

NFS On CLUSTERPRO for Linux

1 はじめに

この文書はLinuxのNFS機能（サーバ部分）をCLUSTERPRO for Linuxと組み合わせてフェイルオーバ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思われます。

- MIRACLE LINUX Standard Edition V2.1
- CLUSTERPRO SE for Linux Ver2.1
- kmountd nfs-utils 0.3.1

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。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。
- (2) 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。
- (3) 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/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、portmaper等すでに起動しておくようにしておいて下さい。

※ 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リソースをグループに登録しないでください。スクリプト中にfip活性/非活性コマンドを記述します。nfsd起動よりも前、nfsd停止よりも後に活性/非活性する必要があります。

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。

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。 CLUSTERPROのスクリプトで使用します。 例) 192.168.1.100
共有 (ミラー) ディスク上のパーティション (切替パーティション)	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fstabや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グループに複数の切替パーティションを登録することも可能です。 例) /dev/sdb2
マウントポイント	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。 CLUSTERPROのfstabやトレッキングツールでも指定します。 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/終了スクリプト (start.bat/stop.bat) にてnfsサービスの起動/停止も行います。 (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/終了スクリプト (start.bat/stop.bat) 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。

※ MiracleLinux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.
        action $"Starting NFS services: "##/usr/sbin/exportfs -r
        echo -n $"Starting NFS quotas: "
        daemon rpc.rquotad
        echo
        echo -n $"Starting NFS mountd: "
        daemon rpc.mountd $RPCMOUNTDOPTS
        echo
        echo -n $"Starting NFS daemon: "
        daemon rpc.nfsd $RPCNFSDCOUNT
        echo
        touch /var/lock/subsys/nfs
        ;;
    stop)
        # Stop daemons.
        echo -n $"Shutting down NFS mountd: "
        killproc rpc.mountd
        echo
        echo -n $"Shutting down NFS daemon: "
        killproc nfsd
        echo
        action $"Shutting down NFS services: "##/usr/sbin/exportfs -au
        echo -n $"Shutting down NFS quotas: "
        killproc rpc.rquotad
        echo
        rm -f /var/lock/subsys/nfs
        ;;
    ;;
    echo $"Usage: $0 {start | stop | status | restart | reload}"
    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.BAT *
#*****
```

```
if [ "$ARMS_EVENT" = "START" ]
then
    if [ "$ARMS_DISK" = "SUCCESS" ]
    then
        armlog "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)
        armfipcmd -a 192.168.1.100 >/dev/null 2>&1----(c)
        /bin/chmod 777 /mnt/sdb2 >/dev/null 2>&1----(d)

    if [ "$ARMS_SERVER" = "HOME" ]
    then
        armlog "NORMAL2"
    else
        armlog "ON_OTHER1"
    fi

    else
        armlog "ERROR_DISK from START"
    fi
elif [ "$ARMS_EVENT" = "RECOVER" ]
then
    armlog "RECOVER"
elif [ "$ARMS_EVENT" = "FAILOVER" ]
then
    if [ "$ARMS_DISK" = "SUCCESS" ]
    then
        armlog "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)
        armfipcmd -a 192.168.1.100 >/dev/null 2>&1----(c)
        /bin/chmod 777 /mnt/sdb2 >/dev/null 2>&1----(d)

    if [ "$ARMS_SERVER" = "HOME" ]
    then
        armlog "FAILOVER2"
    else
```

```

    armlog "ON_OTHER2"
    fi

    else
        armlog "ERROR_DISK from FAILOVER"
    fi
else
    armlog "NO_ARM"
fi
    armlog "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) fipを活性化させる。
- (d) 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.BAT *
#####

~~~中略~~~

if [ "$ARMS_EVENT" = "START" ]
then
    if [ "$ARMS_DISK" = "SUCCESS" ]
    then
        armlog "NORMAL1"

        armfipcmd -d 192.168.1.100 >/dev/null 2>&1----(e)
        /usr/sbin/exportfs -u /mnt/sdb2/ >/dev/null 2>&1----(f)
        /sbin/service nfs stop >/dev/null 2>&1-----(g)
        armsleep 30 >/dev/null 2>&1-----(h)

        if [ "$ARMS_SERVER" = "HOME" ]
        then
            armlog "NORMAL2"
        else
            armlog "ON_OTHER1"
        fi

    else
        armlog "ERROR_DISK from START"
    fi
elif [ "$ARMS_EVENT" = "FAILOVER" ]
then
    if [ "$ARMS_DISK" = "SUCCESS" ]
    then
        armlog "FAILOVER1"

        armfipcmd -d 192.168.1.100 >/dev/null 2>&1----(e)
        /usr/sbin/exportfs -u /mnt/sdb2/ >/dev/null 2>&1----(f)
        /sbin/service nfs stop >/dev/null 2>&1-----(g)
        armsleep 30 >/dev/null 2>&1-----(h)
    fi
fi

```

```

if [ "$ARMS_SERVER" = "HOME" ]
then
    armlog "FAILOVER2"
else
    armlog "ON_OTHER2"
fi
else
    armlog "ERROR_DISK from FAILOVER"
fi

else
    armlog "NO_ARM"
fi
    armlog "EXIT"
exit 0

```

＜終了スクリプトの説明＞

- (e) `fip`を非活性にさせる。
- (f) `/usr/sbin/exportfs`にて`export`していたmountpointを消す。
- (g) `/sbin/service` で`nfs`を停止させる。
- (h) `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,no_root_squash,insecure_locks`

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 ...`

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 (turbolinux server 8) 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

- `portmaper`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。
- `armsleep`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。フェイルオーバが完了し、NFSクライアントが`retry`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`armsleep`の時間より短い時間に設定してください。
- `fip`はフェイルオーバグループのリソースに登録しないでください。CLUSTERPRO上で制御されるスクリプトに記述します。
- 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),(d)を必要数分記述し、呼応して(f)を増やしてください。