

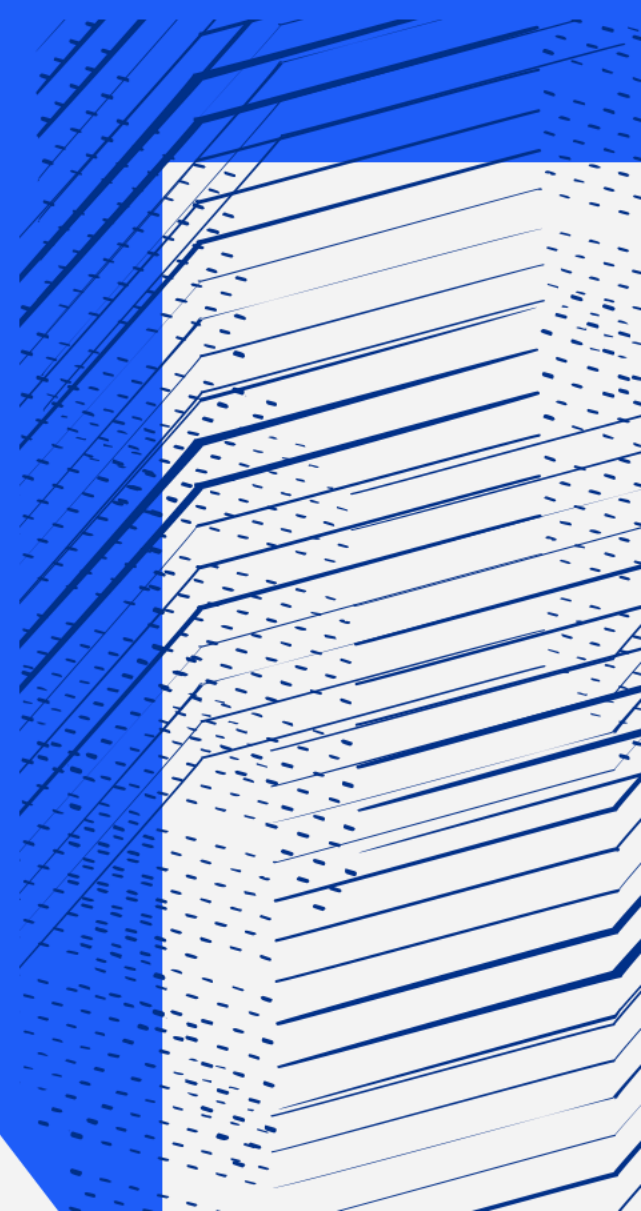


Science and
Technology
Facilities Council

Scientific Computing

On the use of XRootD in StorageD at the Rutherford Appleton Laboratory, UK

Emmanuel Bejide
and
Christopher Prosser



Agenda

1 Introduction to StorageD

2 Use of XRootD in StorageD.

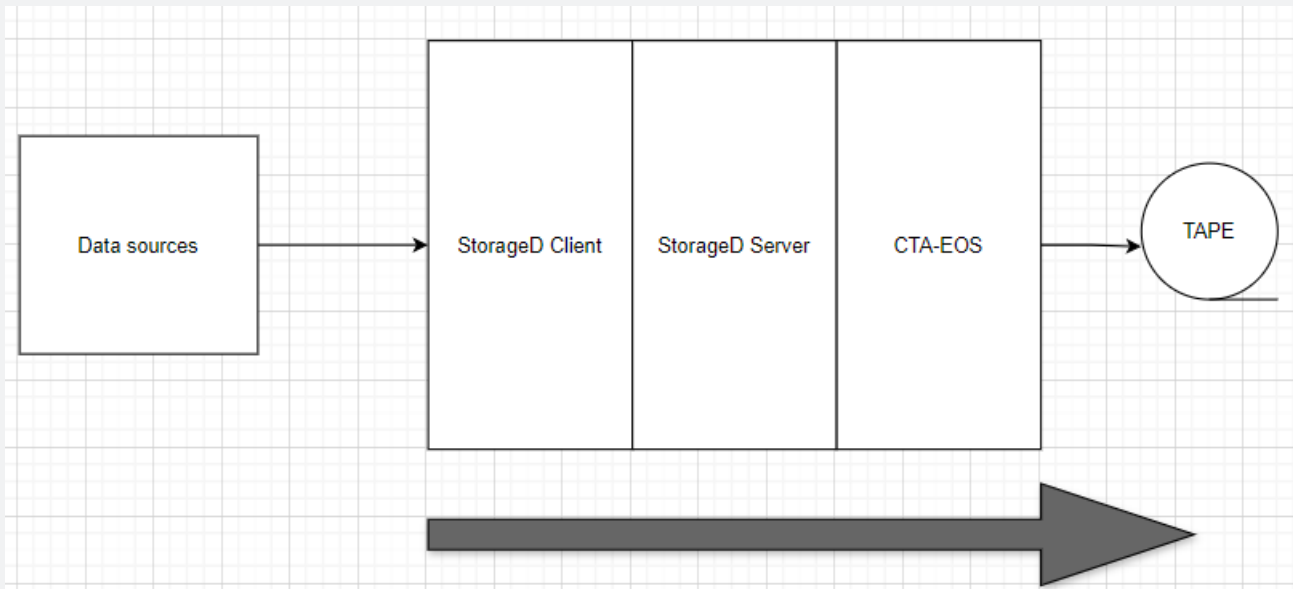
3 StorageD Management and Monitoring



Image © STFC Alan Ford

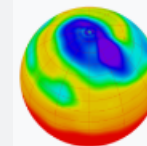
Introduction to StorageD

- StorageD is the data aggregator and archiving systems that supports the work of the Diamond Light Source (DLS) and the Centre for Environmental Data Analysis (CEDA) at the Rutherford Appleton Laboratory (RAL).



Diamond Light Source is the UK's national synchrotron. It works like a giant microscope, harnessing the power of electrons to produce bright light that scientists can use to study anything from fossils to jet engines to viruses and vaccines.

<https://www.diamond.ac.uk/Home/About.html>



**Centre for Environmental
Data Analysis**

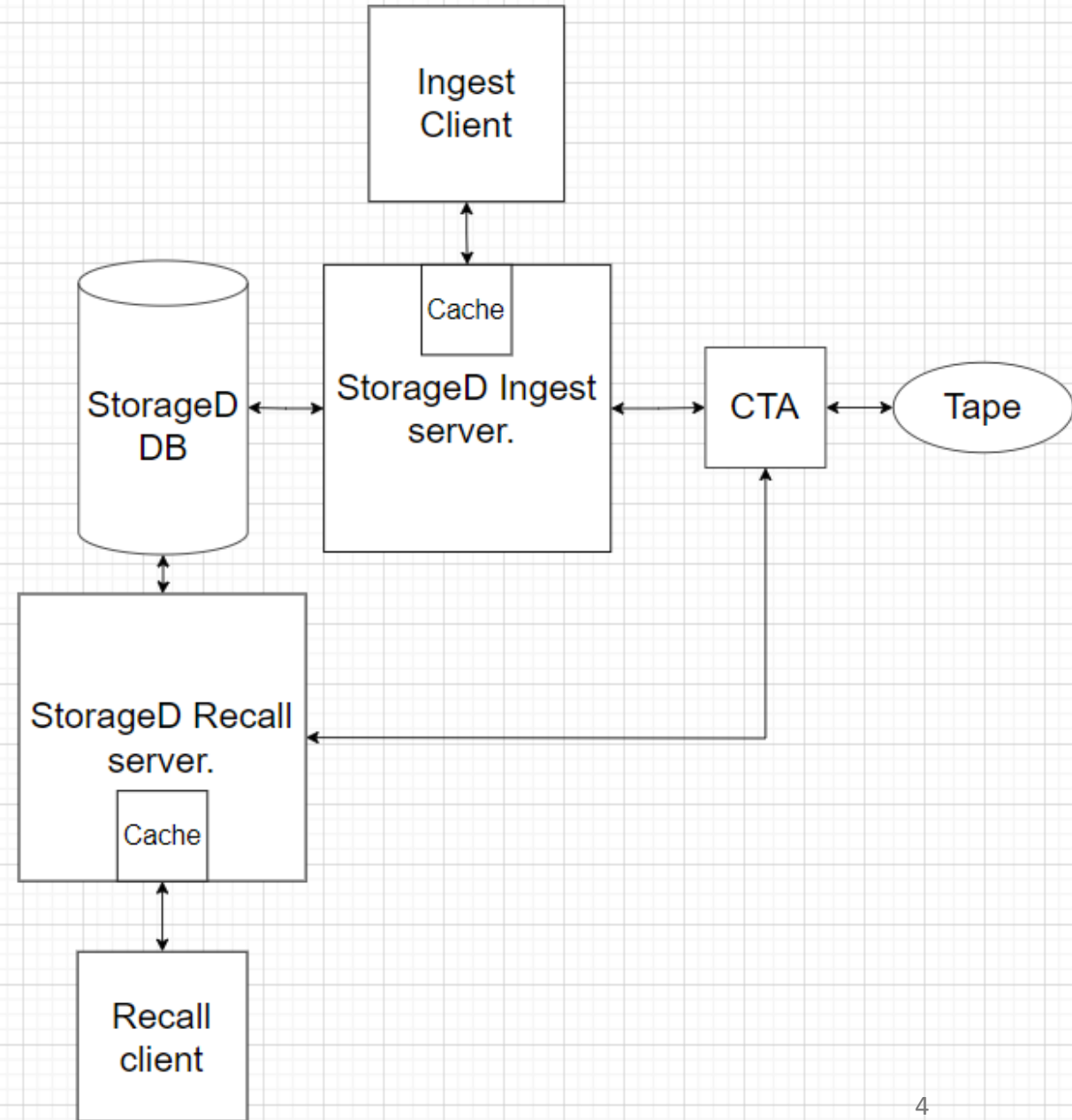
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

CEDA aims to support environmental science, further environmental data archival practices, and develop and deploy new technologies to enhance access to data.

<https://www.ceda.ac.uk/>

Introduction to StorageD

- Current production instance of StorageD works on single client-to-server connection



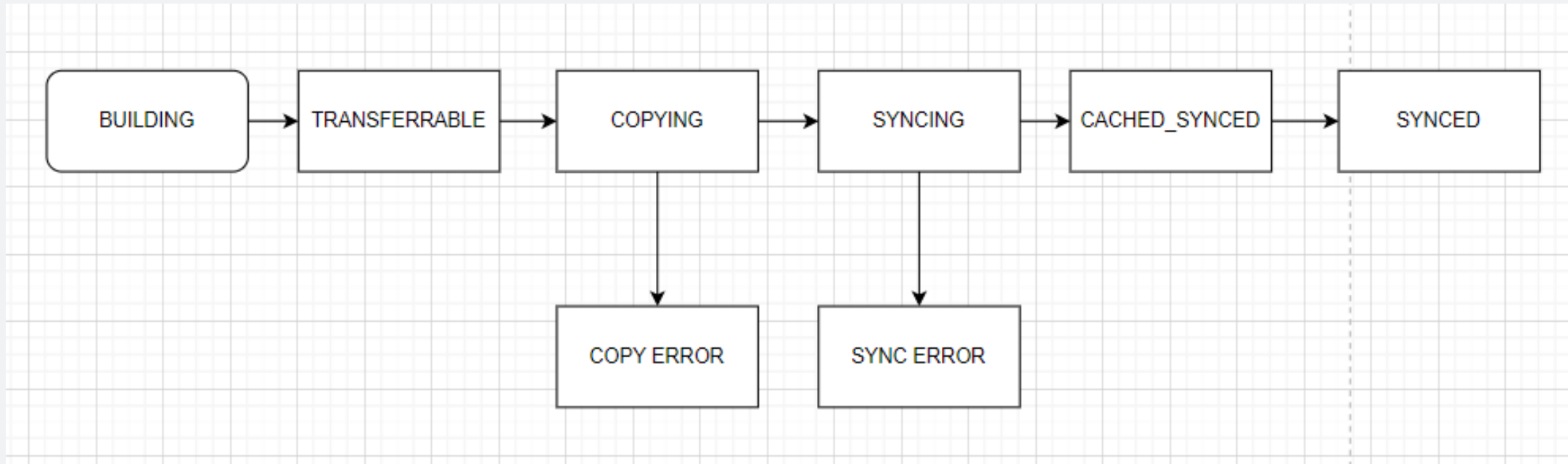
Introduction to StorageD (Aggregation)

	FILE_ID	TRANSFER_ID	FILE_NAME	FILE_SIZE	TRANSFER_ORDINAL	TRANSFER_OFFSET
1	2920836465	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50777.nxs	265735	1	0
2	2920836466	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50774.nxs	265735	2	265735
3	2920836467	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50774/i08-50774-PANDBOX.h5	1814350	3	531470
4	2920836468	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50778.nxs	265735	4	2345820
5	2920836469	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50791.nxs	265735	5	2611555
6	2920836470	162199497	/dls/i14/data/2019/cm22977-5/scan/i14-69614.nxs	286923	1	0
7	2920836471	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50782.nxs	265735	6	2877290
8	2920836472	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50782/i08-50782-PANDBOX.h5	6381192	7	3143025
9	2920836473	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50788.nxs	265735	8	9524217
10	2920836474	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50798.nxs	265735	9	9789952
11	2920836475	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50806.nxs	265735	10	10055687
12	2920836476	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50783.nxs	265735	11	10321422
13	2920836477	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50807.nxs	265735	12	10587157
14	2920836478	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50779.nxs	265735	13	10852892
15	2920836479	162199498	/dls/i11/data/2019/cm22960-5/869828-mac-001.raw	7441832	1	0
16	2920836480	162199498	/dls/i11/data/2019/cm22960-5/869817-mythen_1.raw	196252	2	7441832
17	2920836481	162199498	/dls/i11/data/2019/cm22960-5/869817-mythen_1.dat	337515	3	7638084
18	2920836482	162199498	/dls/i11/data/2019/cm22960-5/869817.dat	298	4	7975599
19	2920836483	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50805.nxs	265735	14	11118627
20	2920836484	162199497	/dls/i14/data/2019/cm22977-5/scan/i14-69613.nxs	286923	2	286923
21	2920836485	162199497	/dls/i14/data/2019/cm22977-5/scan/i14-69615.nxs	286923	3	573846
22	2920836486	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50804.nxs	265735	15	11384362
23	2920836487	162199497	/dls/i14/data/2019/cm22977-5/scan/i14-69616.nxs	286923	4	860769
24	2920836488	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50785.nxs	265735	16	11650097
25	2920836489	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50811.nxs	265735	17	11915832
26	2920836490	162199498	/dls/i11/data/2019/cm22960-5/869815-mythen_1.raw	198674	5	7975897
27	2920836491	162199498	/dls/i11/data/2019/cm22960-5/869815-mythen_1.dat	339639	6	8174571
28	2920836492	162199498	/dls/i11/data/2019/cm22960-5/869815.dat	298	7	8514210
29	2920836493	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50784.nxs	265735	18	12181567
30	2920836494	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50810.nxs	265735	19	12447302
31	2920836495	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50790.nxs	265735	20	12713037
32	2920836496	162199496	/dls/i08/data/2019/cm22973-5/nexus/i08-50790/i08-50790-PANDBOX.h5	112053429	21	12978772

Ingest Pipeline

- State transition

- Aggregates go through a series of states



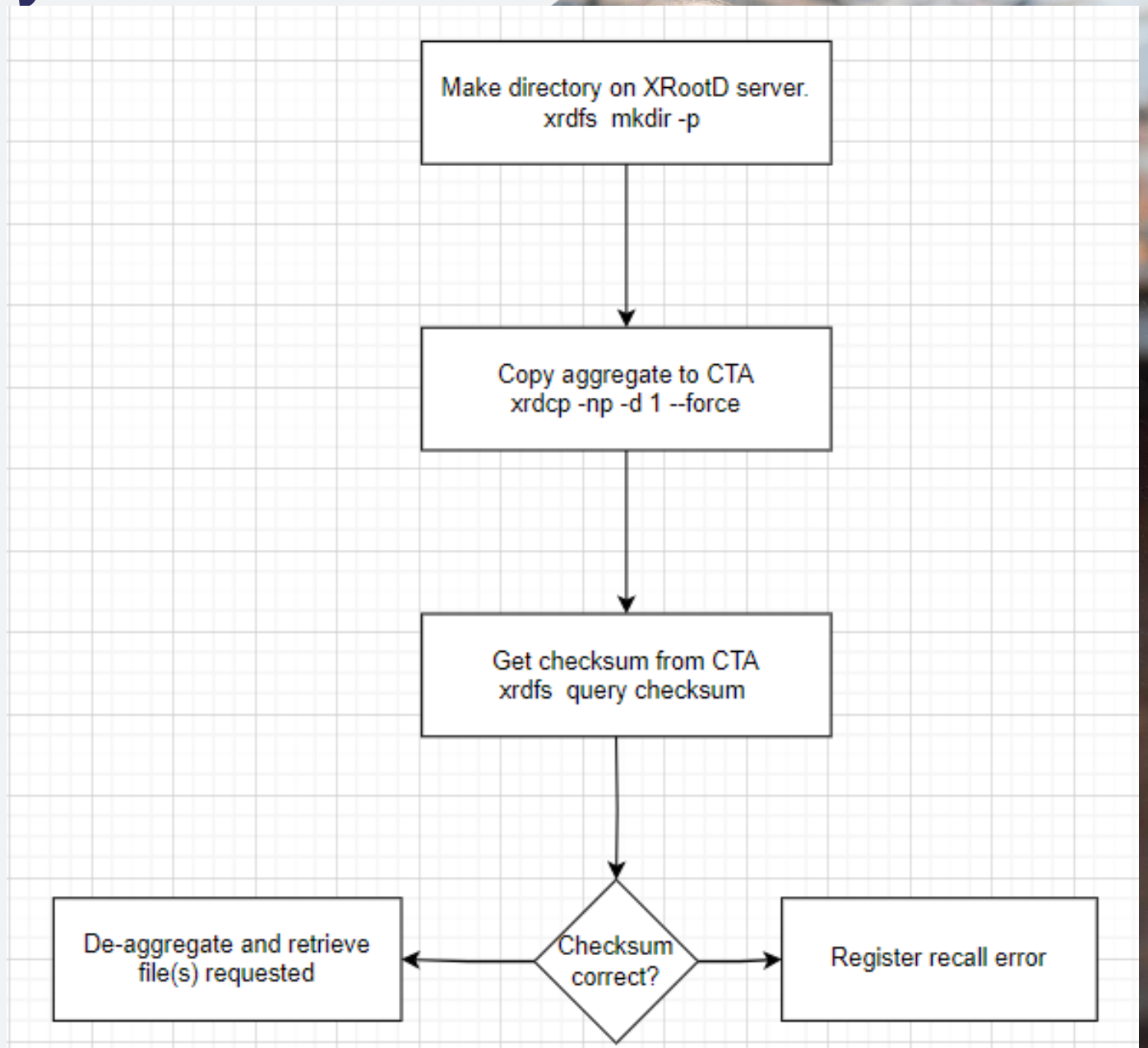
Ingest Pipeline

- Files registration (ADD_NEW_FILES)
 - Ingest is initiated by the client with a list of files
 - Files' offsets are calculated.
 - Files' information are sent by the client to the server.
 - Protobuf communication
 - The server registers files' information on the database.

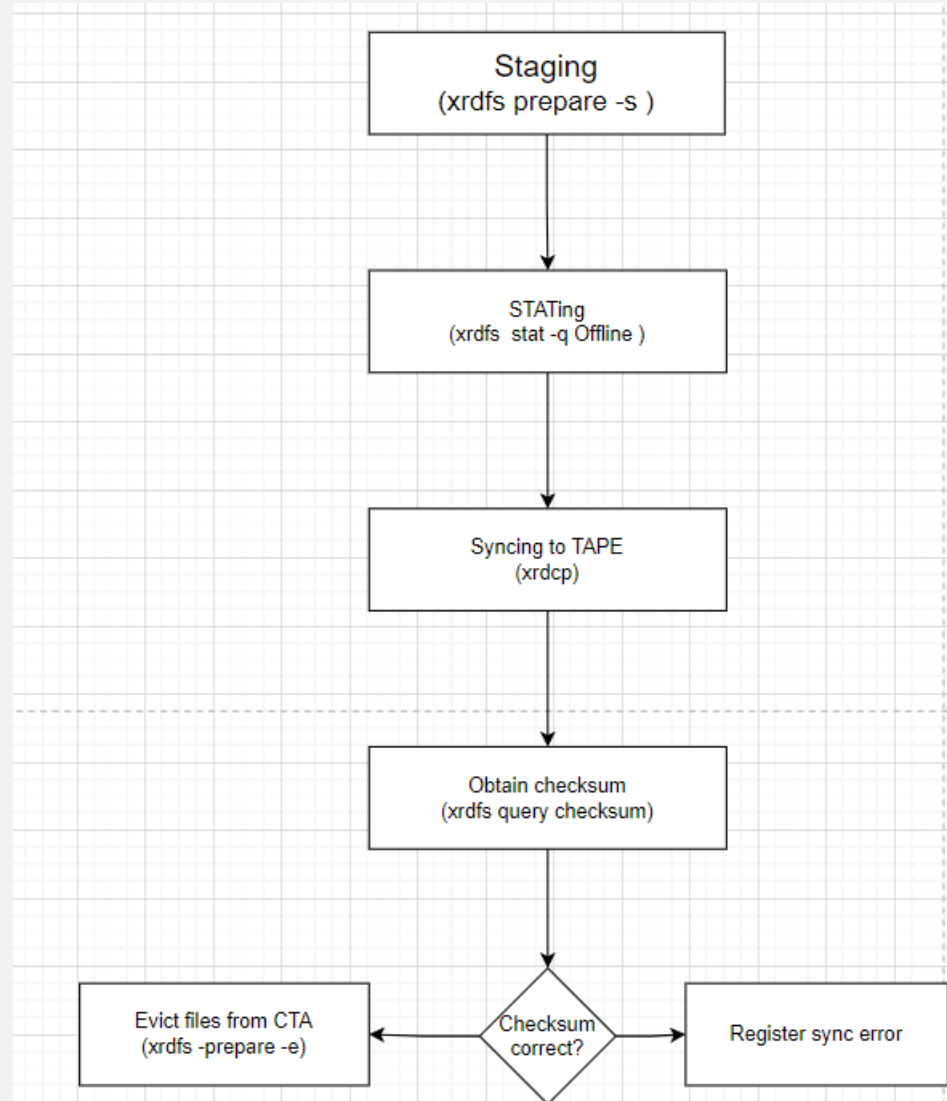
Ingest Pipeline

- Files transfer to TAPE (GET_TRANSFERRABLES)
 - At random times, the client asks the server for next transferrable
 - Protobuf communication
 - Aggregate copied to server's cache.
 - Checksums compared
 - Aggregate copied from server's cache to TAPE (CTA-EOS)
 - This is where we use XrootD.
 - Checksum compared
 - Ingest is completed.

Ingest Pipeline - sync to TAPE



Recall Pipeline – The role of XRootD





Science and
Technology
Facilities Council

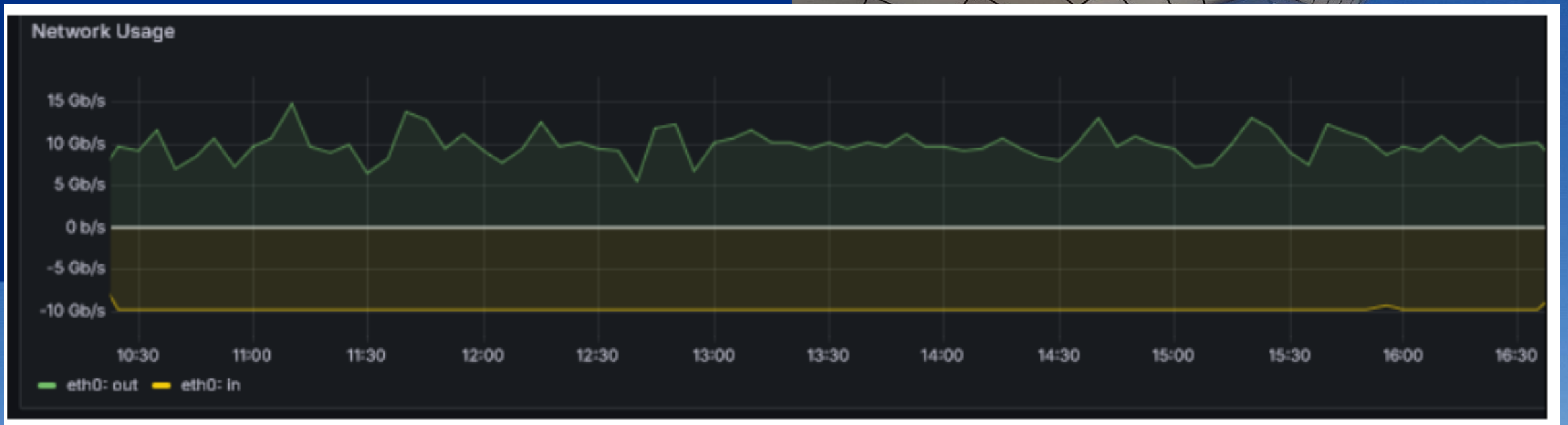
Scientific Computing

Management and monitoring of StorageD

Management of StorageD

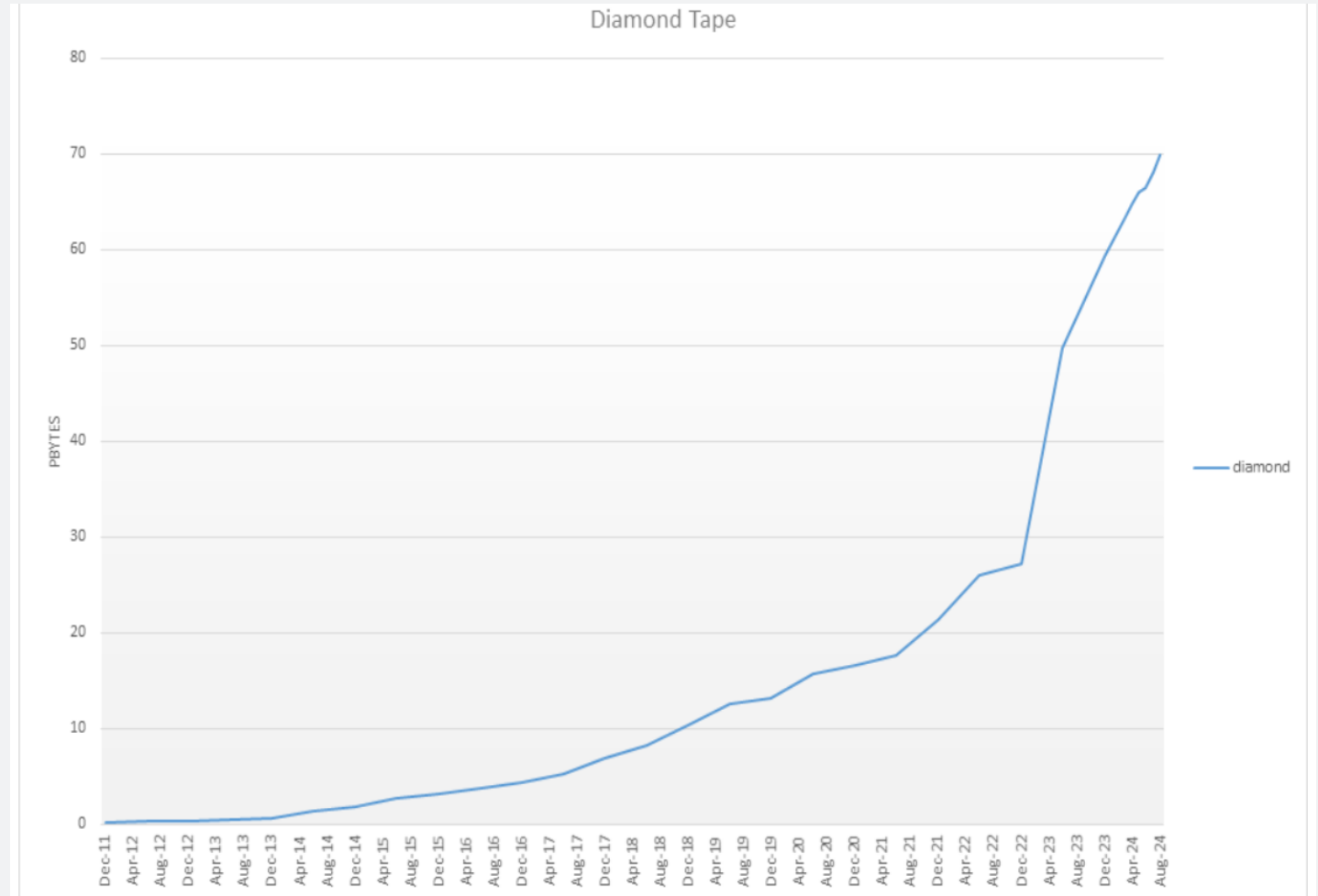
- Throughput

- this is not the limit of SD but the line rate of the client that's reading the data and is sufficient at the moment ... but we plan to scale it up.



Management of StorageD

- We have just over 70PB on tape / 5 Billion files
- We typically ingest at a rate of 10Gb/s at the moment which equates to 100TB in 24 hours.



Monitoring of StorageD

– journey of an aggregate to TAPE

DLS Aggregation Details - 247738285 - produced 2024-09-12 17:43:21

Aggregation data summary							
State	File count	Size (Gb)	Size (Bytes)	First file Loaded	Last file Loaded	Checksum	Time to TAPE
CACHED_SYNCED	1000	.94	941547089	2024-09-06 10:08:06	2024-09-06 10:08:06	3037ed24	2024-09-06 10:38:36

Aggregation History		
Initial State	Final State	Transition at
CREATION	BUILDING	2024-09-06 10:08:05
BUILDING	CLOSED	2024-09-06 10:38:09
CLOSED	TRANSFERABLE	2024-09-06 10:38:09
TRANSFERABLE	COPYING	2024-09-06 10:38:09
COPYING	CACHED	2024-09-06 10:38:32
CACHED	SYNCING	2024-09-06 10:38:34
SYNCING	CACHED_SYNCED	2024-09-06 10:38:36
CACHED_SYNCED	FLUSHING	2024-09-06 10:40:02
FLUSHING	SYNCED	2024-09-06 10:40:03
RETRIEVING	CACHED_SYNCED	2024-09-12 16:03:43
SYNCED	RETRIEVING	2024-09-12 16:03:43

Monitoring of StorageD

– journey of an aggregate to TAPE

Data throughput in last 24 hours

StorageD Data Collected (Gb)	StorageD File count	ICAT Data Collected (Gb)	ICAT File count
37614.83	415219	38063.25	415718

Data Ingested in Last 24hrs (by hour)

Data throughput in last day by hour

Month	Day	Hour	StorageD Data Input (Mb)	StorageD File Count	ICAT Data Input (Mb)	ICAT File Count
09	12	22	44755.15	113	225116.74	247
09	12	21	461635.89	478	496562.34	519
09	12	20	464863.04	535	458930.68	558
09	12	19	592864.1	1043	596851.73	1315
09	12	18	586928.86	1150	2747393.23	16627
09	12	17	252703.83	814	4339385.46	49723
09	12	16	927242.97	22437	4875808.12	89614
09	12	15	7010326.58	130514	2441699.47	45093
09	12	14	2596181.88	57193	3486098.86	54745
09	12	13	3039633.02	60175	5320563.71	100756
09	12	12	3001690.17	76762	3895448.32	7379
09	12	11	240666.75	807	4611771.52	25956
09	12	10	15504066.77	55292	1210164.76	16998
09	12	09	214765.46	2816	294921.68	649
09	12	08	245944.1	483	175021.15	408
09	12	07	214469.77	490	268581.63	568
09	12	06	209681.17	441	164252.47	383
09	12	05	205020.38	417	174281.94	301
09	12	04	180333.65	471	218675.49	672
09	12	03	177265	523	154876.18	447
09	12	02	181372.11	479	230077.26	611
09	12	01	244167.07	589	296495.77	491
09	12	00	478583.67	591	538233.66	746
09	11	23	539668.48	606	440290.28	488
09	11	22	n/a	n/a	401742.84	424



Science and
Technology
Facilities Council

Scientific Computing

Future Plans

Future Plans

- Rework staging on CTA-EOS
- Distributed StorageD
 - Flexible architectures
 - Scheduling and resource management
- Improve the dashboard
- Error handling
- Security



Science and
Technology
Facilities Council

Scientific Computing

Thank You