CREDENTIAL BASED DYNAMIC GROUPING IN CLOUD COMPUTING

[1]P. Kavitha [2]P. Ashok

[1]Associate Professor
[2]Assistant Professor
[1]Computer Science and Engineering,
[1]Sri Sai Ram Institute of Technology, Chennai, India

Abstract: When any data is stored in a private cloud it is visible and under access to all the internal members of the organization this may create threat to the organization by the internal hackers. So, our project provides a way of making the data visible only to members in a group who satisfy some credential measure. And makes the data hidden to others thus enhancing security and avoiding data theft. A dynamic grouping algorithm which gives good performance is obtained to group the data immediately once it is uploaded in Cloud. Data stored in a private or hybrid cloud of any organization remains secure from the internal hackers, thus by making it visible only to employees who satisfy the credentials. It also helps in secure data transmission in a group that has some internal hacker, by identifying them in that possible group.

***Index Terms*:** *Credential, Grouping Algorithm, Internal Hacker, Cloud*

1. INTRODUCTION

Cloud Computing, *"The actual practice of using collection or a network of remote servers that is hosted on the internet to store, manage and process data, rather than in a local server or in a personal computer* "the actual meaning given to it by Google when you type in Cloud Computing. When you see about the types of cloud it is been classified into three, (i) Private Cloud, (ii) Public Cloud,(iii) Hybrid Cloud.

(i)Private Cloud: This architecture is operated solely for a single organization; it may be industrial, educational or anything which is limited to the users of that organization and no one else.

(ii)Public Cloud: This architecture renders its services over the wide range of public use. Public cloud services may be free, and they differ with private cloud architecture based on security parameters.

(iii)Hybrid Cloud: It is a collaboration of private and public clouds together that offers the benefits of both the models.

Clouds are getting more and more widespread due to their pay-as-you-go model that attracts several tiny and medium businesses. A number of them, due to their success, have grownup terribly giant, every containing lots of thousands of servers and hosting up too many virtual machines. To support versatile and economical segment communication in these large-scale cloud knowledge centers, researchers have projected several novel styles. One such flexible communication is that the data being uploaded by the centralized controller must be visible only to the group of people whom he wishes to give access for that particular file, for the rest, the access and the visibility must be denied. This sort of communication is proposed for inducing security to the data that is being uploaded by the admin from the internal hackers of the organization.

2. VISIBILITY CONSTRAINTS TO THE DATA

However, it has been demonstrated that the full control and visibility over all the data's that have been uploaded in the cloud are not always necessary for all the employees or the account holders of the private cloud. By giving the complete visibility of data's in the cloud to all the employees, it may lead to a risk of data theft or any other privacy breaches. And also, the right granularity of the data must be handled by the admin clearly. All this risk occurs due to one major reason that the data in cloud is made visible to all employees, who are related to it and those who are not too. It doesn't seem to be a problem when the data is visible and under access to the employees who are related to it, but the actual problem arises when the access of data are not limited to the employees who are actually not related to it. So, in this paper we advocate a new solution for maintaining the control and visibility of data being uploaded in private clouds. From the observations, it is clear that the data theft by the internal hackers arises due to the visibility of the non-related data to them. So, we came to a conclusion that, only if the data is visible to them, then only they tend to rob it, if it made invisible to them then the data theft can be prevented. That is making the data available to the employees those who are related to it. So, even if the data theft occurs in this situation it is easy to identify the internal suspect as the visibility constraints of the data is made limited.

3. DYNAMICGROUPINGOFDATA

The availability of the data to related employees is made possible by grouping them dynamically while they are being uploaded in the cloud. The idea of grouping the data dynamically overcomes the disadvantage of internal traffic of giving data access to employees. This was the major drawback in the present existing system of grouping the employees and giving the particular groups to access the related data. But by doing like this, it can lead to internal traffic in the system of grouping the employees periodically, thereby degrading the performance and throughput of the system. So, to overcome this, a dynamic grouping algorithm is required to group the data instead. An algorithm that has high performance, throughput and efficiency is chosen to group the data based on credential points that has been generated by the data attributes. These attributes of the data include, (i) Domain the data belongs to and (ii) Project that the data belongs to.So, the grouping is like this, the data of a particular domain is grouped under a particular project respectively with its credential point as generated in a hierarchical structure.

TheFig.1 clearly shows how the data are grouped hierarchically by the domains and projects they actually belong to. The grouping data is mainly based on the credential point that is been generated while mentioning its domain and the project it belongs to. The corresponding folder is created for each credential point and data is stored respectively. This hierarchical structure helps in easy access of data to its related employees and be out of reach to those who are not related to it. As the algorithm makes the grouping Dynamic in nature, that increases performance, throughput and efficiency of the system more when compared to other systems.

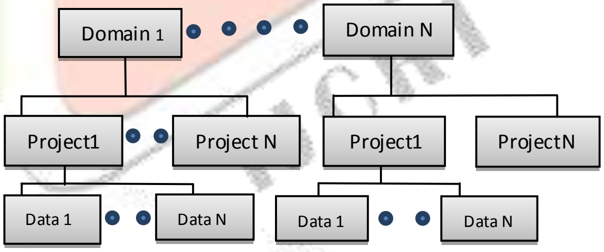


Fig.1. Hierarchical structure shows how the data are grouped by its attributes while it's been uploaded in the cloud.

4. CREDENTIALS TO EMPLOYEES

The main aim of this paper is to provide a way to prevent data theft from internal hackers who are being masked behind the overall employees of the organization. Since the organization's growth, performance and reputation totally depends on its employees, it is difficult to find the minority who betray the organization. So, here we propose an idea to prevent data theft from the internal hackers, by assigning a credential point to each and every employee in the organization. As mentioned earlier, the data's being uploaded are grouped dynamically by their generated credential points. Andnow the employees who all satisfy the credential point of the data with their own credential points are allowed to access the data and others don't. This is made possible by limiting the visibility constraints of the data that is been uploaded. In simple words, the data is visible to the all employees who satisfy the credential point for others the data remain invisible, leaving no clue for its existence in cloud. This method is being used because of the fact, that only if the data is made visible to the internal hacker he tends to misuse it, if the data not related to him is been restricted to his visibility then the chance of data theft reduces, and data is secured from the internal hackers. Prevention of data theft from internal hackers of an organization can be accomplished by providing credential points to employees in two possible cases. **Case(i): If credential point of the internal hacker does not satisfy the credential point of the data**, he is not allowed to access the data by restricting the data's visibility constraint. By this data theft is controlled. **Case(ii): If credential point of the internal hacker does satisfy the credential point of the data,** the possibility of data theft is more prone in this case. But the advantage is that, even if the data theft occurs it is easier to find that internal hacker as the number of employees who satisfy the credential point of the data is less than the total number of employees in the organization. And so, can prevent this mishap in near future.

5. CREDENTIAL BASED DYNAMICGROUPING

The system architecture comprised of the data being uploaded with the credential point in the centralized storage, the employees of the organization with their own credential points. In this topic we clearly going to see the internal working architecture of the system of dynamic grouping and credential point generation.

The below given architecture shows the existing system with its visibility issue as a drawback.

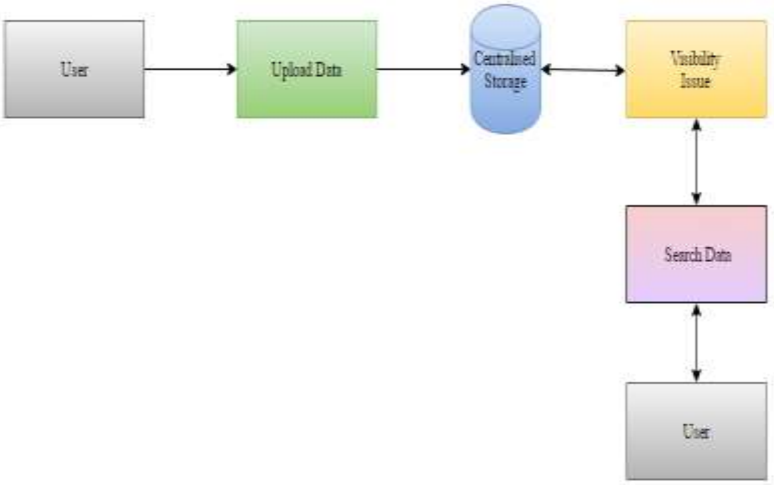


Fig.2. Illustrating the existing system that has a drawback in the visibility issues.

To overcome the above-mentioned drawback in the existing system, we introduce credential based dynamic grouping in the proposed system in this paper. So,in the above-mentioned architecture's drawback is eradicated with the credential point generation for both data and employee side to overcome the visibility issue to prevent the data theft from internal hackers of the organization. Below the modified architecture diagram is shown, that overcomes the disadvantage of existing system.

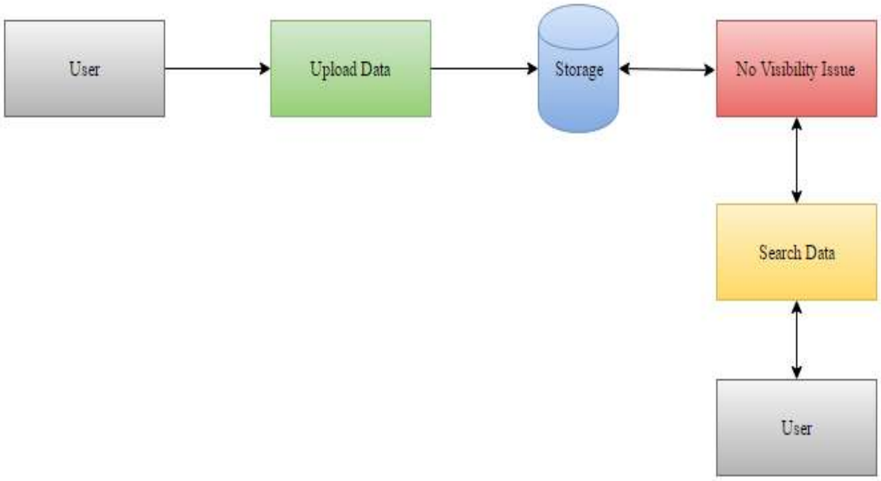


Fig.3. Illustrating the proposed system architecture that has no visibility issues due to credential based dynamic grouping.

To explain more in depth about the working of our system, we created a use case diagram explaining the modules clearly. First the admin of the organization feed the employee details into the database, with that the database generates the credential points for all the employees based on their attributes like designation, trust index, performance, project they work on. And the database also stores data that is being uploaded by the admin or the employees in a hierarchical form based on the domain it belongs to and its credential point. The employee of an organization can now view only the files which are satisfied by his credential points, other files are not visible to him and he cannot access those. And in a time of rare case, an employee is given the access to view files who have not satisfied the credential factor, but it is done as per the organization's requirement is also provided. The Fig 4. Shows the use case diagram explains the architecture and the working functionality of the Credential based Dynamic grouping system.

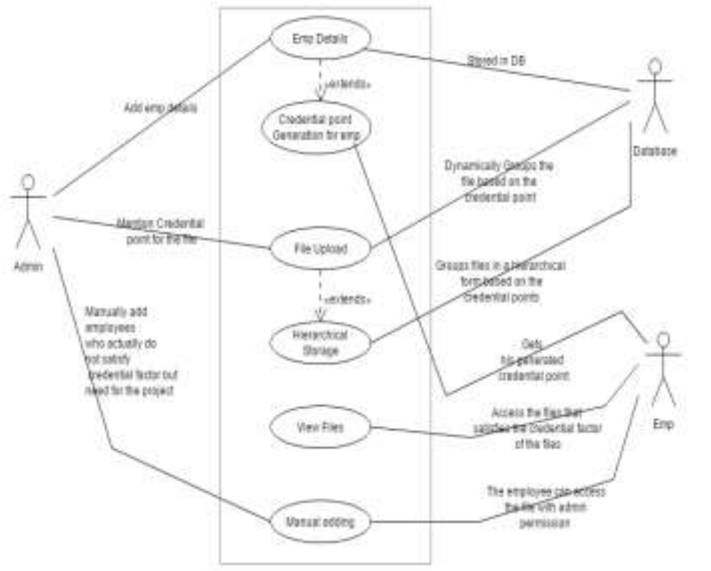

Emp Details
«extends»
Credential point Generation for emp
File Upload
«extends»
Hierarchical Storage
View Files
Manual adding
Admin
Database
Emp
Add emp details
Maintain Credential point for the file
Manually add employees who actually do not satisfy credential factor but need for the project
Stored in DB
Dynamically Groups the file based on the credential point
Groups files in a hierarchical form based on the Credential points
Gets his generated credential point
Access the files that satisfies the Credential factor of the files
The employee can access the file with admin permission

Fig.4. Use case Diagram explains the functionality of Credential based Dynamic Grouping System.

The System Architecture of our project is displayed below to give a clear-cut understanding of what is happening internally to prevent the data theft from internal hackers by generating the credential factor for both data and employees. The main elements are the Admin who adds the employee details and also uploads the data in the cloud. The Employee of the organization for whom the credential point is generated based on his/her cumulative attributes.

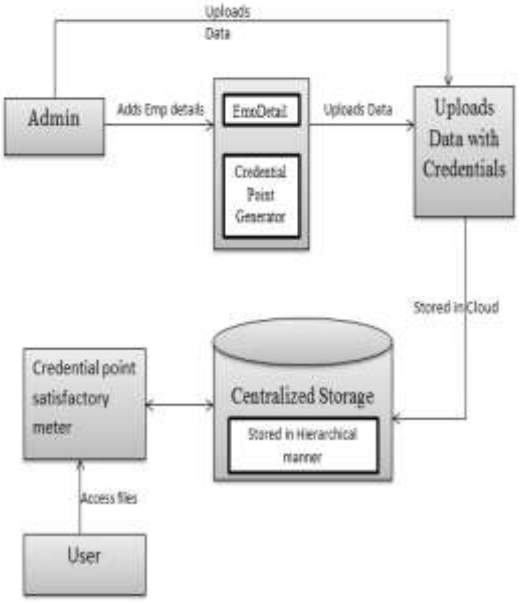

Uploads Data
Admin
Adds Emp details
EmpDetail
Credential Point Generator
Uploads Data
Uploads Data with Credentials
Stored in Cloud
Credential point satisfactory meter
Centralized Storage
Stored in Hierarchical manner
Access files
User

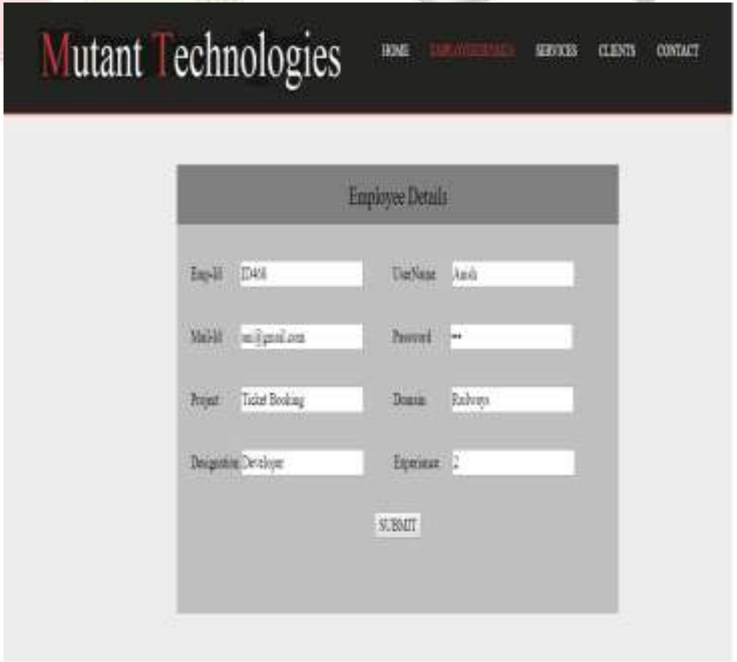
Fig.5. System Architecture explaining the functioning of the credential based dynamic grouping to prevent data theft.

Once the data is being uploaded by the employee or the admin with its credential points, it is stored in the cloud in the hierarchical format for the easy use. So, when an employee with his own credential point satisfy the credential point of the data can access that particular data and cannot access the other data whose credential point is not being satisfied by the employee.

6. IMPLEMENTATION

Initially, the admin logs in to his account he can either add employee details or he can upload the data in the cloud. When an admin adds the employee details he mentions all its attributes that describes him to be a part of the organization in order to generate the credential point for that particular employee for his details.

Below the screenshot shows the fields that have to be filled by the admin to enter the employee details.

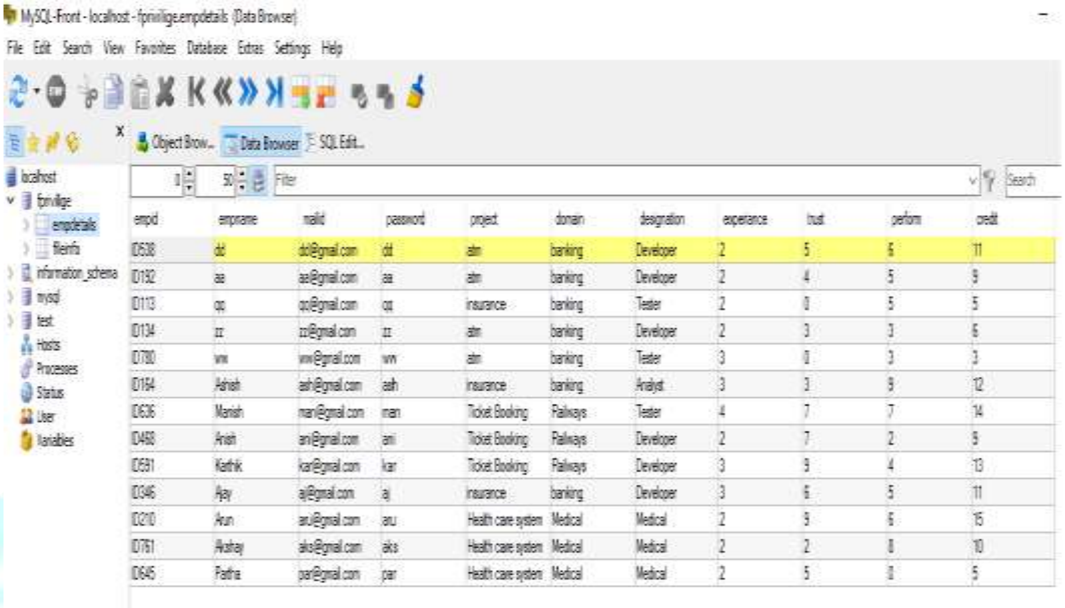


Employee Details

Emp-Id	ID458	UserName	Anish
Mail-Id	an@gmail.com	Password	**
Project	Ticket Booking	Domain	Railways
Designation	Developer	Experience	2

Fig.6. Screenshot shows the details to be entered for the employee by the admin of the organization for the credential generation.

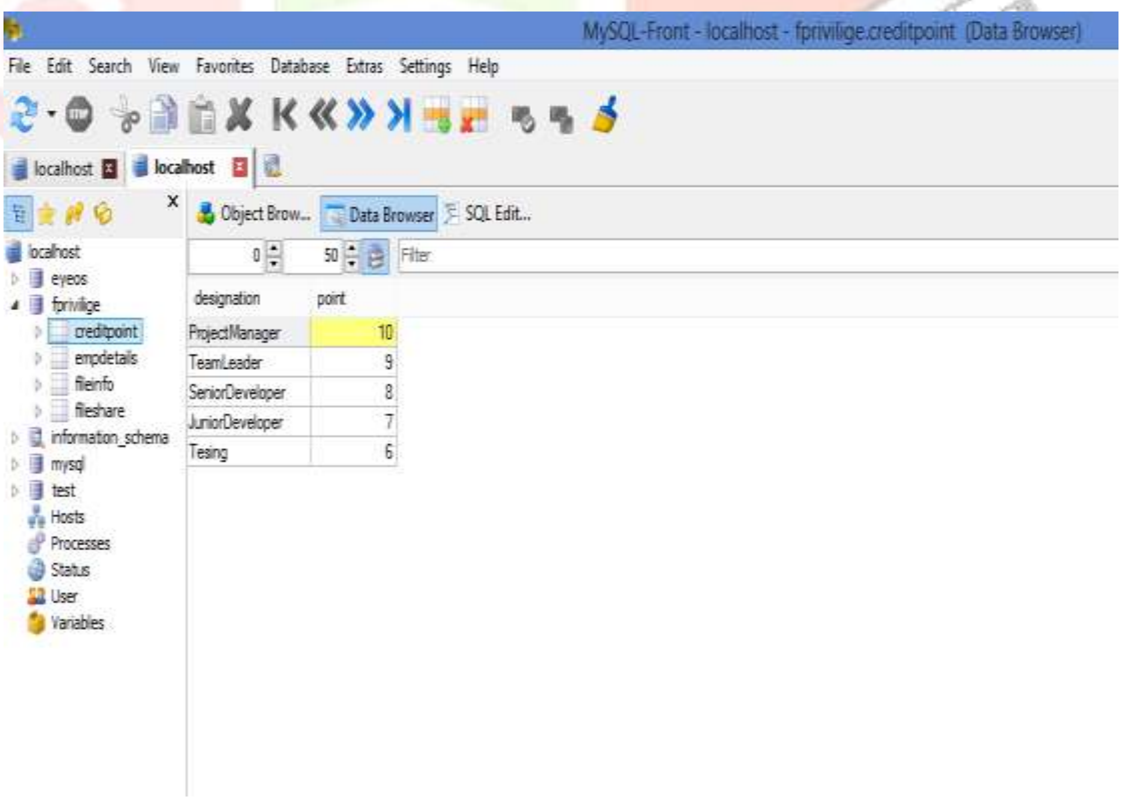
After, entering the details of the employee it gets stored in the database with credential point generation for each employee based on their attributes like designation, performance, trust index, experience. This attributes for generating credential point for employees can be changed according to organizations requirement.


MySQL-Front - localhost - fprivilige.empdetails (Data Browser)
File Edit Search View Favorites Database Extras Settings Help
Object Brow... Data Browser SQL Edit...
localhost
fprivilige
empdetails
fileinfo
information_schema
mysql
test
Hosts
Processes
Status
User
Variables
Filter
Search

empid	empname	mailid	password	project	domain	designation	experiance	trust	perform	credit
ID538	dd	dd@gmail.com	dd	atm	banking	Developer	2	5	6	11
ID192	aa	aa@gmail.com	aa	atm	banking	Developer	2	4	5	9
ID113	qq	qq@gmail.com	qq	insurance	banking	Tester	2	0	5	5
ID134	zz	zz@gmail.com	zz	atm	banking	Developer	2	3	3	6
ID780	ww	ww@gmail.com	ww	atm	banking	Tester	3	0	3	3
ID184	Ashish	ash@gmail.com	ash	insurance	banking	Analyst	3	3	9	12
ID636	Manish	man@gmail.com	man	Ticket Booking	Railways	Tester	4	7	7	14
ID468	Anish	ani@gmail.com	ani	Ticket Booking	Railways	Developer	2	7	2	9
ID591	Karthik	kar@gmail.com	kar	Ticket Booking	Railways	Developer	3	9	4	13
ID346	Ajay	aj@gmail.com	aj	insurance	banking	Developer	3	6	5	11
ID210	Arun	aru@gmail.com	aru	Health care system	Medical	Medical	2	9	6	15
ID761	Akshay	aks@gmail.com	aks	Health care system	Medical	Medical	2	2	8	10
ID645	Partha	par@gmail.com	par	Health care system	Medical	Medical	2	5	0	5

Fig.7. Screenshot shows details of the employees being stored in the database with their generated Credential point.

The Credential point for each employee is generated by taking their Designation as one of the necessary parameter. As it plays a vital role in organization level. So, we created a credential point for each level of designation in a table format and stored in the database, based on this the employee credential point is generated.



designation	point
ProjectManager	10
TeamLeader	9
SeniorDeveloper	8
JuniorDeveloper	7
Tesing	6

Fig.8.Screenshot shows the Credential point assigned for each level of designation.

Now, speaking about file uploading each employee becomes the admin for the file for which he is uploading. He mentions the credential point for the file he is going to upload.

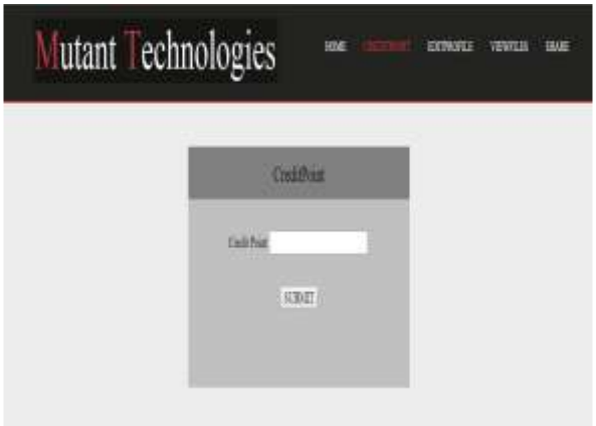
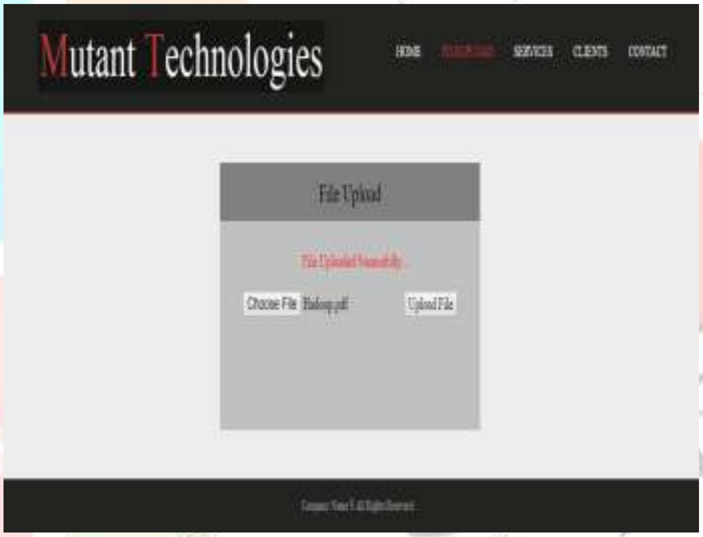


Fig.9. Screenshot shows that the credential point should be mentioned for the data the employee is going to upload.



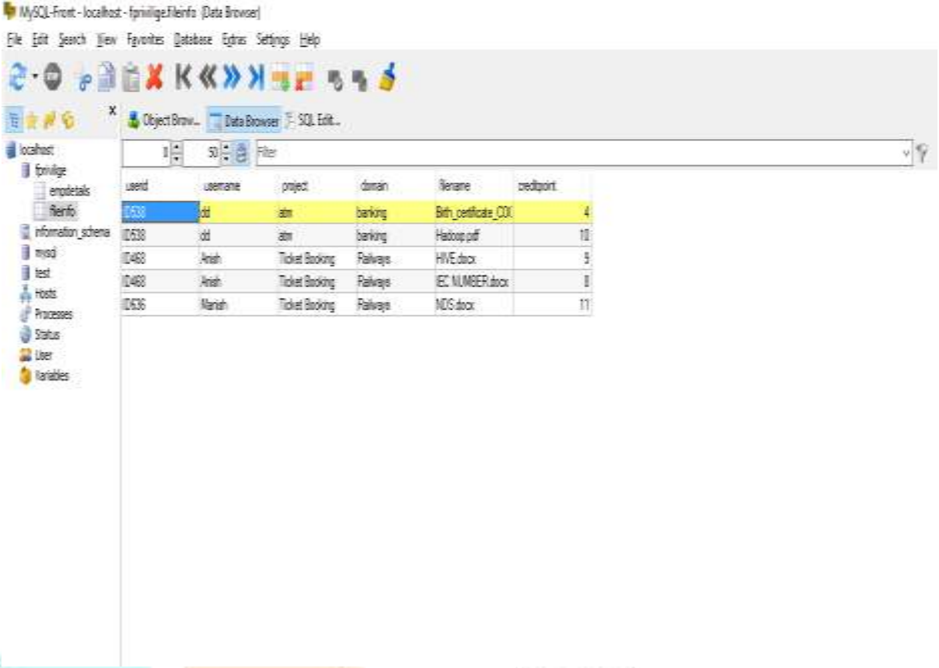
File Uploaded Successfully

Hadoop.pdf

Fig.10. Screenshot shows Successful upload of an file by the employee after mentioning its credential point.

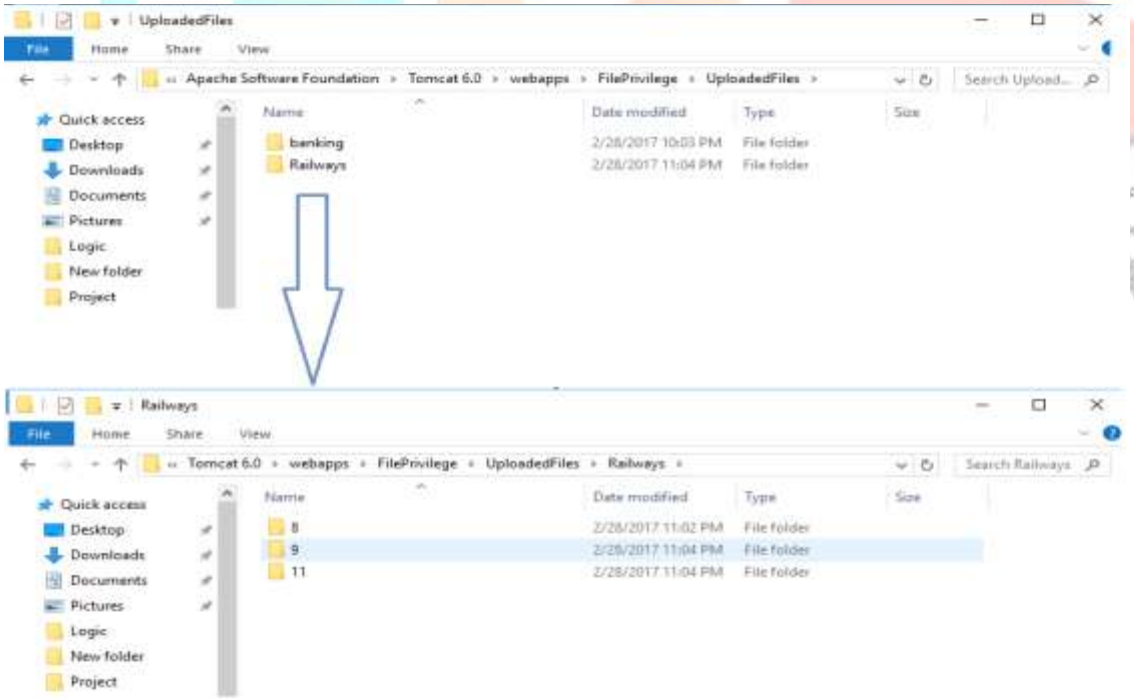
All the files that are uploaded by the employees are stored in the Database with their necessary details. The necessary details include the file name, owner of that particular file, Domain the file belongs to, file owner’s ID, Credential point of the file mentioned by its owner.

The data being uploaded is dynamically grouped and stored in hierarchical manner as mentioned earlier. First a folder regarding the domain is created then the project it belongs to is created at last within the credential point folder is created and the data corresponding to it is stored in it.


MySQL-Front - localhost - fprivilige.fileinfo (Data Browser)
File Edit Search View Favorites Database Extras Settings Help
Object Brow... Data Browser SQL Edit...
localhost
fprivilige
empdetails
fileinfo
information_schema
mysql
test
Hosts
Processes
Status
User
Variables
1
50
Filter

userid	username	project	domain	filename	creditpoint
ID538	dd	atm	banking	Birth_certificate_COI	4
ID538	dd	atm	banking	Hadoop.pdf	10
ID463	Anish	Ticket Booking	Railways	HIVE.docx	9
ID463	Anish	Ticket Booking	Railways	IEC NUMBER.docx	8
ID636	Manish	Ticket Booking	Railways	NOS.docx	11

Fig.11. Screenshot shows all the data which is uploaded being stored in the database with its necessary details.

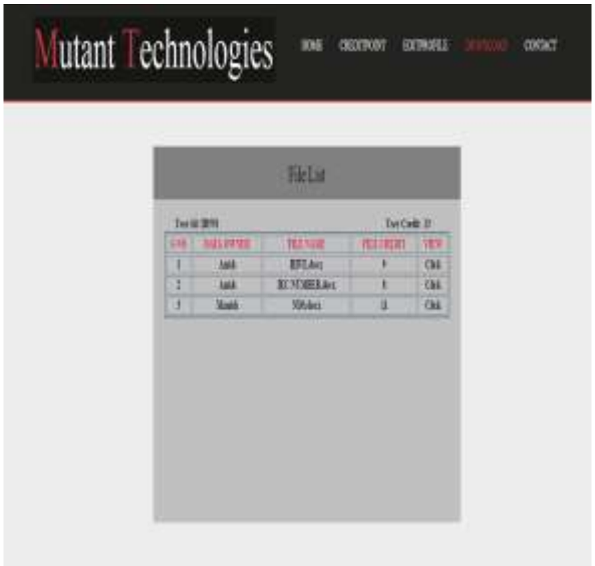


Name	Date modified	Type	Size
banking	2/28/2017 10:03 PM	File folder	
Railways	2/28/2017 11:04 PM	File folder	

Name	Date modified	Type	Size
8	2/28/2017 11:02 PM	File folder	
9	2/28/2017 11:04 PM	File folder	
11	2/28/2017 11:04 PM	File folder	

Fig.12. Screenshot shows the hierarchical storage of data based on its credential point and the domain it belongs to.

The Employee once logs into his account he can either upload the data or can access the data he is allowed to. So, we provide the employee with the list of files that he is allowed to access based on the credential point of the employee with that of the data. The employee can access the data for which he had satisfied its credential points with his credential points and can't access the other data without the data owner's permission.



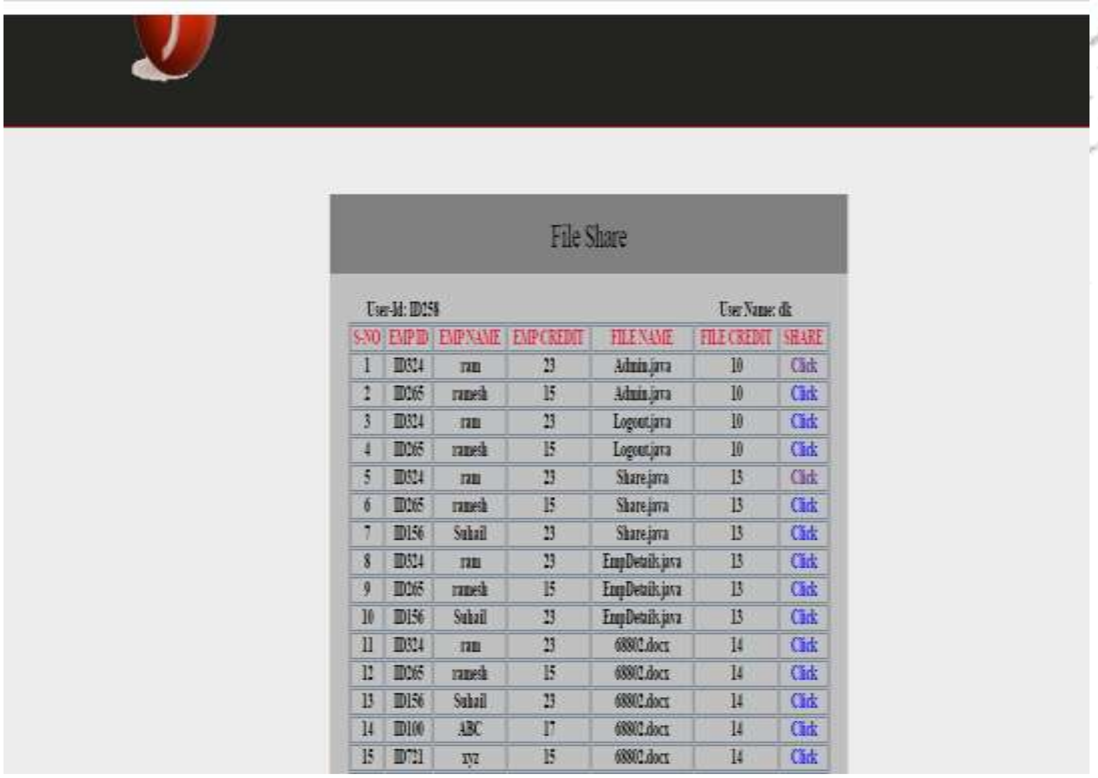
File List

User Id: [illegible] User Credit: [illegible]

S-NO	DATA OWNER	FILE NAME	FILE CREDIT	VIEW
1	Anish	[illegible].docx	9	Click
2	Anish	[illegible].docx	8	Click
3	Manish	[illegible].docx	11	Click

Fig.13. Screenshot shows the list of files that this employee can access because he satisfied the credential points of these files with his own credential point.

In a very rare case when an employee doesn't seems to satisfy the credential of the data but he is related to the project at this case the data owner if he wishes to share his data to that employee he can give the access privileges to employees he wish. This is performed by just clicking a button corresponding to the user. The employee can see who all can view the data that is uploaded by him and also who cannot. So, he can remove the privileges to the one who has before or can give privilege to the one who doesn't have before as the data owner wishes



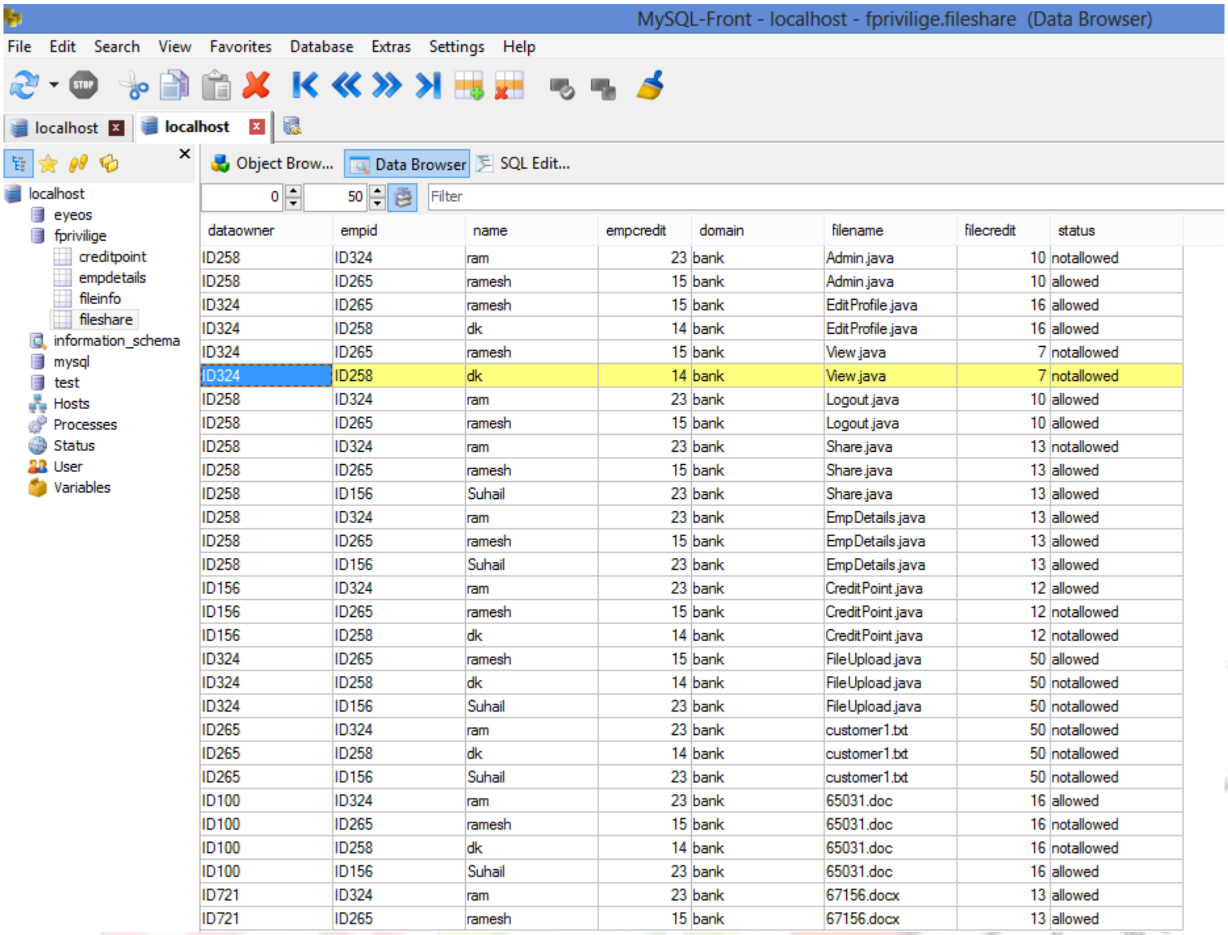
File Share

User-Id: ID258 User Name: dk

S-NO	EMP ID	EMP NAME	EMP CREDIT	FILE NAME	FILE CREDIT	SHARE
1	ID324	ram	23	Admin.java	10	Click
2	ID265	ramesh	15	Admin.java	10	Click
3	ID324	ram	23	Logout.java	10	Click
4	ID265	ramesh	15	Logout.java	10	Click
5	ID324	ram	23	Share.java	13	Click
6	ID265	ramesh	15	Share.java	13	Click
7	ID156	Suhail	23	Share.java	13	Click
8	ID324	ram	23	EmpDetails.java	13	Click
9	ID265	ramesh	15	EmpDetails.java	13	Click
10	ID156	Suhail	23	EmpDetails.java	13	Click
11	ID324	ram	23	68802.docx	14	Click
12	ID265	ramesh	15	68802.docx	14	Click
13	ID156	Suhail	23	68802.docx	14	Click
14	ID100	ABC	17	68802.docx	14	Click
15	ID721	xyz	15	68802.docx	14	Click

Fig.14. Screenshot shows the list of employees who all can view the data that the user uploaded and who all cant. And user can also give or remove privileges to employees for his data as he wish.

Providing all these features, additionally for ease and to avoid confusion a separate table is provided containing the details of files like the file name, its owner, owner's id, the domain they belong to for all files and the information regarding who all can access these files are maintained. So that for a particular file we can know who all can access it (i.e.) who are having privileges to access it and who don't have to access it.



dataowner	empid	name	empcredit	domain	filename	filecredit	status
ID258	ID324	ram	23	bank	Admin.java	10	notallowed
ID258	ID265	ramesh	15	bank	Admin.java	10	allowed
ID324	ID265	ramesh	15	bank	EditProfile.java	16	allowed
ID324	ID258	dk	14	bank	EditProfile.java	16	allowed
ID324	ID265	ramesh	15	bank	View.java	7	notallowed
ID324	ID258	dk	14	bank	View.java	7	notallowed
ID258	ID324	ram	23	bank	Logout.java	10	allowed
ID258	ID265	ramesh	15	bank	Logout.java	10	allowed
ID258	ID324	ram	23	bank	Share.java	13	notallowed
ID258	ID265	ramesh	15	bank	Share.java	13	allowed
ID258	ID156	Suhail	23	bank	Share.java	13	allowed
ID258	ID324	ram	23	bank	EmpDetails.java	13	allowed
ID258	ID265	ramesh	15	bank	EmpDetails.java	13	allowed
ID258	ID156	Suhail	23	bank	EmpDetails.java	13	allowed
ID156	ID324	ram	23	bank	CreditPoint.java	12	allowed
ID156	ID265	ramesh	15	bank	CreditPoint.java	12	notallowed
ID156	ID258	dk	14	bank	CreditPoint.java	12	notallowed
ID324	ID265	ramesh	15	bank	FileUpload.java	50	allowed
ID324	ID258	dk	14	bank	FileUpload.java	50	notallowed
ID324	ID156	Suhail	23	bank	FileUpload.java	50	notallowed
ID265	ID324	ram	23	bank	customer1.txt	50	notallowed
ID265	ID258	dk	14	bank	customer1.txt	50	notallowed
ID265	ID156	Suhail	23	bank	customer1.txt	50	notallowed
ID100	ID324	ram	23	bank	65031.doc	16	allowed
ID100	ID265	ramesh	15	bank	65031.doc	16	notallowed
ID100	ID258	dk	14	bank	65031.doc	16	notallowed
ID100	ID156	Suhail	23	bank	65031.doc	16	allowed
ID721	ID324	ram	23	bank	67156.docx	13	allowed
ID721	ID265	ramesh	15	bank	67156.docx	13	allowed

Fig.14. Screenshot shows the list of files with the details that who all can access it.

The above Screen shot shows that the employees who are all allowed accessing that particular file and who are all not allowed to.

7. CONCLUSION

Credential based dynamic grouping helps to prevent the data theft from the internal hackers in a new way of generating credential points to both the employees and the data being uploaded. The employees who satisfy the credential of the data with their own credential point are allowed to access the data and others are don't. But in case, if the internal hacker satisfies the credential factor of the data and misuses it he can be easily identified from the possible small group and can prevent the future of the organization from these sorts of mishaps. And in a rare case, based on the Organization's requirement an employee wants to access a particular data who did not satisfy the credential factor is given access manually by the organization at that particular situation. So, this Credential based dynamic grouping prevents the data theft from internal hackers and at any case if a data theft occurs it makes easy to identify the internal hacker and prevent the occurrence of these mishaps in near future.

7. REFERENCES

[1]. **A scalable, commodity data center network Architecture**. Conference paper in ACM SIGCOMM Computer Communication Review 38(4):63-74 · October 2008.

[2]. **SprintNet: A high performance server-centric network architecture for datacenter** Ting Wang; Zhiyang Su; Yu Xia; Yang Liu; JogeshMuppala; Mounir Hamdi2014 IEEE International Conference on Communications (ICC) Year: 2014.

[3]. **Hedera: Dynamic Flow Scheduling for Data Center Networks.**Conference proceedings of the 7th USENIX Symposium on Networked Systems Design and implementation, NSDI 2010, April 28-30,2010, San Jose, CA, USA.

[4]. **ElasticTree: Saving Energy in Data Center Networks.** Proceeding NSDI'10 Proceedings of the 7th USENIX conference on Networked systems design and implementation. Pages 17-17.

[5].**MicroTE: Fine Grained Traffic Engineering for Data Centers.** Conference: Proceedings of the 7th CoNEXT.

[6]. **DevoFlow: Scaling Flow Management for High-Performance Networks.**Andrew R. Curtis (University of Waterloo); Jeffrey C. Mogul, Jean Tourrilhes, Praveen Yalagandula, Puneet Sharma, Sujatha Banerjee (HP Labs), SIGCOMM 2011.

[7]. **Floodless in SEATTLE: A Scalable Ethernet Architecture for Large Enterprises.** Conference paper in ACM SIGCOMM Computer Communication Review 38(4):3-14. October 2008.

[8].**VL2: A Scalable and Flexible Data Center Network.**Albert Greenberg James R. Hamilton Navendu Jain SrikanthKandulaChanghoon Kim ParantapLahiri David A. MaltzParveen Patel SudiptaSengupta. Proceeding in SIGCOMM'09 OF ACM SIGCOMM 2009.

[9]. M. Al-Fares, A. Loukissas, and A. Vahdat. A scalable, commoditydata center network architecture. In SIGCOMM, 2008.

[10]. M. Armbrust, A. Fox, R. Gri_th, et al. **Above the Clouds: A Berkeley View of Cloud Computing** UC Berkeley TRUCB/EECS-2009-28