Proteção de dados Oracle

Enterprise applications

NetApp
February 11, 2026

Índice

Proteção de dados Oracle

Proteção de dados com o ONTAP

O NetApp sabe que os dados mais críticos são encontrados em bancos de dados.

Uma empresa não pode operar sem acesso a seus dados e, às vezes, os dados definem o negócio. No entanto, a proteção de dados é mais do que apenas garantir um backup utilizável. No entanto, é preciso realizar os backups de maneira rápida e confiável, além de armazená-los em segurança.

O outro lado da proteção de dados é a recuperação de dados. Quando os dados estão inacessíveis, a empresa é afetada e pode estar inoperante até que os dados sejam restaurados. Este processo deve ser rápido e confiável. Finalmente, a maioria dos bancos de dados deve ser protegida contra desastres, o que significa manter uma réplica do banco de dados. A réplica deve estar suficientemente atualizada. Também deve ser rápido e simples tornar a réplica um banco de dados totalmente operacional.

Esta documentação substitui o relatório técnico publicado anteriormente *TR-4591: Proteção de dados Oracle: Backup, recuperação e replicação.*

Planejamento

A arquitetura de proteção de dados empresariais certa depende dos requisitos de negócios relacionados à retenção de dados, capacidade de recuperação e tolerância de interrupções durante vários eventos.

Por exemplo, considere o número de aplicações, bancos de dados e conjuntos de dados importantes no escopo. Criar uma estratégia de backup para um único conjunto de dados que garanta a conformidade com SLAs típicos é bastante simples, pois não há muitos objetos para gerenciar. À medida que o número de conjuntos de dados aumenta, o monitoramento se torna mais complicado e os administradores podem ser forçados a gastar um tempo cada vez maior lidando com falhas de backup. À medida que um ambiente atinge a nuvem e o provedor de serviços dimensiona, é necessária uma abordagem totalmente diferente.

O tamanho do conjunto de dados também afeta a estratégia. Por exemplo, existem muitas opções para backup e recuperação com um banco de dados 100GB porque o conjunto de dados é tão pequeno. A simples cópia de dados de Mídia de backup com ferramentas tradicionais normalmente fornece um rto suficiente para recuperação. Um banco de dados 100TB normalmente precisa de uma estratégia completamente diferente, a menos que o rto permita uma interrupção de vários dias, caso em que um procedimento tradicional de backup e recuperação baseado em cópia pode ser aceitável.

Finalmente, existem fatores fora do próprio processo de backup e recuperação. Por exemplo, existem bancos de dados que suportam atividades críticas de produção, tornando a recuperação um evento raro que só é realizado por DBAs qualificados? Alternativamente, os bancos de dados fazem parte de um grande ambiente de desenvolvimento em que a recuperação é uma ocorrência frequente e gerenciada por uma EQUIPE DE TI generalista?

Planejamento de rto, RPO e SLA

O ONTAP permite que você personalize facilmente uma estratégia de proteção de dados de banco de dados Oracle de acordo com suas necessidades de negócios.

Esses requisitos incluem fatores como a velocidade de recuperação, a perda máxima de dados permitida e as necessidades de retenção de backup. O plano de proteção de dados também deve levar em consideração

vários requisitos regulatórios para retenção e restauração de dados. Finalmente, diferentes cenários de recuperação de dados devem ser considerados, desde a recuperação típica e previsível resultante de erros de usuário ou aplicativo até cenários de recuperação de desastres que incluem a perda completa de um site.

Pequenas alterações nas políticas de proteção e recuperação de dados podem ter um efeito significativo na arquitetura geral de storage, backup e recuperação. É fundamental definir e documentar padrões antes de iniciar o trabalho de design para evitar complicar uma arquitetura de proteção de dados. Recursos desnecessários ou níveis de proteção levam a custos desnecessários e sobrecarga de gerenciamento, e um requisito inicialmente negligenciado pode levar um projeto na direção errada ou exigir alterações de design de última hora.

Objetivo de tempo de recuperação

O objetivo de tempo de recuperação (rto) define o tempo máximo permitido para a recuperação de um serviço. Por exemplo, um banco de dados de recursos humanos pode ter um rto de 24 horas porque, embora seja muito inconveniente perder o acesso a esses dados durante o dia de trabalho, a empresa ainda pode operar. Em contraste, um banco de dados que suporta o Registro geral de um banco teria um rto medido em minutos ou mesmo segundos. Um rto de zero não é possível, porque deve haver uma maneira de diferenciar entre uma interrupção de serviço real e um evento de rotina, como um pacote de rede perdido. No entanto, um rto quase zero é um requisito típico.

Objetivo do ponto de restauração

O objetivo do ponto de restauração (RPO) define a perda máxima de dados tolerável. Em muitos casos, o RPO é determinado exclusivamente pela frequência de snapshots ou atualizações do SnapMirror.

Em alguns casos, o RPO pode ser tornado mais agressivo, protegendo determinados dados de forma seletiva com mais frequência. Em um contexto de banco de dados, o RPO geralmente é uma questão de quanto dados de log podem ser perdidos em uma situação específica. Em um cenário típico de recuperação em que um banco de dados é danificado devido a um bug de produto ou erro de usuário, o RPO deve ser zero, o que significa que não deve haver perda de dados. O procedimento de recuperação envolve restaurar uma cópia anterior dos arquivos de banco de dados e, em seguida, rereproduzir os arquivos de log para trazer o estado do banco de dados até o ponto desejado no tempo. Os ficheiros de registo necessários para esta operação já devem estar no local original.

Em cenários incomuns, os dados de log podem ser perdidos. Por exemplo, um acidental ou malicioso `rm -rf *` de arquivos de banco de dados pode resultar na exclusão de todos os dados. A única opção seria restaurar do backup, incluindo arquivos de log, e alguns dados seriam inevitavelmente perdidos. A única opção para melhorar o RPO em um ambiente de backup tradicional seria executar backups repetidos dos dados de log. Isso tem limitações, no entanto, devido à constante movimentação de dados e à dificuldade de manter um sistema de backup como um serviço em constante execução. Um dos benefícios dos sistemas avançados de storage é a capacidade de proteger dados de danos acidentais ou mal-intencionados nos arquivos, além de fornecer um RPO melhor sem movimentação de dados.

Recuperação de desastres

A recuperação de desastres inclui a arquitetura, as políticas e os procedimentos DE TI necessários para recuperar um serviço em caso de desastre físico. Isso pode incluir inundações, incêndios ou pessoas agindo com intenção maliciosa ou negligente.

A recuperação de desastres é mais do que apenas um conjunto de procedimentos de recuperação. É o processo completo de identificar os vários riscos, definir os requisitos de recuperação de dados e continuidade do serviço e fornecer a arquitetura certa com os procedimentos associados.

Ao estabelecer os requisitos de proteção de dados, é essencial diferenciar entre os requisitos típicos de RPO e rto e os requisitos de RPO e rto necessários para a recuperação de desastres. Alguns ambientes de aplicações exigem um RPO de zero e um rto quase zero para situações de perda de dados que vão desde um erro de usuário relativamente normal até um incêndio que destrói um data center. No entanto, existem consequências administrativas e de custos para estes elevados níveis de proteçãoão.

Em geral, os requisitos de recuperação de dados que não sejam de desastres devem ser rigorosos por dois motivos. Primeiro, erros de aplicativos e erros de usuário que danificam dados são previsíveis ao ponto de serem quase inevitáveis. Em segundo lugar, não é difícil projetar uma estratégia de backup que possa fornecer um RPO de zero e um rto baixo, contanto que o sistema de storage não seja destruído. Não há motivo para não solucionar um risco significativo que seja facilmente remediado, e é por isso que os destinos RPO e rto para recuperação local devem ser agressivos.

Os requisitos de rto e RPO para recuperação de desastres variam mais amplamente com base na probabilidade de um desastre e nas consequências da perda ou interrupção de dados associada a uma empresa. Os requisitos de RPO e rto devem ser baseados nas necessidades reais dos negócios e não em princípios gerais. Eles devem levar em conta vários cenários de desastre lógico e físico.

Desastres lógicos

Os desastres lógicos incluem corrupção de dados causada por usuários, erros de aplicativos ou sistemas operacionais e falhas de software. Os desastres lógicos também podem incluir ataques maliciosos de terceiros com vírus ou worms ou explorando vulnerabilidades de aplicativos. Nesses casos, a infraestrutura física não está danificada, mas os dados subjacentes não são mais válidos.

Um tipo cada vez mais comum de desastre lógico é conhecido como ransomware, no qual um vetor de ataque é usado para criptografar dados. A criptografia não danifica os dados, mas torna-os indisponíveis até que o pagamento seja feito a terceiros. Um número cada vez maior de empresas está sendo alvo específico de hacks de ransomware. Para essa ameaça, o NetApp oferece snapshots à prova de violações, onde nem mesmo o administrador de storage pode alterar os dados protegidos antes da data de expiração configurada.

Desastres físicos

Os desastres físicos incluem a falha de componentes de uma infraestrutura que excede seus recursos de redundância e resultam em perda de dados ou perda estendida de serviço. Por exemplo, a proteção RAID fornece redundância de unidade de disco e o uso de HBAs fornece redundância de porta FC e cabo FC. Falhas de hardware de tais componentes são previsíveis e não afetam a disponibilidade.

Em um ambiente corporativo, geralmente é possível proteger a infraestrutura de um local inteiro com componentes redundantes até o ponto em que o único cenário de desastre físico previsível é a perda completa do local. O Planejamento de recuperação de desastre depende da replicação local a local.

Proteção de dados síncrona e assíncrona

Em um mundo ideal, todos os dados seriam replicados de forma síncrona em locais geograficamente dispersos. Essa replicação nem sempre é viável ou até possível por várias razões:

- A replicação síncrona aumenta inevitavelmente a latência de gravação porque todas as alterações devem ser replicadas em ambos os locais antes que a aplicação/banco de dados possa prosseguir com o processamento. O efeito de desempenho resultante às vezes é inaceitável, descartando o uso do espelhamento síncrono.
- O aumento da adoção do storage SSD de 100% significa que a latência de gravação adicional provavelmente será notada porque as expectativas de desempenho incluem centenas de milhares de IOPS e latência inferior a milissegundos. Aproveitar todos os benefícios do uso de SSDs de 100% pode

exigir uma nova visita à estratégia de recuperação de desastres.

- Os conjuntos de dados continuam a crescer em termos de bytes, criando desafios com a garantia de largura de banda suficiente para sustentar a replicação síncrona.
- Os conjuntos de dados também crescem em termos de complexidade, criando desafios com o gerenciamento da replicação síncrona de grande escala.
- Estratégias baseadas na nuvem frequentemente envolvem maiores distâncias de replicação e latência, o que impede o uso do espelhamento síncrono.

A NetApp oferece soluções que incluem replicação síncrona para as demandas mais exigentes de recuperação de dados e soluções assíncronas que proporcionam melhor desempenho e flexibilidade. Além disso, a tecnologia NetApp se integra perfeitamente a muitas soluções de replicação de terceiros, como o Oracle DataGuard

Tempo de retenção

O último aspeto de uma estratégia de proteção de dados é o tempo de retenção de dados, que pode variar muito.

- Um requisito típico é de 14 dias de backups noturnos no local principal e 90 dias de backups armazenados em um local secundário.
- Muitos clientes criam arquivos trimestrais autônomos armazenados em Mídias diferentes.
- Um banco de dados constantemente atualizado pode não ter necessidade de dados históricos, e os backups precisam ser mantidos apenas por alguns dias.
- Os requisitos regulatórios podem exigir recuperação até o ponto de qualquer transação arbitrária em uma janela de 365 dias.

Disponibilidade do banco de dados

O ONTAP foi projetado para oferecer a máxima disponibilidade de banco de dados Oracle. Uma descrição completa dos recursos de alta disponibilidade do ONTAP está além do escopo deste documento. No entanto, assim como na proteção de dados, uma compreensão básica dessa funcionalidade é importante ao projetar uma infraestrutura de banco de dados.

Pares HA

A unidade básica de alta disponibilidade é o par de HA. Cada par contém links redundantes para suportar a replicação de dados para o NVRAM. NVRAM não é um cache de gravação. A RAM dentro do controlador serve como cache de gravação. O objetivo do NVRAM é registrar temporariamente os dados como uma salvaguarda contra falhas inesperadas do sistema. A este respeito, é semelhante a um log refazer banco de dados.

Tanto o NVRAM quanto o log refazer de banco de dados são usados para armazenar dados rapidamente, permitindo que as alterações aos dados sejam confirmadas o mais rápido possível. A atualização para os dados persistentes em unidades (ou datafiles) não ocorre até mais tarde durante um processo chamado ponto de verificação em plataformas ONTAP e na maioria dos bancos de dados. Nem os dados do NVRAM nem os logs de refazer do banco de dados são lidos durante operações normais.

Se um controlador falhar abruptamente, é provável que haja alterações pendentes armazenadas no NVRAM que ainda não foram gravadas nas unidades. O controlador do parceiro deteta a falha, controla as unidades e

aplica as alterações necessárias que foram armazenadas no NVRAM.

Takeover e giveback

Takeover e giveback refere-se ao processo de transferência de responsabilidade por recursos de storage entre nós em um par de HA. Há dois aspetos para a aquisição e a giveback:

- Gestão da conetividade de rede que permite o acesso às unidades
- Gestão das próprias unidades

As interfaces de rede que suportam o tráfego CIFS e NFS são configuradas com um local de origem e failover. Uma aquisição inclui mover as interfaces de rede para sua casa temporária em uma interface física localizada na(s) mesma(s) sub-rede(s) do local original. A giveback inclui mover as interfaces de rede de volta para seus locais originais. O comportamento exato pode ser ajustado conforme necessário.

As interfaces de rede que suportam protocolos de bloco SAN, como iSCSI e FC, não são relocadas durante a aquisição e a giveback. Em vez disso, os LUNs devem ser provisionados com caminhos que incluam um par de HA completo, o que resulta em um caminho primário e um caminho secundário.

Caminhos adicionais para controladores adicionais também podem ser configurados para dar suporte à realocação de dados entre nós em um cluster maior, mas isso não faz parte do processo de HA.

O segundo aspeto da aquisição e da giveback é a transferência da propriedade do disco. O processo exato depende de vários fatores, incluindo o motivo da aquisição/giveback e as opções de linha de comando emitidas. O objetivo é realizar a operação da forma mais eficiente possível. Embora o processo geral possa parecer exigir vários minutos, o momento real em que a propriedade da unidade é transferida de nó para nó geralmente pode ser medido em segundos.

Tempo de takeover

A e/S do host tem uma pequena pausa na e/S durante as operações de takeover e giveback, mas não deve haver interrupção do aplicativo em um ambiente configurado corretamente. O processo de transição real no qual a e/S é atrasada geralmente é medido em segundos, mas o host pode exigir tempo adicional para reconhecer a alteração nos caminhos de dados e reenviar operações de e/S.

A natureza da interrupção depende do protocolo:

- Uma interface de rede que suporta tráfego NFS e CIFS emite uma solicitação ARP (Address Resolution Protocol) para a rede após a transição para um novo local físico. Isso faz com que os switches de rede atualizem suas tabelas de endereços de controle de acesso de Mídia (MAC) e retomem a e/S de processamento A interrupção no caso de aquisição planejada e a giveback geralmente é medida em segundos e, em muitos casos, não é detetável. Algumas redes podem ser mais lentas para reconhecer completamente a mudança no caminho da rede, e alguns sistemas operacionais podem colocar em fila muita e/S em um tempo muito curto que deve ser tentado novamente. Isto pode prolongar o tempo necessário para retomar a I/O.
- Uma interface de rede que suporte protocolos SAN não faz a transição para um novo local. Um sistema operacional do host deve alterar o caminho ou os caminhos em uso. A pausa em I/o observada pelo host depende de vários fatores. Do ponto de vista do sistema de armazenamento, o período em que a e/S não pode ser atendida é de apenas alguns segundos. No entanto, diferentes sistemas operacionais de host podem exigir tempo adicional para permitir que uma e/S termine o tempo limite antes de tentar novamente. Os sistemas operacionais mais novos são mais capazes de reconhecer uma mudança de caminho muito mais rapidamente, mas os sistemas operacionais mais antigos geralmente exigem até 30 segundos para

reconhecer uma mudança.

Os tempos de aquisição esperados durante os quais o sistema de storage não pode fornecer dados a um ambiente de aplicativo são mostrados na tabela abaixo. Não deve haver erros em qualquer ambiente de aplicativo, o controle deve aparecer como uma pequena pausa no processamento de e/S.

	NFS	AFF	ASA
Takeover planejado	15 seg	6-10 seg	2-3 seg
Takeover não planejado	30 seg	6-10 seg	2-3 seg

Somas de verificação e integridade de dados

O ONTAP e seus protocolos compatíveis incluem vários recursos que protegem a integridade do banco de dados Oracle, incluindo dados em repouso e dados transmitidos pela rede de rede.

A proteção de dados lógicos no ONTAP consiste em três requisitos principais:

- Os dados devem estar protegidos contra corrupção de dados.
- Os dados devem estar protegidos contra falha da unidade.
- As alterações nos dados devem ser protegidas contra perda.

Essas três necessidades são discutidas nas seções a seguir.

Corrupção de rede: Somas de verificação

O nível mais básico de proteção de dados é a soma de verificação, que é um código especial de deteção de erros armazenado junto com os dados. A corrupção de dados durante a transmissão da rede é detetada com o uso de uma soma de verificação e, em alguns casos, várias somas de verificação.

Por exemplo, um quadro FC inclui uma forma de checksum chamada de verificação de redundância cíclica (CRC) para garantir que a carga útil não esteja corrompida em trânsito. O transmissor envia os dados e o CRC dos dados. O recetor de um quadro FC recalcula o CRC dos dados recebidos para garantir que ele corresponda ao CRC transmitido. Se o CRC recém-calculado não corresponder ao CRC anexado ao quadro, os dados estarão corrompidos e o quadro FC será descartado ou rejeitado. Uma operação de e/S iSCSI inclui somas de verificação nas camadas TCP/IP e Ethernet e, para proteção adicional, também pode incluir proteção CRC opcional na camada SCSI. Qualquer corrupção de bits no fio é detetada pela camada TCP ou camada IP, o que resulta na retransmissão do pacote. Assim como no FC, erros no CRC SCSI resultam em uma rejeição ou rejeição da operação.

Drive corruption: Somas de verificação

As somas de verificação também são usadas para verificar a integridade dos dados armazenados nas unidades. Os blocos de dados gravados nas unidades são armazenados com uma função de checksum que produz um número imprevisível vinculado aos dados originais. Quando os dados são lidos a partir da unidade, a soma de verificação é recalculada e comparada com a soma de verificação armazenada. Se não corresponder, os dados ficaram corrompidos e devem ser recuperados pela camada RAID.

Corrupção de dados: Gravações perdidas

Um dos tipos mais difíceis de detetar é uma gravação perdida ou extraviada. Quando uma escrita é reconhecida, ela deve ser escrita para a Mídia no local correto. A corrupção de dados no local é relativamente fácil de detetar usando uma soma de verificação simples armazenada com os dados. No entanto, se a gravação é simplesmente perdida, então a versão anterior dos dados ainda pode existir e a soma de verificação estaria correta. Se a gravação for colocada no local físico errado, a soma de verificação associada seria mais uma vez válida para os dados armazenados, mesmo que a gravação tenha destruído outros dados.

A solução para este desafio é a seguinte:

- Uma operação de gravação deve incluir metadados que indicam o local onde a gravação deve ser encontrada.
- Uma operação de gravação deve incluir algum tipo de identificador de versão.

Quando o ONTAP grava um bloco, ele inclui dados sobre a localização do bloco. Se uma leitura subsequente identificar um bloco, mas os metadados indicarem que ele pertence ao local 123 quando foi encontrado no local 456, então a gravação foi extraviada.

Detetar uma escrita totalmente perdida é mais difícil. A explicação é muito complicada, mas essencialmente o ONTAP está armazenando metadados de uma forma que uma operação de gravação resulta em atualizações para dois locais diferentes nas unidades. Se uma gravação for perdida, uma leitura posterior dos dados e metadados associados mostrará duas identidades de versão diferentes. Isso indica que a gravação não foi concluída pela unidade.

A corrupção de gravação perdida e extraviada é extremamente rara, mas, à medida que as unidades continuam a crescer e os conjuntos de dados aumentam a escala de exabytes, o risco aumenta. A detecção de gravações perdidas deve ser incluída em qualquer sistema de storage que suporte workloads de banco de dados.

Falhas de unidade: RAID, RAID DP e RAID-teC

Se um bloco de dados em uma unidade for descoberto como corrompido ou toda a unidade falhar e estiver totalmente indisponível, os dados devem ser reconstituídos. Isso é feito no ONTAP usando unidades de paridade. Os dados são distribuídos em várias unidades de dados e, em seguida, os dados de paridade são gerados. Este é armazenado separadamente dos dados originais.

O ONTAP usou originalmente o RAID 4, que usa uma única unidade de paridade para cada grupo de unidades de dados. O resultado foi que qualquer unidade do grupo poderia falhar sem resultar em perda de dados. Se a unidade de paridade falhar, nenhum dado será danificado e uma nova unidade de paridade poderá ser construída. Se uma única unidade de dados falhar, as unidades restantes poderão ser usadas com a unidade de paridade para regenerar os dados em falta.

Quando as unidades eram pequenas, a chance estatística de duas unidades falharem simultaneamente foi insignificante. À medida que as capacidades da unidade crescem, também tem o tempo necessário para reconstruir os dados após uma falha de unidade. Isso aumentou a janela na qual uma segunda falha da unidade resultaria em perda de dados. Além disso, o processo de reconstrução cria muitas e/S adicionais nas unidades sobreviventes. À medida que as unidades envelhecem, o risco de carga adicional que leva a uma segunda falha da unidade também aumenta. Finalmente, mesmo que o risco de perda de dados não tenha aumentado com o uso continuado do RAID 4, as consequências da perda de dados se tornariam mais graves. Quanto mais dados forem perdidos no caso de uma falha do grupo RAID, mais tempo levaria para recuperar os dados, estendendo a interrupção dos negócios.

Esses problemas levaram a NetApp a desenvolver a tecnologia NetApp RAID DP, uma variante do RAID 6.

Essa solução inclui duas unidades de paridade, o que significa que todas as duas unidades em um grupo RAID podem falhar sem criar perda de dados. As unidades continuaram a crescer em tamanho, o que levou a NetApp a desenvolver a tecnologia NetApp RAID-teC, que introduz uma terceira unidade de paridade.

Algumas práticas recomendadas de banco de dados históricos recomendam o uso do RAID-10, também conhecido como espelhamento distribuído. Isso oferece menos proteção de dados do que até mesmo o RAID DP porque há vários cenários de falha de dois discos, enquanto que no RAID DP não há nenhum.

Há também algumas práticas recomendadas de banco de dados históricos que indicam que RAID-10 é preferível às opções RAID-4/5/6 devido a problemas de desempenho. Essas recomendações às vezes se referem a uma penalidade de RAID. Embora essas recomendações geralmente estejam corretas, elas não são aplicáveis às implementações de RAID no ONTAP. O problema de desempenho está relacionado com a regeneração de paridade. Com as implementações tradicionais de RAID, o processamento das gravações aleatórias de rotina executadas por um banco de dados requer várias leituras de disco para regenerar os dados de paridade e concluir a gravação. A penalidade é definida como o IOPS de leitura adicional necessário para executar operações de gravação.

O ONTAP não incorre em uma penalidade de RAID porque as gravações são encenadas na memória em que a paridade é gerada e, em seguida, gravadas no disco como um único stripe RAID. Não são necessárias leituras para concluir a operação de gravação.

Em resumo, quando comparado ao RAID 10, o RAID DP e o RAID-teC oferecem muito mais capacidade utilizável, melhor proteção contra falha de unidade e nenhum sacrifício de performance.

Proteção contra falhas de hardware: NVRAM

Qualquer storage array que atenda a um workload de banco de dados precisa atender às operações de gravação o mais rápido possível. Além disso, uma operação de gravação deve ser protegida contra perda de um evento inesperado, como uma falha de energia. Isso significa que qualquer operação de gravação deve ser armazenada com segurança em pelo menos dois locais.

Os sistemas AFF e FAS contam com a NVRAM para atender a esses requisitos. O processo de escrita funciona da seguinte forma:

1. Os dados de gravação de entrada são armazenados na RAM.
2. As alterações que devem ser feitas nos dados no disco são registradas no NVRAM no nó local e no nó parceiro. O NVRAM não é um cache de gravação; em vez disso, é um diário semelhante a um log de refazer de banco de dados. Em condições normais, não é lido. Ele é usado apenas para recuperação, como após uma falha de energia durante o processamento de e/S.
3. A gravação é então reconhecida para o host.

O processo de gravação nesta fase é concluído do ponto de vista da aplicação, e os dados são protegidos contra perda porque são armazenados em dois locais diferentes. Eventualmente, as alterações são gravadas no disco, mas esse processo está fora da banda do ponto de vista do aplicativo, porque ocorre depois que a gravação é reconhecida e, portanto, não afeta a latência. Este processo é mais uma vez semelhante ao log de banco de dados. Uma alteração ao banco de dados é registrada nos logs de refazer o mais rápido possível, e a alteração é então reconhecida como comprometida. As atualizações para os datafiles ocorrem muito mais tarde e não afetam diretamente a velocidade de processamento.

No caso de uma falha do controlador, o controlador do parceiro assume a propriedade dos discos necessários e replica os dados registados no NVRAM para recuperar quaisquer operações de e/S que estivessem em trânsito quando a falha ocorreu.

Proteção contra falhas de hardware: NVFAIL

Como discutido anteriormente, uma gravação não é reconhecida até que ela tenha sido registrada no NVRAM local e no NVRAM em pelo menos um outro controlador. Essa abordagem garante que uma falha de hardware ou falha de energia não resulte na perda de e/S em trânsito Se o NVRAM local falhar ou a conectividade com o parceiro de HA falhar, esses dados em trânsito não serão mais espelhados.

Se o NVRAM local relatar um erro, o nó será encerrado. Esse desligamento resulta em failover para uma controladora de parceiro de HA. Nenhum dado é perdido porque o controlador que sofre a falha não reconheceu a operação de gravação.

O ONTAP não permite um failover quando os dados estão fora de sincronia, a menos que o failover seja forçado. Forçar uma alteração de condições desta forma reconhece que os dados podem ser deixados para trás no controlador original e que a perda de dados é aceitável.

Os bancos de dados são especialmente vulneráveis à corrupção se um failover for forçado porque os bancos de dados mantêm grandes caches internos de dados no disco. Se ocorrer um failover forçado, as alterações anteriormente confirmadas serão efetivamente descartadas. O conteúdo da matriz de armazenamento salta efetivamente para trás no tempo, e o estado do cache do banco de dados não reflete mais o estado dos dados no disco.

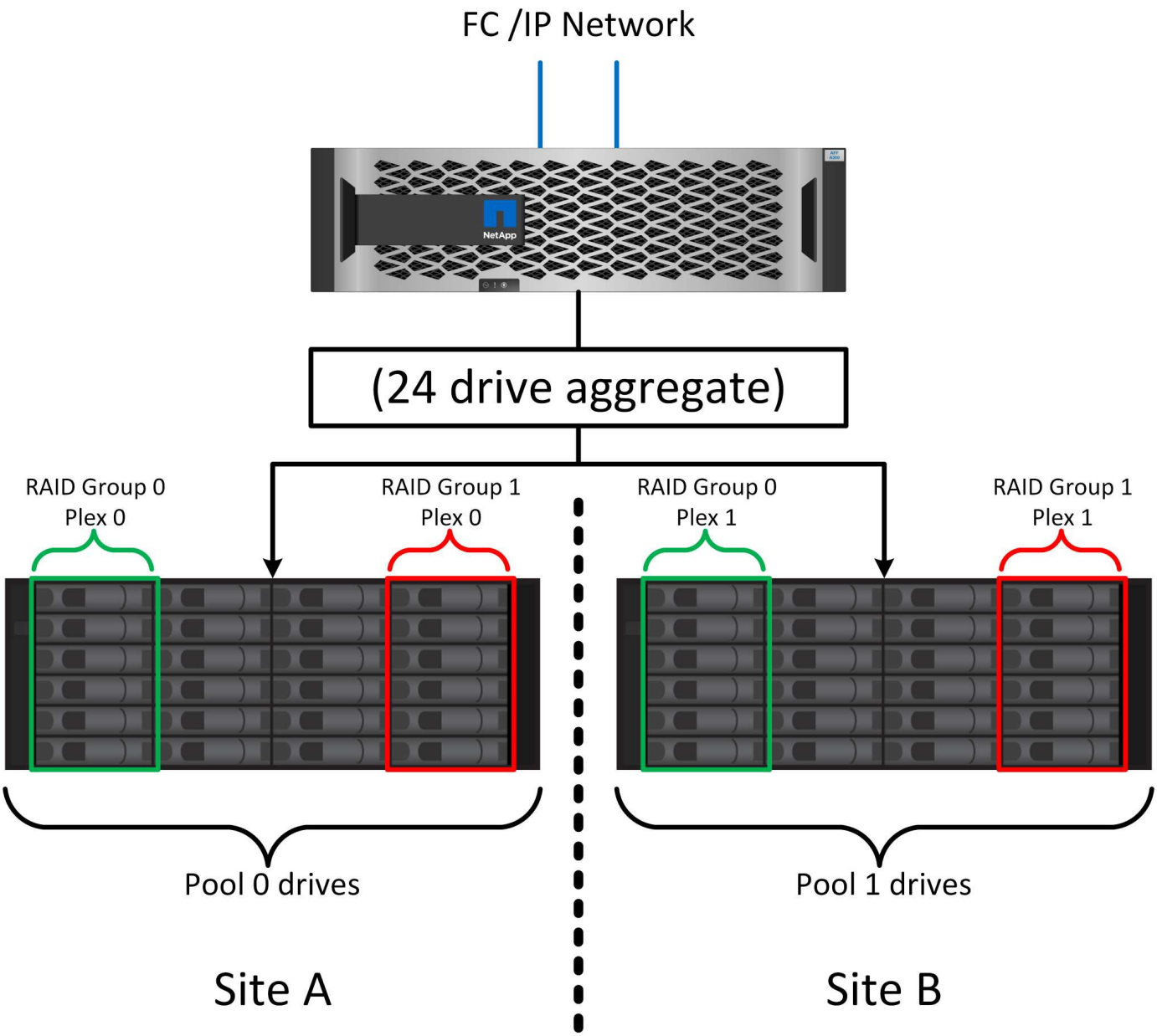
Para proteger os dados contra essa situação, o ONTAP permite que os volumes sejam configurados para proteção especial contra falhas do NVRAM. Quando acionado, esse mecanismo de proteção resulta em um volume entrando em um estado chamado NVFAIL. Esse estado resulta em erros de e/S que causam o desligamento de um aplicativo para que eles não usem dados obsoletos. Os dados não devem ser perdidos porque qualquer gravação reconhecida deve estar presente no storage array.

As próximas etapas usuais são para que um administrador desligue totalmente os hosts antes de colocar manualmente os LUNs e volumes novamente on-line. Embora essas etapas possam envolver algum trabalho, essa abordagem é a maneira mais segura de garantir a integridade dos dados. Nem todos os dados exigem essa proteção, e é por isso que o comportamento do NVFAIL pode ser configurado volume a volume.

Proteção contra falhas no local e no compartimento: SyncMirror e plexos

O SyncMirror é uma tecnologia de espelhamento que aprimora, mas não substitui, o RAID DP ou o RAID-teC. Ele espelha o conteúdo de dois grupos RAID independentes. A configuração lógica é a seguinte:

- As unidades são configuradas em dois pools com base no local. Um pool é composto por todas as unidades no local A, e o segundo pool é composto por todas as unidades no local B..
- Um pool comum de armazenamento, conhecido como agregado, é criado com base em conjuntos espelhados de grupos RAID. Um número igual de unidades é extraído de cada local. Por exemplo, um agregado SyncMirror de 20 unidades seria composto por 10 unidades do local A e 10 unidades do local B..
- Cada conjunto de unidades em um determinado local é configurado automaticamente como um ou mais grupos RAID-DP ou RAID-teC totalmente redundantes, independentemente do uso do espelhamento. Isso fornece proteção contínua de dados, mesmo após a perda de um site.



A figura acima ilustra um exemplo de configuração do SyncMirror. Um agregado de 24 unidades foi criado na controladora com 12 unidades de um compartimento alocado no local A e 12 unidades de um compartimento alocado no local B. as unidades foram agrupadas em dois grupos RAID espelhados. RAID Group 0 inclui um Plex de 6 unidades no local Um espelhado para um Plex de 6 unidades no local B. da mesma forma, RAID Group 1 inclui um Plex de 6 unidades no local Um espelhado para um Plex de 6 unidades no local B.

O SyncMirror normalmente é usado para fornecer espelhamento remoto com sistemas MetroCluster, com uma cópia dos dados em cada local. Ocasionalmente, ele tem sido usado para fornecer um nível extra de redundância em um único sistema. Em particular, ele fornece redundância em nível de prateleira. Um compartimento de unidades já contém fontes de alimentação duplas e controladores e é, no geral, pouco mais do que chapas metálicas, mas em alguns casos, a proteção extra pode ser garantida. Por exemplo, um cliente da NetApp implantou o SyncMirror para uma plataforma móvel de análise em tempo real usada durante testes automotivos. O sistema foi separado em dois racks físicos fornecidos por alimentações de energia independentes de sistemas UPS independentes.

Somas de verificação

O tópico de checksums é de particular interesse para DBAs que estão acostumados a usar backups de streaming Oracle RMAN migram para backups baseados em snapshot. Um recurso do RMAN é que ele executa verificações de integridade durante operações de backup. Embora esse recurso tenha algum valor, seu principal benefício é para um banco de dados que não é usado em um storage array moderno. Quando as unidades físicas são usadas para um banco de dados Oracle, é quase certo que a corrupção eventualmente ocorre à medida que as unidades envelhecem, um problema que é resolvido por somas de verificação baseadas em array em arrays de armazenamento reais.

Com um storage array real, a integridade de dados é protegida pelo uso de somas de verificação em vários níveis. Se os dados estiverem corrompidos em uma rede baseada em IP, a camada TCP (Transmission Control Protocol) rejeita os dados do pacote e solicita a retransmissão. O protocolo FC inclui somas de verificação, assim como os dados SCSI encapsulados. Depois que ele está no array, o ONTAP tem proteção RAID e checksum. A corrupção pode ocorrer, mas, como na maioria dos arrays corporativos, ela é detetada e corrigida. Normalmente, uma unidade inteira falha, solicitando uma reconstrução RAID e a integridade do banco de dados não é afetada. Ainda é possível que bytes individuais em uma unidade sejam danificados por radiação cósmica ou células flash com falha. Se isso acontecer, a verificação de paridade falharia, a unidade falharia e uma reconstrução RAID começaria. Mais uma vez, a integridade dos dados não é afetada. A linha final de defesa é o uso de somas de verificação. Se, por exemplo, um erro de firmware catastrófico em uma unidade corrompeu dados de uma forma que de alguma forma não fosse detetado por uma verificação de paridade RAID, a soma de verificação não corresponderia e o ONTAP impediria a transferência de um bloco corrompido antes que o banco de dados Oracle pudesse recebê-lo.

A arquitetura Oracle datafile e refazer log também foi projetada para fornecer o mais alto nível possível de integridade de dados, mesmo em circunstâncias extremas. No nível mais básico, os blocos Oracle incluem checksum e verificações lógicas básicas com quase todas as I/O. Se o Oracle não travou ou uma tablespace off-line, então os dados estão intactos. O grau de verificação da integridade dos dados é ajustável, e o Oracle também pode ser configurado para confirmar gravações. Como resultado, quase todos os cenários de falha e falha podem ser recuperados e, no caso extremamente raro de uma situação irrecuperável, a corrupção é imediatamente detetada.

A maioria dos clientes NetApp que usam bancos de dados Oracle descontinuam o uso de RMAN e outros produtos de backup após a migração para backups baseados em snapshot. Ainda existem opções nas quais o RMAN pode ser usado para executar a recuperação em nível de bloco com o SnapCenter. No entanto, no dia a dia, o RMAN, o NetBackup e outros produtos só são usados ocasionalmente para criar cópias de arquivamento mensais ou trimestrais.

Alguns clientes optam por executar `dbv` periodicamente para realizar verificações de integridade em seus bancos de dados existentes. NetApp desencoraja essa prática porque cria carga de e/S desnecessária. Como discutido acima, se o banco de dados não estava enfrentando problemas anteriormente, a chance de `dbv` detetar um problema é perto de zero, e este utilitário cria uma carga de e/S sequencial muito alta na rede e no sistema de armazenamento. A menos que haja razão para acreditar que existe corrupção, como a exposição a um bug conhecido da Oracle, não há razão para ser executado `dbv`.

Noções básicas de backup e recuperação

Backups baseados em snapshot

A base da proteção de dados de banco de dados Oracle no ONTAP é a tecnologia NetApp Snapshot.

Os valores-chave são os seguintes:

- **Simplicidade.** Um snapshot é uma cópia somente leitura do conteúdo de um contentor de dados em um determinado momento.
- **Eficiência.** Os instantâneos não requerem espaço no momento da criação. O espaço só é consumido quando os dados são alterados.
- **Capacidade de gerenciamento.** Uma estratégia de backup baseada em snapshots é fácil de configurar e gerenciar, pois os snapshots são uma parte nativa do sistema operacional de storage. Se o sistema de armazenamento estiver ligado, ele estará pronto para criar backups.
- **Escalabilidade.** É possível preservar até 1024 backups de um único contêiner de arquivos e LUNs. Para conjuntos de dados complexos, vários contêineres de dados podem ser protegidos por um único conjunto consistente de snapshots.
- O desempenho não é afetado, independentemente de um volume conter 1024 snapshots ou nenhum.

Embora muitos fornecedores de storage ofereçam tecnologia Snapshot, a tecnologia Snapshot no ONTAP é única e oferece benefícios significativos para ambientes de aplicações e bancos de dados empresariais:

- As cópias snapshot fazem parte do layout de arquivo em qualquer lugar (WAFL) subjacente. Eles não são uma tecnologia adicional ou externa. Isso simplifica o gerenciamento porque o sistema de storage é o sistema de backup.
- As cópias snapshot não afetam a performance, exceto em alguns casos de borda, como quando muitos dados são armazenados em snapshots que o sistema de storage subjacente preenche.
- O termo "grupo de consistência" é frequentemente usado para se referir a um agrupamento de objetos de armazenamento que são gerenciados como uma coleta consistente de dados. Um snapshot de um determinado volume ONTAP constitui um backup de grupo de consistência.

Os snapshots do ONTAP também são mais dimensionados do que a tecnologia da concorrência. Os clientes podem armazenar snapshots 5, 50 ou 500 sem afetar a performance. O número máximo de instantâneos atualmente permitido em um volume é 1024. Se a retenção adicional de snapshot for necessária, há opções para colocar os snapshots em cascata em volumes adicionais.

Como resultado, proteger um conjunto de dados hospedado no ONTAP é simples e altamente dimensionável. Os backups não exigem movimentação de dados, portanto, uma estratégia de backup pode ser adaptada às necessidades da empresa, em vez das limitações de taxas de transferência de rede, grande número de unidades de fita ou áreas de preparo de disco.

Um instantâneo é um backup?

Uma pergunta comumente feita sobre o uso de snapshots como estratégia de proteção de dados é o fato de que os dados "reais" e os dados instantâneos estão localizados nas mesmas unidades. A perda dessas unidades resultaria na perda dos dados primários e do backup.

Este é um problema válido. Os snapshots locais são usados para necessidades diárias de backup e recuperação e, nesse aspeto, o snapshot é um backup. Cerca de 99% de todos os cenários de recuperação em ambientes NetApp dependem de snapshots para atender até aos requisitos mais agressivos de rto.

No entanto, os snapshots locais nunca devem ser a única estratégia de backup. É por isso que a NetApp oferece tecnologia como replicação SnapMirror e SnapVault para replicar snapshots para um conjunto independente de unidades com rapidez e eficiência. Em uma solução de arquitetura adequada com snapshots e replicação de snapshot, o uso da fita pode ser minimizado para talvez um arquivamento trimestral ou eliminado totalmente.

Backups baseados em snapshot

Há muitas opções para o uso de cópias ONTAP Snapshot para proteger seus dados. Além disso, os snapshots são a base de muitos outros recursos da ONTAP, incluindo replicação, recuperação de desastres e clonagem. Uma descrição completa da tecnologia de instantâneos está além do escopo deste documento, mas as seções a seguir fornecem uma visão geral.

Existem duas abordagens principais para criar um instantâneo de um conjunto de dados:

- Backups consistentes com falhas
- Backups consistentes com aplicativos

Um backup consistente com falhas de um conjunto de dados refere-se à captura de toda a estrutura do conjunto de dados em um único ponto no tempo. Se o conjunto de dados for armazenado em um único volume, o processo será simples. É possível criar um Snapshot a qualquer momento. Se um conjunto de dados abranger volumes, é necessário criar um instantâneo de grupo de consistência (CG). Existem várias opções para criar snapshots CG, incluindo o software NetApp SnapCenter, recursos nativos do grupo de consistência do ONTAP e scripts mantidos pelo usuário.

Backups consistentes com falhas são usados principalmente quando a recuperação do ponto de backup é suficiente. Quando uma recuperação mais granular é necessária, backups consistentes com aplicações geralmente são necessários.

A palavra "consistente" em "consistente com aplicativos" é muitas vezes um erro. Por exemplo, colocar um banco de dados Oracle no modo de backup é referido como um backup consistente com aplicativos, mas os dados não são consistentes ou silenciosos de forma alguma. Os dados continuam a mudar durante todo o backup. Em contraste, a maioria dos backups MySQL e Microsoft SQL Server realmente silenciam os dados antes de executar o backup. A VMware pode ou não tornar certos arquivos consistentes.

Grupos de consistência

O termo "grupo de consistência" refere-se à capacidade de um storage array gerenciar vários recursos de armazenamento como uma única imagem. Por exemplo, um banco de dados pode consistir em 10 LUNs. O array deve ser capaz de fazer backup, restaurar e replicar esses 10 LUNs de maneira consistente. A restauração não é possível se as imagens dos LUNs não forem consistentes no ponto de backup. A replicação desses 10 LUNs requer que todas as réplicas sejam perfeitamente sincronizadas umas com as outras.

O termo "grupo de consistência" não é frequentemente usado quando se discute ONTAP porque consistência sempre foi uma função básica da arquitetura de volume e agregado dentro do ONTAP. Muitos outros storage arrays gerenciam LUNs ou sistemas de arquivos como unidades individuais. Eles poderiam, então, ser opcionalmente configurados como um "grupo de consistência" para fins de proteção de dados, mas esta é uma etapa extra na configuração.

O ONTAP sempre foi capaz de capturar imagens de dados locais e replicados consistentes. Embora os vários volumes em um sistema ONTAP não sejam geralmente formalmente descritos como um grupo de consistência, é isso que eles são. Um snapshot desse volume é uma imagem de grupo de consistência, a restauração para esse snapshot é uma restauração de grupo de consistência, e o SnapMirror e o SnapVault oferecem replicação de grupo de consistência.

Instantâneos do grupo de consistência

Os snapshots de grupos de consistência (snapshots cg) são uma extensão da tecnologia básica do ONTAP Snapshot. Uma operação de snapshot padrão cria uma imagem consistente de todos os dados em um único volume, mas às vezes é necessário criar um conjunto consistente de snapshots em vários volumes e até

mesmo em vários sistemas de storage. O resultado é um conjunto de instantâneos que podem ser usados da mesma forma que um instantâneo de apenas um volume individual. Eles podem ser usados para recuperação de dados local, replicados para fins de recuperação de desastres ou clonados como uma única unidade consistente.

O maior uso conhecido de snapshots cg é para um ambiente de banco de dados de aproximadamente 1PB TB de tamanho abrangendo 12 controladoras. Os snapshots cg criados neste sistema foram usados para backup, recuperação e clonagem.

Na maioria das vezes, quando um conjunto de dados abrange volumes e ordem de gravação deve ser preservado, um cg-snapshot é usado automaticamente pelo software de gerenciamento escolhido. Nesses casos, não há necessidade de entender os detalhes técnicos dos instantâneos cg. No entanto, há situações em que requisitos complicados de proteção de dados exigem controle detalhado sobre o processo de replicação e proteção de dados. Fluxos de trabalho de automação ou o uso de scripts personalizados para chamar APIs cg-snapshot são algumas das opções. Compreender a melhor opção e o papel do cg-snapshot requer uma explicação mais detalhada da tecnologia.

A criação de um conjunto de instantâneos cg é um processo de duas etapas:

1. Estabeleça cercas de gravação em todos os volumes de destino.
2. Crie instantâneos desses volumes enquanto estiver no estado cercado.

A esgrima de escrita é estabelecida em série. Isso significa que, à medida que o processo de esgrima é configurado em vários volumes, a e/S de gravação é congelada no primeiro volume da sequência, uma vez que continua a ser comprometida com volumes que aparecem mais tarde. Isso pode inicialmente parecer violar o requisito para que a ordem de gravação seja preservada, mas isso só se aplica a e/S que é emitida assincronamente no host e não depende de outras gravações.

Por exemplo, um banco de dados pode emitir muitas atualizações assíncronas de arquivos de dados e permitir que o sistema operacional reordene a e/S e as complete de acordo com sua própria configuração de agendador. A ordem deste tipo de e/S não pode ser garantida porque a aplicação e o sistema operativo já lançaram a exigência de preservar a ordem de escrita.

Como um exemplo de contador, a maioria das atividades de Registro de banco de dados é síncrona. O banco de dados não prossegue com outras gravações de log até que a e/S seja reconhecida, e a ordem dessas gravações deve ser preservada. Se uma e/S de log chegar em um volume cercado, ela não será reconhecida e a aplicação será bloqueada em outras gravações. Da mesma forma, a e/S de metadados do sistema de arquivos geralmente é síncrona. Por exemplo, uma operação de exclusão de arquivos não deve ser perdida. Se um sistema operacional com um sistema de arquivos xfs excluísse um arquivo e a e/S que atualizasse os metadados do sistema de arquivos xfs para remover a referência a esse arquivo aterrado em um volume cercado, a atividade do sistema de arquivos pausará. Isso garante a integridade do sistema de arquivos durante operações cg-snapshot.

Depois que o grima de gravação é configurado nos volumes de destino, eles estão prontos para a criação de snapshot. Os instantâneos não precisam ser criados exatamente ao mesmo tempo porque o estado dos volumes é congelado de um ponto de vista de gravação dependente. Para se proteger contra uma falha na aplicação criando os instantâneos cg, a esgrima de gravação inicial inclui um tempo limite configurável no qual o ONTAP libera automaticamente a esgrima e retoma o processamento de gravação após um número definido de segundos. Se todos os instantâneos forem criados antes do período de tempo limite expirar, o conjunto de instantâneos resultante será um grupo de consistência válido.

Ordem de escrita dependente

Do ponto de vista técnico, a chave para um grupo de consistência é preservar a ordem de gravação e, especificamente, a ordem de gravação dependente. Por exemplo, um banco de dados gravando em 10 LUNs

grava simultaneamente em todos eles. Muitas gravações são emitidas assincronamente, o que significa que a ordem em que são concluídas não é importante e a ordem real que são concluídas varia de acordo com o sistema operacional e o comportamento da rede.

Algumas operações de gravação devem estar presentes no disco antes que o banco de dados possa continuar com gravações adicionais. Essas operações críticas de gravação são chamadas de gravações dependentes. A e/S de gravação subsequente depende da presença dessas gravações no disco. Qualquer snapshot, recuperação ou replicação desses 10 LUNs deve garantir que a ordem de gravação dependente seja garantida. As atualizações do sistema de arquivos são outro exemplo de gravações dependentes da ordem de gravação. A ordem em que as alterações do sistema de arquivos são feitas deve ser preservada ou todo o sistema de arquivos pode ficar corrompido.

Estratégias

Há duas abordagens principais para backups baseados em snapshot:

- Backups consistentes com falhas
- Backups ativos protegidos por snapshot

Um backup consistente com falhas de um banco de dados refere-se à captura de toda a estrutura do banco de dados, incluindo datafiles, logs de refazer e arquivos de controle, em um único ponto no tempo. Se o banco de dados for armazenado em um único volume, o processo será simples; uma captura Instantânea pode ser criada a qualquer momento. Se um banco de dados abranger volumes, um snapshot de grupo de consistência (CG) deve ser criado. Existem várias opções para criar snapshots CG, incluindo o software NetApp SnapCenter, recursos nativos do grupo de consistência do ONTAP e scripts mantidos pelo usuário.

Os backups Snapshot consistentes com falhas são usados principalmente quando a recuperação do ponto de backup é suficiente. Registros de arquivo podem ser aplicados em algumas circunstâncias, mas quando uma recuperação pontual mais granular é necessária, um backup on-line é preferível.

O procedimento básico para um backup on-line baseado em snapshot é o seguinte:

1. Coloque a base de dados `backup` no modo.
2. Crie um instantâneo de todos os volumes que hospedam datafiles.
3. Sair `backup` do modo.
4. Execute o comando `alter system archive log current` para forçar o arquivamento de logs.
5. Crie instantâneos de todos os volumes que hospedam os logs do arquivo.

Este procedimento produz um conjunto de instantâneos contendo datafiles no modo de backup e os logs críticos de arquivo gerados no modo de backup. Estes são os dois requisitos para recuperar um banco de dados. Arquivos como arquivos de controle também devem ser protegidos por conveniência, mas o único requisito absoluto é a proteção para arquivos de dados e logs de arquivo.

Embora clientes diferentes possam ter estratégias muito diferentes, quase todas essas estratégias são baseadas nos mesmos princípios descritos abaixo.

Recuperação baseada em Snapshot

Ao projetar layouts de volume para bancos de dados Oracle, a primeira decisão é se usar a tecnologia NetApp SnapRestore baseada em volume (VBSR).

O SnapRestore baseado em volume permite que um volume seja revertido quase instantaneamente para um

ponto anterior no tempo. Como todos os dados no volume são revertidos, o VBSR pode não ser apropriado para todos os casos de uso. Por exemplo, se um banco de dados inteiro, incluindo datafiles, refazer logs e Registros de arquivamento, for armazenado em um único volume e esse volume for restaurado com VBSR, os dados serão perdidos porque o Registro de arquivo mais recente e os dados de refazer são descartados.

VBSR não é necessário para restaurar. Muitos bancos de dados podem ser restaurados usando o SnapRestore de arquivo único (SFSR) baseado em arquivo ou simplesmente copiando arquivos do snapshot de volta para o sistema de arquivos ativo.

O VBSR é preferido quando um banco de dados é muito grande ou quando ele deve ser recuperado o mais rápido possível, e o uso do VBSR requer isolamento dos arquivos de dados. Em um ambiente NFS, os arquivos de dados de um determinado banco de dados devem ser armazenados em volumes dedicados que não estejam contaminados por qualquer outro tipo de arquivo. Em um ambiente SAN, os arquivos de dados devem ser armazenados em LUNs dedicados em volumes dedicados. Se um gerenciador de volumes for usado (incluindo Oracle Automatic Storage Management [ASM]), o grupo de discos também deve ser dedicado a arquivos de dados.

Isolar datafiles desta maneira permite que eles sejam revertidos para um estado anterior sem danificar outros sistemas de arquivos.

Reserva do Snapshot

Para cada volume com dados Oracle em um ambiente SAN, o `percent-snapshot-space` deve ser definido como zero porque reservar espaço para um snapshot em um ambiente LUN não é útil. Se a reserva fracionária estiver definida como 100, um instantâneo de um volume com LUNs requer espaço livre suficiente no volume, excluindo a reserva instantânea, para absorver 100% de rotatividade de todos os dados. Se a reserva fracionária for definida para um valor mais baixo, uma quantidade correspondente menor de espaço livre será necessária, mas sempre exclui a reserva instantânea. Isso significa que o espaço de reserva do snapshot em um ambiente LUN é desperdiçado.

Em um ambiente NFS, há duas opções:

- Defina o `percent-snapshot-space` com base no consumo de espaço esperado do instantâneo.
- Defina o `percent-snapshot-space` como zero e gerencie o consumo de espaço ativo e instantâneo coletivamente.

Com a primeira opção, `percent-snapshot-space` é definido para um valor diferente de zero, normalmente em torno de 20%. Este espaço é então escondido do usuário. Esse valor não cria, no entanto, um limite de utilização. Se um banco de dados com uma reserva de 20% sofrer 30% de rotatividade, o espaço instantâneo pode crescer além dos limites da reserva de 20% e ocupar espaço não reservado.

O principal benefício de definir uma reserva para um valor como 20% é verificar se algum espaço está sempre disponível para instantâneos. Por exemplo, um volume 1TB com uma reserva de 20% permitiria apenas que um administrador de banco de dados (DBA) armazenasse 800GB TB de dados. Essa configuração garante pelo menos 200GBMB de espaço para consumo de snapshot.

```
`percent-snapshot-space`Quando está definido como zero, todo o espaço no volume está disponível para o usuário final, o que proporciona melhor visibilidade. O DBA deve entender que, se ele ou ela vir um volume de 1TB TB que aproveita snapshots, esse 1TB TB de espaço será compartilhado entre dados ativos e a rotatividade do Snapshot.
```

Não há preferência clara entre a opção um e a opção dois entre os usuários finais.

ONTAP e snapshots de terceiros

O Oracle Doc ID 604683,1 explica os requisitos para suporte a instantâneos de terceiros e as várias opções disponíveis para operações de backup e restauração.

O fornecedor terceirizado deve garantir que os snapshots da empresa estejam em conformidade com os seguintes requisitos:

- Os snapshots devem ser integrados às operações de restauração e recuperação recomendadas pela Oracle.
- Os snapshots devem ser consistentes com falhas de banco de dados no ponto do snapshot.
- A ordenação de gravação é preservada para cada arquivo dentro de um snapshot.

Os produtos de gerenciamento ONTAP e NetApp da Oracle atendem a esses requisitos.

SnapRestore

A restauração rápida de dados no ONTAP a partir de um snapshot é fornecida pela tecnologia NetApp SnapRestore.

Quando um conjunto de dados essencial não está disponível, as operações de negócios essenciais estão inativas. As fitas podem quebrar, e até mesmo as restaurações de backups baseados em disco podem ser lentas para serem transferidas pela rede. O SnapRestore evita esses problemas fornecendo restauração quase instantânea de conjuntos de dados. Mesmo os bancos de dados em escala de petabyte podem ser completamente restaurados com apenas alguns minutos de esforço.

Existem duas formas de SnapRestore - baseado em arquivo/LUN e baseado em volume.

- Arquivos individuais ou LUNs podem ser restaurados em segundos, seja um arquivo 2TB LUN ou 4KB.
- O volume de arquivos ou LUNs pode ser restaurado em segundos, seja 10GB ou 100TB TB de dados.

Um "contentor de arquivos ou LUNs" normalmente se referiria a um FlexVol volume. Por exemplo, você pode ter 10 LUNs que compõem um grupo de discos LVM em um único volume ou um volume pode armazenar os diretórios base NFS de usuários do 1000. Em vez de executar uma operação de restauração para cada arquivo individual ou LUN, você pode restaurar todo o volume como uma única operação. Esse processo também funciona com contêineres com escalabilidade horizontal que incluem vários volumes, como um FlexGroup ou um Grupo de consistência do ONTAP.

O motivo pelo qual o SnapRestore funciona tão rápido e eficientemente é devido à natureza de um snapshot, que é essencialmente uma visualização paralela somente leitura do conteúdo de um volume em um determinado momento. Os blocos ativos são os blocos reais que podem ser alterados, enquanto o snapshot é uma visualização somente leitura no estado dos blocos que constituem os arquivos e LUNs no momento em que o snapshot foi criado.

O ONTAP só permite acesso somente leitura a dados instantâneos, mas os dados podem ser reativados com o SnapRestore. O instantâneo é reativado como uma visualização de leitura e gravação dos dados, retornando os dados ao seu estado anterior. O SnapRestore pode operar no volume ou no nível do arquivo. A tecnologia é essencialmente a mesma com algumas pequenas diferenças de comportamento.

Volume SnapRestore

O SnapRestore baseado em volume retorna todo o volume de dados para um estado anterior. Essa operação não requer movimentação de dados, o que significa que o processo de restauração é essencialmente instantâneo, embora a operação de API ou CLI possa levar alguns segundos para ser processada. Restaurar 1GB TB de dados não é mais complicado ou demorado do que restaurar 1PB TB de dados. Essa funcionalidade é a principal razão pela qual muitos clientes empresariais migram para os sistemas de storage da ONTAP. Ele fornece um rto medido em segundos, até mesmo para os maiores conjuntos de dados.

Uma desvantagem para o SnapRestore baseado em volume é causada pelo fato de que as alterações dentro de um volume são cumulativas ao longo do tempo. Portanto, cada snapshot e os dados de arquivo ativos dependem das alterações que levam a esse ponto. Reverter um volume para um estado anterior significa descartar todas as alterações subsequentes que foram feitas aos dados. O que é menos óbvio, no entanto, é que isso inclui instantâneos criados posteriormente. Isso nem sempre é desejável.

Por exemplo, um SLA de retenção de dados pode especificar 30 dias de backups noturnos. Restaurar um conjunto de dados para um instantâneo criado há cinco dias com o volume SnapRestore descartaria todos os snapshots criados nos cinco dias anteriores, violando o SLA.

Existem várias opções disponíveis para resolver esta limitação:

1. Os dados podem ser copiados de um snapshot anterior, em vez de executar um SnapRestore de todo o volume. Esse método funciona melhor com conjuntos de dados menores.
2. Um snapshot pode ser clonado em vez de restaurado. A limitação a essa abordagem é que o snapshot de origem é uma dependência do clone. Portanto, ele não pode ser excluído a menos que o clone também seja excluído ou seja dividido em um volume independente.
3. Uso de SnapRestore baseado em arquivos.

File SnapRestore (ficheiro)

O SnapRestore baseado em arquivo é um processo de restauração mais granular baseado em snapshot. Em vez de reverter o estado de um volume inteiro, o estado de um arquivo individual ou LUN é revertido. Não é necessário eliminar instantâneos, nem esta operação cria qualquer dependência de um instantâneo anterior. O ficheiro ou LUN fica imediatamente disponível no volume ativo.

Nenhuma movimentação de dados é necessária durante uma restauração do SnapRestore de um arquivo ou LUN. No entanto, algumas atualizações internas de metadados são necessárias para refletir o fato de que os blocos subjacentes em um arquivo ou LUN agora existem em um snapshot e no volume ativo. Não deve haver efeito no desempenho, mas esse processo bloqueia a criação de snapshots até que ele esteja concluído. A taxa de processamento é de aproximadamente 5Gbps (18TBMB/hora) com base no tamanho total dos arquivos restaurados.

Backups online

Dois conjuntos de dados são necessários para proteger e recuperar um banco de dados Oracle no modo de backup. Note que esta não é a única opção de backup Oracle, mas é a mais comum.

- Um instantâneo dos arquivos de dados no modo de backup
- Os logs de arquivo criados enquanto os datafiles estavam no modo de backup

Se a recuperação completa, incluindo todas as transações confirmadas, é necessário um terceiro item:

- Um conjunto de registos de refazer atuais

Existem várias maneiras de impulsionar a recuperação de um backup on-line. Muitos clientes restauram snapshots usando a CLI do ONTAP e, em seguida, usando o Oracle RMAN ou sqlplus para concluir a recuperação. Isso é especialmente comum com grandes ambientes de produção em que a probabilidade e a frequência de restaurações de banco de dados são extremamente baixas e qualquer procedimento de restauração é Tratado por um DBA qualificado. Para automação completa, soluções como o NetApp SnapCenter incluem um plug-in Oracle com interfaces gráficas e de linha de comando.

Alguns clientes de grande escala adotaram uma abordagem mais simples, configurando scripts básicos nos hosts para colocar os bancos de dados no modo de backup em um momento específico, em preparação para um snapshot agendado. Por exemplo, programe o comando `alter database begin backup` às 23:58, `alter database end backup` às 00:02 e, em seguida, programe instantâneos diretamente no sistema de armazenamento à meia-noite. O resultado é uma estratégia de backup simples e altamente dimensionável que não requer software ou licenças externos.

Layout de dados

O layout mais simples é isolar datafiles em um ou mais volumes dedicados. Eles devem ser não contaminados por qualquer outro tipo de arquivo. Isso é para garantir que os volumes de arquivo de dados possam ser restaurados rapidamente através de uma operação SnapRestore sem destruir um log refazer importante, controlfile ou log de arquivo.

A SAN tem requisitos semelhantes para isolamento de arquivos de dados dentro de volumes dedicados. Com um sistema operacional como o Microsoft Windows, um único volume pode conter vários LUNs de arquivo de dados, cada um com um sistema de arquivos NTFS. Com outros sistemas operacionais, geralmente há um gerenciador de volumes lógico. Por exemplo, com o Oracle ASM, a opção mais simples seria limitar os LUNs de um grupo de discos ASM a um único volume que pode ser feito backup e restaurado como uma unidade. Se forem necessários volumes adicionais por motivos de gerenciamento de performance ou capacidade, a criação de um grupo de discos adicional no novo volume resultará em um gerenciamento mais simples.

Se essas diretrizes forem seguidas, os snapshots poderão ser agendados diretamente no sistema de storage sem a necessidade de realizar um snapshot de grupo de consistência. A razão é que os backups Oracle não exigem que os datafiles sejam copiados ao mesmo tempo. O procedimento de backup on-line foi projetado para permitir que os arquivos de dados continuem sendo atualizados, pois são transmitidos lentamente para a fita ao longo de horas.

Uma complicação surge em situações como o uso de um grupo de discos ASM que é distribuído entre volumes. Nesses casos, um cg-snapshot deve ser executado para garantir que os metadados ASM sejam consistentes em todos os volumes constituintes.

Atenção: Verifique se o ASM `spfile` e `passwd` os arquivos não estão no grupo de discos que hospeda os arquivos de dados. Isso interfere na capacidade de restaurar seletivamente datafiles e apenas datafiles.

Procedimento de recuperação local – NFS

Este procedimento pode ser conduzido manualmente ou através de uma aplicação como o SnapCenter. O procedimento básico é o seguinte:

1. Encerre o banco de dados.
2. Recupere o(s) volume(s) de arquivo de dados para o instantâneo imediatamente antes do ponto de restauração desejado.
3. Reproduza registos de arquivo até ao ponto pretendido.

4. Repita os logs atuais de refazer se a recuperação completa for desejada.

Este procedimento pressupõe que os registos de arquivo desejados ainda estão presentes no sistema de ficheiros ativo. Se não estiverem, os logs do arquivo devem ser restaurados ou rman/sqlplus podem ser direcionados para os dados no diretório instantâneo.

Além disso, para bancos de dados menores, os arquivos de dados podem ser recuperados por um usuário final diretamente `.snapshot` do diretório sem a ajuda de ferramentas de automação ou administradores de armazenamento para executar um `snaprestore` comando.

Procedimento de recuperação local – SAN

Este procedimento pode ser conduzido manualmente ou através de uma aplicação como o SnapCenter. O procedimento básico é o seguinte:

1. Encerre o banco de dados.
2. Quiesce o(s) grupo(s) de discos que hospedam os arquivos de dados. O procedimento varia consoante o gestor de volume lógico escolhido. Com ASM, o processo requer a desmontagem do grupo de discos. Com o Linux, os sistemas de arquivos devem ser desmontados e os volumes lógicos e grupos de volumes devem ser desativados. O objetivo é parar todas as atualizações no grupo de volume alvo a serem restauradas.
3. Restaure os grupos de discos de arquivo de dados para o instantâneo imediatamente antes do ponto de restauração desejado.
4. Reative os grupos de discos recentemente restaurados.
5. Reproduza registos de arquivo até ao ponto pretendido.
6. Repita todos os logs de refazer se a recuperação completa for desejada.

Este procedimento pressupõe que os registos de arquivo desejados ainda estão presentes no sistema de ficheiros ativo. Se não estiverem, os registos de arquivo devem ser restaurados colocando os LUNs de registo de arquivo offline e executando uma restauração. Este também é um exemplo no qual dividir os logs de arquivo em volumes dedicados é útil. Se os logs de arquivo compartilharem um grupo de volumes com os logs de refazer, os logs de refazer devem ser copiados em outro lugar antes da restauração do conjunto geral de LUNs. Esta etapa impede a perda dessas transações finais registradas.

Backups otimizados para Storage Snapshot

Backup e recuperação baseados em snapshot se tornaram ainda mais simples quando o Oracle 12c foi lançado porque não há necessidade de colocar um banco de dados no modo hot backup. O resultado é a capacidade de agendar backups baseados em snapshot diretamente em um sistema de storage e ainda preservar a capacidade de executar recuperação completa ou pontual.

Embora o procedimento de recuperação de hot backup seja mais familiar aos DBAs, há muito tempo foi possível usar snapshots que não foram criados enquanto o banco de dados estava no modo hot backup. Etapas manuais extras foram necessárias com o Oracle 10gi e 11gi durante a recuperação para tornar o banco de dados consistente. Com o Oracle 12ci, `sqlplus` e `rman` conter a lógica extra para reproduzir logs de arquivo em backups de arquivos de dados que não estavam no modo de backup ativo.

Como discutido anteriormente, a recuperação de um hot backup baseado em snapshot requer dois conjuntos de dados:

- Um instantâneo dos arquivos de dados criados no modo de backup
- Os logs de arquivo gerados enquanto os datafiles estavam no modo hot backup

Durante a recuperação, o banco de dados lê metadados dos arquivos de dados para selecionar os logs de arquivo necessários para recuperação.

A recuperação otimizada para snapshot de storage requer conjuntos de dados ligeiramente diferentes para alcançar os mesmos resultados:

- Um instantâneo dos arquivos de dados, além de um método para identificar a hora em que o snapshot foi criado
- Arquive logs do tempo do ponto de verificação mais recente do arquivo de dados até a hora exata do instantâneo

Durante a recuperação, o banco de dados lê metadados dos arquivos de dados para identificar o Registro de arquivo mais antigo necessário. A recuperação completa ou pontual pode ser realizada. Ao executar uma recuperação pontual, é fundamental saber o tempo do snapshot dos arquivos de dados. O ponto de recuperação especificado deve ser após o tempo de criação dos instantâneos. A NetApp recomenda adicionar pelo menos alguns minutos à hora do instantâneo para contabilizar a variação do relógio.

Para obter detalhes completos, consulte a documentação da Oracle sobre o tópico "recuperação usando Otimização de Snapshot de armazenamento" disponível em várias versões da documentação do Oracle 12c. Além disso, consulte Oracle Document ID Doc ID 604683,1 sobre o suporte a snapshots de terceiros da Oracle.

Layout de dados

O layout mais simples é isolar os arquivos de dados em um ou mais volumes dedicados. Eles devem ser não contaminados por qualquer outro tipo de arquivo. Isso é para garantir que os volumes de arquivo de dados possam ser restaurados rapidamente com uma operação SnapRestore sem destruir um log de refazer importante, controlfile ou log de arquivo.

A SAN tem requisitos semelhantes para isolamento de arquivos de dados dentro de volumes dedicados. Com um sistema operacional como o Microsoft Windows, um único volume pode conter vários LUNs de arquivo de dados, cada um com um sistema de arquivos NTFS. Com outros sistemas operacionais, geralmente há um gerenciador de volume lógico também. Por exemplo, com o Oracle ASM, a opção mais simples seria limitar grupos de discos a um único volume que pode ser feito backup e restaurado como uma unidade. Se forem necessários volumes adicionais por motivos de gerenciamento de performance ou capacidade, criar um grupo de discos adicional no novo volume resulta em gerenciamento mais fácil.

Se essas diretrizes forem seguidas, os snapshots poderão ser agendados diretamente no ONTAP sem a necessidade de realizar um snapshot de grupo de consistência. O motivo é que backups otimizados para snapshot não exigem que sejam feitos backup de dados ao mesmo tempo.

Uma complicação surge em situações como um grupo de discos ASM que é distribuído entre volumes. Nesses casos, um cg-snapshot deve ser executado para garantir que os metadados ASM sejam consistentes em todos os volumes constituintes.

[Nota]Verifique se os arquivos ASM spfile e passwd não estão no grupo de discos que hospeda os arquivos de dados. Isso interfere na capacidade de restaurar seletivamente datafiles e apenas datafiles.

Procedimento de recuperação local – NFS

Este procedimento pode ser conduzido manualmente ou através de uma aplicação como o SnapCenter. O

procedimento básico é o seguinte:

1. Encerre o banco de dados.
2. Recupere o(s) volume(s) de arquivo de dados para o instantâneo imediatamente antes do ponto de restauração desejado.
3. Reproduza registos de arquivo até ao ponto pretendido.

Este procedimento pressupõe que os registos de arquivo desejados ainda estão presentes no sistema de ficheiros ativo. Se não estiverem, os registos de arquivo têm de ser restaurados ou `rman` `sqlplus` podem ser direcionados para os dados no `.snapshot` diretório.

Além disso, para bancos de dados menores, os arquivos de dados podem ser recuperados por um usuário final diretamente `.snapshot` do diretório sem a ajuda de ferramentas de automação ou um administrador de armazenamento para executar um comando SnapRestore.

Procedimento de recuperação local – SAN

Este procedimento pode ser conduzido manualmente ou através de uma aplicação como o SnapCenter. O procedimento básico é o seguinte:

1. Encerre o banco de dados.
2. Quiesce o(s) grupo(s) de discos que hospedam os arquivos de dados. O procedimento varia consoante o gestor de volume lógico escolhido. Com ASM, o processo requer a desmontagem do grupo de discos. Com o Linux, os sistemas de arquivos devem ser desmontados e os volumes lógicos e grupos de volumes são desativados. O objetivo é parar todas as atualizações no grupo de volume alvo a serem restauradas.
3. Restaure os grupos de discos de arquivo de dados para o instantâneo imediatamente antes do ponto de restauração desejado.
4. Reative os grupos de discos recentemente restaurados.
5. Reproduza registos de arquivo até ao ponto pretendido.

Este procedimento pressupõe que os registos de arquivo desejados ainda estão presentes no sistema de ficheiros ativo. Se não estiverem, os registos de arquivo devem ser restaurados colocando os LUNs de registo de arquivo offline e executando uma restauração. Este também é um exemplo no qual dividir os logs de arquivo em volumes dedicados é útil. Se os logs do arquivo compartilharem um grupo de volumes com os logs de refazer, os logs de refazer devem ser copiados em outro lugar antes da restauração do conjunto geral de LUNs para evitar perder as transações registradas finais.

Exemplo de recuperação completa

Suponha que os arquivos de dados foram corrompidos ou destruídos e a recuperação completa é necessária. O procedimento para o fazer é o seguinte:

```
[oracle@host1 ~]$ sqlplus / as sysdba
Connected to an idle instance.
SQL> startup mount;
ORACLE instance started.
Total System Global Area 1610612736 bytes
Fixed Size                  2924928 bytes
Variable Size            1040191104 bytes
Database Buffers          553648128 bytes
Redo Buffers               13848576 bytes
Database mounted.
SQL> recover automatic;
Media recovery complete.
SQL> alter database open;
Database altered.
SQL>
```

Exemplo de recuperação pontual

Todo o procedimento de recuperação é um único comando: `recover automatic`.

Se a recuperação pontual for necessária, o carimbo de data/hora do(s) instantâneo(s) deve(m) ser conhecido(s) e pode(m) ser identificado(s) da seguinte forma:

```
Cluster01::> snapshot show -vserver vserver1 -volume NTAP_oradata -fields
create-time
vserver   volume        snapshot   create-time
--------  ------------  ---------  ------------------------
vserver1  NTAP_oradata  my-backup  Thu Mar 09 10:10:06 2017
```

A hora de criação do instantâneo é listada como 9th de Março e 10:10:06. Para estar seguro, um minuto é adicionado à hora do instantâneo:

```
[oracle@host1 ~]$ sqlplus / as sysdba
Connected to an idle instance.
SQL> startup mount;
ORACLE instance started.
Total System Global Area 1610612736 bytes
Fixed Size                  2924928 bytes
Variable Size            1040191104 bytes
Database Buffers          553648128 bytes
Redo Buffers               13848576 bytes
Database mounted.
SQL> recover database until time '09-MAR-2017 10:44:15' snapshot time '09-
MAR-2017 10:11:00';
```

A recuperação agora é iniciada. Ele especificou um tempo instantâneo de 10:11:00, um minuto após o tempo gravado para contabilizar a possível variação do relógio e um tempo de recuperação alvo de 10:44. Em seguida, sqlplus solicita os logs de arquivo necessários para alcançar o tempo de recuperação desejado de 10:44.

```
ORA-00279: change 551760 generated at 03/09/2017 05:06:07 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_31_930813377.dbf
ORA-00280: change 551760 for thread 1 is in sequence #31
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 552566 generated at 03/09/2017 05:08:09 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_32_930813377.dbf
ORA-00280: change 552566 for thread 1 is in sequence #32
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 553045 generated at 03/09/2017 05:10:12 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_33_930813377.dbf
ORA-00280: change 553045 for thread 1 is in sequence #33
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
ORA-00279: change 753229 generated at 03/09/2017 05:15:58 needed for
thread 1
ORA-00289: suggestion : /oralogs_nfs/arch/1_34_930813377.dbf
ORA-00280: change 753229 for thread 1 is in sequence #34
Specify log: {<RET>=suggested | filename | AUTO | CANCEL}
Log applied.
Media recovery complete.
SQL> alter database open resetlogs;
Database altered.
SQL>
```

A recuperação completa de um banco de dados usando snapshots usando o `recover automatic` comando não requer licenciamento específico, mas a recuperação pontual usando `snapshot time` requer a licença Oracle Advanced Compression.

Ferramentas de gerenciamento e automação de banco de dados

O principal valor do ONTAP em um ambiente de banco de dados Oracle vem das principais tecnologias da ONTAP, como cópias Snapshot instantâneas, replicação simples do SnapMirror e criação eficiente de volumes do FlexClone.

Em alguns casos, a configuração simples desses principais recursos diretamente no ONTAP atende aos requisitos, mas as necessidades mais complicadas exigem uma camada de orquestração.

SnapCenter

O SnapCenter é o principal produto de proteção de dados da NetApp. Em um nível muito baixo, ele é semelhante aos produtos SnapManager em termos de como executa backups de bancos de dados. No entanto, ele foi criado do zero para fornecer um painel único para gerenciamento de proteção de dados em sistemas de storage da NetApp.

O SnapCenter inclui funções básicas, como backups e restaurações baseados em snapshot, replicação SnapMirror e SnapVault e outros recursos necessários para operar em escala para empresas de grande porte. Esses recursos avançados incluem funcionalidade de controle de acesso baseado em funções (RBAC) expandida, APIs RESTful para integração com produtos de orquestração de terceiros, gerenciamento central sem interrupções de plug-ins do SnapCenter em hosts de banco de dados e uma interface de usuário projetada para ambientes de escala de nuvem.

DESCANSO

O ONTAP também contém um conjunto de APIs RESTful. Isso permite que fornecedores terceirizados de 3rd criem proteção de dados e outros aplicativos de gerenciamento com profunda integração com o ONTAP. Além disso, a API RESTful é fácil de consumir por clientes que desejam criar seus próprios fluxos de trabalho e utilitários de automação.

Informações sobre direitos autorais

Informações sobre marcas comerciais