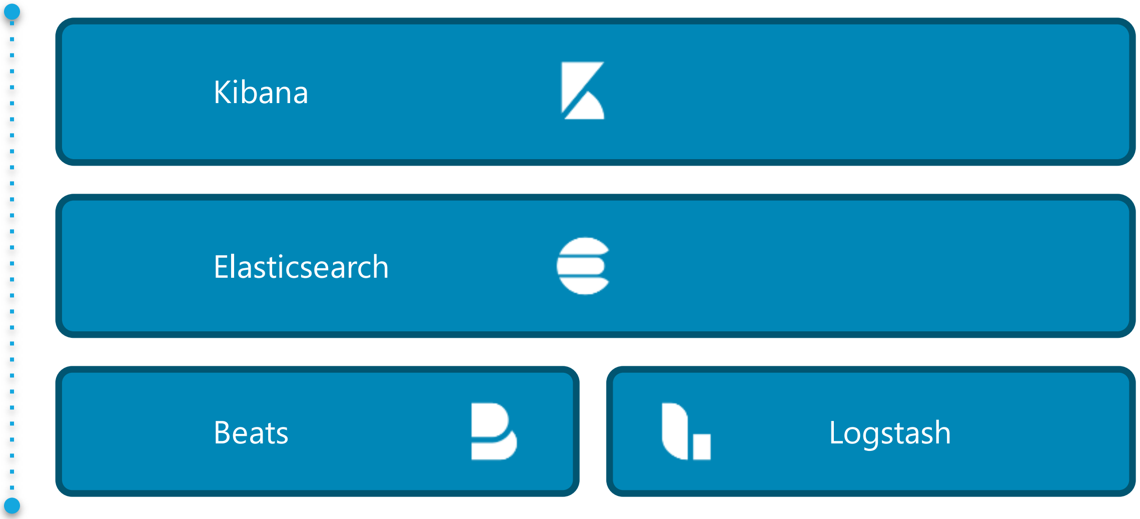
# Elastic Stack Reference Architecture

**Elastic**
**www.elastic.co**

# Elastic Stack

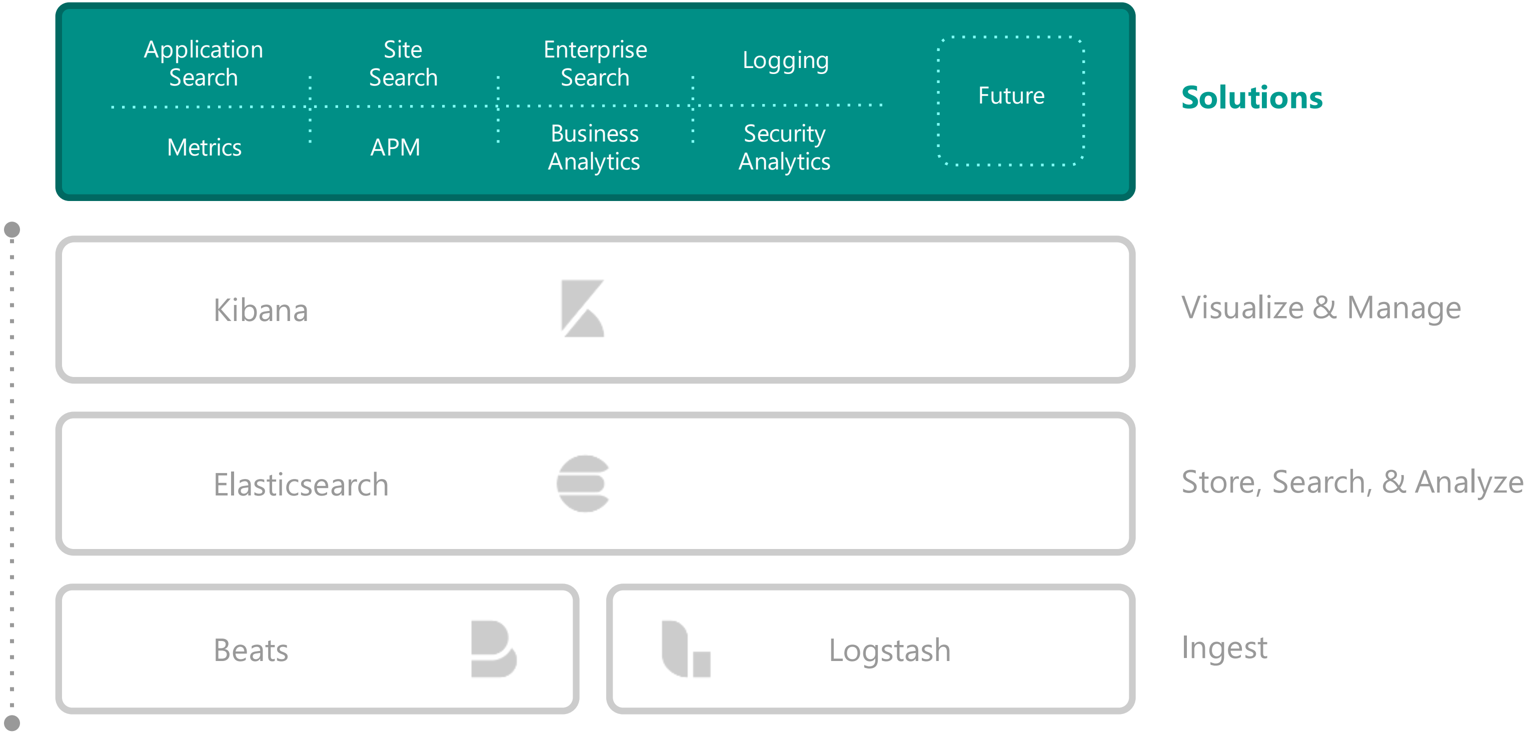


Kibana — Visualize & Manage

Elasticsearch — Store, Search, & Analyze

Beats, Logstash — Ingest

# Elastic Stack



Application Search

Site Search

Enterprise Search

Logging

Metrics

APM

Business Analytics

Security Analytics

Future

**Solutions**

Kibana

Visualize & Manage

Elasticsearch

Store, Search, & Analyze

Beats

Logstash

Ingest

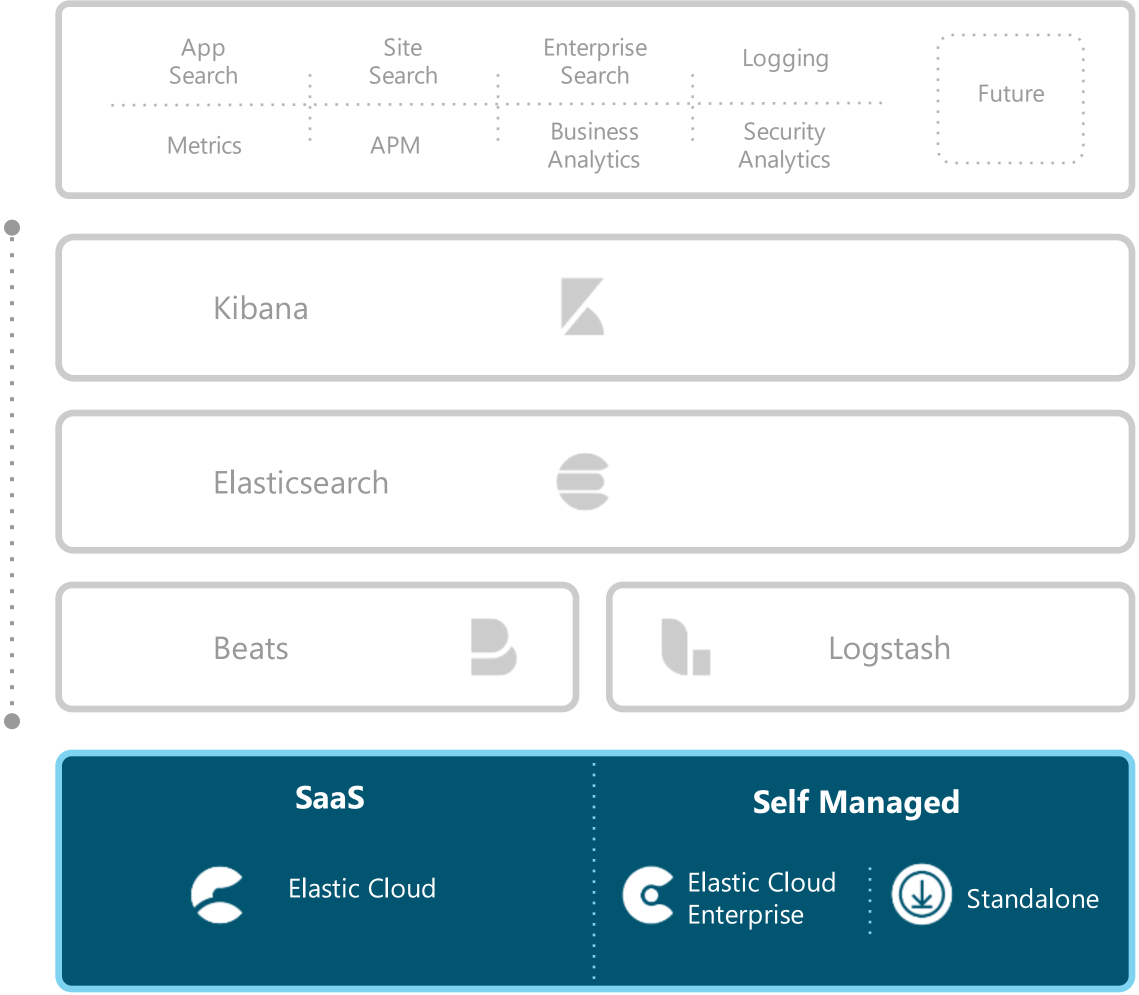
# Solutions

APM

Site Search

App Search

# Elastic Stack



App Search

Site Search

Enterprise Search

Logging

Metrics

APM

Business Analytics

Security Analytics

Future

**Solutions**

Kibana

Visualize & Manage

Elasticsearch

Store, Search, & Analyze

Beats

Logstash

Ingest

**SaaS**

Elastic Cloud

**Self Managed**

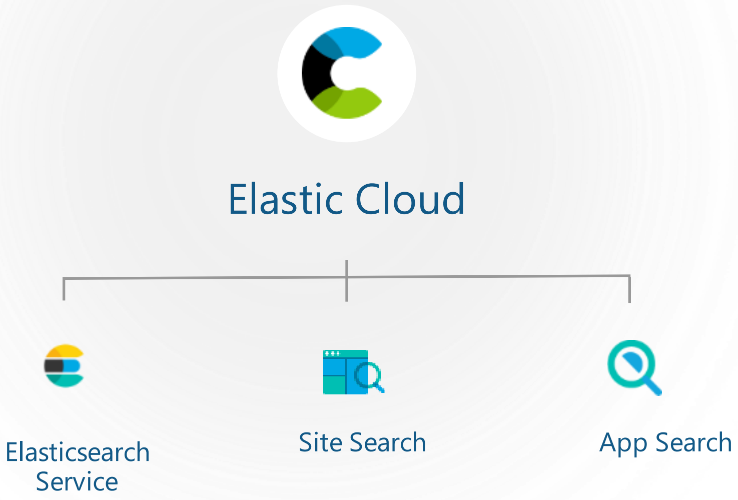
Elastic Cloud Enterprise

Standalone

**Deployment**

elastic

## SaaS



Elastic Cloud

Elasticsearch Service

Site Search

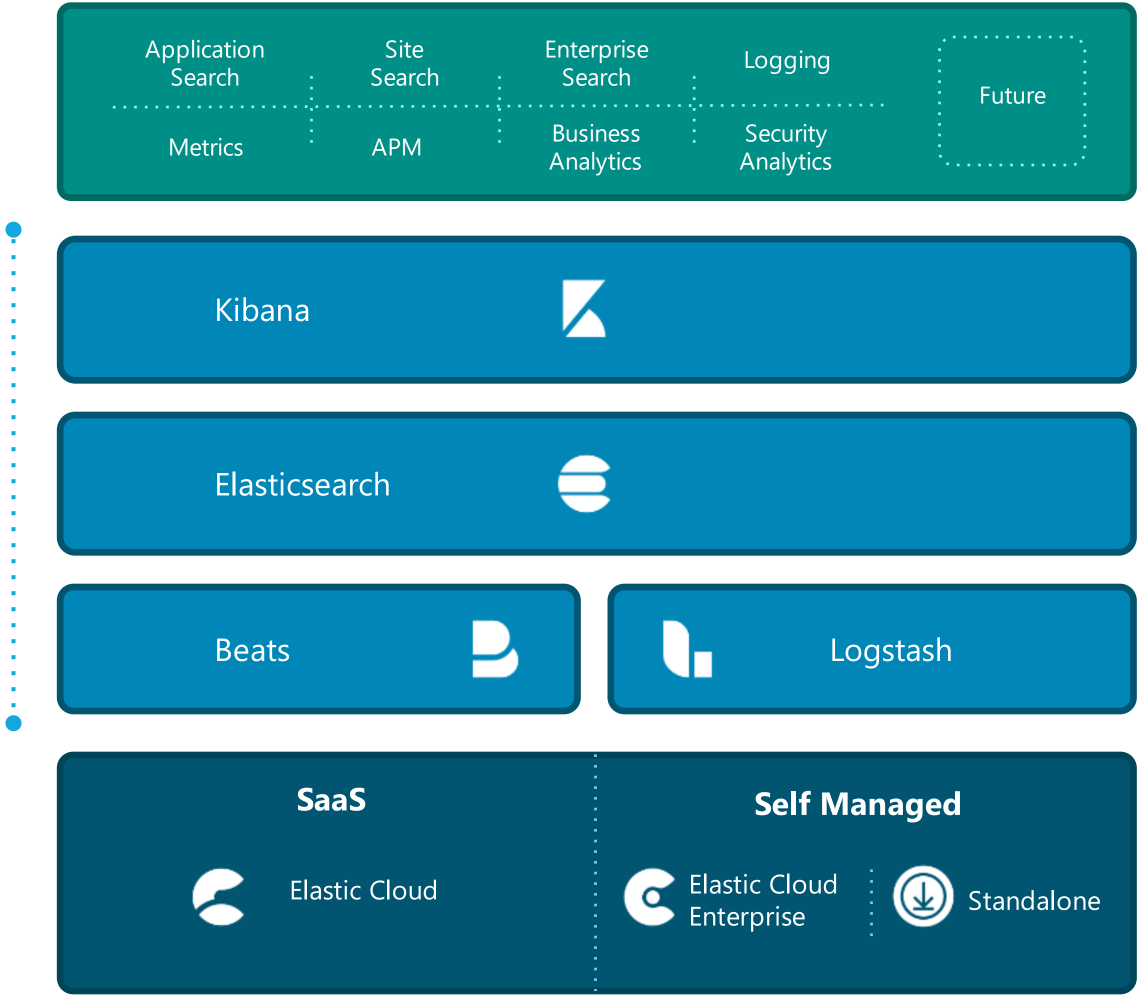
App Search

## Self Managed

Elastic Cloud Enterprise

Standalone

Elastic Stack



Application Search

Site Search

Enterprise Search

Logging

Metrics

APM

Business Analytics

Security Analytics

Future

**Solutions**

Kibana

Visualize & Manage

Elasticsearch

Store, Search, & Analyze

Beats

Logstash

Ingest

**SaaS**

Elastic Cloud

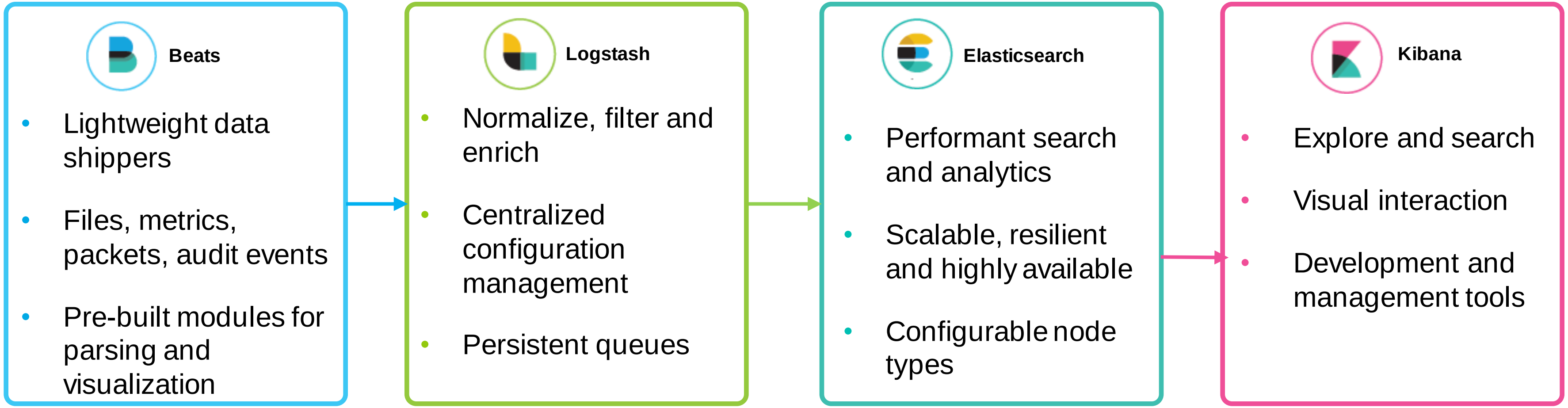
**Self Managed**

Elastic Cloud Enterprise

Standalone

**Deployment**

# Logical Processing Pipeline



**Beats**

- Lightweight data shippers
- Files, metrics, packets, audit events
- Pre-built modules for parsing and visualization

**Logstash**

- Normalize, filter and enrich
- Centralized configuration management
- Persistent queues

**Elasticsearch**

- Performant search and analytics
- Scalable, resilient and highly available
- Configurable node types

**Kibana**

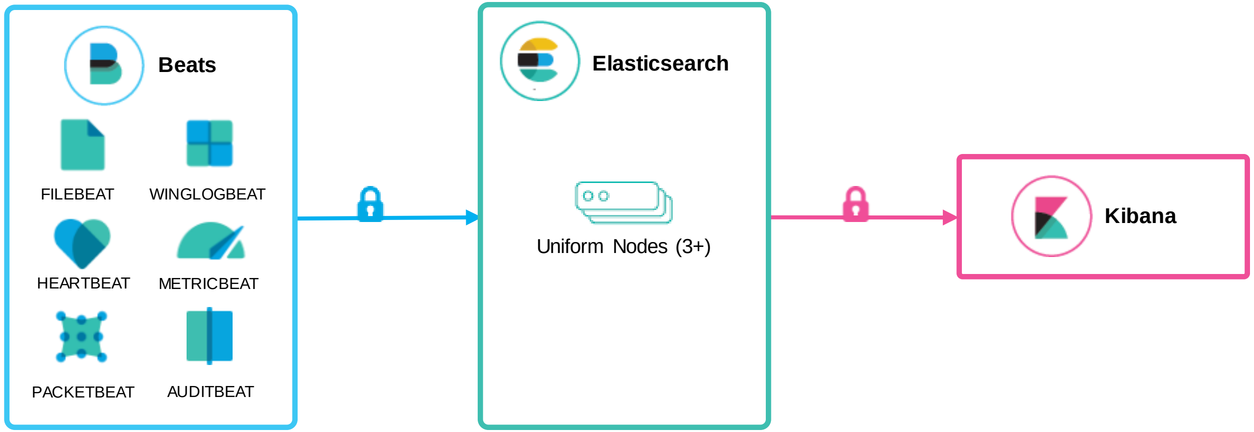
- Explore and search
- Visual interaction
- Development and management tools

All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Logging Architecture

# Basic Processing Pipeline

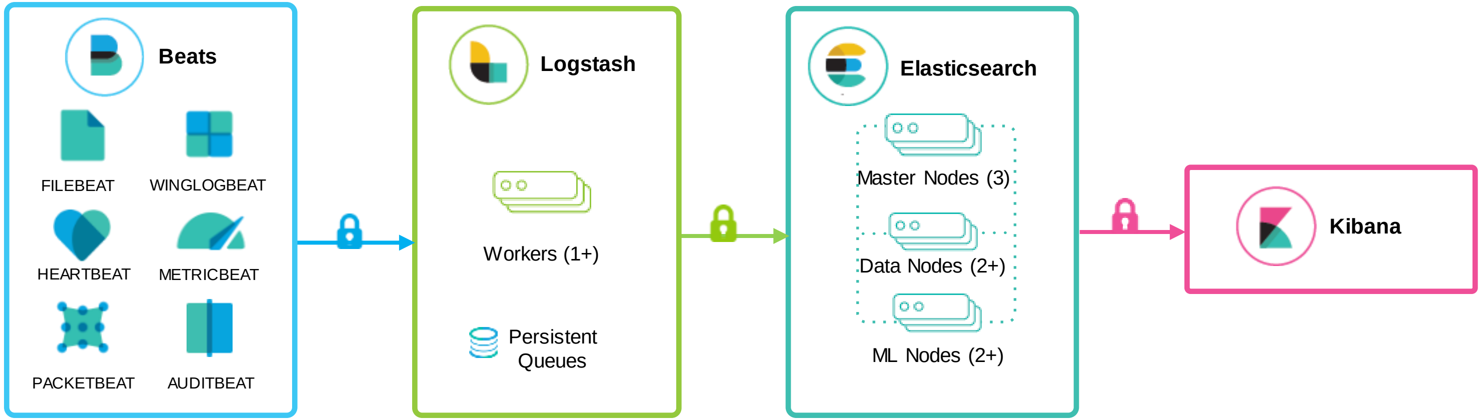
*Beats, Elasticsearch w/ Ingest Node Pipelines and Kibana*



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

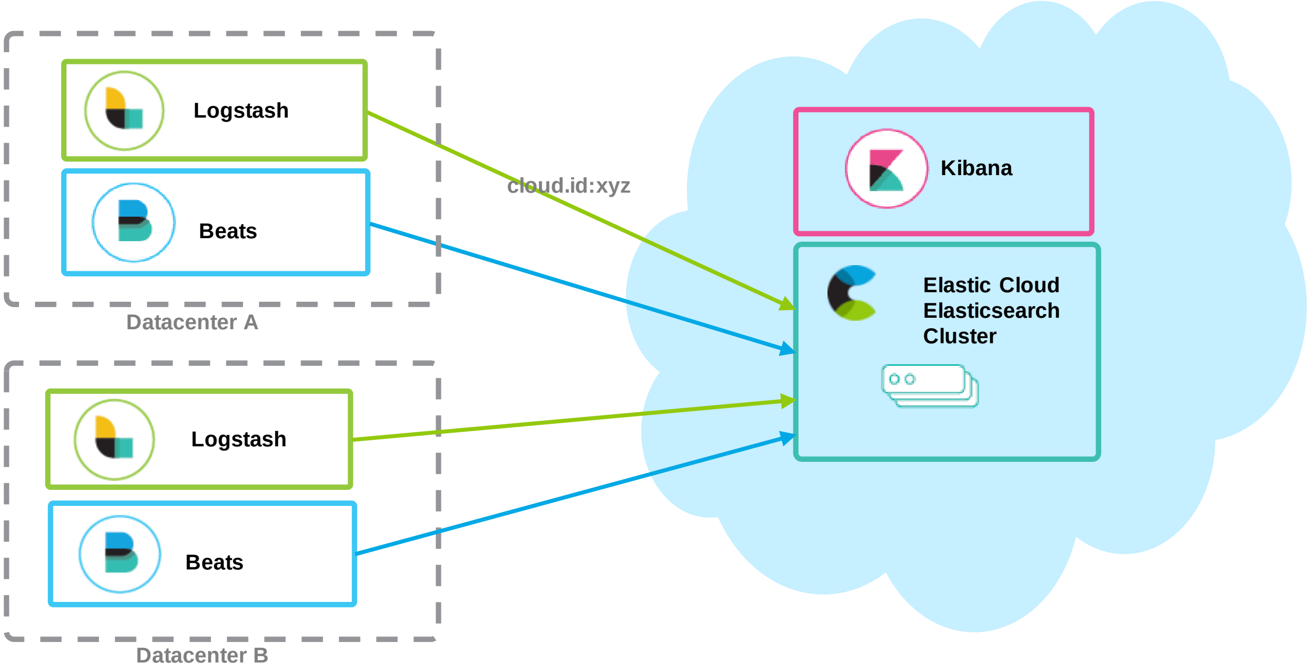
# Advanced Processing and Resiliency

*Adding Logstash processing, differentiated Elasticsearch node types*



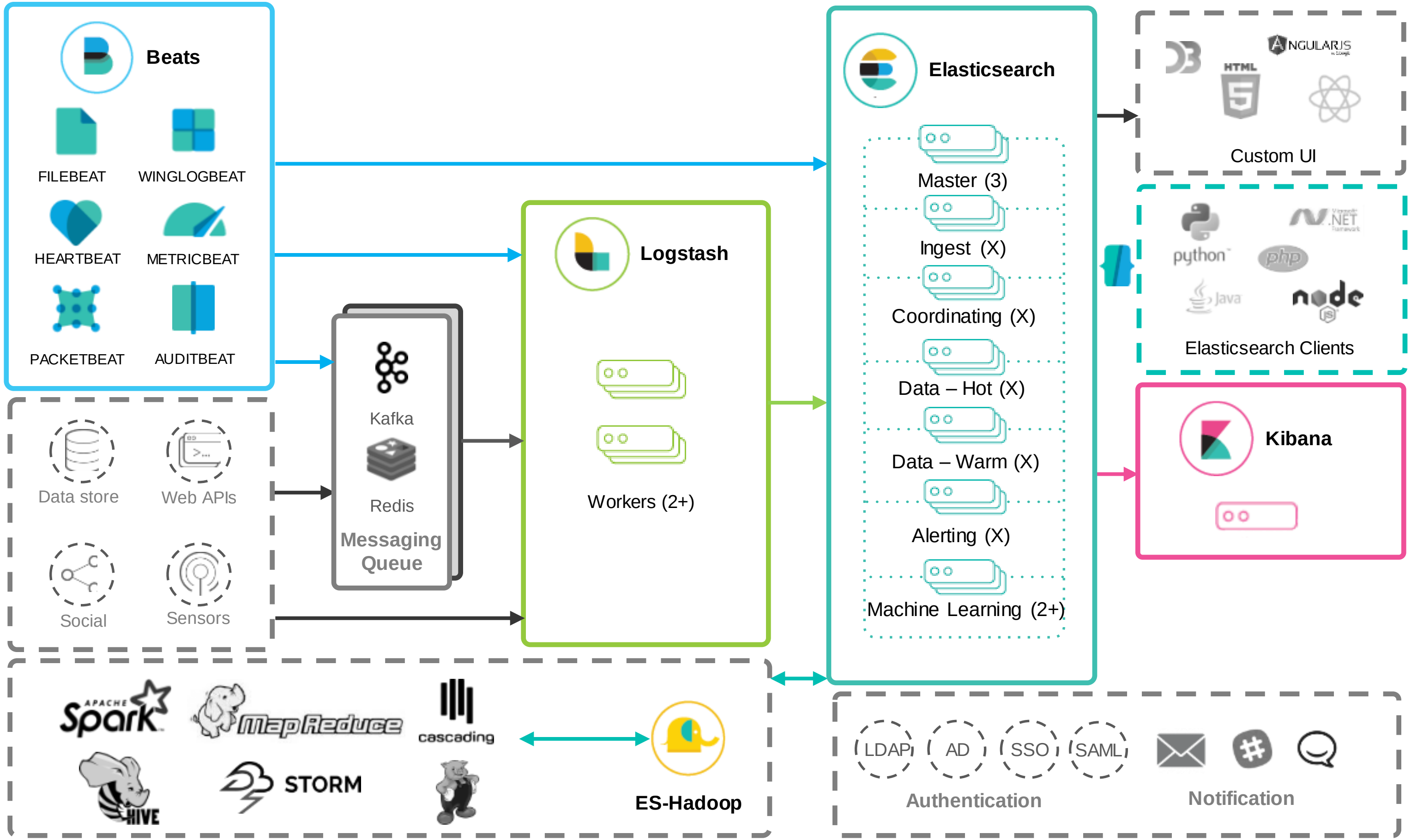
# Elastic Cloud Based Logging

*Capturing events securely to Elastic Cloud with easy configuration*



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

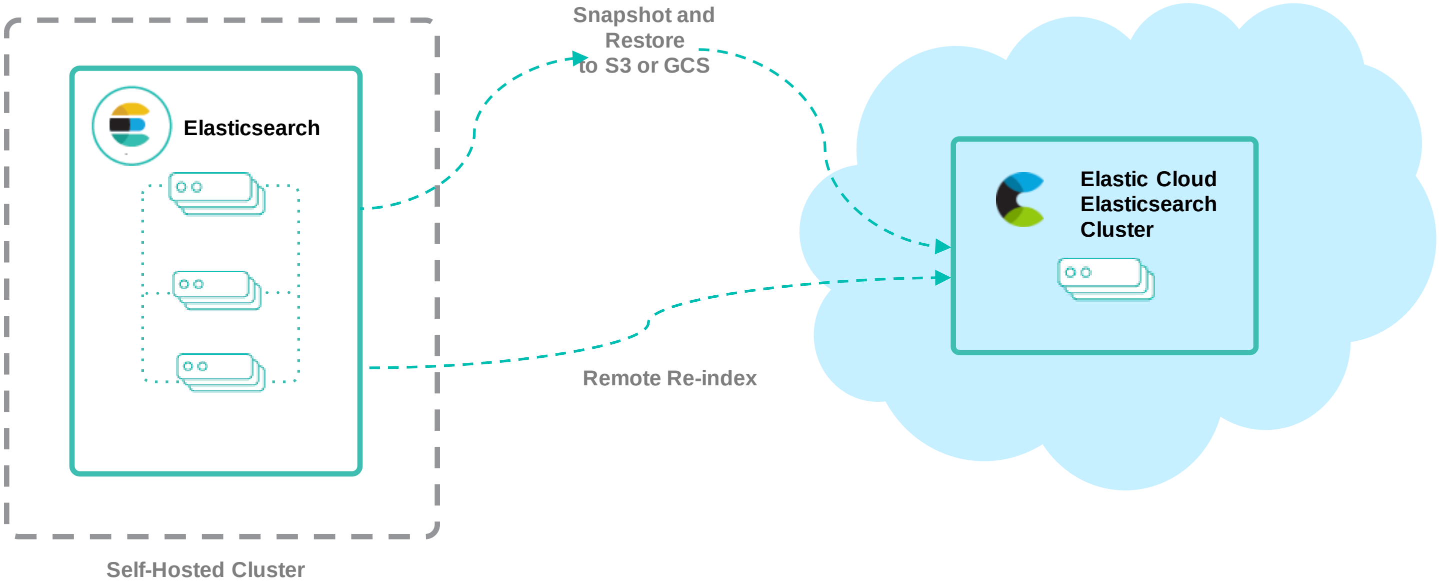
# Deployment in the Enterprise



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Options for Migrating from Self-Managed to Elastic Cloud

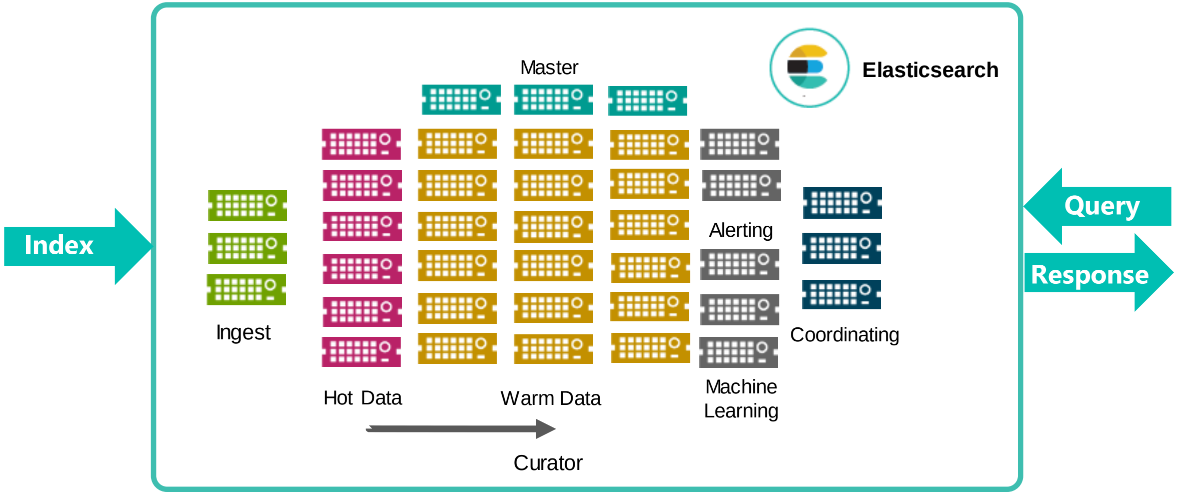
*Multiple options for easily migrating to Elastic Cloud from self-managed*


Elasticsearch
Snapshot and Restore to S3 or GCS
Elastic Cloud Elasticsearch Cluster
Remote Re-index
Self-Hosted Cluster

All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Inside a Large Elasticsearch Logging Cluster

*Reduce infrastructure costs, isolate workloads, and manage data lifecycle*

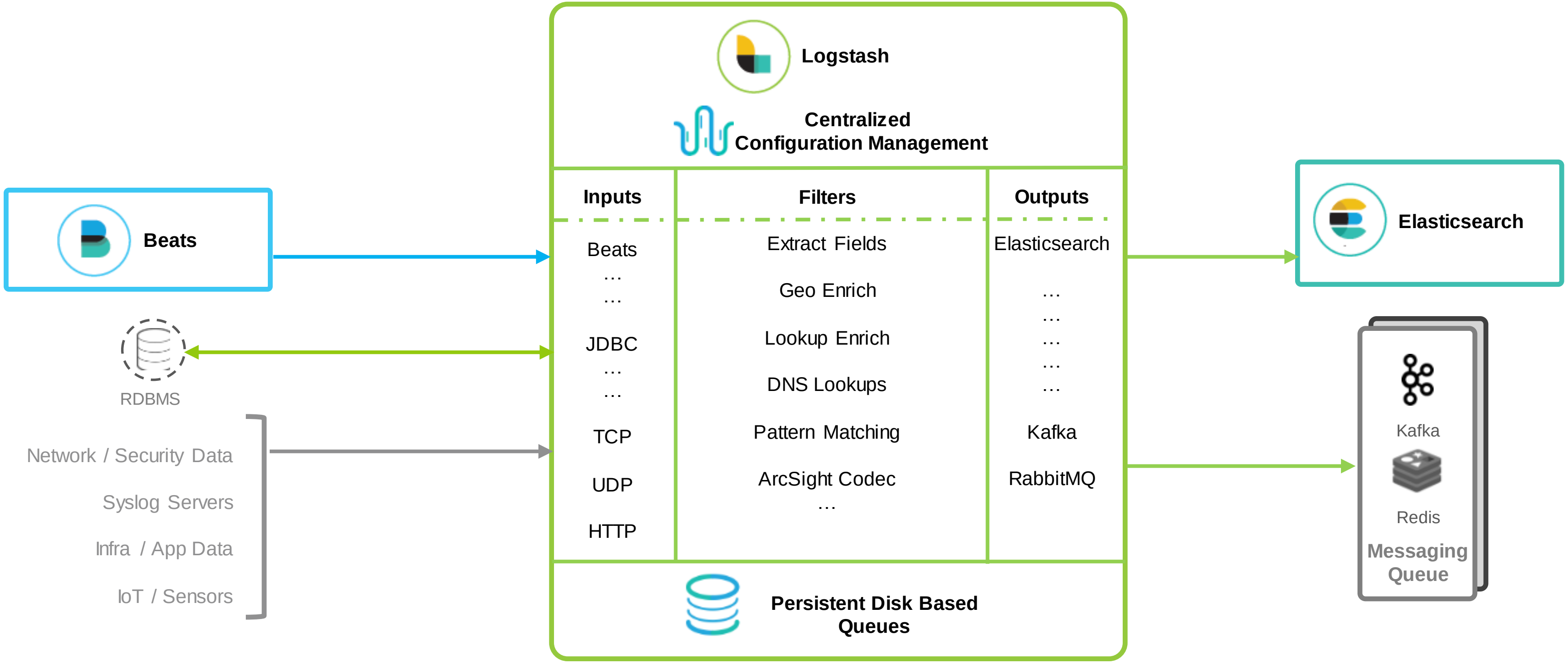


All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Variants of Logging Architectures

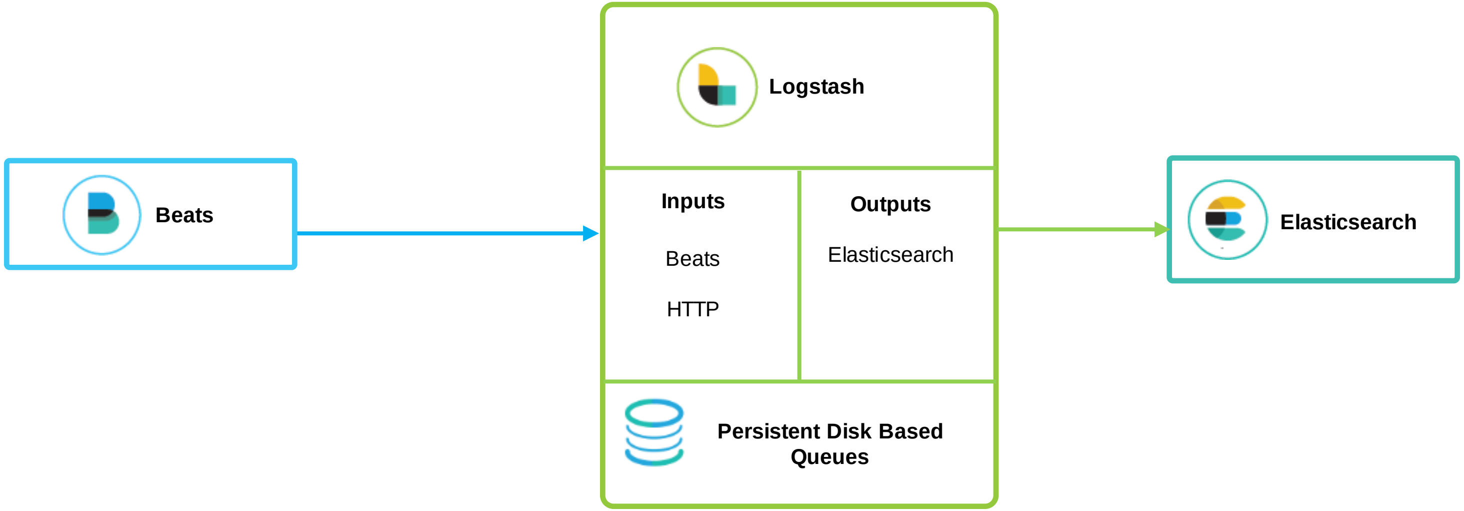
# Inside Logstash

*Example Inputs, Filters and Outputs*



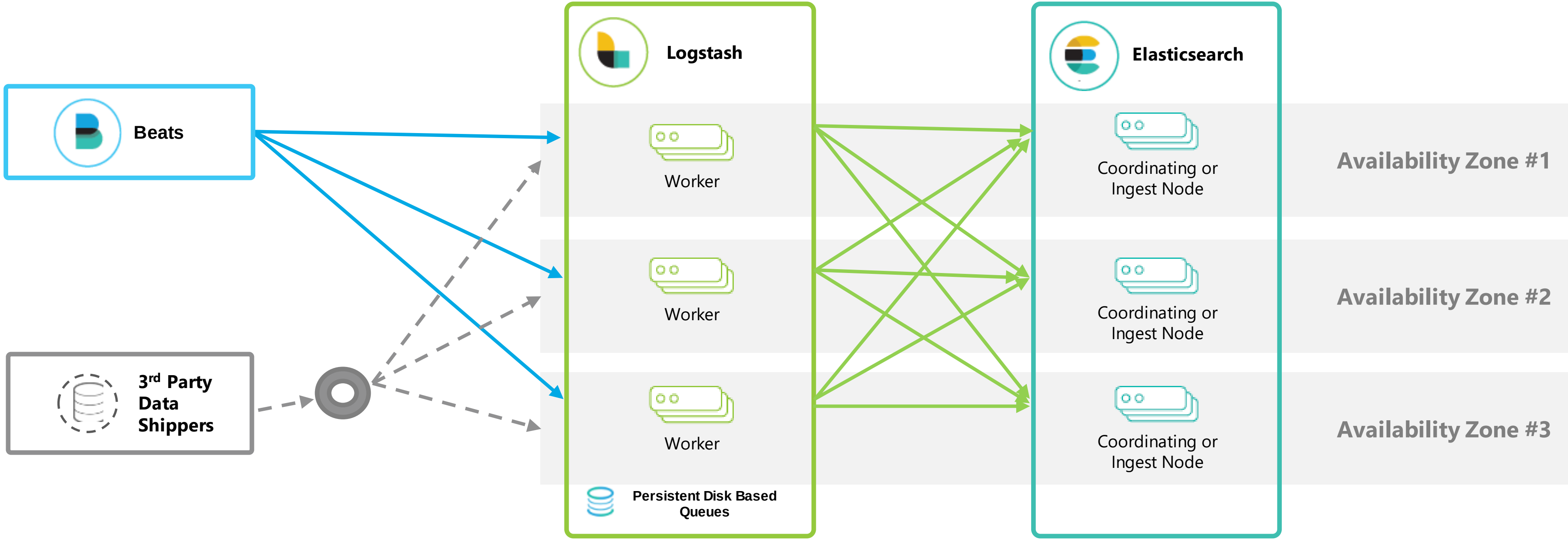
# At-least-once Delivery with Logstash

*Using inputs and outputs with acknowledgement capability*



# Highly Available, Persistent Logstash without an MQ

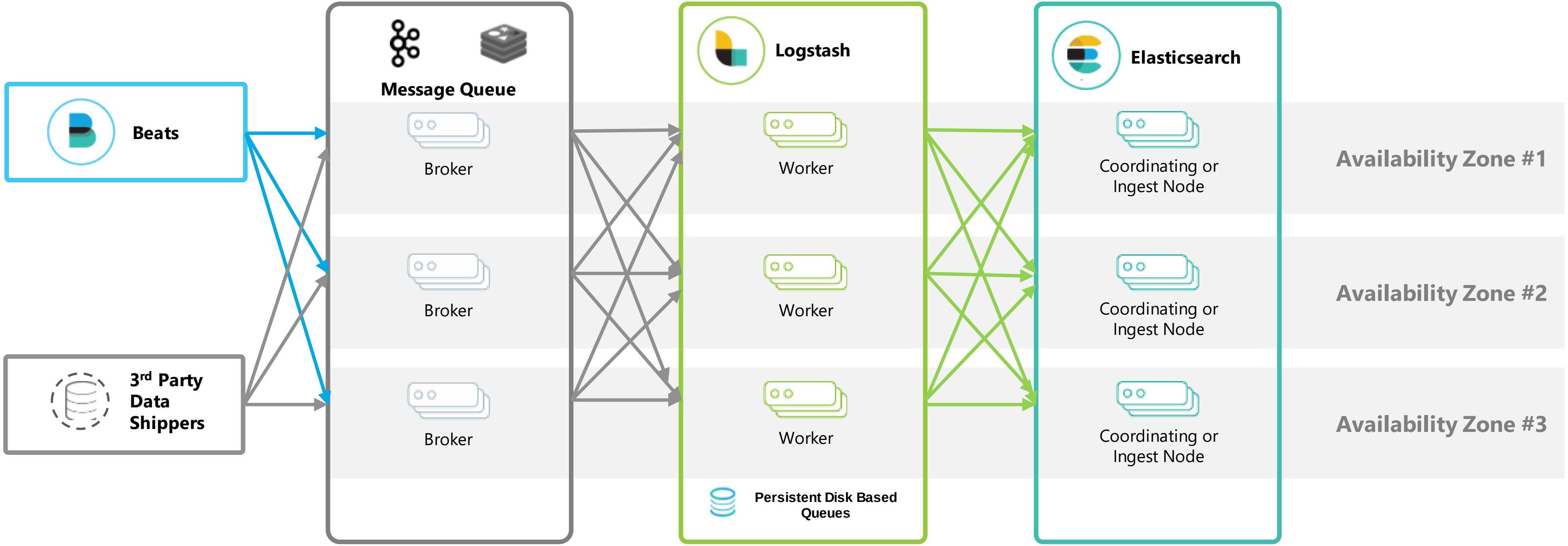
*Round-robin with persistent Logstash queues responds to backpressure and continues operation with AZ interruption.*



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Highly Available, Persistent Logstash with an MQ

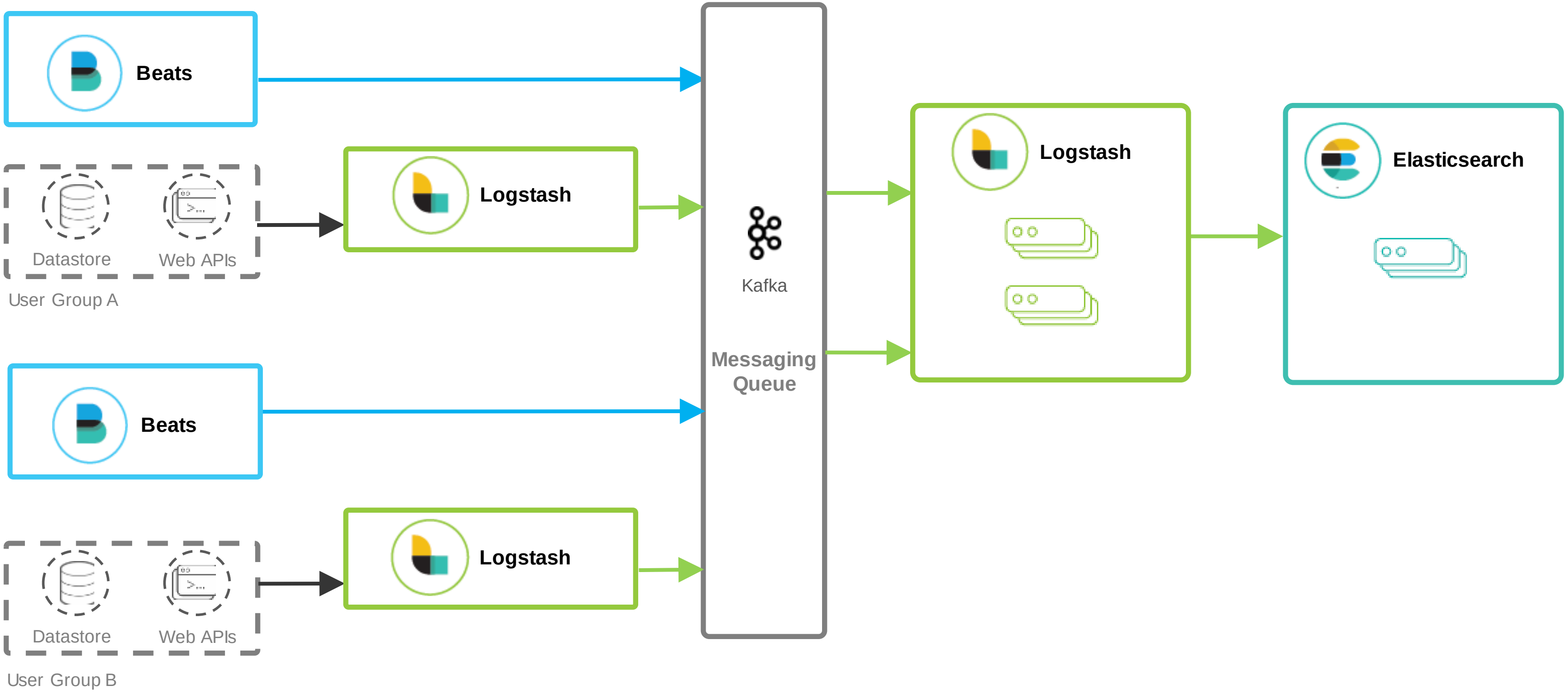
*Adding a message queue allows for replay and queue replication, thereby avoiding message loss in the case of a unrecoverable Logstash disk failure.*



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Control of Logstash Configuration by Group

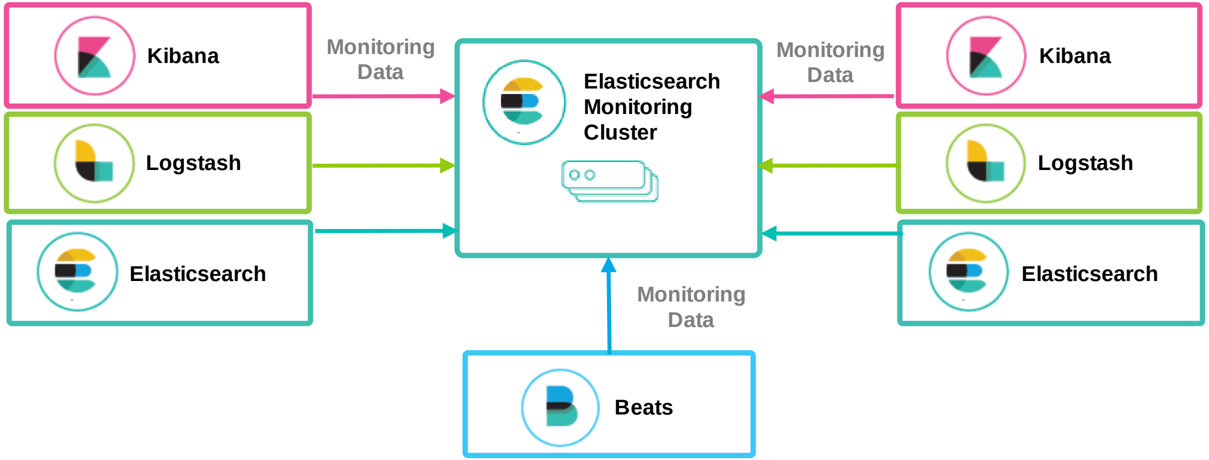
*Distribute Logstash configuration management while controlling central pipeline*



# Deployment Best Practices

# Centralized Monitoring Cluster

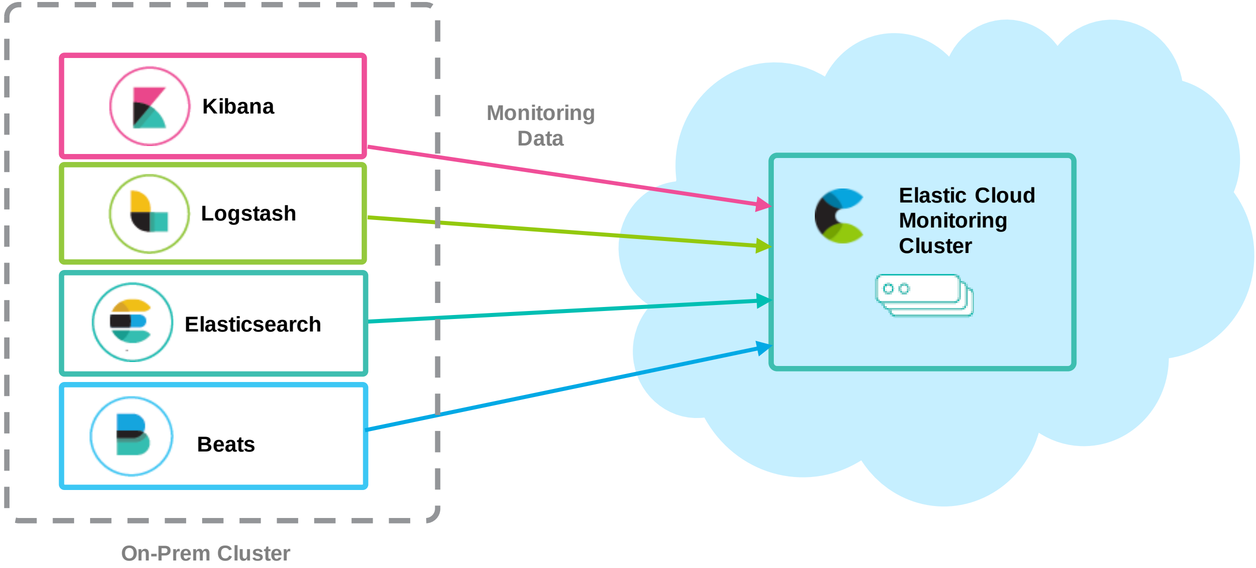
*Maintain isolated monitoring cluster for monitoring workload isolation*



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Cloud Monitoring Cluster

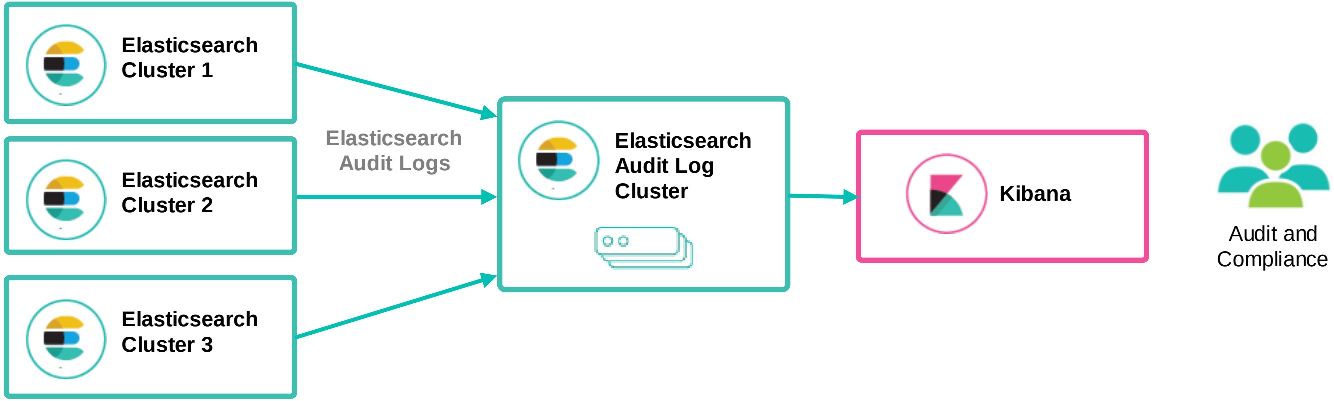
*Opt-in Elastic Cloud cluster for monitoring on-premise stack*



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Isolated Audit Logging Cluster

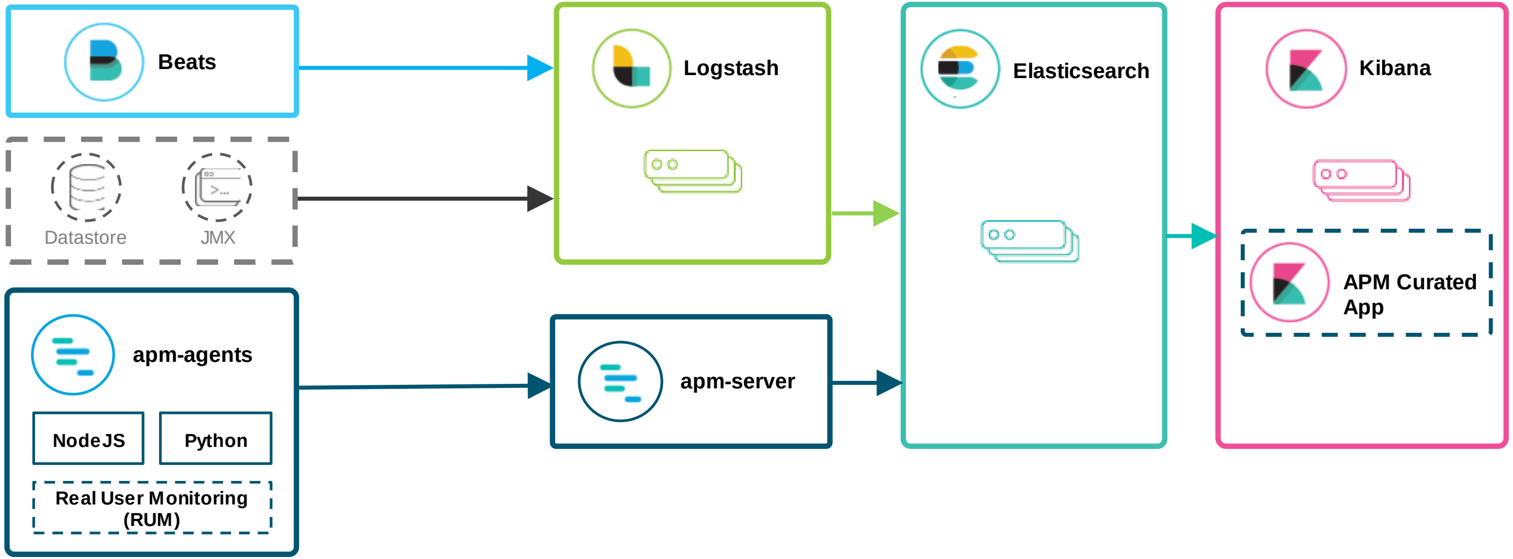
*Maintain isolated audit logging cluster for increased security and compliance*



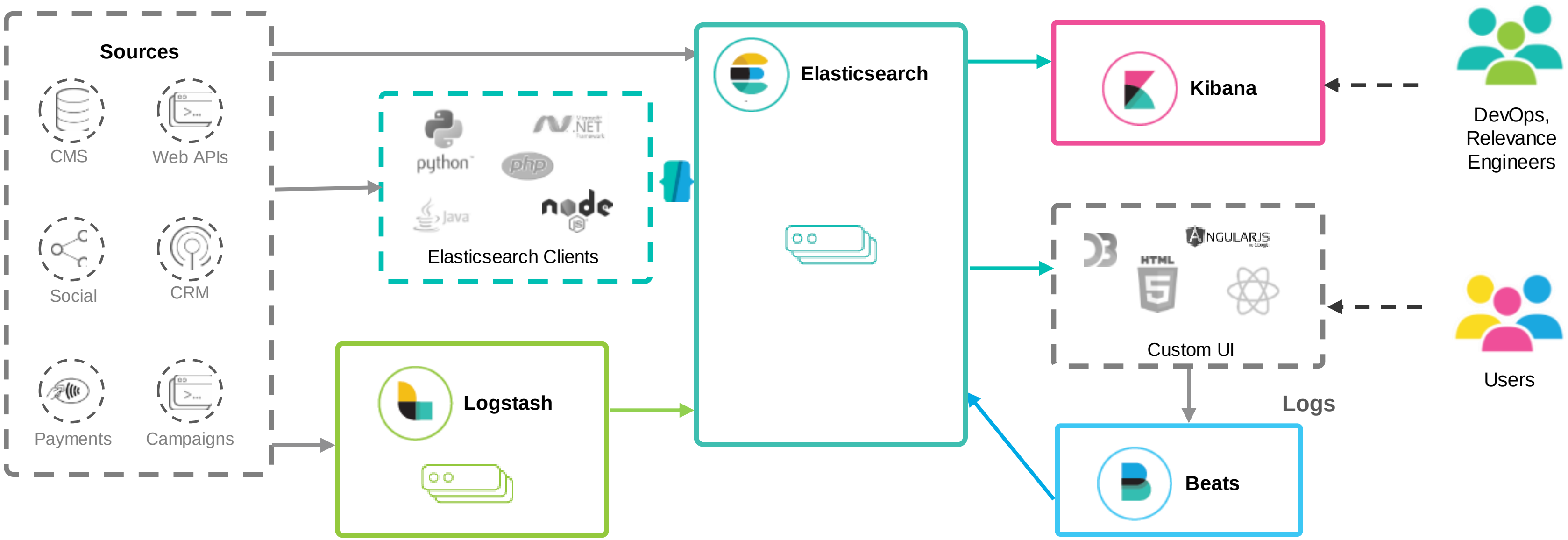
All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Use-Case Architectures

# Application Metric Collection with Elastic APM



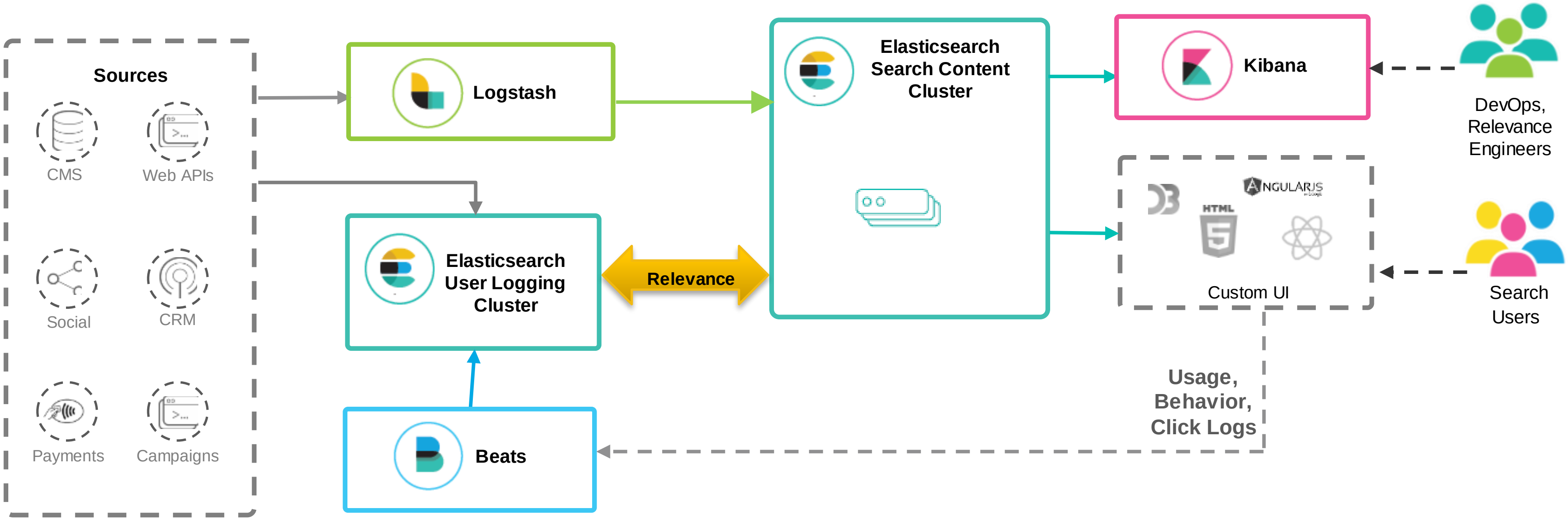
# Application or Enterprise Search: Compact



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

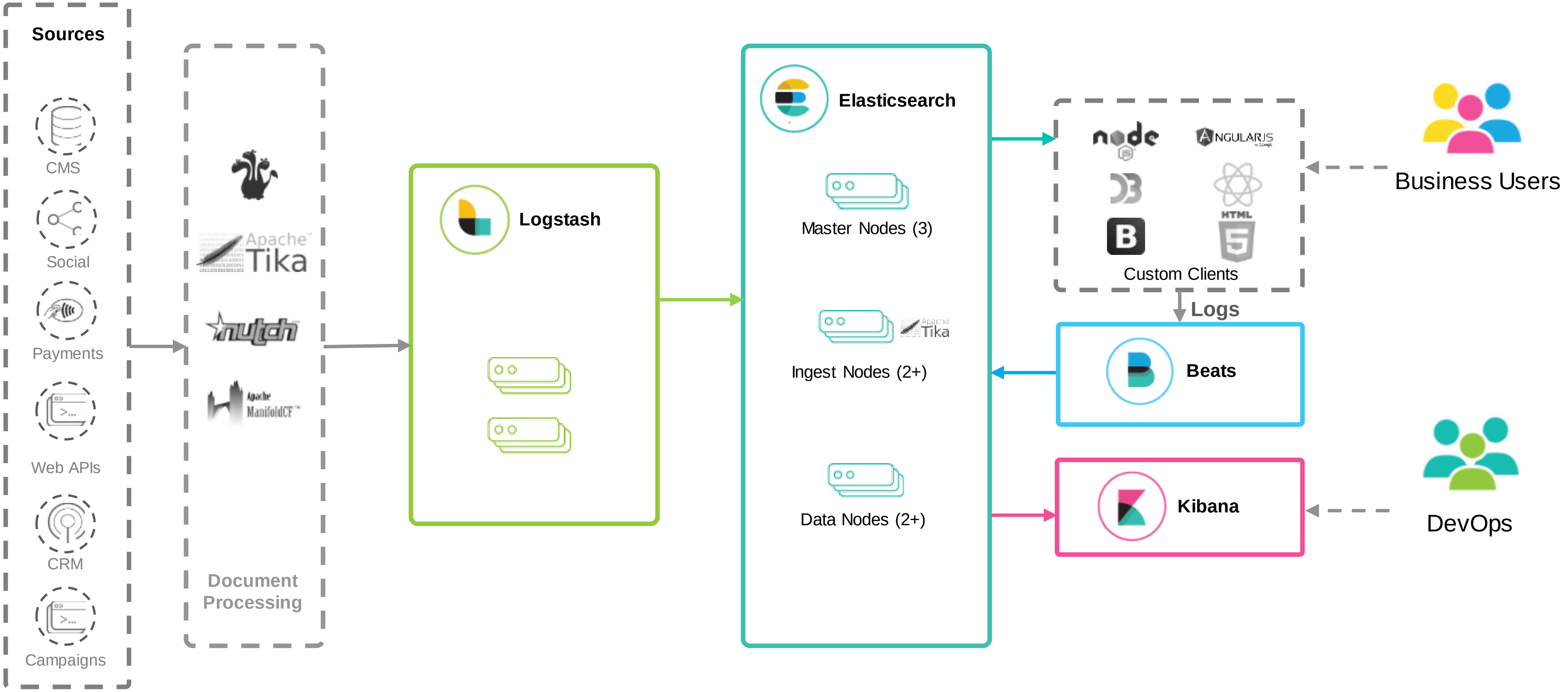
# Application or Knowledge Search: Scaled

*A dedicated cluster for logs/intelligence*


Sources
CMS
Web APIs
Social
CRM
Payments
Campaigns
Logstash
Elasticsearch Search Content Cluster
Kibana
DevOps, Relevance Engineers
Elasticsearch User Logging Cluster
Relevance
ANGULARJS
HTML 5
Custom UI
Search Users
Usage, Behavior, Click Logs
Beats

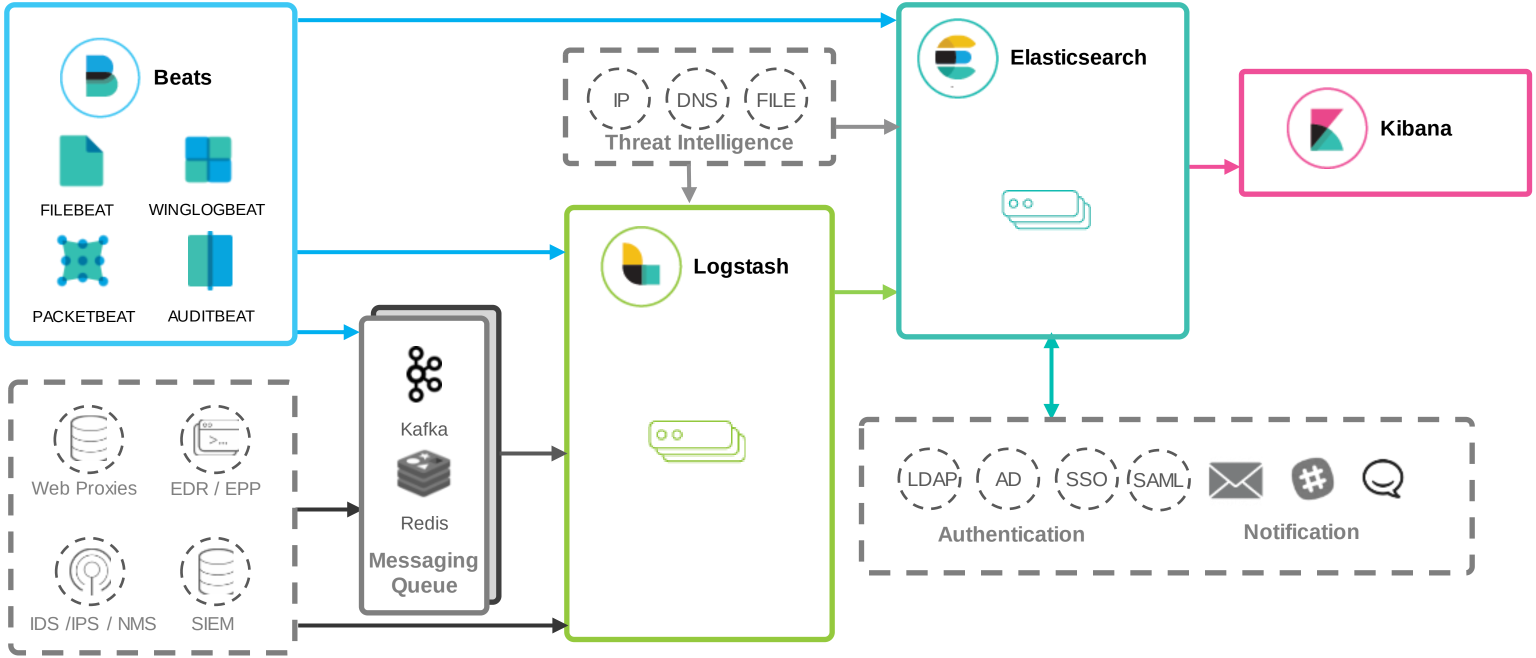
All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Enterprise Search: Scaled



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

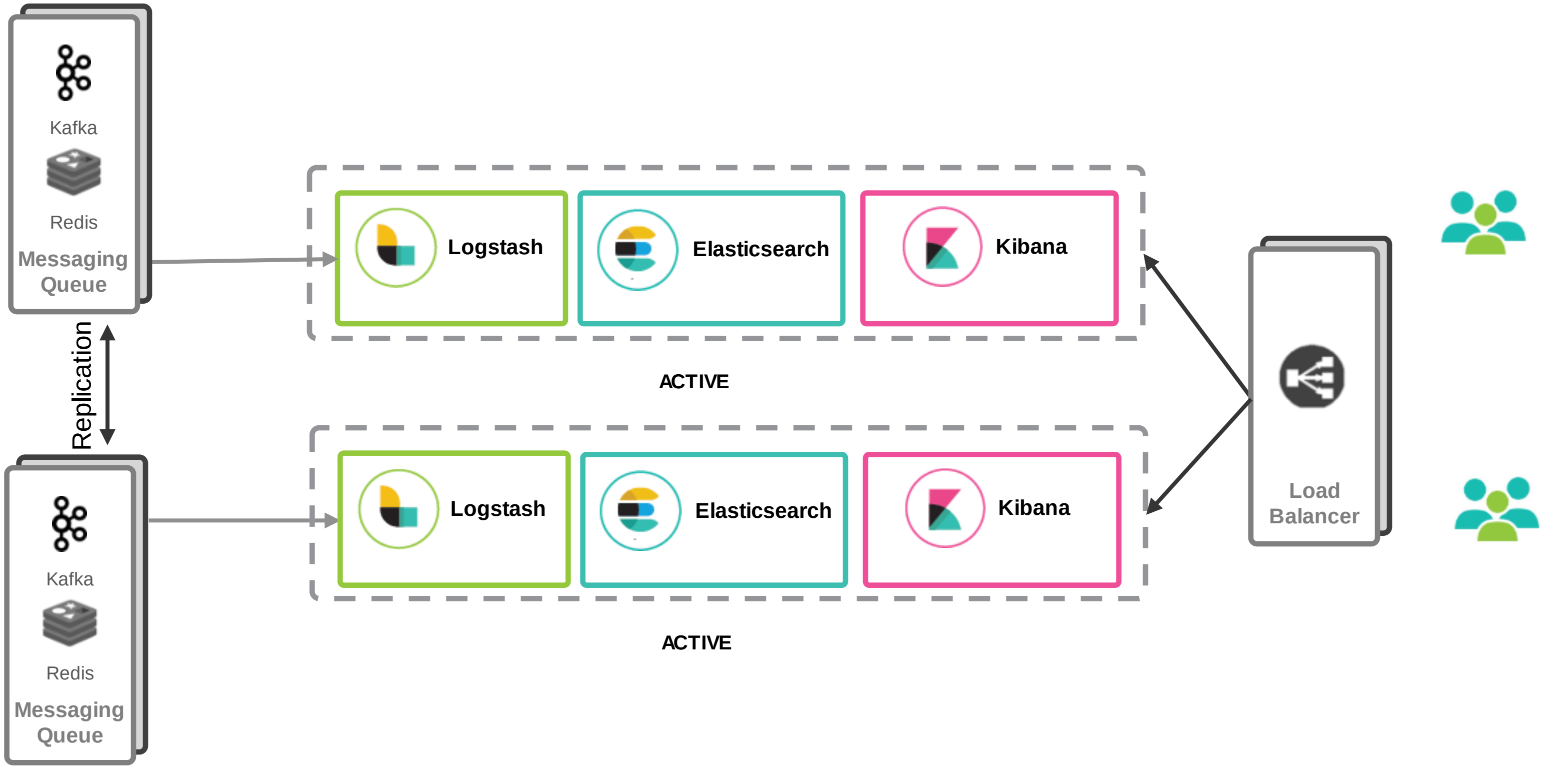
# Security Analytics Enterprise



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

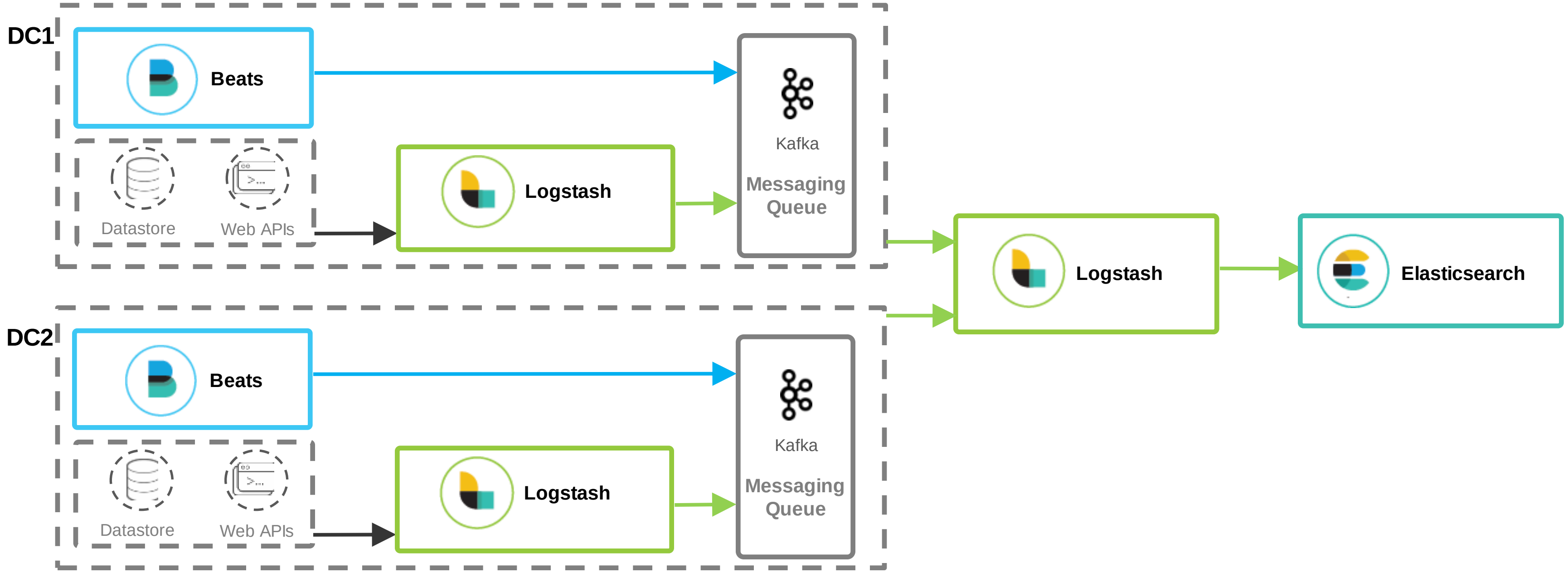
# Multi Data Center

# Multiple Data Centers, Duplicate Data

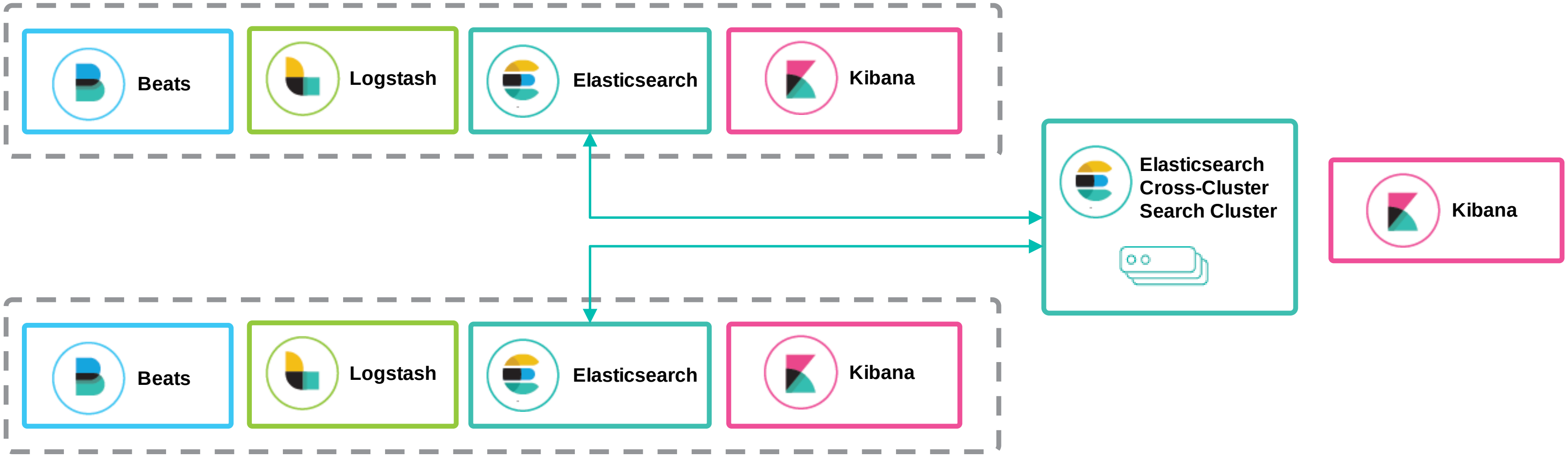


All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Multi Data Centers with a Queue at Each DC



# Multi Data Center, Distinct Data and Cross-Cluster Search

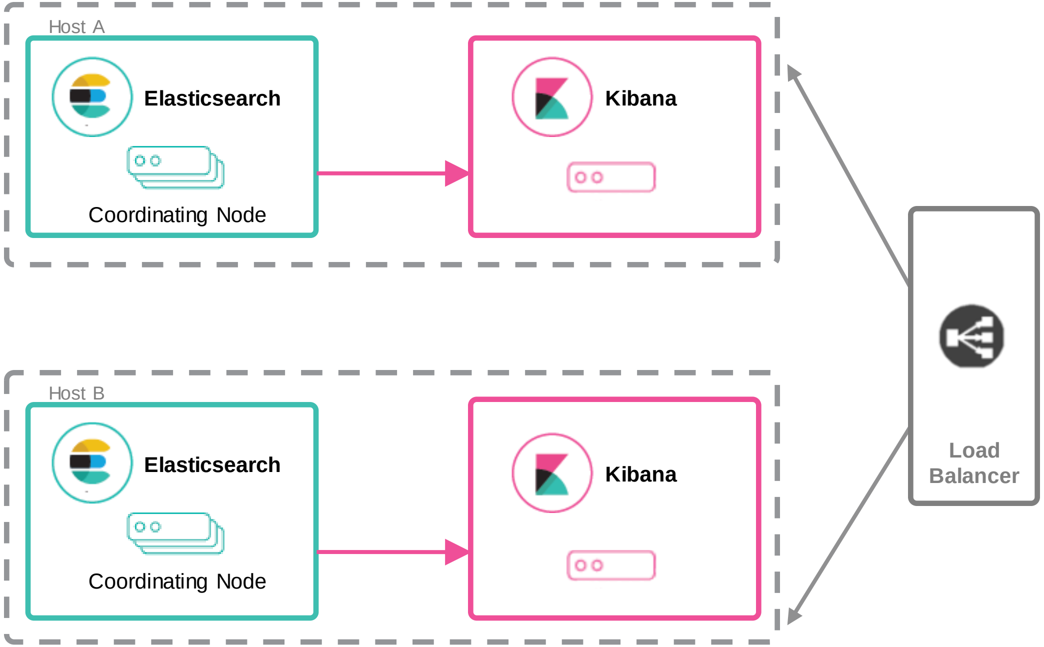


All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Scaling Kibana

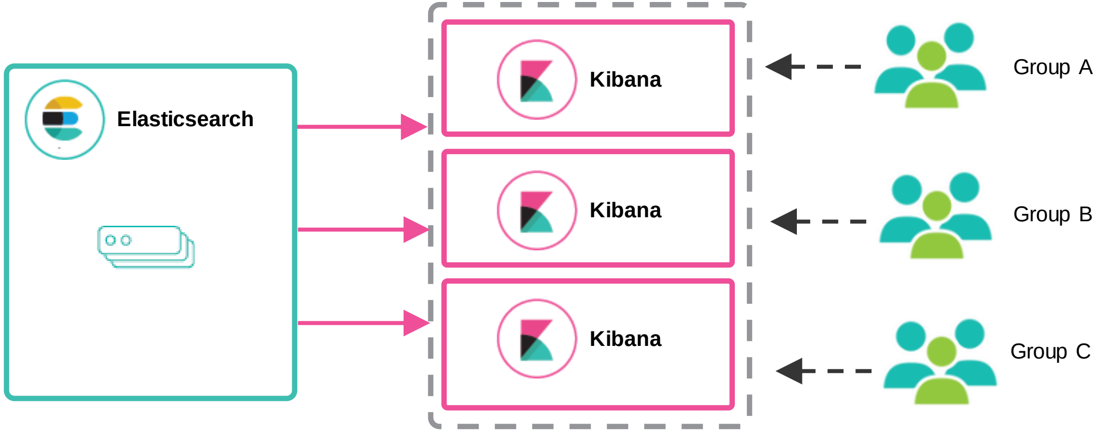
# High Availability

*Pair two coordinating nodes with two independent Kibana nodes*



# Separating Content by Groups

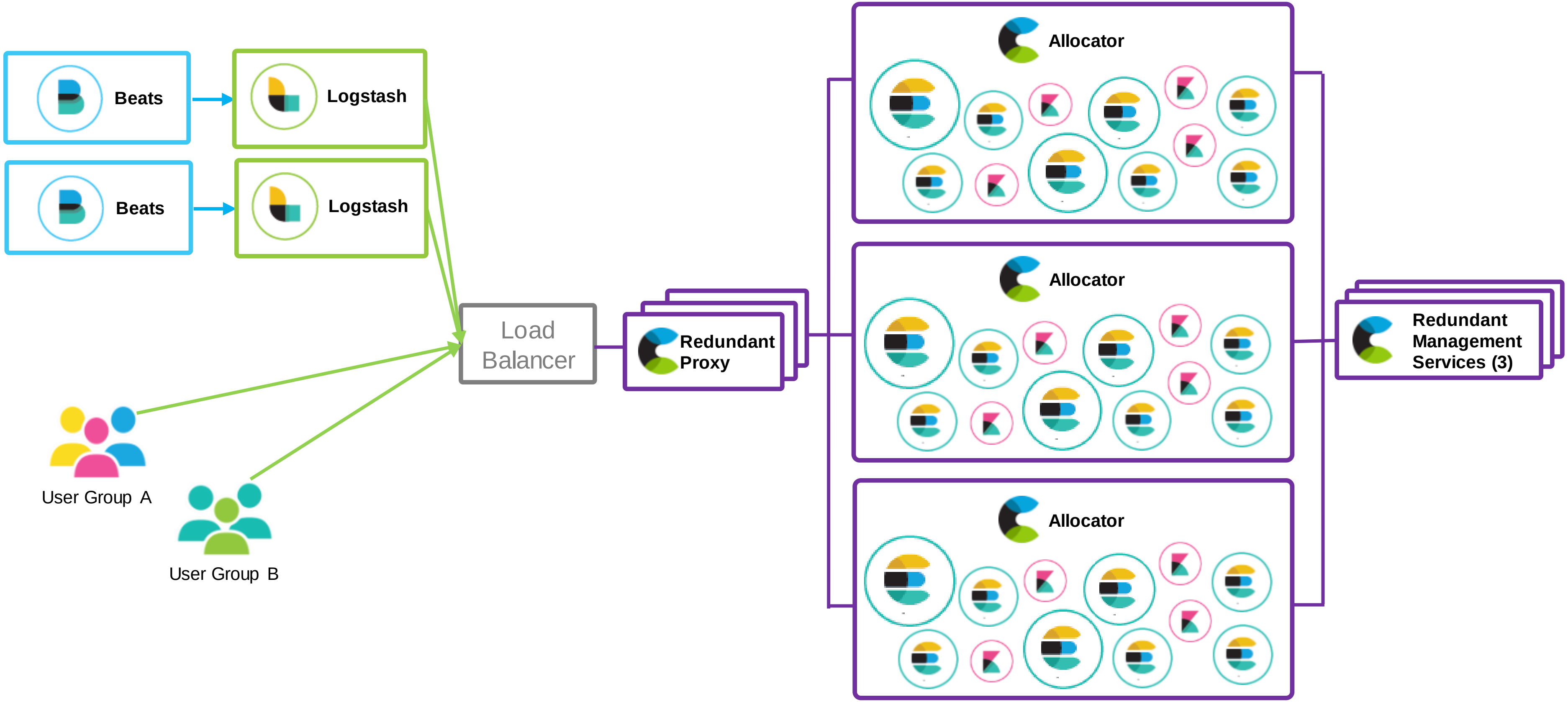
*Isolate user content by group in different Kibana instances*



# Elastic Enterprise

## A SaaS that fits the enterprise

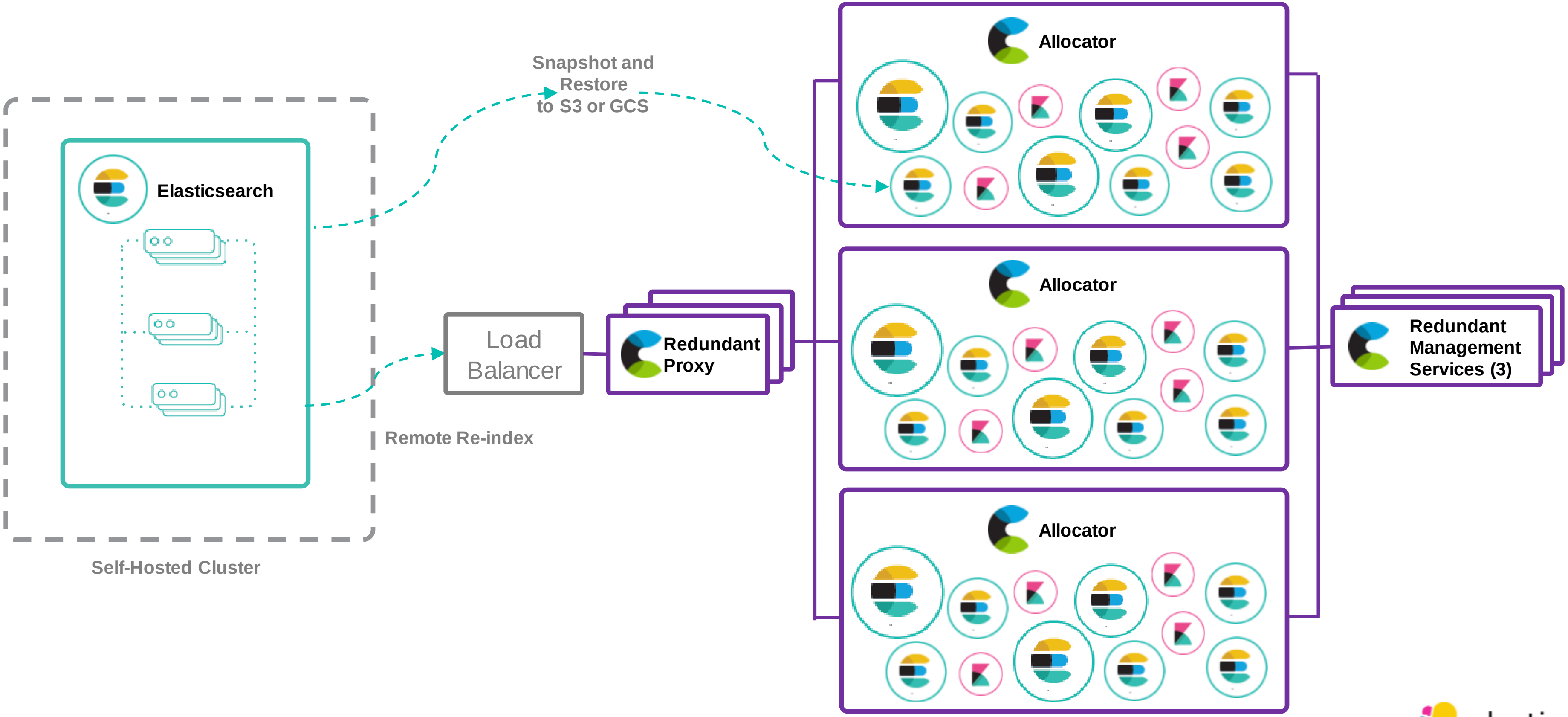
# Elastic Cloud Enterprise



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

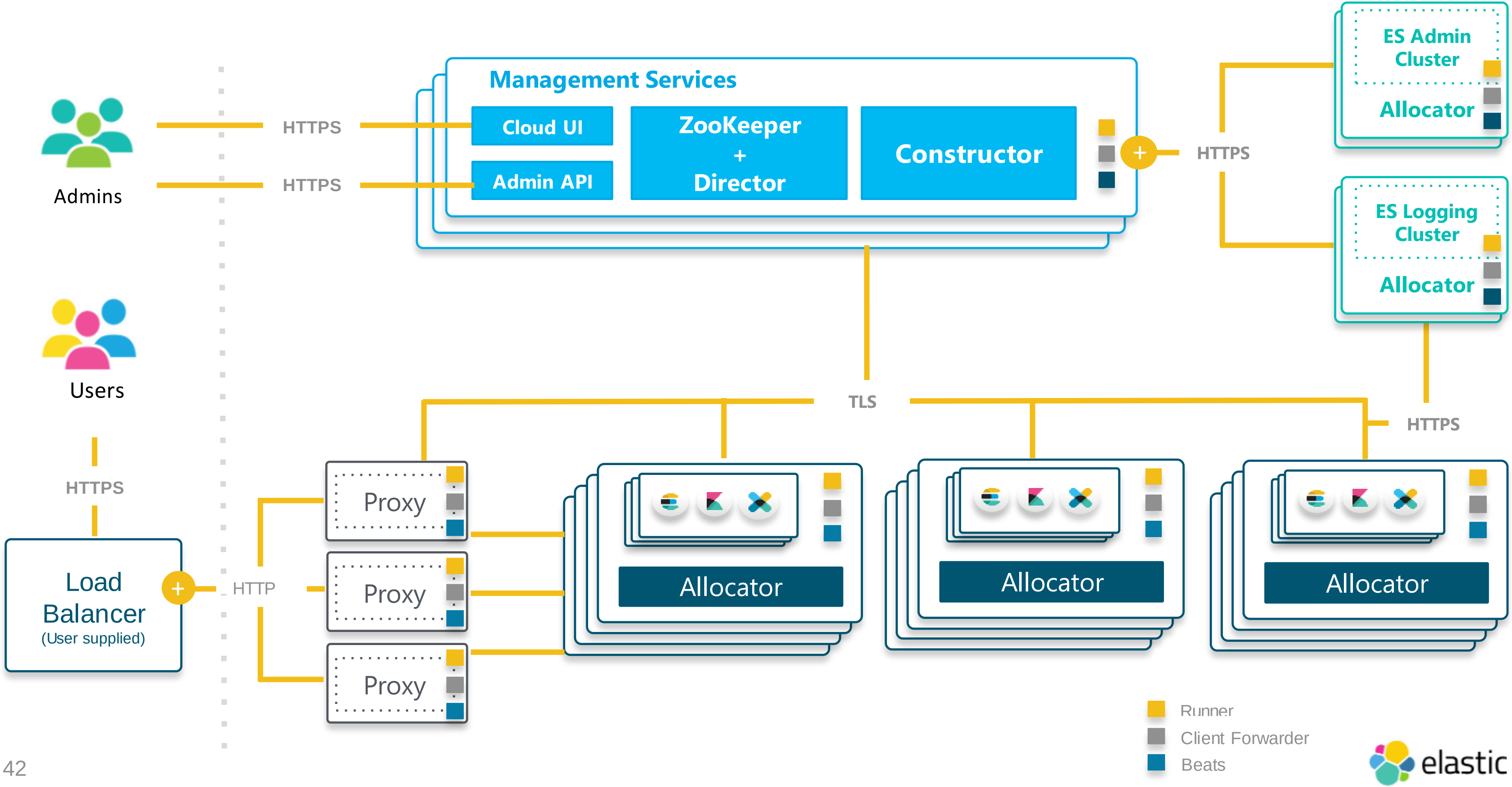
# Moving to Elastic Cloud Enterprise

*Options for easily moving your Elasticsearch cluster*



All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Elastic Cloud Enterprise high level architecture



# Hardware

# Hardware Recommendations

*For On-Premise or IaaS Cloud Deployments*

**Logstash**

**Heavy Worker**

16 CPU cores
32 GB RAM
Any Storage

6 TB

**Elasticsearch**

**Cool Data (Logging)**

8 CPU cores
64 GB RAM
SSD, SAS, Fast SAN

4 TB

**Elasticsearch**

**Warm/Uniform Data (Logging)**

8 CPU cores
64 GB RAM
SSD, SAS, Fast SAN

2 TB

**Elasticsearch**

**Hot Data (Logging)**

8+ CPU cores
64 GB RAM
SSD

**Kibana**

**Heavy Reporting**

2 CPU cores
8 GB RAM
Any Storage

**Logstash**

**Regular Worker**

8 CPU cores
16 GB RAM
Any Storage

**Elasticsearch**

**Ingest**

8 CPU cores
32 GB RAM
Any Storage

**Elasticsearch**

**Machine Learning**

8 CPU cores
64 GB RAM
Any Storage

1 TB

**Elasticsearch**

**High Throughput Search**

16 CPU cores
32 GB RAM
SSD

**Kibana**

**Regular**

2 CPU cores
4 GB RAM
Any Storage

**Logstash**

**Light Worker**

4 CPU cores
8 GB RAM
Any Storage

**Elasticsearch**

**Tribe or Cross-Cluster**

4 CPU cores
16 GB RAM
Any Storage

**Elasticsearch**

**Coordinating**

4 CPU cores
16 GB RAM
Any Storage

**Elasticsearch**

**Master**

4 CPU cores
8 GB RAM
Any Storage

All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

** Data volumes are guidelines*

# Hardware Recommendations

*For Elastic Cloud Enterprise Deployments*

- *Data volumes are guidelines*
- *Multi data tiers available from ECE 1.2*

| | SEARCH 1:16 | HOT 1:32 | WARM OR UNIFORM 1:64 | COOL 1:96 |
|---|---|---|---|---|
| Logstash<br>**Heavy Worker**<br>16 CPU cores<br>32 GB RAM<br>Any Storage | 8 TB<br>Elastic Cloud Enterprise<br>**Heavy Instance (Search)**<br>64+ CPU cores<br>512 GB RAM<br>SSD | 16 TB<br>Elastic Cloud Enterprise<br>**Heavy Instance (Hot)**<br>64+ CPU cores<br>512 GB RAM<br>SSD | 32 TB<br>Elastic Cloud Enterprise<br>**Heavy Instance (Warm)**<br>64 CPU cores<br>512 GB RAM<br>SSD, SAN, SATA | 48 TB<br>Elastic Cloud Enterprise<br>**Heavy Instance (Cool)**<br>64 CPU cores<br>512 GB RAM<br>SSD, SAN, SATA |
| Logstash<br>**Regular Worker**<br>8 CPU cores<br>16 GB RAM<br>Any Storage | 4 TB<br>Elastic Cloud Enterprise<br>**Regular Instance (Search)**<br>32+ CPU cores<br>256 GB RAM<br>SSD | 8 TB<br>Elastic Cloud Enterprise<br>**Regular Instance (Hot)**<br>32+ CPU cores<br>256 GB RAM<br>SSD | 16 TB<br>Elastic Cloud Enterprise<br>**Regular Instance (Warm)**<br>32 CPU cores<br>256 GB RAM<br>SSD, SAN, SATA | 24 TB<br>Elastic Cloud Enterprise<br>**Regular Instance (Cool)**<br>32 CPU cores<br>256 GB RAM<br>SSD, SAN, SATA |
| Logstash<br>**Light Worker**<br>4 CPU cores<br>8 GB RAM<br>Any Storage | 2 TB<br>Elastic Cloud Enterprise<br>**Light Instance (Search)**<br>16+ CPU cores<br>128 GB RAM<br>SSD | 4 TB<br>Elastic Cloud Enterprise<br>**Light Instance (Hot)**<br>16+ CPU cores<br>128 GB RAM<br>SSD | 8 TB<br>Elastic Cloud Enterprise<br>**Light Instance (Warm)**<br>16 CPU cores<br>128 GB RAM<br>SSD, SAN, SATA | 12 TB<br>Elastic Cloud Enterprise<br>**Light Instance (Cool)**<br>16 CPU cores<br>128 GB RAM<br>SSD, SAN, SATA |

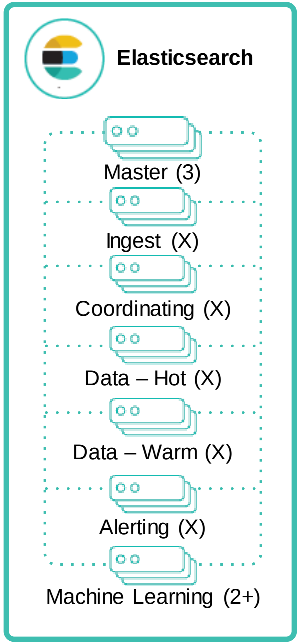
| CONTROL PLANE |
|---|
| Elastic Cloud Enterprise<br>**Dedicated Director + Coordinator + Proxy**<br>4 CPU cores<br>32 GB RAM<br>Any Storage |
| Elastic Cloud Enterprise<br>**Dedicated Proxy**<br>4 CPU cores<br>16 GB RAM<br>Any Storage |

All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Inside a Cluster

# Elasticsearch Node Types

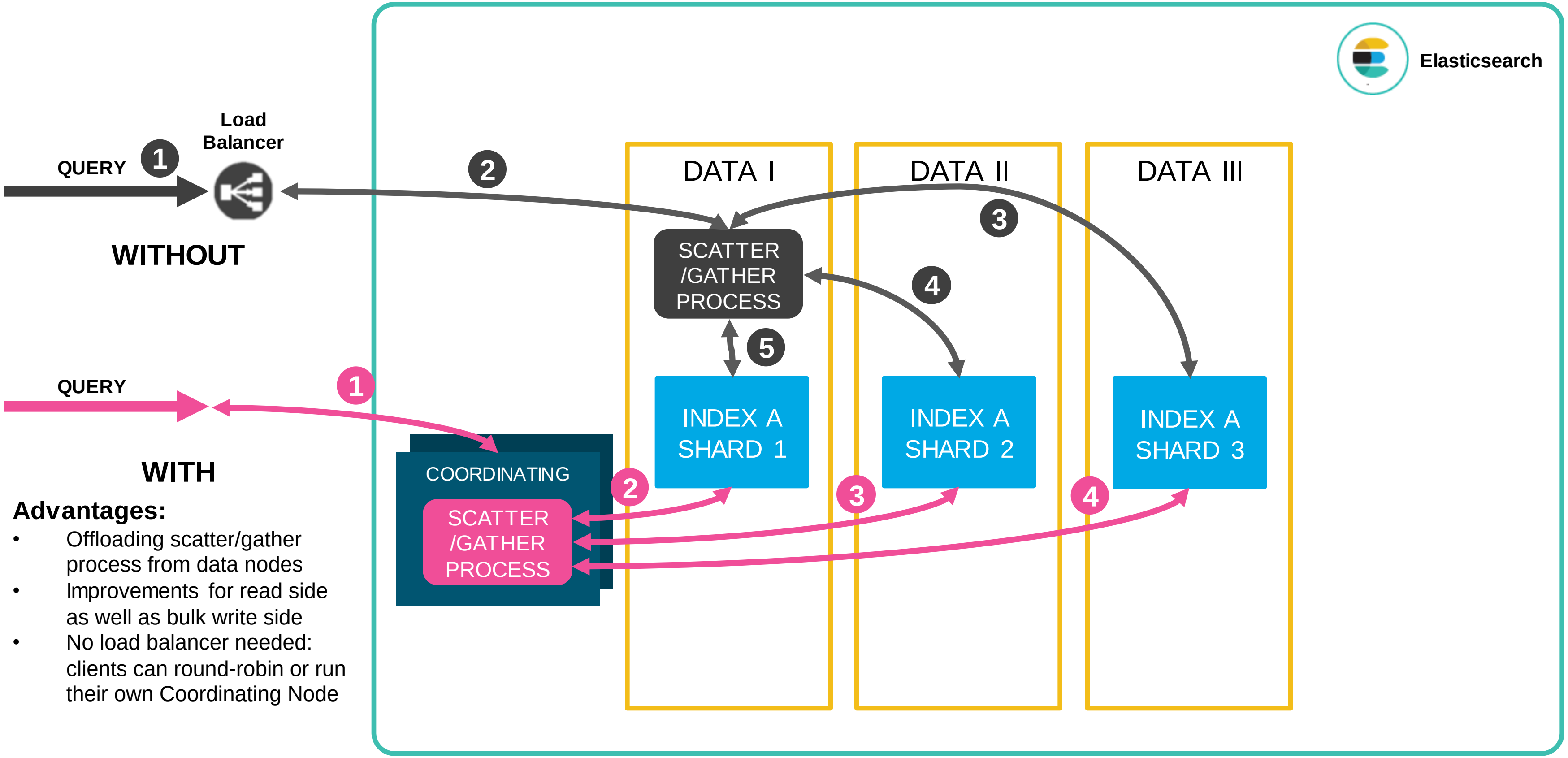
*Nodes can play one or more roles, for workload isolation and scaling*



- 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

All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Elasticsearch Coordinating Nodes Detail

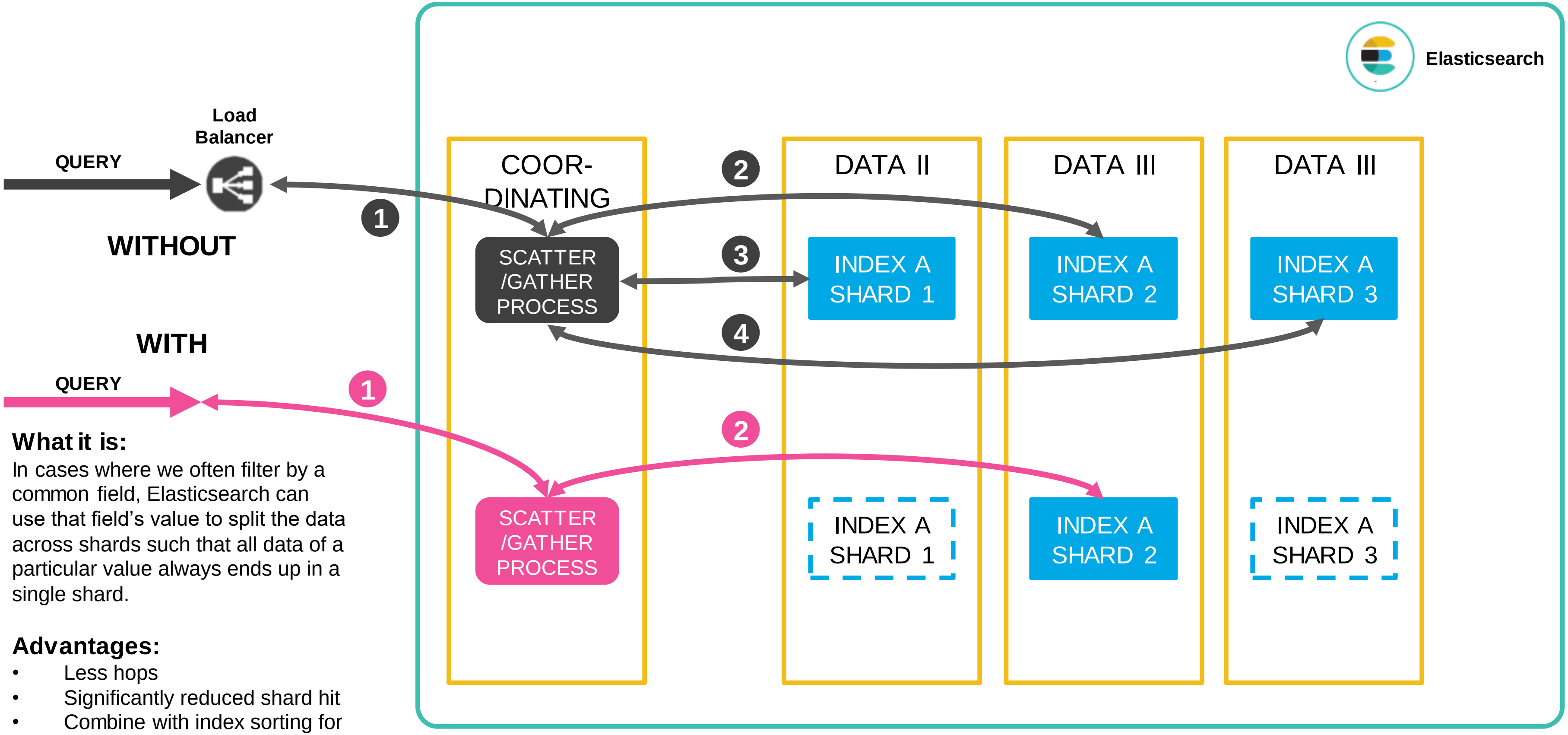


**Advantages:**

- Offloading scatter/gather process from data nodes
- Improvements for read side as well as bulk write side
- No load balancer needed: clients can round-robin or run their own Coordinating Node

All product names, logos, and brands are property of their respective owners and are used only for identification purposes. This is not an endorsement.

# Elasticsearch Custom Routing



**What it is:**
In cases where we often filter by a common field, Elasticsearch can use that field's value to split the data across shards such that all data of a particular value always ends up in a single shard.

**Advantages:**

- Less hops
- Significantly reduced shard hit
- Combine with index sorting for crazy-fast searching