International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Enhancement in Security of Backup and Recovery Based System

Bhavya Bhatia[1], Dr. Shobha G[2]

[1]*Department of CSE, RVCE, Bengaluru, Karnataka, India*
[2]*Professor, Department of CSE, RVCE, Bengaluru, Karnataka, India*

Abstract - *The development of data backup and recovery technology guarantees the reliability and availability of the data. In any case, just reliability and availability cannot meet the security needs of data backup and recovery in significant zones of military, national defense and other high security areas of usage. In this paper, a security layer is proposed over the backup and recovery system. This security layer is designed through the improved basic access authentication, token-based authentication, single sign on authentication and password guessing attack prevention for ensuring that only the authentic users to have access of backup and recovery services. The framework is able to understand and recognize the authenticity of data which is backed up, and ensures prevention from data leak by forbidding the operators to back up data on a illegal destination or restore data on an illegal target machine.*

***Key Words*: backup, recovery, confidentiality, security**

1. INTRODUCTION AND BACKGROUND

Data storage is a significant part of the consistent advancement of information technology. To prevent data against any loss, data backup and recovery plays a crucial role. So as to take care of data availability and reliability, the researchers have come up with various techniques of backup for database, file system, virtual machine and operating system.[1] There are various techniques to implement backup and recovery-based system like the XOR based backup and recovery technique [3], copy-on-write and redirect-on-write snapshot. [2]. While discussing about the implementation details of backup and recovery system, there are different types of backup techniques. According to the amount of data to be backed up, there are different types of backup types namely full backup, incremental backup and synthetic backup. Full backup means taking the backup of the entire data which requires large disk space to store the data. An incremental backup is a sort of backup that lone duplicates information that has been changed since the previous backup activity was completed. Synthetic backup is created by merging the previous and older full and incremental backups. It is evident that the disk space required in incremental backup is less than the disk space required in the full backup. In terms of complexity implementation of incremental backup is more complex as compared to full backup.[4][5] Key techniques such as data compression and data de-duplication [5] have come up in recent years to make the backup process space efficient. These specified technologies have come up in the recent years to provide enhancement in reliability, availability and performance of data backup and restore systems.

Due to demand from the industry, the above mentioned techniques only focus on reliability and availability of data without concentrating much on the confidentiality if the data. In this paper, through the enhanced authentication based on improved Basic access authentication, token-based authentication, single sign on and prevention of password guessing attack, we design a security layer on top of a data backup and recovery system that ensures the legitimacy and authenticity of data and can be used for higher data confidentially operations in military and defense.

This paper is organized as follows. The first part is the introduction and the background. Section 2 explains the design of the system and section 3 talks about the implementation of the system. Finally, we give the summary.

2. DESIGN OF THE SYSTEM

2.1 Architecture of the system

The system consists of various components namely the management console, backup and recovery server, the backup and recovery module and the authentication service. The management console is an interface from where the user can create the backup and recovery tasks and track the status of the tasks accordingly. The management console interface can be web based or command line. The backup and recovery server serve the requests that are received from the management console. These requests are related to backup, recover or fetching the details of the tasks initiated. The basic backup module which is in the server is responsible for start the task of backup in the specified storage place. The basic recovery module recovers the data stored in the storage devices to the target machines. The backup or recovery module can handle data from file system, database, operating system and virtual machine. These four components from where data is to be taken for backup needs to be registered in the backup and recovery server. For ensuring the legitimacy and authenticity of the data, it is made sure that the data is recovered on the registered components only. The authentication service is a web-based service, which is installed on the backup and recovery server. This service is used for registering the authentic modules on to the server. The authentication service provides the backup and recovery system security services such as basic access authentication, token-based authentication,

single sign on and password guessing attack prevention. Basic access authentication verifies the credentials of the user and stored the credentials in an encrypted format in the database. The token-based authentication makes use of signed token generated using private key which provides users to securely connect to the backup and recovery module and the management console. Single sign on is an authentication scheme that allows the user to login with credentials at a time on several related consoles. Password guessing attack prevention prevents guessing of password of an authorized user by several login attempts.

The backup and recovery server when once register the different types of components from which data is to be backed up, then it starts serving the backup or recovery requests. The backup tasks are created by the authorized user in the management console and are received by the server. The server serves the corresponding requests and initiates the backup in the authorized storage spaces. While recovering of the backed-up data, the recover tasks are created in the management console. The server receivers the recover requests and initiates the recovering of data in the authorized machines. The data can be recovered in the same components- Virtual machines or Database or operating systems from where the backup was taken. In case of data loss, the data can be restored the same machines. The overall architecture of the proposed system is described in Figure 1.

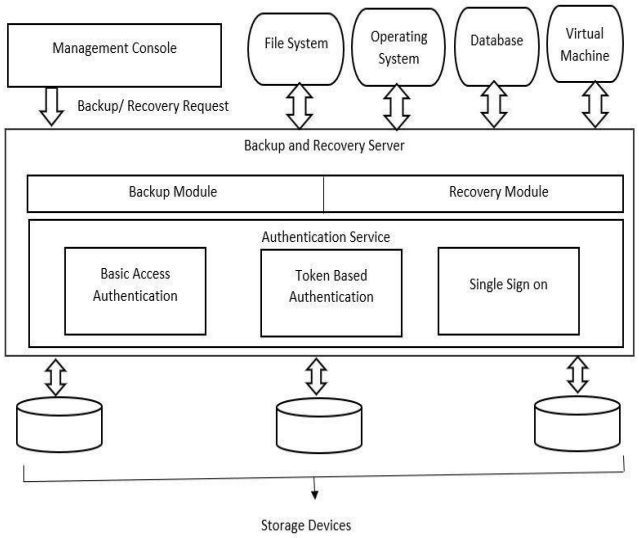

Fig-1: Architecture of the system

2.2 Basic business function

The basic business function of the system is to serve with backup and restore service for the components such as file system, virtual machine, database and operating system. For all these modules apart from basic backup and recovery some advanced features like data-compression, resume backup from the break-point, data de-duplication, consistency detection, and data encryption are also covered. Users can manage all the above features through a web interface or command line which is the management console. This web interface and command line allows users to initiate backup or recovery tasks and also keep a track of the status of the tasks initiated.

2.3 Authentication and added security

While taking recovery of the data, it is to be ensured that the data is not recovered in an illegitimate machine and hence resulting in data loss. The security has to be enhanced in order to ensure that there is no data loss. The authentication service is the main module that handles security in this system. It ensured that the data which is recovered is at a legitimate machine and ensures that there is no data leak. There are three different ways in which security is ensured.

1. **Basic Access Authentication:** It is a simple username password authentication in which a user passes the username and password to the server to authenticate. An unauthorized access to the database can expose user credentials, so there is a need to make this more secure. For ensuring the security, the password is stored using SHA256 encoding [6] in the encrypted format. This one-way secure hash algorithm helps in adding security to this method. This authentication method helps in making the password of the user secure.

2. **Token Based authentication:** It is a more enhanced form of authentication [8] which makes use of private key and public keys. This helps user to securely connect to the backup and recovery application. The machine from where data backup or recover has to be taken is referred to as client machine. A trust is made between the client and the server which the help of singed token. With help of this the server knows on which client the data is recovered and data leak can be prevented.

3. **Single Sign on Authentication:** It is an authentication technique that is used in the architecture of this system which allows the user to login only once with username and password on various management consoles. Single sign on makes use of the authentication verification data and allows the user to login to different platforms of the management console using single sign in [9].

4. **Password guessing attack prevention:** This technique is used for preventing the unauthorized access to the backup and recovery server. An unauthorized access to the server can be made by guessing the user password in several attempts.[10] To prevent this we may prevent the user to make several login attempts and prevent the guessing of password. If more attempts are made to login and the activity looks suspicious then the user is blocked.

3. IMPLEMENTATION OF THE SYSTEM

This section talks about the implementation of each of the authentication and security function which are used in this model.

i. **Basic Access Authentication**: The password of the user is encrypted using the SHA256 algorithm.[6] The password is sent as an input to this hash function to generate a hashed text. This is a very secure one-way hash algorithm which means that the text can be encrypted using this hash algorithm but from the hashed text it is difficult to get the original text. Also, a small change in the input text may completely change the hashed text. So, it is very difficult to get the original input from the hashed text. In the database of the server, the password is stored as hashed text using SHA-256. In this way even if someone get an unauthorized access to the backup and recovery server, still it is not possible to guess the original password from the hashed text.

ii. **Token based authentication:** For implementation of token-based authentication, JSON web token (JWT) [7][8] are used. It is a secure way of transmitting information between client and server using a JSON object. The registered user will make a request with its credentials from the web console. The browser sends a POST request to the server with username and password. The server creates a JSON Web Token (JWT) [8] token secretly. The server returns the JWT token to the browser. The browser then sends the JWT token which is signed with the user information in the authentication header. The server then checks the JWT signature and gets the user information. After verifying the user, the server sends the response back. The establishing of trust between the user and server is displayed in figure 2.

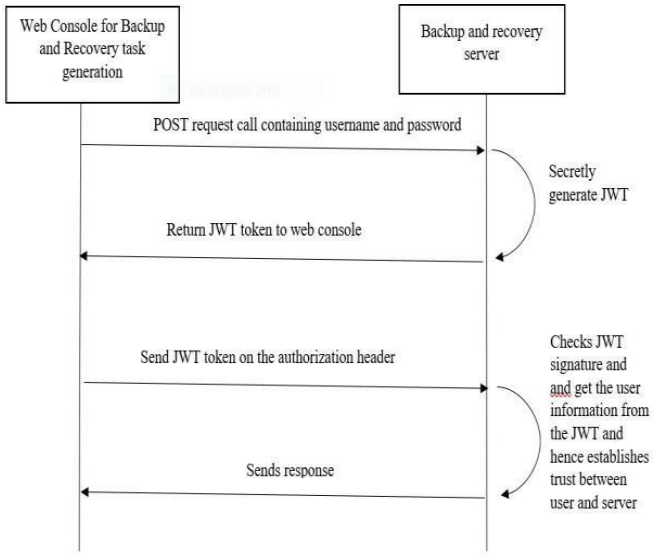

Fig-2 Establishing trust between user and backup server using JWT token

iii. **Single Sign on Authentication:** Single sign on authentication in this model helps user to log in only once for different types of related interfaces of backup and recovery. The user can log in only once for even different management consoles like command line interface and web interface. This process of authentication is done through Security Assertion Markup Language (SAML) token.[11] This token is used authentication the data. In this process there is an authentication server as well between the interaction of user and backup server. The steps of single sign on authentication are as follows:

- The user sends a login request to the authentication server from one of the management consoles.
- Authentication server receives the login request and generates a SAML token which is digitally signed and contains the information of the user such and username and password in encrypted format. Also, this token is digitally signed.
- The SAML token is sent back the user.
- User tries to make a login request from a different application which means a different management console. In this case in the request user sends SAML token in the request.
- Authentication server decodes the request and extracts information such as login credentials, SAML token and the session Id of the user.
- The decoded information is sent to Backup and recovery server, which validates the session id of the user and sends the response to the user.
- Based on the response of the server the login request succeeds or fails.

Figure 3 describes the single sign on process through the use if SAML tokens.

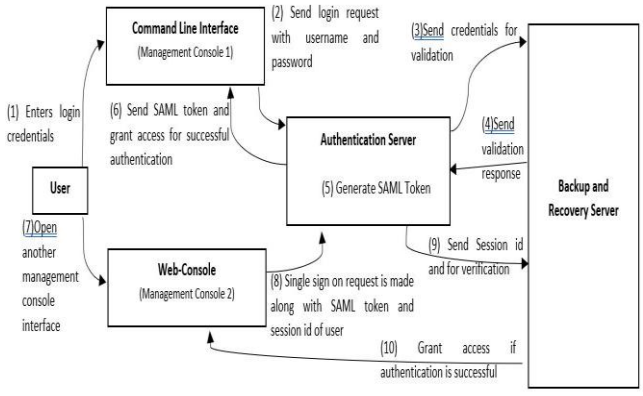

Figure 3 Single sign on process with SAML token

From the described process, the need to login again for different management consoles can be avoided and the user is granted access by logging in just once.

iv. **Password guessing attack prevention:** There can be an attempt to get the user credentials by guessing the password and trying many login attempts. For this the user is limit to a specific number of login attempts. If the user tries to make login attempt more than the set limit, the user is blocked. The logs of the user can be used to get the track of the number of login attempts.

4. CONCLUSION

Based on the current backup and recovery systems which mainly focus on data reliability and availability, this paper puts light on another kind of important aspects. Except to fulfil data reliability and availability many fields require the system that can ensure data confidentiality requirements which may come from military and defense fields. By strengthening the basic authentication method, the security of the system can be advanced. The token-based authentication helps in establishing trust between the user and the backup server, so that in no case the data is recovered on an illegitimate machine which may lead to data loss. The single sign on technique prevents the user to indulge in the monotonous task of validating credentials for each interface of the management console. Lastly, the password guessing attack prevention ensures that there are limited number of login attempts so that suspicious activity of guessing the password can be prevented.

Overall by adopting the security enhancements to the backup and recovery system may help ensuring the data confidentiality by preventing recover of data in the illegal target machine and hence prevent data loss.

ACKNOWLEDGEMENT

This study and research are carried out in R.V. College of Engineering, Bengaluru, Department of Computer Science and Engineering, under the guidance of Dr Shobha G. I also thank our Head of the Department Dr. Ramakanth Kumar P and the Principal of the institution, Dr. K N Subramanya for providing us all the required facilities and suitable environment to successfully complete this research.

REFERENCES

[1] J. Zhang and H. Li, "Research and Implementation of a Data Backup and Recovery System for Important Business Areas," 2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, 2017, pp. 432-437.

[2] W.J. Xiao, Q. Yang, J. Ren, C.S. Xie, H.Y. Li, "Design and analysis of block-level snapshots for data protection and recovery", IEEE, Trans.Comput. 58 (December(12)) (2009) 1615–1625.2.2

[3] Raje, Manali & Mukhopadhyay, Debajyoti. (2015). Algorithm for Back-up and Authentication of Data Stored on Cloud.

[4] Zhu, B., Li, K., Patterson, R.H. "Avoiding the disk bottleneck in the data domain deduplication le system", in Proc. 6th USENIX Conference on File and Storage Technologies, California, 2008, pp.269-282

[5] Quanqing Xu, Liang Zhao, Mingzhong Xiao, etc, "YuruBackup: A Space-Efficient and Highly Scalable Incremental Backup System in the Cloud", International Journal of Parallel Programming, vol 43, Jun. 2015, pp. 316-338, doi:10.1007/s10766-013-0280-7.

[6] W. A. Ali, N. M. Sahib and J. Waleed, "Preservation Authentication and Authorization on Blockchain," 2019 2nd International Conference on Engineering Technology and its Applications (IICETA), Al-Najef, Iraq, 2019, pp. 83-88.

[7] S. Ahmed and Q. Mahmood, "An authentication-based scheme for applications using JSON web token," 2019 22nd International Multitopic Conference (INMIC), Islamabad, Pakistan, 2019, pp. 1-6.

[8] M. Haekal and Eliyani, "Token-based authentication using JSON Web Token on SIKASIR RESTful Web Service," 2016 International Conference on Informatics and Computing (ICIC), Mataram, 2016, pp. 175-179

[9] M. D. Karunanithi and B. Kiruthika, "Single sign-on and single log out in identity," International Conference on Nanoscience, Engineering and Technology (ICONSET 2011), Chennai, 2011, pp. 607-61

[10] R. Kirushnaamoni, "Defenses to curb online password guessing attacks," 2013 International Conference on Information Communication and Embedded Systems (ICICES), Chennai, 2013, pp. 317-322

[11] F. Nie, F. Xu and R. Qi, "SAML-based single sign-on for legacy system," 2012 IEEE International Conference on Automation and Logistics, Zhengzhou, 2012, pp. 470-473.