EMBRY-RIDDLE
Aeronautical University
SCHOLARLY COMMONS

PhD Dissertations and Master's Theses

Spring 2021

Towards a Federated Identity and Access Management Across Universities

Jameel Alsulami
alsulamj@my.erau.edu

Scholarly Commons Citation

Alsulami, Jameel, "Towards a Federated Identity and Access Management Across Universities" (2021). *PhD Dissertations and Master's Theses*. 580.
https://commons.erau.edu/edt/580

Towards a Federated Identity and Access Management Across Universities.

by

Jameel Alsulami

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Cybersecurity Engineering
at Embry-Riddle Aeronautical University

Department of Electrical Engineering and Computer Science
Embry-Riddle Aeronautical University
Daytona Beach, Florida

May 2021

Towards a Federated Identity and Access Management Across Universities.

by Jameel Alsulami

This thesis was prepared under the direction of the candidate’s Thesis Committee Chair, Dr. Laxima Niure Kandel, and has been approved by the members of the thesis committee. It was submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the Degree of Master of Science in Cybersecurity Engineering.

____
Laxima Niure Kandel, Ph.D.
Committee Chair

____
Houbing Song, Ph.D.
Committee Member

____
Shafagh Jafer, Ph.D.
Committee Member

__________________________ ________________
Timothy A. Wilson, Sc.D. Date
Chair, Electrical Engineering and Computer Science

__________________________ ________________
Maj Mirmirani, Ph.D. Date
Dean, College of Engineering

__________________________ ________________
Christopher Grant, Ph.D. Date
Associate Provost of Academic Support

Acknowledgments

I am so grateful and thankful to ALLAH. I would like to take this opportunity to thank my family, my wife, and my children for their support and patience. I thank my advisor, Professor. Laxima Niure Kandel, for her encouragement and advice during my writing of my thesis, which has helped me reach this stage. I really appreciate who has supported me in various ways during my studies and made me feel comfortable.

Table of Contents

List of Figures

1. Abstract

Many research projects are too complex to yield the efforts of a single investigator and require a coordinated effort from interdisciplinary research teams across universities and industries. The research data, documents, experimental testbeds, high-end computing equipment, etc. is a critical component of any large-scale project and hence the cooperation and resource sharing across universities become very important for timely and budget-friendly execution of these projects. However, it is extremely challenging to frequently and effectively access data and other resources across universities without creating new identities for the users. In this thesis, we propose Federated Identity Management (FIM) approach for facilitating secure resource sharing among collaborating associates without creating new identities. We provide a comprehensive literature survey of identity and access management and discuss the privacy issues associated with identity management that can be addressed using FIM. We also provide a comprehensive overview and security features of the OAuth 2.0 framework which is an industry-standard protocol for authorization and user management used by FIM. The proposed scheme can be generalized and used by the student users to access academic libraries and recreate research results easily and securely.

Keyword: federated identity management, OAuth 2.0, cloud computing, identity management, cloud identity, federated cloud identity broker, privacy, protocol.

2. Chapter 1

Introduction

Collaboration among universities has the potential to yield multiple benefits. For example, collaboration could improve research results, lead to more resource gains, enhance problem-solving competence, etc. Due to a multitude of benefits, academic research is becoming increasingly collaborative. Many academics use research as their primary source of information. For academic papers, researchers must cite academic research to make arguments. Most researchers are working hard to discover new information and document it reliably. When researchers cooperate, their study could solve complex problems and ignite innovation and economic growth. Also, the collaboration between universities, researchers, and students has the potential to improve the dissemination of knowledge and teaching methods. In addition, collaboration with world-class education institutions could provide students the opportunity to learn state-of-the-art research methods directly from experts in the field. Unlike a few decades ago, where data and technology sharing were not possible, due to the recent digital revolution, information and data sharing is easily accessible today. A great deal of information is nowadays available online and is accessible to everyone, at any time, and at any place--thanks to the internet's accessibility and technological advancement.

The technological advancement has also led to the migration of several traditional educational services to online modalities, recently. In online modalities, students are able to take tests and assignments and access libraries and lectures via the internet. During the recent Covid-19 pandemic, most institutes and universities moved the face-to-face classes to online, and lectures were delivered using the internet and technologies such as zoom, Microsoft teams, etc. The success and effectiveness of these new modalities relies on robust IT infrastructures of the institution. Thus, a strong IT infrastructure is the foundation of any university’s science, research, and educational activities (Pardeshi, V. H. 2014). However, due to the high cost of modern technologies, universities and colleges take a long time to deploy and use these methods. Since, moving to online modalities entails significant funding and expenditure and given the governments and private institutions' depleted budget reserves (Pardeshi, V. H. 2014), not all organizations can rip the benefits of new technological advancements. The shrinking IT budgets and escalating IT needs can be solved using the Cloud Computing (CC) platform as it provides on-demand, ubiquitous, practical access to computer systems. Thus, cloud computing environments have the potential to save higher education institutions from financial challenges by lowering the cost by means of hardware virtualization. It allows organizations to balance the capacity they need and pay as they go monthly for the services they use. Higher education institutions seeking to reduce IT costs can find solutions in cloud computing (Dillon, T., Wu, C., & Chang, E. 2010). In the last few years, higher education is attempting to capitalize on the opportunities presented by CC in order to provide access to advanced IT infrastructure, data centers, and applications while mitigating associated security and privacy issues (Vaquero, L. M., Rodero-Merino, L., Caceres, J., & Lindner, M. 2008). The cloud-based IT infrastructure could contribute to the standardization and updating of educational content as well as the enhancement of cooperation and collaboration between higher education institutions (Pardeshi, V. H. 2014). Although the CC-based services have advanced massively over the past years, the technology still has not fully matured and has many limitations. Particularly, when accessing the data on cloud-based platforms, private credentials of the user's identity could be exposed putting the user's privacy and security at risk. Thus, authentication and authorization are needed for data access.

As stated above, despite the multitude of benefits of sharing knowledge and resources, collaboration among multiple universities is not easy due to the complex intervention of multiple components. One major component is privacy and security issues because of which universities may not want to share their data even when the use and availability of data resources can benefit every research university. Universities may hesitate since dangers loom that the other stakeholders will exploit the shared data for personal and private needs. In extreme cases, the stakeholder may rob the technology, leave the collaboration, and threaten the survival of the other stakeholder that shared it. Despite these risks, collaboration is needed and sometimes required as the shared data could provide the needed background information about the subject and could aid in formulating the right research questions for collective good. Furthermore, sharing private research data and resources with other stakeholders can lead to the timely and budget-friendly execution of large projects that could impact science, innovation, economic and workforce growth.

To rip the vast benefits of academic collaboration and mitigate the limitations of privacy and security issues in the cloud (Indu, I., Anand, P. R., & Bhaskar, V. 2018), many organizations have adopted Identity and Access Management (IAM) systems. According to Gartner, "identity and access management is the discipline that enables the right individuals to access the right resources at the right times for the right reasons." Furthermore, the IAM systems provide protection to confidential information stored in the cloud enabling reliability and usability of customer access control imperative to any organization's sites. However, IAM is not a panacea, nor it mitigates all the privacy and security issues of the cloud. Also, universities may require different identities (or redundant identities) under the IAM scope. To address the problems of IAM, in this thesis, we propose federated identity management (FIM) as a means of identity access and management to enable the vision of academic collaboration. In this thesis, we particularly focus on academic collaboration; however, the technique can be generalized to other organizations or industries. The goal is to have many access credentials linked i.e., only one identity creation is under FIM. Single identity would allow users from one realm to securely access resources in another realm without the need for redundant logins. The user does not need to remember a long list of complicated passwords enhancing the user experience as well as security and privacy of the user identity. The proposed method has a single interface that allows the user to access data and applications on different systems without different login credentials for each one.

1.1.Motivation

Academic collaboration is extolled as an important feature in the successful execution of large-scale research projects and for the understanding of complex systems (Hörbe, R., & Hötzendorfer, W. 2015). Despite the enthusiasm at the theoretical level, the implementation of collaboration among academic communities is not yet practical. One reason for this is the issue of trust and competitive nature among the stakeholders. Also, nowadays the universities host academic databases on commercial and institutional cloud computing platforms which magnifies the trust and security issues. To mitigate the security and privacy issues and authenticate the access of resources to legitimate stakeholders, we propose FIM-- a framework that aims to hide the identities and confidential information about the user and allows for collaboration within and across organizations using the single sign-in credential. We believe this method would allow easy collaboration and benefit many researchers and students. Students, under this framework, can securely access academic libraries and/or research data which would be used for recreating the results that could in turn help in better knowledge acquisition and learning.

1.2.Methodology

In the first part of the research, we do an extensive literature review and provide background information about the federated identity and how it can benefit the academic community collaboration. Through extensive literature review, we examine past approaches and studies the theoretical background underlying IAM models, which in turn facilitated the identification of essential components that fit into our FIM architecture. Therefore, several methodologies used in our design are drawn on the contribution of researchers who have studied IAB. The second part of the research deals with the adoption of some recent and powerful technology such as OAuth 2.0 and OpenID connect to design the proposed FIM framework.

Chapter2. Background

2.1. Cloud Computing

Cloud computing allows organizations to balance the capacity they need and pay as they go monthly for the services they use. Cloud computing is a model that enables ubiquitous, practical, on-demand network connectivity to a shared pool of configurable computing services that can be easily provided and even released with minimal management or service provider involvement. The Cloud computing model consists of five main characteristics, four deployment models, and three service models (Mell, P., & Grance, T. 2011). Cloud computing seeks to achieve the virtualization of resources and improve the overall computing capacity (Dillon, T., Wu, C., & Chang, E. 2010). Cloud computing has presented a brand-new standard that assists users in dynamically storing or developing applications and gaining access to them anywhere and even at any time by connecting to an application using the network (Vaquero, L. M., Rodero-Merino, L., Caceres, J., & Lindner, M. 2008). By offering computation, storage, and software-based services, cloud computing has obtained broad acceptance for organizations and individuals. Cloud Computing serves to solve the infrastructure shortage of consumers by offering pay-per-use applications on-demand. In cloud computing, cloud providers take responsibility and function to run software and hardware to promote performance. Cloud computing commoditization has resulted in a radical form of vertical disintegration, where physical infrastructure is unbundled from the platform layer and delivered as a service (Kushida, K. E., Murray, J., & Zysman, J. 2015).

2.1.1. Cloud Computing Service Models

A variety of service models are available to access cloud computing. These services are specifically intended to exhibit certain characteristics as well as satisfy organizational needs. An organization can select and customize one of the best-suited possibilities from the list below (Bulusu, S., & Sudia, K. 2013).

- Software as a Service (SaaS).

This functionality is given to the customer to use the applications running on a cloud platform on demand. The consumer can use the software of the provider on a cloud infrastructure. Applications can be accessed from several client devices through a client interface such as a web browser. The consumer does not handle the underlying cloud infrastructure.

- Platform as a Service (PaaS).

The consumer has the capability to deploy onto the cloud infrastructure their developed or acquired applications created using the programming languages and tools provided by the provider. The consumer has control over which applications they install.

- Infrastructure as a Service (IaaS).

The customer has the capability to provide the processing, storage, networking, and software environments in which the consumer can run their applications. The consumer has some control over, can manage operating systems, storage deployed applications and possibly limited control of select networking components.

2.1.2. Cloud Computing Deployment Models

One or more deployment models can be used to deploy the service models explained above (Bulusu, S., & Sudia, K. 2013). The cloud deployment models are used to demonstrate how the cloud services are made available to users. Cloud computing is typically divided into four basic deployment models (Laszewski, T., & Nauduri, P. (2010). The following are explanations of each of these deployment models:

1. Public cloud: All users who want to use a computing resource will use this type of cloud deployment model. Application development and testing, non-mission-critical activities such as file sharing, and e-mail service are the most popular uses of public clouds. This model is the first choice by businesses with low privacy concerns.

2. Private cloud: This model is commonly used for a single organization's infrastructure. Acquisition and maintenance costs for private clouds are more costly than for public clouds. A private cloud is better to meet the security and privacy challenges of organizations. The private cloud model minimizes data security issues. The private cloud allows for more customization of the infrastructure to meet the needs of the company. The private cloud allows for more customization of the infrastructure to meet the needs of the company. A private model is perfect for companies who want to shield their mission-critical operations or businesses with rapidly evolving needs.

3. Community cloud: the only difference between a community deployment model and a private deployment model is the number of users. Access to a community cloud environment is usually limited to community members. Unlike a private cloud server, which a single organization owns, a community cloud is shared by many organizations of similar backgrounds.

4. Hybrid cloud: A hybrid cloud is one in which an organization uses interconnected private and public cloud infrastructure. Hybrid cloud enables organizations to incorporate the aspects of the three models that better serve their needs. This model is commonly used by many organizations when they need to scale up their IT infrastructure quickly.

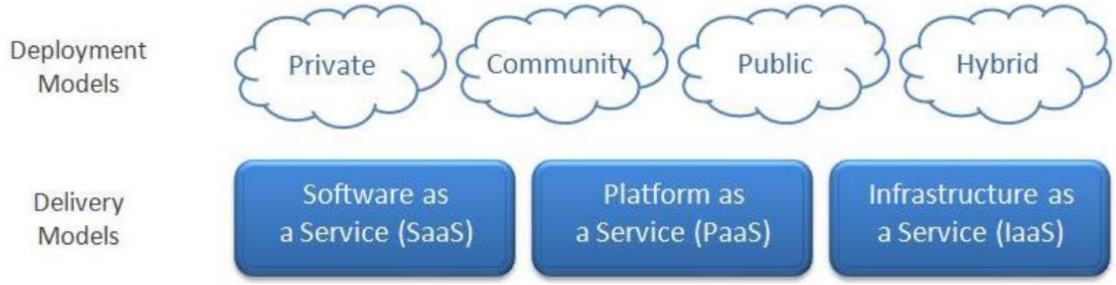
Deployment Models
Private
Community
Public
Hybrid
Delivery Models
Software as a Service (SaaS)
Platform as a Service (PaaS)
Infrastructure as a Service (IaaS)

Figure 1 Cloud Computing Deployment and Delivery Models (Mell & Grance, 2011)

2.1.3. Security Issues in the Cloud Computing

The security field is a significant concern in cloud computing. Many organizations are hesitant to move their data to the cloud environment because there are many threats involved in handling sensitive data (Singh, A., & Chatterjee, K. 2017). The cloud environment is affected by threats and attacks. If Integrity, availability, and confidentiality of the cloud resources and service of different layers are breached, it could raise a new security concern (Singh, A., & Chatterjee, K. 2017). The key to effective security implementation in cloud computing is understanding where the service provider's responsibility ends and where the customer's responsibility begins (Dotson, C. 2019). Cloud security is considered as a part of computer security. It defines a set of policies, technology, and control that helps secure the data and services.

2.1.4. The Benefits of Cloud Computing

The benefits of cloud computing have excited the attention of the information and technology community. Cloud computing technology is proving to be beneficial to many organizations and individuals as opposed to traditional computing methods. Organizations will focus on doing their research and development instead of thinking about patching security holes and coping with other computing problems. Organizations will have the ability to choose from several vendors that offer reliable, flexible services, development environments, and an infrastructure that can be leveraged with no long-term contracts. The most important advantage of cloud computing is cost saving. As it does not require any physical hardware investment, it allows organizations to save significant capital costs. Reduced cost due to operating efficiencies and quicker business services delivery. The organizations do not need substantial training to operate the hardware. The cloud service provider handles the procurement and maintenance of equipment. Cloud computing provides access to hundreds of applications at any time without the need for installation and configuration. Cloud Computing helps organizations to deploy their business quickly. Therefore, deploying rapidly helps in getting the resources needed for a framework.

Data can be available on the cloud and be backed up and retrieved more comfortably from on-premises. Reliability is the most significant advantage of cloud computing.

Organizations always get updated about the changes instantly. Employees at workplaces or operating remotely will use all possible facilities accessible to them. Internet access is what they need. The cloud has nearly unlimited storage space. Organizations can easily extend your storage space with minimal month-to-month fees. The cloud computing platform allows employees to communicate instantly and conveniently regardless of where they are located. Cloud computing allows fast implementation by remote access. Organization whole systems will become operational in a short period.

The overall cost benefits of the cloud are driven primarily by hardware obligations. If a virtual machine, a server, or an entire data center goes down in the cloud, it will be managed by the cloud provider, and the organization can continue doing its business as usual. In contrast, when an organizations' on-premises hardware fails, it will cost a lot of money. In the cloud, running costs will be much smaller than on-premises. Dedicated on-premises servers are still expensive, and they often need more than a one-time expense such as maintenance and updates cost. An incredibly significant benefit of cloud-cost savings can be seen in data center construction and infrastructure. Additionally, there is a need for ancillary services equipment to support on-premises servers like a switch, rack, cooling fans, which will increase the cost. This means that operating costs will reduce significantly in the cloud rather than run on-premises.

2.2 Identity and Access Management (IAM)

Identity and Access Management systems are used to enhance their ability to protect confidential data stored in the cloud and to provide further protection to confidential information stored in the cloud. IAM provides the sorts of reliability and usability of customer access control imperative to any organization's sites these days. How an organization decides to implement IAM in its cloud environment will depend on its requirements. The IAM framework's primary functions are to ensure that users who access enterprise resources are who they say they are and identify and monitor user, device, or service, access rights (Indu, I., Anand, P. R., & Bhaskar, V. 2018) . The vital IAM concepts are access and user.

2.2.1 IAM Processes in a Cloud Computing Environment

In a cloud computing environment, the general processes of adding, modifying, or removing a user remain unchanged compared to the traditional IT environment. Users must be added, modified, or removed from a system before accessing its authorized resources. IAM in the traditional model is handled, managed, and regulated within the on-premises by the organization. Local authentication allows users to access local services such as data and applications (Jansen, W., & Grance, T. 2011).

The organization that uses cloud services is often not responsible for authentication management. The majority of authentication occurs in the cloud. For users to access cloud services, most cloud service providers use their authentication mechanism. In a cloud computing environment, the organization that uses the services determines which resources the users can access. When the organization uses cloud services, both cloud service providers and the organizations that use cloud services have different authorization models. Furthermore, the organization that uses the cloud services does not have the control to enforce its security policies for cloud service providers' services because the cloud service providers oversee access to their services.

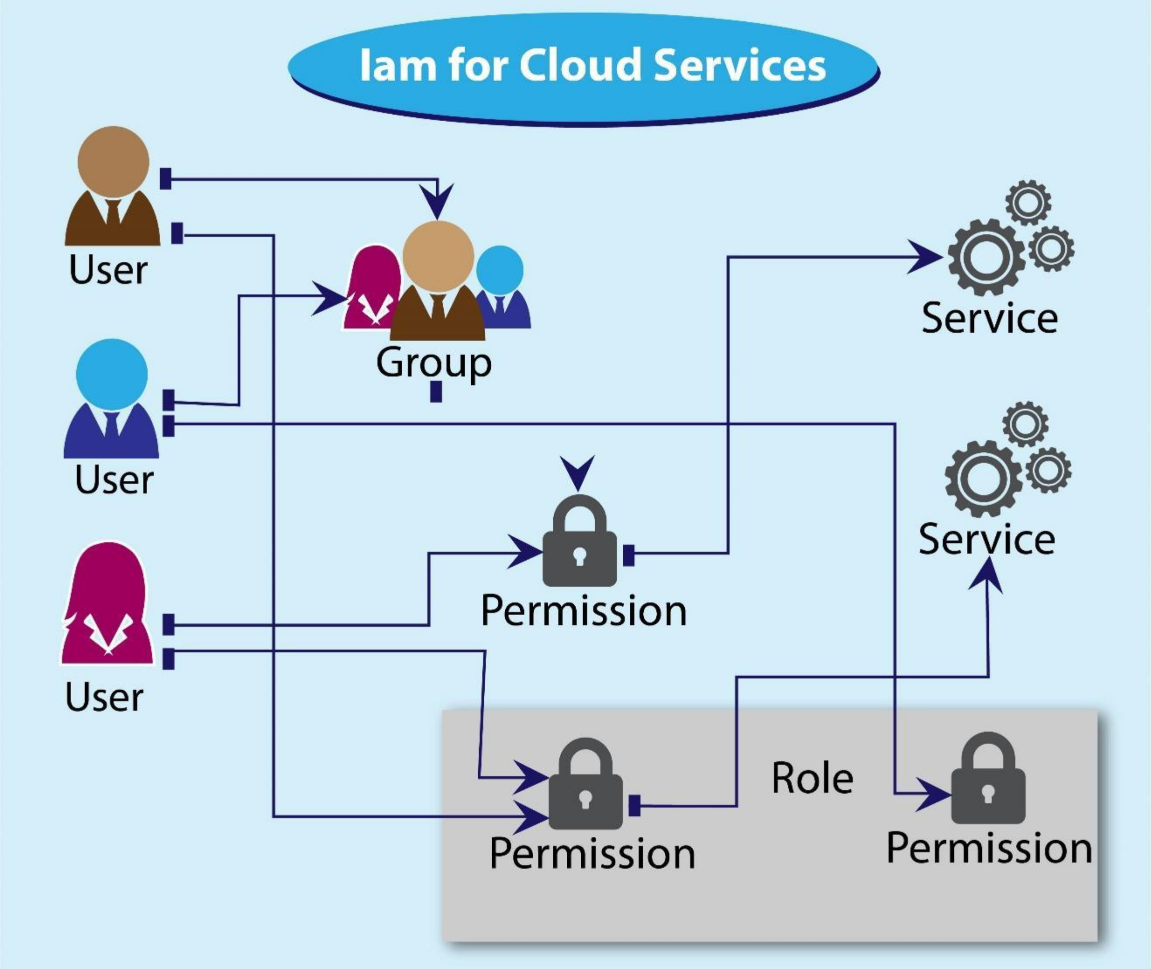

Figure 2 IAM for Cloud Services

2.3 Cloud Identity Management Models

Because cloud computing is becoming more prevalent in the IT industry, secure identity management is becoming increasingly important (Zwattendorfer, B., Zefferer, T., & Stranacher, K. 2014). In a cloud computing environment, there are many several for managing identities and access. The identity can be delivered from the cloud, it can be taken to the cloud, or it can be stored in the cloud (Zwattendorfer, B., Zefferer, T., & Stranacher, K. 2014).

1. Identity in the Cloud-Model: Identity management is the cloud service provider's responsibility that hosts the application. In this model, organizations can rely on an existing identity management system, managed by the cloud service provider, rather than hosting and managing their own. When using this model, an organization's expenses can

be minimized. The organization loses control over the identity data stored and handled in the cloud when this model is used, and the cloud service provider assumes responsibility for security and privacy.

2. Identity to the Cloud-Model: The key distinction of this model is that the service provider and its applications are hosted in the cloud. This model prevents unnecessary identity data disclosure to a cloud service provider because the identity provider is not deployed in the cloud. The organization continues to host the entire user and identity management.

3. Identity from the Cloud-Model: Both the cloud application and the identity provider are hosted in this model's cloud. The separation of cloud service providers is an advantage of this model. Organizations in this model can choose their preferred cloud identity provider. This is especially important because the identity provider responsible for the organization's identity and user management must be trusted.

4. The Cloud Identity Broker-Model: The cloud identity broker links one or more services with one or more identity providers. Using the broker concept, the identity broker conceals the ambiguity of the service provider's individual identity providers. This means that only one interface is required, which is required for the identity broker. This model is even more effective when the broker is deployed in the cloud. One downside is that the cloud identity broker's functionality is based on both the user and the service provider. The service provider would be unable to offer its services to the user if the identity broker does not accept the chosen identity provider for authentication.

5. The Federated Cloud Identity Broker-Model: Users and service providers do not have to rely on the exact identity broker in this federated model. Both the user and the service provider can depend on their preferred individual broker. This removes the downside of being reliant on the exact identity broker for both the user and the service provider.

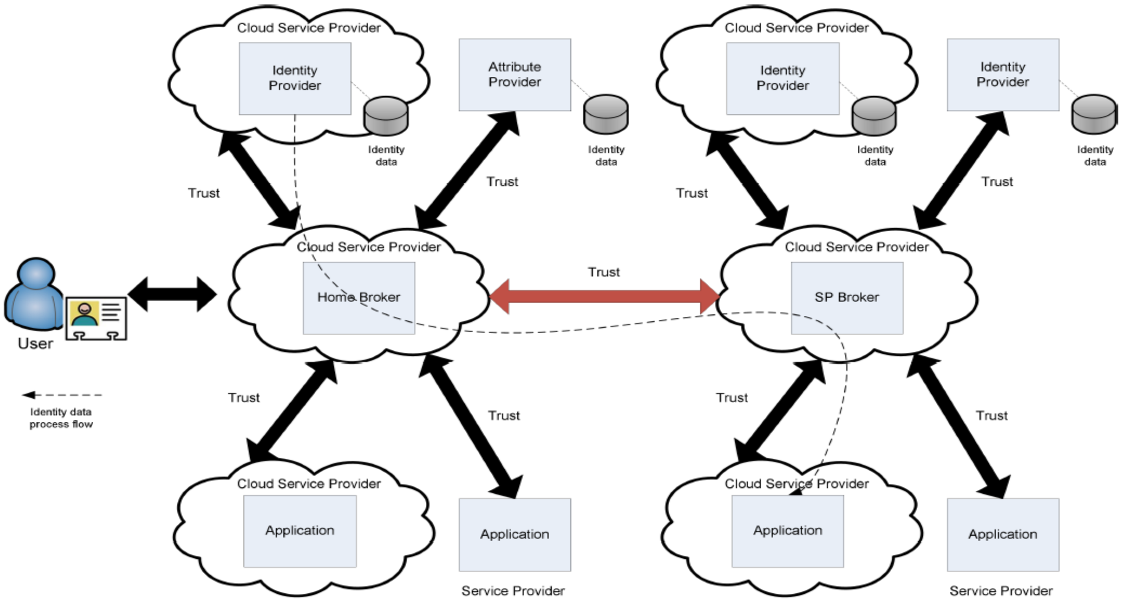

Figure 3 A Federated Cloud Identity Broker-Model (Bernd Zwattendorfer , 2016)

2.4. Access Risks in Cloud Security

Regardless of whether it is a service based on a private cloud, public cloud, or a hybrid cloud, the importance of user access control is a nearly constant challenge. When using cloud computing, organizations should recognize that their sensitive information will be shared with a third-party cloud computing provider. Various security mechanisms are currently used to alleviate the problems in the cloud. Access management, in general, includes three capabilities: the ability to define and authenticate users, grant users access rights, and develop and implement access control policies. Multi Factor authentication can help mitigate the risk of credential compromise, as compromised privileged user credentials give an attacker the ability to access and customize cloud customer services. By requiring an attacker to obtain several, independent authentication components, multiple factors reduce the probability of a compromise.

2.5. Literature Review

Research performed by previous scholars helps current researchers a lot because it allows them to avoid making the same mistakes. A literature review examines published research in a specific subject area to provide an overview of current knowledge. Additionally, literature reviews help to advance the research's cause by examining its significance from the perspective of well-known scholars. Therefore, tens of thousands of books, articles, and studies have been published on the topic. In this chapter, the thesis will concentrate on the research work done by other well-known experts and writers and attempt to provide some robust solutions to this issue.

In the NIST definition of cloud computing paper, Peter Mell and Timothy Grance (2011) defined the service and deployment models and discussed what cloud computing is to best use cloud computing. It is an evident fact that the definition has provided a brief introduction of the service and deployment models in cloud computing that can be useful for a person who does not know much about cloud computing. The authors made a simple taxonomy that is not intended to prescribe or constrain any specific deployment method, service delivery, or business activity.

Vaishali H Pardeshia (2014) proposed a cloud computing architecture for a higher education institution that includes multiple deployment models, service models, and user domains. He presented a five-phase strategy to facilitate the migration from traditional IT to a cloud computing environment and provided suggestions for a smooth and effective transition from a traditional system to a cloud-based system.

A University's management needs to be assured of the benefits, challenges, and cost of adopting a federated identity technology before they invest. As a starting point, we wanted to identify the benefits of deploying federated identity management in scientific literature. the benefits of adopting federated identity management systems from a user and business perspective and a high-level view on adopting federated identity management in an integrated operations environment. The use of Federated Identity Management, according to several studies, will improve the ability to protect personal privacy. Moreover, because of the reduced number of authentication operations, users would choose different and better passwords from their Identity Providers. Federated Identity Management has the potential to enhance user protection. In federated identity management systems, users are relieved from memorizing multiple passwords. With federated identity management systems, users will benefit from increased simplicity because federated identity solutions can reduce complexity. Additionally, improved user experience is accomplished by seamless access to services and the removal of redundant user login processes. According to Jostein Jensen the user and business benefits of adopting Federated Identity Management systems and a high-level view of implementing Federated Identity Management in an Integrated Operations environment.

In potential market situations, Gail-Joon Ahn and John Lam discussed FIM privacy concerns. they suggested systematic mechanisms to specify privacy preferences in FIM and introduced a user preference expression language that is essential to managing users in FIM (Ahn, G. J., & Lam, J. 2005). Without any doubt, federated identity management's critical concerns are information security and privacy, as identity federation requires sharing confidential user information over an insecure and open network.

From A Survey on Security Issues of Federated Identity in the Cloud Computing by Eghbal Ghazizadeh et al (2012). They believe that by establishing the proper trust relationship among participating federated entities, infrastructure components, and cloud platforms, Trusted Computing can reinforce existing security solutions. They addressed federated identity management systems, cloud computing, single sign-on, and SSO protocols like OpenID and OAuth and highlighted the introduced models for addressing identity theft in a federated environment. The aim of this paper was to demonstrate how trusted computing technologies can dramatically reduce phishing attacks on cloud-based user assets (Ghazizadeh, E., Zamani, M., & Pashang, A. 2012).

Vahid Jalili and others (2019) have established a generic and extensible approach for securely accessing biomedical datasets spread through cloud computing platforms. This approach combines OpenID Connect and OAuth2, which are best-practice web protocols for authentication and authorization, with a web-based computational workbench that is used by thousands of scientists worldwide called Galaxy. This approach allows users to access and analyze data spread through various cloud computing providers without special knowledge of these protocols.

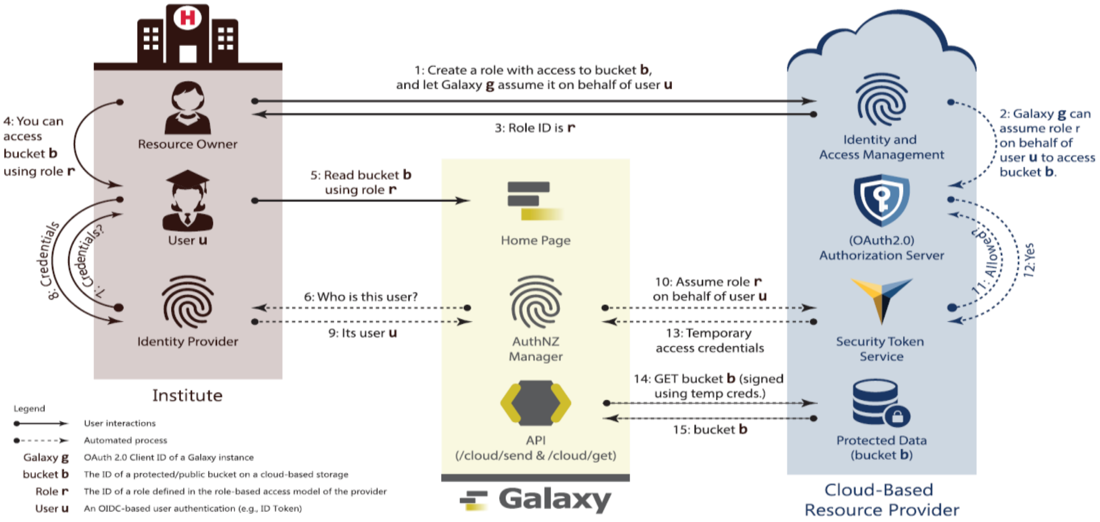
1: Create a role with access to bucket b,
and let Galaxy g assume it on behalf of user u
3: Role ID is r
2: Galaxy g can assume role r on behalf of user u to access bucket b.
4: You can access bucket b using role r
Resource Owner
User u
Identity Provider
Institute
7: Credentials?
8: Credentials
5: Read bucket b using role r
Home Page
6: Who is this user?
9: Its user u
AuthNZ Manager
API
(/cloud/send & /cloud/get)
Galaxy
Identity and Access Management
(OAuth2.0) Authorization Server
11: Allowed?
12:Yes
10: Assume role r on behalf of user u
13: Temporary access credentials
Security Token Service
14: GET bucket b (signed using temp creds.)
15: bucket b
Protected Data (bucket b)
Cloud-Based Resource Provider
Legend
User interactions
Automated process
Galaxy g OAuth 2.0 Client ID of a Galaxy instance
bucket b The ID of a protected/public bucket on a cloud-based storage
Role r The ID of a role defined in the role-based access model of the provider
User u An OIDC-based user authentication (e.g., ID Token)

Figure 4 Cloud bursting galaxy federated identity. Vahid Jalili and others (2019)

Federated Identity Management in the Norwegian Oil and Gas Industry by Jostein Jensen (2014), has offered a brief introduction to all mandatory terminology that can help anyone unfamiliar with federated identity management. He performed a study of companies engaged in developing oil and gas in Norway to ascertain their perceptions of the potential advantages, threats, and additional security risks associated with the adoption of federated identity management. He indicated that a strong emphasis on security is needed across the software development lifecycle when designing identity management solutions. He has documented various access management issues facing the industry and has outlined the benefits and problems of federated identity management from both an academic and a business perspective.

Some authors (2012), who are interested in federated identity management for their research cyberinfrastructure, demand that user bases be expanded through diverse alliances to overcome the challenge of getting access to scientific data across organizational and national boundaries. Motivated by these needs, they established an active community called FIM4R. FIM4R consists of communities and infrastructures involved in enabling federated identity (Broeder, D et al., 2012).

Chapter 3 Federated identity

3.1 Introduction to Federated Identity

Federated identity management is a critical component of digital identity management (Madsen, P., Koga, Y., & Takahashi, K. 2005). Federated Identity management is primarily motivated by the need to improve user experience and privacy (Ahn, G. J., & Lam, J. 2005). Federated Identity management will simplify the user management process (Ahn, G. J., & Lam, J. 2005). The main goal of federated identity management is to address how to leverage organizations' identity management operations to allow partners and customers direct access to their applications. Federated identity refers to the collaborative and interdependent management of identity information across organizations. The federation model allows users from one domain to securely access resources from another domain without having to go through several login processes. Federated identity management removes the need for users to have an account in the organization directory; instead, they can access services by logging in once to their identity provider. Eve Maler and Drummond Reed (2016) presented four logical components to the federated model: the user, the user agent, the service provider, the identity provider (Maler, E., & Reed, D. 2008).

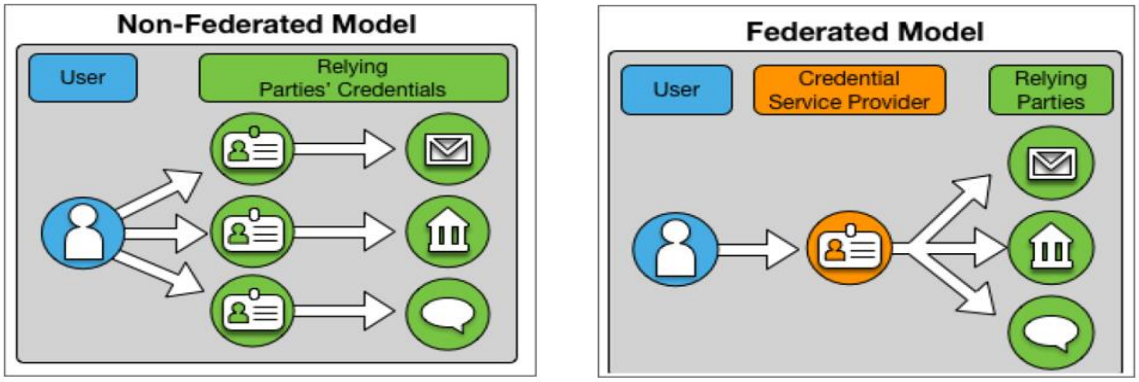

Figure 5 Federated and Non-Federated Identity (Temoshok and Abruzzi, 2018)

3.1.2.The CIA Triad in Federated Identity

The CIA Triad(Confidentiality, Integrity, Availability) is considered the foundation of information security. Confidentiality is based on the ability to identify and implement clear access thresholds for information. People must protect their confidential, private information from unauthorized access in today's world. Access control lists, volume and file encryption, and Unix file permissions are among the most popular methods for maintaining confidentiality. In contrast, integrity is intended to protect data from unwanted deletion or alteration. When an approved individual makes a change that should not have been made, integrity means that the damage can be undone. While availability is to protect information and make it accessible when needed, authentication processes, access networks, and systems must function correctly.

In the federated identity model, Confidentiality is reinforced as follows: third parties do not have plaintext access to user credentials or attributes and will never be able to access decryption keys. A malicious man-in-the-middle attack would not violate the authenticated user's data, and it is impossible to gain unauthorized access to transactional data. In contrast, integrity is reinforced as follows: the relying party has the assurance that the data has not been altered by the hub or a malicious third party. The relying party is confident that the data is being provided by a valid credential service provider and guaranteed that a malicious third party would not impersonate a legitimate user and reuse previously valid assertions (Grassi, P., Lefkovitz, N., & Mangold, K. 2016).

3.2.Trust Relationships in Federated Identity Management

A foundation of trust is needed in a federated business model. A federated model is one in which an organization can grant access to an identity that its internal security mechanisms have not vetted. The organization trusts an identity claimed by a third party. If an organization does not have insight into its business partners' identity and access management systems and processes, it cannot participate in a federated business model.

Trust among members of an identity federation is essential to its operation and is defined through a collection of agreements and associated rules unique to that group. Identity federations are made up of credential service providers (CSPs) and relying parties (RP) that have agreed to engage in a particular form of federated identity management. Users are registered, passwords are established, users are authenticated, and federation RPs are informed of their authentication status by CSPs. RPs use the authentication status information to approve user access to online services and applications based on identity assertions supported by CSPs. According to David Temoshok and Christine Abruzzi, "a trust framework is the set of rules and policies that govern how the federation members will operate and interact (Temoshok, D., Temoshok, D., & Abruzzi, C. 2018)."

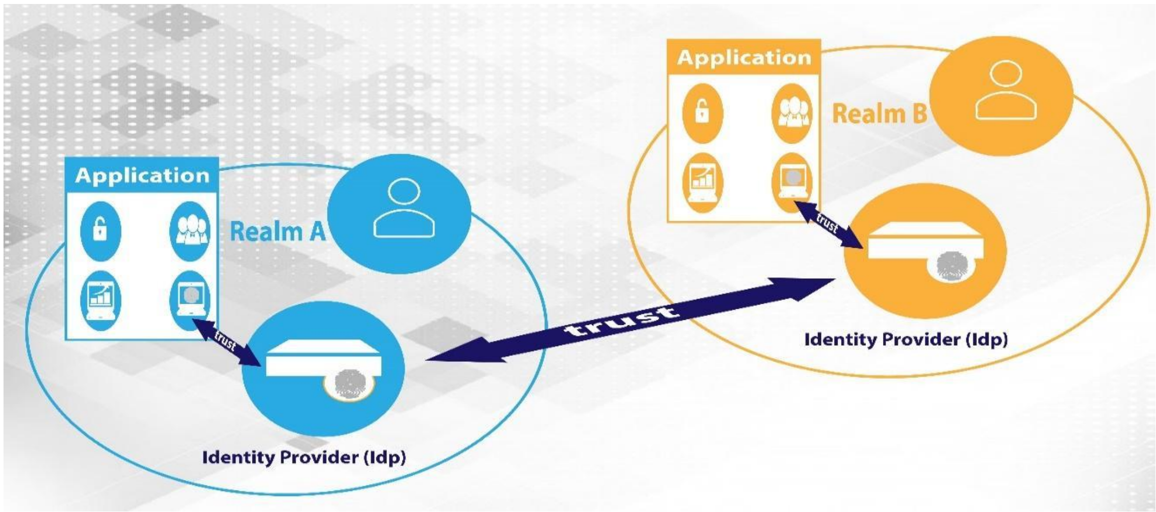

Figure 6 Trust relationships in federated identity management.

3.3.Privacy in Federated Identity Management

When it comes to cloud computing, one of the most important issues is privacy. In terms of managing privacy, protecting data, and adhering to legislation, sharing personally identifiable information is a significant concern. Federated identity management raises new privacy concerns. Although features such as pseudonymous authentication and limited attribute release can help increase privacy, federated identity management can also increase privacy risks for various reasons. NIST Cybersecurity Practice Guide describes the practical steps required to execute an example solution that addresses current challenges in the federated identity market. An identity ecosystem of federated identity solutions can play a pivotal role in achieving a more secure cyberspace. The federation prevents relying party service and identity providers from learning the identities of each other. Other than what is known from their direct relationship with the user, both entities cannot monitor and connect user behaviors (Laszewski, T., & Nauduri, P. 2010).

According to Paul Grassi and Naomi Lefkovitz (2016), although the identity broker may raise individual privacy risks, Privacy-enhancing technologies built into hardware or software eliminate adverse effects on individuals when their personal information is being collected or processed. They presented a set of goals that represent a comprehensive set of privacy goals that federated identity solutions may achieve based on requirements and demand. The following are explanations of each of these goals (Grassi, P., Lefkovitz, N., & Mangold, K. 2016).

Goal 1: The federation prohibits RPs and CSPs from discovering each other's identities. Neither entity may monitor or connect user activities beyond what is learned from their direct relationship with the user.

Goal 2: Participants in the federation, other than those the user approves, cannot access user attributes. Users must first agree to share the attribute from the CSP to the RP. Then, Validated attributes are obtained by RPs from authoritative CSPs. When the RP has the actual attribute value, they will use it to meet their service requirements.

Goal 3: A compromised, or malicious federation participant cannot impersonate a user. To reduce this threat, controls must be put in place.

Goal 4: Attributes are only provided when requested by an RP, not every time a user logs in to use RP services. RPs will only collect the attributes required by the services a user is requesting to satisfy the user's services.

Goal 5: Users must explicitly consent to the disclosure of their attributes to an RP.

Goal 6: User pseudonyms cannot be tracked or linked through transactions by entities that mediate identity transactions. For maintaining privacy, especially when multiple Web services collaborate to provide an aggregated product that necessitates the sharing of user attributes, pseudonyms are an essential technique.

To facilitate the development and operation of privacy-preserving information systems, NIST (2016) has developed three draft privacy engineering objectives. These objectives are intended to assist system designers and developers in developing information systems capable of achieving their practical objectives while also supporting an organization's privacy priorities and risk management (Grassi, P., Lefkovitz, N., & Mangold, K. 2016). These privacy objectives are Predictability, Manageability, and Disassociability.

Predictability is the capacity for individuals, owners, and operators to make accurate decisions about personal information and its processing by an information system.

Manageability is the capacity to administer personal information granularly, including its modification, deletion, and selective disclosure.

Disassociability refers to a system's ability to handle personal data or events without associating them with specific persons or devices outside the system's operational requirements.

3.4.Common Technology in Federated Identity

Several identity management technologies have already appeared over time. Four popular federated identity protocols. The following technologies are widely used in federated identity management:

1. Security Assertion Markup Language (SAML): SAML is a federated authentication protocol. It is most often used to allow users to sign into multiple applications using a single login. In other words, SAML is an XML-based standard for sharing authentication and authorization data between identity providers and service providers to verify the user's identity and permissions before granting or denying access to services. Moreover, SAML is a relatively heavyweight protocol due to the scale of the XML messages sent to and from the SP and IDP.
2. OAuth: OAuth is a protocol for authorization that decides what that user should be allowed to do. In other words, OAuth is a security standard that enables users to grant permission to one application to access their data stored in another application. The OAuth 2.0 allows clients to retrieve user profile information without information about the end user's authentication. Without providing their password, the users authorize one application to access their data. OAuth is not a single sign-on protocol; it lacks default digital signatures and encryptions, making it vulnerable to various security threats and data access breaches. Permission-granting procedures are sometimes referred to as delegated authorization. So, in OAuth, authorization is delegated. A variation of the OIDC and OAuth2 protocols can be used for various authentication purposes, including machine-to-machine and device-to-device authentication.
3. OpenID: OpenID is a protocol for authentication that ensures that the user you are talking to is indeed who he claims to be. Although OpenID 2.0 had some useful security features, it was restricted to web applications and had several other design limitations, such as dependency on XML and custom message signatures, which contributed to adoption difficulties.

4. The InfoCard protocol is only compatible with the WS-* Web service protocols, which means that WS-Trust serves as the foundation for InfoCard. The InfoCard protocol is a credential exchange protocol created to provide users with a consistent digital identity experience using a specialized user agent[19].

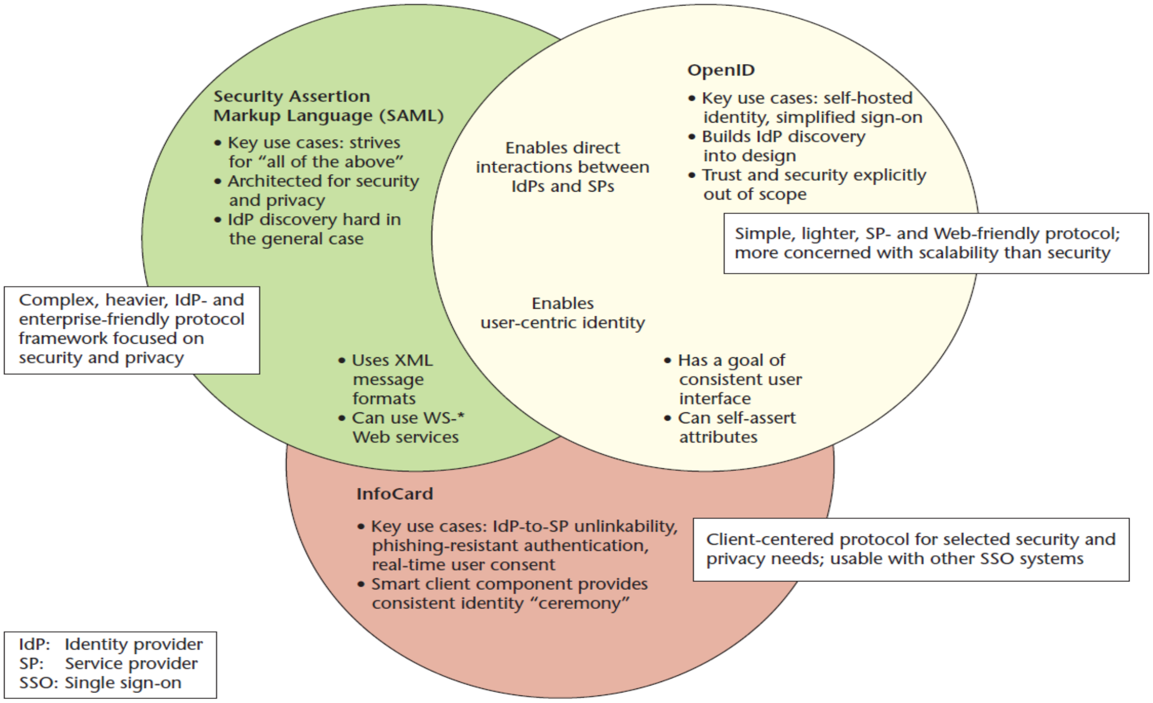
Security Assertion Markup Language (SAML)
• Key use cases: strives for "all of the above"
• Architected for security and privacy
• IdP discovery hard in the general case
OpenID
• Key use cases: self-hosted identity, simplified sign-on
• Builds IdP discovery into design
• Trust and security explicitly out of scope
Enables direct interactions between IdPs and SPs
Simple, lighter, SP- and Web-friendly protocol; more concerned with scalability than security
Complex, heavier, IdP- and enterprise-friendly protocol framework focused on security and privacy
Enables user-centric identity
• Uses XML message formats
• Can use WS-* Web services
• Has a goal of consistent user interface
• Can self-assert attributes
InfoCard
• Key use cases: IdP-to-SP unlinkability, phishing-resistant authentication, real-time user consent
• Smart client component provides consistent identity "ceremony"
Client-centered protocol for selected security and privacy needs; usable with other SSO systems
IdP: Identity provider
SP: Service provider
SSO: Single sign-on

Figure 7. The differences between SAML, OpenID, and The InfoCard. (Eve Maler and Drummond Reed, 2008)

3.6.The OAuth 2.0 Framework

3.6.1.OAuth 2.0 Roles

OAuth 2.0 Framework defines four roles as follows:

Resource Owner: The user owns his or her identity, records, and any acts that can be taken with his or her accounts.

Client: The application that wants to access the Resource Owner's data or perform actions on his or her behalf.

Authorization Server: The application that is used to apply access policies and allows access to the data on the Resource Server on behalf of the Resource Owner where the Resource Owner already owns an account.

Resource Server: The service requested by the Client on behalf of the Resource Owner.

The Resource Server and Authorization Server may be implemented in the same server entity or a separate one (Lodderstedt, T., McGloin, M., & Hunt, P. 2013).

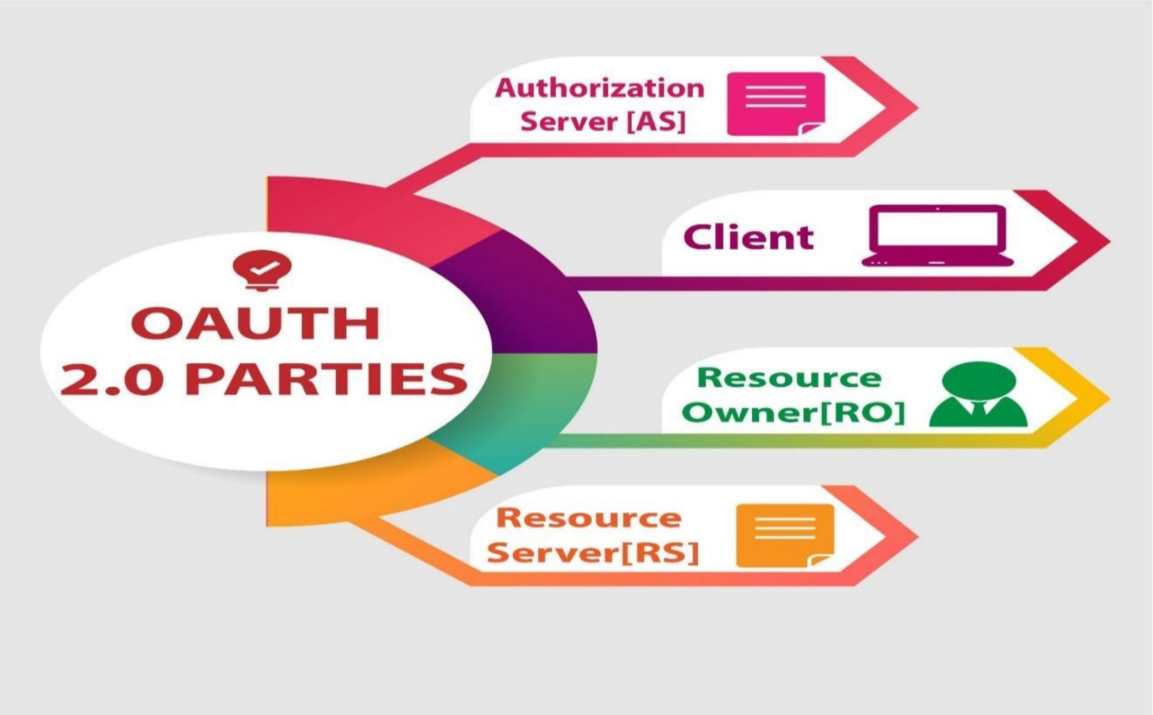

Figure 8 OAuth Roles.

3.6.2.Authorization Grant

An authorization grant is a credential that the client uses to obtain an access token. It represents the resource owner's authorization. According to Hardt, D. (2012), there are four grant types of an authorization grant as follows: (Hardt, D. 2012).

1. The authorization code is obtained by using an authorization server, which serves as a middleman between the client and the resource owner.

2. The implicit grant is a streamlined authorization code flow designed for clients implemented using a scripting language such as JavaScript in a browser.

3. The resource owner's password credentials can be used immediately as an authentication grant to gain an access token.

4. The client's credentials may be used as an authorization grant when the authorization scope is limited to the client's protected resources.

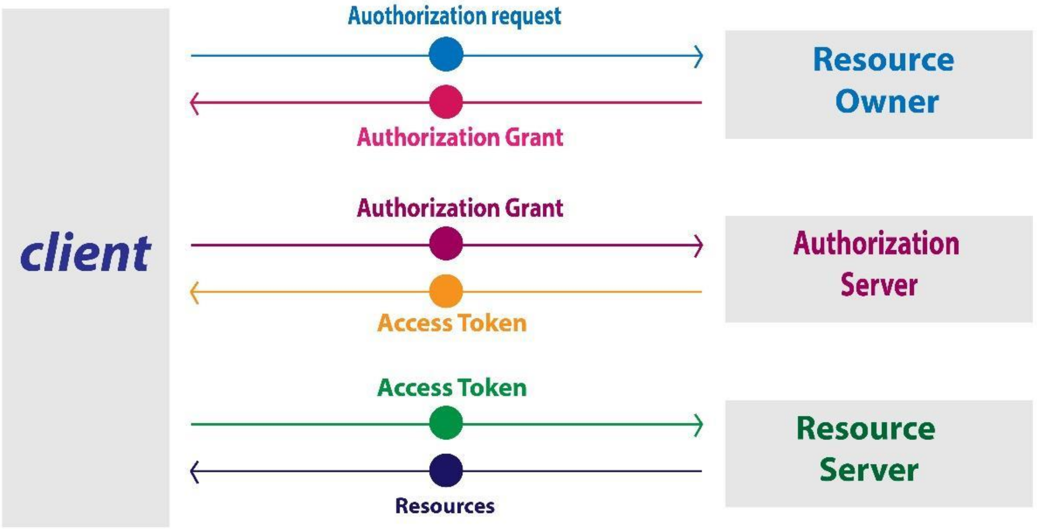

Figure 9 Protocol Flow

3.7.Security Features on The OAuth2.0 Framework

According to Lodderstedt et al. (2013), there are some of the security features built into the OAuth 2.0 protocol to reduce attacks and security issues (Lodderstedt, T., McGloin, M., & Hunt, P. 2013). Tokens such as access tokens, refresh tokens, and authorization codes are used extensively in OAuth. Scope represents the access authorization associated with a token in terms of resource servers, resources, and methods on those resources. The scope is the OAuth framework for handling the power associated with an access token directly. A limited Access Token Lifetime is responsible for limiting the lifetime of an access token and passing this information to the client.

A client uses an access token to gain access to a resource. Because access tokens have short life spans a refresh token allows a specific client to access resources on behalf of the resource owner for an extended period. The client uses an authorization "code" to receive access and refresh tokens. While a redirect URI helps identify malicious clients and avoid phishing attacks from clients trying to trick the user into thinking the phisher is the client, the "state" parameter is used to link requests and callbacks to prevent cross-site request forgery attacks and Client Identifier to increase the level of protection in delegated authorization scenarios (Lodderstedt, T., McGloin, M., & Hunt, P. 2013).

Chapter 4

4.1.The Proposed Approach

After studying several studies about Federated Identity Management and its technology, it would be easier to present the proposed approach that provides access to academic libraries quickly and securely. In our approach in this thesis, we will use Federated Cloud Identity Broker-Model. This cloud identity management model relies on a federated approach that provides the opportunity to use multiple cloud identity brokers to communicate with each other. In this model, users and service providers can authenticate using their preferred cloud identity broker. The user maintains total control over which data is exchanged with the service provider and cloud identity broker and can choose how much information to share with the service provider and cloud identity broker (Zwattendorfer, B., Slamanig, D., Stranacher, K., & Hörandner, F. 2014).

Depending on our use case in this thesis we will use OAuth and OIDC protocols for authentication and authorization. A third-party application can obtain limited access to an HTTP service using the OAuth 2.0 authorization framework (Lodderstedt, T., McGloin, M., & Hunt, P. (2013). The client obtains an access token instead of using the resource owner’s credentials to access protected resources. The OpenID Connect (OIDC) identity layer is developed on top of the OAuth 2.0 framework to enable third-party applications to verify the end-identity users and obtain simple user profile information. OIDC obtains conforming to the OAuth 2.0 specifications by using JSON web tokens (JWTs). JSON Web Token (JWT) is represented using the JWS compact serialization and intended for space-constrained environments. JSON Web Token (JWT) enables the claims to be transferred between two parties and be digitally signed and encrypted. The goal of OIDC is to provide users with one log-in for multiple applications.

This approach enables users to login using their identities with a wide range of identity providers, and it leverages CloudAuthz to obtain authorization to cloud-based resource providers, such as AWS, Azure, and Google Cloud Platform. Users can access academic resources across multiple cloud computing platforms using best practice Web security approaches, thereby minimizing unauthorized data access and credential use risks.

This approach utilizes the authorization code flow of the OIDC protocol to authorize and authenticate a user. First, a user's identity is verified by an Identity Provider. After that, realm (B) receives security tokens from the Identity Provider. The security contains claims about the user's authentication and authorization. There is a two-step procedure for accomplishing this flow, as detailed below:

A. The administrator of realm (B) sets up the realm for the authorization code flow by registering it with the identity provider and obtaining security credentials for realm

(B) and using these credentials to ensure the authenticity of communications between the parties.

B. A realm (A) authenticates a user via sending a request to the Identity Provider. The request includes Security tokens, Redirect URL, Anti-forgery claims.

After successful authentication, an Identity Provider forwards to the realm an authorization code and the state token (B). Then, the realm (B) uses the state token to validate the redirect message's authenticity and associate the authorization code with a user of the realm (B). The realm (B) then exchanges the authorization code for an ID token and a refresh token from the Identity Provider. Each resource provider has an established method for authorizing a client to assume a role.

In This approach, the OIDC protocol is used to federate users' identity and authentication. Using a temporary identity token, an identity provider authenticates a user to the realm (B) in this model. The realm (A) then uses this token to obtain cloud-native credentials and access protected resources through the resource provider's API. The realm (A) automatically refreshes the token as a trusted party to continue operating on the user’s behalf beyond the validity of the initial token.

4.2.The Advantages of Proposed Approach

Individuals gain control over the use and sharing their identity attributes within the federation through federated identity management systems (Jensen, J. 2011). By federating users' identities across several security domains, the user can authenticate to one domain and then access resources in the other domain without authenticating again. Users' security can improve because of federated identities. Additionally, reducing and strengthening authentication operations would also reduce the risk of identity theft (Madsen, P., Koga, Y., & Takahashi, K. 2005).

It is easier for an IT administrator to handle fewer user identities across multiple applications. The many methods used for federated login allow the user to have only one login credentials set, minimizing the amount of administrative effort needed. There would be no need to deal with new users' problems. It will be just about giving secure access. Moreover, there is no longer a need to handle users or identities that are not under their control individually, significantly lowering the cost of identity life cycle management. Having several login credentials opens to several security risks. In such a situation, federated login allows organizations to resolve this difficulty while also mitigating security risks. Federated identity solutions make several security features easier to implement.

Most websites with a registration mechanism discover that some of the users who begin the process never complete it. As websites request additional information, the percentage of registration success drops even more. A federated login mechanism increases success rates by automating most of the access process and eliminating users' need to recall a website password which provides a seamless and trusted user interface with a single registration and sign-on so that users can conveniently navigate between Web pages using a single identity and monitor the release of their data directly. Reduce identity and security protection costs by linking and reusing identities through organizations. There is no longer the need for duplicating data to share the academic resources among the researcher community.

Students have become accustomed to requiring an ever-increasing number of identities to complete their daily academic tasks. Users want to connect their identities once to prove ownership and then be allowed to perform actions using this set of identities (Kushida, K. E., Murray, J., & Zysman, J. 2015). Allowing students to access other universities' resources and learn about new teaching methods will help improve teaching in universities that do not have good learning resources. Some universities have bilateral agreements with search engines to provide more research resources to their students, improving their research quality. If this service is shared, it will help universities with lower budgets deliver it to their students who work in the research field.

Chapter 5

5.1.Conclusion

Academic collaboration has immense potential to improve the research and teaching qualities of universities. By collaborating, the researchers can share research data, documents, experimental testbeds, high-end computing equipment, etc., and can solve complex problems which cannot be solved by a single investigator's efforts. Despite its immense potential and fundamental dependency, collaboration across universities is extremely challenging due to underlying trust issues. In addition, the collaborating entities face the difficulty of creating a new identity for each new collaboration they sign-up for. To mitigate these current limitations, in this thesis, we proposed Federated Identity Management (FIM) approach for facilitating secure resource sharing among collaborating associates without creating new identities. In the first part of the thesis, we provided a comprehensive literature survey of identity and access management and discussed the privacy issues associated with identity management that can be addressed using FIM. The users and service providers in FIM have the advantage to be authenticated using their preferred cloud identity broker which is a unique feature not available in the traditional IAM model. This additional feature in FIM reduces the cost of user management as well as makes the security integration and user experience seamless. Although the identity broker may raise individual privacy risks, an identity ecosystem of federated identity solutions can play a pivotal role in achieving more secure cyberspace by preventing identity providers from collecting information about users' identities. The second part of the thesis provides the comprehensive overview and security features of the OAuth 2.0 framework which is an industry-standard protocol for authorization and user management used by FIM. We also enumerated all the critical implementation steps and detailed design of the proposed framework. Our future work will involve additional methods, such as case studies at organizations using federated identity management and interviews with users who extensively use it. We believe that such case studies would undoubtedly reveal salient requirements and contributed to a better framework design.

5.2.Future Research

The cloud provides several benefits to reduced costs and time to deliver solutions and share data. Federated identity is a maturing part of this revolution. This research only takes the universities that use cloud services into account and focuses on federated identity as an essential factor for collaboration between universities on cloud computing. The proposed approach can be generalized for collaboration to allow university and school systems to provide central shared services worldwide and establish partnerships that allow students from other partner universities to enroll in courses and log in to their campus learning system. More research should be done to develop better use of federated identity models for university cooperation. Cloud service providers must attempt to implement these identity models to protect university resources and user's data adequately.

References

[1] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.

[2] Dillon, T., Wu, C., & Chang, E. (2010, April). Cloud computing: issues and challenges. In 2010 24th IEEE international conference on advanced information networking and applications (pp. 27-33). Ieee.

[3] Vaquero, L. M., Rodero-Merino, L., Caceres, J., & Lindner, M. (2008). A break in the clouds: towards a cloud definition.

[4] Kushida, K. E., Murray, J., & Zysman, J. (2015). Cloud computing: From scarcity to abundance. Journal of Industry, Competition and Trade, 15(1), 5-19.

[5] Singh, A., & Chatterjee, K. (2017). Cloud security issues and challenges: A survey. Journal of Network and Computer Applications, 79, 88-115.

[6] Dotson, C. (2019). Practical Cloud Security: A Guide for Secure Design and Deployment. O'Reilly Media.

[7] Hardt, D. (2012). The OAuth 2.0 authorization framework.

[8] Indu, I., Anand, P. R., & Bhaskar, V. (2018). Identity and access management in cloud environment: Mechanisms and challenges. Engineering science and technology, an international journal, 21(4), 574-588.

[9] Zwattendorfer, B., Zefferer, T., & Stranacher, K. (2014, April). An Overview of Cloud Identity Management-Models. In WEBIST (1) (pp. 82-92).

[10] Jansen, W., & Grance, T. (2011). Sp 800-144. guidelines on security and privacy in public cloud computing.

[11] Hörbe, R., & Hötzendorfer, W. (2015, May). Privacy by design in federated identity management. In 2015 IEEE Security and Privacy Workshops (pp. 167-174). IEEE.

[12] Jensen, J. (2014). Federated Identity Management in the Norwegian Oil and Gas Industry.

[13] Partners, L. C. S. Cloud Federation and Federated Access Control.

[14] Jensen, J. (2011, August). Benefits of federated identity management-A survey from an integrated operations viewpoint. In International Conference on Availability, Reliability, and Security (pp. 1-12). Springer, Berlin, Heidelberg.

[15] Laszewski, T., & Nauduri, P. (2010). Migrating to cloud. Waltham, USA, Elsevier. UNDP (2012), World Population Prospects: The.

[16] Indu, I., Anand, P. R., & Bhaskar, V. (2018). Identity and access management in cloud environment: Mechanisms and challenges. Engineering science and technology, an international journal, 21(4), 574-588.

[17] Bulusu, S., & Sudia, K. (2013). A study on cloud computing security challenges.

[18] Grassi, P., Lefkovitz, N., & Mangold, K. (2016). PRIVACY-ENHANCED IDENTITY FEDERATION.

[19] Maler, E., & Reed, D. (2008). The venn of identity: Options and issues in federated identity management. IEEE security & privacy, 6(2), 16-23.

[20] Jalili, V., Afgan, E., Taylor, J., & Goecks, J. (2020). Cloud bursting Galaxy: federated identity and access management. Bioinformatics, 36(1), 1-9.

[21] Ahn, G. J., & Lam, J. (2005, November). Managing privacy preferences for federated identity management. In Proceedings of the 2005 workshop on Digital identity management (pp. 28-36).

[22] Pardeshi, V. H. (2014). Cloud computing for higher education institutes: architecture, strategy and recommendations for effective adaptation. Procedia Economics and Finance, 11, 589-599.

[23] Ahn, G. J., & Lam, J. (2005, November). Managing privacy preferences for federated identity management. In Proceedings of the 2005 workshop on Digital identity management (pp. 28-36).

[24] Ghazizadeh, E., Zamani, M., & Pashang, A. (2012, December). A survey on security issues of federated identity in the cloud computing. In *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings* (pp. 532-565). IEEE.

[25] Madsen, P., Koga, Y., & Takahashi, K. (2005, November). Federated identity management for protecting users from ID theft. In Proceedings of the 2005 workshop on Digital identity management (pp. 77-83).

[26] Ahmed, M., & Petrova, K. (2020). A Zero-Trust Federated Identity and Access Management Framework for Cloud and Cloud-based Computing Environments.

[27] Temoshok, D., Temoshok, D., & Abruzzi, C. (2018). Developing Trust Frameworks to Support Identity Federations. US Department of Commerce, National Institute of Standards and Technology.

[28] Broeder, D., Wartel, R., Jones, B., Kershaw, P., Kelsey, D., Lüders, S., ... & Weyer, H. J. (2012). Federated identity management for research collaborations (No. CERN-OPEN-2012-006).

[29] Zwattendorfer, B., Slamanig, D., Stranacher, K., & Hörandner, F. (2014, September). A federated cloud identity broker-model for enhanced privacy via proxy re-encryption. In IFIP International Conference on Communications and Multimedia Security (pp. 92-103). Springer, Berlin, Heidelberg.

[30] Lodderstedt, T., McGloin, M., & Hunt, P. (2013). OAuth 2.0 threat model and security considerations. IETF2013.