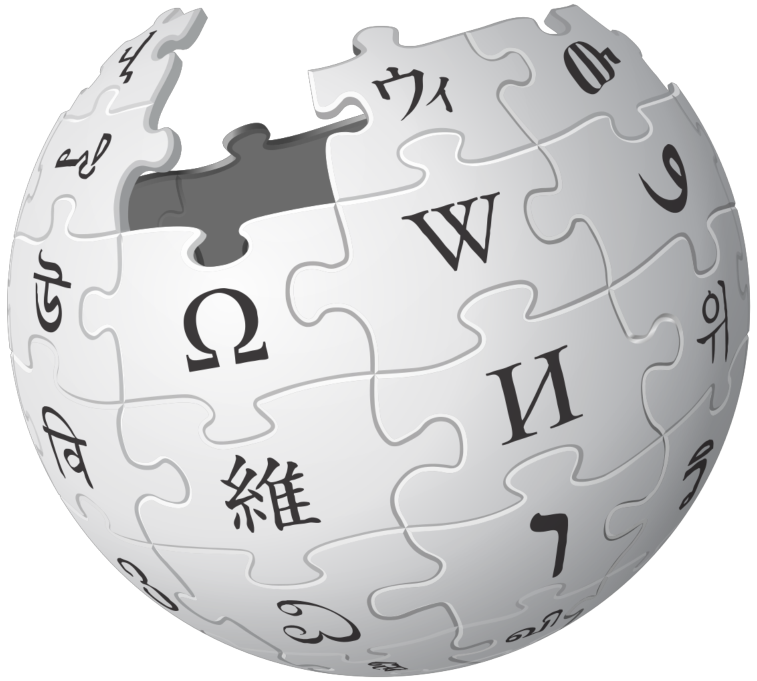
Scale Your Metrics with Elasticsearch

Philipp Krenn @xeraa

elastic

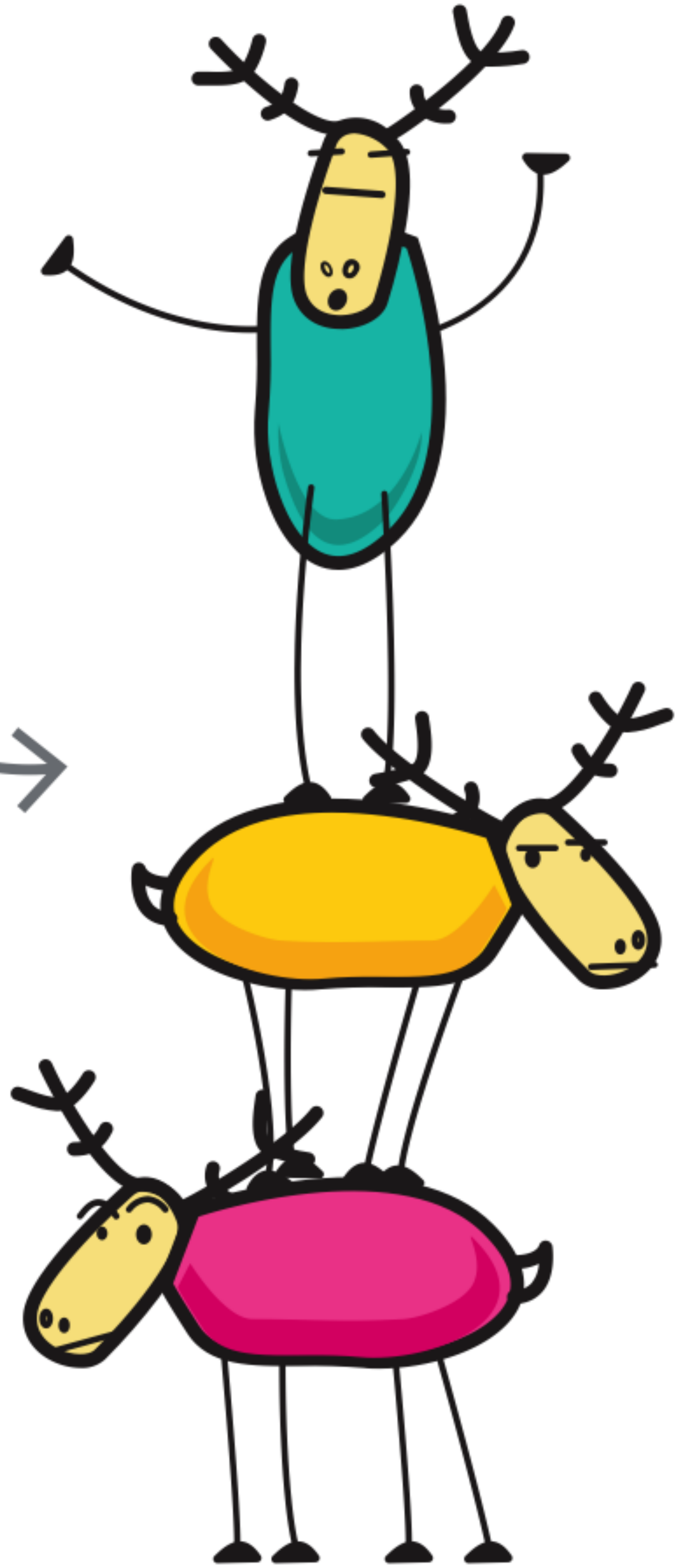
elasticsearch.

```
$ curl http://localhost:9200
{
  "name" : "elasticsearch-hot",
  "cluster_name" : "metrics-cluster",
  "cluster_uuid" : "O6nHPLLgTrmZEpYli6JW5w",
  "version" : {
    "number" : "6.5.0",
    "build_flavor" : "default",
    "build_type" : "tar",
    "build_hash" : "c53b7d3",
    "build_date" : "2018-11-08T21:28:50.577384Z",
    "build_snapshot" : false,
    "lucene_version" : "7.5.0",
    "minimum_wire_compatibility_version" : "5.6.0",
    "minimum_index_compatibility_version" : "5.0.0"
  },
  "tagline" : "You Know, for Search"
}
```


ウィ
W
Ω
И
維

kibana

logstash



ELK Stack!
Get it?

E Elasticsearch

L Logstash

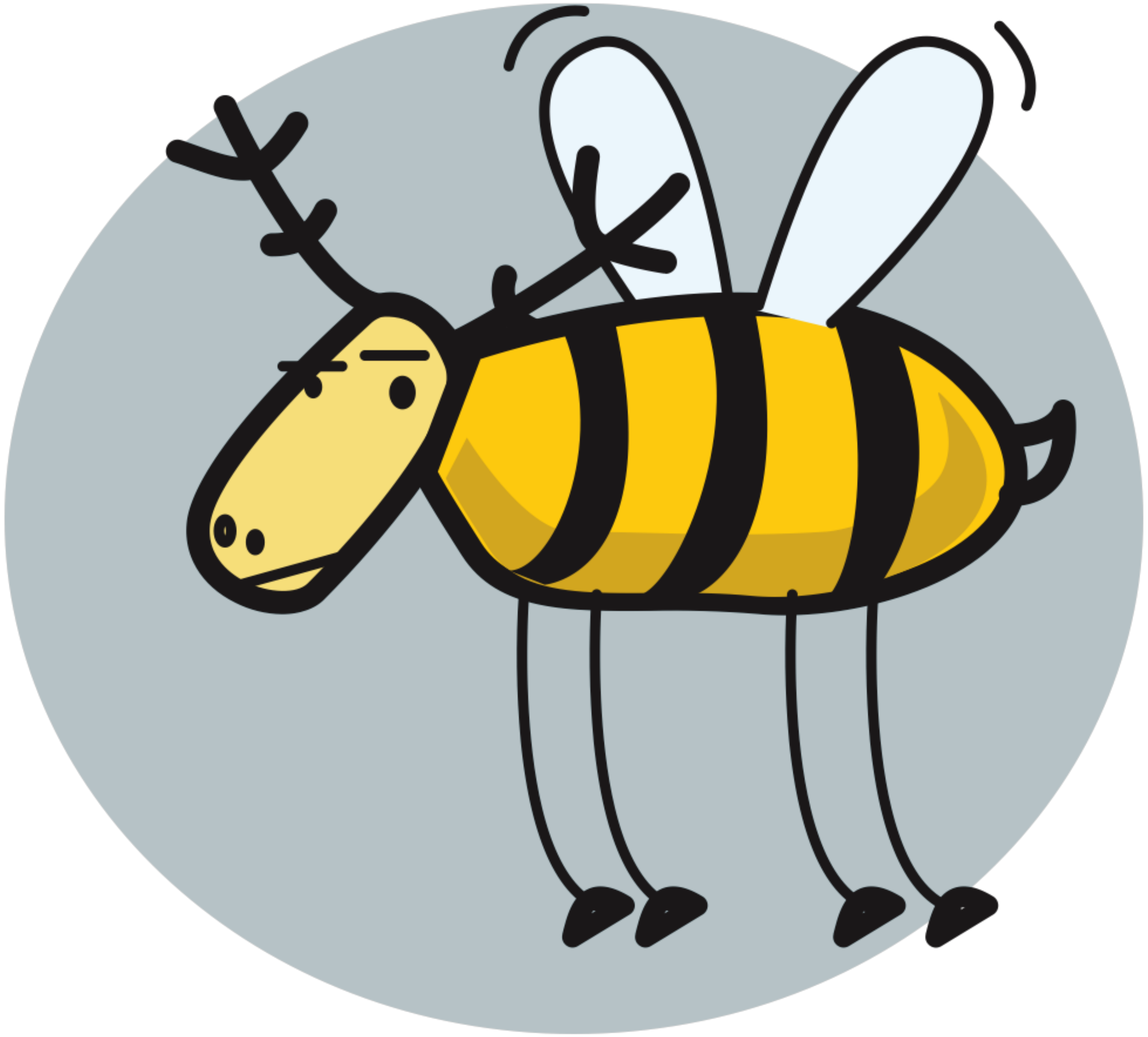
K Kibana

lyft

slack

fitbit

Apparently, I'm an ELKB personality.
Meyers Briggs
Meyers Briggs



elastic

elastic stack

BLIZZARD®
ENTERTAINMENT

CAR
2GO

ebay

> *I'm not going to use a search engine for metrics.*
>
> **— Too often**

elastic

Developer 🥑

Questions: https://sli.do/xeraa

Answers: Live or https://twitter.com/xeraa

Agenda

Building Blocks

Architecture

Demo

Building Blocks

Only accept features that scale.

— **https://github.com/elastic/engineering/blob/master/development_constitution.md**

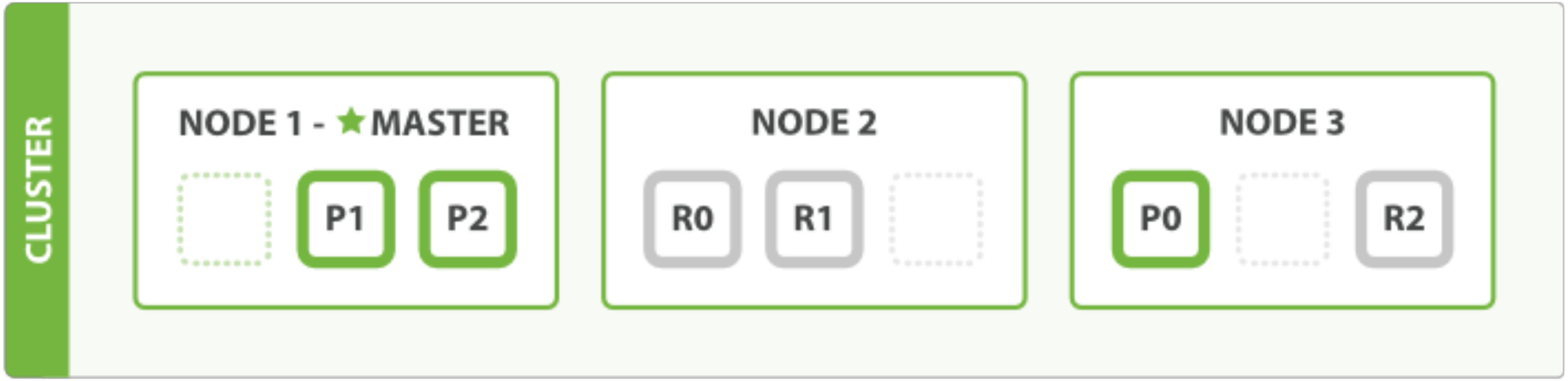
Horizontal Scaling

Shards

Replication

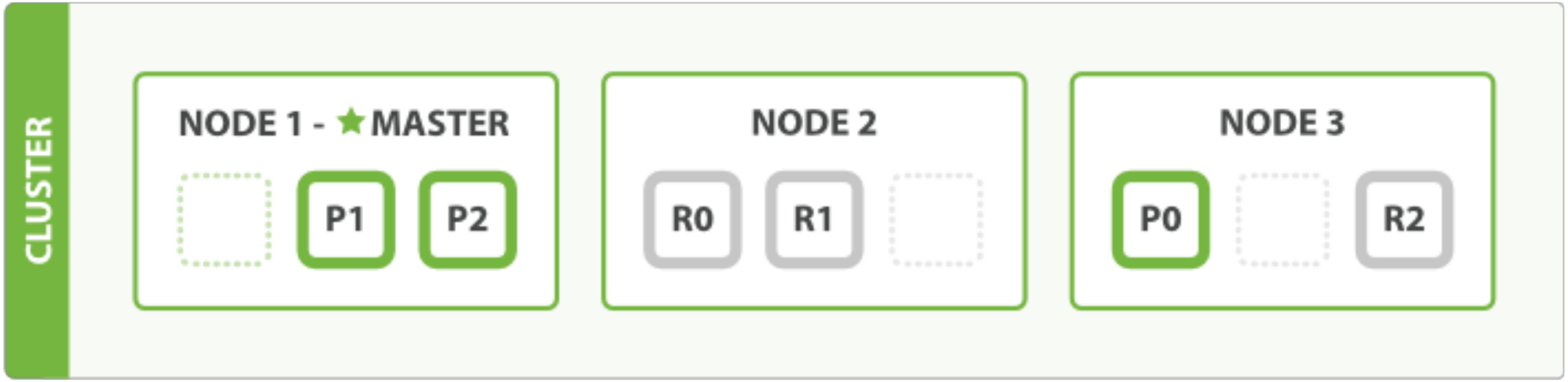
Writes & Reads

Cluster, Node, Index, Shard



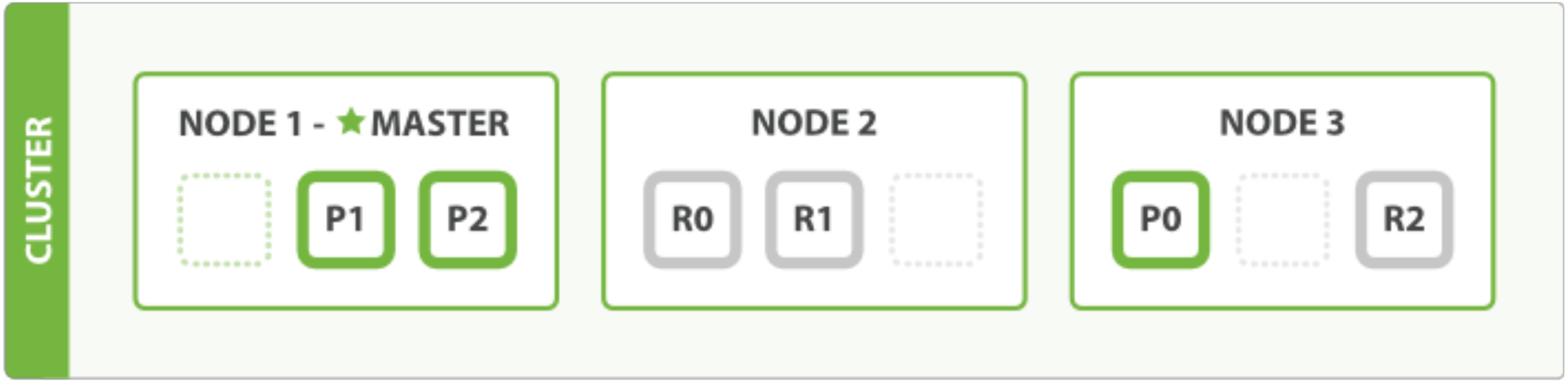
Write

Coordinating Node, Hash, Primary, Replica(s)



Get & Aggregate

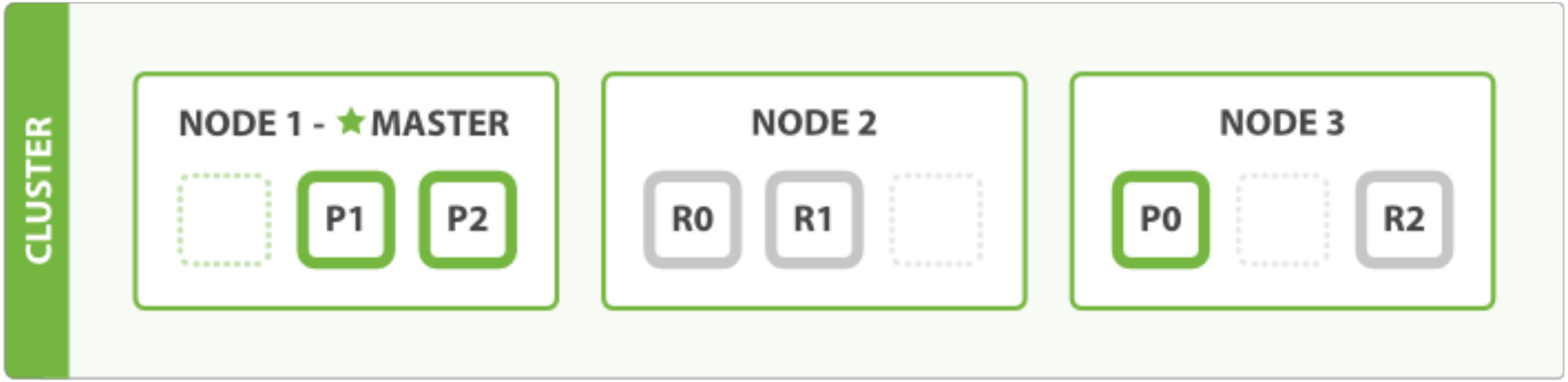
Coordinating Node, Hash, Shard



elastic

Search

Coordinating Node, Query then Fetch



Append-Only Optimization

IDs assigned on coordinating node
Fast add instead of the slow update

Lucene Segments

`index.refresh_interval: 1s`

7.0: `index.search.idle.after`

Storage Compression

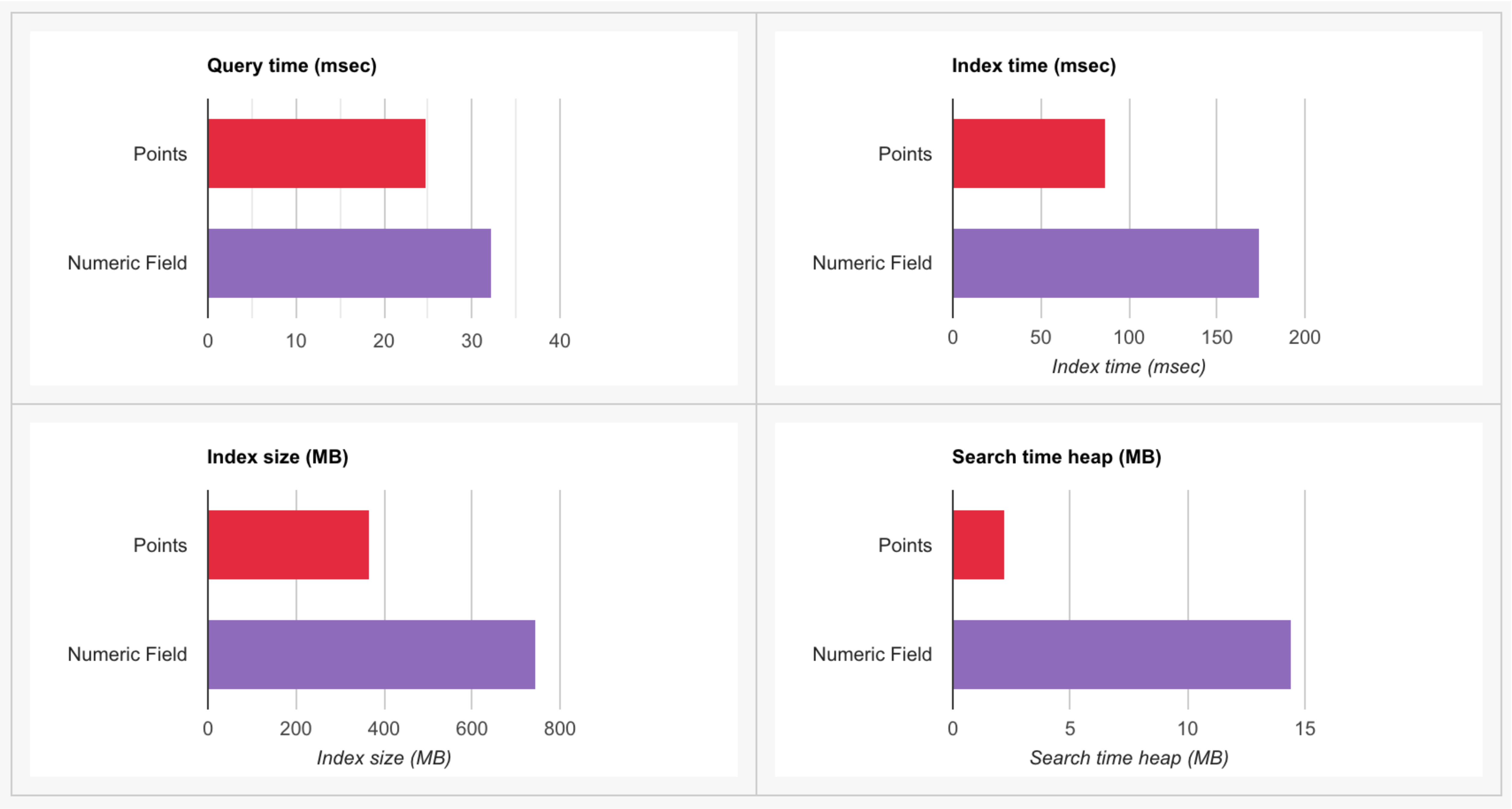
LZ4 (default)

DEFLATE (`best_compression`)

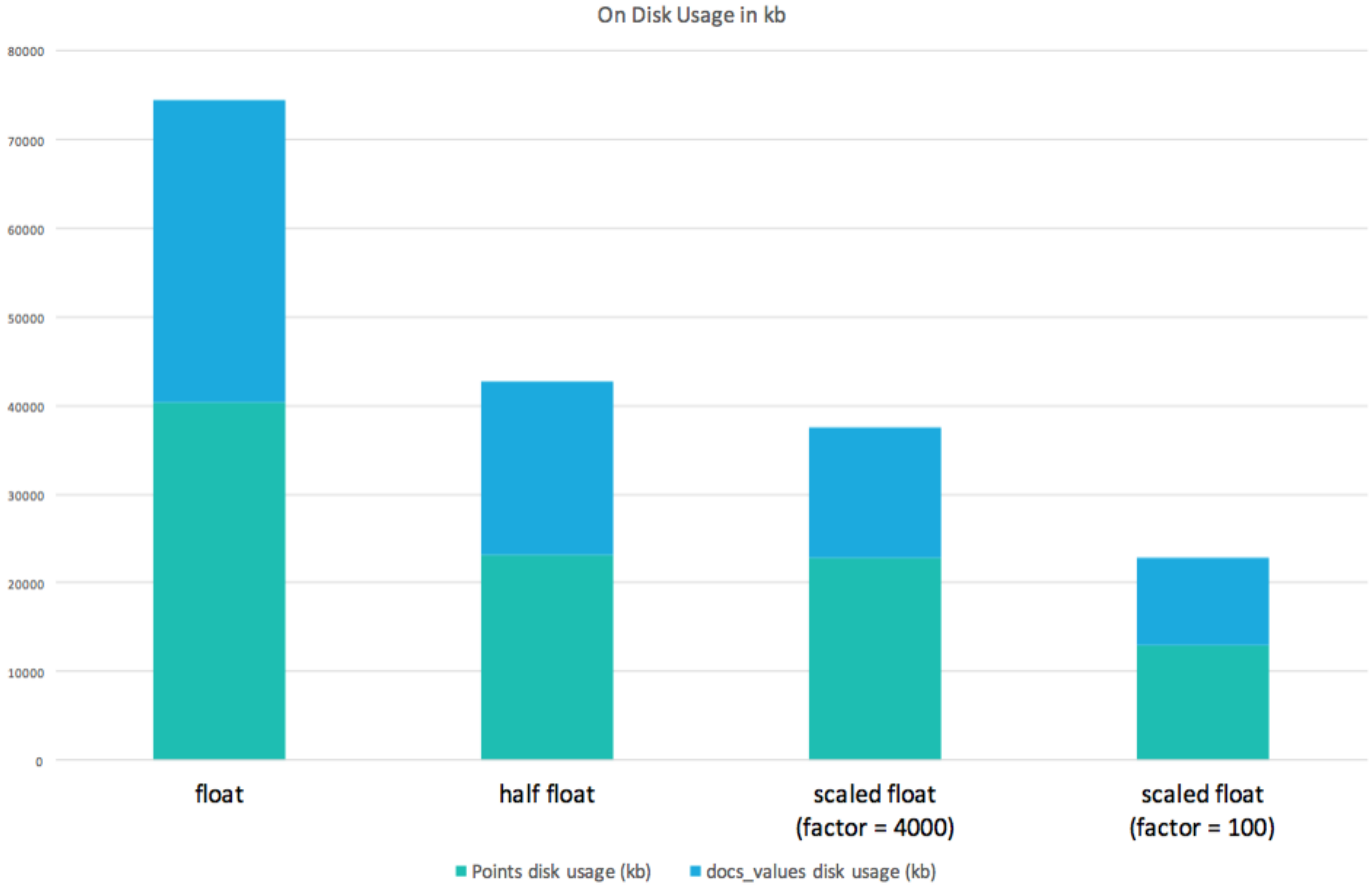
BKD Trees

Points in Lucene

Integer (1D 4 byte point) vs legacy IntField



Half & Scaled Floats



On Disk Usage in kb
80000
70000
60000
50000
40000
30000
20000
10000
0
float
half float
scaled float (factor = 4000)
scaled float (factor = 100)
Points disk usage (kb)
docs_values disk usage (kb)

https://github.com/elastic/beats/blob/master/metricbeat/module/system/load/_meta/fields.yml

```
- name: load
  type: group
  description: >
    CPU load averages.
  release: ga
  fields:
    - name: "1"
      type: scaled_float
      scaling_factor: 100
      description: >
        Load average for the last minute.
    - name: "5"
      type: scaled_float
      scaling_factor: 100
      description: >
        Load average for the last 5 minutes.
    ...
```

`_all` Removal

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

Doc Values

Replaced Fielddata

https://www.elastic.co/guide/en/elasticsearch/guide/current/_deep_dive_on_doc_values.html

Architecture

Time Based Indices

```
index: "metricbeat-%{[beat.version]}-%{+yyyy.MM.dd}"
```

Rollover Indices

Condition when to switch

```
PUT /metricbeat-000001
{
  "aliases": {
    "metricbeat": {}
  }
}

# Add >1000 documents to metricbeat-000001

POST /metricbeat/_rollover
{
  "conditions": {
    "max_age":   "1d",
    "max_docs":  1000,
    "max_size":  "5gb"
  }
}
```

```
{
  "acknowledged": true,
  "shards_acknowledged": true,
  "old_index": "metricbeat-000001",
  "new_index": "metricbeat-000002",
  "rolled_over": true,
  "dry_run": false,
  "conditions": {
    "[max_age: 1d]": false,
    "[max_docs: 1000]": true,
    "[max_size: 5gb]": false,
  }
}
```

Rollups

5MB
25MB
1TB

think of the bytes

```
PUT _xpack/rollup/job/metricbeat
{
  "id": "metricbeat",
  "index_pattern": "metricbeat-*",
  "rollup_index": "metricbeat_rollup",
  "cron": "0 * * * ?",
  "page_size": 1000,
  "groups": {
    "date_histogram": {
      "interval": "5m",
      "delay": "5m",
      "time_zone": "UTC",
      "field": "@timestamp"
    },
    "terms": {
      "fields": [
        "docker.container.id"
      ]
    }
  },
  "metrics": [
    {
      "field": "docker.network.in.bytes",
      "metrics": [
        "sum"
      ]
    },
    {
      "field": "docker.network.out.bytes",
      "metrics": [
        "sum"
      ]
    }
  ]
}
```

Nodes

🔥 🌡️ ❄️

```
$ bin/elasticsearch
  -Enode.attr.rack=rack1
  -Enode.attr.size=hot
```

```
PUT /metricbeat/_settings
{
  "index.routing.allocation.include.size": "hot"
}
```

Cross Cluster Search

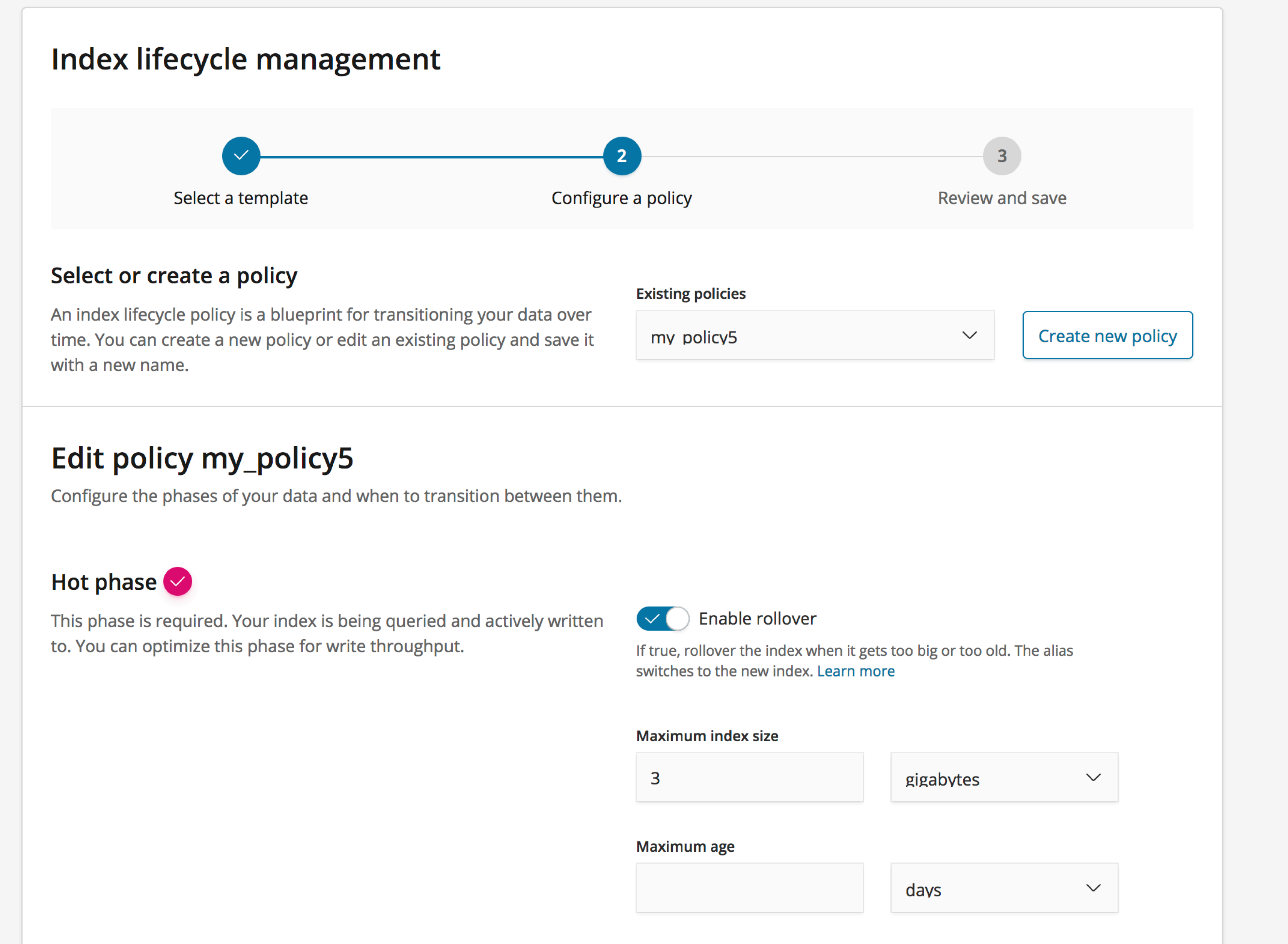
~~Tribe Node~~

Cross Cluster Replication

Demo

Index Lifecycle Management

Currently
https://github.com/elastic/curator



Index lifecycle management

✓ Select a template — 2 Configure a policy — 3 Review and save

Select or create a policy

An index lifecycle policy is a blueprint for transitioning your data over time. You can create a new policy or edit an existing policy and save it with a new name.

Existing policies: my_policy5

Edit policy my_policy5

Configure the phases of your data and when to transition between them.

Hot phase ✓

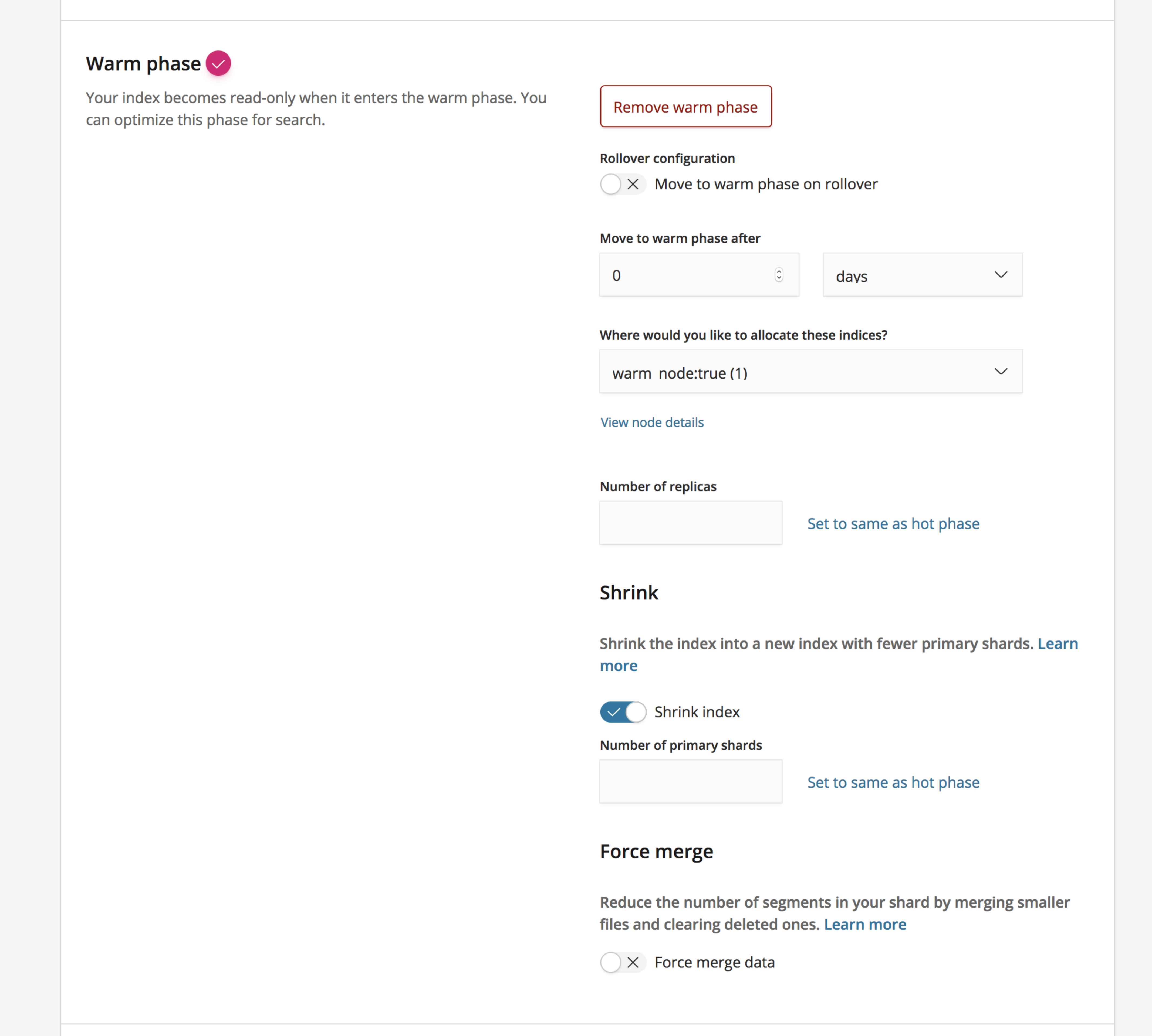
This phase is required. Your index is being queried and actively written to. You can optimize this phase for write throughput.

☑ Enable rollover

If true, rollover the index when it gets too big or too old. The alias switches to the new index. Learn more

Maximum index size: 3 gigabytes

Maximum age: days



Warm phase ✓

Your index becomes read-only when it enters the warm phase. You can optimize this phase for search.

Remove warm phase

Rollover configuration

○ ✕ Move to warm phase on rollover

Move to warm phase after

0 | days

Where would you like to allocate these indices?

warm node:true (1)

View node details

Number of replicas

Set to same as hot phase

Shrink

Shrink the index into a new index with fewer primary shards. Learn more

✓ ● Shrink index

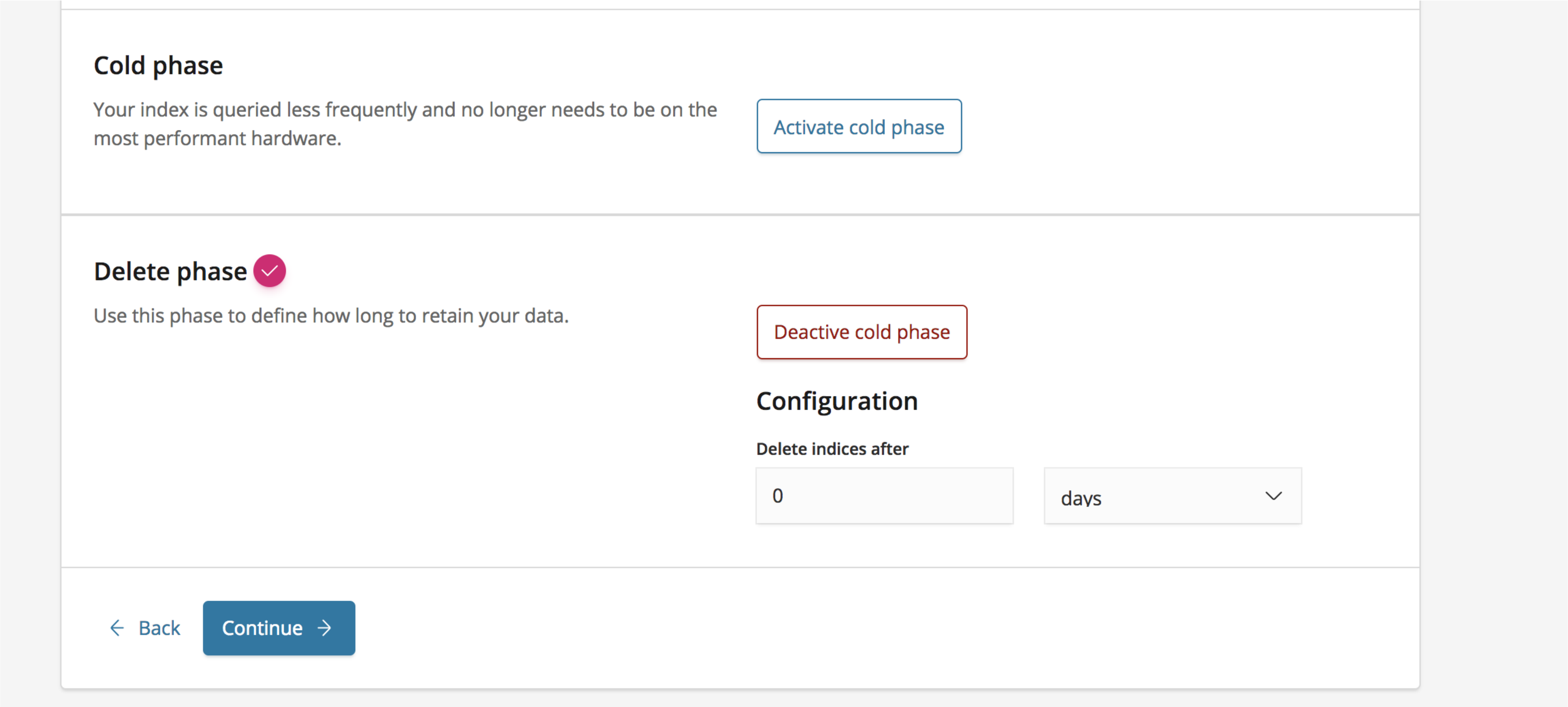
Number of primary shards

Set to same as hot phase

Force merge

Reduce the number of segments in your shard by merging smaller files and clearing deleted ones. Learn more

○ ✕ Force merge data



Cold phase

Your index is queried less frequently and no longer needs to be on the most performant hardware.

Delete phase ✓

Use this phase to define how long to retain your data.

Configuration

Delete indices after

0 days

Frozen Indices

Close + lazy open and release search resources

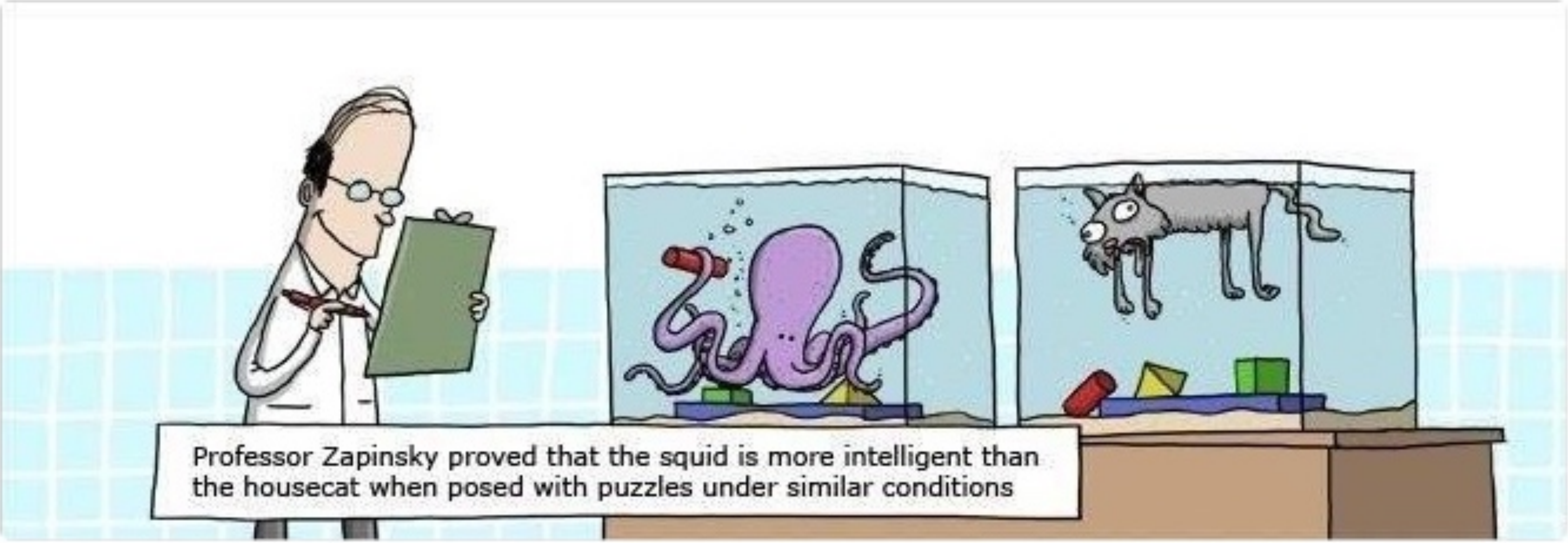
Conclusion

Agenda

Building Blocks
Architecture
Demo

Benchmarks

Fair
Reproducible
Close to Production


Professor Zapinsky proved that the squid is more intelligent than the housecat when posed with puzzles under similar conditions

From 🏝️ to 🗺️

Questions?

Philipp Krenn @xeraa

elastic