NetApp ONTAP 스토리지 컨트롤러를 사용하는 Confluent Kafka

NetApp Solutions

NetApp
April 18, 2024

This PDF was generated from https://docs.netapp.com/ko-kr/netapp-solutions/data-analytics/kafka-ontap-overview.html on April 18, 2024. Always check docs.netapp.com for the latest.

목차

NetApp ONTAP 스토리지 컨트롤러를 사용하는 Confluent Kafka . . . 1
TR-4941: NetApp ONTAP 스토리지 컨트롤러에 대해 잘 알고 있습니다 . . . 1
해결 방법 . . . 1
기술 개요 . . . 3
Confluent 성능 검증 . . . 7
농작물 사용 워크로드 생성기를 사용한 성능 테스트 . . . 11
성능 모범 사례 지침 . . . 12
결론 . . . 13

NetApp ONTAP 스토리지 컨트롤러를 사용하는 Confluent Kafka

TR-4941: NetApp ONTAP 스토리지 컨트롤러에 대해 잘 알고 있습니다

Karthikeyan Nagalingam, Joe Scott, NetApp Rankesh Kumar, Confluent

Confluent Platform의 확장성과 탄력성을 크게 확보하려면 IT에서 워크로드를 매우 빠르게 확장하고 균형을 맞출 수 있어야 합니다. 계층형 스토리지를 사용하면 운영 부담을 줄여 대량의 데이터를 유창한 관리 방식으로 저장할 수 있습니다.

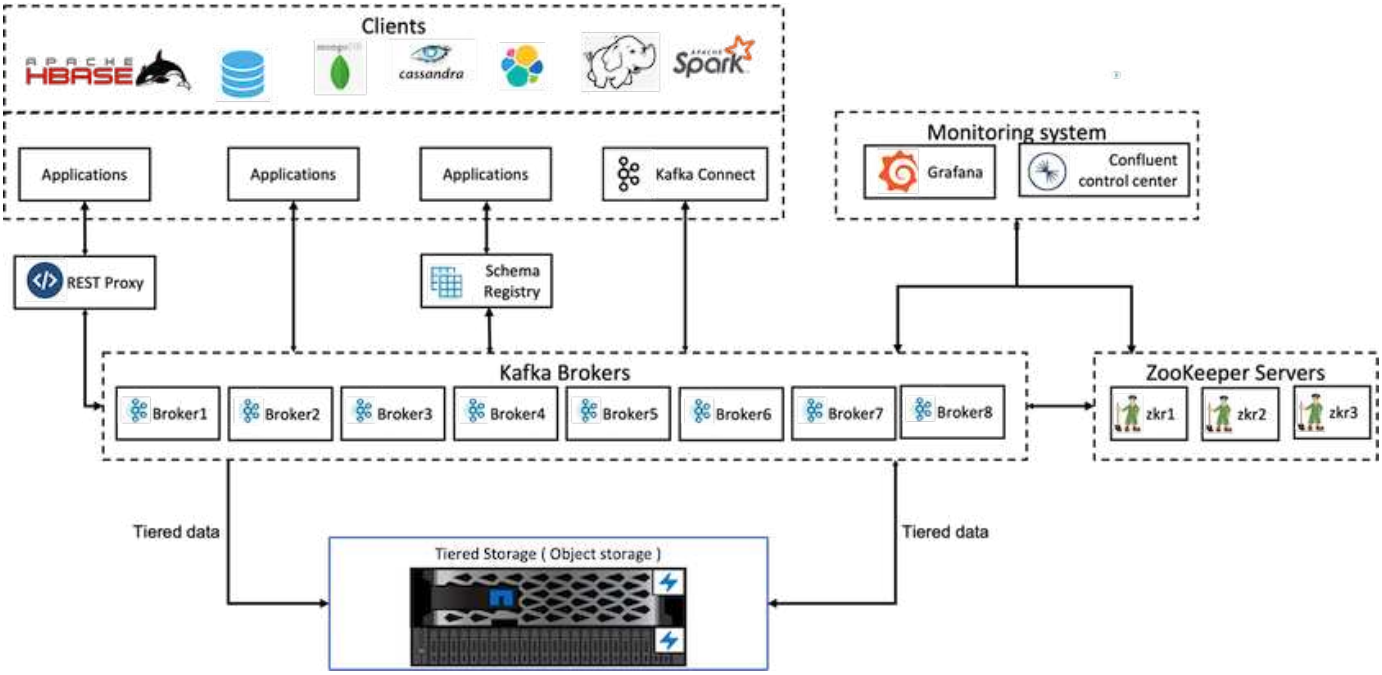
기본 개념은 데이터 스토리지와 데이터 처리를 분리하여 각각을 독립적으로 쉽게 확장할 수 있도록 하는 것입니다.

업계 최고의 혁신 기술을 갖춘 NetApp ONTAP 데이터 관리 소프트웨어는 데이터가 어디에 있든 다양한 이점을 유창하게 제공합니다.

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

해결 방법

ONTAP 기반 Confluent 및 NetApp AFF A900 스토리지 컨트롤러는 데이터 스트림을 위해 설계된 분산 시스템입니다. 두 가지 모두 수평적으로 확장 가능하고 내결함성이 있으며 부하 상태에서 뛰어난 성능을 제공합니다. 데이터 공간을 최소화하는 데이터 축소 기술로 분산된 데이터 스트리밍 및 스트림 처리에서 서로 보완하여 스토리지 비용을 절감합니다. AFF A900 스토리지 컨트롤러는 뛰어난 성능을 제공하는 동시에 컴퓨팅과 데이터 스토리지 리소스를 분리할 수 있도록 합니다. 따라서 시스템 관리가 단순화되고 리소스를 독립적으로 확장할 수 있습니다.



솔루션 아키텍처 세부 정보

이 섹션에서는 계층형 스토리지를 위한 NetApp ONTAP를 사용하여 Confluent Platform 배포에서 성능을 검증하는 데 사용되는 하드웨어 및 소프트웨어에 대해 설명합니다. 다음 표에서는 솔루션 아키텍처와 기본 구성 요소에 대해 설명합니다.

플랫폼 구성 요소	환경 구성
Confluent 플랫폼 버전 6.2	• 주키프 3개 • 브로커 서버 8대 • 도구 서버 5대 • Grafana 1개 • 제어 센터 1개
모든 노드의 운영 체제	Linux(Ubuntu 18.04)
웜 버킷용 NetApp ONTAP	• AFF A900 고가용성(HA) 쌍 1개 • 24 x 800 SSD 4개 • S3 프로토콜 • 100GbE
15 Fujitsu Primergy RX2540 서버	• CPU 2개, 총 물리적 코어 16개 • 인텔 제온 • 256GB 물리적 메모리 • 100GbE 이중 포트

기술 개요

이 섹션에서는 이 솔루션에 사용된 기술에 대해 설명합니다.

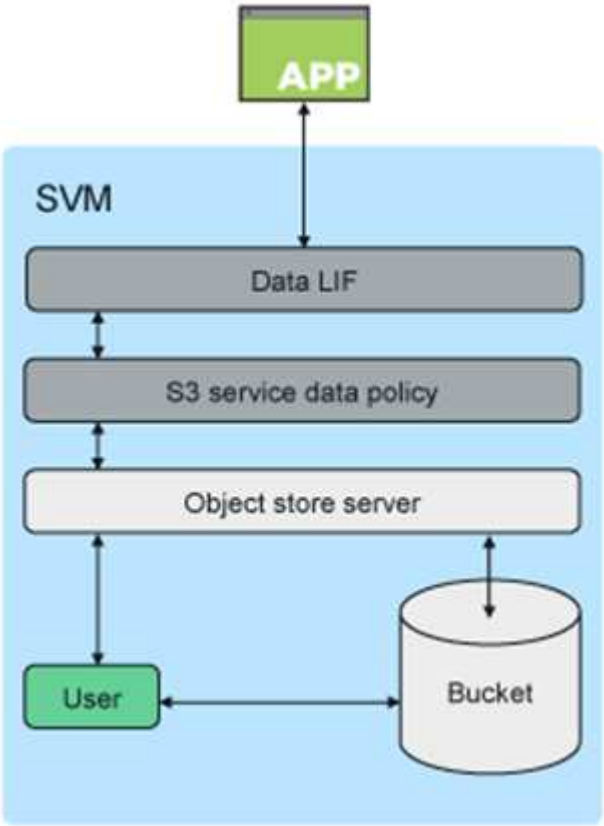
NetApp ONTAP 스토리지 컨트롤러

NetApp ONTAP는 고성능 엔터프라이즈급 스토리지 운영 체제입니다.

NetApp ONTAP 9.8은 Amazon S3(Simple Storage Service) API를 지원합니다. ONTAP은 AWS(Amazon Web Services) S3 API 작업의 서브셋을 지원하며 클라우드 공급자(AWS, Azure 및 GCP) 및 온프레미스 전반에서 데이터를 ONTAP 기반 시스템에서 오브젝트로 나타낼 수 있습니다.

NetApp StorageGRID 소프트웨어는 오브젝트 스토리지를 위한 NetApp의 대표적인 솔루션입니다. ONTAP은 오브젝트 데이터를 위해 NetApp에서 제공하는 Data Fabric을 확장하고 NetApp 제품 포트폴리오의 가치를 높여 에지에서 수집 및 사전 처리 지점을 제공함으로써 StorageGRID를 보완합니다.

S3 버킷에 대한 액세스는 인증된 사용자 및 클라이언트 애플리케이션을 통해 제공됩니다. 다음 다이어그램에서는 S3 버킷에 액세스하는 애플리케이션을 보여 줍니다.



주요 사용 사례

S3 API를 지원하는 주된 목적은 ONTAP에서 오브젝트 액세스를 제공하는 것입니다. ONTAP 유니파이드 스토리지 아키텍처는 이제 파일(NFS 및 SMB), 블록(FC 및 iSCSI), 오브젝트(S3)를 지원합니다.

네이티브 **S3** 애플리케이션

점점 더 많은 애플리케이션에서 S3를 사용하여 오브젝트 액세스에 ONTAP 지원을 활용할 수 있습니다. 고용량 아카이브 워크로드에 적합하지만 네이티브 S3 애플리케이션에서 고성능을 필요로 하는 것은 다음과 같습니다.

- 분석
- 인공 지능
- 엣지-코어 수집
- 머신 러닝

고객은 이제 ONTAP System Manager와 같은 익숙한 관리 효율 툴을 사용하여 ONTAP의 개발 및 운영에 필요한 고성능 오브젝트 스토리지를 신속하게 프로비저닝할 수 있습니다. 또한 ONTAP 스토리지 효율성 및 보안을 활용할 수 있습니다.

FabricPool 엔드포인트

ONTAP 9.8부터 FabricPool는 ONTAP의 버킷 계층화를 지원하므로 ONTAP에서 ONTAP로 계층화할 수 있습니다. 이 옵션은 기존 FAS 인프라를 오브젝트 저장소 엔드포인트로 용도 변경하고자 하는 고객에게 적합합니다.

FabricPool는 다음 두 가지 방법으로 ONTAP에 대한 계층화를 지원합니다.

- * 로컬 클러스터 계층화. * 비활성 데이터는 클러스터 LIF를 사용하여 로컬 클러스터에 있는 버킷으로 계층화합니다.
- * 원격 클러스터 계층화. * 비활성 데이터는 FabricPool 클라이언트의 IC LIF와 ONTAP 오브젝트 저장소의 데이터 LIF를 사용하여 기존 FabricPool 클라우드 계층과 비슷한 방식으로 원격 클러스터에 있는 버킷으로 계층화합니다.

ONTAP S3는 추가 하드웨어 및 관리 없이 기존 클러스터에서 S3 기능을 원하는 경우에 적합합니다. 300TB 이상의 배포를 위해 NetApp StorageGRID 소프트웨어는 계속해서 오브젝트 스토리지를 위한 주력 NetApp 솔루션으로 부상하고 있습니다. ONTAP 또는 StorageGRID를 클라우드 계층으로 사용할 때는 FabricPool 라이센스가 필요하지 않습니다.

Confluent 계층형 스토리지를 위한 **NetApp ONTAP**

모든 데이터 센터는 비즈니스 크리티컬 애플리케이션을 계속 실행하고 중요한 데이터를 안전하게 사용할 수 있어야 합니다. 새로운 NetApp AFF A900 시스템은 ONTAP 엔터프라이즈 에디션 소프트웨어와 복원력이 뛰어난 설계를 기반으로 합니다. NetApp의 새로운 초고속 NVMe 스토리지 시스템은 미션 크리티컬 운영의 중단을 제거하고, 성능 조정을 최소화하고, 랜섬웨어 공격으로부터 데이터를 보호합니다.

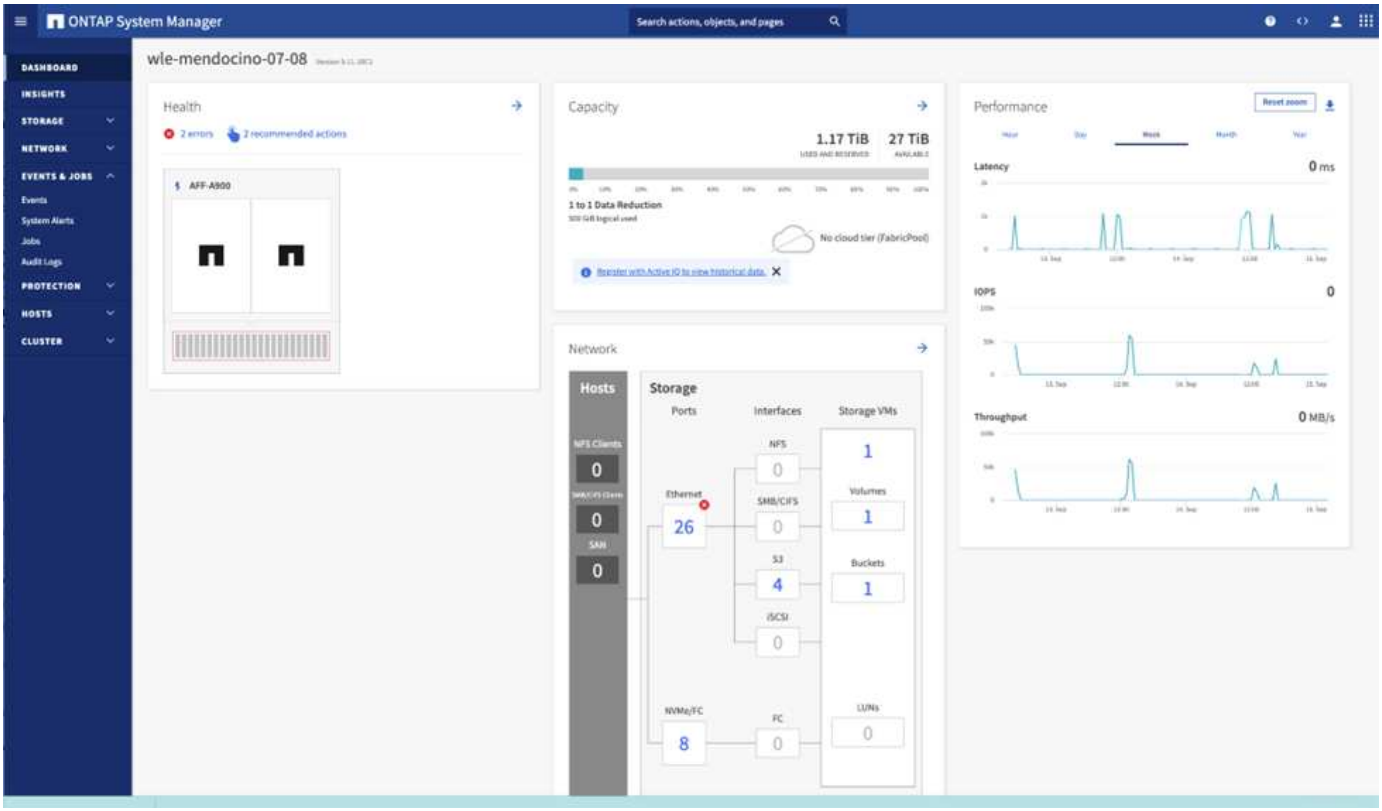
초기 구축부터 Confluent 클러스터의 확장까지, 귀사의 환경에서는 비즈니스 크리티컬 애플리케이션의 중단 없는 변화에 빠르게 적응해야 합니다. ONTAP 엔터프라이즈 데이터 관리, 서비스 품질(QoS) 및 성능 을 통해 고객 환경을 계획 및 조정할 수 있습니다.

NetApp ONTAP와 Confluent Tiered Storage를 함께 사용하면 ONTAP를 스케일아웃 스토리지 대상으로 활용하여 Apache Kafka 클러스터의 관리를 단순화하고 컴퓨팅 및 스토리지 리소스를 Confluent로 독립적으로 확장할 수 있습니다.

ONTAP S3 서버는 ONTAP의 성숙한 스케일아웃 스토리지 기능을 기반으로 구축되었습니다. ONTAP 클러스터는 S3 버킷을 확장하여 ONTAP 클러스터에 새로 추가된 노드를 사용함으로써 원활하게 확장할 수 있습니다.

ONTAP 시스템 관리자를 사용하여 간단하게 관리할 수 있습니다

ONTAP System Manager는 브라우저 기반의 그래픽 인터페이스로, 단일 창에서 전 세계에 분산된 위치에서 ONTAP 스토리지 컨트롤러를 구성, 관리, 모니터링할 수 있습니다.



System Manager 및 ONTAP CLI를 사용하여 ONTAP S3를 구성 및 관리할 수 있습니다. System Manager를 사용하여 S3를 활성화하고 버킷을 생성할 때 ONTAP은 단순한 구성에 대한 모범 사례 기본값을 제공합니다. CLI에서 S3 서버 및 버킷을 구성할 경우에도 원하는 경우 System Manager로 관리하거나 그 반대로 구성할 수 있습니다.

System Manager를 사용하여 S3 버킷을 생성하는 경우, ONTAP은 시스템에서 가장 가용성이 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 대신 사용자 지정 QoS 정책 그룹 또는 정책 그룹을 지정할 수 있습니다.

사전 정의된 적응형 QoS 정책 그룹은 다음과 같습니다.

- * Extreme. * 가장 짧은 지연 시간과 최고의 성능이 필요한 애플리케이션에 사용됩니다.
- 성능. * 적절한 성능 요구사항과 지연 시간이 필요한 애플리케이션에 사용됩니다.
- * 값. * 처리량과 용량이 지연 시간보다 더 중요한 애플리케이션에 사용됩니다.
- * 사용자 정의. * 사용자 정의 QoS 정책을 지정하거나 QoS 정책을 지정하지 않습니다.

계층화에 * 사용을 선택하면 성능 서비스 수준이 선택되지 않으며 시스템은 계층형 데이터에 대해 최적의 성능을 갖춘 저비용 미디어를 선택합니다.

ONTAP는 가장 적합한 디스크가 있는 로컬 계층에서 이 버킷을 프로비저닝하려고 시도하여 선택한 서비스 수준을 충족시킵니다. 그러나 버킷에 포함할 디스크를 지정해야 하는 경우 로컬 계층(애그리게이트)을 지정하여 CLI에서 S3 오브젝트 스토리지를 구성하는 것이 좋습니다. CLI에서 S3 서버를 구성할 경우에도 원할 경우 System Manager로 관리할 수 있습니다.

버킷에 사용할 애그리게이트를 지정할 수 있는 기능은 CLI를 통해서만 지정할 수 있습니다.

유창하게

Confluent Platform은 지속적인 실시간 스트림으로 데이터에 쉽게 액세스, 저장 및 관리할 수 있는 포괄적인 데이터 스트리밍 플랫폼입니다. Apache Kafka를 처음 개발한 Confluent는 Kafka 관리 또는 모니터링의 부담을 덜면서 엔터프라이즈급 기능을 통해 Kafka의 이점을 확장해 줍니다. 현재 Fortune 100대 기업 중 80% 이상이 데이터 스트리밍 기술을 사용하고 있으며 대부분 Confluent를 사용하고 있습니다.

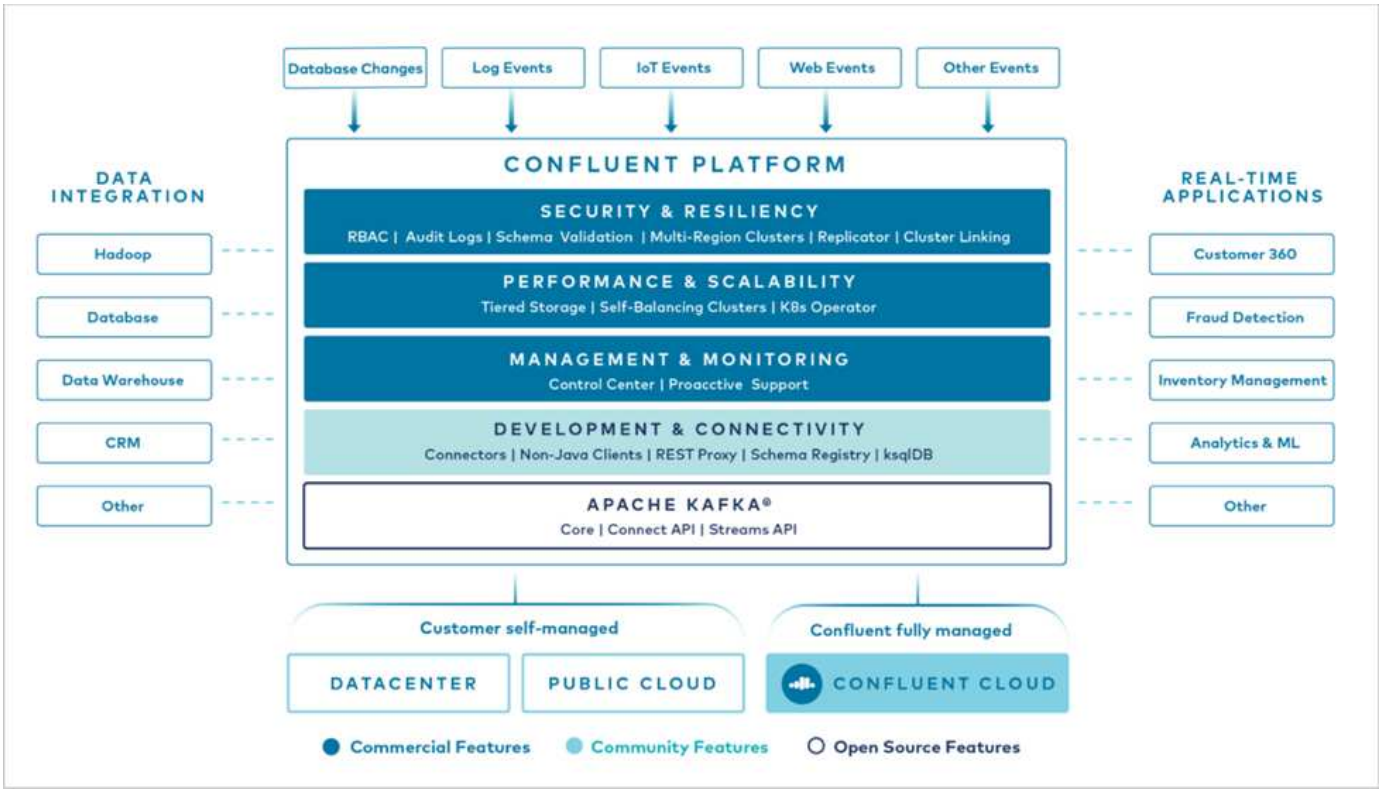
왜 **Confluent**인가**?**

Confluorent는 기록 데이터와 실시간 데이터를 단일 중앙 데이터 소스에 통합하여 완전히 새로운 범주의 최신 이벤트 기반 애플리케이션을 쉽게 구축하고, 범용 데이터 파이프라인을 구축하며, 완전한 확장성, 성능, 안정성으로 강력한 새 사용 사례를 활용할 수 있도록 지원합니다.

Contfluent는 어떤 용도로 사용됩니까**?**

Confluent Platform을 사용하면 데이터가 다른 시스템 간에 어떻게 전송 또는 통합되는지 등의 기본 메커니즘을 걱정하지 않고 데이터에서 비즈니스 가치를 창출하는 방법에 집중할 수 있습니다. 특히 Confluent Platform은 데이터 소스를 Kafka에 연결하고 스트리밍 애플리케이션을 구축하며 Kafka 인프라의 보안, 모니터링 및 관리를 간소화합니다. 현재 Confluent Platform은 금융 서비스, 옴니채널 소매, 자율 자동차, 사기 탐지, 마이크로서비스, IoT 등 다양한 산업 전반의 다양한 사용 사례에 사용됩니다.

다음 그림에서는 Confluent 플랫폼의 구성 요소를 보여 줍니다.



Confluent 이벤트 스트리밍 기술 개요

Confluent Platform의 핵심은 입니다 "카프카"가장 인기 있는 오픈 소스 분산 스트리밍 플랫폼입니다. Kafka의 주요 기능은 다음과 같습니다.

- 레코드 스트림을 게시하고 구독합니다.
- 내결함성이 있는 방식으로 레코드 스트림을 저장합니다.
- 레코드 스트림을 처리합니다.

즉시 사용할 수 있는 Confluorent Platform에는 스키마 레지스트리, REST 프록시, 총 100개 이상의 사전 구축된 Kafka 커넥터 및 ksqlDB도 포함되어 있습니다.

Confluent 플랫폼 엔터프라이즈 기능 개요

- * Confluent Control Center. * Kafka 관리 및 모니터링을 위한 UI 기반 시스템. Kafka Connect를 쉽게 관리하고 다른 시스템에 대한 연결을 생성, 편집 및 관리할 수 있습니다.
- Kubernetes를 위한 * Contfluent. * Kubernetes를 위한 Confluent는 Kubernetes 운영자입니다. Kubernetes 운영자는 특정 플랫폼 애플리케이션에 대한 고유한 기능과 요구 사항을 제공하여 Kubernetes의 오케스트레이션 기능을 확장합니다. Confluent Platform의 경우, Kubernetes에서 Kafka의 구축 프로세스를 크게 간소화하고 일반적인 인프라 라이프사이클 작업을 자동화할 수 있습니다.
- * Kafka 연결 커넥터 * 커넥터는 Kafka Connect API를 사용하여 Kafka를 데이터베이스, 키 값 저장소, 검색 인덱스 및 파일 시스템 등의 다른 시스템에 연결합니다. Confluorent Hub에는 가장 널리 사용되는 데이터 소스 및 싱크에 대한 다운로드 가능한 커넥터가 있습니다. 여기에는 Confluorent Platform이 포함된 이러한 커넥터의 전체 테스트 및 지원 버전이 포함됩니다. 자세한 내용은 을 참조하십시오 "여기".
- * 자체 밸런싱 클러스터 * 는 자동화된 로드 밸런싱, 장애 감지 및 자동 복구를 제공합니다. 또한 수동 조정 없이 필요에 따라 브로커를 추가하거나 해체할 수 있도록 지원합니다.
- * 연결 클러스터. * 직접 클러스터를 연결하고 링크 브리지를 통해 클러스터 간에 주제를 미러링합니다. 클러스터 링크를 사용하면 멀티 데이터 센터, 멀티 클러스터, 하이브리드 클라우드 구축을 간편하게 설정할 수 있습니다.
- * Confluorent auto data balancer. * 브로커 수, 파티션 크기, 파티션 수 및 클러스터 내의 리더 수에 대한 클러스터를 모니터링합니다. 균형 조정을 통해 트래픽을 재조정함으로써 운영 워크로드에 미치는 영향을 최소화하면서 클러스터 전체에서 짝수 워크로드를 생성할 수 있습니다.
- * Confluent Replicator. * 여러 데이터 센터에서 여러 Kafka 클러스터를 훨씬 쉽게 유지 관리할 수 있습니다.
- * 계층형 스토리지. * 즐겨 사용하는 클라우드 공급자를 사용하여 대량의 Kafka 데이터를 저장할 수 있는 옵션을 제공하므로 운영 부담과 비용이 줄어듭니다. 계층형 스토리지를 사용하면 비용 효율적인 오브젝트 스토리지에 데이터를 보관하고 더 많은 컴퓨팅 리소스가 필요할 때만 브로커를 확장할 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. JMS를 사용하는 레거시 애플리케이션이 있고 기존 JMS 메시지 브로커를 Kafka로 교체하려는 경우 유용합니다.
- * Confluent MQTT proxy. * 중간에 MQTT 브로커가 없어도 MQTT 장치 및 게이트웨이에서 Kafka에 직접 데이터를 게시할 수 있는 방법을 제공합니다.
- * Confluent 보안 플러그인 * Confluent 보안 플러그인은 다양한 Confluent 플랫폼 도구 및 제품에 보안 기능을 추가하는 데 사용됩니다. 현재 Confluent REST 프록시에 사용할 수 있는 플러그인이 있어 수신 요청을 인증하고 인증된 보안 주체를 Kafka에 요청에 전파할 수 있습니다. 이렇게 하면 Confluent REST 프록시 클라이언트가 Kafka 브로커의 멀티테넌트 보안 기능을 활용할 수 있습니다.

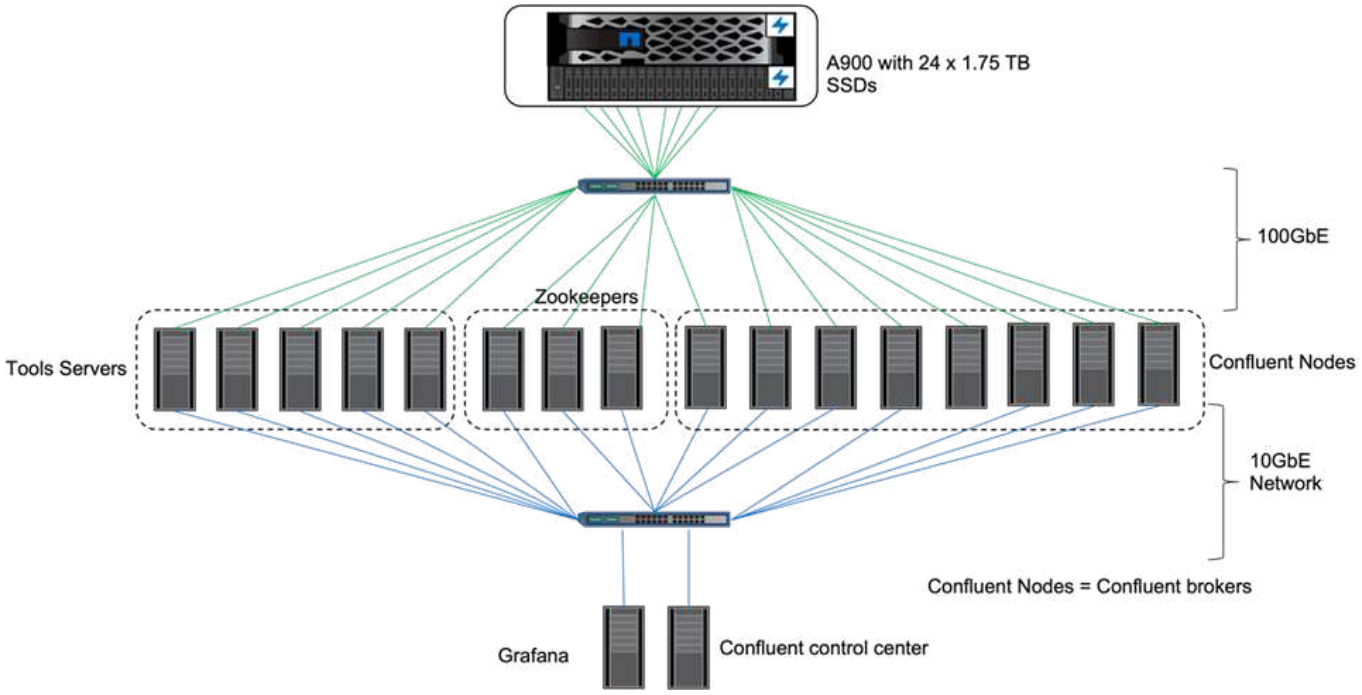
Confluent 성능 검증

NetApp ONTAP의 계층형 스토리지를 위한 Confluent Platform을 사용하여 검증을 수행했습니다. NetApp과 Confluent 팀은 이 검증을 함께 수행하여 IT에 필요한 테스트 사례를 실행했습니다.

Confluent 설정

설치를 위해 256GB RAM과 16개의 CPU를 갖춘 3대의 zookeepers, 5개의 브로커, 5대의 테스트 서버를 사용했습니다. NetApp 스토리지의 경우 ONTAP와 AFF A900 HA 쌍을 사용했습니다. 스토리지와 브로커는 100GbE 연결을 통해 연결되었습니다.

다음 그림에서는 계층형 스토리지 검증에 사용된 구성의 네트워크 토폴로지를 보여 줍니다.



도구 서버는 Confluent 노드로 이벤트를 보내거나 받는 응용 프로그램 클라이언트로 작동합니다.

Confluent 계층형 스토리지 구성

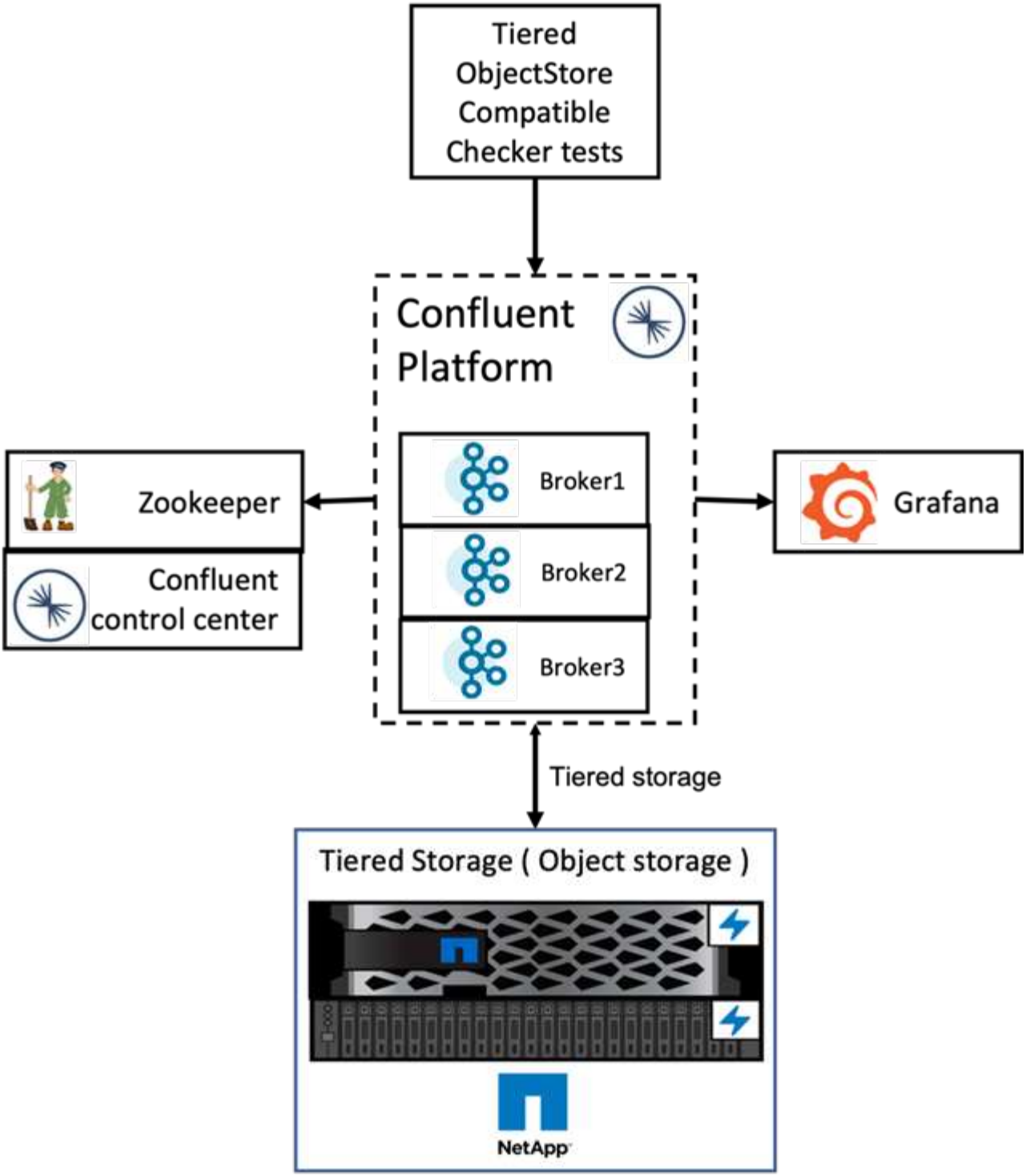
다음과 같은 테스트 매개 변수를 사용했습니다.

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

검증을 위해 HTTP 프로토콜과 함께 ONTAP를 사용했지만 HTTPS도 작동했습니다. 액세스 키와 비밀 키는 confluent.tier.s3.cred.file.path 매개 변수에 제공된 파일 이름에 저장됩니다.

NetApp 스토리지 컨트롤러 – ONTAP

검증을 위해 ONTAP에서 단일 HA 쌍 구성을 구성했습니다.



확인 결과

검증을 위해 다음 5가지 테스트 사례를 완료했습니다. 첫 번째 두 테스트는 기능 테스트이고 나머지 세 가지는 성능 테스트입니다.

객체 저장소 정확도 테스트

이 테스트는 API 호출을 사용하여 계층형 스토리지에 사용되는 객체 저장소에서 GET, PUT 및 DELETE 등의 기본 작업을 수행합니다.

계층화 기능 정확도 테스트

이 테스트에서는 오브젝트 스토리지의 엔드 투 엔드 기능을 검사합니다. 토픽을 생성하고, 새로 생성된 토픽에 대한 이벤트 스트림을 생성하고, 브로커가 세그먼트를 오브젝트 스토리지에 아카이브하고, 이벤트 스트림을 소비하며, 소비된 스트림이 생성된 스트림과 일치하는지 확인합니다. 객체 저장소 결함 주입 여부와 관계없이 이 테스트를 수행했습니다. ONTAP의 노드 중 하나에서 서비스 관리자 서비스를 중지하고 엔드 투 엔드 기능이 오브젝트 스토리지에서 작동하는지 확인하여 노드 장애를 시뮬레이션했습니다.

계층 가져오기 벤치마크

이 테스트에서는 계층형 오브젝트 스토리지의 읽기 성능을 검증하고 벤치마크에 의해 생성된 세그먼트에서 과부하된 읽기 요청 범위를 검사했습니다. 이 벤치마크에서 Confluent는 계층 가져오기 요청을 처리하기 위해 사용자 지정 클라이언트를 개발했습니다.

생산 - 워크로드 생성기 사용

이 테스트에서는 세그먼트 아카이브를 통해 오브젝트 저장소에서 쓰기 워크로드를 간접적으로 생성합니다. 소비자 그룹이 세그먼트를 가져올 때 객체 스토리지에서 읽기 워크로드(세그먼트 읽기)가 생성되었습니다. 이 워크로드는 TOCC 스크립트에 의해 생성되었습니다. 이 테스트에서는 오브젝트 저장소에서 병렬 스레드의 읽기 및 쓰기 성능을 확인했습니다. 계층화 기능 정확도 테스트에서 보았듯이, 객체 저장소 결함 주입을 사용하여 테스트했으며 포함하지 않았습니다.

보존 워크로드 생성기

이 테스트에서는 무거운 주제 보존 워크로드에서 오브젝트 스토리지의 삭제 성능을 검사했습니다. 보존 워크로드는 테스트 주제와 동시에 많은 메시지를 생성하는 TOCC 스크립트를 사용하여 생성되었습니다. 테스트 주제는 공격적인 크기 기반 및 시간 기반 보존 설정으로 구성되었으며, 이로 인해 이벤트 스트림이 개체 저장소에서 지속적으로 제거됩니다. 그런 다음 세그먼트가 아카이브되었습니다. 이로 인해 오브젝트 저장소 삭제 작업의 브로커링 및 컬렉션에 의해 오브젝트 스토리지가 많이 삭제되었습니다.

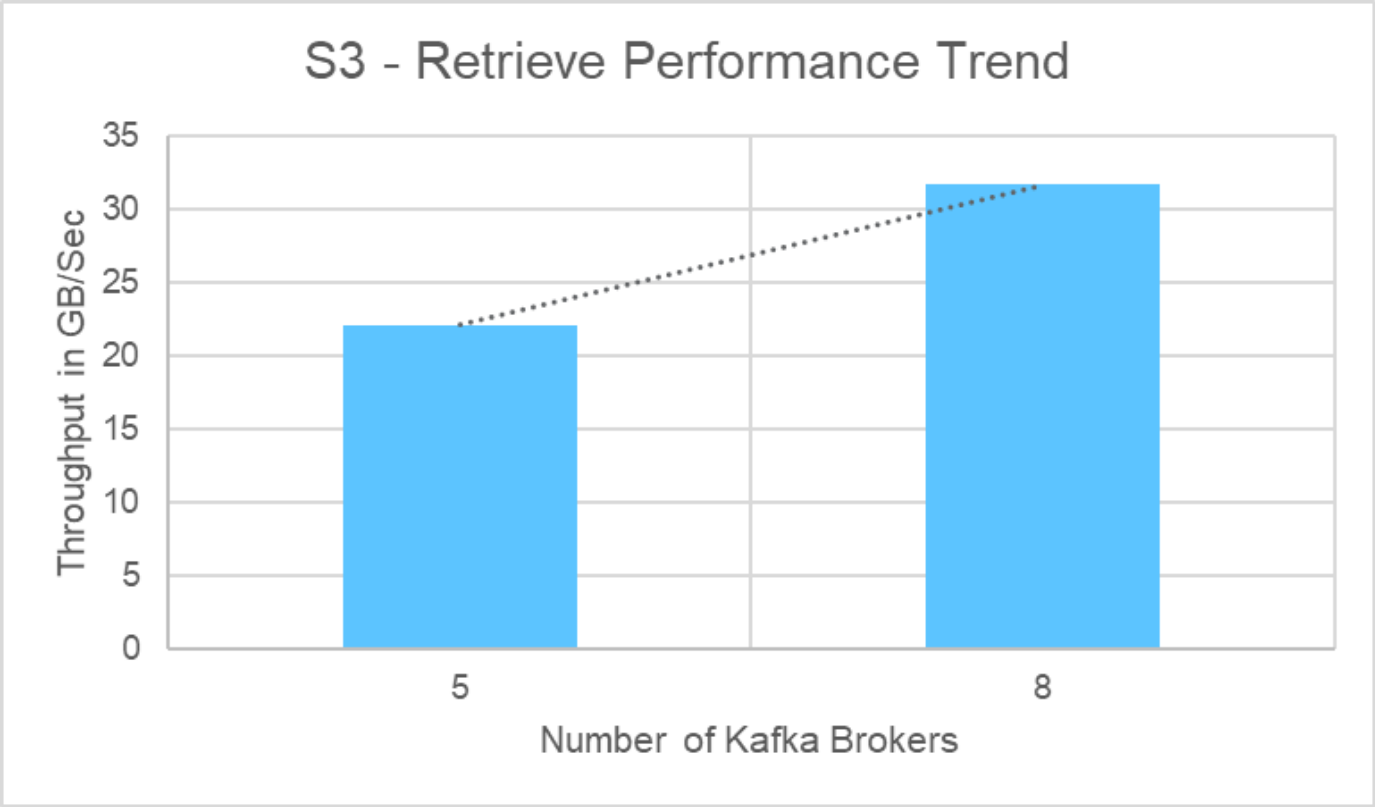
검증 세부 정보는 를 참조하십시오 "유창하게" 웹 사이트.

농작물 사용 워크로드 생성기를 사용한 성능 테스트

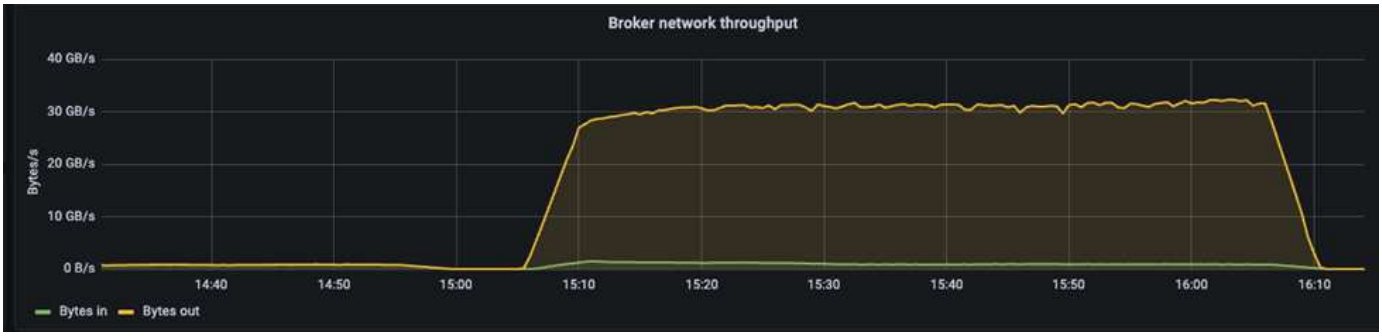
AFF A900 HA 쌍 NetApp 스토리지 컨트롤러 1개를 사용하여 농작물 사용 워크로드 중 5개 또는 8개의 브로커 노드로 계층형 스토리지 테스트를 수행했습니다. 테스트에 따르면 AFF A900 리소스 활용률이 100%에 이를 때까지 브로커 노드의 수에 따라 완료 시간과 성능 결과가 확장되었습니다. ONTAP 스토리지 컨트롤러를 설치하려면 최소 하나의 HA 쌍이 필요했습니다.

S3 검색 작업에 대한 성능은 Confluent 브로커 노드의 수에 따라 선형으로 향상되었습니다. ONTAP 스토리지 컨트롤러는 단일 배포에서 최대 12개의 HA 쌍을 지원합니다.

다음 그래프에는 브로커 노드 5개 또는 8개가 결합된 S3 계층화 트래픽이 나와 있습니다. AFF A900 단일 HA Pair 성능을 극대화했습니다.



다음 그래프는 약 31.74GBps의 Kafka 처리량을 보여 줍니다.



또한 ONTAP 스토리지 컨트롤러 'perfstat' 보고서에서 유사한 처리량도 관찰되었습니다.

```
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.

- ONTAP의 경우 가능하면 GET SIZE >= 1MB 를 사용합니다.
- 브로커 노드의 erver.properties` 에서 num.network.threads` 및 num.io.threads를 증가시키면 계층화 작업이 S3 계층으로 증가하게 됩니다. 이러한 결과는 num.network.threads` 및 num.io.threads가 32로 설정된 것입니다.

- S3 버킷은 구성원 애그리게이트당 8개의 구성요소를 목표로 해야 합니다.
- S3 트래픽을 구동하는 이더넷 링크는 스토리지와 클라이언트 모두에서 가능하면 MTU 9k 를 사용해야 합니다.

결론

이 검증 테스트는 NetApp ONTAP 스토리지 컨트롤러에 유창한 confluent에서 31.74GBps의 계층화 처리량에 도달했습니다.

추가 정보를 찾을 수 있는 위치

이 문서에 설명된 정보에 대해 자세히 알아보려면 다음 문서 및/또는 웹 사이트를 검토하십시오.

- Confluent란 무엇입니까?

  "https://www.confluent.io/apache-kafka-vs-confluent/"

- S3 싱크 매개 변수 세부 정보

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

- 아파치 카프카

  "https://en.wikipedia.org/wiki/Apache_Kafka"

- S3 in ONTAP 모범 사례

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

- S3 오브젝트 스토리지 관리

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

- NetApp 제품 설명서

  "https://www.netapp.com/support-and-training/documentation/"

저작권 정보

Copyright © 2024 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 라이센스와 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 라이센스도 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 라이센스를 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 라이센스는 DFARS 조항 252.227-7015(b)(2014년 2월)에 명시된 권한으로 제한됩니다.

상표 정보

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