

D4Science Infrastructure - Task #4371

Upgrade # 3868 (Completed): D4Science Infrastructure Upgrade to gCube 3.11.0

Upgrade # 3873 (Completed): D4Science Gateways upgrade to gCube 3.11.0

Task # 2516 (Closed): Provide RStudio in the e-Infrastructure VREs

same R configuration (R version and packages) on R-Studio as the one on machine(s) used by Dataminer.

Jun 23, 2016 11:05 AM - Pasquale Pagano

Status:	Closed	Start date:	Jun 23, 2016
Priority:	Normal	Due date:	Jul 23, 2016
Assignee:	Andrea Dell'Amico	% Done:	100%
Category:	Application	Estimated time:	0.00 hour
Target version:	Data Processing		
Infrastructure:	Development, Pre-Production, Production		
Description We should analyze and deliver a solution to automatize the configuration of the Dataminer instances. The idea is that a user configure RStudio and all packages he installs are then automatically installed on the Dataminer instances available in his VRE.			
Related issues:			
Related to D4Science Infrastructure - Task #4865: Installing data analysis ma...		Closed	Aug 22, 2016
Related to D4Science Infrastructure - Task #5448: Dataminer - Need to align R...		Closed	Sep 30, 2016 Nov 04, 2016

History

#1 - Sep 08, 2016 11:38 AM - Andrea Dell'Amico

- Related to Task #4865: Installing data analysis machine to develop prototypes added

#2 - Oct 08, 2016 01:02 PM - Andrea Dell'Amico

The new proposal is the following:

- Three lists are maintained into a subversion repository. One for the debian packages, one for the R packages installed from CRAN or other sites that publish source tarballs, one for the R packages installed from github
- A cron job is installed on all the R enabled servers (dataminer, rstudio). The cron job runs frequently (at 10 minutes intervals?), downloads the files and installs the packages that aren't already present. It also runs once daily to reinstall the packages that have new versions available.

#3 - Oct 08, 2016 01:07 PM - Andrea Dell'Amico

- Related to Task #5448: Dataminer - Need to align R configuration with RStudio simulation environment added

#4 - Oct 18, 2016 02:42 PM - Andrea Dell'Amico

- Status changed from New to In Progress

- % Done changed from 0 to 80

- Infrastructure Development, Pre-Production added

The script and the cron jobs are ready.

We now only need a repository where to store the three lists.

#5 - Oct 21, 2016 02:41 PM - Andrea Dell'Amico

- File `r_github_pkgs.txt` added

- File `r_cran_pkgs.txt` added

- File `r_deb_pkgs.txt` added

I'm attaching the files lists. @gianpaolo.coro@isti.cnr.it would you setup a subversion repository for them?

#6 - Oct 21, 2016 02:51 PM - Gianpaolo Coro

Here is the SVN folder:

<https://svn.d4science.research-infrastructures.eu/gcube/trunk/data-analysis/RConfiguration/RPackagesManagement>

Public link:

<https://svn.research-infrastructures.eu/public/d4science/gcube/trunk/data-analysis/RConfiguration/RPackagesManagement/>

#7 - Oct 24, 2016 03:41 PM - Andrea Dell'Amico

- Status changed from *In Progress* to *Feedback*

- % Done changed from 80 to 100

The lists are now got from subversion. script and the cron jobs work as proposed in [#note-2](#):

- One cron job runs at 10 minutes interval and installs new packages if any
- One cron jobs runs every night at 3am and updates all the listed packages.

Note that a complete reinstall needs slightly less than three hours.

The script relies on locking to avoid parallel runs.

Installed on dataminer1-d-d4s.d4science.org and dataminer1-proto.d4science.org right now.

#8 - Oct 24, 2016 07:28 PM - Gianpaolo Coro

Could you please install it on dataminer1-devnext?

What happens if one of the packages has typo? The complete installation should not stop, I guess.

Is the latest version always installed, by default configuration?

I have added one package on the SVN files. I will go to see if the package has been installed.

#9 - Oct 24, 2016 07:47 PM - Andrea Dell'Amico

Gianpaolo Coro wrote:

Could you please install it on dataminer1-devnext?

Done.

What happens if one of the packages has typo? The complete installation should not stop, I guess.

It doesn't stop.

Is the latest version always installed, by default configuration?

If a package is not already present, yes.

Otherwise it will be updated by the daily cron job.

#10 - Oct 28, 2016 09:58 AM - Andrea Dell'Amico

- Status changed from *Feedback* to *Closed*

Script and cron jobs have been installed on all the dataminer servers.

Files

r_github_pkgs.txt	294 Bytes	Oct 21, 2016	Andrea Dell'Amico
r_cran_pkgs.txt	946 Bytes	Oct 21, 2016	Andrea Dell'Amico
r_deb_pkgs.txt	369 Bytes	Oct 21, 2016	Andrea Dell'Amico