

# INSTALLATION OF OPEN SOURCE R

## ON THE

### IBM PUREDATA FOR ANALYTICS APPLIANCE

This document describes how you can compile and install Open Source R and some prerequisite tools on the IBM PureData for Analytics appliance by using the *installOpenSourceR.sh* shell script.

It is recommended to first install Open Source R on a Netezza emulator (the directions and code are designed for this approach). The compiled code can be easily transferred to other compatible appliances or emulators, so typically, it does not matter where the source code compilation is done. You can download an emulator from the IBM developerWorks *Netezza Developer Network* (NDN) community. You need to register first at developerWorks (<https://www.ibm.com/developerWorks>). Search for “NDN” to locate the Netezza Developer Network community. Follow the instructions in the overview page to get access to the secure part of the community. The secure part contains a bookmark “NZ Emulator Installation” with further instructions.

Before you can use the *installOpenSourceR.sh* script, you must install IBM Netezza Analytics Version 2.5.x or Version 3.0.x. The installation of Open Source R needs several tools that are installed with Netezza Analytics. For each version of Netezza Analytics, a corresponding R version is recommended as follows. Other combinations might be possible, but have not been tested.

- Netezza Analytics 2.5.0 to 2.5.3: R 2.14.x
- Netezza Analytics 2.5.4 to 3.0.0: R 2.15.x
- Netezza Analytics 3.0.1 to 3.0.2: R 3.0.x

**Note:** You need the root password for the SPUs because some code must be compiled on a SPU. The root password for the host is not needed. *All operations on the host are performed by user nz, including the execution of installOpenSourceR.sh.*

Several Linux tools must be installed before you can install Open Source R. All these tools are installed by the script automatically. You just need to download all these source code archives from the corresponding websites. The following table lists the tools and recommended URLs. Store all the downloaded files and the *installOpenSourceR.sh* script in the same directory on the Netezza host.

If the script cannot find one of the tools or OpenSource R, it tries to download the source code itself, using the URLs from the table below. For bash, it also downloads all patches it can find, to avoid bash security problems that were recently found (“shellshock”).

Source Code	Example URL
termcap-1.3.1.tar.gz	<a href="ftp://ftp.gnu.org/pub/gnu/termcap/termcap-1.3.1.tar.gz">ftp://ftp.gnu.org/pub/gnu/termcap/termcap-1.3.1.tar.gz</a>
ncurses-5.9.tar.gz	<a href="ftp://ftp.gnu.org/pub/gnu/ncurses/ncurses-5.9.tar.gz">ftp://ftp.gnu.org/pub/gnu/ncurses/ncurses-5.9.tar.gz</a>

readline-6.2.tar.gz	<a href="ftp://ftp.cwru.edu/pub/bash/readline-6.2.tar.gz">ftp://ftp.cwru.edu/pub/bash/readline-6.2.tar.gz</a>
libiconv-1.14.tar.gz	<a href="ftp://ftp.gnu.org/pub/gnu/libiconv/libiconv-1.14.tar.gz">ftp://ftp.gnu.org/pub/gnu/libiconv/libiconv-1.14.tar.gz</a>
make-3.82.tar.gz	<a href="ftp://ftp.gnu.org/pub/gnu/make/make-3.82.tar.gz">ftp://ftp.gnu.org/pub/gnu/make/make-3.82.tar.gz</a>
bash-4.3.tar.gz	<a href="ftp://ftp.gnu.org/gnu/bash/bash-4.3.tar.gz">ftp://ftp.gnu.org/gnu/bash/bash-4.3.tar.gz</a>
coreutils-8.13.tar.gz	<a href="ftp://ftp.gnu.org/gnu/coreutils/coreutils-8.13.tar.gz">ftp://ftp.gnu.org/gnu/coreutils/coreutils-8.13.tar.gz</a>
findutils-4.4.2.tar.gz	<a href="ftp://ftp.gnu.org/gnu/findutils/findutils-4.4.2.tar.gz">ftp://ftp.gnu.org/gnu/findutils/findutils-4.4.2.tar.gz</a>
grep-2.9.tar.gz	<a href="ftp://ftp.gnu.org/gnu/grep/grep-2.9.tar.gz">ftp://ftp.gnu.org/gnu/grep/grep-2.9.tar.gz</a>
sed-4.2.tar.gz	<a href="ftp://ftp.gnu.org/gnu/sed/sed-4.2.tar.gz">ftp://ftp.gnu.org/gnu/sed/sed-4.2.tar.gz</a>
tar-1.26.tar.gz	<a href="ftp://ftp.gnu.org/gnu/tar/tar-1.26.tar.gz">ftp://ftp.gnu.org/gnu/tar/tar-1.26.tar.gz</a>
libpthread-stubs-0.1.tar.bz2	<a href="http://xcb.freedesktop.org/dist/libpthread-stubs-0.1.tar.bz2">http://xcb.freedesktop.org/dist/libpthread-stubs-0.1.tar.bz2</a>
pkg-config-0.28.tar.gz	<a href="http://pkgconfig.freedesktop.org/releases/pkg-config-0.28.tar.gz">http://pkgconfig.freedesktop.org/releases/pkg-config-0.28.tar.gz</a>
R-2.14.2.tar.gz or R-2.15.3.tar.gz or R-3.0.2.tar.gz	<a href="http://lib.stat.cmu.edu/R/CRAN/src/base/R-2/R-2.14.2.tar.gz">http://lib.stat.cmu.edu/R/CRAN/src/base/R-2/R-2.14.2.tar.gz</a> <a href="http://lib.stat.cmu.edu/R/CRAN/src/base/R-2/R-2.15.3.tar.gz">http://lib.stat.cmu.edu/R/CRAN/src/base/R-2/R-2.15.3.tar.gz</a> <a href="http://lib.stat.cmu.edu/R/CRAN/src/base/R-3/R-3.0.2.tar.gz">http://lib.stat.cmu.edu/R/CRAN/src/base/R-3/R-3.0.2.tar.gz</a>

To see the available options, you can first run the script with the --help option:

```
./installOpenSourceR.sh --help
```

Typically, the only option you need is the option to specify the SPU root password:

```
./installOpenSourceR.sh -p <spu-root-password>
```

Depending on the performance of your emulator machine, the whole compilation and installation process can take more than an hour.

To uninstall Open Source R, call the script by entering the following command:

```
./installOpenSourceR.sh -u
```

## TRANSFER OF THE BINARIES TO ANOTHER NETEZZA SYSTEM

After the binaries are compiled and installed on one machine, you can install the compiled binaries also on another Netezza system, hereafter called *target machine*.

### Prerequisite:

Netezza Analytics must be installed on the target machine.

To install the binaries on the target machine, do the following steps:

1. Create the following directory on the target machine. The directory name must be identical to the directory name on the source machine.  
/nz/extensions/nz/r\_install

2. From the `/nz/extensions/nz/r_install` directory on the source machine, copy the following files to the newly created directory on the target machine:
  - a) `sysroot_overlay.tar.gz`
  - b) `r_<version>_overlay.tar.gz`

Copy also this script *installOpenSourceR.sh* to this directory.

3. Call the *installOpenSourceR.sh* script by entering the following command. The SPU password is not required in this case:

```
./installOpenSourceR.sh -i -b
```

To uninstall Open Source R, call the script by entering the following command:

```
./installOpenSourceR.sh -u
```

## TESTING THE R INSTALLATION

An extensive test of the R installation should be done from an R client that has been extended with the Netezza Analytics R client packages (`nzr`, `nza`, `nzmatrix`). The *demo* functions of these packages also call R functions on the server. The installation and usage of the Netezza Analytics R client packages is described in the client package documentation.

**Copyright and trademark information**

IBM, the IBM logo and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).