TERENA Trusted Cloud Drive for Academic Research

by Peter Szegedi, TERENA

20 March, 2013

The Trusted Cloud Drive (TCD) project [1] aims at piloting an experimental, high-performance, trusted, cloud storage solution for the Research and Education (R&E) community gathered under the Trans-European Research and Education Networking Association (TERENA) [2]. It builds on an open source cloud storage brokering platform [3] that provides federated user access, strong data encryption, supports various storage back-ends, and most importantly ensures the separation of the storage data from the metadata (such as file attributes, encryption keys, etc.) that are kept in a trusted location. TCD can also be considered as a storage middleware that maintains trust and privacy within the user domain and acts as a secure relay towards the private and/or public providers' domains connected.

National Research and Education Networks (NRENs) around the globe - such as the Intrenet2 in the US or SURFnet in the Netherlands - connect universities, university colleges and campuses with high-capacity links, peer with commercial networks at major exchanges, and provide advanced value-added services to the R&E community. They are membership organizations, governed by universities, subsidized by national governments, and work in a nonprofit manner. Undoubtedly, massive data storage is vital for academic research. Individual researchers and students on campus more and more use commercial cloud storage offerings (e.g., Google Drive, iCloud, Dropbox) available on the market. However, these public services are not primarily designed for the needs of sensitive research data sets. Therefore, universities and research institutes are seeking for partnership with private storage solution integrators and application developers (e.g., PowerFolder, SpiderOak, OwnCloud) to build and operate their own storage infrastructure on campus that needs not only capital investment but operational knowledge and experience too. These private storage clusters can then provide the desired performance and data privacy but, due to the lack of standards and sometimes proprietary vendor solutions, cannot always interface with each other or with the public services.

NRENs are in a good position to deliver high-performance data storage infrastructure as a service specially tailored to R&E community over their advanced networks at national scale. Moreover, thanks to the European and global NREN collaboration, they can also aggregate demands and facilitate community provided storage to be shared across TERENA members. Trust is the main asset of NRENs, as they are governed by the universities that are also the major clients of the NRENs. The Trusted Cloud Drive service pilot - the initiative of TERENA – builds on this trust relationship and provides the necessary software tool and know-how at NRENs' hands. The open source, cloud storage brokering platform incorporated by the TCD pilot can be installed at university locations or hosted by NRENs to aggregate demands and broker storage resources. It can also act as a storage middleware layer that separates the underlying trust domain from the storage back-end providers' domain so that maintains the data privacy of the users.

The user is able to authenticate to TCD with his/her federated account provided by the home institution hence rich set of identity attributes are available to determine the actual service offering via the platform. On the front-end a native Web Application or standard WebDAV

access can be used for typical disk operations. Acting as a middleware, it is also possible to integrate the TCD platform with other, feature-rich storage applications provided by commercials or the community. Beyond the scope of the pilot, TERENA has been discussing e.g., with PowerFolder and the OwnCloud Community about potential integration scenarios. Trusted Cloud Drive does the encryption and the separation of the metadata from the storage data. The encryption keys and the sensitive metadata are kept in the local metadata store. The encrypted storage data blob can then be exported to the public cloud using various storage back-end APIs, including Amazon S3, OpenStack Swift, Pithos+ (the Greek NREN's cloud) and soon other APIs supported by Jclouds.

The TERENA TCD pilot started in May 2012 and the final results with the list of potential use cases, service delivery scenarios and legal advices are expected to be published in April 2013. During the open pilot period (last 9 months) the software platform has been installed and tested at NRENs in Greece, Czech Republic, Croatia, Poland, Belgium, Portugal, Spain and Brazil. All together 19 NRENs, 8 Universities and 3 Research Labs have expressed their interest in experimenting with TCD in one way or another. TERENA is eager to maintain this community around the open source code in order to ensure the long term sustainability of the Trusted Cloud Drive. Commercial companies are also interested in exploring the potential service integration scenarios that goes beyond the scope of the pilot.

[1] http://terena.org/clouddrive
[2] http://www.terena.org/
[3] https://github.com/TERENA/CloudDrive

TERENA TF-Storage Task Force participants discussing about the Trusted Cloud Drive pilot in March, 2013 in Berlin, Germany