

Java under the Hood of



Elasticsearch

Philipp Krenn

@xeraa

Free and Open, Distributed, RESTful Search Engine

www.elastic.co/products/elasticsearch

[View license](#)

61.3k stars 22.3k forks

★ Starred

👁 Watch ▼

<> Code

🗨 Issues 3.5k

🔗 Pull requests 502

▶ Actions

📁 Projects 1

⋮

🔗 main ▼

⋮

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

philipp@~/Documents/GitHub/elasticsearch(git:main) ✓> cloc .

22,159 text files.

21,026 unique files.

1,787 files ignored.

Language	files	blank	comment	code
Java	16,061	365,056	316,642	2,351,745
AsciiDoc	1,646	49,416	13,574	227,792
YAML	1,510	19,484	3,938	173,217
JSON	818	27	0	60,644
Text	260	4,364	0	26,083
Gradle	365	3,041	2,632	18,594
XML	72	577	1,145	8,813
Properties	42	179	434	8,712
Groovy	58	1,206	814	7,178
TOML	10	670	965	3,322
CSV	11	0	0	2,840
Markdown	21	681	0	1,946
XSD	6	112	40	1,499
SVG	25	0	18	1,121
ANTLR Grammar	5	207	87	1,059
Bourne Shell	20	204	347	900
...
SUM:	21026	445741	341758	2898545

```
philipp@~/Documents/GitHub/elasticsearch(git:main) ✓> mergestat summarize commits
```

```
Commits          65,763
Non-Merge Commits 61,279
Files Δ          91,842
Unique Authors   2,065
First Commit     13 years ago (2010-02-08)
```

Author	Commits	Commit %	Files Δ	Additions	Deletions	First Commit	Latest Commit
Martijn van Groningen	2,979	4.53%	22,957	1,465,778	689,734	11 years ago (2012-07-30)	5 days ago (2022-09-14)
Jason Tedor	2,858	4.35%	21,334	1,814,307	1,437,611	8 years ago (2015-07-07)	a year ago (2021-10-20)
Simon Willnauer	2,811	4.27%	16,080	1,124,143	1,018,974	11 years ago (2012-06-10)	4 years ago (2019-08-29)
kimchy	2,478	3.77%	5,316	604,412	289,808	13 years ago (2010-02-08)	12 years ago (2011-07-23)
Nik Everett	2,390	3.63%	17,973	896,438	503,010	10 years ago (2013-07-18)	6 days ago (2022-09-13)
Shay Banon	2,303	3.50%	6,968	554,647	526,360	12 years ago (2011-07-24)	8 years ago (2015-07-24)
Ryan Ernst	2,178	3.31%	29,641	2,309,483	1,761,652	9 years ago (2014-07-15)	2 months ago (2022-07-29)
Armin Braun	1,488	2.26%	4,930	105,752	117,232	6 years ago (2017-09-19)	4 days ago (2022-09-15)
David Pilato	1,435	2.18%	9,931	1,092,448	695,985	12 years ago (2011-10-20)	10 months ago (2021-11-23)
javanna	1,380	2.10%	14,474	554,526	507,303	9 years ago (2014-02-19)	4 years ago (2018-11-06)
Adrien Grand	1,364	2.07%	7,468	212,344	270,444	10 years ago (2013-05-21)	4 months ago (2022-06-14)
Boaz Leskes	1,126	1.71%	8,753	461,364	313,507	10 years ago (2013-05-22)	4 years ago (2019-03-27)
Clinton Gormley	1,090	1.66%	1,611	264,508	35,561	12 years ago (2011-04-26)	4 years ago (2019-02-18)
Tanguy Leroux	1,057	1.61%	9,041	285,194	142,622	8 years ago (2015-02-13)	13 days ago (2022-09-06)
David Turner	1,052	1.60%	7,832	389,846	155,187	6 years ago (2017-09-29)	4 days ago (2022-09-15)
...2,050 more authors							

```
$ curl -u elastic "https://localhost:9200"
{
  "name": "instance-0000000000",
  "cluster_name": "bfc4c4e741374efebbde66e1b3ff3a5a",
  "cluster_uuid": "0ZAX3Rm9T8iWV48uH9jSfA",
  "version": {
    "number": "8.5.0",
    "build_flavor": "default",
    "build_type": "docker",
    "build_hash": "6a0e30826aab16fb90341d71815a06ee71ea32de",
    "build_date": "2022-10-05T13:16:46.105697005Z",
    "build_snapshot": false,
    "lucene_version": "9.4.0",
    "minimum_wire_compatibility_version": "7.17.0",
    "minimum_index_compatibility_version": "7.0.0"
  },
  "tagline": "You Know, for Search"
}
```



Developer 🥑

Bundled JDK

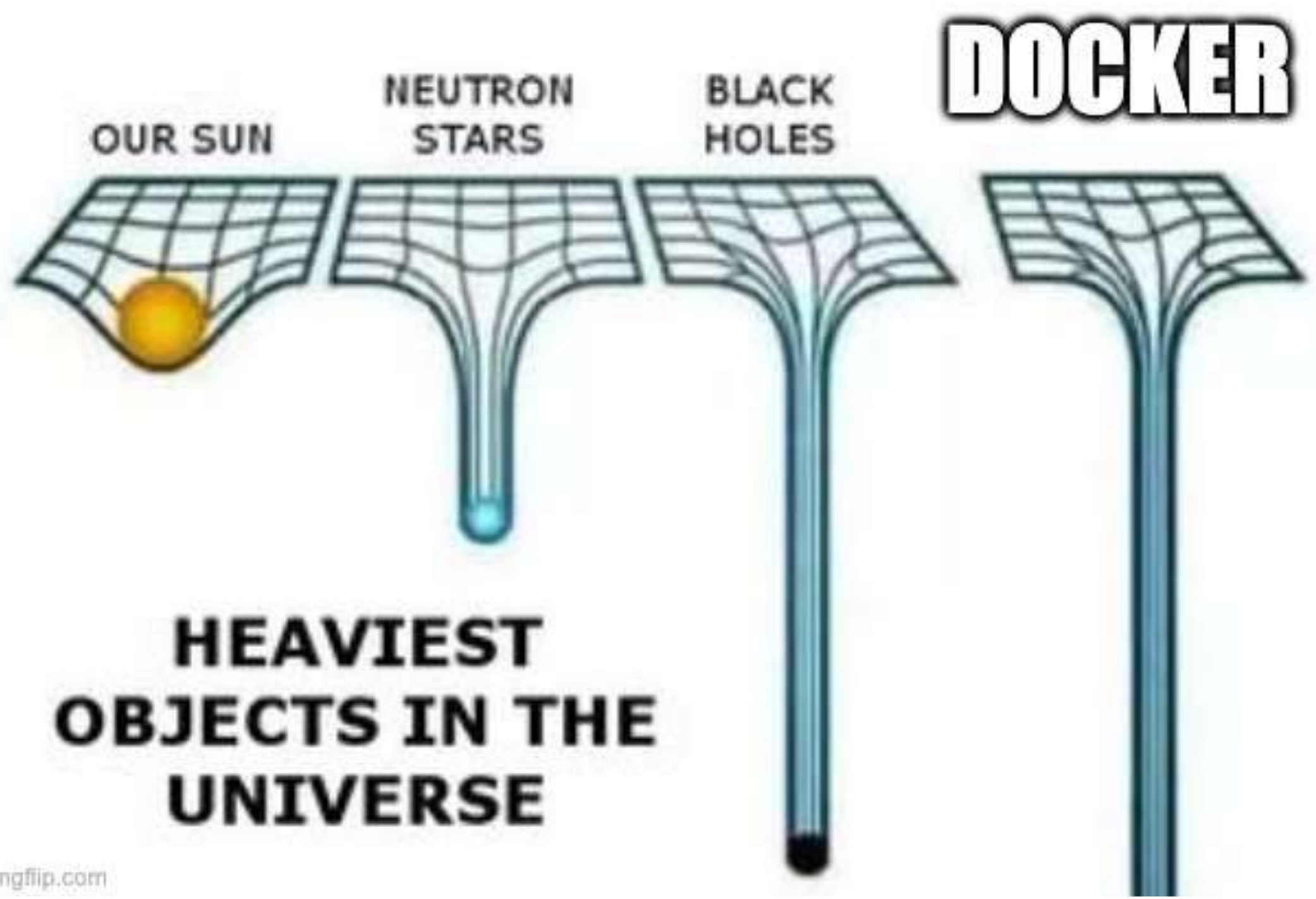
**OpenJDK (GPLv2+CE) bundled in
all binaries**

**Override with ES_JAVA_HOME,
optional delete folder**

Switch back to Oracle OpenJDK (from Adoptium)

<https://github.com/elastic/elasticsearch/pull/86554>

Concerns: security > size



Always latest JDK

Oracle Java SE Support Roadmap ^{*†}		
Release	GA Date	Premier Support Until
7 (LTS)	July 2011	July 2019
8 (LTS)**	March 2014	March 2022
9 (non-LTS)	September 2017	March 2018
10 (non-LTS)	March 2018	September 2018
11 (LTS)	September 2018	September 2023
12 (non-LTS)	March 2019	September 2019
13 (non-LTS)	September 2019	March 2020
14 (non-LTS)	March 2020	September 2020
15 (non-LTS)	September 2020	March 2021
16 (non-LTS)	March 2021	September 2021
17 (LTS)	September 2021	September 2026****
18 (non-LTS)	March 2022	September 2022
19 (non-LTS)***	September 2022	March 2023
20 (non-LTS)***	March 2023	September 2023
21 (LTS)***	September 2023	September 2028

<https://www.oracle.com/java/technologies/java-se-support-roadmap.html>

Blocker: "Upgrade bundled JDK to Java 19"

<https://github.com/elastic/elasticsearch/pull/90571>

Java Features

Elasticsearch baseline

JDK 8 for 7
JDK 17 for 8

Precondition for JDK 17

Split up server and client

Supported versions – it's complicated

https://www.elastic.co/support/matrix#matrix_jvm

Test matrix

Jenkins

Ship tested JVM and settings

-Xmx2g vs 17 : -Xmx2g vs 17-18 : -Xmx2g vs 17- : -Xmx2g

"Java 17 features in Elasticsearch code base"

<https://github.com/elastic/elasticsearch/issues/81820>

- ✓ **Pattern matching for `instanceof`**
- ✓ **Text blocks for JSON and multiline strings**
- ✓ **Improved switch expressions**

🕒 **Pattern matching for switch**

🕒 **Sealed classes**

🕒 **Records**

Lucene 9.4.0

**"Support for Java 19 foreign memory access"
(Project Panama)**

**"After update to Lucene 9.4 use
--enable-preview on Java==19
(exact) to allow mmap use new
JDK APIs"**

<https://github.com/elastic/elasticsearch/issues/90526>

Result

**[https://elasticsearch-ci.elastic.co/view/
Elasticsearch%20main/job/
elastic+elasticsearch+main+periodic+java-preview-
features/](https://elasticsearch-ci.elastic.co/view/Elasticsearch%20main/job/elastic+elasticsearch+main+periodic+java-preview-features/)**

Garbage Collectors

Throughput vs latency

Concurrent Mark Sweep (CMS)

Long-time recommendation, removed in JDK 14

G1GC

Supported since 6.5, default since 7.6

<https://github.com/elastic/elasticsearch/pull/49123>



Ismael Juma

@ijuma



"When running an Apache Cassandra benchmark, Generational ZGC manages to run with a fourth of the heap size and still gets four times the throughput compared to non-generational ZGC (while still not having pause times longer than 1 millisecond)"

OpenJDK **OpenJDK** @OpenJDK · Oct 14

RFC: Draft JEP: Generational ZGC

For details, see the JBS entry:

bugs.openjdk.org/browse/JDK-827...

more @ mail.openjdk.org/pipermail/hots...

#OpenJDK #ZGC

9:23 PM · Oct 14, 2022 · Twitter for Android

ZGC not yet in Elasticsearch

<https://github.com/elastic/elasticsearch/issues/58989>

Only IntelliJ 2020.1+ supported

What about Kotlin?

Java & Security





Anonymous



3



0



24 Jun, 12:02pm



How much Java Security Manager actually helps you and how often it's a pain in the a\$\$?

security.policy

```
// Allow host/ip name service lookups
permission java.net.SocketPermission "*", "resolve";

// Allow reading and setting socket keepalive options
permission jdk.net.NetworkPermission "getOption.TCP_KEEPIDLE";
permission jdk.net.NetworkPermission "setOption.TCP_KEEPIDLE";
permission jdk.net.NetworkPermission "getOption.TCP_KEEPINTERVAL";
permission jdk.net.NetworkPermission "setOption.TCP_KEEPINTERVAL";
permission jdk.net.NetworkPermission "getOption.TCP_KEEPCOUNT";
permission jdk.net.NetworkPermission "setOption.TCP_KEEPCOUNT";
```

<https://github.com/elastic/elasticsearch/blob/7.16/server/src/main/resources/org/elasticsearch/bootstrap/security.policy>

Java Security Manager

Few exceptions for <https://github.com/elastic/elasticsearch/search?q=SocketPermission> like Netty

Java Security Manager

Deprecated in JDK17

<https://openjdk.java.net/jeps/411>

For now

"-Djava.security.manager=allow"

<https://github.com/elastic/elasticsearch/pull/80957>

JSM replacement

Modularization

<https://github.com/elastic/elasticsearch/labels/modularization>

Gradle

Don't try this at home

365 files, 19K LoC

Tests and migrations

Simple

```
./gradlew run
```

Options

```
./gradlew run --debug-jvm
```

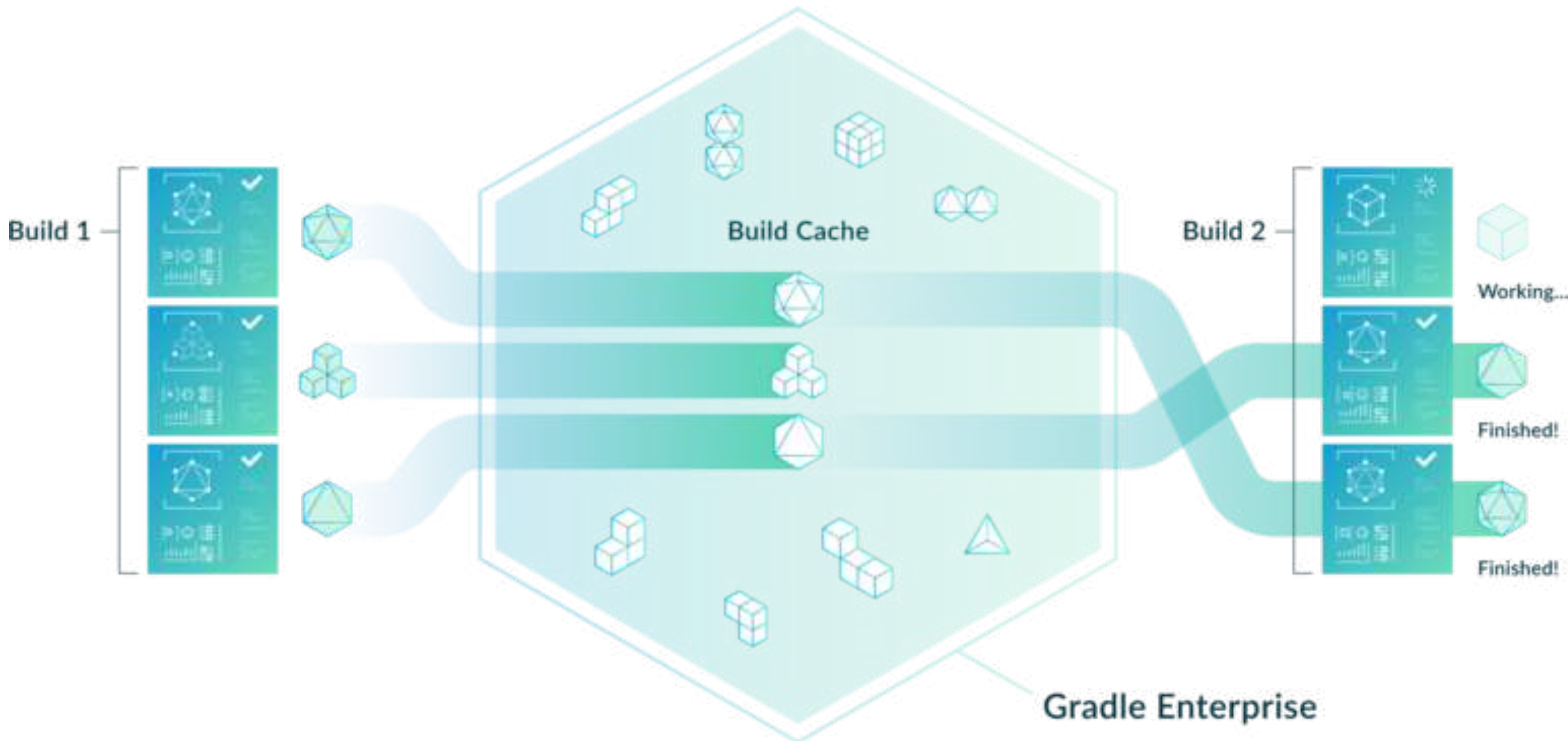
```
./gradlew :server:test --tests  
org.elasticsearch.package.ClassName
```

```
./gradlew test -Dtests.filter=@awaitsfix
```

```
./gradlew test -Dtests.seed=DEADBEEF
```

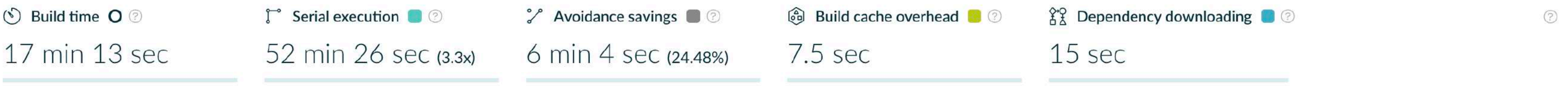
Complex

```
testClusters.register("runTask") {  
  ...  
  numberOfNodes = 2  
  def cluster = testClusters.named("runTask").get()  
  cluster.getNodes().each { node ->  
    node.setting('cluster.initial_master_nodes', cluster.getLastNode().getName())  
    node.setting('node.roles', '[master,data_hot,data_content]')  
  }  
  cluster.getFirstNode().setting('node.roles', '[]')  
  ...  
}
```



User Hostname Project elasticsearch Requested tasks/goals Build tool Build tool version

Start time Last 7 days Custom values Tags Outcome Refresh

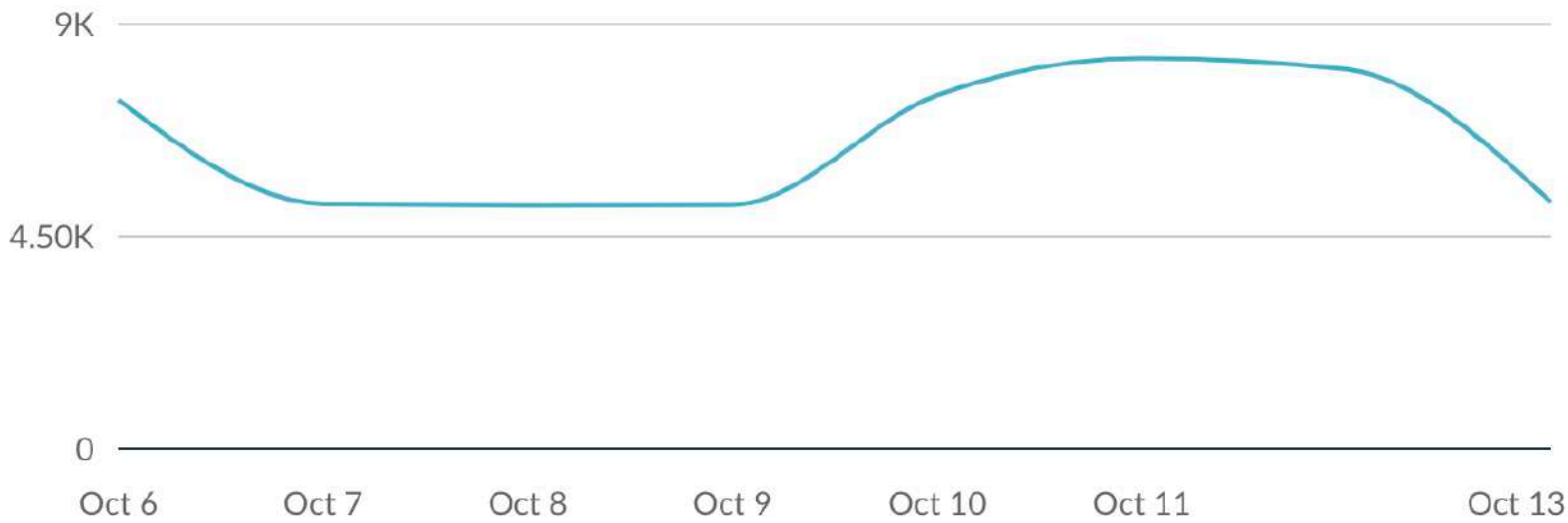


User Hostname Project elasticsearch Requested tasks/goals Build tool Build tool version

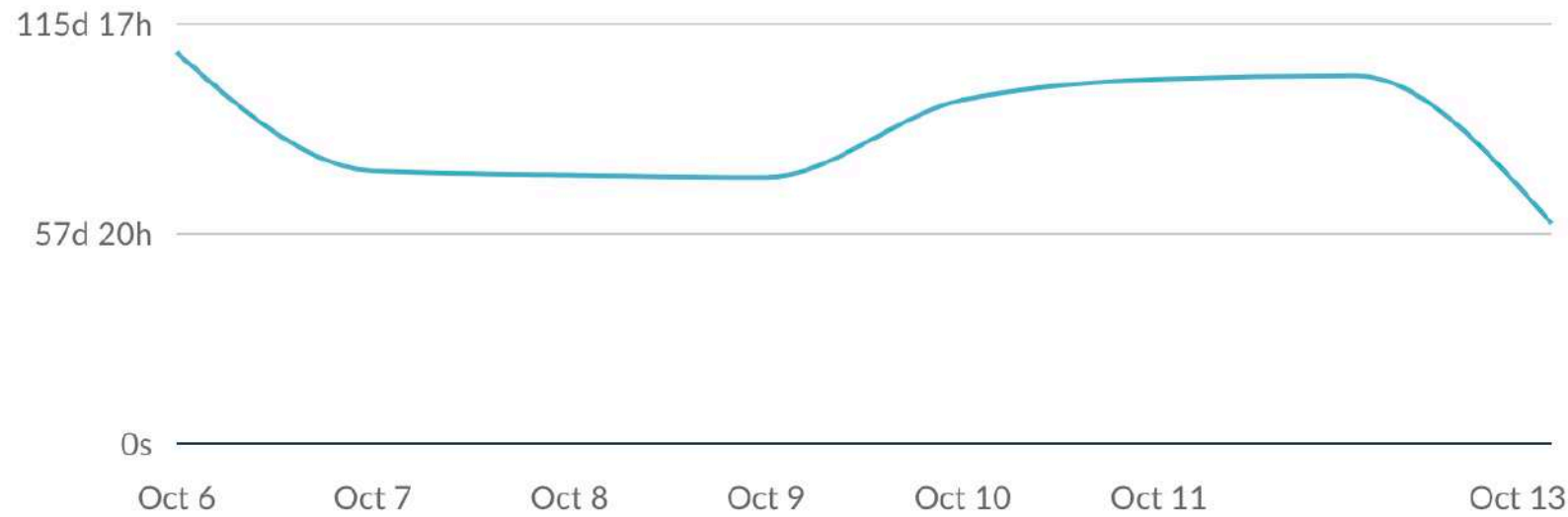
Start time Last 7 days Custom values Tags Outcome Refresh

Overview Build time Serial execution Avoidance savings Build cache overhead Dependency downloading Day Week Month

Build count
51.9K



Cumulative build time
687 d 17 hr



Build time
19 min 4 sec

Non execution 52 sec Execution 18 min 12 sec

Conclusion

When will you rewrite in X?

Elasticsearch's approach: latest JDK bundled

Java under the Hood of



Elasticsearch

Philipp Krenn

@xeraa