



收集数据并监控工作负载性能

Active IQ Unified Manager 9.14

NetApp
November 12, 2024

目录

收集数据并监控工作负载性能	1
Unified Manager 监控的工作负载类型	1
工作负载性能测量值	2
什么是预期性能范围	3
如何在性能分析中使用延迟预测	4
Unified Manager 如何使用工作负载延迟来确定性能问题	5
集群操作如何影响工作负载延迟	6
监控 MetroCluster 配置的性能	6

收集数据并监控工作负载性能

Unified Manager 每 5 分钟收集并分析一次工作负载活动，以确定性能事件，并每 15 分钟检测一次配置更改。它最多保留 30 天的 5 分钟历史性能和事件数据，并使用这些数据预测所有受监控工作负载的预期延迟范围。

Unified Manager 必须至少收集 3 天的工作负载活动，然后才能开始分析，并在 "工作负载分析" 页面和 "事件详细信息" 页面上显示 I/O 响应时间的延迟预测。在收集此活动时，延迟预测不会显示工作负载活动发生的所有更改。在收集3天的活动之后、Unified Manager会每24小时在午夜12: 00调整一次延迟预测、以反映工作负载活动的变化并建立更准确的动态性能阈值。

在 Unified Manager 监控工作负载的前 4 天内，如果自上次数据收集以来已超过 24 小时，则延迟图表将不会显示该工作负载的延迟预测。上次收集之前检测到的事件仍然可用。



夏令时 (DST) 会更改系统时间，从而更改受监控工作负载的性能统计信息的延迟预测。Unified Manager 会立即开始更正延迟预测，大约需要 15 天才能完成。在此期间，您可以继续使用 Unified Manager，但由于 Unified Manager 使用延迟预测来检测动态事件，因此某些事件可能不准确。更改时间之前检测到的事件不受影响。

Unified Manager 监控的工作负载类型

您可以使用 Unified Manager 监控两种类型的工作负载的性能：用户定义的工作负载和系统定义的工作负载。

- * 用户定义的工作负载 _ *

从应用程序到集群的 I/O 吞吐量。这些是读取和写入请求所涉及的进程。卷，LUN，NFS 共享，SMB/CIFS 共享和工作负载是用户定义的工作负载。



Unified Manager 仅监控集群上的工作负载活动。它不会监控应用程序，客户端或应用程序与集群之间的路径。

如果工作负载符合以下一项或多项条件，则 Unified Manager 将无法监控此工作负载：

- 它是处于只读模式的数据保护 (DP) 副本。(对于 DP 卷，系统会监控用户生成的流量。)
- 此卷为脱机数据克隆。
- 它是MetroCluster配置中的镜像卷。

- * 系统定义的工作负载 _ *

与存储效率，数据复制和系统运行状况相关的内部流程，包括：

- 存储效率，例如重复数据删除
- 磁盘运行状况，包括 RAID 重建，磁盘擦洗等
- 数据复制，例如 SnapMirror 副本
- 管理活动

- 文件系统运行状况，包括各种 WAFL 活动
- 文件系统扫描程序，例如 WAFL 扫描
- 副本卸载，例如从 VMware 主机卸载存储效率操作
- 系统运行状况，例如卷移动，数据压缩等
- 不受监控的卷

只有当系统定义的工作负载使用的集群组件处于争用状态时，图形用户界面才会显示这些工作负载的性能数据。例如，您不能搜索系统定义的工作负载的名称以在图形用户界面中查看其性能数据。

工作负载性能测量值

Unified Manager 会根据历史和预期统计值测量集群上的工作负载性能，这些统计值构成工作负载的延迟预测值。它会将实际工作负载统计值与延迟预测进行比较，以确定何时工作负载性能过高或过低。未按预期执行的工作负载将触发动态性能事件以通知您。

在下图中，实际值（红色）表示时间范围内的实际性能统计信息。实际值已超过性能阈值，即延迟预测的上限。峰值是时间范围内的最高实际值。偏差用于衡量预期值（预测）与实际值之间的变化，而峰值偏差则表示预期值与实际值之间的最大变化。



下表列出了工作负载性能测量值。

测量	说明
活动	<p>策略组中的工作负载使用的 QoS 限制百分比。</p> <p>i 如果 Unified Manager 检测到策略组发生更改，例如添加或删除卷或更改 QoS 限制，则实际值和预期值可能会超过设置限制的 100%。如果某个值超过设置限制的 100%，则会显示为 >100%。如果某个值小于设置限制的 1%，则会显示为 <1%。</p>

测量	说明
实际	给定工作负载在特定时间的测量性能值。
偏差	<p>预期值与实际值之间的变化。它是实际值减去预期值与预期范围上限值减去预期值的比率。</p> <div style="display: flex; align-items: center;">  <p>负偏差值表示工作负载性能低于预期，而正偏差值表示工作负载性能高于预期。</p> </div>
预期	预期值基于对给定工作负载的历史性能数据的分析。Unified Manager 将分析这些统计值以确定值的预期范围（延迟预测）。
延迟预测（预期范围）	延迟预测是对特定时间的性能上限和下限值的预测。对于工作负载延迟，上限值构成性能阈值。当实际值超过性能阈值时，Unified Manager 将触发动态性能事件。
峰值	一段时间内测量的最大值。
峰值偏差	一段时间内测量的最大偏差值。
队列深度	在互连组件上等待的待定 I/O 请求数。
利用率	对于网络处理，数据处理和聚合组件，是指一段时间内完成工作负载操作的繁忙时间的百分比。例如，网络处理或数据处理组件处理 I/O 请求或聚合完成读取或写入请求的时间百分比。
写入吞吐量	从本地集群上的工作负载到 MetroCluster 配置中的配对集群的写入吞吐量（以 MB/秒为单位）。

什么是预期性能范围

延迟预测是对特定时间的性能上限和下限值的预测。对于工作负载延迟，上限值构成性能阈值。当实际值超过性能阈值时，Unified Manager 将触发动态性能事件。

例如、在正常工作时间(上午9:00到下午5:00)、大多数员工可能会在上午9:00到上午10:30之间检查电子邮件。对电子邮件服务器的需求增加意味着在此期间后端存储上的工作负载活动增加。员工可能会注意到电子邮件客户端的响应速度较慢。

在中午12:00至下午1:00之间的午餐时间和工作日结束后下午5:00、大多数员工可能会离开计算机。对电子邮件服务器的需求通常会减少，同时也会降低对后端存储的需求。或者、也可能会有计划的工作负载操作(例如、存储备份或病毒扫描)在下午5:00之后开始、并增加后端存储上的活动。

在几天内，工作负载活动的增加和减少决定了活动的预期范围（延迟预测），工作负载的上限和下限。如果某个对象的实际工作负载活动超出上限或下限，并且在一段时间内保持在边界之外，则可能表示该对象已过度使用或未充分使用。

如何形成延迟预测

Unified Manager 必须至少收集 3 天的工作负载活动，然后才能开始分析并在图形用户界面中显示 I/O 响应时间的延迟预测。所需的最低数据收集量不会考虑工作负载活动所发生的所有更改。在收集前3天的活动后，Unified Manager 会每24小时在中午12:00调整一次延迟预测、以反映工作负载活动的变化并建立更准确的动态性能阈值。



夏令时（DST）会更改系统时间，从而更改受监控工作负载的性能统计信息的延迟预测。Unified Manager 会立即开始更正延迟预测，大约需要 15 天才能完成。在此期间，您可以继续使用 Unified Manager，但由于 Unified Manager 使用延迟预测来检测动态事件，因此某些事件可能不准确。更改时间之前检测到的事件不受影响。

如何在性能分析中使用延迟预测

Unified Manager 使用延迟预测来表示受监控工作负载的典型 I/O 延迟（响应时间）活动。当工作负载的实际延迟超过延迟预测的上限时，它会向您发出警报，从而触发动态性能事件，以便您可以分析性能问题描述并采取更正操作来解决此问题。

延迟预测用于设置工作负载的性能基线。随着时间的推移，Unified Manager 会从过去的性能测量结果中学习，以预测工作负载的预期性能和活动级别。预期范围的上限用于建立动态性能阈值。Unified Manager 使用基线来确定实际延迟何时高于或低于阈值或超出其预期范围的界限。实际值与预期值之间的比较可为工作负载创建性能配置文件。

如果由于集群组件上的争用而导致工作负载的实际延迟超过动态性能阈值，则延迟会很高，并且工作负载的运行速度会比预期慢。共享相同集群组件的其他工作负载的性能也可能比预期慢。

Unified Manager 将分析阈值超出事件并确定此活动是否为性能事件。如果高工作负载活动长时间保持一致，例如持续数小时，则 Unified Manager 会将活动视为正常，并动态调整延迟预测以形成新的动态性能阈值。

某些工作负载的活动可能持续较低，而延迟的延迟预测在一段时间内变化率不高。为了最大限度地减少分析性能事件期间的事件数量，Unified Manager 仅会为操作和延迟远远高于预期的低活动卷触发事件。



在此示例中，卷的延迟预测（以灰色显示）最低为每个操作 3.5 毫秒（毫秒 / 操作），最高为 5.5 毫秒 / 操作。如果由于网络流量间歇性峰值或集群组件上的争用，实际延迟（以蓝色显示）突然增加到 10 毫秒 / 操作，则该延迟将高于延迟预测并超过动态性能阈值。

当网络流量减少或集群组件不再处于争用状态时，延迟将返回到延迟预测范围内。如果延迟长时间保持在 10 毫秒 / 操作以上，您可能需要采取更正措施来解决此事件。

Unified Manager 如何使用工作负载延迟来确定性能问题

工作负载延迟（响应时间）是指集群上的卷响应客户端应用程序的 I/O 请求所需的时间。Unified Manager 会使用延迟检测性能事件并向您发出警报。

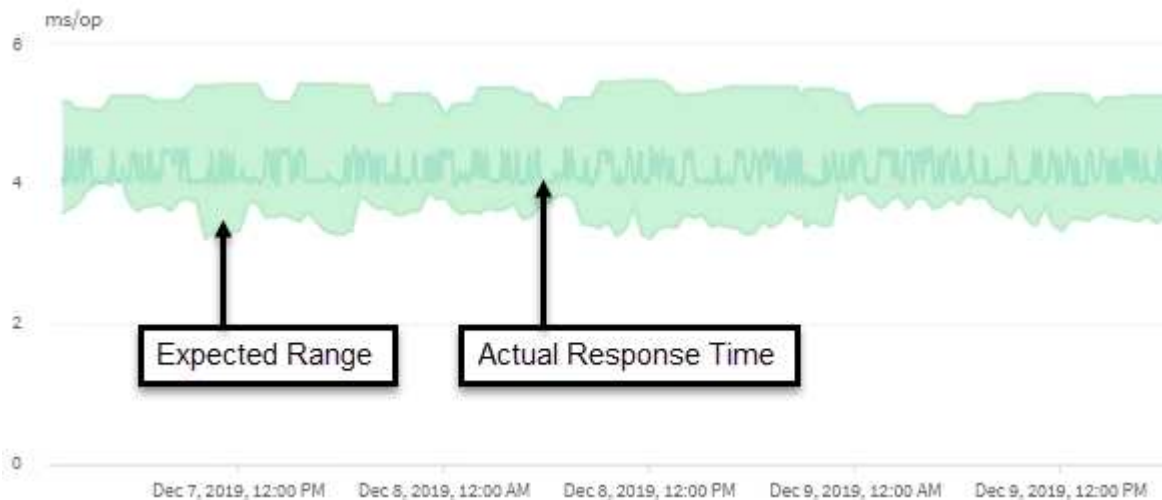
高延迟意味着从应用程序向集群上的卷发出的请求所需时间比平常要长。由于一个或多个集群组件上发生资源争用，高延迟的发生原因可能位于集群本身上。集群外部的因素也可能导致高延迟，例如网络瓶颈，托管应用程序的客户端问题或应用程序本身的问题。



Unified Manager 仅监控集群上的工作负载活动。它不会监控应用程序，客户端或应用程序与集群之间的路径。

集群上的操作（例如创建备份或运行重复数据删除）会增加对其他工作负载共享的集群组件的需求，这也会导致高延迟。如果实际延迟超过预期范围（延迟预测）的动态性能阈值，则 Unified Manager 将分析此事件，以确定它是否为可能需要解决的性能事件。延迟以每操作毫秒数（毫秒 / 操作）为单位。

在“工作负载分析”页面的“总延迟”图表上，您可以查看延迟统计信息的分析，以查看读取和写入请求等单个进程的活动与整体延迟统计信息的比较情况。此比较可帮助您确定哪些操作的活动最多，或者特定操作是否存在影响卷延迟的异常活动。在分析性能事件时，您可以使用延迟统计信息来确定事件是否由集群上的问题描述引起。您还可以确定事件涉及的特定工作负载活动或集群组件。



此示例显示了延迟图表。实际响应时间（延迟）活动为蓝线，延迟预测（预期范围）为绿色。

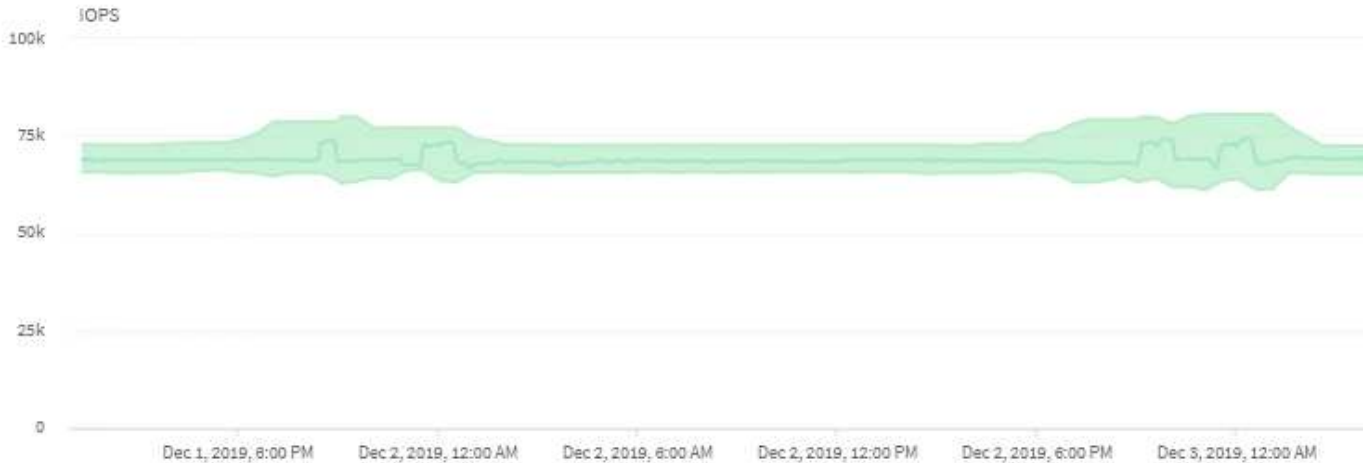


如果 Unified Manager 无法收集数据，蓝线可能会出现空隙。发生这种情况的原因可能是集群或卷无法访问，Unified Manager 在此期间关闭或收集所用时间超过 5 分钟收集时间段。

集群操作如何影响工作负载延迟

操作（IOPS）表示集群上所有用户定义和系统定义的工作负载的活动。IOPS 统计信息可帮助您确定集群进程（例如创建备份或运行重复数据删除）是否正在影响工作负载延迟（响应时间），或者是否可能导致或影响性能事件。

在分析性能事件时，您可以使用 IOPS 统计信息来确定性能事件是否由集群上的问题描述引起。您可以确定可能是性能事件主要促成因素的特定工作负载活动。IOPS 以每秒操作数（操作数 / 秒）为单位。



此示例显示了 IOPS 图表。实际操作统计信息为蓝线，操作统计信息的 IOPS 预测为绿色。



在集群过载的某些情况下、Unified Manager可能会显示消息 `Data collection is taking too long on Cluster cluster_name`。这意味着，没有为 Unified Manager 收集足够的统计信息进行分析。您需要减少集群使用的资源，以便收集统计信息。

监控 MetroCluster 配置的性能

通过 Unified Manager，您可以监控 MetroCluster 配置中集群之间的写入吞吐量，以确定写入吞吐量较高的工作负载。

如果这些高性能工作负载导致本地集群上的其他卷的 I/O 响应时间较长，则 Unified Manager 将触发性能事件以通知您。



Unified Manager 将 MetroCluster 配置中的集群视为单个集群。它不会区分配对集群，也不会关联每个集群的写入吞吐量。

当 MetroCluster 配置中的本地集群将其数据镜像到其配对集群时，数据将写入 NVRAM，然后通过交换机间链路（ISL）传输到远程聚合。Unified Manager 会分析 NVRAM，以确定写入吞吐量较高而过度利用 NVRAM 的工作负载，从而使 NVRAM 处于争用状态。

响应时间偏差超过性能阈值的工作负载称为 `victims_nvram`，而写入 NVRAM 的吞吐量偏差高于正常值并导致出现争用的工作负载称为 `_bulbles`。由于只有写入请求才会镜像到配对集群，因此 Unified Manager 不会分析读取吞吐量。

您可以通过从以下屏幕分析相应 LUN 和卷的工作负载来查看 MetroCluster 配置中任何集群的吞吐量。您可以按集

群筛选结果。从左侧导航窗格中：

- *存储>集群>性能：所有集群*视图。请参见
- *存储>卷>性能：所有卷*视图。
- *存储> LUN >性能：所有LUN*视图。
- 工作负载分析>所有工作负载
- 相关信息 *

"性能事件分析和通知"

"MetroCluster 配置的性能事件分析"

"性能事件中涉及的工作负载的角色"

"确定性能事件中涉及的受影响工作负载"

"确定性能事件中涉及的抢占资源的工作负载"

"确定性能事件中涉及的鲨鱼工作负载"

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。