～ みなさまの投稿をお待ちしております ～

Twitter

#OracleTechNight

Topic#1

RAC 環境構築時の設計ポイントと設定値の検討における考慮点について

ORACLE

- 以下の事項は、弊社の一般的な製品の方向性に関する概要を説明するものです。また、情報提供を唯一の目的とするものであり、いかなる契約にも組み込むことはできません。以下の事項は、マテリアルやコード、機能を提供することをコミットメント（確約）するものではないため、購買決定を行う際の判断材料になさらないで下さい。オラクル製品に関して記載されている機能の開発、リリースおよび時期については、弊社の裁量により決定されます。

RAC 環境構築時に設計ポイント

- 非 RAC → RAC 移行時に考慮すべき設計項目

  RAC 環境とすることでシングル環境よりも必要量が増加が見込まれるリソース

設計ポイント		考慮ポイント	設定値の検討方法
CPU	RAC 固有処理の CPU	キャッシュ・フュージョン等のノード間データ転送処理による CPU リソースの増加	• 転送ブロック数に依存のため、机上での見積りは困難（事前検証を実施して確認できると最も良い） • 事前検証等が困難な場合、LMS プロセスが CPU を 100 % 使用しても枯渇しない CPU リソースを確保（LMS プロセス数はコア数に依存） • ただし転送ブロック数が多いと LMS プロセスを増加させるケースもある点は留意する
メモリ	共有プール(SGA メモリ)	ノード間のブロック整合性を保つ処理で必要となるバッファ・キャッシュ管理用領域分のサイズ追加（gcs resources / gcs shadows）	• gcs resources は11gR2 で 8KB ブロック使用の場合、バッファ・キャッシュ 1 GB あたり 45MB 増加(参考値※) • gcs shadow は環境により、変動が大きく、実機環境を確認しサイズを調整
表領域	SYSAUX 表領域	スナップショットはインスタンス毎の情報となるため、増加するノード数分の情報が増加し SYSAUX に格納	• V$SYSAUX_OCCUPANTS から AWR のサイズを確認 • 上記に RAC 化後のノード数を係数として掛けたサイズを AWR サイズとして SYSAUX 表領域のサイズを追加

※ 2013年の DBA & Developers Day オラクルコンサルが語る共有プール管理の極意を参照ください
http://www.oracle.com/webfolder/technetwork/jp/ondemand/ddd2013/C-4.pdf

RAC 環境構築時に設計ポイント

- インスタンス停止時の業務のF/Oを踏まえた設計項目(1/2)

  生存インスタンスでリソース枯渇起因のエラーや性能劣化を抑制するポイント

設計ポイント		考慮ポイント	設定値の検討方法
CPU	CPU リソース	F/O 後にサーバあたりの使用 CPU リソースが増加を見込んだリソース確保や設定、アプリの制御を実施	• 最小稼動サーバ数でも CPU リソースが枯渇しないよう CPU リソースを確保する • 重要なアプリに十分なリソースを割り当てる必要がある場合、リソース・マネージャの活用を検討 (アプリ毎に使用するDBサービス分離を実施)
メモリ	SGA メモリ	F/O 後の下記等の増加による SGA メモリの必要サイズ増加を見込んだサイズを設定 • カーソル、セッション・プロセス情報(共有プール) • データブロック(バッファ・キャッシュ) • DataPump 同時実行数(Streams プール)	• 事前検証にて最小稼動サーバ数の際に SGA メモリの不足がないことを確認する • 現行の共有プールとバッファ・キャッシュのアドバイザの結果を使用して机上で見積り、後ほど実機環境を確認しサイズを調整
	PGA メモリ	F/O 後にインスタンスあたりプロセス数の増加による PGA 固定領域の増加と SQL 作業領域使用量の増加を見込んだサイズを設定	• 専用サーバ接続で F/O 後にサーバプロセス数が増加する場合、1 プロセスあたり 3MB 増加を見込む • 各インスタンスの作業領域サイズを確認し、最小稼動サーバ数となっても設定値超過や物理メモリ枯渇がないことを確認する
リソース・リミット	processes パラメータ	F/O 後にインスタンスあたりの起動プロセス数の増加を考慮した値を設定	• 専用サーバ接続の場合、最小稼動サーバ数時の最大接続数を考慮した値以上を設定する
	sessions パラメータ	F/O 後にインスタンスあたりの接続数の増加を考慮した値を設定	• 共有サーバ接続の場合、最小稼動サーバ数時の最大接続数を考慮した値以上を設定する • 専用サーバ接続の場合、本設定は基本的に考慮不要

RAC 環境構築時に設計ポイント

- インスタンス停止時の業務のF/Oを踏まえた設計項目(2/2)

  生存インスタンスでリソース枯渇起因のエラーや性能劣化を抑制するポイント

設計ポイント		考慮ポイント	設定値の検討方法
REDO ログ・ファイル	REDO ログ・ファイルサイズ	F/O 後に REDO ログ生成量が増加後も一定のログスイッチ間隔が確保できるサイズとする	• 最小稼動サーバ数の際の REDO ログ生成量でも 10 分に 1 回程度のログスイッチとなるサイズを設定
表領域	UNDO 表領域サイズ	F/O 後に UNDO 生成量が増加後も UNDO 表領域が枯渇しないサイズとする	• V$UNDOSTAT から UNDO 生成量と保持期間を確認し最小稼動サーバ数でも UNDO 表領域が枯渇しないサイズを設定
	一時表領域サイズ	F/O 後に一時表領域使用量が増加後も一時表領域が枯渇しないサイズとする	• V$TEMPSEG_USAGE から一時表領域使用量を確認し最小稼動サーバ数でも一時表領域が枯渇しないサイズを設定
リスナー	接続スループット	F/O 時のリスナー接続処理により CPU リソースが枯渇しない秒間の接続スループットを設定する	• 秒間接続数毎の CPU 使用率を確認し、接続処理にワイ宛可能な CPU リソースとなる秒間接続数をいずれかのパラメータに設定 ✓ CONNECTION_RATE_<リスナー名> ✓ RATE_LIMIT

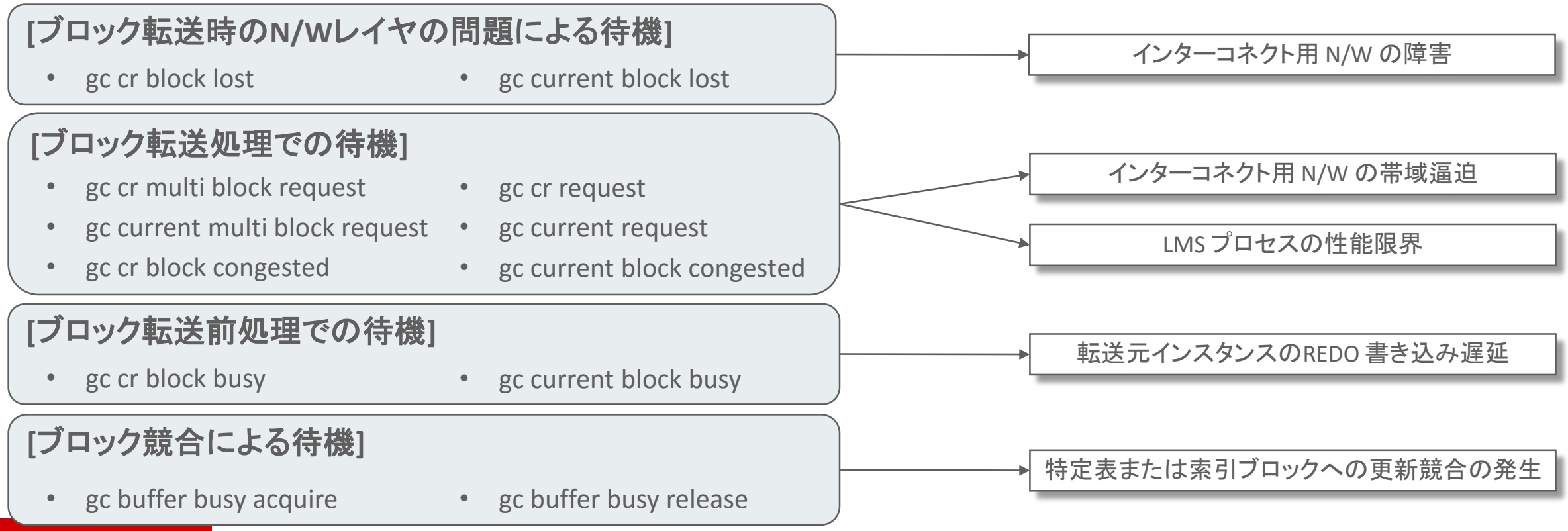
Topic#2

RAC 環境でキャッシュ・フュージョンの性能問題の切り分けとそのために使用する情報について

待機イベントからのボトルネック特定

- 上位待機イベント(Top 10 Foreground Events by Total Wait Time)を基に待機の発生原因を推定して調査

※ 代表的な待機イベントと原因のみ記載



待機イベントからのボトルネック特定

- 待機イベント確認後の原因調査で使用する情報の例

確認ポイント	確認情報	
送受信エラーパケット数	netstat	RX-ERR / RX-DRP / RX-ERR / RX-DRP
送受信パケット数	netstat	RX-OK / TX-OK
送受信データブロック数	AWR	Global Cache Load Profile → Global Cache blocks received の Per Second → Global Cache blocks served の Per Second
平均データブロック受信時間	AWR	Global Cache and Enqueue Services – Workload Characteristics → Avg global cache cr block receive time (ms) → Avg global cache current block receive time (ms)
転送ブロック消失数	AWR	Instance Activity Stats → gc blocks lost の Total
LMS プロセスの稼働状況	top	ora_lms*n*_<*SID*> プロセス %CPU
ノード間ブロック競合が多いセグメント	AWR	Segments by Global Cache Buffer Busy
クラスタ待機が多い SQL	AWR	SQL ordered by Cluster Wait Time
REDOログ書き込み性能 (他インスタンス)	AWR	Foreground Wait Events → log file sync の Avg wait (ms) Background Wait Events → log file parallel write の Avg wait (ms)
LGWR プロセスの稼働状況 (他インスタンス)	top	ora_lgwr_<*SID*> プロセス %CPU

ボトルネック特定と対応例

- 待機イベントを基に原因の切り分け、特定後に対処を進めます

問題箇所	調査の例	対応例
インターコネクト用 N/W の障害	(1) 転送ブロック消失数の確認 (2) 送受信エラーパケット数	• N/W レイヤの調査を進め、問題箇所を解消
インターコネクト用 N/W の帯域逼迫	(1) 平均データブロック受信時間 (2) 送受信データブロック数 (3) 送受信パケット数 (4) クラスタ待機が多い SQL	• ブロック転送処理の抑制するチューニング ✓ 不要フルスキャン実施 SQL のチューニング ✓ 特定処理のノード固定化など処理方式見直し • インターコネクト帯域の増強
LMS プロセスの性能限界	(1) 平均データブロック受信時間 (2) LMS プロセスの稼働状況 (3) クラスタ待機が多い SQL	• LMS プロセス数の増加対応 • ブロック転送処理の抑制するチューニング ✓ 不要フルスキャン実施 SQL のチューニング ✓ 特定表アクセス処理のノード固定化
転送元インスタンスのREDO 書き込み遅延	(1) REDOログ書き込み性能(他インスタンス) (2) LGWR プロセスの稼働状況(他インスタンス)	• REDO 書き込み処理の遅延に対するチューニング (非 RAC の場合と同じチューニング)
特定表または索引ブロックへの更新競合の発生	(1) ノード間ブロック競合が多いセグメント (2) クラスタ待機が多い SQL	• ブロック転送を抑制するチューニング ✓ 不要フルスキャン実施 SQL のチューニング ✓ 対象表のパーティション化とパーティション毎の処理ノード固定化 ✓ 特定処理のノード固定化など処理方式見直し • (表の場合) 1 ブロック内の格納行数の減少 • (索引の場合) 逆キー索引の活用(制限事項有)

Topic#3

非 RAC 環境から RAC 移行時に性能問題の発生を抑えるための事前対応例

RAC 移行により問題となりやすい処理

- 非 RAC で下記に該当する処理は RAC 移行により処理時間が増加する可能性があります
  - ブロック競合による待機が発生している処理

    ブロック転送処理に要する時間は、1処理あたりで見ると非常に短い時間の増加です。

    しかし、ブロック競合のように待ち行列となっている場合、RAC 化により大きな処理時間の増加につながるケースがあります。

  - OLTP 系処理でのフルスキャンや非効率な索引範囲検索の実施処理

    1ブロックあたりの転送時間は 1ミリ秒以下であることが多く、処理ブロック数が少ない場合ほとんど影響はありません。

    しかし、大量のブロックにアクセスする処理では上記の時間が累積することで、RAC 化によって大きな処理時間の増加につながるケースがあります。

    なおバッチ処理は実行回数が少ないため、OLTP 系の処理ほ顕在化しないことが多くありません。

RAC 移行により問題となりやすい処理の調査

- 非 RAC 環境で下記を確認し、事前の対応を検討します

対象処理	調査の例	事前対応例
ブロック競合による待機が発生している処理	AWR のセグメント統計の下記セクションを参照 『Segments by Buffer Busy Waits』	左記に記載のオブジェクトについて下記を検討 • パーティション活用によるブロック競合の低減 • 表の場合、PCTFREE 変更による格納効率の低下対応 • 索引の場合、逆キー索引の活用(レンジ・スキャン不可)
OLTP 系処理でのフルスキャンや非効率な索引範囲検索の実施処理	AWR のセグメント統計の下記セクションを参照 『Segments by Logical Reads』 『Segments by Physical Reads』	左記に記載のオブジェクト情報と SQL 情報を元に該当の可能性がある SQL を特定し、AWR SQL レポートを生成 序期レポートの SQL 実行計画を確認し、絞り込める条件が存在するにも関わらずフルスキャンしている箇所が存在する場合、下記を検討 • 適切な索引が作成されていない場合、索引の作成 • SQL 構造が複雑な場合、ヒント句による索引使用の強制
	AWR の SQL 統計の下記セクションを参照 『SQL ordered by Gets』 『SQL ordered by Reads』	

Topic#4

RAC障害発生時の情報取得について知りたい

ORACLE

1. 前提：事象の見極め方

- 正常な状態とはどのような状態であるかを把握する
  - crsctl stat res –t コマンドの実行結果を確認⇒Grid Infrastracture（GI）のステータスを確認できる

```
[grid@node1 ~]$ crsctl stat res -t
--------------------------------------------------------------------------------
Name           Target  State        Server                   State details
--------------------------------------------------------------------------------
Local Resources
--------------------------------------------------------------------------------
ora.asm
      1        ONLINE  ONLINE       node1                    Started,STABLE
      2        ONLINE  ONLINE       node2                    Started,STABLE
      3        OFFLINE OFFLINE                               STABLE

ora.rac1221.db
      1        ONLINE  INTERMEDIATE node1                    Stuck Archiver,HOME=/u01/app/oracle/product/12.2.0/dbhome_1, STABLE
      2        ONLINE  INTERMEDIATE node2                    Stuck Archiver,HOME=/u01/app/oracle/product/12.2.0/dbhome_1, STABLE
```

・どのリソースが起動しているのか
・ステータスが ONLINE となっているのか

state = INTERMIDIATE

2. まずはログを確保

事象の確認は基本的にログにて行います

- RAC を構成する要素とその要素から出力されるログ
  - RAC データベース、ASM 、Grid Infrastracture（GI）
  - アラートログファイル、トレースファイルを確保

  - GI のログであれば・・・

11gR2 RAC : Oracle Clusterware のログ格納場所について(KROWN:137282)(Doc ID 1747142.1)

12cR1 RAC : Oracle Clusterware のログ格納場所について(R12.1.0.2 ～)(KROWN:168363) (Doc ID 1770194.1)

  - **RACデータベース、ASMのログであれば・・・**

<diagnostic_dest>/diag/rdbms/<db name>/<SID>/trace 下の各種ファイル

※ diagnostic_dest ： 初期化パラメータ diagnostic_dest の値

【参考】RACにおける管理支援ユーティリティ

TFAコレクタ

Trace File
Analyzer Collector
Doc. ID 1609374.1

3.初動対応時のログの読み解き方

- RAC データベース、ASM
  - まずはアラートログを確認、「ORA- 」エラーの出力は無いか（シングル環境と同じアプローチ）？
  - 事象発生時間帯がハッキリしている場合は事象発生時間帯のログを注視

- Grid Infrastracture
  - まずはアラートログを確認
  - 事象発生時間帯がハッキリしている場合は事象発生時間帯のログを注視
  - GI の場合はエラーの内容が多岐にわたるので、別のログを参照する旨の表示がされている場合は、そのログを確認いただくことが重要

2017-06-20 04:27:35.962 [CRSD(9789)]CRS-1028: Oracle Cluster Registryバックアップ場所+DATA:18030523はノードnode1からアクセスできません。詳細は /u01/app/grid/diag/crs/node1/crs/trace/crsd.trcを参照し てください

4. ご紹介 : SRDC

- SRDCとは
  - SR調査のオートメーションを推進し、調査に必要なログ取得の方針をガイド

SRDC - RAC/Cluster ASM/ACFS に関する問題のための診断情報の採取方法 (Doc ID 2170355.1)

SRDC - Grid Infrastructure、ストレージ管理、RAC のデータ収集 (Doc ID 2142748.1)

SRDC - RAC データベース / インスタンスのパフォーマンス問題に関する（ハングではない）情報の収集方法 (Doc ID 2189112.1)

SRDC - RAC インスタンス排除 (Eviction) に関するデータ収集 (Doc ID 2190491.1)

※　一部のみ抜粋

5. 資料送付時の注意点

SR起票時に資料を送付いただく際の注意点を挙げます

- 事象発生時付近のログのみで構いません
- RAC 構成の場合は全ノードのログを送付してください
- 必ずしもログが出力されない場合もあることに注意

Oracle Database Technology Night
～集え！オラクルの力（チカラ）～

次回予告

会社帰りに参加できる夕方開催セミナー

Oracle Database Technology Night
～集え！オラクルの力（チカラ）～

RAC R12.2 インストールから運用までの勘所
～ RACとの付き合い方を考える ～

今宵のテーマは、「Oracle Real Application Clusters(Oracle RAC)」です。
基礎編をマスターしたら全体像と実装・運用が見えてきます。このRACというクラスタ&DBエンジン
をどのようにシステムに組み込んで使っていくか、その勘所を探っていきます。

お申し込み・詳細はこちら
2017年7月28日（金）18:45~20:30 （受付 18:30より）
http://www.oracle.com/goto/jpm170728

ORACLE

こんな時、かけこむ会社が増えています。

Oracle Digitalは、オラクル製品の導入をご検討いただく際の総合窓口。
電話とインターネットによるダイレクトなコニュニケーションで、どんなお問い合わせにもすばやく対応します。
もちろん、無償。どんなことでも、ご相談ください。

Oracle Digital

お問い合わせは電話またはWebフォーム

0120-155-096

受付時間:月~金9:00~12:00 / 13:00~18:00(祝日・年末年始休業日を除く)

http://www.oracle.com/jp/contact-us

ORACLE

Integrated Cloud

Applications & Platform Services


ORACLE®