

# Repeatable Benchmarking: An Exploration of OpenSearch vs Elasticsearch

Evan Downing

Security Engineer | Trail of Bits



# Background

[AWS Open Source Blog](#)

## Introducing OpenSearch

by Carl Meadows, Jules Graybill, Kyle Davis, and Mehul Shah | on 12 APR 2021 | in [Amazon OpenSearch Service, Announcements, Open Distro for Elasticsearch, Open Source](#) | [Permalink](#) | [Comments](#) | [Share](#)

Today, we are introducing the OpenSearch project, [a community-driven, open source fork of Elasticsearch and Kibana](#). We are making a long-term investment in OpenSearch to ensure users continue to have a secure, high-quality, fully open source search and analytics suite with a rich roadmap of new and innovative functionality. This project includes OpenSearch (derived from Elasticsearch 7.10.2) and OpenSearch Dashboards (derived from Kibana 7.10.2). Additionally, the OpenSearch project is the new home for our previous distribution of Elasticsearch (Open Distro for Elasticsearch), which includes features such as enterprise security, alerting, machine learning, SQL, index state management, and more. All of the software in the OpenSearch project is released under the Apache License, Version 2.0 (ALv2). We invite you to check out the code for [OpenSearch](#) and [OpenSearch Dashboards](#) on GitHub, and join us and the growing community around this effort.



# OpenSearch Benchmarks

The screenshot shows the GitHub repository for `opensearch-project/opensearch-benchmark`. At the top, navigation links include Code, Issues (170), Pull requests (1), Actions, Projects (3), Security, and Insights. Repository statistics show 26 branches, 27 tags, 2,417 commits, 106 forks, and 125 stars. A commit by `lanHoang` is highlighted, titled "Update classifier to correct naming". Below this, a table lists repository folders and their descriptions:

Folder	Description	Last Update
<code>.ci</code>	Lint checker for 1.x/2.x terms (#933)	5 days ago
<code>.github</code>	Lint checker for 1.x/2.x terms (#933)	5 days ago
<code>benchmarks</code>	Initial set of Data Reader changes for Strea...	5 days ago
<code>docker</code>	More updates to the OSB build (#635)	11 months ago
<code>docs</code>	Address remaining legacy names	5 days ago
<code>it</code>	Address Integration Tests	5 days ago
<code>jenkins</code>	Address string issue in Groovy (#646)	11 months ago
<code>osbenchmark</code>	Fix integ test failure in 2.0 RC (#942)	5 days ago
<code>samples/ccr</code>	Address Integration Tests	5 days ago
<code>scripts</code>	Address Integration Tests	5 days ago
<code>tests</code>	Initial set of Data Reader changes for Strea...	5 days ago
<code>.coveragerc</code>	Add coverage configuration	9 years ago
<code>.fossa.yml</code>	Replaced occurrences of "Benchmark" with...	7 months ago

On the right, the "About" section describes the project as a community-driven, open source project for performance testing. It includes links to the README, Apache-2.0 license, Code of conduct, Contributing, Security policy, and Activity. The "Releases" section shows the current version as **2.0.0** (Latest), released 5 days ago, with a total of 27 releases.

## High-level Results

- **OpenSearch is 7.82x faster at NYC Taxis**
- **OpenSearch is 16.55x faster at Big5 Data Histogram**
- **OpenSearch is 3.38x faster at Big5 Term Aggregations**
- **Elasticsearch is faster at 55/81 tasks from all of Big5, NYC Taxis, PMC, NOAA, and Vectorsearch**
- **OpenSearch (Lucene) is 258.2% slower than Elasticsearch (Lucene)**



trailofbits opensearch-benchmark Public

Edit Pins Watch 5 Fork 0 Star 3


main 41 Branches 0 Tags

Go to file

Add file

<> Code

About

 <b>evandowning</b>	Add newer versions of OpenSearch and Elasticsearch (#257) ✓	b88ec85 · 3 months ago	🕒 249 Commits
📁 .github	Add newer versions of OpenSearch and Elasticsearch (#2...		3 months ago
📁 infra	Update infra/README.md		7 months ago
📄 LICENSE	Update license copyright (#256)		4 months ago
📄 README-actions.md	Update README		7 months ago
📄 README.md	Update project READMEs (#233)		7 months ago

README Apache-2.0 license Security

✎ ⋮

## Amazon Benchmark

Benchmark OpenSearch and Elasticsearch with the [OpenSearch Benchmark](#) framework.

This project can automatically benchmark OpenSearch and Elasticsearch and report the results.

You can run benchmarks either manually or with Github Actions.

## Benchmarking Overview

This project can run the following [OpenSearch Benchmark workloads](#):

- big5
- noaa
- nmf

Artifacts related to benchmarking, including documentation, benchmarking harness source code and benchmarking results.

- 📖 Readme
- 📄 Apache-2.0 license
- 📄 Security policy
- 📄 Activity
- 📁 Custom properties
- ☆ 3 stars
- 👁 5 watching
- 🍴 0 forks
- Report repository

### Releases

No releases published  
[Create a new release](#)

### Packages

No packages published  
[Publish your first package](#)

### Contributors 8



# The Reproducibility Problem

- **Challenge**

- OpenSearch runs nightly benchmarks on AWS
- Elasticsearch runs nightly benchmarks on bare-metal
- Blogs/Reports compare benchmarks once
- Rapidly changing implementations

- **Our Solution**

- Cloud-based testing framework
- GitHub Action nightly runs
- Open-sourced automation tools
- Multiple workloads, multiple runs daily



# Our Goal



# Methodology

- **OpenSearch v2.17.1 vs Elasticsearch v8.15.4**
- **OpenSearch Benchmark workloads**
- **Same configuration settings**
- **Fresh AWS instances daily for 15 days**
- **90th percentile service time measurements**
- **5 runs per test (first discarded for warm-up)**
- **1k's - 10k's of sample measurements per operation per run**





# Statistical Validation

- **9 weeks of nightly runs**
- **p-value: 0.05**
  - Likelihood that a difference is due to chance
- **Statistical power: 90%**
  - Likelihood of being able to identify a difference
- **Validated results of all operations**



trailofbits/opensearch-bench

Disable nmslib for OS v3 and

Presentation - Google Sheets

github.com/trailofbits/opensearch-benchmark-private

Work

trailofbits / opensearch-benchmark-private

Q Type / to search

+

<> Code 15 Issues Pull requests 1 Actions Projects Wiki Security Insights Settings

opensearch-benchmark-private Internal

Watch 1 Fork 0 Star 0

evan/big5 had recent pushes 34 minutes ago

Compare & pull request

main 46 Branches 0 Tags

Go to file t

Add file

<> Code

evandowning Remove params to sort (#23) f07ff1a · 3 months ago 283 Commits

.github	Disable nmslib for OS v3 and newer (#32)	3 months ago
infra	Remove params to sort (#23)	3 months ago
LICENSE	Update license copyright (#256)	6 months ago
README-actions.md	Update README	8 months ago
README.md	Update project READMEs (#233)	8 months ago

README Apache-2.0 license Security

About

No description, website, or topics provided.

Readme

Apache-2.0 license

Security policy

Activity

Custom properties

0 stars

1 watching















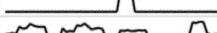










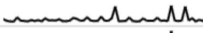














0 forks

Releases

No releases published

Create a new release

# Why this matters

OpenSearch 2.17.1	Elasticsearch 8.15.4	Min	Max
		4.66	23.88
		4.72	20.37
		4.72	20.21
		148.19	469.55
		702.33	2,578.80
		0.89	2.63
		1.04	1,813.50
		214.22	248.16
		309.13	386.47
		4.49	1,526.28
		31.04	1,333.46
		1.18	3.02
		3.05	47.62
		2.61	11.13
		2.60	11.08
		64.79	346.69
		10.77	61.03
		23.35	3,359.73
		19.02	3,315.97
		729.30	1,394.02

# Performance Outliers

- **Both engines showed unexpected performance spikes**
- **OpenSearch: More extreme outliers (up to 1,400x slower)**
- **Elasticsearch: More frequent outliers (19 vs 11 tasks)**
- **Ongoing Investigation**
  - Dedicated EC2 instances
  - Same Region/Zone
  - Same VPC
  - Same CPU hardware spec
  - No correlation on particular day executed



# How are things going now?

Results (Geo Mean)		Elasticsearch	OpenSearch	Ratio
Workload	Category	8.18.1	2.19.1	Elasticsearch 8.18.1 / OpenSearch 2.19.1
big5	Date Histogram	2,302.93	78.86	29.201
big5	Range Queries	1.71	1.70	1.005
big5	Sorting	9.00	6.28	1.434
big5	Term Aggregations	386.90	75.93	5.096
big5	Text Querying	6.28	17.55	0.358
		9.0.1	3.0.0	Elasticsearch 9.0.1 / OpenSearch 3.0.0
big5	Date Histogram	2,204.78	81.97	26.896
big5	Range Queries	1.62	1.31	1.243
big5	Sorting	9.09	6.56	1.385
big5	Term Aggregations	316.06	82.15	3.847
big5	Text Querying	12.21	16.59	0.736



Relative  
Improvement



# Roadmap

- **Refactor Terraform scripts to support more search engines**
  - Milvus
  - Vespa
- **Pin benchmark workload dependencies**
  - Workloads can change, while benchmarking might not
- **Diagnose and remedy outlier results**
- **Update automatic report generation**
- **Publish raw results periodically**



# Takeaways & Contact

- **Bottom Line:**

- Automated, continuous benchmark comparisons
- Open-source and transparent
- User friendly

- **Resources:**

- Framework: <https://github.com/trailofbits/opensearch-benchmark>
- Blog Post: <https://blog.trailofbits.com/2025/03/06/benchmarking-opensearch-and-elasticsearch/>
- Contact: <https://www.trailofbits.com/contact/>

