

Prezentacja systemu plików NFS

Jakub Jonik

15 stycznia 2003

1 Wprowadzenie

NFS (network file system) jest rozproszonym systemem plików opracowanym przez firmę **SUN Microsystems**. Został wprowadzony do użytku w roku 1985 (wersja 2.0). Następną wersję (3.0), dominującą na rynku, wprowadzono w roku 1994. Obecnie rozwijana jest czwarta wersja NFS.

NFS implementuje architekturę klient - serwer. Każda stacja robocza może działać zarówno jako klient i serwer. NFS dostępny na niemal wszystkie platformy programowe i sprzętowe. Umożliwia użytkownikom korzystanie z katalogów i plików znajdujących się fizycznie na różnych stacjach roboczych przyjmując abstrakcyjny model systemu plików i odwzorowując go w lokalny system plików, zależny od systemu operacyjnego. Każdy serwer eksportuje jeden lub więcej swoich katalogów, udostępniając je odległym klientom. Eksportowane są całe drzewa katalogów.

NFS posiada kilka elementów, które różnią go od innych systemów plików sieciowych. Po pierwsze korzysta z protokołu UDP (choć ostatnie wersje wykorzystuje także protokół TCP/IP). Stanowi to dużą przewagę ponieważ UDP jest protokołem nie wymagającym stałego połączenia, systemy plików oparte na NFS potrafią przetrwać awarię sieci bez problemów.

2 Cele projektowe systemu NFS

2.1 Przezroczystość

- **Przezroczystość dostępu** - klient NFS dostarcza interfejs dla aplikacji, umożliwiający przeprowadzanie operacji na zdalnych plikach identyczny z interfejsem dla lokalnego systemu plików. Oznacza, to że aplikacje działające na stacji klienckiej mogą wykonywać operacje na zdalnych plikach bez żadnych zmian w kodzie źródłowym.
- **Przezroczystość położenia** - Umożliwia klientowi NFS dostęp do zdalnych plików bez znajomości ich fizycznej lokalizacji. Klient NFS określa sieciową przestrzeń nazw plików dodając je do swojej lokalnej przestrzeni nazw. Miejsce montowania zdalnych plików w lokalnej hierarchii nazw jest wybierane przez klienta.
- **Przezroczystość awarii** - Ponieważ serwer NFS jest bezstanowy i większość operacji na plikach zdalnych jest powtarzalna sytuacje awaryjne dotyczące zdalnych plików są postrzegane przez klienta jako awarie lokalne. W przypadku wystąpienia awarii usługi świadczone przez serwer zostają zatrzymane do czasu ponownego uruchomienia serwera. Awaria procesu po stronie klienta nie wpływa na działanie żadnego z serwerów.
- **Przezroczystość wydajności** - Klient i serwer stosują buforowanie w celu poprawienia wydajności. Moduły klienta i serwera są instalowane w jądrze systemu UNIX.
- **Przezroczystość wędrówki** - W każdym kliencie zdalne pliki montowane są do lokalnego katalogu przez oddzielny proces. W przypadku gdy zmienia się fizyczne położenia plików należy uaktualnić tablice montowania u każdego klienta.

Przydatnym narzędziem służącym do montowania plików jest **automounter**. Jest to proces, który niewidocznie, na żądanie, montuje i odmontowuje systemy plików. Z punktu widzenia aplikacji działających po stronie klienta zdalny system plików jest cały czas dostępny.

2.2 Bezpieczeństwo

Architektura protokołu NFS umożliwia wykorzystanie wielu mechanizmów bezpieczeństwa takich jak:

- Przekazywanie systemu plików tylko do wybranych grup klientów.
- Udostępnianie systemu plików tylko w trybie do odczytu.
- Przypisywanie numeru identyfikacyjnego jednego użytkownika drugiemu.
- Wyłączenie logowania się jako *root* ze zdalnego komputera.
- Blokowanie plików (tylko wersja czwarta).

3 Implementacja NFS w systemie UNIX

3.1 Klient NFS

Klient NFS działa w jądrze systemu UNIX. Uzyskuje on dostęp do plików poprzez montowanie, czyli dołączenie zdalnego systemu plików do swojego lokalnego drzewa katalogów. Uzyskuje się to poprzez polecenie:

```
mount opcje serwer:katalog /punkt montowania
```

punkt montowania oznacza katalog lokalnego systemu plików, w którym widoczne mają być nazwy montowanych plików.

opcje pozwalają na:

- określenie rodzaju montowania: z prawem zapisu lub bez.
- określenie montowania jako *twardego* lub *miękkiego*. W przypadku montowania *miękkiego* klient po utraceniu łączności z serwerem podejmie kilka prób odzyskania połączenia po czym zgłosi błąd. W przypadku montowania *twardego* klient będzie próbował odzyskać połączenie aż do skutku.

Przykłady:

```
mount -t nfs -o ro,soft omega:home/jj189399 /home/jj189399
```

```
mount -t nfs -o rw,hard omega:home/jj189399 /home/jj189399
```

Do demontowania zasobów służy polecenie:

```
umount serwer:katalog
```

Uwagi dotyczące klienta NFS:

- klient może montować całe eksportowane przez serwery systemy plików lub ich części.
- punkt montowania może znajdować się na wcześniej zamontowanym katalogu z zewnętrznego systemu plików.
- w komendzie *mount* muszą być używane pełne ścieżki.
- nie można odmontować katalogu, w którym znajdują się otwarte pliki.
- pliki lokalne znajdujące się w katalogu będącym punktem montowania stają się niewidoczne.

3.1.1 Program automount

Przydatnym narzędziem jest, wspomniany już wcześniej, **automounter**. Jest to program który:

- W przypadku żądania dostępu do pliku fizycznie położonego na dysku innego komputera, automatycznie, w sposób niewidoczny dla użytkownika, montuje odpowiedni system plików.
- gdy użytkownik przez pewien czas (kilka minut) nie korzysta z pliku, automatycznie demontuje system plików.

Program *automount* uruchamia demona *automountd* dla każdego pliku opisanego w plikach konfiguracyjnych. Punkty montowania i montowane katalogi określa zestaw plików zwany *mapami automountera* .

3.1.2 Pamięć podręczna klienta

Po stronie klienta buforowane są zarówno bloki danych jak i atrybuty. Zwiększa to wydajność klienta NFS, jednak może prowadzić do niezgodności danych. Wykonanie operacji pisania po stronie klienta nie powoduje natychmiastowego uaktualnienia zmian w pliku znajdującym się na serwerze. Co jakiś czas proces klienta NFS sprawdza czy w pliku znajdującym się na serwerze zaszyły jakieś zmiany i jeżeli tak to unieważnia wszystkie bloki danych znajdujące się w pamięci podręcznej. Za uaktualnianie zmian z pamięci podręcznej klienta do plików znajdujących się na serwerze oraz za wyprzedzone czytanie plików z serwera odpowiada *biodemon* .

3.2 Serwer NFS

Moduł serwera, ze względów wydajnościowych, jest dołączony do jądra systemu UNIX. Jego głównym zadaniem jest udostępnianie (eksportowanie) klientom plików znajdujących się na serwerze. Listę udostępnionych zasobów administrator umieszcza w pliku konfiguracyjnym: */etc/exports* (BSD) lub */etc/dfs/dfstab* (SVR4). Katalogi wymienione w pliku konfiguracyjnym będą udostępnione klientom NFS po wykonaniu komendy: */usr/sbin/exportfs -a* Przeważnie podczas startu systemu uruchamiany jest program, który czyta plik konfiguracyjny, następnie przekazuje do jądra systemu operacyjnego informacje o prawach dostępu do drzew katalogów (zdefiniowanych w pliku konfiguracyjnym). Następnie uruchamiane są demony:

- *portmap* - dba o rejestrowanie usług RPC i zamianę numerów usług RPC na numery portów. Po uruchomieniu demona RPC nasłuchuje wszystkie porty, łączy się z programem *portmap* i rejestruje w nim numery portów. Kiedy program klienta wywołuje serwer RPC, otwiera połączenie z programem *portmap* na serwerze i podaje mu numer programu. *Portmap* odpowiada numerem portu, na którym serwer nasłuchuje, dalej klient może już bezpośrednio łączyć się z procesem na serwerze.
- *mountd* - odpowiada za montowanie i demontowanie zasobów. Zapewnia on następujące działania:
 - weryfikuje pliki */etc/host.allow* i *etc/hosts.decy* pod kątem zezwolenia na połączenie klienta z serwerem.
 - sprawdza, czy żądany system plików jest zapisany w pliku *etc/exports*
 - sprawdza, czy klient jest uprawniony do korzystania z serwera, a jeśli tak to w jakim trybie.
 - tworzy uchwyt do systemu plików i zwraca go klientowi.
 - dodaje klienta do pliku *etc/rmtab*.

- `nfsd` - Klient po zamontowaniu systemu plików kontaktuje się z demonem `nfsd` z takimi zadaniami jak otwieranie lub zamykanie plików, sprawdzanie ich statusu i typu oraz czytanie i zapisywanie danych.

Uwagi dotyczące serwera NFS:

- w przypadku gdy drzewo katalogów zawiera elementy zamontowane z innego nośnika, to muszą być one osobno wymienione w pliku konfiguracyjnym.
- nie można eksportować zamontowanego z innego serwera systemu plików.

3.2.1 Pamięć podręczna serwera

Serwery NFS korzystają z podręcznej pamięci buforowanej. Zapewniają czytanie z wyprzedzeniem, ale operacje zapisu są wykonywane natychmiast ponieważ ewentualna awaria serwera mogłaby spowodować utratę danych, która byłaby przez klientów niezauważona.

3.2.2 Ograniczenia

- Serwer może eksportować wyłącznie lokalne systemy plików.
- Eksportowania zasobów lokalnych i montowania zasobów zewnętrznych może dokonywać jedynie administrator.

3.3 Protokół NFS

Protokół NFS działa w oparciu o RPC (*Remote Procedure Call* - zdalne wywołanie procedur). RPC jest usługą, dzięki której działający na komputerze lokalnym program klienta może wywoływać kod, który jest uruchamiany na serwerze w taki sam sposób jakby był uruchamiany lokalnie. Interfejs RPC dostarczany przez serwer NFS pozwala na wykonywanie standardowych operacji na plikach (m.in. zapisywanie, odczytywanie, zmiana nazwy).

Jest jednak problem, który pojawia się, gdy procedury są wywoływane przez komputery z różną architekturą. Ponieważ RPC jest zaprojektowany do transferu struktur danych, elementy takie jak kolejność bajtów (które są różne dla różnych architektur) mają znaczenie dla danych przechowywanych w pamięci i mogą mieć wpływ w przypadku ich rozesłania w sieci. W związku z tym został stworzony standard nazwany *XDR*. Wszystkie dane przesyłane w sieci są najpierw zamieniane na format *XDR* na komputerze lokalnym, a następnie ponownie zamieniane na format wewnętrzny przez komputer odbierający dane.