



NetApp ONTAP 스토리지 컨트롤러가 포함된 Confluent Kafka

NetApp artificial intelligence solutions

NetApp
August 18, 2025

목차

NetApp ONTAP 스토리지 컨트롤러가 포함된 Confluent Kafka	1
TR-4941: NetApp ONTAP 스토리지 컨트롤러와 Confluent	1
해결책	1
솔루션 아키텍처 세부 정보	2
기술 개요	3
NetApp ONTAP 스토리지 컨트롤러	3
주요 사용 사례	3
네이티브 S3 애플리케이션	4
FabricPool 엔드포인트	4
지류	6
Confluent 성능 검증	7
Confluent 설정	8
Confluent 계층형 스토리지 구성	8
NetApp 스토리지 컨트롤러 – ONTAP	9
검증 결과	10
생산-소비 워크로드 생성기를 사용한 성능 테스트	11
성능 모범 사례 지침	12
결론	13
추가 정보를 찾을 수 있는 곳	13

NetApp ONTAP 스토리지 컨트롤러가 포함된 Confluent Kafka

TR-4941: NetApp ONTAP 스토리지 컨트롤러와 Confluent

Karthikeyan Nagalingam, Joe Scott, NetApp Rankesh Kumar, Confluent

Confluent 플랫폼을 더 확장 가능하고 탄력적으로 만들려면 작업 부하를 매우 빠르게 확장하고 균형을 맞출 수 있어야 합니다. 계층형 스토리지를 사용하면 이러한 운영 부담을 줄여 Confluent에 엄청난 양의 데이터를 저장하는 것이 쉬워집니다.

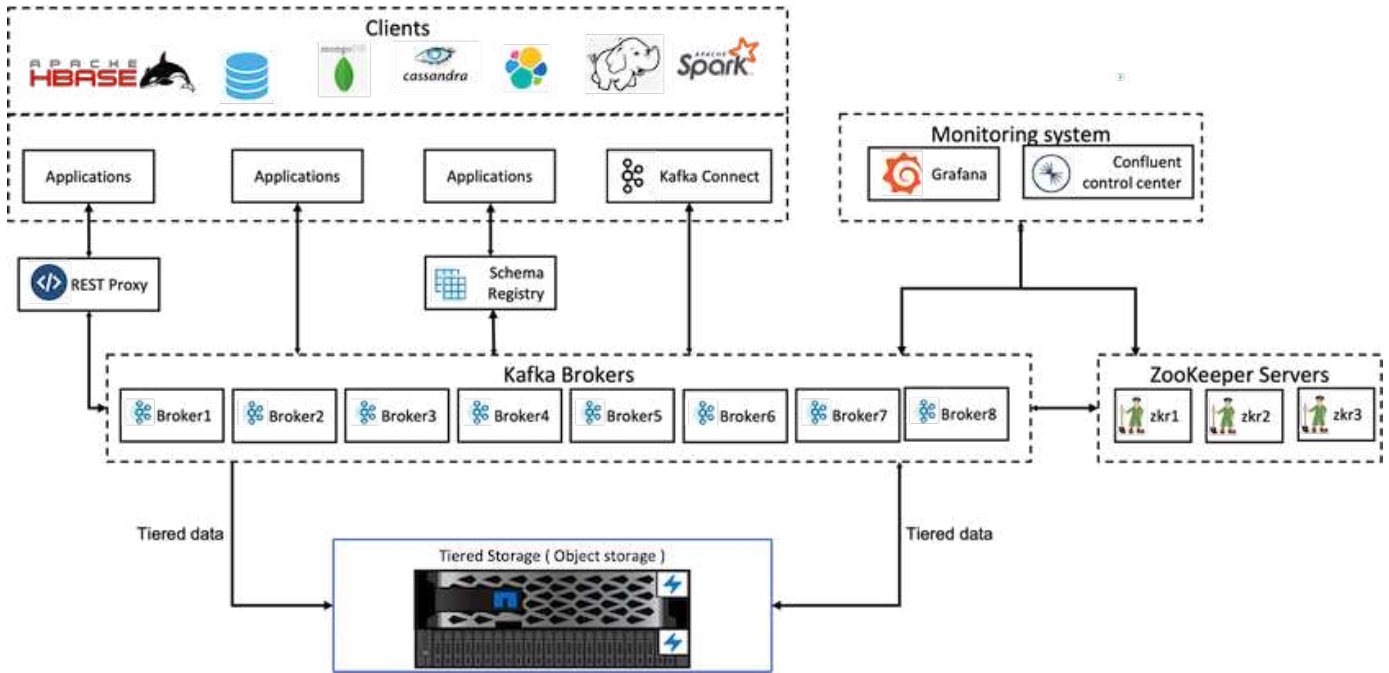
기본적인 아이디어는 데이터 저장과 데이터 처리를 분리하는 것인데, 이를 통해 각각을 독립적으로 확장하기가 훨씬 쉬워집니다.

업계를 선도하는 혁신 기술이 탑재된 NetApp ONTAP 데이터 관리 소프트웨어는 데이터가 있는 곳 어디에서나 Confluent에 많은 이점을 제공합니다.

이 문서에서는 계층형 스토리지 벤치마킹 키트를 사용하여 NetApp ONTAP에서 Confluent 플랫폼의 성능 벤치마크를 설명합니다.

해결책

ONTAP 기반의 Confluent 및 NetApp AFF A900 스토리지 컨트롤러는 데이터 스트림을 위해 설계된 분산 시스템입니다. 두 제품 모두 수평적 확장이 가능하고, 내결함성이 뛰어나며, 부하 상황에서도 뛰어난 성능을 제공합니다. 이러한 기술은 데이터 감소 기술을 통해 데이터 공간을 최소화하고 저장 비용을 낮추어 분산 데이터 스트리밍 및 스트림 처리에서 서로 보완됩니다. AFF A900 스토리지 컨트롤러는 뛰어난 성능을 제공하는 동시에 컴퓨팅 및 데이터 스토리지 리소스를 분리할 수 있습니다. 이를 통해 시스템 관리가 간소화되고 리소스를 독립적으로 확장할 수 있습니다.



솔루션 아키텍처 세부 정보

이 섹션에서는 계층형 스토리지를 위한 NetApp ONTAP 사용하여 Confluent Platform을 구축할 때 성능 검증에 사용되는 하드웨어와 소프트웨어에 대해 설명합니다. 다음 표에서는 솔루션 아키텍처와 기본 구성 요소를 설명합니다.

플랫폼 구성 요소	환경 구성
Confluent 플랫폼 버전 6.2	• 동물원 관리인 3명 • 8개의 브로커 서버 • 5 x 도구 서버 • 1 x 그래파나 • 1 x 제어 센터
모든 노드의 운영 체제	리눅스(우분투 18.04)
따뜻한 버킷을 위한 NetApp ONTAP	• 1 x AFF A900 고가용성(HA) 쌍 • 4 x 24 x 800 SSD • S3 프로토콜 • 100GbE
15개의 Fujitsu PRIMERGY RX2540 서버	• 2개의 CPU, 총 16개의 물리적 코어 • 인텔 제온 • 256GB 물리적 메모리 • 100GbE 듀얼 포트

기술 개요

이 섹션에서는 이 솔루션에 사용된 기술에 대해 설명합니다.

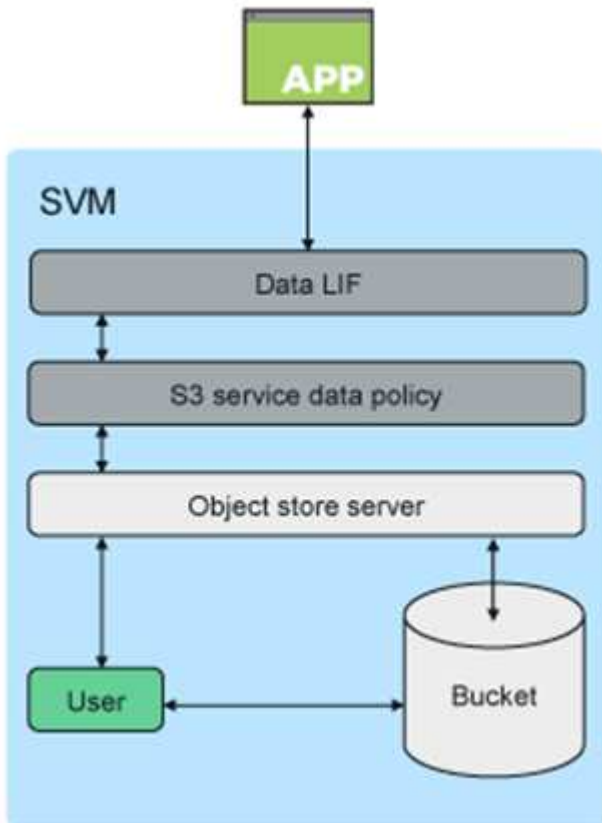
NetApp ONTAP 스토리지 컨트롤러

NetApp ONTAP 은 고성능 엔터프라이즈급 스토리지 운영 체제입니다.

NetApp ONTAP 9.8은 Amazon Simple Storage Service(S3) API에 대한 지원을 도입합니다. ONTAP Amazon Web Services(AWS) S3 API 작업의 하위 집합을 지원하고 클라우드 공급업체(AWS, Azure, GCP)와 온프레미스에서 ONTAP 기반 시스템의 객체로 데이터를 표현할 수 있도록 합니다.

NetApp StorageGRID 소프트웨어는 객체 스토리지를 위한 NetApp 의 대표적인 솔루션입니다. ONTAP 에지에서 수집 및 사전 처리 지점을 제공하고, NetApp 이 제공하는 객체 데이터에 대한 데이터 패브릭을 확장하고, NetApp 제품 포트폴리오의 가치를 높여 StorageGRID 보완합니다.

S3 버킷에 대한 액세스는 권한이 있는 사용자와 클라이언트 애플리케이션을 통해 제공됩니다. 다음 다이어그램은 애플리케이션이 S3 버킷에 액세스하는 것을 보여줍니다.



주요 사용 사례

S3 API를 지원하는 주요 목적은 ONTAP 에서 객체 액세스를 제공하는 것입니다. ONTAP 통합 스토리지 아키텍처는 이제 파일(NFS 및 SMB), 블록(FC 및 iSCSI), 객체(S3)를 지원합니다.

네이티브 S3 애플리케이션

점점 더 많은 애플리케이션이 S3를 사용하여 개체 액세스를 위해 ONTAP 지원을 활용할 수 있습니다. 대용량 보관 작업에는 적합하지만 네이티브 S3 애플리케이션에서 고성능에 대한 요구가 빠르게 증가하고 있으며 여기에는 다음이 포함됩니다.

- 해석학
- 인공지능
- 에지-투-코어 수집
- 머신러닝

이제 고객은 ONTAP System Manager와 같은 익숙한 관리 도구를 사용하여 ONTAP에서 개발 및 운영을 위한 고성능 개체 스토리지를 빠르게 프로비저닝하고, 이를 통해 ONTAP 스토리지 효율성과 보안의 이점을 누릴 수 있습니다.

FabricPool 엔드포인트

ONTAP 9.8부터 FabricPool ONTAP의 버킷 계층화를 지원하여 ONTAP 계층화를 허용합니다. 이는 기존 FAS 인프라를 객체 저장소 엔드포인트로 재활용하려는 고객에게 매우 좋은 옵션입니다.

FabricPool 두 가지 방법으로 ONTAP 계층화를 지원합니다.

- 로컬 클러스터 계층화. 비활성 데이터는 클러스터 LIF를 사용하여 로컬 클러스터에 있는 버킷에 계층화됩니다.
- 원격 클러스터 계층화. 비활성 데이터는 FabricPool 클라이언트의 IC LIF와 ONTAP 개체 저장소의 데이터 LIF를 사용하는 기존 FabricPool 클라우드 계층과 유사한 방식으로 원격 클러스터에 있는 버킷에 계층화됩니다.

ONTAP S3는 추가 하드웨어 및 관리 없이 기존 클러스터에서 S3 기능을 원하는 경우에 적합합니다. 300TB가 넘는 배포의 경우 NetApp StorageGRID 소프트웨어는 여전히 개체 스토리지를 위한 NetApp의 주력 솔루션입니다. ONTAP 또는 StorageGRID 클라우드 계층으로 사용하는 경우 FabricPool 라이선스가 필요하지 않습니다.

Confluent 계층형 스토리지용 NetApp ONTAP

모든 데이터 센터는 비즈니스에 중요한 애플리케이션을 계속 실행하고 중요한 데이터를 안전하게 사용할 수 있어야 합니다. 새로운 NetApp AFF A900 시스템은 ONTAP Enterprise Edition 소프트웨어와 고탄력성 설계를 기반으로 합니다. 당사의 새로운 초고속 NVMe 스토리지 시스템은 미션 크리티컬 운영의 중단을 없애고, 성능 조절을 최소화하며, 랜섬웨어 공격으로부터 데이터를 보호합니다.

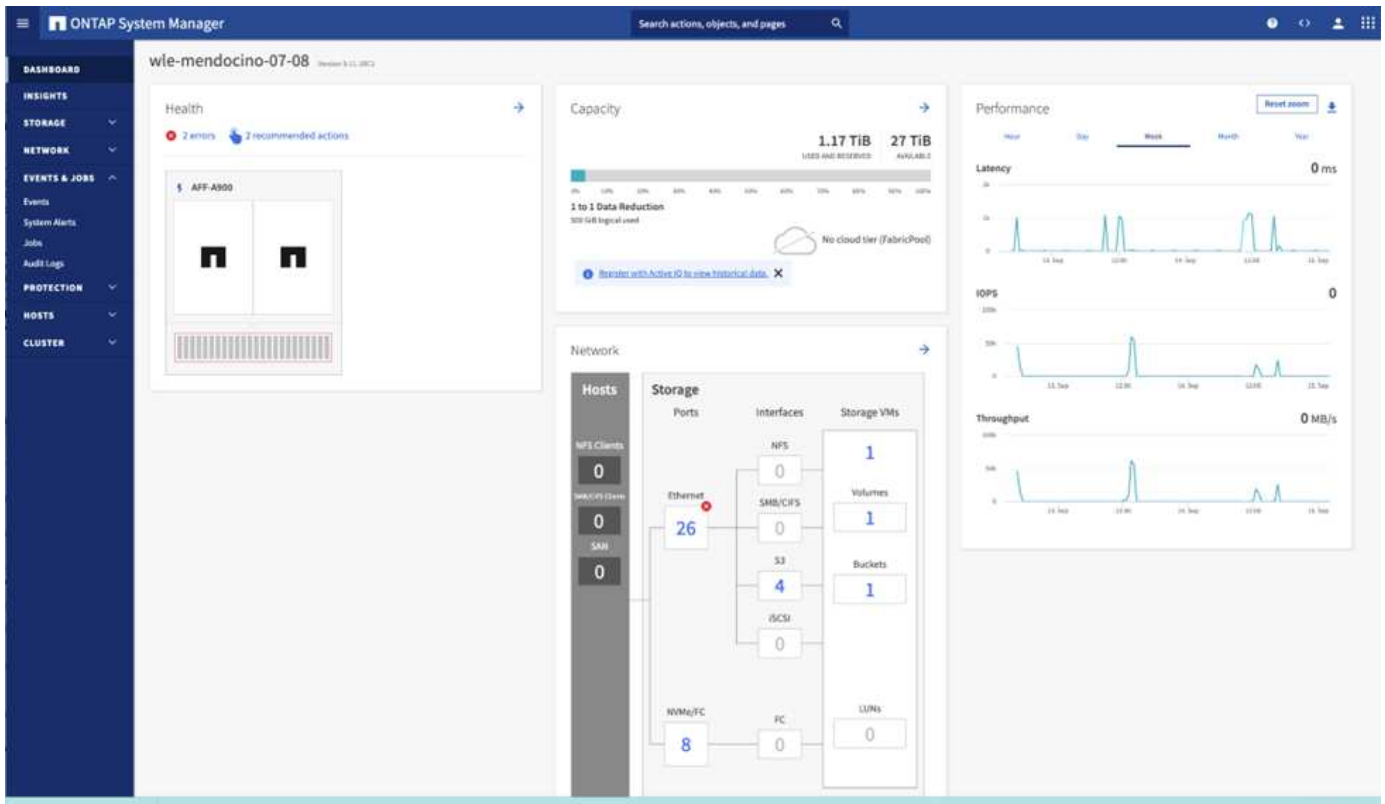
초기 배포부터 Confluent 클러스터 확장까지, 귀사의 환경은 비즈니스에 중요한 애플리케이션을 중단시키지 않는 변화에 신속하게 적응해야 합니다. ONTAP 엔터프라이즈 데이터 관리, 서비스 품질(QoS), 성능을 통해 환경에 맞게 계획하고 적응할 수 있습니다.

NetApp ONTAP과 Confluent Tiered Storage를 함께 사용하면 ONTAP 확장형 스토리지 대상으로 활용하여 Apache Kafka 클러스터 관리를 간소화하고 Confluent의 컴퓨팅 및 스토리지 리소스를 독립적으로 확장할 수 있습니다.

ONTAP S3 서버는 ONTAP의 성숙한 확장형 스토리지 기능을 기반으로 구축되었습니다. ONTAP 클러스터에 새로 추가된 노드를 사용하도록 S3 버킷을 확장하면 ONTAP 클러스터를 원활하게 확장할 수 있습니다.

ONTAP System Manager를 통한 간편한 관리

ONTAP 시스템 관리자는 브라우저 기반 그래픽 인터페이스로, 전 세계에 분산된 위치에 있는 ONTAP 스토리지 컨트롤러를 단일 창에서 구성, 관리 및 모니터링할 수 있습니다.



System Manager와 ONTAP CLI를 사용하여 ONTAP S3를 구성하고 관리할 수 있습니다. S3를 활성화하고 System Manager를 사용하여 버킷을 생성하면 ONTAP 단순화된 구성에 대한 모범 사례 기본값을 제공합니다. CLI에서 S3 서버와 버킷을 구성하는 경우에도 원하는 경우 System Manager를 사용하여 관리할 수 있으며 그 반대의 경우도 가능합니다.

System Manager를 사용하여 S3 버킷을 생성하면 ONTAP 시스템에서 사용 가능한 가장 높은 기본 성능 서비스 수준을 구성합니다. 예를 들어, AFF 시스템에서 기본 설정은 Extreme입니다. 성능 서비스 수준은 사전 정의된 적응형 QoS 정책 그룹입니다. 기본 서비스 수준 중 하나를 지정하는 대신 사용자 지정 QoS 정책 그룹을 지정하거나 정책 그룹을 지정하지 않을 수 있습니다.

미리 정의된 적응형 QoS 정책 그룹에는 다음이 포함됩니다.

- 극심한. 가장 낮은 지연 시간과 가장 높은 성능이 필요한 애플리케이션에 사용됩니다.
- 성능. 적당한 성능 요구와 지연 시간이 필요한 애플리케이션에 사용됩니다.
- 값. 지연 시간보다 처리량과 용량이 더 중요한 애플리케이션에 사용됩니다.
- 관습. 사용자 정의 QoS 정책을 지정하거나 QoS 정책을 지정하지 않습니다.

*계층화에 사용*을 선택하면 성능 서비스 수준이 선택되지 않으며, 시스템은 계층화된 데이터에 대해 최적의 성능을 제공하는 저비용 미디어를 선택하려고 합니다.

ONTAP 선택한 서비스 수준을 충족하면서 가장 적합한 디스크가 있는 로컬 계층에 이 버킷을 프로비저닝하려고 시도합니다. 하지만 버킷에 포함할 디스크를 지정해야 하는 경우 로컬 계층(집계)을 지정하여 CLI에서 S3 개체 스토리지를 구성하는 것을 고려하세요. CLI에서 S3 서버를 구성하는 경우에도 원하는 경우 시스템 관리자를 사용하여 관리할 수 있습니다.

버킷에 어떤 집계를 사용할지 지정하는 기능을 원하는 경우 CLI를 사용해서만 가능합니다.

지류

Confluent Platform은 연속적이고 실시간 스트림으로 데이터에 쉽게 액세스하고 저장하고 관리할 수 있는 본격적인 데이터 스트리밍 플랫폼입니다. Apache Kafka의 최초 개발자가 개발한 Confluent는 Kafka의 이점을 엔터프라이즈급 기능으로 확장하는 동시에 Kafka 관리나 모니터링의 부담을 제거합니다. 오늘날 Fortune 100 기업 중 80% 이상이 데이터 스트리밍 기술을 사용하고 있으며, 대부분이 Confluent를 사용합니다.

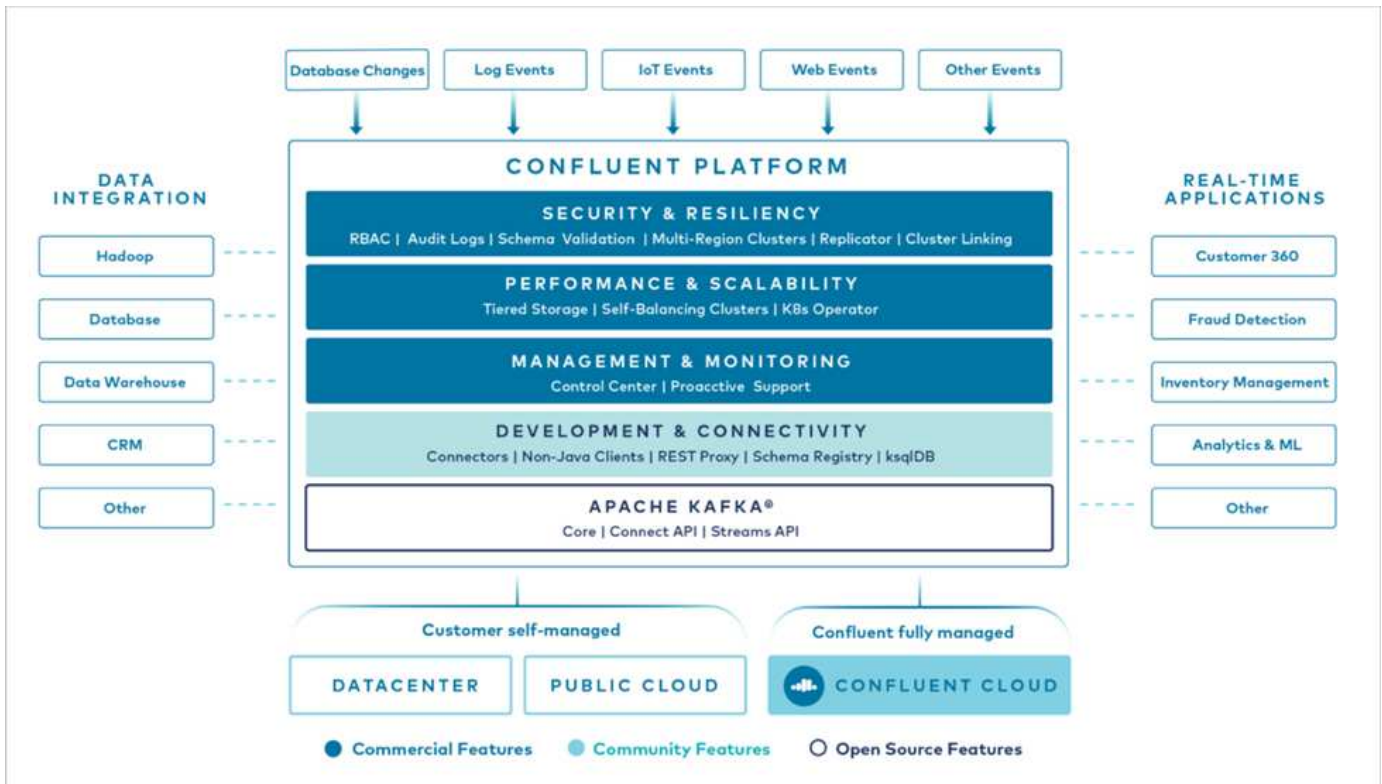
왜 Confluent를 선택해야 할까요?

Confluent는 과거 데이터와 실시간 데이터를 단일의 중앙 진실 소스로 통합하여 완전히 새로운 종류의 최신 이벤트 중심 애플리케이션을 쉽게 구축하고, 간편적인 데이터 파이프라인을 확보하고, 완전한 확장성, 성능, 안정성을 갖춘 강력한 새로운 사용 사례를 창출할 수 있도록 지원합니다.

Confluent는 무엇에 사용되나요?

Confluent Platform을 사용하면 여러 시스템 간에 데이터가 전송되거나 통합되는 방식과 같은 기본적인 메커니즘에 대해 걱정하는 대신 데이터에서 비즈니스 가치를 도출하는 방법에 집중할 수 있습니다. 특히, Confluent Platform은 Kafka에 대한 데이터 소스 연결, 스트리밍 애플리케이션 구축, Kafka 인프라의 보안, 모니터링 및 관리를 간소화합니다. 오늘날 Confluent Platform은 금융 서비스, 옴니채널 소매, 자율주행차부터 사기 탐지, 마이크로서비스, IoT까지 다양한 산업에 걸쳐 광범위한 사용 사례에 사용되고 있습니다.

다음 그림은 Confluent Platform의 구성 요소를 보여줍니다.



Confluent 이벤트 스트리밍 기술 개요

Confluent 플랫폼의 핵심은 다음과 같습니다. "카프카" 가장 인기 있는 오픈소스 분산 스트리밍 플랫폼입니다. 카프카의 주요 기능은 다음과 같습니다.

- 레코드 스트림을 게시하고 구독합니다.

- 장애에 견딜 수 있는 방식으로 레코드 스트림을 저장합니다.
- 레코드 스트림을 처리합니다.

Confluent Platform에는 기본적으로 스키마 레지스트리, REST 프록시, 100개 이상의 사전 구축된 Kafka 커넥터, ksqldb가 포함되어 있습니다.

Confluent 플랫폼 엔터프라이즈 기능 개요

- **Confluent 제어 센터.** Kafka를 관리하고 모니터링하기 위한 UI 기반 시스템입니다. Kafka Connect를 쉽게 관리하고 다른 시스템에 대한 연결을 생성, 편집, 관리할 수 있습니다.
- **Kubernetes용 Confluent.** Kubernetes용 Confluent는 Kubernetes 운영자입니다. 쿠버네티스 운영자는 특정 플랫폼 애플리케이션에 대한 고유한 기능과 요구 사항을 제공하여 쿠버네티스의 오케스트레이션 기능을 확장합니다. Confluent Platform의 경우, 여기에는 Kubernetes에서 Kafka의 배포 프로세스를 크게 단순화하고 일반적인 인프라 수명 주기 작업을 자동화하는 것이 포함됩니다.
- **Kafka Connect 커넥터.** 커넥터는 Kafka Connect API를 사용하여 Kafka를 데이터베이스, 키-값 저장소, 검색 인덱스, 파일 시스템 등의 다른 시스템에 연결합니다. Confluent Hub에는 가장 인기 있는 데이터 소스와 싱크에 대한 다운로드 가능한 커넥터가 있으며, 여기에는 Confluent Platform에서 이러한 커넥터의 완전히 테스트되고 지원되는 버전이 포함됩니다. 더 자세한 내용은 다음을 참조하세요. "[여기](#)".
- **자체 균형 클러스터.** 자동화된 부하 분산, 장애 감지 및 자체 복구 기능을 제공합니다. 또한 필요에 따라 브로커를 추가하거나 해제하는 기능을 제공하며, 수동으로 조정할 필요가 없습니다.
- **합류 클러스터 연결.** 링크 브리지를 통해 클러스터를 직접 연결하고 한 클러스터의 주제를 다른 클러스터로 미러링합니다. 클러스터 연결을 통해 다중 데이터 센터, 다중 클러스터 및 하이브리드 클라우드 배포 설정이 간소화됩니다.
- **Confluent 자동 데이터 밸런서.** 클러스터 내의 브로커 수, 파티션 크기, 파티션 수, 리더 수를 모니터링합니다. 이 기능을 사용하면 클러스터 전체에서 균일한 작업 부하를 생성하기 위해 데이터를 이동할 수 있으며, 재조정하는 동안 프로덕션 작업 부하에 미치는 영향을 최소화하기 위해 재조정 트래픽을 조절할 수 있습니다.
- **합류 복제기.** 여러 데이터 센터에서 여러 Kafka 클러스터를 유지 관리하는 것이 그 어느 때보다 쉬워졌습니다.
- **계층화된 스토리지.** 선호하는 클라우드 공급업체를 사용하여 대용량의 Kafka 데이터를 저장할 수 있는 옵션을 제공하므로 운영 부담과 비용이 줄어듭니다. 계층형 스토리지를 사용하면 비용 효율적인 개체 스토리지에 데이터를 보관하고, 더 많은 컴퓨팅 리소스가 필요할 때만 확장 브로커를 사용할 수 있습니다.
- **Confluent JMS 클라이언트.** Confluent Platform에는 Kafka용 JMS 호환 클라이언트가 포함되어 있습니다. 이 Kafka 클라이언트는 Kafka 브로커를 백엔드로 사용하여 JMS 1.1 표준 API를 구현합니다. 이 기능은 JMS를 사용하는 레거시 애플리케이션이 있고 기존 JMS 메시지 브로커를 Kafka로 교체하려는 경우에 유용합니다.
- **Confluent MQTT 프록시.** 중간에 MQTT 브로커가 필요 없이 MQTT 장치 및 게이트웨이에서 Kafka로 직접 데이터를 게시하는 방법을 제공합니다.
- **Confluent 보안 플러그인.** Confluent 보안 플러그인은 다양한 Confluent Platform 도구와 제품에 보안 기능을 추가하는 데 사용됩니다. 현재 Confluent REST 프록시에 사용할 수 있는 플러그인이 있는데, 이 플러그인은 들어오는 요청을 인증하고 인증된 주체를 Kafka에 대한 요청에 전파하는 데 도움이 됩니다. 이를 통해 Confluent REST 프록시 클라이언트는 Kafka 브로커의 멀티테넌트 보안 기능을 활용할 수 있습니다.

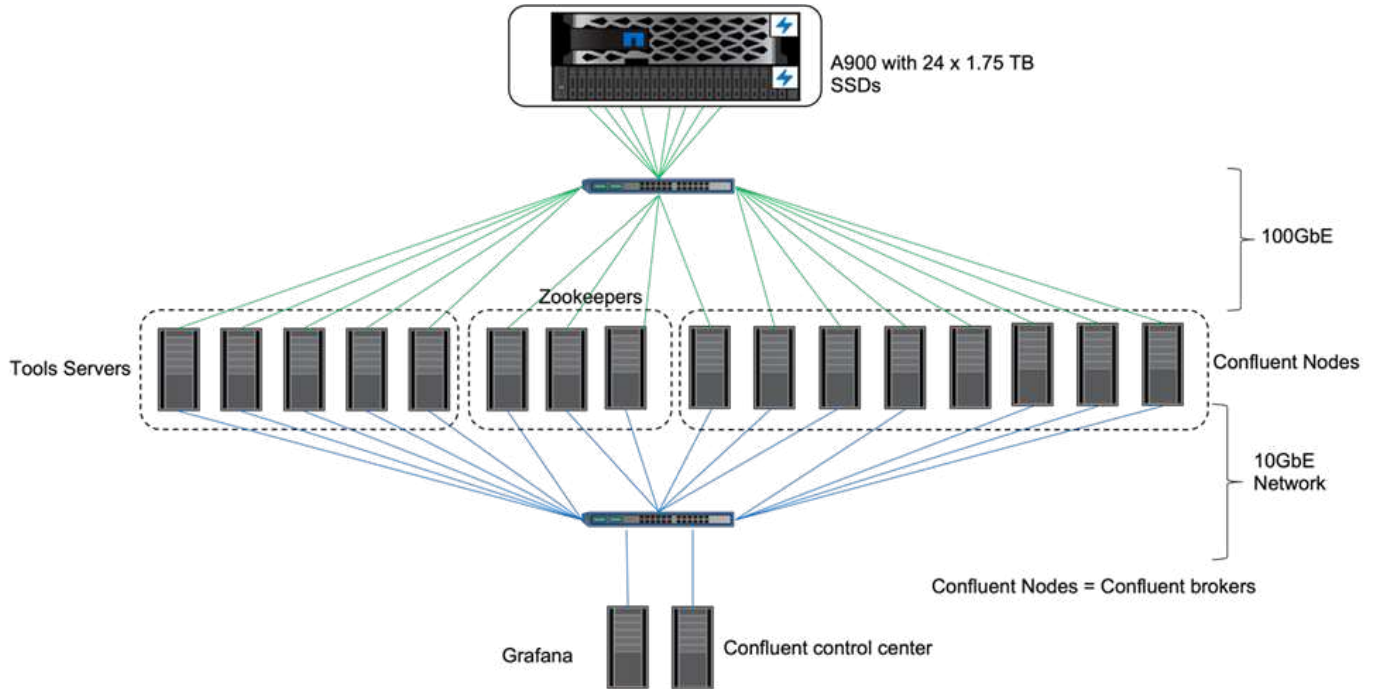
Confluent 성능 검증

NetApp ONTAP의 계층형 스토리지에 대해 Confluent Platform을 사용하여 검증을 수행했습니다. NetApp과 Confluent 팀은 이 검증 작업을 함께 수행했으며 이에 필요한 테스트 사례를 실행했습니다.

Confluent 설정

설정을 위해 256GB RAM과 16개 CPU를 갖춘 3개의 동물원 관리자, 5개의 브로커, 5개의 테스트 서버를 사용했습니다. NetApp 스토리지의 경우 AFF A900 HA 쌍이 있는 ONTAP 사용했습니다. 저장소와 브로커는 100GbE 연결을 통해 연결되었습니다.

다음 그림은 계층형 스토리지 검증에 사용되는 구성의 네트워크 토폴로지를 보여줍니다.



도구 서버는 Confluent 노드와 이벤트를 주고받는 애플리케이션 클라이언트 역할을 합니다.

Confluent 계층형 스토리지 구성

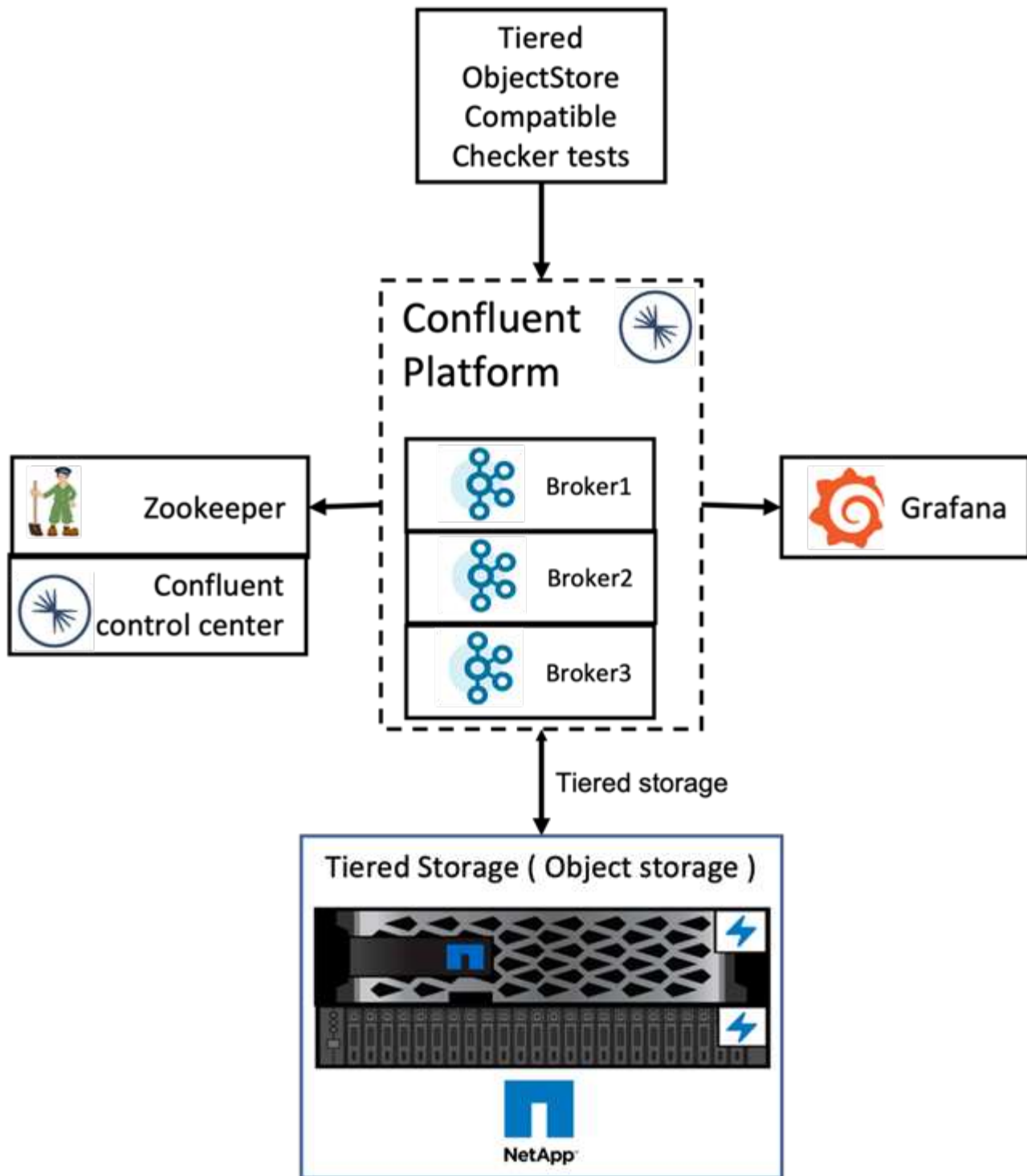
우리는 다음과 같은 테스트 매개변수를 사용했습니다.

```
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
```

검증을 위해 HTTP 프로토콜을 사용하는 ONTAP 사용했지만 HTTPS도 작동했습니다. 액세스 키와 비밀 키는 제공된 파일 이름에 저장됩니다. `confluent.tier.s3.cred.file.path` 매개변수.

NetApp 스토리지 컨트롤러 – ONTAP

검증을 위해 ONTAP 에서 단일 HA 쌍 구성을 구성했습니다.



검증 결과

우리는 검증을 위해 다음의 5가지 테스트 사례를 완료했습니다. 처음 두 가지는 기능 테스트였고 나머지 세 가지는 성능 테스트였습니다.

객체 저장소 정확성 테스트

이 테스트는 API 호출을 사용하여 계층형 스토리지에 사용되는 객체 저장소에서 get, put, delete와 같은 기본 작업을 수행합니다.

계층화 기능 정확성 테스트

이 테스트는 개체 스토리지의 중단 간 기능을 확인합니다. 주제를 생성하고, 새로 생성된 주제에 대한 이벤트 스트림을 생성하고, 브로커가 세그먼트를 개체 스토리지에 보관할 때까지 기다리고, 이벤트 스트림을 사용하고, 사용된 스트림이 생성된 스트림과 일치하는지 검증합니다. 우리는 객체 저장소 오류 주입을 적용한 경우와 적용하지 않은 경우로 이 테스트를 수행했습니다. ONTAP의 노드 중 하나에서 서비스 관리자 서비스를 중지하고 엔드투엔드 기능이 개체 스토리지에서 작동하는지 검증하여 노드 장애를 시뮬레이션했습니다.

티어 페치 벤치마크

이 테스트는 계층형 개체 스토리지의 읽기 성능을 검증하고 벤치마크에서 생성된 세그먼트에서 높은 부하가 걸리는 범위 페치 읽기 요청을 확인했습니다. 이 벤치마크에서 Confluent는 계층별 페치 요청을 처리하기 위해 사용자 정의 클라이언트를 개발했습니다.

생산-소비 작업 생성기

이 테스트는 세그먼트 보관을 통해 개체 저장소에 대한 쓰기 작업 부하를 간접적으로 생성합니다. 소비자 그룹이 세그먼트를 가져올 때 개체 스토리지에서 읽기 작업 부하(세그먼트 읽기)가 생성되었습니다. 이 작업 부하는 TOCC 스크립트에 의해 생성되었습니다. 이 테스트는 병렬 스레드에서 개체 스토리지의 읽기 및 쓰기 성능을 확인했습니다. 우리는 계층화 기능 정확성 테스트에서 했던 것처럼 객체 저장소 오류 주입을 적용한 경우와 적용하지 않은 경우를 테스트했습니다.

보존 작업량 생성기

이 테스트는 주제 보존 작업 부하가 큰 상황에서 개체 스토리지의 삭제 성능을 점검했습니다. 보존 작업 부하는 테스트 주제와 병렬로 많은 메시지를 생성하는 TOCC 스크립트를 사용하여 생성되었습니다. 테스트 주제는 이벤트 스트림이 개체 저장소에서 지속적으로 제거되도록 하는 공격적인 크기 기반 및 시간 기반 보존 설정으로 구성되었습니다. 그런 다음 세그먼트는 보관되었습니다. 이로 인해 브로커가 개체 저장소에서 많은 삭제를 수행하게 되었고 개체 저장소 삭제 작업의 성능이 저하되었습니다.

확인 세부 사항은 다음을 참조하세요. "[지류](#)" 웹사이트.

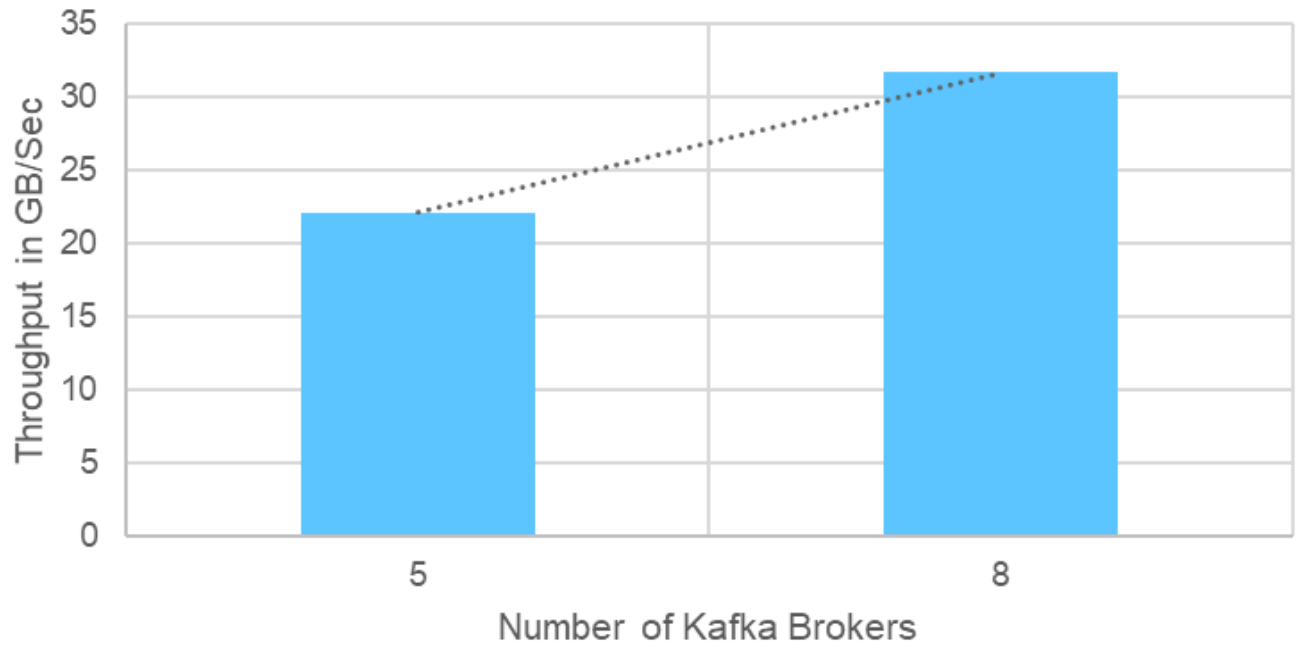
생산-소비 워크로드 생성기를 사용한 성능 테스트

우리는 하나의 AFF A900 HA 쌍 NetApp 스토리지 컨트롤러를 사용하여 생산-소비 워크로드 동안 5개 또는 8개의 브로커 노드로 계층형 스토리지 테스트를 수행했습니다. 테스트에 따르면, AFF A900 리소스 사용률이 100%에 도달할 때까지 완료 시간과 성능 결과는 브로커 노드 수에 따라 조정되었습니다. ONTAP 스토리지 컨트롤러 설정에는 최소 1개의 HA 쌍이 필요했습니다.

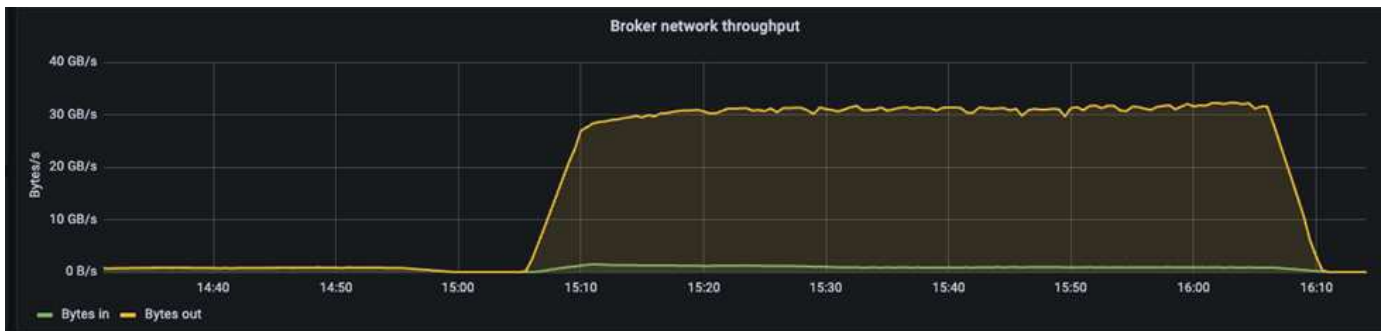
S3 검색 작업의 성능은 Confluent 브로커 노드 수에 따라 선형적으로 증가했습니다. ONTAP 스토리지 컨트롤러는 단일 배포에서 최대 12개의 HA 쌍을 지원합니다.

다음 그래프는 5개 또는 8개의 브로커 노드가 있는 결합된 S3 계층화 트래픽을 보여줍니다. 우리는 AFF A900 단일 HA 쌍의 성능을 극대화했습니다.

S3 - Retrieve Performance Trend



다음 그래프는 약 31.74GBps의 Kafka 처리량을 보여줍니다.



또한 ONTAP 스토리지 컨트롤러에서도 유사한 처리량을 관찰했습니다. perfstat 보고서.

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

성능 모범 사례 지침

이 페이지에서는 이 솔루션의 성능을 개선하기 위한 모범 사례를 설명합니다.

- ONTAP 의 경우 가능하면 GET 크기 $\geq 1\text{MB}$ 를 사용하세요.
- 증가하다 num.network.threads 그리고 num.io.threads ~에 server.properties 브로커 노드에서 S3 계층으로 계층화 활동을 늘릴 수 있습니다. 이러한 결과는 다음과 같습니다. num.network.threads 그리고 num.io.threads 32로 설정.

- S3 버킷은 멤버 집계당 8개의 구성 요소를 타겟으로 해야 합니다.
- S3 트래픽을 구동하는 이더넷 링크는 가능한 경우 저장소와 클라이언트 모두에서 9k의 MTU를 사용해야 합니다.

결론

이 검증 테스트에서는 NetApp ONTAP 스토리지 컨트롤러가 있는 Confluent에서 31.74GBps의 계층화 처리량을 달성했습니다.

추가 정보를 찾을 수 있는 곳

이 문서에 설명된 정보에 대해 자세히 알아보려면 다음 문서 및/또는 웹사이트를 검토하세요.

- Confluent란 무엇인가요?

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

- S3-싱크 매개변수 세부 정보

["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)

- 아파치 카프카

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

- ONTAP 의 S3 모범 사례

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

- S3 객체 스토리지 관리

["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 제품 문서

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

저작권 정보

Copyright © 2025 NetApp, Inc. All Rights Reserved. 미국에서 인쇄된 본 문서의 어떠한 부분도 저작권 소유자의 사전 서면 승인 없이는 어떠한 형식이나 수단(복사, 녹음, 녹화 또는 전자 검색 시스템에 저장하는 것을 비롯한 그래픽, 전자적 또는 기계적 방법)으로도 복제될 수 없습니다.

NetApp이 저작권을 가진 자료에 있는 소프트웨어에는 아래의 라이선스와 고지사항이 적용됩니다.

본 소프트웨어는 NetApp에 의해 '있는 그대로' 제공되며 상품성 및 특정 목적에의 적합성에 대한 명시적 또는 묵시적 보증을 포함하여(이에 제한되지 않음) 어떠한 보증도 하지 않습니다. NetApp은 대체품 또는 대체 서비스의 조달, 사용 불능, 데이터 손실, 이익 손실, 영업 중단을 포함하여(이에 국한되지 않음), 이 소프트웨어의 사용으로 인해 발생하는 모든 직접 및 간접 손해, 우발적 손해, 특별 손해, 징벌적 손해, 결과적 손해의 발생에 대하여 그 발생 이유, 책임론, 계약 여부, 엄격한 책임, 불법 행위(과실 또는 그렇지 않은 경우)와 관계없이 어떠한 책임도 지지 않으며, 이와 같은 손실의 발생 가능성이 통지되었다 하더라도 마찬가지입니다.

NetApp은 본 문서에 설명된 제품을 언제든지 예고 없이 변경할 권리를 보유합니다. NetApp은 NetApp의 명시적인 서면 동의를 받은 경우를 제외하고 본 문서에 설명된 제품을 사용하여 발생하는 어떠한 문제에도 책임을 지지 않습니다. 본 제품의 사용 또는 구매의 경우 NetApp에서는 어떠한 특허권, 상표권 또는 기타 지적 재산권이 적용되는 라이선스도 제공하지 않습니다.

본 설명서에 설명된 제품은 하나 이상의 미국 특허, 해외 특허 또는 출원 중인 특허로 보호됩니다.

제한적 권리 표시: 정부에 의한 사용, 복제 또는 공개에는 DFARS 252.227-7013(2014년 2월) 및 FAR 52.227-19(2007년 12월)의 기술 데이터-비상업적 품목에 대한 권리(Rights in Technical Data -Noncommercial Items) 조항의 하위 조항 (b)(3)에 설명된 제한사항이 적용됩니다.

여기에 포함된 데이터는 상업용 제품 및/또는 상업용 서비스(FAR 2.101에 정의)에 해당하며 NetApp, Inc.의 독점 자산입니다. 본 계약에 따라 제공되는 모든 NetApp 기술 데이터 및 컴퓨터 소프트웨어는 본질적으로 상업용이며 개인 비용만으로 개발되었습니다. 미국 정부는 데이터가 제공된 미국 계약과 관련하여 해당 계약을 지원하는 데에만 데이터에 대한 전 세계적으로 비독점적이고 양도할 수 없으며 재사용이 불가능하며 취소 불가능한 라이선스를 제한적으로 가집니다. 여기에 제공된 경우를 제외하고 NetApp, Inc.의 사전 서면 승인 없이는 이 데이터를 사용, 공개, 재생산, 수정, 수행 또는 표시할 수 없습니다. 미국 국방부에 대한 정부 라이선스는 DFARS 조항 252.227-7015(b)(2014년 2월)에 명시된 권한으로 제한됩니다.

상표 정보

NETAPP, NETAPP 로고 및 <http://www.netapp.com/TM>에 나열된 마크는 NetApp, Inc.의 상표입니다. 기타 회사 및 제품 이름은 해당 소유자의 상표일 수 있습니다.