



AWS
re:Invent

O P N 2 1 2

Analyze your log data with Open Distro for Elasticsearch

Alolita Sharma

Principal Technologist
Amazon Web Services

Lucas Winkelman

Software Engineer
Amazon Web Services

Agenda

What is Elasticsearch?

Introducing Open Distro for Elasticsearch

Open Distro features

Features in development

Deployment options

Get involved in the Open Distro project community

What is Elasticsearch?

Sometimes referred to as the “ELK Stack”—Elasticsearch, Logstash, and Kibana

Distributed search and analytics engine build on Apache Lucene

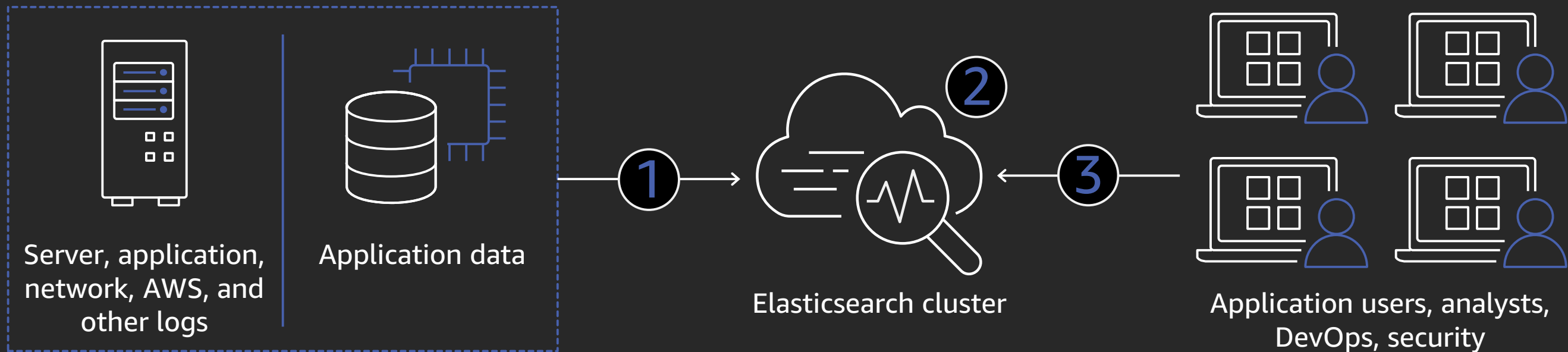
Easy ingestion and visualization

Rank			DBMS	Score		
Nov 2019	Oct 2019	Nov 2018		Nov 2019	Oct 2019	Nov 2018
1.	1.	1.	Oracle +	1336.07	-19.81	+34.96
2.	2.	2.	MySQL +	1266.28	-16.78	+106.39
3.	3.	3.	Microsoft SQL Server +	1081.91	-12.81	+30.36
4.	4.	4.	PostgreSQL +	491.07	+7.16	+50.83
5.	5.	5.	MongoDB +	413.18	+1.09	+43.70
6.	6.	6.	IBM Db2 +	172.60	+1.83	-7.27
7.	7.	↑ 8.	Elasticsearch +	148.40	-1.77	+4.94
8.	8.	↓ 7.	Redis +	145.24	+2.32	+1.06
9.	9.	9.	Microsoft Access	130.07	-1.10	-8.36
10.	10.	↑ 11.	Cassandra +	123.23	+0.01	+1.48

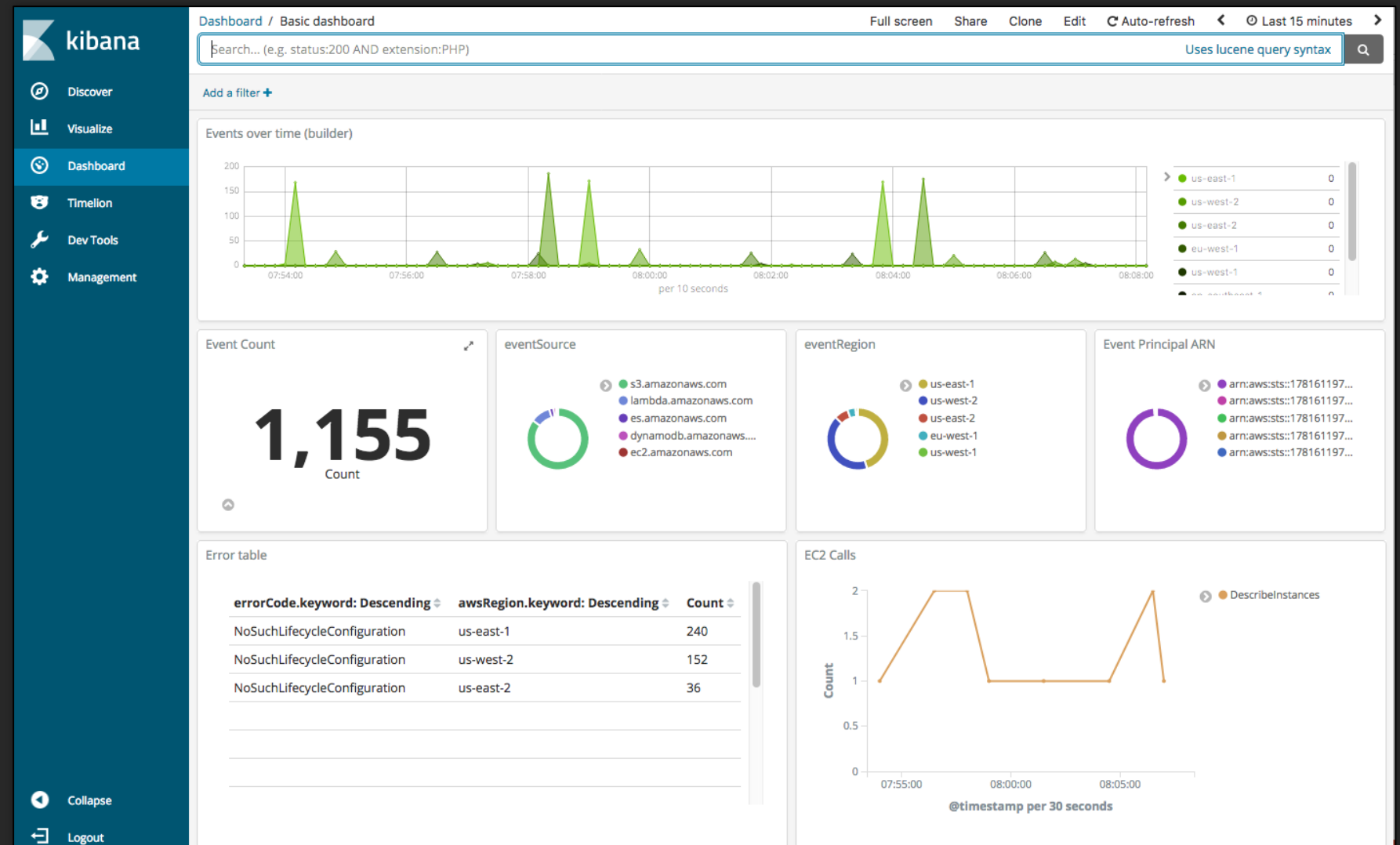
Source: db-engines.com Nov 2019

It is a database

- 1 Send data as JSON via REST APIs
- 2 Data is indexed—all fields searchable, including nested JSON
- 3 Queries, via REST APIs, allow fielded matching, Boolean expressions, include sorting and analysis



Kibana is a
lightweight,
real-time
visualization tool



Machine data driving Elasticsearch growth

Machine-generated data is growing **10x faster** vs. business data... **logs, logs, and more logs**



IT and DevOps:
databases, servers,
storage, networking



**Increase in IoT and
mobile devices: gaming,
sensors, web content**



**Cloud-based
architectures**



Open Distro for Elasticsearch

An Apache 2-licensed distribution of Elasticsearch enhanced with enterprise-grade security, alerting, SQL, performance analyzer, and more

Open Distro for Elasticsearch—features



Security

Achieve encryption in-flight, role-based access control, audit logging, and compliance



Alerting

Monitor your data and send automatic alerts on any changes in your data



SQL

Easily interact with your Elasticsearch cluster and extract insights using the familiar SQL query syntax



Performance Analyzer

Get deep visibility into system bottlenecks even when your Elasticsearch cluster is under duress

Security

KEEP YOUR DATA SECURE

Encryption

Keep your data secure at rest and in transit

Authentication

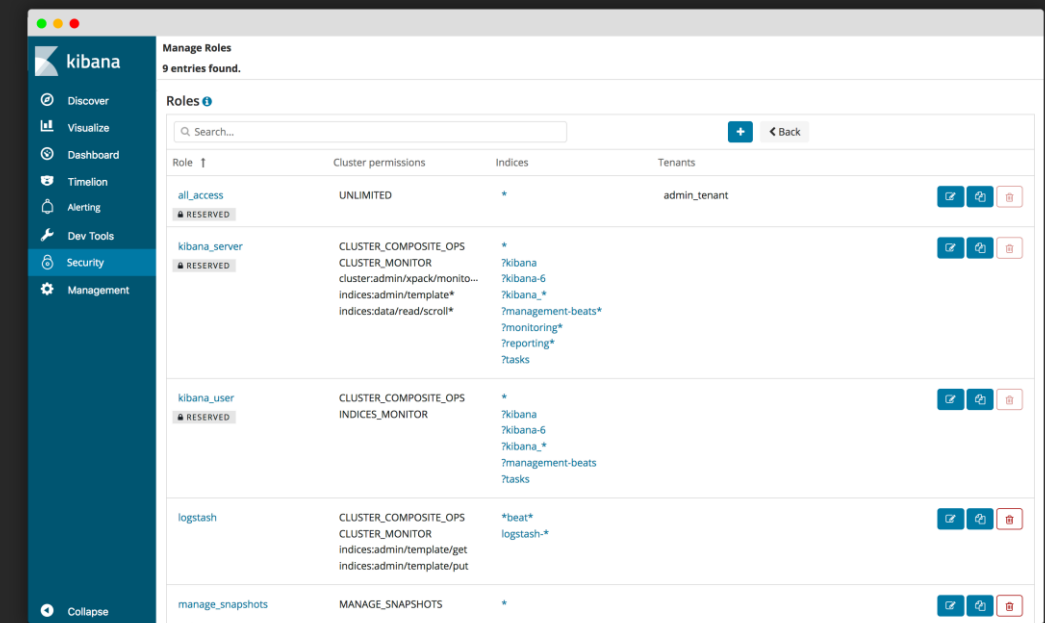
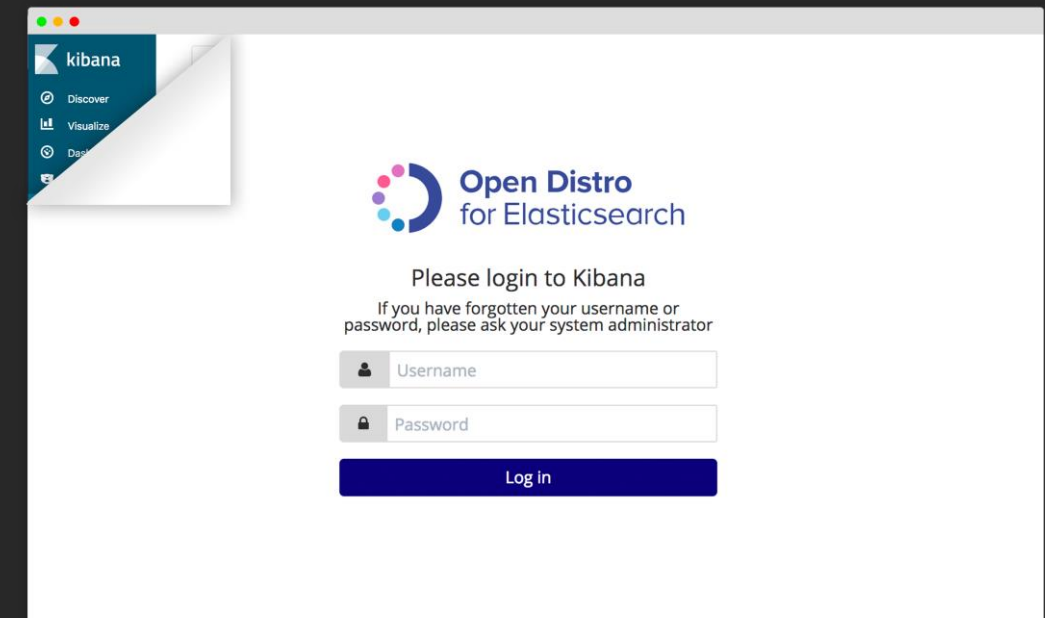
Leverage your existing authentication infrastructure

RBAC

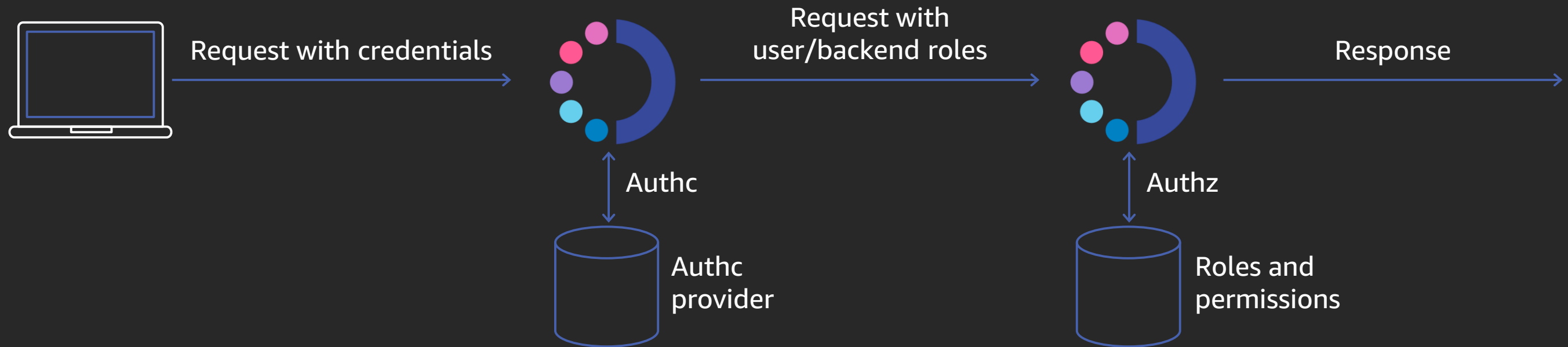
Granular access control to control the user actions on your cluster

Audit logging

Track and record all user actions and meet HIPAA, PCI compliance



Security access control flow



Authc — via basic HTTP auth, LDAP, AD, SAML, web tokens, SSL

Authz — backend identities mapped to Open Distro roles

Permissions — allow a role to perform an action against a cluster/index/document/field

Action groups — groups of permissions

Security audit logs

Cluster access by authenticated user

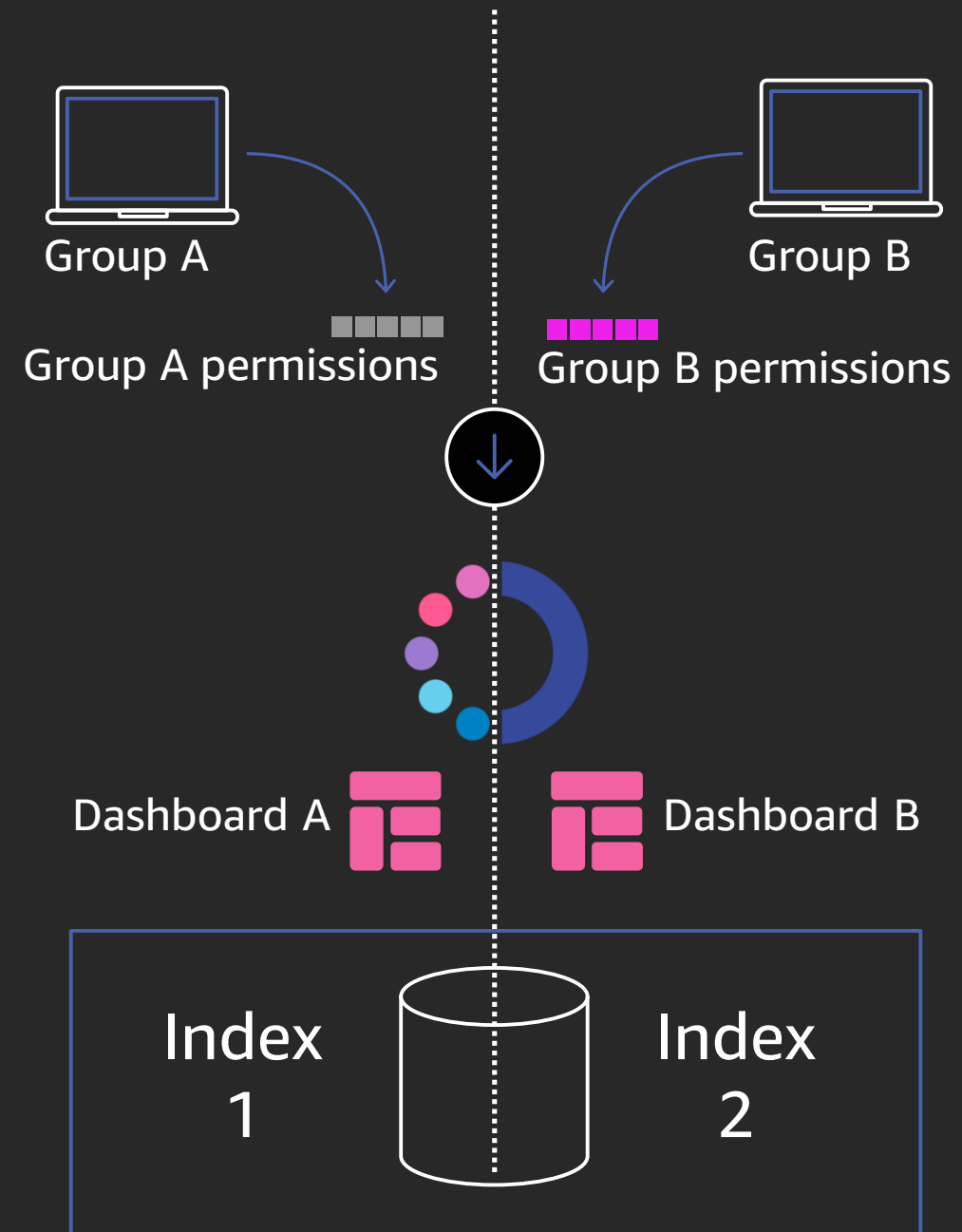
Request path—expose who did what

Combine with the Alerting plugin for security notifications

```
{
  "_index" : "security-auditlog-2019.05.03",
  "_type" : "auditlog",
  "_id" : "JUZQf2oB36QWxanFlTXp",
  "_score" : 1.0,
  "_source" : {
    "audit_cluster_name" : "odfe-cluster",
    "audit_node_name" : "WBkxX9a",
    "audit_category" : "FAILED_LOGIN",
    "audit_request_origin" : "REST",
    "audit_node_id" : "WBkxX9a5SnK1XmiHQW7M7w",
    "audit_request_layer" : "REST",
    "audit_rest_request_path" : "/_opendistro/_security/authinfo",
    "@timestamp" : "2019-05-03T20:09:08.310+00:00",
    "audit_request_effective_user_is_admin" : false,
    "audit_format_version" : 3,
    "audit_utc_timestamp" : "2019-05-03T20:09:08.310+00:00",
    "audit_request_remote_address" : "172.18.0.2",
    "audit_node_host_address" : "172.18.0.4",
    "audit_rest_request_headers" : {
      "Connection" : [
        "keep-alive"
      ],
      "Host" : [
        "odfe-node1:9200"
      ],
      "Content-Length" : [
        "0"
      ]
    },
    "audit_request_effective_user" : "<NONE>",
    "audit_node_host_name" : "172.18.0.4"
  }
}
```

Kibana multi-tenancy

Enabled out of the box



Alerting

RECEIVE ALERTS ON YOUR DATA

Create monitors

Query the data you want to
and receive alerts on it

Customize alert conditions

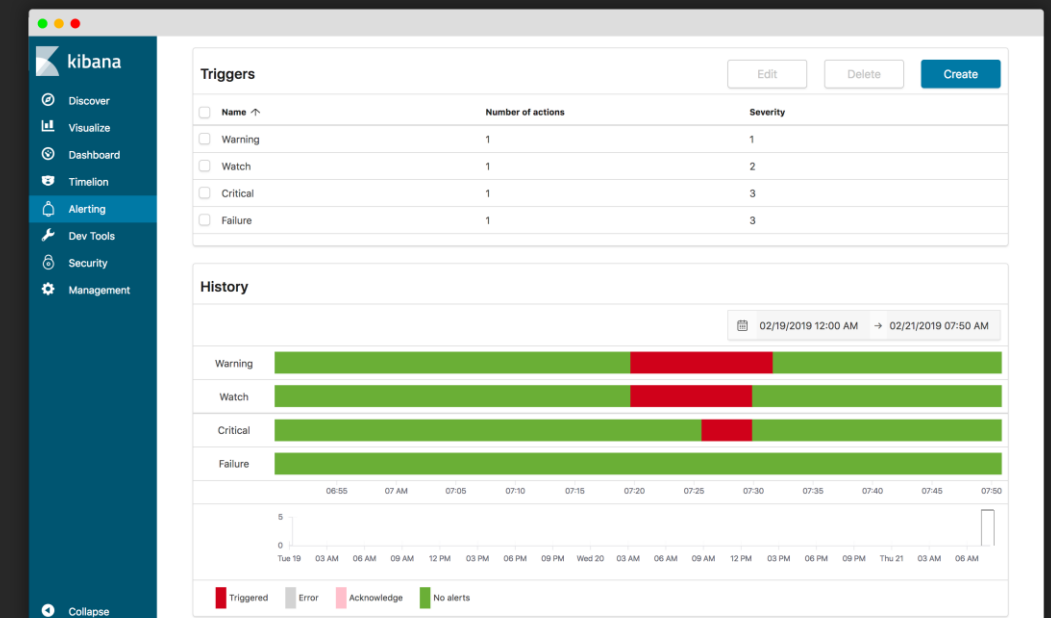
Define alerting threshold and severity
for multiple trigger conditions

Get notifications

Built-in integrations for webhook and Slack
to get notified on the channels you use

View alerts

All alert executions are indexed
for easy tracking and visualization



SQL support

QUERY DATA WITH SQL

Comprehensive SQL support

Supports over 40 functions, data types, and commands including join support

Translate SQL to JSON

Create JSON using SQL to configure sophisticated access control policies

Use existing tools

Provides a JDBC driver so you can use a variety of business intelligence, analytics, and ETL tools

1	# Get all accounts with JSON document response	index	id	score	account_number	balance	firstname	lastname	age	gender
2	GET _sql	accounts	25	1	25	\$ 40,540.00	Virginia	Ayala	39	F
3	{	accounts	44	1	44	\$ 34,487.00	Aurelia	Harding	37	M
4	"query": "SELECT * FROM accounts"	accounts	99	1	99	\$ 47,159.00	Ratliff	Heath	39	F
5	}	accounts	119	1	119	\$ 49,222.00	Laverne	Johnson	28	F
6		accounts	126	1	126	\$ 3,607.00	Effie	Gates	39	F
7	# Get all accounts with CSV response	accounts	145	1	145	\$ 47,406.00	Rowena	Wilkinson	32	M
8	GET _sql?format=csv	accounts	183	1	183	\$ 14,223.00	Hudson	English	26	F
9	{	accounts	190	1	190	\$ 3,150.00	Blake	Davidson	30	F
10	"query": "SELECT * FROM accounts"	accounts	208	1	208	\$ 40,760.00	Garcia	Hess	26	F
11	}	accounts	222	1	222	\$ 14,764.00	Rachelle	Rice	36	M
12		accounts	227	1	227	\$ 19,780.00	Coleman	Berg	22	M
13	# Get average age of employees	accounts	253	1	253	\$ 20,240.00	Melissa	Gould	31	M
14	GET _sql?format=csv	accounts	260	1	260	\$ 2,726.00	Kari	Skinner	30	F
15	{	accounts	265	1	265	\$ 46,910.00	Marion	Schneider	26	F
16	"query": "SELECT AVG(age) as avg, employer, state, city FROM accounts GROUP BY employer .keyword, state.keyword, city.keyword"	accounts	335	1	335	\$ 35,433.00	Vera	Hansen	24	M
17	}	accounts	366	1	366	\$ 42,368.00	Lydia	Cooke	31	M
18		accounts	385	1	385	\$ 11,022.00	Rosalinda	Valencia	22	M
19	# Compose SQL query to Elasticsearch query DSL	accounts	397	1	397	\$ 37,418.00	Leonard	Gray	36	F
20	GET _sql/_explain	accounts	400	1	400	\$ 20,685.00	Kane	King	21	F
21	{	accounts	450	1	450	\$ 2,643.00	Bradford	Nielsen	25	M
22	"query": "SELECT AVG(age) as avg, employer, state, city FROM accounts GROUP BY employer .keyword, state.keyword, city.keyword"	accounts	486	1	486	\$ 35,902.00	Dixie	Fuentes	22	F
23	}									

1	SELECT Avg(age) AS avg,	1	{
2	employer,	2	"from": 0,
3	state,	3	"size": 0,
4	city	4	"_source": {
5	FROM accounts	5	"includes": [
6	GROUP BY employer.keyword,	6	"AVG",
7	state.keyword,	7	"employer",
8	city.keyword	8	"state",
9		9	"city"
10		10],
~		11	"excludes": []
~		12	},
~		13	"stored_fields": [
~		14	"employer",
~		15	"state",
~		16	"city"
~		17],
~		18	"aggregations": {
~		19	"employer.keyword": {
~		20	"terms": {
~		21	"field": "employer.keyword",
~		22	"size": 200,
~		23	"min_doc_count": 1,

Performance Analyzer

GET DEEP DIAGNOSTIC INSIGHTS INTO YOUR CLUSTER

Identify bottlenecks across the stack

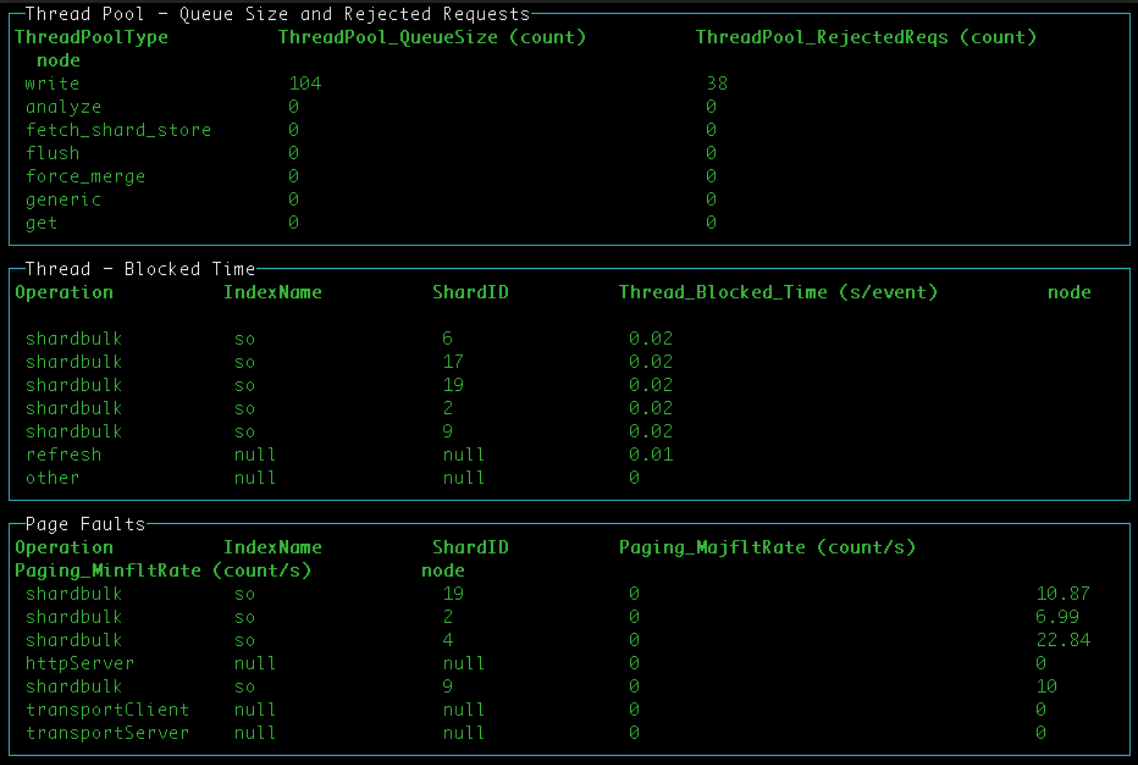
Provides a powerful REST API for querying Elasticsearch metrics to diagnose issues across stack

Runs independent of your cluster

Perform diagnostics even if the cluster is under duress

Analyze hundreds of data points

Supports over 60 metrics across 10 dimensions for instrumentation of your cluster health



PerfTop

PERFORMANCE ANALYZER CLIENT

Provides pre-configured dashboards for analyzing cluster, node, and shard performance

Custom JSON templates to create the dashboards to diagnose your cluster performance



Index Management

NEW! Manage your Elasticsearch Indices

Manage indices

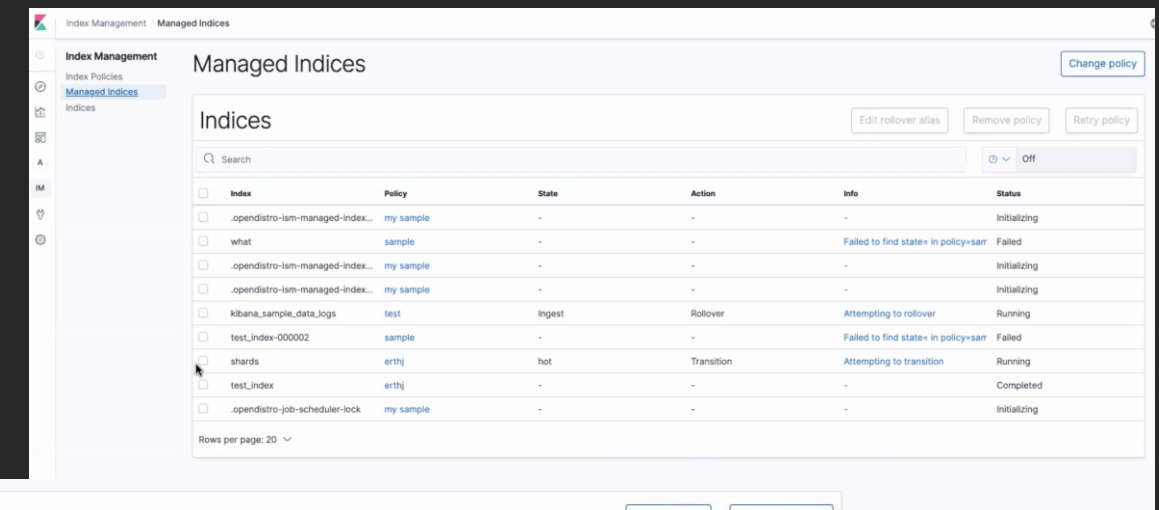
Manage your indices for Elasticsearch through Kibana

Define policies

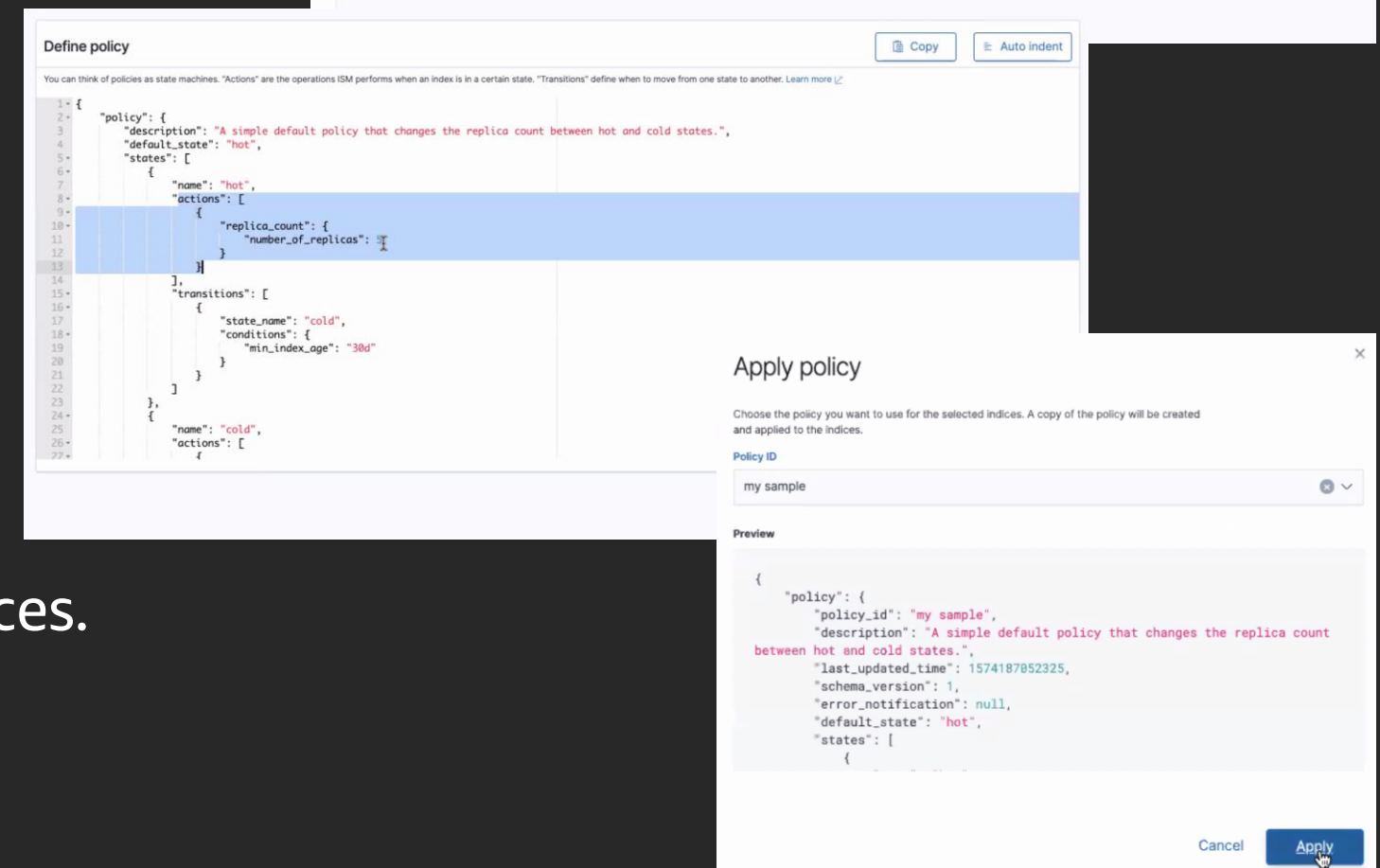
Create policies to manage your indices

Apply policies

Apply policies to monitor and manage your indices. Supported actions include open, read only, read write, replica count, rollover, force merge, close, delete and notification



Index	Policy	State	Action	Info	Status
.opendistro-ism-managed-index...	my sample	-	-	-	Initializing
what	sample	-	-	Failed to find state in policy: sam	Failed
.opendistro-ism-managed-index...	my sample	-	-	-	Initializing
.opendistro-ism-managed-index...	my sample	-	-	-	Initializing
kibana_sample_data_logs	test	Ingest	Rollover	Attempting to rollover	Running
test_index-000002	sample	-	-	Failed to find state in policy: sam	Failed
shards	erthj	hot	Transition	Attempting to transition	Running
test_index	erthj	-	-	-	Completed
.opendistro-job-scheduler-lock	my sample	-	-	-	Initializing



```
{
  "policy": {
    "description": "A simple default policy that changes the replica count between hot and cold states.",
    "default_state": "hot",
    "states": [
      {
        "name": "hot",
        "actions": [
          {
            "replica_count": {
              "number_of_replicas": 3
            }
          }
        ]
      },
      {
        "name": "cold",
        "actions": [
          {
            "replica_count": {
              "number_of_replicas": 1
            }
          }
        ]
      }
    ],
    "transitions": [
      {
        "state_name": "cold",
        "conditions": {
          "min_index_age": "30d"
        }
      }
    ]
  },
  "name": "hot",
  "actions": [
    {
      "replica_count": {
        "number_of_replicas": 3
      }
    }
  ]
}
```

Apply policy

Choose the policy you want to use for the selected indices. A copy of the policy will be created and applied to the indices.

Policy ID

my sample

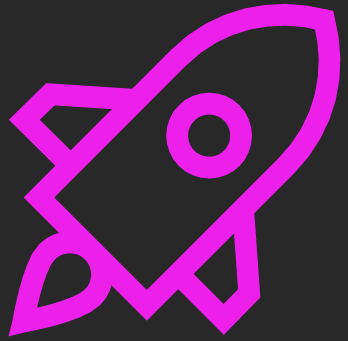
Preview

```
{
  "policy": {
    "policy_id": "my sample",
    "description": "A simple default policy that changes the replica count between hot and cold states.",
    "last_updated_time": 1574187052325,
    "schema_version": 1,
    "error_notification": null,
    "default_state": "hot",
    "states": [
      {
        "name": "hot",
        "actions": [
          {
            "replica_count": {
              "number_of_replicas": 3
            }
          }
        ]
      },
      {
        "name": "cold",
        "actions": [
          {
            "replica_count": {
              "number_of_replicas": 1
            }
          }
        ]
      }
    ],
    "transitions": [
      {
        "state_name": "cold",
        "conditions": {
          "min_index_age": "30d"
        }
      }
    ]
  },
  "name": "hot",
  "actions": [
    {
      "replica_count": {
        "number_of_replicas": 3
      }
    }
  ]
}
```

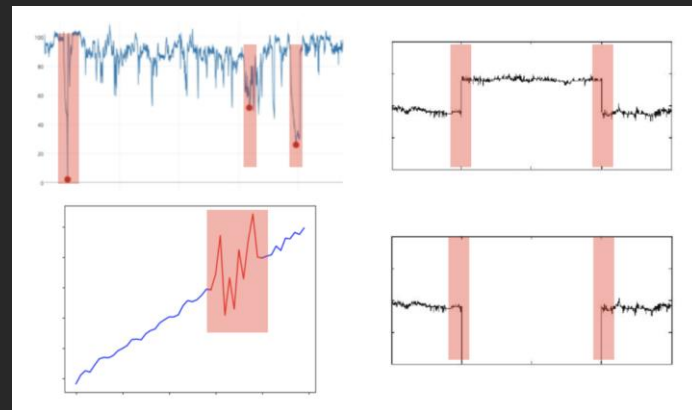
Cancel Apply

New features

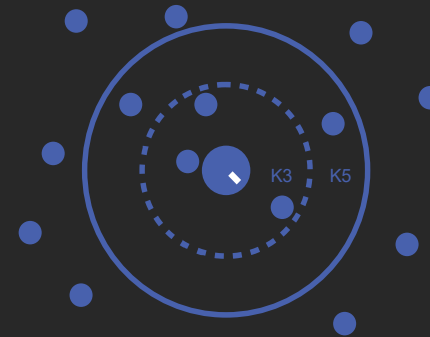
IN DEVELOPMENT



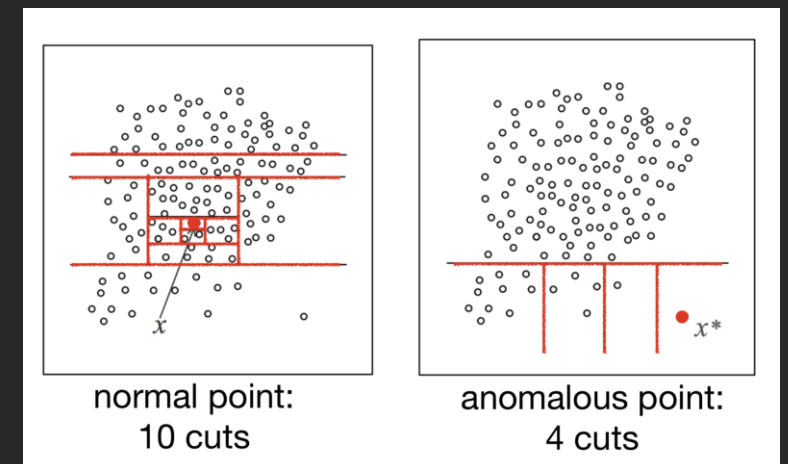
Performance
Root Cause
Analysis



Anomaly Detection



k-NN search



Machine learning
algorithms (RCF)

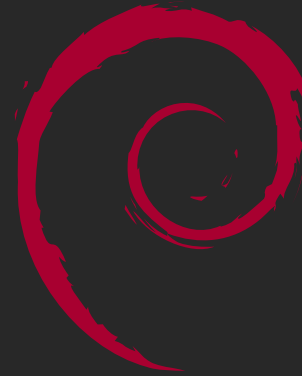
Flexible deployment options



Docker



RPM



Debian



Tarball

Community and Contributions

Open Distro for Elasticsearch's success is driven by the community's participation, contributions, and innovation to the project.

Join in for project discussions, share your knowledge with fellow community members, write a blog post, participate in Open Distro events and meetups, contribute documentation, file PRs, bugs or request a feature at:

Project Discussion Forum (Q&A): discuss.opendistrocommunity.dev

Source Code: github.com/opendistro-for-elasticsearch

Community on GitHub: github.com/opendistro-for-elasticsearch/community/issues

Project Website: opendistro.github.io

Project Blog: opendistro.github.io/blog

Related breakouts

OPN 212 Analyze your log data with Open Distro for Elasticsearch

OPN 204 Secure your Open Distro for Elasticsearch cluster

OPN 302-R, 302-R1 Get started with Open Distro for Elasticsearch

OPN 310-R, 310-R1 Alerting with Open Distro for Elasticsearch

OPN 311-R, 311-R1 Analyze Performance of your workload with Open Distro for Elasticsearch

ANT 346 Know your data with machine learning in Open Distro for Elasticsearch

Thank you!

Alolita Sharma

alolitas@amazon.com



Please complete the session
survey in the mobile app.