

Gerenciamento e alta disponibilidade em armazenamento de banco de dados

Fabio Samuel dos Santos Canedo
Vitor de Oliveira Teixeira
Gustavo Cesar Bruschi (Orientador)

FATEC BAURU

acheroniano@gmail.com

vitor.teixeira2@fatec.sp.gov.br

gustavo@bruschi.net

Abstract. *In this article we discuss concepts and tools capable of providing solutions for a better storage management database with a focus on high availability and high performance. The concepts covered can be applied to stand-alone servers or clustered servers. Defined the concepts we will cover the use of a proprietary tool that provides high availability storage database.*

Resumo. *Neste artigo iremos abordar conceitos e ferramentas capazes de prover soluções para um melhor gerenciamento do armazenamento de banco de dados com foco em alta disponibilidade e alta performance. Os conceitos abordados podem ser aplicados em servidores stand-alone ou servidores em cluster. Definidos os conceitos vamos abordar o uso de uma ferramenta proprietária que prove alta disponibilidade em armazenamento de banco de dados.*

1. Introdução

Um grande desafio na vida de qualquer administrador de banco de dados é gerenciar e prover alto desempenho e alta disponibilidade na infraestrutura tecnológica.

Conforme Shrivastava e Somasundaram (2009, p. 25) a necessidade de armazenamento de informação no mundo corporativo cresce muito.

Para as empresas atingirem o sucesso de seus negócios é vital que utilizem ferramentas que gerenciem o armazenamento de dados, permitam a alta disponibilidade e a continuidade de negócios.

Em um ambiente de alta disponibilidade, paralisar o banco de dados para realizar manutenções de desempenho (*tunning*), realizar novas configurações ou troca de discos rígidos no ambiente tecnológico poderá trazer graves consequências financeiras para a empresa.

No artigo vamos utilizar a sigla de *Database Administrator* (DBA), essa sigla é utilizada no mercado de trabalho para denominar um administrador de banco de dados.

Com o crescimento do banco de dados o DBA necessita cada vez mais de ferramentas que aumentem sua produtividade e o ajude a automatizar as tarefas diárias.

O armazenamento e a alta disponibilidade em sistemas de banco de dados é um grande foco de qualquer DBA.

Quando se procura ter desempenho do acesso aos dados e alta disponibilidade, devem ser utilizadas tecnologias e recursos disponíveis no mercado de maneira adequada, garantindo inúmeras melhorias no gerenciamento e disponibilidade do ambiente.

2. Objetivo

O Objetivo do artigo é introduzir conceitos em alta disponibilidade em armazenamento de banco de dados, enfatizar a responsabilidade que um DBA deve ter com as informações cruciais do banco de dados e avaliar uma ferramenta disponível no mercado que facilita o gerenciamento do armazenamento dessas informações.

Serão criados dois cenários, para demonstrar particularidades de instalação da ferramenta em ambiente Windows e Linux.

Prover alta disponibilidade em armazenamento em banco de dados não é uma tarefa fácil, para que isso ocorra o DBA deverá conhecer seu ambiente de uma forma ampla.

Tais responsabilidades, inerentes ao DBA, definem o tema do artigo “Gerenciamento e alta disponibilidade em armazenamento de banco de dados”

3. Banco de dados relacional

Segundo Date (2004, p. 6): “[...] um sistema de banco de dados é basicamente um sistema computadorizado de manutenção de registros”. Esses registros são armazenados em um banco de dados que é controlado por um Sistema Gerenciador de Banco de Dados (SGDB).

Em conformidade com Elmasri e Navathe (2011, p. 3) o SGBD possui as características de compartilhamento de dados, controle de acesso, controle de redundância, interfaceamento, esquematização, backup e controle de integridade, essas opções garantem os dados seguros, íntegros e disponíveis quando solicitado por qualquer usuário do sistema.

O modelo de dados mais popular é o Entidade-Relacionamento (ER), no modelo ER possuímos entidades, atributos e relacionamentos.

A linguagem padrão do modelo ER é a linguagem *Structured Query Language* (SQL) que foi baseada em álgebra relacional.

A linguagem SQL é de fácil compreensão, e por esse motivo, tornou-se um padrão adotado pelo mercado.

4. Armazenamento

4.1. Sistema de Arquivos

Para Flynn e McHoes (2002, p. 193) um sistema de arquivos “[...] controla todos os arquivos em um sistema e processa os comandos dos usuários que desejam interagir (ler, gravar, modificar, criar, apagar, etc.) [...]”.

Qualquer sistema de arquivos deve ser capaz de armazenar, organizar, decodificar e manipular os dados independentemente da forma de armazenamento utilizada.

Outro ponto abordado por Flynn e McHoes (2002) é que um sistema de arquivo deve gerenciar e manter a integridade e segurança dos arquivos.

Os sistemas de arquivos devem permitir a identificação, organização, compartilhamento, acesso, proteção e operações de Entrada/Saída (E/S) nos dados armazenados, além de garantir um acesso rápido aos dados solicitados pelo sistema operacional.

Os sistemas de arquivos mais comuns são:

- Apple: HFS+;
- Unix: UFS, Ext2, Ext3 e Ext4;
- IBM: HPFS;
- Microsoft: FAT 32 e NTFS;

Em resumo um sistema de arquivo deve ser extremamente seguro, rápido e de fácil manutenção.

4.2. Storage

Shrivastava e Somasundaram (2009, p. 33) definem um storage como “[...] um dispositivo que armazena dados de forma persistente para uso posterior.”.

Um storage deve ser gerenciável, prover disponibilidade, segurança, escalabilidade, desempenho e integridade dos dados armazenados.

Destacam-se os tipos:

- *Direct-Attached Storage* (DAS): um dispositivo de armazenamento tradicional, não utiliza nenhum tipo de rede.
- *Network-Attached Storage* (NAS): um dispositivo conectado a um computador que é acessado através de uma rede, normalmente é associado ao protocolo *Network File System* (NFS).
- *Storage Area Network* (SAN): uma rede especializada, que permite a outros computadores terem acesso ao armazenamento de dados.

A diferença crucial entre o NAS e o SAN é a forma como o sistema de arquivos acessa a SAN.

Pela SAN o acesso é ao nível de endereçamento de bloco (*block-addressing raw level data*) dos discos da SAN, esse modo normalmente está associado ao protocolo *Fibre Channel*.

4.3. RAID

Segundo Shrivastava e Somasundaram (2009, p. 73) RAID é uma tecnologia que possibilita o uso de múltiplos discos em conjunto, o qual fornece proteção aos dados contra falhas de *Hard Disk Drives* (HDDs), essa tecnologia também pode ser aplicada em *storage*.

De forma geral o RAID gera um aumento de desempenho de E/S no sistema de armazenamento, devido ao uso de vários HDDs.

Em 1987, o artigo “*A Case for Redundant Arrays of Inexpensive Disks (RAID)*”, publicado por Patterson, Gibson e Katz na Universidade da Califórnia, em *Berkeley*, descreve o uso de discos baratos e de pequena capacidade na implementação de RAID.

Com o tempo o termo RAID foi redefinido para *Redundant Array of Independent Disks* refletindo o avanço em tecnologia de armazenamento.

Shrivastava e Somasundaram (2009, p. 74) afirmam que a tecnologia RAID deixou de ser um conceito acadêmico para se tornar um padrão da indústria.

Para Shrivastava e Somasundaram (2009, p. 74) “Há dois tipos de implementação de RAID: *software* ou *hardware*”, conseguindo-se maior confiabilidade, compatibilidade e desempenho em implementações feitas via *hardware*.

O RAID via *hardware* pode ser implementado através de placas controladoras ou através de *storages* externos que apresentam os volumes para o *host*, e estes comunicam-se através de um protocolo suportado pelo conjunto de discos.

- RAID 0 - *Striping*, os dados são distribuídos entre os discos que estão alocados na RAID, não possui segurança, a perda de um disco representa a perda do *array*.
- RAID 1 - *Mirroring*, os dados são gravados simultaneamente, identicamente em dois ou mais discos, a falha de um disco não acarreta a perda de dados.
- RAID 1+0 - Também conhecido como RAID 10 é a combinação de *striping* e *mirroring*. Os dados são primeiro espelhados e depois distribuídos.
- RAID 5 - Mínimo de três discos para ser implementado, distribui os dados e os códigos de paridade desses dados entre os discos da *array*. Se um disco falhar não há perda de informações. Os dados e a paridade desses dados são usados para reconstruir o disco que falhou. O RAID 5 é o preferido para mensagens de e-mail, *data mining*, serviços de *media* além de SGBD pois otimiza os acessos aos dados.

Discos únicos podem falhar isso representa uma ameaça constante para o DBA. Utilizando técnicas de RAID aumenta-se a disponibilidade através do espelhamento e paridade, além de melhorar o desempenho, pois ao distribuir os dados por múltiplos HDDs incrementa-se a performance de E/S.

RAID é uma tecnologia fundamental para diversos avanços em armazenamento e oferece alto desempenho e alta disponibilidade.

4.4. *Raw devices*

Um conceito importante em armazenamento é o do *Raw devices* (Dispositivo de armazenamento de dados em estado bruto).

Em um ambiente crítico, bancos de dados precisam do melhor desempenho de gravação e de recuperação de dados em disco.

A melhor performance de leitura e gravação no banco de dados, é um dos objetivos mais almejados por qualquer DBA.

Raw devices são partições na forma bruta, ou seja, sem formatação, nesse caso é necessário um aplicativo específico para seu gerenciamento, pois como a partição não está formatada ela não possui um sistema de arquivos.

5. Ambientes críticos e alta disponibilidade

5.1. Cluster

Pitanga (2008, p. 33) justifica a utilização de cluster com o seguinte argumento “[...] um grande problema computacional em que o processamento paralelo se considerado uma vantagem, pode ser indicado para utilização em um *cluster*”.

Um *cluster* é um agrupamento de dois ou mais computadores ou sistemas que compartilham basicamente memória, armazenamento, processamento e rede.

Em um sistema de cluster os computadores são denominados como nós e conectados entre si através de uma *Local Área Network* (LAN).

As principais características de um cluster são: desempenho, balanceamento de carga e redundância de dados.

Para o usuário um cluster resume-se em um único computador bem mais potente.

Alguns tipos de clusters:

- *High Availability*: de alta disponibilidade, utilizado em base de dados de missão crítica.
- *Load Balancing*: distribui o tráfego entrante e recursos dos nós, utilizado em *web farms* e também em cluster de processamento distribuído. Esse modelo aumenta o desempenho e a disponibilidade de aplicações principalmente de grandes tarefas computacionais, também são usados em aplicações financeiras e científicas.

5.2. Continuidade de Negócios

Business Continuity (BC), implica em preparo, resposta e recuperação de uma falha sistêmica, envolve medidas pró-ativas, análises e avaliações de risco, proteção de dados, segurança e também medidas reativas em caso de falhas.

Segundo Shrivastava e Somasundaram (2009, p. 251) a “Continuidade do Negócio” tem que garantir a “disponibilidade de informações” garantindo as operações da empresa.

Indisponibilidade de dados, tempo inativo ou paradas não planejadas resultam em perda de produtividade e receita, desempenho financeiro fraco e danos à reputação.

O impacto do tempo de inatividade sobre o negócio é a soma de todas essas perdas geradas como resultado de determinada interrupção.

O custo médio por hora de tempo inativo fornece uma estimativa chave para estabelecer soluções de BC.

5.3. Tolerância a Falhas

De acordo com Shrivastava e Somasundaram (2009, p. 260) “Para atenuar um ponto único de falha, sistemas são projetados com redundância, de modo que só falhem se todos os componentes do grupo de redundância falharem.”

Essa estrutura assegura que a falha de um único componente não afete a disponibilidade de dados.

Diretrizes rigorosas são aplicadas para implementar a estrutura de tolerância a falhas em *Data Centers*, dentre as quais podemos citar:

- Configuração de *Storage Array*.
- Configuração de RAID assegura uma operação contínua em caso de falha do *Hard Disk* (HD)
- Configuração de *Clusters*.

Novas tecnologias levam a um conjunto variado de opções em termos de dispositivos e soluções de armazenamento que atendam a requisitos de alta disponibilidade e continuidade de negócios.

Analisar configurações de *hardware* e *software* e o seu impacto nas operações da empresa, são essenciais, pois em um ambiente de negócios em constante mudança, o BC se torna uma tarefa exigente.

Em seguida discutimos uma solução específica de alta disponibilidade em armazenamento.

6. Alta disponibilidade em armazenamento

A utilização de ferramentas específicas que garantam integridade, alta disponibilidade, segurança e alto desempenho no armazenamento de dados em um ambiente crítico devem ser cuidadosamente escolhidos pelo DBA, pois um planejamento que não aborde corretamente o crescimento da empresa afetará sua necessidade de armazenamento, o acesso íntegro e seguro ao longo do tempo.

Um gerenciamento incorreto das informações disponibilizadas em seu banco de dados afeta drasticamente a continuidade dos negócios.

Todo DBA procura ferramentas onde ele possa ter um melhor desempenho aliado com maior segurança. Os conceitos básicos acima irão nos ajudar a entender uma tecnologia desenvolvida pela Oracle que foi implementada na versão Oracle 10g release 1 chamada de Oracle *Automatic Storage Management* (ASM).

A tecnologia ASM tem seu próprio sistema de arquivos, chamado de Oracle ACFS, (*Automatic Storage Management Cluster File System*) ele é um sistema de arquivos escalável e multi-plataforma.

O Oracle ASM é capaz de gerenciar o armazenamento em servidores *stand-alone* ou servidores em *cluster*.

A ferramenta dispõe de funções integradas que permitem ao DBA a análise completa do seu ambiente de armazenamento, garantindo a integridade das informações disponibilizadas.

É possível visualizar a carga de E/S de informações em uso no grupo de discos, e essa visualização permite ao DBA planejar de maneira mais eficiente as manutenções técnicas, além de escalonar adequadamente a expansão do ambiente de armazenamento.

7. Banco de dados Oracle

Oracle *Corporation* mais conhecida como Oracle é uma empresa americana com sede na cidade de *Redwood City* no estado da Califórnia.

Seu principal produto é Oracle *Database* que possui as versões *Express Edition*, *Standard Edition*, *Standard Edition One* e *Enterprise Edition* atualmente esta na versão 11G *release 2*.

O banco de dados Oracle tem como sua principal característica seguir o modelo relacional além de evoluir a cada versão oferecendo ferramentas para um melhor gerenciamento, provendo escalabilidade, segurança e alto desempenho para o armazenamento de dados.

Segundo Bryla e Loney (2007, p. 30) “Os arquivos de banco de dados Oracle são agrupados em uma ou mais tablespaces.”.

Enquanto uma *tablespace* é uma divisão lógica, um *datafile* é uma divisão física que por sua vez pode fazer parte de uma ou mais *tablespaces*.

Uma instância Oracle, conceito fundamental na estrutura do banco de dados Oracle é uma combinação de área de memória reservada e processos em *background* usados para gerenciar o acesso às informações armazenadas no banco de dados.

8. Oracle ASM

Um dos recursos que permite alta disponibilidade ao acesso aos dados é o *Oracle ASM*.

O Oracle ASM é um gerenciador de volumes e um sistema de arquivos que utiliza o recurso de multiplexação automatizada que distribui entre os discos disponíveis arquivos de dados, arquivos de redo log e arquivos de controle, ele suporta instâncias simples e configurações com múltiplas instâncias através do Oracle *Real Application Cluster* (RAC).

Essa solução de gerenciamento e armazenamento utiliza dispositivos *Raw devices* separados por grupos de discos, utilizados para armazenar os arquivos de dados.

Sendo os grupos de discos uma coleção de discos gerenciados como uma única unidade.

A tecnologia Oracle ASM permite o gerenciamento, adição e remoção de novos discos ao grupo de discos enquanto o banco de dados não para de funcionar, desta forma eliminamos paradas não programadas.

O Oracle ASM pode coexistir com outros tipos de sistemas de arquivos e tecnologias de gerenciamento de armazenamento de discos, facilitando assim sua implantação.

Segundo Bryla e Loney (2007, p. 102) uma instância Oracle ASM é similar ao conceito de instância de um banco de dados, possui uma área chamada de *System Global Area* (SGA) e processos de *background*, mas por realizar uma quantidade menor de tarefas a SGA do Oracle ASM tem menor impacto no desempenho do servidor.

Instâncias do Oracle ASM disponibilizam os arquivos e montam os grupos de discos para a instância do banco de dados.

Em ambientes com Oracle RAC é criado uma instância ASM para cada nó do cluster. Apenas uma instância ASM é necessária para cada nó, não importa quantas instâncias de banco de dados estejam no nó.

A Figura 1 mostra múltiplas instâncias de banco de dados acessando uma única instância Oracle ASM.

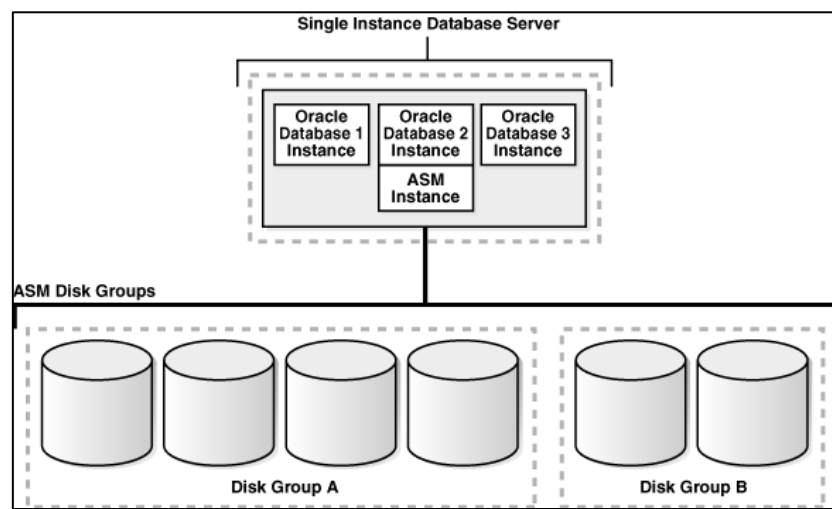


Figura 1: Oracle ASM for Single-Instance Oracle Databases
Fonte: Oracle (2010a).

A Figura 2 mostra o Oracle RAC acessando um pool de *storage*, uma instância Oracle ASM para cada nó, servindo múltiplas ou simples instâncias de banco de dados. Todos os bancos podem acessar e compartilhar os mesmos dois grupos de discos.

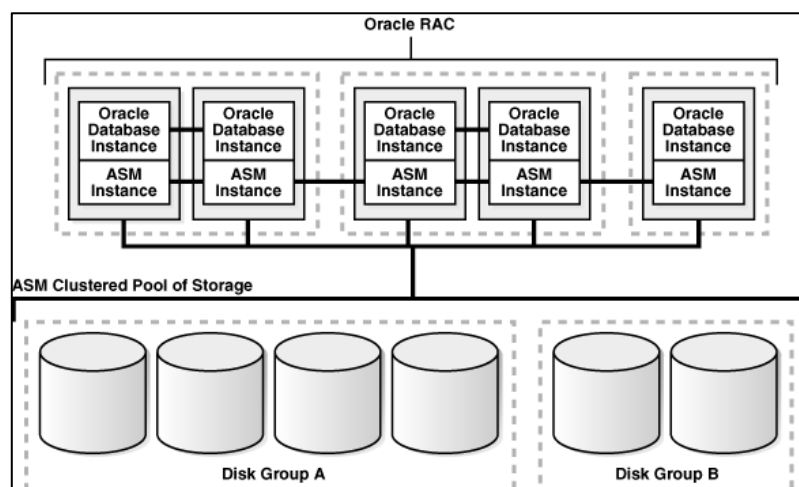


Figura 2: Oracle ASM Cluster Configuration with Oracle RAC
Fonte: Oracle (2010a).

Um grupo de discos é um conjunto de vários discos.

É o objeto fundamental que o Oracle ASM gerencia, inclui discos, arquivos e as unidades de alocação.

Um banco de dados pode fazer uso de arquivos espalhados em vários grupos de discos pelo fato que os arquivos fazem parte de vários bancos de dados.

Oracle ASM *Disks* são dispositivos de armazenamento disponibilizados para o grupo de discos do Oracle ASM, eles podem conter:

- Um disco ou uma partição de uma *array* de discos;
- Um disco inteiro ou uma partição de um disco;
- *Logical Volumes*;
- NFS;

O Oracle ASM espalha os arquivos proporcionalmente para o grupo de discos.

Esse padrão de armazenamento mantém cada disco com a mesma capacidade e garante que cada disco tenha a mesma carga de E/S.

Esse balanceamento de carga desencoraja a configuração de diferentes ASM *Disks* em um mesmo disco físico.

8.1. Oracle ASM Files

São arquivos armazenados nos grupos de disco, o banco de dados comunica-se com esses arquivos.

É como o banco de dados se relaciona a um arquivo armazenado em qualquer sistema de arquivos.

Quando um arquivo é criado no banco de dados ele é dividido e distribuído (*striped*) através do grupo de discos, por exemplo, um grupo de seis discos com Oracle ASM *disks* tem seu espaço de alocação distribuído igualmente nos seis discos. Quando acessado, ele será lido nos seis discos em paralelo, aumentando assim seu desempenho.

9. Implementação do Oracle ASM

Para conceituar melhor a ferramenta Oracle ASM efetuamos a instalação do banco de dados Oracle *Database 11G release 1* no sistema operacional Linux CentOS 5.4 32 bits e Oracle *Database 11G release 2* em *Microsoft Windows Server 2012 Standard* 64 bits. Observamos que em ambos os sistemas foi necessário à preparação do grupo de discos no modo RAW *device* antes da instalação do banco de dados.

Na instalação efetuada no Linux CentOS foi necessário configurar e habilitar o serviço chamado *rawdevice*, foi feito a instalação e configuração de três pacotes para a preparação do ambiente, além da criação de mais três grupos de usuários específicos para a instalação do Oracle ASM.

Durante a instalação do Oracle *Database 11G release 1* foi possível ver o momento onde é criado a instância ASM.

No *Microsoft Windows Server 2012* existe uma ferramenta incorporada no sistema operacional que ajuda no processo de gerenciamento de disco, sendo possível criar a partição primária no modo RAW.

Foi necessário à instalação do Oracle *Grid Infrastructure* para a criação do grupo de discos, após isso efetuamos a instalação do *Oracle Database 11G release 2*.

Em ambos os casos utilizamos as ferramentas *Enterprise Manager* e *SQL Plus* para verificar o *status* da instância do banco de dados e ASM.

Apesar da versão dos sistemas citados acima não serem homologados pela Oracle a instalação ocorreu de forma satisfatória.

Durante os testes para a instalação utilizamos também o sistema *Microsoft Windows Server 2008 Standard 32 bits*. Na etapa treze para a instalação do Oracle *Database 11G release 2* ocorreu o erro INS-35210, foi apurado que sua causa ocorre pelo fato do Oracle ASM não poder ser instalado em versão 32 bits do *Microsoft Windows*.

Para implementar o Oracle ASM no ambiente Linux é necessária uma melhor preparação do ambiente em comparação ao ambiente *Microsoft*.

A instalação foi efetuada em ambiente virtual utilizando o Virtual Box 4.2.12 que é de propriedade da Oracle.

A ferramenta Oracle ASM dispõem de recursos que facilitam o gerenciamento do grupo de discos, ao precisar expandir o grupo de discos não é necessário à parada do banco de dados.

Essa tecnologia possui uma estrutura flexível, é possível gerenciar instancias ASM e grupo de discos através do aplicativo *Enterprise Manager* ou *SQL Plus*. O Oracle ASM fornece balanceamento de carga de E/S nos grupos de discos e LUNs.

Utilizando o aplicativo *Enterprise Manager* visualizamos o grupo de discos o estado de cada disco, histórico de uso dos discos, possui também um gráfico de tempo de reposta de E/S, esses recursos provem gerenciamento além dos recursos de alta disponibilidade como a redundância de discos.

10. Conclusão

A necessidade de gerenciamento e alta disponibilidade de armazenamento devem unir diversas técnicas e ferramentas.

Pontos falhos como discos únicos e ambientes sem redundância afetam diretamente a continuidade dos negócios.

Um sistema de arquivos em conjunto com um gerenciador de volumes deve utilizar técnicas de alta disponibilidade e gerenciamento de uma forma eficiente e segura.

As ferramentas disponíveis no mercado garantem ao DBA soluções altamente eficientes no gerenciamento do banco de dados, e quando implementadas garantem desempenho, escalabilidade, segurança e alta disponibilidade das informações armazenadas.

Ao longo da instalação, configuração e utilização do Oracle ASM foi possível verificar que a ferramenta traz opções para gerenciamento e alta disponibilidade no armazenamento do banco de dados de uma forma prática, ajudando no trabalho do DBA.

Referências

- BRYLA, B; LONEY K. Manual do DBA, Bookman, 2007.
- DATE, C. J. Introdução a Sistemas de Banco de Dados, 8 ed. Elsevier, 2004.
- ELMASRI, R; NAVATHE, S. B. Sistemas de banco de Dados, 6 ed. Pearson, 2011.
- FLYNN, I. M.; MCHOES, M. A. Introdução aos Sistemas Operacionais, Thomson, 2002.
- ORACLE. 2010a. Automatic Storage Management E16102-05. Disponível em: <http://docs.oracle.com/cd/E18283_01/server.112/e16102.pdf>. Acesso em: 02 abr. 2013
- ORACLE. 2010b. Automatic Storage Management E18951-03. Disponível em: <http://docs.oracle.com/cd/E11882_01/server.112/e18951/asmfilesystem.htm>. Acesso em: 02 abr. 2013
- ORACLE Oracle Database 10g - Automatic Storage Management. Disponível em: <<http://www.oracle.com/technetwork/database/asmov-134266.pdf>>. Acesso em: 02 abr. 2013
- PATTERSON, D. A; GIBSON, G; KATZ, R. H. A Case for Redundant Arrays of Inexpensive Disks (RAID). Disponível em: <<http://www.eecs.berkeley.edu/Pubs/TechRpts/1987/CSD-87-391.pdf>> Acesso em: 02 abr. 2013
- PITANGA, M. Construindo Supercomputadores com Linux, 3 ed. Brasport, 2008.
- PITANGA, M. Computação em Cluster. Disponível em: <<http://www.rozero.host22.com/disciplinas/unatec/arquitetura/Cluster.pdf>>. Acesso em: 02 abr. 2013
- SHRIVASTAVA, A; SOMASUNDARAM, G. Armazenamento e Gerenciamento de Informações, Bookman, 2009.