Cluster Sizing (Azure Databricks)

Cluster Types

- **All-purpose:** Shared by multiple users, ideal for ad-hoc analysis, data exploration and development.
  -**Job cluster:** ETL jobs.

Cluster Modes

- **Standard:** Ideal for processing large amounts of data with Spark.
- **Single node:** Jobs that use smaller amounts of data or single-node machine learning libraries.
- **High Concurrency:** Groups of users that need to share resources.

In Azure, you can create clusters using a combination of on-demand and spot instances, to reduce costs.

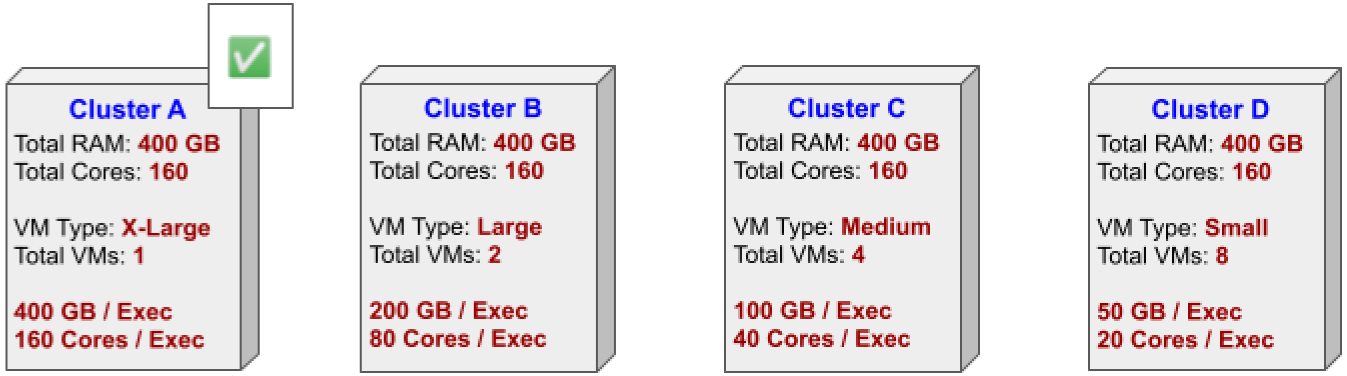
Autoscaling

- Another way to reduce costs is autoscaling (for instance, in high concurrency clusters).
- If the minimum number of nodes is too small, users might complain that it is very slow.
- Autoscaling is not compatible with spark-submit jobs nor with Delta caching (Delta lake). If you need this features, consider a fixed size cluster.
- Autoscaling can also be applied to local storage.
- Clusters are by default terminated after 120 minutes of inactivity.

Cluster sizing

- Besides the number of workers, one needs to consider:
    - Total executor cores: Sum of cores across all worker nodes /executors.
    - Total executor memory: Sum of RAM across all executors.
    - Executor local storage: local disk is used in the case of memory spills during shuffles or caching.
- Workers with high amount of RAM can help jobs perform more efficiently, but also lead to delays during garbage collection.
- To minimize impact of garbage collection sweeps, it is better to have many instances with smaller RAM sizes (exceptions apply).

Cluster sizing examples: Data Analysis



✅

Cluster A
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **X-Large**
Total VMs: **1**

400 GB / Exec
160 Cores / Exec

Cluster B
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Large**
Total VMs: **2**

200 GB / Exec
80 Cores / Exec

Cluster C
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Medium**
Total VMs: **4**

100 GB / Exec
40 Cores / Exec

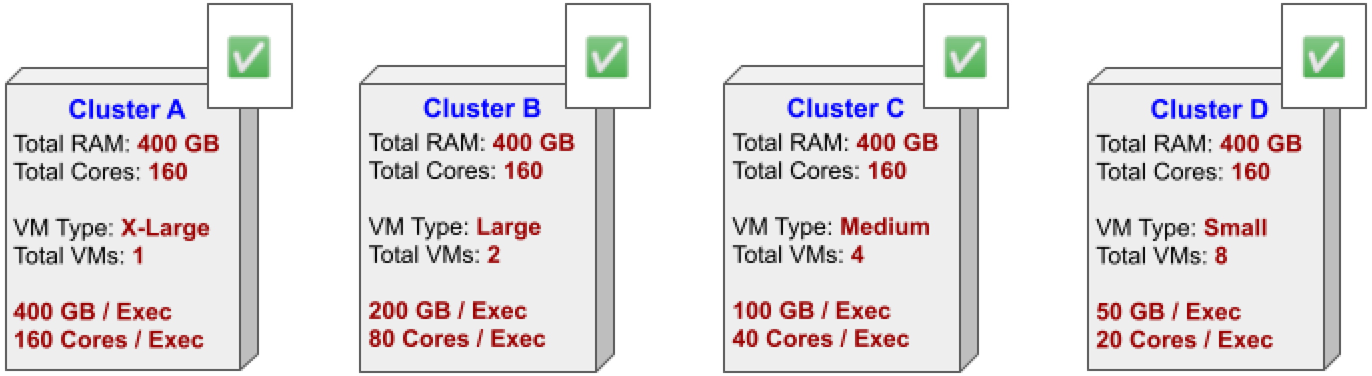
Cluster D
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Small**
Total VMs: **8**

50 GB / Exec
20 Cores / Exec

- Smaller number of nodes to reduce network + I/O operations.
- A cluster with a large number of nodes with less memory and storage will require more data shuffling.
- For training machine learning models, type B or C are recommended.

Cluster sizing examples: Basic batch ETL jobs



✓ **Cluster A**
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **X-Large**
Total VMs: **1**

400 GB / Exec
160 Cores / Exec

✓ **Cluster B**
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Large**
Total VMs: **2**

200 GB / Exec
80 Cores / Exec

✓ **Cluster C**
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Medium**
Total VMs: **4**

100 GB / Exec
40 Cores / Exec

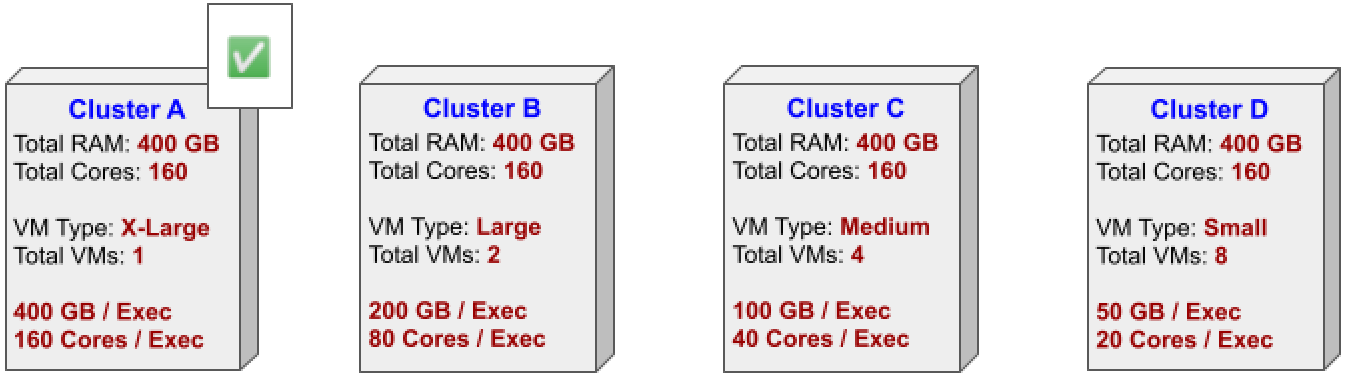
✓ **Cluster D**
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Small**
Total VMs: **8**

50 GB / Exec
20 Cores / Exec

- Delta Caching is probably not useful, as data re-reading is not expected.

Cluster sizing examples: Complex batch ETL jobs



✅

Cluster A
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **X-Large**
Total VMs: **1**

400 GB / Exec
160 Cores / Exec

Cluster B
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Large**
Total VMs: **2**

200 GB / Exec
80 Cores / Exec

Cluster C
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Medium**
Total VMs: **4**

100 GB / Exec
40 Cores / Exec

Cluster D
Total RAM: **400 GB**
Total Cores: **160**

VM Type: **Small**
Total VMs: **8**

50 GB / Exec
20 Cores / Exec

- Jobs with many joins and unions across multiple tables.