

ADC からのお知らせ No.207 (2011/10/04)

■ NFS 領域のファイル情報不整合問題について

天文データセンター解析システムのユーザーの皆様

7月18日にNFS領域において、長時間にわたってファイル情報の不整合が発生したとの報告があり、調査を行ってまいりました。その現状および対処方法などについてご連絡いたします。

[現象]

NFS 領域にて、ファイル情報 (md5 や ctime など) に不整合が発生し、その状態が長時間継続していたことが報告されました。

具体的には、/lfs04c 領域にあるファイルに対して、ana09 と ana11 から "md5sum" や "ls -alc" を実行した場合に、返ってくる結果が異なるという現象が発生し、それが14分間ほど継続しました。

[現象についての解説]

NFS サーバ/クライアント環境において、特定のオプションを指定していない場合は、NFS 上のファイル情報は NFS クライアント上にあるキャッシュを参照して取得します。

そのため、キャッシュの更新タイミングによっては、最新のファイル情報を参照出来ず、NFS クライアント間でファイル情報の整合性が、短時間 (60 秒) 取れなくなる場合があります。

しかし、今回のように長時間にわたってファイル情報に不整合が発生する問題は、不具合として Red Hat コミュニティで認識されています。ただし、その発生自体が稀であり、現状では原因や再現方法が判明していません。

[対処方法]

上記のように再現方法が判明していないので、近いうちに解決方法が RedHat から提示されることは期待できません。したがって、ここでは同問題が発生した際の対処方法について説明いたします。

前述の特定のオプション ("noac" オプション) を指定することで、同問題を回避できることは確認していますが、それをしてしまうと、NFS 領域の read/write 性能が大幅に低下してしまうので、その対応はとりません。

現状の設定において同現象が発生した場合、回避策として touch コマンド等でディレクトリ情報を更新することで、キャッシュが最新の情報に更新されます。具体的には以下の通りです。

(1) 事象が発生しているサーバ上で、該当ディレクトリに任意の

ファイルを作成する。

(例)

```
% touch / 該当ディレクトリ /test
```

- (2) その後、"ls -l / 該当ディレクトリ " を実行すれば、
該当ディレクトリの情報を再度サーバへ問い合わせるため
キャッシュが更新され最新の情報になる。

ユーザの皆様にはご不便をおかけしますが、
ご理解ご協力をお願いいたします。

また、何か不具合等を感じましたら、consult[at]adc.nao.ac.jp
までご連絡ください ([at] を @ に変換)。