

SuSE Linux クラスタ計算機の利用・運用手引き

11/01/2009 M.Tanaka

<http://dphysique.nifs.ac.jp/>

はじめに

ここでは、オペレーティングシステムとして SuSE Linux を用いたクラスタ計算機の利用と運用について、最低限必要となる事柄について記述しています。詳細や疑問の点は、必ず解説書（最終ページに参考文献）を読んで理解してください。

一連の操作を行うためには、まず Linux がインストールされた計算機において、X-Window と呼ばれるコマンド入力窓を、開始メニューから選択して呼び出します。 様々な操作をグラフィカルな対話形式で行える YaST メニューがありますが、これは個別コマンド群を集めたものです（システムの問題解決のときは、普通個別コマンドを使います）。

以下の説明で、コマンドの前につく記号 # は管理者権限での実行、% はユーザーモードでの実行を表すプロンプトです（システム状態に関するコマンドは、**管理者権限がないと実行できない設計になっています**）。

ネットワークを通して遠隔の Linux マシンからログインするときは、secure shell である SSH を使います（デフォルトで telnet 接続は拒否されます）：

```
% ssh -l user_name host_ip_address <- ハイフン l は同じユーザ名ならば省略可能
```

MS Windows からも（無料）ソフトウェアをインストールすることでアクセスやファイル転送が可能です。ただし、利用するプログラムによっては、行末の改行文字(^M)など、不可解な問題が起きることがあります（MS Windows 上で編集、またはダウンロードしたファイルを Linux システムへ転送すると、**行末に改行文字 2 種 CR, LF が埋め込まれ、誤動作の原因に!**）。

[A] システムを利用する

● テキストの編集法

X-Window 上でのテキストファイルの編集には、動作が軽快な「vi エディター」を推奨します。機能が豊富な Emacs などにも広く利用されています。

代表的な vi の操作法

以下のコマンドを知っていればで、普通は十分でしょう。ただし 1 冊は解説書を読むことを勧めます。

編集の開始

```
% vi file_name
```

file_name という名前のファイルの編集を開始する。この直後、すでにファイルが存在

する場合は、その中身が表示される。新規ファイルの場合は空白画面になる。入力を始めるには以下の操作を行う。

- i i (input モードの意味) をタイプ (表示されない) して、入力モードに入る。これ以後、タイプした文字が画面上に表示される。
- [escape キー] 一般に、直前の 1 操作を取り消す (間違えたときに便利なキー。何度押ししても、1 回押したときと同じ)。

編集の終了

- :wq (コロンの後に wq) セーブした後、編集を終了する。
- :q! (コロンの後に q!) 編集を放棄して、終了する

コピーと消去

文字の挿入 (input) : 挿入したい文字 (の前) にカーソルを移動し、i をタイプする。

そこから挿入が開始される。挿入を終了する時は、[escape キー] を押す。

文字の消去 (cross out?) : カーソルを文字にあわせ、x をタイプすることで 1 文字が消去できる。「文字数 xx」とタイプすることで複数の文字が消せる。

行の消去 (delete) : 消去する行にカーソルをあわせ、dd をタイプ。行数に続いて「行数 dd」をタイプすることで、複数行が消去できる。消去した内容はバッファーに保存されるので、以下のコピーで再利用できる。

行のコピー (yank and paste) :

コピーしたい最初の行にカーソルを合わせ、yy または「行数 yy」とタイプすることで、バッファーに内容を保存する。続いて、コピー先の行にカーソルを移動し、p をタイプすると、この行の直後にバッファーに内容が挿入される。

行の移動 : 「行の消去」と「行のコピー」機能を使う。

検索と置換

文字の置換 : 置き換えたい文字にカーソルを移動し、r をタイプし、希望の文字をそこにタイプする。

文字の検索 : / (スラッシュ) に続けて文字をタイプし、[return キー] を押す。再検索は n (next) をタイプする。

文字の一括置換 : パターンを検索して置換することができる。まずコロンをタイプし、%s/ に続けて以下のように置換する前の文字列、置換後の文字列を/ で区切って入力する。最後の g (任意) は、合致するすべてを置換することを指示する。
%s/patter_in/pattern_pit/g

ジャンプ

次のページ : ctrl キー f (同時に押す)

前のページへ : ctrl キー b (同時に押す)

ある行へジャンプ : 「行番号 G」とタイプする。ファイルの終端へジャンプするときは、「G」をタイプする。先頭へもとるときは「1G」。

外部ファイルの利用

連続する複数行を外部のファイルへ書き出す (write) :

コロンに続き、開始行と終了行の行番号を入力 (, で分離する)、続けて w をタイプ、その次にスペースを空けてファイル名をタイプする、
:line1,lin2w file_name
すでに存在するファイルへ上書きするときは、w でなく「w!」をタイプする。

テキストファイルの読み込み (read) :

挿入する直前の行にカーソルを移動し、コロンの直後に r、スペースを空けてファイル名をタイプする : :r file_name

● Linux コマンド

X-Window で使用する多くの Linux (Unix) コマンドは 2 文字で表されます。以下のコマンドは、可能な限りユーザーモードで実行します (システムに悪影響する操作間違いを防ぐため)。

コピー、削除、移動

ファイルのコピー :

%cp の後、スペースで区切り、コピー元、コピー先のファイル名をタイプする。
-v (任意) を指定すると、すでにコピー先のファイルが存在する場合、コピーの動作確認を求められる。
% cp -v file_in file_out

ファイルの削除 : % rm -v file_name

ファイル名の変更、または移動 : % mv -v file_in file_out

%mv のあと、スペースを空けて、元のファイル名、変更後のファイル名をタイプ。
ファイル名 file_out の代わりに領域名 dir_out を指定すると、その領域へファイルを移動する。

表示

ファイルのリスト表示 : % ls -l -l を指定すると、属性、サイズ情報などが表示される
別の領域への移動 : %cd dir_name

1 ページずつ表示 : % ls -l | less コマンドの後に、| less をタイプする

ファイル操作

ファイル属性の変更 : % chmod 750 file_name

各ファイルはアクセス属性を持つ (上でのべた% ls -l で表示されるリスト参照)。
r, w, x はそれぞれ read, write, execute に対応し、数値の 4, 2, 1 が対応する。最初桁の d は、それが領域であることを示しており、属性が x のときだけ、その領域に移ることができる。

r+w+x の結合した属性は数字 7 (4+2+1) で表される。

chmod のあとの 3 桁の数字は、それぞれ自分自身、同じグループのユーザー、どちらでもない第 3 者のアクセス権限を示す。

750 の意味：自分はこのファイルに対して read, write, execute ができ、同じグループのユーザーは read, execute ができるが、第 3 者は何もできない。

所有者の変更：

```
% chown user_id, user_group file_name
```

```
% chown -R dir_name
```

所有するユーザーとグループを指定し、続けてファイルの名前をタイプする。

file_name の代わりに領域名 dir_name を与え、-R を指定すると、領域の内部まで遡及して所有書を変更する。

情報を見る

ヘルプ： % command_name -help

ハイフン 2 個の後に help で、簡単な使用説明が見られる（備考録程度の説明）

オンラインマニュアル： %man command_name

上のヘルプコマンドより詳しい説明が読める。すべてのコマンドに付随するわけではないが、かなり役に立つ情報が示される。

Linux (Unix) コマンドは非常に多岐にわたります。Linux の成り立ち、構造を含めて系統的に理解するためには、解説書を読んで勉強してください。

[B] システムを運用する

これらは管理者権限で実行します

● Firewall を設定する

SuSE のファイアーウォールは RSH を通さない設定になっています。もし rsh での内部通信が必要なときは、rsh-server をインストールした上で、以下のスクリプトを、たとえば SuSE-Firewall という名前のファイルに格納し、管理ノードの上で以下を行います。

```
% chmod 700 SuSE-Firewall    実行できるように変更
# ./ SuSE-Firewall          管理者権限で実行する
```

セキュアシェルである SSH により、ローカルシステム内でパスワードなしで相互にアクセス可能とすることも可能です。

---- SuSE-Firewall の内容（あくまでも一例） -----

```
#!/bin/sh
```

```
# ここでは、# は 1 行目を除き、コメント行の意味
```

```
/sbin/modprobe iptable_nat
```

```
echo 1 > /proc/sys/net/ipv4/ip_forward
```

```
# LOCAL1 - tcp/ip
```

```
LOCAL1='192.168.1.0/24' # 自分の環境に合わせて修正する
```

```

GLOBAL='123.123.1.100' # 自分の環境に合わせて修正する

# Initialize (flush rules)
/usr/sbin/iptables -t filter -F INPUT
/usr/sbin/iptables -t filter -F OUTPUT
/usr/sbin/iptables -t filter -F FORWARD
/usr/sbin/iptables -t nat -F POSTROUTING

# Basic policies
/usr/sbin/iptables -t filter -P OUTPUT ACCEPT
/usr/sbin/iptables -t filter -P INPUT DROP
/usr/sbin/iptables -t filter -P FORWARD DROP
#
/usr/sbin/iptables -t filter -A INPUT -m state --state ✕
ESTABLISHED,RELATED -j ACCEPT
/usr/sbin/iptables -t filter -A FORWARD -m state --state ✕
ESTABLISHED,RELATED -j ACCEPT

# Masquerade
#/usr/sbin/iptables -t nat -A POSTROUTING -o $GLOBAL -j MASQUERADE

# Allow DNS
/usr/sbin/iptables -t filter -A INPUT -d $GLOBAL -p udp --dport 53 -j ACCEPT

# Allow loopback
/usr/sbin/iptables -t filter -A INPUT -i lo -j ACCEPT

/usr/sbin/iptables -t filter -A INPUT -s $LOCAL1 -p tcp -j ACCEPT
/usr/sbin/iptables -t filter -A INPUT -s $LOCAL1 -p udp -j ACCEPT

# Allow ping
# % ping ip-address で通信可能かテストできる。安全上は無効がよい。
#/usr/sbin/iptables -t filter -A INPUT -p icmp -j ACCEPT

# Allow ssh
/usr/sbin/iptables -t filter -A INPUT -d $GLOBAL -p tcp --dport ssh -j ACCEPT

# Allow http
#/usr/sbin/iptables -t filter -A INPUT -d $GLOBAL -p tcp --dport 80 -j ACCEPT

# verify
/usr/sbin/iptables -L
-----

```

● ネットワークファイルシステム(NFS)の起動

サーバノードにある領域を、ネットワークを介して利用するため

(a) サーバノードで

ファイル/etc/exports に、NFS を許す内容を記述する。その例は、
/home *(rw,no_root_squash)

<- 領域/home をすべてのノードに、read/write 可能状態で開放する
(詳細は解説書を参照してください)

```
# /etc/init.d/nfsserver start    サーバデーモンを起動する  
# ps ax | grep nfs             nfs (という名前の手続き) の起動を確認するため
```

(b) 計算ノードで

```
# mkdir dir_name              マウントポイント dir_name を作成する (初めの 1 回だけ実行)  
# mount s100: /dir_name0 /dir_name  
    サーバ-s100 の領域 dir_name0 を、dir_name という名前でこのノードに  
    マウントすることで、遠隔利用を可能とする。
```

```
# df                            正しくマウントしたことを確認するため、領域リストを表示する。
```

なお、領域を外すときは、# umount dir_name で行います。

/etc/fstab にマウント情報を記載すると、システム立ち上げ時に自動的に NFS 領域がマウントされます。ただし、システムに問題が発生したとき、NFS 参照ができずにシステムが「永久待ち状態」に陥ることがあります。

● システムの再起動、停止

```
# /sbin/shutdown -r now        -r の指定をすると再起動、-h はシャットダウン。
```

Ctrl, Alt, Delete キーの同時押しで、システムを再起動できます。

電源ボタンを 4 秒以上押し続けることで電源が切れますが、これは最後の手段としてください (ハードディスクが壊れることもあるようです、幸い経験はないが)。

● サービスの有効化

初期設定では rsh や xinetd は無効 Off になっているので、有効にします。

また、/etc/hosts.equiv に rsh を許すホスト名を列記します (再起動が必要)。

```
# /sbin/chkconfig -l          (マイナス記号を 2 回タイプ) 現在のサービス状態の確認  
# /sbin/ifconfig rsh on      たとえば、rsh を有効にする
```

● ネットワーク状態の確認

ネットワークがつながらないときの、最初のチェックポイント。ネットワークカードの差し方が悪い (通信ランプが点滅しているか?) ほか、ファイアーウォールによりアクセスがブロックされている場合もあります。ブート時に、Esc キーか PF-1 キーを押すと、eth0 ...

の情報が表示されるので、それを見るとハードウェア、ソフトウェアのいずれに問題があるか見当がつけられます。

(a) ネットワークの設定を確認する

```
# /sbin/ifconfig eth0, eth1 などの設定状態が表示される。状態変更もできる  
(詳細は、解説書を見てください)
```

(b) ネットワークは存在するか?

```
# /sbin/route -n 存在しているネットワークの状態が表示される
```

(c) (低レベルで) 通信できるか?

```
% ping host_name 接続されていれば、反応が返ってくる
```

● ホスト名のリスト

ローカルなホスト名を/etc/hosts に列挙します。よく使うホスト名を書いておくと、ネームサーバがないシステムで、短縮名が使えます。

```
192.168.1.100 s100.cluster1 s100 <- これで s100 という短縮名が有効に
```

● コンパイラー、NQS の導入

実際に理工学計算を高速に行うためには、商用のコンパイラー（個人的には高速性と安定性の理由で PGI コンパイラーを推奨）で、C/ C++/ Fortran がセットになったものを導入します。ライセンスは通常ネットワークカード固有の MAC アドレスで管理されます。インストールの後、ライブラリなどの参照 PATH を与えるなど、システムおよびユーザーレベルでいくつかの設定が必要です。

ユーザーと使用マシンが限定されるとき以外は、ジョブは、NQS (Network Queuing System) を介して、サブミットして実行します。空いているノードが自動的に割り当てられます。無料で TORQUE や、その改良版がダウンロード、利用できます（利用規程が記載されています）。

● ホームページの運用

セキュリティの高いホームページの運用は、Linux 上で、Apache を利用することを推奨します。概要とインストール法は <http://www.apache.org/> に記載されています。コンテンツの作成は熟練すればマニュアルでもできますが、既成のソフトウェアを利用すると容易です。

なお、URL 名は www で始まる必要はなく、固定アドレスで取得できる名前であれば使用できます。また、安全対策上、クラスター計算機とは分離して、専用サーバーにより運用すべきでしょう。

さいごに

Linux システムに関する最新の情報に関しては、システムの/doc 領域に含まれる、またはオンライン検索で見つかる Linux フォーラムの Howto 文書などが有用です。進歩の早い部分や特殊技法は、ネットワークで見つけることができます。しかし、真偽の判断は忘れずに、そして実施は各自の責任で行ってください。

参考文献

Unix 入門 (トッパン、小山ほか、1996)

Linux ネットワーク (トッパン、小山ほか、1998)

どちらも Linux(Unix)の利用法をを系統的に解説した名著 (絶版?)。出版から 10 年が経過し、個々の点 (とくにネットワーク関係) で記述が古くなった部分もありますが、根幹は不変です。

入門 vi (O'Reilly ジャパン、Lamb & Robinson, 2002。日本語訳)

vi スクリーンエディターのよくわかる入門書。

Apache ハンドブック (O'Reilly ジャパン、第 3 版 2003)

第 3 版から、Apache ver.2 の解説があります。