



## 管理性能服务级别

### Active IQ Unified Manager 9.14

NetApp  
March 07, 2024

# 目录

- 性能服务级别 ..... 1
  - 通过分配PSL管理工作负载 ..... 1
  - 创建和编辑性能服务级别 ..... 5

# 性能服务级别

通过性能服务级别(PSL)、您可以定义工作负载的性能和存储目标。您可以在最初创建工作负载时为工作负载分配PSL、也可以稍后通过编辑工作负载来分配PSL。

存储资源的管理和监控基于服务级别目标 (Service Level Objective, SLO)。SLO 由基于所需性能和容量的服务级别协议定义。在 Unified Manager 中, SLO 是指在 NetApp 存储上运行的应用程序的 PSL 定义。存储服务会根据底层资源的性能和利用率进行区分。PSL 是存储服务目标的问题描述。通过 PSL, 存储提供程序可以为工作负载指定性能和容量目标。在为工作负载分配PSL时、ONTAP 上的相应工作负载将通过其性能和容量目标进行管理。每个PSL均按峰值、预期和绝对最小IOPS以及预期延迟进行管理。

Unified Manager具有以下类型的PSL:

- 系统定义: Unified Manager提供了一些无法更改的预设策略。这些预定义的PSL包括:
  - 极致性能
  - 性能
  - 价值

极致性能, 性能和价值 PSL 适用于数据中心中的大多数常见存储工作负载。

Unified Manager还为数据库应用程序提供了三个性能服务级别。这些是性能极高的 PSL, 支持突发 IOPS, 适用于吞吐量需求最高的数据库应用程序。

- 适用于数据库日志的至高性能
- 适用于数据库共享数据的至高性能
- 适用于数据库数据的至高性能
- 用户定义: 如果预定义的性能服务级别不符合您的要求、则可以根据您的需求创建新的PSL。有关信息, 请参见 ["创建和编辑性能服务级别"](#)。
- 超出至高性能: 超出至高性能PSL是系统建议的PSL、适用于要求IOPS高于至高性能的工作负载。系统会根据工作负载的IOPS、容量和延迟对这些工作负载进行内部分析、并在\*存储>工作负载>所有工作负载\*屏幕上为每个工作负载建议使用"超出至高PSL"。您可以将PSL应用于工作负载、以确保最佳性能。

根据工作负载行为、系统会动态生成工作负载的IOPS参数、并以`超过至高<number-(峰值IOPS/TB)><数字(预期IOPS/TB)>`的格式附加到"超出至高" PSL的名称中。例如、如果系统确定工作负载的峰值和预期IOPS分别为`106345`和`37929`、则为该工作负载生成的"超出极高PSL"将命名为`超过极高106345 37929`。虽然系统建议使用这些PSL、但在将其分配给工作负载时、这些PSL的类型会标记为`User-Defined`。

## 通过分配PSL管理工作负载

您可以从\*策略\*>\*性能服务级别\*页面以及使用存储提供程序API访问PSL。通过为存储工作负载分配 PSL 来管理存储工作负载非常方便, 因为您不必单独管理存储工作负载。也可以通过重新分配另一个 PSL 来管理任何修改, 而不是单独管理这些修改。Unified Manager可帮助您根据内部评估和建议为工作负载分配PSL。

有关为工作负载分配系统建议的PSL的信息、请参见 ["为工作负载分配系统建议的 PSL"](#)

"性能服务级别" 页面列出了可用的 PSL 策略, 并可用于添加, 编辑和删除这些策略。



您不能修改系统定义的 PSL 或当前分配给工作负载的 PSL。您不能删除分配给工作负载的 PSL，也不能删除唯一可用的 PSL。

此页面显示以下信息：

字段	Description
Name	PSL的名称。
Type	策略是系统定义的还是用户定义的。
预期 IOPS/TB	应用程序应在 LUN 或文件共享上执行的最小 IOPS 数。预期 IOPS 用于根据存储对象分配的大小指定最小预期分配的 IOPS。
峰值 IOPS/TB	<p>应用程序可以对 LUN 或文件共享执行的最大 IOPS 数。峰值 IOPS 用于根据存储对象的已分配大小或存储对象的已用大小指定可能分配的最大 IOPS。</p> <p>峰值 IOPS 基于分配策略。分配策略为 <code>allocated-space</code> 或 <code>used-space</code>。如果将分配策略设置为 <code>allocated-space</code>，则会根据存储对象的大小计算峰值 IOPS。如果将分配策略设置为 <code>used-space</code>，则峰值 IOPS 将根据存储对象中存储的数据量并考虑存储效率进行计算。默认情况下，分配策略设置为 <code>used-space</code>。</p>

字段	Description
绝对最小 IOPS	<p>当预期 IOPS 小于此值时，绝对最小 IOPS 将用作覆盖。系统定义的 PSL 的默认值如下：</p> <ul style="list-style-type: none"> <li>• 至高性能：如果预期 IOPS <math>\geq 6144/\text{TB}</math>，则绝对最小 IOPS = 1000</li> <li>• 性能：如果预期 IOPS <math>\geq 2048/\text{TB}</math> 且 <math>&lt; 6144/\text{TB}</math>，则绝对最小 IOPS = 500</li> <li>• 值：如果预期 IOPS <math>\geq 128/\text{TB}</math> 且 <math>&lt; 2048/\text{TB}</math>，则绝对最小 IOPS = 75</li> </ul> <p>系统定义的数据库 PSL 的默认值如下：</p> <ul style="list-style-type: none"> <li>• 适用于数据库日志的至高性能：如果预期 IOPS <math>\geq 22528</math>，则绝对最小 IOPS = 4000</li> <li>• 数据库共享数据的至高性能：如果预期 IOPS <math>\geq 16384</math>，则绝对最小 IOPS = 2000</li> <li>• 适用于数据库数据的至高性能：如果预期 IOPS <math>\geq 12288</math>，则绝对最小 IOPS = 2000</li> </ul> <p>自定义 PSL 的绝对最小 IOPS 值越高，最大值可以为 75000。较低的值计算如下：</p> <p><math>1000 / \text{预期延迟}</math></p>
预期延迟	存储 IOPS 的预期延迟，以每操作毫秒数（毫秒 / 操作）为单位。
Capacity	集群中的总可用容量和已用容量。
工作负载	已分配 PSL 的存储工作负载数量。

有关峰值 IOPS 和预期 IOPS 如何帮助在 ONTAP 集群上实现一致的差异化性能的信息，请参见以下知识库文章：[https://kb.netapp.com/Advice\\_and\\_Troubleshooting/Data\\_Infrastructure\\_Management/Active\\_IQ\\_Unified\\_Manager/What\\_is\\_Performance\\_Budgeting%3F](https://kb.netapp.com/Advice_and_Troubleshooting/Data_Infrastructure_Management/Active_IQ_Unified_Manager/What_is_Performance_Budgeting%3F)["什么是性能预算？"]

## 为违反 PSL 定义的阈值的工作负载生成的事件

请注意、如果工作负载在前一小时 30% 的时间内超过预期延迟值、则 Unified Manager 将生成以下事件之一、以通知您可能存在性能问题描述：

- 已违反性能服务级别策略定义的工作负载卷延迟阈值
- 已违反性能服务级别策略定义的工作负载 LUN 延迟阈值。

您可能需要分析工作负载，以查看可能导致延迟值增加的原因。

有关详细信息，请参见以下链接：

- "卷事件"
- "违反性能阈值策略时会发生什么情况"
- "Unified Manager 如何使用工作负载延迟来确定性能问题"
- "什么是性能事件"

## 系统定义的PSL

下表提供了有关系统定义的 PSL 的信息：

性能服务级别	问题描述和用例	预期延迟（毫秒 / 操作）	峰值 IOPS	预期 IOPS	绝对最小 IOPS
极致性能	提供极高的吞吐量和极低的延迟  非常适合延迟敏感型应用程序	1.	12288	6144	1000
性能	提供高吞吐量和低延迟  非常适合数据库和虚拟化应用程序	2.	4096	2048	500
价值	提供高存储容量和中等延迟  非常适合高容量应用程序，例如电子邮件，Web 内容，文件共享和备份目标	17	512	128.	75
适用于数据库日志的至高性能	以最低延迟提供最大吞吐量。  非常适合支持数据库日志的数据库应用程序。此 PSL 可提供最高的吞吐量，因为数据库日志非常突发，并且日志记录始终是按需的。	1.	45056	22528	4000

性能服务级别	问题描述和用例	预期延迟（毫秒 / 操作）	峰值 IOPS	预期 IOPS	绝对最小 IOPS
适用于数据库共享数据的至高性能	以最低延迟提供极高的吞吐量。  非常适合存储在通用数据存储库中但在数据库之间共享的数据库应用程序数据。	1.	32768	16384	2000 年
适用于数据库数据的至高性能	以最低延迟提供高吞吐量。  非常适合数据库应用程序数据，例如数据库表信息和元数据。	1.	24576	12288	2000 年


## 创建和编辑性能服务级别

如果系统定义的性能服务级别与您的工作负载要求不匹配，您可以创建自己的性能服务级别，这些服务级别针对您的工作负载进行了优化。

- 您需要的内容 \*
- 您必须具有应用程序管理员角色。
- 性能服务级别名称必须唯一，并且不能使用以下保留关键字：

主，至尊，性能，价值，未分配，Learning，Idle，Default 和 None。

您可以通过为要访问存储的应用程序定义所需的服务级别目标，从性能服务级别页面创建和编辑自定义性能服务级别。

 如果性能服务级别当前已分配给工作负载，则无法修改此级别。

### 步骤

1. 在左侧导航窗格中的 \* 设置 \* 下，选择 \* 策略 \* > \* 性能服务级别 \*。
2. 在 \* 性能服务级别 \* 页面中，根据您是要创建新的性能服务级别还是要编辑现有性能服务级别，单击相应的按钮。

至 ...	请按照以下步骤操作 ...
创建新的性能服务级别	单击 * 添加 *。
编辑现有性能服务级别	选择现有性能服务级别，然后单击 * 编辑 *。

此时将显示用于添加或编辑性能服务级别的页面。

3. 通过指定性能目标来自定义性能服务级别，然后单击 \* 提交 \* 以保存性能服务级别。

您可以从工作负载页面或在配置新工作负载时将新的或更改的性能服务级别应用于工作负载（LUN，NFS 文件共享，CIFS 共享）。



## 版权信息

版权所有 © 2024 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

## 商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。