGE.ALIGNMENT.DB.HOSTS>
```

```

        <OPENCGA.STORAGE.ALIGNMENT.DB.USER></OPENCGA.STORAGE.ALIGNMENT.DB.
USER>
        <OPENCGA.STORAGE.ALIGNMENT.DB.PASSWORD></OPENCGA.STORAGE.ALIGNMENT.
DB.PASSWORD>

        <!-- Storage-mongodb -->
        <OPENCGA.STORAGE.MONGODB.VARIANT.DB.AUTHENTICATION_DATABASE><
/OPENCGA.STORAGE.MONGODB.VARIANT.DB.AUTHENTICATION_DATABASE>
        <OPENCGA.STORAGE.MONGODB.VARIANT.DB.CONNECTIONS_PER_HOST>20<
/OPENCGA.STORAGE.MONGODB.VARIANT.DB.CONNECTIONS_PER_HOST>

        <!-- Storage-hadoop -->
        <!--If empty, will use the ZOOKEEPER_QUORUM read from the hbase
configuration files-->
        <OPENCGA.STORAGE.HADOOP.VARIANT.DB.HOSTS></OPENCGA.STORAGE.HADOOP.
VARIANT.DB.HOSTS>
        <OPENCGA.STORAGE.HADOOP.VARIANT.DB.USER></OPENCGA.STORAGE.HADOOP.
VARIANT.DB.USER>
        <OPENCGA.STORAGE.HADOOP.VARIANT.DB.PASSWORD></OPENCGA.STORAGE.
HADOOP.VARIANT.DB.PASSWORD>
        <OPENCGA.STORAGE.HADOOP.VARIANT.HBASE.NAMESPACE></OPENCGA.STORAGE.
HADOOP.VARIANT.HBASE.NAMESPACE>
        <OPENCGA.STORAGE.HADOOP.VARIANT.ARCHIVE.TABLE.PREFIX>${OPENCGA.DB.
PREFIX}_study</OPENCGA.STORAGE.HADOOP.VARIANT.ARCHIVE.TABLE.PREFIX>

        <!-- Email server -->
        <OPENCGA.MAIL.HOST></OPENCGA.MAIL.HOST>
        <OPENCGA.MAIL.PORT></OPENCGA.MAIL.PORT>
        <OPENCGA.MAIL.USER></OPENCGA.MAIL.USER>
        <OPENCGA.MAIL.PASSWORD></OPENCGA.MAIL.PASSWORD>

        <!-- cellbase -->
        <OPENCGA.CELLBASE.VERSION>v4</OPENCGA.CELLBASE.VERSION>
        <OPENCGA.CELLBASE.REST.HOST>http://bioinfo.hpc.cam.ac.uk/cellbase/<
/OPENCGA.CELLBASE.REST.HOST>
        <OPENCGA.CELLBASE.DB.HOST>localhost:27017</OPENCGA.CELLBASE.DB.
HOST>
        <OPENCGA.CELLBASE.DB.USER></OPENCGA.CELLBASE.DB.USER>
        <OPENCGA.CELLBASE.DB.PASSWORD></OPENCGA.CELLBASE.DB.PASSWORD>
        <OPENCGA.CELLBASE.DB.AUTHENTICATION_DATABASE></OPENCGA.CELLBASE.DB.
AUTHENTICATION_DATABASE>
        <OPENCGA.CELLBASE.DB.READ_PREFERENCE>secondaryPreferred</OPENCGA.
CELLBASE.DB.READ_PREFERENCE>
    </properties>
</profile>

```