

# Overview description of Facilities Archive and Recall for CEDA and DLS

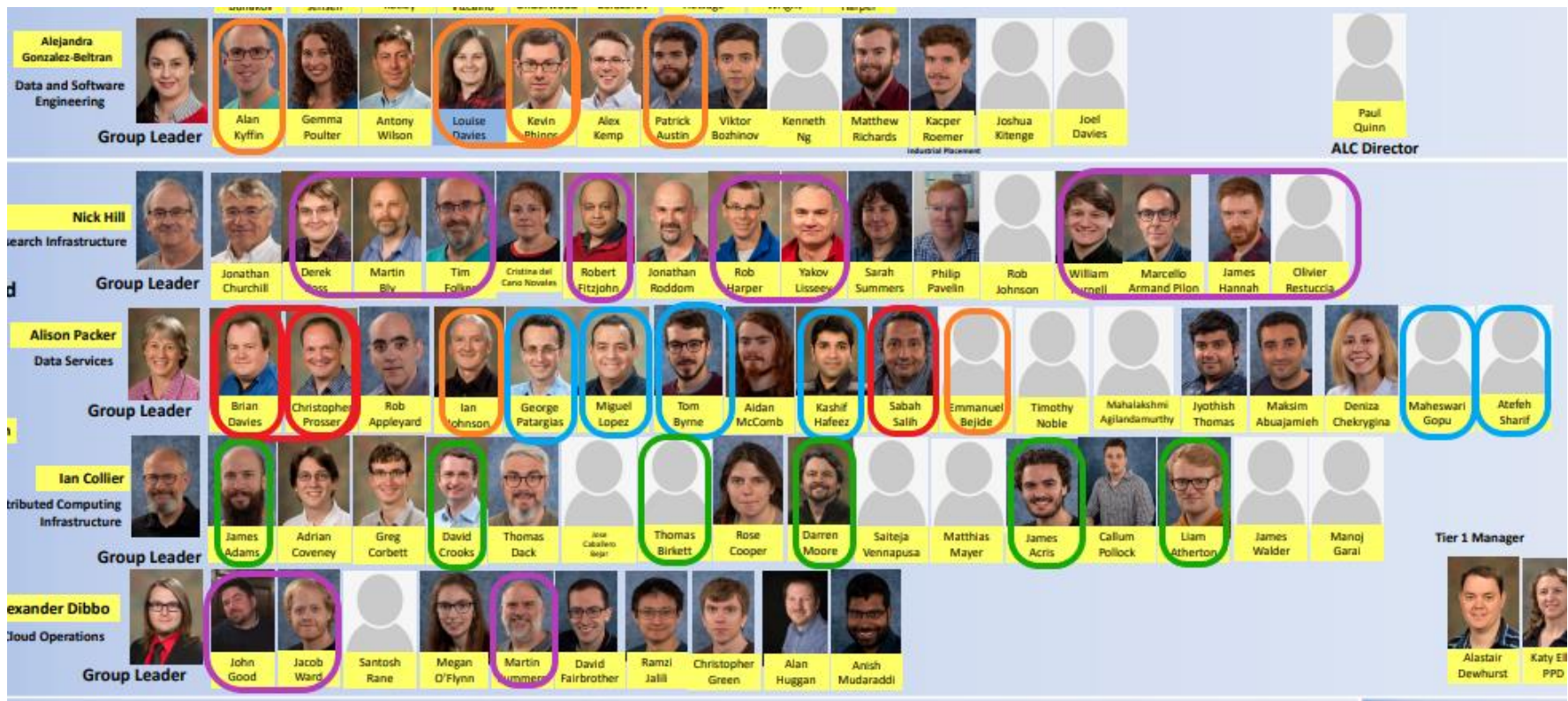
SCD AHM May 2024

# Service offering

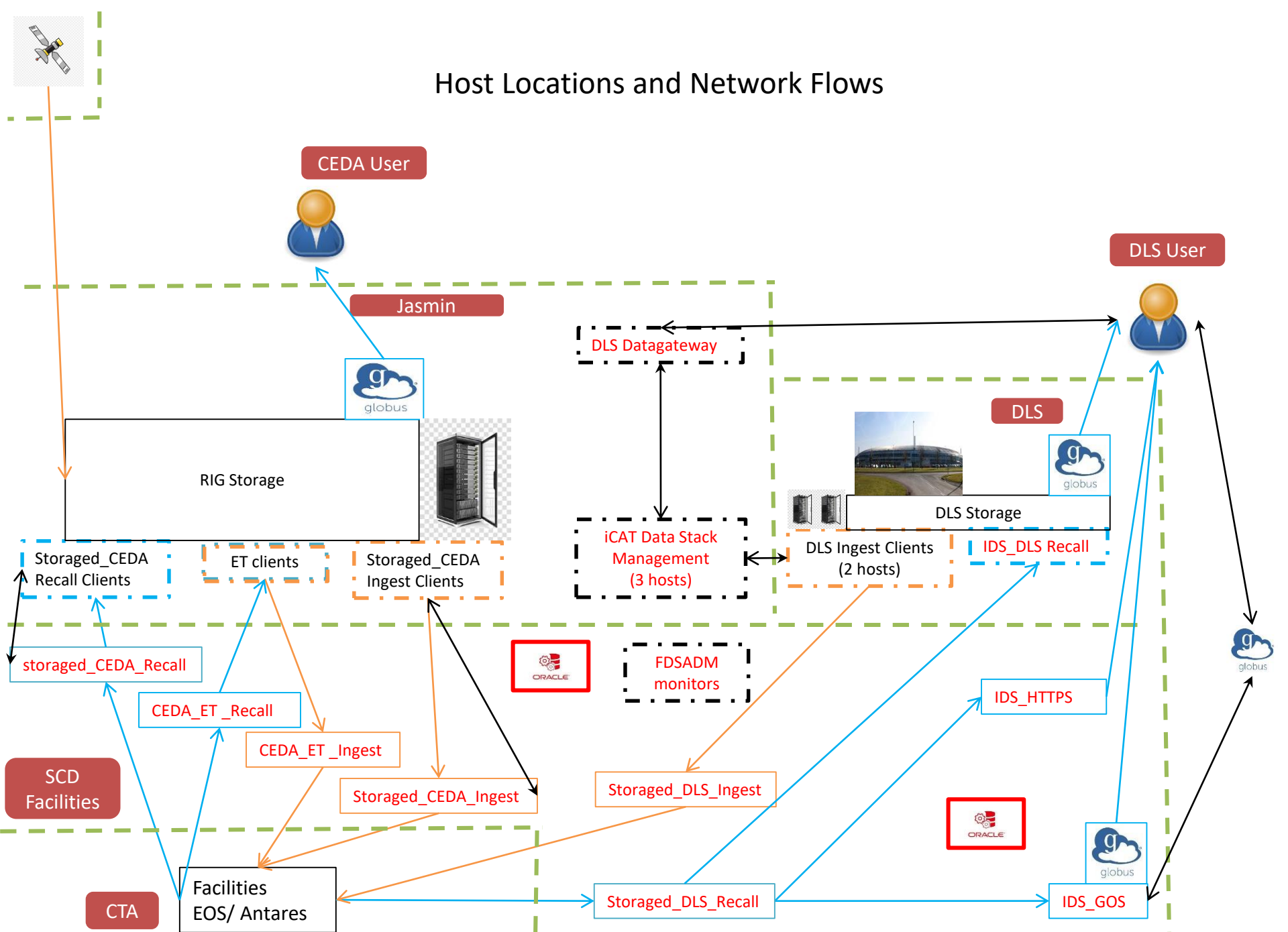
- Archive Storage For Both DLS and CEDA
  - (Akin to WLCG Tier-0)
  - Data Management
    - (ICAT/DATAGATEWAY)
  - Data movement
    - (StorageD/ET/IDS/Globus)
      - Aggregation via StorageD component
      - ET “ElasticTape”
- Ingest from both onsite and offsite
- Provide Recall of Data to offsite locations
  - Small and Large volumes
- Provide Recall of Data to Onsite Facilities
  - SCARF/Jasmin, Openstack Cloud, DLS work areas

# Many People Involved

( plus DLS (~7) personnel and  
CEDA(RALSspace) (~5) personnel )



# Host Locations and Network Flows



# Diamond Ingest and Recall Example

## DLS View

Village -> Instrument-> Visit->Dataset-> Files

The screenshot shows the Diamond Instruments website. The navigation bar includes 'diamond', 'About Us', 'For Users', 'Industry', 'Public', 'Science', 'Instruments', 'Careers', and 'More'. The main content area is titled 'Instruments by Science Group' and features icons for various techniques: Macromolecular Crystallography, Soft Condensed Matter, Imaging and Metrology, Biological Crystallography, Magnetic Materials, Structures and Surfaces, Crystallography, and Spectroscopy. Below this is a 'Quick links' section with a grid of instrument IDs. The 'All Instruments' section is active, showing a filter button and a list of 43 instruments. Three instrument cards are visible: 102-1 VMXm (Commissioning), 102-2 VMXI (Operational in optimisation mode), and 103 MX (High throughput and highly automated). Each card provides technical specifications and lists associated techniques.

## SCD view

- Set of Files in a directory on a Filesystem-> 1 Aggregate uploaded into Antares
- Contents of a Visit goes into a single EOS directory
- Set of directories related to a Village goes into single TapePool
- Closed/Open Visits

# Examples of Ingest Monitoring Pages (DLS)



## Summary of current state

StorageD current state summary

Metric	Topic	Value
Last state change at	.	2024-04-09 09:34:32
Throughput: files in last 24 hrs	File count: 1544	Tb registered: 0.01
Throughput: transfers to TAPE in last 24 hrs	Transfer count: 38	0.01Tb
Transfer state = <a href="#">CACHED_SYNCED:130.246.184.122</a>	total count: 396 transfers; 669426 files	3101.48Gb
Transfer state = <a href="#">DELETED:DEFAULT</a>	total count: 2 transfers; 2 files	1134.14Gb
Transfer state = <a href="#">FLUSHING:130.246.184.122</a>	total count: 1 transfers; 73 files	5.11Gb
Transfer state = <a href="#">SYNCED:DEFAULT</a>	total count: 10391928 transfers; 4776685496 files	64118096.88Gb

## Dropfile queue(s)

...and getting that info took 0 seconds

## Current Transfer Errors

Complete history - errored transfer state changes

Beamline	Visit	Transfer ID	File count	Transfer size	Current state	Last error report time	Last error type
----------	-------	-------------	------------	---------------	---------------	------------------------	-----------------

Data throughput in last 24 hours

StorageD Data Collected (Gb)	StorageD File count	ICAT Data Collected (Gb)	ICAT File count
10.76	1544	12.08	1545

Data registered in last 24hrs by beamline/visit

Station	Visit	StorageD Data Collected (Gb)	StorageD File Count	StorageD Last File	ICAT Data Collected (Gb)	ICAT File Count	ICAT Last File
B18	<a href="#">cm37265-2</a>	0	2	2024-04-08 14:49:50	0	2	2024-04-08 15:32:43
I05-1	<a href="#">cm37239-2</a>	.02	12	2024-04-08 13:51:41	.02	12	2024-04-08 14:29:19
I06	<a href="#">cm37240-2</a>	2.75	46	2024-04-08 17:05:32	2.62	46	2024-04-08 17:39:41
I06-2	<a href="#">mm34486-1</a>	4.71	122	2024-04-08 19:42:12	4.49	122	2024-04-08 20:28:56
I07	<a href="#">cm37245-2</a>	0	6	2024-04-08 10:54:41	0	6	2024-04-08 11:40:02
I10	<a href="#">cm37251-2</a>	0	2	2024-04-08 14:10:12	0	2	2024-04-08 14:50:25
I10-1	<a href="#">cm37252-2</a>	0	1	2024-04-08 11:08:12	0	1	2024-04-08 11:40:04
I13	<a href="#">cm37258-2</a>	.14	2	2024-04-08 23:42:27	.14	2	2024-04-09 00:21:38
I14	<a href="#">cm37259-2</a>	0	5	2024-04-08 12:21:29	1.33	6	2024-04-08 13:04:41
I15	<a href="#">cm37260-2</a>	.41	17	2024-04-08 17:21:33	.39	17	2024-04-08 18:00:49
I18	<a href="#">cm37264-2</a>	.03	50	2024-04-08 16:21:38	.03	50	2024-04-08 16:57:20
I19-1	<a href="#">cm37266-2</a>	3.11	1250	2024-04-08 12:44:15	2.97	1250	2024-04-08 13:26:08
I20	<a href="#">cm37268-2</a>	.09	27	2024-04-08 16:20:47	.09	27	2024-04-08 16:57:23
I21	<a href="#">cm37269-2</a>	0	2	2024-04-08 15:13:22	0	2	2024-04-08 15:53:55

# Examples of recall Monitoring Pages (DLS)

Admin

Maintenance Admin Download

ID	Full Name	Username	Prepared ID	Access Method	Availability	Size	Requested Date	Deleted
13796907			e5f15361-3672-43a4-96...	https	Restoring from tape	16.19 MB	2024-04-09 09:39:12	No
13796163			30866090-3113-4a3e-a7...	https	Available	49.94 MB	2024-04-09 21:52:08	No
13789198			4cc41736-9922-413b-bc...	ds	Available	69.24 GB	2024-04-08 20:17:56	No
13768780			c71d7528-be50-4d67-ab...	ds	Available	111.05 GB	2024-04-08 19:27:22	No
13755725			9036a3ca-a572-4925-8b...	https	Available	10.91 GB	2024-04-08 18:28:25	No
13757004			1fb0328c-8f0c-457e-8d1...	https	Available	1.97 MB	2024-04-08 18:22:38	No
13755251			2a11697f-57a9-4804-b17...	https	Available	188 MB	2024-04-08 15:09:26	No
13755243			ae915129-be1b-4419-915...	https	Available	18.49 GB	2024-04-08 15:02:49	No
13755237			ae80164e-322d-465e-911...	https	Available	18.07 KB	2024-04-08 15:01:53	No
13755225			c3d39e61-5299-4165-9f6...	https	Available	62.86 MB	2024-04-08 14:42:39	No
13754711			8d06eeed-848c-4889-e2...	https	Available	48.58 MB	2024-04-08 14:11:47	No
13754702			91729a2f-002d-4d2c-b6...	https	Available	23.29 MB	2024-04-08 14:04:46	No
13754684			2a921c53-8a06-4014-05a...	https	Restoring from tape	733.25 KB	2024-04-08 13:42:56	Yes
13754578			a7718da9-7c5e-41c7-a3...	https	Available	733.25 KB	2024-04-08 13:42:39	No
13752821			89c71761-1d6c-4039-ae1...	https	Available	906.14 MB	2024-04-08 13:41:38	No

## Current retrievals summary

StorageD current retrievals summary

RR ID	RR Start Time	RR State	R State	R State Count	File Count	File Size (GB)
789754	20240408 02:20:25	BUILDING	COMPLETED	49	2614	16.29
			PROCESSABLE	512	17121	106.66
789755	20240408 02:35:07	BUILDING	SERVING	1	86	.54
			PROCESSABLE	6	680	.01
789774	20240408 05:36:08	BUILDING	SERVING	1	114	0
			PROCESSABLE	91	5170	32.21
789774	20240408 07:18:52	BUILDING	SERVING	1	86	.54
			PROCESSABLE	6	680	.01
789774	20240408 10:35:05	BUILDING	SERVING	1	114	0
			PROCESSABLE	6	680	.01
789784	20240408 12:11:22	BUILDING	SERVING	1	114	0
			PROCESSABLE	512	17121	106.66
789784	20240408 13:20:59	BUILDING	SERVING	1	86	.54
			COMPLETED	1615	2242	578.9
789784	20240408 13:20:59	BUILDING	PROCESSABLE	89	55	1.39
			SERVING	1	2	.23
789786	20240409 08:10:33	BUILDING	SYNCED	9	11	.29
			PROCESSABLE	91	5170	32.21
789835	20240409 08:10:33	BUILDING	SERVING	1	86	.54
			PROCESSABLE	60	87	2.38
789836	20240409 08:12:40	BUILDING	SERVING	1	2	.03
			PROCESSABLE	9	11	.29
789836	20240409 08:12:40	BUILDING	COMPLETED	61	88	2.41
			SERVING	9	11	.29
789837	20240409 08:21:04	BUILDING	COMPLETED	40	56	1.42
			SYNCED	9	11	.29

Home Proposals

Results: 16

Display as cards

Title	Name
As Analysis of Internal Strains and Grain Rotation in Commercially Pure Nickel Subjected to Creep Loading	SM03775
Determining how mechanical cycling impacts residual strains, structural order and fracture strength in photo-cured polymer dental filling materials	SM03702
Determining how rate dependent residual strains impact the visco-elastic creep response of photo-cured polymer dental filling materials	SM03577
Determining how residual strains and structural order impact fracture strength in photo-cured polymer dental filling materials (Copy)	SM03916
Extracting nano-scale strains in intact arthritic joints during loading via combined 4D Synchrotron tomography and digital volume correlation (Copy)	MT14292
Investigating how Printing Parameters Impact Interfacial Strains and Failure in 3D Printed Dental Fillings	SM03532
Investigating reverse lattice strains in zirconium hydroxides (Copy)	SM03694
Investigating how Fluid Processing Affects Residual Impact Residual Strains within 3D Printed Dental Fillings	SM05842
Investigation of halide-missing induced local strains in halide perovskite semiconductors	CY28352
Investigation of halide-missing induced local strains in halide perovskite semiconductors (Copy)	CY30299
Mapping of 3D strains during micro-indentation of alloys for energy applications	MT8431
Measurement of Depth Resolved Strains in Protective Oxide Coatings on Aluminum	EE0384
Multi-scale Analysis of Internal Strains and Residual Stresses in Commercially Pure Nickel Subjected to Creep Loading	SM03719
Operando Strains in ZrO <sub>2</sub> Catalysts	MT3864
Operando strains in Cu/SiO <sub>2</sub> /Si as a function of humidity	MT16101
Validation of New X-Ray Diffraction Imaging and 3D Surface Modelling Techniques to Map Strains and Warpage in Advanced IC Packages	MT81026

File Manager

Collector DLS Globus Collector

Path /reportstages/dsls/dsl56453/

Start

Transfer 3 Timer Options

NAME	LAST MODIFIED	SIZE
...	...	...
DLS_2023-12-13_30-33-32	2023-12-13 02:38 AM	-
DLS_2023-12-13_30-8-38	03:04 PM '23 02:37 AM	-
DLS_2024-2-28_10-52-32	2024-02-28 10:52 AM	-
DLS_2024-2-28_9-45-21	2024-02-28 09:45 AM	-
DLS_example_journal_1	2024-03-01 02:30 AM	-
DLS_msc_to_Atlantis	2024-03-01 02:30 AM	-
DLS_2022-11-14_9-49-38	22-11-14 09:49 AM	-
DLS_2022-12-12_9-48-52	22-12-12 09:48 AM	-
DLS_2022-5-02_almem_1	03-05-2022 10:05 AM	-
DLS_2022-3-25_posenwhor_1	03-03-2022 03:02 AM	-
DLS_2022-7-12_12-47-58	22-07-12 12:47 PM	-
DLS_2023-10-18_posenwhor_1	23-10-18 03:02 PM	-
DLS_2023-12-13_30-33-32	23-12-13 02:38 AM	-
DLS_2023-12-6_9-48-24	23-12-06 09:48 AM	-
DLS_2023-7-18-11-57-13	23-07-18 11:57 AM	-
DLS_msc_to_Atlantis	2024-03-01 02:30 AM	-
Newcom2022_1	2023-08-02 02:22 PM	-

# Scale ( by numbers)

## INGEST

- DLS
  - 65.5PB , 4.8B files
- CEDA (13 years)
  - 27.9PB , 417M files
    - Daily record 77TB, 6.4M files
- ET (10 years)
  - 12.6PB , 256M files
    - Daily record 82TB, 1.9M files

## Recall

- DLS (7 years)
  - 3.1PB ,
- CEDA (13 years)
  - 82.5PB , 403M files
    - Monthly record 788TB, 82M files
- ET (10 years)
  - 2.7PB , 7.5M files