

VRE & RDM: $6 + 6 = 11$



SURF

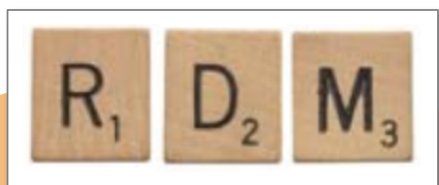


- Hylke Koers
- Groepsleider Data Management Services, SURFsara

RDM & VRE: Meerwaarde door integratie



RDM & VRE: Meerwaarde door integratie

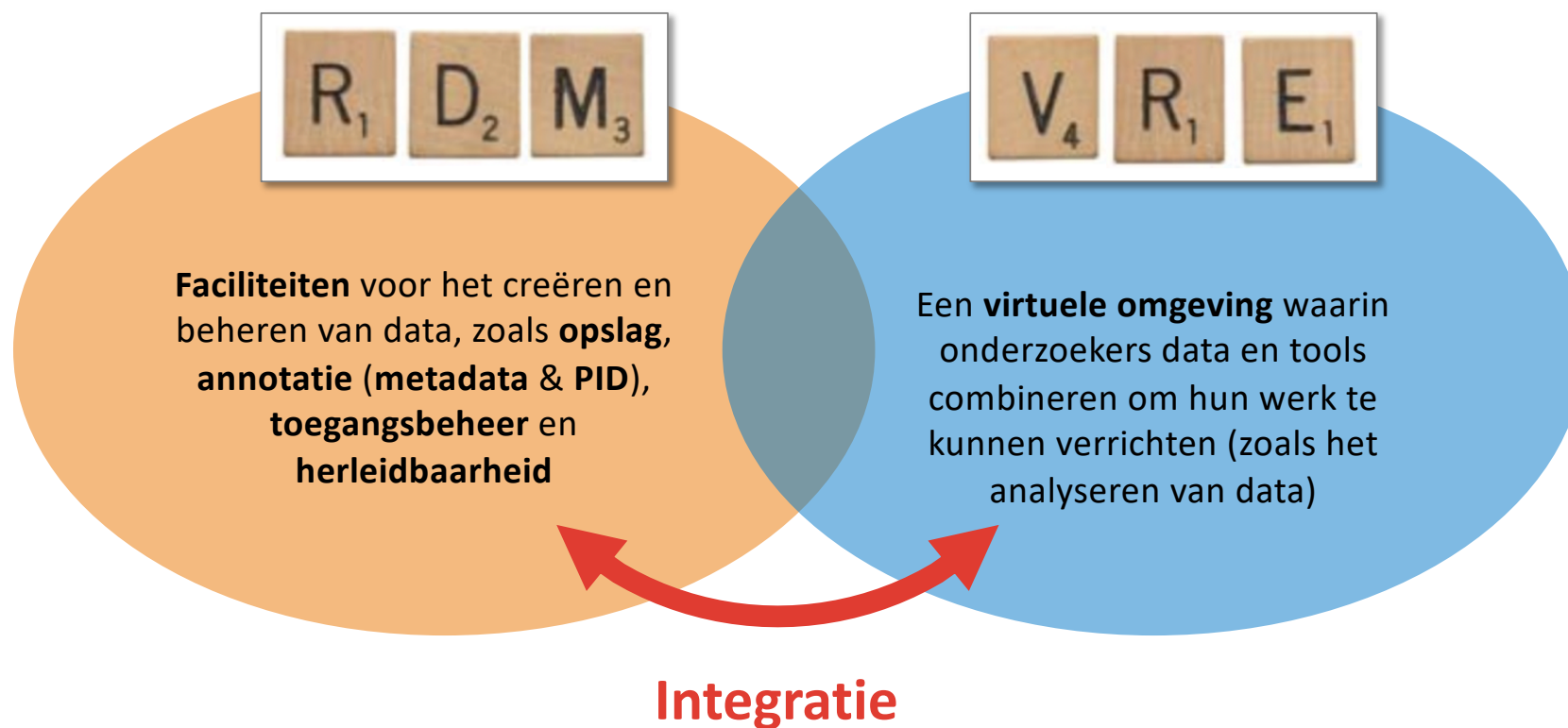


Faciliteiten voor het creëren en beheren van data, zoals **opslag**, **annotatie (metadata & PID)**, **toegangsbeheer** en **herleidbaarheid**

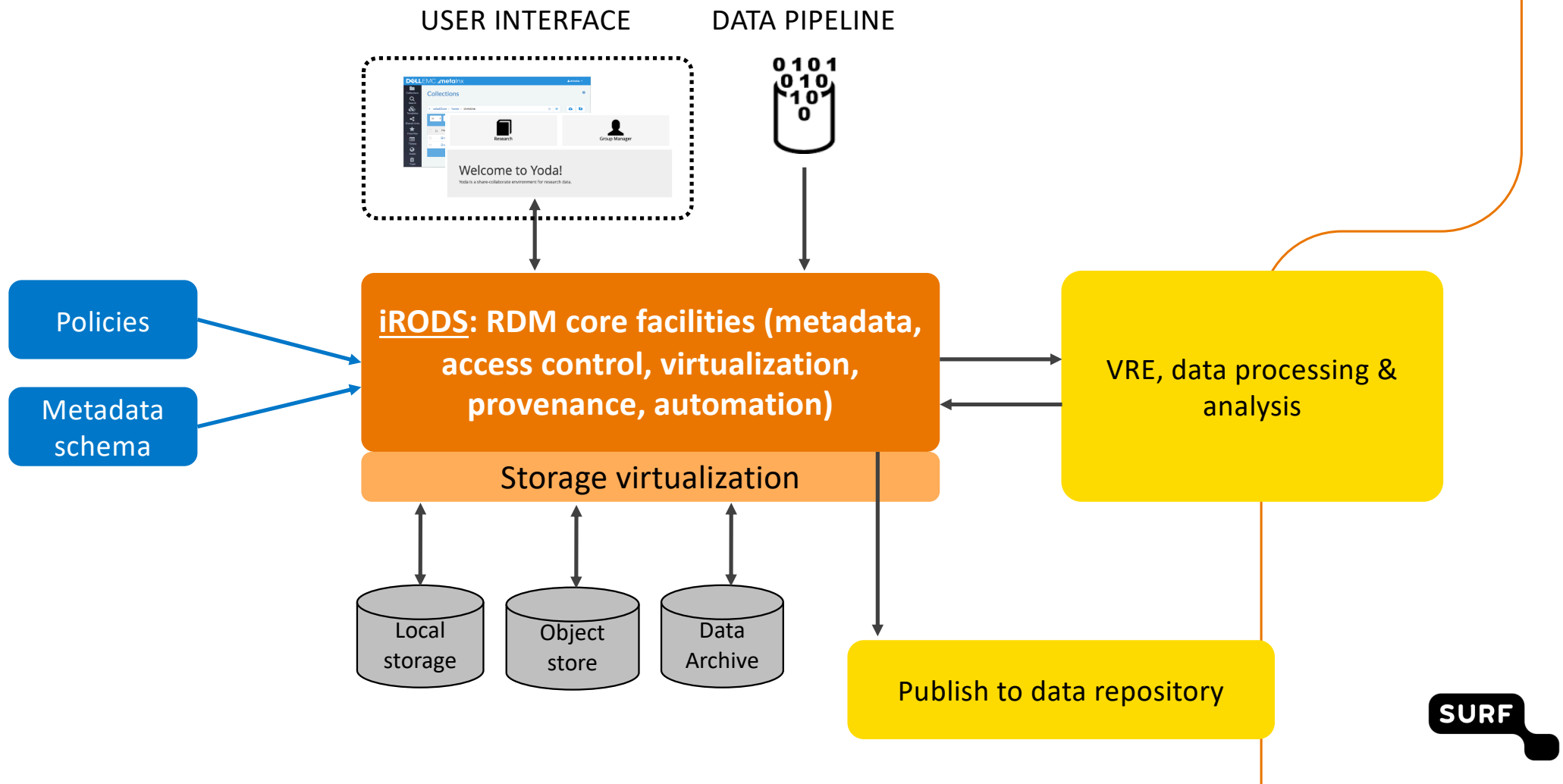


Een **virtuele omgeving** waarin onderzoekers data en tools combineren om hun werk te kunnen verrichten (zoals het analyseren van data)

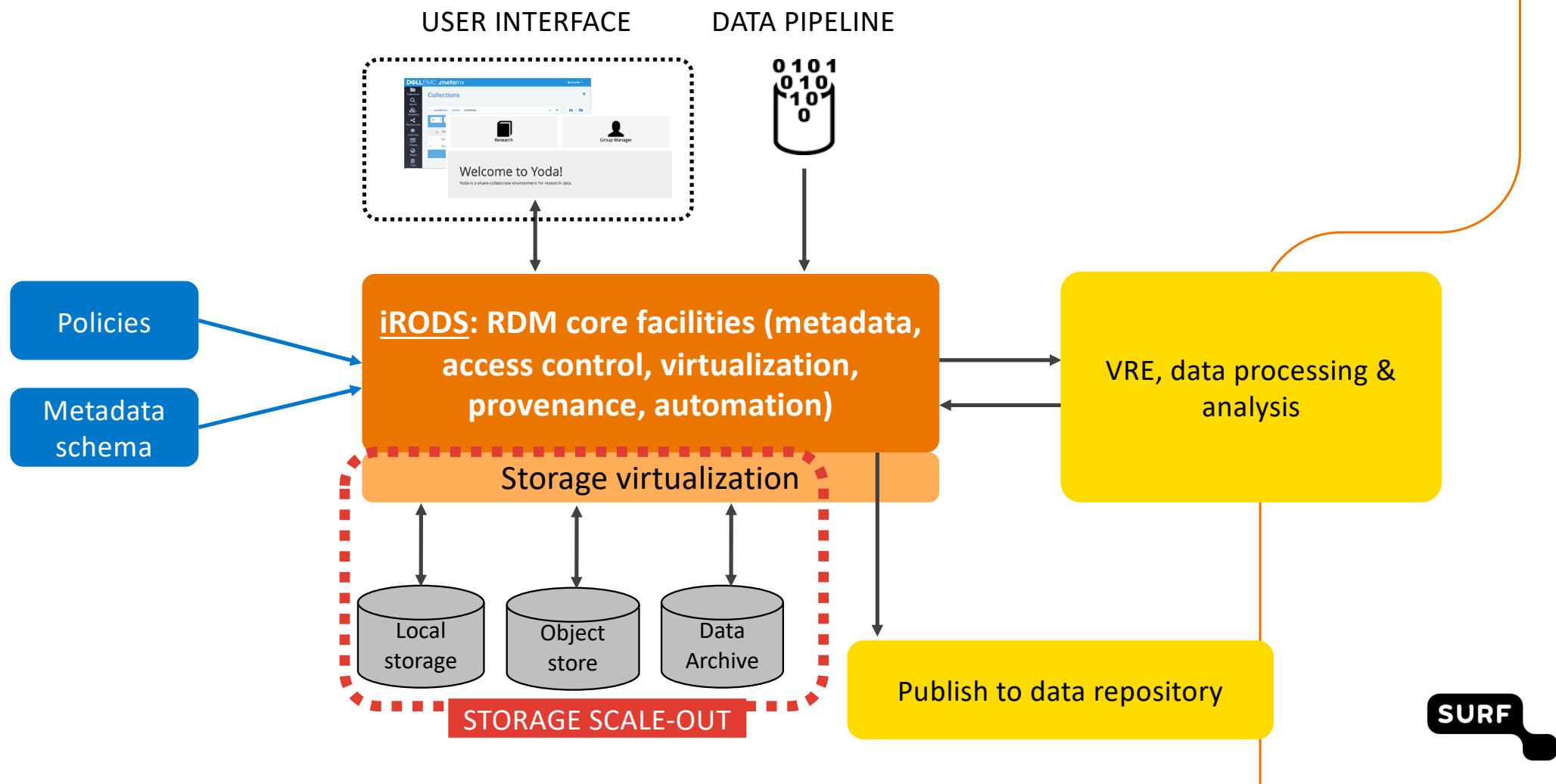
RDM & VRE: Meerwaarde door integratie



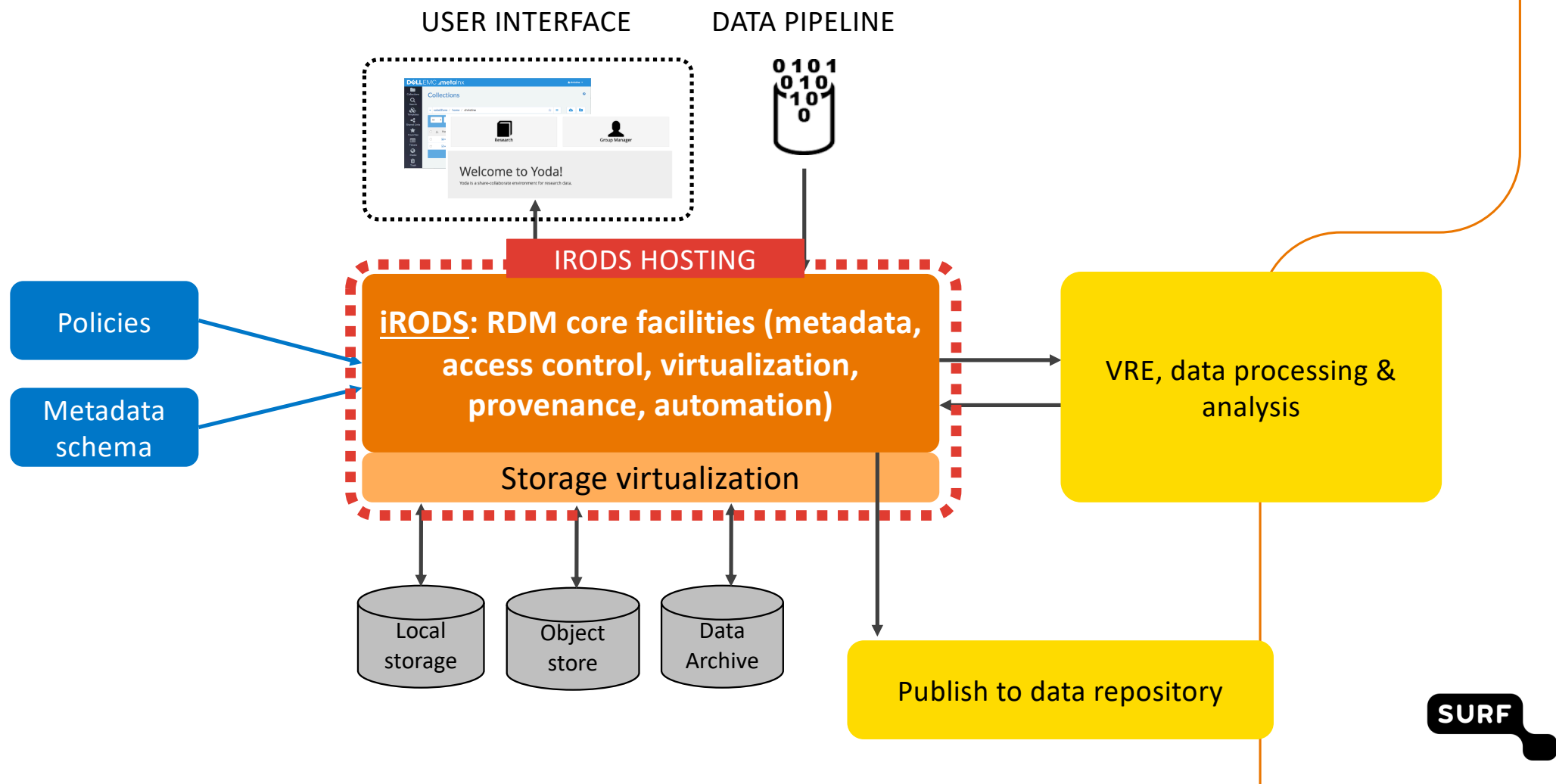
Dienstontwikkeling: RDM modules met iRODS als 'spin in het web'



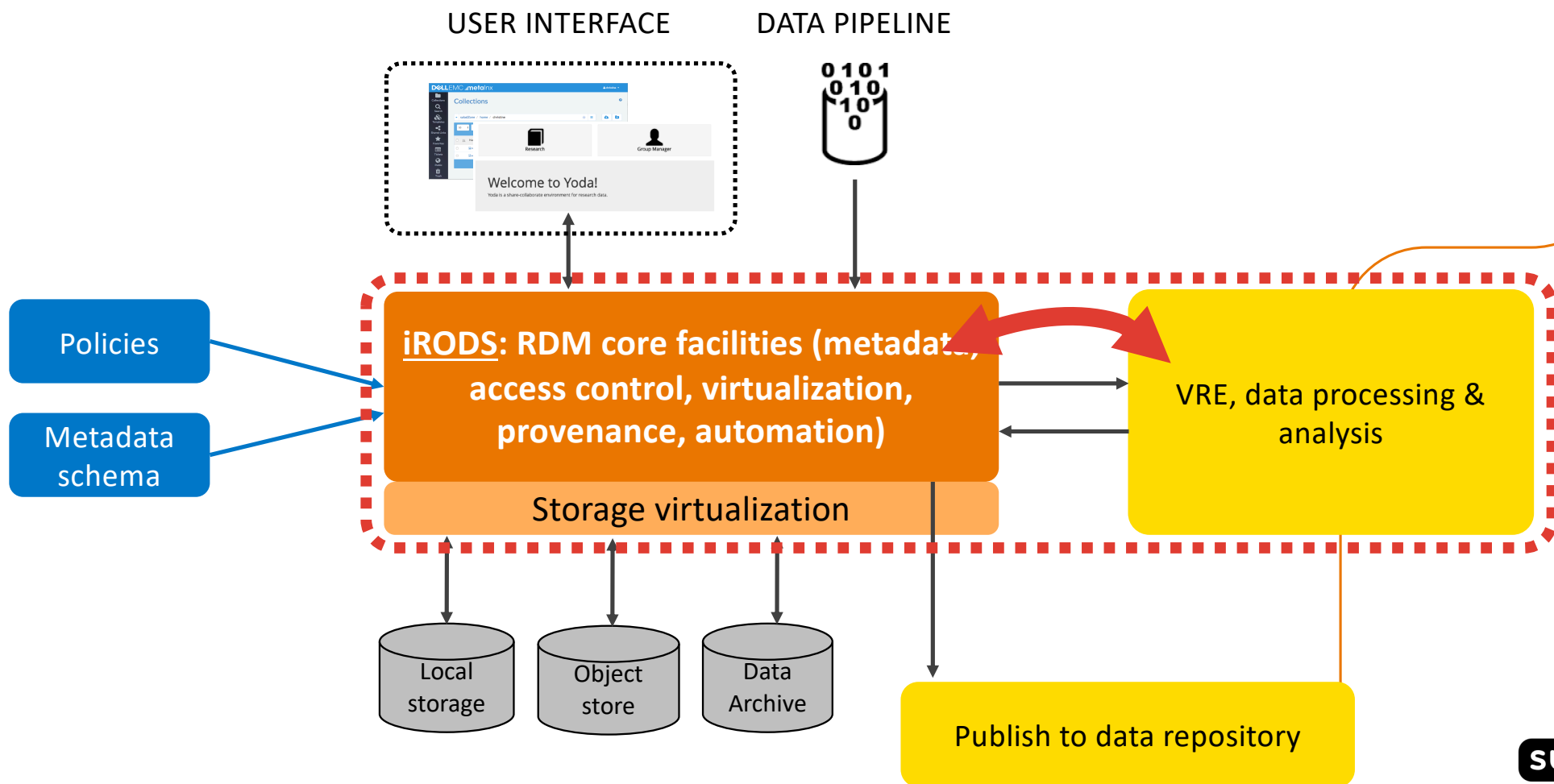
SURF dienst: RDM storage scale-out



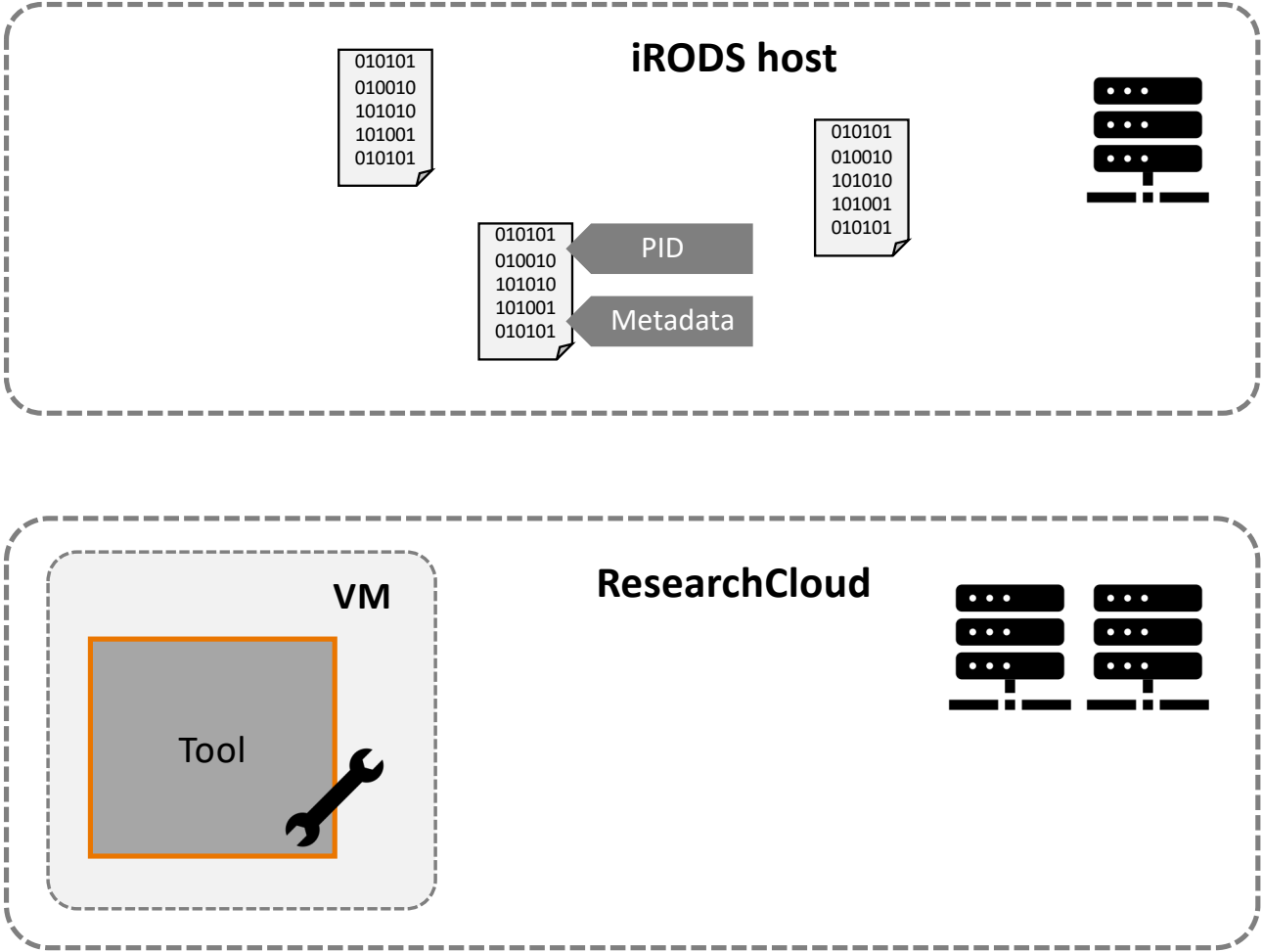
SURF dienst: iRODS hosting



Dienstontwikkeling: integratie iRODS & VRE omgeving

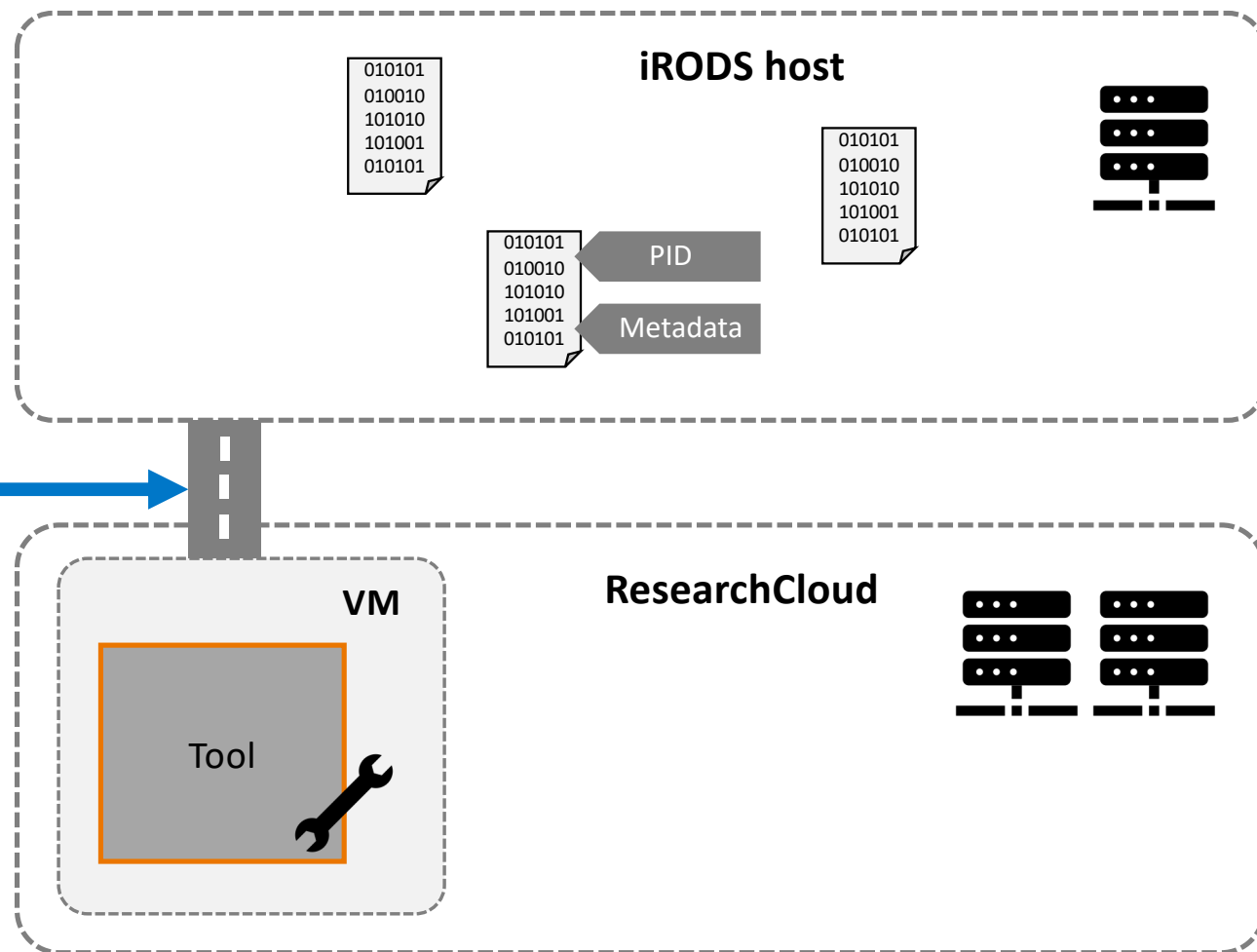


Demo time!

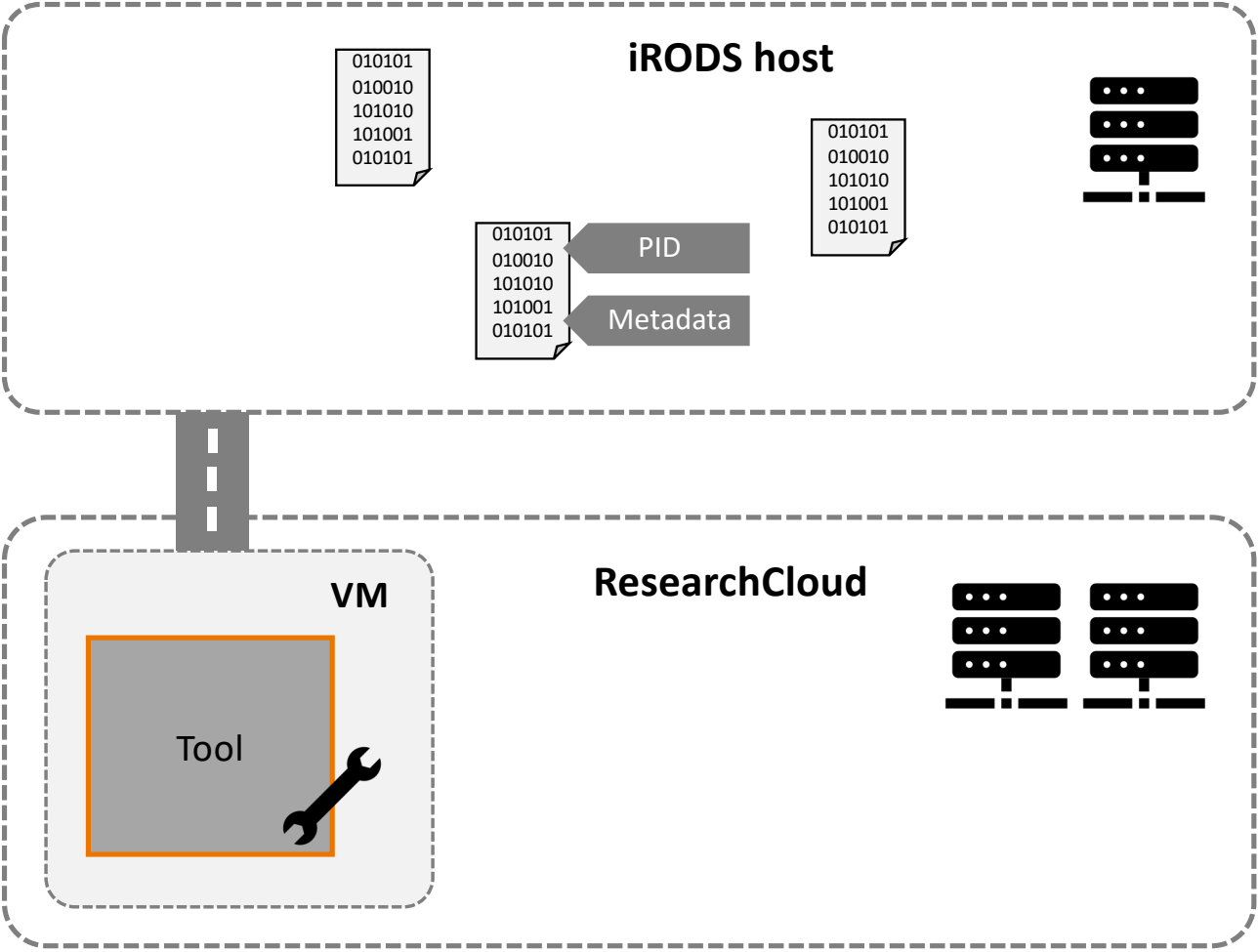


Demo time!

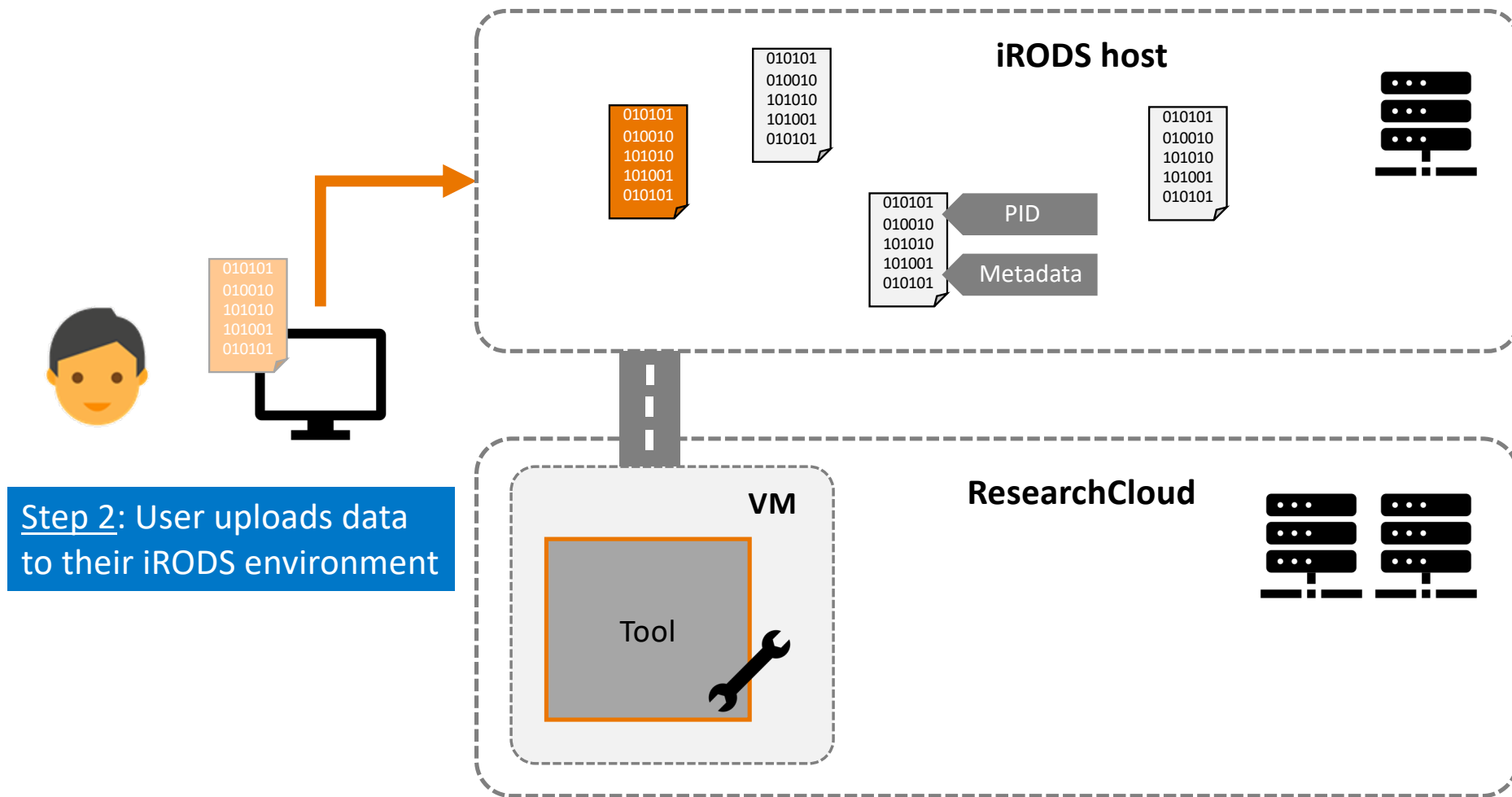
Set up: Integration
iRODS – RSC (through
WebDAV mount)



Demo time!



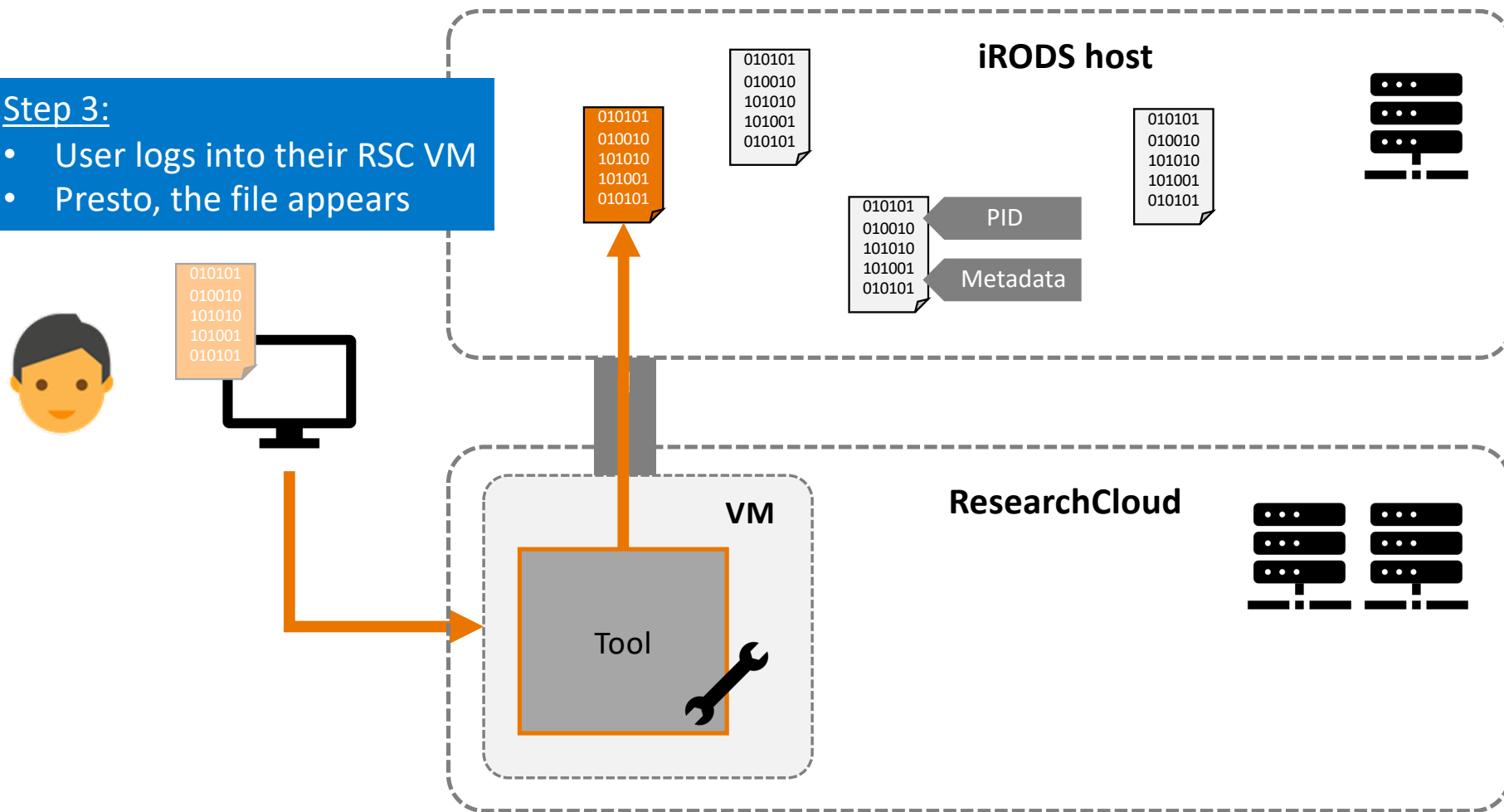
Demo time!



Demo time!

Step 3:

- User logs into their RSC VM
- Presto, the file appears



SET-UP:

RSC VM CONNECTED WITH IRODS HOST



SURF

Quick actions

☒ Show deleted workspaces

Create your workspace [Configure manually](#)

- 1 Collaborative organisation
- 2 Application
- 3 Dataset(s)
- 4 Power required
- 5 Cloud solution
- 6 Name

Are you going to work in a collaborative organisation, or a private space?

A workspace and its applications needs to be connected to a collaborative organisation. Please select the collaborative organisation where your workspace will belong to.

Tom Test

You already belong to this collaborative organisation. In order to start a new workspace and invite others to collaborate, please choose this option.

CHOOSE

iRODS environment for Claudio

You already belong to this collaborative organisation. In order to start a new workspace and invite others to collaborate, please choose this option.

CHOOSE

Private - no collaborative organisation

This means you are starting a workspace without any other members, and it is privately funded.

CHOOSE

CONTINUE

Create your workspace

[Restart workspace creation](#)

✓ Collaborative organisation

✓ Application

3 Dataset(s)

4 Power required

5 Cloud solution

6 Name

Choose the dataset(s) from the catalog that fits your needs

The following datasets are available for your workspace, based on your approved requests for datasets in the catalog. Choosing more than one selection is possible.

Ionosphere Data Set

Ionosphere dataset

CHOOSE

AMC cloud workshop 2019

AMC cloud workshop datasets and scripts

CHOOSE

BACK

CONTINUE WITHOUT SELECTION

Continue without selection means you want to use your own data. You can import your data in the chosen application.



Create your workspace

[Restart workspace creation](#)

- ✓ Collaborative organisation
- ✓ Application
- ✓ Dataset(s)
- 4 Power required
- 5 Cloud solution
- 6 Name

Choose the power you want to use

What is the size of your dataset? Your size requirements determine the power required from the VM.

Small

VM specifications

- Up to 2 cores
- 2 Gigabytes of memory

CHOOSE

Medium

VM specifications

- Up to 4 cores
- 8 Gigabytes of memory

CHOOSE

Large

VM specifications

- Up to 8 cores
- 32 Gigabytes of memory

CHOOSE

Extra Large

VM specifications


- Up to 64 cores
- 256 Gigabytes of memory

CHOOSE

BACK

CONTINUE

Choose one of the available budgets

 Claudio Cacciari irod pilot budget
Credits: 200


SELECTED ✓

Choose the solution that best fits your needs


HPC Cloud @ SURFsara

Netherlands

Small

 Total cost 5 days
4.80 credits

2% of the budget


 Power
1 core, 2 GB RAM

CHOOSE


Azure on UVA-HVA subscription

Public

Small

 Total cost 5 days
6.00 credits

3% of the budget

 Power
1 core, 2 GB RAM

CHOOSE

BACK

CONTINUE



Create your workspace

[Restart workspace creation](#)

✓ Collaborative organisation — ✓ Application — ✓ Dataset(s) — ✓ Power required — ✓ Cloud solution — 6 Name

Almost there! Some final details

Name your workspace so you can recognise it later!

Workspace name *

MyWonderfulWorkSpace

✓ I have read and agree to the [Terms and Conditions](#).

BACK

SUBMIT



Ubuntu VM 18.04 with iRods - 15

State: running

MyWonderfulWorkSpace

State: running

Workspace Details

Owner

Claudio Cacciari

Collaborative Organisation

iRODS environment for Claudio

Date Created

6th of April (14:31)

Cost so far

0 credits (0.96 credits /day)

Will be removed

11th of April (14:25)

Projected total cost

\$4.8

Application Details

Application and datasets used

Ubuntu 18.04 with iRods with no dataset

Cloud

HPC Cloud @ SURFsara

Size

small (1 cpu, 2 GB RAM)

IP address

145.100.58.31

DELETE

ACCESS

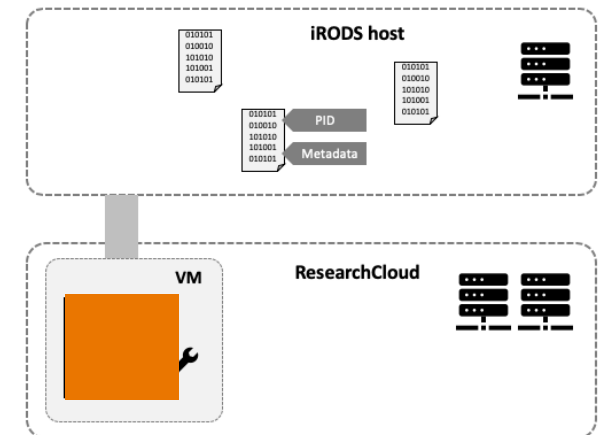
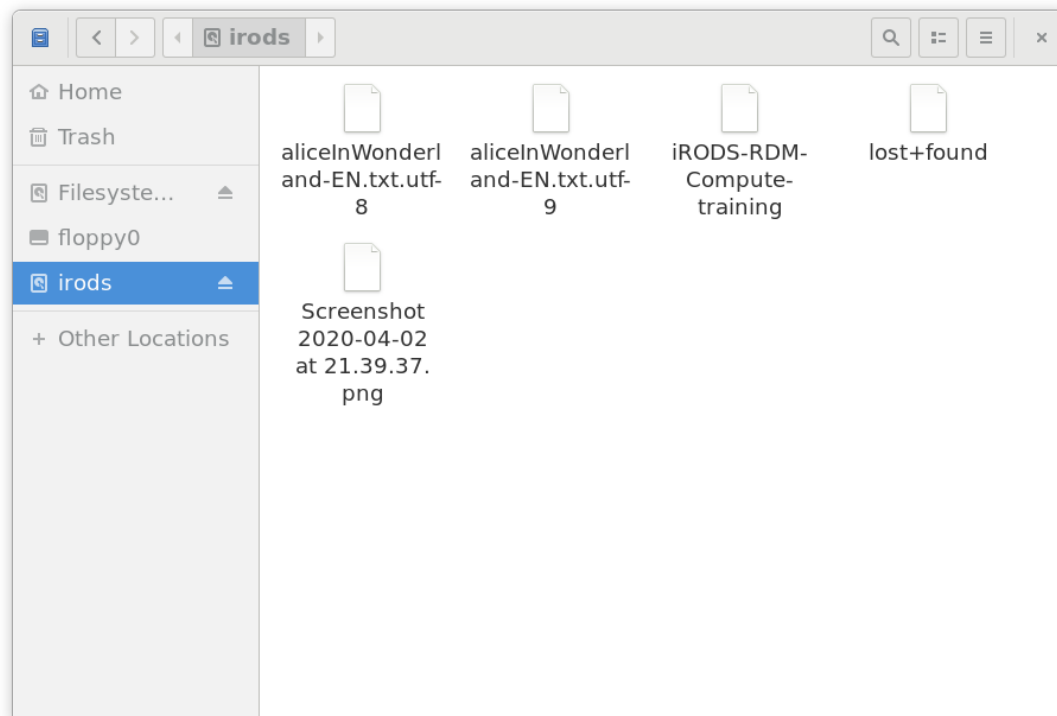
USE

RSC VM & IRODS, INTEGRATED



SURF


File viewer on RSC VM



WebDAV to iRODS host

Cancel

Connect

 **Enter password for SURFDAV**

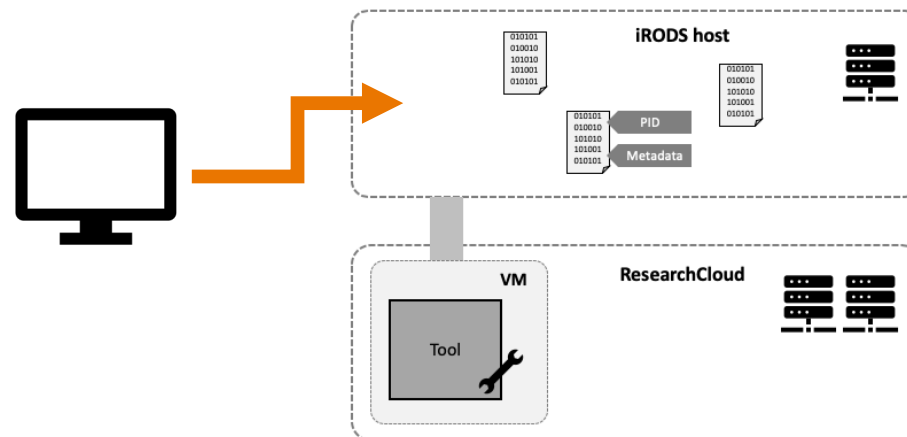
Username

Password

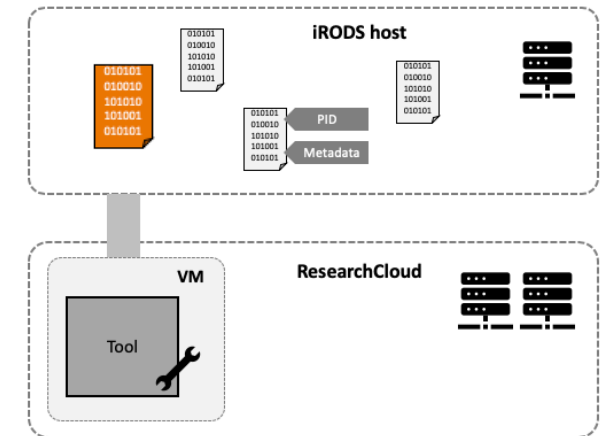
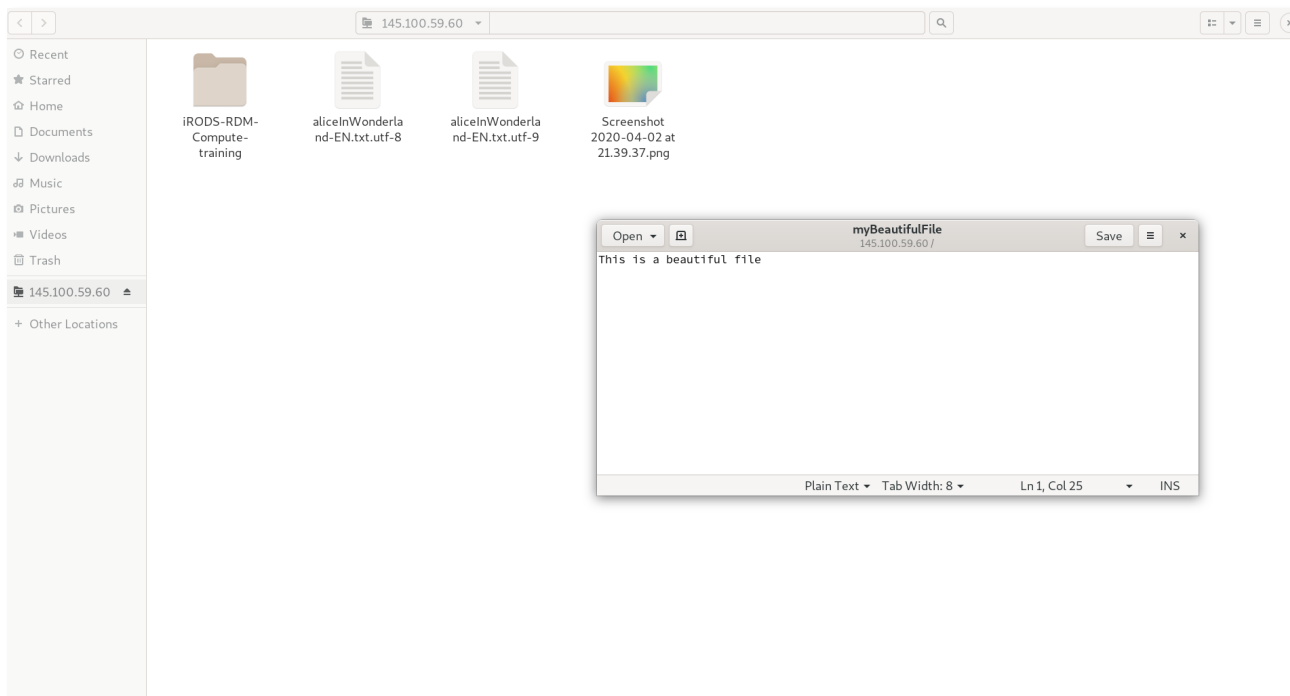
☐ Forget password immediately

☒ Remember password until you logout

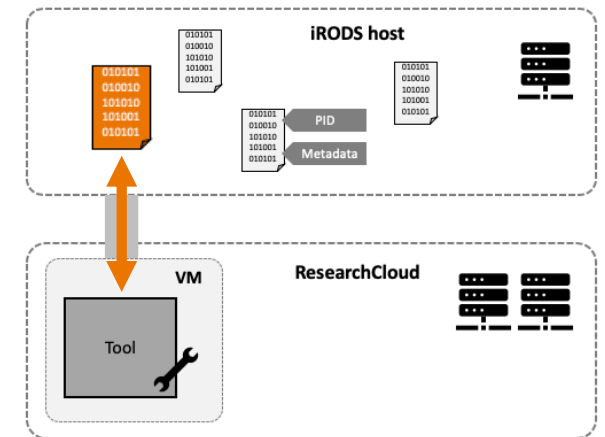
☐ Remember forever



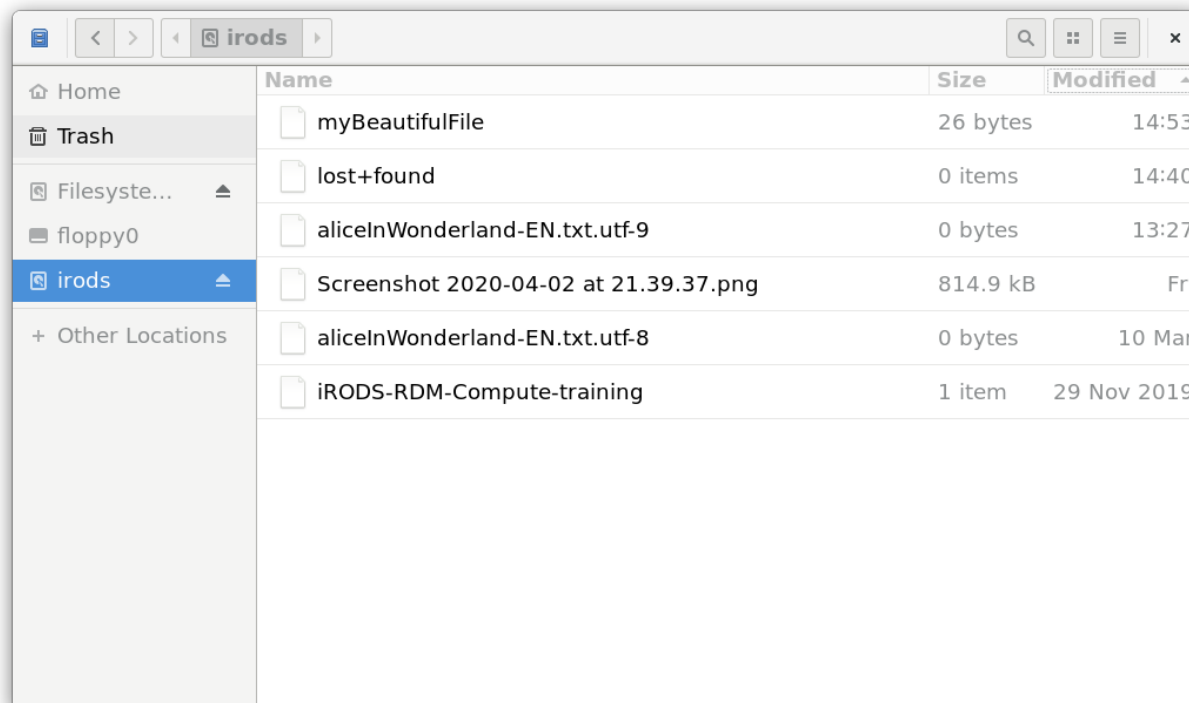
Upload file to iRODS host



Behind the scenes...

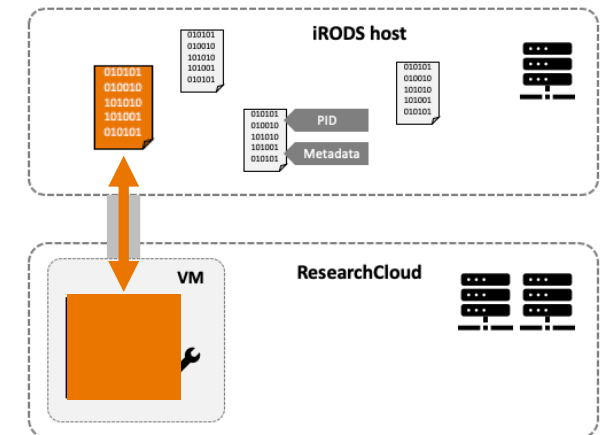


File viewer on RSC VM (145.100.58.31)



The screenshot shows a file viewer interface with a sidebar on the left and a main table on the right. The sidebar contains navigation options: Home, Trash, Filesystem..., floppy0, irods (selected), and + Other Locations. The main table displays a list of files and folders with columns for Name, Size, and Modified.

Name	Size	Modified
myBeautifulFile	26 bytes	14:53
lost+found	0 items	14:40
aliceInWonderland-EN.txt.utf-9	0 bytes	13:27
Screenshot 2020-04-02 at 21.39.37.png	814.9 kB	Fri
aliceInWonderland-EN.txt.utf-8	0 bytes	10 Mar
iRODS-RDM-Compute-training	1 item	29 Nov 2019



Volgende stappen

- Testen integratie met real-life use cases met focus op realiseren waarde van iRODS: metadata, PID, provenance, etc.
- Automatiseren van set-up
- Integratie metadata
- Koppelen met Federated Identify & Access Management (SURF Research Access Management)
- Koppelen met ResearchDrive

ANNEX

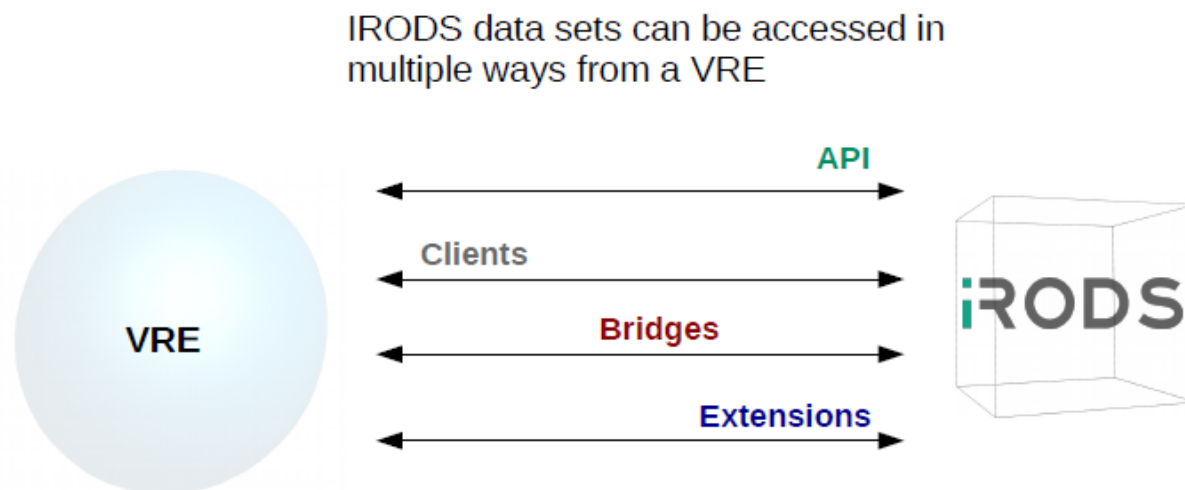


SURF

Welke bouwblokken biedt een RDM systeem?

- Vindbaarheid van data (metadata en meer)
- “Track-and-trace” (provenance) van data van experiment tot analyse tot publicatie/archivering
- Automatisering van data opslag en de bijbehorende trade-offs zoals:
 - Data dicht bij compute nodes voor snelle bewerking; vs.
 - Data op tape voor kosteneffectieve lange-termijn opslag.
- Beheren van metadata volgens community standaarden
- Toegangsrechten op data
- Data (automatisch) beschikbaar maken in de virtuele werkomgeving
- Data publiceren in een data repository vanuit de werkomgeving

Integratie iRODS en VRE



Bridges are software clearly connecting iRODS with a specific framework/platform. While **Extensions** are software which extend iRODS interoperability, but are not targeting a specific platform.

Work in progress!

API:

- **GoRODS**: <https://github.com/ijacquay712/GoRODS>
- **Minio iRODS Gateway (S3)**: <https://github.com/bioteam/minio-irods-gateway>
- **Perl-irods-wrap**: <https://github.com/wtsi-npg/perl-irods-wrap>
- **C/C++**: <https://github.com/irods/irods>
- **Java jargon**: <https://github.com/DICE-UNC/jargon>
- **REST API**: <https://github.com/DICE-UNC/irods-rest>
- **iRODS-PHP**: <https://github.com/UtrechtUniversity/irods-php/tree/development>
- **python-irodsclient**: <https://github.com/irods/python-irodsclient>

Bridges:

- **Ri rods**: https://github.com/irods/irods_client_library_r_cpp
- **Parrot VFS**: <http://ccl.cse.nd.edu/software/parrot>
- **irods_client_library_r_rest**: https://github.com/irods/irods_client_library_r_rest
- **jupyterlab_irods**: <https://github.com/towicode/IJab>
- **irods_tools_lustre**: https://github.com/irods-contrib/irods_tools_lustre
- **Samba**: https://github.com/davorvusic/vfs_stor
- **GridFTP**: <https://github.com/EUDAT-B2STAGE/B2STAGE-GridFTP>

Clients:

- **Kanki**: <https://github.com/ilarik/kanki-irodsclient>
- **Tears**: <https://github.com/whitwham/tears>
- **ishell**: <https://github.com/niess/ishell>
- **Cloud Browser**: <https://github.com/DICE-UNC/irods-cloud-browser>
- **Metalnx**: <https://github.com/irods-contrib/metalnx-web>
- **iDrop**: <https://github.com/DICE-UNC/idrop>
- **icommand**: https://github.com/irods/irods_client_icommands
- **NFSRODS**: https://github.com/irods/irods_client_nfsrods
- **Baton**: <https://github.com/wtsi-npg/baton>
- **Fuse**: https://github.com/irods/irods_client_fuse
- **Davrods**: <https://github.com/UtrechtUniversity/davrods>
- **Cyberduck**: <https://cyberduck.io/irods/>

Extensions:

- **iRODS Automated Ingest Framework**: https://github.com/irods/irods_capability_automated_ingest
- **S3 resource**: https://github.com/irods/irods_resource_plugin_s3