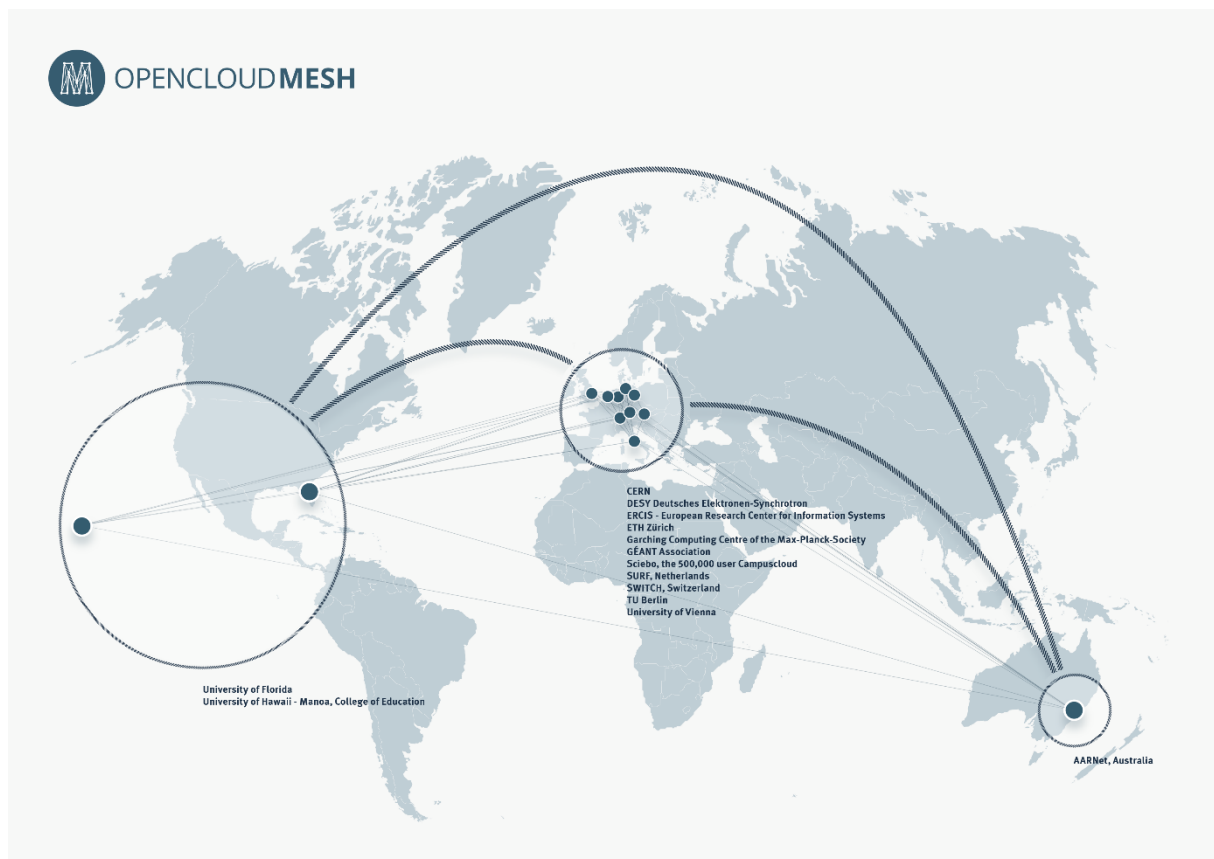# OpenCloudMesh

An open standard to interconnect private clouds globally.

Version: 23th July 2015 v 0.002

Christian Schmitz // Frank Karlitschek

Index

1. Preamble

OpenCloudMesh, even as v0.002, as name, project and concept is the result of lengthy projects, discussions and plain old work. The issues at hand described in this document are real, and a real path to a next generation of what is now considered "cloud" and may well be considered "mesh" at some point in the future. The challenges are numerous and require a broad, global approach to work out a tangible solution that will be robust enough to be valid for the next decade or beyond. With that said, I suppose that the initial team of the willing can be considered the best bunch of people to undertake such a project.

To quote Barack Obama: If you run you stand a chance of losing, but if you don't run you've already lost". So let's start running.

1. Preamble

2. Cloud or Local Silo or !?

The core of any IT Infrastructure is the aim to increase productivity and efficiency of the end user by providing them with a specific service, function or information requirement. This mandates that a vast majority of IT related activities have a human interface at some aspect of their operation. Advances in actual usability of aforementioned human interface is one of the primary shaping forces of IT landscapes with landmark events being the evolution from the command line to the graphical user interface to touch interfacing.

A less "visual revolution" in terms of end user usability is the current trend to ubiquitous availability of unstructured and structured data. A very typical early example of this trend is email, practically no user really knows where an email is located apart. This concept has seen rapid advances since smartphones entered the stage since the 2007 launch of the iPhone and various successor technologies. The following consumerization of ubiquitous storage and compute has seen almost unprecedented pickup with "Cloud-Services" that allow end-users a centrally provided global service for various services such as (unstructured) data (i.e. Dropbox), messaging (i.e. WhatsApp) or even media (i.e. Spotify).

While aforementioned services all allow rapid end user adoption, they are generally aimed at end users and do not meet the legal or integrative requirements of professional end users. As the end users are inherently unwilling to go back to "classic" usability, IT is faced with either offering their legacy services and information in a new way or deal with a situation where end users, often illegally, outsource data to third party services on their own initiative.
The solution would be hosting your own Cloud software, delivering modern services to end users directly from your existing infrastructure and utilizing existing data and processes. This very popular private cloud approach solves privacy, legal, integration and performance issues of public clouds but falls short of ubiquity in globalized collaborations. To avoid private clouds being "fancier ftp servers" and "Yet Another Data Silo", a vendor agnostic way of interoperating between such private clouds would solve the YADS problem in an elegant way: allowing the best of both private and public cloud.

3. OpenCloudMesh, solving YADS globally

OpenCloudMesh has a very simple goal: to be an open and vendor agnostic standard for private cloud interoperability.
The problems, concepts and solution approaches to solving this are absolutely cutting edge as of 2015, hence they offer both practical and research challenges. Science and research in its open and peer reviewed nature has become a truly globalized project. To address the YADS problem, a working group under the umbrella of the GÉANT Association is being created with the goal of ensuring neutrality and a clear context for this project.

The problem OpenCloudMesh is addressing is not a solution only applicable to the science and research community. Just as the world wide web was designed as scientific information browsing tool, it quickly turned into the de facto way to access information globally. We see data silo creation in every datacenter and even rack. Every entity, be it governmental, ngo or private venture, has an ever growing set of data that is used and utilized by its own internal users and at an equally growing pace with external collaborators.

Typcial examples would be engineering teams in globalized enterprises creating automotive or aerospace products with components from dozens of suppliers, teams from various nations and organizations working hand in hand to provide a coordinated relief in case of catastrophe or even just the very typical exchange between different branches of the government within a state or across state borders.

4. Open organization and collaboration

The project is run under the neutral umbrella of the GÉANT Association, a European NGO representing 41 countries and almost 50m users. Collaboration and coordination will be done under supervision of a GÉANT Association representative, currently Peter Szegedi.

Specific task and project leads will be elected at the first conference to kick off the project, due to happen at the ownCloud Workshop at TU-Berlin, Friday 28th of August 2015, invites will go out ASAP.

In addition to that, it is important to point out that there will also be active involvement welcome explicitly by non ownCloud related private cloud software vendors and operating facilities. Initial invites to such vendors will be made until the public release on August 3rd, 2015

Official communication is handled via dedicated mailing lists covering specific tracks:

- ocm-all@terena.org as generic mailing list covering
- ocm-admin@terena.org covering administrative and conceptual topics
- ocm-tech@terena.org covering technical and engineering topics

5. Plentiful challenges

The scope of the problem suggests that despite having an existing draft of an existing federated cloud sharing API, plenty of interesting and potentially challenging aspects are already at hand that could be interesting for the overall OCM standard. A preliminary, not complete list of such topics could be:

- storage brokering between nodes
- automatic tagging of data streams (i.e. automatedingest, localonly, secure, etcpp.) to allow controlled and granular mesh buildup
- integration into other federated services (i.e. EUDAT, Hydra, Eduroam)
- integration and adaption of workflow and visualization (i.e. HPC Resources, .root or other files)
- local caching
- redundancy through global replication
- act as global Science CDN (see A. Griffin / G. Aben [1])

5. Plentiful challenges

6. Technical References

All the cloud servers that support federated sharing have to implement two different kinds of public APIs. The first is an API for the actual file transfer. We use WebDAV for that. The second API is a specific federated sharing API that is used to initiate sharing requests and exchange all kinds of other meta and process data.

Both APIs are http/REST base and have to be accessible via the internet. It is strongly recommended to use https/SSL/TLS for optimal security.

The WebDAV API is based on the standard WebDAV protocol. A few extensions are needed for performance improvements. Authentication is done via sharing keys that are passed as http basic auth.

The federated sharing API is used to initiate sharing requests. If a sharing request is accepted or denied by the user on the second server than the accepted or denied codes are sent back via the API to the first server. Part of a sharing request is a sharing key which is used by the second server to access the shared file or folder. This sharing key can later be revoked by the sharer if the share has to be terminated.

The complete API specification is available as a nonpublic preview under the link [2] in the attachments. It will be available as a public API release on the 3rd of August. We appreciate your feedback and collaboration on this preview and of course also the full public release. The idea is to implement this API as an open federated sharing standard that can be implemented by different file sync and share services and providers.

7. Quotes

„In the research area data must not only be under control of scientists. We have to be able to share them with our partners easily and secure without creating data silos. For this reason, we are helping other research institutions to implement their private cloud storage and encourage everyone to participate in the construction of an open standard for data exchange."
Dr. Thomas Hildmann, Project-Lead of Project „ownCloud-Implementation an DFN-Cloud at TU Berlin"

„The ERCIS - European Research Center for Information Systems - is an international network of universities, organisations, and scientists, conducting cooperative research in the field of Information Systems. For such a network, the exchange of sensitive data is crucial. OpenCloudMesh is ideal to support our collaboration!"
Dr. Armin Stein, Managing Director, ERCIS

"SWITCH, the Swiss research and education network, welcomes open standards. We think that cross platform and cross institutional sync and share services are an important step towards connecting researchers across institutions and countries and welcome the OpenCloudMesh initiative as a viable step in that direction. SWITCHdrive, an ownCloud based service, already connects 27 Swiss Universities today and its users will further benefit from this development."
Dr. Andreas Dudler, Managing Director of SWITCH

"We are at a critical juncture in cloud computing. There is no longer a need to choose between privacy and security and collaboration and ease of use. We believe OpenCloudMesh will redefine the way people use the cloud to share their important files."
Peter Szegedi, Project Development Officer, Management Team, GÉANT Association

"OpenCloudMesh is ideally suited to improve the storage cloud service offerings within the Max Planck Society given its distributed nature with more than 80 individual research institutes throughout Germany and beyond."
Dr. Raphael Ritz, Head of Data Division, Garching Computing Centre of the Max Planck Society

"Sciebo is one of Germany's biggest sync&share projects and implemented jointly by 22 universities in North Rhine Westphalia. Since a fast data exchange in teaching and research becomes more important in many modern scientific fields, we are striving for a reliable and secure backbone for this task. We are applying a multiple server site concept and therefore we are relying on the server-to-server sharing feature. Our aim is to offer a transparent and easy user experience which also should be compatible with alternative solutions of external co-operation partners. Thus, we are calling for an open and standardized solution as promised by the OpenCloudMesh protocol."
Dr. Raimund Vogl, Head of ZIV, Univeristiy of Münster

"For many years' scientists at DESY have been exchanging data with partners worldwide. The systems used to do that change with time and we feel that sharing data over a network of federated private clouds is fitting the needs of users today. This is why we are excited to work on the Open Cloud Mesh"
Peter van der Reest, Head of Information Fabrics Group, Information Technology, Deutsches Elektronen-Synchrotron DESY, A Research Centre of the Helmholtz Association

"Enabling collaboration between scholars and researchers is crucial to make progress on the problems of today. For this reason the universities in the state of Florida are working on a collaborative storage infrastructure and the ownCloud software is part of that infrastructure. The functionality was demonstrated recently at SC14 http://news.it.ufl.edu/research/new-storage-system-simplifies-bigdata-sharing/"
Dr. Erik Deumens, Director UF Research Computing, University of Florida

8. Involved Institutions:

- [CERN](#)
- [Sciebo , the 500k user Campuscloud](#)
- [University of Florida](#)
- [SWITCH](#)
- [SURF](#)
- [University of Hawaii – Manoa, College of Education](#)
- [University of Vienna](#)
- [AARNet](#)
- [Garching Computing Centre of the Max-Planck-Society](#)
- [GÉANT Association](#)
- [DESY Deutsches Elektronen-Synchrotron](#)
- [ETH Zürich](#)
- [ERCIS – European Research Center for Information Systems](#)
- [TU Berlin](#)

Attachments

[1] Creating a Research Cloud with Things You Already Have in Your Kitchen
https://tnc2013.terena.org/core/presentation/3

[2] Open Collaboration Services v2 (WIP) pw "OpenCloudMesh" available under
https://s3.owncloud.com/owncloud/index.php/s/wbVQGyucRigQ155