

HP Operations Orchestration Software

ソフトウェアバージョン : 7.50

Oracle Real Application Clusters 用 HP OO インストールの 手引き

ドキュメントリリース : 2009 年 3 月

ソフトウェアリリース : 2009 年 3 月



ご注意

保証

HP の製品およびサービスの保証は、かかる製品およびサービスに付属する明示的な保証の声明において定められている保証に限ります。本ドキュメントに記載されたいかなる内容も、追加の保証を構成するものではありません。当社は、本ドキュメントに技術上の誤り、編集上の誤り、記載漏れがあった場合でも責任を負わないものとします。

本ドキュメントに記載した情報は、予告なしに変更することがあります。

制限付き権利

本ドキュメントで取り扱っているコンピューターソフトウェアは秘密情報であり、その保有、使用、または複製には、HP からの有効なライセンスが必要です。FAR 12.211 および 12.212 に従って、商業用コンピューターソフトウェア、コンピューターソフトウェア資料、および商業用製品の技術データは、ベンダー標準の商業用ライセンス条件に基づいて米国政府にライセンスされています。

著作権

© Copyright 2009 Hewlett-Packard Development Company, L.P.

商標

本ドキュメントで言及されているすべての商標は、それぞれの所有者の財産です。

ヘルプ、チュートリアルなどの場所

HP Operations Orchestration Software (HP OO) のドキュメントセットは次の内容で構成されています。

- Central のヘルプ
Central のヘルプには次の情報が記載されています。
 - フローの検索と実行
 - HP OO の機能の設定 (HP OO 管理者向け)
 - フローの実行結果から利用可能な情報の生成と表示Central のヘルプシステムは PDF ドキュメントとしても参照できます。このファイルは、HP OO のホームディレクトリ (¥Central¥docs 内) にあります。
- Studio のヘルプ
Studio のヘルプは、プログラミングの初心者から上級者まで幅広く対応する、フロー作成のための手引きです。Studio のヘルプシステムは PDF ドキュメントとしても参照できます。このファイルは、HP OO のホームディレクトリ (¥Studio¥docs ディレクトリ内) にあります。
- Central 用および Studio 用の動画チュートリアル
HP OO のチュートリアルはどちらも 30 分未満で終わるもので、次の内容について基本的な説明をします。
 - Central : フローに基づく情報の検索、実行、表示
 - Studio : フローの編集チュートリアルは、次の場所で利用できます。
 - [Studio Welcome] ペイン
 - HP OO¥Studio ホームディレクトリ内の Tutorials サブディレクトリ
 - Opware Network
- HP OO のオペレーション、フローおよび Accelerator Pack についてのセルフドキュメンテーション
セルフドキュメンテーションは、フローに含まれているオペレーションやステップの説明で見ることができます。

インターネット上でのドキュメントの検索または更新

ヒューレット・パッカーのソフトウェアのドキュメントは継続的に拡充が図られています。HP OO のドキュメントセットやチュートリアルは、HP サポートの Web サイトでいつでも入手または更新できます。

HP OO のドキュメントおよびチュートリアルを入手するには

1. HP Software 製品マニュアルの Web サイト (<http://support.openview.hp.com/selfsolve/manuals>) に移動します。
2. HP パスポートのユーザー名とパスワードを入力してログインします。
もしくは
HP パスポートを取得していない場合は、[New users — please register] をクリックし、HP パスポートを作成してからこのページに戻ってログインします。
HP パスポートの取得についてご不明な点がある場合は、HP OO の窓口でご確認ください。
3. [製品] リストボックスで下にスクロールし、[Operations Orchestration] を選択します。

4. [製品バージョン] リストで、目的のマニュアルのバージョンをクリックします。
5. [オペレーティングシステム] リストで、該当するオペレーティングシステムをクリックします。
6. [検索] ボタンをクリックします。
7. [検索結果] リストで、必要なファイルのリンクをクリックします。

サポート

パッチ、トラブルシューティング支援、サポート契約の管理、製品マニュアル、その他のサポート情報に関しては、次のサイトを参照してください。

- <http://support.openview.hp.com>

目次

保証	ii
制限付き権利	ii
商標	ii
ヘルプ、チュートリアルなどの場所	iii
インターネット上でのドキュメントの検索または更新	iii
サポート	iv
概要	1
HP OO の設定	2
Linux	2
Windows	3
接続フェイルオーバー	3
パフォーマンス	3
HP OO Scheduler	4
トラブルシューティング	4
JDBC 接続テスト	4
RAC データベース	4
ONS	4
接続ロードバランシング	5
有用な情報	5

概要

Oracle Real Application Clusters (RAC) テクノロジーは、複数の Oracle データベースノードによって作業負荷を分散し、障害時に自動フェイルオーバーを実行する高可用性ソリューションです。これは、クラスタ内のすべてのノードを最新の状態に維持するための共有ストレージと高速のプライベートインターコネクต์に依存したホット-ホット設定です。

Oracle 自体は共有ストレージの高可用性を管理せず、SAN や NAS などのほかの手段でストレージの可用性を高める必要があります。また、クラスタ内のすべてのノードは互いに近接しているため、RAC テクノロジーはディザスタリカバリーに適していません。Oracle DataGuard などのその他のテクノロジーを使用すると、お互いに離れた場所に設置されたデータベース間でホットスタンバイ設定が可能です。

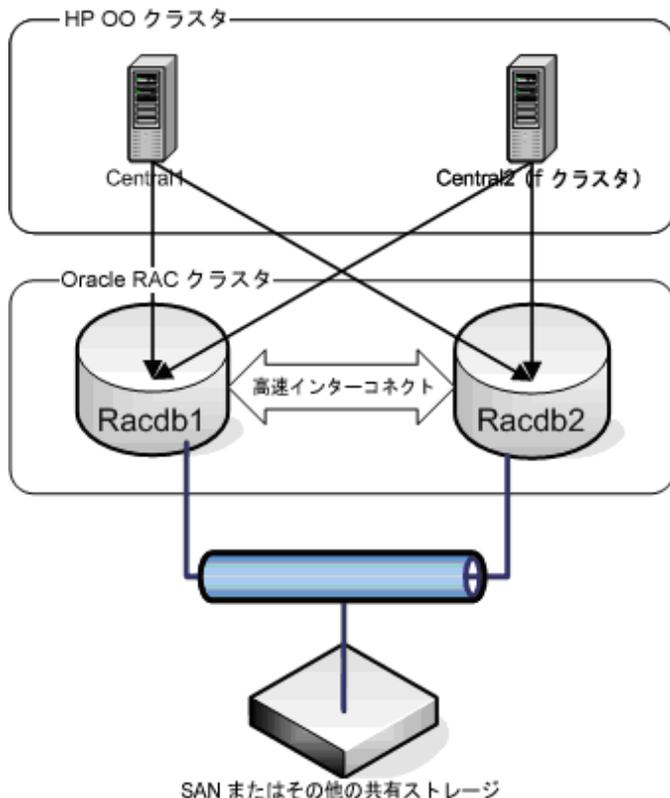


図 1 – RAC 構成の例

HP OO は、データベースとして Oracle クラスタを使用し、高速接続フェイルオーバーを実行するように構成できます。HP OO では、ロードバランスされた接続プールとして Oracle JDBC ドライバーを使用し、失敗したトランザクションの再試行を新しい接続を介して行い、データベースノードの障害を処理します。これによりフローを最小の時間遅延で実行し、完了できます。

HP OO の設定

Oracle クラスタとしてデータベースを認識し、高速接続フェイルオーバーを可能にするために、HP OO に必要な追加の設定ステップがいくつかあります。

標準的な Oracle インストールとの設定上の重要な相違点は、接続するホストに標準的なホスト名の代わり仮想 IP アドレスを指定する必要があることです。仮想 IP アドレスはクラスタソフトウェアによって管理され、ノードの仮想 IP アドレスはノード障害時に自動的に別のノードへ転送されます。したがって、データベースホストを設定する際は仮想 IP アドレスのみを使用する必要があります。

クラスタ環境にとって特に重要な項目の 2 番目は、Oracle Notification Service (ONS) イベントに対するサブスクリプトを HP OO に設定することです。ONS とは、高速アプリケーション通知 (Fast Application Notification、FAN) イベントと呼ばれるイベントを生成するクラスタ上で実行されるサービスです。これらのイベントは、クラスタ内のノードのサービスステータスに変更が発生するたびに生成されます (たとえば、ノードの停止や再起動など)。また ONS は、レシーバーがクラスタ内の最適なサービスに要求をリダイレクトするのに利用できるようロードバランシングアダプザリイベントをパブリッシュします。これらのイベントの利点は、クライアントがクラスタノードのステータスを判断するのに、標準的な TCP タイムアウトイベント (数分掛かる場合があります) を待つ必要がないことです。

Linux

install.config ファイルを使用して Linux ディストリビューションを設定します。標準的 (1 つのノード) なインストールでは、DB_HOST および DB_PORT で単一インスタンスの Oracle のホストおよびポートを指定します。RAC のインストールでは、これらの設定は使用されず、値を入力した場合も無視されます。その代わりに、ORACLE_RAC_NODES および ORACLE_RAC_ONS_CONFIG を設定する必要があります。これらの例は次のとおりです。

```
ORACLE_RAC_NODES=racdb1-vip:1521,racdb2-vip:1521
ORACLE_RAC_ONS_CONFIG=racdb1-vip:6200,racdb2-vip:6200
```

racdb1-vip には 1 つ目のクラスタノードの仮想 IP アドレス、また racdb2-vip には 2 つ目のクラスタノードの仮想 IP アドレスを入力し、3 つ以上のノードを設定する場合は各ノードの仮想 IP アドレスをさらに追加します。Oracle (10g リリース 2) のデフォルトのポートは、インスタンスが 1521、ONS が 6200 です。ONS ポートは各ノードの \$ORA_CRS_HOME/opmn/conf/ons.config ファイルの remoteport という設定で確認できます。

RACDB などクラスタデータベースの Oracle サービス名には、DB_NAME 設定を使用する必要があります。

これらの設定に基づき、インストーラーが作成する JDBC URL は次のようになります。

```
jdbc:oracle:thin:@
(DESCRIPTION=
  (LOAD_BALANCE=on)\
  (ADDRESS_LIST=
    (ADDRESS=(PROTOCOL=TCP) (HOST=racdb1-vip) (PORT=1521))\
    (ADDRESS=(PROTOCOL=TCP) (HOST=racdb2-vip) (PORT=1521))\
  (CONNECT_DATA=(SERVICE_NAME=RACDB)))
```

このタイプの接続文字列によって、アドレスリスト内のノード間でロードバランシングを使用して、RACDB サービスを使用するように JDBC ドライバーに伝達されます。

Windows

Windows のインストーラーは、リリース 7.50 の RAC インストレーションをサポートしていません。ただし、ノードの 1 つでインストーラーをターゲットにしてから上記のような接続文字列を作成し、このような接続文字列が必要になるたびに置き換えることが可能です。1 つのノードのみに製品をインストールするとデータベースが共有されるため上手く動作します。その結果、表は両方のノードに表示されます。難点としては、JDBC URL が使用されていて現在の値を長い接続文字列に置き換える場所をすべて特定することです。対象となるファイルは次のとおりです。

```
%ICONCLUDE_HOME%\Central\WEB-INF\classes\hibernate.properties
```

```
%ICONCLUDE_HOME%\Central\conf\Central.properties
```

接続フェイルオーバー

HP OO が RAC と連携するように設定されている場合、JDBC 接続プールは FAN イベントを受信して障害の発生したノードやインスタンスに対して作成された接続を終了します。ノードやインスタンス障害が発生した時点で実行中であったトランザクションはすべて除外されます。

HP OO は JDBC 接続プールによって送信されたエラーコードを認識し、失敗したトランザクションを 10 回まで、`dharma.hibernate.persistor.connectionRetryTimeSeconds` という設定変数によって指定された再試行間のタイムラグを使用して再試行します。このタイムラグは、デフォルトでは 10 秒です。フェイルオーバーに 100 秒以上かかると予想される場合は、フェイルオーバーに必要な時間に合わせてこのパラメーターの値を大きくできます（完全にノードが停止していること検出するには、Oracle CRS は 1 秒間隔で約 60 回再試行する必要があります）。

JDBC 接続チャンネルの通信障害を検出するたびに、HP OO がトランザクションを再試行することに注意してください。これにはネットワーク接続障害（ノード障害の場合とは異なる）の場合も含まれます。この場合、接続障害が発生した時間によっては、再試行されたトランザクションが成功しないことがあります。たとえば、コミットが開始され、データベースノードに対して成功したのに、そのリプライが返らなかった場合、繰り返されるトランザクションはデータベースの整合性制約違反またはバージョン競合が原因で失敗している可能性が高いといえます。もう 1 つのリカバリ不能障害の例としては、ノードのトランザクションが失敗して、正常なノードに接続している別のユーザーが、リプレイされるトランザクションで不整合となる更新を行った場合、同じ状況が単一ノード構成で発生すると同様に、HP OO で実行可能な処理はありません。

パフォーマンス

RAC を使用すると接続ロードバランシングによるスケーラビリティを得られるという利点があります。しかし、クラスタ内のノードが共通のデータベース（ファイルのセット）を共有して、プライベートインターコネクトを介した通信量が増加することが原因となり、問題が生じます。

クラスタ内の各インスタンス（Oracle プロセスのセット）は共通のデータベースから読み取りおよび書き込みができます。各インスタンスは別個の SGA（System Global Area）を持ち、個別のシステムのようにシステムの実行に必要なメモリ内データ構造を含んでいます。ただし、個別の設定とは異なり、キャッシュフュージョンと呼ばれるテクノロジーによって各ノードのバッファキャッシュを「融合」し、分散型のグローバルバッファキャッシュを作成できます。クラスタノードがブロックを要求する場合、このグローバルキャッシュでブロックを検索します。要求ノードで利用できないブロックは、ブロックを保持しているノードからプライベートインターコネクトを介して転送する必要があります。ブロックの移動によって生成されるトラフィックに加え、複雑な同期とロックが発生します。

したがって、パフォーマンスは単に共有クラスタファイルシステムとプライベートインターコネクトのパフォーマンスで測定できません。その他の複雑な相互作用があります。たとえば、順序番号の生成など一見すると小さなタスクでも、ノードがジェネレーターに対してロックを取得するのに競合するため、複雑になります。

HP OO Scheduler

HP OO Scheduler は Quartz サードパーティアプリケーション (<http://www.opensymphony.com/quartz/>) を管理します。
<http://jira.opensymphony.com/browse/QUARTZ-341> で特定されたバグの解決策として、Scheduler を多数のノード上でロードバランスしないことが推奨されます。

したがって、RAC 仮想 IP (たとえば racdb1-vip) を 1 つだけ使用して Scheduler を設定する必要があります。
仮想 IP は障害時に利用可能なノードへ転送されるため、Scheduler は高可用性を維持します。

前の Oracle 用スケジューラー実装からのもう 1 つの変更点は、過去のバージョンでは接続が必要になるたびに新しい接続を作成したのに対して、接続プールが使用されていることです。これにより、データベースアクセスの効率が高まりますが、プールサイズを適切に設定する必要があります (\$SCHEDULER_HOME/Scheduler/WEB-INF/jdbcDatasource.xml ファイルを編集して、MaxLimit プロパティを変更します)。

Quartz スケジューラーは現在、トランザクションの途中でノードが失敗した場合、Central のように自動フェイルオーバーを実行するようコーディングされていません。したがって、ノード障害時にスケジューラーが失敗する場合があります。しかし、次のトランザクションのために稼働中のノードへ VIP が自動的に転送されるため、Scheduler は動作し続ける必要があります。

トラブルシューティング

JDBC 接続テスト

TestOracleConnection.jar ツールを使用すると、クラスタ内の各ノードに対する接続に加え、ロードバランス接続をテストできます。たとえば、nodes=racdb1-vip:1521,racdb2-vip:1521 のように nodes パラメーターを指定します。これにより、ツールは racdb1-vip, racdb2-vip および両方のホストで構成される接続文字列に対する接続を個別にテストします。

また、このツールは次のパラメーターを実行することにより ONS イベントへのサブスクライブに使用できます。

```
$ java -Doracle.ons.debug=true -jar TestOracleConnection.jar\ subscribe_ons  
nodes="racdb1-vip:6200,racdb2-vip:6200"
```

RAC データベース

各 RAC ノードで Oracle ユーザーとして次のコマンドを実行し、システムヘルスを確認できます。

```
srvctl status database -d REDVMRAC
```

上記のコマンドによって、クラスタ内の各ノードのインスタンスの状態がレポートされます。

```
crs_stat [-t]
```

上記のコマンドによって、ONS、VIP、GSD、リスナー、インスタンス、データベースなどすべてのクラスタサービスの状態がレポートされます。-t オプションを使用すると、すべてが表形式でコンパクトに表示されます。

ONS

各 RAC ノードで、「onsctl ping」コマンドの (Oracle ユーザーによる) 発行により ONS が実行中であることを確認し、実行中でないとレポートされた場合は、「onsctl start」を使用して開始します。次のことに注意してください。「onsctl ping」により ons は実行中でないとレポートされ、一方「crs_stat」コマンドにより ons サービスはオンラインであるとレポートされた場合は、一部の環境変数が対立している可能性があります。すなわち、onsctl が \$ORA_CRS_HOME/bin の代わりに \$ORACLE_HOME/bin から参照されている場合は \$ORACLE_CRS_HOME/bin の前に \$ORA_CRS_HOME/bin をロードするように Oracle ユーザーのプロファイルを変更してノードを再起動します。

接続ロードバランシング

sqlplus または SYSDBA として RAC データベースにログインできるその他のプログラムを使用すると、次のクエリを実行してアクティブセッションおよび実行中のクラスタノードを表示できます。

```
select i.instance_name,
       i.host_name,
       s.username,
       to_char(sysdate, ' Day HH24:MI') current_time,
       s.LAST_CALL_ET seconds_since_active,
       s.failed_over,
       t.sql_text || ' ...' sql_fragment
from gv$instance i,
     gv$session s,
     gv$sqltext t
where s.inst_id = i.instance_number
     and s.status = 'ACTIVE'
     and s.username = 'HP_OO_USER'
     and rawtohex(s.sql_address) <> '00'
     and s.sql_address = t.address
     and t.piece = 0;
```

「hp_oo_user」の場所には HP OO ユーザーのユーザー名が入ります（SQL クエリでは大文字）。

負荷がかかったシステムでは、異なるインスタンスが同じユーザーのクエリを処理することも可能です。

有用な情報

Oracle、Oracle Downloads、Oracle Base の Web サイトには、次のような有用な情報が豊富に掲載されています。

- Oracle Downloads Web サイト
 - 『Oracle Database High Availability Overview』
 - 『Introduction to Oracle Clusterware and Oracle Real Application Clusters』
- Oracle Web サイト
 - 『Oracle Application Server 10 g Fast Connection Failover 構成ガイド – オラクル・ホワイト・ペーパー 2006 年 10 月』
このホワイトペーパーには、FCF テクノロジーの詳しい説明に加え、FAN イベントのデバッグとトラブルシューティングのヒントが記載されています。これは Oracle Application Server 向けに書かれていますが、概念は HP OO 環境でも有用です。
 - 『Oracle Real Application Clusters の自動ワークロード管理 – Oracle テクニカル・ホワイト・ペーパー』
オンライン版。
このホワイトペーパーには、サービスの詳細な定義、ロードバランシングアドバイザリの説明、FAN イベントの構造の詳細、ロードバランシングアドバイザリイベントの監視方法についての有用なヒントが記載されています。
- Oracle Base Web サイト
このサイトは VM ウェアでの RAC のインストール方法（パフォーマンスのベンチマーク用ではなく、その他のテスト用）が掲載されています。