International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 4, Issue 09, September-2018

USAGE OF DES ALGORITHM TO PROVIDE SECURITY FOR THE DATA IN BIG DATA

Merum Sivaprasad[1], Dr. S. Jassica Saritha[2]

[1]M.Tech Student, Department of Computer Science & Engineering, JNTUA College of Engineering, Pulivendula, Pulivendula 516390, Andhra Pradesh, India.
[2]Assistant Professor, Department of Computer Science & Engineering, JNTUA College of Engineering, Pulivendula,
Pulivendula 516390, Andhra Pradesh, India.

Abstract-In recent years, huge data became a hot analysis topic. The increasing quantity of huge data additionally will increase the prospect of breaching the privacy of people. Since huge data need high process power and huge storage, distributed systems are used. As multiple parties are concerned in these systems, the chance of privacy violation is enhanced. There are varieties of privacy-preserving mechanisms developed for privacy protection at totally different stages (e.g., data generation, data storage, and knowledge processing) of an enormous knowledge life cycle. During this paper we have a tendency to use DES formula to perform the cryptography strategy. Really data encryption standard (DES) may be a symmetric-key formula for the encryption of electronic knowledge. Though insecure, it absolutely was extremely prestigious within the advancement of recent cryptography. By exploitation this formula we will establish the encryption and execution time of every file. This approach is intended to maximize the privacy protection scope by employing a selective cryptography strategy inside the specified execution time needs.

Keywords: - Data Encryption Standard (DES), big data, symmetric-key algorithm, Encryption.

I. INTRODUCTION

Because of ongoing modern improvement, the measure of records produced by long variety informal verbal exchange locales, sensor systems, Internet, human services applications, and several specific organizations, is absolutely increasing little by little. All the enormous measure of statistics produced from various assets in numerous configurations with speedy is alluded as massive records. Huge facts have been a really dynamic studies territory for maximum current few years. The records age rate is developing so fast that it is finishing up to a terrific diploma difficult to address it utilizing conventional strategies or frameworks. In the period in-between, huge information may be prepared, semi-organized, or unstructured, which includes extra problems while acting information stockpiling and preparing errands. Thusly, to this end, we require better methods to save and look into statistics continuously. Enormous statistics, if caught and broke down in an auspicious way, may be modified over into noteworthy bits of knowledge which can be of essential esteem. It can inspire agencies and institutions to decorate the inward fundamental management manipulate and may make new open doorways thru data examination. It can likewise increase the logical studies and financial system with the aid of changing customary plans of motion and logical qualities. Huge information can be characterized in unique ways. For the extent of this paper we utilize the definition given through International Data Corporation (IDC). In, the term large statistics is characterized as "any other age of improvements and fashions, meant to monetarily separate an incentive from large volumes of a huge collection of facts, by empowering high-velocity trap, disclosure, in addition to research". In view of this definition, the properties of large records are reflected by way of three V's, which are, extent, velocity and assortment, as appeared Volume alludes to the measure of information created. With the improvement of lengthy variety interpersonal verbal exchange destinations, we've seen a sensational increment within the measure of the information. The fee at which new facts are created is regularly described as pace. A typical subject of large statistics is that the facts are assorted, i.E., they may contain content, sound, picture, or video and so forth. This diverse kind of statistics is signified by using assortment.

Despite massive insights is probably feasibly utilized for us to all the extra likely hold close the area and decorate in uncommon segments of human endeavors; the exploding measure of insights has quickened potential coverage break. For example, Amazon and Google can drench up our buying attitudes and browsing inclinations. Long variety relational correspondence dreams, as an instance, Face book save each one of the statistics kind of our very own fact and social institutions. All round perceived video sharing web sites, for example, YouTube recommends us chronicles in mellow of our hobby data. With all of the quality pushed with the guide of tremendous statistics, gathering, securing and reusing our own facts to gain enterprise undertaking favors, have set an possibility to our health and protection. In 2006, AOL launched 20 million appearance request for 650 clients via clearing the AOL distinguishing evidence and IP cope with for check out capacities. Regardless, it took researchers just couple of days to re-apprehend the customers. Customers' protection can be cracked under the going with situations Personal insights at the same time as joined with out of doors datasets may also likewise prompt the inference of late convictions across the customers. Those substances can be protected up and ought now not be supplied to others. Singular facts is in more than one examples collected and used to expand the estimation of business challenge. For instance, man or female's purchasing penchants may also likewise monitor a massive quantity of person actualities. The sensitive information are situated away and organized in an area in no way again moored surely and insights spillage may additionally likewise show up inside the midst of ability and managing stages. Remembering the end cause to make certain extensive insights warranty, more than one structures have been delivered as of overdue. These parts might be collected in mild of the degrees of tremendous information approaches of lifestyles cycle, i.E., measurements age, storing, and adapting to. In statistics age diploma, for the guarantee of coverage, get the opportunity to challenge and contorting measurements techniques are linked. While get to control systems endeavor to compel the passageway to humans's close to home actualities, defiling insights techniques trade the crucial measurements formerly they is probably released to a non-depended on in party. The ways to cope with adapt to wellness safety in information collecting level are transcendently in mellow of encryption structures. Encryption primarily based frameworks is probably additionally segregated into characteristic based surely encryption (ABE), Identity mainly based encryption (IBE), and ability manner encryption. Moreover, to make certain the complicated facts, hybrid fogs are related in which sensitive records are secured in non-public cloud. The certainties taking care of level comprises of security sparing facts administering (PPDP) and acing extraction from the realities. In PPDP, anonymization methodologies, as an example, speculation and hide are used to ensure the wellbeing of information. Ensuring the utilization of the statistics whilst defending the coverage is a first rate look at PPDP. In the thinking of eliminating method, there exist multiple devices to isolate useful realities from considerable scale and complicated records. These structures may be moreover apportioned into clustering, portrayal and affiliation lead mining based truely techniques. While grouping and portrayal separate the statistics into distinct social events, association oversee mining based totally honestly procedures find the prized institutions and patterns within the actualities.

This paper is an multiplied portray of our exam and prior work centered on the overall facts encryption method of full-size facts in cloud systems. Differentiation and our beforehand of time artistic endeavors, the good sized blanketed estimation of this paintings is to improve the use adaption of the proposed method via moreover solidifying the subtle variables of the framework. Our beyond works essentially impart to the operating directing rule of the dynamic statistics encryption contraption and the use estimation. In this paper, we've widened our works of artwork via upgrading the device plot for each precise mode degree by means of methods for utilizing the DES remember.

II. PROPOSED SYSTEM

For wellness, safety saving and numerous mists processing, DES calculation is actualized normally within the cloud device to offer the safety using large facts is proposed in our work. This calculation is applied for Encryption. To get the most severe security to the information, concealing records is given and DES calculation to encryption is attached. For getting ready the large records system safety and protection saving are the methodologies in the proposed work. In disbursed garage the data is placed away whilst a proprietor transfers it. The records encoded within the dispensed storage are completed through the DES calculation. In various open mists the facts need to be sheltered and relaxed. Information can be shielded therefore it won't have any statistics befuddle and cannot be gotten to by way of a few different client in reducing facet encryption. So in this paper makes use of encryption method is applied for cozy capability and correspondence. At least customers or numerous consumer contact the database subsequent to growing the interpreting key which gives checking affirmation points of hobby and enlistment and restrictive sharing that is the ultimate tool. Secrecy is applied to shroud the factors of hobby of client and test the profile of consumer verification.

In the wake of sending the unscrambling key the beneficiary can ready to see the first facts. Enabling it to get the guardianship of personal documents, with an intruder expert co-op sharing of the data, solid stockpiling of huge volumes of information, stores of database management, fee of the data proprietor are every one of the workable consequences that are supplied by using the distributed computing however even inquiries concerning security coverage is being raised. Once, within the wake of accumulating the facts, it's miles encoded using Encryption calculation called Password Based Encryption with DES. In this manner encryption of the data of patient information is finished utilizing DES. In this way, setting away Encrypted statistics into HDFS Location that is completed in Hadoop. The Hadoop File System became created utilizing the disseminated document framework plan. HDFS holds plenty of statistics and gives less demanding get entry to shop such super statistics, at that point the documents are placed away over specific machines. After Encrypt the dataset, the produced parent content material is put away inside the HDFS. MD5 and DES calculation entails utilizing a watchword, a byte show off called salt, and a cycle take a look at along a MD5 message method to create a DES mystery key, this secret's then used to carry out DES encryption or decoding. It is basically used to encode non-public keys, in spite of the truth that it might be utilized to scramble any discretionary information. A secret key based totally determine cannot be instated without unusual records this is passed with the aid of the calculation detail. This statistics is referred to as the salt and cycle tally.

DES Algorithm

The DES (Data Encryption Standard) is an encryption figuring to scramble the facts.DES tackles bits, or parallel numbers as 1s. DES is a champion a few of the maximum comprehensively expressed, openly handy cryptographic systems. It transformed into made via IBM in the Seventies yet converted into later procured by the National Institute of Standards and Technology (NIST), as Federal Information Processing Standard forty six (FIPS PUB forty six). The Data Encryption Standard (DES) is a rectangular Cipher that should scramble and Although the information key for DES is sixty four bits in time period, the genuine key utilized by DES is handiest 56 bits big. The scarcest sizable (proper-most) piece in every byte is a correspondence bit, and must be set in order that there are constantly an atypical assortment of 1s in each byte. These fairness bits are forgotten, so basically the seven maximum basic bits of each byte are linked, bringing around a key period of fifty six bits. The count encounters 16 cycles that weave squares of plaintext with values received from the key. The figuring adjustments sixty 4-piece dedication to an development of endeavors right into a sixty 4-piece yield. Comparative steps, with a comparable key are used for translating. There are various attacks and structures recorded till now those enterprise the inadequacies of DES, which made it an indeterminate rectangular determine. Notwithstanding the growing worries round its weak point, DES continues to be commonly utilized by cash related businesses and stand-out firms worldwide to make sure fragile on-line applications.

The circulate of DES Encryption computation is showed up in Figure 1. The computation printed cloth with a hidden level, sixteen adjusts rectangular determine and an tremendous alternate (i.E. Exchange beginning level).

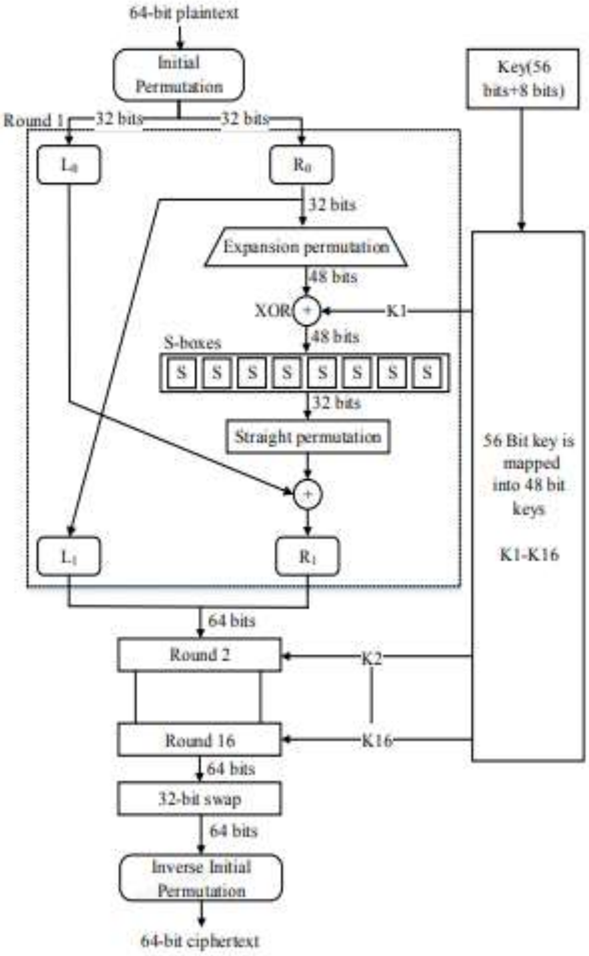
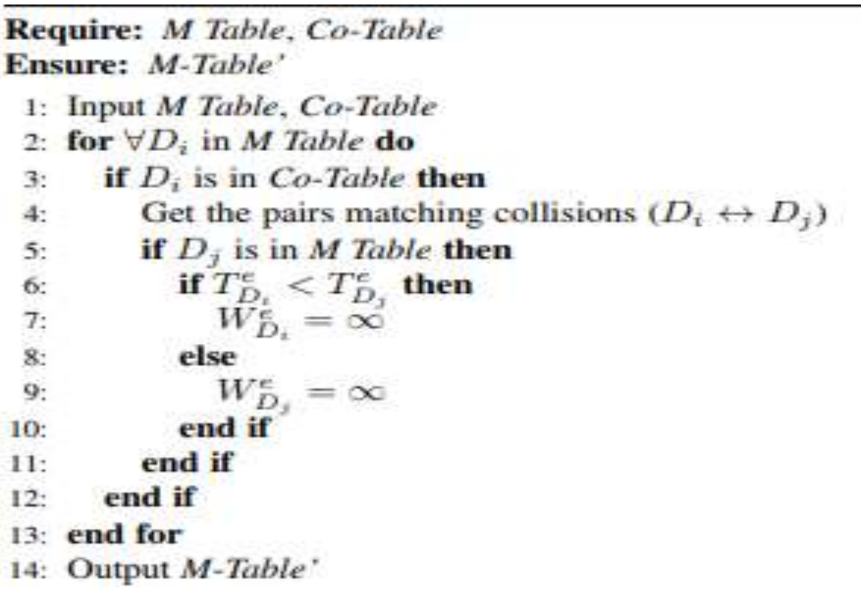


Fig. 1 General Depiction of DES

Weight Modelization (WM) Algorithm:
The WM calculation is created for adjusting M Table making use of weight esteems. The motive for this calculation is to check whether an facts package deal is an absolute necessity scrambled goal, whilst considering the family members between bundles. Subsequently, the units coordinating crashes are linked on this calculation with a specific cease aim to apprehend the matched facts. Data assets incorporate a M Table and a Co-Table. The yield of this calculation is an altered M Table, that's spoken to as a M-Table'. MTable' is a contribution for the two Algorithms five.1 and five.3. Additionally, a Co-Table alludes to a table mapping every single matched datum, which is pre-characterized by way of security techniques or engineers. The Co-Table is utilized to govern units coordinating impacts. Calculation five.2 affords the pseudo codes of WM calculation.
The underneath calculation Presents the pseudo codes of WM calculation
1) Input the primary mapping desk M Table and the predefined Co-Table.
2) For all statistics Di in M Table, decide if records Di is related to desk Co-Table. Discover the matched facts Dj whilst Di is in Co-Table and this blending procedure is spoken to as Di ↔ Dj . 3) Judge whether information Dj is within the

mapping desk M Table with a specific give up purpose to decide if the load esteem need to be altered. The weight esteem must be modified whilst Dj is in M Table.
4) Compare the encryption time lengths among Di and Dj . Dole out an endlessness incentive to De Di whilst the execution time Di is shorter than D0 j s. Something else, hire a limitlessness incentive to De Dj , which implies that we bear in mind this data the most noteworthy encryption need.
5) After all data are labored and refreshed, yield the altered table M-Table'.
The time multifaceted nature of WM calculation is T(n) = O(n). As a factor of reference work of the essential calculation, WM calculation expands the security guarantee level via using a blanketed machine. The following section depicts the technique for producing S Table.

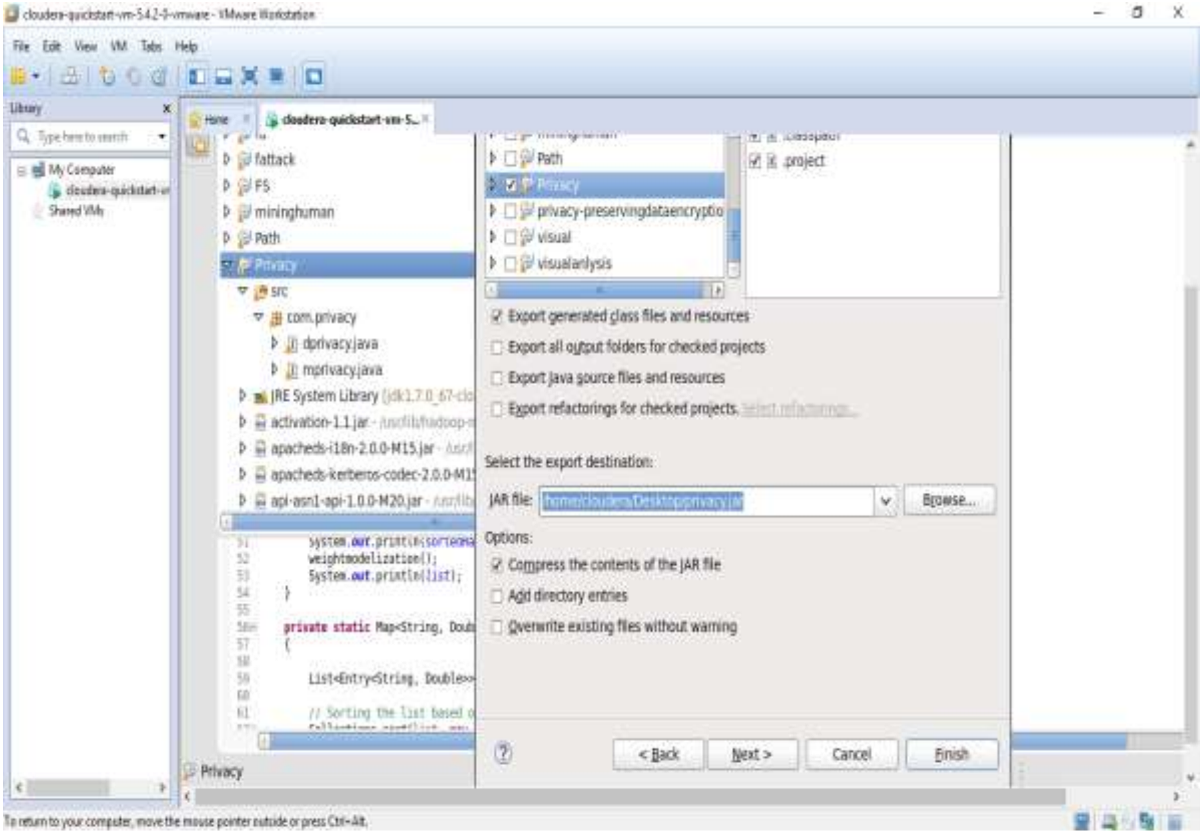
Weight Modelization (WM) Algorithm:-



```
Require: M Table, Co-Table
Ensure: M-Table'
Input M Table, Co-Table
for $\forall D_i$ in M Table do
    if $D_i$ is in Co-Table then
        Get the pairs matching collisions ($D_i \leftrightarrow D_j$)
        if $D_j$ is in M Table then
            if $T^e_{D_i} < T^e_{D_j}$ then
                $W^e_{D_i} = \infty$
            else
                $W^e_{D_j} = \infty$
            end if
        end if
    end if
end for
Output M-Table'
```

III. RESULTS AND DISCUSSION

In the below screen we will create a jar file.



cloudera-quickstart-vm-5.4.2-0-vmware - VMware Workstation
File Edit View VM Tabs Help
Library
Type here to search
My Computer
cloudera-quickstart-vm
Shared VMs
Home
cloudera-quickstart-vm-5...
fattack
FS
mininghuman
Path
Privacy
src
com.privacy
dprivacy.java
mprivacy.java
JRE System Library [jdk1.7.0_67-clo
activation-1.1.jar - /usr/lib/hadoop-
apacheds-i18n-2.0.0-M15.jar - /usr/
apacheds-kerberos-codec-2.0.0-M15
api-asn1-api-1.0.0-M20.jar - /usr/lib
Path
Privacy
privacy-preservingdataencryptio
visual
visualanlysis
.classpath
.project
Back
Next >
Cancel
Finish
Privacy
To return to your computer, move the mouse pointer outside or press Ctrl+Alt.

```
        System.out.println(sortedMa
        weightmodelization();
        System.out.println(list);
    }

    private static Map<String, Doub
    {

        List<Entry<String, Double>>

        // Sorting the list based o
```

☑ Export generated class files and resources
☐ Export all output folders for checked projects
☐ Export Java source files and resources
☐ Export refactorings for checked projects. Select refactorings...

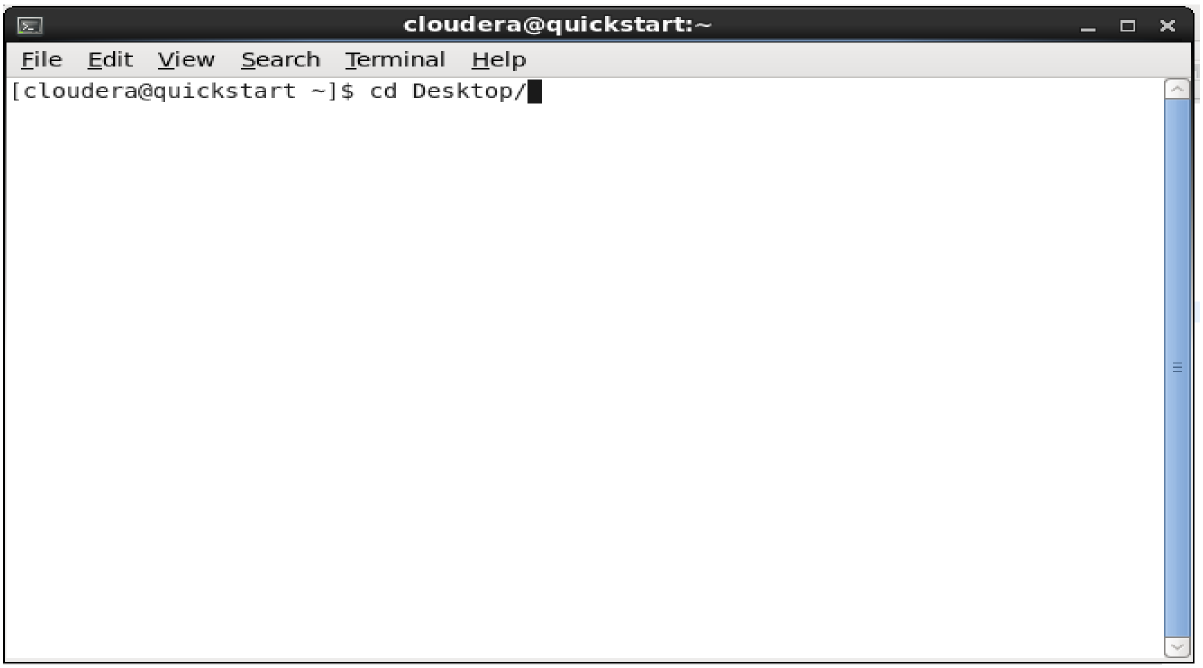
Select the export destination:

JAR file: /home/cloudera/Desktop/privacy.jar Browse...

Options:

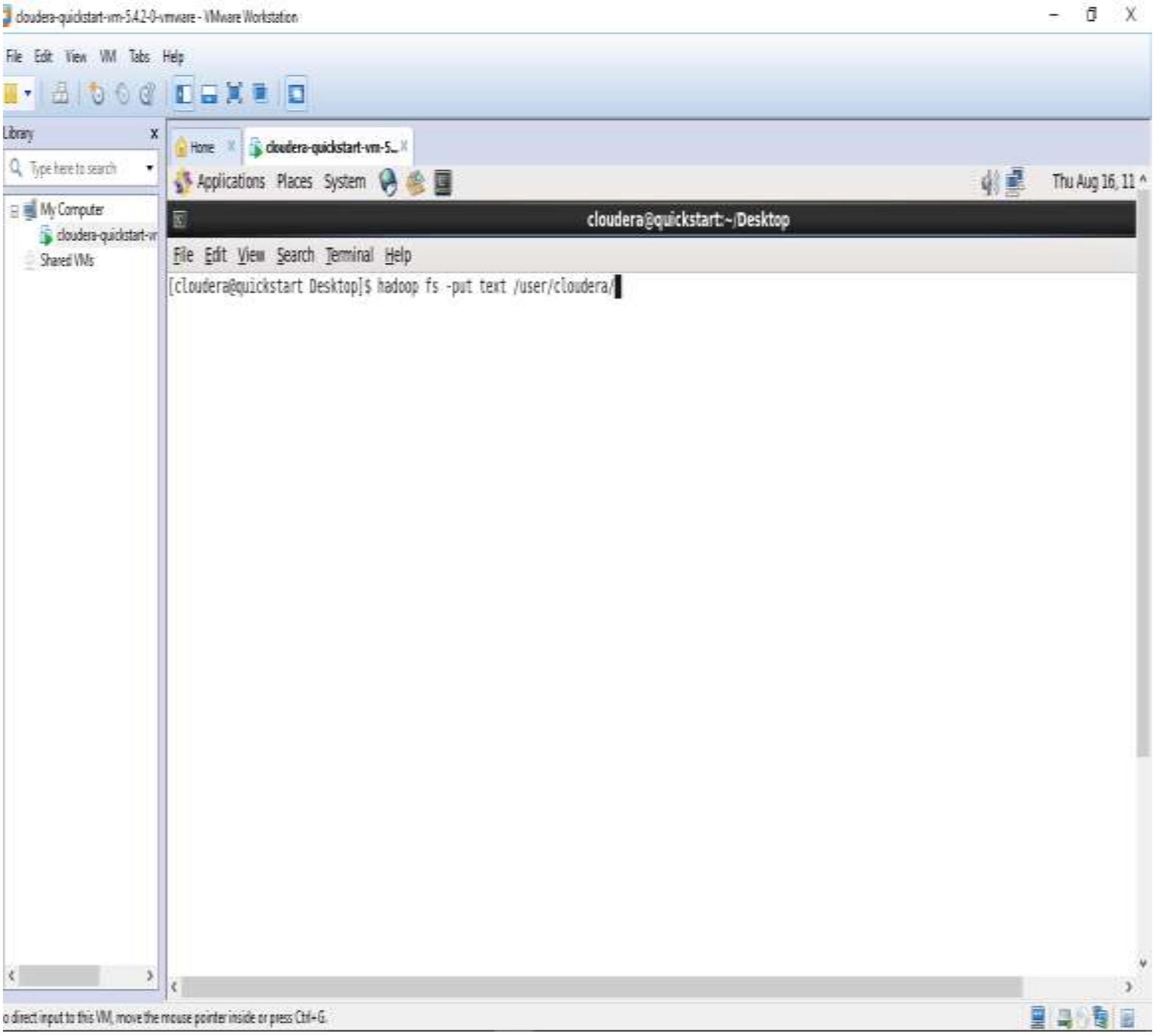
☑ Compress the contents of the JAR file
☐ Add directory entries
☐ Overwrite existing files without warning

Initially we have different path on the screen now we will change this to desktop path in the below screen



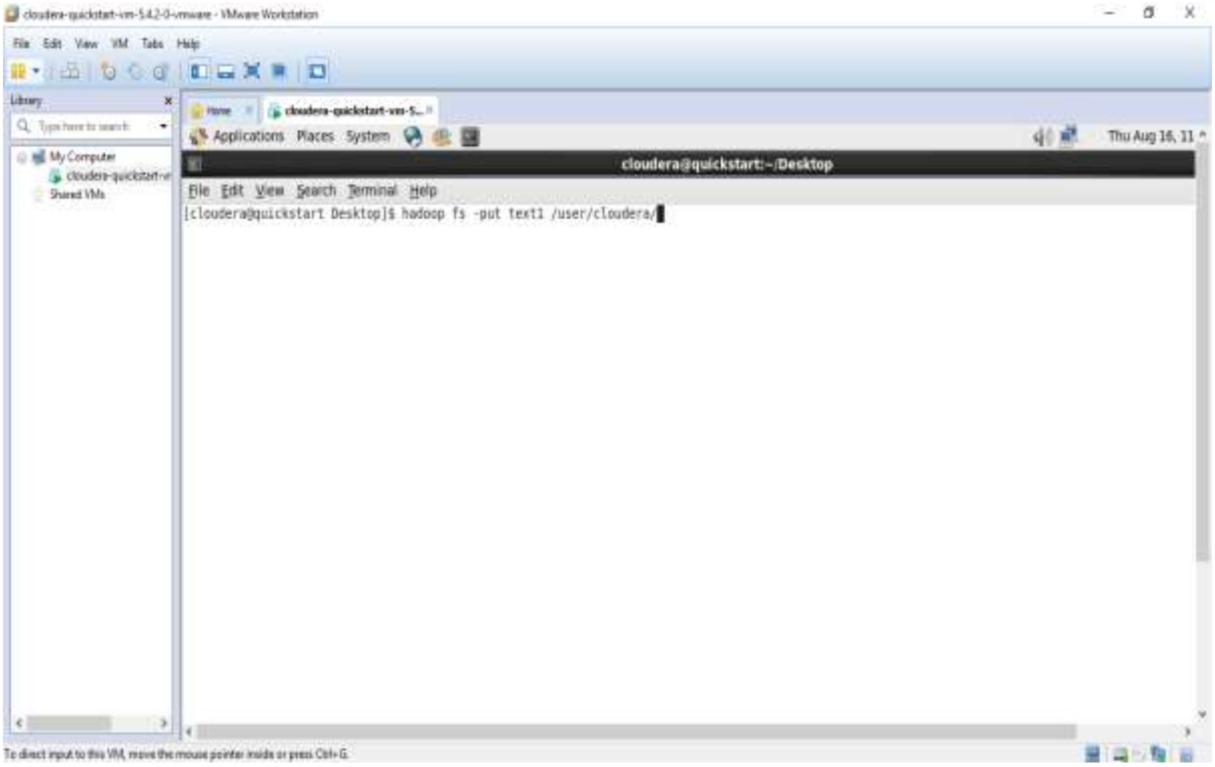
```
[cloudera@quickstart ~]$ cd Desktop/
```

In this below screen we will take dataset from desktop and store it in the HDFS storage Location

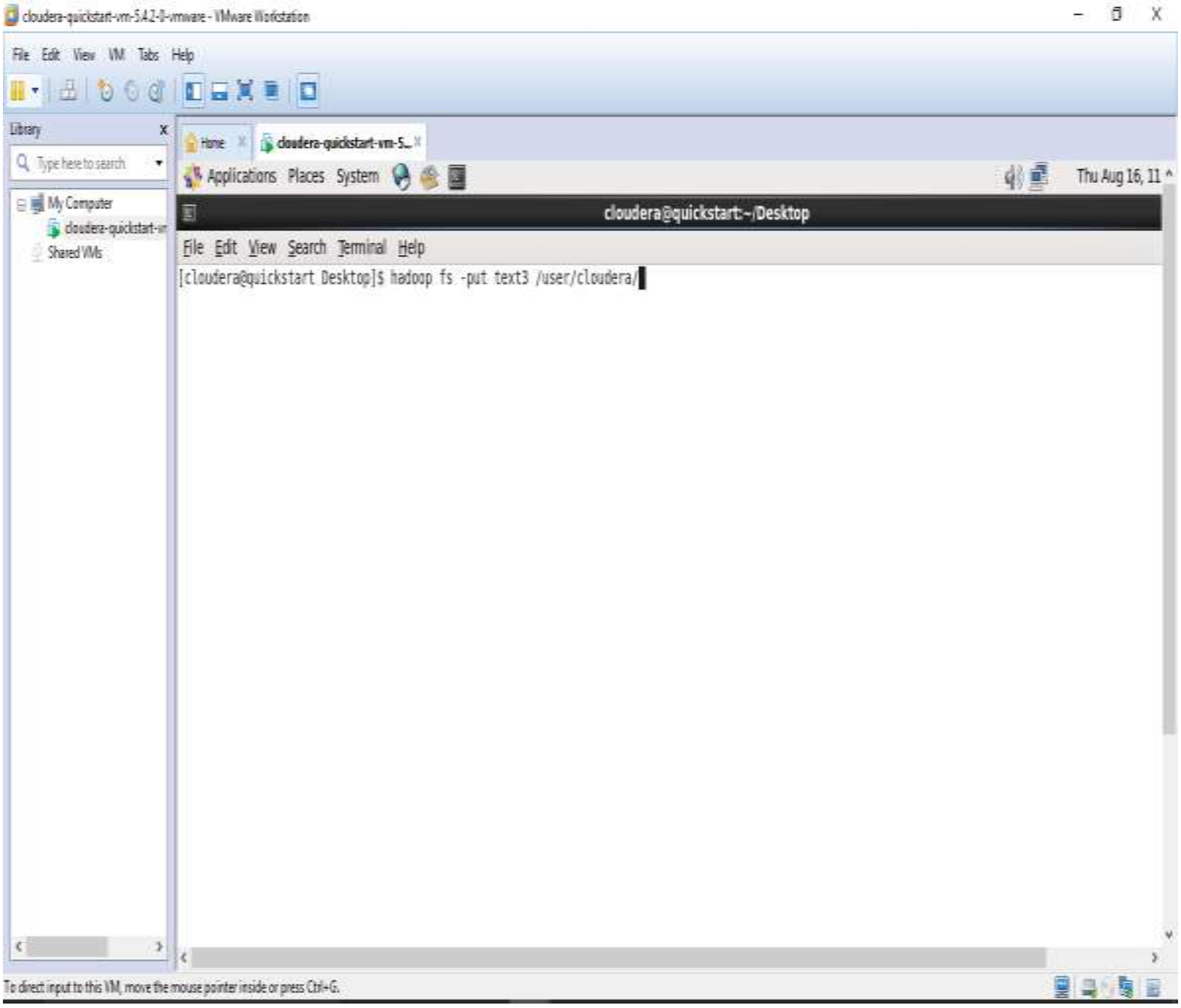

cloudera-quickstart-vm-5.4.2-0-vmware - VMware Workstation
File Edit View VM Tabs Help
Library
Type here to search
My Computer
cloudera-quickstart-vi
Shared VMs
Home
cloudera-quickstart-vm-5...
Applications Places System
Thu Aug 16, 11
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help

```
[cloudera@quickstart Desktop]$ hadoop fs -put text /user/cloudera/
```

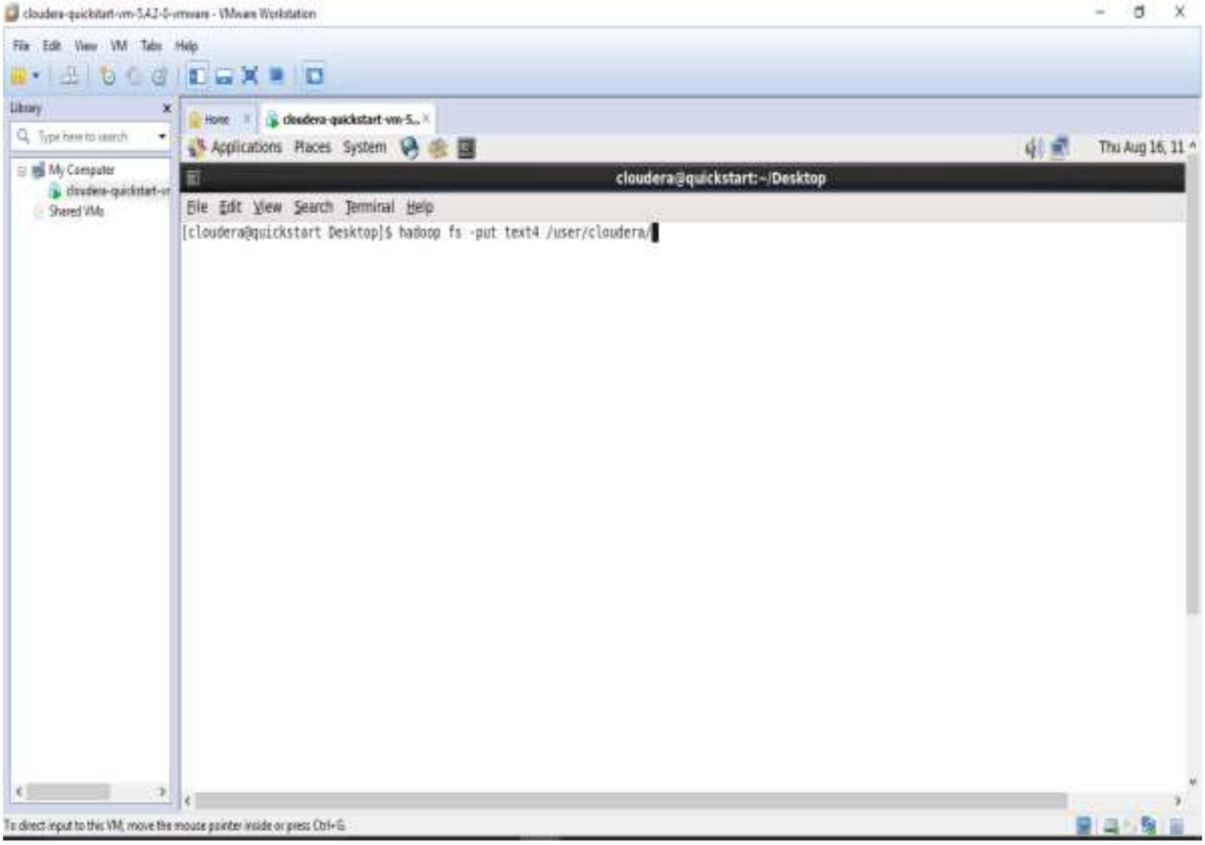
o direct input to this VM, move the mouse pointer inside or press Ctrl+G.


cloudera-quickstart-vm-5.4.2-0-vmware - VMware Workstation
File Edit View VM Tabs Help
Library
Type here to search
My Computer
cloudera-quickstart-v
Shared VMs
Home
cloudera-quickstart-vm-5...
Applications Places System
Thu Aug 16, 11
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
To direct input to this VM, move the mouse pointer inside or press Ctrl+G.

```
[cloudera@quickstart Desktop]$ hadoop fs -put text1 /user/cloudera/
```

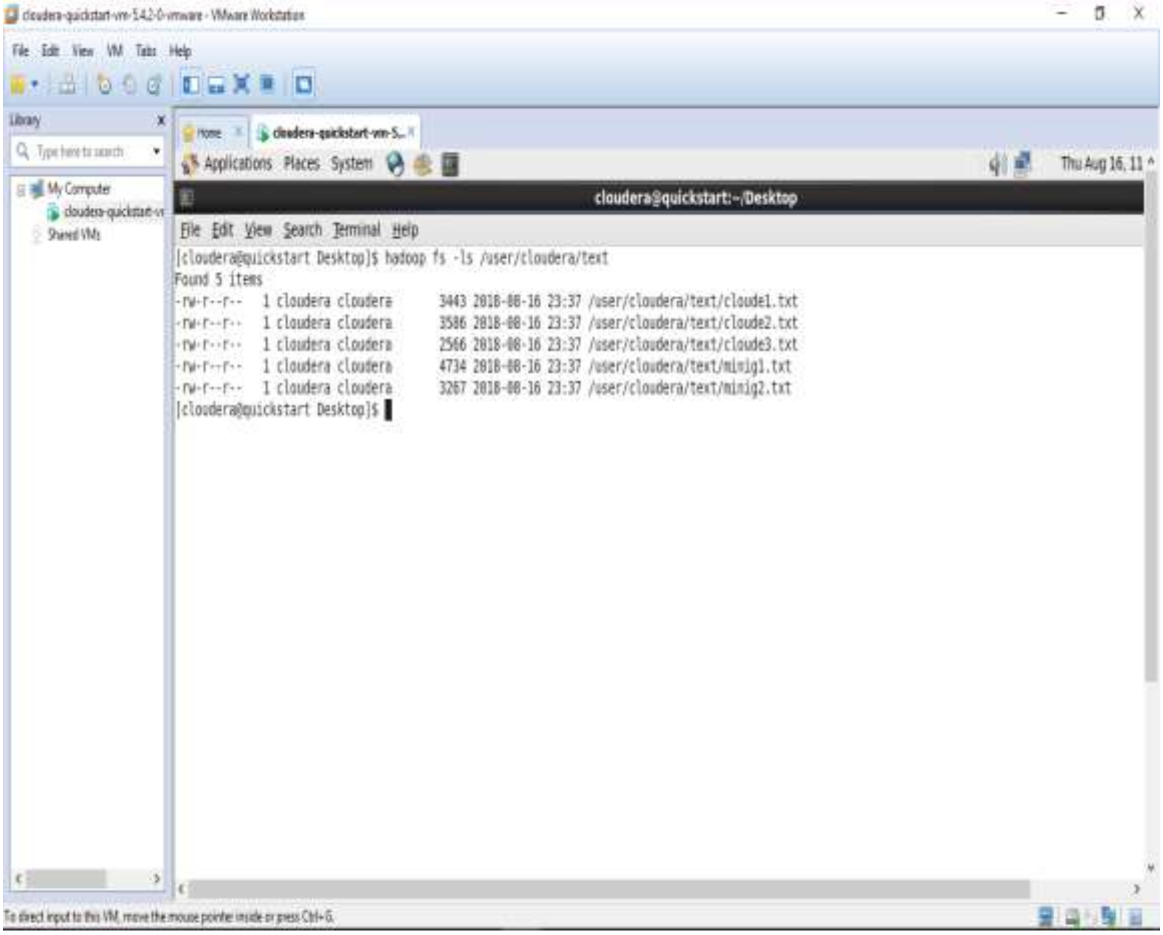

cloudera-quickstart-vm-5.4.2-0-vmware - VMware Workstation
File Edit View VM Tabs Help
Library
Type here to search
My Computer
cloudera-quickstart-v
Shared VMs
Home
cloudera-quickstart-vm-5...
Applications Places System
Thu Aug 16, 11
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
To direct input to this VM, move the mouse pointer inside or press Ctrl+G.

```
[cloudera@quickstart Desktop]$ hadoop fs -put text3 /user/cloudera/
```



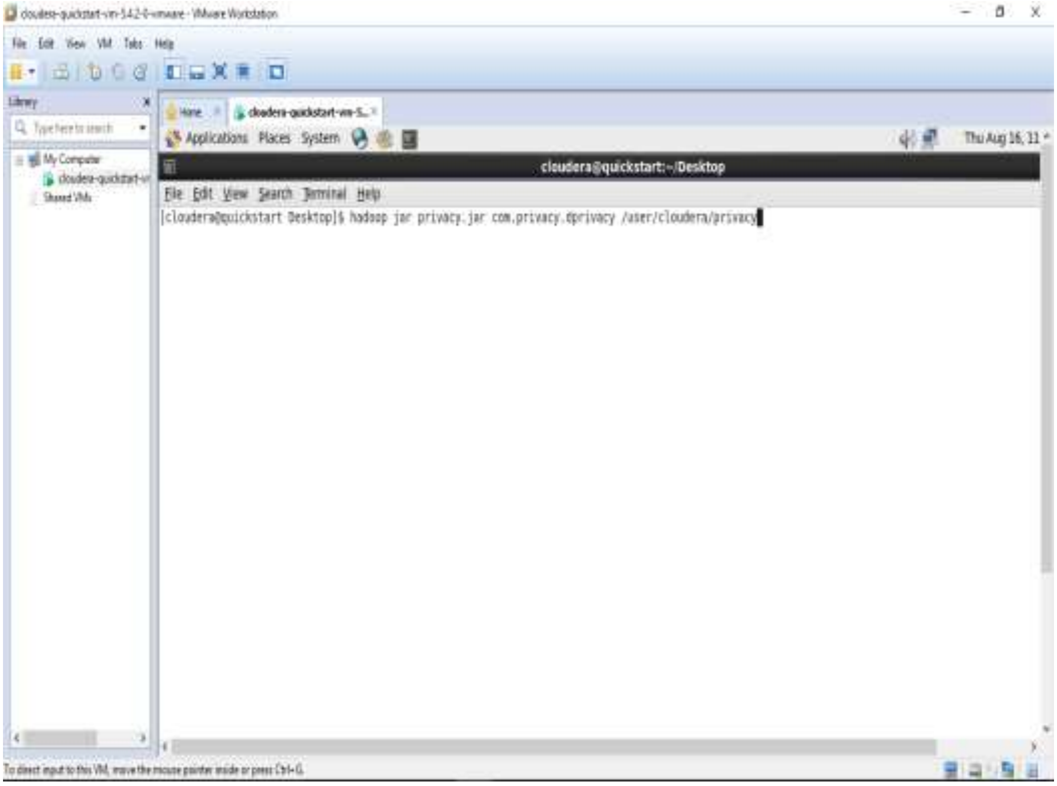
```
[cloudera@quickstart Desktop]$ hadoop fs -put text4 /user/cloudera/
```

In this screen we will check whether the dataset is present in HDFS storage or not.

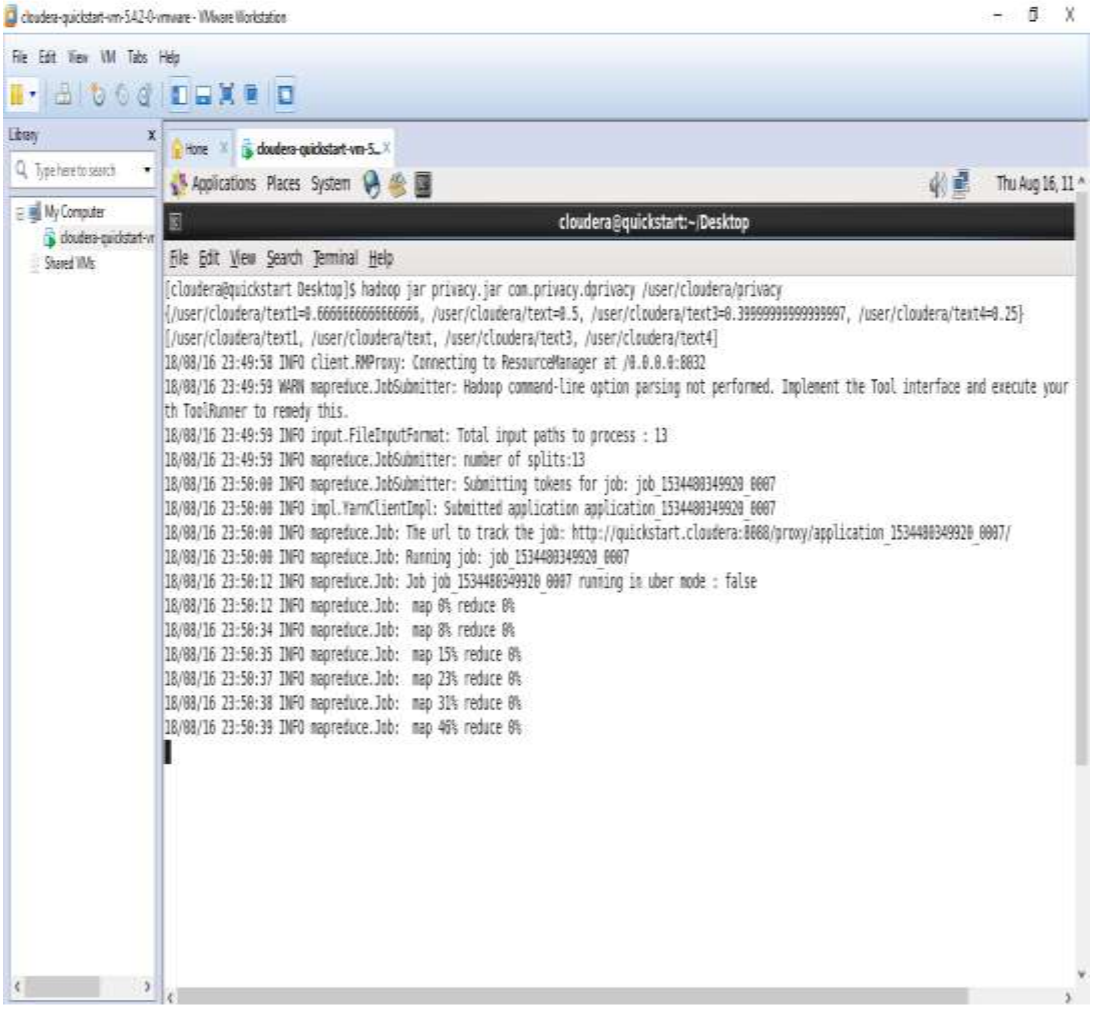


```
[cloudera@quickstart Desktop]$ hadoop fs -ls /user/cloudera/text
Found 5 items
-rw-r--r--   1 cloudera cloudera       3443 2018-08-16 23:37 /user/cloudera/text/cloude1.txt
-rw-r--r--   1 cloudera cloudera       3586 2018-08-16 23:37 /user/cloudera/text/cloude2.txt
-rw-r--r--   1 cloudera cloudera       2566 2018-08-16 23:37 /user/cloudera/text/cloude3.txt
-rw-r--r--   1 cloudera cloudera       4734 2018-08-16 23:37 /user/cloudera/text/minig1.txt
-rw-r--r--   1 cloudera cloudera       3267 2018-08-16 23:37 /user/cloudera/text/minig2.txt
[cloudera@quickstart Desktop]$
```

In this screen we will run the Jar file:

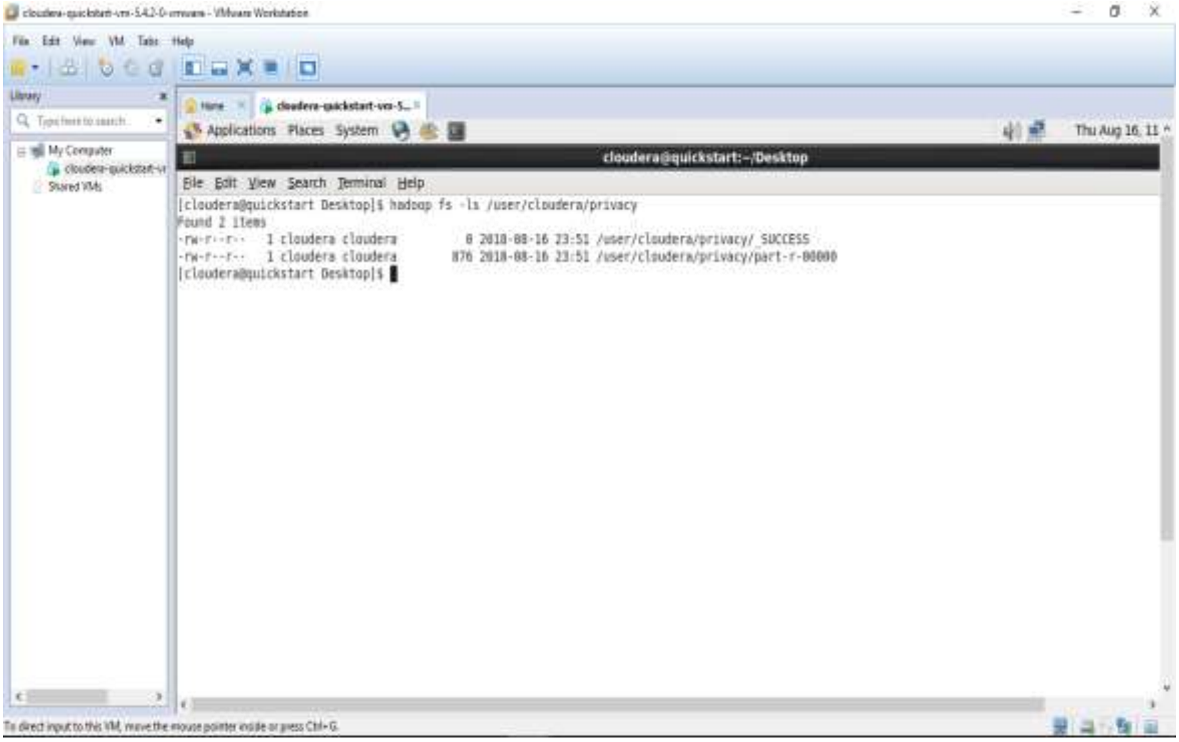

cloudera-quickstart-vm-5.4.2-0-vmware - VMware Workstation
File Edit View VM Tabs Help
Library
Type here to search
My Computer
cloudera-quickstart-vi
Shared VMs
Home
cloudera-quickstart-vm-5...
Applications Places System
Thu Aug 16, 11
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
To direct input to this VM, move the mouse pointer inside or press Ctrl+G.

```
[cloudera@quickstart Desktop]$ hadoop jar privacy.jar com.privacy.dprivacy /user/cloudera/privacy
```


cloudera-quickstart-vm-5.4.2-0-vmware - VMware Workstation
File Edit View VM Tabs Help
Library
Type here to search
My Computer
cloudera-quickstart-vi
Shared VMs
Home
cloudera-quickstart-vm-5...
Applications Places System
Thu Aug 16, 11
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help

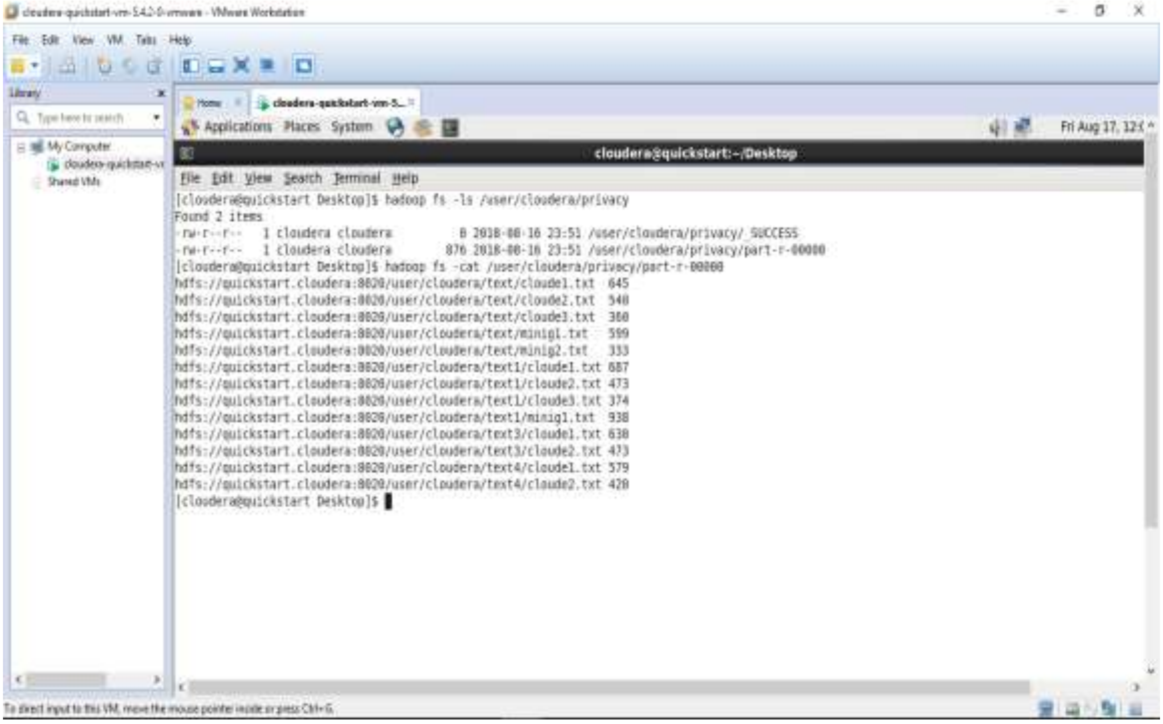
```
[cloudera@quickstart Desktop]$ hadoop jar privacy.jar com.privacy.dprivacy /user/cloudera/privacy
{/user/cloudera/text1=0.6666666666666666, /user/cloudera/text=0.5, /user/cloudera/text3=0.39999999999999997, /user/cloudera/text4=0.25}
[/user/cloudera/text1, /user/cloudera/text, /user/cloudera/text3, /user/cloudera/text4]
18/08/16 23:49:58 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
18/08/16 23:49:59 WARN mapreduce.JobSubmitter: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your
th ToolRunner to remedy this.
18/08/16 23:49:59 INFO input.FileInputFormat: Total input paths to process : 13
18/08/16 23:49:59 INFO mapreduce.JobSubmitter: number of splits:13
18/08/16 23:50:00 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1534480349920_0007
18/08/16 23:50:00 INFO impl.YarnClientImpl: Submitted application application_1534480349920_0007
18/08/16 23:50:00 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application_1534480349920_0007/
18/08/16 23:50:00 INFO mapreduce.Job: Running job: job_1534480349920_0007
18/08/16 23:50:12 INFO mapreduce.Job: Job job_1534480349920_0007 running in uber mode : false
18/08/16 23:50:12 INFO mapreduce.Job:  map 0% reduce 0%
18/08/16 23:50:34 INFO mapreduce.Job:  map 8% reduce 0%
18/08/16 23:50:35 INFO mapreduce.Job:  map 15% reduce 0%
18/08/16 23:50:37 INFO mapreduce.Job:  map 23% reduce 0%
18/08/16 23:50:38 INFO mapreduce.Job:  map 31% reduce 0%
18/08/16 23:50:39 INFO mapreduce.Job:  map 46% reduce 0%
```

In the below screen we view the List of Output Files:



```
[cloudera@quickstart Desktop]$ hadoop fs -ls /user/cloudera/privacy
Found 2 items
-rw-r--r--   1 cloudera cloudera          0 2018-08-16 23:51 /user/cloudera/privacy/_SUCCESS
-rw-r--r--   1 cloudera cloudera        876 2018-08-16 23:51 /user/cloudera/privacy/part-r-00000
[cloudera@quickstart Desktop]$
```

View the encryption execution time



```
[cloudera@quickstart Desktop]$ hadoop fs -ls /user/cloudera/privacy
Found 2 items
-rw-r--r--   1 cloudera cloudera          0 2018-08-16 23:51 /user/cloudera/privacy/_SUCCESS
-rw-r--r--   1 cloudera cloudera        876 2018-08-16 23:51 /user/cloudera/privacy/part-r-00000
[cloudera@quickstart Desktop]$ hadoop fs -cat /user/cloudera/privacy/part-r-00000
hdfs://quickstart.cloudera:8020/user/cloudera/text/cloude1.txt  645
hdfs://quickstart.cloudera:8020/user/cloudera/text/cloude2.txt  540
hdfs://quickstart.cloudera:8020/user/cloudera/text/cloude3.txt  360
hdfs://quickstart.cloudera:8020/user/cloudera/text/minig1.txt   599
hdfs://quickstart.cloudera:8020/user/cloudera/text/minig2.txt   333
hdfs://quickstart.cloudera:8020/user/cloudera/text1/cloude1.txt 687
hdfs://quickstart.cloudera:8020/user/cloudera/text1/cloude2.txt 473
hdfs://quickstart.cloudera:8020/user/cloudera/text1/cloude3.txt 374
hdfs://quickstart.cloudera:8020/user/cloudera/text1/minig1.txt  938
hdfs://quickstart.cloudera:8020/user/cloudera/text3/cloude1.txt 638
hdfs://quickstart.cloudera:8020/user/cloudera/text3/cloude2.txt 473
hdfs://quickstart.cloudera:8020/user/cloudera/text4/cloude1.txt 579
hdfs://quickstart.cloudera:8020/user/cloudera/text4/cloude2.txt 420
[cloudera@quickstart Desktop]$
```

IV. CONCLUSION

The amount data is growing everyday and it's not possible to imagine ensuing generation applications while not manufacturing and execution data driven algorithms. During this paper, we've conducted a comprehensive survey on the privacy problems once coping with massive data. This paper targeted on the privacy problems with massive information and regarded the sensible implementations in cloud computing. The planned approach, DES rule, was designed to maximize the potency of privacy protections. This DES rule is especially used for encryptions underneath completely different temporal arrangement constraints. The experimental evaluations showed the planned approach had associate degree adjustive and superior performance. By victimization this approach the privacy of the massive data content is improved extremely.

REFERENCES

[1] Ren, Yulong, and Wen Tang. "A Service Integrity Assurance Framework for Cloud Computing Based On Mapreduce."Proceedings OfIeee Ccis2012. Hangzhou: 2012, Pp- 240 –244, Oct. 30 2012-Nov. 1 2012. [2] Hao, Chen, and Ying Qiao. "Research Of Cloud Computing Based On The Hadoop Platform.". Chengdu, China: 2011, Pp. 181 – 184, 21-23 Oct 2011.

[3] A, Katal, Wazid M, And Goudar R.H. "Big Data: Issues, Challenges, Tools and Good Practices.". Noida: 2013, Pp. 404 – 409, 8-10 Aug. 2013.

[4] Wie, Jiang, Ravi V.T, And Agrawal G. "A Map-Reduce System With An Alternate Api For Multi-Core Environments.". Melbourne, Vic: 2010, Pp. 84-93, 17-20 May. 2010. International Journal Of Network Security & Its Applications (Ijnsa), Vol.6, No.3, May 2014.

[5] K, Chitharanjan, And Kala Karun A. "A Review On Hadoop — Hdfs Infrastructure Extensions.".JejuIsland: 2013, Pp. 132-137, 11-12 Apr. 2013.

[6] Lu, Huang, Ting-tin Hu, and Hai-shan Chen. "Research on Hadoop Cloud Computing Model and its Applications." Hangzhou, China: 2012, pp. 59 – 63, 21-24 Oct. 2012.

[7] Wie, Jiang, Ravi V.T, and Agrawal G. "A Map-Reduce System with an Alternate API for Multi-core Environments." Melbourne, VIC: 2010, pp. 84-93, 17-20 May. 2010. International Journal of Network Security & Its Applications (IJNSA), Vol.6, No.3, May 2014 56.

[8] K, Chitharanjan, and Kala Karun A. "A review on Hadoop — HDFS infrastructure extensions.".JeJu Island: 2013, pp. 132-137, 11-12 Apr. 2013.

[9] F.C.P, Muhtaroglu, Demir S, Obali M, and Girgin C. "Business model canvas perspective on big data applications." Big Data, 2013 IEEE International Conference, Silicon Valley, CA, Oct 6-9, 2013, pp. 32 - 37. [10] Zhao, Yaxiong , and Jie Wu. "Dache: A data aware caching for big-data applications using the MapReduce framework." INFOCOM, 2013 Proceedings IEEE, Turin, Apr 14-19, 2013, pp. 35 - 39.

[11] Y. Yu, M. Au, G. Ateniese, X. Huang, W. Susilo, Y. Dai, and G. Min. Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage. IEEE Transactions on Information Forensics and Security, PP (99):1, 2016.

[12] L. Weng, L. Amsaleg, A. Morton, and S. Marchand-Maillet. A privacy-preserving framework for large-scale content-based information retrieval. IEEE Transactions on Information Forensics and Security, 10(1):152–167, 2015.

[13] K. Gai, M. Qiu, H. Zhao, and J. Xiong. Privacy-aware adaptive data encryption strategy of big data in cloud computing. In The 2nd IEEE International Conference of Scalable and Smart Cloud (SSC 2016), pages 273–278, Beijing, China, 2016. IEEE.

[14] Y. Zhang, C. Xu, S. Yu, H. Li, and X. Zhang. SCLPV: Secure certificate less public verification for cloud-based cyber-physical social systems against malicious auditors. IEEE Transactions on Computational Social Systems, 2(4):159–170, 2015.

[15] C. Wang, S. Chow, Q. Wang, K. Ren, and W. Lou. Privacy-preserving public auditing for secure cloud storage. IEEE Transactions on Computers, 62(2):362–375, 2013.

[16] Y. Li, W. Dai, Z. Ming, and M. Qiu. Privacy protection for preventing data over-collection in smart city. IEEE Transactions on Computers, PP:1, 2015.

[17] K. Gai, M. Qiu, L. Chen, and M. Liu. Electronic health record error prevention approach using ontology in big data. In 17thIEEE International Conference on High Performance Computing and Communications, pages 752–757, New York, USA, 2015.

[18] M. Qiu, K. Gai, B. Thuraisingham, L. Tao, and H. Zhao. Proactive user-centric secure data scheme using attribute-based semantic access controls for mobile clouds in financial industry. Future Generation Computer Systems, PP:1, 2016.

[19] K. Gai, M. Qiu, B. Thuraisingham, and L. Tao. Proactive attribute based secure data schema for mobile cloud in financial industry. In The IEEE International Symposium on Big Data Security on Cloud;17th IEEE International Conference on High Performance Computing and Communications, pages 1332–1337, New York, USA, 2015.