Policy for a CERN Solid server

Summary

Solid (**So**cial **Li**nked **D**ata) is Sir Tim Berners-Lee's project including standards, missing from the original Web specifications, that give back to the data owners the power to decide *where* their data are stored and *who* has access to them.
This is a proposal for installing and running a Community Solid server (CSS) at CERN before the end of 2021. In this way the CERN Single Sign-On (SSO) integration is ensured, as well as the location of personal data. Hence all CERN Office of Data Protection (ODP) requirements are covered.
All the details on *Reasons, Definitions,* **and** *Open Questions* below.

Reasons

It is in the interest of CERN to run its *own instance of an existing open source implementation* of a Solid server. Reasons:

1. It is important for the image of CERN and its role in the future of the web and its members' personal data sovereignty.

    - A CERN-Solid affiliation, technically materialising via a solid.cern.ch server, concretely shows the Organisation's support and promotion of the world of open source tools. This is in-line with the HEP community and CERN tradition, having always found open source tools valuable and having contributed to them.

2. Several CERN application owners and service managers will be relieved to know that personal data of their users are not stored in files/databases of their own application but on the users' Solid pods.

    - It will be easy for new utility apps of all sorts to flourish when they don't have to build a back-end for each one, the Solid server serving as a universal back-end.

3. Existing Solid servers outside CERN can store our data in resources we can't control and might not comply with CERN Personal Data policy rules.

Which Solid Server flavour

Solid Poc period
During the CERN-Solid Proof of Concept (PoC) Project [1], we recommended the [solidcommunity.net (https://solidcommunity.net)](https://solidcommunity.net) server, a Node Solid Server (NSS) implementation, to obtain a few pods for the CERN volunteer testers.

As of the autumn 2021

The [solid.cern.ch (http://solid.cern.ch)](http://solid.cern.ch) reserved name (now a re-direct to our events' and documents' index) will host the CERN instance of an **open source** Solid server. The question of Solid server flavour to adopt for the future CERN-hosted Solid server shall inevitably be

constantly re-evaluated because the products evolee. Technical alternatives are listed below. Project [2] will take care of the deployment of the adopted solution.

1. [solidcommunity.net (https://solidcommunity.net)](https://solidcommunity.net) and several other instances (e.g. [solidweb.org (https://solidweb.org)](https://solidweb.org) hosted in Germany) are Node.js servers. This is the longest-existing Solid server implementation. It also gives good results when submitted to the test-suite. Nevertheless, it dates since 2016 and lacks funding and support for the future, while the Solid specifications evolved a lot since its implementation, which **remains experimental**. It is unlikely that effort will be invested by the community to enhance its functionality for the users. Therefore, users could stay with NSS (solidcommunity.net), while knowing that its UI/UX functionality will not improve and the service will remain experimental.

2. Depending on its maturity, the Community Solid Server (CSS) (version 1.0 due in the summer 2021) will be used. The UI is decoupled from the server. [These recipes (https://github.com/solid/community-server-recipes)](https://github.com/solid/community-server-recipes) offer 2 CSS server configurations with different UIs.
   Reasons: It is a new Server written from scratch in TypeScript, it has funding (donations by Inrupt), it is open source, the project is based in Europe (Ghent University). See the *Technical notes* section for details on issues to be addressed.

3. The Enterprise Solid Server (ESS) being developed by the company [Inrupt (http://inrupt.com)](http://inrupt.com) is closed code, commercial, tailored to customer needs, the data store is USA-based and fails several tests of the original test-suite. This is why this option is will not be adopted.

Definitions

- **Solid** is a standard, an ecosystem, a movement and a community initiated by Sir Tim Berners-Lee, that allows people to take control over their own data, i.e. where they are stored and who has access to them. It combines existing W3C standards and is built on top of the existing Web.
- **A Solid pod** is a decentralised data store for one's personal data. Pods are like secure personal Web servers for all kinds of data.
- **A Solid server** is a Web server that stores users' pods, with support for access control. [3]

Details

Why we can use existing Solid server solutions for the PoC:

The PoC, by definition, is not an operational, long-living, sensitive-data service. It is internal, short-lived, with a limited number of participants, who are technical and security-aware.

We recommend [solidcommunity.net (https://solidcommunity.net)](https://solidcommunity.net) [5], [6] as a pod provider **for the PoC**. Still CERN users can choose any provider from [4]. All Solid server flavours and instances are supposed to be compatible.

Indeed, each of the following people are hosted on different instances of the same Solid server implementation, NSS:
[Sir TimBL (https://timbl.inrupt.net/profile/card#me)](https://timbl.inrupt.net/profile/card#me)
[Jan (https://janschill.net/profile/card#me)](https://janschill.net/profile/card#me)
[Maria (https://dimou.solidcommunity.net/profile/card#me)](https://dimou.solidcommunity.net/profile/card#me)

Why we need a CERN Solid server on-site after the PoC:

The CERN Office of Data Privacy (ODP) [7] requires processing of personal data to take place in CERN's member states for the Organisation to have full control over the data it is responsible for. Data processing outside the CERN member states [8] *is* possible under restrictions and a quite heavy overhead. For example the data must be encrypted in transit and in rest and the encryption key is held exclusively by CERN. There is no encryption today for Solid pod data. It is recommended, but not mandatory, by the Solid specification, that pod data transfer *should* use TLS connections through the https URI scheme in order to secure the communication between clients and servers [9].

CERN services using Cloud solutions *do* exist today, e.g. the CERN Zoom contract. The ODP and the Cloud License Office [10] ensure relevant contract clauses by which all CERN data is stored in a EU country, the provider will exclusively use this data for the provision of the service and that they will abide to the principles of OC11 [11] and GDPR [12].

Existing Solid servers [4] may store pod data outside Europe, via Cloud solutions, the physical storage of which we can't always control.

Given the above rules, this is a risk we shouldn't take.

Technical notes for a future solid.cern.ch server

Technology questions to decide before turning solid.cern.ch into a CERN instance of a Solid server:

1. Is the Solid server a standalone server
2. What is the software stack it needs
3. What storage do we need
4. Are the data stored in a database or in files?
5. Is it wise to have a single server or several servers for redundancy?
6. An official Docker image for a Solid server doesn't exist.

Possibilities depending on the answers to the above:

1. Standalone server fully managed by a CERN-Solid sys. admin: In that case we would need to request a virtual machine with the name solid.cern.ch on OpenStack and install the Solid software on it.
2. Use one of CERN Webserver VMs and install Solid on top of it.
3. Docker/Container version: in that case it could possibly be installed on OpenShift (PaaS for Web Application).

Option 3 above, i.e. using a VM/OpenShift and some pre-existing docker images [13], seems promising.

What makes CSS attractive

CSS development is funded, in form of donations by Inrupt. Digita (https://www.digita.ai/), a belgian company is an active co-developer of CSS, interested receiving requirements from a CERN, because of the lab's large and sophisticated user environment. CSS at this moment doesn't exist as a URI, from which to get a pod and try the server's functionality. Still, as the software approaches v.1.0 **in summer 2021**, this possibility should be there soon. **Advantages for CSS, as Solid server implementation and Digita, as a young company**: the image of CERN linked with the company's work, CERN users with web applications' expertise provide functionality input.

Advantages for CERN, as a lab, which lacks resources and does mostly physics: an open source, european pod development, with maintenance and support. We can submit requirements and have an eager team to enhance and fix.
Open questions:

1. **The current UI** of NSS-flavour servers is very unsatisfactory. Several *issues* were submitted by us in github [13]. However, this approach doesn't scale. The UI has to be radically re-designed.
2. **Quality of service**. CERN users are used to 24/7 reliable services. CSS is being developed partially by University volunteers. If the service is not of operational quality, the CERN users will *not* embrace Solid, as we wish and as they should.
3. **Required in-house development**. CSS comes with no ID provider and no storage included. This will require development work for CERN, to integrate the CERN SSO and local storage. Project [2] in the References' list addresses the resources required for these important aspects.
4. **Few Solid applications** exit for now. This makes it hard for organisations with limited resources to invest in Solid. Nevertheless, some, like liqid.chat are cool and promising and several others are presented every month [14].

Conclusion

Solid is today like the web in 1992... Nevertheless, it is here to stay. It is the future. It is the *desirable* future for rescuing personal data sovereignty. The day will come when Solid will be everywhere. Then CERN will be happy to have joined earli..ish.

References

1. The CERN-Solid Proof of Concept project (https://it-student-projects.web.cern.ch/projects/cern-solid-code-investigation)
2. A project including a CERN-own Solid server (https://it-student-projects.web.cern.ch/projects/cern-solid-and-slides-app)
3. A good presentation on Solid by Ruben Verborgh, CSS project leader (https://rubenverborgh.github.io/ECA-2021/#)
4. Solid Pod options (https://solidproject.org/users/get-a-pod)
5. The NSS instance we recommend **during the CERN-Solid Proof of Concept** (https://solidcommunity.net)
6. The storage solidcommunity.net uses (https://www.digitalocean.com/trust/)
7. CERN Office of Data Privacy (https://privacy.web.cern.ch/office-data-privacy-odp)
8. CERN Member states (https://home.cern/about/who-we-are/our-governance/member-states)
9. Pod data transfer encryption rules - not mandatory (https://solidproject.org/TR/protocol#http-server)
10. CERN Cloud License Office (https://cloud-licence-office.web.cern.ch/)
11. CERN Operational Circular 11 (OC11) (https://cds.cern.ch/record/2651311/files/Circ_Op_Angl_No11_Rev0.pdf)
12. General Data Protection Regulation (https://gdpr-info.eu/)
13. A project allowing to run NSS in Docker (https://github.com/angelo-v/docker-solid-server). NSS has many different dependencies (JavaScript. TLS certificates and more). To avoid installing all necessary dependencies on the physical machine, with Docker you can write a configuration and scripts that allow a virtualization and a one-command-setup.

14. [Issues with the NSS UI (https://github.com/solid/solid-ui/issues)](https://github.com/solid/solid-ui/issues)
15. [Solid World monthly webinars (https://solidproject.org/events)](https://solidproject.org/events)
16. [Links from Solid notes/presentations/papers (http://solid.cern.ch)](http://solid.cern.ch)

Credits

This note was written thanks to input from:

Sir Tim Berners-Lee, Jan Schill, Michiel de Jong, Thomas Baron, Tim Smith, Gabrielle Thiede, Jakub Moscicki, Andreas Wagner, Ruben Verborgh.

Maria Dimou / CERN-Solid Collaboration manager
Last update: 2021/05/10
Created: 2021/03/15