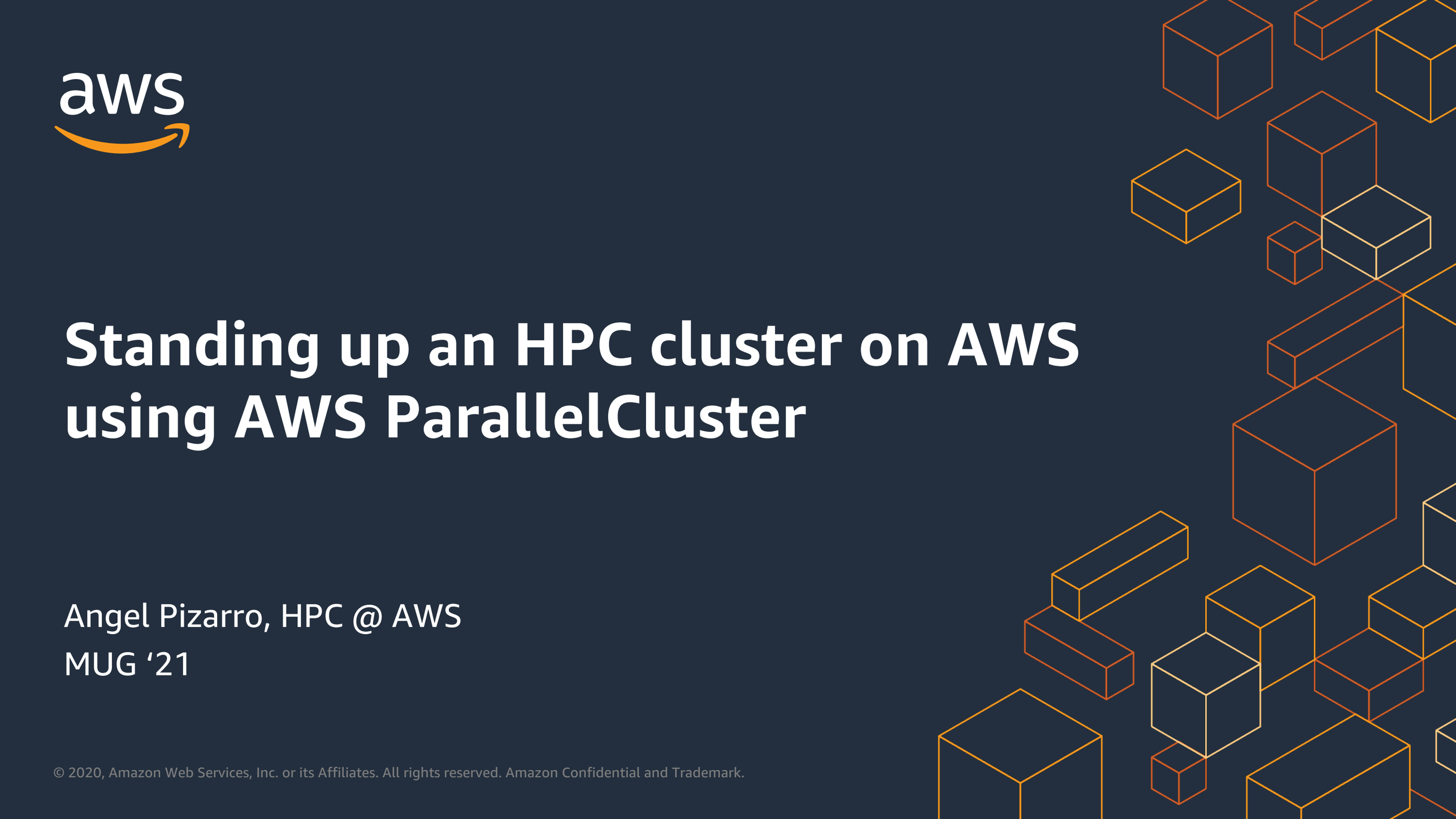
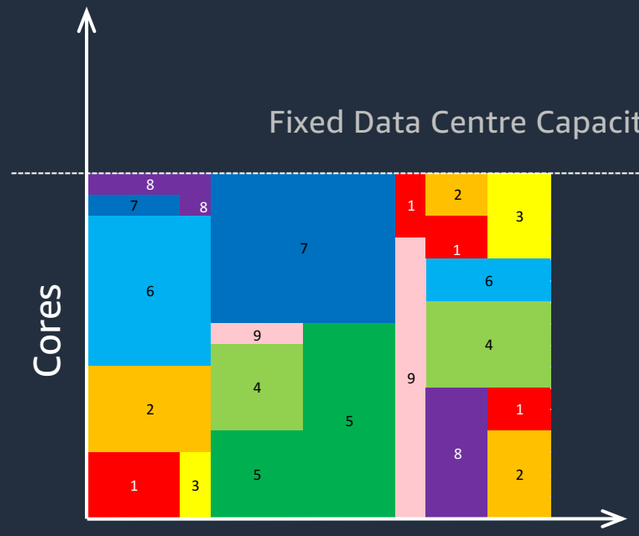


Standing up an HPC cluster on AWS using AWS ParallelCluster

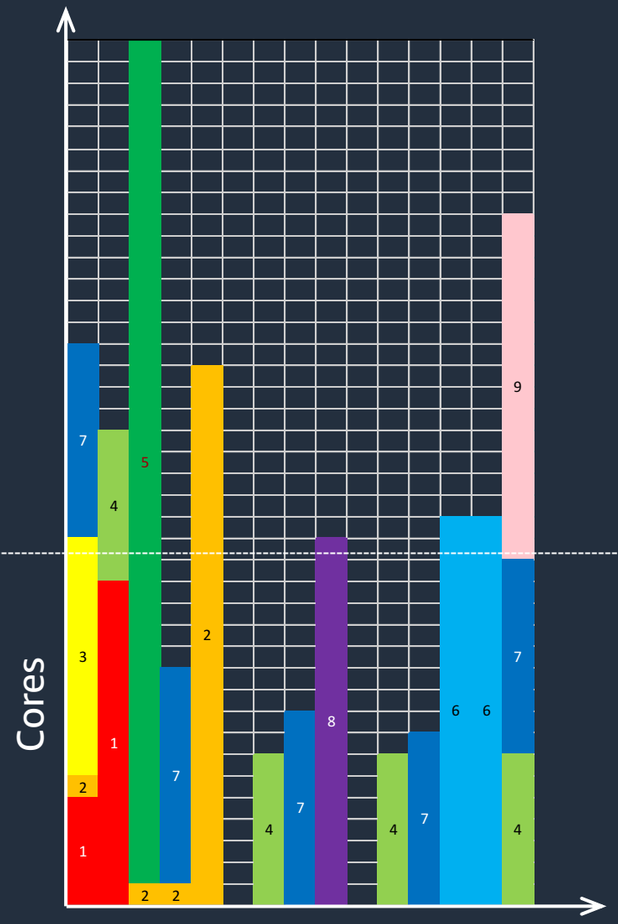
Angel Pizarro, HPC @ AWS

MUG ‘21

The metric for success should be time-to-results



Finite capacity, usually with long queues to wait in.

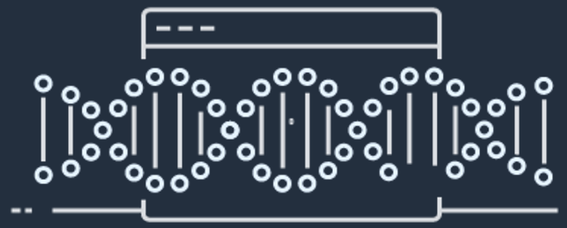


Massive capacity when needed to speed up time to results, and agile environment when additional hardware and software experimentation is needed.

"For every $1 spent on HPC, businesses see $463 in incremental revenues and $44 in incremental profit."*

Source: Hyperion ROI Study (http://www.hyperionresearch.com/roi-with-hpc/)

HPC workloads across industries



Life Sciences

Financial Services

Oil & Gas

Design & Engineering

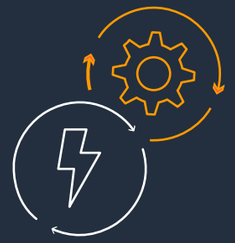
Climate & Geosciences

Autonomous Vehicles

HPC workloads with different compute and throughput characteristics

Tightly-coupled workloads

Loosely-coupled workloads



Accelerated computing

Visualization

AI/ML

High volume data analytics

What do HPC customers need

High Performance

Reliability

Low Latency/Overhead

Consistency & Jitter free network

Fairness

Scalable Reliable Datagram (SRD)

A reliable high-performance lower-latency network transport

Guaranteed delivery

- Does not consume any resources on EC2 instances

Network aware multipath routing

- Optimally utilizes all network paths (ECMP), no hot-spots

Orders of magnitude lower tail latency & jitter

- Fast recovery from network events

No ordering guarantees

- No head-of-line blocking

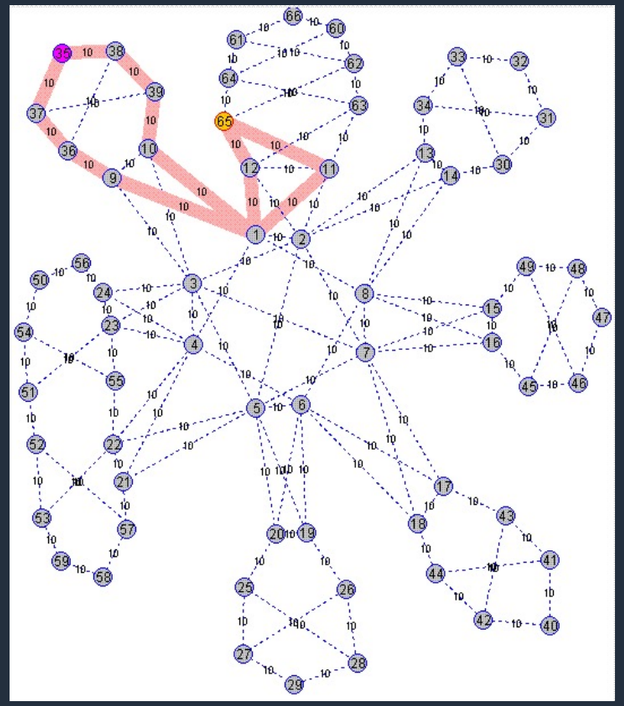
https://ieeexplore.ieee.org/document/9167399

aws

Elastic Fabric Adapter – Networks built to scale

Up to 400 Gbps networking bandwidth

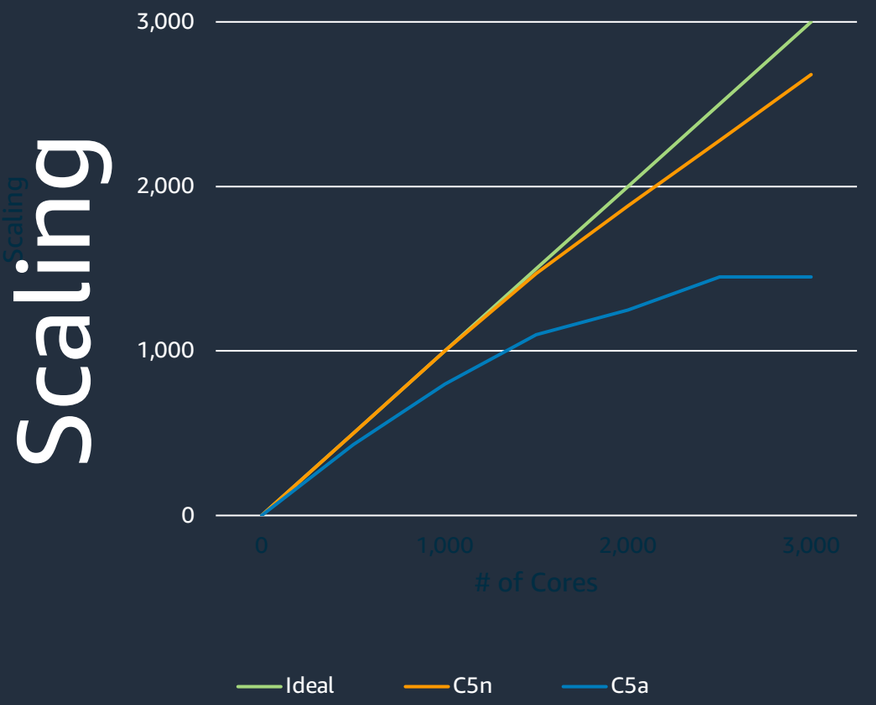
- ✓ OS bypass
- ✓ GPUdirect and RDMA
- ✓ Libfabric core supports wide array of MPIs and NCCL



OCMP-enabled packet spraying and cloud-scale congestion control

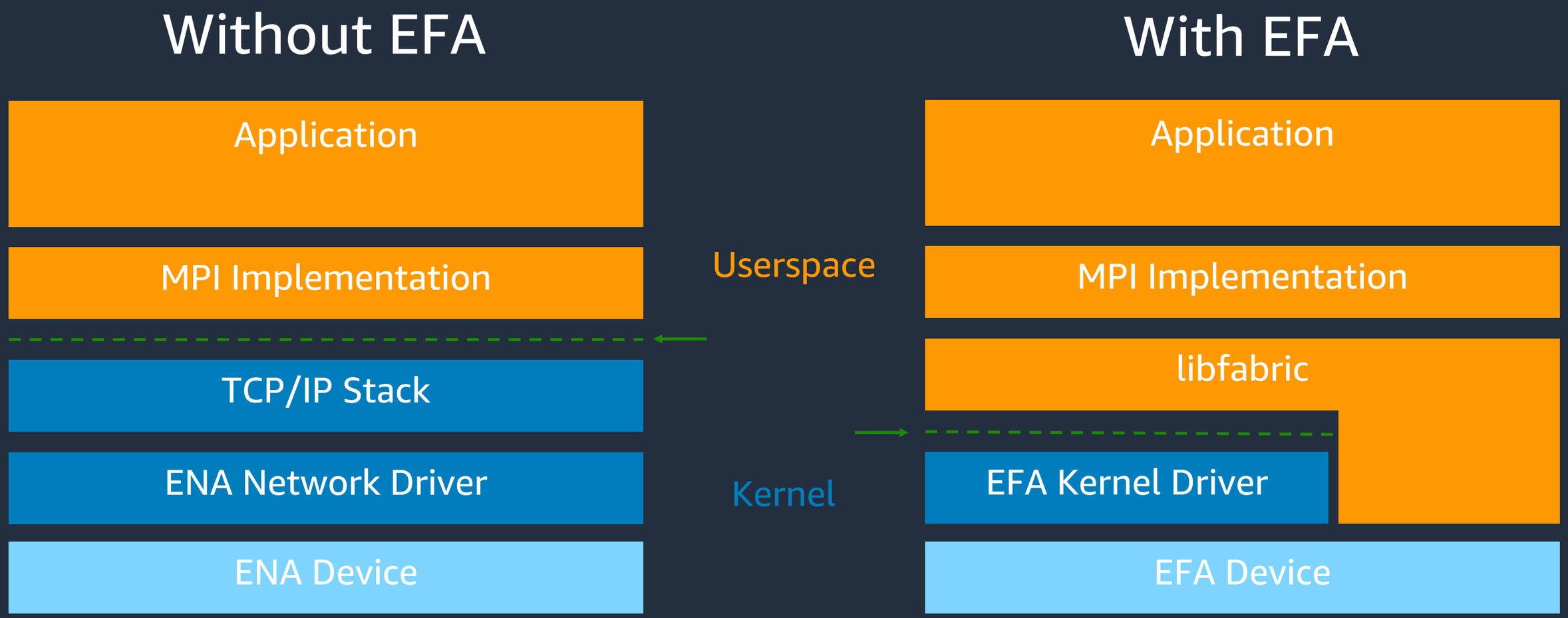
ANSYS Fluent

External flow over F1 race car (140M cell mesh)



At ~3,000 cores (~83 nodes), C5n+EFA shows ~89% scaling efficiency vs ~48% using C5 w/o EFA

HPC software stack on Amazon EC2



Elastic Fabric Adapter (EFA)

Scale **tightly coupled** HPC applications on AWS

Intel

Custom Intel Xeon Scalable processor

NVIDIA

A100, V100, and T4 GPUs

Arm

AWS Graviton2

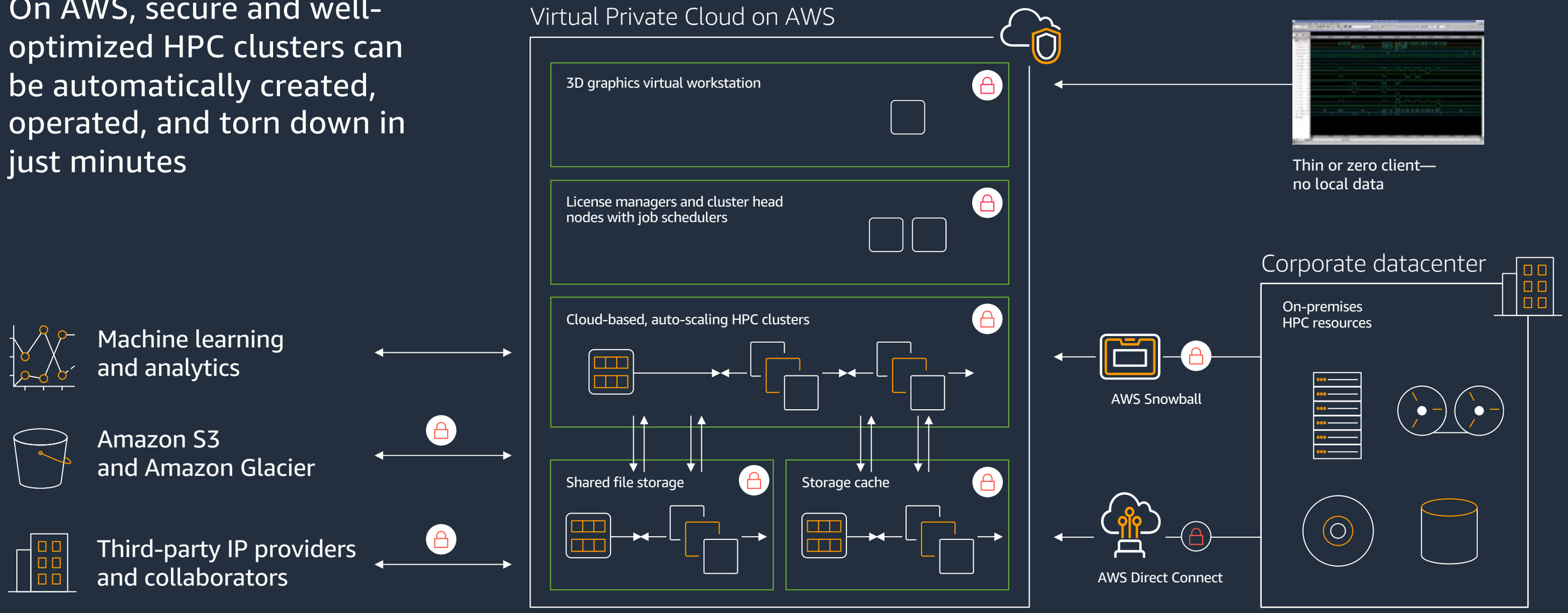
EFA

AWS HPC/ML Network Interface

- Instance flexibility
- Infrastructure elasticity
- High data throughput
- Low-latency message passing
- Faster application time-to-completion

High Performance Computing (HPC) on AWS

On AWS, secure and well-optimized HPC clusters can be automatically created, operated, and torn down in just minutes



AWS ParallelCluster

intel select solution

One-stop shop to set up your HPC cluster

Integrated with AWS services you need

Amazon FSx for Lustre

Amazon EC2 instances

EFA

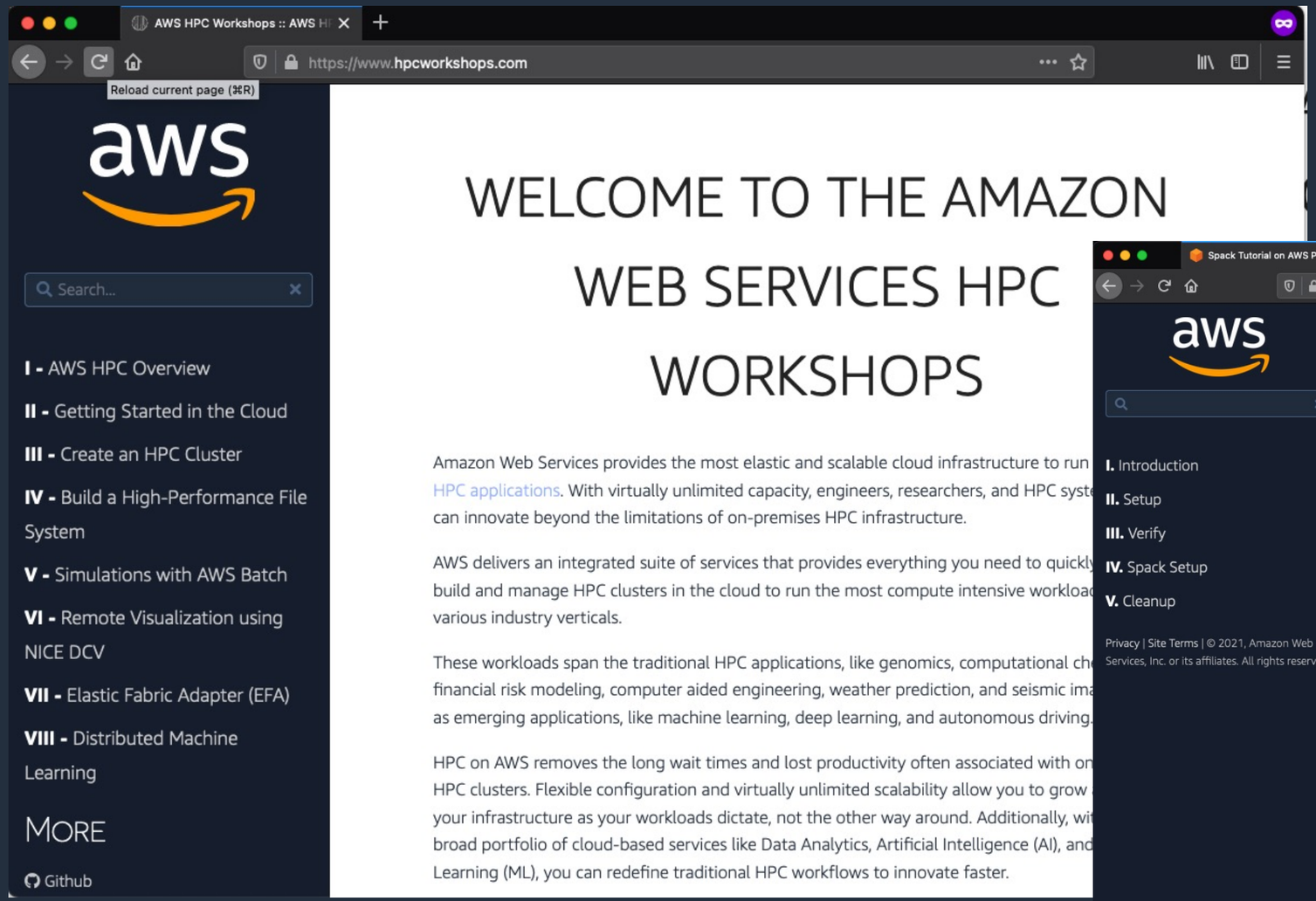
NICE DCV

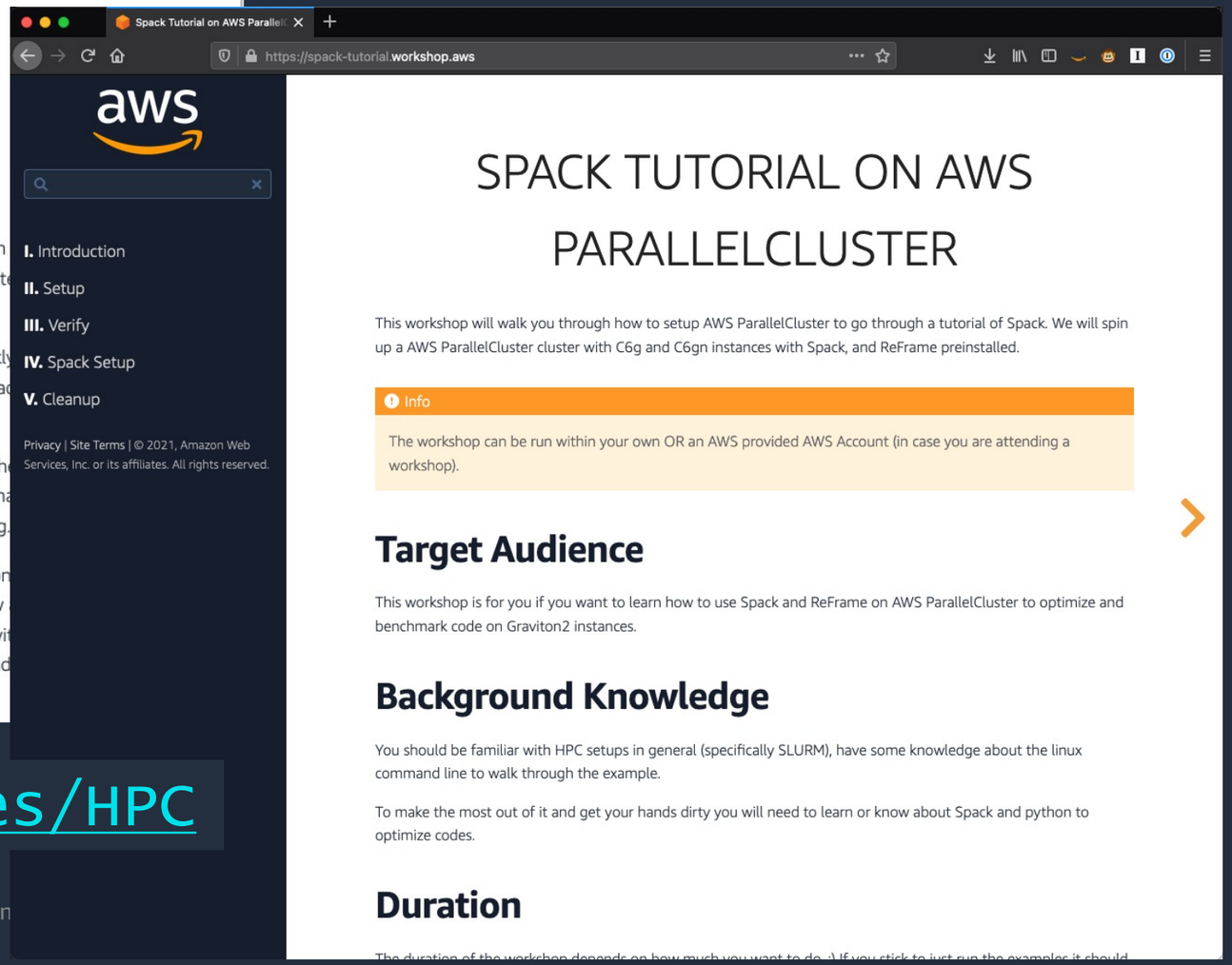
AWS Batch

Running EFA using AWS ParallelCluster

Subtitle

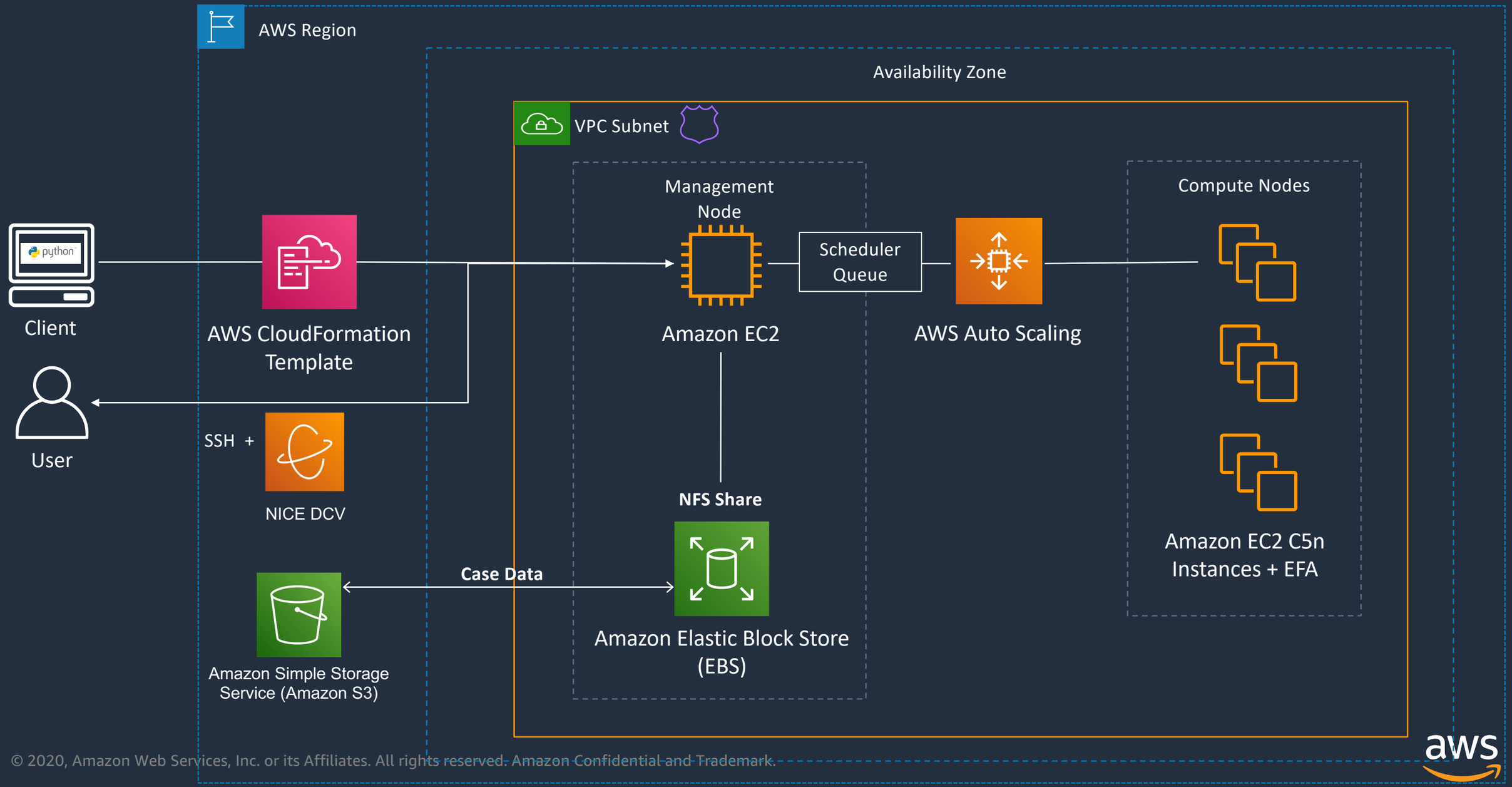
Getting started using Spack with AWS ParallelCluster



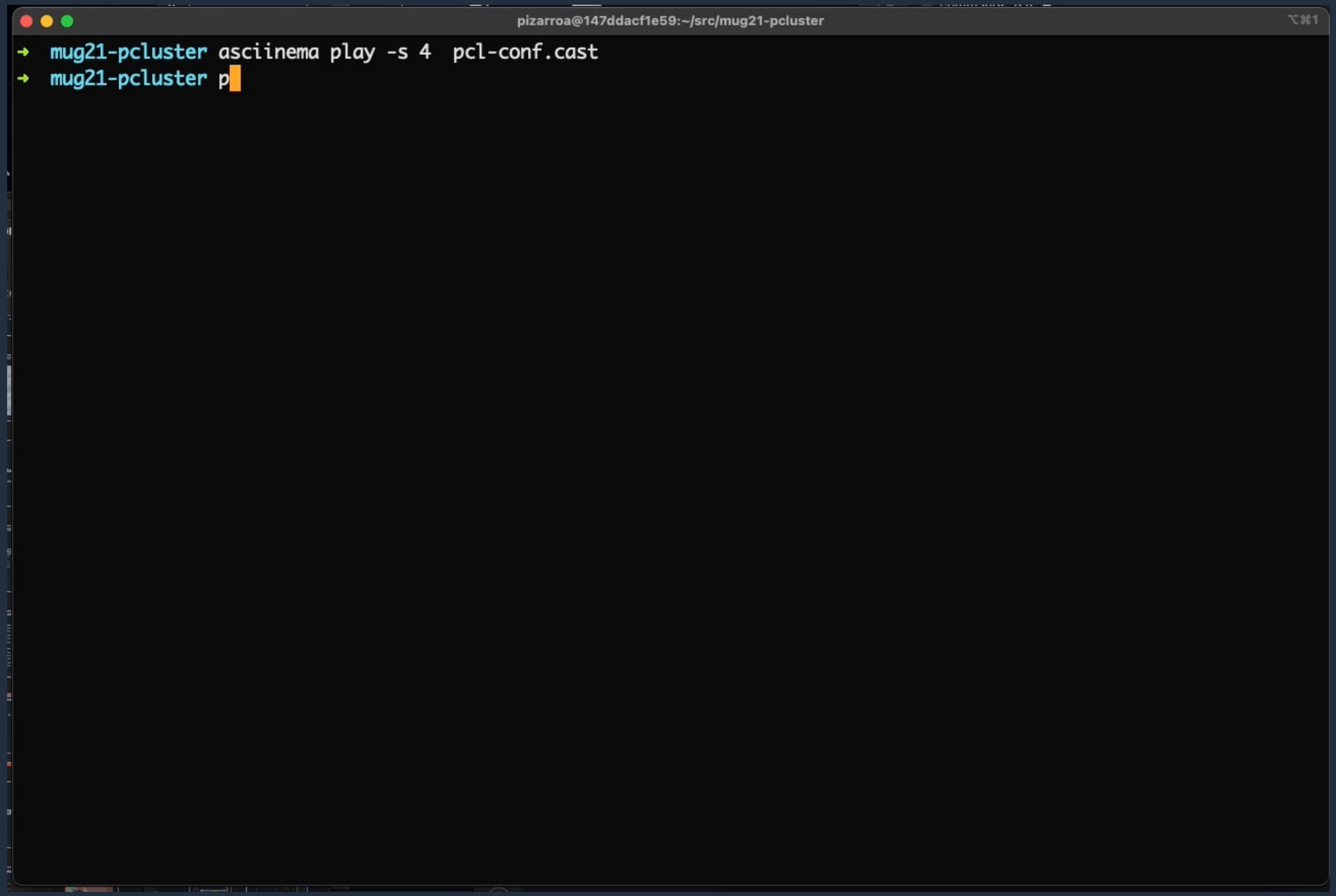


https://workshops.aws/categories/HPC

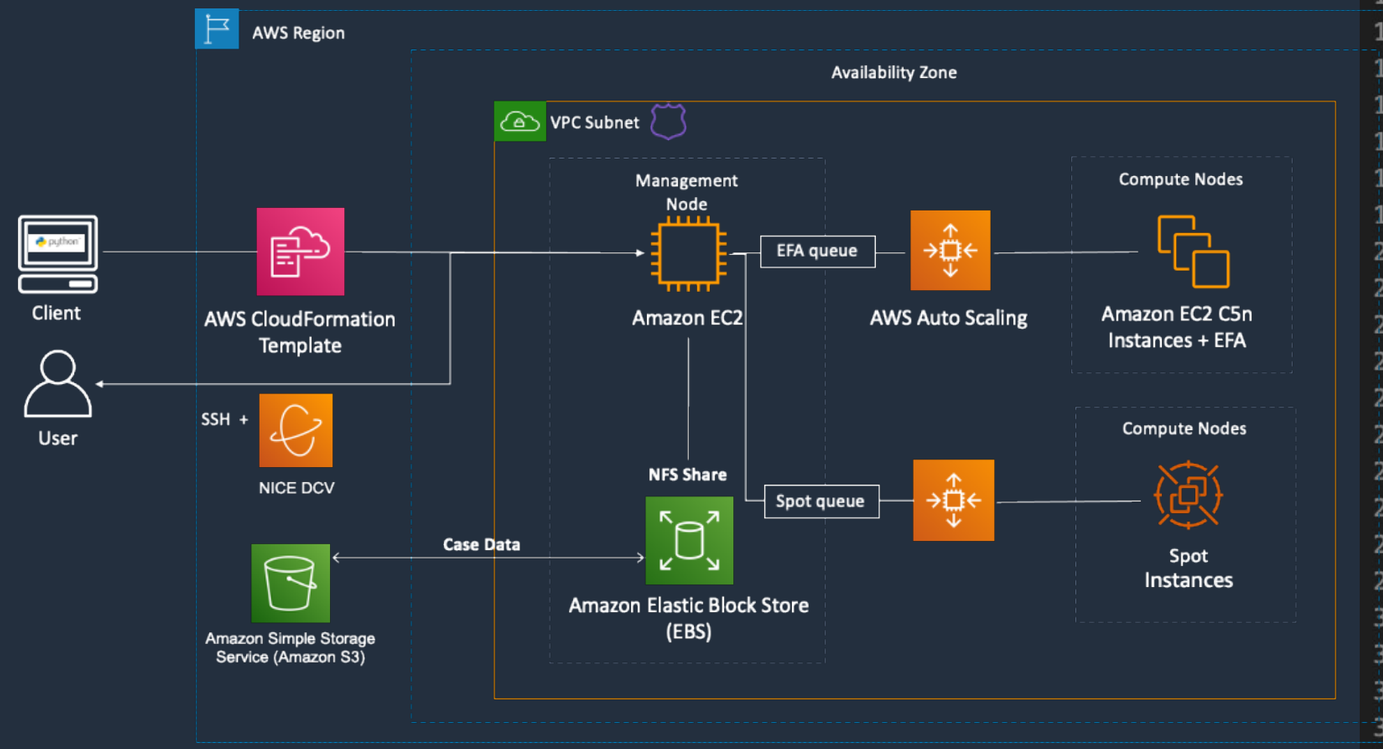
AWS ParallelCluster architecture



DEMO TIME (sort of 😅)



Feature: Multiple Slurm Queues

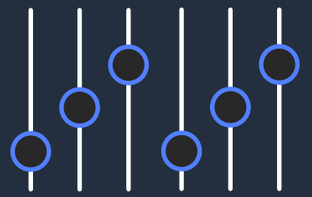



```
12  [scaling demo]
13  scaledown_idletime = 5              # optional, defaults to 10 minutes
14
15  [cluster multi-queue]
16  key_name = pc-key-mug21
17  base_os = alinux2
18  scheduler = slurm
19  master_instance_type = c5.2xlarge
20  vpc_settings = default
21  scaling_settings = demo
22  queue_settings = spot,ondemand
23
24  [queue spot]
25  compute_resource_settings = spot_i1,spot_i2
26  compute_type = spot                 # optional, defaults to ondemand
27
28  [compute_resource spot_i1]
29  instance_type = c5.xlarge
30  min_count = 0                       # optional, defaults to 0
31  max_count = 10                      # optional, defaults to 10
32
33  [compute_resource spot_i2]
34  instance_type = t3.medium
35  min_count = 1
36  initial_count = 2
37
38  [queue ondemand]
39  compute_resource_settings = ondemand_i1
40  disable_hyperthreading = true       # optional, defaults to false
```

https://docs.aws.amazon.com/parallelcluster/latest/ug/tutorial-mqm.html

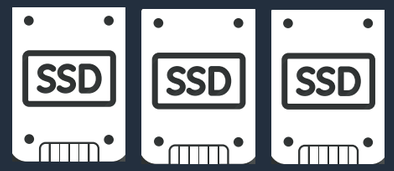
Feature: Amazon FSx for Lustre

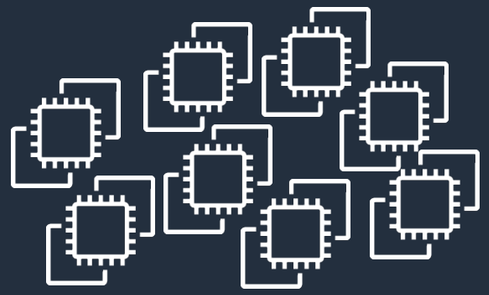
Parallel file system



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

SSD-based





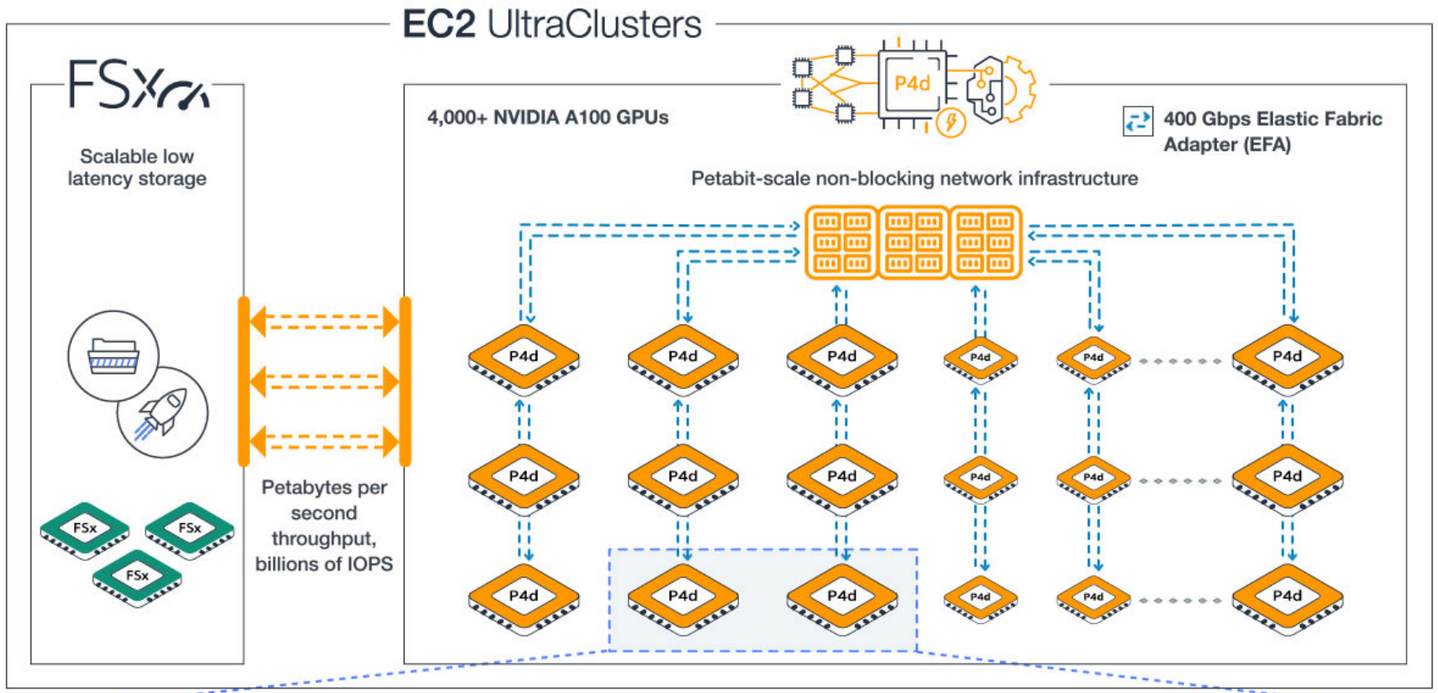
Supports hundreds of
thousands of cores

```
15  [cluster default]
16  key_name = pc-key-mug21
17  scheduler = slurm
18  master_instance_type = c5.2xlarge
19  base_os = alinux2
20  vpc_settings = default
21  queue_settings = compute
22  scaling_settings = demo
23  fsx_settings = fs-mug21
24
25  [fsx fs-mug21]
26  shared_dir = /fsx
27  storage_capacity = 3600
28  imported_file_chunk_size = 1024
29  export_path = s3://bucket/folder
30  import_path = s3://bucket
31  weekly_maintenance_start_time = 1:00:00
```


aws

EC2 UltraClusters of P4d

Supercomputing-class performance for deep learning workflows



- Based on NVIDIA A100 GPUs
- Availability to scale-out to large clusters for distributed training
- 2.5x better deep learning performance and 60% lower cost to train
- EC2 UltraClusters with EFA enables 400 Gbps and allows you to scale to over 4,000 GPUs

aws

Thank you!

https://docs.aws.amazon.com/parallelcluster/latest/ug/what-is-aws-parallelcluster.html

https://workshops.aws/categories/HPC
https://spack-tutorial.workshop.aws/
https://www.hpcworkshops.com/

https://mvapich.cse.ohio-state.edu/userguide/mv2x-aws/

https://mvapich.cse.ohio-state.edu/userguide/userguide_spack/

aws