Confiabilidade e disponibilidade do armazenamento de dados

Apresentação

Falar em disponibilidade e confiabilidade sem citar o armazenamento de dados não faz sentido, pois todo sistema armazena dados. Então, este é um item de grande importância em todo sistema que busca uma alta disponibilidade em serviços de armazenamento, tratando da forma e dos locais onde os dados serão armazenados. Informações e dados precisam ser protegidos e tratados de modo a garantir integridade e confiabilidade, visto que podem ser perdidos de inúmeras maneiras, desde danificação de sistemas operacionais e erros humanos até problemas causados por vírus.

Nesta Unidade de Aprendizagem, você vai aprender sobre o armazenamento de dados, abordando técnicas de confiabilidade e tolerância a falhas, focando sempre em alta disponibilidade, buscando reconhecer os riscos de perdas de dados e informações, além de identificar aspectos de redundância e replicação de dados, explicando também o funcionamento e a importância de *backups* e restaurações.

Bons estudos.

Ao final desta Unidade de Aprendizagem, você deve apresentar os seguintes aprendizados:

- Reconhecer os riscos de perdas de dados e informações.
- Identificar aspectos de redundância e replicação de dados.
- Explicar o funcionamento e a importância de *backups* e restaurações.

Infográfico

Cerca de metade dos casos de interrupções em operações de negócios se deve a falhas de *software* ou de rede, podendo danificar dados e levando a paralisação de operações. Essas paralisações podem ser rápidas ou demoradas, dependendo do plano de ação da empresa para retomar as atividades em casos de incidentes. Todas as organizações já passaram por algum problema desse tipo e, normalmente, muitas acabam percebendo apenas nesse momento o quanto suas soluções atuais são frágeis.

Nesse contexto, o *Disaster Recovery as a Service* (DRaaS) surge como uma ferramenta muito importante para impedir que problemas de paralisação de serviços e perda de dados se repitam. O DRaaS consiste na cópia de toda a sua estrutura principal em nuvem, ou seja, a replicação de dados, sistemas e servidores físicos ou virtuais. Isso possibilita uma recuperação rápida em casos de catástrofes de caráter acidental, como, por exemplo, incêndios, rompimento de fibra, falha no provedor de *cloud*, entre outros; ou, ainda, quando ocorrem catástrofes naturais, como inundações, raios, etc.

Veja, no Infográfico a seguir, um pouco mais sobre o DRaaS, seus benefícios e como esse serviço é faturado.

DISASTER RECOVERY AS A SERVICE (DRAAS)

O DRaaS visa a manter a estrutura da empresa em funcionamento, criando sistemas resilientes, que **garantam o uso de servidores, redes e dados durante o processo de recuperação, possibilitando que o funcionamento não seja interrompido de forma alguma**. Dessa forma, a empresa consegue gerenciar melhor seus riscos, reduzindo os prejuízos financeiros que uma paralisação pode causar.

Como é o funcionamento de um DRaaS?

O DRaaS é contratado com base na **quantidade de servidores e espaço em disco consumido**. Sendo assim, o custo financeiro varia de acordo com as necessidades da empresa e só vai ser alterado quando, realmente, for necessário utilizar o serviço.

DRaaS
Disaster Recovery as a Service

Estrutura

Sua estrutura é adaptável, de acordo com a necessidade da empresa. Sendo assim, o valor da mensalidade considera o **espaço necessário para a estrutura, que deve ser igual à principal, visto que os dados são replicados em *real time***.

Caso a empresa opte por não adicionar todos à estrutura, o número de servidores pode depender de sua criticidade. Além disso, também são considerados outros recursos necessários para que a solução de recuperação de desastres fique mais eficiente e transparente.

Exemplo de uso

Se uma empresa tem um *data center* local ou mesmo na nuvem e, por algum motivo, ele ficar inoperante por um tempo, o DRaaS assume seu lugar. Isso acontece porque este funciona com servidores auxiliares, em um *data center* externo, onde fica armazenado o *backup* em tempo real desses servidores.

E, como há consumo de memória, banda e CPU, seu custo financeiro aumenta, **mas só no mês de uso**. Esse valor depende também do tempo em que o DRaaS precisará ficar em operação como se fosse a estrutura principal.

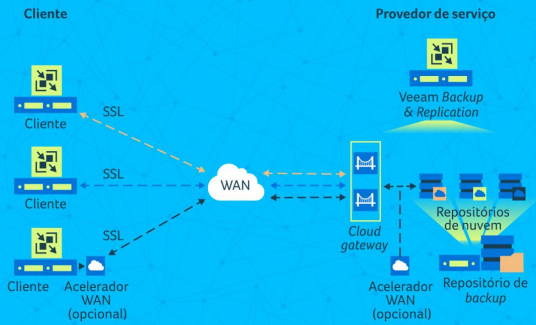
DRaaS # BaaS

O *backup as a service* (BaaS) nada mais é do que a **cópia dos arquivos em dispositivos físicos ou na nuvem**, importante para garantir o acesso e a recuperação quando dados forem danificados ou perdidos. Já o DRaaS **replica sistemas operacionais e de rede, servidores, páginas de Internet** e o que mais for necessário e fizer parte da infraestrutura da empresa, para ficar disponível quando houver necessidade de uso.

Um detalhe importante sobre a diferença entre ambos é que, no *backup*, os **dados precisam ser baixados para serem recuperados**, enquanto, no DRaaS, a **entrega dos dados e estrutura é em *real time***, não necessitando de *downloads* ou novas configurações.

Garantia de continuidade de negócios

Mesmo com benefícios semelhantes aos do BaaS, o DRaaS se diferencia por **manter a empresa sempre em funcionamento**, trocando os servidores principais por servidores de *backup* em um *data center* em nuvem. Tal funcionalidade não pode ser alcançada, muito menos considerando o *real time*, com o uso de *backup* dos arquivos, mesmo que em nuvem.



Aponte a câmera para o código e acesse o link do conteúdo ou clique no código para acessar.

Conteúdo do Livro

Organizações e usuários são dependentes da tecnologia e, com isso, o armazenamento de dados é um dos principais componentes de TI. As empresas armazenam todos os dados, aplicativos e informações referentes a seus negócios em dispositivos de armazenamento. O ponto principal é saber se esses dados estão realmente seguros e se existe um plano que garanta a disponibilidade e a confiabilidade necessária para que não ocorra uma perda inesperada e, muitas vezes, até mesmo irreparável, o que pode ser o fim para uma organização.

No capítulo Confiabilidade e disponibilidade do armazenamento de dados, da obra *Redes de alta disponibilidade*, base teórica desta Unidade de Aprendizagem, você vai estudar os riscos de perdas de dados e informações e aspectos de redundância e replicação de dados, além do funcionamento e importância de *backups* e restaurações.

Boa leitura.

REDES DE ALTA DISPONIBILIDADE

Juliane Adélia Soares

SAGAH
SOLUÇÕES EDUCACIONAIS INTEGRADAS

Confiabilidade e disponibilidade do armazenamento de dados

Objetivos de aprendizagem

Ao final deste texto, você deve apresentar os seguintes aprendizados:

- Reconhecer os riscos de perdas de dados e informações.
- Identificar aspectos de redundância e replicação de dados.
- Explicar o funcionamento e a importância de *backup* e restauração.

Introdução

O armazenamento de dados é muito importante para empresas de todos os portes, seja para salvar informações de clientes, de estoque, fornecedores, funcionários, ou quaisquer outros dados necessários. A perda desses dados pode trazer inúmeras consequências negativas, pois eles são a base para o funcionamento dos negócios.

Neste capítulo, você conhecerá os riscos de perdas de dados e informações. Além disso, conhecerá os aspectos de redundância e replicação de dados, bem como o funcionamento e a importância da realização de *backup* e restauração.

1 Riscos de perdas de dados e informações

A perda de dados é um problema extremamente sério para as organizações de qualquer tamanho e para os usuários finais, uma vez que causa a interrupção de qualquer negócio baseado em informações. Quando arquivos e documentos importantes são perdidos, a empresa deve gastar tempo e recursos para recriar e recuperar esses dados, quando possível, pois, em muitos casos, a perda é irreparável. Essa inconveniência também atrasa os prazos de produtividade e pode causar a perda de clientes.

Um artigo da Consolidated Technologies, Inc. (CTI CONSOLIDATED TECHNOLOGIES, 2018) cita alguns dados referentes às consequências que uma grande perda de dados pode causar às organizações, são eles:

- 94% das empresas que experimentam uma severa perda de dados não se recuperam;
- 51% dessas empresas fecharam em até dois anos após a perda de dados;
- 43% dessas empresas não reabriram novamente;
- 70% das pequenas empresas encerraram as suas atividades dentro de um ano após um grande incidente de perda de dados.

A Figura 1, a seguir, apresenta as estatísticas em relação às principais causas de perdas de dados.

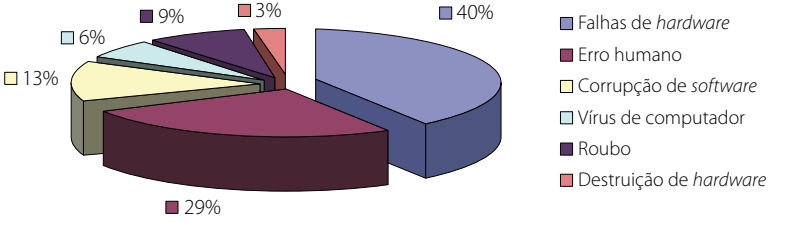


Figura 1. Causas de perdas de dados.
Fonte: Adaptada de Chue, Pereira e Zhao (2019).

Falhas de *hardware* (40%)

Os discos são responsáveis pelo armazenamento de dados, programas, aplicativos e sistemas operacionais, por isso, tornam-se os componentes de *hardware* mais importantes em uma infraestrutura de tecnologia da informação (TI), uma vez que os processos de negócios em TI são totalmente dependentes do armazenamento de dados.

Os discos rígidos (HD) possuem discos móveis, onde os dados ficam armazenados, e um braço mecânico, que realiza a leitura e a escrita. Em virtude dessas partes móveis, eles são os componentes com a menor confiabilidade, e, consequentemente, a maioria dos erros de *hardware* são erros de disco. Os SSDs (*solid-state drive*) também são opções de armazenamento e, como

não possuem partes móveis, são menos sensíveis que os HDs. No entanto, com o tempo, o *floating gate* (i.e., a parte onde os dados são armazenados em um SSD) vai perdendo a sua capacidade de reter cargas, de modo que a vida útil de um SSD tende a ser menor do que a de um HD.

Outros componentes de *hardware*, exceto o HD/SSD, não causam diretamente a perda de dados, porém, em um computador em funcionamento, qualquer dano pode atingir o disco. Por exemplo, se um líquido for acidentalmente derramado no equipamento, ele pode atingir o dispositivo de armazenamento. Portanto, qualquer forma de dano a um sistema de computador pode causar potencialmente a perda de dados.

Erro humano (29%)

É muito importante possuir uma equipe de TI especializada e bem-treinada, porém é importante lembrar que ninguém está livre de cometer erros, não só entre a equipe de TI, mas entre os funcionários no geral. O caso mais comum de erro humano é quando os dados são excluídos acidentalmente. Assim como esses arquivos podem ter ido parar apenas na lixeira do sistema, eles também podem ter se tornado irrecuperáveis.

Outras causas frequentes de erro humano são desligar o sistema antes de salvar os dados e sobrescrever arquivos importantes. Contudo, esses erros podem ser minimizados por meio do uso de *software*, visto que a automação dos sistemas reduz a necessidade de interação humana com os dados e, consequentemente, o risco de erros humanos.

O erro humano também pode desempenhar um papel importante em muitas outras causas de perda de dados, como: danos ao dispositivo de armazenamento, derramamentos de líquidos, corrupção de *software* e formatação do disco.

Corrupção de *software* (13%)

A maioria dos dados armazenados em sistemas de computador está vinculada a um *software* que ajuda a gerar esses dados. Contudo, uma falha no funcionamento desse *software* pode causar a perda de dados. Por exemplo, desligamentos inesperados, que podem ser causados por quedas de energia ou ser impróprios do *software*, podem causar sérios problemas aos dados, visto que podem os corromper ou excluir o seu progresso. Quando o *software* está corrompido, pode não ser possível executá-lo novamente, o que significa que os dados armazenados nele ficarão inacessíveis.

Vírus e *malware* (6%)

Os vírus e *malware* também são uma grande ameaça, pois eles têm a capacidade de destruir, roubar ou criptografar os dados e, em alguns casos, corromper completamente o *hardware* do computador, ou até mesmo se infiltrar em toda a rede. Em geral, esses ataques ocorrem por meio de *e-mails* maliciosos ou *phishing*, na tentativa de fazer os funcionários clicarem em *links* corrompidos para conseguir roubar os seus dados. Por exemplo, existem *phishing* que imitam o *site* inteiro de um banco para enganar os usuários e roubar os seus dados bancários.

Contudo, esse risco pode ser minimizado por meio do uso de *software* de antivírus. Portanto, deve-se manter os sistemas de antivírus atualizados e executar verificações regularmente para detectar a presença de um vírus antes que eles possam causar danos sérios.

Roubo de computadores ou *notebooks* (9%)

Muitas organizações contam com funcionários trabalhando externamente, utilizando *notebooks*, *tablets* ou *smartphones*. O roubo de equipamentos é um risco sério, pois, se o equipamento possuir dados confidenciais, pode ocorrer uma violação de dados. O recomendado é possuir um meio de limpar remotamente os dados desses equipamentos. Além disso, se não for feito *backup* em outros locais, todos os dados armazenados nesse equipamento, sejam eles pessoais ou organizacionais, serão perdidos.

Destruição de *hardware* (3%)

Desastres (p. ex., tornados, incêndios, alagamentos) podem acontecer quando menos se espera e têm impactos devastadores para as empresas, pois podem destruir o negócio como um todo. Como os desastres naturais podem destruir completamente a tecnologia de uma empresa, é crucial fazer *backup* de dados em um local remoto que não possa ser comprometido por um desastre local.

Todos os itens citados são riscos reais que podem ocorrer, e a perda de dados traz inúmeros outros problemas consigo. No entanto, apesar de algumas vezes os problemas serem inevitáveis, se a empresa possuir uma estrutura de *backup* de dados, seja internamente ou na nuvem, as chances de os dados serem perdidos definitivamente serão menores.

Saiba mais

Para conhecer os relatos de cinco casos de perdas de dados que marcaram a história, acesse o *blog* Backup Garantido e pesquise por "Sem backup: 5 casos impressionantes de perda de dados que marcaram a história" para ler o artigo.

2 Redundância e replicação de dados

A redundância refere-se à duplicação de dados ou serviços críticos com a intenção de aumentar a confiabilidade do sistema. Por exemplo, se houver apenas uma cópia de um arquivo armazenado em um servidor e este falhar, esse arquivo poderá ser perdido. Para isso, existe a redundância de dados, que permite criar cópias duplicadas do arquivo para resolver esse problema. Da mesma forma, a redundância pode ser útil para serviços críticos que precisam estar em constante execução, nos quais é necessário garantir que várias cópias ou versões estejam em execução simultaneamente para impedir a falha de um nó único.

A redundância é um método de proteção de dados e pode ser proposta como uma medida de proteção contra falhas em tempo real. O RAID (Redundant Array of Inexpensive Disks; ou Matriz Redundante de Discos Independentes, em português) é um recurso de redundância comum nos servidores, pois ele cria várias cópias de arquivos em diversos discos rígidos, independentes entre si. Se um deles falhar, outro disco assumirá a carga de trabalho.

Segundo Laan (2017), além do RAID, existem outras técnicas de redundância. Por exemplo, para aumentar a disponibilidade em uma SAN (Storage Area Network; ou Rede de Área de Armazenamento, em português), os componentes HBAs (adaptadores de barramento de *host*) e *switches* podem ser instalados de forma redundante. Os HBAs são placas de interface implementadas em servidores. Com a SAN, é possível contar com caminhos redundantes de acesso e espelhamento dos dados em tempo real. Sendo assim, por meio do uso de vários caminhos entre os HBAs e os comutadores SAN, também conhecidos como caminhos múltiplos, o *failover* pode ser instanciado automaticamente quando ocorre uma falha, de modo que os dados estarão sempre disponíveis.

As empresas também podem optar por utilizar vários sistemas de armazenamento, a fim de aumentar ainda mais a redundância, os quais são instalados em diferentes locais para garantir a segurança desses dados. Por meio da replicação, os blocos de discos alterados do sistema de armazenamento primário são enviados continuamente para o sistema de armazenamento secundário, onde eles também são armazenados (Figura 2).

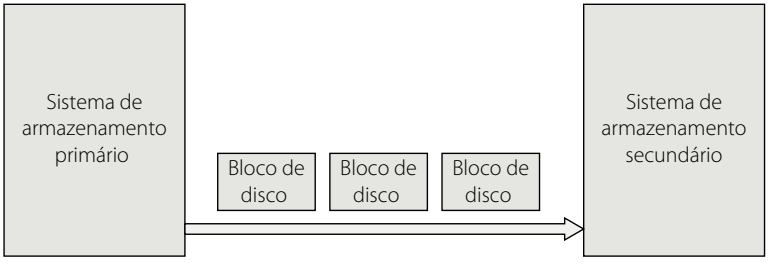


Figura 2. Replicação de armazenamento.
Fonte: Adaptada de Laan (2017).

A replicação de dados consiste em manter continuamente uma cópia secundária dos dados de um volume primário, possivelmente de modo remoto, com o objetivo de fornecer alta disponibilidade e redundância. Ela pode ser executada por meio de redes SAN, LAN (Local Area Network; ou Rede de Área Local, em português) ou WAN (Wide Area Network; ou Rede de Longa Distância, em português).

Em suma, a replicação de dados é um modo de *backup* em que as alterações da fonte de dados primária são refletidas em outro lugar em tempo real, para a replicação síncrona, ou em modo de armazenamento e encaminhamento atrasado, para a replicação assíncrona. O modo utilizado dependerá da distância pela qual a replicação precisa ocorrer. A replicação síncrona geralmente opera a distâncias inferiores a 100 Km, ao passo que a replicação assíncrona é essencialmente ilimitada (CRITCHLEY, 2014).

Replicação síncrona

Na replicação síncrona, além da replicação no sistema de armazenamento secundário, cada gravação no sistema de armazenamento primário deve ser concluída antes da confirmação da gravação no sistema operacional. Um bloco de dados é gravado nos discos de armazenamento primário e enviado pela rede ao disco secundário. Quando os dados são recebidos pelo disco secundário, este envia a confirmação do recebimento dos dados, enviando um pacote de volta pela rede para o disco primário. Em suma, a replicação síncrona garante que os dados nos dois sistemas de armazenamento sejam sincronizados o tempo todo e nunca sejam perdidos.

Com o uso do *cache* de gravação do sistema de armazenamento secundário, as gravações podem ser confirmadas quando os dados são armazenados no *cache*, em vez de armazenados no disco, o que acelera o processo de gravação. Esse tipo de replicação é recomendado para dados críticos e geralmente opera a distâncias inferiores a 100 Km, uma vez que, se o comprimento físico do cabo for superior a essa distância, a latência fica muito alta, diminuindo a velocidade das aplicações, que precisam aguardar a conclusão da gravação no sistema de armazenamento secundário.

Com a replicação síncrona, em caso de falha do local de origem, as aplicações podem fazer *failover* para o local remoto e retomar as operações, com a garantia de que não ocorrerá a perda de dados. As etapas da replicação síncrona (Figura 3) são as seguintes:

- a aplicação (p. ex., aplicação de sistemas bancários) realiza a gravação de dados no armazenamento primário;
- os dados de *logs* são gravados e replicados para o armazenamento secundário. O envio de *logs* envolve a aplicação de um *log* de transações de todas as inserções, atualizações ou exclusões realizadas no armazenamento primário;
- os dados de *logs* são gravados no armazenamento secundário;
- o armazenamento secundário envia um pacote de confirmação de recebimento dos dados;
- o armazenamento secundário confirma a gravação da aplicação;
- (t) e (t^1) são liberados para o volume, e os *logs* sempre realizam a gravação.

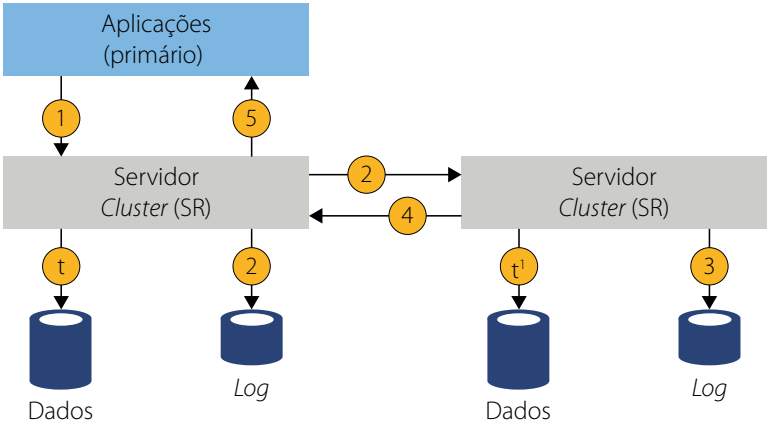


Figura 3. Replicação síncrona.
Fonte: Adaptada de Microsoft (2019).

Replicação assíncrona

Na replicação assíncrona, após os dados serem gravados no armazenamento primário, a gravação é imediatamente confirmada no sistema operacional, sem a necessidade de se esperar que o armazenamento secundário termine as gravações. Esse modelo permite um tempo de resposta mais rápido, bem como uma solução de recuperação de desastre, que funciona geograficamente.

As etapas da replicação assíncrona (Figura 4) são as seguintes:

- a aplicação realiza a gravação dos dados no armazenamento primário;
- os dados de *logs* são gravados;
- a gravação dos dados é confirmada;
- os dados são replicados para o armazenamento secundário;
- os dados dos *logs* são gravados no armazenamento secundário;
- o armazenamento secundário confirma a gravação dos dados;
- (t) e (t^1) são liberados para o volume, e os logs sempre realizam a gravação.

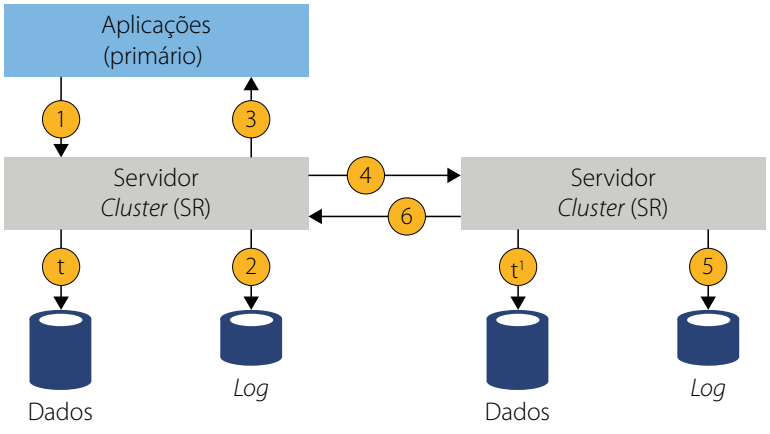


Figura 4. Replicação assíncrona.
Fonte: Adaptada de Microsoft (2019).

Apesar de não possuir o impacto da latência da replicação síncrona, a replicação assíncrona tem a desvantagem de possível perda de dados caso ocorra uma falha no armazenamento primário antes de os dados serem gravados no sistema de armazenamento secundário. Portanto, essa replicação é menos adequada para serviços críticos, como os *clusters* de *failover*.

É importante ressaltar que a réplica de armazenamento não é uma solução de *backup*, pois ela replica todas as alterações realizadas para todos os blocos de dados no volume, independentemente do tipo de alteração. Assim, se os dados forem excluídos no armazenamento primário, eles serão excluídos igualmente no armazenamento secundário. Portanto, a replicação é uma solução de confiabilidade e disponibilidade, pois garante que, caso o servidor principal falhe, a réplica possa assumir sem causar indisponibilidade e perda de dados aos usuários.

3 *Backup* e restauração

Os *backups* são cópias de dados (p. ex., informações, sistemas, servidores) utilizadas para restaurar os dados em caso de perda, corrupção ou situação de desastre. Segundo Marcus e Stern (2003), os *backups* são o coração de qualquer projeto de sistemas crítico, os quais, se tratados adequadamente, representam a última linha de defesa contra praticamente qualquer catástrofe.

Portanto, os *backups* são o último recurso, sendo utilizados apenas se tudo falhar, com o objetivo de salvar a organização em caso de alguma perda drástica de dados. Quando um sistema é bem-projetado, ele deve possuir opções para reparar os dados incorretos de dentro do sistema ou utilizar ferramentas de gerenciamento de sistemas, como ferramentas de banco de dados.

Fique atento

É importante destacar que arquivamento e *backup* são coisas diferentes. O **arquivamento** é o armazenamento de dados a longo prazo, a fim de cumprir leis e regulamentos. Por exemplo, um hospital deve manter o registro e o histórico de seus pacientes por muitos anos, mesmo depois de seu falecimento. Já o ***backup*** é uma solução contra a perda de dados, mas não deve ser mantido por muito tempo, pois, além de ocupar muito espaço de armazenamento como solução para a recuperação de desastre, ele não se torna realmente útil com dados muito antigos. Desse modo, ambos — arquivamento e *backup* — devem ser feitos regularmente. Na maioria das organizações, eles são realizados diariamente, porém, em alguns casos, são feitos a cada hora ou continuamente, em ambientes críticos.

É importante levar em consideração que os dados em discos sincronizados em um *site* de recuperação de desastre não fornecem proteção suficiente, já que, nesse caso, como os dados são sincronizados imediatamente, se os arquivos forem atingidos por um vírus, os arquivos nos discos de recuperação também serão atingidos. Portanto, é recomendado ter várias cópias de *backup* e, pelo menos, uma *off-line*.

Quando os *backups* são realizados em mídias físicas, sempre há riscos de elas serem danificadas. Desse modo, ideal é manter cópias de *backups* em lugares distantes ao local original, para que, em casos de desastres que atinjam a região, os *backups* não sejam destruídos. Além disso, é recomendado possuir *backups* em servidores em nuvem que garantam a disponibilidade, pois, assim, as chances de os dados serem perdidos serão menores.

De acordo com Laan (2017), os *backups* podem ser feitos por meio de vários esquemas, descritos a seguir. Cada um deles descreve quais dados são armazenados em *backup* e quando e como podem se tornar muito complexos em grandes ambientes com muitas aplicações.

Backup completo

Como o nome diz, é um *backup* completo dos dados (Figura 5). Em ambientes grandes, apesar de oferecerem maior proteção, devido à garantia de que todos os dados serão salvos, os *backups* completos são criados apenas em intervalos maiores, como a cada semana ou mês. A criação desses *backups* leva muito tempo, exige grande capacidade de armazenamento e largura de banda, pois é feito o *backup* de todos os arquivos em uma área ou serviço de armazenamento, independentemente das configurações de bits, arquivo morto, data da última alteração, e assim por diante. As cópias podem ser transferidas para servidores, nuvem, fitas magnéticas, discos, entre outros.

Apesar das desvantagens, a restauração de um *backup* completo apresenta maior tolerância a falhas e menor tempo de recuperação, considerando-se que qualquer parte dos dados será localizada em um único conjunto de *backups*.

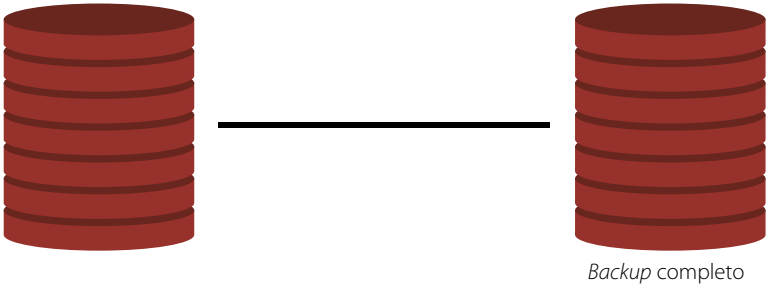


Figura 5. *Backup* completo.
Fonte: Adaptada de Controle Net (2020).

Backup incremental

Os *backups* incrementais (Figura 6) surgiram como complemento para os *backups* completos, pois eles salvam apenas os dados recém-criados ou alterados desde o último *backup*, seja ele completo ou incremental, o que economiza espaço de armazenamento. Contudo, a desvantagem desse esquema é que a sua restauração é demorada, pois ela deve ocorrer a partir do *backup* completo mais

recente e a partir de cada *backup* incremental subsequente. Por exemplo, se for necessária a recuperação dos dados de *backup* de sábado e o último *backup* completo foi realizado no domingo, será necessária a recuperação do *backup* completo de domingo e de todos os *backups* incrementais de segunda-feira até sábado. Entretanto, se um dos arquivos incrementais apresentar problema, toda a restauração estará comprometida.

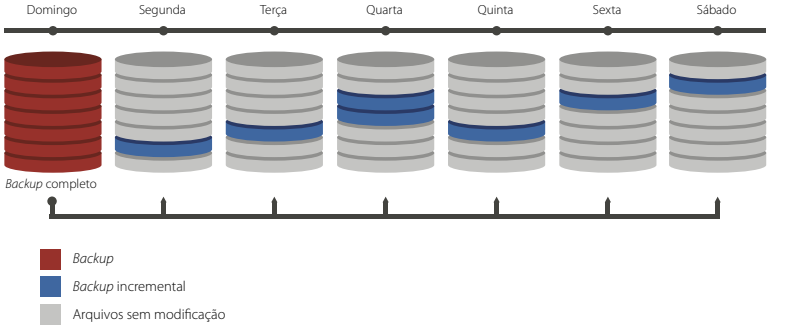


Figura 6. *Backup* incremental.

Fonte: Adaptada de Controle Net (2020).

Backup diferencial

No *backup* diferencial (Figura 7), são salvos os dados recém-criados ou alterados desde o último *backup* completo. A diferença desse *backup* para o incremental é que cada *backup* diferencial armazena todos os dados alterados desde o último *backup* completo, o que utiliza um maior espaço de armazenamento.

A restauração de um *backup* diferencial é bastante eficiente, pois implica apenas a restauração do *backup* completo e do *backup* diferencial mais recente. Contudo, a desvantagem desse esquema é que, dependendo da quantidade de dados que a empresa cria ou altera, cada processo poderá gerar arquivos de *backup* diferenciais cada vez maiores, superando, inclusive, o tamanho do *backup* completo.

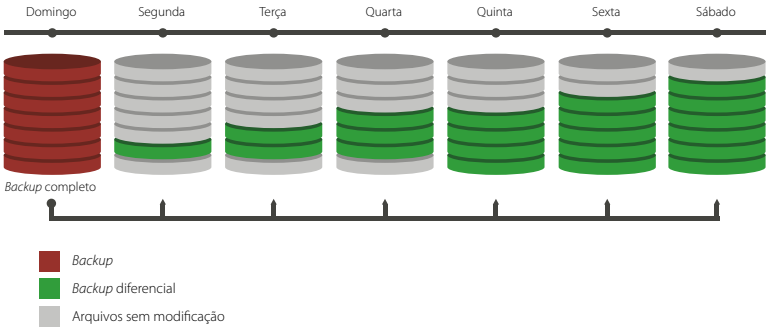


Figura 7. *Backup* diferencial.

Fonte: Adaptada de Controle Net (2020).

Os dispositivos de armazenamento são os componentes de *hardware* mais importantes de uma infraestrutura de TI dentro de uma organização, pois eles são os responsáveis por armazenar todos os dados e informações referentes aos negócios. Se uma empresa perde os seus dados, seja por falhas de *hardware*, desastres, erros humanos, ou demais causas possíveis, isso pode levá-la à falência.

A redundância e a replicação de dados são soluções de confiabilidade e disponibilidade de armazenamento de dados. A replicação garante que todos os dados de um servidor primário sejam replicados em um servidor de armazenamento secundário, para que, caso o primário falhe, o segundo assuma os serviços, garantindo a disponibilidade dos dados aos usuários. Além disso, os *backups* e as restaurações são extremamente importantes e necessários, pois são uma maneira de garantir que, se ocorrer uma perda de dados por exclusão acidental, por exemplo, com a exclusão de dados na réplica também, estes possam ser recuperados através de *backup*.

Portanto, a continuidade dos negócios depende do armazenamento seguro de dados e informações, de modo que o investimento nessas soluções se faz necessário para garantir que uma perda não seja fatal para as organizações.

Referências

CHUE, S.; PEREIRA, D. M. O.; ZHAO, D. *Survey research paper:* data loss restoration. 2019. Disponível em: https://medium.com/@danielmop/survey-research-paper-data-loss--restoration-4fae9fa38685. Acesso em: 06 jul. 2020.

CONTROLE NET. *Tipos de backup:* completo ou full, incremental e diferencial. [2020]. Disponível em: https://www.controle.net/faq/tipos-de-backup-o-que-e-backup-full--incremental-e-diferencial. Acesso em: 06 jul. 2020.

CRITCHLEY, T. *High availability IT services*. Boca Raton: CRC, 2014.

CTI CONSOLIDATED TECHNOLOGIES. *10 common causes of data loss.* 2018. Disponível em: https://consoltech.com/blog/10-common-causes-of-data-loss/. Acesso em: 06 jul. 2002.

LAAN, S. *IT infrastructure architecture:* infrastructure building blocks and concepts. 3rd ed. Morrisville: Lulu, 2017.

MARCUS, E.; STERN, H. *Blueprints for high availability.* 2nd ed. Indianapolis: John Wiley & Sons, 2003.

MICROSOFT. *Visão geral da réplica de armazenamento.* 2019. Disponível em: https://docs.microsoft.com/pt-br/windows-server/storage/storage-replica/storage-replica--overview. Acesso em: 06 jul. 2020.

Leitura recomendada

BACKUP EVERYTHING. *What are the causes of data loss?* [2020]. Disponível em: https://backupeverything.co.uk/what-are-the-causes-of-data-loss/. Acesso em: 06 jul. 2020.

Fique atento

Os *links* para *sites* da *web* fornecidos neste capítulo foram todos testados, e seu funcionamento foi comprovado no momento da publicação do material. No entanto, a rede é extremamente dinâmica; suas páginas estão constantemente mudando de local e conteúdo. Assim, os editores declaram não ter qualquer responsabilidade sobre qualidade, precisão ou integralidade das informações referidas em tais *links*.

Encerra aqui o trecho do livro disponibilizado para esta Unidade de Aprendizagem. Na Biblioteca Virtual da Instituição, você encontra a obra na íntegra.

Conteúdo:

SAGAH
SOLUÇÕES EDUCACIONAIS INTEGRADAS

Dica do Professor

O *backup* em nuvem está sendo cada vez mais utilizado, por proporcionar maior confiabilidade em relação aos demais. No entanto, qualquer tipo de sistema pode falhar, inclusive os de nuvem. Sendo assim, o mais recomendado é investir em um procedimento de *backup* em diferentes locais, potencializando a confiabilidade, a disponibilidade e a segurança.

Empresas que oferecem esse serviço no formato de SaaS podem não prever a restauração ou, então, cobrar altas taxas para incluí-la em contrato, fazendo com que surja uma nova tendência de *backup*, o *backup cloud-to-cloud* (em português, *backup* de nuvem para nuvem).

Descubra, nesta Dica do Professor, um pouco mais sobre essa nova tendência, suas vantagens, os pontos de atenção e as diferenças entre o *backup cloud-to-cloud* e o *backup* tradicional.

Aponte a câmera para o código e acesse o link do conteúdo ou clique no código para acessar.

Exercícios

1) **A perda de dados é um grande problema para as empresas, pois causa interrupção em qualquer negócio e, em alguns casos, pode ser irreparável.**

Levando em consideração as formas mais comuns de perda de dados, escolha a opção correta.

A) O erro humano é considerado a forma mais comum de perda de dados.

B) A corrupção de *software* é o segundo erro mais comum de perda de dados.

C) Vírus e *malware* estão no topo, considerando os motivos de perda de dados.

D) O roubo de computadores e demais dispositivos de armazenamento de dados está em terceiro lugar de motivos de perdas de dados.

E) As falhas de *hardware* são consideradas a forma mais comum de perda de dados, estando definitivamente no topo da pirâmide.

2) **A redundância de dados nada mais é do que a duplicação de dados ou de serviços críticos para a empresa, sendo extremamente importante para aumentar a confiabilidade do sistema oferecido. Para um serviço ou sistema que busca disponibilidade e confiabilidade, não é recomendado ter apenas um único local de armazenamento, pois corre um sério risco de ficar sem sistema a qualquer momento.**

Considerando a redundância de dados, escolha a opção correta.

A) A redundância é um método de proteção de dados e pode ser proposta como medida de proteção contra falhas em tempo real.

B) Para aumentar a redundância, há também a opção de utilizar sistemas de armazenamento instalados em um mesmo local.

C) A redundância é um método de proteção de dados e pode ser proposta como medida de proteção contra falhas, mas não é realizada em tempo real.

D) O RAID não é um recurso de redundância, mas um recurso de *backup* comum em servidores. Ele cria uma cópia de si mesmo e a deixa guardada.

E) A redundância não é um método de proteção de dados e não pode ser usada como medida de proteção contra falhas.

3) **A replicação de dados consiste em manter, de forma contínua, uma cópia secundária dos dados de um volume primário para oferecer alta disponibilidade e redundância. A replicação pode ser executada por meio de redes SAN, LAN ou WAN. É uma forma de *backup* e seu modo vai depender da distância pela qual a replicação precisa ocorrer.**

Com base nessas considerações sobre replicação de dados, marque a opção correta.

A) Na replicação assíncrona, um bloco de dados é gravado nos discos de armazenamento primário e enviado pela rede ao secundário.

B) Na replicação assíncrona, depois que os dados são gravados no armazenamento primário, a gravação é imediatamente confirmada no sistema operacional.

C) A replicação assíncrona é recomendada para dados críticos e, por isso, geralmente opera a distâncias inferiores a 100 km.

D) A replicação assíncrona garante que os dados nos dois sistemas de armazenamento sejam sincronizados o tempo todo.

E) A replicação assíncrona tem o impacto da latência da replicação síncrona, porém a desvantagem de possível perda de dados.

4) **Os *backups* consistem em cópias de dados, informações, sistemas e servidores que são usadas para restaurar dados em caso de perda, corrupção de arquivos ou sistemas e situações de desastre. São considerados como o coração de sistemas críticos e, se tratados adequadamente, representam a última linha de defesa contra praticamente qualquer catástrofe.**

Considerando o *backup*, marque a opção correta.

A) O *backup* é uma solução contra a perda de dados e deve ser mantido por tempo indeterminado.

B) O arquivamento é o armazenamento de dados a curto prazo e só ocorre devido a algumas políticas de empresas que necessitam desse armazenamento.

C) O *backup* é uma solução contra a perda de dados e não deve ser mantido por muito tempo.

D) *Backup* é o primeiro recurso usado para salvar a organização quando o sistema falha e há alguma perda drástica de dados.

E) Não é necessário manter cópias de *backups* em lugares distantes do local original, pois, normalmente, não há riscos de destruição.

5) **Dados armazenados em discos sincronizados em um *site* de recuperação de desastre não fornecem proteção suficiente, já que, nesse caso, como os dados são sincronizados imediatamente, há algumas complicações.**

Considerando esse contexto, escolha a opção correta.

A) Se os arquivos forem atingidos por um vírus, não é necessário ter várias cópias de *backup*, muito menos uma *off-line*.

B) Se os arquivos forem atingidos por um *ransomware*, os arquivos nos discos de recuperação não serão atingidos, bastando habilitar os dados do *backup*.

C) Se os arquivos forem atingidos por um vírus, é só habilitar o *backup* para funcionar, pois eles não serão atingidos.

D) Se os arquivos forem atingidos por um vírus, os arquivos nos discos de recuperação também serão atingidos.

E) Se os arquivos forem atingidos por um vírus, é recomendado ter outras cópias de *backup*, mas não uma *off-line*.

Na prática

Um dos maiores desafios para as empresas é definir onde seus dados serão armazenados, pois entre os fatores primordiais em suas buscas estão a confiabilidade e a alta disponibilidade. Uma das vantagens é que, no mercado, há vários produtos e serviços para esse fim, com valores e tecnologias diversificados. No entanto, para a escolha da melhor solução, são necessários conhecimentos técnicos e das necessidades da empresa, considerando os benefícios desejados e o orçamento estabelecido.

Acompanhe, Na Prática, as necessidades da Bem Atendimento quanto a armazenamento e disponibilidade de dados. Por ser da área de prestação de serviços de limpeza, essa empresa necessita da disponibilidade de seus dados o maior tempo possível e de forma confiável, para se manter competitiva. Veja a escolha da melhor tecnologia para a Bem Atendimento, baseando-se nos requisitos levantados e nos demais benefícios contemplados pela solução escolhida, o Storage NAS (*Network Attached Storage*).

Conteúdo interativo disponível na plataforma de ensino!

Saiba mais

Para ampliar o seu conhecimento a respeito desse assunto, veja abaixo as sugestões do professor:

Como fazer *backup* em nuvem: boas práticas

Acesse este conteúdo para saber um pouco mais sobre boas práticas de *backup* em nuvem, que também podem ser utilizadas para qualquer outro tipo de *backup*. São práticas essenciais para manter seus dados protegidos, caso necessite usá-los em algum momento, garantindo a disponibilidade e a confiabilidade de seus sistemas.

Aponte a câmera para o código e acesse o link do conteúdo ou clique no código para acessar.

Backup em nuvem – vantagens e importância

Neste artigo, você verá que a manutenção de *backup* local, normalmente, é mais sensível a diversas formas de perdas de dados a qualquer momento. Além disso, também pode se desgastar com o tempo, bem como depender de mais espaço para armazenamento e mais trabalho para restauração. Portanto, nesse contexto, entram a importância e as vantagens do *backup* em nuvem.

Aponte a câmera para o código e acesse o link do conteúdo ou clique no código para acessar.

Entenda a diferença entre armazenamento e *backup* na nuvem

Você conhece a diferença entre dados armazenados em nuvem e *backup* em nuvem? Essas funcionalidades são bem distintas entre si, já que, no armazenamento, você mantém seus dados na nuvem, mas não realiza cópia deles, ao contrário do que acontece no *backup* em nuvem.

Aponte a câmera para o código e acesse o link do conteúdo ou clique no código para acessar.