

Best Practice Authentication and Authorization

in Distributed Business Scenarios with OpenID Connect and OAuth 2.0

Master Thesis

submitted in conformity with the requirements for the degree of

Master of Science in Engineering (MSc)

Master's degree programme **IT & Mobile Security**

FH JOANNEUM (University of Applied Sciences), Kapfenberg

Supervisor: Elmar Krainz, FH JOANNEUM Kapfenberg

submitted by: Cornelia Rauch
personnel identifier: 1610419026

August 2018

**Assignment for the master thesis of
Cornelia Rauch
Matr. no. 1610419026**

**Subject:
“Best Practice Authentication and Authorization”**

Abstract

The pressure for developers to find appropriate authentication and authorization methods is extreme. The selected architecture for the authentication and authorization of a system should protect sensible data reliable while not compromising the user experience. This thesis is dedicated to helping developers find an identity management architecture together with comprehensive security controls suitable for a specific use case. Therefore, the identity management system is split up in the categorize identity proofing, authentication, and assertion. The categorize are analyzed, and possible Level of Assurance are depicted. To find the most appropriate LOA and the corresponding security controls for a use case an risk assessment is contacted. For a best practice example, an identity management architecture featuring the security controls suggested by the risk assessment is implemented. The best practice example also considers the requirements of modern application architectures with a light weighted frontend and an API that gives access to private data. After initial implementation of developers should consider an additional risk assessment.

Kapfenberg, 28.08.2018

Academic adviser:

Elmar Krainz

Cornelia Rauch

Formal declaration

I hereby declare that the present master's thesis was composed by myself and that the work contained herein is my own. I also confirm that I have only used the specified resources. All formulations and concepts taken verbatim or in substance from printed or unprinted material or from the Internet have been cited according to the rules of good scientific practice and indicated by footnotes or other exact references to the original source.

The present thesis has not been submitted to another university for the award of an academic degree in this form. This thesis has been submitted in printed and electronic form. I hereby confirm that the content of the digital version is the same as in the printed version.

I understand that the provision of incorrect information may have legal consequences.

Kapfenberg, 28.09.2018

Cornelia Rauch

Acknowledgement

Thanks to everybody who supported me and made it possible to complete this work.

I would like to thank my thesis advisor Elmar Krainz of the FH Joanneum at Kapfenberg. He consistently allowed this paper to be my own work and supported me with useful comments, remarks, and engagement throughout the process.

Finally, I must express my very profound gratitude to my parents and to my sister for providing me with support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

1	Introduction	1
1.1	Problem Statement	1
1.2	Research Question	2
1.3	Hypothesis	2
1.4	Method	3
2	Related Work to Identity Management	4
2.1	Security Considerations	5
2.2	Identity Proofing	8
2.3	Authentication	10
2.4	Assertion	14
2.5	Token-based Authentication	16
2.6	Single Sign-On Federation Systems	24
3	Practical Part	35
3.1	Use Case	35
3.2	Risk Assessment	37
3.3	Solution	47
4	Conclusion and Outlook	60
	References	70

List of Tables

2.1	Identity Assurance Levels (Grassi, Garcia, and L., 2017, p.18)	10
2.2	Authenticator Assurance Levels (Grassi, Garcia, and L., 2017, p.19)	13
2.3	Federated Assurance Levels (Grassi, Garcia, and L., 2017, p.19)	15
3.1	Maximum Potential Impacts for Each Assurance Level (Grassi, Garcia, and L., 2017, p. 25)	39

List of Figures

2.1	"On the Internet, nobody knows you are a dog"(Steiner, 2018)	4
2.2	Componets Authentication	9
2.3	OAuth API	19
2.4	JWT Inheritance	21
2.5	Federation	25
2.6	OAuth Authorization Process (LeBlanc, 2011, p. 353)	30
2.7	OAuth 2.0 and OpenID Connect Lodderstedt, 2014	31
3.1	Identity Proofing Assurance Flow Chart (Grassi, Garcia, and L., 2017, p 27)	45
3.2	Authentication Assurance Flow Chart (Grassi, Garcia, and L., 2017, p 30)	46
3.3	Federation Assurance Flow Chart (Grassi, Garcia, and L., 2017, p 32)	47
3.4	Architecture IdentityServer4	48
3.5	Architecture IdentityServer4 (Brock and Baier, 2018)	48
3.6	IdentityServer Middleware	50
3.7	Self signed Certicate with Powershell	50
3.8	OpenID Discovery Document	51
3.9	JWKS Endpint	52
3.10	API documentation with Swagger	53
3.11	Implicit Code Flow - Identity Server Configuration	55
3.12	Implicit Code Flow - Angular	56
3.13	Secure Page Login	56
3.14	Login Mask Identity Server	57
3.15	Consent Screen	58
3.16	Claims	58

Introduction

With the rise of social networks, online services quite recently experienced a revolution and thus have had a significant impact on the way private information gets propagated on the Internet. Developers split up solutions into backend and frontend applications and build their backend solutions as Application Programming Interfaces (API) which can be used to propagate services online. Frontend applications or even third-party applications can then consume published APIs. This approach especially gained popularity with the rise of mobile and Single Page Applications (SPA), which heavily depend on a light-weighted frontend. However using APIs to access the business logic means, that the trust relationship is not just between two parties, it also can include third-party applications as well. This uncertain trust relationship raises concerns of users and developers regarding security and privacy of personal information. Building trust is a substantial part for architects and developers when designing a system. In order to efface the security and privacy concerns of users, advanced access and identity management system are needed [cf. (Cirani et al., 2015), (Tkalec, 2015), (Røssvoll, 2013)].

1.1 Problem Statement

In traditional architectures, a third party receives the user's credentials from the user, to provide the user with access to information from a protected resource. Prasad (2016) points out that sharing the credentials with a third party might lead to security problems. Third parties might be responsible for fatal security gaps, for example by storing passwords in plain-text. Prasad (2016) further states that the compromise of one third-party will then lead to the compromise of the credentials of the end user. As a result,

the safety of the secure resources cannot be granted anymore. Moreover, third-parties will receive the comparatively more greater amount of access to the user data than needed. Another concern is pointed out by Sakimura et al. (2014), which is that users often are required to create new accounts for each service they want to use. On the one hand, the process of creating a new account for a service can be annoying for the user because they have the burden of keeping track of multiple accounts and remembering multiple credentials. On the other hand, the complex task of managing this accounts can lead to problems like password fatigue and in the worst case to identity theft or the compromise of services [cf. (Sakimura, Bradley, Jones, Medeiros, et al., 2014), (Prasad, 2016)].

This common problems with a modern application setup, where the API represents the backend application raise the question of, how to find the best way to secure these applications and which of the existing approaches is best for a particular case. Furthermore, what risks come with the use of a specific approach and what is the state of privacy?

1.2 Research Question

What are feasible ways to implement authentication and authorization for enterprises that depend on adaptability and focus on modern single page applications? How can businesses find out which identity management system is best suited for their use case? Furthermore, what are benefits of the token-based authentication with OAuth2 and OpenID Connect versus traditional authentication methods?

1.3 Hypothesis

In modern architectures, a third party often is responsible for the implementation of the process of authentication and authorization of the user. Considering the setup of a company and the use case of the authentication and authorization of users, the requirements towards identity management can be very different. Grassi, Garcia, & L. (2017) suggest for identity management, to rely on standards and specifications that are widespread and documented. Example for tools and standards include Internet protocols for accessing services like REST, SOAP, and XML and federate identity standards for service authentication such as SAML, OAuth, and OpenID Connect.

The focus of this thesis is on modern distributed architectures and ways to provide users secure access to protected resources and confidential management of the user's credentials for a specific scenario. The hypothesis of this thesis is that the best-suited identity management system for a given use case can be found by conducting a risk assessment and finding accurate Level of Assurance (LOA) and security controls to reduce common identity threats and mitigate risks while keeping a balance between security and usability.

1.4 Method

In order to prove the described hypothesis the following method is applied:

- Background and related work about authentication and authorization methods
- Risk assessment of a use case
- Development of a best practice example

The priority was to make extensive research on the topic to get an overview of the current state of authentication and authorization methods. After the research, a case for an assessment is defined and described in detail. Based on this case a risk assessment is executed, and the appropriate level of assurance is determined. The level of assurance need for this case allows choosing the appropriate technologies to mitigate risks and the impact of errors. With this set of technologies and example solution for a modern enterprise single page application is created to give a best practice example based on the theoretical gathered knowledge.

Related Work to Identity Management



"On the Internet, nobody knows you're a dog."

Figure 2.1: "On the Internet, nobody knows you are a dog"(Steiner, 2018)

The cartoon of (Steiner, 2018) in figure 2.1 showing two dogs sitting in front of a computer with the iconic title cited above became an illustration of how people view anonymity on the Internet. Being anonymous on the Internet, however, is not that easy anymore. The use of mobile devices has changed the way how we access information, interact with each other and share content. Digital services are used to access information, which is opening fraught opportunities to attackers who are trying to impersonate someone and access confidential information. With this change of user behavior,

the way we think of authentication and authorization methods has to adjust. Identity management is mandatory to provide a seamless user experience, including identity proofing, authentication, and assertions in federated environments [cf. (Grassi, Garcia, and L., 2017), (Corre et al., 2017)].

Users find themselves struggling using multiple devices, accounts, and services. The user's burden of this site-by-site account management is putting security at risk. The goal of new authentication and authorization solutions is to help the user managing his accounts by providing single sign-on (SSO), based on an exchange of identity-related assertion across security domains in a scalable way [cf. (ibid.)].

In this chapter, existing work and methods to authenticate and authorize users are described. The part is split up in the chapters Security Considerations, Identity Proofing, Authentication, Assertion, Token-based Authentication and SSO Federation Systems.

2.1 Security Considerations

Before getting further into the topic of authentication and authorization, this section will shed light on some basic security principles, concerning authentication and authorization, which help to understand the need for authentication and authorization mechanisms.

Some basic design principles formulated by Saltzer & Schroeder (1975) were paraphrased by Neumann (2013) and are still relevant today. The principles give an underlying overview of what should be the focus when designing a secure system. The first of the ten basic security principles formulated by Saltzer & Schroeder (1975) is the economy of mechanism which means to keep the design as simple as possible. The next principle describes that access should not be explicitly denied; instead it should be explicitly permitted. For example, when using Access Control Lists (ACLs), all access should be explicitly denied by default. This kind of access control is called Fail-safe defaults. Furthermore, Saltzer & Schroeder (1975) states the complete mediation principle, which states that every access to every object has to be checked for authority without exceptions. A fundamental concept that is part of the basic security principles is open design. The design of an application should not be secret. It can not be assumed that design secrecy will enhance security. This principle is often applied in cryptography. For example, the design of cryptographic algorithms is available for the public, and just the keys remain secret. A broadly used principle in authentication is the separation of privileges. Separation of privileges means that two keys should be

used to protect resources if feasible and privileges should be separated. Every application and user should be provided with least privileges they need to complete their job. The existence of overly powerful mechanisms such as superuser is inherently dangerous - this is called least privilege. The least common mechanism principles compel to minimize the number of mechanisms that common to more than one user are used by all users. In authentication users often get frustrated with the complicated sign-in processes. Therefore, the psychological acceptability should be kept in mind. Keep it simple. The design of the interface for the user should be easy to understand so that the user routinely and automatically applies the protection mechanism correctly. The attack factor or work factor is essential to protect sensitive resources. Cost-to-protect should commensurate with threats and expected risks. It should not be possible to circumvent the mechanism with the resources of the attacker. The last one is a recording of compromises which means to provide trails of evidence which are tamper-proof and difficult to bypass. All of these principles are important when choosing an authentication system and need to be considered carefully [cf. (Neumann, 2013), (Saltzer and Schroeder, 1975)].

Besides formulating these fundamental principles which will be discussed in various forms, Saltzer & Schroeder (1975) also discuss the terms "privacy" and "security". Those terms get frequently used by authors writing about information storing systems, like in this paper. However, the terms "privacy" and "security" are often used very differently. Saltzer & Schroeder (1975) for example, defines "privacy" as the ability of an individual to specify whether, when and to whom sensible information is released, and "security" is described as a technique that can control who can modify resources on a computer. Another more recent description of the terms "security" and "privacy" is defined by Brooks et al. (2017). Brooks et al. (2017), states the importance of the distinction between privacy and security. This distinction between privacy and security are essential because there are security issues unrelated to privacy, just as there are privacy issues that are unrelated to security. While security concerns arise from illegal system behavior, privacy concerns arise from byproducts of authorized personally identifiable information (PII) processing. Even byproducts that are considered to protect PII can raise security concerns, for example, it can be questioned to which degree a tool for persistent activity monitoring should reveal information about individuals that are related to security purposes. However, security and privacy have in common that they want to protect personal information and resources or PII [cf. (Brooks et al., 2017), (Saltzer and Schroeder, 1975)].

These security issues and privacy issues, of course, raise particular concerns for users as well as for companies offering authentication services. When it comes to protecting

personal resources, there are three primary concerns. According to Todorov (2007), those three concerns are Confidentiality, Integrity, and Availability. The term confidentiality means that personal information is protected from disclosure to unauthorized individuals and organizations. Integrity or integrity of information is protecting information from accidental or intentional tampering. Modification of confidential data may affect the data validity. Availability is the need to be able to access information at the time a user requests it. The availability of the services that exposes information has to be given as well. In an ideal world companies offering authentication and authorization services will do everything to use the best technologies regarding countermeasures to protect confidentiality, integrity, and availability. Establishing countermeasures, however, can be costly leading to a trade-off between costs and level of production of information. A typical approach to establishing information security management is to analyze risks first and then from counter measurements [cf. (Todorov, 2007)].

The paper Digital Identity Guidelines by Grassi, Garcia & L. (2017), splits the risk assessment into different sections, rather than combining proofing, authentication, and federation of digital services into a single Level of Assurance (LOA). In order to provide the most effective approach each of the section is analyzed regarding risk and impacts of failure. The assessment should help to choose security controls and mitigation strategies that help to avoid errors. Errors that can be avoided with an extensive risk assessment among others are identity proofing, authentication and federation errors. Identity proofing risks include a malicious applicant successfully poses as someone else or for example the impacts of collecting more information about an applicant then required. Choosing the right level of assurance in the identity proofing process gives robustness and confidence to the determination of identity. The right authentication assurance level is chosen to mitigate problems in the authentication process, and the binding between an authenticator and a specific individual. An example of a potential error is the applicant may use credentials that not rightfully belong to them. Last but not least, the accurate federation assessment helps to mitigate potential federation errors. Federation errors might occur when an identity assertion is compromised. Analyzing this risk makes it easier to choose an assurance level for each category. Therefore, the risk gets categorized of harm according to the paper Digital Identity Guidelines by Grassi, Garcia & L. (2017):

- Inconvenience, distress, or damage to standing or reputation
- Financial loss or agency liability
- Harm to agency programs or public interests

- Personal safety
- Civil or criminal violations

Each of the category can then be rated with a potential impact level which can be either low, moderate or high. The result of this risk assessment can then be the basis for choosing the appropriate identity, authentication and federation assurance level. The assurance level determination is just relevant for analyzing the risk of online transaction offered by digital services and should not be used for the complete business process which may include offline processing in an entirely segmented system [cf. (Grassi, Garcia, and L., 2017)].

The result of all design principles, security and privacy considerations should be a system, network or component that is trustworthy. According to Neumann (2018), trustworthiness is given if an entity satisfies its specified requirement, after a reliable assessment. The requirements that deserve special consideration are, those that are critical to an enterprise, mission, system, network or other entity. One of the requirements to make a system trustworthy is a reliable authentication and authorization process. These processes are discussed in more detail in the next sections [cf. (Neumann, 2013)].

2.2 Identity Proofing

“Digital identity is the unique representation of a subject engaged in an online transaction. The process used to verify a subject’s association with their real-world identity is called identity proofing” (Grassi, Garcia, and L., 2017)

A digital identity as explained above is the result of what is called the authentication process. It is a way of identifying the user as who they claim to be. In order to prove the user’s identity presentation, validation and verification of a minimum set of attributes are necessary [cf. (Boyd, 2012)].

To become authenticated, the user has to go through certain steps on the way. These steps are enabled by typical components unique to the authentication process. Todorov (2007) identifies three typical components that are part of the authentication of the Supplicant, the Authenticator, and the Security Database. The Supplicant is the party that provides the evidence to prove the identity of a user or client. The result of the authentication process should be the authenticated user or client. The Authenticator, also called server, is responsible for ascertaining the user identity. After proving of

identity, the authenticator can authorize or audit the user access to resources. The Security Authority Database is a storage or a mechanism to check the user's credentials. The storage can be represented by as much as a flat file, a server on the network providing centralized user authentication or a distributed authentication server [cf. (Todorov, 2007)].

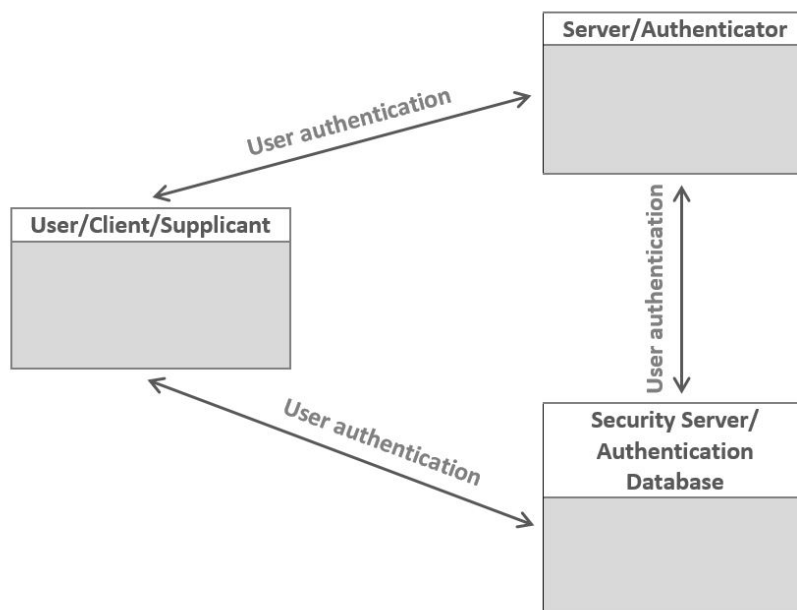


Figure 2.2: Components User Authentication System (Todorov, 2007)

It is vital that all the components as shown in figure 2.2, of a user authentication system can communicate independently of each other. Whether or not all communication channels are used depends on the authentication mechanism and the model of trust that it implements. For example, the Kerberos authentication protocol does not feature direct communication between the authenticator and security server [cf. (ibid.)].

The paper Digital Identity Guidelines - Enrollment and Identity Proofing by Grassi, J.L. Fenton, et al. (2017) also describes components that are similar to the ones described by Todorov (2007) in figure 2.2. Furthermore The paper Digital Identity Guidelines - Enrollment and Identity Proofing by Grassi, J.L. Fenton, et al. (2017) describes three different identity assurance level shown in (table 2.1). These assurance level state to which degree identity assurance can or should be provided in a system [cf. (Grassi, J. L. Fenton, et al., 2017), (Todorov, 2007)].

Furthermore the paper 'Digital Identity Guidelines - Enrollment and Identity Proofing' by Grassi, J. L. Fenton, et al. (2017), describes additional expectations of the identity proofing process. The expected outcome of the identity proofing process should be that

Identity Assurance Level	
IAL1	There is no need to link the applicant to a real-life identity. All attributes provided are self asserted and are neither validated nor verified.
IAL2	The evidence provided supports the real-world existence of the claimed identity. The identity proofing has to be either remote or physical present.
IAL3	The identity proofing requires physical presence. Attributes must be verified by a authorized and trained CSP.

Table 2.1: Identity Assurance Levels (Grassi, Garcia, and L., 2017, p.18)

the claimed identity is linked to a single, unique identity. The provided evidence for the claimed identity has to be validated. The validation should prove that the evidence is correct, genuine and exists in the real world. This section describes certain aspects of the authentication process and gives a very general description of components that can be part of an authentication process. Furthermore, identity proofing assurance level is provided that are relevant for the risk assessment. However, the chapter describes that the user has to provide proof of the claimed identity but lacks to describes how this provided proof might look and how it is validated – this is described in the next (section 2.3) Authentication [cf. (Grassi, J. L. Fenton, et al., 2017)].

2.3 Authentication

For a very typical authentication process, the user provides its username and password when the application demands it. If the user provides a correct username and password, an application assumes the user is indeed the owner of the account they want to log on. The evidence provided by the user in the authentication process is called credentials. Most of the time as mentioned above credentials get provided in the form of username and password. Nevertheless, credentials also may take other forms like PIN's, key cards, eye scanners and so on [cf. (Todorov, 2007), (Boyd, 2012)].

Credentials, which prove the identity of an entity and find use as authenticators in authentication systems, are called factors. The Digital Identity Guidelines by Grassi, Garcia & L. (2017) categorize following types of factors:

- Something, the user, knows - Cognitive information the user has to remember. Examples include passwords, PIN, answers to secret questions.
- What the user has - something the user owns. Examples include a security token, driving license, one-time password (OTP).
- What the user is - biometric information of the user. Examples include fingerprint, voice, and face.

Other types of information which are not considered authentication factors but can be used to enrich the authentication process according to Dasgupta (2017) are:

- Where the user is - the location of the user can be used as a fourth factor of authentication. Examples include GPS, IP addresses.
- When the user logs on - Time can also be extracted as a separate factor. Verification of employee's identification in different office hours can prevent many kinds of grave data breaches. The time factor can easily prevent online banking fraud events to a great extent.

The authenticators are based on secrets that can be either public key pairs (asymmetric keys) or shared secrets (symmetric keys). Public key pairs consist of a private and a public key. The private key gets stored on the authenticator, the holder of the private key can use it to prove they control the authenticator. The verifier of the authenticator can then use the public key - most likely received with the help of a public key certificate - together with the authentication protocol to verify the identity of the user. Shared secrets can be either symmetric keys or memorized secrets. Both keys and passwords can be used with similar protocols. However, there is a huge difference for the user. Symmetric keys are generally stored on hardware; secrets have to be memorized by the subscriber which can lead to multiple vulnerabilities. Cryptographic keys are typically long enough to make network-based guessing attacks untenable but if the user chooses short, memorable passwords the system may be vulnerable [cf. (Grassi, Garcia, and L., 2017), (Dasgupta, Arunava, and Abhijit, 2017)].

To secure a solution properly, it should at least use two factors of the three listed above. To make use of more than one factor of a pool of potential credentials to verify the identity of a user is referred to as Multi-factor Authentication (MFA). The goal of multi-factor authentication is to provide a layered defense and make it harder for unauthorized individuals to gain access. If one of the factors breaks, the service can still rely on the non-compromised authentication factors [cf. (ibid.)].

Using just one factor is called Single Factor Authentication(SFA). Dasgupta (2017)

clearly describes the drawbacks SFA has compared to MFA, primarily the universal used password-based authentication. The user needs to remember different passwords for multiple accounts. Therefore, the user often reuses one password also known as password fatigue [cf. (Dasgupta, Arunava, and Abhijit, 2017)].

In an Interview by Tomkins (2009) with Jon Brody, he explains Password Fatigue like the following. An average user has 15 accounts; some people might even have up to 30 accounts - far too much to manage appropriately. Users then tend to adopt specific password patterns like using simple passwords for nontransactional sites and complex passwords for banking sites. Since many complex passwords are hard to remember users also often reuse passwords for different services at one point - this is called password fatigue [cf. (Tomkins, 2009)].

Besides password fatigue Todorov (2007) draws attention to one of the significant challenges of secure user authentication represented by default passwords. Vendors often ship their devices with pre-configured standard passwords. Although vendors recommend changing default passwords, system architects and engineers often fail to do so because they are more focused on the business logic than on security causing security issues. Systems with default passwords are more straightforward to attack, for example by knowing or guessing the device the attacker has an easy time authenticating with the system accordingly [cf. (Todorov, 2007)].

Due to the problem of password fatigue and default passwords, one factor like a password might easily gets compromised. The user can then no longer use the service until the repair of the system which can lead to a delay of the user when trying to access necessary information. Also, there is a risk that the user does not notice the compromise of the single factor which can lead to devastating effects [cf. (Dasgupta, Arunava, and Abhijit, 2017)].

When designing an authentication process by using multiple factors the designers of the process should be acutely aware of the type of application and the information that has to be secured. For example, a solution for an international bank should have different standards than an application for a maintaining a grocery list. On the one hand, challenging and complex authentication processes for trivial applications might scare away users. On the other hand, simple methods for applications protecting sensitive data might drive users away as well. Application designers should always try to find a middle way that suits both parties, the application owners and the users [cf. (Grassi, Garcia, and L., 2017)].

Furthermore the paper Digital Identity Guidelines - Authentication and Lifecycle Man-

Authenticator Assurance Level	
AAL1	The first level defines that a claimant has to provide some kind of assurance that they control an authenticator bound to a subscriber's account through a secure authentication protocol. The guarantee can be either in the form of a single-factor or multi-factor authentication.
AAL2	The second level requires the use of at least two distinct authentication factors. The confidence that the claimant controls the authenticators of the subscriber's account is high at this level. Appropriate cryptographic techniques are required.
AAL3	The third level provides a very high level of confidence that the claimant is in possession of the authenticators bound to the subscriber. Claimants need to prove possession of the key with two distinct authenticators through a secure authentication protocol. One of the authenticators needs to be hardware-based.

Table 2.2: Authenticator Assurance Levels (Grassi, Garcia, and L., 2017, p.19)

agement by Grassi, J. Fenton, et al. (2017) defines three different authenticator assurance levels to make sure the system a company is designing uses the appropriate authenticators for their purpose. Each of the authenticator assurance levels has to fulfill all the requirements of their successor. The paper Digital Identity Guidelines - Authentication and Lifecycle Management by Grassi, J. Fenton, et al. (2017) describes three levels shown in (table 2.2) [cf. (Grassi, J. Fenton, et al., 2017)].

Authentication is a critical part of every application. As users are getting more concerned on security, the pressure on developers grows to provide a solution that secures sensible data while keeping up usability standards, which can often be a trade-off. Complex applications need complex security which can mean high costs for individually developed solutions; therefore application developers should also think about using federated authentication solutions [cf. (ibid.)].

2.4 Assertion

The authentication and authorization process are very closely related to each other and for users often hard to separate. After the authentication process of a user, the application has now proved that the user is who they claim to be, but not every user is the same. After the user authenticates, the user may want to access data or services. Based on the information provided by the authentication process the application can allow or deny the user to access information or services. In other words, it needs to be evaluated if the user is authorized to access data from a service. Furthermore, authorization offers granular control to distinguish between read, write or execute access to individual resources, typically access control lists (ACL) are used for each operation [cf. (Todorov, 2007), (Boyed, 2012)].

Systems offer a user login process or sign-in process, in order to receive the information necessary to make authorization decisions. The login process initiates the authentication process between the user and the system. As a result of this process where the user has to prove their identity, the user receives a system or application specific structure. In a federated identity system, this is called an assertion. The assertion holds an identifier and identification information about the user which can indicate for example what kind of resources the user can access. For every action, the user then has to provide the assertion and based on the information provided within the assertion the user is then either granted or denied access. Another example of the usage of the assertion is the personalization of websites [cf. (Todorov, 2007), (Grassi, Garcia, and L., 2017)].

In federated systems, the verifier of authentication information of the user is called Identity Provider (IdP), and the party that receives and uses information is called the Relying Party (RP). In the context of the federated identification systems the user, or the one that is trying to access information from a system is called the subscriber. The IdP generates an assertion for the verifier associated with the subscriber. This process allows subscribers to access multiple RPs without maintaining separate credentials and the process also supports SSO [cf. Grassi, Richer, et al., 2017].

The paper Digital Identity Guidelines - Federation and Assertion by Grassi, Richer, et al. (2017), defines different levels of federation assurance that define how assertion should be constructed and secured for a given transaction. Each successive level has to fulfill the requirements of all lower level. Federation assurance levels defined by the paper Digital Identity Guidelines - Federation and Assertion by Grassi, Richer, et al. (2017) are shown in (table 2.3).

For any assertion level, the IdP has to make sure that an RP cannot impersonate the

Federation Assurance Level	
FAL1	The first level requires a bearer assertion, signed with approved cryptography by an IdP, for example, OpenID Connect Basic Client profile or Security Assertion Markup Language (SAML) Web SSO Artifact Binding profile with no additional feature.
FAL2	The second level requires an assertion like the OpenID Connect ID Token or SAML Assertion, to be encrypted with the public key of the RP, whereas the RP is the only party that can decrypt the bearer assertion.
FAL3	Additionally to the first two levels the last federation assertion level requires the subscriber to prove their possession of a key that is bound to the assertion (holder of key assertion) and initially was used to authenticate to the IdP.

Table 2.3: Federated Assurance Levels (Grassi, Garcia, and L., 2017, p.19)

IdP at another RP by signing the assertion. The signing of the assertion can be done by either a MAC using a shared key or a digital signature using an asymmetric key [cf. (Grassi, Garcia, and L., 2017), (Grassi, Richer, et al., 2017)].

One of the factors that differentiate the federation assurance level is the usage of assertion binding. An assertion binding can be chosen based on the RP requirements. It may be required for example that the RP needs additional proof of the binding of an assertion to a particular subscriber. The two different kinds of assertion binding to chose from is the bearer and holder-of-key assertions. Any party can present a bearer assertion in order to prove they have the identity of the bearer. Based on this handling of bearer assertions it can be assumed, that if an attacker is successful in capturing a bearer assertion, the attacker can represent the subscriber that was previously associated with this bearer assertion. The attacker could present the assertion or reference to the RP and impersonate the subscriber. The holder-of-key assertion in comparison holds a reference to a key which indicates which subscriber is representing the assertion. The key is signed and asserted by the issuer of the assertion [cf. (ibid.)].

Typically tokens are used for assertions, in this context, tokens are software tokens which are used to provide access control for systems. The assertion binding described before can also be referred to as token profiles. Typically there are two kind of tokens that are frequently used, which is the access token and the refresh token. An access

token, when it is valid can be used to access resources. If the access token is expired and not valid anymore a refresh token can be used to reclaim a new access token. OpenID Connect additionally defines an ID token which will be discussed later [cf. (Spencer, 2018)].

Furthermore, tokens can be passed either by reference or by value. On the one hand, if tokens are passed by reference, they have to be resolved in another instance. On the other hand, a token can be passed by value and already has all the necessary information about the user to establish for example a session. Those two ways of passing a token are significant when it comes to security considerations. Also, tokens have different types. The OAuth specification, for example, does not specify which kind of token has to be used. Spencer (2018), lists some of the typical tokens:

- WS-Security token (SAML)
- JSON Web Tokens (JWT)
- Legacy tokens
- Custom tokens

In praxis developers often use JWT's because they are very flexible and allow to add additional claims. Also, they use JSON and not XML like SAML tokens, which is much easier to process. More about JSON Tokens can be read in the upcoming chapter Token-based Authentication [cf. (ibid.)].

2.5 Token-based Authentication

The way web developers write back-end applications has changed significantly with the rising popularity of single page and mobile applications. Backend-developers no longer spend a lot of time building markup. Instead, they build APIs for front-end applications to consume. The split up of front-end and back-end allows the back-end to focus on business logic and data management while the front-end solely focus on the representation of the content. The number one way single page and mobile web applications are authenticating users according to Tkalec (2015) is token-based authentication [cf. (Tkalec, 2015)].

Serilleja (2015) shares the view that token-based authentication is the modern way to handle authentication. Token-based authentication should be considered because of various factors. It is optimal for mobile applications which work in a stateless way and

need to adapt to the sudden change of demand in a scalable way. Furthermore, token-based authentication provides extra security and applications can pass on authenticated users to other applications. However, before taking a closer look at token-based authentication, it should be taken into account how Serilleja (2015) and Tkalec (2015) concluded that token-based authentication is the best alternative for modern web applications to authenticate their users. Therefore, Serilleja (2015), for example, examines how authentication was done in the past. One approach to authenticate users used in the past that puts token-based authentication in perspective is server-based authentication [cf. (Sevilleja, 2015), (Tkalec, 2015)].

2.5.1 Server-Based Authentication.

A lot of modern-day API's built on the Representational State Transfer (REST) programming paradigm which basis is the HTTP protocol which is stateless as well. A protocol that is stateless does not recall the actions that were taken beforehand which mean for example that if we authenticate the user, in the next request we have to authenticate the user again because the application will not know the user anymore [cf. (Sevilleja, 2015)].

The aim of server-based authentication is for the application to remember the user that logged on at the application. The application has to store the information on the server, which can be done in a few different ways on the session, usually in memory on the disk. The workflow of a server-based authentication starts with the server delivering the website and the user logging in with username and password. The server saves the information from the user login info in a session. After the session is established, the session is checked on the server for every request. If the session is valid, the server returns the requested data to the server. However, since modern single-application and mobile applications are on the rise, this method to authenticate user shown some problems, especially when it comes to scalability [cf. (ibid.)].

The session handling with server-based authentication is especially hard on the server's bandwidth. Most of the time the session gets established in memory on the server when the user authenticates. This approach leads to an enormous overhead. The second problem with this approach is that the information of the user is held in memory on the server. Since more and more companies are moving servers to the cloud, this is not only a security issue but replicating servers to scale is limited. Also nowadays users want to access their data at any moment from every device. Providing the user with the possibility to access data across multiple mobile devices is vital, which means

cross-origin resource sharing has to be enabled. With server-based authentication, it is possible to run into problems with the forbidden request when the user tries to access data from another domain [cf. (Sevilleja, 2015)].

2.5.2 Token-based authentication with JSON Web Token

The most important thing about token-based authentication is that it is stateless, much like HTTP. The server does not have to hold the session of the user over an extended period on the server. Instead, the user can request resources by offering a token generated by an authentication server. The token send in the query string or Authorization header can then be validated at the resource server, and the secure resource will be returned to the user. The approach of using JSON Web Tokens gives one the ability to scale applications without considering on which domain the user logged on. Another advantage besides being scalable is that token-based authentication gives one the possibility to reuse the same token for authenticating the user. Therefore, it gets easier to build applications that share permission with other application because many separate servers, running on multiple platforms and domains can reuse the same token. The approach also gives performance advantages compared to server-side authentication because there is no need to find and deserialize the session on each request. However, since it is best practice to encrypt the token, the token still needs to be validated, and the content needs to be parsed. One way to implement token-based authentication is with the help of JSON Web Tokens. JSON Web Tokens are gaining popularity fast and are backed by huge companies like Google and Microsoft. Also, the Internet Engineering Task Force defines a standard specification. OpenID and OAuth also use the JSON Web Token as a standard; therefore, the usage of the JSON Web Token will be explained in detail [cf. (Tkalec, 2015)].

JSON Web Token (JWT) is a compact structure that holds information about the authentication of a user or claims. The structure is indented for space-constrained environments such as HTTP Authorization headers. The payload of the JSON Web Token is of JSON, enabling the claims to be digitally signed or integrity protected with a Message Authentication Code (MAC) and encrypted [cf. (Jones, Bradley, and Sakimura, 2015)].

The standard is used to transport data between interested parties. The transferred data can be for example the identity of a user or user's entitlements. Furthermore, with the possibility of digital signatures and encryption data can be transferred securely over an unsecured channel. The signatures also allow asserting the identity of a user if the

recipient trusts JWT is asserting party [cf. (Siriwardena, 2016)].

An example of a JWT Token is an `id_token`. Google provides for Developers an OAuth 2.0 Playground, where developers can choose a scope and try it out against the Google API. To get an excellent example of a JWT, OAuth2 API v2 shown in figure 2.3 was selected and authorized the API [cf. (Google, 2018)].

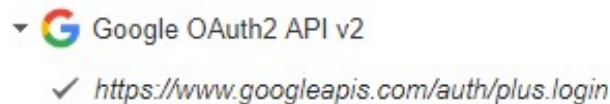


Figure 2.3: Google Developers Playground OAuth 2.0 API

After choosing and authorizing, the API Google returns an Authorization Code. This Authorization Code is specific for a certain Authentication Flow defined by OAuth and can then be exchanged for the tokens. As a result, all information for the scopes that were selected beforehand is returned. Because OAuth 2 API v2 was selected, the returned value is an `id_token`. This `id_token` is a nice representation of a JWT and holds the standard information that is required according to the OAuth specification [cf. (ibid.)].

```

1 eyJhbGciOiJSUzI1NiIsImtpZCI6ImRhZDQ0Nm5NTc2NDg1ZWZMGQyMjg4NDJlNzNh
2 Y2UwYmMzNjdiYzQifQ
3 .
4 eyJhenAiOiI0MDc0MDg3MTgxOTIuYXBwcy5nb29nbGV1c2VyY29udGVudC5jb20iLCJh
5 dWQiOiI0MDc0MDg3MTgxOTIuYXBwcy5nb29nbGV1c2VyY29udGVudC5jb20iLCJzdWIi
6 OiIxMTIzMDE5MzgzMjI4MTAyMzk3MTIiLCJlbWVpYCI6ImNvcn5lbGhcmF1Y2hAZ214
7 LmF0IiwiaWlhaWxfdmVyaWZpZWQiOnRydWUsImF0X2hhc2giOiJSdJlRQjZlZjZFdDZR
8 aHRJWG5laWVnIiwiaXhwIjoxNTI5NzQ3MTk2LCJpc3MiOiJodHRwczovL2FjY291bnRz
9 Lmdvb2dsZS5jb20iLCJpYXQiojE1Mjk3NDM1OTYsIm5hbWUiOiJDb25ueSBSYXVjaCI
10 InBpY3RlcmUiOiJodHRwczovL2x0NC5nb29nbGV1c2VyY29udGVudC5jb20vLWJfBt3
11 aDVaNUJnL0FBQUBQUBQUBQJL0FBQUBQUBQk5rL0p5Qm1XTW9uYU1JL3M5Ni1jL3Bo
12 b3RvLmpwZyIsImdpdmVuX25hbWUiOiJDb25ueSIsImZhbWlseV9uYW1lIjo1UmF1Y2gi
13 LCJsb2NhGUiOiJkZSJ9
14 .
15 yibOtrXy9_-cfOWytzwGuE4zqLv-MK_2-PYIKR_xecJt9ACnMnNMSmio6i8Vu7U061wF
16 OTb-qRennHbvy3lTRZLcTXttIrIUl-NdnZs2BrTSGWrw9aRZeJIHAXiY4fGRHj9VZXs_
17 _J3Nn0EoBmT7Cnua2hb4U_X3hAyGpEv1SGKc5HvbyzoAtNh081Cyj1TI-AidCPTu5vh
18 68C55tLJ87PWNm8WU1rCPOPbDVTjhYlqJKCpgUJ39_p_MXL_uHBZXRRvbOyV_tZVlw47
19 rjd8GFnbQlQqsYAR-6wrFbNL1pY6tPyriqZnQdi5KqYWPuWgPxbFDUfhZAmWXT8-PTsc
20 gQETEP8o3RvRHtfSu8Gx4UOhukt9_VxVdHmpFw

```

Listing 2.1: JWT Token

The JWT Token in the listing 2.1 is presented as a sequence of URL-safe base64url-encoded values. The different values are separated by a (‘.’) character. How many parts a JWT has is dependent on how the JWT is serialized. Either by using the JWS Compact Serialization or JWE Compact Serialization [cf. (Jones, Bradley, and Sakimura, 2015)].

To make sense of this JWT token in the listing 2.1 it is best to look at each of the three parts of the JWT separately. When decoding the first part of the JWT, we receive a JSON object. Each part can be decoded individually but if a quick representation of a token is needed developers are the best advised to use <https://jwt.io/>. The website not only decodes the information of the token it also verifies the token. The website shows a warning if the token is not flawless. The verification of the token and the signing is done with multiple libraries that inform about certain vulnerabilities of the JSON Web Token. The first decoded part of the JWT gives us the following JSON [cf. (ibid.)].

```
1 {  
2   "alg": "HS256",  
3   "typ": "JWT"  
4 }
```

Listing 2.2: JOSE Header

The JSON object in listing 2.2 is the JOSE Header, representing the type of the token, the cryptographic operations applied and optionally additional properties of the JWT. Based on the information of the JOSE Header it can explain if the JWT is a JWS or a JWE. When speaking of JWT, we speak of one of the implementations of JWT because in fact, JWT does not exist itself. Concrete implementation of the JWT is JSON Web Signature (JWS) or JSON Web (Encryption). The figure 2.4 gives an optical representation of the structure [cf. (Siriwardena, 2016)].

The ‘type’ Header Parameter we can examine in listing 2.2 indicates the kind of token, considering the example it is a JSON Web Token. The second parameter is the ‘alg’ Header Parameter and is specified in either the ‘JSON Web Security (JWS)’ reference documentation written by Jones, Bradley & N. (2015) or the JSON Web Encryption (JWE) reference documentation written by Jones & Hilbrand (2015). In this case, the JWT is a JWS since a JWE needs more specific Header Parameter. In the JWS the ‘alg’ Header Parameter gives information about the algorithm which was used to create the signature. Which kind of arguments are accepted by the ‘alg’ Header Parameter are explained in yet another specification called ‘JSON Web Signature and Encryption Algorithms’ written by Jones (2015). In the (listing 2.2 the ‘HS256’ algorithm is used, meaning that the JWS was MACed using the HMAC SHA-256 algorithm. The def-

inition of the 'alg' Header Parameter in the 'JSON Web Signature (JWS)' reference documentation is very similar to the definition in the 'JSON Web Encryption (JWE)' reference documentation by Jones & Hilbrand (2015) except the cryptographic algorithm is used to encrypt or determine the value of the CEK. Jones & Hilbrand (2015), also defines an 'enc' Encryption Algorithm Header, which is used for content encryption on plaintext. The authenticated encryption performed produces a ciphertext and the authentication tag. Also for the 'enc' Header Parameter, the valid arguments can be found in the 'JSON Web Signature and Encryption Algorithms' written by Jones (2015). The algorithm used must be an AEAD algorithm with a specified key length. There is further Header Parameter that can be defined, here are just the essential JOSE Header Parameter for this document listed [cf. (Jones and Hildebrand, 2015),(Jones, 2015), (Jones, Bradley, and N., 2015)].

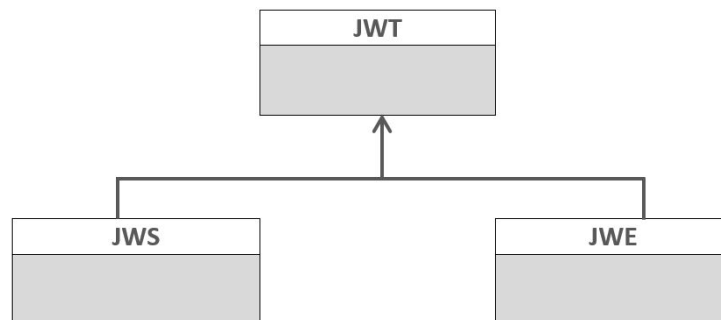


Figure 2.4: JWT Inheritance

The second decoded part of the example in listing 2.1 JWT Token, represents the claim set. The representation of the claim set is a JSON, where each key has to be unique. If the keys are duplicated one can either end up with a JSON parsing error or the last one of the duplicated keys is returned.

```

1 {
2   "azp": "407408718192.apps.googleusercontent.com",
3   "aud": "407408718192.apps.googleusercontent.com",
4   "sub": "112301938322810239712",
5   "email": "corneliarauch@gmx.at",
6   "email_verified": true,
7   "at_hash": "Rv2QB6hg1jt6QhtIXneieg",
8   "exp": 1529747196,
9   "iss": "https://accounts.google.com",
10  "iat": 1529743596,
  
```

```

11 "name": "Conny Rauch",
12 "picture": "https://lh4.googleusercontent.com/-bElKwh5Z5Bg/
13 AAAAAAAAAAI/AAAAAAAAABNk/JyBmWMonaII/s96-c/photo.jpg",
14 "given_name": "Conny",
15 "family_name": "Rauch",
16 "locale": "de"
17 }

```

Listing 2.3: Claim Set

The decoded value from the example in listing 2.1 of the Google OAuth 2.0 Playground by Google (2018), returns a JSON object shown in the listing 2.3. The claim set we receive is composed of mandatory and optional claims. Specifically requested in this example were the login, email and profile scope which affected the claims returned from the Google OAuth 2.0 API. Mandatory claims for the login with OpenID are iss, iat, aud, sub and exp and an optional claim that was returned from the API is azp. This claims will be discussed in more detail in the chapter on OpenID and OAuth 2.0. For the profile scope, the specific claims that were returned are name, picture, given_name, family_name, and locale. Email specific claims are email and email_verified. Furthermore, identity providers can include additional elements that are neither mandatory or optional claims [cf. (Google, 2018), (Siriwardena, 2016)].

The last and third part from the example in listing 2.1 represents the base64url-encoded signature. To know which kind of signature the decoded code is resembling a peek at the JOSE Header will help. The JOSE Header as mentioned before gives information about the cryptographic elements that were used related to the signature. In case of listing 2.4 the Google OAuth 2.0 API uses RASSSA-PKCS1-V1_5 with the SHA-256 hashing algorithm. However, the <https://jwt.io/> website does not provide us with the decoded public key.

```

1 {
2 RSASHA256(
3 base64UrlEncode(header) + "." +
4 base64UrlEncode(payload),
5 Public Key or Certificate.
6 )
7 }

```

Listing 2.4: Signature JWT

To serialize an encrypted message, one has to follow either the JWS or the JWE specification. Each of the specifications has the type's compact serialization and serialization.

Google OpenID Connect uses the compact serialization. The OpenID Connect specification suggest using the JWS compact serialization or the JWE compact serialization. In this paper only the compact serialization will be discussed, the specification of the other serialization method can be either looked up in the JWE specification or the JWS specification. To call a JWS or a JWE, a JWT it has to follow the compact serialization [cf. (Jones, Bradley, and N., 2015), (Jones and Hildebrand, 2015)].

The aim of the JWS Compact Serialization is it to present content as a compact, URL-safe string, which is either digitally signed or MACed. It is not possible to use multiple signature or MAC in a JWS Compact Serialization and furthermore it is not allowed to use JWS Unprotected Headers. An unprotected header is a JSON object which includes the header element that is not integrity protected, which concludes that a protected header is a JSON object that is integrity protected by using MAC or digital signatures. A JWS Compact Serializations is represented as a concatenated string shown in listing 2.5 [cf. (Jones, Bradley, and N., 2015)].

```

1  {
2  BASE64URL(UTF8(JWS Protected Header)) '.'
3  BASE64URL(JWS Payload) '.'
4  BASE64URL(JWS Signature)
5  }
```

Listing 2.5: JWS Compact Serialization

The first element is called the JOSE header which contains all the information that advertises the public key corresponding to the private key that was used to sign the message. Elements of the header include the jku, jwk, kid, x5u, x5c, x5t and x5t#s256. The second element is the JWS Payload or the content to be signed, which does not have to be JSON. The following approach is used to construct the message: ASCII(BASE64URL-ENCODE(UTF8(JOSE Header))) '.' BASE64URL-ENCODE(JWS Payload)). The last element is the JWS signature, which is computed over the message that was constructed beforehand, using the algorithm defined in the JOSE Header. These three key components together are called a JWS token. When using the JWE Compact Serialization, the output is called JWE token. Compared to the JWS token, which we saw above the JWS token consists of 5 different key components [cf. (ibid.), (Jones and Hildebrand, 2015)].

```

1  {
2  BASE64URL-ENCODE(UTF8(JWE Protected Header)) '.'
3  BASE64URL-ENCODE(JWE Encrypted Key) '.'
4  BASE64URL-ENCODE(JWE Initialization Vector) '.'
```

```
5  BASE64URL-ENCODE(JWE Ciphertext) '.'  
6  BASE64URL-ENCODE(JWE Authentication Tag)  
7  }
```

Listing 2.6: JWE Compact Serialization

Both digital signatures and MACs can be used to provide integrity checking. However, specification warns that there significant differences that have to be considered when designing a protocol. MACs only provide the origination of the identification under specific circumstances. It is normally assumed that the private key used for the signature is only known by a single entity. Although in the case of MAC keys all the entities that use it for integrity computation need to know the MAC key in order to validate the message. That means that with MAC validation one can tell if a message is generated from one of the entities that know the symmetric MAC key and not where the message originated [cf. (Jones, Bradley, and Sakimura, 2015)].

2.6 Single Sign-On Federation Systems

Application security is a complex task and developing a customized siloed identity solution can be expensive. A Stand-alone identity store can besides being expensive also causes information assurance and administrative problems for organizations [cf.(JerichoSystems, 2018)].

Boyed (2012) says that a federate authentication or identity federation, is a system that is maintaining its accounts, for example, username and password databases, with the help of third-party service. Often big corporate IT environments already use such solutions. Environment applications, for example, may trust an Active Directory server, an LDAP server or a SAML provider [cf. (Birrell and Schneider, 2013), Boyed, 2012].

As shown in figure 2.5 an identity system usually consists of three parties. The user or subscriber, service provider or resource provider, and the identity provider. Depending on the protocol which is used, different information passes between the three components at different times. While the user or subscriber communicates with the IdP and the RP over a browser the IdP and the RP can communicate over a front channel which works through redirects involving the subscriber or a back channel which is a direct connection not including the subscriber [cf. (Grassi, Richer, et al., 2017)].

The user or subscriber is usually associated with a real user, and their identity is represented by a set of attributes. A single user can be associated with multiple identities.

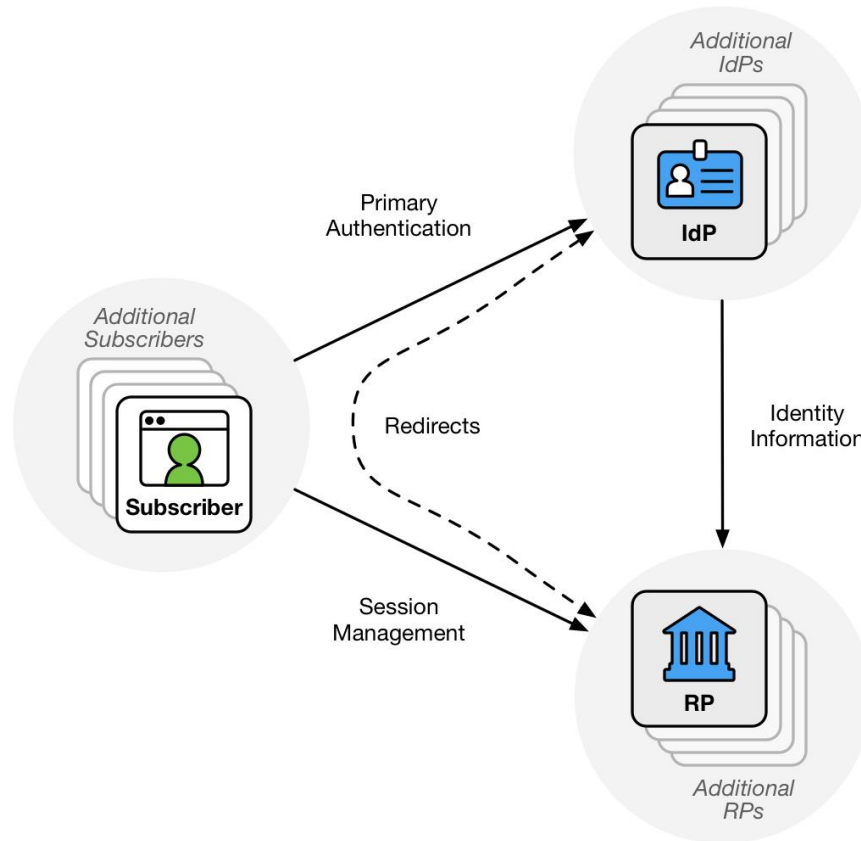


Figure 2.5: Federation

A service/resource provider is responsible for making an authorization decision based on authentication assertion and the attributes derived from it. Many of the service providers implement their identity management, but in federated systems, the identity management is outsourced to the identity provider. An identity provider can be either a stand-alone service or be a part of the service provider. The responsibility of the IdP is it to authenticate the user, issuing authentication assertions, manage identities whereas the IdP has the right to create, update, release and delete attributes associated with the identity [cf. (Birrell and Schneider, 2013)].

The 'Digital Identity Guidelines' by Grassi, Garcia & L. (2017) claim that identity federation is preferred over some siloed identity solution that each serve a single agency or Relying Party (RPs). Furthermore The 'Digital Identity Guidelines' by Grassi, Garcia & L. (2017), lists specific benefits that come with using federated architectures, as can be examined before.

- Enhanced user experience. For example, an individual can be identity proofed once and reuse the issued credential at multiple RPs.
- Cost reduction to both the user (reduction in authenticators) and the agency (re-

duction in information technology infrastructure).

- Data minimization as agencies do not need to pay for collection, storage, disposal, and compliance activities related to storing personal information.
- Pseudonymous attribute assertions as agencies can request a minimized set of attributes to include claims or to fulfill service delivery.
- Mission enablement as agencies can focus on the mission, rather than the business of identity management.

Essentially switching to a federated identity solution should help to reduce the management burden that comes from managing multiple accounts and can lead to potential points of failure. Furthermore, users are giving control over their attributes' dissemination which leads to privacy violations and identity theft. Most of the existing identity federation solution focuses on the individual system, each of which focuses on one of three general types of functionality - SSO, federated identity or anonymous credentials. The SSO system can issue authentication assertion to multiple service providers. Examples for single-sign-on are Passport, OpenID or Shibboleth. Federated identity systems are focused on managing multiple identities for a single user. Examples for federated identity are Project Liberty (<http://projectliberty.org>), Higgins (www.eclipse.org/higgins), PRIME (www.prime-project.eu), CardSpace, and Client-Side Federation. Anonymous credentials systems provide authentication assertions while not revealing the user's identity to a service provider. Examples include Idemix, U-Prove, and P-IMS [cf. (Birrell and Schneider, 2013), (Boyed, 2012), (Grassi, Garcia, and L., 2017)].

2.6.1 Single Sign-On

The focus of this thesis is on mobile SSO solutions. SSO aims to design an authentication system that serves the interests of the user as well as the interests of the service provider. Whereas the user prefers a simple process, the service provider requires a complicated authentication procedure. Ironically trying to make the authentication procedure more save often leads to weakening the whole system due to the user always finding new ways to bypass it. An example mentioned before is password fatigue. Another challenge that SSO is trying to tackle is that standard web authentication solutions that require the user to login with a password, only authenticate the user and are not capable of providing access control or revealing additional information about the user. Most SSO solutions, therefore, try to combine the authentication process and

authorization [cf. (Procházka, Kouřil, and Matyska, 2010)].

The paper 'Taxonomy of Single Sign-On System' by Pashalidis & Mitchell (2003) identifies four generic architectures for SSO systems. An SSO system has to authenticate a user to an Service Provider (SP). Because authentication also implies identification, SSO systems have to incorporate the lifecycle management of identifiers that can take various forms. The paper distinguishes between two main types of SSO systems. The first type is 'pseudo-SSO' and the second type is 'true SSO'. Typical pseudo-SSO are providing automatic authentication for all SP specific authentication methods after the user initially authenticated with the pseudo-SSO component, which is called primary authentication. The responsibility of a pseudo-SSO service is to manage all identities. A user may have multiple SSO identities for a single SP, but in principle, one SSO identity corresponds to one SP. Compared to the pseudo-SSO system in a true-SSO system, the user has to initially authenticate with the Authentication Service Provider (ASP) that is required to have an established relationship with all SPs. The relationship between the ASP and the SP has to be trustworthy. Key functionality of the true-SSO service is that the only authentication that includes the user occurs between the user and the ASP. The SP will get notified of the authentication status, including information about the identity of the user, with so-called authentication assertions. The categorization of SSO architectures can be further distinguished by the location of the ASP/pseudo-SSO component. This component can be either local to the user platform or offered as a third-party service also called SSO-proxy. This further categorization leads Pashalidis & Mitchell (2003) to the four generic architectures mentioned above Local pseudo-SSO systems, Proxy-based pseudo-SSO systems, Local true SSO systems and Proxy-based true SSO systems [cf. (Pashalidis and Mitchell, 2003)].

When opting for a particular SSO architecture, one has to consider the strength and weaknesses of each system carefully. The paper, therefore, analyses the four generic architectures regarding Pseudonymity and Unlinkability, Anonymous Network Access, Support for User Mobility, Deployment Costs, Maintenance Costs, Running Cost and Trust Relationships. Pseudonymity and Unlinkability refer to the fact that the user provides sensible information that needs to be protected and it should not be possible to correlate distinct identities of the same user and personal information, and potential properties should not be linked to the user. The unlinkability cannot be guaranteed for pseudo-SSO because the identities for the SSO are SP specific. To improve unlinkability Pashalidis & Mitchell (2003), suggest to use an 'anonymizing proxy'. For local SSO system additionally services are needed to implement a anonymizing proxy. In proxy-based SSO systems the anonymizing proxy can be integrated. User mobility is supported for proxy-based SSO, for locale SSO there need to be further services in

place. The deployment cost is lower for pseudo-SSO systems than for true-SSO systems. However, the maintenance cost is higher because if any SPs change the whole logic of the pseudo-SSO system has to change. The running cost of pseudo-SSO is likely to be lower than for true-SSO systems. For the pseudo-SSO systems, the trust relationship between users and SP may be dynamically changing depending on the implementation, for true-SSO the trust relationship is established between ASP and SPs and is always consistent. Generally speaking pseudo-SSO systems are more suitable for a closed system where identity management is just for managing the life cycle of remaining credentials. For an open system, it is not enough to maintain credentials. Appropriate privacy protection services and privacy-aware Identity Management schemes should be therefore integrated with true SSO schemes [cf. (Pashalidis and Mitchell, 2003)].

Therefore two very popular assertion technologies are discussed: Shibboleth (SAML) and OpenID Connect. Those technologies are not the only possible SSO assertion technologies. However, they present the most commonly used in federated identity systems. According to Lynch (2011), two SSO solutions gained broad acceptance. The two SSO solutions are SAML-based federations using SOAP with focus on large enterprises, governments and educational networks and the Web Authorization Protocol is a combination of the Protocols OpenID and OAuth. SAML federations have been customized to address the security concerns of those institutions that typically have a large user base, significantly protected resources, complex authorization patterns and data and services spread across multiple domains. However, in a Web 2.0 world, the SAML solutions were seen as too rigid and too severe to maintain; a lightweight SSO was needed, therefore the Web Authorization Protocol was designed. The approach of this protocol is taking advantage of the lightweight RESTful APIs which are reusing the existing HTTP architecture features and the JavaScript Object Notation (JSON) [cf. (Lynch, 2011)].

SAML One of the well-known solutions based on Security Assertion Markup Language version 2 (SAML2) is Shibboleth. Shibboleth is one of the leading middlewares for building identity federations in a higher education sphere. It offers authentication, authorization and attribute assertion between entities. Procházka, Kouřil & Matyska (2010), identifies the following entities Identity Provider (IdP), Service Provider (SP), Discovery Service, Metadata Operator and Federation Operator [cf. (Procházka, Kouřil, and Matyska, 2010)].

The discovery service is used to find the users organization IdP. The IdP defines an

attribute release policy and releases different sets of the user's attributes to different SPs. Users are only able to agree or disagree with the whole set of attributes because the decision comes from the IdP. When the user tries to log on, to a service of a Service Provider, the user gets redirected to a page where an IdP previously found by the Discovery Service can be chosen. If an SP wants to provide their service to multiple federations it has to negotiate policy and technical detail with each federation operator. The federation operator manages the federation policies and introduces SPs and IdPs. Furthermore, the federations operator has to maintain and manage the information about all entities in a federation which is contained in the Metadata. A problem with this architecture is that the SP has to keep track of changes of the technical specification of various federation operators to maintain the configuration for each federation operator. A solution would be a significant federation registration, but this is indeed not possible because of technical and administrative severity and political will. This solution is also somewhat misleading for users since they have to maintain multiple credentials, select from multiple identifiers and so on. According to Procházka, Kouřil, and Matyska, 2010 Shibboleth is too restrictive, a solution with a centrally managed point of IdPs and SPs is preferred. Also, users should not have to deal with redundant accounts [cf. (ibid.)].

OAuth and OpenID Connect The second solution that is introduced is OAuth 2.0 with OpenID Connect. OAuth 2.0 is a protocol that is used for access control. OAuth2.0 is the newest version of OAuth. Compared to the previous version OAuth1.0a, HTTPS is used for all request, and the complicated signature process of tokens was replaced. Furthermore, complexity was reduced, and the usage is much simpler than in the first version. The new version OAuth 2.0 is not compatible with older versions [cf. (LeBlanc, 2011)].

The protocol adds an authorization layer – separating the Client from the Resource Owner. The Client is not bound to any particular implementation characteristics and can make requests to protected resources on behalf of the resource owner. The OAuth 2.0 specification distinguishes between two Client Types, Confidential Clients, and Public Clients. Confidential clients are capable of maintaining the confidentiality and public clients are incapable of maintaining confidentiality [cf. (Hardt, n.d., p.6-14)].

The Resource Owner is a role according to Sakimura et al. (2014) which can grant access to a protected resource also referred to as End-User, if the resource owner is a person. Furthermore, scopes are used to specify to which resources the resource owner wants to grant access. Requests which are sent to an authorization endpoint include

scopes that can be mandatory or optional parameters. Declaring the scopes is up to the provider. An access token is used to gain access instead of the resource owner's credentials [cf. (Sakimura, Bradley, Jones, Mederios, et al., 2014, p.6)].

Access Tokens are issued by an Authorization Server and can be used by the Client to gain access to protected resources hosted by the resource server. The Authorization Server is responsible for issuing the access token and may be the same as the Resource Server. In literature, different terms are used to describe what an identity server does. Among these are Security Token Service, Identity Provider, Authorization Server and IP-STs. OAuth 2.0 uses a token service, which centralizes this logic and reduces complexity for the Client and the APIs [cf. (ibid., p. 6), (Brock and Baier, 2018), (Boyd, 2012)]

A typical request of a Client to access a resource on the Resource Server includes some requests. In the end, the Client should be able to access the protected resource. The process is shown in figure 2.6.

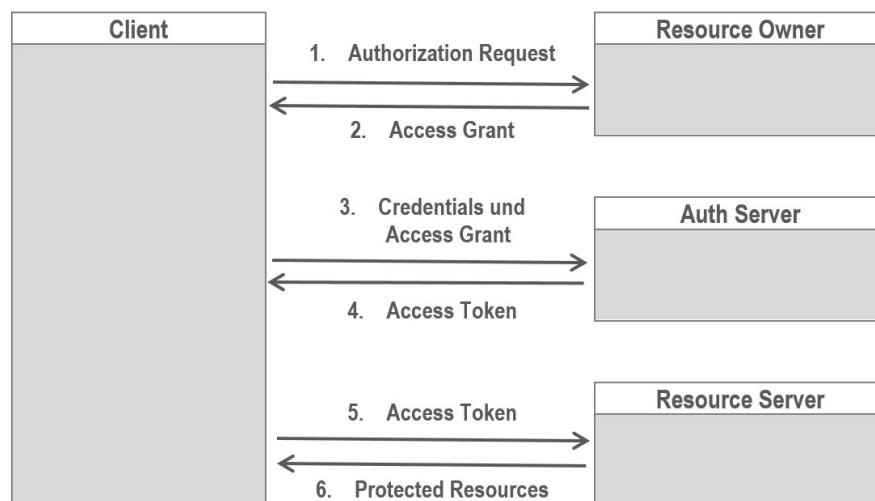


Figure 2.6: OAuth Authorization Process (LeBlanc, 2011, p. 353)

In the first two steps, the Client redirects the User to the Resource Owner, to give the Client permission to access the protected resources. The User is then redirected to the callback location which is included in the original request and adds a verification code which confirms that the Client is entitled to access the requested resource. After the redirection to the callback location, the Client sends a request with the verification code to the Authorization Server. If the request is a success, the Authorization Server replies in step 4 with an access token and eventually a refresh token. With the received access token the client can now make signed requests to the resource server and receive in return protected resources as shown in step five and six in figure 2.6 [cf. (LeBlanc, 2011, p. 353)]

OpenID Connect and OAuth 2.0 seem very similar. OpenID Connect is a framework built on top of OAuth 2.0. Together they focus on Authentication and Access Control in a way that is suitable for modern applications. OpenID can verify the identity of an End-User and obtain basic profile information about the user. It uses Claims to communicate information to the End-User [cf. (Sakimura, Bradley, Jones, Mederios, et al., 2014)]

OpenID Connect performs authentication and determines if the user is already logged on or to log on the user. The result is an authentication token which is returned to the user. A precondition of OpenID Connect is that the client, as well as the user, have to trust the IdP. The functions of the IdP are the same as in OAuth. The reason for that is that the newest version of OpenID is implemented on top of OAuth and extends the authentication process of OAuth [cf. (ibid.), (Boyd, 2012, p. 51)]

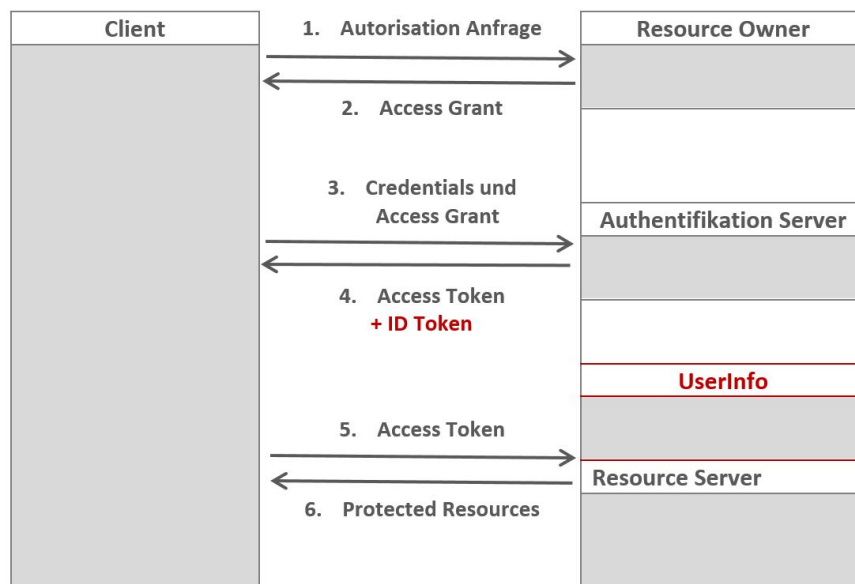


Figure 2.7: OAuth 2.0 and OpenID Connect Lodderstedt, 2014

In figure 2.7 the added Element of the OpenID Connect standard is marked with red. In order to implement the OpenID Connect Standard, the Client has to request the scope 'openid'. After the request, the progress is similar to the transitional OAuth authentication process, but when it comes to returning of the access token, an additional ID Token is returned. This ID Token holds information about the user and the authentication. The only purpose of the ID Token is to log in the client. The only thing left after this is for the client to verify the ID Token [cf. (Lodderstedt, 2014)].

The ID Token is specified by Sakimura et al. (2014) as a JWT compared to OAuth where the type of token is irrelevant. The token is signed and has some particular claims.

Further individual claims could also be added to the ID token. The ID token has claims that are required for the authentication process. Also, standard claims can be added which are predefined by OpenID Connect. Those claims could also be requested from UserInfo Endpoint shown as an additional endpoint in the figure 2.7. The user info endpoint offers an API where additional user info can be requested [cf. (Sakimura, Bradley, Jones, Mederios, et al., 2014)].

Lodderstedt (2014), says that OpenID Connect is a contemporary advancement from protocols like SAML. The protocol considers the ongoing development of web services and mobile applications. Furthermore, the implementation of the protocol is quite easy and uses little HTTP requests with minimal cryptographic overhead to accomplish a complete login process [cf. (Lodderstedt, 2014)].

Sakimura et al. (2014) defines three different way authentication can flow:

1. **Authorization Code Flow** - The Authorization Code Flow returns an Authentication Code which can be used to be exchanged for an id_token or an access_token. The Authorization Code is used for Clients that can maintain a secret securely between themselves and the Authorization Server. Typically this approach is used for server-to-server communication when no interactive user is present. In this thesis, the focus is on user interaction [cf. (Sakimura, Bradley, Jones, Mederios, et al., 2014, p. 10)]

To use the Authorization Code Flow according to Sakimura et al. (2014) the following request_type indicates a Authorization Code Flow:

- code: The response includes an Authorization Code.
2. **Implicit Code Flow** - In the Implicit Code Flow the access token and the ID id token are directly returned to the Client. This behavior may expose information to the End-User or anyone who has access to the End-User's Agent. Tokens are only returned from the Token Endpoint; the Authorization Endpoints is not used. Unlike the Authorization Code Flow where the separate request is used for authorization and authentication, here the Client will receive the access token as a result of the authentication which also means that there is no refresh token. This flow is used for Clients who are implemented in a browser using scripting language [cf. (ibid., p. 20), (Hardt, n.d.)]

According to Sakimura et al. (2014), the following request_types indicate an Implicit Code Flow:

- id_token: The response must include the parameter id_token. The response

does not include an Authorization Code, Access Token, or Access Token Type.

- `id_token` token: The response includes an Access Token, an Access Token Type and an `id_token`.

3. **Hybrid Code Flow** - In the hybrid code flow, tokens are returned from the Authorization Endpoint, and others are returned from the Token Endpoint. The Token Server will return not only the id token but also the code. The Client then has to use the code in order to gain the access token through a back channel call which means there is an additional round trip compared to the Implicit Code Flow. Using the code to receive the access token should minimize the risk of exposing the access token [cf. (Sakimura, Bradley, Jones, Mederios, et al., 2014)].

To use the Hybrid Code Flow according to Sakimura et al. (2014) the following request_types indicate a Hybrid Code Flow:

- `code id_token`: The response includes an Authorization Code and an `id_token`.
- `code id_token token`: The response includes an Authorization Code, an Access Token, an Access Token Type and an `id_token`.

Both OpenID Connect and SAML SSO solution involve different layers of communication and exchange credentials using HTTP redirection and Javascript which leads to vulnerabilities. Vulnerabilities include Cross-Site Request Forgery and Cross Site Scripting. The authentication process for SSO solution includes redirects which can be exploited by attackers. As the SP redirects the User to the IdP, an attacker could intervene and send the user to a fake IdP. If the user does not recognize to swap from the real IdP to the fake IDP, the attacker ends up with the user's credentials. This problem is called a malicious IdP, also known as "phishing" and a very well-known attack against SSO. Another problem that a federation SSO system could be facing is a malicious RP or SP. After a user successfully authenticates on the IdP site, the user is redirected back to the SP with a link. An attacker obtains a link with methods like sniffing and uses the link to get logged into the SP as the User [cf. (Xu, n.d.)].

OpenID Connect and SAML both provide SSO solutions. However, there is the view of the differences between the two protocols. First, in SAM an SP and OpenID, the RP redirects the user to the IdP. The assertions distinguish themselves by the markup language which is used. SAML uses a signed XML document and OpenID uses a signed JSON document. The JSON format makes tokens easier to parse with Javascript. Fur-

thermore, OpenID Connect was designed to work with the web while SAML was readjusted to work on top of the web [cf. (Schwartz, 2016), (KeyCloack, 2018)].

Schwartz (2016), suggest using OpenID Connect for mobile applications, API's and any new applications in general. SAML is just suggested for applications that already use SAML. Nevertheless, KeyCloack (2018) expresses that choosing between this SSO protocol is not straightforward. Although they recommend the use of OpenID Connect, they notice the gradual development of OpenID Connect and how more and more features are implemented that SAML had for years. Therefore some people tend to pick SAML over OIDC, based on the perception that it is more mature [cf. (Schwartz, 2016), (KeyCloack, 2018)].

Practical Part

The practical part of this work consists of three parts: the use case, a risk assessment and an example solution based on what was learned from the risk assessment. In the first part a use case is described. The use case tries to picture a typical scenario that requires thoughtful consideration of authentication and authorization technologies depending on the requirements of the use case. Based on the use case a small risk assessment is done. This risk assessment should help to choose an appropriate level of assurance for identity proofing, authentication and federation. Based on this risk assessment a small application is implemented. The procedure of conducting a risk assessment first should help developers to choose appropriate authentication and authorization methods for a particular use case.

3.1 Use Case

This use case is based on an similar use case of the company "ACP Business Applications" where I am currently employed. This should allow the reader to get a better understanding of the circumstances that led to the design of a federated authentication solution.

The use case features a medium to large company with up to 1000 employees. The mission of the company is to provide a network and analyzing toolset which allows the customer of the product suite to improve security within his network and keep an overview of potential security risks. Previously the customers received monthly reports about their current state of the network. Each of the products and tools resulted in a separate report which was sent via e-mail. To modernize the approach, the reports

should be replaced by a modern single-page application that could be accessed by customers and by employees for administrative reasons. In this example, we call the solution 'Security Assessment Portal.' The Security Assessment Portal should give the customers an overview of current products and the performance of the network. An example of the toolset could be E-Mail-Filtering. The example will be focused on an E-Mail-Filtering-Report and the data which is exposed via the report. It is very important to have a good overview of the data that is used within the application in order to establish appropriate security controls. The Email-Filtering-Report will display a chart per domain. The chart includes data about the amount of e-mails which were blocked, received or carried a virus.

The company already has a particular set of internal applications that are used by the employees within the company network. The company internally uses active directory for authorization on local machines. All applications until now are accessible within the companies network and are using windows authentication. The application should be accessible by employees as well as customers whereas customers should not receive access to the company network. The data of the customers is sensitive and therefore worth protecting. However it should be kept in mind that because of the sensitive data, certain customer do not want to authenticate with an external provider. The application should be hosted within the company and a solution for customers to access the application from outside is needed. The company needs an authentication solution that meets all the outlined requirements while balancing usability and security.

The description gives an overview of the company and the expectations for the new application. This leads to the following requirements for the system architecture:

- Access from outside the company network is enabled
- Application is secured with accounts
- Accounts should be managed within the company
- Only authenticated users have access to secure contents
- Single sign-on for convenience
- API for data access
- Data can be hosted within the company
- The application should be a light weighted Angular solution

The application and its requirements are taken into account in the risk assessment conducted in the next section 3.2. The risk assessment should lead to an better under-

standing which security measurements are needed for this use case.

3.2 Risk Assessment

Risk Assessment is one of the essential elements when it comes to managing risks of an organization. The aim of risk assessment is it to identify, estimate, and prioritize risk to operations, assets, individuals, other organizations and use of the information system. The main task of a risk assessment is to help making decisions on how to respond to certain risk. The steps to a risk assessment are identifying potential threats, internal and external vulnerabilities, the impact to organizations given the potential for threats exploiting vulnerabilities and the likelihood of that harm to occur. There are three tiers according to the 'Guide for Conducting Risk Assessment by NIST (2012) in risk management:

- Tier 1 - organizational level
- Tier 2 - mission/business process level
- Tier 3 - information system level

The first two tiers are focusing on risks related to organizational governance and management activities, mission/business processes, enterprise architecture or the funding of information security programs. The third tier focuses more on how to implement a risk management framework successfully [cf. (NIST, 2012, p.1)].

The result of the risk assessment should state which assurance levels are appropriate. The available assurance levels are identity proofing assurance level, authentication assurance level and the federate assurance level. Identity proofing assurance level describes the robustness of the process to determine the identity of an individual and should help to avoid identity proofing errors. To determine the robustness of the authentication process and the binding of an authenticator and a specific individual's identifier the authentication assurance level is used. Furthermore choosing an appropriate authentication assurance level should help to avoid authentication errors. The last one of the assurance level is the federated assurance level which is optional since not all identity systems need a federated identity solution. The federated assurance level determines the robustness of the assertion protocol the federation uses to communicate authentication and attribute information and should help to avoid federation errors. All of the assurance levels a described in more detail in the sections 2.2 Identity Proofing, 2.3 Authentication and 2.4 Assertion [cf. (Grassi, Garcia, and L., 2017, p.

19)].

To determine assurance levels one needs to identify the potential risk and which measures exist to minimize the impact of those risks. Two factors are crucial for this task. The first factor is the potential harm or impact, and the second the likelihood of such harm or impact [cf. (Bolton, 2003, pp. 3)]. Bolton (2003), splits harm and impact into these categories:

- Inconvenience, distress, or damage to standing or reputation
- Financial loss or agency liability
- Harm to agency programs or public interests
- Unauthorized release of sensitive information
- Personal safety
- Civil or criminal violations

Assurance levels are determined by assessing the potential impact of these categories. The 'Federal Information Processing Standard Publication' by NIST (2004), defines three levels of potential impact concerning security objectives confidentiality, integrity, and availability. Those levels of potential impacts can affect in case of a security breach either the organization or individuals. The levels of impact defined in 'Federal Information Processing Standard Publication' by NIST (2004) are:

- Low
- Moderate
- High

If the level is low the loss of confidentiality, integrity or availability has a limited effect. This can mean for example minor damage to organizational assets, short-term inconvenience, short-term distress, short-term embarrassment, limited reveal of personal or sensitive information to unauthorized parties with minor impact, minor financial loss, risk of civil or criminal violation which would not ordinary be subject to enforcement efforts, minor harm to individuals or the capability of the company to fulfill their mission is degraded. If the level of impact is moderate the effects of loss of confidentiality, integrity or availability are serious. Impacts can be significant damage to organizational assets, significant financial loss, serious short-term or limited long-term inconvenience, serious short-term or limited long-term distress, serious short-term or limited long-term damage to standing reputation, release of personal or

sensitive information to unauthorized parties with moderate impact, risk of civil or criminal violations which may be subject to enforcement efforts, significant harm to individuals or significant degradation of mission fulfillment. The last level of potential impact is high, meaning that the effects can be severe or catastrophic. This means for example major damage to organizational assets, major financial loss, severe or serious long-term inconvenience, severe or serious long-term distress, severe or serious long-term damage to standing reputation, release of personal or sensitive information to unauthorized parties with high impact, risk of civil or criminal violation with special importance to enforcement programs, harm to individuals involving loss of life or serious life-threatening injuries and loss of capability to fulfill the mission [cf. (NIST, 2004, pp. 6), (Grassi, Garcia, and L., 2017, pp. 21)].

The results of the risk assessment are needed to determine the minimum LOA. Once the risk assessment is complete the risk assessment impact profile can then be compared to the impact profiles associated with each assurance level. The table 3.1 can help to determine the required assurance level by finding the lowest level whose impact profile meets or exceeds the potential impact for every category analyzed in the risk assessment [cf. (ibid., p. 25)].

Impact Categories	Assurance Level		
	1	2	3
Inconvenience, distress or damage to standing or reputation	Low	Mod	High
Financial loss or agency liability	Low	Mod	High
Harm to agency programs or public intrests	N/A	Low/Mod	High
Unauthorized release of sensitive information	N/A	Low/Mod	High
Personal safty	N/A	Low	Mod/High
Civil or criminal violation	N/A	Low/Mod	High

Table 3.1: Maximum Potential Impacts for Each Assurance Level (Grassi, Garcia, and L., 2017, p. 25)

When analyzing the risk, it should be considered that there could be more than one error and multiple parties could be involved in the process. Furthermore, the potential impacts description often include relative terms like 'minor' and 'serious' which meaning will depend on the context. Over time and with experience the meaning of these relative terms becomes more definite [cf. (ibid., p. 26)].

One of the most useful documents is NIST's 'Guide for Conducting Security Risk assessment'. The guide follows activities related to NIST'S guide of 'Risk Management Framework' which can be either applied to new or legacy information systems. Typically a risk assessment is a very complex task that takes much time and many external sources to be representative, for example expert opinions. If put together correctly one should end up with a Risk Management Framework (RMF). An RMF provides a process that integrates security and risk management activities into the system development life cycle. The system can be applied to both new and legacy systems. 'Risk Management' by NIST (2018) defines 6 Steps in this RMF. The steps include the categorization of the information system, selection of baseline security controls, implementation of security controls, assessment of the security controls, authorize information system operation based on the determination of risk and monitoring security controls. Many industrial standard risk assessment methods across a wide array of fields and industries are based on these guides. In this practical part, an initial risk assessment is done which should provide a general insight on how to do a risk assessment. The guide to success of a real-life risk assessment is documentation, review, and improvement. However, since a full RMF is out of scope, a scaled down version that is distinctive of digital identity risk management is conducted [cf. (NIST, 2018)].

To prepare for a risk assessment, it helps to identify the purpose, scope, assumptions, and constraints beforehand. The purpose of this risk assessment is to make authorization-related decisions and conducting an initial assessment to identify potential threats, internal and external vulnerabilities which impact the organizations. The scope addresses the 3 Tier - information system level with a focal point on the risk assessment of a single sign-on federation system. Furthermore, the scope is defined by the result expected from the risk assessment. In this risk assessment, the result should be an authentication and authorization solution for a specific use case. The assessment result should usually be reevaluated after the initial draft. [cf. (NIST, 2012, pp. 24)].

Before starting with the risk assessment a risk model has to be chosen. The model can be translated into risk factors which then can be assessed; these risk factors consist of threat, vulnerability, impact, likelihood and predisposing conditions. A threat is a circumstance with the potential to impact the organization. Threat events are caused by threat sources. Threat sources that should be considered in the model can address broad threat sources or particular threat sources. Potential threat sources can be adversarial like Insider, Outsider, Trusted Insider, Privileged Insider, Customer, or Competitor. A thread source can also be accidental like User, Privileged User/Administrator or structural like IT-Equipment (Storage, Communication, Processing), Environmental (Temperature/Humidity, Power Supply) or Natural (Fire, Bombing, Earthquake,

Sunspots). These potential sources can then cause threat events which can be expressed very general or can also be very specific. Which threat are considered depend highly on the company. A company might choose to only consider previously observed events or all possible events. Examples of potential threat sources are cause integrity loss by injecting false but believable data into organizational information systems, obtain unauthorized access, obtain sensitive data/information from publicly accessible information systems, cause disclosure of critical and/or sensitive information by authorized users, conduct externally-based session hijacking, conduct simple Denial of Service (DoS) attack, craft phishing attacks and so on. The likelihood is a weighted risk factor based on the probability that a given threat is capable of exploiting a given vulnerability. A vulnerability is an existing weakness in an information system, system security procedures, internal controls, or implementation that could be exploited by a threat source. Most vulnerabilities are caused by unimplemented or not correctly implemented security controls. Some recommended security controls are described in 'Recommended Security Controls for Federal Information Systems and Organizations'. The implementation of appropriate security controls is an important task that can have a major impact on the operations and assets of a company. Security controls are used to protect the confidentiality, integrity and availability of a system and its information. Examples are Identifier Management, Authenticator Management or Authenticator Feedback. Closely related to vulnerabilities is a predisposed condition which is a condition within a company which affects the likelihood that threat events once initiated, result in adverse impact on the organization like the involvement of PII. The consequences that a successfully exploited vulnerability by a threat source has, is then described by the impact [cf. (NIST, 2012, pp. 17, pp. 35, pp. 70), (Grassi, J. Fenton, et al., 2014)].

Last but not least before stating a risk assessment based on a risk model, an analytic approach and an assessment approach should be considered. The assessment approach can be quantitative, qualitative or semi-quantitative each of the factors including three assessment scales with different corresponding representations. A quantitative assessment is based on the use of numbers whereas the qualitative assessment typically uses nonnumerical categories or levels like very low, low, moderate, high, very high. Finally, semi-quantitative assessment is a combination of a quantitative and a qualitative assessment using bins, scales or representative numbers. For low impact information systems, the qualitative values are used while for moderate-and high-impact systems; the most granular semi-quantitative values are used. The analysis approach is either threat-oriented, asset/impact-oriented or vulnerability-oriented. With the analysis approach the starting point of the risk assessment can be chosen. The risk assessment

starting point can either be based on existing vulnerabilities or assets in the company or possible threats [cf. (NIST, 2012, pp. 17)].

Based on this knowledge a few examples considering the use case above were conducted. Based on the remarks above and on an example from Hudson (2015) the following steps were followed 'Identify Threat', 'Identify Vulnerability', 'Current Controls', 'Likelihood of Impact', 'Effect of Impact' and 'Risk Determination'. Since it is an initial risk assessment, there are no 'Current Controls'. Therefore, this step is left out. The 'Likelihood of Impact' can be assigned a value from low to high whereas high means it is very likely that a vulnerability is exploited by a threat. The lowest value of 'Likelihood of Impact' is 0.1, and the highest is 1. The 'Effect of the Impact' is assigned a value between 0 to 100. A low 'Effect of Impact' means that the consequences are low and high Effect of Impact means that the consequences are severe. To determine the overall risk of a threat, the 'Likelihood' is multiplied with the 'Impact' (Likelihood x Impact) resulting in the risk levels below [cf. (Hudson, 2015)].

- Low = 0-33
- Medium = 34-66
- High = 67-100

Since this analysis is about identity proofing, authentication and federation errors a few threats were chosen to be analyzed with the risk assessment. It should be noted that this does not represent a complete risk assessment since there are more factors to be considered. The risk analysis sees outsiders who want to harm the company and users and identifies them as potential threat sources.

Identity Proofing Risk 1

1. Technical Threat: An attacker successfully proofs as someone else
2. Vulnerability: Phishing over unencrypted network communication
3. Likelihood: 0.5
4. Impact: 80 (Inconvenience, distress)
5. **Risk Determination:** $0.5 \times 80 = 40$ (Medium Risk)

Identity Proofing Risk 2

1. Technical Threat: Collecting and securely storing more information about a person that is required to successfully provide the digital service
2. Vulnerability: Identity Provider is collecting more information than needed from the user on login. Users might not trust the application or not use it
3. Likelihood: 0.8
4. Impact: 30 (Unauthorized release of sensitive information)
5. **Risk Determination** : $0.8 \times 30 = 24$ (Low Risk)

Authentication Risk 1

1. Technical Threat: False claimant using credentials which is not rightfully theirs
2. Vulnerability: User keeps credentials at an insecure place
3. Likelihood: 0.9
4. Impact: 60 (Inconvenience, distress and Unauthorized release of sensitive information)
5. **Risk Determination**: $0.9 \times 70 = 64$ (Medium Risk)

Authentication Risk 2

1. Technical Threat: Account is compromised and user is using same identifier and authenticator for other accounts
2. Vulnerability: User uses same credentials for multiply accounts
3. Likelihood: 0.8
4. Impact: 70 (Inconvenience, distress and unauthorized release of sensitive information)
5. **Risk Determination**: $0.8 \times 70 = 56$ (Medium Risk)

Federation Risk 1

1. Technical Threat: An identity assertion is compromised
2. Vulnerability: Transaction involving third party not over protected channel
3. Likelihood: 0.6

4. Impact: 40 (Inconvenience, distress)
5. **Risk Determination:** $0.6 \times 40 = 24$ (Low Risk)

Federation Risk 2

1. Technical Threat: Identity Server is unavailable
2. Vulnerability: Environmental influences
3. Likelihood: 0.1
4. Impact: 100 (Inconvenience, distress and financial loss)
5. **Risk Determination:** $0.1 \times 100 = 10$ (Low Risk)

An initial risk assessment can take a long time but once conducted the following risk assessments will be much quicker. Modern digital services often combine identity proofing, authentication and federation requirements in one single bundle. It is better to look at each of this components separately to delivery the best identity service. The outcome of the separate component can then be compared to categorize of the table 3.1 and which assurance Level is associated with the risk outcome. Therefore the highest risk is taken and associated with the appropriate assurance level. After the initial risk assessment with just a view threats that were analyzed the associated risk level for identity proofing would be two, for authentication risk, it would be two, and for federated risk, the assurance level would be one [cf. (Grassi, Garcia, and L., 2017), (NIST, 2018), (Hudson, 2015)].

To make it easier to chose an initial assurance level the 'Digital Identity Guidelines' by Grassi, Garcia, & L. (2017) provide an additional assessment on how to choose a appropriate assurance level. Therefore three flowcharts are provided that need to be followed and will then point to an appropriate assurance level. The flow through the cart is represented with the color red. The first flow chart is figure 3.1. Choosing a Identity Assurance Level however does not mean that the identity proofing has to be done by the party who does the assessment - it can also be federate [cf. (Grassi, Garcia, and L., 2017)].

In figure 3.1 the first question is if PII is needed to provide the service. In our use case, probably sensitive company information is provided therefor yes - personal information is needed. Some of the attributes that are provided have to be validated and can not be self-asserted attributes only. The next step covers potential impacts of identity proofing. The two risk that is considered here are the most primary identity proofing

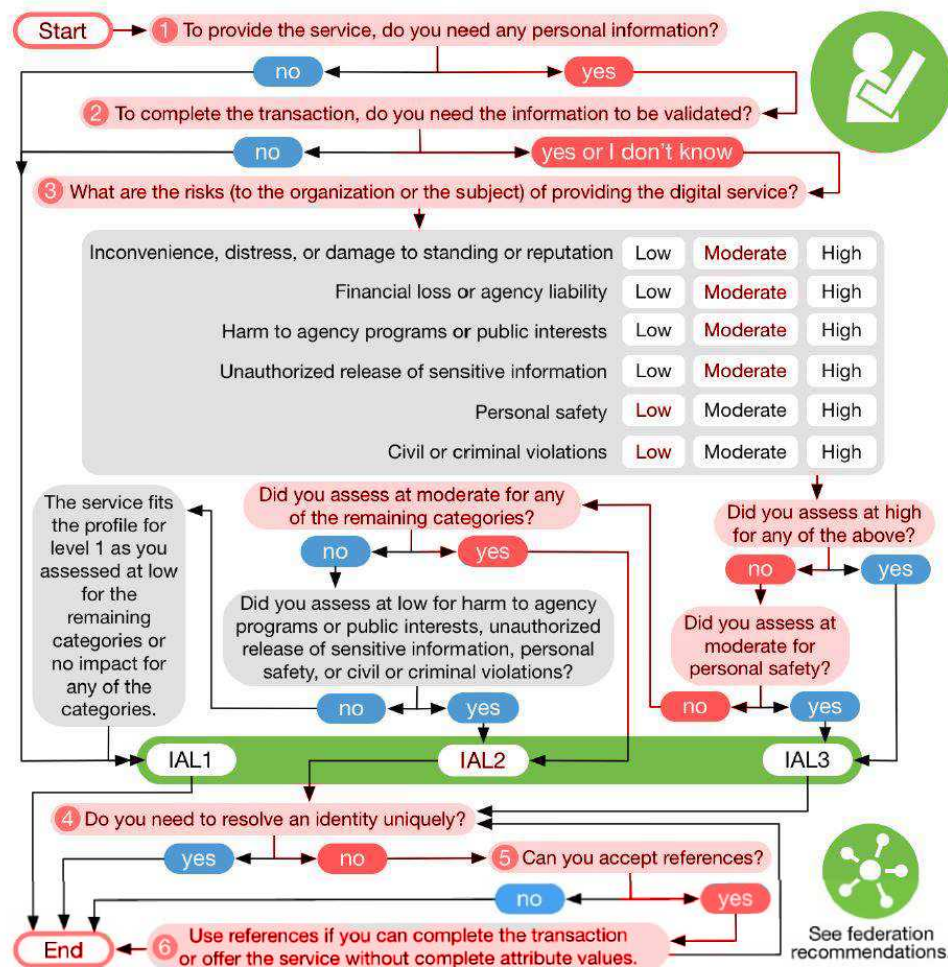


Figure 3.1: Identity Proofing Assurance Flow Chart (Grassi, Garcia, and L., 2017, p 27)

failures from the view of the user and the company: accepting a falsified identity as true and collecting more information as needed. Question 5 focuses on whether the service should be able to access full attribute values. That does not mean that all attributes have to be delivered in claims, but it should be considered. If yes was picked at the question the service is an excellent candidate for a federated infrastructure [cf. (Grassi, Garcia, and L., 2017, pp. 28)].

The next flowchart is figure 3.1 Authentication Assurance Flow Chart which should help choose a company deciding on the authenticators necessary for the service. First, the risks of letting an unauthorized user access information from the perspective of the organization and the user are considered. In this use case company, sensitive information of the customer are endangered an e-mail filtering report however does not have devastating impacts on the customers business. Therefore the assessment ends with the AAL2 [cf. (ibid., pp. 30)].

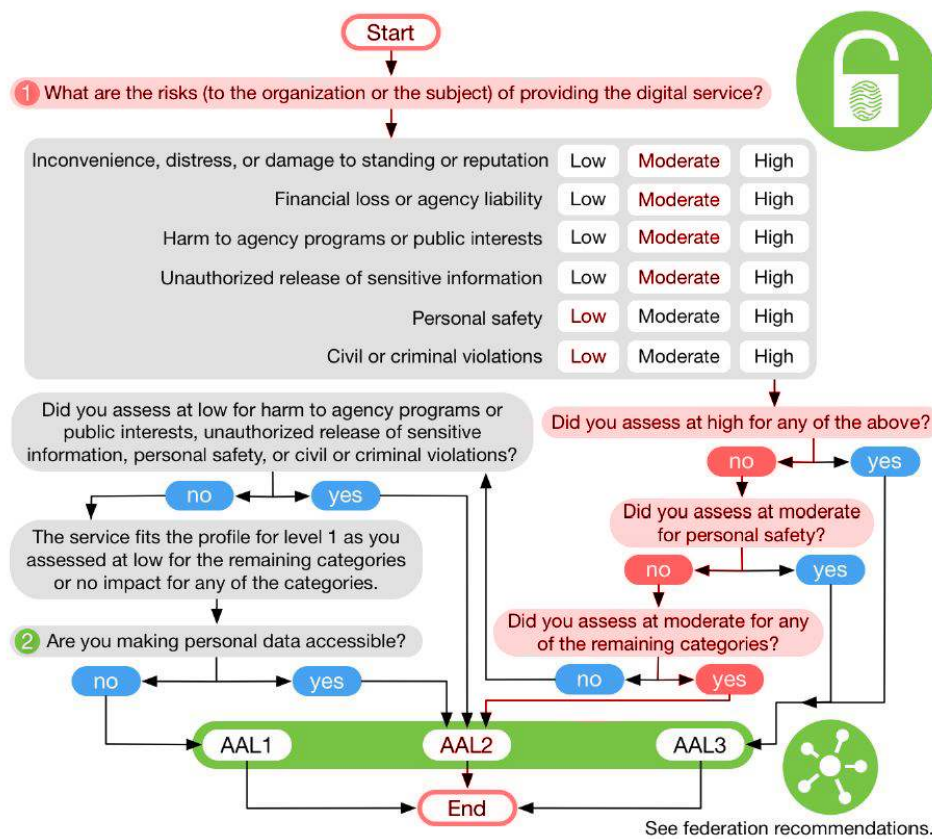


Figure 3.2: Authentication Assurance Flow Chart (Grassi, Garcia, and L., 2017, p 30)

Since it was evaluated in the figure 3.1 Identity Proofing Flow Chart that the service is recommended for a federated solution the last flowchart is the figure 3.3 Federated Assurance Flow Chart. The first thing that is considered is the impact of an unauthorized user compromising an assertion from the organization and the subscriber view. For example assertion replay to impersonate a valid user or leakage of assertion information through the browser. Therefore the assurance level FAL2 is chosen [cf. (Grassi, Garcia, and L., 2017, pp32)].

This initial assertion should give a general direction in which security controls should be in place passed on the different LOA. The risk assessment helps to balance out security and usability, to give the user the best experience. After the first realization of the project, the risk assessment should be done again and refined including the current security controls which are in place.

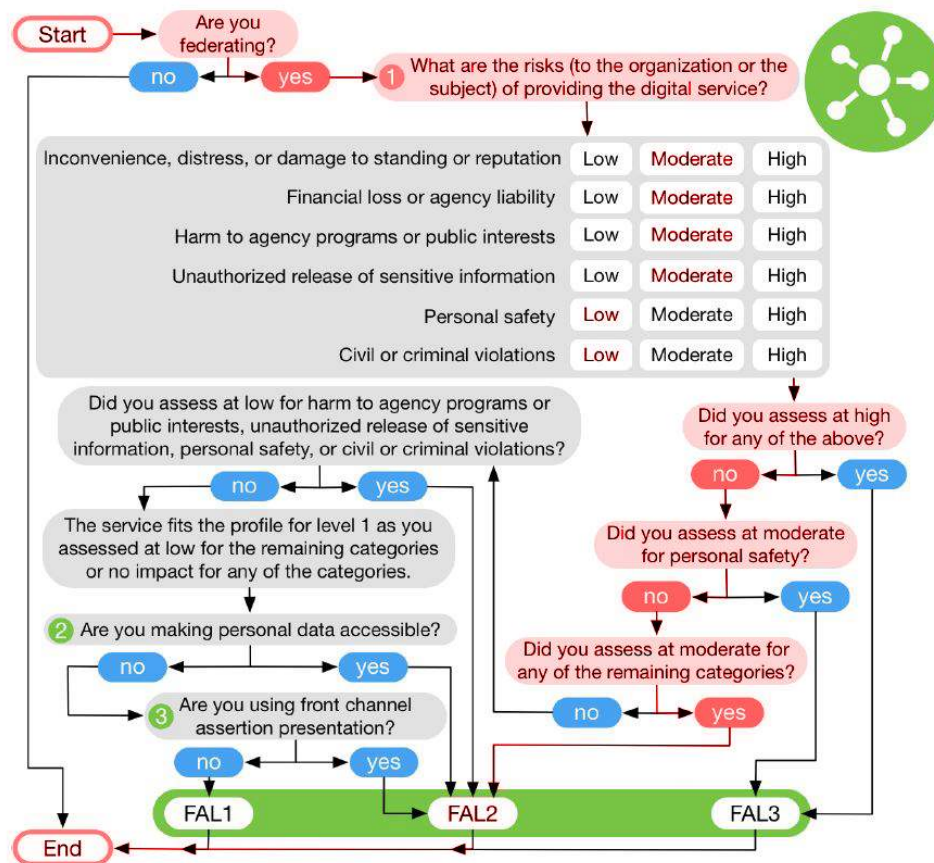


Figure 3.3: Federation Assurance Flow Chart (Grassi, Garcia, and L., 2017, p 32)

3.3 Solution

The Sakimura et al. (2014) explains that modern applications have different requirements than their predecessors, due to new distributed architectures that allow enterprises to be more flexible. Cloud or microservice architectures require different interactions between applications. The below (figure 3.5) shows the most common interactions for modern applications [cf. (Sakimura, Bradley, Jones, Mederios, et al., 2014)].

For example, a browser might call a Web App, and a Web App calls a Web API or perhaps a Native App calls a Web API, which is calling another Web API. Each application has to implement security functions to maintain a secure flow throughout these interactions. Implementing this security features for each involved application leads to a lot of duplicated code and inconsistencies. A different approach to implement security throughout these flows is using a token service. A token service brings the benefit of being able to encapsulate these security functions. Security functions can be updated and hosted at a single point which prevents duplicated functions across

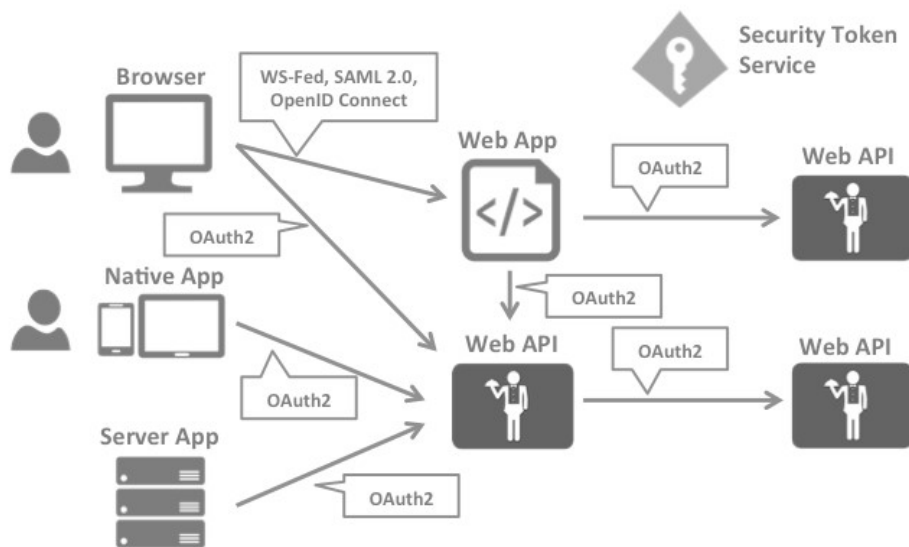


Figure 3.5: Architecture IdentityServer4 (Brock and Baier, 2018)

applications and security flaws. The advantages of a token service or token based application are also outlined in the section 2.5. Furthermore the chapter 2.5 hold examples of JSON Tokens and how they are asserted. The concept of JSON Tokens is very important in token-based authentication systems and ensures confidentiality and integrity [cf. (Sakimura, Bradley, Jones, Mederios, et al., 2014)].

Based on the chapter 3.2 the architecture of the authentication system for the use case described in chapter 3.1 is chosen. The LOA for Identity Proofing should be according to the risk assessment the second one - other LOA are covered in chapter 2.2. The second level of assurance for identity proofing suggests that no real-world existence of the claimed identity has to be present. The other identity proofing assurance levels can be looked up in chapter 2.2 Identity Proofing. The presentation of the real-existence can be either remote or physical. Furthermore the figure 3.1 suggest that since a identity should be resolved uniquely in our use case and it is ok to accept references therefore a federation solution is recommended. In order to successfully proof an identity the client or user has to provide valid credentials to authenticate at the service. The authentication assurance level is also analyzed in the chapter 3.2. According to this initial risk assessment the level of assurance should be two other LOA are described in chapter 2.3. The second level of authentication assurance suggests to use at least two distinct authentication factors and the use of strong cryptographic techniques. However since this was the first risk assessment conducted, the first level for authentication will be used because two distinct authentication factors would be normally used for really sensitive personal data like in a bank app. The first level of assurance for au-

thentication suggest the use of just one authentication factor. Furthermore, since the first identity proofing flow chart also suggested to use federation, the according federation assurance level based on the risk assessment should be the first. The first level of assurance states to use a bearer assertion signed with approved cryptography by an IdP like OpenID Connect Basic Client profile. Therefor the solution will use an token service which login functionality is accepting one distinct authentication factor and asserts the user with bearer tokens.

The languages chosen for the example application are C# and Angular 2. The solution is created with a very useful ASP.NET framework called Identity Server4. Identity Server4 uses the specifics of OpenID Connect and OAuth 2.0 to enable authentication and authorization related features. Features include authentication as a Service, which provides a centralized login logic for all applications, Single sign-on, Access Control for APIs and Federation Gateway. The Identity Server 4 will be used as a Token Server in the practical part [cf. (Brock and Baier, 2018)]

The project illustrating the use of Identity Server and the benefits of OpenID Connect an OAuth is a Visual Studio Project mostly written in C# and Typescript. The project example for Identity Server includes four different projects:

- Identity Server - The Identity Server is a ASP.NET Core Project with basic implementation of an Identity Server which serves as a Token Server.
- ProjectApiNetCore - This application is a ASP.NET Core 2.0 API Application with basic API that returns the users claims if he is authenticated.
- Angular Client - The Angular Client is a ASP.NET Core MVC Angular Project used as a client that can be accessed by an End-User who can authenticate with the Identity Server and requests protected resources of the API ProjectApiNetCore. This project uses the Implicit Code Flow.
- MVC Client - The MVC Client is a ASP.NET Core MVC Project used as a Client that can be accessed by an End-User and can authenticate with the Identity Server and requests protected resources of the API ProjectApiNetCore. This project uses the Hybrid Code Flow.

Identity Server This project is using the IdentityServer4 library by Brock Allen and Dominick Baier. The Project works as a Token Server which means that the Identity Server Project is responsible for the authentication of the user, managing the identities and provide and approve tokens. Furthermore, the Identity Server implements different Authentication Flows.

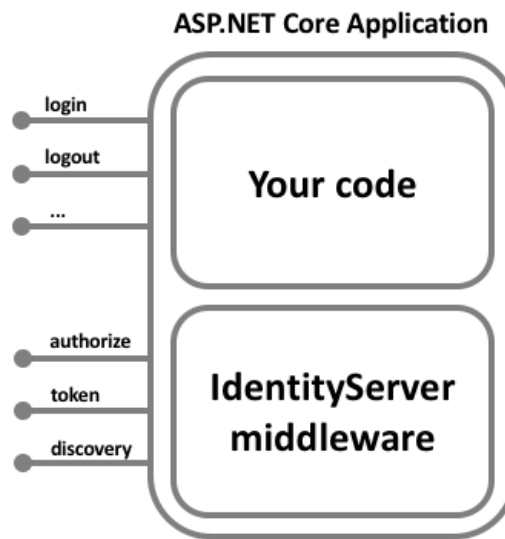


Figure 3.6: IdentityServer Middleware

IdentityServer4 adds endpoints of OAuth 2.0 and OpenID Connect to an arbitrary ASP.NET Core application as shown in figure 3.6. The identity server can be as complex as the developer wants but Brock & Baier (2018) recommend to keep the attack surface as small as possible by just including authentication related UI only [cf. (Brock and Baier, 2018)].

For the basic setup, the Identity Server has to be added to the StartUp class and to ensure a secure communication a certificate is created to sign the request. For the basic setup, Identity Server provides us with `DeveloperSigningCredentials` which provides a dummy certificate. In this example application, a real certificate is included.

```
PS C:\WINDOWS\system32> New-SelfSignedCertificate -Type Custom -Subject "CN=newsts, OU=UserAccounts, DC=acp, DC=acp, DC=at" -TextExtension @(
2.5.29.37={text{1.3.6.1.5.5.7.3.2.1.2.5.29.17={text{upn=cornelia.rauch@acp.at}}} -keyUsage DigitalSignature -keyAlgorithm RSA -keyLength 2048 -CurveExport CurveName -CertStoreLocation "Cert:\CurrentUser\My"

PSParentPath: Microsoft.PowerShell.Security\Certificate::CurrentUser\My
Thumbprint      Subject
-----
88397639321E28E5168BFCA7195A451B1613DCF5 CN=newsts, OU=UserAccounts, DC=acp, DC=acp, DC=at
```

Figure 3.7: Self signed Certificate with Powershell

The certificate used in this application was created with Powershell as shown in figure 3.7. The algorithm used in this certificate is RSA. Depending on the algorithm that is used in the certificate a similar hashing algorithm has to be used for the Client Secret.

The API Resources include the available APIs that can be called and included in the scope of a request. The Identity Resources are Resources that can be included in the returning `id_token`. The clients are the available clients that can be configured and

can use the Identity Server. Those client Configurations indicate which Authorization Flow is used and how to retrieve id_token, access_token, and a possible refresh_token. For simplifying reasons, Identity Server provides an in Memory User Storage which can be used for testing reasons. Other ways to define users are via .NET Core Identity or IdentityServer4 EntityFramework. For the purpose of this implementation this test users were included in the Identity Server application:

1. User1:
 - Username: bob
 - Password: bob
2. User2:
 - Username: alice
 - Password: alice

After the basic setup of the Identity Server, it can be examined if the Identity Server is running correctly by calling the discovery document. The discovery document can be used by the Clients and APIs to retrieve necessary configuration data. The Identity Server has to run at local port 5000 and the navigation path for the browser is: <http://localhost:5000/.well-known/openid-configuration>.

```
{
  issuer: "http://localhost:5000",
  jwks_uri: "http://localhost:5000/.well-known/openid-configuration/jwks",
  authorization_endpoint: "http://localhost:5000/connect/authorize",
  token_endpoint: "http://localhost:5000/connect/token",
  userinfo_endpoint: "http://localhost:5000/connect/userinfo",
  end_session_endpoint: "http://localhost:5000/connect/endsession",
  check_session_iframe: "http://localhost:5000/connect/checksession",
  revocation_endpoint: "http://localhost:5000/connect/revocation",
  introspection_endpoint: "http://localhost:5000/connect/introspect",
  frontchannel_logout_supported: true,
  frontchannel_logout_session_supported: true,
  backchannel_logout_supported: true,
  backchannel_logout_session_supported: true,
  - scopes_supported: [
    "openid",
    "profile",
    "api",
    "offline_access"
  ],
  - claims_supported: [
    "sub",
    "name".
  ]
}
```

Figure 3.8: OpenID Discovery Document

The discovery document shown in figure 3.8 is located at a well-known location and is containing key-value pairs as a JSON structure. The key-value pairs in the discovery

document represent multiple endpoints that are used to authenticate a user including URIs of the authorization, token, userinfo, and public-keys endpoint. The use of a discovery document brings more flexibility to the protocol. The application should have the discovery URL hard-coded in the application according to (I. Google, n.d.) the discovery document URL can then be used to fetch endpoints from the document and use them to send an authentication request for example [cf. (ibid.)].

The metadata that is seen in 3.8 is a mixture of required elements with some recommended ones that are implemented by the Identity Server. The required elements are issuer, authorization_endpoint, token_endpoint, jwks_uri, response_types_supported, subject_types_supported and id_token_signing_alg_supported. The issuer, in this case, is this Identity Servers address. The authorization_endpoint is the URL OAuth 2.0 Authorization Endpoint. The user is redirected to the authorization server's authorization endpoint for authentication and authorization. The authentication request that is sent to the authorization server can have different request parameters. These request parameters are defined by OAuth 2.0 and additional request parameters defined by OpenID Connect. The token_endpoint is the URL of the OAuth 2.0 Token Endpoint. The RP can request access or optionally refresh tokens from the token endpoint. The user_endpoint can be used to request additional information concerning the user. The response_types_supported is a important information about what kind of response is supported and ultimately what kind of authentication flow can be used based on that information described in chapter 2. The subject_types_supported is a list of subject identifier types. Moreover, last but not least id_token_signing_alg_supported is used to find out which JWS signing algorithm is used to encode the ID token and get the claims of the JWT. The algorithm RS256 must be included [cf. (Sakimura, Bradley, Jones, Medeiros, et al., 2014, (Sakimura, Bradley, Jones, and Jay, 2014))].

```

- keys: [
  - {
    kty: "RSA",
    use: "sig",
    kid: "8707eec8504cf2aaaf366585ab82fb54",
    e: "AQAB",
    n: "4E4Auc6FWI_HMxHSU30pbL92PwK8jH4nC47KXzTVHcI-u6r63J
    GZl8mvfn954Bw7jpI59rG4Z0bSyp1gdbxMwzLdkh9Cud__euL0Badl
    alg: "RS256"
  }
]

```

Figure 3.9: JWKS Endpint

The jwks_endpoint is worth a closer look. JWKS SET is published to the JWKS endpoint. The keys shown in figure 3.9 can be rolled over by periodically adding new

keys to the JWK Set. This also indicates if the cryptographic algorithms are configured correctly. For example the message is using the kid of signing key in the JOSE Header to indicate which key has to be used to validate the signature. The algorithm used by the signing party has to be supported by the recipient and can be either an Asymmetric Signature or a Symmetric Signature. When using RSA or ECDSA signatures, the alg Header Parameter has to be set to the correct algorithm, and the private key used to sign must be associated with the public key published by the sender. When using MAC-based signatures, the alg Header Parameter has to be set to the correct algorithm, and the MAC key is the octets of the UTF-8 representation of the client_secret. MAC-based signatures cannot be used by public Clients because they cannot keep secrets. All metadata types are listed in the specification of OpenID Discovery Document from Sakimura et al. (2014) [cf. (Sakimura, Bradley, Jones, and Jay, 2014))].

After calling the OpenID discovery document and the JWKS endpoint, one can be ensured that the basic setup of the IdentityServer works. After that a resource that is worth being protected by the Identity Server has to be created.

ProjectApiNetCore This ProjectApiNetcore is an ASP.NET Core application. This application represents the Resource Server and serves the Protected Resources. In this particular example, the API provides a list of Products the user bought from the company and the possibility to download the report that gives information about the current state of the product. The example API provides dummy data for the particular request. For easy testing of the API and designing the API, swagger.io is used. Furthermore, swagger.io (<https://swagger.io/>) gives a nice developer experience with an interactive API documentation. In addition, it is possible to generate code out of the swagger documentation for different programming languages with tools like NSwag (<https://github.com/RSuter/NSwag>).

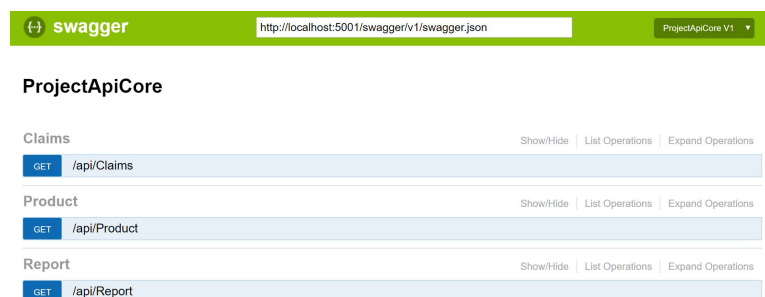


Figure 3.10: API documentation with Swagger

All of the routes shown in figure 3.10, are secured with the help of the IdentityServer4.AccessTokenValidation NuGet package. Each Controller gets an Authorize

annotation. To configure the right authentication method, ASP.NET Core 2.0 offers the possibility to add the authentication to the pipeline with Dependency Injection via the method `Configure Services` in the `Startup.class`. In the `StartUp` class of the API Project the `AddAuthentication` with the Value `"JwtBearerDefaults.AuthenticationScheme"` makes "Bearer" the default authentication scheme. With the `AddIdentityServerAuthentication` it is defined which Token server can be used to validate the incoming token and make sure that this token is from a trusted issuer. It adds the Identity Server access token validation handler into DI and also it is checked if the token is valid to be used with this API (aka scope). Adding `UseAuthentication` to the `Configure` method in the `Startup` class adds the authentication middleware to the pipeline so authentication will be performed automatically on every call into the host. Multiple authentication schemes can be used. A list of authentication schemes is then passed to the `Authorize` attribute separated with a comma. The default scheme results in the `HttpContext.User` property is set to that identity. After adding the Authentication logic to the API Project, the resources should be protected now. To check if the resources are protected the Swagger documentation can be used. Therefore navigate to <http://localhost:5001/swagger/v1/swagger.json> and try out one of the methods shown in figure 3.10. This should result in a 401 Unauthorized HTTP response code.

Client The client represents the UI of the 'Security Assessment Portal' with the initial E-Mail-Filtering functionality. The project is an Angular Solution that follows the Implicit Code Flow described in chapter 2.6.1. The Implicit Code Flow is chosen over the other Code Flows based on the type of application, level of trust and the user experience that should be provided. The first decision is based on rather the machine that hosts the application is also the resource owner. If this is the case no end-user authorization is needed. In this use-case the resource owner is accessible over the API we provided which is hosted on a different server. If the application is a regular web app the Authorization Code Flow would be recommended. However since the client hosted is a Angular SPA app the Implicit Code Flow should be used. The application directly receives an access token and no additional round trips are needed to exchange an received authorization code. The implicit code flow does not return a refresh token because it can not be kept secret but there is the possibility of a silent authentication. The silent authentication is initiated with a request to the authentication request with the parameter `'prompt=none'`. On success the client is redirected to the `redirect_uri` with an successful authentication response. If the user is not already logged on the client is redirected to the `redirect_uri` with an error response [cf. Auth0, n.d.].

For the Implicit Code Flow, the Identity Server has to define a Client that Allows the Grant Type Implicit as shown in figure 3.11. The defined Client can be identified with the ClientId. This particular Client also requests the explicit consent of the Client for the requested Scopes. In this case, the Client is allowed to request the well known Scopes OpenId and Profile. In order to receive an id_token and therefore configure OpenId Connect the scope OpenId has to be requested. The Profile scope gives general information about the user. Also, an API scope is added for 'securityPortalApi', which represents the API Project holding the Resources.

```
// OpenID Connect Implicit Flow Client (Angular)
new Client
{
    ClientId = "ng",
    ClientName = "Angular Client",
    AllowedGrantTypes = GrantTypes.Implicit,
    //Token are send to the browser - the Implicit Code Flow se
    AllowAccessTokensViaBrowser = true,
    RequireConsent = true,

    RedirectUri = { "http://localhost:4427/callback" },
    PostLogoutRedirectUri = { "http://localhost:4427/home" },
    AllowedCorsOrigins = { "http://localhost:4427" },

    AllowedScopes =
    {
        IdentityServerConstants.StandardScopes.OpenId,
        IdentityServerConstants.StandardScopes.Profile,
        "securityPortalApi"
    },
}
```

Figure 3.11: Implicit Code Flow - Identity Server Configuration

AllowAccessTokensViaBrowser makes sure that the access_token can be directly returned to the browser after the client authenticates itself. AllowCorsOrigins allows the given resource to make Cross-Origin-Requests to the Identity Server. Cross-Origin-Request is a request that is made from another domain outside the domain from which the first resource was served. Furthermore, the RedirectUri have to be present, in order for the Identity Server to know where to return after the login attempt was successful.

In the Angular Solution, a service to handle authentication was created. The service handles requests to the API which is hosting the protected resources. The Angular Solution uses the 'oidc-client' library that is suggested by the Identity Server framework to handle authentication requests. To call the Identity Server the following setting are used, shown in figure 3.12.


```
authority: 'http://localhost:5000',
client_id: 'ng',
redirect_uri: 'http://localhost:4427/callback',
post_logout_redirect_uri: 'http://localhost:4427/home',
response_type: 'id_token token',
scope: 'openid profile securityPortalApi',

silent_redirect_uri: 'http://localhost:4427/silent-renew.html',
automaticSilentRenew: true,
accessTokenExpiringNotificationTime: 4,

filterProtocolClaims: true,
loadUserInfo: true
```

Figure 3.12: Implicit Code Flow - Angular

The `client_id` in the Angular Solution has to be the same as the `ClientId` configured in the Identity Server Project. The same is true for the `redirect_uri` and `post_logout_redirect_uri`. Otherwise, the request would result in an error. For example, if a scope is requested which does not exist accordingly to the Identity Server, an invalid scope exception is thrown. Furthermore, the `response_type` has to be according to the used Authorization Flow in the IdentityServer's client configuration. In the Identity Server Solution, the Implicit Code Flow was configured, therefore "id_token token" was defined. The response of these response types will include an access token, an access token type and an id_token.

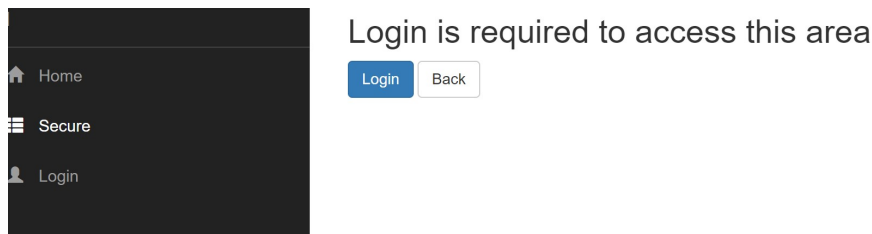


Figure 3.13: Secure Page Login

For the Client to receive an id_token and in this example an access_token from the Authorization Server (Identity Server Project), the client has to prepare an Authentication Request and send it to the Authorization Server. The Angular Client has two possible ways to trigger the login with the Identity Server. In the navigation on the left-hand side there is a Home Page that can be accessed by everyone, then there is a Secure Page that just can be accessed by logged in users and the Login Page, which automatically triggers a request to the Authorization Server.

The figure 3.13, shows the Menu of the application. If the user tries to access the Secure Page without being logged in he is redirected to the "Login is required" Page. In

order to access the secure area, the End-User has to authenticate with the Authorization Server. To authenticate with the Authorization Server the user has to press the Login button. After the Login button is pressed, the End-User is redirected to the Authorization Server Login Page. The design of this screen can be adjusted in the Identity Server Project.

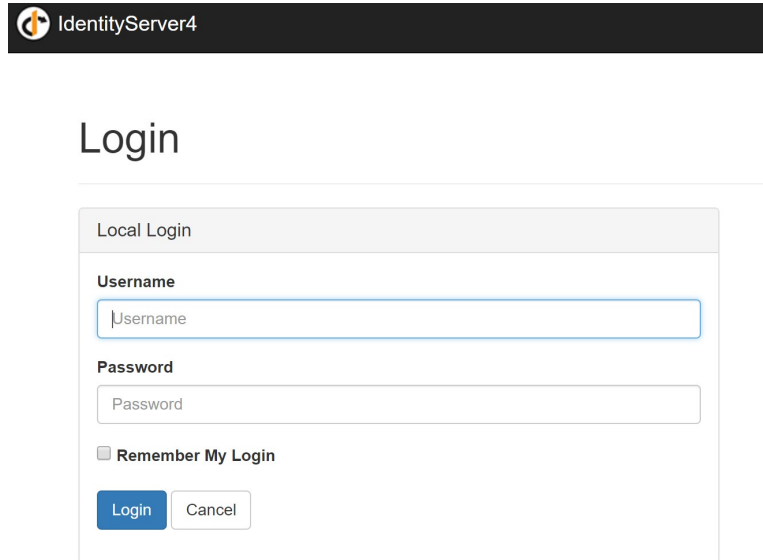


Figure 3.14: Login Mask Identity Server

The (figure 3.14), shows the form where the end-user can log in with the end-user's credentials. In the Identity Server Project, two test-user are registered. Therefore either Alice or Bob can currently log in to the application and view the secured Page. The Authorization Server can also request consent from the User for the requested Claims. In the Identity Server Project, it is configured that the consent of the End-User is required. In consequence of this configuration, the user is prompted an additional screen. This screen is shown in figure 3.15, it would not appear if the consent was not required by the identity server.

The Angular Application is requesting consent for the requested scopes `openId`, `profile` and `securityPortalApi`. The `openId` Scope is required because otherwise, we would not be able to receive an `id_token`. User Profile gives additional information about the User, and the Security Assessment Portal API is a custom API that was added to the scope and allows the user to make requests to this API and the protected resources.

After the End-User has given its consent, the End-User gets redirected back to the Client with an `id_token` and an `access_token`. The `id_token` can be validated and an End-User's Subject Identifier is retrieved.

Angular Client is requesting your permission

Uncheck the permissions you do not wish to grant.

Personal Information

☒ **Your user identifier** *(required)*

☒ **User profile** ⓘ
Your user profile information (first name, last name, etc.)

Application Access

☒ **Security Assessment Portal API**

☒ **Remember My Decision**

Yes, Allow
No, Do Not Allow

Figure 3.15: Consent Screen

If the authentication with the Authorization Server was successful the End-User is now authenticated. As an authenticated user the End-User can now also access the Secure Page. With access_token the received access token calls to the custom API Project can be executed. For example, the claims of the user that are correlating to the scopes can be received and put in a useful context as seen in figure 3.16.

Type	Value
nbf	1534843688
exp	1534847288
iss	http://localhost:5000
aud	http://localhost:5000/resources
aud	securityPortalApi
client_id	ng
sub	818727
auth_time	1534841555
idp	local
scope	openid
scope	profile

Figure 3.16: Claims

This complete example could be a possible implementation of the authentication for the described use case in section 3.1. Now it would also be possible for an external user (customers) to access the application without a domain account. If customers also choose to use other products of the companies, no new account is required because

the same identity server can be used. The personal data of the customers stays in the company and is not given to any third party. With the use of token-based authentication, it is easy to manage the session of end-users, and the user can stay logged in the application (single-sign-on). The use of bear tokens gives appropriate security and the implicit flow used to deliver the token makes it very light-weighted for single page application. The separated API is also convenient because it might be used for multiple purposes.

Conclusion and Outlook

Taking on a new project takes a lot of consideration and planning. The planning includes thoughtful consideration of the security and privacy measures which should be established to create trustworthiness. An essential part of trustworthiness is identity management. Identity management includes ensuring Confidentiality, Integrity, and Availability.

Chapter two of this thesis focuses on a careful examination of identity management, especially under three different angles identity proofing, authentication, and assertion. Furthermore, the difference between using token-based and server-based authentication is analyzed. In token-based authentication systems, it is by far more comfortable to handle sessions, while in server-based authentication systems, this can be challenging tasks. Particular for a single-page application which typically uses an Application Programming Interface hosted on a different server session management can lead to overhead in server-based authentication systems. The thesis also explains the benefits of using a federation system for single sign-on and introduces two possible federation systems for single sign-on. The two systems which are explained are SAML and OpenID Connect. Both of the systems provide the service of Single sign-on, but ultimately OpenID Connect provides a lighter weighted approach for mobile and single page applications.

The second part undertakes the task of finding the best available authentication and authorization methods for a defined use case that requires a single page application and a separated API. A mature part of reaching this goal was the risk assessment of a use case which was described beforehand. The risk assessment examines different technical threats rates and determines the risk of each threat according to an existing approach described in Guide for Conducting Risk Assessment by (NIST, 2012).

Furthermore, flowcharts were consulted which should help to decide which level of assurance would be appropriated and if federating the solution should be considered. This risk assessment lead to three different assurance level which can be categorized into identity proofing, authentication, and assertion.

Based on the level of assurance determinate by the conducted risk assessment a solution was created which is an example of the authentication system developed for an use case. The solution was based on the knowledge gained from the related work part and considered the outcome of the risk assessment. This leads to a solution with a stand-alone identity server which can provide a token service, a separate application programming interface and is able to host private information and a single page application.

The risk assessment was helpful but should be conducted to a greater extent for a real-life application. Furthermore carrying out a risk assessment together with somebody that has had experience in this field would have lead to a more accurate result. Nevertheless, the risk assessment gave some excellent pointers and made decisions more manageable. For example, it was easier to decide rather to use bearer assertions or holder of key assertions.

If the example would be further developed, another risk assessment should be conducted, and the differences should be analyzed, because the key to a successful risk management is documentation review and improvement. In additional risk assessment or in a risk assessment of an legacy application also security controls that are already in place should be included in the analyzes.

Acronyms

AEAD	Authenticated Encryption and Associated Data
API	Application Programming Interface
ASP	Authentication Service Provider
CEK	Content Encryption Key
CSP	Credential Service Provider
DI	Dependency Injection
GPS	Global Positioning System
HMAC	Keyed-Hash Message Authentication Code
HTTP	Hypertext Transfere Protocol
HTTPS	Hypertext Transfere Protocol Secure
IdP	Identity Provider
IP	Internet Protocol
IT	Information Technology
JOSE	Javascript Object Signing and Encryption
JSON	JavaScript Object Notation
JWE	JSON Web Encryption
JWK	JSON Web Key
JWKS	JSON Web Key Set
JWS	JSON Web Signature
JWT	JSON Web Token
LDAP	Lightweight Directory Access Protocol
LOA	Level of Assurance
MAC	Media Access Control

MFA	Multi-factor Authentication
MVC	Multi View Controller
OTP	One-time Password
PII	Personally Identifiable Information
PIN	Personal Identification Number
PKCS	Public-Key Cryptography Standard
REST	Representational State Transfere
RMF	Risk Management Framework
RP	Relying Party
SAML	Security Assertion Markup Language
SFA	Single-factor Authentication
SHA	Secure Hash Algorithm
SOAP	Simple Object Access Protocol
SP	Service Provider
SSO	Single-sign On
STS	Security Token Service
UI	User Interface
URL	Uniform Resource Locator
UTF-8	8-bit USC Transformation Format
WS	Webservice
XML	Extensible Markup Language

Refernce Books

- Boyd, R. (2012). *Getting Started with OAuth 2.0*. Oreilly and Associate Series. O'Reilly Media, Incorporated. ISBN: 9781449311605. Available from: <<https://books.google.at/books?id=qcsoLHusAFsC>>.
- Boyed, Ryan (2012). *Getting Started with OAuth 2.0 - Programming Clients for Secure Web API Authorization and Authentication*. O'Reilly Media.
- Dasgupta, Dipankar, Roy Arunava, and Nag Abhijit (2017). *Advances in User Authentication*. Springer International Publishing AG. ISBN: 978-3-319-58808-7.
- LeBlanc, Jonathan (2011). *Programming Social Applications*. Ed. by Mary Tresler. Sebastopol, CA 95472: O'Reilly Media, Inc. ISBN: 978-1-449-39491-2.
- Prasad, Prakhar (Oct. 2016). *Mastering Modern Web Penetration Testing*. Ed. by Julian Ursell. Ed. by Rahul Nair. Ed. by Amrita Noronha. Ed. by Shweta H. Birwatkar. Birminigham: Packt Publishing Ltd. ISBN: 978-1-78528-458-8.
- Procházka, Michal, Daniel Kouřil, and Luděk Matyska (May 2010). *User centric authentication for web applications*. Chicago, IL, USA: IEEE. ISBN: 978-1-4244-6622-1.
- Røssvoll, Till Halbach (2013). *Reducing the User Burden of Identity Management: A Prototype Based Case Study for a Social-Media Payment Application*. Ed. by Lothar Fritsch, pp 364-370. ISBN: 978-1-61208-250-9.
- Todorov, Dobromir (2007). *Mechanics of User Identification and Authentication - Fundamentals of Identity Managment*. Auerbach Publications. ISBN: 978-1-4200-5219-0.

Refernce Articles

- Birrell, E. and F. B. Schneider (Sept. 2013). “Federated Identity Management Systems: A Privacy-Based Characterization”. In: *IEEE Security Privacy* 11.5, pp. 36–48.
- Cirani, Simone, Marco Picone, Pietro Gonizzi, Luca Veltri, and Gianluigi Ferrari (Feb. 2015). “IoT-OAS: An OAuth-Based Authorization Service Architecture for Secure Services in IoT Scenarios”. In: 15, pp. 1224–1234.
- Lynch, Lucy (Sept. 2011). “Inside the Identity Game”. In: *IEEE Internet Computing* 15.5, pp. 78–82.
- Steiner, Peter (2018). “On the Internet, nobody knows you’re a dog”. In: *The New Yorker* (). Available from: <<http://archives.newyorker.com/?iid=15713&startpage=page0000063#folio=CV1>> [July 2018].
- Tomkins, Benjamin (2009). “Dealing With Password Fatigue”. In: *Forbes*.

Refernce Online Sources

- Auth0 (n.d.). *Which OAuth 2.0 flow should I use?* Available from: <<https://auth0.com/docs/api-auth/which-oauth-flow-to-use>>.
- Corre, Kevin, Oliver Barais, Gerson Sunyé, Frey Vincent, and Jean-Michel Crom (2017). *Why can't users choose their identity provider*. Available from: <<https://petsymposium.org/2017/papers/issue3/paper18-2017-3-source.pdf>> [May 2018].
- Google (July 2018). *Google Developers. OAuth 2.0 Playground*. Available from: <<https://developers.google.com/oauthplayground/>> [July 2018].
- Google, Inc. (n.d.). *Developers Google. Google Identity Platform*. Google Inc. Available from: <<https://developers.google.com/identity/protocols/OpenIDConnect#discovery>>.
- Hudson, Harris (2015). *Targeted Security Risk Assessments Using NIST Guidelines*. Tripwire. Available from: <<https://www.tripwire.com/state-of-security/risk-based-security-for-executives/risk-management/targeted-security-risk-assessments-using-nist-guidelines/>>.
- JerichoSystems (2018). *Identity Silo*. Available from: <https://www.jerichosystems.com/technology/glossaryterms/identity_silo.html> [June 2018].
- KeyCloack (2018). *Sever Administration Guide - KEYCLOAK. Keycloak 1.9.8.Final*. published with GitBook. Keycloak <http://www.keycloak.org>. Available from: <https://www.keycloak.org/docs/1.9/server_admin_guide/topics/sso-protocols/saml-vs-oidc.html>.
- Lodderstedt, Torsten (2014). *OpenID Connect: Login mit OAuth, Teil 1 – Grundlagen*. heise.de. Available from: <<https://www.heise.de/developer/artikel/OpenID-Connect-Login-mit-OAuth-Teil-1-Grundlagen-2218446.html?seite=all>>.

- Neumann, Peter G. (2013). *Principled Assuredly Trustworthy Composable Architectures*. Available from: <<http://www.csl.sri.com/users/neumann/chats4.pdf>> [May 2018].
- NIST, U.S. Department of Commerce (2018). *Risk Management. Risk Management Framework (RMF) Overview*. Available from: <[https://csrc.nist.gov/Projects/Risk-Management/Risk-Management-Framework-\(RMF\)-Overview#1](https://csrc.nist.gov/Projects/Risk-Management/Risk-Management-Framework-(RMF)-Overview#1)>.
- Saltzer, Jerome H. and Michael D. Schroeder (1975). *The Protection of Information in Computer Systems*. Available from: <<http://www.cs.virginia.edu/~evans/cs551/saltzer/>> [June 2018].
- Schwartz, Mike (2016). *OAuth vs. SAML vs. OpenID Connect*. Available from: <<https://www.gluu.org/blog/oauth-vs-saml-vs-openid-connect/>>.
- Sevilleja, Chris (Jan. 21, 2015). *The ins and outs of Token Based Authentication*. Scotch.io. Available from: <<https://scotch.io/tutorials/the-ins-and-outs-of-token-based-authentication>> [July 2018].
- Siriwardena, Prabath (Apr. 2016). *JWT, JWS and JWE for Not So Dummies! (Part I)*. Available from: <<https://medium.facilelogin.com/jwt-jws-and-jwe-for-not-so-dummies-b63310d201a3>> [July 2018].
- Spencer, Travis (Apr. 2018). *API Security: Deep Dive into OAuth and OpenID Connect*. Nordic APIs. Available from: <<https://nordicapis.com/api-security-oauth-openid-connect-depth/>>.
- Tkalec, Tino (2015). *JSON Web Token Tutorial. An Example in Laravel and AngularJS*. Toptal. Available from: <<https://www.toptal.com/web/cookie-free-authentication-with-json-web-tokens-an-example-in-laravel-and-angularjs>> [July 2018].

Reference Manuals

- Brock, Allen and Dominick Baier (2018). *IdentityServer4Documentation. Release 1.0.0*. Available from: <<https://media.readthedocs.org/pdf/identityserver4/release/identityserver4.pdf>>.
- Brooks, Sean, Michael Garcia, Naomi Lefkovitz, Suzanne Lightman, and Ellen Nadeau (Jan. 2017). *An Introduction to Privacy Engineering and Risk Management in Federal Systems*. NISTR 8062. NIST (National Institute of Standards and Technology).
- Grassi, Paul A., James L. Fenton, Naomi B. Lefkovitz, Jamie M. Danker, Yee-Yin Choong, Kirsten K. Greene, and Mary F. Theafanos (July 2017). *Digital Identity Guidelines. Enrollment and Identity Proofing*. Special Publication 800-63A. NIST (National Institute of Standards and Technology).
- Grassi, Paul A., JamesL. Fenton, M. Newton Elain, Ray A. Perner, Andrew R. Regenscheid, William E. Burr, Justin P. Richer, Naomi B. Lefkovitz, Jamie M. Danker, Yee-Yin Choong, Kirsten K. Greene, and Mary F. Theafanos (Apr. 2014). *Recommended Security Controls for Federal Information Systems and Organizations*. Special Publication 800-53 Revision. NIST (National Institute of Standards and Technology).
- (July 2017). *Digital Identity Guidelines. Authentication and Lifecycle Management*. Special Publication 800-63B. NIST (National Institute of Standards and Technology).
- Grassi, Paul A., Michael E. Garcia, and Fenton James L. (July 2017). *Digital Identity Guidelines*. Special Publication 800-63-3. NIST (National Institute of Standards and Technology).
- Grassi, Paul A., Justin P. Richer, Sarah K. Squire, James L. Fenton, Ellen M. Nadeau, Naomi B. Lefkovitz, Jamie M. Danker, Yee-Yin Choong, Greene K. Kirsten, and Mary F. Theafonos (July 2017). *Digital Identity Guidelines. Federation and Asser-*

- tion. Special Publication 800-63C. NIST (National Institute of Standards and Technology).
- Hardt, D. Ed. (n.d.). *The OAuth 2.0 Authorization Framework*. Internet Engineering Task Force (IETF). Available from: <<https://tools.ietf.org/html/rfc6749>>.
- Jones, M. (May 2015). *JSON Web Algorithms (JWA)*. Internet Engineering Task Force (IETF). Available from: <<https://tools.ietf.org/html/rfc7518>> [July 2018].
- Jones, M., J. Bradley, and Sakimura N. (May 2015). *JSON Web Signatures (JWS)*. Internet Engineering Task Force (IETF). Available from: <<https://www.rfc-editor.org/rfc/pdf/rfc7515.txt.pdf>> [July 2018].
- Jones, M., J. Bradley, and N. Sakimura (2015). *JSON Web Token (JWT)*. Internet Engineering Task Force (IETF). Available from: <<https://tools.ietf.org/html/rfc7519>> [July 2018].
- Jones, M. and J. Hildebrand (May 2015). *JSON Web Encryption (JWE)*. Internet Engineering Task Force (IETF). Available from: <<https://www.rfc-editor.org/rfc/pdf/rfc7516.txt.pdf>> [July 2018].
- NIST, U.S. Department of Commerce (Sept. 2012). *Guide for Conducting Risk Assessment. Information Security*. Special Publication 800-30. NIST (National Institute of Standards and Technology).
- Sakimura, N., J. Bradley, M. Jones, and E. Jay (Nov. 8, 2014). *OpenID Connect Core 1.0 incorporating errata set 1*. Available from: <http://openid.net/specs/openid-connect-discovery-1_0.html> [Aug. 8, 2018].
- Sakimura, N., J. Bradley, M. Jones, B. de Medeiros, and C. Mortimore (Nov. 8, 2014). *OpenID Connect Core 1.0 incorporating errata set 1*. Available from: <http://openid.net/specs/openid-connect-core-1_0.html#toc> [June 22, 2018].
- Sakimura, N., J. Bradley, M. Jones, B. de Medeiros, and C. Mortimore (2014). *OpenID Connect Core. 1.0 incorporating errata set 1*. Available from: <http://openid.net/specs/openid-connect-core-1_0.html#toc>.

Refernce Misc

Bolton, Joshua b. (2003). *E-Authentication Guidance for Federal Agencies*. Washington D.C.: OMB (Office of Management and Budget). Available from: <<https://georgewbush-whitehouse.archives.gov/omb/memoranda/fy04/m04-04.pdf>>.

NIST, U.S. Department of Commerce (2004). *Federal Information Processing Standard Publication. Standards for Security Categorization of Federal Information and Information Systems*. NIST. Available from: <<https://csrc.nist.gov/csrc/media/publications/fips/199/final/documents/fips-pub-199-final.pdf>>.

Xu, Fangyuan (n.d.). *Security and Privacy Concern for Single Sign-on Protocols*. Project. Available from: <<http://www.cs.tufts.edu/comp/116/archive/fall2015/fxu.pdf>>.