

# ELK Stack





Somkiat Puisungnoen

Somkiat Puisungnoen

Update Info 1 View Activity Log 10+ ...

Timeline About Friends 3,138 Photos More

When did you work at Opendream? X

... 22 Pending Items

Intro

Software Craftsmanship

Software Practitioner at สยามชัมนาณกิจ พ.ศ. 2556

Agile Practitioner and Technical at SPRINT3r

Post Photo/Video Live Video Life Event

What's on your mind?

Public Post

Somkiat Puisungnoen 15 mins · Bangkok · ⚙️

Java and Bigdata





Page

Messages

Notifications 3

Insights

Publishing Tools

Settings

Help ▾



somkiat.cc

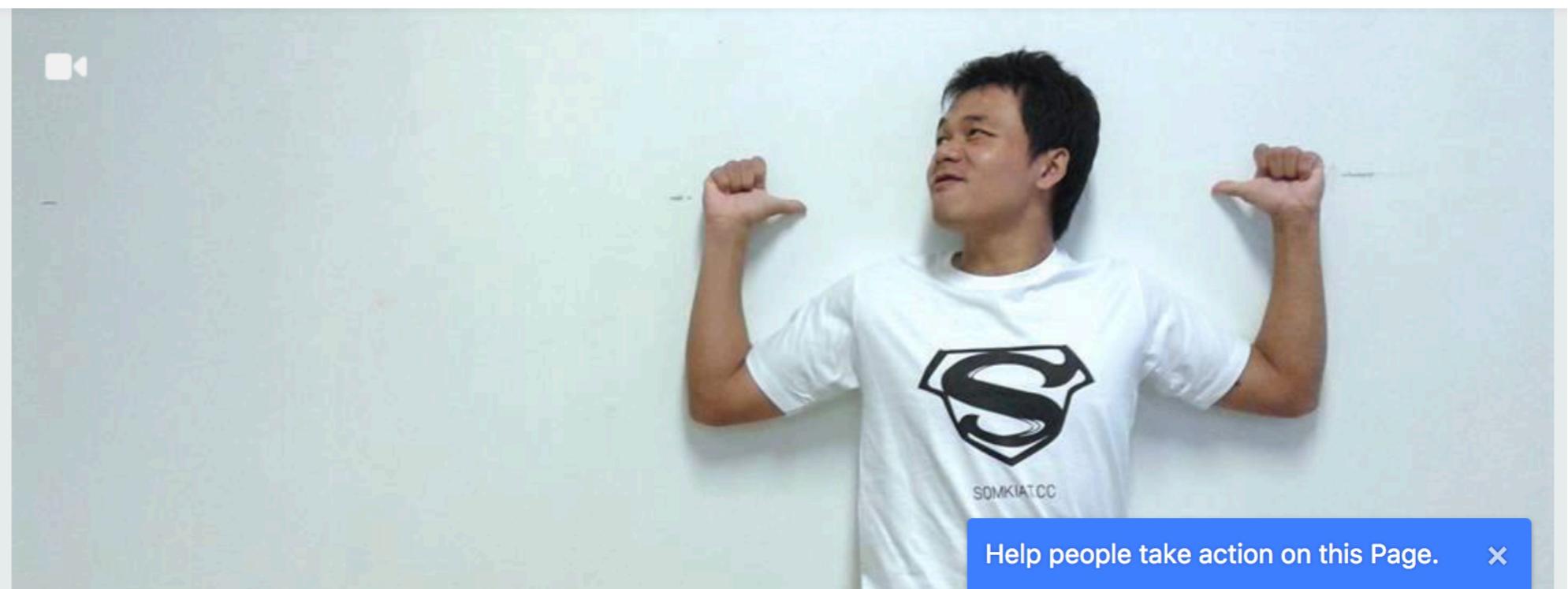
@somkiat.cc

Home

Posts

Videos

Photos



ELK Stack

4

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Agenda

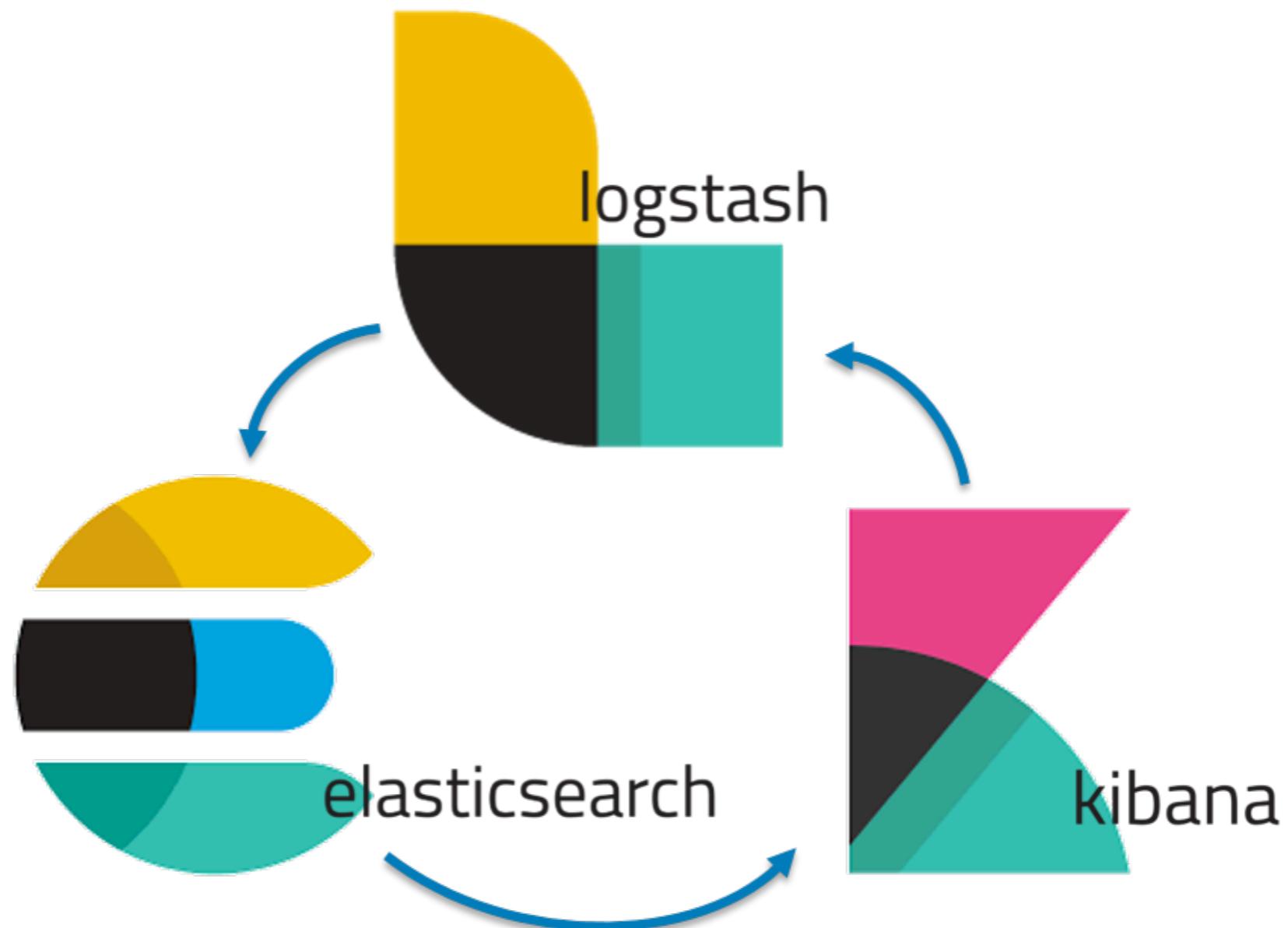
- ELK stack
- Introduction to Elasticsearch
- CRUD (Create, Read, Update, Delete)
- Search DSL (Domain Specific Language)
- Analyzer
- Mapping
- Aggregation



# Agenda

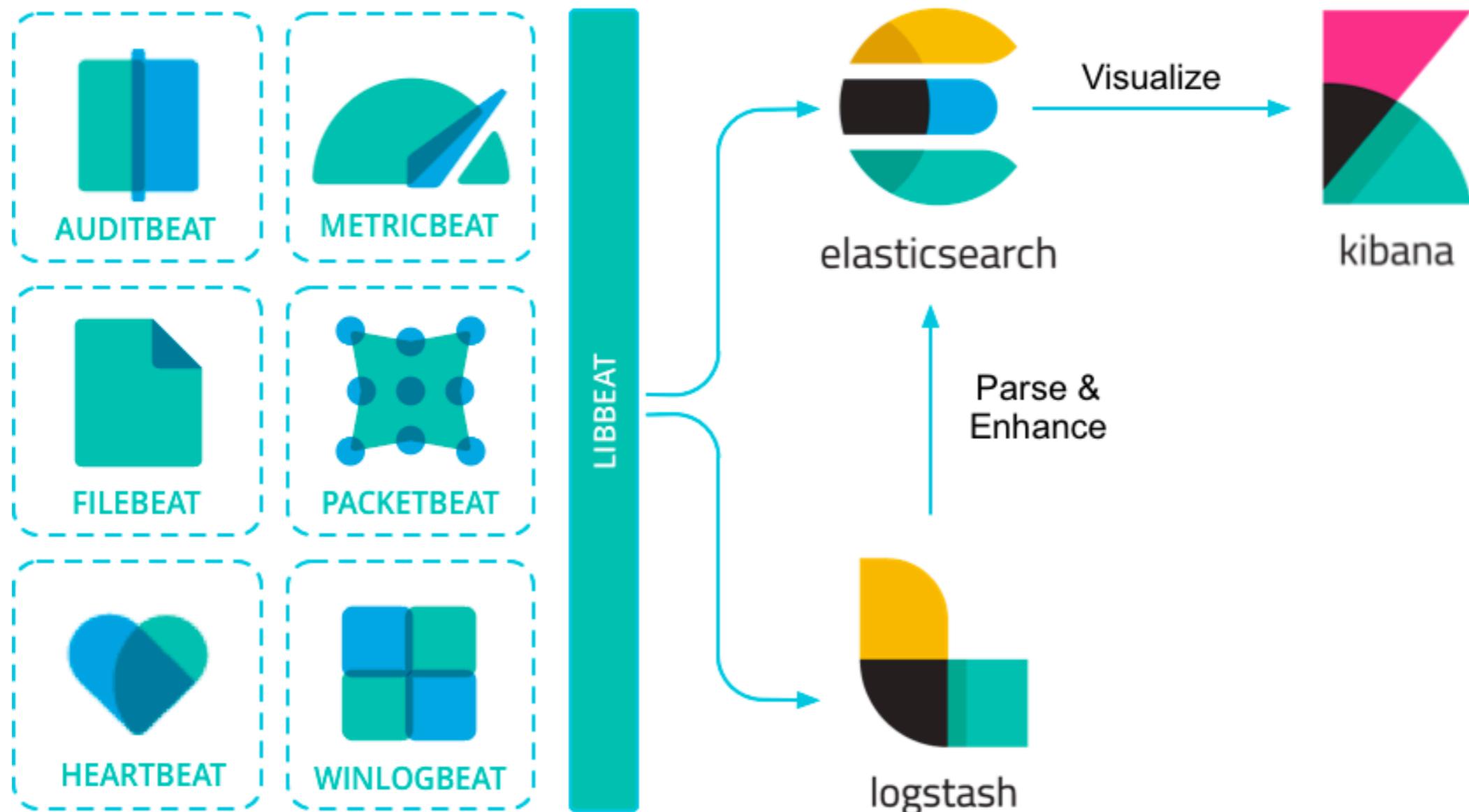
- Working with Kibana
- Useful features
  - Auto-suggestion
  - ngram algorithm
- Clustering management
- Design for scaling
- Working with Beat and Logstash







# Beat



<https://www.elastic.co/guide/en/beats/libbeat/current/index.html>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

# Beat

Purpose	Library
Audit data	Auditbeat
Log files	Filebeat
Cloud data	Functionbeat
Availability	Heartbeat
Metrics	Metricbeat
Network traffic	Packetbeat
Windows event logs	Winlogbeat

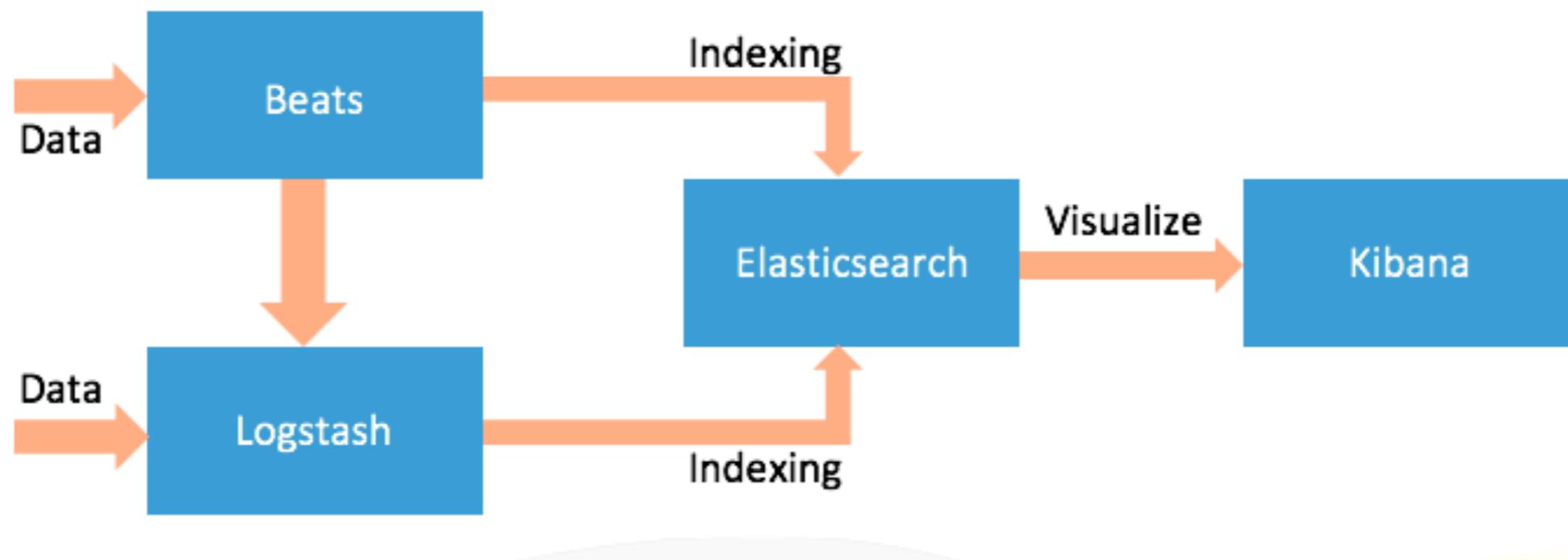
<https://www.elastic.co/guide/en/beats/libbeat/current/beats-reference.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# ELK stack



# Elasticsearch ?



# Elasticsearch

Search  
Analytic  
Real-time  
Distributed  
Scalability



# Distributed Search Engine

Open Source  
Document-based  
Based on **Apache Lucene**  
JSON over HTTP



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Distributed Search Engine



Compass



elasticsearch



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Apache Lucene

The screenshot shows the official Apache Lucene website. At the top, there's a navigation bar with a search bar containing "Search with Apache So" and a dropdown menu "select provider". Below the search bar are three buttons: "CORE (JAVA)", "SOLR", and "PYLUCENE". The main content area features the Apache Lucene logo (a green feather) on the left and the Solr logo (a red sunburst icon next to the word "Solr") on the right. A large green banner in the center reads "Ultra-fast Search Library and Server". Below this banner, a dark grey box contains the text "Apache Lucene and Solr set the standard for search and indexing performance". In the lower-left section, there's a heading "Welcome to Apache Lucene" and a list of what the project develops, including "Lucene Core", "Solr™", and "PyLucene". To the right of this text are two large download buttons: a green one for "Apache Lucene 7.5.0" and an orange one for "Apache Solr 7.5.0". At the bottom right, there's a link labeled "Projects".

Apache Lucene and Solr set the standard for search and indexing performance

## Welcome to Apache Lucene

The Apache Lucene™ project develops open-source search software, including:

- [Lucene Core](#), our flagship sub-project, provides Java-based indexing and search technology, as well as spellchecking, hit highlighting and advanced analysis/tokenization capabilities.
- [Solr™](#) is a high performance search server built using Lucene Core, with XML/HTTP and JSON/Python/Ruby APIs, hit highlighting, faceted search, caching, replication, and a web admin interface.
- [PyLucene](#) is a Python port of the Core project.

[DOWNLOAD](#)

Apache Lucene 7.5.0

[DOWNLOAD](#)

Apache Solr 7.5.0

[Projects](#)

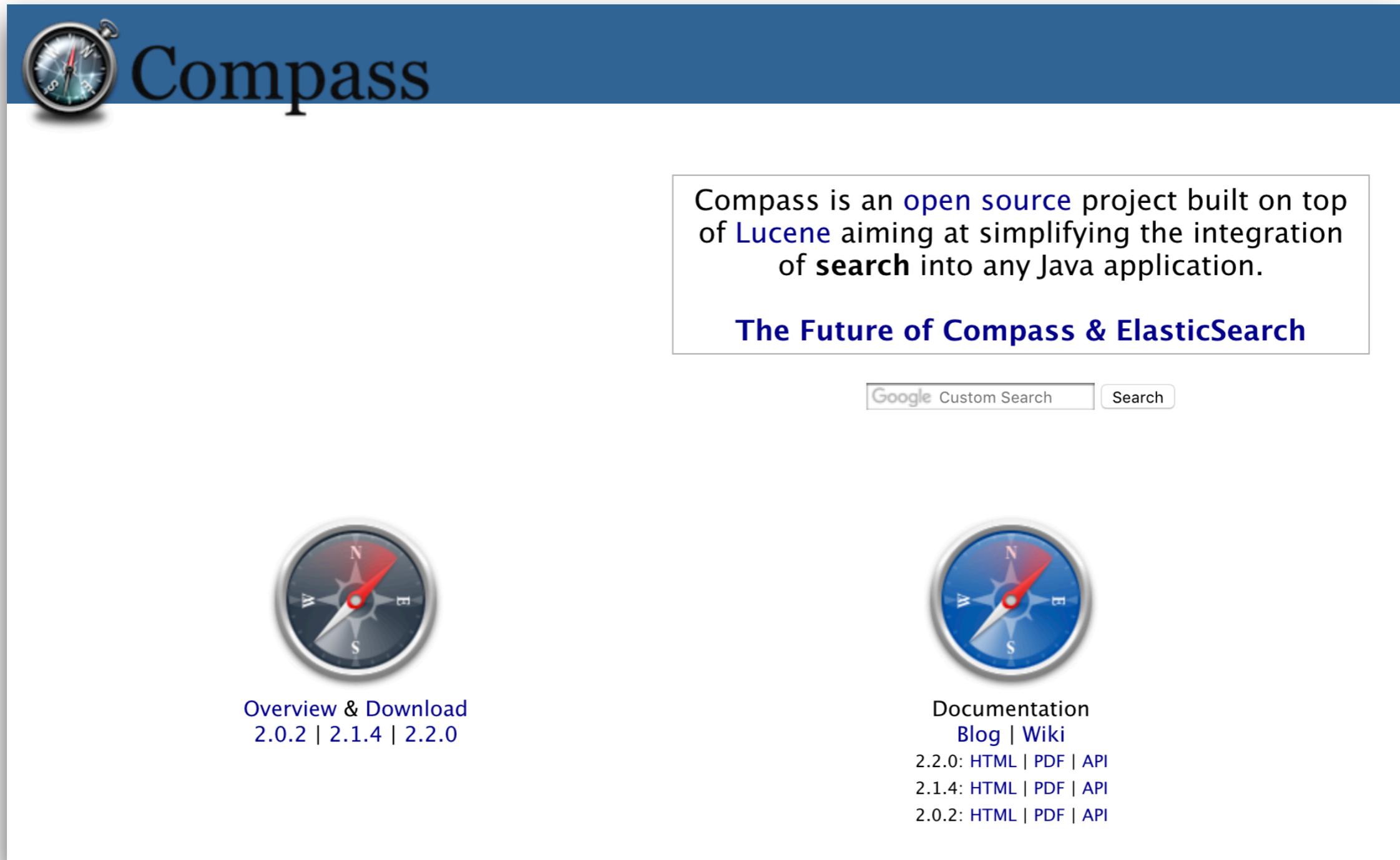
<http://lucene.apache.org/>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Compass



The screenshot shows the official website for the Compass project. At the top, there's a dark blue header bar with a compass icon on the left and the word "Compass" in white. Below the header is a large white main area. In the upper right of this area, there's a box containing text about the project. Further down, there are two large compass icons, one on each side of a central vertical line. Below each icon is a section of text.

Compass is an [open source](#) project built on top of [Lucene](#) aiming at simplifying the integration of [search](#) into any Java application.

**The Future of Compass & ElasticSearch**

[Google Custom Search](#) [Search](#)

 Overview & Download  
2.0.2 | 2.1.4 | 2.2.0

 Documentation  
[Blog](#) | [Wiki](#)  
2.2.0: [HTML](#) | [PDF](#) | [API](#)  
2.1.4: [HTML](#) | [PDF](#) | [API](#)  
2.0.2: [HTML](#) | [PDF](#) | [API](#)

<http://www.compass-project.org/>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# **Document based**

**JSON (JavaScript Object Notation)**

**Dynamic Schema (Schema-less)**

**Some relationship (nested, parent/child)**



# StackOverflow Question

```
{  
  "items": [  
    {  
      "owner": {  
        "reputation": 13,  
        "user_id": 9796344,  
        "user_type": "registered",  
        "profile_image": "",  
        "display_name": "Cherry",  
        "link": "https://stackoverflow.com/users/9796344/cherry"  
      },  
      "score": 0,  
      "last_activity_date": 1528986761,  
      "creation_date": 1528986761,  
      "post_type": "question",  
      "post_id": 50859951,  
      "link": "https://stackoverflow.com/q/50859951"  
    }  
  ],  
  "has_more": false,  
  "quota_max": 10000,  
  "quota_remaining": 9986  
}
```

<https://api.stackexchange.com/docs/posts-by-ids>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Ranking from DB Engine (2018)

350 systems in ranking, June 2019

Rank			DBMS	Database Model	Score		
Jun 2019	May 2019	Jun 2018			Jun 2019	May 2019	Jun 2018
1.	1.	1.	Oracle	Relational, Multi-model	1299.21	+13.67	-12.04
2.	2.	2.	MySQL	Relational, Multi-model	1223.63	+4.67	-10.06
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	1087.76	+15.57	+0.03
4.	4.	4.	PostgreSQL	Relational, Multi-model	476.62	-2.27	+65.95
5.	5.	5.	MongoDB	Document	403.90	-4.17	+60.12
6.	6.	6.	IBM Db2	Relational, Multi-model	172.20	-2.24	-13.44
7.	7.	8.	Elasticsearch	Search engine, Multi-model	148.82	+0.20	+17.78
8.	8.	7.	Redis	Key-value, Multi-model	146.13	-2.28	+9.83
9.	9.	9.	Microsoft Access	Relational	141.01	-2.77	+10.02
10.	10.	10.	Cassandra	Wide column	125.18	-0.54	+5.97

<https://db-engines.com/en/ranking>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

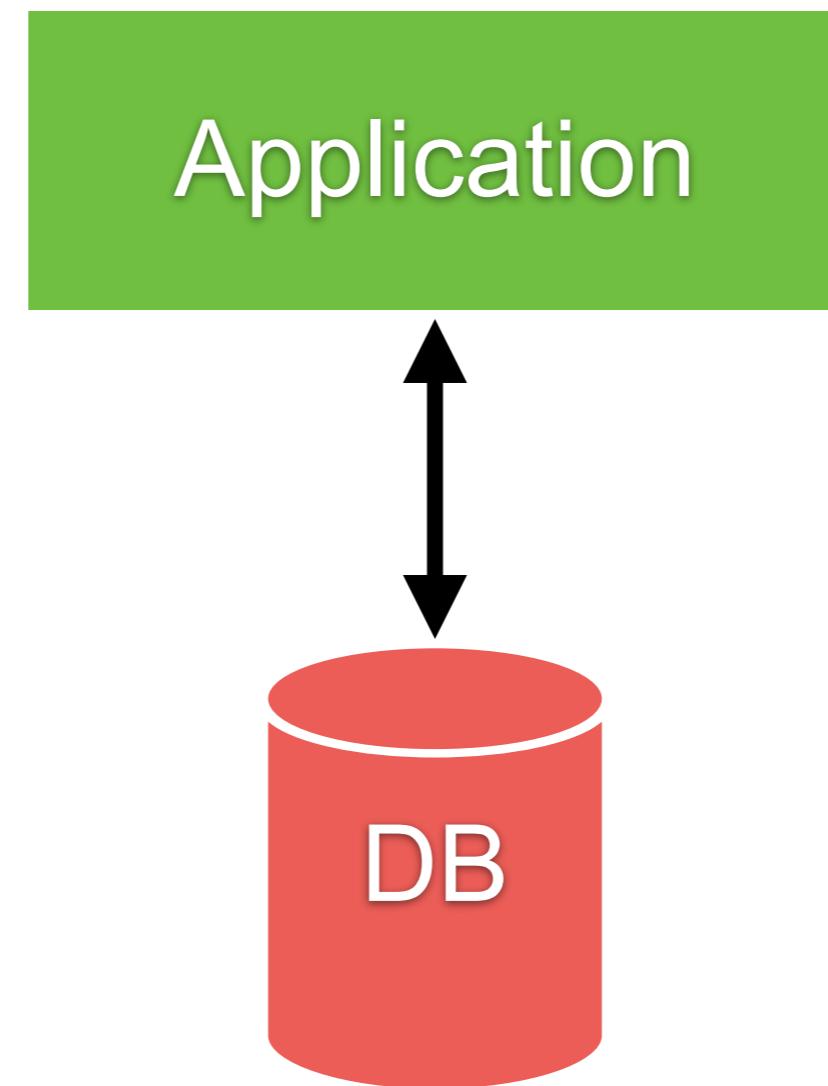
# Use cases

Security/log analytics  
Marketing  
Operations  
Searching data



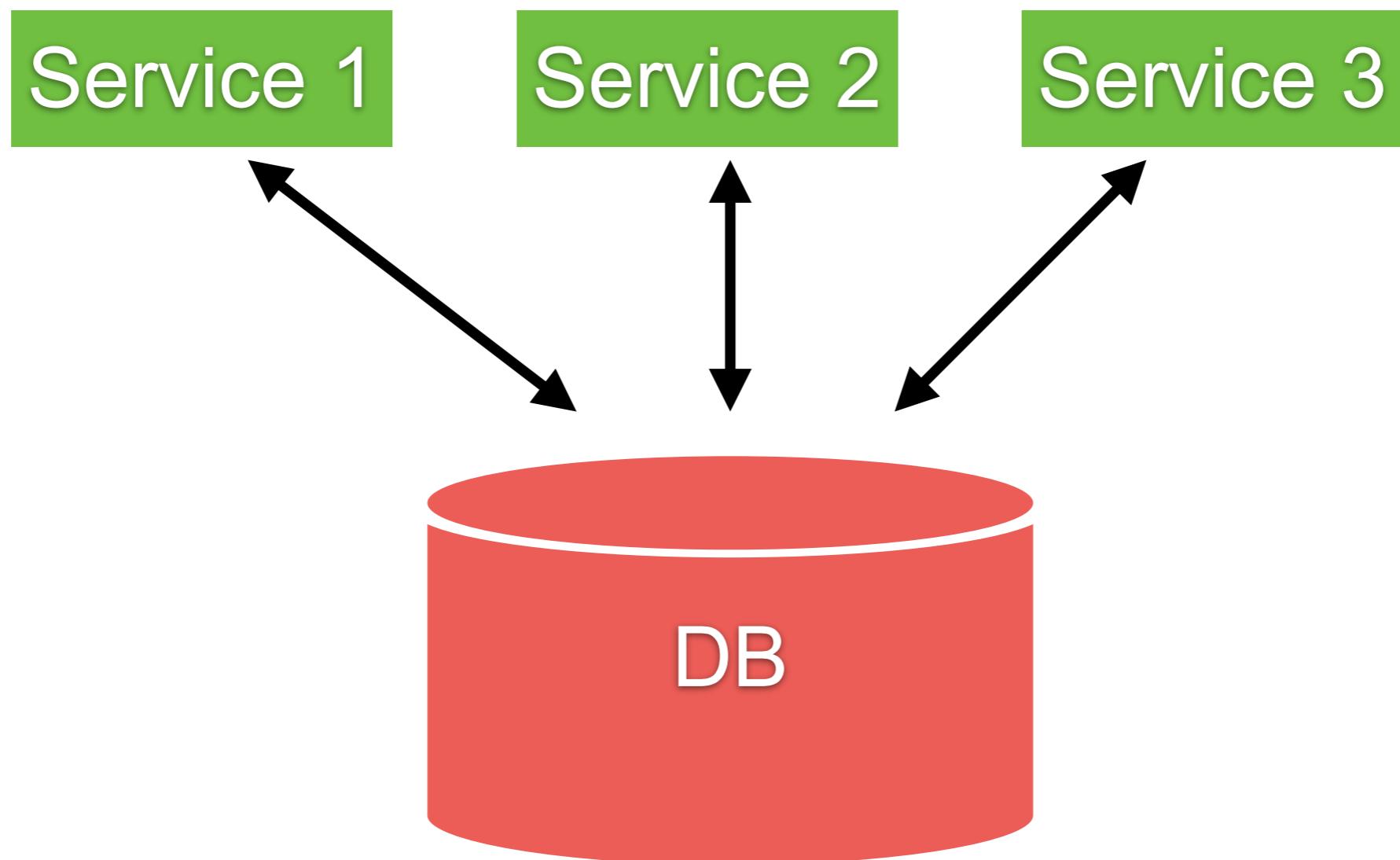
# Problem ?

Single/Centralize database



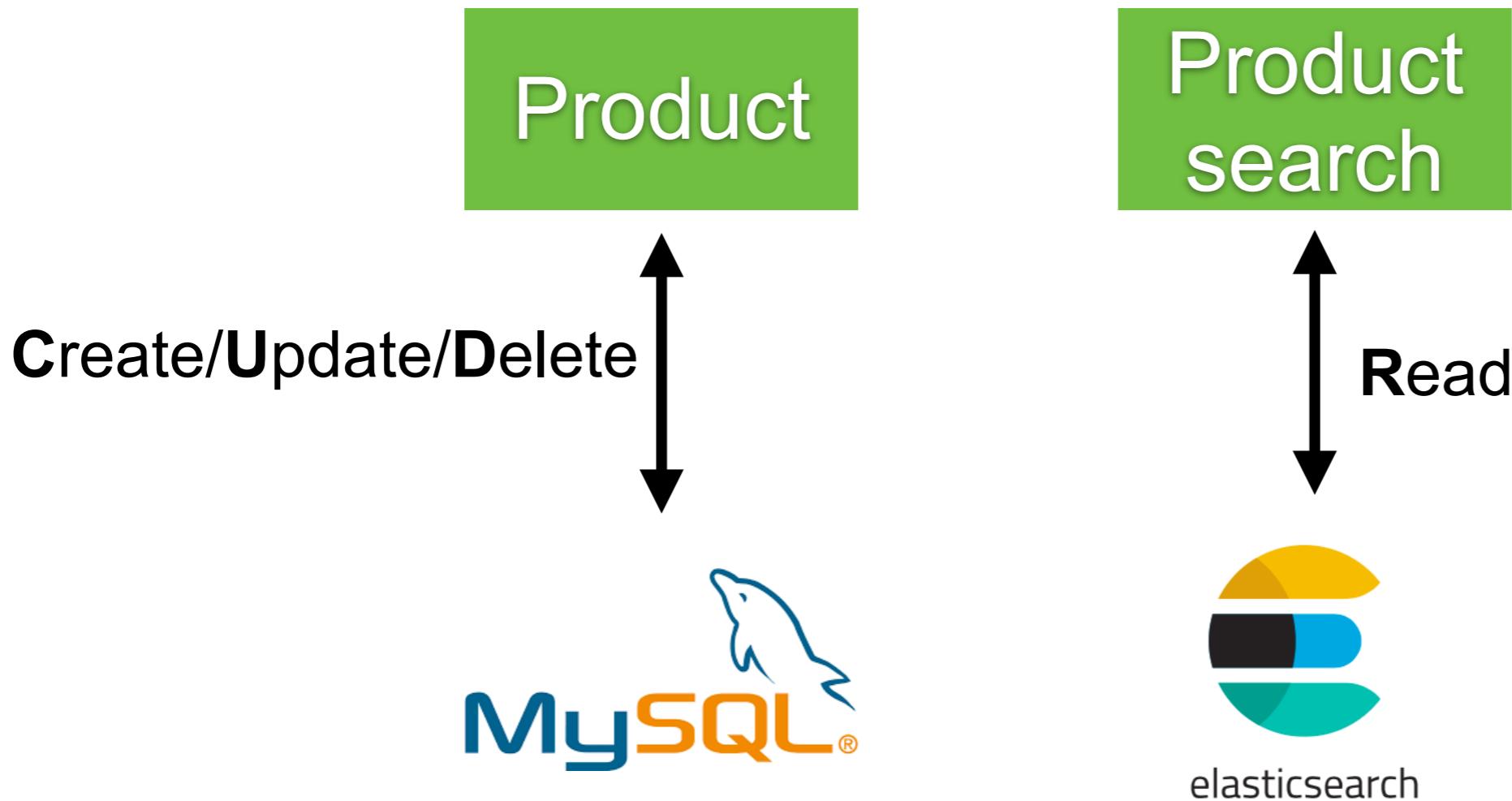
# Problem ?

Single/Centralize database



# Separate data for read and write

For example MySQL to write, Elasticsearch to search



# Let's start



# Installation

Elasticsearch  
Kibana



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Install Elasticsearch



# Elasticsearch

Required Java 8  
JDK and Open JDK  
Need \$JAVA\_HOME



# JAVA\_HOME

\$echo %JAVA\_HOME% //For Windows

\$echo \$JAVA\_HOME // for Linux/Mac



# Start Elasticsearch

./bin/elasticsearch

```
[0g8-71W] loaded module [reindex]
[0g8-71W] loaded module [repository-url]
[0g8-71W] loaded module [transport-netty4]
[0g8-71W] loaded module [tribe]
[0g8-71W] no plugins loaded
[0g8-71W] using discovery type [zen]
initialized
[0g8-71W] starting ...
[0g8-71W] publish_address {127.0.0.1:9300},
[0g8-71W] recovered [0] indices into cluster_state
transport] [0g8-71W] publish_address {127.0.0.1:9200},
```



# Configuration files

`elasticsearch.yml`

`jvm.options`

`log4j2.properties`

Default : `$ES_HOME/config`

Custom config path : `$ES_PATH_CONF`



# Default of Memory

1 GB !!! (Java need more memory)

```
 ] [DW5j42N] JVM arguments [-Xms1g, -Xmx1g, -  
ction=75, -XX:+UseCMSInitiatingOccupancyOnly, -XX:  
Dfile.encoding=UTF-8, -Djna.nosys=true, -XX:-Omit  
.netty.noKeySetOptimization=true, -Dio.netty.recy  
led=false, -Dlog4j2.disable.jmx=true, -Djava.io.t  
T/elasticsearch.G4kbTLZn, -XX:+HeapDumpOnOutOfMem  
s_err_pid%p.log, -Xlog:gc*,gc+age=trace,safepoint  
ize=64m, -Djava.locale.providers=COMPAT, -XX:UseA
```



# Config of JVM

`$ES_HOME/config/jvm.options`

```
# Xms represents the initial size of total heap space  
# Xmx represents the maximum size of total heap space  
  
-Xms1g  
-Xmx1g
```



# Default plugins

```
[o.e.p.PluginsService      ] [DW5j42N] loaded module [aggs-matrix-stats]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [analysis-common]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [ingest-common]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [lang-expression]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [lang-mustache]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [lang-painless]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [mapper-extras]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [parent-join]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [percolator]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [rank-eval]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [reindex]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [repository-url]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [transport-netty4]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [tribe]
```



# Install X-Pack by default

```
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-core]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-deprecation]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-graph]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-logstash]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-ml]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-monitoring]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-rollup]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-security]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-sql]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-upgrade]
[o.e.p.PluginsService      ] [DW5j42N] loaded module [x-pack-watcher]
[o.e.p.PluginsService      ] [DW5j42N] no plugins loaded
```

<https://www.elastic.co/guide/en/elasticsearch/reference/current/installing-xpack-es.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# X-Pack ?

Elastic Stack Extension  
Security  
Monitoring  
Alerting  
Reporting  
Machine Learning



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Licence

	FREE OPEN SOURCE	BASIC	GOLD	PLATINUM
Elasticsearch	<a href="#">Download</a>		<a href="#">Request Info</a>	<a href="#">Request Info</a>
✓ Scalability & Resiliency	✓	✓	✓	✓
✓ Query & Analytics	✓	✓	✓	✓
✓ Data Enrichment	✓	✓	✓	✓
✓ Management & Tooling	✓	✓	✓	✓
✓ Security			✓	✓
✓ Alerting			✓	✓
✓ Machine Learning				✓
Kibana				
✓ Explore & Visualize	✓	✓	✓	✓
✓ Stack Management & Tooling	✓	✓	✓	✓
✓ Stack Monitoring		✓	✓	✓

<https://www.elastic.co/subscriptions>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Hello Elasticsearch

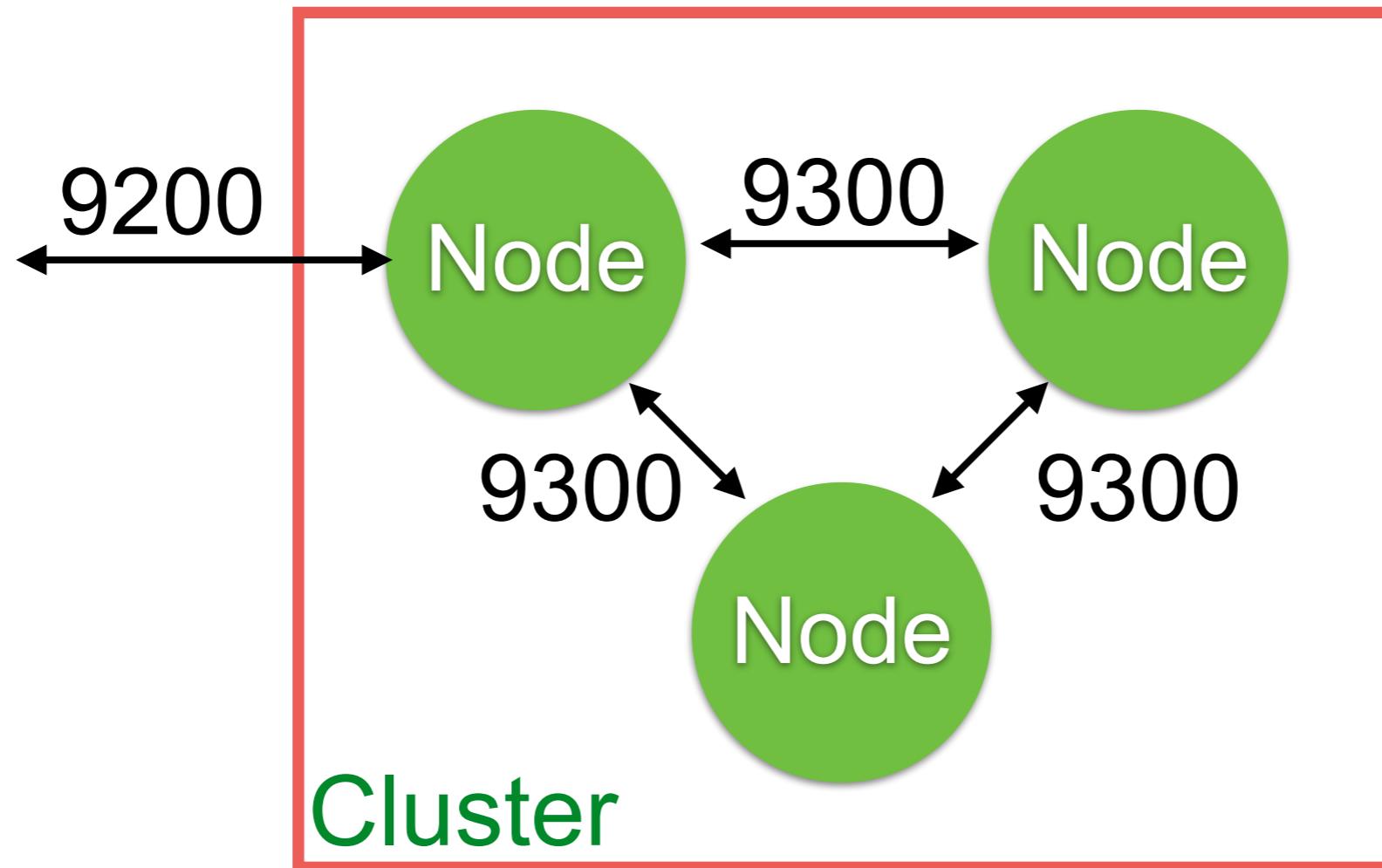
<http://localhost:9200/>

```
{  
    name: "Somkiats-MacBook-Pro",  
    cluster_name: "elasticsearch",  
    cluster_uuid: "gglAIg0NRHyn4AefFR6law",  
    - version: {  
        number: "7.1.1",  
        build_flavor: "default",  
        build_type: "tar",  
        build_hash: "7a013de",  
        build_date: "2019-05-23T14:04:00.380842Z",  
        build_snapshot: false,  
        lucene_version: "8.0.0",  
        minimum_wire_compatibility_version: "6.8.0",  
        minimum_index_compatibility_version: "6.0.0-beta1"  
    },  
    tagline: "You Know, for Search"  
}
```



# Ports of Elasticsearch

RESTful API with JSON Over HTTP (9200)  
Java API (9300)



# Name of node and cluster

```
{  
  name: "Somkiats-MacBook-Pro",  
  cluster_name: "elasticsearch",  
  cluster_uuid: "gglAIg0NRHyn4AefFR61aw",  
  - version: {  
      number: "7.1.1",  
      build_flavor: "default",  
      build_type: "tar",  
      build_hash: "7a013de",  
      build_date: "2019-05-23T14:04:00.380842Z",  
      build_snapshot: false,  
      lucene_version: "8.0.0",  
      minimum_wire_compatibility_version: "6.8.0",  
      minimum_index_compatibility_version: "6.0.0-beta1"  
    },  
  tagline: "You Know, for Search"  
}
```



# Name of node and cluster

\$ES\_HOME/config/elasticsearch.yml

```
# ----- Cluster -----  
#  
# Use a descriptive name for your cluster:  
#  
cluster.name: my-application  
#  
# ----- Node -----  
#  
# Use a descriptive name for the node:  
#  
node.name: node-1  
#  
# Add custom attributes to the node:  
#  
#node.attr.rack: r1  
#
```



# Change in Elasticsearch 7.x

Default name = Hostname

<https://www.elastic.co/guide/en/elasticsearch/reference/master/breaking-changes-7.0.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Try to change and restart !!!



```
{  
  name: "Somkiats-MacBook-Pro",  
  cluster_name: "elasticsearch",  
  cluster_uuid: "gglAIg0NRHyn4AefFR61aw",  
  - version: {  
      number: "7.1.1",  
      build_flavor: "default",  
      build_type: "tar",  
      build_hash: "7a013de",  
      build_date: "2019-05-23T14:04:00.380842Z",  
      build_snapshot: false,  
      lucene_version: "8.0.0",  
      minimum_wire_compatibility_version: "6.8.0",  
      minimum_index_compatibility_version: "6.0.0-beta1"  
    },  
  tagline: "You Know, for Search"  
}
```



# Compatibility of DSL and Index



```
{  
  name: "Somkiats-MacBook-Pro",  
  cluster_name: "elasticsearch",  
  cluster_uuid: "gglAIg0NRHyn4AefFR61aw",  
  - version: {  
      number: "7.1.1",  
      build_flavor: "default",  
      build_type: "tar",  
      build_hash: "7a013de",  
      build_date: "2019-05-23T14:04:00.380842Z",  
      build_snapshot: false,  
      lucene_version: "8.0.0",  
      minimum_wire_compatibility_version: "6.8.0",  
      minimum_index_compatibility_version: "6.0.0-beta1"  
    },  
  tagline: "You Know, for Search"  
}
```

DSL version

Index version



```
{  
  name: "Somkiats-MacBook-Pro",  
  cluster_name: "elasticsearch",  
  cluster_uuid: "gglAIg0NRHyn4AefFR61aw",  
  - version: {  
      number: "7.1.1",  
      build_flavor: "default",  
      build_type: "tar",  
      build_hash: "7a013de",  
      build_date: "2019-05-23T14:04:00.380842Z",  
      build_snapshot: false,  
      lucene_version: "8.0.0",  
      minimum_wire_compatibility_version: "6.8.0",  
      minimum_index_compatibility_version: "6.0.0-beta1"  
    },  
  tagline: "You Know, for Search"  
}
```

Index version



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Health of cluster

[http://localhost:9200/\\_cluster/health](http://localhost:9200/_cluster/health)

```
{  
  "cluster_name": "elasticsearch",  
  "status": "green",  
  "timed_out": false,  
  "number_of_nodes": 1,  
  "number_of_data_nodes": 1,  
  "active_primary_shards": 0,  
  "active_shards": 0,  
  "relocating_shards": 0,  
  "initializing_shards": 0,  
  "unassigned_shards": 0,  
  "delayed_unassigned_shards": 0,  
  "number_of_pending_tasks": 0,  
  "number_of_in_flight_fetch": 0,  
  "task_max_waiting_in_queue_millis": 0,  
  "active_shards_percent_as_number": 100.0  
}
```



# Health of cluster

Status	Meaning
Green	All shards are allocated
Yellow	Primary shard is allocated, but replicas are not
Red	Shard not allocated in the cluster



# cat APIs

`http://localhost:9200/_cat`

```
=^.^=
/_cat/allocation
/_cat/shards
/_cat/shards/{index}
/_cat/master
/_cat/nodes
/_cat/tasks
/_cat/indices
/_cat/indices/{index}
/_cat/segments
/_cat/segments/{index}
/_cat/count
/_cat/count/{index}
/_cat/recovery
/_cat/recovery/{index}
/_cat/health
/_cat/pending_tasks
/_cat/aliases
/_cat/aliases/{alias}
```

<https://www.elastic.co/guide/en/elasticsearch/reference/current/cat.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# cat APIs

`http://localhost:9200/_cat/nodes?v`

ip	heap.percent	ram.percent	cpu	load_1m	load_5m	load_15m	node.role	master	name
127.0.0.1	20	100	7	1.98			mdi	*	DW5j42N



# Install Kibana



# Start Kibana

```
[status][plugin:xpack_main@6.4.2] Status changed from yellow to green - Ready
[status][plugin:searchprofiler@6.4.2] Status changed from yellow to green - Ready
[status][plugin:ml@6.4.2] Status changed from yellow to green - Ready
[status][plugin:tilemap@6.4.2] Status changed from yellow to green - Ready
[status][plugin:watcher@6.4.2] Status changed from yellow to green - Ready
[status][plugin:index_management@6.4.2] Status changed from yellow to green - Ready

[status][plugin:graph@6.4.2] Status changed from yellow to green - Ready
[status][plugin:grokdebugger@6.4.2] Status changed from yellow to green - Ready
[status][plugin:logstash@6.4.2] Status changed from yellow to green - Ready
[status][plugin:reporting@6.4.2] Status changed from yellow to green - Ready
[kibana-monitoring][monitoring-ui] Starting monitoring stats collection
[status][plugin:security@6.4.2] Status changed from yellow to green - Ready
[license][xpack] Imported license information from Elasticsearch for the [monitor]
tatus: active
[listening][server][http] Server running at http://localhost:5601
```



# Hello Kibana

<http://localhost:5601/>

The image shows the Kibana landing page. On the left is a vertical sidebar with icons for Kibana, APM, Metrics, Security, Visualize, Discover, and Saved Objects. The main content area is divided into several sections:

- Add Data to Kibana**:
  - APM**: APM automatically collects in-depth performance metrics and errors from inside your applications. [Add APM](#)
  - Logging**: Ingest logs from popular data sources and easily visualize in preconfigured dashboards. [Add log data](#)
  - Metrics**: Collect metrics from the operating system and services running on your servers. [Add metric data](#)
  - Security analytics**: Centralize security events for interactive investigation in ready-to-go visualizations. [Add security events](#)
- Data already in Elasticsearch?**: [Set up index patterns](#)
- Visualize and Explore Data**:
  - Dashboard**: Display and share a collection of visualizations and saved searches.
  - Timelion**: Use an expression language to analyze time series data
  - Discover**: Interactively explore your data by querying and filtering raw documents.
  - Visualize**: Create visualizations and aggregate data stores in your
- Manage and Administer the Elastic Stack**:
  - Console**: Skip cURL and use this JSON interface to work with your data directly.
  - Index Patterns**: Manage the index patterns that help retrieve your data from Elasticsearch.
  - Saved Objects**: Import, export, and manage your saved searches,



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Using Dev Tools

The screenshot shows the Kibana interface with a sidebar on the left and a main content area on the right.

**Left Sidebar:**

- Kibana logo
- Discover
- Visualize
- Dashboard
- Timelion
- APM
- Dev Tools** (highlighted with a red box)
- Monitoring
- Management

**Main Content Area:**

## Dev Tools

### Welcome to Console

#### Quick intro to the UI

The Console UI is split into two panes: an editor pane (left) and a response pane (right). Use the editor to type requests and submit them in the response pane on the right side.

Console understands requests in a compact format, similar to cURL:

```
1 # index a doc
2 PUT index/type/1
3 {
4   "body": "here"
5 }
6
7 # and get it ...
8 GET index/type/1
```

While typing a request, Console will make suggestions which you can then accept by hitting Enter/Tab. These suggestions are made based on types.

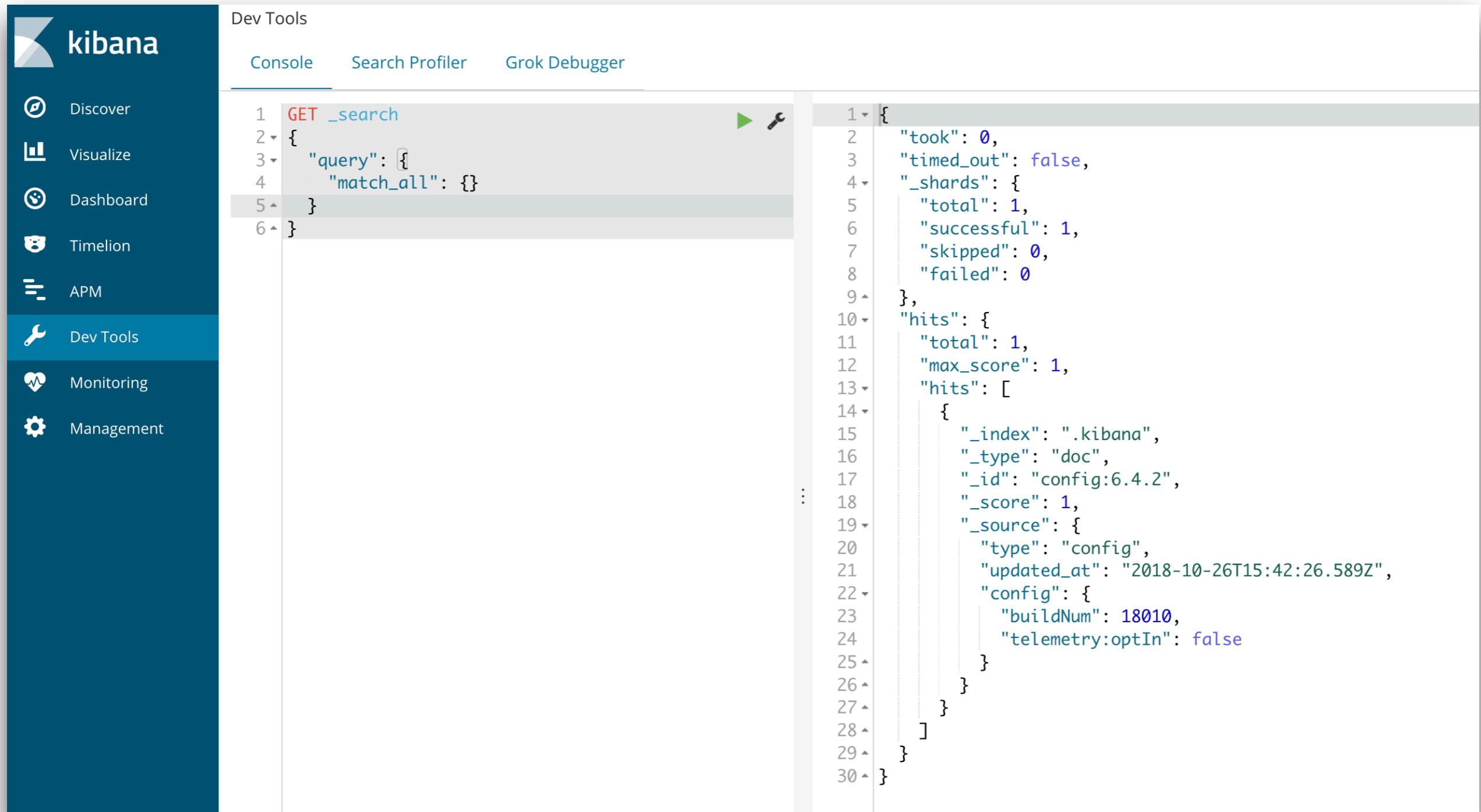
#### A few quick tips, while I have your attention

- Submit requests to ES using the green triangle button.
- Use the wrench menu for other useful things.
- You can paste requests in cURL format and they will be translated to the Console syntax.
- You can resize the editor and output panes by dragging the separator between them.
- Study the keyboard shortcuts under the Help button. Good stuff in there!

[Get to work](#)



# Ready to start



The screenshot shows the Kibana Dev Tools interface. On the left is a sidebar with icons for Discover, Visualize, Dashboard, Timelion, APM, Dev Tools (which is selected), Monitoring, and Management. The main area is titled "Dev Tools" and contains three tabs: Console, Search Profiler, and Grok Debugger. The "Console" tab is active, displaying a code editor with a GET \_search request and its JSON response. The request is:1 GET \_search
2 {
3 "query": {
4 "match\_all": {}
5 }
6 }The response is:1 {
2 "took": 0,
3 "timed\_out": false,
4 "\_shards": {
5 "total": 1,
6 "successful": 1,
7 "skipped": 0,
8 "failed": 0
9 },
10 "hits": {
11 "total": 1,
12 "max\_score": 1,
13 "hits": [
14 {
15 "\_index": ".kibana",
16 "\_type": "doc",
17 "\_id": "config:6.4.2",
18 "\_score": 1,
19 "\_source": {
20 "type": "config",
21 "updated\_at": "2018-10-26T15:42:26.589Z",
22 "config": {
23 "buildNum": 18010,
24 "telemetry:optIn": false
25 }
26 }
27 }
28 ]
29 }
30 }



# Elasticsearch architecture



# Basic concepts

Cluster

Node

Shard

Replica

Gateway

Index

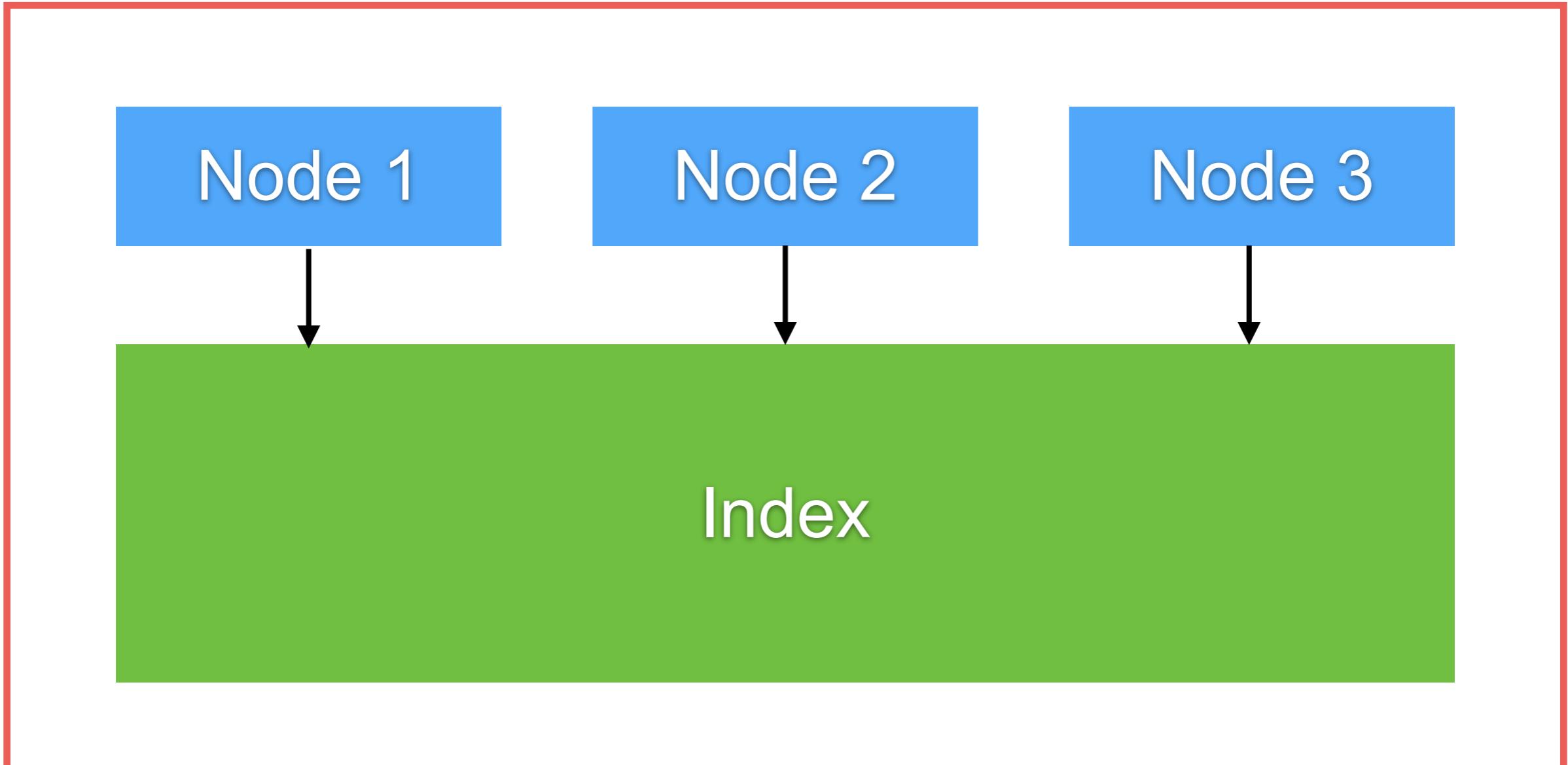
Document

Type

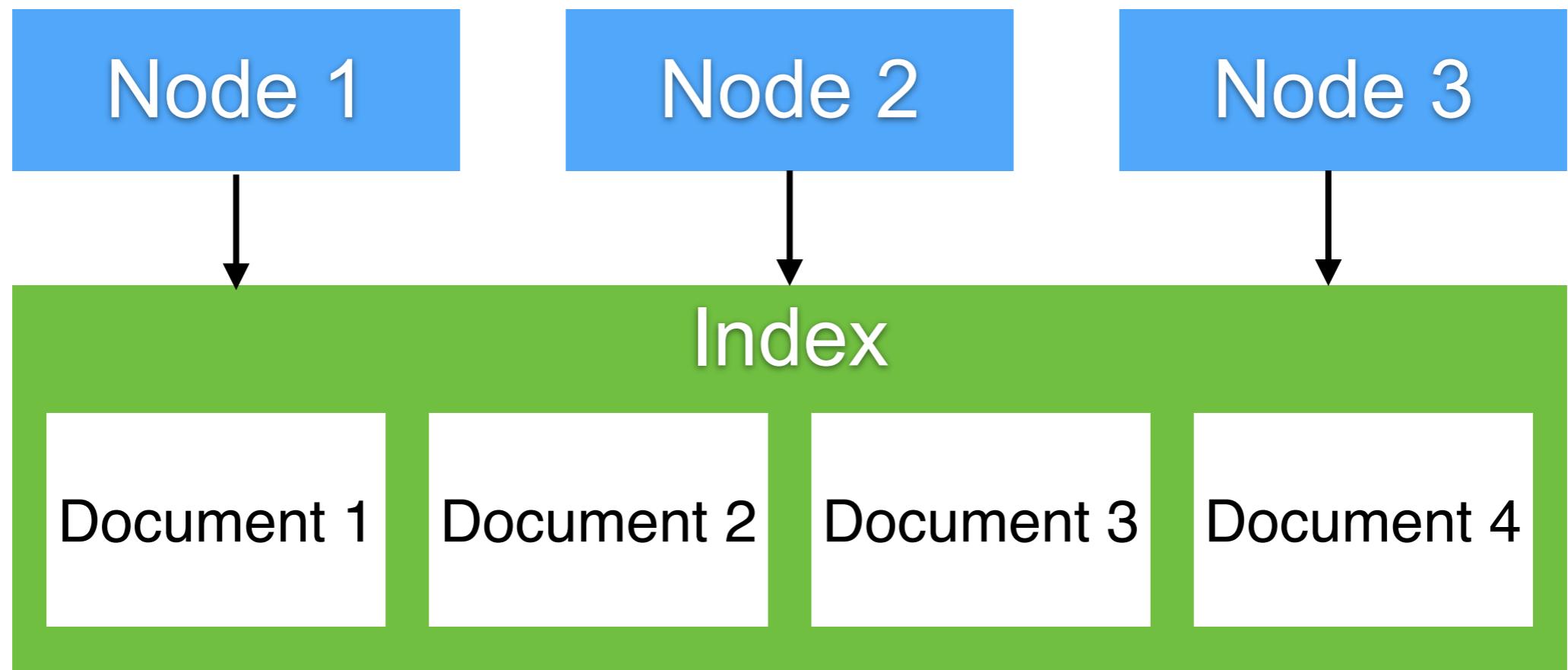
Mapping



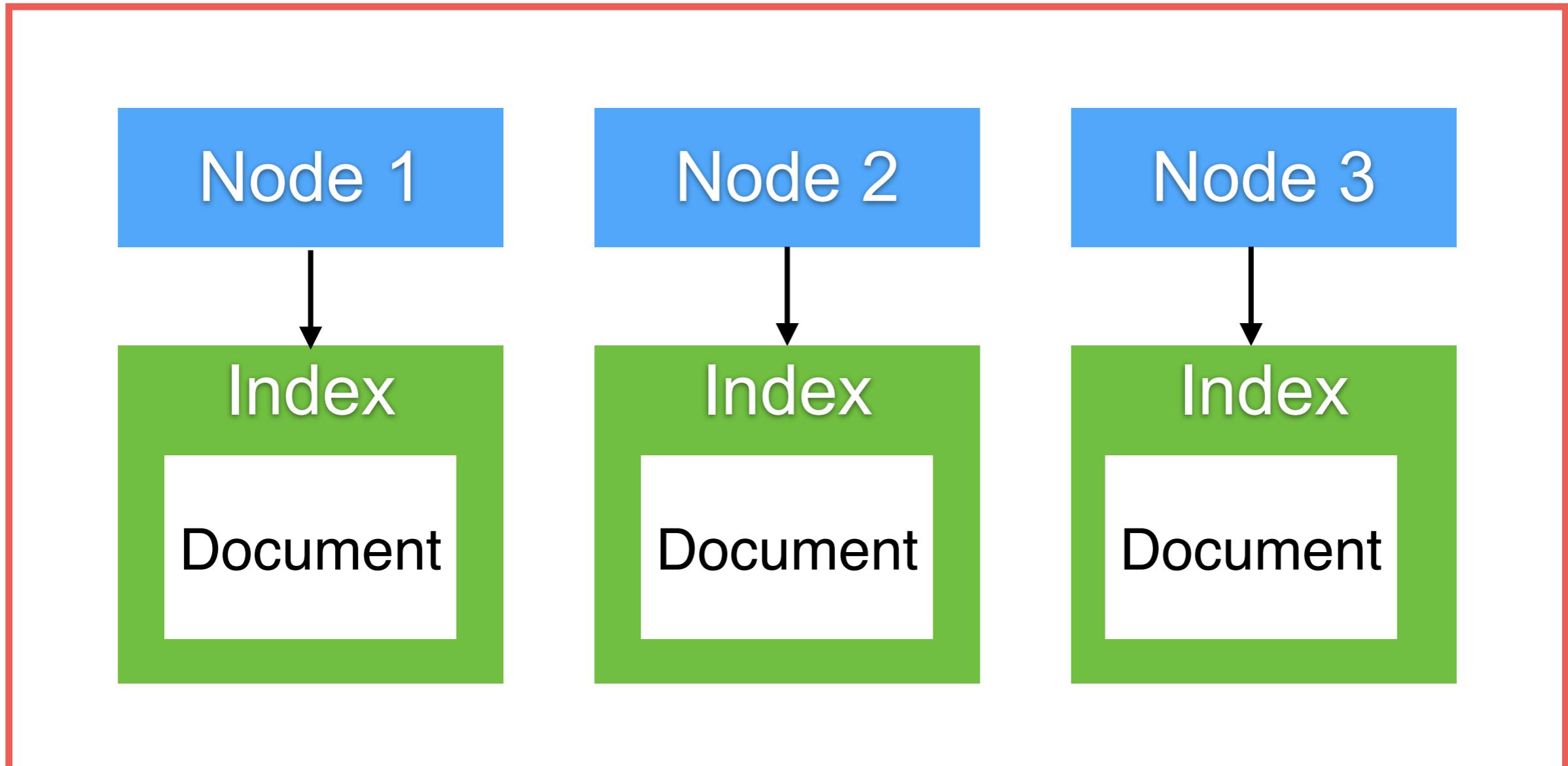
# Cluster



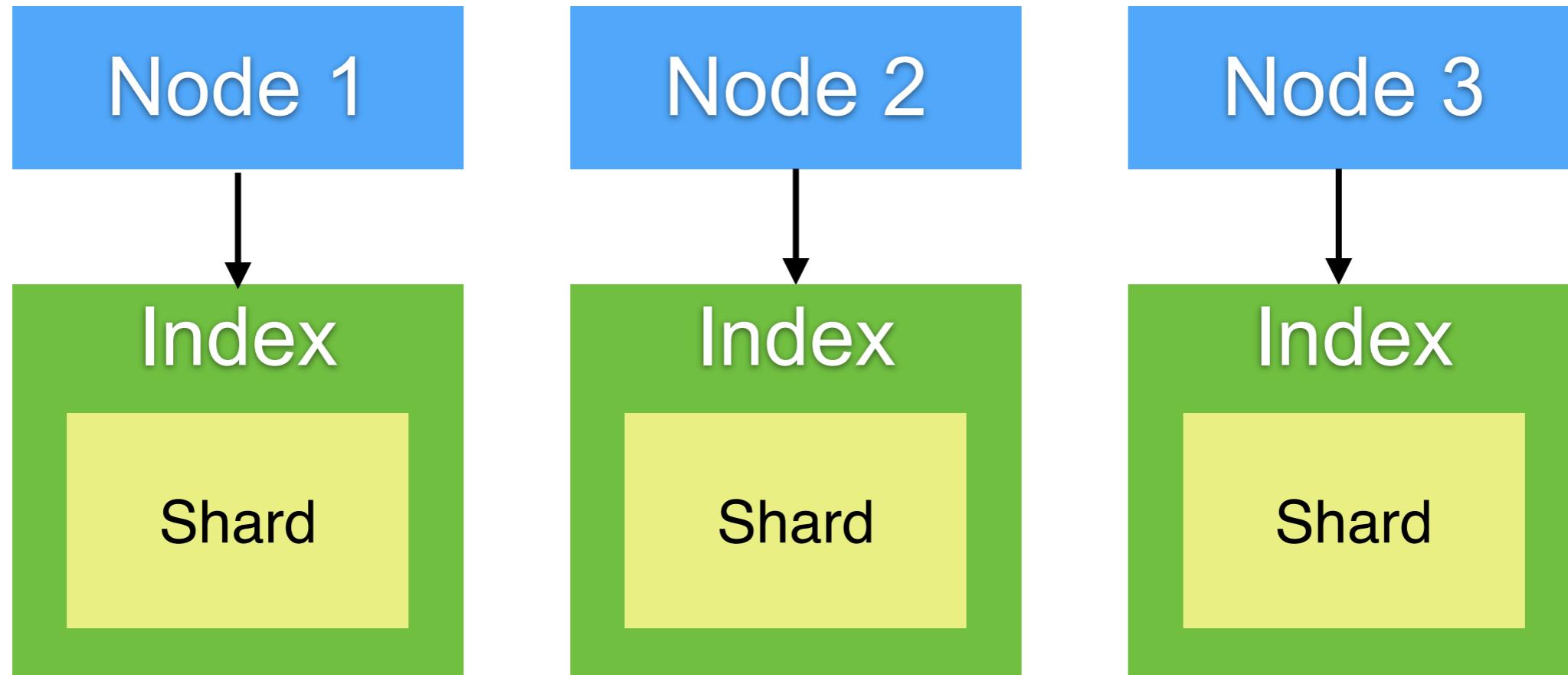
# Documents !!



# Distributed database



# Design for scale (Shard)



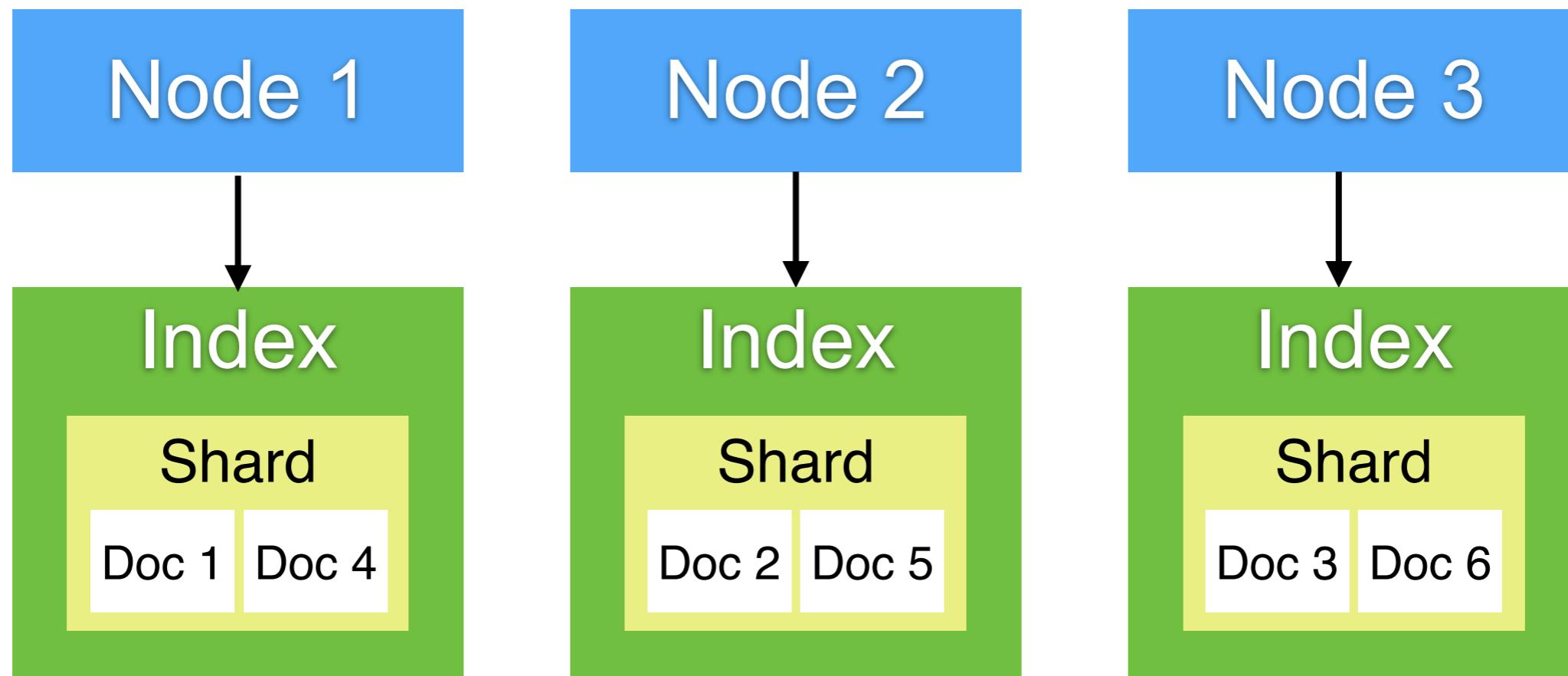
# Index is split into shards

Each shard may be on a different node in cluster

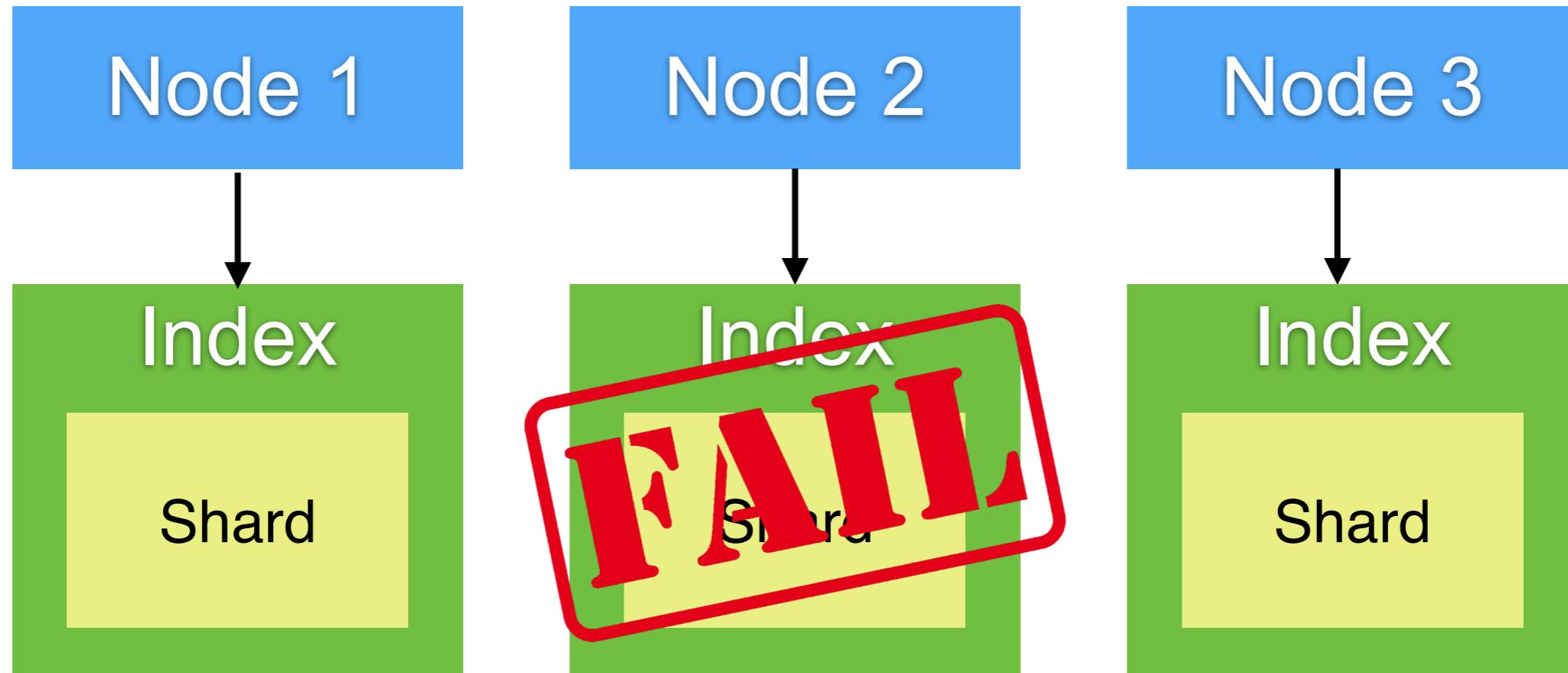
Every shard is a self-contained *Lucene index*



# Documents are hashed



# Design for fail (Replica)



# Replica = copy of shards

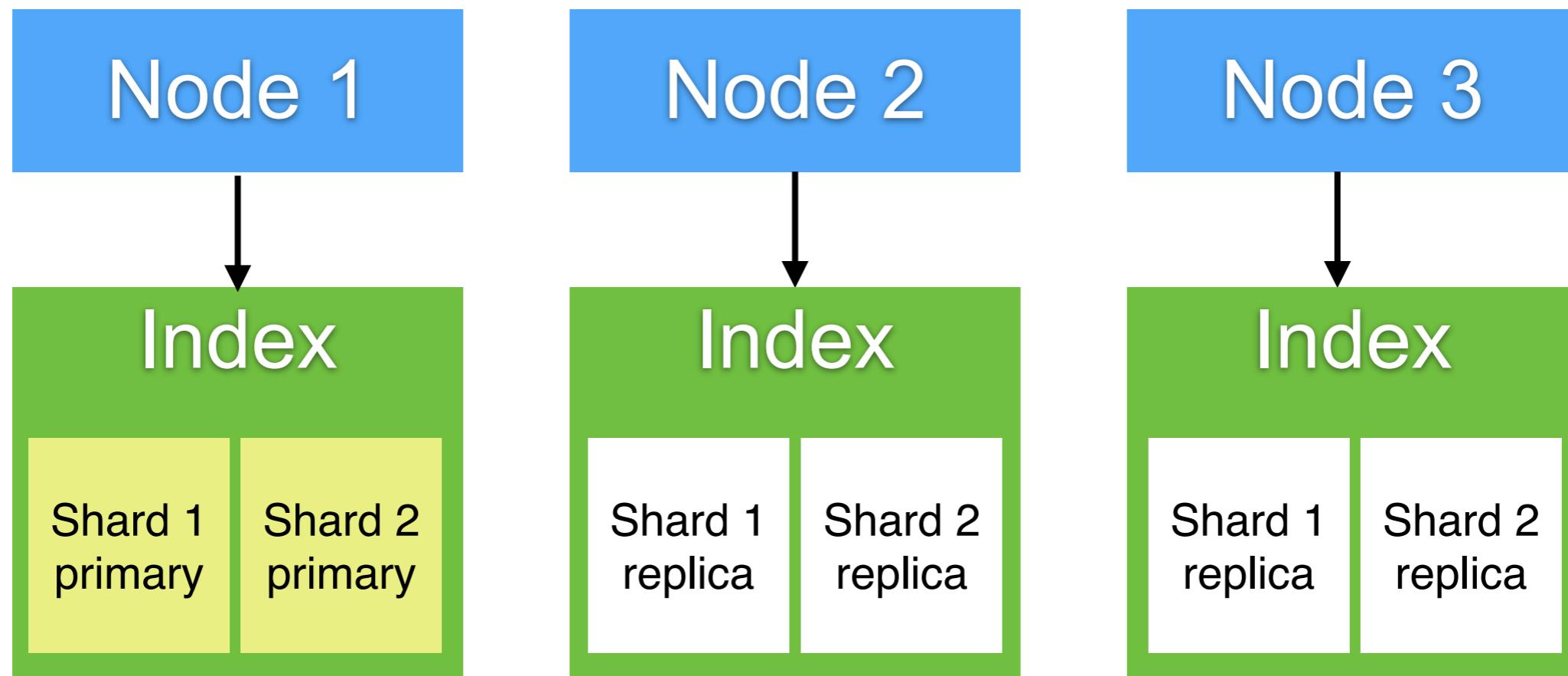
1 replica = 1 Primary + 1 Replica

2 replicas = 1 Primary + 2 Replicas

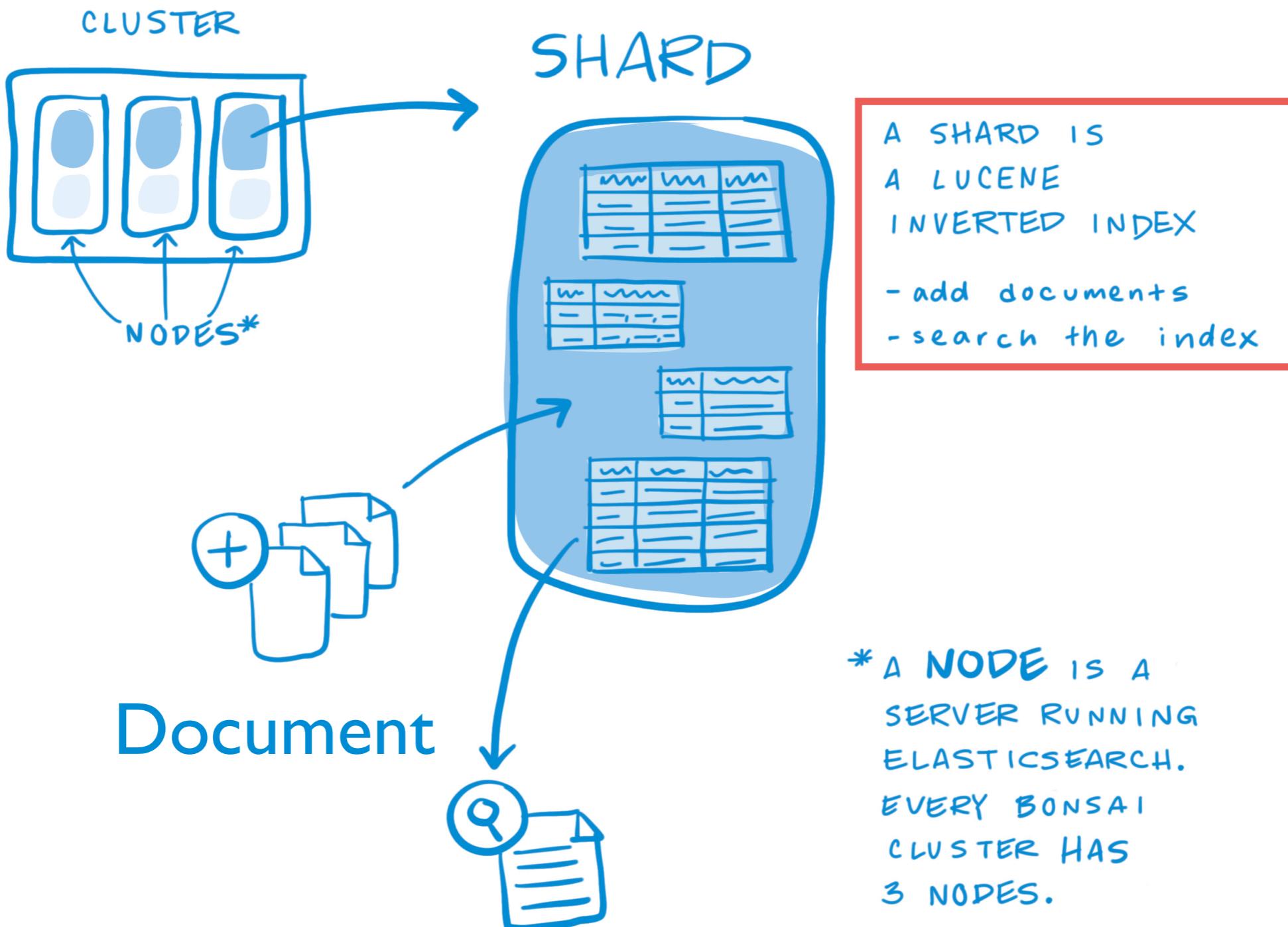
3 replicas = 1 Primary + 3 Replicas



# Replica = 2



# Basic concepts

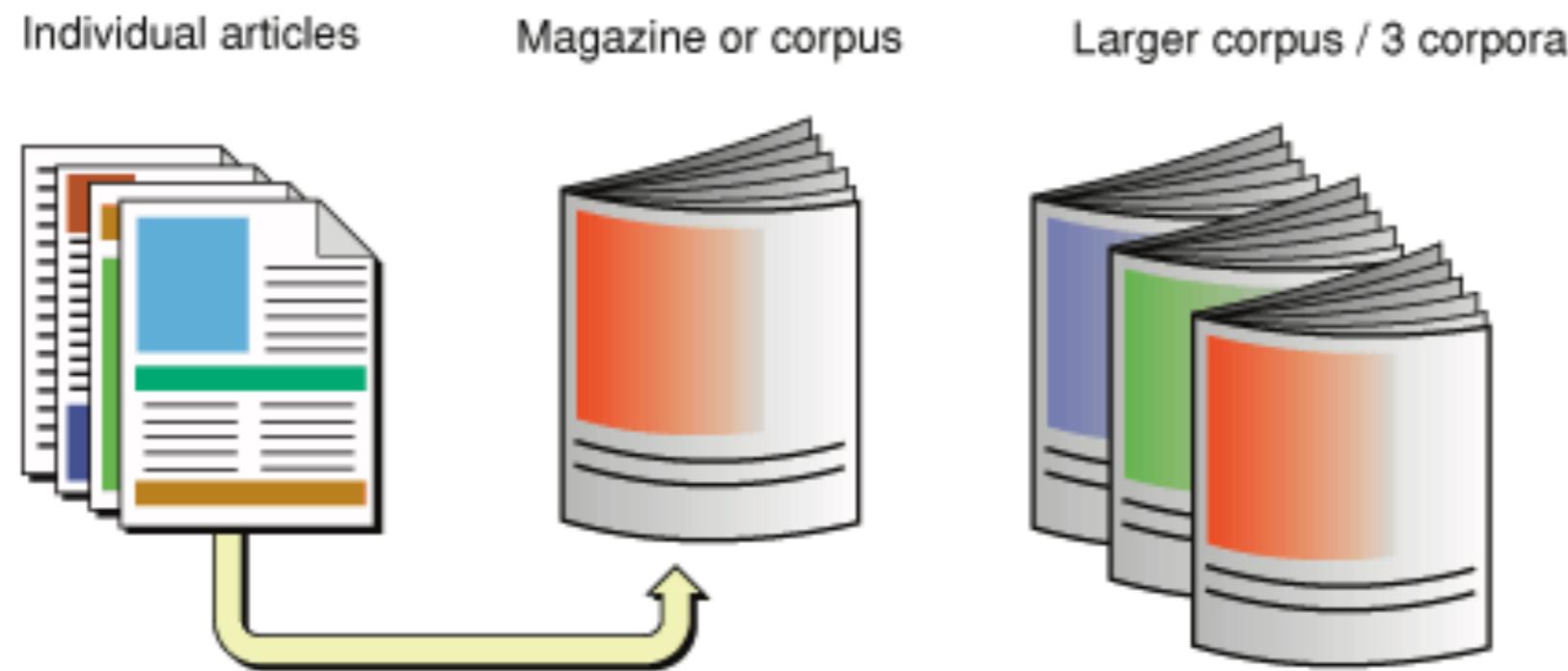


# Inverted Index



# Inverted Index

Corpus is a collection of documents



[https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SearchKitConcepts/searchKit\\_basics/searchKit\\_basics.html#/apple\\_ref/doc/uid/TP40002843-TPXREF101](https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SearchKitConcepts/searchKit_basics/searchKit_basics.html#/apple_ref/doc/uid/TP40002843-TPXREF101)

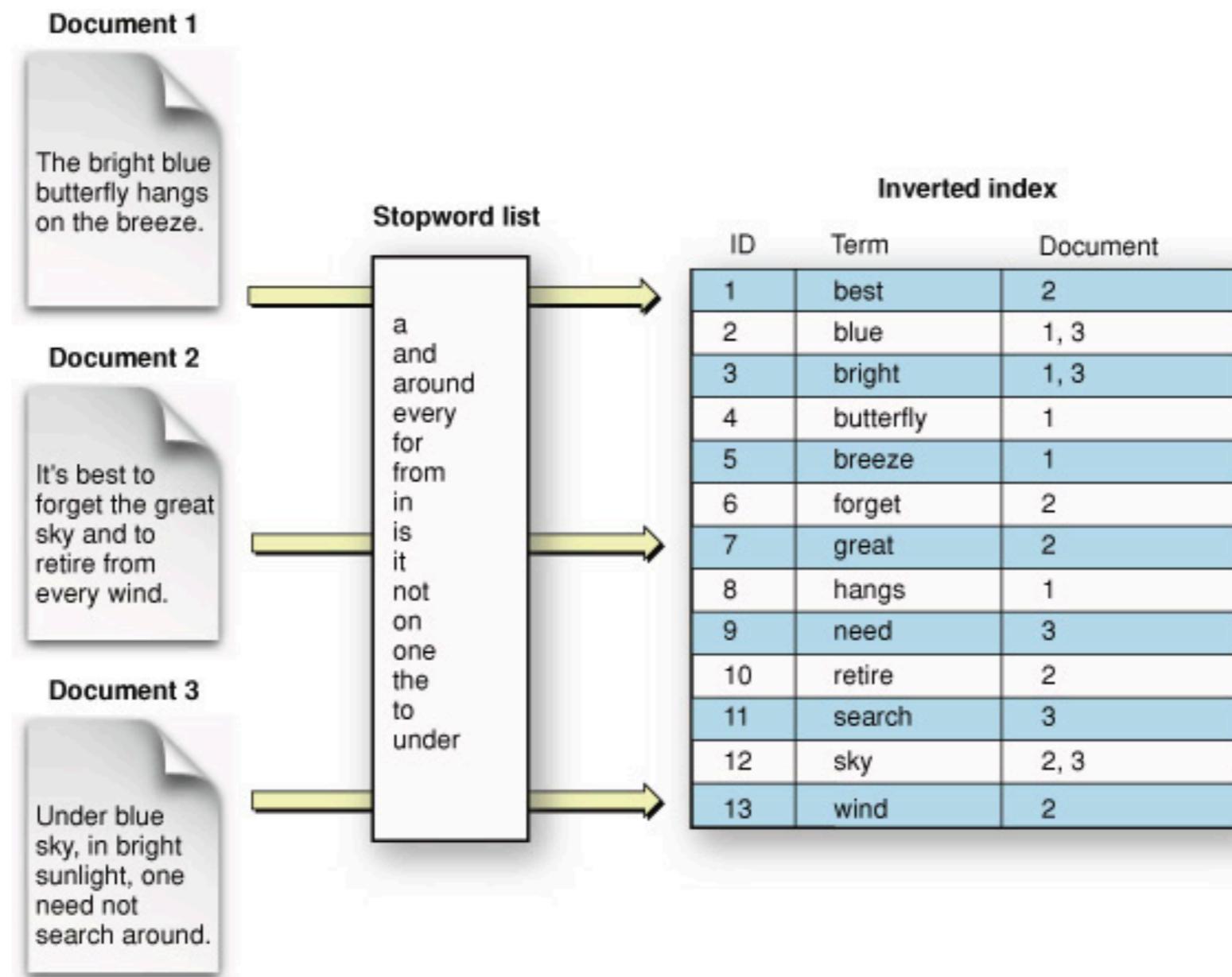


ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Inverted Index

Try to construct index



# Apache Lucene

Elasticsearch

Apache Lucene

JVM

<https://lucene.apache.org/>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Basic concept

Apache Lucene writes all the information to a structure called “**Inverted Index**”



# Basic concept

Document  
Field  
Term  
Token



# Example

Document no.	Data
1	Elasticsearch Server
2	Mastering Elasticsearch Second Edition
3	Apache Solr Cookbook Third Edition



# Token

Token	Document no.
Elasticsearch	1
Elasticsearch	2
Server	1
Mastering	2
Second	2
Edition	2
Edition	3
Apache	3
Solr	3
Cookbook	3
Third	3



# Term

Token	Count	Document no.
elasticsearch	2	1,2
server	1	1
mastering	1	2
second	1	2
edition	2	2,3
apache	1	3
solr	1	3
cookbook	1	3
third	1	3



# Lucene inverted index

Token	Count	Document no.
elasticsearch	2	1,2
server	1	1
mastering	1	2
second	1	2
edition	2	2,3
apache	1	3
solr	1	3
cookbook	1	3
third	1	3



# Lucene inverted index

Write-once and read-many-times structure

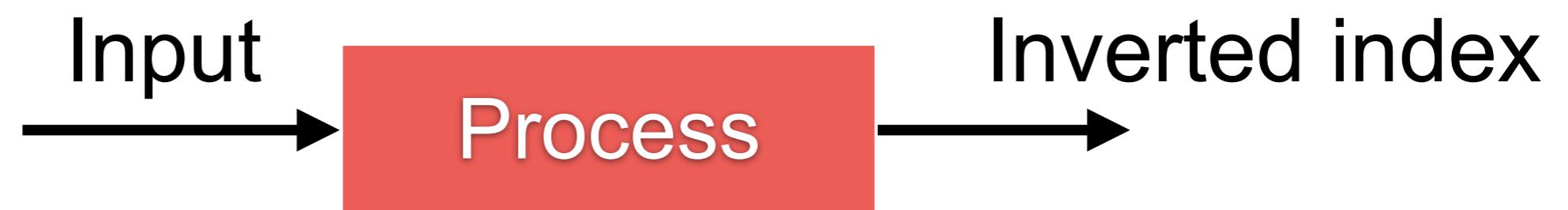
Called “**Segment**”

Can't be delete (just marked to deleted)



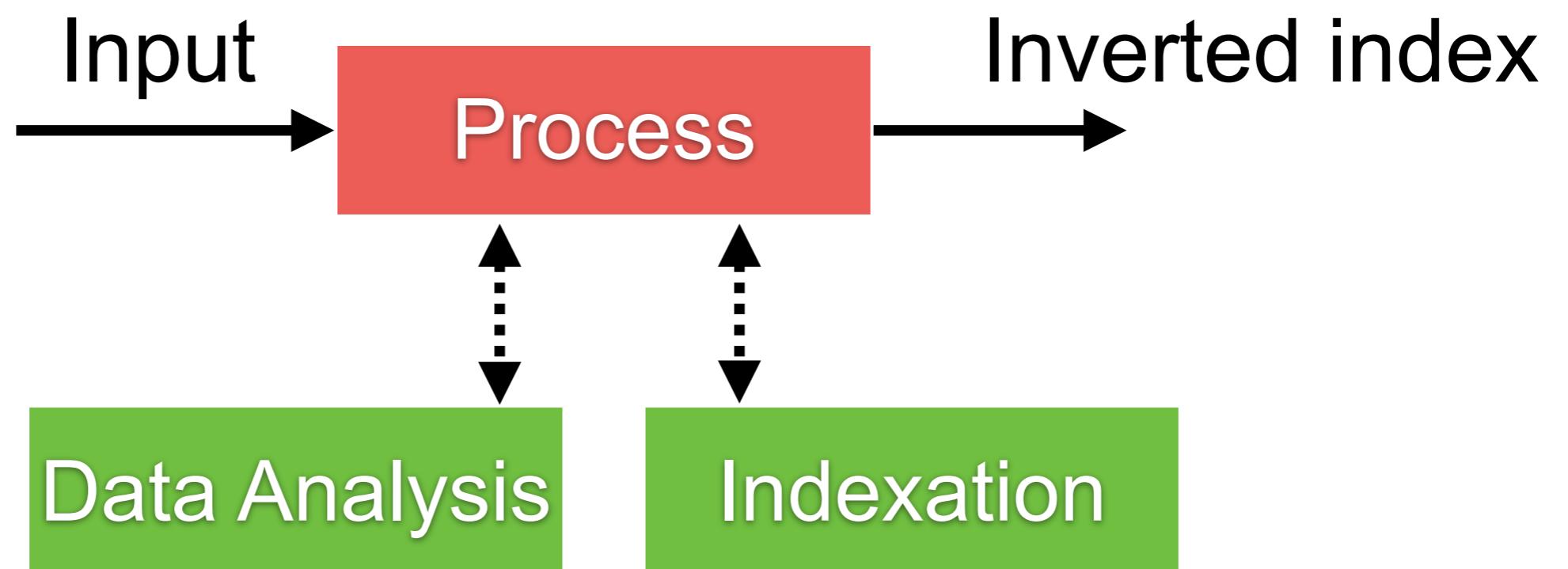
# Input data analysis

Write-once and read-many-times structure

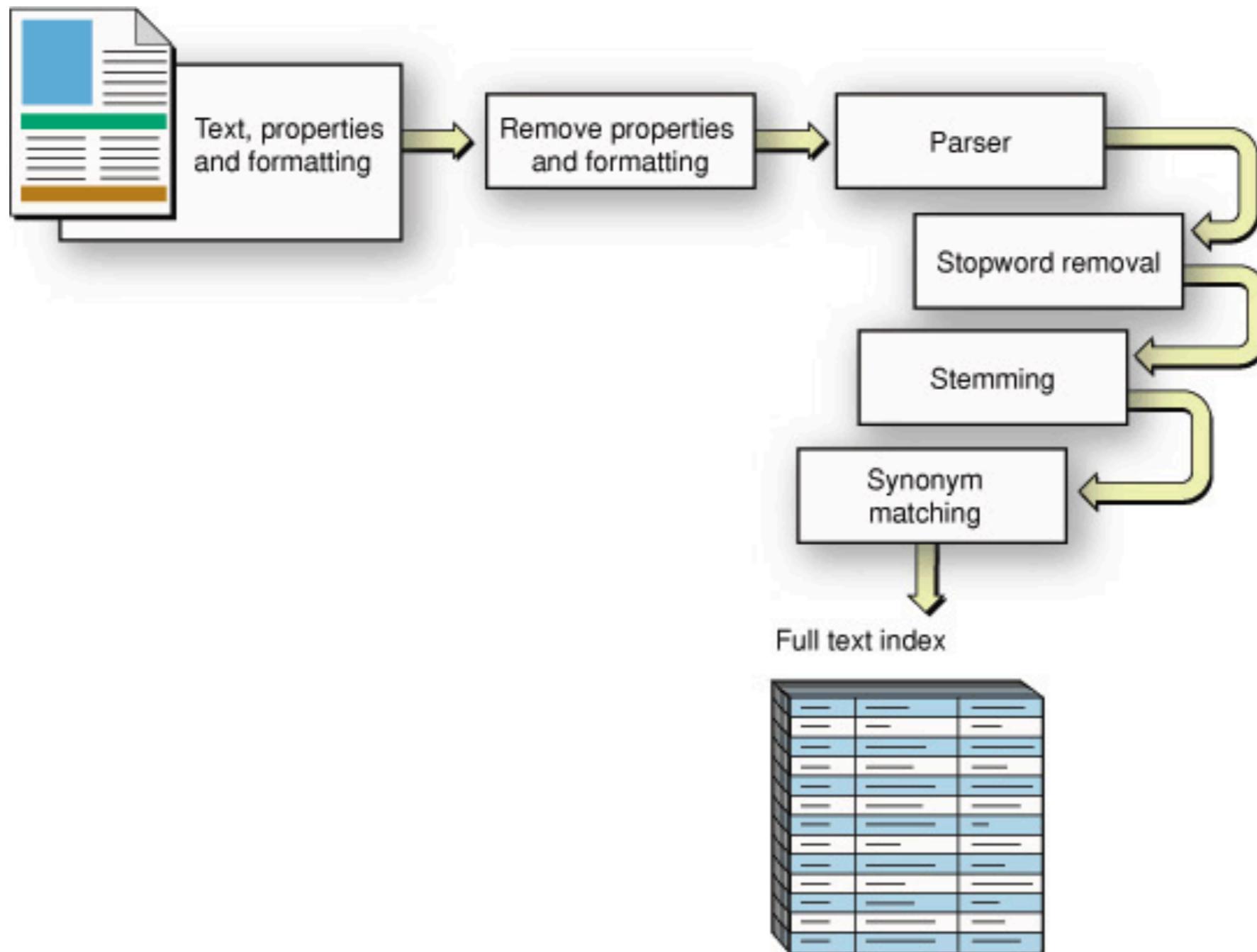


# Input data analysis

Write-once and read-many-times structure



# Process ?



# **CRUD with Elasticsearch**

**02-crud/book\_document.json**



# CRUD with Elasticsearch

Create document

Read document

Update document

Delete document



# Compare with RDBMS

Database

Table

Row

Column

Index

Type\*

Document

Field

\* Only 1 type per index



# Create a document

PUT /store/book/1

```
{  
  "title": "Elasticsearch: The Definitive Guide",  
  "author_name": [  
    "Clinton Gormley",  
    "Zachary Tong"  
,  
  "tag": [  
    "search",  
    "computer"  
,  
  "isbn-13": "978-1449358549",  
  "isbn-10": "1449358543",  
  "price": 44.3,  
  "page": 724,  
}
```



# Create document

PUT **/store/book/1**

Index name

Type name

Document ID



# 1 Type per Index

## Removal of mapping types



IMPORTANT

Indices created in Elasticsearch 6.0.0 or later may only contain a single [mapping type](#). Indices created in 5.x with multiple mapping types will continue to function as before in Elasticsearch 6.x. Mapping types will be completely removed in Elasticsearch 7.0.0.

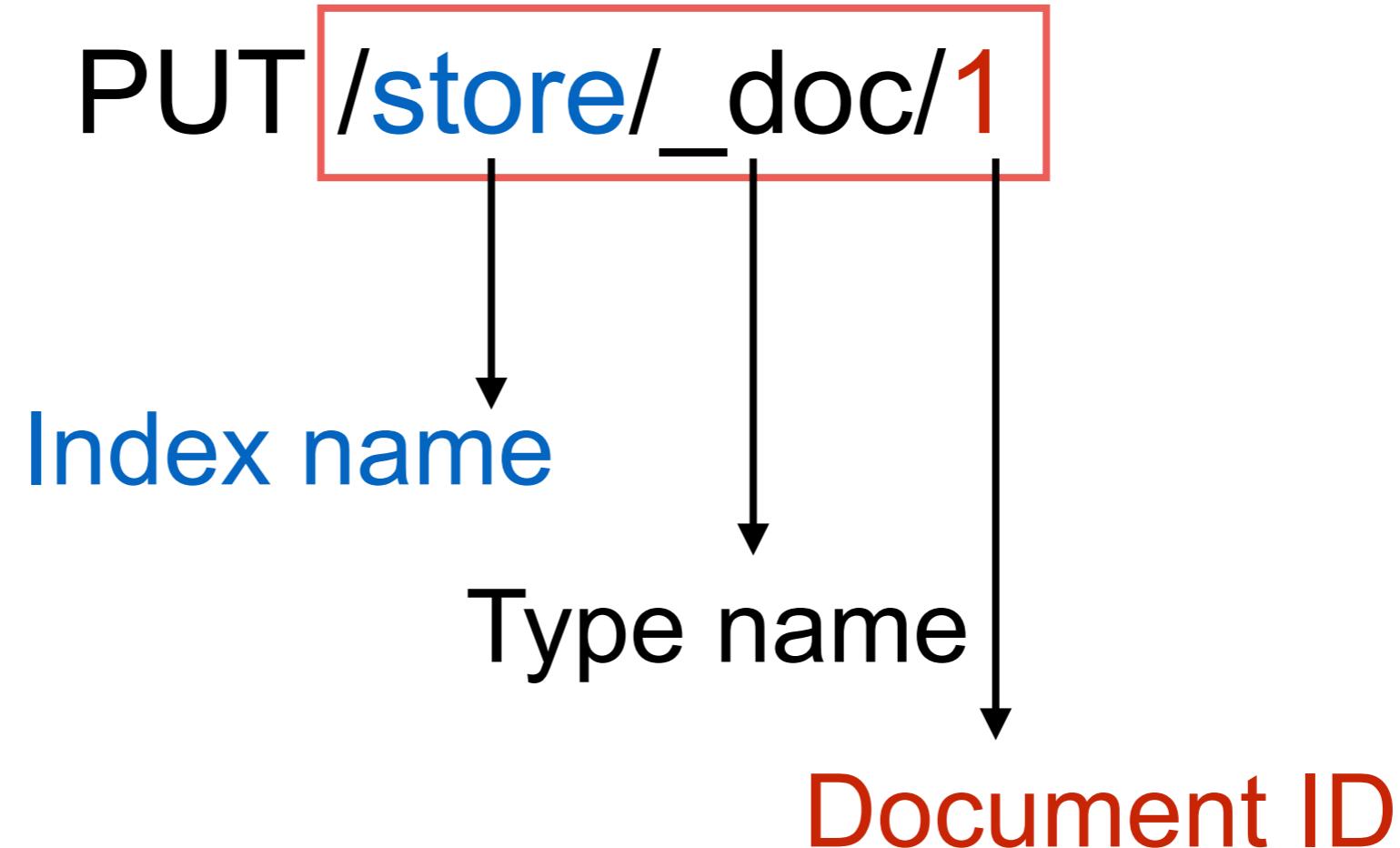
<https://www.elastic.co/guide/en/elasticsearch/reference/6.5/removal-of-types.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Create document (1 type)



# Change in Elasticsearch 7.x

# of shard of index change from 5 to 1

#! Deprecation: the default number of shards will change from [5] to [1] in 7.0.0; if you wish to continue using the default of [5] shards, you must manage this on the create index request or with an index template

```
{  
  "_index": "store1",  
  "_type": "book",  
  "_id": "2",  
  "_version": 1,  
  "result": "created",  
  "_shards": {  
    "index": "store1",  
    "status": "CREATED",  
    "shards": 1,  
    "primary": true  
  }  
}
```



# Read document

GET /store/book/1

```
{  
  "_index": "store",  
  "_type": "book",  
  "_id": "1",  
  "_version": 1,  
  "found": true,  
  "_source": {  
    "title": "Elasticsearch: The Definitive Guide",  
    "author_name": [  
      "Clinton Gormley",  
      "Zachary Tong"  
    ],  
    "tag": [  
      "search",  
      "computer"  
    ]  
  }  
}
```

Information of document



# Update document

Whole document  
Partial document



# Update whole document

PUT /store/\_doc/123

```
{  
  "title": "Update",  
  "author_name": [  
    "user1",  
    "user2"  
  ],  
  "tag": [  
    "update",  
    "book"  
  ]  
}
```



# Update partial document

**POST /store/\_update/123**

```
{  
  "doc": {  
    "title": "partial update",  
    "tag": [  
      "test",  
      "computer"  
    ],  
    "views": 0  
  }  
}
```



# Delete document

DELETE /store/\_doc/1

```
{  
  "_index": "store",  
  "_type": "book",  
  "_id": "1",  
  "_version": 2,  
  "result": "deleted",  
  "_shards": {  
    "total": 2,  
    "successful": 1,  
    "failed": 0  
  },  
  "_seq_no": 1,  
  "_primary_term": 1  
}
```

*Not delete document !!  
Marked deleted only*



# Delete index (delete from disk)

DELETE /store



# More features

Update by query

Delete by query

Partial update document



# Routing

<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping-routing-field.html#>

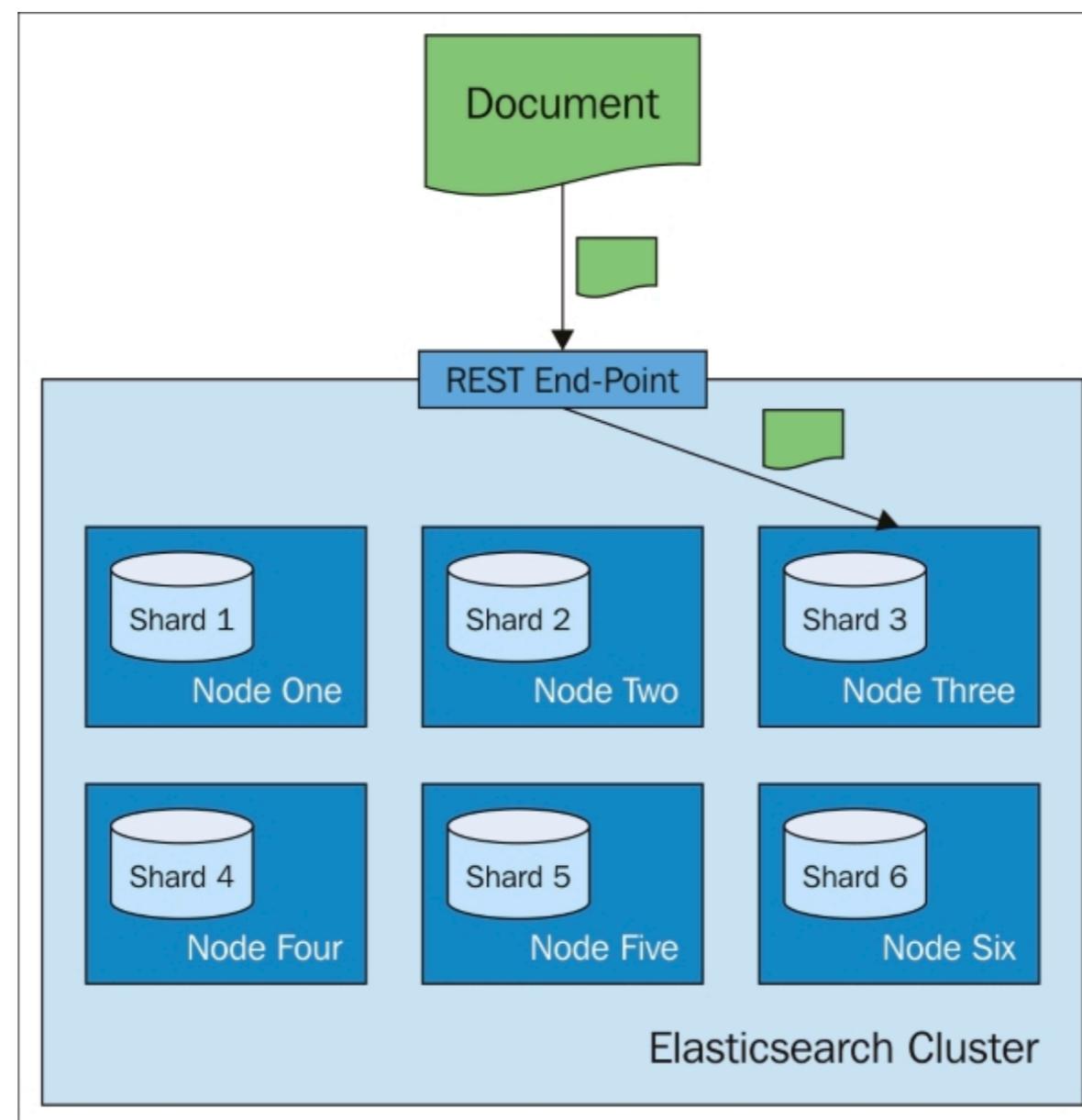


ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

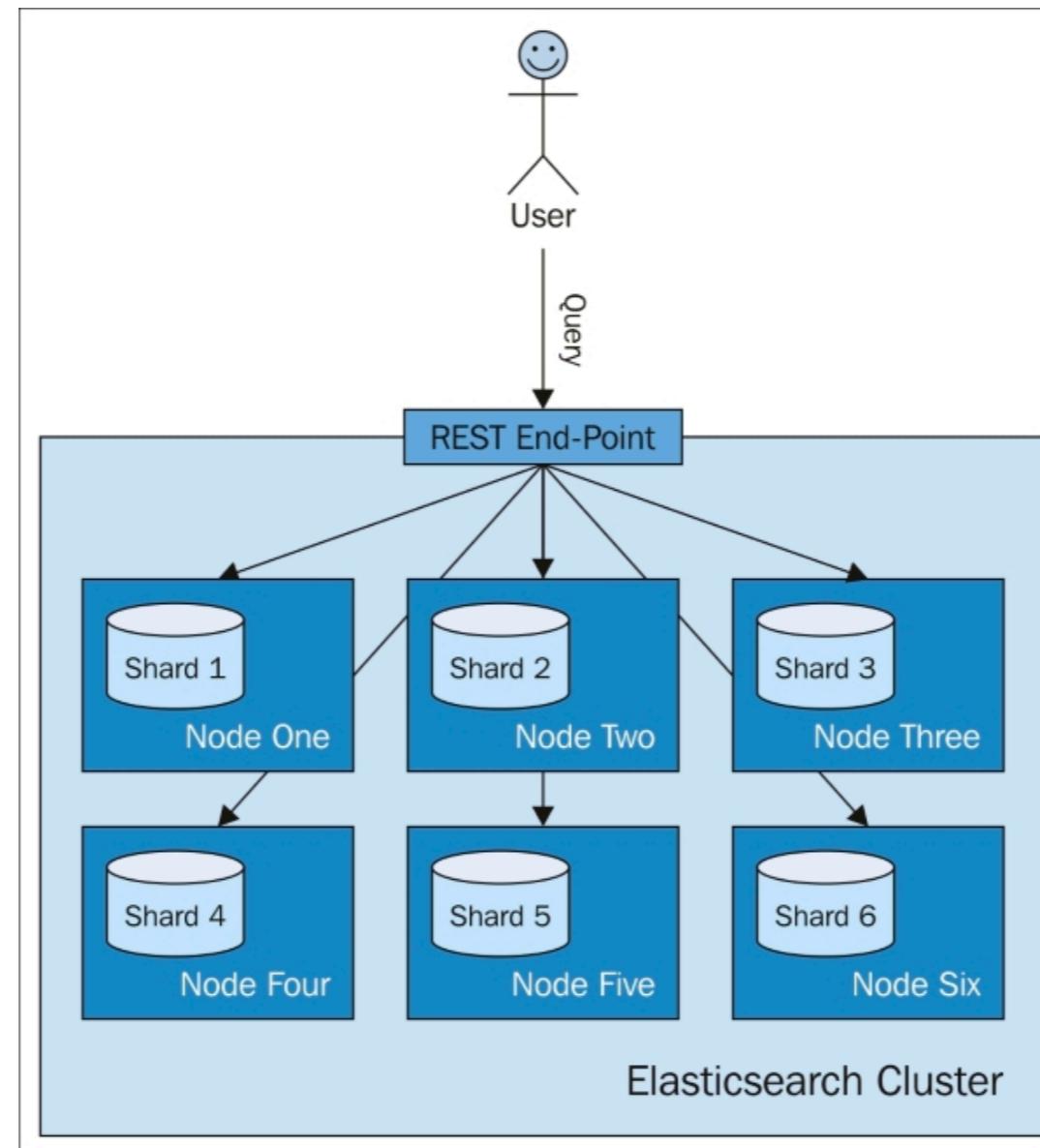
# Default indexing

ES calculate the hash value of the doc id



# Default searching/query

Query all the shards to get data  
(depend on search type)



# Custom routing

Routing field  
Routing to index partition



# Routing field

```
shard_num =  
hash(_routing) % num_primary_shards
```



# Routing to index partition

```
shard_num  
=  
(hash(_routing) + hash(_id)) % routing_partition_size)  
%
```

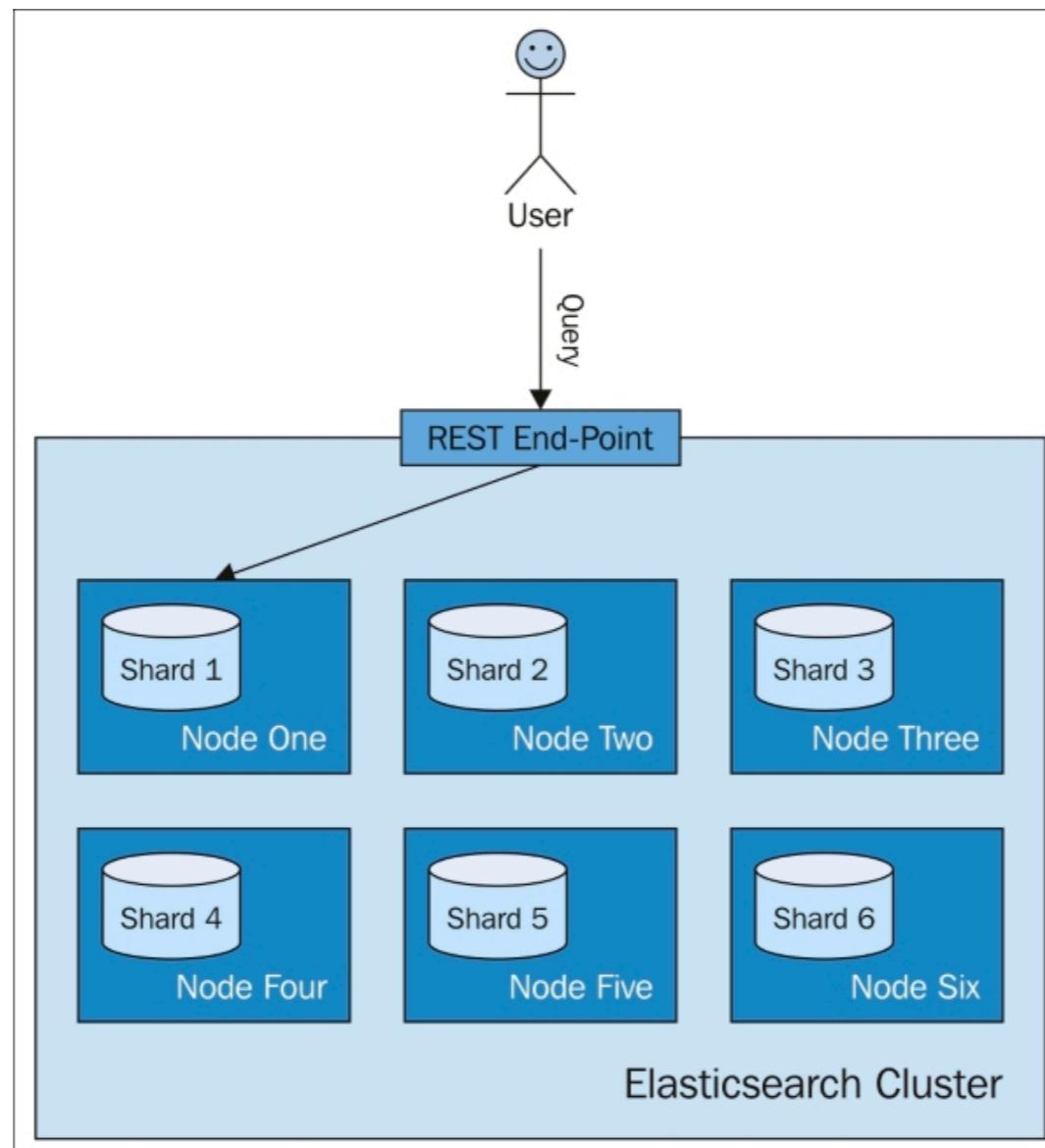
num\_primary\_shards

*routing\_partition\_size = 1 (default)*



# Custom routing

ES will send our query to a single shard



# Bulk API

03-bulk/book\_bulk.json

<https://www.elastic.co/guide/en/elasticsearch/reference/current/docs-bulk.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

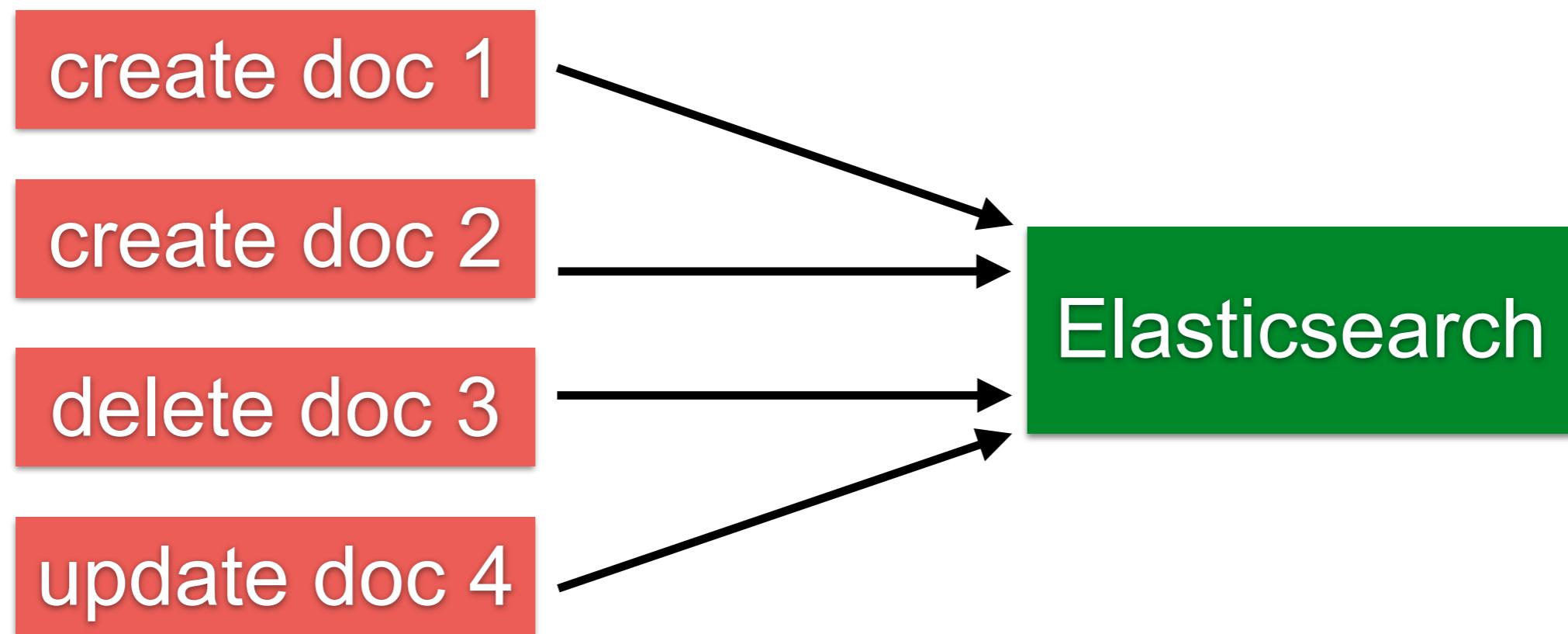
# Bulk API

Perform many index/delete operation in single API call

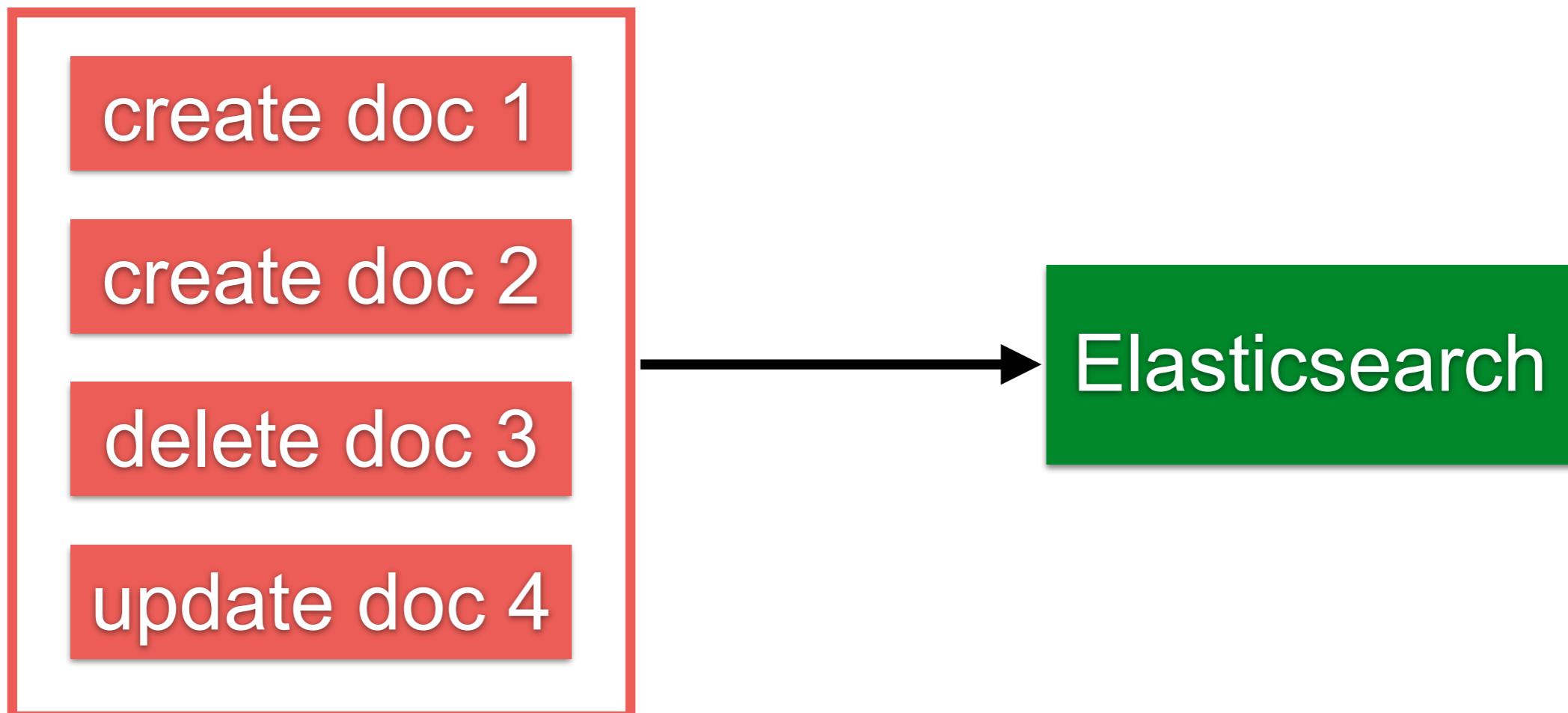
Increase indexing speed



# Without Bulk API



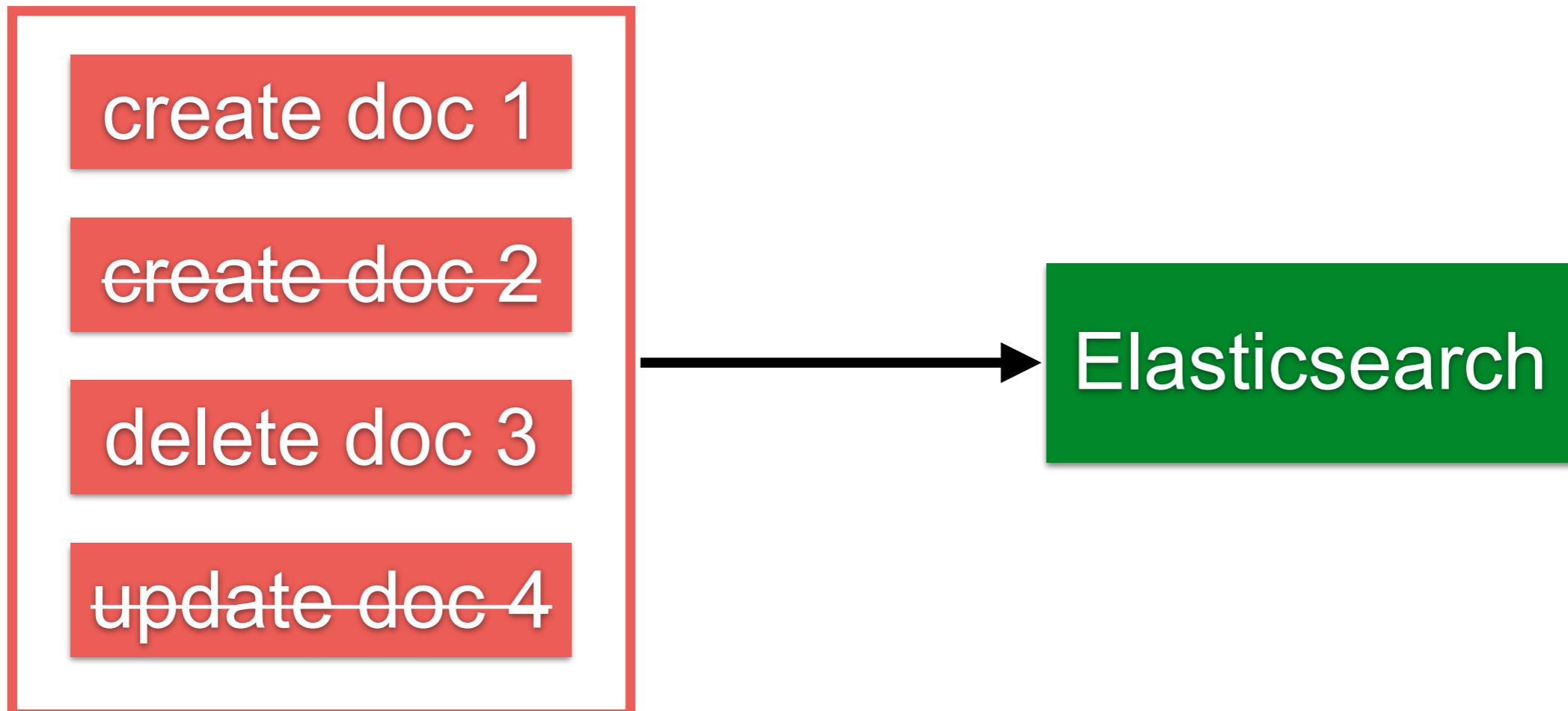
# With Bulk API



Store in memory 5-15 MB



# No transaction in bulk api



# Create a document

**POST /store/book/\_bulk**

```
{"create":{"_id":"1001"}  
{"title": "new book 1000", "description": "my new book"}}
```



# Response from Bulk API

```
{  
  "took": 89,  
  "errors": false,  
  "items": [  
    {  
      "create": {  
        "_index": "store",  
        "_type": "book",  
        "_id": "1001",  
        "_version": 1,  
        "result": "created",  
        "_shards": {  
          "total": 2,  
          "successful": 1,  
          "failed": 0  
        },  
        "_seq_no": 0,  
        "_primary_term": 1,  
        "status": 201  
      }  
    }  
  ]  
}
```

Time in milliseconds

HTTP Status 201 = Created



# Searching / Query

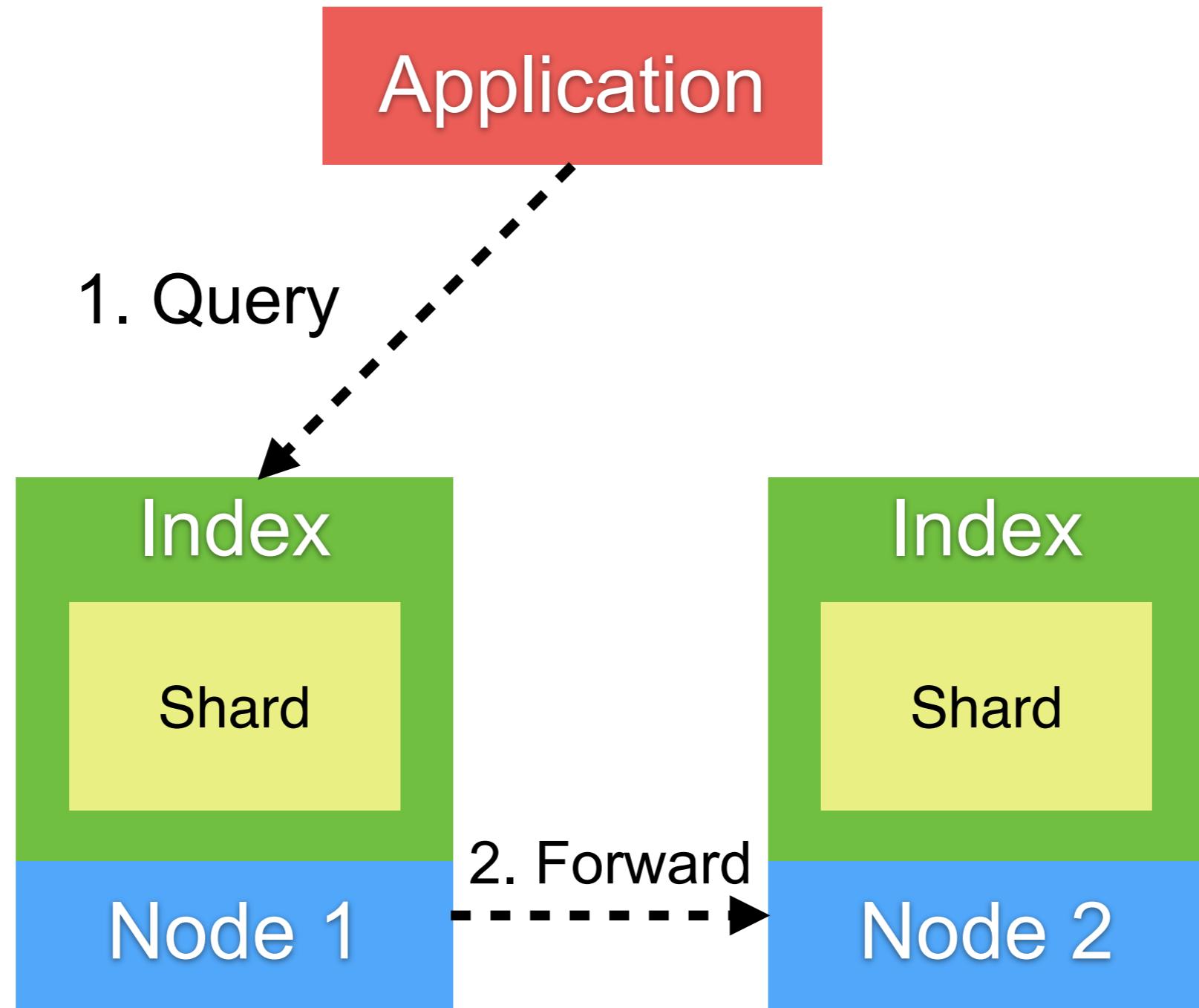


# Searching

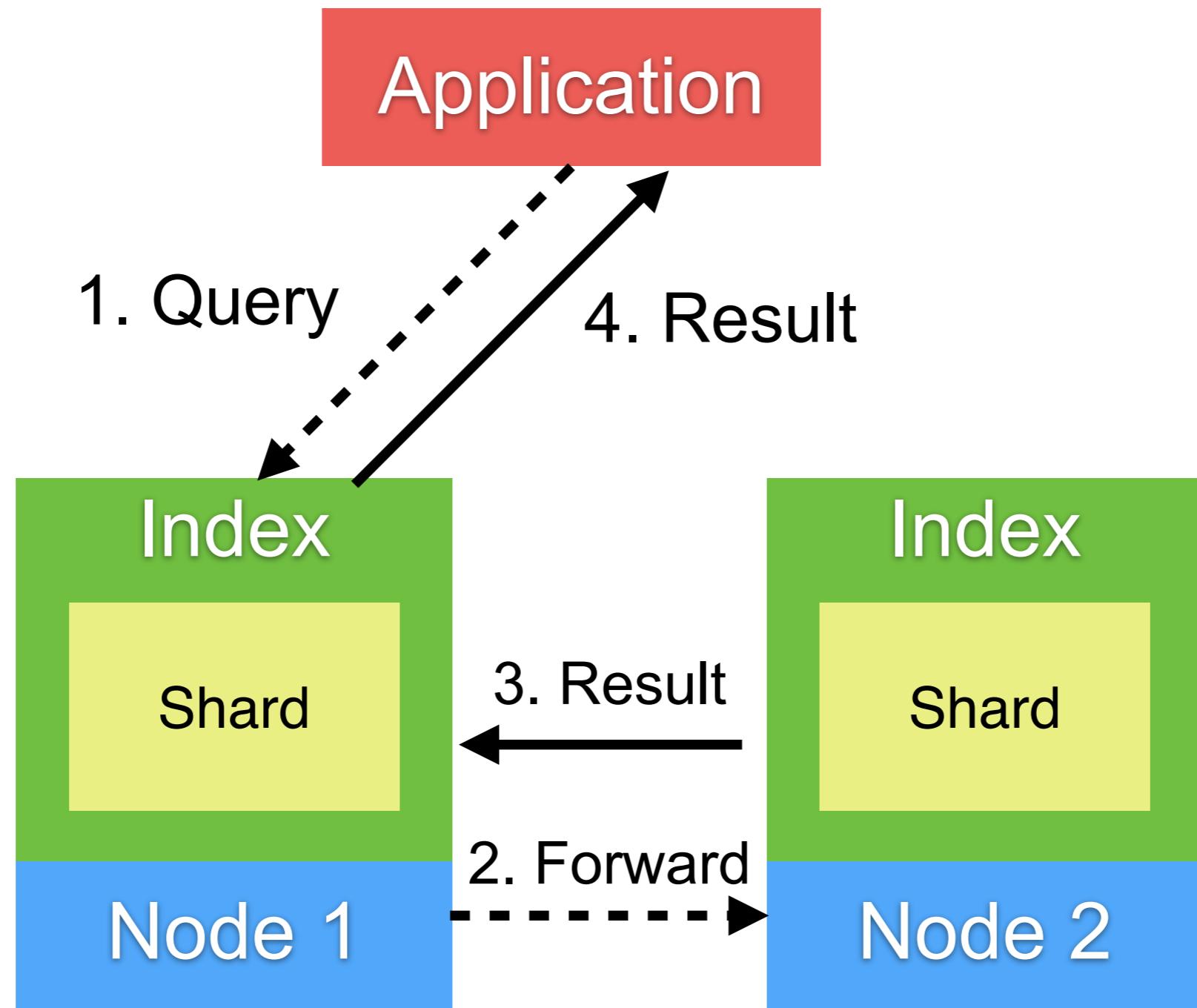
Scattering phase  
Gathering phase



# Scattering phase



# Gathering phase



# Query DSL

04-search/book\_search.json



# Query DSL

**Domain Specific Language for query data**  
**Flexible query language**  
**Based on JSON format**

<https://www.elastic.co/guide/en/elasticsearch/reference/current/query-dsl.html>

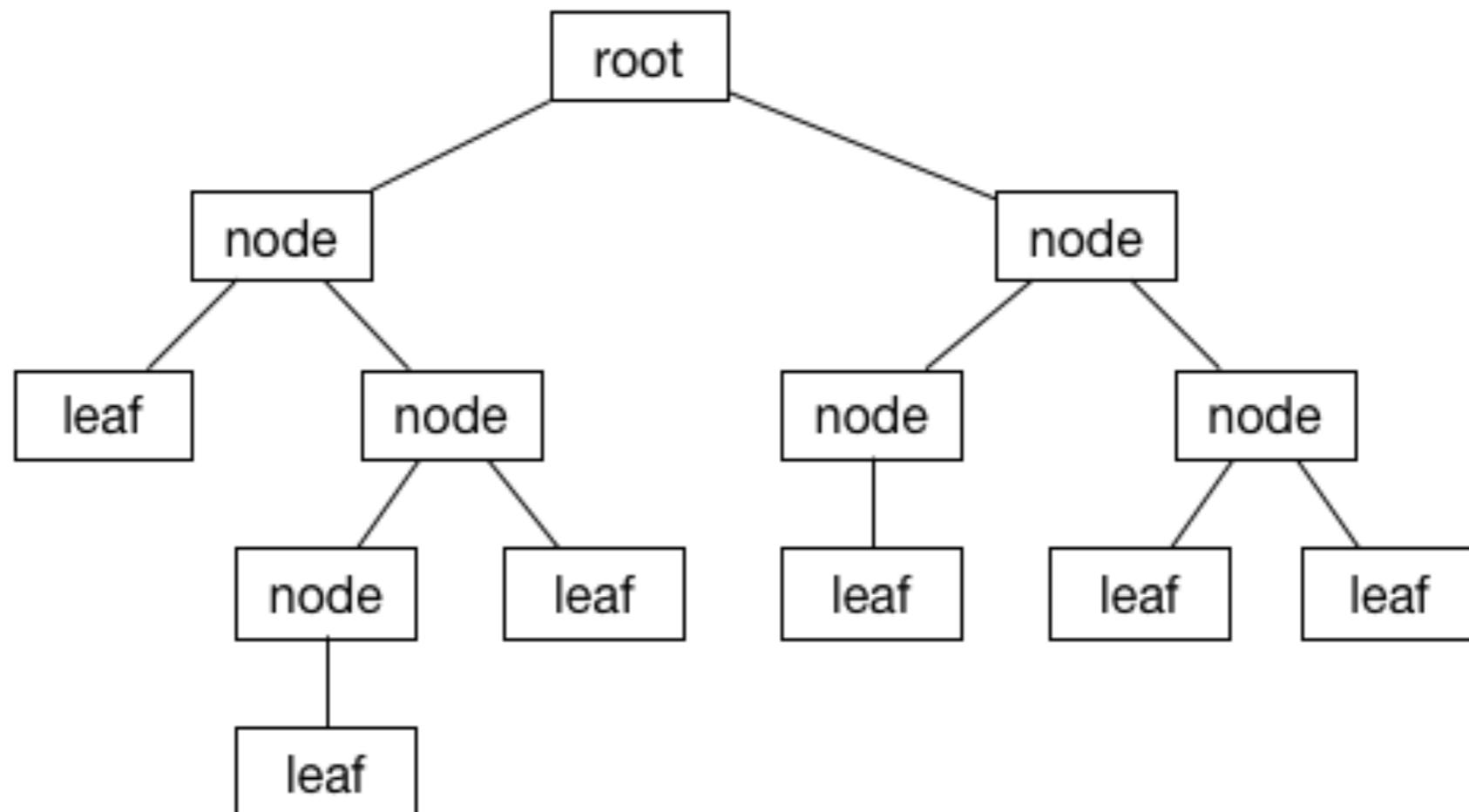


ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

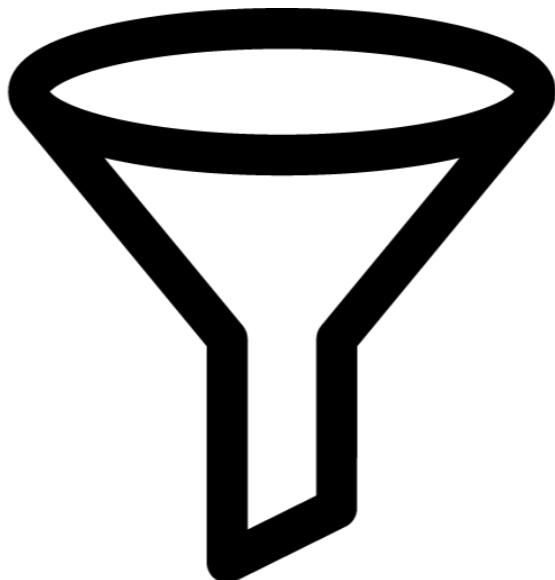
# Query DSL

1. Leaf query clause
2. Compound query clause



# Query DSL

Query (unstructured data)  
Filter (structured data)



Query



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

119

# Query DSL

Query	Filter
Relevance	Boolean, yes/no
Full text search	Exact values
Not cached	Cached
Slower	Faster

***Filter first, then query remaining documents***



# Query DSL

Full text query  
Term level query  
Compound query  
Joining query  
Geo query  
Specialized query  
Span query



# Leaf query clause

GET /store/book/\_search

```
{  
  "query": {  
    "match_all": {}  
  }  
}
```



# Compound query clause

GET /store/book/\_search

```
{  
  "query": {  
    "bool": {  
      "must": [{}],  
      "should": [{}],  
      "must_not": [{}]  
    }  
  }  
}
```



# \_all field is disabled (ES 6.0+)

## \_all field



**WARNING**

Deprecated in 6.0.0.

`_all` may no longer be enabled for indices created in 6.0+, use a custom field and the mapping `copy_to` parameter

Space issue  
Analyzer issue  
Highlighting issue

<https://www.elastic.co/guide/en/elasticsearch/reference/6.5/mapping-all-field.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Workshop

amazon

All ▾ elasticsearch

New to Amazon? Click here to learn more

Deliver to Thailand

Departments ▾ Your Amazon.com Today's Deals Gift Cards Sell

EN ▾ Hello. Sign in Account & Lists ▾ Orders Cart 0

1-16 of 119 results for "elasticsearch"

Show results for

**Books**

- Computers & Technology
- Data Processing
- Web Development & Design
- Online Internet Searching
- Databases & Big Data
- ▼ See more

**Kindle Store**

- Computers & Technology
- Business Software
- Search Engines
- Application Development
- Computer Databases
- ▼ See more
- ▼ See All 8 Departments

Refine by

**Book Language**

- English

**Book Format**

- Paperback

Packt

SPONSORED BY PACKT PUBLISHING

Complete Database Solutions with PostgreSQL

Shop now ▾

SQL Server 2017 Administrator's Guide

PostgreSQL 9.6 High Performance

SQL Server 2017 Administrators Guide

PostgreSQL 9.6 High Performance: Optimize your...

Advertisement

Sponsored ⓘ

Learning Elastic Stack 6.0: A beginner's guide to distributed search, analytics, and visualization using Elasticsearch, Logstash and Kibana

Dec 22, 2017

by Pranav Shukla and Sharath Kumar M N

Eligible for Shipping to Thailand

Get to grips with the new features introduced in Elastic Stack 6.0, and deliver end-to-end real-time distributed data processing solutions.

Paperback

\$34<sup>99</sup>

In Stock

★★★★★ 7

Previous Page 1 2 3 ... 8 Next Page



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

125

# Query

amazon

All ▾ elasticsearch 

Departments ▾ Your Amazon.com Today's Deals Gift Cards Sell

New to Amazon? EN Hello. Sign in Account Lists Orders Cart

1-16 of 119 results for "elasticsearch"

## Filter

**Books**

- Computers & Technology
- Data Processing
- Web Development & Design
- Online Internet Searching
- Databases & Big Data
- ▼ See more

**Kindle Store**

- Computers & Technology
- Business Software
- Search Engines
- Application Development
- Computer Databases
- ▼ See more
- ▼ See All 8 Departments

Refine by

**Book Language**

- English

**Book Format**

- Paperback

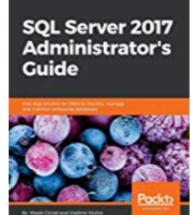
**SPONSORED BY PACKT PUBLISHING**

## Packt

Complete Database Solutions with PostgreSQL

Shop now ›

**SQL Server 2017 Administrator's Guide**  
By Waheed Ahmad and Imanuele Pollicino



SQL Server 2017 Administrators Guide  
★★★★★ 3  
prime

**PostgreSQL 9.6 High Performance**  
By Waheed Ahmad, Gregory Smith



PostgreSQL 9.6 High Performance: Optimize your...  
★★★★★ 1  
prime

Sponsored ⓘ

**Learning Elastic Stack 6.0: A beginner's guide to distributed search, analytics, and visualization using Elasticsearch, Logstash and Kibana** Dec 22, 2017

by Pranav Shukla and Sharath Kumar M N

Eligible for Shipping to Thailand  
Get to grips with the new features introduced in Elastic Stack 6.0, and deliver end-to-end real-time distributed data processing solutions.

**Paging**

Previous Page 1 2 3 ... 8 Next Page

Sort by Featured



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Aggregation API

05-aggregation/book\_aggregation.json

<https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

```
SELECT count(1), sum(price)  
FROM some_table  
GROUP BY some_column
```



# Aggregation Types

Bucketing  
Metric  
Matrix  
Pipeline



# Structure

```
"aggregations" : {  
    "<aggregation_name>" : {  
        "<aggregation_type>" : {  
            <aggregation_body>  
        }  
        [ , "meta" : { [ <meta_data_body> ] } ] ?  
        [ , "aggregations" : { [ <sub_aggregation> ]+ } ] ?  
    }  
    [ , "<aggregation_name_2>" : { ... } ] *  
}
```



# Count by category

GET /store/\_search

```
{  
  "aggs": {  
    "all_book_title": {  
      "terms": {  
        "field": "category.keyword"  
      }  
    }  
  }  
}
```



# Count by category

GET /store/\_search

```
{  
  "aggs": {  
    "all_book_title": {  
      "terms": {  
        "field": "category.keyword"  
      }  
    }  
  }  
}
```



# Count by category

```
GET /store/_search
{
  "aggs": {
    "all_book_title": {
      "terms": { Aggregation type
        "field": "category.keyword"
      }
    }
  }
}
```



# Result of aggregation

```
{  
  "hits": {  
    "total": 5,  
    "max_score": 1,  
    "hits": [  
      {  
        "_source": {  
          "title": "The Logstash Book"  
        }  
      },  
      {  
        "_source": {  
          "title": "Elasticsearch Server: Second Edition"  
        }  
      }  
    ]  
  }  
}
```

Search result



# Result of aggregation

```
"aggregations": {  
    "all_book_title": {  
        "doc_count_error_upper_bound": 0,  
        "sum_other_doc_count": 0,  
        "buckets": [          Aggregation result  
            {  
                "key": "Computer & Technology",  
                "doc_count": 5  
            },  
            {  
                "key": "Online Searching",  
                "doc_count": 3  
            },  
            {  
                "key": "Java Programming",  
                "doc_count": 2  
            }  
        ]  
    }  
}
```



# Show only aggregation result

GET /store/\_search

```
{  
  "size": 0, Set search result size = 0  
  "aggs": {  
    "all_book_title": {  
      "terms": {  
        "field": "category.keyword"  
      }  
    }  
  }  
}
```



# Range of price

GET /store/\_search

```
{  
  "size": 0,  
  "aggs": {  
    "price_range": {  
      "range": {  
        "field": "price",  
        "ranges": [  
          { "from": 0,"to": 10 },  
          { "from": 11,"to": 20 },  
          { "from": 21,"to": 50 }  
        ]  
      }  
    }  
  }  
}
```



# Result of aggregation

```
"buckets": [
  {
    "key": "0.0-10.0",
    "from": 0,
    "to": 10,
    "doc_count": 1
  },
  {
    "key": "11.0-20.0",
    "from": 11,
    "to": 20,
    "doc_count": 0
  },
  {
    "key": "21.0-50.0",
    "from": 21,
    "to": 50,
    "doc_count": 3
  }
]
```



# Range of price and ordering

GET /store/\_search

```
{  
  "size": 0,  
  "aggs": {  
    "price_range": {  
      "range": {  
        "field": "price",  
        "ranges": [  
          { "from": 0,"to": 10 },  
          { "from": 11,"to": 20 },  
          { "from": 21,"to": 50 }  
        ]  
      }  
    }  
  }  
}
```



# Workshop aggregation with car

05-aggregation/car.json



# Try by yourself

Best seller by color

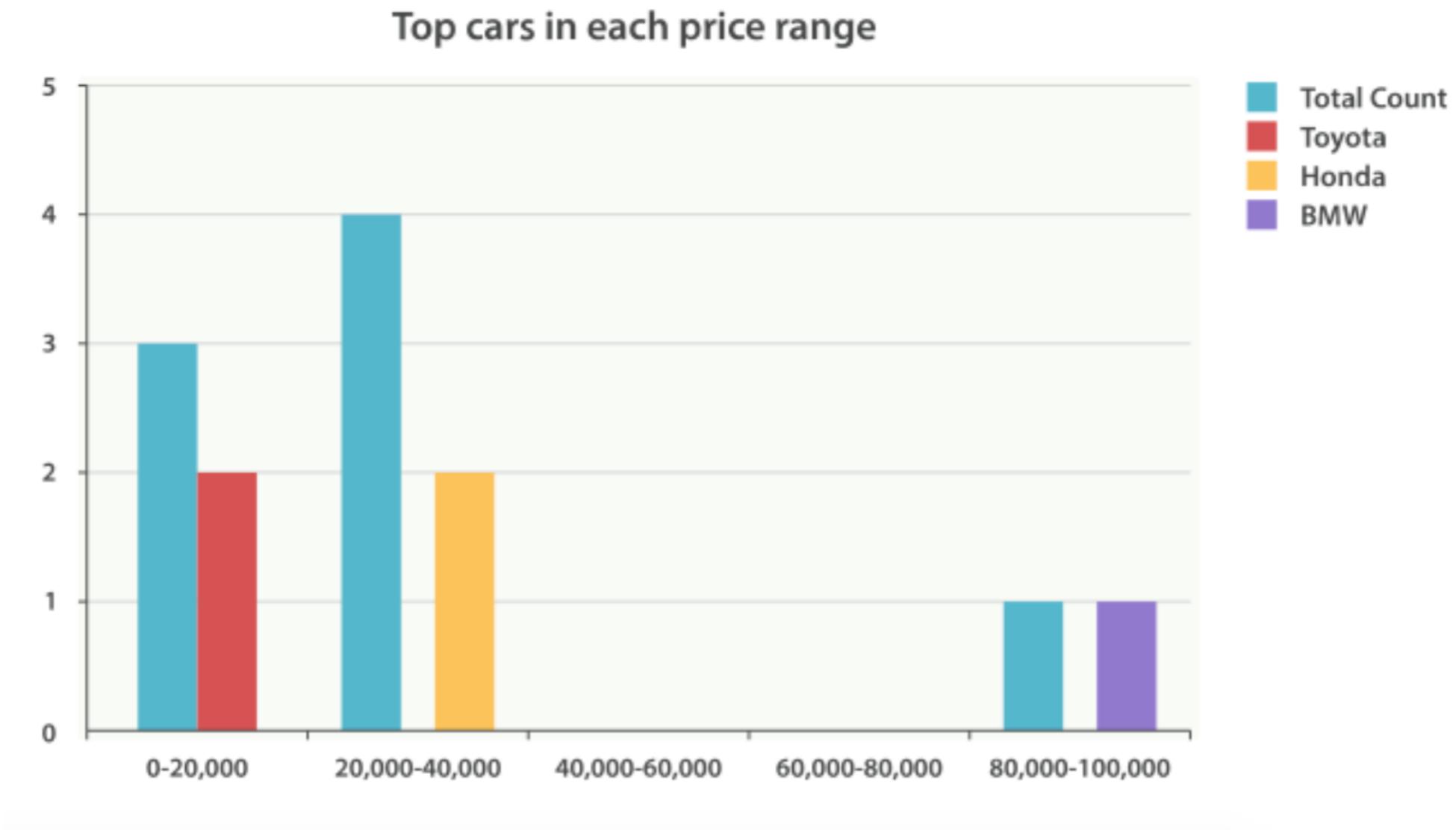
Statistic of best seller by color

Detail of car in each color

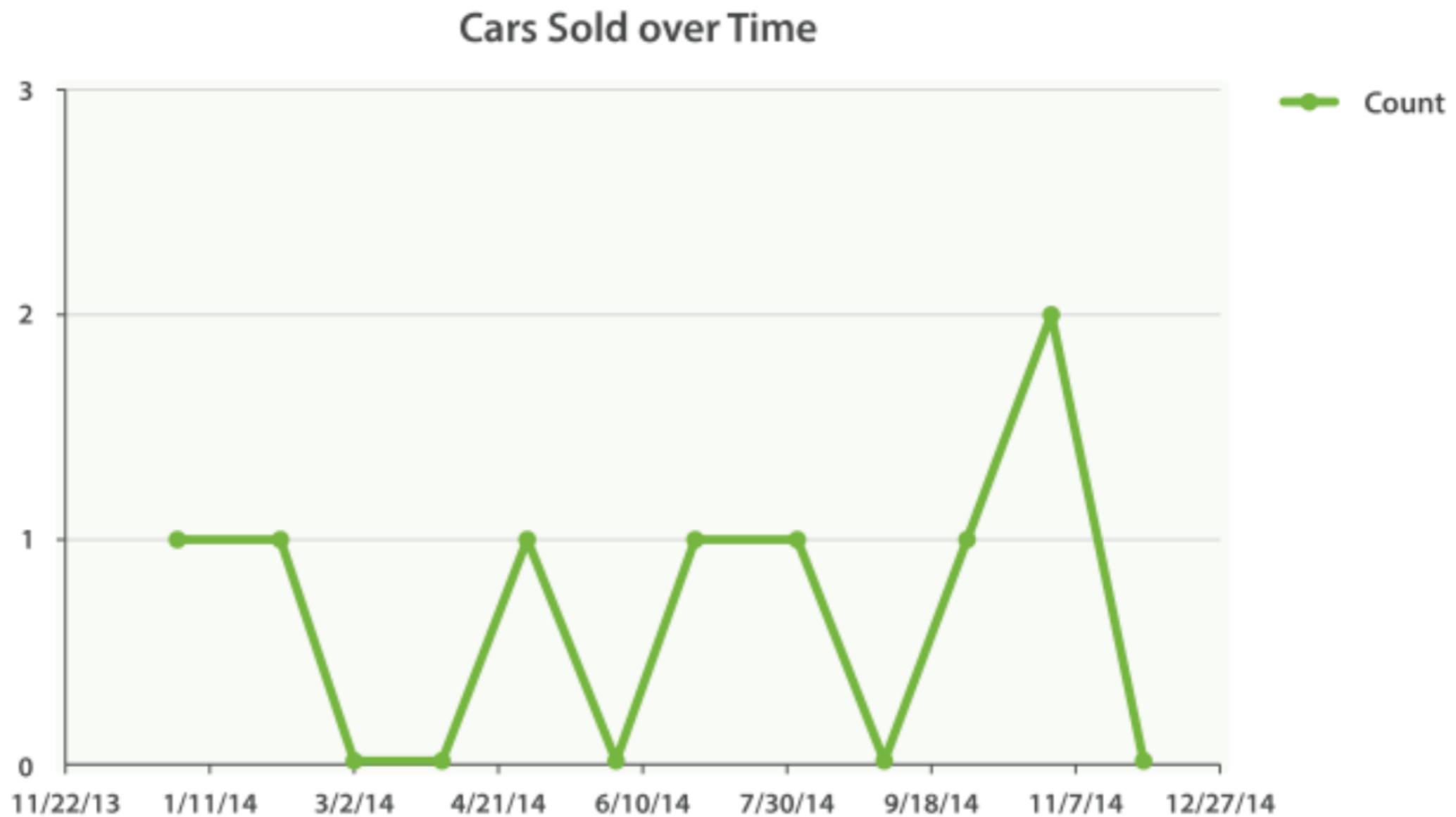
Min/max of price by make



# Top cars in each price range ?



# Cars sold over Time ?



# Mapping

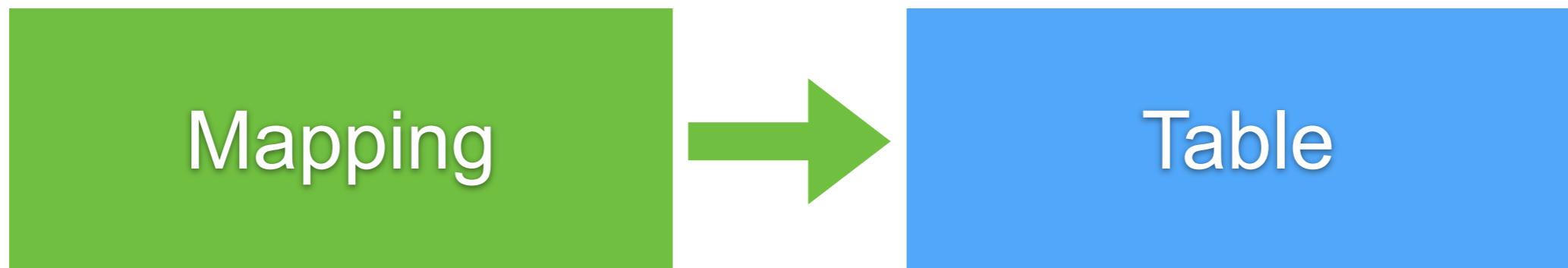
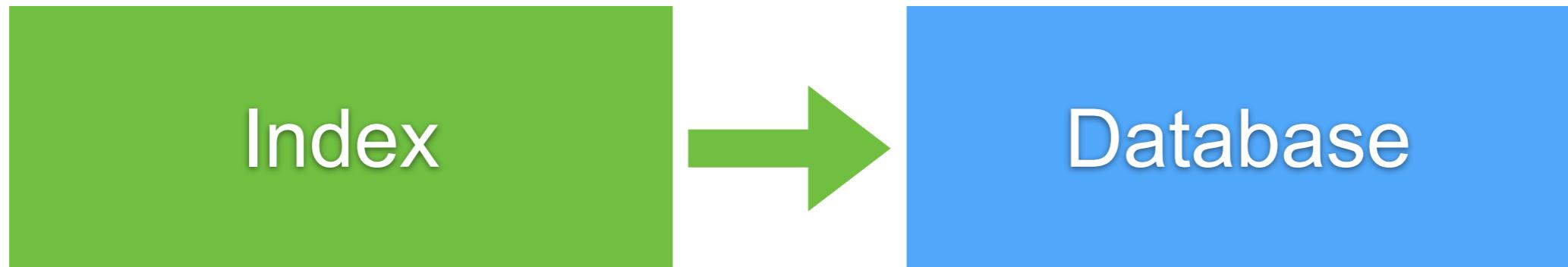
<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Explicit Mapping



# Mapping type

Meta-fields

Field or properties



# Meta-field

Metadata of document  
`_index, _type, _id, _source`



# Field or properties

List of fields or properties of document



# Mapping/Schema of document

GET /store/\_mapping/book

```
"mappings": {  
    "book": {  
        "properties": {  
            "author name": {  
                "type": "text",  
                "fields": {  
                    "keyword": {  
                        "type": "keyword",  
                        "ignore_above": 256  
                    }  
                }  
            }  
        }  
    }  
}
```



# Mapping/Schema of document

GET /store/\_mapping/book

```
"mappings": {  
    "book": {  
        "page": {  
            "type": "long"  
        },  
        "price": {  
            "type": "float"  
        },  
        "published_date": {  
            "type": "date"  
        }  
    }  
}
```



# Field Datatypes

Core

Complex

Geo

Specialized



# Field Datatypes

<b>text</b>	<b>date</b>
<b>keyword</b>	<b>ip</b>
<b>long</b>	<b>boolean</b>
<b>double</b>	<b>completion</b>
<b>geo_point</b>	<b>geo_shape</b>
<b>array</b>	<b>object</b>
<b>nested</b>	<b>binary</b>

<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping-types.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Field Datatypes

<b>text</b>	<b>date</b>
<b>keyword</b>	<b>ip</b>
<b>long</b>	<b>boolean</b>
<b>double</b>	<b>completion</b>
<b>geo_point</b>	<b>geo_shape</b>
<b>array</b>	<b>object</b>
<b>nested</b>	<b>binary</b>

<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping-types.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Array

No data type **array** in Elasticsearch

["name", "title"]	array of string
[1, 2, 3]	array of integer
[ {"name": "up1", "age": 30} ]	array of object



# Mapping configuration

Maximum number of fields = 1,000

Maximum depth of fields = 20

Maximum depth of nested fields = 50



# Dynamic/Explicit mapping

Fields and mapping types not need to defined before being used



# Mapping

## Explicit mapping

Quick data insertion

## Manual mapping

Better search Results

Increased performance

Concerned of the field types



# Custom Mapping in ES 7

Don't specify the type name !!

```
PUT project/_mapping
{
  "properties": {
    "user_id": {
      "type": "keyword"
    },
    "image_name": {
      "type": "keyword"
    }
  }
}
```



# Analyzer

## 07-analyzer

<https://www.elastic.co/guide/en/elasticsearch/reference/current/analysis.html>

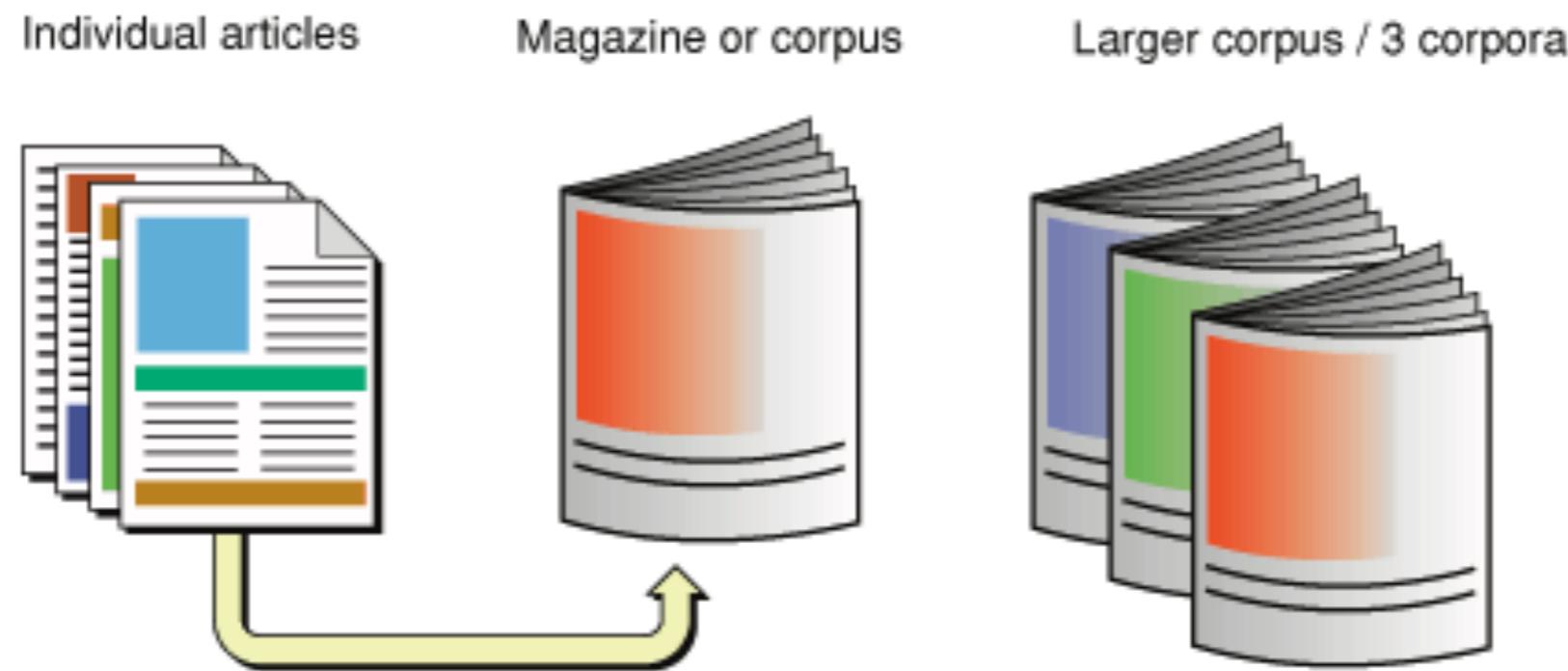


ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Inverted Index

Corpus is a collection of documents



[https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SearchKitConcepts/searchKit\\_basics/searchKit\\_basics.html#/apple\\_ref/doc/uid/TP40002843-TPXREF101](https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SearchKitConcepts/searchKit_basics/searchKit_basics.html#/apple_ref/doc/uid/TP40002843-TPXREF101)

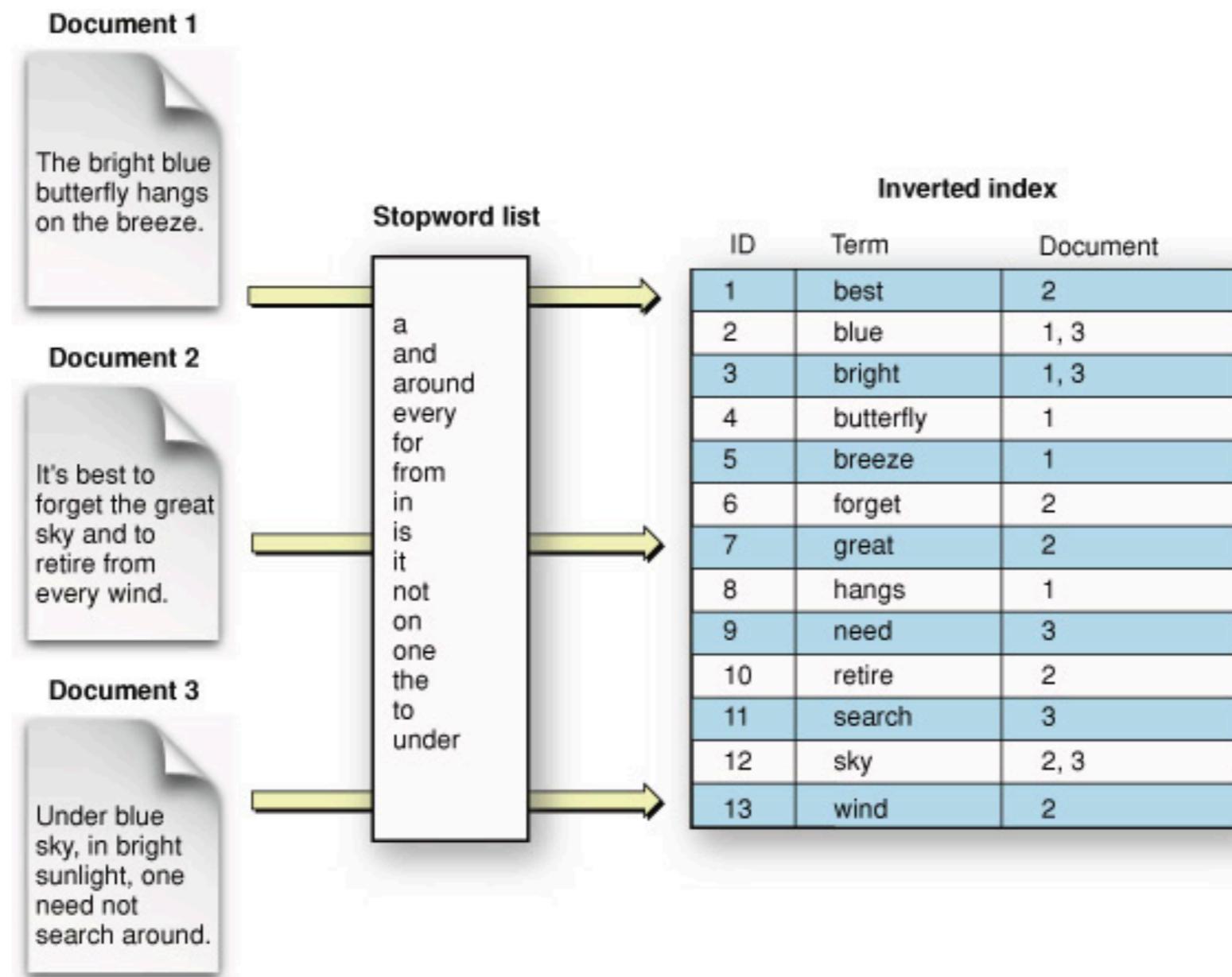


ELK Stack

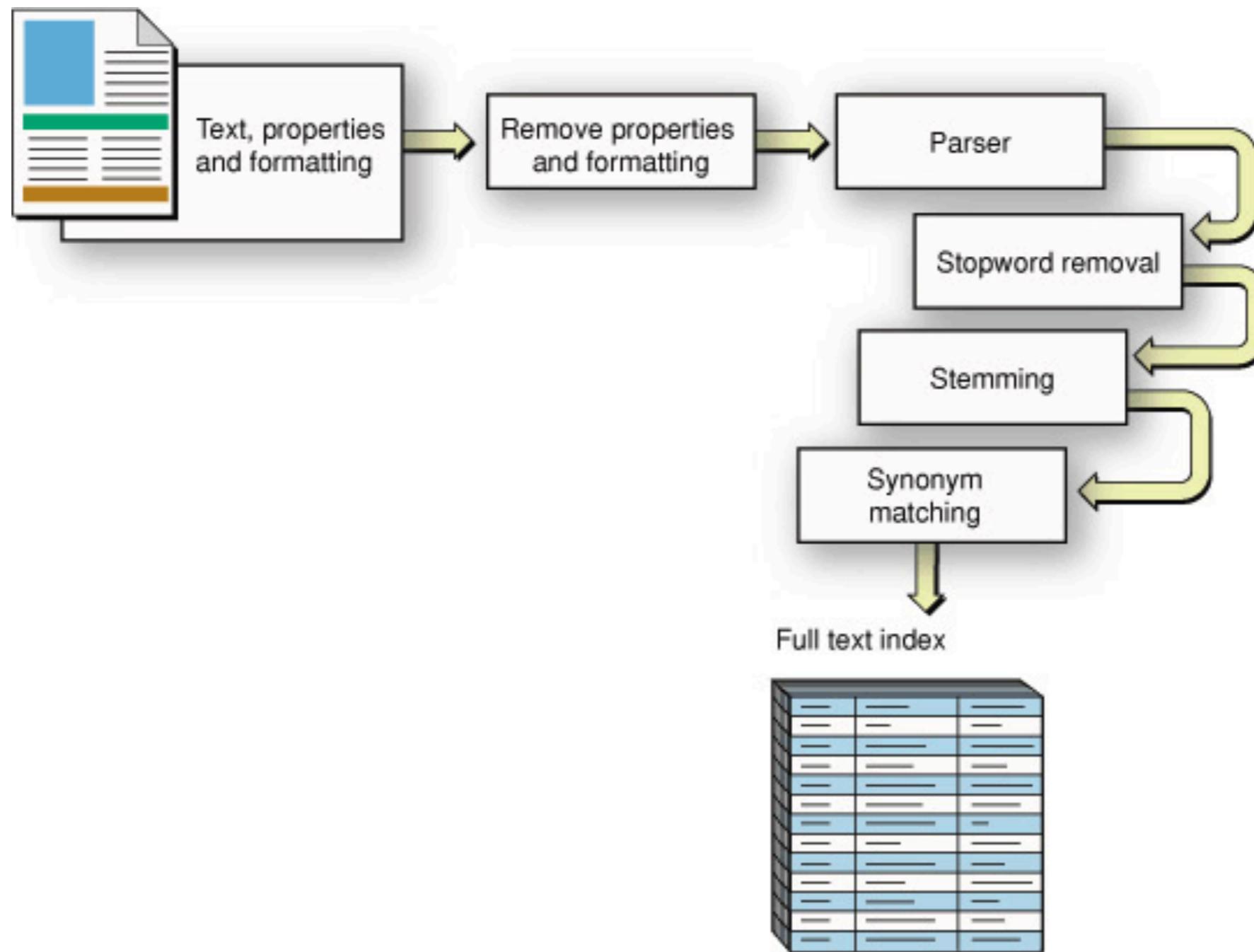
© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Inverted Index

Try to construct index



# Text extraction



# Analyzer in ES

Analyzer  
Tokenizer  
Filter



# Testing analyzer !!

Very important



# Testing analyzer !!

```
POST _analyze
{
  "analyzer": "whitespace",
  "text":      "The quick brown fox."
}
```



# Default analyzer !!

```
POST _analyze
{
  "text": "The quick brown fox."
}
```



# Thai analyzer !!

```
POST _analyze
{
  "analyzer": "thai",
  "text":      "สวัสดีประเทศไทย"
}
```



# Tokenizer and filter

```
POST _analyze
{
  "tokenizer": "standard",
  "filter": [ "lowercase", "asciifolding" ],
  "text":      "Is this déjà vu?"
}
```



# Analyze by index

```
POST my_index/_analyze
{
  "analyzer": "your_analyzer",
  "text":      "your text"
}
```



# Analyze by field

```
POST my_index/_analyze
{
  "field": "my_text",
  "text": "your text"
}
```



# Working with Suggester

<https://www.elastic.co/guide/en/elasticsearch/reference/current/search-suggesters.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Suggesters in ES

Term  
Phrase  
Completion  
Context



# Basic knowledge

N-gram tokenizer  
Edge-ngram tokenizer

<https://www.elastic.co/guide/en/elasticsearch/reference/7.1/analysis-ngram-tokenizer.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# N-Gram

Terms as a sequence of n words

search

Unigram	s,e,a,r,c,h
Bigram	se, ea, ar,rc, ch
Trigram	sea, ear, arc, rch
4-gram	sear, earc, arch
5-gram	searc, earch



# Workshop

## ngram/ngram.json



# More tools



# Elasticsearch Head

 **ElasticSearch Head**  
offered by travistx

★★★★★ (75) | [Developer Tools](#) | 45,312 users

[OVERVIEW](#) | [REVIEWS](#) | [SUPPORT](#) | [RELATED](#)

**ElasticSearch** http://192.168.7.8:9200/ Connect Rick cluster health: yellow (6, 18)

Overview Browser Structured Query Any Request Info Status Nodes Stats Cluster Nodes Cluster State Cluster Health

Cluster Overview New Index

	cu_docs	bnvil	cu_msg	anvil
Leon	size: 180Gb (540Gb) docs: 995131 (995131)	size: 80kb (480kb) docs: 90 (90)	size: 313Gb (1.56Tb) docs: 10047450 (10140915)	index: close
Pris	Info Actions	Info Actions	Info Actions	Info Actions
Rick	0 1	0 1	0 1 2 3 4	0 1 2 3 4
Rachel	1 2	0 1	0 1 2 3 4	0 1 2 3 4
Zhora	1 2	0 1	0 1 2 3 4	0 1 2 3 4
Roy	0 2	0 1	0 1 2 3 4	0 1 2 3 4
Unassigned	0 0	0 1	0 1 2 3 4	0 1 2 3 4

Compatible with your device

**ElasticSearch Head**  
Chrome Extension containing the excellent ElasticSearch Head application.

[Website](#) | [Report Abuse](#)

**Additional Information**  
Version: 0.1.3  
Updated: December 4, 2017  
Size: 434KiB  
Language: English (United States)



<https://github.com/mobz/elasticsearch-head>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Elasticsearch Dump



<https://github.com/taskrabbit/elasticsearch-dump>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Make Logs

Simple generator used to push fake HTTP traffic logs into elasticsearch

*npm install -g @elastic/makelogs*

<https://github.com/elastic/makelogs>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Alias Index

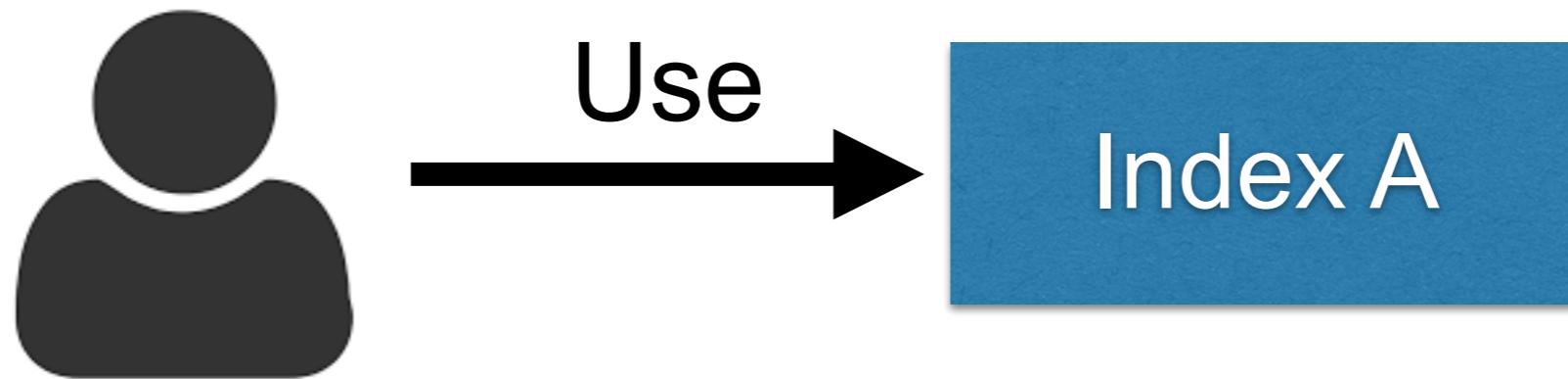


ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

180

# Common usage



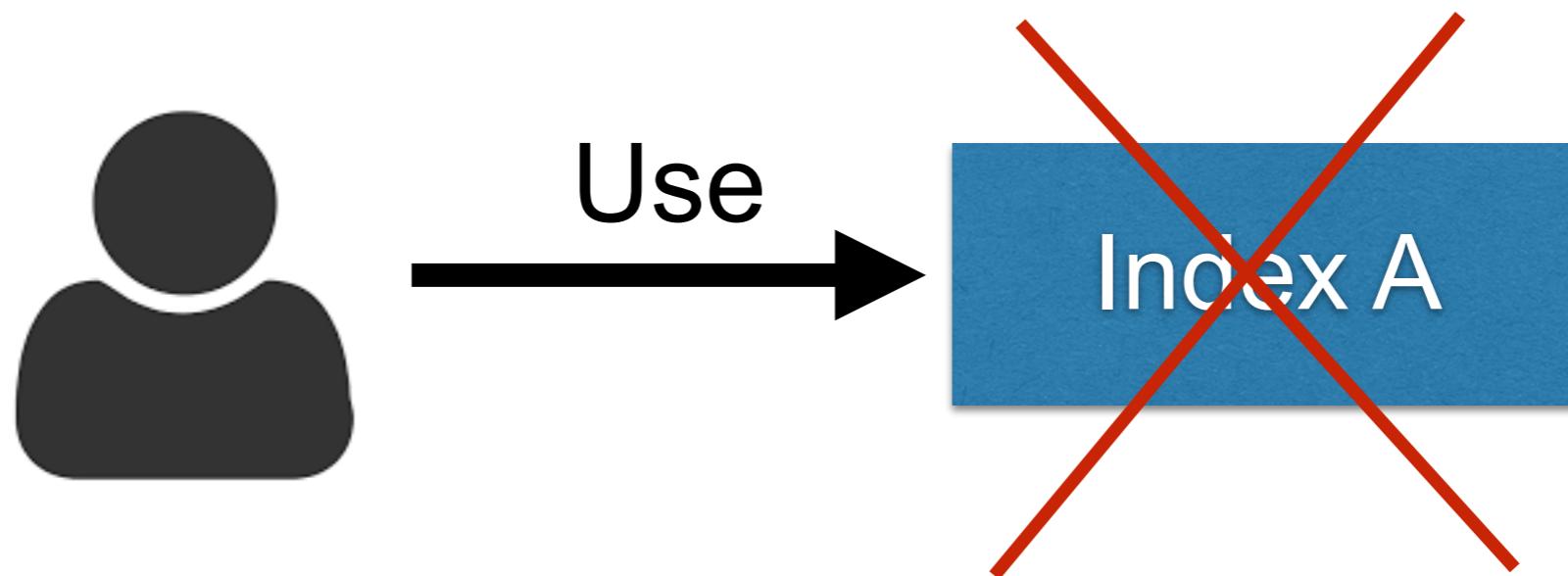
<https://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Problem ?



<https://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

# Using Alias Index



<https://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

# Using Alias Index



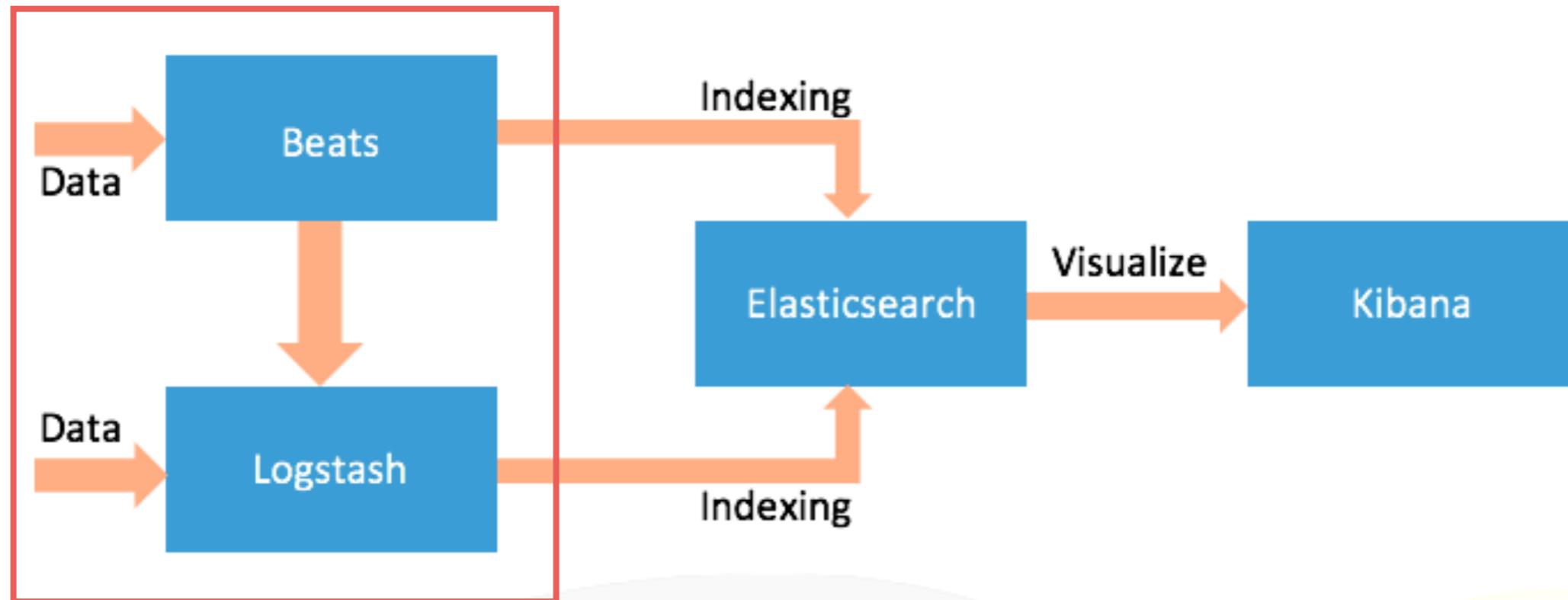
<https://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

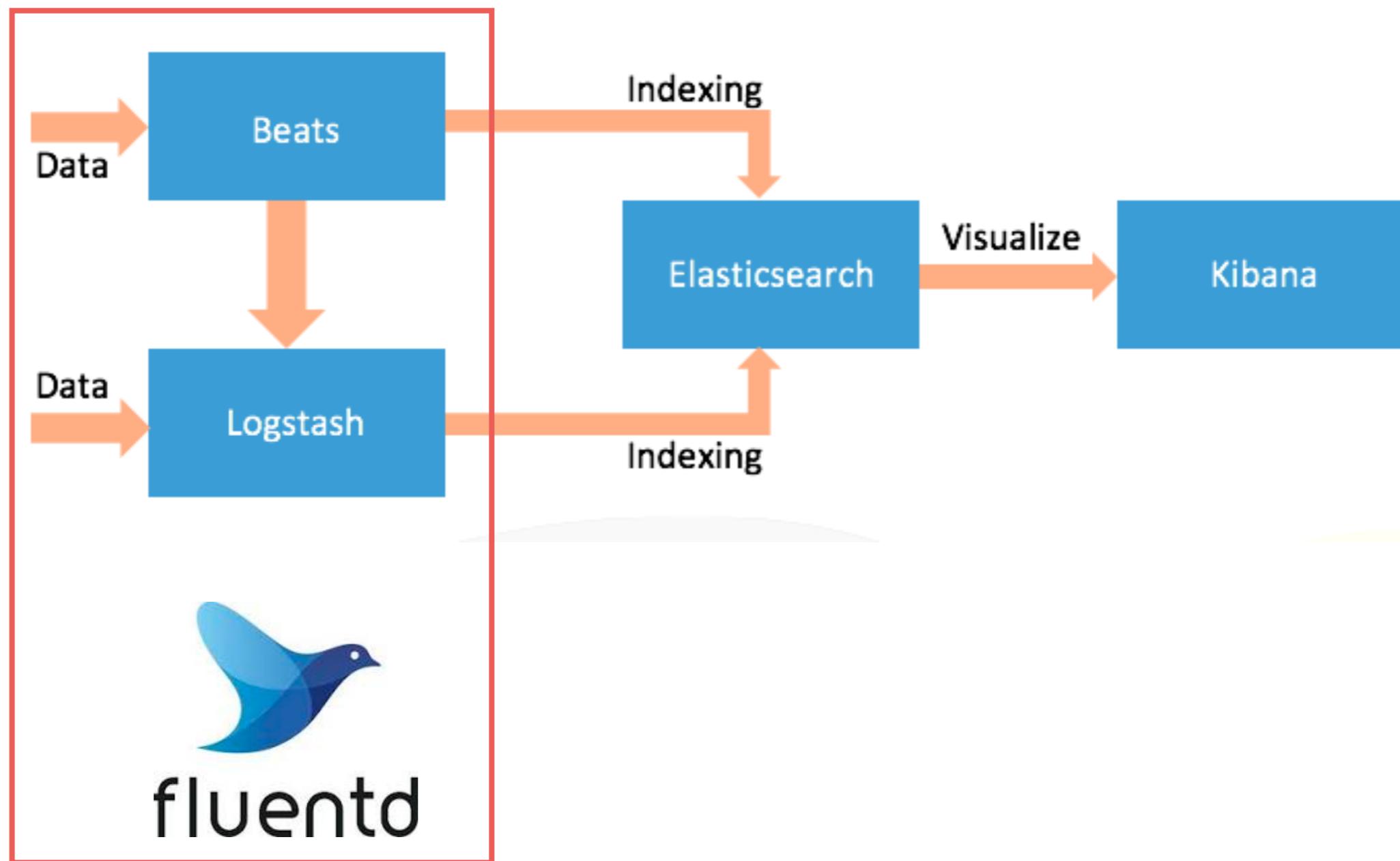
# ELK stack



fluentd



# EFK stack



# Working with Logstash

<https://www.elastic.co/guide/en/logstash/current/index.html>



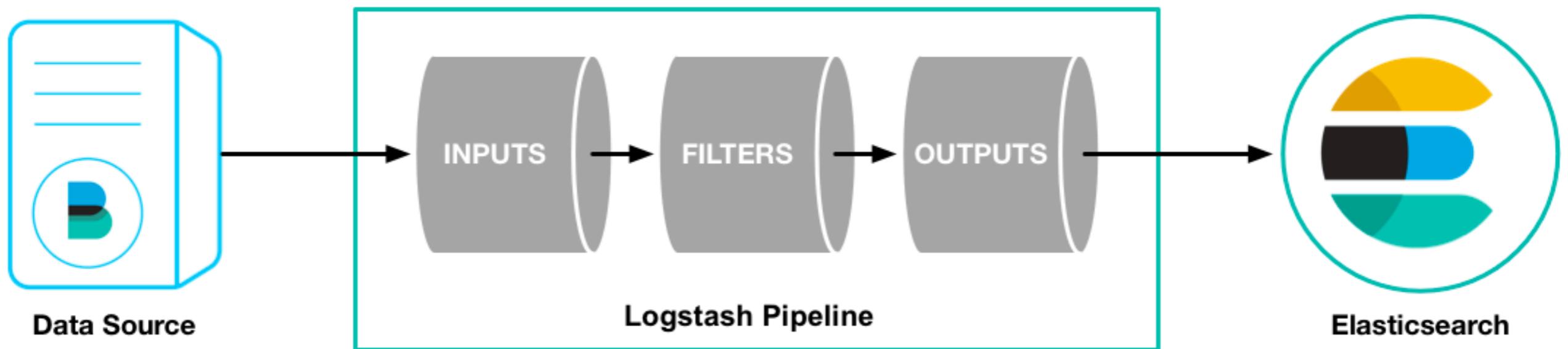
ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Logstash



# Logstash



# Example



# sample.conf

```
input {  
    stdin{  
}  
  
output {  
    stdout {  
        codec => rubydebug  
    }  
}
```



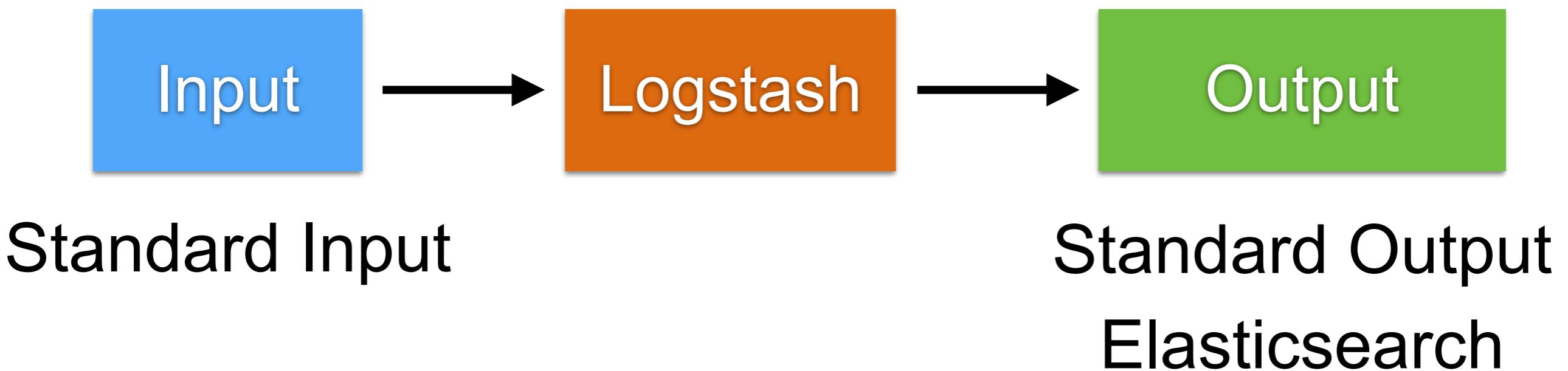
# Example

\$logstash -f sample.conf

```
hello world
{
    "message" => "hello world",
    "@timestamp" => 2019-06-20T06:01:30.048Z,
    "@version" => "1",
    "host" => "Somkiats-MacBook-Pro"
}
```



# Change output to ES



<https://www.elastic.co/guide/en/logstash/current/plugins-outputs-elasticsearch.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# sample.conf

```
input {  
    stdin{  
}  
  
output {  
    stdout {  
        codec => rubydebug  
}  
  
    elasticsearch {  
}  
}
```



# Working with Filter



<https://www.elastic.co/guide/en/logstash/current/filter-plugins.html>



# sample.conf

```
input {
    stdin{}
}

filter {
    grok {
        match => { "message" => "%{WORD:firstname} %
{WORD:lastname}" }
    }
}

output {
    stdout {
        codec => rubydebug
    }
}
```



# Workshop



File system

Standard Output

Elasticsearch

`workshop/logstash-beat-fluentd/demo.conf`



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# try.conf

```
input {  
    stdin{}  
}  
  
output {  
    stdout {  
        codec => rubydebug  
    }  
  
    elasticsearch {  
    }  
}
```



# Design your input first !!



# Use beats is better

<https://www.elastic.co/products/beats>

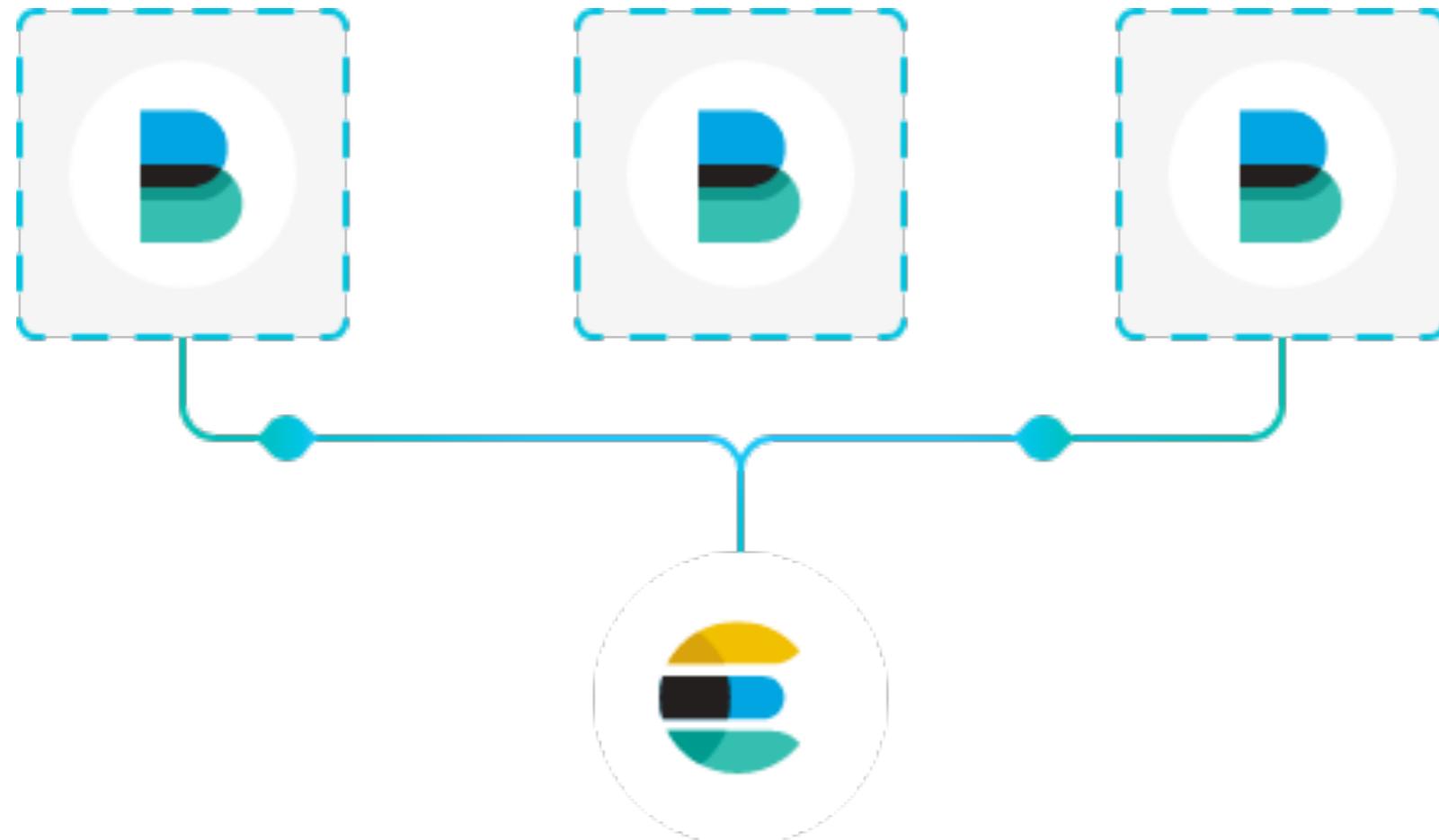


ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

200

# Beat



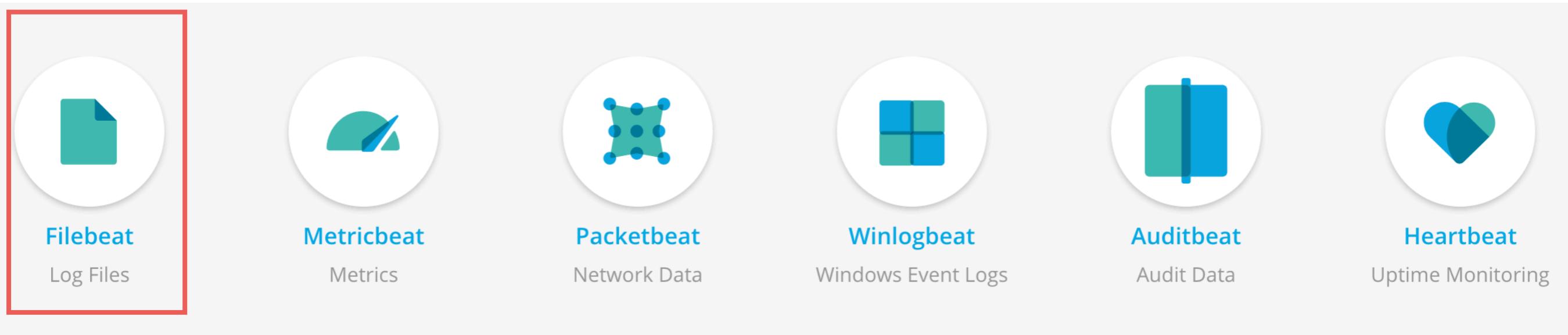
<https://www.elastic.co/guide/en/beats/filebeat/current/index.html>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

# Beat

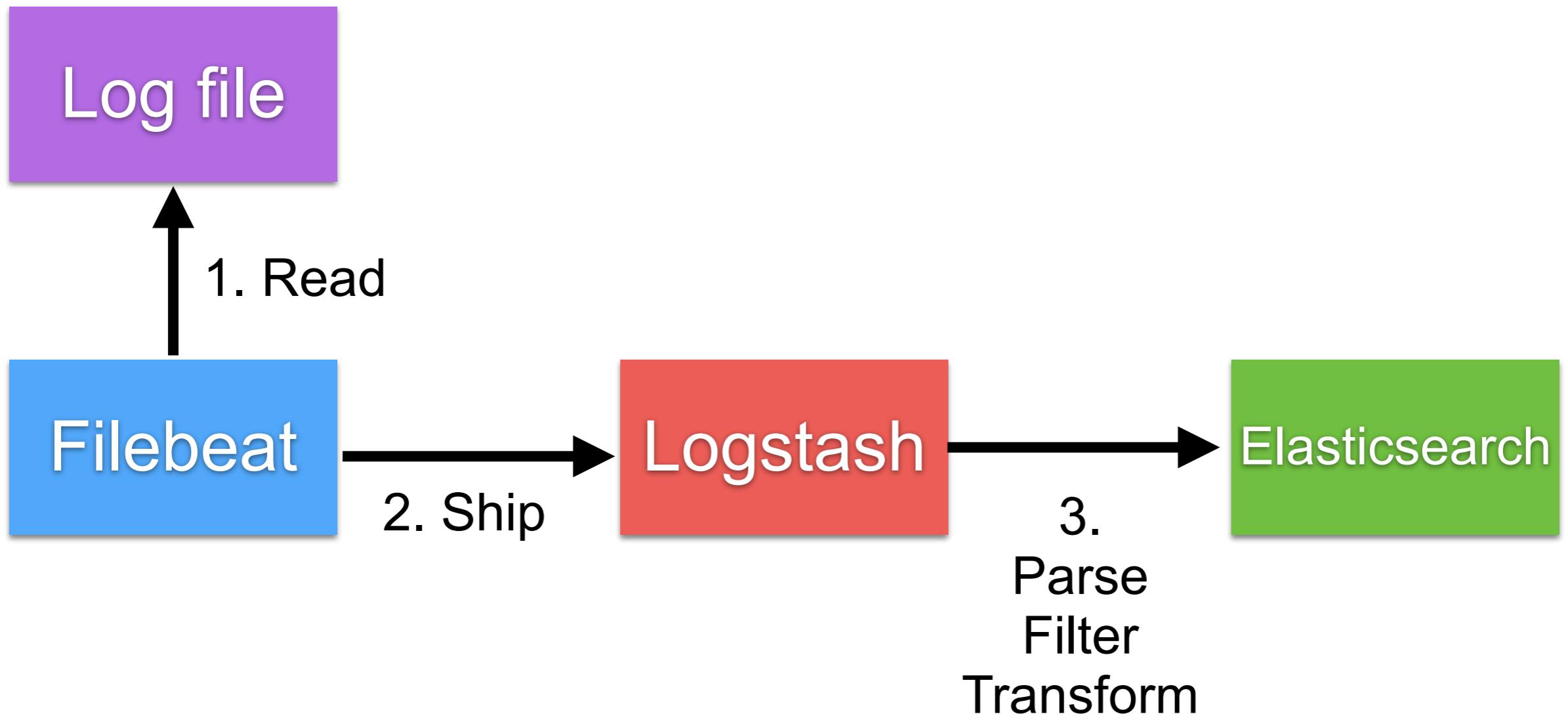


ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

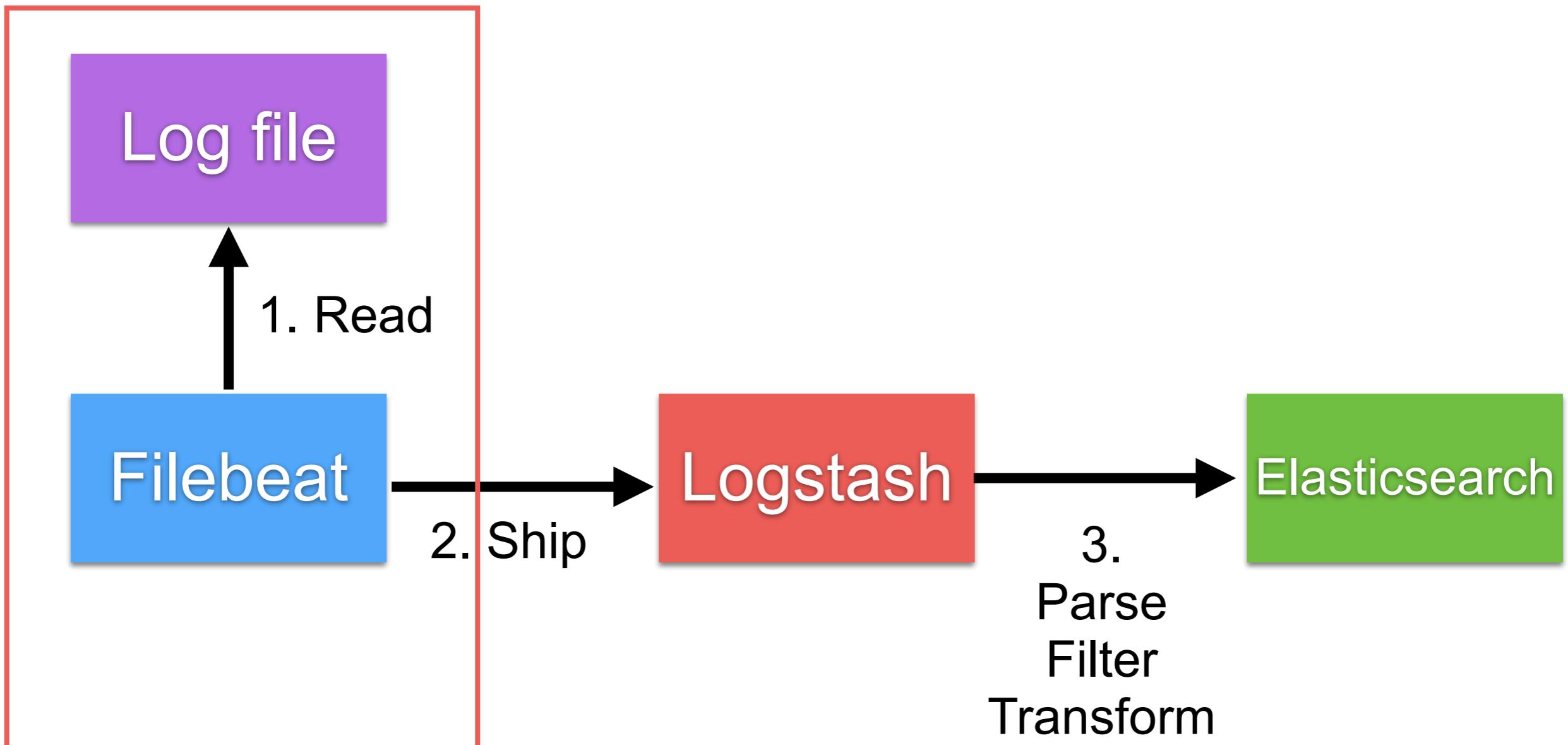
202

# Example of filebeat



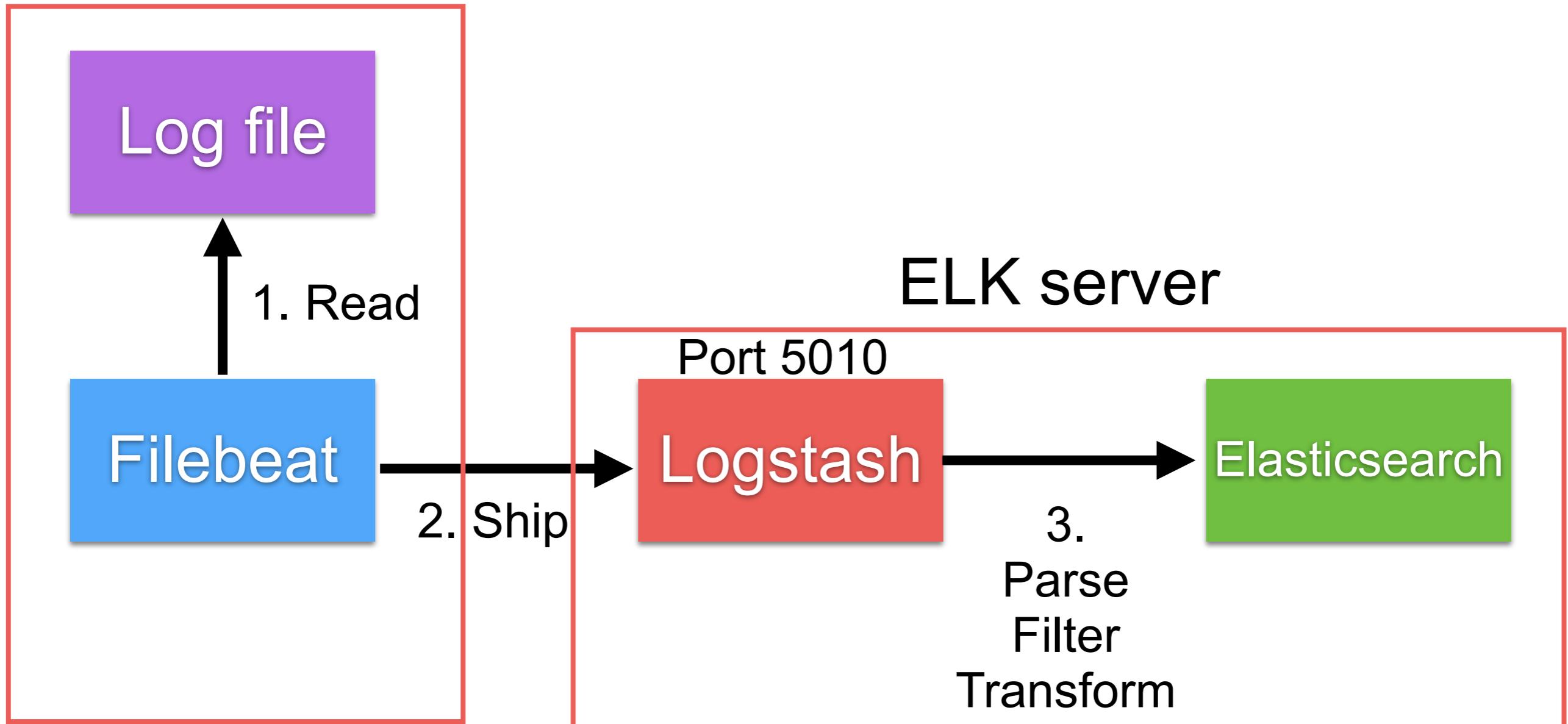
# Example of filebeat

Server A

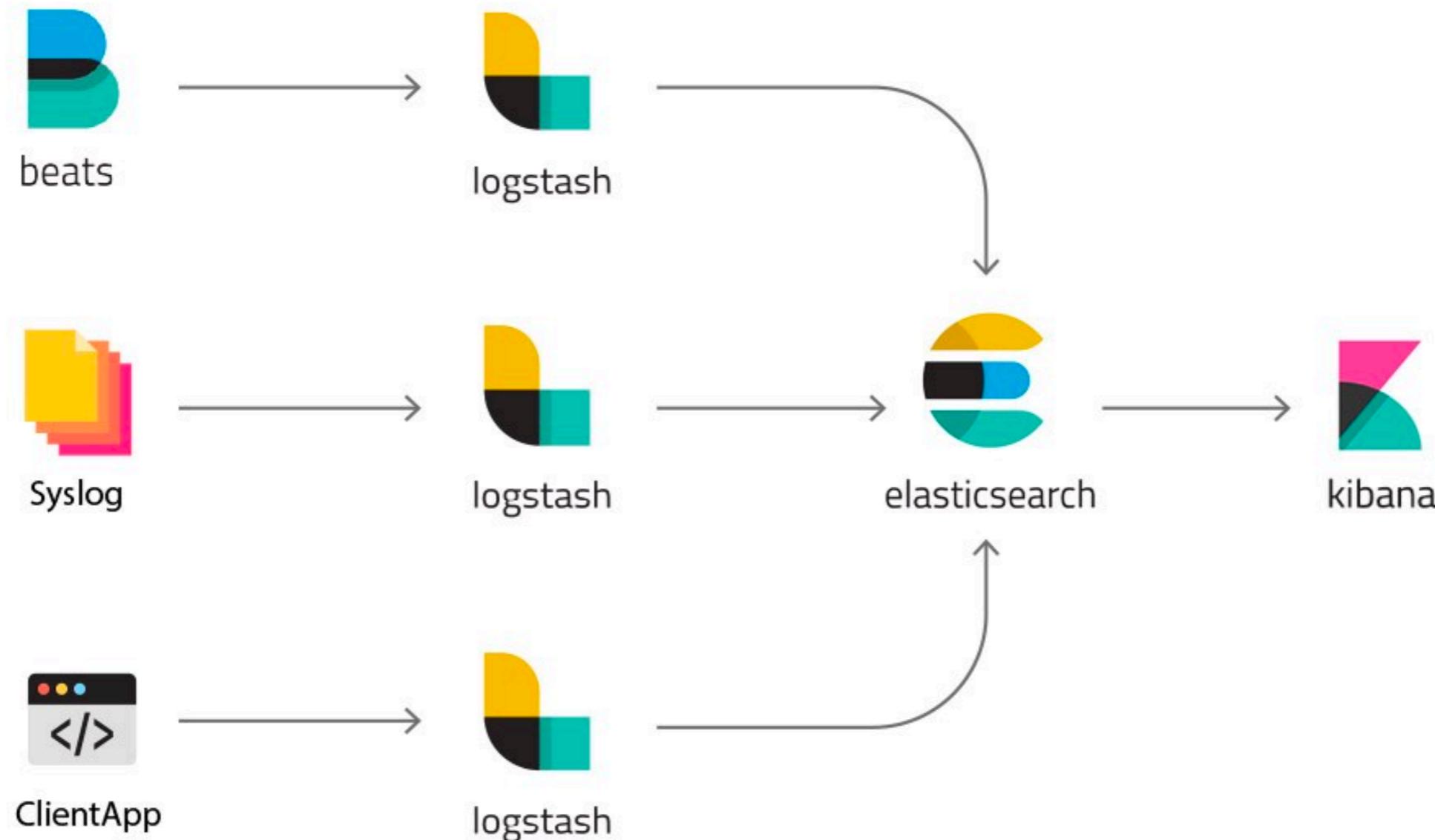


# Example of filebeat

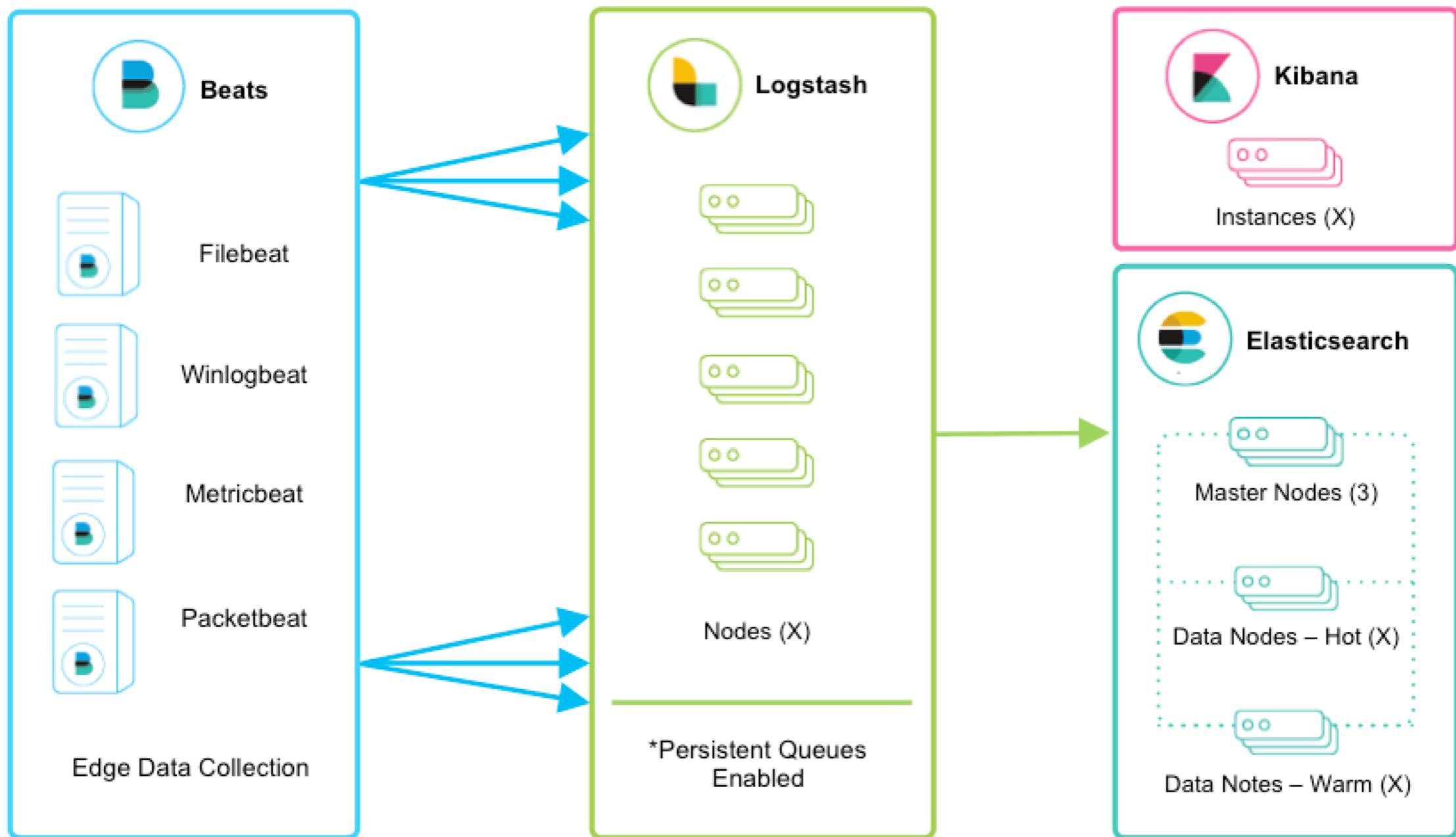
Server A



# Beat and Logstash



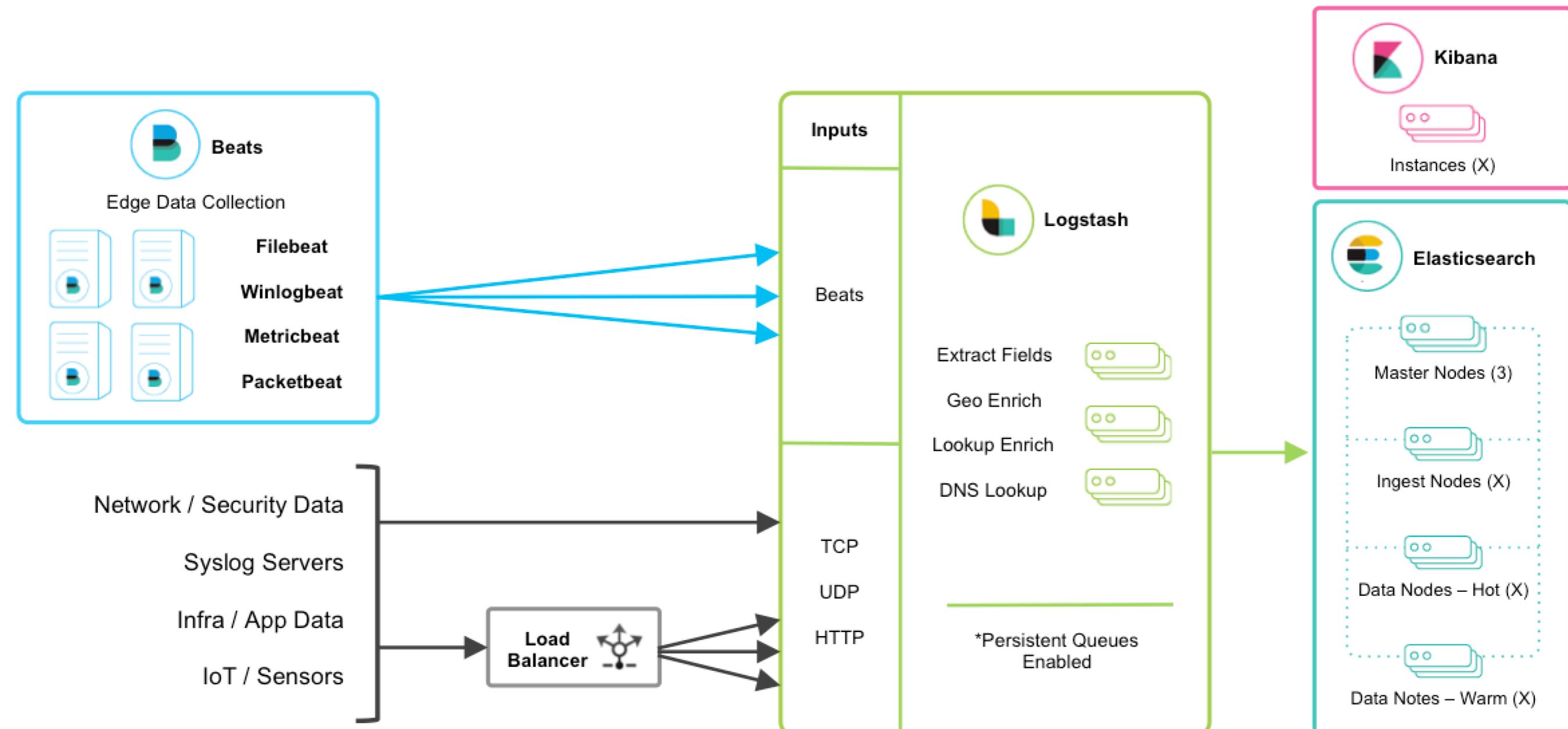
# Scaling



<https://www.elastic.co/guide/en/logstash/current/deploying-and-scaling.html>



# More data sources



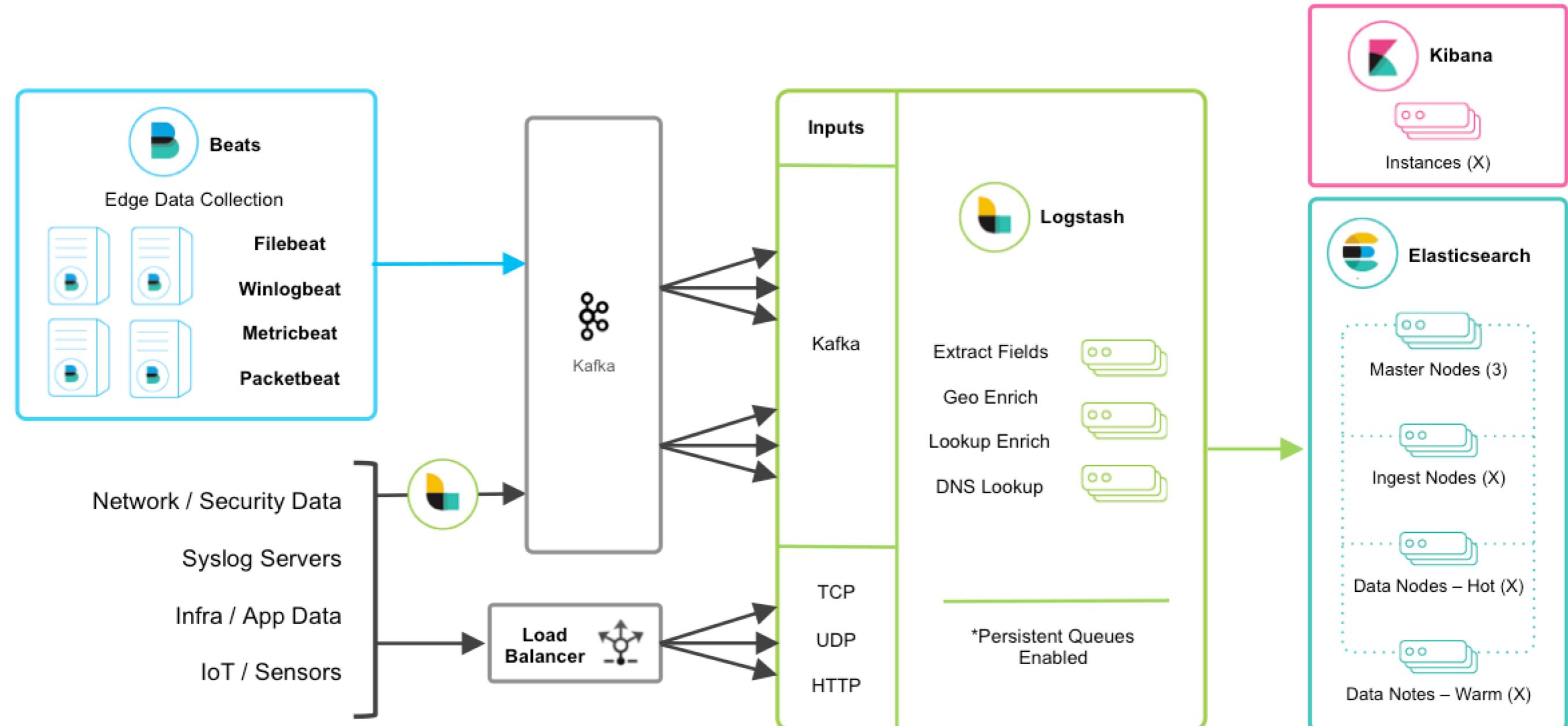
<https://www.elastic.co/guide/en/logstash/current/deploying-and-scaling.html>



ELK Stack

208

# Use messaging Queue



<https://www.elastic.co/blog/logstash-persistent-queue>

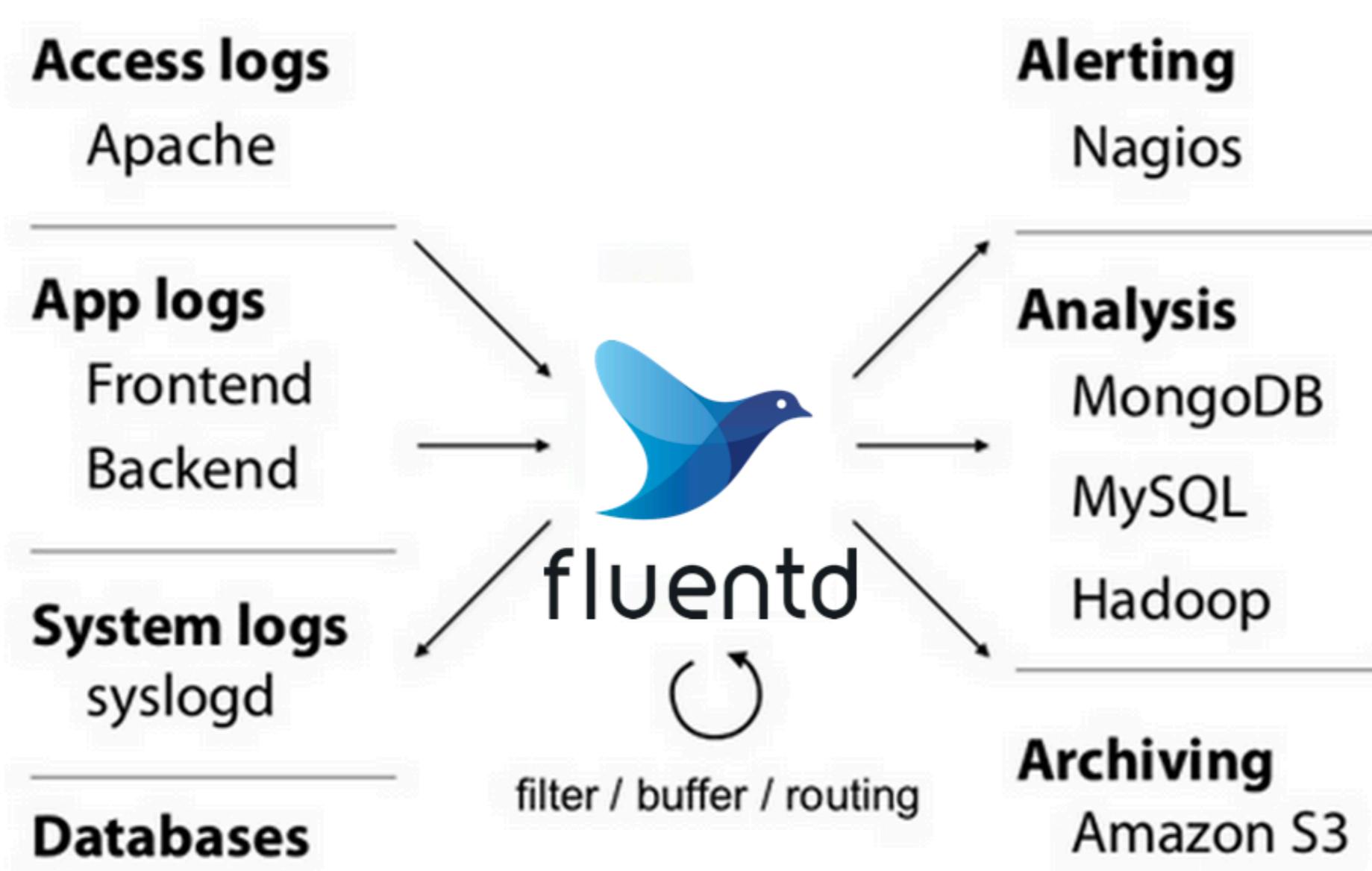


# Working with Fluentd

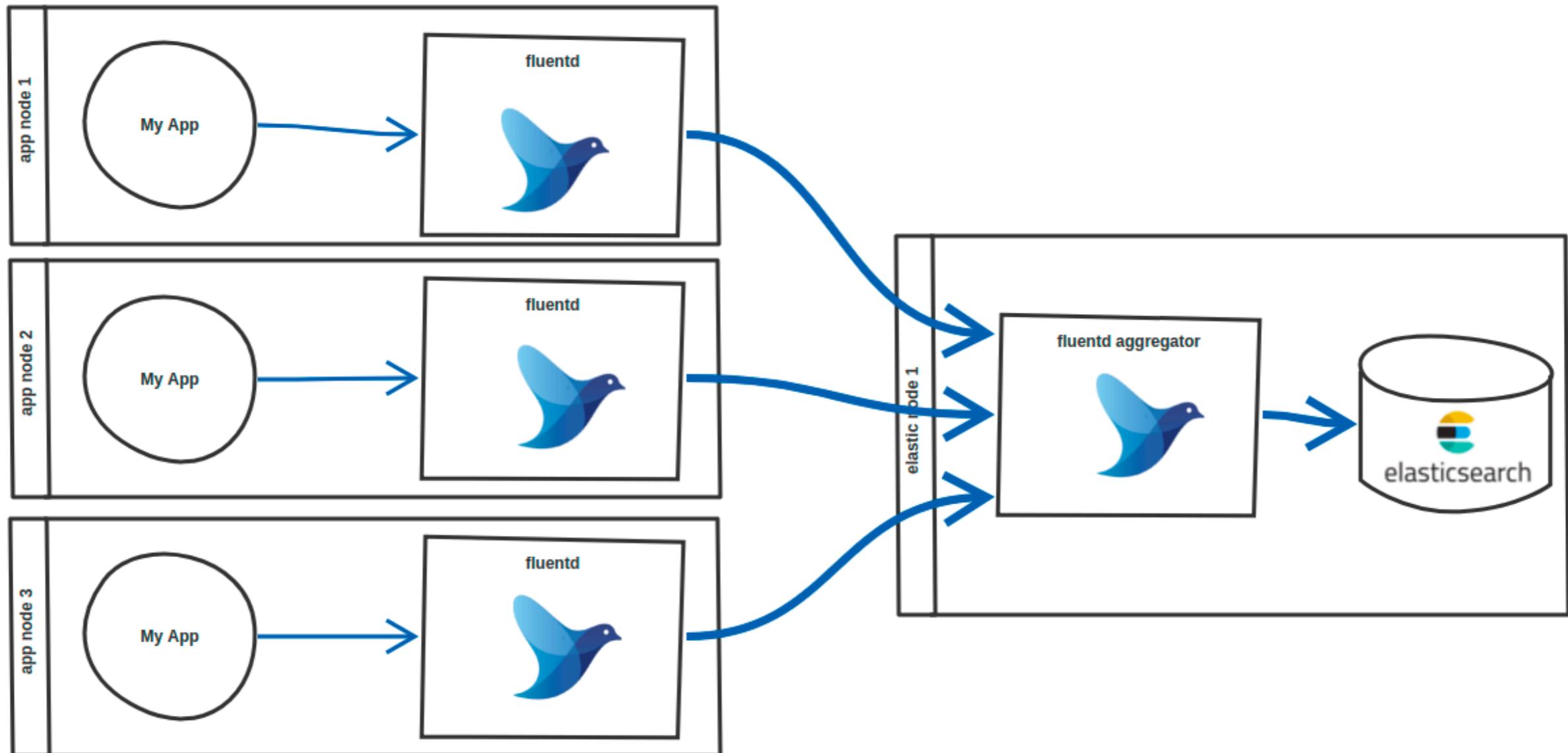
<https://www.fluentd.org/>



# Fluentd

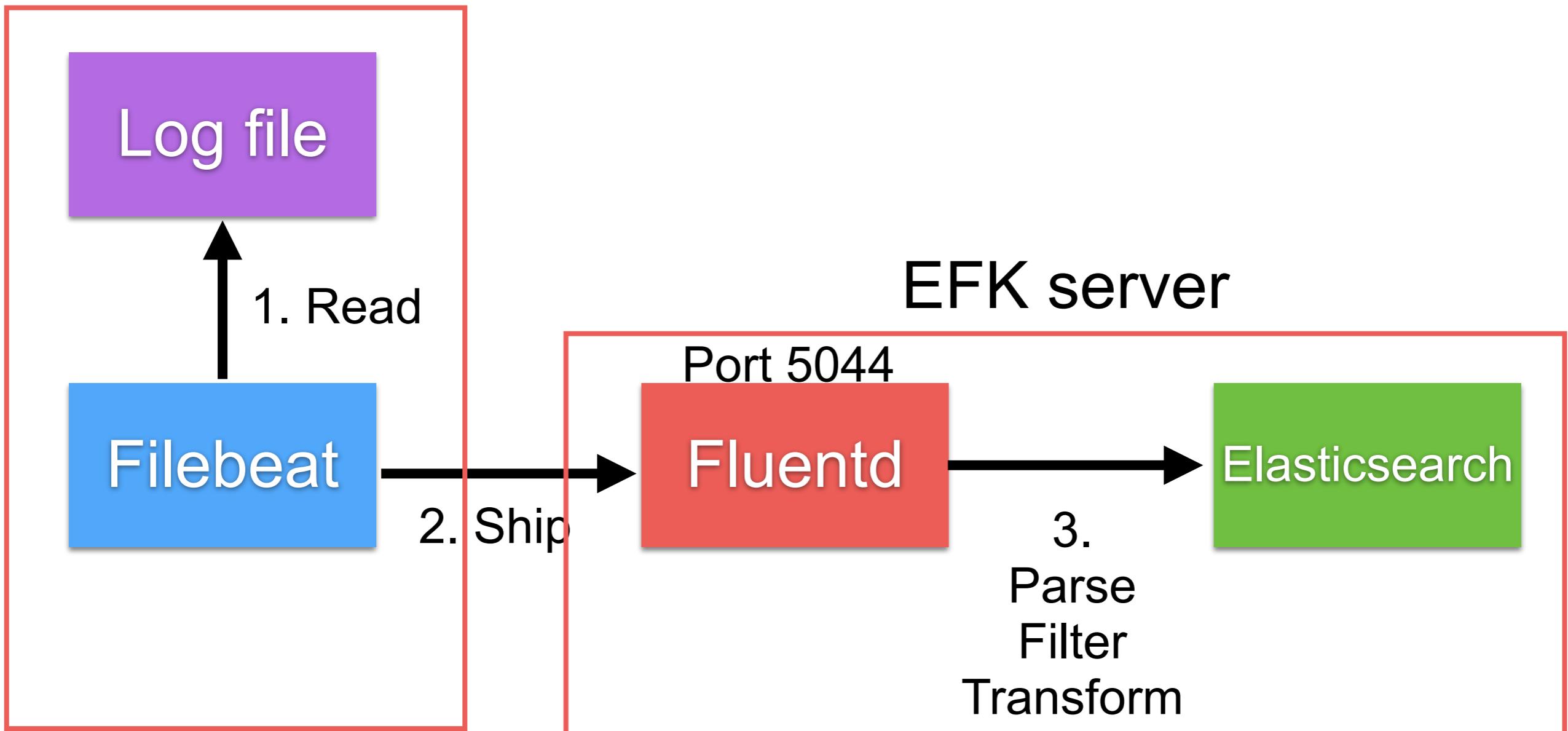


# EFK stack



# Example of fluentd

Server A



# Fluentd regex editor

**Fluentular** fluentd v1.5.2

a Fluentd regular expression editor

</> Regular Expression

“ Test String

⌚ Custom Time Format (See also ruby document; [strftime](#))

Parse

## Example (Apache)

Regular expression:

```
^(?<host>[ ^ ]*) [ ^ ]* (?<user>[ ^ ]*) \[ (?:<time>[^\\ ]*)\] " (?<method>\\S+)(?: +(?<path>[ ^ ]*) +\\S*)?" (?<code>[ ^ ]*) (?<size>[ ^ ]*)(?: "(?<referer>[ ^ \\"]*)" "(?<agent>[ ^ \\"]*)")? $
```

Time Format:

```
%d/%b/%Y:%H:%M:%S %z
```

<http://fluentular.herokuapp.com/>

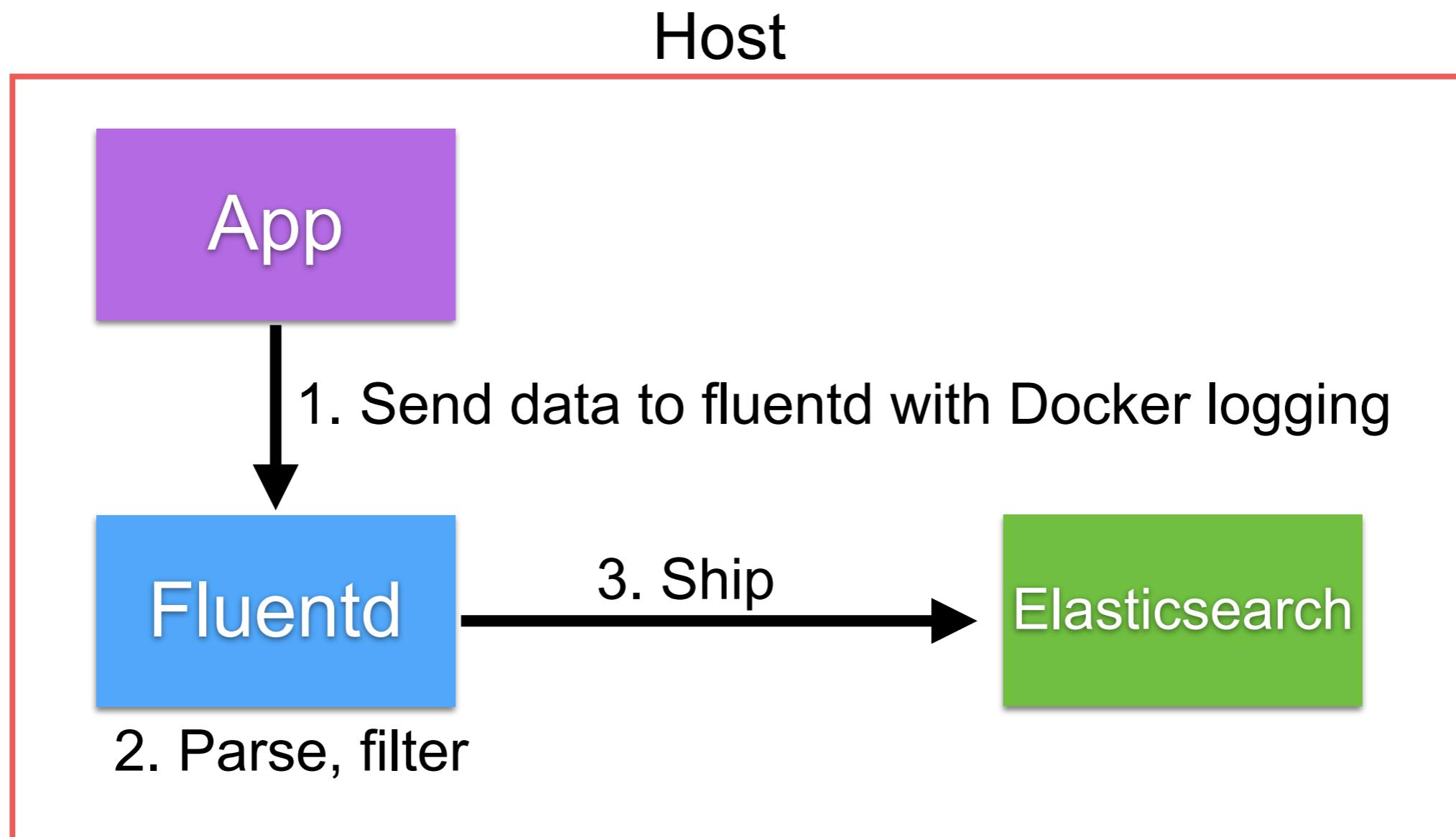


ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Fluentd with Docker

## Using docker-compose



# Design for Failure

09-cluster



# Elasticsearch Nodes

Node Type	Description
Master	Control the cluster
Data	Keep/store data
HTTP/Query	Run your query
Coordinating	Smart Load Balancer
Ingest	Pre-processing documents before indexing
Machine Learning	Required subscription !!

<https://www.elastic.co/guide/en/elasticsearch/reference/current/modules-node.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

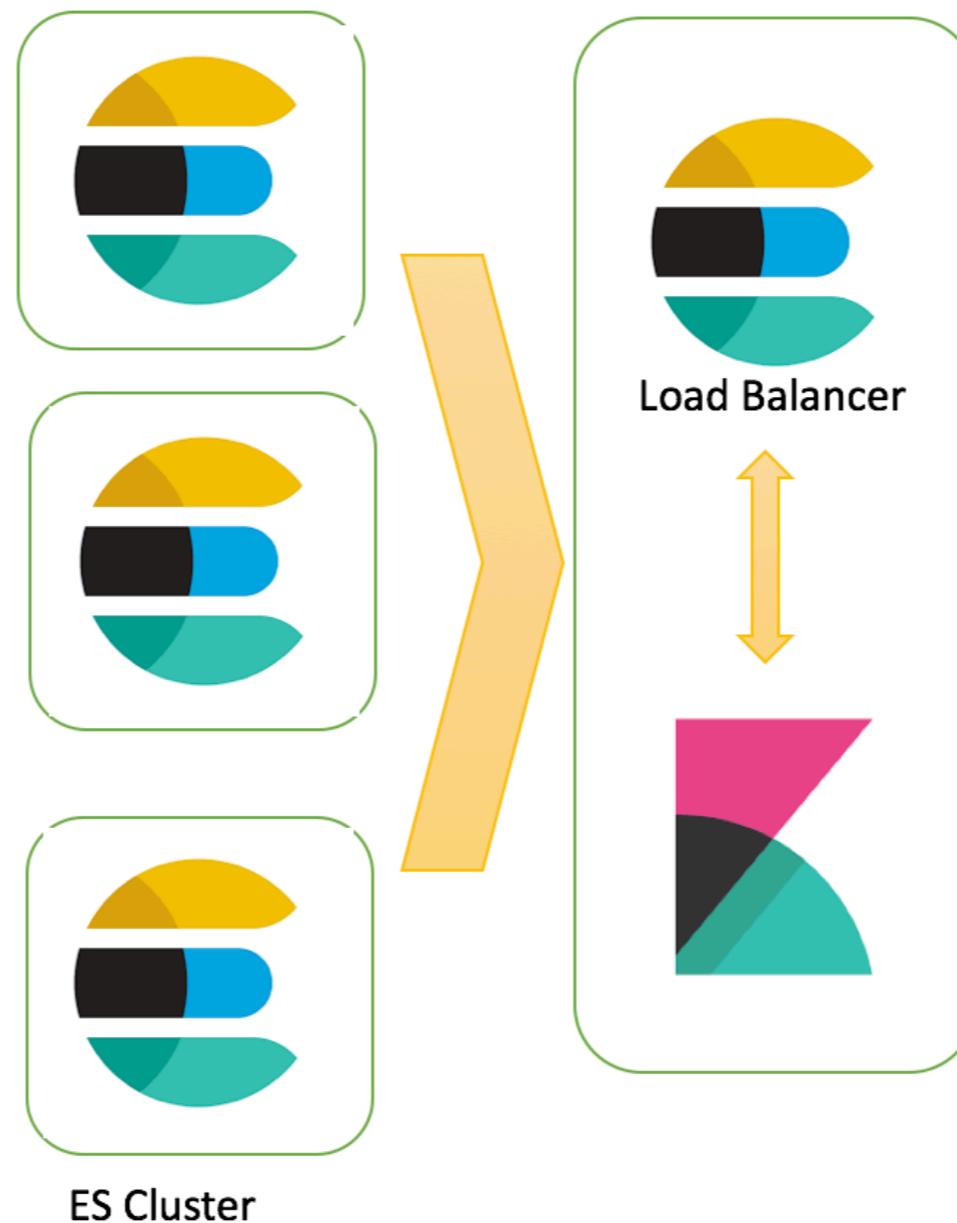
# Elasticsearch Nodes

← → ⌂ ⓘ Not Secure | 35.240.161.188:9200/\_cat/nodes?v&h=ip,name,node.role,master,heap.percent,ram.percent

ip	name	node.role	master	heap.percent	ram.percent
10.148.0.2	master	m	*	17	33
10.148.0.4	query	-	-	10	63
10.148.0.5	coordinator	-	-	9	78
10.148.0.3	data	d	-	13	63



# Elasticsearch Nodes



<https://www.elastic.co/guide/en/kibana/current/production.html#load-balancing>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Elasticsearch Nodes

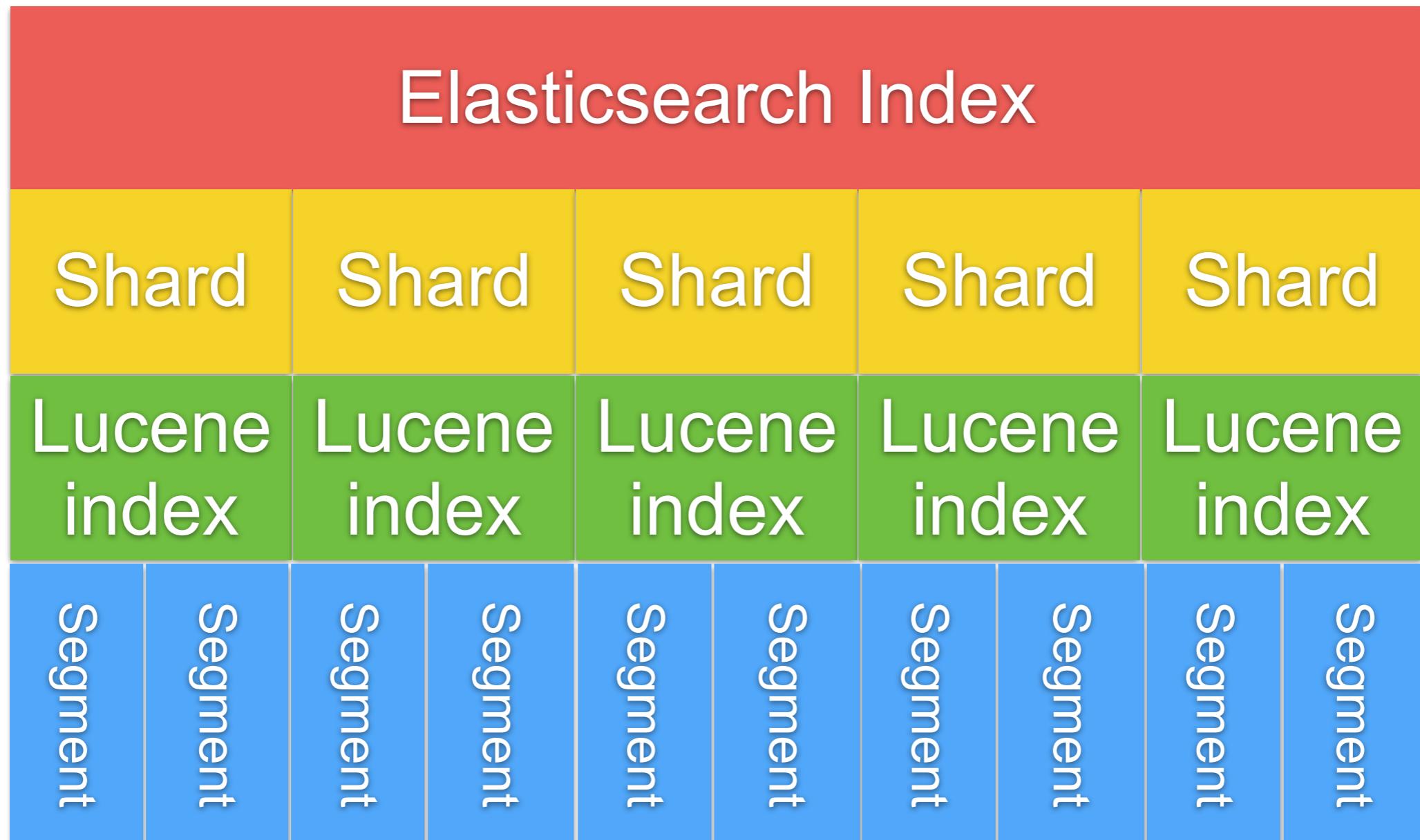


# Apache Lucene

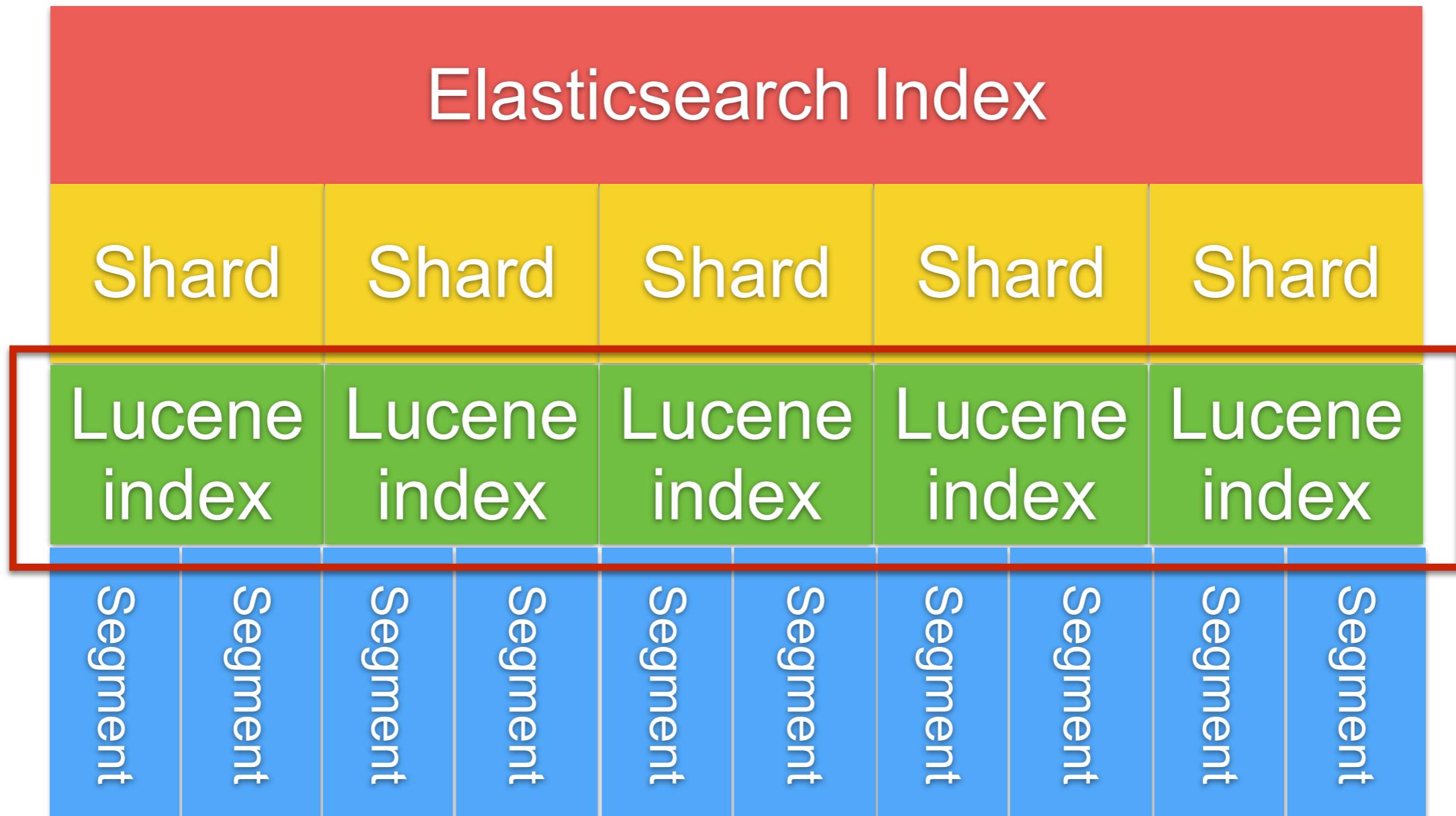
<http://lucene.apache.org/>



# Apache Lucene



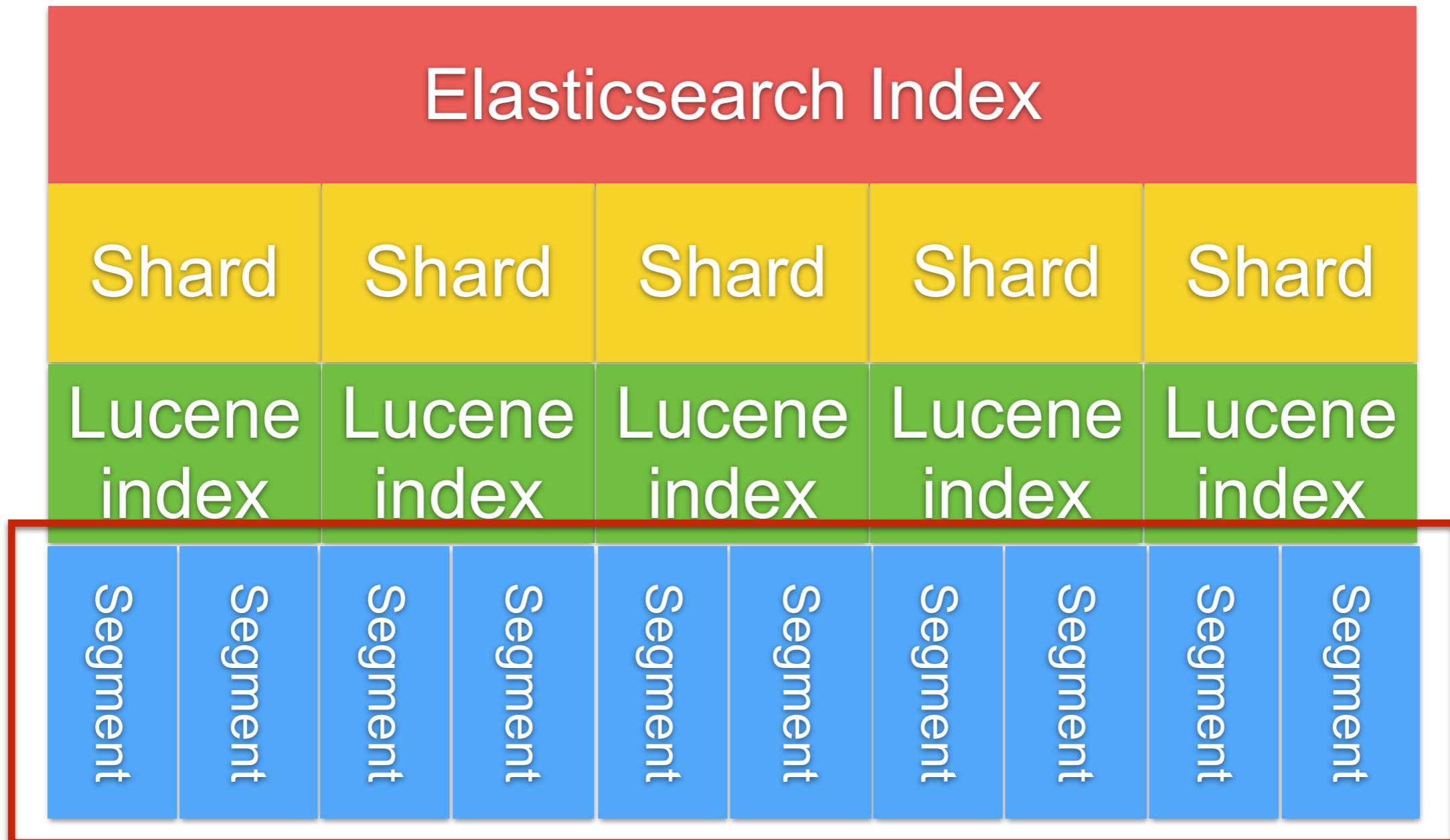
# Apache Lucene



Max # of document of Lucene index = 2,147,483,519



# Apache Lucene



Segments are immutable

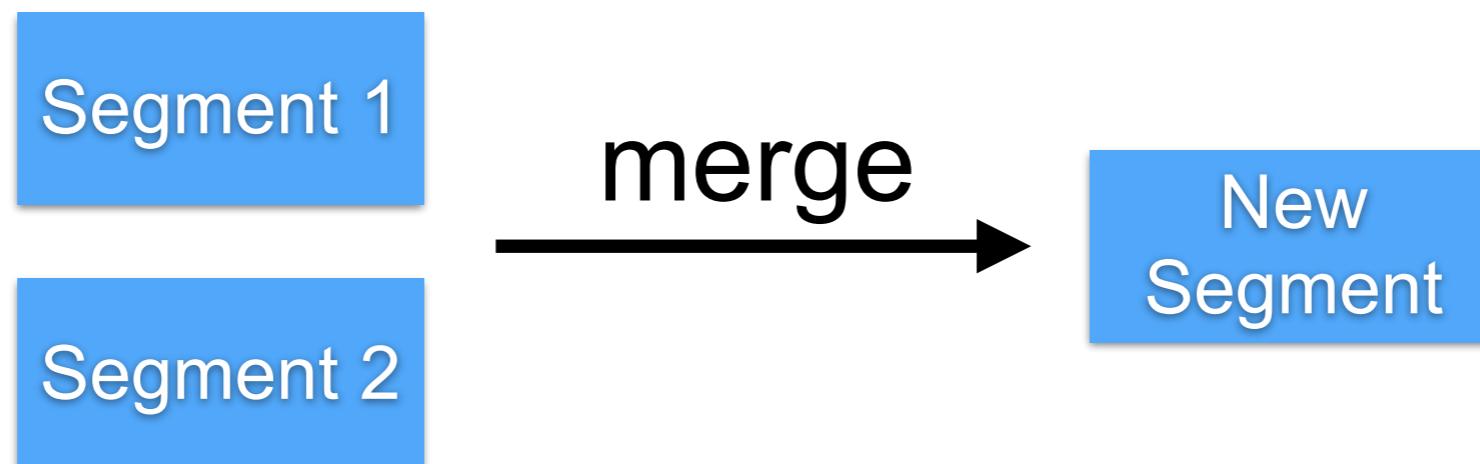


# Segment

More shards, more segments

Documents are never delete !!

Lucene segment **merge** use more CPU/IO  
Segments are immutable



# Hardware



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

226

# **Hardware**

CPU  
**Memory**  
Network  
Storage



# Memory

Enable bootstrap.memorylock

Disable all swap files

Change **ES\_HEAP\_SIZE** (default 1G)

<https://www.elastic.co/guide/en/elasticsearch/reference/current/setup-configuration-memory.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Design your index



# Design your index

Sharding  
Replication



# Sharding

Elasticsearch divides the data in **logical** parts  
# of sharding is define when index created



# How many shard ?



# Need to know your size of data

Data Size	# of shard
< 3M	1
>3M <5M	2
>5M	(# of document / 5M) + 1



# Sharding

**Small shards** on multiple nodes make the cluster recovery faster

**Small shards** on a lot of nodes solve memory mgt problem when query on large data



# More shard, more Segment !!

Elasticsearch Index				
Shard	Shard	Shard	Shard	Shard
Lucene index	Lucene index	Lucene index	Lucene index	Lucene index
Segment	Segment	Segment	Segment	Segment

Need to config file descriptor

<https://www.elastic.co/guide/en/elasticsearch/reference/current/system-config.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# **Don't create more shard than you need !!**



# Replication

Prevent data loss

Default = 1

```
# nodes = [(primary + # replication) /2 ] + 1
```



# Problems with scaling

CPU consumption  
Load average  
Request rate  
Search latency



# Slow log

```
PUT /myindex/_settings
```

```
{  
  "index.search.slowlog.threshold.query.warn: 1s",  
  "index.search.slowlog.threshold.query.info: 500ms",  
  "index.search.slowlog.threshold.query.debug: 1500ms",  
  "index.search.slowlog.threshold.query.trace: 300ms",  
  "index.search.slowlog.threshold.fetch.warn: 500ms",  
  "index.search.slowlog.threshold.fetch.info: 400ms",  
  "index.search.slowlog.threshold.fetch.debug: 300ms",  
  "index.search.slowlog.threshold.fetch.trace: 200ms"  
}
```

*If can't optimize then add more resources or rewrite*

<https://www.elastic.co/guide/en/elasticsearch/reference/current/index-modules-slowlog.html>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Indexing Data



# Indexing data

Must be define data schema for your need  
Default mapping == more cost (Memory/Disk)  
Default for data is “text” + “keyword”  
Understand analyzer and tokenizer  
Use auto generated IDs if possible



# Indexing data

Prefer bulk indexing

**Change refresh interval**

Time based index for log data

<https://www.elastic.co/guide/en/elasticsearch/reference/current/indices-update-settings.html>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

# For Large data

Increase refresh interval  
Decrease replica number

```
PUT /logstash-2015.05.20/_settings
{
  "index" : {
    "refresh_interval" : "-1",
    "number_of_replicas" : 0
  }
}
```

<https://www.elastic.co/guide/en/elasticsearch/reference/current/indices-update-settings.html>



# Query Data



# Query data

Use filters as much as possible  
Use scan and scroll for dumping large data  
    Node query cache  
    Shard query cache  
Retrieve only necessary fields

<https://www.elastic.co/guide/en/elasticsearch/reference/current/query-cache.html>



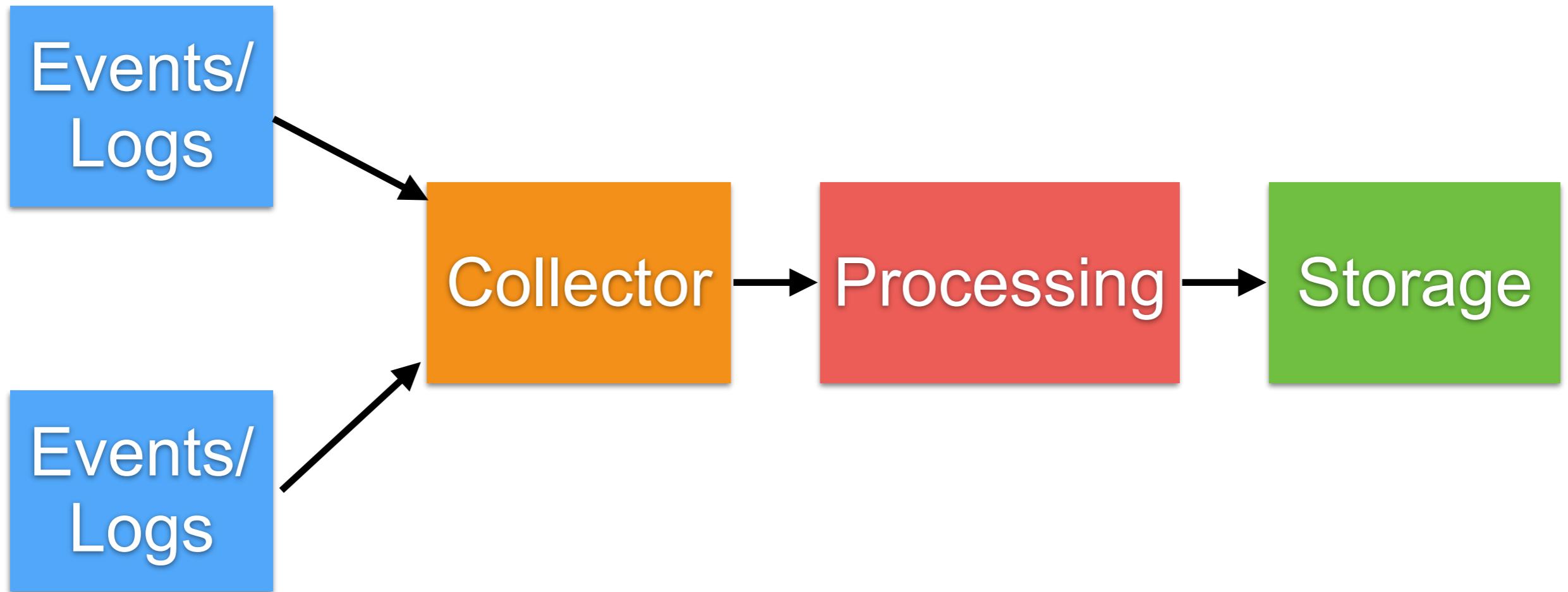
ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

# Use cases



# Event or Logging from Servers



# Event or Logging from Servers



Data  
Collection

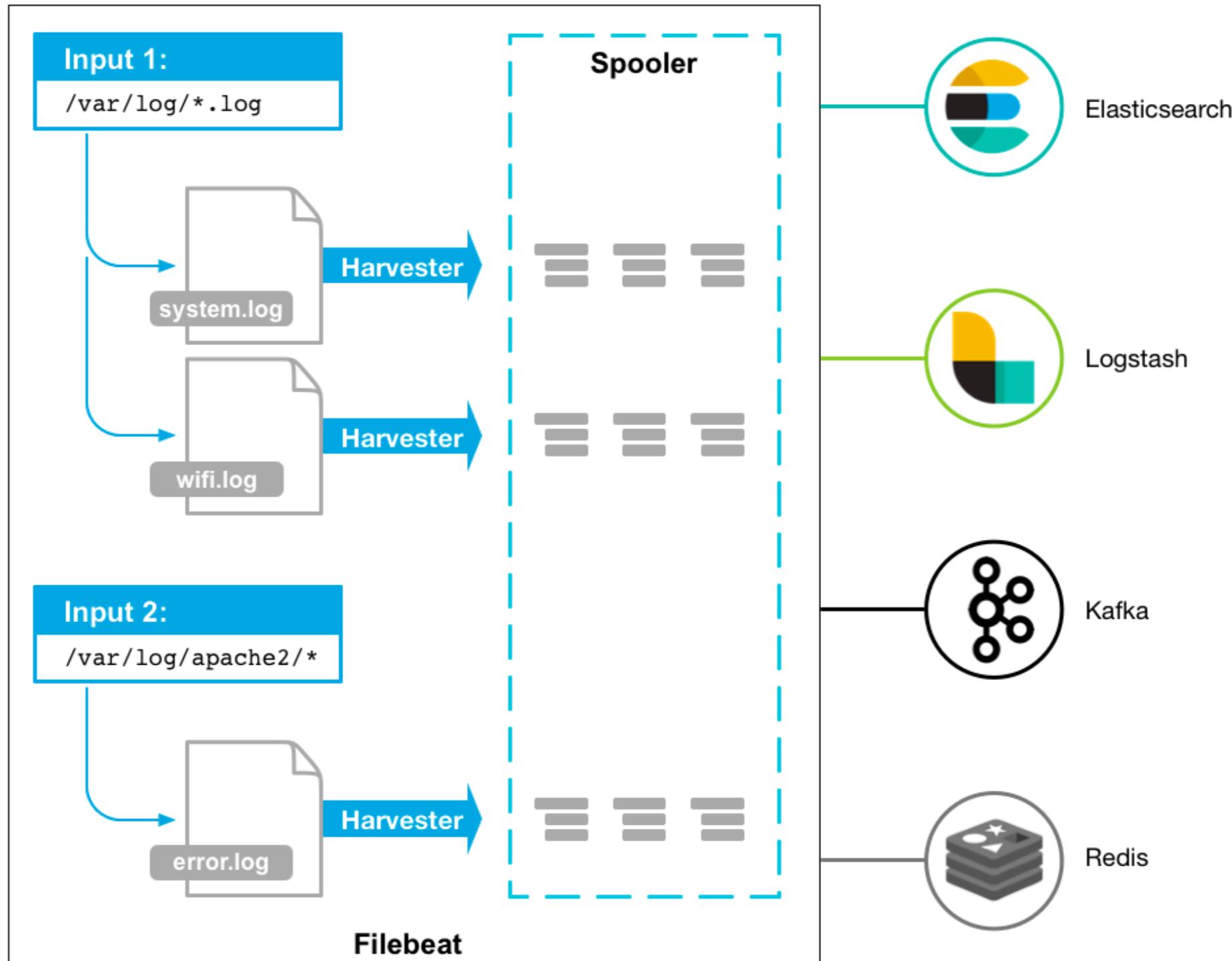
Data  
Aggregation  
& Processing

Indexing &  
storage

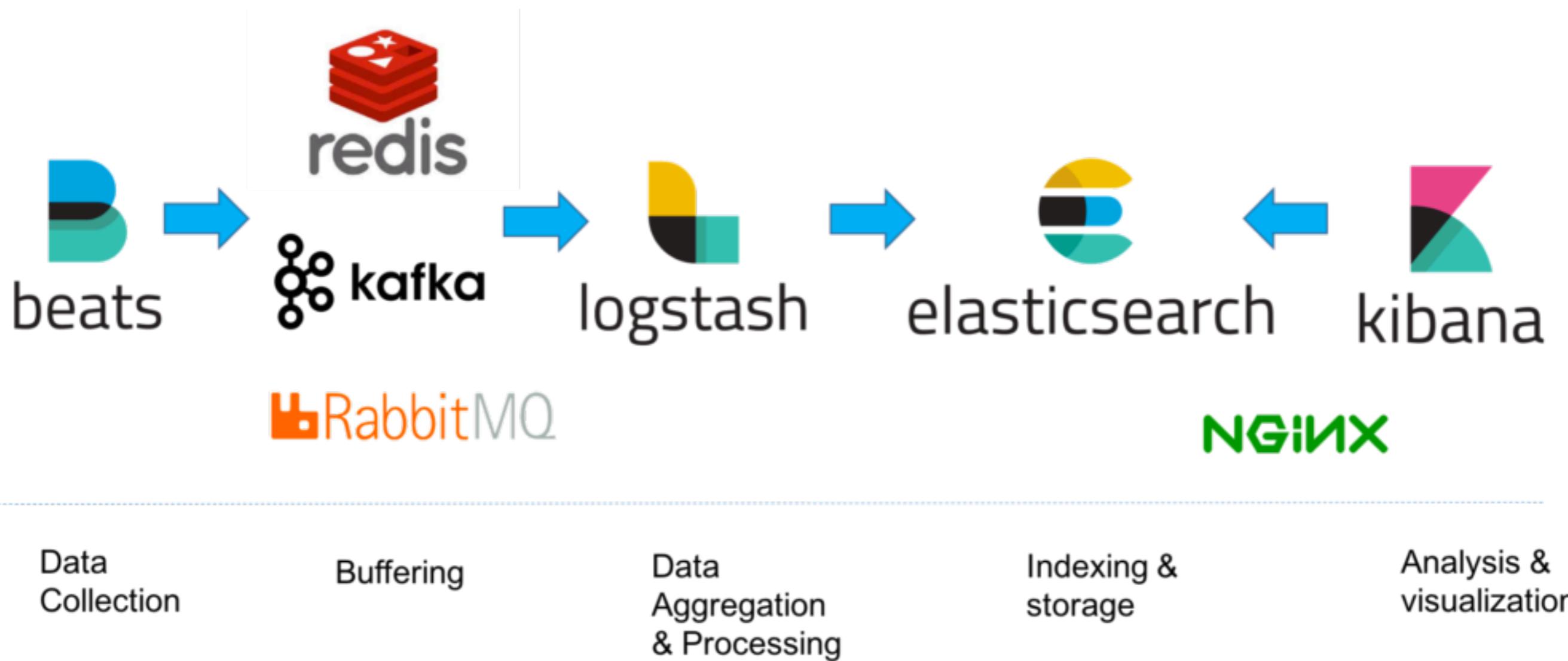
Analysis &  
visualization



# Working with Filebeat



# Event or Logging from Servers



# Monitoring



# Collect data from ?

Elasticsearch nodes

Logstash nodes

Kibana instances

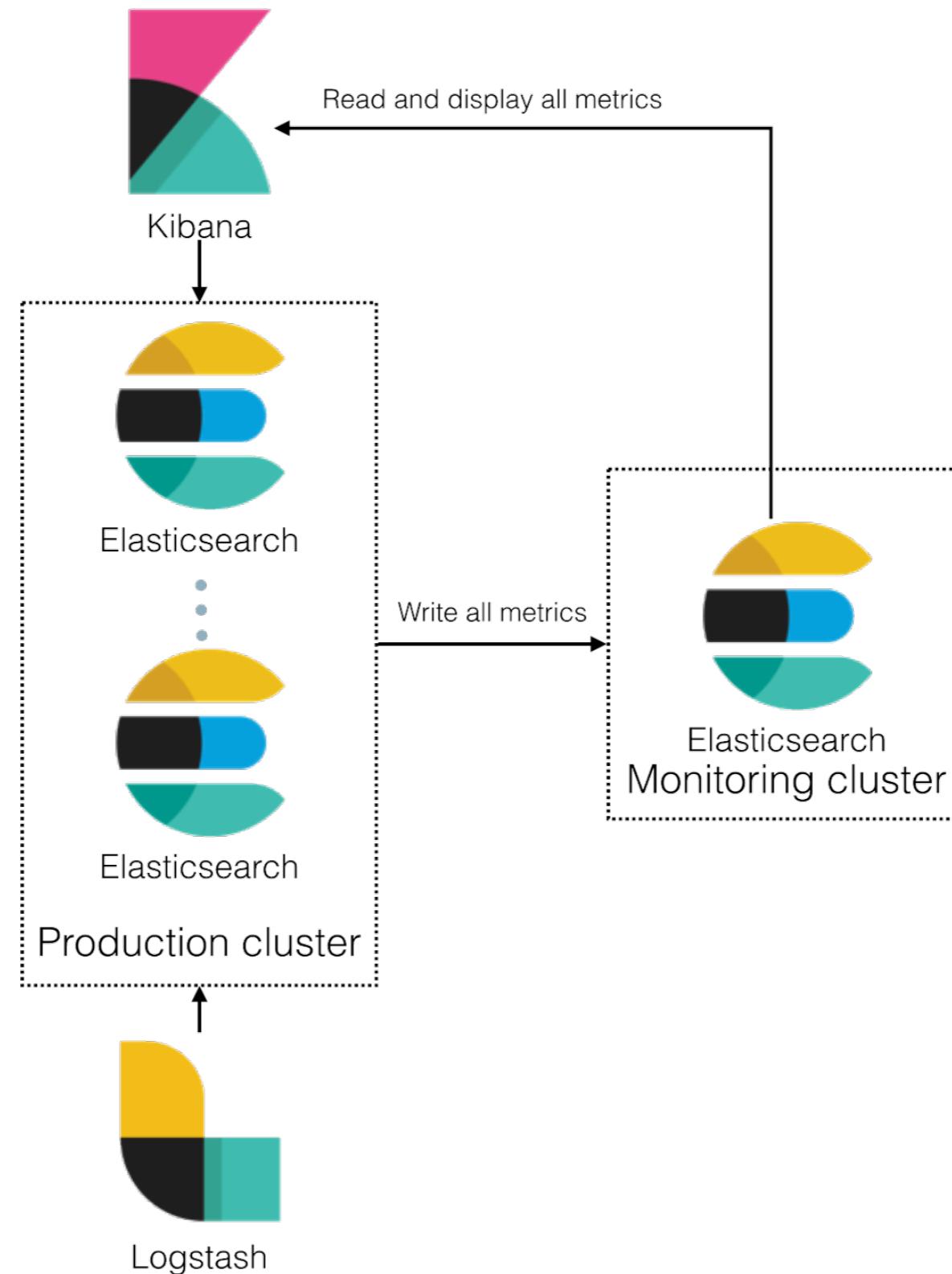
<https://www.elastic.co/guide/en/elasticsearch-stack-overview/6.5/xpack-monitoring.html>



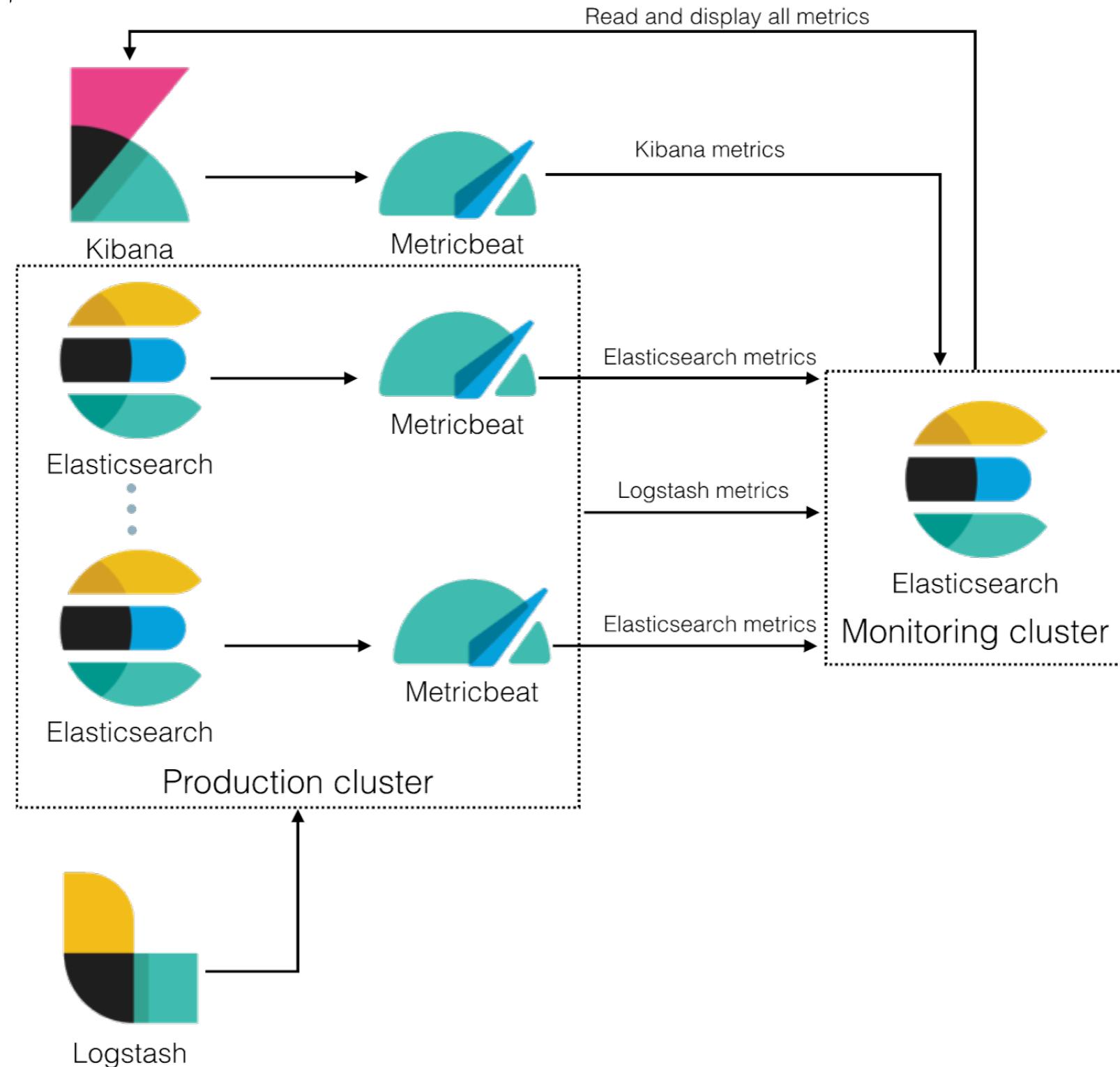
ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

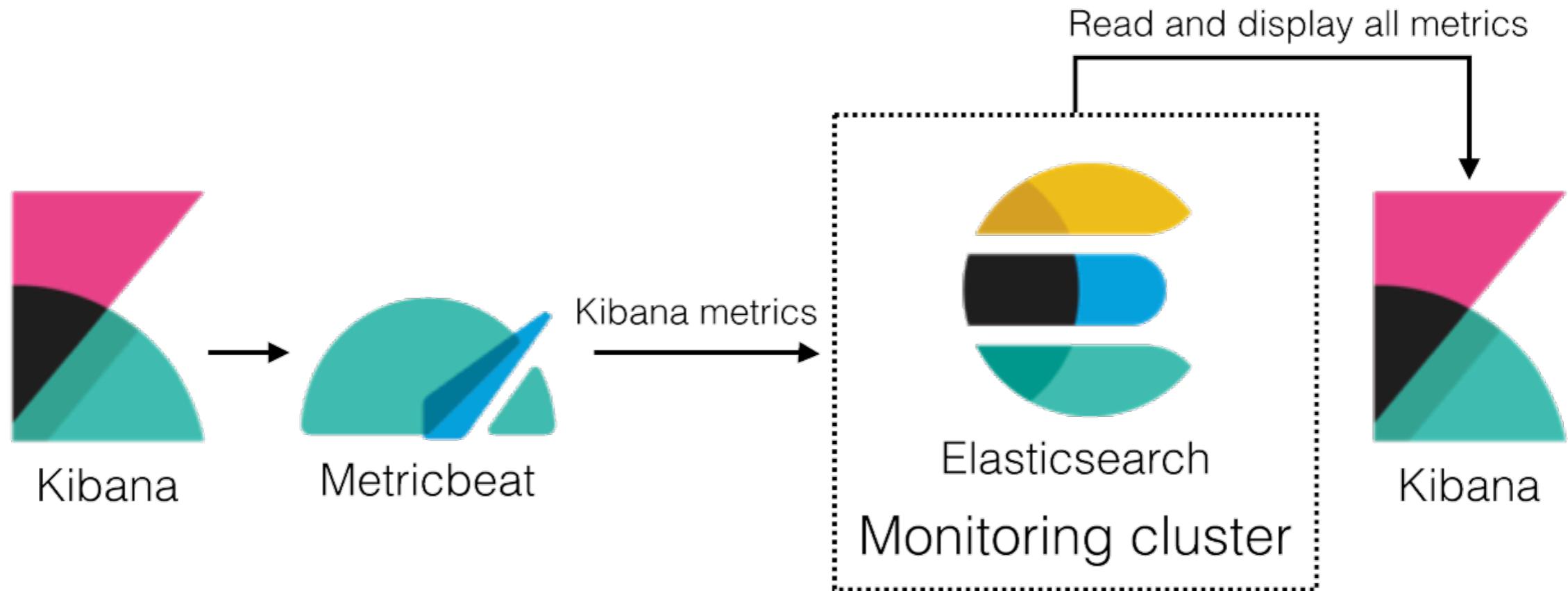
# Recommend architecture



# Elasticsearch 6.4 + (beta)



# Try to separate kibana



# Agenda

- Introduction to Prometheus
- Introduction to Grafana
- Install and configuration
- Working with Prometheus plugins
- Working with Grafana
  - Authentication
  - Datasource
  - Visualization and dashboard
  - Alert system



# Prometheus



# Prometheus



Prometheus

DOCS

DOWNLOAD

COMMUNITY

BLOG



## From metrics to insight

Power your metrics and alerting with a leading  
open-source monitoring solution.

GET STARTED

DOWNLOAD

Prometheus v2.0 is available now — [Read the announcement blog post!](#)

<https://prometheus.io/>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

258

# Prometheus

Metrics monitoring system

Gather **fine grained data** at frequent interval

Make them useful by **label**

Analyse then to understand what is going on



# Prometheus

Easy to config, deploy and maintain

Designed in multiple services

Container ready

Dynamic configuration



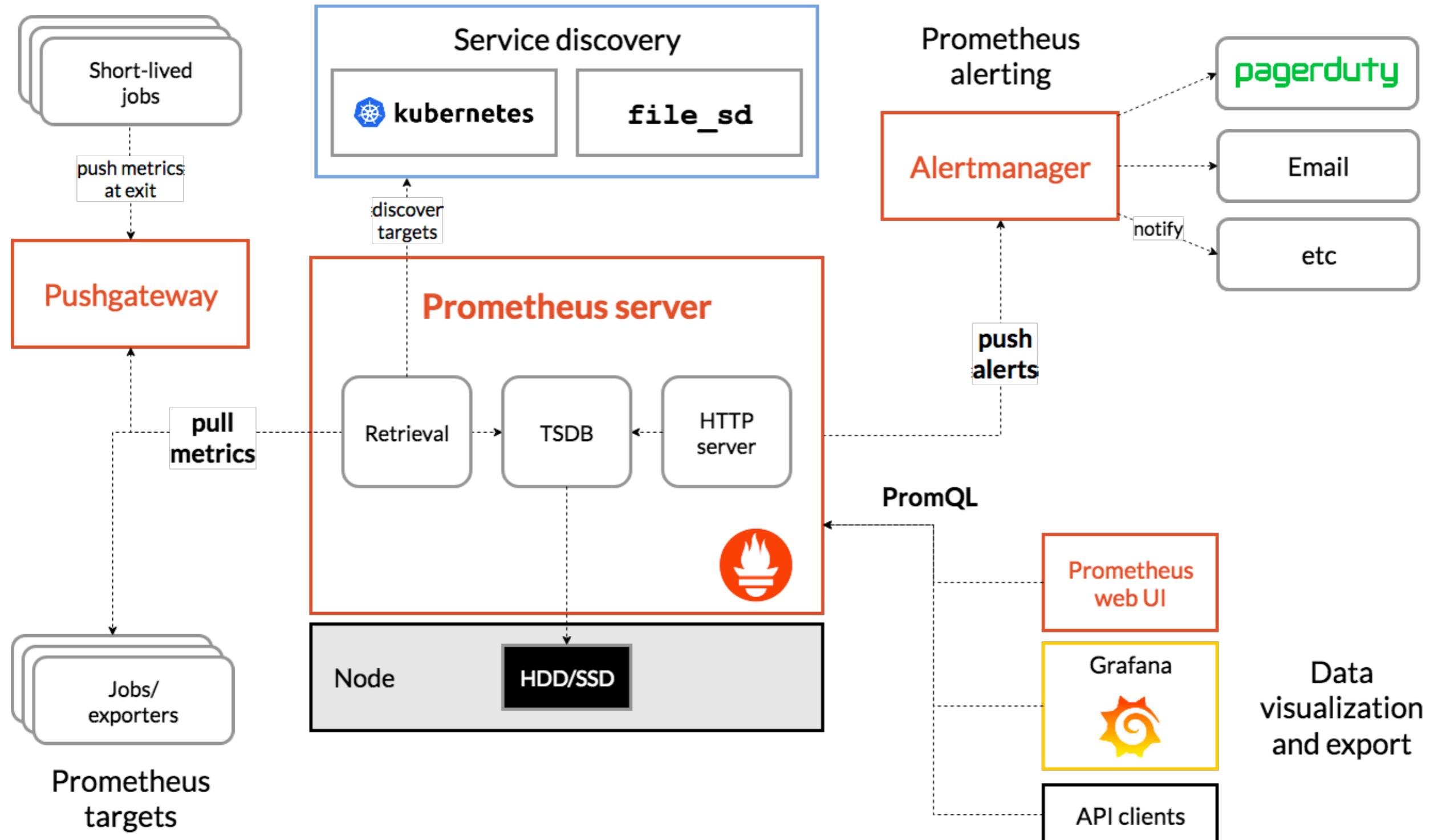
# Prometheus

Open source  
Develop with Go language  
Support for multiple OS  
Time series database  
**Many exporters**  
Alerting

<https://prometheus.io/docs/instrumenting/exporters/>



# What Prometheus do ?



# Prometheus exporters

Expose metrics with an HTTP API  
Binding available for many languages  
Exporters do not save data

<https://prometheus.io/docs/instrumenting/exporters/>



# Prometheus exporters

Node exporter

Blackbox exporter

Mysql exporter

<https://prometheus.io/docs/instrumenting/exporters/>



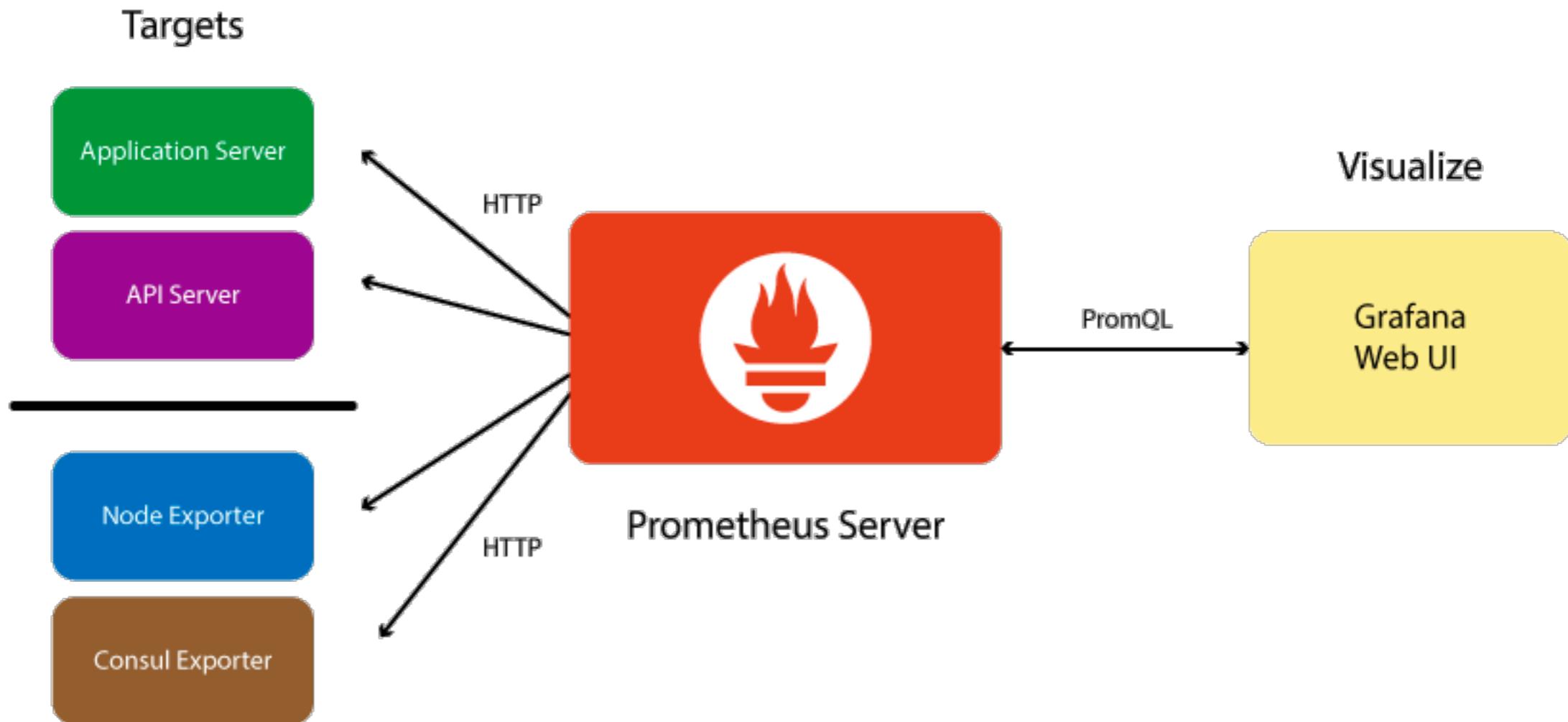
# Alerting

Prometheus has recording rules  
Record frequent queries or alerting  
Alert base on any **PromQL** queries

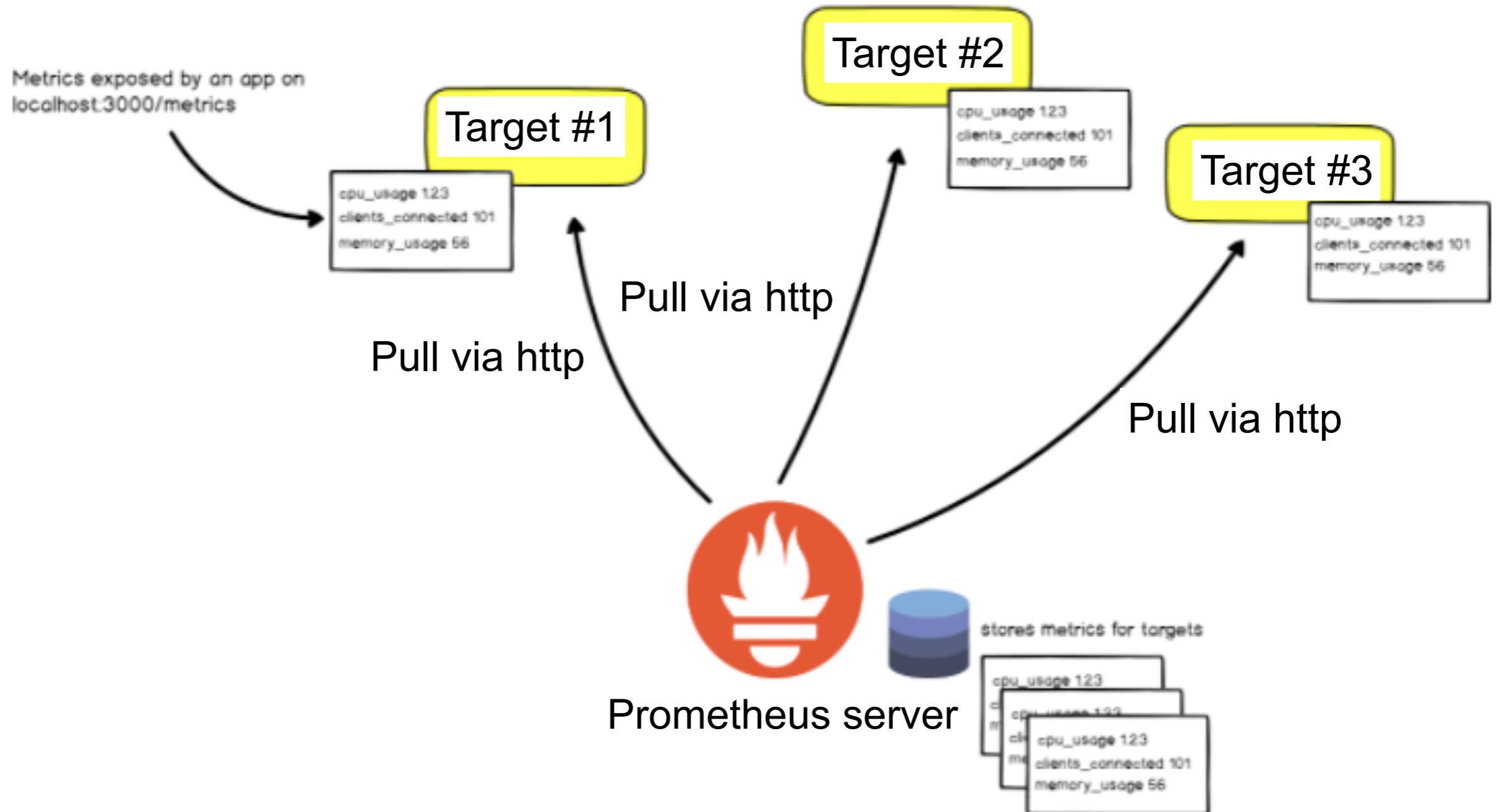
<https://prometheus.io/docs/alerting/latest/overview/>



# What Prometheus do ?



# What Prometheus do ?



# Example



<https://prometheus.io/docs/instrumenting/exporters/>

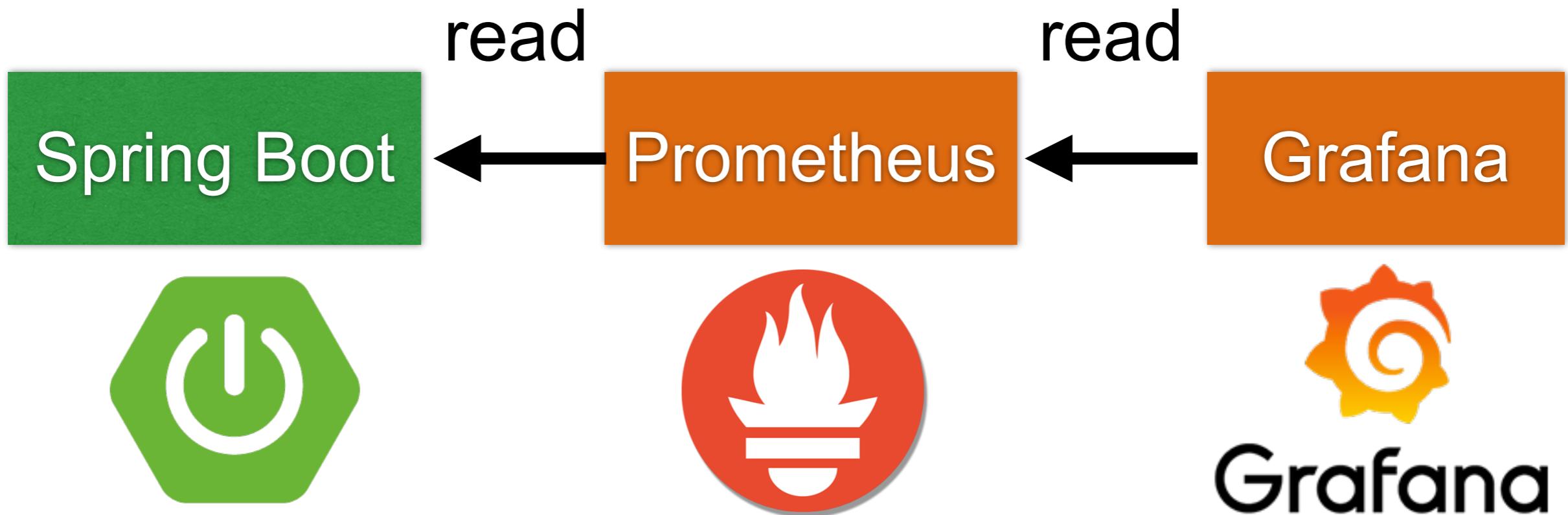


# Workshop

## workshop/spring-prometheus-grafana



# Sample Architecture



# Enable Prometheus

New endpoint = actuator/prometheus

```
← → ⌂ ⓘ localhost:8080/actuator/prometheus

# HELP jvm_memory_used_bytes The amount of used memory
# TYPE jvm_memory_used_bytes gauge
jvm_memory_used_bytes{area="nonheap",id="Code Cache",} 1.49056E7
jvm_memory_used_bytes{area="nonheap",id="Metaspace",} 5.6766712E7
jvm_memory_used_bytes{area="nonheap",id="Compressed Class Space",} 7617096.0
jvm_memory_used_bytes{area="heap",id="PS Eden Space",} 1.7135864E7
jvm_memory_used_bytes{area="heap",id="PS Survivor Space",} 1.6235192E7
jvm_memory_used_bytes{area="heap",id="PS Old Gen",} 2.1936456E7
# HELP hikaricp_connections_idle Idle connections
# TYPE hikaricp_connections_idle gauge
hikaricp_connections_idle{pool="HikariPool-1",} NaN
# HELP tomcat_threads_config_max
# TYPE tomcat_threads_config_max gauge
tomcat_threads_config_max{name="http-nio-8080",} 200.0
# HELP tomcat_servlet_error_total
# TYPE tomcat_servlet_error_total counter
tomcat_servlet_error_total{name="default",} 0.0
# HELP jvm_threads_peak The peak live thread count since the Java virtual machine start
# TYPE jvm_threads_peak gauge
jvm_threads_peak 28.0
# HELP hikaricp_connections_pending Pending threads
# TYPE hikaricp_connections_pending gauge
hikaricp_connections_pending{pool="HikariPool-1",} NaN
# HELP system_cpu_count The number of processors available to the Java virtual machine
```



# Prometheus

PUBLIC | AUTOMATED BUILD

[prom/prometheus](#) 

Last pushed: 17 hours ago

[Repo Info](#) [Tags](#) [Dockerfile](#) [Build Details](#)

## Short Description

Short description is empty for this repo.

## Full Description

Prometheus 

   
   
docker pulls 

Visit [prometheus.io](https://prometheus.io) for the full documentation, examples and guides.

Prometheus is a systems and service monitoring system. It collects metrics

## Docker Pull Command

`docker pull prom/prometheus`

## Owner



prom

## Source Repository

 [prometheus/prometheus](#)

<https://hub.docker.com/r/prom/prometheus/>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

272

# Check Data in Prometheus

http://localhost:9090/

The screenshot shows the Prometheus web interface at the URL `http://localhost:9090/graph`. The interface has a dark header bar with the Prometheus logo and navigation links for Alerts, Graph, Status, and Help. Below the header is a checkbox labeled "Enable query history". A text input field is labeled "Expression (press Shift+Enter for newlines)" with the placeholder "- insert metric at cursor -". There are two tabs: "Graph" (which is selected) and "Console". A table below shows one row with the element "no data" and value "Value". At the bottom left is a blue button labeled "Add Graph".



# Check Target in Prometheus

Status -> Targets

The screenshot shows the Prometheus web interface at the URL `localhost:9090/targets`. The top navigation bar includes links for Prometheus, Alerts, Graph, Status, and Help. The main section is titled "Targets". A checkbox labeled "Only unhealthy jobs" is present. Below this, a summary box displays "spring-boot (1/1 up)" with a "show less" button. A table lists one target endpoint:

Endpoint	State	Labels	Last Scrape	Error
<a href="http://10.10.99.59:8080/actuator/prometheus">http://10.10.99.59:8080/actuator/prometheus</a>	UP	instance="10.10.99.59:8080"	2.355s ago	



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

274

# Show data in Grafana

<https://grafana.com/>



# Grafana

The open platform for beautiful analytics and monitoring

The leading open source software for time series analytics

Get Grafana

APP Grafana TestData By Grafana Project

APP kubernetes By Raintank Inc.

APP Kentik Connect Pro By Kentik

APP NS1 for Grafana By NS1.

Docs Community Events GrafanaCon Blog Log In

Grafana GrafanaCloud Services Dashboards Plugins Get Grafana

Find more plugins on [grafana.com/plugins](#)

**Grafana**

<https://grafana.com/>



ELK Stack

© 2017 - 2018 Siam Chamnkit Company Limited. All rights reserved.

# Grafana

Open source  
Specialized in visualization  
Plugable  
Multiple datasources  
Has an API/ Web app



# Grafana

Fork from Kibana 3  
Develop by JavaScript



# Grafana

PUBLIC REPOSITORY

[grafana/grafana](#) 

Last pushed: 25 minutes ago

[Repo Info](#) [Tags](#)

## Short Description

The official Grafana docker container

## Full Description

### Grafana Docker image

This project builds a Docker image with the latest master build of Grafana.

## Running your Grafana container

Start your container binding the external port 3000 .

```
docker run -d --name=grafana -p 3000:3000 grafana/grafana
```

## Docker Pull Command

```
docker pull grafana/grafana
```

## Owner



grafana

<https://hub.docker.com/r/grafana/grafana/>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.

# Grafana

<https://grafana.com/dashboards/4701>

All dashboards » [JVM \(Micrometer\)](#)



**JVM (Micrometer)** by [mweirauch](#)

[DASHBOARD](#)

Dashboard for Micrometer instrumented applications (Java, Spring Boot)  
Last updated: 21 days ago

[Overview](#) [Revisions](#)



A dashboard for Micrometer instrumented applications (Java, Spring Boot).

**Features**

- JVM memory
- Process memory (provided by [micrometer-jvm-extras](#))
- CPU-Usage, Load, Threads, File Descriptors, Log Events
- JVM Memory Pools (Heap, Non-Heap)
- Garbage Collection

**Get this dashboard:**

[4701](#) [Copy ID to Clipboard](#)

[Download JSON](#) [How do I import this dashboard?](#)

**Dependencies:**

 [GRAFANA 4.6.3](#)

 [GRAPH](#)



# Workshop

<https://grafana.com/>



ELK Stack

© 2017 - 2018 Siam Chamnankit Company Limited. All rights reserved.