

R sur OpenBSD

<https://buskvekster.moood.com/r-sur-openbsd.pdf>

Andre BUSKVEKSTER

andre@buskvekster.moood.com
Omega Verksted

Rencontres R 2023

Pourquoi

Thèmes des posters et présentations Rencontres R 2023

- Shiny
- Bases de données
- Administration
- Gestion de versions

Pourquoi

Avantages d'OpenBSD

Administration facile

Pourquoi

Inconvénients d'OpenBSD

Certains packages R ne fonctionnent pas.

1 Pourquoi

2 Installation

- Difficultés
- Patches
- Ultime recours : vmm et Rserve

3 Fonctionnalités uniques

- Documentation
- rc.subr
- pf

Installation

- Système de base : `pkg_add R`
- RStudio : `pkg_add rstudio`
- Autres : `install.packages`

Installation

Difficultés

- Surtout de problèmes de compilation (C, C++, fortran, etc.)
- Exemples depuis le tuto statspatial
 - `_tthread_timespec_get`
 - `gfortran`

Installation

Porter les packages

- Je lis les messages d'erreur
- J'essaie à trouver les bugs
- Exemples depuis le tuto statspatial
 - `_tthread_timespec_get`
(<https://github.com/r-lib/later/issues/115>)
 - `ln -sf /usr/local/bin/egfortran
/usr/local/bin/gfortran`
- Si je peux pas, j'écris au développeur ou à Ingo Feinerer
- La meilleur option : Je fais un patch moi-même

Installation

Comment contribuer

- ❶ J'essaye d'installer un package.
- ❷ Je vois un bug.
- ❸ Je corrige le package.
- ❹ J'envoie les changements, **mais pas avec GitHub**

Installation

Le cultu_e propriétai_e de certains packages

- GitHub
- Slack
- autres logiciels propriétaires

Installation

Le cultuRe libéRé d'R

- GNU R
- Critères éthiques de GNU concernant l'hébergement de logiciel
(<https://www.gnu.org/software/repo-criteria.fr.html>)
- Patches à R-base : Mail/Bugzilla
(<https://www.r-project.org/bugs.html>)

Installation

Des droits de la femme et de l'homme

- Exemple : impossible pour moi de faire un patch pour un package d'RStudio (Posit)
- Un choix
 - Garder vos droits ou les jeter
 - Liberté ou l'esclavage

Installation

Des droits de la femme et de l'homme

- Exemple : impossible pour moi de faire un patch pour un package d'RStudio (Posit)
- Un choix
 - Garder vos droits ou les jeter
 - Liberté ou l'esclavage

Installation

Ultime recours : Ne pas porter les packages

- Ne pas utiliser ce package (Essayer l'année prochaine)
- Installer GNU/Linux dans un machine virtuel avec communication par Rserve

Installation

Ultime recours : Ne pas porter les packages

- Ne pas utiliser ce package (Essayer l'année prochaine)
- Installer GNU/Linux dans un machine virtuel avec communication par Rserve

vmm

Machines virtuels

- 1 Suivre le FAQ OpenBSD
(<https://www.openbsd.org/faq/faq16.html>)
- 2 Installer Alpine Linux dans le machine virtuel
- 3 Installer R dans Alpine
- 4 Installer Rserve dans Alpine, RScient dans OpenBSD
- 5 `vm.conf` et `rcctl enable vmd` (facultatif)


```
# Basé sur https://gist.github.com/voutilad  
vmctl create -s 6G rserve.qcow2  
vmctl start -d alpine-virt-3.6.0-x86_64.iso \  
    -d rserve.qcow2 -n local -c rserve  
setup-alpine
```

```
# vm.conf
vm "rserve" {
    memory 4G
    disk "/foo/bar/rserve.qcow2" format "qcow2"
    local interface
}
```

```
# Serveur non-OpenBSD  
library(Rserve)  
Rserve()  
  
# Client OpenBSD  
library(RSclient)  
c <- RSConnect(host)
```

1 Pourquoi

2 Installation

- Difficultés
- Patches
- Ultime recours : vmm et Rserve

3 Fonctionnalités uniques

- Documentation
- rc.subr
- pf

Outils OpenBSD

- Machines virtuels (vmm)
- Documentation
- Exécution d'applications en arrière-plan (rc.subr)
- Pare-feu (pf)
- Proxy d'applications Shiny (relayd)
- DevOps (autoinstall)
- Gestion de la configurations (fichiers dans /etc)
- Mises à jour (syspatch, sysupgrade)

Outils OpenBSD

- Machines virtuels (vmm)
- Documentation
- Exécution d'applications en arrière-plan (rc.subr)
- Pare-feu (pf)
- Proxy d'applications Shiny (relayd)
- DevOps (autoinstall)
- Gestion de la configurations (fichiers dans /etc)
- Mises à jour (syspatch, sysupgrade)

Outils OpenBSD

- Machines virtuels (vmm)
- Documentation
- Exécution d'applications en arrière-plan (rc.subr)
- Pare-feu (pf)
- Proxy d'applications Shiny (relayd)
- DevOps (autoinstall)
- Gestion de la configurations (fichiers dans /etc)
- Mises à jour (syspatch, sysupgrade)

Outils OpenBSD

- Machines virtuels (vmm)
- Documentation
- Exécution d'applications en arrière-plan (rc.subr)
- Pare-feu (pf)
- Proxy d'applications Shiny (relayd)
- DevOps (autoinstall)
- Gestion de la configurations (fichiers dans /etc)
- Mises à jour (syspatch, sysupgrade)

Documentation

- <https://www.openbsd.org/faq/>
- `man man`, etc.
 - Ça existe sur GNU/Linux aussi
 - Mais la qualité est beaucoup élevée sur OpenBSD

rc.subr

Installer RStudio comme service

- Exécuter le program en arrière-plan
- `rcctl start rstudio, rcctl check rstudio, &c.`
- La même idée que `systemctl start rstudio, systemctl status rstudio, &c.`

```
pkg_add rstudio  
install -o utilisateur -d \  
    /var/run/rstudio-server rstudio-server
```

rc.subr

Un service d'RStudio

le fichier `/etc/rc.d/rstudio`

```
#!/bin/ksh
daemon="/usr/local/lib/rstudio/bin/rserver"
daemon_flags="--www-port 8787"
daemon_user=moi
. /etc/rc.d/rc.subr
rc_reload=NO
rc_cmd $1
```

```
#!/bin/ksh
```

```
. /etc/rc.d/rc.subr  
rc_reload=NO  
rc_cmd $1
```

```
daemon="/usr/local/lib/rstudio/bin/rserver"  
daemon_flags="--www-port 8787"  
daemon_user=moi
```

```
#!/bin/ksh
daemon="/usr/local/lib/rstudio/bin/rserver"
daemon_flags="--www-port 8787"
daemon_user=moi
. /etc/rc.d/rc.subr
rc_reload=NO
rc_cmd $1
```


pf

Règles pare-feu pour R (/etc/pf.conf)

- Interdire l'accès à l'internet et au réseau local
- Permettre l'accès à CRAN
- Permettre l'accès à une base de données PostgreSQL
- Permettre le serveur RStudio

```
table <cran> { cran.biotoools.fr }  
table <db> { 192.168.98.23 }  
utilisateurR = "moi"  
port_rstudio = "8787"  
port_postgresql = "5432"  
  
block return  
pass
```

```
# interdire les serveurs
block return in proto { tcp udp } \
  to self user $utilisateurR

# interdire les clients
block return out log proto { tcp udp } \
  from self user $utilisateurR
```

```
# permettre les client DNS, CRAN, et PostgreSQL
# et le serveur RStudio
pass out on egress proto udp from self \
  to any port domain user $utilisateurR
pass out proto tcp from self to <cran> \
  port { http https } user $utilisateurR
pass out proto tcp from self to <db> port $port_postgresql
pass in proto tcp from self to self \
  port $port_rstudio user $utilisateurR
```

Résumé

- Avantage l'OpenBSD : l'administration facile
 - Qualité de documentation
 - Culture
 - Architecture : Grand système de base
- Inconvénients : Certain packages ne fonctionnent pas
- Solutions
 - Attendre, utiliser autres packages
 - Faire des patches
 - Communiquer par Rserve avec un système GNU/Linux

Résumé

- Avantage l'OpenBSD : l'administration facile
 - Qualité de documentation
 - Culture
 - Architecture : Grand système de base
- Inconvénients : Certain packages ne fonctionnent pas
- Solutions
 - Attendre, utiliser autres packages
 - Faire des patches
 - Communiquer par Rserve avec un système GNU/Linux

Résumé

- Avantage l'OpenBSD : l'administration facile
 - Qualité de documentation
 - Culture
 - Architecture : Grand système de base
- Inconvénients : Certain packages ne fonctionnent pas
- Solutions
 - Attendre, utiliser autres packages
 - Faire des patches
 - Communiquer par Rserve avec un système GNU/Linux

Attention

- Des fois l'OpenBSD fonctionne trop bien.
- Vos collègues vont voir que vous ne travaillez pas.

Merci

<https://buskvekster.mooo.com/r-sur-openbsd.pdf>

Bonbons scandinaves

- Pour ceux que posent des questions
- Chocolat et massepain
- Trois goûts
 - La prune dans le Madère
 - Le bleuet dans le vodka
 - La fraise dans le vin

Merci

<https://buskvekster.mooo.com/r-sur-openbsd.pdf>

Pourquoi R sur OpenBSD

- Pourquoi les logiciels libres
- Pourquoi l'R
- Pourquoi l'OpenBSD

Pourquoi la liberté

Qu'est-ce que la liberté

- la liberté de faire fonctionner le programme comme vous voulez, pour n'importe quel usage (liberté 0) ;
- liberté d'étudier le fonctionnement du programme, et de le modifier pour qu'il effectue vos tâches informatiques comme vous le souhaitez (liberté 1) ; l'accès au code source est une condition nécessaire ;
- la liberté de redistribuer des copies, donc d'aider les autres (liberté 2) ;
- la liberté de distribuer aux autres des copies de vos versions modifiées (liberté 3) ; en faisant cela, vous donnez à toute la communauté une possibilité de profiter de vos changements ; l'accès au code source est une condition nécessaire.

Pourquoi la liberté

- Pour éliminer l'esclavage
- Pour respecter les droits de la femme et de l'homme
- Parce que les logiciels libres fonctionnent mieux
- Pour le support à long terme

Pourquoi la liberté

- Pour éliminer l'esclavage
- Pour respecter les droits de la femme et de l'homme
- Parce que les logiciels libres fonctionnent mieux
- Pour le support à long terme

Pourquoi la liberté

- Pour éliminer l'esclavage
- Pour respecter les droits de la femme et de l'homme
- Parce que les logiciels libres fonctionnent mieux
- Pour le support à long terme

Pourquoi la liberté

Exemples

- R (et pas S)
- OpenBSD
- Certaines distributions GNU/Linux
- Le système réseau de Windows (NetBSD)
- Certains composants de Mac OS X (FreeBSD)
- OpenSSH (OpenBSD)

Pourquoi la liberté

Exemples

- R (et pas S)
- OpenBSD
- Certaines distributions GNU/Linux
- Le système réseau de Windows (NetBSD)
- Certains composants de Mac OS X (FreeBSD)
- OpenSSH (OpenBSD)

Pourquoi l'R

- Portabilité (du système base au moins)
- Standardisation (parmi les statisticiens)
- Exactitude (en ce qui concerne le statistique)
- Documentation
-

Pourquoi l'R

- Portabilité (du système base au moins)
- Standardisation (parmi les statisticiens)
- Exactitude (en ce qui concerne le statistique)
- Documentation
-

Pourquoi l'OpenBSD

- Portabilité
- Standardisation (protocoles informatiques)
- Exactitude
- Documentation
- Sécurité proactive et cryptographie intégrée

Pourquoi l'OpenBSD

- Portabilité
- Standardisation (protocoles informatiques)
- Exactitude
- Documentation
- Sécurité proactive et cryptographie intégrée

Pourquoi l'OpenBSD

Sécurité proactive

- Exactitude d'architecture
- Recherche
- Installation avec de bonnes valeurs par défaut (security by default)
- Configuration simple

Pourquoi l'OpenBSD

Exemples de sécurité proactive

- W^X (écrire OU-exclusif exécuter) : Si un programme peut écrire dans une zone mémoire des données, cette zone de données ne peut pas être exécuté
- pledge : Facilement limiter les droits d'un programme
- unveil : Limiter les répertoires qu'un programme peut voir

Non-portabilité

En bref

- De différents systèmes
- Beaucoup d'utilisateurs des systèmes GNU/Linux.
- Peu d'utilisateurs d'OpenBSD.
- **Mais** l'OpenBSD est plus simple et les problèmes sont donc plus faciles à résoudre.

Compilation

- Un package écrit uniquement dans R devrait fonctionner.
- Incompatibilité
 - Plusieurs développeurs ne respectent pas les standards
 - Des fois il n'y a pas de standard
- Surtout de problèmes de compilation (C, C++, fortran, etc.)

Culture de logiciels propriétaires

- En principe
 - Liberté ou l'esclavage
 - Respect des droits de la femme et de l'homme, ou pas
- En pratique
 - Système d'exploitation : *BSD et GNU *ou* Mac et Windows
 - Messagerie instantanée : IRC *ou* Slack
 - Forge : Mail et cvs/svn (OpenBSD et R) *ou* GitHub (tidyverse, Posit)
- Exemple : impossible pour moi de faire un patch pour un package d'RStudio (Posit)

Culture de logiciels propriétaires

- En principe
 - Liberté ou l'esclavage
 - Respect des droits de la femme et de l'homme, ou pas
- En pratique
 - Système d'exploitation : *BSD et GNU *ou* Mac et Windows
 - Messagerie instantanée : IRC *ou* Slack
 - Forge : Mail et cvs/svn (OpenBSD et R) *ou* GitHub (tidyverse, Posit)
- Exemple : impossible pour moi de faire un patch pour un package d'RStudio (Posit)

Culture de logiciels propriétaires

- En principe
 - Liberté ou l'esclavage
 - Respect des droits de la femme et de l'homme, ou pas
- En pratique
 - Système d'exploitation : *BSD et GNU *ou* Mac et Windows
 - Messagerie instantanée : IRC *ou* Slack
 - Forge : Mail et cvs/svn (OpenBSD et R) *ou* GitHub (tidyverse, Posit)
- Exemple : impossible pour moi de faire un patch pour un package d'RStudio (Posit)

Non-portabilité

Ultime recours : Ne pas porter les packages

- Ou attendre quelques années (Le tidyverse peut enfin s'installer!)
- Minimiser les packages que vous utilisez
- Ou installer GNU/Linux dans un machine virtuel avec communication par Rserve

Non-portabilité

Ultime recours : Ne pas porter les packages

- Ou attendre quelques années (Le tidyverse peut enfin s'installer!)
- Minimiser les packages que vous utilisez
- Ou installer GNU/Linux dans un machine virtuel avec communication par Rserve

Non-portabilité

Ultime recours : Ne pas porter les packages

- Ou attendre quelques années (Le tidyverse peut enfin s'installer!)
- Minimiser les packages que vous utilisez
- Ou installer GNU/Linux dans un machine virtuel avec communication par Rserve


```
# ~/.ssh/config ou /etc/ssh/ssh_config (OpenBSD)
```

```
ForwardX11 yes
```

```
# /etc/ssh/ssh_config (autre système)
```

```
X11Forwarding yes
```

Non-portabilité

Solution : ssh sans X

- Exécuter : `ssh machine.virtuel.example.com R`
- Copier les resultats : `scp machine.virtuel.example.com resultat.png .`

Non-portabilité

Solution : ssh avec X

Les mises à jour

- Deux fois par an
- Il fonctionnent très bien : sysupgrade

- ❶ La même idée mais
- ❷ le “daemon” sera Rscript
- ❸ le “daemon_flags” comprendra le fichier R.
 - Shiny
 - Rserve

Facile à gérer

- Qualité de la documentation
- Grand système de base
 - Comme R
 - Pas comme GNU/Linux et le tidyverse
- Documentation claire
- Mises à jour faciles

rc.subr

Comparaison avec systemctl

- Langage
 - rc.subr (ksh) : plus standardisé, plus flexible
 - systemd (systemd unit) : plus simple
- Options (variables)
 - rc.subr : moins d'options, plus facile à apprendre
 - systemd : plus d'options, plus flexible mais plus facile de faire une faute
- Documentation
 - rc.subr : facile à trouver (`man rc.subr`) et courte (977 mots)
 - systemd : difficile à trouver
(<https://www.freedesktop.org/software/systemd/man/systemd>)
et longue (13022 mots)

[Unit]

Description=RStudio Server

After=network-online.target

Wants=network-online.target

[Service]

Type=forking

PIDFile=/var/run/rstudio-server.pid

ExecStart=/usr/local/lib/rstudio/bin/rserver

ExecStop=/usr/bin/killall -TERM rserver

KillMode=none

Restart=on-failure

[Install]

WantedBy=multi-user.target

rc.subr

Résumé

- 1 Lire le mode d'emploi `man rc.subr`
- 2 Créer le fichier `/etc/rc.d/rstudio`
- 3 Exécuter `rcctl enable rstudio`
- 4 Pour sauvegarder : `/etc/rc.conf.local`

```
block return  
pass
```

```
utilisateur = "moi"

block return
pass
block return in proto { tcp udp } \
  to self user $utilisateur # interdire les serveurs
block return out log proto { tcp udp } \
  from self user $utilisateur # interdire les clients
```

```
utilisateurR = "moi"
port_rstudio = "8787"

block return
pass
block return in proto { tcp udp } \
  to self user $utilisateurR # interdire les serveurs
block return out log proto { tcp udp } \
  from self user $utilisateurR # interdire les clients

# permettre le serveur RStudio
pass in proto tcp from self \
  to self port $port_rstudio user $utilisateurR
```

Le fichier entier

Pourquoi

Avantages d'OpenBSD

- Exécution d'applications en arrière-plan (rc.subr)
- Machines virtuels (vmm)
- Pare-feu (pf)
- Documentation (man)
- Proxy d'applications Shiny (relayd)
- DevOps (autoinstall)
- Gestion de la configuration (/etc, site.tgz)
- Mises à jour (syspatch, sysupgrade)