



Confluent Kafka mit NetApp ONTAP Storage Controllern

NetApp Solutions

NetApp
May 14, 2024

Inhalt

- Confluent Kafka mit NetApp ONTAP Storage Controllern 1
 - TR-4941: Fließend mit den NetApp ONTAP Storage Controllern 1
 - Nutzen 1
 - Technologieüberblick 3
 - Validierung der Performance in einem fließenden Verhältnis 8
 - Performance-Tests mit dem Produce-Consume Workload Generator 12
 - Leitfaden zu Best Practices für die Performance 13
 - Schlussfolgerung 14

Confluent Kafka mit NetApp ONTAP Storage Controllern

TR-4941: Fließend mit den NetApp ONTAP Storage Controllern

Karthikeyan Nagalingam, Joe Scott, NetApp Rankesh Kumar, Confluent

Damit die Confluent Plattform skalierbarer und flexibler wird, muss sie in der Lage sein, Workloads sehr schnell zu skalieren und auszugleichen. Durch Tiered Storage können große Datenmengen in Confluent gemanagt werden, indem diese Betriebskosten reduziert werden.

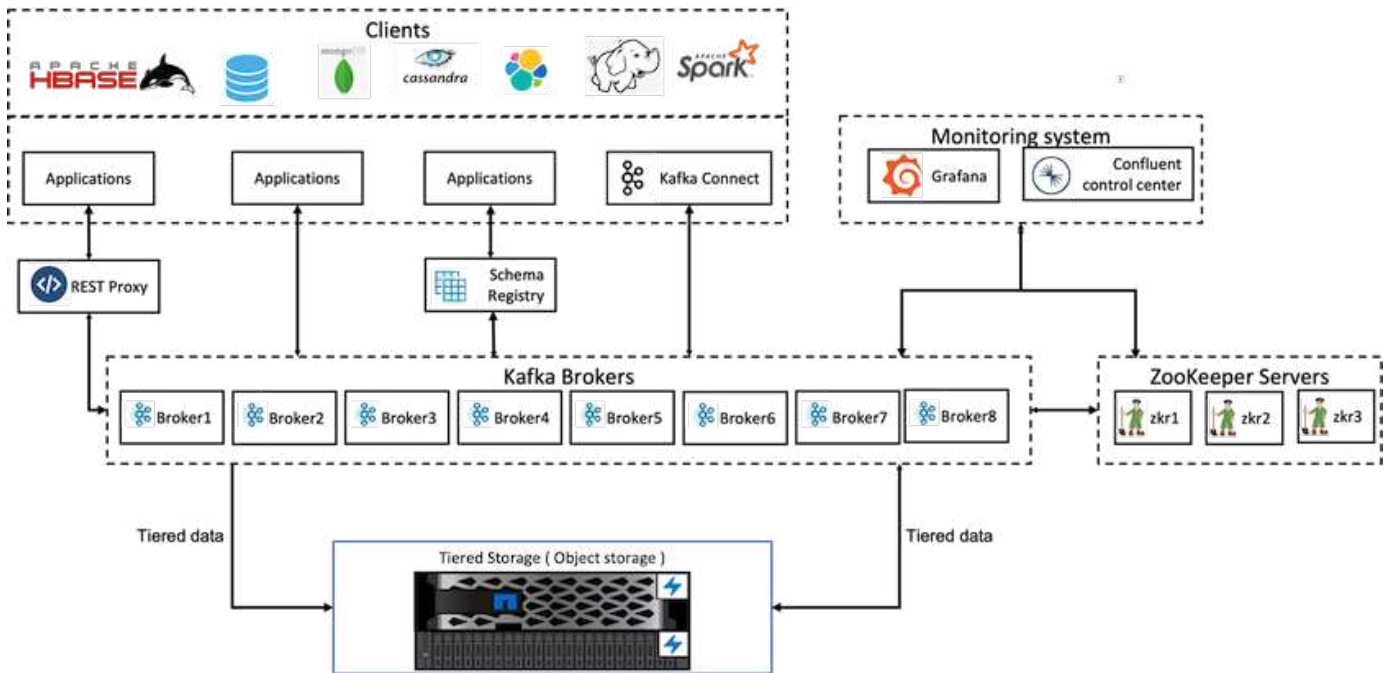
Eine wesentliche Idee besteht darin, den Storage von der Datenverarbeitung zu trennen, wodurch jede unabhängig voneinander skaliert werden kann.

Mit zahlreichen branchenführenden Innovationen vereint die Datenmanagement-Software NetApp ONTAP vielfältige Vorteile in jedem Bereich der Daten.

Dieses Dokument beschreibt die Performance-Benchmarks für die Confluent Plattform auf NetApp ONTAP mithilfe eines Tiered Storage Benchmark Kits.

Nutzen

Der AFF A900 Storage Controller von ONTAP, der zusammen mit dem fließenden Speicher betrieben wird, sind verteilte Systeme für Datenströme. Beide Systeme sind horizontal skalierbar und fehlertolerant und bieten unter Last eine hervorragende Performance. Sie ergänzen sich gegenseitig in der verteilten Datenstreaming und der Stream-Verarbeitung und senken dabei die Storage-Kosten mithilfe von Datenreduzierungstechnologien, die den Platzbedarf von Daten minimieren. Der AFF A900 Storage-Controller bietet eine hervorragende Performance und ermöglicht gleichzeitig die Abkopplung von Computing- und Storage-Ressourcen. Dies vereinfacht die Systemadministration und ermöglicht die unabhängige Skalierung von Ressourcen.



Details zur Lösungsarchitektur

Dieser Abschnitt behandelt die Hardware und Software, die für die Performance-Verifizierung in der Confluent Platform Implementierung mit NetApp ONTAP für Tiered Storage verwendet werden. In der folgenden Tabelle werden die Lösungsarchitektur und die Basiskomponenten behandelt.

Plattformkomponente	Umgebungskonfiguration
Confluent Platform Version 6.2	<ul style="list-style-type: none"> • 3 x Zookeeper • 8 x Broker Server • 5-mal Tools-Server • 1 x Grafana • 1 x Kontrollzentrum
Betriebssystem auf allen Knoten	Linux (ubuntu 18.04)
NetApp ONTAP für Buckets	<ul style="list-style-type: none"> • 1 AFF A900 HA-Paar (High Availability, Hochverfügbarkeit) • 4 x 24 x 800 SSDs • S3-Protokoll • 100 GbE
15 Fujitsu PRIMERGY RX2540 Server	<ul style="list-style-type: none"> • 2 CPUs; 16 physische Kerne insgesamt • Intel Xeon • 256 GB physischer Speicher • Dual-Port mit 100 GbE

Technologieüberblick

In diesem Abschnitt wird die in dieser Lösung verwendete Technologie beschrieben.

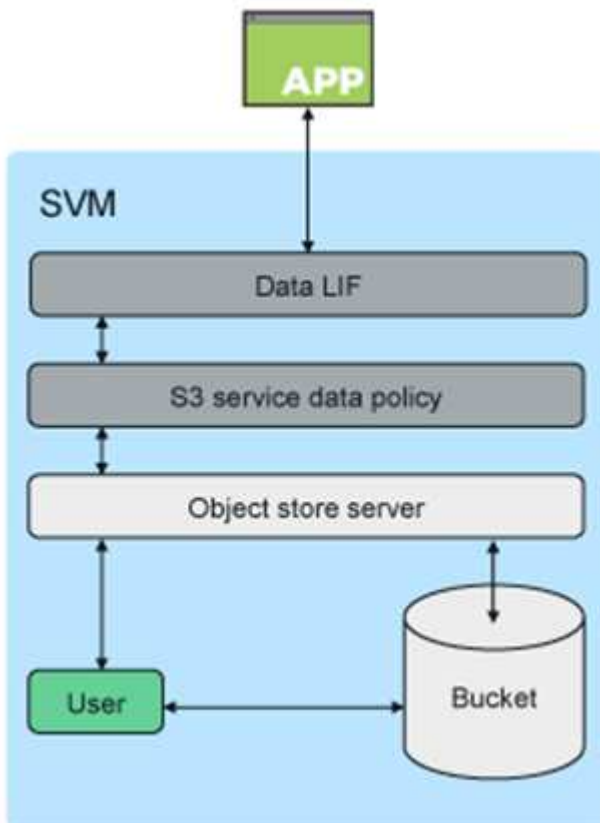
NetApp ONTAP Storage Controller

NetApp ONTAP ist ein hochperformantes Storage-Betriebssystem der Enterprise-Klasse.

NetApp ONTAP 9.8 unterstützt die Amazon Simple Storage Service (S3) APIs. ONTAP unterstützt einen Teilbereich der S3-API-Aktionen von Amazon Web Services (AWS) und ermöglicht die Darstellung von Daten in ONTAP-basierten Systemen über Cloud-Provider (AWS, Azure und GCP) und On-Premises hinweg.

NetApp StorageGRID Software ist die Flaggschiff der NetApp Lösung für Objekt-Storage. ONTAP ergänzt StorageGRID durch die Bereitstellung eines On-the-Edge-Aufnahmepunkts und einen Vorverarbeitungspunkt auf der Basis von NetApp für Objektdaten, eine Erweiterung der Data Fabric Strategie bei der Erweiterung des NetApp Produktportfolios und eine Steigerung des Mehrwerts des NetApp Produktportfolios.

Der Zugriff auf einen S3-Bucket wird über autorisierte Benutzer- und Client-Applikationen bereitgestellt. Im folgenden Diagramm ist die Applikation, die auf einen S3-Bucket zugreift, dargestellt.



Primäre Anwendungsfälle

Durch die Unterstützung von S3-APIs soll primär der Zugriff auf Objekte auf ONTAP bereitgestellt werden. Die Unified Storage-Architektur von ONTAP unterstützt jetzt Dateien (NFS und SMB), Blöcke (FC und iSCSI) und Objekte (S3).

Native S3-Applikationen

Eine wachsende Anzahl von Applikationen kann ONTAP-Unterstützung für Objektzugriff über S3 nutzen. Obwohl sich der Bedarf an hoher Performance in nativen S3-Applikationen für Archivierungs-Workloads eignet, wächst auch hier schnell der Bedarf an:

- Analysen
- Künstliche Intelligenz
- Edge-to-Core-Aufnahme
- Machine Learning

Kunden können jetzt vertraute Management-Tools wie ONTAP System Manager verwenden, um hochperformanten Objekt-Storage für Entwicklung und Betrieb in ONTAP bereitzustellen und so die Effizienz und Sicherheit von ONTAP Storage zu nutzen.

FabricPool-Endpunkte

Ab ONTAP 9.8 unterstützt FabricPool Tiering für Buckets in ONTAP, sodass auch ONTAP-zu-ONTAP-Tiering genutzt werden kann. Dies ist eine hervorragende Option für Kunden, die eine vorhandene FAS-Infrastruktur als Objektspeicher-Endpunkt nutzen möchten.

FabricPool unterstützt Tiering zu ONTAP auf zwei Arten:

- **Lokales Cluster Tiering.** inaktive Daten werden mithilfe von Cluster LIFs zu einem Bucket auf dem lokalen Cluster verschoben.
- **Remote Cluster Tiering.** inaktive Daten werden auf einem Bucket auf einem Remote-Cluster ähnlich wie auf einem herkömmlichen FabricPool Cloud Tier mit IC LIFs auf dem FabricPool Client und Daten-LIFs auf dem ONTAP Objektspeicher verschoben.

ONTAP S3 ist die richtige Lösung, wenn Sie S3-Funktionen auf vorhandenen Clustern ohne zusätzliche Hardware und Management wünschen. Für Implementierungen mit über 300 TB ist NetApp StorageGRID immer noch die Vorzeige-NetApp Lösung für Objekt-Storage. Bei der Nutzung von ONTAP oder StorageGRID als Cloud-Tier ist keine FabricPool Lizenz erforderlich.

NetApp ONTAP für Confluent Tiered Storage

Jedes Datacenter muss geschäftskritische Applikationen unterbrechungsfrei laufen lassen und wichtige Daten verfügbar und sicher halten. Das neue NetApp AFF A900 System wird durch ONTAP Enterprise Edition Software und einem äußerst Resilience Design unterstützt. Unser neues, blitzschnelles NVMe-Storage-System beseitigt Unterbrechungen geschäftskritischer Prozesse, minimiert das Performance-Tuning und schützt Ihre Daten vor Ransomware-Angriffen.

Von der ersten Implementierung bis zur Skalierung Ihres Confluent Clusters muss Ihre Umgebung schnell an Änderungen angepasst werden, die unterbrechungsfrei auf Ihre geschäftskritischen Applikationen abgestimmt sind. Mit ONTAP Lösungen für Enterprise-Datenmanagement, Quality of Service (QoS) und Performance können Unternehmen ihre Umgebung planen und anpassen.

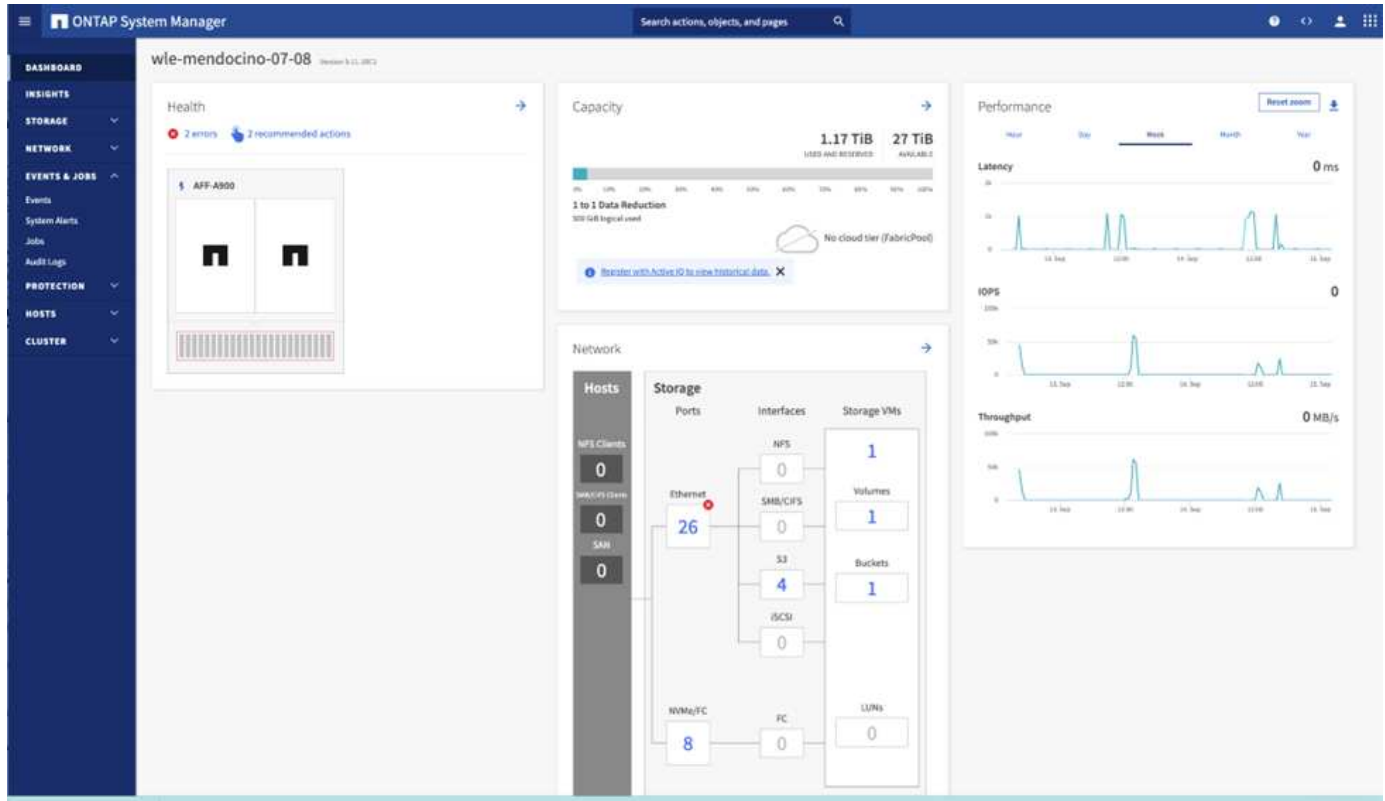
Der Einsatz von NetApp ONTAP und Confluent Tiered Storage vereinfacht zusammen das Management von Apache Kafka Clustern durch die Nutzung von ONTAP als Scale-out Storage-Ziel und ermöglicht eine unabhängige Skalierung von Computing- und Storage-Ressourcen für Confluent.

Ein ONTAP S3 Server baut auf den ausgereiften Scale-out-Storage-Funktionen von ONTAP auf. Die ONTAP Cluster lassen sich nahtlos skalieren, indem die S3-Buckets erweitert werden, um neu hinzugefügte Nodes im

ONTAP Cluster zu verwenden.

Einfaches Management mit ONTAP System Manager

Bei ONTAP System Manager handelt es sich um eine browserbasierte grafische Oberfläche, mit der Sie Ihren ONTAP Storage Controller an weltweit verteilten Standorten über eine zentrale Konsole konfigurieren, managen und überwachen können.



ONTAP S3 lässt sich mit System Manager und der ONTAP CLI konfigurieren und verwalten. Wenn Sie S3 aktivieren und Buckets mithilfe von System Manager erstellen, bietet ONTAP Best-Practice-Standards für eine vereinfachte Konfiguration. Wenn Sie den S3-Server und die Buckets aus der CLI konfigurieren, können Sie sie nach Bedarf auch mit System Manager managen oder umgekehrt.

Wenn Sie mit System Manager einen S3-Bucket erstellen, konfiguriert ONTAP ein Service-Level für die Standard-Performance, das auf Ihrem System am höchsten verfügbar ist. Bei einem AFF-System wäre die Standardeinstellung „Extreme“. Performance-Service-Level sind vordefinierte, anpassungsfähige QoS-Richtliniengruppen. Anstelle eines der Standard-Service-Level können Sie eine benutzerdefinierte QoS-Richtliniengruppe oder keine Richtliniengruppe angeben.

Zu den vordefinierten anpassungsfähigen QoS-Richtliniengruppen gehören:

- **Extreme.** wird für Anwendungen verwendet, die die niedrigste Latenz und höchste Performance benötigen.
- **Performance.** wird für Applikationen mit geringen Performance-Anforderungen und Latenz verwendet.
- **Wert** wird für Applikationen verwendet, bei denen Durchsatz und Kapazität wichtiger sind als die Latenz.
- **Benutzerdefiniert.** Geben Sie eine benutzerdefinierte QoS-Richtlinie oder keine QoS-Richtlinie an.

Wenn Sie **für Tiering** verwenden auswählen, werden keine Leistungsservicelevel ausgewählt und das System versucht, kostengünstige Medien mit optimaler Leistung für die Tiered Data auszuwählen.

ONTAP versucht, diesen Bucket auf lokalen Tiers bereitzustellen, die über die am besten geeigneten Festplatten verfügen und dem ausgewählten Service-Level gerecht werden. Wenn Sie jedoch angeben müssen, welche Festplatten in den Bucket enthalten sind, sollten Sie S3-Objekt-Storage aus der CLI konfigurieren, indem Sie die lokalen Tiers (Aggregat) angeben. Wenn Sie den S3-Server über die CLI konfigurieren, können Sie ihn bei Bedarf weiterhin mit System Manager managen.

Wenn Sie angeben können, welche Aggregate für Buckets verwendet werden, können Sie dies nur über die CLI tun.

Fließend

Confluent Platform ist eine umfassende Daten-Streaming-Plattform, auf die Sie als kontinuierliche Echtzeit-Streams auf einfache Weise auf Daten zugreifen, diese speichern und managen können. Confluent wurde von den ursprünglichen Schöpfern von Apache Kafka erbaut und erweitert die Vorteile von Kafka mit Funktionen der Enterprise-Klasse, ohne Kafka-Management oder -Monitoring zu belasten. Heute basieren mehr als 80 % der Fortune 100 auf Data Streaming-Technologie, und die meisten nutzen Confluent.

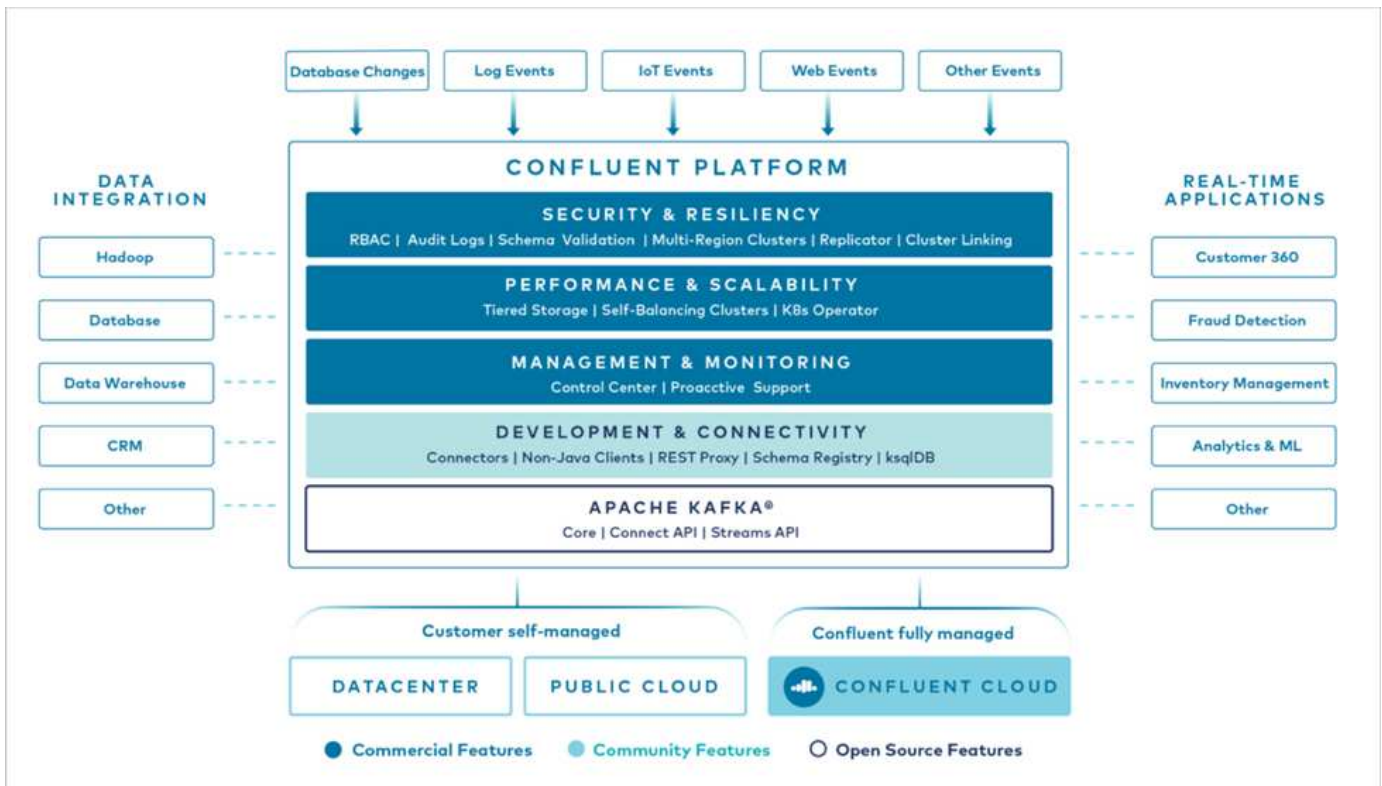
Warum Confluent?

Durch die Integration von historischen und Echtzeit-Daten in eine einzige, zentrale Quelle der Wahrheit erleichtert Confluent den Aufbau einer völlig neuen Kategorie moderner, ereignisgesteuerter Anwendungen, die Erstellung einer universellen Datenpipeline und die Nutzung leistungsstarker neuer Anwendungsfälle mit voller Skalierbarkeit, Leistung und Zuverlässigkeit.

Wofür wird Confluent verwendet?

Mit der Confluent Platform können Sie sich darauf konzentrieren, wie Sie aus Ihren Daten einen geschäftlichen Nutzen ziehen können, statt sich um die zugrunde liegenden Mechanismen sorgen zu müssen, wie beispielsweise der Transport oder die Integration von Daten zwischen verschiedenen Systemen. Confluent Platform vereinfacht insbesondere die Anbindung von Datenquellen an Kafka, die Erstellung von Streaming-Applikationen sowie die Sicherung, Überwachung und das Management der Kafka Infrastruktur. Heute kommt die Confluent Platform für eine Vielzahl von Anwendungsfällen in zahlreichen Branchen zum Einsatz, von Finanzdienstleistungen über Omnichannel-Einzelhandel, autonomen Fahrzeugen bis hin zu Betrugserkennung, Microservices und IoT.

Die folgende Abbildung zeigt die Komponenten der Confluent Platform.



Überblick über die Confluent Event Streaming Technologie

Der Kern der Confluent Plattform ist "Kafka", Die beliebteste verteilte Open Source Streaming-Plattform. Kafka bietet u. a. folgende zentrale Funktionen:

- Veröffentlichen und abonnieren Sie Datenströme.
- Fehlertolerante Speicherung von Datenströmen
- Verarbeiten von Datensätzen.

Die Confluent Plattform umfasst außerdem Schema Registry, REST Proxy, insgesamt 100+ vordefinierte Kafka-Anschlüsse und ksqlDB.

Überblick über die Enterprise-Funktionen der Confluent Plattform

- **Confluent Control Center.** ein UI-basiertes System zur Verwaltung und Überwachung von Kafka. Damit können Sie Kafka Connect ganz einfach verwalten und Verbindungen zu anderen Systemen erstellen, bearbeiten und verwalten.
- **Fließend für Kubernetes.** der fließende für Kubernetes ist ein Kubernetes Operator. Kubernetes-Betreiber erweitern die Orchestrierungsfunktionen von Kubernetes um spezielle Funktionen und Anforderungen für eine spezifische Plattform-Applikation. Bei Confluent Plattform müssen dazu die Implementierung von Kafka auf Kubernetes erheblich vereinfacht und typische Aufgaben im Infrastruktur-Lebenszyklus automatisiert werden.
- **Kafka Connect Connectors.** Steckverbinder verbinden Kafka Connect mit anderen Systemen wie Datenbanken, Schlüsselwertspeicher, Suchindizes und Dateisystemen. Confluent Hub verfügt über herunterladbare Anschlüsse für die beliebtesten Datenquellen und Waschbecken, einschließlich vollständig getestete und unterstützte Versionen dieser Anschlüsse mit Confluent Plattform. Weitere Details finden Sie hier "[Hier](#)".
- **Self-Balancing Cluster.** bietet automatisches Load Balancing, Fehlererkennung und Selbstheilung. Auch

das Hinzufügen oder Dekommissionierung von Vermittlern nach Bedarf ohne manuelles Tuning ist möglich.

- **Fließende Cluster-Verknüpfung.** verbindet Cluster direkt miteinander und spiegelt Themen von einem Cluster zum anderen über eine Link-Bridge. Die Cluster-Verknüpfung vereinfacht die Einrichtung von Implementierungen mit mehreren Rechenzentren, mehreren Clustern und Hybrid Clouds.
- **Confluent Auto Data Balancer.** überwacht Ihren Cluster für die Anzahl der Broker, die Größe der Partitionen, die Anzahl der Partitionen und die Anzahl der Führer innerhalb des Clusters. Auf diese Weise können Sie Daten verschieben, um einen geraden Workload über Ihr Cluster zu erstellen, und gleichzeitig den Datenverkehr neu verteilen, um die Auswirkungen auf die Produktions-Workloads bei der Ausbalancierung zu minimieren.
- **Confluent Replikator.** macht es einfacher als je zuvor, mehrere Kafka Cluster in mehreren Rechenzentren zu pflegen.
- **Tiered Storage.** bietet Optionen zur Speicherung großer Kafka-Datenmengen mit Ihrem bevorzugten Cloud-Provider und reduziert so die Betriebskosten und die Kosten. Mit Tiered Storage können Sie Daten auf kostengünstigem Objekt-Storage und Vermittlern nur dann aufbewahren, wenn Sie mehr Computing-Ressourcen benötigen.
- **Confluent JMS Client.** Confluent Platform enthält einen JMS-kompatiblen Client für Kafka. Dieser Kafka-Client implementiert die JMS 1.1 Standard-API und verwendet Kafka-Broker als Backend. Dies ist nützlich, wenn vorhandene Anwendungen JMS verwenden und Sie den vorhandenen JMS-Nachrichten-Broker durch Kafka ersetzen möchten.
- **Confluent MQTT Proxy.** bietet eine Möglichkeit, Daten direkt an Kafka von MQTT-Geräten und Gateways zu veröffentlichen, ohne dass ein MQTT-Broker in der Mitte nötig ist.
- **Confluent Security Plugins.** Confluent Security Plugins werden verwendet, um Sicherheitsfunktionen zu verschiedenen Tools und Produkten der Confluent Platform hinzuzufügen. Derzeit gibt es ein Plugin für den Confluent REST Proxy, das hilft, die eingehenden Anfragen zu authentifizieren und den authentifizierten Principal an Anfragen an Kafka zu verbreiten. Auf diese Weise können Confluent REST Proxy-Clients die mandantenfähigen Sicherheitsfunktionen des Kafka-Brokers nutzen.

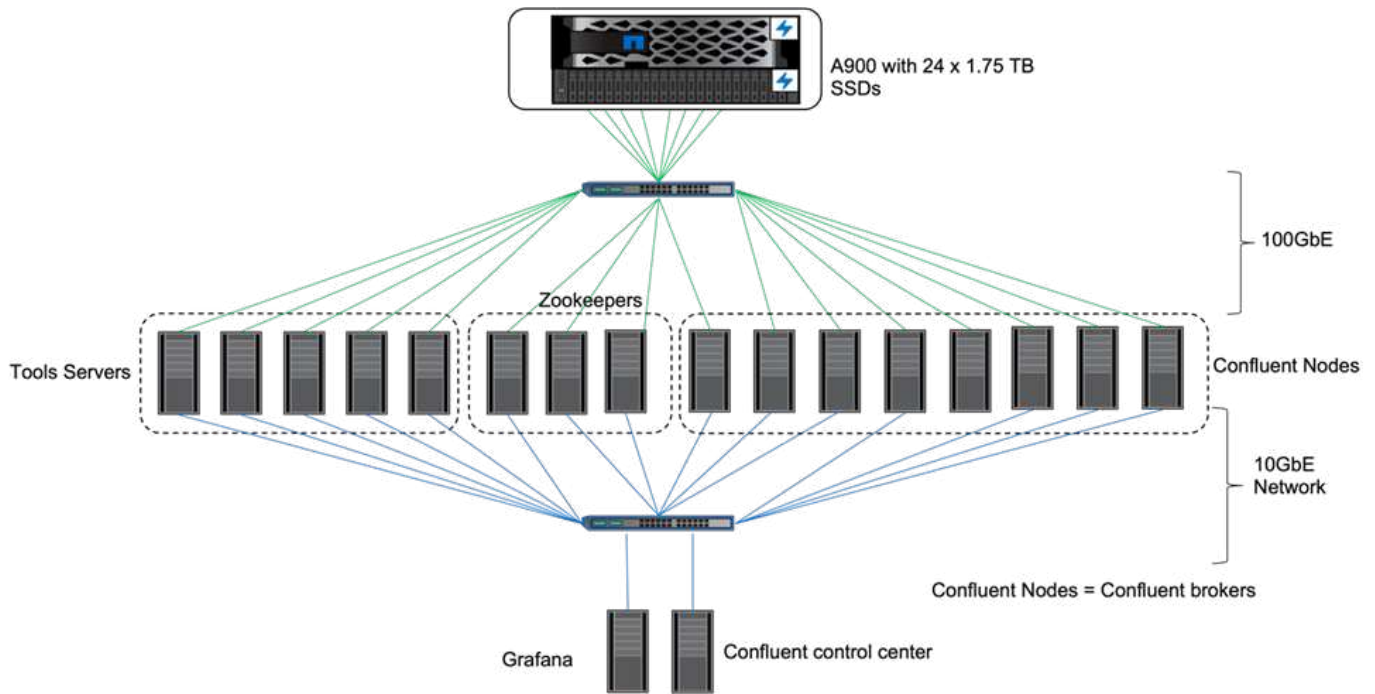
Validierung der Performance in einem fließenden Verhältnis

Wir haben die Verifizierung mit der Confluent Plattform für Tiered Storage auf NetApp ONTAP durchgeführt. Die NetApp und Confluent Teams haben diese Verifizierung gemeinsam bearbeitet und die erforderlichen Testfälle ausgeführt.

Fließende Einrichtung

Für das Setup wurden drei Zookeeper, fünf Broker und fünf Testserver mit 256 GB RAM und 16 CPUs verwendet. Für NetApp Storage haben wir ONTAP mit einem AFF A900 HA Paar verwendet. Der Storage und die Broker waren über 100-GbE-Verbindungen verbunden.

Die folgende Abbildung zeigt die Netzwerktopologie der Konfiguration für die Tiered Storage-Verifizierung.



Die Tools-Server fungieren als Anwendungsclients, die Ereignisse an oder von Confluent Nodes senden oder empfangen.

Fließende Tiered Storage-Konfiguration

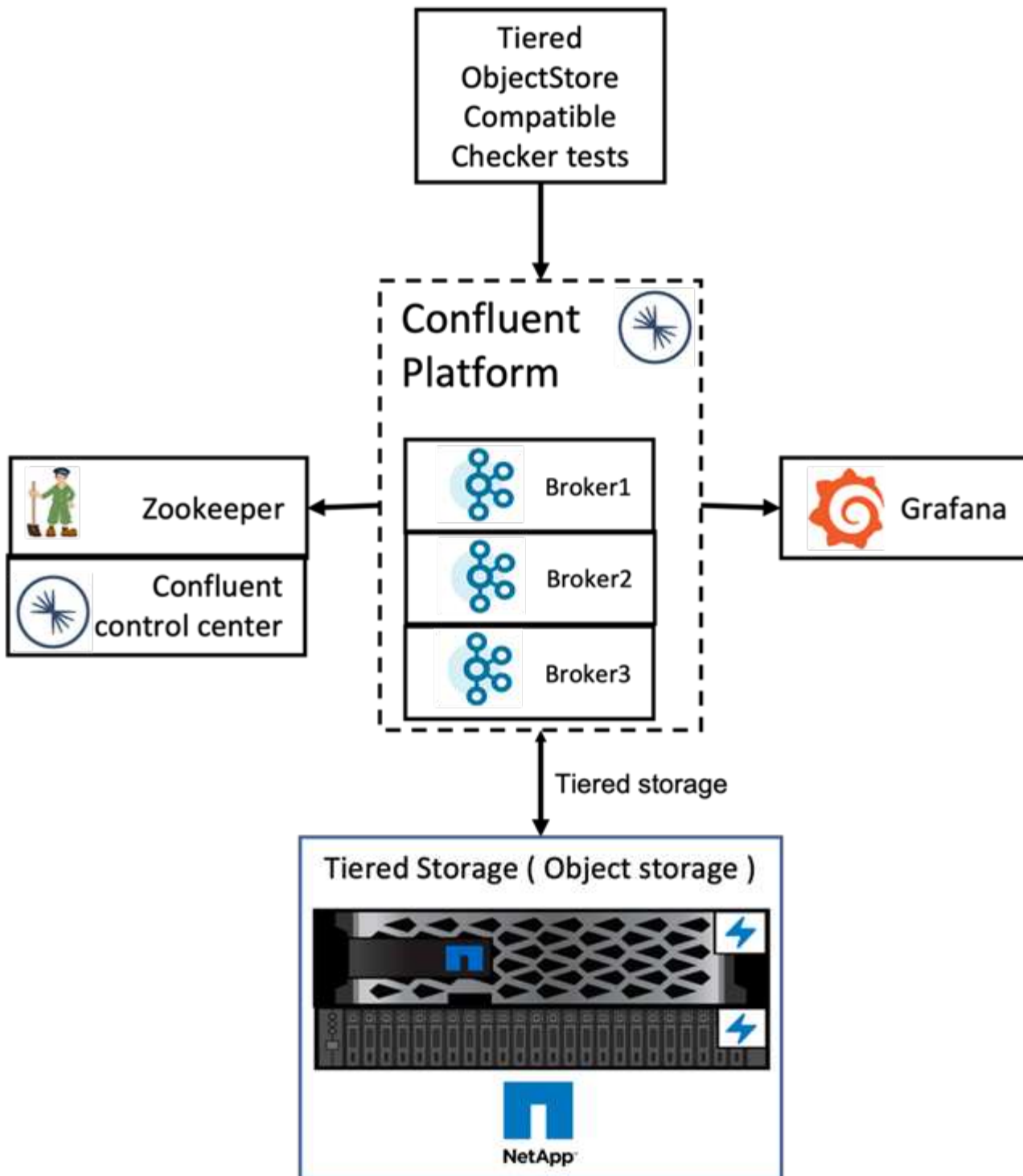
Dabei wurden die folgenden Testparameter verwendet:

```
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
:9092,192.168.150.198:9092,192.168.150.109:9092,192.168.150.165:9092,192.1
68.150.119:9092,192.168.150.133:9092
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
```

Zur Überprüfung haben wir ONTAP mit dem HTTP-Protokoll verwendet, aber HTTPS haben auch funktioniert. Der Zugriffsschlüssel und der Geheimschlüssel werden im Dateinamen im gespeichert `confluent.tier.s3.cred.file.path` Parameter.

NetApp Storage Controller – ONTAP

Wir haben zur Verifizierung eine Single HA-Paar-Konfiguration in ONTAP konfiguriert.



Überprüfungsergebnisse

Wir haben die folgenden fünf Testfälle für die Verifizierung abgeschlossen. Die ersten beiden waren Funktionstests und die verbleibenden drei waren Performance-Tests.

Korrektheit des Objektspeichers

Dieser Test führt grundlegende Vorgänge wie Abrufen, Put und Löschen im Objektspeicher durch, der mithilfe von API-Aufrufen für den Tiered Storage verwendet wird.

Funktionstest der Tiering-Funktion

Dieser Test überprüft die End-to-End-Funktionen des Objekt-Storage. Es erstellt ein Thema, erstellt einen Ereignisstrom zum neu erstellten Thema, wartet auf die Broker, um die Segmente im Objektspeicher zu archivieren, verbraucht den Ereignisstrom und validiert die verbrauchten Streams mit dem erzeugten Stream. Diesen Test haben wir mit und ohne Objektspeicherfehlerinspritzung durchgeführt. Wir haben einen Node-Ausfall simuliert, indem wir den Service Manager Service in einem der Nodes in ONTAP unterbrechen und validieren, dass die End-to-End-Funktionalität mit dem Objekt-Storage kompatibel ist.

Tier-Fetch-Benchmark

Dieser Test bestätigte die Lese-Performance des Tiered Objekt-Storage und prüfte den Bereich, in dem Leseanforderungen unter hoher Last aus vom Benchmark erzeugten Segmenten abgerufen werden. In dieser Benchmark hat Confluent benutzerdefinierte Clients entwickelt, um die Tier-fetch-Anforderungen zu erfüllen.

Produzieren-Workload-Generator zur Nutzung

Dieser Test generiert über die Archivierung von Segmenten indirekt einen Schreib-Workload auf dem Objektspeicher. Der Lese-Workload (Segmente Lesen) wurde aus dem Objekt-Storage generiert, wenn Verbrauchergruppen die Segmente abgerufen haben. Dieser Workload wurde von einem TOCC-Skript generiert. In diesem Test wurde die Performance von Lese- und Schreibvorgängen im Objekt-Storage in parallelen Threads überprüft. Wir haben wie bei dem Tiering-Funktionstest mit und ohne Objektspeicher-Fehlereinspritzung getestet.

Retention Workload Generator

Im Rahmen dieses Tests wurde die Löschleistung eines Objektspeichers unter einem anspruchsvollen Thema Aufbewahrung Workload überprüft. Der Retention Workload wurde mit einem TOCC-Skript erzeugt, das viele Nachrichten parallel zu einem Testthema erzeugt. Das Testthema wurde mit einer aggressiven größenbasierten und zeitbasierten Aufbewahrungseinstellung konfiguriert, die dazu führte, dass der Ereignisstrom kontinuierlich aus dem Objektspeicher gelöscht wurde. Die Segmente wurden anschließend archiviert. Dies führte dazu, dass der Vermittler zahlreiche Löschungen im Objekt-Storage erfasst und bei der Performance der Objektspeicher-Löschvorgänge erfasst wurde.

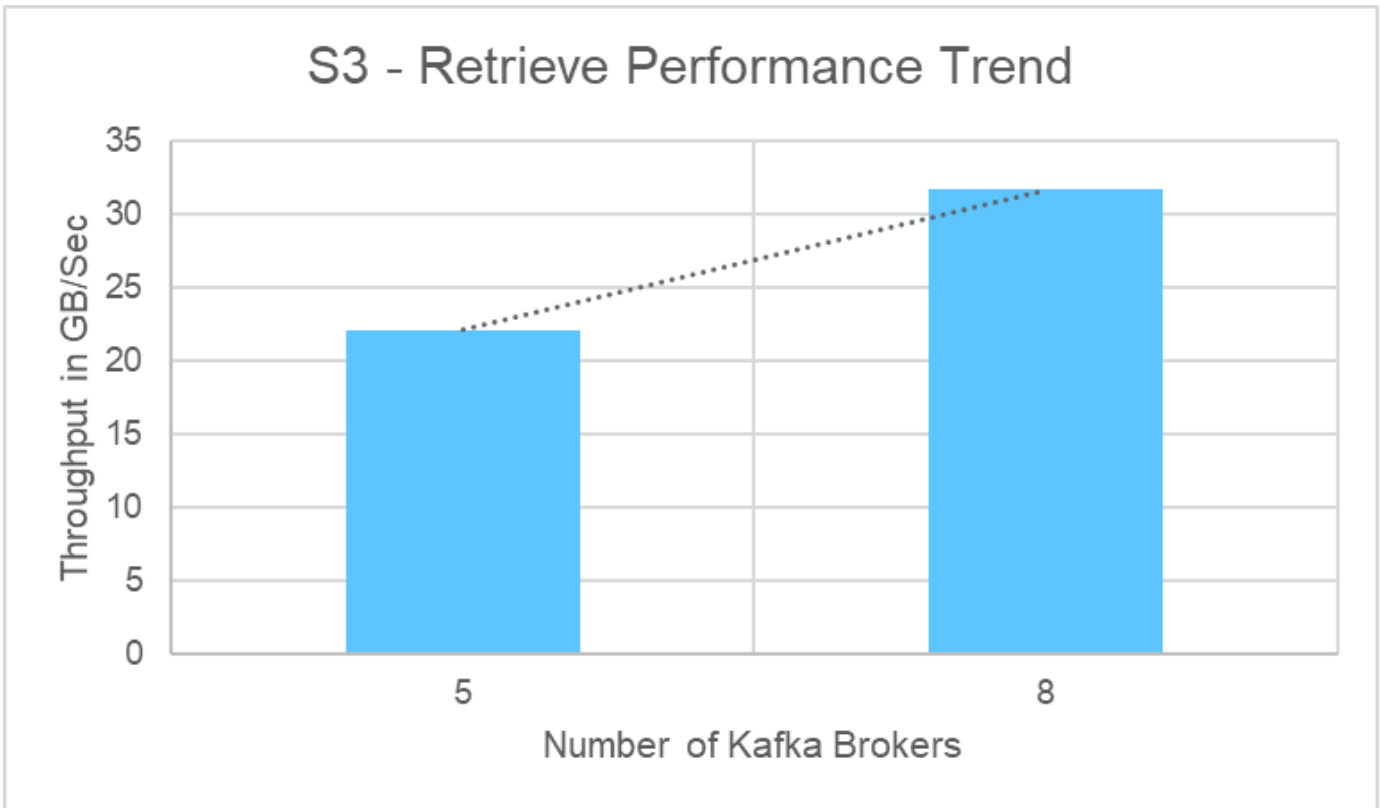
Informationen zur Überprüfung finden Sie im ["Fließend"](#) Website.

Performance-Tests mit dem Produce-Consume Workload Generator

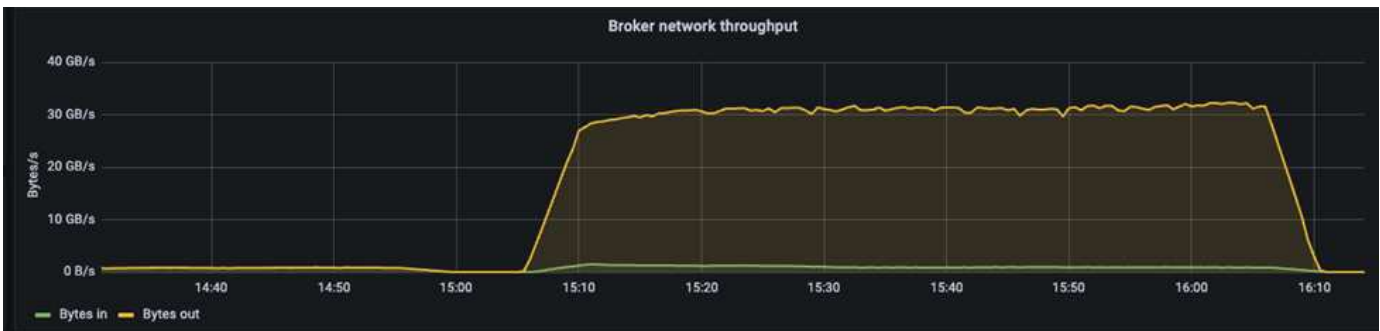
Wir haben während eines Produzieren-Workloads mit einem AFF A900 HA-Paar NetApp Storage Controller Tiered Storage-Tests mit fünf oder acht Broker Nodes durchgeführt. Unsere Tests ergaben, dass sich die Zeit bis zur Fertigstellung und die Performance-Ergebnisse mit der Anzahl der Broker Nodes skalieren lassen, bis die Auslastung von AFF A900 Ressourcen hundertprozentig erreichte. Das Setup des ONTAP Storage Controllers benötigte mindestens ein HA-Paar.

Die Performance für den S3-Abruffvorgang ist basierend auf der Anzahl der Confluent Broker Nodes linear gestiegen. Der ONTAP Storage Controller unterstützt bis zu 12 HA-Paare in einer einzelnen Implementierung.

Das folgende Diagramm zeigt kombinierten S3-Tiering-Datenverkehr mit fünf oder acht Broker-Nodes. Wir maximierter die Performance eines AFF A900 Single HA-Paars.



Im folgenden Diagramm wird der Kafka-Durchsatz bei etwa 31,74 GB/s dargestellt.



Außerdem beobachteten wir einen ähnlichen Durchsatz im ONTAP Storage Controller `perfstat` Bericht.

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

Leitfaden zu Best Practices für die Performance

Auf dieser Seite werden die Best Practices zur Verbesserung der Performance in dieser Lösung beschrieben.

- Verwenden Sie für ONTAP, wenn möglich, EINE GET-Größe $\geq 1\text{MB}$.
- Zunehmende `num.network.threads` Und `num.io.threads` In `server.properties` Mit Broker Nodes können Sie größere Tiering-Aktivitäten zur S3 Tier drängen. Diese Ergebnisse sind mit

`num.network.threads` Und `num.io.threads` Auf 32 einstellen.

- S3-Buckets sollten für acht Komponenten pro Mitgliedaggregat verwendet werden.
- Ethernet-Links für den S3-Datenverkehr sollten auf Storage und Client möglichst eine MTU von 9.000 verwenden.

Schlussfolgerung

Dieser Verifizierungstest erreichte einen Tiering-Durchsatz von 31.74 GB/s bei Confluent mit dem NetApp ONTAP Storage Controller.

Wo Sie weitere Informationen finden

Sehen Sie sich die folgenden Dokumente und/oder Websites an, um mehr über die in diesem Dokument beschriebenen Informationen zu erfahren:

- Was ist Confluent?

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

- S3-Sink Parameterdetails

["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)

- Apache Kafka

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

- S3 in ONTAP Best Practices

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

- S3-Objekt-Storage-Management

["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 Produktdokumentation

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

Copyright-Informationen

Copyright © 2024 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFT SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGEND EINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

Markeninformationen

NETAPP, das NETAPP Logo und die unter <http://www.netapp.com/TM> aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.