CARMAweb 1.3 installation guide

Johannes Rainer

October 18, 2007

Installation instructions

CARMAweb version 1.3.x with R 2.5 and Bioconductor 2.0

The archive `carmaweb-1.3.tar.gz` contains the following files:

- `R-Install`: contains required packages and scripts to install the Bioconductor packages needed for CARMAweb (see section).

- `bin`: contains simple bash scripts to start Rserve and CARMAweb.

- `jboss-4.0.2`: the application server that runs CARMAweb (jboss application server).

- `data`: sub directories where user data will be stored into (CARMAweb can also be configured to store data in other places.)

- `users.xml`: file that will be used for the usermanagement to store user information into.

This installation documentation describes the installation of CARMAweb on a linux (UNIX) system. It is assumed that *carma* is the user on the server that installs and runs CARMAweb (all path settings in the configuration files point to the users home directory */home/carma*, if CARMAweb is installed into another directory all paths have to be configured accordingly).

First of all unpack the CARMAweb archive (copy it to */home/carma* and unpack it `tar -xzf carmaweb-1.3.tar.gz`).

Requirements

Unfortunately Microsoft Windows is not supported yet as OS because some BioConductor packages are not available for Windows.

graphviz

`graphviz` is available for most Linux distributions as an rpm package (note that also the development package is needed!), otherwise it can be downloaded from http://www.graphviz.org/. The graphviz software is needed for the `Rgraphviz` BioConductor package that is used to draw the GO (GeneOntology) graphs.

postgresql

The postgresql libraries and headers have to be installed on the system where CARMAweb should be installed onto. For most linux distributions these are provided by rpm packages.

Installing R and Bioconductor

Installing R

Download R source code from the R home page (http://r-project.org), unpack the source code:

```
tar -xzf R-2.5.0.tar.gz
```

cd into the source directory (`cd R-2.5.0`) and configure the source code:

```
./configure --prefix=/home/carma/R-2.5 --enable-R-shlib
```

with this configuration R will be installed in the home directory of the *carma* user (*/home/carma/R-2.5*) and this makes the R and the Bioconductor libraries independent of other R installations on the server (since different users might want to use another R / Bioconductor version).

Now compile R

```
make
```

and finally install it (not required to be super user, since R is installed to the home directory of *carma*)

```
make install
```

Note: if you are updating an earlyer version of R/CARMAweb, it is suggested to unset the environment variables R_HOME and R_LIBS (i.e. by calling `unset R_HOME` and `unset R_LIBS`).

Installing Rserve

Note: the server needs internet access in order to download and install R and Bioconductor packages!

Start the freshly installed R:

```
/home/carma/R-2.5/bin/R
```

Note: if you encounter error during R startup, call R with the parameter `--vanilla` (i.e. `/home/carma/R-2.5/bin/R --vanilla`), which prevents to read any *old* configuration files (i.e. `.Rprofile` or `.Renviron` files).

install Rserve:

```
install.packages("Rserve",repos="http://cran.r-project.org")
```

copy `Rserve` from */home/carma/R-2.5/lib64/R/library/Rserve* to */home/carma/R-2.5/R/lib64/bin* (if your server is not a 64bit system exchange lib64 by lib).

```
cp /home/carma/R-2.5/lib64/R/library/Rserve/Rserve /home/carma/R-2.5/R/lib64/bin/
```

Installing Bioconductor

Cd into the *R-Install* directory in *carmaweb-1.3*

```
cd /home/carma/carmaweb-1.3/R-Install
```

start R (the newly installed R) in this directory

```
/home/carma/R-2.5/bin/R
```

load the small install script

```
source("install.R")
```

and start the installation procedure

```
install()
```

This install script installs required Bioconductor packages and some metadata and annotation packages. If you need other/more annotation packages install them using the `getBioC` function (for a full list of avaialble metadata packages check the Bioconductor homepage).

quit R

```
q("no")
```

copy the file *.Rprofile* from the *R-Install* directory to the users home directory:

```
cp .Rprofile /home/carma
```

this will prevent Rserve to open X11 windows for every graph that is produced.

For some Bioconductor pacakges it is on some systems required to make some further adjustements. For the *Rgraphviz* package it is recommended to set the environment variable `LD_LIBRARY_PATH` to point also to the `graphviz` installation. This can be done by adding the line

```
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/usr/lib/graphviz
```

(if graphviz was installed in `/usr/lib/graphviz`, on 64bit systems this could be `/usr/lib64/graphviz`) to the `.bash_profile` file in the users home directory (remember also to export this variable by adding the line `export LD_LIBRARY_PATH` in the file and to reload the profile (`source ~/.bash_profile`)).

It is important to note that there should be no errors during the installation of required Bioconductor packages! It is not suggested to go on with the installation of CARMAweb if not all required packages where installed correctly!

This has been successfully tested on a server running CentOS 4.4 (http://www.centos.org).

Info: since there have been reported errors that the `xtable` package could not be found in any repository (although it exists), i recommend installing the `xtable` and also the `maDB` package manually.

Again in the `/home/carma/carmaweb-1.3/R-Install` folder install the packages::

```
/home/carma/R-2.5/bin/R CMD INSTALL maDB_1.9.6.tar.gz
```

and

```
/home/carma/R-2.5/bin/R CMD INSTALL xtable_1.4-6.tar.gz
```

Java

CARMAweb requires Sun's Java Development Kit (JDK) $>= 1.4.2$ (version 1.5 is suggested, which can be downloaded from the page http://java.sun.com). The environment variable JAVA_HOME has to point to the JDK installation directory, JAVA_HOME/bin has to be added to the PATH environment variable. Note that jboss will not run properly with the gnu java versions, it is strongly recommended that a java version from sun is installed! Check which java version will be used by default after installation of the jdk (by typing java -version, or with which java and make sure that the java from sun is used).

Installation of CARMAweb

The web application CARMAweb and all other components (e.g. the cluster analysis module GenesisWeb) run within the application server jboss (http://www.jboss.org). CARMAweb is highly configurable, but this also means that a lot of settings and variables have to be adjusted for CARMAweb to run without errors.

The *bin* directory in the directory *carmaweb* contains two shell scripts to start the application server (and therefore the web application) and the Rserve service.

The file jboss_init.sh has to be edited to be able to start CARMAweb. The JBOSSDIR attribute in the jboss_init.sh file has to be set correctly to point to the bin directory of jboss (e.g. to */home/carma/carmaweb-1.3/jboss-4.0.2/bin* if the carmaweb archive has been extracted to the home directory of a user called *carma*). Another attribute that should be adjusted is the LOGFILE attribute that specifies where the application should write log messages into.

The startRserve.sh is used to start the Rserve service. The R_HOME attribute has to be adapted in this file (to point to the R installation directory (i.e. */home/carma/R-2.5/lib64/R* or */home/carma/R-2.5/lib/R*).

What ports are used?

The default ports that are already set for CARMAweb in the jboss application server are:

- 38080 to answer HTTP requests

- 31099 for the JNDI port

If you are using a firewall only the 38080 port has to be opened, as it is used to answer all incoming requests.

Configuration of CARMAweb

In this section we will go through all CARMAweb modules and see how and where these modules have to be configured. All modules are located within the directory jboss-4.0.2/server/Rweb/deploy, therefore all files described in this section are relative to this path.

CheckRServe.sar

This module controls that the `Rserve` is running properly and that it accepts connections. Therefore it tests periodically if `Rserve` is up and running. The configuration file for this service is *CheckRServe.sar/META-INF/jboss-service.xml* (located in the *deploy* directory of the jboss application server). This file has to be configured by specifying where the shell script file to start `Rserve` is located (edit the *StartCall* attribute to point to the correct location (by default it assumes that the shell script to start `Rserve` is located at */home/carma/carmaweb-1.3/bin/startRserve.sh*)).

ClusterService.sar

This service allows to perform different ways of load balancing by assigning the various jobs to different computers in a cluster. The configuration file for this service is located at *ClusterService.sar/META-INF/jboss-service.xml*. For simple (not load balancing purpose) the only parameter that has to be set correctly is the *PathToSavedFiles* attribute. Set this correctly to a directory where temporary cluster analysis results should be stored into
(the default setting is */home/carma/carmaweb-1.3/data/genesisweb/SavedJOBS*).

RServer.ear

This module is used to establish asynchroneous connections between the CARMAweb web application and the `Rserve` service. The configuration file to configure this service is *RServer.ear/RServer.jar/RServe.properties*. The parameters that have to be adjusted are:

- `workspace.path`: used to specify the directory where the workspaces for SOAP calls should be stored to.

- `R.path`: the directory where R has been installed to (*/home/carma/R-2.5/lib/R* or */home/carma/R-2.5/lib64/R*).

- `maxtime`: SOAP workspaces older than the time specified (in milliseconds) will be deleted automatically. Defaults to 1 week (*maxtime=302400000*).

carma.war

The CARMAweb web application. This module has to be configured via the *carma.war/WEB-INF/web.xml* file. The parameters that have to be adjusted in this file are:

- `CONFIG_FILE_PATH`: has to point to the R installation directory (e.g. */home/carma/R-2.5/lib/R* or */home/carma/R-2.5/lib64/R*)

- `PATH_TO_R`: same as *CONFIG_FILE_PATH*

- `USER_DIRECTORY`: the directory where all user data should be stored into (workspaces, data...). Defaults to */home/carma/carmaweb-1.3/data/carma*.

To enable the SOAP interface to allow also other programs to use features of CARMAweb the *carma.war/WEB-INF/server-config.wsdd* file has to be edited. The `attachments.Directory` has to be adjusted if carmaweb was not unpacked to */home/carma*. Actually just the first part (*/home/carma/*) of the default value has to be replaced by the directory where the carmaweb archive has been unpacked. If this parameter is not configured correctly, SOAP clients can not upload files to the CARMAweb web application.

genesis.war

The GenesisWeb web application. Again the *web.xml* file located in the directory *genesis.war/WEB-INF* has to be adjusted to configure this module. The only parameter to edit is the `PATH_TO_USER_DIR` attribute that has to point to the directory where the genesisweb user data should be stored into.

Configuring the usermanagement

CARMAweb comes with its own simple usermanagement.To configure it the files *bugblatterbeast-usermanagement-utils.properties* in the directories *carma.war/WEB-INF/classes* and *genesis.war/WEB-INF/classes* have to be renamed into $< HOSTNAME >$-*usermanagement-utils.properties*, where $< HOSTNAME >$ is the hostname of the server where CARMAweb has been installed into (just type *hostname* in a console to get the machines hostname). Additionally in both files the parameter `sum_users_filepath` has to be adjusted to point to the *users.xml* file that is usually located in the *carmaweb* directory.

Starting CARMAweb

To start CARMAweb (again assumed you are logged in as the correct user) change to the directory *carmaweb/bin* and type

```
./jboss_init.sh start
```

This should start the application server (and therefore also CARMAweb). Have a look at the log file (can be specified in). If there are no `ERROR` messages everything started successfully. It is not needed to start also `Rserve` manually, because it is started automatically during the CARMAweb startup. The startup of CARMAweb depends on the machine where it is installed, but usually it takes about 1 to 2 minutes.

To see if CARMAweb runs properly open a web browser and go to the url *http://localhost:38080/carma* and create your first account using the *Create account* link.

To stop CARMAweb type `./jboss_init.sh stop.`