CDS Migration

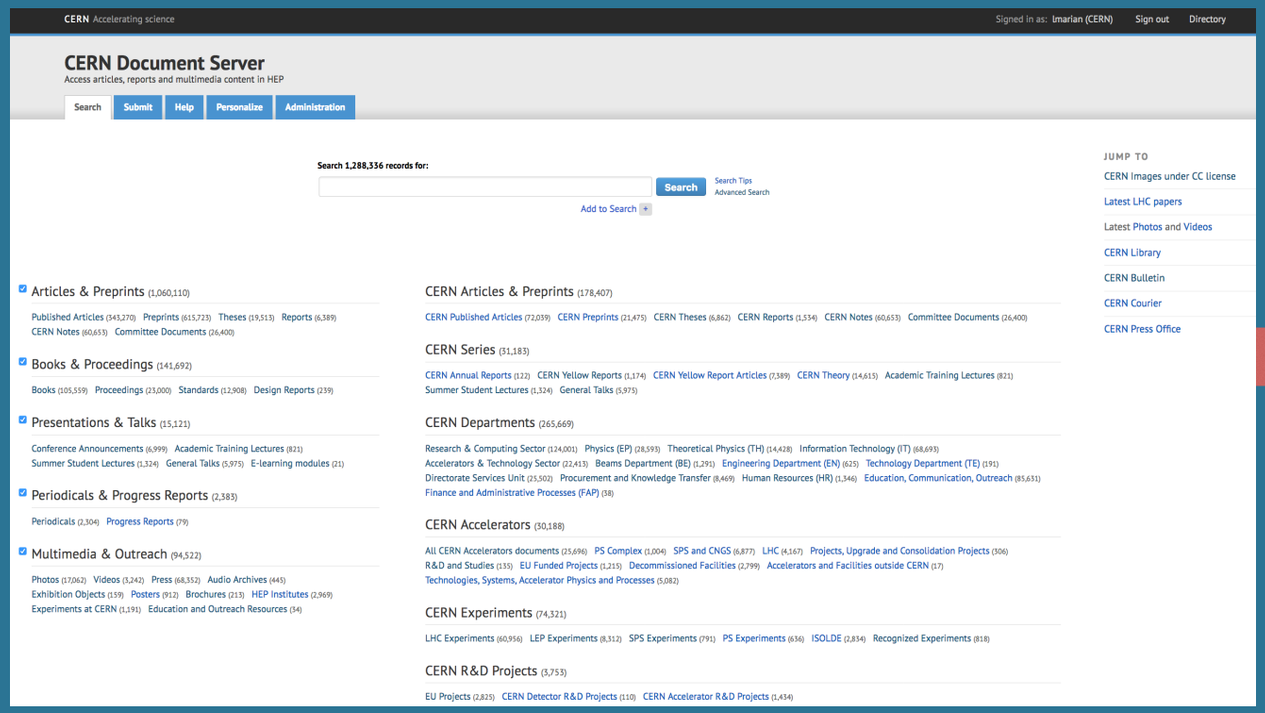
v1.2 to v3.0

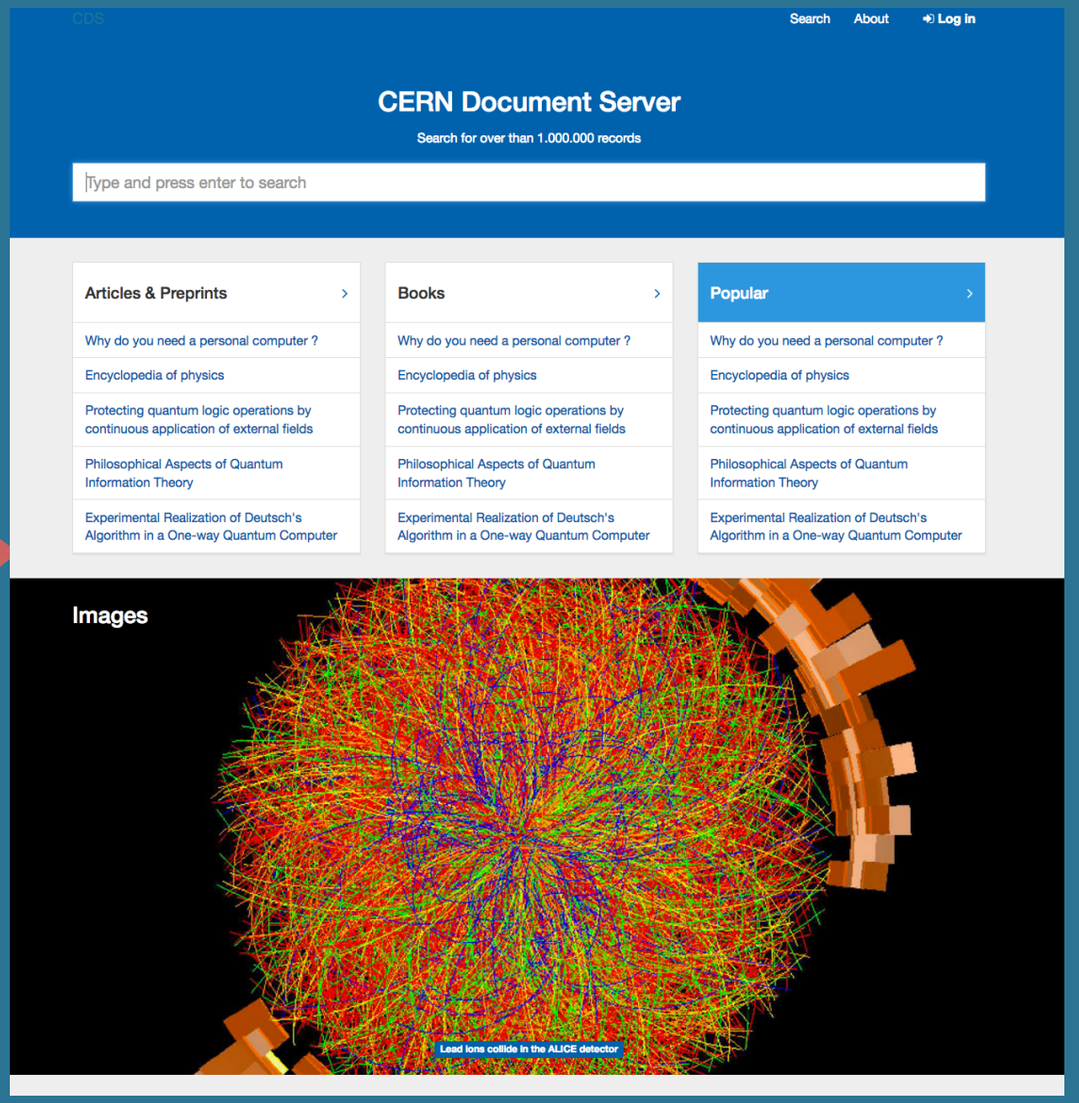
Ludmila Marian, on behalf of the CDS Team
CERN

INVENIO

MLZ
Heinz Maier-Leibnitz Zentrum

CDSLabs (base)





CDSLabs

CDS Migration in Numbers

- **1.2M** records (+ revisions, files, file revisions)
- one-off operation (no synchronisation between CDS and CDSLabs)

CDS Migration in Numbers

- ★ **1.2M** records (+ revisions, files, file revisions)
- ★ one-off operation (no synchronisation between CDS and CDSLabs)

Dump

- duration: a few hours
- data model: MARC21 (CERN version)
- 1 worker node
- out-of-the-box invenio-migrator

CDS Migration in Numbers

- ⭑ **1.2M** records (+ revisions, files, file revisions)
- ⭑ one-off operation (no synchronisation between CDS and CDSLabs)

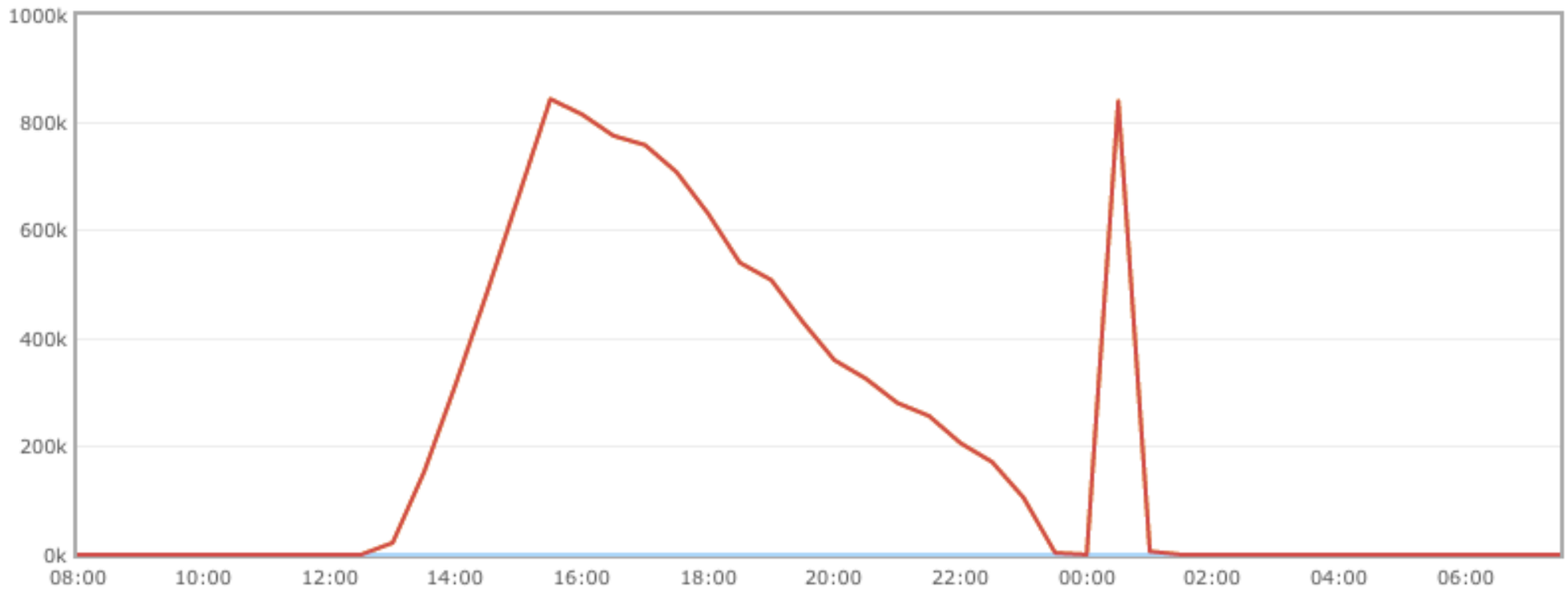
Dump

- duration: a few hours
- data model: MARC21 (CERN version)
- 1 worker node
- out-of-the-box invenio-migrator

Load

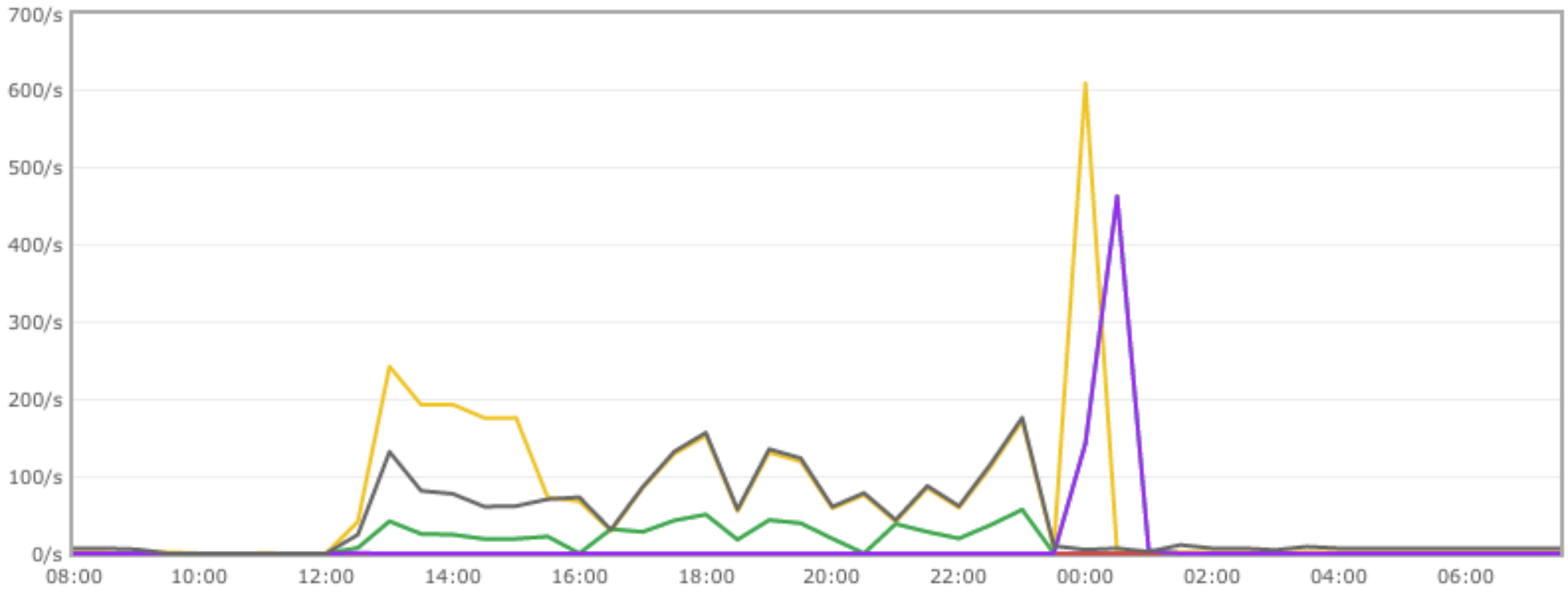
- duration: ~11 hours
- data model: custom (JSON-MARC21 standard + small customisations)
- 3 worker nodes (each running 4 parallel task)
- non-compliant past versions: kept as MARC in JSON

Queued messages (chart: last day) (?)



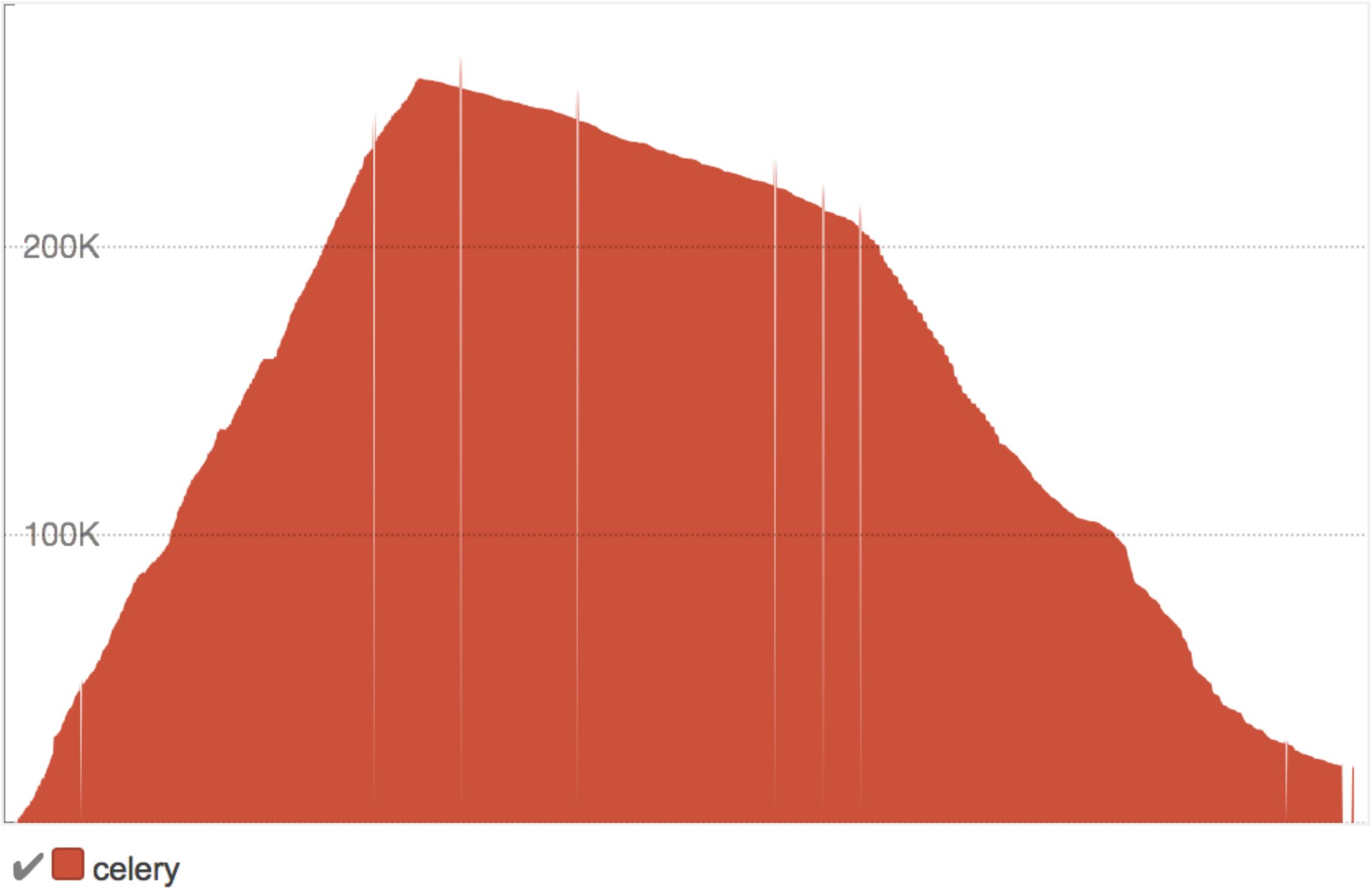
Ready	0 msg
Unacknowledged	0 msg
Total	0 msg

Message rates (chart: last day) (?)

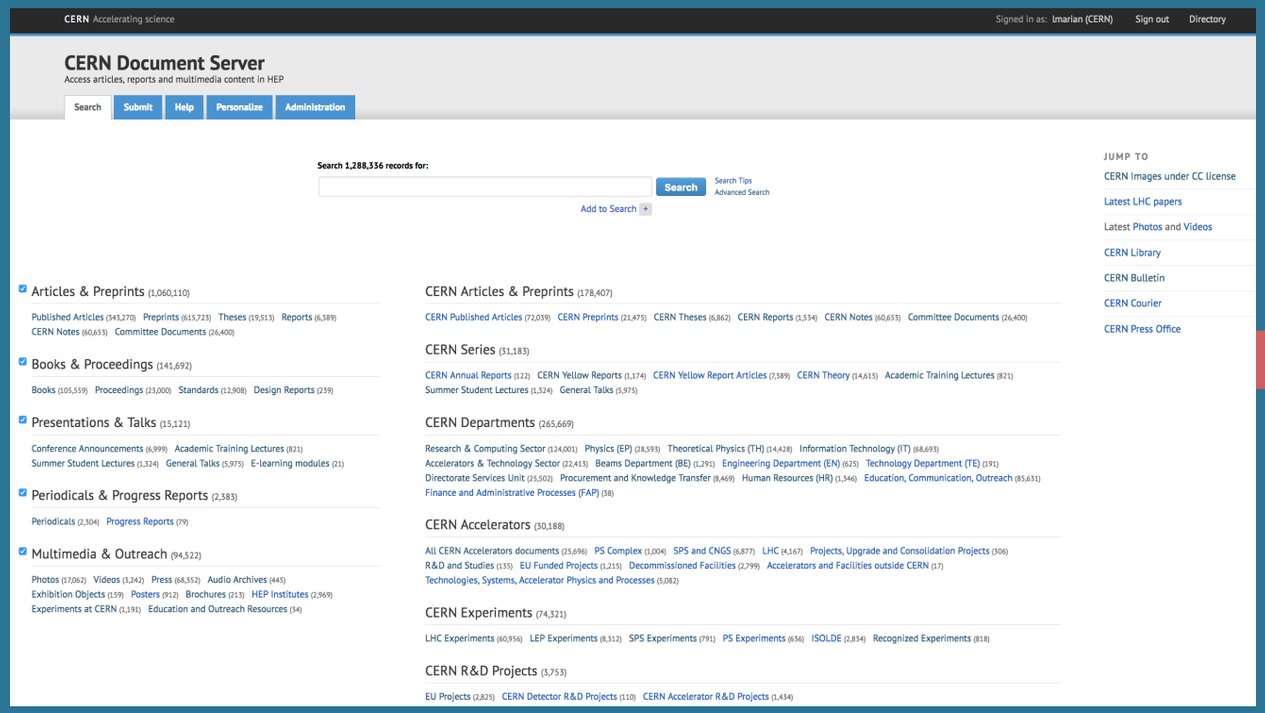


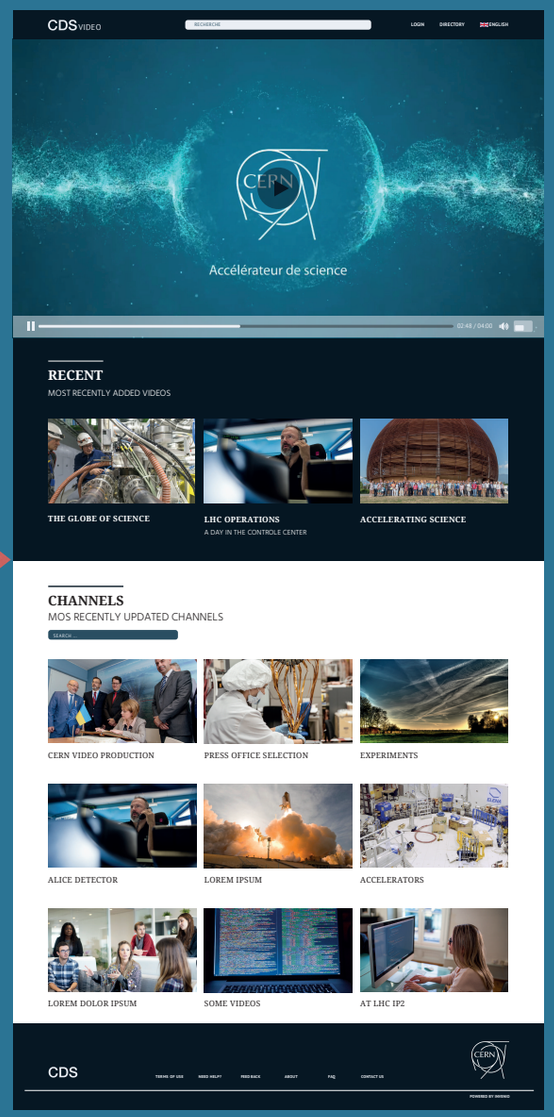
Publish	1.8/s
Deliver	0.00/s
Redelivered	0.00/s
Acknowledge	0.00/s
Get	0.00/s
Deliver (noack)	7.2/s

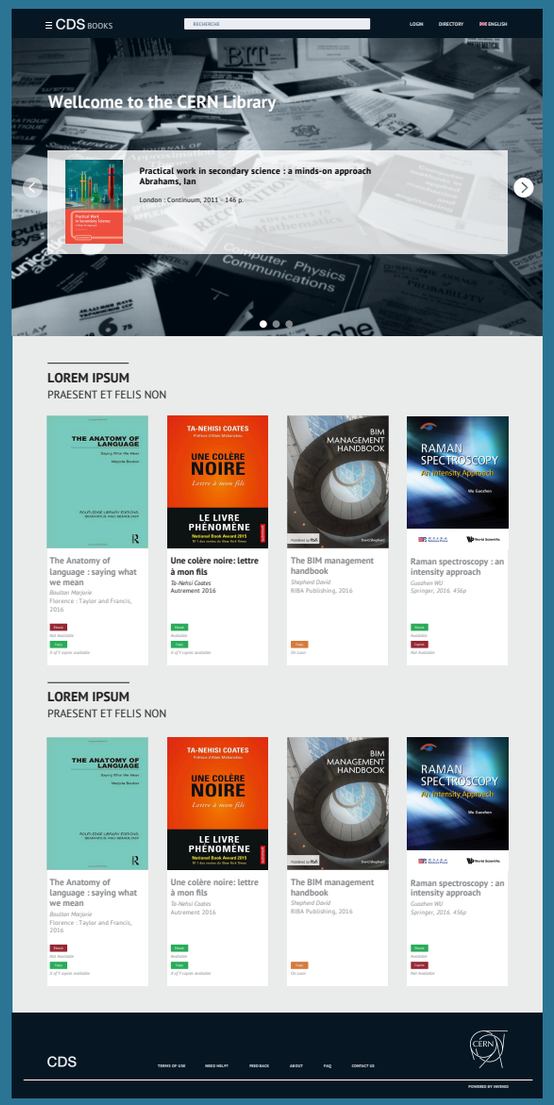
Queued tasks



CDS Migration Plan







CDS Migration Plan

- opportunity to better define our data models
- no synchronisation (except for records IDs)
- no actual migration of the current collection tree (>1’000 collections)
- no actual migration of the submission workflows (>100 workflows)

cds.cern.ch

cdslabs.cern.ch

/CERNDocumentServer/cds

cds-admin@cern.ch

INVENIO