

# ADVANCED TECHNOLOGY GROUP (ATG)



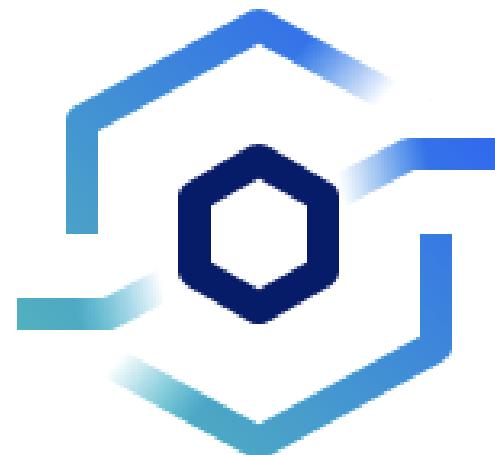
## Accelerate with ATG Webinar

### IBM Storage Ceph File System (CephFS) Deep Dive

Feature updates and live demo

John Shubeck – ATG Storage Technical Specialist

Date: April 8, 2025



## Accelerate with ATG Technical Webinar Series

---

Advanced Technology Group experts cover a variety of technical topics.

**Audience:** Clients who have or are considering acquiring IBM Storage solutions. Business Partners and IBMers are also welcome.

To automatically receive announcements of upcoming Accelerate with ATG webinars - Clients, Business Partners and IBMers are welcome to send an email request to [accelerate-join@hursley.ibm.com](mailto:accelerate-join@hursley.ibm.com).

### 2025 Upcoming Webinars – Register Here!

[Building Affordable, Resilient Storage Solutions with IBM FlashSystem and Brocade SAN](#) - April 24<sup>th</sup>, 2025

[AI Data Assistance with IBM Content Aware Storage \(CAS\)](#) - April 29<sup>th</sup>, 2025

[IBM Storage Virtualize 8.7.3 Technical Update & the New FlashSystem C200](#) - May 13th, 2025



### Important Links to Bookmark:

**Accelerate with ATG** - Click here to access the Accelerate with ATG webinar schedule for 2025, view presentation materials, and watch past replays dating back two years. <https://ibm.biz/BdSUFN>

**ATG MediaCenter Channel** - This channel offers a wealth of additional videos covering a wide range of storage topics, including IBM Flash, DS8, Tape, Ceph, Fusion, Cyber Resiliency, Cloud Object Storage, and more. <https://ibm.biz/BdfEgQ>

## Offerings

---

### Client Technical Workshops

- **Cyber Resilience with IBM Storage Defender: April 9, 2025 (Virtual)**
- **IBM DS8000 G10 Advanced Functions: April 30 & May 1, 2025 (Coppell, TX)**
- **IBM Fusion & Ceph: May 14-15, 2025 (Durham, NC)**
- **IBM FlashSystem Deep Dive & Advanced Functions: May 21-22, 2025 (Chicago, IL)**
- **IBM Storage Scale & Storage Scale Functions: June 4-5, 2025 (NYC)**

### TechZone Test Drive / Demo's

- IBM Storage Scale and Storage Scale System GUI
- IBM Storage Virtualize Test Drive
- IBM DS8900F Storage Management Test Drive
- Managing Copy Services on the DS8000 Using IBM Copy Services Manager Test Drive
- IBM DS8900F Safeguarded Copy (SGC) Test Drive
- IBM Cloud Object Storage Test Drive - (Appliance based)
- IBM Cloud Object Storage Test Drive - (VMware based)
- IBM Storage Protect Live Test Drive
- IBM Storage Ceph Test Drive - (VMware based)

Please reach out to your IBM Representative or Business Partner for more information.

**\*IMPORTANT\* The ATG team serves clients and Business Partners in the Americas, concentrating on North America.**

## Accelerate with ATG Survey

---

Please take a moment to share your feedback with our team!

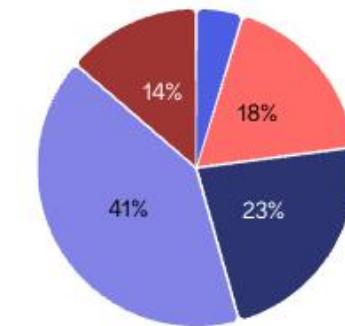
You can access this 6-question survey via [Menti.com](https://Menti.com) with code 5151  
0447 or

Direct link <https://www.menti.com/alhsf3bgvxu6>  
Or

QR Code



What's your favorite way to embrace the "April showers bring May flowers" vibe?



- 5% Jumping in puddles with colorful rain boots!
- 18% Curled up with a good book and a warm drink, listening to the rain.
- 23% Planting seeds and watching them sprout.
- 41% Admiring the blooming flowers and taking nature walks.
- 14% Hoping for a surprise rainbow after the rain.

# ADVANCED TECHNOLOGY GROUP (ATG)



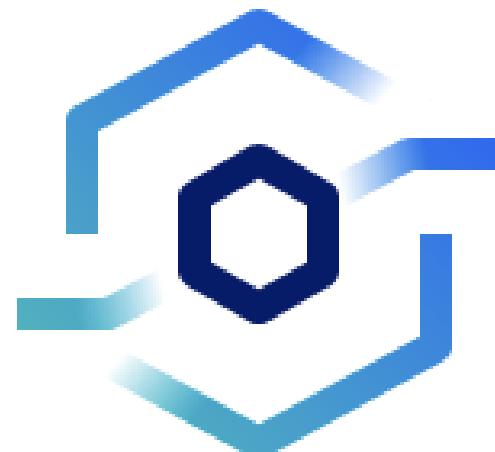
## Accelerate with ATG Webinar

### IBM Storage Ceph File System (CephFS) Deep Dive

Feature updates and live demo

John Shubeck – ATG Storage Technical Specialist

Date: April 8, 2025



## About the Presenter

---



**John Shubeck** is an information technology professional with over 42 years of industry experience spanning both the customer and technology provider experience. John is currently serving as a Senior Storage Technical Specialist on IBM Object Storage platforms across all market segments in the Americas.

## Introducing our panelists



**Shu Mookerjee** is a Level 2 Certified Technical Specialist with over twenty years at IBM, working in a variety of roles including sales, management and technology. For the last decade, he has focused exclusively on storage and has been the co-author of four (4) Redbooks. Currently, Shu is part of the Advanced Technology Group where he provides education, technical guidance, Proofs of Concept and Proofs of Technology to IBMers, business partners and clients.

## Introducing our panelists

---



**Jerrod Carr** is an IBM Principal Storage Technical Specialist in IBM Storage Solutions. Jerrod Carr has been in the Storage industry for over 21 years selling hardware and software for various large technology companies. With beginnings in the Cleversafe IBM team for 8 years providing expertise in Cloud Object Storage, the last 3 years working on the Americas SWAT team as a Senior Storage Specialist providing unstructured data experience to the various markets.

## Summary of topics

---



- Overview of the Ceph File System (CephFS)
- The CephFS Metadata Service (MDS)
- Day 1 operations: Configuring our first Ceph file system
- Day 2 operations: The case for subvolumes
- Day 2 operations: Subvolume snapshots and snapshot schedules
- Day 2 operations: File sharing via NFS
- Configuring multiprotocol (SMB tech preview)

## Ceph File System (CephFS) topology

---

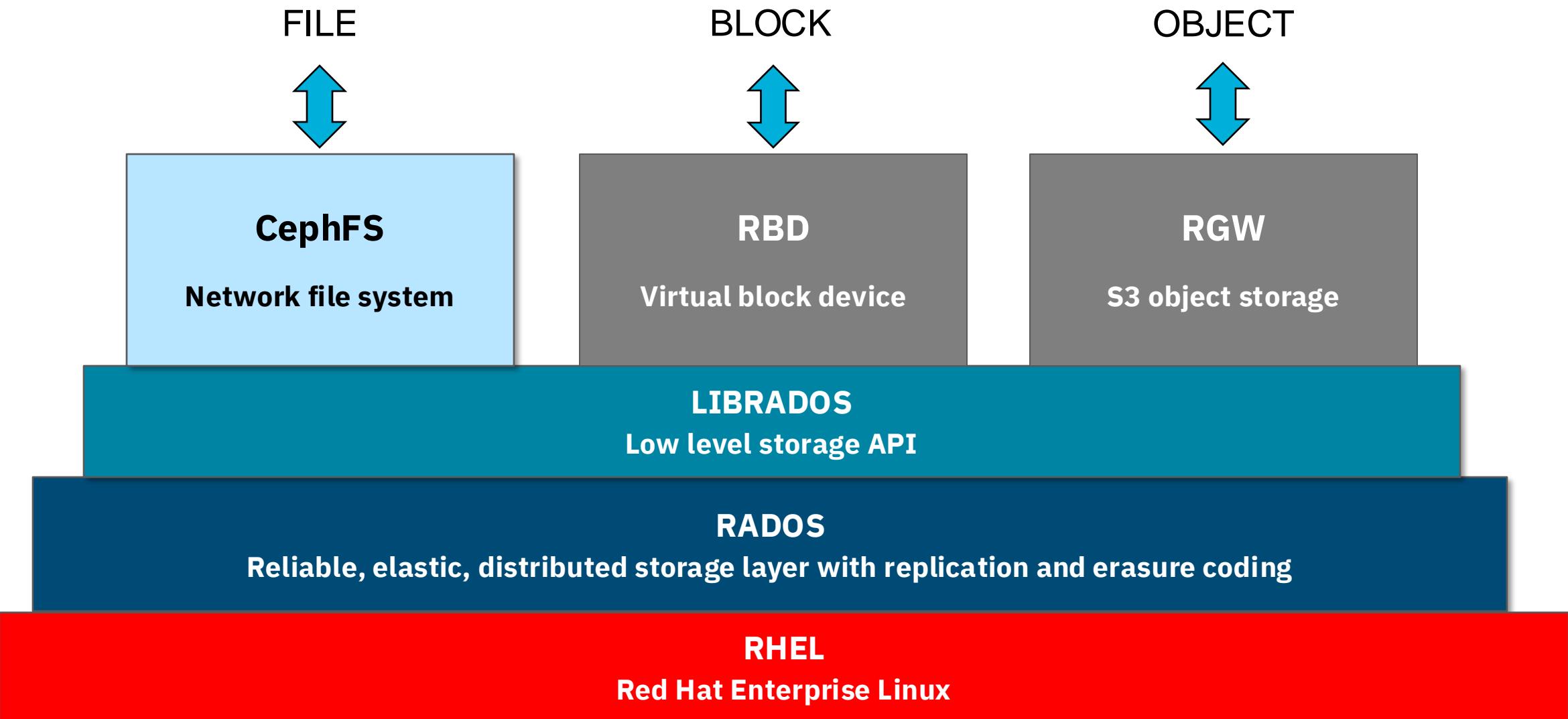


## What is CephFS?

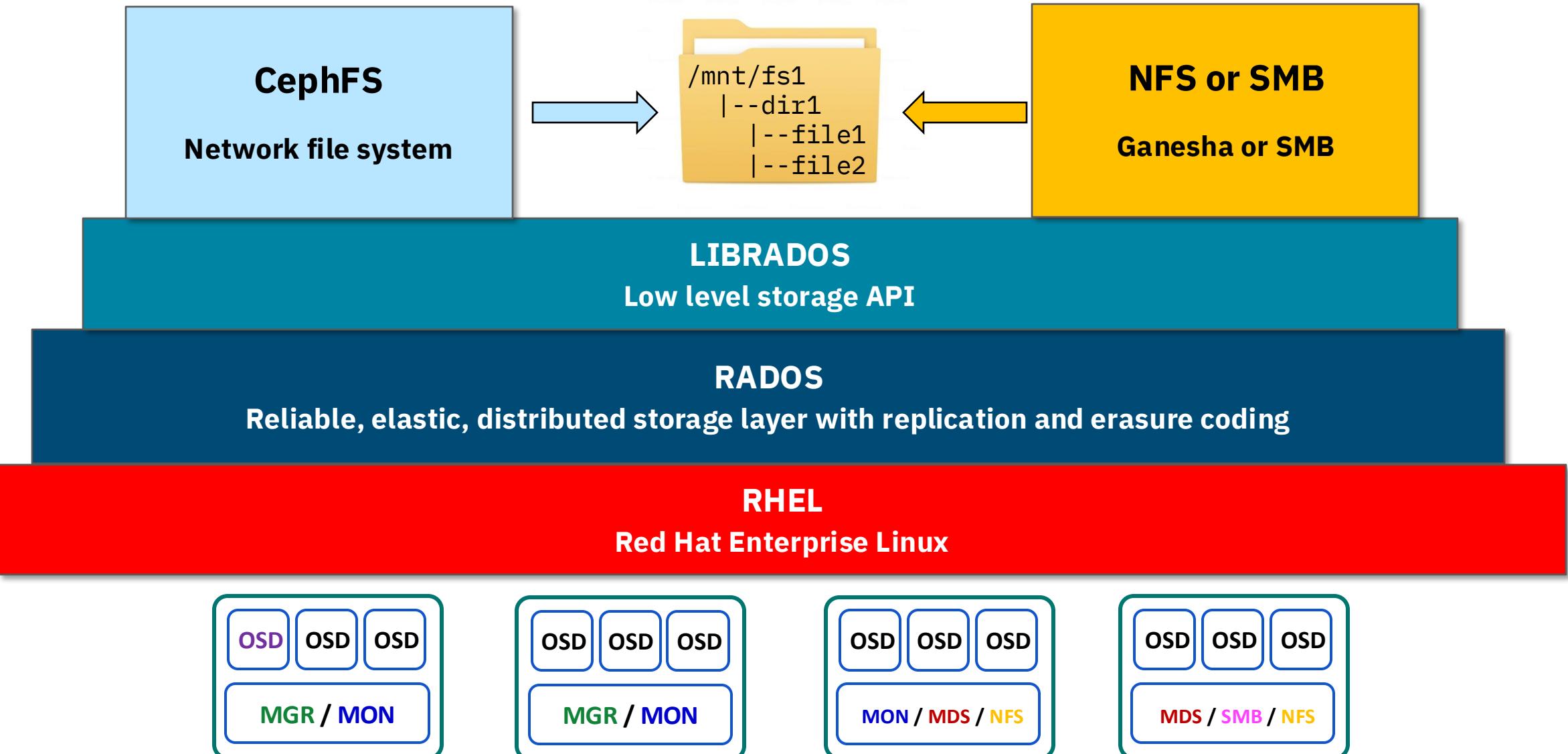
---

- The Ceph File System (*CephFS*) is a POSIX-compatible file system that is build on top of RADOS (i.e. The Ceph distributed object store).
- File-based storage organizes your data as a traditional file system.
  - Data is saved as files with a name and associated metadata, such as modification time stamps, an owner, and access permissions
  - File-based storage uses a directory tree hierarchy to organize how files are stored
- The Metadata Server (*MDS*) manages metadata for CephFS clients.
  - MDS manages the directory hierarchy and stores file metadata, such as the owner, time stamps, and permission modes.
  - MDS is also responsible for access caching, client caches, and maintaining cache coherence.

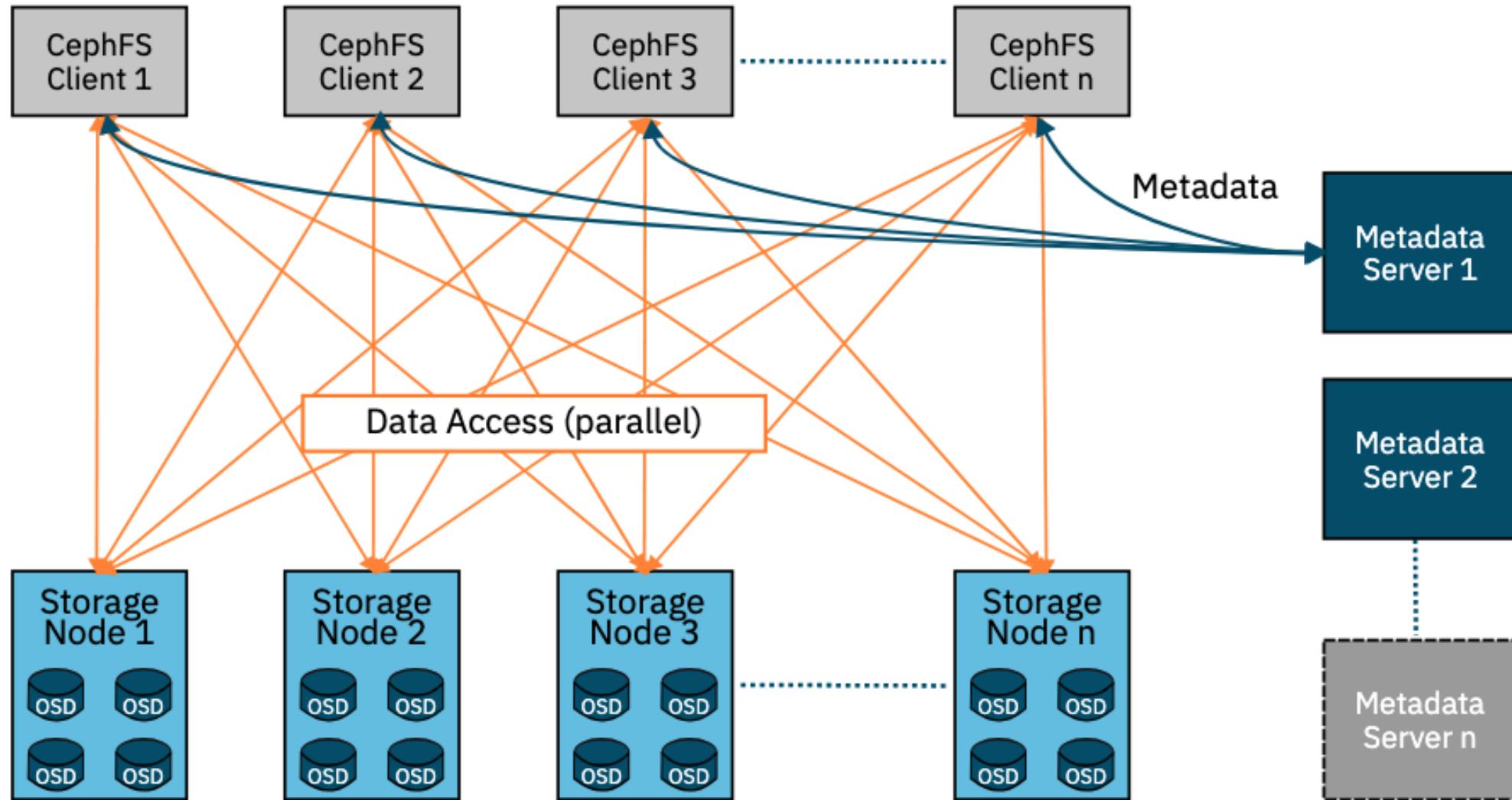
## IBM Storage Ceph File System components (CephFS)



## IBM Storage Ceph File System components (CephFS)

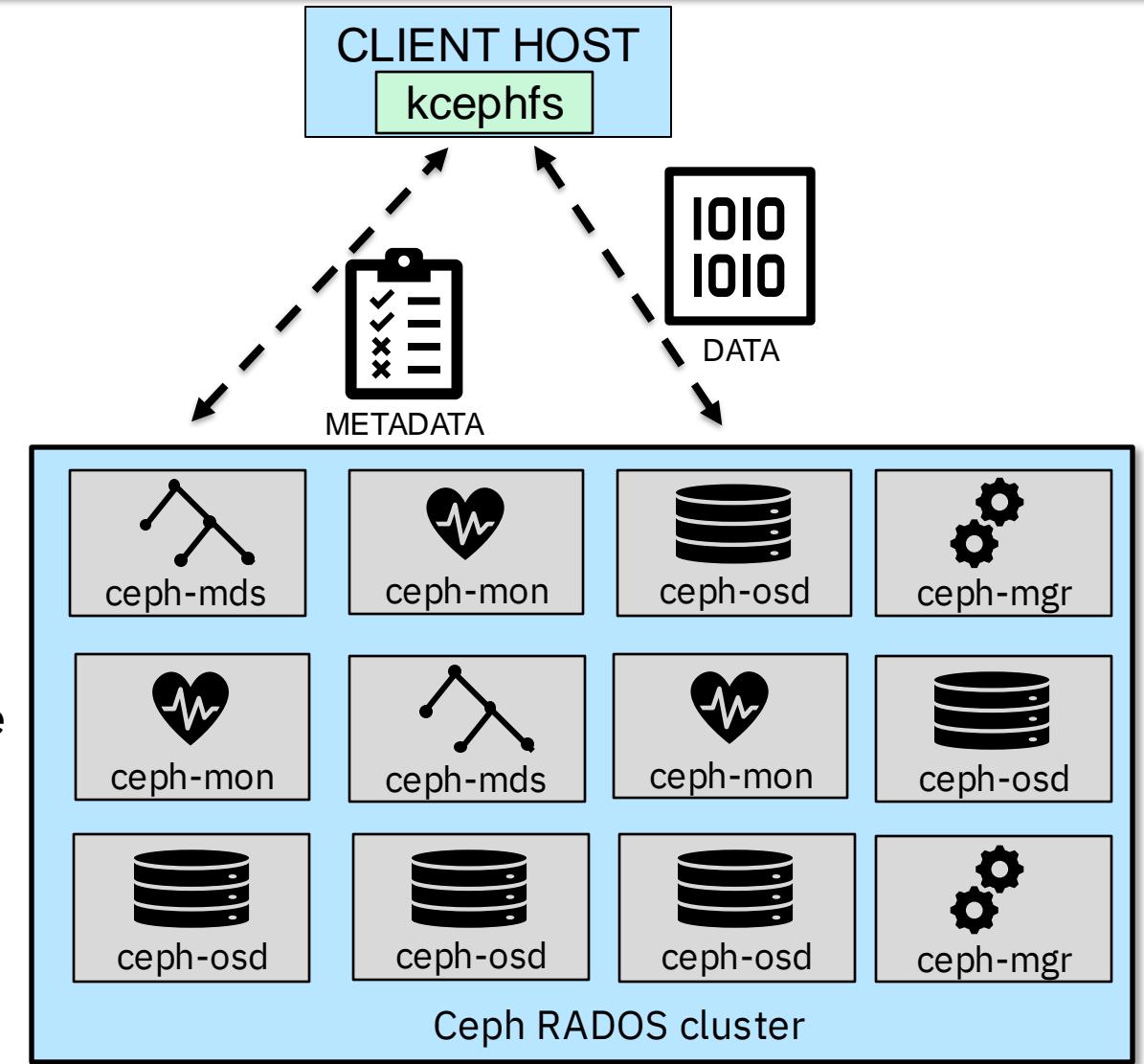
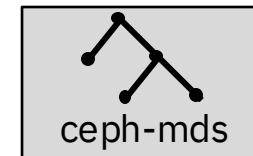


## CephFS data flow at Level 101



## Metadata is stored in RADOS through *ceph-mds*

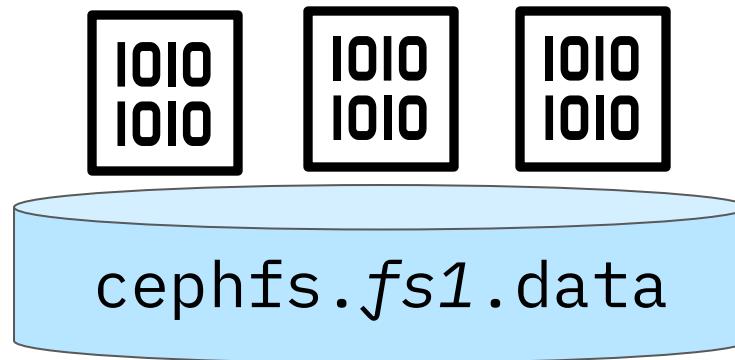
- Distributed network file system
  - Files, directories, rename, hard links, etc.
  - Concurrent shared access from many clients
- Strong consistency and coherent caching
  - Updates from one node visible elsewhere immediately
- Scale metadata and data independently
  - Storage capacity and I/O throughput scale with the number of OSDs
  - Namespace (e.g. number of files) scales with the number of MDS daemons



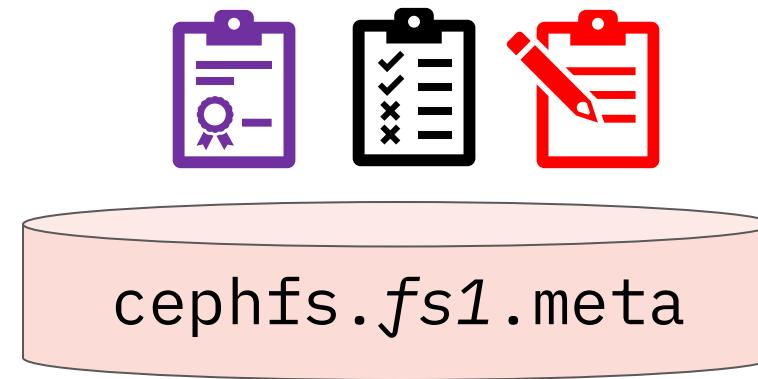
## Ceph Pools for file systems

---

- The Data Pool (*cephfs.data*) stores the file data.



- The Meta Data Pool (*cephfs.meta*) stores the file system meta data.



## Ceph File System level 201 (ceph-mds)

---



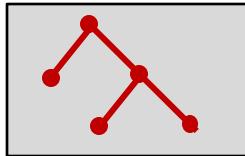
## CEPH-MDS (The Ceph Metadata Daemon)

---

- The Metadata Server ([MDS](#)) manages metadata for CephFS clients. The MDS daemons operate in two modes: active and standby. An active MDS manages the metadata on the CephFS file system. A standby MDS serves as a backup, and switches to the active mode if the active MDS becomes unresponsive. As a best practice, at least one standby MDS in the cluster to ensure high availability

## Ceph-MDS: Metadata Server

---



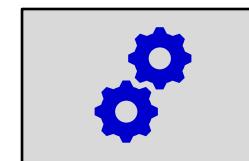
ceph-mds

### MDS (Metadata Server)

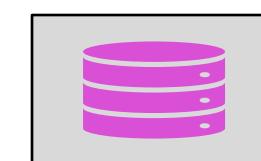
- Manages file system namespace
- Stores file system metadata in RADOS objects
  - File and directory metadata (names, inodes)
- Coordinates file access between clients
- Manages client cache consistency, locks, leases
- Not part of the data path
- 1s – 10s active, plus standbys



ceph-mon

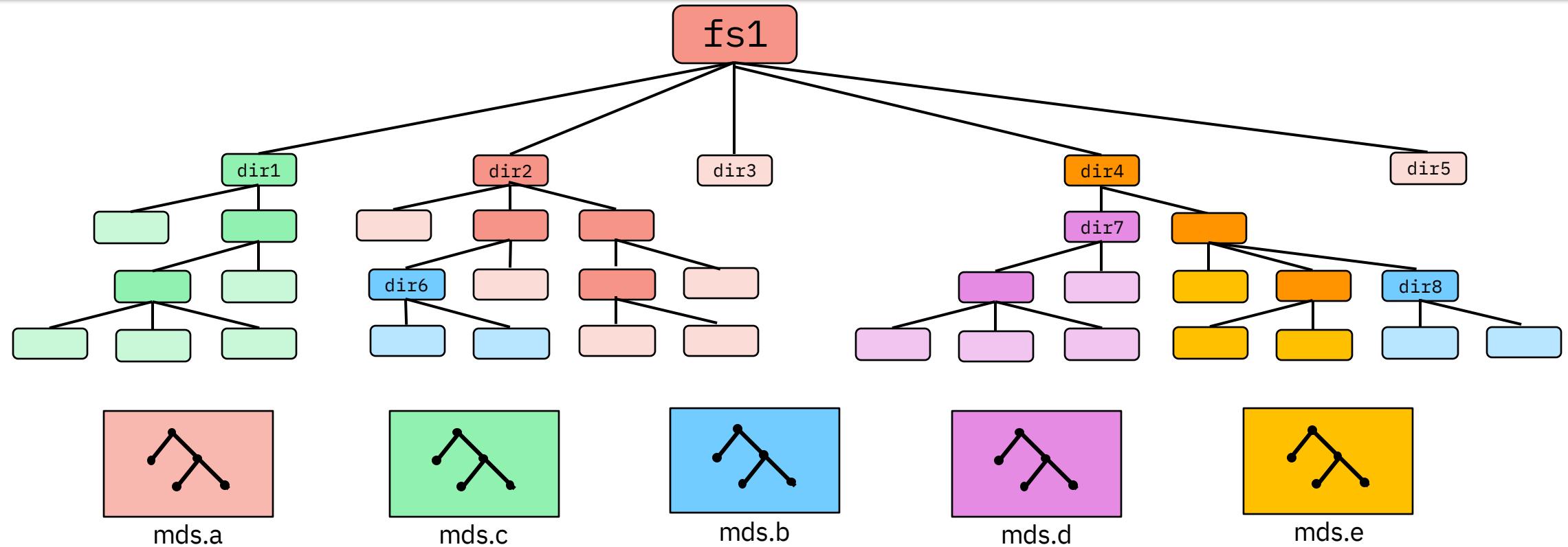


ceph-mgr



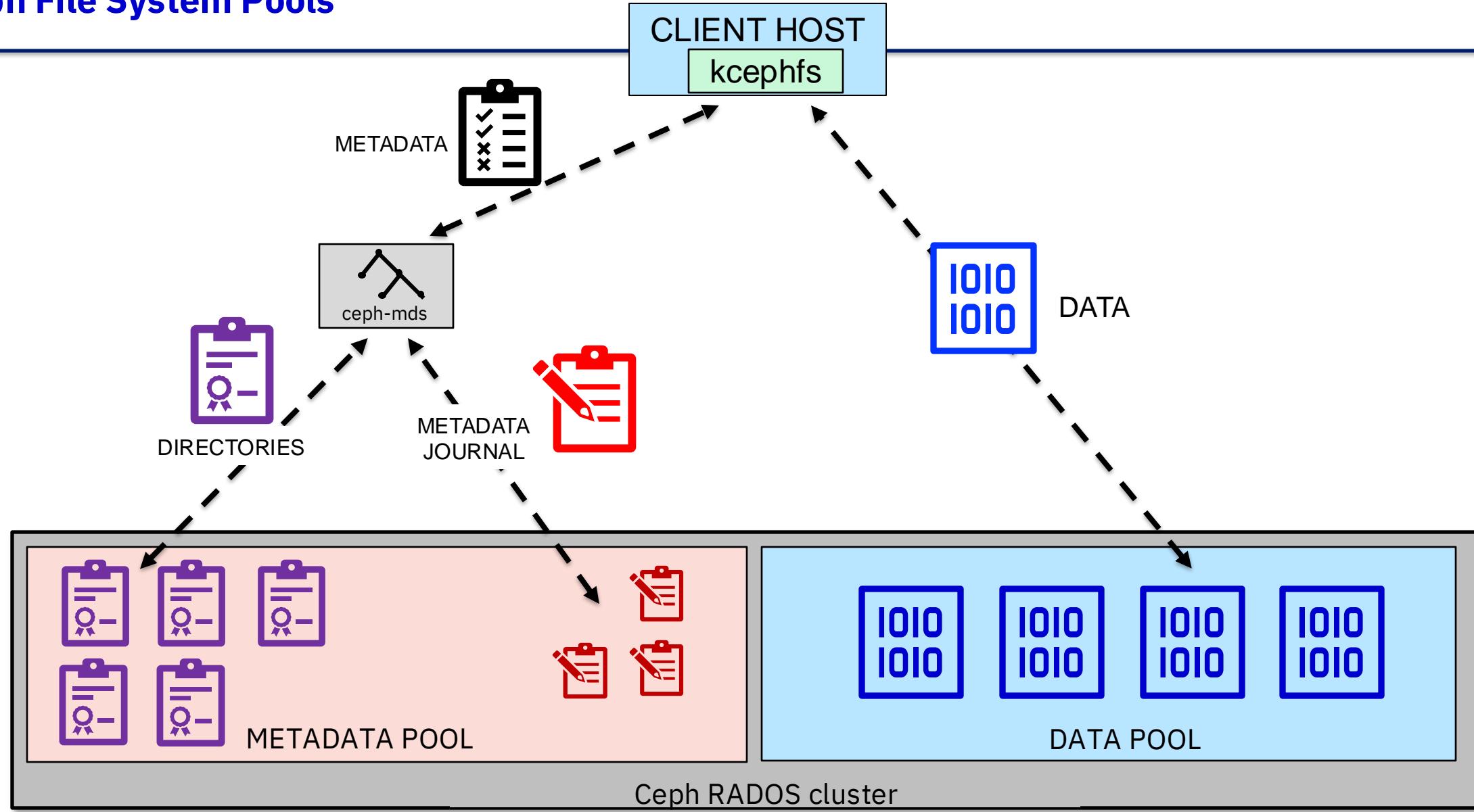
ceph-osd

## Scalable Namespace – Dynamic subtree partitioning

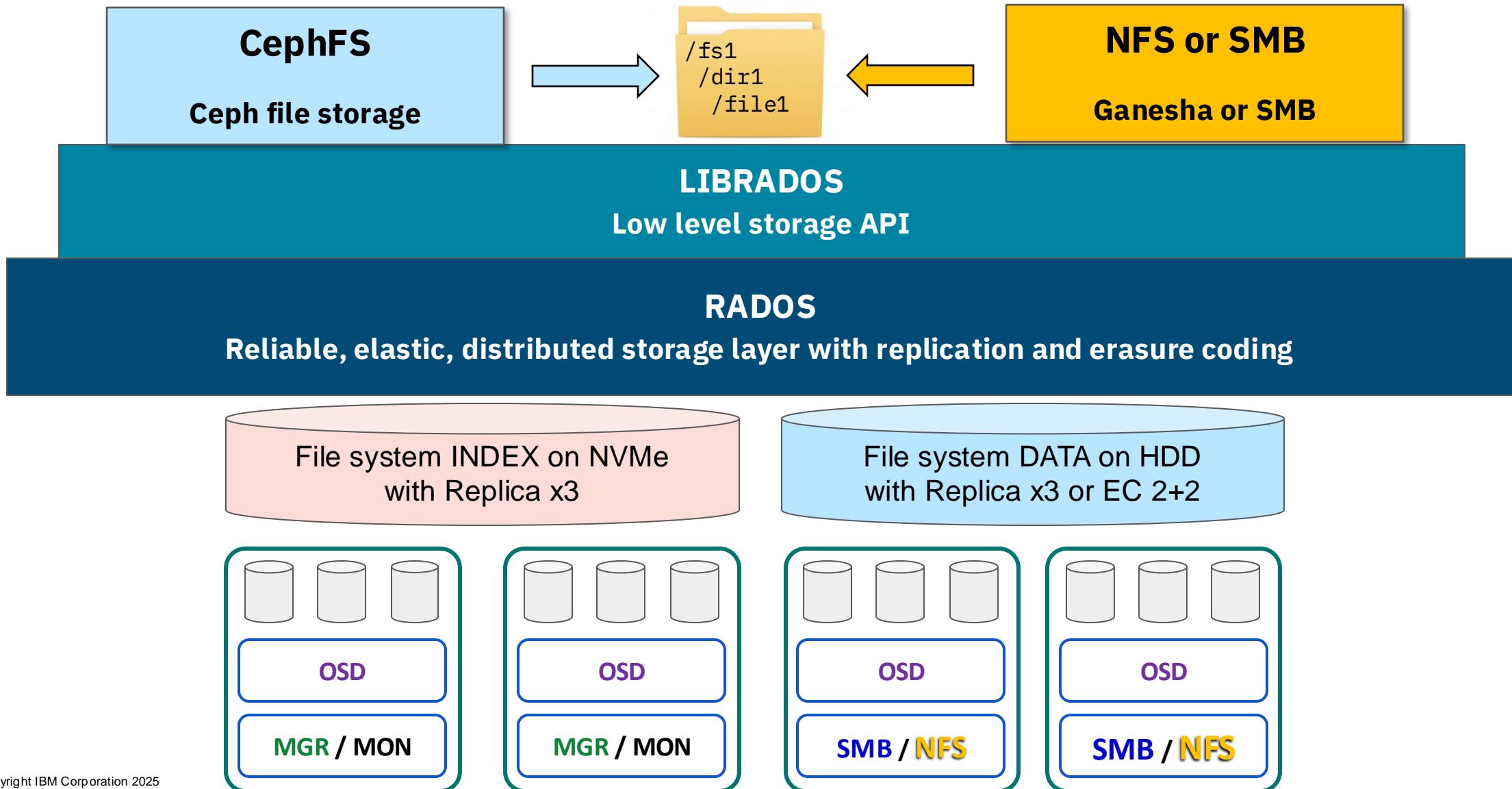


- Partition hierarchy across MDSs based on workload
- Fragment huge directories across MDSs
- Clients learn the overall partition as they navigate the namespace
- Subtree partition maintains directory locality
- Arbitrarily scalable by adding more MDSs

## Ceph File System Pools



## IBM Storage Ceph File System components (CephFS)



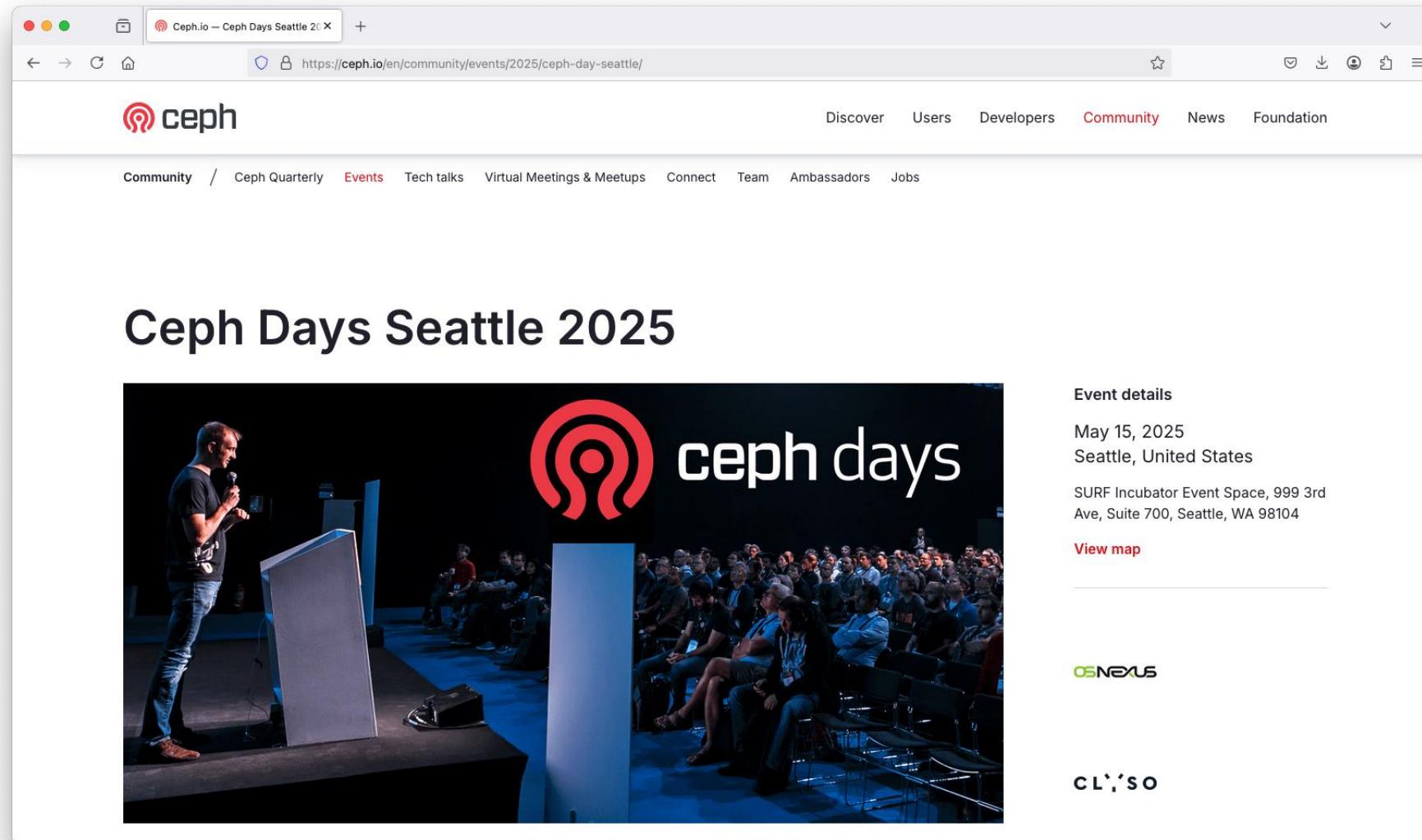
## A message from the Ceph Community

---

. . . 30 SECOND COMMERCIAL MESSAGE . . .

Join the Ceph Community in Seattle!

(<https://www.surfincubator.com/>)



The screenshot shows a web browser displaying the Ceph Days Seattle 2025 event page. The page features the Ceph logo and navigation links for Discover, Users, Developers, Community (which is highlighted in red), News, and Foundation. Below the navigation, there are links for Community, Ceph Quarterly, Events (which is highlighted in red), Tech talks, Virtual Meetings & Meetups, Connect, Team, Ambassadors, and Jobs. The main content area is titled "Ceph Days Seattle 2025" and includes a large image of a speaker on stage. To the right of the image, there is an "Event details" section with the date (May 15, 2025), location (Seattle, United States), and address (SURF Incubator Event Space, 999 3rd Ave, Suite 700, Seattle, WA 98104). A "View map" link is also present. At the bottom of the page, there are logos for OS NEXUS and CL'sO.

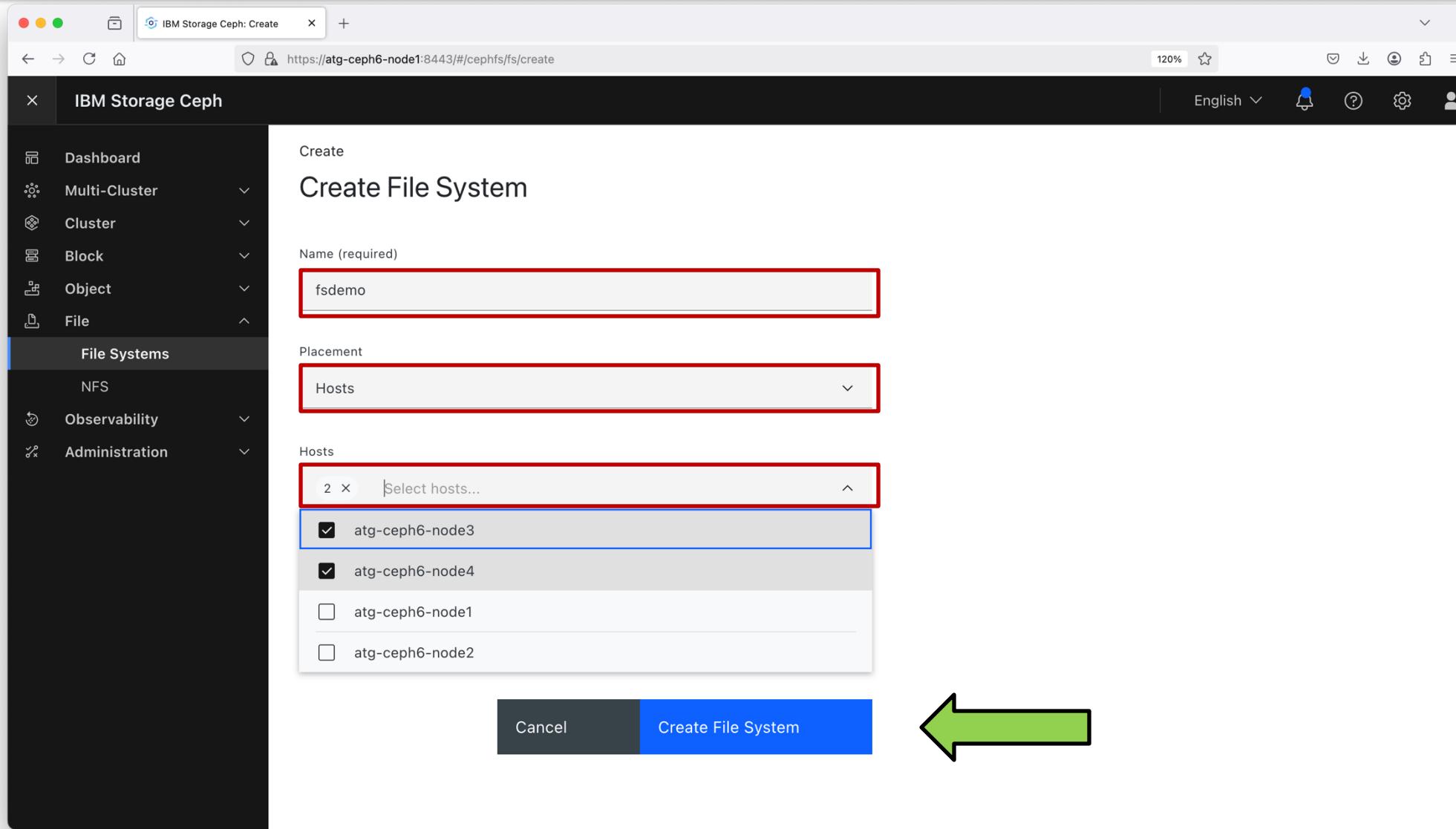
<https://ceph.io/en/community/events/2025/ceph-day-seattle/>

## Ceph File System (CephFS) day 1 operations

---



## Three click configuration in the *Ceph Dashboard* to create a new File System



## Single command in *Cephadm* to create a new File System

```
[ceph: /]# ceph fs volume create fsdemo
```

<https://www.ibm.com/docs/en/storage-ceph/8.0?topic=volumes-creating-cephfs-volume>

## Simple CephFS configuration in the *Dashboard* .and. *Cephadm* (imperative method)<sup>1</sup>

Create

### Create File System

Name (required)

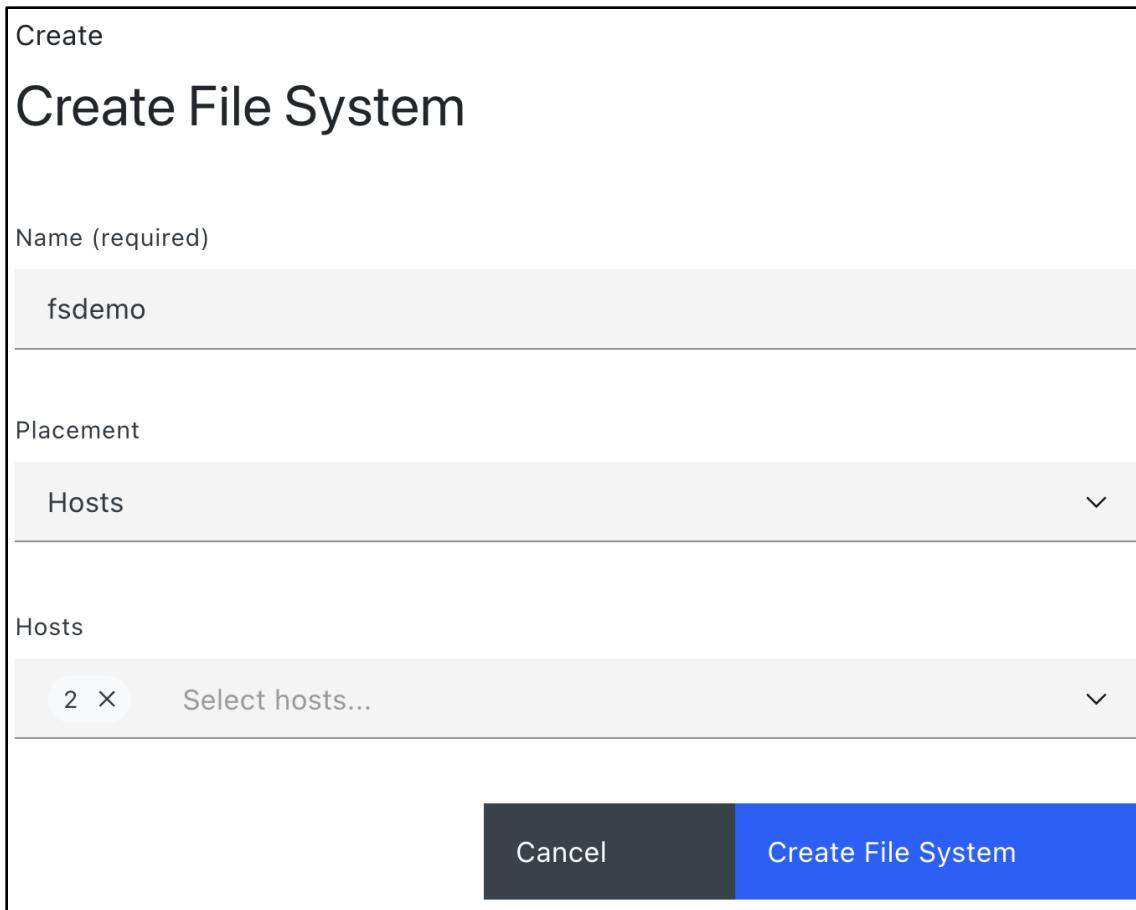
Placement

Hosts

Hosts

2 × Select hosts...

**Cancel** **Create File System**



```
[ceph]# ceph fs volume create fsdemo
```

```
[ceph]# ceph orch ls mds
```

NAME	RUNNING	REFRESHED	AGE	PLACEMENT
mds.fsdemo	2/2	4m ago	4m	count:2

1 <https://www.ibm.com/docs/en/storage-ceph/8.0?topic=volumes-creating-cephfs-volume>

## The CephFS client experience

```
[root@local ~]# mkdir /mnt/fsdemo
[root@local ~]# mount -t ceph atg-ceph6-node2:/ /mnt/fsdemo -o fs=fsdemo,name=admin

[root@local ~]# df /mnt/fsdemo
Filesystem      1K-blocks  Used Available Use% Mounted on
192.168.65.231:/  84520960      0  84520960    0% /mnt/fsdemo

[root@local ~]# mkdir /mnt/fsdemo/dir1 && mkdir /mnt/fsdemo/dir2
[root@local ~]# touch /mnt/fsdemo/dir1/atestfile
[root@local ~]# echo "Hello world!" > /mnt/fsdemo/dir1/hello-world.txt

[root@local ~]# tree /mnt/fsdemo
/mnt/fsdemo
└── dir1
    ├── atestfile
    └── hello-world.txt
    └── dir2
```

## CephFS recursive accounting

- MDS maintains recursive stats across the file hierarchy
  - File and directory counts
  - File size (summation)
  - Latest *ctime*
- Visible via virtual xattrs
- Recursive bytes as directory size
  - If mounted with *rbytes* option
  - Similar to *du*, but more detail and file size rather than disk space used

```
$ sudo mount -t ceph ceph-node:/ /mnt/fsdemo \
-o fs=fsdemo,name=admin,rbytes
$ cd /mnt/fsdemo/dir1
$ getfattr -d -m - .
# file: atg/zipcodes/sample.csv
security.selinux="unconfined_u:object_r:unlabeled_t:s0"
```

```
$ ls -alh
total 0
drwxr-xr-x. 5 root root 47M Mar 21 17:51 .
drwxr-xr-x. 7 root root 87 Mar 24 19:55 ..
drwxr-xr-x. 2 root root 10M Mar 21 16:13 dir1
drwxr-xr-x. 3 root root 23M Mar 29 09:59 dir2
drwxr-xr-x. 3 root root 14M Mar 21 17:52 volumes
```

## Ceph File System (CephFS) day 1 and ½ operations

---



# Day 1 and ½ clicks in the *Ceph Dashboard* to view Ceph File System properties

The screenshot shows the IBM Storage Ceph File System dashboard. The left sidebar is dark with white text, showing navigation links for Dashboard, Multi-Cluster, Cluster, Block, Object, File, and File Systems. The 'File Systems' link is highlighted with a blue border. The main content area has a light gray background. At the top, there is a search bar and a 'Create' button. Below it, a table shows the 'fsdemo' file system with columns: Name, Enabled, and Created. The 'Enabled' column shows a checkmark, and the 'Created' column shows 'An hour ago'. Below the table, a navigation bar has tabs for Details, Directories, Subvolumes, Subvolume groups, Snapshots, Snapshot schedules, Clients (0), and Performance Details. The 'Details' tab is selected and highlighted with a blue border. The 'Ranks' section shows a table with columns: Rank, State, Daemon, Activity, Dentries, Inodes, and Dirs. The first row shows '0' as the rank, 'active' as the state, 'fsdemo.atg-ceph6-node1.uqbbaj' as the daemon, 'Reqs: 0 /s' as activity, '10' as dentries, '13' as inodes, and '12' as dirs. The 'Standbys' section shows a table with columns: Key and Value. It shows 'Standby daemons' as the key and 'fsdemo.atg-ceph6-node2.czexcy' as the value. Below these sections are pagination controls for 'Items per page' (10) and '1-1 of 1 item'. The 'Pools' section shows a table with columns: Pool, Type, Size, and Usage. It lists two pools: 'cephfs.fsdemo.data' (data type, 80.6 GiB size, 0% usage) and 'cephfs.fsdemo.meta' (metadata type, 80.6 GiB size, 0% usage). Below the pools are similar pagination controls.

Rank	State	Daemon	Activity	Dentries	Inodes	Dirs
0	active	fsdemo.atg-ceph6-node1.uqbbaj	Reqs: 0 /s	10	13	12

Key	Value
Standby daemons	fsdemo.atg-ceph6-node2.czexcy

Pool	Type	Size	Usage
cephfs.fsdemo.data	data	80.6 GiB	0%
cephfs.fsdemo.meta	metadata	80.6 GiB	0%

Day 1 and ½ commands in the *Cephadm* for CephFS visibility

```
[ceph: /]# ceph fs ls
name: fsdemo, metadata pool: cephfs.fsdemo.meta, data pools: [cephfs.fsdemo.data ]

[ceph: /]# ceph df
--- RAW STORAGE ---
CLASS      SIZE      AVAIL      USED      RAW USED  %RAW USED
hdd      256 GiB  255 GiB  1.1 GiB  1.1 GiB      0.44
TOTAL    256 GiB  255 GiB  1.1 GiB  1.1 GiB      0.44

--- POOLS ---
POOL          ID  PGS  STORED  OBJECTS  USED  %USED  MAX AVAIL
mgr            1    1  449 KiB      2  1.3 MiB      0  81 GiB
cephfs.fsdemo.meta  6   16  2.4 KiB     22  96 KiB      0  81 GiB
cephfs.fsdemo.data  7  512    0 B        0  0 B      0  81 GiB
```

<https://www.ibm.com/docs/en/storage-ceph/8.0?topic=ceph-object-gateway>

## Day 1 and ½ commands in the *Cephadm* for CephFS visibility

```
[ceph: /]# ceph orch ls mds
NAME      PORTS  RUNNING  REFRESHED  AGE  PLACEMENT
mds.fsdemo  2/2    4m ago    4m    count:2

[ceph: /]# ceph orch ps --daemon-type=mds
NAME          HOST          STATUS  REFRESHED  AGE  MEM  USE
mds.fsdemo.atg-ceph6-node1.uqbbaj atg-ceph6-node1  running (8m)  8m ago  8m  15.4M
mds.fsdemo.atg-ceph6-node2.czexcy atg-ceph6-node2  running (8m)  8m ago  8m  13.5M
. . . output omitted . . .

[ceph: /]# rados lspools
.mgr
cephfs.fsdemo.meta
cephfs.fsdemo.data
```

<https://www.ibm.com/docs/en/storage-ceph/8.0?topic=ceph-object-gateway>

## Day 1 and ½ commands – Get the MDS configuration using *Cephadm*

```
[ceph: /]# ceph orch ls mds --export > mds.yaml
```

```
[ceph: /]# cat mds.yaml
```

```
service_type: mds
service_id: fsdemo
service_name: mds.fsdemo
placement:
  count: 2
  hosts:
    - atg-ceph3-node1
    - atg-ceph3-node2
```

<https://www.ibm.com/docs/en/storage-ceph/8.0?topic=service-deploying-mds-using-specification>

## Day 1 and ½ commands – Apply a configuration using a service specification (declarative)

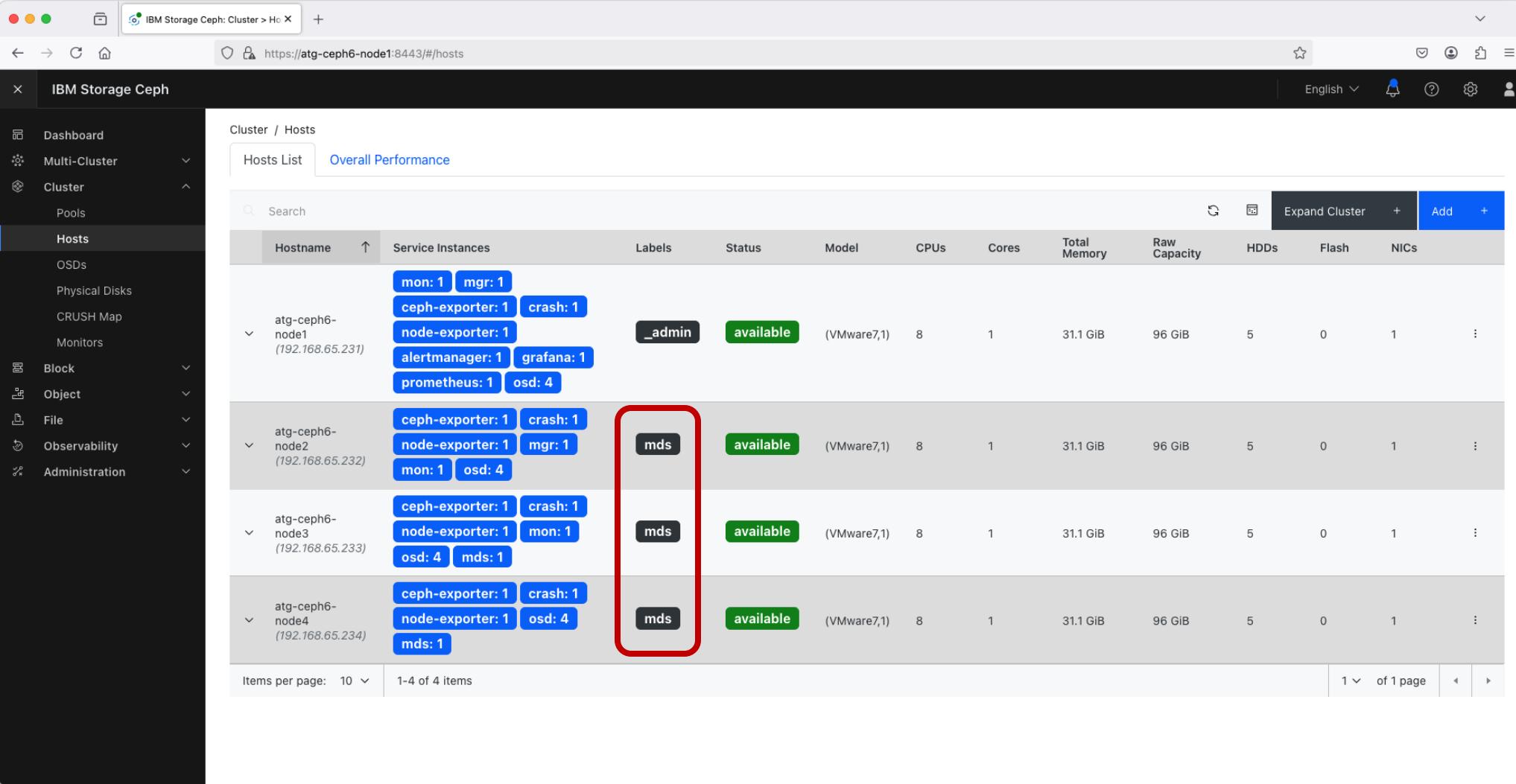
MDS service placement using named Ceph hosts

```
[ceph: /]# cat mds.yaml
service_type: mds
service_id: fsdemo
service_name: mds.fsdemo
placement:
  count: 2
  hosts:
    - atg-ceph6-node3
    - atg-ceph6-node4
```

```
[ceph: /]# ceph orch apply -i mds.yaml
Scheduled mds.fsdemo update...
```

<https://www.ibm.com/docs/en/storage-ceph/8.0?topic=service-deploying-mds-using-specification>

## Day 1 and ½ commands – Apply a label “mds” to selected Ceph hosts/nodes



The screenshot shows the IBM Storage Ceph Hosts List interface. The left sidebar navigation includes: Dashboard, Multi-Cluster, Cluster (with Pools and Hosts selected), OSDs, Physical Disks, CRUSH Map, Monitors, Block, Object, File, Observability, and Administration. The main content area displays a table of hosts with the following columns: Hostname, Service Instances, Labels, Status, Model, CPUs, Cores, Total Memory, Raw Capacity, HDDs, Flash, and NICs. The table lists four hosts:

Hostname	Service Instances	Labels	Status	Model	CPUs	Cores	Total Memory	Raw Capacity	HDDs	Flash	NICs
atg-ceph6-node1 (192.168.65.231)	mon: 1 mgr: 1 ceph-exporter: 1 crash: 1 node-exporter: 1 alertmanager: 1 grafana: 1 prometheus: 1 osd: 4	_admin	available	(VMware7,1)	8	1	31.1 GiB	96 GiB	5	0	1
atg-ceph6-node2 (192.168.65.232)	ceph-exporter: 1 crash: 1 node-exporter: 1 mgr: 1 mon: 1 osd: 4	mds	available	(VMware7,1)	8	1	31.1 GiB	96 GiB	5	0	1
atg-ceph6-node3 (192.168.65.233)	ceph-exporter: 1 crash: 1 node-exporter: 1 mon: 1 osd: 4 mds: 1	mds	available	(VMware7,1)	8	1	31.1 GiB	96 GiB	5	0	1
atg-ceph6-node4 (192.168.65.234)	ceph-exporter: 1 crash: 1 node-exporter: 1 osd: 4 mds: 1	mds	available	(VMware7,1)	8	1	31.1 GiB	96 GiB	5	0	1

Items per page: 10 | 1-4 of 4 items | 1 of 1 page

## Day 1 and ½ commands – Deploy the MDS service to hosts using a label (declarative)

---

MDS service placement following nodes labeled “mds”

```
[ceph: /]# cat mds.yaml
service_type: mds
service_id: fsdemo
service_name: mds.fsdemo
placement:
  count: 3
  label: mds

[ceph: /]# ceph orch apply -i mds.yaml
Scheduled mds.fsdemo update...
```

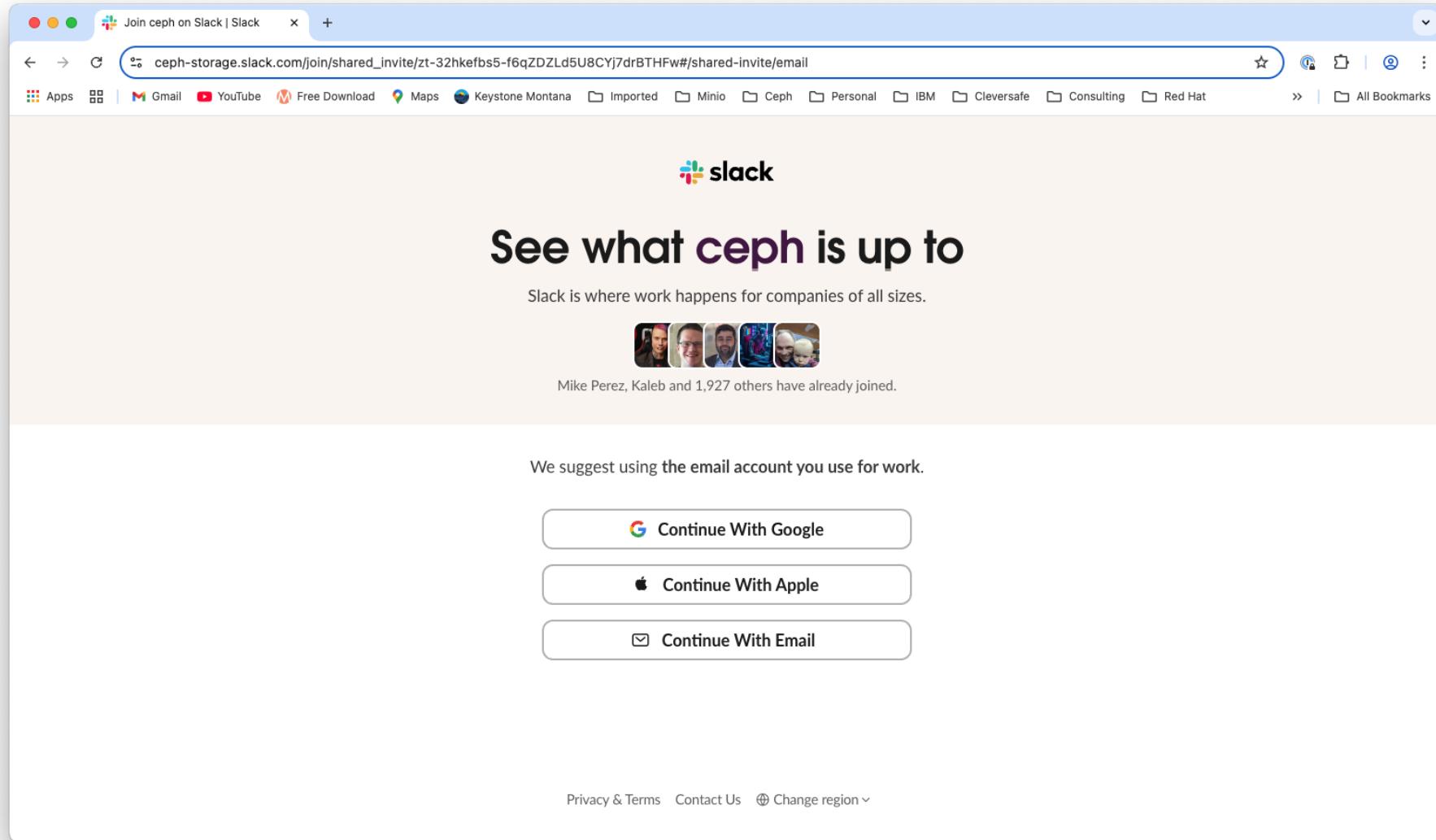
<https://www.ibm.com/docs/en/storage-ceph/8.0?topic=service-deploying-mds-using-specification>

## A message from the Ceph Community

---

. . . 30 SECOND COMMERCIAL MESSAGE . . .

## Join the Ceph Community in Slack

A screenshot of a web browser displaying the Slack sign-up page for the Ceph community. The browser window has a light blue header with the title "Join ceph on Slack | Slack" and a URL bar showing "ceph-storage.slack.com/join/shared\_invite/zt-32hkefbs5-f6qZDZLd5U8CYj7drBTHFw#/shared-invite/email". The Slack logo is at the top. The main heading is "See what ceph is up to" with the subtext "Slack is where work happens for companies of all sizes." Below this is a grid of five small user profile pictures. The text "Mike Perez, Kaleb and 1,927 others have already joined." is displayed. A note says "We suggest using the email account you use for work." followed by three sign-in buttons: "Continue With Google" (with Google logo), "Continue With Apple" (with Apple logo), and "Continue With Email" (with envelope icon). At the bottom are links for "Privacy & Terms", "Contact Us", and "Change region".

## Ceph File System (CephFS) day 2 operations

---



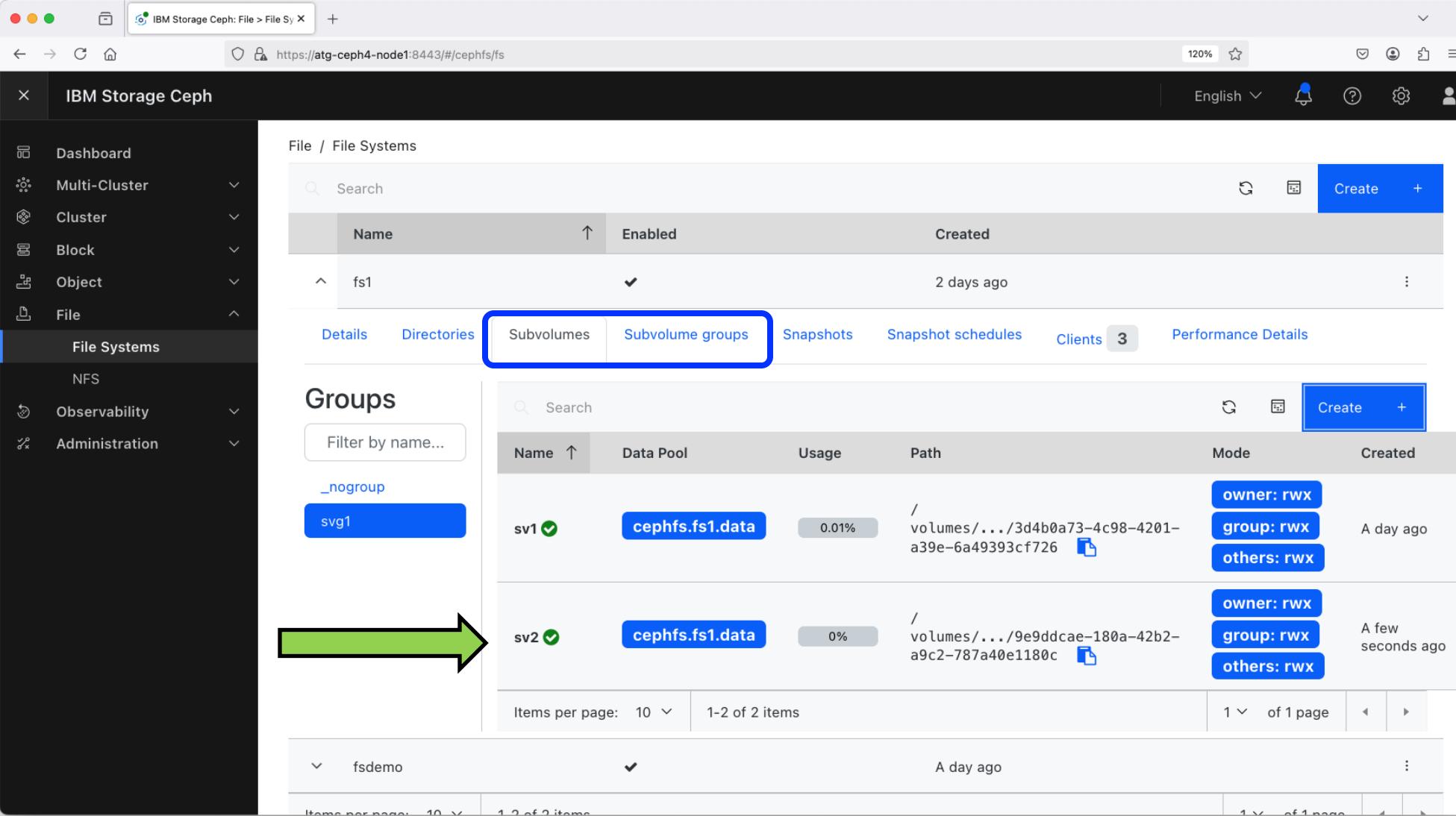
## The case for Ceph File System Subvolumes

---

Ceph recommends using CephFS volumes and subvolumes for managing Ceph File System (CephFS) exports, offering a structured approach for managing directory trees and policies across subvolumes, and enabling features like cloning and snapshots.

- *Structured Management:* They provide a structured way to manage CephFS exports, making it easier to organize and manage data.
- *Policy Enforcement:* Subvolume groups allow you to enforce policies across multiple subvolumes.
- *Cloning and Snapshots:* Subvolumes support cloning and snapshots, enabling efficient data management and disaster recovery.
- *Isolation:* CephFS volumes and subvolumes can be used to isolate tenants or different applications within the same Ceph cluster.

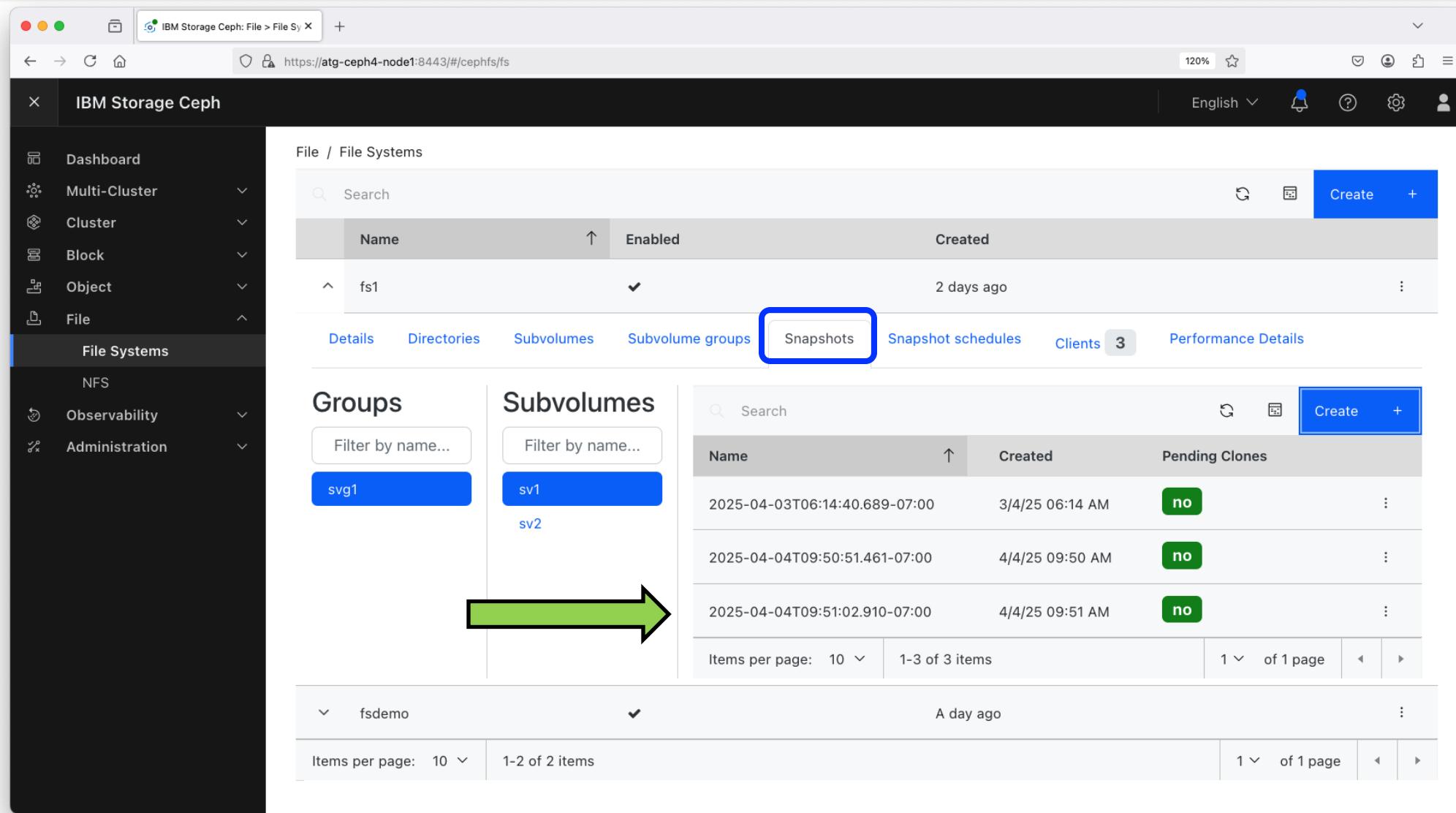
## Expand the file system (Volume) properties for Subvolume property pages



The screenshot shows the IBM Storage Ceph File System Management interface. The left sidebar is dark with white text, showing the navigation path: File > File Systems. The main area has a light background. At the top, there is a search bar and a 'Create' button. Below the search bar, there are tabs: Details, Directories, Subvolumes (which is highlighted with a blue border), Subvolume groups, Snapshots, Snapshot schedules, Clients (3), and Performance Details. The 'Subvolume groups' tab is active. On the left, under 'Groups', there is a table with one item: 'nogroup' with a subvolume 'svg1'. A large green arrow points from the 'Groups' section to the 'Subvolume groups' table on the right. The 'Subvolume groups' table has columns: Name, Data Pool, Usage, Path, Mode, and Created. It shows two subvolumes: 'sv1' and 'sv2'. Subvolume 'sv1' is associated with 'cephfs.fs1.data', has 0.01% usage, and was created a day ago. Subvolume 'sv2' is also associated with 'cephfs.fs1.data', has 0% usage, and was created a few seconds ago. The 'Mode' column for both shows 'owner: rwx', 'group: rwx', and 'others: rwx'.

Name	Data Pool	Usage	Path	Mode	Created
sv1	cephfs.fs1.data	0.01%	/volumes/.../3d4b0a73-4c98-4201-a39e-6a49393cf726	owner: rwx group: rwx others: rwx	A day ago
sv2	cephfs.fs1.data	0%	/volumes/.../9e9ddcae-180a-42b2-a9c2-787a40e1180c	owner: rwx group: rwx others: rwx	A few seconds ago

## Subvolume snapshots



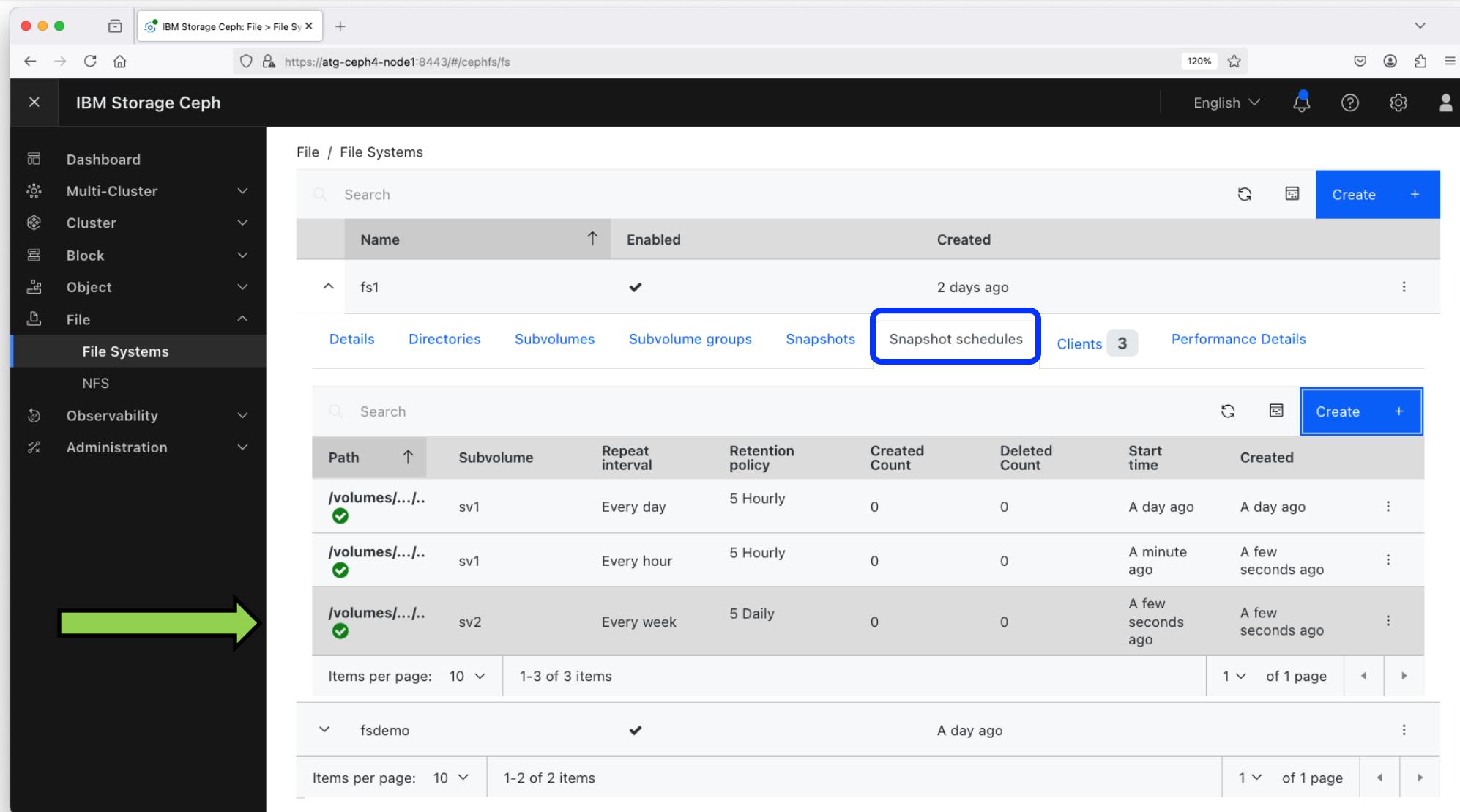
The screenshot shows the IBM Storage Ceph web interface. The left sidebar is dark with white text, showing the navigation structure. The main area is titled 'File / File Systems' and displays a table of file systems. The 'File Systems' tab is selected. Below the table are two sections: 'Groups' and 'Subvolumes'. A green arrow points from the 'Subvolumes' section to the 'Snapshots' tab in the top navigation bar, which is highlighted with a blue box. The 'Snapshots' tab is active, showing a table of snapshots for the selected subvolume 'sv1'. The table includes columns for Name, Created, and Pending Clones. The first snapshot was created on 2025-04-03T06:14:40.689-07:00. The second was on 2025-04-04T09:50:51.461-07:00. The third was on 2025-04-04T09:51:02.910-07:00. All three snapshots have a 'no' status in the Pending Clones column. The bottom of the interface shows pagination and item count information.

Name	Created	Pending Clones
2025-04-03T06:14:40.689-07:00	3/4/25 06:14 AM	no
2025-04-04T09:50:51.461-07:00	4/4/25 09:50 AM	no
2025-04-04T09:51:02.910-07:00	4/4/25 09:51 AM	no

Items per page: 10 1-3 of 3 items

Items per page: 10 1-2 of 2 items

## Subvolume snapshot schedule



The screenshot shows the IBM Storage Ceph File System interface. A green arrow on the left points to the 'File Systems' section in the sidebar. A blue box highlights the 'Snapshot schedules' tab in the top navigation bar of the main content area. The main table displays three snapshot schedules for subvolumes sv1 and sv2 under the path /volumes/.../.., with retention policies of 5 Hourly, 5 Hourly, and 5 Daily respectively. The bottom table shows two items for fsdemo.

Path	Subvolume	Repeat interval	Retention policy	Created Count	Deleted Count	Start time	Created
/volumes/.../..	sv1	Every day	5 Hourly	0	0	A day ago	A day ago
/volumes/.../..	sv1	Every hour	5 Hourly	0	0	A minute ago	A few seconds ago
/volumes/.../..	sv2	Every week	5 Daily	0	0	A few seconds ago	A few seconds ago

Items per page: 10 ▾ 1-3 of 3 items

Path	Subvolume	Repeat interval	Retention policy	Created Count	Deleted Count	Start time	Created
/volumes/.../..	fsdemo	Every day	5 Daily	0	0	A day ago	A day ago

Items per page: 10 ▾ 1-2 of 2 items

## The Ceph File System snapshot client experience

```
[root@local ~]# cd /mnt/fs1/dir1
[root@local ~]# ls -al
total 10241
drwxr-xr-x. 2 root root 3 Apr 3 09:14 .
drwxrwxrwx. 4 root root 2 Apr 3 09:13 ..
-rw-r--r--. 1 root root 10485760 Apr 3 09:14 10MB.dat
-rw-r--r--. 1 root root 0 Apr 3 09:13 atouchfile
-rw-r--r--. 1 root root 14 Apr 3 09:14 hello-world.txt
[root@local ~]# cd .snap
[root@local ~]# ls -al
total 0
drwxr-xr-x 2315609098 root root 0 Apr 3 09:14 .
drwxr-xr-x. 2 root root 3 Apr 3 09:14 ..
drwxr-xr-x. 2 root root 3 Apr 3 09:14 _2025-04-03T06:14:40.689-07:00_1099511627778
drwxr-xr-x. 2 root root 3 Apr 3 09:14 _2025-04-04T09:50:51.461-07:00_1099511627778
drwxr-xr-x. 2 root root 3 Apr 3 09:14 _2025-04-04T09:51:02.910-07:00_1099511627778
drwxr-xr-x. 2 root root 3 Apr 3 09:14 _scheduled-2025-04-04-
17_53_36_UTC_1099511627778
```

## The Ceph File System snapshot client experience

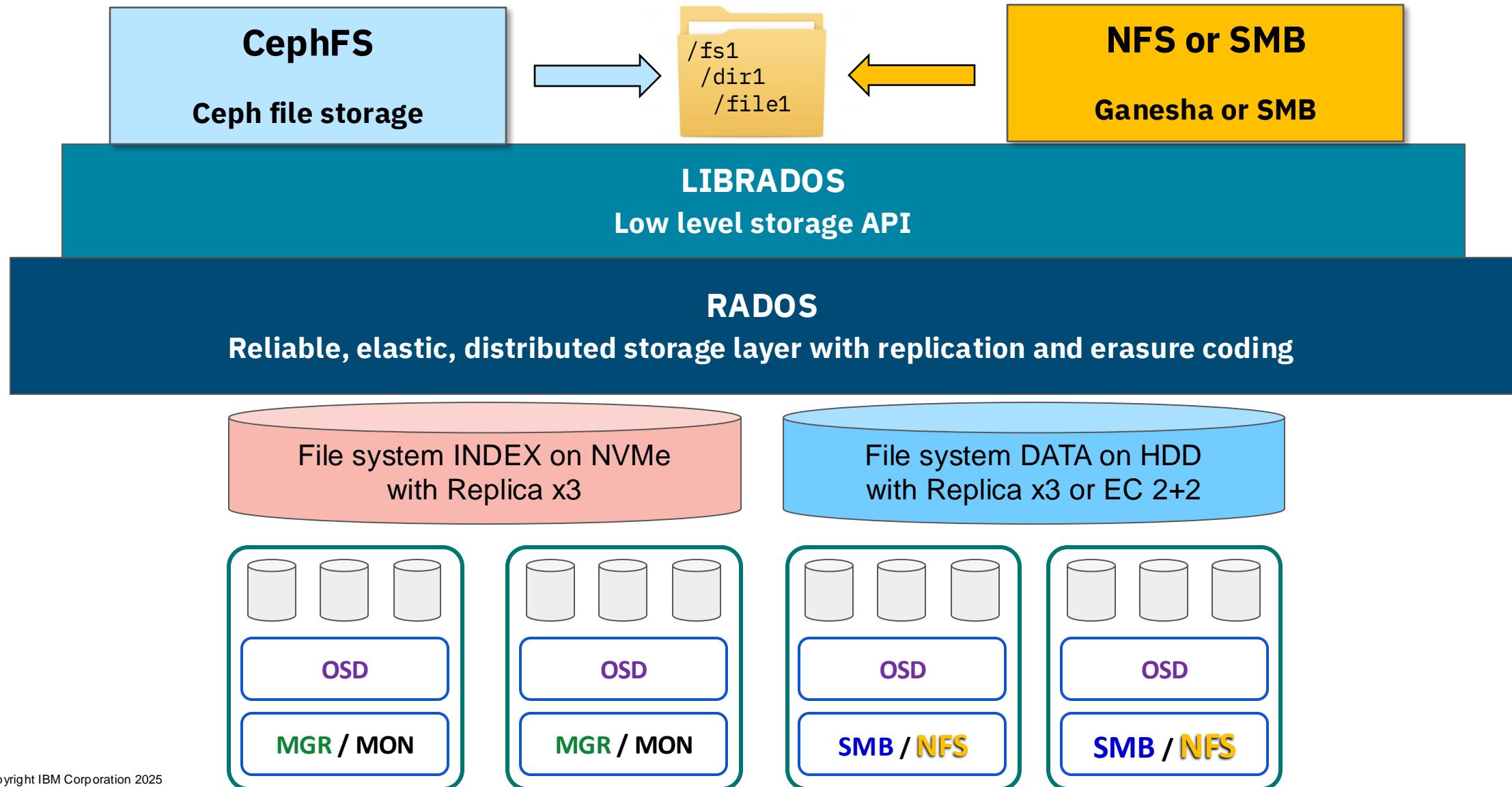
```
[root@local ~]# cd /mnt/fs1/dir1
[root@local ~]# cd .snap
[root@local ~]# tree
.
├── _2025-04-03T06:14:40.689-07:00_1099511627778
│   ├── 10MB.dat
│   ├── atouchfile
│   └── hello-world.txt
└── _2025-04-04T09:50:51.461-07:00_1099511627778
    ├── 10MB.dat
    ├── atouchfile
    └── hello-world.txt
└── _2025-04-04T09:51:02.910-07:00_1099511627778
    ├── 10MB.dat
    ├── atouchfile
    └── hello-world.txt
```

## Ceph file system (CephFS) day 2 operations

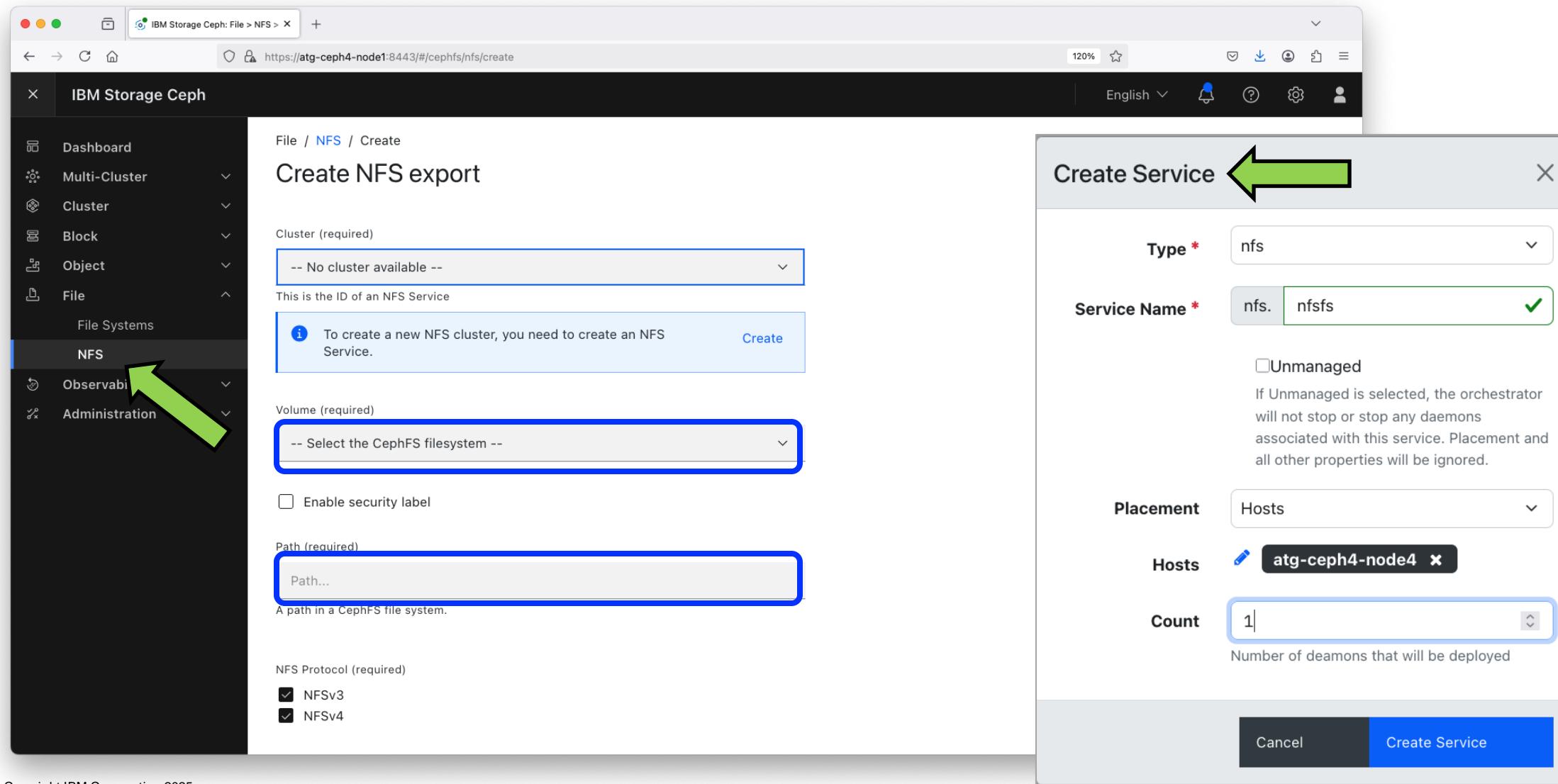
---



## Day 2: Expanding the Ceph File System access via NFS and SMB



## Create an NFS export



IBM Storage Ceph

File / NFS / Create

### Create NFS export

Cluster (required)

-- No cluster available --

This is the ID of an NFS Service

To create a new NFS cluster, you need to create an NFS Service. [Create](#)

Volume (required)

-- Select the CephFS filesystem --

Enable security label

Path (required)

Path...

A path in a CephFS file system.

NFS Protocol (required)

NFSv3

NFSv4

Create Service

Type \* nfs

Service Name \* nfs nfsfs

Unmanaged

If Unmanaged is selected, the orchestrator will not stop or start any daemons associated with this service. Placement and all other properties will be ignored.

Placement

Hosts

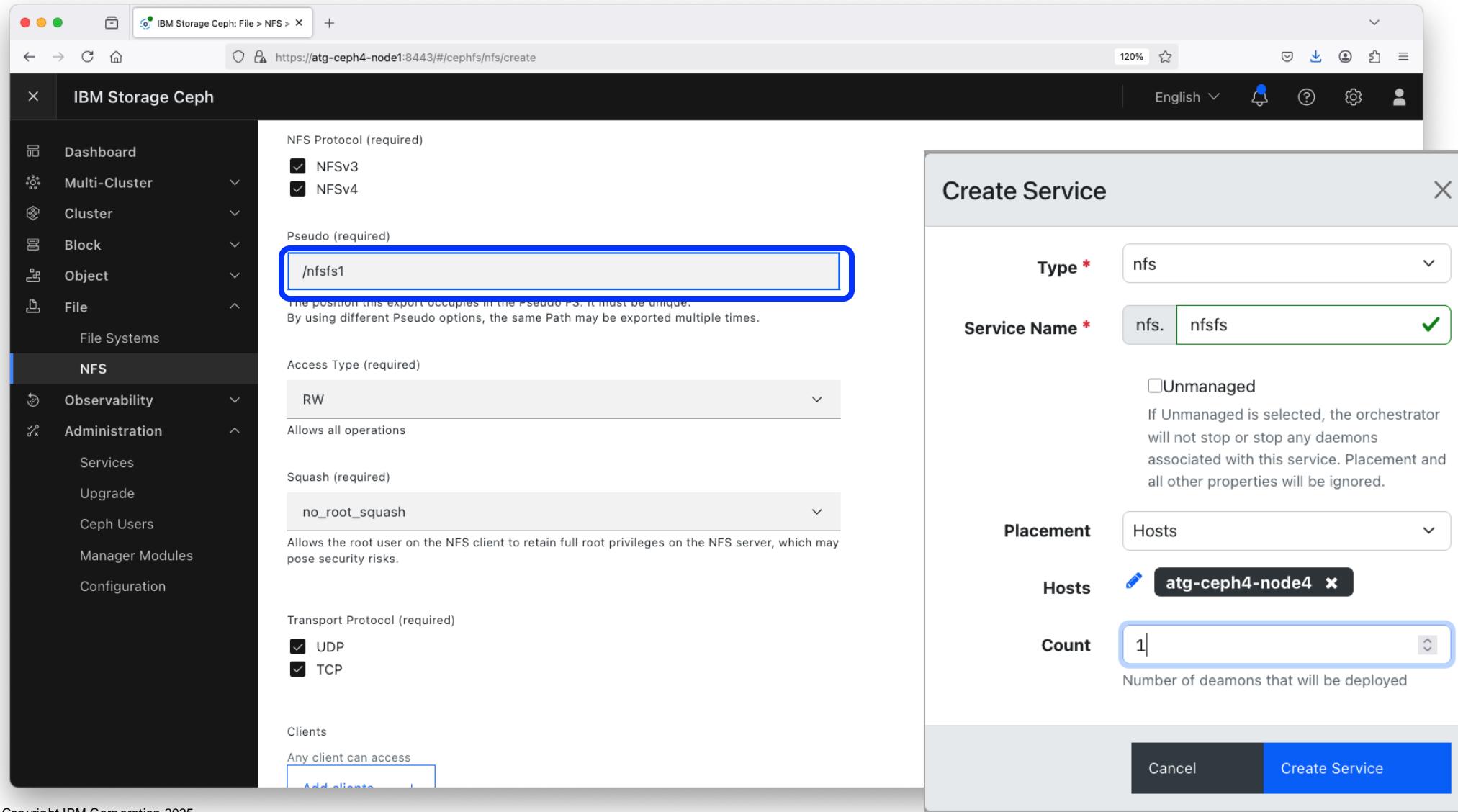
Hosts atg-ceph4-node4

Count 1

Number of daemons that will be deployed

Cancel Create Service

## Ceph NFS export creation dialogue



The image shows a screenshot of the IBM Storage Ceph interface. On the left, the navigation sidebar is visible with options like Dashboard, Multi-Cluster, Cluster, Block, Object, File, and NFS. The NFS option is currently selected. The main content area shows the 'Create NFS Export' form. The 'Pseudo' field is highlighted with a blue border and contains the value '/nfsfs1'. A note below it states: 'The position this export occupies in the Pseudo FS. It must be unique. By using different Pseudo options, the same Path may be exported multiple times.' Other fields include 'Access Type' (set to 'RW'), 'Squash' (set to 'no\_root\_squash'), 'Transport Protocol' (checkboxes for 'UDP' and 'TCP' are checked), and a 'Clients' section with a 'Add clients' button. On the right, a separate 'Create Service' dialog is open. It has fields for 'Type' (set to 'nfs'), 'Service Name' (set to 'nfs.', with 'nfsfs' selected and checked), 'Unmanaged' (unchecked), 'Placement' (set to 'Hosts'), 'Hosts' (set to 'atg-ceph4-node4'), and 'Count' (set to '1'). At the bottom are 'Cancel' and 'Create Service' buttons.

IBM Storage Ceph: File > NFS > +

https://atg-ceph4-node1:8443/#/cephfs/nfs/create

English

Dashboard

Multi-Cluster

Cluster

Block

Object

File

File Systems

NFS

Observability

Administration

Services

Upgrade

Ceph Users

Manager Modules

Configuration

NFS Protocol (required)

NFSv3

NFSv4

Pseudo (required)

/nfsfs1

The position this export occupies in the Pseudo FS. It must be unique. By using different Pseudo options, the same Path may be exported multiple times.

Access Type (required)

RW

Allows all operations

Squash (required)

no\_root\_squash

Allows the root user on the NFS client to retain full root privileges on the NFS server, which may pose security risks.

Transport Protocol (required)

UDP

TCP

Clients

Any client can access

Add clients

Create Service

Type \* nfs

Service Name \* nfs. nfsfs

Unmanaged

If Unmanaged is selected, the orchestrator will not stop or start any daemons associated with this service. Placement and all other properties will be ignored.

Placement Hosts

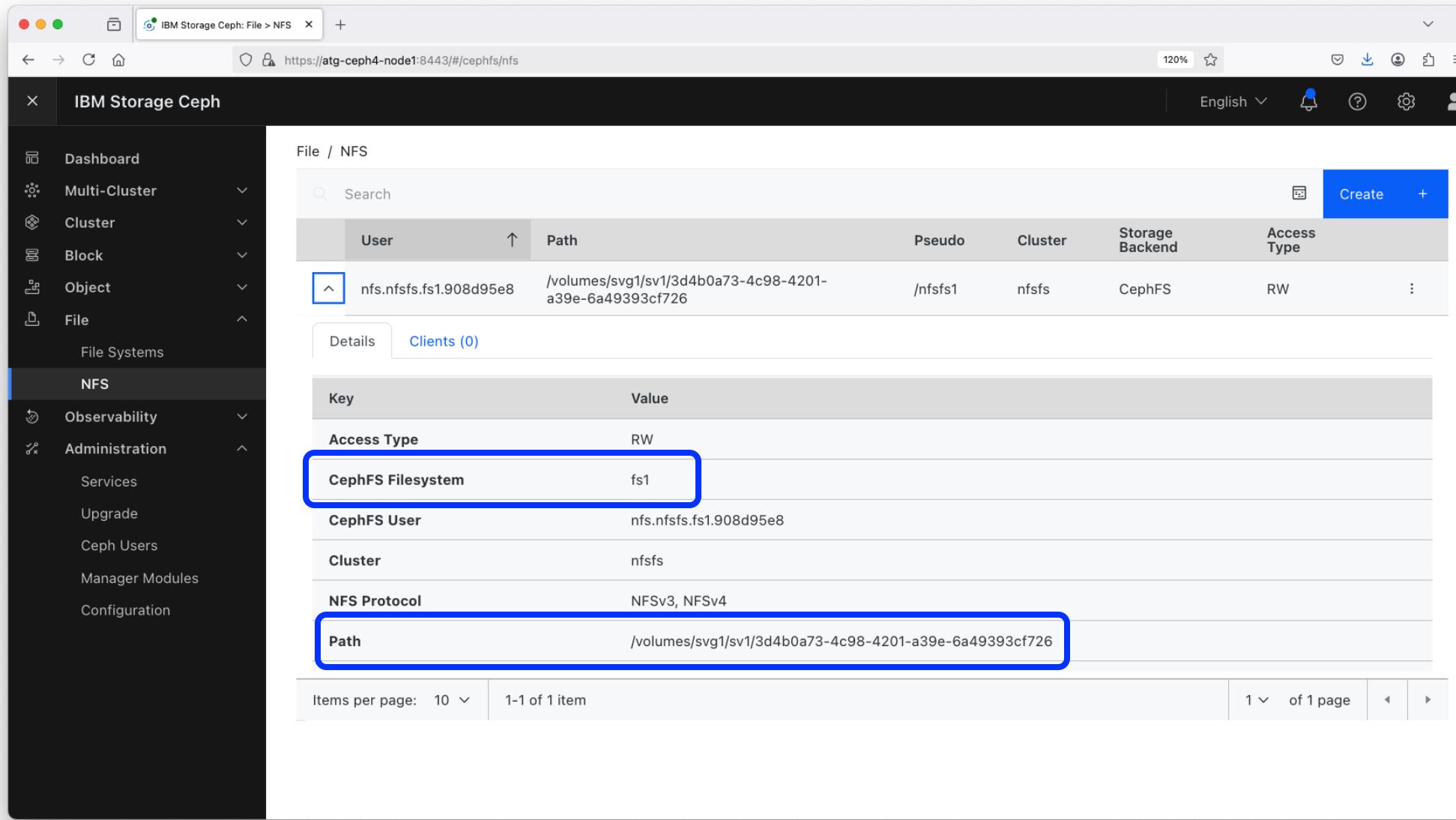
Hosts atg-ceph4-node4

Count 1

Number of daemons that will be deployed

Cancel Create Service

## Ceph File System NFS export properties



The screenshot shows the IBM Storage Ceph web interface for managing NFS exports. The left sidebar is a navigation menu with the following structure:

- Dashboard
- Multi-Cluster
- Cluster
- Block
- Object
- File
  - File Systems
  - NFS** (selected)
  - Observability
  - Administration
  - Services
  - Upgrade
  - Ceph Users
  - Manager Modules
  - Configuration

The main content area is titled "File / NFS" and shows a table of exports. One export is listed:

User	Path	Pseudo	Cluster	Storage Backend	Access Type
nfs.nfsfs.fs1.908d95e8	/volumes/svg1/sv1/3d4b0a73-4c98-4201-a39e-6a49393cf726	/nfsfs1	nfsfs	CephFS	RW

Below the table, there are two tabs: "Details" (selected) and "Clients (0)". The "Details" tab displays the following properties:

Key	Value
Access Type	RW
CephFS Filesystem	fs1
CephFS User	nfs.nfsfs.fs1.908d95e8
Cluster	nfsfs
NFS Protocol	NFSv3, NFSv4
Path	/volumes/svg1/sv1/3d4b0a73-4c98-4201-a39e-6a49393cf726

The "Path" and "CephFS Filesystem" fields are highlighted with a blue border.

## The Ceph NFS client experience

```
[root@local ~]# mkdir /mnt/nfsfs1
[root@local ~]# mount -t nfs atg-ceph4-node4:/nfsfs1 /mnt/nfsfs1
[root@local ~]# tree /mnt/nfsfs1
/mnt/nfsfs1
└── dir1
    ├── 10MB.dat
    ├── atouchfile
    └── hello-world.txt
└── dir2

[root@local ~]# ls -al /mnt/nfsfs1
total 10242
drwxr-xr-x. 2 root root 10485774 Apr  3 09:14 .
drwxrwxrwx. 4 root root 10485774 Apr  3 09:13 ..
-rw-r--r--. 1 root root 10485760 Apr  3 09:14 10MB.dat
-rw-r--r--. 1 root root          0 Apr  3 09:13 atouchfile
-rw-r--r--. 1 root root        14 Apr  3 09:14 hello-world.txt2 directories, 3 files
```

## The Ceph File System snapshot client experience

```
[root@local ~]# cd /mnt/fs1/dir1
[root@local ~]# cd .snap
[root@local ~]# tree
.
├── _2025-04-03T06:14:40.689-07:00_1099511627778
│   ├── 10MB.dat
│   ├── atouchfile
│   └── hello-world.txt
└── _2025-04-04T09:50:51.461-07:00_1099511627778
    ├── 10MB.dat
    ├── atouchfile
    └── hello-world.txt
└── _2025-04-04T09:51:02.910-07:00_1099511627778
    ├── 10MB.dat
    ├── atouchfile
    └── hello-world.txt
```

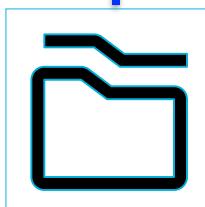
## SMB access on CephFS tech preview



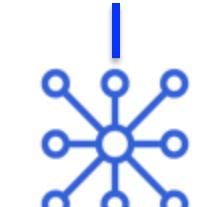
MS Windows  
File Sharing



SMB  
for CephFS



CephFS  
File Storage



IBM Storage Ceph  
Cluster

### Server Message Block file share service

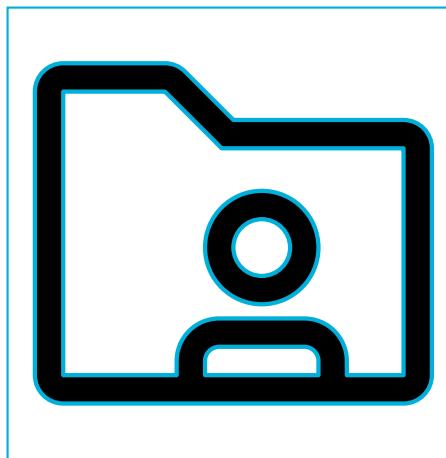
- IBM Storage Ceph CephFS volume file access by SMB protocol.
- SMB v2 and v3
- Management of shares, permissions, etc. through Ceph Manager
- Support Active Directory integration and
- Windows ACLs

### Server Message Block service properties

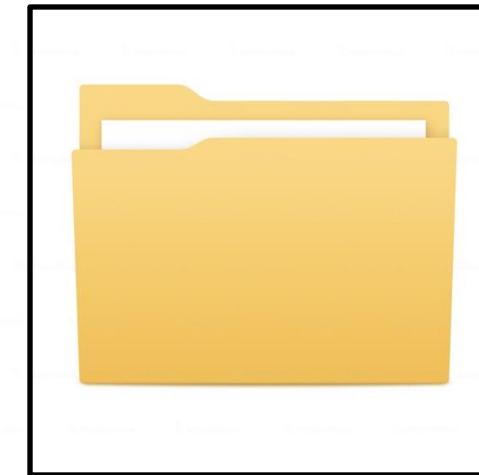
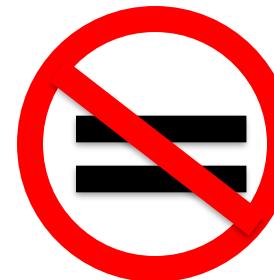
- Provides compatibility with the SMB protocol to allow Windows applications to store file data on IBM Storage Ceph.
- Expose CephFS snapshots via SMB through a dedicated ".snapshot" directory
- Resilient against storage node failure.

## Ceph SMB positioning

SMB is *FILE SHARING* for CephFS targeting application modernization and expansion



SMB  
for CephFS



Enterprise Network  
Attached Storage  
(NAS)

# IBM Storage Ceph SMB demonstration video

IBM MediaCenter [Explore More](#) [Learn & Support](#) [Product & Solutions](#) [Services and Consulting](#)

## Ceph SMB Demonstration

Advanced Technology Group

IBM

**ATG Level Up Video**

Ceph SMB Demonstration

John Shubeck – Senior Storage Technical Specialist

Date: December, 2024

0:01 / 9:50

Details Share

Search

ACTIONS

Related Media

ATG Level Up Video

IBM Storage Ceph HAProxy for Object Storage – Day 1

John Shubeck – Senior Storage Technical Specialist

11:06

Ceph HAProxy for Object - Day 1

ATG Level Up Video

IBM Storage Ceph HAProxy for Object Storage – Day 2

John Shubeck – Senior Storage Technical Specialist

10:07

Ceph HAProxy for Object - Day 2

Advanced Technology Group

IBM

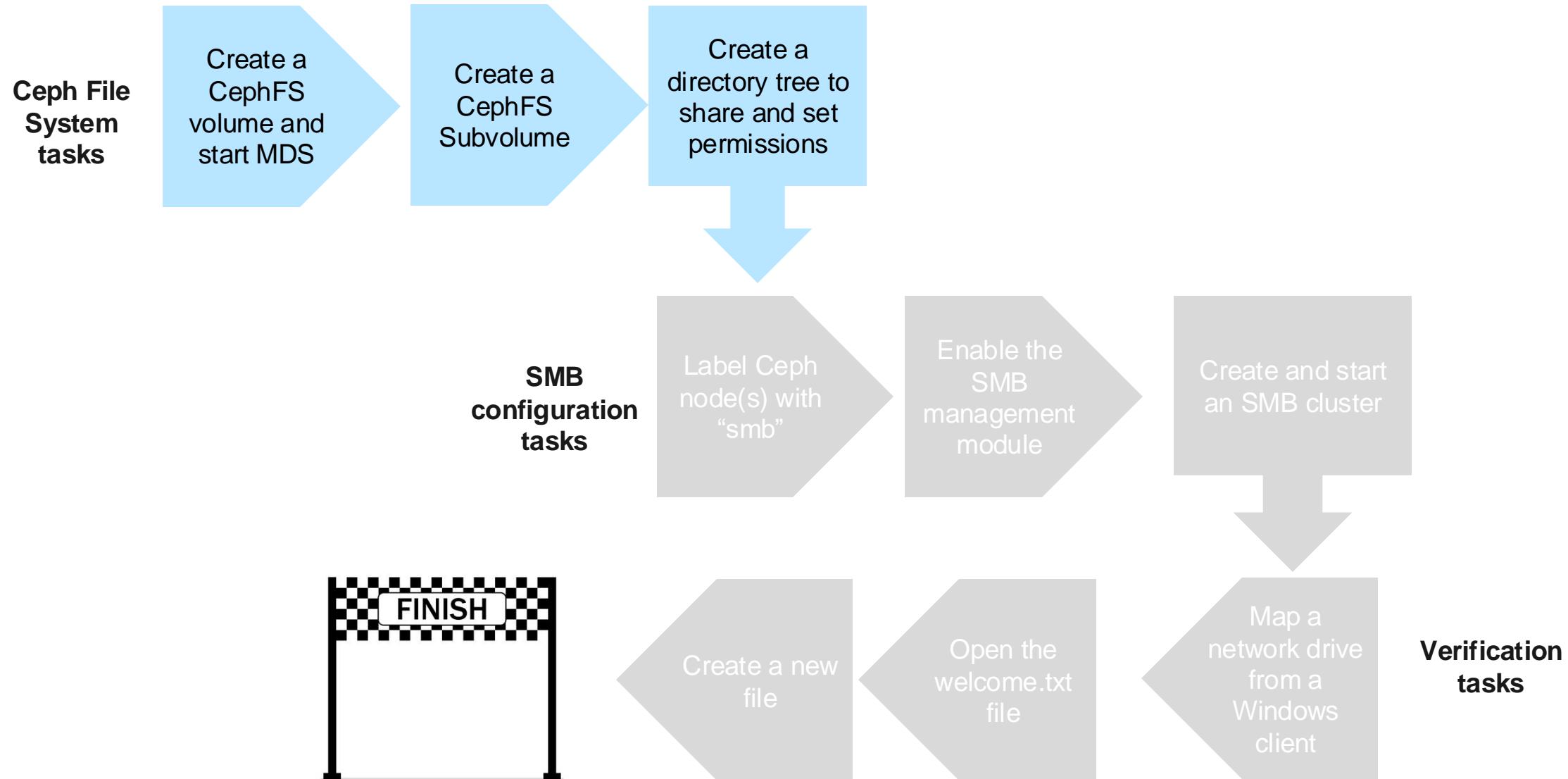
IBM

IBM

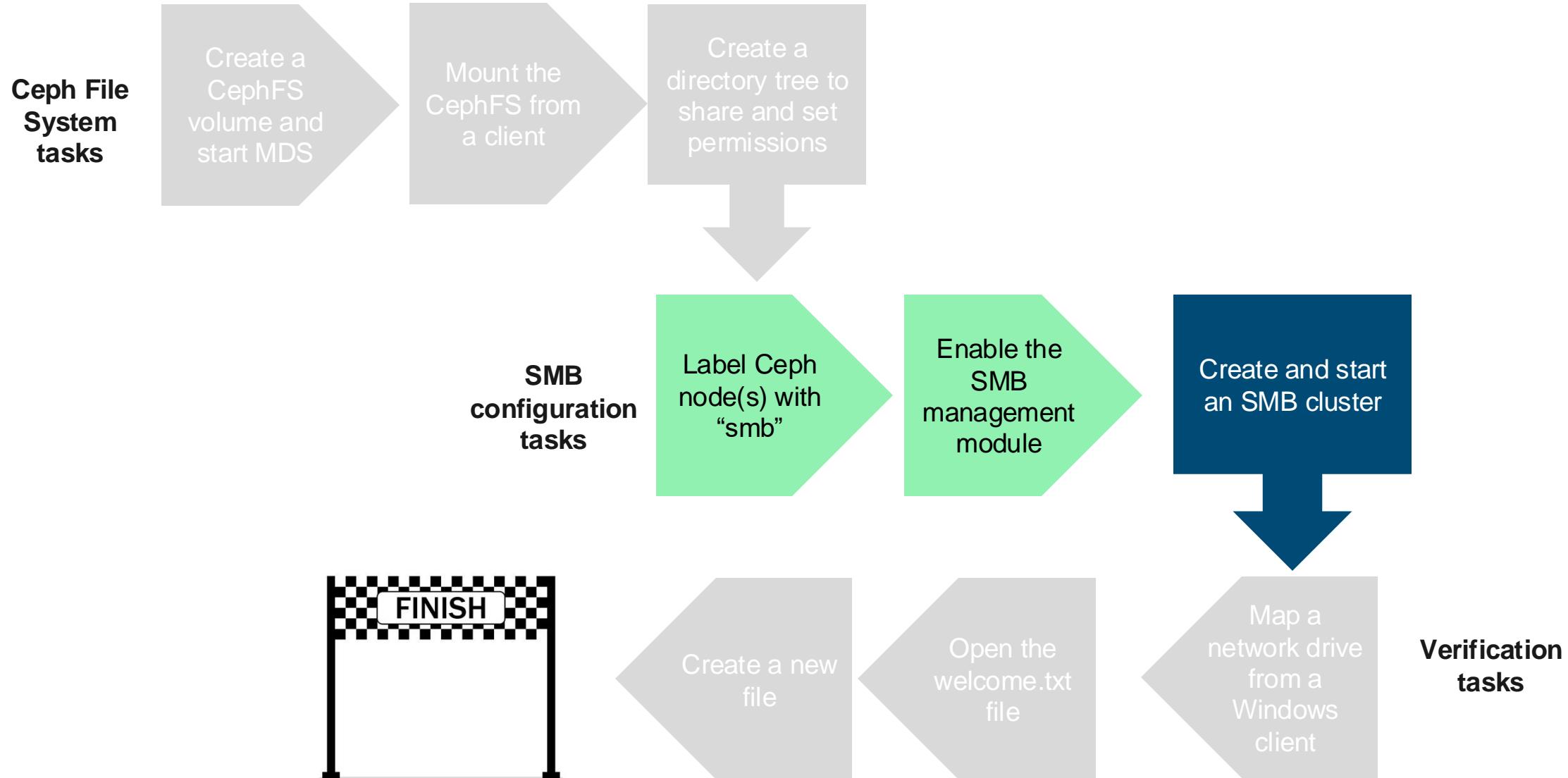
IBM

[https://mediacenter.ibm.com/media/1\\_hpdtq0tn](https://mediacenter.ibm.com/media/1_hpdtq0tn)

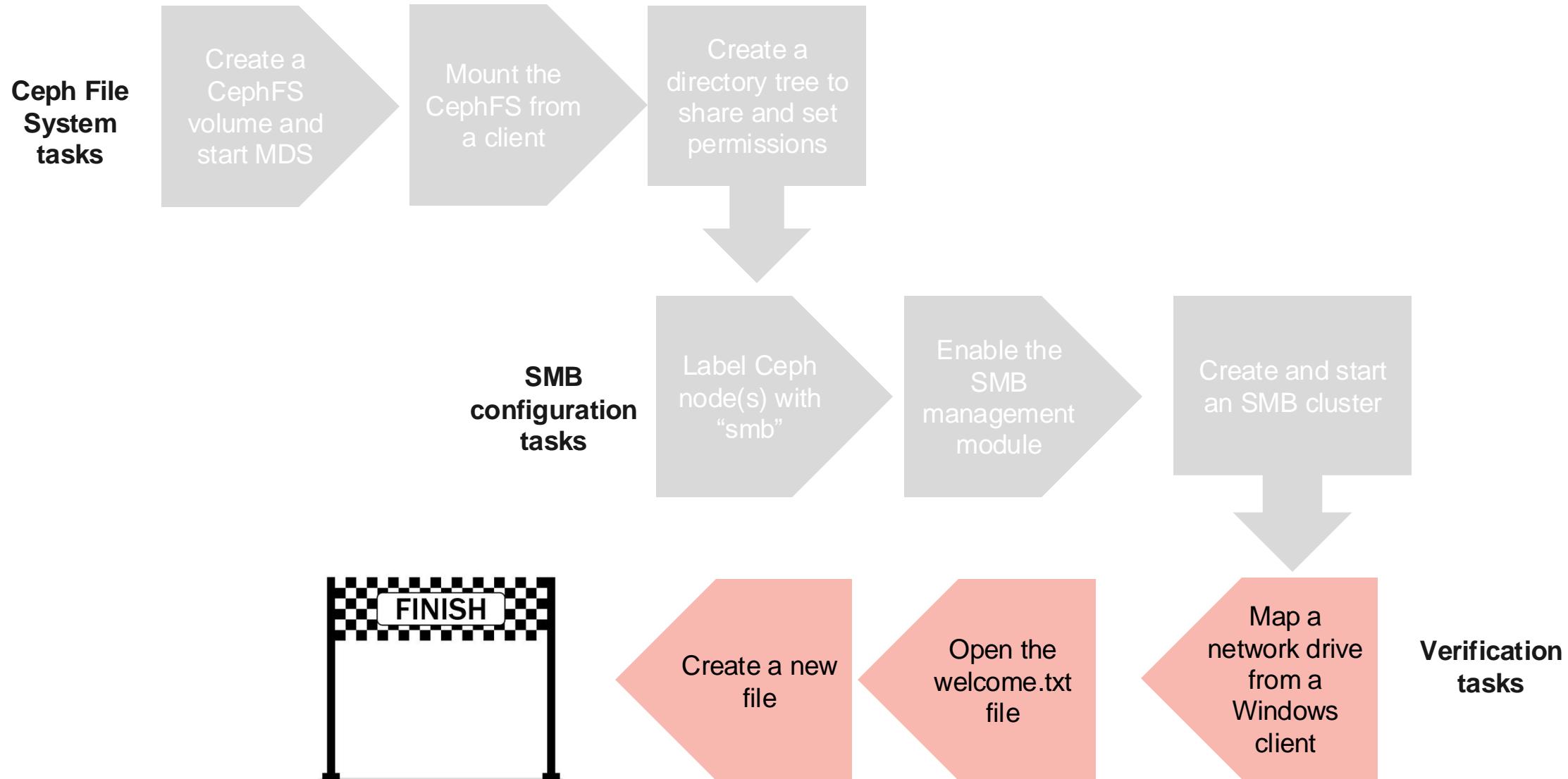
## IBM Storage Ceph SMB configuration phases and tasks



## IBM Storage Ceph SMB configuration phases and tasks



## IBM Storage Ceph SMB configuration phases and tasks



## SMB cluster management – imperative method

```
# Label nodes
# ceph orch host label add atg-ceph4-node2 smb
# 

# Enable the smb management module
# ceph mgr module enable smb
# 

# Create the smb cluster instance
# ceph smb cluster create smbdemo active-directory \
--domain_realm sdi.dmz \
--domain_join_user_pass domain-admin%domain-password \
--placement label:smb \
--clustering default

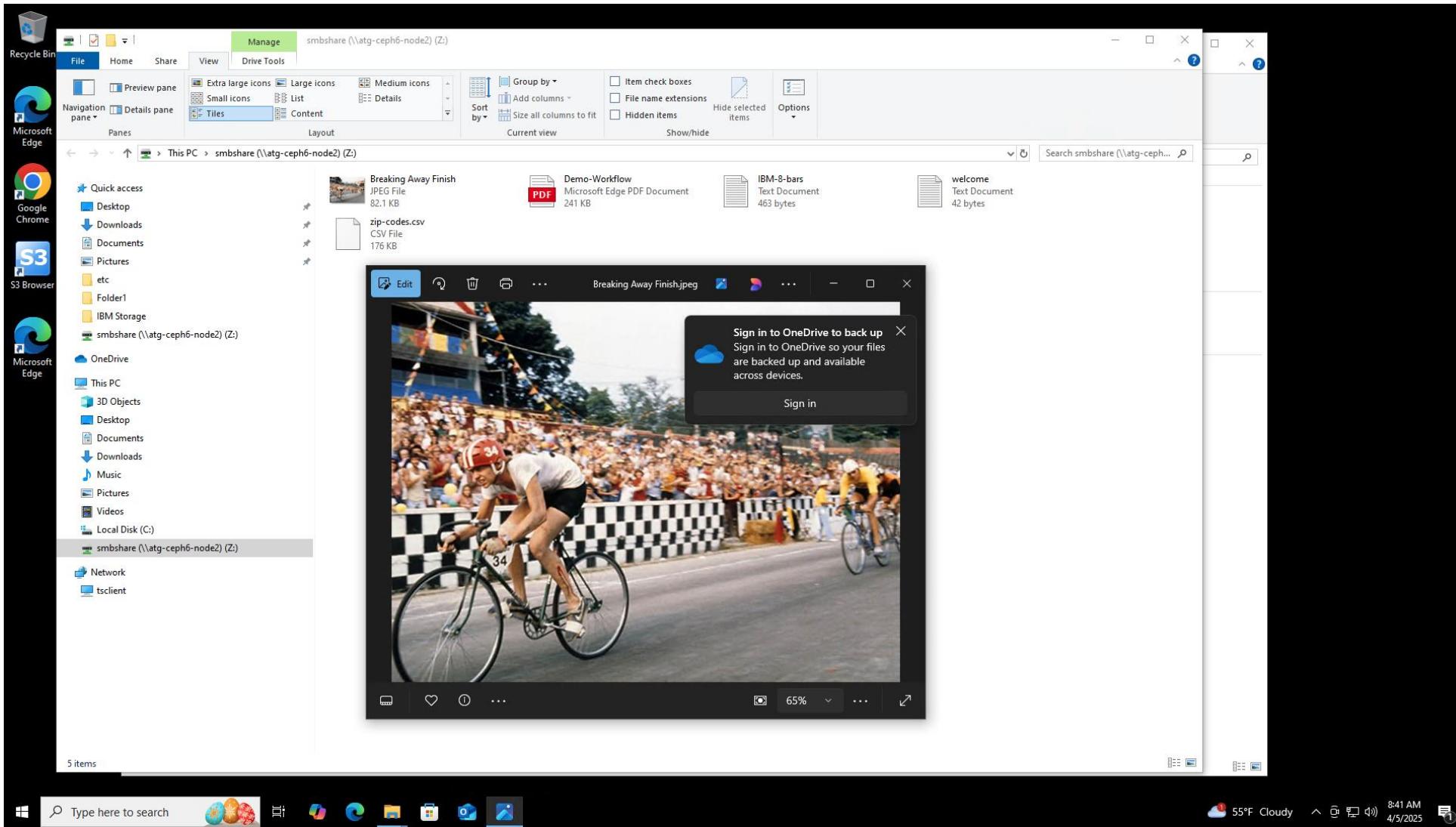
"state": "created",
"success": true
```

## SMB cluster management – create a share

---

```
# ceph smb share create smbdemo smbshare fs1 \
/volumes/svg1/sv1/3d4b0a73-4c98-4201-a39e-6a49393cf726
{
  "resource": {
    "resource_type": "ceph.smb.share",
    "cluster_id": "smbdemo",
    "share_id": "smbshare",
    "intent": "present",
    "name": "smbshare",
    "readonly": false,
    "browseable": true,
    "cephfs": {
      "volume": "fs1",
      "path": "/volumes/svg1/sv1/3d4b0a73-4c98-4201-a39e-6a49393cf726",
      "provider": "samba-vfs"
    }
  },
  "state": "created",
  "success": true
}
```

## The Ceph SMB client experience (Microsoft Windows)



## SMB cluster management – display share information (workaround)

```
# ceph smb share create smbdemo smbshare fs1 myshare
{
  "resource": {
    "resource_type": "ceph.smb.share",
    "cluster_id": "smbdemo",
    "share_id": "myshare",
    "intent": "present",
    "name": "myshare",
    "readonly": false,
    "browseable": true,
    "cephfs": {
      "volume": "fsdemo",
      "path": "myshare",
      "provider": "samba-vfs"
    }
  },
  "msg": "a resource with the same ID already exists",
  "success": false
}
```

## Ceph SMB configuration commands

```
# Ceph smb configuration property listings

# ceph smb cluster ls
# ceph smb show

# ceph smb share ls smbdemo
# ceph smb share create smbdemo smbshare fs1 myshare # Error response but displays share info

# ceph smb dump cluster-config smbdemo

# ceph smb dump service-spec smbdemo

# ceph smb dump service-spec smbdemo --format yaml
```

## Day 3: Ceph File System advanced commands (cephadm)

---



## Creating CephFS with a placement specification

---

For more control over the deploying process, manually create the pools that are associated to the CephFS, start the MDS service on the hosts, and create the CephFS file system.

A CephFS file system requires at least two pools, one to store CephFS data, and another to store CephFS metadata. These can be created individually on different OSDs using different protection policies

```
[ceph: root@node1 /]# ceph osd pool create cephfs_data
[ceph: root@node1 /]# ceph osd pool create cephfs_metadata
[ceph: root@node1 /]# ceph osd pool set cephfs_metadata crush_rule crush_nvme
```

## Creating CephFS with a placement specification (continued)

Create a new Ceph File System mapping to the custom metadata and data pools

```
[ceph: root@node1 /]# ceph fs new mycephfs cephfs_metadata cephfs_data
new fs with metadata pool 14 and data pool 16
```

Discrete control of MDS daemons using host placement

```
[ceph: root@node1 /]# ceph orch apply mds mycephfs --placement="2 ceph-node3 ceph-node4"
Scheduled mds.mycephfs update...
```

## Creating Ceph file system snapshots (part 1)

---

CephFS enables asynchronous snapshots by default when deploying Ceph. These snapshots are stored in a hidden directory called `.snap`.

Use `cephfs` set to enable user-initiated snapshot creation for an existing CephFS file system.

```
[ceph: root@node1 /]# ceph fs set fsdemo allow_new_snaps true
enabled new snapshots
```

To create a snapshot, first mount the CephFS file system on your client node. Then, create a subdirectory inside the `.snap` directory. The snapshot name is the new subdirectory name. This snapshot contains a copy of all the current files in the CephFS file system.

```
[root@client ~]# mkdir /mnt/fsdemo/.snap/snapshot-1
```

## Creating Ceph file system snapshots (conclusion)

1. Authorize a particular client to make snapshots for the CephFS file system (optional)

```
[ceph: root@node1 /]# ceph fs authorize mycephfs client.john / rws
```

2. To restore a file, copy the file or a wild card from the snapshot directory to another directory.

```
[root@client ~]# cp /mnt/fsdemo/.snap/snapshot-1/dir1/IBM-8-bars.txt .
```

3. To fully restore a snapshot from the .snap directory tree, replace the normal entries with copies from the chosen snapshot using the “-a” option.

```
[root@client ~]# rm -rf *
[root@client ~]# cp -rf /mnt/fsdemo/.snap/snapshot-1/dir1/* .
```

## Automated snapshot scheduling

---

You can use CephFS to schedule snapshots. The `snap_schedule` module manages the scheduled snapshots. You can use this module to create and delete snapshot schedules. Snapshot schedule information is stored in the CephFS metadata pool.

To create a snapshot schedule, first enable the `snap_schedule` module on the MGR node.

```
[ceph: root@node1 /]# ceph mgr module enable snap_schedule
```

Then, add the new snapshot schedule

```
[ceph: root@node1 /]# ceph fs snap-schedule add / 1d -fs fs1
Schedule set for path /
```

You can list the snapshot schedules for a path with the list option

```
[ceph: root@node1 /]# ceph fs snap-schedule list / --recursive
/volumes/svg1/sv1/0a392abb-0677-4fda-9f80-9b202191655f/.. 1h
/volumes/svg1/sv1/0a392abb-0677-4fda-9f80-9b202191655f/.. 1d
/volumes/svg1/sv1/0a392abb-0677-4fda-9f80-9b202191655f/.. 1w
```

## Automated snapshot scheduling (conclusion)

On the client node, review the snapshots in the .snap folder on your mounted CephFS

```
[root@client ~]# ls /mnt/mycephfs/.snap
scheduled-2021-10-06-08_00_00
scheduled-2021-10-06-09_00_00
scheduled-2021-10-06-10_00_00
```

Remove a snapshot schedule by specifying the path

```
[ceph: root@node1 /]# ceph fs snap-schedule remove /mycephfs
```

## Removing a Ceph File System

Down the file system

```
[ceph: /]# ceph fs set fsdemo down true
fsdemo marked down.
```

Review file system status

```
[ceph: /]# ceph fs status
fsdemo - 0 clients
=====
      POOL      TYPE    USED   AVAIL
cephfs.fsdemo.meta  metadata  7920k  79.4G
cephfs.fsdemo.data    data      0  79.4G
. . . output omitted . . .
```

Remove the file system

```
[ceph: /]# ceph fs rm fsdemo --yes-i-really-mean-it
```

## Day 3 and ½: Topics for a future Accelerate with ATG webinar

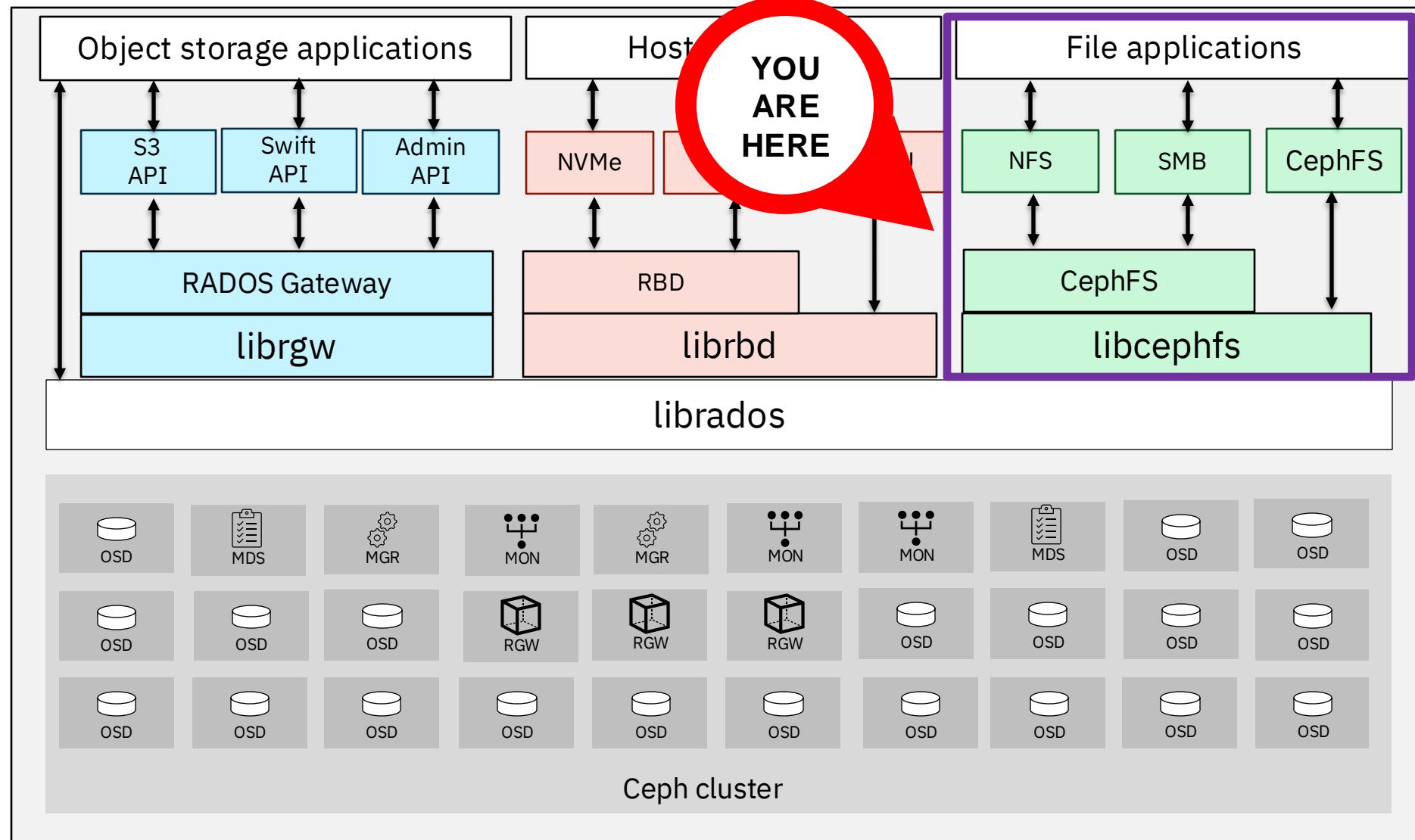
---

- CephFS device classes end to end
- CephFS snapshot clones
- CephFS snapshot mirror
- CephFS performance tips
- Performance visualization
  
- And. . . .

IN CONCLUSION



## Building out an IBM Storage Ceph ecosystem for file services



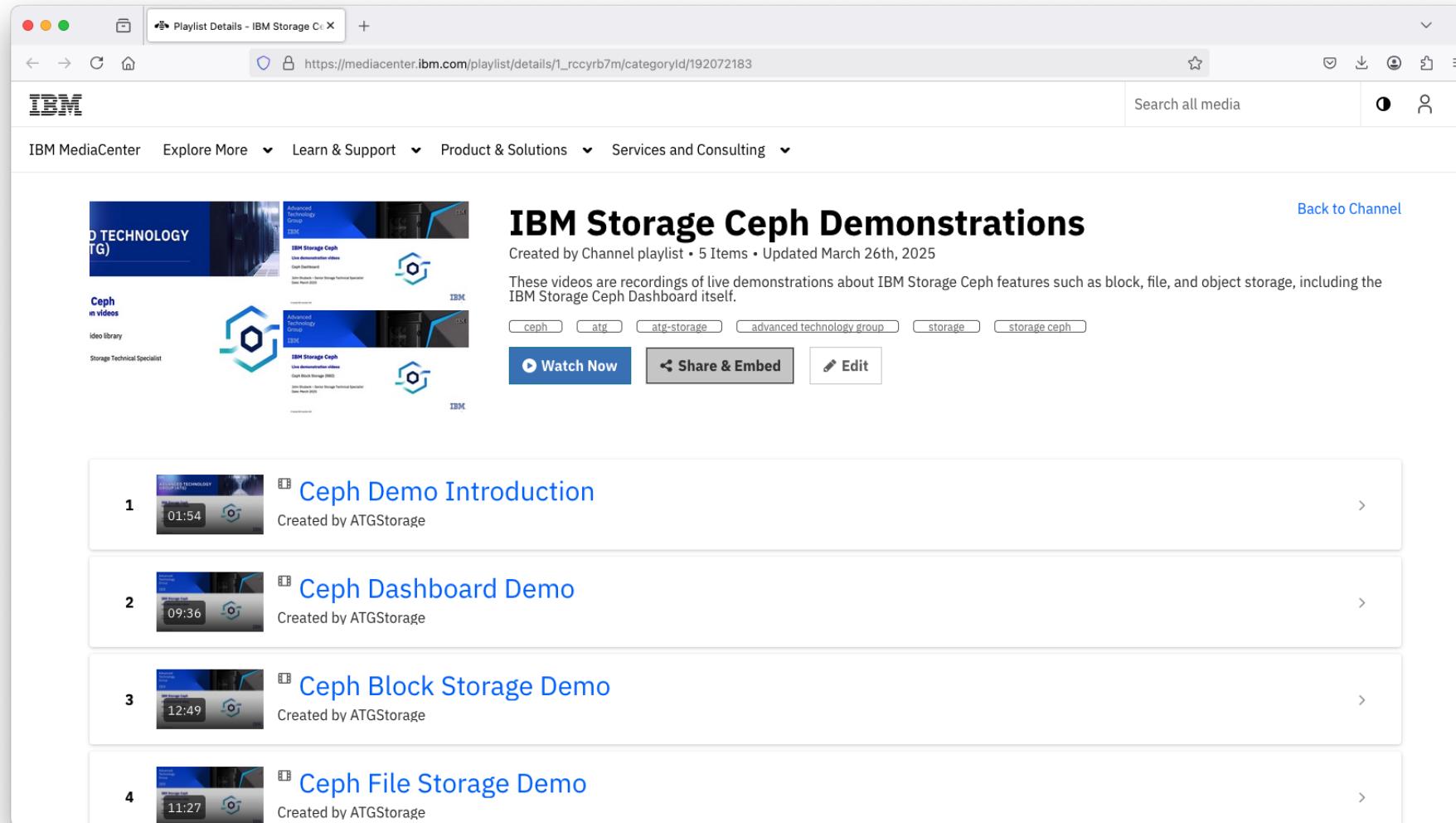
## IBM Storage Ceph learning resources

---



## New IBM Storage Ceph demonstrations in IBM Mediabcenter

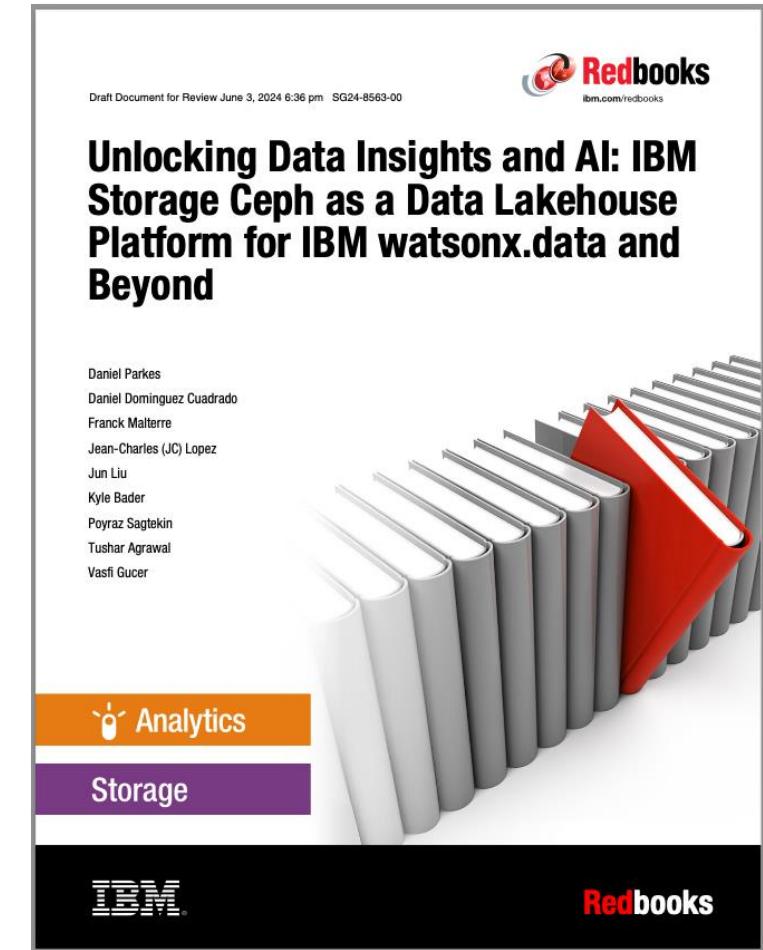
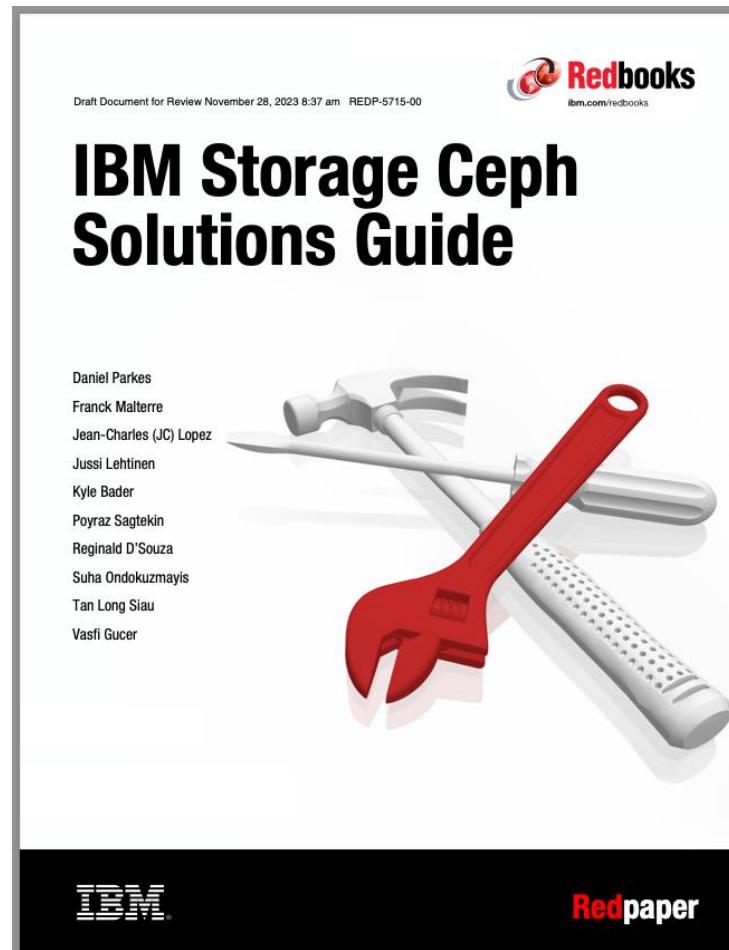
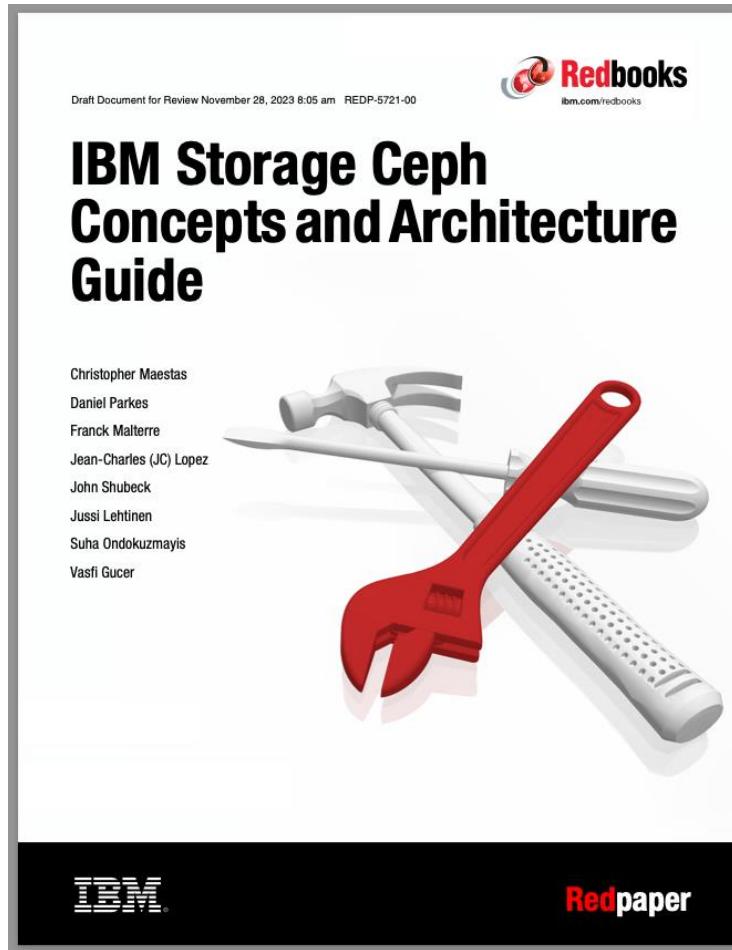
DELIVERED



The screenshot shows a web browser window for the IBM Mediabcenter. The URL in the address bar is [https://mediabcenter.ibm.com/playlist/details/1\\_rccyrb7m/categoryId/192072183](https://mediabcenter.ibm.com/playlist/details/1_rccyrb7m/categoryId/192072183). The page title is "IBM Storage Ceph Demonstrations". The page content includes a thumbnail for the "Ceph Demo Introduction" video, which is 1:54 long and created by ATGStorage. Below it are three more video thumbnails for "Ceph Dashboard Demo" (09:36), "Ceph Block Storage Demo" (12:49), and "Ceph File Storage Demo" (11:27), all created by ATGStorage. The page also features a sidebar with navigation links for "IBM MediaCenter", "Explore More", "Learn & Support", "Product & Solutions", and "Services and Consulting".

[https://mediabcenter.ibm.com/playlist/details/1\\_rccyrb7m/categoryId/192072183](https://mediabcenter.ibm.com/playlist/details/1_rccyrb7m/categoryId/192072183)

## IBM Redbooks for IBM Storage Ceph



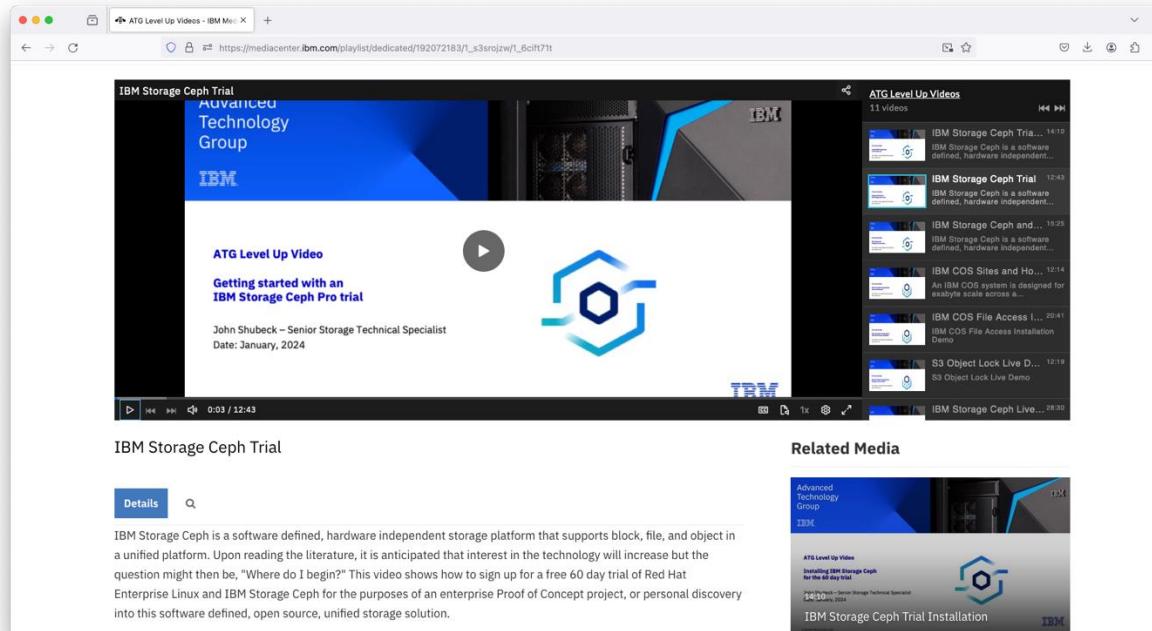
<https://www.redbooks.ibm.com/abstracts/redp5721.html>

<https://www.redbooks.ibm.com/abstracts/redp5715.html>

<https://www.redbooks.ibm.com/abstracts/sg248563.html>

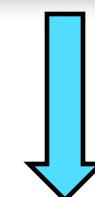
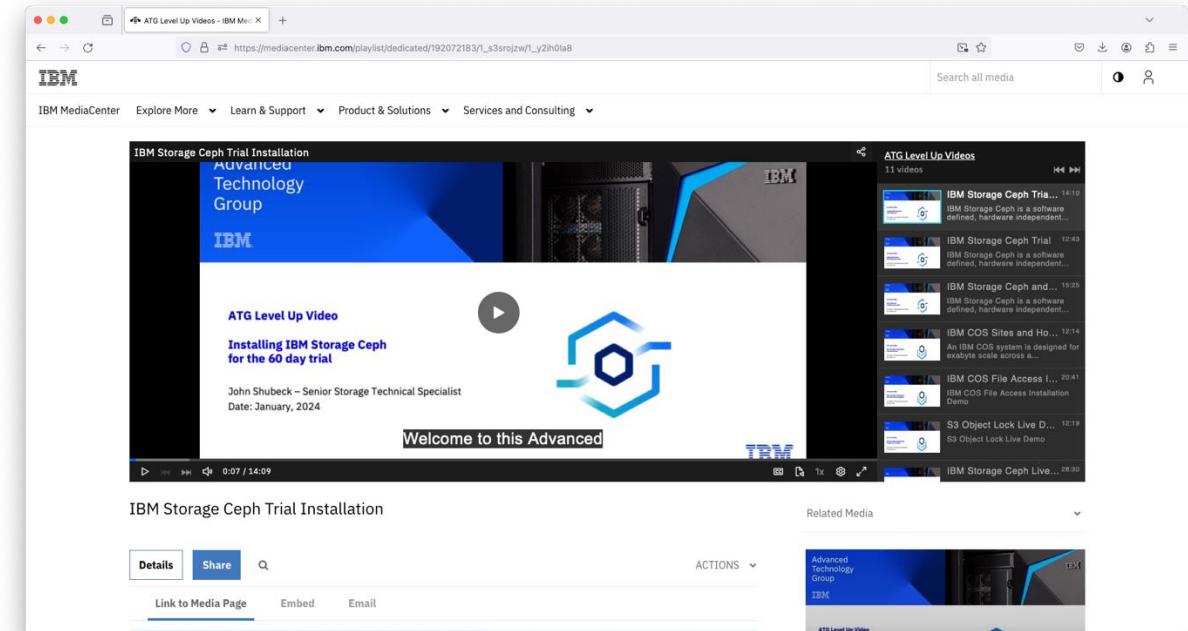
## IBM Storage Ceph Trial videos

### How to sign up and get subscriptions



[https://mediacenter.ibm.com/media/IBM+Storage+Ceph+Trial/1\\_6cift71t](https://mediacenter.ibm.com/media/IBM+Storage+Ceph+Trial/1_6cift71t)

### How to install a POC cluster



[https://mediacenter.ibm.com/media/IBM+Storage+Ceph+Trial+Installation/1\\_y2ih0la8](https://mediacenter.ibm.com/media/IBM+Storage+Ceph+Trial+Installation/1_y2ih0la8)

Where to get help within minutes for IBMers

---



#ceph-help

## IBM TechZone for IBM Storage Ceph Test Drive

North America ATG Storage - IBM Storage Ceph Test Drive - (VMware based) Overview

Rating: ☆☆☆☆☆ (0) Rate this resource

The IBM Storage Ceph (VMware based) cluster is a virtual cluster designed for testing and demonstration purposes. The cluster is set up with various components including Object Storage, Placement Targets, and Storage Pools. The baseline configuration can also be used as a starting point for a production system. Therefore, after each demo, it is restored to its initial state.

\*\*\* NEW \*\*\* This demo has been updated to use the latest version of the Ceph software.

If you need further assistance, please read the documentation or contact the support team.

**Author**  
John Shubeck (jshubeck@us.ibm.com)

**Collaborators**  
BethAnn Stugis (bethann.stugis@us.ibm.com), John Sing (jmsing@us.ibm.com), Carl Brown (carltb@us.ibm.com)

**Ceph**

**Edit**

**Cookie Preferences**

## Accelerate with ATG Survey

---

Please take a moment to share your feedback with our team!

You can access this 6-question survey via [Menti.com](https://Menti.com) with code 5151 0447 or

Direct link <https://www.menti.com/alhsf3bgvxu6>

Or

QR Code



# Thank you!

