BITS Pilani presentation

BITS Pilani
Pilani Campus

Dr. Vivek V. Jog
Dept. Of Computer Engineering

innovate achieve lead

BITS Pilani
Pilani Campus

Big Data Systems (S1-24_CCZG522)
Lecture No.8

Sizing parameters

Storage Considerations

The number of data nodes you need is determined by the size of the data, how it will be analyzed, and the number of replicas you will have. By default, Apache Hadoop has 3 copies.
In this case, if we want to store X GB of data we need X*3 GB of storage for the forecasted period.

Processing Considerations

In addition to having enough space to store your data, you will need room for data processing, computing, and miscellaneous other tasks.

We can assume that, on an average day, only 10% of data is being processed, and a data process creates three times temporary data.

Therefore, you need to account for around 30% of your total storage as extra space.

Contd…

Number of Data Nodes Required

The final calculation for the number of data nodes required for your system will be dependent on your JBOD (“just a bunch of disks”) capacity.

For example: Let's say that you need 500GB of space. If you have a JBOD of 12 disks, and each disk can store 6TB of data, then the data node capacity, or the maximum amount of data that each node can store, will be 72 TB. Data nodes can be added as the data grows, so to start with its better to select the lowest number of data nodes required.

In this case, the number of data nodes required to store 500GB of data equals 500/72, or approximately 7.

Note: Number of Data nodes* = (no. of disks*)

Summarizing

While setting up the cluster, we need to know the below parameters:

- What is the volume of data for which the cluster is being set? (For example, 100 TB.)
- The retention policy of the data. (For example, 2 years.)
- The kinds of workloads you have — CPU intensive, i.e. query; I/O intensive, i.e. ingestion, memory intensive, i.e. Spark processing. (For example, 30% jobs memory and CPU intensive, 70% I/O and medium CPU intensive.)
- The storage mechanism for the data — plain Text/AVRO/Parque/Jason/ORC/etc. or compresses GZIP, Snappy. (For example, 30% container storage 70% compressed.)

Estimating Data Node Requirement

Data Nodes Requirements

With the above parameters in hand, we can plan for commodity machines required for the cluster. (These might not be exactly what is required, but after installation, we can fine tune the environment by scaling up/down the cluster.) The nodes that will be required depends on data to be stored/analyzed.

By default, the Hadoop ecosystem creates three replicas of data. So if we go with a default value of 3, we need storage of *100TB *3=300 TB* for storing data of one year. We have a retention policy of two years, therefore, the storage required will be *1 year data* retention period=300*2=600 TB*.

Assume 30% of data is in container storage and 70% of data is in a Snappy compressed Parque format. From various studies, we found that Parquet Snappy compresses data to 70-80%.

We have taken it 70%. Here is the storage requirement calculation:

Total storage required for data =total storage % in container storage + total storage * %in compressed format*expected compression*

600.30+600*.70*(1-.70)=180+420*.30=180+420*.30=306 TB.*

Contd…

In addition to the data, we need space for processing/computation the data plus for some other tasks. We need to decide how much should go to the extra space. We also assume that on an average day, only 10% of data is being processed and a data process creates three times temporary data. So, we need around 30% of total storage as extra storage.

Hence, the total storage required for data and other activities is *306+306*.30=397.8 TB*.

As for the data node, JBOD is recommended. We need to allocate 20% of data storage to the JBOD file system. Therefore, the data storage requirement will go up by 20%. Now, the final figure we arrive at is *397.8(1+.20)=477.36 ~ 478 TB*.

Let's say *DS=478 TB*.

Additional considerations

Now, we need to calculate the number of data nodes required for 478 TB storage. Suppose we have a JBOD of 12 disks, each disk worth of 4 TB. Data node capacity will be 48 TB.

The number of required data nodes is *478/48 ~ 10*.

In general, the number of data nodes required is *Node= DS/(no. of disks in JBOD*diskspace per disk)*.

Note: We do not need to set up the whole cluster on the first day. We can scale up the cluster as data grows from small to big. We can start with 25% of total nodes to 100% as data grows.

Now, let's discuss data nodes for batch processing (Hive, MapReduce, Pig, etc.) and for in-memory processing.

As per our assumption, 70% of data needs to be processed in batch mode with Hive, MapReduce, etc.

10.70=7* nodes are assigned for batch processing and the other 3 nodes are for in-memory processing with Spark, Storm, etc.

CPU Cores and Tasks per Node

For batch processing, a *2*6*-core processor (hyper-threaded) was chosen, and for in-memory processing, a *2*8* cores processor was chosen. For batch processing nodes, while one core is counted for CPU-heavy processes, .7 core can be assumed for medium-CPU intensive processes. As we have assumption, 30% heavy processing jobs and 70% medium processing jobs, Batch processing nodes can handle [(no. of cores* %heavy processing jobs/cores required to process heavy job)+ (no. of cores* %medium processing jobs/cores required to process medium job)]. Therefore tasks performed by data nodes will be;

12.30/1+12*.70*/.7=3.6+12=15.6 ~15* tasks per node.

As hyperthreading is enabled, if the task includes two threads, we can assume 15*2~30 tasks per node.

RAM Requirements for Data Node

Now, let's calculate RAM required per data node. RAM requirements depend on the below parameters.

*RAM Required=DataNode process memory+DataNode TaskTracker memory+OS memory+CPU's core number *Memory per CPU core*

At the starting stage, we have allocated four GB memory for each parameter, which can be scaled up as required. Therefore, RAM required will be *RAM=4+4+4+12*4=60* GB RAM for batch data nodes and *RAM=4+4+4+16*4=76* GB for in-memory processing data nodes.

Cluster Calculations

innovate achieve lead

Estimate the number of data nodes based on data size and increase pattern :

Estimating the hardware requirement is always challenging in Hadoop environment because we never know when data storage demand can increase for a business.

Understanding following factors in detail to conclude for the current scenario of adding right numbers to the cluster.

Scenario : If 8TB is the available disk space per node (10 disks with 1 TB, 2 disk for operating system/System Logs etc. were excluded.). **Assuming initial data size is 600 TB. How will you estimate the number of data nodes (n)?**

Parameters

innovate achieve lead

Cluster Size : Hadoop

- The actual size of data to store – **600 TB**
- **Data trending analysis-prediction**: At what pace the data will increase in the future (per day/week/month/quarter/year)
- **Replication Factor** plays an important role – default 3x replicas.
- **Logs:** Hardware machine overhead (OS, logs etc.) – 2 disks were considered
- **Disk I/O:** Intermediate mapper and reducer data output on hard disk - 1x
- **Storage capacity**: Space utilization between 60 % to 70 %, disk utilization.
- **Compression ratio**

Cluster Size

innovate achieve lead

Cluster Size : Hadoop

Let's do some calculation to find the number of data nodes required to store 600 TB of data:

Base calculation:

Data Size – 600 TB
Replication factor – 3
Intermediate data – 1
Total Storage requirement – (3+1) * 600 = 2400 TB
Available disk size for storage – 8 TB
Total number of required data nodes (approx.): 2400/8 = 300 machines
Actual Calculation: Base Calculation + Disk space utilization + Compression ratio

Actual Calculations

innovate achieve lead

Cluster Size : Hadoop

Disk space utilization – 65 % (differ business to business)

Compression ratio – 2.3

Total Storage requirement – 2400/2.3 = 1043.5 TB

Available disk size for storage – 8*0.65 = 5.2 TB

Total number of required data nodes (approx.): 1043.5/5.2 = 201 machines

Actual usable cluster size (100 %): (201*8*2.3)/4 = 925 TB

Case: Business has predicted 20 % data increase in a quarter, and we need to predict the new machines to be added in a year

Contd..

Data increase – 20 % over a quarter

Additional data:

1st quarter: 1043.5 * 0.2 = 208.7 TB

2nd quarter: 1043.5 * 1.2 * 0.2 = 250.44 TB

3rd quarter: 1043.5 * (1.2)^2 * 0.2 = 300.5 TB

4th quarter: 1043.5 * (1.2)^3 * 0.2 = 360.6 TB

Available disk size for storage – 8*0.65 = 5.2 TB

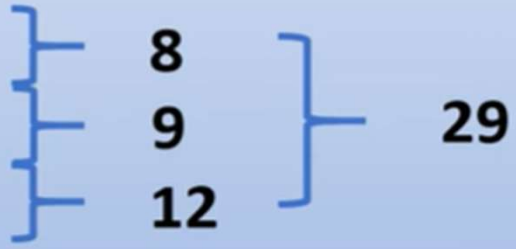
Additional data nodes requirement (approx.):

1st quarter: 208.7/5.2 = **41 machines**

2nd quarter: 250.44/5.2 = **49 machines**

3rd quarter: 300.5/5.2 = **58 machines**

4th quarter: 360.6/5.2 = **70 machines**



With these numbers we can predict next year additional machines requirement for the cluster :

(last quarter + 29) and so on.

That is all for the day