

Implementação de Alta Disponibilidade para Aplicações com Recursos Geograficamente Dispersos

Alexandre Fagundes Biscaio¹, André Peres²

¹ Acadêmico do Curso de Sistemas de Informação da ULBRA Guaíba
< biscaio@terra.com.br >

² Professor Orientador do Curso de Sistemas de Informação da ULBRA Guaíba
< peres@guaiba.ulbra.tche.br >

Resumo: Este artigo traz como proposta duas soluções para entrega de áreas de dados a aplicações, provendo redundância e/ou alta disponibilidade aplicáveis a cada necessidade de solução. São abortados os fatores mais relevantes para a solução sem, no entanto, citar produtos ou fabricantes específicos.

Abstract: This article brings as two solutions for delivering storage area from to applications, providing redundancy and/or high availability applicable to each requirement solution. Are aborted factors most relevant for the solution without, however, quote products or manufactures.

1 - INTRODUÇÃO

Storage, ou área de armazenamento, é onde os dados são armazenados de forma não volátil, ou seja, mesmo que um computador seja desligado, ao ligá-lo novamente os dados armazenados estarão acessíveis, pois os mesmos são gravados de forma magnética (ou óptica) a dispositivos chamados discos rígidos (ou HD- *Hard Drives*) ou discos ópticos, como mídias de DVD. Sua importância para sistemas de informação está na capacidade de manter os dados providos por esses sistemas aos usuários que os utilizam a fim de suportar uma determinada área de negócio de uma empresa ou mesmo como usuários finais no caso, por exemplo, de sistemas online, como a própria *Internet* (rede mundial de computadores).

Este artigo apresenta uma proposta para o desenvolvimento de uma solução capaz de fornecer alta disponibilidade à camada de acesso aos dados (*storage*) por aplicações que necessitem de alta disponibilidade. Considerando que todas as demais camadas de infra-estrutura são facilmente contingenciáveis, apenas a camada de *storage* possui poucas soluções com custo acessível que possam ser implementadas por pequenas ou médias empresas devido ao custo envolvido em

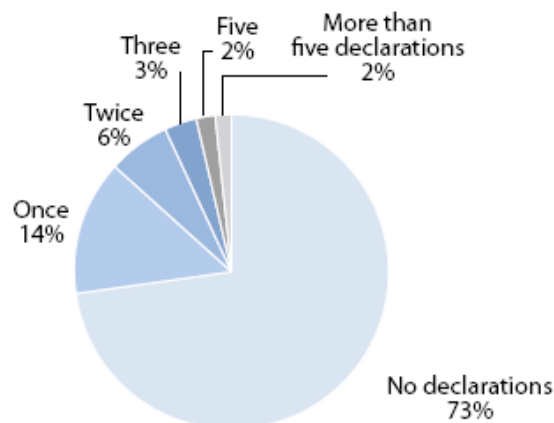
tais soluções, já que usualmente são considerados *storages* e *softwares* dos maiores *players* desta área no mercado.

Alinhado a essa necessidade agrega-se outros motivadores para a busca de uma solução que possa prover redundância à camada de *storage*: a) o percentual de empresas que já sofreram alguma interrupção dos serviços de Tecnologia da Informação e b) os motivos que levam as interrupções do funcionamento de aplicações por falhas de infra-estrutura.

De acordo com o relatório *Disaster Recovery Journal-October 2007-Global Disaster Recovery Preparedness Online Survey*, da *Forrester Research, Inc*, 3 em cada 4 empresas já experimentaram interrupções em seus serviços de TI nos últimos 5 anos de operação, ocasionando conseqüentemente interrupções nos negócios da empresa, trazendo danos ao faturamento, imagem e serviços oferecidos (muitas vezes, serviços essenciais a sociedade), conforme demonstrado na Figura 1. (BALAOURAS, 2008)

Figure 1 More Than A Quarter Of Companies Have Declared A Disaster In The Past Five Years

“How many times have you had to declare a ‘disaster’ and recover operations at your recovery site in the past five years?”



Base: 250 disaster recovery decision-makers and influencers at businesses globally (percentages may not total 100 because of rounding)

Source: Forrester/Disaster Recovery Journal October 2007 Global Disaster Recovery Preparedness Online Survey

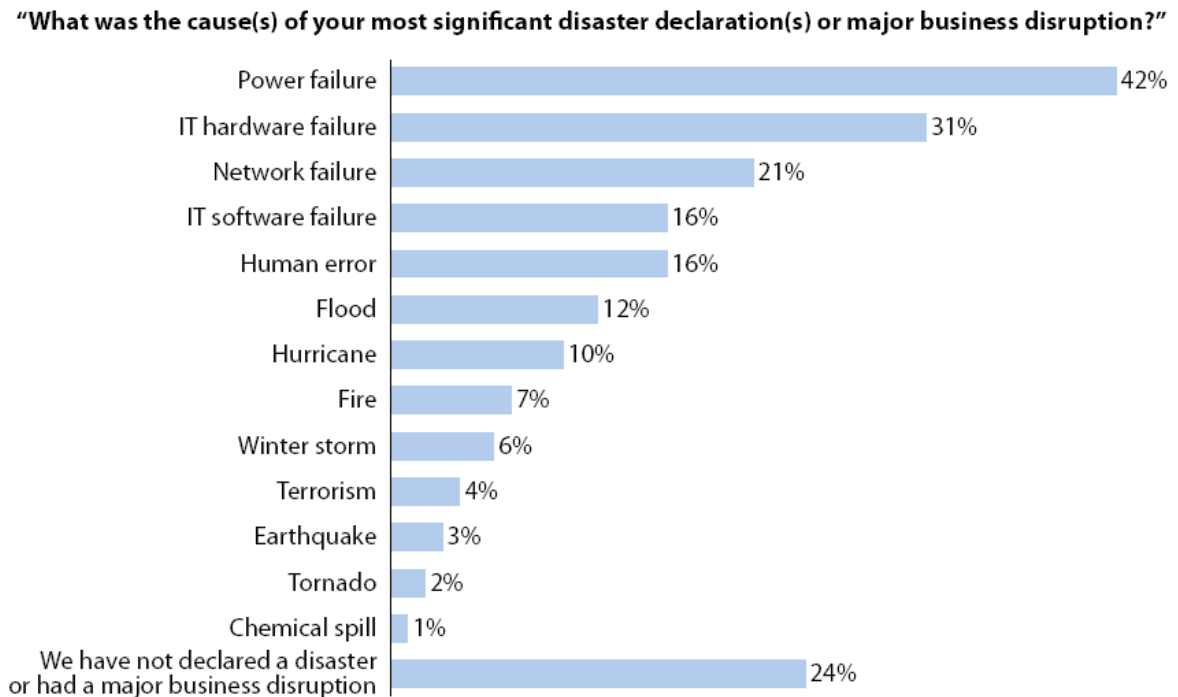
Source: Forrester Research, Inc.

Figura 1: Mais de um quarto das companhias declararam um desastre nos últimos cinco anos

Agregado a isso temos também os motivos destas falhas e o quanto cada uma representa, percentualmente, como pontos a serem tratados a fim de diminuir-se essas falhas, ou propor cenários de infra-estrutura que possam

suportar a maioria destas falhas melhorando assim a disponibilidade de sistemas de Tecnologia de Informação. Tais fatores são mostrados abaixo na Figura 2. (BALAOURAS, 2008)

Figure 2 Power Failures Are The Most Common Cause Of Declared Disasters And Downtime



Base: 250 disaster recovery decision-makers and influencers at businesses worldwide (multiple responses accepted) (Does not include those who answered “Other” or “Don’t know”)

Figura 2: motivos das falhas em infra-estrutura que interromperam serviços de Tecnologia da Informação nas empresas pesquisadas

OBJETIVOS DO TRABALHO

O objetivo deste trabalho é propor cenários, utilizando-se de técnicas e tecnologias disponíveis para dispor de forma redundante ou contingente, por parte de Aplicações de Tecnologia da Informação, acesso à camada de dados da infra-estrutura utilizada.

2 - REFERENCIAL TEÓRICO

Para embasar as propostas que serão apresentadas neste trabalho foram levantadas as técnicas e tecnologias existentes para disponibilizar redundância e/ou alta disponibilidade à camada de dados de uma aplicação (*storage*).

2.1 MATRIZ REDUNDANTE DE DISCOS INDEPENDENTES (RAID)

Técnica que visa criar redundância e melhorar o desempenho no acesso aos dados armazenados, o RAID (*redundant array of independent disks*), concebido pela IBM em 1978, combina a gravação e leitura em dois ou mais discos onde um algoritmo determina como será feita a gravação e a leitura dos dados (por espelhamento ou paridade). Atualmente existem diversos níveis de RAID, dependendo da aplicação a se utilizar deles uns são mais seguros do que outros, bem como uns possuem melhor desempenho do que outros. Este trabalho tratará apenas de duas destas técnicas, mais comumente utilizadas no mercado corporativo, que são RAID 1 e RAID 5. (GOUVEIA, 2005)

O RAID 1 duplica a quantidade de discos a fim de criar um espelho dos dados entre os discos utilizados, conforme demonstrado na Figura 3. Obviamente a capacidade de armazenamento necessária deverá possuir o dobro dos discos para que esta técnica possa ser adotada.

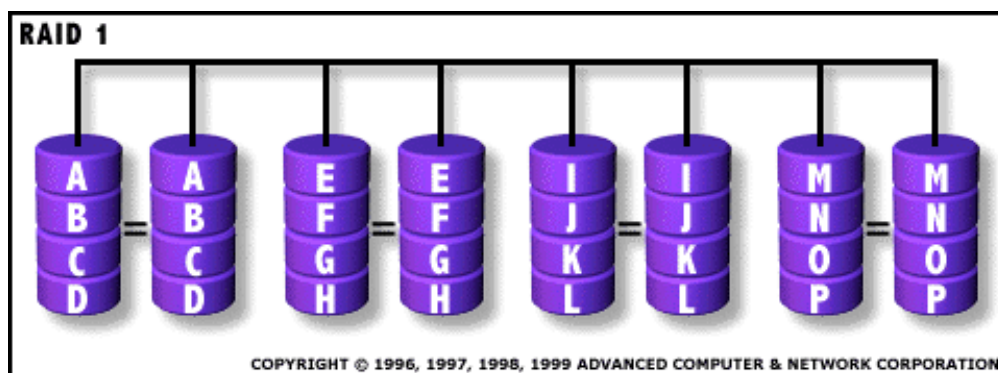


Figura 3: diagrama mostrando disponibilização de discos em RAID 1

Estando os discos distribuídos em pares, com técnica RAID 1, caso um dos discos do par venha a falhar, o outro disco continua provendo acesso aos dados, pois estão replicados, contendo a mesma informação. Esta técnica pode ser implementada contendo-se um mínimo de dois discos.

O RAID 5, técnica um pouco mais avançada e menos onerosa financeiramente pode ser implementada contendo no mínimo 3 discos, onde os dados são distribuídos entre todos os discos, mas em um deles (e não sempre no mesmo disco) informações de paridade são calculadas e gravadas, onde em caso de falha de um disco o dado necessário possa ser calculado a partir de uma operação XOR entre as informações armazenadas nos demais discos, conforme demonstrado na Figura 4 abaixo.

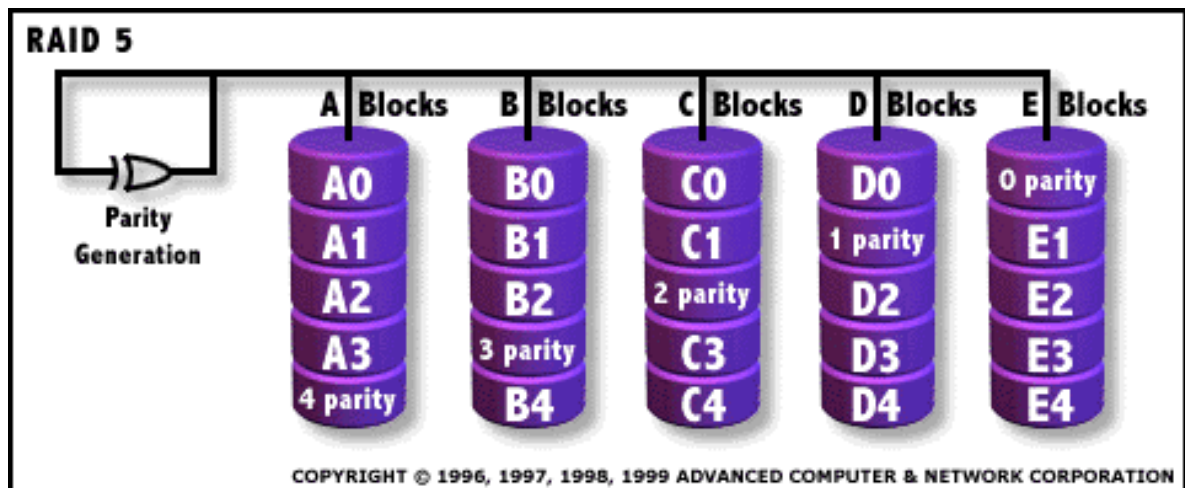


Figura 4: diagrama mostrando disponibilização de discos em RAID 5

Essa técnica além de menos onerosa financeiramente, já que não duplica a quantidade de discos para prover redundância um a um possui um melhor desempenho já que os dados são distribuídos entre os discos do RAID e em um processo de leitura e gravação o esforço é compartilhado entre os discos, deixando de sobrecarregar apenas um conjunto menor de discos (como no RAID 1, por exemplo).

Mesmo com esse nível de redundância provido pela técnica de RAID, um ponto de falha continua existindo no acesso aos dados: falha da controladora que provê acesso aos discos. Para este problema, foi criado um conceito chamado *Duplexing*, onde o número de controladoras de acesso a discos são colocadas de forma redundante a fim de prover contingência no acesso aos discos, conforme demonstrado na Figura 5 abaixo.

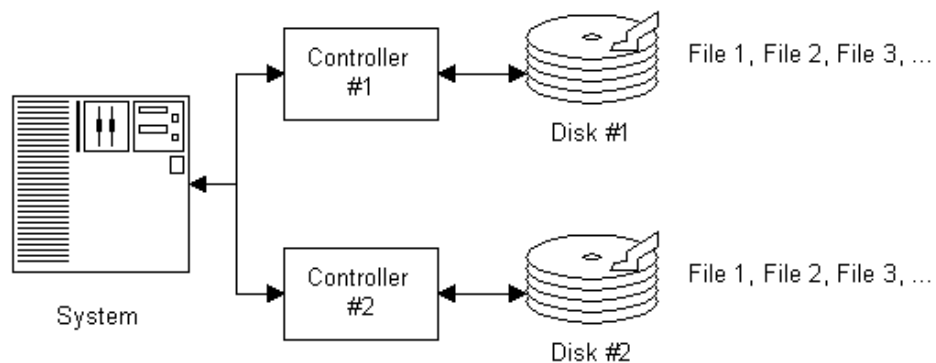


Figura 5: disposição de RAID utilizando duplexing

Neste exemplo, um conjunto de discos poderia ser, por exemplo, dispostos em RAID 1 onde os pares de discos de cada conjunto deveria estar conectados um a cada controladora. Desta forma, caso uma controladora falhe e os discos por ela atendidos não possam ser acessados, os dados continuarão sendo providos pela outra controladora e demais discos.

2.2 REPLICAÇÃO DE STORAGES (OU LUNS E META-LUNS)

A fim de manter os dados íntegros e acessíveis, sem necessitar executar nenhum procedimento de restore de fitas magnéticas (que é extremamente demorado) em caso de perda do acesso aos dados de um *storage*, existe a possibilidade de replicar as unidades lógicas (LUN- *Logical Unit Name*) para um *storage* que não atenda um ambiente de produção. Desta forma, a cada operação de atualização dos dados (gravação ou exclusão) os dados terem sido gravados em um *storage* de produção são replicados para um *storage* de *backup* (replicação síncrona). Também é possível, a fim de minimização do investimento em infraestrutura estabelecer um delta entre uma replicação e outra, com horários pré-determinados (replicação assíncrona), onde um histórico de atualização dos dados é armazenado em uma área específica e posteriormente é atualizado no *storage* de *backup*, conforme mostra a Figura 6 abaixo, porém esta solução exige *software* especializado.

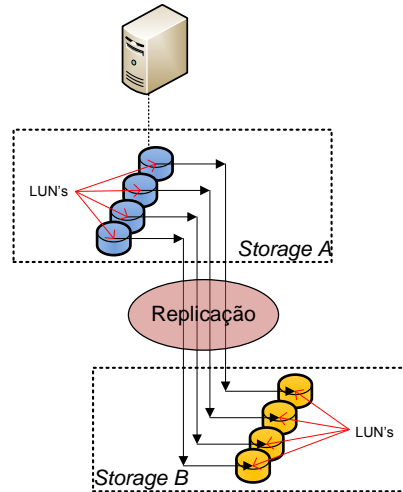


Figura 6: replicação de LUNs entre *storages*

Com o diagrama acima pode-se verificar que no caso de o *storage* de produção ficar indisponível e precisar-se utilizar os dados existentes no *storage* de *backup* existirá outro problema, desta vez não técnico, mas procedimental que é o tempo necessário para disponibilizar um servidor com acesso ao *storage* de *backup*: a reconfiguração do acesso do servidor às LUN's do outro *storage* precisará ser refeita (*zoning*). Isso, dependendo do número de LUNs existentes pode levar alguns minutos, ou até mesmo horas, cruciais para muitas aplicações e áreas de negócios. Isso ocorre, pois existe uma questão de associação lógica na administração dos dados do *storage* que indica quais servidores podem acessar quais áreas de dados (LUNs ou Meta LUNS).

Essa replicação pode, em alguns cenários, ser realizada por *hardware* especializado (Figura 6), ou ser realizada pelo próprio sistema operacional que se utiliza dos *storages*, utilizando-se de ferramentas de replicação específicas para este fim ou contando com tecnologias como o RAID1, já mostrado anteriormente. No diagrama abaixo, Figura 7, está um exemplo de como seria a topologia de um servidor com replicação por *software*.

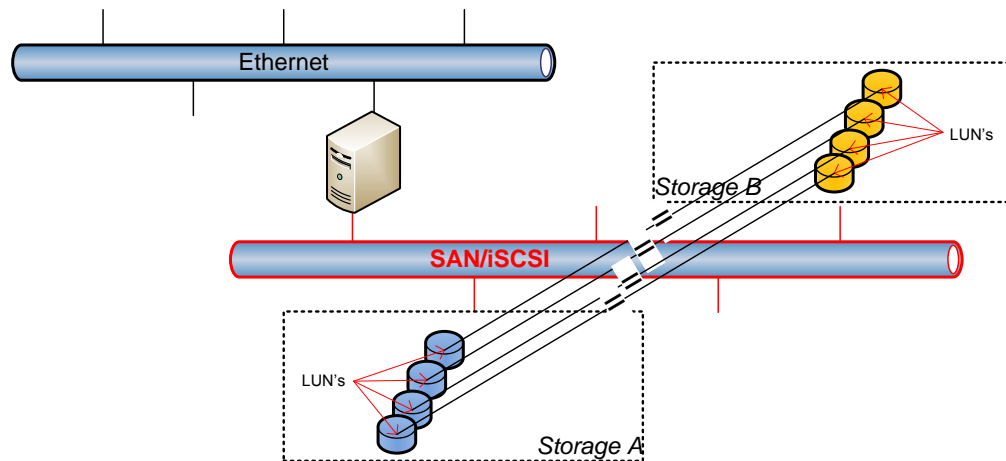


Figura 7: replicação de dados por *software* (dados)

2.3 VIRTUALIZAÇÃO DE *STORAGES*

Virtualização de *storage* é a técnica existente atualmente que pode ser utilizada para minimizar ou extinguir o tempo gasto em reconfiguração dos *zonnings* necessária para fazer um servidor acessar outro *storage* que possui as LUNs replicadas a partir de um *storage* de produção (*storage A* como mostrado anteriormente). Com isso aplicado, mesmo que um *storage* deixe de funcionar, os dados fornecidos ao servidor continuarão disponíveis, já que eles não estão armazenados unicamente em um *storage* e sim, em mais de um. Neste cenário o sistema operacional não acessa fisicamente os discos (LUNs) e sim de uma camada de *software* que provê acesso virtual aos discos (LUNs ou Meta LUNs), conforme mostrado na Figura 8, onde esta deverá também ser redundante conforme demonstrado na Figura 9. (WESLEY, 2005)

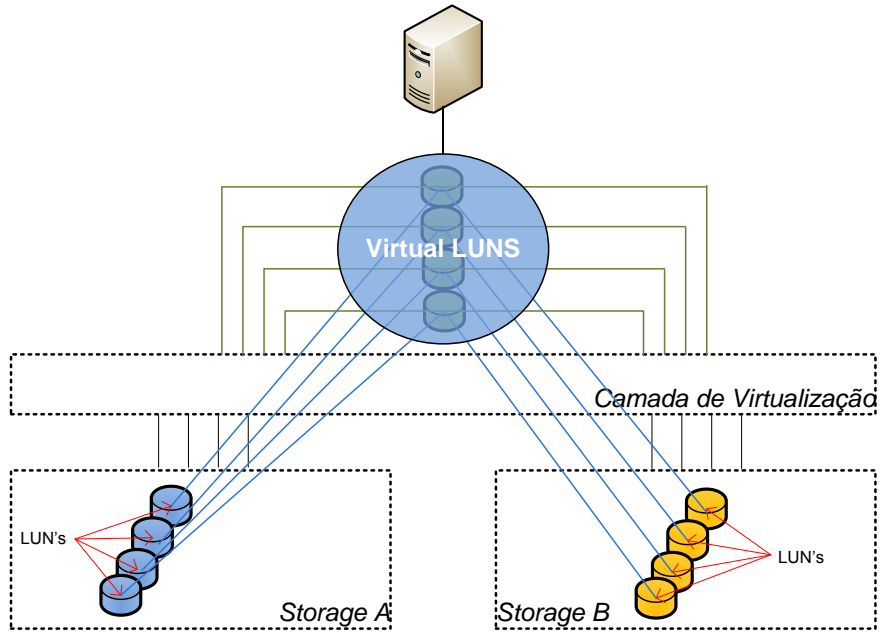


Figura 8: virtualização de storage

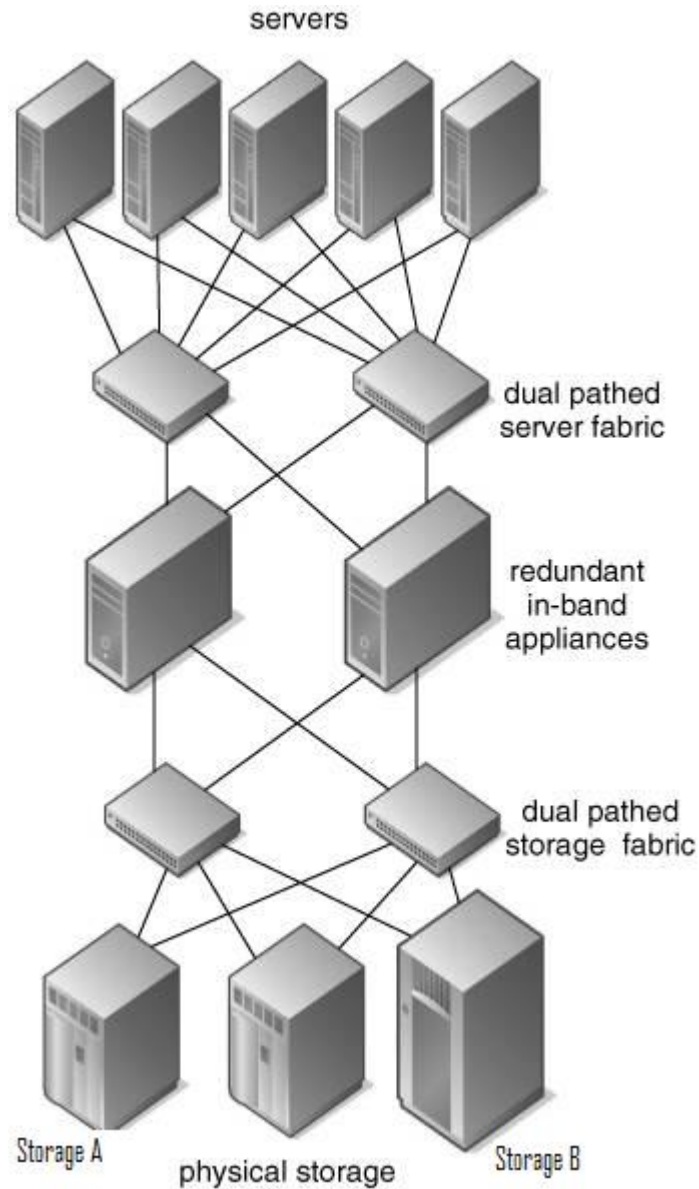


Figura 9: provimento de alta disponibilidade para camada de virtualização de *storage* (*redundant in-band appliances*)

Com essa solução, mesmo que o *Storage A* venha a ficar indisponível, seja qualquer o motivo, os dados continuarão acessíveis ao servidor a partir da camada de virtualização que acessa os dados necessários do *storage B*.

Com estes dois cenários expostos temos soluções aplicáveis a diferentes níveis de RPO (*recovery point objective*-objetivo do ponto de recuperação) e RTO (*recovery time objective*-objetivo do tempo de recuperação) (DOLEWSKI, 2008), para diferentes níveis de investimento, onde o que balizará o investimento será o custo agregado à solução associado ao retorno financeiro suportado pelas aplicações disponibilizadas pela infra-estrutura envolvida em cada solução. Desta

forma, aplicações que não são tão críticas ao negócio (visto o retorno financeiro que provêm) podem ter sua disponibilidade de *storage* mantida por uma solução menos onerosa do que a solução adotada às aplicações mais críticas (visto também o retorno financeiro).

3 - SOLUÇÃO PROPOSTA

De acordo com o contexto apresentado e analisando as requisições existentes no mercado de minimizar investimentos maximizando disponibilidade não se pode deixar de realizar primeiramente uma análise de cada aplicação e associado a cada área de negócios que as utiliza a fim de criar uma matriz em que seja explicitada a real necessidade de alta disponibilidade e em qual nível esta deva ser provida. Conforme a Figura 10, deve ser considerada em que tempo (RTO) e que possíveis perdas de informações (RPO) essas aplicações podem suportar (assim como suas áreas usuárias). (BALOURAS, 2008)

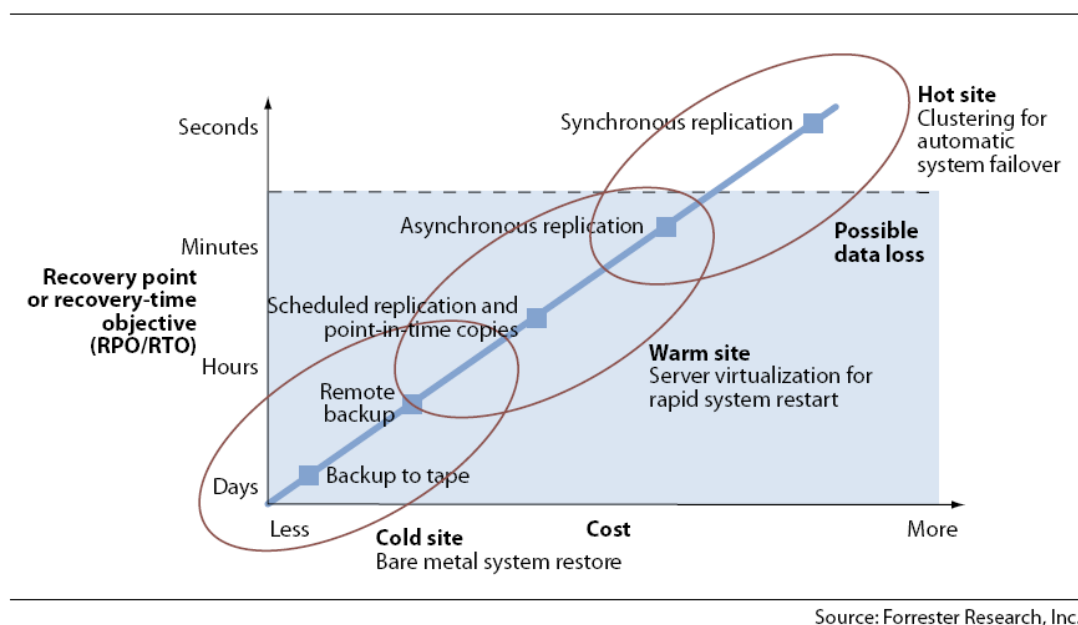


Figura 10: Matriz de recuperação contínua

Com esse gráfico fica evidente que quanto menos informações são perdidas e onde a recuperação é mais rápida (ou inexistente o tempo de *down-time*) mais cara é a solução. Baseado nisso, a primeira coisa a ser feita é determinar que aplicações devam utilizar qual técnica de recuperação ou alta disponibilidade. Como o tema

deste trabalho está focado em alta disponibilidade, será abordada apenas a situação mais crítica, ou seja: RPO e RTO próximos a zero. (DOLEWSKI, 2008)

Como primeiro cenário será proposto um ambiente para a camada de *storage*, de baixo custo, que proverá redundância através de replicação de dados. Para este primeiro cenário a proposta é utilizar como plataforma de *hardware* servidores de baixo custo ou ainda estações de trabalho, porém com bom desempenho principalmente no acesso a discos (atualmente existem soluções com discos de 15000 rpms). Um dimensionamento de memória em torno de 4 Gb de RAM ou mais e uma placa de rede compatível com a solução de rede existente na infra-estrutura deve ser considerada a fim de maximizar o acesso aos recursos do *hardware* evitando gargalos.(SCHIMIDT, 2006)

Para montar este primeiro cenário deverão ser dispostos estes dois servidores com um sistema operacional que disponha de soluções de replicação de dados, assim como uma solução de compartilhamento de arquivos, como CIFS (*common internet file system*), NFS (*network file system*), etc., ou *software* que disponibilize um *target* iSCSI (*internet small computers systems interface*, que é um protocolo de rede que possibilita o envio de comandos SCSI de acesso a dados por redes TCP/IP-*transmission control protocol / internet protocol*) a fim de prover aos servidores de aplicação acesso (compartilhado ou não) ao sistema de arquivos. Os servidores de aplicação poderão estar dispostos sob forma de *clusters* ou de balanceamento de carga (onde o tráfego pode ser balanceado entre mais de um servidor com, por exemplo, *Round Robin DNS* no caso de balanceamento de carga).(GOUVEIA, 2005)

Para prover um melhor tempo de recuperação para servidores de aplicações nos casos de problemas de *hardware* destes a unidade de armazenamento que suportará o sistema operacional (além é claro da unidade de armazenamento que proverá acesso aos dados e a própria aplicação) deverá estar armazenada nesta solução de *storage* e no caso de falha uma simples troca de *hardware* e reconfiguração do iSCSI *initiator* tornará operacional novamente a aplicação e a disponibilidade dos serviços providos por este servidor. A Figura 11, mostra através de um diagrama a arquitetura de *hardware* e a topologia a ser adotada nesta solução de replicação. A camada de Dados abaixo dos servidores de aplicação são na verdade *targets* iSCSI ou compartilhamentos NFS ou CIFS que são providos aos servidores pela camada de servidores de dados (camada inferior).

Na ocorrência de problemas que impeçam o correto funcionamento de um dos servidores de dados (servidores A ou B), que possuem os dados replicados entre si, a única manutenção necessária para disponibilizar novamente os dados para os servidores de aplicação seria a reconfiguração dos *shares* (no caso de utilizar-se CIFS), *exports* (no caso de NFS) ou *initiators* iSCSI (no caso de *targets* iSCSI).

Além disso, adotar uma solução de *backup* que tem como origem os servidores de dados tende a maximizar a utilização dos recursos de *backup* e minimizar as janelas para o mesmo (tempo necessário para o *backup*).

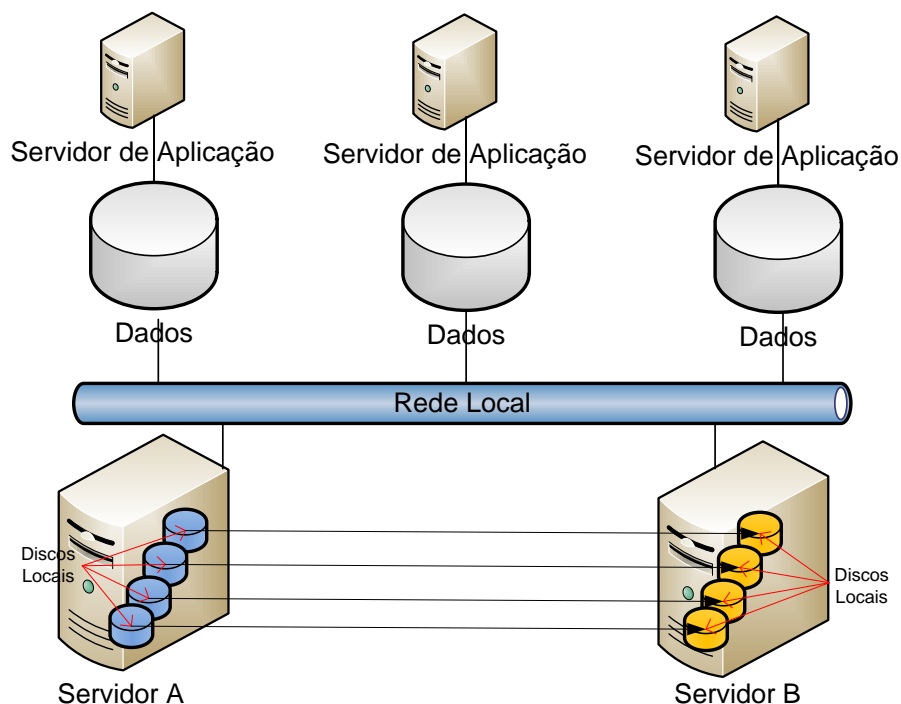


Figura 11: solução de replicação de dados para vários servidores

Como segundo cenário, a fim de atender ininterruptamente o fornecimento de dados a servidores a proposta é a virtualização da camada de acesso a dados, estendendo a topologia anterior com uma camada a mais de *software* e *hardware*. Desta forma os dados não estarão presentes unicamente em um *storage* (servidor ou ainda em um *desktop*) e sim distribuídos em mais de um *hardware*. (WESLEY, 2005)

A solução de virtualização está fundamentada em prover através de serviços de cluster acesso aos *targets* iSCSI ou ainda a compartilhamentos CIFS ou NFS,

dependendo da necessidade, porém a solução baseada em iSCSI tende a atender melhor qualquer cenário, visto a melhoria a nível de tratamento do protocolo SCSI dentro do protocolo de transporte. Na Figura 12 está um diagrama que demonstra como uma infra-estrutura para este fim deverá ser configurada. (WESLEY, 2005)

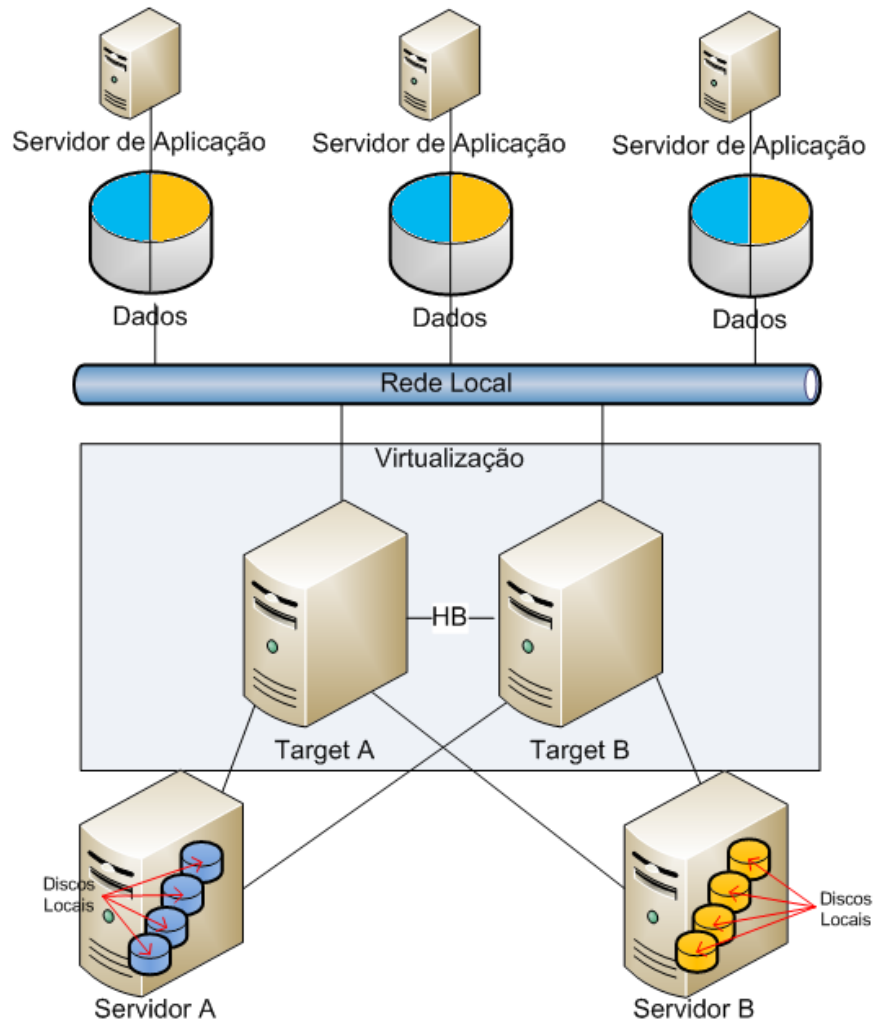


Figura 12: cluster para virtualização de *storage*

Deve-se considerar que, quanto maior a necessidade do negócio frente à tecnologia, melhores *hardwares* deverão ser considerados para suportar esta proposta, a fim de garantir os níveis de disponibilidades exigidos pelo negócio. Para esta análise, indicadores como MTBF (*mean time between failures*) e MTTR (*mean time to repair*) de cada componente deve ser fortemente considerados. (OGGERINO, 2001)

4 - CONCLUSÃO

Diante da proposta apresentada podemos concluir que este trabalho propõe uma solução viável para se prover alta disponibilidade para aplicações de negócios a um custo ajustável, frente às reais necessidades das empresas.

PROPOSTA PARA TCC-II

Para o TCC-II serão realizadas as seguintes atividades:

Cenário um (replicação):

- a) Configuração de dois servidores de dados, virtuais, com sincronização da área de dados dos mesmos
- b) Configuração de um servidor de páginas *web* que se utilizará dos recursos de dados providos pelos servidores de dados

Cenário dois (virtualização de *storage*):

- a) Configuração de dois servidores de dados, virtuais
- b) Configuração de dois servidores, em *cluster*, para prover a camada de virtualização de acesso aos dados providos pelos servidores de dados
- c) Configuração de um servidor de páginas *web* que se utilizará dos recursos de dados providos pela camada de virtualização
- d) Configuração de um servidor de banco de dados que se utilizará dos recursos de dados providos pela camada de virtualização

Para ambos os cenários, estará contemplada as fases de documentação da configuração e plano de testes para disponibilidade em casos de falha dos componentes envolvidos na solução.

5 - BIBLIOGRAFIA

BALAOURAS, Stephanie. ***Building The Business Case For Disaster Recovery Spending, Forrester*** – 03/04/2008. 18p.

Forester Research, Inc. Disaster Recovery Journal. Global Disaster Recovery Preparedness Online Survey. 10/2007

DOLEWSKI, Richard. **System i Disaster Recovery Planning**. Auxilia no mapeamento de informações para apoio a decisão de Planejamento de Recuperação em Desastres. ISBN: 978-1583470671, 2008. 350p.

SCHMIDT, Klaus. **High Availability and Disaster Recovery: Concepts, Design, Implementation**. Alta Disponibilidade e Recuperação de Desastres: conceitos, Design e implementação. ISBN: 978-3540244608, 2006. 410p.

GOUVEIA, José. **Redes de Computadores**. Locais e Wireless. ISBN: 9789727224739, 2005. 312p.

WESLEY, Adilson. **Storage Virtualization: Technologies for Simplifying Data Storage and Management**. ISBN: 0321262514, 2005. 264p.

OGGERINO, Chris. **High Availability Network Fundamentals**. ISBN: 9781587130175, 2001. 250p.