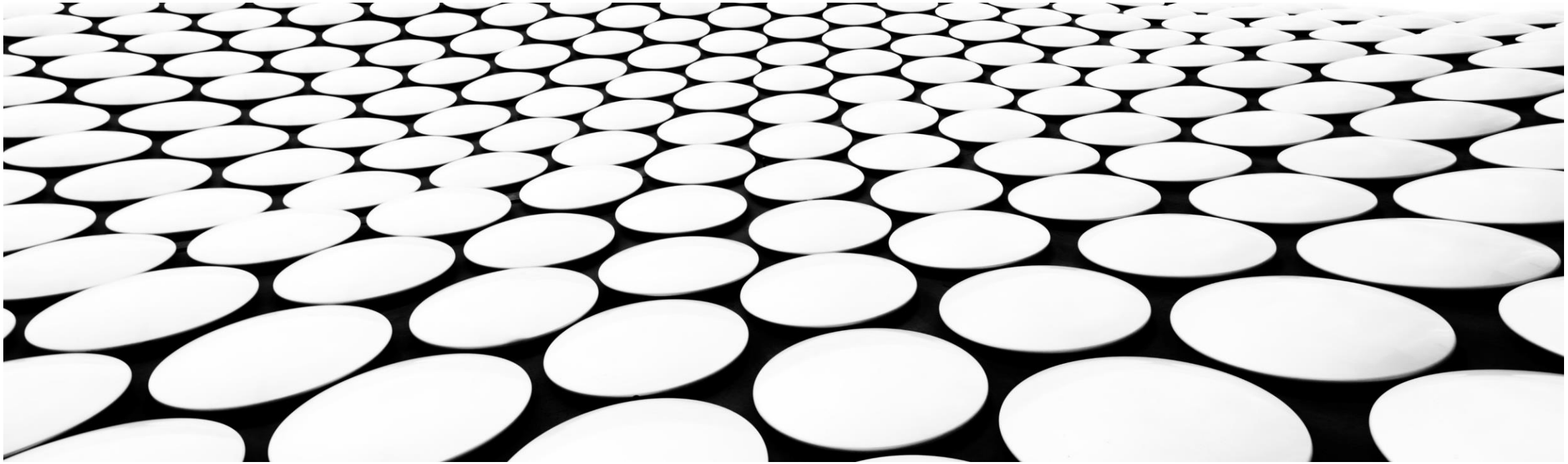
REAL TIME ANALYTICS

BIG DATA USE CASES : OBSERVABILITY & ENTERPRISE SEARCH



WHAT IS ELASTIC STACK & THE USE CASES?

LOOKING AT WHERE WE CAN USE REAL TIME MONITORING AND ML

ELASTIC USE CASES

- Enterprise Search 

Easily implement powerful, modern search experiences for your busy business teams.

- Observability 

Elastic platform offers capability to bring your logs, metrics, and APM traces together at scale in a single stack so your can monitor and react to events happening anywhere in your environments.

- Security 

Elastic Security features equips analysts to prevent, detect, and respond to threats. Elastic supports SIEM, endpoint security, threat hunting, cloud monitoring and more.

ELASTIC ENTERPRISE SEARCH

- Workplace Search 

Make all your teams' content findable, fast. Unify your content platforms (Google Drive, Slack, Salesforce and many more) into a personalized natural search experience

- App Search 

Leverage the focused power of elasticsearch in your app, complete with a web crawler, refined set of APIs and intuitive dashboards, tunable relevance controls, well maintained clients, and robust analytics

- Site Search 

Everything you need to add powerful search to your website.

Key Features to consider...

- Elasticsearch is incredibly fast, optimized relevance scoring models designed for real life, natural search.
- Get started quickly and index easily with the web crawlers, flexible APIs, dynamic schemas, and cloud source connectors
- Highly scalable platform. Elasticsearch can be horizontally scaled very easily.

ELASTIC OBSERVABILITY

- Logs

Elastic Stack is the most popular free and open logging platform. With out of the box support for common data sources and default dashboards to boot. It is capable to ship logs from kubernetes, MySQL, PostgreSQL, AWS and many more data sources. Index your data into elasticsearch and visualize it all in Kibana in minutes.

- Metrics

Using elasticsearch, you can easily monitor your infrastructure compute power and health in real-time.

- APM

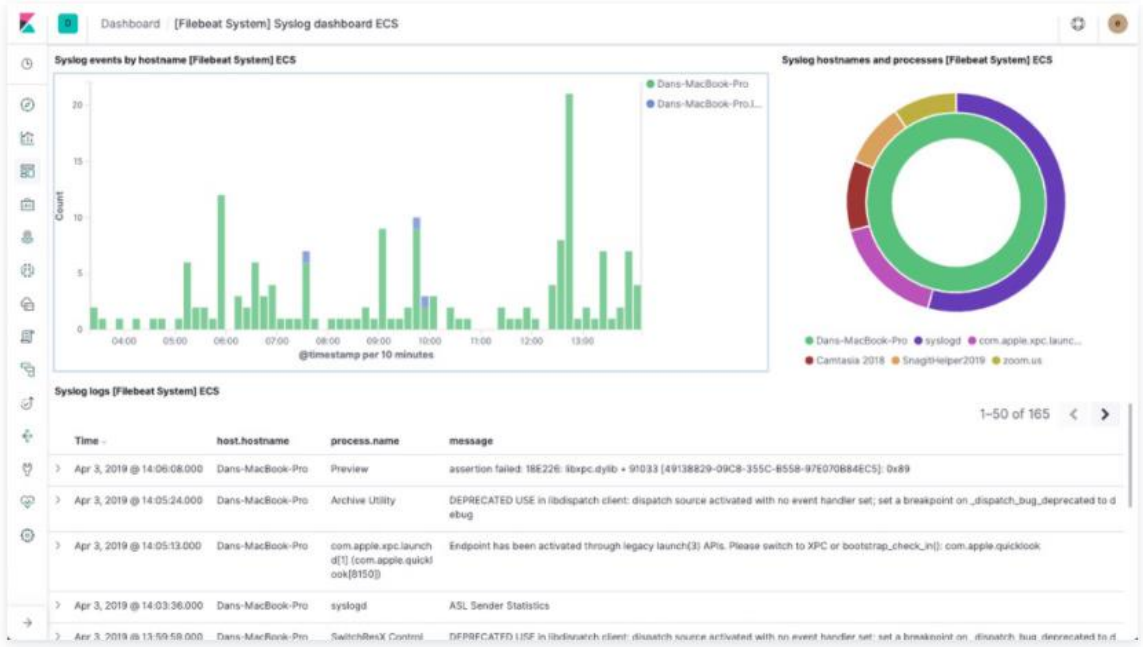
Easily diagnose root cause analysis of issues and performance problems using the Application Performance Monitoring module of elasticsearch.

- Uptime

ELASTIC OBSERVABILITY WITH LOGS

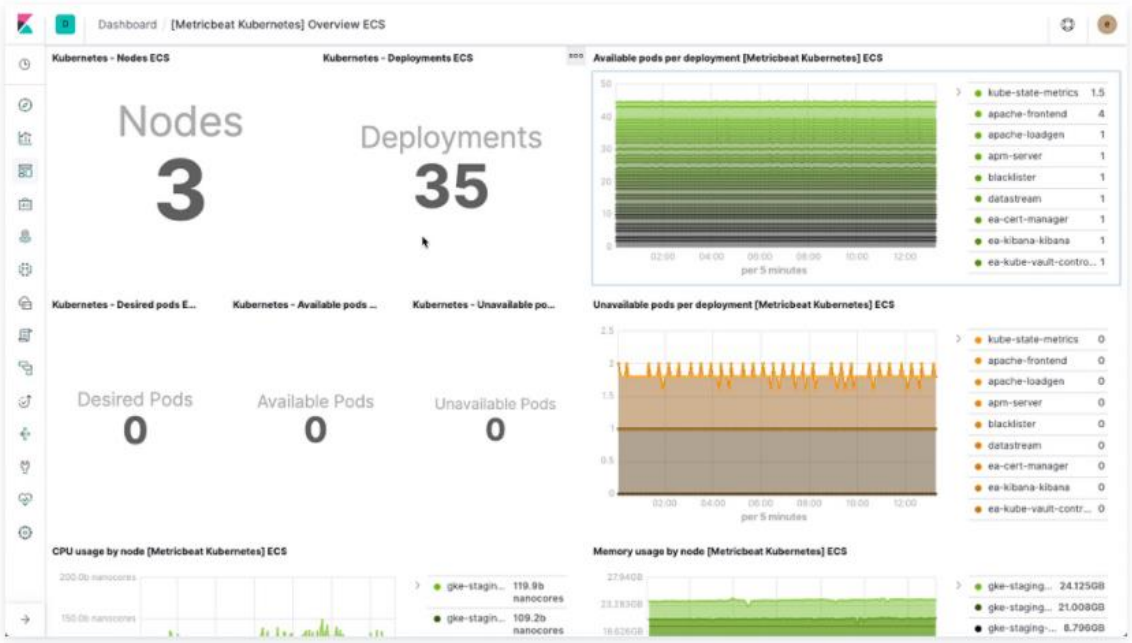
System

Analyze a holistic view of your systems and view activity by host in just click.



Kubernetes

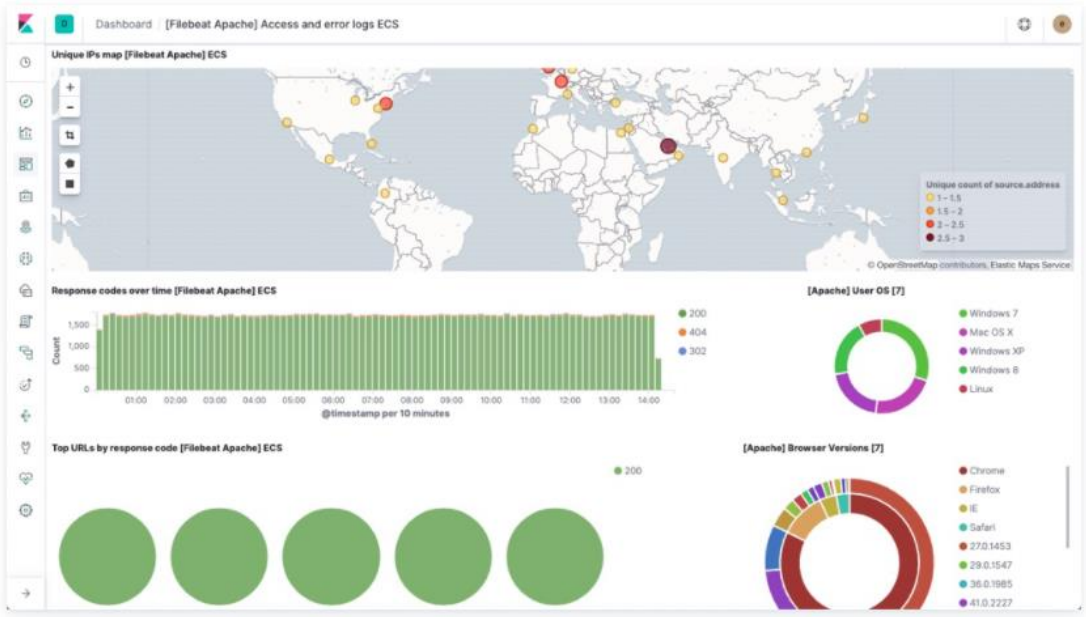
Get visibility into your orchestration environment and filter your kubernetes logs by hosts, pod or custom metadata



ELASTIC OBSERVABILITY WITH LOGS

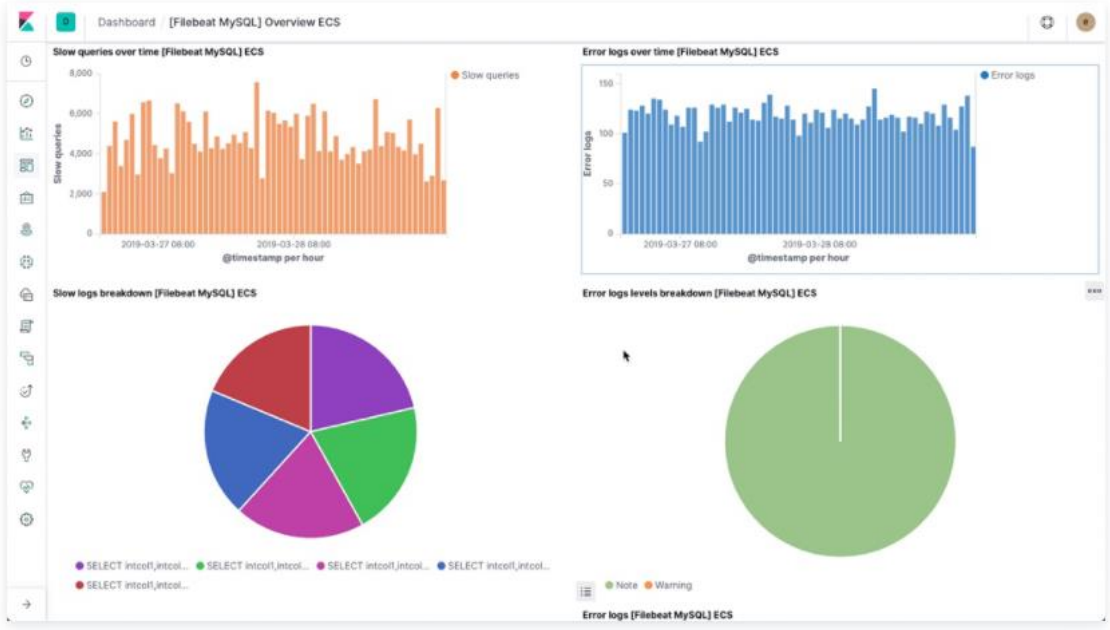
Apache & NGINX

Keep an eye on your web services by parsing and ingesting logs with Apache and NGINX Modules



MySQL & PostgreSQL

Monitor MySQL & PostgreSQL Databases and use error logs to identify instances that might be having trouble



ELASTIC OBSERVABILITY WITH LOGS

Tail a file directly in the UI

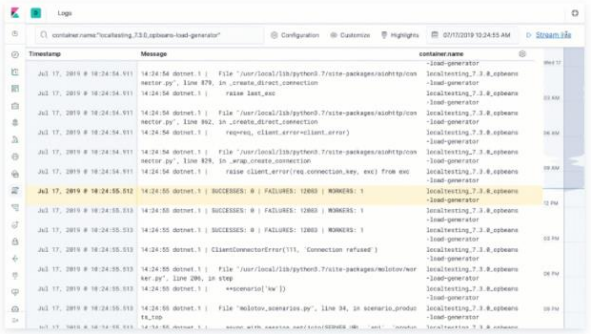
Keep a pulse on all of the logs flowing in from your servers, virtual machines, and containers, applications in a centralized view built for infrastructure operations. Pin structured fields like IP or event type, and explore related logs without leaving your current screen

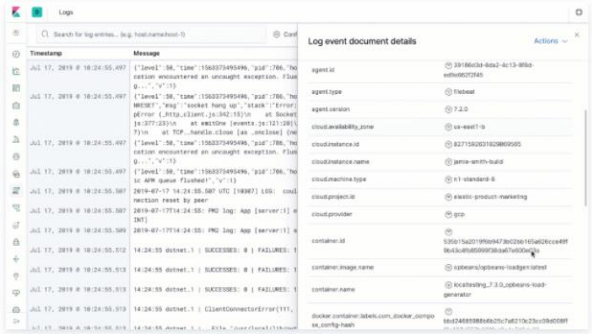
Analyze trends with categorized logs

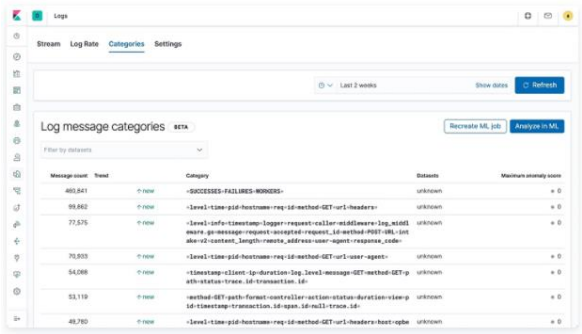
You can easily analyze for patterns in your event data using the log categorization view.

Add machine learning to automate anomaly detection

Elastic's ML feature enables to automatically model the behavior of your elasticsearch data and alert you on issues in real time.



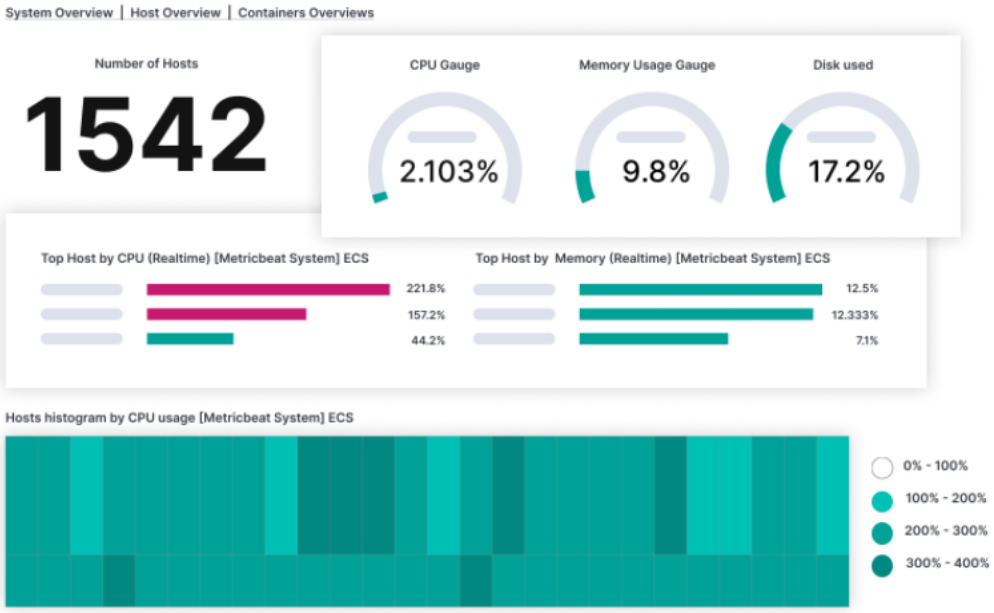




ELASTIC OBSERVABILITY WITH METRICS

Servers

Find hotspots, diagnose problematic spikes, and customize everything from colors to alerts



Docker

Monitor containers with high resource utilization and identify containers that have stopped.

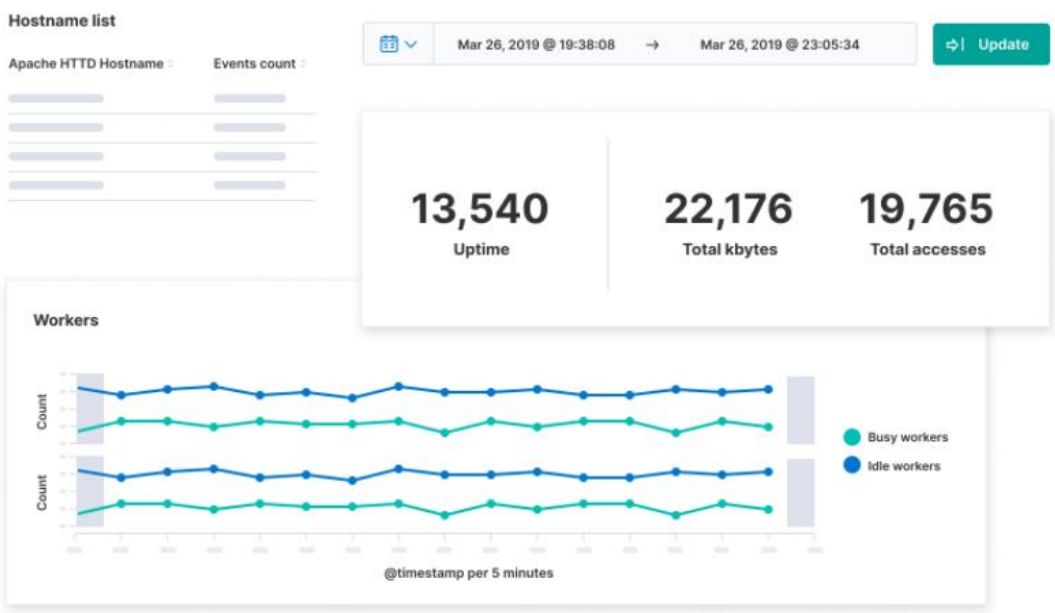
ELASTIC OBSERVABILITY WITH METRICS

Kubernetes

Keep an eye on your pods and use auto discovery to track and adapt to the dynamic environment

Applications

Visualize application metrics in a flash with a vast array of prebuilt Metric beat modules.

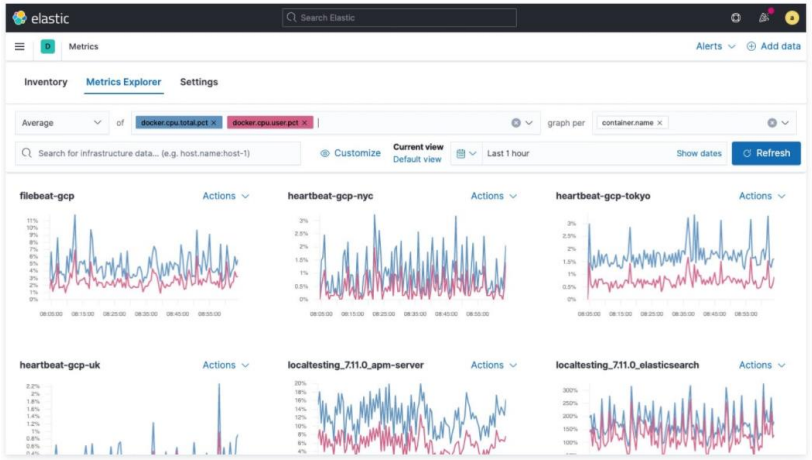


ELASTIC OBSERVABILITY WITH METRICS

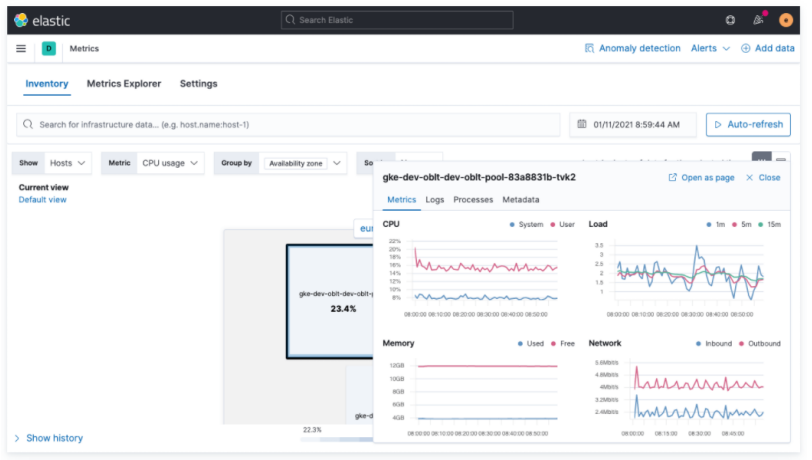
Deep dive into your metrics

Get a different view of your infrastructure with perspectives aligned with your topology. Dig into current and historical performance by CPU, memory, or network traffic.

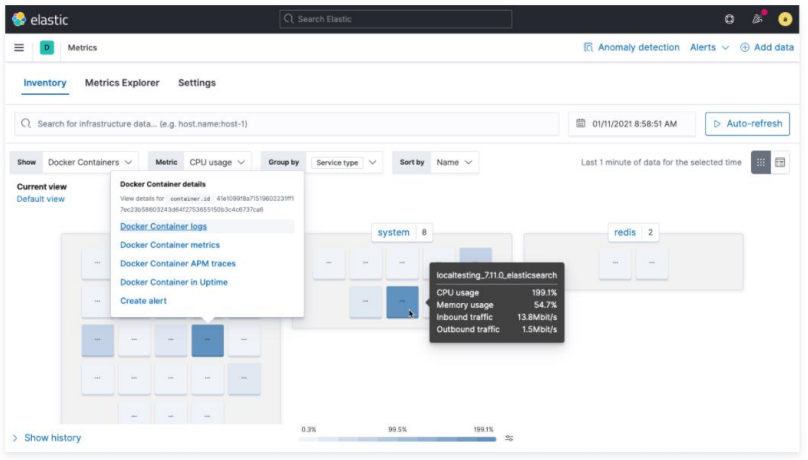
Metrics Explorer



Historical Details



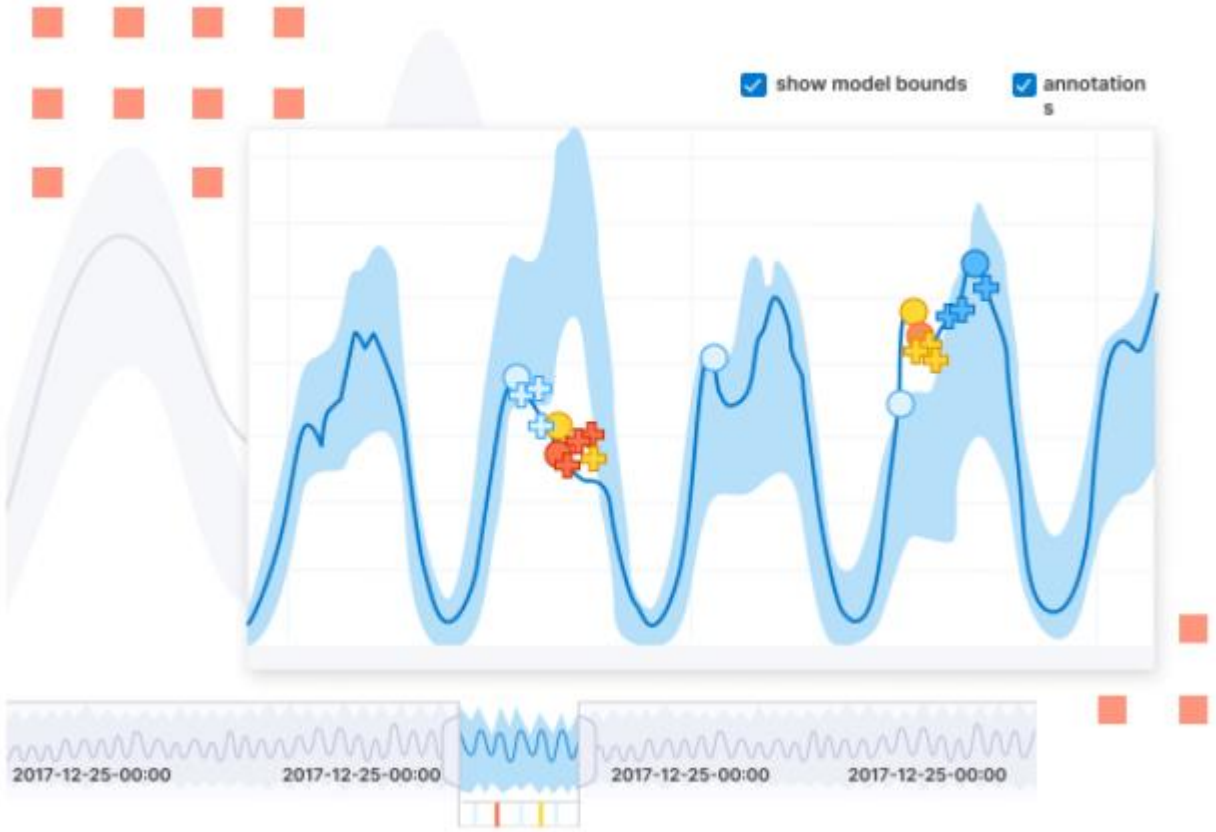
Inventory View



ELASTIC OBSERVABILITY WITH METRICS

Combine Machine Learning and Alerts

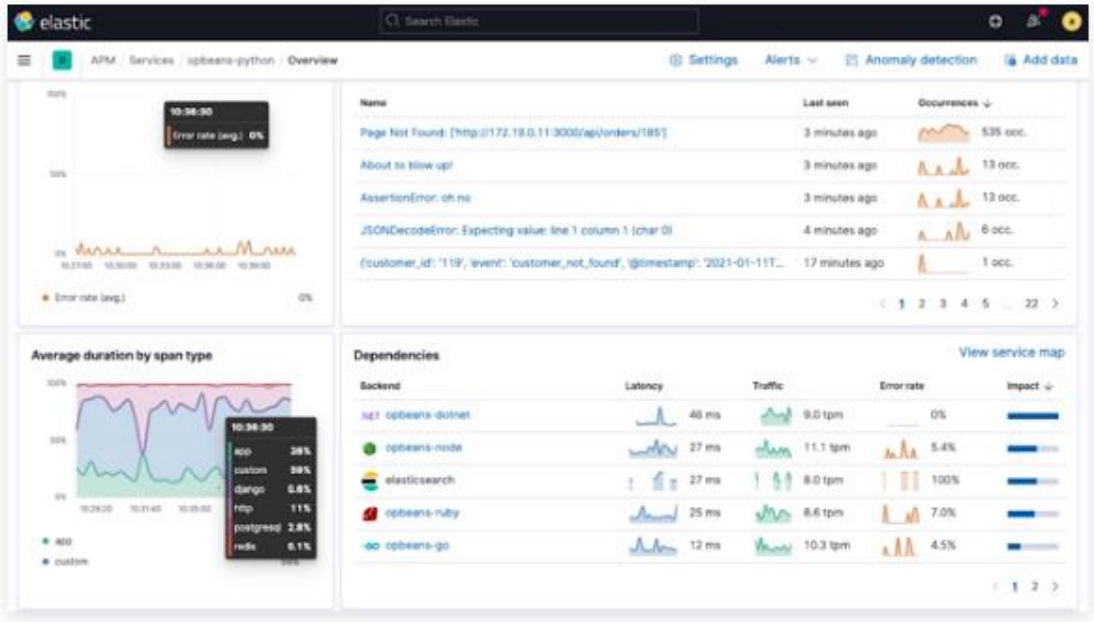
As data scales, its easy to lose errant data points among the streaming averages, measurements, and totals. Define machine learning logs in a few clicks to start detecting anomalies in your data. Then create alerts so you are ready take action when something is awry.



ELASTIC OBSERVABILITY WITH APM

Root cause investigations

Using the APM Module of ELK stack we can easily diagnose the root cause of issues and performance bottlenecks.



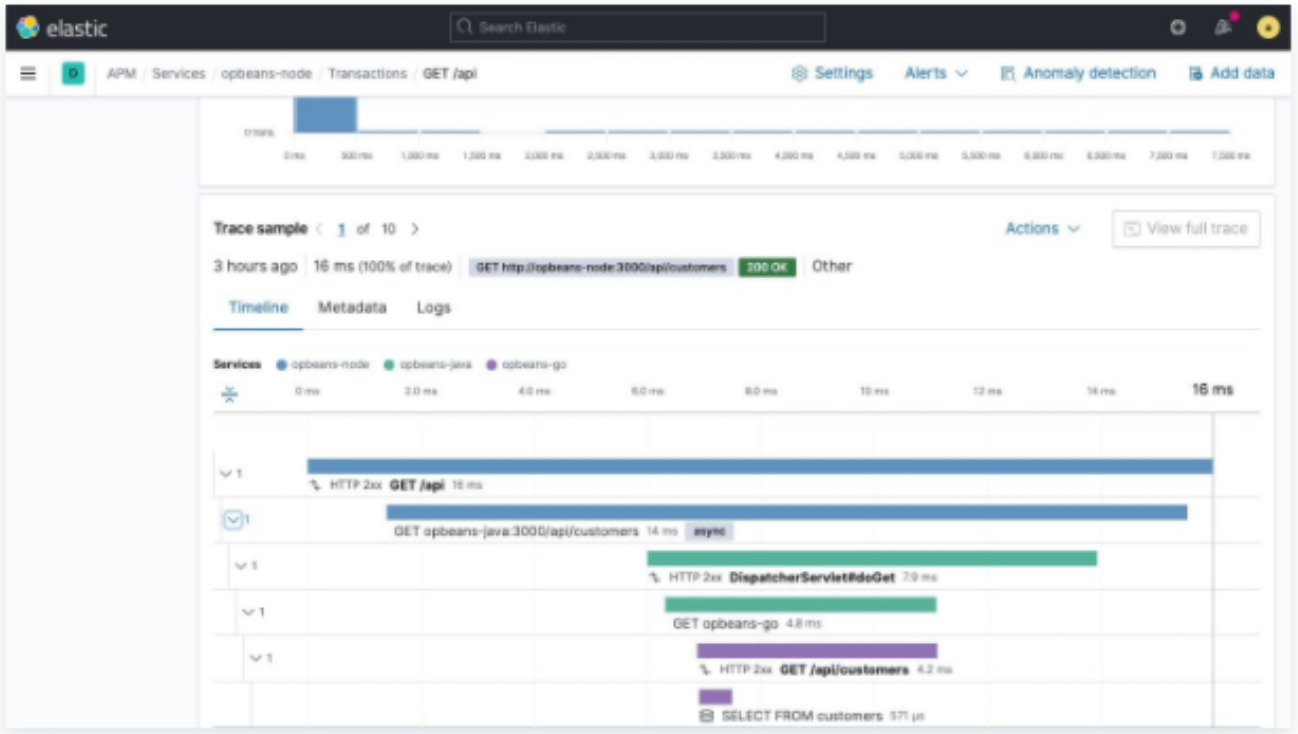
Connect the dots with service maps

Get a visual representation of how your services are connected in one, clear cut view. See how they are performing with highlighted KPIs, identify potential problems and health issues powered by ML.

ELASTIC OBSERVABILITY WITH APM

Distributed Tracing

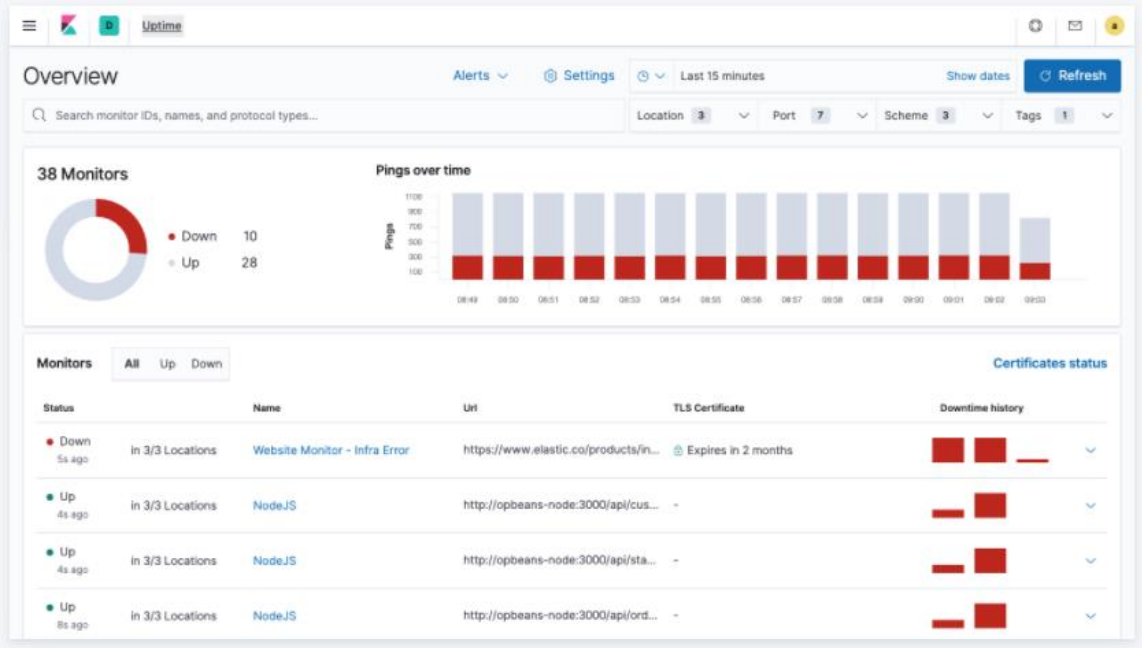
Easily analyze how requests are flowing through your entire infrastructure using the “Distributed Tracing”. It enables to see which messaging queues were utilized and visualize service calls across them, find where latency issues are arising in the path, and pinpoint the components that need optimizing.



ELASTIC OBSERVABILITY WITH UPTIME

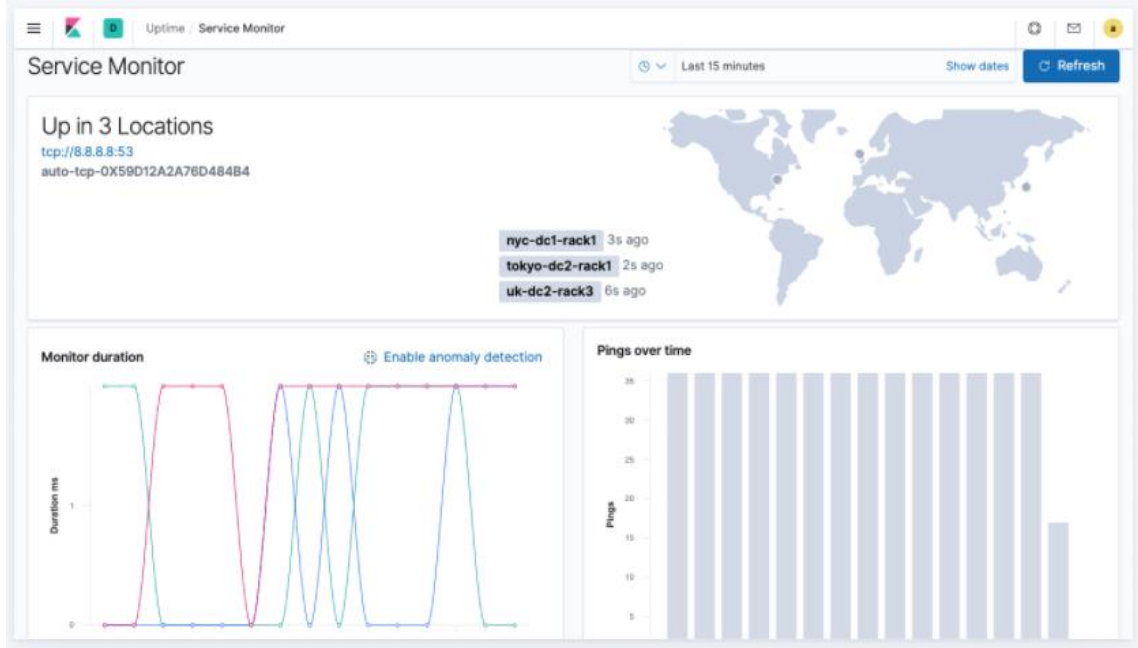
Host availability

Enables you to keep track of your host, network device or even third-party services availability using basic ICMP ping



Service Monitoring

Enables you to monitor your critical services using TCP checks using ports. Example: DNS, FTP



ELASTIC SECURITY

Elastic Security features equips analysts to prevent, detect, and respond to threats. Elastic supports SIEM, endpoint security, threat hunting, cloud monitoring and more.

SIEM **(Security Information and Event Management)**

Threat detection and response on the Elastic Stack. Detect complex threats with prebuilt anomaly detection jobs and publically available detection rules.

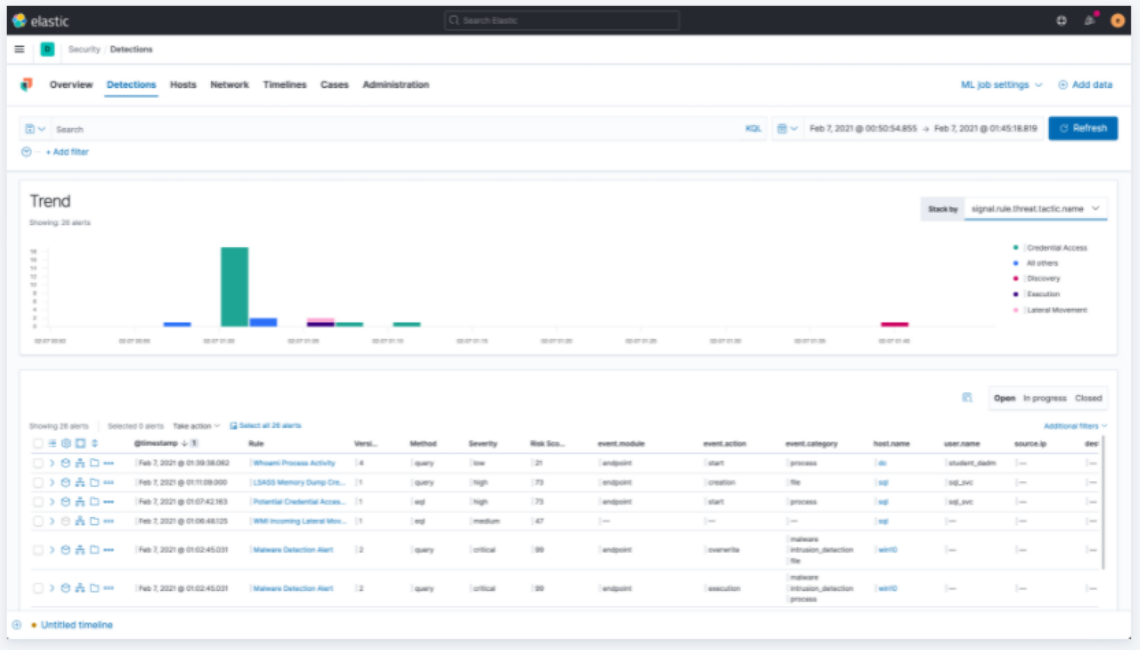
Endpoint Security

Anti-malware, ransomware prevention, and data collection.

ELASTIC SECURITY SIEM CONTINUED

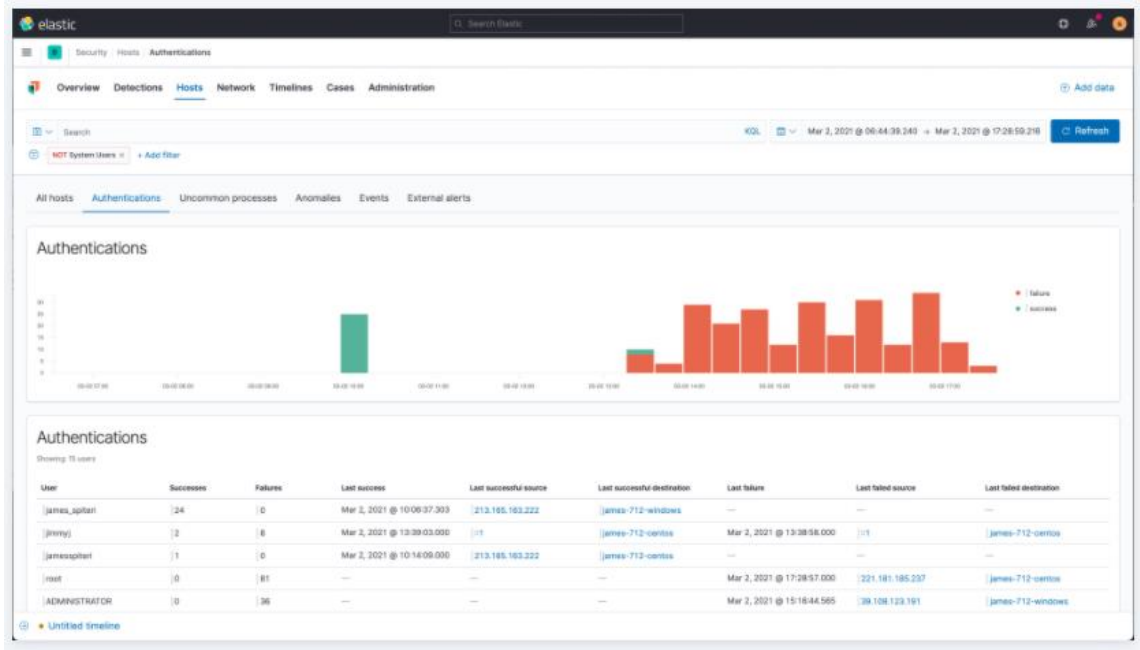
Endpoint activity

Automate detection across your endpoint data to find uncommon processes, anomalies and more



Authentication Logs

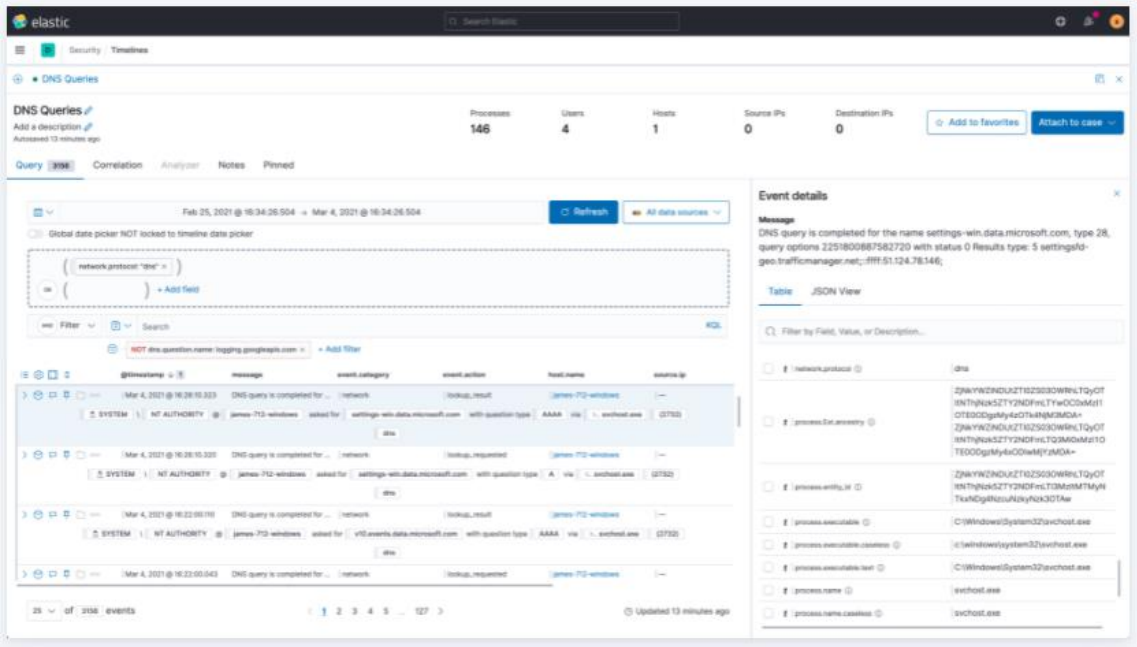
Investigate attempted logins and related activity with authentication data.



ELASTIC SECURITY SIEM CONTINUED

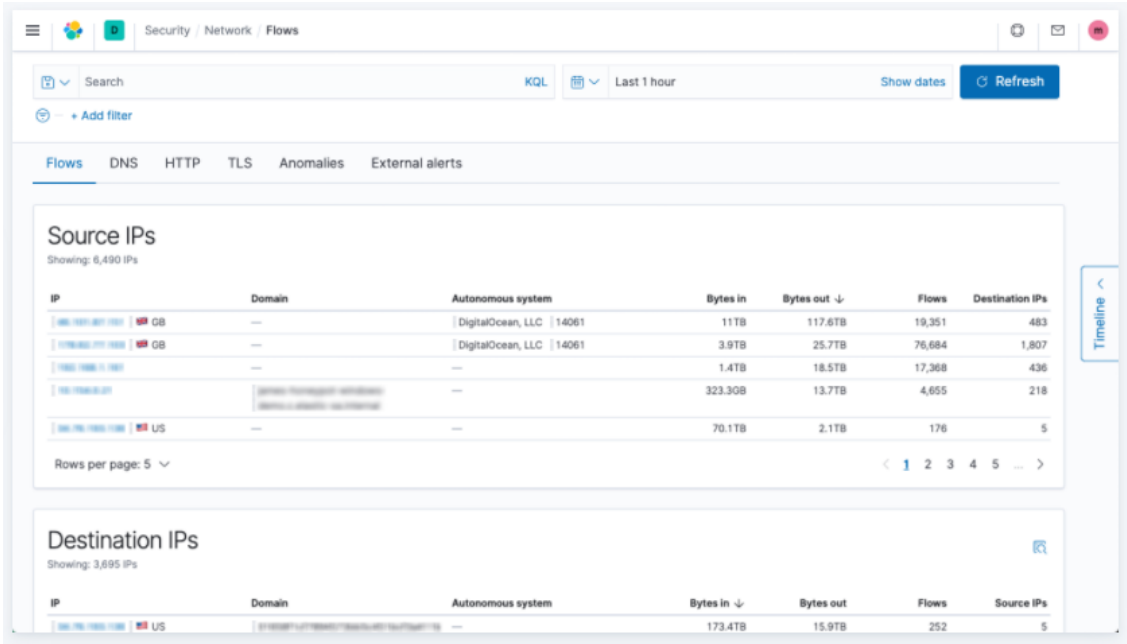
DNS Traffic

Easily analyze vast volumes of DNS data: user access patterns, domain activity, query trends and more.



NetFlow

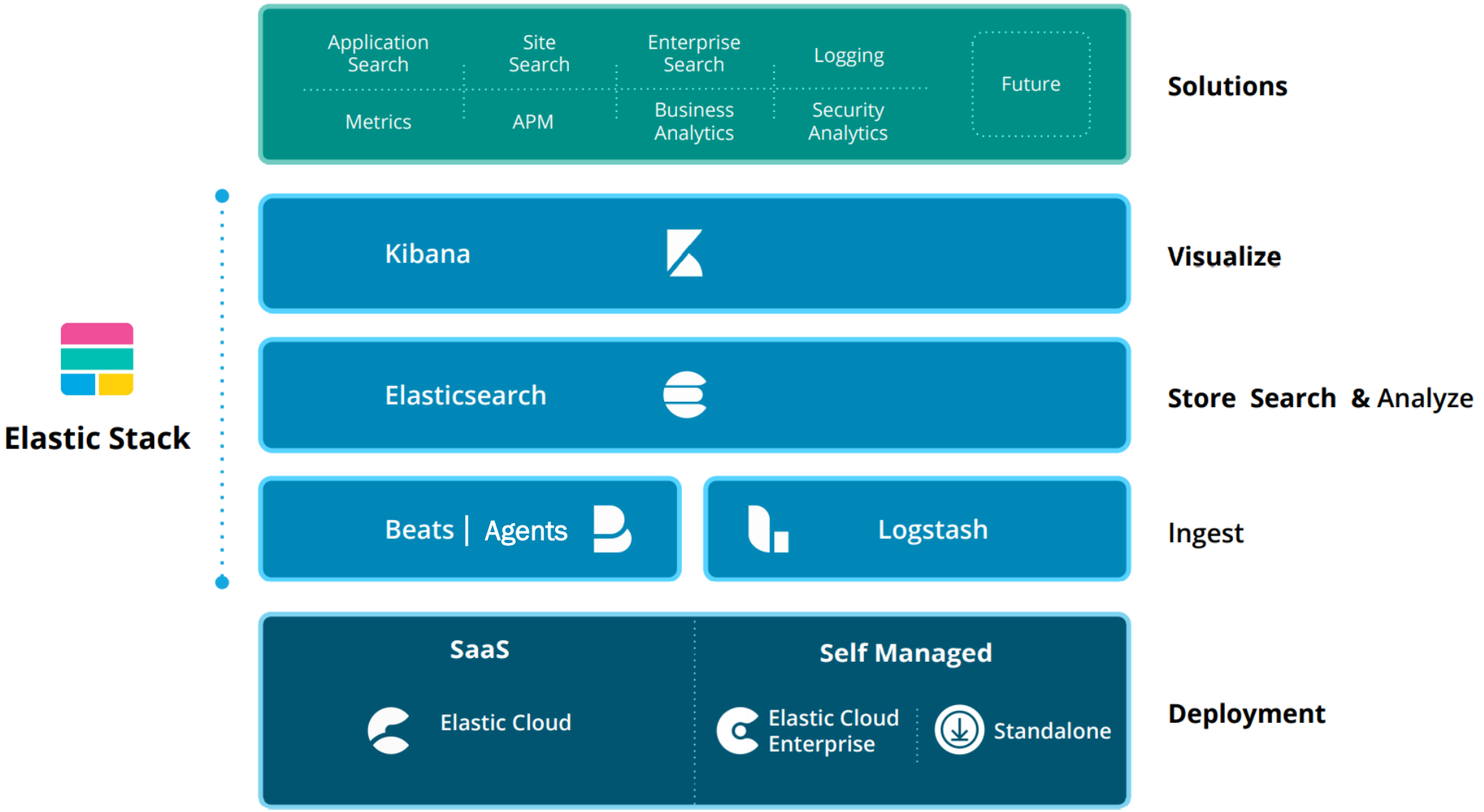
Establish environmental visibility by analyzing flow data at massive scale



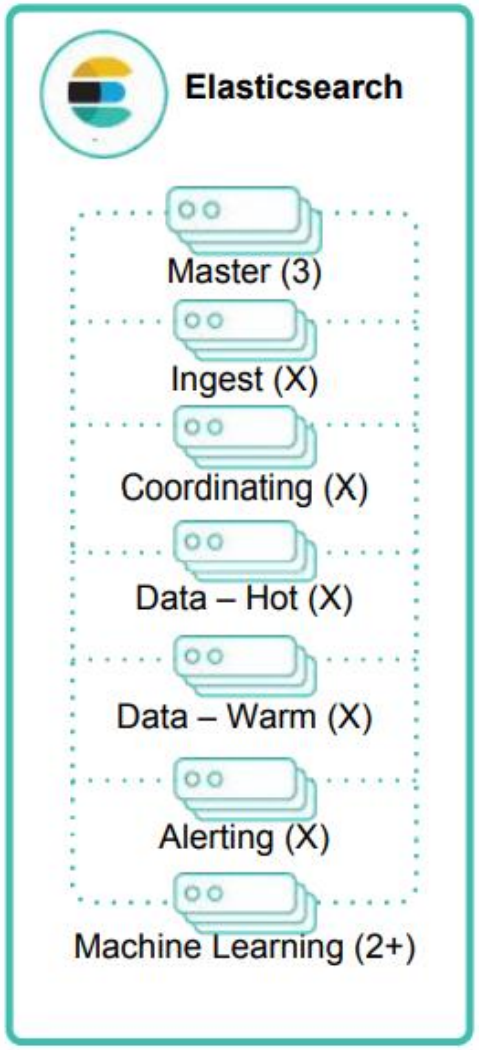
LET'S LOOK AT ELASTIC SEARCH ARCHITECTURE

WHAT ARE THE BUILDING BLOCKS OR ELASTIC STACK AND THE ENTERPRISE DEPLOYMENT ARCHITECTURES

ELASTIC STACK



ELASTICSEARCH NODE TYPES



- Master Nodes
  - Control the cluster, requires a minimum of 3, one is active at any given time
- Data Nodes
  - Hold indexed data and perform data related operations
  - Differentiated hot and warm data nodes can be used
- Ingest Nodes
  - Use ingest pipelines to transform and enrich before indexing
- Coordinating Nodes
  - Route requests, handle search reduce phase, distribute bulk indexing
  - All nodes function as coordinating nodes
- Alerting Nodes
  - Run alerting jobs
- Machine Learning Nodes
  - Run machine learning jobs

ELASTICSEARCH BEATS TYPES

- Filebeat
  - Lightweight shipper of logs and other data
- Metricbeat
  - Lightweight shipper of metric data
- Packetbeat
  - Lightweight shipper of network data
- Winlogbeat
  - Lightweight shipper of windows even logs
- Auditbeat
  - Lightweight shipper of audit data
- Heartbeat
  - Lightweight shipper for uptime monitoring
- Functionbeat
  - Serverless shipper of cloud data

ELASTICSEARCH AGENT

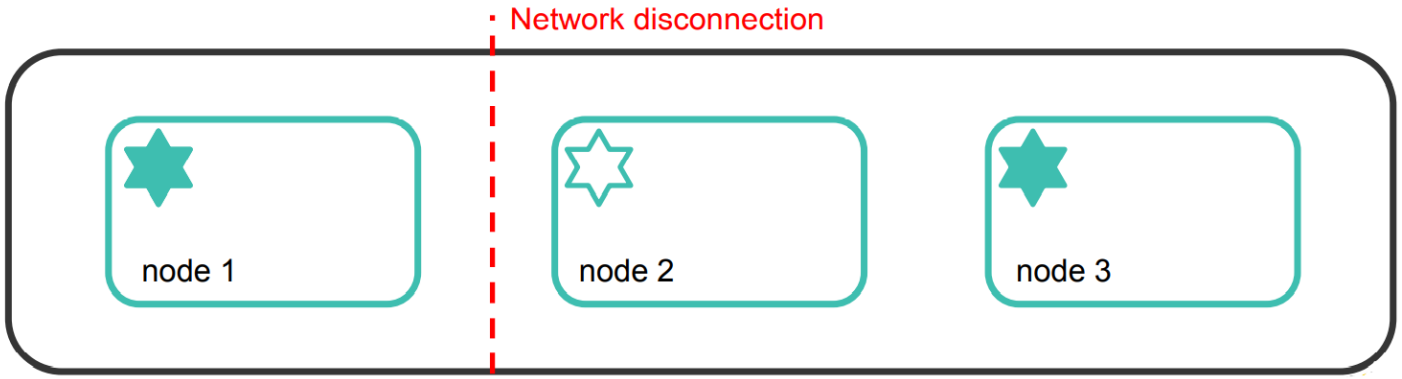
Elastic-Agent is a new solution from elastic where a single **elastic-agent** is has built-in Filebeat, Metricbeat, Auditbeat etc. capabilities. Therefore, a elastic-agents provides easier data collection by installing just one module. You need a additional module called "**fleet servers**" in the cluster to manage the "elastic-agents" the "fleet server" module just like other cluster nodes can be installed and run on the same elastic cluster nodes if required.

LET'S BUILD A MULTI NODE ELASTIC CLUSTER

MULTI NODE ELASTIC CLUSTER WITH FILE BEATS METRIC BEATS ON APACHE, NGINX, MYSQL & POSTGRESQL

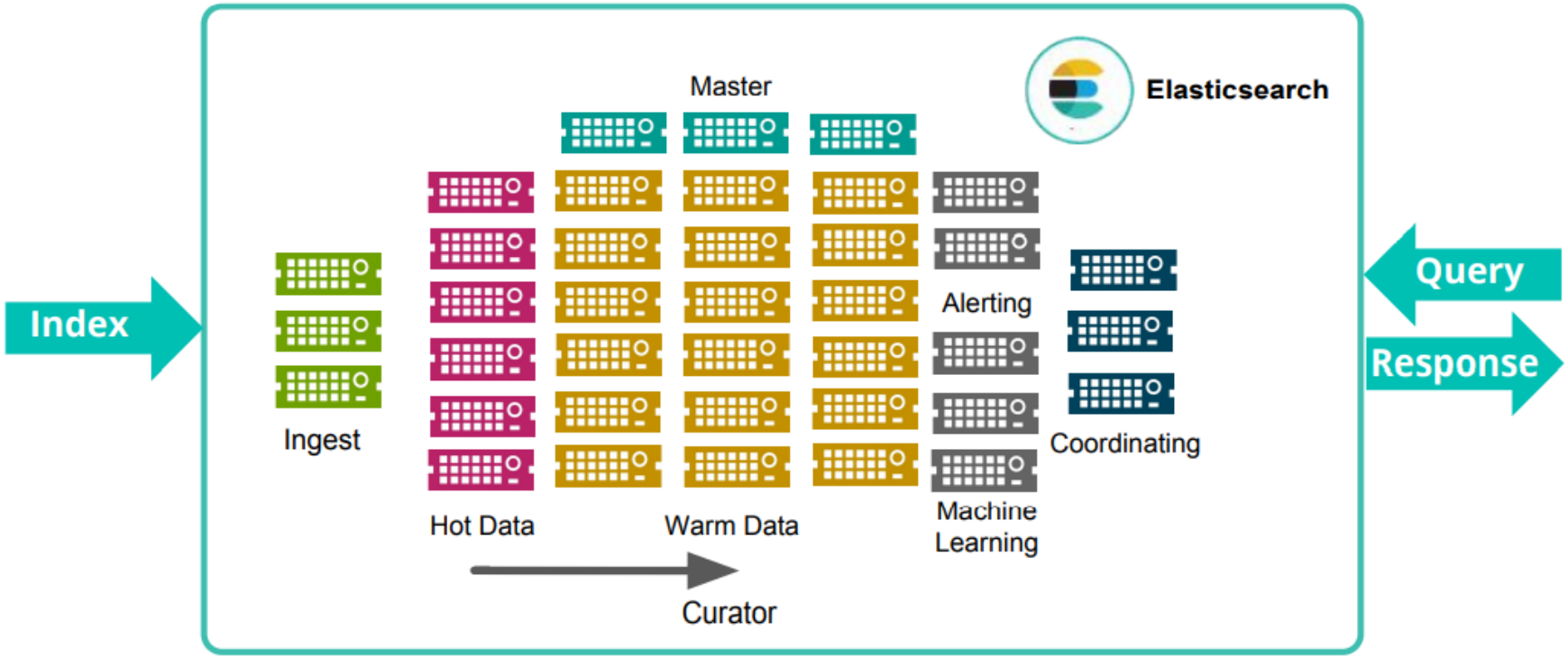
ELASTICSEARCH CLUSTER : MASTER NODES

- Cluster with 3 master eligible nodes
- Concern if network becomes partitioned
- The cluster would inadvertently elect two masters, which is referred to as “Split Brain”
- A master eligible node needs at least “minimum_master_nodes” votes to win an election
  - Setting it to a quorum prevents the split brain scenario
- Recommendation for production clusters is to have 3 dedicated master eligible nodes
  - With the setting “minimum_master_nodes = 2”



INSIDE A LARGE ELASTICSEARCH LOGGING CLUSTER

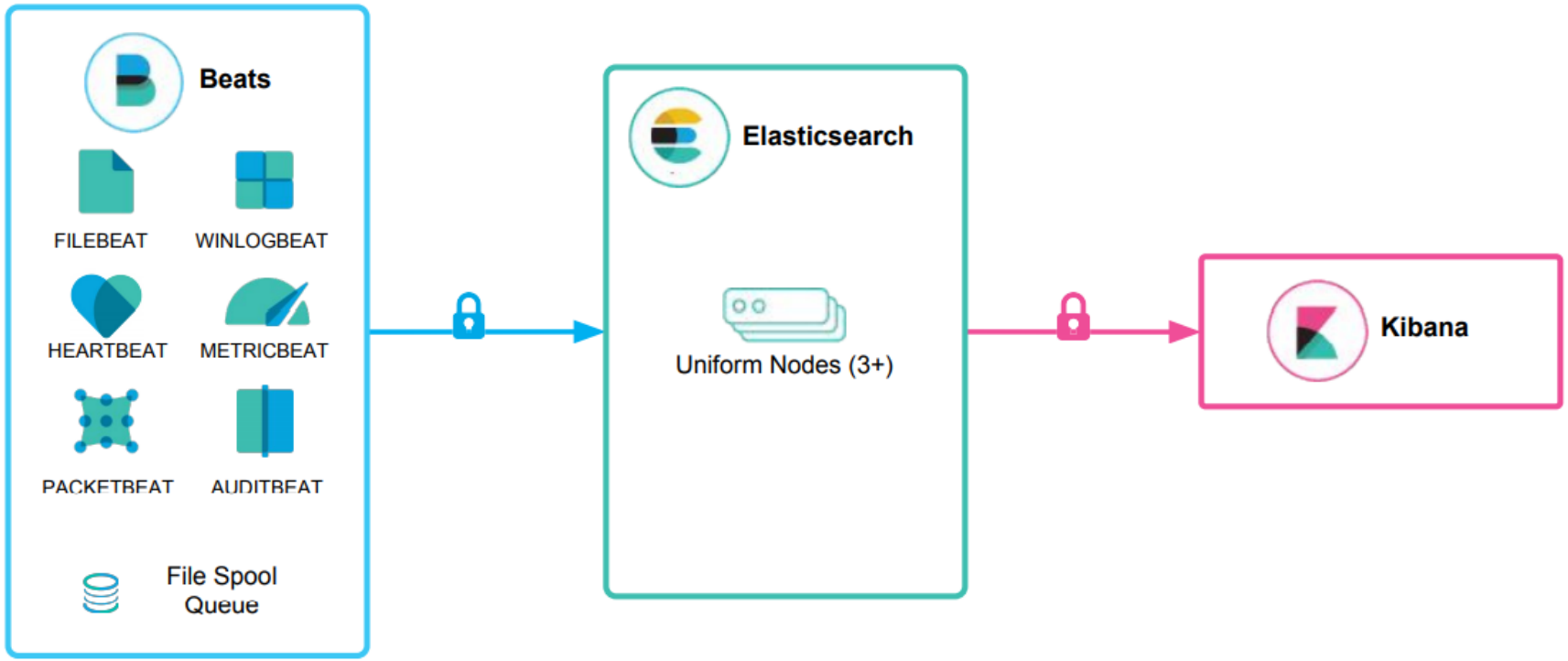
- A multi node elasticsearch logging cluster with dedicated nodes for different roles.
- Storage Data Life Cycle feature of Elasticsearch
  - Data moved between different data nodes according to their current value (hot data, warm data and cold data)



ELASTICSEARCH ENTERPRISE ARCHITECTURES

Quick Start

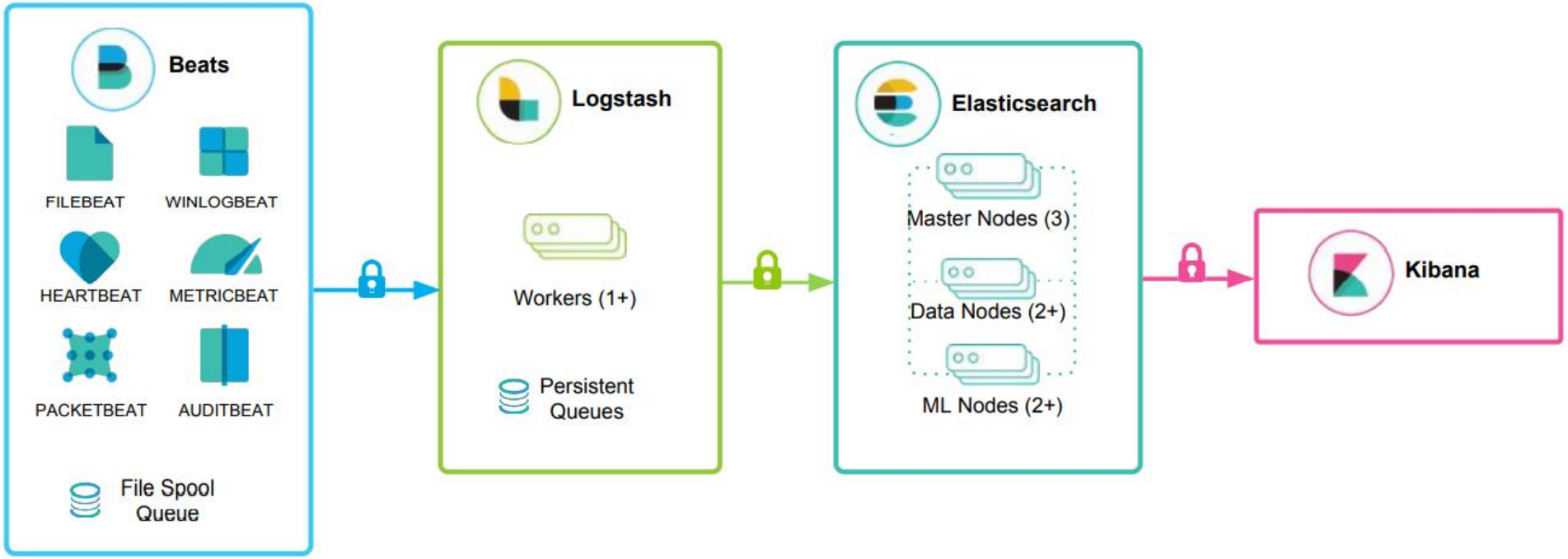
Beats, Elasticsearch and Kibana



ELASTICSEARCH ENTERPRISE ARCHITECTURES

Advanced Processing and Resiliency

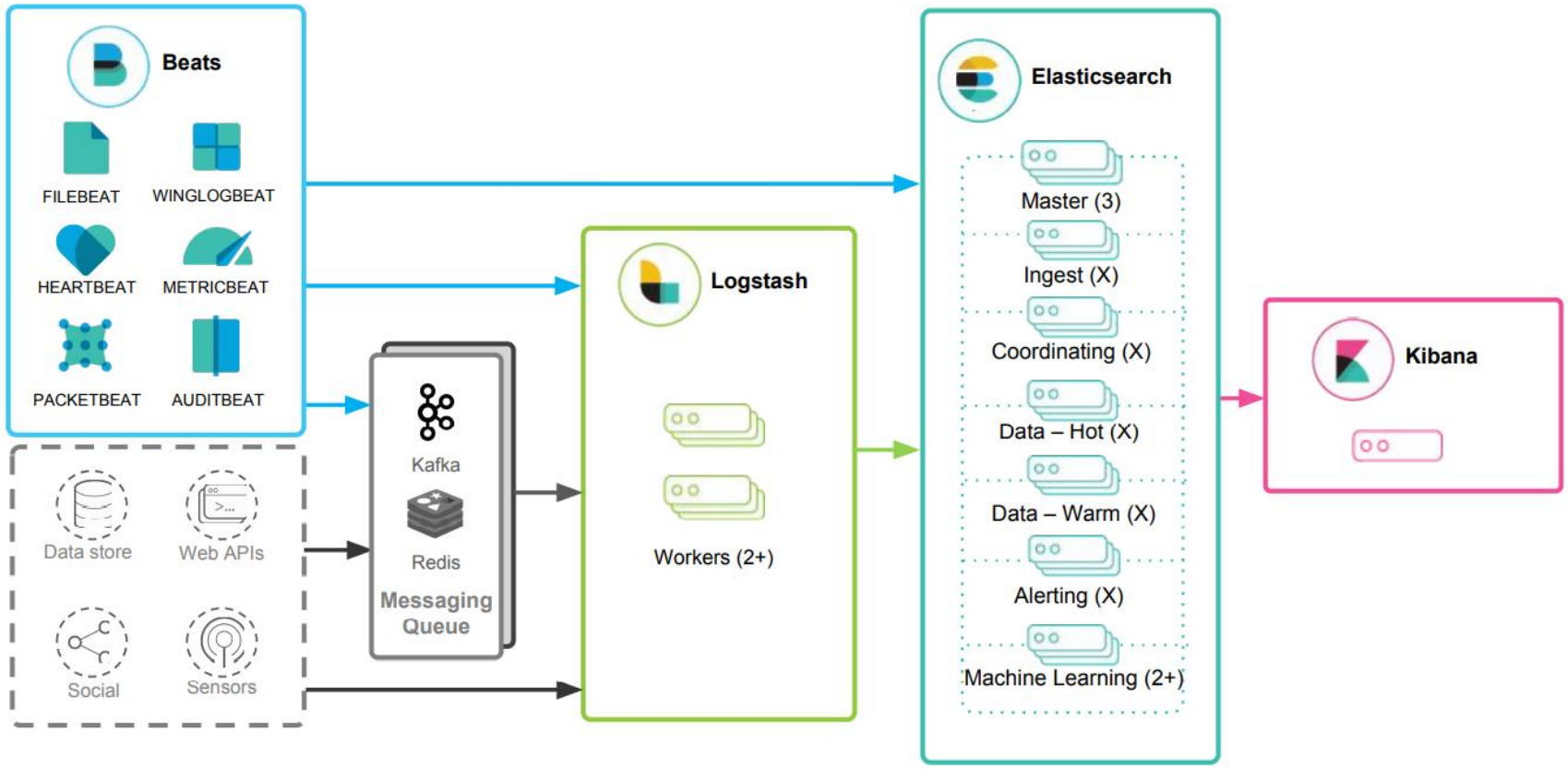
Adding Logstash processing, differentiated Elasticsearch node types



ELASTICSEARCH ENTERPRISE ARCHITECTURES

Flexible Ingestion and Input Sources

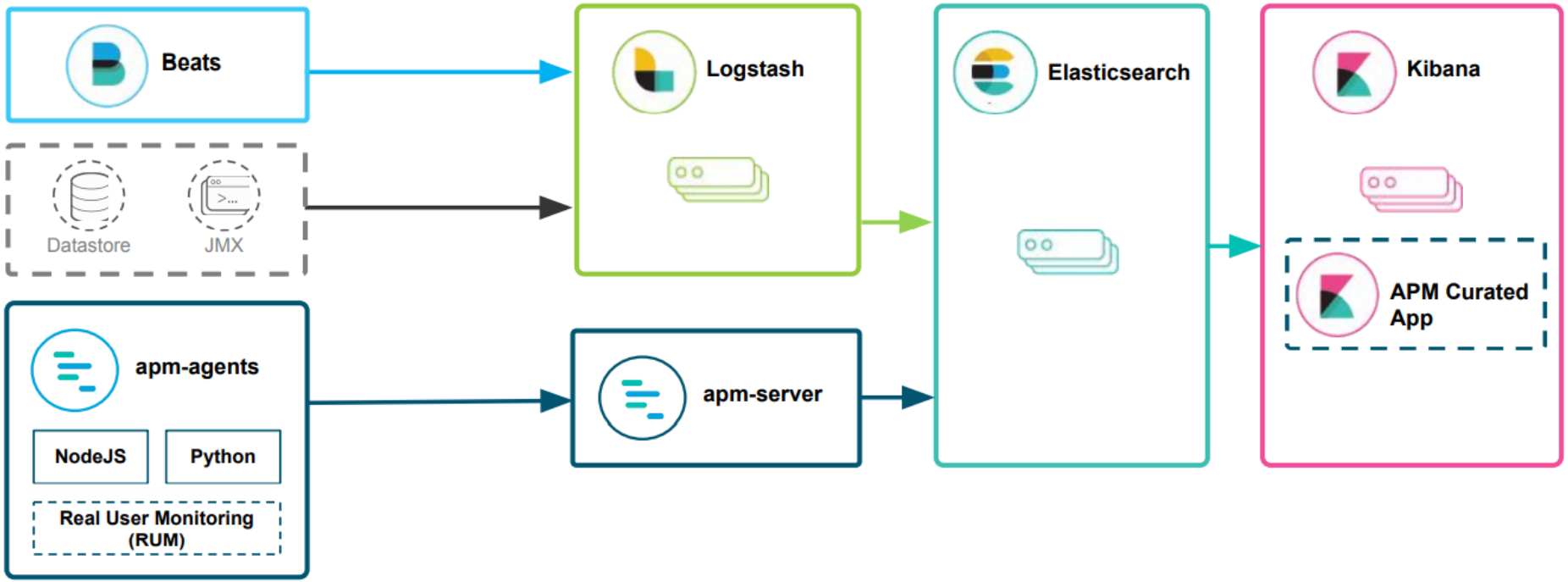
Adding messaging queues to handle velocity



ELASTICSEARCH ENTERPRISE ARCHITECTURES

Application Metric Collection with Elastic APM

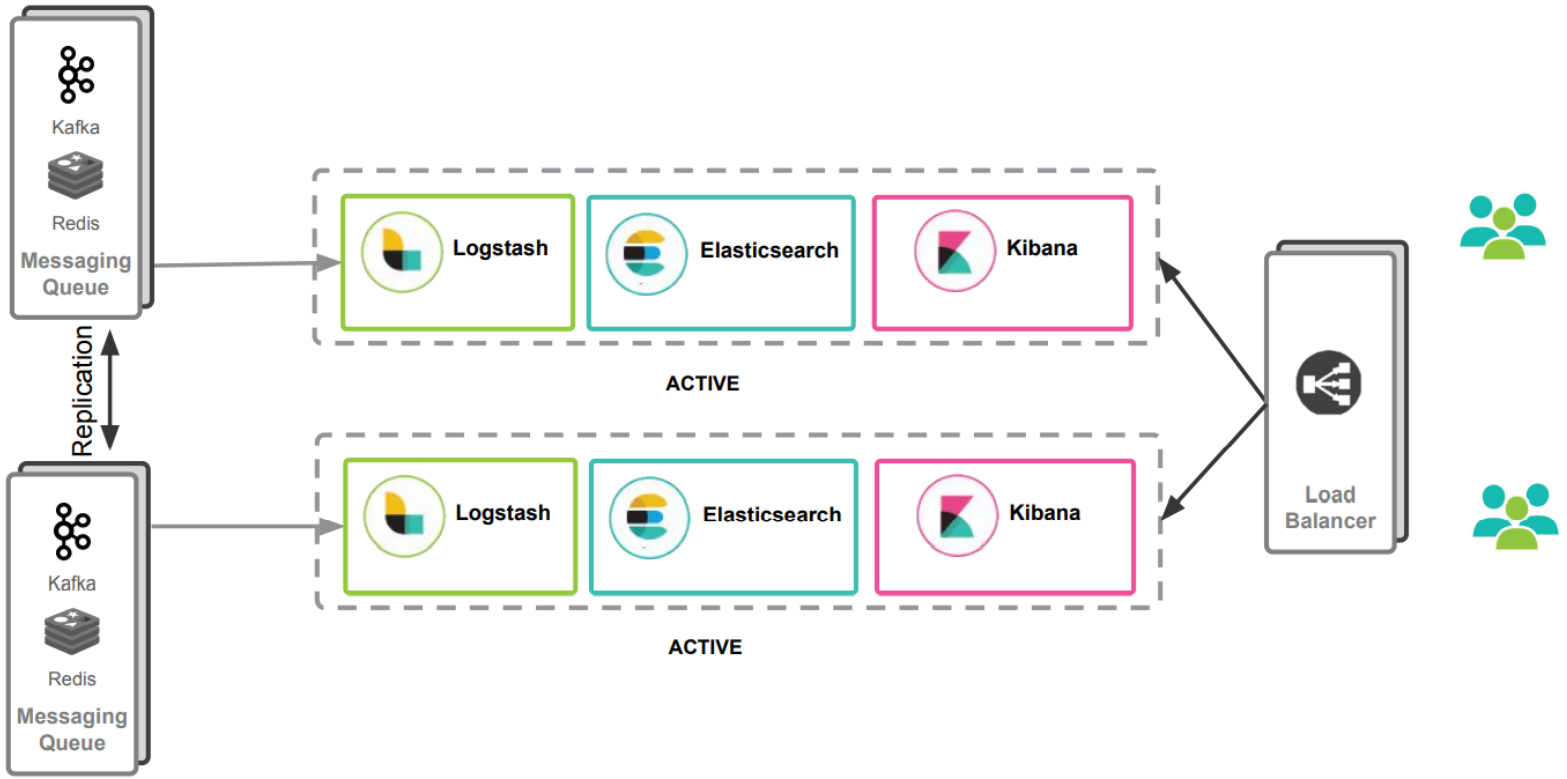
Adding messaging queues to handle velocity



ELASTICSEARCH ENTERPRISE ARCHITECTURES

Multiple Data Centers, Duplicate Data

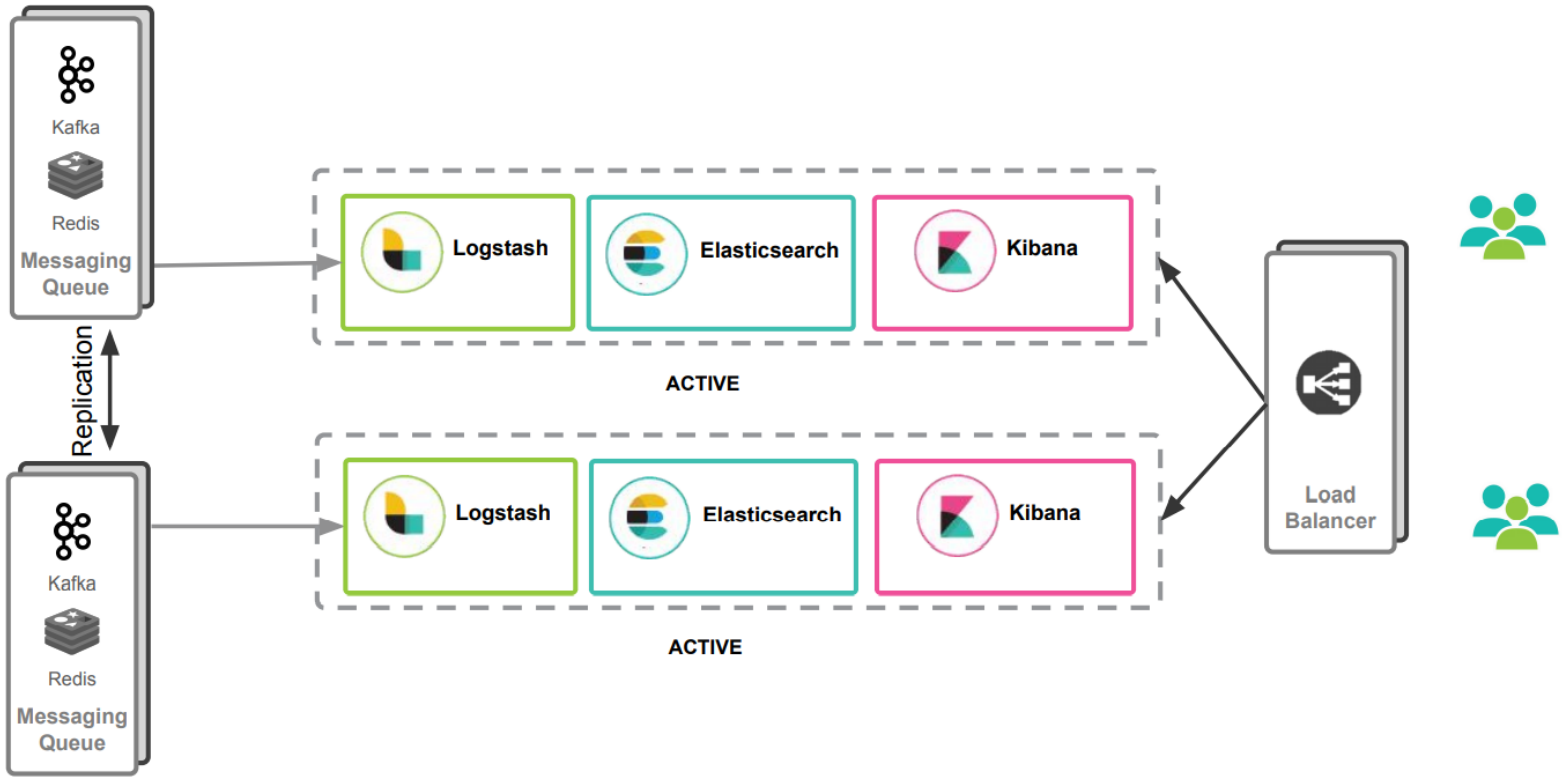
Elastic Cross Cluster Replication enables to replicate the data across sites which are geographically separated



ELASTICSEARCH ENTERPRISE ARCHITECTURES

Multiple Data Centers, Distinct Data and Cross Cluster Search

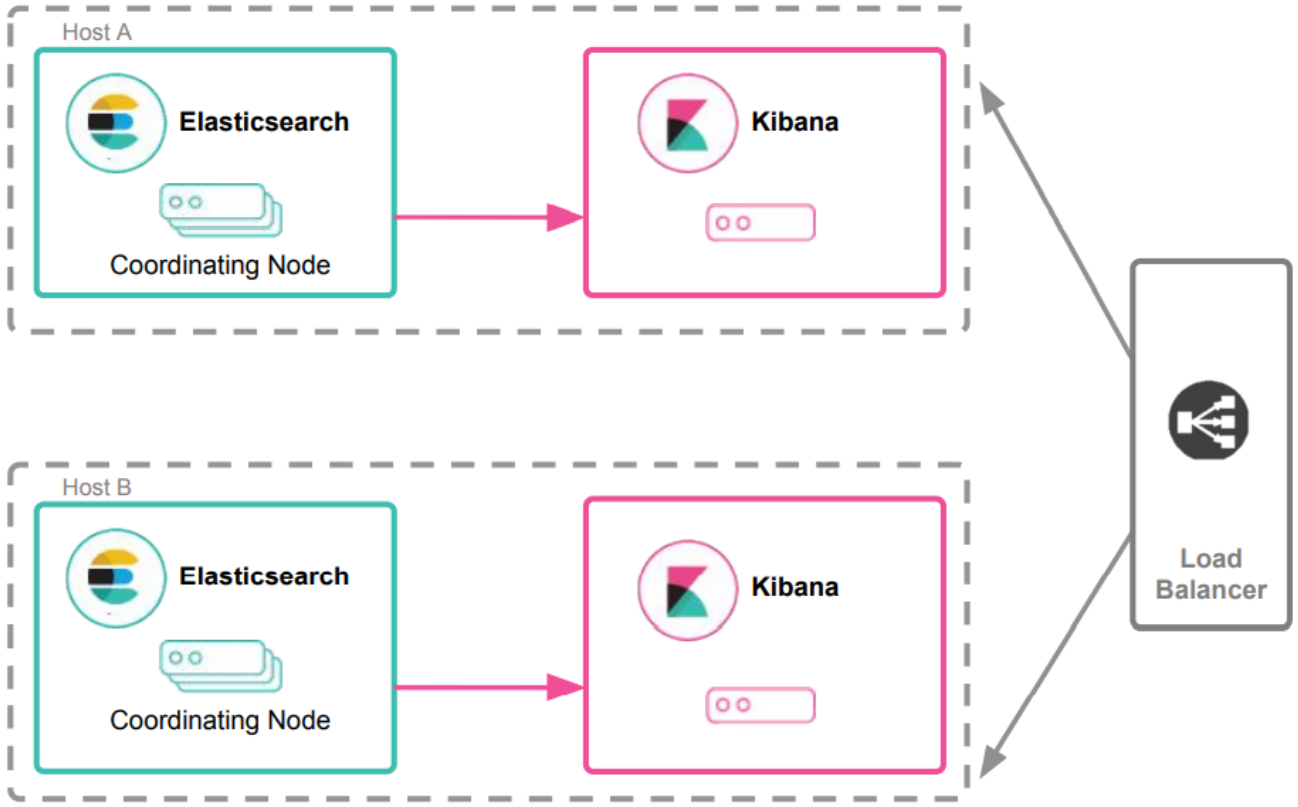
Elastic Cross Cluster Replication enables to replicate the data across sites which are geographically separated



ELASTICSEARCH ENTERPRISE ARCHITECTURES

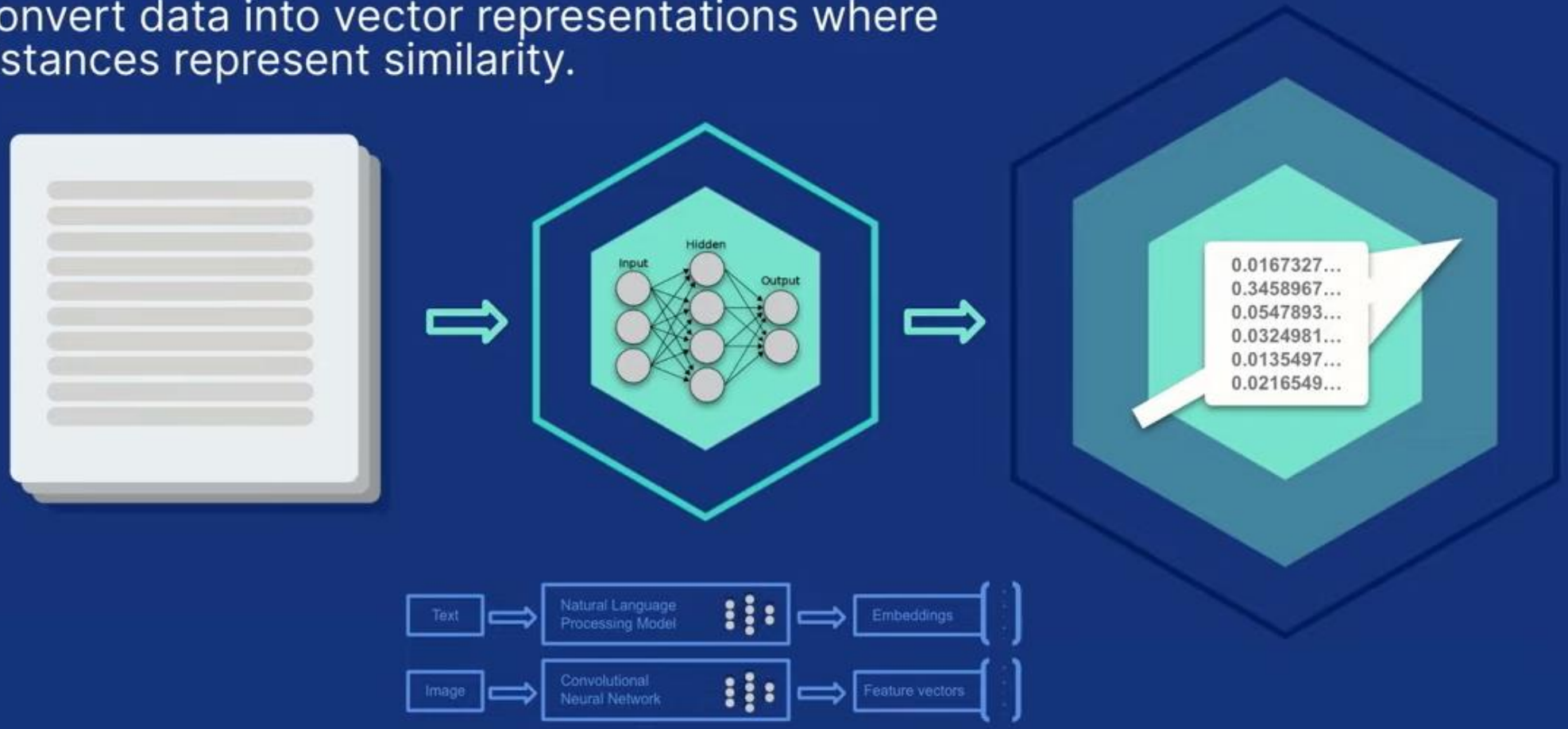
High Availability for Kibana

Pairing two coordinating nodes with two independent Kibana Nodes



What is vector similarity?

Convert data into vector representations where distances represent similarity.



elastic

Embeddings represent aspects of the data

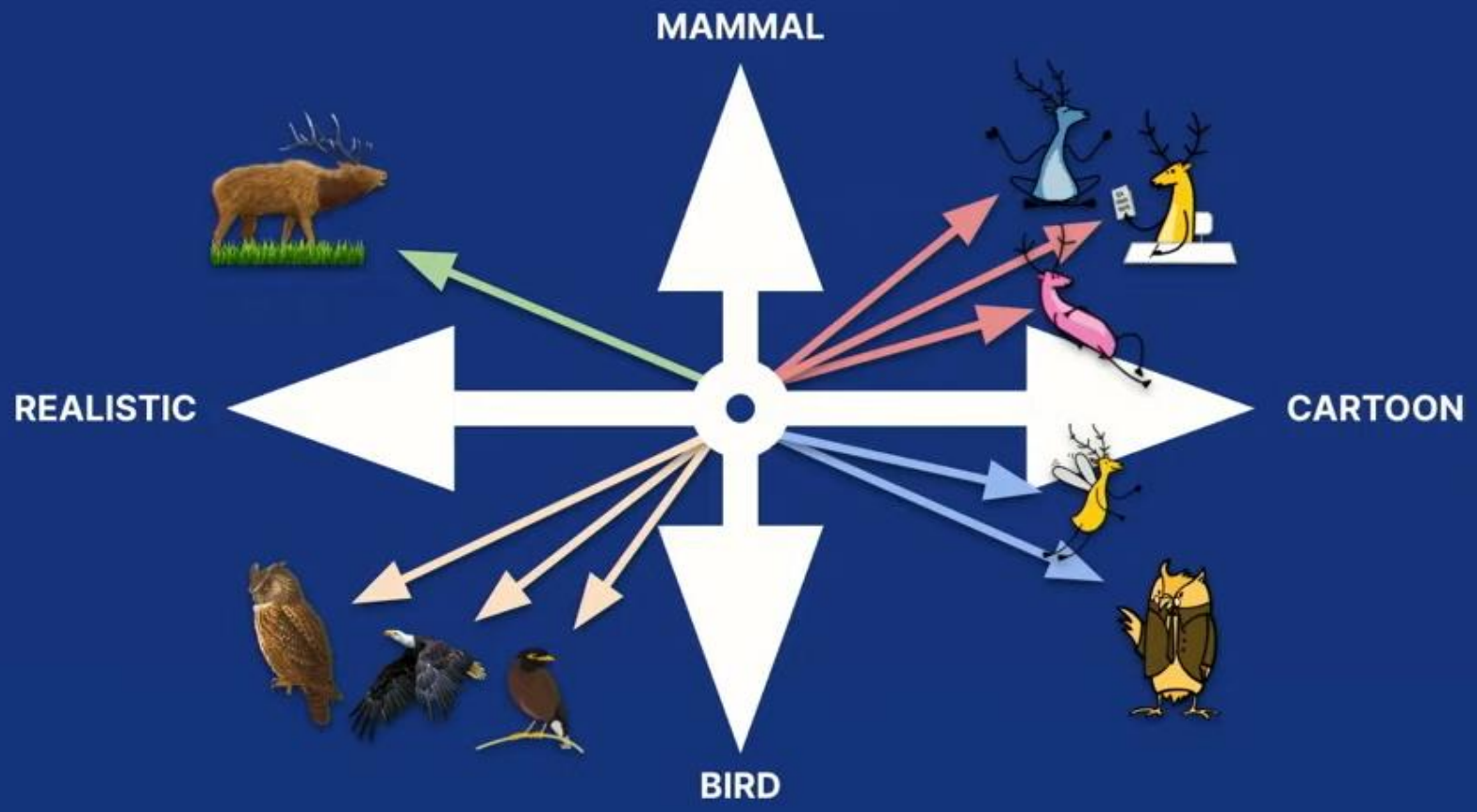
Example: 1-dimensional vector

REALISTIC

CARTOON

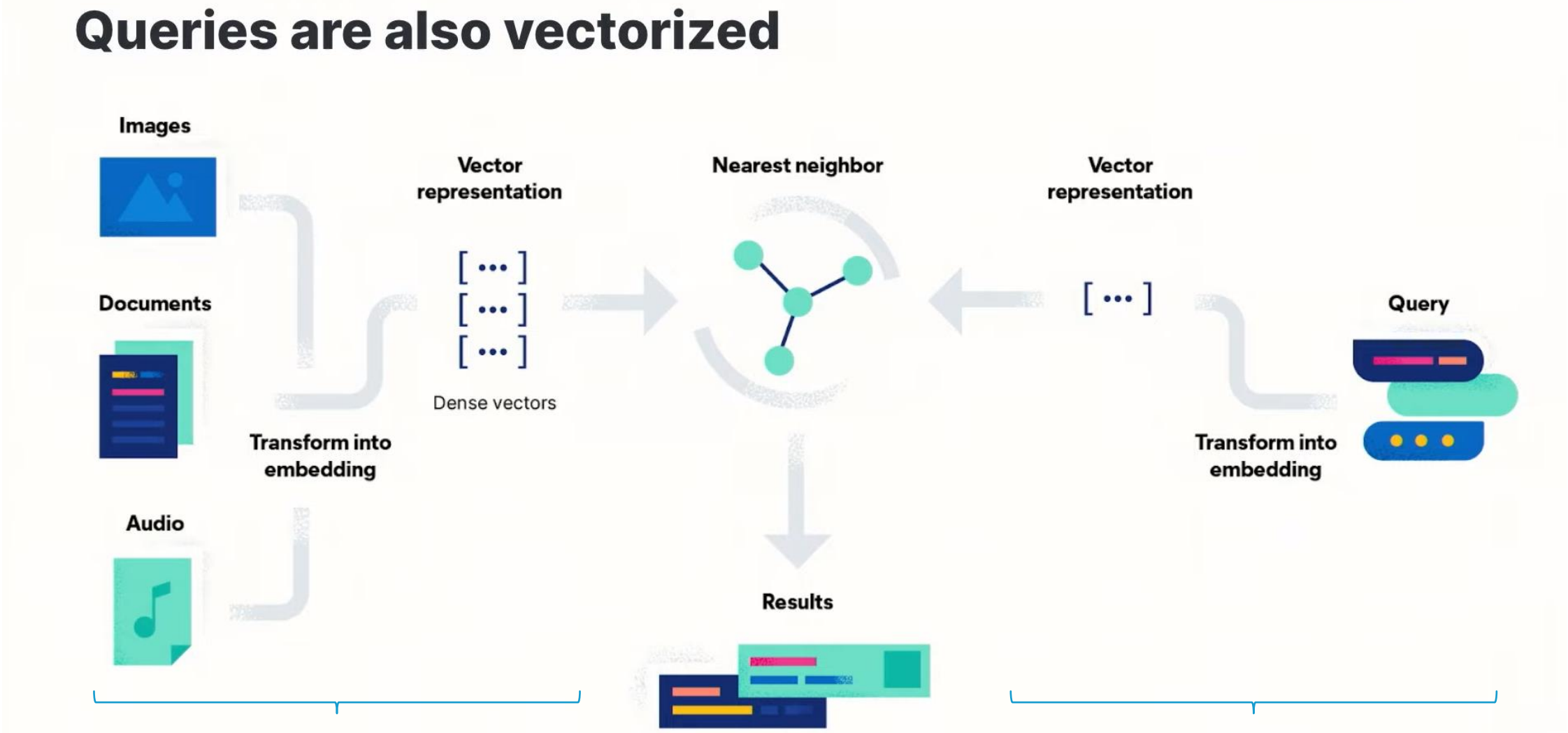
elastic

In the “Embedding space”, similar data are grouped together



elastic

Process of implementing “Semantic Search”

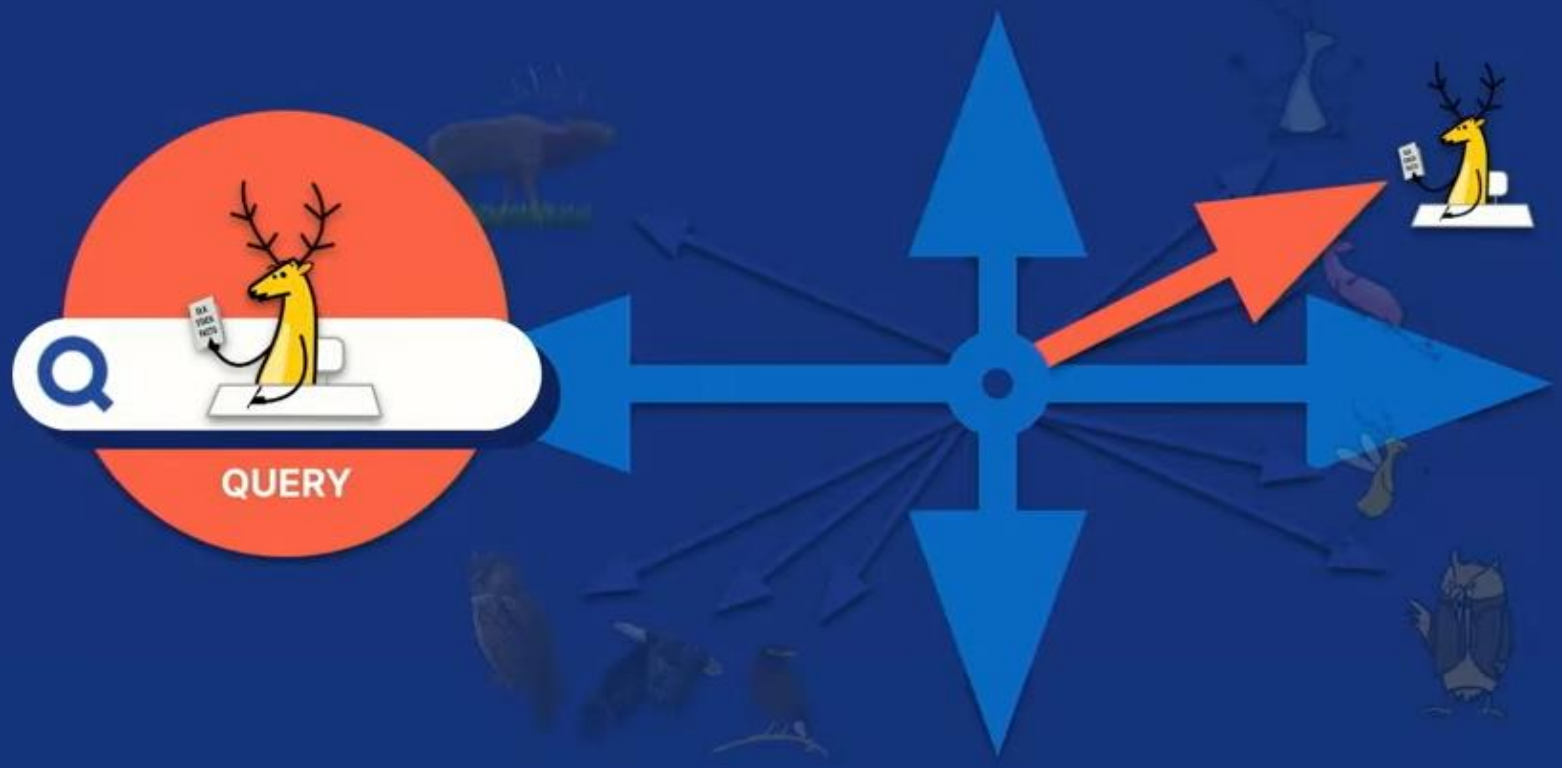


Setting up our data to support “Semantic Search”

Results given to the user

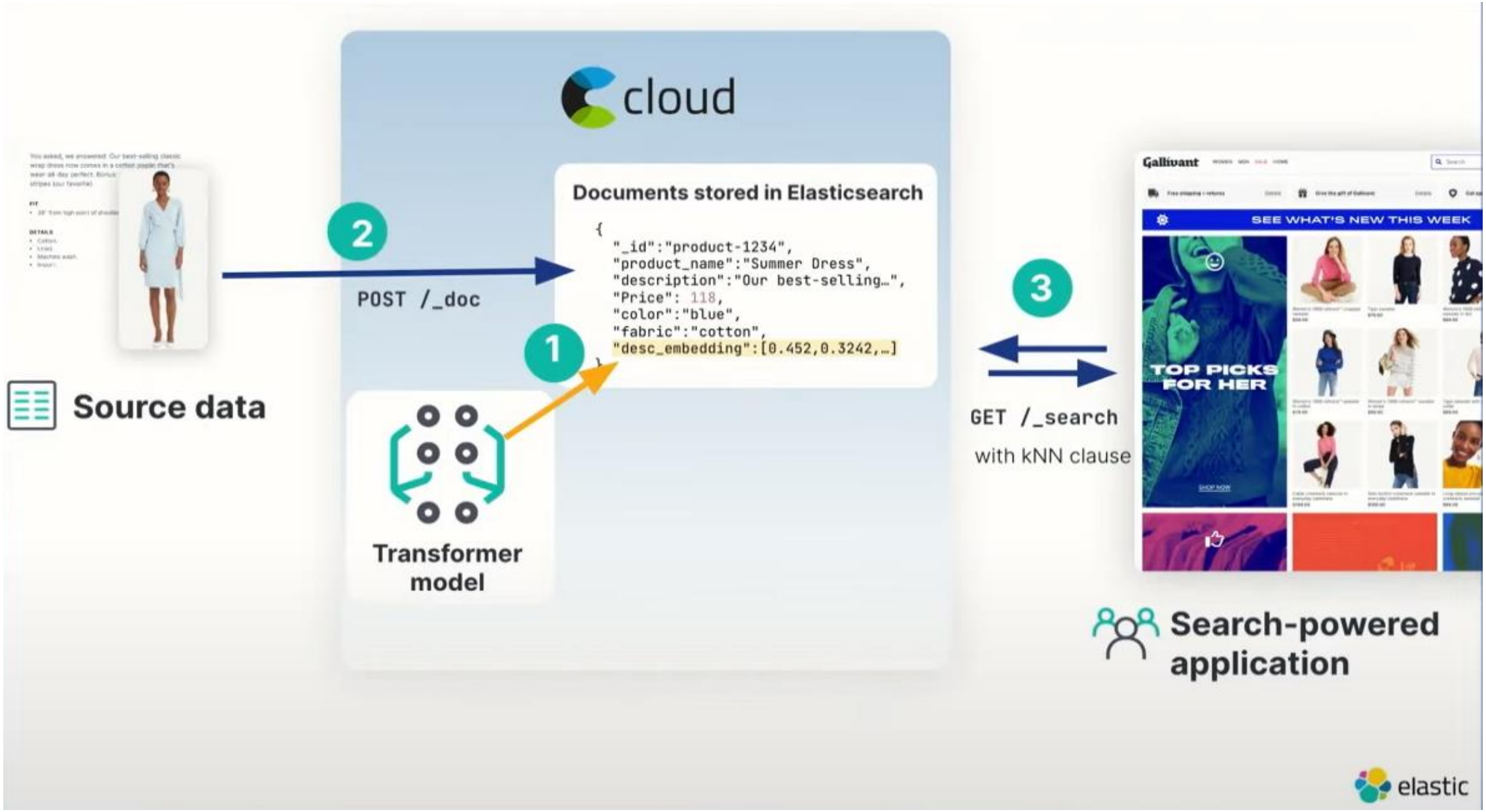
How to handle user search queries

Vector search ranks objects by similarity (relevance) to the query



Relevance	Result
Query	
1	
2	
3	
4	
5	

elastic


cloud
Documents stored in Elasticsearch
{
"_id":"product-1234",
"product_name":"Summer Dress",
"description":"Our best-selling…",
"Price": 118,
"color":"blue",
"fabric":"cotton",
"desc_embedding":[0.452,0.3242,…]
}
2
POST /_doc
1
Source data
Transformer model
3
GET /_search
with kNN clause
Gallivant
SEE WHAT'S NEW THIS WEEK
TOP PICKS FOR HER
Search-powered application
elastic

Step 1: Setting up the machine learning model


```
$ eland_import_hub_model
--url https://cluster_URL
--hub-model-id BERT-MiniLM-L6
--task-type text_embedding
--start
```

PyTorch → eland →

Select the appropriate model

Load the model to the cluster

Manage models

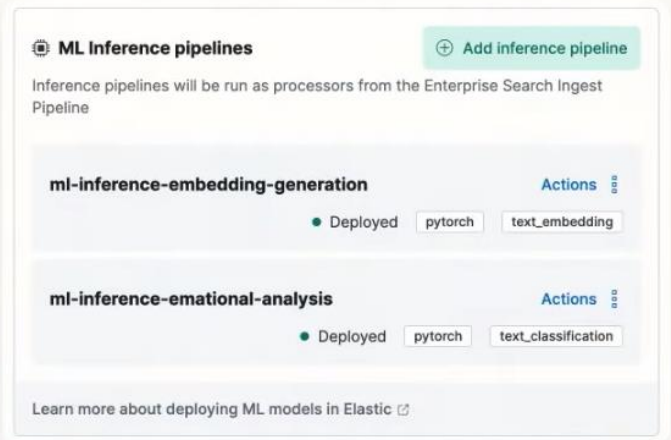
Step 2: Data ingestion and embedding generation

Source data

POST /_doc
Standard field indexing for non-vector types

```
{
  "_id":"product-1234",
  "product_name":"Summer Dress",
  "description":"Our best-selling…",
  "Price": 118,
  "color":"blue",
  "fabric":"cotton",
  "desc_embedding":[0.452,0.3242,…]
}
```

POST /_doc
Encoding via Inference Processor



elastic

Step 3: Issuing a vector query

a Query is submitted to the search-powered application


```
GET product-catalog/_search

{
  "query": {
    "match": {
      "description": {
        "query": "summer clothes",
        "boost": 0.9
      }
    }
  },
  "knn": {
    "field": "desc_embbeding",
    "query_vector": [0.123, 0.244,...],
    "k": 5,
    "num_candidates": 50,
    "boost": 0.1,
    "filter": {
      "term": {
        "department": "women"
      }
    }
  },
  "size": 10
}
```

elastic

Step 3: Issuing a vector query

a Query is submitted to the search-powered application

summer clothes

b Query embedding is generated

```
POST /_ml/trained_models/my-model/_infer

{
  "docs": {
    "description": "summer clothes"
  }
}
```

Transformer model

PyTorch

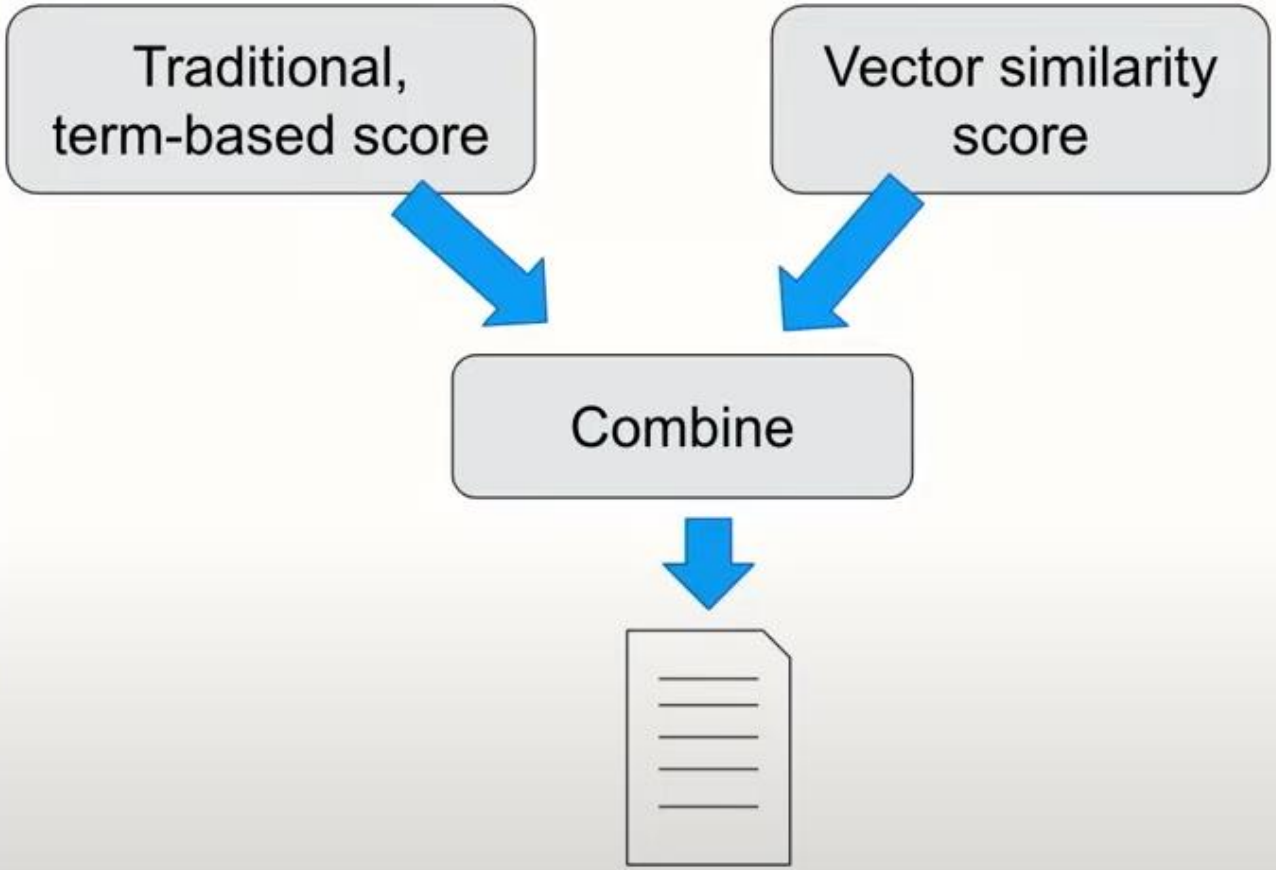
c Issue query using the _search endpoint, with a kNN clause, using the previously generated embedding

```
GET product-catalog/_search

{
  "query": {
    "match": {
      "description": {
        "query": "summer clothes",
        "boost": 0.9
      }
    }
  },
  "knn": {
    "field": "desc_embbeding",
    "query_vector": [0.123, 0.244, ...],
    "k": 5,
    "num_candidates": 50,
    "boost": 0.1,
    "filter": {
      "term": {
        "department": "women"
      }
    }
  },
  "size": 10
}
```

elastic

Hybrid scoring gets you the best of both worlds



We must choose, which search method to use depending on the use case. Some cases “term based or full-text search” will work best and for some cases “semantic search” will work best. But sometimes combining both and using a “hybrid search” will give the best results.

elastic

Actual Implementation of Hybrid Search in Python

```python
@app.post('/')
def handle_search():
    query = request.form.get('query', '')
    filters, parsed_query = extract_filters(query)
    from_ = request.form.get('from_', type=int, default=0)

    if parsed_query:
        search_query = {
            'must': {
                'multi_match': {
                    'query': parsed_query,
                    'fields': ['name', 'summary', 'content'],
                }
            }
        }
    else:
        search_query = {
            'must': {
                'match_all': {}
            }
        }

    results = es.search(
        query={
            'bool': {
                **search_query,
                **filters
            }
        },
        knn={
            'field': 'embedding',
            'query_vector': es.get_embedding(parsed_query),
            'k': 10,
            'num_candidates': 50,
            **filters,
        },
        rank={
            'rrf': {}
        },
        size=5,
        from_=from_,
    )
    return render_template('index.html', results=results['hits']['hits'],
```

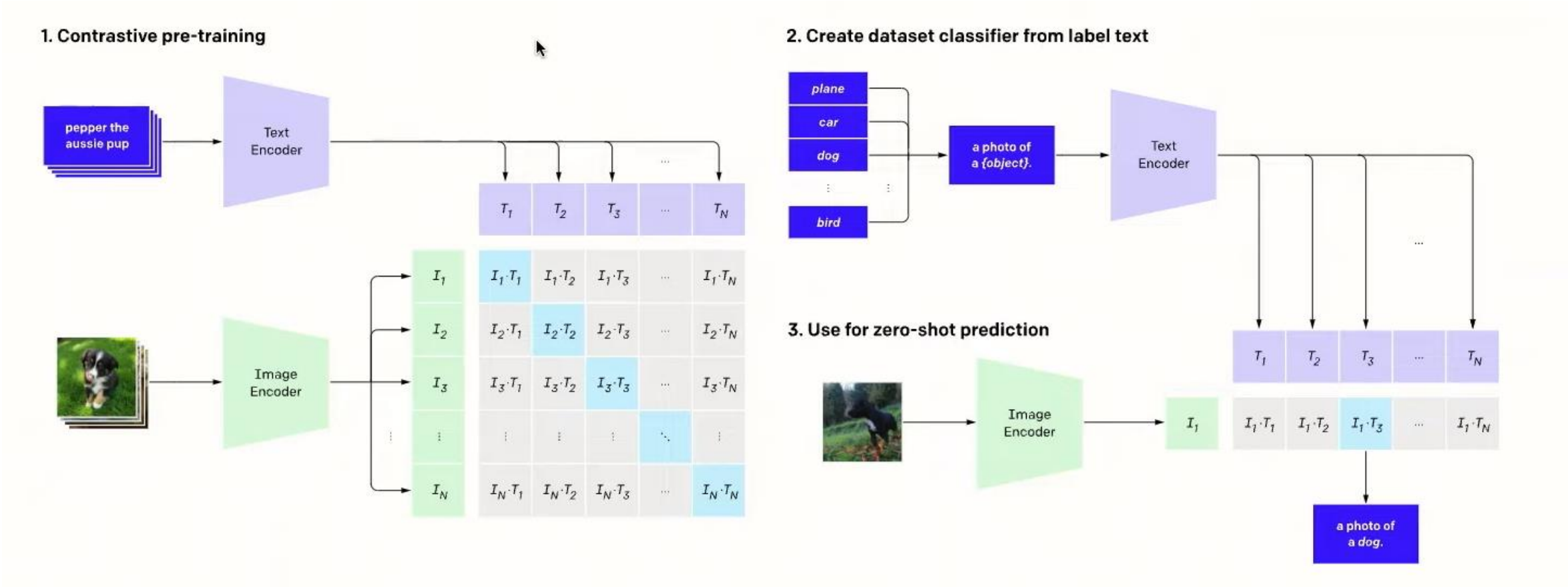
Dynamically setting up the traditional "term based or full text search query"

Then in elasticsearch search function , under "query" section we give the "term based or full text search query"

Additionally, under "knn" section we give the "vector or semantic search query"

We use elastic's "**reciprocal rank fusion**" algorithm to combine the search results of "term based or full text search" and "semantic search" based on the "relevance indicator value"

IMAGE TO TEXT using OPENAI Pre-Trained Model



How will you apply vector search?

Examples of Elastic customer projects

Product similarity search

"Do you sell black v-neck shirts that look like this?"

Answer technical support

"What are the troubleshooting steps for ___?"

Query medical knowledge

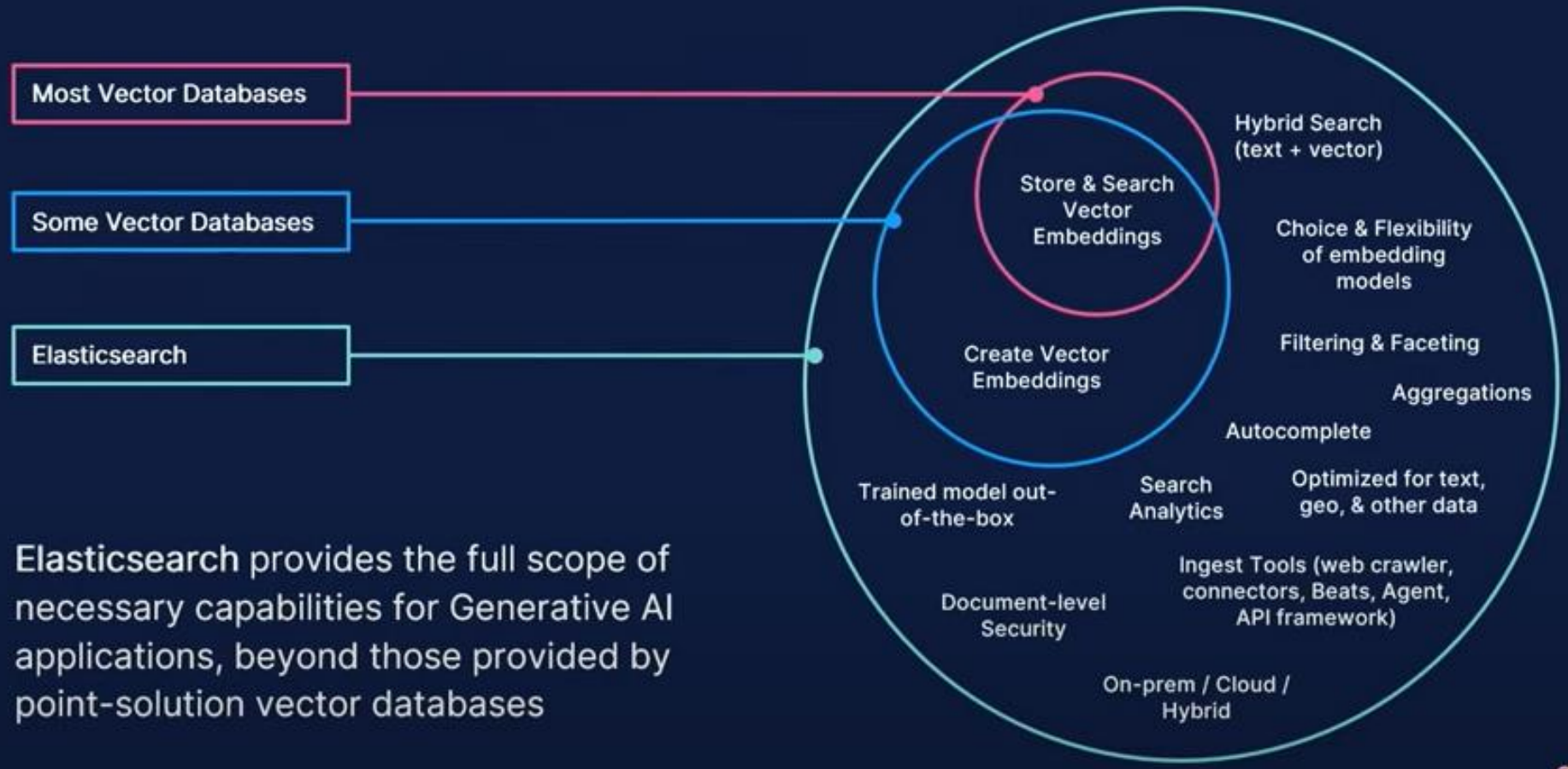
"Is lithium used to treat bipolar disorder?"

React to user sentiments

Identify poor customer interactions before they lead to escalations

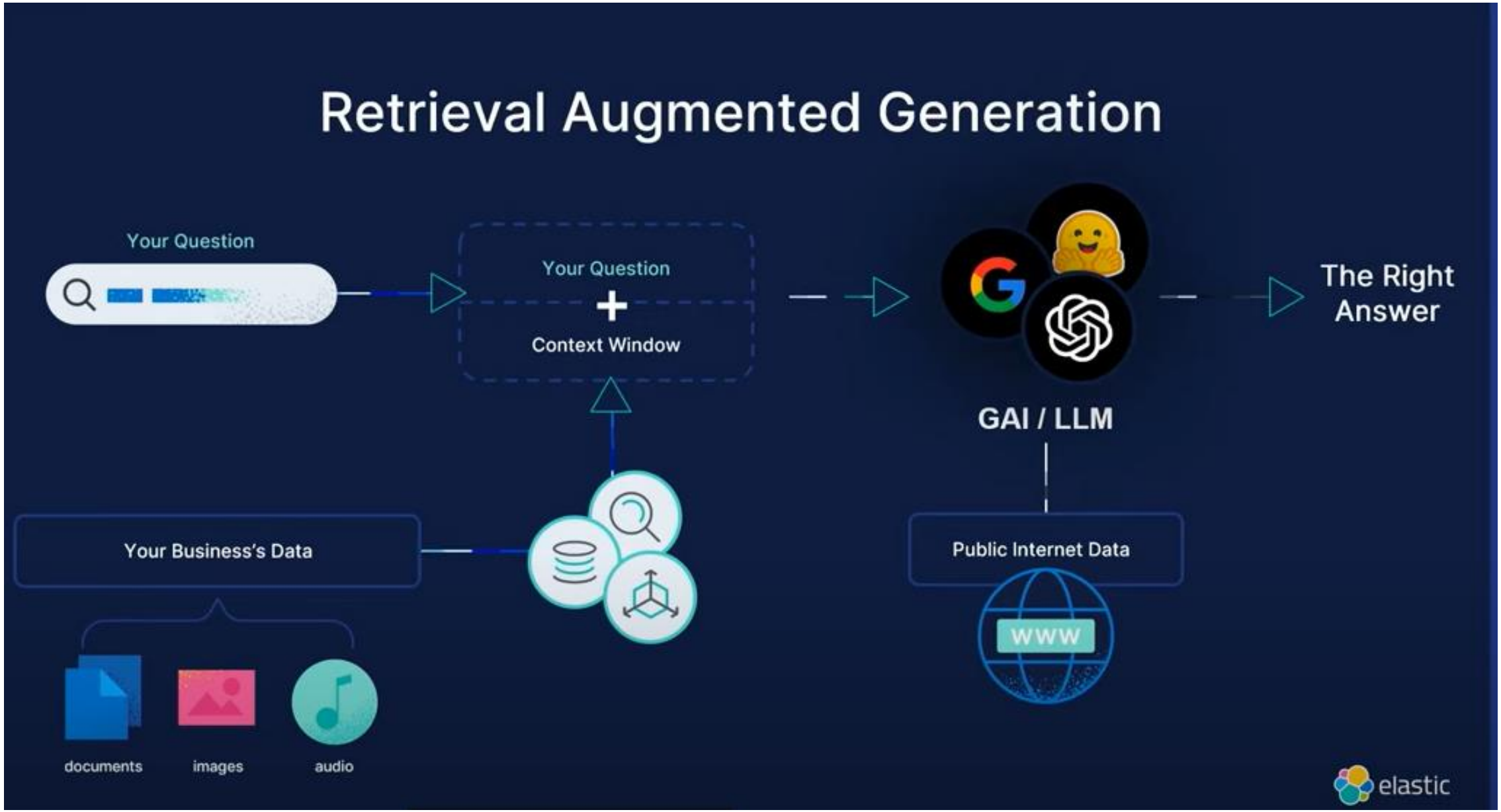
elastic

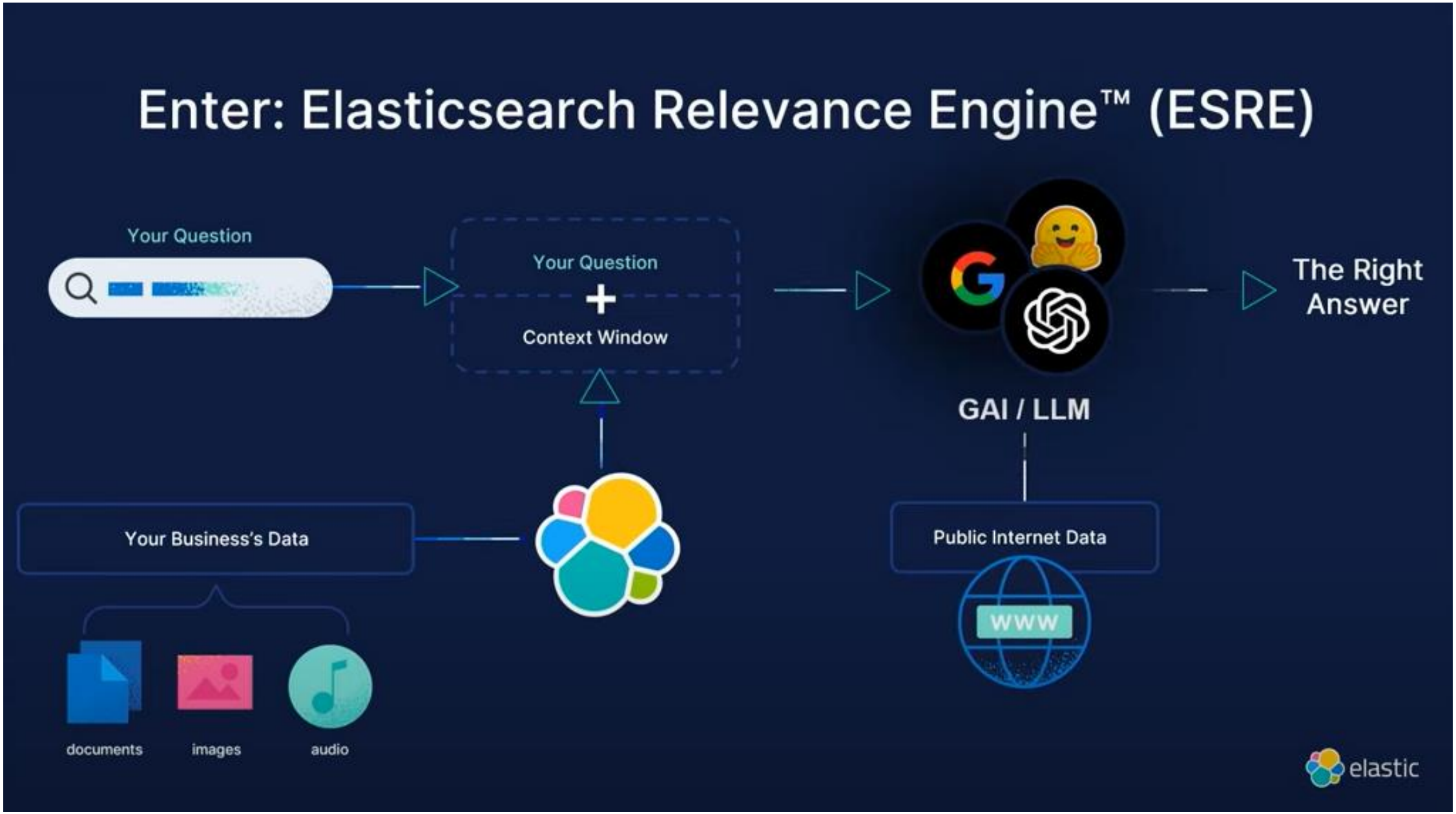
Elastic has ALL the capabilities you need



Elasticsearch provides the full scope of necessary capabilities for Generative AI applications, beyond those provided by point-solution vector databases

elastic


Retrieval Augmented Generation
Your Question
Your Question
+
Context Window
GAI / LLM
The Right Answer
Your Business's Data
Public Internet Data
WWW
documents
images
audio
elastic


Enter: Elasticsearch Relevance Engine™ (ESRE)
Your Question
Your Question
+
Context Window
GAI / LLM
The Right Answer
Your Business's Data
Public Internet Data
WWW
documents
images
audio
elastic

Imagine all the different use cases

HR helpdesk

What are key aspects of the company's 401k policy for an employee in my location? How do I enroll?

Customer Success

Are customers in San Francisco buying products with the deepest discounts? Which promotions are doing well?

Ecommerce

What material list and tools do I need to build an irrigation system for 1 acre back yard in Gilroy, CA?

Regulatory Compliance

Which recent transactions are not compliant with upcoming 2025 regulatory requirement X?

Corporate Finance

Summarize changes to revenue recognition and expense categorization for FY 2024

Predictive Maintenance

Based on sensor data and recent customer reviews, when do I schedule repair unit Y?

they all will require proprietary information

elastic

Thank You!