



Configurazione di ONTAP

Enterprise applications

NetApp
May 09, 2024

Sommario

- Configurazione di ONTAP 1
 - RAID e database Oracle 1
 - Gestione della capacità dello storage e dei database Oracle 1
 - Database Oracle e Storage Virtual Machine 2
 - Gestione delle performance dei database Oracle con QoS ONTAP 3
 - Database Oracle e funzionalità di efficienza ONTAP 4
 - Thin provisioning con database Oracle 8
 - Failover/switchover dei database Oracle e del controller ONTAP 11

Configurazione di ONTAP

RAID e database Oracle

RAID si riferisce all'utilizzo della ridondanza per proteggere i dati dalla perdita di un'unità.

Occasionalmente sorgono domande riguardanti i livelli RAID nella configurazione dello storage NetApp utilizzato per i database Oracle e altre applicazioni aziendali. Molte Best practice Oracle precedenti relative alla configurazione degli array di storage contengono avvisi sull'utilizzo del mirroring RAID e/o sull'eliminazione di determinati tipi di RAID. Sebbene sollevino punti validi, questi sorgenti non si applicano a RAID 4 e alle tecnologie NetApp RAID DP e RAID-TEC utilizzate in ONTAP.

RAID 4, RAID 5, RAID 6, RAID DP e RAID-TEC utilizzano tutti la parità per garantire che il guasto al disco non determini una perdita di dati. Queste opzioni RAID offrono un utilizzo dello storage migliore rispetto al mirroring, ma la maggior parte delle implementazioni RAID presenta uno svantaggio che influisce sulle operazioni di scrittura. Il completamento di un'operazione di scrittura su altre implementazioni RAID potrebbe richiedere letture di più unità per rigenerare i dati di parità, un processo comunemente chiamato penalizzazione RAID.

ONTAP, tuttavia, non subisce questa penalizzazione del RAID. Ciò è dovuto all'integrazione di NetApp WAFL (Write Anywhere file Layout) con il livello RAID. Le operazioni di scrittura vengono unite nella RAM e preparate come uno stripe RAID completo, inclusa la generazione della parità. ONTAP non ha bisogno di eseguire una lettura per completare una scrittura, il che significa che ONTAP e WAFL evitano la penalizzazione RAID. Le performance per le operazioni critiche in termini di latenza, come il logging di redo, vengono mantenute e le scritture random dei file di dati non comportano penalizzazioni RAID dovute alla necessità di rigenerare la parità.

Per quanto riguarda l'affidabilità statistica, anche RAID DP offre una protezione migliore rispetto al mirroring RAID. Il problema principale è la richiesta fatta sui dischi durante una ricostruzione del RAID. Con un set RAID con mirroring, il rischio di perdita di dati causata da un guasto al disco e durante la ricostruzione nel partner nel set RAID è molto maggiore del rischio di un guasto a tre dischi in un set RAID DP.

Gestione della capacità dello storage e dei database Oracle

La gestione di un database o di un'altra applicazione aziendale con storage aziendale prevedibile, gestibile e ad alte prestazioni richiede spazio libero sulle unità per la gestione di dati e metadati. La quantità di spazio libero richiesta dipende dal tipo di unità utilizzata e dai processi aziendali.

Lo spazio libero viene definito come lo spazio non utilizzato per i dati effettivi e include lo spazio non allocato dell'aggregato e lo spazio non utilizzato all'interno dei volumi costituenti. È importante prendere in considerazione anche il thin provisioning. Ad esempio, un volume potrebbe contenere un LUN da 1TB GB, di cui solo il 50% viene utilizzato dai dati reali. In un ambiente con thin provisioning, questo sembra consumare correttamente 500GB GB di spazio. Tuttavia, in un ambiente con provisioning completo, la capacità completa di 1TB TB sembra essere in uso. I 500GB GB di spazio non allocato sono nascosti. Questo spazio non è utilizzato dai dati effettivi e deve quindi essere incluso nel calcolo dello spazio libero totale.

Di seguito sono riportate le raccomandazioni NetApp per i sistemi storage utilizzati per le applicazioni aziendali:

Aggregati SSD, inclusi i sistemi AFF



NetApp consiglia almeno il 10% di spazio libero. Ciò comprende tutto lo spazio inutilizzato, compreso lo spazio libero all'interno dell'aggregato o di un volume ed eventuale spazio libero allocato a causa dell'utilizzo del provisioning completo, ma non utilizzato dai dati effettivi. Lo spazio logico non è importante, la domanda è quanto spazio fisico libero effettivo è disponibile per lo storage dei dati.

Il consiglio di liberare il 10% dello spazio è molto conservativo. Gli aggregati SSD possono supportare i carichi di lavoro a livelli di utilizzo ancora più elevati senza influire sulle performance. Tuttavia, con l'aumento dell'utilizzo dell'aggregato, aumenta anche il rischio di esaurimento dello spazio se l'utilizzo non viene monitorato con attenzione. Inoltre, mentre si utilizza un sistema al 99% della capacità potrebbe non verificarsi un peggioramento delle performance, tuttavia si verificherebbe un sforzo di gestione che impedirebbe il riempimento completo del sistema mentre si ordina hardware aggiuntivo e potrebbe essere necessario del tempo per l'acquisto e l'installazione di dischi aggiuntivi.

Aggregati HDD, compresi gli aggregati Flash Pool



NetApp consiglia almeno il 15% di spazio libero quando si utilizzano unità rotanti. Ciò comprende tutto lo spazio inutilizzato, compreso lo spazio libero all'interno dell'aggregato o di un volume ed eventuale spazio libero allocato a causa dell'utilizzo del provisioning completo, ma non utilizzato dai dati effettivi. Le prestazioni saranno influenzate dagli approcci di conversazione libera al 10%.

Database Oracle e Storage Virtual Machine

La gestione dello storage del database Oracle è centralizzata su una Storage Virtual Machine (SVM)

Una SVM, nota come vserver nell'interfaccia a riga di comando di ONTAP, è un'unità funzionale di base dello storage ed è utile confrontare una SVM con un guest su un server VMware ESX.

Quando viene installato per la prima volta, ESX non dispone di funzionalità preconfigurate, come l'hosting di un sistema operativo guest o il supporto di un'applicazione per l'utente finale. Si tratta di un container vuoto fino a quando non viene definita una macchina virtuale (VM). ONTAP è simile. Quando viene installata per la prima volta, ONTAP non dispone di funzionalità di servizio dati fino a quando non viene creata una SVM. È il linguaggio della SVM che definisce i servizi dati.

Come per altri aspetti dell'architettura dello storage, le migliori opzioni per il design di SVM e interfaccia logica (LIF) dipendono in gran parte dai requisiti di scalabilità e dalle esigenze di business.

SVM

Non esistono Best practice ufficiali per il provisioning di SVM per ONTAP. Il giusto approccio dipende dai requisiti di gestione e sicurezza.

La maggior parte dei clienti utilizza una SVM primaria per la maggior parte delle loro esigenze quotidiane, quindi crea un piccolo numero di SVM per esigenze speciali. Ad esempio, è possibile creare:

- Una SVM per un database aziendale critico gestita da un team di specialisti
- Una SVM per un gruppo di sviluppo al quale è stato assegnato un controllo amministrativo completo in modo da poter gestire il proprio storage in maniera indipendente

- Una SVM per i dati di business sensibili, come le risorse umane o i dati di reporting finanziario, per cui il team di amministrazione deve essere limitato

In un ambiente multi-tenant, è possibile assegnare a ciascun tenant una SVM dedicata. Il limite per il numero di SVM e LIF per cluster, coppia ha e nodo dipende dal protocollo in uso, dal modello di nodo e dalla versione di ONTAP. Consultare "[NetApp Hardware Universe](#)" per questi limiti.

Gestione delle performance dei database Oracle con QoS ONTAP

La gestione sicura ed efficiente di più database Oracle richiede un'efficace strategia di QoS. Il motivo è rappresentato dalle funzionalità di performance in costante aumento offerte da un sistema storage moderno.

Nello specifico, la maggiore adozione dello storage all-flash ha permesso il consolidamento dei carichi di lavoro. Gli storage array che si affidano a supporti rotanti tendevano a supportare solo un numero limitato di workload i/o-intensive a causa delle limitate funzionalità IOPS della tecnologia delle unità rotazionali meno recente. Uno o due database altamente attivi saturerebbero i dischi sottostanti molto prima che gli storage controller raggiungano i loro limiti. Questo è cambiato. La capacità di performance di un numero relativamente contenuto di dischi SSD è in grado di saturare anche gli storage controller più potenti. Ciò significa che è possibile sfruttare tutte le funzionalità dei controller senza la paura di un improvviso crollo delle performance con picchi di latenza dei supporti rotanti.

Come esempio di riferimento, un semplice sistema ha AFF A800 a due nodi è in grado di fornire fino a un milione di IOPS casuali prima che la latenza superi un millisecondo. Ci si aspetta che pochissimi carichi di lavoro singoli raggiungano tali livelli. L'utilizzo completo di questo array di sistema AFF A800 implicherà l'hosting di più carichi di lavoro, per questo motivo in modo sicuro, garantendo al contempo la prevedibilità dei requisiti di qualità del servizio.

Esistono due tipi di qualità del servizio (QoS) in ONTAP: IOPS e larghezza di banda. È possibile applicare controlli di qualità del servizio a SVM, volumi, LUN e file.

QoS (IOPS)

Un controllo della qualità del servizio IOPS si basa ovviamente sugli IOPS totali di una data risorsa, ma esistono alcuni aspetti della qualità del servizio IOPS che potrebbero non essere intuitivi. Alcuni clienti sono rimasti colpiti dall'apparente aumento della latenza al raggiungimento di una soglia IOPS. L'aumento della latenza è il risultato naturale della limitazione degli IOPS. Logicamente, funziona in modo simile a un sistema token. Ad esempio, se un dato volume contenente file di dati ha un limite di 10K IOPS, ogni i/o che arriva deve prima ricevere un token per continuare l'elaborazione. Fino a quando non sono stati consumati più di 10K gettoni in un dato secondo, non sono presenti ritardi. Se le operazioni i/o devono attendere per ricevere il token, questa attesa viene visualizzata come latenza aggiuntiva. Più un carico di lavoro supera il limite di qualità del servizio, più a lungo ogni i/o deve attendere in coda per l'elaborazione del proprio turno, che appare all'utente come una latenza più elevata.



Prestare attenzione nell'applicazione dei controlli QoS ai dati dei log di transazione/ripristino del database. Mentre le richieste di performance del logging di redo sono in genere molto, molto più basse dei data afiles, l'attività del log di redo è molto bursty. L'io avviene in brevi impulsi e un limite QoS che appare appropriato per i livelli di io di redo medi potrebbe essere troppo basso per i requisiti effettivi. Il risultato può essere una serie di limitazioni delle performance, mentre la qualità del servizio viene associata a ogni burst dei log di ripristino. In generale, il redo e la registrazione dell'archivio non devono essere limitati dalla QoS.

QoS della larghezza di banda

Non tutte le dimensioni i/o sono uguali. Ad esempio, un database potrebbe eseguire un elevato numero di piccoli blocchi di lettura con il raggiungimento della soglia IOPS, tuttavia, è possibile che i database eseguano anche un'operazione di scansione completa della tabella, che consisterebbe in un numero molto ridotto di letture di blocchi di grandi dimensioni, consumando una grande quantità di larghezza di banda ma con un numero relativamente basso di IOPS.

Allo stesso modo, un ambiente VMware potrebbe gestire un numero molto elevato di IOPS casuali durante l'avvio, ma eseguirebbe un numero minore di io, ma più grande, durante un backup esterno.

Una gestione efficace delle performance a volte richiede limiti di qualità del servizio (QoS) IOPS o larghezza di banda, o anche entrambi.

Qualità del servizio minima/garantita

Molti clienti cercano una soluzione che includa QoS garantita, che sia più difficile da raggiungere di quanto possa sembrare e che sia potenzialmente abbastanza dispendiosa. Ad esempio, collocare 10 database con una garanzia di 10K IOPS richiede il dimensionamento di un sistema per uno scenario in cui tutti i 10 database vengono eseguiti contemporaneamente a 10K IOPS, per un totale di 100K.

L'utilizzo ottimale per i controlli minimi della qualità del servizio è la protezione dei carichi di lavoro critici. Ad esempio, prendi in considerazione un controller ONTAP con un numero massimo di IOPS possibile di 500K e un mix di workload di produzione e sviluppo. È consigliabile applicare policy QoS massime ai carichi di lavoro di sviluppo per impedire a qualsiasi database di monopolizzare il controller. Quindi, ai carichi di lavoro di produzione si applicano policy minime di qualità del servizio per assicurarsi che dispongano sempre degli IOPS richiesti, quando necessario.

QoS adattiva

La qualità del servizio adattiva fa riferimento alla funzionalità ONTAP, in cui il limite della qualità del servizio si basa sulla capacità dell'oggetto storage. Viene utilizzata raramente con i database perché di solito non esiste alcun collegamento tra le dimensioni di un database e i relativi requisiti prestazionali. I database di grandi dimensioni possono essere quasi inerti, mentre quelli di dimensioni inferiori possono utilizzare un numero elevato di IOPS.

La qualità del servizio adattiva può rivelarsi molto utile con i datastore di virtualizzazione, perché i requisiti di IOPS di tali set di dati tendono a correlare le dimensioni totali del database. Un datastore più recente, che contiene 1TB TB di file VMDK, avrà probabilmente bisogno di circa la metà delle performance rispetto a un datastore da 2TB TB. La qualità del servizio adattiva ti consente di aumentare automaticamente i limiti della qualità del servizio, man mano che il datastore viene popolato con i dati.

Database Oracle e funzionalità di efficienza ONTAP

Le funzionalità di efficienza dello spazio di ONTAP sono ottimizzate per i database Oracle. In quasi tutti i casi, l'approccio migliore è quello di lasciare le impostazioni predefinite con tutte le funzioni di efficienza attivate.

Le funzionalità di efficienza in termini di spazio, come compressione, compaction e deduplica, sono progettate per aumentare la quantità di dati logici applicabili a una determinata quantità di storage fisico. Il risultato è una riduzione dei costi e dell'overhead di gestione.

Ad un livello elevato, la compressione è un processo matematico in cui gli schemi nei dati vengono rilevati e

codificati in modo da ridurre i requisiti di spazio. La deduplica, invece, rileva i blocchi di dati effettivi e ripetuti e rimuove le copie estranee. La tecnologia di compaction consente a più blocchi logici di dati di condividere lo stesso blocco fisico sui supporti.



Per una spiegazione dell'interazione tra efficienza dello storage e prenotazione frazionata, vedere le sezioni seguenti sul thin provisioning.

Compressione

Prima della disponibilità dei sistemi storage all-flash, la compressione basata su array aveva un valore limitato, perché la maggior parte dei carichi di lavoro con i/o-intensive richiedeva un numero molto elevato di spindle per fornire performance accettabili. I sistemi storage contenevano invariabilmente una capacità superiore rispetto a quella richiesta come effetto collaterale dell'elevato numero di dischi. La situazione è cambiata con l'ascesa dello storage a stato solido. Non è più necessario effettuare un provisioning in eccesso significativo dei dischi solo per ottenere buone prestazioni. Lo spazio su disco di un sistema di storage può essere adattato alle effettive esigenze di capacità.

L'aumento della capacità degli IOPS dei dischi a stato solido (SSD) offre quasi sempre risparmi sui costi rispetto ai dischi rotanti, ma la compressione può ottenere ulteriori risparmi aumentando la capacità effettiva dei supporti a stato solido.

Esistono diversi modi per comprimere i dati. Molti database includono proprie funzionalità di compressione, sebbene raramente queste vengano osservate negli ambienti dei clienti. Il motivo è solitamente la penalizzazione delle prestazioni per una **modifica** dei dati compressi, mentre con alcune applicazioni vi sono elevati costi di licenza per la compressione a livello di database. Infine, ci sono le conseguenze globali delle performance sulle operazioni di database. Ha poco senso pagare un costo elevato di licenza per CPU per una CPU che esegue la compressione e la decompressione dei dati piuttosto che un vero lavoro di database. Un'opzione migliore è trasferire il lavoro di compressione sul sistema storage.

Compressione adattiva

La compressione adattiva è stata testata accuratamente con carichi di lavoro Enterprise senza effetti osservati sulle performance, anche in un ambiente all-flash in cui la latenza viene misurata in microsecondi. Alcuni clienti hanno anche segnalato un aumento delle performance con l'utilizzo della compressione, perché i dati rimangono compressi nella cache, aumentando di fatto la quantità di cache disponibile in un controller.

ONTAP gestisce i blocchi fisici in 4KB unità. La compressione adattiva utilizza dimensioni predefinite dei blocchi di compressione di 8KB KB, il che significa che i dati sono compressi in unità da 8KB KB. Corrisponde alle dimensioni dei blocchi di 8KB KB utilizzate più spesso dai database relazionali. Gli algoritmi di compressione diventano più efficienti con la compressione di un numero maggiore di dati come una singola unità. Una dimensione dei blocchi di compressione da 32KB KB sarebbe più efficiente in termini di spazio rispetto a un'unità dei blocchi di compressione da 8KB KB. Ciò significa che la compressione adattiva che utilizza le dimensioni predefinite dei blocchi di 8KB KB produce tassi di efficienza leggermente inferiori, ma esiste anche un vantaggio significativo nell'utilizzo di dimensioni inferiori dei blocchi di compressione. I carichi di lavoro dei database includono un'elevata attività di sovrascrittura. La sovrascrittura di un 8KB di un blocco di dati 32KB compresso richiede la lettura dell'intero 32KB di dati logici, la decompressione, l'aggiornamento della regione 8KB richiesta, la ricompressione e quindi la riscrittura dell'intero 32KB sui dischi. Si tratta di un'operazione molto costosa per un sistema storage ed è il motivo per cui alcuni storage array concorrenti basati su dimensioni dei blocchi di compressione più grandi implicano anche una significativa penalizzazione delle performance con i carichi di lavoro dei database.



Le dimensioni dei blocchi utilizzate dalla compressione adattiva possono essere aumentate fino a 32KB KB. Questo può migliorare l'efficienza di archiviazione e dovrebbe essere considerato per i file inattivi come i log delle transazioni e i file di backup quando una quantità sostanziale di tali dati è memorizzata nell'array. In alcune situazioni, i database attivi che utilizzano dimensioni blocco 16KB KB o 32KB KB possono anche trarre vantaggio dall'aumento delle dimensioni blocco della compressione adattiva per adeguarsi. Consulta un NetApp o un rappresentante del partner per ottenere indicazioni relative all'adeguatezza del tuo carico di lavoro.



Le dimensioni dei blocchi di compressione superiori a 8KB KB non devono essere utilizzate insieme alla deduplica nelle destinazioni di backup in streaming. Il motivo è che piccole modifiche ai dati di backup influiscono sulla finestra di compressione 32KB. Se la finestra si sposta, i dati compressi risultanti differiscono per l'intero file. La deduplica si verifica dopo la compressione, il che significa che il motore di deduplica vede ogni backup compresso in modo diverso. Se è richiesta la deduplica dei backup in streaming, è consigliabile utilizzare solo la compressione adattiva per blocchi da 8KB KB. La compressione adattiva è preferibile, perché funziona a blocchi di dimensioni inferiori e non interrompe l'efficienza di deduplica. Per motivi simili, la compressione lato host interferisce anche con l'efficienza della deduplica.

Allineamento delle compressioni

La compressione adattiva in un ambiente di database richiede alcune considerazioni sull'allineamento dei blocchi di compressione. Ciò rappresenta solo una preoccupazione per i dati che sono soggetti a sovrascritture casuali di blocchi molto specifici. Questo approccio è simile in teoria all'allineamento complessivo del file system, dove l'inizio di un file system deve essere allineato al limite di un dispositivo 4K e la dimensione di blocco di un file system deve essere un multiplo di 4K.

Ad esempio, una scrittura 8KB in un file viene compressa solo se si allinea con un limite 8KB all'interno del file system stesso. Questo punto significa che deve rientrare nel primo 8KB del file, nel secondo 8KB del file e così via. Il modo più semplice per garantire un corretto allineamento è utilizzare il tipo di LUN corretto, ogni partizione creata dovrebbe avere un offset dall'inizio del dispositivo che è un multiplo di 8K, e utilizzare una dimensione del blocco del file system che è un multiplo della dimensione del blocco del database.

Dati come backup o log delle transazioni sono operazioni scritte in sequenza che coprono più blocchi, tutti compressi. Pertanto, non è necessario considerare l'allineamento. L'unico modello di i/o che desta preoccupazione sono le sovrascritture casuali dei file.

Compaction dei dati

La data compaction è una tecnologia che migliora l'efficienza di compressione. Come indicato in precedenza, la sola compressione adattiva può garantire risparmi 2:1:1 al meglio, perché è limitata alla memorizzazione di un i/o da 8KB KB in un blocco WAFL da 4KB KB. I metodi di compressione con dimensioni dei blocchi maggiori garantiscono una maggiore efficienza. Tuttavia, non sono adatte per i dati che sono soggetti a piccole sovrascritture dei blocchi. La decompressione di 32KB unità di dati, l'aggiornamento di una porzione 8KB, la ricomprensione e la riscrittura sui dischi crea overhead.

La data compaction opera consentendo di memorizzare più blocchi logici all'interno dei blocchi fisici. Ad esempio, un database con dati altamente comprimibili come testo o blocchi parzialmente completi può comprimere da 8KB a 1KB. Senza la compaction, quei 1KB PB di dati continuerebbero ad occupare un intero blocco da 4KB KB. Inline data compaction per memorizzare 1KB TB di dati compressi in sole 1KB:1 di spazio fisico insieme ad altri dati compressi. Non si tratta di una tecnologia di compressione, ma semplicemente di un metodo più efficiente per allocare spazio sulle unità e quindi non dovrebbe creare alcun effetto rilevabile sulle prestazioni.

Il grado di risparmio ottenuto varia. I dati già compressi o crittografati non possono in genere essere ulteriormente compressi, e pertanto tali set di dati non traggono vantaggio dalla compattazione. Al contrario, i file di dati appena inizializzati contenenti poco più dei metadati dei blocchi e la compressione di zeri fino a 80:1.

Efficienza di conservazione sensibile alla temperatura

L'efficienza dello storage sensibile alla temperatura (TSSE) è disponibile in ONTAP 9,8 e versioni successive e si basa sulle mappe termiche di accesso ai blocchi per identificare i blocchi a cui si accede raramente e comprimerli con una maggiore efficienza.

Deduplica

La deduplica consiste nella rimozione di dimensioni dei blocchi duplicate da un set di dati. Ad esempio, se lo stesso blocco 4KB esistesse in 10 file diversi, la deduplica reindirizzerebbe quel blocco 4KB in tutti i file 10 allo stesso blocco fisico da 4KB KB. Il risultato sarebbe un miglioramento di 10:1 volte in efficienza per quei dati.

Dati come i LUN di avvio guest di VMware si deduplicano in genere in modo estremamente efficace poiché sono costituiti da più copie degli stessi file del sistema operativo. Sono state osservate un'efficienza pari o superiore a 100:1.

Alcuni dati non contengono dati duplicati. Ad esempio, un blocco Oracle contiene un'intestazione univoca a livello globale per il database e un trailer quasi univoco. Di conseguenza, la deduplica di un database Oracle raramente offre un risparmio superiore al 1%. La deduplica con i database MS SQL è leggermente migliore, ma i metadati univoci a livello di blocco rimangono un limite.

In pochi casi, sono stati osservati risparmi di spazio fino al 15% nei database con blocchi di dimensioni grandi e 16KB. Il 4KB iniziale di ciascun blocco contiene la testata unica a livello globale, mentre il 4KB finale contiene il rimorchio quasi unico. I blocchi interni sono candidati per la deduplica, sebbene in pratica ciò sia quasi interamente attribuito alla deduplica di dati azzerati.

Molti array della concorrenza rivendicano la capacità di deduplicare i database sulla base del presupposto che un database venga copiato più volte. Anche in questo caso è possibile utilizzare la deduplica NetApp, ma ONTAP offre un'opzione migliore: La tecnologia FlexClone di NetApp. Il risultato finale è lo stesso; vengono create più copie di un database che condividono la maggior parte dei blocchi fisici sottostanti. L'utilizzo di FlexClone è molto più efficiente della necessità di dedicare tempo alla copia e alla deduplica dei file di database. In effetti, non viene effettuata alcuna duplicazione piuttosto che deduplica, poiché al primo posto non viene mai creato un duplicato.

Efficienza e thin provisioning

Le funzionalità di efficienza sono forme di thin provisioning. Ad esempio, una LUN da 100GB GB che occupa un volume da 100GB GB potrebbe comprimere fino a 50GB GB. Non ci sono risparmi effettivi ancora realizzati perché il volume è ancora 100GB. Le dimensioni del volume devono essere innanzitutto ridotte in modo che lo spazio salvato possa essere utilizzato in un'altra posizione del sistema. Se successivamente le modifiche apportate al LUN da 100GB GB rendono i dati meno comprimibili, il LUN aumenta le dimensioni e il volume potrebbe riempirsi.

Il thin provisioning è vivamente consigliato in quanto consente di semplificare la gestione, offrendo al contempo un sostanziale miglioramento della capacità utilizzabile con conseguenti risparmi sui costi. Il motivo è semplice: Gli ambienti di database includono spesso molto spazio vuoto, un elevato numero di volumi e LUN e dati comprimibili. Il thick provisioning crea la riserva di spazio sullo storage per volumi e LUN, nel caso in cui un giorno raggiungano il 100% di riempimento e contengano dati non comprimibili al 100%. È improbabile che ciò accada mai. Il thin provisioning consente di recuperare lo spazio e di utilizzarlo altrove e consente la

gestione della capacità basata sul sistema storage stesso piuttosto che su molti volumi e LUN più piccoli.

Alcuni clienti preferiscono utilizzare il thick provisioning, per carichi di lavoro specifici o generalmente basato su pratiche operative e di approvvigionamento consolidate.

Attenzione: se un volume viene sottoposto a thick provisioning, è necessario fare attenzione a disattivare completamente tutte le funzioni di efficienza per quel volume, inclusa la decompressione e la rimozione della deduplica tramite `sis undo` comando. Il volume non dovrebbe essere visualizzato in `volume efficiency show output`. In tal caso, il volume è ancora parzialmente configurato per le funzioni di efficienza. Di conseguenza, la sovrascrittura garantisce un funzionamento diverso, aumentando le possibilità che le sovrascritture causino l'esaurimento inaspettato dello spazio del volume, con conseguenti errori di i/o del database.

Best practice di efficienza

NetApp consiglia di:

Valori predefiniti AFF

I volumi creati su ONTAP in esecuzione su un sistema AFF all-flash vengono sottoposti a thin provisioning con tutte le funzionalità di efficienza inline abilitate. Sebbene in genere i database non beneficino della deduplica e possano includere dati non comprimibili, le impostazioni predefinite sono comunque appropriate per quasi tutti i carichi di lavoro. ONTAP è progettato per elaborare in modo efficiente tutti i tipi di dati e gli schemi i/o, indipendentemente dal fatto che comportino risparmi. Le impostazioni predefinite devono essere modificate solo se le ragioni sono pienamente comprese e se vi è un vantaggio a deviare.

Raccomandazioni generali

- Se i volumi e/o le LUN non sono dotati di thin provisioning, è necessario disabilitare tutte le impostazioni di efficienza perché queste funzioni non offrono risparmi e la combinazione del thick provisioning con l'efficienza dello spazio può causare comportamenti imprevisti, inclusi errori di spazio esaurito.
- Se i dati non sono soggetti a sovrascritture, ad esempio con i backup o i log delle transazioni dei database, puoi ottenere una maggiore efficienza abilitando TSSE con un periodo di raffreddamento ridotto.
- Alcuni file potrebbero contenere una quantità significativa di dati non comprimibili, ad esempio quando la compressione è già abilitata a livello di applicazione dei file sono crittografati. Se uno di questi scenari è vero, considerare la possibilità di disattivare la compressione per consentire un funzionamento più efficiente su altri volumi che contengono dati comprimibili.
- Non utilizzare sia la compressione 32KB che la deduplica con i backup del database. Vedere la sezione [Compressione adattiva](#) per ulteriori informazioni.

Thin provisioning con database Oracle

Il thin provisioning per un database Oracle richiede un'attenta pianificazione, perché ne consegue che è possibile configurare più spazio su un sistema di storage rispetto a quello necessariamente fisicamente disponibile. Vale la pena di fare tutto questo perché, se eseguito correttamente, il risultato è un notevole risparmio sui costi e un miglioramento della gestibilità.

Il thin provisioning è disponibile in molte forme e rappresenta parte integrante di molte funzionalità offerte da ONTAP a un ambiente applicativo aziendale. Il thin provisioning è inoltre strettamente correlato alle tecnologie di efficienza per lo stesso motivo: Le funzionalità di efficienza consentono di memorizzare dati più logici

rispetto a quanto tecnicamente esistente nel sistema storage.

Quasi tutti gli utilizzi delle snapshot implicano il thin provisioning. Ad esempio, un tipico database da 10TB TB su storage NetApp include circa 30 giorni di snapshot. Questa disposizione risulta in circa 10TB di dati visibili nel file system attivo e 300TB dedicati agli snapshot. In genere, il 310TB GB di storage totale risiede su un totale di circa 12TB - 15TB GB di spazio. Il database attivo consuma 10TB e i restanti 300TB di dati richiedono solo da 2TB a 5TB di spazio, in quanto vengono memorizzate solo le modifiche apportate ai dati originali.

Anche il cloning è un esempio di thin provisioning. Un importante cliente NetApp ha creato 40 cloni di un database da 80TB TB per l'utilizzo da parte dello sviluppo. Se tutti i 40 sviluppatori che utilizzano questi cloni sovrascrivono ogni blocco in ogni file dati, sarebbero necessari oltre 3,2PB TB di storage. In pratica, il turnover è basso e il requisito di spazio collettivo è più vicino a 40TB, perché solo le modifiche sono memorizzate sui drive.

Gestione dello spazio

È necessario prestare particolare attenzione al thin provisioning di un ambiente applicativo, perché la velocità di modifica dei dati può aumentare inaspettatamente. Ad esempio, il consumo di spazio dovuto agli snapshot può crescere rapidamente se le tabelle di database vengono riindicizzate o se viene applicata una patch su larga scala ai guest VMware. Un backup posizionato in modo errato può scrivere una grande quantità di dati in un tempo molto breve. Infine, può essere difficile recuperare alcune applicazioni se un file system esaurisce inaspettatamente lo spazio libero.

Fortunatamente, questi rischi possono essere risolti con un'attenta configurazione di `volume-autogrow` e `snapshot-autodelete` criteri: Come indicato dai nomi, queste opzioni consentono a un utente di creare policy in grado di liberare automaticamente lo spazio occupato dalle snapshot o di far crescere un volume per ospitare dati aggiuntivi. Sono disponibili molte opzioni e le esigenze variano in base al cliente.

Vedere "[documentazione per la gestione logica dello storage](#)" per una discussione completa di queste funzioni.

Prenotazioni frazionarie

Riserva frazionaria si riferisce al comportamento di un LUN in un volume rispetto all'efficienza dello spazio. Quando l'opzione `fractional-reserve` è impostato al 100%, tutti i dati nel volume possono subire un turnover del 100% con qualsiasi modello di dati senza esaurire lo spazio sul volume.

Ad esempio, si consideri un database su una singola LUN da 250GB GB in un volume da 1TB GB. La creazione di uno snapshot comporterebbe immediatamente la riserva di ulteriori 250GB GB di spazio nel volume per garantire che il volume non esaurisca lo spazio per alcun motivo. L'utilizzo di riserve frazionarie comporta in genere uno spreco di risorse poiché è estremamente improbabile che ogni byte nel volume del database debba essere sovrascritto. Non c'è motivo di riservare spazio per un evento che non si verifica mai. Tuttavia, se un cliente non è in grado di monitorare il consumo di spazio in un sistema di storage e deve essere certo che lo spazio non si esaurisce mai, sarebbero necessarie prenotazioni frazionarie del 100% per utilizzare gli snapshot.

Compressione e deduplica

Compressione e deduplica sono entrambe forme di thin provisioning. Ad esempio, un impatto dei dati di 50TB:1 potrebbe comprimere fino a 30TB:1, ottenendo un risparmio di 20TB:1. Affinché la compressione possa produrre vantaggi, alcuni di questi 20TB TB devono essere utilizzati per altri dati, altrimenti il sistema storage deve essere acquistato con meno di 50TB TB. In questo modo è possibile memorizzare una quantità di dati superiore rispetto a quella tecnicamente disponibile sul sistema storage. Dal punto di vista dei dati, i dati sono 50TB, anche se occupano solo 30TB sulle unità.

Esiste sempre la possibilità che la compressibilità di un set di dati cambi, con conseguente aumento del consumo di spazio reale. Questo aumento dei consumi implica che la compressione deve essere gestita come con altre forme di thin provisioning in termini di monitoraggio e utilizzo `volume-autogrow` e `snapshot-autodelete`.

Compressione e deduplica sono descritte in dettaglio nella sezione <xref:./oracle/efficiency.html>

Compressioni e prenotazioni frazionarie

La compressione è una forma di thin provisioning. Le prenotazioni frazionarie influiscono sull'utilizzo della compressione, con una nota importante; lo spazio viene riservato prima della creazione dell'istantanea. Normalmente, la riserva frazionaria è importante solo se esiste uno snapshot. Se non è presente uno snapshot, la riserva frazionaria non è importante. Questo non è il caso della compressione. Se viene creata una LUN su un volume con compressione, ONTAP preserva lo spazio per ospitare uno snapshot. Questo comportamento può creare confusione durante la configurazione, ma è previsto.

Ad esempio, consideriamo un volume da 10GB GB con una LUN da 5GB GB compressa a 2,5GB GB senza snapshot. Considerare questi due scenari:

- Riserva frazionaria = 100 risultati in 7,5GB utilizzo
- Riserva frazionaria = 0 risultati in 2,5GB utilizzo

Il primo scenario include 2,5GB di consumo di spazio per i dati attuali e 5GB di spazio per rappresentare il 100% di fatturato della fonte in previsione dell'utilizzo di snapshot. Il secondo scenario non riserva spazio aggiuntivo.

Sebbene questa situazione possa sembrare confusa, è improbabile che si verifichi nella pratica. La compressione implica thin provisioning e il thin provisioning in un ambiente LUN richiede prenotazioni frazionarie. È sempre possibile sovrascrivere i dati compressi con qualcosa di non comprimibile, il che significa che un volume deve essere sottoposto a thin provisioning per la compressione per consentire qualsiasi risparmio.

NetApp consiglia le seguenti configurazioni riservate:



- Impostare `fractional-reserve` a 0 quando è in atto il monitoraggio della capacità di base con `volume-autogrow` e `snapshot-autodelete`.
- Impostare `fractional-reserve` a 100 se non vi è alcuna capacità di monitoraggio o se è impossibile scaricare lo spazio in qualsiasi circostanza.

Spazio libero e allocazione di spazio LVM

L'efficienza del thin provisioning delle LUN attive in un ambiente di file system può andare persa nel tempo quando i dati vengono eliminati. A meno che i dati eliminati non vengano sovrascritti con zero (vedere anche "[ASMRU](#)") Oppure lo spazio viene liberato con il recupero dello spazio TRIM/UNMAP, i dati "cancellati" occupano sempre più spazi vuoti non allocati nel file system. Inoltre, in molti ambienti di database il thin provisioning delle LUN attive è limitato, in quanto i file di dati vengono inizializzati alle dimensioni massime al momento della creazione.

Un'attenta pianificazione della configurazione LVM può migliorare l'efficienza e ridurre al minimo la necessità di provisioning dello storage e di ridimensionamento delle LUN. Quando si utilizza un LVM come Veritas VxVM o Oracle ASM, le LUN sottostanti vengono suddivise in estensioni che vengono utilizzate solo quando necessario. Ad esempio, se un set di dati inizia a 2TB TB ma potrebbe crescere fino a 10TB TB con il passare del tempo, è possibile inserire il set di dati in 10TB LUN con thin provisioning organizzati in un gruppo di dischi

LVM. Occupa solo 2TB GB di spazio al momento della creazione e richiederebbe spazio aggiuntivo solo se le estensioni sono allocate per ospitare la crescita dei dati. Questo processo è sicuro finché lo spazio è monitorato.

Failover/switchover dei database Oracle e del controller ONTAP

Le operazioni di takeover e switchover dello storage devono garantire l'integrità delle operazioni dei database Oracle. Inoltre, gli argomenti utilizzati dalle operazioni di takeover e switchover possono influire sull'integrità dei dati se utilizzati in modo errato.

- In condizioni normali, le scritture in arrivo su un dato controller vengono mirrorate in modo sincrono per il partner. In un ambiente NetApp MetroCluster, le scritture vengono anche mirrorate su un controller remoto. Fino a quando non viene memorizzata in un supporto non volatile in tutte le posizioni, la scrittura non viene riconosciuta all'applicazione host.
- Il supporto che memorizza i dati di scrittura è chiamato memoria non volatile o NVMEM. Viene anche talvolta indicata come memoria non volatile ad accesso casuale (NVRAM, nonvolatile Random Access Memory), e può essere considerata come una cache di scrittura anche se funziona come un journal. In condizioni normali, i dati provenienti da NVMEM non vengono letti; vengono utilizzati solo per proteggere i dati in caso di guasti software o hardware. Quando i dati vengono scritti sulle unità disco rigido, i dati vengono trasferiti dalla RAM nel sistema, non da NVMEM.
- Durante un'operazione di takeover, un nodo di una coppia ha (high Availability) assume il controllo delle operazioni dal partner. Lo switchover è praticamente identico, ma si applica alle configurazioni MetroCluster in cui un nodo remoto assume le funzioni di un nodo locale.

Durante le normali operazioni di manutenzione, un'operazione di takeover o switchover dello storage deve essere trasparente, ad eccezione di una potenziale breve pausa nelle operazioni in base al cambiamento dei percorsi di rete. Il networking può rivelarsi complesso, tuttavia, ed è facile commettere errori, pertanto NetApp consiglia vivamente di eseguire accuratamente le operazioni di takeover e switchover prima di mettere in produzione un sistema storage. In questo modo, è possibile verificare che tutti i percorsi di rete siano configurati correttamente. In un ambiente SAN, controllare attentamente l'output del comando `sanlun lun show -p` per assicurarsi che tutti i percorsi primario e secondario previsti siano disponibili.

Occorre prestare attenzione quando si rilascia un'acquisizione forzata o uno switchover. Imporre una modifica alla configurazione dello storage con queste opzioni significa che lo stato del controller proprietario delle unità non viene preso in considerazione e il nodo alternativo assume forzatamente il controllo delle unità. Una forzatura non corretta di un takeover può causare la perdita o il danneggiamento dei dati. Questo perché un takeover o uno switchover forzato può scartare il contenuto di NVMEM. Una volta completato il takeover o lo switchover, la perdita dei dati potrebbe riportare i dati memorizzati nelle unità a uno stato leggermente più vecchio dal punto di vista del database.

Raramente dovrebbe essere necessario un takeover forzato con una normale coppia ha. In quasi tutti gli scenari di errore, un nodo si arresta e informa il partner in modo che si verifichi un failover automatico. In alcuni casi, ad esempio in caso di guasto permanente che causa la perdita dell'interconnessione tra i nodi e la perdita di un controller, è necessario eseguire un takeover forzato. In una situazione del genere, il mirroring tra i nodi viene perso prima del guasto del controller, il che significa che il controller rimasto non avrebbe più una copia delle scritture in corso. Il takeover deve quindi essere forzato, il che significa che potenzialmente i dati vengono persi.

La stessa logica si applica a uno switchover MetroCluster. In condizioni normali, lo switchover è quasi trasparente. Tuttavia, un disastro può causare una perdita di connettività tra il sito rimasto e il sito disastroso. Dal punto di vista del sito sopravvissuto, il problema potrebbe essere nient'altro che un'interruzione della

connettività tra i siti, e il sito originale potrebbe ancora elaborare i dati. Se un nodo non è in grado di verificare lo stato del controller primario, è possibile solo uno switchover forzato.

NetApp raccomanda adottare le seguenti precauzioni:



- Prestare molta attenzione a non forzare accidentalmente un'acquisizione o uno switchover. In genere, non è necessario forzare e forzare la modifica può causare la perdita di dati.
- Se è necessario un takeover o uno switchover forzato, assicurarsi che le applicazioni vengano arrestate, che tutti i file system vengano dismontati e che i gruppi di volumi LVM (Logical Volume Manager) siano diversi. I gruppi di dischi ASM devono essere smontati.
- In caso di switchover MetroCluster forzato, scollegare il nodo guasto da tutte le risorse di storage rimaste. Per ulteriori informazioni, consultare la Guida alla gestione e al ripristino di emergenza di MetroCluster per la versione pertinente di ONTAP.

MetroCluster e aggregati multipli

MetroCluster è una tecnologia di replica sincrona che passa alla modalità asincrona in caso di interruzione della connettività. Questa è la richiesta più comune da parte dei clienti, perché la replica sincrona garantita significa che l'interruzione della connettività del sito porta a uno stallo completo dell'i/o del database, mettendo il database fuori servizio.

Con MetroCluster, gli aggregati vengono sincronizzati rapidamente dopo il ripristino della connettività. A differenza di altre tecnologie di storage, MetroCluster non dovrebbe mai richiedere un reindirizzamento completo in seguito a un guasto del sito. È necessario spedire solo le modifiche delta.

Nei set di dati estesi agli aggregati c'è solo un piccolo rischio che occorran passaggi aggiuntivi di recovery dei dati in uno scenario di emergenza regolare. In particolare, se (a) la connettività tra siti viene interrotta, (b) la connettività viene ripristinata, (c) gli aggregati raggiungono uno stato in cui alcuni sono sincronizzati e alcuni non lo sono, quindi (d) il sito primario viene perso, il risultato è un sito sopravvissuto in cui gli aggregati non sono sincronizzati tra loro. In tal caso, parti del set di dati vengono sincronizzate tra loro e non è possibile ripristinare applicazioni, database o datastore senza un recovery. Se un set di dati si estende agli aggregati, NetApp consiglia vivamente di sfruttare i backup basati su snapshot con uno dei molti strumenti disponibili, per verificare la possibilità di recupero rapido in questo scenario insolito.

Informazioni sul copyright

Copyright © 2024 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIALIZZABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEGUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina <http://www.netapp.com/TM> sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.