Ismail Syed          Data Science, Vector & ML Specialist EMEA          Oracle

# Configuring Oracle Database 23ai
# In Memory Vector Index

## Pre-Requisites

- Oracle Database 23ai Service Up and Running
- Oracle Database 23ai Node IP Address
- Oracle Database 23ai Node SSH Private Key

## Description

HNSW In-Memory Indexes are stored within the SGA. The total memory size to store In-Memory Indexes are defined by setting the VECTOR_MEMORY_SIZE parameter. VECTOR_MEMORY_SIZE is an object within SGA_MAX_SIZE therefore VECTOR_MEMORY_SIZE < SGA_MAX_SIZE. Ideally you want VECTOR_MEMORY_SIZE to be <= 70 or 80% of SGA_MAX_SIZE. The HNSW Vector Indexes need to fit entirely within the VECTOR_MEMORY_SIZE.  Therefore, you want to make VECTOR_MEMORY_SIZE > the space needed by the vector index.

## Configure the use of In-Memory Vector Indexes

1. SSH into the Oracle Database 23ai Node: `ssh –i <private–key>.key opc@<host–ip–address>`

2. Switch to the root user: `sudo su`

3. Switch to the oracle user: `su – oracle`

4. Login to SQLPlus using sysdba: `sqlplus / as sysdba`

5. Create a copy of the parameter file: `create pfile from spfile;`

6. Alter ***vector_memory_size*** parameter, ensure that the size that is set is at max 70-80% of the ***sga_max_size*** parameter, where ***sga_max_size*** must be about 60-70% of physical RAM on the machine which is normally set by default. In the following example we have set the ***vector_memory_size*** to 500MB: `ALTER SYSTEM SET vector_memory_size = 500M SCOPE=SPFILE;`

7. Shutdown the Database: `shutdown`

8. Start Up the Database: `startup`

9. Verify the ***vector_memory_size*** parameter has been set: `show parameter vector_memory_size;`

10. Open all pluggable databases: `alter pluggable database all open;`

11. Exit SQLPlus: `exit`

12. Exit oracle user: `exit`

13. Exit root user: `exit`

14. Exit SSH session: `exit`