

NFS on CLUSTERPRO for Linux HOWTO

1 はじめに

この文書はLinuxのNFS機能（サーバ部分）をCLUSTERPRO X for Linuxと組み合わせてフェイルオーバークラスタ化するために必要な情報を記述したものです。たとえば、ファイルサーバ等をフェイルオーバークラスタ化することにより、単一Linuxサーバでは達成し得ない可用性の高い高信頼なファイルサーバを構築することができます。

この文書を書くにあたって次のバージョンのソフトウェアを使用しました。この他のバージョンのソフトウェアを使用した場合でも、いくつかの設定項目の読み替えでクラスタ化できると考えられます。

- Red Hat Enterprise Linux AS release 4 (Nauant Update2)
- CLUSTERPRO X 1.0 for Linux
- nfs-utils 1.0.6-65.EL4

2 責任範囲

この文書は、クラスタシステム上の共有ディスク等の切り替えパーティションにNFSの設定をするための注意点や具体的な設定例を示すことにより、システムインテグレーション時の参考情報となるものであり、NFSの動作保証を行うものではありません。

3 どのようにクラスタ化するか

- (1) 基本的にシングル状態のサーバと同様です。NFSサービスを起動し、`exportfs`を用いて、`export`されるマウントポイントを設定し、NFSサービスに反映させます。
- (2) マウントされる対象が切替パーティション（あるいはその配下のディレクトリ）になります。
- (3) すべてのサーバで切り替えパーティションを切り替えながら設定を行います。動作確認をおこないます。

4 クラスタ化のための設定

※ 例は2サーバの共有ディスクタイプのクラスタシステムです。ミラーディスクタイプの場合は共有ディスクの部分をミラーディスクに読み替えてください。

マウントポイント : /mnt/sdb2/
FIP (フローティングIP) : 192. 168. 1. 100

あらかじめ、NFSが動作できる環境を整えておいてください。今回は/sbin/serviceを使用します。

nfsd等の起動/終了の制御はCLUSTERPROの起動スクリプト/終了スクリプトで行います。OSで自動起動設定にしないでください。なお、portmapper等は、すでに起動しておくようにしておいてください。

※ NFSやexportfsなどの作成方法、設定方法は別の文献などを参照してください。

4.1 リソースの設計

FIPを用いてNFSを動作させる場合には、FIP活性がnfsd起動よりも後、FIP非活性がnfsd停止よりも前に行う必要があります。

そのためグループリソースの依存関係で、活性順序を設定する必要があります。

全体の依存関係を、diskリソース、execリソース、FIPリソースの順番で活性するように設定してください。

非活性の時は活性時の逆になりますので、FIPリソース、execリソース、diskリソースの順番で非活性されます。

【全体の依存関係】

深度	名前	依存リソース名	タイプ
0	disk1	none	
1	exec1	disk1	disk resource
2	fip1	disk1	disk resource
		exec1	execute resource

詳しくは CLUSTERPRO X 1.0 for Linux リファレンスガイド

第5章 グループリソースの詳細を参照してください。

スクリプトの具体的な記述方法は「4.2nfsスクリプトの編集」や「4.3サンプルスクリプト」を参考にしてください。また、マウントポイントは事前に作成しておいてください。

必要情報	説明
FIP (フローティングIP アドレス)	NFSクライアントからシームレスに接続するために使用するIPアドレスです。サーバの実IPアドレスと同じネットワークアドレスで、ほかと重複しないIPアドレスを使用してください。 FIPリソースに設定してください。 例) 192.168.1.100
共有 (ミラー) ディスク上のパーティション	ディスクリソースに設定してください。 NFSでシェアするファイルシステムを構築しているパーティションデバイスのスペシャルファイル名 CLUSTERPROのトレッキングツールで指定します。 1つのフェイルオーバーグループに複数のディスクリソースを登録することも可能です。 例) /dev/sdb2
マウントポイント	ディスクリソースに設定してください。 上記ファイルシステムをマウントするマウントポイントです。 CLUSTERPROのスクリプトで使用します。 使用する分だけ用意してください。 例) /mnt/sdb2

4.2 nfsスクリプトの編集

CLUSTERPROの起動スクリプト、終了スクリプト (start.sh、stop.sh) にてNFSサービスの起動/停止を行います。(4.3サンプルスクリプト)

OS起動/終了時に自動的にNFSサービスを起動/終了しないようにしてください。

CLUSTERPROの起動スクリプト、終了スクリプト (start.sh、stop.sh) にてexportfsするために、/etc/rc.d/init.d/nfsを編集します。

Startと**Stop**でexportfsを実行している部分のみコメントアウト、あるいは削除します。下記の修正例を参考にしてください。

※ Red Hat Enterprise Linux AS release 3の例です。斜体字がコメントアウトする部分です。

```
#!/bin/sh
#
# nfs          This shell script takes care of starting and stopping
#             the NFS services.

~~~~中略~~~~

# See how we were called.
case "$1" in
start)
    # Start daemons.

    # Set the ports lockd should listen on
    if [ -n "LOCKD_TCPPOINT" ]; then
        /sbin/sysctl -w fs.nfs.nlm_tcpport=$LOCKD_TCPPOINT >/dev/null
2>&1
    fi
    if [ -n "LOCKD_UDPOINT" ]; then
        /sbin/sysctl -w fs.nfs.nlm_udpoint=$LOCKD_UDPOINT >/dev/null
2>&1
    fi

    action $"Starting NFS services: " #/usr/sbin/exportfs -r
    if [ -n "$RQUOTAD" -a "$RQUOTAD" != "no" ]; then
        echo -n $"Starting NFS quotas: "
        daemon rpc.rquotad
        echo
    fi

~~~~中略~~~~

stop)
    # Stop daemons.
    echo -n $"Shutting down NFS mountd: "
    killproc rpc.mountd
    echo
    echo -n $"Shutting down NFS daemon: "
    killproc nfsd
    echo
    if [ -n "$RQUOTAD" ]; then
        echo -n $"Shutting down NFS quotas: "
        killproc rpc.rquotad
```

```

echo
fi
# Reset the lockd ports if they were set
if [ -n "LOCKD_TCPPORT" ]; then
    /sbin/sysctl -w fs.nfs.nlm_tcpport=0 >/dev/null 2>&1
fi
if [ -n "LOCKD_UDPPORT" ]; then
    /sbin/sysctl -w fs.nfs.nlm_udpport=0 >/dev/null 2>&1
fi
# Do it the last so that clients can still access the server
# when the server is running.
action $"Shutting down NFS services: " #/usr/sbin/exportfs -au
rm -f /var/lock/subsys/nfs
;;
~~~~中略~~~~
;;
echo $"Usage: nfs {start | stop | status | restart | reload | condrestart}"
exit 1
esac
exit 0

```

※ 以上の修正により、サーバ起動時にnfsdがfaildと表示されることがありますが問題ありません。

4.3 サンプルスクリプト

※ 斜体になっているところが追加されたところです。

※ 網掛け部分は環境に合わせて変更してください。

開始スクリプト

```

#!/bin/sh
#*****
#*                start.sh                *
#*****

if [ "$CLP_EVENT" = "START" ]
then
    if [ "$CLP_DISK" = "SUCCESS" ]
    then
        echo "NORMAL1"

        /sbin/service nfs start > /dev/null 2>&1----(a)
        /usr/sbin/exportfs -o rw,no_root_squash,insecure_locks:/mnt/sdb2/ >
        /dev/null 2>&1----(b)
        /bin/chmod 777 /mnt/sdb2 > /dev/null 2>&1----(c)

    if [ "$CLP_SERVER" = "HOME" ]
    then
        echo "NORMAL2"
    else
        echo "ON_OTHER1"
    fi
else

```

```

    echo "ERROR_DISK from START"
    fi
elif [ "$CLP_EVENT" = "FAILOVER" ]
then
    if [ "$CLP_DISK" = "SUCCESS" ]
    then
        echo "FAILOVER1"

        /sbin/service nfs start > /dev/null 2>&1----(a)
        /usr/sbin/exportfs -o rw,no_root_squash,insecure_locks:/mnt/sdb2/ >
/dev/null 2>&1----(b)
        /bin/chmod 777 /mnt/sdb2 > /dev/null 2>&1----(c)

    if [ "$CLP_SERVER" = "HOME" ]
    then
        echo "FAILOVER2"
    else
        echo "ON_OTHER2"
    fi
    else
        echo "ERROR_DISK from FAILOVER"
    fi
else
    echo "NO_CLP"
fi
echo "EXIT"
exit 0

```

<開始スクリプトの説明>

- (a) /sbin/service でnfsを起動させる。
- (b) /usr/sbin/exportfsにてexportすべきmountpointとoptionを指定する。
- (c) 共有ディスク上のファイルシステムのmountpointのパーミッション変更。

終了スクリプト

```

#!/bin/sh
#*****
#*                stop.sh                *
#*****

if [ "$CLP_EVENT" = "START" ]
then
    if [ "$CLP_DISK" = "SUCCESS" ]
    then
        echo "NORMAL1"

        /usr/sbin/exportfs -u r:/mnt/sdb2/ > /dev/null 2>&1----(d)
        /sbin/service nfs stop > /dev/null 2>&1----(e)
        sleep 30 > /dev/null 2>&1----(f)

    if [ "$CLP_SERVER" = "HOME" ]
    then
        echo "NORMAL2"
    else
        echo "ON_OTHER1"
    fi

```

```

else
    echo "ERROR_DISK from START"
fi
elif [ "$CLP_EVENT" = "FAILOVER" ]
then
    if [ "$CLP_DISK" = "SUCCESS" ]
    then
        echo "FAILOVER1"

        /usr/sbin/exportfs -u r:/mnt/sdb2/ > /dev/null 2>&1----(d)
        /sbin/service nfs stop > /dev/null 2>&1----(e)
        sleep 30 > /dev/null 2>&1----(f)

        if [ "$CLP_SERVER" = "HOME" ]
        then
            echo "FAILOVER2"
        else
            echo "ON_OTHER2"
        fi

    else
        echo "ERROR_DISK from FAILOVER"
    fi
else
    echo "NO_CLP"
fi
echo "EXIT"
exit 0

```

<終了スクリプトの説明>

- (d) /usr/sbin/exportfsにてexportしていたmountpointを消す。
- (e) /sbin/serviceでnfsを停止させる。
- (f) nfs関連プロセスの停止待ち。

4.4 NFSサーバ側のオプション

以下のオプションを使用しました。本オプションはCLUSTERPROとの関連性はありませんが使用目的等にあわせて調整してください。

rw リードライト可
no_root_squash クライアントにローカルのroot権限を与える
insecure_locks ロック要求時に承認を必要としない

4.5 クライアントのマウントオプション

以下のオプションを使用しました。本オプションはCLUSTERPROとの関連性はありませんが使用目的等にあわせて調整してください。

```
mount -t nfs -o hard,intr ...
```

-tオプション
ファイルタイプ設定 (NFSを指定)

-oオプション
hard : ファイル操作の再試行 intr : シグナルによるファイル操作中断許可

4.6 動作確認

手動mountにて同一ネットワーク上のLinuxマシン (Red Hat Linux 8.0) からmountし、ファイル作成/削除、ディレクトリ作成/削除をそれぞれのサーバから確認しました。

4.7 注意事項・その他

- portmap等のあらかじめNFSを動作させるに当たり必要なものは事前に起動させるか、スクリプトに記述するようにしてください。
- sleepの待ち時間は環境により調整が必要です。NFSクライアントからファイル等をアクセスした状態でフェイルオーバーグループを停止して“ディスクの切り離し失敗”が発生しないようにしてください。
- フェイルオーバー時にクライアント側でNFSに対する反応がなくなります。フェイルオーバーが完了し、NFSクライアントがリトライできた時点で再度使えるようになります。
- フェイルオーバーした直後にクライアントより接続すると失敗する場合があります。NFSクライアントのセッション終了をsleepの時間より短い時間に設定してください。
- 複数のファイルシステムをexportする場合は、「4.3サンプルスクリプト」の(b)を必要数分記述し、呼応して(d)を増やしてください。