
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

NFX301

Hybrid cloud: Cataloging objects across storage systems

Priyesh Narayanan
Software Engineer
Netflix

Kishore Kasi
Software Architect
Netflix

Denise Mace
Strategic Account Manager
AWS

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Netflix studio storage

Catalog service requirements

Architecture

Data model and API

Summary

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Netflix Studio storage

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Netflix Studio ecosystem



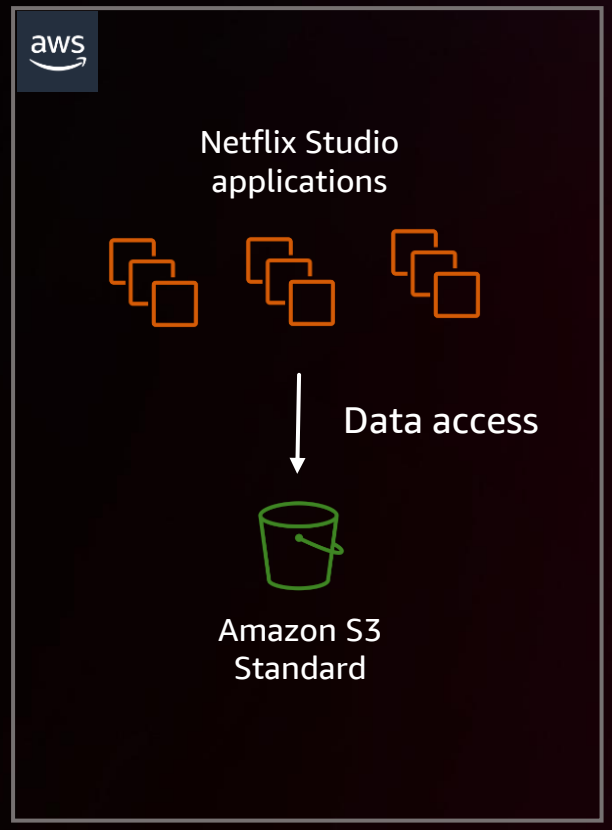
Data processing

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Netflix Studio ecosystem

Netflix Studios

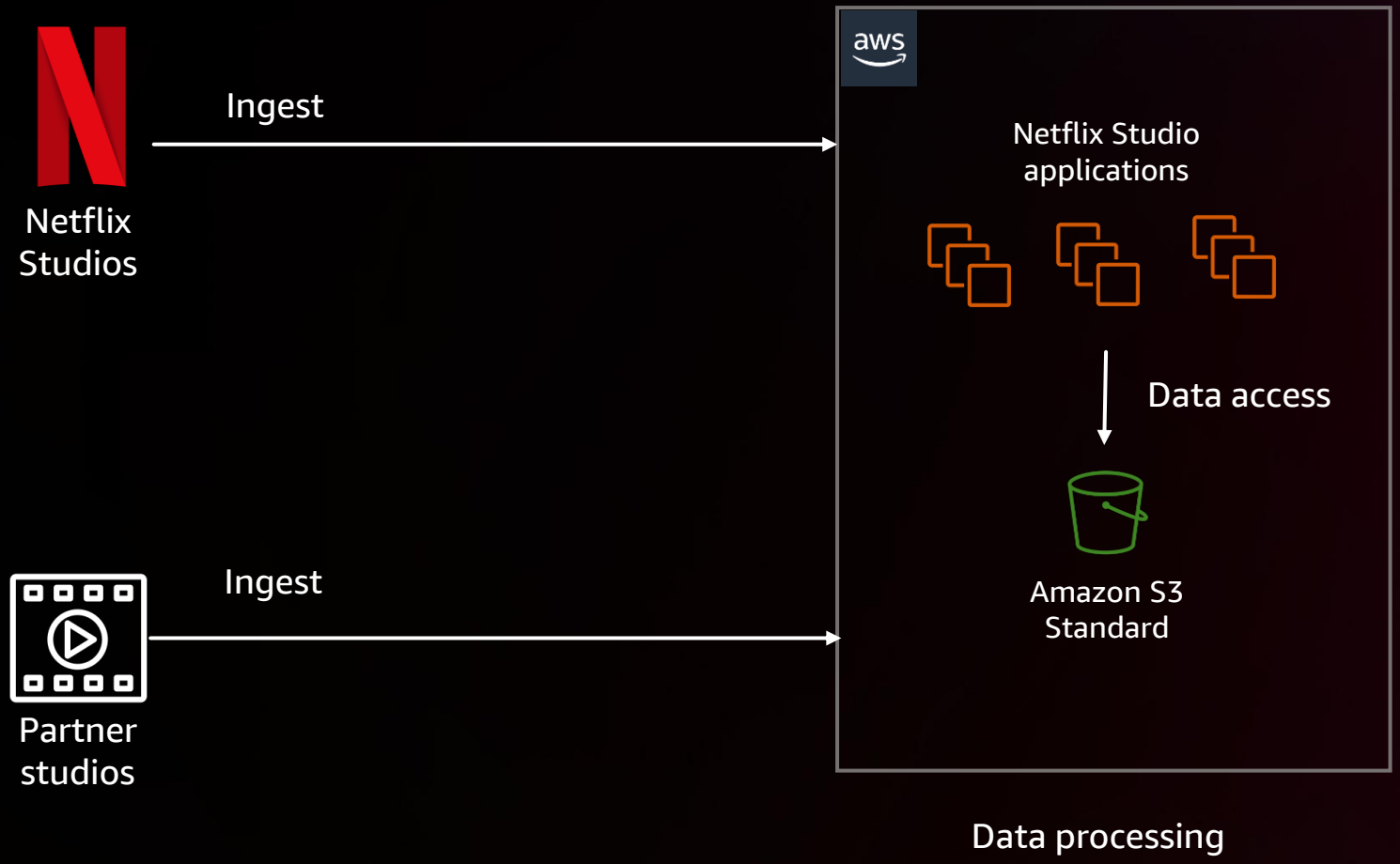
Partner studios



Data processing

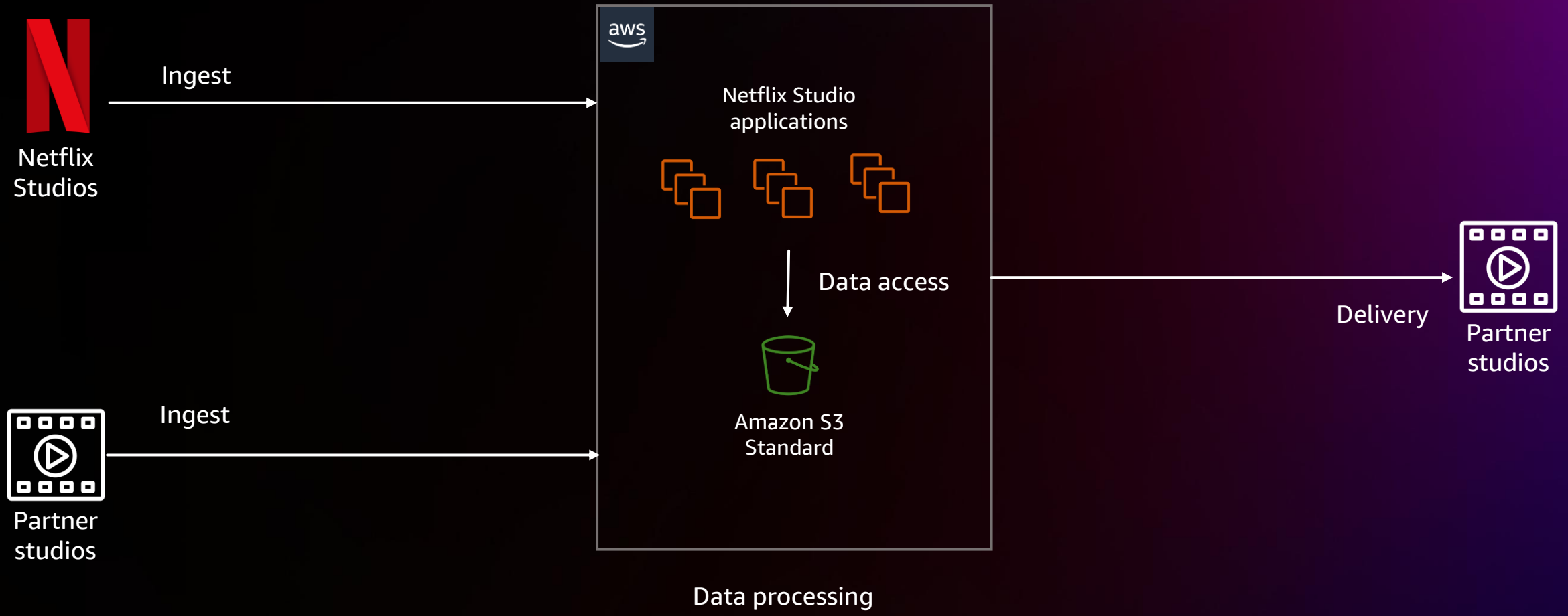
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Netflix Studio ecosystem



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Netflix Studio ecosystem



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

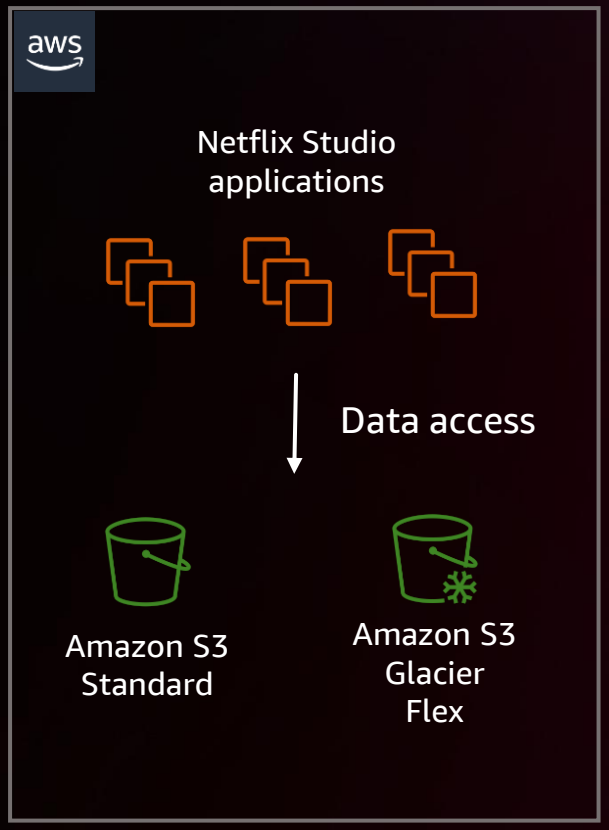
Data processing in the cloud



Data processing

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

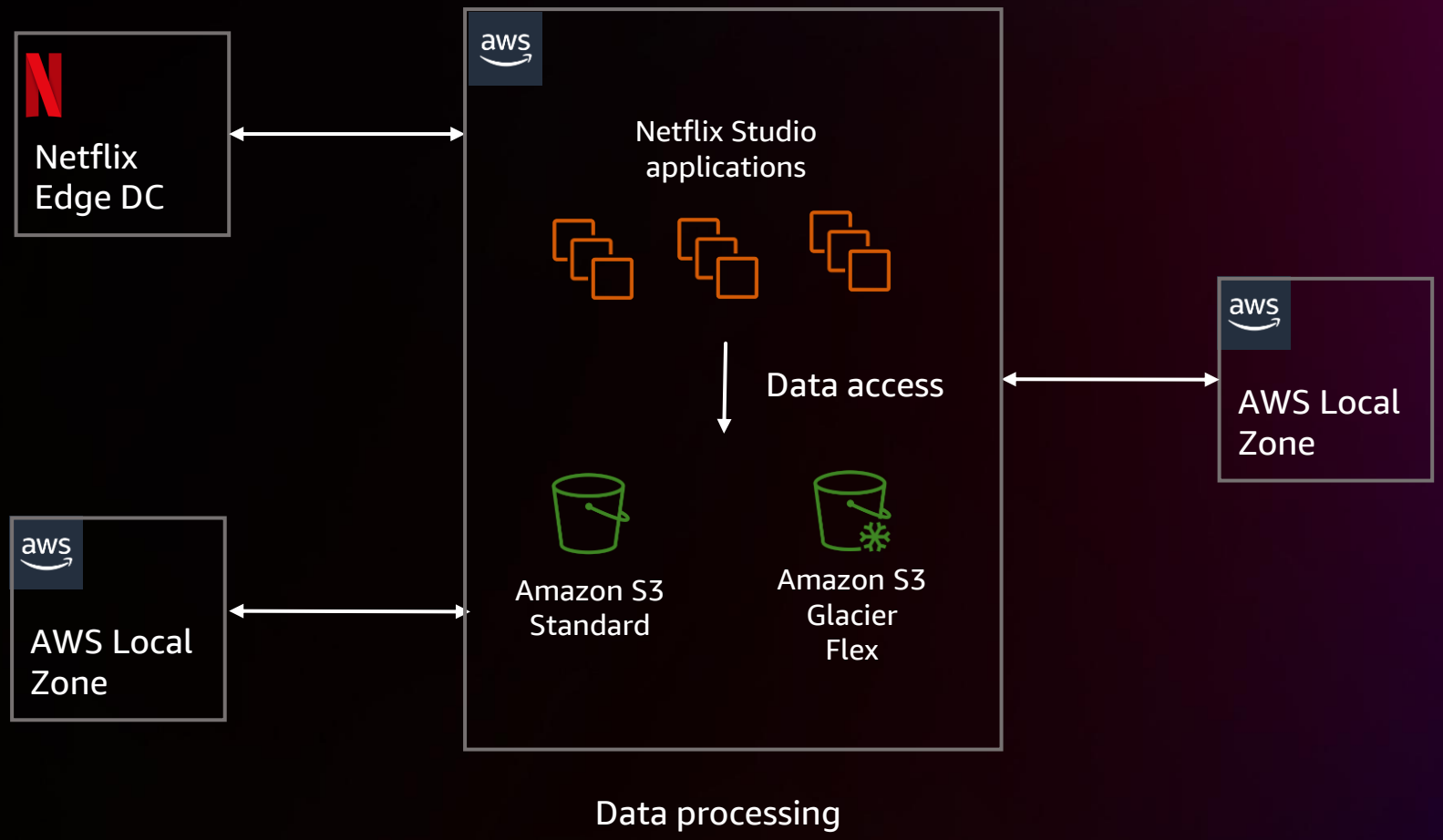
Data management



Data processing

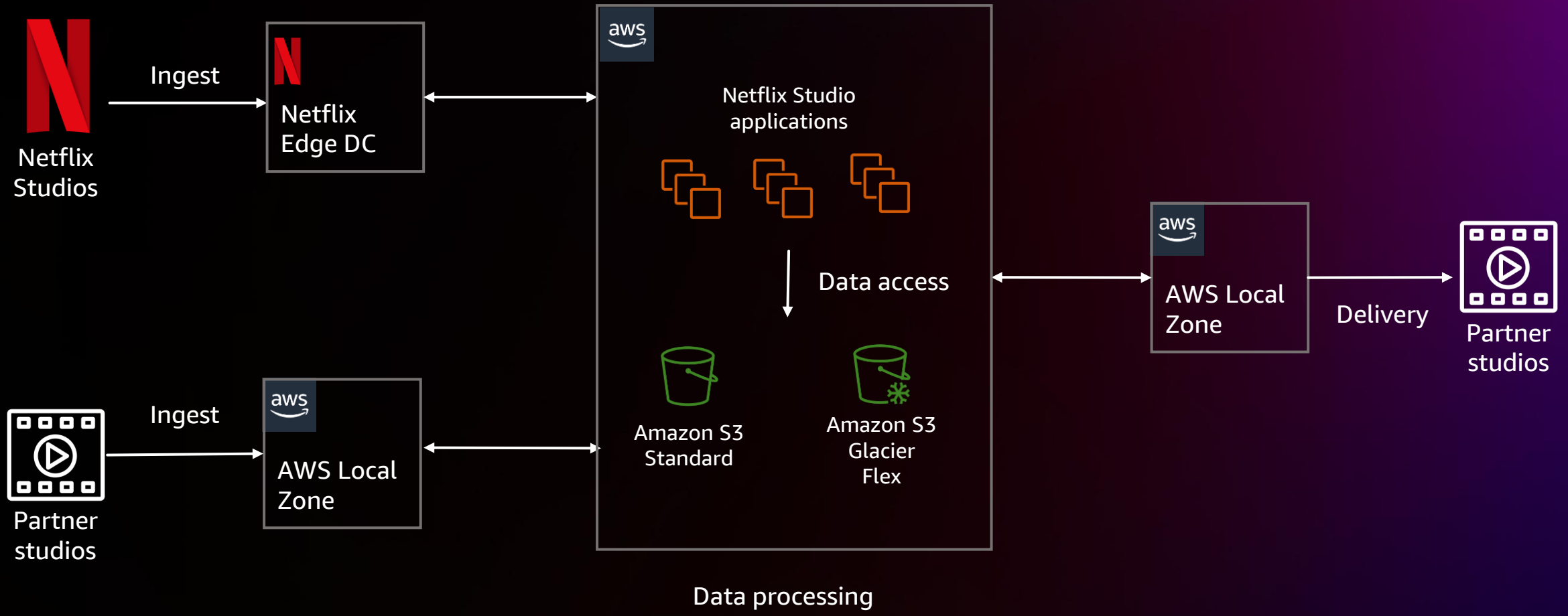
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hybrid cloud



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Netflix Studio in hybrid cloud



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data management in hybrid cloud

- Centralized control and management
- Decentralized data storage
- Transparent data access
- Data catalog

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Catalog service

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Requirements

Global registry for objects

Namespace management

Flexibility

Existing data

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Storage classes and locations



Amazon S3 Standard



Amazon S3 Glacier Flexible Retrieval



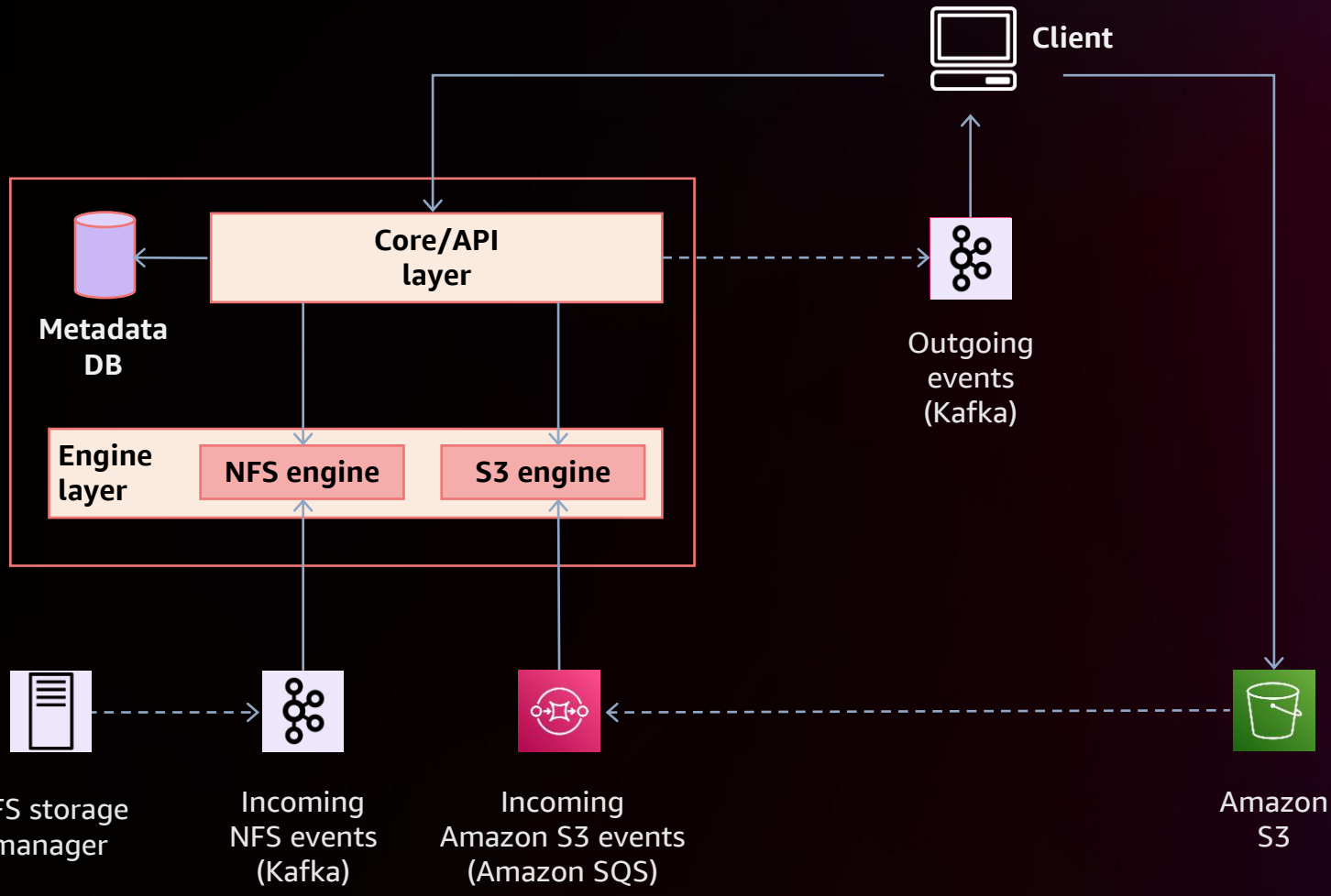
Object cache (self managed)



Filers (over NFS)

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Catalog service architecture



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

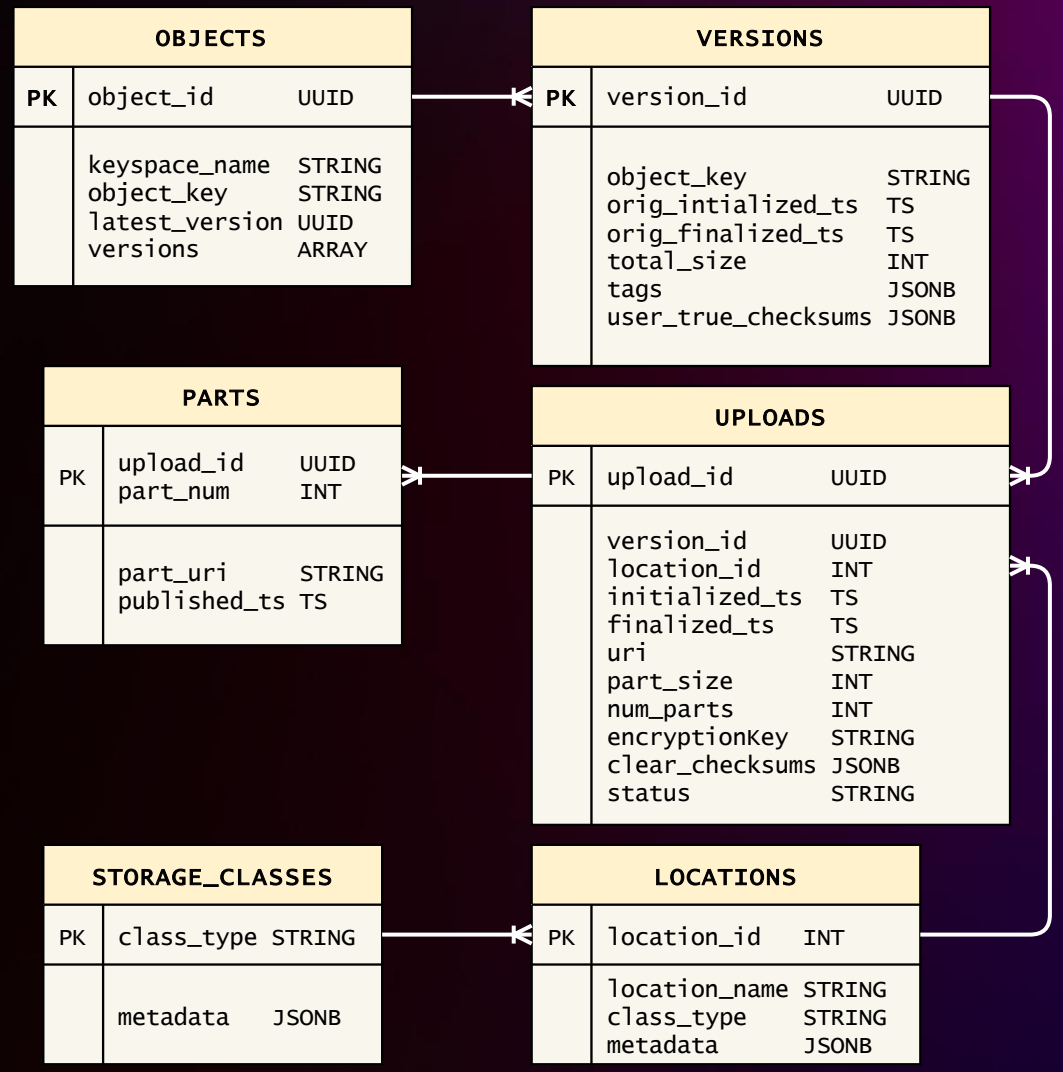
Data model

Objects

Versions

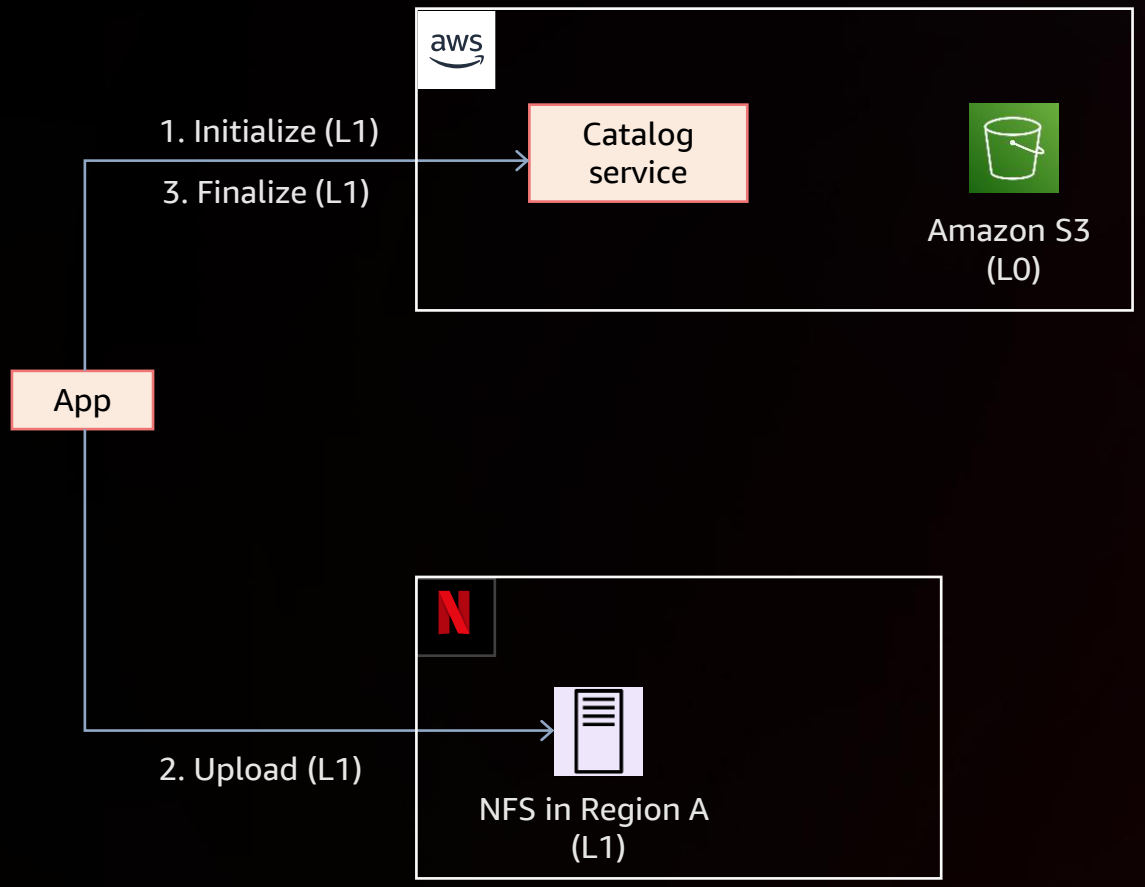
Uploads

Parts



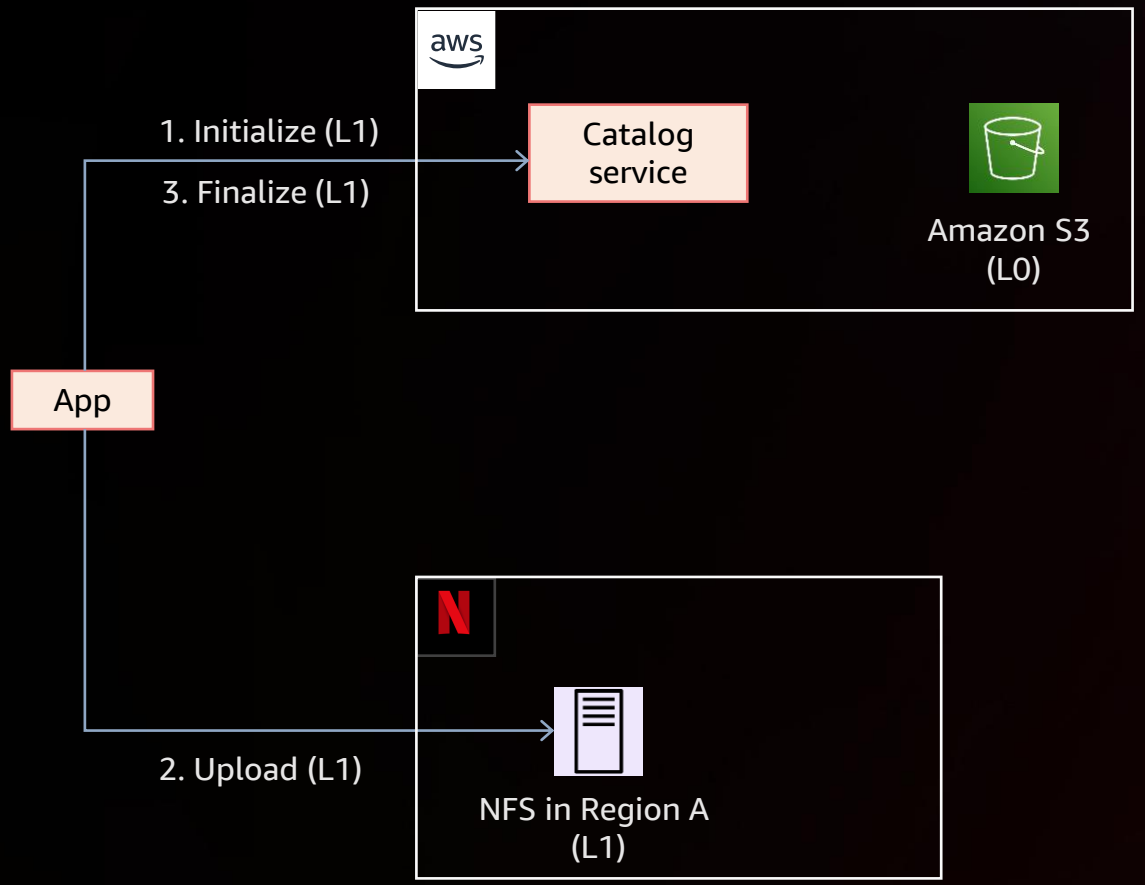
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Example flow: Ingest



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Example flow: Ingest



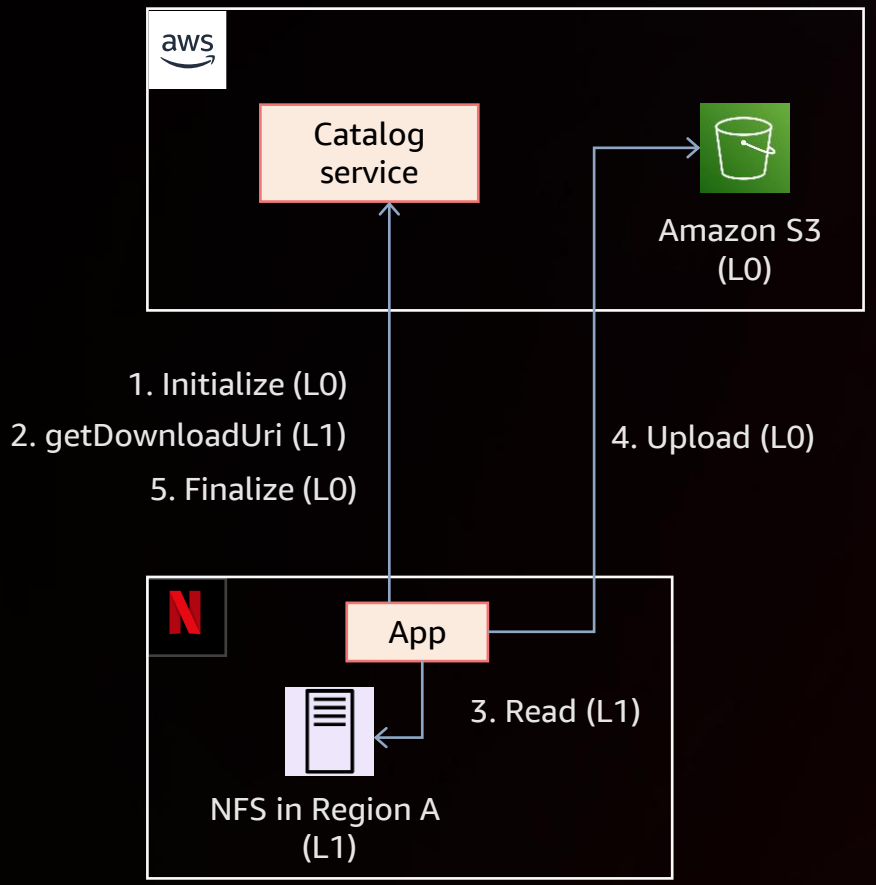
OBJECTS		
object_id	object_key	keyspace_name
o1	key1	ksp1

VERSIONS			
version_id	object_id	total_size	true_checksums
v1	o1	3923	{md5: b356ce6e23bb}

UPLOADS						
upload_id	version_id	location_id	finalized_ts	num_parts	part_size	status
u1	v1	L1	1667349917	10	400	CHECKSUM_MATCHED

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Example flow: Transfer



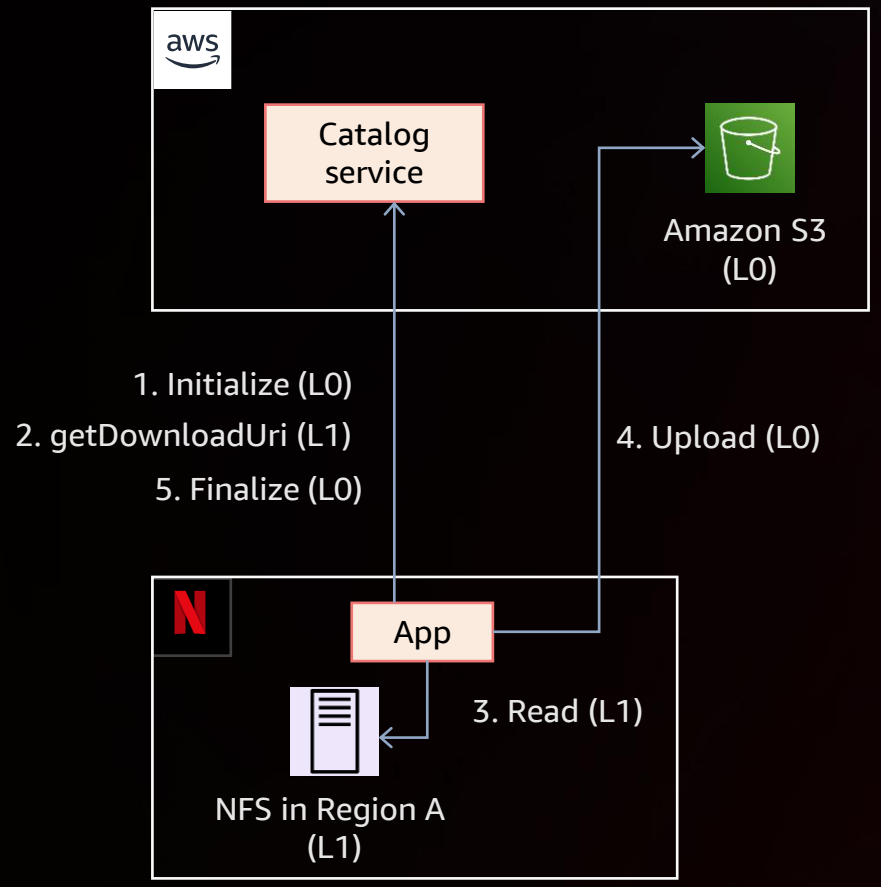
OBJECTS		
object_id	object_key	keyspace_name
o1	key1	ksp1

VERSIONS			
version_id	object_id	total_size	true_checksums
v1	o1	3923	{md5: b356ce6e23bb}

UPLOADS						
upload_id	version_id	location_id	finalized_ts	num_parts	part_size	status
u1	v1	L1	1667349917	10	400	CHECKSUM_MATCHED

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Example flow: Transfer



OBJECTS		
object_id	object_key	keyspace_name
o1	key1	ksp1

VERSIONS			
version_id	object_id	total_size	true_checksums
v1	o1	3923	{md5: b356ce6e23bb}

UPLOADS						
upload_id	version_id	location_id	finalized_ts	num_parts	part_size	status
u1	v1	L1	1667349917	10	400	CHECKSUM_MATCHED
u2	v1	L0	1667349954	4	1000	FINALIZED

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Events

Incoming events from storage locations

Outgoing events for clients

Event types: finalization, checksums, deletion

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

API

- initializeUpload()
- finalizeUpload()
- deleteObjects()

- getDownloadUri()
- getUploadUri()
- getObjectMetadata()

- addStoreLocation()
- getStoreLocation()

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

API example

```
rpc GetDownloadUri (GetDownloadUriRequest) returns (GetDownloadUriResponse) {
};

message GetDownloadUriRequest {
  ObjectInfo object_info;
  repeated LocationFilter location_filters; // List of preferred locations
}

message LocationFilter {
  string location;
  repeated int32 part_nums; // list of part numbers to get
}

message GetDownloadUriResponse {
  ObjectInfo objectInfo;
  string upload_id;
  oneof response {
    map<int32, string> s3_signed_url; // response associated with S3 Engine
    map<int32, string> nfs_uri;       // response associated with NFS Engine
  }
}
```

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Summary

Catalog service global registry

Metadata management (namespace and key)

Location preferences

Part addressability

Support for lifecycle management

Smart client

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.