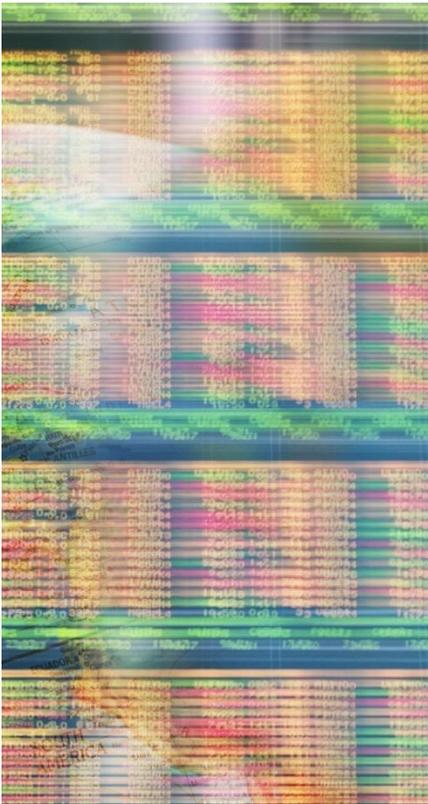


MongoDB* Tuning Guide on 3rd Generation Intel® Xeon® Scalable Processors



Revision Record	2
1. Introduction.....	3
2. BIOS Settings.....	4
3. System Set-up and Optimizations.....	4
3.1 Network configuration	4
3.2 Memory Configuration.....	4
3.3 Disk Configuration	4
3.4 CPU Configuration	4
3.5 Kernel Configuration.....	5
4. MongoDB Settings	5
5. Performance Monitoring Tools.....	7
6. Conclusion.....	7
7. Feedback.....	7

Revision Record

Date	Rev.	Description
05/27/21	1.1	Added conclusion and feedback section
04/06/2021	1.0	Initial public release.

1. Introduction

This guide is targeted towards users who are already familiar with MongoDB* and is designed to provide pointers and setting suggestions for BIOS, OS, MongoDB, and system settings that will provide the best performance for most situations. The settings apply to running MongoDB on a 2-socket Intel® Server System with a 3rd Generation Intel® Xeon® Scalable Processors. For this guide we will be using Ubuntu* 20.04.1 LTS, MongoDB v4.4.1. Acquire and install the target software first. Dependent software binaries or licenses are not distributed with the workloads.

MongoDB is a document-based, general purpose, distributed database built for modern application developers including the cloud. MongoDB is a document database storing data in JSON-like documents. MongoDB is developed by MongoDB Inc. and licensed under the Server Side Public License (SSPL). Learn more about MongoDB here: <https://www.mongodb.com/>

Yahoo! Cloud Serving Benchmark* (YCSB) is an open-source specification and program suite for evaluating retrieval and maintenance capabilities of computer programs. The goal of the Yahoo Cloud Serving Benchmark (YCSB) project is to develop a framework and common set of workloads for evaluating the performance of different "key-value" and "cloud" serving stores. Learn more about YCSB here: <https://research.yahoo.com/news/yahoo-cloud-serving-benchmark/>

3rd Gen Intel® Xeon® Scalable processors deliver industry-leading, workload-optimized platforms with built-in AI acceleration, providing a seamless performance foundation to help speed data's transformative impact, from the multi-cloud to the intelligent edge and back. Here are some of the features in these new processors:

- Enhanced Performance
- Intel® Deep Learning Boost with VNNI
- More Intel® Ultra Path Interconnect
- Increased DDR4 Memory Speed & Capacity
- Intel® Advanced Vector Extensions
- Intel® Security Essentials and Intel® Security Libraries for Data Center
- Intel® Speed Select Technology
- Support for Intel® Optane™ Persistent Memory 200 series

2. BIOS Settings

BIOS can optimize configuration items and recommended values are as follows:

Configuration Item	Recommended Value
Hyper-Threading	Enabled
Hardware Prefetcher	Enable
L2 RFO Prefetch Disable	Enable
Adjacent Cache Prefetch	Enable
DCU Streamer Prefetcher	Enable
DCU IP Prefetcher	Enable
LLC Prefetch	Enable
Turbo Mode	Enable
Freq Throttling Lock (ICX system)	Disable
Power & Perf Policy	Performance
Frequency Driver	intel_pstate

3. System Set-up and Optimizations

3.1 Network configuration

Fast network is critical for the MongoDB performance. To make sure that network is not the performance bottleneck, recommendation is to use 10Gb/sec or higher network bandwidth between server and client.

3.2 Memory Configuration

- Use identical DIMM types throughout the platform:
 - Same size, speed, and number of ranks
- Maximize the same number of channels populated in each memory controller
- Use a "balanced" platform configuration:
 - All available memory channels populated equally
 - Identical DIMMs in all locations (size/speed/rank)

3.3 Disk Configuration

MongoDB performance gets limited if the system-under-test (SUT) is using SSD for storing the database. The recommendation is to use the high disk bandwidth (i.e. NVME drives) that help achieve maximum performance. For instance: 4 x NVME drives in RAID 0 configuration of 6 TB (4x1.5TB)

3.4 CPU Configuration

This guide was developed using 8380 3rd Generation Intel® Xeon® Scalable Processor(s).

- Set the corresponding CPU governor to performance mode


```
sudo cpupower frequency-set --governor performance
```
- Disable the IRQBalance


```
sudo service irqbalance stop
```

- IRQ Affinity

IRQs can be affinityized to one or more specific CPU cores. Therefore, a particular IRQ will be handled only by a specific set of CPU cores. The recommended setting is to affinityize IRQ queues to all cores to improve the workload performance.

3.5 Kernel Configuration

- Set the transparent hugepage

```
echo "never" | sudo tee /sys/kernel/mm/transparent_hugepage/defrag
```

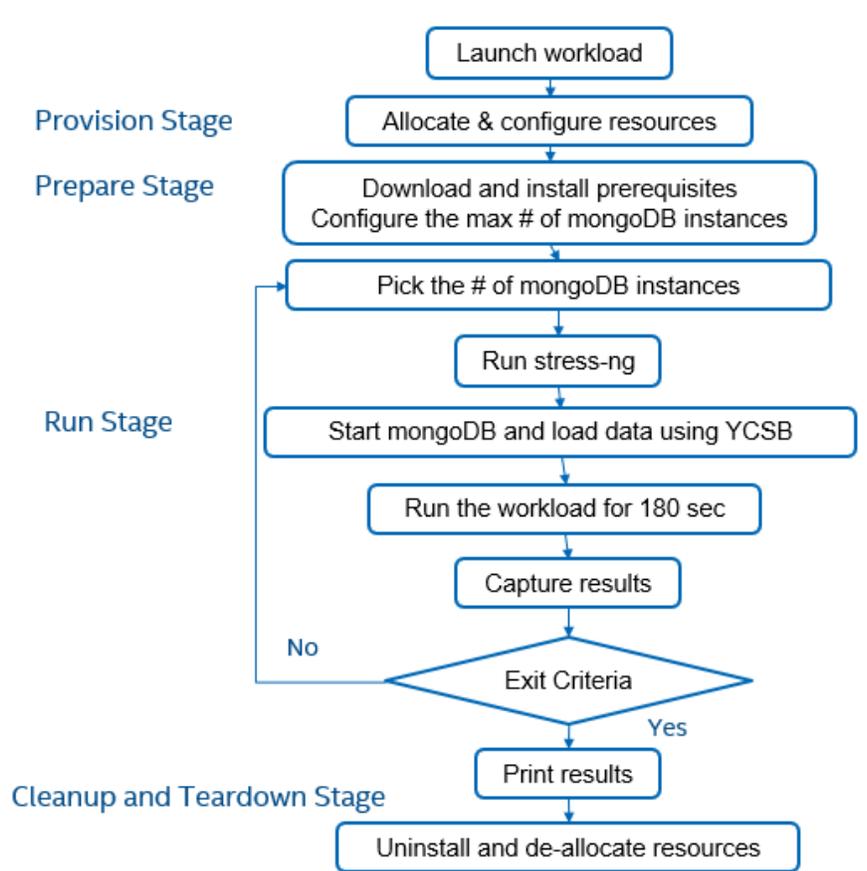
- Set the transparent zone_reclaim_mode

```
echo "0" | sudo tee /proc/sys/vm/zone_reclaim_mode
```

4. MongoDB Settings

4.1 Methodology

In doing our testing, we followed the flow shown below:



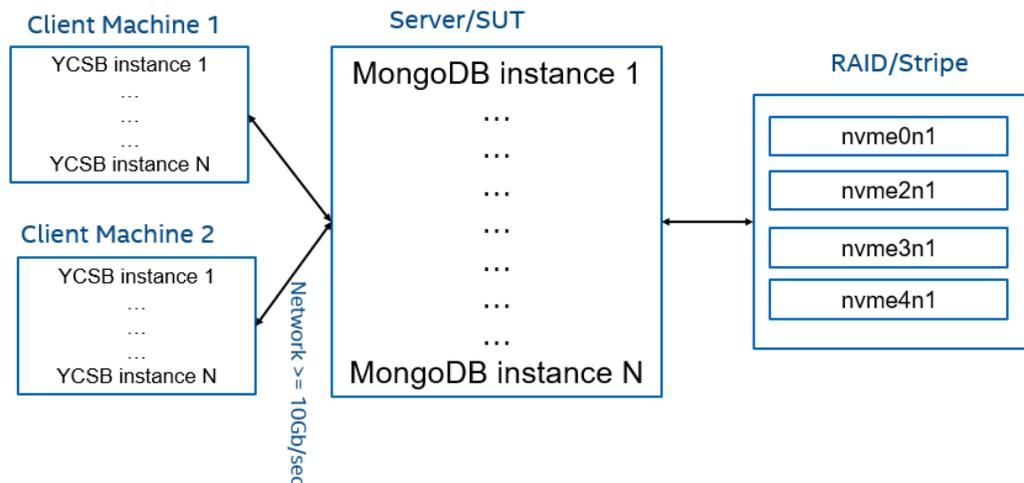
Follow the below-mentioned steps to get mongod running:

- Allocate and configure resources on server and client machines as specified by your workload configuration files
- Download and install the prerequisites (libcurl4) and the workload (mongoDB, YCSB)
- Configure the maximum number of mongoDB instances one can run on the target system

MongoDB* Performance Tuning Guide on 3rd Generation Intel® Xeon® Scalable Processors

- Engage the storage subsystem by limiting memory available to mongoDB instances and run stress-ng in parallel to occupy the remaining memory
- Start the mongoDB and use YCSB to load the data “90Read10Update”
- Run the workload for 180 sec
- Capture the results and check the exit criteria
- Uninstall and de-allocate the resources on server and client
- Print results, throughput being the primary metric

Below is a diagram of our testing setup:



4.2 Application version and URL

MongoDB Enterprise Version 4.4.1 - https://downloads.mongodb.com/linux/mongodb-linux-x86_64-enterprise-ubuntu2004-4.4.1.tgz

- Install with default settings

4.2 Strategy for running MongoDB

MongoDB is an open-source scale-out NoSQL database. In the test scenario, this server-client application has been set up to run multiple MongoDB instances on the SUT (server under test) and automatically increasing the load on the target system incrementally until a maximum throughput and maximum throughput under SLA is achieved. The input database size impacts the performance of the workload. Small databases fit in DRAM and users can get better performance. However, for the more realistic scenario, the database size should be larger than the DRAM size. The latter scenario has been used for the test setup where the storage system has been engaged by limiting memory available to the MongoDB instances. The goal is to utilize the system to its full potential to achieve the max throughput and max throughput under SLA (3msec p99 read latency).

4.3 Storage Engine

WiredTiger is the default storage engine with MongoDB and what we have tested with. Use default settings. MongoDB recommends cache size of 50% of (RAM -1GB). Default settings have been used for WiredTiger.Clients

Two (2) clients minimum are required. Client CPU can be slower speed than the SUT. On the clients, be sure to install YCSB v0.17 (or higher), Java v8 (or higher), and Maven v3.6.3 (or higher)

Filesystem mount options commands used are as follows:

```
sudo mdadm --create /dev/md0 --level=stripe --raid-devices=4 /dev/nvme1n1
/dev/nvme2n1 /dev/nvme3n1 /dev/nvme4n1
sudo mke2fs -F -E lazy_itable_init=0,discard -O ^has_journal -t ext4 -b 4096
/dev/md0
```

5. Performance Monitoring Tools

Software tools that can be used to analyze the workload performance and to generate and monitor system resource statistics while the workload runs are as follows:

Open Source Tools:

- Sar
- Dstat
- Collectd

6. Conclusion

We understand every application is unique. We shared many of our experiences with MongoDB, hoping that some of our learnings could be applied to your specific application. MongoDB has been well tested on Intel platforms. With 3rd Generation Intel® Xeon® Scalable processor, Intel takes it even further by optimizing the platform as a whole -- CPU, memory, storage, and networking working together for the best user experience.

7. Feedback

We value your feedback. If you have comments (positive or negative) on this guide or are seeking something that is not part of this guide, please reach out to us here:

<https://community.intel.com/t5/Software-Tuning-Performance/bd-/software-tuning-perf-optimization>

Notices & Disclaimers

Intel technologies may require enabled hardware, software or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

Code names are used by Intel to identify products, technologies, or services that are in development and not publicly available. These are not "commercial" names and not intended to function as trademarks

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.