

ファイルサーバ向け仮想化機能の設計と実装 (3)

亀井 仁志[†] 松沢 敬一[†] 揚妻 匡邦[†] 中野 隆裕[†]

(株) 日立製作所 システム開発研究所

1. はじめに

HA(High Availability)クラスタは、正常なサーバが障害で停止したサーバのサービスを引き継ぐことによって、サービスの可用性を向上させる技術である。複数のサーバを用いて、障害で停止したことを相互に監視し合うクラスタを構成する。

ファイルサーバにおいても、HA クラスタ技術を用いた高可用性が可能である。このようなファイルサーバでは、障害でファイルサーバが停止しても、障害で停止したファイルサービスと同じホスト名や共有名を別の正常なファイルサーバに引継ぎ、ユーザの業務継続を可能にする。

一方、近年、運用コスト削減や消費電力削減を目的に、複数のファイルサーバを一台へ集約する、ファイルサーバ集約が提案されている。集約する個々のファイルサーバのファイルサービスを 1 つのファイルサーバへ移行する。その際、HA クラスタ技術を備えるファイルサーバへ移行することによって、高可用ファイルサーバへの集約を可能とする。しかし、集約対象のファイルサーバ群が、アカウント情報を、passwd ファイルを用いて管理したり、NIS を用いて管理するなど、アカウント情報の管理方式が異なる場合がある。アカウント管理方式が異なるファイルサーバを集約する場合は、アカウント管理方式を変更する必要があり、ファイルサーバ集約が困難になるケースがあった。

本稿では、仮想化方式 VNAS(a Virtualization method for NAS) [1]を用いて、アカウント管理方式が異なるファイルサーバを集約できるクラスタ構成方式(VNAS HA 方式)を提案する。

2. ファイルサーバの高可用性方法

本章は、HA クラスタ技術の高可用性方法と動作例、ファイルサーバ集約における問題を述べる。

2.1 ファイルサーバ高可用性の概要

HA クラスタ技術を用いたファイルサーバの高可用性方式(HA クラスタ方式)では、クラスタを構成するファイルサーバ間で、ファイルサービスの設定情報を、ファイルサービス引継ぎのために共有する。一般に、HA クラスタ方式では、共有ディスクを用いて、ファイルサービス設定情報を共有する。また、ファイルサービス設定情報をリソースグループと呼ぶ。リソースグループには主に以下

の情報が設定される。

- (1) ホスト名
- (2) 利用可能なディスク情報
- (3) ファイルサーバプログラム設定情報
- (4) ファイルシステム公開ポイント情報

2.2 HA クラスタ方式の動作例

2 台のファイルサーバを用いた HA クラスタ方式の構成例を図 1 に示し、その動作を述べる。

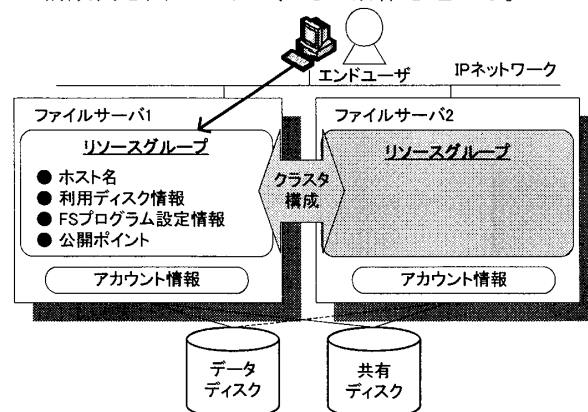


図 1 HA クラスタ方式の構成

図 1 の構成では、リソースグループによって定義されるファイルサービスがファイルサーバ 1 で動作し、エンドユーザのデータをデータディスクへ保存する。また、リソースグループによって定義された設定情報を共有ディスクへ保存している。

ファイルサーバ 1 が障害によって停止した場合、ファイルサーバ 2 がリソースグループによって定義された設定情報を共有ディスクから読み出し、ファイルサービスを動作させることにより、ファイルサービスを引き継ぐ。

2.3 ファイルサーバ集約と高可用性の問題

HA クラスタ方式を用いて高可用性したファイルサーバに対してファイルサービスを集約する場合、各ファイルサーバのファイルサービスを、リソースグループによって定義されるファイルサービスへ移行する方法がある。

しかし、アカウント管理方式は OS 毎に 1 つしか設定できないため、複数の集約対象ファイルサーバのアカウント管理方式が異なる場合、アカウント管理方式を変更する必要があった。

例えば、アカウント情報を、passwd ファイルを用いて管理しているファイルサーバと NIS を用いて管理しているファイルサーバのファイルサービスを集約する場合、passwd ファイルを用いて管理しているファイルサーバのアカウント情報を NIS

A Design and Implementation of the Linux Virtualization for File Server

[†] Systems Development Laboratory, Hitachi, Ltd.

へ登録するといった作業が必要となる。

3. VNAS HA 方式の提案

本章は、VNAS を用いたファイルサーバの高可用化方式(VNAS HA 方式)について述べる。VNAS HA 方式の概要と実現方式、構成と動作を述べる。

3.1 VNAS HA 方式の概要

VNAS HA 方式の前提である VNAS は、ファイルサーバ上で 1 つ動作可能な物理サーバ(Psv)と、Psv 上で複数動作可能な仮想サーバ(Vsv)から成る。

Vsv はエンドユーザへファイルサービスを提供する。アカウント管理設定情報を持つルートファイルシステム(ルート FS)を Vsv 毎に持ち、個々にアカウント管理方式が設定される。VNAS HA 方式では、集約対象ファイルサーバ群のファイルサービスを、個別の Vsv へ移行することによって、アカウント管理方式が異なる複数のファイルサーバを集約可能とする。Psv は装置全体の管理者によって操作され、Vsv の起動・停止やクラスタリングなどを制御する。

VNAS HA 方式の高可用化では、HA クラスタ方式と同様に、複数のファイルサーバを用いてクラスタを構成する。正常なファイルサーバが、障害によって停止したファイルサーバの Vsv を起動することで、ファイルサービスを引継ぐ。

3.2 VNAS HA 方式の実現

VNAS では、Psv が Vsv のホスト名やディスク割当て、ファイルサービスを設定する。VNAS HA 方式では、これらの設定情報を共有ディスクへ保存する。サービス引継ぎ時に Vsv を起動する場合、Psv は Vsv の起動に必要な設定情報を共有ディスクから読み出し、Vsv はファイルサービスに必要な設定情報を共有ディスクから読み出す必要がある。しかし、VNAS はディスクを Vsv 毎に割当てることにより、Vsv 間の干渉を防止しており、Vsv が共有ディスクを持つことができなかつた。一方、Psv はクラスタリングを制御するため、ディスクを共有できる。

そこで、共有ディスクの情報を Vsv 制御情報(リソースグループの(1)(2))とサービス制御情報(リソースグループの(3)(4))に分割した。そして、Vsv 制御情報を共有ディスクへ保存し、サービス制御情報は、Vsv のルート FS へ保存した。サービス引継ぎ時、Psv は共有ディスクから Vsv 制御情報を読み出して Vsv を起動し、Vsv は自らのルート FS からサービス制御情報を読み出してファイルサービスを起動する。これにより、Vsv は共有ディスクへアクセスすることなく、ファイルサービスを起動できる。また、本方式により、共有ディスクへ保存する情報が Vsv 制御情報のみとなり、クラスタを構成するファイルサーバが共有する情報を削減でき、数十台といった多数のファイルサーバを

集約しても、共有ディスクの大きさを小さくできる。

3.3 VNAS HA 方式の構成と動作

VNAS HA 方式の構成例を図 2 に示し、動作を説明する。

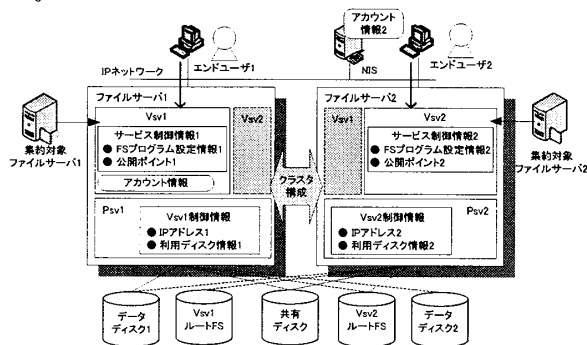


図 2 VNAS HA 方式の構成

図 2 では、集約対象ファイルサーバ 1 のファイルサービスをファイルサーバ 1 の Vsv1 へ移行し、集約対象ファイルサーバ 2 のファイルサービスをファイルサーバ 2 の Vsv2 へ移行している。

Vsv1 はエンドユーザ 1 のデータをデータディスク 1 へ保存する。Vsv2 はエンドユーザ 2 のデータをデータディスク 2 へ保存する。アカウント管理方式は、Vsv1 が passwd 管理を採用し、Vsv2 が NIS 管理を採用している。Vsv1 と Vsv2 のそれぞれにルート FS を割当て、サービス制御情報を保存する。また、Vsv1 と Vsv2 の Vsv 制御情報を共有ディスクへ保存してファイルサーバ間で共有している。

ファイルサーバ 1 が障害によって停止した場合、ファイルサーバ 2 の Psv2 が、共有ディスクから Vsv1 の Vsv 制御情報を読み出して Vsv1 を起動する。ファイルサーバ 2 で起動した Vsv1 は、ルート FS1 からサービス制御情報を読み出してファイルサービスを起動する。

4. おわりに

本稿は、アカウント管理方式が異なるファイルサーバ集約を可能とするクラスタ構成方式である VNAS HA 方式について述べた。VNAS HA 方式によって、クラスタを構成する物理計算機が共有する情報の削減も可能となった。

5. 参考文献

[1] 中野隆裕, "ファイルサーバ向け仮想化機能の設計と実装(1)", 情処 71 回全国大会.

Linux は、Linus Torvalds の米国およびその他の国における登録商標あるいは商標である。