Distributed Computing in Big Data Frame Work : A Review

M. Bhargavikrishna
Research scholar
DeptofCSE
Spmvv,Tirupati
bhargavimandara@gmail.com

Prof. S. Jyothi
Dept of Computer science
Spmvv,Tirupati
jyothi.spmvv@gmail.com

ABSTRACT
Big data analytics has attracted close attention from both industry and academic because of i ts great benefits in cost reduction and better decision making. As the fast growth of various global services, there is an increasing need for big data analytics across multiple data centers (DCs) located in different countries or regions. The data processing platform optimized for t he geo-distributed computing environment.It supports a cross Dc data processing platform optimisiedgeodistributed computing environment.To evaluate the performance the extensive simulations using real traces generated by a set of queries on Hive. The results show that pro posal can reduce 55% interDC traffic compared with centralized processing by aggregating a ll data to a single data center. Public distributed computing is a type of distributed computin g in which socalled volunteers provide computing resources to projects. It introduces a distr ibuted and selforganizing algorithm to build a management systems like health care .Resear ch show that public distributed computing has the required potential and capabilities to handl e big data mining tasks .It provides the foundation for future research required to bring back attention to this lowcost public distributed computing method and make it a suitable platfor m for big data and data mining analysis.

Keywords --
Big data analytics , geodistributed, Distributed algorithm,Self Organization,Data mining

INTRODUCTION

Many companies and organizations in today's world are interested in gathering data making tasks. It leads companies to capturing, storing and processing huge data sets, which i n turn refers to a term called big data mining. Over time data sets become large and connect ed to many data points making data difficult to store and process. Big data mining is a comp utational process of discovering patterns in large data sets. It takes use of methods at the inte rsection of artificial intelligence, machine learning, statistics and database systems. Unfortun ately, internal highperformance computing environments or similar traditional data managem ent solutions are no longer capable of handling such amounts of data . Organizations do not usually have enough internal computational resources to satisfy the demand. More and more companies start to provide global services by deploying data centers (DCs) in different count ies and regions. For example, Google runs its service across several geodistributed data cent ers connected by a dedicatedWAN. Other companies, e.g., Netflix, deploy their services at A mazon's global cloud infrastructure EC2 that spreads across 11 regions over the world. Thes

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

e companies conduct big data analytics across the geodistributed computing and storage envi ronment for risk evaluation, cost reduction, and new product creation. To deal with geo-distributed big data analytics, several recent efforts have been made to create a virtual cluste r across multiple DCs for big data processing. For example, Mandal et al have implemented and evaluated a Hadoop cluster across multiple clouds. Iridium has been proposed for low la tency queries on geo-distributed big data.

BIG DATA

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

Fig 1 : Ten Big Data Applications in Real Life

Ref link : https://images.app.goo.gl/QLxkM7gwFj963wus6

Sources of Big Data

These data come from many sources like

- **Social networking sites:** Facebook, Google, LinkedIn all these sites generates huge amount of data on a day to day basis as they have billions of users worldwide.
- **E-commerce site:** Sites like Amazon, Flipkart, Alibaba generates huge amount of logs from w hich users buying trends can be traced.
- **Weather Station:** All the weather station and satellite gives very huge data which are stored and manipulated to forecast weather.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

- **Telecom company:** Telecom giants like Airtel, Vodafone study the user trends and accordin gly publish their plans and for this they store the data of its million users.
- **Share Market:** Stock exchange across the world generates huge amount of data through its daily transaction.

Big Data has played a pivotal role in the business environment today. We can understand t his by looking at the aspects enlisted below,

- Cost Savings: Some tools of Big Data like Hadoop and Cloud-Based Analytics convey cost favorable circumstances to businesses when a lot of informati on is to be put away and these tools additionally help in distinguishing more proficient met hods for working together.
- Time Reduction: The speedy nature of tools like Hadoop and in-memory analytics can undoubtedly recognize new sources of data which helps organizatio ns in breaking down information instantly and identifying the most suitable decision.
- Comprehend the economic situations: Dissecting Big Data gives a clearer picture of the cu rrent economic scenario. For instance, by breaking down clients' buying practices, an orga nization can discover the items that are sold the most and deliver items as per this pattern. By this, it can stretch out beyond its rivals.

THE HISTORY OF BIG DATA

Although the concept of big data itself is relatively new, the origins of large data sets go bac k to the 1960s and '70s when the world of data was just getting started with the first data ce nters and the development of the relational database.

Around 2005, people began to realize just how much data users generated through Facebook , YouTube, and other online services. Hadoop (an open-source framework created specifically to store and analyze big data sets) was developed that same year. NoSQL also began to gain popularity during this time.

The development of open-source frameworks, such as Hadoop (and more recently, Spark) was essential for the growth of big data because they make big data easier to work with and cheaper to store. In the years since then, the volume of big data has skyrocketed. Users are still generating huge amounts of data—but it's not just humans who are doing it.

With the advent of the Internet of Things (IoT), more objects and devices are connected to t he internet, gathering data on customer usage patterns and product performance. The emerge nce of machine learning has produced still more data.

While big data has come far, its usefulness is only just beginning. Cloud computing has expa nded big data possibilities even further. The cloud offers truly elastic scalability, where deve lopers can simply spin up ad hoc clusters to test a subset of data. And graph databases are be coming increasingly important as well, with their ability to display massive amounts of data in a way that makes analytics fast and comprehensive.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

BIG DATA CHALLENGES

While big data holds a lot of promise, it is not without its challenges.

First, big data is…big. Although new technologies have been developed for data storage, data volumes are doubling in size about every two years. Organizations still struggle to keep pace with their data and find ways to effectively store it.

But it's not enough to just store the data. Data must be used to be valuable and that depends on curation. Clean data, or data that's relevant to the client and organized in a way that enables meaningful analysis, requires a lot of work. Data scientists spend 50 to 80 percent of their time curating and preparing data before it can actually be used.

Finally, big data technology is changing at a rapid pace. A few years ago, Apache Hadoop was the popular technology used to handle big data. Then Apache Spark was introduced in 2014. Today, a combination of the two frameworks appears to be the best approach. Keeping up with big data technology is an ongoing challenge.

Big data best practices

Here are some guidelines for building a successful big data foundation.

Align big data with specific business goals

More extensive data sets enable you to make new discoveries. To that end, it is important to base new investments in skills, organization, or infrastructure with a strong business-driven context to guarantee ongoing project investments and funding. To determine if you are on the right track, ask how big data supports and enables your top business and IT priorities. Examples include understanding how to filter web logs to understand ecommerce behavior, deriving sentiment from social media and customer support interactions, and understanding statistical correlation methods and their relevance for customer, product, manufacturing, and engineering data.

Ease skills shortage with standards and governance

One of the biggest obstacles to benefiting from your investment in big data is a skills shortage. You can mitigate this risk by ensuring that big data technologies, considerations, and decisions are added to your IT governance program. Standardizing your approach will allow you to manage costs and leverage resources. Organizations implementing big data solutions and strategies should assess their skill requirements early and often and should proactively identify any potential skill gaps. These can be addressed by training/crosstraining existing resources, hiring new resources, and leveraging consulting firms.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Optimize knowledge transfer with a center of excellence

Use a center of excellence approach to share knowledge, control oversight, and manage proj ect communications. Whether big data is a new or expanding investment, the soft and hard c osts can be shared across the enterprise. Leveraging this approach can help increase big data capabilities and overall information architecture maturity in a more structured and systematic way.

Top payoff is aligning unstructured with structured data

It is certainly valuable to analyze big data on its own. But you can bring even greater busine ss insights by connecting and integrating low density big data with the structured data you ar e already using today.

Whether you are capturing customer, product, equipment, or environmental big data, the goa l is to add more relevant data points to your core master and analytical summaries, leading t o better conclusions. For example, there is a difference in distinguishing all customer sentim ent from that of only your best customers. Which is why many see big data as an integral ext ension of their existing business intelligence capabilities, data warehousing platform, and inf ormation architecture.

Plan your discovery lab for performance

Discovering meaning in your data is not always straightforward. Sometimes we don’t even k now what we’re looking for. That’s expected. Management and IT needs to support this “lac k of direction” or “lack of clear requirement.”At the same time, it’s important for analysts an d data scientists to work closely with the business to understand key business knowledge gap s and requirements. To accommodate the interactive exploration of data and the experimenta tion of statistical algorithms, you need highperformance work areas. Be sure that sandbox en vironments have the support they need—and are properly governed.

Align with the cloud operating model

Big data processes and users require access to a broad array of resources for both iterative ex perimentation and running production jobs. A big data solution includes all data realms inclu ding transactions, master data, reference data, and summarized data. Analytical sandboxes sh ould be created on demand. Resource management is critical to ensure control of the entire a data flow including pre- and post-
processing, integration,indatabase summarization, and analytical modeling.

A wellplanned private and public cloud provisioning and security strategy plays an integral r ole in supporting these changing requirements.

Use case

An e-
commerce site XYZ (having 100 million users) wants to offer a gift voucher of 100$ to its to

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

p 10 customers who have spent the most in the previous year.Moreover, they want to find th e buying trend of these customers so that company can suggest more items related to them.

Issues

Huge amount of unstructured data which needs to be stored, processed and analyzed.

Solution

Storage: This huge amount of data, Hadoop uses HDFS (Hadoop Distributed File System) which uses commodity hardware to form clusters and store data in a distributed fashion. It w orks on Write once, read many times principle.

Processing: Map Reduce paradigm is applied to data distributed over network to find the r equired output.

Analyze: Pig, Hive can be used to analyze the data.

Cost: Hadoop is open source so the cost is no more an issue.

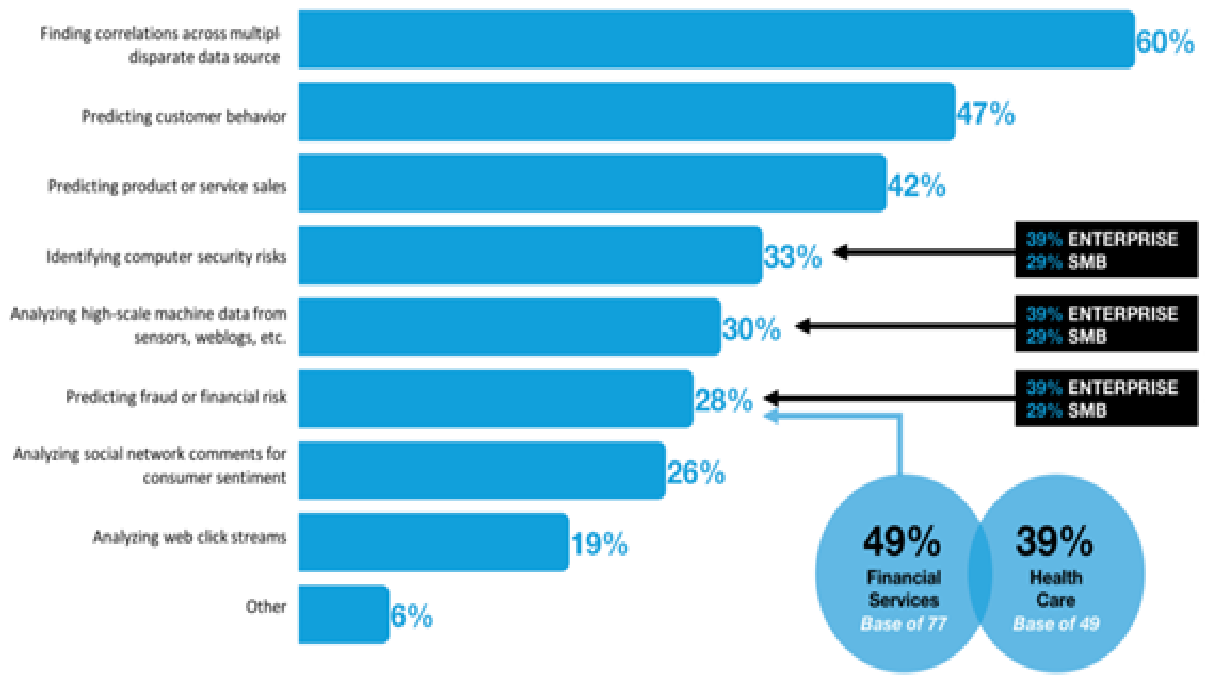


FIG 2 : Finding correlations across multiple disparate data source Predicting custome r behavior.

Reflink:https://www.bing.com/images/search?view=detailV2&insightstoken=bcid_S1MdcS.JDlkDZkX7ZQ.0LmzoC1jg.....2g*ccid_Ux1xL8kO&form=SBIWEB&iss=SBIUPLOADGET&sbisrc=ImgPicker&idpbck=1&sbifsz=584+x+292+%c2%b7+21.54+kB+%c2%b7+png&sbifnm=Untitled.png&thw=584&thh=292&ptime=24&dlen=29404&expw=584&exph=292&selectedindex=0&id=1317468968&ccid=Ux1xL8kO&vt=3&sim=11&cal=0.05&cab=0.95&cat=0.05&car=0.95.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Distributed Computing

A distributed computer system consists of multiple software components that are on multiple computers, but run as a single system. The computers that are in a distributed system can be physically close together and connected by a local network, or they can be geographically d istant and connected by a wide area network. A distributed system can consist of any numbe r of possible configurations, such as mainframes, personal computers, workstations, minicom puters, and so on. The goal of distributed computing is to make such a network work as a sin gle computer.
Distributed systems offer many benefits over centralized systems, including the following:

Scalability

The system can easily be expanded by adding more machines as needed.

Redundancy

Several machines can provide the same services, so if one is unavailable, work does not st op. Additionally, because many smaller machines can be used, this redundancy does not ne ed to be prohibitively expensive.
Distributed computing systems can run on hardware that is provided by many vendors, and c an use a variety of standardsbased software components. Such systems are independent of th e underlying software. They can run on various operating systems, and can use various com munications protocols. Some hardware might use UNIX or Linux as the operating system, w hile other hardware might use Windows operating systems. For intermachine communication s, this hardware can use SNA or TCP/IP on Ethernet or Token Ring.

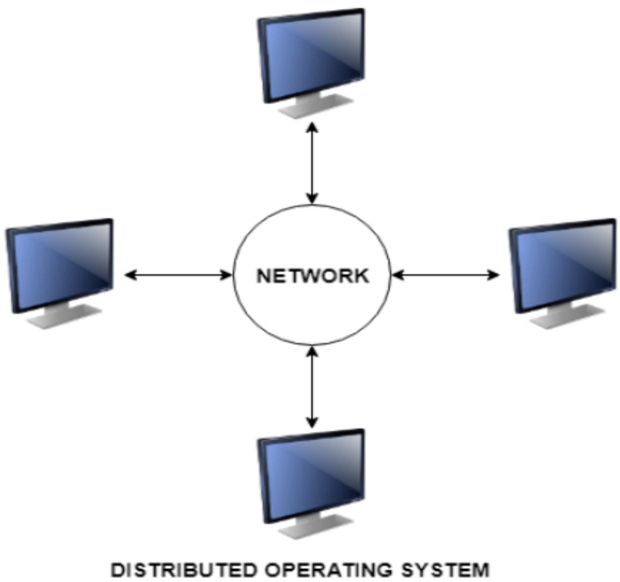


FIG: Distributed Operating System

Reflink:https://www.tutorialspoint.com/Distributed-Systems

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

In a distributed database management system, the database is not stored at a single location. Rather, it may be stored in multiple computers at the same place or geographically spread fa r away. Despite all this, the distributed database appears as a single database to the user.

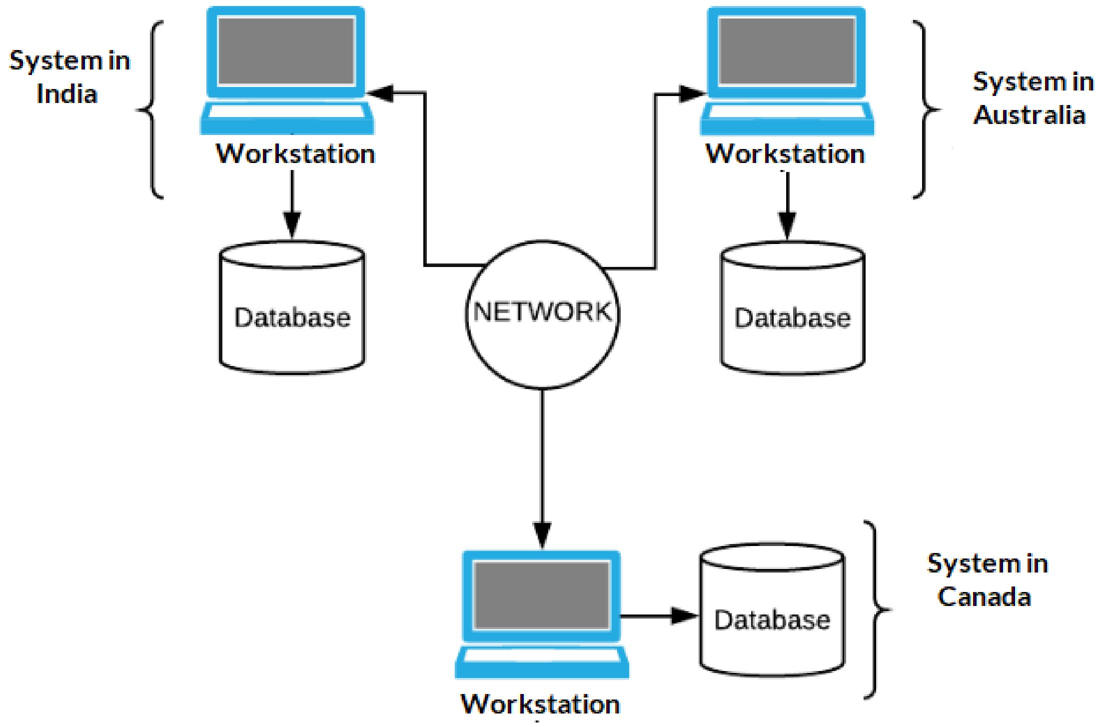


Fig:Network of different countries data bases

As seen in the figure, the components of the distributed database can be in multiple loc ations such as India, Canada, Australia, etc. However, this is transparent to the user i.e the d atabase appears as a single entity.

Big Data with Distributed Computing Frame work

Distributed Computing has a great role in the success of Big Data. Big Data requires very low costing storage space and infrastructure, which is provided by cloud computing. Cl oud Computing is a branch of Distributed Computing . In order to process a huge quantity of data at a very high speed, we required the power of cluster computing, which is also a branc h of Distributed Computing. Thus, to enhance the processing speed of Big Data, there are tw o features: batch processing and stream processing . By default, Hadoop MapReduce is a bat ch processing system, but with the invent of social media inception of data in real-time, there is a need for a stream processing framework for Big Data. MapReduce processin g frame work is designed to handle large data sets and split datasets into small batches. In co ntrast other frameworks process data in an uninterrupted stream, as it flows into the system.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

MapReduce workings as its default processing engine but, for instance, Apache Spark and ot her framework are hooked into the Hadoop Ecosystem to handle real-time streaming data.

Comparison of different computing techniques considering different function

Grid Computing	Utility Computing	Cluster Computing	Cloud Computing
Loosely coupled	On-demand pricing	Tightly coupled systems	On-demand self-service
Diversity and dynamism	Uniform utility computing services	Single system image	Broad network access
Distributed job management and scheduling	Share the resources in the shared pool of machines	Centralized job management and scheduling system	Resources pooling and rapid elasticity
High-end computers (servers, clusters)	High-end computers (servers)	Commodity computers	Commodity computers, high-speed network and high-end servers and NAS

Batch processing frameworks split the large data jobs into small chunks and distribute these chunks on a large number of nodes according to the size of computer cluster to process Big Data. The execution time of a batch processing is determined with the number of active nod es in a cluster and the size of the job. The batch processing model is inappropriate to satisfy real-
time constraints due to having high latency to process Big Data. Stream processing is a mod el for handling real-
time stream synchronized with the data flow and returns the results in a low-
latency fashion. Stream processing also has some features of batch processing such as fault t olerance, high availability, and resource utilization. Real-
time stream processing systems give guarantee to be up and available all the time for real-
time data. Stream processing achieves incremental scalability automatically by distributing p rocessing power as well as storage capacity across multiple computers without any human in teraction. The following frameworks are hooked in Hadoop environment:

- MapReduce (Batch Processing frameworks)
- Spark (Stream and Batch Processing frameworks)
- Flink (Stream and Batch Processing frameworks)
- Storm (Stream processing frameworks)
- Samza (Stream Processing frameworks)

RELATED WORKS

Jun Ni, Ying Chen, Jie Sha, and Minghuan Zhang, Presented review on the demands an d application potentials using big data technology with an emphasis on common challenges. After briefly addressing the Hadoop/MapReduce code components and modules, we use a si mple clinic data to demonstrate how to map and reduce on small dataset with illustrated wor kflow. We give simple scenario of using other MapReduce calculation modules for countin g and classification. This serves as a basic step into future utilization of big data to healthca re domain.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Yaping Chi , Yintan Yang,Ping Xu,Gefei Li,Shuhao Li,Presented review on addressing the limitations of SoftwareDefined Networking (SDN) offers the means to dynamically conf igure the network parameters, dynamically provision network itself can be sliced in an on-demand manner. This research aims to characterize SDN with respect to the demands of big data analytics in Cluster, Grid, and Cloud Computing resources. The main motivation behind this research study is to design and develop an intelligent framework named as Big Data An alytics Management System (BDAMS) for collectively managing the compute, storage, and network resources in Cluster, Grid, and Cloud infrastructure for big data analytics.

J.Lozano, N.Aginako, M.Quartulli, I.Olaizola, E.Zulueta, P. Iriondo Presented review o n contributionof describing a parallelized data processing approach for EO image analysis t hat is based on the MapReduce paradigm and implemented on the Apache Spark framewor k.Existing algorithms for e.g. thematic mapping need to be re-defined in order to exploit distributed execution capabilities to run on large coverage data.

Mohammed S. Al-kahtani1 Lutful Karim2 and Jalal Almhana Presented review on CEDA for big data pr ocessing based on a framework that comprises data processing both at the data collection en d and data processing server end. The proposed CEDA algorithm is application independent and scalable, i.e., using as many nodes as necessary. The proposed CEDA scheme supports both parallel and sequential implementation based on the amount of data to process. If the a mount of data exceeds a certain threshold, it works in parallel mode. Otherwise, it works in s equential mode to reduce processing overhead. Moreover, the algorithm works on different t ypes of nodes ranging from low powered RFID tags and sensors to high speed/powered com puters. The performance of the CEDA scheme was evaluated in terms of data size and data p rocessing time. Simulation results show that the data size processed at the central server usin g CEDA is much smaller compared to that processed by existing approaches

Ivan E. Villalon-Turrubiates presented review paper on An intelligent post-processing paradigm based on the use of a dynamical filtering technique modified to enhan ce the reconstruction quality of remote sensing indexes using multitemporal images and dist ributed computing techniques is proposed. As a matter of particular study, a robust algorithm is reported for the analysis of the dynamic behavior of geophysical signatures extracted fro m remotely sensed scenes. The simulation results prove the efficiency of the proposed techni que along with the computational implementation based on a big-data frame work using distributed Processing.

SandeepKumar Hegde , Dr. Srinivasa K.G Presented review paper on Assigning the tas k to the several cluster of node in Map Reduce is an interesting problem, because efficient ta sk assignment can significantly reduce the processing runtime, or improve hardware utilizati on. Many research work are proposed in Map Reduce on resource management to improve the performance of Hadoop with respect to the user of the application. A fair schedule strate

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

gy provides a delay Scheduling between jobs in pool to improve data locality. The heterogen eous nodes are balanced using load balancing algorithm. In a heterogeneous computing envir onment the performance of Hadoop framework may be reduced due to lack of load manage ment. In order to optimally allocate resources The taskbased algorithms and workflow based algorithms are used. For the data intensive application the Workflowbased approaches are i mplemented which will work even when the estimation about future tasks are not accurate b ut the problem with this technique is it is not suitable when load in the distributed environme nt is not uniform.The job scheduling was also done using Bayesian classification algorithm and the characteristic of this technique was it works based on the principle of Bayesian prob ability hence result is not optimal with this approach. To handle issue of local data the conce pt of wait scheduling is used and length of waitingtime is used as logic in order to schedule t he data with time. Here the task are executed selectively which is not ordered strictly with re spect to queue, so the problem of local data was improved. To improve the fairness of the ta sk in Map Reduce cluster ,weight fair queuing scheduling algorithm is proposed . To allocat e a weight to each and every queue and to schedule tasks of the sub- queue Weighted.

Alfredo Cuzzocrea and Ernesto Damiani presented review paper on theoretical data-driven privacypreserving big data framework in distributed environments can be designed, p roved and extended, it is mandatory to set the different target data domains where the frame work applies. Indeed, numerous specialized "vertical" realizations of the framework are poss ible, each one for each particular data setting. Among others, in this paper we consider the m ultidimensional data case because of, not only multidimensional data arise in a wide spectru m of relevant occurrences , but also they wellmarry with actual emerging big data analytics t ools and systems where multidimensional analysis e.g., OLAP plays a major role.

Jun Ni, Ying Chen, Jie Sha, and Minghuan Zhang Presented review on the demands a nd application potentials using big data technology with an emphasis on common challenges . After briefly addressing the Hadoop/MapReduce code components and modules, we use a s imple clinic data to demonstrate how to map and reduce on small dataset with illustrated wor kflow. We give simple scenario of using other MapReduce calculation modules for counting and classification. This serves as a basic step into future utilization of big data to healthcare domain

Tian-en Huang and Qinglai Guo Presented review on a framework for a distributed computi ng platform is designed. Then, distributed algorithms are developed, including a distributed massive sampling simulation method and a distributed feature selection method. Next, the so ftware platform and hardware platform for the distributed computing platform are establishe d. Finally, the platform is applied to the Guangdong Province Power System in China to eval uate its accuracy and efficiency. The simulation results show that the distributed computing platform can improve computing efficiency and perform better than a centralized platform.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Lin Gu, Deze Zeng, Song Guo, Yong Xiang, Presented review on Big data stream processing (BDSP), has become a crucial requirement for many scientific and industrial applications in recent years. By offering a pool of computation, communication and storage resources,

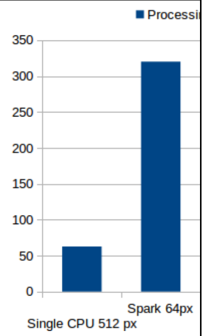
public clouds, like Amazon's EC2, are undoubtedly the most efficient platforms to meet the ever-growing needs of BDSP. Public cloud service providers usually operate a number of geo-distributed datacenters across the globe. Different datacenter pairs are with different inter-datacenter network costs charged by Internet Service Providers (ISPs). While, inter-datacenter traffic in BDSP constitutes a large portion of a cloud provider's traffic demand over the Internet and incurs substantial communication cost, which may even become the dominant operational expenditure factor.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

OVERVIEW OF ALL THE TECHNIQUES USED FOR THE DETECTION AND ANALYSIS OF BIG DATA AND DISTRIBUTED COMPUTING

Sno	Authors	Algorithm/Techinque	Data sets	Accuracy	Result
1	**Jun Ni, Ying Chen, Jie Sha, and Minghuan Zhang**	Hadoop/MapReduce big data processing technique	Student Clinical Data	237.2	MapReduce mechanism **StudentNo Clinic Fee** 13 12.80 14 6.40 15 6.40 16 6.40 17 12.80 18 12.80 19 5.60 20 0.00 21 6.40 22 5.60 23 6.40 24 12.80
2	**Yaping Chi , Yintan Yang,Ping Xu,Gefei Li,Shuhao Li**	MapReduce Processing Algorithm	traffic logs data	**5GB** Standalone Processing 363287 Ms Distributed Processing 20256 ms	COMPARISON OF STYLESDALONE PROCESSING AND DISTRIBUTED PROCESSING PERFORMANCE. HBase read performance test **Data** **Way 0.2GB 0.5GB 1GB 3GB 5GB** Standalone processing 26459ms 61049ms 150547ms 239785ms 363287 ms Distributed processing 35378 ms 54397ms 139045ms 181145 ms 202569 ms

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

3	**J.Lozano, N.Agi nako, M.Quartu lli, I.Olaizola, E. Zulueta, P. Irio ndo**	Gaussian naive Bay esian supervised cla ssification algorithm	original large scale dataset im ages	128 by 128 pixel si zed tiles sh ow	
4	**Mohammed S. Al- kahtani Lutful Karim2 and Jal al Almhana**	Dyanamic Distribute d Algorithm	Number of clust ers at the data c ollecting side M ax. 10 Number of colle cting nodes in e ach cluster at da ta collecting side Max. 1024 Number of grou ping nodes in ea ch cluster at dat a collecting side Max. 20 Number of com modity compute rs Max. 60	results also showe d that the a mount of d ata to be pr ocessed by the central serv er and data processing time in the CEDA app roach	Simulation pa rameters and t heir values **Parameter V alue** Number of cl usters at the d ata collecting side Max. 10 Number of co llecting nodes in each clust er at data collecting sid e

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

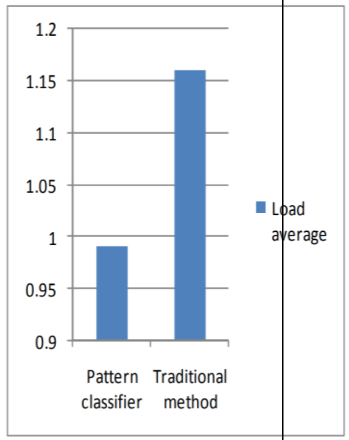
			Size of data coll ected by a node at data collectin g side 128 - 256 Bytes Duration of time data collected a t collecting side 60 Minutes Number of grou pers at central se rver side 4 Processing time of each data byt e by commodity computer 0.25-0.5 nanosecond		Max 102 4 Number of gr ouping nodes in each cluste r at data collecting sid e Max. 2 0 Number of co mmodity com puters Max. 60 Size of data c ollected by a node at data c ollecting side 128 – 256 Bytes

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

5	**Ivan E. Villalon-Turrubiates**	Robust algorithm	distributed processing of large data sets across clusters of computers	MDA framework in order to increase its accuracy is a matter	A RSS map is extracted from a set of multitemporal SPOT-5 images using the WPS method ie Black regions, Dark-

					Duration of time data collected at collecting side 60 Minutes Number of groupers at central server side 4 Processing time of each data byte by commodity Computer 0.25-0.5 nanosecond

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

					gray regions, Light-gray regions
6	**SandeepKumar Hegde ,Dr. Srinivasa K.G**	new pattern classifier algorithm	Parameter For Running Job With Hadoop Cluster Parameter Description Job arrival distribution Exponential Job arrival rate 2 minutes Job Tracker Heart Beat Interval 1 Seconds Admission Interval 1 minutes Job Tracker map slots 60 Job tracker Reduce slots 30	Accuracy of computation	 .Comparison on Load Achieved
7	**Alfredo Cuzzocrea and Ernesto Damiani**	privacy preserving big Multidimensional algorithm	privacy of big data sets sources		Emerging big multidimensional Data
8	**Jun Ni, Ying Chen, Jie Sha, and Minghuan Zhang**	Traffic statistics analysis algorithm flow based on MapReduce		accuracy of the overall analytics process	DRIPROM, along with experimental evaluation and analysis; (ii) focusing on other types of emerging (big) data domains such as graph-like data and textual data

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

<table>
<tr>
<td>9</td>
<td>Tian-
en Huang and Qinglai Gu
o</td>
<td>distributed algorithms</td>
<td>Platform Model

Centralized

8-distributed

Training
Time (s)

4,287

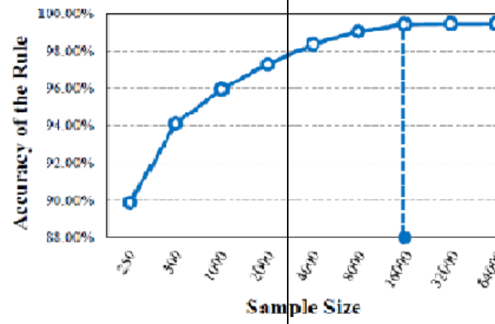
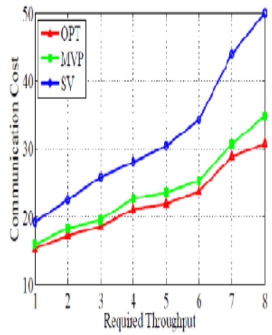
754</td>
<td>Platform Model

Centralized

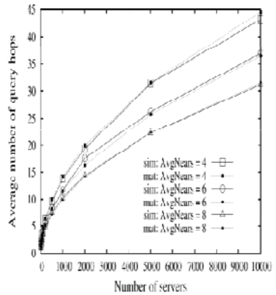
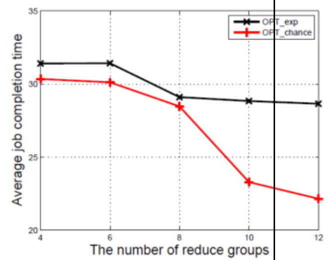
8-distributed

Training
Time (s)

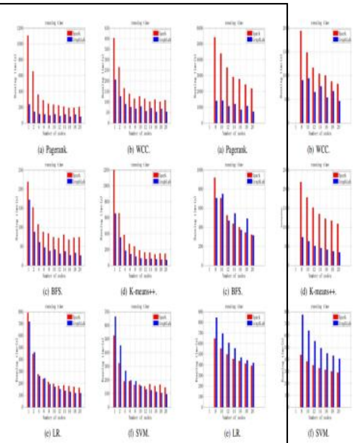
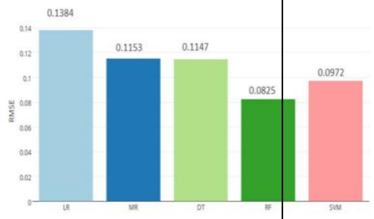
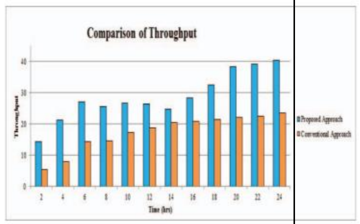
2200

16000</td>
<td>The accuracy of the fin
e operational rule in sc
enarios with different s
ample sizes.

</td>
</tr>
<tr>
<td>10</td>
<td>Lin Gu, Deze Zeng, Song
Guo, Yong Xiang</td>
<td>Multiple Virtual Placement algor
ithm</td>
<td><table><tr><th>Local server</th><th>Selectable linked server</th><th>Distance from target</th><th>Selected server</th></tr><tr><td>101</td><td>(40, 134, 180)</td><td>(182, 88, 42)</td><td>180</td></tr><tr><td>180</td><td>(2, 101, 134, 200)</td><td>(220, 121, 88, 22)</td><td>200</td></tr><tr><td>200</td><td>(2, 180, 222)</td><td>(220, 42, 0)</td><td>222 (Target)</td></tr></table></td>
<td></td>
<td>
</td>
</tr>
</table>

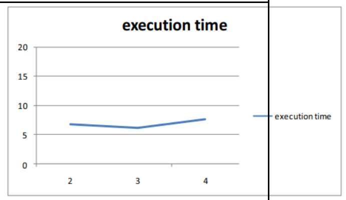
ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

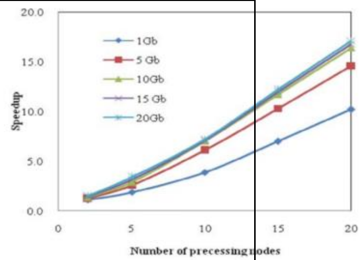
11	**A. Forestiero and G.Papuzzo**	Distributed and Self organising algorithm.	Local server | Selectable linked server | Distance from target | Selected server 101 | (40, 134, 180) | (182, 88, 42) | 180 180 | (2, 101, 134, 200) | (220, 121, 88, 22) | 200 200 | (2, 180, 222) | (220, 42, 0) | 222 (Target)		 The average number of query hops performed by the algorithm to locate the target server. Simulation results, with solid lines, and mathematical results, with dot lines.
12	**P. Li et al ,Song Guo, Toshiaki Miyazaki**	Chance-constrained optimization technique.	Average performance of 30 MapReduce jobs SDA OPT taskOnly OPT exp OPT chance Inter-DC traffic 640 352.8 286.8 289.4 Jobcompletion time 244.5 29.3172 31.4197 29.0284	55% inter-DC traffic	 The average shuffling time under different number of reduce groups.
13	**YinanXu HuiLiu ZhihaoLong**	A hybrid distributed computing framework	parallelized calculating of wind speed data	modified predictor on Spark	Improving percentages in MAE, MAPE and RMSE of IEWT reconstruction component on Spark can reach up to 38.778%, 39.438% and 37.227%, respectively

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

14	**J. Wei, K. Chen, Y. Zhou, Q. Zhou and J. He**	Testing graph and generic analytic algorithms			 Fig. 1. Running time for algorithms with small dataset over local VMs. Fig. 2. Running time for algorithms with large dataset over local VMs.
15	**Mohit Ved, Rizwanahmed B.**	Random Forest algorithm		Multiple regression results in 62.6% accuracy with root mean squared error of 0.11531, while linear regression predicts with an accuracy of 40.2% with root mean squared error of 0.13839	 . RMSE comparison of various Predictive Algorithms. LR: Linear Regression; MR: Multiple Regression; DT: Decision Tree; RF: Random Forest; SVM: Support Vector Machines
16	**Kannan Govindarajan, Thamarai Selvi Somasundaram , David Boulanger , Vivekanandan Suresh Kumar , Kinshuk.**	Software-defined Networking Technology , BigDataApplicationScheduler Algorithm	the submitted user requests	throughput of the submitted big data application requests.	 Comparison of Throughput

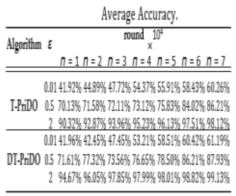
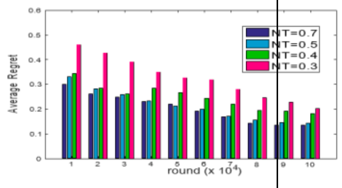
ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

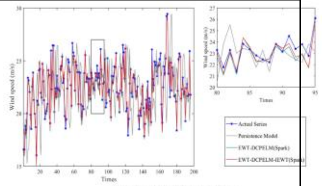
<table>
<tr><th></th><th></th><th></th><th></th><th></th><th></th></tr>
<tr><td>17</td><td>Zhi Yang ,Chunping Zha ng ,Mu Hu ,Feng Lin</td><td>Objectification Parallel Computin g (OPC).</td><td>Yarn based data sets</td><td>cluster changes fro m 2 to 4</td><td>

. Expansibility of system.</td></tr>
<tr><td>18</td><td>J. Yang, Y. Cao, B. Huang and Y. Zhao</td><td>Ditributed Algorithm</td><td>Hiseq2500, Hiseq2500</td><td>greater than 5Gb an d 10 processors.</td><td>

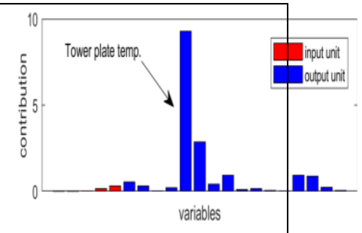
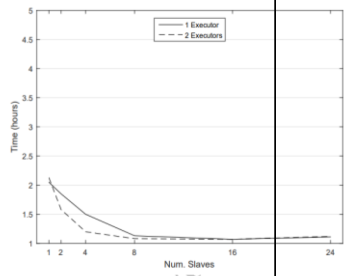
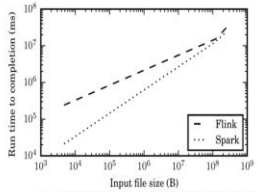
Speedup ratio of distributed algorithm</td></tr>
<tr><td>19</td><td>M. Cavallo, G. Di Modica, C. Polito and O. Tomarchi o</td><td>SICP Mechanism</td><td>Geographically distant sites</td><td>test-
bed was implement ed to prove the via-bility of the approa ch
test-
bed was implement ed to prove the via-bility of the approa ch
test-
bed was implement ed to prove the via-bility of the approa ch</td><td>Table 1. Experiment results.
<table>
<tr><th>Data block location</th><th>Global Reducer</th><th>Real Execution Time [s]</th><th>Predicted Execution Time [s]</th><th>Error [%]</th></tr>
<tr><td>S_1, S_3, S_5</td><td>S_2</td><td>753</td><td>698</td><td>7%</td></tr>
<tr><td>S_1, S_5, S_5</td><td>S_2</td><td>883</td><td>812</td><td>8%</td></tr>
<tr><td>S_3, S_5, S_5</td><td>S_5</td><td>901</td><td>818</td><td>9%</td></tr>
<tr><td>S_5, S_5, S_5</td><td>S_5</td><td>998</td><td>911</td><td>9%</td></tr>
</table></td></tr>
</table>

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

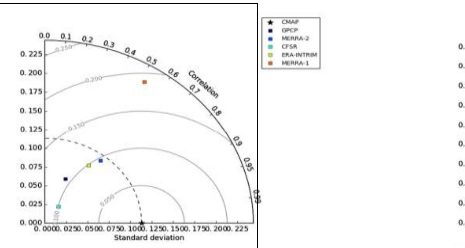
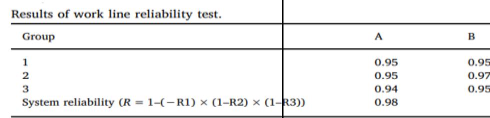
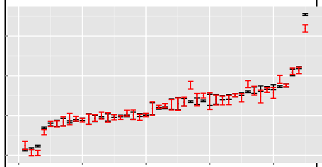
<table>
<tr><td></td><td></td><td></td><td></td><td>Test bed proving the viability of the approach</td><td></td></tr>
<tr><td>20</td><td>S. Dolev, P. Florissi, E. Gudes, S. Sharma and I. Singer</td><td>All Machine Learning algorithms</td><td>Geographical data sets</td><td>JetStream in SQL</td><td></td></tr>
<tr><td>21</td><td>S. Bruce, Z. Li, H. Yang and S. Mukhopadhyay</td><td>A nonparametric two-sample inference algorithm</td><td></td><td></td><td></td></tr>
<tr><td>22</td><td>P. Zhou, K. Wang, L. Guo, S. Gong and B. Zheng</td><td>Novel distributed federated online learning algorithm, T-PriDO Algorithm</td><td>predicting on fixed-size and small-scale datasets</td><td>
<table>
<tr><th colspan="9">Average Accuracy.</th></tr>
<tr><th>Algorithm</th><th>ε</th><th colspan="7">round × 10^4</th></tr>
<tr><th></th><th></th><th>n = 1</th><th>n = 2</th><th>n = 3</th><th>n = 4</th><th>n = 5</th><th>n = 6</th><th>n = 7</th></tr>
<tr><td></td><td>0.01</td><td>41.92%</td><td>44.89%</td><td>47.72%</td><td>54.37%</td><td>55.91%</td><td>58.43%</td><td>60.26%</td></tr>
<tr><td>T-PriDO</td><td>0.5</td><td>70.13%</td><td>71.58%</td><td>72.11%</td><td>73.12%</td><td>75.83%</td><td>84.02%</td><td>86.21%</td></tr>
<tr><td></td><td>2</td><td>90.32%</td><td>92.87%</td><td>93.96%</td><td>95.23%</td><td>96.13%</td><td>97.51%</td><td>98.12%</td></tr>
<tr><td></td><td>0.01</td><td>41.96%</td><td>42.45%</td><td>47.45%</td><td>53.21%</td><td>58.51%</td><td>60.42%</td><td>61.19%</td></tr>
<tr><td>DT-PriDO</td><td>0.5</td><td>71.61%</td><td>77.32%</td><td>73.56%</td><td>76.65%</td><td>78.50%</td><td>86.21%</td><td>87.93%</td></tr>
<tr><td></td><td>2</td><td>94.67%</td><td>96.05%</td><td>97.85%</td><td>97.99%</td><td>98.01%</td><td>98.82%</td><td>99.13%</td></tr>
</table></td><td>

: Average Regret-NT levels.</td></tr>
<tr><td>23</td><td>YinanXu HuiLiu ZhihaoLong</td><td>optimization algorithm</td><td>Resilient Distributed Dataset</td><td>performance of the modified predictor on Spark</td><td>
. The forecasting results of Experiment 3 in Step-S.</td></tr>
<tr><td>24</td><td>J. Wei, K. Chen, Y. Zhou, Q. Zhou and J. He</td><td>generic analytic algorithms</td><td>VM’s data sets</td><td>100% accuracy compared to over local computer clusters.</td><td></td></tr>
</table>

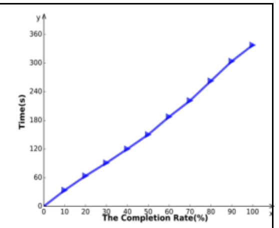
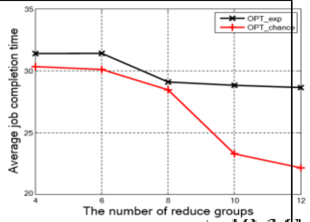
ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

25	**Q. Jiang, S. Yan, H. Cheng and X. Yan**	Distributed Modeling and Computing Frame work for Nonlinear Process Monitoring.	Availability in industrial data	Optimal performance	 Variable contribution plots of the distillation process fault 2.
26	**R. Talavera-Llames, R. Pérez-Chacón, A. Troncoso , F. Martínez-Álvarez**	distributed algorithm	clustering techniques for small and medium datasets		 Speed up depending on the number of slaves in a cluster.
27	**Milad Makkie, Xiang Li, Shannon Quinn, Binbin Lin, Jieping Ye, Geoffrey Mon, Tianming Liu**	D-r1DL algorithm	tfMRI datasets	20% sampled data	 . Run time comparison of D-r1DL using Flink and Spark with varying input data sizes.
28	**Hu, J**	Mathematical statistics and mining algorithms	All input variables	98.27%	

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

29	**FeiHu[a]ChaoweiYang[a] John L.Schnase[b]DanielQ.Duffy[b] MengchaoXu[a]Michael K.Bowen[b]TsengdarLee[c]Weiwei Song[a]**	in-memory, distributed computing framework	Segments of data sets	0.98%	 : The Taylor diagram result for the monthly precipitation fr (−45°, − 90°)~(45°,90
30	**Shiow-LuanWangYung-TsungHou**	Approximation algorithm	All Input variables	0.1 set of performance	 Results of work line reliability test. Group A B 1 0.95 0.95 2 0.95 0.97 3 0.94 0.95 System reliability (R = 1–(−R1) × (1–R2) × (1–R3)) 0.98
31	**Scott Bruce, , Zeda Li, Hsiang-Chieh Yang, SubhadeepMukhopadhyay**	nonparametric two sample inference algorithm	statistical modeling tools of large datasets	95% confidence intervals	
32	**Rohyoung Myung# , Heonchang Yu# , Daewon Lee**	Machine learning algorithm	data analytics programming data sets	1.5 to 3.3 times improvement of execution time	
33	**Ivan E. Villalon-Turrubiates**	Big-Data Technique	data sets across clusters of computers		

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

34	**Le Dong, Zhiyu Lin, Yan Liang, Ling He, Ning Zhang, Qi Chen, Xiaochun Cao and Ebroul Izquierdo**	SICP & DICP Mechanism	ImageNet Dataset		 Pressure resistance test on DICP.
35	**Peng Li, Song Guo, Toshiaki Miyazaki, Xiaofei Liao,**	linearization and relaxation algorithm	shuffling time of 30 MapReduce job instances		 [0.1GB, 0.6GB]

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

						[0.1GB, 0.8GB]
						[0.1GB, 1GB]

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Conclusion

There is an increasing trend of analyzing big data distributed over several data centers located in different countries and regions. Big data has received significant attention from researchers, business industries, education, and scientific communities. It consists of both unstructured and structured data that should be properly extracted, processed, and analyzed in order to obtain meaningful information. It requires large amount of high performance compute cycles, storage, and network bandwidth. The proposed research work first aims to aggregate the high performance computing resources from cluster, grid, and cloud as a collective infrastructure using distributed algorithms. In future, we will continue to study the system implementation by integrating proposed algorithms into popular data processing platform

References

[1] Roger Magoulas and Lorica Ben, "Big Data: Technologies and Techniques for Large-Scale Data", Release 2.11, O'Reilly Media Inc., 2011.
[2] P. Zikopoulos and C. Eaton, "Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data," New York, NY, USA: McGraw-Hill, 2011.
[3] H. Barwick. The "four Vs" of big data. Implementing Information Infrastructure Symposium. [Online]. Available: http://www.computer world.com.au/article/396198/iiis_four_vs_big_data
[4] Hong Bin, Peng Fuyang, Deng Bo, Wang Dongxia. A survey on state monitoring of computational resources in cloud[J]. Computer Applications and Software, 2016(6).
[5] Q. Wang, H. Wang, C. Zhang, W. Wang, Z. Chen and F. Xu, "A Parallel Implementation of Idea Graph to Extract Rare Chances from Big Data", *IEEE Intl Conference on Data Mining Workshop*, Shenzhen, 2014, pp. 503-510.
[6] J. P. Verma, B. Patel, and A. Patel, "Big data analysis: recommendation system with Hadoop framework", IEEE Proc. of CICT, pp. 92-97, 2015.
[7] Crespo,A..,Garcia-Molina, H: Routing indices for peer-to-peer systems.In:Proc.of the 22nd International conference on Distributed Computing systems ICDCS '02.pp. 23-33 (2002)
[8] S. K. Hegde and Srinivasa K.G, "A novel pattern classifier approach towards the performance optimization of Big Data analysis in distributed environment," 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 2017, pp. 100-104, doi: 10.1109/AEEICB.2017.7972391.
[9] A. Cuzzocrea and E. Damiani, "Pedigree-ing Your Big Data: Data-Driven Big Data Privacy in Distributed Environments," 2018 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), 2018, pp. 675-681, doi: 10.1109/CCGRID.2018.00100
[10] Y. Chi, Y. Yang, P. Xu, G. Li and S. Li, "Design and implementation of monitoring data storage and processing scheme based on distributed computing," 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA), 2018, pp. 206-211,doi: 10.1109/ICBDA.2018.8367678.
[11] J. Lozano, N. Aginako, M. Quartulli, I. Olaizola, E. Zulueta and P. Iriondo, "Large scale thematic mapping by supervised machine learning on 'big data' distributed cluster computing frameworks," 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 1504-1507, doi: 10.1109/IGARSS.2015.7326065.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

[12] Z. Yang, C. Zhang, M. Hu and F. Lin, "An Electric Power Big Data Deployment Solution for D istributed Memory Computing," 2015 Seventh International Symposium on Parallel Architectures, Al gorithms and Programming (PAAP), 2015, pp. 155-161, doi: 10.1109/PAAP.2015.38.
[13] Z. Yang, C. Zhang, M. Hu and F. Lin, "OPC: A Distributed Computing and Memory Computin g-
Based Effective Solution of Big Data," 2015 IEEE International Conference on Smart City/SocialCo m/SustainCom (SmartCity), 2015, pp. 50-53, doi: 10.1109/SmartCity.2015.46.
[14] S. Jain, A. Kumar, S. Mandal, and etc., "B4: experience with globally-
deployed software defined wan," in *Proc. of ACM SIGCOMM*, Hong Kong, China, August 2013, pp. 3–14.
[15] J. Yang, Y. Cao, B. Huang and Y. Zhao, "A Ditributed Algorithm for Quality Assessment of Bi ological Sequencing Based on MapReduce," 2019 IEEE 5th International Conference on Computer a nd Communications (ICCC), 2019, pp. 188-192, doi: 10.1109/ICCC47050.2019.9064159.
[16] N. Bahareva, Y. Ushakov, M. Ushakova, D. Parfenov, L. Legashev and I. Bolodurina, "Researc hing a Distributed Computing Automation Platform for Big Data Processing," 2020 International Co nference Engineering and Telecommunication (En&T), 2020, pp. 1-
5, doi: 10.1109/EnT50437.2020.9431254.
[17] S. K. Hegde and Srinivasa K.G, "A novel pattern classifier approach towards the performance o ptimization of Big Data analysis in distributed environment," 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-
Informatics (AEEICB), 2017, pp. 100-104, doi: 10.1109/AEEICB.2017.7972391.
[18] Y. Chi, Y. Yang, P. Xu, G. Li and S. Li, "Design and implementation of monitoring data storag e and processing scheme based on distributed computing," 2018 IEEE 3rd International Conference o n Big Data Analysis (ICBDA), 2018, pp. 206-1,doi:10.1109/ICBDA.2018.8367678.
[19] L. Dong et al., "A Hierarchical Distributed Processing Framework for Big Image Data," in IEEE Transactions on Big Data, vol. 2, no. 4, pp. 297309, 1 Dec. 2016, doi: 10.1109/TBDATA.2016.2613 992
[20] B. R. Chang, H. Tsai, Y. Wang and C. Huang, "Resilient distributed computing platforms for bi g data analysis using Spark and Hadoop," 2016 International Conference on Applied System Innovati on (ICASI), 2016, pp. 1-4, doi: 10.1109/ICASI.2016.7539859.
[21] M. Cavallo, G. Di Modica, C. Polito and O. Tomarchio, "H2F: A Hierarchical Hadoop Framew ork for Big Data Processing in GeoDistributed Environments," 2016 IEEE/ACM 3rd International C onference on Big Data Computing Applications and Technologies (BDCAT), 2016, pp. 27-35.
[22] P. Zhou, K. Wang, L. Guo, S. Gong and B. Zheng, "A PrivacyPreserving Distributed Contextual Federated Online Learning Framework with Big Data Support in Social Recommender Systems," in IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 3, pp. 824-
838, 1 March 2021, doi: 10.1109/TKDE.2019.2936565.
[23] A. Forestiero and G. Papuzzo, "Distributed Algorithm for Big Data Analytics in Healthcare," 20 18 IEEE/WIC/ACM International Conference on Web Intelligence (WI), 2018, pp. 776-
779, doi: 10.1109/WI.2018.00015
[24] J. Lozano, N. Aginako, M. Quartulli, I. Olaizola, E. Zulueta and P. Iriondo, "Large scale themati c mapping by supervised machine learning on 'big data' distributed cluster computing frameworks," 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 1504-
1507, doi: 10.1109/IGARSS.2015.7326065

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

[25] R. S. Gargees and G. J. Scott, "MultiStage Distributed Computing for Big Data: Evaluating Con nective Topologies," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), 2020, pp. 0626-0633, doi: 10.1109/CCWC47524.2020.9031227
[26] P. Li et al., "TrafficAware GeoDistributed Big Data Analytics with Predictable Job Completion Time," in IEEE Transactions on Parallel and Distributed Systems, vol. 28, no. 6, pp. 1785-1796, 1 June 2017, doi: 10.1109/TPDS.2016.2626285
[27] Z. Yang, C. Zhang, M. Hu and F. Lin, "An Electric Power Big Data Deployment Solution for D istributed Memory Computing," 2015 Seventh International Symposium on Parallel Architectures, Al gorithms and Programming (PAAP), 2015, pp. 155-161, doi: 10.1109/PAAP.2015.38
[28] S. Dolev, P. Florissi, E. Gudes, S. Sharma and I. Singer, "A Survey on Geographically Distribut ed BigData Processing Using MapReduce," in IEEE Transactions on Big Data, vol. 5, no. 1, pp. 60-80, 1 March 2019, doi: 10.1109/TBDATA.2017.2723473
[29] S. Bruce, Z. Li, H. Yang and S. Mukhopadhyay, "Nonparametric Distributed Learning Architect ure for Big Data: Algorithm and Applications," in IEEE Transactions on Big Data, vol. 5, no. 2, pp. 166-179, 1 June 2019, doi: 10.1109/TBDATA.2018.2810187
[30] M. Makkie et al., "A Distributed Computing Platform for fMRI Big Data Analytics," in IEEE Tr ansactions on Big Data, vol. 5, no. 2, pp. 109119, 1 June 2019, doi: 10.1109/TBDATA.2018.281150 8.
[31] L. Gu, D. Zeng, S. Guo, Y. Xiang and J. Hu, "A General Communication Cost Optimization Fra mework for Big Data Stream Processing in GeoDistributed Data Centers," in IEEE Transactions on Computers, vol. 65, no. 1, pp. 1929, 1 Jan. 2016, doi: 10.1109/TC.2015.2417566.
[32] M. Ved and R. B., "Big Data Analytics in Telecommunication using State-of-the-art Big Data Framework in a Distributed Computing Environment: A Case Study," 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC), 2019, pp. 411-416, doi: 10.1109/COMPSAC.2019.00066
[33] J. Ni, Y. Chen, J. Sha and M. Zhang, "HadoopBased Distributed Computing Algorithms for Hea lthcare and Clinic Data Processing," 2015 Eighth International Conference on Internet Computing for Science and Engineering (ICICSE), 2015, pp. 188193, doi: 10.1109/ICICSE.2015.41
[34] M. S. Alkahtani, L. Karim and J. Almhana, "Computationally Efficient, Dynamic distributed Al gorithm of sensorbased Big Data," 2017 13th International Wireless Communications and Mobile Co mputing Conference (IWCMC), 2017, pp. 759-763, doi: 10.1109/IWCMC.2017.7986380
[35] K. Govindarajan, T. S. Somasundaram, D. Boulanger, V. S. Kumar and Kinshuk, "A framework for scheduling and managing big data applications in a distributed infrastructure," 2015 Seventh Inte rnational Conference on Advanced Computing (ICoAC), 2015, pp. 1-6, doi: 10.1109/ICoAC.2015.7562784.
[36] A. Cuzzocrea and E. Damiani, "Pedigreeing Your Big Data: DataDriven Big Data Privacy in Di stributed Environments," 2018 18th IEEE/ACM International Symposium on Cluster, Cloud and Gri d Computing (CCGRID), 2018, pp. 675681, doi: 10.1109/CCGRID.2018.00100
Z. Yang, C. Zhang, M. Hu and F. Lin, "OPC: A Distributed Computing and Memory ComputingBas ed Effective Solution of Big Data," 2015 IEEE International Conference on Smart City/SocialCom/S ustainCom (SmartCity), 2015, pp. 50-53, doi: 10.1109/SmartCity.2015.46
J. Yang, Y. Cao, B. Huang and Y. Zhao, "A Ditributed Algorithm for Quality Assessment of Biologi cal Sequencing Based on MapReduce," 2019 IEEE 5th International Conference on Computer and C ommunications (ICCC), 2019, pp. 188-192, doi: 10.1109/ICCC47050.2019.9064159

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731