Progress Highlights on Enabling service integration and sensitive data workflows in DICE

EUDAT Conference September 13th – 15th, 2022, Athens, Greece

Chris Ariyo, CSC IT Center for Science, Finland
DICE T4.1 Leader & T4.4 member

The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207

What is SecureB2SHARE

- Secure B2SHARE is a framework actively developed, evaluated and started during EOSC-Hub by
  - Sigma2 / University of Oslo in Norway
  - CSC – IT Center for Science in Finland
  - Based on EUDAT B2SHARE service
    - With the addition of sensitive data specific changes and functionality
    - Allows storing, describing and sharing sensitive datasets
    - Controlling access to the datasets without jeopardizing privacy or security

SecureB2SHARE

Challenge: provide the sensitive data resources without violating privacy, yet maintaining FAIR principles in a user-friendly way.

Metadata stored in B2SHARE.

Sensitive data stored in access restricted secure storage.

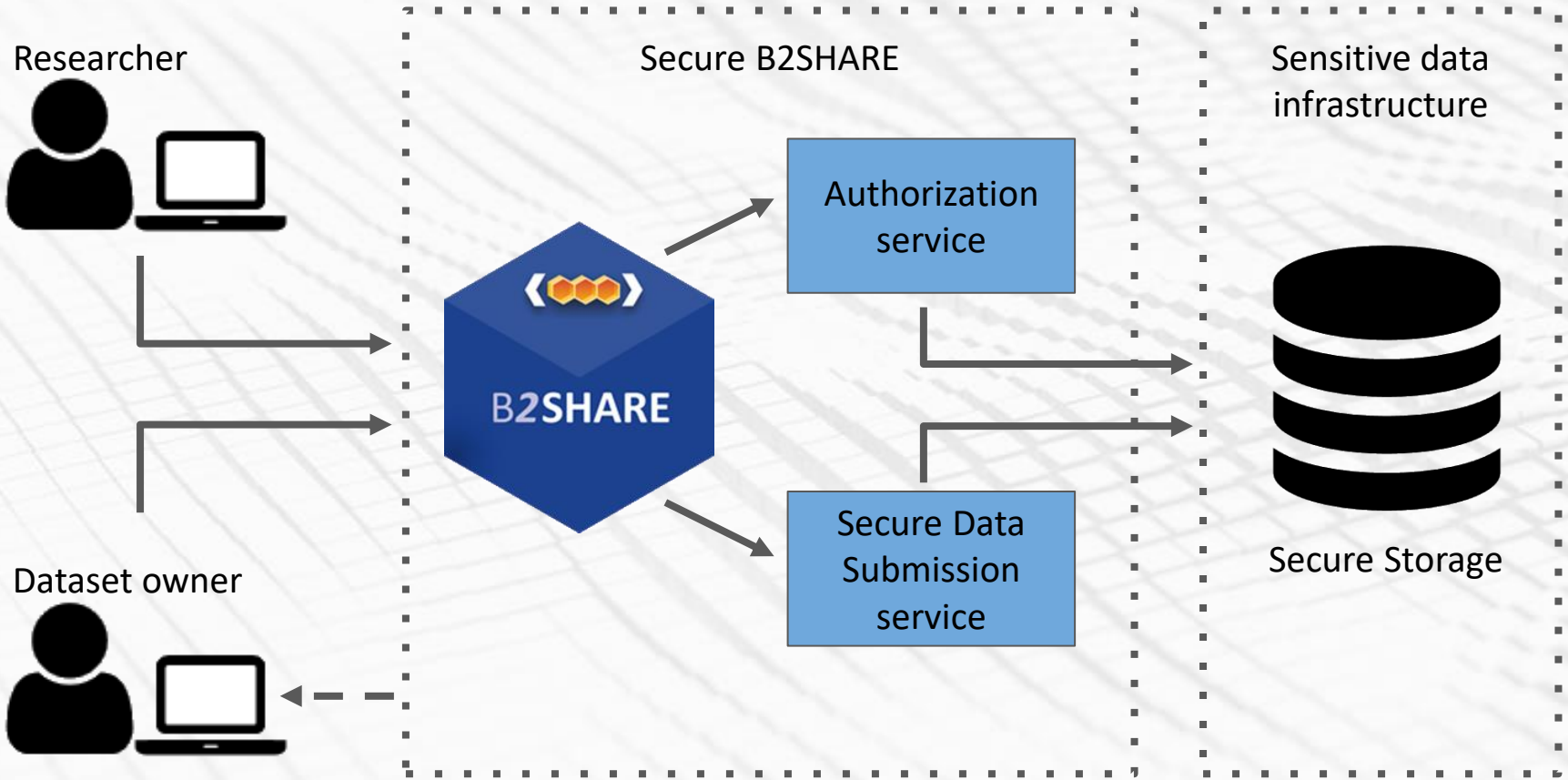
Data access either inside secure environment or through file API.
(implementation specific policy decision)

Pilot instance at UiO, Norway.
Work-in-progress instance at CSC, Finland.

SecureB2SHARE interaction model

- Discover datasets
- View dataset metadata
- Request access to dataset

- Submit datasets
- Describe dataset metadata
- Authorize access to datasets



- Stores metadata of published sensitive data
- Handles access requests to data

- Stores sensitive data
- Manages access to data

Secure B2SHARE

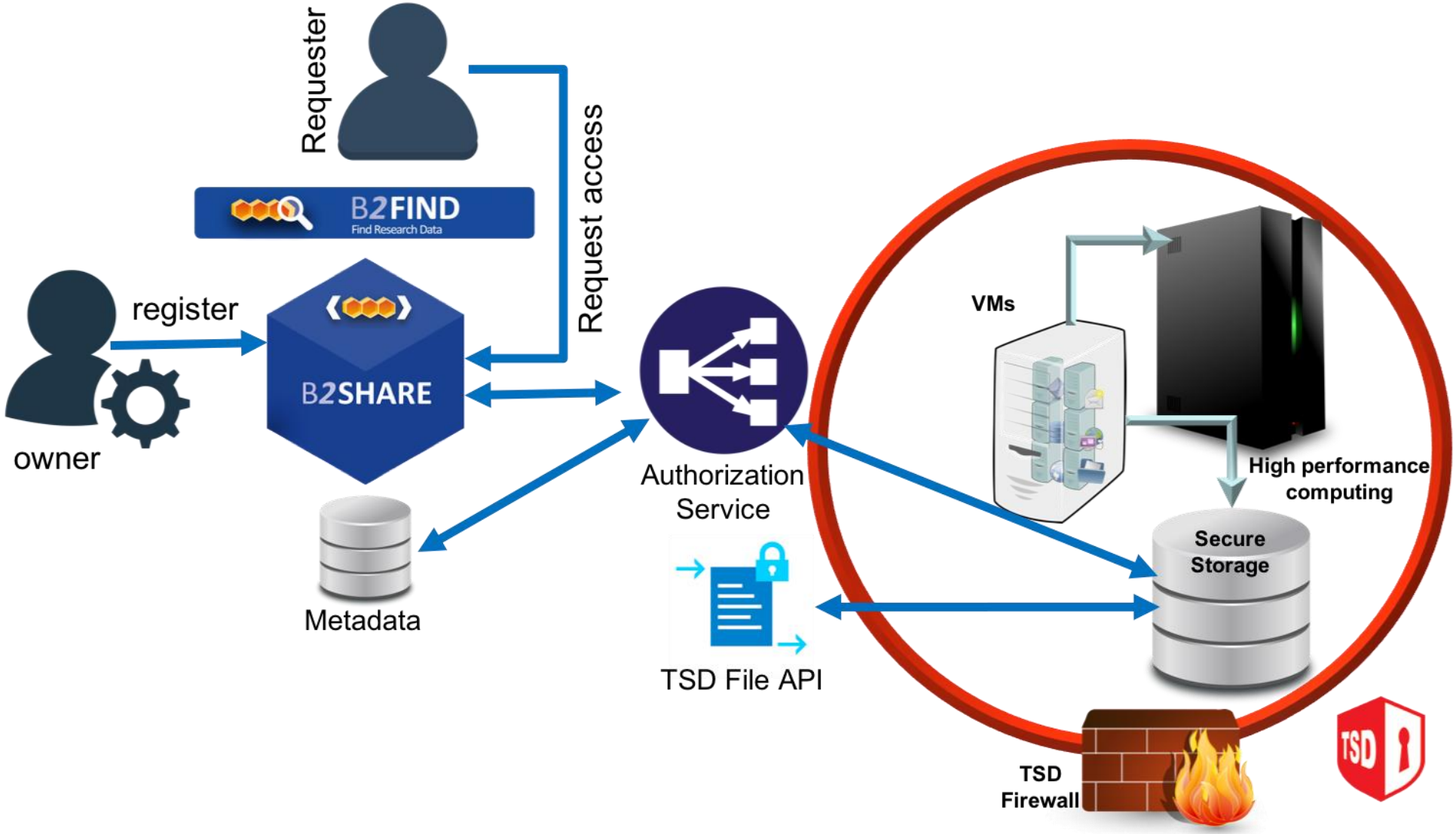
- Development effort is maintaining a shallow fork of EUDAT B2SHARE where
  - B2SHARE is acting as user-friendly metadata browser
  - File upload functionality is removed
  - Records are pointing towards an access request interface

TSD - Services for Sensitive data (University of Oslo)

- In production since 2014
- Norway's largest PaaS e-Infrastructure for sensitive data
- 6400+ unique users, 1600+ active projects
- 3500+ user support cases per year
- 4 PiB of storage and growing
- National service via Sigma2: financing High Performance Computing and storage

Data registration / access model of SecureB2SHARE in TSD

EUDAT Collaborative Data Infrastructure



TSD isolation & security

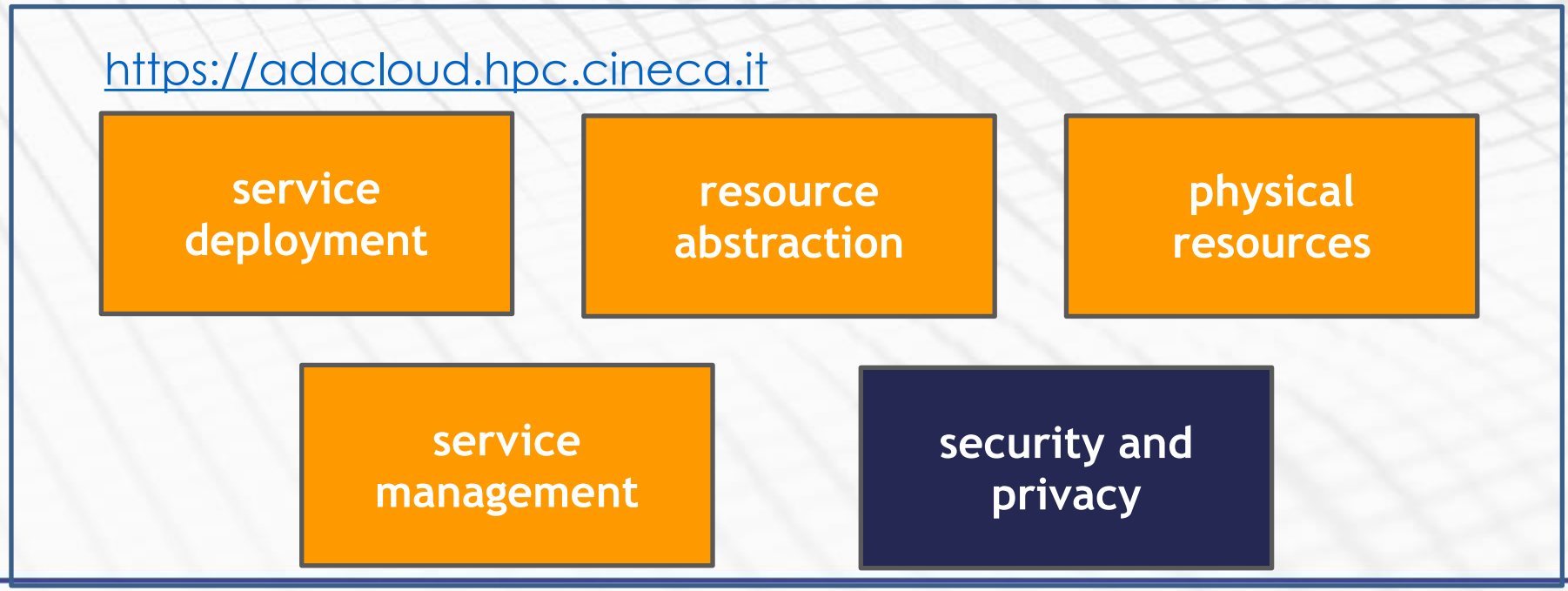
- All projects are isolated on network level
  - Assures no cross-flow of data between tenants
- Users work within the secure environment of their project
- All logins with enforced 2FA

SecureB2SHARE for TSD

- Datasets are registered by data owners via internal web portal in TSD, and assigned dataset identifiers
- The dataset files get archived in secure data storage
- B2SHARE metadata records are created based on registered information
- Access requests are submitted to TSD via Nettskjema (UiO developed solution for secure data collection) by following links from the B2SHARE records
- Requests get ingested by internal system, notifying data owners
- Data owners can then approve/deny the request via a web portal available within TSD
- If a request is approved, TSD's publication system is utilised to make the dataset available for the requestor via the external TSD Publication Portal

Sensitive Data Processing in CINECA

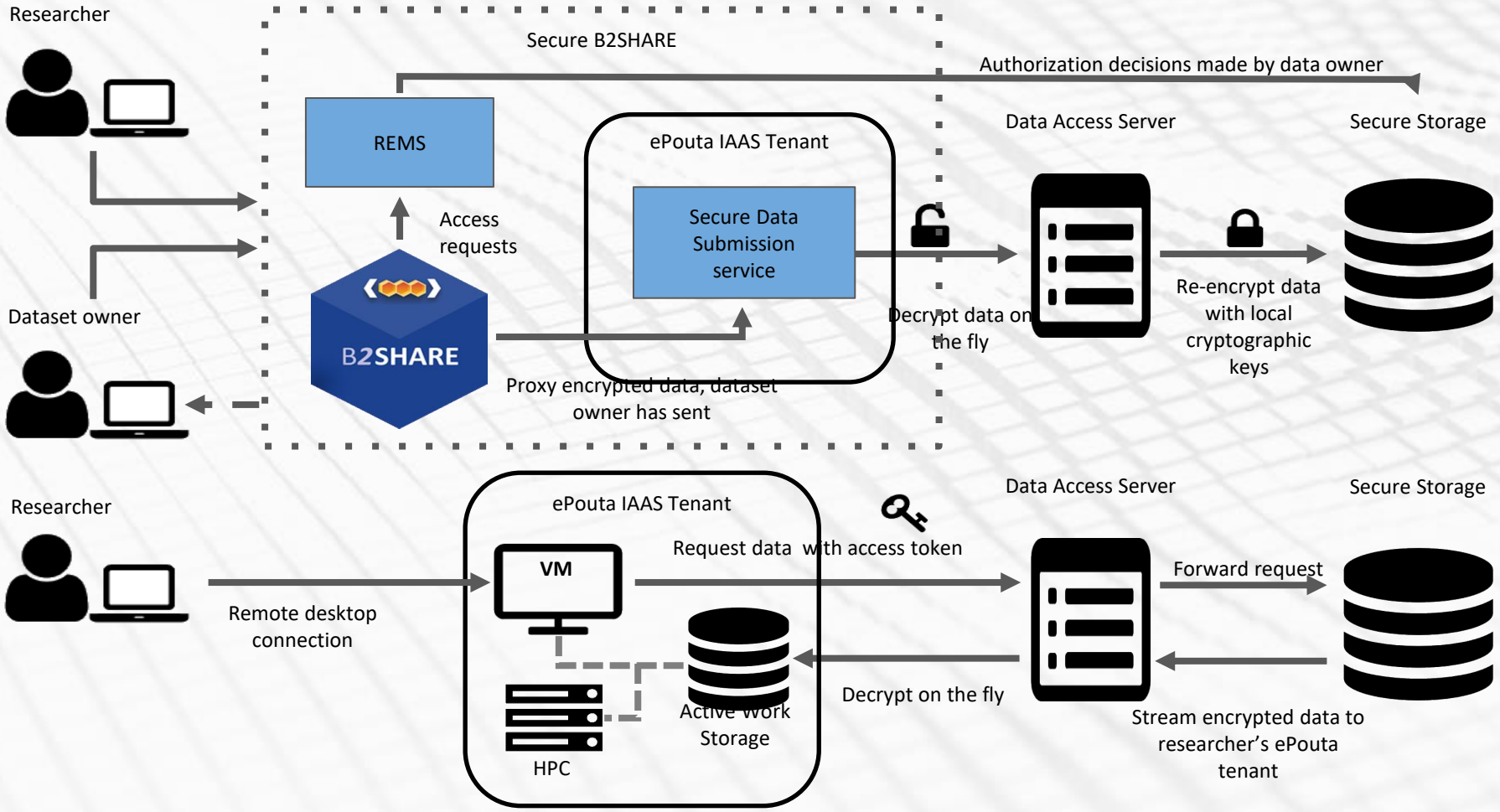
- Cineca relies for this service on its own Cloud HPC solution
  - Trade-off choice
    - Loose extreme power computing from traditional HPC systems (multi-user, multi-purpose)
    - Provide more control over data and access
  - Cineca is pursuing ISO27001 certification for HPC cloud services (audit end of 2022)



CINECA: Security and Privacy factors

Security factor	Description
Authentication and Authorization / Identity and Access Management	Done through an internal web tool (**UserDB**) integrated with the IdP (**keycloak** with 2FA)
	Done through **Openstack** cli/dashboard
Confidentiality, integrity and Availability	Cloud has a **dedicated and separated storage**, **Encryption, access control**, periodic/on demand vulnerability assessment penetration testing (**VAPT**)
Monitoring and Incident Response	**Cloud monitoring** tool, **network monitoring** tool. Incident management follows the **ISO27001 standard** procedures
Security Policy Management	Implemented through **ISO27001** procedures
Privacy-preservation	**Encryption, pseudonymization/anonymization**

SecureB2SHARE @ CSC (WIP)



Sensitive data use cases

- IAAS
  - Users build their own platform using services available on the system
  - Admin users prepare VM resources
  - They can leverage encrypted volumes

- PAAS
  - CINECA develops the platform using services available on the Cloud system

DICE
Data Infrastructure Capacity for EOSC

JOIN OUR COMMUNITY

DICEosc

/company/diceosc

?

Chris Ariyo
CSC – IT Center for Science, Finland