



Confluent Kafka 与 NetApp ONTAP 存储控制器

NetApp artificial intelligence solutions

NetApp
August 18, 2025

目录

Confluent Kafka 与NetApp ONTAP存储控制器	1
TR-4941: 与NetApp ONTAP存储控制器融合	1
解决方案	1
解决方案架构细节	1
技术概述	2
NetApp ONTAP存储控制器	2
主要用例	3
原生 S3 应用程序	3
FabricPool端点	3
汇合	5
Confluent 性能验证	7
Confluent 设置	7
Confluent 分层存储配置	8
NetApp存储控制器 – ONTAP	8
验证结果	9
使用生产-消费工作负载生成器进行性能测试	10
性能最佳实践指南	11
结束语	12
在哪里可以找到更多信息	12

Confluent Kafka 与 NetApp ONTAP 存储控制器

TR-4941: 与 NetApp ONTAP 存储控制器融合

Karthikeyan Nagalingam、Joe Scott、NetApp Rankesh Kumar、Confluence

为了使 Confluent 平台更具可扩展性和弹性，它必须能够非常快速地扩展和平衡工作负载。分层存储通过减少这种操作负担，使得在 Confluent 中存储大量数据变得易于管理。

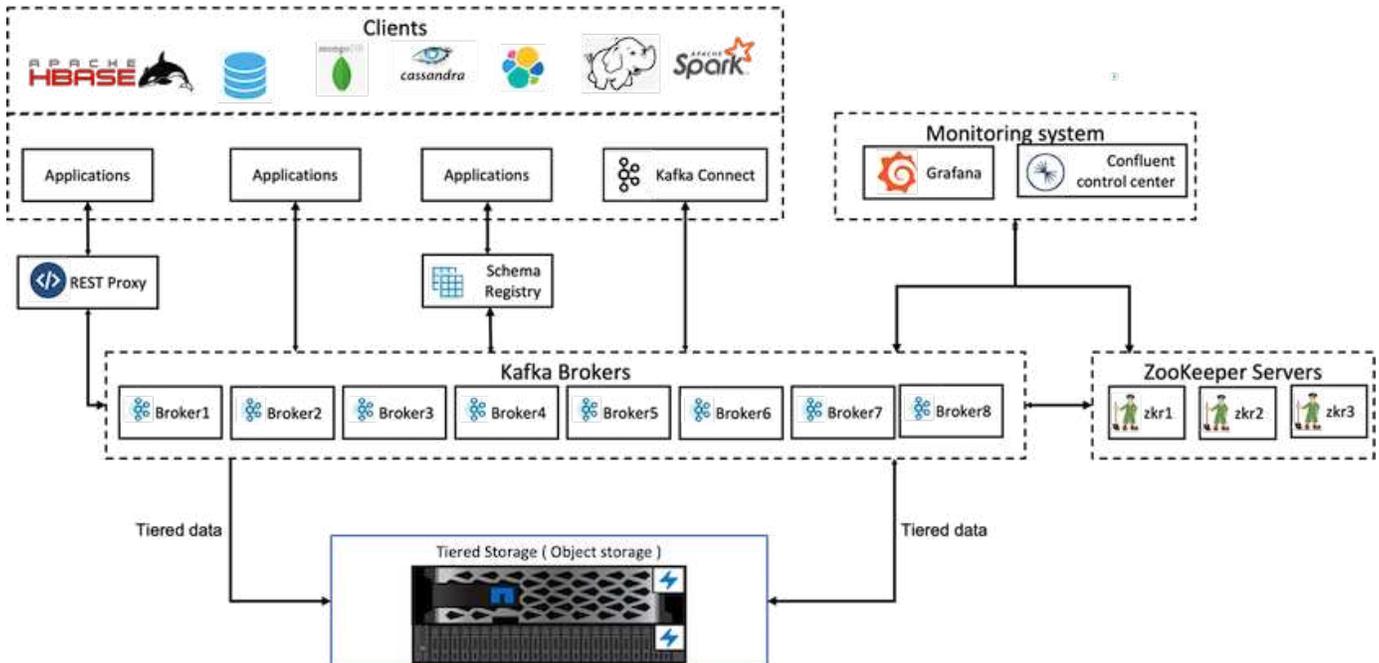
其基本思想是将数据存储与数据处理分开，这使得独立扩展变得更加容易。

NetApp ONTAP 数据管理软件搭载业界领先的创新技术，无论数据位于何处，都能为 Confluent 带来诸多优势。

本文档概述了使用分层存储基准测试套件对 NetApp ONTAP 上的 Confluent 平台进行的性能基准测试。

解决方案

Confluent 和由 ONTAP 提供支持的 NetApp AFF A900 存储控制器是专为数据流设计的分布式系统。两者都具有水平可扩展性和容错性，并且在负载下提供出色的性能。它们在分布式数据流和流处理方面相互补充，通过数据缩减技术最大限度地减少数据占用空间，从而降低存储成本。AFF A900 存储控制器提供出色的性能，同时允许计算和数据存储资源分离。这简化了系统管理并允许独立扩展资源。



解决方案架构细节

本节介绍使用 NetApp ONTAP 进行分层存储的 Confluent Platform 部署中用于性能验证的硬件和软件。下表涵盖了解决方案架构和基本组件。

平台组件	环境配置
Confluent 平台版本 6.2	<ul style="list-style-type: none"> • 3名动物园管理员 • 8 个经纪服务器 • 5 个工具服务器 • 1 个 Grafana • 1 x 控制中心
所有节点上的操作系统	Linux (Ubuntu 18.04)
NetApp ONTAP用于热存储桶	<ul style="list-style-type: none"> • 1 个AFF A900高可用性 (HA) 对 • 4 x 24 x 800 SSD • S3 协议 • 100GbE
15台富士通PRIMERGY RX2540服务器	<ul style="list-style-type: none"> • 2 个 CPU; 总共 16 个物理核心 • 英特尔至强 • 256GB物理内存 • 100GbE 双端口

技术概述

本节介绍此解决方案所使用的技术。

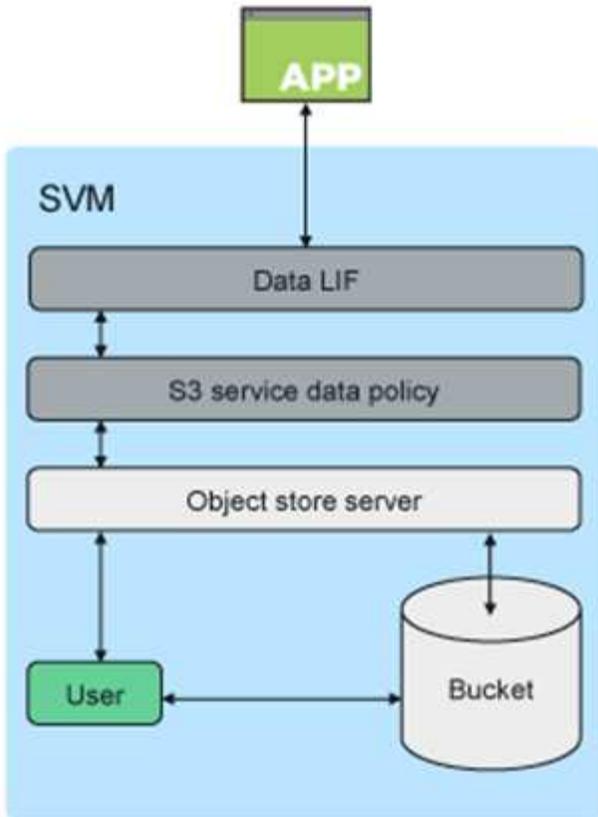
NetApp ONTAP存储控制器

NetApp ONTAP是一款高性能企业级存储操作系统。

NetApp ONTAP 9.8 引入了对 Amazon Simple Storage Service (S3) API 的支持。ONTAP支持 Amazon Web Services (AWS) S3 API 操作的子集，并允许将数据表示为跨云提供商（AWS、Azure 和 GCP）和本地的基于ONTAP的系统中的对象。

NetApp StorageGRID软件是NetApp 的旗舰对象存储解决方案。ONTAP通过在边缘提供摄取和预处理点、扩展由NetApp支持的对象数据数据结构以及增加NetApp产品组合的价值来补充StorageGRID。

通过授权用户和客户端应用程序可以访问 S3 存储桶。下图显示了应用程序访问 S3 存储桶的情况。



主要用例

支持 S3 API 的主要目的是提供 ONTAP 上的对象访问。ONTAP 统一存储架构现在支持文件（NFS 和 SMB）、块（FC 和 iSCSI）和对象（S3）。

原生 S3 应用程序

越来越多的应用程序能够利用 ONTAP 支持通过 S3 进行对象访问。尽管非常适合大容量存档工作负载，但对原生 S3 应用程序的高性能需求正在快速增长，包括：

- 分析
- 人工智能
- 从边缘到核心的采集
- 机器学习

客户现在可以使用熟悉的管理工具（例如 ONTAP 系统管理器）来快速配置用于 ONTAP 中的开发和操作的高性能对象存储，同时充分利用 ONTAP 存储的效率和安全性。

FabricPool 端点

从 ONTAP 9.8 开始，FabricPool 支持对 ONTAP 中的存储桶进行分层，从而允许 ONTAP 到 ONTAP 分层。对于希望将现有 FAS 基础设施重新用作对象存储端点的客户来说，这是一个绝佳的选择。

FabricPool 通过两种方式支持分层到 ONTAP：

- 本地集群分层。使用集群 LIF 将非活动数据分层到位于本地集群上的存储桶中。
- 远程集群分层。非活动数据以类似于传统FabricPool云层的方式分层到位于远程集群上的存储桶中，使用FabricPool客户端上的 IC LIF 和ONTAP对象存储上的数据 LIF。

如果您希望在现有集群上使用 S3 功能而无需额外的硬件和管理，则ONTAP S3 是合适的。对于大于 300TB 的部署， NetApp StorageGRID软件仍然是NetApp对象存储的旗舰解决方案。使用ONTAP或StorageGRID作为云层时不需要FabricPool许可证。

NetApp ONTAP for Confluent 分层存储

每个数据中心都需要保持关键业务应用程序的运行以及重要数据的可用和安全。全新NetApp AFF A900系统采用ONTAP Enterprise Edition 软件和高弹性设计。我们全新的闪电般快速的 NVMe 存储系统可消除对关键任务操作的中断、最大限度地减少性能调整并保护您的数据免受勒索软件攻击。

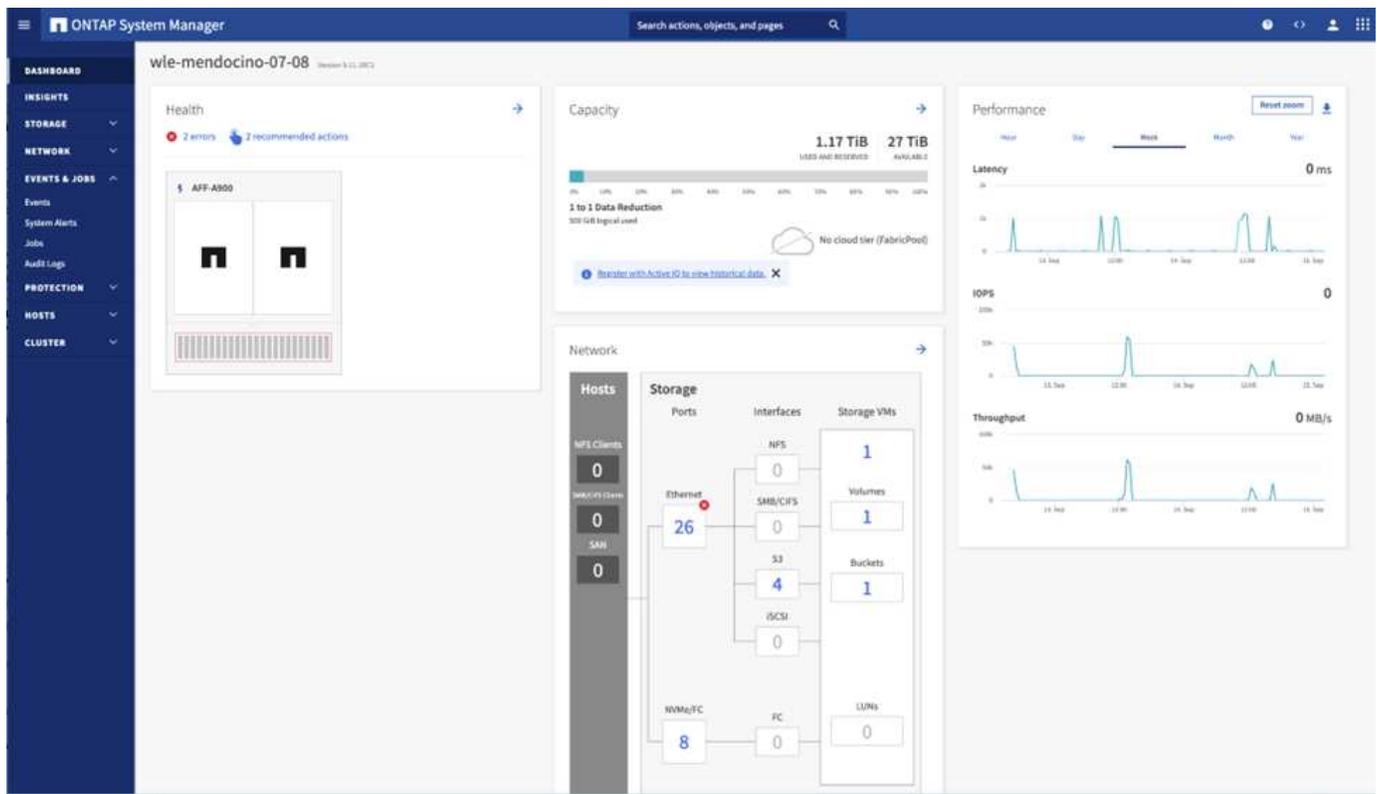
从初始部署到扩展 Confluent 集群，您的环境需要快速适应您的关键业务应用程序无干扰的变化。 ONTAP企业数据管理、服务质量 (QoS) 和性能使您能够规划和适应您的环境。

将NetApp ONTAP与 Confluent 分层存储结合使用，可以利用ONTAP作为横向扩展存储目标，从而简化 Apache Kafka 集群的管理，并支持 Confluent 计算和存储资源的独立扩展。

ONTAP S3 服务器建立在ONTAP成熟的横向扩展存储功能之上。通过扩展 S3 存储桶以使用新添加到ONTAP集群的节点，可以无缝地扩展ONTAP集群。

使用ONTAP系统管理器进行简单管理

ONTAP系统管理器是一个基于浏览器的图形界面，可让您在单一管理平台中配置、管理和监控全球分布位置的ONTAP存储控制器。



您可以使用系统管理器和ONTAP CLI 配置和管理ONTAP S3。当您使用系统管理器启用 S3 并创建存储桶时，ONTAP会提供最佳实践默认设置以简化配置。如果您从 CLI 配置 S3 服务器和存储桶，您仍然可以根据需要使用系统管理器管理它们，反之亦然。

当您使用系统管理器创建 S3 存储桶时，ONTAP会配置系统上可用的最高默认性能服务级别。例如，在AFF系统上，默认设置是 Extreme。性能服务级别是预定义的自适应 QoS 策略组。您可以指定自定义 QoS 策略组或不指定策略组，而不是指定默认服务级别之一。

预定义的自适应 QoS 策略组包括以下内容：

- *极端。*用于需要最低延迟和最高性能的应用程序。
- *表现。*适用于具有中等性能需求和延迟的应用程序。
- *价值。*用于吞吐量和容量比延迟更重要的应用程序。
- *风俗。*指定自定义 QoS 策略或不指定 QoS 策略。

如果选择“用于分层”，则不会选择任何性能服务级别，系统会尝试为分层数据选择具有最佳性能的低成本媒体。

ONTAP尝试在具有最合适磁盘的本地层上配置此存储桶，以满足所选的服务级别。但是，如果您需要指定要包含在存储桶中的磁盘，请考虑通过指定本地层（聚合）从 CLI 配置 S3 对象存储。如果您从 CLI 配置 S3 服务器，您仍然可以根据需要使用系统管理器进行管理。

如果您希望能够指定用于存储桶的聚合，则只能使用 CLI 来实现。

汇合

Confluent 平台是一个全方位的数据流平台，使您能够轻松地以连续的实时流形式访问、存储和管理数据。Confluent 由 Apache Kafka 的原始创建者构建，它通过企业级功能扩展了 Kafka 的优势，同时消除了 Kafka 管理或监控的负担。如今，财富 100 强企业中超过 80% 都采用数据流技术，其中大多数都使用 Confluent。

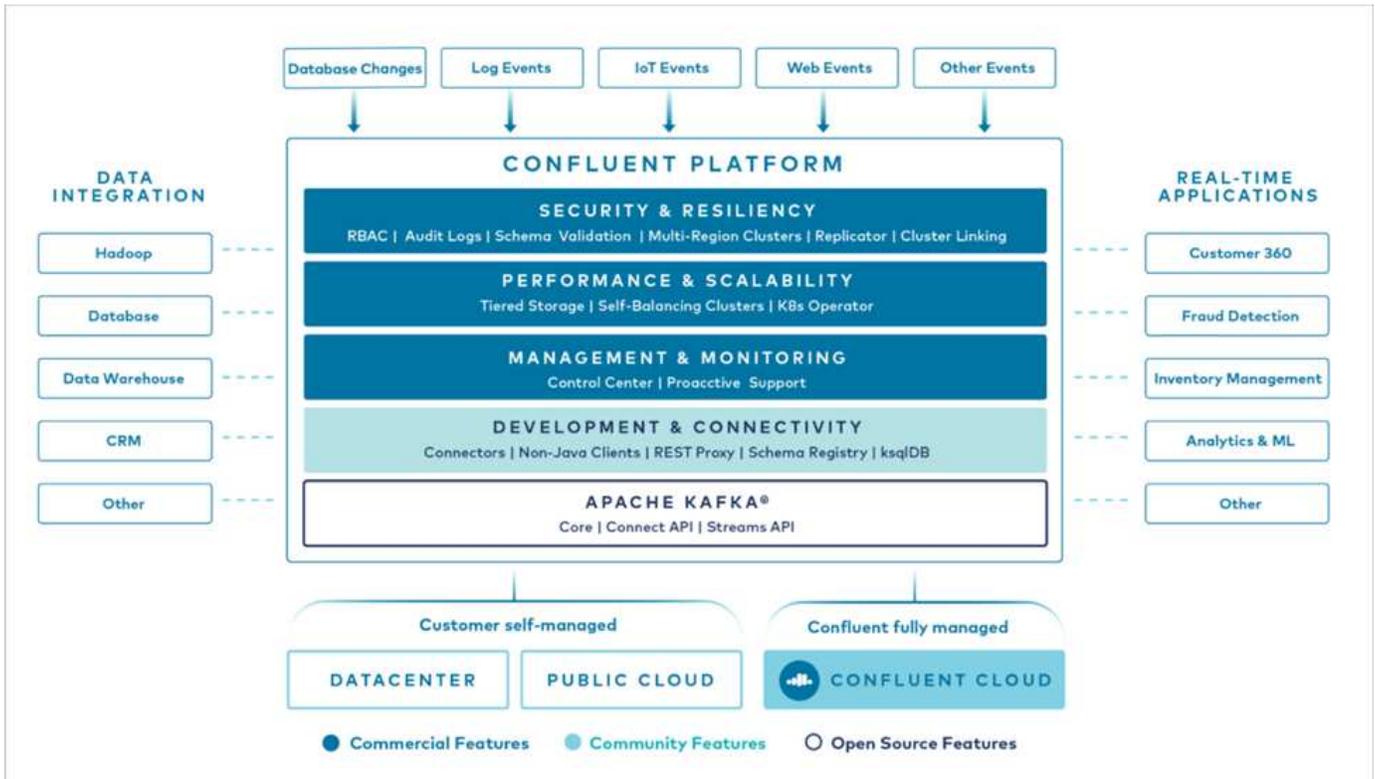
为什么选择 **Confluent**？

通过将历史数据和实时数据集成到单一的中央事实来源，Confluent 可以轻松构建全新类别的现代事件驱动应用程序，获得通用数据管道，并解锁具有完全可扩展性、性能和可靠性的强大新用例。

Confluent 的用途是什么？

Confluent 平台让您专注于如何从数据中获取商业价值，而不必担心底层机制，例如如何在不同的系统之间传输或集成数据。具体来说，Confluent Platform 简化了数据源与 Kafka 的连接、流应用程序的构建以及 Kafka 基础设施的保护、监控和管理。如今，Confluent 平台已广泛应用于众多行业，从金融服务、全渠道零售、自动驾驶汽车到欺诈检测、微服务和物联网。

下图展示了 Confluent Platform 的组件。



Confluent 事件流技术概述

Confluent 平台的核心是 "卡夫卡"，最受欢迎的开源分布式流媒体平台。Kafka 的主要功能包括：

- 发布和订阅记录流。
- 以容错的方式存储记录流。
- 处理记录流。

开箱即用的 Confluent Platform 还包括 Schema Registry、REST Proxy、总共 100 多个预构建的 Kafka 连接器 和 ksqlDB。

Confluent 平台企业功能概述

- *汇合控制中心。*用于管理和监控 Kafka 的基于 UI 的系统。它允许您轻松管理 Kafka Connect 以及创建、编辑和管理与其他系统的连接。
- 适用于 **Kubernetes** 的 **Confluent**。Confluent for Kubernetes 是一个 Kubernetes 操作员。Kubernetes 操作员通过为特定平台应用程序提供独特的功能和要求来扩展 Kubernetes 的编排功能。对于 Confluent 平台，这包括大大简化 Kafka 在 Kubernetes 上的部署过程，并自动执行典型的基础设施生命周期任务。
- Kafka Connect 连接器。连接器使用 Kafka Connect API 将 Kafka 连接到其他系统，例如数据库、键值存储、搜索索引和文件系统。Confluent Hub 具有适用于最流行的数据源和接收器的可下载连接器，包括使用 Confluent Platform 对这些连接器进行全面测试和支持的版本。更多详情请见 "[此处](#)"。
- *自平衡集群。*提供自动负载平衡、故障检测和自我修复。它还支持根据需要添加或停用代理，无需手动调整。
- *汇合簇连接。*直接将集群连接在一起，并通过链接桥将主题从一个集群镜像到另一个集群。集群链接简化了多数据中心、多集群和混合云部署的设置。
- *汇合自动数据平衡器。*监控集群中的代理数量、分区大小、分区数量以及集群内的领导者数量。它允许您

转移数据以在整个集群中创建均匀的工作负载，同时限制重新平衡流量以最大限度地减少重新平衡时对生产工作负载的影响。

- ***汇合复制器。***使得在多个数据中心维护多个 Kafka 集群变得比以往更加简单。
- ***分层存储。***提供使用您最喜欢的云提供商存储大量 Kafka 数据的选项，从而减少运营负担和成本。通过分层存储，您可以将数据保存在经济高效的对象存储中，并且仅在需要更多计算资源时才扩展代理。
- **Confluent JMS 客户端。** Confluent Platform 包含一个与 JMS 兼容的 Kafka 客户端。该 Kafka 客户端实现了 JMS 1.1 标准 API，使用 Kafka 代理作为后端。如果您有使用 JMS 的遗留应用程序并且想要用 Kafka 替换现有的 JMS 消息代理，这将非常有用。
- ***Confluent MQTT 代理。***提供一种从 MQTT 设备和网关直接向 Kafka 发布数据的方法，无需中间的 MQTT 代理。
- **Confluent 安全插件。** Confluent 安全插件用于为各种 Confluent 平台工具和产品添加安全功能。目前，有一个可用于 Confluent REST 代理的插件，可帮助验证传入的请求并将经过验证的主体传播到对 Kafka 的请求。这使得 Confluent REST 代理客户端能够利用 Kafka 代理的多租户安全功能。

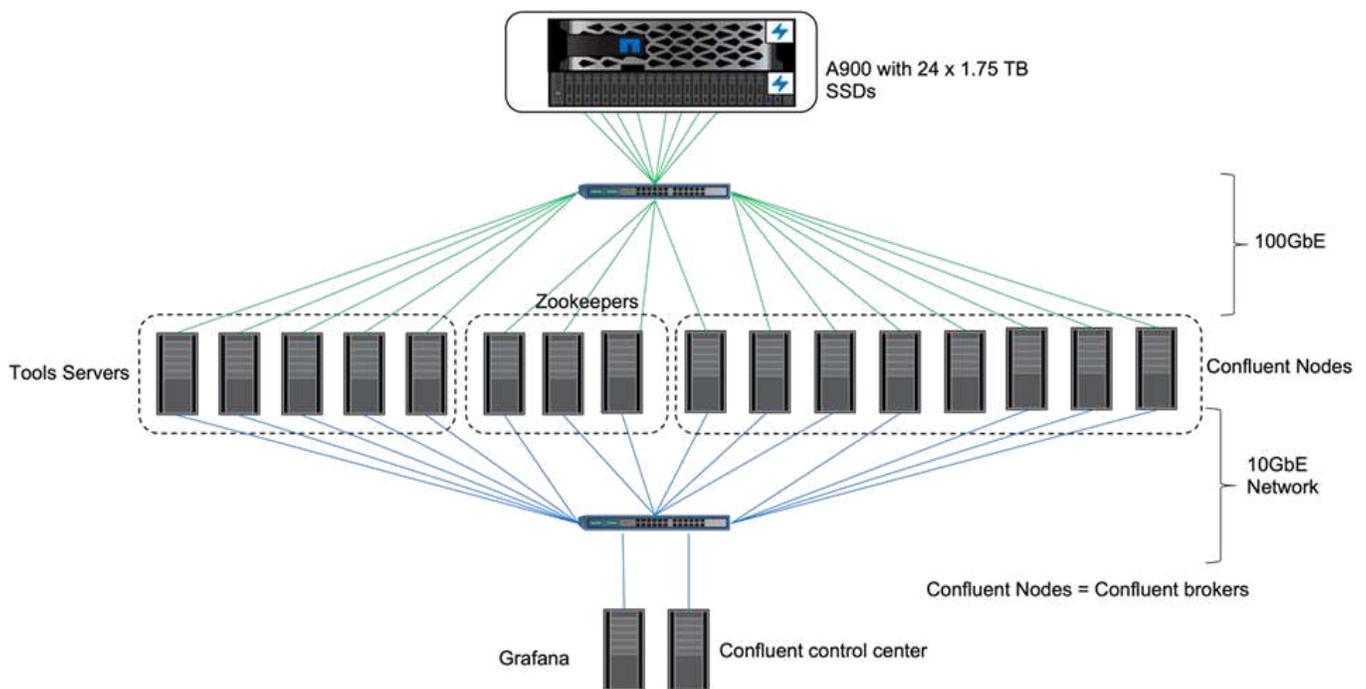
Confluent 性能验证

我们已经使用 Confluent Platform 对 NetApp ONTAP 上的分层存储进行了验证。NetApp 和 Confluent 团队共同进行了此次验证，并运行了所需的测试用例。

Confluent 设置

在设置中，我们使用了三个 Zookeeper、五个代理和五个测试服务器，配备 256GB RAM 和 16 个 CPU。对于 NetApp 存储，我们使用了带有 AFF A900 HA 对的 ONTAP。存储和代理通过 100GbE 连接进行连接。

下图为分层存储验证配置的网络拓扑图。



工具服务器充当应用程序客户端，向 Confluent 节点发送事件或从 Confluent 节点接收事件。

Confluent 分层存储配置

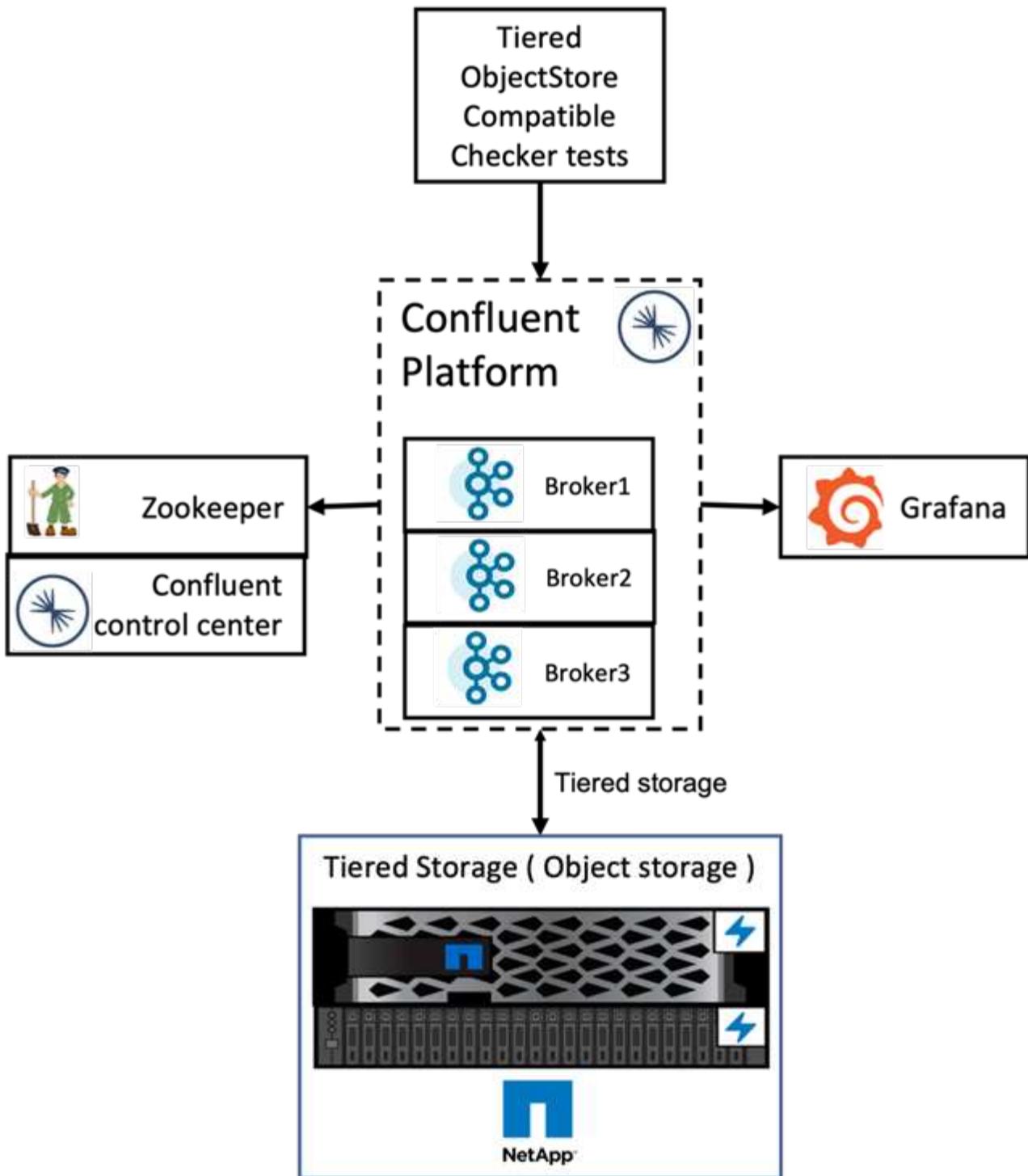
我们使用了以下测试参数：

```
confluent.tier.fetcher.num.threads=80
confluent.tier.archiver.num.threads=80
confluent.tier.enable=true
confluent.tier.feature=true
confluent.tier.backend=S3
confluent.tier.s3.bucket=kafkabucket1-1
confluent.tier.s3.region=us-east-1
confluent.tier.s3.cred.file.path=/data/kafka/.ssh/credentials
confluent.tier.s3.aws.endpoint.override=http://wle-mendocino-07-08/
confluent.tier.s3.force.path.style.access=true
bootstrap.server=192.168.150.172:9092,192.168.150.120:9092,192.168.150.164:9092,192.168.150.198:9092,192.168.150.109:9092,192.168.150.165:9092,192.168.150.119:9092,192.168.150.133:9092
debug=true
jmx.port=7203
num.partitions=80
num.records=200000000
#object PUT size - 512MB and fetch 100MB - netapp
segment.bytes=536870912
max.partition.fetch.bytes=1048576000
#GET size is max.partition.fetch.bytes/num.partitions
length.key.value=2048
trogdor.agent.nodes=node0,node1,node2,node3,node4
trogdor.coordinator.hostname.port=192.168.150.155:8889
num.producers=20
num.head.consumers=20
num.tail.consumers=1
test.binary.task.max.heap.size=32G
test.binary.task.timeout.sec=3600
producer.timeout.sec=3600
consumer.timeout.sec=3600
```

为了验证，我们使用了带有 HTTP 协议的 ONTAP，但 HTTPS 也可以运行。访问密钥和密钥存储在 `confluent.tier.s3.cred.file.path` 范围。

NetApp 存储控制器 – ONTAP

我们在 ONTAP 中配置了单个 HA 对配置以进行验证。



验证结果

我们完成了以下五个测试用例进行验证。前两个是功能测试，其余三个是性能测试。

对象存储正确性测试

此测试使用 API 调用对用于分层存储的对象存储执行获取、放置和删除等基本操作。

分层功能正确性测试

此测试检查对象存储的端到端功能。它创建一个主题，为新创建的主题生成一个事件流，等待代理将段存档到对象存储，使用事件流，并验证使用的流是否与生成的流匹配。我们已经在有和没有对象存储故障注入的情况下执行了此测试。我们通过停止ONTAP中某个节点的服务管理器服务来模拟节点故障，并验证端到端功能是否与对象存储兼容。

层级获取基准

该测试验证了分层对象存储的读取性能，并检查了基准测试生成的段在高负载下的范围提取读取请求。在这个基准测试中，Confluent 开发了自定义客户端来满足层级获取请求。

生产-消费工作负载生成器

该测试通过段的归档间接在对象存储上生成写入工作负载。当消费者群体获取段时，从对象存储中生成读取工作负载（读取的段）。此工作负载由 TOCC 脚本生成。该测试检查了并行线程对对象存储的读写性能。我们对分层功能正确性测试进行了测试，测试了有和没有对象存储故障注入的情况。

保留工作量生成器

该测试检查了对象存储在高主题保留工作负载下的删除性能。保留工作负载是使用 TOCC 脚本生成的，该脚本与测试主题并行生成许多消息。测试主题是使用基于大小和基于时间的激进保留设置进行配置，这会导致事件流不断从对象存储中清除。然后将这些片段存档。这导致代理在对象存储中进行多次删除，并收集对象存储删除操作的性能。

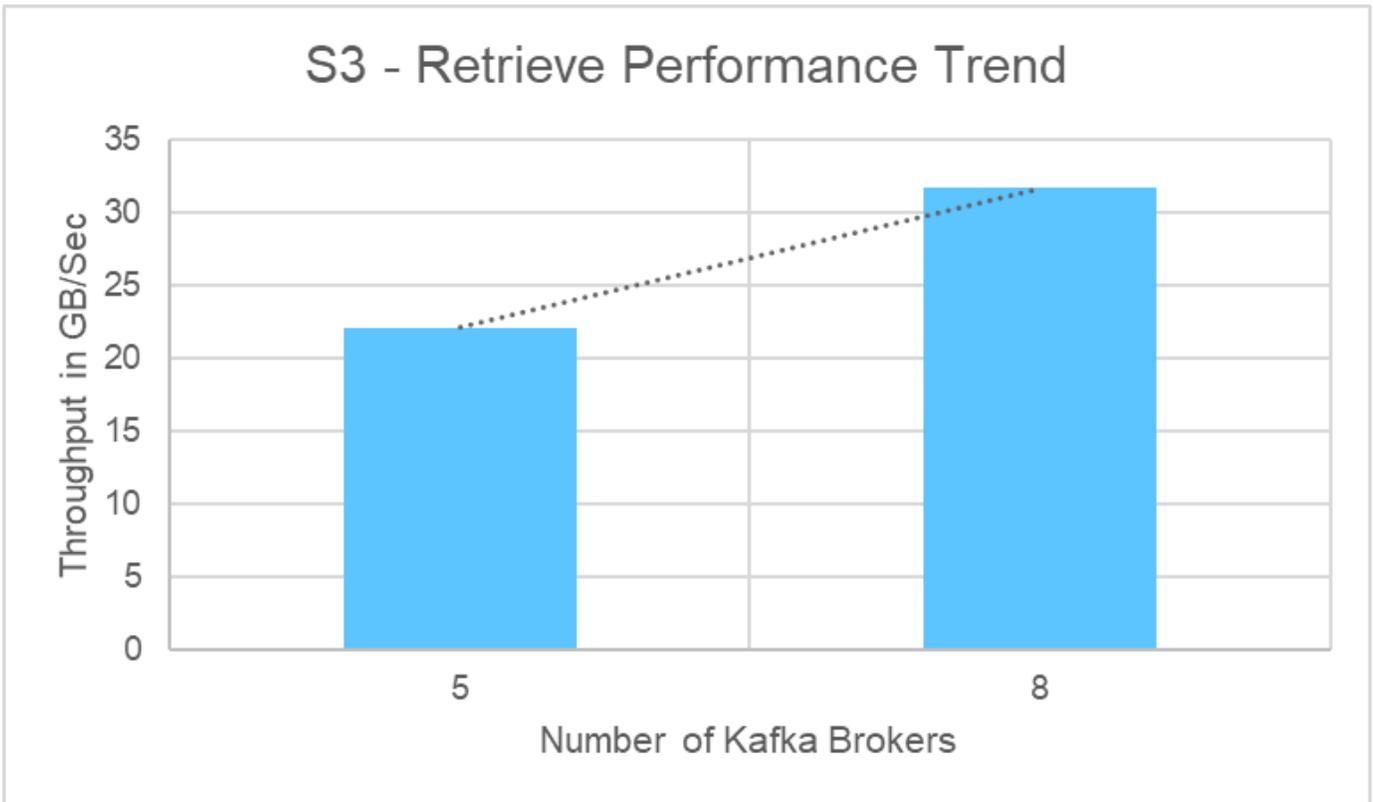
有关验证详细信息，请参阅 ["汇合"网站](#)。

使用生产-消费工作负载生成器进行性能测试

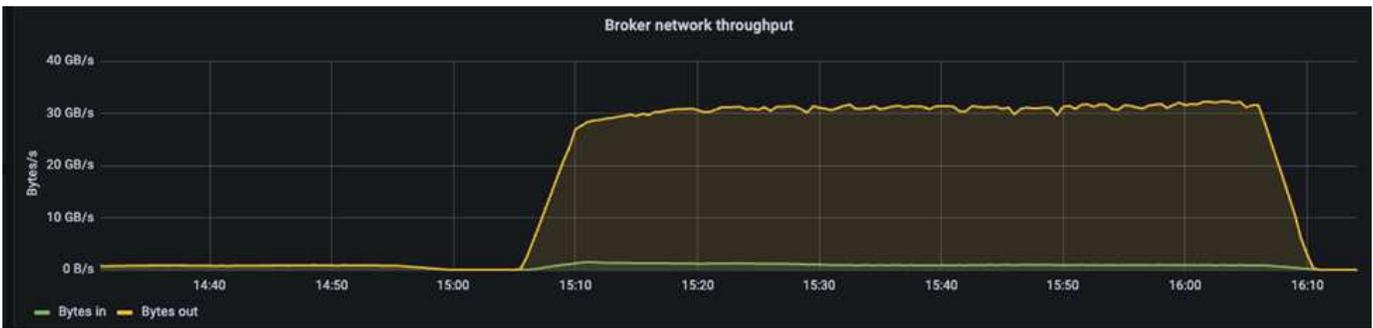
我们在生产-消费工作负载期间使用一个AFF A900 HA 对NetApp存储控制器对五个或八个代理节点执行了分层存储测试。根据我们的测试，完成时间和性能结果随着代理节点的数量而变化，直到AFF A900资源利用率达到百分之百。ONTAP存储控制器设置至少需要一个HA对。

S3 检索操作的性能随着 Confluent 代理节点的数量而线性增加。ONTAP存储控制器在单个部署中最多支持 12 个HA对。

下图显示了具有五个或八个代理节点的组合 S3 分层流量。我们最大限度地提高了AFF A900单 HA 对的性能。



下图显示 Kafka 吞吐量约为 31.74GBps。



我们还观察到ONTAP存储控制器的吞吐量类似 `perfstat` 报告。

```
object_store_server:wle-mendocino-07-08:get_data:34080805907b/ s
object_store_server:wle-mendocino-07-08:put_data:484236974b/ s
```

性能最佳实践指南

本页介绍了提高此解决方案性能的最佳实践。

- 对于ONTAP，如果可能，请使用 $\geq 1\text{MB}$ 的 GET 大小。
- 增加 `num.network.threads` 和 `num.io.threads` 在 `server.properties` 在代理节点上，您可以将增加的分层活动推送到 S3 层。这些结果与 `num.network.threads` 和 `num.io.threads` 设置为 32。

- S3 存储桶应该针对每个成员聚合的八个组成部分。
- 驱动 S3 流量的以太网链路在存储和客户端上都应尽可能使用 9k 的 MTU。

结束语

此次验证测试在配备NetApp ONTAP存储控制器的 Confluent 上达到了 31.74GBps 的分层吞吐量。

在哪里可以找到更多信息

要了解有关本文档中描述的信息的更多信息，请查看以下文档和/或网站：

- 什么是 Confluent?

["https://www.confluent.io/apache-kafka-vs-confluent/"](https://www.confluent.io/apache-kafka-vs-confluent/)

- S3-sink 参数详情

["https://docs.confluent.io/kafka-connect-s3-sink/current/configuration_options.html#s3-configuration-options"](https://docs.confluent.io/kafka-connect-s3-sink/current/configuration_options.html#s3-configuration-options)

- 阿帕奇卡夫卡

["https://en.wikipedia.org/wiki/Apache_Kafka"](https://en.wikipedia.org/wiki/Apache_Kafka)

- ONTAP中的 S3 最佳实践

<https://www.netapp.com/pdf.html?item=/media/17219-tr4814.pdf>

- S3 对象存储管理

["https://docs.netapp.com/us-en/ontap/s3-config/s3-support-concept.html"](https://docs.netapp.com/us-en/ontap/s3-config/s3-support-concept.html)

- NetApp产品文档

["https://www.netapp.com/support-and-training/documentation/"](https://www.netapp.com/support-and-training/documentation/)

版权信息

版权所有 © 2025 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。