Obelix Documentation

Release 0.1

Fredrik Carlsen

May 29, 2015

Contents

Obelix Personalization Search Engine

CHAPTER 1

Quickstart

1.1 Installation

The installation is as simple as downloading the jar file Note: This jar-file require java8 to be installed.

```
wget http://obelix.io/obelix.jar
```

1.2 Dependencies

Obelix require the following software to be installed

- Redis
- Java8

1.3 Usage

Using Obelix can be as little effort as calling `java -jar obelix.jar`. However, for most projects it is required to configure more options.

```
java -jar obelix.jar

# Configure where to store the graph database, the default location is the same folder as the jar fi
--neo4jstore /path/to/store/graph.db

# Set a maximum number of relationships per node, this will remove the oldest when the limit is reach
--max-relationships 100

# Set the number of workers to read from the log queue
--workers 1

# Set the name of the redis queue to look for new log entries
--redis-queue-name logentries

# Set the http port for the Obelix HTTP API
--web-port 4000

# Set the recommendations depth, this means how deep the graph will be traversed
--recommendation-depth 3
```

```
# Set Obelix in batch import mode, this means that it will import all entries in the logentries queue
--batch-import-all

# Tell Obelix to rebuild all recommendations
--build-cache-for-all-users-on-startup

# Enable Metrics
--enable-metrics
```

1.4 Running Obelix as a Daemon (background service) on Ubuntu 14.04

In a production environment it is wise to run Obelix as a background service. You can do this easily by using supervisor.

To set up supervisor, first you need to install the package

```
sudo apt-get install supervisor
```

Then you need to create a configuration file named /etc/supervisor/conf.d/obelix.conf with the following content

```
[program:obelix]
user = someUserWithAccessToTheDirectory
autostart = true
autorestart = true
command = java -jar /mnt/obelix/obelix.jar --option1 value1 --option2 value2...
stdout_logfile = /var/log/obelix.log
stderr_logfile = /var/log/obelix.error.log
```

Then you simply restart the supervisor service

```
sudo service supervisor restart
```

Then you can tail the log to see that Obelix is running

```
sudo tail -f /var/log/obelix.log
sudo tail -f /var/log/obelix.error.log
```

1.5 Recommended JVM settings

For Obelix to perform well, it is recommended to enable the `-XX:+UseConcMarkSweepGC` option on the JVM.

It is also recommended to set your `-Xmx` and `Xms` settings to apropriate values for your host.

An example of a configuration may be:

```
java -Xmx5000m -Xms5000m -XX:+UseConcMarkSweepGC -jar /mnt/obelix/obelix.jar --neo4jstore /mnt/obelix
```

CHAPTER 2

How it works

CHAPTER 3

Metrics

3.1 Introduction

The metrics are produced and stored in an ObelixStore object, this may be an internal ObelixStoreImpl or Redis. By default, this is stored in redis.

3.2 Enable metrics

For metrics to be collected and stored, the –enable-metrics argument needs to be passord to the jar.

```
java -jar obelix.jar

# Enable Metrics
--enable-metrics
```

3.3 Stored Metrics

By default the metrics from Obelix are gathered from different modules and stored as json in redis.

An example of the stored data:

```
"metric" : {
    "total_feeded" : 736,
    "feeded" : 7,
    "total_recommendations_built" : 651,
    "recommendations_built" : 7,
    "all_relationships_count" : 9837,
    "all_users_count" : 2442,
    "all_items_count" : 5154,
    "cache_queue_size" : 3,
    "logentries_queue_size" : 1
    "timestamp" : "2015-05-25T02:20:45.637367",
}
```

total_feeded: Number of interactions sent to Obelix (user x viewed item y) since the beggning.

feeded: The same as **total_feeded**, but the number represent the number of feeded since last time checked (typically a 5 minute interval).

total_recommendations_built: Number of recommendations built, it will be close to the number of total_feeded, but if a user view several items in a short amount of time, Obelix will try to only build the recommendations for that user once.

recommendations_built: The same as total_recommendations_built, but the number represent the number of recommendations since last time checked (typically a 5 minute interval).

all_relationships_count: The current number of relationships in Obelix

all_users_count: The current number of users in Obelix

all_items_count: The current number of items in Obelix

cache_queue_size: The number of items in the cache queue (to build recommendations).

logentries_queue_size: The number of items in the queue ready for feeding.

CHAPTER 4

Quickstart

4.1 Installation

The installation is as simple as downloading the jar file Note: This jar-file require java8 to be installed.

```
wget http://obelix.io/obelix.jar
```

4.2 Dependencies

Obelix require the following software to be installed

- Redis
- Java8

4.3 Usage

Using Obelix can be as little effort as calling `java -jar obelix.jar`. However, for most projects it is required to configure more options.

```
java -jar obelix.jar

# Configure where to store the graph database, the default location is the same folder as the jar fi
--neo4jstore /path/to/store/graph.db

# Set a maximum number of relationships per node, this will remove the oldest when the limit is reac
--max-relationships 100

# Set the number of workers to read from the log queue
--workers 1

# Set the name of the redis queue to look for new log entries
--redis-queue-name logentries

# Set the http port for the Obelix HTTP API
--web-port 4000

# Set the recommendations depth, this means how deep the graph will be traversed
--recommendation-depth 3
```

```
# Set Obelix in batch import mode, this means that it will import all entries in the logentries queue
--batch-import-all

# Tell Obelix to rebuild all recommendations
--build-cache-for-all-users-on-startup

# Enable Metrics
--enable-metrics
```

4.4 Running Obelix as a Daemon (background service) on Ubuntu 14.04

In a production environment it is wise to run Obelix as a background service. You can do this easily by using supervisor.

To set up supervisor, first you need to install the package

```
sudo apt-get install supervisor
```

Then you need to create a configuration file named /etc/supervisor/conf.d/obelix.conf with the following content

```
[program:obelix]
user = someUserWithAccessToTheDirectory
autostart = true
autorestart = true
command = java -jar /mnt/obelix/obelix.jar --option1 value1 --option2 value2...
stdout_logfile = /var/log/obelix.log
stderr_logfile = /var/log/obelix.error.log
```

Then you simply restart the supervisor service

```
sudo service supervisor restart
```

Then you can tail the log to see that Obelix is running

```
sudo tail -f /var/log/obelix.log
sudo tail -f /var/log/obelix.error.log
```

4.5 Recommended JVM settings

For Obelix to perform well, it is recommended to enable the `-XX:+UseConcMarkSweepGC` option on the JVM.

It is also recommended to set your `-Xmx` and `Xms` settings to apropriate values for your host.

An example of a configuration may be:

```
java -Xmx5000m -Xms5000m -XX:+UseConcMarkSweepGC -jar /mnt/obelix/obelix.jar --neo4jstore /mnt/obelix
```