SSO versus MFA: A comprehensive study on Big Data Security

Raheem Qaiser[1], Aqeel Khalique[1], Imran Hussain[1]
[1]*Department of CSE, SEST, JamiaHamdard, New Delhi 110062, INDIA*
qaiser.rahim9@gmail.com
aqeelkhalique@gmail.com
ihussain@jamiahamdard.ac.in

***Abstract*—**In cloud environment, access to cloud services is authenticated using several authentication mechanisms. Two different mechanisms are widely used namely single sign-on (SSO) and multi-factor authentication (MFA). SSO removes the need to sign-in to individual services of a company or system by providing a single sign-in interface managed through access control to utilize the authorized services. MFA asks users to provide two or more types of information to verify access into the authorize service. In this paper, we studied MFA and SSO as enabling security service providers and highlights their characteristics based on volume of data being used in cloud environment. Further, we study SSO and MFA implemented by major cloud based service providers handling big data. The paper also aims to indicate the preference of the authentication method to be used depending on the level of security required and how it affects security in cloud environment and big data applications. The information thus reviewed can act as recommendations to deliver secure access to information systems, by taking care of the configurations and protocols at the user and product's end.

Keywords*—*Single Sign On, SSO, Multi-factor Authentication, MFA, Big Data Security, Cloud Security

I. INTRODUCTION

"Big data refers to data sets whose size is beyond the ability of typical database soft-ware tools to capture, store, manage and analyze "[1]. This data is collected over time in companies and it provides business insights. More than 2.5 quintillion bytes of data are created every day around the world [2]. Enormous amount of data is shared every day on mobile phones, laptops, PC's through the Internet. Vast amounts of data is collected through banking systems, hospital and other medical systems, ISP's and, various smartphone and desktop applications. However the collection of such enormous amounts of data can lead to privacy intrusions, thus bringing into consideration security measures for such Big Data. Big Data Security is a big concern for many organizations and therefore these threats are being acknowledged and measures are being taken for prevention. According to the global report, World Quality Report 2015-1016, the most highly ranked priority in IT strategies is Security.

Rising security issues for storing and using Big data, has led to the use of cloud infra-structure due to its vast storage capacity. Cloud computing is being endorsed by many organizations because of the various benefits it provides. It allows higher productivity, improved efficiency, availability of services at all times, cost- effective solutions, and quick deployment. For big data it brings new methods of big data manipulation and provision of online services in the form of Software as a service (SaaS), Platform as a service (PaaS) and Infrastructure as a service (IaaS). But there is a risk to place the big data in a public space. A single organization opting for a private cloud making use of its own resources has a rather low level of security concern, except for attacks from the inside. Whereas organizations opting for a public cloud, where all the data is maintained by a Cloud Service Provider (CSP) and stored in their server, will be vulnerable to data breaches, data loss, account hijacking, denial of service attacks and phishing sites. So technologies like single sign on (SSO) and multi-factor authentication (MFA) are used as primary security standard for cloud services for their highly secure methods of providing access.

II. SINGLE-SIGN ON (SSO)

Random Single sign-on type of authentication seeks to remove the need to sign-in to individual products of a company or system by providing a single sign-in gateway to utilize all of its products or sub-systems. Let us take the help of an example to understand the basic functioning of SSO. Single sign on utilizes a central service to authenticate access into multiple clients, for example Google, where this central service is Google accounts. When a user logs into a Google account, a cookie is generated that stays with the user as it moves from one Google service to another.

This feature is not necessarily confined under a system or a company, SSO can also enable a user to login once and get access to various resources and applications over different domains and network. For example, take into consideration a customer's network/domain and an application (say

Edumatic)[15]. A user authenticating into the customer's domain while logging in to his/her computer, will give the user access to the different network resources(network shares, printers, webpage logins, et cetera) within the domain without requiring the user to authenticate access through entering his credentials each time.

A. Mechanism then(without SSO) vs now(with SSO)

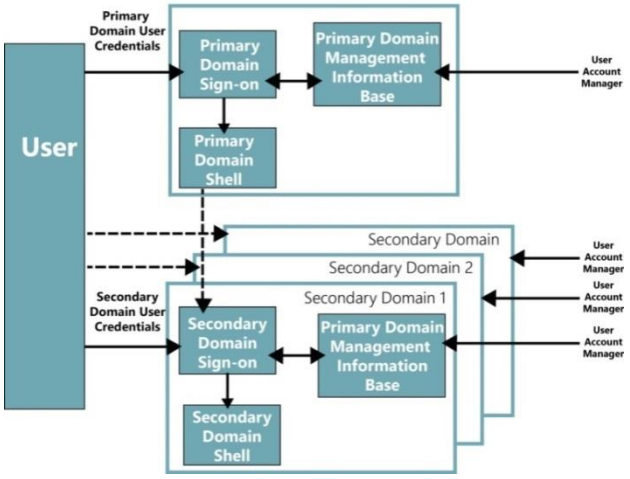

Fig 1,Authenticating an account in the usual manner without the use of SSO

The components in the above diagram act as domains which are independent in a sense such that the end-user has to authenticate his identity independently. The user first transact with a Primary Domain for establishing a session with the primary do-main. The above diagram terms this as the Primary Domain Sign-on, and it needs to be given some credentials like a username and password applicable to the primary domain. An operating system session shell represents the primary domain session that is executed on the work station of the end user in an environment representative of the end-user. The services of the other domains, such as applications or platforms are then ready to be invoked from the primary domain session shell [3].

Further to access the services of the secondary domain an end user has to perform a similar Secondary Domain Sign-on, making the user provide further a new set of user credentials. This scenario makes the user remember separate credentials for both the domains, also requiring management of each domain independently [3].

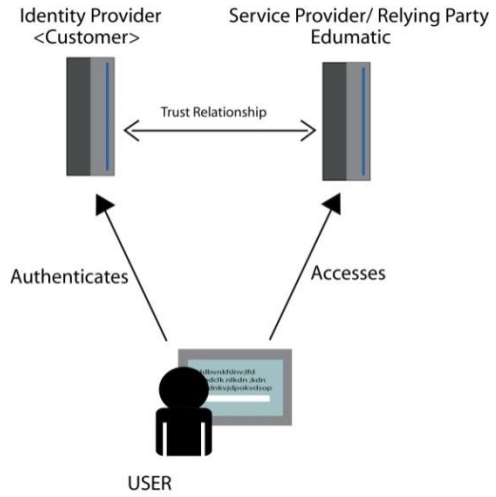

Fig2, User accessing various websites using SSO service with the help of Identity Provider

Authentication in SSO is based on the trust between the domains. In Single sign-on approach the credentials accepted by the user for the primary domain sign-on which are also supported for authentication into the secondary domains which the user might wish to access. The SSO service then authenticates the user's presence through the identity provider or authentication system used by the company. After logging in, authentication verification data is passed by the website as the user moves through the site to authenticate each time he or she moves to a new page [4].

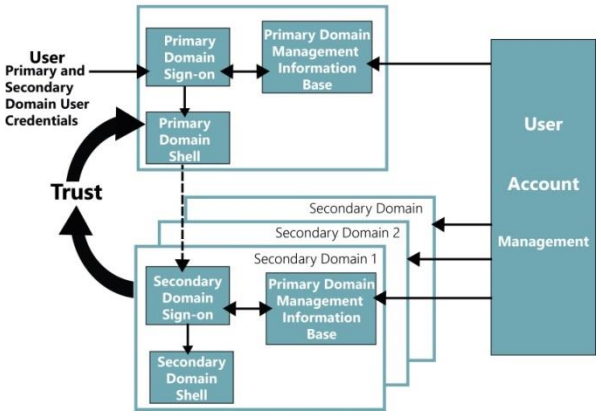

Fig 3, Authenticating access into several websites using SSO

B. Features of SSO

- Ease of use:
  The user experience is improved because the user is not interrupted by credential requests. Once the user is authenticated to the primary domain, access to its other services is also authenticated. The access to all

the applications is seamless because of the sign-on automation.

- Mobile Support:
  With the evolution of technology, mobile phones have become very important and users usually access work related applications through their mobile phones. Thus SSO software enables users to utilize its functions on a mobile phone [5].
- Pay-As-You-Go pricing:
  Usual SSO products charge based on users, based on pay-as-you-go pricing making SSO cost effective and helps manage the investment in the tool [5].
- Reports and analytics:
  SSO products takes a note on which users uses which application, their signing in pattern etc. which can be utilized to identify unusual patterns thereby avoiding unauthorized access to applications. The reports and analytics help the member's part of the IT department in a company to observe important company insights [5].
- Increases User's productivity:
  Signing into a website takes significant time; also remembering the credentials is a cumbersome task. More websites means more management of credentials and more signing in time. With no need to sign in individually to independent services or applications, the productivity of users increases through the use of Single sign-on.
- SSO increases IT department's productivity:
  Large organizations have thousands of users, thus if each user had an account for each individual service of the organization, it would be a headache for the user to maintain the list of credentials. Also if the user forgets his credentials, the people held in charge to manage and reset passwords are imposed with many reset password requests. But with the help of SSO, the many number of credentials are replaced by one set of credentials, thus reducing work on both ends.
- Secure**:**
  During the implementation of SSO, the authentication processes and elements are carried out by the identity provider. These providers are usually reputed big companies (Yahoo, Google, and Amazon) which have tight security measures. So a cyber crook trying to gain illegal access to a website associated with SSO authentication mechanism, will have to go through the robust security measures of the reputed companies and not the individual company. Therefore SSO is very secure and reliable. Strong policies are associated with the authentication which asks the user to use strong passwords that make use of a combination of uppercase and lowercase letters, numbers, and non-alphanumeric characters. In usual centralized SSO, there is trust, because only one company and one security domain is involved. In other SSO systems, security is dependent upon strong encryption for authentication or on trust relationships [6].

- Facilitates B2B collaboration:
  Not all products are made by just one company, some are a collaborative work. So in such scenarios where user has to log in to applications belonging to their own organization or their business partners, maintaining the authentication and authorization mechanism can be cumbersome. With SSO, businesses are able to centralize the management of authentication and users just have to login once and gain access to all the applications of the organization as well as the business partners [7].
- Features Based on Volume of data used:
  The level of security and the volume of data to be used depend upon the kind of business we are taking under consideration. We take SSO solutions available online as examples to understand the features that an SSO provides to its users.
  OneLogin SSO service is best suited for small businesses. This service is easy to use, supports multiple directory types, has a user-friendly access portal, and offers extensive control to administrators. The pricing of the service is what determines which business shall use it. Since this service costs $4 per user per month with a minimum of 10 users, is suitable for small and mid-sized company. This SSO offers a good level of control to those monitoring the system and also provides features like self-registration pages and group management mapping tools.
  Okta is a service suitable for big enterprise, providing integration of wide variety of directory types, detailed reports and a user-friendly interface. All these features are grand to provide more administrative control over the applications and manage huge amounts of data, thus suitable for big enterprise [8].

C. Why SSO

Due to the rise in SaaS cloud based applications, the need for different usernames and passwords has increased. With more and more credentials to learn the user's password hygiene worsens, some start using the same password everywhere and often with time the passwords become less complex compromising their accounts. From the employee's point of view, a lot of time is wasted by logging in and logging out of services and losing a password means resetting it, making it a vicious cycle of never remembering a certain password thus affecting employee productivity.

With the mechanism of Single Sign-on, with the help of single token containing credentials the IT companies could enable or disable a user's access to its multiple systems, applications and various resources. Since the need for many usernames and passwords is replaced by one set of credentials, there is a decrease in risk of lost, forgotten or weak passwords. Less password resets reduces the cost related to the transaction for the same and also reduces the number of helpdesk calls for password resets. It improves the security as less number of credentials is at risk. A big problem for a company is when an employee leaves, it has to revoke his/her access from multiple platforms, but with the help of SSO companies just have to do it once. Adding new users is also simple with the help of SSO.

D. How SSO affects security?

For looking for the drawbacks of utilizing SSO scheme we see the trust relationship between the company and the SSO service provider, and who the SSO service provider is. People who do not want their personal information to be shared would take a step back before utilizing a third-party application. SSO's are prominent targets for hackers and any data theft could be disastrous for the users. Since, the company puts its entire trust of its resources and applications on the SSO software, SSO's have big responsibilities so they are backed up with excellent security measures. If the SSO provider goes down, users would be unable to utilize the websites associated with this service [9].

There are further more risks involved. Just like it is a bad habit to give the same credentials to different web services, it is also dangerous to have one username and password authenticates all the services and resources of an individual. If a hacker is able to get access to these credentials then he can access everything that the employee can. To avoid such situations, having a separate credential for each web service is advantageous since limited amount of data will be breached if the password gets stolen. If the employee forgets his password, he would not be able to access his resources which would mean loss of productivity and he or she would have to wait until he is able to contact a helpdesk for a password reset and regain access to the resources [10].

III. MULTI-FACTOR AUTHENTICATION (MFA)

MFA stands for multi-factor authentication in or sometimes it is usually referred to as two-factor authentication or 2FA which is a subset of MFA. It's a mechanism to enhance security that allows a user to present two pieces of information (their credentials) when they log in to an account. The credentials that could be asked for can be categorized in three categories:

- something the user knows (like a password or a PIN)
- something that the user has (like a smart card)
- something that the user is (like the user's fingerprint)

The policy is that the credentials should come from two different categories, so authenticating using two different passwords will not be considered as multi-factor [11].

Let us take a simple example from our day to day life of logging in to a bank account to explain how MFA happens. If the user has turned on multi-factor authentication or the bank has turned it on for the user, then the process of logging in would go differently. Typically the user would have type in his or her username and password. Then the user will be welcomed with an authenticator app, which will generate a one-time code to be entered on the next screen as a second factor of authentication. Fulfilling these steps will authenticate the user's access to the account. Most MFA services will remember the device used to log-in, so the next time the user uses the same phone or computer the MFA remembers the device as the second factor [11].

A. Authentication Factors

- Knowledge factor ("something only the user knows"):
  It is the most commonly used form of authentication. In this form of authentication the user is required to make use of a string of characters as a password which is secret to his knowledge only. It can also be a PIN (A personal identification number) or a pattern (Pattern is a regular or stochastic sequence or array of sets of information as) [12].

- Possession factor ("something only the user has"):
  This factor is used usually used as the second form of authentication along with a password, where an item possesses by the user is used to authenticate his presence. The item can be a pocket-sized authentication token which display a changing passcode on an LCD or e-ink display, which must be typed for authentication, avoiding the need for an electronic connection. These tokens can be time-based tokens or sequence based tokens [12].

  Some of the common tokens are:
  - Connected tokens: Magnetic stripe cards, Smartcards, Wireless RFID-based tokens, USB tokens and Audio Port tokens.
  - Soft tokens (computer-simulated software-based tokens): the functionality of a token can be emulated by a PC or a smartphone where software installed could enable the device to become the possession factor.
  - One-time pads: Itis a password used only once.
  - Mobile phones: Mobile phone can act as a token device using SMS messaging, phone call or an application on the smartphone.
  - SMS one time password: SMS service could be utilized to send one time passwords.

- Smartphone push: Push notification services like Android's C2DM/GCM could be used to produce a token.
- Mobile signature: Mobile signatures are digital signatures which are produced on a SIM card on a mobile device securely using a user's private key. The text to be signed is securely sent to the SIM card used by the mobile phone. The SIM then displays the text to the end-user who checks it before entering a PIN code which creates a signature. This signature is then sent back to the service provider. The signature can be verified using standard PKI systems.

- Inherence factor ("something only the user is"):

1. Biometrics:

Biometric authentication is the strongest form of authentication as it is unique to its user. Users can authenticate biometrically via their fingerprint, voiceprint, or iris scan provided that suitable hardware is available [12].

B. *Features of MFA*

- Strengthens Security:
  The principle of MFA is that each factor is compensated for its lack in security by the other factors. For example, username and password entered by a user could be susceptible to brute-force attacks or social engineering attacks. So supplement of another authentication factor like "something that a user has" by authenticating through their mobile device or "something the user is" like a biometric factor like fingerprint or voice can be used. So now the hacker has to have all the factors covered up to hack, increasing the difficulty of hacking [13].
- Compliance:
  A lot of compliance standards – federal, state or otherwise – usually specify that organizations require implementation of MFA for certain situations where personal information or financial information is involved [13].

- Simplifies login process:
  MFA might seem like a time consuming task but most companies opting for MFA also opt for SSO authentication mechanism. So users signing in are given the advantage of utilising many services at once [13].
- Increase flexibility and productivity:
  MFA provides flexibility of choosing the factor types and productivity is improved because the burden of passwords is replaced with alternatives.

C. *Why use MFA*

Identify theft is a crime that is easy, has low risk with high rewards and is a major threat to all businesses. It is one of the fastest growing types of crime. From 2013 to 2014, the number of successful breaches went up by 27.5 percent [14]. MFA helps protect users by adding an addition layer of security, making it hard for hackers to breach user's accounts. Stopping online crimes completely is impossible but simple steps can be taken to reduce the chances of being hacked. MFA should be opted for whenever possible especially when it comes to one's personal data like primary mail, financial accounts and health records. While some requires its user to use MFA, many offer it as an extra option that can be enabled. It is advisable that users should take an initiative to turn on MFA. It is also advisable to use services which give access to personal information online, to be used only if it offers MFA.

D. *How it affects security*

Despite all the benefits MFA offers, it is still not 100% secure. Authentication via text message is vulnerable to interception and spoofing by hackers. Sometimes sophisticated malware that has infected the devices of users can redirect the authentication messages and prompts to the hacker. If the user using the MFA authentication requires SMS service to receive the OTP, and the user doesn't have his phone, then authentication becomes impossible. The user must have his phone charged, in range of cellular network for authenticating. MFA in some cases might be very costly to establish like biometric authentication or security keys require specialized hardware. In such a case resource allocation and hardware maintenance becomes costly. The possibility of smartphones and tokens being stolen pose an immense threat allowing the thief to gain access to the user's accounts.

IV. SSO AND MFA IMPLEMENTED BY MAJOR CLOUD BASED SERVICE PROVIDERS

TABLE 1.SSO and MFA used by major cloud based service providers

Company Name	**Authentication type**	
	SSO	*MFA*
Google	Google apps offer a SAML (Security Assertion Markup Language) - based Single sign-on authentication service. It enables user full control of authorization and authentication of hosted user accounts that are accessible through Gmail or Google Calender.	Google Authenticator is software token that implements multifactor authentication (in this case 2FA) using Time-based One-time Password Algorithm and HMAC-based One-time password algorithm for authenticating users of mobile applications by Google. To use Authenticator, app must be installed on a smartphone and must be set up for each site to be used. The site provides a

		shared secret key to the user that is stored in the authenticator app. This secret key is then used for all future logins to the sites.
Facebook	OAuth is a service that can be utilized for Single Sign-on to websites.. It is utilized by Facebook and goes by the name "Facebook Connect". This enables the user's Facebook account to authenticate access into other websites.	Facebook provides two-factor authentication. It provides modes to do the same, through SMS or through a third-party authentication app like Google Authenticator or LastPass.
Amazon	Amazon web services provides the service of Single Sign-on that allows user to use Microsoft Active Directory credentials to access applications that are cloud-based, like AWS accounts and business applications like Salesforce, office 365.	An additional factor for signing in to AWS SSO can be provided using Remote Authentication Dial-In User Service (RADIUS) server. This RADIUS server is then configured to work with AD connector.

V. CONCLUSION

In this paper we discuss the need and ways to avoid security threats by comprehensively studying two modes of authentication, MFA (multi-factor authentication) and SSO (single sign-on). We discuss how they are implemented as well as the features that they encompass. Moving on we study the reasons why we should opt for MFA and SSO authentication techniques and what they might lack in terms of security. We also discuss the characteristics of the authentication technique based on the volume of data that a company deals with. Finally we state how majorcloudbased service providers handle big data security using these techniques.

ACKNOWLEDGEMENT

We thank our friends for so much of motivation and support for completing it successfully and presenting this paper.

REFERENCES

[1] Brown, B., Chui, M., and Manyika, J.: Are you ready for the era of 'big data'. McKinsey Quarterly 4.1, 24-35 (2011).

[2] Ward, J. S., Barker, A.: Undefined by data: a survey of big data definitions. arXiv preprint arXiv:1309.5821(2013).

[3] OpengroupHomepage,http://www.opengroup.org/security/sso/sso_intro.htm,last accessed 2019/02/15.

[4] One Login Homepage, https://www.onelogin.com/learn/how-single-sign-on-works, last accessed 2019/02/15.

[5] QuickLaunch SSO Homepage, http://www.quicklaunchsso.com/single-sign-on-software-essential-features.html, last accessed 2019/02/15.

[6] Jscape Homepage, https://www.jscape.com/blog/bid/104558/SSO-Single-Sign-On-Simplified, last accessed 2019/02/15.

[7] Jscape Homepage, https://www.jscape.com/blog/bid/104856/5-Big-Business-Benefits-of-Using-SSO-Single-Sign-On. last accessed 2019/02/15.

[8] Business news daily Homepage, https://www.businessnewsdaily.com/9766-single-sign-on-solutions-best-identity-access-management.html, last accessed 2019/02/15.

[9] futurehosting Homepage, https://www.futurehosting.com/blog/the-pros-and-cons-of-single-sign-on-for-web-services/, last accessed 2019/02/15.

[10] vassit Homepage, http://blog.vassit.co.uk/8-advantages-of-single-sign-on-sso-technology-and-3-flaws, last accessed 2019/02/15.

[11] Nist Homepage, https://www.nist.gov/itl/tig/back-basics-multi-factor-authentication, last accessed 2019/02/15.

[12] Panse, D. and Haritha, P.: Multi-factor authentication in cloud computing for data storage security. International Journal of Advanced Research in Computer Science and Software Engineering, 4(8), pp.629-634 (2014).

[13] globalsign Homepage, https://www.globalsign.com/en-in/blog/benefits-of-multi-factor-authentication/, last accessed 2019/02/15.

[14] idtheftcenter Homepage, http://www.idtheftcenter.org/ITRC-Surveys-Studies/2014databreaches.html., last accessed 2019/02/15.

[15] https://support.televic-education.com/hc/en-us/articles/360000621093-Single-Sign-on-SSO-?mobile_site=true

Raheem Qaiser is a student of B.Tech. (Final Year) in Computer Science & Engineering at Jamia Hamdard, New Delhi. Raheem has done many C programming based projects and pursued the fields of cloud computing and information security. Raheem has academic interests in the field of Cloud Computing, Network Security, Artificial Intelligence etc.

Aqeel Khalique is Assistant Professor in Jamia Hamdard, New Delhi. Aqeel has done several researches in the area of Information Security, Pervasive Computing, Cloud Computing & Cryptography. Aqeel has completed his M.Tech. from IIT Roorkee and worked in IT and Software Development Companies.

Dr. Imran is working as an Assistant Professor in the Jamia Hamdard. Dr. Imran major research is in the field of e-learning technologies which includes designing, development, implementation and administration of e-learning courses and its integration with open source e-learning tools.