

Alta disponibilidade em Banco de Dados Oracle usando Real Application Cluster (RAC)

Giovana P. Canali¹, Thiago H. C. Caires¹, Gustavo C. Bruschi¹, Luís A. da Silva¹

¹Curso de Tecnologia em Banco de Dados - Faculdade de Tecnologia de Bauru (FATEC)

Rua Manoel Bento da Cruz, nº 30 Quadra 3 - Centro - 17.015-171 - Bauru, SP - Brasil

giovanapignatti@gmail.com, thiago.cupri.aires@gmail.com,
gustavo.bruschi@fatec.sp.gov.br, luis.silva51@fatec.sp.gov.br

Abstract. *In order to be competitive, companies need high availability to store their data, which cannot be susceptible to faults and must be available whenever it is necessary. High availability tools perform these functions to meet the needs. It was concluded that if Oracle Real Application Cluster Tool (RAC) was well used, it will prevent a disaster in recovery, making the computing environment highly available in the company.*

Resumo. Para alavancar no mercado competitivo, empresas precisam de alta disponibilidade para armazenar seus dados, e estes não podem estar susceptíveis a falhas e devem estar disponíveis sempre que for necessário. Ferramentas de alta disponibilidade desempenham essas funções para que as necessidades sejam supridas. Concluiu-se que se bem aplicada a ferramenta Oracle *Real Application Cluster* (RAC) pode auxiliar na recuperação a falhas tornando o ambiente computacional das empresas altamente disponível.

1. Introdução

Um dos recursos que permite alta disponibilidade ao acesso aos dados é o Oracle *Real Application Cluster* (RAC). Segundo Shrivastava e Somasundaram (2009), a informação no mundo corporativo cresce muito e cada vez mais as empresas, para atingir o sucesso de seus negócios, precisam de alta disponibilidade no armazenamento de dados.

Para Canedo, Teixeira e Bruschi (2014), em um ambiente de alta disponibilidade, paralisar o banco de dados para realizar manutenções de desempenho, alterar configurações ou trocar discos rígidos do ambiente poderá trazer graves consequências financeiras para a empresa.

De acordo com Kinghost (2015), para os Banco de Dados que exigem alta disponibilidade, performance e escalabilidade, *clusterização* pode ser uma solução, neste caso dois ou mais servidores são capazes de trabalhar juntos sob um controle de um sistema operacional, com baixa perda de performance. Uma consulta complexa pode levar uma quantidade significativa de tempo para ser processado, porém, ao final dessa consulta os dados devem estar íntegros referente aos dados do começo da ação, sem influências de atualizações ou exclusões que podem ter ocorrido durante o período gasto com a consulta.

Kinghost (2015), afirma que o ambiente de *clusterização* Oracle RAC, além da redundância a nível de *software*, proporciona escalabilidade em termos de *hardware*,

onde é possível adicionar novos nós sem impactar na disponibilidade do serviço. Caso ocorra alguma dificuldade no acesso a algum nó, os outros servidores automaticamente assumem os processos executados, evitando erros de conexão, sendo uma ótima opção para Bancos de Dados com essa carência.

Neste artigo a alta disponibilidade foi analisada e através de um ambiente de teste foi verificado o funcionamento da ferramenta Oracle RAC mediante a uma falha forçada.

2. Cluster computacional

Para Pitanga (2003), um *cluster* é um sistema que compreende dois ou mais computadores, ligados em rede, na qual trabalham em conjunto para executar aplicações ou realizar outras tarefas, de tal forma para que os usuários que os utilizam tenham a impressão que somente um único sistema responde para eles, criando assim uma ilusão de um recurso. Segundo Reis et al. (2011), geralmente são utilizados para processar conteúdos críticos, garantir a disponibilidade de serviços e manter funcionando o sistema com tolerância às falhas e sem interrupções.

Já acordo com Veras (2013), a *clusterização* resolve o problema da falta de recursos em servidores físicos. Idealmente o *cluster* deve permitir que uma aplicação rode em mais de uma máquina física, incrementando a performance ao mesmo tempo em que aumenta a disponibilidade. Para Costa (2009), a alta disponibilidade e alto desempenho são algumas das funcionalidades que *clusters* podem conter, variam de acordo com cada tipo de *clusters* que podem fornecer outras características tais como gerenciabilidade, escalabilidade e acessibilidade. São essas variabilidades que tornam o *cluster* mais adequado para cada ambiente.

Segundo Reis et al. (2011), os clusters podem ser classificados em quatro tipos:

Balanceamento de carga (HS- Load Balancing): neste modelo de cluster todos os nós do cluster estão responsáveis em controlar o tráfego direcionado ao cluster, distribuindo as requisições entre os nós disponíveis. Em caso de falha em um nó, as requisições são redistribuídas entre os nós disponíveis no momento. São normalmente muito utilizados por servidores web que demandam muito acesso.

Combinação Alta Disponibilidade e Balanceamento de carga: combina os dois Clusters acima, aumentando a disponibilidade e escalabilidade dos serviços e recursos. Muito utilizado em servidores web e e-mail.

Processamento Distribuído ou Processamento Paralelo (HPC): este cluster é utilizado em tarefas em que se tem a necessidade de desempenho no processamento. Estes clusters são usados em computação científica. São normalmente associados ao projeto Beowulf.

Alta Disponibilidade (HA - High Availability): este tipo de cluster têm como objetivo prover uma disponibilidade de serviços e recursos de maneira mais ininterrupta possível, através do uso da redundância implícita ao sistema. Caso uma máquina do cluster vier a falhar, aplicações ou serviços estarão disponíveis em outro nó (failover). São muito utilizados para banco de dados de missões críticas, correio, servidores de arquivos e aplicações.

3. Alta disponibilidade

De acordo com Tanenbaum [2007 apud SOUZA e Campos 2008] disponibilidade é uma propriedade do sistema estar pronto para ser usado assim que for requisitado. Existem disponibilidades dos sistemas mesmo com períodos de inoperabilidade, desde que eles sejam curtos. Para HP (2013), este conceito possui relação direta com o *downtime*, tempo de inatividade não planejado do sistema.

Uma métrica para dimensionar a alta disponibilidade é apresentada na tabela 1 que relaciona percentuais de alta disponibilidade, mais conhecido como tabela de quantidade de "noves" que representa a porcentual do tempo em que seu ambiente está operacionalmente ativo. Quanto maior o tempo de paralisação planejado ou não planejado, menor a porcentual. Porcentagem de 99% é considerado como "dois noves", a taxa de 99,9% vale "três noves" e assim por diante.

<i>DOWNTIME</i> por SEMANA	<i>DOWNTIME</i> por ANO	Porcentagem de <i>DOWNTIME</i>	Porcentagem de <i>UPTIME</i>
3 horas e 22 minutos	7,3 dias	2%	98%
1 hora e 41 minutos	3,65 dias	1%	99%
20 minutos e 10 segundos	17horas e 30 minutos	0,2%	99,8%
10 minutos e 5 segundos	8 horas e 45 minutos	0,1%	99,8%
1 minuto	52,5 minutos	0,01%	99,99%
6 segundos	5,25 minutos	0,001%	99,999%

Tabela 1 – Percentuais de Alta Disponibilidade.

Fonte: Marcus [2003 apud SOUZA e CAMPOS 2008]

Segundo Souza e Campos (2008), percentual de 100% de disponibilidade ou Disponibilidade Contínua ocorre quando até mesmo as paradas para manutenção são mascaradas, é um conceito apenas teórico.

Souza e Campos (2008), ainda afirma que as soluções de alta disponibilidade são custosas cabendo ao gestor, usar as técnicas de gestão de risco, avaliar onde aplicar os recursos disponíveis de forma ótima para o negócio. Quando bem aplicado o investimento será compensado pelo *uptime* (maior tempo de funcionamento) que poderá acarretar em aumento de produtividade, menos ociosidade entre os usuários por falta do sistema e, no fim, economia financeira. Dentre os recursos para aumentar a disponibilidade estão as redundâncias nos equipamentos, para que na falha de um, essa falha seja mascarada por outro equipamento. O mascaramento pode ocorrer:

- ColdStandby*, os usuários são desconectados e perdem o trabalho corrente;
- WarmStandby*, usuários são desconectados e perdem o trabalho corrente e o sistema redundante fica em funcionamento para operar em caso de falha, usa-se replicação passiva, ou seja, técnica de replicação assíncrona.
- HotStandby*, o usuário não é desconectado, portanto a falha fica transparente ao usuário. Usa-se replicação ativa, ou seja, técnica de replicação síncrona, dados são gravados simultaneamente nos "sites backups" e na produção.

Para HP (2013), alcançar dois noves de disponibilidade pode ser desafiador, mas o terceiro nove é ainda mais difícil. A infraestrutura envolvida deve ser cuidadosamente planejada, ajustada e administrada para manter as aplicações e os serviços altamente disponíveis. Porém segundo Reis et al. (2011), falhas no setor da TI são praticamente inevitáveis, e um dos principais recursos é a utilização de *clusters* para proporcionar meios de tolerância a estas falhas.

4. Oracle RAC

Segundo Morales (2007), Oracle *Real Application Cluster* (RAC), é um recurso de Banco de Dados Oracle em *Cluster*, descrito como um conjunto de servidores denominados nós, independentes que se comunicam e operam como um único sistema. Sua arquitetura é totalmente compartilhada provendo alta disponibilidade, havendo uma falha de sistema, ele provê alta disponibilidade. Tornando com as redundâncias de: hardware, servidores, conexões de rede e discos, resistente a falhas e evitando assim um ponto único de falha (SPOF- *Single Point of Failure*).

Segundo a Oracle (2010), é utilizada por milhares de clientes em todos os setores para cada tipo de aplicação. Para Morales (2007), o Oracle RAC tornou-se uma solução que permite fornecer serviços de alta disponibilidade e balanceamento de carga com banco de dados Oracle, sendo formado por um *cluster*, onde dois ou mais computadores denominados nós trabalham em conjunto para atender as requisições dos usuários e/ou aplicações. De acordo com Reis et al. (2011), do lado do usuário (ou da aplicação) a conexão ao *cluster* é transparente, ou seja, não se percebe diferença em relação a uma conexão, a um banco de dados não *clusterizado*.

Para Morales (2007), para que o RAC funcione corretamente é necessária uma infraestrutura de *hardware* e *software*. Sendo os componentes mais críticos para o seu bom funcionamento: os discos, que devem ser compartilhados entre todos os nós do *cluster*, e a interconexão de rede de alta velocidade privada, conexão à rede pública. Segundo Souza e Campos (2008), tem como recomendado que todos os componentes tenham redundância, esta recomendação não se aplica apenas ao RAC, mas a qualquer ambiente de produção.

De acordo com Morales (2007), Oracle *Grid* é o *software* que gerencia o *cluster*, sendo o componente essencial na configuração e implementação do Oracle RAC. Deve estar instalado em todos os nós, é responsável por iniciar as instâncias e averiguar a sua disponibilidade e desempenho, ativar e desativar IPs virtuais, realizar a associação entre os nós, o serviço de banco de dados como se fosse uma única instância, registrar os serviços junto aos *listeners*, que é o serviço responsável por “escutar” as requisições.

Para a Oracle (2010), a escolha de armazenamento deve ser compatível, ou seja, homologado pela empresa do software que garante o funcionamento de acordo com o hardware do servidor. A chave para escolher um conjunto apropriado de armazenamento é escolher um sistema de armazenamento que irá fornecer *input/output* (I/O) de maneira escalável para a sua aplicação e, um sistema de I/O que irá escalar os servidores futuramente adicionados ao *pool*.

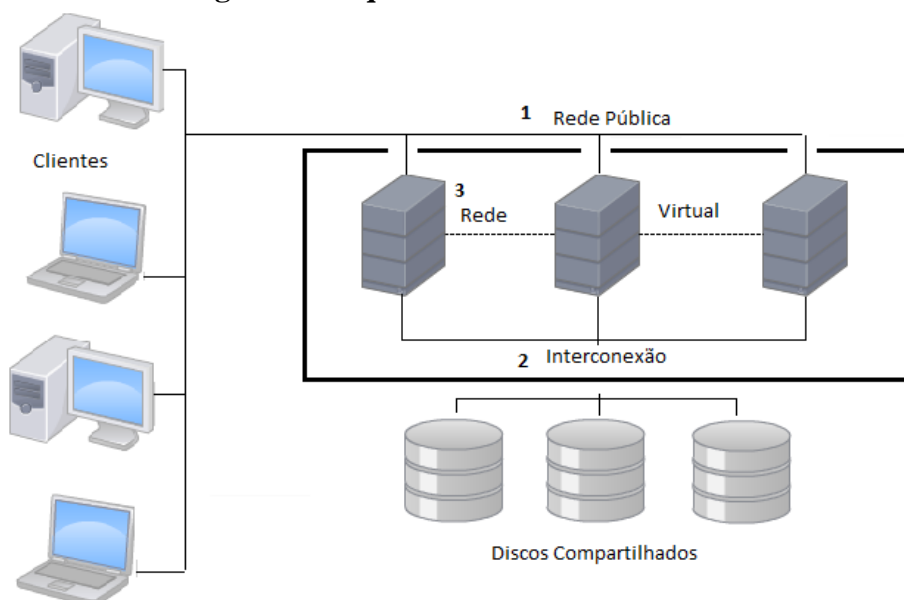
O acesso ao banco de dados em uma estrutura Oracle RAC é feito de forma paralela por todas as instâncias, sendo necessário que todos os nós estejam conectados fisicamente aos discos onde o banco está armazenado, retirando a necessidade de

replicar os dados nos vários servidores. Esse compartilhamento de discos tipicamente se dá através de uma *Storage Area Network* (SAN) usando protocolo *FibreChannel*, podendo também ser alcançado usando *Network Attached Storage* (NAS), ou até mesmo discos SCSI comuns cabeados para as várias máquinas, outras formas mais simples de armazenamento podem ser utilizadas em ambientes menos complexos.

Segundo Morales [2007 apud Oracle 2007] uma outra opção ao sistema de arquivos *clusteraware*, são os *rawdevices* ou o Automatic Storage Management (ASM), possibilitando as instâncias ter um acesso mais direto aos discos usando ASM, recomendado pela Oracle.

Segundo Canedo, Teixeira e Bruschi (2014), o Oracle ASM é um gerenciador de volumes e um sistema de arquivos que utiliza o recurso de multiplexação automatizada. Seu mecanismo é distribuir entre os discos disponíveis, os arquivos de dados, arquivos de *redo log* e arquivos de controle. Suporta instâncias simples e configurações com múltiplas instâncias através do Oracle *Real Application Cluster* (RAC).

Figura 1: Arquitetura do Oracle RAC



Fonte: IMATERS (2012)

De acordo com Morales (2007), analisando a figura 1 o *cluster* possui pelos menos três endereços IP em cada instância:

- a) IP Público, usado para o acesso administrativo ao servidor, representado na figura com o número 1;
- b) Rede de interconexão usados para comunicação entre os nós, número 2;
- c) IP Virtual (VIP – *Virtual IP Address*) usado para conexão dos clientes, que conhecem os VIPs de todos os servidores do *cluster* através de um arquivo de configuração, número 3;

Morales (2007), ainda afirma que é necessário que o VIP esteja na mesma sub-rede do IP público, para caso um nó do *cluster* falhe, o seu endereço virtual é reativado em outro nó, assim redirecionando as novas conexões para as instâncias que estão funcionando. O responsável por esse recurso de ativação/desativação não é o sistema

operacional, mas o *Grid*. Para a Oracle (2010), um *pool* de servidores também requer uma rede privada vulgarmente conhecida como “InterConnect”.

Segundo Morales (2007), a interconexão é uma rede privada, adicional a rede pública (LAN) utilizada pelos nós para receber e trocar requisições. Normalmente a interconexão é implementada usando interfaces de rede *Ethernet* 1Gbps ligadas a um *switch* ou VLAN exclusivos para a rede privada. A utilização de uma interface gigabit deverá ser suficiente para as necessidades de desempenho do nó. Mas é recomendado utilizar uma segunda interface para fornecer redundância a rede privada.

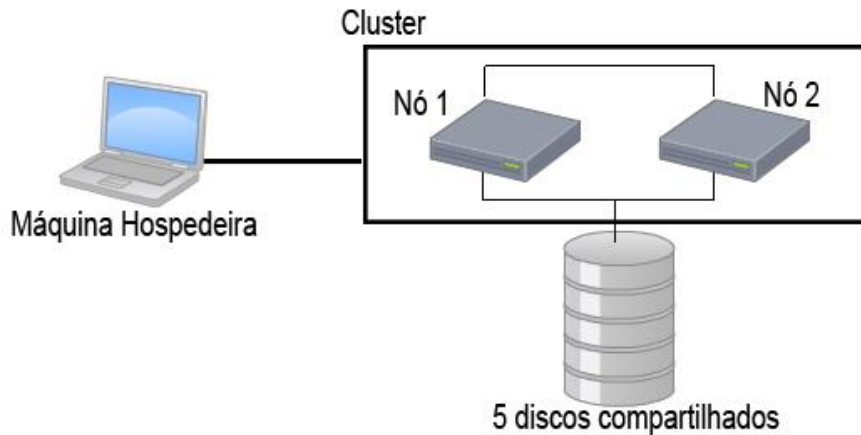
Para Morales (2007), o *Grid*, é responsável por monitorar as instâncias, sendo que cada instância monitora todas as outras através dos sinais de *heartbeat*. O *heartbeat* é uma mensagem enviada periodicamente por cada instância a todas as outras, funcionando como uma batida de coração, checando se as outras instâncias estão funcionando. Quando uma instância percebe que alguma das outras parou de enviar o *heartbeat* (excedido um tempo limite), ela tenta reorganizar o *cluster*, eliminando a instância do nó do *cluster* que não está respondendo e recuperando suas transações. A recuperação de falhas de uma instância é feita através de reaplicação do *log* de transações, *rollback* das transações não terminadas, sendo que o *cluster* continuará a responder enquanto houver ao menos um nó sobrevivente, e qualquer instância viva pode fazer a recuperação das transações da instância que falhou.

5. Materiais e Método

Para o teste prático foi utilizado o *Virtualbox*, *software* de virtualização usado para emular um ambiente, em que as duas máquinas virtuais rodando o Sistema Operacional Windows 2008 *Release 2*, enquanto compartilha os recursos computacionais da máquina hospedeira, um notebook com um processador modelo i7, seis gigabytes de memória *ram* e um *terabyte* de HD com o sistema operacional Windows 8.1.

Cada nó do *cluster* foi configurado com dois *gigabytes* de memória *ram* cada, deixando dois *gigabytes* restantes para a máquina hospedeira, essa arquitetura do ambiente pode ser observada na figura 2.

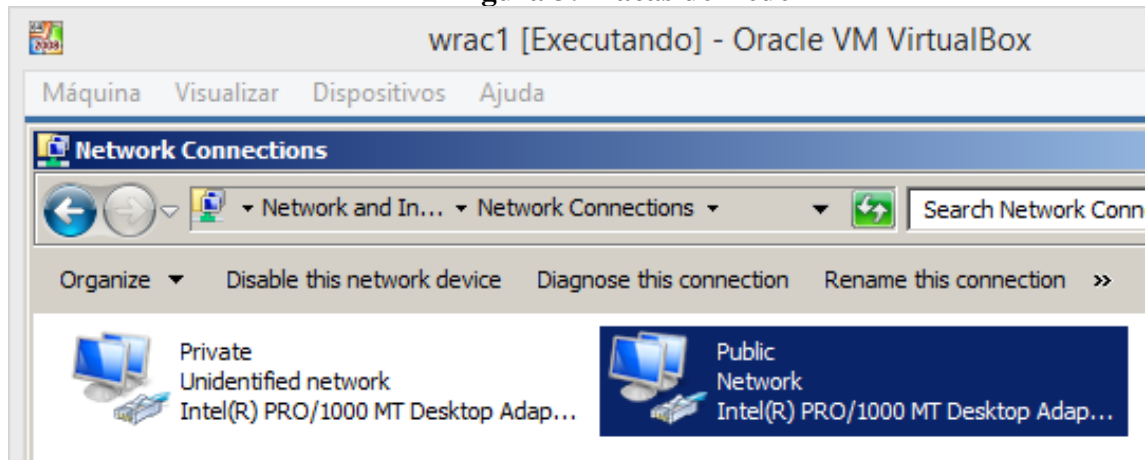
Figura 2: Arquitetura do teste prático



Fonte: Giovana Canali e Thiago Henrique Caires

Duas placas de rede foram adicionadas em cada nó, todas configuradas em modo *Bridge*. Uma das placas de rede foi utilizada para a interface pública enquanto a outra foi utilizada para a interface privada, analise a figura 3, assim separando a comunicação privada entre os nós e a comunicação pública em qualquer computador da mesma rede pode ter acesso.

Figura 3: Placas de Rede

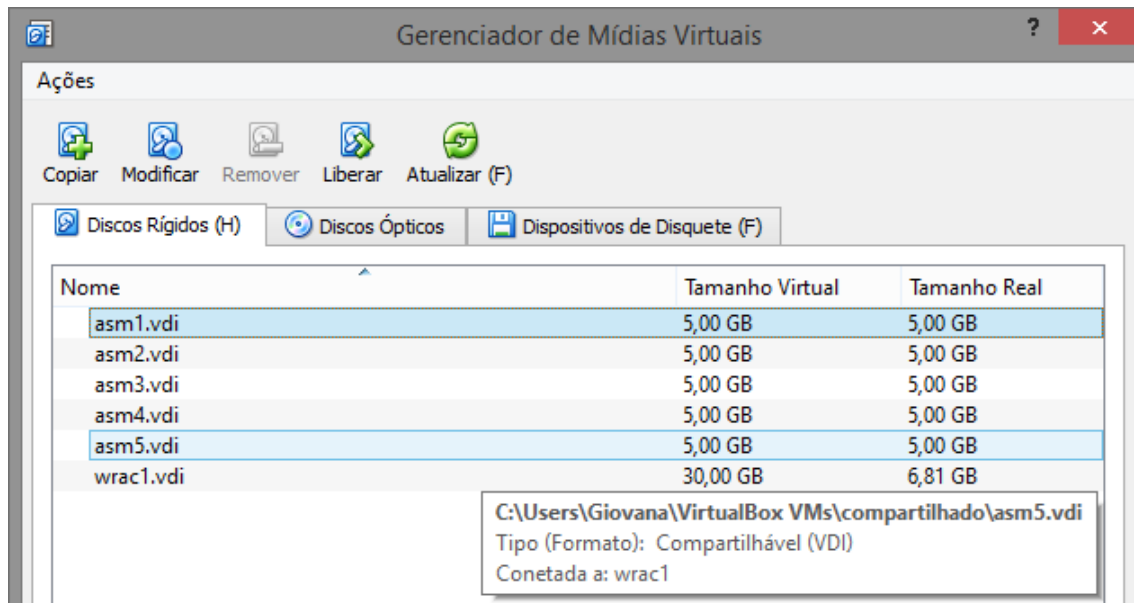


Fonte: Giovana Canali e Thiago Henrique Caires

O armazenamento compartilhado utilizado pelo Oracle RAC, como pode ser analisado na figura 4, foi criado utilizando o gerenciador de mídias virtuais, uma ferramenta incluída no *Virtualbox* com o propósito de automatizar e facilitar a criação, gerenciamento e compartilhamento dos discos nas máquinas virtuais.

Um total de cinco discos compartilhados de cinco gigabytes de armazenamento cada, utilizado como banco de dados para acesso simultâneo de cada nó do *cluster*.

Figura 4: Discos Compartilhados



Fonte: Giovana Canali e Thiago Henrique Caires

6. Resultados Obtidos

Após os dois nós do *cluster* estarem ligados e a instância do banco de dados estar ativa, com o usuário *SYSTEM* *logado*, foi executado o comando abaixo, que verifica qual nó do *cluster* está respondendo as requisições naquele momento, como pode ser observado na figura 5.

```
SELECT host_name FROM gv$instance WHERE instance_number=userenv('instance');
```

Caso a resposta seja *wrac1*, o que está respondendo as requisições naquele momento é o nó um, caso seja o *wrac2* é o nó dois.

Figura 5: Teste de qual nó responde a requisição

```
Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 - 64bit Production
With the Partitioning, Real Application Clusters, Automatic Storage Management,
OLAP,
Data Mining and Real Application Testing options

SQL> SELECT host_name FROM gv$instance WHERE instance_number=userenv('instance')
;

HOST_NAME
-----
WRAC1

SQL>
```

Fonte: Giovana Canali e Thiago Henrique Caires

O próximo passo, demonstrado na figura 6, foi conectar no *cluster* utilizando uma aplicação do computador hospedeiro e acessar através do usuário *HR*.

```
SQLPLUS HR/FATEC@RAC
```


Figura 6: Conexão através do usuário HR

```
C:\Users\Giovana>sqlplus hr/fatec@rac
SQL*Plus: Release 11.2.0.1.0 Production on Dom Mai 24 00:39:13 2015
Copyright (c) 1982, 2010, Oracle. All rights reserved.

Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 - 64bit Production
With the Partitioning, Real Application Clusters, Automatic Storage Management,
OLAP,
Data Mining and Real Application Testing options

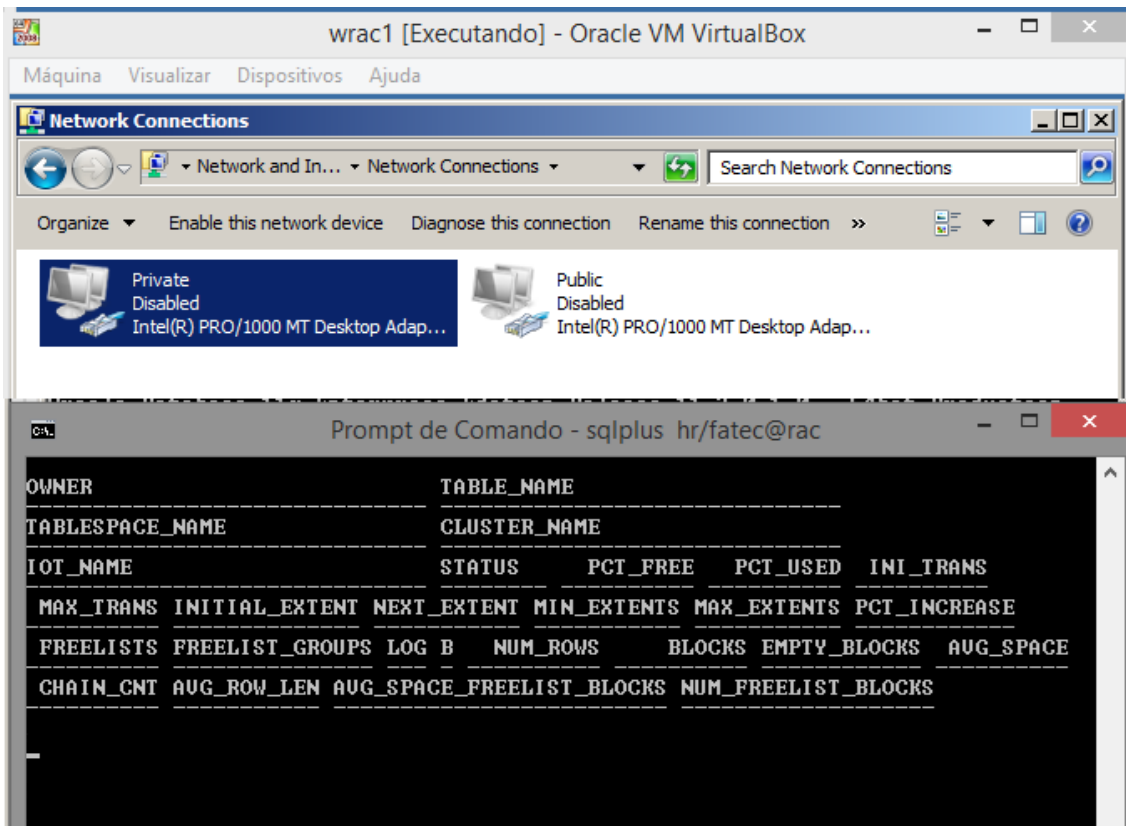
SQL>
```

Fonte: Giovana Canali e Thiago Henrique Caires

Foi realizado uma consulta no banco de dados, utilizando o comando descrito abaixo, onde foi retornado os dados sendo possível fazer o teste de alta disponibilidade mediante a uma falha. Pode se observar na figura 7, que essa falha foi forçada, desabilitando a interface de rede do nó que estava retornando à requisição.

*SELECT * FROM all_tables;*

Figura 7: Failover



Fonte: Giovana Canali e Thiago Henrique Caires

Obteve como resultado que ao derrubar o nó que estava respondendo a requisição, no recebimento dos dados houve um travamento de cinco segundos antes que requisição continuasse normalmente.

Evidenciando positivamente o funcionamento do Oracle RAC e o Processo de *Failover*, pois, como observado na figura 8, o outro nó do *cluster* assumiu a requisição deixada pelo nó que ficou inoperante, entregando assim a requisição solicitada e garantindo o recebimento da informação e o funcionamento do ambiente mesmo após a falha de um componente do mesmo.

Figura 8: Resultado da consulta

```

OWNER          TABLE_NAME
-----
TABLESPACE_NAME  CLUSTER_NAME
-----
IOT_NAME        STATUS      PCT_FREE    PCT_USED    INI_TRANS
-----
MAX_TRANS      INITIAL_EXTENT NEXT_EXTENT MIN_EXTENTS MAX_EXTENTS PCT_INCREASE
-----
FREELISTS      FREELIST_GROUPS LOG_B      NUM_ROWS      BLOCKS EMPTY_BLOCKS  AVG_SPACE
-----
CHAIN_CNT     AVG_ROW_LEN  AVG_SPACE_FREELIST_BLOCKS NUM_FREELIST_BLOCKS
-----
DEGREE
-----
INSTANCES      CACHE      TABLE_LO
-----
SAMPLE_SIZE    LAST_ANAL  PAR_IOT_TYPE TSNES        BUFFER_FLASH_C CELL_FLASH_ROW_MOVE
-----
GLOB_USE_DURATION SKIP_COR_MON CLUSTER_OWNER DEPENDENT
-----
COMPRESS      COMPRESS_FOR DRO_REA_SEG  RESULT
-----
106 rows selected.

```

Fonte: Giovana Canali e Thiago Henrique Caires

7. Conclusão

Alta disponibilidade do Banco de Dados para as empresas é uma necessidade para o mercado competitivo, trazendo recursos para atender melhor os clientes, quando bem programada as tecnologias como o Oracle RAC tendem a retornar muitos benefícios, como confidencialidade, integridade e disponibilidade que são os três pilares da segurança de dados e sistemas.

Com a instalação, configuração e utilização do Oracle RAC é possível analisar que a ferramenta traz opções para mascaramento de falhas dos nós, alta disponibilidade, escabilidade, integridade dos dados, e baixa perda de performance, trazendo recursos para o administrador de Banco de Dados (DBA).

9. Referências

- Canedo, F; Teixeira, V; Bruschi, G. (2014) “Gerenciamento e alta disponibilidade em armazenamento de banco de dados”
- Costa, H. (2009) “Alta disponibilidade e balanceamento de carga para melhoria de sistemas computacionais críticos usando software livre: Um estudo de caso.” (Dissertação em Magister Scientiae) Universidade Federal de Viçosa. Minas Gerais.
- HP. (2013) “Por que preciso me preocupar com a alta disponibilidade?”, http://www.hp.com/latam/br/pyme/solucoes/apr_solucoes_02.html, October
- Kinghost (2015) “Oracle RAC”, <http://www.kinghost.com.br/oracle-rac>, April
- Marcus, E; stern, H. (2003) “Blueprints for high availability.” 2ª ED. Indianapolis. Wiley Publishing, Inc:
- Morales, D. (2007) “Arquitetura dos Real Application Clusters da Oracle, e análise do seu Mecanismo de Cache Distribuído: o Cache Fusion. Instituto de Informática” Universidade Federal do Rio Grande do Sul (UFRGS)
- Oracle. (2010) “Utilizando clusters em aplicações de bancos de dados para reduzir os custos de TI.”, <http://www.oracle.com/technetwork/pt/database/clustering/documentation/utilizando-clusters-bancos-de-dados-1729081-ptb.pdf>, October
- Oracle. (2007) “Oracle clusterware and Oracle Real Application Clusters administration and deployment guide”, <http://www.oracle.com/pls/db102/homepage>, October
- Reis, A; Junior, C; Ferreira, J; Almeida, J; Silva, M; Gavinier, S. (2011) “Cluster de Alta Disponibilidade.” Faculdade de Tecnologia de Guaratinguetá (FATEC-GT) Guaratinguetá.
- Shrivastava, A Somasundaram, G; (2011) “Armazenamento e Gerenciamento de Informações.”, Bookman, p.25.
- Souza, C; Campos, P. (2008) Alta disponibilidade em Banco de Dados utilizando tecnologia Oracle 147f. (Monografia em Especialista em Gestão de Banco de Dados) Associação educativa do Brasil faculdades integradas ICESP. Brasília.
- Tanenbaum, A; STEEN, M. (2007) “Sistemas distribuídos-princípios e paradigmas.” 2ª Ed. São Paulo. Pearson Education.
- Veras, M. (2013) “Cloud Computing: Nova Arquitetura da TI.” Brasport.