

ONE DATA

summer webinars 2026

Chapter 3: advanced data and replica management

Lukasz Opiola



15.07.2026

AGENDA & GOALS

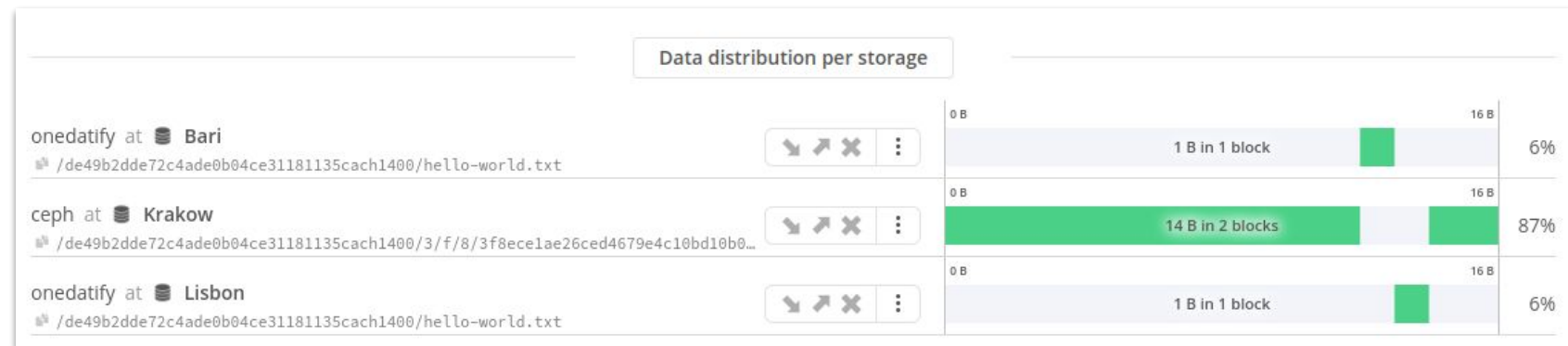
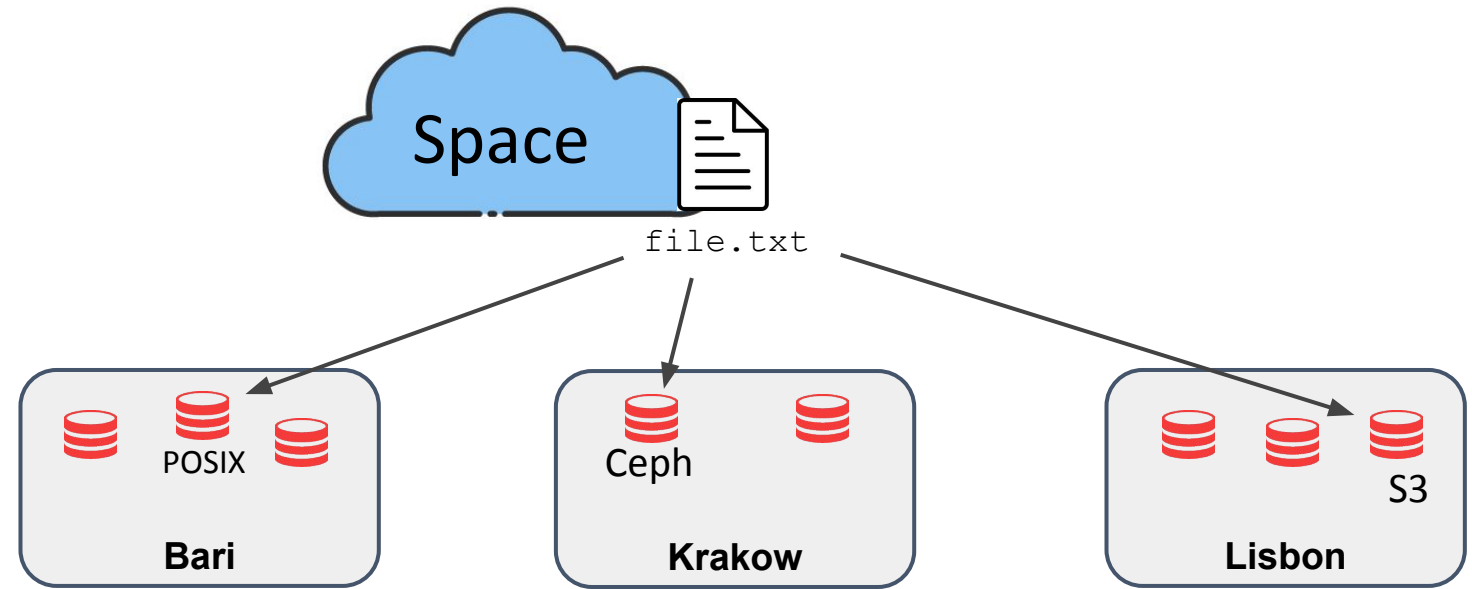
Onedata GUI — advanced data and replica management:

- Data distribution
- Transfers (replication, migration, eviction)
- On-the-fly data delivery
- Chunk-based data management
- QoS (replication rules)
- Directory size statistics
- Copies, symlinks, and hardlinks

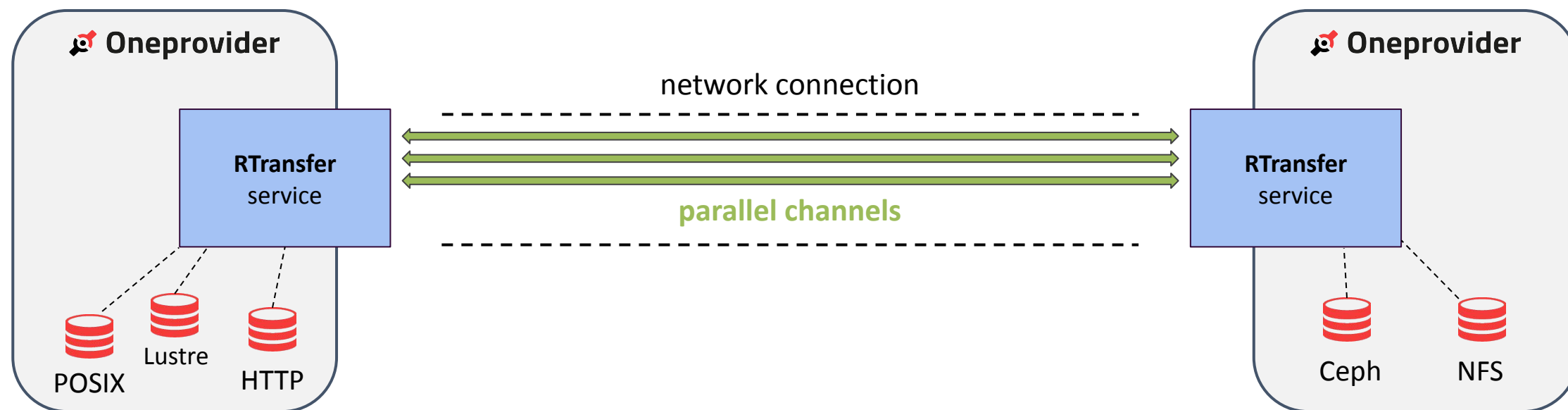


DATA DISTRIBUTION & REPLICA MANAGEMENT

- Chunk-based replica management
- Missing chunks delivered on-the-fly
- Handled automatically, but controllable manually -
prestage, transfer, migrate, define QoS rules
(policy-based replica mgmt)



THE RTRANSFER PROTOCOL

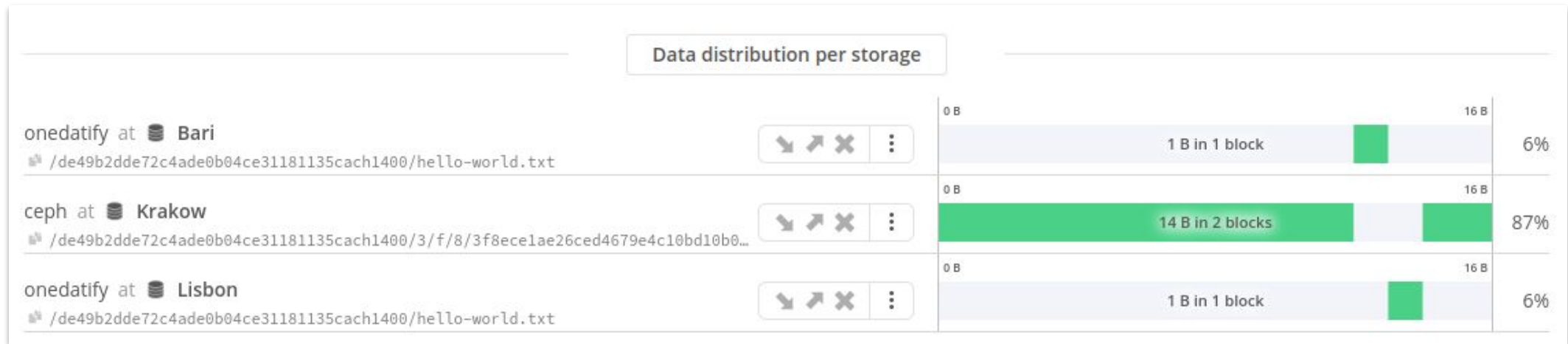


Web UI

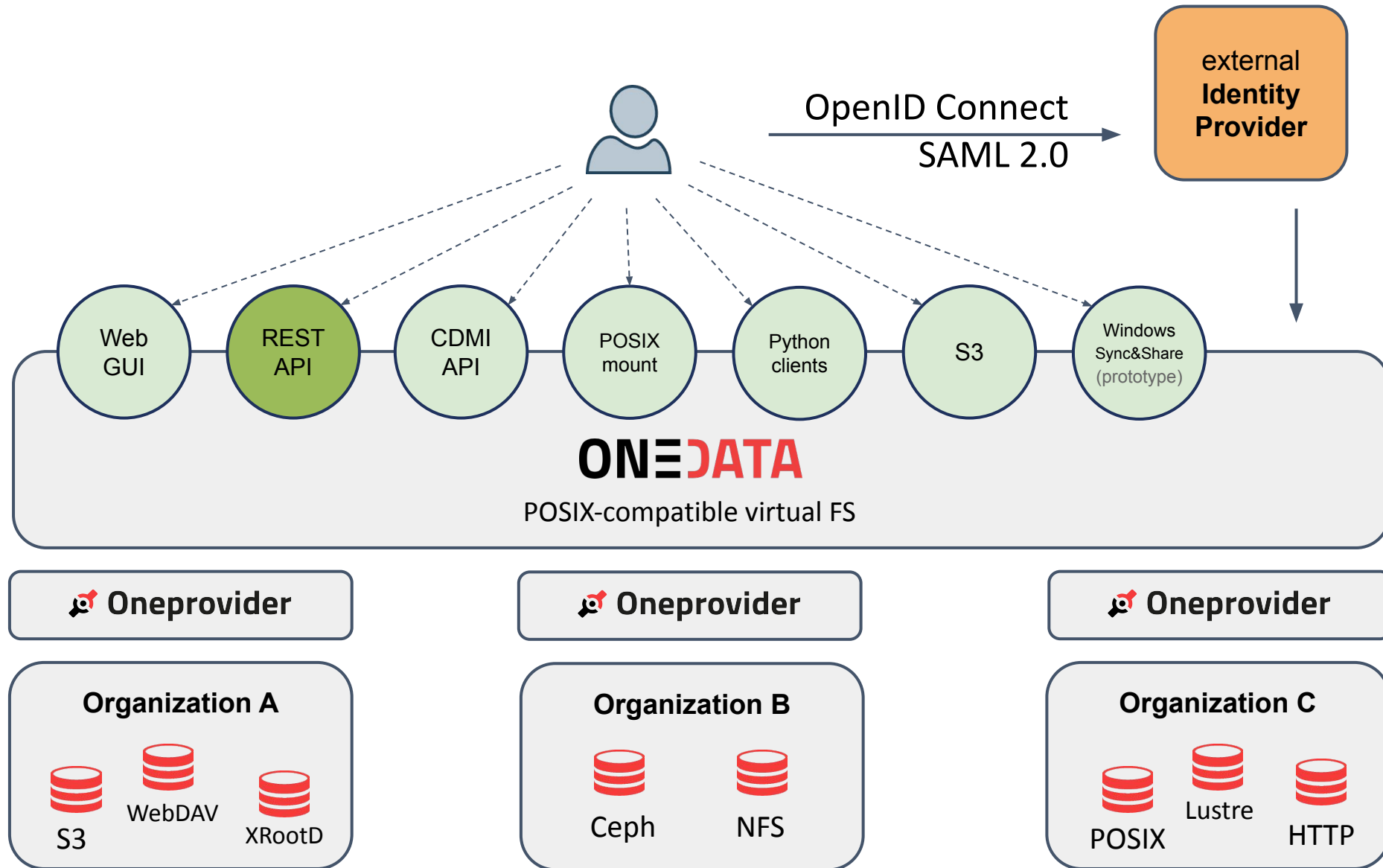
<https://onedata.cloud-esr.fr>

*data distribution,
transfers, on-the-fly delivery*

CHUNK-BASED REPLICA MANAGEMENT



VIRTUALIZED NAMESPACE, ABSTRACTED STORAGE



Web UI

<https://onedata.cloud-esr.fr>

chunk-based replica management

“the chessboard”

QUALITY OF SERVICE (QoS) — RULE-BASED REPLICA MANAGEMENT

A rule engine used to **automatically** manage file replica distribution and redundancy between providers supporting a space. QoS provides a **declarative** approach to replica management.

Users set **QoS requirements** on files or directories to express the expected replication state.

QoS REQUIREMENT

A **QoS requirement** is essentially a rule defining:

- the desired data redundancy (**replica count**),
- the placement of **replicas** (QoS expression).
- Storage backends matching the QoS expression are selected for data replication in a continuous re-evaluation process until the target replica count is satisfied. If there are more matching backends than the target replica count, a random subset is selected.

QoS EXPRESSION

- a **declarative way** of specifying desired storage parameters in a unified format,
 - used to **determine the replica placement** — the target storage backends where the replicas of the data will be stored,
 - refers to **storage backend parameters** that are assigned by Oneprovider admins.
-
- `geo=PL` — storage backends in Poland,
 - `timeout<8` — storage backends with timeout parameter set to less than 8,
 - `timeout=8` — storage backends with timeout parameter set to exactly 8,
 - `geo=PL & type=disk` — storage backends of disk type in Poland,
 - `geo=PL | type=disk` — storage backends in Poland or storage backends of disk type anywhere,
 - `anyStorage \ type=disk` — storage backends that are not of disk type.

Web UI

<https://onedata.cloud-esr.fr>

QoS — Quality of Service

Web UI

<https://onedata.cloud-esr.fr>

directory size statistics

copies, symlinks, and hardlinks

THANK YOU

Visit our website: <https://onedata.org>