

高信頼性を実現する資源管理機構とイベント管理機構*

1D-2

阿部 敏浩[†]、高田 実[‡]、三浦 博史[§]、中山 博之^{**}富士通（株）^{††}、（株）富山富士通^{‡‡}、（株）PFU^{§§}

abe@yk.fujitsu.co.jp

1. はじめに

オープンシステムの普及にともない、UNIX上に大規模な基幹業務システムを構築する事例が増加傾向にある。これにともない、大規模化するシステムの構築／運用の簡易化と、故障発生による業務停止時間を最小限にする高信頼性クラスタシステムが求められている。

従来のクラスタシステムでは、これまでの主流であった小中規模システム構成向けの導入／運用ツールで、システムを構成する膨大なハードウェア／ソフトウェア資源に対して1つずつ設定を行う必要があった。さらに、各資源の管理体系も統一化されていないため、環境設定作業が複雑になっていた。

また、故障検出時の業務引継ぎ処理においても、以下の問題により、業務切替え処理時のオーバーヘッドが大きかった。

- ・故障発生時に引継ぎ処理が記述されたプロシジャをシーケンシャルに起動していく方式であるため、並列動作することができない。
- ・システムに登録されたすべての引継ぎ処理を起動するため、故障に関連の無い資源まで引継ぎ処理を行っている。

これらの問題を解決するために、資源管理機構およびイベント管理機構を開発し、GRANPOWER 7000 クラスタシステムの基盤機能として提供している。本稿では、その機能概要と効果について述べる。

2. 資源管理機構とイベント管理機構

資源管理機構とイベント管理機構の開発目的は、クラスタシステムを構成する資源の管理方

式を統一し、クラスタシステム構築／運用を簡易化することと、業務引継ぎ処理のオーバーヘッドを削減し、業務再開までの時間を最小限にすることにある。

以下に、資源管理機構とイベント管理機構の概要について説明する。

2. 1 資源管理機構

資源管理機構は、ハードウェア／ソフトウェア資源に対して、資源登録時に指定された情報を基に資源を生成し、資源管理データベースに格納する。サブシステムは、資源管理APIを利用して資源の操作を行い、資源情報は、クラスタシステムを構成するノード間で一致化される。

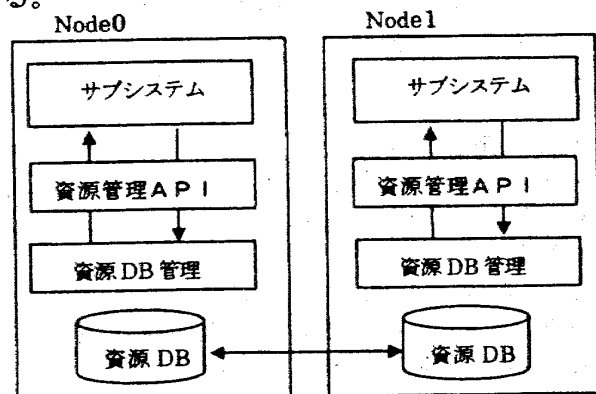


図1 資源管理機構

2. 2 資源管理データベース

資源管理データベースでは、図2で示す通り、個々の資源情報と資源間の依存関係を階層構造で管理している。

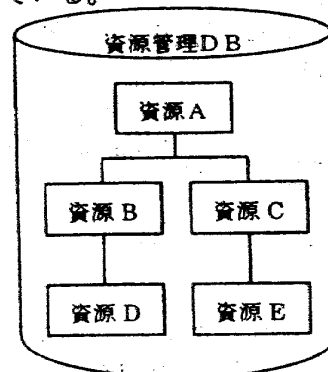


図2 資源DBによる資源の階層管理

* The Resource Management And Event Management
Realized Highly Available

[†] Toshihiro Abe

[‡] Minoru Takata

[§] Hiroshi Miura

^{**} Hiroyuki Nakayama

^{††} Fujitsu Ltd.

^{‡‡} Toyama Fujitsu Ltd.

^{§§} PFU Ltd.

資源管理データベースは、クラスタシステムの各ノード間で一致化を行い、一元性を保証するグローバルデータベースと、ノード固有の情報を格納するローカルデータベースを持つ。

2. 3 イベント管理機構

イベント管理機構は、リソースモニタ群とイベントレシーバ群がイベントの送受信を行うことによって、故障発生から業務再開までの一連の処理を効率的に行うための基盤機能である。

リソースモニタ

資源を監視し、状態変化を通知するイベントを送信するプロセス。

クラスタシステムの基本イベントとしては、各ノード、インタコネクト（ノード間通信路）の状態通知イベントなどがある。

イベントレシーバ

監視対象とする資源の資源名、属性、およびイベント発生時のアクションをあらかじめイベント管理に登録しておき、監視対象の資源の状態変化に応じた処理を行う。

また、分散サービスを利用し、ノード間のイベントレシーバ間で、資源の排他、同期、メッセージ送受信を行うことが可能である。

イベント通知サービス

イベント通知サービスでは、資源管理に登録された資源の状態変化を、イベント通知する機構を提供する。

このイベント通知サービスによって、各サブシステムはリアルタイムに監視対象資源の状態変化に関する事象を知ることができる。

図3に示す通り、資源の故障はリソースモニタからイベントレシーバに通知され、イベントレシーバは故障した資源を切り離し、利用可能な資源に切り替えることができる。

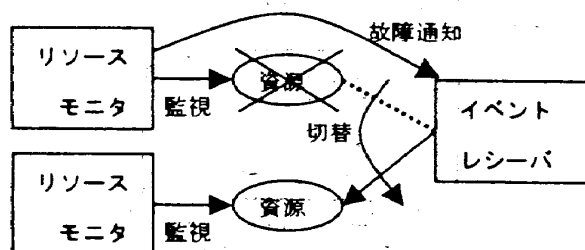


図3 イベント通知サービス

分散サービス

分散サービスは、クラスタシステムを構成するノード間における資源の排他処理、同期処理を行う機構を提供する。

3. 効果

GRANPOWER 7000 シリーズのクラスタシステムに資源管理機構とイベント管理機構を適用することにより、以下に述べる効果を得ることができた。

3. 1 システム構築／運用の簡易化

資源管理方式の統一化と、クラスタシステムを構成する個々の資源定義や資源間の関係を定義する階層をエンドユーザから隠蔽することにより、システム構築／運用が簡易化され、作業コストを低減することができた。

また、将来的に監視対象の資源を追加する場合も、資源管理APIを使用することにより、システムで統一された資源管理体系に、容易にアドインすることができる。

3. 2 業務引継ぎ時間の短縮

資源管理機構とイベント管理機構を利用することによって、以下に示すオーバーヘッドの削減を実現し、故障発生から業務再開までの時間を従来比の1/2～1/3に短縮することができた。

- ・事象検出／通知が高速に、かつ、各資源の引継ぎ処理が並行動作可能となり、故障通知に伴う業務引継ぎ処理が高速に行われる。
- ・引継ぎが必要な資源のみを自動的に判断して引き継ぐ。
- ・引継ぎ時間が長いノード切替えを行わずに、必要な資源のみ切替え、業務を続行。

4. 課題

将来的にクラスタシステムの多数台ノード化／地域分散化が予想されるため、今回開発した資源管理機構とイベント管理機構をベースに、以下の機能拡張を行う必要がある。

- ・現状のクラスタシステムでは構成ノード数に制限があるが、高速なノード間イベント通知機能及びノード間同期機能を利用し、数百といったオーダのノード構成でも短時間で業務引継ぎを可能とする。
- ・資源DB管理機能をベースに資源のグループ化機能を強化し、ノード間／クラスタシステム間通信で資源情報を交換することで、多数台ノード化／地域分散化したクラスタシステムを1つのシステムとして管理できるSSI (Single System Image) を実現し、さらなる操作性の向上を図る。