# 10 Tips on How to Improve Your Elasticsearch Indexing Performance

Jochen Kressin · 2024-02-27

Elasticsearch is a powerful search and analytics engine that allows you to store, search, and analyze large amounts of data quickly. One of the core components of Elasticsearch is its indexing performance. The indexing process is responsible for adding data to Elasticsearch, and it can make or break the performance of your search queries. In this blog post, we will discuss ten tips on how to improve your Elasticsearch indexing performance.

**1. Choose the Right Hardware**

The first step in improving your Elasticsearch indexing performance is to ensure you have the right hardware in place. Elasticsearch relies heavily on RAM, CPU, and disk I/O, so choosing the right hardware is crucial. A machine with sufficient RAM and CPU resources will help speed up the indexing process. SSDs can also help improve disk I/O and reduce indexing times.

**2. Use Bulk Indexing**

Elasticsearch provides a bulk API that allows you to index multiple documents in a single request. Using this API can significantly improve your indexing performance. Instead of indexing one document at a time, you can index hundreds or even thousands of documents at once. This reduces the overhead of network round-trips and improves the efficiency of your indexing process.

**3. Disable Refreshing During Bulk Indexing**

Elasticsearch refreshes its index every second by default. This means that every new document added to the index is immediately searchable. However, this can significantly slow down your indexing process, especially during bulk indexing. You can disable refreshing during bulk indexing to improve your indexing performance.

**4. Decrease Refresh Interval**

Apart from bulk indexing, you can increase indexing performance by generally adjusting the default refresh interval. By increasing the refresh interval, you can reduce the number of refresh operations Elasticsearch needs to perform every second, thereby reducing operational overhead. The downside is that newly added documents may not be searchable immediately. However, in many scenarios, this is probably acceptable.

**5. Optimize Document Design**

The way documents are structured can heavily impact indexing performance. Avoid large and complex documents if possible. Smaller documents are indexed faster. Also, reduce unnecessary nested fields and avoid deep nesting structures, as they can degrade performance.

**6. Reduce the Number of Indexed Fields**

Indexing more fields than you need can slow down your indexing process. More fields mean more analysis and mapping work for Elasticsearch to do, which can increase indexing times. You should only index the fields you need for your search queries.

**7. Disk Speed**

Elasticsearch primary shards write their data to a node's file system, hence disk speed, which is often overlooked, can play a significant role in Elasticsearch indexing performance. For optimal performance, Elasticsearch needs to have fast, reliable, and large-capacity disks. Use SSDs whenever possible.

**8. Index Management**

When an index grows too large, it becomes cumbersome for cluster management, results in low search performance, and reduces indexing speed. If possible, create daily or weekly indexes to mitigate these issues and simplify the management of Elasticsearch. This is especially useful when working with time-series-based data, like log events.

**9. Prevent Mappings Explosion**

An excessive number of mappings can lead to performance issues. Be careful with dynamic mapping and consider defining explicit mappings to control the number of fields indexed.

**10. Translog**

The Translog comes into play when Elasticsearch performs a write operation to the indices. The Transactions Log (Translog) serves as an append-only write-ahead log where every change that affects the index is recorded. Reducing I/O translog operations when feasible will lead to improved indexing performance.

## Conclusion

In conclusion, Elasticsearch is an essential tool for anyone working with large-scale data. However, to get the most out of Elasticsearch, you need to ensure your indexing performance is up to par. By following the ten tips we've outlined above, you can improve your Elasticsearch indexing performance and boost your search and analytics capabilities.

This article was published on the Search Guard blog: https://search-guard.com/blog/10-tips-on-how-to-improve-your-elasticsearch-indexing-performance/
Search Guard is Security and Alerting for Elasticsearch and Kibana — search-guard.com · Start a free trial