

Linux の選択

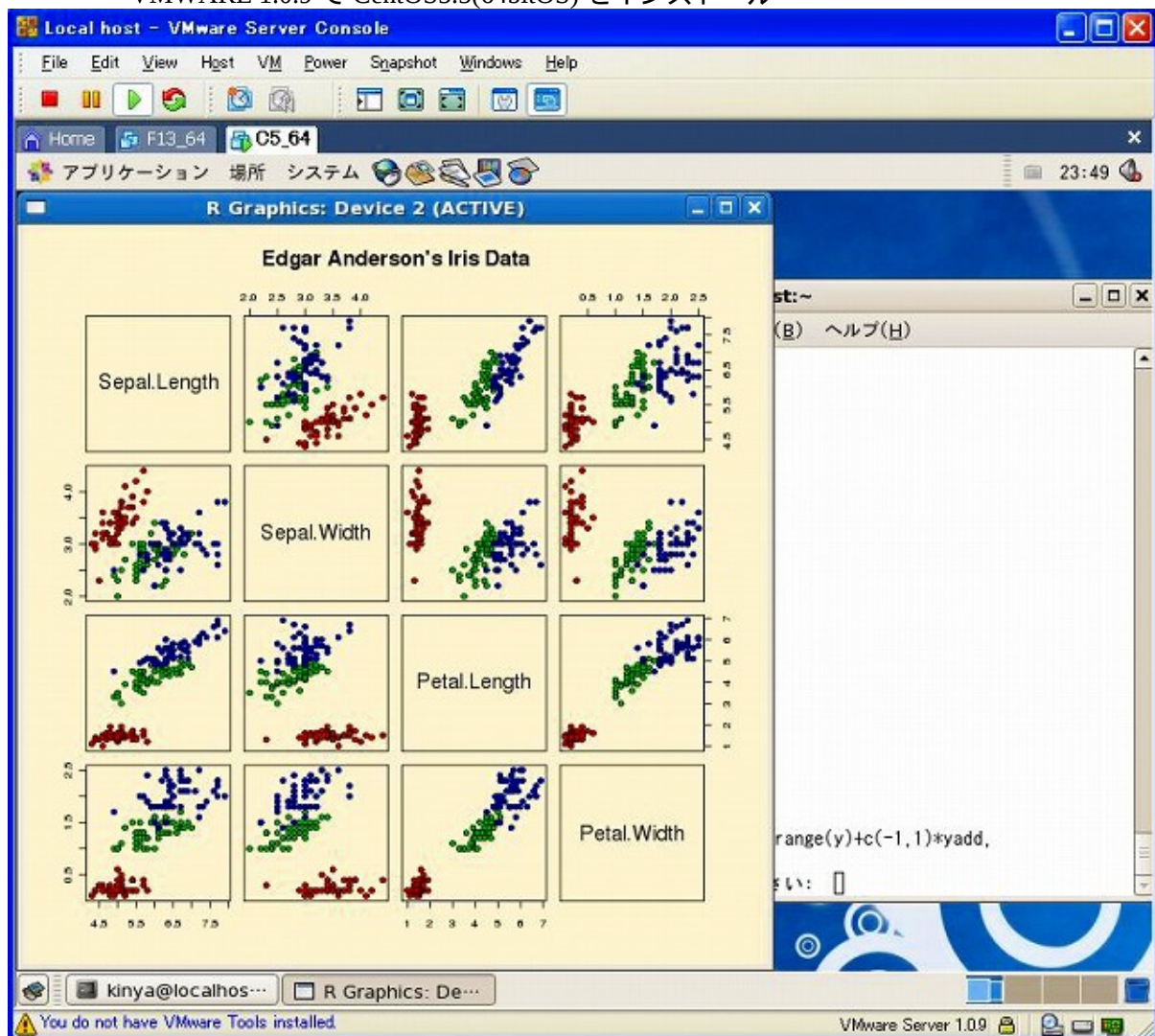
利用する Linux を選択する。

CentOS 5.5(RedHat Linux Clone OS) の 64bit 版を利用する。

- CentOS : <http://www.centos.org/>

インストール作業

- テキストマイニングで扱うデータが多い為、R で計算するとメモリエラーが表示される。
 - この改善策は、64bitOS と R(64bit 版) を利用することにより、扱えるメモリ容量 (2 GB 超) が増加すると知ったからである。
 - OS は、32bit 版 OS ではなく、64bit 版 OS を利用する。
 - R も 32bit 版ではなく、64bit 版を利用する。
- 動作環境
 - Panasonic Let's Note CF-W8(CPU:Core2Duo MEM:4GB)
 - VMWARE 1.0.9 で CentOS5.5(64bitOS) をインストール



Linux に R をインストール

R(64bit 版) をインストール

- ・かなり強引なインストール方法です。きちんと整理し、シェル形式で書くつもりです。
 - ・以下、CentOS5.5(64bit) をインストール後、R(64bit) をダウンロードし、インストールを行ったログです。

```
[root@localhost kinya]# cd /usr/local/src/

[root@localhost src]# wget http://cran.md.tsukuba.ac.jp/bin/linux/redhat/el5/x86_64/R-2.10.0-2.el5.x86_64.rpm
[root@localhost src]# wget http://cran.md.tsukuba.ac.jp/bin/linux/redhat/el5/x86_64/R-core-2.10.0-2.el5.x86_64.rpm
[root@localhost src]# wget http://cran.md.tsukuba.ac.jp/bin/linux/redhat/el5/x86_64/R-devel-2.10.0-2.el5.x86_64.rpm
[root@localhost src]# wget http://cran.md.tsukuba.ac.jp/bin/linux/redhat/el5/x86_64/libRmath-2.10.0-2.el5.x86_64.rpm
[root@localhost src]# wget http://cran.md.tsukuba.ac.jp/bin/linux/redhat/el5/x86_64/libRmath-devel-2.10.0-2.el5.x86_64.rpm

[root@localhost src]# ls
R-2.10.0-2.el5.x86_64.rpm      libRmath-2.10.0-2.el5.x86_64.rpm
R-core-2.10.0-2.el5.x86_64.rpm  libRmath-devel-2.10.0-2.el5.x86_64.rpm
R-devel-2.10.0-2.el5.x86_64.rpm

[root@localhost src]# rpm -ivh libRmath-2.10.0-2.el5.x86_64.rpm
[root@localhost src]# rpm -ivh libRmath-devel-2.10.0-2.el5.x86_64.rpm

[root@localhost src]# yum -y install bzip2-devel*
[root@localhost src]# yum -y install gcc*
[root@localhost src]# yum -y install perl*
[root@localhost src]# yum -y install tetex-latex*
[root@localhost src]# yum -y install xdg-utils*
[root@localhost src]# yum -y install libX*
[root@localhost src]# yum -y install pcre-devel*
[root@localhost src]# yum -y install tcl-devel*
[root@localhost src]# yum -y install tk-devel*

[root@localhost src]# wget http://packages.sw.be/perl-File-Recursive/perl-File-Recursive-0.11-1.2.el5.rf.noarch.rpm
[root@localhost src]# wget ftp://ftp.pbone.net/mirror/centos.karan.org/el5/extras/testing/i386/RPMS/perl-File-Copy-Recursive-0.35-1.el5.kb.noarch.rpm
[root@localhost src]# wget http://packages.sw.be/perl-File-Copy-Recursive/perl-File-Copy-Recursive-0.38-1.el5.rf.noarch.rpm

[root@localhost src]# ls
R-2.10.0-2.el5.x86_64.rpm      libRmath-devel-2.10.0-2.el5.x86_64.rpm
R-core-2.10.0-2.el5.x86_64.rpm  perl-File-Copy-Recursive-0.35-1.el5.kb.noarch.rpm
R-devel-2.10.0-2.el5.x86_64.rpm  perl-File-Copy-Recursive-0.38-1.el5.rf.noarch.rpm
libRmath-2.10.0-2.el5.x86_64.rpm  perl-File-Recursive-0.11-1.2.el5.rf.noarch.rpm

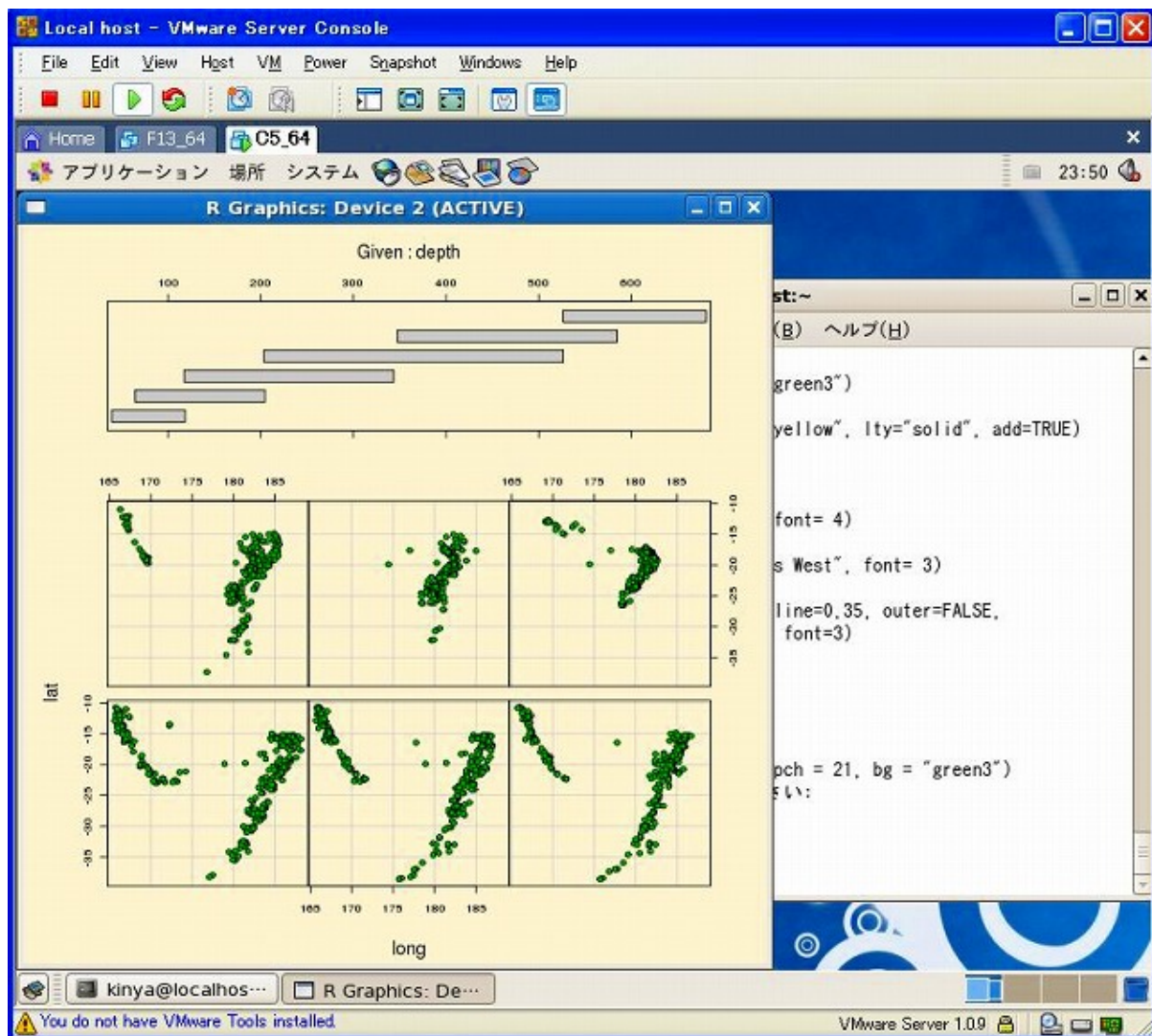
[root@localhost src]# rpm -ivh perl-File-Recursive-0.11-1.2.el5.rf.noarch.rpm
[root@localhost src]# rpm -ivh perl-File-Copy-Recursive-0.38-1.el5.rf.noarch.rpm

[root@localhost src]# rpm -ivh R-core-2.10.0-2.el5.x86_64.rpm
[root@localhost src]# rpm -ivh R-devel-2.10.0-2.el5.x86_64.rpm
[root@localhost src]# rpm -ivh R-2.10.0-2.el5.x86_64.rpm
```

- ・ここまで終了すると、GNOME の端末コンソールより、R を立ち上げる。
 - ・demo(graphics) とコマンドを打つと、以下のような綺麗なグラフのデモを見ることができる。

- ・R のコマンドラインから次のように入力する。

```
> demo(graphics)
```



MeCab のインストール

- ・形態素分析器 MeCab をインストール

```

[root@localhost kinya]# cd /usr/local/src/
[root@localhost src]# ls
R-2.10.0-2.el5.x86_64.rpm      libRmath-devel-2.10.0-2.el5.x86_64.rpm
R-core-2.10.0-2.el5.x86_64.rpm  perl-File-Copy-Recursive-0.35-1.el5.kb.noarch.rpm
R-devel-2.10.0-2.el5.x86_64.rpm  perl-File-Copy-Recursive-0.38-1.el5.rf.noarch.rpm
libRmath-2.10.0-2.el5.x86_64.rpm  perl-File-Recurse-0.11-1.2.el5.rf.noarch.rpm

[root@localhost src]# wget http://downloads.sourceforge.net/project/mecab/mecab/0.98/mecab-0.98
.tar.gz?use_mirror=jaist&ts=1280729466

[root@localhost src]# wget http://downloads.sourceforge.net/project/mecab/mecab-ipadic/2.7.0
-20070801/mecab-ipadic-2.7.0-20070801.tar.gz?use_mirror=jaist&ts=1280729552

[root@localhost src]# ls
R-2.10.0-2.el5.x86_64.rpm      mecab-0.98.tar.gz
R-core-2.10.0-2.el5.x86_64.rpm  mecab-ipadic-2.7.0-20070801.tar.gz
R-devel-2.10.0-2.el5.x86_64.rpm  perl-File-Copy-Recursive-0.35-1.el5.kb.noarch.rpm
libRmath-2.10.0-2.el5.x86_64.rpm  perl-File-Copy-Recursive-0.38-1.el5.rf.noarch.rpm
libRmath-devel-2.10.0-2.el5.x86_64.rpm  perl-File-Recurse-0.11-1.2.el5.rf.noarch.rpm

[root@localhost src]# tar xvfz mecab-0.98.tar.gz
[root@localhost src]# cd mecab-0.98
[root@localhost mecab-0.98]# ./configure --with-charset=utf8 --enable-utf8-only
[root@localhost mecab-0.98]# make
[root@localhost mecab-0.98]# make install

```

```
[root@localhost mecab-0.98]# cd ..
[root@localhost src]# tar xvfz mecab-ipadic-2.7.0-20070801.tar.gz
[root@localhost src]# cd mecab-ipadic-2.7.0-20070801

[root@localhost mecab-ipadic-2.7.0-20070801]# ./configure --with-charset=utf8
[root@localhost mecab-ipadic-2.7.0-20070801]# make
[root@localhost mecab-ipadic-2.7.0-20070801]# make install
```

- ・ MeCab の動作確認

- ・ インストールした後，以下のサンプルプログラムを入力し，うまく表示されれば OK。

```
[root@localhost src]# echo "今日はいい天気です。" | mecab
今日 名詞,副詞可能,*,*,* 今日,キョウ,キョー
は 助詞,係助詞,*,*,* は,ハ,ウ
いい 形容詞,自立,*,* 形容詞・イイ,基本形,いい,イイ,イイ
天気 名詞,一般,*,*,* 天気,テンキ,テンキ
です 助動詞,*,*,* 特殊・デス,基本形,です,デス,デス
。 記号,句点,*,*,* ,。 ,。 ,。
EOS
```

- ・ 茶筌（ちゃせん）オプションを使うと次の通りになります。

```
[root@localhost src]# echo "今日はいい天気です。" | mecab -Ochasen
今日 キョウ 今日 名詞 - 副詞可能
は ハ は 助詞 - 係助詞
いい イイ いい 形容詞 - 自立 形容詞・イイ 基本形
天気 テンキ 天気 名詞 - 一般
です デス です 助動詞 特殊・デス 基本形
。 。 記号 - 句点
EOS
```

RMeCab のインストール

- ・ RMeCab は，うまく wget でできなかったので，GNOME 上のブラウザからダウンロードする。
 - ・ <http://groups.google.com/group/rmecab/files>
 - ・ ファイル名は，次の通り。
 - ・ RMeCab_0.86_R_x86_64-unknown-linux-gnu.tar.gz

root 権限で，ダウンロードしてきたので，以下のフォルダに該当ソフトがあるので，移動する。

```
[root@localhost src]# cd /root/
[root@localhost ~]# ls
Desktop anaconda-ks.cfg install.log install.log.syslog

[root@localhost ~]# cd DEesktop/
[root@localhost Desktop]# ls
RMeCab_0.86_R_x86_64-unknown-linux-gnu.tar.gz
[root@localhost Desktop]# mv RMeCab_0.86_R_x86_64-unknown-linux-gnu.tar.gz /usr/local/src/
[root@localhost Desktop]# cd /usr/local/src/
```

- ・ /usr/local/src の現在の中身は，次の通り

```
[root@localhost src]# ls
R-2.10.0-2.el5.x86_64.rpm
R-core-2.10.0-2.el5.x86_64.rpm
R-devel-2.10.0-2.el5.x86_64.rpm
RMeCab_0.86.tar.gz
RMeCab_0.86_R_x86_64-unknown-linux-gnu.tar.gz
libRmath-2.10.0-2.el5.x86_64.rpm
libRmath-devel-2.10.0-2.el5.x86_64.rpm
mecab-0.98
mecab-0.98.tar.gz
mecab-ipadic-2.7.0-20070801
mecab-ipadic-2.7.0-20070801.tar.gz
perl-File-Copy-Recursive-0.35-1.el5.kb.noarch.rpm
perl-File-Copy-Recursive-0.38-1.el5.rf.noarch.rpm
```

perl-File-Recurse-0.11-1.2.el5.rf.noarch.rpm

- ・ R に先ほどダウンロードしてきたパッケージを読み込ませると・・・

下記のエラーが R 上で表示されるため、/etc/ld.so.conf に追記する。

(エラー内容)
Error in dyn.load(file, DLLpath = DLLpath, ...) :
共有ライブラリ '/usr/lib64/R/library/RMeCab/libs/RMeCab.so' を読み込めません
libmecab.so.1: 共有オブジェクトファイルを開けません: そのようなファイルやディレクトリはありません
エラー: 'RMeCab' に対するパッケージもしくは名前空間のロードが失敗しました

```
[root@localhost library]# vi /etc/ld.so.conf.d/mecab.conf  
/usr/local/lib を追加する。
```

```
[root@localhost library]# /sbin/ldconfig ←設定を反映させる。
```

- ・ R に先ほどダウンロードしてきたパッケージを読み込ませる。

```
[root@localhost src]# R
```

```
R version 2.10.0 (2009-10-26)  
Copyright (C) 2009 The R Foundation for Statistical Computing  
ISBN 3-900051-07-0
```

R は、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してください。

R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形式については
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

[以前にセーブされたワークスペースを復帰します]

```
> install.packages("RMeCab_0.86_R_x86_64-unknown-linux-gnu.tar.gz", destdir = ".", repos = NULL)  
install.packages("RMeCab_0.86.tar.gz", destdir = ".", repos = NULL) 中で警告がありました:  
引数 'lib' が欠けています: /usr/lib64/R/library を使います  
* installing *binary* package 'RMeCab' ...  
* DONE (RMeCab)
```

- ・ RMeCab の動作確認

```
[root@localhost library]# R  
R version 2.10.0 (2009-10-26)  
Copyright (C) 2009 The R Foundation for Statistical Computing  
ISBN 3-900051-07-0
```

R は、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してください。

R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形式については
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

```
>  
> library(RMeCab)  
> RMeCabC(" コーヒーに 300 円使いました。")  
[[1]] 名詞 " コーヒー "  
[[2]] 助詞 " に "  
[[3]] 名詞 "300"
```

```

[[4]] 名詞 "円"
[[5]] 動詞 "使い"
[[6]] 助動詞 "まし"
[[7]] 助動詞 "た"
[[8]] 記号 "。"

>
> res <- RMeCabC(" これはペンです。")
> res
[[1]] 名詞 "これ"
[[2]] 助詞 "は"
[[3]] 名詞 "ペン"
[[4]] 助動詞 "です"
[[5]] 記号 "。"
> unlist(res)
名詞 助詞 名詞 助動詞 記号
"これ" "は" "ペン" "です" "。"
> q()
Save workspace image? [y/n/c]: y

```

Windows と VMWARE 上の R(Linux) とのデータ共有

Samba の設定

■ Samba のインストール

```
[root@localhost ~]# yum -y install samba
```

■ Samba サーバーアクセス用ユーザ作成

```

[root@localhost ~]# useradd user
[root@localhost ~]# passwd user

Changing password for user user.
New UNIX password:
Retype new UNIX password:
passwd: all authentication tokens updated successfully.

```

■ ホームディレクトリを作成（新規ユーザ）

```
[root@localhost ~]# mkdir /etc/skel/samba
```

■ ホームディレクトリを作成（既存ユーザ）

```

[root@localhost ~]# vi mkhomedir.sh

#!/bin/bashfor user in `ls /home`
do
id $user > /dev/null 2>&1
[ $? -eq 0 ] && ¥
[ ! -d /home/$user/samba ] && ¥
mkdir /home/$user/samba && ¥
chown $user:$user /home/$user/samba && ¥
echo "/home/$user/samba create"
done

[root@localhost ~]# sh mkhomedir.sh
/home/user/samba create

```

■ 共有ディレクトリを作成

```

[root@localhost ~]# mkdir /home/samba
[root@localhost ~]# chown nobody:nobody /home/samba

```

■ Samba 設定

```
[root@localhost ~]# vi /etc/samba/smb.conf
```

参考サイト <http://centosrv.com/samba.shtml>

■ Samba がうまく動かない場合 , 次の作業も行った。(1)

- ・ ネットワーク設定 (設定内容は , 参考 URL から拝借) ・ nat する。 dhcp は使わない。

```
$ cat /etc/sysconfig/network-scripts/ifcfg-eth0
# Advanced Micro Devices [AMD] 79c970 [PCnet32 LANCE]
DEVICE=eth0
BOOTPROTO=static
HWADDR=00:50:56:7E:06:58
IPADDR=192.168.249.10
NETMASK=255.255.255.0
GATEWAY=192.168.249.2
ONBOOT=yes
```

- ・ 変更を反映させる為に次のコマンドで更新。

```
# /sbin/service network restart
```

参考 URL <http://d.hatena.ne.jp/ZIGOROu/20090111/1231673135>

■ Samba がうまく動かない場合 , 次の作業も行った。(2)

- ・ SELinux が原因 ?

```
# getenforce で状態確認
# setenforce 0 SELinux 無効化
```

- ・ iptables が原因 ?

- ・ 確認方法

```
/etc/init.d/iptables stop
```

- ・ これでアクセスできたら iptables が原因

- ・ samba を許可

```
iptables -I INPUT 3 -p tcp --dport 445 -j ACCEPT
iptables -I INPUT 3 -p tcp --dport 139 -j ACCEPT
iptables -I INPUT 3 -p udp --dport 138 -j ACCEPT
iptables -I INPUT 3 -p udp --dport 137 -j ACCEPT
```

参考 URL (サンプル事例をコピーさせて頂きました) http://d.hatena.ne.jp/red_snow/20100805/1280984405

■ Samba がうまく動かない場合，次の作業も行った。(3)

・Windows は，samba の home directry を認識できなかった為，次の作業を行った。原因は，Linux 側の設定が，localhost.localdomain となっている事。

・ network ファイルを開いて設定を変更

```
# vi /etc/sysconfig/network  
  
NETWORKING=yes  
NETWORKING_IPV6=yes  
HOSTNAME=localhost.localdomain
```

・ 変更例

```
NETWORKING=yes  
NETWORKING_IPV6=yes  
HOSTNAME=vm-host
```

・ Linux を再起動させる

```
# reboot
```

参考 URL <http://wopose.blogspot.com/2009/10/vmware-centos-samba.html>

Linux(Samba)-Windows での不具合

・ 文字コードの不具合

- ・ Linux (Samba) 側 UTF-8 , Windows 側 SJIS
- ・ 資料 (.TXT 処理済) を Samba にアップロード
- ・ Teraterm で R を起動させる。 → Windows 側からプログラムをコピー & ペーストをして処理をさせる。
- ・ 以上の処理で資料の文字を読み取らず，テキストマイニングできないエラーが生じた。

・ 理由

- ・ Windows で作成した資料の文字コードは，SJIS (Samba では，CP932)。 Linux (Samba) は，UTF-8。
- ・ このような理由から，Linux 側では，SJIS のデータは，文字化けをしていた模様。

・ 解決策

- ・ Linux (Samba) 側にアップしたデータをエディタで開き，文字コードを UTF-8 で保存しなおす必要あり。

・ その他の解決策？

- ・ Linux(Samba) 側の設定 (/etc/samba/smb.conf) の文字コードを変更すれば OK？
- ・ 参考 URL
- ・ <http://d.hatena.ne.jp/rsyudo/20070621/p1>

LinuxPC の複数台活用

PC Cluster を参照