Splunk SmartStore가 포함된 NetApp StorageGRID

NetApp artificial intelligence solutions

NetApp
August 18, 2025

This PDF was generated from https://docs.netapp.com/ko-kr/netapp-solutions-ai/data-analytics/stgr-splunkss-introduction.html on August 18, 2025. Always check docs.netapp.com for the latest.

목차

Splunk SmartStore가 포함된 NetApp StorageGRID . . . 1
TR-4869: Splunk SmartStore를 탑재한 NetApp StorageGRID . . . 1
개요 . . . 1
NetApp StorageGRID 정보 . . . 1
Splunk Enterprise 소개 . . . 3
Splunk SmartStore 소개 . . . 3
솔루션 개요 . . . 3
NetApp StorageGRID . . . 3
스플렁크 엔터프라이즈 . . . 3
스플렁크 스마트스토어 . . . 4
이 솔루션의 이점 . . . 4
Splunk 아키텍처 . . . 5
주요 정의 . . . 5
Splunk 분산 배포 . . . 6
스플렁크 스마트스토어 . . . 7
Splunk SmartStore 데이터 흐름 . . . 7
소프트웨어 요구 사항 . . . 8
단일 및 다중 사이트 요구 사항 . . . 9
하드웨어 요구 사항 . . . 11
스플렁크 디자인 . . . 13
Splunk SmartStore를 위한 유연한 StorageGRID 기능 . . . 15
Grid Manager를 통한 간편한 관리 . . . 15
Splunk용 NetApp StorageGRID 앱 . . . 16
ILM 정책 . . . 17
성능 . . . 17
로드 밸런서 및 엔드포인트 구성 . . . 17
지능형 계층화 및 비용 절감 . . . 17
단일 사이트 SmartStore 성능 . . . 18
구성 . . . 21
SmartStore 원격 매장 성능 검증 . . . 21
StorageGRID 성능 . . . 26
StorageGRID 하드웨어 사용 . . . 27
NetApp 스토리지 컨트롤러를 탑재한 SmartStore - 고객 혜택 . . . 28
결론 . . . 29
추가 정보를 찾을 수 있는 곳 . . . 29

Splunk SmartStore가 포함된 NetApp StorageGRID

TR-4869: Splunk SmartStore를 탑재한 NetApp StorageGRID

Splunk Enterprise는 보안, IT, DevOps 팀 전반에서 성과를 이끌어내는 시장 선도적인 보안 정보 및 이벤트 관리(SIEM) 솔루션입니다.

개요

데이터 양은 기하급수적으로 증가하고 있으며, 이 방대한 리소스를 활용할 수 있는 기업에게는 엄청난 기회가 창출되고 있습니다. Splunk Enterprise는 더욱 다양한 사용 사례에서 채택이 확대되고 있습니다. 사용 사례가 증가함에 따라 Splunk Enterprise가 수집하고 처리하는 데이터 양도 늘어납니다. Splunk Enterprise의 기존 아키텍처는 뛰어난 데이터 접근성과 가용성을 제공하는 분산형 확장형 디자인입니다. 그러나 이러한 아키텍처를 사용하는 기업은 급증하는 데이터 양을 수용하기 위해 확장하는 데 드는 비용 증가에 직면하게 됩니다.

NetApp StorageGRID 탑재된 Splunk SmartStore는 컴퓨팅과 스토리지를 분리하는 새로운 배포 모델을 제공하여 이러한 과제를 해결합니다. 이 솔루션은 고객이 단일 및 여러 사이트에 걸쳐 확장할 수 있도록 하여 Splunk Enterprise 환경에 대한 탁월한 확장성과 탄력성을 제공하는 동시에 컴퓨팅과 스토리지를 독립적으로 확장하고 비용 효율적인 클라우드 기반 S3 개체 스토리지에 지능형 계층화를 추가하여 비용을 절감합니다.

이 솔루션은 검색 성능을 유지하는 동시에 로컬 스토리지의 데이터 양을 최적화하여 컴퓨팅과 스토리지를 수요에 따라 확장할 수 있도록 합니다. SmartStore는 데이터 액세스 패턴을 자동으로 평가하여 실시간 분석을 위해 어떤 데이터에 액세스해야 하는지, 어떤 데이터를 비용이 덜 드는 S3 개체 스토리지에 저장해야 하는지 판별합니다.

이 기술 보고서는 NetApp 이 Splunk SmartStore 솔루션에 제공하는 이점을 간략하게 설명하고 사용자 환경에서 Splunk SmartStore를 설계하고 크기를 조정하기 위한 프레임워크를 보여줍니다. 그 결과, 매력적인 TCO를 제공하는 간단하고 확장 가능하며 탄력적인 솔루션이 탄생했습니다. StorageGRID 확장 가능하고 비용 효율적인 S3 프로토콜/API 기반 개체 스토리지(원격 스토리지라고도 함)를 제공하여 조직이 복원력을 높이는 동시에 더 낮은 비용으로 Splunk 솔루션을 확장할 수 있도록 합니다.

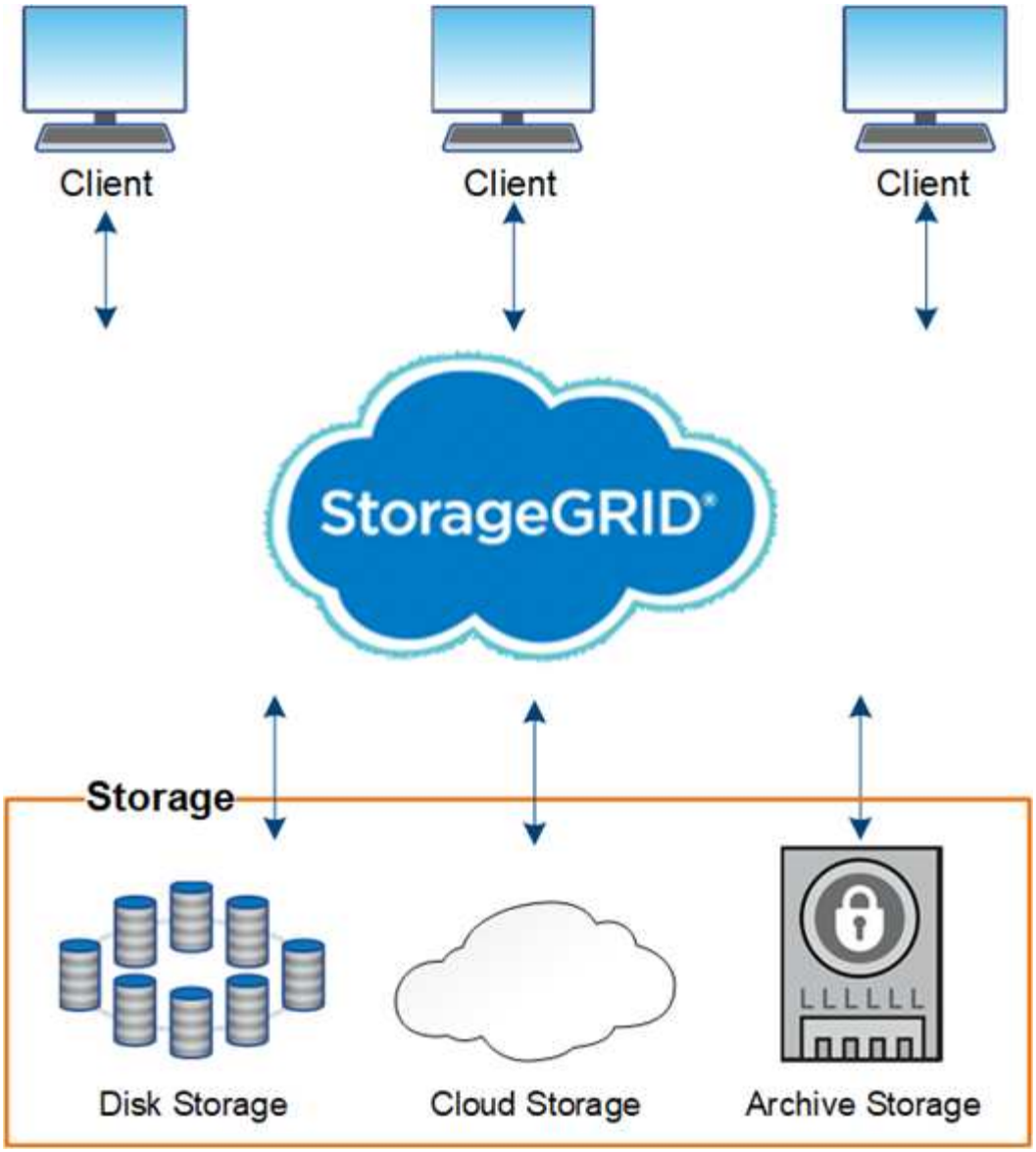
> Splunk SmartStore는 개체 스토리지를 원격 저장소 또는 원격 스토리지 계층이라고 합니다.

NetApp StorageGRID 정보

NetApp StorageGRID 대규모 아카이브, 미디어 저장소, 웹 데이터 저장소를 위한 소프트웨어 정의 객체 스토리지 솔루션입니다. NetApp StorageGRID 통해 업계를 선도하는 혁신과 데이터 관리 솔루션을 제공하는 데 있어 20년의 경험을 활용하고 있으며, 조직이 온프레미스와 퍼블릭, 프라이빗 또는 하이브리드 클라우드 배포에서 정보의 가치를 관리하고 극대화할 수 있도록 지원합니다.

StorageGRID 대규모 비정형 데이터를 위한 안전하고 내구성 있는 스토리지를 제공합니다. 통합된 메타데이터 기반 수명 주기 관리 정책은 데이터 수명 전반에 걸쳐 데이터가 저장되는 위치를 최적화합니다. 비용을 줄이기 위해 콘텐츠를 적절한 위치, 적절한 시간, 적절한 저장 계층에 배치합니다. 단일 네임스페이스를 사용하면 StorageGRID 스토리지의 지리적 위치에 관계없이 단일 호출을 통해 데이터에 액세스할 수 있습니다. 고객은 데이터 센터와 클라우드 인프라 간에 여러 StorageGRID 인스턴스를 배포하고 관리할 수 있습니다.

StorageGRID 시스템은 기존 및 차세대 클라이언트 애플리케이션과 통합 가능한 글로벌 분산형, 중복성, 이기종 노드로 구성됩니다.



IDC MarketScape는 최근 최신 보고서인 IDC MarketScape: Worldwide Object-Based Storage 2019 Vendor Assessment에서 NetApp 리더로 선정했습니다. 가장 까다로운 산업 분야에서 20년 가까이 생산 배포 경험을 바탕으로 StorageGRID 비정형 데이터 분야에서 인정받는 선두주자로 자리매김했습니다.

StorageGRID 사용하면 다음과 같은 이점을 얻을 수 있습니다.

- 단일 네임스페이스를 통해 데이터 센터와 클라우드 사이의 모든 위치에서 데이터에 액세스하기 위해 여러 StorageGRID 인스턴스를 배포하면 수백 페타바이트까지 쉽게 확장할 수 있습니다.
- 인프라 전반에 걸쳐 배포하고 중앙에서 관리할 수 있는 유연성을 제공합니다.
- 계층화된 Erasure Coding(EC)을 활용하여 15나인 내구성으로 탁월한 내구성을 제공합니다.
- Amazon S3 Glacier 및 Azure Blob과의 검증된 통합을 통해 더욱 다양한 하이브리드 멀티클라우드 기능을 구현하세요.
- 독점 API나 공급업체 종속 없이 변조 방지 데이터 보존을 통해 규제 의무를 충족하고 규정 준수를 용이하게 합니다.

StorageGRID 가장 복잡한 비정형 데이터 관리 문제를 해결하는 데 어떻게 도움이 될 수 있는지에 대한 자세한 내용은 다음을 참조하세요. "NetApp StorageGRID 홈페이지" .

Splunk Enterprise 소개

Splunk Enterprise는 데이터를 실제 업무로 전환하는 플랫폼입니다. 로그 파일, 웹사이트, 장치, 센서, 애플리케이션 등 다양한 소스에서 생성된 데이터는 Splunk 인덱서로 전송되어 구문 분석되므로 데이터에서 풍부한 통찰력을 얻을 수 있습니다. 이를 통해 데이터 침해를 식별하고, 고객 및 제품 동향을 파악하고, 인프라를 최적화할 수 있는 기회를 찾거나 다양한 사용 사례에 걸쳐 실행 가능한 통찰력을 창출할 수 있습니다.

Splunk SmartStore 소개

Splunk SmartStore는 Splunk 아키텍처의 이점을 확장하는 동시에 비용 효율적인 확장 기능을 단순화합니다. 컴퓨팅 및 스토리지 리소스를 분리하면 I/O에 최적화된 인덱서 노드가 생성되고, 캐시로 데이터의 하위 집합만 저장하므로 스토리지 요구 사항이 크게 줄어듭니다. 해당 리소스 중 하나만 필요한 경우 별도의 컴퓨팅이나 스토리지를 추가할 필요가 없으므로 상당한 비용 절감 효과를 얻을 수 있습니다. 비용 효율적이고 쉽게 확장 가능한 S3 기반 객체 스토리지를 사용하면 환경이 더욱 간소화되고 비용이 절감되며 더 방대한 데이터 세트를 유지 관리할 수 있습니다.

Splunk SmartStore는 다음을 포함한 조직에 상당한 가치를 제공합니다.

- 따뜻한 데이터를 비용 최적화된 S3 개체 스토리지로 이동하여 스토리지 비용 절감
- 스토리지와 컴퓨팅을 분리하여 원활하게 확장
- 탄력적인 클라우드 기반 스토리지를 활용하여 비즈니스 연속성을 간소화합니다.

솔루션 개요

이 페이지에서는 NetApp StorageGRID, Splunk Enterprise, Splunk SmartStore를 비롯하여 이 솔루션을 완성하는 데 사용된 구성 요소에 대해 설명합니다.

NetApp StorageGRID

NetApp StorageGRID 는 고성능이고 비용 효율적인 객체 스토리지 플랫폼입니다. 분산형 노드 기반 그리드 아키텍처를 사용하여 지능적이고 정책 기반의 글로벌 데이터 관리를 제공합니다. 유비쿼터스 글로벌 객체 네임스페이스와 정교한 데이터 관리 기능을 결합하여 페타바이트 규모의 비정형 데이터와 수십억 개의 객체를 간편하게 관리할 수 있습니다. 단일 호출 객체 액세스는 여러 사이트로 확장되어 고가용성 아키텍처를 단순화하는 동시에 사이트나 인프라 중단에 관계없이 지속적인 객체 액세스를 보장합니다.

멀티테넌시를 통해 여러 클라우드 및 엔터프라이즈 비정형 데이터 애플리케이션을 동일한 그리드 내에서 안전하게 서비스할 수 있어 StorageGRID 의 ROI와 사용 사례가 늘어납니다. 메타데이터 기반 개체 수명 주기 정책을 사용하면 여러 서비스 수준을 생성하여 여러 지역에 걸쳐 내구성, 보호, 성능 및 지역성을 최적화할 수 있습니다. 사용자는 요구 사항이 변경됨에 따라 중단 없이 정책을 조정하고 데이터 환경을 재정비할 수 있습니다.

SmartStore는 원격 스토리지 계층으로 StorageGRID 활용하고 고객이 지리적으로 분산된 여러 사이트를 배포하여 강력한 가용성과 내구성을 확보할 수 있도록 하며, 이는 단일 개체 네임스페이스로 제공됩니다. 이를 통해 Splunk SmartStore는 StorageGRID 의 고성능, 고밀도 용량, 단일 URL을 사용하여 여러 물리적 사이트에 걸쳐 수백 개의 노드로 확장하는 기능을 활용하여 개체와 상호 작용할 수 있습니다. 이 단일 URL을 사용하면 단일 사이트를 넘어서도 중단 없이 저장소 확장, 업그레이드 및 수리를 수행할 수 있습니다. StorageGRID 의 고유한 데이터 관리 정책 엔진은 최적화된 수준의 성능과 내구성을 제공하며 데이터 지역성 요구 사항을 준수합니다.

스플렁크 엔터프라이즈

기계에서 생성된 데이터의 수집 및 분석 분야를 선도하는 Splunk는 운영 분석 기능을 통해 IT를 간소화하고 현대화하는 데 도움을 줍니다. 또한 비즈니스 분석, 보안, IoT 사용 사례로 확장됩니다. 스토리지는 Splunk 소프트웨어를 성공적으로

배포하는 데 중요한 요소입니다.

기계에서 생성된 데이터는 가장 빠르게 성장하는 빅데이터 유형입니다. 이 형식은 예측 불가능하며 다양한 출처에서 나오는 경우가 많고, 종종 빠른 속도로 대량으로 발생합니다. 이러한 작업 부하 특성은 종종 디지털 배기라고 불립니다. Splunk SmartStore는 이러한 데이터를 이해하는 데 도움이 되며, 가장 비용 효율적인 스토리지 계층에 핫 데이터와 웜 데이터를 최적으로 배치하기 위한 스마트한 데이터 계층화를 제공합니다.

스플렁크 스마트스토어

Splunk SmartStore는 StorageGRID 와 같은 개체 스토리지(원격 스토리지 또는 원격 스토리지 계층이라고도 함)를 사용하여 S3 프로토콜을 사용하여 웜 데이터를 저장하는 인덱서 기능입니다.

배포된 데이터 볼륨이 증가함에 따라 일반적으로 저장소에 대한 수요가 컴퓨터 리소스에 대한 수요를 앞지릅니다. SmartStore를 사용하면 컴퓨팅과 스토리지를 별도로 확장하여 인덱서 스토리지와 컴퓨팅 리소스를 비용 효율적으로 관리할 수 있습니다.

SmartStore는 S3 프로토콜과 캐시 관리자를 사용하여 원격 스토리지 계층을 도입했습니다. 이러한 기능을 사용하면 데이터를 인덱서나 원격 저장소에 로컬로 저장할 수 있습니다. 인덱서에 있는 캐시 관리자는 인덱서와 원격 스토리지 계층 간의 데이터 이동을 관리합니다. 데이터는 버킷 메타데이터와 함께 버킷(핫 및 웜)에 저장됩니다.

SmartStore를 사용하면 인덱서 저장소 공간을 최소한으로 줄이고 대부분의 데이터가 원격 저장소 계층에 있으므로 I/O 최적화된 컴퓨팅 리소스를 선택할 수 있습니다. 인덱서는 요청되고 예측된 결과를 반환하는 데 필요한 최소한의 데이터 양을 나타내는 로컬 캐시를 유지 관리합니다. 로컬 캐시에는 핫 버킷, 활성 또는 최근 검색에 참여하는 웜 버킷의 사본 및 버킷 메타데이터가 포함되어 있습니다.

StorageGRID 탑재된 Splunk SmartStore를 사용하면 고객은 고성능 및 비용 효율적인 원격 스토리지를 통해 환경을 점진적으로 확장할 수 있으며, 동시에 전체 솔루션에 높은 수준의 탄력성을 제공합니다. 이를 통해 고객은 더 많은 인덱서가 필요하거나, 데이터 보존 기간을 변경하거나, 중단 없이 수집 속도를 높여야 하는 경우, 언제든지 원하는 수량만큼 구성 요소(핫 스토리지 및/또는 웜 S3 스토리지)를 추가할 수 있습니다.

이 솔루션의 이점

이 솔루션을 사용하면 단일 및 다중 사이트 배포에서 사용자 수나 수집 속도 측면에서 증가하는 수요를 충족하기 위해 컴퓨팅, 핫 스토리지 또는 S3 리소스를 추가할 수 있습니다.

- 성능**.** Splunk SmartStore와 NetApp StorageGRID 결합하면 객체 스토리지를 사용하여 핫 버킷과 웜 버킷 간에 데이터를 빠르게 마이그레이션할 수 있습니다. StorageGRID 대규모 객체 작업 부하에 대해 빠른 성능을 제공하여 마이그레이션 프로세스를 가속화합니다.
- 다중 사이트 준비 완료**.** StorageGRID 분산 아키텍처를 사용하면 Splunk SmartStore가 단일 글로벌 네임스페이스를 통해 단일 사이트와 여러 사이트에 배포를 확장할 수 있으며, 데이터가 어디에 있든 관계없이 모든 사이트에서 데이터에 액세스할 수 있습니다.
- 확장성이 향상되었습니다**.** Splunk 환경에서 변화하는 요구 사항과 수요를 충족하기 위해 컴퓨팅 리소스와 별도로 스토리지 리소스를 확장하여 TCO를 개선합니다.
- 용량**.** StorageGRID 사용하여 단일 네임스페이스를 560PB 이상으로 확장하여 Splunk 배포에서 빠르게 증가하는 볼륨을 처리하세요.
- 데이터 가용성**.** 데이터의 비즈니스 가치가 변화함에 따라 동적으로 조정할 수 있는 메타데이터 기반 정책을 통해 데이터 가용성, 성능, 지리적 분포, 보존, 보호 및 저장 비용을 최적화하세요.

로컬(핫) 및 원격(웜) 스토리지 간 버킷 복사본 전송을 처리하는 인덱서의 구성 요소인 SmartStore 캐시를 사용하여 성능을 향상시킵니다. 이 솔루션에 대한 Splunk 크기 조정은 다음을 기반으로 합니다. "Splunk에서 제공하는

가이드라인" . 이 솔루션을 사용하면 단일 및 다중 사이트 배포에서 사용자 수나 수집 속도 측면에서 증가하는 수요를 충족하기 위해 컴퓨팅, 핫 스토리지 또는 S3 리소스를 추가할 수 있습니다.

Splunk 아키텍처

이 섹션에서는 Splunk 아키텍처에 대해 설명합니다. 여기에는 주요 정의, Splunk 분산 배포, Splunk SmartStore, 데이터 흐름, 하드웨어 및 소프트웨어 요구 사항, 단일 및 다중 사이트 요구 사항 등이 포함됩니다.

주요 정의

다음 두 표에는 분산형 Splunk 배포에 사용되는 Splunk 및 NetApp 구성 요소가 나열되어 있습니다.

이 표에는 분산형 Splunk Enterprise 구성을 위한 Splunk 하드웨어 구성 요소가 나열되어 있습니다.

Splunk 구성 요소	일
인덱서	Splunk Enterprise 데이터 저장소
유니버설 포워더	데이터 수집 및 인덱서에 데이터 전달을 담당합니다.
검색 헤드	인덱서에서 데이터를 검색하는 데 사용되는 사용자 프런트 엔드
클러스터 마스터	인덱서 및 검색 헤드의 Splunk 설치를 관리합니다.
모니터링 콘솔	전체 배포에 사용되는 중앙 모니터링 도구
라이센스 마스터	라이선스 마스터는 Splunk Enterprise 라이선스를 처리합니다.
배포 서버	구성을 업데이트하고 앱을 처리 구성 요소에 배포합니다.
저장 구성요소	일
NetApp AFF	핫 티어 데이터를 관리하는 데 사용되는 올플래시 스토리지입니다. 로컬 스토리지라고도 함.
NetApp StorageGRID	웜 티어 데이터를 관리하는 데 사용되는 S3 개체 스토리지입니다. SmartStore에서 핫 티어와 웜 티어 간에 데이터를 이동하는 데 사용됩니다. 원격 저장소라고도 함.

이 표는 Splunk 스토리지 아키텍처의 구성 요소를 나열합니다.

Splunk 구성 요소	일	책임 구성 요소
스마트스토어	인덱서에게 로컬 스토리지에서 개체 스토리지로 데이터를 계층화하는 기능을 제공합니다.	스플렁크
더운	유니버설 포워더가 새로 작성된 데이터를 저장하는 도착 지점입니다. 저장소는 쓰기가 가능하며, 데이터는 검색이 가능합니다. 이 데이터 계층은 일반적으로 SSD나 빠른 HDD로 구성됩니다.	ONTAP

Splunk 구성 요소	일	책임 구성 요소
캐시 관리자	인덱싱된 데이터의 로컬 캐시를 관리하고, 검색이 발생하면 원격 저장소에서 웜 데이터를 가져오고, 가장 덜 자주 사용되는 데이터를 캐시에서 제거합니다.	스마트스토어
따뜻한	데이터는 논리적으로 버킷에 롤링되고, 먼저 핫 티어에서 웜 티어로 이름이 변경됩니다. 이 계층의 데이터는 보호되며, 핫 계층과 마찬가지로 더 큰 용량의 SSD나 HDD로 구성될 수 있습니다. 일반적인 데이터 보호 솔루션을 사용하여 증분 백업과 전체 백업이 모두 지원됩니다.	StorageGRID

Splunk 분산 배포

많은 컴퓨터에서 데이터가 생성되는 대규모 환경을 지원하려면 대량의 데이터를 처리해야 합니다. 많은 사용자가 데이터를 검색해야 하는 경우 Splunk Enterprise 인스턴스를 여러 컴퓨터에 분산하여 배포를 확장할 수 있습니다. 이를 분산 배포라고 합니다.

일반적인 분산 배포에서 각 Splunk Enterprise 인스턴스는 특수화된 작업을 수행하며 주요 처리 기능에 해당하는 3개의 처리 계층 중 하나에 상주합니다.

다음 표에는 Splunk Enterprise 처리 계층이 나열되어 있습니다.

층	요소	설명
데이터 입력	포워더	포워더는 데이터를 사용한 후 해당 데이터를 인덱서 그룹에 전달합니다.
인덱싱	인덱서	인덱서는 일반적으로 여러 포워더로부터 수신하는 수신 데이터를 인덱싱합니다. 인덱서는 데이터를 이벤트로 변환하고 이벤트를 인덱스에 저장합니다. 인덱서는 검색 헤드의 검색 요청에 응답하여 인덱싱된 데이터를 검색합니다.
검색 관리	검색 헤드	검색 헤드는 검색을 위한 중앙 리소스 역할을 합니다. 클러스터의 검색 헤드는 상호 교환이 가능하며 검색 헤드 클러스터의 모든 멤버에서 동일한 검색, 대시보드, 지식 개체 등에 액세스할 수 있습니다.

다음 표는 분산형 Splunk Enterprise 환경에서 사용되는 중요한 구성 요소를 나열합니다.

요소	설명	책임
인덱스 클러스터 마스터	인덱서 클러스터의 활동 및 업데이트를 조정합니다.	인덱스 관리

요소	설명	책임
인덱스 클러스터	서로 데이터를 복제하도록 구성된 Splunk Enterprise 인덱서 그룹	인덱싱
검색 헤드 배포자	클러스터 마스터에 대한 배포 및 업데이트를 처리합니다.	검색 헤드 관리
검색 헤드 클러스터	검색을 위한 중앙 리소스 역할을 하는 검색 헤드 그룹	검색 관리
로드 밸런서	클러스터된 구성 요소에서 검색 헤드, 인덱서 및 S3 대상의 증가하는 수요를 처리하여 클러스터된 구성 요소 전반에 걸쳐 부하를 분산하는 데 사용됩니다.	클러스터된 구성 요소에 대한 로드 관리

Splunk Enterprise 분산 배포의 이점은 다음과 같습니다.

- 다양하거나 분산된 데이터 소스에 접근
- 모든 규모와 복잡성을 갖춘 기업의 데이터 요구 사항을 처리할 수 있는 기능을 제공합니다.
- 데이터 복제 및 다중 사이트 배포를 통해 고가용성을 달성하고 재해 복구를 보장합니다.

스플렁크 스마트스토어

SmartStore는 Amazon S3와 같은 원격 객체 저장소가 인덱싱된 데이터를 저장할 수 있도록 하는 인덱서 기능입니다. 배포의 데이터 볼륨이 증가함에 따라 일반적으로 스토리지에 대한 수요가 컴퓨팅 리소스에 대한 수요를 앞지릅니다. SmartStore를 사용하면 리소스를 별도로 확장하여 인덱서 스토리지와 컴퓨팅 리소스를 비용 효율적으로 관리할 수 있습니다.

SmartStore는 원격 스토리지 계층과 캐시 관리자를 소개합니다. 이러한 기능을 사용하면 데이터를 인덱서의 로컬 저장소나 원격 저장소 계층에 저장할 수 있습니다. 캐시 관리자는 인덱서와 인덱서에 구성된 원격 스토리지 계층 간의 데이터 이동을 관리합니다.

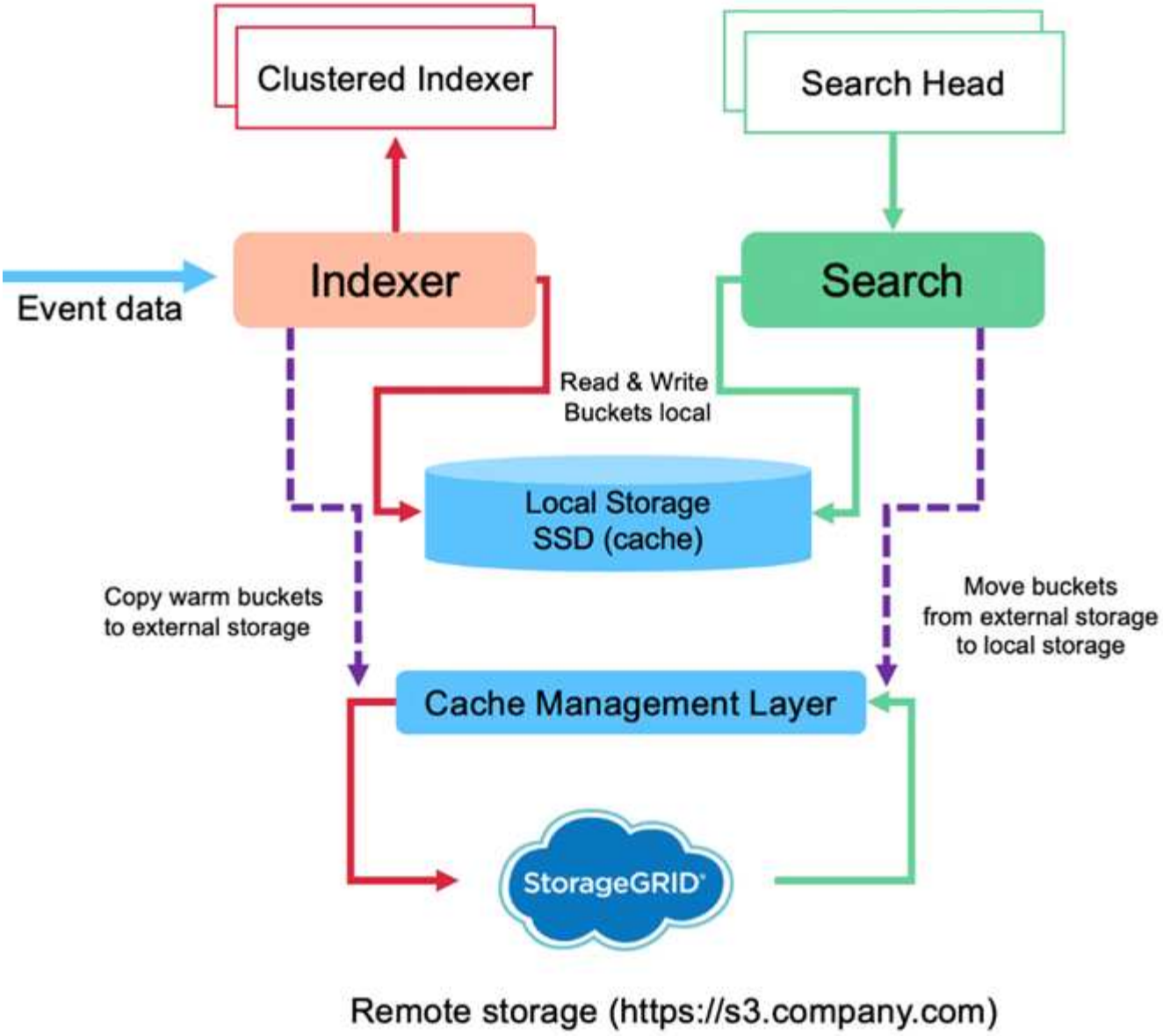
SmartStore를 사용하면 인덱서 저장소 공간을 최소한으로 줄이고 I/O 최적화된 컴퓨팅 리소스를 선택할 수 있습니다. 대부분의 데이터는 원격 저장소에 저장됩니다. 인덱서는 최소한의 데이터(핫 버킷, 활성 또는 최근 검색에 참여하는 웜 버킷의 복사본, 버킷 메타데이터)가 포함된 로컬 캐시를 유지 관리합니다.

Splunk SmartStore 데이터 흐름

다양한 소스에서 들어오는 데이터가 인덱서에 도달하면 데이터는 인덱싱되어 핫 버킷에 로컬로 저장됩니다. 인덱서는 핫 버킷 데이터를 대상 인덱서에 복제하기도 합니다. 지금까지 데이터 흐름은 SmartStore가 아닌 인덱스의 데이터 흐름과 동일합니다.

뜨거운 양동이가 따뜻해지는 경우, 데이터 흐름이 갈라집니다. 소스 인덱서는 기존 복사본을 캐시에 남겨두고 따뜻한 버킷을 원격 개체 저장소(원격 스토리지 계층)에 복사합니다. 이는 검색이 최근에 인덱싱된 데이터를 대상으로 실행되는 경향이 있기 때문입니다. 그러나 원격 저장소는 여러 개의 로컬 복사본을 유지하지 않고도 높은 가용성을 제공하므로 대상 인덱서는 자신의 복사본을 삭제합니다. 버킷의 마스터 사본은 이제 원격 저장소에 있습니다.

다음 이미지는 Splunk SmartStore 데이터 흐름을 보여줍니다.



인덱서의 캐시 관리자는 SmartStore 데이터 흐름의 핵심입니다. 검색 요청을 처리하기 위해 필요에 따라 원격 저장소에서 버킷 사본을 가져옵니다. 또한, 검색에 참여할 가능성이 시간이 지남에 따라 감소하기 때문에 오래되었거나 검색 빈도가 낮은 버킷 사본을 캐시에서 제거합니다.

캐시 관리자의 역할은 사용 가능한 캐시의 사용을 최적화하는 동시에 검색에서 필요한 버킷에 즉시 액세스할 수 있도록 하는 것입니다.

소프트웨어 요구 사항

아래 표에는 솔루션을 구현하는 데 필요한 소프트웨어 구성 요소가 나열되어 있습니다. 솔루션 구현에 사용되는 소프트웨어 구성 요소는 고객 요구 사항에 따라 달라질 수 있습니다.

제품군	제품명	제품 버전	운영 체제
NetApp StorageGRID	StorageGRID 객체 스토리지	11.6	해당 없음

제품군	제품명	제품 버전	운영 체제
센트OS	센트OS	8.1	센트OS 7.x
스플렁크 엔터프라이즈	SmartStore가 포함된 Splunk Enterprise	8.0.3	센트OS 7.x

단일 및 다중 사이트 요구 사항

데이터가 여러 컴퓨터에서 생성되고 많은 사용자가 데이터를 검색해야 하는 Enterprise Splunk 환경(중간 및 대규모 배포)에서는 Splunk Enterprise 인스턴스를 단일 및 여러 사이트에 분산하여 배포를 확장할 수 있습니다.

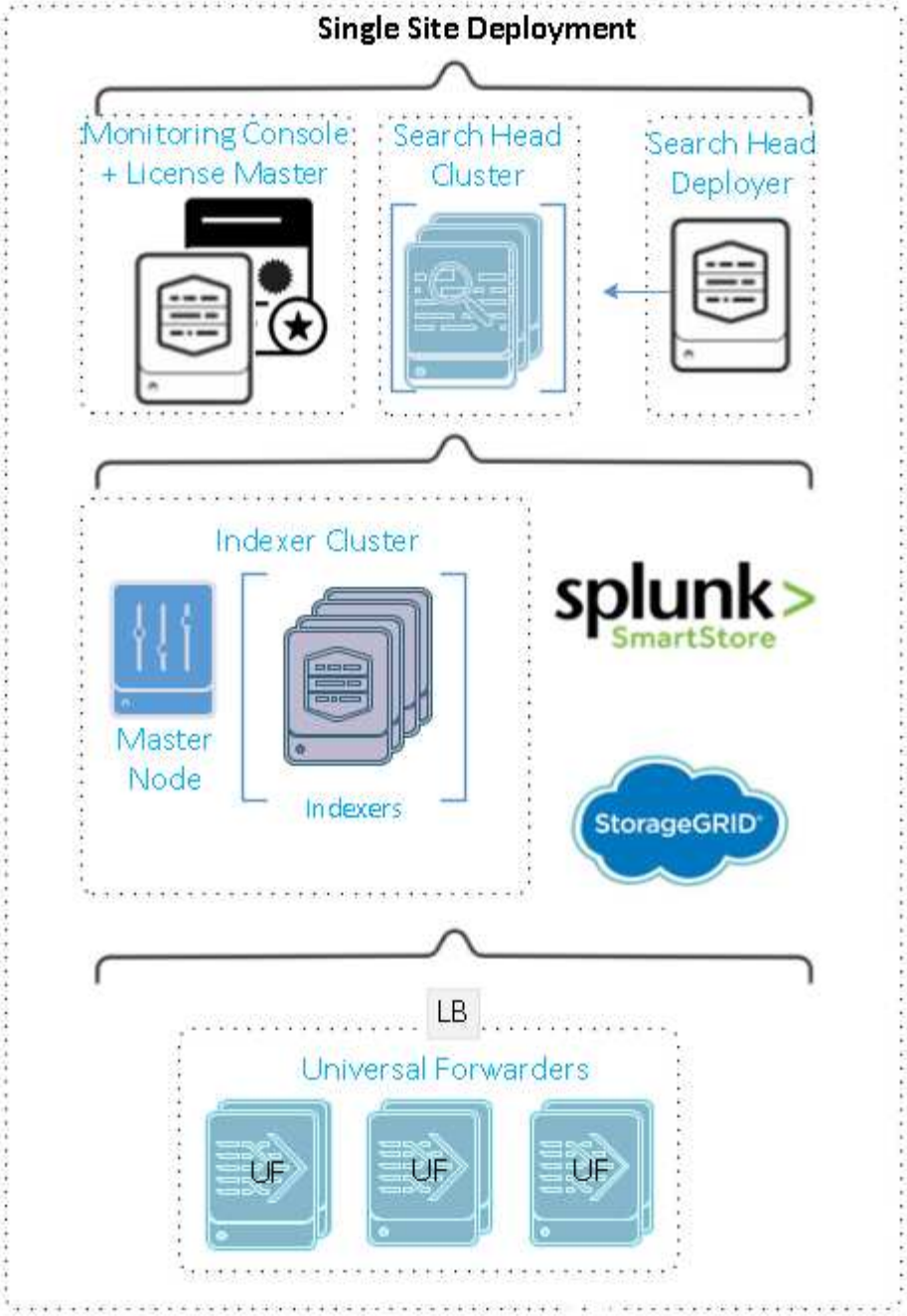
Splunk Enterprise 분산 배포의 이점은 다음과 같습니다.

- 다양하거나 분산된 데이터 소스에 접근
- 모든 규모와 복잡성을 갖춘 기업의 데이터 요구 사항을 처리할 수 있는 기능을 제공합니다.
- 데이터 복제 및 다중 사이트 배포를 통해 고가용성을 달성하고 재해 복구를 보장합니다.

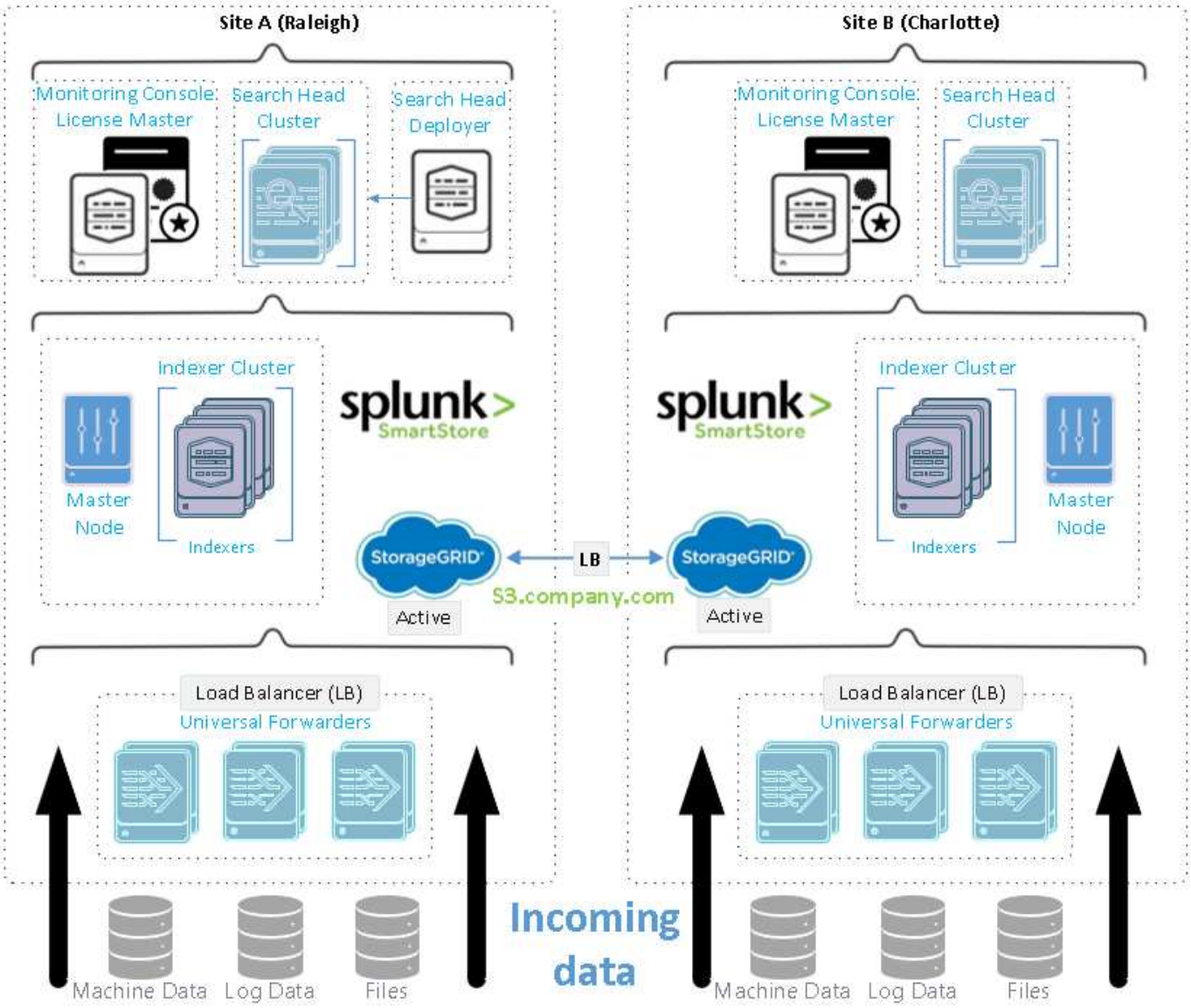
다음 표는 분산형 Splunk Enterprise 환경에서 사용되는 구성 요소를 나열합니다.

요소	설명	책임
인덱스 클러스터 마스터	인덱서 클러스터의 활동 및 업데이트를 조정합니다.	인덱스 관리
인덱스 클러스터	서로의 데이터를 복제하도록 구성된 Splunk Enterprise 인덱서 그룹	인덱싱
검색 헤드 배포자	클러스터 마스터에 대한 배포 및 업데이트를 처리합니다.	검색 헤드 관리
검색 헤드 클러스터	검색을 위한 중앙 리소스 역할을 하는 검색 헤드 그룹	검색 관리
로드 밸런서	클러스터된 구성 요소에서 검색 헤드, 인덱서 및 S3 대상의 증가하는 수요를 처리하여 클러스터된 구성 요소 전반에 걸쳐 부하를 분산하는 데 사용됩니다.	클러스터된 구성 요소에 대한 로드 관리

이 그림은 단일 사이트 분산 배포의 예를 보여줍니다.



이 그림은 다중 사이트 분산 배포의 예를 보여줍니다.



하드웨어 요구 사항

다음 표에는 솔루션을 구현하는 데 필요한 최소 하드웨어 구성 요소 수가 나열되어 있습니다. 솔루션의 특정 구현에 사용되는 하드웨어 구성 요소는 고객 요구 사항에 따라 달라질 수 있습니다.

> Splunk SmartStore와 StorageGRID 단일 사이트에 배포했든 여러 사이트에 배포했든 모든 시스템은 단일 창에서 StorageGRID GRID Manager를 통해 관리됩니다. 자세한 내용은 "Grid Manager를 사용한 간편한 관리" 섹션을 참조하세요.

이 표는 단일 사이트에 사용된 하드웨어를 나열합니다.

하드웨어	수량	디스크	사용 가능 용량	참고
StorageGRID SG1000	1	해당 없음	해당 없음	관리 노드 및 로드 밸런서
StorageGRID SG6060	4	x48, 8TB(NL-SAS HDD)	1PB	원격 저장소

이 표는 다중 사이트 구성에 사용되는 하드웨어를 나열합니다(사이트별).

하드웨어	수량	디스크	사용 가능 용량	참고
StorageGRID SG1000	2	해당 없음	해당 없음	관리 노드 및 로드 밸런서
StorageGRID SG6060	4	x48, 8TB(NL-SAS HDD)	1PB	원격 저장소

NetApp StorageGRID 로드 밸런서**: SG1000**

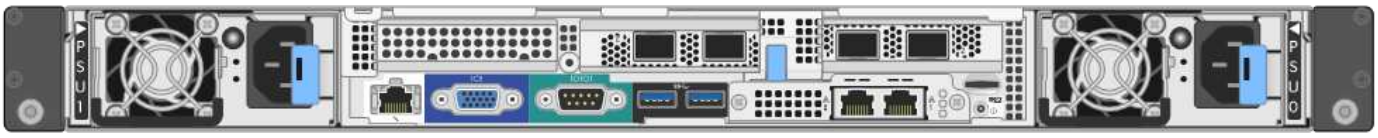
개체 스토리지에는 클라우드 스토리지 네임스페이스를 제공하기 위해 로드 밸런서를 사용해야 합니다. StorageGRID F5 및 Citrix와 같은 주요 공급업체의 타사 로드 밸런서를 지원하지만 많은 고객은 단순성, 복원력 및 고성능을 위해 엔터프라이즈급 StorageGRID 밸런서를 선택합니다. StorageGRID 로드 밸런서는 VM, 컨테이너 또는 특수 목적 어플라이언스로 사용할 수 있습니다.

StorageGRID SG1000은 S3 데이터 경로 연결을 위한 고가용성(HA) 그룹과 지능형 부하 분산을 쉽게 사용할 수 있도록 해줍니다. 다른 온프레미스 개체 스토리지 시스템은 사용자 정의형 로드 밸런서를 제공하지 않습니다.

SG1000 어플라이언스는 다음과 같은 기능을 제공합니다.

- StorageGRID 시스템을 위한 로드 밸런서 및 선택적으로 관리 노드 기능
- 노드 배포 및 구성을 단순화하는 StorageGRID Appliance 설치 프로그램
- S3 엔드포인트 및 SSL의 간소화된 구성
- 전용 대역폭(다른 애플리케이션과 타사 로드 밸런서를 공유하는 것과 대비)
- 최대 4 x 100Gbps의 총 이더넷 대역폭

다음 이미지는 SG1000 Gateway Services 어플라이언스를 보여줍니다.



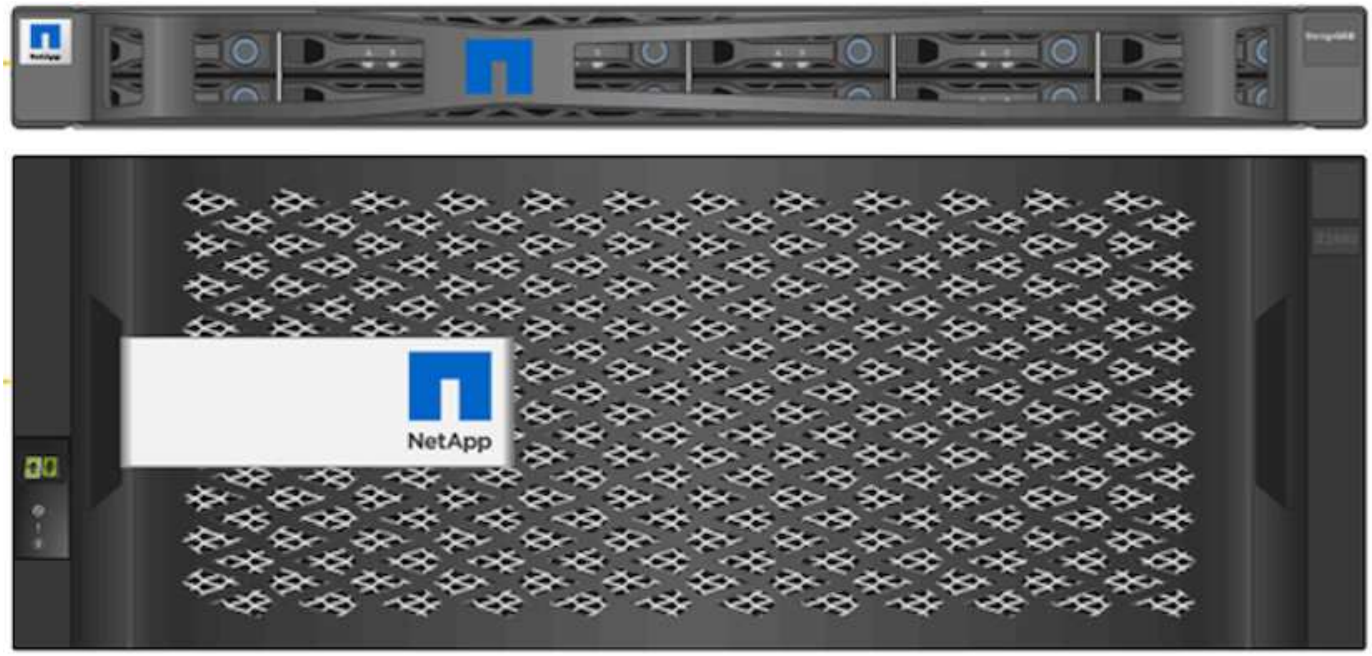
SG6060

StorageGRID SG6060 어플라이언스에는 컴퓨팅 컨트롤러(SG6060)와 스토리지 컨트롤러 셸프(E-Series E2860)가 포함되어 있으며, 스토리지 컨트롤러 셸프에는 스토리지 컨트롤러 2개와 드라이브 60개가 들어 있습니다. 이 기기는 다음과 같은 기능을 제공합니다.

- 단일 네임스페이스에서 최대 400PB까지 확장 가능합니다.
- 최대 4x 25Gbps의 총 이더넷 대역폭.
- 노드 배포 및 구성을 단순화하기 위해 StorageGRID Appliance Installer가 포함되어 있습니다.

- 각 SG6060 어플라이언스는 1~2개의 추가 확장 선반을 장착하여 총 180개의 드라이브를 장착할 수 있습니다.
- 스토리지 컨트롤러 장애 조치 지원을 제공하기 위한 2개의 E-시리즈 E2800 컨트롤러(듀플렉스 구성)
- 60개의 3.5인치 드라이브(2개의 솔리드 스테이트 드라이브와 58개의 NL-SAS 드라이브)를 보관할 수 있는 5개 서랍 드라이브 선반입니다.

다음 이미지는 SG6060 기기를 보여줍니다.



스플렁크 디자인

다음 표에는 단일 사이트에 대한 Splunk 구성이 나열되어 있습니다.

Splunk 구성 요소	일	수량	코어	메모리	운영 체제
유니버설 포워더	데이터 수집 및 인덱서에 데이터 전달을 담당합니다.	4	16개의 코어	32GB 램	센트OS 8.1
인덱서	사용자 데이터를 관리합니다	10	16개의 코어	32GB 램	센트OS 8.1
검색 헤드	사용자 프런트 엔드는 인덱서에서 데이터를 검색합니다.	3	16개의 코어	32GB 램	센트OS 8.1
검색 헤드 배포자	검색 헤드 클러스터에 대한 업데이트를 처리합니다.	1	16개의 코어	32GB 램	센트OS 8.1
클러스터 마스터	Splunk 설치 및 인덱서를 관리합니다.	1	16개의 코어	32GB 램	센트OS 8.1
모니터링 콘솔 및 라이센스 마스터	Splunk 배포 전체에 대한 중앙 모니터링을 수행하고 Splunk 라이선스를 관리합니다.	1	16개의 코어	32GB 램	센트OS 8.1

다음 표에서는 다중 사이트 구성을 위한 Splunk 구성을 설명합니다.

이 표는 다중 사이트 구성(사이트 A)에 대한 Splunk 구성을 나열합니다.

Splunk 구성 요소	일	수량	코어	메모리	운영 체제
유니버설 포워더	데이터를 수집하고 인덱서에게 데이터를 전달하는 역할을 담당합니다.	4	16개의 코어	32GB 램	센트OS 8.1
인덱서	사용자 데이터를 관리합니다	10	16개의 코어	32GB 램	센트OS 8.1
검색 헤드	사용자 프런트 엔드는 인덱서에서 데이터를 검색합니다.	3	16개의 코어	32GB 램	센트OS 8.1
검색 헤드 배포자	검색 헤드 클러스터에 대한 업데이트를 처리합니다.	1	16개의 코어	32GB 램	센트OS 8.1

Splunk 구성 요소	일	수량	코어	메모리	운영 체제
클러스터 마스터	Splunk 설치 및 인덱서를 관리합니다.	1	16개의 코어	32GB 램	센트OS 8.1
모니터링 콘솔 및 라이센스 마스터	Splunk 배포 전체에 대한 중앙 모니터링을 수행하고 Splunk 라이선스를 관리합니다.	1	16개의 코어	32GB 램	센트OS 8.1

이 표는 다중 사이트 구성(사이트 B)에 대한 Splunk 구성을 나열합니다.

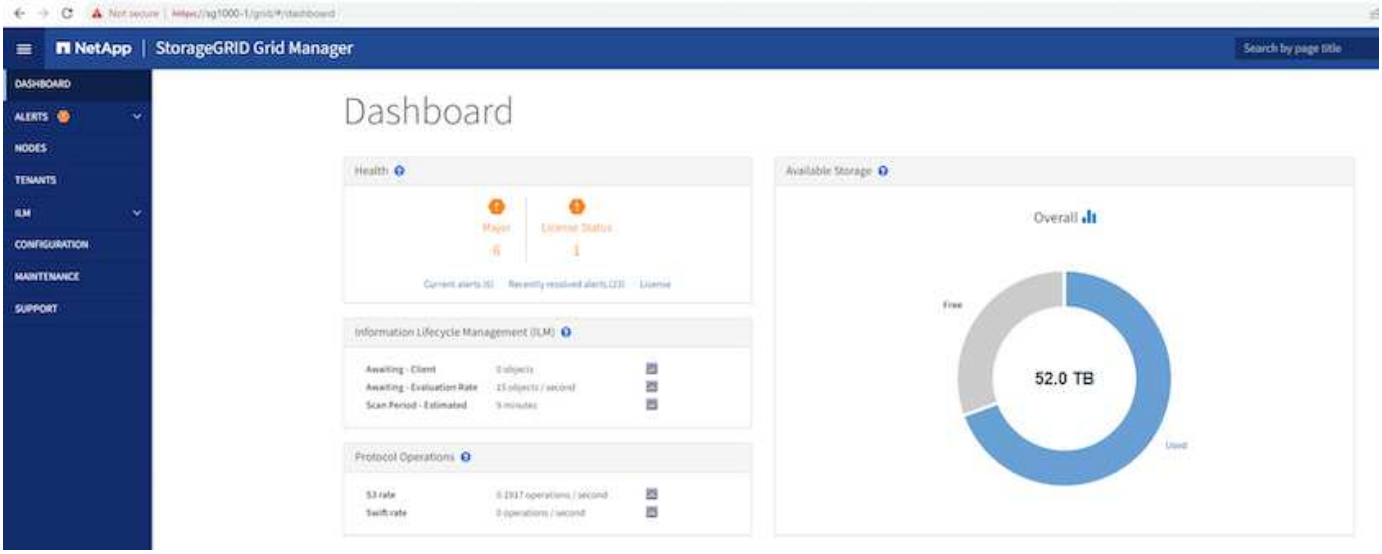
Splunk 구성 요소	일	수량	코어	메모리	운영 체제
유니버설 포워더	데이터 수집 및 인덱서에 데이터 전달을 담당합니다.	4	16개의 코어	32GB 램	센트OS 8.1
인덱서	사용자 데이터를 관리합니다	10	16개의 코어	32GB 램	센트OS 8.1
검색 헤드	사용자 프런트 엔드는 인덱서에서 데이터를 검색합니다.	3	16개의 코어	32GB 램	센트OS 8.1
클러스터 마스터	Splunk 설치 및 인덱서를 관리합니다.	1	16개의 코어	32GB 램	센트OS 8.1
모니터링 콘솔 및 라이센스 마스터	Splunk 배포 전체에 대한 중앙 모니터링을 수행하고 Splunk 라이선스를 관리합니다.	1	16개의 코어	32GB 램	센트OS 8.1

Splunk SmartStore를 위한 유연한 StorageGRID 기능

StorageGRID 사용자가 끊임없이 변화하는 환경에 맞게 활용하고 사용자 정의할 수 있는 다양한 기능을 제공합니다. Splunk SmartStore를 배포하거나 확장하는 과정에서 환경에는 변화에 대한 신속한 적응이 필요하며 Splunk를 중단시키지 않아야 합니다. StorageGRID 유연한 데이터 관리 정책(ILM)과 트래픽 분류기(QoS)를 사용하면 환경에 맞게 계획을 세우고 적응할 수 있습니다.

Grid Manager를 통한 간편한 관리

Grid Manager는 다음 이미지에서 볼 수 있듯이 단일 창에서 전 세계에 분산된 위치에 있는 StorageGRID 시스템을 구성, 관리 및 모니터링할 수 있는 브라우저 기반 그래픽 인터페이스입니다.



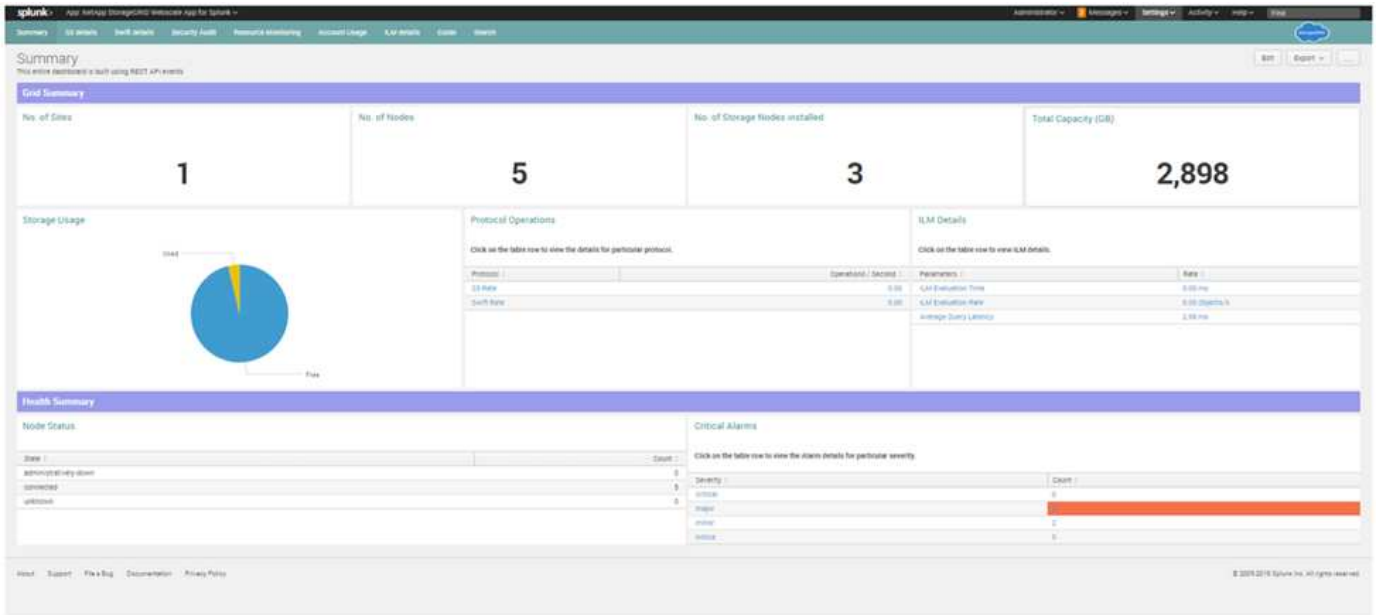
Grid Manager 인터페이스를 사용하여 다음 작업을 수행합니다.

- 이미지, 비디오, 기록 등의 객체가 담긴 페타바이트 규모의 글로벌 분산 저장소를 관리합니다.
- 그리드 노드와 서비스를 모니터링하여 객체 가용성을 보장합니다.
- 정보 수명 주기 관리(ILM) 규칙을 사용하여 시간 경과에 따른 개체 데이터의 배치를 관리합니다. 이러한 규칙은 객체 데이터가 수집된 후 어떻게 처리되는지, 데이터가 손실로부터 어떻게 보호되는지, 객체 데이터가 어디에 저장되는지, 얼마 동안 저장되는지를 관리합니다.
- 시스템 내에서 거래, 성능 및 운영을 모니터링합니다.

Splunk용 NetApp StorageGRID 앱

Splunk용 NetApp StorageGRID 앱은 Splunk Enterprise에 특화된 애플리케이션입니다. 이 앱은 Splunk용 NetApp StorageGRID 애드온과 함께 작동합니다. StorageGRID 상태, 계정 사용 정보, 보안 감사 세부 정보, 리소스 사용 및 모니터링 등에 대한 가시성을 제공합니다.

다음 이미지는 Splunk용 StorageGRID 앱을 보여줍니다.



ILM 정책

StorageGRID 객체의 여러 사본을 보관하고 2+1, 4+2(및 기타 여러 방식)와 같은 EC(삭제 코딩) 방식을 사용하여 특정 성능 및 데이터 보호 요구 사항에 따라 객체를 저장하는 등 유연한 데이터 관리 정책을 제공합니다. 시간이 지남에 따라 작업 부하와 요구 사항이 바뀌므로 ILM 정책도 시간이 지남에 따라 변경해야 하는 것이 일반적입니다. ILM 정책을 수정하는 것은 StorageGRID 고객이 끊임없이 변화하는 환경에 빠르고 쉽게 적응할 수 있도록 하는 핵심 기능입니다.

성능

StorageGRID SG5712, SG5760, SG6060 또는 SGF6024와 같은 VM이나 베어메탈 또는 특수 목적 어플라이언스를 더 추가하여 성능을 확장합니다. 테스트 결과, SG6060 어플라이언스를 사용하여 최소 크기의 3노드 그리드로 SmartStore의 주요 성능 요구 사항을 초과했습니다. 고객이 추가 인덱서를 사용하여 Splunk 인프라를 확장하면 더 많은 스토리지 노드를 추가하여 성능과 용량을 늘릴 수 있습니다.

로드 밸런서 및 엔드포인트 구성

StorageGRID 의 관리 노드는 StorageGRID 시스템을 보고, 구성하고, 관리할 수 있는 Grid Manager UI(사용자 인터페이스)와 REST API 엔드포인트를 제공하며, 감사 로그를 통해 시스템 활동을 추적합니다. Splunk SmartStore 원격 스토리지에 고가용성 S3 엔드포인트를 제공하기 위해 관리 노드와 게이트웨이 노드에서 서비스로 실행되는 StorageGRID 로드 밸런서를 구현했습니다. 또한, 로드 밸런서는 로컬 트래픽을 관리하고 GSLB(글로벌 서버 로드 밸런싱)와 통신하여 재해 복구를 지원합니다.

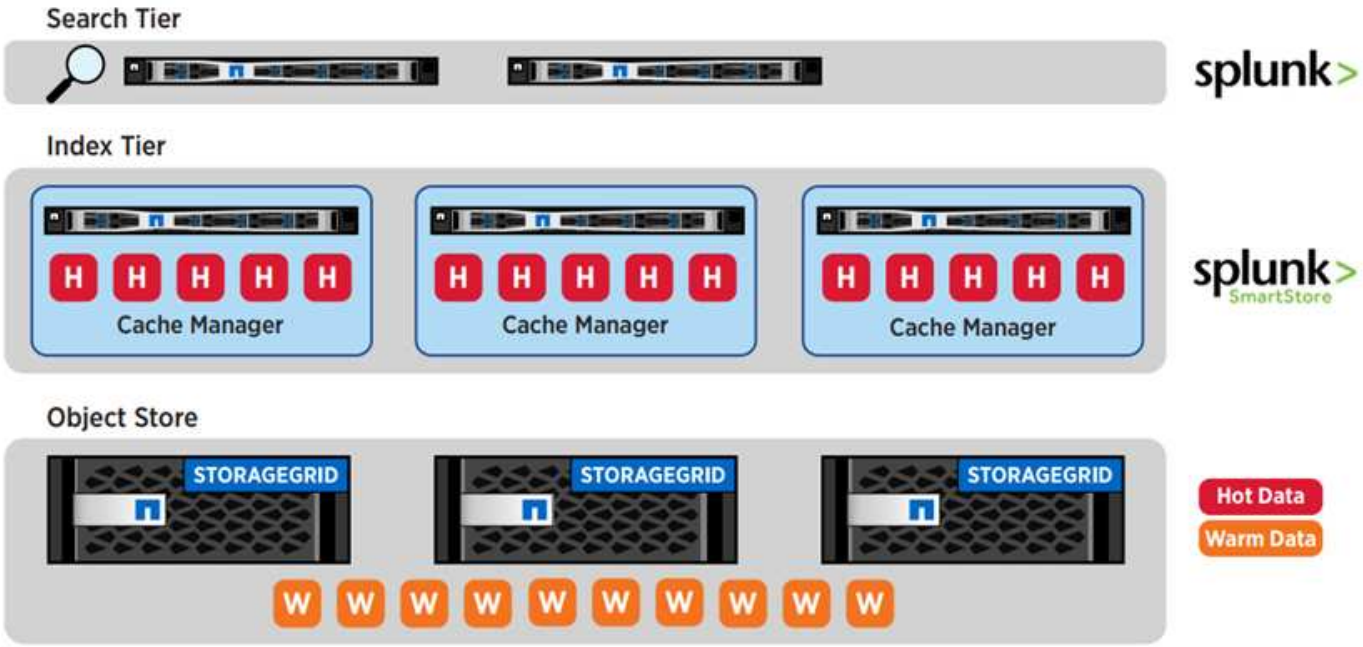
엔드포인트 구성을 더욱 향상시키기 위해 StorageGRID 관리 노드에 내장된 트래픽 분류 정책을 제공하고, 이를 통해 워크로드 트래픽을 모니터링하고, 워크로드에 다양한 서비스 품질(QoS) 제한을 적용할 수 있습니다. 트래픽 분류 정책은 게이트웨이 노드와 관리 노드에 대한 StorageGRID 부하 분산 서비스의 엔드포인트에 적용됩니다. 이러한 정책은 교통 제한 및 모니터링에 도움이 될 수 있습니다.

지능형 계층화 및 비용 절감

고객들은 Splunk 데이터 분석의 강력함과 편리함을 깨닫고, 자연스럽게 점점 더 많은 양의 데이터를 인덱싱하고자 합니다. 데이터 양이 증가함에 따라 이를 처리하는 데 필요한 컴퓨팅 및 스토리지 인프라도 함께 증가합니다. 오래된 데이터는 참조 빈도가 낮아지기 때문에 동일한 양의 컴퓨팅 리소스를 투입하고 값비싼 기본 저장소를 사용하는 것은 점점 더 비효율적이 됩니다. 대규모로 운영하기 위해 고객은 따뜻한 데이터를 더 비용 효율적인 계층으로 이동하여 컴퓨팅과 기본 스토리지를 뜨거운 데이터에 사용할 수 있는 이점을 얻습니다.

StorageGRID 탑재된 Splunk SmartStore는 조직에 확장 가능하고 성능이 뛰어나며 비용 효율적인 솔루션을 제공합니다. SmartStore는 데이터를 인식하므로 데이터 액세스 패턴을 자동으로 평가하여 실시간 분석을 위해 액세스해야 하는 데이터(핫 데이터)와 비용이 적게 드는 장기 저장소에 저장해야 하는 데이터(웜 데이터)를 판별합니다. SmartStore는 업계 표준인 AWS S3 API를 동적이고 지능적으로 사용하여 StorageGRID 제공하는 S3 스토리지에 데이터를 저장합니다. StorageGRID 의 유연한 확장형 아키텍처를 통해 필요에 따라 웜 데이터 계층을 비용 효율적으로 확장할 수 있습니다. StorageGRID 의 노드 기반 아키텍처는 성능 및 비용 요구 사항이 최적으로 충족되도록 보장합니다.

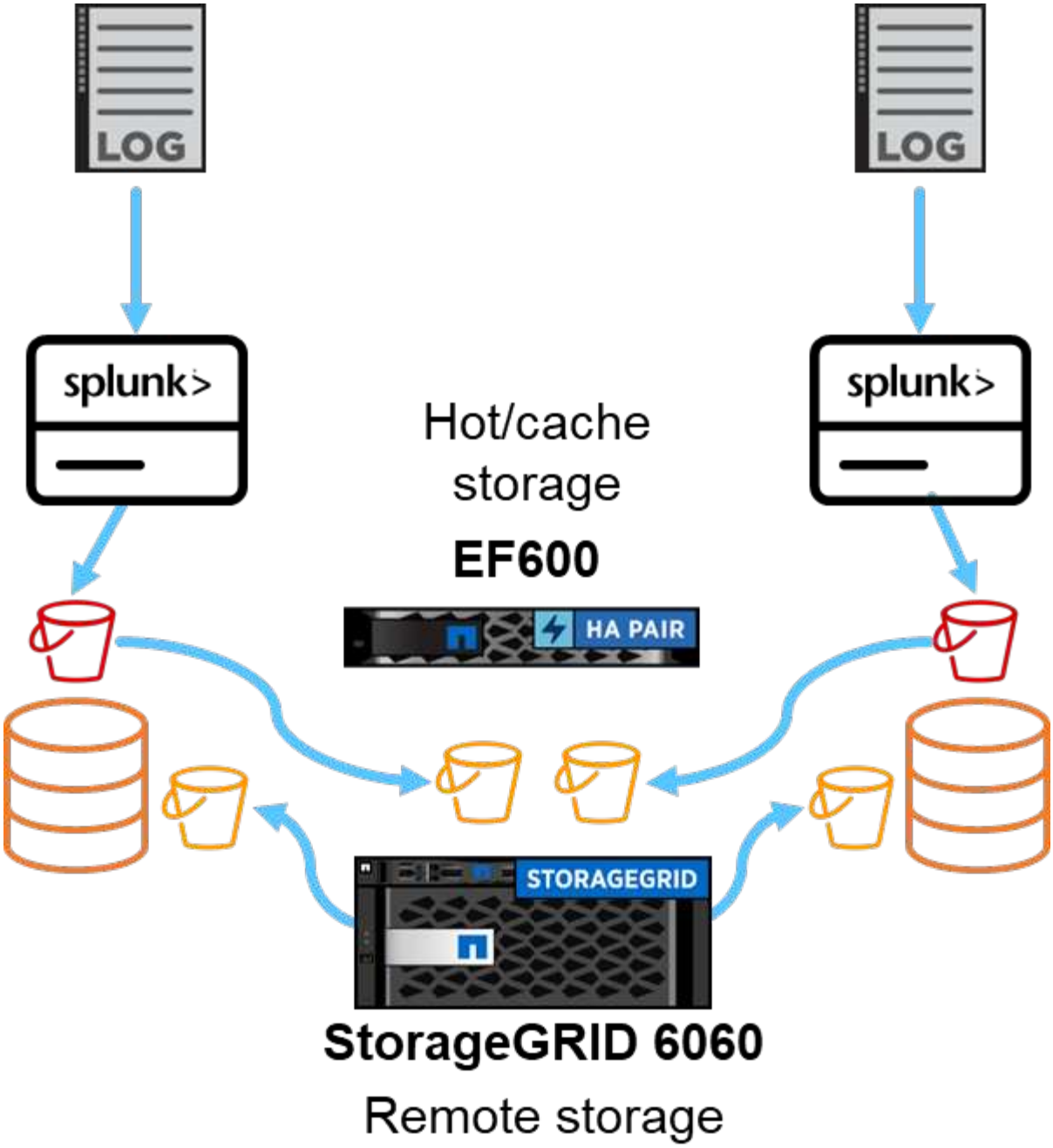
다음 이미지는 Splunk와 StorageGRID 계층화를 보여줍니다.



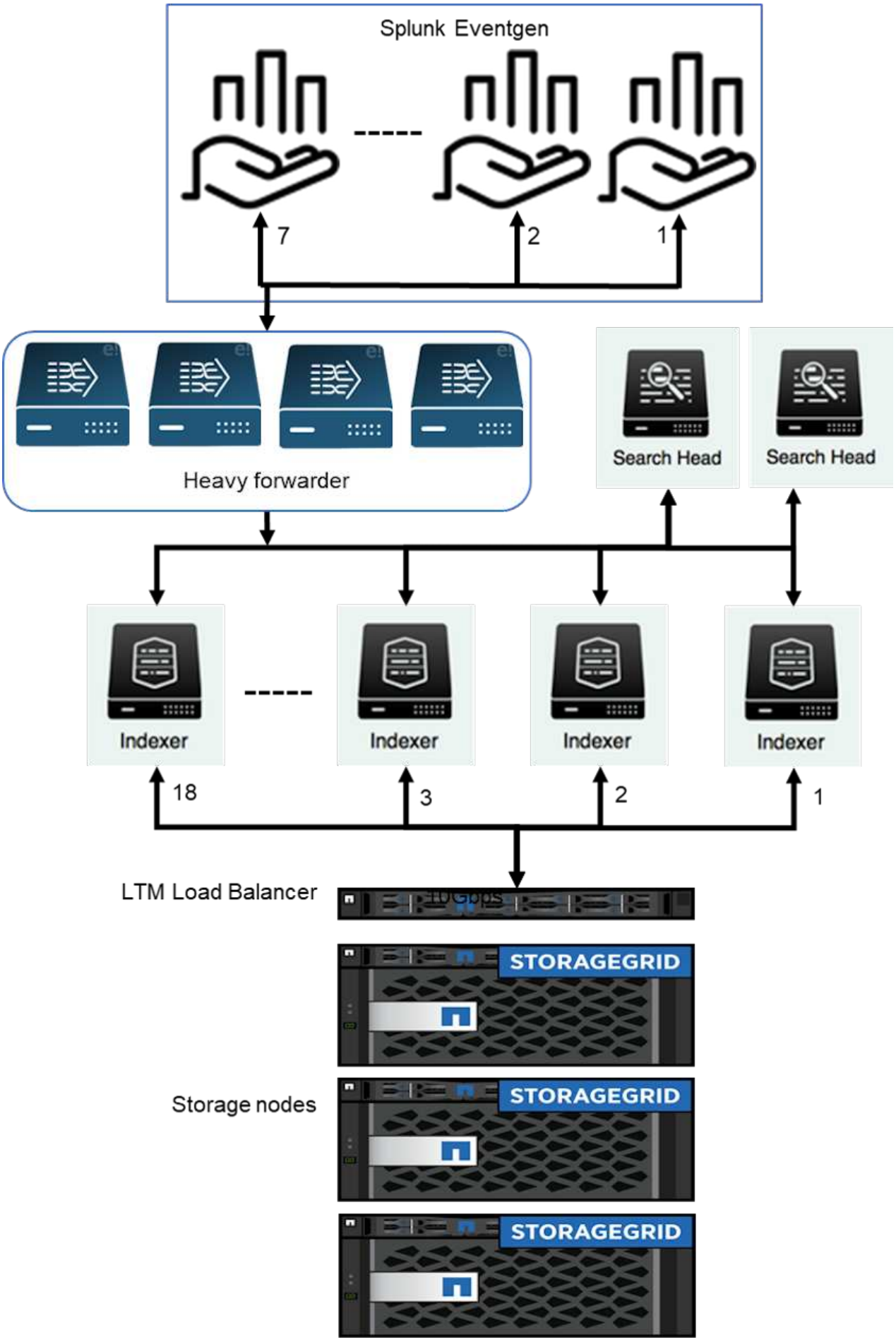
Splunk SmartStore와 NetApp StorageGRID 의 업계 최고 조합은 풀스택 솔루션을 통해 분리된 아키텍처의 이점을 제공합니다.

단일 사이트 **SmartStore** 성능

이 섹션에서는 NetApp StorageGRID 컨트롤러에서의 Splunk SmartStore 성능에 대해 설명합니다. Splunk SmartStore는 따뜻한 데이터를 원격 스토리지로 이동합니다. 이 경우 성능 검증에서 이는 StorageGRID 개체 스토리지입니다.



핫/캐시 스토리지에는 EF600을, 원격 스토리지에는 StorageGRID 6060을 사용했습니다. 성능 검증을 위해 다음과 같은 아키텍처를 사용했습니다. 우리는 2개의 검색 헤드와 4개의 대형 포워더를 사용하여 데이터를 인덱서로 전달하고, 7개의 Splunk 이벤트 생성기(Eventgens)를 사용하여 실시간 데이터를 생성하고, 18개의 인덱서를 사용하여 데이터를 저장했습니다.


Splunk Eventgen
7
2
1
Heavy forwarder
Search Head
Search Head
Indexer
Indexer
Indexer
Indexer
18
3
2
1
LTM Load Balancer
10Gbps
STORAGEGRID
Storage nodes
STORAGEGRID
STORAGEGRID

구성

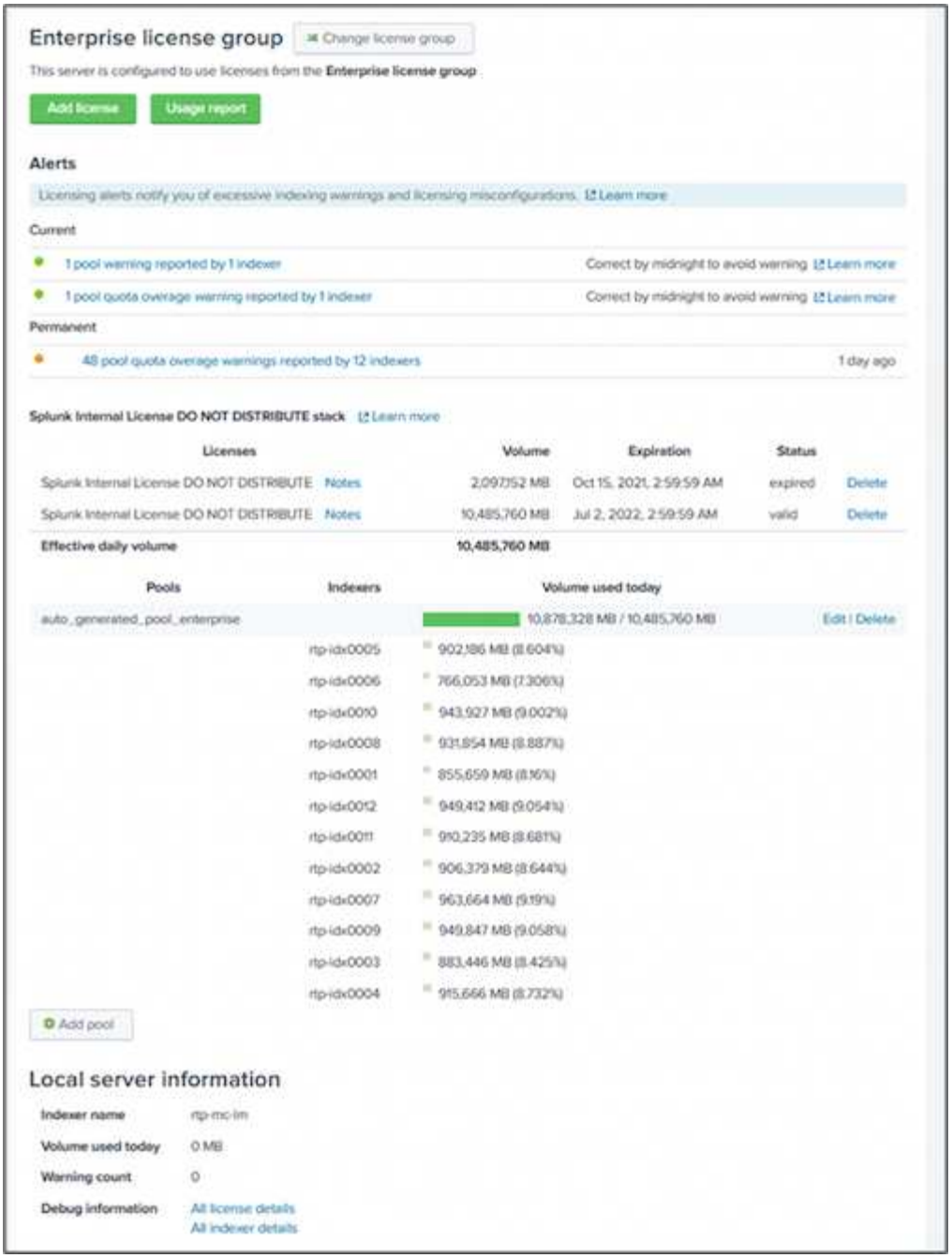
이 표는 SmartStorage 성능 검증에 사용된 하드웨어를 나열합니다.

Splunk 구성 요소	일	수량	코어	메모리	운영 체제
헤비 포워더	데이터 수집 및 인덱서에 데이터 전달을 담당합니다.	4	16개의 코어	32GB 램	슬레드 15 SP2
인덱서	사용자 데이터를 관리합니다	18	16개의 코어	32GB 램	슬레드 15 SP2
검색 헤드	사용자 프런트엔드는 인덱서에서 데이터를 검색합니다.	2	16개의 코어	32GB 램	슬레드 15 SP2
검색 헤드 배포자	검색 헤드 클러스터에 대한 업데이트를 처리합니다.	1	16개의 코어	32GB 램	슬레드 15 SP2
클러스터 마스터	Splunk 설치 및 인덱서를 관리합니다.	1	16개의 코어	32GB 램	슬레드 15 SP2
모니터링 콘솔 및 라이센스 마스터	Splunk 배포 전체에 대한 중앙 모니터링을 수행하고 Splunk 라이선스를 관리합니다.	1	16개의 코어	32GB 램	슬레드 15 SP2

SmartStore 원격 매장 성능 검증

이 성능 검증에서는 모든 인덱서의 로컬 스토리지에 10일 분의 데이터를 저장하는 SmartStore 캐시를 구성했습니다. 우리는 가능하게 했습니다 `maxDataSize=auto` (버킷 크기 750MB) Splunk 클러스터 관리자에서 모든 인덱서에 변경 사항을 푸시했습니다. 업로드 성능을 측정하기 위해 10일 동안 매일 10TB를 수집하고 모든 핫 버킷을 동시에 워밍업 버킷으로 전환한 다음 SmartStore 모니터링 콘솔 대시보드에서 인스턴스별 및 배포 전반의 최대 및 평균 처리량을 파악했습니다.

이 이미지는 하루에 수집된 데이터를 보여줍니다.



클러스터 마스터에서 다음 명령을 실행했습니다(인덱스 이름은 다음과 같습니다. `eventgen-test`). 그런 다음 SmartStore 모니터링 콘솔 대시보드를 통해 인스턴스별 및 배포 전반의 최대 및 평균 업로드 처리량을 파악했습니다.

```
for i in rtp-idx0001 rtp-idx0002 rtp-idx0003 rtp-idx0004 rtp-idx0005 rtp-
idx0006 rtp-idx0007 rtp-idx0008 rtp-idx0009 rtp-idx0010 rtp-idx0011 rtp-
idx0012 rtp-idx0013011 rtdx0014 rtp-idx0015 rtp-idx0016 rtp-idx0017 rtp-
idx0018 ; do  ssh $i "hostname;  date; /opt/splunk/bin/splunk _internal
call /data/indexes/eventgen-test/roll-hot-buckets -auth admin:12345678;
sleep 1  "; done
```

클러스터 마스터는 모든 인덱서(rtp-idx0001…rtp-idx0018)에 대해 암호 없는 인증을 갖습니다.

다운로드 성능을 측정하기 위해 다음 명령을 사용하여 evict CLI를 두 번 실행하여 캐시에서 모든 데이터를 제거했습니다.

> 클러스터 마스터에서 다음 명령을 실행하고 StorageGRID 의 원격 저장소에 있는 10일 분의 데이터를 기반으로 검색 헤드에서 검색을 실행했습니다. 그런 다음 SmartStore 모니터링 콘솔 대시보드를 통해 인스턴스별 및 배포 전반의 최대 및 평균 업로드 처리량을 파악했습니다.

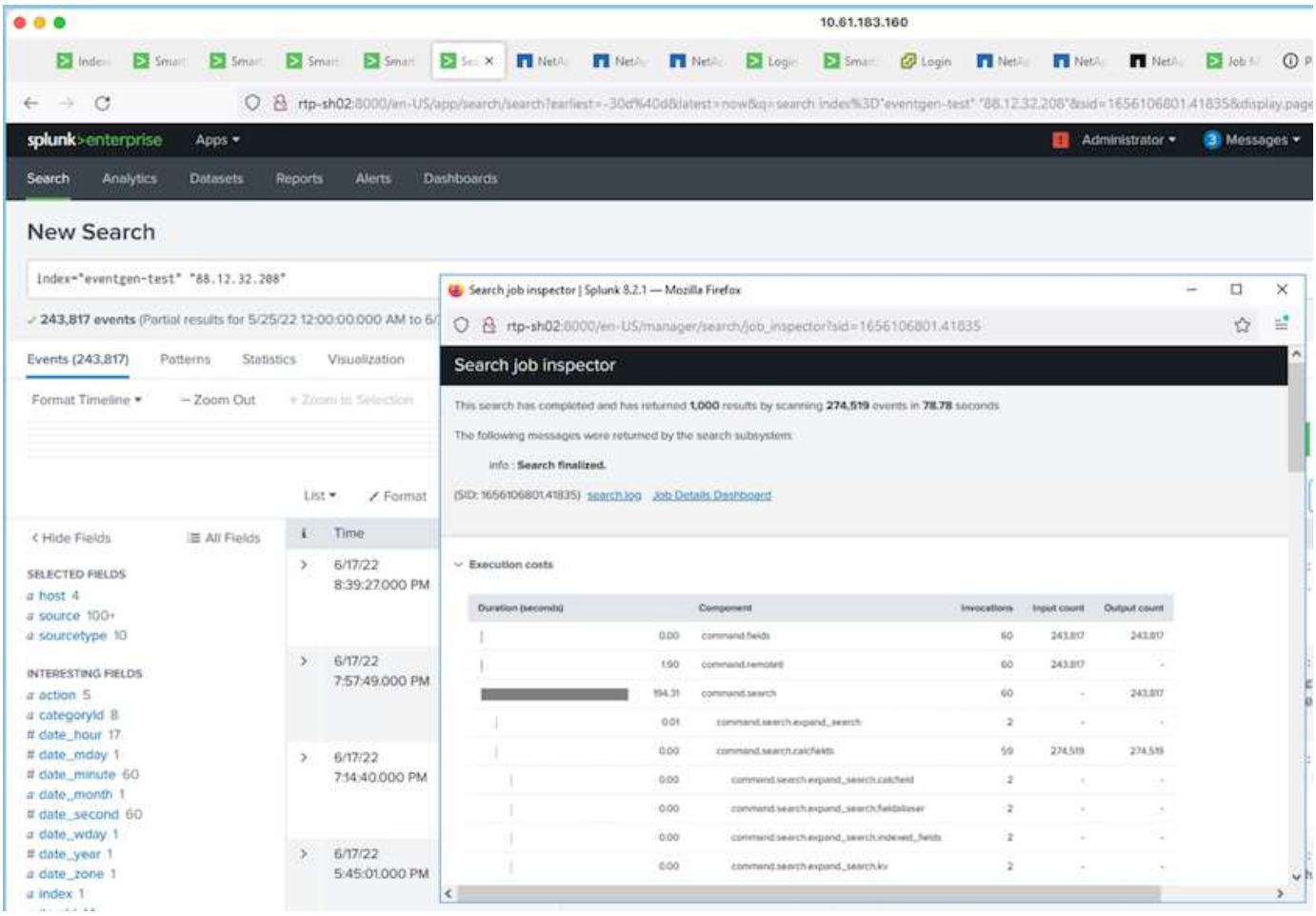
```
for i in rtp-idx0001 rtp-idx0002 rtp-idx0003 rtp-idx0004 rtp-idx0005 rtp-
idx0006 rtp-idx0007 rtp-idx0008 rtp-idx0009 rtp-idx0010 rtp-idx0011 rtp-
idx0012 rtp-idx0013 rtp-idx0014 rtp-idx0015 rtp-idx0016 rtp-idx0017 rtp-
idx0018 ; do  ssh $i " hostname;  date; /opt/splunk/bin/splunk _internal
call /services/admin/cacheman/_evict -post:mb 1000000000 -post:path
/mnt/EF600 -method POST  -auth admin:12345678;   "; done
```

인덱서 구성은 SmartStore 클러스터 마스터에서 푸시되었습니다. 클러스터 마스터는 인덱서에 대해 다음과 같은 구성을 가졌습니다.

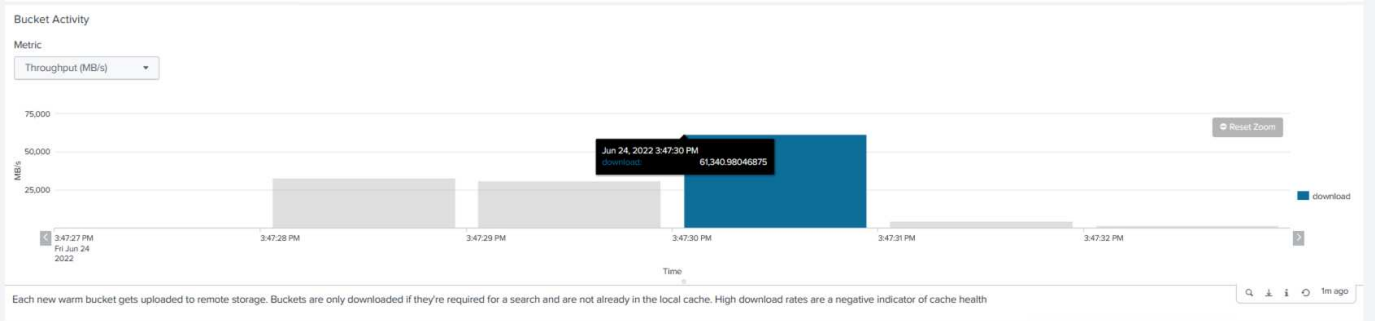
```
Rtp-cm01:~ # cat /opt/splunk/etc/master-apps/_cluster/local/indexes.conf
[default]
maxDataSize = auto
#defaultDatabase = eventgen-basic
defaultDatabase = eventgen-test
hotlist_recency_secs = 864000
repFactor = auto
[volume:remote_store]
storageType = remote
path = s3://smartstore2
remote.s3.access_key = U64TUHONBNC98GQGL60R
remote.s3.secret_key = UBoXNE0jmECie05Z7iCYVzbSB6WJFckiYLcdm2yg
remote.s3.endpoint = 3.sddc.netapp.com:10443
remote.s3.signature_version = v2
remote.s3.clientCert =
[eventgen-basic]
homePath = $SPLUNK_DB/eventgen-basic/db
coldPath = $SPLUNK_DB/eventgen-basic/colddb
thawedPath = $SPLUNK_DB/eventgen-basic/thawed
[eventgen-migration]
homePath = $SPLUNK_DB/eventgen-scale/db
coldPath = $SPLUNK_DB/eventgen-scale/colddb
thawedPath = $SPLUNK_DB/eventgen-scale/thaweddb
[main]
homePath = $SPLUNK_DB/$_index_name/db
coldPath = $SPLUNK_DB/$_index_name/colddb
thawedPath = $SPLUNK_DB/$_index_name/thaweddb
[history]
```

```
homePath = $SPLUNK_DB/$_index_name/db
coldPath = $SPLUNK_DB/$_index_name/colddb
thawedPath = $SPLUNK_DB/$_index_name/thaweddb
[summary]
homePath = $SPLUNK_DB/$_index_name/db
coldPath = $SPLUNK_DB/$_index_name/colddb
thawedPath = $SPLUNK_DB/$_index_name/thaweddb
[remote-test]
homePath = $SPLUNK_DB/$_index_name/db
coldPath = $SPLUNK_DB/$_index_name/colddb
#for storagegrid config
remotePath = volume:remote_store/$_index_name
thawedPath = $SPLUNK_DB/$_index_name/thaweddb
[eventgen-test]
homePath = $SPLUNK_DB/$_index_name/db
maxDataSize=auto
maxHotBuckets=1
maxWarmDBCount=2
coldPath = $SPLUNK_DB/$_index_name/colddb
#for storagegrid config
remotePath = volume:remote_store/$_index_name
thawedPath = $SPLUNK_DB/$_index_name/thaweddb
[eventgen-evict-test]
homePath = $SPLUNK_DB/$_index_name/db
coldPath = $SPLUNK_DB/$_index_name/colddb
#for storagegrid config
remotePath = volume:remote_store/$_index_name
thawedPath = $SPLUNK_DB/$_index_name/thaweddb
maxDataSize = auto_high_volume
maxWarmDBCount = 5000
rtp-cm01:~ #
```

성과 매트릭스를 수집하기 위해 검색 헤드에서 다음 검색 쿼리를 실행했습니다.



우리는 클러스터 마스터로부터 성능 정보를 수집했습니다. 최대 성능은 61.34GBps였습니다.



평균 성능은 약 29GBps였습니다.

StorageGRID 성능

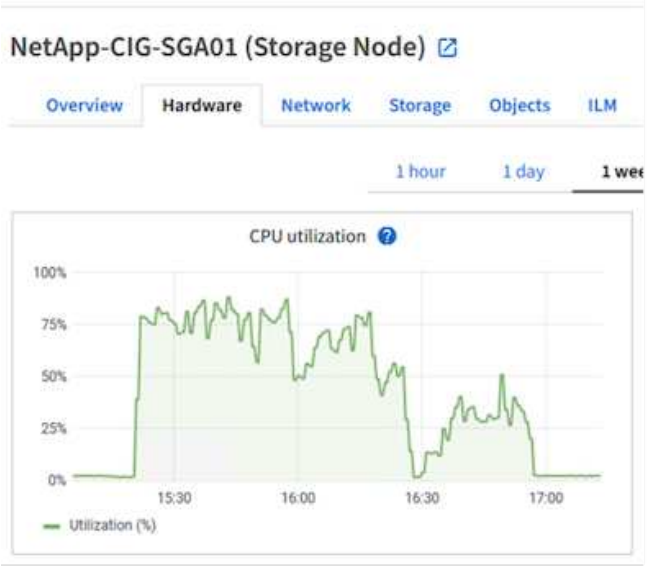
SmartStore의 성능은 대량의 데이터에서 특정 패턴과 문자열을 검색하는 데 기반합니다. 이 검증에서는 이벤트가 다음을 사용하여 생성됩니다. "이벤트젠" 검색 헤드를 통해 특정 Splunk 인덱스(eventgen-test)에 대한 검색이 수행되고, 대부분의 쿼리에 대한 요청은 StorageGRID 로 이동합니다. 다음 이미지는 쿼리 데이터의 적중과 미적중을 보여줍니다. 히트 데이터는 로컬 디스크에서 가져오고, 미스 데이터는 StorageGRID 컨트롤러에서 가져옵니다.

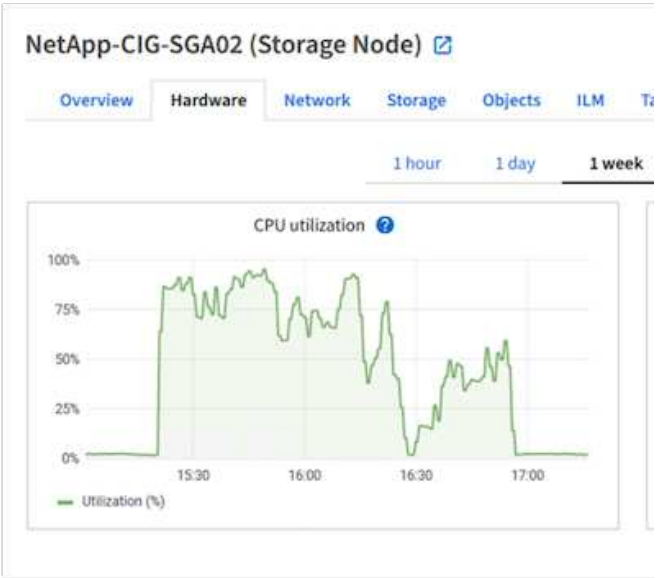
ⓘ 녹색은 적중 데이터를 보여주고, 주황색은 미스 데이터를 보여줍니다.

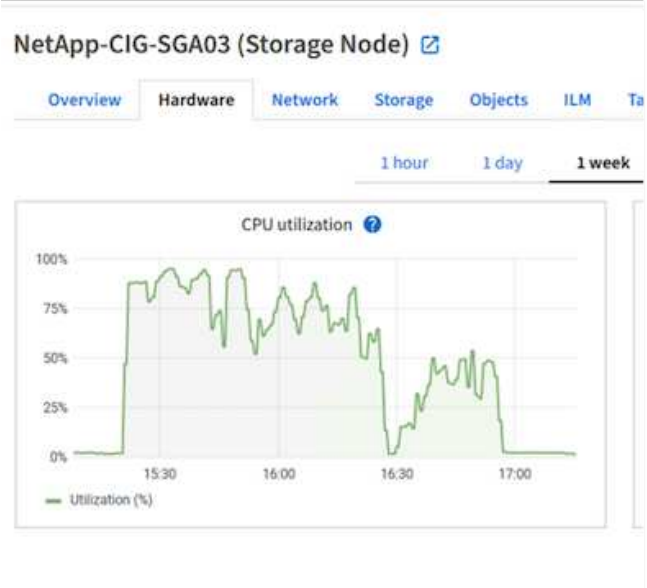
StorageGRID 에서 검색을 위해 쿼리를 실행하면 StorageGRID 에서 S3를 검색하는 속도가 다음 이미지에 표시됩니다.

StorageGRID 하드웨어 사용

StorageGRID 인스턴스에는 로드 밸런서 1개와 StorageGRID 컨트롤러 3개가 있습니다. 세 개의 컨트롤러 모두의 CPU 사용률은 75%에서 100%입니다.







NetApp 스토리지 컨트롤러를 탑재한 SmartStore - 고객 혜택

- 컴퓨팅과 스토리지 분리. Splunk SmartStore는 컴퓨팅과 스토리지를 분리하여 독립적으로 확장할 수 있도록 도와줍니다.
- 주문형 데이터. SmartStore는 필요에 따라 데이터를 컴퓨팅에 가깝게 가져오고 컴퓨팅 및 스토리지 탄력성과 비용 효율성을 제공하여 대규모로 더 오랫동안 데이터를 보존할 수 있도록 해줍니다.
- **AWS S3 API** 호환. SmartStore는 AWS S3 API를 사용하여 AWS S3 및 S3 API 호환 객체 저장소(예: StorageGRID) 인 복원 스토리지와 통신합니다.
- 보관 요구 사항과 비용이 줄어듭니다. SmartStore는 오래된 데이터(웜/콜드)에 대한 저장 요구 사항을 줄여줍니다. NetApp 스토리지는 데이터 보호, 장애 처리 및 고가용성 처리를 제공하므로 데이터의 단일 사본만 필요합니다.
- 하드웨어 오류. SmartStore 배포에서 노드 장애가 발생해도 데이터에 액세스할 수 없는 것은 아니며 하드웨어 장애나 데이터 불균형으로 인해 인덱서가 훨씬 빠르게 복구됩니다.
- 애플리케이션 및 데이터 인식 캐시.
- 필요에 따라 인덱서를 추가-제거하고 클러스터를 설정-해제합니다.

- 스토리지 계층은 더 이상 하드웨어에 구속되지 않습니다.

결론

Splunk Enterprise는 보안, IT, DevOps 팀 전반에 걸쳐 성과를 이끌어내는 시장 선도적인 SIEM 솔루션입니다. 우리 고객의 조직 전반에서 Splunk 사용이 상당히 증가했습니다. 따라서 더 많은 데이터 소스를 추가하는 동시에 더 오랜 기간 동안 데이터를 보관해야 하며, 이로 인해 Splunk 인프라에 부담이 가해집니다.

Splunk SmartStore와 NetApp StorageGRID 의 조합은 조직이 SmartStore와 StorageGRID 개체 스토리지를 통해 수집 성능을 개선하고 여러 지리적 지역에 걸쳐 Splunk 환경에 대한 확장성을 높일 수 있는 확장 가능한 아키텍처를 제공하도록 설계되었습니다.

추가 정보를 찾을 수 있는 곳

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

- "NetApp StorageGRID 문서 리소스"
- "NetApp 제품 문서"
- "Splunk Enterprise 문서"
- "Splunk Enterprise SmartStore 정보"
- "Splunk Enterprise 분산 배포 매뉴얼"
- "Splunk Enterprise 인덱서 및 인덱서 클러스터 관리"

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 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.