

# 最新クラスタシステム技術

恵下田 秀作・飼手 郁夫・神野 絵理

## 要旨

ユビキタス社会の実現に向けて、コンピュータシステムには、今まで以上の高い信頼性が求められています。その要求に対する1つの答えとして、クラスタ技術でサーバシステムの可用性を向上するソフトウェア CLUSTERPROがあります。本稿では最新バージョンであるCLUSTERPRO Xの機能・性能、また最新技術への取り組み、将来への展望について説明します。

## キーワード

●クラスタリング ●フェイルセーフ ●強制停止 ●フェイルオーバー ●高可用性 ●高信頼性 ●HA

## 1. まえがき

本格的なユビキタス社会の到来に向けて、IT基盤の重要性はますます高まり、今まで以上の高い信頼性、つまり「安心」が求められています。NECのITプラットフォームビジョン「REAL IT PLATFORM」において「『安心』を実現するテクノロジー」として、クラスタリングソフトウェア CLUSTERPROがあります。

CLUSTERPROは、1996年の初出荷以来、市場ニーズや最新技術を積極的に取り込み、システムの可用性を向上し続けることで、国内トップクラスの地位を確保しています。

本稿では、最新バージョンであるCLUSTERPRO Xの機能・性能、最新技術への取り組み、将来への展望について説明します。

## 2. CLUSTERPROの基本機能・構成

クラスタシステムは、2台以上の複数台のサーバで構成されます。CLUSTERPROを含むフェイルオーバー型クラスタソフトウェアの役割は、大別すれば以下の2つであり、これによってシステムの可用性を向上させ、フェイルセーフを実現しています。

- ・クラスタシステム内で発生する異常を検出する。
- ・業務サービスを異常が発生したサーバから健全なサーバへ移動・再開(フェイルオーバー)する。

CLUSTERPROはWindows、Linuxプラットフォームに対応し、共有データを保存するストレージの種別によって、共有ディスク型(図1)、データミラー型(図2)があります。

2006年10月、CLUSTERPROの最新バージョンである、CLUSTERPRO Xがリリースされました。以下、

CLUSTERPRO Xの新機能、強化ポイントを交えつつ、以下のトピックを追っていきます。

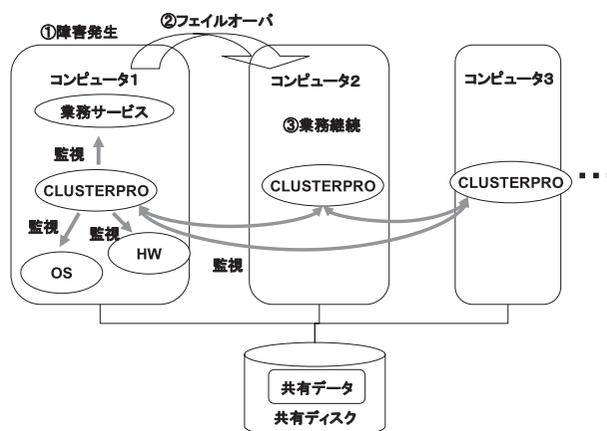


図1 共有ディスク型

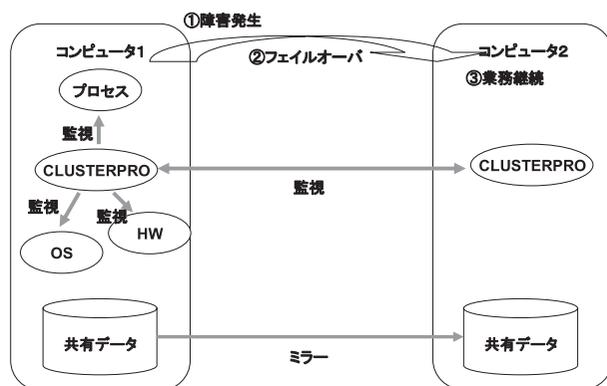


図2 データミラー型

- ・ 複雑多様化する構成
- ・ DR (Disaster Recovery: 災害対策) 領域へ
- ・ メインフレームレベルの高可用性
- ・ VMレベルの切り替え
- ・ プロセスレベルの切り替え
- ・ 効果的な運用の将来像

### 3. 複雑多様化する構成

CLUSTERPROの基本的な構成に共有ディスク型、データミラー型があることは述べましたが、CLUSTERPRO Xでは、データミラー型の拡張構成として、新たに複数のミラー元サーバに対して、ミラー先サーバを1台に集約するN:1のデータミラー型構成(図3)を可能にしました。

また、サイトAに共有ディスク型クラスタを、サイトBに待機系サーバを配置し、サイトAの共有データをサイトBへミラーする共有ディスク型とデータミラー型のミックス構成<sup>1</sup>なども追加し、複雑、多様化する構成へのニーズに応えるべく強化されています。

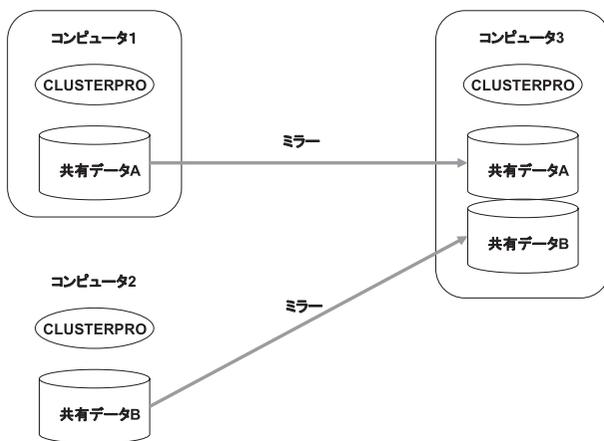


図3 N:1データミラー型

### 4. DR領域へ

CLUSTERPRO Xでは、データミラー型クラスタに関する強化を主に3つ行いました。

- 1) ミラー処理を独自プロトコルからTCP/IPへ変更
- 2) ミラー処理の高速化
- 3) 非同期I/O機能の搭載

プロトコルをTCP/IPにすることで、クラスタを構成する両サーバを異なるネットワークに配置することが可能になり、ミラー処理の高速化と非同期I/O機能の搭載によって、両サーバを遠隔地に配置することがより容易になりました。

ミラー処理の高速化では、ミラー対象領域の全面コピー処理に対するテストで1.5倍前後の高速化が実現されています(表)。

非同期I/O機能の実証実験として、東京-広島間をIP-VPNで接続し、計測を行った結果を図4に示します。このデータから、東京での書き込み量(Write KB/秒)を徐々に上げていくと、発生した書き込みを広島へ転送する通信量(Send KB/秒)は、約1,500KB/秒(12Mbps)までは書き込み量に対して十分な転送通信量が得られていることが分かります。また1,500KB/秒を超えた辺りから書き込み量に転送通信が追いつかなくなり、未転送のデータが履歴データとして蓄積され始めています。つまり、実証環境は1,500KB/秒、1日に換算して約13GB弱の

表 全面コピー所要時間

製品	所要時間	高速化率
CLUSTERPRO Ver8.0	0:16:38	—
CLUSTERPRO X Ver1.0 (同期I/O)	0:09:38	173%
CLUSTERPRO X Ver1.0 (非同期I/O)	0:11:34	144%

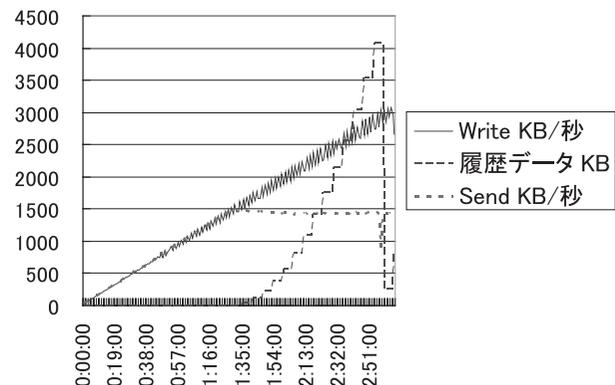


図4 非同期I/O性能 (東京-広島)

<sup>1</sup> 現在検討中の構成です。

更新系業務に耐え得る、という結果が得られました。

また、北米では2004年のCeBITで遠隔地に配置されたftServerどうしをCLUSTERPROでクラスタ化する展示で“Best of CeBIT America 2004”を受賞したことを皮切りに、着々とDR領域での実績を上げています。

## 5. メインフレームレベルの高可用性

CLUSTERPROは1996年の出荷以来、NECがメインフレームで培ったDNAをもとにクラスタソフトとしての機能を充実させています。特に異常状態に陥ったサーバによるデータ破壊を防ぐ観点から、異常が発生したサーバをいかに速やかに停止させるか、という課題に対しては、今なお継続的に取り組んでいます。

現在、CLUSTERPRO Xが備えているサーバの強制停止機能には、以下があります。

- ・ ウォッチドッグドライバによる強制停止、パニック、リセット
- ・ BMC (ベースボードマネジメントコントローラ) に搭載されているハードウェア (H/W) ウォッチドッグタイマと連動した強制停止、リセット

これらは、「異常が発生したサーバに自身を強制停止させる方式」といえます。これに加えて、以下のような「異常が発生したサーバを外部のサーバに強制停止させる方式」を今後サポートする予定です。

- ・ RDMAを利用した強制停止、パニック、リセット
- ・ BIOSをリモート制御することによる強制停止、リセット

## 6. VMレベルの切り替え

VMware社のVMware Infrastructure 3については、企業ユーザでの活用が広がってきており、最近ではXenを始めとするOSSのVMM (Virtual Machine Monitor: 仮想マシンモニタ) も注目を集めるなど、VM (仮想マシン) 技術はエンタープライズ領域で実用的なものとなりつつあり、導入事例も多数生まれています。

CLUSTERPRO XではVMware Infrastructure 3と組み合わせることで、多様な構成を可能としますが、以下に代表的な例を2つ挙げます。

### (1) VM-VMクラスタ

VMどうしでクラスタを構成します (図5)。一般的なクラスタ構成をそのまま仮想環境へ適用する場合に使用されています。H/W保守の期限切れ等に起因する、既存システムの維持 (レガシーマイグレーション) を目的として使用されることもあります。

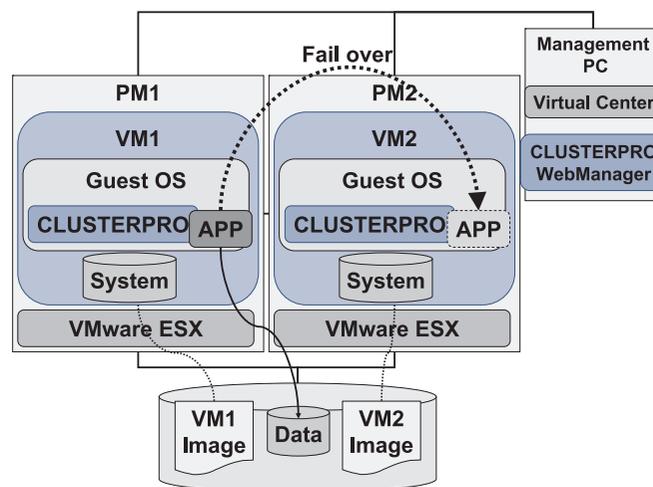


図5 VM-VMクラスタ (共有ディスク型)

### (2) PM-PMクラスタ

VMware ESX のインストールされたPM (Physical Machine: 物理マシン) どうしをクラスタ化し、PM間をVMが移動・

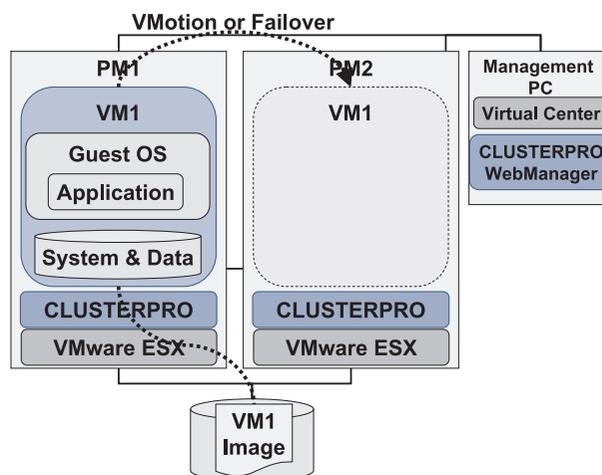


図6 PM-PMクラスタ

フェイルオーバーする構成です（図6）。VMそのものに対して特別な設定をしなくても可用性が向上できる利点があります。その反面、VMがクラスタ化されていないため、VM上で実行される業務サービスの異常検出や、VMのメンテナンス<sup>2</sup>が課題になります。

## 7. プロセスレベルの切り替え

前述のとおりCLUSTERPRO XはVMレベルの切り替えを可能にしていますが、我々は現在、「切り替える対象の粒度」をさらに細かくとらえ、プロセスレベルの切り替えを検討しています。

サーバAで動作しているプロセスをサスペンド状態にし、そのメモリイメージをサーバBへ転送、レジュームすることで、プロセスの状態をそのまま移動させる技術はプロトタイプレベルではすでに実現されています。

将来、リソースの最適配置技術と組み合わせることで、リソースの逼迫しているサーバから、余裕のあるサーバへユーザに気づかせることなく、かつ、シームレスにプロセスを移動させ全体最適を図るようなシステムをめざしています。

## 8. 効果的な運用の将来像

すでにVM技術の連携が実現されていることは述べましたが、このように現在のCLUSTERPRO XはブレードサーバやSSC（Sigma System Center：社製統合管理ツール）と連携した効果的な運用を実現する仕組みを実現しつつあります。

今後SSCのリソース最適配置機能やH/Wプロビジョニング機能と連携することで以下の効果が得られるようになることをめざしています（図7）。

- ・ リソース最適配置機能から得た情報をもとに、CLUSTERPRO Xが業務サービスを適切なサーバへ自動的に移動させて全体最適を図る。
- ・ CLUSTERPRO Xが異常を検出し、クラスタから切り離れたサーバを、H/Wプロビジョニング機能が自動的に交換・修復する。

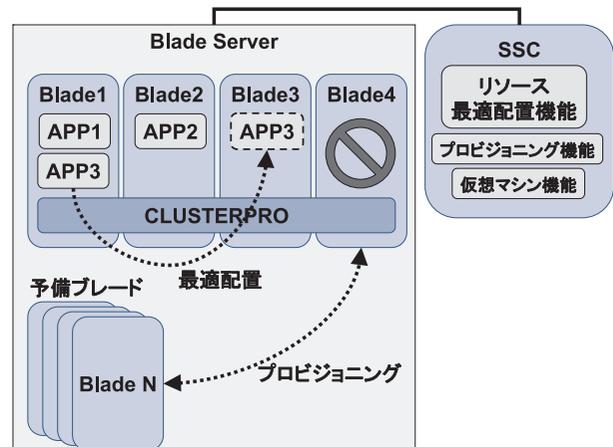


図7 ブレードサーバ/SSC連携

## 9. むすび

CLUSTERPROは、1996年のリリース以来、最新テクノロジーをいち早く採用、あるいは連携することで、常にPCサーバクラスタ業界をリードしてきました。

今後も新たな技術の採用、連携を積み重ね、幅広いお客様に「『安心』を実現するテクノロジー」をお届けし、「REAL IT PLATFORM」のビジョンを実現するべく、研究開発によりいっそうの努力をまいります。

\*本稿に記載している会社名、製品名は、各社の商標または登録商標です。

## 執筆者プロフィール

恵下田 秀作  
コンピュータソフトウェア事業本部  
第二コンピュータソフトウェア事業部  
グループマネージャー

飼手 郁夫  
コンピュータソフトウェア事業本部  
第二コンピュータソフトウェア事業部  
マネージャー

神野 絵理  
NECシステムテクノロジー  
プラットフォーム事業本部  
ネットワークソフトウェア事業部

<sup>2</sup> 現用系VMでサービスを継続させたまま、待機系VMにパッチをあてるといったローリングメンテナンスができず、業務停止となるパターンがあります。