

## An OCM protocol extension to support data transfer in the ScienceMesh

**Antoon Prins (SURF)** 



**CS3MESH4EOSC** has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.





#### Ron Trompert, Thirsa de Boer, Antoon Prins

#### On-demand Data Transfers (SURF, PSNC, CESNET, AARNET, CERN)



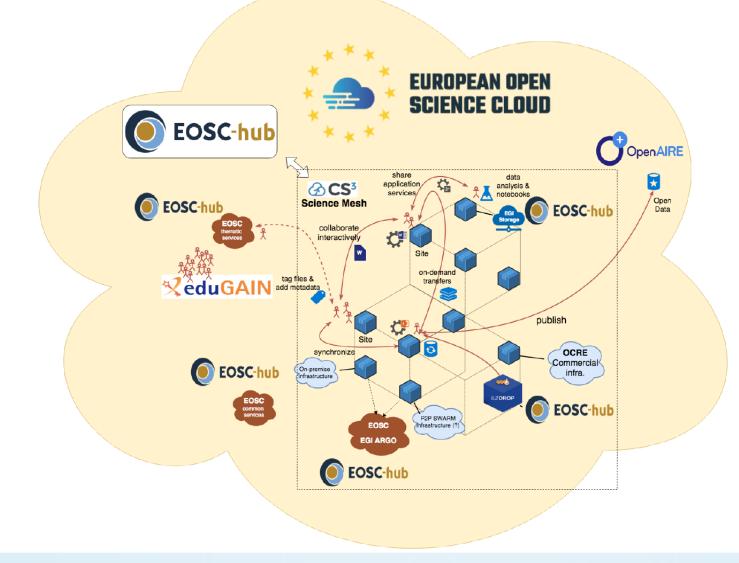
## CS3MESH4EOSC is a H2020 Project (2020-2022)

- Interconnects existing sustainable storage and research services
- Create an innovative collaborative mesh of application services
- \* Supports open source software and FAIR, open data principles
- Build an operational, decentralised infrastructure (with a lightweight central component) based on the protocols:
  - CS3APIs (links apps with an EFSS)
  - Note: No

https://cordis.europa.eu/project/id/863353

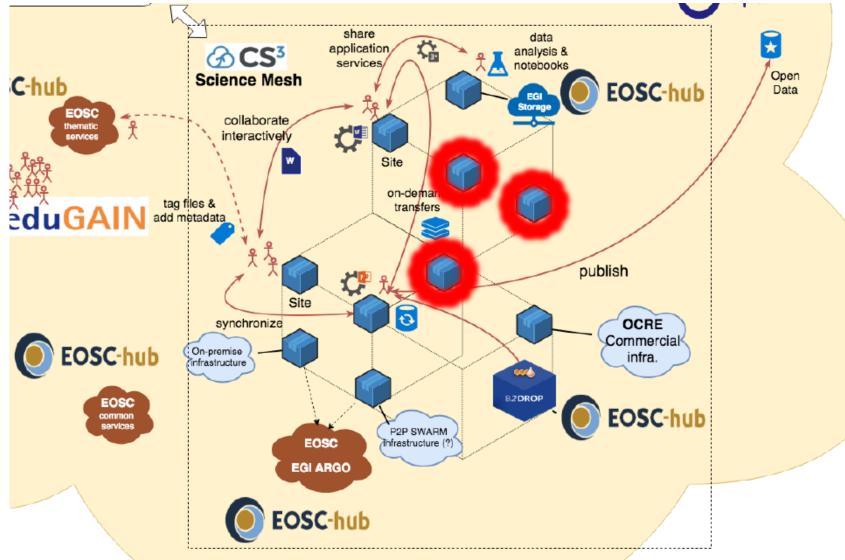


#### The CS3MESH4EOSC ScienceMesh



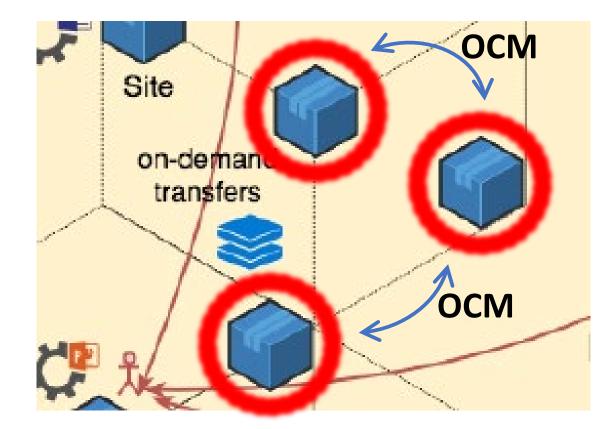


#### The CS3MESH4EOSC ScienceMesh





#### OCM in the ScienceMesh





## Why on-demand data transfers?

#### Data locality may be important for some use-cases

- Sharing of data is not good enough
  - Remote processing may cause overload on the source EFSS system, processing at source is not possible
- Larger data sets

Support for data transfers between EFSS systems and between EFSS systems and other storage systems

Transfer mechanism agnostic



## OCM is instrumental in data sharing between EFSS' invitation flow & data sharing

# Use OCM for data transfer as well ... Yes it can

Same steps, as in data sharing
 Invite, share, accept



#### **POST /shares request**

```
"shareWith": "peter.szegedi@geant.org",
"name": "spec.yaml",
"description": "This is the Open API Specification file (in YAML format) of the Open Cloud Mesh API.",
"providerId": "7c084226-d9a1-11e6-bf26-cec0c932ce01",
"owner": "dimitri@apiwise.nl",
"sender": "john@apiwise.nl",
"ownerDisplayName": "Dimitri",
"senderDisplayName": "John Doe",
"shareType": "user",
"resourceType": "file",
"protocol": {
 "name": "webdav",
  "options": {
    "sharedSecret": "hfiuhworzwnur98d3wjiwhr",
    "permissions": "{http://open-cloud-mesh.org/ns}share-permissions"
```



#### **POST /shares request**

```
"shareWith": "peter.szegedi@geant.org",
"name": "spec.yaml",
"description": "This is the Open API Specification file (in YAML format) of the Open Cloud Mesh API.",
"providerId": "7c084226-d9a1-11e6-bf26-cec0c932ce01",
"owner": "dimitri@apiwise.nl",
"sender": "john@apiwise.nl",
"ownerDisplayName": "Dimitri",
"senderDisplayName": "John Doe",
"shareType": "user",
"resourceType": "file",
"protocol": {
  "name": "webdav",
  "options": {
    "sharedSecret": "hfiuhworzwnur98d3wjiwhr",
    "permissions": "{http://open-cloud-mesh.org/ns}share-permissions"
```



#### POST /shares request

"shareWith": "peter.szegedi@geant.org", "name": "spec.yaml",			
"description": "This is the Open API Specification file (in YAML format) of the Open Cloud Mesh API.",			
"providerId": "7c084226-d9a1-11e6-bf26-cec0c932ce01",			
"owner": "dimitri@apiwise.nl",			
"sender": "john@apiwise.nl",			
"ownerDisplayName": "Dim	itri",		
"senderDisplayName":			
"shareType": "user",	protocol > required	object	
"resourceType": "fil	required	The protocol which is used to establish synchronisation. At the moment only webdav is supported, but other (custom) protocols might be added in the future.	
"protocol": {			
"name": "webdav",			
"options": {			
"sharedSecret": "hfiuhworzwnur98d3wjiwhr",			
"permissions": "{http://open-cloud-mesh.org/ns}share-permissions"			
}			
}			



#### POST /shares request (Data Transfer)

```
"shareWith": "peter.szegedi@geant.org",
"name": "spec.yaml",
"description": "This is the Open API Specification file (in YAML format) of the Open Cloud Mesh API.",
"providerId": "7c084226-d9a1-11e6-bf26-cec0c932ce01",
"owner": "dimitri@apiwise.nl",
"sender": "john@apiwise.nl",
"ownerDisplayName": "Dimitri",
"senderDisplayName": "John Doe",
"shareType": "user",
"resourceType": "file",
"protocol": {
  "name": "datatx",
  "options": {
   "sharedSecret": "hfiuhworzwnur98d3wjiwhr",
    "permissions": "{http://open-cloud-mesh.org/ns}share-permissions"
```



#### Implementing OCM based Data Transfer

#### Data Transfer is a case

```
switch protocol.Value {
    case "webdav":
        ... regular file share
    case "datatx":
        ... data transfer
    case default:
        ... protocol not recognized
}
```

Services to manage the data transfer

Data Transfer tool (rclone, FTS, ...)

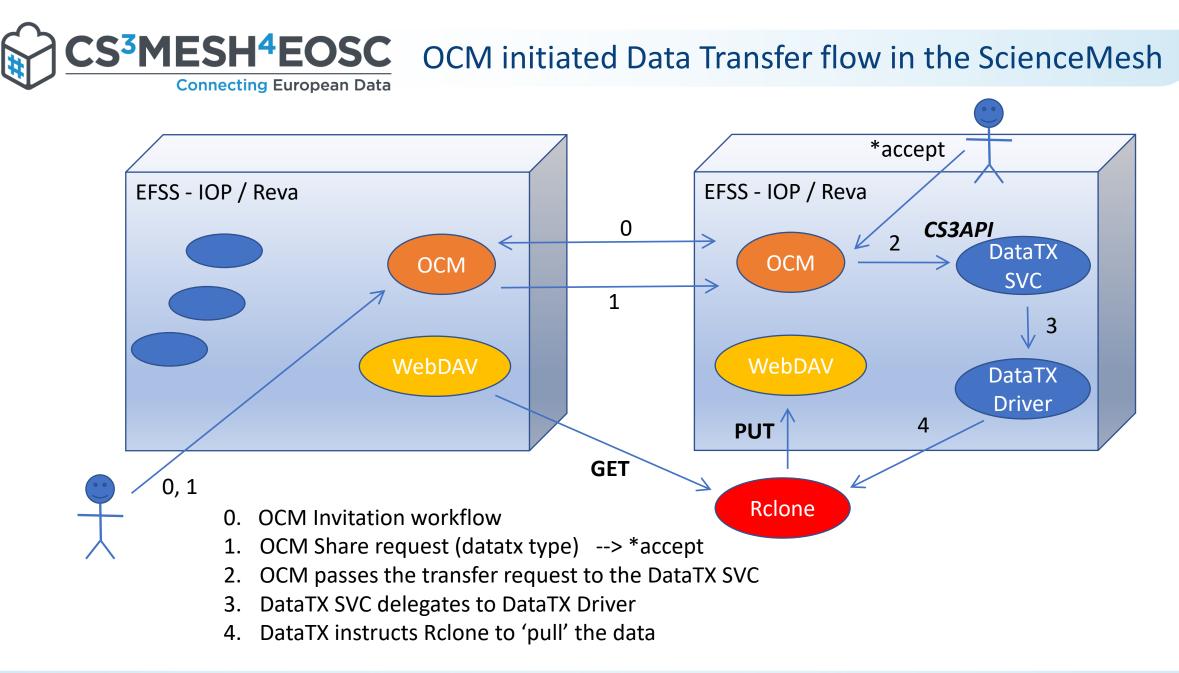


### \* EFSS' on top of the IOP layer

- Build on CS3APIs (collaboration API)
- \* The implementation of this is the Reva project (<u>https://github.com/cs3org/reva</u>)
- Reva is installed alongside an EFSS

#### ScienceMesh of application services

- OCM for linking EFSS'
- CS3APIs in the IOP for linking apps, storages to EFSS





### Data transfer service in Reva (CS3API)

- PullTransfer
- \$ GetTransferStatus, ListTransfers, CancelTransfers, RetryTransfers

For the management methods we have created CLI commands
 PullTransfer is never called explicitly by the user → the user accepts the share

rclone driver

Some essential rclone changes since v1.55.0



#### CCM can support data transfer through a simple protocol extension POST /shares protocol: name: datatx

## OCM flow for Data Transfer is very similar to regular Data Share OCM Invite & Share & Accept

In Reva the services to support this have been implemented
 To be merged into the Reva master branch
 checkout the screencast at <u>https://youtu.be/xfhP7AigLtw</u>





#### **Connecting European Data**

## Thank you! Discover more on...

Cs3mesh4eosc.eu

in company/cs3mesh4eosc

CS3org

CS3MESH4EOSC Project

https://www.youtube.com/channel/UCHKcZEkMqXjCvc3MLFjFxbw

