



ArcSDE 高可用性テクノロジー概要

ESRI システム・インテグレーション
技術資料 (2004 年 12 月)

(原文: 米国 ESRI 社、翻訳: ESRI ジャパン株式会社)

Copyright © 2004 ESRI

All rights reserved.


Printed in Japan.

本書に記載されている内容は、ESRI社 (Environmental Systems Research Institute, Inc.) が独占的に所有するものです。本書は、米国著作権法およびその他の国際著作権条約・協定で保護されています。本書のいかなる部分もESRI社の明確な書面による許可なしには、電子的、機械的 (コピー、記録を含む) などのいかなる手段によっても、またはいかなる情報格納、取得システムにおいても、再生、伝送することはできません。ご質問はすべて以下の担当までお問合せください。

Contracts Manager, ESRI, 380 New York Street, Redlands, CA 92373-8100, USA.

U.S. GOVERNMENT RESTRICTED/LIMITED RIGHTS

Any software, documentation, and/or data delivered hereunder is subject to the terms of the License Agreement. In no event shall the U.S. Government acquire greater than RESTRICTED/LIMITED RIGHTS. At a minimum, use, duplication, or disclosure by the U.S. Government is subject to restrictions as set forth in FAR §52.227-14 Alternates I, II, and III (JUN 1987); FAR §52.227-19 (JUN 1987) and/or FAR §12.211/12.212 (Commercial Technical Data/Computer Software); and DFARS §252.227-7015 (NOV 1995) (Technical Data) and/or DFARS §227.7202 (Computer Software), as applicable. Contractor/Manufacturer is ESRI, 380 New York Street, Redlands, CA 92373-8100, USA.

ESRI, SDE, ArcView, MapObjects, ESRI globe , ArcInfo, ArcSDE, ArcCatalog, ArcEditor, ArcMap, ArcGIS, ArcIMS, ArcStorm, ArcObjects, StreetMap, ArcInfo Librarian, GIS by ESRI, www.esri.com は、米国、欧州およびその他の管轄区におけるESRI社の商標または登録商標です。

その他の社名、商品名、製品名は、各所有者の商標または登録商標です。

本書の一部または全部を無断で転用または複製することを禁じます。

本書に記載されている内容に誤りがあった場合にも、本書を運用した結果の影響についてはいかなる責任も負いかねます。

本書に記載されている内容は、予告無く変更される場合があります。

はじめに

ArcSDE データベースは GIS システムにおいて最も重要な要素の 1 つです。分散したファイルベース (ArcInfo カバレッジやシェープファイル) から中央集中型のジオデータベースへ移行する組織では、ArcSDE データベースの重要性は益々増大しています。空間データおよび属性データの配信ニーズによっては、ArcIMS や ArcGIS Server を経由し、何百もの ArcGIS ユーザが 1 つのデータベースにアクセスすることもあります。ArcSDE サーバもしくは ArcSDE プロセスの障害は、接続しているユーザの業務に大きな影響を与えます。例えば、緊急車両を事故現場までルート案内する 119 番緊急システムの場合、データサーバの障害の影響は多大なものとなります。このような理由から、障害の程度に応じてフォールトトレラントを実現するコンティンジェンシー・プラン (緊急時対策) を、多くの組織が ArcSDE システムの設計時に検討しています。

ArcGIS/ArcSDE 環境の設計の際には、高可用性 (HA: High-Availability) 要件を満たすために考慮すべき事項が多くあります。本技術資料では、ArcSDE で高可用性を実現するために必要な要素、および実装時のオプションについて考察します。また、読者が十分な情報を得た上で構成を決定できるように、オプションの検討に密接に関連する事項についても説明します。エンタープライズ・レベルのシステムでは、トータルのダウンタイムを減らすためには相応の費用がかかります。本技術資料は高可用性を実現するための実手順書ではなく、高可用性システムに関連する概念、および一般的な実装方法について紹介しています。

可用性の高い ArcSDE の実装は、データベース管理ソフトウェアで使用可能な機能に依存しています。DBMS 製品の高可用性オプションについての最新情報を DBMS ベンダーから取得することは良いスタート・ポイントとなります。実装計画に関するより詳細な情報については、DBMS ベンダーに確認することを強くお勧めします。Web サイトには優良な情報が掲載されており、本技術資料の最後でもいくつか紹介しています。ただし、ArcSDE ジオデータベースが単なるデータベースではなく、高可用性の実装を成功させるためには、ArcSDE 独自の制約および検討すべき事項が存在することに注意する必要があります。本資料では特に、これらの事項に注目して説明しています。

本資料では、以下のトピックを対象としています。

- 用語の定義
- 高可用性設計戦略
- ArcSDE/DB 高可用性設計要素
 - データ・ストレージ
 - データベース管理ソフトウェア
 - ArcSDE ミドルウェア
 - ArcSDE/DBMS ネットワーク可用性
- アプリケーション・サーバの考慮事項
 - ArcIMS
 - ArcGIS Server
 - Windows Terminal Server/Citrix
- まとめ
- 高可用性計画に関する参考資料

用語の定義

まず、高可用性に関連する用語の定義を下記に取りまとめます。詳細については、<http://webopedia.com/> などをご参照ください。

フェイルオーバー

プライマリ・システムの障害、もしくはサービスがシャットダウンした際に、スタンバイ・データベース、スタンバイ・サーバ、もしくはスタンバイ・ネットワークへ自動的にスイッチするバックアップ動作。フェイルオーバーは、絶え間ないアクセスが必要なミッション・クリティカル・システムで重要なフォールトトレランス機能です。フェイルオーバーは自動的かつユーザに意識させることなく、障害もしくはダウンしたシステムからバックアップ・システムへプライマリ・システムでの処理をリダイレクトします。

フェイルバック

プライマリ・システムの完全復旧後、スタンバイ・データベース、スタンバイ・サーバ、もしくはスタンバイ・ネットワークから、自動的にスイッチ・バックするバックアップ動作。

フォールトトレランス

ハードウェア、もしくはソフトウェアに予期しない障害が発生した際に正常な動作を保持し続ける能力。多くのフォールトトレランス・レベルがあり、最も低レベルなものとして停電時に処理を続ける能力があります。多くのフォールト・トレラント・コンピュータ・システムはすべての処理をミラー化しています。すなわち、すべての処理が2つ以上の重複システムで実行されるため、1つに障害が発生した際でも他が処理を引き継ぎます。

クラスタリング

複数のコンピュータを相互に接続し、ユーザや他のコンピュータに対して全体で1台のコンピュータであるかのように振舞わせる技術。クラスタリングは、パラレル処理、負荷分散、およびフォールトトレランスを実現するために使用されます。クラスタリングはパラレル処理のアプリケーションで一般的な戦略であり、既存のサーバおよびワークステーションを活用することが可能です。さらに、クラスタに新規サーバを追加して負荷分散させることにより、比較的容易に性能を向上させることができます。

Active/Passive クラスタリング

Active/Passive クラスタリングはデータベース・システムの冗長性を提供する方法です。セカンダリ・システムはホットスタンバイ・モードで、プライマリ・システムの障害時にのみデータベース・トランザクション処理のすべてを開始します。これは ArcSDE ジオデータベースで高可用性を提供する場合も共通の方法で、Microsoft Cluster Server (MSCS)、Veritas Cluster Software、Sun Cluster Software) などのクラスタリング・ソフトウェアを利用して実装します。ArcSDE でテストされておらず広く使用されてはいませんが、いくつかのインプリメンテーション (Informix IDS) ではホットスタンバイを讀取り専用の処理で使用することも可能です。

Active/Active クラスタリング

Active/Active クラスタリングも Active/Passive と同様にシステムの冗長性を提供します。セカンダリ・インスタンスもデータベース (もしくはその他の) サービスを GIS 以外のアプリケーションへ提供します。Active/Active クラスタリングには種類が2つあります。最も共通して使用されている製品では、セカンダリ・プラットフォームは普段、GIS 以外のデータを (ERP やその他の DB アプリケーションへ) サービスしています。この場合、プライマリ・サーバがアクティブな間は、セカンダリ・プラットフォームはプライマリ・データベースへアクセスすることができず、他のデータベースもしくはアプリケーションをサポートするために使用します。もう一方の最新の Active/Active テクノロジー (Oracle RAC) では、複数のインスタンスが同時に同じデータベースへアクセスすることが可能です。これらに関する詳細につきましては、後ほど説明します。

高可用性設計戦略

高可用性を実現するための機能は、ハードウェア、OS ソフトウェア、専用のクラスタリング・ソフトウェア、およびアプリケーションに実装することができます。ミッション・クリティカルなシステムでは、フェイルオーバー機能を備えた冗長な要素を多くの部分で組み込みます。そこまでクリティカルではないアプリケーションでは、ハードウェアもしくはクラスタリング・ソフトウェアのみの実装、もしくは実装しない場合もあります。システム設計内に含める高可用性機能の範囲は、システム障害および計画停止のリスク、もしくはリスク軽減の費用によって変化します。高可用性アーキテクチャについて議論する前に、ジオデータベースのダウンタイムが与える影響について検討し、ターゲットとする高可用性ソリューションの費用対効果分析を行ってください。

高可用性データベース・ソリューション設計のゴールは、妥当な費用で多くの潜在的な単一障害点 (Single point of failure) を排除することです。以下のセクションでは、利用可能な技術の概要をまとめており、比較的シンプルかつ安価なものから説明し、その後、より複雑なソリューションについて紹介しています。各オプションを利用することで可用性と費用は増加します。また、一般的に各オプションは前に紹介する技術の上に構築されます。

冗長サーバ・コンポーネント

ArcSDE の停止を防ぐ第一のラインはサーバ・プラットフォームです。冗長ハードウェア・コンポーネント (二重化した電源装置、ストレージ・コントローラ、ネットワーク・カードなど)、マルチ CPU、および RAID ディスクによりフォールトトレラント DB サーバを構築することができれば、多くの潜在的な単一障害点が排除されます。このような構成では、1 つのコンポーネントに障害があった場合でもサーバは稼動し続けることが可能であり、少ない費用で多くのメリットを得ることができます。ArcSDE/DB 環境ではこの構成を強くお勧めします。

冗長 ArcSDE/DB サーバ

第二のラインは、プライマリ・サーバの障害時に備え、バックアップとして動作するセカンダリ DB プラットフォームを追加することです。このセカンダリ・プラットフォームとして DBMS および ArcSDE がインストールされている検証用 / 開発用サーバを使用することも可能で、プライマリ・システムの障害時にクリティカルなユーザ負荷に耐えることができるサーバを選定しておきます。より堅牢な実装では、セカンダリをプライマリ・サーバと同じハードウェアで構成し、ミラーリングします。この構成の場合、プライマリ・システムの障害時には、ジオデータベースに対するすべての負荷に耐え得ることが可能です。システム障害時の復旧シナリオは、データベースで使用するストレージ・サブシステムへのアクセスをセカンダリ・プラットフォームへ付与、DBMS の起動、プライマリ・システムの IP アドレスをセカンダリ・システムへ移行、適切に動作しているかどうかのテスト、システムをユーザへ開放という手順になります。この処理はハードウェア、ソフトウェア、およびシステム要員の準備状況に依存し、数分から数時間かかります。この構成ではサーバ・ハードウェアの費用は 2 倍かかりますが、システム障害時には数時間以内で復旧することが可能です。冗長サーバが無い場合には、サーバ部品の調達、修理およびテストにより、ダウンタイムが莫大になります。

シェアード・ナッシング型クラスタ ArcSDE/DB サーバ

次に高レベルな構成は、プライマリ・プラットフォームおよびセカンダリ・プラットフォームでシェアード・ナッシング型クラスタを構成することです。このクラスタリングではジオデータベース・サーバの障害を防ぐことはできませんが、自動でプラットフォーム障害を検知してシステム負荷をスタンバイしているセカンダリ・プラットフォームへ移行します。これにより、ArcSDE/DB がすぐにユーザに利用可能であることを保障します。クラスタリングの復旧シナリオは、すべてのフェイルオーバー・タスクが自動で行われることを除き、上記で説明したものに非常に類似しています。一般的に、このフェイルオーバー処理には 20 ~ 90 秒かかります。Active/Active、もしくは Active/Passive のシェアード・ナッシング構成を構築可能で、一度冗長ハードウェアを設置すれば、ソフトウェアおよびハードウェアの追加費用は大きく発生しません。

シェアード・ディスク型クラスタ ArcSDE/DB サーバ

シェアード・ナッシング型クラスタリング (Active/Active および Active/Passive) に加え、異なるレベルの可用性が求められる場合があります。シェアード・ディスク型クラスタでは、各ノードはデータ・ストア内のすべての情報に均等にアクセスします。データ自体ではなく、処理のみが各ノードに分割されます。1 つ以上のサーバに障害が発生した場合でも、すべてのアプリケーションデータは引き続き他のノードで使用可能であり、稼働し続けることができます。複数のサーバがデータベースを更新する場合、データベースの不整合を防ぐため、すべてのノードはメモリ内のデータ構造、システム・テーブル、および一時データ構造を同期する必要があります。この構成を GIS 環境で実現するのは一般的には不可能で、追加費用も発生するため、シェアード・ディスク型クラスタは広く利用されていません。Oracle RAC はこの技術の一例です。

ArcSDE でフェイルオーバー・クラスタ・システムを設計する前に、ArcSDE を構成するシステム要素をレビューすることをお勧めします。

- ArcSDE サービス
 - GIOMGR
 - GSRVR
- DBMS (サードパーティ製: Oracle, DB2, SQL Server など)
- OS (MS Windows 2003 Enterprise Edition, Linux, HP-UX など)
- クラスタリング・ソフトウェア (Veritas Cluster Server, Sun Clustering Software, Oracle TAF and RAC, MS Clustering Service, Linux Clustering など)
- サーバ・プラットフォーム
- ネットワーク接続
- ストレージ・サブシステム
- ストレージ・サブシステムへの接続
- クライアント・ワークステーション
- ArcIMS Server、プロセス、および DB 接続
- ArcGIS Server、プロセス、および DB 接続
- ESRI ライセンス・マネージャ

上記のように、ArcSDE からユーザへのデータ供給には多くの要素が関与しており、高可用性システムの設計時にはこれらの各々を考慮する必要があります。ここでは、本資料の範囲を超えているため、ハードウェアとソフトウェアのすべての組み合わせを対象とはしません。したがって、ハードウェアおよびソフトウェア固有の実装に関する部分については、ハードウェアおよびソフトウェアのベンダーにご相談ください。本資料では、最新の負荷分散技術およびスケールアウト技術について考察することで、主にフェイルオーバー・クラスタリングに焦点を合わせて説明します。

ArcSDE/DB 高可用性設計要素

ArcSDE を構成する上で、保護対象となる主要な要素は次の通りです。

1. データ(データ・ストレージ・サブシステムおよびストレージ内のデータ)
2. ArcSDE が依存している実行中のデータベース管理ソフトウェア、サーバ、OS など
3. ArcSDE サービスと接続
4. ArcSDE/DBMS のネットワークへのアクセス

以下のセクションでは、まず DBMS および ArcSDE サービスに焦点を絞り、上記の分野について説明します。

ArcGIS Desktop および ESRI ArcSDE ソフトウェア・コンポーネントはクラスタ対応(Cluster Aware)ではなく、フェイルオーバー機能が無いことに注意してください。すなわち、クラスタリングやフェイルオーバー機能を実装した場合、その他のシステム要素がこれらの機能を行います。

ここでは災害計画 / 復旧に関する直接的な対応についての説明は控えます。高可用性を実現するためのシステム設計は大惨事に関連するいくつかのリスクを減少し、災害復旧をサポートするために使用することが可能です。バックアップ・ノードが離れた場所に設置してあるシェアード・ナッシング型クラスタは、災害復旧を実装するためには効果的な方法であり、GIS コミュニティ内、特に大規模なインターネットサイトで使用されています。このシナリオでは、データは複数のサイトで複製されており、インターネット・サービス・プロバイダ (ISP: Internet Service Provider) に設置してあるルータを介してクラスタ操作がサポートされています。

1) 最も重要なシステム要素 - データ

ディスク・キャッシュもしくはその他のストレージ・サブシステムの障害による GIS データの損失は大惨事になります。幸いにも、この種の壊滅的な損失を防ぐためのいくつかの戦略が存在します。ホワイトペーパー「System Design Strategies」(<http://www.esri.com/systemsint/kbase/strategies.html>)のセクション 7.6.4 では、ディスク・ストレージ保護のための様々なオプションについて紹介していますが、一般的には RAID5 と RAID1/0 の組み合わせが、システム・パフォーマンスを最大化しコストを最小化する、十分なストレージ保護を提供します。RAID ストレージはデータベース管理システムのためのベストプラクティスとなります。

データベース・レプリケーションもしくはディスク・レプリケーションを、データ保護に使用することも可能です。上記の災害復旧にて示しているように、データのレプリケーションによりデータベースの複製を、ローカルおよび遠隔地で複数管理することができます。レプリケーションはデータ保護の提供に加え、データ・パブリッシング、データ・ウェアハウジング、データ・ディストリビューション、およびその他の様々なビジネス・ファンクションをサポートします。また、ハードウェア・レベルのディスク・ミラーリングも、データベースの複製を維持するために使用されます。ハードウェア・レベルのミラーリングは、ジオデータベースの複製時における特有の潜在的な問題を回避し

ます。

RAID テクノロジーに加え、データ・ストレージ・ハードウェア構成のベストプラクティスには通常、ストレージ・サブシステムと各 DBMS / ファイル・サーバ間の冗長な配線 / 接続、および透過的フェイルオーバー (Transparent Failover) 機能が含まれます。実際のストレージ保護における推奨は、システムの可用性要件とストレージ・アーキテクチャに依存しており、また、GIS インフラストラクチャと GIS アプリケーション・アーキテクチャの全体に大きく依存します。ストレージ保護は災害復旧計画でも検討する必要があります。

2) データベース管理システムの保護

このセクションでは以下を説明します。

- DBMS プラットフォーム構成
- 透過的フェイルオーバー (Transparent Failover)
- フェイルオーバーのシナリオ
- ESRI ソフトウェアの機能
- ESRI ライセンス・マネージャ

DBMS プラットフォーム構成

データの次に最も重要な要素はデータベース管理システムです。サーバやデータベースの再起動(もしくは再構築)、およびすべてのデータベース接続の再確立は複雑な操作であり、多くの時間がかかります。DBMS インスタンスを保護するためのベストプラクティスはシステム・クラスタリングであり、複数の個別のサーバ・プラットフォームを設定に応じて、Active/Active もしくは Active/Passive モードで構成します。クラスタリングを適切に構成するためには、2 つ以上のサーバが必要になります。また、各サーバには DBMS ソフトウェアをインストールし、データベース・テーブルが格納されているストレージ・サブシステムへアクセスできるようにする必要があります。ストレージ・サブシステムとして SAN (Storage Area Network) もしくは DAS (Direct Attached Storage) を使用することができます。クラスタリング環境では NAS (Network Attached Storage) はサポートされていません。複数のプラットフォームが同一の物理ストレージにアクセスする必要があるフェイルオーバー構成において、一般的に最も柔軟性があるのは SAN になります。

Windows 2003 Enterprise Edition、Windows 2003 Datacenter、Windows 2000 Advanced Server、および Windows 2000 Datacenter に含まれる Microsoft Cluster Service (MSCS) は Windows OS と共に動作し、複数の Windows システムをクラスタ (Active/Active もしくは Active/Passive) に構成します。Veritas Cluster Server は UNIX 環境で、Sun Cluster Software は Solaris 環境で同様の機能を提供します。これら製品は基本的なクラスタ環境 (シェアード・ナッシング型) をディスク、IP アドレス、および同時に 1 つのクラスタ・ノードを介してアクセスされるその他のクラスタ・リソースに提供します。データベースもしくはアプリケーションが 1 つのクラスタ・ノードから他ノードへフェイルオーバーした場合、ディスク、IP アドレス、およびデータベースやアプリケーションに関連するその他のクラスタ・リソースは、即座かつ自動で新規ノードへ移行されます。2 つのクラスタ・ノードが同時に同一のディスクに書き込むことはできず (すべてのアクセスがディスクを所有している単一のノードで排他的であるため)、与えられたアプリケーション負荷は複数のクラスタ・ノードへスケール・アウトされることはありません。すなわち、データベース・テーブルはプライマリ・プラットフォームとセカンダリ・プラットフォームから同時にアクセスすることはできません。この種のクラスタリングは GIS コミュニティ内で広く利用されており、フェイルオーバー機能を提供する最も共通の方法です。

クラスタリングには 2 つの特徴があります。最も共通で使用されている製品では、セカンダリ・プラットフォームは普段、GIS 以外のデータやアプリケーション (ERP やその他の DB アプリケーションなど) を供給するために使用されています。プライマリ・プラットフォームがアクティブな間は、セカンダリ・プラットフォームはプライマリ・サーバのジオデータベースにアクセスすることはできませんが、他のデータベースやアプリケーションをサポートするために使用することができます。最新の Active/Active テクノロジである Oracle RAC では、複数の DB インスタンスが同じデータベースに同時にアクセスすることを許可することが可能です。理論的には、Oracle RAC はフェイルオーバー機能に加えてデータベースのスケラビリティと負荷分散機能を提供し、パフォーマンス要件の増加に伴ってサーバ構成を追加することができます。しかしながら、このテクノロジーはジオデータベース環境では広く利用されておらず、ArcSDE 構成にメリットがあるかどうか明確になっていません。このテクノロジーの主要な懸案事項はパフォーマンスです。各サーバはすべてのサーバ上での処理のコピーをメモリ内に維持する必要があります。これにより、特にジオデータベース環境では、多くのオーバーヘッドが生じます。すなわち、軽量のトランザクション・アプリケーションでは良好に動作しますが、GIS 環境ではその限りではありません。Active/Active ソリューションで考慮しなければならないもう 1 つの事項として、Active/Passive では必要の無い DBMS およびその他のソフトウェアのソフトウェア・ライセンス費用があります。

IBM Informix などのその他のデータベース・システムでは、フェイルオーバーやクラスタ機能を提供するいくつかの異なる方法が存在します。例えば、最新の IBM Informix IDS の Active/Active クラスタリングでは、セカンダリ・プラットフォームと同期するために、ディスクのミラーリング、ログ・.shipping、および読取り機能を提供しています。この構成では負荷分散および水平的スケラビリティは利用できず、GIS インプリメンテーションでは広く使用されていません。また、ESRI ではテストを行っていないため、実現可能性と機能性をユーザ自身で評価する必要があります。

透過的フェイルオーバー (Transparent Failover、具体的には Oracle TAF)

透過的フェイルオーバーは Active/Active クラスタリングの形態をとっており、データベース障害後のユーザへの影響を最小限にしつつ、ユーザに業務を継続させることが可能です。これは、高速かつ自動でユーザのデータベース・サービスを存続しているアクティブなデータベース・インスタンスに移行させることで実現されています。透過度 (Degree of Transparency) は、障害発生時に処理されていたサービスのタイプにより異なります。データベース・トランザクションが処理中の場合は、トランザクションはロールバックされ、ユーザはその更新処理を再度実行する必要があります。読み取り専用ユーザはエラーメッセージの受信後、データへのリクエストを再実行します。障害発生時にデータベース・トランザクションの処理中でない読み取り専用ユーザは、障害が発生したことに気づくことはありません。理論的には、処理中のサーバサイドのデータベース・クエリは存続しているデータベース・インスタンスへ移行され、初めから再スタートされます。

このテクノロジーはデータベースに対するクエリが小さく、トランザクションの単位が小さい、軽量のトランザクション環境で効率よく動作します。ArcGIS Desktop ソフトウェア・アーキテクチャでは、複雑なロングトランザクションを利用します。さらに、編集者もしくは参照者は永続的なセッション、およびデータベース接続を確立するため、データベース障害発生時にこれらを行き移すことは容易ではありません。前に簡単に説明しましたが、複数のアクティブなインスタンスを使用した場合のパフォーマンスを妨げるいくつかの技術的な問題があります。このような理由により、RAC/TAF ソリューションは ESRI GIS 環境では広く利用されておらず、ESRI では正式にはサポートしていません。より詳細な情報につきましては、ArcSDE サポートページの下記情報をご覧ください。

<http://www.esri.com/support/arcsde/faq/documents/asfaq00221.html>

他でも説明しましたように、これらベンダー技術は時間と共に移り変わっていくため、採用するアーキテクチャを決定する前にハードウェアおよびソフトウェアのベンダーへ確認し、ソリューションを実装する前に GIS 環境での機能検証に関するベンダーのマーケティング資料を十分に評価することをお勧めします。

フェイルオーバーのシナリオ

Microsoft Cluster Service などの最も一般的に使用されているクラスタリング・ソリューションでは、クラスタがプライマリ・データベースで障害を検知すると、ユーザ作成の自動化されたスクリプトが初期化され、ストレージ・サブシステム上のデータベース、およびプライマリ・データベースの IP アドレスの所有がフェイルオーバー・サーバへ移行されます。その後、データベースの再起動、および完了していないトランザクションのロールバックを行います。障害時には、ArcGIS Desktop ソフトウェアはデータベースへの再接続を試みます。この時点でデータベースおよびその他のリソースが使用可能であれば、ArcGIS Desktop ソフトウェアは通常、自動でデータベースへの接続を再確立します。ArcSDE データベース・ノードでの一般的なフェイルオーバー時間は、20～90 秒の範囲になります。開発者(ユーザ)は、各々の環境における自動化されたフェイルオーバー・スクリプトの作成、およびテストを実行する責務を負います。

障害時のユーザのジオデータベースに対する編集情報は編集セッションから失われ、障害時にデータベースにコミットしていないすべての編集も失われます。この作業を復旧するためには、編集セッションを再スタートし、再度同じ処理を行う必要があります。データベース障害後、参照専用の ArcMap ユーザは MXD への接続を失い、次の処理を実行しようとした場合にエラーメッセージを受信します。フェイルオーバー・インスタンスが利用可能になった後で、MXD ファイルを再ロードする必要があります。

ここで示したすべての情報は Oracle RAC でも同様に関連し、Oracle RAC の場合にはデータベースのフェイルオーバー時間は 20～60 秒から 2～5 秒に短縮されます。

ESRI ソフトウェアの機能

前述のように、ArcGIS Desktop および ESRI ArcSDE ソフトウェアはクラスタ対応(Cluster Aware)ではなく、フェイルオーバーや高可用性機能は組み込まれていません。したがって、クラスタによるアプリケーション機能の監視は OS サービスレベル上で行うか、特別なアプリケーション監視プロセスを介して行います。データベースが正常に動作しているかどうかを、クラスタ監視ソフトウェアにより確認することが可能です。これは、単純に、データベースの状態をデータベース・アプリケーションのプロセスが存在しているかどうかにより判定します。DBMS のクラスタリングおよびその他のフォールトトレラント構成は OS、およびベンダーより提供されているデータベース固有のクラスタリング実装に依存しています。

ESRI ライセンス・マネージャ

高可用性ソリューションを提供する場合には、ESRI ライセンス・マネージャを考慮に入れる必要があります。ArcGIS 8.3 以前では、ユーザがデータベースに接続するためにはライセンス・マネージャ・ファイルがサーバ上に必要であったため、一般的に、アクティブなサーバへのソフトウェアのインストールに加え、アクティブでないセカンダリ・サーバ(フェイルオーバー・サーバ)にもそのコピーを維持する必要がありました。ESRI では通常、一方のみをアクティブに使用するという要件に合意いただければ、ArcSDE の 2 番目のコピーには追加のライセンス使用料を必要としません。このライセンス使用料の詳細につきましては、ESRI ジャパンまでお問い合わせください。クラスタリング・ソフトウェアをライセンス・マネージャのフェイルオーバーを管理するために使用することができます。この構成を選択する場合にも、一度は実環境での検証を行ってください。フェイルオーバー・クラスタでは、FlexLM サービスをフェイルオーバー・スクリプト内に含めることにより、フェイルオーバー・プラットフォームに移行させます。

ArcGIS 9.0 では、ライセンス情報はデータベース内に格納されます。クラスタリング環境では、ライセンス・マネージャ情報は、フェイルオーバー後にセカンダリ・プラットフォームで有効になります。

3) ArcSDE サービスと接続

ArcSDE は、DBMS が動作しているサーバ上のミドルウェア・プロセス、異なるサーバ上のミドルウェア・プロセス (いずれもアプリケーション・サーバ接続モード)、もしくは各クライアント・ワークステーション (ダイレクト・コネクション・モード) やアプリケーション・サーバ (ArcIMS、ArcGIS Server、Windows Terminal Server) 上のプロセスとして動作させることができます。ArcSDE を異なるミドルウェア・プラットフォーム上で動作させる構成は、維持管理がより難しく、構成上の障害点を増すため高可用性ソリューションをより複雑にします。詳細につきましては、下記 URL の技術資料「ArcSDE 階層ハードウェア構成」をご覧ください。

http://www.esri.com/support/document/arcsde/arcsde_tiered_hardwareconfigurations_J.pdf

シェアード・ナッシング型クラスタでは、いずれの接続モード (アプリケーション・サーバおよびダイレクト・コネクション) もデータベース接続で利用でき、フェイルオーバー時のエンドユーザの体感も同等になります。フェイルオーバー・データベースが再起動に成功して利用可能になると、クライアント・ソフトウェアは自動で接続を再確立します。

Oracle9i/10g データベースが Oracle RAC に参加している環境では、クライアント・ワークステーションは ArcSDE データベースへの接続を確立するためにはダイレクト・コネクションを使用する必要があります。図 1 に一般的なシェアード・ナッシング型構成の例を示します。

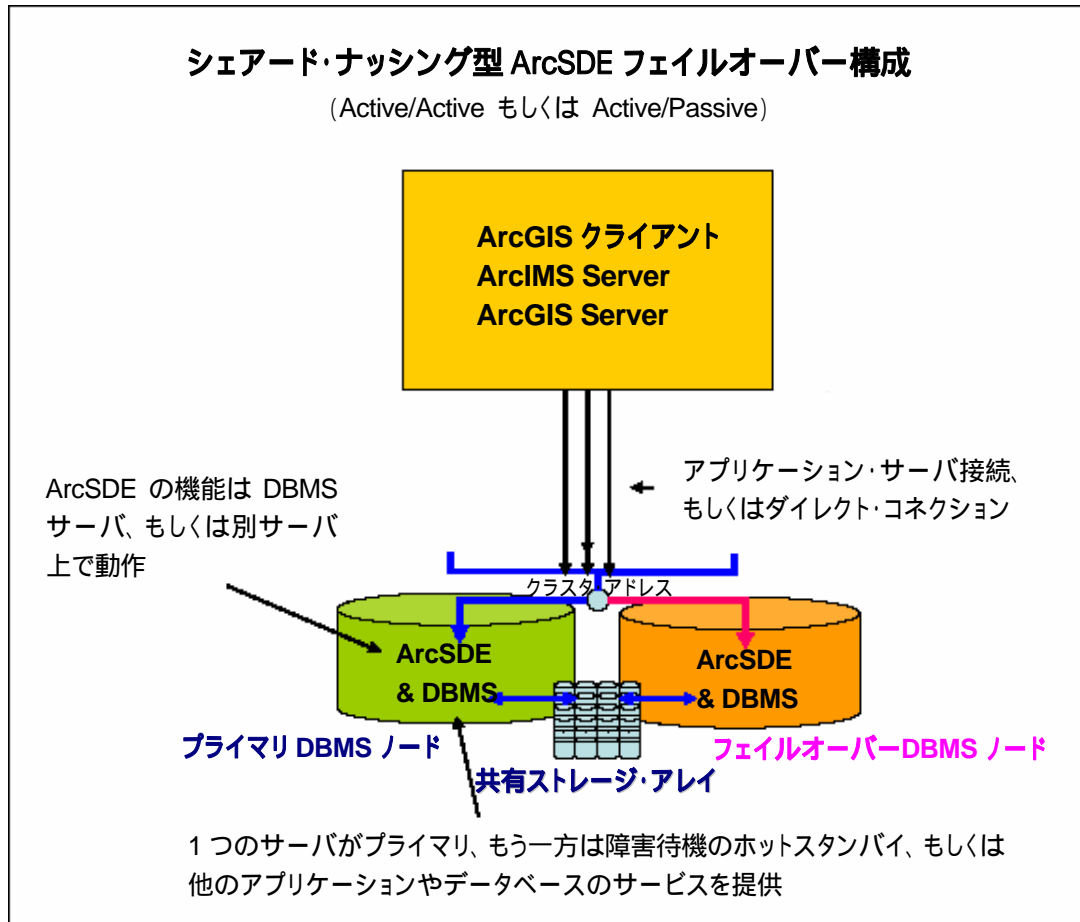


図1 シェアード・ナッシング型 ArcSDE フェイルオーバー構成

ダイレクト・コネクション構成は、ArcSDE/DBMS サーバ構成(アプリケーション・サーバ構成)と比較し、DBMS サーバの CPU 負荷を 20～40 パーセント削減しますが、クライアント・ワークステーション(もしくはアプリケーション・サーバ)の負荷が増え、幾分、全体的なクライアントのレスポンス・タイムが増加します。しかしながら、最新の多くのワークステーションでは、追加の負荷をパフォーマンスに影響を与えることなく高速に処理することができます。負荷の大きなデータベース・サーバと比較的処理能力の高いクライアント・ワークステーションから構成されるサイトでは、ダイレクト・コネクション・モードを使用することで全体的なレスポンス・タイムを改善することができます。以下に、上記で紹介した「ArcSDE 階層ハードウェア構成」に記載されているダイレクト・コネクション・モードのメリットとデメリットを示します。

<メリット>

- GSRVR プロセスをクライアント GIS アプリケーションへ移行させることにより、データベース・サーバの CPU 負荷を 20～40%削減し、データベース・サーバのスケーラビリティを向上
- データベース・サーバのメモリ要件を縮小
- データベース・サーバ上でアプリケーション・ミドルウェア(ArcSDE)を動作させる必要性が無い

<デメリット>

- ArcSDE サーバ/クライアント通信により、アプリケーション・サーバ接続と比較し、ネットワーク・トラフィックが増加(ArcGIS 9.0 では改善)
- 各クライアント PC 上にデータベース・クライアント・ソフトウェアのインストールと構成が必要
- ArcGIS 8.x では、使用するデータベースとデータ型に依存し、クライアントのレスポンス・タイムは約 10～30 パーセント増加

4) ArcSDE/DBMS のネットワークへのアクセス

ArcSDE サーバの高可用性には、共有接続上の冗長なネットワーク・パス(バックボーン、WAN など)が必要になります。ネットワーク要素を利用して、クラスタ障害サイトを構成することも可能です。データベースの可用性で最も重要なネットワーク・リソースは、ArcSDE サーバと他 GIS 関連サーバ(ArcIMS、ArcGIS Server、Windows Terminal Server)間、および ArcSDE とエンタープライズ LAN(Local Area Network)間の接続です。ベストプラクティスは、2 つのネットワーク・インタフェース・カード(NIC)を各サーバにインストールしてフェイルオーバー機能を実装することです。一方の NIC が各ノードの心臓部を提供するためにクラスタ構成で使用され、障害発生時にはフェイルオーバーのトリガーとなります。

ベストプラクティスは、ArcSDE サーバで 100 Mbps や 1 Gbps などの広帯域幅ネットワーク接続を使用することです。

アプリケーション・サーバの考慮事項

ArcIMS

ArcIMS は ArcSDE への接続で、ダイレクト・コネクション・モード、およびアプリケーション・サーバ・モードのいずれも使用することができます。いずれの場合でも、ArcIMS はフェイルオーバー・データベース・サーバが動作を開始すると、データベース接続を再確立します。ArcIMS 自身を Microsoft Cluster Server を用いてクラスタリング環境にセットアップすることも可能で、アプリケーション・サーバおよび空間サーバのフェイルオーバーと負荷分散を行うことができます。ArcIMS の高可用性構成につきましては、ホワイトペーパー「System Design Strategies」(<http://www.esri.com/systemsint/kbase/strategies.html>)の Chapter 4 を参照してください。ArcIMS 保護はネットワーク負荷分散で実装することも可能で、特別なクラスタリング・ソフトウェアを必要としません。技術資料「ArcIMS 4.0.1 High Availability Configuration Testing Using Network Load Balancing」(<http://www.esri.com/systemsint/kbase/archive.html>)では、ArcIMS のネットワーク負荷分散、およびクラスタリングに関する詳細について紹介しています。

ArcGIS Server

ArcGIS Server のような機能を備えたサーバベース・アプリケーションは、高可用性を提供するための構成計画が必要になります。ArcGIS Server は、ArcGIS Server サービスを使用するクライアントに代わり、データベースとサーバ・プラットフォーム間でセッションを確立します。クラスタ・フェイルオーバーを介したパフォーマンスはトランザクションベースの ArcIMS サービスに似ており、ArcGIS Server のユーザ・ワークフロー・セッションは ArcGIS Desktop の接続に類似しています。ArcGIS Server を、高可用性およびアプリケーション負荷分散を行うように構成することも可能です。ArcGIS Server の高可用性構成につきましては、ホワイトペーパー「System Design Strategies」(<http://www.esri.com/systemsint/kbase/strategies.html>)の Chapter 4 を参照してください。

Windows Terminal Server/Citrix

ArcGIS Desktopソフトウェアが動作するWindows Terminal Server(WTS)は、ダイレクト・コネクション・モードとアプリケーション・サーバ・モードのいずれも利用することが可能です。WTSサーバは通常、広帯域LAN上にデータベース・サーバと共に設置されているため、ダイレクト・コネクション利用時のネットワーク遅延はほとんど問題となりません。ダイレクト・コネクションを使用することでWTSサーバに追加の負荷を与えることとなりますが、ユーザはパフォーマンスの劣化をほとんど感じることはありません。WTSおよびCitrixを用いた高可用性および負荷分散構成についての詳細情報につきましては、ベンダーの資料を参照してください。

まとめ

フェイルオーバー機能を提供するテクノロジーの選択は、特定のハードウェア、OS およびデータベース・システムの環境、およびフェイルオーバーで必要とされる機能に大きく依存します。特定のソリューションの採用を決定する前に、これらの各分野について十分調査し、よく理解しておくことが重要です。ESRI ジオデータベース環境における(低リスクで)最適なソリューションを証明しました。最新テクノロジーの検討時には、ArcSDE 環境でその機能が動作することを保障するために、システム設計を終了する前に十分に検証する必要があります。シェアード・ナッシング型クラスタリングは、現状、高可用性ソリューションの実装で最もリスクの少ないテクノロジーです。しかしながら、常に最新テクノロジーが生まれ、提供されているため、時間と共にジオデータベースを使用する際の最適な高可用性ソリューションは変化しています。

高可用性 ArcSDE 構成は以下の要素を含みます。

- DBMS サーバには、冗長な電源、NIC、RAID ストレージなどのフォールトトレラント機能を組み込みます。
- DBMS サーバのインスタンスは、少なくとも2つのサーバから構成されるサーバ・クラスタの一部とします。
- DBMS サーバを高速で冗長なネットワーク・パスに接続します。
- 各クラスタ・サーバには同じDBMSソフトウェアをインストールします。
- データベース・テーブル格納のための共有ストレージ・サブシステムを、クラスタの各ノードから利用可能にします。
- 各クラスタ・サーバを冗長な高速パス(例えば、デュアル・ホスト・バス・アダプタ)を用いてストレージ・サブシステムに接続します。
- クラスタリングの使用時には、セカンダリ・プラットフォームに自動でリソースの移行、およびプロセスの再起動を行うフェイルオーバー・スクリプトを開発者(ユーザ)が作成します。

高可用性計画に関する参考資料

- <http://www.microsoft.com/ntserver/productinfo/enterprise/>
Microsoft Cluster Service(MSCS)、MS SQL Server、ネットワーク負荷分散、およびその他のWindowsベースの高可用性テクノロジーに関する情報
- <http://h18000.www1.hp.com/solutions/enterprise/highavailability/oracle/rac-overview.html>
HP/Oracle RAC アーキテクチャとその考慮事項に関する情報
- <http://www.oracle.com/technology/deploy/availability/pdf/MAAoverview.pdf>
Oracle の高可用性に関する資料

- http://www-306.ibm.com/software/data/informix/pubs/whitepapers/HA_f.pdf
IBM Informix IDS の高可用性オプションについての情報
- <http://www-106.ibm.com/developerworks/db2/library/techarticle/0304wright/0304wright.html>
IBM DB2 可用性および災害復旧についての概要
- http://www.esri.com/systemsint/kbase/docs/sql_server_2000_mscs_configuration.pdf
ESRI 検証レポート - MS SQL Server 2000 MSCS 構成

ArcSDE 高可用性テクノロジー概要

2004 年 12 月

発行/編集所 ESRI ジャパン株式会社

<http://www.esrij.com/>

- ・ 本書は、ESRI 社(米国 カリフォルニア州)が作成した「ArcSDE High-Availability Overview, ESRI Systems Integration Technical Brief」を ESRI ジャパン株式会社(日本 東京)が翻訳した技術文書です。
- ・ 原本および本書の著作権、著作権は ESRI 社に帰属しております。
- ・ 本書の無断複製は禁じます。
- ・ 本書に記載されている社名、商品名は、各社の商標および登録商標です。
- ・ 本書に記載されている内容は改良のため、予告なく変更される場合があります。