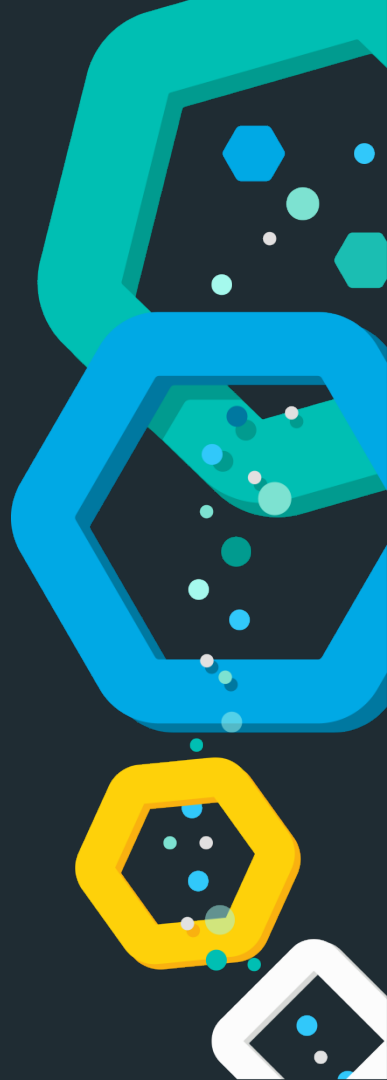
elastic{on}TOUR

Ingest

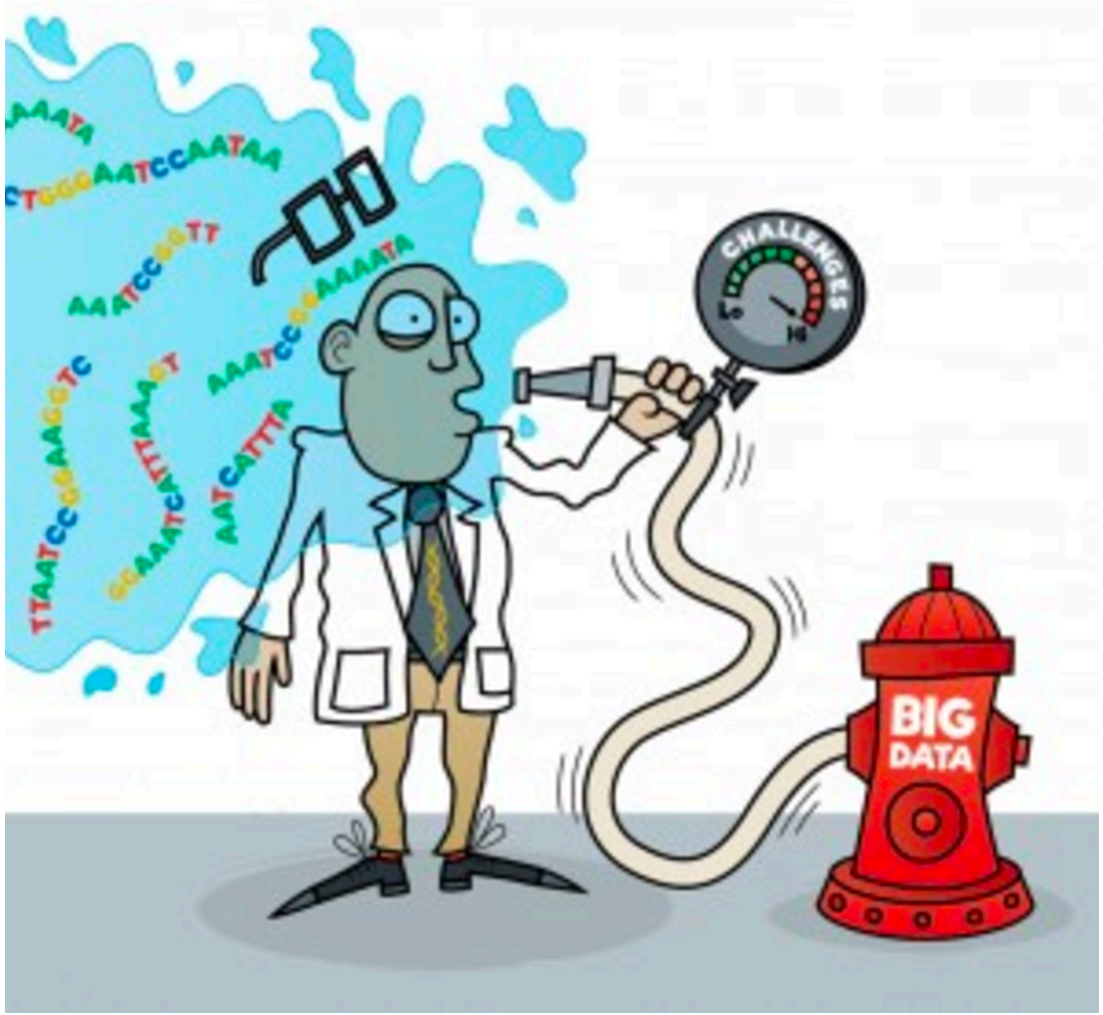
David Pilato, Developer | Evangelist

Paris, 31 Janvier 2017



Data Ingestion

The process of collecting and importing data for immediate use in a datastore



elastic


simple things
should be simple
elastic

Ingest Technologies

Elasticsearch

APIs

Ingest Node

Beats

Lightweight Data Shippers

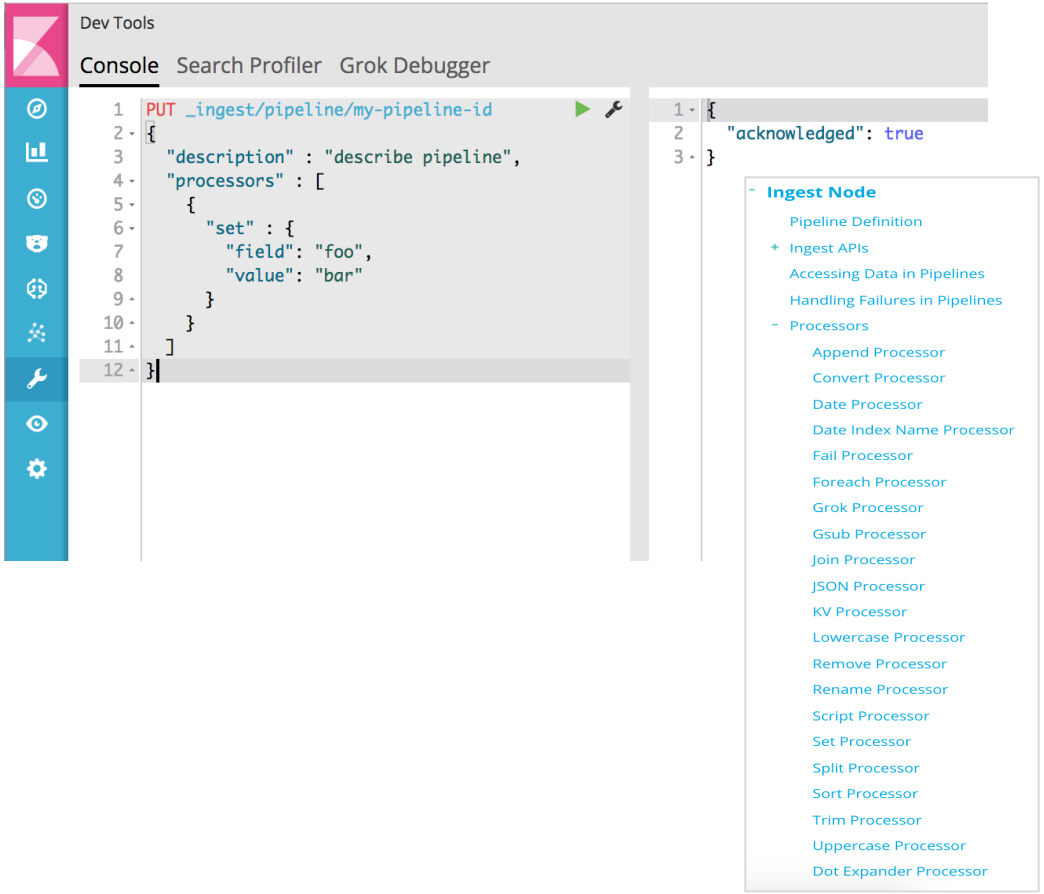
Logstash

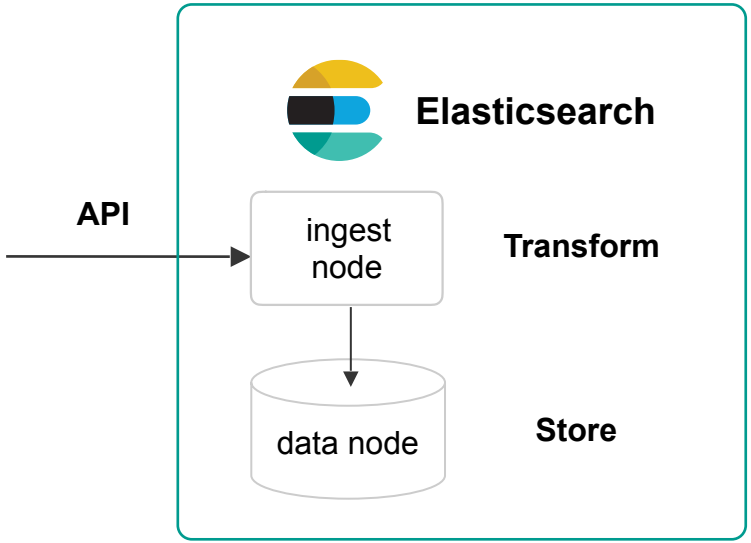
Centralized Data Collection Engine

ES-Hadoop

Hadoop Ecosystem Connector

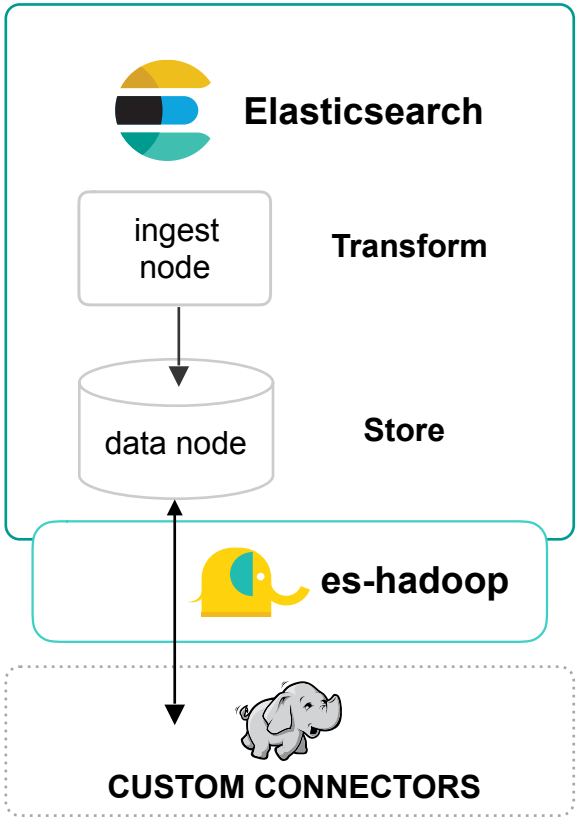
Elastic Ingestion Technologies





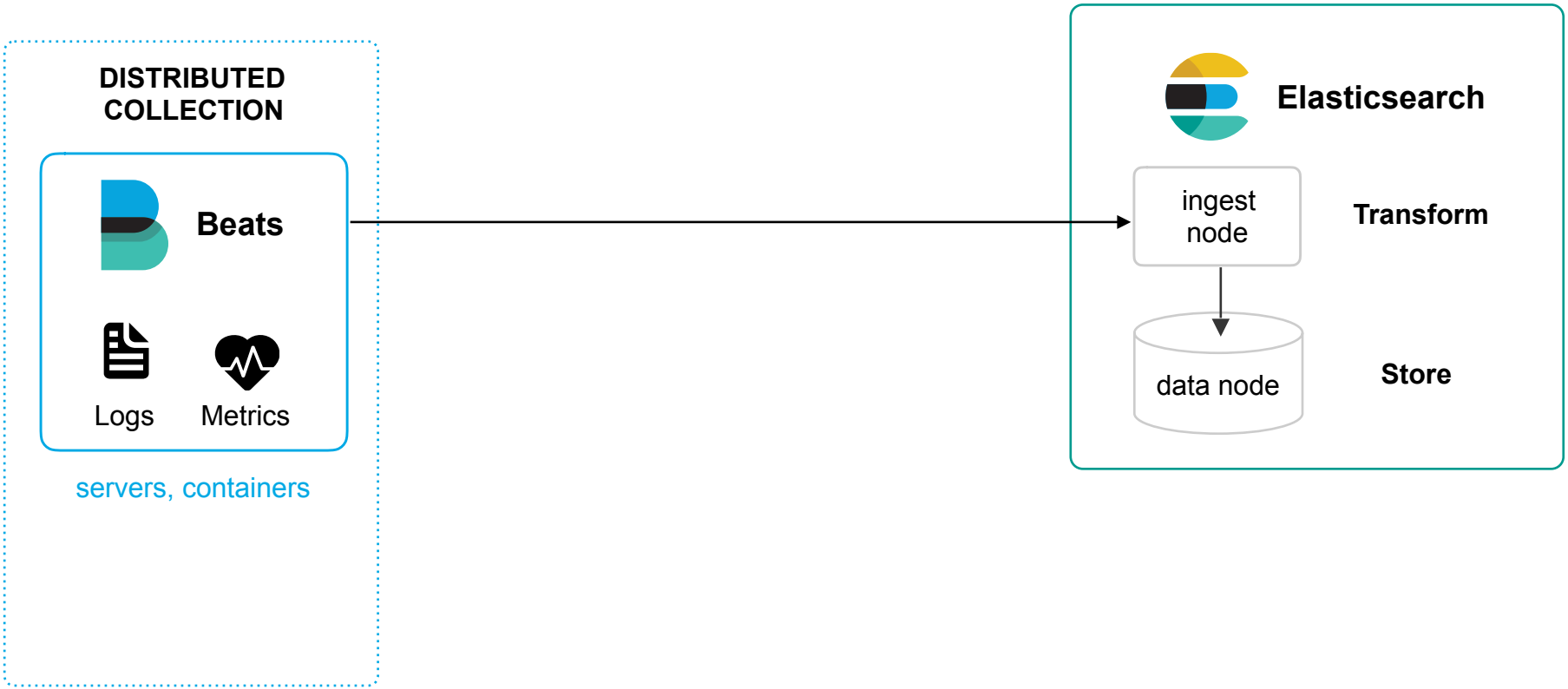
elastic

Elastic Ingestion Technologies



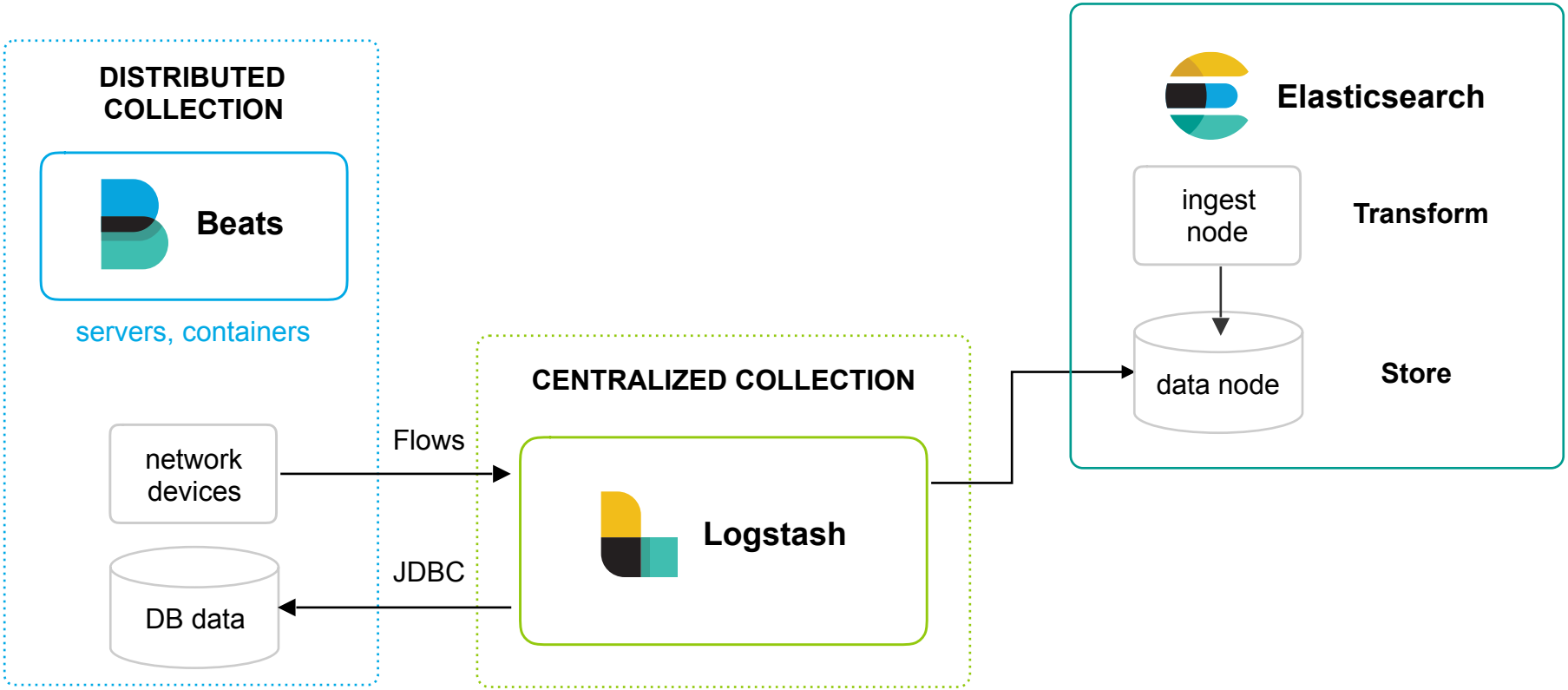
elastic

Elastic Ingestion Technologies



elastic

Elastic Ingestion Technologies

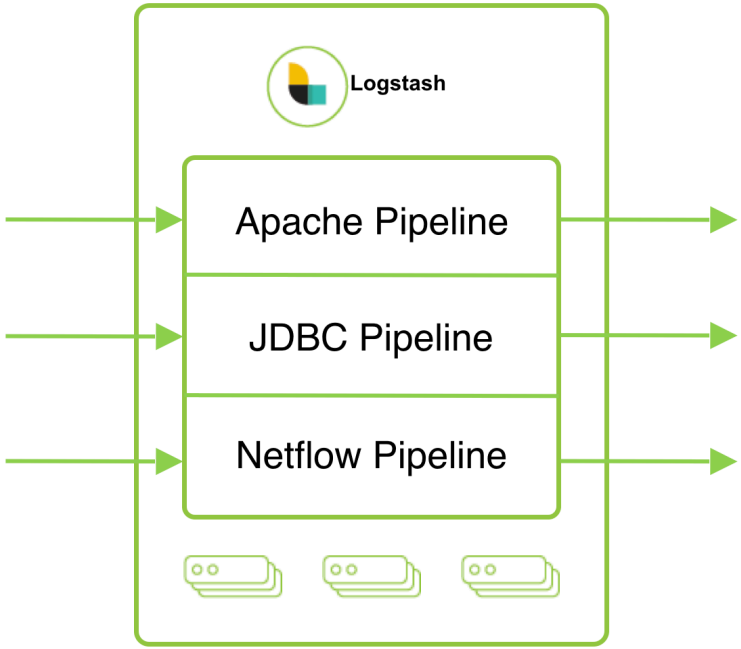


elastic

Ingestion Architecture

Scalable and robust centralized ETL

- Java event rewrite
- Multiple pipelines

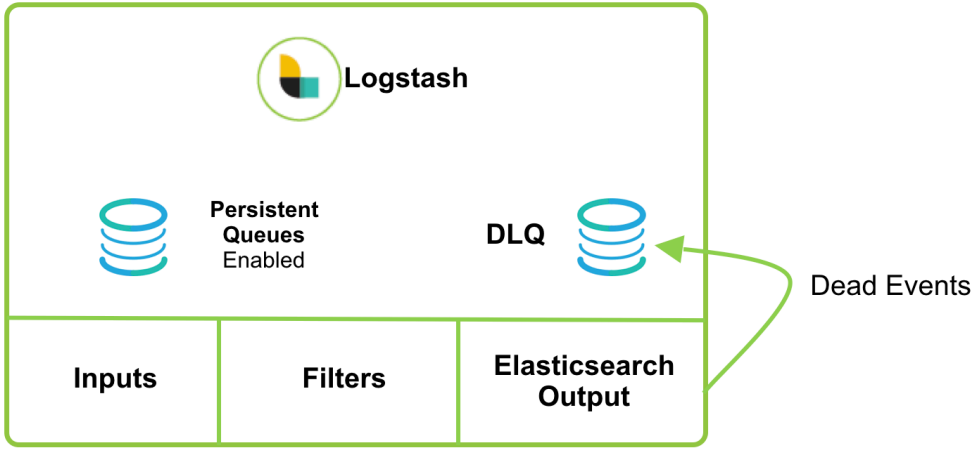


elastic

Ingestion Architecture

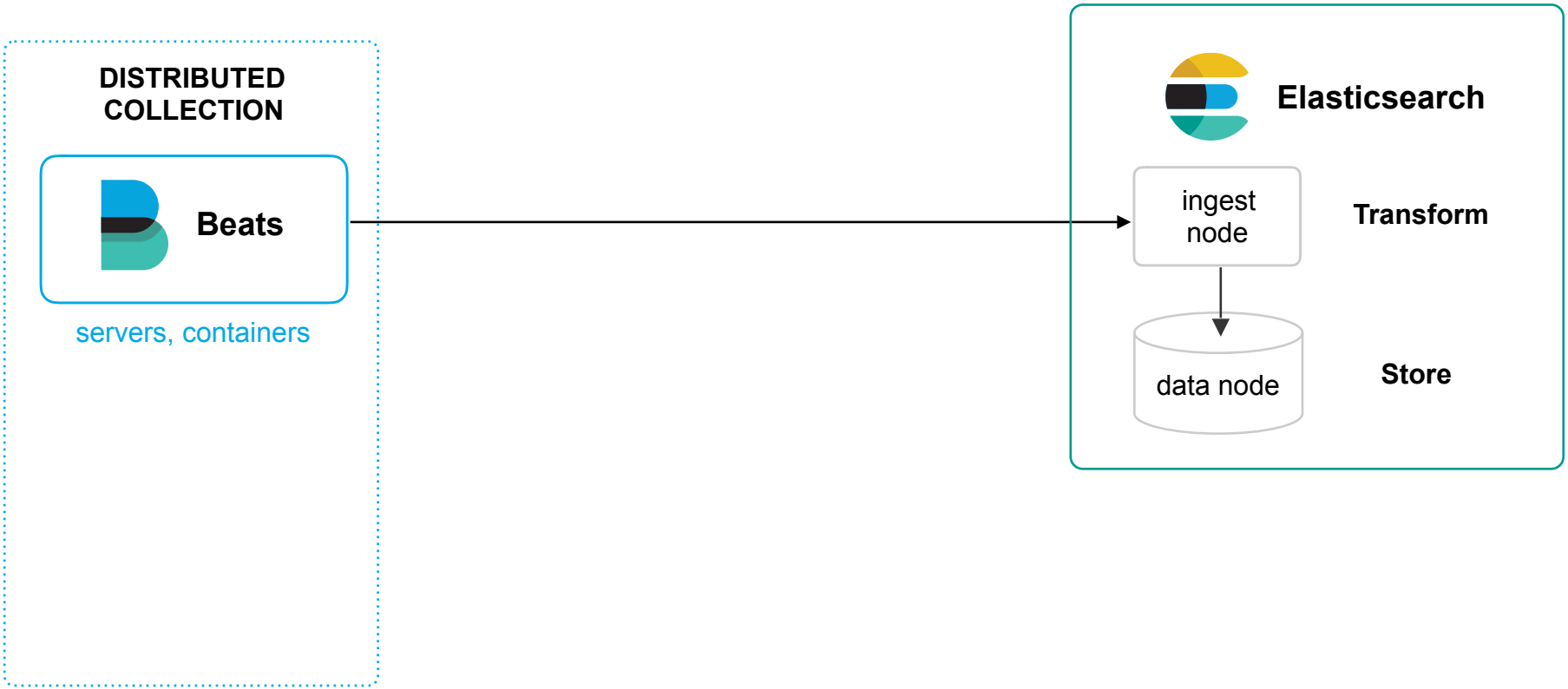
Scalable and robust centralized ETL

- Persistent queues
- Dead letter queues



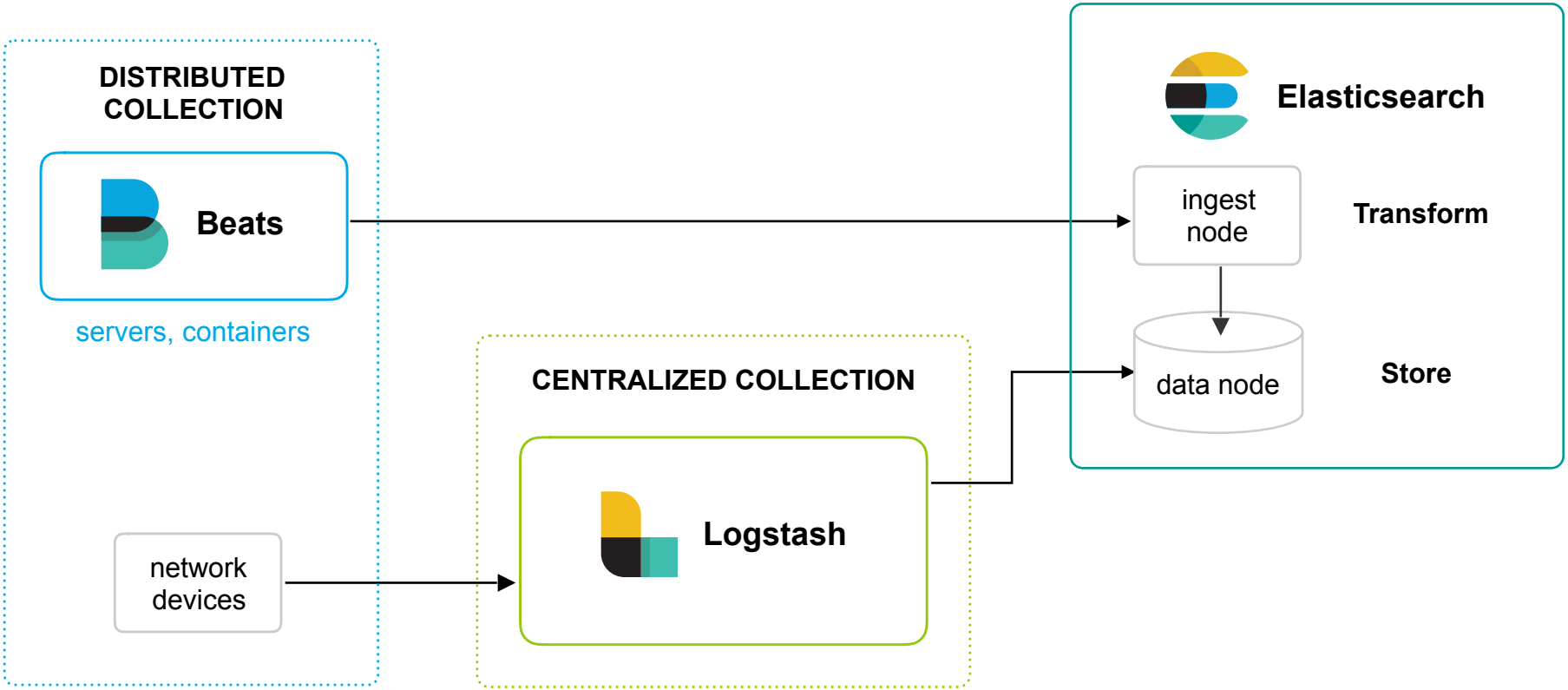
elastic

Cooperative Ingestion



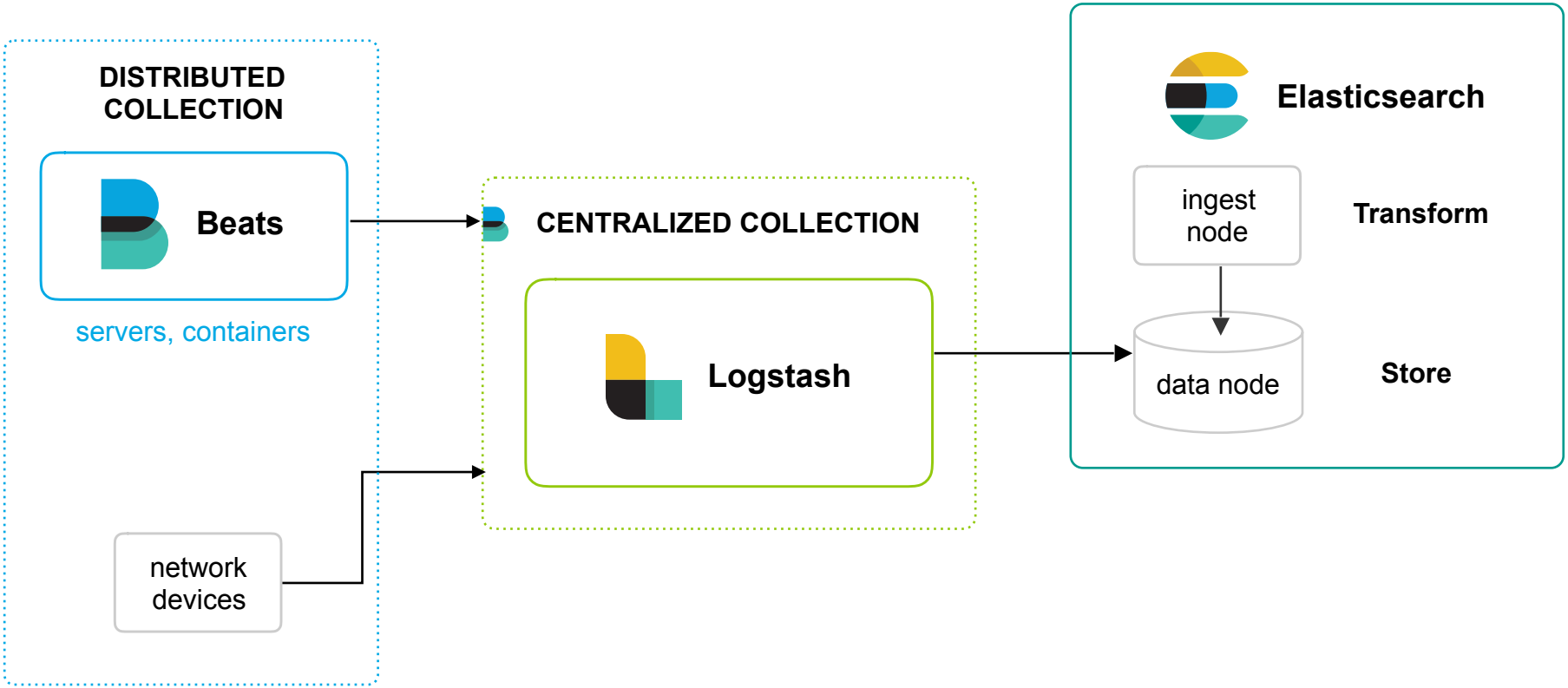
elastic

Elastic Ingestion Technologies



elastic

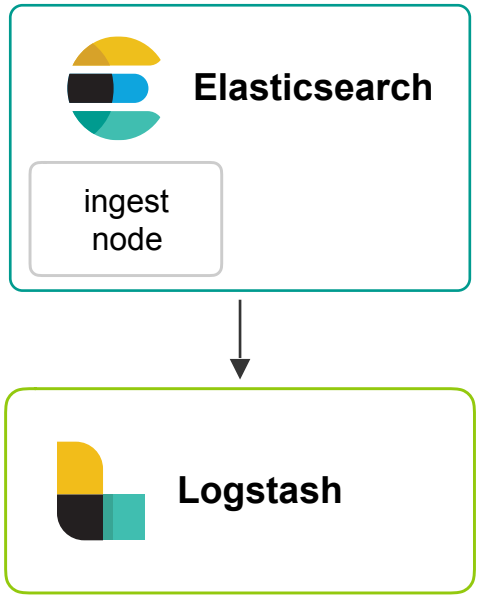
Elastic Ingestion Technologies



elastic

Easy migration between ingest technologies

Ingest Node to Logstash conversion tool



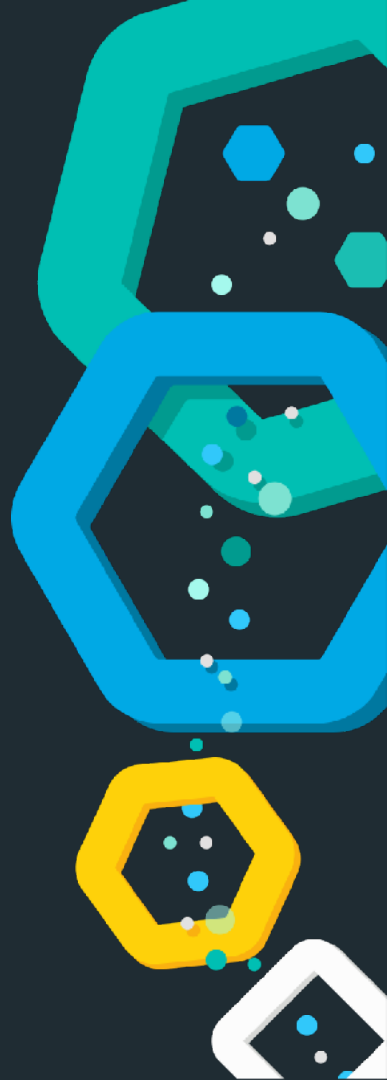

```
$ bin/ingest-convert.sh --input=file:///tmp/ingest-nginx.json --output=file:///tmp/logstash-nginx.conf
$ cat /tmp/logstash-nginx.conf
filter {
   grok {
      match => {
         "request" => "%{WORD:request_action} %{DATA:request1} HTTP/%{NUMBER:http_version}"
      }
   }

   geoip {
      source => "remote_ip"
      target => "geoip"
   }
   date {
      match => [
         "time",
         "dd/MMM/YYYY:HH:mm:ss Z"
      ]
      target => "@timestamp"
      locale => "ENGLISH"
   }
}
output {
   elasticsearch {
      hosts => "localhost"
   }
}
$
```


elastic

elastic

Use Cases & Data Sources



Use Cases & Data Sources

Logging

Common Log Formats

System
Web Servers
Queues

Metrics

Turnkey Monitoring

Infrastructure
Containers
Databases

Security

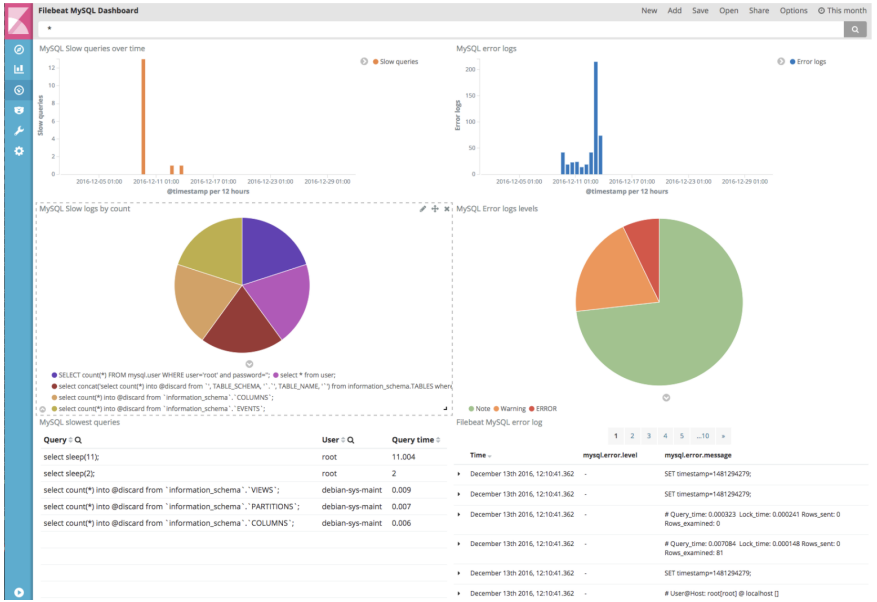
SecOps Dashboards

Audit
Firewalls, IDS/IPS
SIEM Augmentation

Modules: Data sources made easy

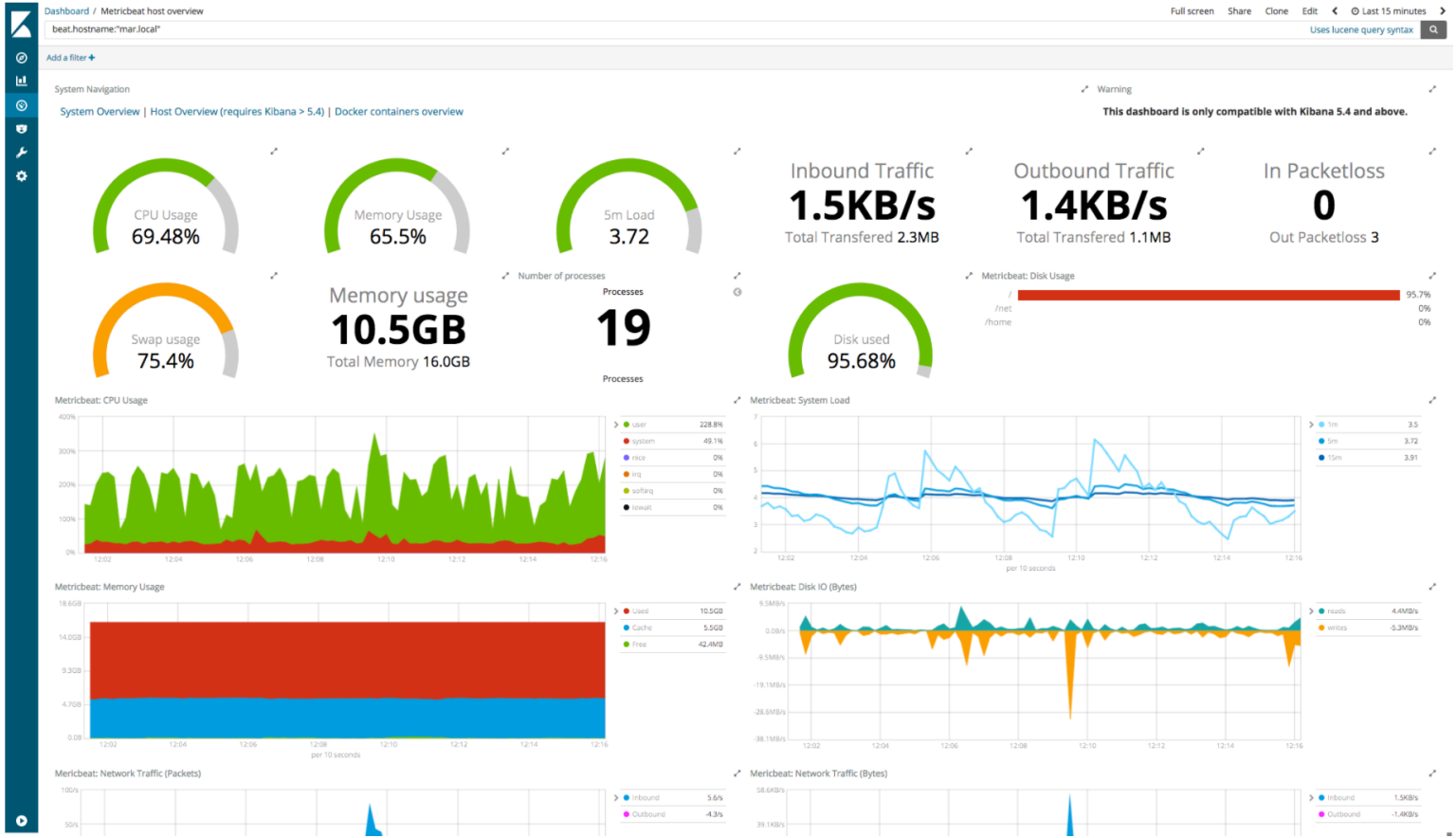
- Collect specific type of data
- Parse and enrich it
- Default dashboards, alerts, ML jobs

./filebeat -e -modules=system -setup



elastic

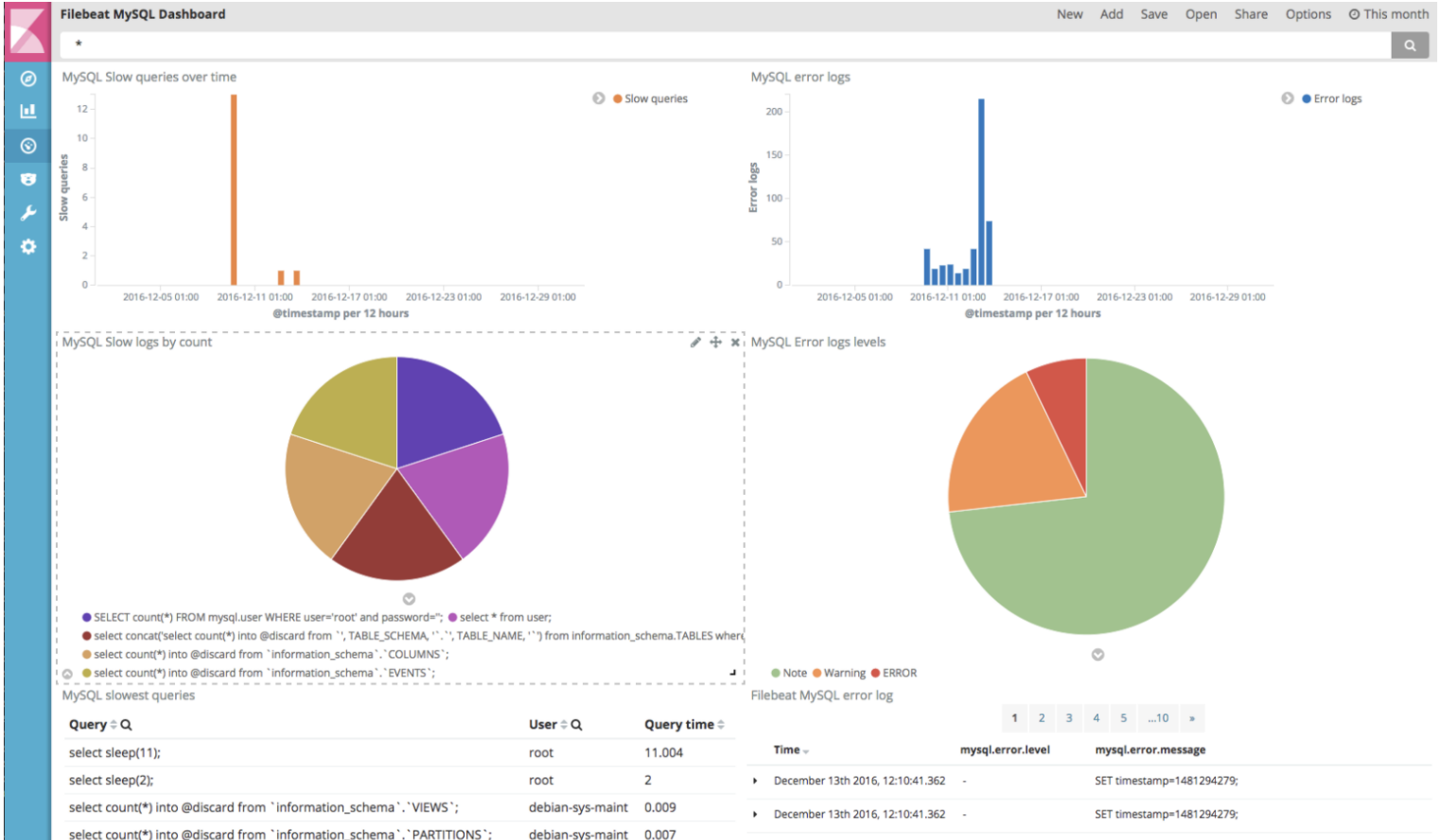
Metricbeat modules (introduced in 5.0)



Aerospike
Apache
Ceph
Couchbase
Docker
Dropwizard
Elasticsearch
Golang
Graphite
HAProxy
HTTP
Jolokia
Kafka
Kibana
Kubernetes
Memcached
MongoDB
MySQL
Nginx
PHP_FPM
PostgreSQL
Prometheus
RabbitMQ
Redis
System
vSphere
Windows
ZooKeeper

elastic

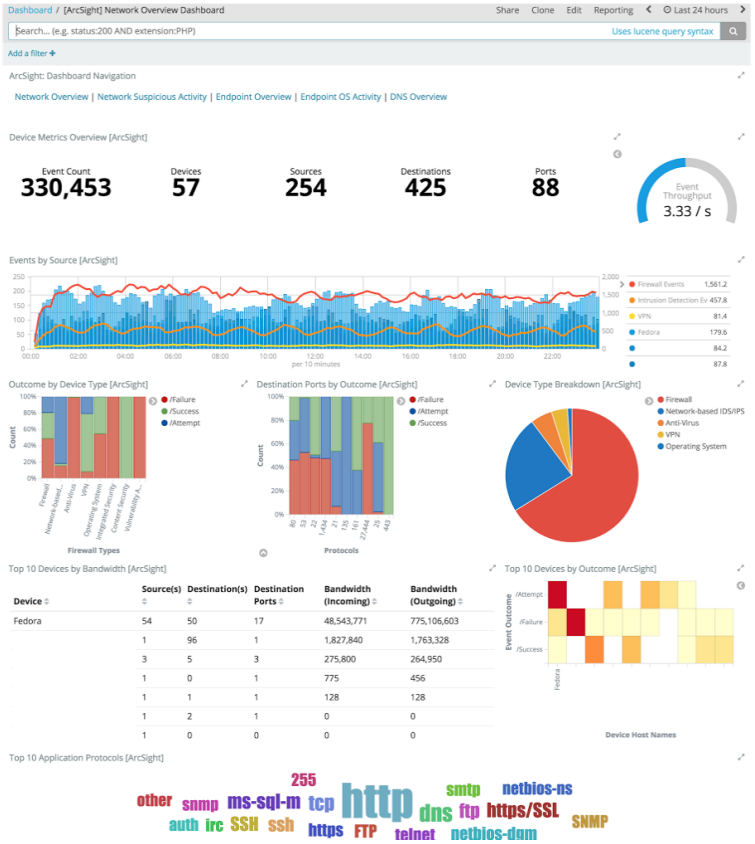
Filebeat modules (introduced in 5.3)



Apache2
Auditd
Icinga
Kafka
MySQL
Nginx
PostgreSQL
Redis
System

elastic

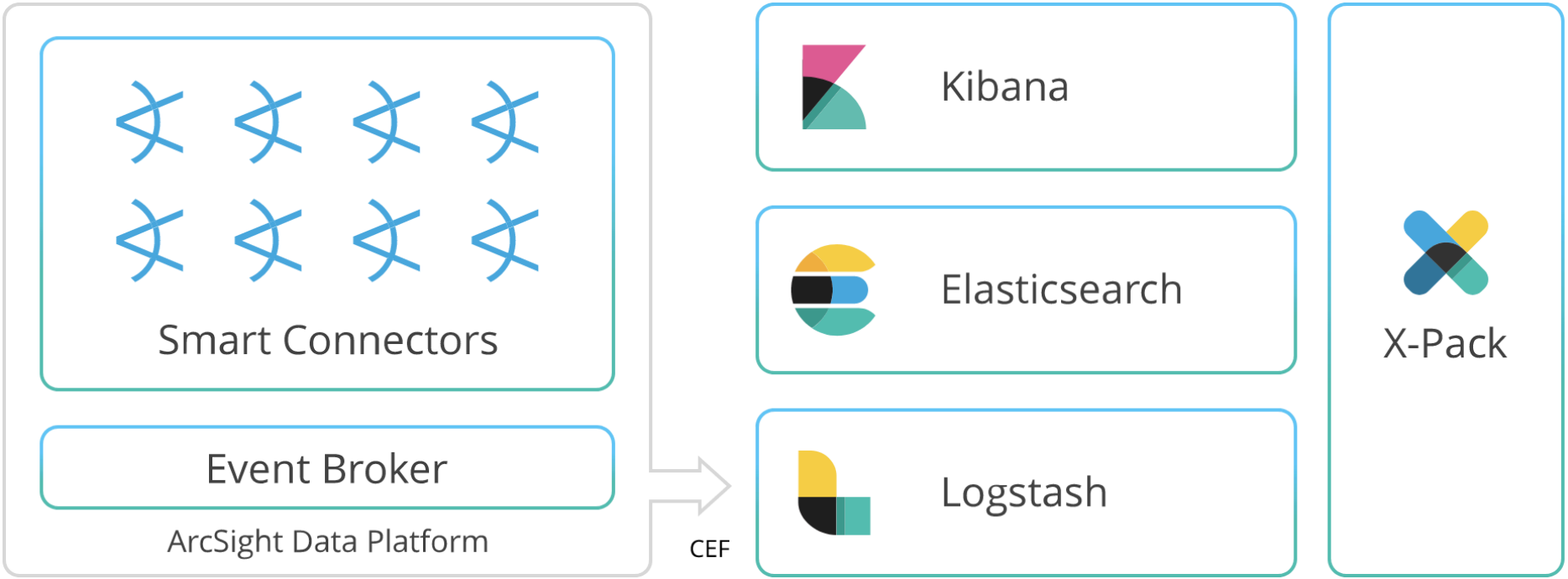
Logstash modules (introduced in 5.6)



ArcSight Netflow

elastic

ArcSight Module (Introduced in 5.6)



elastic

DEMO

Logging Data Sources

FILEBEAT

WINLOGBEAT

Infrastructure

System

- Linux / MacOS
- Windows Events

Containers

- Docker (6.0)
- Kubernetes (6.0)

Applications

Databases

- MySQL
- PostgreSQL (6.1)

Queues

- Kafka (6.1)
- Redis (6.0)

Web servers

- Apache
- Nginx

Other

- HAProxy*
- Zookeeper*

** Near-term roadmap*

Metrics & Event Data

Infrastructure

System
- Linux
- MacOS
- Windows
- Perfmon (6.0)
- WMI*

Containers
- Docker
- Kubernetes (6.0)

Virtualization
- vSphere (6.0)

Cloud
- AWS
- GCP
- Azure*
- DigitalOcean
- ….

Network
- Netflow (5.6)
- Packets

Storage
- Ceph

** Near-term roadmap*

elastic

Metrics & Event Data

METRICBEAT HEARTBEAT LOGSTASH



Applications

Datastores
- MySQL
- PostgreSQL
- MongoDB
- Couchbase
- Aerospike (6.0)
- Graphite (6.1)

Queues
- Kafka
- Redis
- RabbitMQ (6.0)

Caches
- Memcached (6.0)

Uptime
- Heartbeat

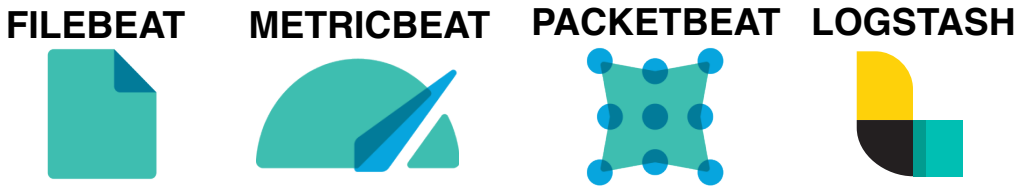
Custom apps
- JMX/Jolokia
- PHP-FPM
- Golang (6.0)
- Dropwizard (6.0)

Web servers
- Apache
- Nginx

Other
- HAProxy
- Zookeeper
- Prometheus

** Near-term roadmap*

Security Data Sources



Security

SIEM Augmentation

- ArcSight (5.6)
- more*

Audit

- Auditd
- Auditbeat (6.0)

Activity

Systems

- Access
- SSH

Applications

- Connections
- Users

Network

- IPs / GeoIP
- DNS Packets
- Netflow (5.6)
- Firewalls*
- IDS/IPS*

** Near-term roadmap*

elastic

Business Analytics

LOGSTASH

Structured

Databases

- JDBC input
- JDBC filter

Activity

SaaS services

- Salesforce
- Heroku
- Github
- Azure*

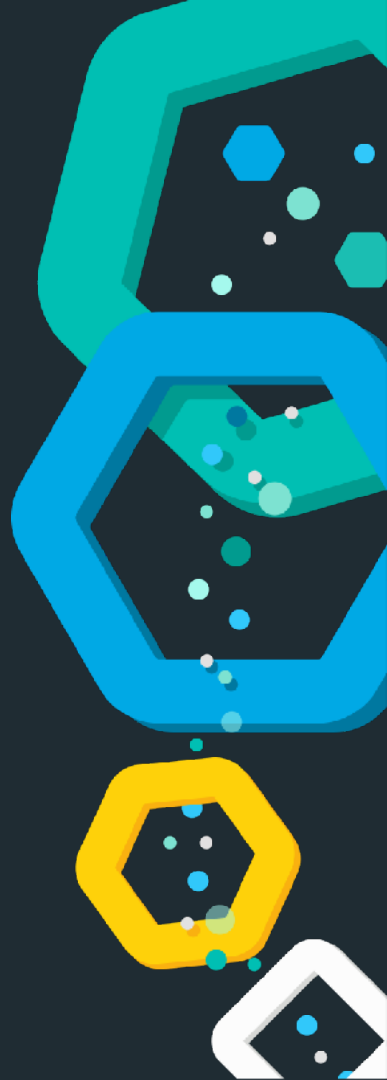
Social media

- Twitter

** Near-term roadmap*

elastic

Administration

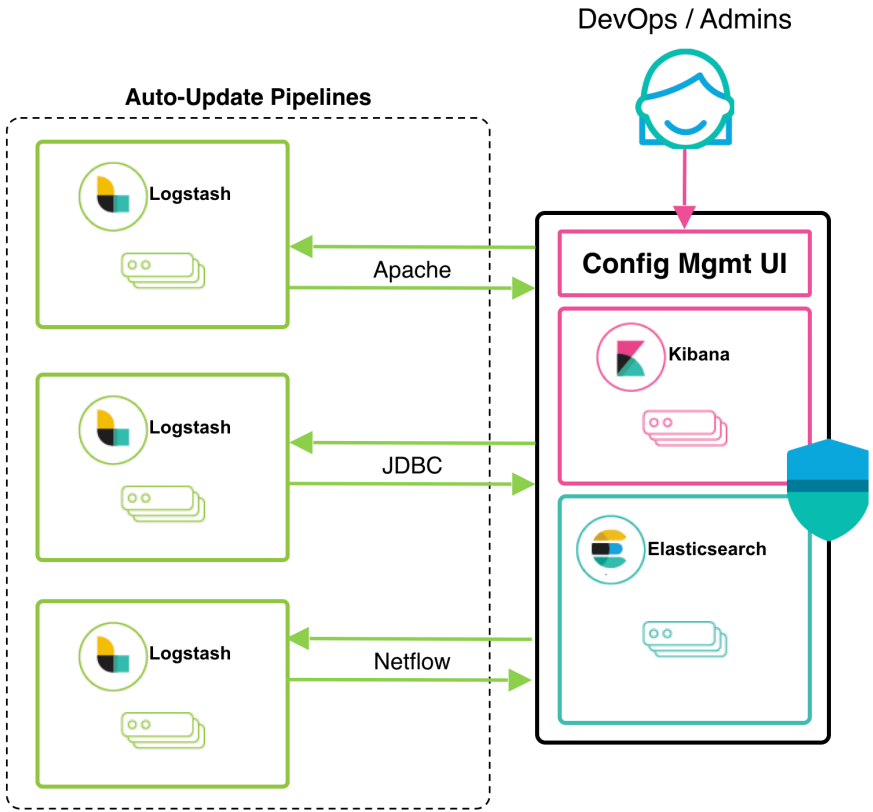


Monitoring & Management

Logstash

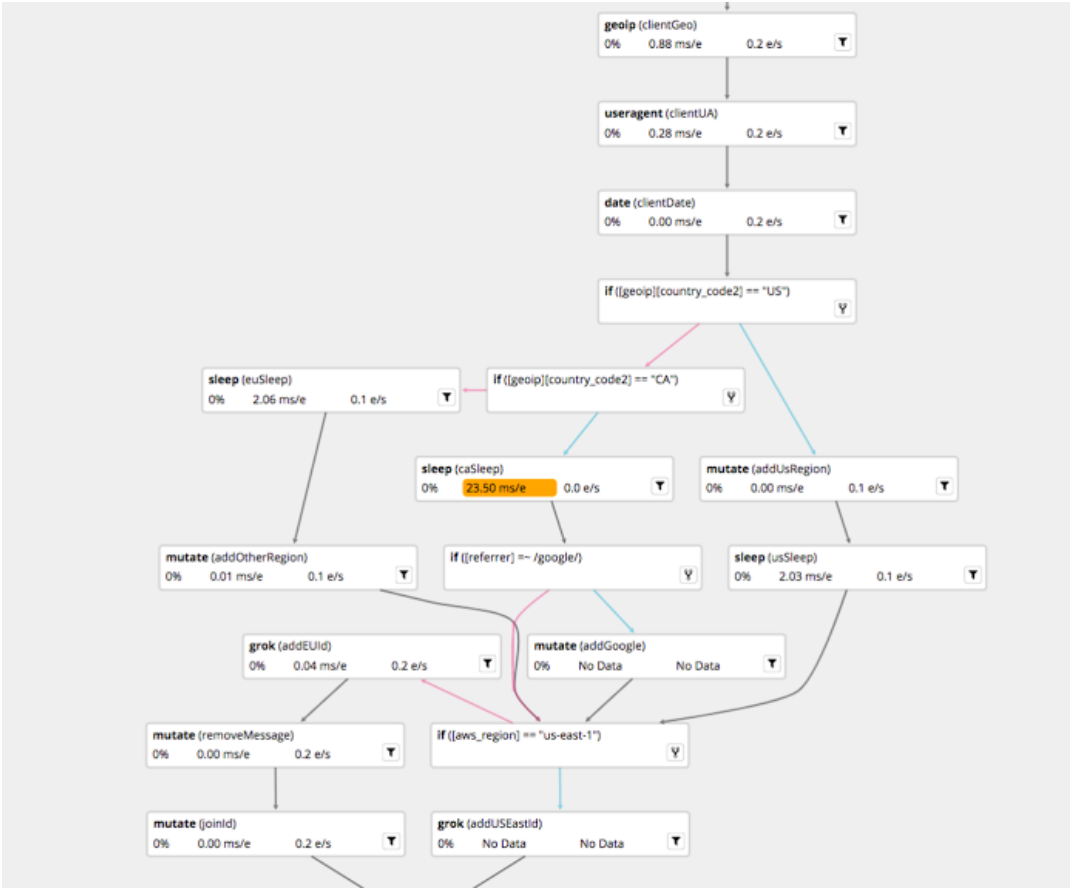
- Centralized monitoring (5.3)
- Centralized management (6.0)

** Near-term roadmap*



elastic

Logstash Monitoring



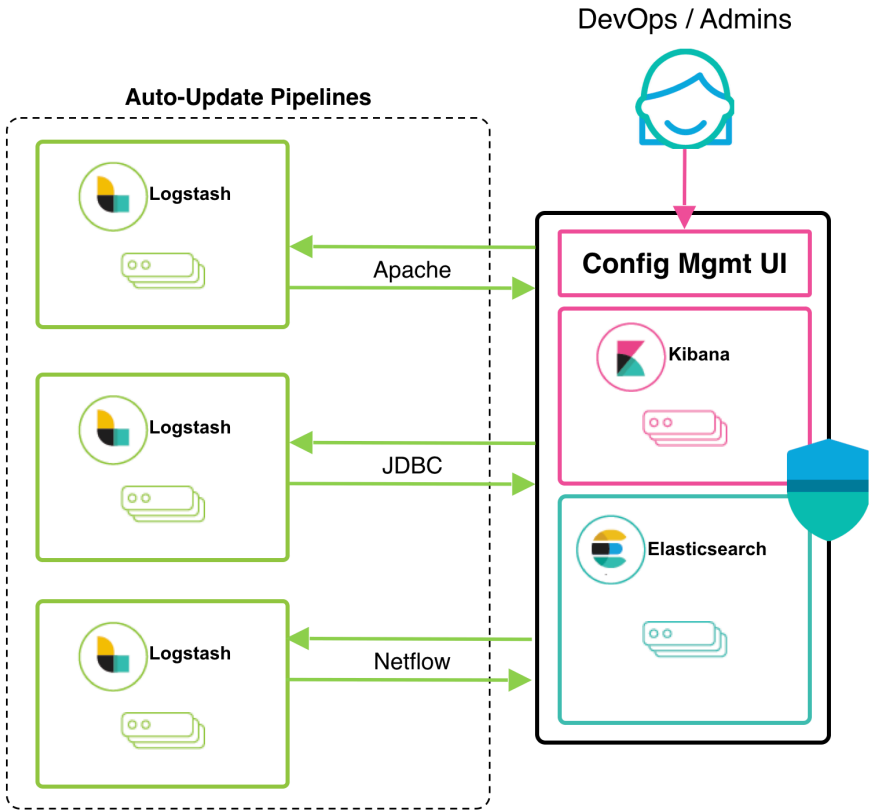
elastic

Monitoring & Management

Logstash

- Centralized monitoring (5.3)
- Centralized management (6.0)

** Near-term roadmap*



elastic

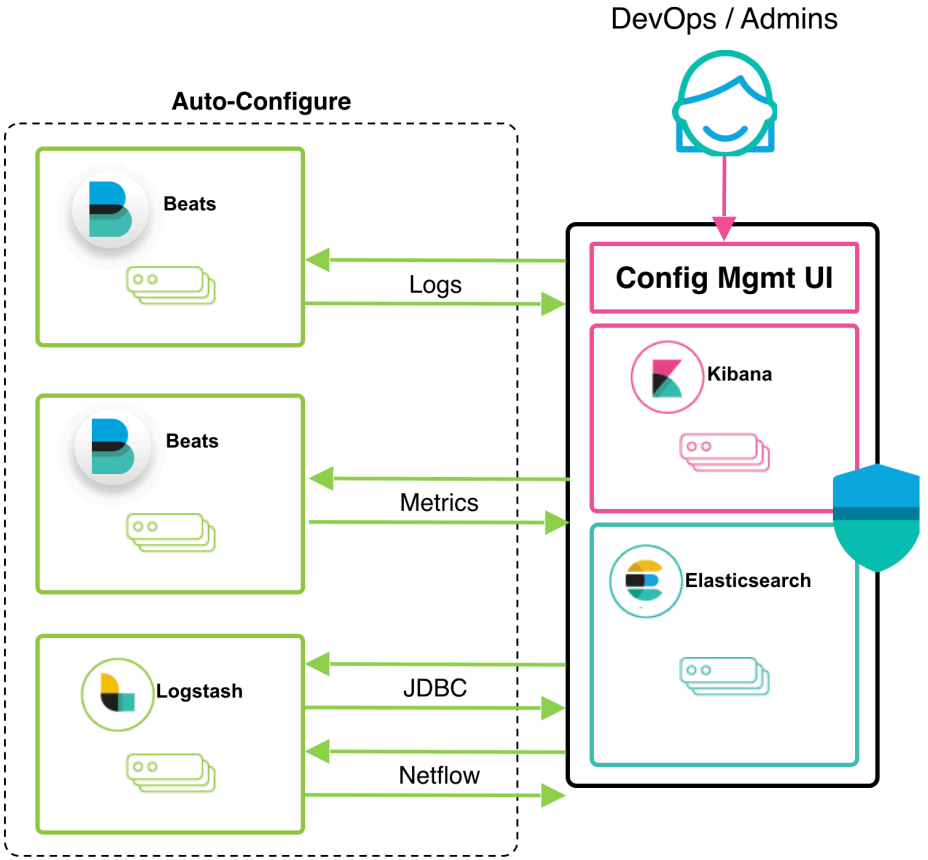
Monitoring & Management

Logstash

- Centralized monitoring (5.3)
- Centralized management (6.0)

Beats (Roadmap)

- Centralized monitoring
- Centralized management



elastic

Next steps

- Familiarize yourself with latest integrations (including in X-Pack)
- Watch UI roadmap for additional add-data workflows
- Come talk to us at the AMA booth

elastic{on}TOUR

Thank You

Find me at AMA booth or email david@elastic.co