



# Configurazione del database

## Enterprise applications

NetApp  
May 09, 2024

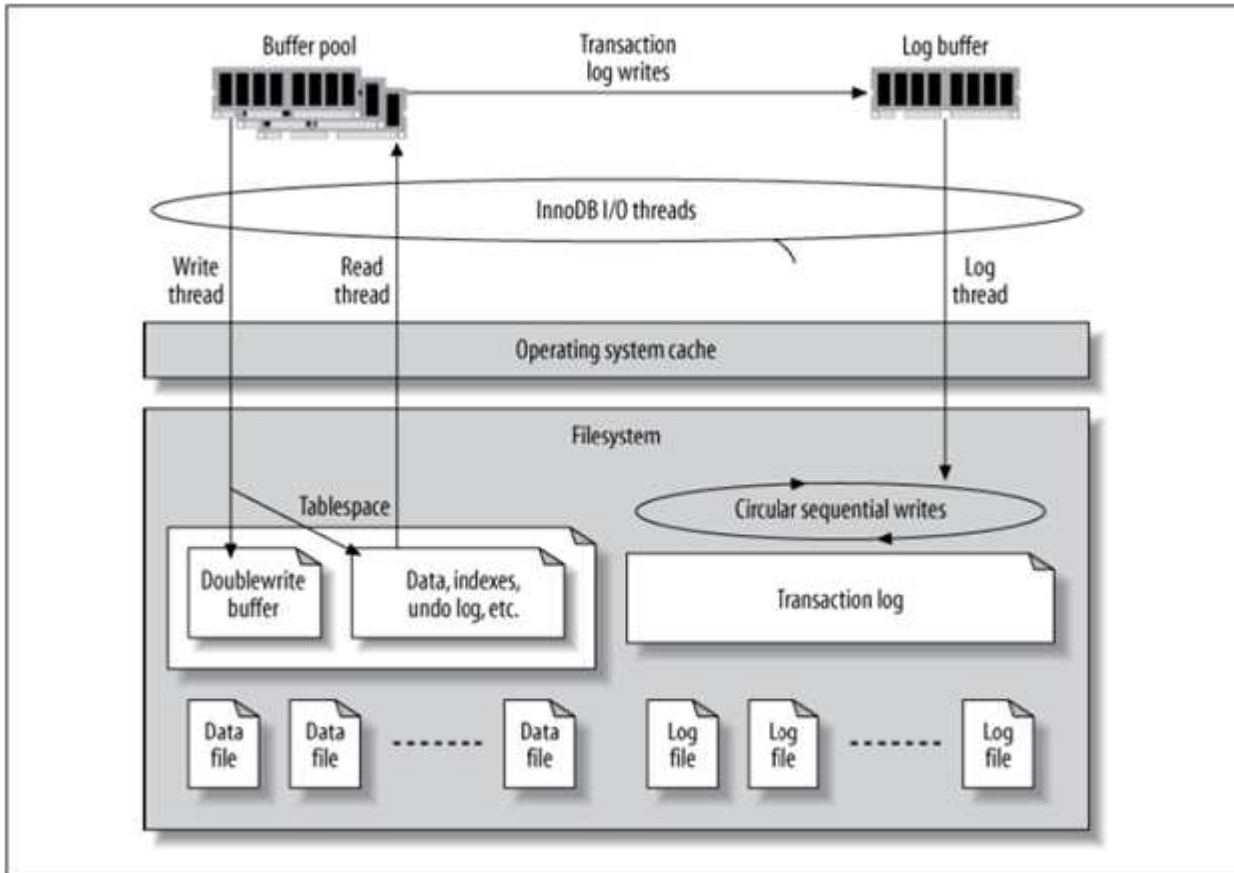
# Sommario

Configurazione del database .....	1
MySQL e InnoDB .....	1
Parametri di configurazione MySQL .....	3
innodb_log_file_size .....	4
innodb_flush_log_at_trx_commit .....	4
innodb_doublewrite .....	5
innodb_buffer_pool_size .....	5
innodb_flush_method .....	5
innodb_io_capacity .....	6
innodb_lru_scan_depth .....	7
open_file_limits .....	7

# Configurazione del database

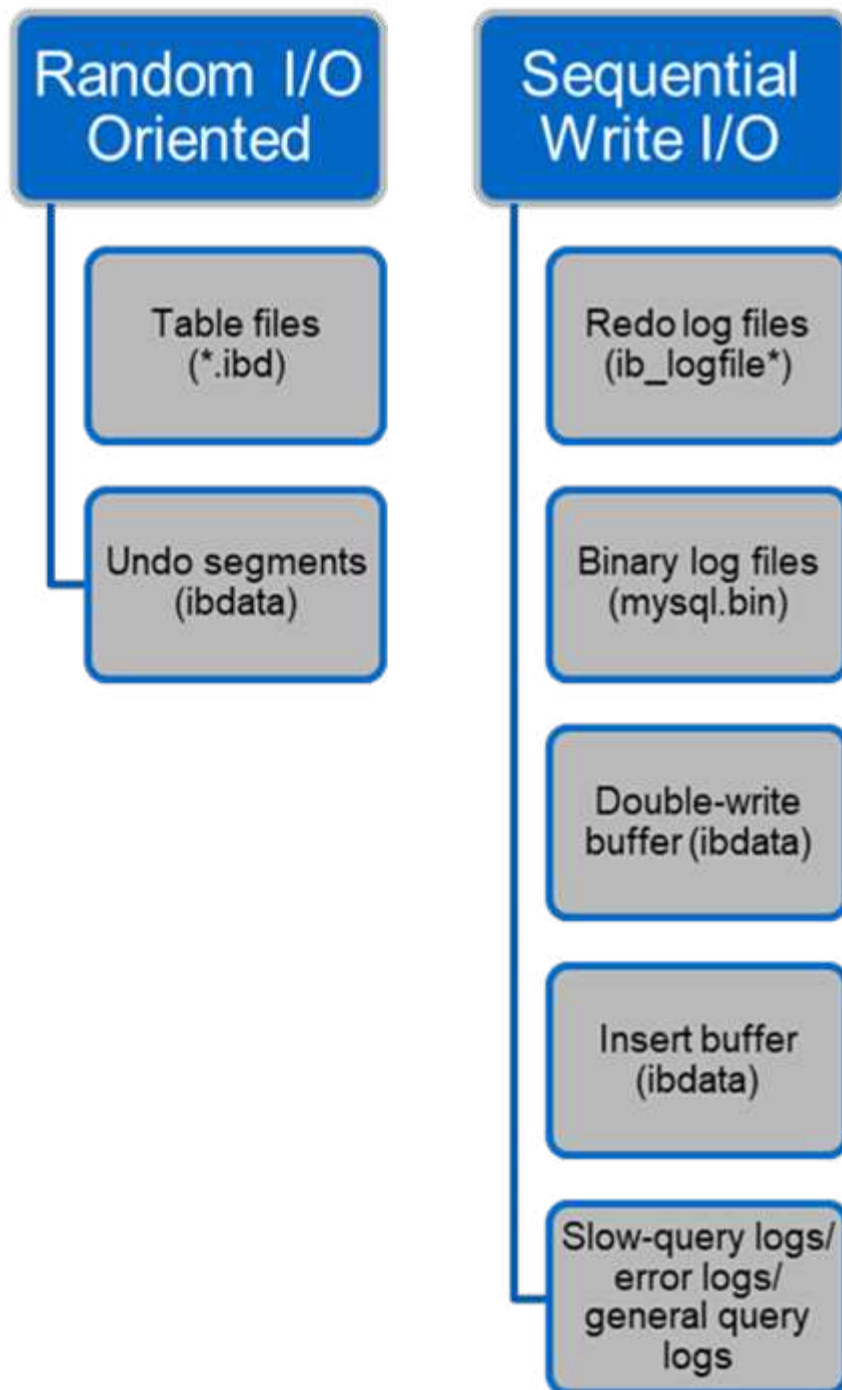
## MySQL e InnoDB

InnoDB funge da livello intermedio tra lo storage e il server MySQL, e memorizza i dati nelle unità.



I/o MySQL è suddiviso in due tipi:

- I/o di file casuali
- I/o di file sequenziale



I file di dati vengono letti e sovrascritti in modo casuale, con conseguente aumento degli IOPS. Pertanto, si consiglia di utilizzare l'unità SSD.

I file di log di ripristino e i file di log binari sono registri transazionali. Vengono scritti in sequenza, così potrai ottenere buone performance sul disco HDD con cache in scrittura. Al momento del ripristino si verifica una lettura sequenziale, che raramente causa problemi di prestazioni, poiché le dimensioni dei file di registro sono in genere inferiori ai file di dati e le letture sequenziali sono più veloci delle letture casuali (che si verificano sui file di dati).

Il buffer double-write è una caratteristica speciale di InnoDB. InnoDB prima scrive le pagine svuotate nel buffer di doppia scrittura e poi scrive le pagine nelle posizioni corrette sui file di dati. Questo processo impedisce il

danneggiamento della pagina. Senza il buffer di scrittura doppia, la pagina potrebbe danneggiarsi se si verifica un'interruzione dell'alimentazione durante il processo di scrittura su unità. La scrittura nel buffer double-write è sequenziale, pertanto è altamente ottimizzato per gli HDD. Al momento del ripristino vengono eseguite letture sequenziali.

Poiché la NVRAM ONTAP fornisce già la protezione in scrittura, non è necessario il doppio buffer in scrittura. MySQL ha un parametro, `skip_innodb_doublewrite`, per disattivare il buffer di doppia scrittura. Questa funzione può migliorare notevolmente le prestazioni.

Il buffer insert è anche una caratteristica speciale di InnoDB. Se i blocchi di indice secondari non univoci non sono in memoria, InnoDB inserisce le voci nel buffer di inserimento per evitare operazioni di i/o casuali. Periodicamente, il buffer di inserimento viene Unito agli alberi di indice secondari nel database. Il buffer di inserimento riduce il numero di operazioni di i/o unendo le richieste di i/o allo stesso blocco; le operazioni di i/o casuali possono essere sequenziali. Anche il buffer di inserimento è altamente ottimizzato per gli HDD. Durante le normali operazioni, vengono eseguite operazioni di scrittura e lettura sequenziali.

I segmenti di annullamento sono orientati all'i/o casuale. Per garantire la concorrenza multi-versione (MVCC), InnoDB deve registrare le vecchie immagini nei segmenti di annullamento. La lettura delle immagini precedenti dai segmenti di annullamento richiede letture casuali. Se si esegue una transazione lunga con letture ripetibili (come `mysqldump`, una singola transazione) o si esegue una query lunga, è possibile che si verifichino letture casuali. Pertanto, in questo caso è preferibile memorizzare i segmenti di annullamento negli SSD. Se si eseguono solo transazioni o query brevi, le letture casuali non costituiscono un problema.

**NetApp consiglia** il seguente layout di progettazione dello storage a causa delle caratteristiche i/o di InnoDB.



- Un unico volume per memorizzare i file di MySQL orientati ai/o casuali e sequenziali
- Un altro volume per memorizzare i file di MySQL orientati a i/o puramente sequenziali

Questo layout aiuta inoltre a progettare politiche e strategie di protezione dei dati.

## Parametri di configurazione MySQL

NetApp consiglia alcuni importanti parametri di configurazione di MySQL per ottenere prestazioni ottimali.

Parametri	Valori
<code>innodb_log_file_size</code>	256M
<code>innodb_flush_log_at_trx_commit</code>	2
<code>innodb_doublewrite</code>	0
<code>innodb_flush_method</code>	<code>fsync</code>
<code>innodb_buffer_pool_size</code>	11G
<code>innodb_io_capacity</code>	8192
<code>innodb_buffer_pool_instances</code>	8
<code>innodb_lru_scan_depth</code>	8192
<code>open_file_limit</code>	65535

Per impostare i parametri descritti in questa sezione, è necessario modificarli nel file di configurazione MySQL (my.cnf). Le Best practice di NetApp sono il risultato di test eseguiti internamente.

## innodb\_log\_file\_size

La scelta della dimensione corretta per il file di log InnoDB è importante per le operazioni di scrittura e per avere un tempo di ripristino decente dopo un arresto anomalo del server.

Poiché molte transazioni sono registrate nel file, la dimensione del file di registro è importante per le operazioni di scrittura. Quando i record vengono modificati, la modifica non viene immediatamente riscritta nello spazio di tabella. La modifica viene invece registrata alla fine del file di registro e la pagina viene contrassegnata come sporca. InnoDB utilizza il proprio registro per convertire l'i/o casuale in i/o sequenziale

Quando il log è pieno, la pagina sporca viene scritta nello spazio di tabella in sequenza per liberare spazio nel file di log. Ad esempio, si supponga che un server si blocchi nel corso di una transazione e che le operazioni di scrittura vengano registrate solo nel file di registro. Prima che il server possa tornare attivo, deve passare attraverso una fase di recupero in cui vengono riprodotte le modifiche registrate nel file di registro. Maggiore è il numero di voci presenti nel file di registro, maggiore sarà il tempo necessario al server per il ripristino.

In questo esempio, la dimensione del file di registro influisce sia sul tempo di ripristino che sulle prestazioni di scrittura. Quando si sceglie il numero giusto per la dimensione del file di registro, bilanciare il tempo di ripristino rispetto alle prestazioni di scrittura. In genere, qualsiasi valore compreso tra 128M e 512M è un buon valore.

## innodb\_flush\_log\_at\_trx\_commit

In caso di modifica dei dati, la modifica non viene immediatamente scritta nell'archivio.

I dati vengono invece registrati in un buffer di registro, che è una porzione di memoria allocata da InnoDB alle modifiche del buffer registrate nel file di registro. InnoDB svuota il buffer nel file di registro quando viene eseguito il commit di una transazione, quando il buffer diventa pieno o una volta al secondo, a seconda dell'evento che si verifica per primo. La variabile di configurazione che controlla questo processo è `innodb_flush_log_at_trx_commit`. Le opzioni valore includono:

- Quando si imposta `innodb_flush_log_trx_at_commit=0`, InnoDB scrive i dati modificati (nel pool di buffer InnoDB) nel file di log (`ib_logfile`) e scarica il file di log (write to storage) ogni secondo. Tuttavia, non fa nulla quando la transazione è impegnata. Se si verifica un'interruzione dell'alimentazione o un arresto anomalo del sistema, nessuno dei dati non scaricati è recuperabile perché non vengono scritti né nel file di registro né nelle unità.
- Quando si imposta `innodb_flush_log_trx_commit=1`, InnoDB scrive il buffer di log nel log delle transazioni e lo svuota nello storage durevole per ogni transazione. Ad esempio, per tutti i commit delle transazioni, InnoDB scrive nel registro e quindi nello storage. La lentezza dello storage influisce negativamente sulle performance, ad esempio riducendo il numero di transazioni InnoDB al secondo.
- Quando si imposta `innodb_flush_log_trx_commit=2`, InnoDB scrive il buffer di log nel file di log ad ogni commit; tuttavia, non scrive dati nell'archivio. InnoDB scarica i dati una volta al secondo. Anche in caso di interruzione dell'alimentazione o arresto anomalo del sistema, i dati dell'opzione 2 sono disponibili nel file di registro ed è recuperabile.

Se l'obiettivo principale è la prestazione, impostare il valore su 2. Poiché InnoDB scrive sui dischi una volta al secondo, non per ogni commit delle transazioni, le performance migliorano in modo significativo. Se si verifica un'interruzione dell'alimentazione o un arresto anomalo, i dati possono essere recuperati dal registro delle transazioni.

Se l'obiettivo principale è la sicurezza dei dati, impostare il valore su 1 in modo che per ogni commit di transazione, InnoDB si scarichi sulle unità. Tuttavia, le prestazioni potrebbero risentirne.



**NetApp recommended** impostare il valore `innodb_flush_log_trx_commit` su 2 per ottenere prestazioni migliori.

## innodb\_doublewrite

Quando `innodb_doublewrite` È attivato (impostazione predefinita), InnoDB memorizza tutti i dati due volte: Prima nel buffer di doppia scrittura e poi nei file di dati effettivi.

È possibile disattivare questo parametro con `--skip-innodb_doublewrite` per i benchmark o quando siete più preoccupati per le prestazioni superiori che l'integrità dei dati o possibili guasti. InnoDB utilizza una tecnica di scaricamento file chiamata double-write. Prima di scrivere le pagine nei file di dati, InnoDB le scrive in un'area contigua denominata buffer double-write. Una volta completata la scrittura e lo scarico nel buffer di doppia scrittura, InnoDB scrive le pagine nelle posizioni corrette nel file di dati. Se il sistema operativo o un processo `mysqld` si blocca durante la scrittura di una pagina, InnoDB può in seguito trovare una buona copia della pagina dal buffer di doppia scrittura durante il recupero del crash.



**NetApp recommended** disabilitare il buffer double-write. La NVRAM ONTAP svolge la stessa funzione. Il doppio buffering danneggia inutilmente le prestazioni.

## innodb\_buffer\_pool\_size

Il pool di buffer InnoDB è la parte più importante di qualsiasi attività di ottimizzazione.

InnoDB si affida in gran parte al pool di buffer per la memorizzazione nella cache degli indici e il reaming dei dati, all'indice hash adattivo, al buffer insert e a molte altre strutture di dati utilizzate internamente. Il pool di buffer memorizza inoltre le modifiche ai dati in modo che le operazioni di scrittura non debbano essere eseguite immediatamente nello storage, migliorando così le prestazioni. Il pool di buffer è parte integrante di InnoDB e le sue dimensioni devono essere regolate di conseguenza. Per impostare le dimensioni del pool di buffer, tenere conto dei seguenti fattori:

- Per una macchina dedicata solo InnoDB, impostare la dimensione del pool di buffer su 80% o più della RAM disponibile.
- Se non si tratta di un server dedicato MySQL, impostare la dimensione al 50% della RAM.

## innodb\_flush\_method

Il parametro `innodb_flush_method` specifica come InnoDB apre e svuota i file di log e di dati.

### Ottimizzazioni

Nell'ottimizzazione InnoDB, l'impostazione di questo parametro modifica le prestazioni del database, se applicabile.

Le seguenti opzioni consentono di svuotare i file tramite InnoDB:

- `fsync`. InnoDB utilizza `fsync()` chiamata di sistema per cancellare sia i file di dati che i file di registro. Questa opzione è l'impostazione predefinita.
- `O_DSYNC`. InnoDB utilizza `O_DSYNC` possibilità di aprire e svuotare i file di log e `fsync()` per svuotare i file di dati. InnoDB non utilizza `O_DSYNC` Direttamente, perché ci sono stati problemi con esso su molte varietà di UNIX.
- `O_DIRECT`. InnoDB utilizza `O_DIRECT` (oppure `directio()` Su Solaris) per aprire i file di dati e gli usi `fsync()` per cancellare sia i file di dati che i file di registro. Questa opzione è disponibile su alcune versioni di GNU/Linux, FreeBSD e Solaris.
- `O_DIRECT_NO_FSYNC`. InnoDB utilizza `O_DIRECT` Durante lo spurgo dell'i/o, tuttavia, salta `fsync()` chiamata di sistema successiva. Questa opzione non è adatta per alcuni tipi di file system (ad esempio, XFS). Se non si è certi che il file system richieda un `fsync()` chiamata di sistema, ad esempio per conservare tutti i metadati dei file, utilizzare `O_DIRECT` invece.

## Osservazione

Nei test di laboratorio di NetApp, il `fsync` L'opzione predefinita è stata utilizzata su NFS e SAN ed è stata un'improvvisazione per le prestazioni eccezionale rispetto a `O_DIRECT`. Mentre si utilizza il metodo di lavaggio come `O_DIRECT` Con ONTAP, abbiamo osservato che il client scrive molte scritte a byte singolo al margine del blocco 4096 in modo seriale. Queste operazioni di scrittura hanno aumentato la latenza sulla rete e degradato le performance.

## innodb\_io\_capacity

Nel plug-in InnoDB è stato aggiunto un nuovo parametro chiamato `innodb_io_Capacity` da MySQL 5,7.

Controlla il numero massimo di IOPS eseguiti da InnoDB (che include la velocità di scaricamento delle pagine sporche e la dimensione batch del buffer di inserimento [`ibuf`]). Il parametro `innodb_io_Capacity` imposta un limite massimo per le IOPS da parte delle attività in background di InnoDB, come il lavaggio delle pagine dal pool di buffer e l'Unione dei dati dal buffer di modifica.

Impostare il parametro `innodb_io_Capacity` sul numero approssimativo di operazioni di i/o che il sistema può eseguire al secondo. Idealmente, mantenere l'impostazione più bassa possibile, ma non così bassa che le attività in background rallentano. Se l'impostazione è troppo alta, i dati vengono rimossi dal pool di buffer e il buffer di inserimento troppo rapidamente per la memorizzazione nella cache, per fornire un vantaggio significativo.



**NetApp consiglia** che, se si utilizza questa impostazione su NFS, analizzi il risultato del test di IOPS (SysBench/FiO) e imposti il parametro di conseguenza. Utilizzare il valore più piccolo possibile per lo spurgo e lo spurgo per continuare a meno che non vengano visualizzate pagine modificate o sporche di quanto si desidera nel pool di buffer InnoDB.



Non utilizzare valori estremi come 20.000 o più a meno che non si sia dimostrato che valori inferiori non sono sufficienti per il carico di lavoro.

Il parametro `InnoDB_io_Capacity` regola le velocità di lavaggio e i/o correlati



È possibile danneggiare seriamente le prestazioni impostando questo parametro o il parametro `innodb_io_Capacity_max` troppo alto e spreco delle operazioni di i/o con il lavaggio prematuro.



## innodb\_lru\_scan\_depth

Il `innodb_lru_scan_depth` Parametro influenza gli algoritmi e le euristiche dell'operazione di scaricamento per il pool di buffer InnoDB.

Questo parametro è principalmente di interesse per gli esperti di performance che ottimizzano i carichi di lavoro i/o-intensive. Per ogni istanza del pool di buffer, questo parametro specifica fino a che punto nell'elenco di pagine LRU (Last Recently Used) il thread di pulizia della pagina deve continuare la scansione, cercando le pagine sporche da eliminare. Questa operazione in background viene eseguita una volta al secondo.

È possibile regolare il valore verso l'alto o verso il basso per ridurre al minimo il numero di pagine libere. Non impostare un valore molto superiore al necessario, poiché le scansioni possono avere un costo significativo in termini di prestazioni. Inoltre, è consigliabile regolare questo parametro quando si modifica il numero di istanze del pool di buffer, perché `innodb_lru_scan_depth * innodb_buffer_pool_instances` definisce la quantità di lavoro eseguito dal filo del pulitore di pagina ogni secondo.

Un'impostazione più piccola di quella predefinita è adatta per la maggior parte dei carichi di lavoro. Considerare l'aumento del valore solo se si dispone di capacità i/o di riserva con un workload tipico. Per contro, se un carico di lavoro con un numero elevato di operazioni di scrittura satura la capacità i/o, diminuirne il valore, soprattutto se si dispone di un pool di buffer di grandi dimensioni.

## open\_file\_limits

Il `open_file_limits` parametro determina il numero di file che il sistema operativo consente a mysqld di aprire.

Il valore di questo parametro in fase di esecuzione è il valore reale consentito dal sistema e potrebbe essere diverso dal valore specificato all'avvio del server. Il valore è 0 sui sistemi in cui MySQL non può modificare il numero di file aperti. L'efficace `open_files_limit` il valore si basa sul valore specificato all'avvio del sistema (se presente) e sui valori di `max_connections` e `table_open_cache` utilizzando queste formule:

- $10 + \text{max\_connections} + (\text{table\_open\_cache} \times 2)$
- $\text{max\_connections} \times 5$
- Limite del sistema operativo se positivo
- Se il limite del sistema operativo è infinito: `open_files_limit` il valore viene specificato all'avvio; 5.000 se nessuno

Il server tenta di ottenere il numero di descrittori di file utilizzando il massimo di questi quattro valori. Se non è possibile ottenere molti descrittori, il server tenta di ottenere il numero di descrittori consentito dal sistema.

## 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.