



Configuração de host com sistemas AFF/ FAS

Enterprise applications

NetApp

February 11, 2026

Índice

Configuração de host com sistemas AFF/ FAS	1
AIX	1
E/S simultânea	1
Opções de montagem AIX NFS	1
AIX jfs/jfs2 Opções de montagem	2
HP-UX	2
Opções de montagem NFS HP-UX	3
Opções de montagem HP-UX VxFS	4
Linux	4
Tabelas de slots TCP do Linux NFSv3	4
Opções de montagem em NFS do Linux	5
Espelhamento ASM	7
Opções de montagem Linux xfs, ext3 e ext4	8
ASMLib/AFD (controlador de filtro ASM)	8
Tamanhos de blocos ASMLib	8
Tamanhos de bloco do Acionamento do filtro ASM (AFD)	9
Microsoft Windows	10
NFS	10
SAN	10
Solaris	10
Opções de montagem do Solaris NFS	11
Opções de montagem do Solaris UFS	12
Solaris ZFS	12
Kernel	12
Configuração LUN	12

Configuração de host com sistemas AFF/ FAS

AIX

Tópicos de configuração para banco de dados Oracle no IBM AIX com ONTAP.

E/S simultânea

Alcançar o desempenho ideal no IBM AIX requer o uso de e/S simultâneas. Sem I/o concorrente, as limitações de desempenho provavelmente são porque o AIX executa e/S atômica serializada, o que incorre em sobrecarga significativa.

Originalmente, o NetApp recomendou o uso da `cio` opção de montagem para forçar o uso de e/S concorrente no sistema de arquivos, mas esse processo teve desvantagens e não é mais necessário. Desde a introdução do AIX 5,2 e do Oracle 10gR1, o Oracle no AIX pode abrir arquivos individuais para IO simultâneo, em vez de forçar e/S simultânea em todo o sistema de arquivos.

O melhor método para ativar e/S concorrente é definir o `init.ora` parâmetro `filesystemio_options` como `setall`. Isso permite que a Oracle abra arquivos específicos para uso com e/S concorrente.

Usar `cio` como opção de montagem força o uso de e/S concorrente, o que pode ter consequências negativas. Por exemplo, forçar e/S concorrente desativa o readahead em sistemas de arquivos, o que pode danificar o desempenho de e/S que ocorre fora do software de banco de dados Oracle, como copiar arquivos e executar backups de fita. Além disso, produtos como Oracle GoldenGate e SAP BR*Tools não são compatíveis com o uso da `cio` opção de montagem com certas versões do Oracle.

A NetApp recomenda o seguinte:

- Não use a `cio` opção de montagem no nível do sistema de arquivos. Em vez disso, ative a e/S concorrente através do uso `filesystemio_options=setall`.
- Utilize apenas a `cio` opção de montagem se não for possível definir `filesystemio_options=setall`.

Opções de montagem AIX NFS

A tabela a seguir lista as opções de montagem NFS AIX para bancos de dados de instância única Oracle.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144</code>
Registros do Redo ControlFiles Datafiles	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144</code>
ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, intr</code>

A tabela a seguir lista as opções de montagem NFS AIX para RAC.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144</code>
Registros do Redo ControlFiles Datafiles	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr,noac</code>
CRS/Voting	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr,noac</code>
Dedicado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144</code>
Compartilhado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr</code>

A principal diferença entre as opções de montagem de instância única e RAC é a adição `noac` das opções de montagem. Essa adição tem o efeito de desabilitar o cache do sistema operacional do host, que permite que todas as instâncias do cluster RAC tenham uma visão consistente do estado dos dados.

Embora o uso da `cio` opção de montagem e do `init.ora` parâmetro `filesystemio_options=setall` tenha o mesmo efeito de desabilitar o cache do host, ainda é necessário usar `noac`o . `noac`. É necessário para implantações compartilhadas ORACLE_HOME para facilitar a consistência de arquivos, como arquivos de senha Oracle e spfile arquivos de parâmetros. Se cada instância em um cluster RAC tiver um dedicado ORACLE_HOME, esse parâmetro não será necessário.

AIX jfs/jfs2 Opções de montagem

A tabela a seguir lista as opções de montagem do AIX jfs/jfs2.

Tipo de ficheiro	Opções de montagem
ADR Home	Predefinições
Registros do Redo ControlFiles Datafiles	Predefinições
ORACLE_HOME	Predefinições

Antes de usar dispositivos AIX `hdisk` em qualquer ambiente, incluindo bancos de dados, verifique o parâmetro `queue_depth`. Este parâmetro não é a profundidade da fila HBA; em vez disso, ele se relaciona com a profundidade da fila SCSI do indivíduo `hdisk` device. Depending on how the LUNs are configured, the value for `queue_depth` pode ser muito baixa para um bom desempenho. Os testes mostraram que o valor ideal é 64.

HP-UX

Tópicos de configuração para banco de dados Oracle no HP-UX com ONTAP.

Opções de montagem NFS HP-UX

A tabela a seguir lista as opções de montagem HP-UX NFS para uma única instância.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, suid</code>
Arquivos de controle Datafiles Redo logs	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, forcedirectio, nointr,suid</code>
ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, suid</code>

A tabela a seguir lista as opções de montagem HP-UX NFS para RAC.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, noac,suid</code>
Arquivos de controle Datafiles Redo logs	<code>rw, bg,hard, [vers=3,vers=4.1],proto=tcp, timeo=600,rsize=262144,wsize=262144, nointr,noac,forcedirectio,suid</code>
CRS/votação	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr,noac, forcedirectio,suid</code>
Dedicado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, suid</code>
Compartilhado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr,noac,suid</code>

A principal diferença entre as opções de montagem de uma única instância e RAC é a adição `noac` de e `forcedirectio` às opções de montagem. Essa adição tem o efeito de desabilitar o cache do sistema operacional do host, o que permite que todas as instâncias do cluster RAC tenham uma visão consistente do estado dos dados. Embora o uso do `init.ora` parâmetro `filesystemio_options=setall` tenha o mesmo efeito de desabilitar o cache do host, ainda é necessário usar `noac` e `forcedirectio`.

O motivo `noac` necessário para implantações compartilhadas ORACLE_HOME é facilitar a consistência de arquivos, como arquivos de senha Oracle e spfies. Se cada instância em um cluster RAC tiver um dedicado ORACLE_HOME, esse parâmetro não será necessário.

Opções de montagem HP-UX VxFS

Use as seguintes opções de montagem para sistemas de arquivos que hospedam binários Oracle:

```
delaylog, nodatainlog
```

Use as seguintes opções de montagem para sistemas de arquivos que contêm datafiles, logs de refazer, logs de arquivamento e arquivos de controle nos quais a versão do HP-UX não suporta e/S simultânea:

```
nodatainlog, mincache=direct, convosync=direct
```

Quando a e/S simultânea for suportada (VxFS 5.0.1 e posterior, ou com o ServiceGuard Storage Management Suite), utilize estas opções de montagem para sistemas de ficheiros que contenham ficheiros de dados, registos refeitos, registos de arquivo e ficheiros de controlo:

```
delaylog, cio
```

O parâmetro `db_file_multiblock_read_count` é especialmente crítico em ambientes VxFS. A Oracle recomenda que esse parâmetro permaneça desconfigurado no Oracle 10g R1 e posterior, a menos que especificamente direcionado de outra forma. O padrão com um tamanho de bloco Oracle 8KB é 128. Se o valor deste parâmetro for forçado a 16 ou menos, remova a `convosync=direct` opção de montagem porque pode danificar o desempenho sequencial de e/S. Esta etapa prejudica outros aspectos do desempenho e só deve ser tomada se o valor de `db_file_multiblock_read_count` tiver de ser alterado do valor padrão.

Linux

Tópicos de configuração específicos para o sistema operacional Linux.

Tabelas de slots TCP do Linux NFSv3

As tabelas de slot TCP são equivalentes a NFSv3 mm de profundidade de fila do adaptador de barramento do host (HBA). Essas tabelas controlam o número de operações NFS que podem ficar pendentes de uma só vez. O valor padrão é geralmente 16, o que é muito baixo para um desempenho ideal. O problema oposto ocorre em kernels Linux mais recentes, que podem aumentar automaticamente o limite da tabela de slots TCP para um nível que satura o servidor NFS com solicitações.

Para um desempenho ideal e para evitar problemas de desempenho, ajuste os parâmetros do kernel que controlam as tabelas de slots TCP.

Executar o `sysctl -a | grep tcp.*.slot_table` comando e respeitar os seguintes parâmetros:

```
# sysctl -a | grep tcp.*.slot_table
sunrpc.tcp_max_slot_table_entries = 128
sunrpc.tcp_slot_table_entries = 128
```

Todos os sistemas Linux devem incluir `sunrpc.tcp_slot_table_entries`, mas apenas alguns incluem `sunrpc.tcp_max_slot_table_entries`. Ambos devem ser definidos para 128.

A falha em definir esses parâmetros pode ter efeitos significativos no desempenho. Em alguns casos, o desempenho é limitado porque o sistema operacional linux não está emitindo e/S suficiente. Em outros casos, as latências de e/S aumentam à medida que o sistema operacional linux tenta emitir mais e/S do que pode ser reparado.

Opções de montagem em NFS do Linux

A tabela a seguir lista as opções de montagem NFS do Linux para uma única instância.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144</code>
Arquivos de controle Datafiles Redo logs	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr</code>
ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr</code>

A tabela a seguir lista as opções de montagem NFS do Linux para RAC.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, actimeo=0</code>
Arquivos de controle arquivos de dados Redo logs	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr,actimeo=0</code>
CRS/votação	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr,noac,actimeo=0</code>
Dedicado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144</code>
Compartilhado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr,actimeo=0</code>

A principal diferença entre as opções de montagem de instância única e RAC é a adição `actimeo=0` das opções de montagem. Essa adição tem o efeito de desabilitar o cache do sistema operacional do host, o que permite que todas as instâncias do cluster RAC tenham uma visão consistente do estado dos dados. Embora o uso do `init.ora` parâmetro `filesystemio_options=setall` tenha o mesmo efeito de desabilitar o cache do host, ainda é necessário usar `actimeo=0`.

O motivo `actimeo=0` necessário para implantações compartilhadas ORACLE_HOME é facilitar a consistência

de arquivos, como os arquivos de senha Oracle e spfiles. Se cada instância em um cluster RAC tiver um dedicado ORACLE_HOME, esse parâmetro não será necessário.

Geralmente, arquivos que não sejam de banco de dados devem ser montados com as mesmas opções usadas para datafiles de instância única, embora aplicativos específicos possam ter requisitos diferentes. Evite as opções de montagem noac e actimeo=0, se possível, porque essas opções desativam a leitura e o buffer no nível do sistema de arquivos. Isso pode causar problemas graves de desempenho para processos como extração, tradução e carregamento.

ACESSO e GETATTR

Alguns clientes observaram que um nível extremamente alto de outros IOPS, como O ACCESS e GETATTR, pode dominar suas cargas de trabalho. Em casos extremos, operações como leituras e gravações podem ser tão baixas quanto 10% do total. Este é um comportamento normal com qualquer banco de dados que inclua o uso actimeo=0 e/ou noac no Linux porque essas opções fazem com que o sistema operacional Linux recarregue constantemente metadados de arquivos do sistema de armazenamento. Operações como ACCESS e GETATTR são operações de baixo impacto que são atendidas a partir do cache ONTAP em um ambiente de banco de dados. Não devem ser consideradas IOPS originais, como leituras e gravações, que criem verdadeira demanda em sistemas de storage. No entanto, esses outros IOPS criam alguma carga, especialmente em ambientes RAC. Para resolver esta situação, ative o DNFS, que ignora o cache do buffer do sistema operacional e evita essas operações desnecessárias de metadados.

Linux Direct NFS

Uma opção de montagem adicional, chamada nosharecache, é necessária quando (a) o DNFS está ativado e (b) um volume de origem é montado mais de uma vez em um único servidor (c) com uma montagem NFS aninhada. Essa configuração é vista principalmente em ambientes compatíveis com aplicações SAP. Por exemplo, um único volume em um sistema NetApp pode ter um diretório localizado em /vol/oracle/base e um segundo em /vol/oracle/home. Se /vol/oracle/base for montado em /oracle e /vol/oracle/home for montado em /oracle/home, o resultado serão montagens NFS aninhadas que se originam na mesma fonte.

O sistema operacional pode detectar o fato de que /oracle e /oracle/home residir no mesmo volume, que é o mesmo sistema de arquivos de origem. Em seguida, o SO usa o mesmo identificador de dispositivo para acessar os dados. Isso melhora o uso do cache do sistema operacional e outras operações, mas interfere com o DNFS. Se o DNFS tiver de aceder a um ficheiro, como o spfile, ligado /oracle/home, poderá tentar, erroneamente, utilizar o caminho errado para os dados. O resultado é uma operação de e/S com falha. Nessas configurações, adicione a nosharecache opção de montagem a qualquer sistema de arquivos NFS que compartilhe um volume de origem com outro sistema de arquivos NFS nesse host. Isso força o sistema operacional Linux a alocar um identificador de dispositivo independente para esse sistema de arquivos.

Linux Direct NFS e Oracle RAC

O uso do DNFS tem benefícios especiais de desempenho para o Oracle RAC no sistema operacional Linux porque o Linux não tem um método para forçar e/S direto, o que é necessário com RAC para coerência entre os nós. Como solução alternativa, o Linux requer o uso da actimeo=0 opção de montagem, que faz com que os dados de arquivo expirem imediatamente do cache do sistema operacional. Essa opção, por sua vez, força o cliente NFS Linux a reler constantemente os dados de atributos, o que danifica a latência e aumenta a carga no controlador de armazenamento.

A ativação do DNFS ignora o cliente NFS do host e evita esse dano. Vários clientes relataram melhorias significativas no desempenho em clusters RAC e reduções significativas na carga do ONTAP (especialmente em relação a outros IOPS) ao habilitar o DNFS.

Linux Direct NFS e arquivo oranfstab

Ao usar DNFS no Linux com a opção multipathing, várias sub-redes devem ser usadas. Em outros sistemas operacionais, vários canais DNFS podem ser estabelecidos usando as LOCAL opções e DONTROUTE para configurar vários canais DNFS em uma única sub-rede. No entanto, isso não funciona corretamente no Linux e problemas de desempenho inesperados podem resultar. Com o Linux, cada NIC usada para o tráfego DNFS deve estar em uma sub-rede diferente.

Programador de e/S.

O kernel Linux permite um controle de baixo nível sobre a maneira como e/S para bloquear dispositivos é agendada. Os padrões em várias distribuições do Linux variam consideravelmente. Testes mostram que o prazo geralmente oferece os melhores resultados, mas ocasionalmente o NOOP foi um pouco melhor. A diferença de desempenho é mínima, mas teste ambas as opções se for necessário extrair o máximo desempenho possível de uma configuração de banco de dados. O CFQ é o padrão em muitas configurações e demonstrou problemas significativos de desempenho com cargas de trabalho de banco de dados.

Consulte a documentação relevante do fornecedor do Linux para obter instruções sobre como configurar o agendador de e/S.

Multipathing

Alguns clientes encontraram falhas durante a interrupção da rede porque o daemon multipath não estava sendo executado em seu sistema. Em versões recentes do Linux, o processo de instalação do sistema operacional e do daemon multipathing podem deixar esses sistemas operacionais vulneráveis a esse problema. Os pacotes são instalados corretamente, mas não são configurados para inicialização automática após uma reinicialização.

Por exemplo, o padrão para o daemon multipath no RHEL5,5 pode aparecer da seguinte forma:

```
[root@host1 iscsi]# chkconfig --list | grep multipath
multipathd      0:off    1:off    2:off    3:off    4:off    5:off    6:off
```

Isso pode ser corrigido com os seguintes comandos:

```
[root@host1 iscsi]# chkconfig multipathd on
[root@host1 iscsi]# chkconfig --list | grep multipath
multipathd      0:off    1:off    2:on     3:on     4:on     5:on     6:off
```

Espelhamento ASM

O espelhamento ASM pode exigir alterações nas configurações de multipath do Linux para permitir que o ASM reconheça um problema e alterne para um grupo de falhas alternativo. A maioria das configurações ASM no ONTAP usa redundância externa, o que significa que a proteção de dados é fornecida pelo array externo e o ASM não espelha dados. Alguns sites usam ASM com redundância normal para fornecer espelhamento bidirecional, normalmente em diferentes sites.

As configurações do Linux mostradas na "[Documentação dos utilitários de host do NetApp](#)" incluem parâmetros multipath que resultam em filas indefinidas de e/S Isso significa que uma e/S em um dispositivo LUN sem caminhos ativos aguarda o tempo necessário para que a e/S seja concluída. Isso geralmente é

desejável porque os hosts Linux esperam que as alterações de caminho SAN sejam concluídas, que os switches FC sejam reiniciados ou que um sistema de storage conclua um failover.

Esse comportamento ilimitado de enfileiramento causa um problema com o espelhamento ASM porque o ASM deve receber uma falha de e/S para que ele tente novamente e/S em um LUN alternativo.

Defina os seguintes parâmetros no arquivo Linux `multipath.conf` para LUNs ASM usados com espelhamento ASM:

```
polling_interval 5  
no_path_retry 24
```

Essas configurações criam um tempo limite de 120 segundos para dispositivos ASM. O tempo limite é calculado como `polling_interval * no_path_retry` como segundos. O valor exato pode precisar ser ajustado em algumas circunstâncias, mas um tempo limite de 120 segundos deve ser suficiente para a maioria dos usos. Especificamente, 120 segundos devem permitir que uma tomada de controle ou giveback ocorra sem produzir um erro de e/S que resultaria em que o grupo de falha fosse colocado offline.

Um valor menor `no_path_retry` pode reduzir o tempo necessário para que o ASM alterne para um grupo de falhas alternativo, mas isso também aumenta o risco de um failover indesejado durante atividades de manutenção, como um controle de controle. O risco pode ser atenuado por um monitoramento cuidadoso do estado de espelhamento do ASM. Se ocorrer um failover indesejado, os espelhos podem ser resynced rapidamente se a resincronização for executada de forma relativamente rápida. Para obter informações adicionais, consulte a documentação Oracle sobre ASM Fast Mirror Resync para a versão do software Oracle em uso.

Opções de montagem Linux xfs, ext3 e ext4



A NetApp recomenda usando as opções de montagem padrão.

ASMLib/AFD (controlador de filtro ASM)

Tópicos de configuração específicos para o sistema operacional Linux usando AFD e ASMLib

Tamanhos de blocos ASMLib

ASMLib é uma biblioteca de gerenciamento ASM opcional e utilitários associados. Seu valor principal é a capacidade de carimbar um LUN ou um arquivo baseado em NFS como um recurso ASM com uma etiqueta legível por humanos.

Versões recentes do ASMLib detetam um parâmetro LUN chamado Logical Blocks per Physical Block exponent (LBPPBE). Esse valor não foi reportado pelo destino SCSI ONTAP até recentemente. Ele agora retorna um valor que indica que um tamanho de bloco 4KB é preferido. Esta não é uma definição de tamanho de bloco, mas é uma dica para qualquer aplicativo que usa LBPPBE que I/Os de um determinado tamanho podem ser manipulados de forma mais eficiente. No entanto, o ASMLib interpreta LBPPBE como um tamanho de bloco e marca persistentemente o cabeçalho ASM quando o dispositivo ASM é criado.

Esse processo pode causar problemas com atualizações e migrações de várias maneiras, tudo com base na incapacidade de misturar dispositivos ASMLib com diferentes tamanhos de bloco no mesmo grupo de discos ASM.

Por exemplo, arrays mais antigos geralmente relataram um valor LBPPBE de 0 ou não relataram esse valor de todo. ASMlib interpreta isso como um tamanho de bloco de 512 bytes. Matrizes mais recentes seriam interpretadas como tendo um tamanho de bloco 4KB. Não é possível misturar dispositivos de 512 bytes e 4KB no mesmo grupo de discos ASM. Isso bloquearia um usuário de aumentar o tamanho do grupo de discos ASM usando LUNs de dois arrays ou utilizando ASM como uma ferramenta de migração. Em outros casos, o RMAN pode não permitir a cópia de arquivos entre um grupo de discos ASM com um tamanho de bloco de 512 bytes e um grupo de discos ASM com um tamanho de bloco de 4KBMB.

A solução preferida é corrigir o ASMlib. O ID do bug Oracle é 13999609, e o patch está presente no oracleasm-support-2.1.8-1 e superior. Este patch permite que um usuário defina o parâmetro ORACLEASM_USE_LOGICAL_BLOCK_SIZE como true no /etc/sysconfig/oracleasm arquivo de configuração. Isso impede que o ASMlib use o parâmetro LBPPBE, o que significa que os LUNs na nova matriz agora são reconhecidos como dispositivos de bloco de 512 bytes.

A opção não altera o tamanho do bloco em LUNs que foram previamente carimbados pelo ASMlib. Por exemplo, se um grupo de discos ASM com blocos de 512 bytes precisar ser migrado para um novo sistema de armazenamento que relata um bloco 4KB, a opção ORACLEASM_USE_LOGICAL_BLOCK_SIZE deve ser definida antes que os novos LUNs sejam carimbados com ASMlib. Se os dispositivos já tiverem sido carimbados por oracleasm, eles devem ser reformatados antes de serem carimbados com um novo tamanho de bloco. Primeiro, desfigure o dispositivo com oracleasm deletedisk`e, em seguida, limpe os primeiros 1GB do dispositivo com `dd if=/dev/zero of=/dev/mapper/device bs=1048576 count=1024. Por fim, se o dispositivo tiver sido particionado anteriormente, use o kpartx comando para remover partições obsoletas ou simplesmente reiniciar o sistema operacional.

Se o ASMlib não puder ser corrigido, o ASMlib pode ser removido da configuração. Esta alteração é disruptiva e requer a remoção de carimbo de discos ASM e certificar-se de que o asm_diskstring parâmetro está definido corretamente. No entanto, essa alteração não requer a migração de dados.

Tamanhos de bloco do Acionamento do filtro ASM (AFD)

O AFD é uma biblioteca de gerenciamento ASM opcional que está se tornando a substituição do ASMlib. Do ponto de vista do armazenamento, ele é muito semelhante ao ASMlib, mas inclui recursos adicionais, como a capacidade de bloquear e/S não-Oracle para reduzir as chances de erros de usuário ou aplicativo que poderiam corromper dados.

Tamanhos de bloco de dispositivos

Como o ASMlib, o AFD também lê o parâmetro LUN blocos lógicos por expoente de bloco físico (LBPPBE) e, por padrão, usa o tamanho do bloco físico, não o tamanho do bloco lógico.

Isso pode criar um problema se AFD for adicionado a uma configuração existente onde os dispositivos ASM já estejam formatados como dispositivos de bloco de 512 bytes. O driver AFD reconheceria o LUN como um dispositivo 4K e a incompatibilidade entre o rótulo ASM e o dispositivo físico impediria o acesso. Da mesma forma, as migrações seriam afetadas porque não é possível misturar dispositivos de 512 bytes e 4KB no mesmo grupo de discos ASM. Isso bloquearia um usuário de aumentar o tamanho do grupo de discos ASM usando LUNs de dois arrays ou utilizando ASM como uma ferramenta de migração. Em outros casos, o RMAN pode não permitir a cópia de arquivos entre um grupo de discos ASM com um tamanho de bloco de 512 bytes e um grupo de discos ASM com um tamanho de bloco de 4KBMB.

A solução é simples - AFD inclui um parâmetro para controlar se usa os tamanhos de blocos lógicos ou físicos. Este é um parâmetro global que afeta todos os dispositivos no sistema. Para forçar o AFD a usar o

tamanho do bloco lógico, defina options oraclefd oraclefd_use_logical_block_size=1 no /etc/modprobe.d/oraclefd.conf arquivo.

Tamanhos de transferência multipath

As recentes alterações do kernel do linux impõem restrições de tamanho de e/S enviadas para dispositivos multipath, e o AFD não honra essas restrições. Os I/os são então rejeitados, o que faz com que o caminho LUN fique offline. O resultado é uma incapacidade de instalar o Oracle Grid, configurar ASM ou criar um banco de dados.

A solução é especificar manualmente o comprimento máximo de transferência no arquivo multipath.conf para LUNs ONTAP:

```
devices {  
    device {  
        vendor "NETAPP"  
        product "LUN.*"  
        max_sectors_kb 4096  
    }  
}
```

 Mesmo que não existam problemas atualmente, este parâmetro deve ser definido se AFD for usado para garantir que uma futura atualização do linux não cause problemas inesperadamente.

Microsoft Windows

Tópicos de configuração para banco de dados Oracle no Microsoft Windows com ONTAP.

NFS

A Oracle suporta o uso do Microsoft Windows com o cliente NFS direto. Esse recurso oferece um caminho para os benefícios de gerenciamento do NFS, incluindo a capacidade de exibir arquivos entre ambientes, redimensionar volumes dinamicamente e utilizar um protocolo IP menos caro. Consulte a documentação oficial da Oracle para obter informações sobre como instalar e configurar um banco de dados no Microsoft Windows usando DNFS. Não existem boas práticas especiais.

SAN

Para uma ótima eficiência de compressão, certifique-se de que o sistema de ficheiros NTFS utilize uma unidade de alocação de 8K GB ou maior. O uso de uma unidade de alocação 4K, que geralmente é o padrão, afeta negativamente a eficiência da compressão.

Solaris

Tópicos de configuração específicos do Solaris os.

Opções de montagem do Solaris NFS

A tabela a seguir lista as opções de montagem do Solaris NFS para uma única instância.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144</code>
Registros do Redo ControlFiles Datafiles	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr, llock, uid</code>
ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, uid</code>

Provou-se que o uso do `llock` melhora significativamente a performance nos ambientes dos clientes, eliminando a latência associada à aquisição e liberação de bloqueios no sistema de storage. Use essa opção com cuidado em ambientes nos quais vários servidores são configurados para montar os mesmos sistemas de arquivos e o Oracle está configurado para montar esses bancos de dados. Embora esta seja uma configuração altamente incomum, ela é usada por um pequeno número de clientes. Se uma instância for iniciada accidentalmente uma segunda vez, a corrupção de dados pode ocorrer porque a Oracle não consegue detetar os arquivos de bloqueio no servidor estrangeiro. Os bloqueios NFS não oferecem proteção de outra forma; como no NFS versão 3, eles são apenas consultivos.

Como `llock` os parâmetros `e forcedirectio` são mutuamente exclusivos, é importante que `filesystemio_options=setall` esteja presente no `init.ora` arquivo para que `directio` seja usado. Sem esse parâmetro, o cache do buffer do sistema operacional do host é usado e o desempenho pode ser afetado negativamente.

A tabela a seguir lista as opções de montagem do Solaris NFS RAC.

Tipo de ficheiro	Opções de montagem
ADR Home	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, noac</code>
Arquivos de controle arquivos de dados Redo logs	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr, noac, forcedirectio</code>
CRS/votação	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr, noac, forcedirectio</code>
Dedicado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, uid</code>
Compartilhado ORACLE_HOME	<code>rw,bg,hard,[vers=3,vers=4.1],proto=tcp, timeo=600, rsize=262144, wsize=262144, nointr, noac, uid</code>

A principal diferença entre as opções de montagem de uma única instância e RAC é a adição `noac` de e

`forcedirectio` às opções de montagem. Essa adição tem o efeito de desabilitar o cache do sistema operacional do host, o que permite que todas as instâncias do cluster RAC tenham uma visão consistente do estado dos dados. Embora o uso do `init.ora` parâmetro `filesystemio_options=setall` tenha o mesmo efeito de desabilitar o cache do host, ainda é necessário usar `noac` e `forcedirectio`.

O motivo `actimeo=0` necessário para implantações compartilhadas `ORACLE_HOME` é facilitar a consistência de arquivos, como arquivos de senha Oracle e spfiles. Se cada instância em um cluster RAC tiver um dedicado `ORACLE_HOME`, esse parâmetro não será necessário.

Opções de montagem do Solaris UFS

A NetApp recomenda fortemente o uso da opção de montagem de log para que a integridade dos dados seja preservada no caso de uma falha de host do Solaris ou a interrupção da conectividade FC. A opção de montagem de log também preserva a usabilidade dos backups Snapshot.

Solaris ZFS

O Solaris ZFS deve ser instalado e configurado cuidadosamente para oferecer o melhor desempenho.

mvector

O Solaris 11 incluiu uma mudança na forma como processa grandes operações de e/S, o que pode resultar em graves problemas de desempenho em matrizes de armazenamento SAN. O problema está documentado no relatório de bug de rastreamento do NetApp 630173, "regressão de desempenho do Solaris 11 ZFS".

Este não é um bug do ONTAP. É um defeito do Solaris que é rastreado sob os defeitos Solaris 7199305 e 7082975.

Você pode consultar o suporte Oracle para saber se sua versão do Solaris 11 é afetada ou testar a solução alternativa alterando `zfs_mvector_max_size` para um valor menor.

Você pode fazer isso executando o seguinte comando como root:

```
[root@host1 ~]# echo "zfs_mvector_max_size/W 0t131072" | mdb -kw
```

Se surgir algum problema inesperado dessa alteração, ela pode ser facilmente revertida executando o seguinte comando como root:

```
[root@host1 ~]# echo "zfs_mvector_max_size/W 0t1048576" | mdb -kw
```

Kernel

O desempenho confiável do ZFS requer um kernel Solaris corrigido contra problemas de alinhamento de LUN. A correção foi introduzida com o patch 147440-19 no Solaris 10 e com o SRU 10,5 para Solaris 11. Utilize apenas o Solaris 10 e posterior com o ZFS.

Configuração LUN

Para configurar um LUN, execute as seguintes etapas:

1. Crie um LUN do tipo solaris.
2. Instale o Kit de Utilitário do host (HUK) apropriado especificado pelo "[Ferramenta de Matriz de interoperabilidade NetApp \(IMT\)](#)".
3. Siga as instruções no HUK exatamente como descrito. Os passos básicos estão descritos abaixo, mas consulte o "[documentação mais recente](#)" para obter o procedimento adequado.
 - a. Execute o `host_config` utilitário para atualizar o `sd.conf`/`sdd.conf` arquivo. Isso permite que as unidades SCSI descubram corretamente LUNs ONTAP.
 - b. Siga as instruções dadas pelo `host_config` utilitário para ativar a entrada/saída multipath (MPIO).
 - c. Reinicie. Esta etapa é necessária para que quaisquer alterações sejam reconhecidas em todo o sistema.
4. particione os LUNs e verifique se eles estão alinhados corretamente. Consulte o "Apêndice B: Verificação do alinhamento do WAFL" para obter instruções sobre como testar e confirmar diretamente o alinhamento.

zpools

Um zpool só deve ser criado após as etapas no "[Configuração LUN](#)" serem executadas. Se o procedimento não for feito corretamente, pode resultar em degradação grave do desempenho devido ao alinhamento de e/S. O desempenho ideal no ONTAP requer que a e/S seja alinhada a um limite de 4K mm numa unidade. Os sistemas de arquivos criados em um zpool usam um tamanho de bloco efetivo que é controlado por meio de um parâmetro `ashift` chamado , que pode ser visualizado executando o comando `zdb -C`.

O valor `ashift` padrão é 9, o que significa 2⁹, ou 512 bytes. Para um desempenho ideal, o `ashift` valor deve ser 12 (2¹² 4K). Esse valor é definido no momento em que o zpool é criado e não pode ser alterado, o que significa que os dados em zpools com `ashift` outros que não 12 devem ser migrados copiando dados para um zpool recém-criado.

Depois de criar um zpool, verifique o valor de `ashift` antes de continuar. Se o valor não for 12, os LUNs não foram detetados corretamente. Destrua o zpool, verifique se todas as etapas mostradas na documentação relevante dos Utilitários do host foram executadas corretamente e recrie o zpool.

Zpools e Solaris LDOMs

Os Solaris LDOMs criam um requisito adicional para garantir que o alinhamento de e/S esteja correto. Embora um LUN possa ser encontrado corretamente como um dispositivo 4K, um dispositivo vdisk virtual em um LDOM não herda a configuração do domínio de e/S. O vdisk baseado nesse LUN retorna para um bloco de 512 bytes.

É necessário um ficheiro de configuração adicional. Primeiro, os LDOM individuais devem ser corrigidos para o bug Oracle 15824910 para habilitar as opções de configuração adicionais. Este patch foi portado para todas as versões usadas atualmente do Solaris. Uma vez que o LDOM é corrigido, ele está pronto para a configuração dos novos LUNs corretamente alinhados da seguinte forma:

1. Identifique o LUN ou LUNs a serem usados no novo zpool. Neste exemplo, é o dispositivo `c2d1`.

```
[root@LDM01 ~]# echo | format
Searching for disks...done
AVAILABLE DISK SELECTIONS:
0. c2d0 <Unknown-Unknown-0001-100.00GB>
   /virtual-devices@100/channel-devices@200/disk@0
1. c2d1 <SUN-ZFS Storage 7330-1.0 cyl 1623 alt 2 hd 254 sec 254>
   /virtual-devices@100/channel-devices@200/disk@1
```

2. Recupere a instância vdc dos dispositivos a serem usados para um pool ZFS:

```
[root@LDM01 ~]# cat /etc/path_to_inst
#
# Caution! This file contains critical kernel state
#
"/fcoe" 0 "fcoe"
"/iscsi" 0 "iscsi"
"/pseudo" 0 "pseudo"
"/scsi_vhci" 0 "scsi_vhci"
"/options" 0 "options"
"/virtual-devices@100" 0 "vnex"
"/virtual-devices@100/channel-devices@200" 0 "cnex"
"/virtual-devices@100/channel-devices@200/disk@0" 0 "vdc"
"/virtual-devices@100/channel-devices@200/pciv-communication@0" 0 "vpci"
"/virtual-devices@100/channel-devices@200/network@0" 0 "vnet"
"/virtual-devices@100/channel-devices@200/network@1" 1 "vnet"
"/virtual-devices@100/channel-devices@200/network@2" 2 "vnet"
"/virtual-devices@100/channel-devices@200/network@3" 3 "vnet"
"/virtual-devices@100/channel-devices@200/disk@1" 1 "vdc" << We want
this one
```

3. `/platform/sun4v/kernel/driv/vdc.conf` Editar :

```
block-size-list="1:4096";
```

Isso significa que a instância 1 do dispositivo recebe um tamanho de bloco de 4096MB.

Como exemplo adicional, suponha que as instâncias 1 a 6 do vdisk precisem ser configuradas para um tamanho de bloco 4K e `/etc/path_to_inst` lerem da seguinte forma:

```
"/virtual-devices@100/channel-devices@200/disk@1" 1 "vdc"
"/virtual-devices@100/channel-devices@200/disk@2" 2 "vdc"
"/virtual-devices@100/channel-devices@200/disk@3" 3 "vdc"
"/virtual-devices@100/channel-devices@200/disk@4" 4 "vdc"
"/virtual-devices@100/channel-devices@200/disk@5" 5 "vdc"
"/virtual-devices@100/channel-devices@200/disk@6" 6 "vdc"
```

4. O arquivo final vdc.conf deve conter o seguinte:

```
block-size-list="1:8192", "2:8192", "3:8192", "4:8192", "5:8192", "6:8192";
```

Cuidado

O LDOM deve ser reinicializado depois que o vdc.conf é configurado e o vdsk é criado. Este passo não pode ser evitado. A alteração do tamanho do bloco só entra em vigor após uma reinicialização. Prossiga com a configuração do zpool e certifique-se de que o ashift está corretamente configurado para 12, conforme descrito anteriormente.

Registo intenção ZFS (ZIL)

Geralmente, não há razão para localizar o ZFS Intent Log (ZIL) em um dispositivo diferente. O log pode compartilhar espaço com a piscina principal. O uso principal de um ZIL separado é quando se usa unidades físicas que não têm os recursos de armazenamento em cache de gravação em arrays de armazenamento modernos.

logbias

Defina logbias o parâmetro em sistemas de arquivos ZFS que hospedam dados Oracle.

```
zfs set logbias=throughput <filesystem>
```

O uso desse parâmetro reduz os níveis gerais de gravação. Sob os padrões, os dados escritos são comprometidos primeiro com o ZIL e depois para o pool de armazenamento principal. Essa abordagem é apropriada para uma configuração usando uma configuração de unidade simples, que inclui um dispositivo ZIL baseado em SSD e Mídia giratória para o pool de armazenamento principal. Isso ocorre porque permite que um commit ocorra em uma única transação de e/S na Mídia de menor latência disponível.

Ao usar um storage array moderno que inclua sua própria funcionalidade de armazenamento em cache, essa abordagem geralmente não é necessária. Em circunstâncias raras, pode ser desejável submeter uma gravação com uma única transação ao log, como uma carga de trabalho que consiste em gravações aleatórias altamente concentradas e sensíveis à latência. Há consequências na forma de amplificação de gravação porque os dados registrados são gravados no pool de armazenamento principal, resultando em uma duplicação da atividade de gravação.

E/S direta

Muitas aplicações, incluindo produtos Oracle, podem ignorar o cache de buffer do host habilitando a e/S direta

Esta estratégia não funciona como esperado com sistemas de arquivos ZFS. Embora o cache do buffer do host seja ignorado, o próprio ZFS continua a armazenar dados em cache. Essa ação pode resultar em resultados enganosos ao usar ferramentas como fio ou sio para realizar testes de desempenho, pois é difícil prever se a e/S está chegando ao sistema de armazenamento ou se está sendo armazenada em cache localmente no sistema operacional. Essa ação também torna muito difícil usar esses testes sintéticos para comparar o desempenho do ZFS com outros sistemas de arquivos. Na prática, há pouca ou nenhuma diferença no desempenho do sistema de arquivos em workloads reais do usuário.

Vários zpools

Backups, restaurações, clones e arquivamento baseados em snapshot de dados baseados em ZFS devem ser executados no nível do zpool e, geralmente, exigem vários zpools. Um zpool é análogo a um grupo de discos LVM e deve ser configurado usando as mesmas regras. Por exemplo, um banco de dados provavelmente é melhor definido com os datafiles que residem em `zpool1` e os logs de arquivo, arquivos de controle e logs de refazer que residem em `zpool2`. Essa abordagem permite um hot backup padrão no qual o banco de dados é colocado no modo hot backup, seguido por um snapshot de `zpool1`. O banco de dados é então removido do modo hot backup, o arquivo de log é forçado e um snapshot de `zpool2` é criado. Uma operação de restauração requer a desmontagem dos sistemas de arquivos zfs e a remoção do zpool em sua totalidade, seguindo-se uma operação de restauração do SnapRestore. O zpool pode então ser colocado online novamente e o banco de dados recuperado.

sistema de arquivos_options

O parâmetro Oracle `filesystemio_options` funciona de forma diferente com o ZFS. Se `setall` ou `directio` for usado, as operações de gravação são síncronas e ignoram o cache do buffer do sistema operacional, mas as leituras são armazenadas em buffer pelo ZFS. Essa ação causa dificuldades na análise de desempenho, pois às vezes, a e/S é intercetada e atendida pelo cache ZFS, tornando a latência do armazenamento e a e/S total menos do que parece ser.

Informações sobre direitos autorais

Copyright © 2026 NetApp, Inc. Todos os direitos reservados. Impresso nos EUA. Nenhuma parte deste documento protegida por direitos autorais pode ser reproduzida de qualquer forma ou por qualquer meio — gráfico, eletrônico ou mecânico, incluindo fotocópia, gravação, gravação em fita ou storage em um sistema de recuperação eletrônica — sem permissão prévia, por escrito, do proprietário dos direitos autorais.

O software derivado do material da NetApp protegido por direitos autorais está sujeito à seguinte licença e isenção de responsabilidade:

ESTE SOFTWARE É FORNECIDO PELA NETAPP "NO PRESENTE ESTADO" E SEM QUAISQUER GARANTIAS EXPRESSAS OU IMPLÍCITAS, INCLUINDO, SEM LIMITAÇÕES, GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO E ADEQUAÇÃO A UM DETERMINADO PROPÓSITO, CONFORME A ISENÇÃO DE RESPONSABILIDADE DESTE DOCUMENTO. EM HIPÓTESE ALGUMA A NETAPP SERÁ RESPONSÁVEL POR QUALQUER DANO DIRETO, INDIRETO, INCIDENTAL, ESPECIAL, EXEMPLAR OU CONSEQUENCIAL (INCLUINDO, SEM LIMITAÇÕES, AQUISIÇÃO DE PRODUTOS OU SERVIÇOS SOBRESSALENTES; PERDA DE USO, DADOS OU LUCROS; OU INTERRUPÇÃO DOS NEGÓCIOS), INDEPENDENTEMENTE DA CAUSA E DO PRINCÍPIO DE RESPONSABILIDADE, SEJA EM CONTRATO, POR RESPONSABILIDADE OBJETIVA OU PREJUÍZO (INCLUINDO NEGLIGÊNCIA OU DE OUTRO MODO), RESULTANTE DO USO DESTE SOFTWARE, MESMO SE ADVERTIDA DA RESPONSABILIDADE DE TAL DANO.

A NetApp reserva-se o direito de alterar quaisquer produtos descritos neste documento, a qualquer momento e sem aviso. A NetApp não assume nenhuma responsabilidade nem obrigação decorrentes do uso dos produtos descritos neste documento, exceto conforme expressamente acordado por escrito pela NetApp. O uso ou a compra deste produto não representam uma licença sob quaisquer direitos de patente, direitos de marca comercial ou quaisquer outros direitos de propriedade intelectual da NetApp.

O produto descrito neste manual pode estar protegido por uma ou mais patentes dos EUA, patentes estrangeiras ou pedidos pendentes.

LEGENDA DE DIREITOS LIMITADOS: o uso, a duplicação ou a divulgação pelo governo estão sujeitos a restrições conforme estabelecido no subparágrafo (b)(3) dos Direitos em Dados Técnicos - Itens Não Comerciais no DFARS 252.227-7013 (fevereiro de 2014) e no FAR 52.227- 19 (dezembro de 2007).

Os dados aqui contidos pertencem a um produto comercial e/ou serviço comercial (conforme definido no FAR 2.101) e são de propriedade da NetApp, Inc. Todos os dados técnicos e software de computador da NetApp fornecidos sob este Contrato são de natureza comercial e desenvolvidos exclusivamente com despesas privadas. O Governo dos EUA tem uma licença mundial limitada, irrevogável, não exclusiva, intransferível e não sublicenciável para usar os Dados que estão relacionados apenas com o suporte e para cumprir os contratos governamentais desse país que determinam o fornecimento de tais Dados. Salvo disposição em contrário no presente documento, não é permitido usar, divulgar, reproduzir, modificar, executar ou exibir os dados sem a aprovação prévia por escrito da NetApp, Inc. Os direitos de licença pertencentes ao governo dos Estados Unidos para o Departamento de Defesa estão limitados aos direitos identificados na cláusula 252.227-7015(b) (fevereiro de 2014) do DFARS.

Informações sobre marcas comerciais

NETAPP, o logotipo NETAPP e as marcas listadas em <http://www.netapp.com/TM> são marcas comerciais da NetApp, Inc. Outros nomes de produtos e empresas podem ser marcas comerciais de seus respectivos proprietários.