bash script to rip/download rpms from a repo

there are many ways hwo to clone a rpm repo. most notably, one can use reposync which does exactly that, or if the repository is available via rsync, rsync —list-only could be used to find the right directory to download via rsync.

another option is using wget for http repos which show a directory listing.

my script here is another (quick and dirty) method with only a few dependencies, ment to work also on linux systems that don't have yum utilities and it is able to download repos from http even without a directory listing, by parsing the repodata necessary to get the list of files. as a little addition, my script downloads the rpms by parallelizing 10 wget threads using xargs, which can significantly speed up the download process.

CAUTION this is a quick and dirty scirpt, i'm not doing much as far as input checking goes, so make sure you enter the right parameters! use this at your own risk.

repofetch.sh

```
#!/bin/bash
set -e
targetDir=$1
url=$2
if [ -z "$targetDir" -o -z "$url" ]; then
    echo
    echo "repofetch.sh downloads all rpm's from a rpm repostiroy and
creates a local repo out of it"
    echo "
            usage: repofetch.sh <target dir> <repo url>"
    echo
    echo " <target dir>: target directory on local filesystem where
the rpms should go (repodata will be one directory higher)"
    echo " <repo url>: url of the repo including the path to the
directory where the repodata directory is located"
    echo
    exit 1
echo "creating target directory $targetDir"
mkdir -p $targetDir
cd $targetDir
echo "fetch repomd.xml"
wget -nd $url/repodata/repomd.xml
echo "fetch primary.xml"
primaryxml=$(grep -A 20 '<data type="primary">' repomd.xml | grep
"<location href=" | head -1 | sed -e
's/^.*href="repodata\/\([^"]*\)".*$/\1/')
wget -nd "$url/repodata/$primaryxml"
```

```
echo "remove all pre-existing rpm's and download all rpms found in
primary.xml"
rm -f *.rpm
files=$(zgrep "<location href" $primaryxml | sed -e</pre>
's/^.*href="\([^"]*\)".*$/\1/')
echo -n -e "$files\0" | tr \\n \\0 | xargs -0 -I DDD -P 10 wget -nd
"$url/DDD"
echo "clean up"
rm -f $primaryxml repomd.xml
cd ../
createrepo.
echo "## DONE ##"
echo
echo "content template for repo file:"
echo
echo "[repo-name]
name=Long repo description - \$basearch
baseurl=file://$targetDir
enabled=1
gpgcheck=0"
echo
```

From:

http://wiki.psuter.ch/ - pswiki

Permanent link:

http://wiki.psuter.ch/doku.php?id=bash script to rip download rpms from a repo





http://wiki.psuter.ch/ Printed on 06.11.2025 06:12