Mid-Project Report
August 14th, 2012

Nils Dussart
0961540
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Proposal</td>
<td>3</td>
</tr>
<tr>
<td>Project title</td>
<td>3</td>
</tr>
<tr>
<td>Faculty Advisor</td>
<td>3</td>
</tr>
<tr>
<td>Project Scope and Individual Student Learning Goals</td>
<td>3</td>
</tr>
<tr>
<td>Proposed Product Activity Grid</td>
<td>4</td>
</tr>
<tr>
<td>Architecture diagrams</td>
<td>5</td>
</tr>
<tr>
<td>High level architecture</td>
<td>6</td>
</tr>
<tr>
<td>Authentication Sequence</td>
<td>7</td>
</tr>
<tr>
<td>Literature reviews/synopses</td>
<td>8</td>
</tr>
<tr>
<td>RFC959 - FILE TRANSFER PROTOCOL (FTP)</td>
<td>8</td>
</tr>
<tr>
<td>RFC2228 - FTP Security Extensions</td>
<td>8</td>
</tr>
<tr>
<td>OAuth Web Authorization Protocol</td>
<td>9</td>
</tr>
<tr>
<td>Inside the Identity Management Game</td>
<td>10</td>
</tr>
<tr>
<td>The National Strategy for Trusted Identities in Cyberspace</td>
<td>11</td>
</tr>
<tr>
<td>Taking Steps to Secure Web Services</td>
<td>11</td>
</tr>
<tr>
<td>The venn of identity: Options and issues in federated identity management</td>
<td>12</td>
</tr>
<tr>
<td>Systematically breaking and fixing OpenID security: Formal analysis, semi-automated empirical evaluation, and practical countermeasures</td>
<td>14</td>
</tr>
<tr>
<td>A Billion Keys, but Few Locks: the Crisis of Web Single Sign-on.</td>
<td>14</td>
</tr>
<tr>
<td>Bibliography of your literature reviews</td>
<td>16</td>
</tr>
<tr>
<td>Technical documentation</td>
<td>16</td>
</tr>
<tr>
<td>Appendices/Additional Documentation</td>
<td>17</td>
</tr>
</tbody>
</table>
PROJECT PROPOSAL

PROJECT TITLE

Federated Login with Client Server Applications.

FACULTY ADVISOR

Professor Munehiro Fukuda

PROJECT SCOPE AND INDIVIDUAL STUDENT LEARNING GOALS

My project goal is to analyze the possibility of using cloud identity providers as potential identity provider to achieve authentication with on premise client-server applications.

The main objectives are to:

• Identify how to make client-server applications leverage federated login
• Build a proof of concept using and FTP client & server developed in Java
• Identify the necessary modifications to the FTP RFC to make this possible
• Identify the pros and cons of such designs and their limitations if any
# Proposed Product Activity Grid

The general format of the Project Activity Grid is as follows:

<table>
<thead>
<tr>
<th>General Activity Category</th>
<th>Proposed Documentation</th>
<th>Activity Percent</th>
<th>Delivery Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Project Management/Communications (PMC)</td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Literature research</td>
<td></td>
<td></td>
<td>Every week</td>
</tr>
<tr>
<td>Status Reports and meeting with Advisor</td>
<td></td>
<td></td>
<td>Every week</td>
</tr>
<tr>
<td>2) Technical Approach (TA)</td>
<td></td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Code some prototypes using different protocols</td>
<td></td>
<td></td>
<td>September 14th</td>
</tr>
<tr>
<td>Continue reading on the protocols themselves</td>
<td></td>
<td></td>
<td>September 28th</td>
</tr>
<tr>
<td>Decide on one protocol and token</td>
<td></td>
<td></td>
<td>October 12th</td>
</tr>
<tr>
<td>Finalize design document and plan</td>
<td></td>
<td></td>
<td>October 26th</td>
</tr>
<tr>
<td>Implement authentication library for the FTP Client</td>
<td></td>
<td></td>
<td>November 9th</td>
</tr>
<tr>
<td>Implement token exchange between Client and Server</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implement authentication library for the FTP Server</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) Testing Strategy/Results (TSR)</td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Test solution with specific identity provider</td>
<td></td>
<td></td>
<td>November 9th</td>
</tr>
<tr>
<td>Test solution with multiple identity providers</td>
<td></td>
<td></td>
<td>November 23rd</td>
</tr>
<tr>
<td>4) Lessons Learned/Next Steps (LL)</td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>LL Report</td>
<td></td>
<td></td>
<td>November 30th</td>
</tr>
</tbody>
</table>
The following diagram describes the usual flow from a web service hosted in a cloud environment. In this case the token is basically relayed between the identity provider (STS) and the service itself. Since this flow happens through the client browser, the client itself has no operation or task to perform in the authentication process itself since everything happens in the backend.

This flow is important to my project since it provides a description if the usual architecture used between 2 different web services roles. Our objectives as part of this project will be to intercept the token and take action on it in the premise of the client rather than in the cloud.
The project consists into developing a proof of concept where a native client server application will authenticate against an identity provider located in the Cloud. The client will have to obtain an authentication token from that identity provider and then transfer it to the FTP server. The server once in possession of the token will have to verify the authenticity of the token and then make an authorization decision based on the token if it’s coming from a trusted identity provider.
The following diagram provides the authentication sequence that will be used by all the different players in this project. As mentioned in class the token validation still needs to be investigated and this will be done in the next couple of weeks and will result in this diagram to be updated.
Description of the document

This document is the official specification of the File Transfer Protocol (FTP). It contains all the original commands and expected states that are required to build a compliant FTP client and FTP server.

The document describes 2 specific commands of interest to my project, the USER command and the PASS command. These 2 commands provide a way to pass a user and password argument from the client to the server to provide authentication.

The user name command allows passing a string identifying the user to the server. The user identification is required by the server for access to its file system. This command will normally be the first command transmitted by the user after the control connections are made.

The password command (PASS) allows passing a string which specifies the user’s password. This command must be immediately preceded by the user name command.

Specific interest

Interestingly enough the document states the following:

Since password information is quite sensitive, it is desirable in general to "mask" it or suppress type out. It appears that the server has no foolproof way to achieve this. It is therefore the responsibility of the user process to hide the sensitive password information.

The original rfc959 didn’t provide a way to protect the password at all. The password was simply transferred in the clear. In the case of my project the token we will pass between the client and server will not need to be protected since it does not contain any secret. Also the token will be signed which protects against alteration.

There are no specific command in the RFC that allows to pass something else than a password. This means we will have to re-use the existing command but find a way to pass a flag to tell the server we will pass a token or create a new non-standard FTP command.

Description of the document

This document defines extensions to the FTP protocol describe in the previous page. These extensions provide strong authentication, integrity, and confidentiality on both the control and data channels with the introduction of new optional commands, replies, and file transfer encodings. The following new optional commands are introduced in this specification:

- AUTH (Authentication/Security Mechanism),
- ADAT (Authentication/Security Data),
- PROT (Data Channel Protection Level),
- PBSZ (Protection Buffer Size),
- CCC (Clear Command Channel),
- MIC (Integrity Protected Command),
- CONF (Confidentiality Protected Command), and
- ENC (Privacy Protected Command).

**Specific interest**

It is clear from the specification that these commands were not meant to provide a secure way to transfer the password nor provide new methods to authenticate. Instead the document defines authentication as the action of establishing a secure connection between the client and the server and authorization as the action of verifying the user name password to gain access to the file system.

This RFC provides clear security improvements but based on the provided definitions it proposes only to the authentication not the authorization. This means that no improvements are provided as part of this RFC to improve the authorization of users.

Like previously mentioned the transfer of the token in the clear is not an issue since it doesn’t contain any secret. This means that this RFC and the commands it provides are not required to implement an authorization model based on a cloud identity provider.

---

**OAUTH WEB AUTHORIZATION PROTOCOL**

**Description of the document**

The document provides a quick historical overview of the OAuth protocol which began as a community effort among numerous companies that provide Internet services. These organizations recognized the need to solve a specific type of identity management problem and developed the first version of a mechanism for doing so. The document describes the use of the protocol in the context of web services and specifically in the scenario where one service needs to interact with another on the behalf of the user.

The goal was to avoid having to ask user to turn their credentials to every service that needed to interact with a specific service. Turning over credentials gives that service full access to the user’s credential without any limitations. The OAuth protocol addresses this problem by providing an alternative mechanism through which users can authorize specific actions, and only those specific actions, without giving unrestricted or permanent access.

**Specific Interest**

Unfortunately the document seems to be mostly about the OAuth delegation model. No mentions are made about the authentication itself and how it can be achieved. The article does mention that the protocol is indeed used by Google, Yahoo, Facebook, and Twitter which makes it a protocol of choice for my project. I will still research Open ID since it seems it is still used by different major services including Google and Yahoo and understand what the clear difference where.
Interesting in the articles was the definition provided for client, authorization server and access token which will be useful for my research.

Authorization server — the service that verifies the resource owner’s credentials and performs the authorization checks. This is often the same as the resource server (as it is in both examples), and is always in its trust domain.

Access token — a piece of data the authorization server creates that lets the client request access from the resource server. This is the authorization credential the client will use in place of the resource owner’s own credentials.

Client — the service asking for authorization. Note that this is the OAuth transaction’s client, not the end user’s client program.

**INSIDE THE IDENTITY MANAGEMENT GAME**

**Description of the document**

This document describes the identity space as fragmented and incoherent. Multiple standards exists for managing authentication and authorization but so far no global solution exist from identity standpoint that scale all the necessary needs of the Internet.

As the web is evolving so are the need of new generation of web application such cloud-based services, social websites and mobile platforms. The document starts by providing an historical view of the identity space. Different standards exists but the main ones are:

OASIS supplies SAML and Web services (WS-*) as well as several other standards

IETF provides several relevant standards such as HTTP, TLS and OAuth.

The open source community is all also contributing to this space. The best example is OpenID.

There are also a multitude of other projects in both the private and public sector with specific requirements and needs.

The document then provides descriptions of several key protocol like OpenID and OAuth. The high level goal of both protocols is to let users create and asset an identity that is accepted by a wide variety of service. This would provide SSO since the user only needs to remember one password. The document define some key roles:

The end user (data subject)

The service provider (relying party, or RP)

The identity provider (IdP)

**Specific interest**

The document describes the limitation and current issues associated with OpenID and OAuth. Several implementation of OAuth exists which creates implementation issues. OpenID while still being worked seems to be less and less used. On the other hand, no mention is made in this
document of how OAuth is used in authentication scenario and how credentials are passed to the identity provider.

It seems that OAuth is clearly the protocol I should use for my project but unfortunately this document is once again focus specifically on the delegation scenario between 2 web services which is not aligned with the requirements of my project.

**THE NATIONAL STRATEGY FOR TRUSTED IDENTITIES IN CYBERSPACE**

**Description of the document**

The document is making the case that password need to be retired since password centric attacks are increasingly common. The author also explains that reliance on weak password technology has been a growing attack vector that threatens to erode confidence in the online world. Human factors are also contributing to his arguments since users have a tendency to use the less complex password possible. Finally passwords don’t scale well since the average Internet user in the USA needs to keep track on average of 30 passwords.

The article explains that alternative technologies are needed to replace passwords as the primary method of online authentication. An example is made of the USA DoD and their implementation of the Common Access Card (CAC) to securely manage identities and provide highly secure credentials.

Finally the author concludes by explaining the US government’s National Strategy for Trusted Identities in Cyberspace (NSTIC) focuses on working in partnership with the private sector to facilitate the adoption of online identification technologies that are secure and trusted.

**Specific interest**

This document confirms that Internet and government based identity providers provide large benefits from both a security standpoint and from a user experience standpoint. While the document doesn’t provide specific details for the technical implementation of my project, it confirms the assumption that Identity based providers are a technology that will continue to be used widely for years to come. Finding ways for legacy application to interact with those new solutions will allow leveraging their advantages.

**TAKING STEPS TO SECURE WEB SERVICES**

**Description of the document**

This document describes at the high level the standards and protocols available to organization to protect their services. The article covers securing web services but only from an enterprise point of view. It states that SOAP functions as transport mechanism for XML messages by letting programs running on one OS communicate with programs running on another, using HTTP and XML to exchange data. However, the World Wide Web original SOAP version provided no security.

To feel that gap, the Oasis organization created a new standard called SAML “SAML defines a vendor-neutral way to express security information in an XML format”. “It defines the schemas for the structure of documents that include information related to user identity and access or
authorization rights.” Basically SAML defines a specific token format that be used to prove identity, contain identity information and other security attributes. SAML can also indicate the authentication that must be used with a message, such as a password, Kerberos ticket, hardware token or X.509 certificate. Finally SAML allows single sign-on, since it possible for a web services to transfer the token to another site for authentication as long as the other site recognize the same authentication authority.

In addition to defining a token format, the Oasis organization created a set of new protocols to allow Web services to work with several different security models via SOAP extensions. These protocols allow applications to attach security data to the headers of a SOAP messages. These protocols, for example, provide ways of encrypting or signing messages including SAML tokens. This is critical in ensuring that the token was not altered. One specific protocol is worth mentioned for my project, WS-Trust provides a model for establishing both direct and brokered trust relationships for secure interaction. It basically defines how to acquire a token, validate it and exchange it.

The article concludes that one major obstacle for the adoption of those protocols and standards is the fact they rely on a PKI infrastructure or some enterprise directory services. This is not ideal consumer based scenarios.

**Specific interest**

This document was very interesting for my project since it provides the overview of how enterprise web services are enabling authentication using the specific Oasis protocols and standards. The goal of my project was mainly to interact with a consumer based identity provider such as Google but with time the difference between enterprise directory and public identity provider seems to get blurrier and blurrier. I will look into those standards especially SAML tokens and WS-Trust as possible solutions to my project.

---

**THE VENN OF IDENTITY: OPTIONS AND ISSUES IN FEDERATED IDENTITY MANAGEMENT**

**Description of the document**

This article provides an explanation of what a federation is and provides details about the architectural challenges and the associated security and privacy issues. The main goal of federation is to provide SSO to end users while using services or resources located elsewhere. While there are many benefits, there are significant security risks because the architectures share information across domains.

The author then describes several attacks that are plausible with federation. For example, he describes that user authentication is one of those weak links in the identity chain. Most site relies on username and an associated password pair which are well known to be weak and susceptible to phishing attacks.

The author provides and other example based on the fact that for web services provider’s federated identity is less expensive than implementing a high-quality authentication infrastructure because it offloads the authentication task to an identity provider. However, SSO can magnify the costs of a stolen password because it expands the scope of malicious activity. Most SSO protocols offer ways
to mitigate this risk. For example, they might limit to a minute or less the valid lifetime of the security token; some protocols also offer a single logout (SLO) feature that offers users near-simultaneous sign-out of all SSO-accessed Web sites.

The article then list several Architectural challenges:

- Identity provider discovery: services need a method to discovery an identity provider that they should trust and use. The configuration might be static or manual in which case the user must specify the location of the identity provider.
- Identifier schemes: this describes how to represent the same identity across different identity providers.

There are currently 3 main federation protocols, SAML, Open ID and Info Card. The author provides a quick summary of the protocols with some pros and cons.

**Specific interest**

This article was interesting from a thread modeling standpoint. The different threats it describe will influence my choice of protocol for my project. Man in the middle attacks and replay attacks are clearly possible and should be considered as part of my design.

The article doesn’t provide enough information to make design decision on Open ID versus SAML. While it lists some pros and cons the article does not go deep enough in the reasons why one group of users might use Open ID while the other might use SAML.

**IDENTITY MANAGEMENT AND TRUST SERVICES: FOUNDATIONS FOR CLOUD COMPUTING.**

**Description of the document**

This specific article describes how a campus or an educational organizational should deal with Identity Management & Trust Services. The author makes a case that IT organizations will move from providing IT services locally to becoming an integrator of IT services. Those services will be provided locally but also outside of the institution. Because of this, the article states that institutions should plan for shared services and must understand the essential role that identity management and trust services play in making integration possible.

The challenge is to integrate these different services in a coherent and effective manner. Issues such as authentication, access control, and the user experience when moving from one service to another need to be accounted for to be successful. The author then describes that institution have been missing a way to integrate these external service offerings and offer a comprehensive approach that should focus on the following three activities:

1. Developing an identity management system;
2. Creating a standard set of attributes for each person;
3. Enabling external access through a federation.

Completing those 3 specific activities will allow institutions to develop a common standard to authentication that will address their needs.
**Specific interest**

While this article didn't have a direct impact on my project, it does bring interesting background on the subject that I'm researching. The author makes a good case at defending the fact that each institution should use a common set of standards to enable a common identity management for their services. The article mentions briefly the Oasis set of standards and more specifically SAML which is on the protocols I am considering using.

---

**SYSTEMATICALLY BREAKING AND FIXING OPENID SECURITY: FORMAL ANALYSIS, SEMI-AUTOMATED EMPIRICAL EVALUATION, AND PRACTICAL COUNTERMEASURES.**

**Description of the document**

This document does an in depth analysis of the OpenID protocol and provides a threat modeling exercise. The analysis reveals that the protocol does not guarantee the authenticity and integrity of the authentication request. The results of the investigation suggest that many OpenID-enabled websites are vulnerable to a series of cross-site request forgery attacks. The author proposes as part of the article some simple and scalable mitigation technique for OpenID-enabled websites, and an alternative man-in-the-middle defense mechanism for deployments of OpenID without SSL.

**Specific interest**

This article provides a good overview of OpenID and its limitations and weaknesses. Since OpenID is one of the protocols I'm contemplating for my project, the article did provide a great of information that I've found useful especially in the first parts of the article.

---

**A BILLION KEYS, BUT FEW LOCKS: THE CRISIS OF WEB SINGLE SIGN-ON.**

**Description of the document**

This research document does a deep analysis into the different solutions that provides WEB SSO on the web and why there are not more adopted widely. The paper provides a technical description of several solution offered currently on the Web of interest to us is the mention of federated scenarios through SAML and the OpenID protocol.

The article then lists the incentives that are currently lacking to make this attractive to the different players in this field:

- Lack of incentives for Relying Parties
- Lack of driving forces from users.
- Lack of driving forces from Identity providers

The author provides several recommendations to address the lack of adoption of those technologies. One of them is to build a business incentive for relying parties to trust identity providers. The second is improve the user experience when using such technologies.

**Specific interest**
This document provided with very good insight at the current space and as to why some of the technologies I want to use as part of my project are not leveraged more. The recommendation to build identity support into the browser is a very good idea and is well aligned with my goal of analyzing the user experience when leveraging my proof of concept.
BIBLIOGRAPHY OF YOUR LITERATURE REVIEWS


TECHNICAL DOCUMENTATION


SAML 2.0 Session Token Profile Version 1.0 - http://docs.oasis-open.org/security/saml/Post2.0/saml-session-token/v1.0/saml-session-token-v1.0.html

Federated Login for Google Account Users - [https://developers.google.com/accounts/docs/OpenID](https://developers.google.com/accounts/docs/OpenID)

UX research on Desktop Apps using federated login and/or OAuth - [https://sites.google.com/site/oauthgoog/UXFedLogin/desktopapps](https://sites.google.com/site/oauthgoog/UXFedLogin/desktopapps)


SAML Single Sign-On (SSO) Service for Google Apps - [https://developers.google.com/google-apps/sso/saml_reference_implementation](https://developers.google.com/google-apps/sso/saml_reference_implementation)