Gestire i carichi di lavoro

ONTAP 9

NetApp
February 12, 2026

This PDF was generated from https://docs.netapp.com/it-it/ontap/performance-admin/identify-remaining-performance-capacity-task.html on February 12, 2026. Always check docs.netapp.com for the latest.

Sommario

Gestire i carichi di lavoro . . . 1
- Identifica la capacità di performance rimanente in ONTAP . . . 1
- Identificare i client o i file ad alto traffico in ONTAP . . . 3
- Throughput garantito con QoS . . . 4
  - Throughput garantito con panoramica della qualità del servizio in ONTAP . . . 4
  - Abilita o disabilita i piani di throughput ONTAP v2 . . . 10
  - Flusso di lavoro QoS di archiviazione ONTAP . . . 11
  - Imposta un limite massimo di throughput con la qualità del servizio in ONTAP . . . 12
  - Impostazione di un limite minimo di throughput con la qualità del servizio in ONTAP . . . 15
  - Utilizza i gruppi di criteri QoS adattivi in ONTAP . . . 17
  - Impostare un modello di gruppo di criteri adattivo in ONTAP . . . 21

Gestire i carichi di lavoro

Identifica la capacità di performance rimanente in ONTAP

La capacità delle performance, o *headroom*, misura la quantità di lavoro che è possibile posizionare su un nodo o su un aggregato prima che le performance dei carichi di lavoro sulla risorsa comincino ad essere influenzate dalla latenza. La conoscenza della capacità di performance disponibile nel cluster consente di eseguire il provisioning e bilanciare i carichi di lavoro.

Prima di iniziare

Per questa attività sono necessari comandi avanzati del livello di privilegio.

A proposito di questa attività

È possibile utilizzare i seguenti valori per `-object` opzione per raccogliere e visualizzare le statistiche di headroom:

- Per CPU, `resource_headroom_cpu`.
- Per gli aggregati, `resource_headroom_aggr`.

È inoltre possibile completare questa attività utilizzando Gestione di sistema e Active IQ Unified Manager.

Fasi

1. Passare al livello di privilegio avanzato:

   ```
   set -privilege advanced
   ```

2. Avvia la raccolta di statistiche in tempo reale:

   ```
   statistics start -object resource_headroom_cpu|aggr
   ```

   Ulteriori informazioni su `statistics start` nella "Riferimento al comando ONTAP".

3. Visualizzare in tempo reale le informazioni statistiche di headroom:

   ```
   statistics show -object resource_headroom_cpu|aggr
   ```

4. Tornare al privilegio amministrativo:

   ```
   set -privilege admin
   ```

Esempio

Nell'esempio seguente vengono visualizzate le statistiche medie orarie del headroom per i nodi del cluster.

È possibile calcolare la capacità di performance disponibile per un nodo sottraendo `current_utilization` contatore da `optimal_point_utilization` contatore. In questo esempio, la capacità di utilizzo per `CPU_sti2520-213` È -14% (72%-86%), il che suggerisce che la CPU è stata in media utilizzata in eccesso nell'ultima ora.

Potrebbe essere stato specificato `ewma_daily`, `ewma_weekly`, o. `ewma_monthly` ottenere le stesse informazioni in media per periodi di tempo più lunghi.

```
sti2520-2131454963690::*> statistics show -object resource_headroom_cpu
-raw -counter ewma_hourly
  (statistics show)

Object: resource_headroom_cpu
Instance: CPU_sti2520-213
Start-time: 2/9/2016 16:06:27
End-time: 2/9/2016 16:06:27
Scope: sti2520-213

    Counter                                                     Value
    -------------------------------- --------------------------------
    ewma_hourly                                                     -
                         current_ops                             4376
                     current_latency                            37719
                 current_utilization                               86
                   optimal_point_ops                             2573
               optimal_point_latency                             3589
             optimal_point_utilization                             72
     optimal_point_confidence_factor                                1

Object: resource_headroom_cpu
Instance: CPU_sti2520-214
Start-time: 2/9/2016 16:06:27
End-time: 2/9/2016 16:06:27
Scope: sti2520-214

    Counter                                                     Value
    -------------------------------- --------------------------------
    ewma_hourly                                                     -
                         current_ops                                0
                     current_latency                                0
                 current_utilization                                0
                   optimal_point_ops                                0
               optimal_point_latency                                0
             optimal_point_utilization                             71
     optimal_point_confidence_factor                                1
2 entries were displayed.
```

Ulteriori informazioni su `statistics show` nella "Riferimento al comando ONTAP".

Identificare i client o i file ad alto traffico in ONTAP

È possibile utilizzare la tecnologia ONTAP Active Objects per identificare client o file responsabili di una quantità sproporzionata di traffico cluster. Una volta identificati questi file o client "top", è possibile ribilanciare i carichi di lavoro del cluster o intraprendere altre azioni per risolvere il problema.

Prima di iniziare

Per eseguire questa attività, è necessario essere un amministratore del cluster.

Fasi

1. Visualizzare i principali client che accedono al cluster:

   `statistics top client show -node` *`node_name`* `-sort-key` *`sort_column`* `-interval` *`seconds_between_updates`* `-iterations` *`iterations`* `-max` *`number_of_instances`*

   Il seguente comando visualizza i principali client che accedono `cluster1`:

   ```
   cluster1::> statistics top client show

   cluster1 : 3/23/2016 17:59:10


                                                 *Total
           Client Vserver           Node Protocol    Ops
   -------------- ------- -------------- -------- ------
   172.17.180.170     vs4 siderop1-vsim4      nfs    668
   172.17.180.169     vs3 siderop1-vsim3      nfs    337
   172.17.180.171     vs3 siderop1-vsim3      nfs    142
   172.17.180.170     vs3 siderop1-vsim3      nfs    137
   172.17.180.123     vs3 siderop1-vsim3      nfs    137
   172.17.180.171     vs4 siderop1-vsim4      nfs     95
   172.17.180.169     vs4 siderop1-vsim4      nfs     92
   172.17.180.123     vs4 siderop1-vsim4      nfs     92
   172.17.180.153     vs3 siderop1-vsim3      nfs      0
   ```

   Ulteriori informazioni su `statistics top client show` nella "Riferimento al comando ONTAP".

2. Visualizzare i file principali a cui si accede dal cluster:

   `statistics top file show -node` *`node_name`* `-sort-key` *`sort_column`* `-interval` *`seconds_between_updates`* `-iterations` *`iterations`* `-max` *`number_of_instances`*

   Il seguente comando visualizza i file principali a cui si accede `cluster1`:

```
cluster1::> statistics top file show

cluster1 : 3/23/2016 17:59:10

                                    *Total
                     File Volume Vserver          Node    Ops
------------------------ ------ ------- -------------- ------
/vol/vol1/vm170-read.dat   vol1     vs4 siderop1-vsim4     22
/vol/vol1/vm69-write.dat   vol1     vs3 siderop1-vsim3      6
     /vol/vol2/vm171.dat   vol2     vs3 siderop1-vsim3      2
     /vol/vol2/vm169.dat   vol2     vs3 siderop1-vsim3      2
      /vol/vol2/p123.dat   vol2     vs4 siderop1-vsim4      2
      /vol/vol2/p123.dat   vol2     vs3 siderop1-vsim3      2
     /vol/vol1/vm171.dat   vol1     vs4 siderop1-vsim4      2
     /vol/vol1/vm169.dat   vol1     vs4 siderop1-vsim4      2
     /vol/vol1/vm169.dat   vol1     vs4 siderop1-vsim3      2
      /vol/vol1/p123.dat   vol1     vs4 siderop1-vsim4      2
```

Ulteriori informazioni su `statistics top file show` nella "Riferimento al comando ONTAP".

Throughput garantito con QoS

Throughput garantito con panoramica della qualità del servizio in ONTAP

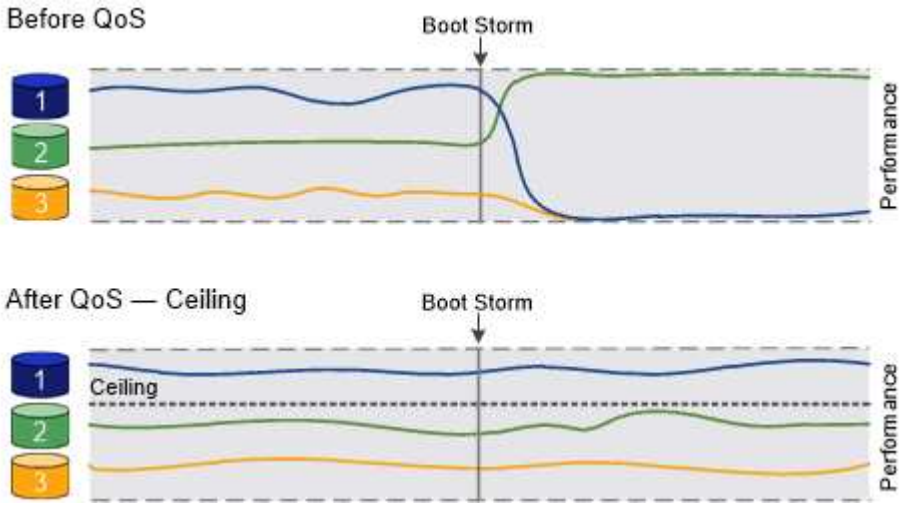
È possibile utilizzare la qualità del servizio (QoS) dello storage per garantire che le performance dei carichi di lavoro critici non vengano degradate dai carichi di lavoro concorrenti. È possibile impostare un *soffitto* di throughput su un carico di lavoro concorrente per limitarne l'impatto sulle risorse di sistema o impostare un *piano* di throughput per un carico di lavoro critico, garantendo che soddisfi gli obiettivi di throughput minimi, indipendentemente dalla domanda dei carichi di lavoro concorrenti. È anche possibile impostare un soffitto e un pavimento per lo stesso carico di lavoro.

Limiti massimi di throughput (qualità del servizio massima)

Un limite massimo di throughput limita il throughput per un carico di lavoro a un numero massimo di IOPS o Mbps o IOPS e Mbps. Nella figura riportata di seguito, il limite massimo di throughput per il carico di lavoro 2 garantisce che i carichi di lavoro 1 e 3 non siano "ingombrati".

Un *gruppo di policy* definisce il limite massimo di throughput per uno o più carichi di lavoro. Un carico di lavoro rappresenta le operazioni di i/o per un *oggetto storage:* volume, file, qtree o LUN o tutti i volumi, file, qtree o LUN di una SVM. È possibile specificare il limite massimo quando si crea il gruppo di criteri oppure attendere che i carichi di lavoro vengano monitorati per specificarlo.

Il throughput per i carichi di lavoro potrebbe superare il limite massimo specificato fino al 10%, soprattutto se un carico di lavoro subisce rapidi cambiamenti nel throughput. Il limite massimo potrebbe essere superato fino al 50% per gestire i burst. I burst si verificano su singoli nodi quando i token accumulano fino al 150%



Limiti minimi di throughput (qualità del servizio minima)

Un piano di throughput garantisce che il throughput per un carico di lavoro non scenda al di sotto di un numero minimo di IOPS o Mbps o IOPS e Mbps. Nella figura riportata di seguito, i livelli di throughput per il carico di lavoro 1 e il carico di lavoro 3 garantiscono il raggiungimento degli obiettivi di throughput minimi, indipendentemente dalla domanda per carico di lavoro 2.

Come suggeriscono gli esempi, un limite di throughput rallenta direttamente il throughput. Un piano di throughput rallenta indirettamente il throughput, dando priorità ai carichi di lavoro per i quali è stato impostato il piano.

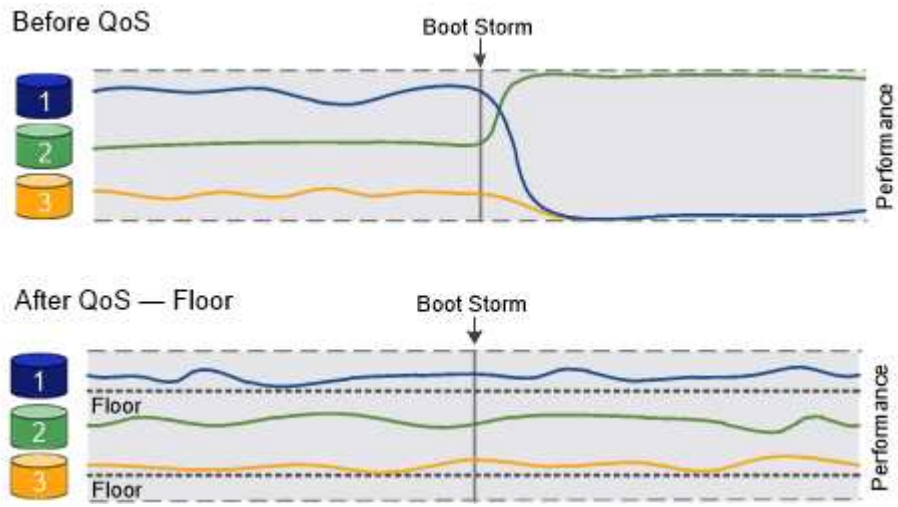
È possibile specificare il piano di lavoro quando si crea il gruppo di policy oppure attendere fino a quando non si monitorano i carichi di lavoro per specificarlo.

A partire da ONTAP 9.13.1, è possibile impostare limiti minimi di throughput nell'ambito SVM mediante [adaptive-qos-templates]. Nelle versioni di ONTAP precedenti alla 9.13.1, un gruppo di criteri che definisce un piano di throughput non può essere applicato a una SVM.

Nelle versioni precedenti a ONTAP 9.7, i piani di throughput sono garantiti quando è disponibile una capacità di performance sufficiente.

In ONTAP 9.7 e versioni successive, è possibile garantire il throughput anche quando la capacità delle performance è insufficiente. Questo nuovo comportamento si chiama Floors v2. Per soddisfare le garanzie, floors v2 può comportare una latenza maggiore sui carichi di lavoro senza un piano di throughput o sul lavoro che supera le impostazioni di base. Floors v2 si applica sia alla QoS che alla QoS adattiva.

L'opzione di attivazione/disattivazione del nuovo comportamento dei piani v2 è disponibile in ONTAP 9.7P6 e versioni successive. Un carico di lavoro potrebbe scendere al di sotto del piano specificato durante operazioni critiche come `volume move trigger-cutover`. Anche quando è disponibile una capacità sufficiente e non si svolgono operazioni critiche, il throughput di un workload potrebbe scendere al di sotto del piano specificato fino al 5%. Se il provisioning dei piani è eccessivo e non esiste una capacità di performance, alcuni carichi di lavoro potrebbero scendere al di sotto del piano specificato.



Gruppi di criteri QoS condivisi e non condivisi

A partire da ONTAP 9.4, è possibile utilizzare un gruppo di policy di qualità del servizio *non-shared* per specificare che il limite di throughput definito o il piano si applica a ogni singolo carico di lavoro membro. Il comportamento dei gruppi di policy *shared* dipende dal tipo di policy:

- Per i limiti di throughput, il throughput totale per i carichi di lavoro assegnati al gruppo di criteri condivisi non può superare il limite massimo specificato.
- Per i piani di throughput, il gruppo di policy condiviso può essere applicato solo a un singolo workload.

QoS adattiva

Normalmente, il valore del gruppo di criteri assegnato a un oggetto di storage è fisso. È necessario modificare il valore manualmente quando la dimensione dell'oggetto di storage cambia. Un aumento della quantità di spazio utilizzata su un volume, ad esempio, richiede solitamente un aumento corrispondente del limite di throughput specificato per il volume.

QoS adattiva scala automaticamente il valore del gruppo di policy in base alle dimensioni del carico di lavoro, mantenendo il rapporto tra IOPS e TB|GB in base alle dimensioni del carico di lavoro. Si tratta di un vantaggio significativo quando si gestiscono centinaia o migliaia di carichi di lavoro in un'implementazione di grandi dimensioni.

In genere, si utilizza la QoS adattiva per regolare i limiti di throughput, ma è anche possibile utilizzarla per gestire i piani di throughput (quando le dimensioni del carico di lavoro aumentano). La dimensione del carico di lavoro viene espressa come spazio allocato per l'oggetto di storage o come spazio utilizzato dall'oggetto di storage.

Lo spazio utilizzato è disponibile per i piani di throughput in ONTAP 9.5 e versioni successive. Non è supportato per i piani di throughput in ONTAP 9.4 e versioni precedenti.

- Una policy di *spazio allocato* mantiene il rapporto IOPS/TB|GB in base alle dimensioni nominali dell'oggetto di storage. Se il rapporto è di 100 IOPS/GB, un volume da 150 GB avrà un limite di throughput di 15,000 IOPS, a condizione che il volume rimanga tale. Se il volume viene ridimensionato a 300 GB, la QoS adattiva regola il limite di throughput a 30,000 IOPS.
- Una policy *used space* (predefinita) mantiene il rapporto IOPS/TB|GB in base alla quantità di dati effettivi memorizzati prima dell'efficienza dello storage. Se il rapporto è di 100 IOPS/GB, un volume da 150 GB con 100 GB di dati memorizzati avrebbe un limite massimo di throughput di 10,000 IOPS. Man mano che la quantità di spazio utilizzato cambia, la QoS adattiva regola il limite di throughput in base al rapporto.

A partire da ONTAP 9.5, è possibile specificare una dimensione del blocco i/o per l'applicazione in uso che consenta di esprimere un limite di throughput in IOPS e Mbps. Il limite Mbps viene calcolato moltiplicando le dimensioni del blocco per il limite IOPS. Ad esempio, una dimensione del blocco i/o di 32K per un limite IOPS di 6144 IOPS/TB produce un limite di Mbps di 192 MBps.

È possibile prevedere il seguente comportamento sia per i limiti di throughput che per i piani:

- Quando un carico di lavoro viene assegnato a un gruppo di policy QoS adattivi, il soffitto o il piano vengono aggiornati immediatamente.
- Quando un carico di lavoro in un gruppo di policy QoS adattiva viene ridimensionato, il soffitto o il piano viene aggiornato in circa cinque minuti.

Il throughput deve aumentare di almeno 10 IOPS prima di eseguire gli aggiornamenti.

I gruppi di policy di QoS adattivi non sono sempre condivisi: Il limite di throughput definito o il piano si applica a ciascun carico di lavoro membro singolarmente.

A partire da ONTAP 9.6, i piani di throughput sono supportati da ONTAP Select Premium con SSD.

Modello di gruppo di policy adattive

A partire da ONTAP 9.13.1, puoi impostare un modello di qualità del servizio adattivo su una SVM. I modelli di gruppi di policy adattivi consentono di impostare i livelli e i limiti di throughput per tutti i volumi in una SVM.

È possibile impostare i modelli di gruppi di criteri adattivi solo dopo la creazione di SVM. Utilizzare `vserver modify` con il `-qos-adaptive-policy-group-template` parametro per impostare il criterio.

Quando si imposta un modello di gruppo di criteri adattativi, i volumi creati o migrati dopo l'impostazione del criterio ereditano automaticamente il criterio. Gli eventuali volumi presenti nella SVM non vengono influenzati quando si assegna il modello di policy. Se si disattiva il criterio su SVM, qualsiasi volume successivamente migrato o creato su SVM non riceverà il criterio. La disattivazione del modello di gruppo di criteri adattivi non influisce sui volumi che hanno ereditato il modello di criteri, poiché conservano il modello di criteri.

Per ulteriori informazioni, vedere Impostare un modello di gruppo di criteri adattativi.

Supporto generale

La seguente tabella mostra le differenze nel supporto per i limiti di throughput, i piani di throughput e la QoS adattiva.

Risorsa o funzione	Limite di throughput	Piano di throughput	Throughput floor v2	QoS adattiva
Versione di ONTAP 9	Tutto	9.2 e versioni successive	9.7 e versioni successive	9.3 e versioni successive
Piattaforme	Tutto	• AFF • C190 [1] • ONTAP Select premium con SSD [1]	• AFF • C190 • ONTAP Select Premium con SSD	Tutto
Protocolli	Tutto	Tutto	Tutto	Tutto
FabricPool	Sì	Sì, se la policy di tiering è impostata su "nessuno" e non ci sono blocchi nel cloud.	Sì, se la policy di tiering è impostata su "nessuno" e non ci sono blocchi nel cloud.	No
SnapMirror sincrono	Sì	No	No	Sì

[1] C190 e il supporto ONTAP Select sono iniziati con la versione ONTAP 9,6.

Carichi di lavoro supportati per i limiti di throughput

La tabella seguente mostra il supporto dei workload per i limiti di throughput per la versione di ONTAP 9. I volumi root, i mirror di condivisione del carico e i mirror di protezione dei dati non sono supportati.

Supporto dei carichi di lavoro	ONTAP 9.8 e versioni successive	ONTAP da 9,7 a 9,4	ONTAP 9.3 e versioni precedenti
Volume	sì	sì	sì
File	sì	sì	sì
LUN	sì	sì	sì
SVM	sì	sì	sì
Volume FlexGroup	sì	sì	sì (solo ONTAP 9,3)
qtree [1]	sì	no	no

Supporto dei carichi di lavoro	ONTAP 9.8 e versioni successive	ONTAP da 9,7 a 9,4	ONTAP 9.3 e versioni precedenti
Carichi di lavoro multipli per gruppo di policy	sì	sì	sì
Gruppi di criteri non condivisi	sì	sì	no

[1] A partire da ONTAP 9.9.1, l'accesso SMB è supportato anche nei qtree nei volumi FlexVol e FlexGroup con SMB abilitato. A partire da ONTAP 9.8, l'accesso NFS è supportato nei qtree nei volumi FlexVol e FlexGroup con NFS abilitato.

Carichi di lavoro supportati per i piani di throughput

La seguente tabella mostra il supporto dei workload per i piani di throughput in base alla versione di ONTAP 9. I volumi root, i mirror di condivisione del carico e i mirror di protezione dei dati non sono supportati.

Supporto dei carichi di lavoro	ONTAP 9.13.1 e versioni successive	ONTAP da 9,8 a 9.13.0	ONTAP da 9,4 a 9,7	ONTAP 9.3
Volume	sì	sì	sì	sì
File	sì	sì	sì	sì
LUN	sì	sì	sì	sì
SVM	sì	no	no	no
Volume FlexGroup	sì	sì	sì	no
qtree [1]	sì	sì	no	no
Carichi di lavoro multipli per gruppo di policy	sì	sì	sì	no
Gruppi di criteri non condivisi	sì	sì	sì	no

[1] a partire da ONTAP 9,8, l'accesso NFS è supportato in qtree in volumi FlexVol e FlexGroup con NFS abilitato. A partire da ONTAP 9.9.1, l'accesso SMB è supportato anche nei qtree dei volumi FlexVol e FlexGroup con SMB attivato.

Carichi di lavoro supportati per QoS adattiva

La seguente tabella mostra il supporto dei carichi di lavoro per la QoS adattiva in base alla versione di ONTAP 9. I volumi root, i mirror di condivisione del carico e i mirror di protezione dei dati non sono supportati.

Supporto dei carichi di lavoro	ONTAP 9.13.1 e versioni successive	ONTAP da 9,4 a 9.13.0	ONTAP 9.3
Volume	sì	sì	sì
File	sì	sì	no

Supporto dei carichi di lavoro	ONTAP 9.13.1 e versioni successive	ONTAP da 9,4 a 9.13.0	ONTAP 9.3
LUN	sì	sì	no
SVM	sì	no	no
Volume FlexGroup	sì	sì	no
Carichi di lavoro multipli per gruppo di policy	sì	sì	sì
Gruppi di criteri non condivisi	sì	sì	sì

Numero massimo di workload e gruppi di policy

La seguente tabella mostra il numero massimo di workload e gruppi di policy per versione di ONTAP 9.

Supporto dei carichi di lavoro	ONTAP 9.4 e versioni successive	ONTAP 9.3 e versioni precedenti
Carichi di lavoro massimi per cluster	40,000	12,000
Carichi di lavoro massimi per nodo	40,000	12,000
Numero massimo di gruppi di criteri	12,000	12,000

Abilita o disabilita i piani di throughput ONTAP v2

È possibile attivare o disattivare il throughput floors v2 su AFF. L'impostazione predefinita è Enabled (attivato). Con FLOors v2 abilitato, è possibile soddisfare i piani di throughput quando i controller vengono utilizzati in modo pesante a scapito di una maggiore latenza su altri carichi di lavoro. Floors v2 si applica sia a QoS che a QoS adattivo.

Fasi

1. Passare al livello di privilegio avanzato:

   ```
   set -privilege advanced
   ```

2. Immettere uno dei seguenti comandi:

Se si desidera…	Utilizzare questo comando:
Disattiva piani v2	`qos settings throughput-floors-v2 -enable false`
Abilitare i piani v2	`qos settings throughput-floors-v2 -enable true`

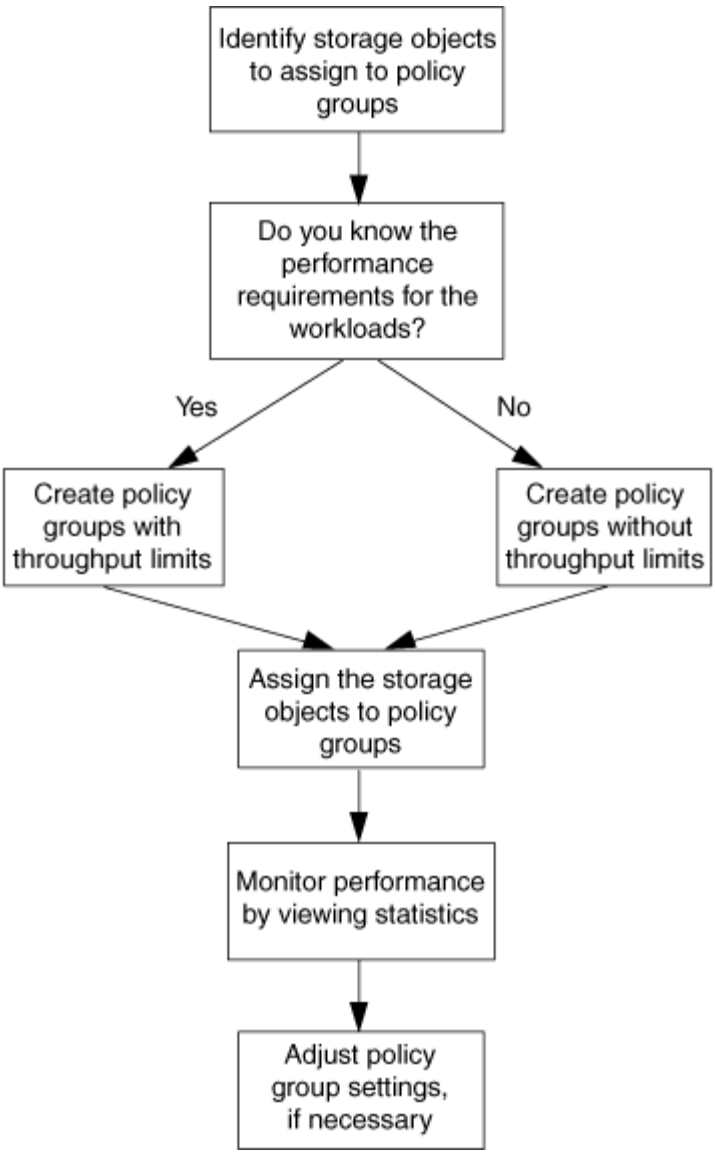
Per disattivare il throughput floors v2 in un cluster MetroCluster, è necessario eseguire `qos settings throughput-floors-v2 -enable false` comando sui cluster di origine e di destinazione.

```
cluster1::*> qos settings throughput-floors-v2 -enable false
```

Ulteriori informazioni su `qos settings throughput-floors-v2` nella "Riferimento al comando ONTAP".

Flusso di lavoro QoS di archiviazione ONTAP

Se si conoscono già i requisiti di performance per i carichi di lavoro che si desidera gestire con QoS, è possibile specificare il limite di throughput quando si crea il gruppo di policy. In caso contrario, è possibile attendere fino a quando non si monitorano i carichi di lavoro per specificare il limite.



Imposta un limite massimo di throughput con la qualità del servizio in ONTAP

È possibile utilizzare `max-throughput` Campo per un gruppo di criteri per definire un limite massimo di throughput per i carichi di lavoro degli oggetti di storage (QoS Max). È possibile applicare il gruppo di criteri quando si crea o si modifica l'oggetto di storage.

Prima di iniziare

- Per creare un gruppo di criteri, è necessario essere un amministratore del cluster.
- Per applicare un gruppo di criteri a una SVM, è necessario essere un amministratore del cluster.

A proposito di questa attività

- A partire da ONTAP 9.4, è possibile utilizzare un gruppo di policy di qualità del servizio *non-shared* per specificare che il limite di throughput definito si applica a ogni singolo carico di lavoro membro. In caso contrario, il gruppo di criteri è *shared:* il throughput totale per i carichi di lavoro assegnati al gruppo di criteri non può superare il limite massimo specificato.

  Impostare `-is-shared=false` per `qos policy-group create` per specificare un gruppo di criteri non condiviso.

- È possibile specificare il limite di throughput per il limite massimo in IOPS, MB/s o IOPS, MB/s. Se si specificano IOPS e MB/s, viene applicato il limite raggiunto per primo.

> Se si impostano un soffitto e un pavimento per lo stesso carico di lavoro, è possibile specificare il limite di throughput per il soffitto solo in IOPS.

- Un oggetto storage soggetto a un limite di QoS deve essere contenuto dalla SVM a cui appartiene il gruppo di criteri. Più gruppi di criteri possono appartenere alla stessa SVM.
- Non è possibile assegnare un oggetto di storage a un gruppo di criteri se l'oggetto contenente o i relativi oggetti figlio appartengono al gruppo di criteri.
- È consigliabile applicare un gruppo di criteri allo stesso tipo di oggetti di storage.

Fasi

1. Creare un gruppo di criteri:

   `qos policy-group create -policy-group` *`policy_group`* `-vserver` *`SVM`* `-max -throughput` *`number_of_iops`*`|Mb/S|iops,Mb/S -is-shared true|false`

   Ulteriori informazioni su `qos policy-group create` nella "Riferimento al comando ONTAP".

   È possibile utilizzare il `qos policy-group modify` comando per regolare i limiti massimi di throughput.

   Il comando seguente crea il gruppo di criteri condivisi `pg-vs1` Con un throughput massimo di 5,000 IOPS:

   ```
   cluster1::> qos policy-group create -policy-group pg-vs1 -vserver vs1
   -max-throughput 5000iops -is-shared true
   ```

   Il comando seguente crea il gruppo di criteri non condivisi `pg-vs3` Con un throughput massimo di 100 IOPS e 400 Kb/S:

   ```
   cluster1::> qos policy-group create -policy-group pg-vs3 -vserver vs3
   -max-throughput 100iops,400KB/s -is-shared false
   ```

   Il comando seguente crea il gruppo di criteri non condivisi `pg-vs4` senza un limite di throughput:

   ```
   cluster1::> qos policy-group create -policy-group pg-vs4 -vserver vs4
   -is-shared false
   ```

   Ulteriori informazioni su `qos policy-group modify` nella "Riferimento al comando ONTAP".

2. Applicare un gruppo di criteri a una SVM, a un file, a un volume o a un LUN:

   `storage_object` `create -vserver` *`SVM`* `-qos-policy-group` *`policy_group`*

   Per ulteriori informazioni sui comandi descritti in questa procedura, consultare la "Riferimento al comando ONTAP". È possibile utilizzare il *`storage_object`* `modify` comando per applicare un gruppo di criteri

diverso all'oggetto di archiviazione.

Il seguente comando applica il gruppo di criteri `pg-vs1` A SVM `vs1`:

```
cluster1::> vserver create -vserver vs1 -qos-policy-group pg-vs1
```

I seguenti comandi applicano il gruppo di criteri `pg-app` ai volumi `app1` e. `app2`:

```
cluster1::> volume create -vserver vs2 -volume app1 -aggregate aggr1
-qos-policy-group pg-app
```

```
cluster1::> volume create -vserver vs2 -volume app2 -aggregate aggr1
-qos-policy-group pg-app
```

3. Monitorare le performance dei gruppi di policy:

   `qos statistics performance show`

   Ulteriori informazioni su `qos statistics performance show` nella "Riferimento al comando ONTAP".

   > Monitorare le performance dal cluster. Non utilizzare uno strumento sull'host per monitorare le prestazioni.

   Il seguente comando mostra le prestazioni del gruppo di criteri:

```
cluster1::> qos statistics performance show
Policy Group           IOPS      Throughput   Latency
-------------------- -------- --------------- ----------
-total-                 12316       47.76MB/s  1264.00us
pg_vs1                   5008       19.56MB/s     2.45ms
_System-Best-Effort        62       13.36KB/s     4.13ms
_System-Background         30           0KB/s        0ms
```

4. Monitorare le performance dei carichi di lavoro:

   `qos statistics workload performance show`

   > Monitorare le performance dal cluster. Non utilizzare uno strumento sull'host per monitorare le prestazioni.

   Il seguente comando mostra le performance del carico di lavoro:

```
cluster1::> qos statistics workload performance show
Workload           ID     IOPS      Throughput    Latency
--------------- ------ -------- ---------------- ----------
-total-             -    12320        47.84MB/s  1215.00us
app1-wid7967     7967     7219        28.20MB/s   319.00us
vs1-wid12279    12279     5026        19.63MB/s     2.52ms
_USERSPACE_APPS    14       55        10.92KB/s   236.00us
_Scan_Backgro..  5688       20            0KB/s        0ms
```

Ulteriori informazioni su `qos statistics workload performance show` nella "Riferimento al comando ONTAP".

Puoi utilizzare `qos statistics workload latency show` il comando per visualizzare statistiche dettagliate sulla latenza per i workload di qualità del servizio. Ulteriori informazioni su `qos statistics workload latency show` nella "Riferimento al comando ONTAP".

Impostazione di un limite minimo di throughput con la qualità del servizio in ONTAP

È possibile utilizzare `min-throughput` Campo per un gruppo di policy per definire un piano di throughput per i carichi di lavoro degli oggetti storage (QoS min). È possibile applicare il gruppo di criteri quando si crea o si modifica l'oggetto di storage. A partire da ONTAP 9.8, è possibile specificare il volume di throughput in IOPS o Mbps o IOPS e Mbps.

Prima di iniziare

- Per creare un gruppo di criteri, è necessario essere un amministratore del cluster.
- A partire dall'ONTAP 9.13.1, è possibile imporre limiti minimi di throughput a livello di SVM utilizzando un modello di gruppo di policy adattive. Non è possibile impostare un modello di gruppo di criteri adattativi su una SVM con un gruppo di criteri QoS.

A proposito di questa attività

- A partire da ONTAP 9.4, è possibile utilizzare un gruppo di policy di qualità del servizio *non-shared* per specificare che il piano di throughput definito deve essere applicato a ogni singolo carico di lavoro membro. Questa è l'unica condizione in cui un gruppo di policy per un piano di throughput può essere applicato a più carichi di lavoro.

  Impostare `-is-shared=false` per `qos policy-group create` per specificare un gruppo di criteri non condiviso.

- Il throughput di un carico di lavoro potrebbe scendere al di sotto del piano specificato se la capacità delle performance (spazio di crescita) del nodo o dell'aggregato è insufficiente.
- Un oggetto storage soggetto a un limite di QoS deve essere contenuto dalla SVM a cui appartiene il gruppo di criteri. Più gruppi di criteri possono appartenere alla stessa SVM.
- È consigliabile applicare un gruppo di criteri allo stesso tipo di oggetti di storage.
- Un gruppo di criteri che definisce un piano di throughput non può essere applicato a una SVM.

Fasi

1. Controllare che le prestazioni sul nodo o sull'aggregato siano adeguate, come descritto nella "Identificazione della capacità di prestazioni rimanente".
2. Creare un gruppo di criteri:

   `qos policy-group create -policy group` *`policy_group`* `-vserver` *`SVM`* `-min -throughput` *`qos_target`* `-is-shared true|false`

   Ulteriori informazioni su `qos policy-group create` nella "Riferimento al comando ONTAP".

3. Il comando consente `qos policy-group modify` di regolare i limiti minimi di throughput.

   Il seguente comando crea il gruppo di policy condivise `pg-vs2` con una velocità minima di 1.000 IOPS:

   ```
   cluster1::> qos policy-group create -policy group pg-vs2 -vserver vs2
   -min-throughput 1000iops -is-shared true
   ```

   Il comando seguente crea il gruppo di criteri non condivisi `pg-vs4` senza un limite di throughput:

   ```
   cluster1::> qos policy-group create -policy group pg-vs4 -vserver vs4
   -is-shared false
   ```

   Ulteriori informazioni su `qos policy-group modify` nella "Riferimento al comando ONTAP".

4. Applicare un gruppo di criteri a un volume o a un LUN:

   `storage_object` `create -vserver` *`SVM`* `-qos-policy-group` *`policy_group`*`Puoi usare il` `_storage_object_modify` comando per applicare un gruppo di policy diverso all'oggetto di archiviazione.

   Il seguente comando applica il gruppo di criteri `pg-app2` al volume `app2`:

   ```
   cluster1::> volume create -vserver vs2 -volume app2 -aggregate aggr1
   -qos-policy-group pg-app2
   ```

   Per ulteriori informazioni sui comandi descritti in questa procedura, consultare la "Riferimento al comando ONTAP".

5. Monitorare le performance dei gruppi di policy:

   `qos statistics performance show`

   > Monitorare le performance dal cluster. Non utilizzare uno strumento sull'host per monitorare le prestazioni.

   Il seguente comando mostra le prestazioni del gruppo di criteri:

```
cluster1::> qos statistics performance show
Policy Group           IOPS      Throughput   Latency
-------------------- -------- --------------- ----------
-total-                 12316       47.76MB/s  1264.00us
pg_app2                  7216       28.19MB/s   420.00us
_System-Best-Effort        62       13.36KB/s     4.13ms
_System-Background         30           0KB/s        0ms
```

Ulteriori informazioni su `qos statistics performance show` nella "Riferimento al comando ONTAP".

6. Monitorare le performance dei carichi di lavoro:

`qos statistics workload performance show`

> Monitorare le performance dal cluster. Non utilizzare uno strumento sull'host per monitorare le prestazioni.

Il seguente comando mostra le performance del carico di lavoro:

```
cluster1::> qos statistics workload performance show
Workload          ID     IOPS      Throughput    Latency
--------------- ------ -------- ---------------- ----------
-total-              -    12320        47.84MB/s  1215.00us
app2-wid7967      7967     7219        28.20MB/s   319.00us
vs1-wid12279     12279     5026        19.63MB/s     2.52ms
_USERSPACE_APPS     14       55        10.92KB/s   236.00us
_Scan_Backgro..   5688       20            0KB/s        0ms
```

Ulteriori informazioni su `qos statistics workload performance show` nella "Riferimento al comando ONTAP".

> Puoi utilizzare `qos statistics workload latency show` il comando per visualizzare statistiche dettagliate sulla latenza per i workload di qualità del servizio. Ulteriori informazioni su `qos statistics workload latency show` nella "Riferimento al comando ONTAP".

Utilizza i gruppi di criteri QoS adattivi in ONTAP

È possibile utilizzare un gruppo di policy *QoS adattivo* per adattare automaticamente un limite massimo o minimo di throughput alle dimensioni del volume, mantenendo il rapporto tra IOPS e TB/GB al variare delle dimensioni del volume. Si tratta di un vantaggio significativo quando si gestiscono centinaia o migliaia di carichi di lavoro in una distribuzione di grandi dimensioni.

Prima di iniziare

- È necessario eseguire ONTAP 9.3 o versione successiva. I gruppi di policy QoS adattivi sono disponibili a

partire da ONTAP 9.3.

- Per creare un gruppo di criteri, è necessario essere un amministratore del cluster.

A proposito di questa attività

Un oggetto storage può essere membro di un gruppo di criteri adattivi o non adattivi, ma non di entrambi. La SVM dell'oggetto di storage e il criterio devono essere identici. L'oggetto di storage deve essere in linea.

I gruppi di policy di QoS adattivi non sono sempre condivisi: Il limite di throughput definito o il piano si applica a ciascun carico di lavoro membro singolarmente.

Il rapporto tra i limiti di throughput e le dimensioni degli oggetti di storage è determinato dall'interazione dei seguenti campi:

- `expected-iops`è il minimo IOPS previsto per TB/GB allocato.

> `expected-iops` È garantita solo sulle piattaforme AFF. `expected-iops` È garantita per FabricPool solo se la policy di tiering è impostata su "nessuno" e nessun blocco è nel cloud. `expected-iops` È garantito per i volumi che non sono in una relazione sincrona SnapMirror.

- `peak-iops`è il massimo IOPS possibile per TB/GB allocato o utilizzato.
- `expected-iops-allocation` specifica se per gli iops previsti viene utilizzato lo spazio allocato (impostazione predefinita) o lo spazio utilizzato.

> `expected-iops-allocation` È disponibile in ONTAP 9.5 e versioni successive. Non è supportato in ONTAP 9.4 e versioni precedenti.

- `peak-iops-allocation` specifica se viene utilizzato lo spazio allocato o lo spazio utilizzato (impostazione predefinita) per `peak-iops`.
- `absolute-min-iops` È il numero minimo assoluto di IOPS. È possibile utilizzare questo campo con oggetti di storage molto piccoli. Sovrascrive entrambi `peak-iops` e/o. `expected-iops` quando `absolute-min-iops` è maggiore del valore calcolato `expected-iops`.

  Ad esempio, se si imposta `expected-iops` Fino a 1,000 IOPS/TB e le dimensioni del volume sono inferiori a 1 GB, il valore calcolato `expected-iops` Sarà un IOP frazionale. Il valore calcolato `peak-iops` sarà una frazione ancora più piccola. Per evitare questo problema, impostare `absolute-min-iops` a un valore realistico.

- `block-size` Specifica la dimensione del blocco i/o dell'applicazione. L'impostazione predefinita è 32K. I valori validi sono 8K, 16K, 32K, 64K, QUALSIASI. QUALSIASI indica che la dimensione del blocco non viene applicata.

Gruppi di policy QoS adattive predefinite

Sono disponibili tre gruppi di criteri QoS adattivi predefiniti, come mostrato nella tabella seguente. È possibile applicare questi gruppi di criteri direttamente a un volume.

Gruppo di criteri predefinito	IOPS/TB previsti	IOPS/TB di picco	IOPS minimo assoluto
`extreme`	6,144	12,288	1000
`performance`	2,048	4,096	500
`value`	128	512	75

Restrizioni di assegnazione del gruppo di criteri degli oggetti di archiviazione

In alcuni casi non è possibile assegnare un oggetto di archiviazione a un gruppo di policy se l'oggetto che lo contiene o i suoi oggetti figlio appartengono a un gruppo di policy.

A partire da ONTAP 9.18.1, è possibile utilizzare criteri QoS nidificati, che consentono di assegnare gruppi di criteri sia all'oggetto contenitore, ad esempio una SVM, sia al suo oggetto figlio, ad esempio un volume. In un ambiente multi-tenant, i criteri QoS annidati consentono agli amministratori di suddividere i limiti QoS sulle SVM tra i volumi e i qtree all'interno della SVM e di fornire la possibilità di bilanciare le risorse di archiviazione tra gli ambienti di elaborazione, consentendo al contempo la definizione delle priorità dei carichi di lavoro critici per la missione.

I criteri QoS annidati sono supportati per le seguenti coppie di oggetti:

- SVM e volumi FlexVol o FlexGroup contenuti nell'SVM.
- Volumi FlexVol o FlexGroup e qtree all'interno dei volumi.

Per le policy QoS nidificate, viene utilizzata la policy applicabile più restrittiva.

Nella tabella seguente sono elencate le restrizioni.

Se si assegna…	**Quindi non è possibile assegnare a un gruppo di policy…**
SVM a un gruppo di criteri	Tutti gli oggetti di archiviazione contenuti nell'SVM. Se si esegue ONTAP 9.18.1, i volumi FlexVol e FlexGroup contenuti nelle SVM *possono* essere assegnati a un gruppo di policy.
Su un gruppo di criteri	L'SVM contenente il volume o eventuali LUN figlio. Se si esegue ONTAP 9.18.1 e versioni successive, l'SVM contenente il volume *può* essere assegnato a un gruppo di policy. Inoltre, è possibile assegnare qtree nei volumi FlexVol o FlexGroup .
LUN a un gruppo di criteri	Il volume o SVM contenente i LUN

Se si assegna…	Quindi non è possibile assegnare a un gruppo di policy…
Su un gruppo di criteri	Il volume o SVM contenente il file

Fasi

1. Creare un gruppo di criteri QoS adattivi:

   `qos adaptive-policy-group create -policy group` *`policy_group`* `-vserver` *`SVM`* `-expected-iops` *`number_of_iops`*`/TB|GB -peak-iops` *`number_of_iops`*`/TB|GB -expected-iops-allocation-space|used-space -peak-iops-allocation allocated-space|used-space -absolute-min-iops` *`number_of_iops`* `-block-size 8K|16K|32K|64K|ANY`

   Ulteriori informazioni su `qos adaptive-policy-group create` nella "Riferimento al comando ONTAP".

   > `-expected-iops-allocation` e. `-block-size` È disponibile in ONTAP 9.5 e versioni successive. Queste opzioni non sono supportate in ONTAP 9.4 e versioni precedenti.

   Il seguente comando crea un gruppo di criteri QoS adattivi `adpg-app1` con `-expected-iops` Impostato su 300 IOPS/TB, `-peak-iops` Impostato su 1,000 IOPS/TB, `-peak-iops-allocation` impostare su `used-space`, e. `-absolute-min-iops` Impostato su 50 IOPS:

   ```
   cluster1::> qos adaptive-policy-group create -policy group adpg-app1
   -vserver vs2 -expected-iops 300iops/tb -peak-iops 1000iops/TB -peak-iops
   -allocation used-space -absolute-min-iops 50iops
   ```

2. Applicare un gruppo di criteri QoS adattivi a un volume:

   `volume create -vserver SVM -volume` *`volume`* `-aggregate` *`aggregate`* `-size` *`number_of`* `TB|GB -qos-adaptive-policy-group` *`policy_group`*

   Ulteriori informazioni su `volume create` nella "Riferimento al comando ONTAP".

   Il seguente comando applica il gruppo di criteri QoS adattivi `adpg-app1` al volume `app1`:

   ```
   cluster1::> volume create -vserver vs1 -volume app1 -aggregate aggr1
   -size 2TB -qos-adaptive-policy-group adpg-app1
   ```

   I seguenti comandi applicano il gruppo di criteri QoS adattivi predefinito `extreme` al nuovo volume `app4` e al volume esistente `app5`. Il limite di throughput definito per il gruppo di criteri si applica ai volumi `app4` e. `app5` singolarmente:

   ```
   cluster1::> volume create -vserver vs4 -volume app4 -aggregate aggr4
   -size 2TB -qos-adaptive-policy-group extreme
   ```

```
cluster1::> volume modify -vserver vs5 -volume app5 -qos-adaptive-policy
-group extreme
```

Impostare un modello di gruppo di criteri adattivo in ONTAP

A partire dall'ONTAP 9.13.1, è possibile imporre limiti e limiti di throughput a livello di SVM utilizzando un modello di gruppo di policy adattivo.

A proposito di questa attività

- Il modello di gruppo di criteri adattivi è un criterio predefinito `apg1`. Il criterio può essere modificato in qualsiasi momento. Può essere impostato solo con l'API REST CLI o ONTAP e può essere applicato solo alle SVM esistenti.
- Il modello di gruppo di policy adattive influisce solo sui volumi creati o migrati sulla SVM dopo aver impostato il criterio. I volumi esistenti sulla SVM mantengono lo stato esistente.

  Se si disattiva il modello di gruppo di criteri adattivi, i volumi su SVM conservano i criteri esistenti. Solo i volumi successivamente creati o migrati sulla SVM saranno influenzati dalla disabilitazione.

- Non è possibile impostare un modello di gruppo di criteri adattativi su una SVM con un gruppo di criteri QoS.
- I modelli di gruppi di policy adattivi sono progettati per le piattaforme AFF. È possibile impostare un modello di gruppo di policy adattivo su altre piattaforme, ma il criterio potrebbe non applicare un throughput minimo. In modo simile, puoi aggiungere un modello di gruppo di policy adattivo a una SVM di un aggregato FabricPool o di un aggregato senza supportare un throughput minimo, ma non verrà applicato il limite minimo di throughput.
- Se la SVM si trova in una configurazione MetroCluster o in una relazione SnapMirror, il modello di gruppo di criteri adattativi verrà applicato alla SVM mirrorata.

Fasi

1. Modificare la SVM per applicare il modello di gruppo di criteri adattativi: `vserver modify -qos -adaptive-policy-group-template apg1`
2. Verificare che il criterio sia stato impostato: `vserver show -fields qos-adaptive-policy-group`

Informazioni sul copyright

Copyright © 2026 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 COMMERCIABILITÀ 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 CONSEQUENZIALE (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.