워크로드 관리
ONTAP 9

NetApp
February 12, 2026

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

목차

워크로드 관리 . . . 1
- ONTAP에 남아 있는 성능 용량을 식별합니다 . . . 1
- ONTAP에서 트래픽이 높은 클라이언트 또는 파일을 식별합니다 . . . 2
- QoS로 처리량 보장 . . . 4
  - ONTAP에서 QoS 개요를 통해 처리량을 보장합니다 . . . 4
  - ONTAP 처리량 바닥 v2 활성화 또는 비활성화 . . . 9
  - ONTAP 스토리지 QoS 워크플로 . . . 10
  - ONTAP에서 QoS를 사용하여 처리량의 한계를 설정합니다 . . . 11
  - ONTAP에서 QoS로 처리량 기준을 설정합니다 . . . 14
  - ONTAP에서 적응형 QoS 정책 그룹 사용 . . . 16
  - ONTAP에서 적응형 정책 그룹 템플릿을 설정합니다 . . . 19

워크로드 관리

ONTAP에 남아 있는 성능 용량을 식별합니다

성능 용량, 즉_FLO여유_는 리소스의 워크로드 성능이 지연 시간의 영향을 받기 전에 노드나 애그리게이트에 배치할 수 있는 작업의 양을 측정합니다. 클러스터에서 사용 가능한 성능 용량을 파악하면 워크로드를 프로비저닝하고 조정할 수 있습니다.

시작하기 전에

이 작업에는 고급 권한 레벨 명령이 필요합니다.

이 작업에 대해

'-object' 옵션에 다음 값을 사용하여 여유 공간 통계를 수집하고 표시할 수 있습니다.

- CPU의 경우 resource_refLO여유_cpu'입니다.
- Aggregate의 경우 RESOURCE_LOBLO여유_aggr입니다.

System Manager 및 Active IQ Unified Manager를 사용하여 이 작업을 완료할 수도 있습니다.

단계

1. 고급 권한 레벨로 변경:

   세트 프리빌리지 고급

2. 실시간 여유 공간 통계 수집 시작:

   '통계 시작 – object resource_refo여유_cpu|aggr'

   에 대한 자세한 내용은 `statistics start` "ONTAP 명령 참조입니다"을 참조하십시오.

3. 실시간 여유 공간 통계 정보 표시:

   'tortistics show-object resource_fre여유_cpu|aggr'

4. 관리 권한으로 돌아가기:

   'Set-Privilege admin'입니다

예

다음 예제에는 클러스터 노드의 평균 사용 시간 여유 공간 통계가 표시됩니다.

Optimal_point_Utilization 카운터에서 Current_Utilization 카운터를 빼서 노드에 대해 사용 가능한 성능 용량을 계산할 수 있습니다. 이 예에서는 CPU_sti2520-213의 사용률 용량이 -14%(72%-86%)로, 이는 CPU가 지난 1시간 동안 평균 초과 사용되었음을 나타냅니다.

같은 정보를 더 오랜 기간 평균한 값으로 얻기 위해 'ewma_daily', 'ewma_weekly', 'ewma_monthly'를 지정할 수도 있습니다.

```
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.
```

에 대한 자세한 내용은 `statistics show` "ONTAP 명령 참조입니다"을 참조하십시오.

ONTAP에서 트래픽이 높은 클라이언트 또는 파일을 식별합니다

ONTAP 활성 개체 기술을 사용하여 불균형적으로 많은 양의 클러스터 트래픽을 담당하는

클라이언트 또는 파일을 식별할 수 있습니다. 이러한 "상위" 클라이언트 또는 파일을 식별한 후에는 클러스터 워크로드를 재조정하거나 다른 단계를 수행하여 문제를 해결할 수 있습니다.

시작하기 전에

이 작업을 수행하려면 클러스터 관리자여야 합니다.

단계

1. 클러스터에 액세스하는 상위 클라이언트 보기:

   'tortistics top client show -node_name_-sort-key_sort_column_-interval_seconds_between_updates_ -iterations_-max_number_of_instances_'

   다음 명령을 실행하면 'cluster1'에 액세스하는 상위 클라이언트가 표시됩니다.

```
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
```

   에 대한 자세한 내용은 `statistics top client show` "ONTAP 명령 참조입니다"을 참조하십시오.

2. 클러스터에서 액세스하는 상위 파일을 봅니다.

   'tortistics top file show -node_node_name_-sort-key_sort_column_-interval_seconds_between_updates_ -iterations_-max_number_of_instances_'

   다음 명령을 실행하면 'cluster1'에서 액세스되는 최상위 파일이 표시됩니다.

```
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
```

에 대한 자세한 내용은 `statistics top file show` "ONTAP 명령 참조입니다"을 참조하십시오.

QoS로 처리량 보장

ONTAP에서 QoS 개요를 통해 처리량을 보장합니다

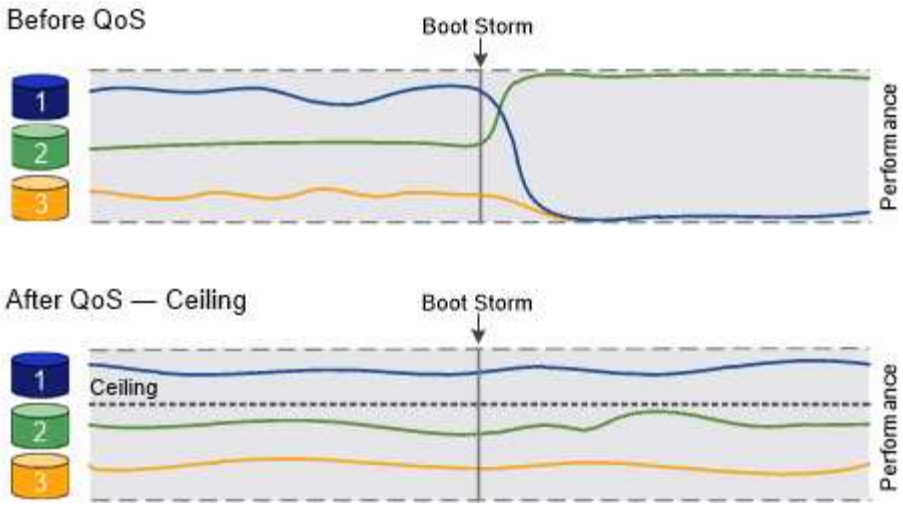
스토리지 QoS(서비스 품질)를 사용하여 주요 워크로드의 성능이 다른 워크로드에 의해 저하되지 않도록 보장할 수 있습니다. 경쟁 워크로드에 대한 throughput_ceiling_을 설정하여 시스템 리소스에 미치는 영향을 제한하거나 중요한 워크로드에 대한 throughput_floor_를 설정하여 경쟁 워크로드의 요구에 관계없이 최소 처리량 목표를 달성할 수 있습니다. 동일한 워크로드에 대해 천장과 바닥을 설정할 수도 있습니다.

처리량 한도(QoS 최대)

처리량 상한은 워크로드의 처리량을 최대 IOPS 또는 MBps, 즉 IOPS 및 MBps로 제한합니다. 아래 그림에서는 워크로드 2의 처리량 상한을 통해 워크로드 1과 3이 "괴롭지" 않도록 합니다.

policy group_은 하나 이상의 워크로드에 대한 처리량 한도를 정의합니다. 워크로드는 _ 스토리지 객체에 대한 입출력 작업을 나타냅니다. _ 볼륨, 파일, qtree 또는 LUN 또는 SVM의 모든 볼륨, 파일, Qtree 또는 LUN입니다. 정책 그룹을 생성할 때 상한을 지정하거나 워크로드를 모니터링하여 지정할 때까지 기다릴 수 있습니다.

> 워크로드에 대한 처리량이 지정된 상한을 최대 10% 초과할 수 있습니다. 특히, 작업 부하에 대한 처리량이 급격하게 변경될 경우 더욱 그렇습니다. 천장은 버스트를 처리하기 위해 최대 50%까지 초과될 수 있습니다. 토큰이 최대 150%까지 누적되면 단일 노드에서 버스트가 발생합니다



처리량 최저액**(QoS Min)**

처리량 한도 에서는 워크로드의 처리량이 최소 IOPS 또는 MBps, 즉 IOPS 및 MBps 미만으로 떨어지지 않도록 보장합니다. 아래 그림에서 워크로드 1과 워크로드 3의 처리량 플로어는 워크로드 2의 수요에 관계없이 최소 처리량 목표를 충족할 수 있도록 합니다.

예를 들어, 처리량 상한은 처리량을 직접 조절합니다. 처리량 플로어에서는 플로어가 설정된 워크로드에 우선 순위를 부여하여 간접적으로 처리량을 조절합니다.

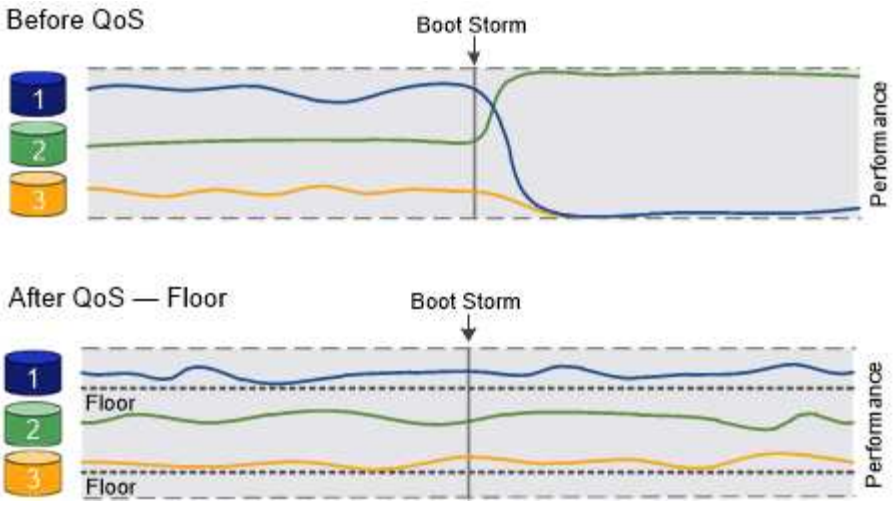
정책 그룹을 생성할 때 층을 지정하거나 워크로드를 모니터링하여 지정할 때까지 기다릴 수 있습니다.

ONTAP 9.13.1부터 를 사용하여 SVM 범위에서 처리량 플로트를 설정할 수 있습니다[adaptive-qos-templates]. 9.13.1 이전의 ONTAP 릴리즈에서는 처리량 플로어를 정의하는 정책 그룹을 SVM에 적용할 수 없습니다.

ONTAP 9.7 이전의 릴리즈에서는 사용 가능한 성능 용량이 충분한 경우 처리량 바닥이 보장됩니다.

ONTAP 9.7 이상에서는 사용 가능한 성능 용량이 부족하더라도 처리량 플로어를 보장할 수 있습니다. 이러한 새로운 바닥 동작을 바닥 v2라고 합니다. 보장 사항을 충족하기 위해 v2층은 처리량 공간 또는 바닥 설정을 초과하는 작업 환경에서 작업 부하에 대한 대기 시간이 더 길어질 수 있습니다. 버전 v2는 QoS 및 적응형 QoS에 모두 적용됩니다.

ONTAP 9.7P6 이상에서 새로운 바닥 v2의 동작을 활성화/비활성화하는 옵션을 사용할 수 있습니다. 과 같은 중요한 작업 중에는 워크로드가 지정된 설치 공간 아래로 떨어질 수 있습니다 `volume move trigger-cutover`. 충분한 용량을 사용할 수 있고 중요한 작업이 마련되어 있지 않더라도 워크로드의 처리량은 지정된 바닥 미만으로 최대 5% 떨어질 수 있습니다. 바닥이 초과 프로비저닝되고 성능 용량이 없는 경우 일부 워크로드가 지정된 설치 공간 아래로 떨어질 수 있습니다.



공유 및 비공유 **QoS** 정책 그룹

ONTAP 9.4부터 _non-shared_QoS 정책 그룹을 사용하여 정의된 처리량 상한 또는 최저가 각 구성원 워크로드에 개별적으로 적용되도록 지정할 수 있습니다. shared_policy 그룹의 동작은 정책 유형에 따라 달라집니다.

- 처리량 천장의 경우 공유 정책 그룹에 할당된 워크로드의 총 처리량은 지정된 한도를 초과할 수 없습니다.
- 처리량 플로어의 경우 공유 정책 그룹을 단일 워크로드에만 적용할 수 있습니다.

적응형 **QoS**

일반적으로 스토리지 객체에 할당한 정책 그룹의 값은 고정됩니다. 스토리지 오브젝트의 크기가 변경되면 값을 수동으로 변경해야 합니다. 예를 들어, 볼륨에 사용된 공간의 양을 늘리려면 일반적으로 볼륨에 지정된 처리량 상한을 늘려야 합니다.

_Adaptive QoS_는 워크로드 크기에 따라 정책 그룹 값을 자동으로 확장하며 워크로드 크기 변화에 따라 IOPS와 TB|GB의 비율을 유지합니다. 이는 대규모 구축 환경에서 수백 또는 수천 개의 워크로드를 관리할 때 큰 이점입니다.

일반적으로 적응형 QoS를 사용하여 처리량 한도를 조정할 수 있지만 워크로드 크기가 증가하는 경우 이를 사용하여 처리량 플로어를 관리할 수도 있습니다. 워크로드 크기는 스토리지 객체에 할당된 공간 또는 스토리지 객체가 사용하는 공간으로 표시됩니다.

ONTAP 9.5 이상의 처리량 층에서는 사용된 공간을 사용할 수 있습니다. ONTAP 9.4 및 이전 버전의 처리량 바닥재에서는 지원되지 않습니다.

- allocated space_policy는 스토리지 객체의 공칭 크기에 따라 IOPS/TB|GB 비율을 유지합니다. 비율이 100 IOPS/GB인 경우, 볼륨이 해당 크기로 유지되는 한 150 GB 볼륨의 처리량은 최대 15,000 IOPS입니다. 볼륨의 크기를 300GB로 변경하면 적응형 QoS는 처리량의 상한을 30,000 IOPS로 조정합니다.
- a_used space_policy(기본값)는 스토리지 효율성 이전에 저장된 실제 데이터의 양에 따라 IOPS/TB|GB 비율을 유지합니다. 비율이 100 IOPS/GB인 경우 100GB 데이터가 저장된 150GB 볼륨의 처리량은 최대 10,000 IOPS입니다. 사용된 공간의 양이 변경되면 적응형 QoS는 비율에 따라 처리량 상한을 조정합니다.

ONTAP 9.5부터 IOPS 및 MBPS 모두에서 처리량 제한을 나타낼 수 있도록 응용 프로그램에 대한 I/O 블록 크기를 지정할 수 있습니다. MBPS 제한은 블록 크기에 IOPS 제한을 곱하여 계산됩니다. 예를 들어, 6144IOPS/TB의 IOPS 한계 32K의 I/O 블록 크기는 192MBps의 MBPS 제한을 생성합니다.

처리량 천장과 바닥에 대해 다음과 같은 동작을 예상할 수 있습니다.

- 워크로드가 적응형 QoS 정책 그룹에 할당되면 상한 또는 하한 이 즉시 업데이트됩니다.
- 적응형 QoS 정책 그룹의 워크로드 크기를 조정하면 최대 또는 최저값이 약 5분 내에 업데이트됩니다.

업데이트를 적용하기 전에 처리량이 최소 10 IOPS 이상 증가해야 합니다.

적응형 QoS 정책 그룹은 항상 공유되지 않습니다. 정의된 처리량 상한 또는 최저값은 각 구성원 워크로드에 개별적으로 적용됩니다.

ONTAP 9.6부터 SSD가 장착된 ONTAP Select 프리미엄에서 처리량 플로어를 지원합니다.

적응형 정책 그룹 템플릿입니다

ONTAP 9.13.1부터 SVM에 적응형 QoS 템플릿을 설정할 수 있습니다. 적응형 정책 그룹 템플릿을 사용하면 SVM에서 모든 볼륨의 처리량 플로어 및 한도를 설정할 수 있습니다.

적응형 정책 그룹 템플릿은 SVM이 생성된 후에만 설정할 수 있습니다. 를 사용합니다 `vserver modify` 명령과 함께 `-qos-adaptive-policy-group-template` 매개 변수를 사용하여 정책을 설정합니다.

적응형 정책 그룹 템플릿을 설정하면 정책을 설정한 후 생성되거나 마이그레이션된 볼륨이 정책을 자동으로 상속합니다. 정책 템플릿을 할당할 때 SVM에 존재하는 볼륨은 영향을 받지 않습니다. SVM에서 정책을 사용하지 않도록 설정하면 이후에 SVM으로 마이그레이션되거나 SVM에서 생성된 모든 볼륨에서 정책을 받지 않습니다. 적응형 정책 그룹 템플릿을 사용하지 않도록 설정해도 정책 템플릿을 상속한 볼륨에는 영향을 주지 않습니다.

자세한 내용은 을 참조하십시오 적응형 정책 그룹 템플릿을 설정합니다.

일반 지원

다음 표에는 처리량 천장, 처리량 바닥 및 적응형 QoS 지원 차이의 차이가 나와 있습니다.

리소스 또는 기능	처리량 한도	처리량 플로어	처리량 플로어 **v2**	적응형 **QoS**
ONTAP 9 버전	모두	9.2 이상	9.7 이상	9.3 이상
플랫폼	모두	• AFF • C190(1) • ONTAP Select 프리미엄 및 SSD 1	• AFF • C190 • SSD가 포함된 ONTAP Select 프리미엄	모두
프로토콜	모두	모두	모두	모두
FabricPool	예	예. 계층화 정책이 "없음"으로 설정되고 클라우드에 블록이 없는 경우	예. 계층화 정책이 "없음"으로 설정되고 클라우드에 블록이 없는 경우	아니요
SnapMirror Synchronous	예	아니요	아니요	예

ONTAP 9.6 릴리스부터 C190 및 ONTAP Select 지원이 시작되었습니다.

처리량 상한에 대해 지원되는 워크로드

다음 표에서는 ONTAP 9 버전별 처리량 천장에 대한 워크로드 지원을 보여 줍니다. 루트 볼륨, 로드 공유 미러 및 데이터 보호 미러는 지원되지 않습니다.

워크로드 지원	**ONTAP 9.8** 이상	**ONTAP 9.7~9.4**	**ONTAP 9.3** 및 이전 버전
볼륨	예	예	예
파일	예	예	예
LUN을 클릭합니다	예	예	예
SVM	예	예	예
FlexGroup 볼륨	예	예	예(ONTAP 9.3만 해당)
qtree(1)	예	아니요	아니요
정책 그룹당 워크로드가 여러 개일 수 있습니다	예	예	예
비공유 정책 그룹입니다	예	예	아니요

[1] ONTAP 9.9.1부터 SMB가 활성화된 FlexVol 및 FlexGroup 볼륨의 qtree에서도 SMB 액세스가 지원됩니다. ONTAP 9.8부터 NFS가 활성화된 FlexVol 및 FlexGroup 볼륨의 qtree에서 NFS 액세스가 지원됩니다.

처리량 플로어에 대해 지원되는 워크로드

다음 표에는 ONTAP 9 버전별 처리량 플로어에 대한 워크로드 지원이 나와 있습니다. 루트 볼륨, 로드 공유 미러 및 데이터 보호 미러는 지원되지 않습니다.

워크로드 지원	**ONTAP 9.13.1** 이상	**ONTAP 9.8 ~ 9.13.0**	**ONTAP 9.4 - 9.7**	**ONTAP 9.3**
볼륨	예	예	예	예
파일	예	예	예	예
LUN을 클릭합니다	예	예	예	예
SVM	예	아니요	아니요	아니요
FlexGroup 볼륨	예	예	예	아니요
qtree(1)	예	예	아니요	아니요
정책 그룹당 워크로드가 여러 개일 수 있습니다	예	예	예	아니요

워크로드 지원	**ONTAP 9.13.1** 이상	**ONTAP 9.8 ~ 9.13.0**	**ONTAP 9.4 - 9.7**	**ONTAP 9.3**
비공유 정책 그룹입니다	예	예	예	아니요

ONTAP 9.8부터 NFS 액세스가 FlexVol 및 FlexGroup 볼륨에서 지원됩니다. ONTAP 9.9.1부터 SMB가 활성화된 FlexVol 및 FlexGroup 볼륨의 qtree에서도 SMB 액세스가 지원됩니다.

적응형 **QoS**에 지원되는 워크로드

다음 표는 ONTAP 9 버전별 적응형 QoS에 대한 워크로드 지원을 보여줍니다. 루트 볼륨, 로드 공유 미러 및 데이터 보호 미러는 지원되지 않습니다.

워크로드 지원	**ONTAP 9.13.1** 이상	**ONTAP 9.4 ~ 9.13.0**	**ONTAP 9.3**
볼륨	예	예	예
파일	예	예	아니요
LUN을 클릭합니다	예	예	아니요
SVM	예	아니요	아니요
FlexGroup 볼륨	예	예	아니요
정책 그룹당 워크로드가 여러 개일 수 있습니다	예	예	예
비공유 정책 그룹입니다	예	예	예

최대 워크로드 및 정책 그룹 수

다음 표에는 ONTAP 9 버전별 최대 워크로드 및 정책 그룹 수가 나와 있습니다.

워크로드 지원	**ONTAP 9.4** 이상	**ONTAP 9.3** 및 이전 버전
클러스터당 최대 워크로드	40,000개	12,000
노드당 최대 워크로드	40,000개	12,000
최대 정책 그룹 수	12,000	12,000

ONTAP 처리량 바닥 **v2** 활성화 또는 비활성화

AFF에서 Throughput Floor v2를 활성화 또는 비활성화할 수 있습니다. 기본값은 ENABLED입니다. v2를 지원하는 경우, 컨트롤러가 다른 워크로드에 지연 시간이 길어지는 대신 사용량이 많은 경우 처리량 플로어를 충족할 수 있습니다. 버전 v2는 QoS 및 Adaptive QoS 모두에 적용됩니다.

단계

1. 고급 권한 레벨로 변경:

세트 프리빌리지 고급

2. 다음 명령 중 하나를 입력합니다.

원하는 작업	다음 명령을 사용합니다.
Floor v2를 비활성화합니다	QoS 설정 처리량-층-v2-활성화 거짓
Floor v2를 활성화합니다	QoS 설정 처리량-층-v2-활성화 true

MetroCluster 클러스터에서 처리량 층 v2를 비활성화하려면 를 실행해야 합니다

QoS 설정 처리량-층-v2-활성화 거짓

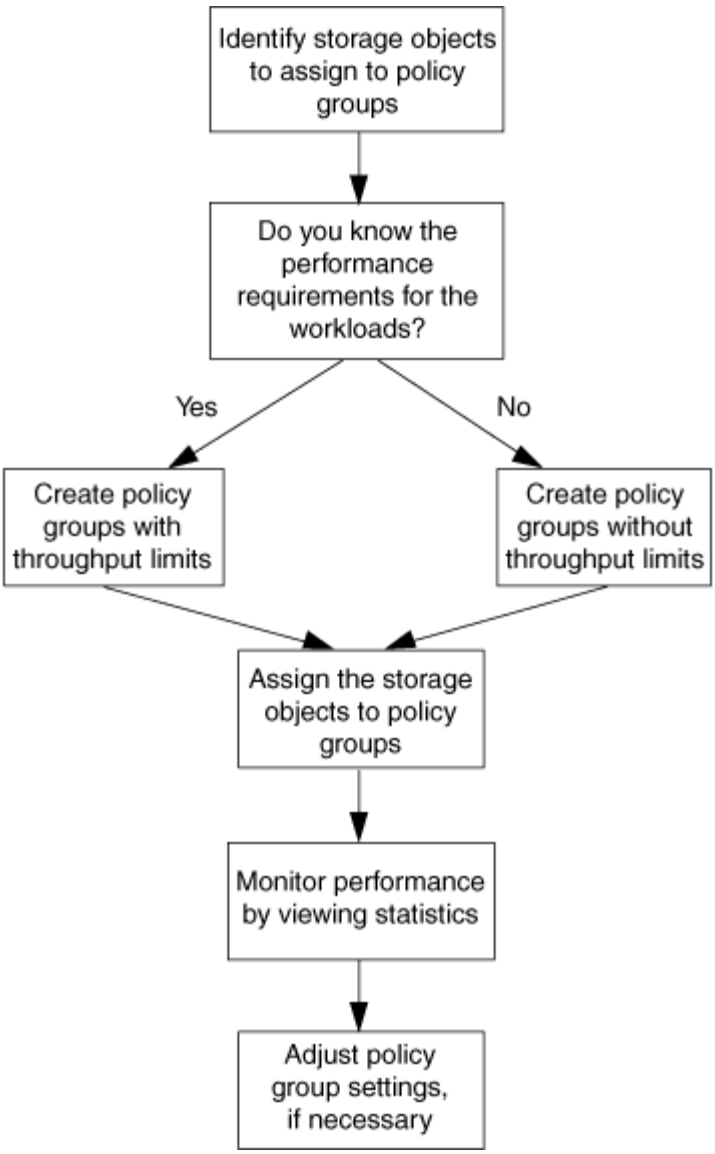
소스 클러스터와 대상 클러스터 모두에서 명령을 내립니다.

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

에 대한 자세한 내용은 `qos settings throughput-floors-v2` "ONTAP 명령 참조입니다"을 참조하십시오.

ONTAP 스토리지 QoS 워크플로

QoS를 사용하여 관리할 워크로드의 성능 요구사항을 이미 알고 있는 경우 정책 그룹을 생성할 때 처리량 한도를 지정할 수 있습니다. 그렇지 않으면 워크로드를 모니터링하여 한도를 지정할 때까지 기다릴 수 있습니다.



ONTAP에서 QoS를 사용하여 처리량의 한계를 설정합니다

정책 그룹에 대해 '최대 처리량' 필드를 사용하여 스토리지 오브젝트 워크로드의 처리량 한도(QoS Max)를 정의할 수 있습니다. 스토리지 객체를 생성하거나 수정할 때 정책 그룹을 적용할 수 있습니다.

시작하기 전에

- 정책 그룹을 생성하려면 클러스터 관리자여야 합니다.
- SVM에 정책 그룹을 적용하려면 클러스터 관리자여야 합니다.

이 작업에 대해

- ONTAP 9.4부터 _non-shared_QoS 정책 그룹을 사용하여 정의된 처리량 상한이 각 구성원 워크로드에 개별적으로 적용되도록 지정할 수 있습니다. 그렇지 않으면 정책 그룹이 _ 공유: _ 입니다. 정책 그룹에 할당된 워크로드의 총 처리량은 지정된 한도를 초과할 수 없습니다.

  비공유 정책 그룹을 지정하기 위해 QoS policy-group create 명령에 대해 -is-shared=false를 설정합니다.

- 천장에 대한 처리량 제한을 IOPS, MB/s 또는 IOPS, MB/s로 지정할 수 있습니다 IOPS와 MB/s를 모두 지정하는 경우, 어느 한쪽의 제한에 먼저 도달하더라도 적용됩니다.

> 동일한 워크로드에 대해 천장과 바닥을 설정하는 경우 IOPS만 사용하여 천장에 대한 처리량 제한을 지정할 수 있습니다.

- QoS 제한이 적용되는 스토리지 개체는 정책 그룹이 속한 SVM에 포함되어 있어야 합니다. 여러 정책 그룹이 동일한 SVM에 속할 수 있습니다.
- 포함하는 객체 또는 해당 하위 객체가 정책 그룹에 속해 있는 경우 스토리지 객체를 정책 그룹에 할당할 수 없습니다.
- 정책 그룹을 동일한 유형의 스토리지 객체에 적용하는 것은 QoS 모범 사례입니다.

단계

1. 정책 그룹 생성:

   "QoS policy-group create-policy-group_policy_group_-vserver_SVM_-max-throughput_number_of_IOPS_|MB/S|IOPS, MB/S-is-shared true|false

   에 대한 자세한 내용은 `qos policy-group create` "ONTAP 명령 참조입니다"을 참조하십시오.

   명령을 사용하여 처리량 한도를 조정할 수 `qos policy-group modify` 있습니다.

   다음 명령을 실행하면 최대 5,000 IOPS의 처리량을 갖는 공유 정책 그룹 'pg-vs1'이 생성됩니다.

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

   다음 명령을 실행하면 비공유 정책 그룹 'pg-vs3'이 생성되고 최대 처리량은 100IOPS, 400KB/S입니다.

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

   다음 명령을 실행하면 처리량 제한 없이 비공유 정책 그룹 'pg-vs4'가 생성됩니다.

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

   에 대한 자세한 내용은 `qos policy-group modify` "ONTAP 명령 참조입니다"을 참조하십시오.

2. SVM, 파일, 볼륨 또는 LUN에 정책 그룹 적용:

   '*storage_object_create-vserver_SVM*-QoS-policy-group_policy_group_'

   이 절차에서 설명하는 명령에 대한 자세한 내용은 를 "ONTAP 명령 참조입니다"참조하십시오. 명령을 사용하여 스토리지 객체에 다른 정책 그룹을 적용할 수 *`storage_object`* `modify` 있습니다.

다음 명령은 SVM 'VS1'에 정책 그룹 'pg-vs1'을 적용합니다.

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

다음 명령은 정책 그룹 pg-app을 볼륨 app1과 app2에 적용합니다.

```
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. 정책 그룹 성능 모니터링:

   QoS 통계 성능 표시

   에 대한 자세한 내용은 `qos statistics performance show` "ONTAP 명령 참조입니다"을 참조하십시오.

   > 클러스터의 성능 모니터링 호스트의 툴을 사용하여 성능을 모니터링하지 마십시오.

   다음 명령을 실행하면 정책 그룹 성능이 표시됩니다.

```
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. 워크로드 성능 모니터링:

   QoS 통계 워크로드 성능 표시

   > 클러스터의 성능 모니터링 호스트의 툴을 사용하여 성능을 모니터링하지 마십시오.

   다음 명령을 실행하면 워크로드 성능이 표시됩니다.

```
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
```

에 대한 자세한 내용은 `qos statistics workload performance show` "ONTAP 명령 참조입니다"을 참조하십시오.

명령을 사용하여 QoS 워크로드에 대한 자세한 지연 시간 통계를 볼 수 `qos statistics workload latency show` 있습니다. 에 대한 자세한 내용은 `qos statistics workload latency show` "ONTAP 명령 참조입니다"을 참조하십시오.

ONTAP에서 QoS로 처리량 기준을 설정합니다

정책 그룹의 '처리량' 필드를 사용하여 스토리지 오브젝트 워크로드의 처리량(QoS Min)을 정의할 수 있습니다. 스토리지 객체를 생성하거나 수정할 때 정책 그룹을 적용할 수 있습니다. ONTAP 9.8부터는 처리량(IOPS 또는 MBps, IOPS 및 MBps)을 지정할 수 있습니다.

시작하기 전에

- 정책 그룹을 생성하려면 클러스터 관리자여야 합니다.
- ONTAP 9.13.1부터 를 사용하여 SVM 레벨에서 처리량 바닥글을 적용할 수 있습니다.적응형 정책 그룹 템플릿입니다 QoS 정책 그룹을 갖는 SVM에서는 적응형 정책 그룹 템플릿을 설정할 수 없습니다.

이 작업에 대해

- ONTAP 9.4부터 _non-shared_QoS 정책 그룹을 사용하여 각 구성원 작업 부하에 대해 정의된 처리 층을 개별적으로 적용할 수 있습니다. 이는 처리량 플로어의 정책 그룹을 여러 워크로드에 적용할 수 있는 유일한 조건입니다.

  비공유 정책 그룹을 지정하기 위해 QoS policy-group create 명령에 대해 -is-shared=false를 설정합니다.

- 노드나 애그리게이트에 성능 용량(여유 공간)이 충분하지 않은 경우 워크로드에 대한 처리량이 지정된 플로어에 아래로 떨어질 수 있습니다.
- QoS 제한이 적용되는 스토리지 개체는 정책 그룹이 속한 SVM에 포함되어 있어야 합니다. 여러 정책 그룹이 동일한 SVM에 속할 수 있습니다.
- 정책 그룹을 동일한 유형의 스토리지 객체에 적용하는 것은 QoS 모범 사례입니다.
- 처리량 플로어를 정의하는 정책 그룹은 SVM에 적용할 수 없습니다.

단계

1. 에 설명된 대로 노드 또는 애그리게이트에서 적절한 성능 용량이 있는지 확인합니다 "남은 성능 용량 식별".
2. 정책 그룹 생성:

'QoS policy-group create-policy group_policy_group_-vserver_SVM_-min-throughput_QoS_target_-is-shared true|false'

에 대한 자세한 내용은 `qos policy-group create` "ONTAP 명령 참조입니다"을 참조하십시오.

3. 명령을 사용하여 처리량 최저치를 조정할 수 `qos policy-group modify` 있습니다.

   다음 명령을 실행하면 처리량이 최소 1,000 IOPS인 공유 정책 그룹이 `pg-vs2` 생성됩니다.

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

   다음 명령을 실행하면 처리량 제한 없이 비공유 정책 그룹 'pg-vs4'가 생성됩니다.

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

   에 대한 자세한 내용은 `qos policy-group modify` "ONTAP 명령 참조입니다"을 참조하십시오.

4. 볼륨 또는 LUN에 정책 그룹 적용:

   `storage_object` `create -vserver` *`SVM`* `-qos-policy-group` *`policy_group`*`당신은 사용할 수 있습니다 `_storage_object_modify` 저장소 개체에 다른 정책 그룹을 적용하는 명령입니다.

   다음 명령은 정책 그룹 pg-app2를 볼륨 app2에 적용합니다.

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

   이 절차에서 설명하는 명령에 대한 자세한 내용은 를 "ONTAP 명령 참조입니다"참조하십시오.

5. 정책 그룹 성능 모니터링:

   QoS 통계 성능 표시

   > ⓘ 클러스터의 성능 모니터링 호스트의 툴을 사용하여 성능을 모니터링하지 마십시오.

   다음 명령을 실행하면 정책 그룹 성능이 표시됩니다.

```
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
```

에 대한 자세한 내용은 `qos statistics performance show` "ONTAP 명령 참조입니다"을 참조하십시오.

6. 워크로드 성능 모니터링:

QoS 통계 워크로드 성능 표시

> 클러스터의 성능 모니터링 호스트의 툴을 사용하여 성능을 모니터링하지 마십시오.

다음 명령을 실행하면 워크로드 성능이 표시됩니다.

```
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
```

에 대한 자세한 내용은 `qos statistics workload performance show` "ONTAP 명령 참조입니다"을 참조하십시오.

> 명령을 사용하여 QoS 워크로드에 대한 자세한 지연 시간 통계를 볼 수 `qos statistics workload latency show` 있습니다. 에 대한 자세한 내용은 `qos statistics workload latency show` "ONTAP 명령 참조입니다"을 참조하십시오.

ONTAP에서 적응형 QoS 정책 그룹 사용

적응형 *QoS* 정책 그룹을 사용하면 볼륨 크기에 맞게 처리량 상한선이나 하한선을 자동으로 조정하여 볼륨 크기가 변경되더라도 IOPS 대 TB/GB 비율을 유지할 수 있습니다. 이는 대규모 배포에서 수백 또는 수천 개의 작업 부하를 관리할 때 상당한 이점입니다.

시작하기 전에

- ONTAP 9.3 이상을 실행해야 합니다. 적응형 QoS 정책 그룹은 ONTAP 9.3부터 사용할 수 있습니다.
- 정책 그룹을 생성하려면 클러스터 관리자여야 합니다.

이 작업에 대해

스토리지 개체는 적응형 정책 그룹 또는 비적응 정책 그룹의 구성원일 수 있지만 둘 다 속할 수는 없습니다. 스토리지 오브젝트 및 정책의 SVM은 동일해야 합니다. 스토리지 객체가 온라인 상태여야 합니다.

적응형 QoS 정책 그룹은 항상 공유되지 않습니다. 정의된 처리량 상한 또는 최저값은 각 구성원 워크로드에 개별적으로 적용됩니다.

스토리지 오브젝트 크기에 대한 처리량 제한의 비율은 다음 필드의 상호 작용에 의해 결정됩니다.

- `expected-iops`할당된 TB/GB당 예상 최소 IOPS입니다.

> `expected-iops` 는 AFF 플랫폼에서만 보장됩니다. `expected-iops` 계층화 정책이 "없음"으로 설정되어 있고 클라우드에 블록이 없는 경우에만 FabricPool에 대해 보장됩니다. `expected-iops` SnapMirror 동기식 관계에 없는 볼륨에 대해 보장합니다.

- `peak-iops`할당되거나 사용된 TB/GB당 가능한 최대 IOPS입니다.
- 'expected-IOPS-allocation'은 할당된 공간(기본값)이나 사용된 공간을 예상 IOPS에 사용할지 여부를 지정합니다.

> ONTAP 9.5 이상에서 '예상 IOPS 할당'을 사용할 수 있습니다. ONTAP 9.4 이하 버전에서는 지원되지 않습니다.

- peak-IOPS-allocation은 peak-IOPS를 위해 할당된 공간과 사용된 공간(기본값)을 사용할지 여부를 지정한다.
- 절대 최소 IOPS는 절대 최소 IOPS의 수입니다. 이 필드는 매우 작은 스토리지 객체와 함께 사용할 수 있습니다. 절대분 IOPS가 계산된 예상 IOPS보다 크면 peak-IOPS와 expected-IOPS를 모두 재정의합니다.

  예를 들어, 'expected-IOPS'를 1,000 IOPS/TB로 설정하고 볼륨 크기가 1GB 미만인 경우 계산된 'expected-IOPS'는 IOP가 분수 값이 됩니다. 계산된 피크 IOPS는 이보다 훨씬 적은 비율입니다. 절대-최소-IOPS를 실제 값으로 설정하면 이러한 문제를 방지할 수 있습니다.

- 블록 크기 는 애플리케이션 입출력 블록 크기를 지정합니다. 기본값은 32K입니다. 유효한 값은 8K, 16K, 32K, 64K, any입니다. Any는 블록 크기가 적용되지 않음을 의미합니다.

기본 적응형 **QoS** 정책 그룹

다음 표에 나와 있는 것처럼 세 가지 기본 적응형 QoS 정책 그룹을 사용할 수 있습니다. 이러한 정책 그룹을 볼륨에 직접 적용할 수 있습니다.

기본 정책 그룹입니다	예상 IOPS/TB	최대 IOPS/TB	절대 최소 IOPS
"익스트림"	6,144	12,288	1000입니다
'퍼포먼스'	2,048	4,096개	500입니다
값	128	512	75를

스토리지 개체 정책 그룹 할당 제한 사항

어떤 경우에는 정책 그룹에 저장 객체를 할당할 수 없습니다. 저장 객체를 포함하는 객체나 해당 객체의 자식 객체가 정책 그룹에 속해 있는 경우입니다.

ONTAP 9.18.1부터 중첩된 QoS 정책을 사용할 수 있습니다. 이를 통해 SVM과 같은 포함 개체와 볼륨과 같은 자식 개체 모두에 정책 그룹을 할당할 수 있습니다. 멀티 테넌트 환경에서 중첩된 QoS 정책을 사용하면 관리자가 SVM 내의 볼륨과 Q트리에서 SVM의 QoS 제한을 세분화하고, 컴퓨팅 환경 전반에서 스토리지 리소스의 균형을 맞추는 동시에 미션 크리티컬 워크로드의 우선 순위를 지정할 수 있습니다.

다음 개체 쌍에 대해 중첩된 QoS 정책이 지원됩니다.

- SVM에 포함된 SVM 및 FlexVol 또는 FlexGroup 볼륨.
- 볼륨 내의 FlexVol 또는 FlexGroup 볼륨과 qtree.

중첩된 QoS 정책의 경우 적용 가능한 가장 제한적인 정책이 사용됩니다.

다음 표에는 제한 사항이 나열되어 있습니다.

다음을 할당하는 경우...	그러면 정책 그룹에 할당할 수 없습니다.
SVM을 정책 그룹으로 이동합니다	SVM에 포함된 모든 저장 객체. ONTAP 9.18.1을 실행 중인 경우 SVM에 포함된 FlexVol 및 FlexGroup 볼륨을 정책 그룹에 할당할 수 있습니다.
볼륨을 정책 그룹에	볼륨이나 자식 LUN이 포함된 SVM입니다. ONTAP 9.18.1 이상을 실행 중인 경우 볼륨이 포함된 SVM을 정책 그룹에 할당할 수 있습니다. 또한 FlexVol 또는 FlexGroup 볼륨의 qtree를 할당할 수 있습니다.
정책 그룹에 LUN을 지정합니다	LUN을 포함하는 볼륨 또는 SVM
파일을 정책 그룹에 저장합니다	파일이 포함된 볼륨 또는 SVM

단계

1. 적응형 QoS 정책 그룹을 생성합니다.

   'QOS adaptive-policy-group create-policy group_group_-vserver_SVM_-expected-IOPS_number_of_IOPS_/TB|GB-peak-IOPS_number_of_IOPS_/TB|GB-expected-IOPS-allocation-space|used-space-peak-peak-IOPS-allocation-space|used-space-space-absolute-min-64K_K_number_K_number_K_number_K_K_number|K_K_K_number|K_MB|K_

   에 대한 자세한 내용은 `qos adaptive-policy-group create` "ONTAP 명령 참조입니다"을 참조하십시오.

ONTAP 9.5 이상에서 '-expected-IOPS-allocation' 및 '-block-size를 사용할 수 있다. 이러한 옵션은 ONTAP 9.4 이전 버전에서는 지원되지 않습니다.

다음 명령을 실행하면 300 IOPS/TB로 설정된 adpg-app1 Adaptive QoS 정책 그룹, 1,000 IOPS/TB로 설정된 -peak-IOPS, 사용된 공간으로 설정된 -peak-IOPS-allocation, 50 IOPS로 설정된 -absolute-min-IOPS 정책 그룹,

```
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. 적응형 QoS 정책 그룹을 볼륨에 적용합니다.

'volume create-vserver SVM-volume_volume_-aggregate_aggregate_-size_number_of_TB|GB-QoS-adaptive-policy-group_policy_group_'

에 대한 자세한 내용은 `volume create` "ONTAP 명령 참조입니다"을 참조하십시오.

다음 명령은 Adaptive QoS 정책 그룹 'adpg-app1'을 볼륨 'app1'에 적용합니다.

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

다음 명령은 새 볼륨 app4와 기존 볼륨 app5에 기본 적응형 QoS 정책 그룹 "extreme"을 적용합니다. 정책 그룹에 대해 정의된 처리량 상한은 볼륨 app4 및 app5에 개별적으로 적용됩니다.

```
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
```

ONTAP에서 적응형 정책 그룹 템플릿을 설정합니다

ONTAP 9.13.1부터 적응형 정책 그룹 템플릿을 사용하여 SVM 레벨에서 처리량 하한과 상한을 적용할 수 있습니다.

이 작업에 대해

- 적응형 정책 그룹 템플릿은 기본 정책입니다 `apg1`. 정책은 언제든지 수정할 수 있습니다. CLI 또는 ONTAP REST API에서만 설정할 수 있고 기존 SVM에만 적용할 수 있습니다.
- 적응형 정책 그룹 템플릿은 정책을 설정한 후 SVM에서 생성되거나 SVM으로 마이그레이션된 볼륨에만 영향을 줍니다. SVM의 기존 볼륨은 기존 상태를 유지합니다.

적응형 정책 그룹 템플릿을 비활성화하면 SVM의 볼륨은 기존 정책을 그대로 유지합니다. 이후에 SVM에서 생성하거나 SVM으로 마이그레이션한 볼륨만 지원 해제에 의해 영향을 받습니다.

- QoS 정책 그룹을 갖는 SVM에서는 적응형 정책 그룹 템플릿을 설정할 수 없습니다.
- 적응형 정책 그룹 템플릿은 AFF 플랫폼용으로 설계되었습니다. 적응형 정책 그룹 템플릿은 다른 플랫폼에 설정할 수 있지만 정책이 최소 처리량을 적용하지 않을 수 있습니다. 마찬가지로, FabricPool 최소 처리량을 지원하지 않는 애그리게이트의 SVM에 적응형 정책 그룹 템플릿 을 추가할 수도 있지만, 처리량 수준이 적용되지 않습니다.
- SVM이 MetroCluster 구성 또는 SnapMirror 관계에 있는 경우 적응형 정책 그룹 템플릿이 미러링된 SVM에 적용됩니다.

단계

1. SVM을 수정하여 적응형 정책 그룹 템플릿을 적용합니다. `vserver modify -qos-adaptive-policy -group-template apg1`
2. 정책이 설정되었는지 확인합니다. `vserver show -fields qos-adaptive-policy-group`

저작권 정보

Copyright © 2026 NetApp, Inc. All Rights Reserved. 미국에서 인쇄됨 본 문서의 어떠한 부분도 저작권 소유자의 사전 서면 승인 없이는 어떠한 형식이나 수단(복사, 녹음, 녹화 또는 전자 검색 시스템에 저장하는 것을 비롯한 그래픽, 전자적 또는 기계적 방법)으로도 복제될 수 없습니다.

NetApp이 저작권을 가진 자료에 있는 소프트웨어에는 아래의 라이센스와 고지사항이 적용됩니다.

본 소프트웨어는 NetApp에 의해 '있는 그대로' 제공되며 상품성 및 특정 목적에의 적합성에 대한 명시적 또는 묵시적 보증을 포함하여(이에 제한되지 않음) 어떠한 보증도 하지 않습니다. NetApp은 대체품 또는 대체 서비스의 조달, 사용 불능, 데이터 손실, 이익 손실, 영업 중단을 포함하여(이에 국한되지 않음), 이 소프트웨어의 사용으로 인해 발생하는 모든 직접 및 간접 손해, 우발적 손해, 특별 손해, 징벌적 손해, 결과적 손해의 발생에 대하여 그 발생 이유, 책임론, 계약 여부, 엄격한 책임, 불법 행위(과실 또는 그렇지 않은 경우)와 관계없이 어떠한 책임도 지지 않으며, 이와 같은 손실의 발생 가능성이 통지되었다 하더라도 마찬가지입니다.

NetApp은 본 문서에 설명된 제품을 언제든지 예고 없이 변경할 권리를 보유합니다. NetApp은 NetApp의 명시적인 서면 동의를 받은 경우를 제외하고 본 문서에 설명된 제품을 사용하여 발생하는 어떠한 문제에도 책임을 지지 않습니다. 본 제품의 사용 또는 구매의 경우 NetApp에서는 어떠한 특허권, 상표권 또는 기타 지적 재산권이 적용되는 라이센스도 제공하지 않습니다.

본 설명서에 설명된 제품은 하나 이상의 미국 특허, 해외 특허 또는 출원 중인 특허로 보호됩니다.

제한적 권리 표시: 정부에 의한 사용, 복제 또는 공개에는 DFARS 252.227-7013(2014년 2월) 및 FAR 52.227-19(2007년 12월)의 기술 데이터-비상업적 품목에 대한 권리(Rights in Technical Data -Noncommercial Items) 조항의 하위 조항 (b)(3)에 설명된 제한사항이 적용됩니다.

여기에 포함된 데이터는 상업용 제품 및/또는 상업용 서비스(FAR 2.101에 정의)에 해당하며 NetApp, Inc.의 독점 자산입니다. 본 계약에 따라 제공되는 모든 NetApp 기술 데이터 및 컴퓨터 소프트웨어는 본질적으로 상업용이며 개인 비용만으로 개발되었습니다. 미국 정부는 데이터가 제공된 미국 계약과 관련하여 해당 계약을 지원하는 데에만 데이터에 대한 전 세계적으로 비독점적이고 양도할 수 없으며 재사용이 불가능하며 취소 불가능한 라이센스를 제한적으로 가집니다. 여기에 제공된 경우를 제외하고 NetApp, Inc.의 사전 서면 승인 없이는 이 데이터를 사용, 공개, 재생산, 수정, 수행 또는 표시할 수 없습니다. 미국 국방부에 대한 정부 라이센스는 DFARS 조항 252.227-7015(b)(2014년 2월)에 명시된 권한으로 제한됩니다.

상표 정보

NETAPP, NETAPP 로고 및 http://www.netapp.com/TM에 나열된 마크는 NetApp, Inc.의 상표입니다. 기타 회사 및 제품 이름은 해당 소유자의 상표일 수 있습니다.