

ETICS Log Analyzer

Creating a Report

Queries should be defined in an XML file, the path of which should be passed as an argument to ServerLogReader.

Verbose mode (providing a full list of each user's actions) is disabled by default - to enable it pass 'verbose' as the second argument

The XML file should consist of one report, which contains several queries.

The report element is simply defined : `<report>...</report>`

Query objects have both a type and a filter. The only filter currently available is date, which consists of start and end attributes. The date format is YYYY-MM-DD.

e.g. `<query><type>users</type><filter><date><start>2011-06-15</start><end>2011-06-21</end></date></filter>`

Query Definitions

The following query types are implemented:

- **users**: Lists the number of distinct users for each day in the range (`<query><type>users</type><filter><date><start>2011-06-15</start><end>2011-06-21</end></date></filter></query>`)
- **countries**: Lists the numbers of users from each country active during the range (`<query><type>countries</type><filter><date><start>2011-06-15</start><end>2011-06-21</end></date></filter></query>`)
- **sessions**: Lists the number of unique sessions for each day in the range (`<query><type>sessions</type><filter><date><start>2011-06-15</start><end>2011-06-21</end></date></filter></query>`)
- **configurations**: Lists the number of configuration creation/deletion/modification performed in the range (`<query><type>configurations</type><filter><date><start>2011-06-15</start><end>2011-06-21</end></date></filter></query>`)
- **operations**: Lists the total number of operations performed on each day in the range (`<query><type>operations</type><filter><date><start>2011-06-15</start><end>2011-06-21</end></date></filter></query>`)
- **hosts**: Lists all active hosts, the number of seconds they were utilized for and their utilisation overall through the period (`<query><type>hosts</type><filter><date><start>2011-06-15</start><end>2011-03-21</end></date></filter></query>`)
- **jobs**: Lists the numbers of build and debug jobs on each day of the range (`<query><type>jobs</type><filter><date><start>2011-06-15</start><end>2011-06-21</end></date></filter></query>`)
- **data**: Lists the total amount of data transferred, broken down by version, type and platform (`<query><type>data</type><filter><date><start>2011-06-05</start><end>2011-04-21</end></date></filter></query>`)

Log Sources

The log analyser uses three sources : the textual logs of BuildSystemSoapBindingImplementation, database tables obtained from the repository web interface and and jobs information from the NMI database.

To configure:

- **text logs** - modify the array dates[] in LogImporter to contain the date suffixes (e.g. .2011-06-25) of the file you wish to import. Note that only BuildSystemServiceSoapBindingImpl.log files are scanned, as these are the only ones that contain relevant logs.
- **Jobs information** - these logs can be obtained by command line from the server nmi_history, then imported to the main database
- **Software repository info** - this should be downloaded from the web interface in **csv** format and imported into MySQL by executing the SQL command given in database_commands.txt (changing the name of the file by the correct one)

Building

The software has the following dependencies:

- jcommon-1.0.16.jar
- mysql-connector-java-5.1.16-bin.jar

To build on the command line simply compile all the .java files in /src

(e.g. `javac *.java`)

Depending on your settings it may be necessary to explicitly include the jars on the build path - if so, use the argument `-classpath /lib`, where lib

is a folder containing the jars.

Running

To run the report generator execute `java ServerLogReader <path to xml> [verbose]`

e.g. `java ServerLogReader.java weekly.xml verbose`

The generated file (data.csv) is placed in the main `org.etics.utilities.log-analyser` directory

To run the log importer simply execute `java LogImporter` - it takes no arguments. Make sure that you have placed the new log files in the `org.etics.utilities.log-analyser` directory, and that you have set the date suffix correctly - as in **Log Sources a)**

Some useful possible outputs

- Number of users per week
- Breakdown of jobs by type
- Daily users per country
- Amount of data transferred for each version
- Average utilisation of virtual hosts
- Number of operations performed each day
- Number of sessions per day
- Number of sessions created and deleted each week

Deployment Notes

This software is currently configured to use a local MySQL database.

Before running the software on a new machine you must create a new database that can be accessed from that machine and execute the three build table commands saved in database `commands.txt`. You are then free to run LogImporter or import database dumps using MySQL.

Potential Future Work

- Move some definitions and configuration (like the text file pattern matchers) out of the code and into external configuration files.
- Improve LogImporter so that it automatically imports all files in a given directory.
- Write scripts (potentially bash) to automatically transfer log files from server and automatically dump and import databases (will require additional permissions)