
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

ANT221

Provision and scale OpenSearch resources with serverless

Carl Meadows

Director, Product Management
AWS

Pavani Baddepudi

Senior Product Manager
AWS

Ryan Skousen

CTO, riskCanvas
Genpact

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

OpenSearch use cases

Amazon OpenSearch Serverless

Demo

Peek into serverless architecture

Genpact use cases with OpenSearch Serverless

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Popular use cases







Text search

Natural language

Boolean queries

Relevance

Streaming data

High-volume ingest

Near real time

Distributed storage

Analysis

Time-based visualizations

Nestable statistics

Time series tools

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Amazon OpenSearch Service

Amazon OpenSearch Service securely unlocks real-time search, monitoring, and analysis of operational data



Managed: Increase operational excellence by using a popular open source solution



Secure: Audit and secure your data with a data center and network architecture and built-in certifications



Observability: Systematically detect potential threats and react to a system's state through an open source solution for machine learning, alerting, and visualization



Cost conscious: Optimize time and resources for strategic work

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



How does OpenSearch relate to Amazon OpenSearch Service?

- Community-driven, open source search and analytics suite derived from Apache 2.0 licensed Elasticsearch 7.10.2 and Kibana 7.10.2
- Distributed search engine powered by **Apache Lucene**, **OpenSearch**, and a data visualization and user interface, **OpenSearch Dashboards**
- Also includes a series of functionality adding **tools** and **plugins**
- Includes all advanced functionality ported over from **Open Distro for Elasticsearch**

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



OpenSearch project growth

>100M
OpenSearch project downloads since launch

Top 4 search engine
DB-Engines ranking

40+
Partners and growing

10K+ pull request merged
200%+ growth

Multiple service providers
AWS, Oracle, Aiven-Azure, Bonsai-GCP

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon OpenSearch Service: Managed Clusters

- **Dedicated clusters** configured and optimized for your workloads in minutes
- **Wide range** of **Amazon EC2 instances** and **storage options** to meet the needs of different workloads
- Pre-deployed for peak usage with sufficient capacity to meet cost and availability goals
- Requires expertise in OpenSearch internals (e.g., shards, index lifecycle) to size and optimize
- Regular monitoring to update sizing and configurations as workload demands change



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Making OpenSearch even easier

Empowering users to leverage OpenSearch keeping consistently high performance

Users want to focus on their business insights

Access data without thinking about provisioning, sizing, scaling, and tuning

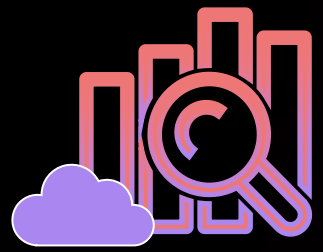
Unpredictable, *on-demand,* varying *workloads*

Diverse use cases with heterogeneous compute and storage requirements



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Introducing

Amazon OpenSearch Serverless

preview



Easy to administer

No sizing, scaling, and tuning of clusters, and no shard and index lifecycle management



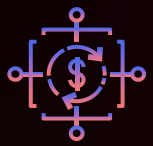
Fast

Automatically scale resources to maintain consistently fast data ingestion rates and query response times



Ecosystem

Get started in seconds using the same OpenSearch clients, pipelines, and APIs



Cost-effective

Pay only for the resources consumed

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

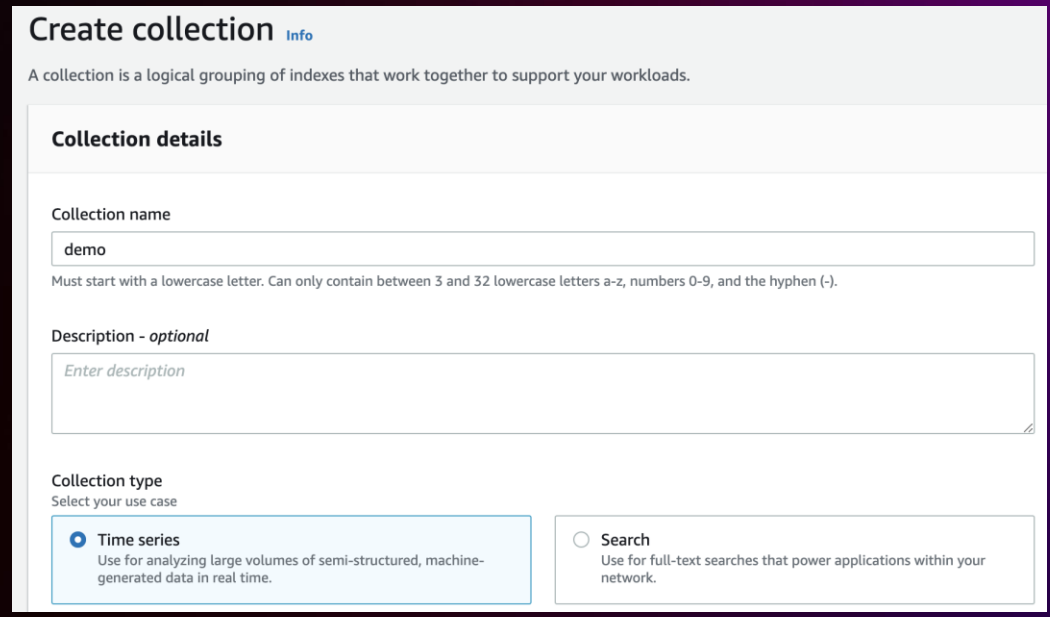
OpenSearch Serverless

Collections: A set of indexes that work together

- Separate endpoints for OpenSearch and OpenSearch Dashboards
- Can have specific or inherited access, network, and encryption policies
- Optimized for ”time series” or “search”

OpenSearch Compute Units (OCUs): Used to index and search collections

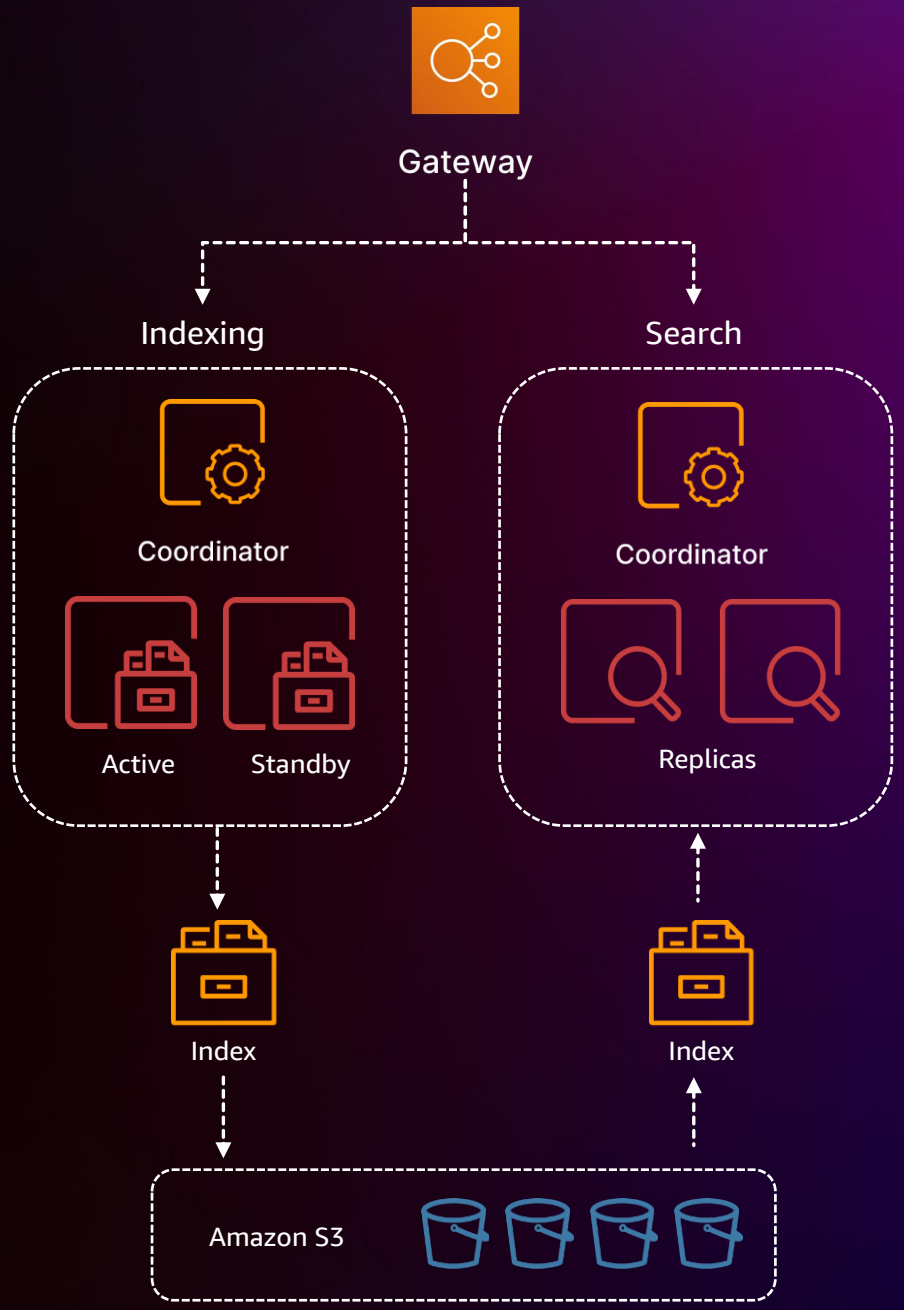
- 6 GB RAM increments (min 4 per account)
- Max OCUs can be set to control costs
- Automatically provisioned for the workload
- Shared across collections



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Technical innovations

- Storage and compute decoupled
- Separate indexing and search pipelines
- Built-in hot-warm tier
- Active-standby data nodes
- Secure service mesh
- Serverless Dashboards



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Security made easy

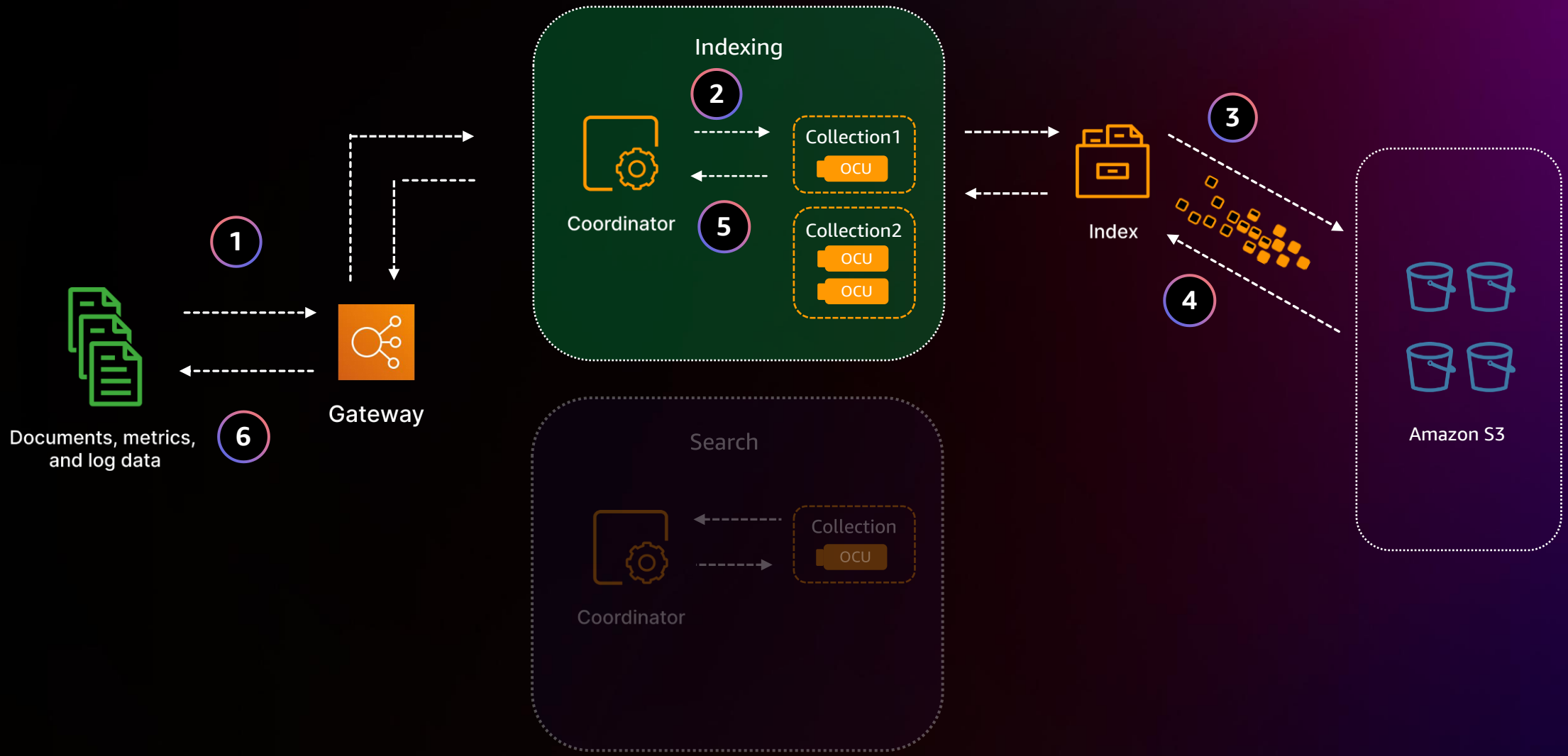
- **Encryption at rest** and **in flight** by default
- Support for **unique KMS keys** by **collection**
- Hierarchical policies
- **Role-based access control** by collection
- Network access configurable by collection
  - **VPC PrivateLink** and **Public Access**
  - Separate policies for OpenSearch API and OpenSearch Dashboards endpoints
- **SAML auth** for OpenSearch Dashboards



aws

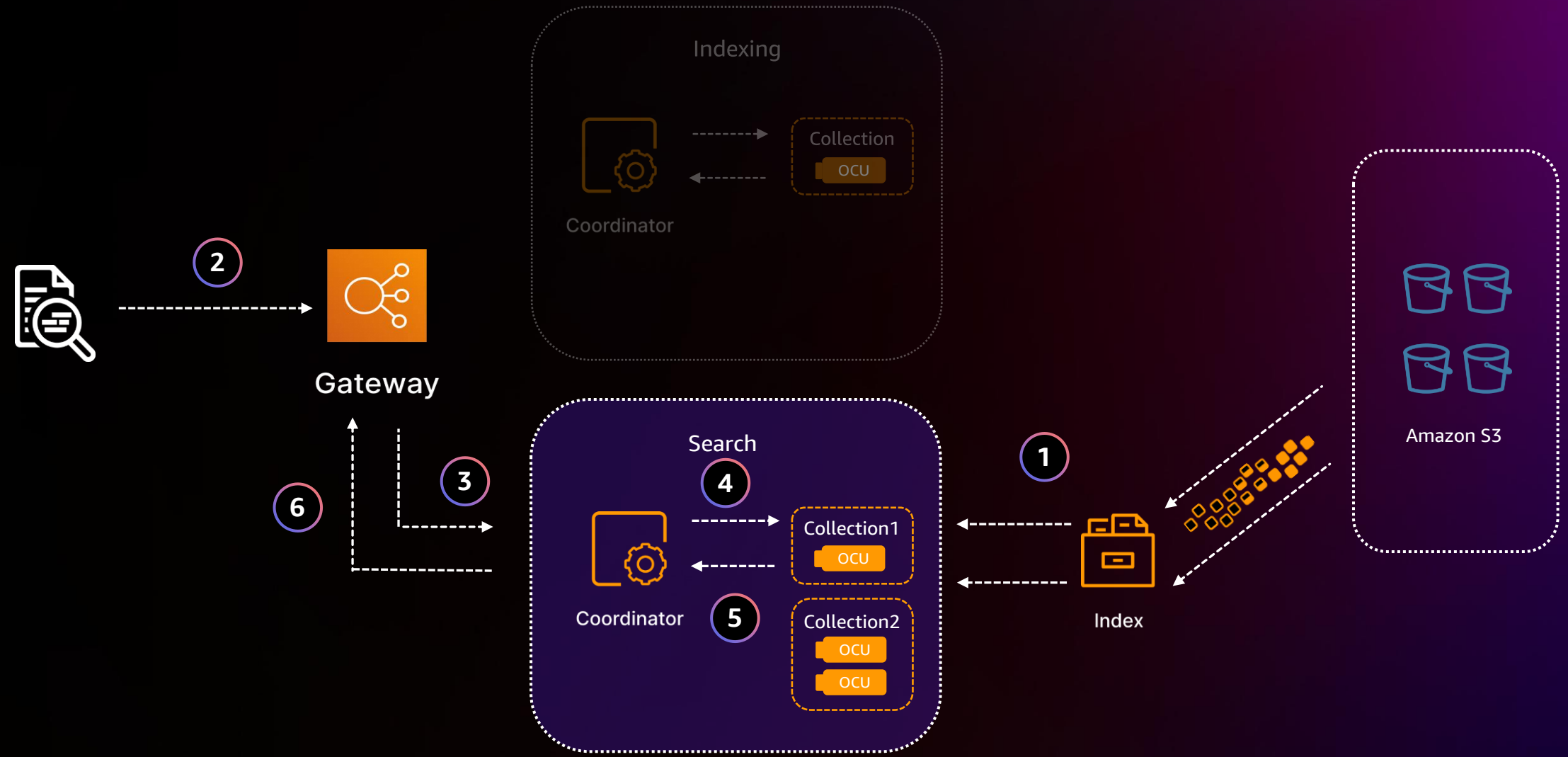
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Indexing



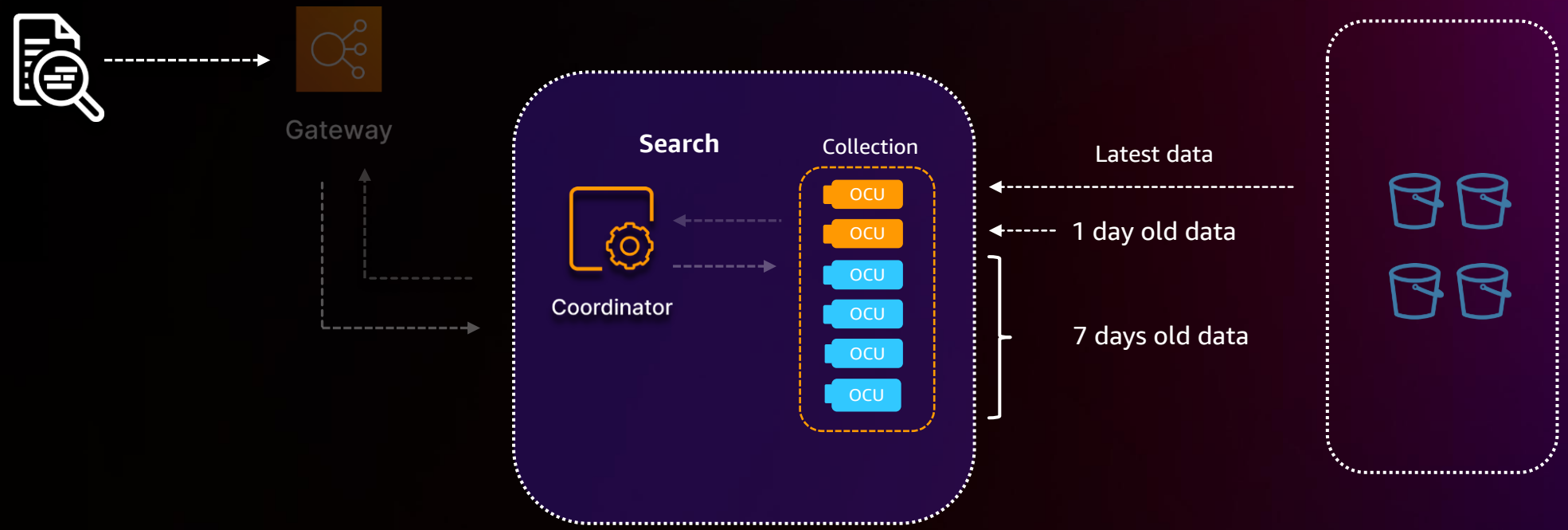
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Search



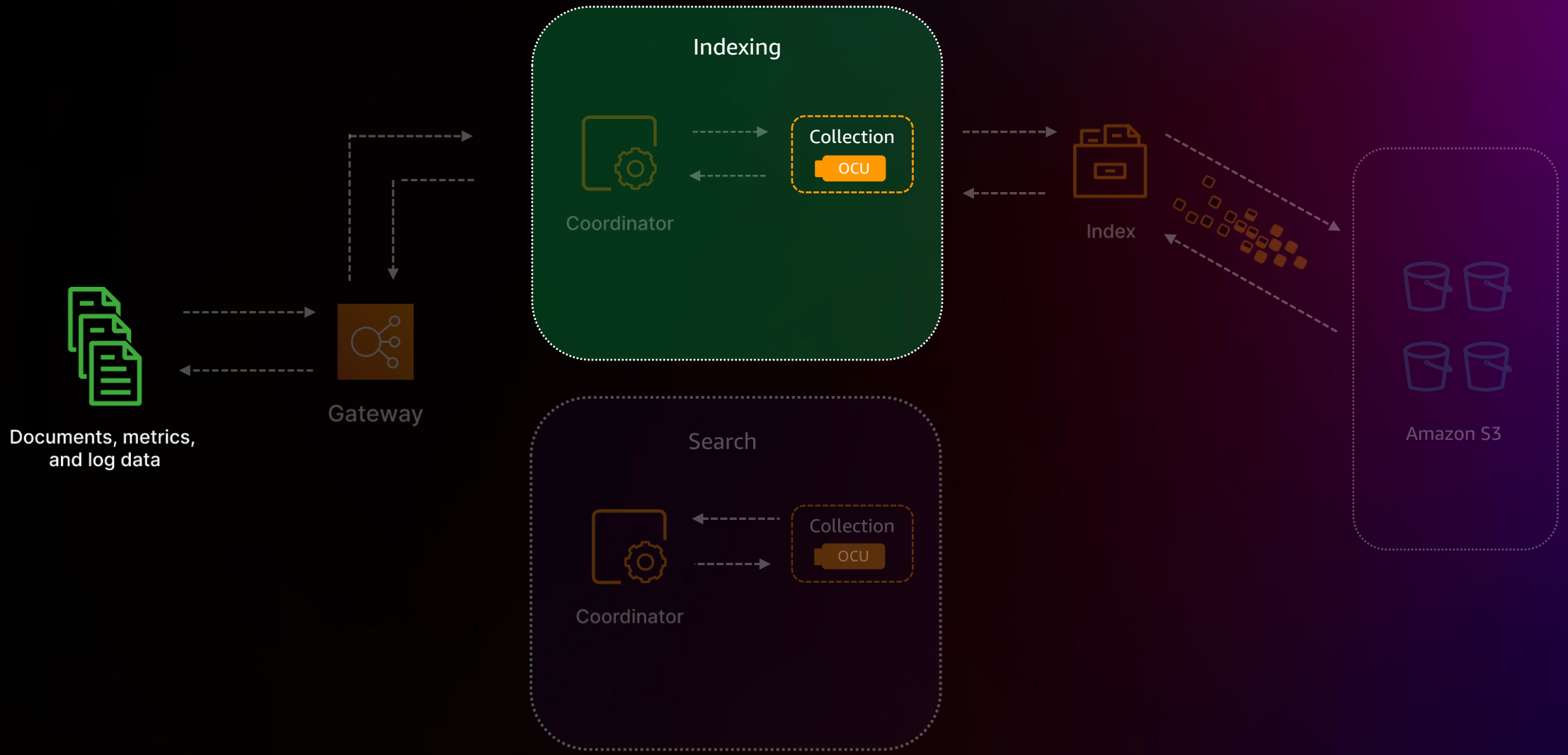
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hot-warm architecture



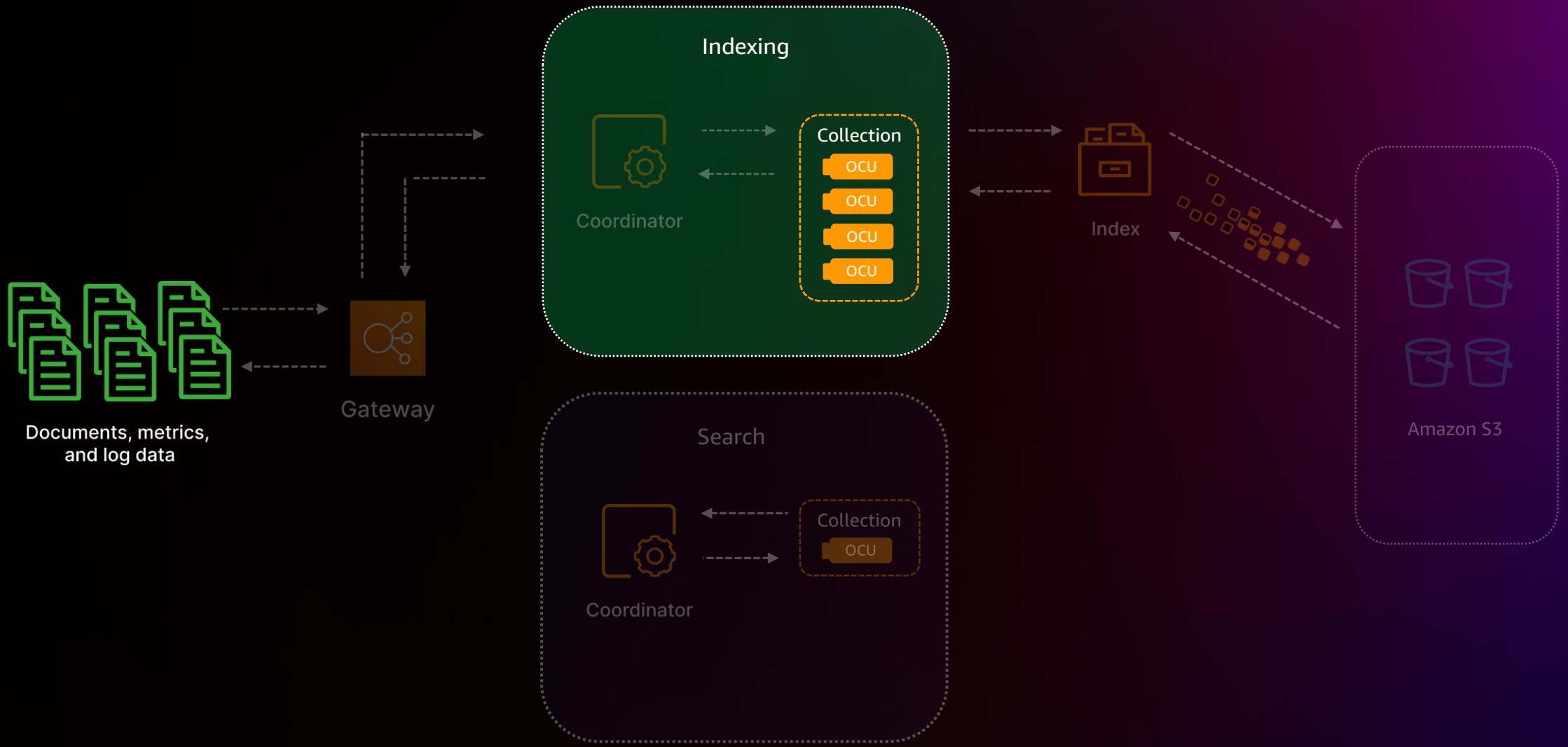
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Indexing auto-scaling



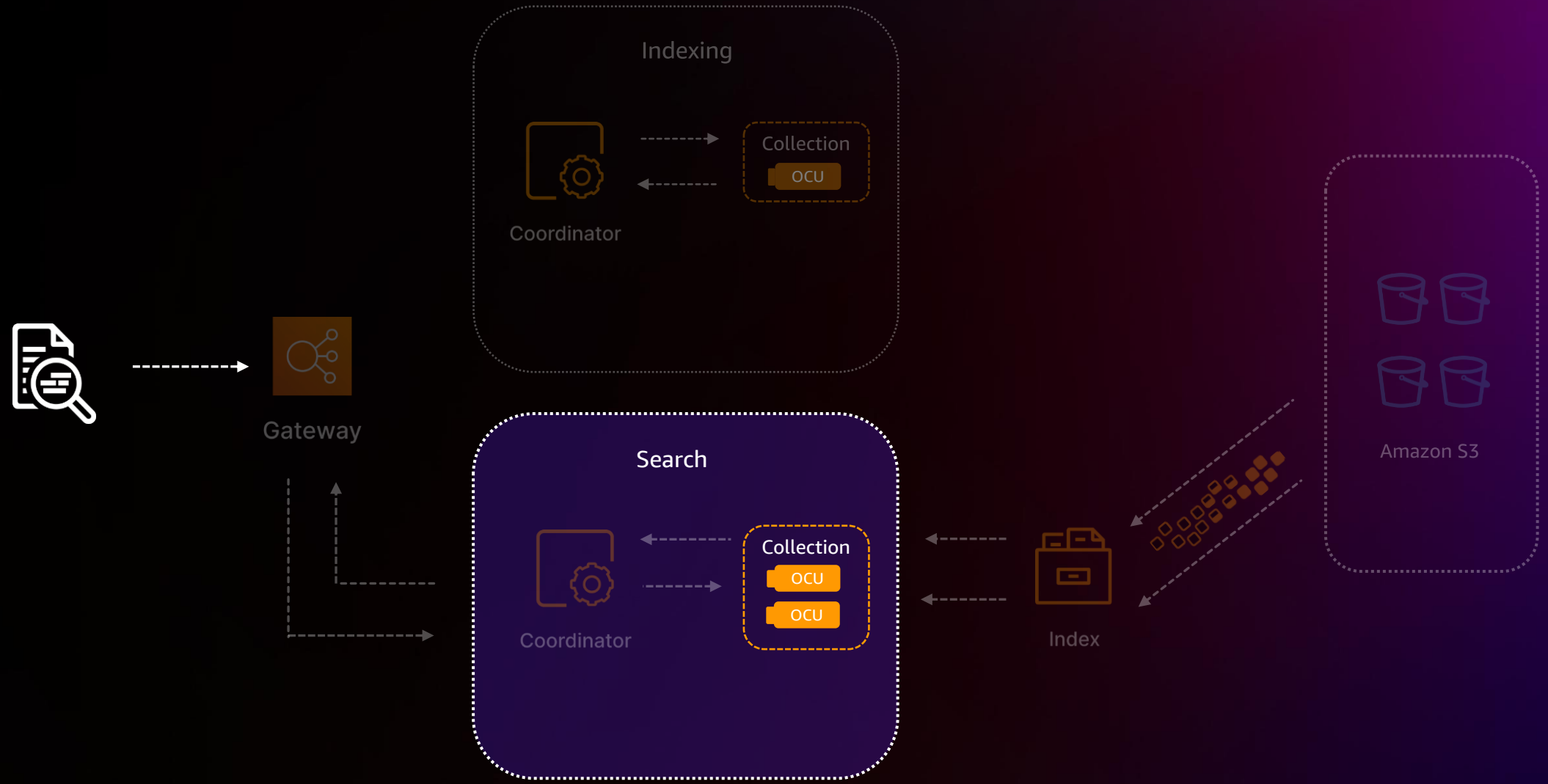
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Indexing auto-scaling



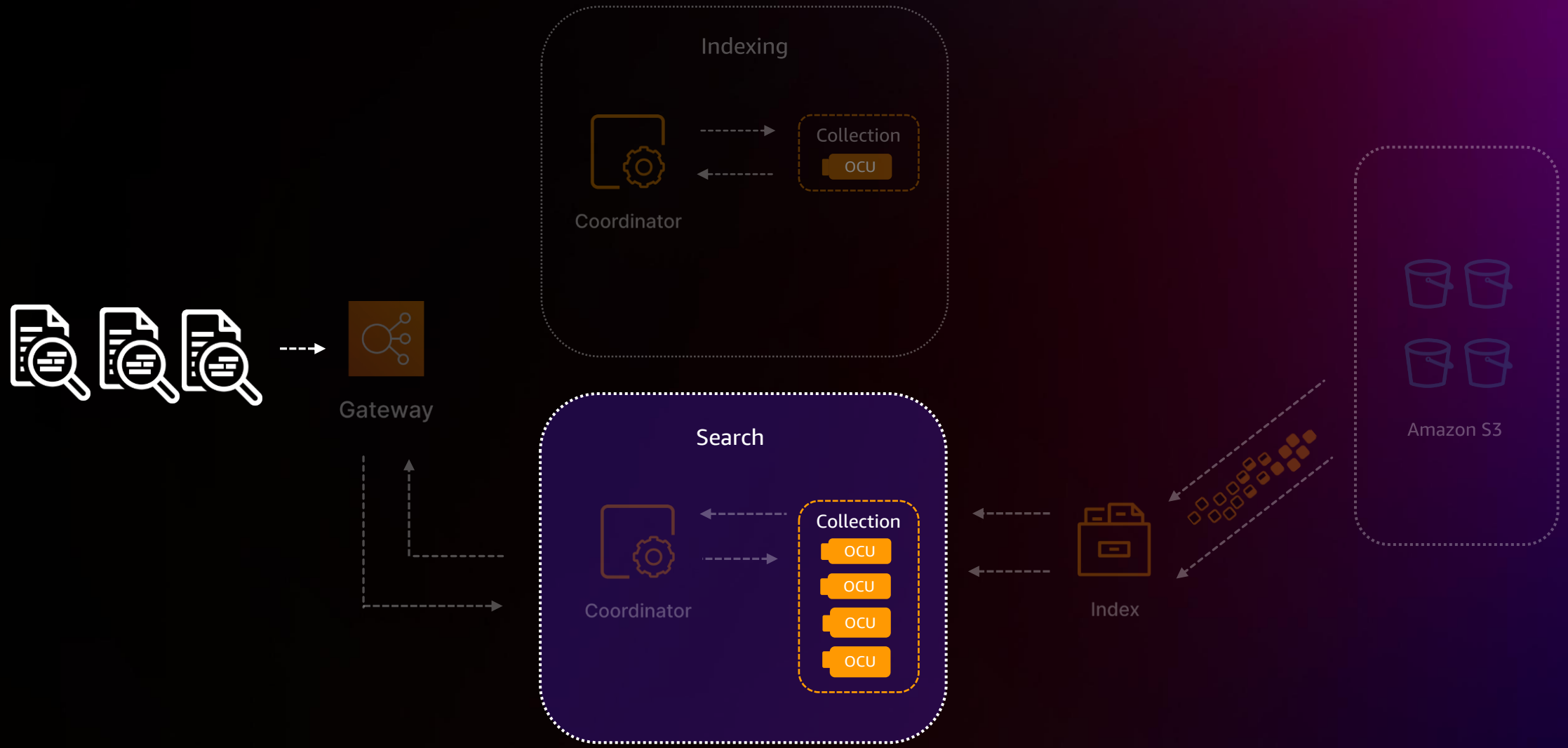
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Search auto-scaling



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Search auto-scaling



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Genpact use cases with OpenSearch Serverless

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Problem: AI/ML != AML

Transparency preferred over black box approach

Fraud

- Bank LOSES money
- Clear feedback loop to train models
- Regulatory detection criteria is less prescriptive
- Heavy use of AI/ML – build the best mouse trap

VS

Anti-money laundering (AML)

- Bank MAKES money
- No feedback loop
- Regulators make the rules
- Unsupervised ML possible but not required

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Genpact: Who are we?



1st big-data analytic system on government networks



Ingest **thousands** of data feeds and **millions** of records (90%+ unstructured)



NLU processing to extract features and enable **SUB-SECOND** query



Founded by Members
US DoD, Intel, OFAC, and FBI



TENS of THOUSANDS develop ENTITY-CENTRIC intelligence in defense of our country and allies

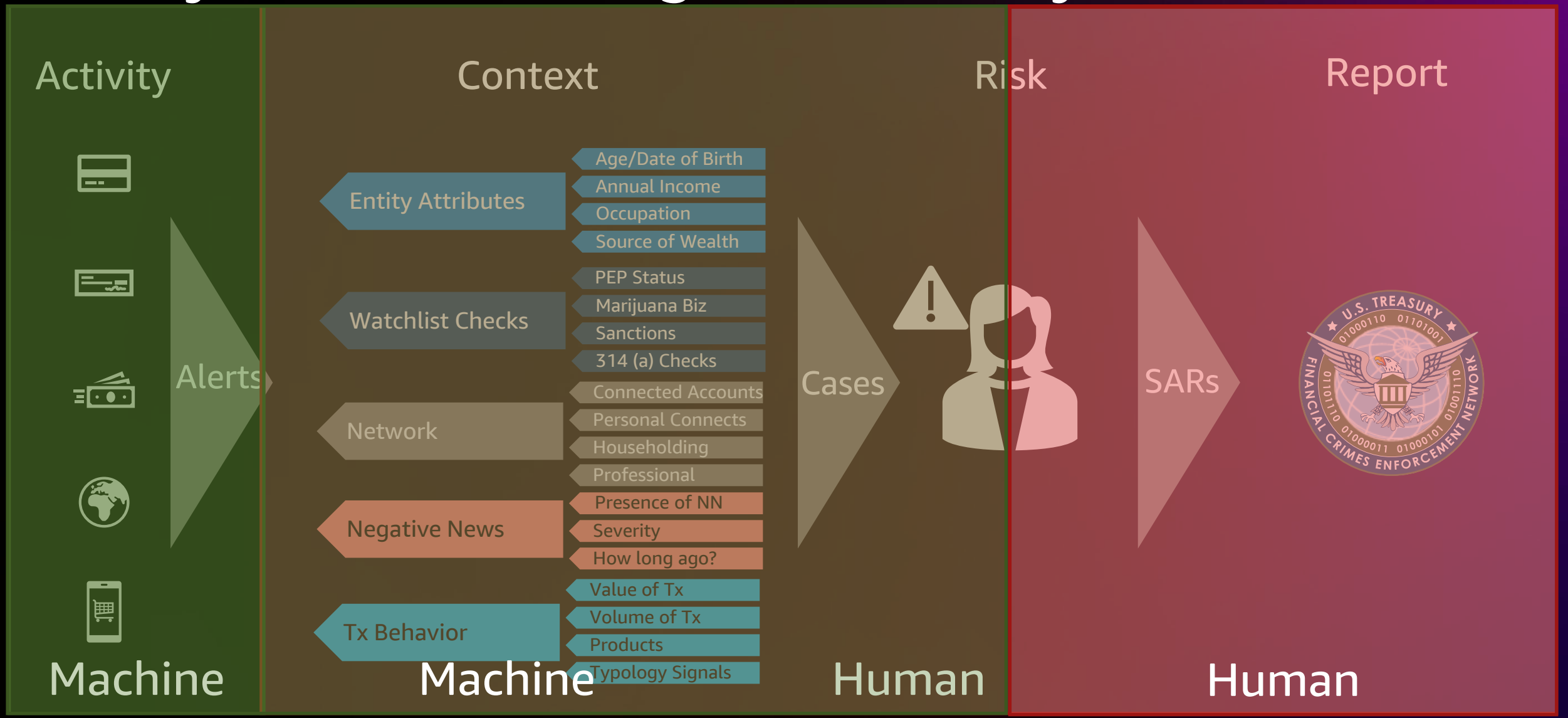


Analyzed **BILLIONS** of features to assemble entities

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Entity-based investigations today



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

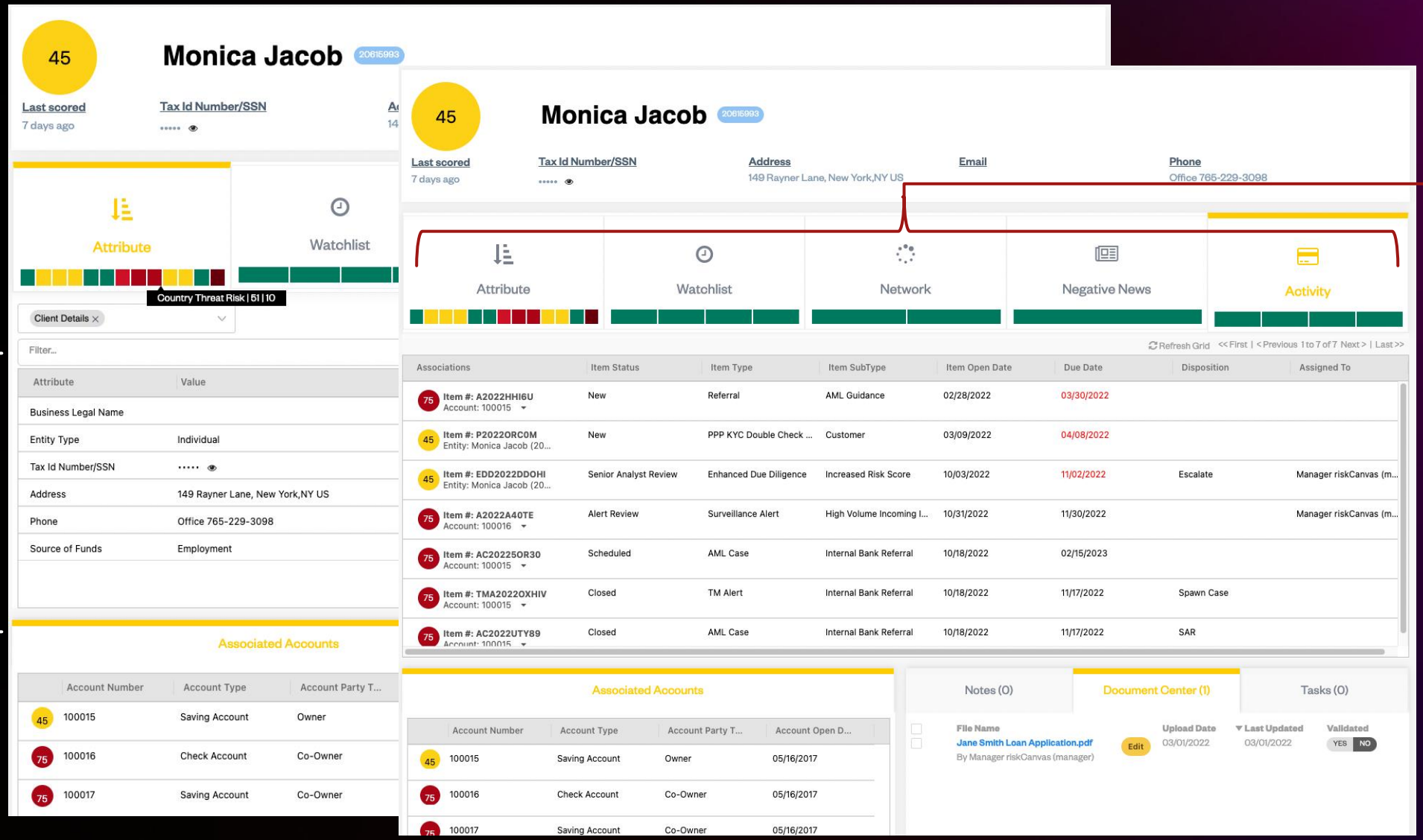
riskCanvas Entity 360°

Customer key info

Customer details

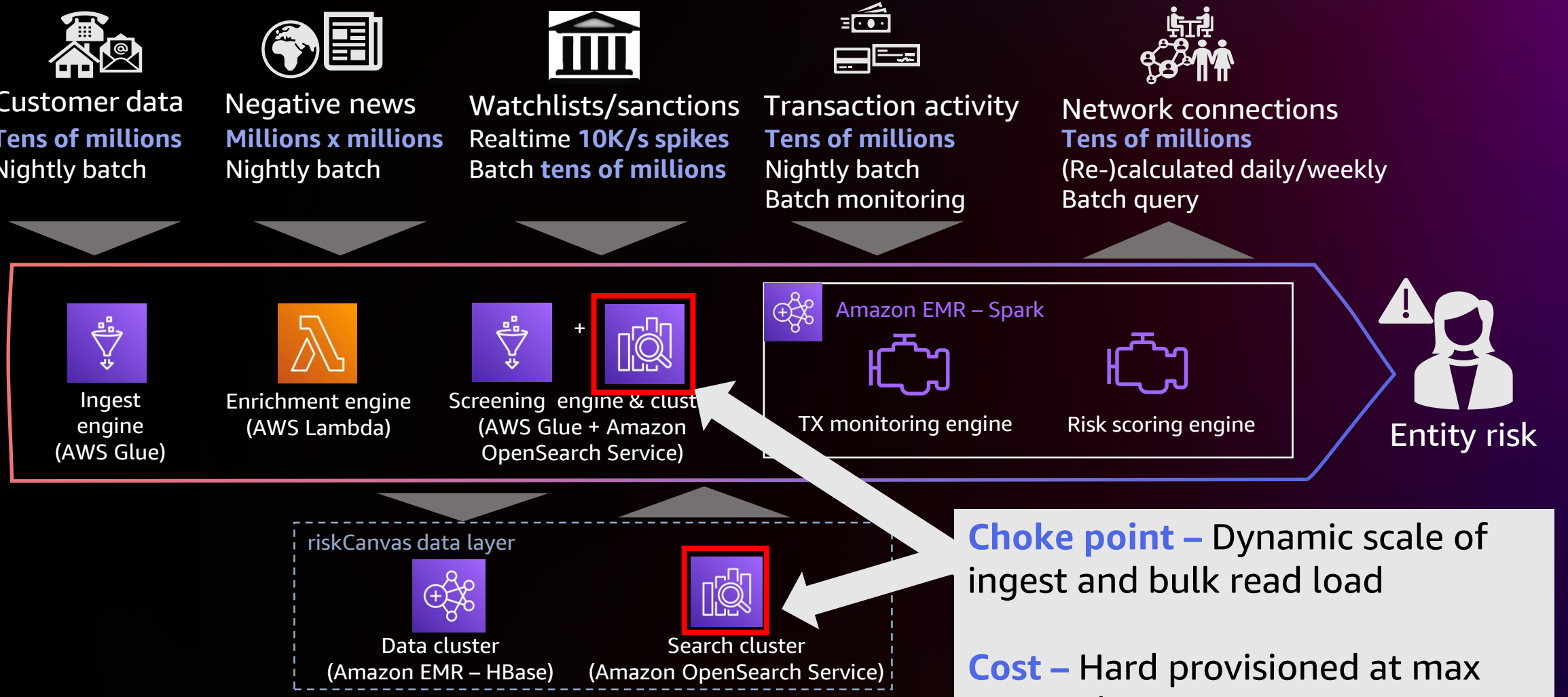
Customer "riskDNA"

Customer/ transaction activity



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

riskCanvas ECM Compute

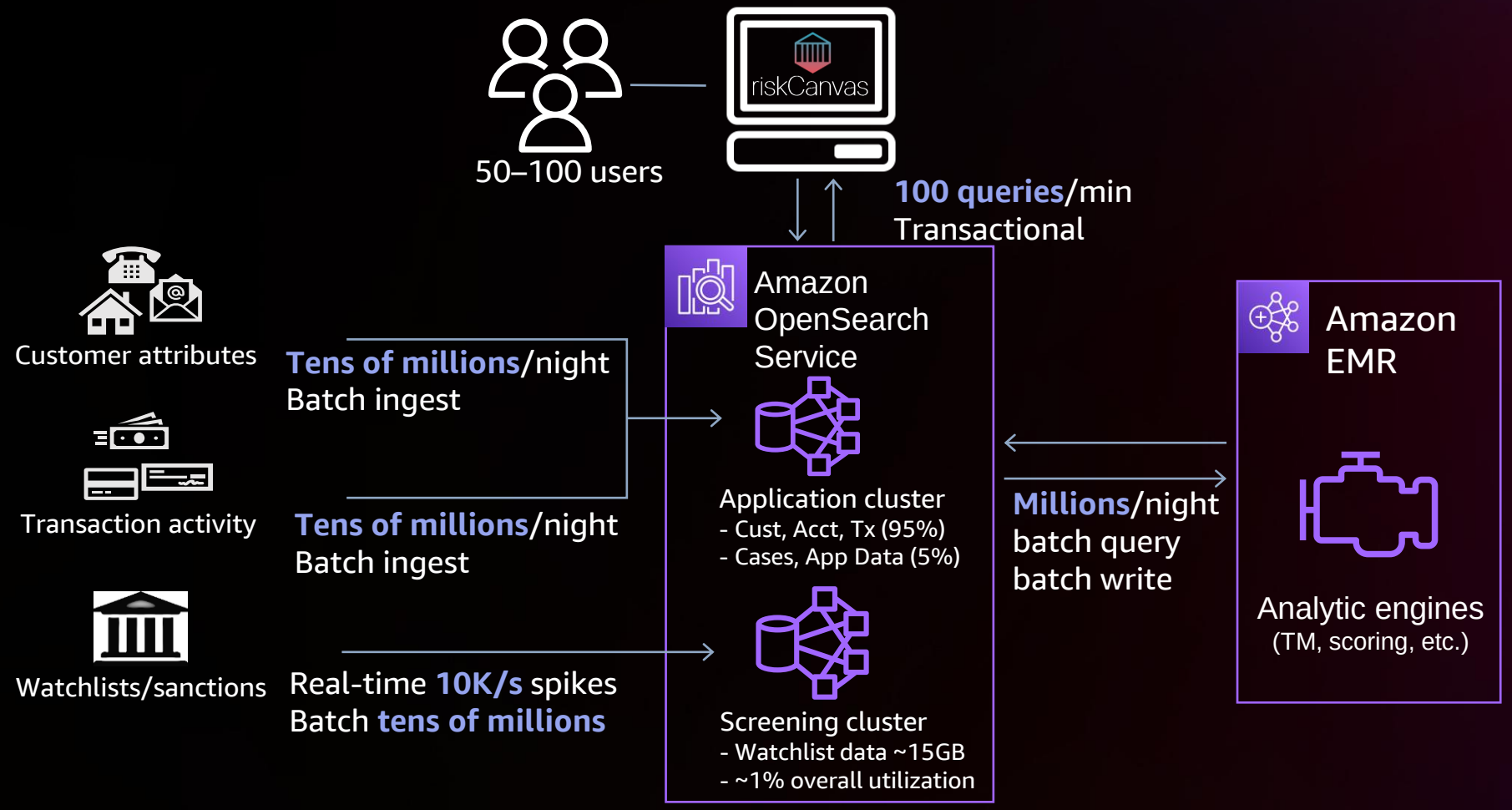


Choke point – Dynamic scale of ingest and bulk read load

Cost – Hard provisioned at max required capacity

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

riskCanvas architecture – Current



Challenges

Right sizing

Provisioned for peak, underutilized

Scaling

App, ingest, analysis

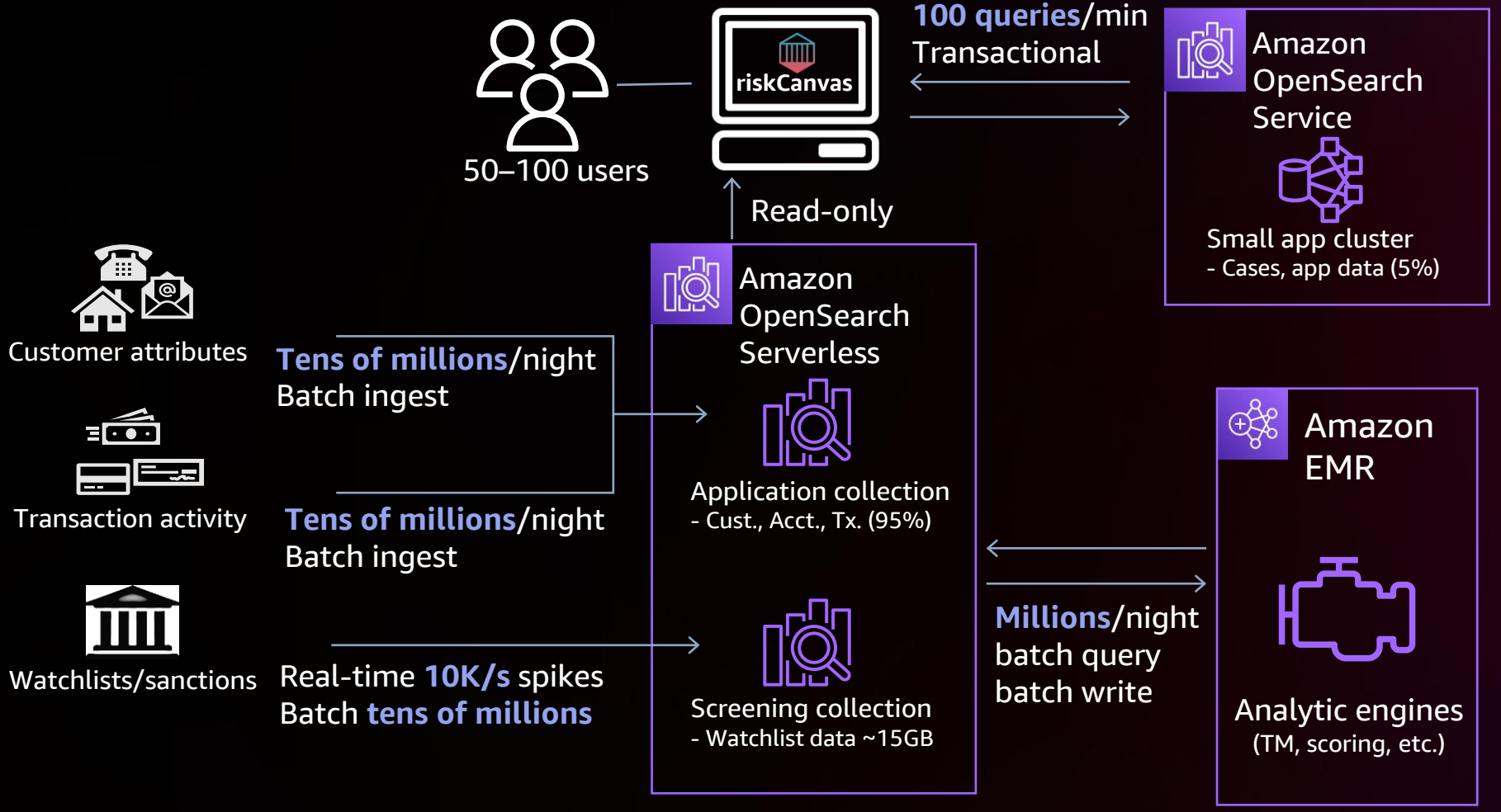
Responsiveness

Speed under load

Cost

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

riskCanvas architecture – OpenSearch Serverless



Opportunities

Right Sizing
Dynamically scale

Scale
Ingest and query independently

Responsive
Separation of loads

Cost
60-80% reduction and only pay for usage

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

OpenSearch Serverless testing – Integration

OpenSearch Serverless integration effort

1. Create a “Collection” – **5 mins** to spin up
2. Security config – **30 mins**
3. Test with Postman – **5 mins**
4. Client code (Java) – ~20 LOC to add a SIGv4 header to requests – **1 day**
5. Update config files → OpenSearch Serverless endpoint – **5 mins**
6. A bit of trial and error (C’mon, it’s beta) – 2 weeks . . .

Done!

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

OpenSearch Serverless testing – Coverage

What is working now?

- ☑ _bulk
- ☑ _cat (indices, etc.)
- ☑ Index/_search
- ☑ Index/_doc (PUT,GET)
- ☑ Mappings (text, dates, objects, hidden fields, analyzers) – NOT tested IP, Geo

What DOESN'T work?

- ✕ _cluster (huzzah!!!)
- ✕ _refresh (there's a trick)
- ✕ _scripts (named/stored scripts)
- ✕ _scroll
- ✕ Auto scaling (yet . . .)

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How is OpenSearch Serverless priced?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pricing

Managed Cluster

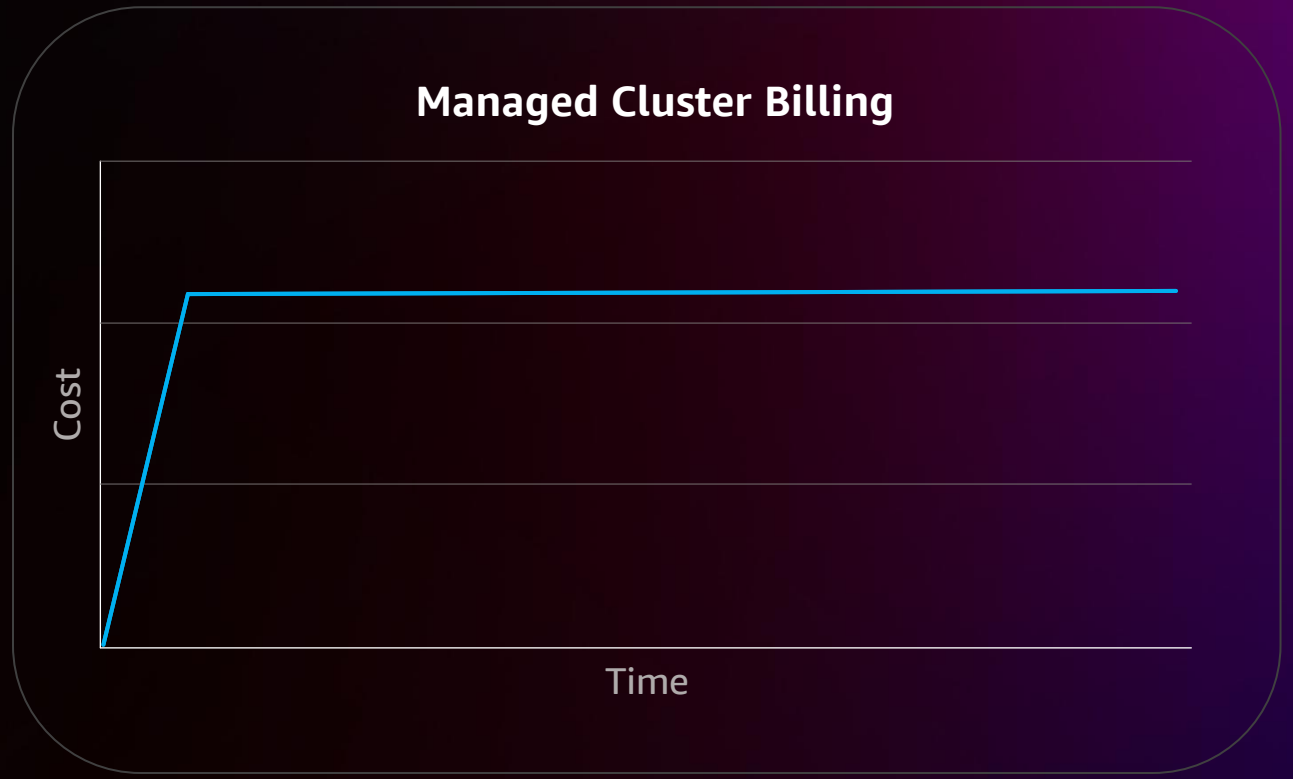
Configure your instances

+ Add instances for high availability

+ Add instances for UltraWarm

+ Add some buffer for peak workloads

Node instance



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pricing

Serverless

OCU

Storage



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pricing

Managed Cluster

Configure your instances

+ Add instances for high availability

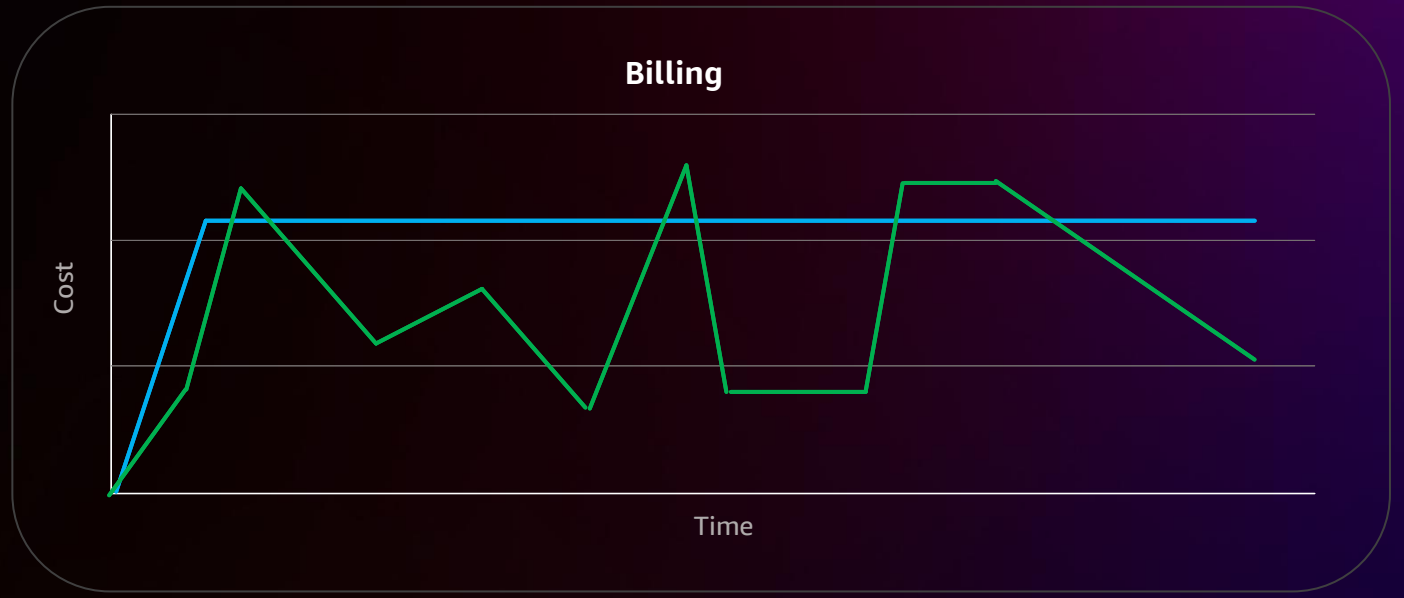
+ Add instances for UltraWarm

+ Add some buffer for peak workloads

Serverless

OCU

Storage



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon OpenSearch Serverless

Search and analyze data in seconds

Fast data ingestion rates

Millisecond query response times

Pay only for the resources used



Secure and compliant

Compute provisioning

Automatic scaling

High durability

Built-in resiliency

Automated patch management

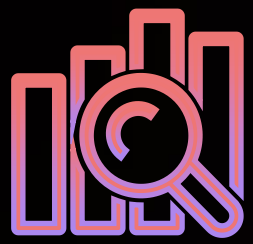
Advanced Monitoring

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Flexibility and choice in search and analytics

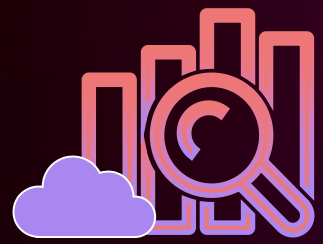
Managed Clusters



Full control on compute and storage size

Enables custom cluster settings and routing

Serverless



Automatic provisioning, scaling, and tuning

Built in shard and index lifecycle management

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What to do next

Visit the OpenSearch Serverless page

Sign up for migration assessment and assistance

Read more customer success stories

Find an AWS Partner to help

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Carl Meadows
carlmead@amazon.com
Twitter @Carl_F_Meadows

Pavani Baddepudi
bpavani@amazon.com
Twitter @pavanibharat

Ryan Skousen
ryan.skousen@genpact.com



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.