Given your cluster with **746 cores**, **3 TB memory**, and **5 nodes**, and assuming you're processing **2 TB input data** with each executor having **16 GB memory**, here’s how to configure the spark-submit command and estimate the number of iterations.

### **Step 1: Cluster Configuration Analysis**

#### **1. Memory Per Node**

Total memory: 3 TB=3072 GB3 \, \text{TB} = 3072 \, \text{GB}3TB=3072GB.

Memory per node:

Memory per Node=3072 GB5 Nodes=614.4 GB/Node.\text{Memory per Node} = \frac{3072 \, \text{GB}}{5 \, \text{Nodes}} = 614.4 \, \text{GB/Node}.Memory per Node=5Nodes3072GB​=614.4GB/Node.

#### **2. Executor Memory**

* Each executor is assigned **16 GB memory**, plus overhead memory.
* Overhead memory is typically **10% of executor memory**: Overhead Memory=16 GB×0.1=1.6 GB.\text{Overhead Memory} = 16 \, \text{GB} \times 0.1 = 1.6 \, \text{GB}.Overhead Memory=16GB×0.1=1.6GB. Total memory per executor: 16 GB+1.6 GB=17.6 GB.16 \, \text{GB} + 1.6 \, \text{GB} = 17.6 \, \text{GB}.16GB+1.6GB=17.6GB.

#### **3. Executors Per Node**

Executors per Node=Memory per NodeMemory per Executor=614.417.6≈34 Executors/Node.\text{Executors per Node} = \frac{\text{Memory per Node}}{\text{Memory per Executor}} = \frac{614.4}{17.6} \approx 34 \, \text{Executors/Node}.Executors per Node=Memory per ExecutorMemory per Node​=17.6614.4​≈34Executors/Node.

#### **4. Total Executors**

Total Executors=Executors per Node×Nodes=34×5=170 Executors.\text{Total Executors} = \text{Executors per Node} \times \text{Nodes} = 34 \times 5 = 170 \, \text{Executors}.Total Executors=Executors per Node×Nodes=34×5=170Executors.

#### **5. Executor Cores**

* With **746 cores** across 170 executors:

Cores per Executor=Total CoresTotal Executors=746170≈4 Cores/Executor.\text{Cores per Executor} = \frac{\text{Total Cores}}{\text{Total Executors}} = \frac{746}{170} \approx 4 \, \text{Cores/Executor}.Cores per Executor=Total ExecutorsTotal Cores​=170746​≈4Cores/Executor.

### **Step 2: Parallelism**

To process **2 TB of input data**, we need sufficient parallelism:

* Recommended parallelism is **2-3x total cores**:

Parallelism=746×2=1492 Tasks.\text{Parallelism} = 746 \times 2 = 1492 \, \text{Tasks}.Parallelism=746×2=1492Tasks.

Set spark.default.parallelism and spark.sql.shuffle.partitions to **1492**.

### **Step 3: Spark-Submit Command**

Here’s the spark-submit command for this configuration:

bash

Copy code

spark-submit \

--master yarn \

--deploy-mode cluster \

--name "Spark Job for 2TB Data" \

--num-executors 170 \ # Total executors (34 per node x 5 nodes)

--executor-cores 4 \ # Cores per executor

--executor-memory 16G \ # Memory per executor

--conf spark.executor.memoryOverhead=1.6G \ # Overhead memory

--conf spark.default.parallelism=1492 \ # Parallelism

--conf spark.sql.shuffle.partitions=1492 \ # Shuffle partitions

/path/to/your\_spark\_application.py

### **Step 4: Estimating Iterations**

#### **1. Tasks per Stage**

With **2 TB data** and **1492 tasks (parallelism)**:

* Each task processes:

Data per Task=2 TB1492≈1.34 GB/Task.\text{Data per Task} = \frac{2 \, \text{TB}}{1492} \approx 1.34 \, \text{GB/Task}.Data per Task=14922TB​≈1.34GB/Task.

#### **2. Iterations per Stage**

* **Total Executors**: 170 executors with 4 cores each can run:

Tasks Running Simultaneously=170×4=680 Tasks.\text{Tasks Running Simultaneously} = 170 \times 4 = 680 \, \text{Tasks}.Tasks Running Simultaneously=170×4=680Tasks.

* To complete 1492 tasks:

Iterations per Stage=TasksSimultaneous Tasks=1492680≈2.2 Iterations (round up to 3).\text{Iterations per Stage} = \frac{\text{Tasks}}{\text{Simultaneous Tasks}} = \frac{1492}{680} \approx 2.2 \, \text{Iterations (round up to 3)}.Iterations per Stage=Simultaneous TasksTasks​=6801492​≈2.2Iterations (round up to 3).

#### **3. Number of Stages**

The number of stages depends on your Spark job's operations:

* Simple ETL job (e.g., map, filter, reduceByKey): **3-4 stages**.
* Complex jobs (e.g., joins, aggregations): **5-8 stages**.

#### **4. Total Iterations**

Total Iterations=Iterations per Stage×Number of Stages.\text{Total Iterations} = \text{Iterations per Stage} \times \text{Number of Stages}.Total Iterations=Iterations per Stage×Number of Stages.

* For a simple job with 4 stages:

3 Iterations per Stage×4 Stages=12 Iterations.3 \, \text{Iterations per Stage} \times 4 \, \text{Stages} = 12 \, \text{Iterations}.3Iterations per Stage×4Stages=12Iterations.

* For a complex job with 8 stages:

3 Iterations per Stage×8 Stages=24 Iterations.3 \, \text{Iterations per Stage} \times 8 \, \text{Stages} = 24 \, \text{Iterations}.3Iterations per Stage×8Stages=24Iterations.

### **Final Estimate**

* **Simple Job**: 12 iterations.
* **Complex Job**: 24 iterations.

### **Monitoring Execution**

1. **Spark UI**:
   * Use the **Stages tab** to view the number of stages and tasks executed.
2. **YARN UI**:
   * Monitor resource allocation and task distribution across the cluster.