Archive Strategy

Bonitasoft

November 2020

Sommaire

[1. Introduction 3](#_Toc57020607)

[2. Principle 3](#_Toc57020608)

[3. Archive all 3](#_Toc57020609)

[4. Configure what you want to keep 4](#_Toc57020610)

[5. Delete cases based on a date and processes 4](#_Toc57020611)

[5.1. The Bonita Purge Tool 5](#_Toc57020612)

[5.2. Delete Case plugin (Truckmilk) 6](#_Toc57020613)

[6. Move case information to another Bonita Server 7](#_Toc57020614)

[7. Move case information to an external Database 10](#_Toc57020615)

[8. Move information to a file 11](#_Toc57020616)

Content

# Introduction

Your system has different processes. Then, you produce cases, and cases are ended, they are archived. What is the different strategy to use for the archive data?

# Principle

When a case is executed, Bonita starts to archive information:

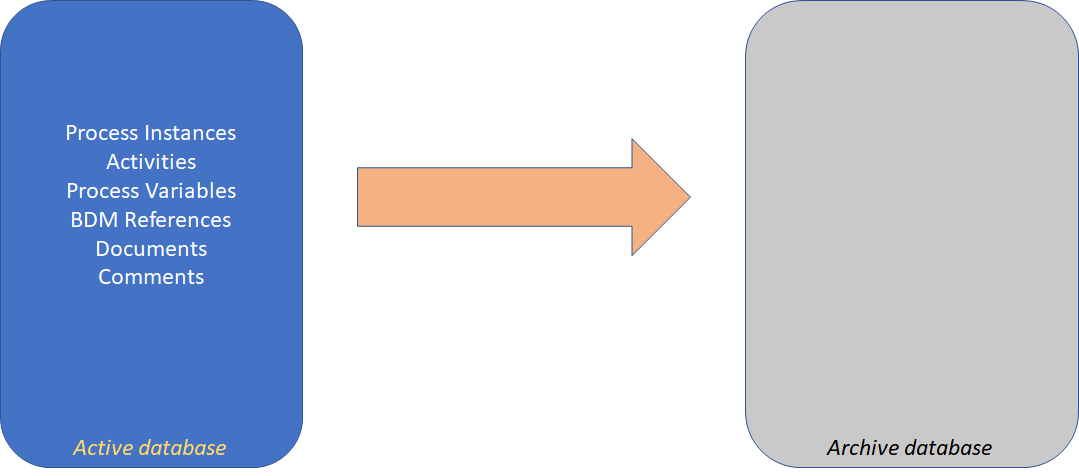
* variable history,
* task history (when a task started, who executed it).
* Documents
* Comments

Behind the scene, the achievement is running immediately. For example, when a task is executed, history is saved in an archived table at this moment (the case is still active). When the case is finally done, only one record is moved in the archived table.

To keep performance on the active cases, Bonita chooses to use the second set of tables, tables started by ARCH.

# Archive all

In the first strategy, keep your engine with the default configuration. Then, all information is kept in the Bonita Engine.

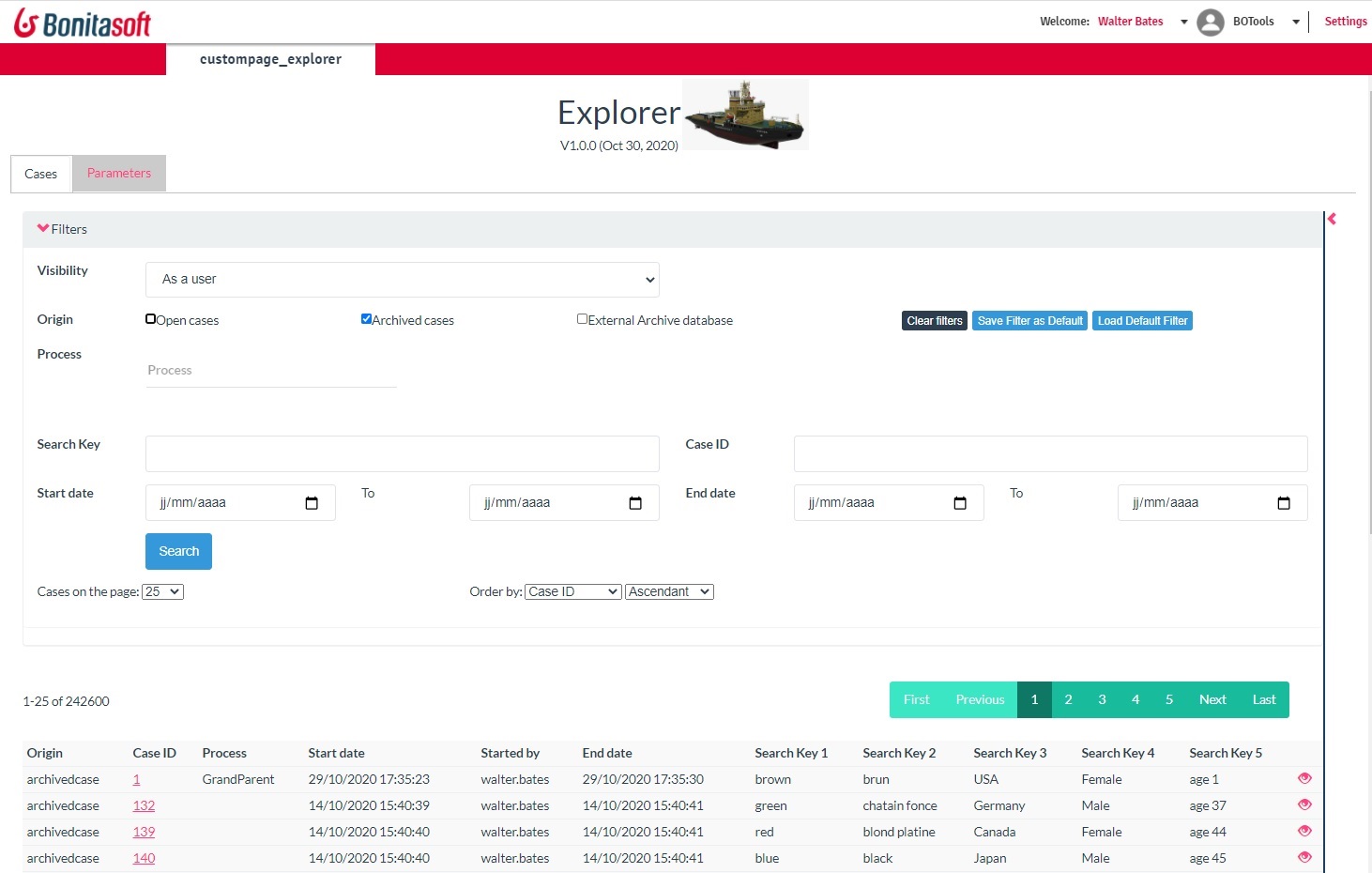


Note: to accelerate the search, some index must be created in the archive table.

Via the Portal, the search is possible only on the search key (1 to 5). Because the Portal can order the result on multiple criteria (case ID, Process Name, Start date, Started by, End date), you should create an index on each attribute.

To have a multi-search tool on an archived case, for user or administrator, use the Explorer Page.

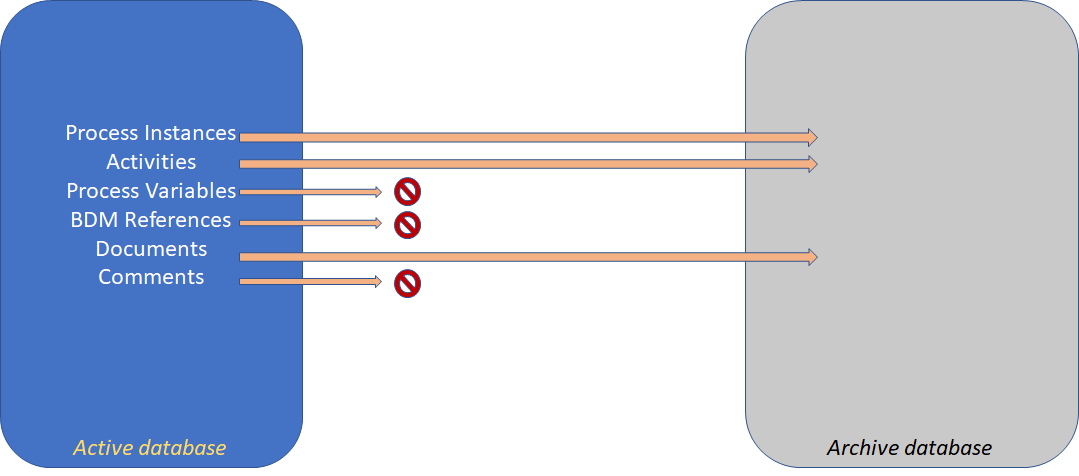
<https://community.bonitasoft.com/project/page-explorer-seach-cases-multiple-criteria-multiple-origin>



# Configure what you want to keep

Via the Configurable service, you can decide what you want to archive.

Visit https://documentation.bonitasoft.com/bonita/7.11/configurable-archive



The policy is applied on all processes. You can’t decide to archive process “Review Expense” and to not archive “Daily report”.

#### Concerns:

No archive is saved. If you want to delete archived after a while, this is not an option.

This is a general strategy, it’s not possible to manage itprocess per process

#### Advantages***:***

This method is efficient

# Delete cases based on a date and processes

You want to keep some archive data for a while, then purge them. Purge may be done on a delay criteria (purge all data older than 4 years) or a mixed set of criteria, based on process and delay (process “Review Expense”, purge after 4 years; process “Daily report”, purge after 1 day.

Two options are available in the Community:

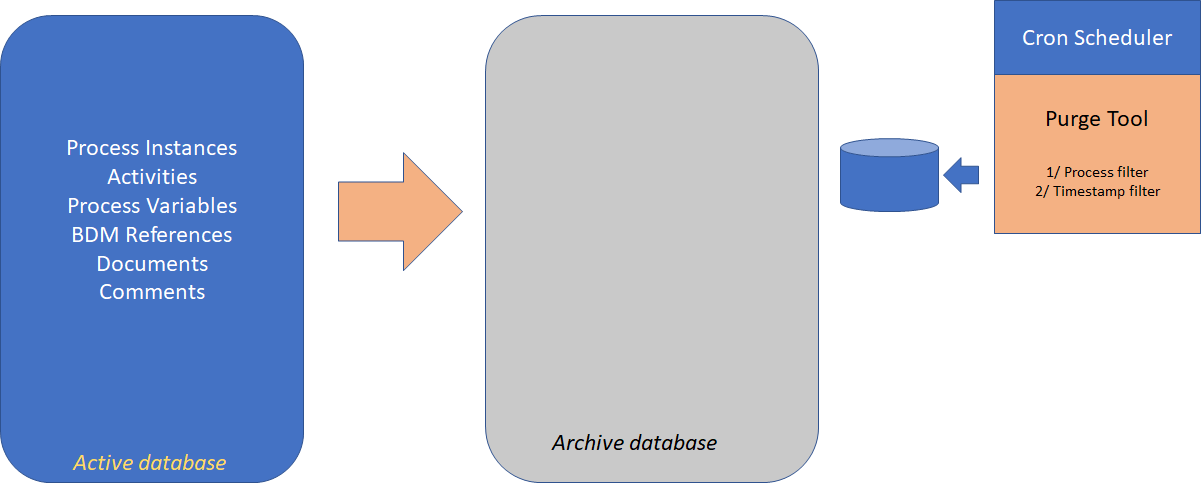
## The Bonita Purge Tool

<https://community.bonitasoft.com/project/bonita-purge-tool>

This is a script-based tool, which delete process per process, with a Timestamp.

It execute requests directly in the database, which is a Bonitasoft bad practice.

It use one transaction to identify and delete all process instances. This strategy is efficient, but need a large roolback in the database engine. If there is too much cases to delete (especialy too many documents), the request failed.



#### Concerns:

Parameters are complexes to manipulate: ProcessId and a TimeStamp (not a delay). Your script must be encapsulated by a way to calculate a timestamp from a delay.

If you have multiple processes, you have to run the scripts for each.

When you deploy a new version, a new command must be added (argument is based on a ProcessDefinitionId)

The command runs a transaction to delete in one shoot all cases, tasks, documents, comments. This is very efficient but need a large transaction memory, especially the first execution if you have a lot of data to delete. An example demonstrate that you need 5 Gb memory to deletes 137 000 cases (without documents) in MySQL.

Execution can’t be limited in time. When you run it, it will purge all in one transaction and can’t be limited like “execution-only between 2 am to 5 am.”

To automate the tool, the Cron system can be used on Unix. On Windows, the configuration is complex.

The execution of the tool is outside the Bonita engine. Connection requirement is not explained in the tool. It requires a connection to the Bonita engine or directly to the database? If it requires a connection to the Bonita Engine, an SSO connection prevents the execution; If it requires a Database connection, some cache issue can be detected (it’s discouraged to update directly the Bonita database).

There is no history, to keep track of all deletes case

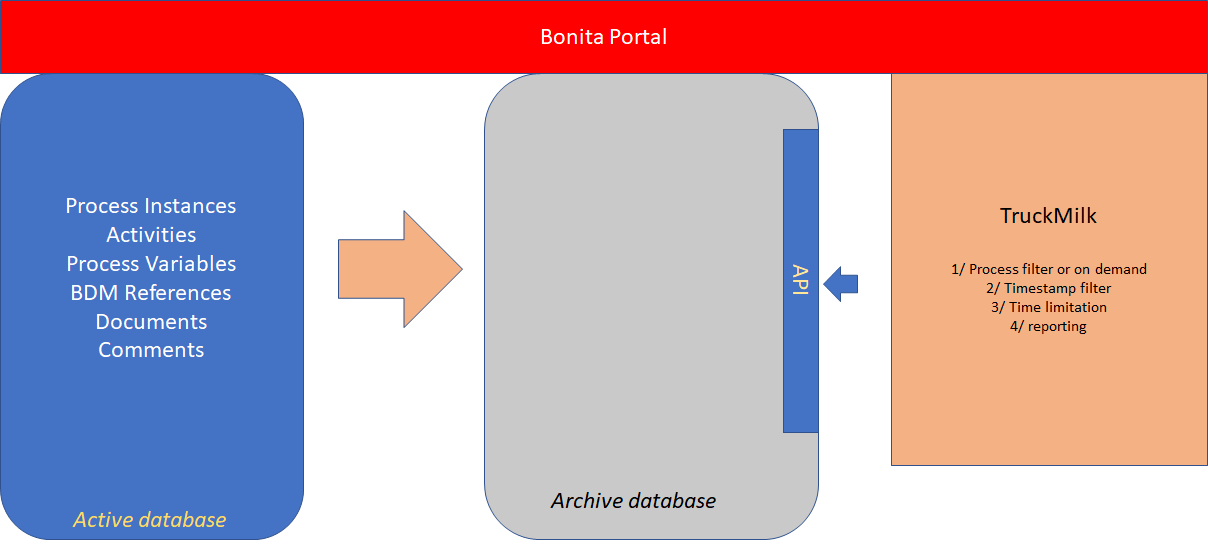
Because this tool deletes directly in the database, which may change between version, the correct version of the tool must be used for a special Bonita Version.

#### Advantages

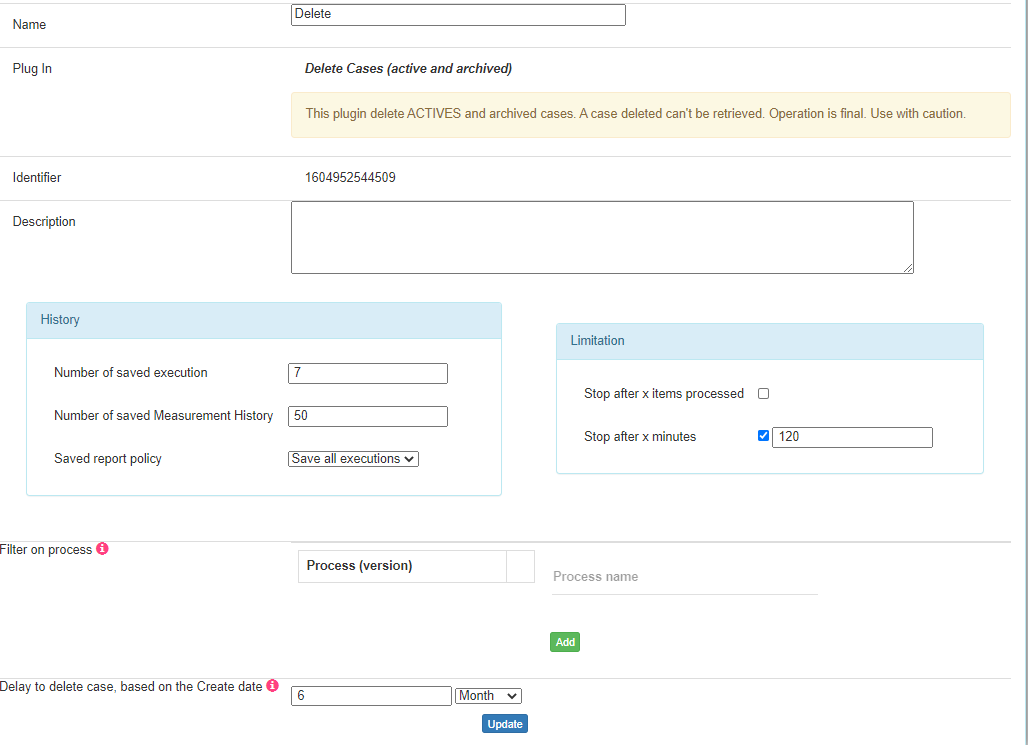
Due to the implementation, the method has the best performance for the speed (not for the disk or memory usage).

## Delete Case plugin (Truckmilk)

This plugin tool deletes cases, based on a set of processes and a delay. Execution use the Bonita API, which is a Bonitasoft Good Practice. then respect the Bonita Engine (cache invalidation, security, compliance).



Truckmilk job to configure deletion:



Thank Truckmilk, different functions are available:

* Time limitation: it’s possible to limit the deletion on a timeslot. Experiences show to delete 1 Million cases (with no documents), limited execution 3 hours per night requires 20 days, but this tool can handle the volume without any special configuration.
* The report, to collect the list of cases deleted
* Truckmilk has a scheduler; no additional deployment is required whatever the platform
* At any time, the administrator can monitor jobs: advancement percentage, stop the job, reschedule it, number of cases deleted, and number of cases still in the scope
* A Truckmilk is executed on the Bonita Server. Security is verified

#### Concerns:

The plugin uses the BonitaAPI, which deletes cases per packets (the plugin uses packs of 50). The BonitaAPI is slower than a direct SQL in the database, especially in version under 7.9

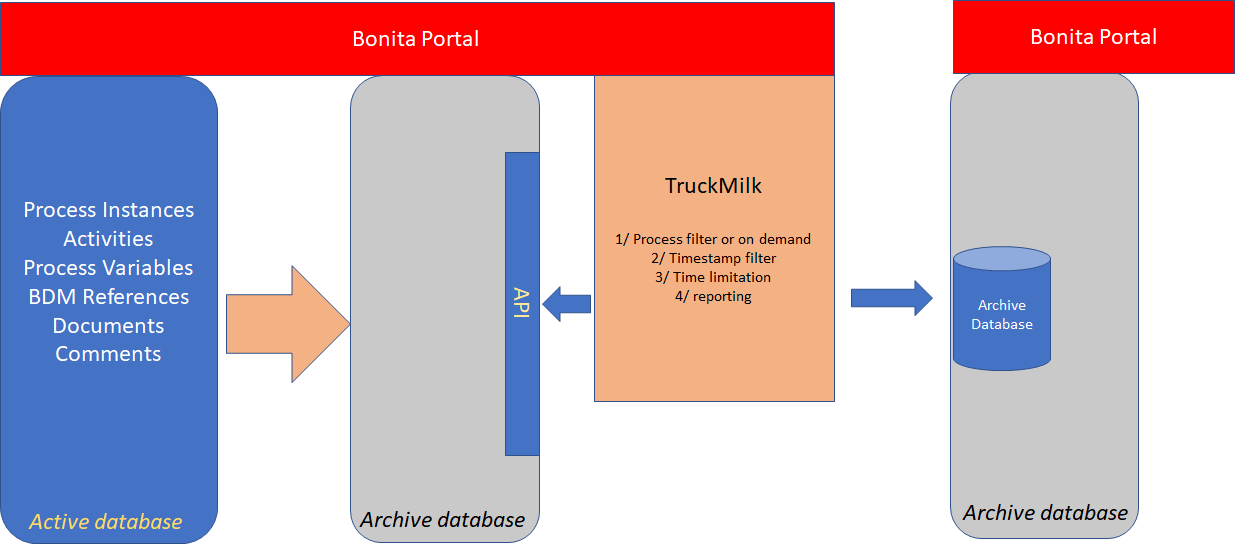
#### Advantages

All advantages are given by Truckmilk (security, feedback, interface).

This is the only tool available if you have a lot of cases to remove or with large documents.

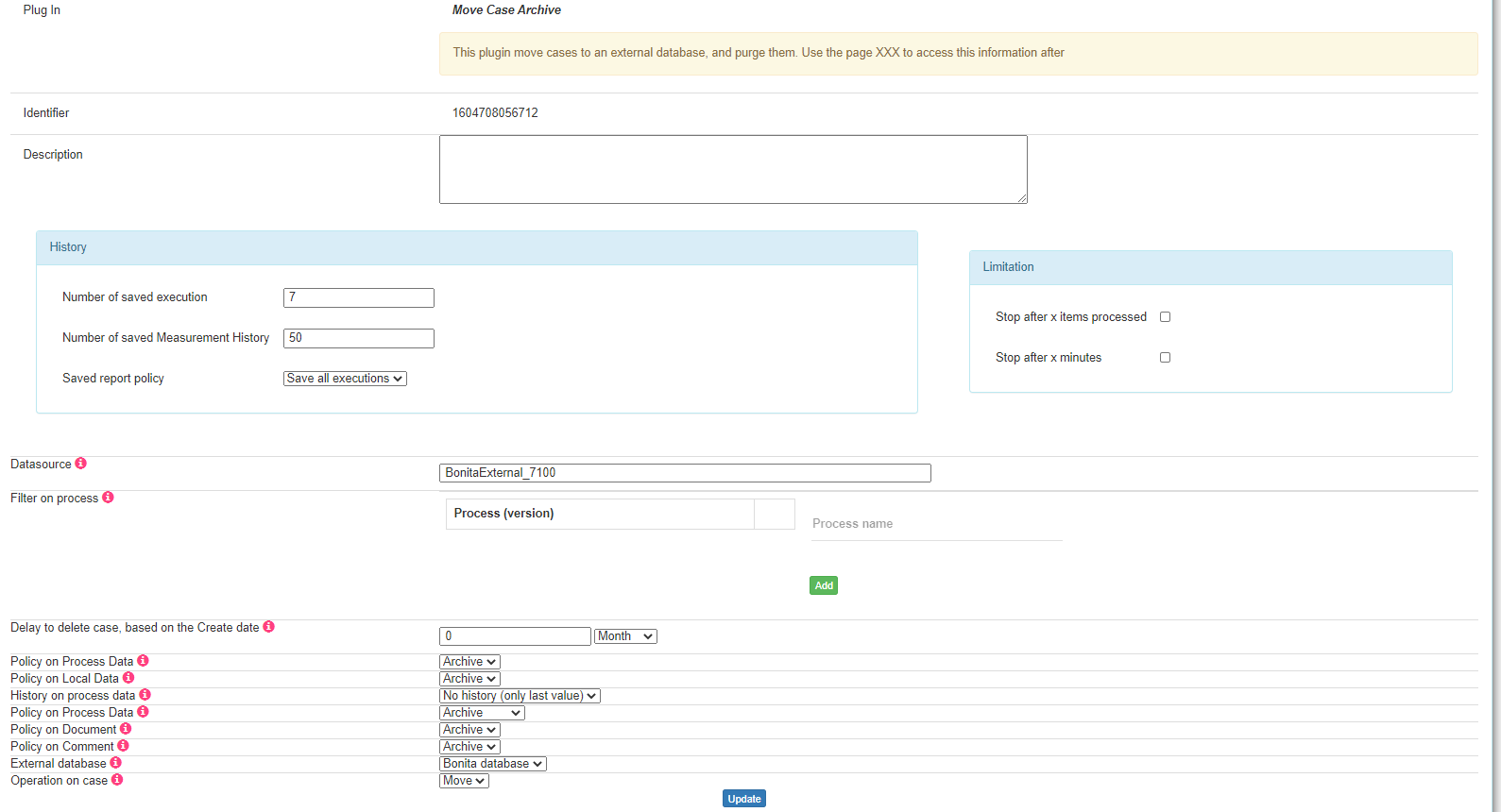
# Move case information to another Bonita Server

To keep the server with low number of information, it is possible to move cases to a second Bonita Server. This second Bonita Server supports only archives cases.



To be consistent, this server must have the same environment in case of users and processes. For example, to move a case from the process “Review Expense”, that implied that this process is deployed on the second server. Else, Portal will not show cases. Same with users: if the case was created by Jan.Fisher, this user must exist on the second server. They may have different ID: during operation, Truckmilk jobs search the local ID based on named (process name+ version, username).

If the second server contains all different information, then it’s possible to move cases to a second Server via the Truckmilk plugin “Move Case Archive.”



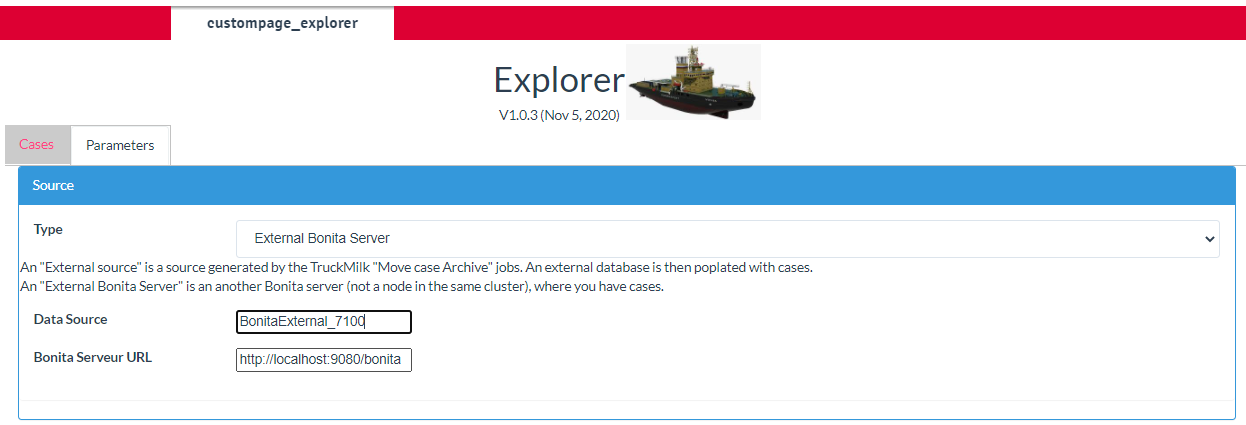
With this plugin, a set of processes and delays can be used. The scope to move is part of the configuration.

A Bonita Server does not have any API to creates archive cases. Secondly, the API prevents a application to connect to two different Bonita Server, by construction. So, the plugin makes some direct SQL request to copy the data to the external database. When a case is moved, the administrator can decide to purge it on the current server (Move policy) or just do a duplication (Copy policy). Deletion is made than via the Bonita API to respect the Bonita server.

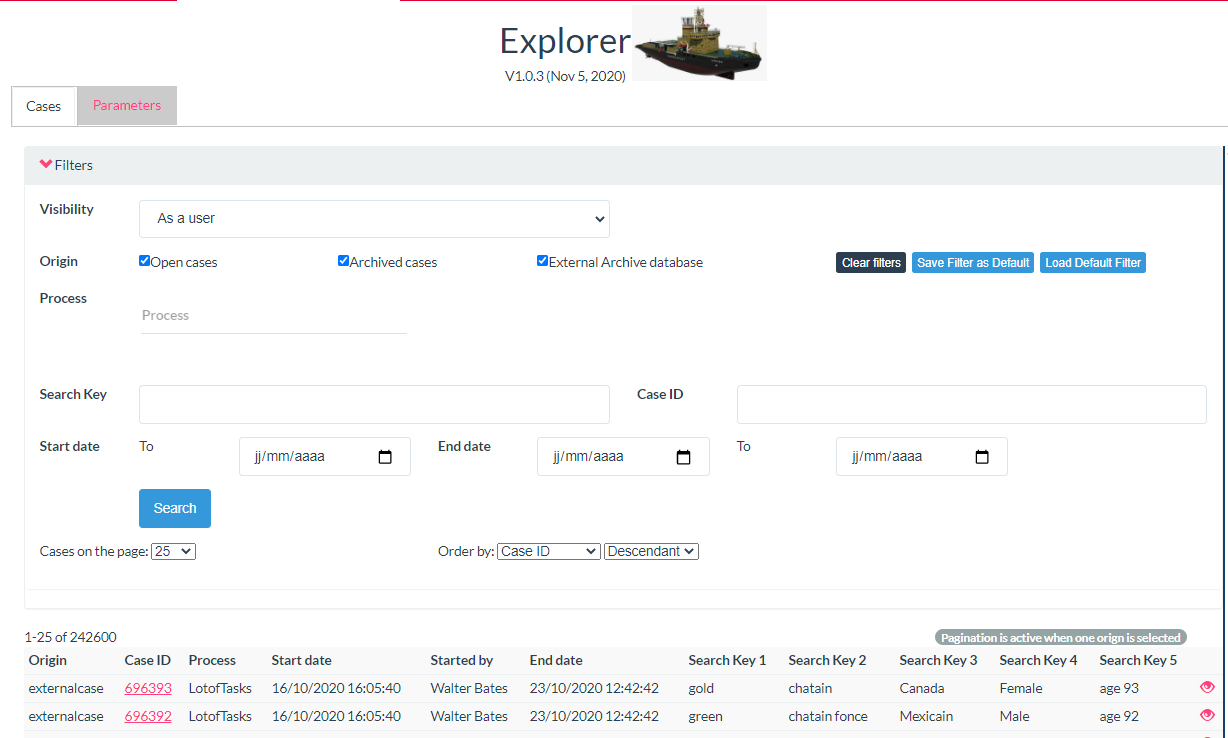
#### Access information

To access information, users must connect them on both Bonita Server. This can be simplified by:

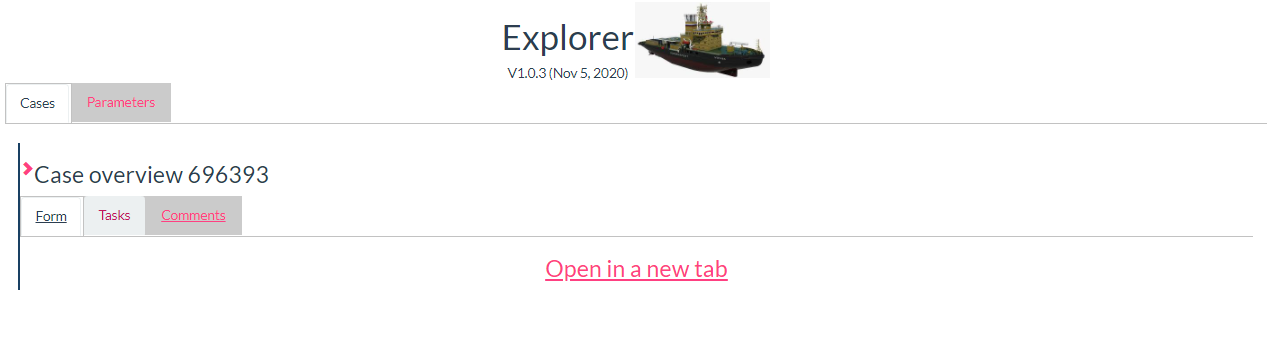
* An SSO to simplify the connection on both servers, but user need to execute a search twice
* Usage of the Explorer Page (<https://community.bonitasoft.com/project/page-explorer-seach-cases-multiple-criteria-multiple-origin>) . On this page, an “external Bonita Server” can be registered.



Then, any search can be executed on different servers automatically.



To access the overview page, a link is given, which redirect to the second server (user need then to connect on this second server)



#### Bonita Community as the second server

For the second server, a Bonita Community edition can be used. In this edition, the organization must be loaded with all processes.

A Subscription Edition (BAR file) can be uploaded on a Community server but can’t be executed: that’s exactly what we look for here.

#### Concerns:

The copy is made in SQL. That implied that the Bonita Server must have a connection to the second server via a Data source

Database structure may change, then the tool needs to be upgraded

The second server must be on the same line as the original server when a new process is deployed, for example, or users are created.

The second server must have a license; even it doesn’t execute any cases.

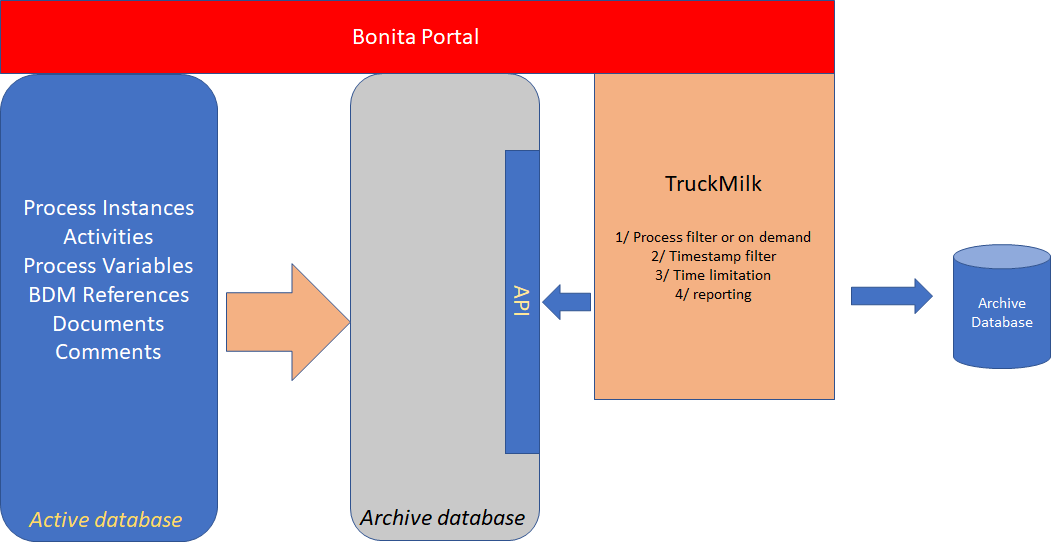
If the case overview needs to access some Business Data information, then the second server must reference the same Business Data. The plugin does not copy the Business Data (because they are not part of the process data – only references are copied).

#### Advantages:

It takes advantage of all TruckMilk possibility (reporting, limited execution), and respect the Bonita Server to purge cases.

# Move case information to an external Database

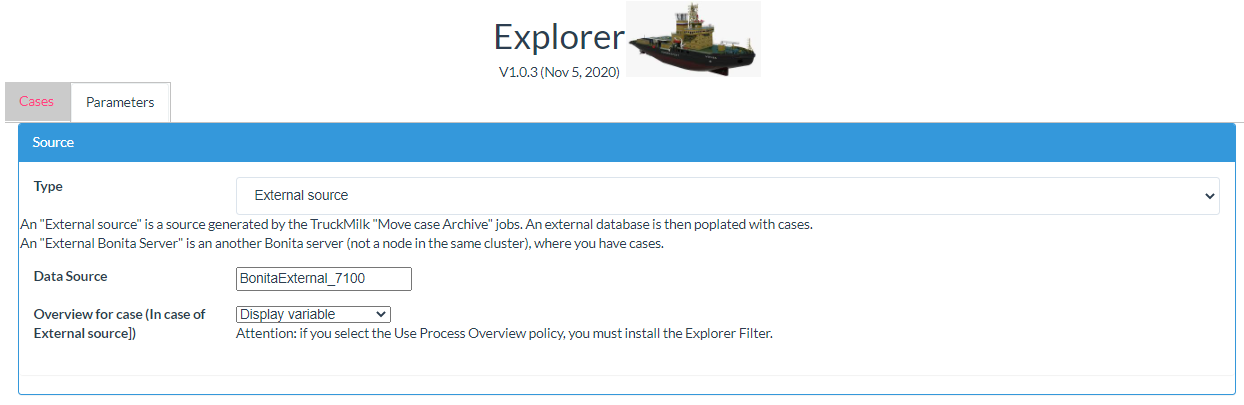
Cases can be move in an external database. This database contains tables to save process, tasks, comments, documents.



The Truckmilk plugin “Move Case Archive” does that job. The plugin creates all tables in the target database.

This external database is not under the control of a Bonita server and can’t be accessed directly for a user.

The Explorer pages can query this database via the parameter “External source”.



The point is about the Case Overview. There is then no case overview form.

#### Case overview options

Different options are:

* A case overview is not mandatory. Have a list of all variable is enough. Then, select the Display Variables option is enough.
* A document generated in the case is the overview. Then, this document is selected in the plugin, and in the external database, this document is flag as the overview. When the Explorer page access the case overview, it displays the document
* An overview document is generated during the plugin “move case”. This generation used Bonita connectors “Office” and then “PDF” to generate the document. In the Explorer page, this document is accessible.

#### Concerns

Display variable may be not enough, especially when the case references some Business Data. First because user will just see the link, second because the data may be deleted in the meantime.

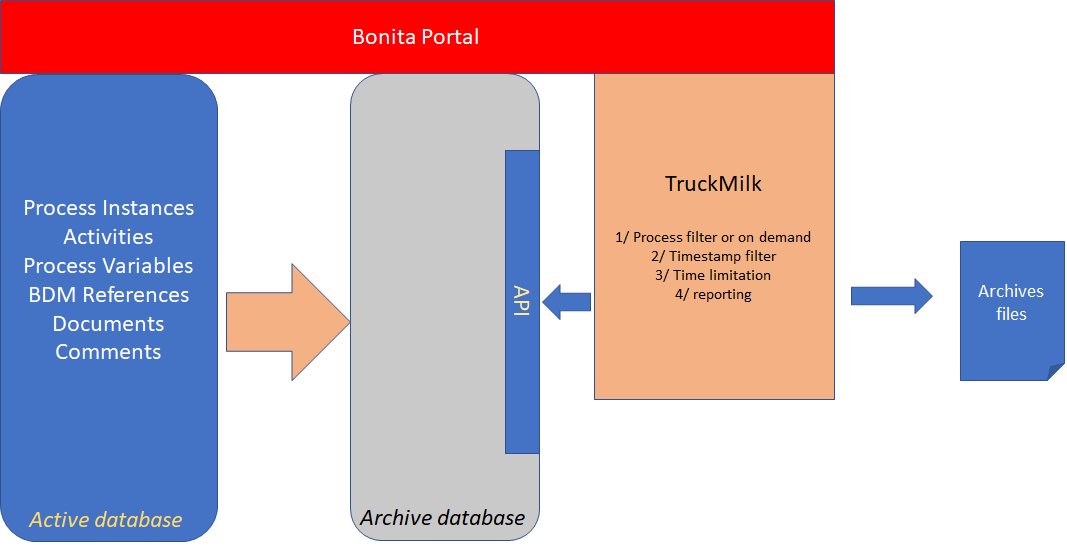
Asking the plugin to generate a document is not simple, because this implied to generate the Word document with Velocity (this is not a simple job).

#### Advantages

The data are saved outside Bonita scope, and can be reusable for any application, because the structure is stable.

# Move information to a file

Same as the previous function, except that information are saved in files, which is simplest for configuration. Then, no request can be realized, and, because information is saved on a file system, no security on data integrity (a file may be deleted).



Then, one file (a ZIP file) is created per case. The ZIP file contains a Case Description (JSON format) and all documents on cases. Files are saved under a folder (on folder per process).

Note: this jobs does not exists yet at this moment.

#### Concerns

No security, and no search are possible in files.

#### Advantages

Simple configuration.

HEADQUARTERS

PARIS, FRANCE

76 boulevard de la République

92100 Boulogne-Billancourt

EMEA, ASIA & LATIN AMERICA

GRENOBLE, FRANCE

32, rue Gustave Eiffel  
38000 Grenoble

NORTH AMERICA

SAN FRANCISCO, USA

44 Tehama Street  
San Francisco, CA 94105

NEW YORK, USA

33 Nassau Avenue

Brooklyn NY 11222