Backup e Restauração

Apresentação

O backup nada mais é do que a cópia de segurança das informações armazenadas. É um item fundamental na tecnologia da informação, independentemente do nicho. Deve ser aplicado tanto para um usuário comum quanto para uma corporação, seja qual for seu porte. A fundamentação do backup, basicamente, é a restauração dos dados em casos de problemas, bugs ou danificação da informação original. Nesta Unidade de Aprendizagem você vai conhecer os processos de backup e recuperação.

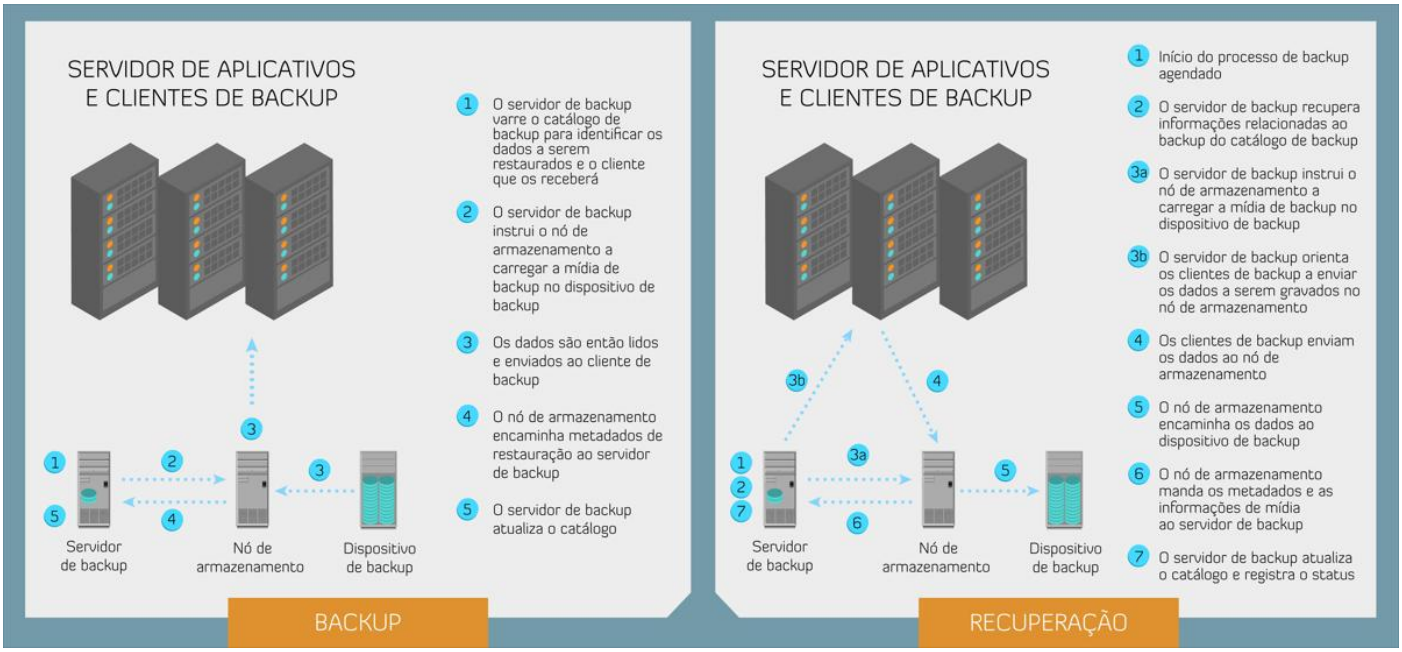
Bons estudos.

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

- Selecionar a melhor alternativa para backups em organizações.
- Identificar as topologias de backup.
- Reconhecer a importância do backup e restore de informações.

Infográfico

No infográfico você vai conhecer o processo de backup e restauração de informações.



Conteúdo do Livro

Os dados das empresas são de vital importância para elas. Os backups fornecem proteção contra perda de dados, bem como facilitam os testes em estruturas maiores.

Acompanhe um trecho do capítulo Backup e Recuperação, da obra *Armazenamento e gerenciamento de informações: como armazenar, gerenciar e proteger informações digitais*, que aborda o estudo do backup e da restauração.

EMC Education Services

Armazenamento e Gerenciamento de Informações

Como armazenar, gerenciar e proteger informações digitais

EMC²
where information lives®

S693a Somasundaram, G.
Armazenamento e gerenciamento de informações : como armazenar, gerenciar e proteger informações digitais / G. Somasundaram, Alok Shrivastava, EMC Education Services ; tradução: Acauan Pereira Fernandes ; revisão técnica: EMC Brasil. – Porto Alegre : Bookman, 2011.
480 p. ; 25 cm.

ISBN 978-85-7780-750-5

1. Computação. 2. Armazenamento e gerenciamento de informação digital. I. Shrivastava, Alok. II. EMC Education Services. III. Título.

CDU 004.658

Catalogação na publicação: Ana Paula M. Magnus – CRB-10/Prov-009/10

Capítulo 12

Backup e Recuperação

Backup é a cópia dos dados de produção criada e mantida com o único propósito de recuperar dados apagados e corrompidos. Com o crescimento dos negócios e demandas normativas de armazenamento, preservação e disponibilidade de dados, as organizações se deparam com a tarefa de fazer backup de um volume cada vez maior de dados. Essa tarefa se torna mais desafiadora à medida que a demanda por backups consistentes e restauração rápida aumenta em toda a empresa – o que pode se espalhar por vários locais. Além disso, as organizações precisam executar backups a um custo mais baixo e com recursos mínimos.

CONCEITOS-CHAVE

- Backup operacional
- Arquivamento
- Período de retenção
- Recuperação Bare-Metal
- Arquitetura de backup
- Topologias de backup
- Biblioteca de fita virtual

As organizações devem assegurar que os dados certos estejam nos lugares certos no tempo certo. Avaliar tecnologias de backup, assim como requisitos de recuperação e de manutenção de dados e aplicativos, é um passo essencial para garantir a implementação bem-sucedida da solução de backup e recuperação. Esse recurso deve facilitar a recuperação fácil dos backups e arquivos conforme necessitado pela empresa.

Este capítulo contém detalhes sobre os objetivos dos backups, estratégias para operações de backup e recuperação, métodos de backup, a arquitetura de backup e a mídia de backup.

12.1 Objetivos do backup

Backups são feitos por três motivos: recuperação de desastres, backup operacional e arquivamento.

12.1.1 Recuperação de desastres

Backups podem ser executados para abordar necessidades de recuperação de desastres. As cópias de backup são usadas para restaurar dados em um local alternativo quando o local principal está incapacitado em virtude de um desastre. Com base nos requisitos de RPO e RTO, as organizações utilizam diferentes estratégias de backup para recuperação de desastres. Quando um método de backup baseado em fitas é explorado como estratégia de recuperação de desastres, a mídia de fita do backup é enviada e armazenada em outro local. Essas fitas podem ser trazidas para restauração no local em que ocorrerá a recuperação. Organizações com requisitos rígidos de RPO e RTO dispõem de tecnologia de replicação remota para replicar dados em um local de recuperação de desastres. Isso permite a elas trazer sistemas de produção de volta on-line em um período de tempo relativamente curto no caso de um desastre. A replicação remota é examinada no Capítulo 14.

12.1.2 Backup operacional

Dados no ambiente de produção mudam a cada transação e operação comercial. Um *backup operacional* é um backup de dados em determinado momento e é usado para restaurá-los no caso de perda de dados ou corrupções lógicas que podem ocorrer durante o processamento de rotina. A maioria das solicitações de restauraçãoem muitas organizações cai nesta categoria. Por exemplo, é comum que um usuário apague acidentalmente um e-mail importante ou que um arquivo seja corrompido, o que pode ser restaurado a partir do backup operacional.

Backups operacionais são criados para as informações ativas de produção a partir de técnicas de backup incrementais ou diferenciais, detalhadas posteriormente neste capítulo. Um exemplo de backup operacional é um executado para um banco de dados de produção antes de uma grande atualização em lote. Isso garante a disponibilidade de uma cópia limpa do banco de dados de produção se a atualização em lote o corromper.

12.1.3 Arquivamento

Backups também são executados para tratar de necessidades de arquivamento. Embora o CAS tenha surgido como solução principal para arquivamento, os backups tradicionais ainda são usados por empresas de pequeno e mé-

dio porte para preservar a longo prazo registros de transações, mensagens de e-mail e outros lançamentos empresariais necessários para a observação de regulamentações.

Além de abordar a recuperação de desastres, arquivamento e requisitos operacionais, os backups servem como proteção contra a perda de dados devido a danos físicos no dispositivo de armazenamento, falhas de software ou ataques de vírus. Os backups também podem ser usados para proteção contra acidentes, como uma eliminação ou destruição intencional de dados.

12.2 Considerações sobre backup

O montante de perda de dados e o tempo inativo que uma empresa pode suportar em termos de RTO e RPO são as preocupações primárias na seleção e implementação de uma estratégia específica de backup. Outro fator é o período de retenção, que define por quanto tempo uma empresa precisa manter os backups. Alguns dados são preservados por anos e outros por apenas alguns dias. Por exemplo, dados gravados para arquivamento são conservados por mais tempo do que os gravados para recuperação operacional.

Também é importante considerar o tipo da mídia de backup com base no período de conservação e acessibilidade dos dados. As organizações também devem levar em conta a granularidade dos backups, explicada posteriormente neste capítulo. O desenvolvimento de uma estratégia de backup deve incluir uma decisão sobre o horário mais apropriado para a execução do backup, a fim de minimizar qualquer interrupção nas operações de produção. De forma semelhante, o local e o momento da operação de restauração devem ser considerados, junto com as características dos arquivos e a compactação de dados, que influenciam o processo de backup.

O local, o tamanho e o número de arquivos também devem ser considerados, já que podem igualmente afetar o processo de backup. O local é um fator importante para os dados a serem gravados. Muitas organizações têm dezenas de plataformas heterogêneas que suportam soluções complexas. Considere um ambiente de data warehouse que use dados de backup de muitas fontes. O processo de backup deve abordar essas fontes em termos de integridade transacional e de conteúdo. Esse processo deve ser coordenado em todas as plataformas heterogêneas nas quais os dados estão localizados.

O tamanho dos arquivos também afeta o processo de backup. Fazer backup de arquivos de tamanho grande (como dez arquivos de 1 MB) pode precisar de menos recursos de sistema do que um volume de dados igual em um grande número de arquivos pequenos (dez mil arquivos de 1 KB, por exemplo). A operação de backup e recuperação demora mais quando um sistema de arquivos contém muitos arquivos pequenos.

Assim como o tamanho, o número de arquivos a serem gravados também influencia o processo de backup. Por exemplo, no backup incremental, um sistema de arquivos que contém um milhão de arquivos com uma taxa de alteração diária de 10% terá de criar 100.000 entradas no catálogo de backup, que abarca a tabela de conteúdo do conjunto de dados gravados e informações sobre a sessão de backup. Esse grande número de entradas no sistema de arquivos afeta o desempenho do processo de backup e restauração porque é muito demorado pesquisar em todo o sistema de arquivos.

O desempenho do backup também depende da mídia usada. A lenta operação de iniciar e parar em sistemas baseados em fita determina o desempenho do backup, especialmente quando se fazem backups de um grande número de arquivos pequenos.

A compactação de dados é amplamente usada em sistemas de backup porque economiza espaço na mídia. Muitos dispositivos de backup, como drives de fita, têm suporte interno para a compactação de dados baseada em hardware. Para usá-la eficientemente, é importante entender as características dos dados. Alguns, como binários de aplicativos, não se compactam bem. Dados de texto compactam-se bem, enquanto outros, como arquivos JPEG e ZIP, já são compactados.

12.3 Granularidade do backup

A granularidade do backup depende das necessidades da empresa e dos RTO/RPO requeridos. Com base na granularidade, os backups podem ser classificados como completos, cumulativos e incrementais. A maioria das organizações utiliza uma combinação desses três tipos de backup para satisfazer seus requisitos de backup e recuperação. A Figura 12-1 mostra as categorias de granularidade de backup.

O *backup completo* é feito sobre os dados integrais nos volumes de produção em determinado momento. Um backup completo é criado a partir da cópia dos dados dos volumes de produção para um dispositivo de armazenamento secundário. *Backups incrementais* duplicam os dados que foram alterados desde o último backup completo ou incremental, o que tiver ocorrido mais recentemente. Esse processo é muito mais rápido (porque o volume de dados gravados é restrito aos dados alterados), mas demora mais tempo para restaurar. O *backup cumulativo* copia os dados que foram alterados desde o último backup completo. Este método demora mais do que o backup incremental, mas é mais rápido de restaurar.

O *backup completo sintético* é outro tipo de backup utilizado em implementações nas quais os recursos dos volumes de produção não podem ser reservados exclusivamente a um processo de backup por períodos extensos para executar um backup completo. Geralmente é criado a partir do backup completo mais recente e todos os backups incrementais executados após esse backup completo. Um backup completo sintético permite que um backup completo seja criado off-line, sem interromper a operação de I/O no volume de produção. Isso também libera do processo de backup os recursos da rede, disponibilizando-os para outros usos de produção.

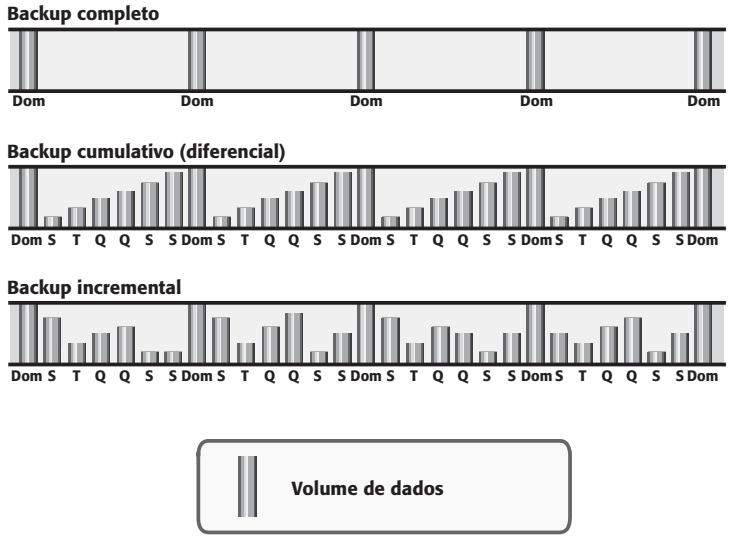


Figura 12-1 Níveis de granularidade de backup.

Operações de restauração variam com a granularidade do backup. Um backup completo fornece um repositório único a partir do qual os dados podem ser facilmente restaurados. O processo de restauração por meio de um backup incremental requer o último backup completo e todos os backups incrementais disponíveis até o ponto de restauração. Uma restauração com base em um backup cumulativo necessita do último backup completo e do backup cumulativo mais recente. A Figura 12-2 ilustra um exemplo de um backup incremental e a restauração.

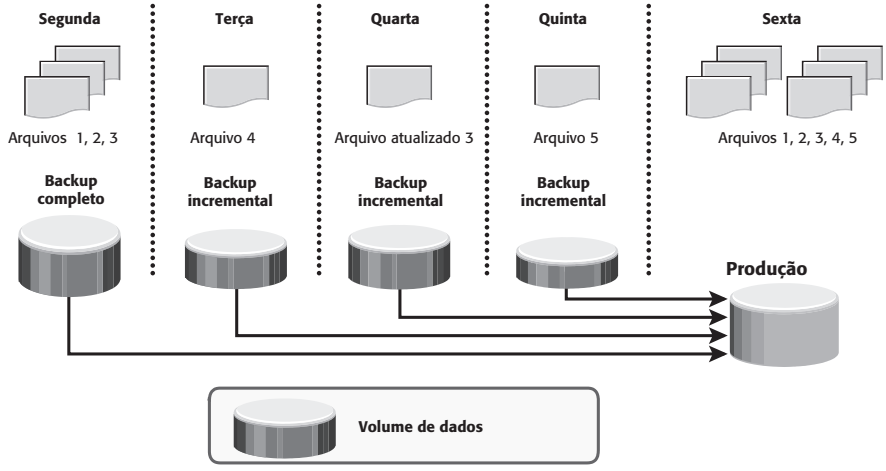


Figura 12-2 Restauração a partir de um backup incremental.

Neste exemplo, um backup completo é executado na noite de segunda-feira. Depois disso, a cada dia um backup incremental é executado. Na terça, um novo arquivo (o Arquivo 4 na figura) é adicionado e nenhum outro arquivo foi alterado. Consequentemente, apenas o Arquivo 4 é copiado durante o backup incremental na terça à noite. Na quarta-feira, nenhum arquivo novo foi adicionado, mas o Arquivo 3 foi modificado. Portanto, apenas o Arquivo 3 modificado é copiado durante o backup incremental na noite de quarta. De forma semelhante, o backup incremental na quinta reproduz apenas o Arquivo 5. Na manhã de sexta-feira, há uma corrupção de dados, o que requer sua restauração a partir do backup. O primeiro passo na direção dessa restauração é restaurar todos os dados do backup completo da noite da segunda-feira. O próximo passo é aplicar os backups incrementais da terça, quarta e quinta-feira. Desta forma, os dados podem ser restaurados com sucesso, voltando a seu estado anterior, conforme existiam na noite da quinta-feira. A Figura 12-3 ilustra um exemplo de backup cumulativo e sua restauração.

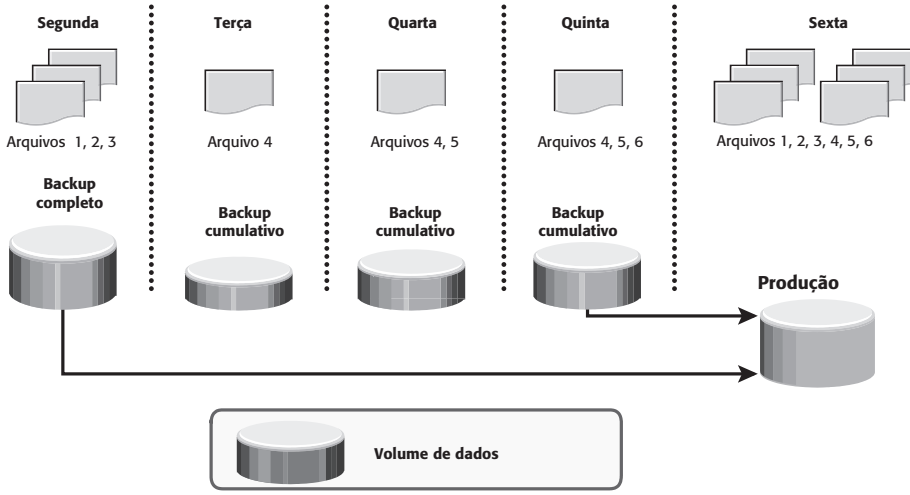


Figura 12-3 Restauração a partir de um backup cumulativo.

Neste exemplo, um backup completo dos dados da empresa é feito na noite de segunda-feira. Depois disso, a cada dia é feito um backup cumulativo. Na terça, o Arquivo 4 é adicionado e nenhum outro arquivo é modificado desde o último backup completo da noite de segunda-feira. Consequentemente, o backup cumulativo da noite de terça-feira copia apenas o Arquivo 4. Na quarta, o Arquivo 5 é adicionado. O backup cumulativo que ocorre na noite da quarta-feira copia o Arquivo 4 e o 5 porque eles foram adicionados ou modificados desde o último backup completo. De forma semelhante, na quinta, o Arquivo 6 é adicionado. Portanto, o backup cumulativo na noite da quinta-feira copia todos os três arquivos: 4, 5 e 6.

Na sexta-feira pela manhã, acontece uma corrupção nos dados que requer sua restauração usando cópias do backup. O primeiro passo na restauração dos dados de um backup cumulativo é restabelecer todos os dados do backup completo da noite de segunda-feira. O próximo passo é aplicar apenas o último backup cumulativo – o de quinta-feira à noite. Desta forma, o volume de dados de produção pode ter seu estado anterior restaurado, o da noite da quinta-feira.

12.4 Considerações sobre recuperação

O RPO e o RTO são considerações importantes ao se planejar uma estratégia de backup. O RPO define o limite tolerável de perda de dados para uma empresa e especifica o intervalo de tempo entre dois backups. Em outras palavras, o RPO determina a frequência dos backups. Por exemplo, se o aplicativo A requerer um RPO de um dia, será necessário que os dados sejam gravados pelo menos uma vez ao dia.

O período de retenção de um backup também é derivado de um RPO especificado para recuperação operacional. Por exemplo, os usuários de um aplicativo A podem requerer a restauração dos dados desse aplicativo a partir da sua cópia de backup operacional, que foi criada há um mês. Isso determina o período de retenção do backup. O RPO do aplicativo A pode, portanto, variar de um dia a um mês, conforme as necessidades operacionais de recuperação.

Entretanto, a empresa pode escolher reter o backup por um período mais longo em virtude de políticas internas ou fatores externos, como políticas regulatórias. Todavia, esse procedimento requer um espaço grande de armazenamento, o que se traduz em um custo maior. Por consequência, é importante definir o período de retenção com base em uma análise de todas as solicitações de restauração no passado e no orçamento alocado. Se períodos curtos de retenção de backup forem especificados, talvez não seja possível recuperar todos os dados necessários para o ponto de recuperação exigido, na medida em que os dados podem ser mais antigos que o período de retenção. Períodos de retenção longos podem ser definidos para todos os backups, a fim de alcançar qualquer RPO no período de retenção estabelecido.

O RTO se relaciona ao tempo gasto pelo processo de recuperação. Para satisfazer a um RTO definido, a empresa talvez queira usar uma combinação de diferentes soluções de backup para minimizar o tempo de recuperação. Em um ambiente de backup, o RTO influencia o tipo de mídia de backup que deve ser empregado. Por exemplo, a recuperação de fluxos de dados multiplexados em fitas demora mais do que uma recuperação a partir de fitas sem multiplexação.

Devido a restrições de recuperação, as organizações executam mais backups completos do que realmente precisam. Backups cumulativos e incrementais dependem de um backup completo prévio. Ao restaurar a partir de mídia de fita, diversas fitas são necessárias para recuperar integralmente o sistema. Com um backup completo, a recuperação pode ser feita com um RTO mais baixo e em menos passos.

12.5 Métodos de backup

Backup dinâmico e backup estático são os dois métodos implantados para backup. Eles são baseados no estado do aplicativo quando o backup é executado. Em um *backup dinâmico*, o aplicativo está em execução, sendo que os usuários estão acessando seus dados durante o processo. Em um *backup estático*, o aplicativo não fica ativo no decurso de backup.

O backup de *dados de produção* on-line se torna mais desafiador porque os dados estão sendo usados e alterados ativamente. Um arquivo aberto é bloqueado pelo sistema operacional e não é copiado durante o processo de backup até que o usuário o feche. O aplicativo de backup pode gravar arquivos abertos tentando executar a operação novamente sobre arquivos que estavam abertos anteriormente no processo de backup. Durante esse procedimento, pode ser que os arquivos abertos anteriormente estejam fechados, de modo que uma nova tentativa será bem-sucedida. Contudo, este método não é considerado robusto, pois em alguns ambientes certos arquivos estão sempre abertos.

Em tais situações, o aplicativo de backup fornece *agentes de arquivos abertos*, que interagem diretamente com o sistema operacional e permitem a criação de cópias consistentes de arquivos abertos. Em alguns ambientes, o uso de agentes de arquivos abertos não é suficiente. Por exemplo, um banco de dados é composto de muitos arquivos de diversos tamanhos, ocupando vários sistemas de arquivos. Para assegurar um backup consistente do banco de dados, todos os arquivos precisam ser gravados no mesmo estado. Isso não significa necessariamente que todos os arquivos precisam ser gravados ao mesmo tempo, mas todos devem estar sincronizados de modo que o banco de dados possa ser restaurado com consistência.

Backups consistentes de bancos de dados também podem ser feitos a partir de um backup estático. Isso requer que o banco de dados permaneça inativo durante o backup. É claro que a desvantagem de um backup estático é que o banco de dados fica inacessível para os usuários durante o processo de backup.

O backup dinâmico é usado em situações nas quais não é possível desligar o banco de dados. Isso é facilitado pelos *agentes de backup de bancos de dados*, que podem executar um backup enquanto o banco de dados está ativo. A desvantagem associada a um backup dinâmico é que os agentes geralmente afetam o desempenho geral do aplicativo.

Um método de cópia *point-in-time (PIT)* é implantado em ambientes nos quais o impacto do tempo inativo causado por um backup estático ou o desempenho resultante de um backup dinâmico sejam inaceitáveis. Uma cópia PIT baseada em ponteiro consome apenas uma fração do espaço de armazenamento e pode ser criada muito rapidamente. Tal cópia é implementada em uma solução baseada em disco na qual uma LUN virtual é criada e armazena ponteiros para os dados armazenados na LUN de produção ou local de gravação. Nesse método de backup, o banco de dados é parado ou congelado momentaneamente enquanto a cópia PIT é criada. A cópia PIT é então montada em um servidor secundário e o backup ocorre no servidor primário. Essa técnica é detalhada no Capítulo 13.

Para assegurar consistência, não basta gravar dados de produção para a recuperação. Certos atributos e propriedades vinculados a um arquivo, como per-

missões, proprietário e outros metadados, também precisam ser gravados. Esses atributos são tão importantes quanto os próprios dados e devem ser gravados por motivo de consistência. O backup do setor de inicialização e as informações de layout das partições também são críticos para uma recuperação bem-sucedida.

Em um ambiente de recuperação de desastres, a *recuperação bare-metal (BMR)*, que não depende do software original, se refere a um backup no qual todos os metadados, informações de sistema e configurações de aplicativos são gravados apropriadamente para gerar uma recuperação de sistema completa. A BMR cria o sistema básico, que abrange particionamento, layout do sistema de arquivos, sistema operacional, aplicativos e todas as configurações relevantes. A BMR recupera primeiro o sistema básico antes de iniciar a recuperação dos arquivos de dados. Algumas tecnologias de BMR podem restaurar um servidor em hardware diferente.

12.6 Processo de backup

Um sistema de backup usa arquitetura cliente/servidor com um servidor de backup e vários clientes de backup. O servidor de backup gerencia as operações e mantém o catálogo de backup, que contém informações sobre o processo e os metadados do backup. O servidor de backup se baseia em clientes de backup para coletar os dados a serem gravados. Os clientes podem ser locais ao servidor ou ficar em outro servidor, presumivelmente para fazer backup dos dados visíveis a esse servidor. O servidor de backup recebe metadados do backup desses clientes para executar suas atividades.

A Figura 12-4 ilustra o processo de backup. O nó de armazenamento é responsável pela gravação dos dados no dispositivo de backup (em um ambiente de backup, um nó de armazenamento é um host que controla dispositivos de backup). Geralmente, o nó de armazenamento é integrado com o servidor de backup e ambos são hospedados na mesma plataforma física. Um dispositivo de backup é anexado diretamente à plataforma que hospeda o nó de armazenamento. Algumas arquiteturas de backup se referem ao nó de armazenamento como *servidor de mídia*, porque ele se conecta ao dispositivo de armazenamento. Nós de armazenamento desempenham um papel importante no planejamento de backups porque podem ser usados para consolidar servidores de backup.

O processo de backup é baseado nas políticas definidas no servidor de backup, como o horário do dia ou o término de um evento. O servidor de backup então inicia o processo enviando uma solicitação a um cliente de backup (backups também podem ser iniciados por um cliente). A solicitação instrui o cliente de backup a enviar seus metadados ao servidor de backup e os dados a serem gravados ao nó de armazenamento apropriado. No recebimento da solicitação, o cliente de backup envia os metadados ao servidor de backup. O servidor de backup grava seus metadados em seu catálogo de metadados. O cliente de backup também envia os dados ao nó de armazenamento, que grava os dados no dispositivo de armazenamento.

Depois de todos os dados serem gravados, o nó de armazenamento encerra a conexão com o dispositivo de backup, e o servidor de backup grava no catálogo de metadados o status de backup encerrado.

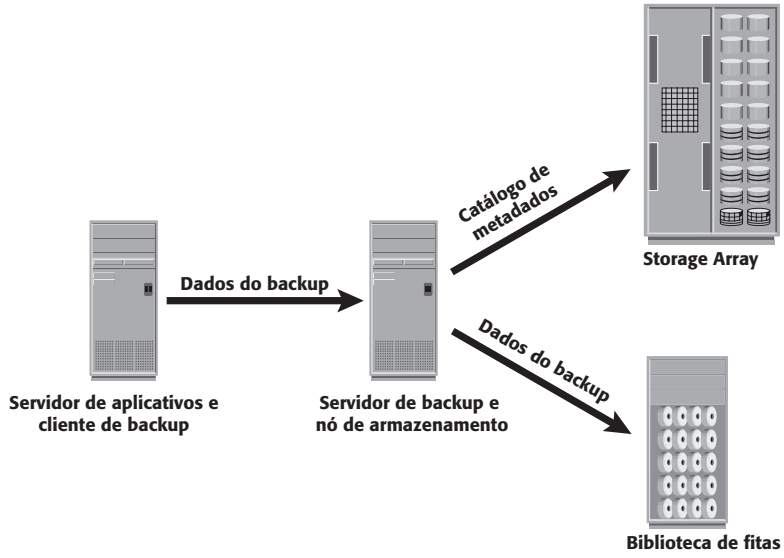


Figura 12-4 Arquitetura e processo de backup.

O software de backup também fornece capacidades extensas de relatórios baseadas no catálogo de backup e nos arquivos de registro. Esses relatórios podem incluir informações como o volume de dados gravados, o número de backups completados, o número de backups incompletos e os tipos de erro que podem ter ocorrido. Os relatórios podem ser personalizados dependendo do software de backup específico usado.

12.7 Operações de backup e restauração

Quando um processo de backup é iniciado, ocorre uma significativa comunicação em rede entre os diferentes componentes de uma infraestrutura de backup. O servidor de backup inicia o processo de backup para diferentes clientes com base no cronograma de backup configurado para eles. Por exemplo, o processo de backup para um grupo de clientes pode ser agendado para começar todos os dias às 3h da manhã.

O servidor de backup coordena o processo com todos os componentes da configuração (veja a Figura 12-5). O servidor de backup mantém as informações sobre os clientes de backup a serem contactados e nós de armazenamento a serem usados em uma operação de backup. O servidor de backup recupera as informações relacionadas ao backup do catálogo de backup e, baseado nelas, instrui o nó de armazenamento a carregar a mídia de backup apropriada nos dispositivos de backup. Simultaneamente, ele orienta os clientes de backup a iniciar a varredura dos dados, seu empacotamento e envio pela rede até o nó de armazenamento atribuído. O nó de armazenamento, por sua vez, envia os metadados ao servidor de backup para mantê-lo atualizado sobre a mídia utilizada no processo de backup.

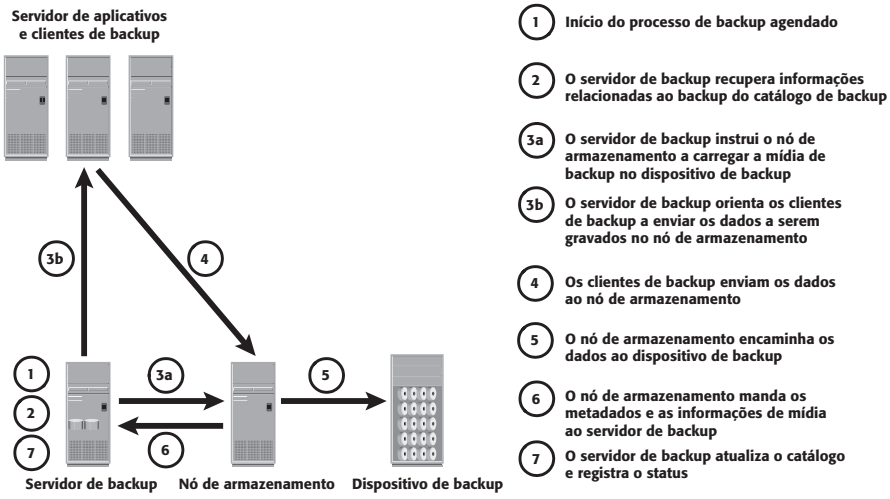


Figura 12-5 Operação de backup.

O servidor de backup atualiza continuamente o catálogo de backup com essas informações.

Depois de os dados serem gravados, podem ser restaurados quando necessário. Um processo de restauração deve ser iniciado manualmente. Alguns produtos de software de backup têm um aplicativo separado para operações de restauração. Os aplicativos de restauração são acessíveis apenas para os administradores. A Figura 12-6 ilustra um processo de restauração.

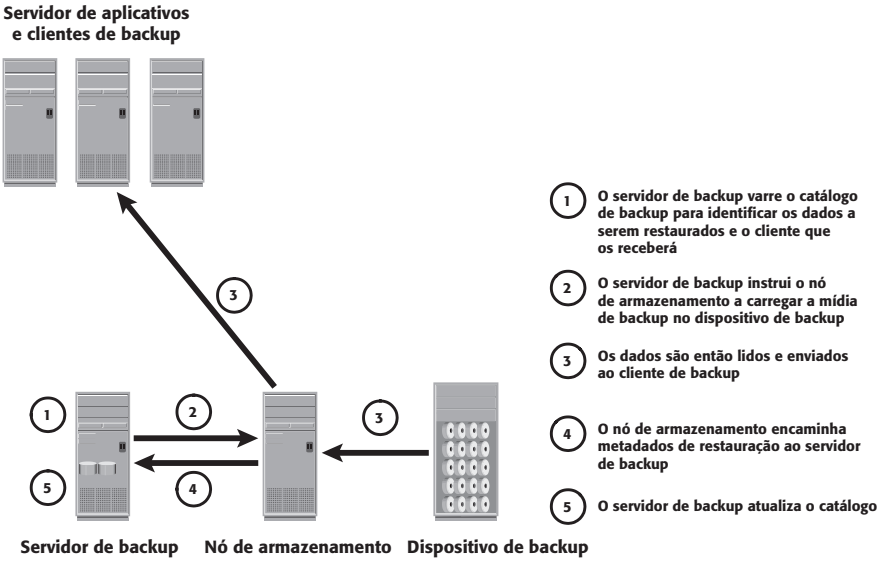


Figura 12-6 Operação de restauração.

No recebimento de uma solicitação de restauração, um administrador abre o aplicativo de restauração para visualizar a lista de clientes que foram gravados. Ao selecionar o cliente para o qual a solicitação de restauração foi feita, o administrador também precisa identificar o cliente que receberá os dados restaurados. Os dados podem ser restaurados no mesmo cliente para o qual a solicitação de restauração foi feita ou em outro cliente. O administrador então determina os dados a serem restaurados e o ponto específico do tempo no qual os dados têm de ser restaurados com base no RPO. Como todas essas informações vêm do catálogo de backup, o aplicativo de restauração também deve se comunicar com o servidor de backup.

O administrador primeiro seleciona os dados a serem restaurados e inicia o processo de restauração. O servidor de backup, por sua vez, usando o nó de armazenamento apropriado, identifica a mídia de backup que precisa ser montada no dispositivo de backup. Os dados são então lidos e enviados ao cliente que foi identificado para receber os dados restaurados.

Algumas restaurações são executadas com sucesso por meio da recuperação apenas dos dados de produção solicitados. Por exemplo, o processo de recuperação de uma planilha eletrônica é completado quando o arquivo específico é restaurado. Em restaurações de bancos de dados, dados adicionais como arquivos de registro e dados de produção devem ser restaurados. Nestes casos, o RTO é estendido devido aos passos adicionais no processo de restauração.

12.8 Topologias de backup

Três topologias básicas são usadas em um ambiente de backup: backup de conexão direta, backup baseado em LAN e backup baseado em SAN. Uma topologia mista também é utilizada com a combinação das topologias baseadas em LAN e em SAN.

Em um *backup de conexão direta*, um dispositivo de backup é ligado imediatamente ao cliente. Apenas os metadados são enviados ao servidor de backup através da LAN. Essa configuração libera a LAN do tráfego do backup. O exemplo da Figura 12-7 mostra o uso de um dispositivo de backup que não é com-

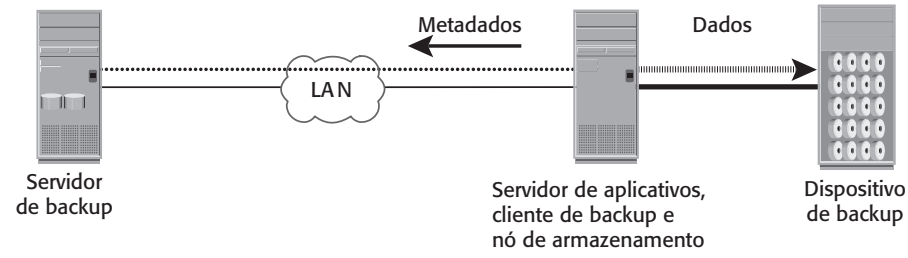


Figura 12-7 Topologia de backup de conexão direta.

partilhado. À medida que o ambiente cresce, entretanto, surge uma necessidade de gerenciamento central de todos os dispositivos de backup e de compartilhamento de recursos para otimizar custos. Neste exemplo, o cliente também age como um nó de armazenamento que grava dados no dispositivo de backup.

Em um *backup baseado em LAN*, todos os servidores são conectados à LAN e todos os dispositivos de armazenamento são anexados diretamente ao nó de armazenamento (veja a Figura 12-8). Os dados a serem gravados são transferidos do cliente de backup (origem) para o dispositivo de backup (destino) através da LAN, o que pode afetar o desempenho da rede. O fluxo pela LAN também incluencia o desempenho de rede de todos os sistemas conectados ao mesmo segmento do servidor de backup. Recursos de rede são severamente restringidos quando vários clientes acessam e compartilham a mesma unidade de biblioteca de fitas (TLU).

Esse impacto pode ser minimizado pela adoção de certas medidas, como a configuração de redes separadas para backup e a instalação de nós de armazenamento dedicados para alguns servidores de aplicativos.

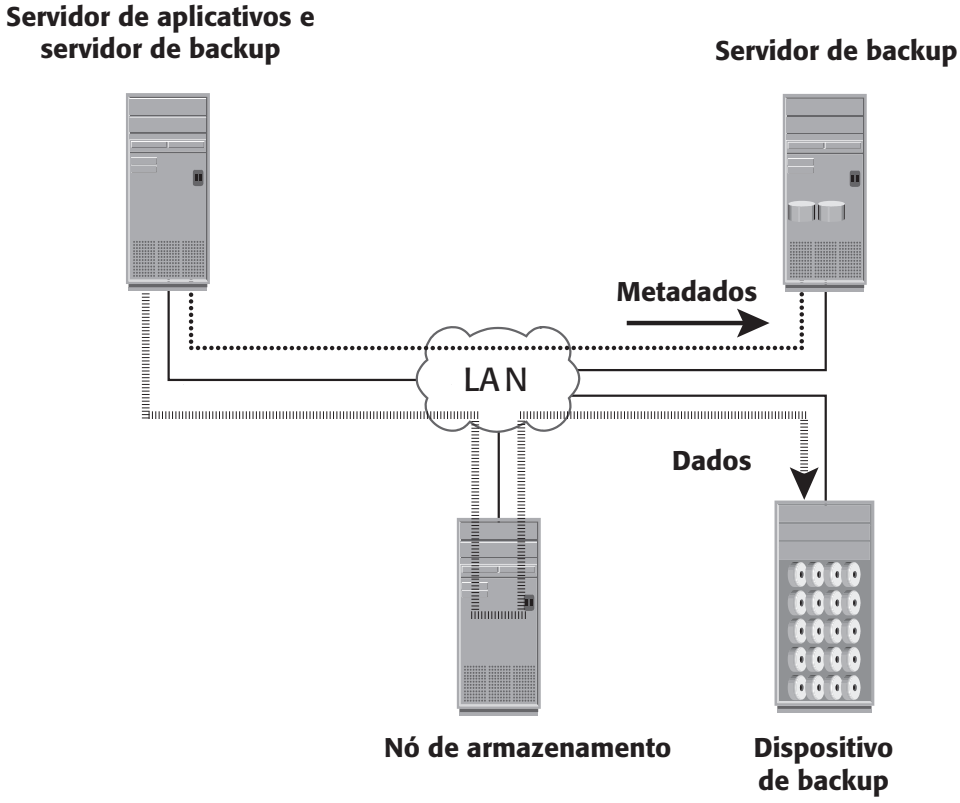


Figura 12-8 Topologia de backup baseado em LAN.

O *backup baseado em SAN* também é conhecido como *backup sem LAN*. A Figura 12-9 ilustra um backup baseado em SAN. A topologia de backup baseado em SAN é a solução mais apropriada quando um dispositivo de backup precisa ser compartilhado entre os clientes. Neste caso, o dispositivo de backup e os clientes são anexados à SAN.

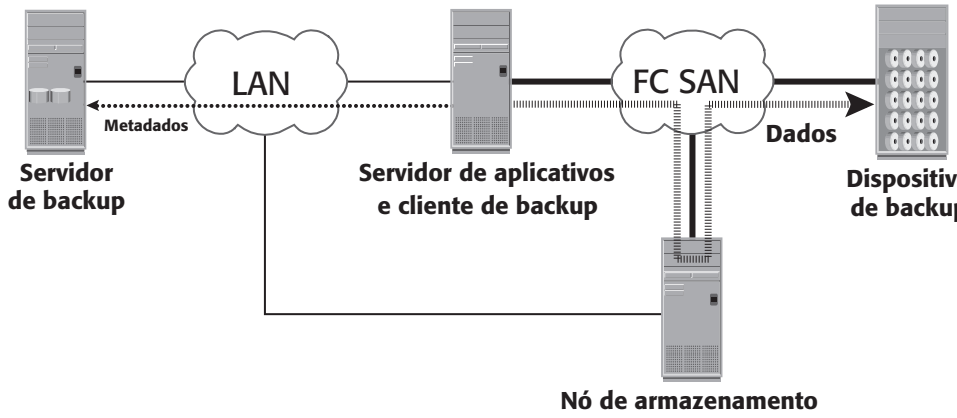


Figura 12-9 Topologia de backup baseado em SAN.

Neste exemplo, os clientes leem os dados dos servidores de e-mail na SAN e gravam no dispositivo de backup anexado à SAN. O tráfego dos dados de backup fica restrito à SAN e os metadados do backup são transportados pela LAN. Entretanto, o volume dos metadados é insignificante quando comparado aos dados de produção. O desempenho da LAN não é degradado nesta configuração.

Removendo o gargalo da rede, a SAN melhora o desempenho do backup em fita porque libera a LAN do tráfego de backup. Ao mesmo tempo, backups sem LAN podem afetar o host e o aplicativo, já que eles consomem largura de banda de I/O, memória e recursos de CPU do host.

O surgimento de discos de baixo custo como mídia de backup permitiu aos arrays de disco serem anexados à SAN e usados como dispositivos de backup. Um backup em fita desses backups de dados nos discos pode ser criado e enviado para outro local para recuperação de desastres e retenção a longo prazo.

A *topologia mista* usa topologias baseadas em LAN e em SAN, conforme mostrado na Figura 12-10. Esta topologia poderia ser implementada por diversos motivos, inclusive custo, localização do servidor, redução do gasto administrativo e preocupações com o desempenho.

12.8.1 Backup sem servidor

Um *backup sem servidor* é uma metodologia de backup sem LAN que não envolve um servidor de backup para duplicar os dados. A cópia pode ser criada por um

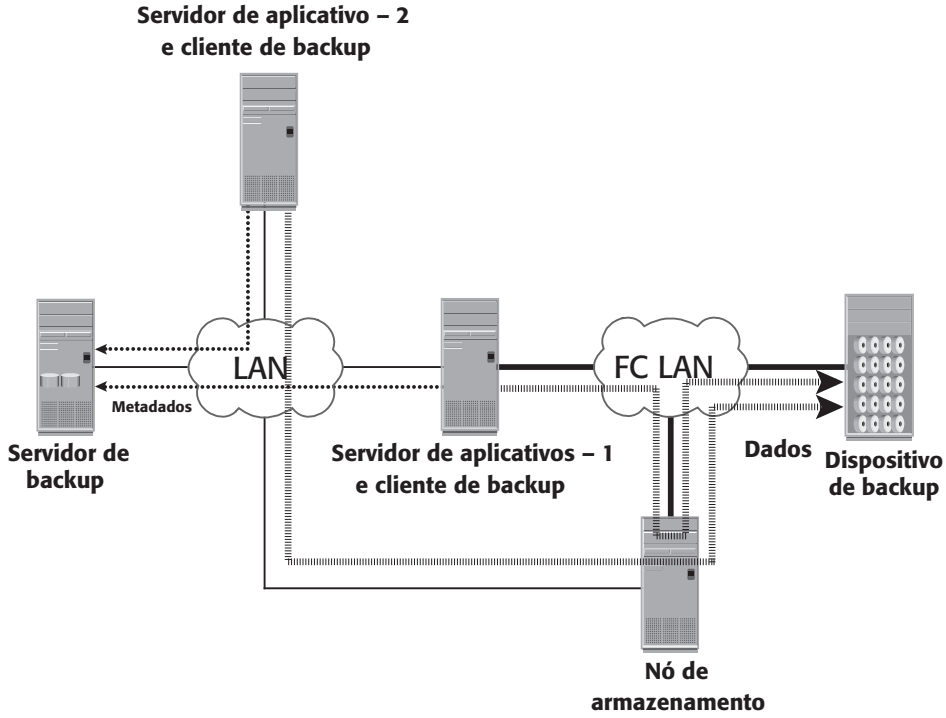


Figura 12-10 Topologia de backup misto.

controlador anexado à rede com a utilização de uma cópia estendida SCSI ou um dispositivo dentro da SAN. Esses backups são chamados de "sem servidor" porque usam recursos da SAN em vez de recursos do host para transportar dados de backup de sua origem para o dispositivo de backup, reduzindo o impacto sobre o servidor de aplicativos.

Outro método amplamente empregado para executar backup sem servidor é utilizar tecnologias de replicação local e remota. Neste caso, uma cópia consistente dos dados de produção é replicada dentro do mesmo array ou do array remoto, o qual pode ser movido para o dispositivo de backup pelo uso de um nó de armazenamento. Tecnologias de replicação são examinadas em detalhes no Capítulo 13 e no Capítulo 14.

12.9 Backup em ambientes NAS

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

Dica do Professor

Assista ao vídeo para conhecer os conceitos, a granularidade e as topologias de backup.

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

Exercícios

1) **São motivos para efetuar backup:**

A) Recuperação de desastres, backup operacional e arquivamento.

B) Recuperação de dados, backup diferencial e arquivamento.

C) Recuperação de desastres, backup operacional e snapshot.

D) Recuperação de desastres, backup operacional e compartilhamento.

E) Recuperação de I/O, backup operacional e arquivamento.

2) **O backup por arquivamento deverá utilizar, preferencialmente, qual tecnologia local?**

A) Cloud backup.

B) Snapshot.

C) CAS.

D) Dentro do próprio disco em LVM.

E) SAN-IP.

3) **Qual dos itens abaixo não faz parte da granularidade de backup?**

A) Backup completo.

B) Backup incremental.

C) Backup cumulativo.

D) Backup retrospectivo.

E) Backup sistemático.

4) **O RPO determina:**

A) A frequência dos backups.

B) A tolerância dos backups.

C) Tempo de restauração.

D) Consumo de CPU.

E) Tráfego de I/O.

5) Qual das topologias abaixo NÃO faz parte de uma topologia de backup?

A) SAN.

B) NAS.

C) CAS.

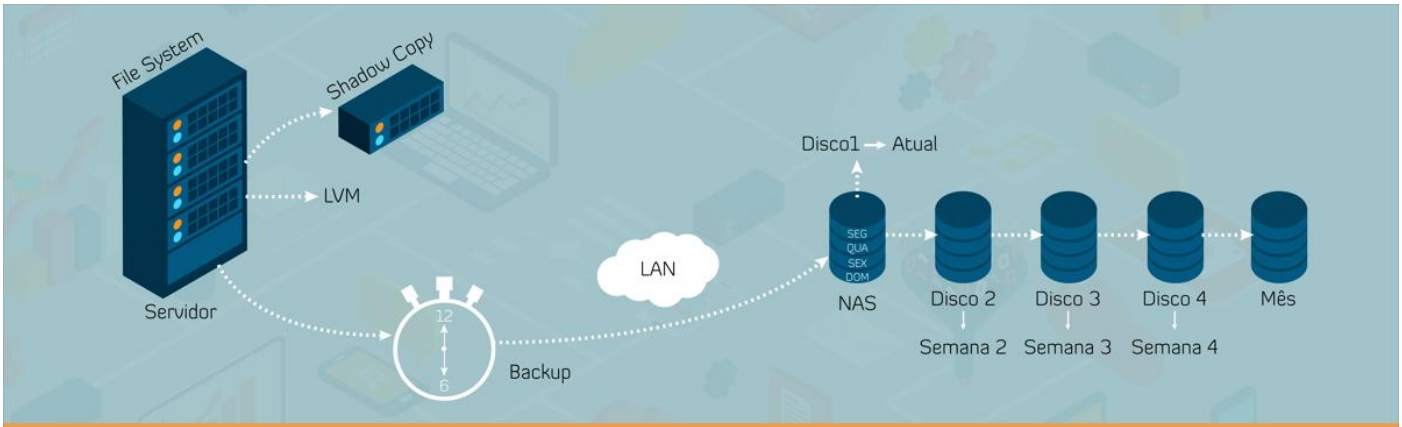
D) LAN.

E) SAN-IP.

Na prática

Rogério é gerente administrativo de uma empresa de pequeno porte no ramo de consertos de eletrodomésticos. Como ele tinha receio de perder informações importantes da empresa e dos clientes, decidiu chamar um profissional que apresentasse uma solução de backup e recuperação de dados.

Acompanhe as sugestões do profissional de TI para a empresa de Rogério.



Para uma empresa com poucos recursos financeiros, pode-se utilizar sistema de arquivos em LVM, habilitando o recurso de cópias sombra, tendo, assim, um rápido recovery em caso de desastres. Pode-se somar à estrutura um NAS, em que a cópia total dos dados será efetuada após o horário de expediente da empresa. A granularidade poderá ser total e diária, porém sobrescrevendo o volume total do dia a dia durante a semana, e gerando, a cada domingo, um backup total semanal, que será guardado separadamente. Após um mês, ou seja, quatro semanas, serão quatro backups, sendo um para cada semana. Assim, pode-se agrupar esses quatro backups separadamente em um backup mensal, que será enviado para uma estrutura em nuvem. A cada mês, a estrutura se renova em um rodízio, garantindo, assim, um histórico de recuperações e um bom sistema de tolerância a falhas.

Saiba mais

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

Fazer backup ou restaurar dados no dispositivo Android

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

Desenvolvendo uma estratégia eficiente de backup e restore em ambientes transacionais

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

Conceitos básicos de backup

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

Configurar um arquivo ou backup do Tivoli Storage Manager (TSM)

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

Qual a importância do backup?

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