Special Issue Published in International Journal of Trend in Research and Development (IJTRD), ISSN: 2394-9333, www.ijtrd.com

A Survey on Comparison of Job Scheduling Algorithms in Hadoop

Sneha. M[1], Shoney Sebastian[2] and Rajeshwari C N[3],
[1]PG Scholar, [2, 3]Associate Professor,
Department of Computer Science, Christ University, Bangalore, India

***Abstract*:** Cloud computing is emerging as a new trend to store and analyze data, Hadoop is a very good and efficient framework that can be used to process the data through map reduce programming paradigm. This paper provides the insight for the rise of big data and what role does the hadoop plays to handle this. The structure and architecture view of hadoop with all the components and Yarn architecture is discussed. Scheduling is allot the set of jobs to process in the corresponding slots based on the few job scheduling algorithms, comparative study of all the scheduling algorithms are made which would be helpful while processing the data.

Keywords: *Map Reduce, Yarn, Hadoop, slots, HDFS.*

I. INTRODUCTION

Cloud computing is a network of remote servers hosted on the internet to store, manage and process data. To meet every change in business needs organization needs to invest time and budget to scale up their IT infrastructure such as hardware, software and services. However with unpremises IT infrastructure the scaling process can be slow and organizations are frequently unable to achieve optimal utilization of IT infrastructure. [1]Cloud computing is a paradigm shift that provides computing over the internet, cloud computing service consists of highly optimized data centers that provide various software, hardware and information resources to use for needed. Organizations can simply connect to the cloud and can use the available resources on a pay per use process. This way company avoids capital expenditure on additional and premise infrastructure services and instantly scale up or scale down according to the business requirements.

Cloud computing [2] can be deployed by using three different models that is private, public and hybrid cloud. Private cloud functions solely for one organization and is highly secure. Public cloud is owned by the cloud service provider and offers the highest level of efficiency and shared resources. Hybrid cloud is a combination of private and public deployment models, this provides increased efficiency. By using infrastructure as a service model organization gets infrastructure components such as computing power and storage capacity. Here the organization has the control over entire IT infrastructure including the hosting environment and their applications, however the organization needs to allocate the additional staffs to maintain and manage the infrastructure and the application. Microsoft has proposed remote desktops and virtual machines to Microsoft windows Azure. The Paas model provides organization with an offering or a run time environment to create and deploy applications. Here the organization is only responsible for maintenance and management of the applications, Microsoft provides windows azure platform as Paas offering. The Saas model provides organization with ready to use application, user combination of cloud based computing and storage services. Microsoft provides various online services such as business productivity online service and Microsoft dynamics CRM online as a saas offering. Nowadays to produce any of the new products we need to store and process the data which requires high volume of storage, consequently the company is looking for cost effective compute and storage solution. So they can use window azure platform service like pay per use and they can be stored and analyzed the data rather than focusing on physical set up of the software.

Big data and its tools

Big data refers [3] to huge volume of data that cannot be stored and processed within a given timeframe. So how huge this data needs to be? Basically big data will be in the size of GB/TB/PB/Ex/ > this size, this does not refers to big data completely even a small amount of data is referred to as big data depending on the context being used. Some real world examples are Facebook, twitter, you tube, Google, linked in etc., Facebook alone reports 100TB of data everyday which is been reported by some of the reported block whereas twitter processes 400 million tweets every day and in you tube for each minute 48 hours of fresh videos are uploaded, you can just imagine how much volume of data are stored and processed in these sites. But as the number of users keep growing on these sites storing and processing data becomes a challenging task, since these data holds a valuable information these data needs to be processed in a short span of time, by using this valuable information companies can boost their sales and increase revenues. By making use of traditional computing system we would not be able to be able to accomplish this task within the given time frame. This is where big data tools comes into picture, with that storing and processing can be done in feasible way.

Few big data tools are like windows azure and all kind of Nosql databases. Windows azure is Microsoft cloud computing platform which is designed to run your applications at scale on internet, windows azure takes an application center in cloud computing and that means they manage the entire lifecycle of an application and also they monitor the application when it is running out in the internet.

Challenges in big data

In big data most of the data's generated are unstructured and semi structured data [4]. It takes time to derive meaning out of it to place and store the data for further processing. Creating a code that converts data from unstructured or semi structured format to usable data code is a complex process, once a hand coded proof of concept solution is developed then protection raises unexpected issues more than half of hand coded solutions never make it to protection. It's a tedious job to find people with right skills to staff big data projects when you do find the resources they are both expensive and hard to retain and also technology changes where today's solution would not be the best way to solve tomorrows problem.

Hadoop for big data analytics

For connecting to unstructured and structured data like file transactions and flat files, social and interaction data ,sensor

and machine data.[4] A developer will create same integration rules or mappings which can now run on hadoop for faster and less expensive processing. The results of that processing can be then exported anywhere like for ware housing, transactions, reporting and even for alerts. Bigdata editions visual development interface is already familiar to current power center developers that interface enables them to move straight to developing on hadoop without knowing hadoop. Moreover hadoop provides Scalability and not expensive compared to other tools used to handle bigdata.

Framework of hadoop

Hadoop distributed file system is a cluster on which the data is stored it has master slave kind of architecture where Name node is the master which contains information about Metadata and also manages all the data nodes where the actual data gets stored in terms of blocks. There is 1 Name node in the cluster and there can be thousands of DataNode in the cluster, NameNode is a not a high availability machine i.e. it has high memory and redundant power supplies also if it goes down the performance is affected. Secondary NameNode takes the backup of NameNode which will be useful in times of failure like if the NameNode fails with the log of secondary NameNode which has a configuration file of NameNode we can setup the NameNode.

Map reduce programming paradigm

Map Reduce is a framework - a programming(java) model of hadoop, to run any of the batch processing which is distributed across the cluster the appropriate program should be written in map reduce where it takes cares of scheduling of jobs, monitoring of jobs, allocating resources, mining of jobs and managing failures. Job tracker is the core service or the thread running all the times which is used for scheduling the jobs on the data nodes and as well as to monitor these tasks. [5] The jobs which should be run on cluster are first submitted to the job tracker which has a queue which will keep on submitting or scheduling these tasks to the various data nodes through a task tracker and also it monitors the submitted jobs. The map reduce program will run on the data nodes and the data is fetched from the data node based on map and reduce program through the task tracker. Both of these components are necessary to run the jobs on data node and program can be executed parallelly on thousands of machine [5].

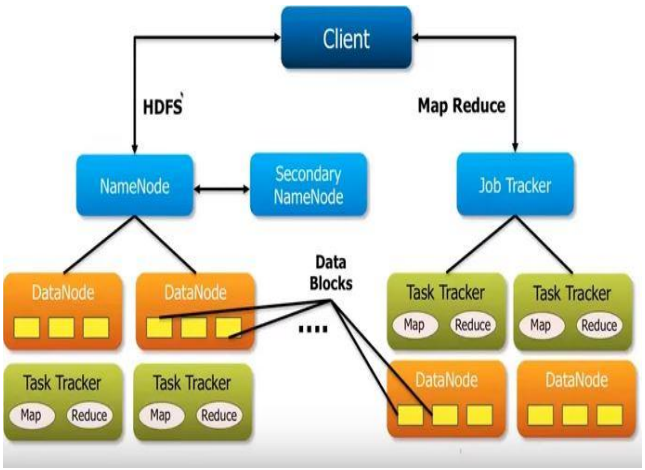
The file which we wants to be stored and processed into the database that will be splitted into different chunks of the size 64MB. This data operates according to the key value pairs (appropriate program is written here) only the required data is selected from this rest and all is filtered out. This is the mapper phase maps the required data based on the programming model, this mapper function will run on all the fragmented blocks.

Yarn architecture of Hadoop

The reducer will combine the output of mapper and used for the further processing. This is very efficient way to process the cluster, where the same code can be applied to all the required and necessary clusters.
This is hadoop 2.0 or MapReduce 2, yarn is introduced instead of job tracker and task tracker and also slight variations in Name Node. This architecture is used to overcome the following challenges

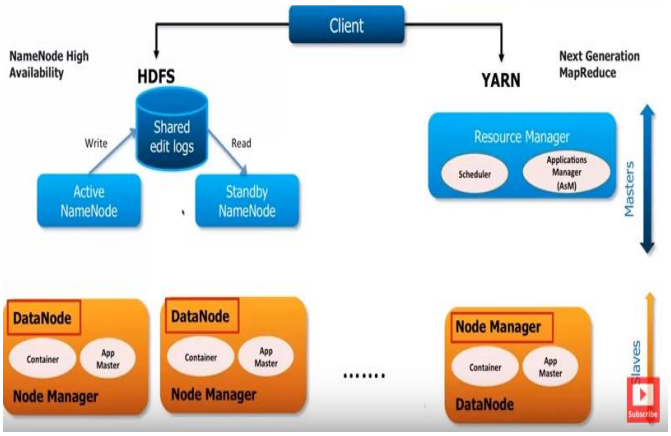
- Horizontal scalability of NameNode
- NameNode high availability
- Overburdened job tracker
- To run non-MapReduce bigdata applications on HDFS
- Support multi tenancy

Here, in HDFS active NameNode and standby NameNode are used to avoid single point of failure instead of NameNode. Active and standby NameNode share data with the help of shared edit logs which is network storage, active NameNode writes into shared edit logs and in case of failure standby NameNode takes over manually restart is not required.



Yarn contains resource manager instead of JobTracker which was responsible for scheduling and monitoring the jobs, also to distribute the resource based on the availability. Resource manager has two components: scheduler (responsible for allocating various resources to run the applications) performs allocation according to the constraints and application manager (for managing jobs) which communicates with node resource manager (negotiates the resource) and manager (assigns the given resources to nodes). Task tracker is replaced with node manager which is responsible to run any kind of jobs with the help of container.

With the help of Yarn (yet another resource negotiator) any kind of jobs can be implemented not only the batch processing, this is not the case with former hadoop 1.0 and also with the help of storm hadoop is used for real time processing.



Job scheduling in hadoop

Scheduling is allocating resources in a best possible way among multiple tasks. Since hadoop is a very large clustered environment many tasks will be parallelly executing scheduling should take care of allocating resources on the priority based and on what basis the priority should be given.[6] Scheduling happens in task tracker in case of

MapReduce 1 and resource manager in case of MapReduce 2, Few such scheduling algorithms are as follows:

FIFO Scheduler

Fist in first out scheduling, as the name indicates whoever comes in first priority is given to that job and placed in a queue. Sequentially the resources are used by the jobs which are in a queue, this scheduling is a default scheduler for MapReduce 1. [7]The job that is submitted first would take up all the resources and would be executed first, in this case of scheduling if a large job was submitted just before a small but a higher priority job a small job should wait for a long time unreasonably. This situation is little improved by deploying priority scheme in the queue, here jobs could be prioritized as very high, high, low, very low and normal. So this situation improved later where the high priority jobs moved high up the order but still in this case the preemption was not possible and so the smaller jobs want to wait for the long time if the longer process was already taken up and was under process. This scheme gives the clients unequal share of the cluster and random turnaround time.

Capacitive scheduler

This is a default scheduler that comes with map reduce2 or yarn setup this takes a little different stands to multi user scheduling. [8]In this case queues are divided on the basis of users or groups of users which is termed as organizations and there are multiple queues specific for the organization, each queue is given portion of resources in a cluster. Like if a new job enters into organization A, so it would be picked up as there is no job running. Moreover this would take up as many resources as much as available and cluster is effectively utilized. When job in organization B appears tasks of the first job will be killed to free up slots for the new job. [10] The main features of capacitive scheduling is capacity guarantees, elasticity, multi tenancy, hierarchical queues and security.

Fair Scheduler

Conceptually it is similar to capacitive scheduler with minor differences like the capacitive schedulers the queues are divided and termed as pools, so the jobs would be picked up from the pools and would be given their portions of the resources. [4]Suppose if another job comes into the pool then this is processed like First in first out or FIFO with priority, in this case a small high priority job should wait for a long time this is improved in fair scheduler that the jobs which was waiting the queue is picked up and processed parallel so as to give a better user experience.

Deadline constraint scheduler

By default the jobs are scheduled based on the priorities assigned to the job by the user. [4] But in deadline constraint scheduler user submits the job with deadline. Before allotting the job to the task tracker, job tracker computes the minimum number of slots required by the heartbeat message sent by TaskTracker to the JobTracker. If the minimum number of slots meet the requirement the jobs are allocated for processing else user gets a notification saying slots are not available. [9] So that the user can modify the deadline of job and can be resubmitted and the job can be scheduled. Here the jobs are not killed and freed the slots rather deadline is modified which doesn't affect the currently executing jobs. The throughput of the job is not taken into the picture and if the minimum number of slots is not available then the user should wait for a longer time.

Resource aware scheduler

Whenever the user submits job, job tracker checks whether the resources (CPU space, time, disk space etc.,) are available in the slots which is required to execute the jobs [8]. If the resources are available in the task tracker then the jobs are allotted else either the jobs are in waiting state until the resources are freed up or the jobs are fragmented depending on the required size and then jobs are allotted to each slots.

II. CONCLUSION AND FUTUREWORK

Hadoop framework is scalable upto 40,000 nodes, the problem arises when the cluster increases and all slots should be used effectively. So algorithms should be selected in such a way that output should be effective and high throughput should be produced. The default schedulers like FCFS, Fair, Capacitive scheduling suits good when we consider the normal jobs but the problem arises when all the slots are full and there is no free slot to allocate for a new job. In this case with the deadline constraint scheduling the required map reduce slots can be calculated and a fair share can be given through fair scheduling algorithm.

Table 1: Comparative study on all the algorithms

Properties	FCFS Scheduling	Capacitive Scheduling	Fair Scheduling	Deadline constraint Scheduling
When slots are free, if a new job enters	Effective, all the slots will be used solely by that job.	Effective, all the slots will be used solely by that job.	Jobs will get the slots which are allotted to it.	If the number of slots meet the requirement, then the job is allocated for processing
When the slots are not free and if a new job enters	Job should wait until the current execution of job is completed.	If the pool is not free, it checks with other pool and the job is allotted.	The job will get its fair share.	The job should wait until it meets the requirement constraint
Performance	Satisfactory	Effective	Satisfactory	Will be effective in some circumstances.
Complexity	Not complex	Not complex	Not complex	Complex and less throughput

References

[1] Mladen A. Vouk, "Cloud Computing – Issues, Research and Implementations" Journal of Computing and Information Technology - CIT 16, 2008, 4, 235–246.

[2] Alexandru Adrian TOLE, "Big Data Challenges" Database Systems Journal vol. IV, no. 3/2013.

[3] http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf

[4] B.Thirumala Rao, Dr. L.S.S.Reddy , "Survey on Improved Scheduling in Hadoop MapReduce in Cloud Environments" international Journal of Computer Applications (0975 – 8887) Volume 34– No.9, November 2011.

[5] Mr.A.U.Patil, Mr T.I Bagban, Mr.A.P.Pande, "Recent Job Scheduling Algorithms in Hadoop Cluster Environments: A Survey" International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015.

[6] Shekhar Gupta, Christian Fritz, Bob Price, Roger Hoover, and Johan de Kleer, "ThroughputScheduler : Learning to Schedule on Heterogeneous Hadoop Clusters" Palo Alto Research Center, Palo Alto, CA, USA.

[7] Yang XIA†, Lei WANG1, Qiang ZHAO, Gongxuan ZHANG, "Research on Job Scheduling Algorithm in Hadoop, " *Journal of Computational Information Systems 7: 16 (2011) 5769-5775.*

[8] C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.

[9] Matei Zaharia Dhruba Borthakur Joydeep Sen Sarma Khaled Elmeleegy Scott Shenker Ion Stoica, "Job Scheduling for Multi-User MapReduce Clusters," *†University of California, Berkeley Facebook Inc Yahoo! Research.*

[10] Job Scheduling in hadoop – an expose by joydeep sensarma, https://www.youtube.com/watch?v=fH7SYCpF-d Q.