Towards a Disaster Resilient Self-Sovereign Identity Research Plan

Kalin Kostadinov

April 25, 2021

Background of the Research

Humans, receiving services over the Internet, always need to use some form of identity. It is a representation of their true self, and is referred to as their "digital identity". It is required for trust to be established between users and their service providers.

Since the invention of the Internet, many different digital identity management systems have been deployed [1]. The problem with all of them is that they only function if their users' identities are kept on some kind of a centrally controlled server. Thus, owners of the identities are the users, but the identity management system itself possesses full control over those identities.

In recent years, identity management has become a major concern for governments. This has led to a substantial amount of research and regulations in the field [2]. The goal of this research has been the formal description of a novel identity management system that does not take control over an identity from its rightful owner [3]. It promises to become the new benchmark in identity management systems by providing the properties of "Self Sovereign Identity" (SSI) [4]. It is meant to give control back to users over their data. It accomplishes this by allowing every identity holder to store and manage their data, using their resources or resources under their jurisdiction.

There are already several implementations that cover part of SSI's properties [5]. One of them, called IPv8 [6], has been developed by the Delft Blockchain Lab and is arguably the most sophisticated SSI management system. However, the issue with IPv8 is that it does not offer long-term data resilience because it does not use a global blockchain that commits transactions by all users into one structure, thus not offering a mechanism for recovery from identity loss. Instead, every single user has its blockchain, called TrustChain [7], for managing their identity. The idea behind this design decision is that users have generally more control over their own identity if they are the only ones physically possessing their data blocks. The problem is however that since mobile applications are the most effective way of hosting an identity management system like IPv8, it is not clear how the identities of users are supposed to be recovered when access to them is lost.

Research Question

Self-sovereign identity management systems have matured over the past couple of years. The biggest obstacle which prevents them from going mainstream, however, is the problem of adoption. All these SSI [5] applications need to be designed with any type of user in mind. Thus, the complexity of those applications should be as low as possible, for them to be understood and used by every single person in the world or at least the countries where regulations require the use of SSI management systems.

IPv8 alongside other SSI implementations have one of two problems in this regard. They either rely on public blockchains that contain the entire transaction history of all users, which prevents such systems to be used in several situations (e.g. offline or when transaction times need to be as low as possible) because of their need for global consensus and thus hinders availability. Or they do not employ any mechanism for disaster resilience, like IPv8. There is a need for a solution to the latter problem because it will allow for systems that do not use any type of central authority and global consensus to outweigh the other SSI management systems when it comes to service availability. The problem of availability as a sub-problem of adoption is a research area that is worthwhile exploring.

It is clear that to make a system resilient to data loss, there needs to be some kind of a protocol that adds redundancy. The data needs to be physically stored in at least two separate locations. This, however, adds some complexity and overhead to the system. Also, there needs to be a caching mechanism that allows transactions to be temporarily stored, when in an offline setting, before they can be synchronized with other deployments of the system. And if there are multiple nodes having control over the same identity, there needs to be a mechanism that can revoke access to the identity by its owner.

The following research question will be laid at the center of this research: How to make fully distributed Self-Sovereign Identity management systems disaster resilient? To answer this question the following sub-questions have to be answered as well:

- How to add redundancy to the system without any overhead or with as little overhead as possible?
- How can access to the identity management system be revoked?
- How can a cache be implemented for transactions in an offline setting?

Depending on the progress, more sub-questions might be defined as well.

Method

Two ideas emerge from the above observations. Either keep the SSI management system on a master remote server or reproduce the blockchain from other users, which know something about the identity being reproduced. The first one will add some overhead which is not desirable for the adoption of this technology. The second one increases complexity since one has to keep in mind that not all identities, one has communicated with, still exist, will be honest about previous transactions, or are online during the process of rebuilding the chain. With either solution, a protocol will be developed with the same purpose but different implementations. This protocol will in effect allow for the creation of "terminals". These terminals will be control units for someone's digital identity and every device should be able to become one. With them, two new concerns emerge: access revocation for terminals that a user has lost control over, and caching for keeping the SSI management system operational in case of offline transactions.

My research will begin with a literature search for systems in other fields which have already solved the problem of availability for self-contained applications. If I can discover more ideas, I will use them during the later parts of the research. Furthermore, two different solutions will be designed and evaluated. The one which adds less complexity and overhead to the usage of IPv8 will be picked and later implemented and integrated with IPv8. The research will conclude with a paper that will explain the problem, I am trying to solve, in more detail and describe the solution of choice, how it was implemented and whether it was able to solve the problem of disaster resilience.

Since I am going to try and use IPv8 as a platform for improvements, I have as a task to explore its implementation in Kotlin for the super app [8], the Delft Blockchain Lab is currently working on. My goal is to assess the application and find a suitable approach for integrating my implementation.

I was previously involved in the development of another application that also uses IPv8 for storing COVID-19 immunity certificates [9]. The main objective of that project was to create a user-friendly GUI. I can use that project as an experimental environment for my findings since it already uses REST API to communicate with the IPv8 deployment and can be easily turned into an application that controls a remote deployment of someone's identity. It also relies on React Native [10] for its GUI. That means IPv8 can get an implementation for iOS which the lab is lacking.

Depending on how the protocol is going to be designed, it might be useful to keep the core back-end that is going to run on the mobile device in Python as is the implementation of IPv8 [11] itself. Since there is still no publicly available tool that efficiently builds Python applications for Android, iOS, and other mobile or embedded operating systems, I might make use of my tool. It is called Porthon and its goal is to make Python applications portable and resource-efficient. Currently, there is a working version of the IPv8 implementation in Python for Android and one for iOS is underway.

A further development, that goes beyond the goals of this research project, will be the creation of emergency terminals that will be available at border control for instance. They will allow someone to still access their identity manager with restricted controls in the case that their other "terminals" are not available. Those emergency terminals should only allow for verification of attestations.

Planning of the Research Project

For week 1 of the research project, I have to submit a project plan. To achieve this, I aim to read through the following research from Uwe Der, Stefan Jähnichen, and Jan Sürmeli [12], Alexander Mühle, Andreas Grüner, Tatiana Gayvoronskaya, and Christoph Meine [13], Quinten Stokkink, Dick Epema, and Johan Pouwelse [6], Q. Stokkink and J. Pouwelse [14], A. Tobin and D. Reed [15], D.S. Baars [16]. I will also look into other relevant research, following the citations from the papers mentioned above and finding potentially others that focus more on the issues of availability. Given my findings, I will write my research plan. Deadlines: April 19th - Planning Week 1, April 20th - Information Literacy, April 25th - Research Plan.

During week 2, I plan on searching for best practices in solving availability issues in the field of mobile applications. Any such ideas will be used during week 3. In week 2, I have to also present my research plan to my supervisors and peers. Deadlines: May 2nd - Research Plan Presentation.

Week 3 will be spent on deciding which of the two proposed solutions is better feasible, given the time and technology constraints. Deadlines: May 6th - Academic Communication Skills Assignment 1: First 300 Words.

In week 4, I will work on formally defining a solution for the problem. This week is supposed to produce some parts of the final research paper as well. Deadlines: May 16th - Academic Communication Skills Assignment 2: Midterm Poster for Feedback.

Week 5 will set the beginning of the implementation process of the protocol. During it, I will work on adding a recovery mechanism for IPv8. Since this is the halfway mark, I will have a midterm presentation as well. Deadlines: May 19th - Midterm Poster.

During week 6, access revocation will be at the center of most of my efforts. Deadlines: May 27th - Academic Communication Skills Assignment 3: Improve First 300 Words, and Add Section of 300 Words.

For week 7, I plan on working on the caching algorithm that will allow for offline transactions to still be possible. There are no deadlines for week 7.

In week 8, integration of the protocol with IPv8 will start. Deadlines: June 7th - Paper Draft V1, June 10th - Peer Reviews.

Week 9 will consist of the integration and testing of the protocol. Deadlines: June 16th - Paper Draft V2.

During week 10, most of the work will be devoted to finalizing the poster and the research paper. Deadlines: June 27th - Academic Communication Skills Assignment 4: Final Poster for Feedback, June 27th - Final Paper.

Week 11 is devoted to preparations for the final poster presentation. Deadlines: June 29th - Final Poster.

Meetings with my supervisors are scheduled every Tuesday at 10:00. During them, details and progress of the project will be discussed.

References

[1] Gergely Alpár, Jaap-Henk Hoepman, and Johanneke Siljee. The identity crisis. security, privacy and usability issues in identity management. 2011.

[2] European Commission. Regulation (eu) 2016/679 of the european parliament and of the council. 2016.

[3] Md Sadek Ferdous, Farida Chowdhury, and Madini O. Alassafi. In search of self-sovereign identity leveraging blockchain technology. *IEEE Access*, 7:103059–103079, 2019.

[4] Christopher Allen. The path to self-sovereign identity. 2016.

[5] Dirk van Bokkem, Rico Hageman, Gijs Koning, Tat Luat Nguyen, and Naqib Zarin. Self-sovereign identity solutions: The necessity of blockchain technology. 04 2019.

[6] Quinten Stokkink, Dick Epema, and Johan Pouwelse. A truly self-sovereign identity system. 2020.

[7] Pim Otte, Martijn de Vos, and Johan Pouwelse. Trustchain: A sybil-resistant scalable blockchain. *Future Generation Computer Systems: the international journal of grid computing: theory, methods and applications*, 107:770–780, June 2020.

[8] Tribler. Trustchain super app.

[9] Tribler. Immune: Building a critical infrastructure for the nation-wide identification of recovered covid-19 patients.

[10] Facebook. React native.

[11] Tribler. Ipv8.

[12] Uwe Der, Stefan Jähnichen, and Jan Sürmeli. Self-sovereign identity - opportunities and challenges for the digital revolution. *CoRR*, abs/1712.01767, 2017.

[13] Alexander Mühle, Andreas Grüner, Tatiana Gayvoronskaya, and Christoph Meinel. A survey on essential components of a self-sovereign identity. *Computer Science Review*, 30:80–86, 2018.

[14] Quinten Stokkink and Johan Pouwelse. Deployment of a blockchain-based self-sovereign identity. *CoRR*, abs/1806.01926, 2018.

[15] A. Tobin and D. Reed. The inevitable rise of self-sovereign identity. 2016.

[16] D.S. Baars. Towards self-sovereign identity using blockchain technology. October 2016.