
aws
SUMMIT
INDIA | MAY 25, 2023

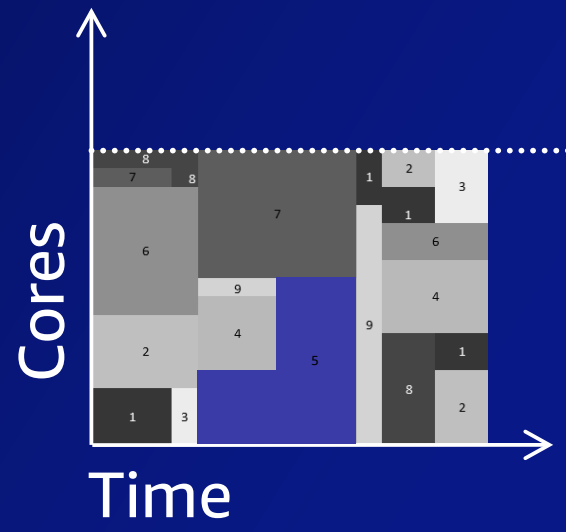
SDB008

Optimize storage using Amazon Fsx for Lustre

Vibhu Pareek (he/him)
Solutions Architect
AWS India

Accelerate time to results for compute-heavy workloads

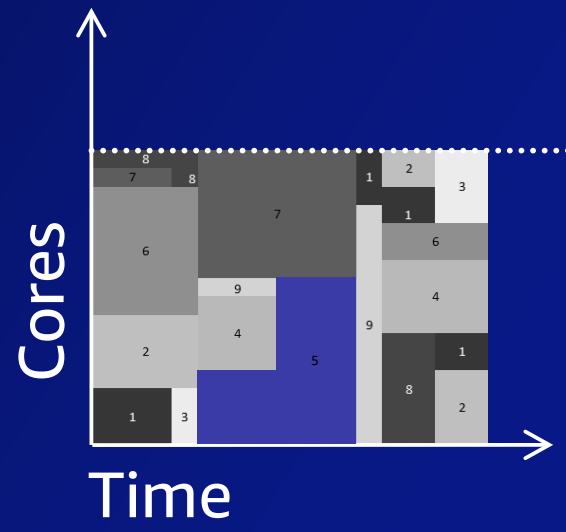
Finite capacity on-premises



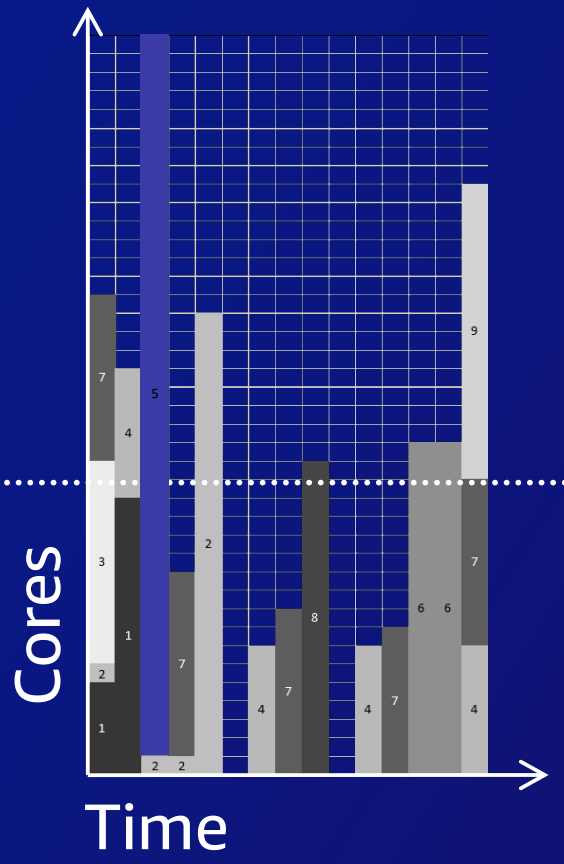
Fixed data center capacity limit

Accelerate time to results for compute-heavy workloads

Finite capacity on-premises

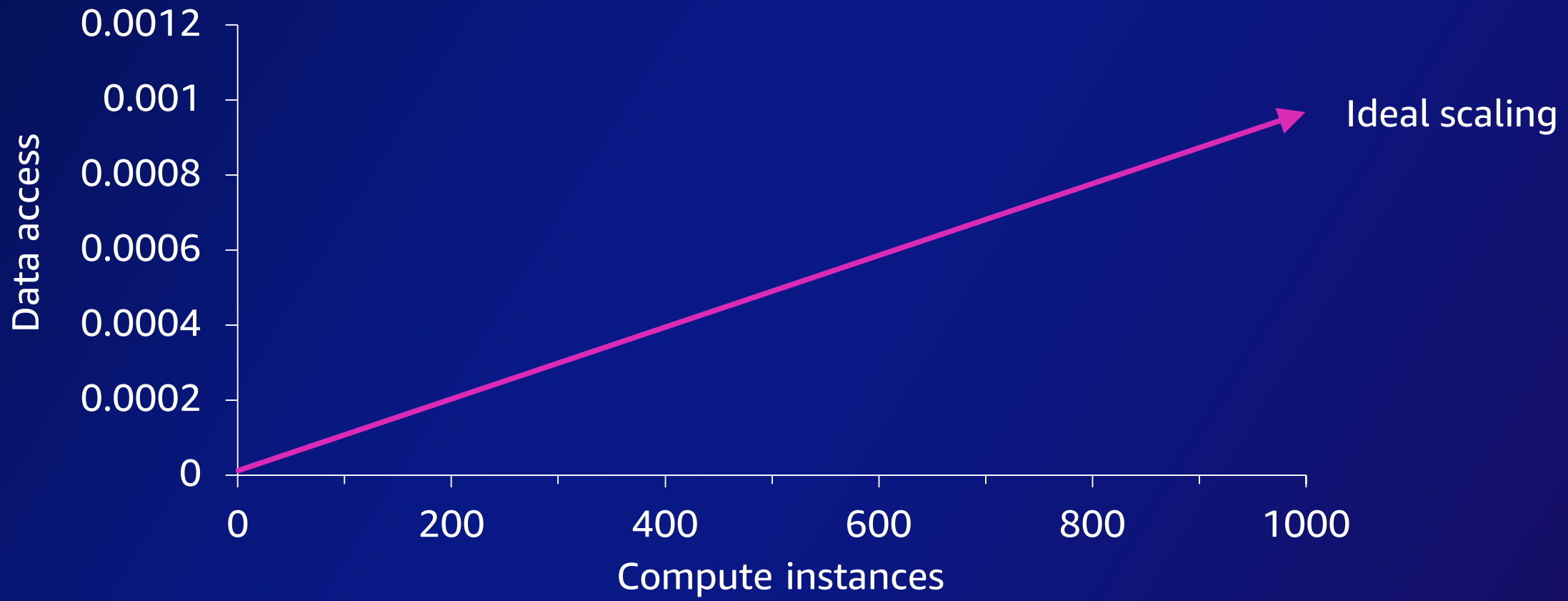


Virtually unlimited capacity on cloud

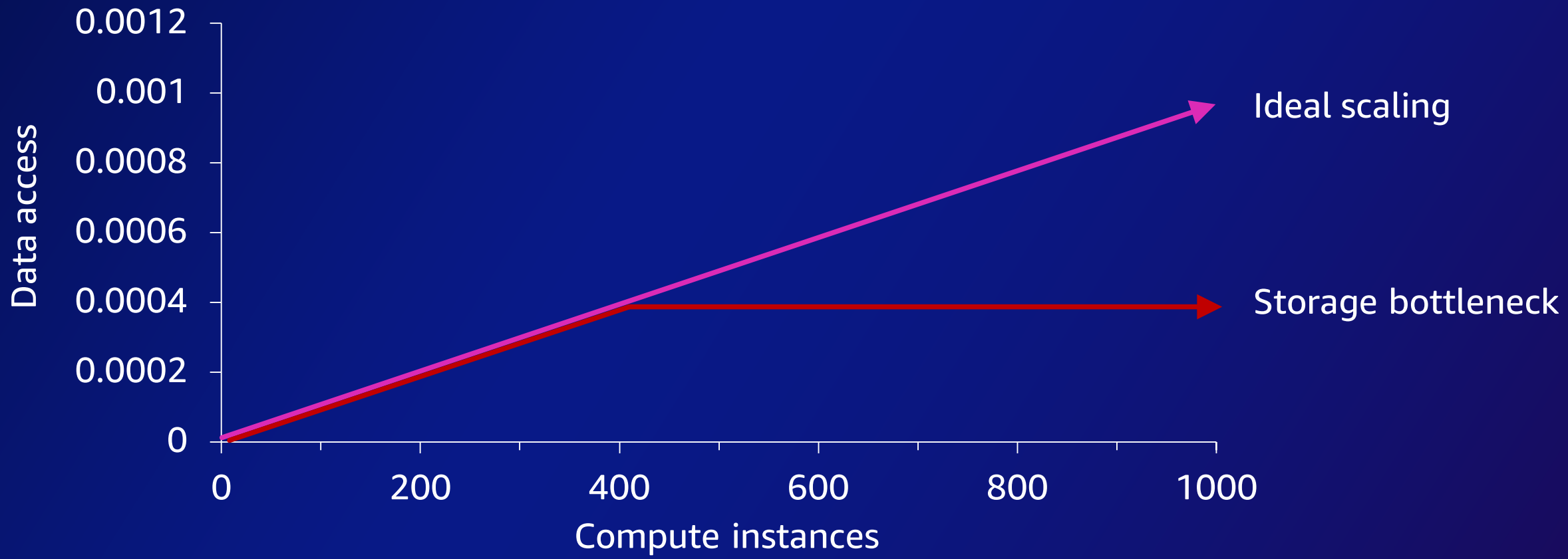


Fixed data center capacity limit

Scalable compute requires scalable data access



Scalable compute requires scalable data access



Common architecture patterns

On EBS or instance storage

Common architecture patterns

On EBS or instance storage

Self-managed file systems

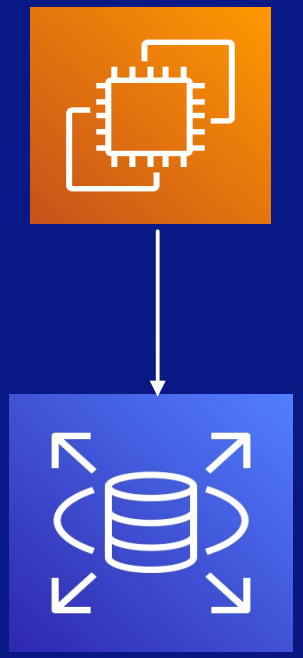
Common architecture patterns

On EBS or instance storage

Self-managed file systems

Directly on S3

Common architecture patterns



Amazon EC2 + RDS

Amazon FSx

Launch, run, and scale . . .

Feature-rich, high-performance, cost-effective, and fully managed storage . . .

Lustre

Windows File Server

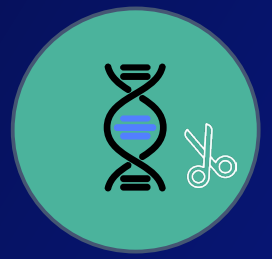
NetApp ONTAP

OpenZFS

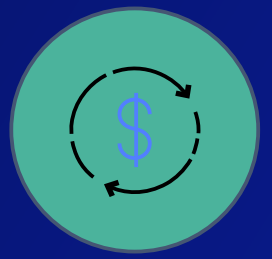
Powered by the world's most popular file systems

Amazon Fsx for Lustre usecases across industry verticals and application areas

Industries and example use cases



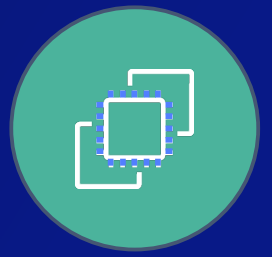
Life sciences:
Genome analysis



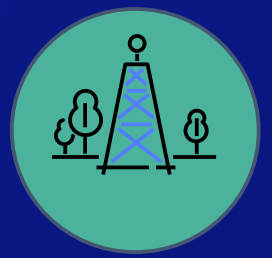
Financial services:
Modeling and analytics

Media and entertainment:
Rendering and transcoding

Automotive:
ECU simulations and object detection



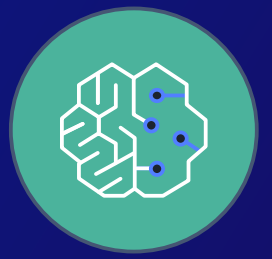
Semiconductor:
Electronic design automation



Oil and gas:
Seismic data processing

Application areas

Big data analytics



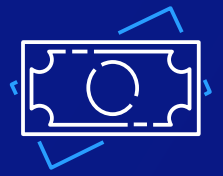
Machine learning

High performance computing

[1] High Performance Computing on AWS Redefines What is Possible

Why consider Amazon Fsx for Lustre?

Highly scalable throughput capacity



Storage options optimized for price-performance

Access to Amazon S3 data through a fast file interface

Why consider Amazon Fsx for Lustre?

Highly scalable
throughput
capacity

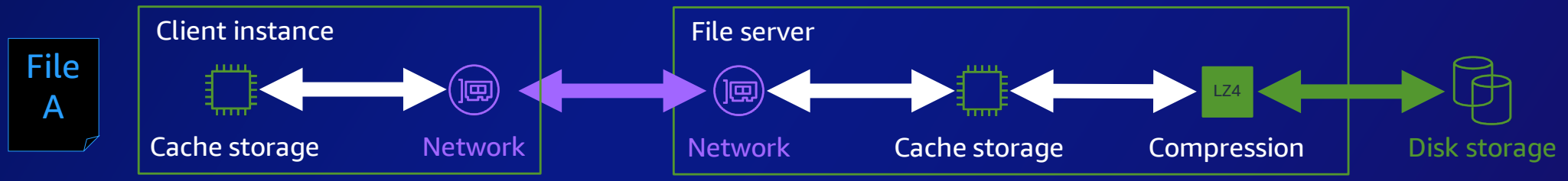
Hundreds of GB/s

of throughput. Can scale upto
1000 GB/s per file system

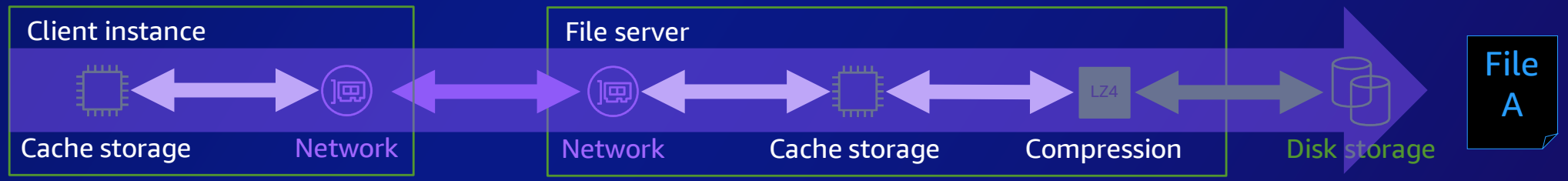
Millions

of IOPS

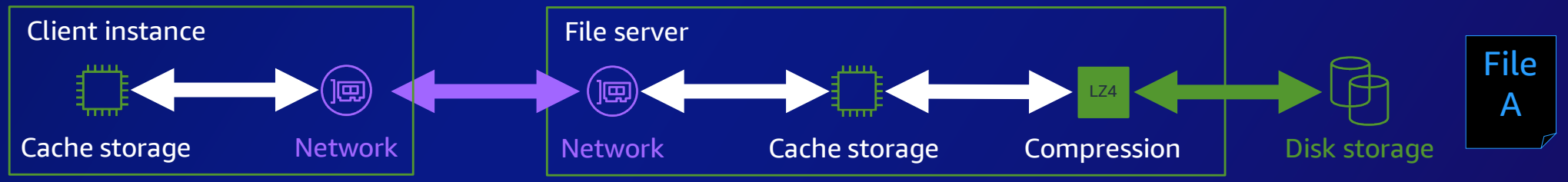
Highly scalable throughput capacity: How it works



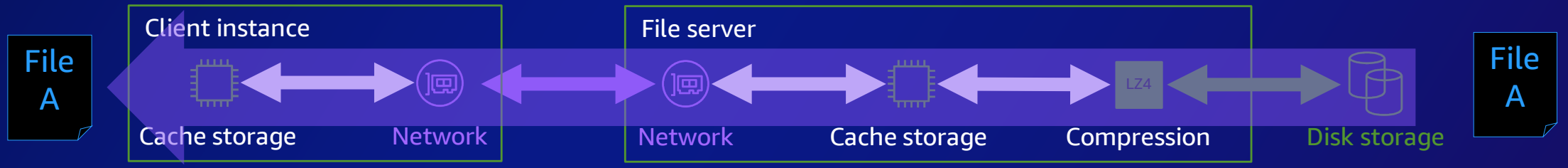
Highly scalable throughput capacity: How it works



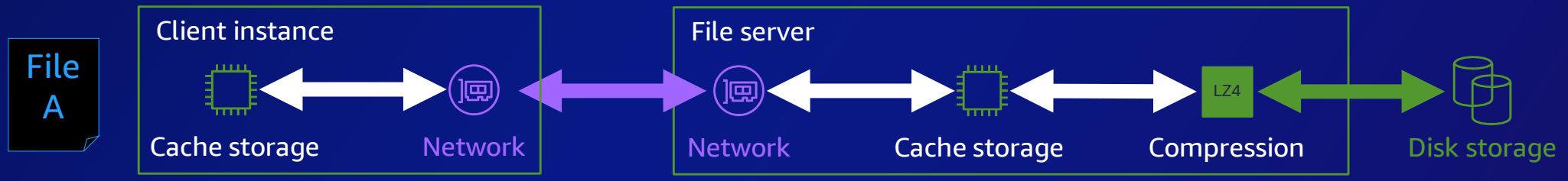
Highly scalable throughput capacity: How it works



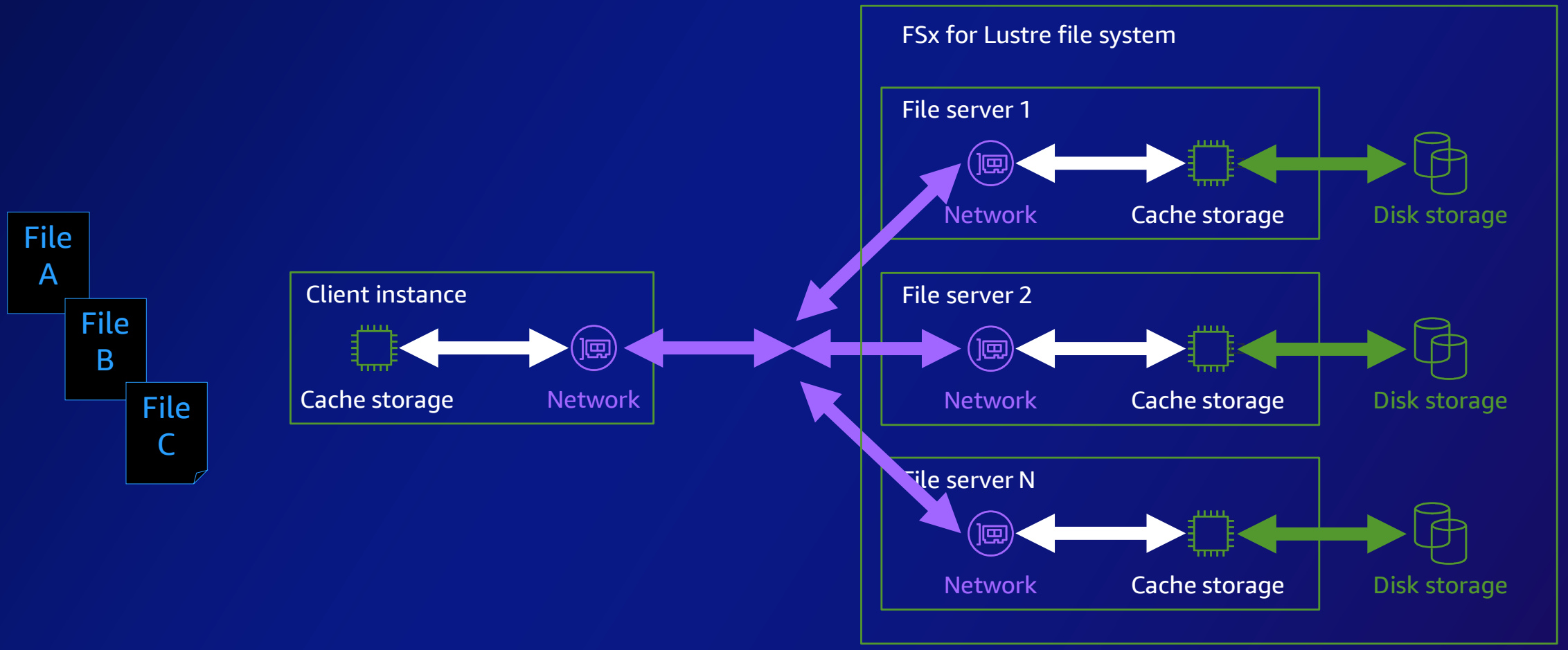
Highly scalable throughput capacity: How it works



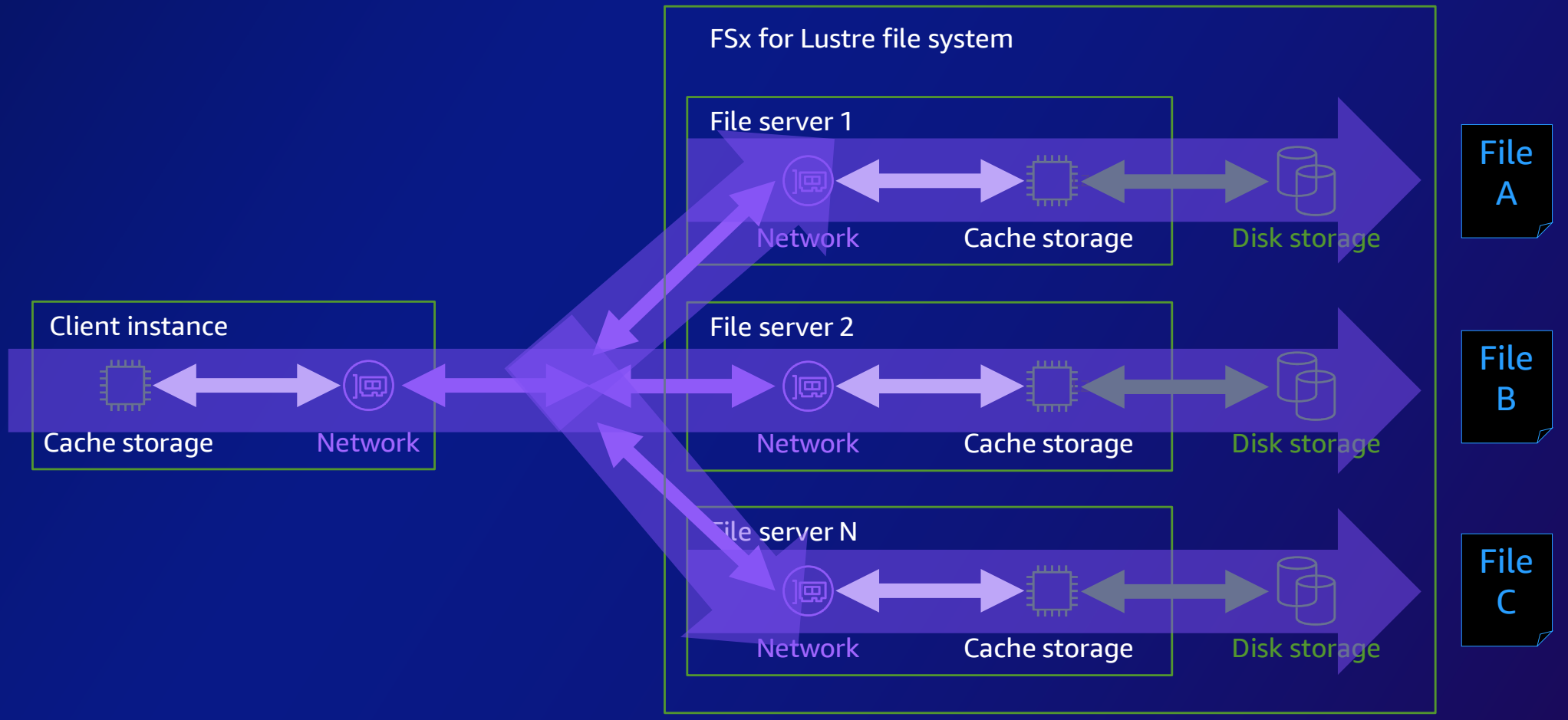
Highly scalable throughput capacity: How it works



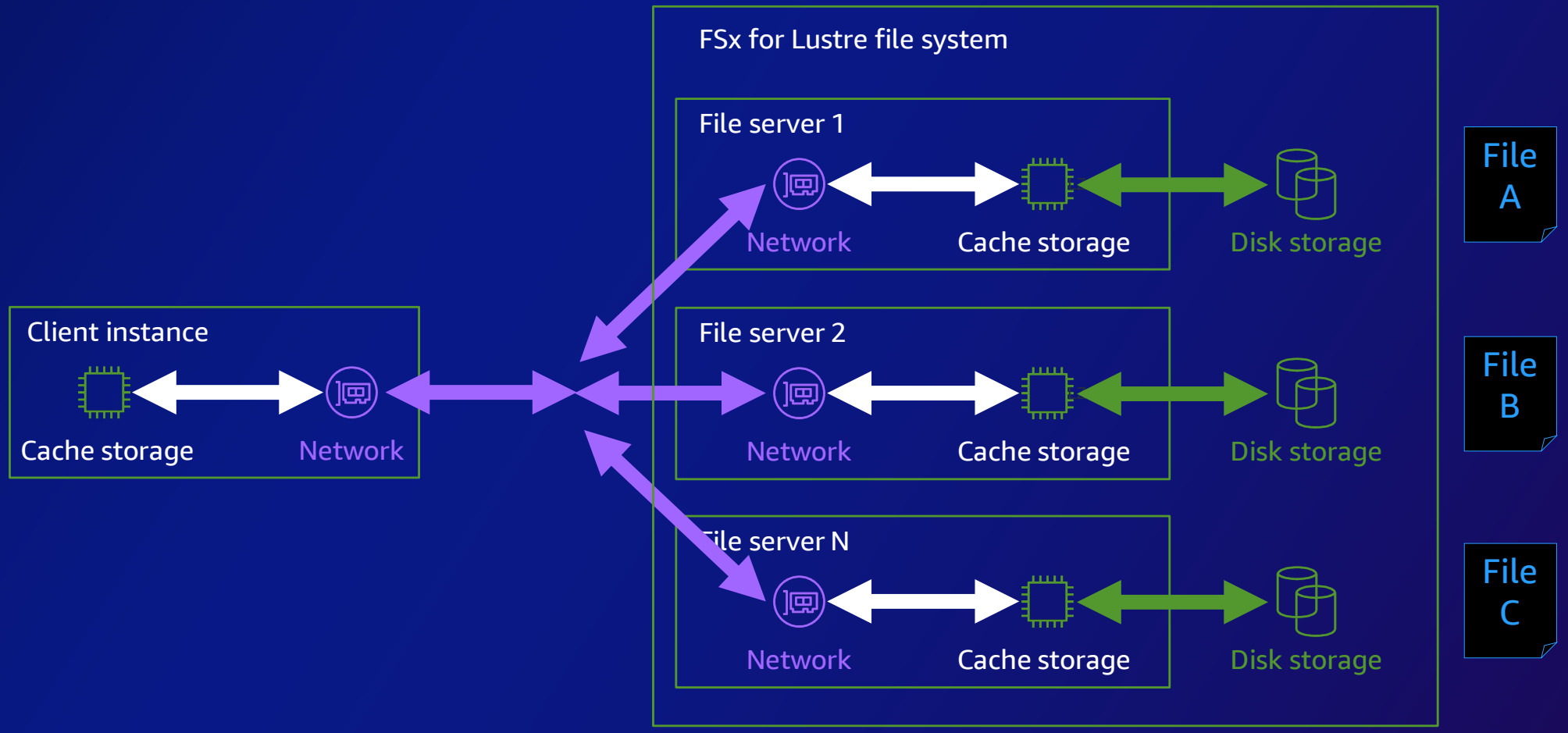
Highly scalable throughput capacity: How it works



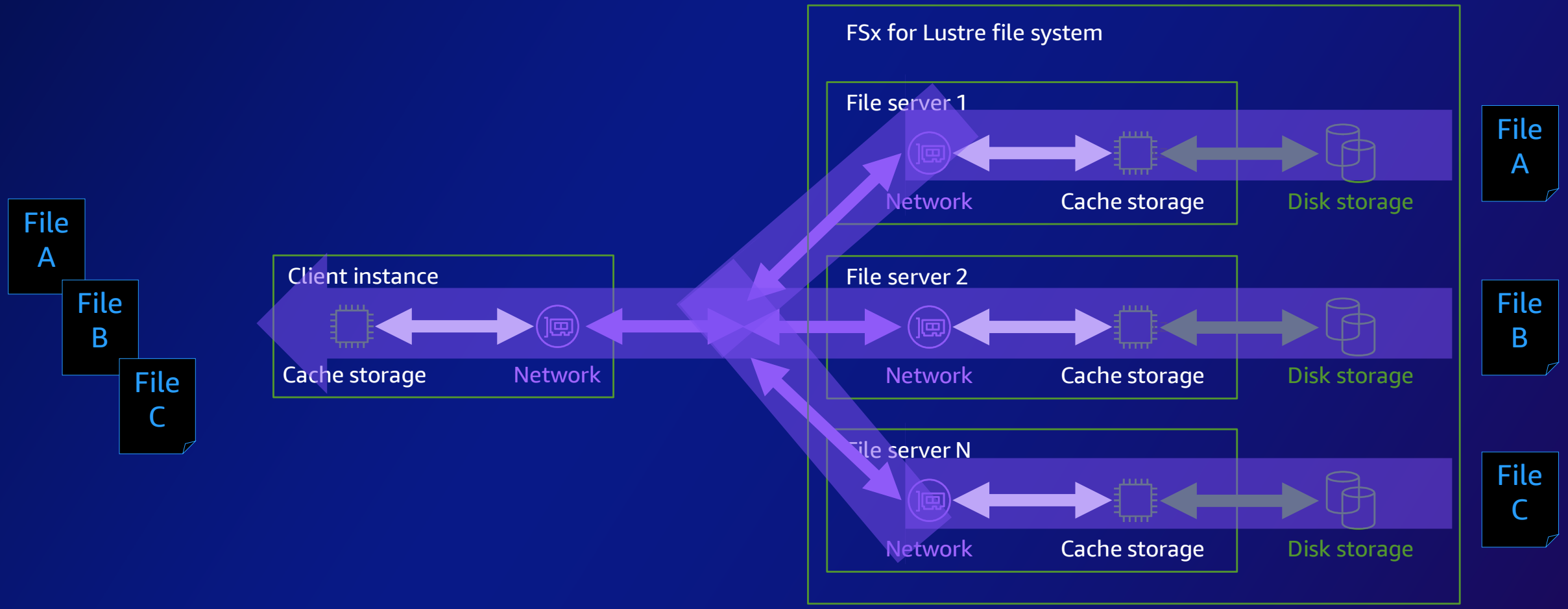
Highly scalable throughput capacity: How it works



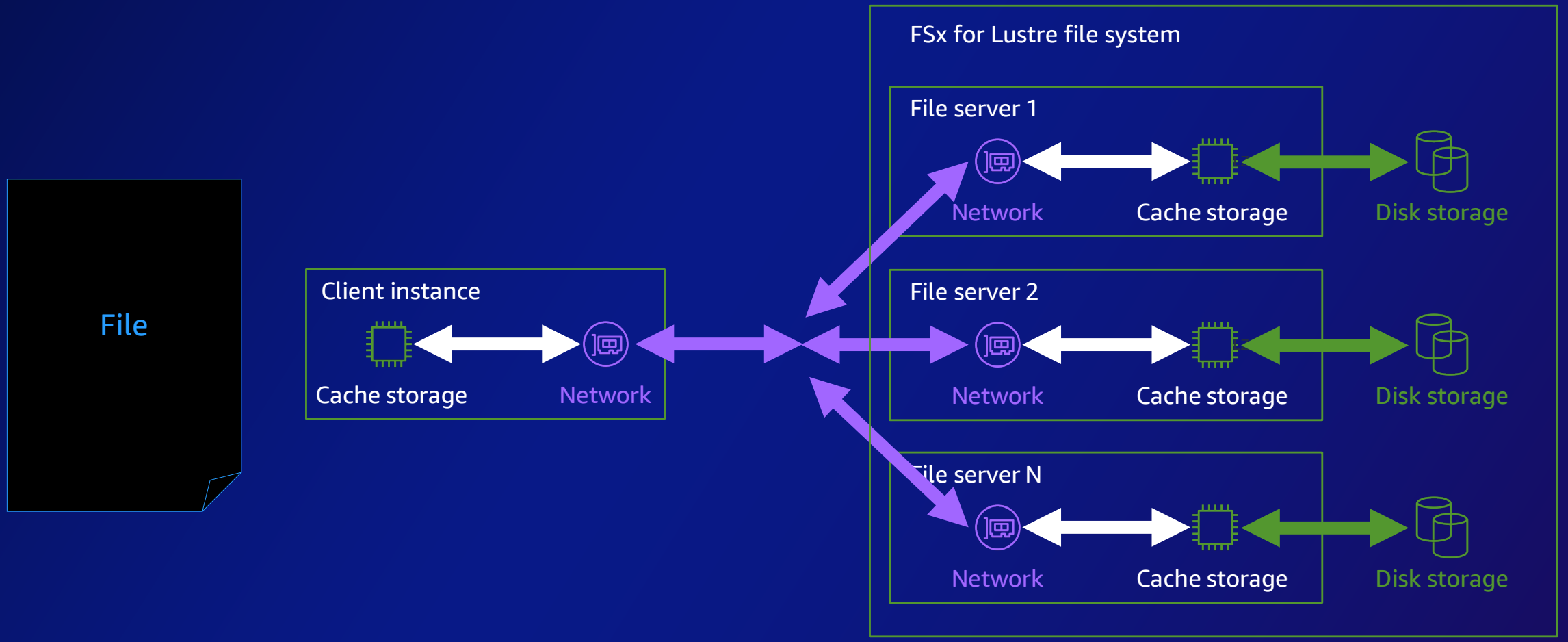
Highly scalable throughput capacity: How it works



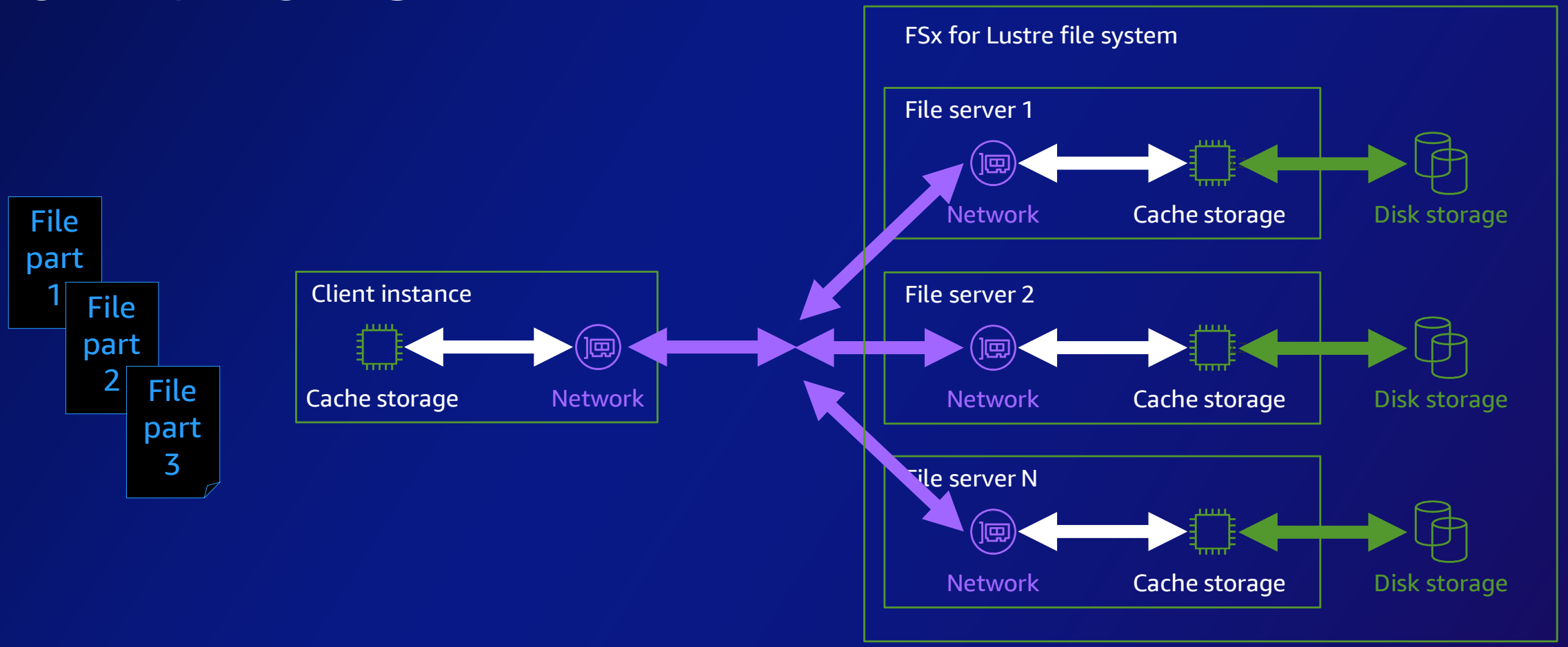
Highly scalable throughput capacity: How it works



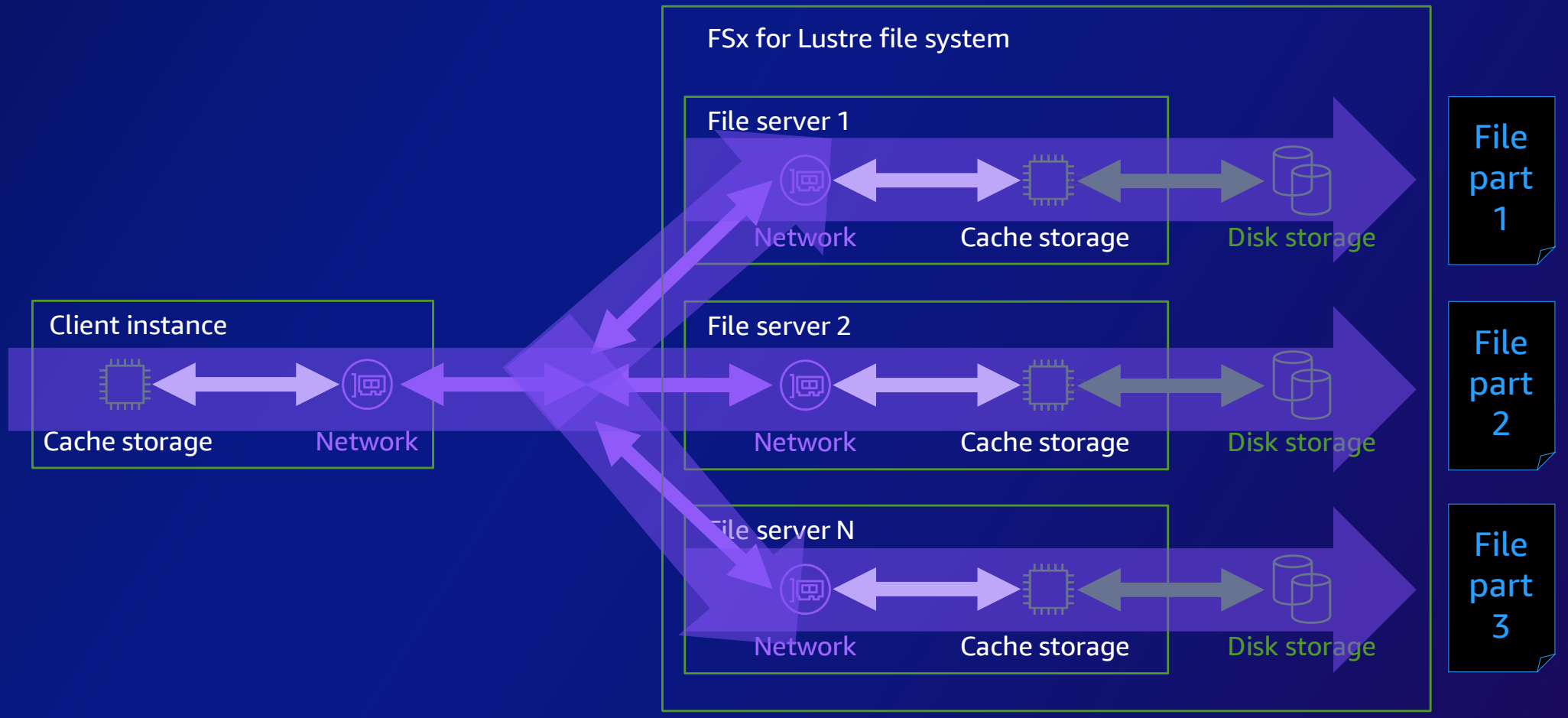
Highly scalable throughput capacity: How it works



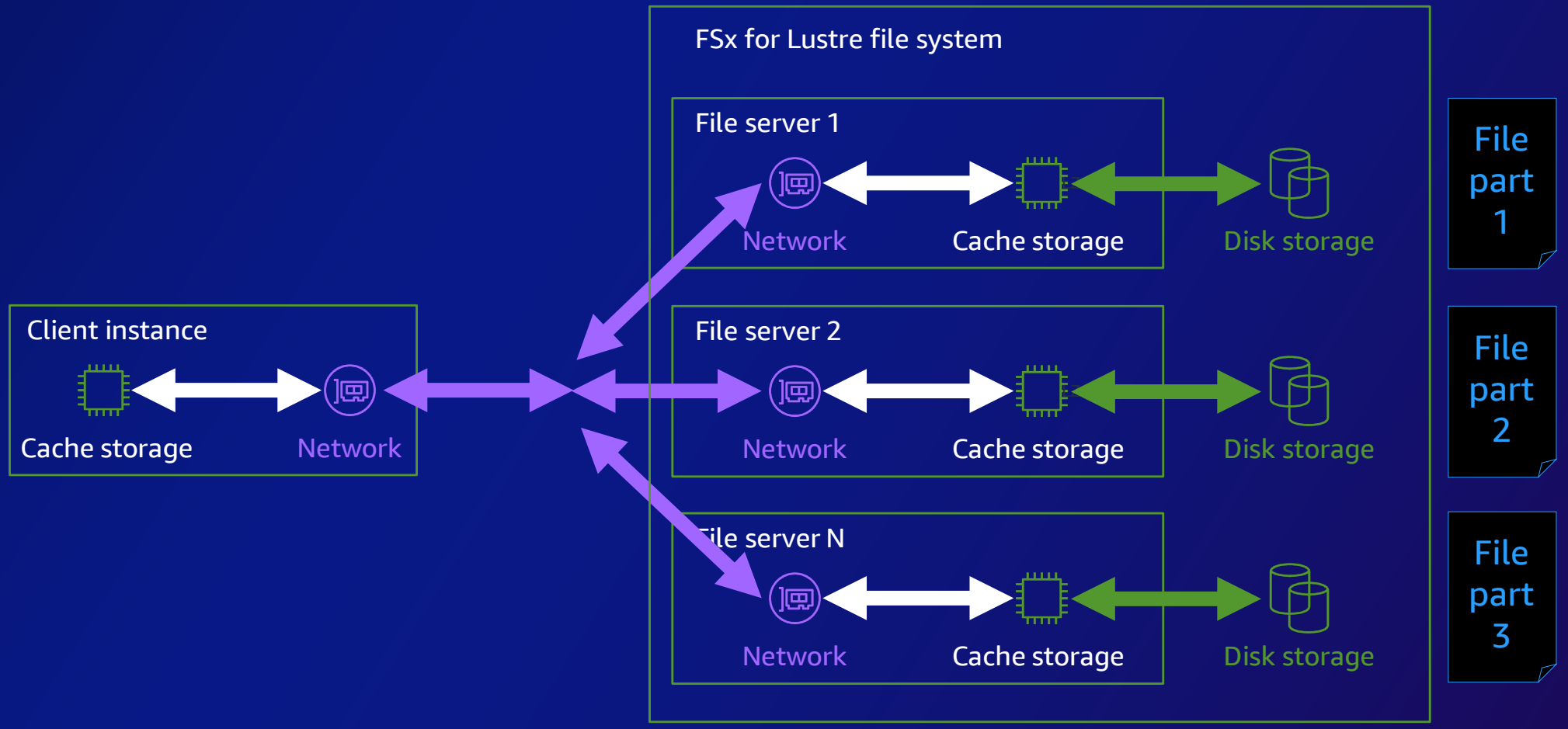
Highly scalable throughput capacity: How it works



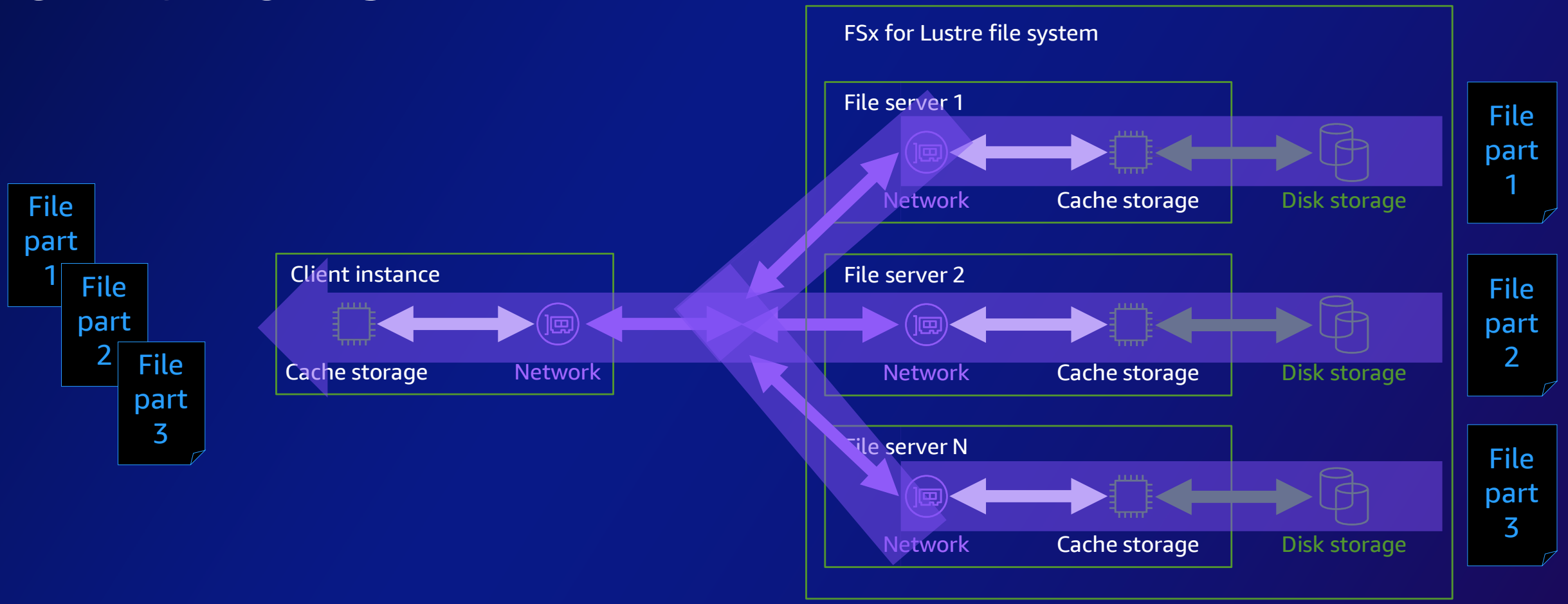
Highly scalable throughput capacity: How it works



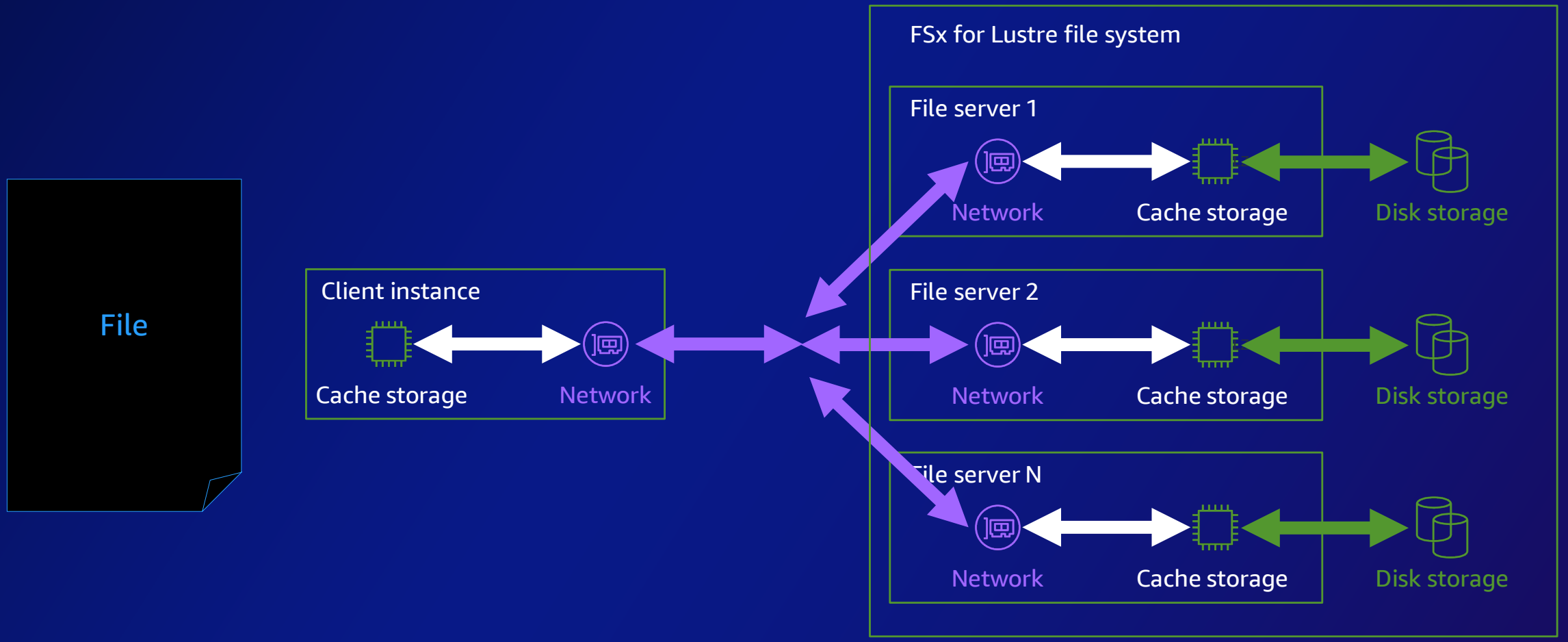
Highly scalable throughput capacity: How it works



Highly scalable throughput capacity: How it works

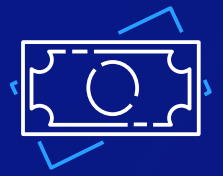


Highly scalable throughput capacity: How it works



Why consider Amazon Fsx for Lustre?

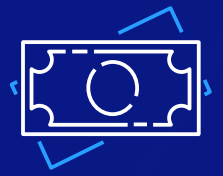
Highly scalable throughput capacity



Storage options optimized for price-performance

Access to Amazon S3 data through a fast file interface

Why consider Amazon Fsx for Lustre?



Storage options
optimized for
price-performance

Amazon FSx for Lustre deployment options

High and scalable performance

Amazon FSx for Lustre **SSD** Scratch file system

Amazon FSx for Lustre **SSD** Persistent file system

Amazon FSx for Lustre **HDD** Persistent file system

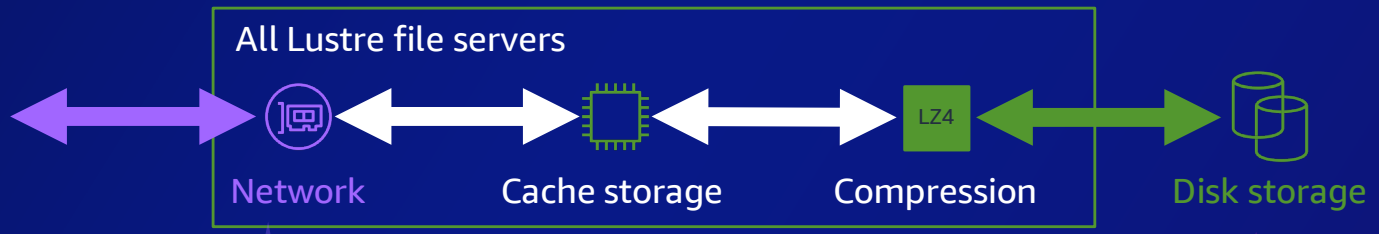
In all options, we support encryption at-rest and in-transit

Storage options optimized for price-performance

(*BOM REGION PRICING)

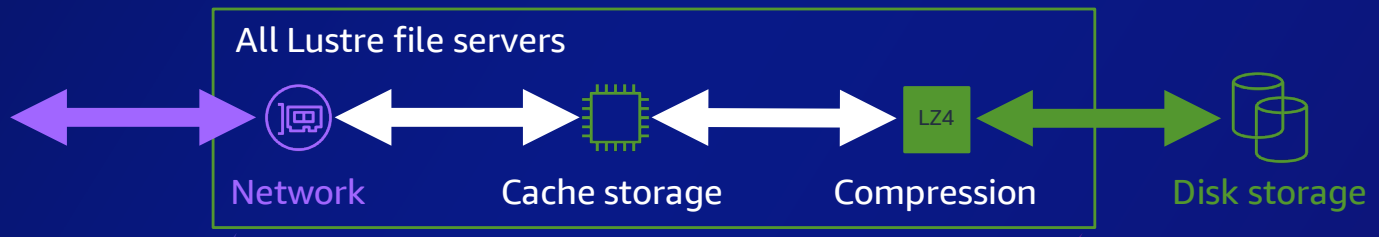
Storage type	Baseline throughput per TiB	Price per GB/s-hour
HDD Persistent	12 MB/s	$3.5
	40 MB/s	$3.5
SSD Persistent	125 MB/s	$2.17
	250 MB/s	$1.63
	500 MB/s	$1.37
	1000 MB/s	$1.24

Data compression: How it works



Deployment type	Network throughput	Cache storage (RAM)	Disk throughput
PERSISTENT-250	640 MB/s	6.8 GB	250 MB/s

Data compression: How it works

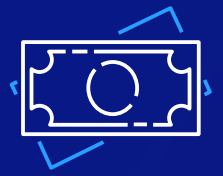


Deployment type	Network throughput	Cache storage (RAM)	Effective disk throughput	Disk throughput
PERSISTENT-200	640 MB/s	6.8 GB	500 MB/s	250 MB/s

2:1 compression ratio = 2x throughput

Why consider Amazon Fsx for Lustre?

Highly scalable throughput capacity

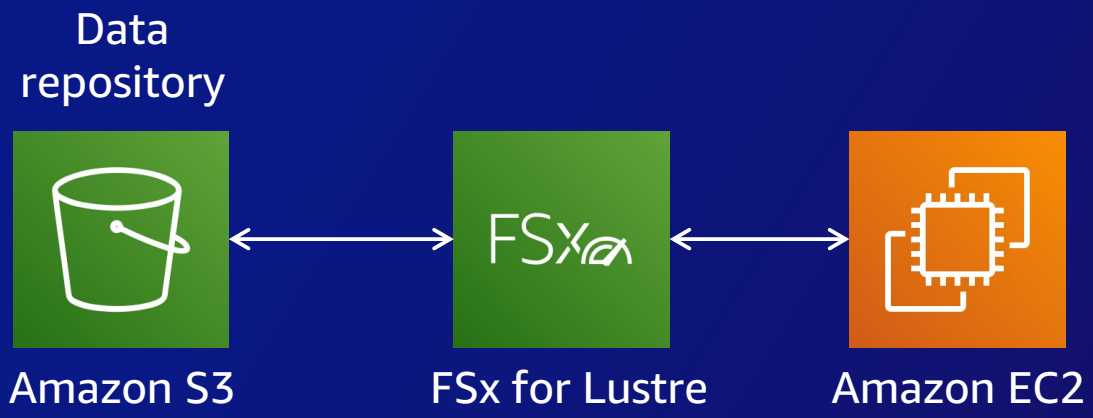


Storage options optimized for price-performance

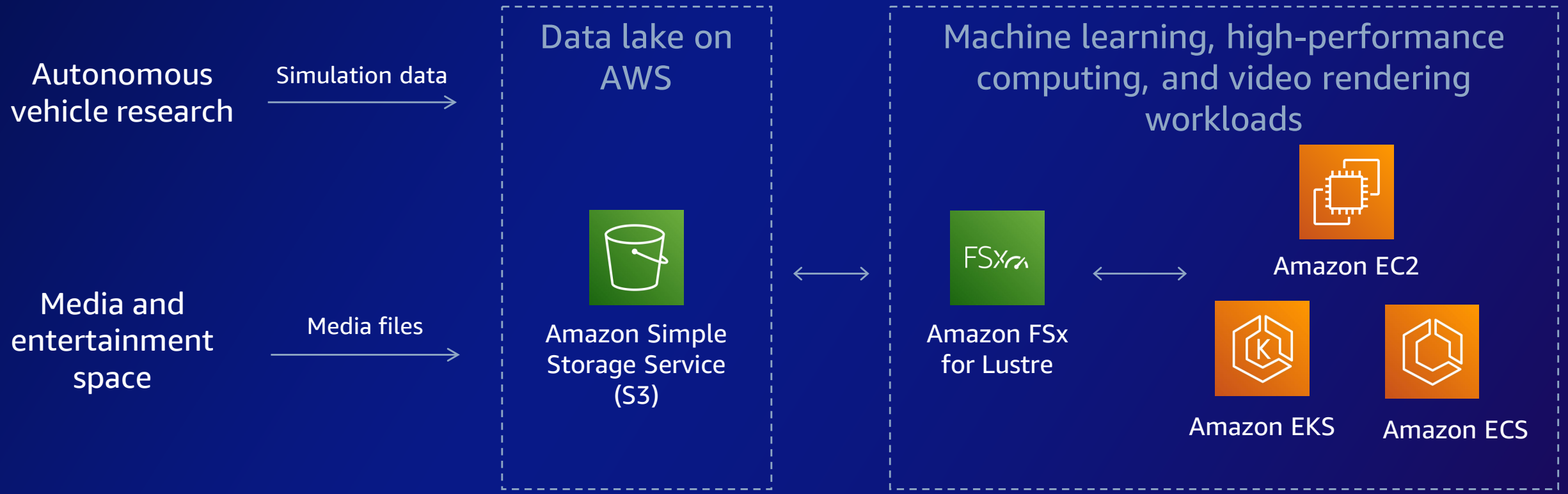
Access to Amazon S3 data through a fast file interface

Why consider Amazon Fsx for Lustre?

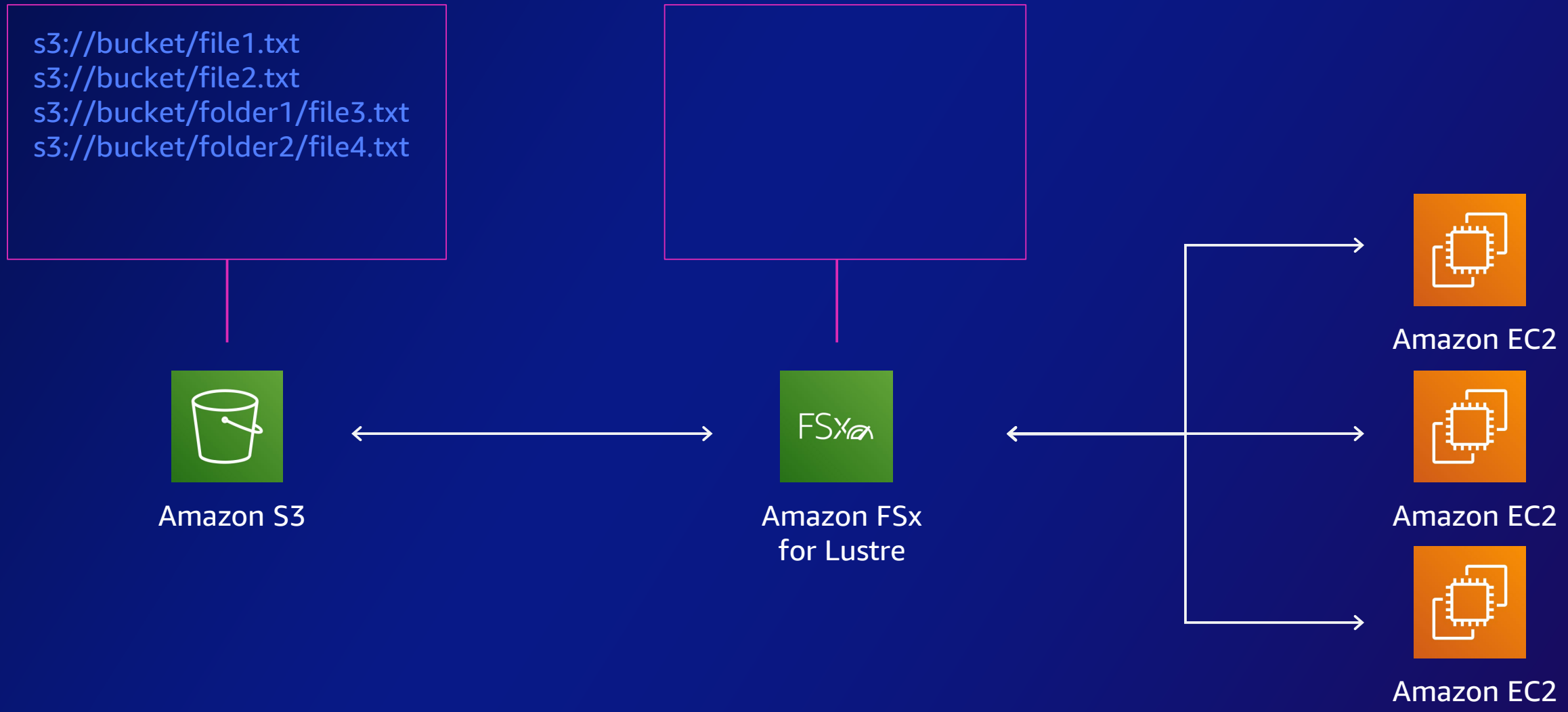
Access to Amazon S3 data through a fast file interface



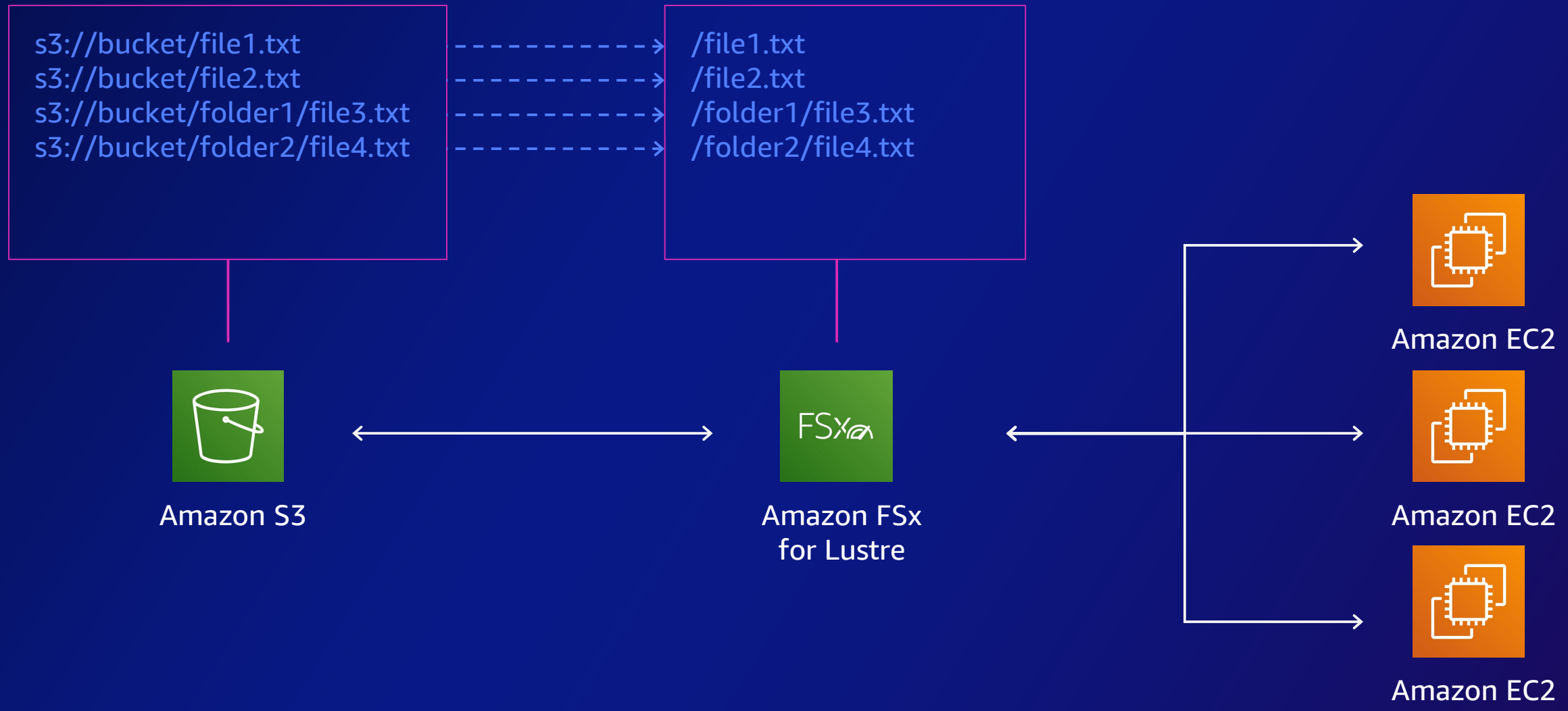
Example use cases: Process S3 data on FSx



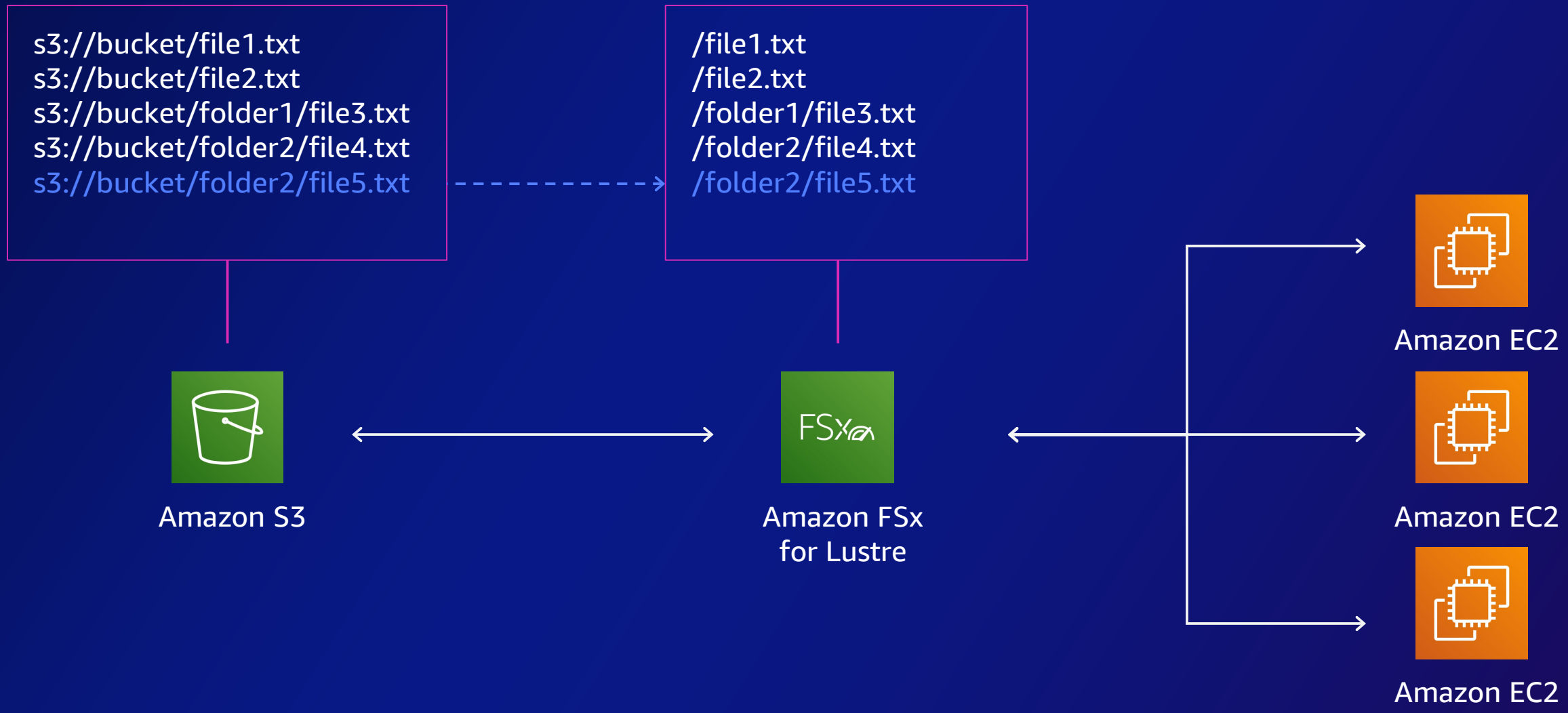
Fast file interface for data on S3: How it works



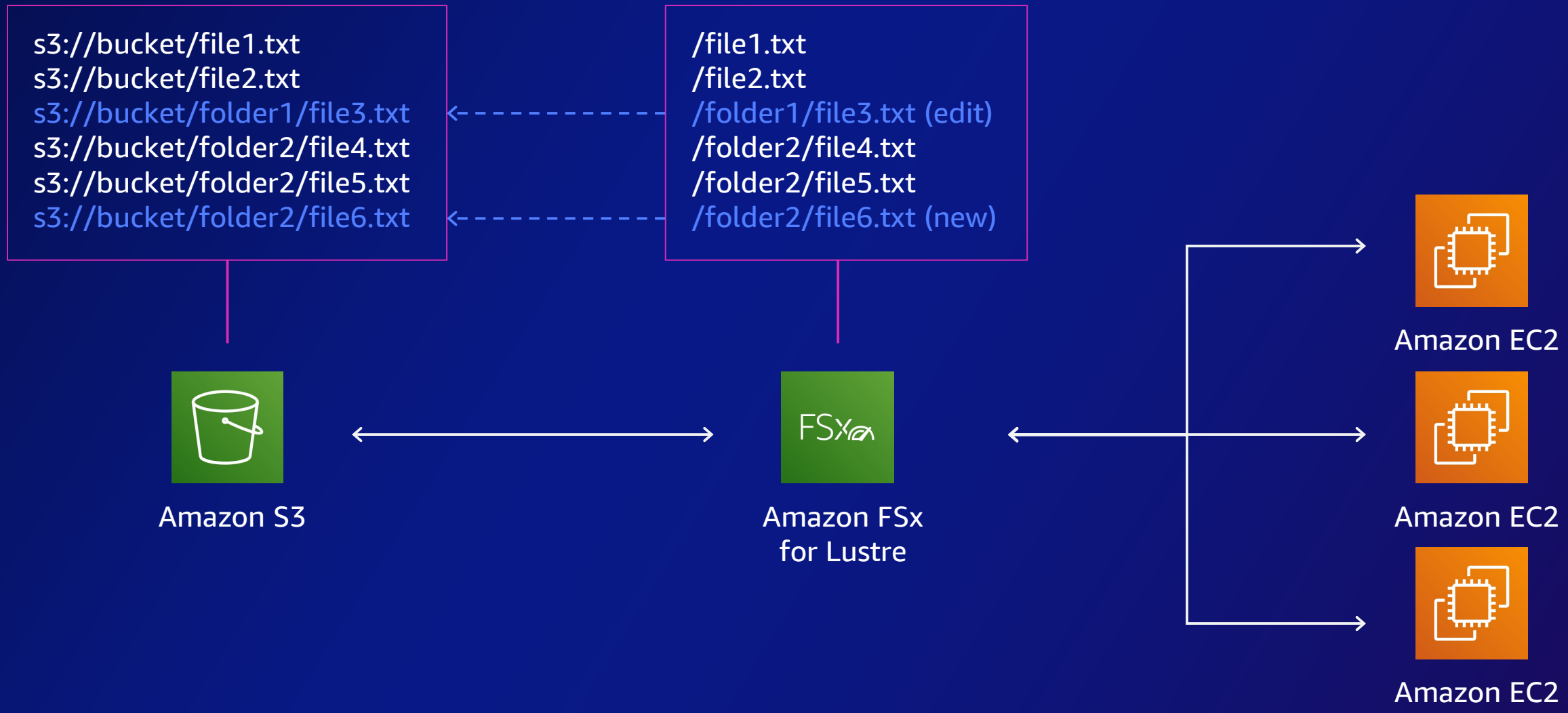
S3 objects appear on FSx file system



Updates on S3 are imported to FSx for Lustre



Updates on FSx for Lustre are exported to S3



Updates on FSx for Lustre are exported to S3

s3://bucket/file1.txt
s3://bucket/file2.txt
s3://bucket/folder1/file3.txt
s3://bucket/folder2/file4.txt
s3://bucket/folder2/file5.txt

/file1.txt
/file2.txt
/folder1/file3.txt (edit)
/folder2/file4.txt
/folder2/file5.txt
/folder2/file6.txt (new)

Amazon S3

Amazon FSx
for Lustre

Amazon EC2

Amazon EC2

Amazon EC2

Spin up / spin down with compute resources

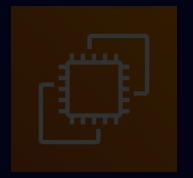
s3://bucket/file1.txt
s3://bucket/file2.txt
s3://bucket/folder1/file3.txt
s3://bucket/folder2/file4.txt
s3://bucket/folder2/file5.txt
s3://bucket/folder2/file6.txt

Amazon S3

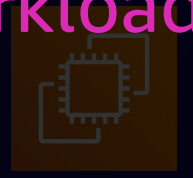
/file1.txt
/file2.txt
/folder1/file3.txt (edit)
/folder2/file4.txt
/folder2/file5.txt
/folder2/file6.txt (new)

Spin down resources between workloads

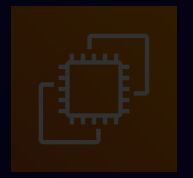
Amazon FSx
for Lustre



Amazon EC2



Amazon EC2



Amazon EC2

Demo

What all goes behind creating self managed Lustre Filesystem?

skillbuilder.aws

Your time is now

Build in-demand cloud skills your way

Thank you!

Please complete the session survey

Vibhu Pareek
Solutions Architect
AWS India