AWS re:Invent

STG323-R

Amazon FSx for Lustre: High-performance file system with Amazon S3

Fabrizio Manfredi

Principal Solutions Architect

Amazon Web Services

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Agenda

Amazon FSx for Lustre overview and key features

Amazon FSx for Lustre in action

- Plan your installation
- Create your first file system
- Administration tasks and monitoring
- Performance tuning

Q/A

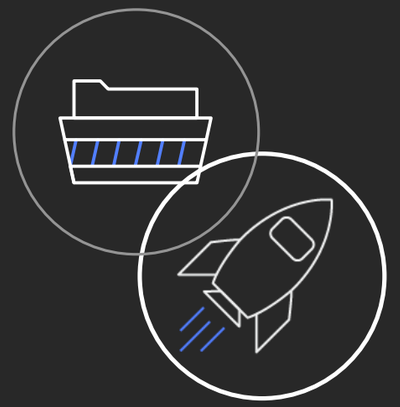
Amazon FSx for Lustre overview & key features

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

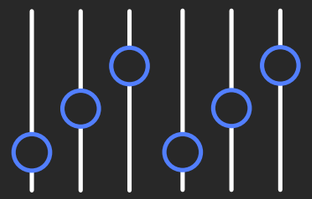
aws

High and scalable performance



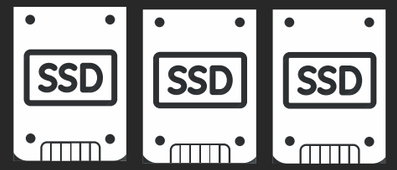
High and scalable performance

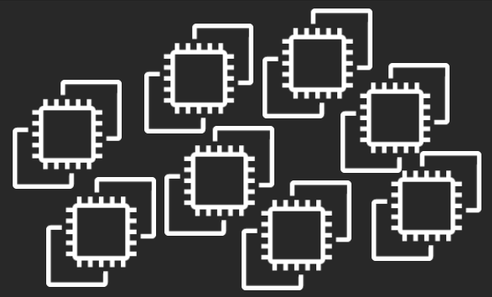
Parallel file system



100+ GiB/s throughput
Millions of IOPS
Consistent sub-millisecond latencies

SSD-based





Supports concurrent access from hundreds of thousands of cores

Each terabyte (TB) of storage provides 200 MB/second of file system

Amazon FSx for Lustre

Optimized for processing large datasets at high performance and low cost

Super-fast file system that keeps compute resources optimally utilized

Long-term data stored in Amazon Simple Storage Service (Amazon S3) or on-premises

Launch and delete files systems in minutes

Pay only for the resources you use

Amazon FSx availability

US
US West (Oregon)
US West (N. California)
US East (N. Virginia)
US East (Ohio)

EU
Ireland
Frankfurt
London
Stockholm

APAC
Sydney
Singapore
Tokyo

Amazon FSx for Lustre in action

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Build you first Fsx for Lustre file system

Plan your installation

- Estimate the file system size
- Define the life cycle of the data
- Define the client instance type
- Select a proper operating system
- Define network layout (AZ, CIDR, …)

Create your first file system: Easy to start

```
subnet=<subnet>
bucket=<bucket>
region=<region>

aws fsx create-file-system \
--file-system-type LUSTRE \
--storage-capacity 3600 \
--subnet-ids ${subnet} \
--lustre-configuration ImportPath=s3://{bucket} \
 --region ${region} \
 --output json
```

Mount file system

1. Start an EC2 instance, with the proper networking configuration
2. Install the Lustre client (documentation available for major distros)
3. Mount the file system using standard command:

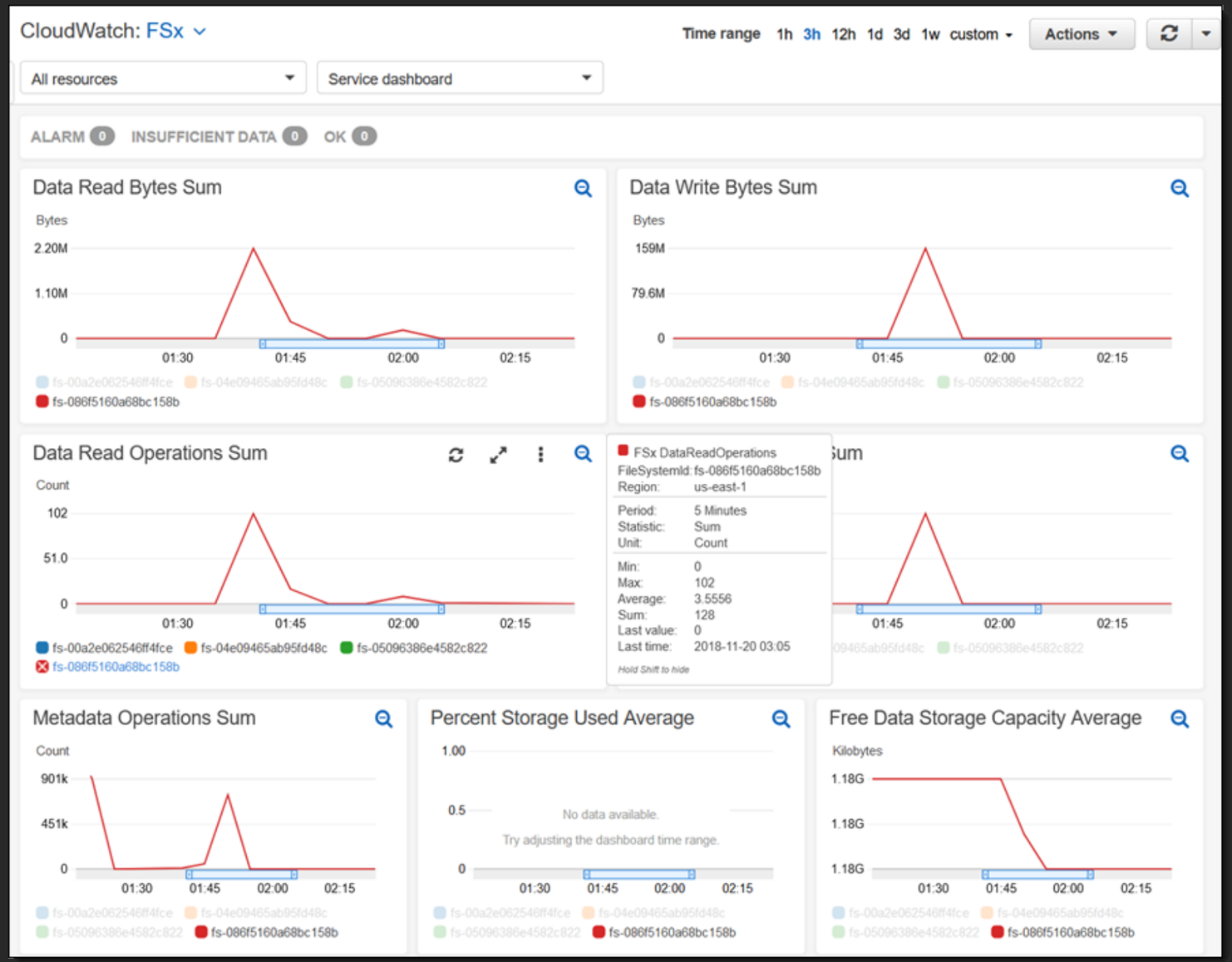
```
sudo mount -t lustre ${dnsname}@tcp:/fsx /mnt/fsx

lfs df
UUID 1K-blocks Used Available Use% Mounted on
fsx-MDT0000_UUID 107838464 4578048 103258368 4% /mnt/fsx[MDT:0]
fsx-OST0000_UUID 1182566272 4608 1182559616 0% /mnt/fsx[OST:0]
fsx-OST0001_UUID 1182566272 4608 1182559616 0% /mnt/fsx[OST:1]
fsx-OST0002_UUID 1182566272 4608 1182559616 0% /mnt/fsx[OST:2]
filesystem_summary: 3547698816 13824 3547678848 0% /mnt/fsx
```

Administration monitoring

Several metrics available:

- Space usage
- Bandwidth
- IOPS
- …



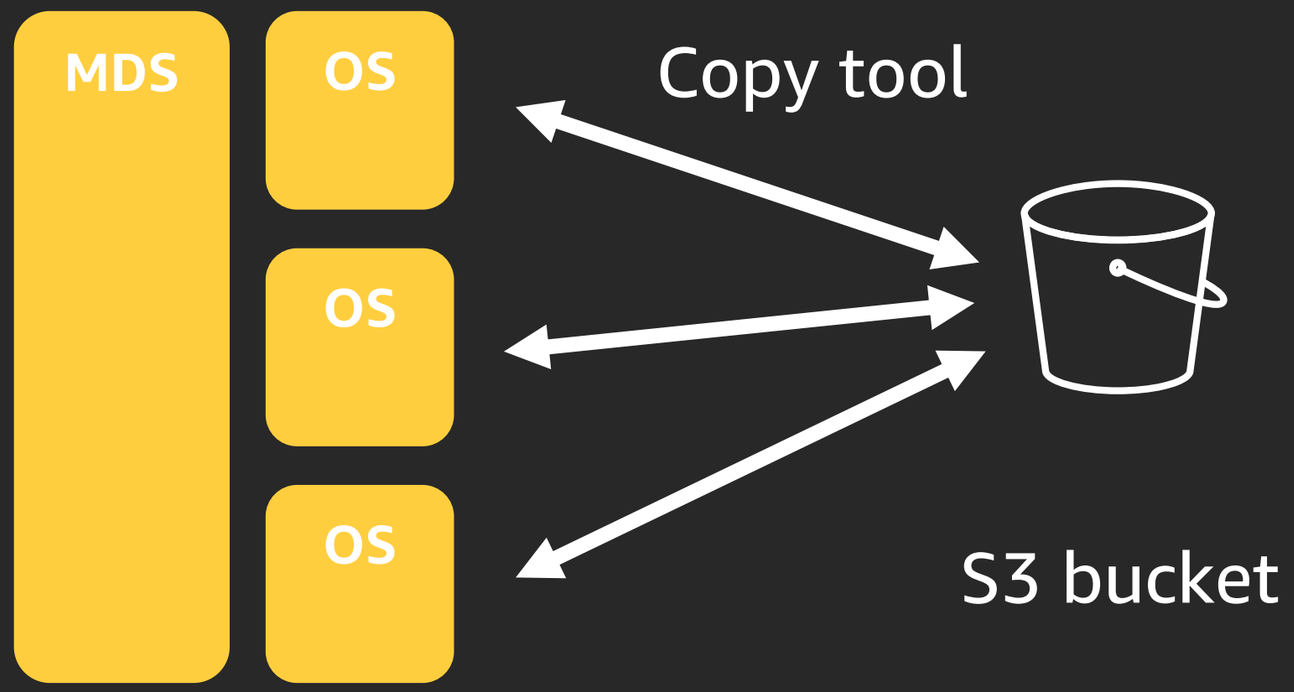
Administration tasks: Lustre/S3

- Lazy load
- Change files
- Archive them
- Release
- Read back from Amazon S3
- Show the changes are preserved

Administration tasks: Lustre/S3

Amazon FSx for Lustre can be connected to Amazon S3 as an HSM backend for:

- Read operation
- Archive operation



Administration tasks: Lustre/S3

- Files are accessed using standard mechanisms:
  - Opening a released file
  - Using standard Lustre lfs hsm_restore commands
- Convenient to restore all files, prior to start working; periodically, or at the end, archive all their files back to Amazon S3
- Using standard Lustre lfs commands, you can enqueue thousands of requests into the coordinator queue
- Workload is distributed across all file system servers to maximize bandwidth

Administration tasks: Preload

Preload bucket or subtree:

```
Preload      : lfs hsm_restore path/to/export/file
Verify Status: lfs hsm_action path/to/export/file
```

Restore bucket or subtree:

```
nohup  find local/directory -type f -print0 | xargs -0 -n 1 sudo lfs hsm_restore &
```

Verify status:

```
find local/directory -type f -print0 | xargs -0 -n 1 -P 8 sudo lfs hsm_action | grep " NOOP " | wc -l
```

Administration tasks: Archive

Archive to S3 bucket :

```
Archive      : lfs hsm_archive path/to/export/file
Verify Status: lfs hsm_action path/to/export/file
```

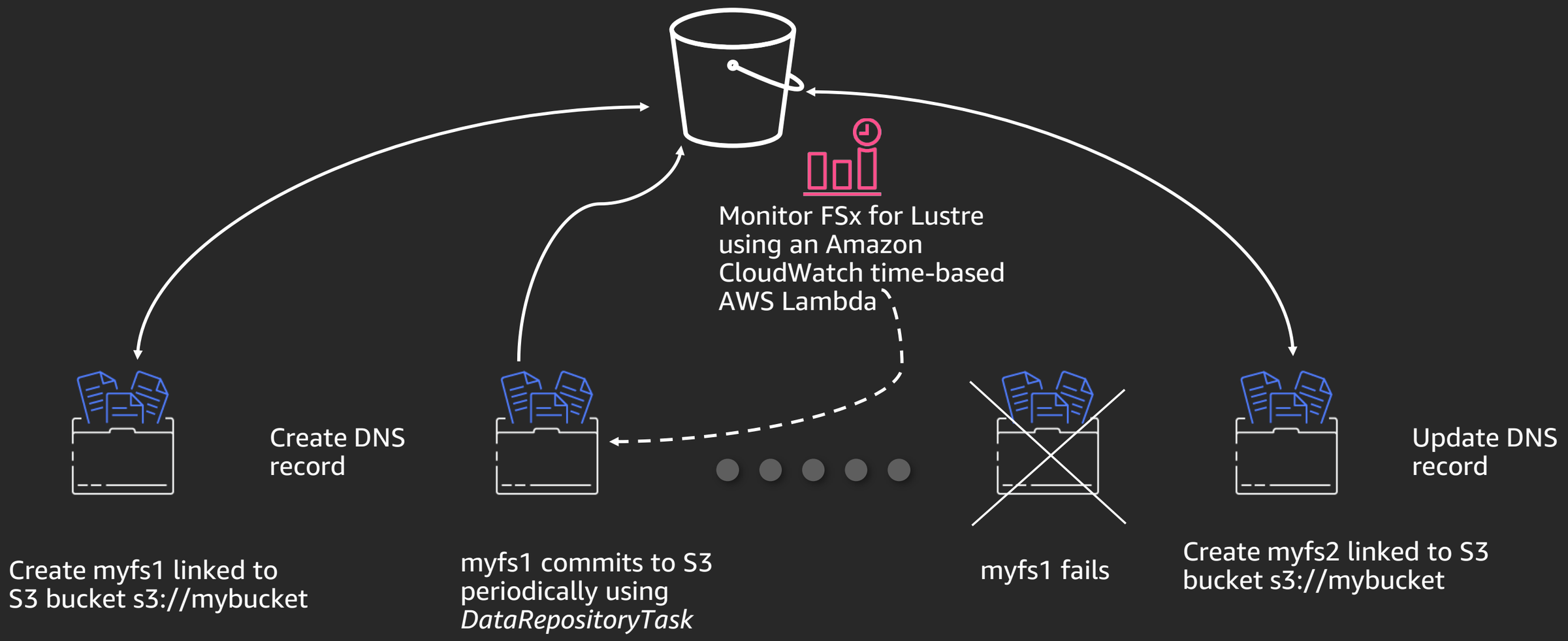
Archive bucket or subtree:

```
nohup  find local/directory -type f -print0 | xargs -0 -n 1 sudo lfs hsm_archive &
```

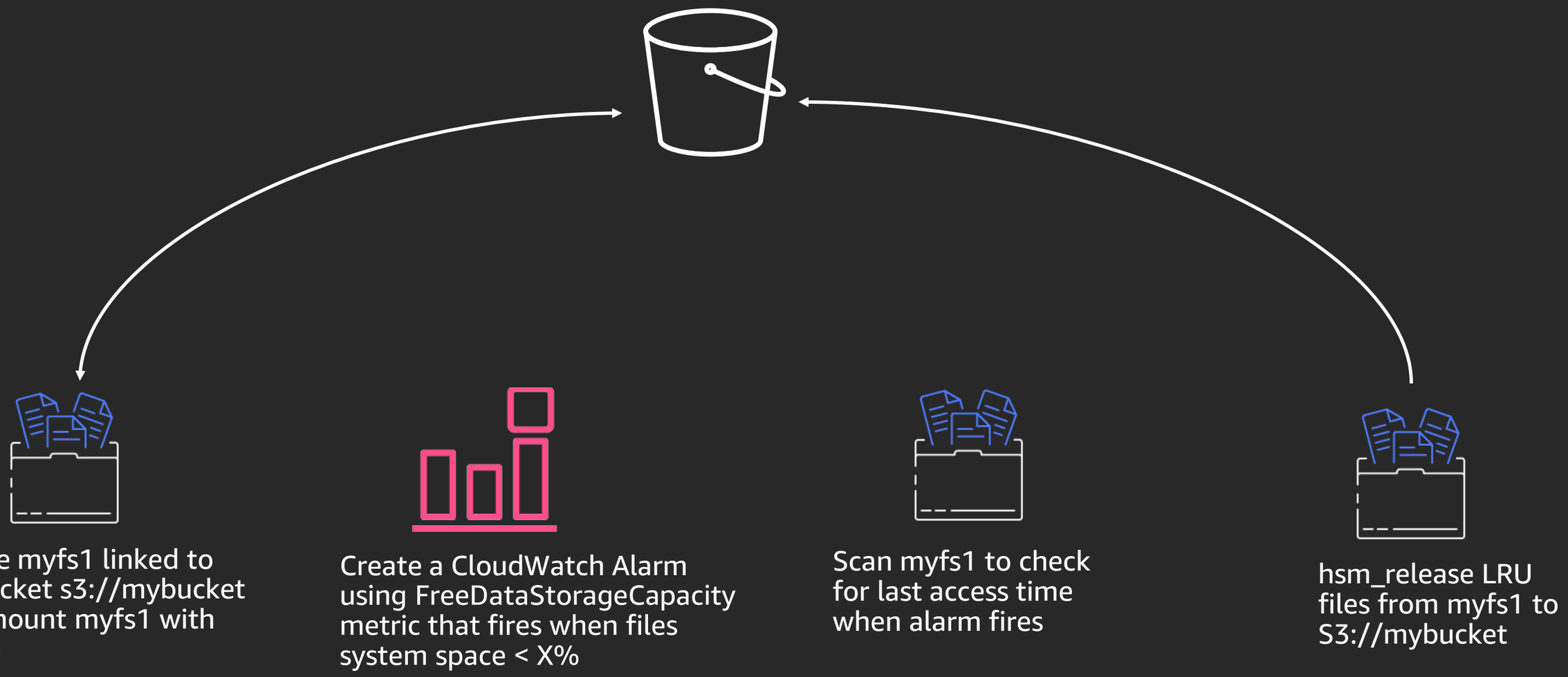
Verify status:

```
find path/to/export/file -type f -print0 | xargs -0 -n 1 -P 8 sudo lfs hsm_action | grep "ARCHIVE" | wc -l
```

Availability/Durability management of FSx Lustre



Evict inactive data sets to S3—free up space

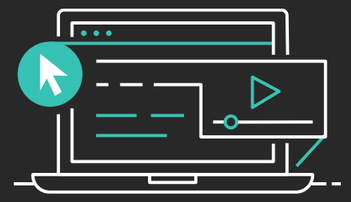


Optimizing I/O performance on FSx Lustre

- Striping filesystem data
  - Stripe bigger files for better I/O performance when concurrent access is the dominant access pattern
  - Stripe files across disks based on CloudWatch Max metric
  - Set ImportedFileChunkSize = (dominant file size / # of disks)
- Average I/O size
  - Throughput increases with higher average IO size
- Client selection
  - Choose EC2 instance type with enough memory, CPU, & bandwidth

Learn storage with AWS Training and Certification

Resources created by the experts at AWS to help you build cloud storage skills



45+ free digital courses cover topics related to cloud storage, including:

- Amazon S3
- AWS Storage Gateway
- Amazon S3 Glacier
- Amazon Elastic File System (Amazon EFS)
- Amazon Elastic Block Store (Amazon EBS)

Classroom offerings, like Architecting on AWS, feature AWS expert instructors and hands-on activities

Visit aws.amazon.com/training/path-storage/

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws training and certification

Thank you!

Fabrizio Manfredi

fabman@amazon.com

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Please complete the session survey in the mobile app.

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws