SuiteCRM Analytics

SuiteCRM Data Integration Installation

Revision: **1.3**

Table of Contents

Prerequisites

SuiteCRM Data Integration requires the following pre-requisites installed on your server

- Linux environment
- OpenJDK 8 JRE (Java)
- wget
- unzip
- zip
- Mysql server + an empty MySQL database

Downloading SuiteCRM Data Integration

There are two ways of downloading the SuiteCRM Data Integration:

• [Download from SuiteCRM website](): a pre-built package containing installation scripts, web server and all the third party open source libraries required.

• Clone the SuiteCRM Data Integration repo: a script to build the SuiteCRM data integration package (generating the same available package via the SuiteCRM website)

Manually Build the Package (Cloning from Repo)

Clone repo via terminal:

```
git clone https://github.com/ivylabs/suitecrm-data-integration
```

Navigate into the newly cloned repository's directory and run:

```
./build.sh
```

This should output on your terminal all the files that the solution downloads and zips up into a package.
The end of the script will ask if you wish to delete the installation files.

```
Would you like to remove the installation files? This will save disk
space. [y/N] y


-------------------------------------------------------------

 Build is complete!


-------------------------------------------------------------
```

This will generate a package named "***suitecrm-data-integration-server.zip***" within the root of your directory. This is the same package available via the SuiteCRM website.

Configuring SuiteCRM Data Integration

After downloading or generating the "***suitecrm-data-integration-server.zip"*** file, Upload and Extract the file on a hosted web server.

Navigate into the newly extracted "***suitecrm-data-integration-server"*** folder and edit the the "***install.properties"*** file:

```
# Java Virtual Memory allocation
JVM_SIZE=1024
# SuiteCRM Source Database Connection Details
SUITECRM_HOST=127.0.0.1
SUITECRM_PORT=3306
SUITECRM_DATABASE=suitecrm_testdata
SUITECRM_USERNAME=suitecrmrootuser
SUITECRM_PASSWORD=suitecrmrootuserpassword
# SuiteCRM Analytis Target Database Connection
SUITECRM_ANALYTICS_HOST=127.0.0.1
SUITECRM_ANALYTICS_PORT=3306
SUITECRM_ANALYTICS_DATABASE=suitecrm_dwh
SUITECRM_ANALYTICS_USERNAME=suitecrmrootuser
SUITECRM_ANALYTICS_PASSWORD=suitecrmrootuserpassword
# Enable SMTP email functionality
SMTP_ENABLED=0
# SMTP Server Details
SMTP_HOSTNAME=smtp.gmail.com
SMTP_PORT=587
SMTP_USE_AUTHENTICATION=1
SMTP_USERNAME=user@suitecrm-analytics.co.uk
SMTP_PASSWORD=password
#SMTP_SECURE_AUTHENTICATION=
#SMTP_SECURE_AUTHENTICATION=SSL
SMTP_SECURE_AUTHENTICATION=TLS
# Email address that SuiteCRM Analytics sends from
SMTP_FROM_EMAIL_ADDRESS=donotreply@suitecrm-analytics.co.uk
SMTP_FROM_EMAIL_NAME=SuiteCRM Analytics
# Email addresses seperated by spaces that error emails are delivered
to
SMTP_SEND_ERROR_EMAILS_TO=user@suitecrm-analytics.co.uk
# Set this to 1 or zero. Setting it to 1 means that you will get an
email for every sucessful run on the ETL
SMTP_SEND_SUCCESS_EMAILS=1
# Email addresses seperated by spaces that success emails are delivered
to
SMTP_SEND_SUCCESS_EMAILS_TO=user@suitecrm-analytics.co.uk
user123@suitecrm-analytics.co.uk
```

These config variables from "*install.properties*" define the connection to your SuiteCRM database to extract information from:

```
# SuiteCRM Source Database Connection Details
SUITECRM_HOST=127.0.0.1
SUITECRM_PORT=3306
SUITECRM_DATABASE=suitecrm_testdata
SUITECRM_USERNAME=suitecrmrootuser
SUITECRM_PASSWORD=suitecrmrootuserpassword
```

Update the above values in your "*install.properties*" file to reflect the following:

"SUITECRM_DATABASE" = The Database name of your SuiteCRM Instance

"SUITECRM_USERNAME" = The username used to connect to the SuiteCRM Database

"SUITECRM_PASSWORD" = The password used to connect to the SuiteCRM Database

These config variables define the connection to your SuiteCRM Data Integration to transform and load the extracted SuiteCRM data into:

```
# SuiteCRM Analytics Target Database Connection
SUITECRM_ANALYTICS_HOST=127.0.0.1
SUITECRM_ANALYTICS_PORT=3306
SUITECRM_ANALYTICS_DATABASE=suitecrm_dwh
SUITECRM_ANALYTICS_USERNAME=suitecrmrootuser
SUITECRM_ANALYTICS_PASSWORD=suitecrmrootuserpassword
```

Note: For this step, it is recommended to create a new MySQL database as the SuiteCRM Data Integration Data Warehouse

Update the above values in your "*install.properties*" file to reflect the following:

"SUITECRM_ANALYTICS_DATABASE" = The name of the Database to be used as the SuiteCRM Data Warehouse

"SUITECRM_ANALYTICS_USERNAME" = The username used to connect to the Database used as the SuiteCRM Data Warehouse

"SUITECRM_ANALYTICS_PASSWORD" = The password used to connect to the Database used as the SuiteCRM Data Warehouse

The SMTP config variables will define the connection to an email mail client to send out notifications on the status of the extraction script (success or failure).
These are optional:

```
# Enable SMTP email functionality
SMTP_ENABLED=0
# SMTP Server Details
SMTP_HOSTNAME=smtp.gmail.com
```

```
SMTP_PORT=587
SMTP_USE_AUTHENTICATION=1
SMTP_USERNAME=user@suitecrm-analytics.co.uk
SMTP_PASSWORD=password
#SMTP_SECURE_AUTHENTICATION=
#SMTP_SECURE_AUTHENTICATION=SSL
SMTP_SECURE_AUTHENTICATION=TLS
# Email address that SuiteCRM Analytics sends from
SMTP_FROM_EMAIL_ADDRESS=donotreply@suitecrm-analytics.co.uk
SMTP_FROM_EMAIL_NAME=SuiteCRM Analytics
# Email addresses seperated by spaces that error emails are delivered
to
SMTP_SEND_ERROR_EMAILS_TO=user@suitecrm-analytics.co.uk
# Set this to 1 or zero. Setting it to 1 means that you will get an
email for every sucessful run on the ETL
SMTP_SEND_SUCCESS_EMAILS=1
# Email addresses seperated by spaces that success emails are delivered
to
SMTP_SEND_SUCCESS_EMAILS_TO=user@suitecrm-analytics.co.uk
user123@suitecrm-analytics.co.uk
```

Configuring Timezone

The SuiteCRM Data Integration extraction logic is based on time and it uses the operating systems configured timezone to perform its processes.

Its because of this that the operating system timezone of our source SuiteCRM Instance should match that of the system hosting SuiteCRM Data Integration. We are not always able to edit the Operating system timezone itself but we can instead configure SuiteCRM Data Integration to use a custom timezone from the OS.

Lets assume our SuiteCRM Instance was hosted in Kiev but our SuiteCRM Data Integration instance was hosted somewhere in the UK. To align the timezones of the two systems we must perform the following steps.

- Find a valid timezone for your operating system. You can use the following command:
  - **timedatectl list-timezones**
- Navigate to the suitecrm-data-integration-server/client folder
- Open the spoon.sh in a text editor
- Add the following argument to the end of the existing $OPTS variable
  - **-Duser.timezone=Europe/Kiev**

For example the following line:

OPT="$OPT $PENTAHO_DI_JAVA_OPTIONS -Dhttps.protocols=TLSv1,TLSv1.1,TLSv1.2 -
Djava.library.path=$LIBPATH -DKETTLE_HOME=$KETTLE_HOME -
DKETTLE_REPOSITORY=$KETTLE_REPOSITORY -DKETTLE_USER=$KETTLE_USER -
DKETTLE_PASSWORD=$KETTLE_PASSWORD -
DKETTLE_PLUGIN_PACKAGES=$KETTLE_PLUGIN_PACKAGES -
DKETTLE_LOG_SIZE_LIMIT=$KETTLE_LOG_SIZE_LIMIT -
DKETTLE_JNDI_ROOT=$KETTLE_JNDI_ROOT"

becomes:

OPT="$OPT $PENTAHO_DI_JAVA_OPTIONS -Dhttps.protocols=TLSv1,TLSv1.1,TLSv1.2 -
Djava.library.path=$LIBPATH -DKETTLE_HOME=$KETTLE_HOME -
DKETTLE_REPOSITORY=$KETTLE_REPOSITORY -DKETTLE_USER=$KETTLE_USER -
DKETTLE_PASSWORD=$KETTLE_PASSWORD -
DKETTLE_PLUGIN_PACKAGES=$KETTLE_PLUGIN_PACKAGES -
DKETTLE_LOG_SIZE_LIMIT=$KETTLE_LOG_SIZE_LIMIT -
DKETTLE_JNDI_ROOT=$KETTLE_JNDI_ROOT -Duser.timezone=Europe/Kiev"

Installing SuiteCRM Data Integration

Once the configuration has been defined, now run the setup script in the root of your
'suitecrm-data-integration' directory, using the command:

```
./setup-suitecrm-data-integration.sh
```

This will check the connection of your databases and create the tables required to extract
your SuiteCRM data.

You will also be asked if you wish to install the **Demo Data**. Selecting **Y** will populate the
Data
Warehouse with Demo Data allowing you to test the solution. If you wish to run the
solution
against your live instance of SuiteCRM then type **N**

If you decide to extract data from your live instance of SuiteCRM you simply have to rerun
the **setup** script and select **N** when asked if you want to install the demo data. This will
delete the demo data and prepare a fresh Data Warehouse. You could also just execute the
run-suitecrm-data-integration.sh command and this will also wipe the demo data.

Then you can run the SuiteCRM Data Integration script to extract and transform your data:

```
./run-suitecrm-data-integration.sh
```

This should output on your terminal all logging referring to the extracting, transforming and loading data into the data warehouse tables.

Note: The above script will process any records that have been added/modified between the previous run and the current system time.
Due to this, it is recommended that both the server in which the Data Warehouse resides and the current system be on the same timezone setting, to allow the smoothest transition of data.

The solution currently extracts data from the following modules:

- Accounts
- Campaigns
- Cases
- Contacts
- Invoices
- Leads
- Opportunities
- Products/Product Categories
- Users
- Custom Fields from Cases, Leads

This is the end of the SuiteCRM Data Integration installation.

If you wish to know more about Pentaho ETL solutions you can visit the Spoon user guide.

This solution includes a pre-packaged spoon client that you can run via terminal:

```
./data-integration-client.sh
```