



Oracleの構成

Enterprise applications

NetApp
December 17, 2024

目次

Oracleの構成	1
概要	1
Oracleシングルインスタンス	1
Oracle拡張RAC	2
RAC Tiebreaker	4

Oracleの構成

概要

SnapMirrorアクティブ同期を使用しても、データベースの運用に関するベストプラクティスが追加または変更されとは限りません。

最適なアーキテクチャは、ビジネス要件によって異なります。たとえば、データ損失に対するRPO=0の保護が目標であるにもかかわらず、RTOが緩和されている場合は、Oracleのシングルインスタンスデータベースを使用し、SM-ASでLUNをレプリケートすれば十分であり、Oracleのライセンスの問題からより安価になる可能性があります。リモートサイトに障害が発生しても運用は中断されず、プライマリサイトが停止すると、サブサイートのLUNはオンラインで使用可能な状態になります。

RTOの方が厳しい場合は、スクリプトやPacemakerやAnsibleなどのクラスタウェアを使用した基本的なアクティブ/パッシブ自動化を使用すると、フェイルオーバー時間が短縮されます。たとえば、プライマリサイトでVMの障害を検出し、リモートサイトでVMをアクティブにするようにVMware HAを設定できます。

最後に、きわめて高速なフェイルオーバーを実現するために、Oracle RACを複数のサイトに導入できます。データベースは常にオンラインで両方のサイトで利用できるため、RTOは基本的にゼロになります。

Oracleシングルインスタンス

以下に説明する例は、SnapMirrorアクティブ同期レプリケーションを使用してOracleシングルインスタンスデータベースを導入するための多数のオプションの一部を示しています。

[不統一アクセスを使用するOracle SI]

事前設定されたOSを使用したフェイルオーバー

SnapMirror Active Syncはディザスタリカバリサイトにデータの同期コピーを作成しますが、そのデータを利用できるようにするには、オペレーティングシステムと関連するアプリケーションが必要です。基本的な自動化により、環境全体のフェイルオーバー時間を大幅に短縮できます。PacemakerなどのClusterware製品は、サイト間でクラスタを作成するためによく使用されます。多くの場合、フェイルオーバープロセスは単純なスクリプトで実行できます。

プライマリノードが失われると、クラスタウェア（またはスクリプト）によって代替サイトでデータベースがオンラインになります。1つは、データベースを構成するSANリソース用に事前設定されたスタンバイサーバを作成する方法です。プライマリサイトに障害が発生すると、クラスタウェアまたはスクリプト化された代替サイトが次のような一連の処理を実行します。

1. プライマリサイトの障害を検出
2. FCまたはiSCSI LUNの検出の実行
3. ファイルシステムのマウント、ASMディスクグループのマウント
4. データベースの起動

このアプローチの主な要件は、リモートサイトでOSを実行することです。Oracleバイナリを使用して事前に

設定する必要があります。つまり、Oracleのパッチ適用などのタスクをプライマリサイトとスタンバイサイトで実行する必要があります。また、災害が発生した場合は、Oracleバイナリをリモートサイトにミラーリングしてマウントすることもできます。

実際のアクティベーション手順は簡単です。LUN検出などのコマンドでは、FCポートあたりのコマンド数が少なくて済みます。ファイルシステムのマウントはコマンドにすぎませ `mount` ン。データベースとASMの両方を、1つのコマンドでCLIから開始および停止できます。

仮想OSによるフェイルオーバー

データベース環境のフェイルオーバーを拡張して、オペレーティングシステム自体を含めることができます。理論的には、このフェイルオーバーはブートLUNで実行できますが、ほとんどの場合、仮想OSで実行されます。手順の手順は次のようになります。

1. プライマリサイトの障害を検出
2. データベースサーバ仮想マシンをホストするデータストアのマウント
3. 仮想マシンの起動
4. データベースを手動で起動するか、仮想マシンでデータベースが自動的に起動するように設定します。

たとえば、ESXクラスタが複数のサイトにまたがっているとします。災害が発生した場合は、スイッチオーバー後にディザスタリカバリサイトで仮想マシンをオンラインにすることができます。

ストレージ障害からの保護

上の図は"[不均一なアクセス](#)"、の使用方法を示しています。SANが複数のサイトにまたがっているわけではありません。これは設定が簡単で、現在のSAN機能では唯一の選択肢となる場合もありますが、プライマリストレージシステムに障害が発生すると、アプリケーションがフェイルオーバーされるまでデータベースが停止します。

耐障害性を高めるために、このソリューションをとともに導入することもでき"[均一なアクセス](#)"ます。これにより、アプリケーションは、反対側のサイトからアドバタイズされたパスを使用して動作を継続できます。

Oracle拡張RAC

多くのお客様が、Oracle RACクラスタを複数のサイトにまたがって構成し、完全なアクティブ/アクティブ構成を実現することで、RTOを最適化しています。Oracle RACのクォーラム管理を含める必要があるため、設計全体が複雑になります。

従来の拡張RACクラスタでは、ASMミラーリングを使用してデータを保護していました。このアプローチは機能しますが、多くの手動設定手順が必要になり、ネットワークインフラストラクチャにオーバーヘッドが発生します。一方、SnapMirrorのアクティブな同期機能でデータレプリケーションを実行できるようにすることで、ソリューションが大幅に簡易化されます。同期、中断後の再同期、フェイルオーバー、クォーラム管理などの操作が容易になります。また、SANの設計と管理を簡素化するために、SANをサイト間に分散させる必要もありません。

レプリケーション

SnapMirrorアクティブ同期のRAC機能を理解するには、ストレージをミラーリングされたストレージでホストされている単一のLUNセットとして表示することが重要です。例：

[Oracle論理アクセス]

プライマリコピーまたはミラーコピーはありません。論理的には、各LUNのコピーは1つだけで、そのLUNは2つの異なるストレージシステム上にあるSANパスで使用できます。ホストから見ると、ストレージフェイルオーバーは発生せず、代わりにパスが変更されます。さまざまな障害イベントが発生すると、LUNへの特定のパスが失われても、他のパスはオンラインのままになる可能性があります。SnapMirrorのアクティブな同期により、すべての運用パスで同じデータを利用できるようになります。

ストレージ構成

この構成例では、ASMディスクは、エンタープライズストレージの単一サイトRAC構成と同じように設定されています。ストレージシステムはデータ保護を提供するため、ASM外部冗長性が使用されます。

ユニフォームアクセスと非インフォームアクセス

SnapMirrorアクティブ同期上のOracle RACで最も重要な考慮事項は、均一アクセスと非均一アクセスのどちらを使用するかです。

アクセスが統一されているため、各ホストは両方のクラスタのパスを認識できます。一様でないアクセスとは、ホストがローカルクラスタへのパスのみを認識できることを意味します。

どちらのオプションも特に推奨または推奨されないものではありません。ダークファイバを使用してサイトを接続しているお客様もいれば、そのような接続を利用していないお客様や、SANインフラで長距離ISLをサポートしていないお客様もいます。

不均一なアクセス

アクセスが一様でない場合、SANの観点からはより簡単に設定できます。

[Oracle RACの不均一アクセス]

このアプローチの主な欠点"不均一なアクセス"は、サイト間のONTAP接続が失われたり、ストレージシステムが失われたりすると、一方のサイトのデータベースインスタンスが失われることです。これは明らかに望ましくありませんが、シンプルなSAN構成と引き換えに許容可能なリスクになる可能性があります。

均一なアクセス

アクセスを統一するには、SANをサイト間に拡張する必要があります。主なメリットは、ストレージシステムが停止してもデータベースインスタンスが失われないことです。その結果、現在使用されているパスがマルチパスに変更されます。

不均一アクセスを設定するには、いくつかの方法があります。



次の図には、単純なコントローラ障害時に使用されるアクティブだが最適化されていないパスもありますが、この図では省略しています。

近接設定を使用したAFF

サイト間のレイテンシが大きい場合は、ホストとの近接設定を使用してAFFシステムを設定できます。これにより、各ストレージシステムはどのホストがローカルでどのホストがリモートであるかを認識し、パスの優先順位を適切に割り当てることができます。

[アクセスが統一されたRAC]

通常運用時は、各Oracleインスタンスがローカルのアクティブ/最適パスを優先的に使用します。その結果、すべての読み取りはブロックのローカルコピーによって処理されます。これにより、レイテンシが最小限に抑えられます。書き込みIOも同様に、ローカルコントローラへのパスに送信されます。IOは確認される前にレプリケートする必要がありますが、その場合もサイト間ネットワークを通過するための追加のレイテンシが発生しますが、同期レプリケーションソリューションではこれを回避することはできません。

近接設定なしのASA / AFF

サイト間のレイテンシがそれほど高くない場合は、ホストとの近接を設定せずにAFFシステムを構成するか、ASAを使用できます。

[アクセスが統一されたRAC]

各ホストが両方のストレージシステムのすべての動作パスを使用できるようになります。これにより、各ホストが1つだけでなく2つのクラスタの潜在的なパフォーマンスを利用できるようになるため、パフォーマンスが大幅に向上します。

ASAでは、両方のクラスタへのすべてのパスがアクティブで最適化されているとみなされるだけでなく、パートナーコントローラのパスもアクティブになります。その結果、常にクラスタ全体でオールアクティブなSANパスが作成されます。



ASAシステムは、不均一なアクセス設定でも使用できます。サイト間パスは存在しないため、IOがISLを経由してもパフォーマンスに影響はありません。

RAC Tiebreaker

SnapMirrorアクティブ同期を使用した拡張RACはIOに関して対称アーキテクチャですが、スプリットブレイン管理に接続される例外が1つあります。

レプリケーションリンクが失われ、どちらのサイトにもクォーラムがない場合はどうなりますか。何が起るべきでしょうか? この質問は、Oracle RACとONTAPの両方の動作に当てはまります。サイト間で変更をレプリケートできない場合に運用を再開するには、一方のサイトを停止し、もう一方のサイトを使用できなくなる必要があります。

は、**"ONTAPメディアエーター"**ONTAPレイヤでこの要件に対応します。RACのタイブレークには複数のオプションがあります。

Oracleタイブレーカー

Oracle RACのスプリットブレインリスクを管理する最善の方法は、奇数のRACノードを使用すること（できれば3つ目のサイトのTiebreakerを使用すること）です。3つ目のサイトが使用できない場合は、Tiebreakerインスタンスを2つのサイトの一方のサイトに配置して、優先サバイバーサイトに指定できます。

OracleおよびCSS_CRITICAL

ノード数が偶数の場合、Oracle RACのデフォルトの動作では、クラスタ内のいずれかのノードの重要度が他のノードよりも高くなります。優先度の高いノードを含むサイトはサイト分離を継続し、もう一方のサイトのノードは削除されます。優先順位付けは複数の要因に基づいて行われますが、設定を使用してこの動作を制御することもできます `css_critical`。

"例"アーキテクチャでは、RACノードのホスト名はjfs12およびjfs13です。の現在の設定は`css_critical`次とおりです。

```
[root@jfs12 ~]# /grid/bin/crsctl get server css_critical
CRS-5092: Current value of the server attribute CSS_CRITICAL is no.

[root@jfs13 trace]# /grid/bin/crsctl get server css_critical
CRS-5092: Current value of the server attribute CSS_CRITICAL is no.
```

jfs12が設定されたサイトを優先サイトにする場合は、サイトAのノードでこの値をyesに変更し、サービスを再起動します。

```
[root@jfs12 ~]# /grid/bin/crsctl set server css_critical yes
CRS-4416: Server attribute 'CSS_CRITICAL' successfully changed. Restart
Oracle High Availability Services for new value to take effect.

[root@jfs12 ~]# /grid/bin/crsctl stop crs
CRS-2791: Starting shutdown of Oracle High Availability Services-managed
resources on 'jfs12'
CRS-2673: Attempting to stop 'ora.crsd' on 'jfs12'
CRS-2790: Starting shutdown of Cluster Ready Services-managed resources on
server 'jfs12'
CRS-2673: Attempting to stop 'ora.ntap.ntappdb1.pdb' on 'jfs12'
...
CRS-2673: Attempting to stop 'ora.gipcd' on 'jfs12'
CRS-2677: Stop of 'ora.gipcd' on 'jfs12' succeeded
CRS-2793: Shutdown of Oracle High Availability Services-managed resources
on 'jfs12' has completed
CRS-4133: Oracle High Availability Services has been stopped.

[root@jfs12 ~]# /grid/bin/crsctl start crs
CRS-4123: Oracle High Availability Services has been started.
```

著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を随時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5252.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および/または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用权を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用权については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/TM>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。