IOSR Journal of Mobile Computing & Application (IOSR-JMCA)
e- ISSN: 2394-0050, P-ISSN: 2394-0042.Volume 7, Issue 1 (Jan. – Feb. 2020), PP 13-24
www.iosrjournals.org

Cloud Computing Resources Impacts on Heavy-Load Parallel Processing Approaches

ZainabSalih Ageed[1], Mayyadah R. Mahmood[2], Mohammed M. A. Sadeeq[3], Maiwan B. Abdulrazzaq[4], HiviIsmat Dino[5]

[1]*Computer Science & IT, college of Science, Nawroz University, Kurdistan region – Iraq*
[2]*Computer Science dept., college of Science/ University of Zakho, Kurdistan region – Iraq*
[3]*Duhok Polytechnic University presidency / Duhok Polytechnic, Kurdistan region – Iraq*
[4]*Computer Science dept., college of Science/ University of Zakho, Kurdistan region – Iraq*
[5]*Computer Science dept., college of Science/ University of Zakho, Kurdistan region – Iraq*

Abstract
One of the most important subject which many researchers depending on it by applying many algorithms and methods is Cloud Computing. Some of these methods were used to enhance performance, speed, and advantage of task level parallelism and some of these methods used to deal with big data and scheduling. Many others decrease the computation's quantity in the process of implementation; specially decrease the space of the memory. Parallel data processing is one of the common applications of infrastructure, which is classified as a service in cloud computing. The purpose of this paper is to review parallel processing in cloud. However, the results and methods are inconsistent; therefore, the scheduling concepts give easy method to use the resources and process the data in parallel and decreasing the overall implementation time of processing algorithms. Overall, this review give us and open new doors for using the suitable technique in parallel data processing filed. As a result our work show according to many factors which strategies is better.

Key Word*:Cloud Computing, Cloud Resources, Distributed Systems, Parallel Processing, Heavy-Load.*

I. Introduction

Currently, the cloud computing providing computational abilities for facility depend on prototypical therefore it is described as an effective paradigm. In addition, there are many other researches that managed with varied fog calculating exploration parts like: quality of service (QoS) and fault tolerance [1]–[5]. Cloud computing is a style described by (NIST) that is a model for permitting the resources like: providers, storages, stowage, networks, besides applications. Which may supplied besides unrestricted through smallest organization efforts before facility provider contact. It's also permitting suitable, ubiquitous, on demand right of entry network, Cloud calculating considered as a style/way to virtualized resources offered over the internet and it's dyna1mically scalable style [6]–[8].

Our needing for more processing with the continuously expanding interest has been amongst us and it has been developing from this point forward. The researchers turn parallel processing in cloud before 10 years ago since it has a unique nature and involved time during the development of cloud. In addition, its offer during processing a maximum guaranteed power and all its servers are hybrid in terms of computer science [9]–[11]. One big challenge for data analysis programming is during the processing of these data, the time of generating it should be faster. Addressing the issues of performance may investigate the development of classical techniques like scheduling, computing, data intensive, cloud computing [12]–[14].

II. Parallel Processing

The subject of parallel processing is correlated with algorithms strategies and structural for improving consistency, cost, performance and many other attributes of computer through many sorts of concurrency[15]–[17]. Since the computer become faster, it can be assume that they will become fast enough and the computing power enhancement will be sated. Since history suggests that as a particular technology satisfies known application, there will be many of application demand the enhancement and development of new technology, and as some additional technology area[18]–[20]. Equivalent besides programs structures investigation needs inspiration; Researchers have started working toward to use parallel processing to increase speed/performance of chips by a few variables [21]–[23].The reasons of depending this technology is shortening as follows [24], [25]:

a. Advanced quickness, or tackling problem quicker as critical while requests take soft/rigid limits. In this case more than one hour of time to calculate and continuously predicting besides yielding notices.

b. Greater quantity/taking care of difficulties more cases. Once it a significant frequent comparable assignments necessity to presented.

c. Control of higher computational, taking care of bigger difficulties besides capability of usage precise comprehensive, and along these lines extra truthful, prototypes before reproduction keeps running for longer periods of time like five days as opposed to daily predicting.

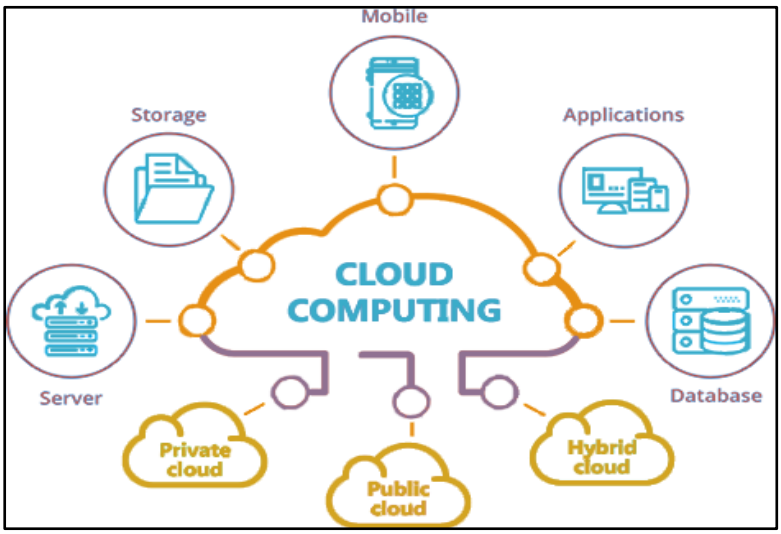
III. Cloud Computing

It can be defined as converting data besides procedures, where computation is done also how it is done, it is a transforming information technology and investigated in many fields. It is also gradually changed the working environment of IT professional and cracks various troubles for conservative calculation containing treatment highest works, fixing bring up-to-date of software besides by means of extra calculating series. In addition, cloud computing created new challenges like security, data ownership and trans-cods data storage [26]–[28].

The concept of "network as computer" proposed by Sun Computer in 1983, and this was the door for the intelligent/expansion track. The new era trendy computer development was by the executive chief of Google *Eric Schmidt* when he put forward the concept "cloud computing" and that was in 2006. After that by one year, Google and IBM work together with universities in USA and initiated to develop the services of could hardware and software technology on campus [29]–[33].

With cloud computing there is one fundamental principle provided which is no need to be anxiety about in what way the service you are purchasing is offered. With the services of web forget difficult for given that reliable calculating another person and you essentially focus on whatever your task is[34], [35]. We are depended toward consuming whole regulator for PC organizations besides comprehensive accountability, also cloud calculating variations. Deriving dual fundamental ideas, private besides public (i.e. cloud counterparts Intranets besides Internet, Online electronic message plus permitted facilities similar those Google gives which considered utmost acquainted cases of global one)[36]–[38]. Cloud contains of three fundamental segments (components), which are Client computers, Distributed Servers and Datacenters. Figure1. shows the components of cloud[39].

Figure 1. Cloud components proposed by[39]



IV. Cloud Service Models

There are three main services model in cloud depending on NIST definition, programming, PaaS, besides Organization[40]. The following 3 initial cloud service prototypes. However, one thing to memorize is that as we are dealing with service providers, almost everything is negotiable. Existing services are changed and new services are added to meet customer needs (SaaS, PaaS and IaaS)[41]–[45].

V. Literature Review

In 2017, Chen et al. [46], described an application contains more than one task with precedence constraints. The application was described using directed (DAG) and implemented on heterogeneous cloud computing frameworks. The requirements of services for cloud suppliers are one of the most important characteristics which is: reducing the length of the schedule during achieving the constraints of the budget of an

application. This study discover the undertaking of the re assignment with reducing the length of schedule and minimize the cost. The selected algorithm was MSLBL for spending plan parallel application. The study described in two phase, at the first the correctness was confirmed using proof and experiments by proposing an algorithm simply satisfy the limitation of the spending. Second, the complexity task scheduling with low time and proficient was implemented by proposing MSLBL algorithm to minimize the length of the schedule. The final result after applying the above steps approve that the proposed algorithm fulfilling the budget constraints with reducing as possible the length of the schedule on many real parallel applications.

In 2018, Li. [47], presented a method for planning besides analyzing empirical routine procedures which utilize equivalent rapidity technique. Two procedures produced in this study, post power and pre power determination for both time and vitality compelled arrangement numerous various logical processors for superiority constraints simultaneous jobs through unceasing rapidity stages. Therefore, embedding the ES strategy into algorithms was the fundamentals technique; this strategy makes the analysis and the performance of the algorithms good. Where 1- the algorithms developed appropriate for assortment surroundings containing: grid, cluster, many cores CPUs utilized. 2- Algorithm has rigorous recital guarantee and practical implication where the comparison of the performance had been with the optimal algorithms.

In 2017, Vijayakumar and Arun[48], assessed the applications online for vulnerabilities at regular intervals. The vulnerabilities checking tools will be triggered by Webhook if there is any changed made in the code included the software development lifecycle. This tools based on Hashing algorithm in the updated application. So, vulnerabilities' checking as incessant incorporation portion besides incessant positioning procedure are the main aims of this system that are deployed in the cloud plus scanning the application constantly.

In 2017,Nandra et al. [49], worked for processing big earth observation EO data task, compute intensive by providing a solution for scalable case and high performance. They focusing on offering the user algorithm up to date; also study expects to improve the general time of executing the processing algorithms. Their plan done by proposing a methodology with distributed execution environment to demonstrate of combining modular description; also facilitate the processing of EO data volumes. The description of the algorithm employee the representation of the workflow and offer a simple and easy way from of the user with no programming knowledge with a particular arrangement, easy of utilize and understand. This model enable the framework over its computing nodes to evenly distributed and significantly encourage the idea of parallelizable tasks. Capabilities and abilities was demonstrated by this study to take the advantages of the inherit task-level parallelism; also, reduce the time of execution.

In 2018, Cao et al. [50], depending on the MapReduce algorithm was the viewpoint of this study by focusing on exploring the processing of data of high speed railway fault signal diagnosis; also, improving the data flow of strategy of partitioning. For classifying and modeling data, they use Bias classification algorithm. The parallelization's process of MapReduce, implemented by the matrix of the data partition (Tk) had been saved in segmentation line. In every node of cluster, the load of computing was been distributed. In addition, calculating the consumption of partitioned matrix and time of mobile data matrix. Final stage showed that, the consumption of space of memory was reduced and the speed of counting in the railway signal system was been improved. Plus that, the proposed algorithm reduces the computations amount of execution processing.

In 2019, Haussmann et al. [51], they found that Implementing parallel environments by the user which give the ability to optimize the costs of monetary of parallel computation if used cloud computing as platform. They investigated how the total costs of computations can be impact by the scalability characteristics of parallel application. Also, their study provided how optimizing the cost of monetary of individual parallel cloud computations by providing an approach for facilitating data scale automatically. In addition, by focusing on the irregularly problems and setting precise group for expandability be influenced by contribution information. For minimizing the computation costs a dynamic optimization methods are required. They presented a charge prototypical reflects cost of simultaneous organization working plus cost after delay result. All of this just to measuring the total cost of monetary of individual parallel computations. As a result for this approach, they gives details into the performance characteristics and they discussed the total cost depends on cost model for no. of processors which was minimal.

In 2018, Silva et al.[52], deriving an application in HPC shelf by introducing framework SAFe. It's give the role to SWfMS clarify which applications can monitor besides organize the huge gage equivalent calculating during execution. Where SWfMS is the Scientific Workflow Management System. Deeper evaluation is possible by implementing HPC shelf application on top of SAFe and SAFeSWL workflows. The orchestration code of workflow is generated by the provider of application by using subset of SAFeSWL. Controlling the use of resources can explicitly done by controlling the deployment time and releasing of components, and normally applying a mechanism for study with nonfunctional concerns. To proof the concept validation of SAFe and SAFeSWL in this study was been applied by using two frameworks: MapReduce and Montage. The study includes too the orchestration and architecture of three workflows: Pleiades, M101 and

Word Counter. The main point applied in this study was: implementing the HPC shelf applications on top of SAFe and as mentioned above, possibility of deeper evaluation of SAFeSWL and SAFe.

In 2018, Ramantha and Latha[53], proposed a method to obtain and get the accurate prediction of job completions time which was the strategy of scalability of scale out methods. The performance level of MapReduce benchmark was shown through their results in the private cloud. The execution time of five common MapReduce benchmark applications was predicating and evaluating by the regression based performance model. The process was over their private cloud environment and got an accuracy result through well supply consumption that represents ninety nine percentage. The correctness of using the regression based performance model shown by running MapReduce jobs using strategy of scale out in private cloud.

In 2018, Kannadasan et al. [54], analyzed techniques and software of cloud system by comparing between them according to processing, nature of use friendliness and capacity of handling data. They used different implementation on large scale data analysis like HEP, Cloudburst, and Cap3. Their study found that CGL MapReduce and Hadoop when used with large data sc ale are better than dryad because of the mechanism storing of dryad. They also found that Hadoop is easy to using it n comparison and more user friendliness. Allowing faster transfer and better handling of data computation times and analyzing purpose can be improved with more advancement done is the mechanism of online storage.

In 2018, Zheng et al.[55], suggested a mechanism with parallel optimization which was multi ant colony parameter. The mechanism of correcting the parameter improves the stability and convergence of the ACO algorithm.

In 2019, Rashid et al. [56], proposed a system which remotely can solve the complicated problems by using mobiles or PCs as a client side. The proposed system contains unlimited no. of servers which can participate in an easy way within the process of solving the problem, the problem which the study can be depended on it with extended complexity requirement. The proposed system covered two republics on diverse spaces deprived of at complexity since planned depending on fundamentals of shared system. It can be extended to cover unlimited no. of countries. Accomplishment capability through large dispensation can be provided by addressing the real system and by focusing on steps of design's mechanism. Also, by the cloud domain, the ability of utilizing processing's power depending on distributed system's principles.

In 2019, Chen et al.[57], proposed an algorithm applied on Apache platform aimed at great data gage interval sequence via PPTSP. Corresponding optimization methods are considered because of many issued appeared such as data communicating waiting for synchronization and balancing the workload. As a result, the proposed algorithm if it's compared with the other algorithms achieved significant advantages depending on accuracy and performance of prediction. Keeping high scalability and enhancing the performance with low costs can effectively improve by using PPTSP algorithm.

In 2020, Tchendjiet al.[58], described new efficient parallel algorithm on CGM to solve the problem of none parallel solutions. They proposed an consecutive explanation of active software design method which needs O(nmr) running period. Avoiding redundancy because the overlap they proposed Multi DAG, in this step the evaluation of difficulties precise sequence will be determines. Then, they proposed two algorithms based on the DAG above which was CGM parallel algorithms. Now, the first algorithm requires O(nmr) implementation period through O(p) communiqué sequences. Because of dependencies in the DAG between the nodes, there is a problem found which the high of idle processors is so overtime with numerous shiftless CPUs. In other side, the reducing the idleness time by making the processor active as possible by using irregular partition of DAG. This requires O(nmrp) with O(kp) is communication round and O(nmrp) execution time. Hence, decreasing inactivity period and reduce time of communiqué. The result performed well covenant through hypothetical forecasts and reduces time execution foe second algorithm.

VI. Discussion

The summary of all previous literature review is presented in Table 1. As we noticed, there are many methods and algorithms used to provide and got the perfect results in a short time with reducing the execution time depending on which technique has been used. The comparison is done in terms of improving performance, speed, and advantage of task-level parallelism. Some of these methods used to deal with big data and scheduling.Chen et al. [46], proposed an active way to reducing the length of the schedule by using the level budget of MSLBL to choose the processor. There are two problems appeared here, the first one is pleasing the constraints of the budgets and the second problem is reducing the length of the schedule. Solving the first problem through moving constraints of financial plan for presentation to each application-task, while solving the second problem by depending on task's scheduling with low complexity in time in a heuristic manners for each task. The algorithm which [46]proposed may get shorter lengths of the schedule and at the same time pleasing the budget application's constraint than existing methods in different cases. So, the length of the schedule of their algorithm has is shorter than HBCS of financial plan limited equivalent submissions, which result point to that the MSLBL is more active in many.

Li. [47], investigated in the environment of cloud computing the issue of time and energy constrained scheduling on multiple many core processors. The following algorithms are designed with discrete or continuous level's speed: algorithms of pre-power resolve besides post-power resolve. Representation these algorithms is rated in an analytical way and experimental way. In addition, The task model and processor model are described too. While Vijayakumar and Arun[48], system was for checking if the vulnerability of the application was been added into the SDLC after each change or upgrade. The result of this step saved as XML document and this is make it easy to read and easy to be understand. The data of the xml document analysis which can be adopted or chosen for JSON for the future. Many tools can be used here such as Tensor Flow to apply scalability at the cloud interface to execute the load of processing. The overall performance improved whenever the entire problem from n gram approach with distributed algorithm of MapReduce. There are automation's importance in CI/CD pipeline, therefore the security and compliance deployed in to cloud. An enhancement ca be applied to this framework by adopting model like PUB/SUB or by implementing solution which make it more scalable. Nandra et al. [49], take the advantage of the level of the task parallelism by demonstrating the BigEarth platforms and this will lead to inheriting ability within a processing flow. The strength of the system puts in executing repetitive, independent and tasks of patching process. Something should be mentioned here which is batch processing tasks in the end characteristics to the stage of executing the huge globe reflection information groups. Cao et al. [50], used to classify and model the data "Bias Classification Algorithm" and the strategy of partitioning of data flow was improved. The project has a set running time test of Comparison between: A SON and SON algorithms in A workstation. The computation workload of single node can be reduces be the algorithm but sometimes the workload reduction is not clear because the data set are small. The execution time for this algorithm comparing to other algorithms is better especially for the huge data set, also it takes low space in the memory comparing to SON algorithm.

Haussmann et al. [51], study the determination of costs of monetary calculations that implemented in the parallel environment of cloud. Also, featuring pay per use billing. By consideration the problem which have high scalable is more productive than the problem with poor scalable in utilize additional processors. They are focusing in two things: high scaling which is meaning the size of problem still constant instead of increasing the no. of processors, and they focusing too on the input data which the highly scaling depends on it. The computation of the parallel can be in the following of calculation's phases: scaling/stable phase. Java based prototype was used to put their work into practice during implementation and architecture. Silva et al. [52]for enhancing improving capability, they accomplished the objective of HCP shelf of technologies meeting with its requirements related to parallel processing of large scale heterogeneous behind the software engineering of components, workflow management systems and cloud computing. Its prominent characteristics are 1- software's opinion uniform, hardware, many other types, 2- contextual contracts system with integration of HPC Shelf. 3- Controlling of explicitly of the time of deployment with all possibilities plus instantiation and components releasing and allowing the fine control of resource's using. 3- Studying non-functional concerns with extensible mechanisms.Ramantha and Latha[53], the researchers in this approach focusing on: the model of performance for foretelling the completion time of job, and the execution time evaluations of Hadoop MapReduce jobs in private cloud environment. They foretelling the time of execution by using regression based performance and they got accuracy after applying this model. They test the model n Hadoop cluster in OpenStack cloud environment, explored the accuracy performance and the utilization of effective resources of an approach and the model which was used in ARIA project was SLO. The model tested in cluster which was Hadoop in OpenStack cloud environment. The cloud workload prediction have used a module for SaaS applications which was ARIMA module. The intensive enterprise jobs of storage was decided by the network latency and the storage of the disk. The Hadoop Tera-Gen program creates the input data, and it's also by default generate a billion lines and the length of each line equal to a hundred bytes. The result of this study was helping the users of the private cloud to deploying before in their Hadoop MapReduce application. Kannadasan et al. [54], compared and analyzed cloud computing technologies: Hadoop, Dryad and CGL MapReduce. Then they adjusted the procedure for it by using three ways for map reducing: at the first, they take the data collected seeds. After that, they calculate alignment for each seed and used it aimed at calculation through different t. After the above step, they got the pit-aligned with finest conceivable and to be upsurge running-time spent in reduction phase by the algorithm which is greatly reduced. Then, they start calculating the alignment for any new seed with all the others seeds which will give better alignment. Finally the data sent to the reducer to be analyzed and it is cleared now that Dryad is slower than Hadoop.

Zheng et al. [55], the parameters of ant colony is a perfect way to enhance the ACO algorithm's performance. The researchers proposed easy strategy of equivalent finest of multi ant colony. According to the scale of the problem, the solution of the average stagnation optimal all the time basic pointer for the evaluation of ACO algorithm. Its run under each TSP instance 50 times, after that the solution is selected and got by its lag and average it. The proposed algorithm increases the opportunities of ant colony to getting more active parameter setting, and enhance the algorithm performance. Rashid et al. [56],established the important of the

study with new contribution could be made by identifying a place for it. The different methods were been evaluating by the bulk section which is used toward recognize suitable method aimed at inspecting study questions. They clarify solution by getting the benefits from the diagram modeling of unified language that displays the diverse facts for suggested structure vision. The plan of the system show the amount for creating product of work, also specified cost of the system. Chen et al. [57], implemented a PPTSP algorithm in cloud computing of apache spark for large scale times series, to effectively handle the massive historical data set. Therefore, to reducing the scale of data a TSDCA algorithm is presented and extracting the characteristics. Depending on this, the researchers proposed a (MTSPPR) using FSA method. In addition, a (PTSP) algorithm is proposed. The DStream and RDDs are used for calculating and saving these the datasets, where DStream is the distributed streams and RDDs is the resilient distributed datasets.

Table 1: Overview of Parallel Processing on Cloud Computing

Ref. / Year	The Problem Statement	Methodology	Strategy/ Algorithm	Machine properties	Tools	Significant Results
Chen et al. [46], 2017	Propose an efficient algorithm to fulfilling the budget restraint and reducing the schedule size of an application	DAG directed acyclic graph With combination of Gaussian and Fast Fourier transform elimination applications with budget deadline	(MSLBL) algorithm	The simulation of the experiments are conducted on a PC platform Intel Core i5 2.60 GHz CPU and4 GB memory. The simulation system contains 128 processors which was a heterogeneous system of cloud computing with different computing abilities and unit prices, where the types and the prices of processors are based on the Amazon Elastic Compute Cloud (EC2) environment	Execute a emulator in Java language	To minimize the length of the schedule MSLBL algorithm implements efficient and low time complexity task scheduling
Li. [47], 2018	-optimization problems- The difficulties of scheduling energy and time constrained of precedence constrained parallel jobs in a cloud computing environment with multiple many-core processors.	(ES) Equal Speed model with the processor model	Plan andstudy theperformanceo f heuristicalgorith msthat utilizethecorresp onding speedmethod	Not declared.	C++ programming language and runningin a Linux environment, algorithms had been compared with optimal algorithms //embed the ES strategy into algorithms	Algorithms developedapplicable to a varietyof environments,includi ng parallel,distributed and cloudcomputing wheremany-core processorsare used.It's haveGuarantee in performance andpractical implication
Vijayakumar and Arun[48], 2017	-vulnerable state after updating – The evaluation of vulnerability travelled into the cloud if the code has been changed or the application updated by the client.	Checking if the structure of the code has been changed or the code itself has an alteration or any upgrading is given for the application that is already traveled to the cloud for vulnerability after every update is made	Secure Socket Layer (SSL) approach Hashing algorithmLinear kernelPolynomi al kernelCosine algorithmSigmo id algorithm Laplacian kernel	Not declared.	WebhookRIPS, Brakeman, FindBugs, PMDLambda functions through web hooks and API Gateway PHP and Ruby on RailsHTML, XML	- monitor the applications that are employed in the cloud and will run a vulnerability Checking if the structure of the code has been changed or the code itself. - The entire processing will performed at the cloud by using tolls like tensor flow for scalability purpose.

Nandra et al. [49], 2017	Enlarge ofData availabilitygives rise tonewchances of examiningand gainingknowledg efrom publiclyavailable datastores	Proposed todemonstrateof joiningmodularpr ocessingexplanati on methodology with adistributedexecu tionenvironment	EO-Earth Observationdat a tasks	Using 1.31 GBOf storage space.	MSAVIand NDVI, imagefiles(a near-infrared andred band),dimensions 40,980 x30,978 pixels, Big Earth platform configured in private cloud	take advantage ofthe inherent task-levelparallelism Capabilities andskills of systemto improve the time of task execution andscalable, high performance solution for the processing and compute intensive, big EO data tasks
Cao et al. [50], 2018	Data flows are categorized by high speed, continuity and boundlessness. If there are also many dissimilar elements or condition for processing many data flows at once, it is difficult to store All information of data flows in memory.	A parallelprocessing algorithm is projected for fault diagnosis of railway Signal systems	the Bias classification algorithm is used to model and categorize the data and MapReduce separating strategy of data flow	The hardware device is: 6 PC machines of win7 systems, Master nodes use 16 GB memory and Intel i7 processor, and Slave nodes use 3.6 GHz main frequency, 4 GB memory and Intel i7 processor.	The data flow is conducted with batch processing Using: MapReduce the Hadoop 2.0.4 stable version is adopted in the test. The data test take the Chengdu Railway Bureau fault database as the chief data source, plus the faults of existing railway signal system and high-speed railways monitored by CSM	The test of dissimilar data sets displaysthat the algorithm projected is superior to the overall algorithmin the statistical computation of the frequent item sets of big dataflow.Mean:Migh t reduce the quantity of computation in the execution process, greatly decrease the memory space consumption, and enhance the counting speed in railway signal system.
Haussmann et al. [51], 2019	Erratically structured troubles, where the scalability obviously depends on the input data and quantifying the financial costs of calculations	Execute an extensive experimental calculation by using exemplary problems of the presented concepts that exhibit a high degree of erratically quantifying the financial costs of computations, executed in cloud-based parallel environments featuring pay-per-use billing.	Fulfilled a Java based prototype of a distributed task pool and the optimization method Employing two OpenStack IaaS elements for the implementation : The monitoring service (called Ceilometer) and the compute service (called Nova). Ceilometer.	The underlying hardware consists of identically configured servers, each equipped with two Intel Xeon E5-2650v2 CPUs and all CPUs have eight cores operating at 2.6 GHz. 128 GB as main memory.The coordinator is executed on a VM with 2 vCPUs and 4 GB main memory with 1 vCPU and 2 GB main memory for hosting workers	by several network the servers are linked by separating management, storage, and tenant VM communicationUtilizi ng identical virtual machines.By a self-service virtual network all the VMs are linked	The approach lets cost driven auto-scaling for parallel cloud calculations that are considered by an unidentified scalability performance

Silva et al. [52], 2018	Addressing the challenge of dealing with heterogeneous software and hardware resources in a large-scale HPC infrastructure comprising many parallel computing platforms	Handling workflows parallel computing systems as a component kind, making probable to reuse workflows across applications, instead of producing them dynamically.	Using the architectural subset of SAFeSWLMontage and Map Reduce frameworksMontage which have been used for proof-of-concept validation of SAFe and SAFeSWL	Not declared.	XSD grammar that represents the concrete syntax of the orchestration subset of SAFeSWLITaskPort interfaceXML-based architectural file two designs of HPC Shelf applications are presented for demonstrating and evaluating the ideas behind SAFe:Montage and MapReduce	Monitoring the time of deployment, instantiation and releasing of components, allowing the fine control of the use of resources; and a naturally extensible mechanism for dealing with nonfunctional concerns
Ramantha and Latha[53], 2018	The processing of large amounts of data analysis	Performance model for predicting job completion time and the Hadoop MapReduce jobs execution time evaluations in the private cloud environment	Hadoop/ MapReduce The cloud workload prediction has been used ARIMA Module for SaaS / linear regression method and Scale-Out Strategy	Memory proficiency of the OpenStack private cloud Hadoop cluster and to sort 1 TB of data.	Discover performanceaccuracy and effective resource utilization of an alternative bounds-based approach and SLO based model used in ARIA project. 1. Word Count Application is used as a Hadoop MapReduce job processing of input dataset. 2. RandomWriter program used to create the unsorted random sequence of words which generates 10–100 GB size of datasets.3.The TeraSort application is used to test the CPU	Predicting and evaluating the based performance regression model The accuracy results of executing time of 5 common MapReduce benchmark applications over private cloud environment with well resource utilization show 99%.
Kannadasan et al. [54], 2018	Need to store big quantity of data and computations and need to store huge experimental results of the experiment and assess the result	Comparing and analyzing cloud computing technologies Hadoop, Dryad and CGL MapReduce	Cloud-Burst algorithm	Not declared.	- CAP3 as a DNA based Sequence assembly program-apply a function that is programmed in ROOT script to all of the input files-use a homomorphic filter-Message Passing Interface -Root method	The execution time of Hadoop in most application oriented work is quicker than the other techniques Hadoop offered more user friendliness and is easy to use in comparison to other techniques present
Zheng et al. [55], 2018	Lack of performance of traditional ACO algorithm and analyze parameter sensitivity of MMAS algorithm	Suggest an algorithm on dynamic parameter adaptation strategy which is multi-ant colony parallel optimization built,	Ant-Colony optimization algorithm	The computer hardware used: Intel Xeon E3-1230 v2 (8 Threads)and 8G ram.	Parallel mechanism shared memory parallel programming OpenMP joined with VS 2010 to achieve multiple ant colony parallel execution and information exchange The OS was Win7 64bit ultimate..	Increase the chances of the ant colony gaining a more adaptive parameter setting,and increase/ enhance the performance of the algorithm.The performance of the algorithm's real time has been significantly recovered.

Rashid et al. [56],2019	Access huge computing power remotely from light processing devices combine and applying distributed parallel processing via distributed cloud calculating that can solve some specific and huge client problems which needs uncommon processing power and solving them in a smallest time	The system is totally automated, it means the servers can register their selves mechanically on a central server that works like a middle link between all servers and clients. Then the clients can get benefit from the provided processing power by the registered servers. The clients can identify tasks to be processed by the servers, they also be capable to specify number of servers to participate in data processing	The Unified Modeling Language (UML) Diagrams, which displays dissimilar points of opinion of the proposed system. The following diagrams are designed: Class diagram, data-flow diagram, use-case diagram, sequence diagram, and architecture	Install some servers in different countries	Web server database system at cloud-side, adding to VC# tool at both Client-side and Server-side	The client gets results a detailed report about the amount of time and CPUs usages of participated server
Chen et al. [57], 2019	Critical to professionally recognize the potential periodic patterns from massive time-series data and offer exact predictions	a Periodicity-based Parallel Time Series Prediction (PPTSP) algorithm for large-scale time-series data is proposed and executed in the Apache Spark cloud computing environment	TSDCA, MTSPPR, and PTSP algorithms using the Streaming real-time computing module	Not declared.	The parallel solution of the TSDCA, MTSPPR, and PTSP algorithms were implemented on the Apache Spark cloud platform	PPTSP algorithm accomplished significant advantages compared with other algorithms in terms of exactness and performance of prediction. The PPTSP algorithm can successfully enhance the performance and keep high scalability and low data communication costs.

Tchendji et al.[58], 2020	There is no parallel solutions for dynamic programming technique	Proposed a multi-level Direct Acyclic Graph (DAG) that determines the correct evaluation order of sub-problems in order to avoid redundancy due to overlap	Describes new efficient parallel algorithms on Coarse Grained Multicomputer model (CGM) to solve the problem	48 × 128GB of RAM with thin nodes, and 12 named thick nodes with 12 × 512GB of RAM	The C programming language is used, on the operating system CentOS Linux release 7.6.1810. The inter-processor communication is implemented with the MPI library(OpenMPI version). The algorithms implemented on the cluster dolphin of the MATRICS platform of the University of Picardie Jules Verne using 60 computation nodes (48 nodes called thin nodes) with 48 × 128GB of RAM, and 12 named thick nodes with 12 × 512GB of RAM). Each node is made of two Intel Xeon Processor E5-2680 V4 (35M Cache, 2.40 GHz), each of them consists of 14 cores. All nodes are interconnected with 630 Omni-Path links providing 100Gbps throughput.	The experimental results show a good agreement with theoretical predictions. The progressive reduction of size of the blocks allows processors to stay active as long as possible. This decreases their latency time and then minimizes the overall communication time. All this reduces the overall execution time of the second algorithm

VII. Conclusion

In this paper, a review of literatures on parallel processing in cloud environment is presented. As appeared after reviewing that each paper has specific approaches and various tools and measures are available for each approach. Some algorithms used for reducing time of request and others techniques used to improve speed and performance of data. We have systematically analyzed about thirteen previous works on Parallel Processing in cloud computing. And as it can be seen from our view that the implementation of Chen et al. [16] for reducing the length of the schedule of an application by proposing an active algorithm to pleasing the budget constraints offered low time and efficiencies in task scheduling. And another high result got by Ramanthan and Latha [23] which performance model had been used for calculating job completion time. The execution time of five common MapReduce benchmark applications was predicating and calculating by the regression based performance model. The process was over their environment which is private cloud with better resource utilization which depicts 99% of accuracy result. Another previous work Kannadasan et al. [24] found that Hadoop offered more user friendliness and in comparisons with the other techniques it's easier. In the mechanism of online storage, many things can be improved with more advancement like allowing better handling and the time of transferring data of computation and analyzing purpose. Allowing better handling and faster transfer time of the data for computation and analyzing purpose can be improved with more advancement done is the online storage mechanism. Besides the previous result, from our view we found that Zryan et al. [26] proposed a novel system related to remote code-breaking parallel processing approach via cloud computation implementation. The system of [26] provides unlimited number of participants at client- and server-side which provides high processing power to solve complex problem containing heavy loads.

References

[1] J. Lim, T. Suh, J. Gil, and H. Yu, "Scalable and leaderless Byzantine consensus in cloud computing environments," *Information Systems Frontiers*, vol. 16, no. 1, pp. 19–34, 2014.

[2] H. Choi, J. Lim, H. Yu, and E. Lee, "Task classification based energy-aware consolidation in clouds," *Scientific programming*, vol. 2016, 2016.

[3] S. R. Zeebaree, R. R. Zebari, and K. Jacksi, "Performance analysis of IIS10.0 and Apache2 Cluster-based Web Servers under SYN DDoS Attack," *TEST Engineering & Management*, vol. 83, no. March-April 2020, pp. 5854–5863, 2020.

[4] S. R. Zeebaree, R. R. Zebari, K. Jacksi, and D. A. Hasan, "Security Approaches For Integrated Enterprise Systems Performance: A Review," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 8, no. 12, Dec. 2019.

[5] S. R. M. Zeebaree, H. M. Shukur, L. M. Haji, R. R. Zebari, K. Jacksi, and S. M.Abas, "Characteristics and Analysis of Hadoop Distributed Systems," *Technology Reports of Kansai University*, vol. 62, no. 4, pp. 1555–1564, Apr. 2020.

[6] H. H. Song, "Testing and Evaluation System for Cloud Computing Information Security Products," *Procedia Computer Science*, vol. 166, pp. 84–87, 2020.

[7] S. R. Zeebaree, K. Jacksi, and R. R. Zebari, "Impact analysis of SYN flood DDoS attack on HAProxy and NLB cluster-based web servers," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 1, pp. 510–517, 2020.

[8] Ahmed, Omar M., and Amira B. Sallow. "Android security: a review." Academic Journal of Nawroz University 6, no. 3 (2017): 135-140.

[9] National Academies of Sciences and Medicine, *Assessing and responding to the growth of computer science undergraduate enrollments*. National Academies Press, 2018.

[10] S. R. M. Zeebaree *et al.*, "Multicomputer Multicore System Influence on Maximum Multi-Processes Execution Time," *TEST Engineering & Management*, vol. 83, no. May-June 2020, pp. 14921–14931, 24May 2020.

[11] R. Zebari, S. Zeebaree, K. Jacksi, and H. Shukur, "E-Business Requirements for Flexibility and Implementation Enterprise System: A Review," *International Journal of Scientific & Technology Research*, vol. 8, pp. 655–660, Nov. 2019.

[12] F. Magoules, J. Pan, and F. Teng, *Cloud computing: Data-intensive computing and scheduling*. CRC press, 2012.

[13] R. R. Zebari, S. R. Zeebaree, and K. Jacksi, "Impact Analysis of HTTP and SYN Flood DDoS Attacks on Apache 2 and IIS 10.0 Web Servers," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, 2018, pp. 156–161.

[14] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 56–70, May 2020, doi: 10.38094/jastt1224.

[15] B. R. Ibrahim, S. R. Zeebaree, and B. K. Hussan, "Performance Measurement for Distributed Systems using 2TA and 3TA based on OPNET Principles," *Science Journal of University of Zakho*, vol. 7, no. 2, pp. 65–69, 2019.

[16] S. R. Zeebaree and H. Rajab, "Design and Implement a Proposed Multi-Sources to Multi-Destinations Broadcasting Video-Signals," in *2019 4th Scientific International Conference Najaf (SICN)*, 2019, pp. 103–108.

[17] S. R. M. Zeebaree, K. H. Sharif, and R. M. M. Amin, "Application Layer Distributed Denial of Service Attacks Defense Technique: A Reiview," *Academic Journal of Nawroz University (AJNU)*, vol. 7, no. 4, pp. 113–117, 2018.

[18] A. K. Ibrahim, M. H. Abdulwahab, M. B. Abdulrazzaq, and M. R. Mahmood, "A Tree Method for Managing Documents in Mongodb," 2020.

[19] S. Q. Sabri, A. M. Ahmad, and M. B. Abdulrazaq, "Design and Implementation of Student and Alumni Web Portal," *Science Journal of University of Zakho*, vol. 5, no. 3, pp. 272–277, 2017.

[20] H. I. Dino and M. B. Abdulrazzaq, "Facial Expression Classification Based on SVM, KNN and MLP Classifiers," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 70–75.

[21] M. Gong, L. Pan, T. Song, and G. Zhang, *Bio-inspired Computing–Theories and Applications: 11th International Conference, BIC-TA 2016, Xi'an, China, October 28-30, 2016, Revised Selected Papers*, vol. 681. Springer, 2017.

[22] O. H. Jader, S. R. Zeebaree, and R. R. Zebari, "A State Of Art Survey For Web Server Performance Measurement And Load Balancing Mechanisms," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 8, no. 12, pp. 535–543, Dec. 2019.

[23] Ahmed, Omar M., and Wafaa M. Abduallah. "A Review on Recent Steganography Techniques in Cloud Computing." Academic Journal of Nawroz University 6, no. 3 (2017): 106-111.

[24] B. Parhami, *Introduction to parallel processing: algorithms and architectures*. Springer Science & Business Media, 2006.

[25] O. Alzakholi, L. Haji, H. Shukur, R. Zebari, S. Abas, and M. Sadeeq, "Comparison Among Cloud Technologies and Cloud Performance," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 40–47, Apr. 2020, doi: 10.38094/jastt1219.

[26] R. Arora, A. Parashar, and C. C. I. Transforming, "Secure user data in cloud computing using encryption algorithms," *International journal of engineering research and applications*, vol. 3, no. 4, pp. 1922–1926, 2013.

[27] A.-Z. S. R. Zeebaree, A. Z. Adel, K. Jacksi, and A. Selamat, "Designing an ontology of E-learning system for duhok polytechnic university using protégé OWL tool," *J. Adv. Res. Dyn. Control Syst.*, *vol*, vol. 11, pp. 24–37.

[28] L. M. Haji, S. R. Zeebaree, K. Jacksi, and D. Q. Zeebaree, "A State of Art Survey for OS Performance Improvement," *Science Journal of University of Zakho*, vol. 6, no. 3, pp. 118–123, 2018.

[29] C. W. Hung, "Cloud computing," in *Cloud Computing-Technology and Practices*, IntechOpen, 2018.

[30] A. AL-Zebari, S. R. Zeebaree, K. Jacksi, and A. Selamat, "ELMS–DPU Ontology Visualization with Protégé VOWL and Web VOWL," *Journal of Advanced Research in Dynamic and Control Systems*, vol. 11, pp. 478–85, 2019.

[31] Z. N. Rashid, S. R. Zebari, K. H. Sharif, and K. Jacksi, "Distributed Cloud Computing and Distributed Parallel Computing: A Review," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, 2018, pp. 167–172.

[32] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, "Enhance the Mammogram Images for Both Segmentation and Feature Extraction Using Wavelet Transform," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 100–105.

[33] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, "Multi-Level of DNA Encryption Technique Based on DNA Arithmetic and Biological Operations," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, 2018, pp. 312–317.

[34] Z. N. Rashid, K. H. Sharif, and S. Zeebaree, "Client/Servers Clustering Effects on CPU Execution-Time, CPU Usage and CPU Idle Depending on Activities of Parallel-Processing-Technique Operations "," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 7, no. 8, pp. 106–111, 2018.

[35] K. Jacksi, S. Zeebaree, and N. Dimililer, "Design and Implementation of LOD Explorer: A LOD Exploration and Visualization Model," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 31–39, 2020.

[36] N. O. Y. Subhi R. M. Zebari, "Effects of Parallel Processing Implementation on Balanced Load-Division Depending on Distributed Memory Systems," *J. of university of Anbar for pure science*, vol. 5, no. 3, 2011.

[37] A. M. Abdulazeez and S. R. Zeebaree, "Design and Implementation of Electronic Learning System for Duhok Polytechnic University," *Academic Journal of Nawroz University*, vol. 7, no. 3, pp. 249–258, 2018.

[38] A. M. Abdulazeez, S. R. Zeebaree, and M. A. Sadeeq, "Design and Implementation of Electronic Student Affairs System," *Academic Journal of Nawroz University*, vol. 7, no. 3, pp. 66–73, 2018.

[39] D. C. Marinescu, *Cloud computing: theory and practice*. Morgan Kaufmann, 2017.

[40] M. R. Mahmood and A. M. Abdulazeez, "Different Model for Hand Gesture Recognition with a Novel Line Feature Extraction," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 52–57.

[41] D. Rountree and I. Castrillo, *The basics of cloud computing: Understanding the fundamentals of cloud computing in theory and practice*. Newnes, 2013.

[42] A. S. Y. Subhi Rafeeq Mohammed Zebari, "Improved Approach for Unbalanced Load-Division Operations Implementation on Hybrid Parallel Processing Systems," *Journal of University of Zakho*, vol. 1, no. (A) No.2, pp. Pp832-848, 2013.
[43] S. R. Zeebaree, H. M. Shukur, and B. K. Hussan, "Human resource management systems for enterprise organizations: A review," *Periodicals of Engineering and Natural Sciences*, vol. 7, no. 2, pp. 660–669, 2019.
[44] M. A. Mohammed *et al.*, "An anti-spam detection model for emails of multi-natural language," *Journal of Southwest Jiaotong University*, vol. 54, no. 3, 2019.
[45] S. R. Zeebaree, A. B. Sallow, B. K. Hussan, and S. M. Ali, "Design and Simulation of High-Speed Parallel/Sequential Simplified DES Code Breaking Based on FPGA," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 76–81.
[46] W. Chen, G. Xie, R. Li, Y. Bai, C. Fan, and K. Li, "Efficient task scheduling for budget constrained parallel applications on heterogeneous cloud computing systems," *Future Generation Computer Systems*, vol. 74, pp. 1–11, 2017.
[47] K. Li, "Scheduling parallel tasks with energy and time constraints on multiple manycore processors in a cloud computing environment," *Future generation computer systems*, vol. 82, pp. 591–605, 2018.
[48] K. Vijayakumar and C. Arun, "Continuous security assessment of cloud based applications using distributed hashing algorithm in SDLC," *Cluster Computing*, vol. 22, no. 5, pp. 10789–10800, 2019.
[49] C. Nandra, V. Bacu, and D. Gorgan, "Parallel Earth Data Tasks Processing on a Distributed Cloud Based Computing Architecture," in *2017 21st International Conference on Control Systems and Computer Science (CSCS)*, 2017, pp. 677–684.
[50] Y. Cao, P. Li, and Y. Zhang, "Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing," *Future Generation Computer Systems*, vol. 88, pp. 279–283, 2018.
[51] J. Haussmann, W. Blochinger, and W. Kuechlin, "Cost-efficient parallel processing of irregularly structured problems in cloud computing environments," *Cluster Computing*, vol. 22, no. 3, pp. 887–909, 2019.
[52] J. de Carvalho Silva, A. B. de Oliveira Dantas, and F. H. de Carvalho Junior, "A Scientific Workflow Management System for orchestration of parallel components in a cloud of large-scale parallel processing services," *Science of Computer Programming*, vol. 173, pp. 95–127, 2019.
[53] R. Ramanathan and B. Latha, "Towards optimal resource provisioning for Hadoop-MapReduce jobs using scale-out strategy and its performance analysis in private cloud environment," *Cluster Computing*, vol. 22, no. 6, pp. 14061–14071, 2019.
[54] R. Kannadasan, N. Prabakaran, P. Boominathan, A. Krishnamoorthy, K. Naresh, and G. Sivashanmugam, "High performance parallel computing with cloud technologies," *Procedia computer science*, vol. 132, pp. 518–524, 2018.
[55] Y. Zheng, Q. Yang, L. Jin, and L. He, "A Parallel Pre-schedule Max-Min Ant System," in *International Conference on Cloud Computing and Security*, 2018, pp. 183–193.
[56] Z. N. Rashid, S. R. Zeebaree, and A. Shengul, "Design and Analysis of Proposed Remote Controlling Distributed Parallel Computing System Over the Cloud," presented at the 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 118–123.
[57] J. Chen, K. Li, H. Rong, K. Bilal, K. Li, and S. Y. Philip, "A periodicity-based parallel time series prediction algorithm in cloud computing environments," *Information Sciences*, vol. 496, pp. 506–537, 2019.
[58] V. K. Tchendji, A. N. Ngomade, J. L. Zeutouo, and J. F. Myoupo, "Efficient CGM-based parallel algorithms for the longest common subsequence problem with multiple substring-exclusion constraints," *Parallel Computing*, vol. 91, p. 102598, 2020.