AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

CMP312

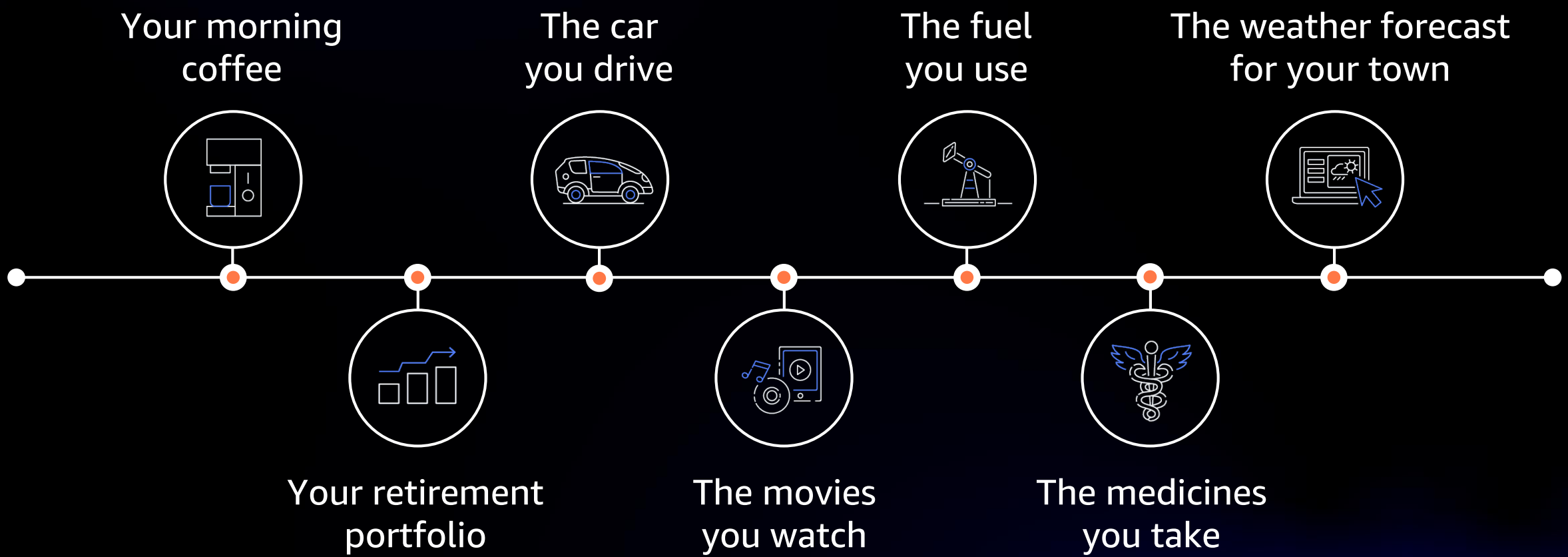
The future of HPC is looking a lot like ML

Barry Bolding (he/him)
Director HPC, Quantum, and Autonomous Computing
AWS

Jyothi Venkatesh (she/her)
HPC Specialist Solutions Architect
AWS

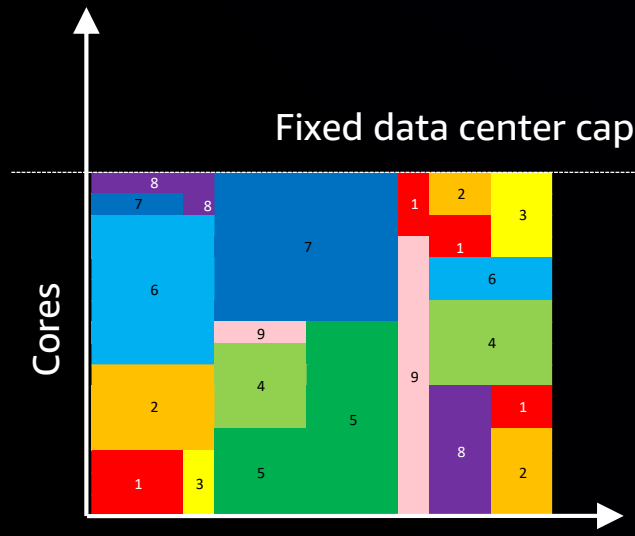
Terry Cunningham (he/him)
CEO
Descartes Labs

High performance computing (HPC) impacts your life every day

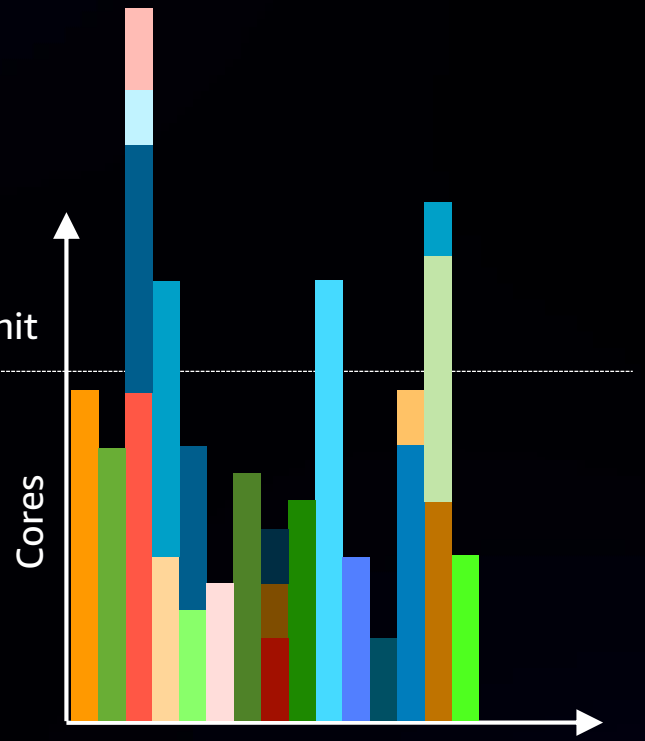


aws

We think the metric for success for any business 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 $507 in incremental revenues and $47 in incremental profit."***

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

Working backward: The Amazonian way

Working backward from customer requirements

Functional asks

Faster development cycles

Tools to derive insights

Easier access to data

Operational asks

Security, data governance

Cost management

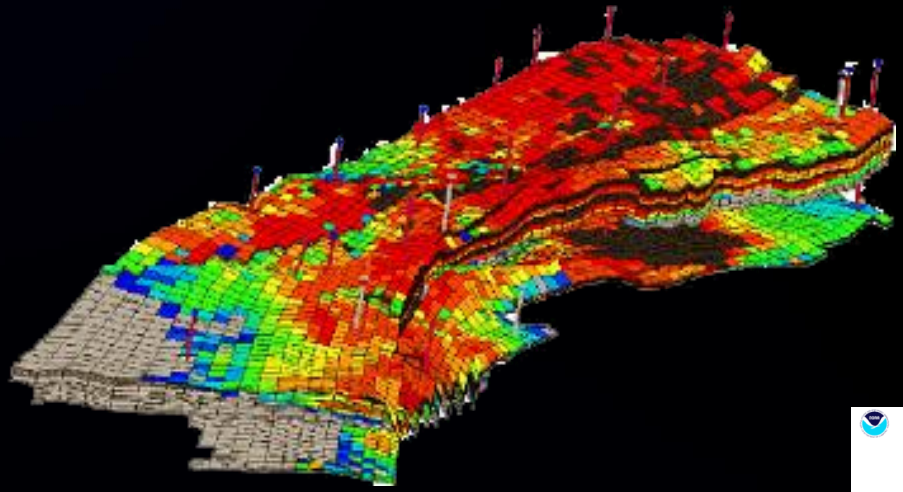
Data transfer and management

New: Can I derive increased business value from machine learning (ML) + HPC?

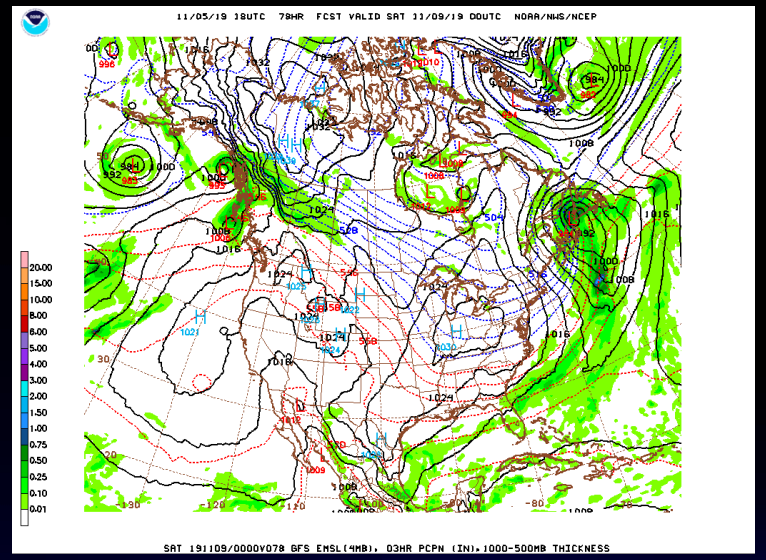
aws

Any type of HPC workload is possible on AWS

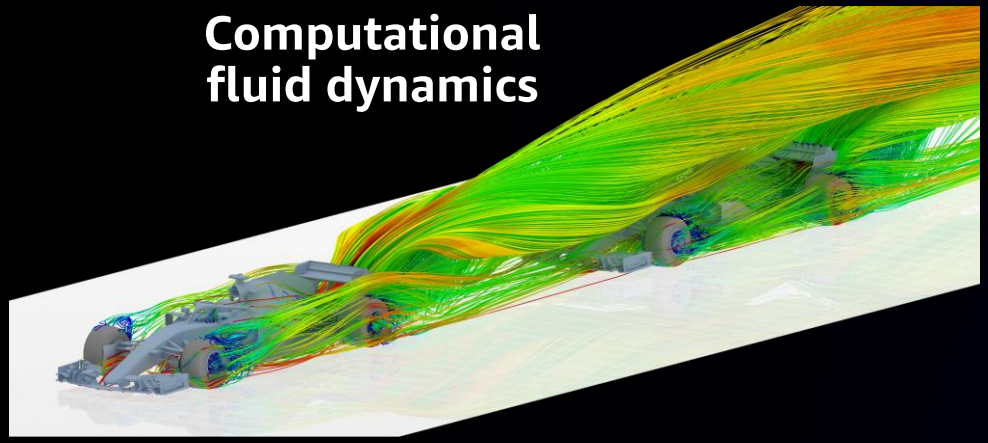
Reservoir simulations



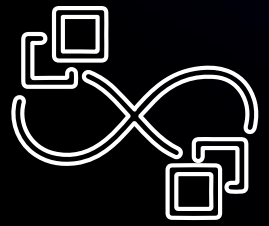
Weather forecasting



Computational fluid dynamics

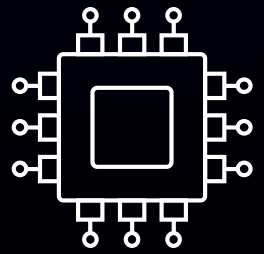


Key services that enable HPC on AWS



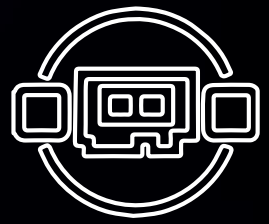
Visualization

AWS ParallelCluster



Compute

Amazon EC2

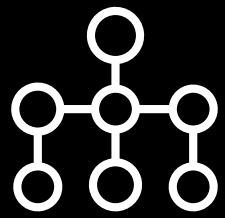


Networking

Elastic Fabric Adapter (EFA)
+
up to 400 Gbps networking

Storage

Amazon FSx for Lustre and Amazon FSx for NetApp ONTAP



Orchestration

AWS Batch

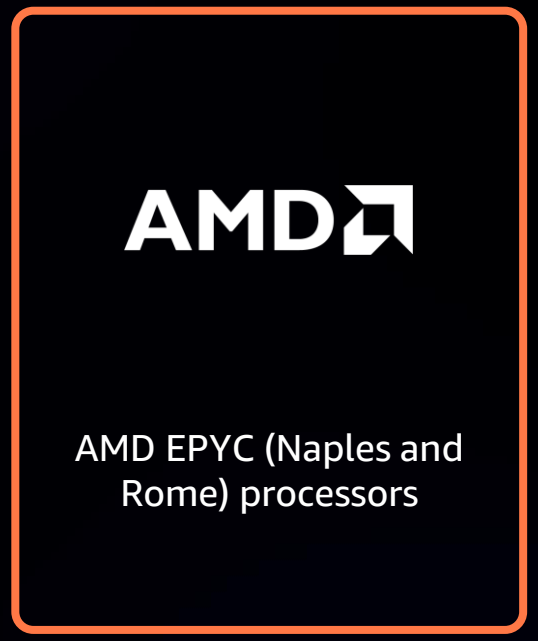
AWS ParallelCluster

EnginFrame

AWS takes a systemic view of workload

aws

Broadest choice of processors in the cloud



Modular building blocks for rapid design and delivery of EC2 instances

aws

Broadest choice of accelerators in the cloud

NVIDIA V100 Tensor Core and T4 Tensor Core GPUs

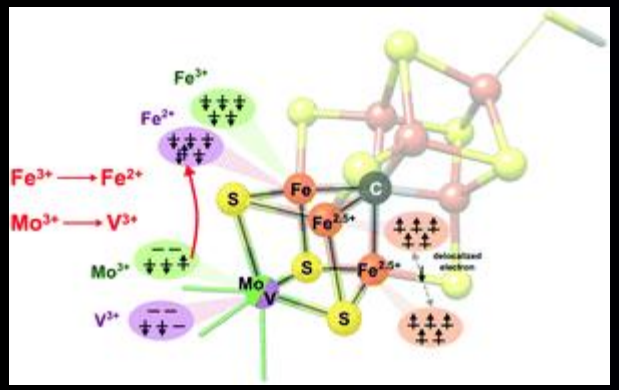
Xilinx FPGAs for custom hardware acceleration

AWS Inferentia for ML inference

Right accelerator for any workload

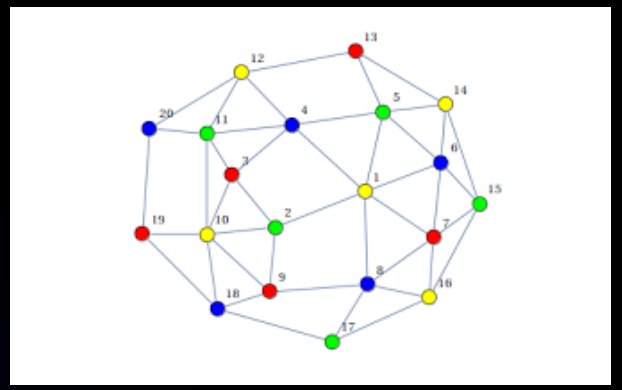
The disruptive potential of quantum computing

Quantum computing holds the promise to **speed up solutions to hard problems**



Computational chemistry

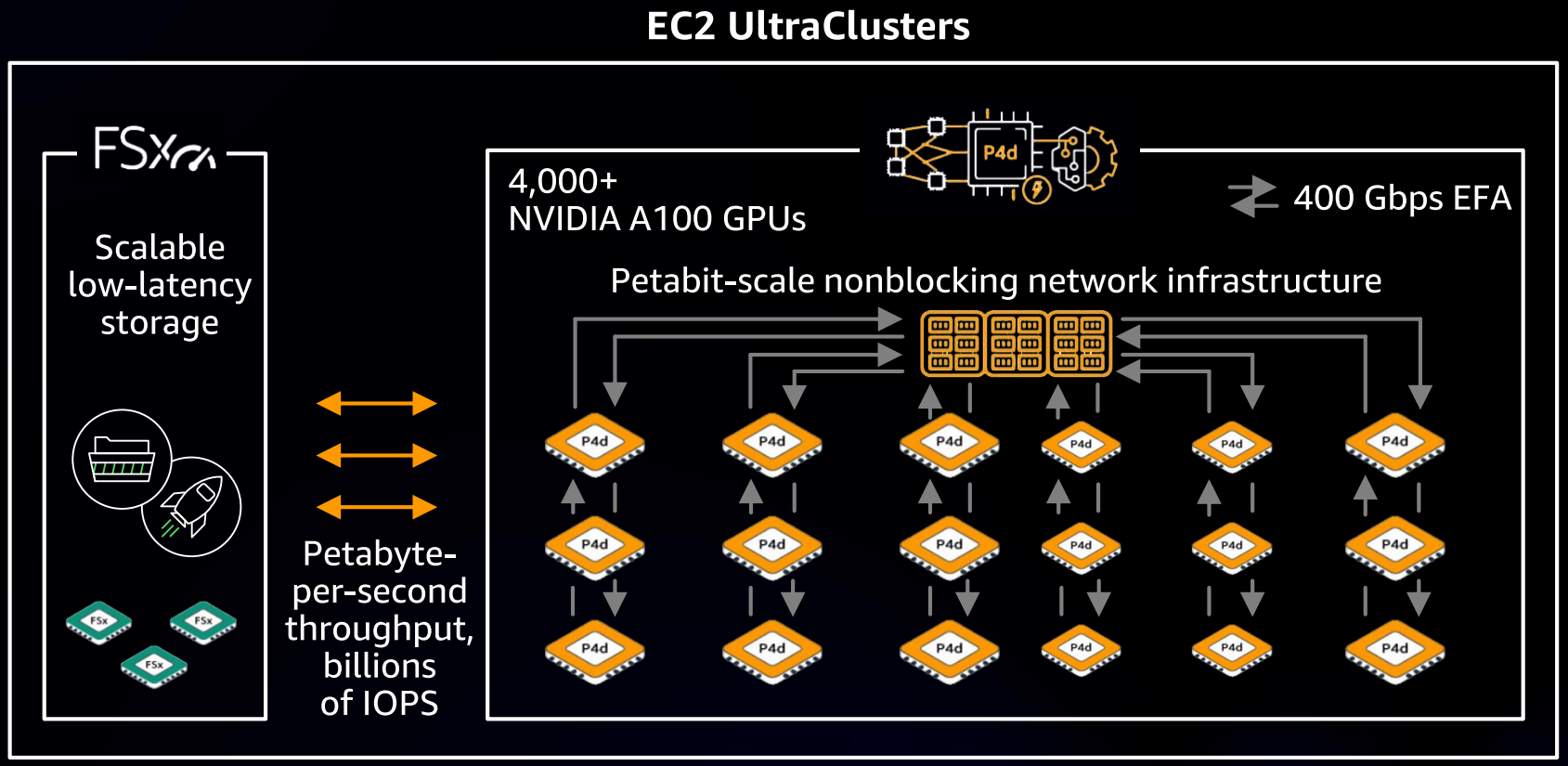
Machine learning



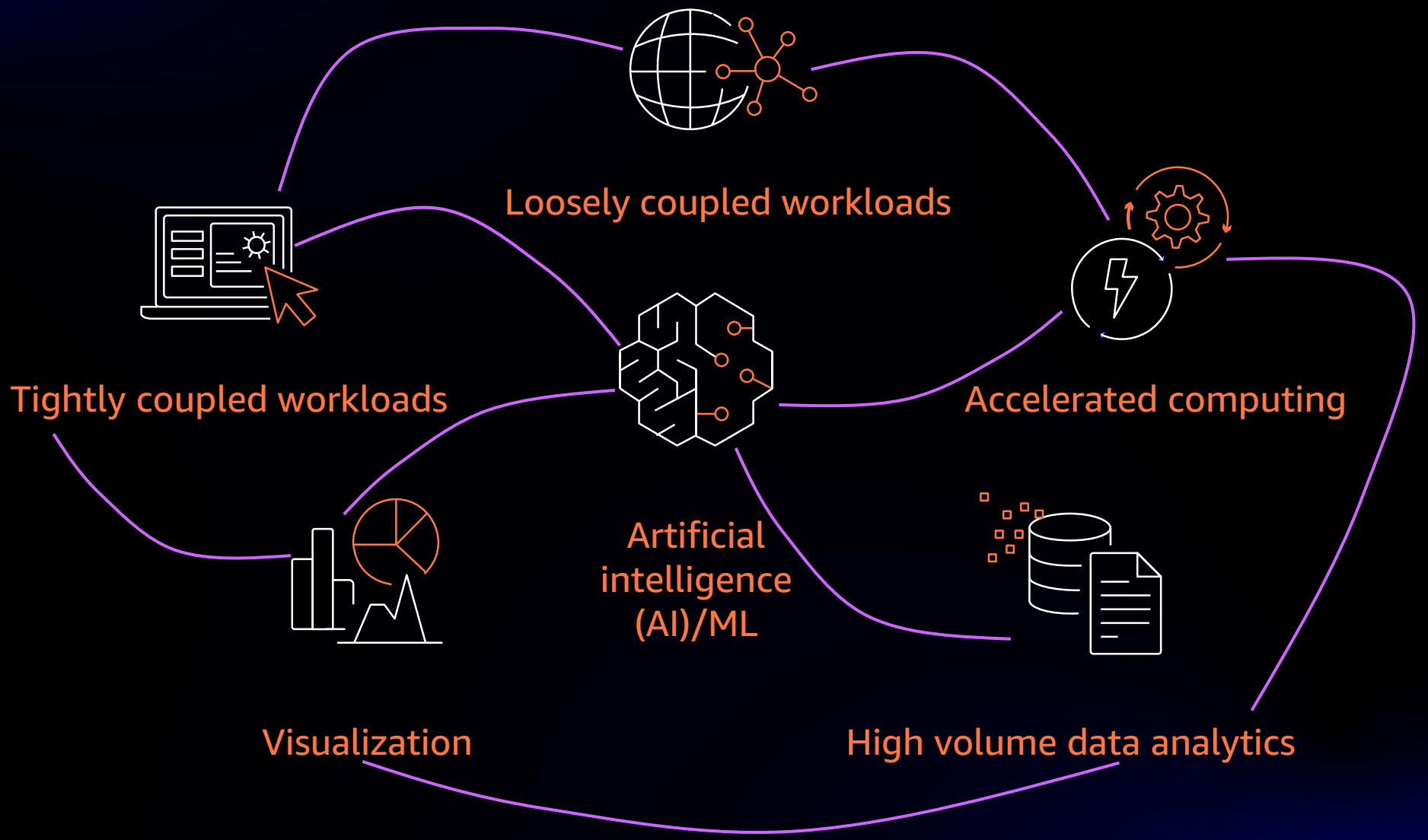
Optimization

EC2 UltraClusters: Supercomputing-class performance for HPC and ML/DL

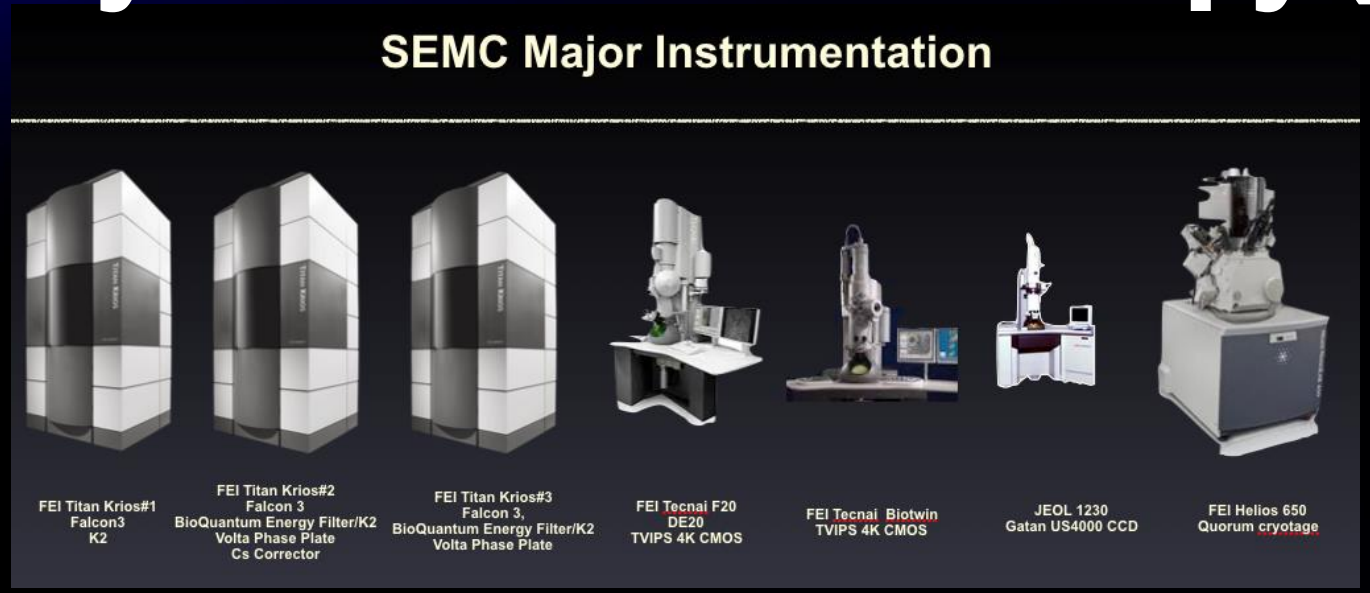
- Based on NVIDIA A100 graphics processing units (GPUs)
- Availability to scale out to large clusters for distributed training
- 2.5x better deep learning (DL) performance and 60% lower cost to train
- EC2 UltraClusters with EFA enable 400 Gbps and allow you to scale to over 4,000 GPUs



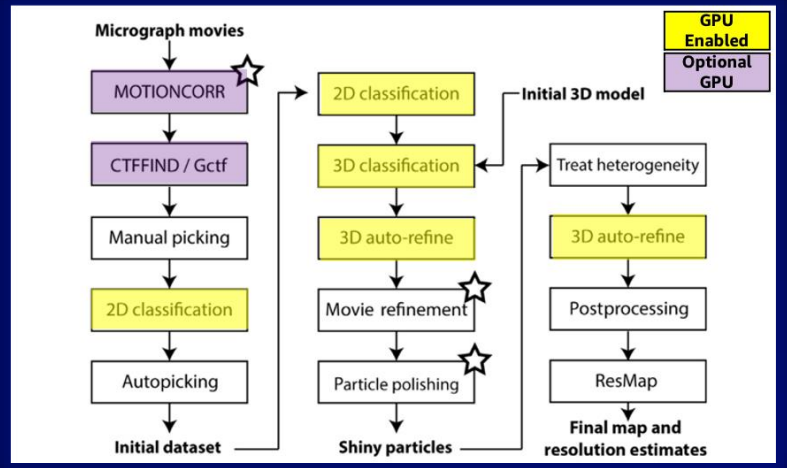
Varied compute and throughput characteristics

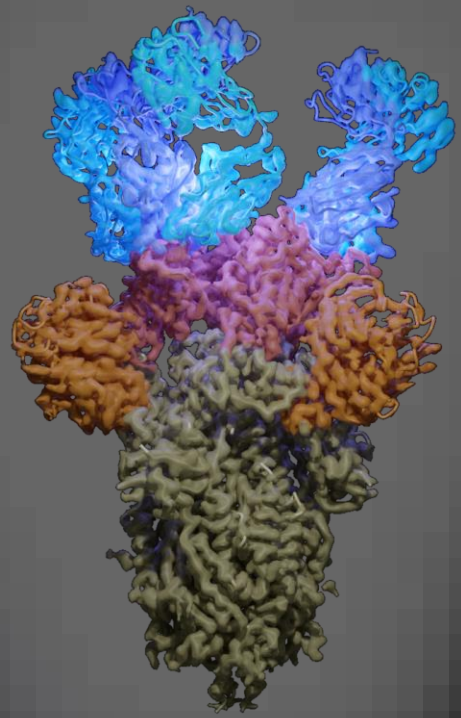


Cryo-Electron Microscopy (Cryo-EM)



https://nramm.nysbc.org/infrastructure/



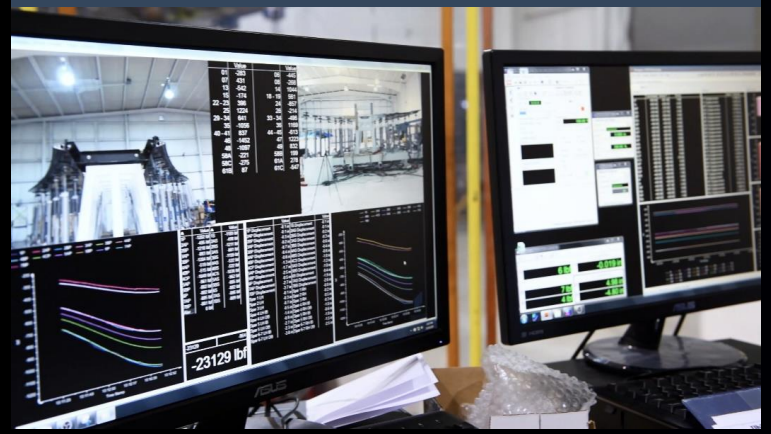


Antibody 2-4
Liu et al., *Nature* 2020
PDB 6XEY
10.1038/s41586-020-2571-7
Structure courtesy of Micah Rapp

https://aws.amazon.com/blogs/industries/building-a-gpu-enabled-cryoem-workflow-on-aws/

Case study: Boom Supersonic

GOES ALL-IN ON AWS



Boom Supersonic used AWS to **accelerate the design and construction** of its supersonic aircraft

Boom can run thousands of advanced computer simulations concurrently, resulting in an estimated **6x increase in productivity versus an on-premises environment**

Boom used more than **53 million compute hours** on AWS to **complete design and testing** of its Overture airliner

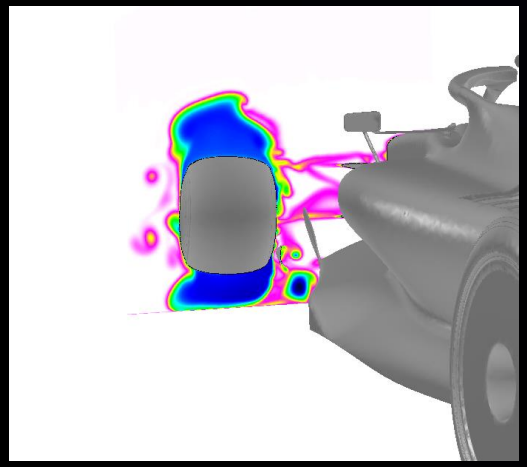
"AWS, the world's leading cloud provider, will help us continuously refine our designs."

Blake Scholl, Founder and CEO, Boom Supersonic

aws

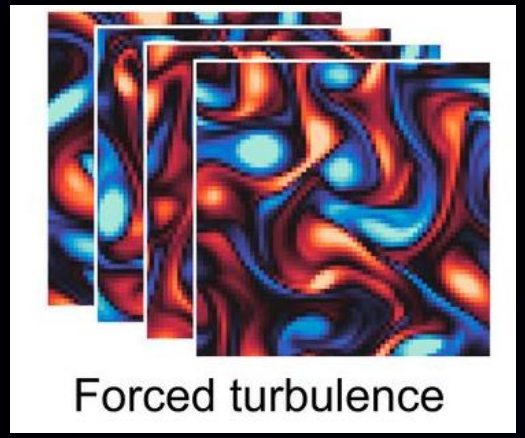
ML + HPC possibilities

Extract relationships from simulation data



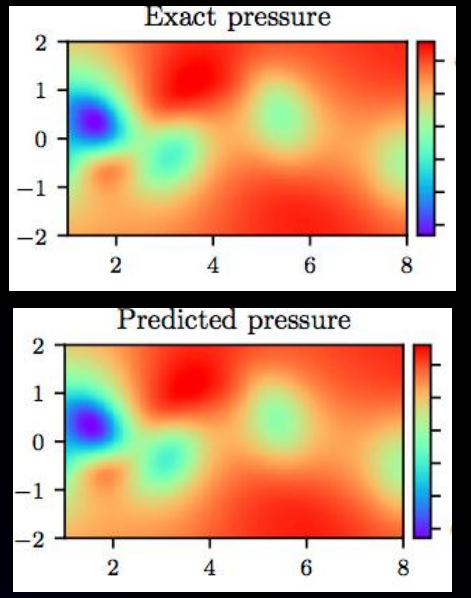
Credit: F1

Accelerate HPC simulations with embedded ML models



Credit: “Machine learning–accelerated computational fluid dynamics,” https://www.pnas.org/content/118/21/e2101784118

Learn solutions to governing equations directly



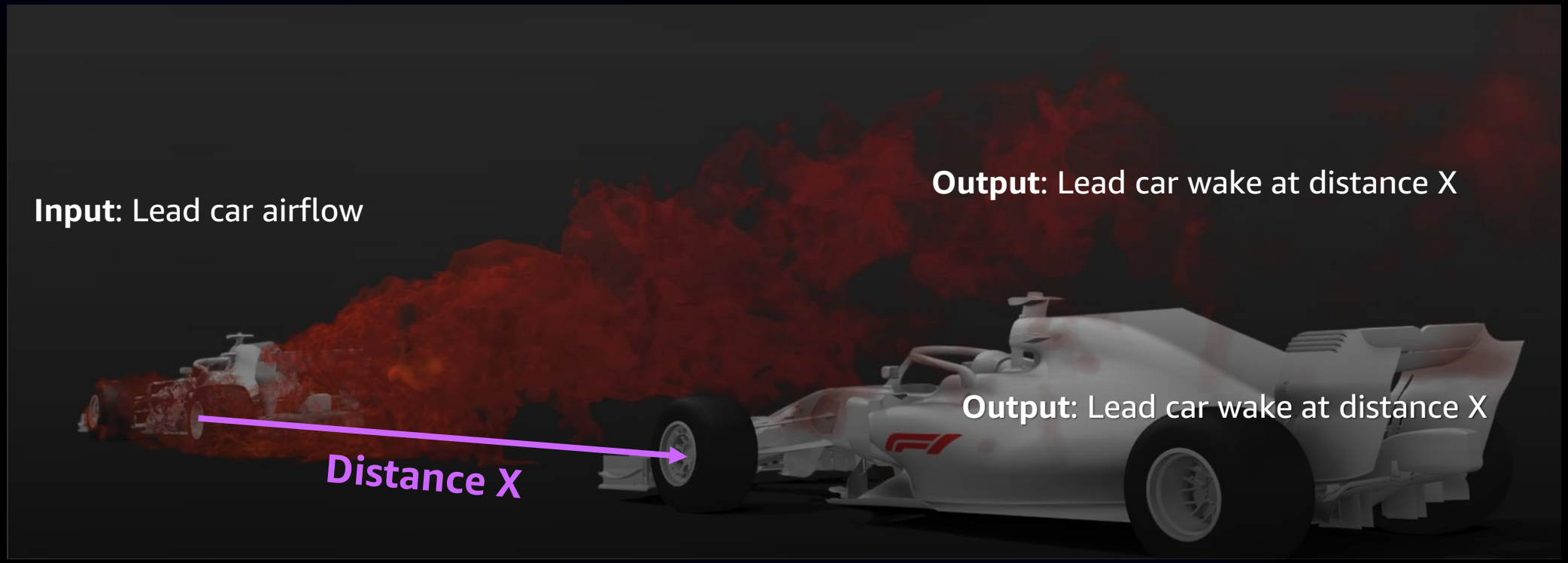
Credit: “Physics Informed Deep Learning,” https://maziarraissi.github.io/PINNs/

AI-driven infrastructure optimization

aws

Demonstrating ML + HPC

EXTRACT RELATIONSHIPS FROM HPC SIMULATION RESULTS



Problem statement:

Determine optimal lead car features to obtain target flow profile at distance X

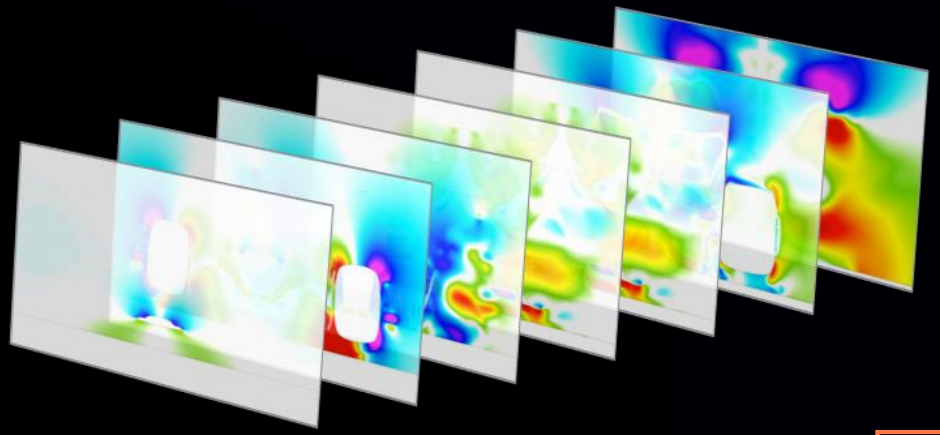
ML + HPC solution

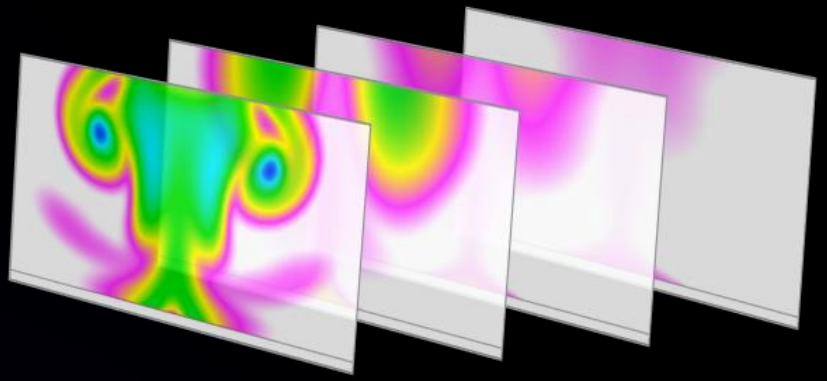
3D geometry

Input flow field

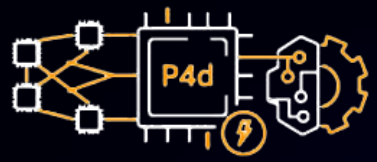
DL model

Output flow field



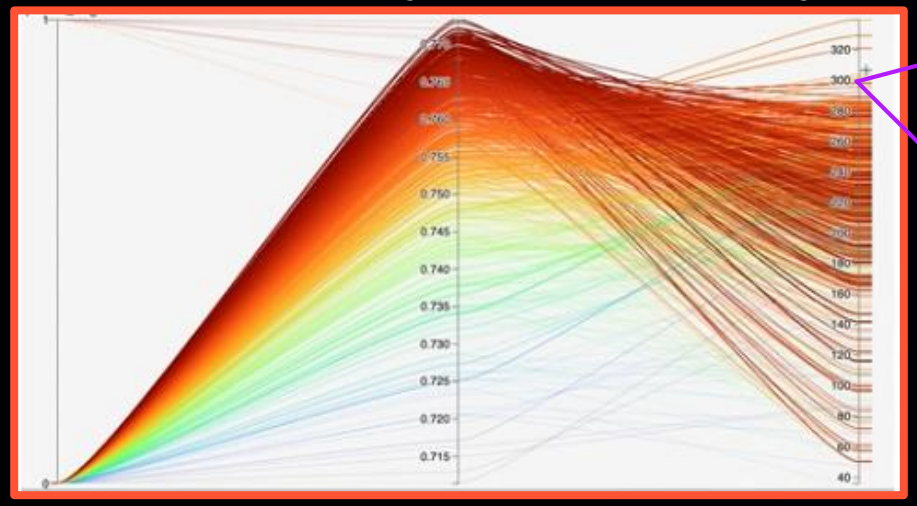


EC2 UltraClusters



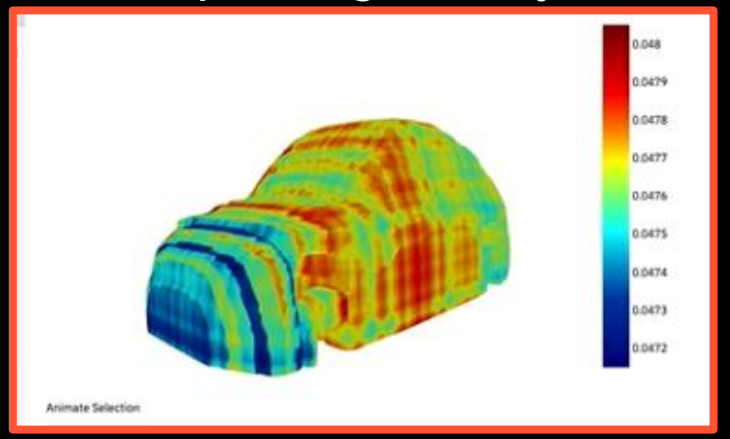
Mapping trajectories to design solutions

Objective 1 Objective 2

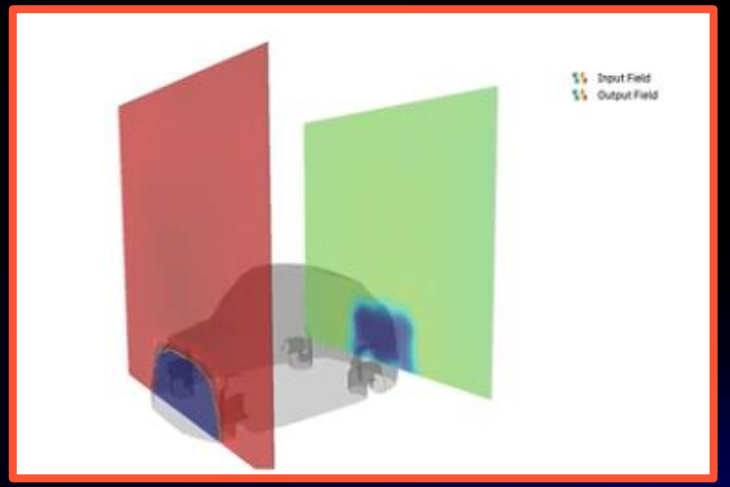


10,000 designs explored

Optimal geometry



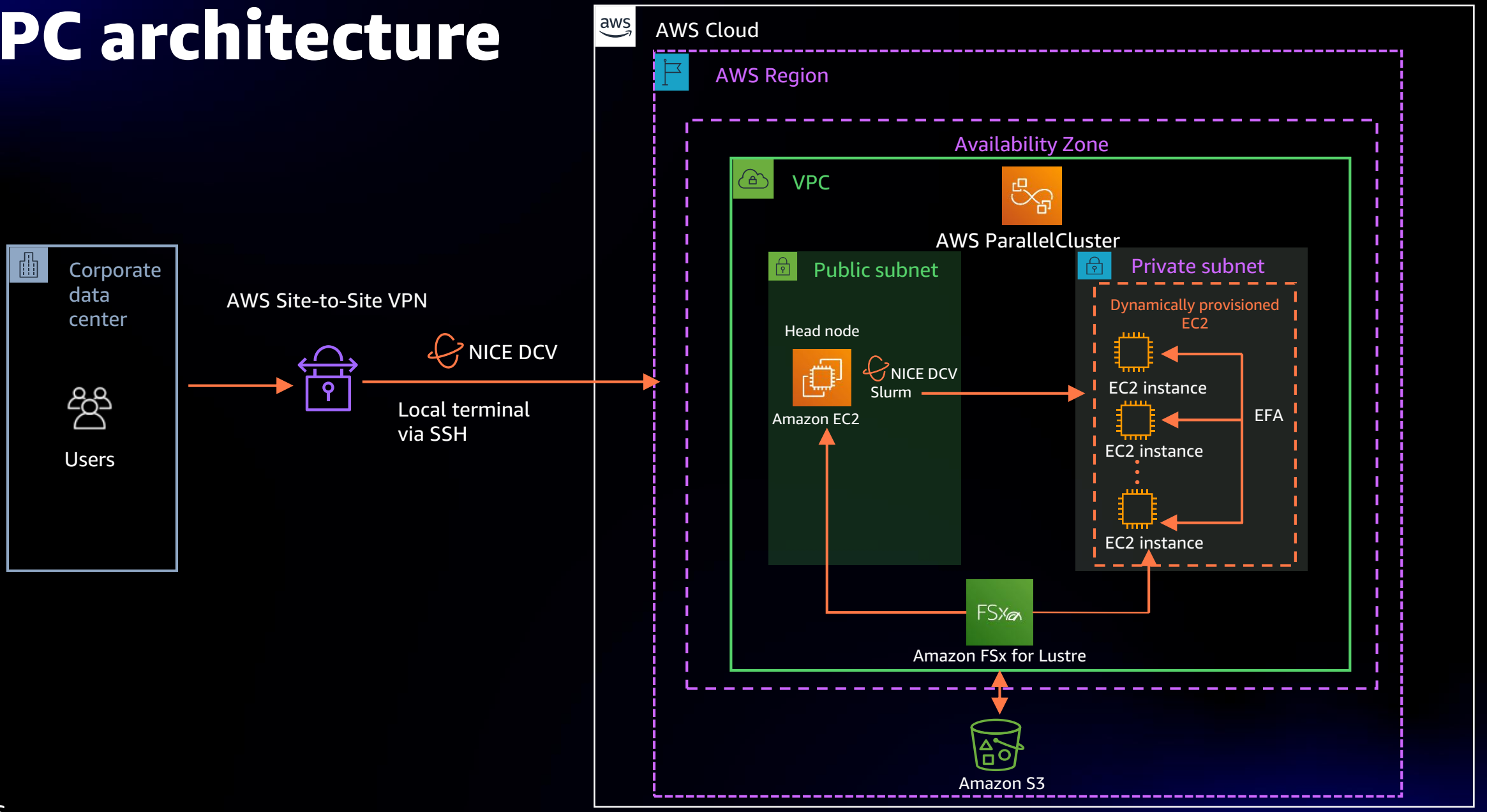
Optimal flow field



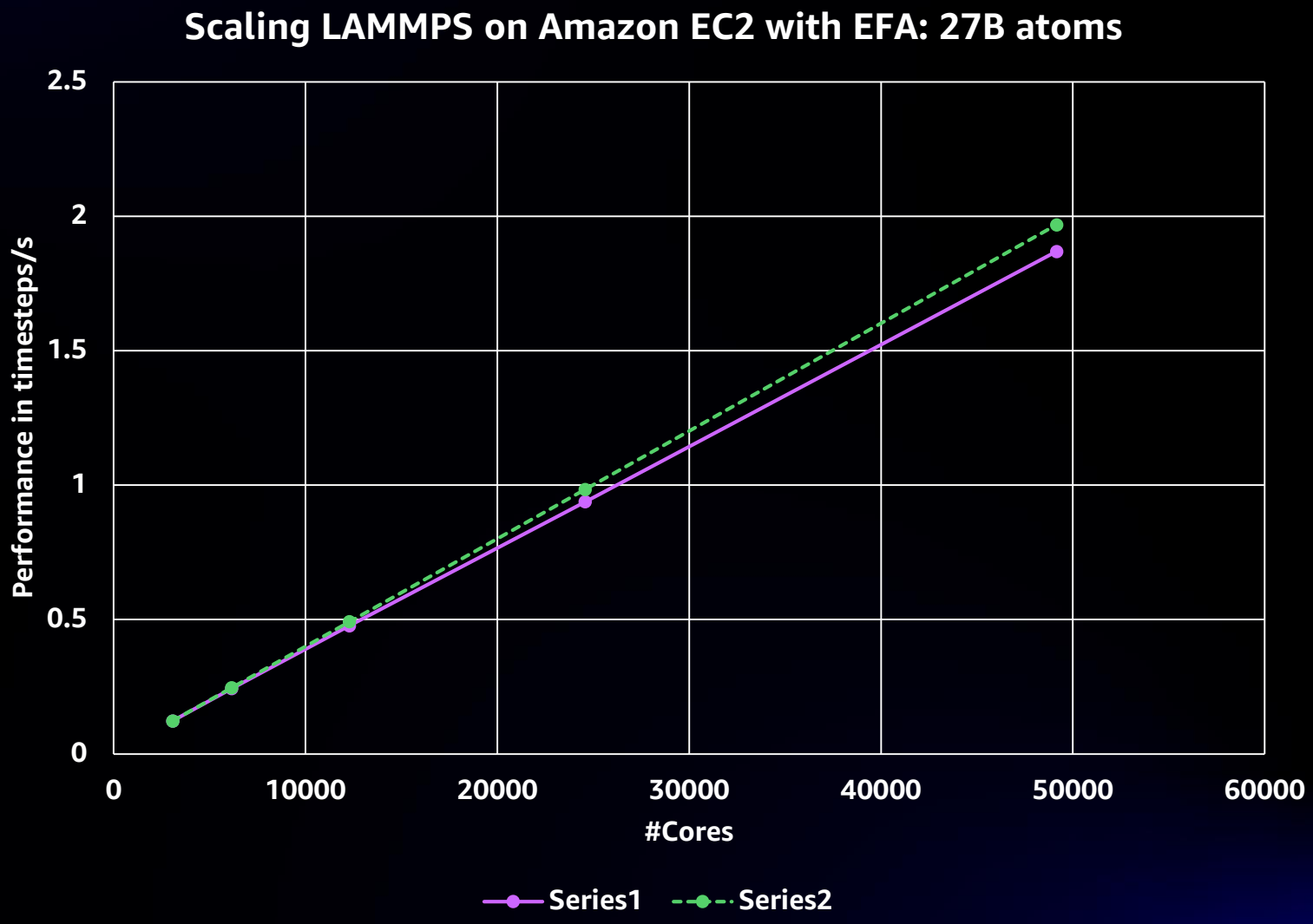
Jyothi Venkatesh

Specialist HPC Solutions Architect, AWS

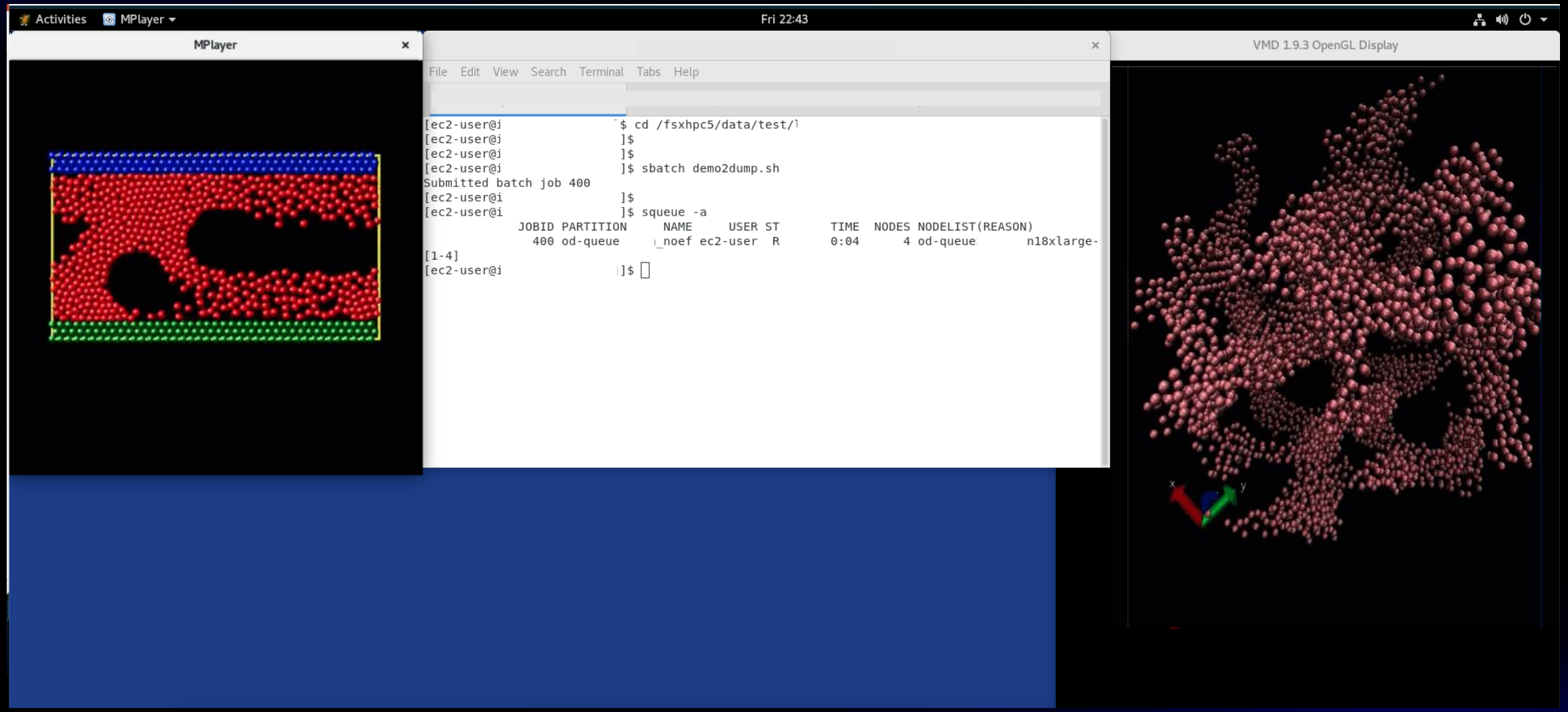
HPC architecture



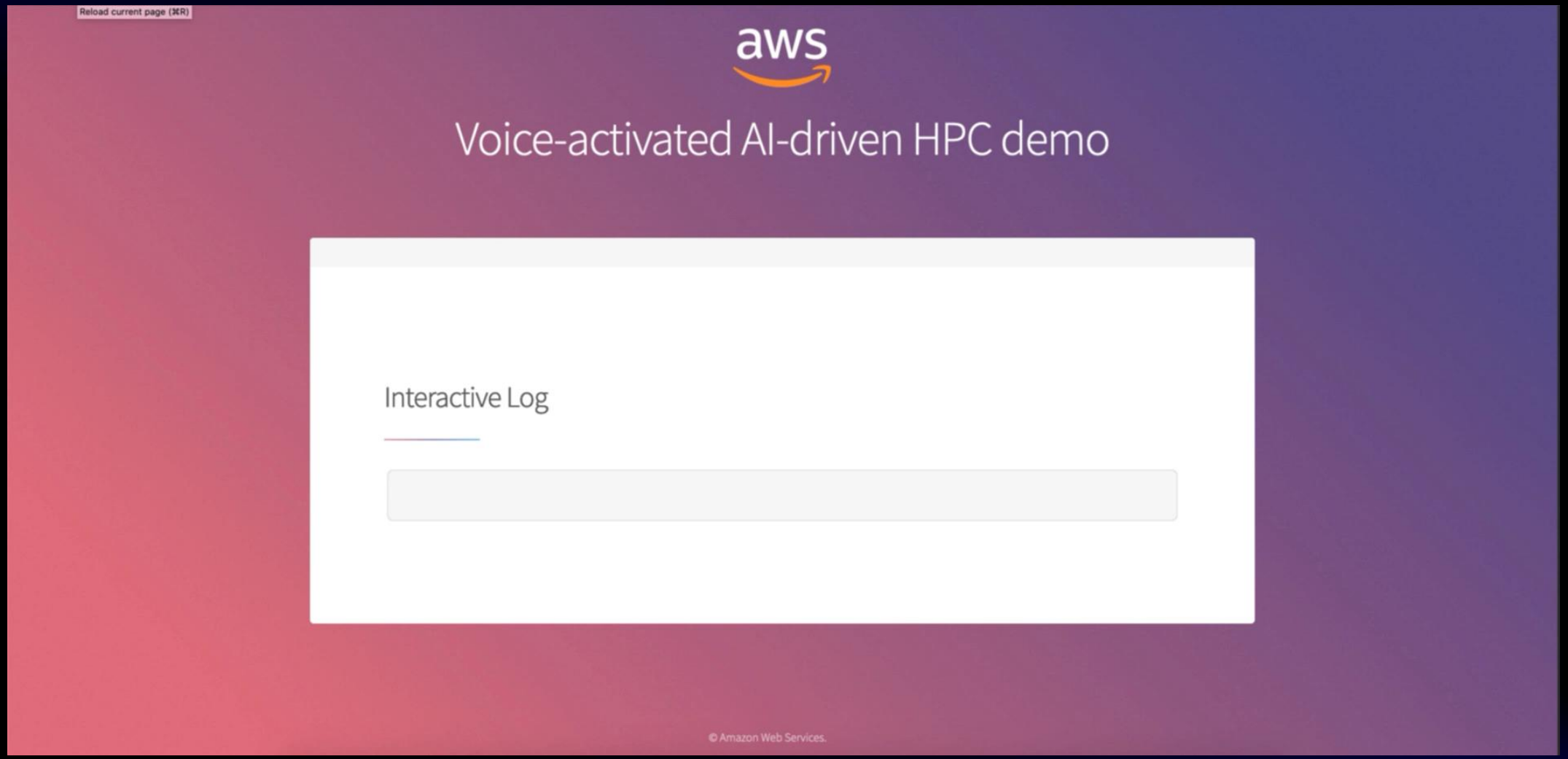
Scaling LAMMPS on AWS



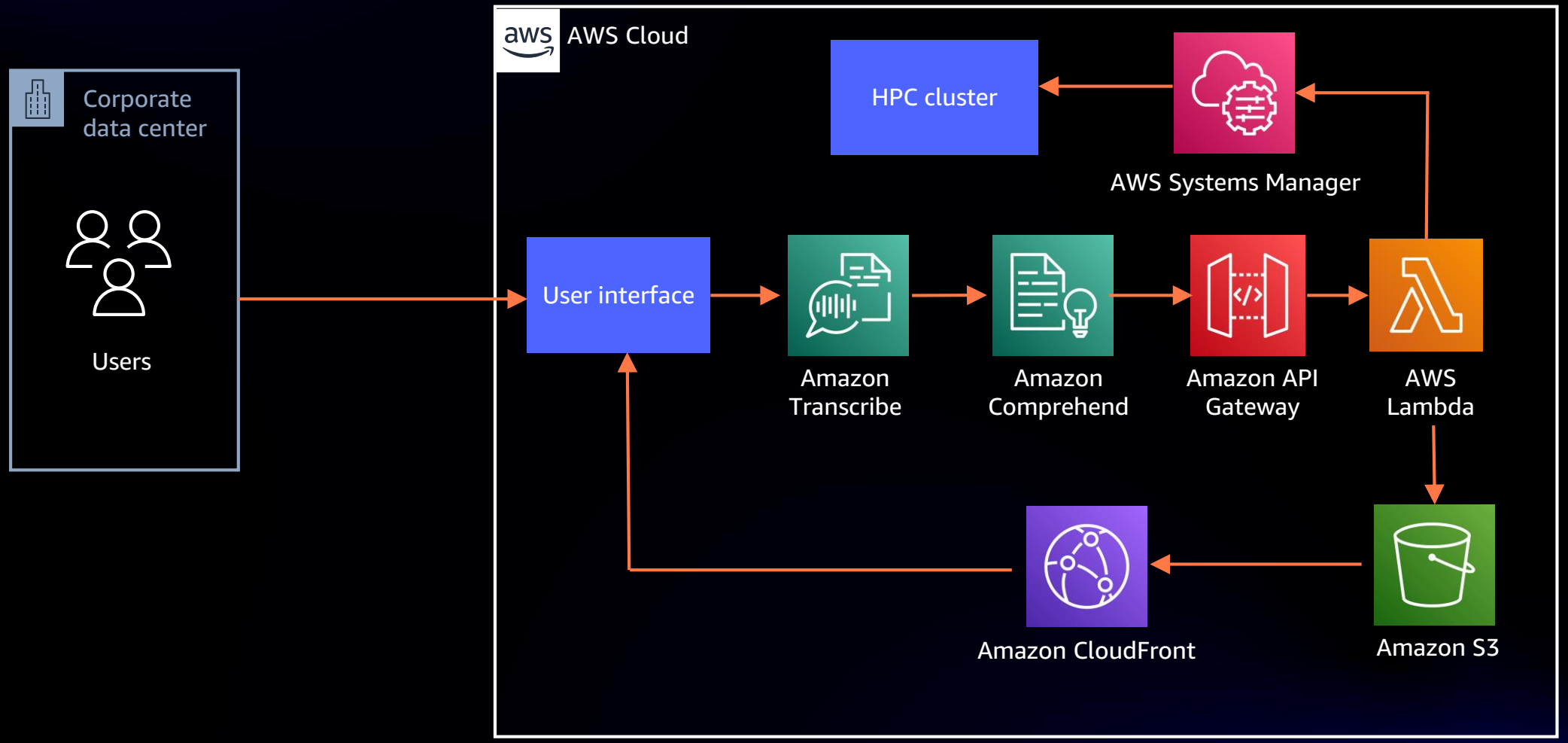
Running HPC applications on AWS



AI-driven HPC: The future is here!



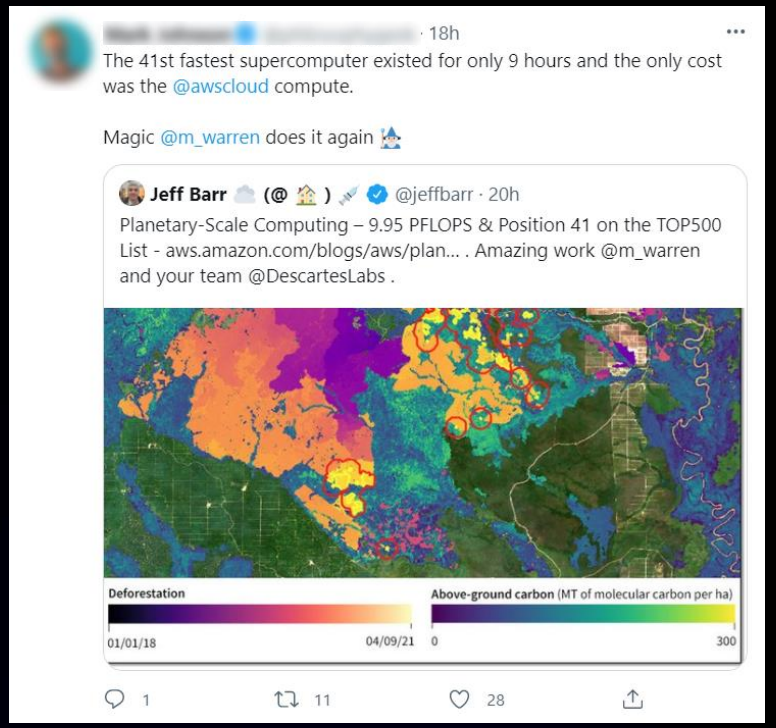
Voice-activated, AI-driven HPC demo



AWS and Descartes Labs: Planetary-scale modeling for sustainability

Descartes Labs

aws



Jeff Barr blog: https://aws.amazon.com/blogs/aws/planetary-scale-computing-9-95-pflops-position-41-on-the-top500-list/

Note: Descartes Labs moved up from #41 to #40 after official rankings were announced due to the removal of the #33 entry

aws

THE FUTURE OF HPC IS LOOKING A LOT LIKE ML

Going beyond net zero commitments

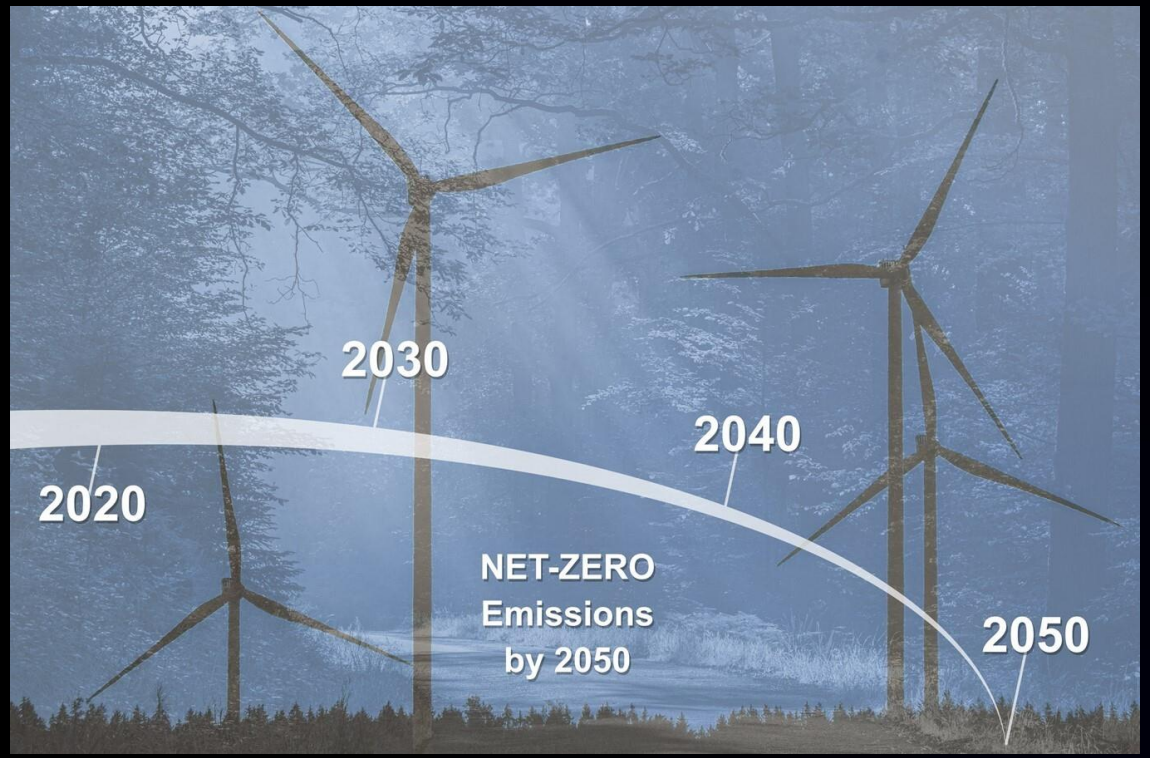
Terry Cunningham (he/him)
Descartes Labs
CEO

Academia
Agriculture
Aviation & aerospace
Chemicals
Consumer goods
Defense & intelligence
Engineering & construction
Financial services
Forest products
Manufacturing
Maritime
Mining & metals
Oil & gas
Public sector & NGOs
Real estate
Retail
Technology
Telecom
Transportation
Utilities

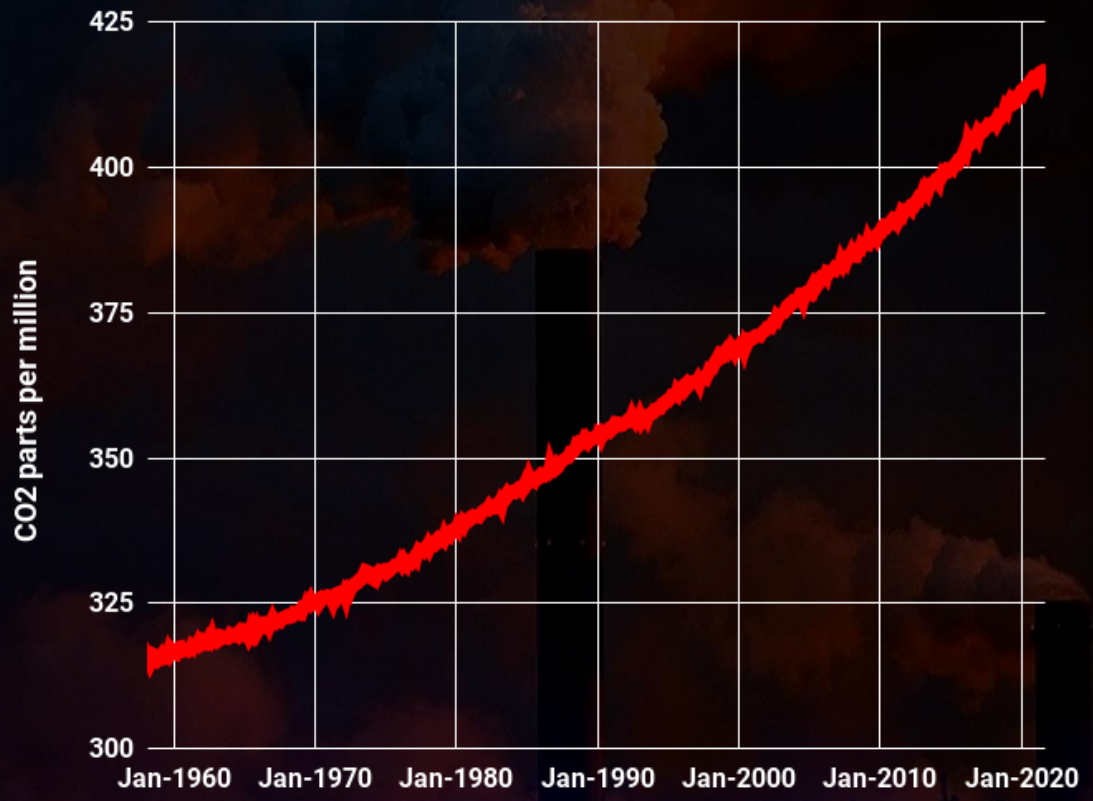
Net zero commitments

HOW IT STARTED

Hundreds of companies pledge to achieve net zero emissions



HOW IT'S GOING



For good, measure!

THERE IS NO 1.5°C WORLD WITHOUT TRANSFORMING THE LAND-USE ECONOMY

25%

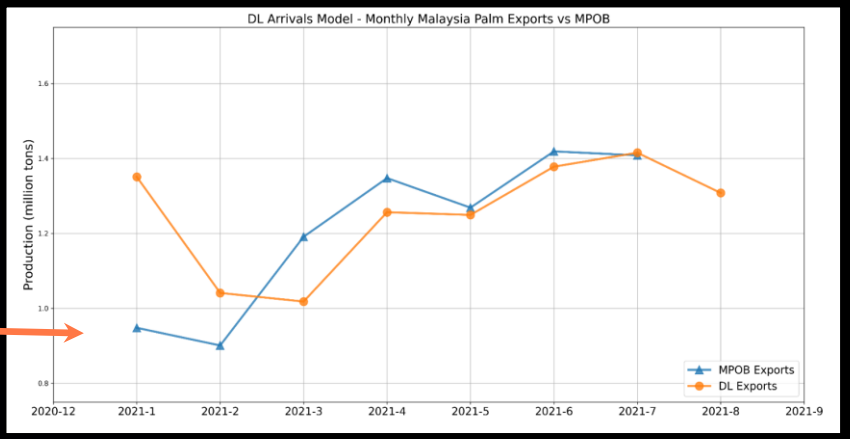
of global emissions are from land use

For example, palm oil


How do we pay
for sustainability?

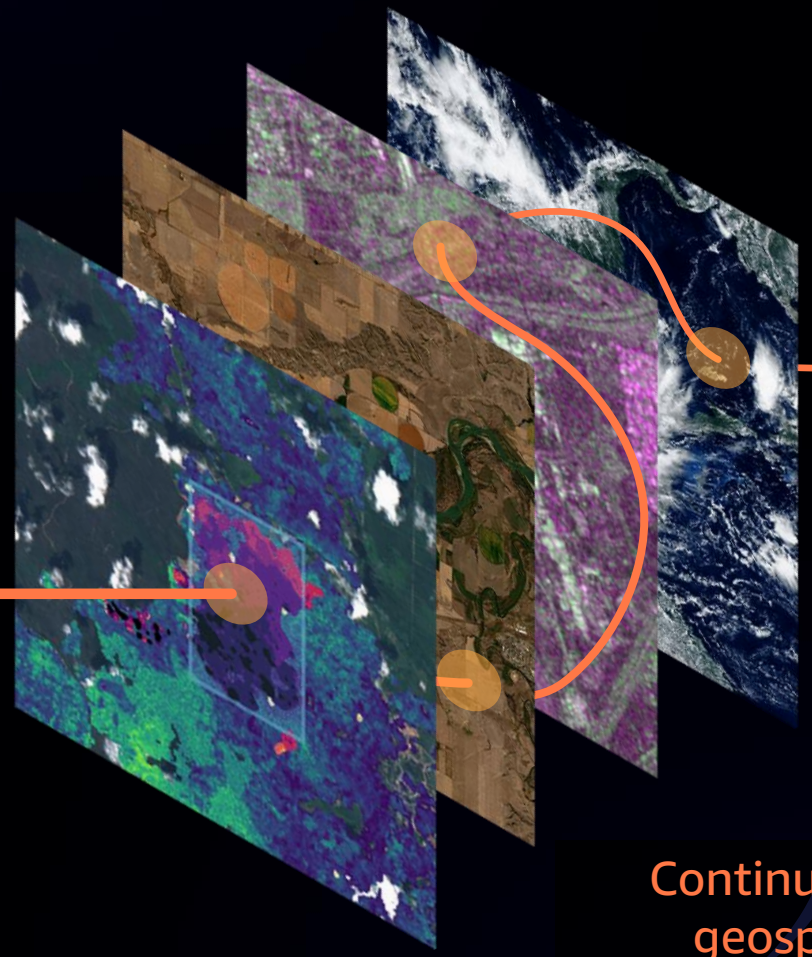
For profit, forecast!

Smart procurement with improved price forecasts



HPC with AWS

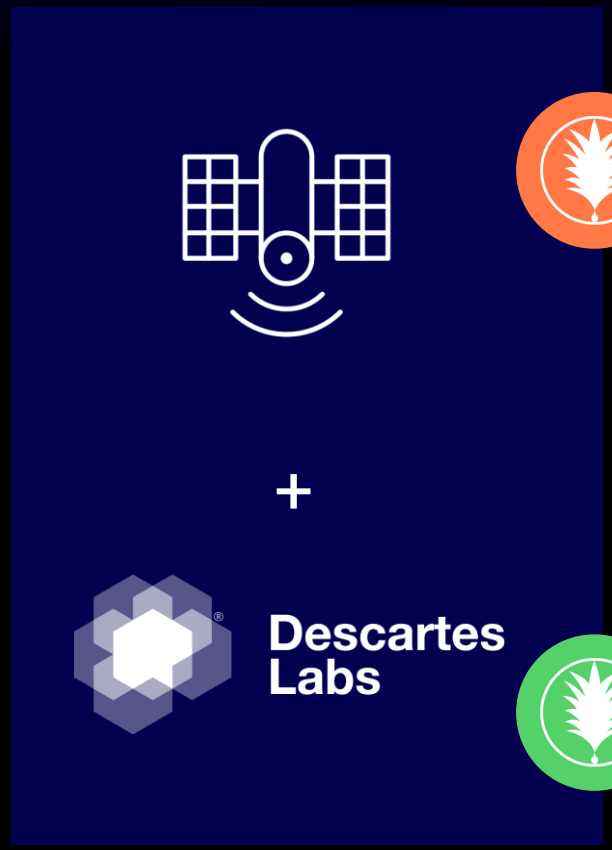
powered by aws



Continuously updating geospatial insights

aws

Price forecasts can help fund sustainability initiatives



+

Descartes Labs

FORECASTING

4% savings

using a more accurate palm price forecast

Funds

SUSTAINABILITY

Deforestation alerts

of particular salience and importance

aws

AWS News Blog

Planetary-Scale Computing – 9.95 PFLOPS & Position 40 on the TOP500 List

by Jeff Barr | on 28 JUN 2021 | in Amazon EC2, High Performance Computing, Launch, News | Permalink | Comments | Share

TOP500 The List. CERTIFICATE

Amazon EC2 Instance Cluster us-east-1a - Amazon EC2 r5.24xlarge, Xeon Platinum 8260 24C 2.4GHz, 25G Ethernet

Descartes Labs, United States

is ranked

No. 40

among the World's TOP500 Supercomputers

with 9.95 PFlop/s Linpack Performance

in the 57th TOP500 List published at the ISC Virtual 2021

Conference on June 28, 2021.

Congratulations from the TOP500 Editors

Erich Strohmaier
NERSC/Berkeley Lab

Jack Dongarra
University of Tennessee

Horst Simon
NERSC/Berkeley Lab

Martin Meuer
Prometeus

This is only possible with AWS

Storage

Amazon Simple Storage Service (Amazon S3)

Amazon S3 Storage Lens

Compute

Amazon Elastic Compute Cloud (Amazon EC2)

AWS Lambda

AWS Batch

Database

Amazon Relational Database Service (Amazon RDS)

Analytics

Amazon Redshift

AWS Glue

Containers

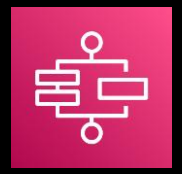
Amazon Elastic Kubernetes Service (Amazon EKS)

Amazon Elastic Container Service (Amazon ECS)

AWS Fargate

Integration

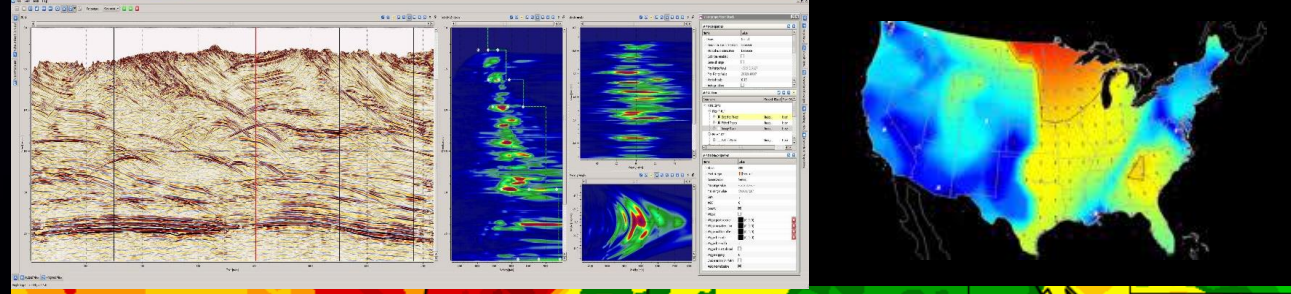
Amazon Simple Queue Service (Amazon SQS)

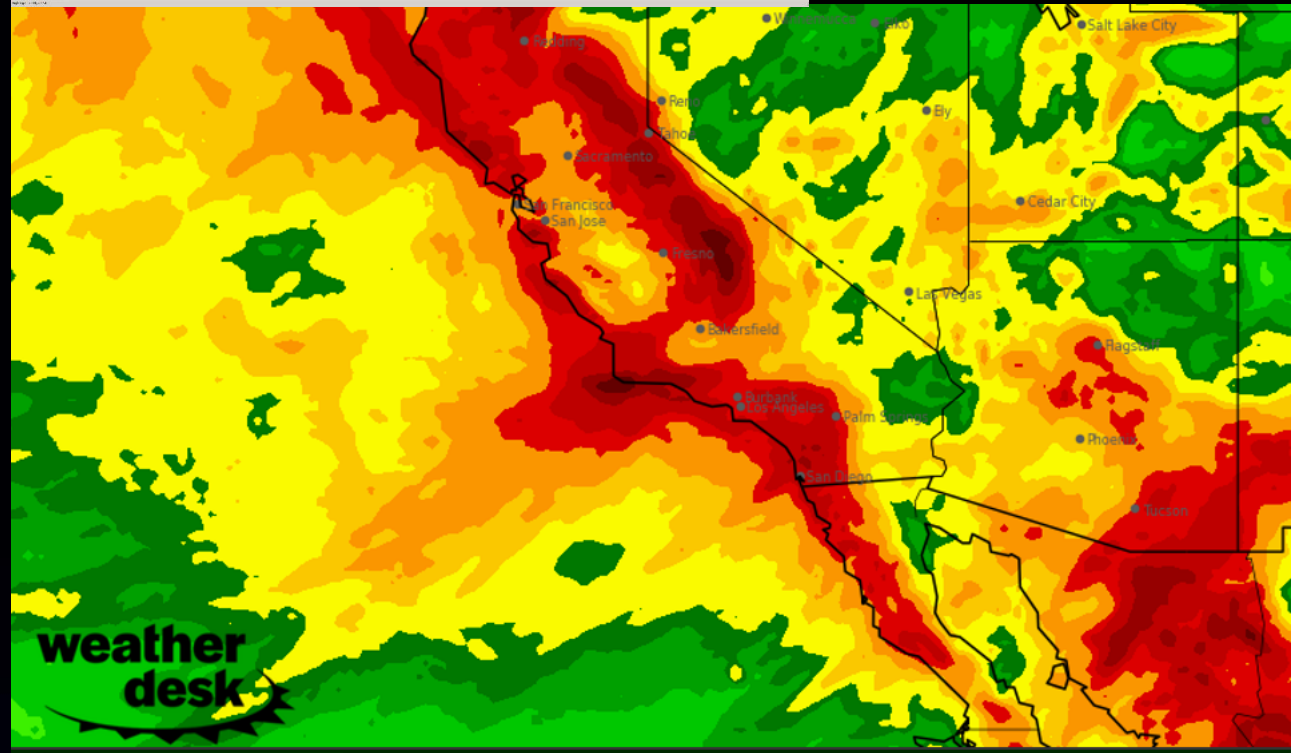


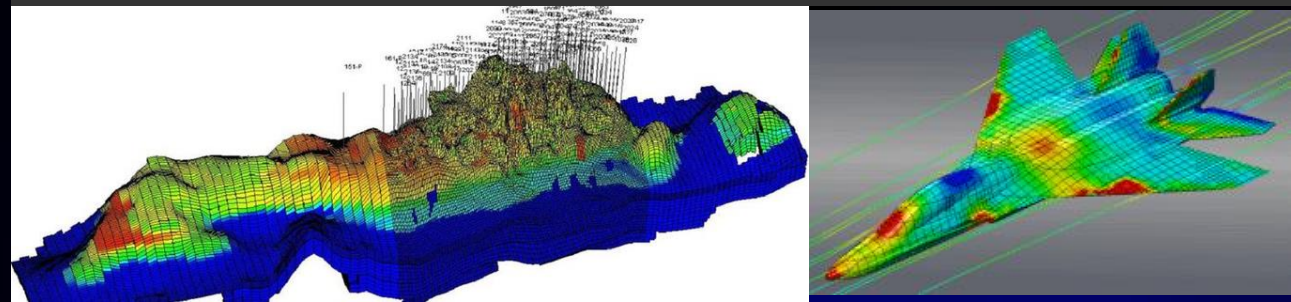
AWS Step Functions

HPC + ML on AWS enables customers to find new business value

Rethink . . .
Reinvent . . .
Results . . .







Thank you!