



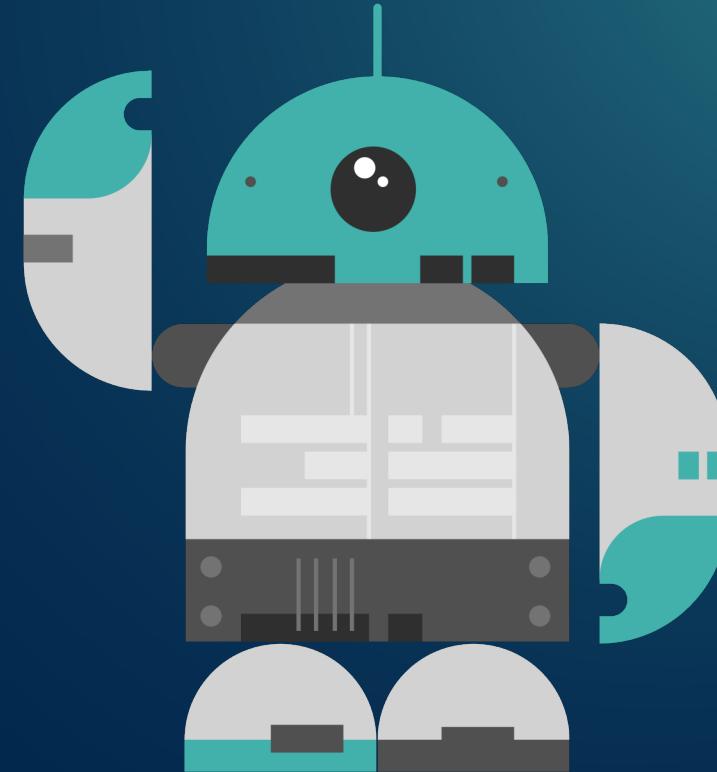
# Amazon OpenSearch

## What's New

Solution Architect  
East, Kim

# Content

- OpenSearch Overview
- What's new 2024





# OpenSearch Project

OpenSearch는 Apache 2.0 라이선스가 부여된 검색 및 분석 제품군으로, [OpenSearch](#), [OpenSearch 대시보드](#), 고급 이상 징후 탐지, 알림, 통합 가시성 및 보안 분석을 제공하는 플러그인 제품군으로 구성되어 있습니다.



> 2억+ 다운로드

OpenSearch project downloads since launch



Top 4 검색 엔진

DB-Engines ranking



55+

Partners and growing



10k+ pull request merged

200%+ growth



다양한 서비스 프로바이더

AWS, Oracle, Aiven-Azure, Bonsai-GCP

# OpenSearch platform

## OpenSearch Core



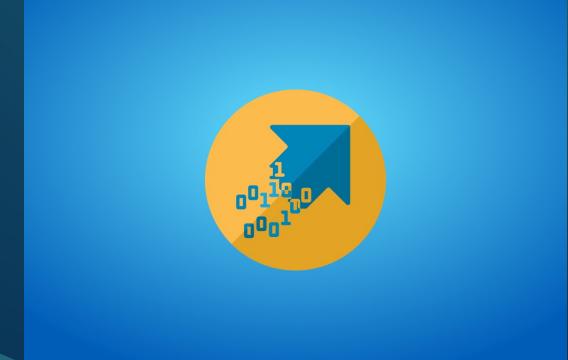
Vector 데이터베이스가 통합된  
강력한 분산 검색 및 분석 엔진

## OpenSearch Dashboards



OpenSearch용 오픈 소스  
데이터 시각화 플랫폼 및 UI

## Data Prepper



데이터 증강, 변환, 집계를 위한  
경량 데이터 수집기



# Amazon OpenSearch Service

Amazon OpenSearch Service는 보안 및 운영 데이터의 실시간 검색, 모니터링, 분석을 안전하게 지원합니다.



## Managed

인기 있는 오픈 소스 솔루션을 사용하여 운영 효율성 향상



## Secure

데이터 센터, 네트워크 아키텍처, 내장된 인증을 통해 데이터를 감사하고 보호



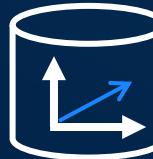
## Cost conscious

더 중요한 업무를 위해 시간 및 리소스 최적화



## Observability

머신 러닝, 경고, 시각화를 위한 오픈 소스 솔루션을 통해 시스템 문제를 감지, 분석, 해결 가능



## Vector search

Semantic search 및 RAG, k-NN 검색과 embeddings, sparse, hybrid, multi-modal 지원

# Amazon OpenSearch Service

## helps Builders ingest, search, visualize, and analyze data



### Data Engineer

데이터를 전처리 하기 위한 파이프라인을 구축하세요.

### Search Engineer

사용자에게 최상의 검색 환경을 제공하세요.

### Data Scientist

Vector 데이터베이스를 활용하여 차세대 AI Application을 구축하세요.

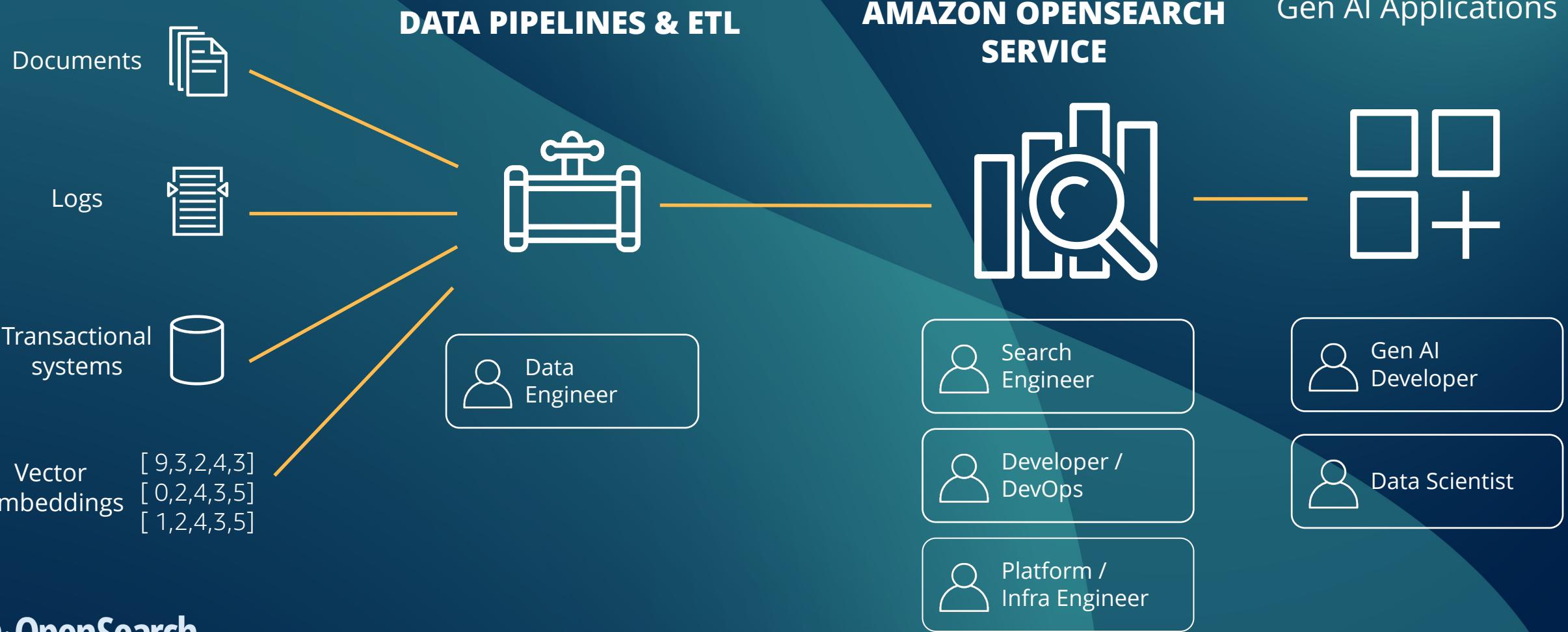
### Platform Engineer

비용 효율적인 관리형 클러스터를 실행하고 관리하세요.

### DevOps Engineer

로그, 트레이스, 메트릭을 시각화 하고 쿼리를 작성하기 위한 도구를 개발하세요.

# Amazon OpenSearch Service landscape





# Amazon OpenSearch Serverless

인프라나 인덱스 및 색드 전략에 대한 걱정  
없이 AWS 클라우드에서 OpenSearch를  
실행할 수 있습니다.



## Easy to administer

클러스터의 크기 조정, 확장, 튜닝, 색드,  
인덱스 수명 주기 관리가 필요 없음



## Fast

리소스를 자동으로 확장하여 일관되게 빠른  
데이터 수집 속도와 쿼리 응답 시간을 유지



## Ecosystem

동일한 OpenSearch 클라이언트, 파이프라인  
및 API를 사용하여 몇 초 만에 시작



## Cost-effective

사용한 리소스에 대해서만 지불



# Amazon OpenSearch Ingestion

POWERED BY  Data Prepper



## Reduce cost

데이터 중복 제거, 샘플링, 필터링하고 노이즈가 있는 데이터를 저렴한 스토리지로 라우팅



## Enforce data quality

Observability, 보안 조사/해결 시간을 가속화하는 스키마를 채택하여 데이터를 변환, 필터링, 강화



## Protect sensitive data

민감한 정보를 수정하고 난독화  
데이터 레지던시 법률 준수를 유지하기 위해  
데이터를 라우팅



## Simplify integrations

통합 가시성을 지원하는 자동 확장 데이터  
파이프라인 및 Zero-ETL

# OpenSearch Ingestion simplifies getting data

## Data sources

- Amazon S3:
  - CSV/JSON/Parquet/Avro/**Iceberg**
- http/https
- FluentBit/FluentD
- OpenTelemetry: Logs/Metrics/Traces
- Amazon DynamoDB
- **Amazon DocumentDB**
- **Apache Kafka**
- **Amazon Managed Streaming for Kafka**
- **Amazon Kinesis Data Streams**
- **Confluent Cloud**
- **Elasticsearch**
- **OpenSearch**
- **Amazon OpenSearch Service**
- **Amazon Security Lake**



Amazon OpenSearch  
Ingestion Service

## Pay-as-you-go Serverless Pipelines

- 수집 요구사항에 맞춰 자동 확장
- 비용 절감을 위한 수동 일시 중지-재개
- Min-Max 임계값으로 비용 관리

## Reliability, Security, Scale

- Multi-AZ Architecture
- 영구적인 버퍼링
- End-to-End
- DLQ 지원

# OpenSearch Ingestion simplifies getting data

## Processors

- Stateless:
  - grok, filter, mutate, drop, parse, substitute, obfuscate, translate, annotate, **AWS Lambda**, truncate, and **GeoIP** processors
- Stateful:
  - 집계, 샘플링 및 조건부 라우팅 처리
  - 로그의 메트릭과 트레이스의 Red 메트릭으로 이동중인 데이터에 대한 신호 필터링
  - 시계열 데이터 및 로그 패턴에 대한 이상 징후(Anomaly detection) 탐색을 사용한 이동중인 데이터와 ML 통합



Amazon OpenSearch  
Ingestion Service

## Sinks



Amazon OpenSearch  
Service



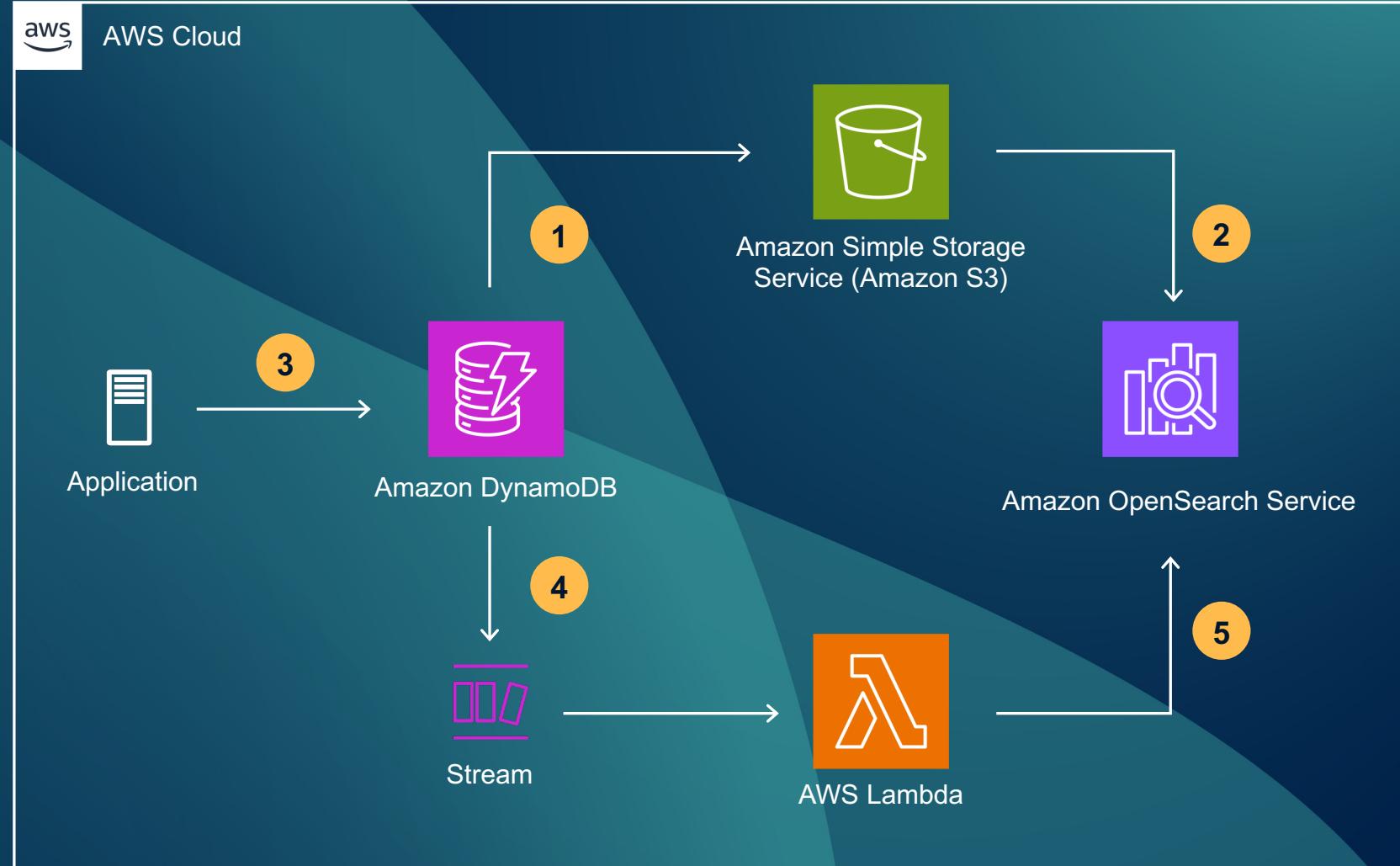
Amazon S3



Amazon Security Lake

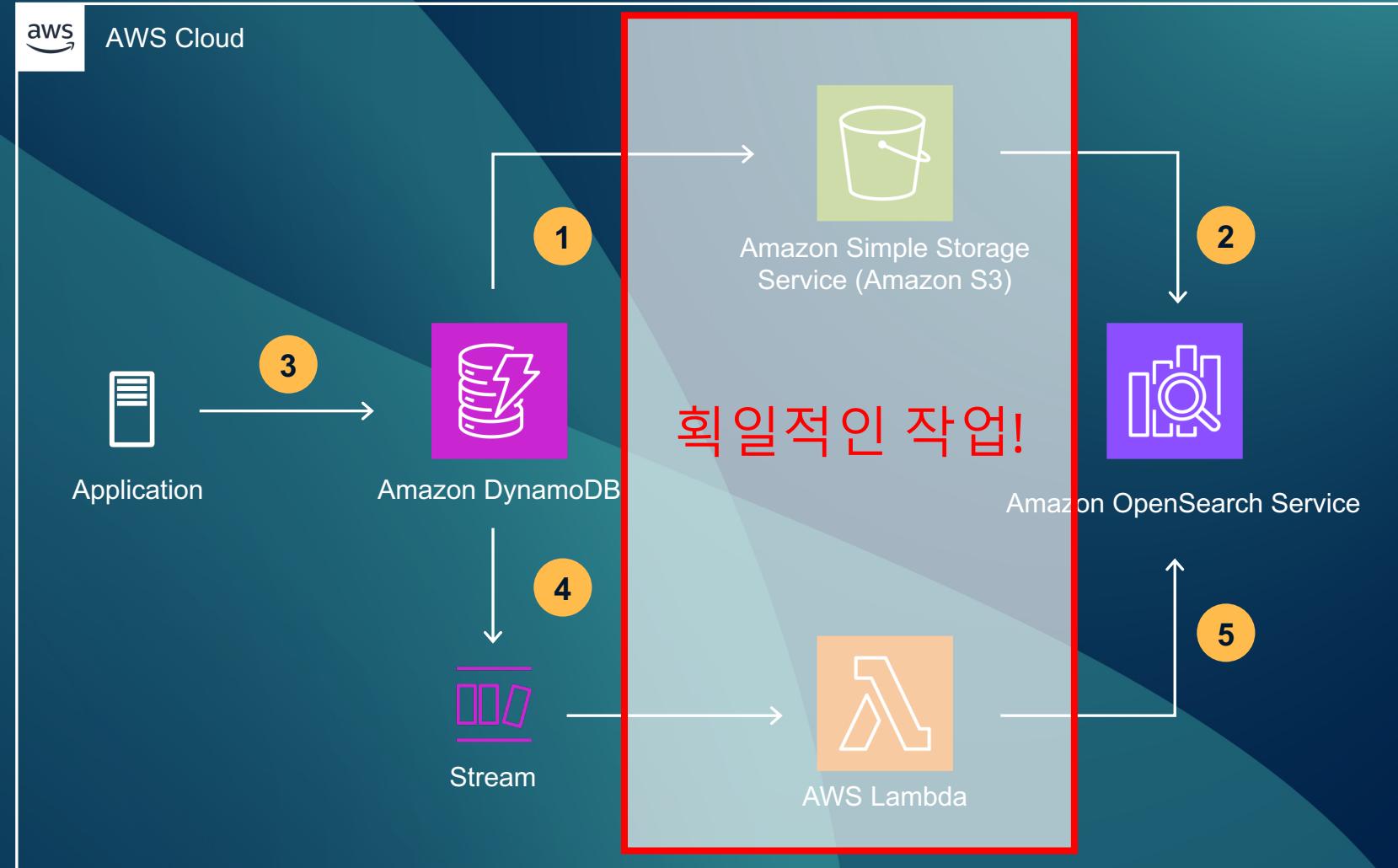
# Searching Data today

1. ExportToPointInTime을 사용하여 스냅샷을 S3로 내보내기
2. S3에서 OpenSearch Service로 스냅샷을 입력
3. 1과 2가 진행되는 동안 애플리케이션이 DynamoDB에 업데이트를 계속 전송
4. DynamoDB에 대한 업데이트가 스트림에 나타납니다.
5. AWS Lambda로 스트림을 처리하고 OpenSearch로 푸시



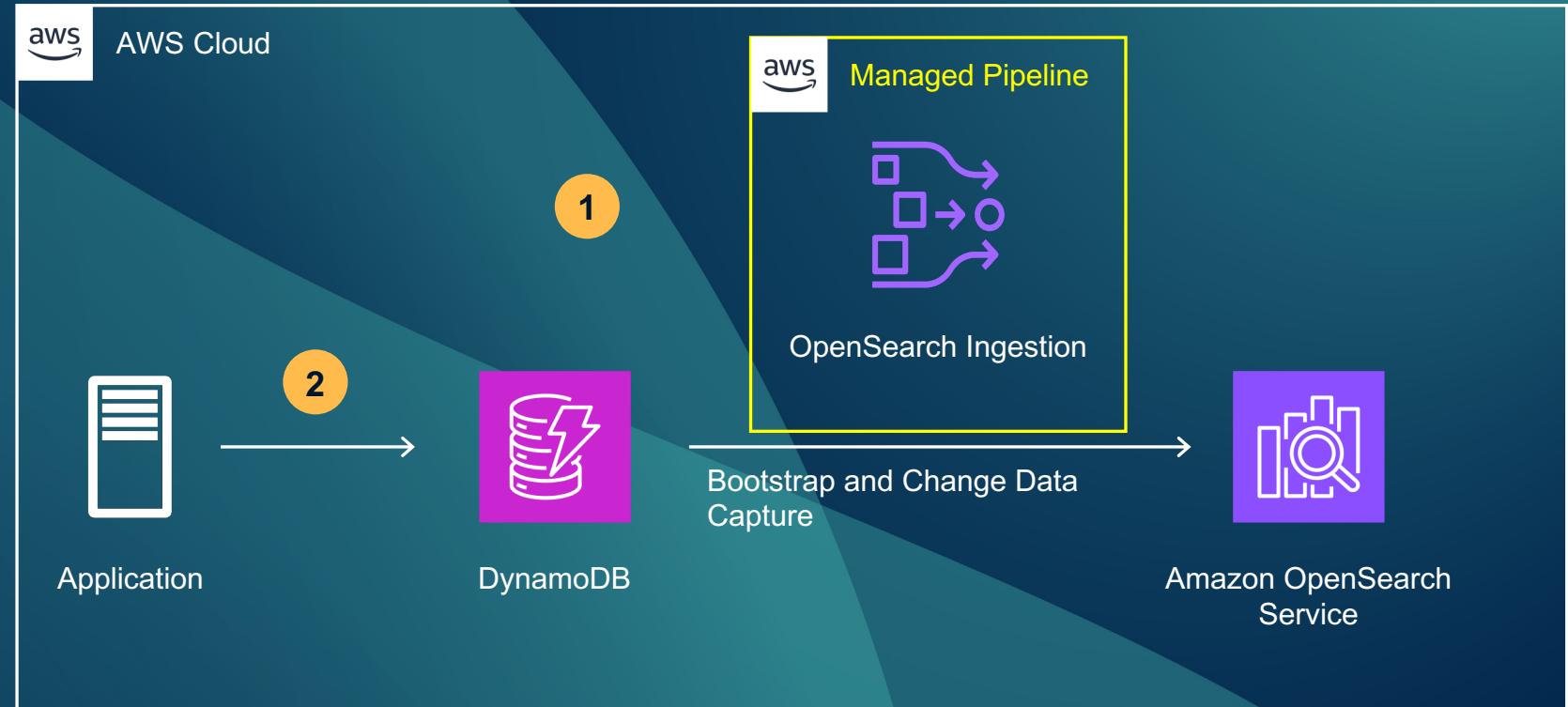
# Searching Data today

1. ExportToPointInTime을 사용하여 스냅샷을 S3로 내보내기
2. S3에서 OpenSearch Service로 스냅샷을 입력
3. 1과 2가 진행되는 동안 애플리케이션이 DynamoDB에 업데이트를 계속 전송
4. DynamoDB에 대한 업데이트가 스트림에 나타납니다.
5. AWS Lambda로 스트림을 처리하고 OpenSearch로 푸시



# Zero ETL integration w/ OpenSearch

1. DynamoDB와 OpenSearch Service 간의 파이프라인 설정
2. 애플리케이션은 계속해서 DynamoDB에 업데이트를 전송하고, 이는 다음과 같이 동기화



# What's new 2024

# What's new 2024

신규 기능	내용
<a href="#">OpenSearch JSON Web Token(JWT) 인증/인가 지원</a>	OpenSearch API 사용시 JWT 토큰을 사용하여 인증/인가를 수행할 수 있습니다.
<a href="#">OpenSearch Ingestion Blueprints 지원</a>	AWS Console에서 전체 텍스트 검색을 사용하여 청사진을 검색할 수 있는 새로운 사용자 인터페이스를 제공합니다. 청사진에서는 Amazon OpenSearch Service로 데이터를 수집할 수 있는 모든 소스를 쉽게 찾을 수 있습니다.
<a href="#">Zero-ETL 통합 지원</a>	Amazon OpenSearch Service와의 Amazon DocumentDB 제로 ETL 통합은 OpenSearch API를 사용하여 Amazon DocumentDB 문서에서 퍼지 검색, 교차 컬렉션 검색 및 다국어 검색과 같은 고급 검색 기능을 고객에게 제공합니다.
<a href="#">OpenSearch New UI 지원</a>	단일 엔드포인트를 통해 관리형 도메인과 서비스 컬렉션 아우르는 현대화된 통합 분석 환경을 지원합니다.
<a href="#">OpenSearch CloudWatch Logs 통합 지원</a>	CloudWatch 고객은 OpenSearch의 Piped Processing Language(PPL)와 OpenSearch SQL을 활용할 수 있도록 통합 환경을 지원합니다.
<a href="#">OpenSearch Security Lake Data 통합 지원</a>	Amazon Security Lake와의 Zero-ETL 통합을 통해 OpenSearch를 활용하여 보안 데이터를 바로 쿼리 할 수 있습니다.

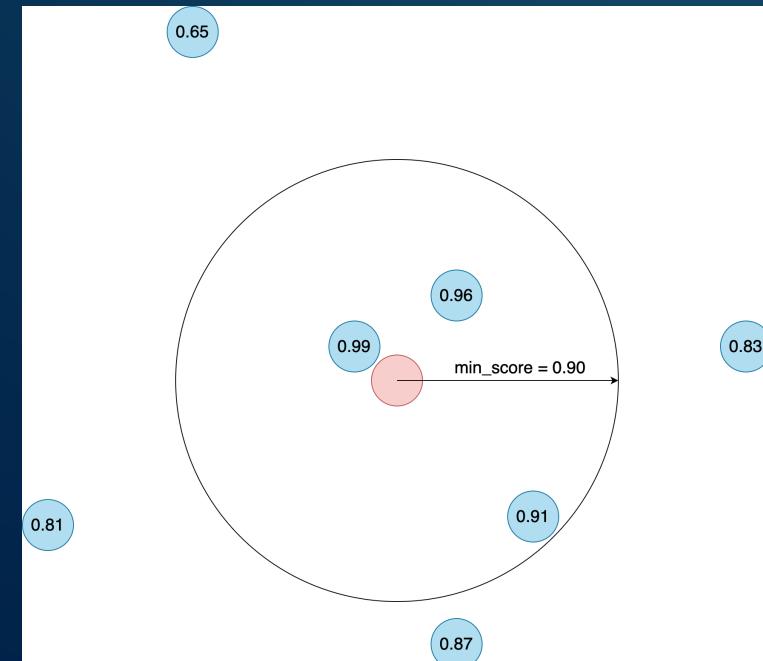
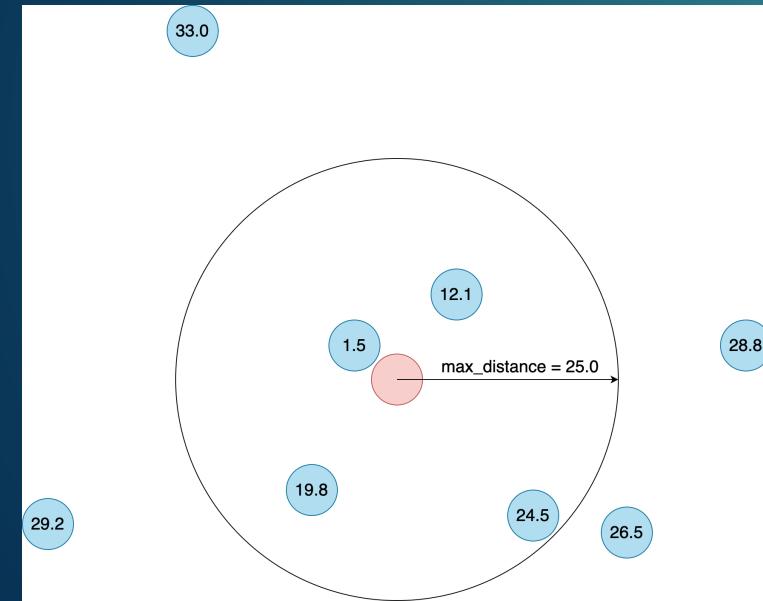
# What's new 2024

신규 기능	내용
<a href="#">OpenSearch 디스크 최적화 벡터 엔진 지원</a>	벡터 엔진의 양자화를 사용하여 인덱스를 압축해 고 정밀도 벡터를 사용하는 메커니즘으로 리콜을 향상시킵니다.
<a href="#">OpenSearch OR1 인스턴스 지원</a>	OR1은 기존 인스턴스보다 최대 30% 향상된 가격 대비 성능을 제공하며(내부 벤치마크 기준) Amazon S3를 사용하여 99.999999999의 내구성을 제공합니다.
<a href="#">OpenSearch Coordinator Node 지원</a>	전용 코디네이터 노드를 사용하면 데이터 노드의 트래픽 조정 및 OpenSearch 대시보드 호스팅 책임이 줄어들어 리소스 사용률을 높이고 클러스터의 전반적인 효율성을 개선할 수 있습니다.
<a href="#">OpenSearch 단일 클러스터로 최대 1000개 노드 스케일링 지원</a>	Amazon Security Lake와의 제로 ETL 통합을 기능을 사용하여 OpenSearch를 통해 보안 데이터를 바로 쿼리하고 분석할 수 있습니다.
<a href="#">OpenSearch Custom Plugin 지원</a>	OpenSearch 기능을 확장하고 웹 사이트 검색, 로그 분석, 애플리케이션 모니터링 및 관찰성과 같은 애플리케이션에 대한 개인화된 경험을 제공할 수 있는 새로운 플러그인 관리 옵션인 사용자 지정 플러그인을 지원합니다.
<a href="#">OpenSearch Serverless 기능 개선</a>	Binary Vector 및 FP16 압축을 지원하여 필요한 메모리 용량을 낮춤에 따라 비용을 절감할 수 있으며 최대 30TB까지 확장을 지원합니다.

# Radial Search

- Vector 검색 결과에 대해 Vector간 거리나 최소 점수 기준 필터링
- Vector 간 거리는 max\_distance로 지정
- 최소 점수는 min\_score로 지정

Engine	Filter	Nested Field	Search Type
Lucene	true	false	approximate
Faiss	true	true	approximate



# Ingestion Pipeline Blueprints

NEW

Search Blueprints

All

Use case

- DLQ Replay
- GeoIP
- Log Analytics
- Metrics Analytics
- Migration
- Security Analytics
- Streaming
- Trace Analytics
- ZeroETL

Service

- ALB Logs
- Apache Kafka
- Apache Logs
- CloudTrail Logs
- Confluent Cloud
- DocumentDB
- DynamoDB
- ELB Logs
- HTTP
- KDS
- MSK
- MSK Serverless
- OpenSearch
- OpenTelemetry
- S3
- Security Lake
- Transit Gateway Logs
- VPC Flow Logs
- WAF Logs

Ingestion pipeline blueprints (39)

Create scalable data processing pipelines for stream and batch processing of large scale data sets.

Blank

ALB access logs

Apache logs

Apache logs sampling

CloudTrail logs

Confluent Cloud

DLQ replay

ELB access logs

GeoIP enrichment for Apache logs

GeoIP enrichment for VPC Flow logs

GeoIP enrichment for WAF Access logs

Log aggregation with conditional routing

Logs to metrics

Logs to metrics anomaly detection

Managed streaming for Apache Kafka

OpenSearch data migration

OpenTelemetry logs

OpenTelemetry metrics

OpenTelemetry trace analytics

Process application logs

S3 logs

S3 Parquet or Avro logs

S3 scan

S3 schedule scan

S3 selective download

S3 sink conditional routing

Security Lake S3 Parquet OCSF

Self managed Apache Kafka

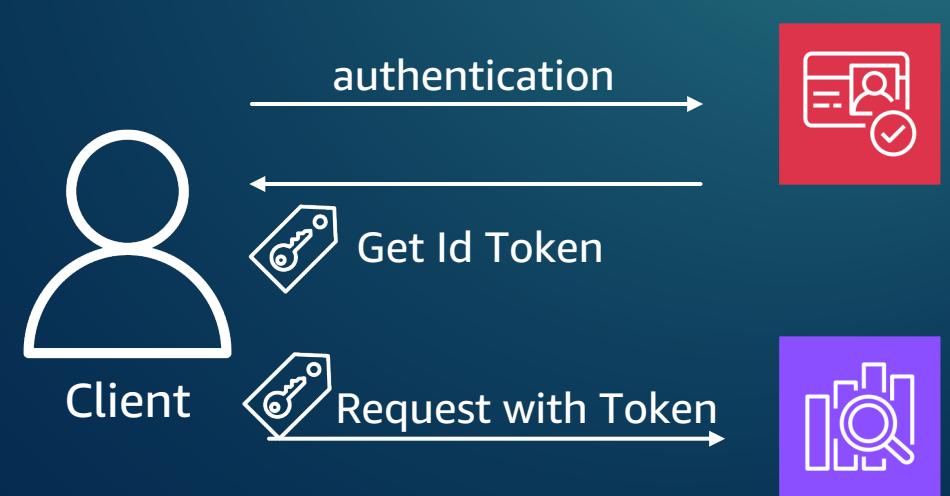
Self managed OpenSearch/Elasticsearch

Sink filter with Avro schema Parquet codec

Cancel Select blueprint

# Json Web Token Support

- JWT를 통한 API 인증 지원
- 대시보드 액세스에서는 JWT 사용  
불가
- Amazon cognito User Pool 연동  
지원



# Zero-ETL integrations with transactional systems



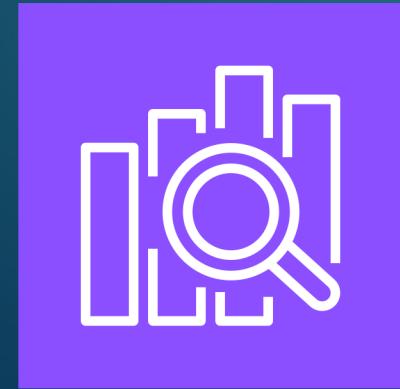
Amazon DynamoDB

일관된 짧은 자연 시간,  
내구성, 유연성을 갖춘 NoSQL  
데이터베이스 서비스



Amazon DocumentDB

복잡한 중첩 문서 구조를 지원하는  
유연한 NoSQL 문서 데이터베이스  
서비스



Amazon  
OpenSearch Service

고객 애플리케이션을 위한  
풍부한 검색 환경 제공

# OpenSearch dashboards(Next-Gen)



Amazon OpenSearch Service  
*Managed Cluster 1*



Amazon OpenSearch Service  
*Managed Cluster 2*



Amazon OpenSearch Service  
*Managed Cluster 3*



Amazon OpenSearch Service  
*Serverless Collection 1*



Amazon OpenSearch Service  
*Serverless Collection 2*



OpenSearch  
Dashboards  
*Endpoint 1*



OpenSearch  
Dashboards  
*Endpoint 2*



OpenSearch  
Dashboards  
*Endpoint 3*



Next-gen  
OpenSearch UI  
*Single Endpoint*



OpenSearch  
Dashboards  
*Endpoint 4*



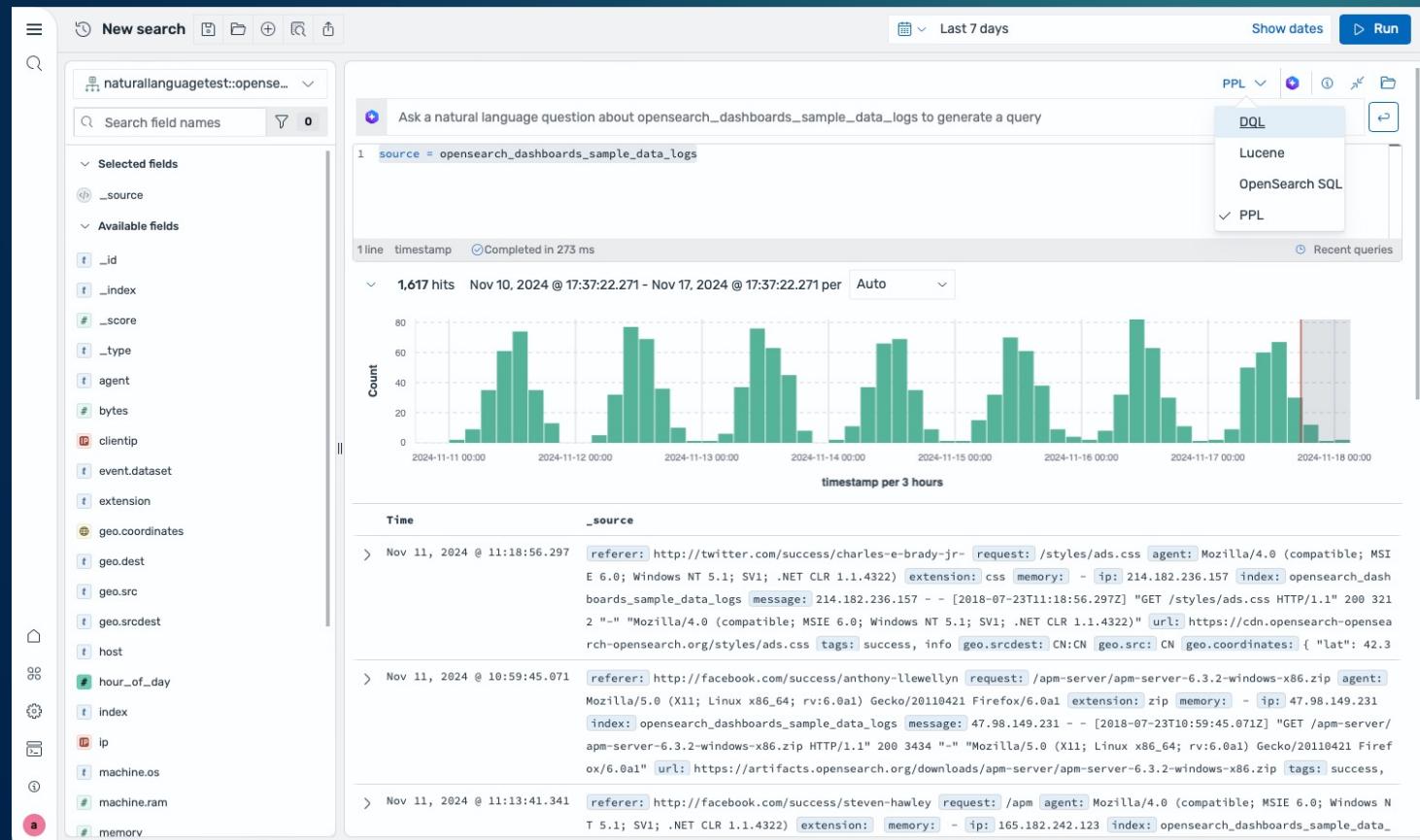
OpenSearch  
Dashboards  
*Endpoint 5*

# Next-Gen OpenSearch UI

NEW

## A SINGLE ENDPOINT FOR OPERATIONAL DATA

- **Multiple data sources** – OpenSearch Service (관리형, 서버리스)와 S3 데이터 소스
  - **Modernized user interface** – 사용 편의 향상
  - **Introducing Workspaces** – 사용자별 격리된 환경 제공
  - **Enhanced data exploration** – 새로운 검색 도구 제공



# Integrated natural language to query skills

Screenshot of the OpenSearch interface showing the integration of natural language queries with PPL (Parsed Precision Language) and visualizations.

**Search Bar:** New search, Last 5 days, Show dates, Run.

**Selected fields:** \_source.

**Available fields:** \_id, \_index, \_score, \_type, agent, bytes, clientip, event.dataset, extension, geo.coordinates, geo.dest, geo.src, geo.srcdest.

**Query:** filter out 200s and url does not contain cdn

**PPL query generated:**

```
source=opensearch_dashboards_sample_data_logs| where NOT response='200' AND NOT LIKE(url, '%cdn%')
```

**Statistics:** 1 line timestamp, Completed in 283 ms, 70 hits, Nov 12, 2024 @ 21:19:48.698 - Nov 17, 2024 @ 21:19:48.698 per Hour.

**Bar Chart:** Count vs timestamp per hour.

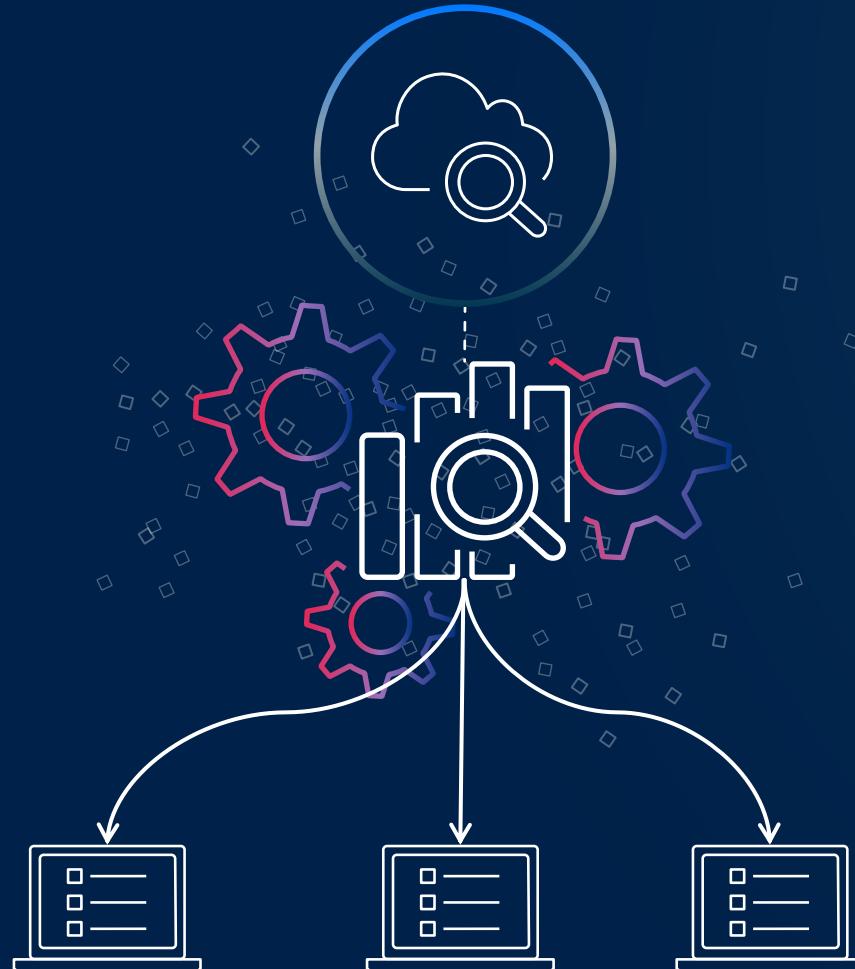
timestamp per hour	Count
2024-11-13 00:00	1
2024-11-13 12:00	3
2024-11-14 00:00	1
2024-11-14 12:00	3
2024-11-15 00:00	1
2024-11-15 12:00	5
2024-11-16 00:00	1
2024-11-16 12:00	4
2024-11-17 00:00	1
2024-11-17 12:00	3

**Log Entry:**

```
> Nov 13, 2024 @ 16:40:29.179 referer: http://www.opensearch-opensearch.com/success/klaus-dietrich-flade request: /beats/filebeat/filebe at-6.3.2-linux-x86.tar.gz agent: Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1 extension: gz
```

# Analyze CloudWatch Logs using OpenSearch Service

A BETTER-TOGETHER EXPERIENCE FOR AWS CUSTOMERS



Zero-ETL integration: 데이터 이동 또는 중복 없이 OpenSearch에서 CloudWatch 로그를 쿼리하고 분석할 수 있습니다.

Enhanced analytics: SQL, PPL, 기본 제공 대시보드를 사용해 VPC, AWS WAF, AWS CloudTrail용 AWS 서비스 로그에 대한 분석이 가능합니다.

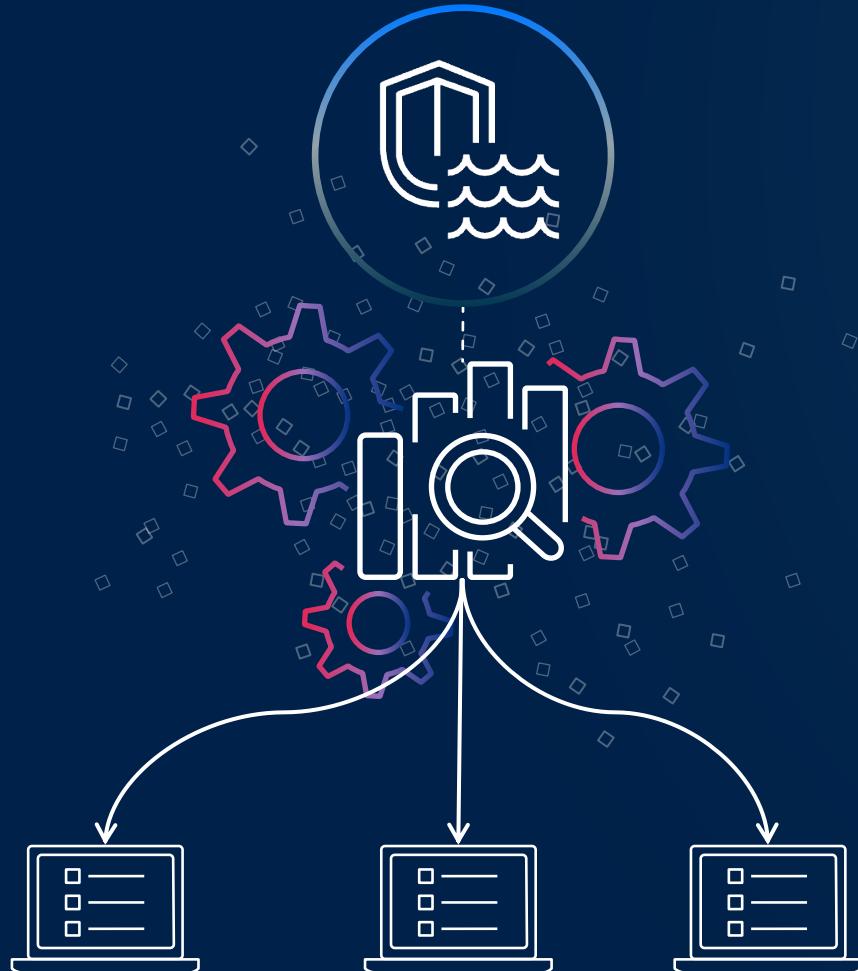
사용자 정의 인덱싱 또는 구체화된 보기를 선택하여 선택한 데이터 세트에서 더 빠른 분석이 가능합니다.

Amazon CloudWatch와 OpenSearch Service 기능의 장점을 모두 활용할 수 있습니다.

통합 로그 관리: 수집, 저장, 분석을 중앙 집중화하면서 운영 오버헤드와 비용을 절감할 수 있습니다.

# Analyze Security Lake data using OpenSearch Service

A BETTER-TOGETHER EXPERIENCE FOR AWS CUSTOMERS



Zero-ETL integration: 데이터 중복이나 사용자 정의 파이프라인 없이 OpenSearch Service에서 직접 Security Lake 데이터를 쿼리하고 분석할 수 있습니다.

유연한 쿼리 및 색인: 온디맨드 색인 및 구체화된 보기 옵션과 함께 직접 쿼리를 위해 SQL(구조화된 쿼리 언어) 또는 PPL(파이프로 처리 언어)을 사용할 수 있습니다.

OCSF 준수 : 사전 구축된 쿼리 및 대시보드를 활용하여 다양한 데이터 소스에서 신속한 보안 인사이트를 얻을 수 있습니다.

비용 효율적인 보안 분석: 특정 데이터 세트를 선택적으로 색인하고 가속화하여 성능과 비용을 최적화합니다.

통합 스키마에서 포괄적인 보안 가시성 제공합니다.

# OpenSearch OR1 optimized instance family

ADDRESSING COST AND DURABILITY



**80% indexing throughput improvement**

높은 색인 처리량



**30% price performance improvement**

전체 비용 절감



**11 9s of durability**

내구성이 뛰어난 Amazon S3에 데이터 색인



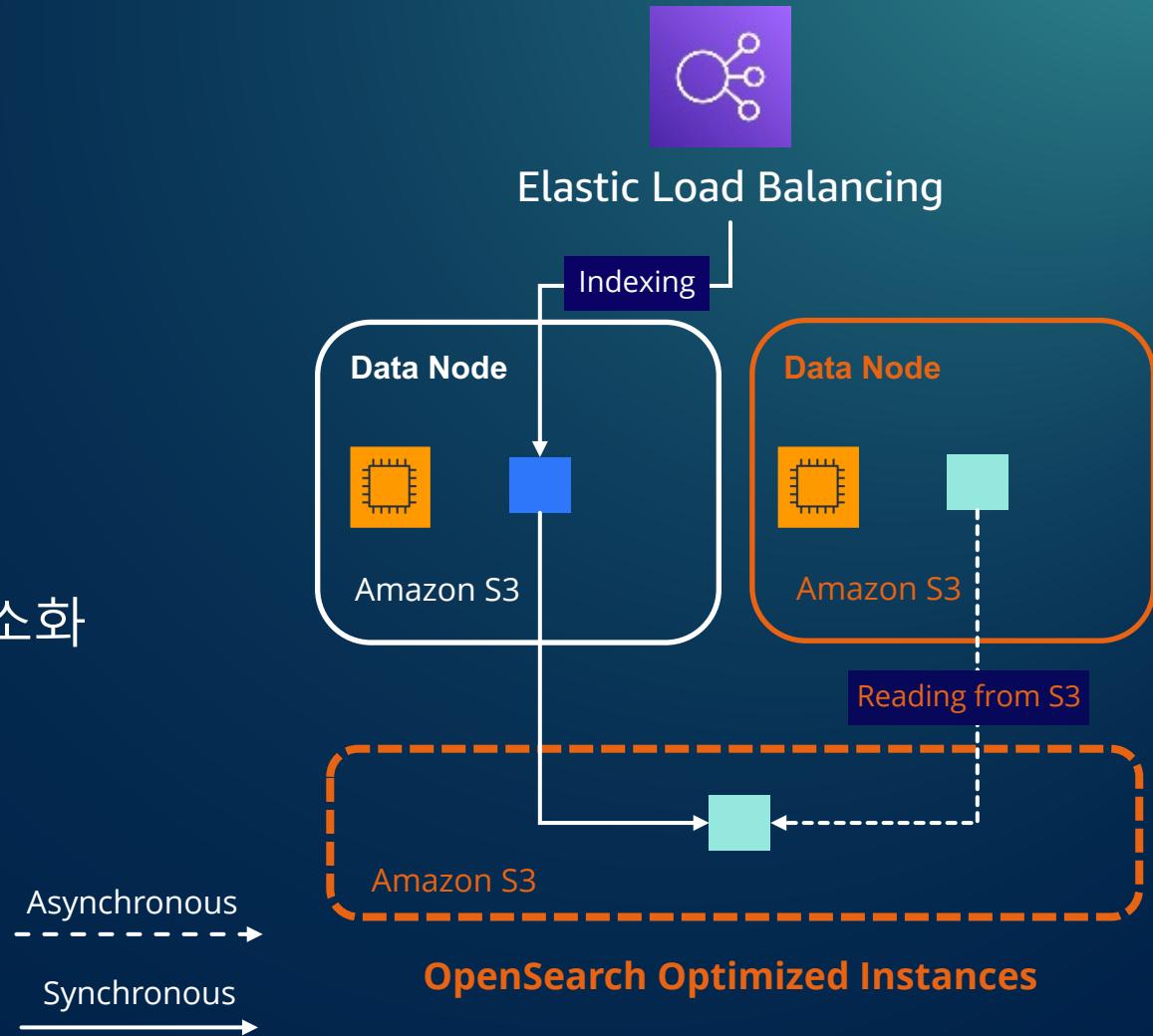
**Automatic recovery**

Red 인덱스에서 자동으로 복구

NEW  
In-place upgrade  
support

# OR1 indexing and replication process

- 물리적 복제
- Amazon S3를 백업 저장소로 활용
- 분리된 인덱싱 및 복제 작업 수행
- Red 인덱스(RP0)에서 자동 복구
- Blue/Green 업데이트시 중복 인덱싱 부하 최소화
- OpenSearch 2.11 이상 사용 가능



# Introducing dedicated coordinator nodes

## Request processing



Requests land on a coordinator

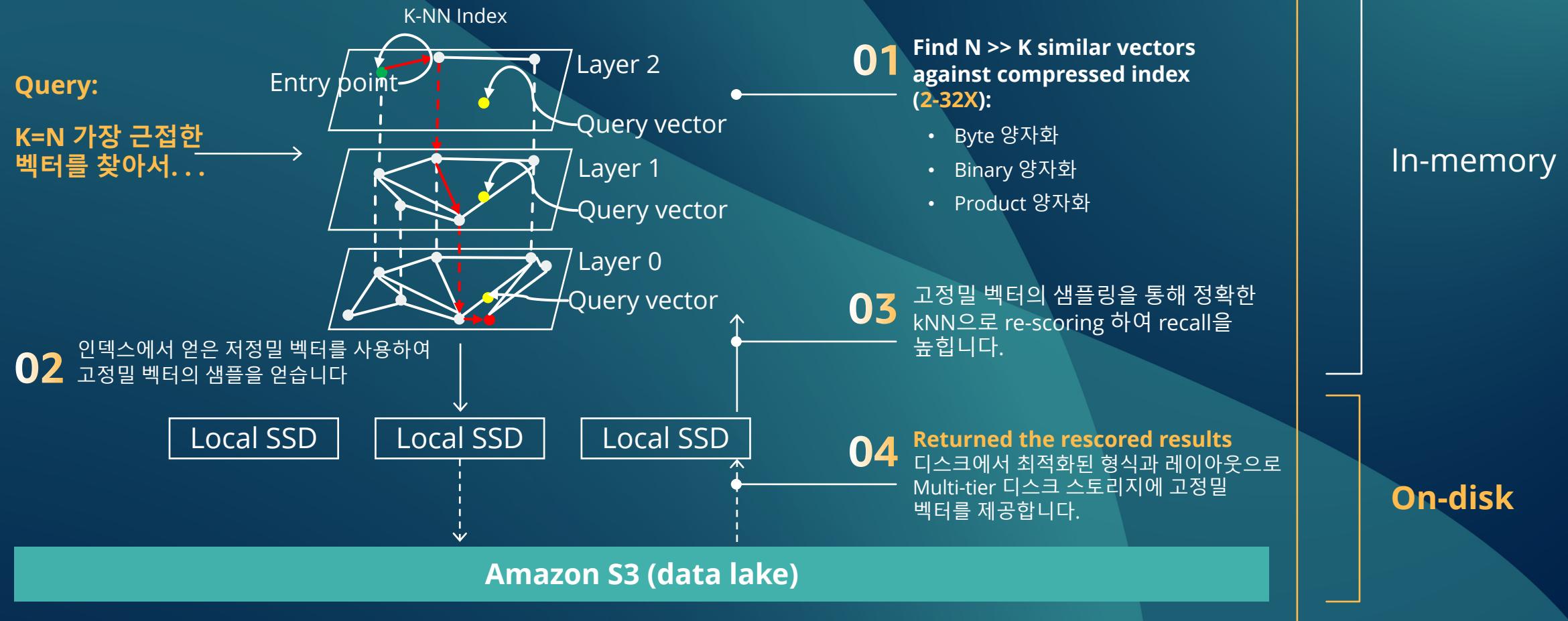
Requests are queued and distributed to data nodes with shards for the request

Data nodes process the requests locally and send results back to the coordinator

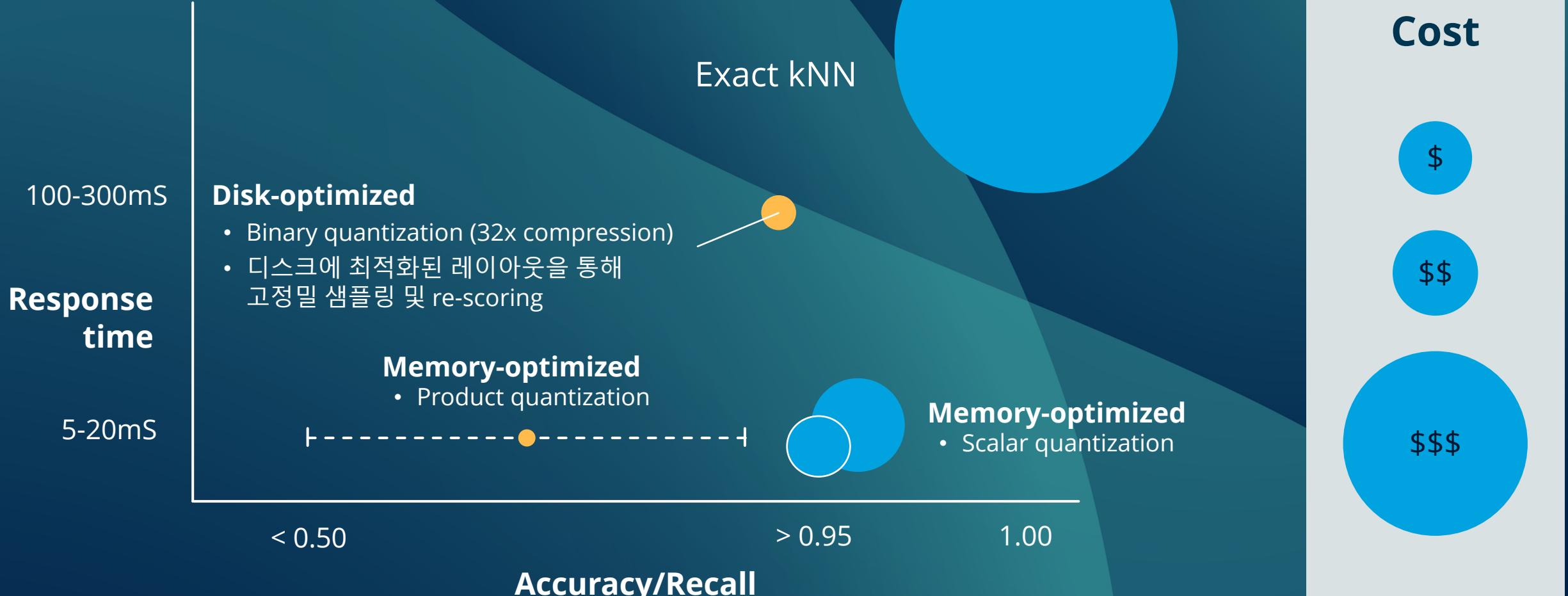
Coordinator re-aggregates and sends the response

- 15% 더 높은 색인 처리량
- 20% 더 높은 검색 처리량
- 클러스터 복원력 향상
- 효율적인 확장
- IP 예약 감소
- Dashboard 컴퓨팅 리소스 절감

# Disk-optimized vector search: Up-to-32x memory reduction



# Disk-optimized vector search



# Key points for vector cost reduction

트랜스포머 모델 선택 및 청크 전략

더 큰 인스턴스(더 적은 JVM 설치 공간)

Circuit breaker > 0.5

Scalar quantization (float16)

워크로드가 너무 큰 경우, NVMe SSD와 함께 k-NN을 사용하여 자연 시간이 짧은 디스크 액세스

예약 인스턴스

고밀도 벡터가 필요한 경우 Sparse neural 사용

# Support up to 25PB in a single cluster

노드 제한을 200에서 1000개로 증설

스토리지 제한이 4PB에서 25PB로 확장

99.99 SLA(Multi-AZ Standby 포함)

향상된 전용 클러스터 관리자로 메타데이터 관리 개선

Requires:

- *OpenSearch 2.17+*
- *UltraWarm 노드를 지원하는 Hot 노드*용 OR1 데이터 노드
- 전용 코디네이터 노드



# Extend your cluster with custom plugins

사용자 지정 언어 분석, 필터링, 집계, 정렬등 사용자 환경에 맞는 커스텀한 설정이 가능함

플러그인 업로드, 연결, 삭제를 위한 API 및 Console 지원

AWS Inspector를 통한 통합 보안 취약성 검사 가능



# New 3<sup>rd</sup> party partner plugins

Portal26 encrypted search plugin:

- NIST FIPS 140-2 인증 암호화
- BYOK 인덱스 수준 키 지원
- 인덱싱 데이터 암호화

Babel Street Match plugin:

- 정확한 이름, 조직, 주소, 날짜 매칭
- AI를 활용한 지능적 대응
- 24개 언어 콘텐츠 지원
- 보안 운영 및 규정 준수 강화



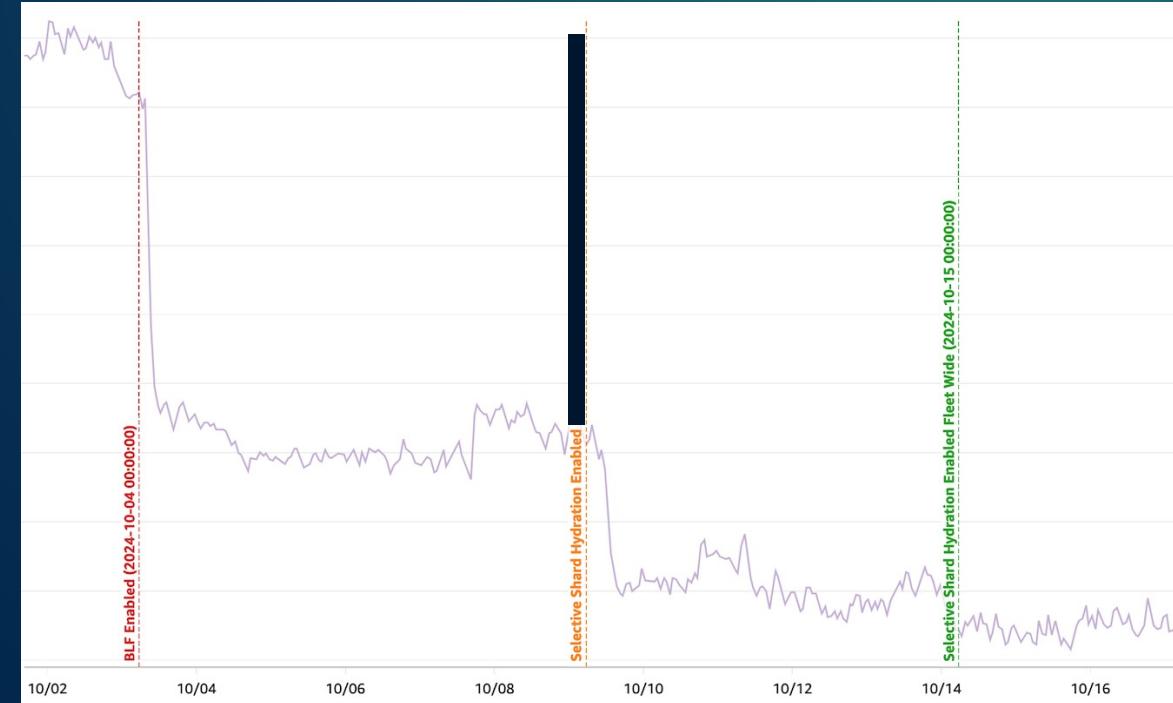
# OpenSearch Serverless Innovation

## Cost reduction:

- 0.5 OCU: 엔트리 가격 50% 절감
- Smart caching: 색인량이 많은 고객을 위한 인덱스 OCU 최대 77% 절감
- Binary vector and FP16 support: 86%의 비용 절감

## New capabilities

- 30TB 시계열 컬렉션 지원
- 새로운 필드 유형 및 집계 지원
- PIT(특정 시점 검색) 지원
- OpenSearch SQL / PPL language support



스마트 캐싱을 배포하는 동안 OCU 사용률 25% 감소

# In Summary

OpenSearch의 **Ingestion** 기능 향상과 **DynamoDB** 그리고 **DocumentDB Zero-ETL 통합**을 통해 보다 쉽게 데이터를 OpenSearch로 가져올 수 있습니다.

통합된 차세대 OpenSearch UI, 자연어 쿼리 기능, CloudWatch Logs 및 Amazon Security Lake와의 Zero-ETL 통합으로 실시간 분석을 대규모로 개선할 수 있습니다.

Sparse, Hybrid, Multi-Modal, Conversational 검색 지원 및 새로운 **디스크 최적화 벡터 엔진**을 포함한 지연 시간, 정확도, 비용 목표를 충족하는 조정 기능 등 최신 AI/ML 기반 검색 기술과 벡터 데이터베이스 기능으로 검색의 수준을 한 단계 끌어올릴 수 있습니다.

1.0\* 버전 대비 6.5배 향상된 OpenSearch 2.17 그리고 OR1 인스턴스 클래스, 전용 코디네티어 노드, 25PB 클러스터 지원, 사용자 정의 플러그인을 통해 최적의 검색 환경을 제공하고 개선된 OpenSearch 서비스로 최고의 가격 대비 선능과 확장성을 경험할 수 있습니다.

