CLUSTERPRO® システム構築ガイド

CLUSTERPRO® for Windows Ver7.0

入門編

第1版　2003.4.28

改版履歴

版数	改版年月日	改版ページ	内　　容
第1版	2003. 4.28		新規作成

はじめに

『CLUSTERPROシステム構築ガイド』は、これからクラスタシステムを設計・導入しようとしているシステムエンジニアや、すでに導入されているクラスタシステムの保守・運用管理を行う管理者や保守員の方を対象にしています。

補足情報

【OSのアップグレードについて】

クラスタサーバのOSをアップグレードする場合、手順を誤ると予期せぬタイミングでフェイルオーバが発生したり、最悪の場合、システムにダメージを与える可能性があります。

必ず製品添付のセットアップカードの手順に沿ってOSをアップグレードしてください。

また、サービスパックの適用も上記に準じます。

CLUSTERPRO® Ver 7.0 FastSync™ Option対応について

CLUSTERPRO® Ver 7.0 FastSync™ Option （以下FastSync Optionと省略）は、CLUSTERPRO Ver 7.0 LEに対応しています。

CLUSTERPROドキュメント体系

CLUSTERPROのドキュメントは、CLUSTERPROをご利用になる局面や読者に応じて以下の通り分冊しています。初めてクラスタシステムを設計する場合は、システム構築ガイド【入門編】を最初にお読みください。

■ セットアップカード　(必須) ~~設計~~・構築・~~運用~~・~~保守~~

製品添付の資料で、製品構成や動作環境などについて記載しています。

■ システム構築ガイド

【入門編】　(必須)設計・~~構築~~・~~運用~~・~~保守~~

クラスタシステムをはじめて設計・構築する方を対象にした入門書です。

【システム設計編(基本/共有ディスク,ミラーディスク)】　(必須)設計・構築・~~運用~~・~~保守~~

クラスタシステムを設計・構築を行う上でほとんどのシステムで必要となる事項をまとめたノウハウ集です。構築前に知っておくべき情報、構築にあたっての注意事項などを説明しています。システム構成が共有ディスクシステムかミラーディスクシステムかで分冊しています。

【システム設計編(応用)】　(選択)設計・構築・~~運用~~・~~保守~~

設計編(基本)で触れなかった CLUSTERPRO のより高度な機能を使用する場合に必要となる事項をまとめたノウハウ集です。

【クラスタ生成ガイド(共有ディスク,ミラーディスク)】　(必須) ~~設計~~・構築・~~運用~~・~~保守~~

CLUSTERPRO のインストール後に行う環境設定を実際の作業手順に沿って分かりやすく説明しています。システム構成が共有ディスクシステムかミラーディスクシステムかで分冊しています。

【運用/保守編】　(必須) ~~設計~~・~~構築~~・運用・保守

クラスタシステムの運用を行う上で必要な知識と、障害発生時の対処方法やエラー一覧をまとめたドキュメントです。

【GUI リファレンス】　(必須) ~~設計~~・~~構築~~・運用・~~保守~~

クラスタシステムの運用を行う上で必要な CLUSTERPRO マネージャなどの操作方法をまとめたリファレンスです。

【コマンドリファレンス】　(選択)設計・構築・運用・~~保守~~

CLUSTERPRO のスクリプトに記述できるコマンドやサーバまたはクライアントのコマンドプロンプトから実行できる運用管理コマンドについてのリファレンスです。

【API リファレンス】　(選択)設計・~~構築~~・~~運用~~・~~保守~~

CLUSTERPRO が提供する API を利用してクラスタシステムと連携したアプリケーションを作成する場合にお使いいただくリファレンスです。

【PP 編】　(選択必須)設計・構築・~~運用~~・~~保守~~

この編に記載されている各 PP は、CLUSTERPRO と連携して動作することができます。
各 PP が、CLUSTERPRO と連携する場合に必要な設定や、スクリプトの記述方法、注意事項などについて説明しています。使用する PP については必ずお読みください。

【注意制限事項集】　(選択)設計・構築・運用・保守

クラスタシステム構築時、運用時、異常動作等障害対応時に注意しなければならない事項を記載したリファレンスです。必要に応じてお読み下さい。

目　次

1 CLUSTERPROの概要

ここ数年来、コストパフォーマンスなどの理由によりWindowsプラットフォームの適用範囲が広がり、システムは複雑化・拡大化の傾向にあります。このため、Windowsサーバシステムにも可用性や拡張性がより一層強く求められ、脚光を浴びているのがクラスタシステムです。

CLUSTERPROは、クラスタシステムを支えるミドルウェアであり、低価格で高可用なシステムから大規模で可用性・拡張性の高いシステムを構築できる幅広い製品を提供しています。

1.1 クラスタシステム導入の効果

クラスタシステムとは、

複数のサーバを協調動作させ、一台のサーバでは達成できない、高い可用性と拡張性を提供するシステムを指し、CLUSTERPROによるクラスタシステムの導入により、次の二つの効果を得られます。

* 高可用性

クラスタを構成するサーバのうち一台が障害などにより停止しても、そのサーバが処理していた業務を他の健全なサーバへ自動的に引き継ぐことにより、障害時の業務停止時間を最小限に抑えます。

* 高拡張性

最大16台までの、パラレルデータベースをサポートすることにより、拡張性の高い高性能なデータベースプラットホームを提供します。

1.2 障害監視とフェイルオーバ

クラスタシステムはサーバ内で発生する種々の障害を監視し、障害発生時に業務を他サーバに移動(フェイルオーバ)します。

1.2.1 障害監視のしくみ

(1) サーバ監視

サーバ監視とはクラスタシステムの最も基本的な監視機能で、クラスタを構成するサーバが停止していないかを監視する機能です。

CLUSTERPROはサーバ監視のために、定期的にサーバ同士で生存確認を行います。この生存確認をハートビートと呼びます。ハートビートは次の四つの通信パスを使用して行います。

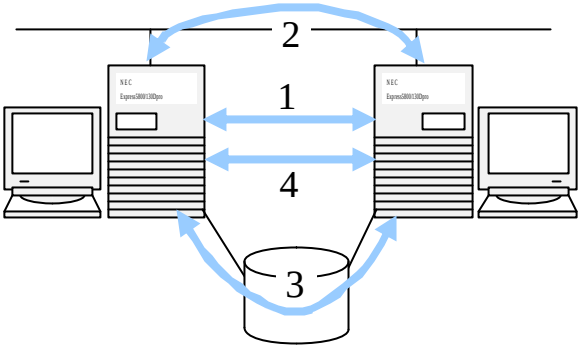
＊　インタコネクト専用LAN

クラスタシステム専用の通信パスで、一般のEthernet NICを使用します。ハートビートを行うと同時にサーバ間の情報交換に使用します。

＊　パブリックLAN

クライアントと接続している通信パスを予備のインタコネクトとして使用します。TCP/IPが使用できるNICであればどのようなものでも構いません。

インタコネクト専用LANの異常時には、サーバ間の情報交換にも使用します。



1　インタコネクト専用LAN
2　パブリックLAN
3　共有ディスク
4　COMポート

＊　共有ディスク[1]

クラスタを構成する全てのサーバに接続されたディスク上に、CLUSTERPRO専用のパーティション(CLUSTERパーティション)を作成し、CLUSTERパーティション上でハートビートを行います。

他サーバに障害が発生した場合にこの経路による最終確認を行うことによってネットワークパーティション症状によるデータ破壊を防ぎます。

＊　COMポート

フェイルオーバ型クラスタを構成するサーバ間を、COMポートを介して通信を行い、他サーバの生存を確認します。ここでの通信は、ネットワークパーティション症状の解決に用います。

ネットワークパーティション症状(Split-brain-syndrome)とは

クラスタサーバ間の全ての通信路に障害が発生しネットワーク的に分断されてしまう状態のこと。
ネットワークパーティション症状に対応できていないクラスタシステムでは、通信路の障害とサーバの障害を区別できず、同一資源を複数のサーバからアクセスしデータ破壊を引き起こす場合があります。詳細については「1.2.4 ネットワークパーティション解決」を参照してください。

これら四つの通信経路を使用することでサーバ間の通信の信頼性は飛躍的に向上し、ネットワークパーティション症状の発生を防ぎます。

(2)　業務監視

業務監視とは業務アプリケーションそのものや業務が実行できない状態に陥る障害要因を監視する機能です。

＊　アプリケーション・サービスの死活監視

CLUSTERPROのARMLOAD[2]コマンドによりアプリケーションやサービスを起動し、定期的にプロセスの生存またはサービスの状態を確認することで実現します。業務停止要因が業務アプリケーションの異常終了である場合に有効です。

---

[1] VERITAS Volume Managerを使用する場合は共有ディスクを使用したサーバ監視は出来ません。
[2] ARMLOADコマンドの詳細については「CLUSTERPROシステム構築ガイド コマンドリファレンス」を参照してください。

<注意>
CLUSTERPROが直接起動したアプリケーション/サービスのみが監視対象です。
他の常駐プロセスを起動し終了してしまうようなアプリケーションでは、常駐プロセスの異常を検出することはできません。
APの内部状態の異常は監視できません。
アプリケーションのストールや結果異常を検出することはできません。

* **リソース監視**

CLUSTERPROの“リソース監視[3]”のリソースにより、クラスタ資源(ディスクパーティション、IPアドレスなど)やパブリックLANの状態を監視することで実現します。業務停止要因が業務に必要な資源の異常である場合に有効です。

(3) 内部監視

CLUSTERPRO内部のモジュール間相互監視です。CLUSTERPROの各監視機能が正常に動作していることを監視します。

次のような監視をCLUSTERPRO内部で行っています。

* CLUSTERPROサービスとサービス監視プロセスとの相互監視
* 各種ハートビートスレッドのストール監視

1.2.2 監視できる障害と監視できない障害

(1) サーバ監視

監視条件: 障害サーバからのハートビートが途絶

* 監視できる障害の例
  + ハードウェア障害(OSが継続動作できないもの)
  + STOPエラー

* 監視できない障害の例
  + OSの部分的な機能障害(マウス/キーボードのみが動作しない等)

(2) 業務監視

監視条件: 障害アプリケーションの消滅、継続的なリソース異常、あるネットワーク装置への通信路切断

* 監視できる障害の例
  + アプリケーションの異常終了
  + 共有ディスクへのアクセス障害(HBA[4]の故障など)
  + パブリックLAN NICの故障

* 監視できない障害の例
  + アプリケーションのストール/結果異常

---

[3] リソース監視については、「CLUSTERPROシステム構築ガイド システム設計編(応用)」「CLUSTERPROシステム構築ガイド GUIリファレンス」を参照してください。

[4] Host Bus Adapterの略で、共有ディスク側ではなく、サーバ本体側のアダプタのことです。

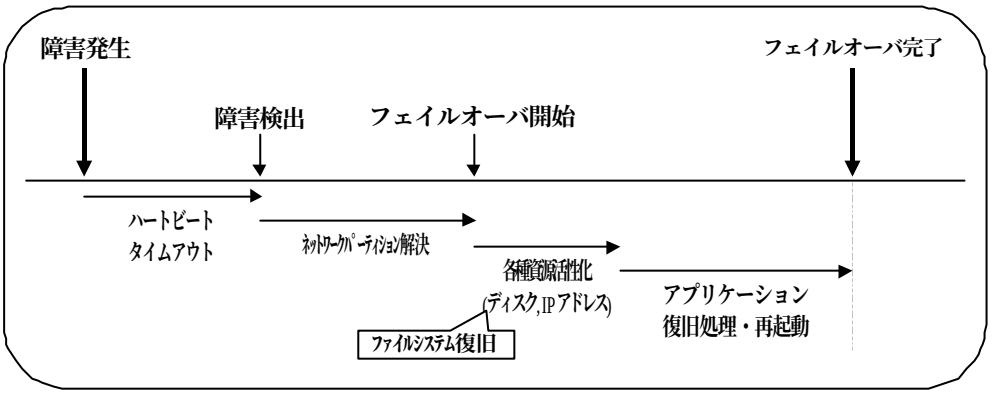
1.2.3 フェイルオーバのしくみ

CLUSTERPROでは、フェイルオーバ開始前に、検出した障害がサーバの障害かネットワークパーティション症状かを判別します。この後、健全なサーバ上で各種資源を活性化し業務アプリケーションを起動することでフェイルオーバを実行します。

このとき、同時に移動する資源の集まりをフェイルオーバグループと呼びます。フェイルオーバグループは利用者から見た場合、仮想的なコンピュータとみなすことができます。

> <注意>
> クラスタシステムでは、アプリケーションを健全なサーバで起動しなおすことでフェイルオーバを実行します。このため、アプリケーションのメモリ上に格納されている実行状態をフェイルオーバすることはできません。

障害発生からフェイルオーバ完了までの時間は数分間必要です。以下にタイムチャートを示します。



* ハートビートタイムアウト
  + 業務を実行しているサーバの障害発生後、待機系がその障害を検出するまでの時間です。
  + 業務の負荷に応じてクラスタプロパティの設定値を調整します。
    (出荷時設定では30秒に設定されています。)

* ネットワークパーティション解決
  + ネットワークパーティション方式として共有ディスク方式が指定されている場合に、検出した相手サーバの障害が、ネットワークパーティション症状によるものか実際に相手サーバが障害を起こしたのかを確認するための時間です。
  + CLUSTERパーティションへのアクセス時間や、ハートビートタイムアウト値などに連動して必要な時間が変化します。
    (出荷時設定では30秒以上60秒以下で解決するように設定されています。)

* 各種資源活性化
  + 業務で必要な資源を活性化するための時間です。
  + 一般的な設定では数秒で活性化しますが、フェイルオーバグループに登録されている資源の種類や数によって必要時間は変化します。
    (詳しくは、「システム構築ガイド システム設計編(基本)」を参照してください。)

* 開始スクリプト実行時間
  + データベースのロールバック/ロールフォワードなどのデータ復旧時間と業務で使用するアプリケーションまたはサービスの起動時間です。
  + ロールバック/ロールフォワード時間などはチェックポイントインターバルの調整である程度予測可能です。詳しくは、各PPのドキュメントを参照してください。

1.2.4 ネットワークパーティション解決

ネットワークパーティション症状(Split Brain syndrome)とはクラスタサーバ間の全ての通信路に障害が発生しネットワーク的に分断されてしまう状態のことです。

ネットワークパーティション症状に対応できていないクラスタシステムでは、通信路の障害とサーバの障害を区別できず、同一資源を複数のサーバからアクセスしデータ破壊を引き起こす場合があります。CLUSTERPROでは、他サーバからのハートビート切れを検出すると、サーバの障害かネットワークパーティション症状かを判別します。サーバダウンと判定した場合は、健全なサーバ上で各種資源を活性化し業務アプリケーションを起動することでフェイルオーバを実行します。ネットワークパーティション症状と判定した場合には、業務継続よりデータ保護を優先させるため、緊急シャットダウンなどの処理を実施します。

ネットワークパーティション解決方式には下記の方法があります。

* COM方式
  + 2ノードクラスタで使用できます。
  + シリアルクロスケーブルが必要です。
  + COM通信路を使用して相手サーバの生存確認を行うことによってネットワークパーティション症状の判定を行います。
  + COM通信路(COMポートやシリアルクロスケーブル)に異常が発生している状態でサーバダウンが発生した場合は、ネットワークパーティションの解決が失敗するため、フェイルオーバできません。正常なサーバも緊急シャットダウンします。
  + COM通信路が正常な状態で全てのネットワーク通信路に障害が発生した場合は、ネットワークパーティションを検出して、最高プライオリティサーバを除いた全てのサーバが緊急シャットダウンします。
  + COM通信路(COMポートやシリアルクロスケーブル)に異常が発生している状態で全てのネットワーク通信路に障害が発生した場合は、全てのサーバが緊急シャットダウンします。
  + 万一、クラスタサーバ間の全てのネットワーク通信路とCOM通信路に同時に障害が発生した場合には、両サーバがフェイルオーバを実行します。この場合は同一資源を複数のサーバからアクセスしてデータ破壊を引き起こす場合があります。

* 共有ディスク方式[5]
  + 共有ディスクを使用するクラスタで使用できます。
  + 共有ディスク上に専用のディスクパーティション(CLUSTERパーティション)が必要です。
  + 共有ディスク上に定期的にデータを書き込み、相手サーバの最終生存時刻を計算することでネットワークパーティション症状の判定を行います。
  + 共有ディスクや共有ディスクへの経路(SCSIバスなど)に異常が発生している状態でサーバダウンが発生した場合は、ネットワークパーティションの解決が失敗するため、フェイルオーバできません。正常なサーバも緊急シャットダウンします。

[5] VERITAS Volume Managerを使用する場合、共有ディスク方式は使用できません。

+ 共有ディスクが正常な状態で全てのネットワーク通信路に障害が発生した場合は、ネットワークパーティションを検出して、最高プライオリティサーバ及び最高プライオリティサーバと通信できるサーバがフェイルオーバ処理を実施します。それ以外のサーバは全て緊急シャットダウンします。
+ 共有ディスクや共有ディスクへの経路(SCSIバスなど)に異常が発生している状態で全てのネットワーク通信路に障害が発生した場合は、全てのサーバが緊急シャットダウンします。
+ 共有ディスクへのIO時間が指定したディスクタイムアウト時間より長くかかる場合にはネットワークパーティション解決処理がタイムアウトしてフェイルオーバできないことがあります。
+ ネットワークパーティション解決のためにハートビートタイムアウトの2倍とディスクIO待ち時間の2倍のうち、長い方の時間が必要となります。

* COM+共有ディスク方式[6]
  + 2ノードで共有ディスクを使用するクラスタで使用できます。
  + シリアルクロスケーブルが必要です。
  + 共有ディスク上に専用のディスクパーティション(CLUSTERパーティション)が必要です。
  + COM通信路(COMポートやシリアルクロスケーブル)に異常が発生している状態でサーバダウンが発生した場合は、共有ディスク方式に切り替わり、ネットワークパーティションの解決を行います。
  + COM通信路が正常な状態で全てのネットワーク通信路に障害が発生した場合は、ネットワークパーティションを検出して、最高プライオリティサーバを除いた全てのサーバが緊急シャットダウンします。
  + COM通信路(COMポートやシリアルクロスケーブル)に異常が発生している状態で全てのネットワーク通信路に障害が発生した場合は、共有ディスク方式に切り替わり、ネットワークパーティションの解決を行います。
  + 万一、クラスタサーバ間の全てのネットワーク通信路とCOM通信路に同時に障害が発生した場合にも、少なくとも一方のサーバが緊急シャットダウンを行いますので、データ破壊を避けることができます。

* 多数決方式
  + 3ノード以上のクラスタで使用できます。
  + ネットワーク障害によって過半数のサーバと通信できなくなったサーバがダウンすることによってネットワークパーティション症状によるデータ破壊を防ぎます。
  + 半数以上のサーバがダウンした場合は、残りの全ての正常サーバもダウンします。
  + ハブの故障などによって全てのサーバが孤立した場合は全サーバダウンとなります。

* データミラー方式
  + 2ノードデータミラーのクラスタで使用できます。
  + データミラー環境では、本方式が自動的に選択され、他のネットワークパーティション方式は選択できません。
  + 万一、クラスタサーバ間の全てのネットワーク通信路に障害が発生した場合には、両サーバがフェイルオーバを実行します。

[6] VERITAS Volume Managerを使用する場合、共有ディスク方式は使用できません。

* ネットワークパーティション解決しない
  + ディスクリソース（共有ディスク）を使用しないクラスタで選択できます。
  + 万一、クラスタサーバ間の全てのネットワーク通信路に障害が発生した場合には、全サーバがフェイルオーバを実行します。

推奨するネットワークパーティション解決方式は下記です。
- データミラー環境ではデータミラー方式が自動的に選択されます。
- 3ノード以上のクラスタには多数決方式を推奨します。
- 2ノードで共有ディスクを使用するクラスタにはCOM+共有ディスク方式を推奨します。
- 2ノードで共有ディスクを使用しないクラスタを使用するクラスタにはCOM方式を推奨します。
- VERITAS Volume Managerを使用する場合は共有ディスク方式が使用できないためCOM方式が必須です。

ネットワークパーティション解決方式	ノード数	必要HW	フェイルオーバ不可のケース	全ネットワーク経路断線時	両サーバがフェイルオーバするケース	ネットワークパーティション解決に必要な時間
COM	2	シリアルケーブル	COM異常	最高優先度のサーバが生存	全ネットワーク断線と同時にCOM異常発生	0
共有ディスク	制限なし	共有ディスク	ディスク異常	最高優先度のサーバが生存	なし	ハートビートタイムアウトとディスクIO待ち時間から計算される時間が必要
COM＋共有ディスク	2	シリアルケーブル共有ディスク	COM異常かつディスク異常	最高優先度のサーバが生存	なし	0
多数決	3以上	なし	過半数サーバダウン	過半数サーバと通信できるサーバが生存	なし	0
データミラー	2	なし	なし	両サーバがフェイルオーバ実施	全ネットワーク断線時	0
なし	制限なし	なし	なし	全サーバがフェイルオーバ実施	全ネットワーク断線時	0

1.2.5 フェイルオーバ資源

CLUSTERPROがフェイルオーバ対象にできる主な資源とそれぞれの活性化/非活性化の方法は以下のとおりです。

* 切替パーティション
  + 業務アプリケーションが引き継ぐべきデータを格納するためのディスクパーティションです。

* フローティングIPアドレス
  + フローティングIPアドレスを使用して業務へ接続することで、フェイルオーバによる業務の実行位置(サーバ)の変化をクライアントは気にする必要がなくなります。
  + パブリックLANアダプタへのIPアドレス動的割り当てとARPパケットの送信により実現しています。ほとんどのネットワーク機器からフローティングIPアドレスによる接続が可能です

> <注意>
> ARPパケットを受信することでARPテーブルを更新できるネットワーク機器であれば、フローティングIPアドレスでの接続が可能です。

* 仮想コンピュータ名
  + 仮想的なコンピュータ名(NetBIOS名)を指定して業務に接続できます。
  + NetBIOS名の動的な追加削除機能により実現しています。Windowsクライアントからの接続に使用します。また、リモートLANから接続する場合は、WINSサーバが必要になります。
  + Windows Server 2003, Windows XP, Windows 2000, Windows Me, Windows 98にて仮想コンピュータ名を使用する場合には、CLUSTERPROクライアントのインストールが必要になる場合があります[7]。

* スクリプト
  + CLUSTERPROでは、業務アプリケーションをスクリプトから起動します。
  + 共有ディスクにて引き継がれたファイルはファイルシステムとして正常であっても、データとして不完全な状態にある場合があります。スクリプトにはアプリケーションの起動のほか、フェイルオーバ時の業務固有の復旧処理も記述します。

[7]詳しくは「CLUSTERPROシステム構築ガイド システム設計編(基本/共有ディスク)」「CLUSTERPROシステム構築ガイド システム設計編(基本/ミラーディスク)」を参照してください。

1.3　クラスタシステムの構成と運用形態

1.3.1　システム構成

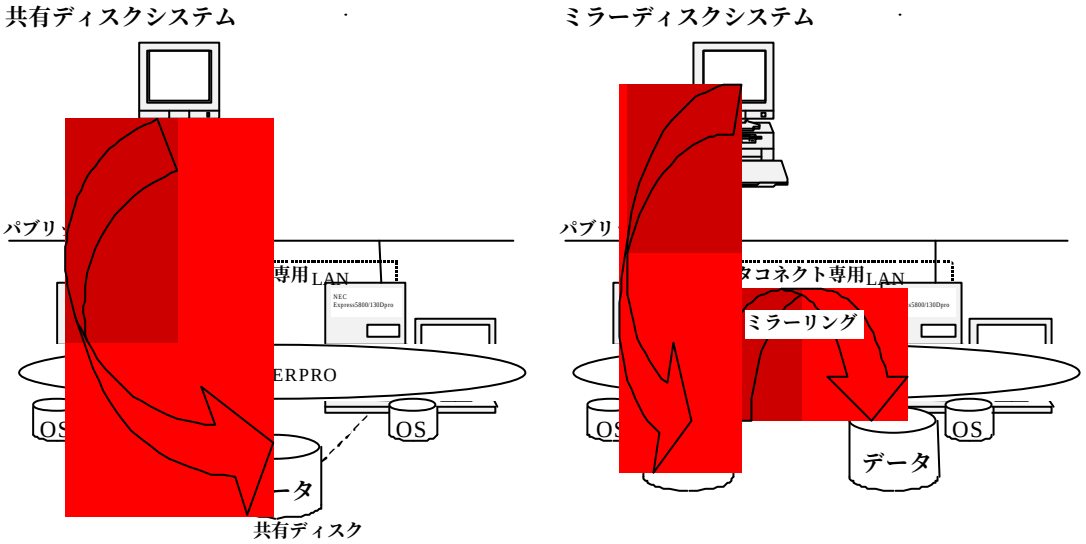
(1)　共有ディスクシステム

ディスクアレイ装置をクラスタサーバ間で共有する構成です。サーバ障害時には待機系サーバが共有ディスク上のデータを使用し業務を引き継ぎます。共有ディスクを使用することにより数十GBのデータ容量から数TBのデータ容量までの範囲をカバーできます。

(2)　ミラーディスクシステム(共有ディスクなし)

共有ディスク装置を使用せず、クラスタサーバ上のデータディスクをネットワーク経由でミラーリングする構成です。サーバ障害時には待機系サーバ上のミラーデータを使用し業務を引き継ぎます。データのミラーリングはI/O単位で行うため上位アプリケーションから見ると共有ディスクと同様に見えます。ミラーディスクのデータ容量は障害復旧後のミラー再構築時間に影響するため、20～30GBまでのデータ容量のシステムに向いています。

> ミラー再構築とは
>
> ミラーディスクシステムでサーバまたはディスクに障害が発生すると一方のサーバのディスクのみデータ更新されるため、サーバ間のミラーディスクに不整合が生じます。障害復旧後にミラー再構築を行うことでこの不整合を解消します。

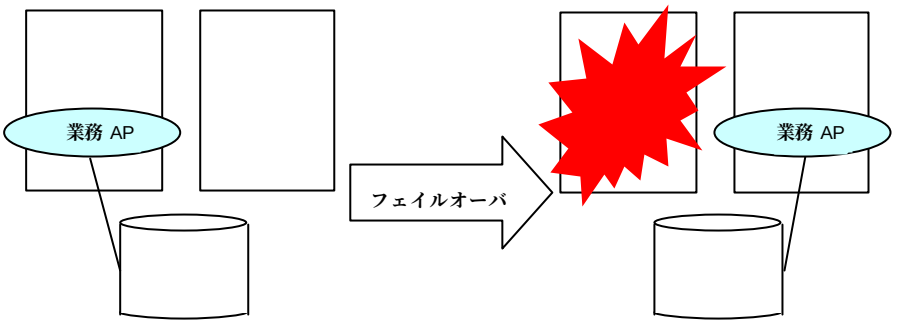


1.3.2　運用形態によるクラスタシステムの分類

(1)　片方向スタンバイクラスタ

一方のサーバを運用系として業務を稼動させ、他方のサーバを待機系として業務を稼動させない運用形態です。

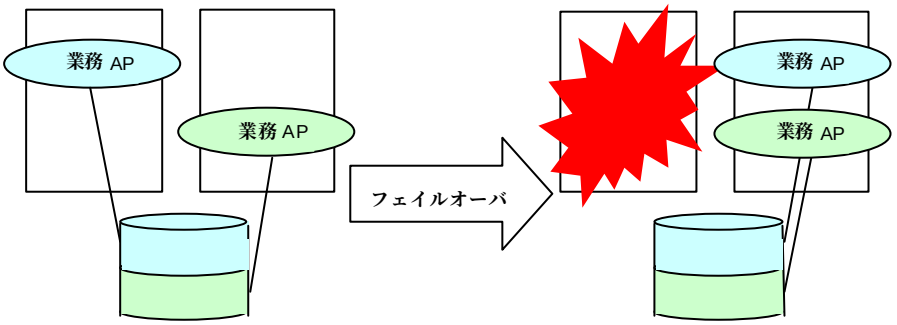
最もシンプルな運用形態でフェイルオーバ後の性能劣化のない可用性の高いシステムを構築できます。



(2)　同一アプリケーション双方向スタンバイクラスタ

複数のサーバである業務アプリケーションを稼動させ相互に待機する運用形態です。

アプリケーションは双方向スタンバイ運用をサポートしているものでなければなりません。ある業務データを複数に分割できる場合に、アクセスしようとしているデータによってクライアントからの接続先サーバを変更することで、データ分割単位での負荷分散システムを構築できます。

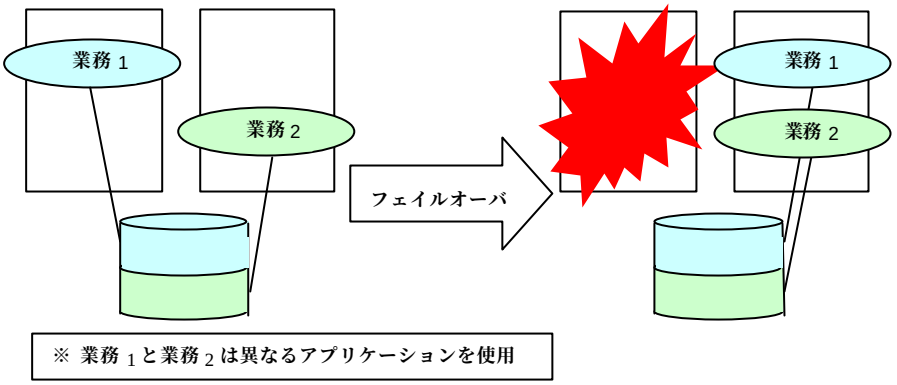


※ 図の業務 AP は同一アプリケーション
※ フェイルオーバ後に一つのサーバ上で複数の業務 AP インスタンスが動く

(3) 異種アプリケーション双方向スタンバイクラスタ

複数の種類の業務アプリケーションをそれぞれ異なるサーバで稼動させ相互に待機する運用形態です。

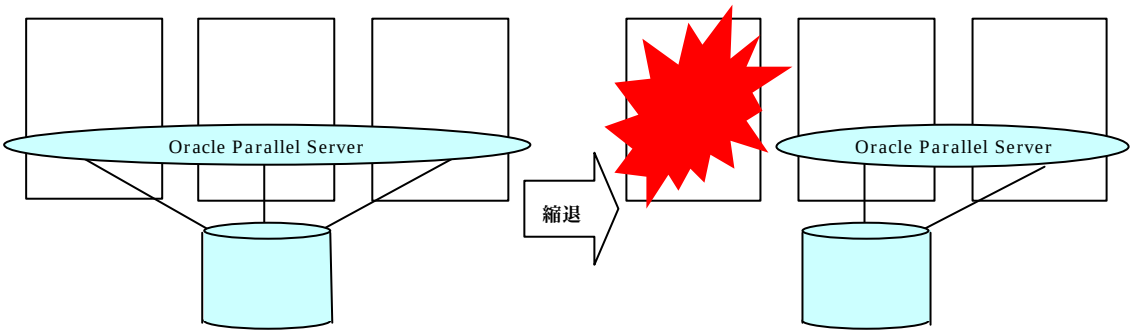
アプリケーションが双方向スタンバイ運用をサポートしている必要はありません。業務単位での負荷分散システムを構築できます



(4) パラレルクラスタ[8]

並列データベースエンジン(Oracle Parallel Server)を使用する運用形態です。

この運用形態では全てのサーバで一つのデータベースにアクセスでき、あるサーバが障害を起こした場合は縮退します。負荷分散システム基盤を構築できると同時に、データベースへの並列クエリー(Oracle Parallel Query)と組み合わせることで高性能データベースとしても機能します。



[8] VERITAS Volume Managerを使用する場合はOracle Parallel Serverは使用できません。

1.3.3 CLUSTERPRO製品構成

CLUSTERPROはシステム規模に対応して製品を選択できるように下表のように製品化しています。

		CLUSTERPRO LE	CLUSTERPRO SE	CLUSTERPRO EE	CLUSTERPRO SX
システム規模 (データ容量)		中小規模 (20～30GB)※	中規模 (TBクラスまで)	大中規模 (TBクラスまで)	大規模 (TBクラスまで)
ディスク		ミラーディスク	共有ディスク	共有ディスク	共有ディスク
サーバ数		2	2	2～16	2～16
運用形態	片方向	○	○	○	○
	同一AP双方向	○	○	○	○
	異種AP双方向	○	○	○	○
	パラレル	×	○	○	○
サポートPP		○	○	○	○

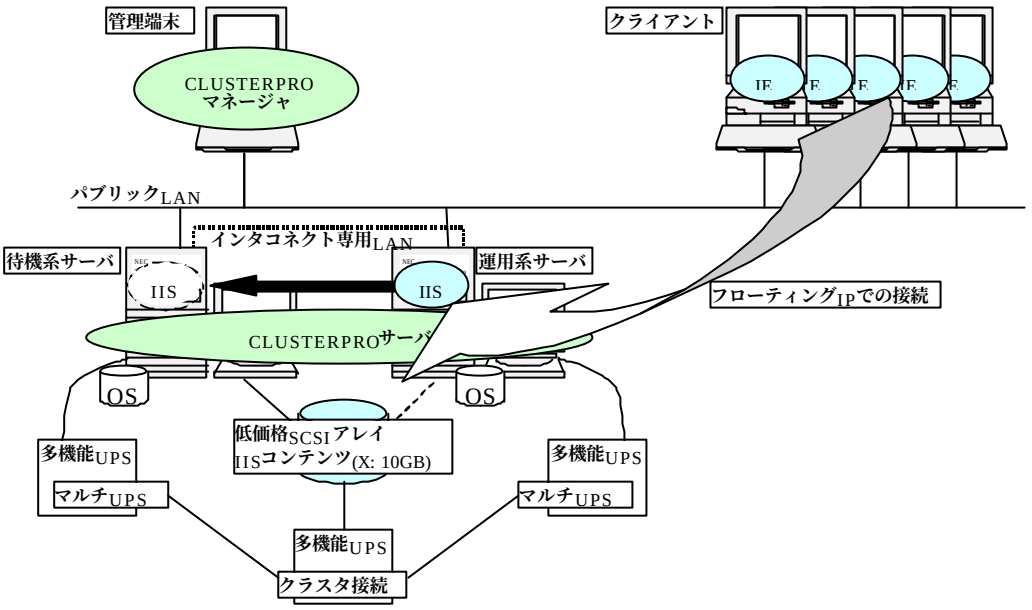
※ ミラー構築・再構築に要する時間の考慮が必要です。
CLUSTERPRO LEにおいて、FastSync Option（有償）をインストールすると、ミラー再構築の時間を短縮できる場合があります。
詳細については「CLUSTERPROシステム構築ガイド システム設計編(基本/ミラーディスク)」を参照してください。

2　クラスタシステムの設計

2.1　システム概要

この章では、実際のシステムの設計を例にシステム設計時に注意しなければならない点について説明します。

IISを利用したWWWによる情報発信サーバをCLUSTERPROによる片方向スタンバイクラスタシステムとして設計します。



2.2 クラスタシステム設計ステップ

2.2.1 システム構成の決定

(1) ハードウェア構成

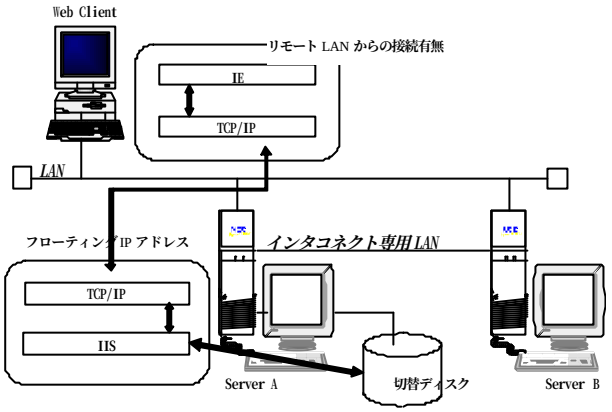
業務に必要な性能要件や必要なディスク容量からハードウェアの構成品を決定します。

(2) ソフトウェア構成

業務規模や運用形態から使用するCLUSTERPROを決定し、使用する業務アプリケーションを洗い出します。

(3) ネットワークおよびドメイン構成

業務アプリケーションのデータの流れを中心にネットワーク構成を決定します。



クラスタシステムでは、Windows 2000, Windows Server 2003 のActive Directory環境においても、クラスタサーバが同一ドメインに属している必要があります。CLUSTERPROサーバをActive Directoryのドメインコントローラとすることは推奨しません。

2.2.2 アプリケーションのCLUSTERPRO対応確認

構築しようとしているシステムで使用する業務アプリケーションがCLUSTERPRO上で動作するか確認します。

(1) 動作確認済みプログラムプロダクト

CLUSTERPROでは、主なプログラムプロダクト(PP)について動作確認を行い、構築方法、注意事項を参考情報として公開しています。

> <注意>
> クラスタシステム上でアプリケーションやサービスを動作させる場合、機能制限や注意事項がある場合があります。
> 詳しくは「CLUSTERPROシステム構築ガイド PP編」を参照してください。

下記は、動作確認済みPPの一部です。

+ IIS (Internet Information Server) Version 4.0/5.0
+ Exchange Server 5.5 Enterprise Edition
+ SQL Server Version 6.5/7.0
+ Oracle R7.3.4/8.0.x/8.1.5
+ ARCserve
+ Backup Exec
+ ESMPRO/ServerAgent, ESMPRO/ServerManager

(2) その他のアプリケーション

動作確認できていないPPを使用したり、新規にアプリケーションを作成したりする場合は、各アプリケーション用のスクリプトを記述することでクラスタ対応できます。

> <注意>
> クラスタシステム上で動作させようとするアプリケーションの性質や構造によってはクラスタ対応できない場合があります。
> 詳しくは「CLUSTERPROシステム構築ガイド システム設計編(基本) – CLUSTERPRO環境下でのアプリケーション/サービス」を参照してください。

2.2.3 フェイルオーバグループの設計

次の点について注意しフェイルオーバグループを設計します。

- フェイルオーバグループの単位についての指針
  - 依存関係のあるアプリケーション/資源群を一つのグループとする
  - 各業務単位でグループを作成する
  - 負荷分散単位でグループを作成する
- フェイルオーバポリシについての考慮
  - フェイルオーバ後のサーバのシステム要件を確認
  - フェイルオーバ先として必要なサーバだけをフェイルオーバポリシに登録

フェイルオーバポリシとは フェイルオーバグループに設定された属性であり、フェイルオーバ先として登録したサーバと優先順位の一覧です。CLUSTERPROは優先順位を元に、フェイルオーバ先サーバを制御します。

例えば、IISの情報発信用のフェイルオーバグループであれば、次のようなフェイルオーバグループを一つ作成することになります。

フェイルオーバグループ名		IIS
資源	切替パーティション	X: (NTFS – 10GB)
	フローティングIPアドレス	192.168.0.3
フェイルオーバポリシ		SERVER1→ SERVER2

2.2.4 スクリプト

システム構築ガイドPP編を参考にスクリプトを作成します。
スクリプト簡易作成支援オプションを使用することによりスクリプトを作成が容易になります。詳しくは下記を参照してください。

http://www.ace.comp.nec.co.jp/CLUSTERPRO/

スクリプトの作成例
(開始スクリプト: IIS片方向スタンバイ, ARMLOADによるAP監視, ARMRSPによる資源監視)

```
REM START.BAT
REM GROUP NAME=IIS; FAILOVER_POLICY=SERVERA->SERVERB
REM DATE: 2000/01/31 REV1
IF “%ARMS_EVENT%” == “START” GOTO NORMAL
IF “%ARMS_EVENT%” == “FAILOVER” GOTO FAILOVER
IF “%ARMS_EVENT%” == “RECOVER” GOTO RECOVER
GOTO NO_ARM

:NORMAL
ARMLOAD RSP /R 9 /H 1 /FOV ARMRSP /A

IF “%ARMS_DISK%” == “FAILURE” GOTO ERROR_DIS

ARMLOG “NORMAL START IIS GROUP SCRIPT”
ARMLOAD WWWPS /S /M “W3SVC”
ARMLOG “WWW STARTED”
GOTO EXIT

:FAILOVER
ARMLOAD RSP /R 9 /H 1 /FOV ARMRSP /A

IF “%ARMS_DISK%” == “FAILURE” GOTO ERROR_DIS

ARMLOG “FAILOVER START IIS GROUP SCRIPT”
ARMLOAD WWWPS /S /M “W3SVC”
ARMLOG “WWW STARTED”
GOTO EXIT

:RECOVER
ARMLOG “RECOVER START IIS GROUP”
GOTO EXIT
:NO_ARM
ARMLOG “ERROR NO_ARM IIS GROUP” /ARM
GOTO EXIT
:ERROR_DISK


ARMLOG “ERROR DISK IIS G

start diskfail.bat
```

- クラスタ資源の監視を行う

- アプリケーションの生存監視を行う
- アプリケーション起動ごとにARMLOGでログ出力することでスクリプトのデバッグを容易にする

- フェイルオーバ後の明示的な復旧処理が必要なければ、通常起動時と同じ記述でよい。

処理概要：
ディスクエラー発生時の処理

追加部分

```
rem ***************************************
rem *            diskfail.BAT             *
rem ***************************************

IF “%ARMS_SERVER%” == “OTHER” GOTO EXIT

START ARMFOVER /F %ARMS_GROUPNAME%


:EXIT
exit
```

- ディスクの接続に失敗した場合、待機系へフェイルオーバを試みる

スクリプトの作成例
(停止スクリプト: IIS片方向スタンバイ, ARMLOADによるAP監視, ARMRSPによる資源監視)

```
REM STOP.BAT
REM GROUP NAME=IIS; FAILOVER_POLICY=SERVERA->SERVERB
REM DATE: 2000/01/31 REV1
IF “%ARMS_EVENT%” == “START” GOTO NORMAL
IF “%ARMS_EVENT%” == “FAILOVER” GOTO FAILOVER
GOTO NO_ARM

:NORMAL
ARMKILL RSP

IF “%ARMS_DISK%” == “FAILURE” GOTO ERROR_DISK

ARMLOG “NORMAL STOP IIS GROUP SCRIPT”
ARMKILL WWWPS
ARMLOG “WWW STOPPED”
GOTO EXIT

:FAILOVER
ARMKILL RSP

IF “%ARMS_DISK%” == “FAILURE” GOTO ERROR_DISK

ARMLOG “FAILOVER STOP IIS GROUP SCRIPT”
ARMKILL WWWPS
ARMLOG “WWW STOPPED”
GOTO EXIT

:NO_ARM
ARMLOG “ERROR NO_ARM IIS GROUP” /ARM
GOTO EXIT
:ERROR_DISK
GOTO EXIT

:EXIT
ARMLOG “END IIS STOP SCRIPT”
```

- ARMLOADで起動したアプリケーションはARMKILLで終了する

2.3 信頼性確保のポイント

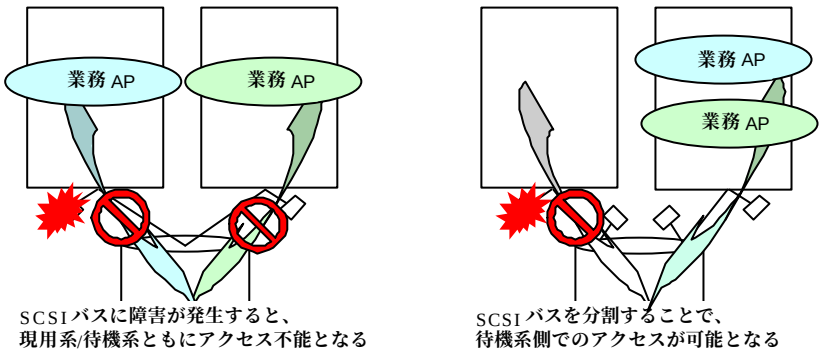
2.3.1 Single Point of Failure(SPF)の排除

クラスタシステムではサーバの多重化を実現しますが、共有ディスクなど、サーバ間で共有する部分もあります。この共有部分を多重化もしくは排除することで信頼性を高めるようクラスタシステム設計してください。

＊ 共有ディスクのSCSIバス/FibreChannelとコントローラ

・ミラーディスク	➔共有部分なし
・SCSIバス分割 + ディスク資源監視 (ARMRSPによるフェイルオーバ)	➔他ノードへフェイルオーバ 障害発生時に行っていた業務の再実行
・SCSI/FibreChannelの二重化 (パス切替SWを使用)	➔健全なパスへ切り替え 業務は継続実行

SCSI バスが SPF となる場合



SCSIバスに障害が発生すると、
現用系/待機系ともにアクセス不能となる

SCSIバスを分割することで、
待機系側でのアクセスが可能となる

＊ クライアントと接続するLAN

・LANの多重化	➔健全なパスへ切り替え 障害発生時に行っていた業務の再実行

＊ その他の共有装置

・多重化可能	➔装置を多重化する
・多重化不可能	➔共有装置の故障によって共有装置を使用しない業務も含めてシステム全体が停止しないようシステム設計する

Single Point of Failure(SPF)とは 単一障害によりシステムの停止を引き起こしてしまうシステム部分のことを言います。クラスタシステムでは、サーバ間で共有している共有ディスクへのアクセスパスが代表的な例です。

2.3.2 障害の検出

業務を継続不能な状態または、業務継続不能になりうる障害を早期に検出できるシステムを設計してください。

* ハードウェア障害
  + 二重障害が発生する前に障害を検出し通報する
    (RAID構成のアレイディスクで二台以上のディスクに障害が発生するなど)
  + フェイルオーバしない障害を検出しフェイルオーバする
    ESMPRO/AlertManagerを利用して障害の前触れとなるイベントログなどを監視しフェイルオーバする
    (詳しくは、「CLUSTERPROシステム構築ガイド システム設計編(応用)」を参照してください)

* ソフトウェア障害
  + アプリケーションの障害を検出しフェイルオーバする
    ARMLOADを使用するかアプリケーションモニタを作成し障害発生時にフェイルオーバする
    (詳しくは、「システム構築ガイド システム設計編(応用)」を参照してください)

2.3.3 フェイルオーバ後の信頼性

フェイルオーバ後に業務アプリケーションが正常に動作できるようシステムを設計してください。

運用形態	注意点
片方向スタンバイクラスタ	待機系サーバのCPU性能やメモリ容量などの諸元は十分か？
同一AP双方向スタンバイクラスタ	サーバが複数の業務を実行するために必要な諸元を確保しているか？ フェイルオーバ後のアプリケーションの動作に問題はないか？
異種AP双方向スタンバイクラスタ	サーバが複数の業務を実行するために必要な諸元を確保しているか？ フェイルオーバ後のアプリケーション間の相性に問題はないか？

3　クラスタシステムの構築から運用

3.1　クラスタシステムの構築

クラスタシステムの構築は下記のように行います。下記に記述した点について注意しクラスタシステムを構築します。インストール手順や設定方法の詳細については、各ソフトウェアのセットアップカードなどのドキュメントを参照してください。

(1)OSのインストール
OSを各サーバにインストールします。このとき、以下の点に注意してください。
* サービスパックを適用する
* クラスタサーバは同一ドメインに属する
* SNMPサービスをインストールする
* サーバサービスのプロパティを
「ネットワークアプリケーションのスループットを最大にする」に変更する
* インタコネクト専用LANアダプタにはTCP/IP以外のプロトコルをバインドしない
* 時刻を合わせる

(2)CLUSTERPROのインストール

① CLUSTERPROサーバ
* クラスタサーバにCLUSTERPROサーバをインストール
* 共有ディスクとして使用するディスクをX-CALLに設定
* 共有ディスク上にパーティション作成
* 作成したパーティションにドライブレター/クラスタ文字設定
* OS起動時間調整
* 「CLUSTERPRO Server」サービスの[スタートアップの種類]を自動に設定

② CLUSTERPROマネージャ
* 管理端末に、CLUSTERPROマネージャをインストール

③ クラスタ生成 － サーバ追加 － グループ追加
* CLUSTERPROマネージャから行う

(3)UPS関連ソフトウェアのインストール
* ESMPRO/UPSController と ESMPRO/AutomaticRunningController をインストール(構築ガイドPP編(ESMPROシリーズ)を参照)

(4)業務アプリケーションのインストール
* 業務で使用するアプリケーションをインストール(構築ガイドPP編を参照)

3.2 運用前の評価と障害復旧マニュアルの作成

本番運用前に本番時の負荷や障害を想定した評価を行い、作成したスクリプトの動作の確認やパラメータ調整を行います。この評価に基づいて障害発生時の復旧マニュアルを作成してください。

3.2.1 障害発生個所と偽証評価

障害発生個所を洗い出し偽証評価を行います。例えば、次のような障害個所が考えられます。

障害個所		主な障害内容	業務継続性
サーバ本体		CPU故障(OS STOPエラー)	△1
		ディスプレイ/キーボード/マウス故障	○
共有ディスク ※ 共有ディスクシステムの場合のみ	ディスクアレイ装置	HDD一台故障	○
		HDD複数台故障	×
		コントローラ故障	△2
	SCSI/FibreChannel	ケーブル断線, 終端故障, HUB/Switch故障	△2
	HBA (Host Bus Adapter)	アダプタ故障(OS STOPエラー)	△1
		アダプタ故障(I/Oエラー)	△3
UPS		サーバ接続UPS故障, RS232C故障	○
		サーバ接続UPS停電	△1
		共有ディスク接続UPS故障, RS232C故障	○
		共有ディスク接続UPS停電	×
		全UPS停電	×
LAN	インタコネクトLAN	NIC故障	○
		ケーブル断線, HUB/Switch故障	○
	パブリックLAN	NIC故障	△4
		ケーブル断線, HUB/Switchのポート故障	△4
ソフトウェア		OS STOPエラー	△1
		業務アプリケーションの停止	△5
		業務アプリケーションのストール	△6

○：業務の継続が可能
△：CLUSTERPROまたはハードウェアの機能により、業務の継続が可能
×：業務の継続はできない

△1: フェイルオーバが発生し待機系で業務を継続する
△2: 共有DISKへのパスを二重化することで正常なパスで業務を継続する
△3: リソース監視を使用するかESMPRO/AlertManagerと連携することでディスク資源の障害を検出、フェイルオーバを実行し待機系で業務を継続する
△4: リソース監視を使用するかESMPRO/AlertManagerと連携することでLANの障害を検出し、フェイルオーバを実行し待機系で業務を継続する
△5: ARMLOADコマンドを使用しアプリケーションの障害を検出し、フェイルオーバを実行し待機系で業務を継続する
△6: 業務アプリケーションに合わせたアプリケーションモニタを作成することでス

トール検出し、フェイルオーバを実行し待機系で業務を継続する

3.2.2 状態遷移評価

運用に即した評価項目を作成しクラスタシステムの状態遷移評価を行います。２サーバのクラスタシステムでは、次のような評価項目が考えられます。

評価項目		操作	確認内容
起動	クラスタ起動	クラスタを構成する全てのサーバを起動する	全てのサーバおよびフェイルオーバグループが正常に起動する
	運用系のみ起動	運用系サーバのみ起動する	OS起動後5分間待機系の起動を待ち合わせた後、フェイルオーバグループが起動する
	待機系のみ起動	待機系サーバのみ起動する	
シャットダウン		クラスタをシャットダウンする (CLUSTERPROマネージャを使用)	全てのサーバがシャットダウンする また再起動後に全てのサーバおよびフェイルオーバグループが正常になることを確認する UPSを使用している場合は、UPSの電源が切断されることも確認する
フェイルオーバ	サーバ切り離し	現用系サーバをクラスタから切り離す (CLUSTERPROマネージャを使用)	現用系サーバがクラスタから切り離され、フェイルオーバが発生する フェイルオーバ後に業務が継続できる
	サーバシャットダウン	現用系サーバをSTARTメニューからシャットダウンする	
	サーバ電源切断	現用系サーバの電源を切断する (シャットダウンを行わず)	
フェイルオーバグループ	移動	フェイルオーバグループを移動/停止/起動する (CLUSTERPROマネージャを使用)	フェイルオーバグループが待機系へ移動する 移動後業務が継続できる
	停止		切替パーティションの切り離しに失敗しない
	起動		フェイルオーバグループが起動する 起動後業務が開始できる
障害復旧	片サーバダウンからの復帰	クラスタから切り離されているサーバをクラスタに復帰する (CLUSTERPROマネージャを使用)	クラスタに組み込まれ待機系になる

	両サーバダウンからの復帰	**適切なサーバをクラスタに強制復帰し、他のサーバを復帰する** (CLUSTERPRO マネージャを使用)	**クラスタ状態が正常状態になる** **この時点ではグループは停止状態となる**

3.2.3 パラメータ調整

下記の項目はシステム構成に応じて調整をする必要があります。

項目	調整方法
ハートビートタイムアウト時間	サーバに業務で発生する最大負荷をかける。 (その状況で不正にフェイルオーバが起きないこと) フェイルオーバ時にディスクパス切替を起こす現象が発生する場合に対応するか否か検討する。
立ち上げ同期の猶予時間	全てのサーバの起動時間差を計測する。
ディスクIO待ち時間	サーバダウン時に共有ディスクにアクセス可能になるまでの目安時間を指定する。 例えば、共有ディスクにNEC FibreChannelディスクやiStorageディスクを使用して、サーバとFC 接続を行う場合には、80 秒以上を設定する。
スクリプトタイムアウト時間	スクリプトの最大実行時間を計測する。
OS起動時間	共有ディスクの起動にかかる最大時間を設定する。

3.3 クラスタシステムの運用

クラスタシステムの運用手順や注意点をまとめます。CLUSTERPROマネージャの具体的な操作については「CLUSTERPROシステム構築ガイド GUIリファレンス」を参照してください。

(1) クラスタ起動

* 自動運転での起動
  + ESMPRO/AutomaticRunningControllerで設定した時間になると、自動的に共有ディスクと2台のサーバの電源が投入されます。

    = 多機能UPSのAUTO/LOCALスイッチがAUTOになっていなければなりません。
    = サーバのOS起動待ち時間が共有ディスクの起動時間より大きい値が設定していなければなりません。

* 手動での起動
  + 多機能UPSのAUTO/LOCALスイッチをLOCALに変更しUPSの電源を投入します。

    = LOCAL状態では自動運転しません。自動運転に戻すためにはAUTO/LOCALスイッチをAUTOに変更します。

(2) クラスタシャットダウン

* 自動運転でのシャットダウン
  + ESMPRO/AutomaticRunningControllerで設定した時間になると、自動的にクラスタシャットダウンを実行し、共有ディスクと2台のサーバの電源が切断されます。

    = 多機能UPSのAUTO/LOCALスイッチがAUTOになっていなければなりません。

* 手動でのシャットダウン
  + CLUSTERPROマネージャにて、クラスタを選択しクラスタシャットダウンを行います。

    = 多機能UPSのAUTO/LOCALスイッチがAUTOになっていなければなりません。

> <注意>
>
> 多機能UPSのAUTO/LOCALスイッチがLOCAL状態では、CLUSTERPROのサービスだけが停止しOSはシャットダウンしません。この場合、AUTO状態に変更することでシャットダウン開始します。
> 保守以外では、AUTO状態で運用してください。

(3) 現用系ダウン

* 自動的にフェイルオーバする場合
  + 現用系サーバがダウンした場合、CLUSTERPROマネージャ画面でダウンしたサーバが赤く表示され、自動的に待機系へのフェイルオーバが発生します。

* 自動的にフェイルオーバしない場合
  + OSの一部機能のストールやアプリケーションのストールなど、CLUSTERPROが自動でフェイルオーバできない場合、現用系サーバのスタートメニューからシャットダウンまたはダンプスイッチの押下などを行い、現用系サーバを完全にダウンさせてください。これによって、CLUSTERPROは相手サーバのダウンに気づき待機系へフェイルオーバを実行します。
  + このとき、できる限り障害サーバの電源は切断してください。

(4) 障害サーバの復帰

CLUSTERPROでは障害発生後にサーバを再起動するとそのサーバは自動的にクラスタから切り離された状態になります。そのサーバを待機状態に戻すためには、障害原因を取り除いた後、クラスタへ復帰させる必要があります。

CLUSTERPROマネージャにて、クラスタから切り離されたサーバをクラスタに復帰します。

> <注意>
>
> クラスタから切り離された状態でもCLUSTERPROの状態確認などの通信を行っています。このため、ネットワークやHBAなど他のサーバと接続している部位を保守する場合は、クラスタシステムからハードウェア的に切り離して(ネットワークおよび共有ディスクが接続されていない状態で)作業してください。

(5) サーバの保守(ローリングアップグレード)

メモリ追加などのサーバ保守は片方のサーバずつアップグレード(ローリングアップグレード)することでシステム停止時間を最小限にとどめることが可能です。

以下の手順で行います。

(1) 待機系サーバの切り離し
CLUSTERPROマネージャから、待機系サーバを選択しサーバ切り離しを行います。
(2) 待機系サーバにメモリを追加し再起動する。
(3) 待機系サーバの復帰
CLUSTERPROマネージャで、待機系サーバを選択しサーバ復帰を行います。
(4) グループ移動
CLUSTERPROマネージャで、運用系サーバで動作しているグループを待機系サーバへ移動します。
(5) 運用系サーバの切り離し
(6) 運用系サーバにメモリを追加し再起動する。

<注意>

ソフトウェアのローリングアップグレードは運用系と待機系でソフトウェアリビジョンが変わることによる弊害が考えられます。
共有ディスク上のデータの互換性など問題が無いことが分かっている場合にのみ行ってください。

(6) 全てのサーバがダウンした場合の復帰

何らかの原因で全てのサーバがダウンしてしまった場合は、信頼できるクラスタ情報を持っているサーバをCLUSTERPROで判断できないためサーバの復帰ができません。

この状態からの復帰は、以下の手順で行います。

(1) CLUSTERPROマネージャから、クラスタ情報のベースとなるサーバを選択してサーバ強制復帰を行います。
(2) CLUSTERPROマネージャから、残りのサーバを復帰させます。
(3) CLUSTERPROマネージャから、フェイルオーバグループを起動します。

4　高度なクラスタ

4.1　さらに信頼性を高めるために

4.1.1　アプリケーション障害への対策

アプリケーションのストールや結果異常など、ARMLOADコマンドでは検出できない障害に対応するためには、アプリケーションごとの内部事情を理解した上で、アプリケーションの状態を監視するようなモニタプログラムを作成します。

例えば、定期的にデータベースへアクセスし一定時間応答がないなどストールしていると判断できる場合に、サーバをシャットダウンすることで、データベースアプリケーションのストール監視を行うことができます。また、アプリケーションの使用メモリ容量や使用ディスク容量を監視し、異常に消費している場合フェイルオーバグループを移動するという方法も考えられます。

このようなモニタプログラムは、サーバのシャットダウンやフェイルオーバのために、CLUSTERPROコマンド (ARMDOWN/ARMFOVER)を利用することができます。

詳しくは、「システム構築ガイド システム設計編(応用)」を参照してください。

4.1.2　ハードウェア障害・OSの部分障害への対策

ハードウェア障害のうちいくつかは業務が停止する前にイベントログなどに警告または異常を登録します。これらを監視することで、CLUSTERPROが直接検出できない障害をフェイルオーバ対象とすることができます。

これは次の方法で、ハードウェア障害の発生を検出しCLUSTERPROへフェイルオーバ指示を出すことで実現します。

* ESMPRO/ServerAgentによる障害監視
* ESMPRO/AlertManagerによる障害時のフェイルオーバ指示

同様の方法によってメモリやディスクの使用量などを監視し、OSが障害に陥る前に事前に障害の種を検出しフェイルオーバを行うことで、OSの部分障害による業務停止を防ぐことが可能です。

詳しくは、「システム構築ガイド システム設計編(応用)」を参照してください。

4.1.3　サーバマネージメントボード

共有ディスクシステム(共有ディスク型のクラスタシステム)では、SCSIバスを複数のサーバで共有しているため、サーバ障害時にSCSIアダプタ(HBA)の故障などによってバスをノイズなどで乱してしまう場合、CLUSTERPROはフェイルオーバに失敗する場合があります。フェイルオーバに成功しても、待機系で共有ディスクへのアクセスが正常にできず業務が実行できない場合もあります。

このような場合、サーバマネージメントボード(SMB)を使用することで、障害サーバの電源を切断することが有効です。

詳しくは、「CLUSTERPROシステム構築ガイド システム設計編(応用)」を参照してください。

なお、上記機能を使用するためには、NECのサーバマネージメントボード (N8503-33 サーバマネージメントボード)あるいは、それと同等のサーバマネージメントボードが必要です。

4.1.4 LAN二重化

LANの二重化とは、クライアントとサーバ間の通信パスを複数用意し一つのパスで障害が発生した場合に他方のパスを使用して業務を継続できるようにすることです。

詳しくは、「CLUSTERPROシステム構築ガイド システム設計編(応用)」を参照してください。

4.1.5 ディスクパスの二重化

共有ディスクへのアクセスパスを二重化します。共有ディスクにデュアルポート機構を増設し、パス切替SWを使用することで実現します。

詳しくは、「システム構築ガイド システム設計編(応用)」を参照してください。

4.2 性能を向上するために

4.2.1 パラレルクラスタ

Oracle Parallel Server(OPS)を使用することで、複数のサーバから同一のデータベースへアクセスできる並列データベースエンジンを構築することができます。OPSによって、可用性とともに拡張性の高い分散処理アプリケーションの構築が容易になります。また、Oracle Parallel Query(OPQ)によって並列検索を行うことで、検索処理の高速化を実現できます。

OPSについては、「Oracle Parallel Serverシステム構築ガイド」を参照してください。

4.3 その他

4.3.1 CLUSTERPROクライアント

通常の使用では、クライアントPCにCLUSTERPROのモジュールをインストールする必要はありません。

CLUSTERPROクライアントをクライアントPCにインストールすることによって、フェイルオーバの発生などをクライアントPC上でポップアップにてユーザに知らせることができます。また、パブリックLANを二重化する場合には、CLUSTERPROクライアントをインストールすることでLANの経路切り替えを自動で行うことができます。

詳しくは、「CLUSTERPROシステム構築ガイド システム設計編(応用)」を参照してください。

4.3.2 CLUSTERPROコマンド

CLUSTERPROでは、ARMLOADなど高可用性を実現するためのコマンド、フェイルオーバグループのスクリプトの記述を容易にするためのコマンドや、クラスタシステムを運用管理するためのコマンドを用意しています。

仮想コンピュータ名を使用する場合には、CLUSTERPROクライアントのインストールが必要になることがあります。

詳しくは、「CLUSTERPROシステム構築ガイド コマンドリファレンス」を参照してください。

4.3.3 CLUSTERPRO API

CLUSTERPRO APIを利用することで、クラスタシステムの状態を把握して動作を行うア

プリケーション(Cluster-Aware AP)を作成することができます。これによって、クラスタの状態遷移を待って自動的に再接続するようなクライアントアプリケーションの開発などが可能です。

詳しくは、「CLUSTERPROシステム構築ガイド APIリファレンス」を参照してください。

4.3.4 回線切替装置

クラスタシステムで、V.24やX.21などの回線を扱う場合、回線切替装置を使用します。

回線リソースをフェイルオーバグループに登録することで、回線ごとにクラスタ内のどちらのサーバに接続するかを切り替えることができます。

詳しくは、「CLUSTERPROシステム構築ガイド システム設計編(応用)」を参照してください。

なお、上記機能を使用するためには、NECの回線切替装置(N8591-01/02 V.24/X.21回線切替装置、N8545-01/03 V.24/X.21回線切替ユニット、N8545-02/04 V.24/X.21回線切替拡張ユニット)あるいは、それと同等の回線切替装置が必要です。また、NECの高速多線回線ボード(N8104-102 高速多回線ボード)あるいは、それと同等の高速多回線ボードが必要です。

4.3.5 VERITAS Volume Manager

VERITAS Volume Managerを使用することにより、共有ディスクにダイナミックディスクが利用できます。

詳しくは、「CLUSTERPROシステム構築ガイド システム設計編(応用)」を参照してください。

5　付録

5.1　用語集

用　語	説　明
あ	
インタコネクト	クラスタサーバ間の通信パス。 （関連）プライベートLAN、パブリックLAN
か	
仮想IPアドレス	フェイルオーバした場合、クライアントのアプリケーションが接続先のサーバ切り替えを意識することなく行うために、CLUSTERPROが使用する仮想的なIPアドレス。 （関連）実IPアドレス
仮想コンピュータ名	フェイルオーバした場合、クライアントのアプリケーションが接続先のサーバ切り替えを意識することなく行うために、CLUSTERPROが使用する仮想的なコンピュータ名。
管理クライアント	CLUSTERPROマネージャが起動されているマシン。
起動属性	クラスタ起動時、自動的にフェイルオーバグループを起動するか手動で起動するかを決定するフェイルオーバグループの属性。 管理クライアントから設定が可能。
共有ディスク	複数サーバよりアクセス可能なディスク。 (関連)クロスコールディスク
共有ディスクシステム	共有ディスクを使用するクラスタシステム。
共有パーティション	複数のコンピュータに接続され、同時に使用可能なディスクパーティション。 (関連)切替パーティション、CLUSTERパーティション
切替パーティション	複数のコンピュータに接続され、切り替えながら使用可能なディスクパーティション。 （関連）共有パーティション、CLUSTERパーティション
切替ミラーディスク	ディスクミラーリングを行うことにより、同一のディスクに接続されているように使用することが可能なディスク。 （関連）CLUSTERパーティション
クラスタシステム	複数のコンピュータをLANなどでつないで、１つのシステムのように振る舞わせるシステム形態。
クラスタ復帰	障害によりクラスタから切り離されたサーバを、復旧後正常なクラスタに戻すこと。
クラスタシャットダウン	CLUSTERPROマネージャより、クラスタシステムを構成しているサーバを、クラスタとして正常にシャットダウンさせること。
クラスタディスクグループ	クラスタ環境で共有されるダイナミックディスクで構成されるディスクグループ。VERITAS Volume Manager下では、クラスタディスクグループが、リソースの単位になる。
クロスコールディスク	複数サーバよりアクセス可能なディスク、関連共有ディスク。
現用系	ある一つの業務セットについて、業務が動作しているサーバ。 (関連)待機系

用　語	説　明
さ	
セカンダリ(サーバ)	通常運用時、フェイルオーバグループがフェイルオーバする先のサーバ。 （関連）プライマリサーバ
実IPアドレス	仮想IPアドレスに対し、各マシンに通信のために設定されたIPアドレス。 （関連）仮想IPアドレス
た	
待機系	現用系ではない方のサーバ。 （関連）現用系
ディスクグループ	VERITAS Volume Managerでは、ディスクをディスクグループに分けて管理している。ディスクグループには以下の3種類がある。 **・ベーシックグループ** プライマリパーティション等を持つ物理ディスク（**ベーシックディスク**）で構成される。 **・ダイナミックグループ** **ダイナミックディスク**で構成されるディスクグループ。 **・クラスタディスクグループ** **クラスタディスク**で構成されるディスクグループ。 クラスタディスクはダイナミックディスクにクラスタ機能を追加したもの。 CLUSTERPROでVERITAS Volume Managerを使用する場合はこれを使用する。
な	
ネットワークパーティション症状	インタコネクトを使用して行うハートビートがLAN障害により両方とも途切れてしまうこと。 （関連）インタコネクト、ハートビート
は	
ハートビート	サーバ間の監視のために、定期的に送信しあうこと。 （関連）インタコネクト、ネットワークパーティション
パブリックLAN	サーバ/クライアント間通信パスのこと。 （関連）インタコネクト、プライベートLAN
フェイルオーバ	待機系が、現用系上の業務アプリケーションを引き取ること。
フェイルバック	フェイルオーバした後に両サーバを再起動させて業務を本来の現用系に戻すこと（Ver3.0のみ使用）。
フェイルオーバグループ	業務を実行するのに必要なクラスタ資源、属性の集合。
フェイルオーバグループの移動	両サーバの再起動をせずにフェイルバックを実行させること。
フェイルオーバポリシ	フェイルオーバ可能なサーバリストとその中でのフェイルオーバ優先順位を持つ属性。
プライベートLAN	インタコネクト専用LANと同じ意味で使用。 （関連）インタコネクト、パブリックLAN
プライマリ(サーバ)	フェイルオーバグループでの基準で主となるサーバ。 （関連）セカンダリ（サーバ）

用　語	説　明
フローティングIPアドレス	フェイルオーバが発生したとき、クライアントのアプリケーションが接続先サーバの切り替えを意識することなく使用できるIPアドレス。クラスタサーバが所属するLANと同一のネットワークアドレス内で、他に使用されていないホストアドレスを割り当てる。
ま	
ミラーディスクシステム	共有ディスクを使用しないクラスタシステム。 サーバのローカルディスクをサーバ間でミラーリングする。

用　語	説　明
A	
Aware AP	CLUSTERPROと連携するためにCLUSTERPRO APIを使用するアプリケーション。
C	
CLUSTERパーティション	相手サーバの監視を行う、CLUSTERPRO専用パーティション。 (関連)共有パーティション、切替パーティション
F	
FastSync Option	FastSync Optionは、CLUSTERPRO LEで構築されたクラスタシステムにおいて、差分データ部分を復旧することにより、ミラー構築の時間を短縮するためのオプション製品。
P	
PP (Program Product)	有償ソフトウェア製品。
S	
Split-brain-syndrome	→ ネットワークパーティション
V	
Volume Manager	米国VERITAS社のストレージ管理ツール。インタフェース使用により、ストレージ リソース、データの追加/移動等のディスク管理タスクを簡素化できる。共有ディスクのソフトウェアミラー、RAID5を可能にする。 (VERITAS Volume Manager に関する詳細情報は http://www2.bs1.fc.nec.co.jp/vxvm/ を参照願います。)
Volume Managerディスクグループ	VERITAS Volume Managerで作成したクラスタディスクグループであり、CLUSTERPROで使用するリソースの単位。