



Tiering-Strategien

Enterprise applications

NetApp
May 09, 2024

Inhalt

- Tiering-Strategien 1
 - FabricPool-Tiering für vollständige Dateien in der Oracle-Datenbank 1
 - Oracle FabricPool Tiering für partielle Dateien 1
 - Protokoll-Tiering für die Oracle-Datenbankarchivierung 2
 - Oracle mit FabricPool Snapshot Tiering 3
 - Backup-Tiering für Oracle Datenbank 4

Tiering-Strategien

FabricPool-Tiering für vollständige Dateien in der Oracle-Datenbank

FabricPool Tiering wird zwar auf Block-Ebene ausgeführt, kann jedoch in einigen Fällen für Tiering auf Dateiebene verwendet werden.

Viele Applikationsdatensätze sind nach Datum geordnet. Solche Daten sind im Allgemeinen immer seltener zugänglich, wenn sie älter werden. Beispielsweise verfügt eine Bank möglicherweise über ein Repository mit PDF-Dateien, die fünf Jahre Kundenabrechnungen enthalten, aber nur die letzten Monate sind aktiv. FabricPool kann verwendet werden, um ältere Datendateien in die Kapazitäts-Tier zu verschieben. Eine Abkühlzeit von 14 Tagen würde dafür sorgen, dass die letzten 14 Tage der PDF-Dateien auf der Performance-Ebene verbleiben. Darüber hinaus würden Dateien, die mindestens alle 14 Tage gelesen werden, „heiß“ bleiben und daher auf der Performance-Ebene verbleiben.

Richtlinien

Um einen dateibasierten Tiering-Ansatz zu implementieren, müssen Sie über Dateien verfügen, die geschrieben und nicht nachträglich geändert werden. Der `tiering-minimum-cooling-days` Richtlinien sollten so hoch eingestellt werden, dass Dateien, die Sie möglicherweise benötigen, auf der Performance-Tier verbleiben. Ein Datensatz, für den die letzten 60 Tage Daten mit optimaler Performance benötigt werden, erfordert beispielsweise die Einstellung `tiering-minimum-cooling-days` Bis 60. Ähnliche Ergebnisse lassen sich auch anhand der Dateizugriffsmuster erzielen. Wenn beispielsweise die Daten der letzten 90 Tage benötigt werden und die Applikation auf diese 90-Tage-Zeitspanne zugreift, verbleiben die Daten in der Performance-Tier. Durch Einstellen der `tiering-minimum-cooling-days` Zeitraum bis 2, erhalten Sie prompt Tiering, nachdem die Daten weniger aktiv werden.

Der `auto` Eine Richtlinie ist für das Tiering dieser Blöcke erforderlich, da nur die `auto` Die Richtlinie wirkt sich auf Blöcke aus, die sich im aktiven Filesystem befinden.



Jeder Zugriff auf Daten setzt die Heatmap-Daten zurück. Virus-Scan, Indizierung und sogar Backup-Aktivitäten, die die Quelldateien lesen, verhindern Tiering, da dies erforderlich ist `tiering-minimum-cooling-days` Schwellenwert wird nie erreicht.

Oracle FabricPool Tiering für partielle Dateien

Da FabricPool auf Block-Ebene arbeitet, können geänderte Dateien teilweise auf Objekt-Storage verschoben werden und dennoch nur teilweise auf Performance-Tier verbleiben.

Dies ist bei Datenbanken üblich. Datenbanken, für die bekanntermaßen inaktive Blöcke enthalten sind, eignen sich auch für das FabricPool Tiering. Beispielsweise kann eine Supply-Chain-Management-Datenbank historische Informationen enthalten, die bei Bedarf verfügbar sein müssen, aber während des normalen Betriebs nicht aufgerufen werden. Mit FabricPool können die inaktiven Blöcke selektiv verschoben werden.

Beispielsweise Datendateien, die auf einem FabricPool Volume mit einem ausgeführt werden `tiering-minimum-cooling-days` Im Zeitraum von 90 Tagen werden sämtliche Blöcke aufbewahrt, auf die in den vorangegangenen 90 Tagen auf der Performance Tier zugegriffen wurde. Alle Daten, auf die 90 Tage lang nicht zugegriffen wird, werden jedoch auf die Kapazitäts-Tier verlagert. In anderen Fällen bleiben bei normalen Applikationsaktivitäten die richtigen Blöcke auf der richtigen Tier erhalten. Wenn beispielsweise eine

Datenbank normalerweise dazu verwendet wird, die Daten der letzten 60 Tage regelmäßig zu verarbeiten, ist dies wesentlich geringer `tiering-minimum-cooling-days` Zeitraum kann festgelegt werden, da die natürliche Aktivität der Anwendung dafür sorgt, dass Blöcke nicht vorzeitig verschoben werden.

Der `auto` Richtlinien sollten mit Vorsicht bei Datenbanken verwendet werden. Viele Datenbanken verfügen über periodische Aktivitäten wie etwa Vorgänge zum Quartalsende oder die Neuindizierung. Wenn der Zeitraum dieser Vorgänge größer ist als der `tiering-minimum-cooling-days` Es können Performance-Probleme auftreten. Wenn zum Quartalsende beispielsweise 1 TB an Daten verarbeitet werden müssen, die ansonsten nicht verarbeitet wurden, befinden sich diese Daten möglicherweise nun auf der Kapazitäts-Tier. Lesezugriffe von der Kapazitäts-Tier sind oft extrem schnell und verursachen möglicherweise keine Performance-Probleme. Die genauen Ergebnisse hängen jedoch von der Objektspeicher-Konfiguration ab.

Richtlinien

Der `tiering-minimum-cooling-days` Die Richtlinie sollte so hoch eingestellt werden, dass Dateien, die auf der Performance-Tier erforderlich sind, aufbewahrt werden. Beispielsweise müsste eine Datenbank, in der die letzten 60 Tage Daten bei einer optimalen Performance benötigt werden, die festlegen `tiering-minimum-cooling-days` Zeitraum bis 60 Tage. Ähnliche Ergebnisse lassen sich auch anhand der Zugriffsmuster von Dateien erzielen. Wenn beispielsweise die Daten der letzten 90 Tage benötigt werden und die Applikation auf diese 90-Tage-Datenspanne zugreift, verbleiben die Daten in der Performance-Tier. Einstellen des `tiering-minimum-cooling-days` Zeitraum bis 2 Tage würde die Daten sofort nach dem Zeitpunkt verschieben, an dem die Daten weniger aktiv sind.

Der `auto` Eine Richtlinie ist für das Tiering dieser Blöcke erforderlich, da nur die `auto` Die Richtlinie wirkt sich auf Blöcke aus, die sich im aktiven Filesystem befinden.



Jeder Zugriff auf Daten setzt die Heatmap-Daten zurück. Daher verhindert die Überprüfung der vollständigen Tabelle der Datenbank und sogar die Backup-Aktivitäten, die die Quelldateien lesen, Tiering, da die erforderlichen `tiering-minimum-cooling-days` Schwellenwert wird nie erreicht.

Protokoll-Tiering für die Oracle-Datenbankarchivierung

Die wahrscheinlich wichtigste Verwendung für FabricPool ist die Verbesserung der Effizienz bekannter, kalter Daten, wie z. B. Transaktions-Logs der Datenbank.

Die meisten relationalen Datenbanken arbeiten im Transaktionsprotokoll-Archivierungsmodus, um Point-in-Time Recovery bereitzustellen. Änderungen an den Datenbanken werden durch die Aufzeichnung der Änderungen in den Transaktionsprotokollen vorgenommen und das Transaktionsprotokoll wird ohne Überschreibung beibehalten. Dies kann zur Anforderung führen, eine enorme Menge an archivierten Transaktions-Logs aufzubewahren. Ähnliche Beispiele gibt es bei vielen anderen Applikations-Workflows, die Daten generieren, die aufbewahrt werden müssen, auf die jedoch mit hoher Wahrscheinlichkeit niemals zugegriffen werden wird.

FabricPool löst diese Probleme mit einer einzigen Lösung mit integriertem Tiering. Dateien werden gespeichert und bleiben an ihrem üblichen Speicherort zugänglich, belegen jedoch praktisch keinen Speicherplatz auf dem primären Array.

Richtlinien

Verwenden Sie A `tiering-minimum-cooling-days` Eine Richtlinie von wenigen Tagen führt zur Aufbewahrung von Blöcken in den kürzlich erstellten Dateien (die Dateien sind, die in naher Zukunft am

wahrscheinlichsten erforderlich sind) auf der Performance-Tier. Die Datenblöcke aus älteren Dateien werden dann auf die Kapazitäts-Tier verschoben.

Der `auto` Erzwingt sofortiges Tiering, wenn der Kühlwellenwert erreicht wurde, unabhängig davon, ob die Protokolle gelöscht wurden oder weiterhin im primären Dateisystem vorhanden sind. Auch das Speichern aller potenziell erforderlichen Protokolle an einer zentralen Stelle im aktiven Filesystem vereinfacht das Management. Es gibt keinen Grund, Snapshots zu durchsuchen, um eine Datei zu finden, die wiederhergestellt werden muss.

Einige Applikationen, wie z. B. Microsoft SQL Server, schneiden Transaktions-Log-Dateien während von Backup-Vorgängen ab, sodass sich die Protokolle nicht mehr im aktiven File-System befinden. Die Kapazität kann mithilfe des gespeichert werden `snapshot-only` tiering-Richtlinie, aber die `auto` Die Richtlinie ist für Protokoll Daten nicht nützlich, da Protokoll Daten im aktiven Dateisystem selten abgekühlt werden sollten.

Oracle mit FabricPool Snapshot Tiering

Die erste Version von FabricPool war auf den Backup-Anwendungsfall ausgerichtet. Als einzige Art von Blöcken, die Tiering ermöglichen konnten, handelte es sich um Blöcke, die nicht mehr mit Daten im aktiven File-System verknüpft waren. Daher können nur die Snapshot Datenblöcke auf diese Kapazitäts-Tier verschoben werden. Dies bleibt eine der sichersten Tiering-Optionen, wenn Sie sicherstellen müssen, dass die Performance nie beeinträchtigt wird.

Richtlinien: Lokale Snapshots

Es gibt zwei Optionen für das Tiering inaktiver Snapshot-Blöcke auf die Kapazitäts-Tier. Zunächst einmal die `snapshot-only` Die Richtlinie zielt nur auf die Snapshot-Blöcke ab. Obwohl der `auto` Die Richtlinie umfasst die `snapshot-only` Blöcke, sondern auch Tiering Blöcke aus dem aktiven File-System. Dies ist möglicherweise nicht wünschenswert.

Der `tiering-minimum-cooling-days` Der Wert sollte auf einen Zeitraum festgelegt werden, in dem Daten, die während einer Wiederherstellung erforderlich sein könnten, auf der Performance-Tier zur Verfügung stehen. So enthalten die meisten Wiederherstellungsszenarien einer kritischen Produktionsdatenbank zu einem bestimmten Zeitpunkt in den letzten Tagen einen Wiederherstellungspunkt. Einstellung `A tiering-minimum-cooling-days` Mit dem Wert 3 würde sichergestellt, dass bei einer Wiederherstellung der Datei eine Datei entsteht, die sofort die maximale Performance liefert. Alle Blöcke in den aktiven Dateien befinden sich immer noch auf schnellem Storage, ohne dass eine Wiederherstellung aus dem Kapazitäts-Tier erforderlich ist.

Richtlinien – replizierte Snapshots

Ein Snapshot, der mit SnapMirror oder SnapVault repliziert wird, der nur für die Wiederherstellung verwendet wird, sollte im Allgemeinen die FabricPool verwenden `all` Richtlinie: Bei dieser Richtlinie werden Metadaten repliziert. Alle Datenblöcke werden jedoch sofort an die Kapazitäts-Tier gesendet, was maximale Performance liefert. Die meisten Recovery-Prozesse arbeiten mit sequenziellem I/O, was von vornherein effizient ist. Die Recovery-Zeit vom Zielort des Objektspeichers ist zu bewerten, in einer gut durchdachten Architektur muss dieser Recovery-Prozess jedoch nicht wesentlich langsamer sein als die Wiederherstellung von lokalen Daten.

Wenn die replizierten Daten auch für das Klonen verwendet werden sollen, wird der verwendet `auto` Die Politik ist angemessener, mit einem `tiering-minimum-cooling-days` Wert, der Daten umfasst, von denen erwartet wird, dass sie regelmäßig in einer Klonumgebung verwendet werden. Der aktive Arbeitsdatensatz einer Datenbank kann beispielsweise Daten enthalten, die in den letzten drei Tagen gelesen oder geschrieben

wurden, aber es können auch weitere Verlaufsdaten von 6 Monaten enthalten sein. Wenn ja, dann die `auto` Durch eine Richtlinie am Ziel von SnapMirror wird das Arbeitsdatensatz auf der Performance-Ebene verfügbar.

Backup-Tiering für Oracle Datenbank

Zu den herkömmlichen Applikations-Backups gehören Produkte wie der Oracle Recovery Manager, die dateibasierte Backups außerhalb des Standorts der Originaldatenbank erstellen.

```
`tiering-minimum-cooling-days` policy of a few days preserves the most recent backups, and therefore the backups most likely to be required for an urgent recovery situation, on the performance tier. The data blocks of the older files are then moved to the capacity tier.
```

Der `auto` Die Richtlinie ist die am besten geeignete Richtlinie für Backup-Daten. Dadurch wird ein sofortiges Tiering sichergestellt, wenn der Kühlenschwellenwert erreicht wurde, unabhängig davon, ob die Dateien gelöscht wurden oder weiterhin im primären Dateisystem vorhanden sind. Das Speichern aller potenziell erforderlichen Dateien an einem zentralen Speicherort im aktiven Dateisystem vereinfacht ebenfalls das Management. Es gibt keinen Grund, Snapshots zu durchsuchen, um eine Datei zu finden, die wiederhergestellt werden muss.

Der `snapshot-only` Richtlinien können zwar funktionieren, sie gelten jedoch nur für Blöcke, die sich nicht mehr im aktiven File-System befinden. Daher müssen Dateien auf einer NFS- oder SMB-Freigabe vor dem Daten-Tiering zuerst gelöscht werden.

Diese Richtlinie wäre bei einer LUN-Konfiguration sogar noch weniger effizient, da beim Löschen einer Datei aus einer LUN nur Dateiverweise aus den Metadaten des Filesystems entfernt werden. Die tatsächlichen Blöcke auf den LUNs bleiben vorhanden, bis sie überschrieben werden. Dies kann zu einer sehr langen Verzögerung zwischen dem Löschen einer Datei und dem Überschreiben der Blöcke führen und zu Tiering-Kandidaten werden. Der Wechsel des bietet einige Vorteile `snapshot-only` Blöcke auf die Kapazitäts-Tier, aber insgesamt funktioniert das FabricPool Management von Backup-Daten am besten mit der `auto` Richtlinie:



Mit diesem Ansatz können Benutzer den für Backups erforderlichen Speicherplatz effizienter managen. FabricPool selbst ist jedoch keine Backup-Technologie. Das Tiering von Backup-Dateien in Objektspeicher vereinfacht das Management, da die Dateien noch auf dem ursprünglichen Storage-System sichtbar sind, die Datenblöcke im Zielspeicherort jedoch vom ursprünglichen Storage-System abhängig sind. Wenn das Quell-Volumen verloren geht, sind die Objektspeicher-Daten nicht mehr nutzbar.

Copyright-Informationen

Copyright © 2024 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFT SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGEND EINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

Markeninformationen

NETAPP, das NETAPP Logo und die unter <http://www.netapp.com/TM> aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.