



使用 QoS 策略群組資訊管理效能

Active IQ Unified Manager

NetApp
January 15, 2026

目錄

使用 QoS 策略群組資訊管理效能	1
儲存QoS如何控制工作負載處理量	1
查看所有叢集上可用的所有 QoS 策略群組	2
查看同一 QoS 策略群組中的磁碟區或 LUN	2
查看應用於特定磁碟區或 LUN 的 QoS 原則群組設置	3
查看效能圖表以比較同一 QoS 策略群組中的磁碟區或 LUN	3
處理量圖表中顯示不同類型的QoS原則	4
在效能資源管理器中查看工作負載 QoS 最小和最大設置	5

使用 QoS 策略群組資訊管理效能

Unified Manager可讓您檢視所有監控叢集上可用的服務品質（QoS）原則群組。這些原則可能是使用ONTAP 支援功能的支援軟體（System Manager或ONTAP 支援功能的CLI）或Unified Manager Performance Service層級原則來定義。Unified Manager也會顯示已指派QoS原則群組的磁碟區和LUN。

如需調整QoS設定的詳細資訊、請參閱 "[效能管理總覽](#)"

儲存QoS如何控制工作負載處理量

您可以建立服務品質（QoS）原則群組、以控制其所含工作負載的每秒I/O（IOPS）或處理量（MB/s）限制。如果工作負載位於沒有設定限制的原則群組中（例如預設原則群組）、或設定的限制不符合您的需求、您可以增加限制、或將工作負載移至具有所需限制的新原則群組或現有原則群組。

「傳統」QoS原則群組可指派給個別工作負載、例如單一磁碟區或LUN。在此情況下、工作負載可以使用完整的處理量限制。QoS原則群組也可以指派給多個工作負載、在此情況下、處理量限制會在工作負載之間「共享」。例如、指派給三個工作負載的QoS限制為9,000 IOPS、會將合併的IOPS限制在超過9,000 IOPS。

也可將「Adaptive」QoS原則群組指派給個別工作負載或多個工作負載。然而、即使指派給多個工作負載、每個工作負載仍會獲得完整的處理量限制、而非與其他工作負載共享處理量值。此外、調適性QoS原則會根據每個工作負載的磁碟區大小、自動調整處理量設定、因此隨著磁碟區大小變更、IOPS與TB的比率也會維持不變。例如、如果調適性QoS原則中的尖峰設定為5,000 IOPS/TB、則10 TB磁碟區的處理量上限為50,000 IOPS。如果磁碟區稍後調整至20 TB、調適性QoS會將最大IOPS調整至10,000次。

從推出《S25 9.5版》開始ONTAP、您可以在定義可調適的QoS原則時納入區塊大小。如此可有效地將原則從IOPS/TB臨界值轉換成MB/s臨界值、適用於工作負載使用非常大的區塊大小、最終使用大量處理量的情況。

對於共享群組QoS原則、當原則群組中所有工作負載的IOPS或MB/s超過設定限制時、原則群組會節流工作負載以限制其活動、進而降低原則群組中所有工作負載的效能。如果原則群組節流產生動態效能事件、則事件說明會顯示相關原則群組的名稱。

在「Performance : All Volumes」（效能：所有磁碟區）檢視中、您可以根據IOPS和MB/s來排序受影響的磁碟區、以查看哪些工作負載的使用率可能是事件的最大原因。在「Performance / Volumes Explorer（效能/磁碟區資源管理器）」頁面中、您可以選取磁碟區上的其他磁碟區或LUN、以比較受影響的工作負載IOPS或Mbps處理量使用量。

原則群組會將過度使用節點資源的工作負載指派給較嚴格的原則群組設定、藉此節流工作負載以限制其活動、進而減少該節點上的資源使用量。不過、如果您希望工作負載能夠使用更多節點資源、可以增加原則群組的值。

您可以使用System Manager、ONTAP 支援各種指令或Unified Manager Performance Service層級來管理原則群組、包括下列工作：

- 建立原則群組
- 新增或移除原則群組中的工作負載
- 在原則群組之間移動工作負載
- 變更原則群組的處理量限制

- 將工作負載移至不同的Aggregate和/或節點

查看所有叢集上可用的所有 QoS 策略群組

您可以顯示Unified Manager正在監控之叢集上可用的所有QoS原則群組清單。這包括傳統的QoS原則、調適性QoS原則、以及由Unified Manager Performance Service層級原則管理的QoS原則。

步驟

1. 在左導覽窗格中、按一下「儲存設備>* QoS原則群組*」。

預設會顯示「效能：傳統QoS原則群組」檢視。

2. 檢視每個可用的傳統QoS原則群組的詳細組態設定。
3. 按一下「Expand (展開)」按鈕 (▼) 在QoS原則群組名稱旁、檢視原則群組的詳細資料。
4. 在「檢視」功能表中、選取其中一個額外選項以檢視所有調適性QoS原則群組、或檢視使用Unified Manager Performance Service層級所建立的所有QoS原則群組。

查看同一 QoS 策略群組中的磁碟區或 LUN

您可以顯示已指派給相同QoS原則群組的磁碟區和LUN清單。

如果傳統的QoS原則群組在多個磁碟區之間「共享」、這有助於查看特定磁碟區是否過度使用原則群組所定義的處理量。它也能協助您決定是否可以將其他磁碟區新增至原則群組、而不會對其他磁碟區造成負面影響。

在調適性QoS原則和Unified Manager效能服務層級原則的情況下、這有助於檢視使用原則群組的所有磁碟區或LUN、以便在變更QoS原則的組態設定時、查看哪些物件會受到影響。

步驟

1. 在左導覽窗格中、按一下「儲存設備>* QoS原則群組*」。

預設會顯示「效能：傳統QoS原則群組」檢視。

2. 如果您對傳統原則群組感興趣、請留在此頁面。否則、請選取其中一個「額外檢視」選項、以顯示Unified Manager Performance Service層級所建立的所有調適性QoS原則群組或所有QoS原則群組。
3. 在您感興趣的QoS原則中、按一下展開按鈕 (▼) 在QoS原則群組名稱旁、檢視更多詳細資料。

The screenshot shows a table of QoS Policy Groups. The columns are: QoS Policy Group, Cluster, SVM, Min Through..., Max Through..., Absolute Min..., Block Size, and Asso. There are two rows visible:

- julia_vs2_cifs_Performance**: Cluster: opm-simplicity, SVM: julia_vs2_cifs, Min Throughput: 2048.0 IOPS/TB, Max Throughput: 4096.0 IOPS/TB, Absolute Minimum: 500IOPS, Block Size: 1.
- julia_vs1_nfs_Performance**: Cluster: opm-simplicity, SVM: julia_vs1_nfs, Min Throughput: 2048.0 IOPS/TB, Max Throughput: 4096.0 IOPS/TB, Absolute Minimum: 500IOPS, Block Size: 2.

Below the table, there is a section titled "Details" with the following information:

- Allocated Capacity: 0.99 TB (progress bar from 0.99 TB to 1.15 TB)
- Associated Objects: 2 Volumes, 0 LUNs (with a red box around "2 Volumes" and "0 LUNs")
- Events: None

- 按一下「Volumes (磁碟區)」或「LUN (LUN)」連結、即可檢視使用此QoS原則的物件。

Volume或LUN的效能詳細目錄頁面會顯示使用QoS原則之物件的已排序清單。

查看應用於特定磁碟區或 LUN 的 QoS 原則群組設置

您可以檢視已套用至磁碟區和LUN的QoS原則群組、也可以連結至效能/QoS原則群組檢視、以顯示每個QoS原則的詳細組態設定。

以下是檢視套用至磁碟區之QoS原則的步驟。檢視LUN資訊的步驟類似。

步驟

- 在左導覽窗格中、按一下「儲存設備>*磁碟區*」。
- 預設會顯示健全狀況：所有磁碟區檢視。
- 在View (檢視) 功能表中、選取* Performance (效能) : Volume in QoS Policy Group* (QoS原則群組*中的Volume)。
- 找到您要檢閱的磁碟區、然後向右捲動、直到看到「* QoS原則群組*」欄為止。
- 按一下QoS原則群組名稱。

系統會顯示對應的「服務品質」頁面、視其為傳統QoS原則、調適性QoS原則或使用Unified Manager Performance Service層級建立的QoS原則而定。

- 檢視QoS原則群組的詳細組態設定。
- 按一下「Expand (展開)」按鈕 (▼) 在QoS原則群組名稱旁、檢視原則群組的詳細資料。

查看效能圖表以比較同一 QoS 策略群組中的磁碟區或 LUN

您可以檢視相同QoS原則群組中的磁碟區和LUN、然後比較單一IOPS、MB/s或IOPS/TB圖

表上的效能、找出任何問題。

比較相同QoS原則群組中磁碟區效能的步驟如下所示。檢視LUN資訊的步驟類似。

步驟

1. 在左導覽窗格中、按一下「儲存設備>*磁碟區*」。

預設會顯示健全狀況：所有磁碟區檢視。

2. 在View（檢視）功能表中、選取* Performance（效能）：Volume in QoS Policy Group*（QoS原則群組*中的Volume）。

3. 按一下您要檢閱的Volume名稱。

隨即顯示Volume的Performance Explorer頁面。

4. 在「檢視與比較」功能表中、選取*「同一個QoS原則群組*中的磁碟區」。

下表列出其他共用相同QoS原則的磁碟區。

5. 按一下「新增」按鈕、將這些磁碟區新增到圖表中、以便比較圖表中所有選定磁碟區的IOPS、MB/s、IOPS/TB及其他效能計數器。

您可以變更時間範圍、以檢視不同時間間隔（預設為72小時）的效能。

處理量圖表中顯示不同類型的QoS原則

您可以在Performance Explorer中檢視ONTAP定義的服務品質（QoS）原則設定、以及工作負載分析IOPS、IOPS/TB和MB/s圖表中、檢視已套用至磁碟區或LUN的設定。圖表中顯示的資訊會因套用至工作負載的QoS原則類型而有所不同。

處理量上限（或「尖峰」）設定可定義工作負載可耗用的最大處理量、進而限制對系統資源競爭工作負載的影響。處理量下限（或「預期」）設定可定義工作負載必須具備的最低處理量、使關鍵工作負載無論競爭工作負載的需求為何、都能達到最低處理量目標。

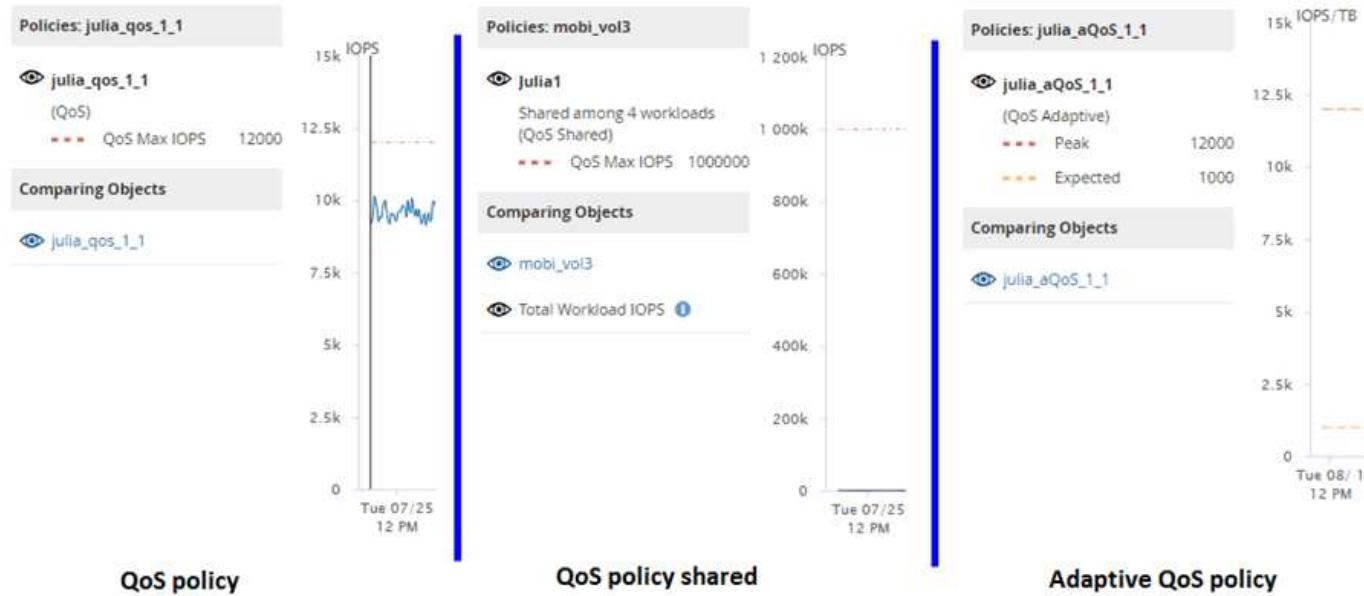
IOPS和MB/s的共享和非共享QoS原則使用「最小」和「最大」兩詞來定義樓層和上限。針對IOPS/TB的調適性QoS原則（在ONTAP《》9.3中推出）、使用「Exp意料之中」和「尖峰」兩詞來定義樓層和上限。

雖然利用VMware可建立這兩種QoS原則、但根據它們在工作負載上的套用方式、QoS原則有三種顯示方式會顯示在效能圖表中。ONTAP

原則類型	功能	Unified Manager介面中的指標
指派給單一工作負載的QoS共用原則、或指派給單一工作負載或多個工作負載的QoS非共用原則	每個工作負載都會耗用指定的處理量設定	顯示「(QoS)」
指派給多個工作負載的QoS共用原則	所有工作負載都會共用指定的處理量設定	顯示「(QoS共享)」

原則類型	功能	Unified Manager介面中的指標
指派給單一工作負載或多個工作負載的QoS原則	每個工作負載都會耗用指定的處理量設定	顯示「(QoS調適性)」

下圖顯示三個選項如何顯示在計數器圖表中的範例。



當工作負載的IOPS/TB圖表中出現以IOPS定義的一般QoS原則時、ONTAP 將IOPS值轉換為IOPS/TB值、而Unified Manager會在IOPS/TB圖表中顯示該原則、並顯示「QoS、以IOPS定義」文字。

在工作負載的IOPS圖表中顯示以IOPS/TB定義的調適性QoS原則時ONTAP、將IOPS/TB值轉換為IOPS值、而Unified Manager會在IOPS圖表中顯示該原則、並顯示「QoS調適性-已使用、定義單位為IOPS/TB」或「QoS調適性-已分配、定義單位為IOPS/TB」、視IOPS分配尖峰設定的設定方式而定。當分配設定設為「分配空間」時、尖峰IOPS會根據磁碟區的大小來計算。當分配設定設為「已用空間」時、尖峰IOPS會根據儲存在磁碟區中的資料量來計算、並考量儲存效率。



IOPS/TB圖表只會在磁碟區使用的邏輯容量大於或等於128 GB時顯示效能資料。當所選時間範圍內使用的容量低於128 GB時、會在圖表中顯示落差。

在效能資源管理器中查看工作負載 QoS 最小和最大設置

您可以在Performance Explorer圖表中檢視Volume或LUN上ONTAP定義的服務品質 (QoS) 原則設定。處理量上限設定可限制競爭工作負載對系統資源的影響。處理量下限設定可確保關鍵工作負載符合最低處理量目標、無論競爭工作負載的需求為何。

QoS處理量「最小」和「最大」IOPS和MB/s設定只有在ONTAP 進行了以供選擇的情況下、才會顯示在計數器圖表中。處理量下限設定僅適用於執行ONTAP 不含更新版本的系統、僅AFF 適用於某些作業系統、目前只能針對IOPS進行設定。

自ONTAP 推出的Adaptive QoS原則以不含IOPS的功能為起點、以IOPS/TB來表示、而非以IOPS來表示。這些原則會根據每個工作負載的磁碟區大小、自動調整QoS原則值、因此隨著磁碟區大小變更、IOPS與TB的比率也會維持不變。您只能將調適性QoS原則群組套用至磁碟區。QoS術語「預期」和「尖峰」用於調適性QoS原則、

而非最小和最大。

當工作負載處理量在前一小時的每個效能收集期間超過定義的QoS最大原則設定時、Unified Manager會針對QoS原則外洩事件產生警告事件。在每個收集期間、工作負載處理量可能只會在短時間內超過QoS臨界值、但Unified Manager會在圖表上顯示收集期間的「平均」處理量。因此、您可能會看到QoS事件、但工作負載的處理量可能未超過圖表中所示的原則臨界值。

步驟

1. 在所選磁碟區或LUN的「效能檔案總管」頁面中、執行下列動作以檢視QoS上限和樓層設定：

如果您想要...	執行此動作...
檢視IOPS上限 (QoS上限)	在「IOPS總計」或「明細」圖表中、按一下「縮放檢視」。
檢視MB/s上限 (QoS上限)	在「MB/s總計」或「明細」圖表中、按一下「縮放檢視」。
檢視IOPS最低 (QoS最低)	在「IOPS總計」或「明細」圖表中、按一下「縮放檢視」。
檢視IOPS/TB上限 (QoS尖峰)	若為磁碟區、請在IOPS/TB圖表中按一下「縮放檢視」。
檢視IOPS/TB樓層 (預期QoS)	若為磁碟區、請在IOPS/TB圖表中按一下「縮放檢視」。

虛線的橫線表示ONTAP 在功能表中設定的最大或最小處理量值。您也可以檢視QoS值的變更實作時間。

2. 若要檢視與QoS設定比較的特定IOPS和MB/s值、請將游標移至圖表區、以查看快顯視窗。

如果您注意到某些磁碟區或LUN的IOPS或MB/s非常高、並在強調系統資源、您可以使用System Manager或ONTAP CLI來調整QoS設定、使這些工作負載不會影響其他工作負載的效能。

如需調整QoS設定的詳細資訊、請參閱 "[效能管理總覽](#)"

版權資訊

Copyright © 2026 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP 「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變更本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 <http://www.netapp.com/TM> 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。