

Oracle Exadataの動作状態および
リソース使用率の監視

Oracle Maximum Availability Architecture

ホワイト・ペーパー

2013年9月

Maximum
Availability
Architecture

Oracle Exadataの動作状態および
リソース使用率の監視

はじめに	1
方法論	1
問題発生前に取るべき措置	2
Exachk	2
AWRベースライン	2
静的ベースライン	2
変動ベースライン	5
構成ファイルのコピー	8
チューニング可能なパラメータのスナップショット	14
環境内で行われた変更の特定	15
初期の簡易チェック	16
コンポーネントの稼働	16
インシデント・マネージャ	21
アラート・ログ	22
OSシステム・ログ	25
CPU使用率	25
メモリ使用率	28
ILOMイベント	29
ネットワーク・ステータス	32
ディスク・ステータス	35
CheckHWnFWProfile	36
サービス	37
Exachk	38
ベースライン・データの使用による問題のトラブルシューティング	39
構成ファイルの比較	39
カーネルのチューニング可能パラメータに対する変更のチェック	39
AWRデータ	40
詳細診断	43
ハードウェアの除外	43
I/Oパフォーマンス	43
セルがI/Oバウンドであるかどうかのチェック	45
セルがI/Oバウンドである場合のI/O消費者の特定	51

問題シナリオ1 - HDDの使用率は高いが、フラッシュはアイドル状態にある	53
問題シナリオ2 - HDDが書込みIOPSバウンドである	56
DWとOLTPの両方でのIORMの使用.....	58
Enterprise Managerを使用したI/O監視	59
データベースのFree Buffer Waits	59
セル・サーバーのターゲット・ページ	59
ファイル・システム情報.....	61
コンピュート・ノードがCPUバウンドであるかどうかのチェック	66
コンピュート・ノードがメモリ・バウンドであるかどうかのチェック	67
VMSTATの調査.....	67
TOP kswapdの調査	67
次のステップ	68
コンピュート・ノードがCPUバウンドまたはメモリ・バウンドである場合.....	68
セル・サーバーがI/Oバウンドではなく、コンピュート・ノードがCPUバウンドでもメモリ・バウンドでもない場合.....	68
データベース・パラメータの設定方法.....	68
HugePagesの構成方法	69
統合環境の構成方法	69
データベース診断ツール.....	70
ADDM	70
AWRレポート	71
予防ツールによる混合ワークロード・システムでのワークロード分散処理.....	74
I/Oリソース・マネージャ	74
Enterprise Managerを使用したExadataのリソース管理.....	75
データベースの上位アクティビティ・ページ	77
AWRレポート	78
データベース・メトリック	79
参考資料	81
MAAベスト・プラクティス	81
まとめ	82

はじめに

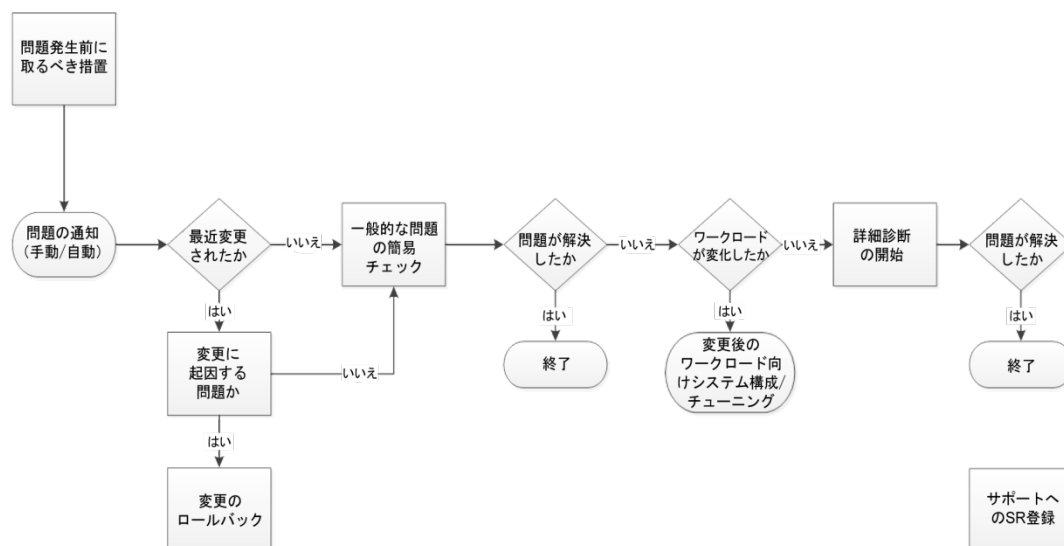
本書では、Oracle Exadata環境の動作状態とリソース使用率を監視するためのエンド・ツー・エンドのアプローチについて説明します。また、技術的な詳細情報の他に、管理者が素早く問題を特定して手際よく修正できるようにするためのトラブルシューティング方法について検証していきます。

方法論

本書の目的は、Oracle Exadataシステムで発生する問題のトラブルシューティング手法を提供することです。本書では、システムの各コンポーネントの動作が正しいことを実証し、発生したインシデントから、動作の正しいコンポーネントの役割を排除していくという"除外"アプローチを採用します。システム診断全体の中で重点を置くべき領域は次の5つです。

1. トラブルシューティングに役立つ、問題が発生する前に取るべき措置
2. システムに対する変更
3. 迅速な分析
4. ベースラインの比較
5. 詳細診断

次のフローチャートは、上記手順のフローをわかりやすくしたものです。



問題発生前に取るべき措置

時間をかけて不測の事態に備えることで、問題の診断にかかる時間を大幅に短縮できます。以下に、情報収集を確実に行うために必要な手順を記載します。

Exachk

ExachkとOracle Enterprise Managerのヘルスチェック・プラグインによる出力を利用することで、収集された最新データの確認と、サポート対象のバージョン・レベルおよび推奨されるOracle Exadata ベスト・プラクティスに対する相互参照が容易になります。Oracle Exadata Database MachineのExachkツールとヘルスチェック・ツールは、手動またはスケジュール設定を介して実行でき、主要なソフトウェア、ハードウェア、ファームウェア、構成に関するデータ収集に役立ちます。

Exachkの実行頻度は要件に合わせて任意で指定でき、定期的な実行（例：月次）が推奨されます。また、Oracle Exadata Database Machineに対する企業の計画保守プログラムの一環として実行すると良いでしょう。最新のベスト・プラクティスや既知の重大な問題に対する追加チェックを含める目的で、**定期的に更新されているExachkおよびヘルスチェック・ツールの最新バージョンを必ず確認してください**。詳しくは、Oracle Supportドキュメント1070954.1：“Oracle Exadata Database Machine Exachk or HealthCheck” - [Exachk MOS Note](#)を参照してください。すべてのExadataベスト・プラクティスを確認するには、MOS Note：[Exadata Best Practices MOS Note](#)を参照してください。

AWRベースライン

AWRベースラインを利用すると、指定した範囲のスナップショットを保持してパフォーマンスを比較するために使用できます。これらのベースラインは、保存期間を過ぎても、通常のAWRスナップショットのように削除されることはありません。

Exadataマシンで参照用に作成するベースラインには、2種類があります。各コンピュータ・ノード上のすべてのデータベース・インスタンスには、静的ベースラインに加えて変動ベースラインがあり、環境内で発生する可能性のある各種ワークロードを取得します。たとえば、週次バッチ・ロードが月曜夜7時～10時に実行される場合、この期間中にインスタンスのベースラインを取得すると非常に有効でしょう。

静的ベースライン

AWR静的ベースラインを作成するには、2つのスナップショットを選択します。これら2つのスナップショットの間に存在するAWRデータが評価用に保存されます。たとえば、バッチ処理ジョブや通常のOLTP処理、またはピーク・ワークロード期間に対してベースラインを作成できます。AWRベースラインは、手動で削除するまで保持されます。静的ベースラインを作成するには、DBA権限を持つユーザーとして次のコマンドを実行します。

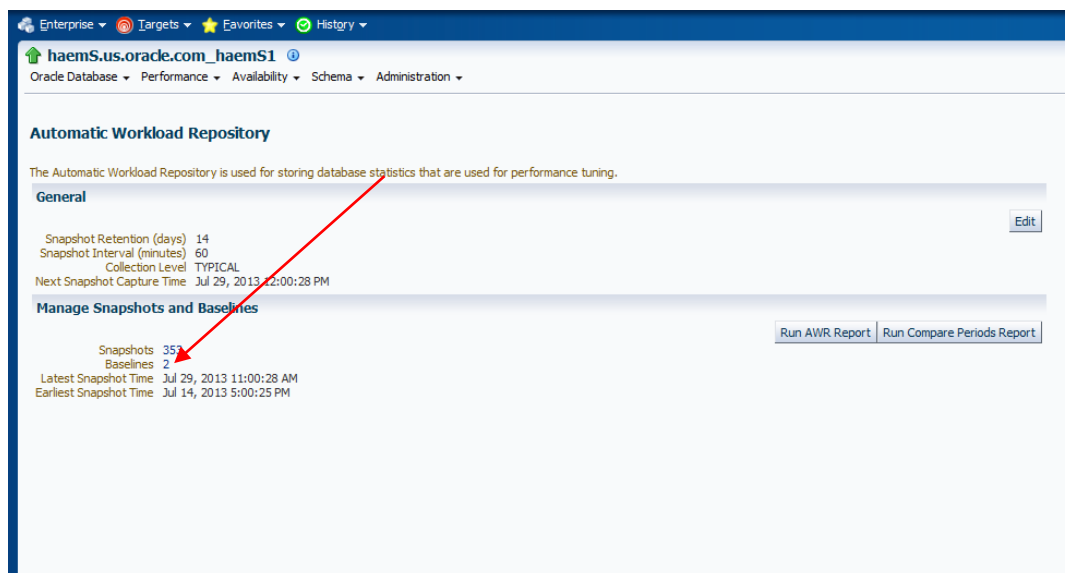
```
exec dbms_workload_repository.create_baseline (start_snap_id => <starting snapshot id>,
end_snap_id => <ending snapshot id, baseline_name => 'Normal Baseline');
```

Enterprise Managerで静的ベースラインを作成するには、データベース・インスタンスのターゲット・ページに移動し、「Performance」→「AWR」→「AWR Administration」の順に選択します。

The screenshot displays the Oracle Enterprise Manager Cloud Control 12c interface. The breadcrumb navigation path is: Enterprise > Targets > Favorites > History > haemS.us.oracle.com_haemS1 > Oracle Database > Performance > Availability > Schema > Administration. The 'Performance' menu is open, showing options like Performance Home, Top Activity, ASH Analytics, SQL Monitoring, SQL, AWR (selected), Advisors Home, Memory Advisors, Emergency Monitoring, Real-Time ADDM, Adaptive Thresholds, Search Sessions, Blocking Sessions, and Database Replay. The 'AWR' sub-menu is also open, showing options: AWR Report, AWR Administration, Compare Period ADDM, and Compare Period Reports. The main dashboard shows various performance metrics and charts, including 'Activity Class' (Active Sessions), 'Resources' (Host CPU), and 'SQL Monitor - Last Hour' (table below).

Status	Duration	SQL ID
✓	17.00 s	9v5rq4jb13htq
✓	5.00 s	73qygu3avgtn
✓	14.00 s	gsbdfku007tup
✓	5.00 s	7x0yk3yrzv3rf
✓	5.00 s	b12g3n0660mc
✓	4.00 s	6ajkhuk78nsr
✓	22.00 s	9v5rq4jb13htq
✓	19.00 s	9v5rq4jb13htq
✓	154.00 s	30cf6ccvbdaz7

次のページで、Baselinesというテキストの横にある数字を選択します。この数字は、設定済みのベースライン数を表します。



次に、「Create」ボタンを選択します。



次の画面では、単一の静的ベースラインまたは繰り返しベースラインのいずれかを作成できます。繰り返しベースラインは、ユーザー定義スケジュールに基づいて作成されます（例：毎週月曜日の午前9時～午後5時）。

変動ベースライン

変動ベースラインは、静的ベースラインと同じ情報を収集します。デフォルトの時間枠は、AWR保存期間（8日間）です。ただし、AWRスナップショットはエージ・アウトされるため、データは変わっていきます。AWR保存期間をできる限り長く設定して、必要な情報すべてを利用できるようにすると良いでしょう。変動時間枠の最大値は、保存期間と同じです。したがって、保存期間を少なくとも30日間に変更することを推奨します。次に例を示します。

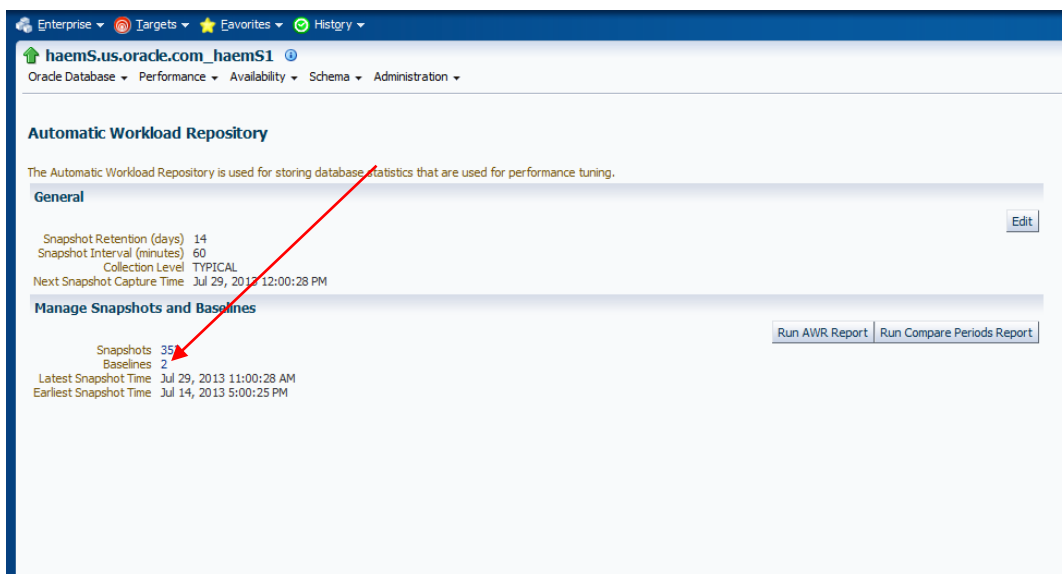
```
exec dbms_workload_repository.modify_baseline_window_size(30);
```

変動時間枠を変更するには、Enterprise ManagerでAWR Baselinesページのベースラインに移動し、「Performance」→「AWR」→「AWR Administration」の順に選択します。

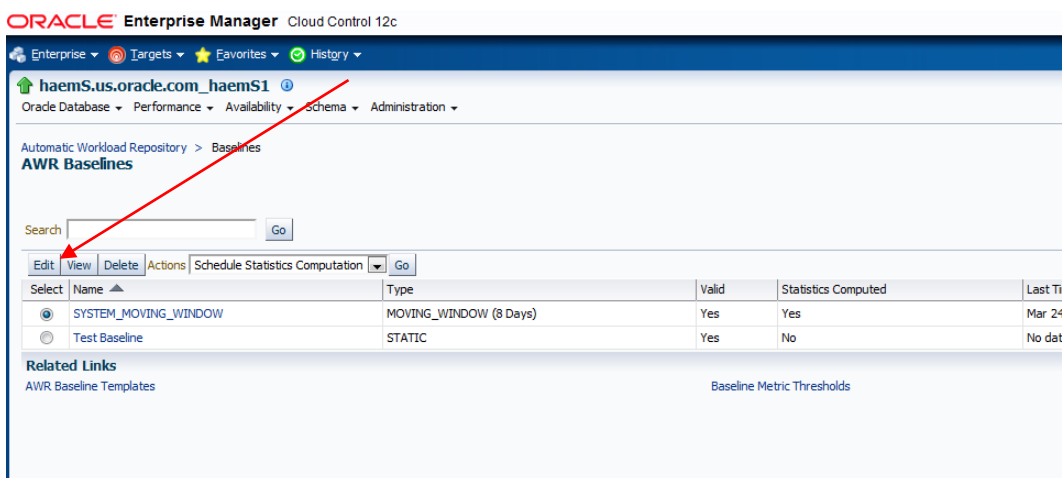
The screenshot shows the Oracle Enterprise Manager Cloud Control 12c interface. The breadcrumb navigation is: Enterprise > Targets > Favorites > History > haemS.us.oracle.com_haemS1. The main menu includes Oracle Database, Performance, Availability, Schema, and Administration. The Performance menu is expanded, showing options like Performance Home, Top Activity, ASH Analytics, SQL Monitoring, SQL, AWR, Advisors Home, Memory Advisors, Emergency Monitoring, Real-Time ADDM, Adaptive Thresholds, Search Sessions, Blocking Sessions, and Database Replay. The AWR option is selected, opening a sub-menu with AWR Report, AWR Administration, Compare Period ADDM, and Compare Period Reports. The main dashboard displays several widgets: Summary (Status, Diagnostics, Compliance Summary, Jobs Running), Performance (Activity Class chart), Resources (Host CPU chart), SQL Monitor - Last Hour (table), and Incidents and Problems.

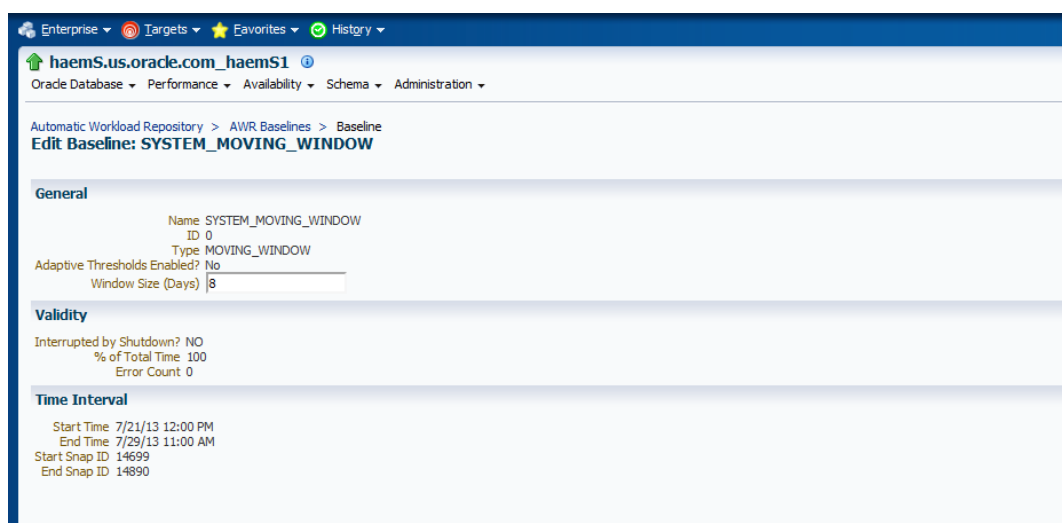
Status	Duration	SQL ID
✓	17.00 s	9v5rq4jb13htq
✓	5.00 s	73qygu3avgtn
✓	14.00 s	gsbdfku007tup
✓	5.00 s	7x0yk3yrzv3rf
✓	5.00 s	b12g3n0660rnc
✓	4.00 s	6ajkhukk78nsr
✓	22.00 s	9v5rq4jb13htq
✓	19.00 s	9v5rq4jb13htq
✓	154.00 s	30cfccvbdaz7

次のページで、Baselinesの横にある数字を選択します。



次に、"SYSTEM_MOVING_WINDOW"ベースラインの横にあるラジオ・ボタンを選択し、「Edit」ボタンをクリックします。





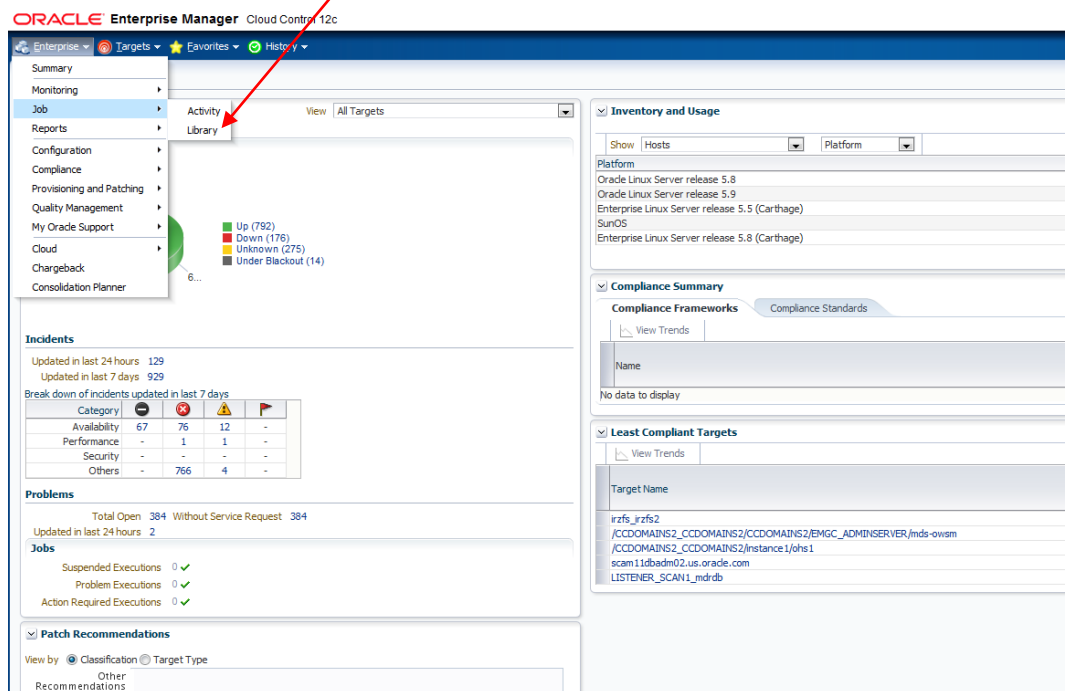
構成ファイルのコピー

重要ファイルについて、タイムスタンプの付いた増分コピーを取得しておく、環境内で行われた変更をより簡単に特定できます。つまり、Exadata環境のさまざまな部分で、変更の監査証跡を取得できます。定期的にコピーを作成してローカル・ファイル・システムに保存するスクリプトを作成することをお勧めします。このようなファイルは、テープに何度もバックアップするようにしているかもしれませんが、さまざまな繰返しをテープから読み込むには時間を要し、非常にコストがかかる可能性があります。次に、最低限保存すべきファイルを示します。

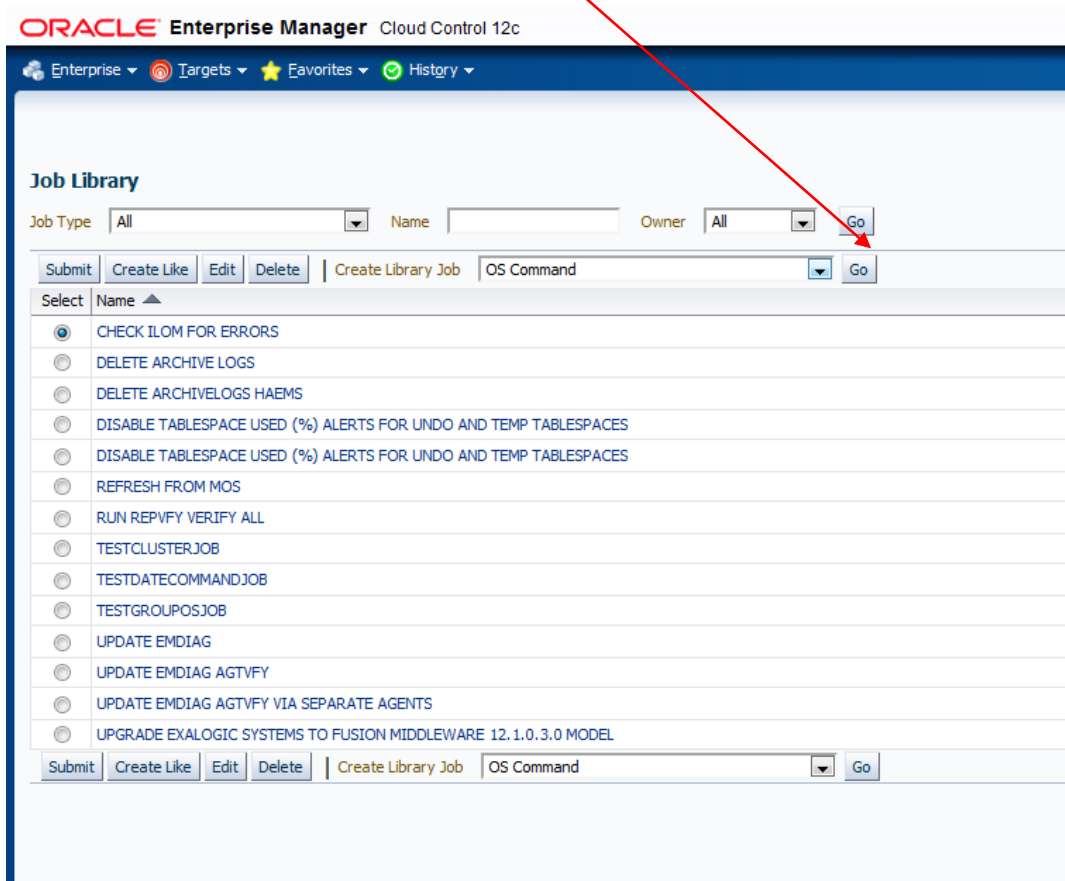
- データベースのすべてのインスタンスのinitファイルとspfile
- パスワード・ファイル
- ASMのinitファイル
- ASMのパスワード・ファイル

Enterprise Managerを使用してこのタスクを実行することもできます。適用するターゲットごとにジョブを作成してスケジュール設定することで、定期的に上記ファイルをコピーできます。例として、データベースのspfileの週次コピーを作成するジョブを作成します。

はじめに、Enterprise ManagerのJob Libraryページに移動します。



次に、新しいジョブを作成します。



The screenshot shows the Oracle Enterprise Manager Cloud Control 12c interface. The page title is "ORACLE Enterprise Manager Cloud Control 12c". The navigation bar includes "Enterprise", "Targets", "Favorites", and "History". The main section is titled "Job Library".

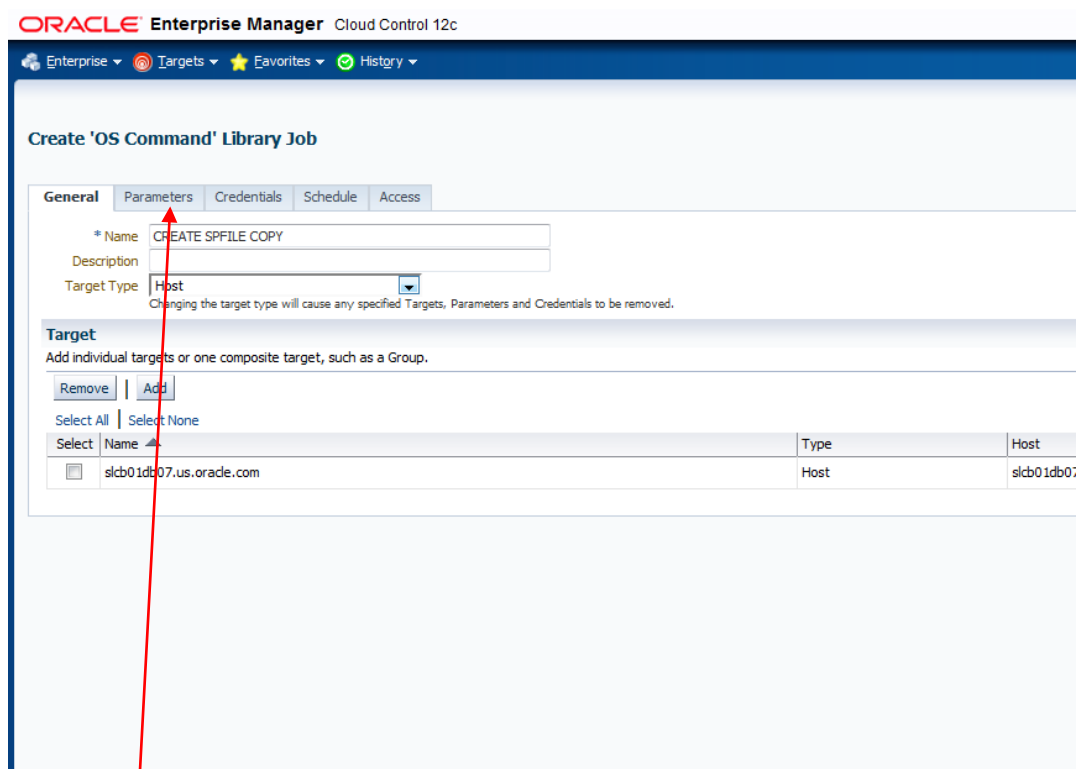
At the top of the Job Library section, there are search filters: "Job Type" (set to "All"), "Name" (empty), and "Owner" (set to "All"). A "Go" button is located to the right of the "Owner" filter, which is highlighted by a red arrow. Below these filters are buttons for "Submit", "Create Like", "Edit", "Delete", and "Create Library Job". A dropdown menu is set to "OS Command", with a "Go" button to its right.

The main content area is a table with a "Select" column and a "Name" column. The table contains the following jobs:

Select	Name
<input checked="" type="radio"/>	CHECK ILOM FOR ERRORS
<input type="radio"/>	DELETE ARCHIVE LOGS
<input type="radio"/>	DELETE ARCHIVELOGS HAEMS
<input type="radio"/>	DISABLE TABLESPACE USED (%) ALERTS FOR UNDO AND TEMP TABLESPACES
<input type="radio"/>	DISABLE TABLESPACE USED (%) ALERTS FOR UNDO AND TEMP TABLESPACES
<input type="radio"/>	REFRESH FROM MOS
<input type="radio"/>	RUN REPVFY VERIFY ALL
<input type="radio"/>	TESTCLUSTERJOB
<input type="radio"/>	TESTDATECOMMANDJOB
<input type="radio"/>	TESTGROUPOSJOB
<input type="radio"/>	UPDATE EMDIAG
<input type="radio"/>	UPDATE EMDIAG AGTVFY
<input type="radio"/>	UPDATE EMDIAG AGTVFY VIA SEPARATE AGENTS
<input type="radio"/>	UPGRADE EXALOGIC SYSTEMS TO FUSION MIDDLEWARE 12.1.0.3.0 MODEL

At the bottom of the table, there are buttons for "Submit", "Create Like", "Edit", "Delete", and "Create Library Job". A dropdown menu is set to "OS Command", with a "Go" button to its right.

次の画面でジョブ名を入力し、ターゲット・タイプをHostに設定し、ターゲットを追加します。

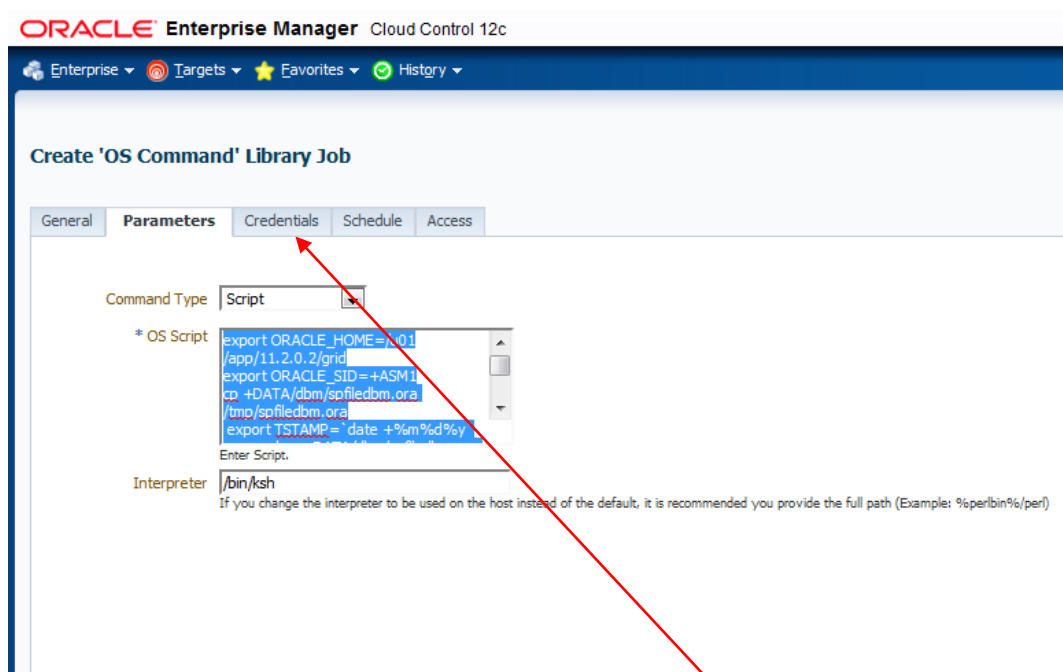


The screenshot shows the Oracle Enterprise Manager Cloud Control 12c interface for creating a new 'OS Command' Library Job. The 'General' tab is selected, and the job name is 'CREATE SPFILE COPY'. The target type is set to 'Host'. A red arrow points to the 'Parameters' tab.

Select	Name	Type	Host
<input type="checkbox"/>	slcb01db07.us.oracle.com	Host	slcb01db07.us.oracle.com

次に「Parameters」タブをクリックします。Command Typeで「Script」を選択し、OS Scriptボックスに次のスクリプト・テキストを入力し、Interpreterに/bin/kshと指定します。この例で使用するスクリプトは非常に簡単なものです。

```
export ORACLE_HOME=/u01/app/11.2.0.2/grid
export ORACLE_SID=+ASM1
cd +DATA/dbm/spfiledbm.ora /tmp/spfiledbm.ora
export TSTAMP=`date +%m%d%y`
asmcmd cp +DATA/dbm/spfiledbm.ora
/u01/app/oracle/spfiledbm.ora_${TSTAMP}
```



次に「Credentials」タブをクリックし、使用する資格証明を選択します。

The screenshot shows the Oracle Enterprise Manager Cloud Control 12c interface. The main title is "Create 'OS Command' Library Job". Below the title are four tabs: "General", "Parameters", "Credentials", "Schedule", and "Access". The "Credentials" tab is currently selected. A red arrow points from the "Schedule" tab to the "Credentials" tab. The "Credentials" section contains a tip: "TIP Select global named credentials. Target instance associated credentials are not supported." Below this is the "Host Credentials" section, which includes the following fields: "Credential" (radio buttons for Preferred, Named, and New, with New selected), "* UserName" (text field with "oracle"), "* Password" (password field with 10 dots), "* Confirm Password" (password field with 10 dots), "Run Privilege" (dropdown menu with "None" selected), and a "Save As" checkbox (checked) with the text "NC_HOST_2013-08-21-080129".

「Schedule」タブを選択し、ジョブを実行する頻度を指定します。少なくとも月に1回は変更を収集することをお勧めしますが、より動的な環境では頻度を高くする必要があります。

「save to library」を選択すると、ジョブがライブラリに保存され、実行スケジュールが設定されます。

The screenshot shows the 'Create OS Command Library Job' configuration page in Oracle Enterprise Manager Cloud Control 12c. The 'Schedule' tab is active, showing the following settings:

- Type:** Repeating
- Frequency Type:** By Weeks
- Repeat Every:** 1 Weeks
- Time Zone:** (UTC-08:00) US Pacific Time (PST)
- Start Date:** Aug 21, 2013
- Start Time:** 8 :29 AM PM
- Grace Period:** Indefinite End After [] Hours [] Minutes
- Repeat Until:** Indefinite Specified Date
 - Date:** [] [] [] (example: Aug 21, 2013)
 - Time:** [] : [] AM PM

チューニング可能なパラメータのスナップショット

UNIXカーネルのチューニング可能なパラメータは、サーバー動作に大きな影響を与えます。何らかの変化が発生した場合に素早く比較できるようにするため、チューニング可能なパラメータのバックアップを取得しておくことをお勧めします。これは手動で実行できますが、上述の手順を使用して、Enterprise Managerで定期的にスナップショットを取得するように設定することもできます。この情報はコンピュータ・ノードごとに収集する必要があります。次のコマンドを実行すると、パラメータのリストが表示されます。

```
# sysctl -a
```

これらをタイムスタンプ付きファイルに保存する場合、次のコマンドを実行します。

```
# sysctl -a > /root/kernel_parameters/sysctl_073013
```

このコマンドをすべてのノードで同時に実行するには、dcliコマンドを使用します。ここでdbs_groupは、すべてのコンピュータ・ノードを含むファイルです。

```
# dcli -l root -g ./dbs_group "sysctl -a > /root/sysctl_`date +%m%d%y`"
```

環境内で行われた変更の特定

環境に対する変更が意図しない影響をもたらすことはよくあります。最近のシステム変更を識別することが、一般的な問題の原因を特定するための重要な第一歩です。正式な変更管理プロセスが実施されている場合、変更の特定は素早く簡単に実行できます。それ以外の場合は、考えられる変更元の調査を開始する必要があるでしょう。変更には次のようなものがあります。

- 最近のOracleパッチ（オペレーティング・システム、データベース、セル・サーバー、クラスタウェアなど）
- 新しくデプロイしたアプリケーション
- 既存アプリケーションのコード変更
- 使用状況におけるその他の変更（新規ユーザーの追加）
- Oracleの構成変更
- オペレーティング・システムの構成変更
- 新規プラットフォームへの移行
- 環境の拡張
- ファブリックへのその他のInfiniBandデバイスの追加
- リソース管理計画の変更

職務分離の状況によっては、1人と話すだけですべての変更元の確認が完了する場合がありますが、サイロ化した大規模組織では、多数のチームに調査する必要があるでしょう。

システム上の変更がある場合、これらの変更が特定済みの問題に関連していないことを確認するための措置を取る必要があります。変更によって悪影響が発生していると確認された場合、分析を行い、最善の方針を割り出す必要があります。これには変更のロールバックや容量の拡張、コードの変更などが含まれます。

初期の簡易チェック

環境の変更が目下の問題の原因ではないと確認できたら、問題の性質を分離するために、一連の簡易チェックを実行します。次に示すチェックリストは、数分以内に完了する簡易チェックです。チェックの実行について、詳しくは以下を参照してください。

	コンピュート・ノード	セル・サーバー	データベース	ASM	クラスタウェア	InfiniBand スイッチ
コンポーネントの稼働	X	X	X	X	X	
アラート・ログ			X	X	X	
OSシステム・ログ	X	X				
CPU	X	X				
メモリ使用率	X	X				
ILOMエラー	X	X				X
すべてのネットワークの稼働	X	X				X
ディスク・ステータス		X				
CheckHWnFWProfile	X	X				
データベース・サービス/リスナーの稼働					X	
Exachk	X					

コンポーネントの稼働

実行がもっとも簡単なチェックは、おそらく、Exadataマシンのコンポーネントが稼働中であるかどうかの確認です。すべてのハードウェア・コンポーネントとソフトウェア・コンポーネントが稼働中で利用可能であることを確認することが、さらに詳しい調査を開始するための堅固な基盤になります。

コンピュート・ノードとセル・サーバーにログインすると、これらが稼働中であるかどうかを確認できます。また、これらの実行レベルが適切であるかどうかを確認することをお勧めします。確認するには、runlevelコマンドを実行します。

```
# runlevel
```

```
N3
```

すべてのコンピュート・ノードとストレージ・サーバーは、実行レベルを3に設定する必要があります。

crsctl stat res -tを実行すると、すべてのCRSリソースを簡単に確認できます。

```
# ./crsctl stat res -t
-----
NAME                TARGET  STATE        SERVER
STATE_DETAILS
-----
Local Resources
-----
ora.DATA.dg
                ONLINE  ONLINE      slcb01db07
                ONLINE  ONLINE      slcb01db08
ora.DBFS_DG.dg
                ONLINE  ONLINE      slcb01db07
                ONLINE  ONLINE      slcb01db08
ora.LISTENER.lsnr
                ONLINE  ONLINE      slcb01db07
                ONLINE  ONLINE      slcb01db08
ora.RECO.dg
                ONLINE  ONLINE      slcb01db07
                ONLINE  ONLINE      slcb01db08
ora.asm
                ONLINE  ONLINE      slcb01db07      Started
                ONLINE  ONLINE      slcb01db08      Started
ora.gsd
                OFFLINE OFFLINE      slcb01db07
                OFFLINE OFFLINE      slcb01db08
ora.net1.network
                ONLINE  ONLINE      slcb01db07
                ONLINE  ONLINE      slcb01db08
ora.ons
                ONLINE  ONLINE      slcb01db07
                ONLINE  ONLINE      slcb01db08
```

コンピュータ・ノードからsvrctlコマンドを実行すると、データベースとASMのステータスを簡単に確認できます。

```
svrctl status database -d <DATABASE NAME> -v
```

```
Instance <INSTANCE NAME> is running on node <SERVER NAME>.Instance status:Open.
```

```
Instance <INSTANCE NAME> is running on node <SERVER NAME>.Instance status:Open.
```

これにより、データベースが稼働中であり、オープン状態にあることが分かります。ASMをチェックするには、次のコマンドを実行します。

```
svrctl status asm-v
```

```
ASM is running on <NODE1>,<NODE2>
```

```
Detailed state on node <NODE1>:Started
```

```
Detailed state on node <NODE2>:Started
```

これらのチェックはすべて、Enterprise Managerからも簡単に実施できます。Database Machineのホームページを表示すると、RAC全体のグラフィカル表示から簡単に停止コンポーネントを特定できます。

Exadataのターゲット・ページを表示します。

The screenshot shows the Oracle Enterprise Manager Cloud Control 12c interface. The 'Targets' dropdown menu is open, listing various target types including 'Exadata'. A red arrow points from the 'Exadata' option in the menu to the 'Exadata' link in the 'Incidents' section below.

Incidents

Detected in last 24 hours 98
 Detected in last 7 days 476
 Trend: Down of incidents updated in last 7 days

Category	🟢	🔴	🟡	🚩
Availability	70	38	21	-
Performance	-	1	-	-
Security	-	-	4	-
Others	-	335	1	-

対象のExadata Database Machineを選択します。

The screenshot displays the Oracle Enterprise Manager interface for 'Oracle Exadata Database Machines'. At the top, there are navigation tabs for 'Enterprise', 'Targets', 'Favorites', and 'History'. Below this is a search section with a 'Search' button and a 'Target Name' input field. A toolbar contains 'View', 'Remove', 'Add', and 'Detach' options. The main area is a table listing database machines with columns for 'Target Name', 'Status', and 'Me'.

Target Name	Status	Me
> DB Machine adc2ardb0506.us.oracle.com	↑	Or
> DB Machine edx2.us.oracle.com	↑	Or
> DB Machine sca.us.oracle.com_2	↑	Or
> DB Machine sca.us.oracle.com_3	↑	Cl
> DB Machine scac01.us.oracle.com	↑	Or
> DB Machine scac01db0506.us.oracle.com	↑	Cl
> DB Machine scac02.us.oracle.com	↑	Or
> DB Machine scac04db0102.us.oracle.com	↑	Or
> DB Machine scam02.us.oracle.com	↑	Or
> DB Machine scam02.us.oracle.com_2	↑	Or
> DB Machine scam02.us.oracle.com_3	↑	Hc
> DB Machine scam09.us.oracle.com	↑	Or
> DB Machine scam11db0102.us.oracle.com	↑	Hc
> DB Machine slcc15.us.oracle.com	↑	Or
> DB Machine slcc17.us.oracle.com	↑	Or
> DB Machine slcc26.us.oracle.com	↑	Or

Exadata Database Machineホームページが表示されます。

The screenshot displays the Exadata Database Machine management interface. On the left, a 'Target Navigation' pane shows a tree view of the database machine components. The main area shows a rack of server components, each with a status indicator (a green dot for 'Up', a red square for 'Down', etc.). The components are organized into sections: 'EXADATA' (top), 'EXTRME PERFORMANCE' (middle), and 'Rack 1' (bottom). A legend on the right explains the status indicators.

Component Name	Status
scac02cel14	Up
scac02cel13	Up
scac02cel12	Up
scac02cel11	Up
scac02cel10	Up
scac02cel09	Up
scac02cel08	Up
scac02db08	Up
scac02db07	Up
scac02db06	Up
scac02db05	Up
scac02sw-ib3	Up
scac02sw-kvm	Up
scac02sw-int	Up
scac02sw-ib2	Up
scac02db04	Up
scac02db03	Up
scac02db02	Up
scac02db01	Up
scac02cel07	Up
scac02cel06	Up
scac02cel05	Up
scac02cel04	Up
scac02cel03	Up
scac02cel02	Up
scac02cel01	Up
scac02sw-ib1	Up

Legend

- Up (Green dot)
- Down (Red square)
- Blackout (Black square)
- Exadata Cell (Dark grey square)
- Compute Node (Light blue square)
- Infiniband Switch (Light grey square)
- Ethernet Switch (White square)
- Keyboard-Video-Mouse (Dark grey square)
- Unallocated (Black square)

インシデント・マネージャ

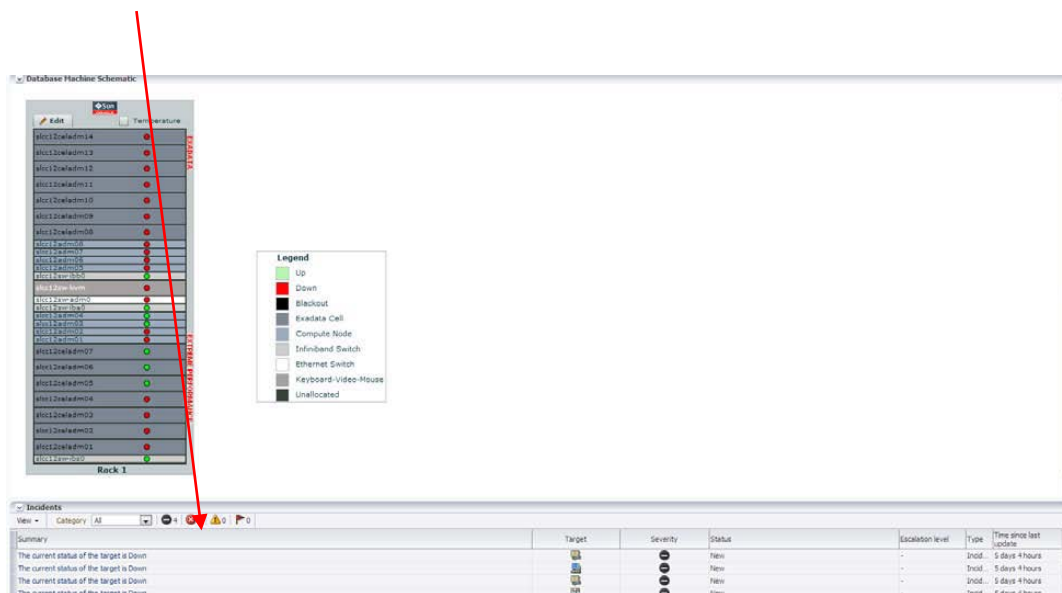
インシデント・マネージャを利用すると、管理者は問題を素早く効率的に特定し、監視し、解決できます。インシデント・マネージャは、3レベルの階層を使用し、問題をグループ化して分類します。

イベント：Enterprise Managerによって検出されたターゲットに影響する重大な事象。

インシデント：一連の重大なイベント、または同じ問題に関係する関連イベントの組合せ。

問題：インシデントの根本原因。現在、これはOracleソフトウェアでの重大なエラーを表し、診断するインシデントの根本原因を意味します。

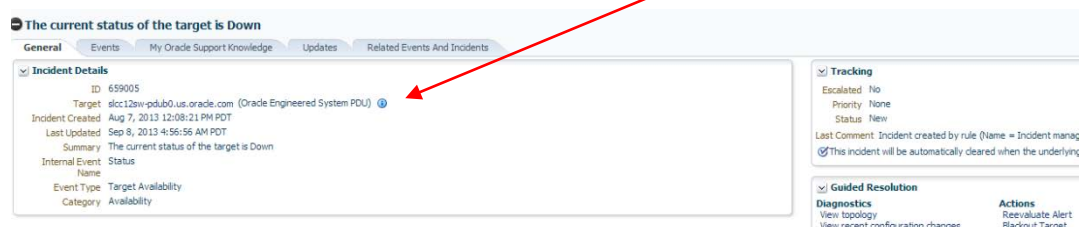
Exadataコンポーネントに対して作成されたインシデントは、次に示すとおり、Database Machineのホームページで参照できます。



The screenshot displays the Oracle Enterprise Manager Database Machine Schematic. The main area shows a rack of database machines, with a legend indicating their status: Up (green), Down (red), Blackout (black), Evacuate Cell (grey), Compute Node (light blue), Infiniband Switch (dark blue), Ethernet Switch (light grey), Keyboard-Video-Mouse (dark grey), and Unallocated (dark blue). A red arrow points from the 'Down' status in the legend to a red 'Down' icon in the schematic. Below the schematic is the 'Incidents' table, which lists several incidents related to the 'Down' status of the database machines.

View	Category	All	Severity	Status	Escalation level	Type	Time since last update
Summary							
The current status of the target is Down			Down	New		Incident	5 days 4 hours
The current status of the target is Down			Down	New		Incident	5 days 4 hours
The current status of the target is Down			Down	New		Incident	5 days 4 hours
The current status of the target is Down			Down	New		Incident	5 days 4 hours

さらに、個々のインシデントをクリックすると、インシデント・マネージャ画面が表示されます。



アラート・ログ

アラート・ログは、データベースやクラスタウェア、ASMなどのOracleコンポーネントによって書き込まれたメッセージとエラーを時系列に並べたファイルです。アラート・ログは次のディレクトリにあります。

- データベース
 - `<DIAGNOSTIC DIRECTORY>/diag/rdbms/<database name>/<INSTANCE NAME>/trace/alert_<INSTANCE NAME>.log`
- ASM
 - `<DIAGNOSTIC DIRECTORY>/diag/asm/+asm/<ASM INSTANCE>/trace/alert_<ASM INSTANCE>.log`
- クラスタウェア
 - `CLUSTWARE HOME>/log/<HOSTNAME>/alert<HOSTNAME>.log`

データベース運用について、詳しくは次のドキュメントを参照してください。

http://docs.oracle.com/cd/E11882_01/server.112/e10897/toc.htm

データベースとASMターゲットのアラート・ログ・メトリックは、インスタンスのターゲット・ページで参照できます。ターゲット・ページに移動し、「Oracle Database」、「Automatic Storage Management」、または「Cluster」から「Monitoring」→「All Metrics」を選択します。

ORACLE Enterprise Manager Cloud Control 12c

Enterprise ▾ Targets ▾ Favorites ▾ History ▾

haemS.us.oracle.com_haemS1

Oracle Database ▾ Performance ▾ Availability ▾ Schema ▾ Administration ▾

- Home
- Monitoring ▾
 - User-Defined Metrics
 - All Metrics
 - Metric and Collection Settings
 - Metric Collection Errors
 - Status History
 - Incident Manager
 - Alert History
 - Blackouts
- Diagnostics ▾
- Control ▾
- Job Activity
- Information Publisher Reports
- Logs ▾
- Provisioning ▾
- Configuration ▾
- Compliance ▾
- Target Setup
- Target Information

1 0

Performance

Activity Class Services

Active Sessions

11:42 AM 11:52 AM 12:02 PM 12:12 PM 12:22 PM

Resources

Host CPU

Active Sessions

Other Instance

SQL Monitor - Last Hour

Status	Duration	SQL ID	Session ID
✓	3.00 s	3vpa1s21sasn2	56
✓	5.00 s	9cjr0dgg12duz	607
✓	18.00 s	9v5rq4jb13htq	208

Compliance Summary

Compliance Standards

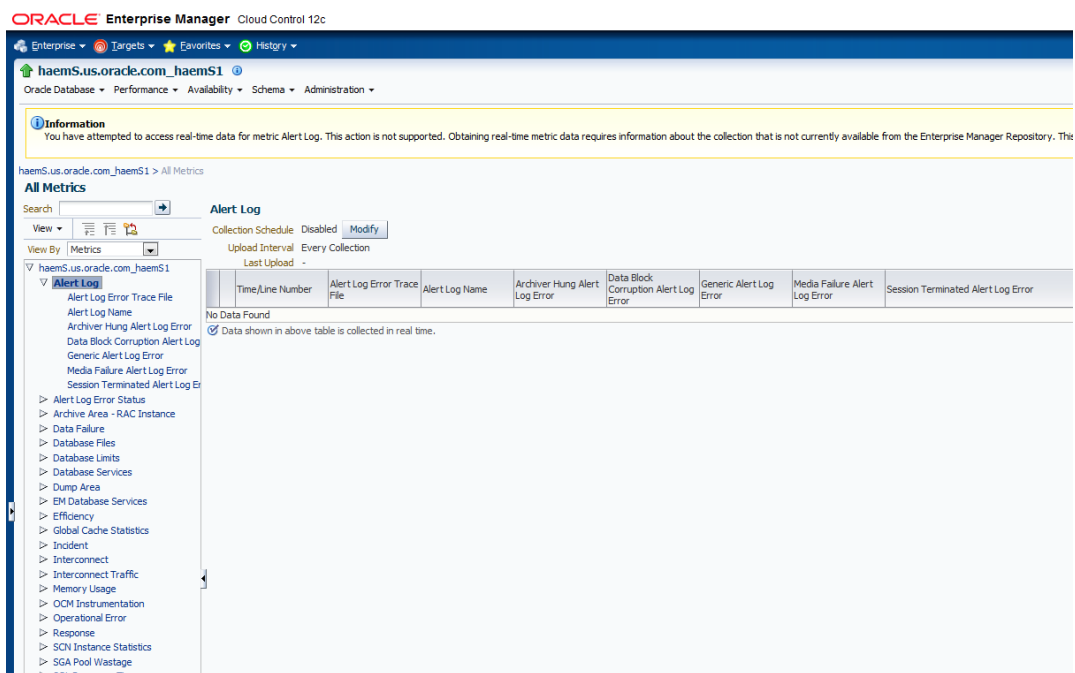
View Trends

Name	Average Score
No data to display	

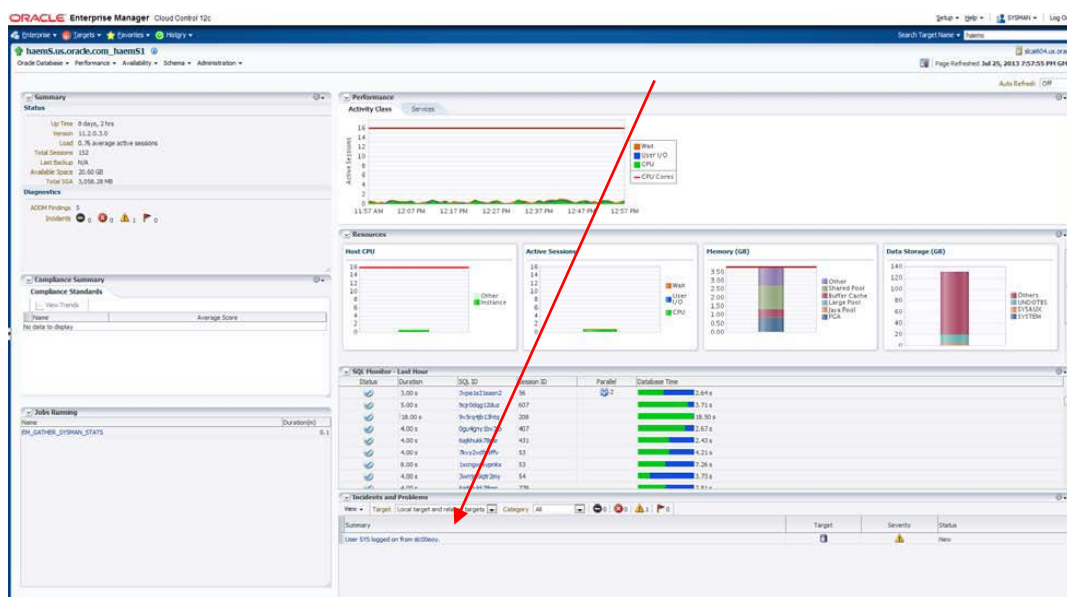
Jobs Running

Name	Duration(m)
------	-------------

この画面から、いくつかのアラート・ログ・メトリックを確認できます。



また、重大な問題が検出されると、Enterprise Managerのインシデントが作成されます。これはメインのターゲット・ページで直接参照できます。



ソフトウェア・レベルの問題がある場合、たいていはこれらのファイルのいずれかに問題が表示されます。加えて、ディスクやネットワークの問題などの一部のハードウェア・レベルのイベントも識別されます。

OSシステム・ログ

システム・ログは、OSに関するほとんどのメッセージとエラーが書き込まれる場所です。ハードウェアまたはオペレーティング・システムに問題がありそうな場合、このログは優れた出発点となります。Linuxの場合、このファイルは通常、次の場所にあります。

`/var/log/messages`

Solarisの場合：

`/var/adm/messages`

ログ・ファイルの場所は設定で変更できます。この情報を確認するには、`/etc/syslog.conf`を参照してください。

CPU使用率

コンピュータ・ノードのCPU使用率は、`top`、`AWR`、`iostat`、`vmstat`などのさまざまなツールを使用して測定できます。すべてのツールで同じ数値が表示されますが、通常、CPU使用率は指定された期間中の平均値です。使い勝手に合わせてツールを選択できますが、インテルCPUのハイパースレッディングを考慮する必要があります。

すべてのExadataモデルで使用されているインテルCPUは、CPUコアごとに2つのスレッドを使用して実行されます。これにより総合的なパフォーマンスが向上しますが、2番目のスレッドは最初のスレッドほど強力ではありません。オペレーティング・システムはすべてのスレッドが同等であることを前提にするため、利用可能なCPU能力が実際より高くなります。これを考慮に入れる必要があります。次に、実際のCPU使用率を推定するために使用できるおおよその目安を示します。ただし、この数値はワークロードによって異なります。

- CPU使用率が50%未満である場合は、1.7を掛けます。
- CPU使用率が50%以上である場合は、 $85\% + (\text{CPU使用率} - 50\%) \times 0.3$ を使用します。

次の表に概算値を示します。

測定された使用率	実際の使用率
10%	17%
20%	34%
30%	51%
40%	68%
50%	85%
60%	88%
70%	91%
80%	94%
90%	97%
100%	100%

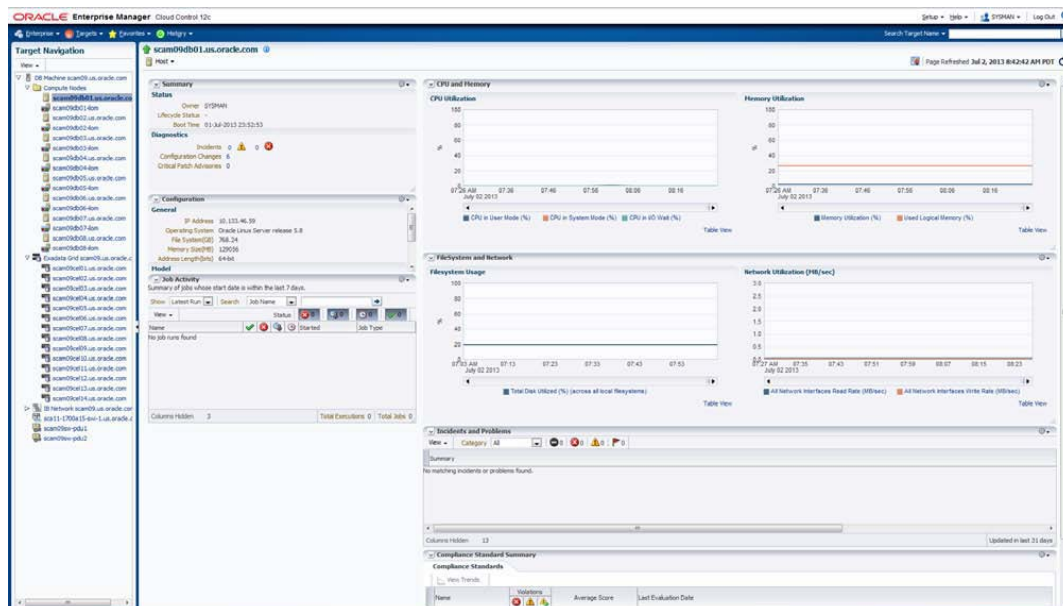
実際の使用率は、利用可能な予備CPUがどれだけあるかを示すおおよその基準になります。

応答時間が重要なワークロードに対しては、実際の使用率を85%未満に維持するように計画します。使用率が85%を超える場合、IORMまたはインスタンス・ケーシングを使用してワークロードに適切な順位を付け、重要プロセスのCPU不足を防止します。

バッチやレポート作成でパフォーマンスの最大化を目指す大規模な並列ワークロードでは、CPUまたはI/Oの使用率が100%に達することは一般的です。この重要な例にはデータウェアハウスがあります。混合使用システムでバッチおよびレポート作成のパフォーマンスを得るには、インタラクティブな応答時間を犠牲にする必要があります。インタラクティブな応答時間が重要である場合、バッチとレポート作成に使用する並列度を制限する必要があります。こうすることで、バッチとレポート作成がシステム・リソースを過剰に消費することを回避できます。混合ワークロード環境でのベスト・プラクティスは、IORMを有効化し、objective=balancedと設定することです。これにより、小さいI/Oが大きいI/Oより優先されますが、DWワークロードが完了しないほどではありません。

Enterprise Managerは、リソースの使用率と容量を測定するための豊富なツール・セットを提供しています。次の情報は、提供される機能のほんの一例にすぎません。Enterprise Manager向けのExadataプラグインについて、詳しくは次のドキュメントを参照してください。『[Managing Oracle Exadata with Oracle Enterprise Manager 12c](#)』

Enterprise Managerでは、コンピュート・ノードとセル・サーバーの両ターゲットに対するCPU使用率を確認できます。Database Machineターゲットのホームページで、対象のターゲットを選択します。ターゲットのホームページが表示され、CPU情報の概要が表示されます。



コンピュータ・ノードの詳細なCPUパフォーマンス情報を表示するには、Hostドロップダウン・リストから「Host」→「Monitoring」→「CPU Details」を選択します。



メモリ使用率

メモリ使用率を監視するには、/proc/meminfoという仮想ファイルを使用します。利用可能な合計空きメモリの目安を得るには、「MemFree:」メトリックと「Cached:」メトリックを合算します。Linuxは必要に応じてキャッシュからメモリを開放するため、これも空きメモリとしてみなされます。Exadataデータベースは、Linuxページ・キャッシュをデータベースI/Oに利用しないため、Linuxページ・キャッシュを比較的小さく抑える必要があります。次に例を示します。

```
cat /proc/meminfo | egrep '^MemTotal:|^MemFree:|^Cached:'
```

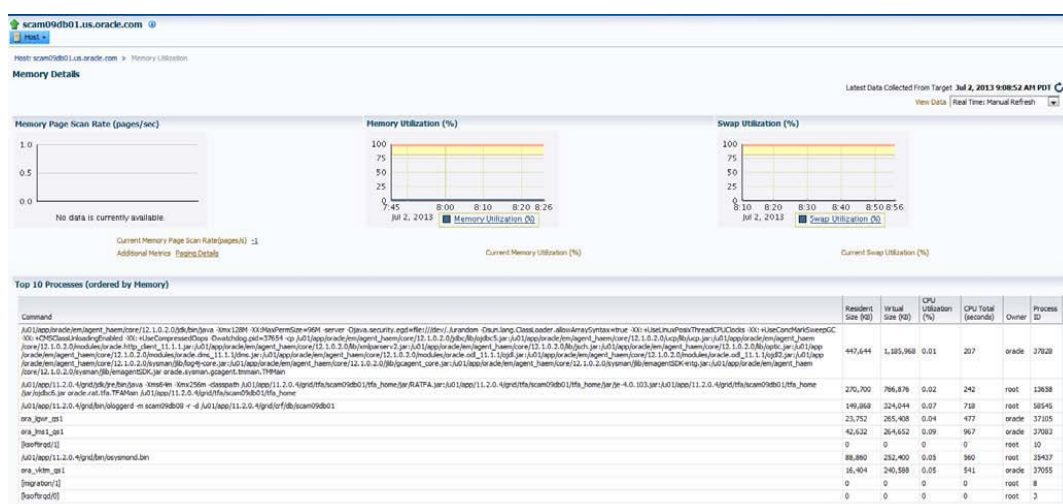
```
MemTotal:      1058421596kB
```

```
MemFree:       488324496kB
```

```
Cached:        972268kB
```

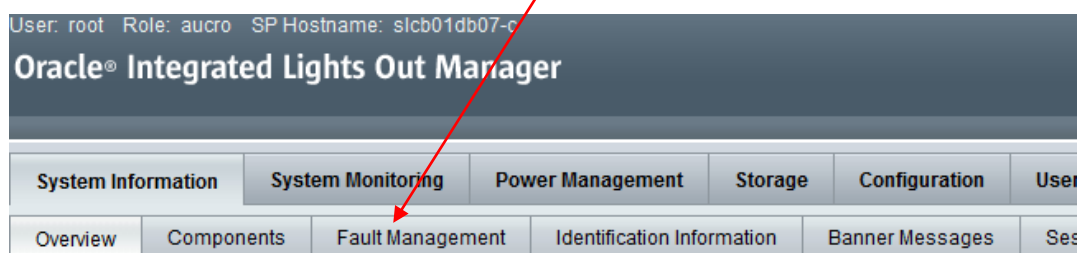
メトリック	サイズ (kB)	割合 (%)
MemTotal:	1,058,421,596	100.0%
MemFree:	488,324,496	46.1%
Cached:	972,268	0.1%
空きメモリ (計算値)	489,296,764	46.2%

メモリ使用率には、上述のCPU使用率と同じEnterprise Managerの画面からアクセスできます。コンピュータ・ノードで、メモリ使用率に関する追加情報にアクセスするには、コンピュータ・ノードのホームページで、「Host」を選択し、Hostドロップダウン・リストから「Monitoring」→「Memory Details」を選択します。



ILOMイベント

ILOM (Integrated Lights Out Manager) は、サーバーの管理と監視に使用される専用サービス・プロセッサです。それぞれのストレージ・サーバー、コンピュータ・ノード、InfiniBandスイッチに専用のILOMが構成されています。ILOMのエラーとメッセージは、複数の場所から参照できます。1番目はWeb管理コンソールです。Webコンソールから「Fault Management」を選択します。



System Overview

View system summary information. You may also change power state and view system status and fault information.

Product Name: SUN FIRE X4170 SERVER

Part/Serial Number: 4384281-1 / 0923XF503F

Host Power: **On**

System Status: Normal

BIOS Version: 07060304

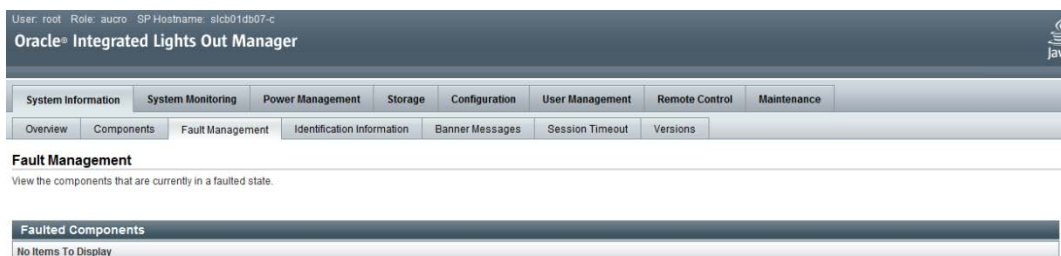
SP Hostname: slcb01db07-c

Uptime: 43 days, 19:09:35

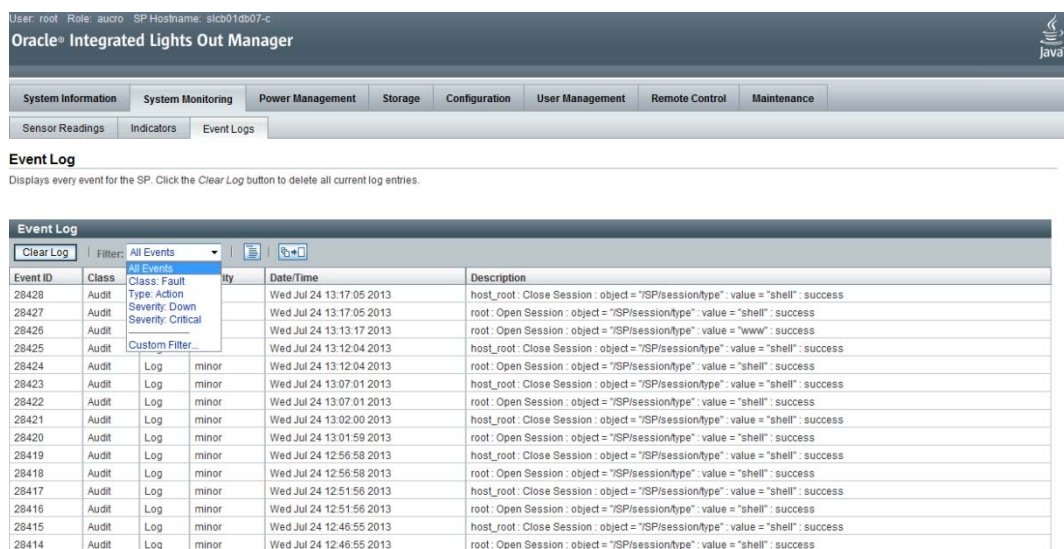
IP Address: 10.245.19.14

ILOM Version: v3.0.16.15.a.r73751

これにより、CPU、DIMM、ファンなどの障害の発生したコンポーネントが表示されます。



システム・イベント・ログを確認すると、障害と分類されていない追加情報（速度超過/不足のファン、温度センサーの警告など）を参照できます。「System Monitoring」タブ→「Event Logs」タブの順に選択します。Filterドロップダウン・リストを利用すると、タイプでフィルタリングできます。



また、ipmitoolを使用して、ILOMホストからシステム・イベント・ログを表示できます。たとえば、ログ・ファイル内の最新イベントを10個リストするには、次のコマンドを実行します。

```
# ipmitool sel list 10
1 | 10/01/2008 | 11:32:43 | Power Supply #0x04 | Failure detected | Asserted
2 | 07/13/2009 | 15:02:31 | Power Supply #0x05 | Failure detected | Asserted
3 | 11/11/2009 | 13:18:20 | Fan #0x09 | Transition to Degraded
4 | 11/11/2009 | 13:18:20 | Fan #0x09 | Transition to Running
5 | 11/11/2009 | 13:22:36 | Fan #0x09 | Transition to Running
6 | 11/11/2009 | 13:22:37 | Fan #0x09 | Transition to Degraded
7 | 11/11/2009 | 13:22:37 | Fan #0x09 | Transition to Running
8 | 11/11/2009 | 13:22:38 | Fan #0x09 | Transition to Running
9 | 11/11/2009 | 13:22:39 | Fan #0x09 | Transition to Degraded
a | 11/11/2009 | 13:22:39 | Fan #0x09 | Transition to Running
```

ネットワーク・ステータス

Exadataシステムに対する完全なネットワーク・パフォーマンスの診断は、このホワイト・ペーパーの対象外です。ただし、すべてのネットワークが稼働中であることを確認するために実行できる簡単なチェックがいくつかあります。

コンピュータ・ノードで、クラスタ・レベルからVIPのステータスを確認します。

```
$ srvctl status vip -n node1
VIP slcb0107-vip is enabled
VIP slcb0107-vip is running on node: node1
```

オペレーティング・システムから次のコマンドを実行すると、すべてのネットワークが稼働中であることを確認できます。

```
dcli -l root -g ./all_group "ifconfig -a | grep DOWN"
```

dcliは、同じコマンドを複数ノードで実行するためのユーティリティです。このコマンドについて、詳しくは次の記事を参照してください。

<http://www.oracle.com/technetwork/articles/oem/exadata-commands-part3-402445.html>

-gフラグは、コマンドを実行する全ノードのリストを含むテキスト・ファイルを指します。このインスタンスでは、コマンドをすべてのコンピュータ・ノードとストレージ・サーバーで実行する必要があります。

コマンドから出力が返されない場合は、定義されたすべてのネットワークが停止していないことを示します。

Oracle Enterprise Managerもまた、ネットワーク・ステータスのチェックに適しています。ホスト・ターゲットのホームページから、「Host」→「Monitoring」→「All Metrics」の順に選択します。

The screenshot displays the Oracle Enterprise Manager Cloud Control 12c interface. The top navigation bar includes 'Enterprise', 'Targets', 'Favorites', and 'History'. The main content area shows a 'Host' dropdown menu with 'Monitoring' selected. A secondary menu is open under 'Monitoring', with 'All Metrics' highlighted by a red arrow. The interface also shows a 'Job Activity' section with a search bar and a table of job runs, and two charts on the right: 'CPU Utilization' and 'Filesystem Usage'.

ORACLE Enterprise Manager Cloud Control 12c

Enterprise ▾ Targets ▾ Favorites ▾ History ▾

dadzab07.us.oracle.com ⓘ

Host ▾

- Home
- Monitoring ▾
 - CPU Details
 - Memory Details
 - Disk Details
 - Program Resource Utilization
 - All Metrics
 - Metric and Collection Settings
 - Metric Collection Errors
 - User-Defined Metrics
- Control ▾
- Job Activity
- Information Publisher Reports
- Administration ▾
- Net Services Administration
- Log File Alerts
- Storage Details
- Remote File Editor
- Execute Host Command
- Privilege Delegation Setting
- Related Targets
- Configuration ▾
- Compliance ▾
- Target Setup ▾
- Target Information

Job Activity

Summary of jobs whose start date is within the last 7 days.

Show Latest Run ▾ Search Job Name ▾

View ▾ Status ✖ 0 🔄 0 🕒 0 ✅ 0

Name	Status	Started	Job Type
No job runs found			

CPU and Memory

CPU Utilization

100
80
60
40
20
0

12:45 PM
July 24

■ CPU in U

Filesystem and I/O

Filesystem Usage

100
80
60
40

次に「Network Interfaces」を選択します。

The screenshot shows the Oracle Enterprise Manager interface for the host `dadzab07.us.oracle.com`. The page displays a list of metrics under the heading "All Metrics". A red arrow points to the "Network Interfaces" metric in the list. The list includes various system metrics such as Buffer Activity, CPU Usage, Compute Node Temperature, Disk Activity, File and Directory Monitoring, Filesystems, HCA Port Errors, HCA Port State, Host Storage Support, Load, Log File Monitoring, Network Interfaces, Network Interfaces Summary, Paging Activity, Process, Inode, File Tables Statistics, Program Resource Utilization, Response, and Response of Reliable Datagram Sockets protocol layer. On the right side of the page, there are two empty tables: "Open Metric Event" and "Top 5 alerting met", both with a "Metric Name" header and "No data to display." content.

すべてのネットワーク・インタフェースとその統計情報がページに表示されます。

Network Interfaces
 Collection Schedule Every 15 Minutes [Modify](#)
 Upload Interval Every Collection
 Last Upload Jul 24, 2013 1:33:03 PM MDT

Network Interface Name	Network Interface Collisions (%)	Network Interface Combined Utilization (%)	Network Interface Input Errors (%)	Network Interface Output Errors (%)	Network Interface Read (MB/s)	Network Interface Read Utilization (%)	Network Interface Total Error Rate (%)	Network Interface Total I/O Rate (MB/sec)	Network Interface Write (MB/s)	Network Interface Write Utilization (%)
> eth0	0	0	0	0	0.01	0	0	0.01	0	0
> b0	0	0	0	0	0	0	0	0	0	0
> b1	0	0	0	0	0	0	0	0	0	0
> bond0	0	0	0	0	0	0	0	0	0	0

Data shown in above table is collected in real time.

ディスク・ステータス

Megacliコマンドを使用すると、すべてのディスクがオンライン状態にあるかどうかを簡単に確認できます。次に、すべてのディスクのステータスを表示するサンプル・コマンドを示します。実行対象のグループは、すべてのコンピュート・ノードとストレージ・ノードのリストを含むテキスト・ファイルです。

```
# dcli -g all_group -l root /opt/MegaRAID/MegaCli/MegaCli64 AdpAllInfo
-aALL | grep "Device Present" -A 8

slcb01db07: Device Present
slcb01db07: =====
slcb01db07: Virtual Drives      : 1
slcb01db07: Degraded            : 0
slcb01db07: Offline              : 0
slcb01db07: Physical Devices     : 5
slcb01db07: Disks              : 4
slcb01db07: Critical Disks       : 0
slcb01db07: Failed Disks          : 0
--
slcb01db08: Device Present
slcb01db08: =====
slcb01db08: Virtual Drives      : 1
slcb01db08: Degraded            : 0
slcb01db08: Offline              : 0
slcb01db08: Physical Devices     : 5
slcb01db08: Disks              : 4
```

```
slcb01db08: Critical Disks : 0
slcb01db08: Failed Disks  : 0
--
slcb01cel12: Device Present
slcb01cel12: =====
slcb01cel12: Virtual Drives : 12
slcb01cel12: Degraded       : 0
slcb01cel12: Offline        : 0
slcb01cel12: Physical Devices : 14
slcb01cel12: Disks          : 12
slcb01cel12: Critical Disks  : 0
slcb01cel12: Failed Disks   : 0
--
slcb01cel13: Device Present
slcb01cel13: =====
slcb01cel13: Virtual Drives : 12
slcb01cel13: Degraded       : 0
slcb01cel13: Offline        : 0
slcb01cel13: Physical Devices : 14
slcb01cel13: Disks          : 12
slcb01cel13: Critical Disks  : 0
slcb01cel13: Failed Disks   : 0
```

いずれかのドライブに対して機能低下またはオフラインが示された場合、追加の措置を講じて、問題のあるディスクを特定、修正する必要があります。

CheckHWnFWProfile

CheckHWnFWProfileは、コンピュータ・ノードとストレージ・ノード上にあるハードウェアとファームウェアがサポート構成になっているかどうかを検証するプログラムです。このプログラムの実行にはわずかに数秒しかかからず、次に示すような、サポートされないディスクなどの問題を特定するために役立ちます。Exachkも問題を確認するためにこのコマンドを実行します。

```
# dcli -l root -g ./all_group
"/opt/oracle.SupportTools/CheckHWnFWProfile"
```

```
slcb01db07: [SUCCESS] The hardware and firmware profile matches one of
the supported profiles
```

```
slcb01db08: [SUCCESS] The hardware and firmware profile matches one of
the supported profiles
```

```
slcb01cel12: [INFO] All drives are not identical
```

```
slcb01cel12: [ERROR] Incompatible mix of disk models. All models must
be from the list: No supported models found.
```

```
slcb01cel13: [INFO] All drives are not identical
```

```
slcb01cel13: [ERROR] Incompatible mix of disk models. All models must
be from the list: No supported models found.
```

サービス

何らかのサービスが定義されている場合、これらもチェックする必要があります。これには、lsnrctl コマンドを使用します。SCANリスナーをチェックすると、すべての該当インスタンスのハンドラが稼働中であるかどうかを検証されます。次の例では、クラスタ内の2つのインスタンス上にサービス haem_dbfs.us.oracle.com のサービス・ハンドラがあることが分かります。

```
$ lsnrctl status LISTENER_SCAN2
```

```
LSNRCTL for Linux: Version 11.2.0.3.0 - Production on 24-JUL-2013 13:37:51
```

```
Copyright (c) 1991, 2011, Oracle. All rights reserved.
```

```
Connecting to (DESCRIPTION=(ADDRESS=(PROTOCOL=IPC)(KEY=LISTENER_SCAN2)))
STATUS of the LISTENER
```

```
-----
Alias          LISTENER_SCAN2
Version        TNSLSNR for Linux: Version 11.2.0.3.0 - Production
Start Date     18-JUN-2013 19:54:18
Uptime         35 days 17 hr. 43 min. 34 sec
Trace Level    off
Security       ON: Local OS Authentication
SNMP           OFF
Listener Parameter File /u01/app/11.2.0/grid/network/admin/listener.ora
Listener Log File
/u01/app/11.2.0/grid/log/diag/tnslsnr/slcai604/listener_scan2/alert/log.xml
Listening Endpoints Summary...
  (DESCRIPTION=(ADDRESS=(PROTOCOL=ipc)(KEY=LISTENER_SCAN2)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=0.0.0.0)(PORT=1525)))
Services Summary...
Service "haem_dbfs.us.oracle.com" has 2 instance(s).
  Instance "haemS1", status READY, has 1 handler(s) for this service...
  Instance "haemS2", status READY, has 1 handler(s) for this service...
```


Exachk

Enterprise ManagerでExadataのヘルスチェック・プラグインを使用すると、Exachkによる最新の出力が表示されます。ヘルスチェック・プラグインの構成について、詳しくは次のドキュメントを参照してください。

http://docs.oracle.com/cd/E11857_01/install.111/e23441/pick.htm

構成の完了したヘルスチェック・ターゲットは、MAAベスト・プラクティスに準拠していない領域を識別するためのクイック・リファレンス・ポイントになります。

The screenshot displays the Oracle Enterprise Manager interface for a target named 'scac02db01_1nc'. The 'Summary' section shows the target type as 'Oracle Engineered System Healthchecks', the version as '2.2.2_20130617', and the Exadata type as 'K2-2'. The 'Incidents and Problems' table lists two problems: 'java.lang.Throwable:...' and 'oracle.sysman.guagen:...', both with a severity of 'New'. The 'Exachk Execution Results Summary' table provides a detailed view of the health check results, including metrics, check names, node and/or database information, DB instance, InnoDB parameters, status, and collection timestamps. A 'Flashback database on standby' section provides additional context and recommendations.

Metric	Check Name	Node and/or Database	DB Instance	InnoDB Parameter	Status	Outfile Path	Collection Timestamp
Database Checks	Flashback database on primary	sp			FAB.		Jul 25, 2013 10:02:35 AM #
Database Checks	Flashback database on primary	spst			FAB.		Jul 25, 2013 10:02:35 AM #
Database Checks	Flashback database on standby	spdev			FAB.		Jul 25, 2013 10:02:35 AM #
Database Checks	Flashback database on standby	spstby			FAB.		Jul 25, 2013 10:02:35 AM #
Database Checks	Logical standby unsupported datatypes	sp			FAB.		Jul 25, 2013 10:02:35 AM #
Database Checks	Not Available	sp			FAB.		Jul 25, 2013 10:02:35 AM #
Database Checks	Not Available	sp			FAB.		Jul 25, 2013 10:02:35 AM #
Database Checks	Not Available	spdev			FAB.		Jul 25, 2013 10:02:35 AM #

Flashback database on standby

Exachk Results From: exachk
 Exachk Results Timestamp: Jul 17, 2013 10:31:48 AM acfsc
 Fail Message: Flashback on STANDBY is not configured
 Risk Message:
 Benefit/Impact Message:
 Recommendation Message: Oracle Flashback Technology enables fast logical failure repair. Oracle recommends that you use automatic undo management with sufficient space to attain your desired undo retention guarantee, enable Oracle Flashback Database, and allocate sufficient space and I/O bandwidth in the fast recovery area. Application monitoring is required for early detection. Effective and fast repair comes from leveraging and rehearsing the most common application specific logical failures and using the different flashback features effectively (e.g. flashback query, flashback version query, flashback transaction query, flashback transaction, flashback drop, flashback table, and flashback database).
 Key HA Benefits:
 With application monitoring and rehearsed repair actions with flashback technologies, application downtime can reduce from hours and days to the time to detect the logical inconsistency...Please refer to MOS note 1274318.1 for more information.

* Exachk 2.2.1以降では、2つのExachk出力を比較する機能が提供されています。これは変更の識別に有効です。

ベースライン・データの使用による問題のトラブルシューティング

構成ファイルの比較

構成ファイルの変更は、システムに問題を引き起こしやすい要素です。diffコマンドを使用するだけで、ファイルへの最近の変更が明らかになります。"問題発生前取るべき措置"の項の推奨事項に従って、重要なファイルのバックアップが取得されているはずですが。

spfileとパスワード・ファイルの情報はバイナリであるため、diffコマンドを使用した比較は実行できません。ただし、stringsコマンドを使用してASCIIデータをエクスポートすることで、比較を実行できます。以下を参照してください。

```
$ strings spfileemrep.ora > spfileemrep.ora.txt
$ strings spfileemrep.ora_072513_0100 >
spfileemrep.ora_072513_0100.txt
$ diff spfileemrep.ora.txt spfileemrep.ora_072513_0100.txt
31c31
< *.processes=300
---
> *.processes=600
35c35
< *.sga_target=2147483648
---
> *.sga_target=1147483648
```

** Enterprise Managerで期間比較レポートを実行すると、spfileやinitファイルの違いも確認できます。ただし、データの粒度は収集されたAWRスナップショットと同じになります。

カーネルのチューニング可能パラメータに対する変更のチェック

カーネル設定を既知の正常なコピーと比較することを推奨します。次のコマンドは、以前のベースライン・コピーと現在の構成の間でカーネル・パラメータを比較します。dbs_groupファイルは、すべてのコンピュート・ノードを含む単純なテキスト・ファイルです。

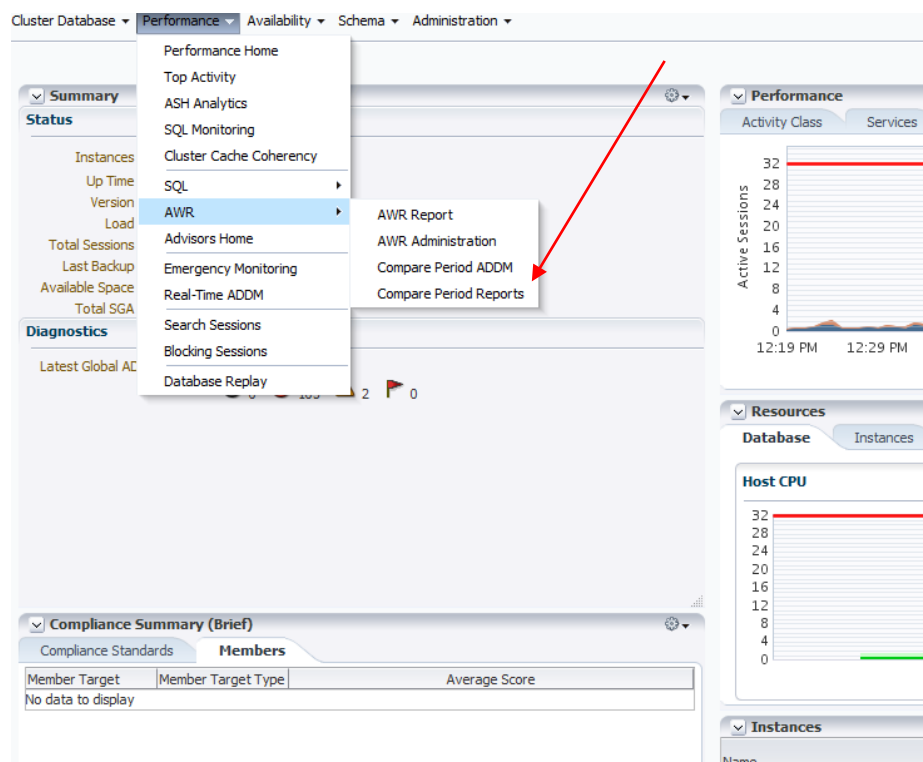
```
# dcli -l root -g ./dbs_group "sysctl -a > /tmp/sysctl.current;diff
/root/<baseline kernel configuration file > /tmp/sysctl.current"
```

*** いくつかのパラメータが動的に変更されるのは正常な処理です。したがって、diff出力の差分が現在の問題に関係あるかどうかを見極めるには、上記の出力を慎重に分析する必要があります。

AWRデータ

AWRデータは、データベースの問題をトラブルシューティングするために必要な重要情報を提供します。AWRの詳細分析については本書の後半で詳しく説明しますが、「問題発生前取るべき措置」で説明したAWRベースラインのデータを使用すると、パフォーマンスやワークロードのずれを簡単に特定できます。ユーザー数、トランザクション数、REDO速度、トランザクションあたりの物理読み取り数と物理書き込み数などの主要メトリックの変化は、システムで発生した変更を素早く特定するために役立ちます。

Enterprise Managerで現在のAWRレポートをベースラインに対して比較するには、データベース・ターゲットのホームページから、「Performance」→「AWR」→「Compare Period Reports」の順に選択します。



次に、比較する期間を選択します。1番目の期間に、パフォーマンス・ベンチマークにもっとも近い静的ベースラインまたは変動ベースラインを選択します。たとえば、期間やワークロード・データの類似したものを比較します。2番目の期間には、パフォーマンスの問題を含むスナップショットを選択します。次に、「Generate Report」ボタンをクリックします。

Cluster Database ▾ Performance ▾ Availability ▾ Schema ▾ Administration ▾

Automatic Workload Repository > Run AWR Report
Run Compare Periods Report

Specify two periods for comparison. For each period, you can either pick one baseline or a pair of snapshots. [Generate Report](#)

First Period	Second Period
<input checked="" type="radio"/> By Baseline Baseline <input type="text" value="Test Baseline"/>	<input type="radio"/> By Baseline Baseline <input type="text"/>
<input type="radio"/> By Snapshot Begin Snapshot <input type="text"/>	<input checked="" type="radio"/> By Snapshot Begin Snapshot <input type="text" value="14879"/>
End Snapshot <input type="text"/>	End Snapshot <input type="text" value="14892"/>
Show Instances	Show Instances

ワークロードやデータベース動作における相違点を示したレポートが生成されます。

Report Summary

Host Configuration Comparison

- CPU and Memory values are from the end snapshot; averaged across all instances
- Other values are averages for all instances

	1st	2nd	Diff	%Diff
Number of CPUs:	32	32	0	0.0
Number of CPU Cores:	16	16	0	0.0
Number of CPU Sockets:	2	2	0	0.0
Physical Memory:	258064.9M	258064.9M	0M	0.0
Load at Start Snapshot:	1.81	1.97	.16	8.8
Load at End Snapshot:	1.9	1.86	-.04	-2.1
%User Time:	1.01	1	-.02	-1.0
%System Time:	.57	.56	-.01	-1.8
%Idle Time:	97.87	97.91	.04	0.0
%IO Wait Time:	.04	.04	0	0.0

Cache Sizes

- Cache Sizes are from the end snapshot; averaged across all instances

	1st (M)	2nd (M)	Diff (M)	%Diff
Memory Target				
SGA Target	2,048.0	2,048.0	0.0	0.0
Buffer Cache	552.0	552.0	0.0	0.0
Shared Pool	1,384.0	1,384.0	0.0	0.0
Large Pool	16.0	16.0	0.0	0.0
Java Pool	24.0	24.0	0.0	0.0
Streams Pool				
PGA Target	1,024.0	1,024.0	0.0	0.0
Log Buffer	10.0	10.0	0.0	0.0

Load Profile

	1st per sec	2nd per sec	%Diff	1st per txn	2nd per txn	%Diff
DB time:	1.1	1.1	-3.6	0.1	0.1	0.0
CPU time:	0.6	0.6	0.0	0.0	0.0	0.0
Redo size:	107,420.0	112,108.8	4.4	5,988.0	6,118.4	2.2
Logical reads:	9,380.8	9,493.7	1.2	522.9	518.1	-0.9
Block changes:	673.4	713.9	6.0	37.5	39.0	3.8
Physical reads:	139.3	119.4	-14.3	7.8	6.5	-16.0
Physical writes:	45.1	38.2	-15.4	2.5	2.1	-17.1
User calls:	88.2	98.1	11.3	4.9	5.4	8.9
Parses:	117.5	119.0	1.2	6.6	6.5	-0.9
Hard parses:	5.2	4.8	-6.2	0.3	0.3	-10.3
W/A MB processed:	3,651,674.2	4,621,325.5	26.6	203,558.7	252,211.0	26.6
Logons:	2.4	2.7	12.7	0.1	0.1	15.4
Executes:	758.5	770.7	1.6	42.3	42.1	-0.5
Transactions:	17.9	18.3	2.1			
				1st	2nd	Diff
% Blocks changed per Read:				7.2	7.5	0.3
Recursive Call %:				95.8	95.4	-0.4
Rollback per transaction %:				66.7	66.0	-0.7
Rollback per Sort:				12.2	11.2	-9.0

詳細診断

ハードウェアの除外

I/Oパフォーマンス

ワークロードが変化した場合の動作予測が難しいのは、ハード・ディスクの特性によるものです。特に、以下が挙げられます。

- ディスク・コントローラにはメモリ・キャッシュがあり、I/Oパフォーマンス（特に書込み）を改善するために使用できます。書込み速度が高くなると、キャッシュが対応できなくなり、書込み待機時間が突然長くなる場合があります。また、ディスク・コントローラの書込みキャッシングのために、オペレーティング・システムが実際のディスク使用率を把握できないため、ディスク使用率は当てにならない場合があります。
- 書込みごとにディスク・ヘッドの位置を変更する必要があるため、ランダム・ディスク処理はシリアル処理よりも時間がかかる可能性があります。
- 大型I/O処理ではデータ送信に時間がかかるため、完了までに長い時間がかかる場合があります。
- 使用率が高くなると、ディスクの空きを待機する時間が長くなるため、I/O待機時間が大幅に伸びる可能性があります。
- ディスクには強いI/O制限が設定されていません。キューイングされるI/Oが増えるほど、ディスクに対するI/Oスループットが高くなります。これは、同時にキューイングされるI/Oが増えると、ディスク・コントローラがよりインテリジェントなI/Oスケジュール設定を実行できるためです。したがって、I/Oスループットと応答時間の間でパフォーマンスの妥協点を見つける必要があります。キューイングするI/Oが増えると、システム・スループットは増えてますが、応答時間が低下します。

x2システムやx3システムのハイパフォーマンス・ディスクでは、応答時間を大幅に伸ばすことなく、1秒あたり約300のSmall I/Oを実行でき（ピーク・パフォーマンスは400IOPS以上）、フルラックでは50,000IOPSを達成できます。Large I/Oには、Small I/Oのおおよそ3倍のコストがかかります。概算のディスク使用率を求めるには、Small I/Oの合計数を計算し、Large I/Oの合計数に3を掛けたものを加えます。次に、この数値を300というIOPSのしきい値と比較すると、使用率が特定できます。3TBの大容量ディスクの場合、IOPSは約120であり、Large I/Oに使用すべき乗数は約2です。

3TBのハイキャパシティディスクでは、IOPSが167前後であり、フルラックの場合は約28,000IOPSになります。また、Large I/Oには2を掛けます。Exadataの容量について、詳しくはOracle Exadata Database Machine X2-2およびX3-2のデータシートを参照してください。

ディスク待機時間の長さは必ずしも問題になるとは限らず、アプリケーションが受ける影響によって決まります。データウェアハウスの場合、問合せ処理時のディスクのスループットと待機時間が最大になることにまったく問題はなりません。

幸いにも、Exadataのフラッシュ・キャッシュとフラッシュ・ログはディスクI/Oを軽減し、ディスク待機時間の拡大による影響の多くからアプリケーションを隔離します。

- Smart Flash logは、ディスクのREDO書き込み待機時間が長くなることがあっても、素早くコミットを完了します。
- Smart Flash cacheはフラッシュから直接データを提供することで、ディスク読取りとディスク書き込みの両方のデータベース待機時間を短縮します。データがロードされ、キャッシュからエージ・アウトされる際、フラッシュ・キャッシュのパフォーマンスは、そのサイズと使用率の影響を受けます。バッチ・ジョブやレポート作成ジョブによってディスク待機時間が長くなることもある場合、フラッシュ・キャッシュを使用することでインタラクティブ・ユーザーを隔離し、適切な応答時間を維持できます。

ディスク使用率が上昇しているか、上昇が予想される場合は、パフォーマンスへの影響を注意深く監視する必要があります。特に、次の事項に注意します。

- "cell single block physical read"と"log file parallel write"の待機イベント時間の増加
- "cell multiblock physical read"、"log file sync"、"db file parallel write"の時間の増加が重要になる場合もありますが、これらのイベントの待機時間はしばしば、通常のワークロードでも大きく変動するため、傾向の監視のためにはそれほど有用ではありません。
- OLTPアプリケーションの応答時間の増加、またはバッチ実行時間の増加。

CPUバウンドのコンピュータ・ノードでは、I/Oが完了しても、I/Oを発行するプロセスがすぐに再スケジュールされないためI/O待機時間が長くなりますが、これは正確なものではないため注意が必要です。このため、CPUのスケジュール時間はI/O待機時間の一部として測定されます。したがって、CPUが上限に達した場合にデータベース・レベルで測定されるI/O応答時間は正確ではありません。

コンピュータ・ノードのCPUが上限に達した場合にI/O使用率を正確に測定するには、セル・サーバーのメトリックを使用します。

cellcli -e list metristoryコマンドを使用すると、セル・サイドの多数のI/Oメトリックを確認できます。完全なリストについてはドキュメントを参照するか、cellcli -e list metricdefinition detailを参照してください。頻繁に使用される時間メトリックの1つに、CD_IO_ST_RQ (Small I/Oに対するリクエストあたりの平均サービス時間) があります。次に例を示します。

```
cellcli -e list metristory CD_IO_ST_RQ where metricObjectName like 'CD.*' and collectionTime > '2013-07-15T14:00:00-07:00' and collectionTime < '2013-07-15T14:30:00-07:00' attributes collectionTime,metricObjectName,metricValue.
```

これは、今日の午後2時から2時30分までの間の各ディスクのサービス時間を1分の収集間隔で表示します。

I/O使用率が高い場合に考えられる改善措置は次のとおりです。

- I/O率の高いSQLのチューニング - 実行I/O数を減らせる、より良い問合せ計画があるかもしれません。ディスクが上限に達している場合、特に、"最適化されていない"読取りの多いSQLに注意を払います。最適化された読取りはフラッシュ・キャッシュからデータを読み取るため、ディスク使用率の上昇の一因にはなりません。
- スマート・フラッシュ・キャッシュとスマート・フラッシュ・ログが構成済みであることを確認します。
- 重要なセグメントやヒット率の高いセグメントは、バッファ・キャッシュまたはフラッシュ・キャッシュ内に"保持"します。
- Exadataバージョン11.2.3.2.1およびOracle Grid Infrastructure 11.2.0.3 BP9以上を実行するシステムの場合、"free buffer waits"を監視して、その発生率が高い場合は、Writeback Flash cacheを有効化します。
- 圧縮を実装することでデータ・サイズを縮小し、結果的にアプリケーション実行に必要なI/O数を削減します。

セルがI/Oバウンドであるかどうかのチェック

パフォーマンスが低いのか、使用率が上限に達しているのかを区別します。

- OS Watcher iostat やセル・ディスクのメトリックを使用して、HDD とフラッシュの合計スループット (MBPS) と IOPS を計算します。
- 最大値については、Exadata Database Machine の[データシート](#)を参照します。
 - I/O が最大値に近づいている場合は、待機時間の拡大に注意します。
 - 待機時間が長ければ、ディスクのパフォーマンスが低いとは限りません。
 - I/O に時間がかかるのは、おもにディスク・キューを待機する時間に起因します。
 - I/O待機時間は100ミリ秒を超えることもあります (ディスクのパフォーマンスは低くありません)。
 - I/O待機時間はディスク・キューの長さに依存するため、ワークロードによって異なります。
 - MBPS と IOPS が同時に最大値に達することはないことに留意してください。
 - 混合ワークロードの評価方法
 - ディスク使用率の調査 - 100%に近づいているかどうか。
 - 必要に応じてキャリブレーションを実行 (セルの停止が必要)。

セルがI/Oバウンドである場合に考慮すべき事項

- SAME (Stripe And Mirror Everywhere)
 - ・ いずれかのディスクまたはセルが上限に達した場合は、ワークロードが平行化されていても、これらのディスクまたはセルによってパフォーマンスが低下します。
 - ・ ラックを小さいクラスタに分割するのではなく、できる限り多くのセルを使用します。
- システム・ディスクとデータ・ディスクでは、潜在的にパフォーマンスが異なる点に注意します。
 - ・ システム・ディスクには、ユーザー・データだけでなくセル自体のファイル・システムも含まれます。
 - ・ システム・ディスクのパフォーマンスは、データ・ディスクを下回る可能性があります。
 - 3TBのハイキャパシティドライブはIOPS容量が低いため、600GBのハイパフォーマンス・ドライブと比較すると違いが顕著になります。

Exadataの最大値

X3-2 X3-8

X2-2 X2-8

システム	デバイスあたりの最大スループット	デバイスあたりの最大IOPS	セルあたりの最大スループット	セルあたりの最大IOPS
ハイパフォーマンスHDD (600G)	152MB/秒	29710/秒	1.8GB/秒	35710/秒
ハイキャパシティHDD (3TB)	108MB/秒	14810/秒	1.3GB/秒	178510/秒
フラッシュ	228MB/秒 (読取り専用)	781210/秒 (読取り専用)	3.657GB/秒 (読取り専用)	125,00010/秒 (読取り専用)

OS Watcher IOSTATがある場所

セル上で/opt/oracle.oswatcher/osw/archive/oswostatへ移動します。

- ・ すべてのデバイス (HDDとフラッシュの両方) に対して、5秒ごとにスナップショットが取得されます。
- ・ スナップショットは1時間ごとに1つのファイルに格納されます。
- ・ スナップショットは7日間保持されます。

OS Watcher IOSTATの解析方法

列名	内容
デバイス	デバイス名 (デバイスからセル・ディスクへマッピング)
r/s	1秒あたりの読取りI/O数
w/s	1秒あたりの書込みI/O数
rsec/s	1秒あたりの読取りセクター (0.5KB) 数
wsec/s	1秒あたりの書込みセクター (0.5KB) 数
avgrq-sz	セクター (0.5KB) 内のI/Oリクエストの平均サイズ
avgqu-sz	平均ディスク・キュー・サイズ
Await	平均I/O待機時間 (ミリ秒、サービス時間とキュー内の時間を含む)
Svctm	平均I/Oサービス時間 (ミリ秒)
%util	デバイス帯域幅の使用率 (0~100%)

OS Watcher IOSTATの分析方法

2TBの大容量セルを使用したHDDスナップショット

```

avg-cpu:  %user   %nice %system %iowait  %steal   %idle
           10.72    0.00    0.34    5.50    0.00   83.44

Device:            rrqm/s   wrqm/s   r/s     w/s    rsec/s   wsec/s  avgrq-sz  avgqu-sz   await  svctm   %util
sda (system)    53.20   20.20 61.00 12.20 119305.60 275.20 1633.62   50.51 704.65 12.94 94.74
sdb (system)    63.00   21.60 67.40 9.60 134588.80 278.40 1751.52   185.82 2424.44 12.99 100.02
sdc                77.80    0.00  80.00   2.60  154174.40    11.60  1866.66    15.03  158.55   9.59  79.24
sdd                77.40    0.00  78.60   0.40  156766.40    12.80  1984.55    13.10  163.34   9.46  74.70
sde                53.00    0.00  64.80   0.20  129342.40     1.60  1989.91    28.14  506.29   9.95  64.68
sdf                67.60    0.00  70.00   0.40  138744.00    32.00  1971.25     9.12  130.45  10.51  73.96
sdg                88.40    0.00  83.00   0.80  159870.40    57.60  1908.45    15.11  148.26   9.76  81.76
sdh                86.80    0.00  83.00   4.60  165675.20  1640.00  1909.99    14.57  139.70   9.28  81.26
sdi                93.60    0.00  94.20   0.00  188516.80     0.00  2001.24    18.27  183.00   9.34  87.96
sdj                83.40    0.00  89.80   3.20  178910.40  1235.20  1937.05    12.79  140.77   8.79  81.76
sdk                78.20    0.00  77.80   0.60  155710.40    19.20  1986.35    11.05  138.07   9.84  77.16
sdl                75.20    0.00  79.60   1.60  155230.40    64.00  1912.49    16.34  172.51   9.67  78.50

```

スナップショットからどのような特性が読み取れるでしょうか。

- 大部分は読取りのワークロード
- 平均リクエスト・サイズが約 1MB の読取り
- 最初の 2 つのデバイス（システム・ディスク）は使用率が 100%近い
- これらのデバイスは、データシートに記載された最大スループット (85MB/秒) に近いスループットを提供している
 - ディスク・サブシステムの上限に達する大きい読取りを含むデータウェアハウスのワークロード

- セル・ディスクおよびフラッシュ・キャッシュのメトリックの使用

OSW IOSTATを使用する以外に、セル・ディスクとフラッシュ・キャッシュのメトリックからもセルのI/O特性に対する新しい洞察が得られます。

セル・ディスクのスループット・メトリック

セル・ディスクのIOPSメトリック

フラッシュ・キャッシュのスループット・メトリック

フラッシュ・キャッシュのIOPSメトリック

次のメトリックは、HDDセル・ディスクのI/Oスループットを示します。

HDDセル・ディスクのI/Oスループット・メトリックについて、詳しくはマニュアル『Oracle Exadata Storage Server Software User's Guide』のMonitoring Cell Disk Metricsを参照してください。

メトリック	内容
CD_IO_BY_R_LG_SEC	Large I/O使用時の1秒あたりの読取りサイズ (MB)
CD_IO_BY_W_LG_SEC	Large I/O使用時の1秒あたりの書込みサイズ (MB)
CD_IO_BY_R_SM_SEC	Small I/O使用時の1秒あたりの読取りサイズ (MB)
CD_IO_BY_W_SM_SEC	Small I/O使用時の1秒あたりの書込みサイズ (MB)

Large I/Oとは、128KBを上回るI/Oです。

Small I/Oとは、128KB以下のI/Oです。

HDDセル・ディスクのI/Oスループット・メトリックの分析方法

4つのメトリックを合計し、[X2](#)または[X3](#)のデータシートの値と比較します。

ワークロードがHDDのI/Oスループットの制約を受ける場合、大きいリクエストが優位に立つ傾向にあります。

```
CELLCLI> list metristory CD_IO_BY_R_LG_SEC where collectionTime <
"2011-10-25T04:00:00-05:00" and collectionTime > "2011-10-25T03:59:00-
05:00" and metricObjectName like 'CD_*';
```

HDDセル・ディスクのIOPSを示すメトリック

メトリック	内容
CD_IO_RQ_R_LG_SEC	1秒あたりのLarge読取りリクエストの数
CD_IO_RQ_R_SM_SEC	1秒あたりのSmall読取りリクエストの数
CD_IO_RQ_W_LG_SEC	1秒あたりのLarge書き込みリクエストの数
CD_IO_RQ_W_SM_SEC	1秒あたりのSmall書き込みリクエストの数

HDDセル・ディスクのI/Oスループット・メトリックについて、詳しくはマニュアル『Oracle Exadata Storage Server Software User's Guide』の Monitoring Cell Disk Metricsを参照してください。

Large I/Oとは、128KBを上回るI/Oです。

Small I/Oとは、128KB以下のI/Oです。

- 4つのメトリックを合計し、データシートの値と比較します。
- ワークロードがHDDのIOPSの制約を受ける場合、小さいリクエストが優位になる傾向にあります。
 - ```
CELLCLI> list metristory CD_IO_RQ_R_SM_SEC where
collectionTime < "2011-10-25T04:00:00-05:00" and
collectionTime > "2011-10-25T03:59:00-05:00" and
metricObjectName like 'CD_*';
```

## フラッシュ・キャッシュのI/OスループットとIOPSを示すメトリック

フラッシュ・ディスクのI/OスループットとIOPSのメトリックについては、マニュアル『Oracle Exadata Storage Server Software User's Guide』のMonitoring Cell Disk Metricsを参照してください。

| メトリック             | 内容                                 |
|-------------------|------------------------------------|
| CD_IO_BY_R_SM_SEC | セル・ディスクから小ブロックで読み取る1秒あたりのサイズ (MB)  |
| CD_IO_RQ_R_SM_SEC | セル・ディスクから小ブロックを読み取るための1秒あたりのリクエスト数 |

フラッシュ・キャッシュには32Kのキャッシュ行が含まれるため、フラッシュ・キャッシュI/Oは常に小さくなります。

### フラッシュ・キャッシュ・メトリックの分析方法

- セル・ディスクのメトリックを使用し、データシートの値と比較します。
  - CELLCLI> list metrichistory CD\_IO\_BY\_R\_SM\_SEC where collectionTime < "2011-10-25T04:00:00-05:00" and collectionTime > "2011-10-25T03:59:00-05:00" and metricObjectName like 'FD.\*';
  - CELLCLI> list metrichistory CD\_IO\_RQ\_R\_SM\_SEC where collectionTime < "2011-10-25T04:00:00-05:00" and collectionTime > "2011-10-25T03:59:00-05:00" and metricObjectName like 'FD.\*';
- その他のフラッシュ・キャッシュ・メトリックについては、マニュアル『Oracle Exadata Storage Server Software User's Guide』の Monitoring Flash Cache Metricsを参照してください。

### セルがI/Oバウンドであるかどうかのチェック

混合ワークロード (DWとOLTP) の場合、上記のチェックでMBPSとIOPSのいずれも最大値に達していなくても、セルがI/Oの制約を受ける可能性があります。

このような場合はどうすれば良いのでしょうか。

- OSW IOSTATのデバイス帯域幅に対する使用率を調べます。
- HDDについては、すべてのデータベースに対してDB\_IO\_UTIL\_LGとDB\_IO\_UTIL\_SMの値を合計し、合計使用率が100%に近いかどうかを確認します。
- 不確かな場合は、キャリブレーションを実行してパフォーマンスの低いディスクを除外します (セルは停止する必要があるため、停止時間を回避するにはローリング方式で実施する必要があります)。

### セルがI/Oバウンドである場合のI/O消費者の特定

上記のチェックからセルがI/Oバウンドであることが判明し、複数のデータベースが同じストレージ・セルを共有している場合、セル・データベースのI/Oメトリックを使用して、I/Oを大量に消費しているユーザーまたはデータベースを特定します。

### データベースHDDのI/O分類を示すメトリック

データベースHDDのI/Oメトリックについては、マニュアル『Oracle Exadata Storage Server Software User's Guide』のMonitoring IORM with Database Metricsを参照してください。

- IORM が使用されていない場合も利用できます。

| メトリック         | 内容                                                            |
|---------------|---------------------------------------------------------------|
| DB_IO_UTIL_LG | このメトリックは、このデータベースの大きいリクエストによって使用されるHDDディスク・リソースの割合 (%) を示します。 |
| DB_IO_UTIL_SM | このメトリックは、このデータベースの小さいリクエストによって使用されるHDDディスク・リソースの割合 (%) を示します。 |

### セル・データベースHDDのI/Oメトリックの分析方法

- データベースごとに両方のメトリックを合計し、もっともI/Oを消費しているデータベースを特定します。

```

- CELLCLI> list metrichistory DB_IO_UTIL_LG where
collectionTime < "2011-10-25T04:00:00-05:00" and
collectionTime > "2011-10-25T03:59:00-05:00 " and
metricObjectName = 'DW_db';

- CELLCLI> list metrichistory DB_IO_UTIL_SM where
collectionTime < "2011-10-25T04:00:00-05:00" and
collectionTime > "2011-10-25T03:59:00-05:00 " and
metricObjectName = 'DW_db';

```

- `_OTHER_DATABASE_` (ASMやクラスタウェア) とその他のI/Oに注意します。
- `_OTHER_DATABASE_` を含むすべてのデータベースに対して両メトリックを合計し、セルのHDD使用率が100%に近づいているかどうかを確認します。

### 例 - セル・データベースHDDのI/O分類

- セル・データベースのメトリックを使用して、HDDのI/Oリソースをもっとも消費しているデータベースを特定します。

```
CELLCLI> list metristory DB_IO_UTIL_LG where collectionTime < "2011-10-25T05:13:00-05:00" and collectionTime > "2011-10-25T05:11:00-05:00"
```

```

DB_IO_UTIL_LG DSS 75 % 2011-10-25T05:11:30-05:00
DB_IO_UTIL_LG DW 10 % 2011-10-25T05:11:30-05:00
DB_IO_UTIL_LG MTXDB 5 % 2011-10-25T05:11:30-05:00
DB_IO_UTIL_LG _OTHER_DATABASE_ 0 % 2011-10-25T05:11:30-05:00
DB_IO_UTIL_LG DSS 43 % 2011-10-25T05:12:30-05:00
DB_IO_UTIL_LG DW 8 % 2011-10-25T05:12:30-05:00
DB_IO_UTIL_LG MTXDB 6 % 2011-10-25T05:12:30-05:00
DB_IO_UTIL_LG _OTHER_DATABASE_ 0 % 2011-10-25T05:12:30-05:00

```

#### データベース・フラッシュ・キャッシュのI/O分類を示すメトリック

データベース・フラッシュ・キャッシュのI/Oメトリックについては、マニュアル『Oracle Exadata Storage Server Software User's Guide』のMonitoring IORM with Database Metricsを参照してください。

- IORM が使用されていない場合も利用できます。

| メトリック           | 内容                                                         |
|-----------------|------------------------------------------------------------|
| DB_FC_IO_BY_SEC | このメトリックは、このデータベースからフラッシュ・キャッシュに対する1秒あたりのI/O容量 (MB) を示します。  |
| DB_FC_IO_RQ_SEC | このメトリックは、データベースからフラッシュ・キャッシュに対して発行される1秒あたりのI/Oリクエスト数を示します。 |

#### データベース・フラッシュ・キャッシュのI/Oメトリックの分析方法

- データベース・フラッシュ・キャッシュのI/Oメトリックをデータベースごとにソートし、もっともI/Oを消費しているデータベースを特定します。
  - スループット (DW) :
    - ```
CELLCLI> list metristory DB_FC_IO_BY_SEC where collectionTime < "2011-10-25T04:00:00-05:00" and collectionTime > "2011-10-25T03:59:00-05:00" and metricObjectName = 'DW_db';
```
 - IOPS (OLTP) :
 - ```
CELLCLI> list metristory DB_FC_IO_RQ_SEC where collectionTime < "2011-10-25T04:00:00-05:00" and collectionTime > "2011-10-25T03:59:00-05:00" and metricObjectName = 'OLTP_db';
```

**セルがI/Oバウンドであると特定された場合に考慮すべき事項**

- もっともI/Oを消費しているデータベースに焦点を絞る
- 典型的な問題シナリオ
  - データウェアハウスのワークロード
    - ディスクの使用率は高くても、フラッシュはアイドル状態にある：オブジェクトを KEEP として指定できるか
  - OLTPワークロード
    - 読取りは問題ないが、ディスクは書き込み IOPS の制約を受ける：チェックポイントが意欲的すぎるか、またはバッファ・キャッシュ・サイズが小さすぎないか
- 一般的な戦略
  - データウェアハウス：
    - 上位SQLのチューニング
    - 余分なCPUがある場合は圧縮を使用 - CPUと引換えにディスクを獲得
  - データウェアハウスと OLTP の両方：
- IORMを使用して、データベースとコンシューマ・グループ間でI/Oリソースの優先順位を付ける

**問題シナリオ1 - HDDの使用率は高いが、フラッシュはアイドル状態にある**

- データウェアハウスでの一般的な問題シナリオ：
  - スマート・スキャンが発行する大きい読取りはフラッシュ・キャッシュをバイパスするため、HDDのディスク使用率は高くても、フラッシュはアイドル状態にあります。
- 解決方法：
  - 大きい読取りをフラッシュにオフロードするため、KEEP オブジェクトに対してフラッシュを使用します。
  - 次のステップを実行します。
    1. I/O集中度レポート (@?/rdbms/admin/spawrio) を実行します。
    2. 2.KEEP オブジェクトの合計サイズがフラッシュ・キャッシュ・サイズを超えないようにします。
      - 許容される KEEP サイズをフラッシュ・キャッシュ・サイズの 80%に制限します。
      - KEEP に対する読取りの多い小規模表を対象にします。
    3. 候補となるすべての表を KEEP に指定します。



- ワークロードを繰り返して、読取り I/O がフラッシュにオフロードされることを確認します。

**ステップ1- 読取りI/O集中度レポートの読み方**

ASH Activity - Estimate of I/O wait times  
 -> # Samples: # of samples in ASH - approximation of DB time  
 -> % Activity: approximation of % Time based on ASH samples  
 -> ID values are based on aggregation type:  
 by Wait Class: Wait Class name  
 by Segments : Owner.Object(SubObject)  
 \* wildcard is used if object or subobject name is too long  
 by File : FileID-FileName

| Space            | IO                            |            | Obj.                    |            |           |
|------------------|-------------------------------|------------|-------------------------|------------|-----------|
| Aggregation      | Id                            |            | Type                    | Tablespace | # Samples |
| GB               | Intensity                     | % Activity |                         |            |           |
| ----- by Segment |                               |            |                         |            |           |
| PART EDW_D02     | 87,083                        | 67.8       | EDW_ATS.TECS_PHC(P2011) | 3.2        | TABLE     |
|                  | EDW_ATS.TECS_PHC(P2011)       |            | TABLE PART EDW_16K_D0   |            |           |
| 87,083           | 67.8                          | 1,284.9    | 3.2                     |            |           |
|                  | EDW_ATS.ENTITY_ADDR           |            | TABLE EDW_D01           |            |           |
| 34,103           | 83.6                          | 408.1      | 1.3                     |            |           |
|                  | EDW_ATS.SHIPMENT_SCORE(P2011) |            | TABLE PART EDW_D07      |            |           |
| 25,543           | 85.2                          | 299.8      | 0.9                     |            |           |
|                  | EDW_ATS.TECS_PHC(P2010)       |            | TABLE PART EDW_16K_D0   |            |           |
| 15,006           | 98.3                          | 152.6      | 0.6                     |            |           |
|                  | EDW_ATS.TECS_PHC(P2010)       |            | TABLE PART EDW_D08      |            |           |
| 15,006           | 98.3                          | 152.6      | 0.6                     |            |           |
|                  | EDW_ATS.RULE_FINDING(P2011)   |            | TABLE PART EDW_D07      |            |           |
| 13,079           | 137.6                         | 95.0       | 0.5                     |            |           |
|                  | EDW_ATS.XBKSHIPMENTSC*(P2011) |            | INDEX PART EDW_IO2      |            |           |
| 12,904           | 153.5                         | 84.1       | 0.5                     |            |           |

IO Intensity - by Tablespace  
 -> I/O Intensity calculated as IO Reqs per sec/GB allocated  
 -> tablespaces with >= 0.25 % of Captured IOs displayed  
 -> %IOPs - Running Total of % of Captured IOPs  
 %Cap - IOs as a percentage of Captured IOPs  
 %Tot - IOs as a percentage of Total sysstat IOPs  
 Sysstat IOs per Sec: 7,532.4  
 -> ordered by Total IOPs desc, IO Intensity desc

| Write         | IO Req  | Rd Req  | Wr Req | Space     | IO        | Read      | Write     |      |      |
|---------------|---------|---------|--------|-----------|-----------|-----------|-----------|------|------|
| Tablespace    | per Sec | Total   | Total  | GB        | Total     | Read MB   | Write MB  |      |      |
| %IOPs         | IO MB   | IO Blks | %Reads | MB/s      | Intensity | Intensity | Intensity | %Cap | %Tot |
| EDW_D07       | 370.8   | 276.7   | 94.1   | 13,818.1  | 0.0       | 0.0       | 0.0       | 11.4 |      |
| 12,942,243.8  | 828.3M  | 83.4    | 16.6   | 13.8      | 11.5      | 2.3       | 11.4      | 4.9  |      |
| EDW_D01       | 353.1   | 213.2   | 139.9  | 110,880.0 | 0.0       | 0.0       | 0.0       | 22.2 |      |
| 27,989,013.2  | 1791.3M | 86.5    | 13.5   | 28.6      | 24.4      | 4.1       | 10.8      | 4.7  |      |
| EDW_D06       | 217.2   | 197.0   | 20.2   | 26,444.0  | 0.0       | 0.0       | 0.0       | 28.8 |      |
| 15,848,043.0  | 1014.3M | 97.1    | 2.9    | 17.2      | 16.7      | 0.5       | 6.7       | 2.9  |      |
| EDW_16K_D0    | 172.9   | 167.3   | 5.6    | 30,693.6  | 0.0       | 0.0       | 0.0       | 34.1 |      |
| 140,063,687.1 | 8964.1M | 99.9    | 0.1    | 139.9     | 139.8     | 0.1       | 5.3       | 2.3  |      |
| EDW_D02       | 120.0   | 114.2   | 5.8    | 7,215.8   | 0.0       | 0.0       | 0.0       | 54.9 |      |
| 11,069,533.8  | 708.5M  | 96.9    | 3.1    | 13.4      | 13.1      | 0.3       | 3.7       | 1.6  |      |

## ステップ2- 合計KEEPサイズの評価方法

| Id                       | Type       | GB   | Intensity |
|--------------------------|------------|------|-----------|
| EDW_ATS.TECS_PHC (P2011) | TABLE PART | 67.8 | 1,284.9   |
| EDW_ATS.TECS_PHC (P2011) | TABLE PART | 67.8 | 1,284.9   |
| EDW_ATS.ENTITY_ADDR      | TABLE      | 83.6 | 408.1     |

Total KEEP size = 67.8 + 67.8 + 83.6 = 219.2 GB

## セルあたりのデフォルト・フラッシュ・キャッシュ・サイズ = 1.6TB

|                 | フルラック (X3) | ハーフラック (X3) | クォーターラック (X3) |
|-----------------|------------|-------------|---------------|
| フラッシュ・キャッシュ・サイズ | ~22.4TB    | ~11.2TB     | ~5.6TB        |

- KEEPを指定する以前のフラッシュ・キャッシュ使用率を分析します。
- 新しく指定したKEEPオブジェクトによって、フラッシュ・キャッシュを効果的に活用しているその他の重要ワークロードが妨げられないようにします。

## ステップ3- オブジェクトに対するKEEPの指定方法

- 次のSQLを実行します。
  - `ALTER TABLE TECS_PHC MODIFY PARTITION P2011 STORAGE (CELL_FLASH_CACHE KEEP);`
  - `ALTER TABLE ENTITY_ADDR STORAGE (CELL_FLASH_CACHE KEEP);`

## ステップ4- フラッシュ利用の検証方法

- 同じワークロードを繰り返して実行します。
- OSW IOSTATメトリックとセル・メトリックを調べて、次を確認します。
  - ディスク I/O 使用の減少 (HDD に対して発行される大きい読取りの減少)
  - フラッシュ I/O 使用の増加 (フラッシュ・キャッシュに対して発行される小さい読取りの増加)
- すべてのデータベースにわたる総合的ワークロードの実行が (低速でなく) 高速になっていることを確認します。
  - 合計 KEEP サイズが大きくなりすぎると、潜在的にフラッシュ・キャッシュのスラッシングが発生することに注意します。
    - 新しいKEEPキャッシュ行によって、古いKEEPキャッシュ行が除外されます。

フラッシュ・キャッシュの20%に含まれるデフォルト・キャッシュ行はそのままです。

- 次のcellcliコマンドを使用して、現在のKEEP使用状況を取得します。
  - list metriccurrent attributes name, metricvalue where name like 'FC\_BYKEEP\_USED'
- 現在のKEEPサイズが合計フラッシュ・キャッシュ・サイズの80%に近い場合、KEEPを削減します。

#### 問題シナリオ2 - HDDが書き込みIOPSバウンドである

- 次は、11.2.3.2.xより前のセルに適用される制約です。
- Exadataフラッシュ・キャッシュは十分な読取りIOPSを提供します。
- 書き込みについてはどうでしょうか。
  - フラッシュ・キャッシュは WriteThrough であるため、書き込みを受け入れることはありません。
  - HDD ディスクが書き込み IOPS の制約を受ける場合はどうすれば良いでしょうか。
    1. MTTRまたはチェックポイントがアグレッシブすぎないかをチェックします。
      - アグレッシブである場合、MTTR ターゲットを緩和します。
    2. バッファ・キャッシュのサイズが小さすぎないかをチェックします。
      - 小さすぎる場合は、バッファ・キャッシュ・サイズを大きくします。

#### ステップ1 - バッファ・プール統計を使用したデータベース書き込みの調査方法

- バッファ・プール統計の調査
  - 「AWR」 → 「Buffer Pool Statistics」へ移動します。
  - 物理書き込みにはチェックポイント書き込み、エージング書き込みなどが含まれます。

#### バッファ・プール統計

- 標準ブロック・サイズのプール D : デフォルト、K : KEEP、R : リサイクル
- その他のブロック・サイズのデフォルト・プール : 2k、4k、8k、16k、32k

| P | Number of Buffers | Pool Hit % | Buffer Gets | Physical Reads | Physical Writes | Free Buff Wait | Writ Comp Wait | Buffer Busy Waits |
|---|-------------------|------------|-------------|----------------|-----------------|----------------|----------------|-------------------|
| D | 4,748,295         | 91         | 202,981,617 | 18,237,002     | 10,511,448      | 0              | 0              | 4,173             |

[バッファ・プール統計に戻る](#)  
[画面上部に戻る](#)

#### チェックポイント・アクティビティ

- 合計物理書き込み : 10,513,291

| MTTR Writes | Log Size Writes | Log Ckpt Writes | Other Settings Writes | Autotune Ckpt Writes | Thread Ckpt Writes |
|-------------|-----------------|-----------------|-----------------------|----------------------|--------------------|
| 1,585,932   | 0               | 0               | 0                     | 897,667              | 0                  |

### ステップ2- MTTRアドバイザリを使用したデータベース書き込みの調査方法

- 「AWR」 → 「MTTR Advisory」へ移動します。
- 'Size for Est (s)'と'Est Total Writes'を比較します。
- fast\_start\_mttr\_target設定を増やして、書き込みを削減します。
  - トレードオフとして、リカバリ時間が長くなる点に注意します。
- [パフォーマンス・チューニング・ガイド](#)の[Recovery Tuning](#)を参照してください。

### MTTRアドバイザリ

- 終了スナップ時の行のみを表示
- 推定書き込み数とI/O数の単位は千

| Size for Est (s) | Dirty Limit | Est Cache Writes | Est Cache Write Fctr | Est Total Writes | Est Total Write Fctr | Est Total IOs | Est Total IO Fctr |
|------------------|-------------|------------------|----------------------|------------------|----------------------|---------------|-------------------|
| 90               | 150,365     | 98,675           | 2.83                 | 98,681           | 2.83                 | 153,985       | 1.71              |
| 450              | 751,829     | 48,173           | 1.38                 | 48,178           | 1.38                 | 103,482       | 1.15              |
| 900              | 1,503,658   | 34,873           | 1.00                 | 34,879           | 1.00                 | 90,182        | 1.00              |
| 1,211            | 2,023,256   | 29,433           | 0.84                 | 29,439           | 0.84                 | 84,742        | 0.94              |
| 1,485            | 2,480,007   | 27,763           | 0.80                 | 27,769           | 0.80                 | 83,072        | 0.92              |

### ステップ3- バッファ・キャッシュ・サイズの評価方法

- チェックポイント書き込みが合計物理書き込みに大きく影響しない場合、バッファ・キャッシュのサイズが小さすぎないかを確認します。このサイズが小さすぎると、エージング書き込みが過剰に発生する場合があります。
- バッファ・キャッシュが小さすぎることを示す兆候：
  - AWR Buffer Pool Advisory (v\$db\_cache\_advisoryに基づく)でサイズが増えた読取りが大幅に減少している場合は、おそらく、エージング書き込みも減少すると考えられますが、保証はありません。
  - "db file parallel write"をチェックして、長い待機時間がないかを確認します。
- 必要に応じて、バッファ・プール・サイズを増やします。
- [パフォーマンス・チューニング・ガイド](#)の[Configuring and Using the Buffer Cache](#)を参照してください。

### DW向けの上位SQLのチューニング

DWワークロードがI/Oバウンドであり、一般的な問題シナリオ1に当てはまらない場合、上位SQLの合計I/Oを削減できるでしょうか。

- AWR から上位の SQL を特定します。
  - 最適化されたI/O = フラッシュ・キャッシュを使用したI/O + ストレージ索引を使用したI/O
  - 最適化されていないI/O = HDDのI/O（構成されている場合は、フラッシュ・グリッドディスクのI/O）
- それぞれの上位 SQL に対して、SQL 監視レポートを生成します。
- [SQL チューニング・ガイド](#)に戻ります。
  - 索引が欠如していないかを確認します。
  - サイズの小さすぎるPGAがないかを確認します。PGAのサイズが小さすぎると、作業領域の代わりに一時セグメントが使用されます。

### DWとOLTPの両方でのIORMの使用

IORMを使用して、データベースとコンシューマ・グループ間でI/Oリソースの優先順位を付けます。

次のリソースを参照してください。

- [Managing I/O Resources](#) (マニュアル Oracle Exadata Storage Server Software User's Guide)
- [Use IORM metric scripts](#) (MOS)
- [IORM recommended patches](#) (MOS)

### 例 - セル・メトリック・サマリーの取得方法

[IORMメトリック・スクリプト](#) (MOS) には、セル・メトリック・サマリーが含まれています。

CELL METRICS SUMMARY

```
Cell Total Utilization: Small=45% Large=0%
Cell Total Flash Cache: IOPS=3630
Cell Total Disk Throughput:MBPS=43
Cell Total Small I/O's: IOPS=3392.1
Cell Total Large I/O's: IOPS=41.4
```

```
Cell Avg small read latency: 64.59 ms
Cell Avg small write latency: 71.19 ms
Cell Avg large read latency: 45.36 ms
```

Cell Avg large write latency: 53.07 ms

Exadataのパフォーマンスが不十分であるとの苦情が寄せられた場合は、

- セルが I/O バウンドであるかどうかをチェックします。
- コンピュート・ノードが CPU バウンドまたはメモリ・バウンドであるかどうかをチェックします。
- どちらにも当てはまらない場合、データベース・パフォーマンスのチューニングと計画に戻ります。

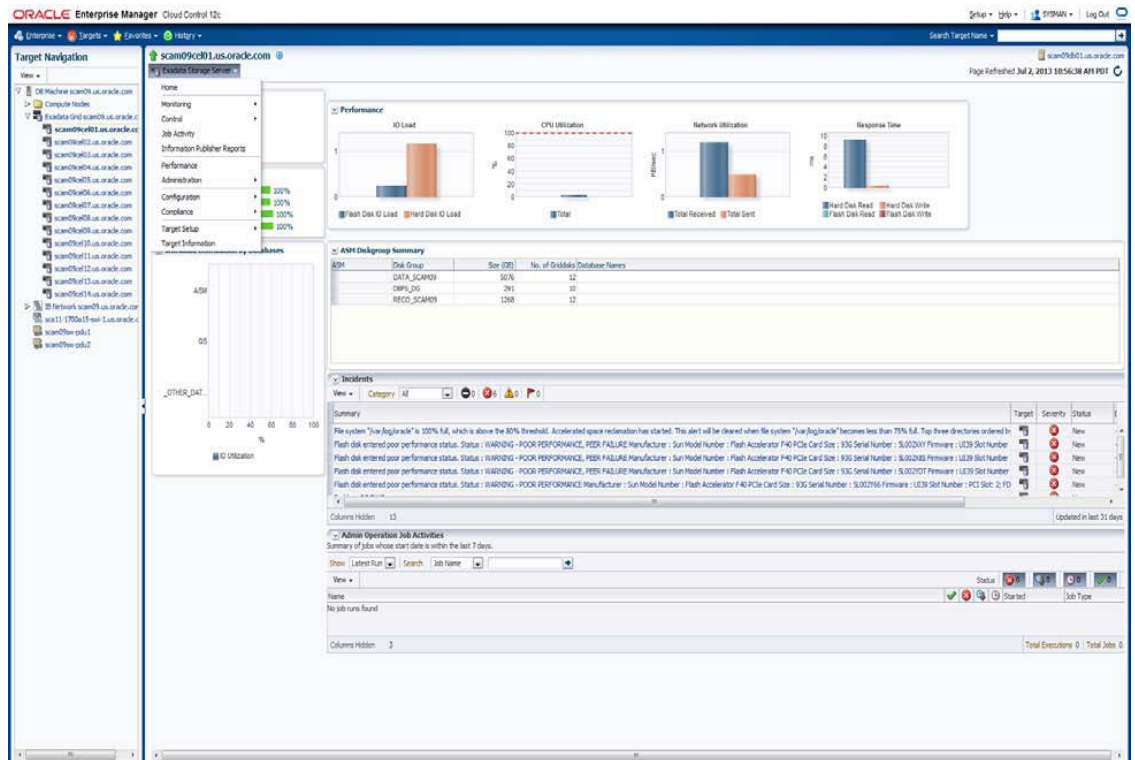
## Enterprise Managerを使用したI/O監視

### データベースのFree Buffer Waits

監視すべき非常に重要なメトリックの1つに、“free buffer waits”待機イベント時間があります。free buffer waitsは、データベース・プロセスが読取り操作を実行するための空きバッファを見つけられなかったことを示します。このイベントは、DBWRプロセスが十分な速度でストレージにブロックを書き込めなかったときに発生します。“free buffer waits”は、I/Oシステム の書き込み速度が上限に達しているか、または上限に近づいていることを示します。上位5つの待機イベントにこのイベントが含まれる場合、事前予防的な措置を講じて書き込み速度を下げるか、ストレージのI/O容量を増やす必要があります。

### セル・サーバーのターゲット・ページ

Enterprise Managerは、ハード・ディスクとフラッシュ・キャッシュの両方に対するディスク・パフォーマンスの集計情報を、セル・サーバー・レベルで提供しています。ここで提供される貴重な情報は、潜在的なボトルネックを特定し、ベースラインとなるI/O統計を確立するために役立ちます。I/Oページにアクセスするには、データベース・マシン・ターゲットのホームページからセル・サーバー・ターゲットを選択し、Exadata Storage Serverページで「Performance」を選択します。



### Exadataセルのパフォーマンス・ビュー

次に示すのは、Exadataセルのパフォーマンス・ビューの出力例です。ここでは、メモリとフラッシュ・キャッシュの使用に関する最新情報と履歴情報に加えて、データベース別のCPUおよびワークロード配分を確認できます。



## ファイル・システム情報

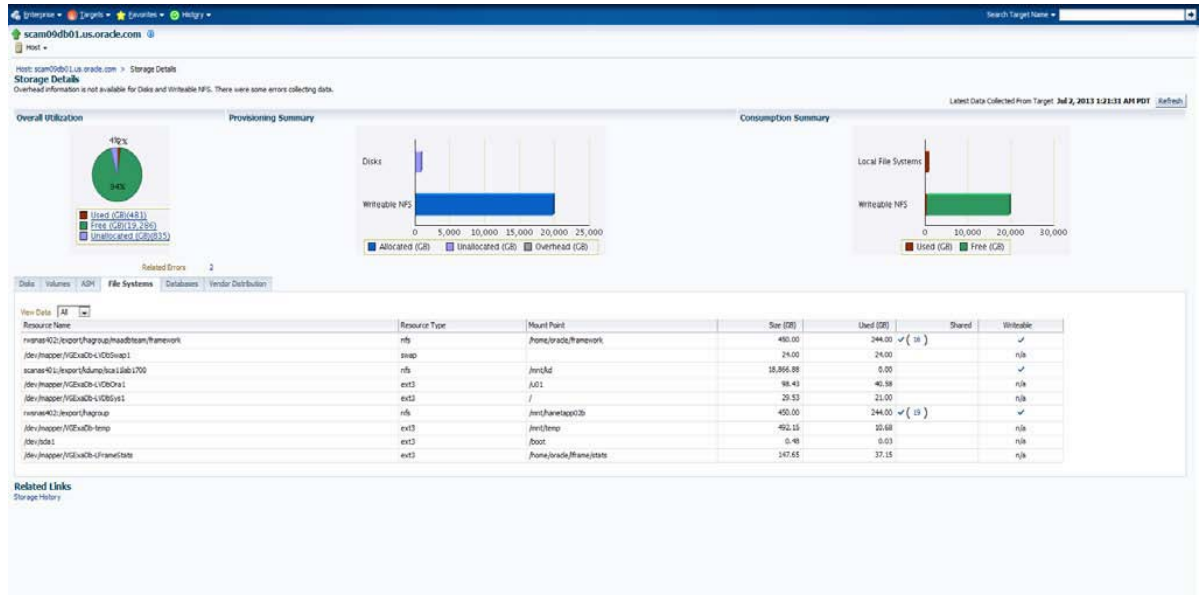
コンピュータ・ノード上にマウントされたファイル・システムの詳細情報を確認できます。コンピュータ・ノードのホームページから「Host」ドロップダウン・メニューを選択し、「Storage Details」を選択します。

The screenshot displays the Oracle Enterprise Manager Cloud Control 12c interface for a host named 'scam09db01.us.oracle.com'. The 'Host' dropdown menu is open, showing various management options. The 'Storage Details' option is highlighted. The main dashboard area shows several monitoring panels: 'CPU and Memory' with a line graph for CPU Utilization, 'Filesystem and Network' with a line graph for Filesystem Usage, and 'Job Activity' with a table of job runs. The 'Job Activity' table is currently empty, showing 'No job runs found'. The 'Incidents and Problems' panel at the bottom shows a summary of incidents.

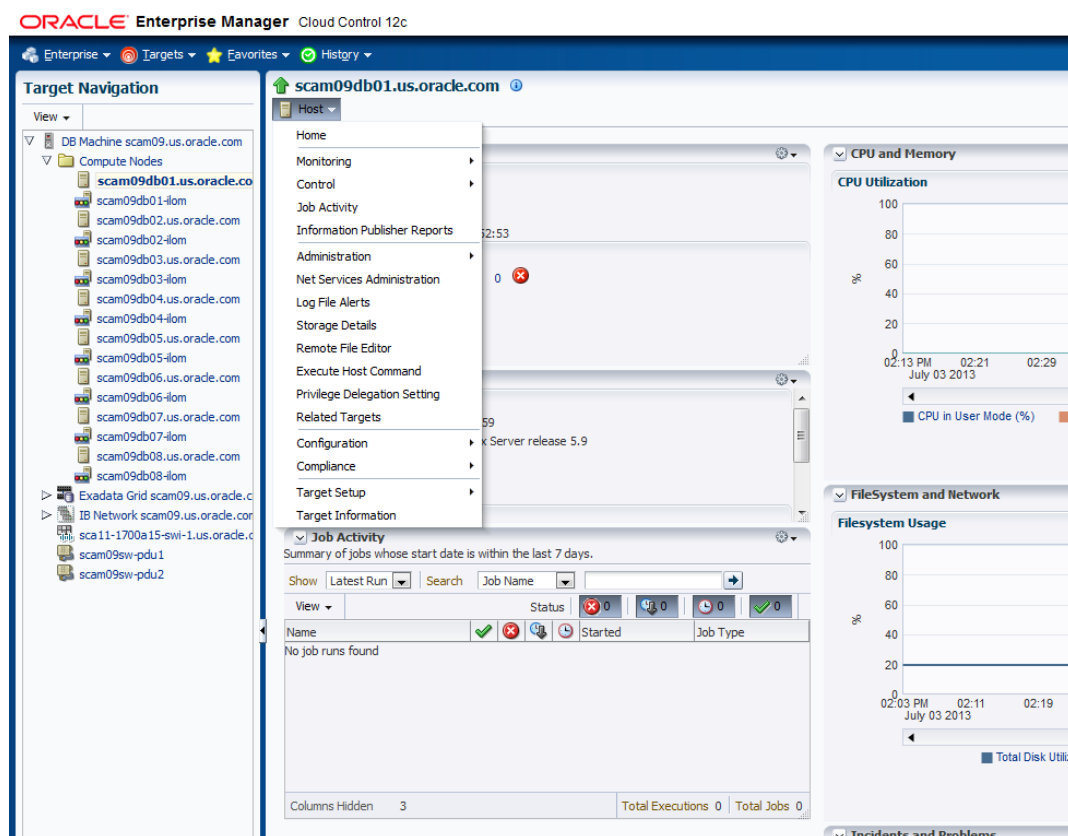
| Name              | Status | Started | Job Type |
|-------------------|--------|---------|----------|
| No job runs found |        |         |          |



表示される画面から、ファイル・システム、ASM、ローカル・ディスク、データベースに関するストレージ情報にアクセスできます。



特定のASMに関するより詳細な情報が必要な場合、ASMターゲット・ページ自体を表示できます。データベース・マシンのホームページからこの情報にアクセスするには、左側のナビゲーション・メニューから、表示するASMターゲットのあるコンピュート・ノードを選択します。コンピュート・ノード・ページが表示されたら、「Host」を選択し、Hostドロップダウン・メニューから「Related Targets」を選択します。

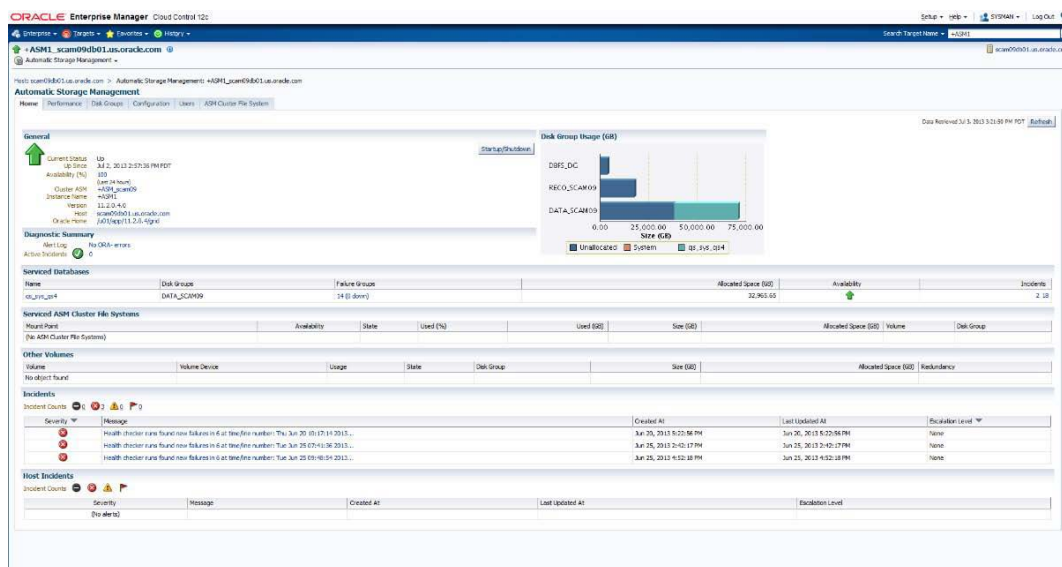


次の画面に、ASMターゲットのリストが表示されます。

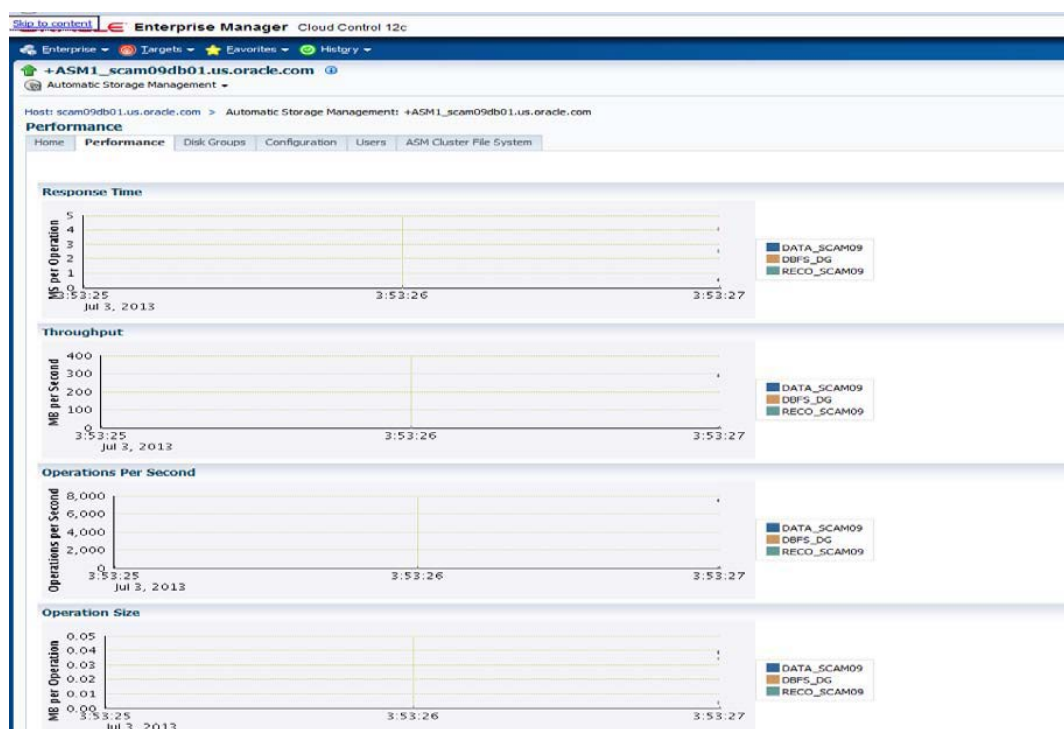
The screenshot shows the Oracle Enterprise Manager interface. On the left is the 'Target Navigation' tree, and on the right is the 'Related Targets' table. The table lists various targets for the host 'scam09db01.us.oracle.com'.

| Select                   | Name                              | Installed location                                         | Availability | Compliance Violations |
|--------------------------|-----------------------------------|------------------------------------------------------------|--------------|-----------------------|
| <input type="checkbox"/> | scam09db01.us.oracle.com:3872     | <a href="#">/u01/app/oracle/agent/haem/core/12.1.0.2.0</a> | ↑            | 0 0 0                 |
| <input type="checkbox"/> | +ASM1_scam09db01.us.oracle.com    | <a href="#">/u01/app/11.2.0.4/grid</a>                     | ↑            | 0 0 0                 |
| <input type="checkbox"/> | qs_sys_qs1                        | <a href="#">/u01/app/oracle/product/12.1.0.0/dbhome_1</a>  | ↑            | 0 0 0                 |
| <input type="checkbox"/> | LISTENER_scam09db01.us.oracle.com | <a href="#">/u01/app/11.2.0.4/grid</a>                     | ↑            | 0 0 0                 |
| <input type="checkbox"/> | DB Machine scam09.us.oracle.com   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | sca11-1700a15-swi-1.us.oracle.com | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db01-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db02-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db03-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db04-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db05-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db06-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db07-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09db08-fcom                   | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09sw-pdu1                     | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09sw-pdu2                     | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09cel01.us.oracle.com         | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09cel02.us.oracle.com         | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09cel03.us.oracle.com         | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09cel04.us.oracle.com         | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09cel05.us.oracle.com         | Not Applicable                                             | ↑            | 0 0 0                 |
| <input type="checkbox"/> | scam09cel06.us.oracle.com         | Not Applicable                                             | ↑            | 0 0 0                 |

ターゲットをクリックすると、ASMホームページが表示されます。



ここで、ディスク・グループやパフォーマンス、構成に関する詳細情報を参照および変更できます。次に示すのは、パフォーマンス・ページです。



## コンピュータ・ノードがCPUバウンドであるかどうかのチェック

- OSWのTOPデータからロード平均を確認します。
  - コンピュータ・ノード上で、`/opt/oracle.oswatcher/osw/archive/oswtop`にアクセスします。
- CPUコア数 = 最大ロードである点に注意します。
- CPUコア数を調べるには、[X3-8](#)と[X3-2](#)のデータシートを参照してください。

|                                                                  |                                                                                |
|------------------------------------------------------------------|--------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>コンピュータ・ノードのモデル</li> </ul> | <ul style="list-style-type: none"> <li>CPUコア数</li> </ul>                       |
| <ul style="list-style-type: none"> <li>X3-2</li> </ul>           | <ul style="list-style-type: none"> <li>2ソケット (ソケットあたり8コア) →16CPUコア</li> </ul>  |
| <ul style="list-style-type: none"> <li>X3-8</li> </ul>           | <ul style="list-style-type: none"> <li>8ソケット (ソケットあたり10コア) →80CPUコア</li> </ul> |

- コアあたりのロード平均 = コアあたりの実行可能プロセス数を計算します。
    - 質問：80というロード平均値は高すぎますか。
      - 回答：モデルによって異なります。
        - X2-2の場合、ロード/コア = 80/12 ≈ コアあたり 6.67 実行可能プロセス →高すぎる
        - X2-8の場合、ロード/コア = 80/64 ≈ コアあたり 1.25 実行可能プロセス →OK

3つのロード平均値は、それぞれ1分間、5分間、15分間の平均値です。

```
top - 08:29:25 up 2 days, 7:34, 0 users, load average: 283.02, 436.25, 422.83
Cpu(s): 1.4%us, 97.1%sy, 0.0%ni, 0.4%id, 1.0%wa, 0.0%hi, 0.1%si, 0.0%st
Mem: 98848968k total, 98407464k used, 441504k free, 916k buffers
 PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
 1401 root 12 -5 0 0 0 R 81.0 0.0 23:50.36 [kswapd0]
```

```
top - 08:29:36 up 2 days, 7:34, 0 users, load average: 259.12, 426.07, 419.66
Cpu(s): 2.5%us, 78.5%sy, 0.0%ni, 4.8%id, 14.1%wa, 0.0%hi, 0.1%si, 0.0%st
Mem: 98848968k total, 98401084k used, 447884k free, 792k buffers
 PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
 1401 root 10 -5 0 0 0 R 123.1 0.0 23:57.74 [kswapd0]
```

コンピュータ・ノード/コア = 283/12 ≈ コアあたり23実行可能プロセス

## コンピュータ・ノードがメモリ・バウンドであるかどうかのチェック

### VMSTATの調査

- OSW の VMSTAT データからメモリ・アクティビティを調べます。
  - コンピュート・ノード上で、`/opt/oracle.oswatcher/osw/archive/oswvmstat` にアクセスします。
  - 安定性のある優れたシステム・パフォーマンスを達成するには、スワッピングを解消する必要があります。

### スワッピングの例

```
procs -----memory----- --swap-- -----io----- --system-- -----cpu-----
r b swpd free buff cache si so bi bo in cs us sy id wa st
29 2 3837960 345492 219940 17162256 28 50166 28 50302 25320 27636 71 11 10 8 0
36 2 4083500 330308 179616 17082456 204 54676 204 54726 18572 28731 62 7 13 18 0
29 1 4084292 337328 117000 16605404 102 42 161 180 32837 28761 79 6 14 1 0
16 0 4084548 331532 116952 16554832 163 51 166 188 11681 17849 47 36 18 0 0
34 0 4085052 331460 116696 16491880 6 102 6 206 13546 13483 37 52 12 0 0
30 0 4086416 332476 116336 16398348 108 290 108 492 2200 4224 8 90 1 0 0
31 2 4087308 329096 116296 16337248 1 178 1 241 1664 3576 6 93 0 0 0
35 1 4087812 329276 116040 16269804 25 102 25 224 1587 2876 6 93 1 0 0
38 1 4088556 329152 115968 16195300 204 154 242 285 1309 3790 7 92 0 0 0
43 1 4089292 342020 114728 16110568 924 128 1187 230 1614 3739 8 92 0 0 0
29 0 4090956 328376 113476 15915564 451 332 525 573 3933 6770 4 94 2 0 0
37 0 4092040 328608 111428 15804568 112 154 117 426 1251 2881 1 98 0 0 0
34 0 4093000 331672 101392 15697876 435 154 479 255 1651 3203 1 98 0 0 0
```

### TOP kswapdの調査

- OSWのTOPデータから、kswapdとシステムCPUを調べます。
  - コンピュート・ノード上で、`/opt/oracle.oswatcher/osw/archive/oswtop` にアクセスします。
  - kswapd がコアを全面的に消費しているかどうか、システム CPU の使用量が高いかどうかをチェックします。
- 実際の顧客事例でのスワッピング例

```
top - 14:29:08 up 7 days, 17:27, 0 users, load average: 33.53, 26.46, 21.14
Tasks: 573 total, 35 running, 538 sleeping, 0 stopped, 0 zombie
Cpu(s): 6.3%us, 93.3%sy, 0.0%ni, 0.3%id, 0.1%wa, 0.0%hi, 0.0%si, 0.0%st
Mem: 74027752k total, 73698984k used, 328768k free, 116300k buffers
Swap: 16771852k total, 4087308k used, 12684544k free, 16334716k cached
 PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
 1049 root 20 -5 0 0 0 R 100.1 0.0 6:33.55 [kswapd0]
 4192 root 20 -5 0 0 0 R 98.7 0.0 47:48.07 [krdsd]
 1953 oracle 25 0 17.2g 1.2g 219m R 86.2 1.6 3:16.17 ora_p061_cbaprdx4
 1929 oracle 25 0 17.2g 1.2g 220m R 84.6 1.6 3:15.81 ora_p049_cbaprdx4
```

## 次のステップ

### コンピュータ・ノードがCPUバウンドまたはメモリ・バウンドである場合

- コンピュータ・ノードのCPUが不足しつつある場合
  - 負荷を減らします。
  - データベースのパラメータ設定が[ベスト・プラクティス MOS Note 757552.1](#)に従っているかをチェックします。
- コンピュータ・ノードのメモリが不足しつつある場合（スワッピング）
  - 負荷を減らします。
  - データベースのパラメータ設定が[ベスト・プラクティス MOS Note 757552.1](#)に従っているかをチェックします。
  - HugePagesを構成します（LinuxでHugePagesが適切に構成されていない場合、メモリの過剰使用につながります。これはExadata上で必須のベスト・プラクティスです）。

### セル・サーバーがI/Oバウンドではなく、コンピュータ・ノードがCPUバウンドでもメモリ・バウンドでもない場合

- セルがCPUバウンドか
  - セルのCPUが上限に達している場合、Exadataは自動的に負荷をコンピュータ・ノードに戻します。
- セルがメモリ・バウンドか
  - Exadataセルは、メモリを自動管理することでスワッピングを防止します。
- Oracleの[パフォーマンス・チューニング・ガイド](#)を再確認
  - AWR、ADDM、ASHなど（[データベース診断ツール](#)の項を参照）

## データベース・パラメータの設定方法

### - DWの容量計画

顧客は、[MOS Note 1274318.1](#)および[MOS Note 1347995.1](#)を参照可能

| パラメータ                  | X3-2   | X3-8   |
|------------------------|--------|--------|
| parallel_max_servers   | 240    | 1280   |
| parallel_min_servers   | 96     | 512    |
| parallel_degree_policy | Manual | Manual |
| parallel_degree_limit  | 16     | 24     |

|                         |     |      |
|-------------------------|-----|------|
| parallel_servers_target | 128 | 512  |
| sga_target              | 16G | 128G |
| pga_aggregate_target    | 16G | 256G |

### - OLTPの容量計画

- 顧客向けに、[MOS Note 1274318.1](#)および[MOS Note 1347995.1](#)も提供されています。

| パラメータ                | X3-2 | X3-8 |
|----------------------|------|------|
| parallel_max_servers | 240  | 1280 |
| parallel_min_servers | 0    | 0    |
| sga_target           | 24G  | 128G |
| Pga_aggregate_target | 16G  | 64G  |

## HugePagesの構成方法

- HugePagesは、ページ表サイズを縮小し、プロセスの起動時間を短縮します。
- /proc/meminfo内のPageTablesが物理メモリ・サイズの2%を超える場合は、オペレーティング・システムのHugePages/パラメータに共有メモリ・セグメントの合計値を設定します (Linuxのみ)。
- すべてのSGAに対応するだけの十分なHugePagesを割り当てます。
- HugePagesの設定方法については、[MOS Note 361323.1](#)を参照してください。
- 使用中の共有メモリ・サイズを特定する方法については、[MOS Note 401749.1](#)を参照してください。

## 統合環境の構成方法

- 複数のデータベースが含まれる場合は注意が必要です。
- ここまでの構成に関する推奨事項は、同じノード上のすべてのデータベース設定を要約したものです。
- 詳しくは、[Exadata Consolidation Best Practice Paper](#)を参照してください。
  - データベース・パラメータの設定方法
  - HugePages の構成方法
  - その他



## データベース診断ツール

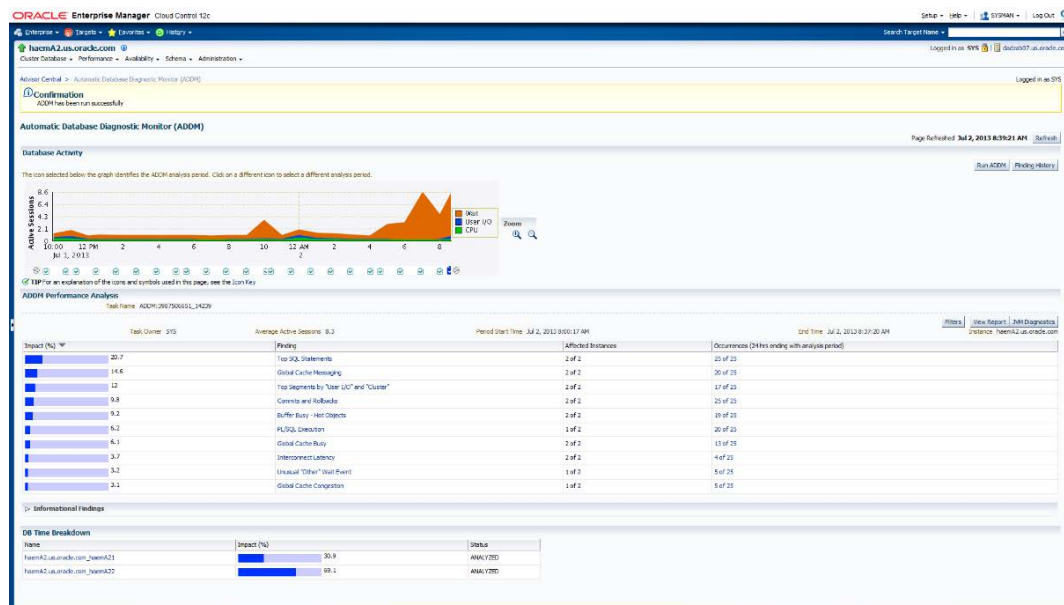
### ADDM

ADDMは、一定期間にわたるシステム・パフォーマンス・メトリックを分析し、その後のチューニングの重点領域を特定するために優れたツールです。場合によっては最適な状態で実行されており、追加のチューニングは必要ないことがあります。ADDMアドバイスは、多くの場合、上位のSQL文と待機イベントに焦点を合わせます。また、過度なSQL解析や過剰なログオン、空きバッファの待機など、パフォーマンスに悪影響を与える可能性のある一般的な問題をチェックします。

Enterprise Managerは、使いやすいADDM用グラフィカル・インタフェースを提供しています。これを利用するには、データベース・ターゲット・ページを開き、Performanceドロップダウン・リストから「Performance Home」を選択します。次に、画面右側の「Run ADDM Now」ボタンを選択します。



Enterprise ManagerからADDMを実行した出力の例：



## 上位SQL

経過時間、CPU使用率、バッファ・キャッシュの取得、その他のメトリックに関連する上位SQLは、AWRレポートで確認できます。

経過時間に関する上位SQLは、アプリケーション・パフォーマンスにもっとも直接的な関係があります。これは数回だけ実行される長時間の間合せの場合もあれば、何度も実行される短時間の間合せの場合もあります。

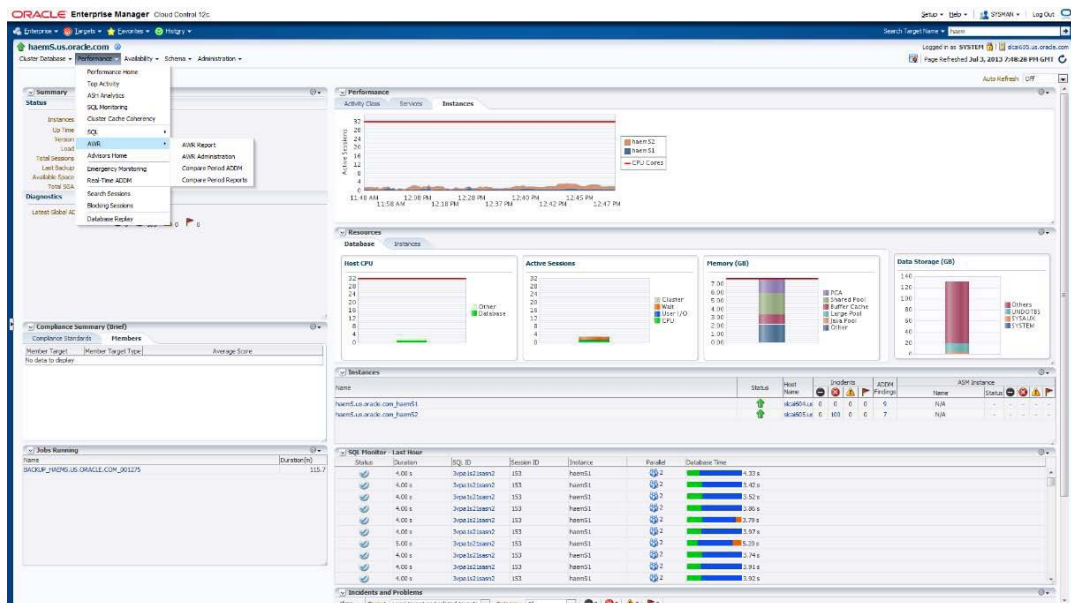
CPU時間とバッファ・キャッシュ取得に関する上位SQLは、不適切な間合せ計画を持つ間合せを見つけるために役立ちます。

I/Oに関する上位SQLは、過剰なI/Oを発行している間合せの特定に有効です。

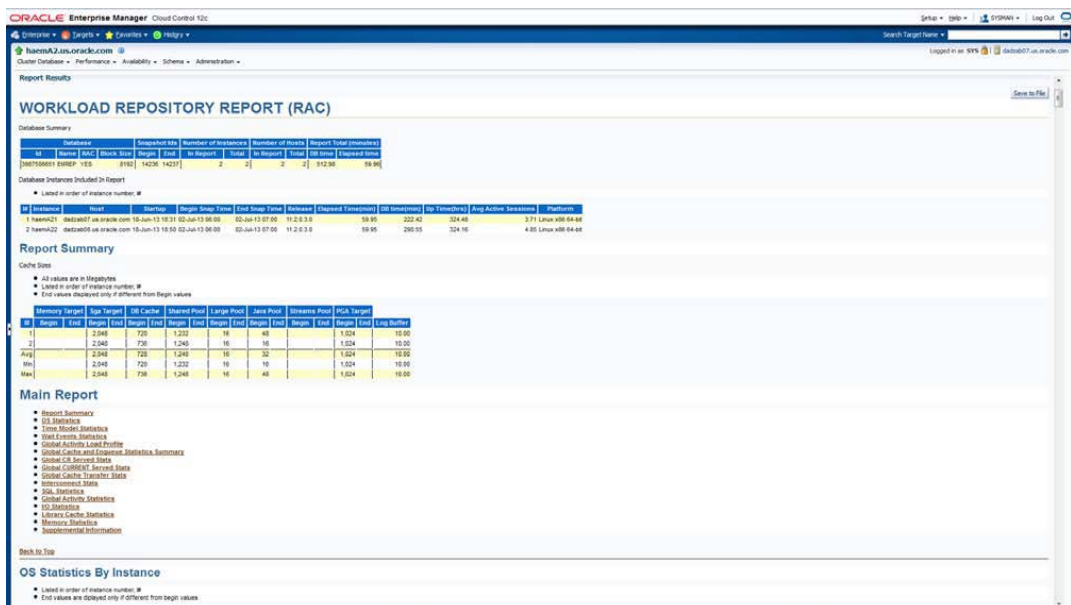
SQLチューニング・アドバイザを使用すると、特定の問合せに対して代替の問合せ計画を評価し、各計画によるパフォーマンスの見積もりを提供できます。また、索引の追加など、パフォーマンスを向上するその他の変更を推奨する場合もあります。

## AWRレポート

AWRレポートはEnterprise Managerから取得できます。データベース・ターゲットのホームページに移動し、「AWR」→「AWR Report」の順に選択します。



レポートの時間範囲を選択するオプションが表示されます。このオプションは、RACデータベースの使用時に、個々のインスタンスに対してAWRレポートを実行する場合にも使用できます。



### 上位の待機イベント

AWRレポートには上位の待機イベントも表示されます。かなりの割合のデータベース時間を消費しているイベントは、データベース・ワークロードのパフォーマンスに大きな影響を与えるため、イベントがアプリケーションのベースライン・パフォーマンスと一致しているか、または注意が必要かについて評価する必要があります。次に挙げる例では、"library cache: mutex X"待機が特に多く見受けられます。

上位5位のフォアグラウンド・イベント

| イベント                          | 待機        | 時間(秒)  | 平均待機(ミリ秒) | データベース時間の割合(%) | 待機クラス       |
|-------------------------------|-----------|--------|-----------|----------------|-------------|
| DB CPU                        |           | 10,442 |           | 81.89          |             |
| library cache: mutex X        | 4,560,452 | 2,149  | 0         | 16.85          | Concurrency |
| buffer busy waits             | 2,556,621 | 96     | 0         | 0.75           | Concurrency |
| SQL*Net more data from client | 3,759,147 | 89     | 0         | 0.70           | Network     |
| log file sync                 | 6,624     | 30     | 5         | 0.23           | Commit      |

このケースでは、通常、このイベントは発生しないため、アプリケーションを変更する必要がありました。しかし、イベントとワークロードによって、評価および潜在的ソリューションは異なります。重要なのは、さまざまなワークロードに何を期待するかという全体像を描いておき、異常や相違があった場合はこれを調査することです。

調査すべきもっとも重要なメトリックは、イベントの合計待機時間です。チューニング作業で、システム上の合計待機時間のごく一部しか消費しないイベントの待機を減らすことに集中しても、通常、効果はほとんど（またはまったく）ありません。

## 予防ツールによる混合ワークロード・システムでのワークロード分散処理

### I/Oリソース・マネージャ

<http://www.oracle.com/webfolder/technetwork/Exadata/MAA-BestP/IORM/IORM.pdf>

Exadataシステムでは、ディスクやディスク・グループ間でワークロードを分割、分散、管理する複雑さを伴うことなく、より優れたデータベース統合を実現しながら、帯域幅と領域の有効利用を推進できます。すべてのデータベースが同じディスクを共有するため、IO Resource Management (IORM) を使用して、統合データベース間で予測可能なレベルのパフォーマンスを維持できます。

IORMを使用すると、システム上の各データベースに対するI/O操作の割合を保証し、必要に応じて上限を設けることができます。ストレージ・セルあたりで利用可能な合計I/Oの割合として割当てが分散され、すべてのデータベース・プロセスはIORMによって管理されます。したがって、それぞれのシステムに対するIORMの有効性を評価する際、アプリケーション・フォアグラウンドのユーザー・プロセスの動作だけでなく、データベース書込みなどのバックグラウンドのデータベース・プロセスも考慮に入れる必要があります。

IORMについて、詳しくは[Master Note for Oracle Database Resource Manager \(Doc ID 1339769.1\)](#)を参照してください。

次に、I/O使用量の監視とリソース管理に使用できるツールを紹介します。

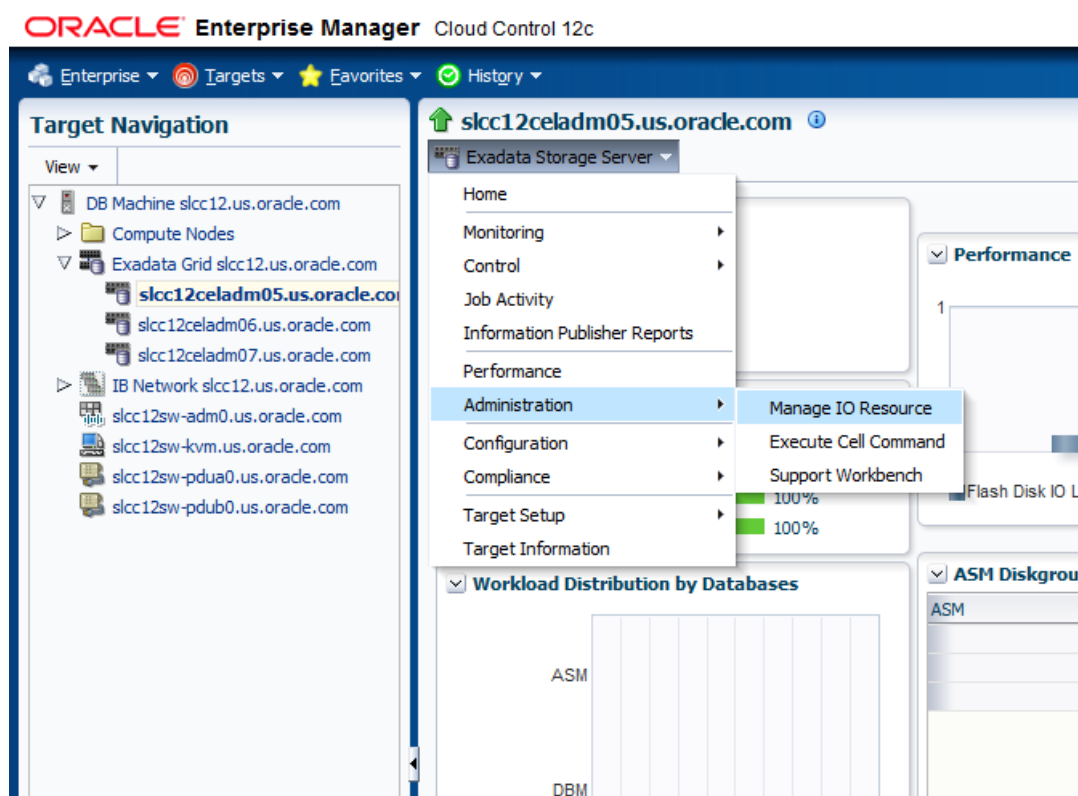
[metric\\_iorm.pl](#)

上記のMaster Noteから取得できるこのスクリプトは、IORMを使用したデータベース・スロットリングの即時効果を把握するために使用できます。データベースのI/O要件とIORM計画の結果を評価するために最適なツールです。次に、大部分は低いI/O使用率を示し、キューでの待機が発生する例を示します。セルのI/Oが上限に達するか、IORMによってデータベースが制限されると、qtimesが増加します。

```
Database: DBM
Utilization: Small=22% Large=5% ← database IO utilization
Flash Cache: IOPS=2068
Disk Throughput: MBPS=9
Small I/O's: IOPS=2154 Avg qtime=0.0ms
Large I/O's: IOPS=0.1 Avg qtime=1.2ms ← time spent waiting in the queue
```

## Enterprise Managerを使用したExadataのリソース管理

Exadataストレージ・セルのI/Oリソースは、Enterprise ManagerのManage IO Resourceページを使用して、管理および監視できます。このページは、次に示すストレージ・セルのメイン・ページから利用できます。



Manage IO Resourceページでは、次に示すようなリソースを設定できます。

slcc12celadm05.us.oracle.com

Exadata Storage Server

### I/O Resource Manager (IORM) Settings:

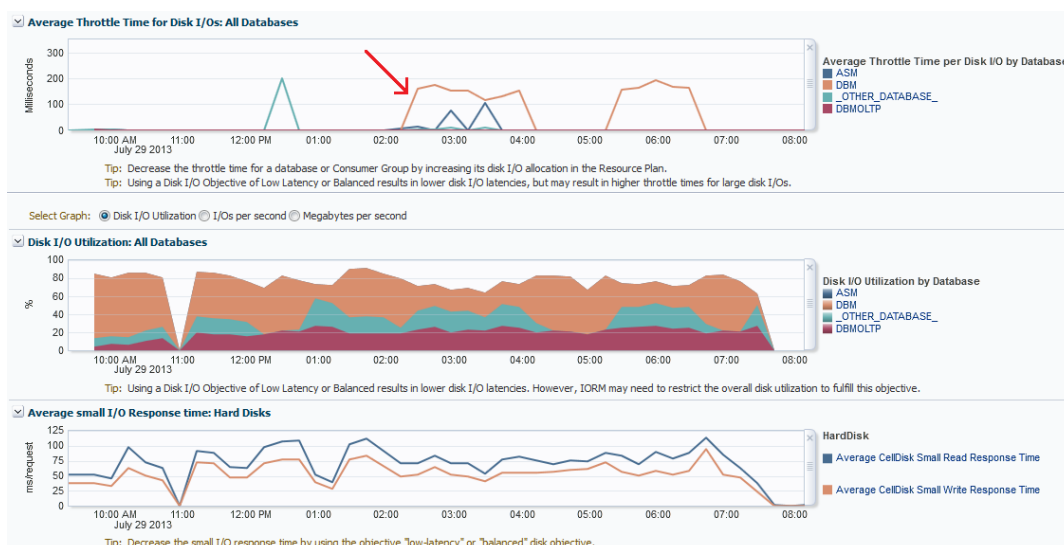
I/O Resource Manager controls how databases utilize the disks and flash cache, based on the settings specified here.

Status:  Active  Inactive      Disk I/O Objective: Auto

Inter-Database Plan:

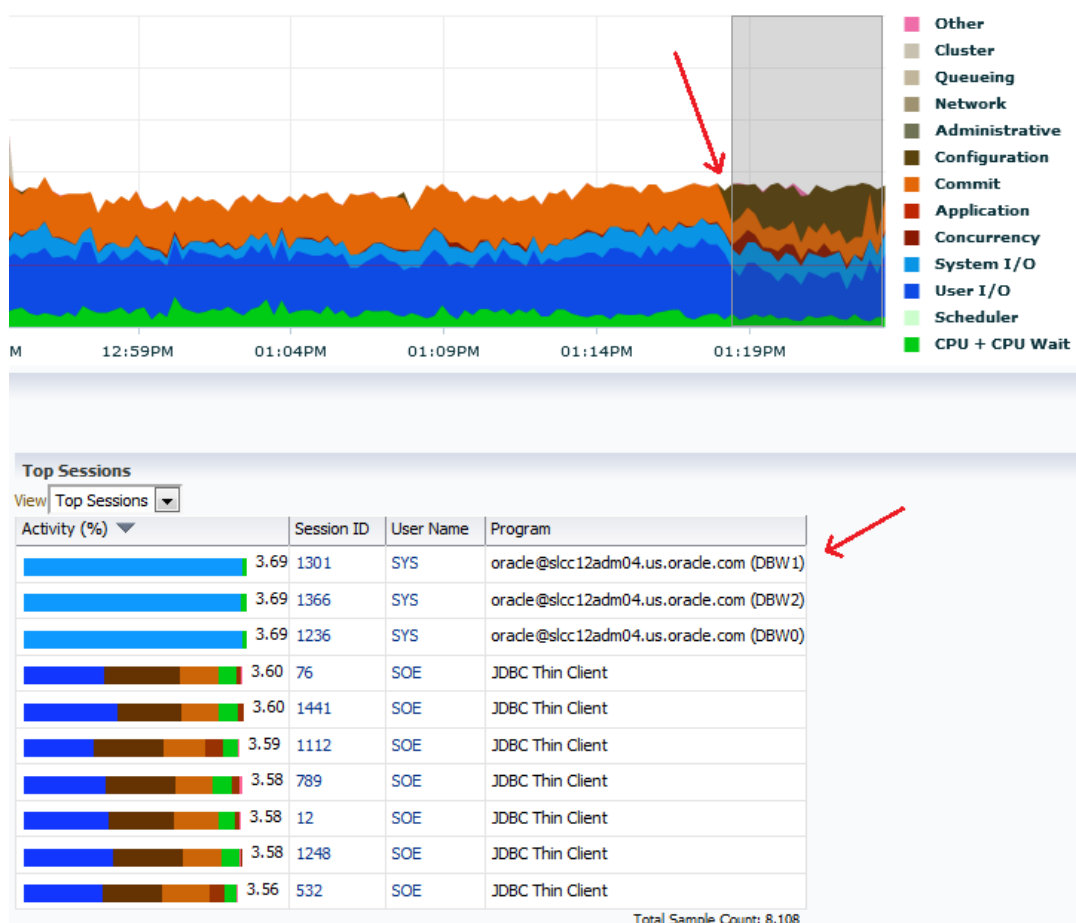
| Database Name | Disk I/O Utilization Limit(%) | Disk I/O Allocation(%) | Use Flash Cache                     | Use Flash Log                       |
|---------------|-------------------------------|------------------------|-------------------------------------|-------------------------------------|
| dbm           |                               | 70                     | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| dbmoltp       |                               | 30                     | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| other         |                               |                        | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |

また、セルのI/O使用とI/Oスロットリングを監視できます。これらのグラフは、ベースラインを特定し、長期的な動向を分析するのに最適です。次に示す例では、IORMによってデータベースが抑制されていることが分かります。



データベースの上位アクティビティ・ページ

Enterprise ManagerのDatabase Target Top Activityページには、データベースのアクティブ・セッション履歴データがグラフィカル表示されています。このページを使用すると、データベースのベースライン・シグネチャの変化を素早く見つけれられます。この例では、Configuration待機の変化が確認できます。これは、データベース・インスタンスがI/Oの制約を受けた結果です。バッファ・キャッシュからディスクへのフラッシュに時間が費やされるにしたがって、データベース・ライターが上位セッションになっていることに注意します。





## AWRレポート

AWRレポートは、ベースラインを定義し、最近のI/Oパフォーマンス・イベントを調査するために便利です。IO Statセクションから、I/Oリソースのスロットリングが発生した場合に、I/O操作の平均時間が増加することがわかります。

次に、ベースラインの平均I/O待機時間を示します。

### ファンクション別のIOStatサマリー

- 'Data'列に接尾辞M、G、T、Pが付く場合は1024の倍数になり、その他の列に接尾辞K、M、G、T、Pが付く場合は1000の倍数になります。
- 順序は（データ読み取り+書き込み）の降順です。

| ファンクション名           | 読み取りデータ | リクエスト数/秒 | データ/秒   | 書き込みデータ | リクエスト数/秒 | データ/秒   | 待機数     | 平均時間(ミリ秒) |
|--------------------|---------|----------|---------|---------|----------|---------|---------|-----------|
| Buffer Cache Reads | 23.2G   | 5055.94  | 39.5183 | 0M      | 0.00     | 0M      | 3039.7K | 0.70      |
| DBWR               | 1M      | 0.11     | .001665 | 21.3G   | 3259.05  | 36.4068 | 68      | 22.50     |
| LGWR               | 0M      | 0.00     | 0M      | 4.4G    | 1381.56  | 7.57219 | 438.2K  | 0.67      |
| Others             | 32M     | 3.48     | .053301 | 11M     | 1.06     | .018322 | 2392    | 3.72      |
| Direct Writes      | 0M      | 0.00     | 0M      | 0M      | 0.00     | 0M      | 0       |           |
| 合計:                | 23.2G   | 5059.53  | 39.5733 | 25.8G   | 4641.68  | 43.9973 | 3480.4K | 0.70      |

データベースのスロットリングに従って、平均待機時間が増加します。

| ファンクション名           | 読み取りデータ | リクエスト数/秒 | データ/秒   | 書き込みデータ | リクエスト数/秒 | データ/秒   | 待機数     | 平均時間(ミリ秒) |
|--------------------|---------|----------|---------|---------|----------|---------|---------|-----------|
| Buffer Cache Reads | 8.7G    | 1729.49  | 14.7892 | 0M      | 0.00     | 0M      | 1038.9K | 1.84      |
| DBWR               | 5M      | 1.25     | .008327 | 6.6G    | 949.37   | 11.2634 | 748     | 155.12    |
| Others             | 2.7G    | 8.71     | 4.55001 | 2.6G    | 5.43     | 4.49672 | 2825    | 35.06     |
| LGWR               | 0M      | 0.00     | 0M      | 1.5G    | 609.20   | 2.64307 | 246.4K  | 1.24      |
| Streams AQ         | 0M      | 0.00     | 0M      | 0M      | 0.00     | 0M      | 2       | 4.00      |
| 合計:                | 11.3G   | 1739.44  | 19.3475 | 10.8G   | 1564.00  | 18.4032 | 1288.9K | 1.88      |

データベースのスロットリングが発生する場合、上位5位の待機イベントも通常変わります。次にベースラインを示します。

### 上位5位のフォアグラウンド・イベント

| イベント                            | 待機        | 時間(秒) | 平均待機(ミリ秒) | データベース時間の割合(%) | 待機クラス       |
|---------------------------------|-----------|-------|-----------|----------------|-------------|
| log file sync                   | 902,294   | 4,677 | 5         | 36.18          | Commit      |
| cell single block physical read | 2,498,568 | 3,677 | 1         | 28.44          | User I/O    |
| DB CPU                          |           | 3,546 |           | 27.43          |             |
| gc current block 2-way          | 1,794,254 | 256   | 0         | 1.98           | Cluster     |
| library cache: mutex X          | 238,070   | 233   | 1         | 1.80           | Concurrency |

データベース・ライターが使用済みバッファをディスクに書き込む時間が長くなるにつれて、アプリケーションのバッファ待機も増加します。

| イベント                            | 待機        | 時間(秒) | 平均待機(ミリ秒) | データベース時間の割合(%) | 待機クラス         |
|---------------------------------|-----------|-------|-----------|----------------|---------------|
| free buffer waits               | 294,844   | 3,329 | 11        | 24.20          | Configuration |
| cell single block physical read | 1,037,919 | 2,458 | 2         | 17.87          | User I/O      |
| log file sync                   | 331,531   | 2,045 | 6         | 14.87          | Commit        |
| write complete waits            | 213       | 1,716 | 8056      | 12.47          | Configuration |
| DB CPU                          |           | 1,416 |           | 10.29          |               |

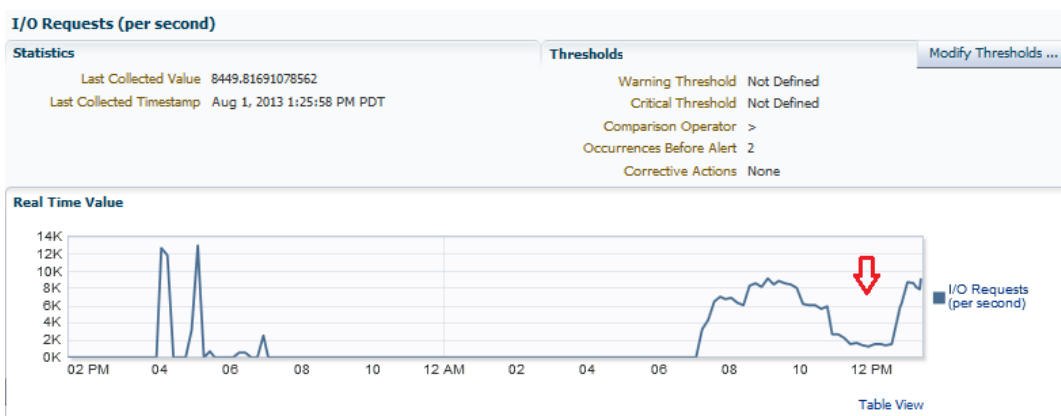
### データベース・メトリック

データベース・メトリックは、定義済みしきい値に達した場合のアラート通知やトレンド分析に使用できます。

次のグラフでは、Small I/O読取りの待機時間がスロットリングの発生時に増えることがわかります。



次に、同じ期間のI/Oリクエストを示します。スロットリングの発生時にリクエストが減少していることを確認します。



### 上限の設定

特に強い制限を使用していない場合、ストレージ・セルは、必要に応じて、未使用I/Oをその他のデータベースに対して自由に再分散できるため、IORM計画による効果がほとんど出ない可能性があります。このような捉えにくい待機時間は、通常短く一時的なイベントであり、一部のグラフには表示されません。上限が使用される場合、ストレージ・セルは追加の空きI/Oリソースをデータベースに割り当てないため、効果はずっと顕著になります。上限は、優先順位の低いデータベースがその他の重要性の高いデータベースを妨げることのないように、これらのデータベースを抑制および分離するために最適です。

## 参考資料

### MAAベスト・プラクティス

- [Exadata Consolidation Best Practice Paper](#)
- <http://www.oracle.com/webfolder/technetwork/Exadata/MAA-BestP/IORM/IORM.pdf>

## まとめ

要約すると、システムの動作状態およびパフォーマンスを総合的に監視するには、事前予防的なトッ  
プダウン・アプローチの維持が不可欠です。

EM GC、Exadataセル・アラート、ASR、Exachkなどのオラクル提供ツールを利用して、システムや  
データベース、アプリケーションに対する事前予防的な状態監視とアラート通知を有効化し、ソフ  
トウェアまたはハードウェアのエラーが発生していないかどうかをチェックすることで、重大なエ  
ラーを自動的かつ速やかに特定できるようになります。これにより、管理者は迅速な修正措置を講  
じることができるため、本番システムのエンドユーザーが影響を受けたり、停止を経験したりする  
ことはありません。

さらに、アプリケーションのサービス・レベルと応答時間が代表的なワークロード・テストに基づ  
いて定義されるように徹底します。これにより、これらのベースラインのデータ点に対してしきい  
値を適切に設定できるため、リソース使用率やユーザー応答時間が増え始めたり、予想外に変化し  
たりした場合に、素早くこれを特定できます。

一定期間にわたってしきい値を超過した場合、監視すべき重要なリソース・インジケータは、CPU、  
メモリ、ディスク領域、ディスクI/Oパフォーマンスです。瞬間的な使用量の急増は、通常、全体的  
なパフォーマンスに大きな影響を与えません。目安としては、即座にパフォーマンスを測定するよ  
りも、定期的な平均値を15分間調査すると良いでしょう。定期的な平均値が15分以上にわたってし  
きい値を超える場合、詳しい調査が必要です。Exadata環境で推奨されるしきい値は、ワークロード  
が瞬間的にピークに達した場合に備えて、システム上に十分な余裕（約20%）を持たせるように意  
図されています。主要なリソースが、想定される設定済みしきい値を一定期間にわたって超えた場  
合、同時に稼働するすべてのデータベースで上位イベント、SQL、I/Oを調査し、システムをチュー  
ニングする必要があります。



Exadataの動作状態およびリソース使用率の監視

2013年8月

著者：Mike Chafin

貢献者：Jia Shi、Shari Yamaguchi、Juan Loaiza

Oracle Corporation

World Headquarters

500 Oracle Parkway

Redwood Shores, CA 94065

U.S.A.

海外からのお問い合わせ窓口：



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2013, Oracle and/or its affiliates. All rights reserved. 本文書は情報提供のみを目的として提供されており、ここに記載される内容は予告なく変更されることがあります。本文書は一切間違いがないことを保証するものではなく、さらに、口述による明示または法律による黙示を問わず、特定の目的に対する商品性もしくは適合性についての黙示的な保証を含み、いかなる他の保証や条件も提供するものではありません。オラクル社は本文書に関するいかなる法的責任も明確に否認し、本文書によって直接的または間接的に確立される契約義務はないものとします。本文書はオラクル社の書面による許可を前もって得ることなく、いかなる目的のためにも、電子または印刷を含むいかなる形式や手段によっても再作成または送信することはできません。

Oracleは米国Oracle Corporationおよびその子会社、関連会社の登録商標です。その他の名称はそれぞれの会社の商標です。

0109

