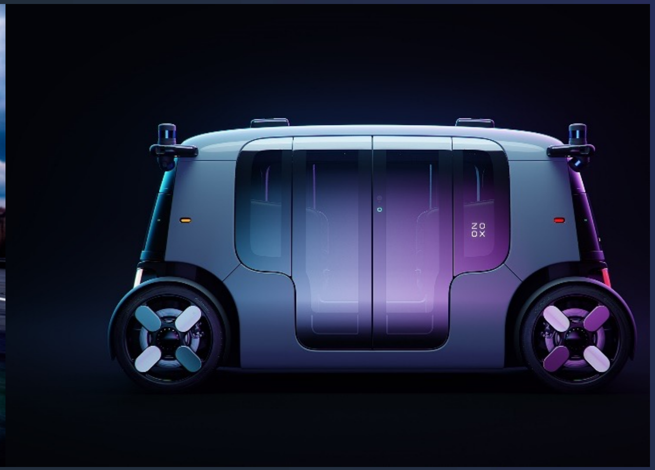
Cloud Computing for Computational Fluid Dynamics

Dr Neil Ashton
Principal, Computational Engineering Solution Architect, AWS

How do you design the next generation of X?

Companies need a way to assess the **performance** and **efficiency** of new designs

They want to go from conceptual design to full design in the **shortest** possible time with the **least** expense

Digital certification is the dream for many industries; **let's explore that further**

Option 1 – Physical tests

- Build it
- Expensive
- Accurate

Option 2 – Wind tunnel tests

Mock up

Expensive

Limited

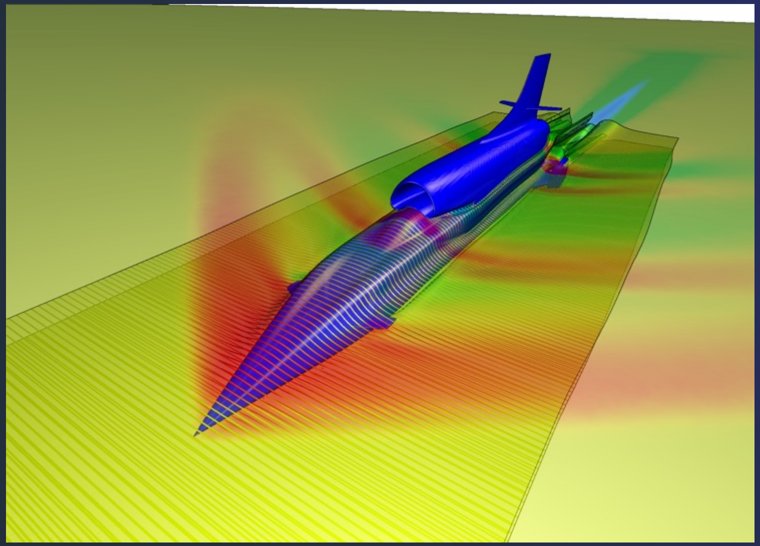
Credit: Andareed at English Wikipedia

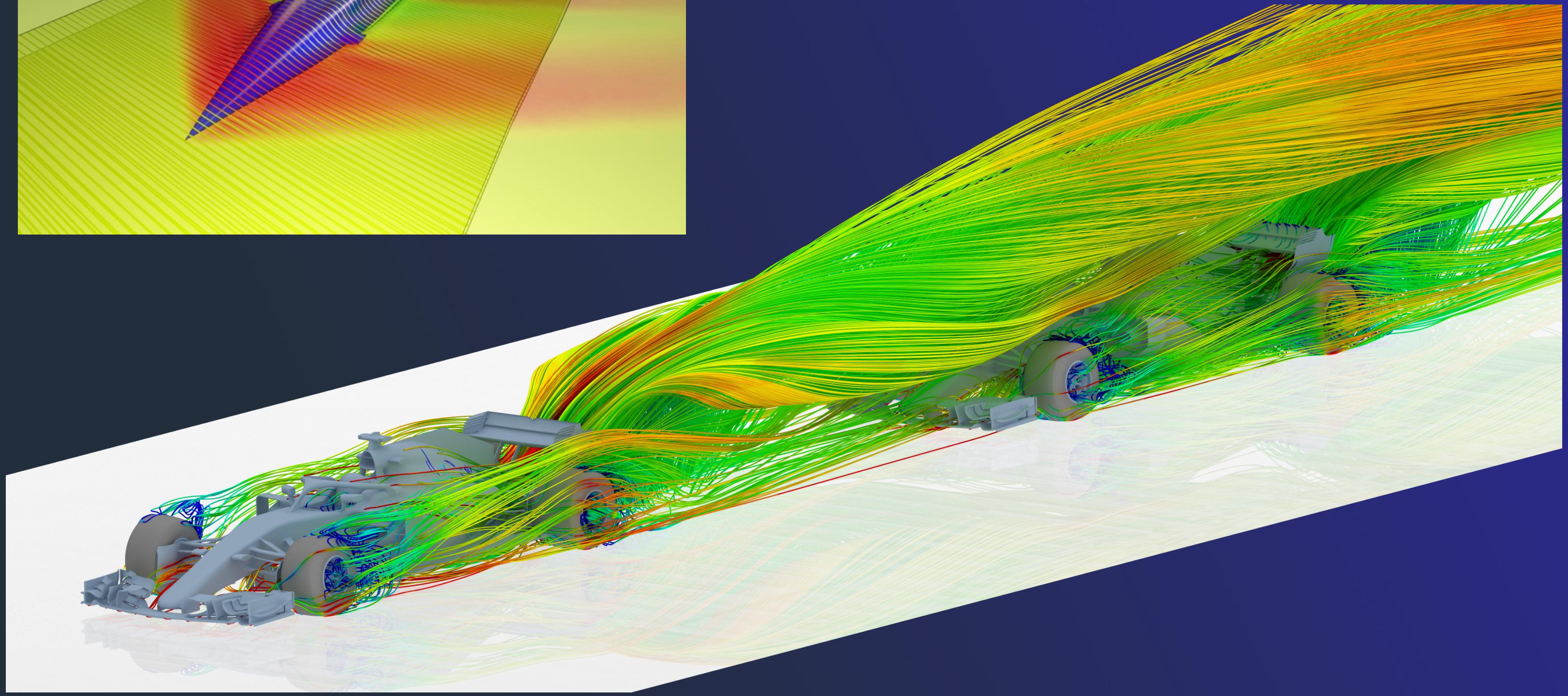
Option 3 – Simulation (computational fluid dynamics)

No Manufacture

Lower cost

Accurate?





Option 3 – Simulation (computational fluid dynamics)

No Manufacture

Lower cost

Accurate?

$$\frac{\partial u_i}{\partial x_i} = 0$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j}$$

Levels of CFD fidelity

DNS

Highest Fidelity - Direct numerical simulation

- Best representation of all the physics; no modeling required
- Extremely expensive; currently just a research tool

N_C=10B

$\Delta t = 1e^{-7}$

N_T=40M (4s)

N_p=200K (92 days)*

~$13M**

** 0.2 s per time-step (50K cells per core)

** $0.03 per core hour

Levels of CFD fidelity

LES

High-Medium Fidelity - Large eddy simulation

- Resolve large scale but model small scale
- Remains expensive; mesh generation

N_C=2B

$\Delta t = 1e^{-6}$

N_T=4M (4s)

N_p=40K (12 days)*

~$350K**

** 0.25 s per time-step
** $0.03 per core hour

Levels of CFD fidelity

RANS

Low-Fidelity - Reynolds-averaged Navier-Stokes

- Required extensive mathematical modeling
- Cheap/fast for engineering applications

N_C=250M

Δt=n/a

N_T=4k

N_p=2.5K (2 hours)*

~$150**

** 2 s per iteration (100K cells per core)
** $0.03 per core hour

Levels of CFD fidelity

Hybrid

Medium/High - Hybrid RANS-LES/WMLES

- Low-fidelity in near-wall regions, LES further away
- "Compromise" for industry

N_C=1B

$\Delta t=1e^{-5}$

N_T=400k

N_p=20K (108 hours)*

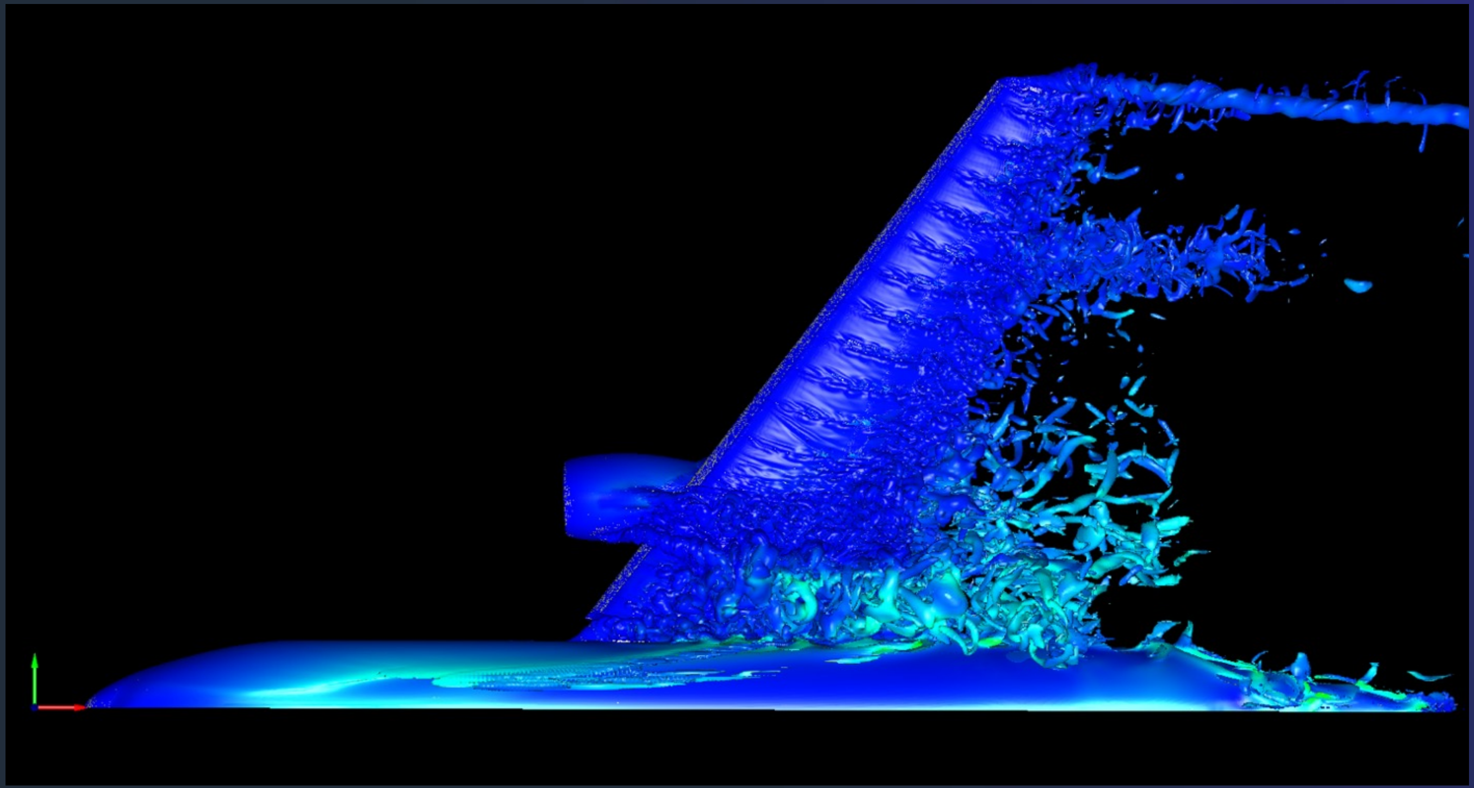
~$64K**

** 1 s per time-step (50K cells per core) – implicit
** $0.03 per core hour

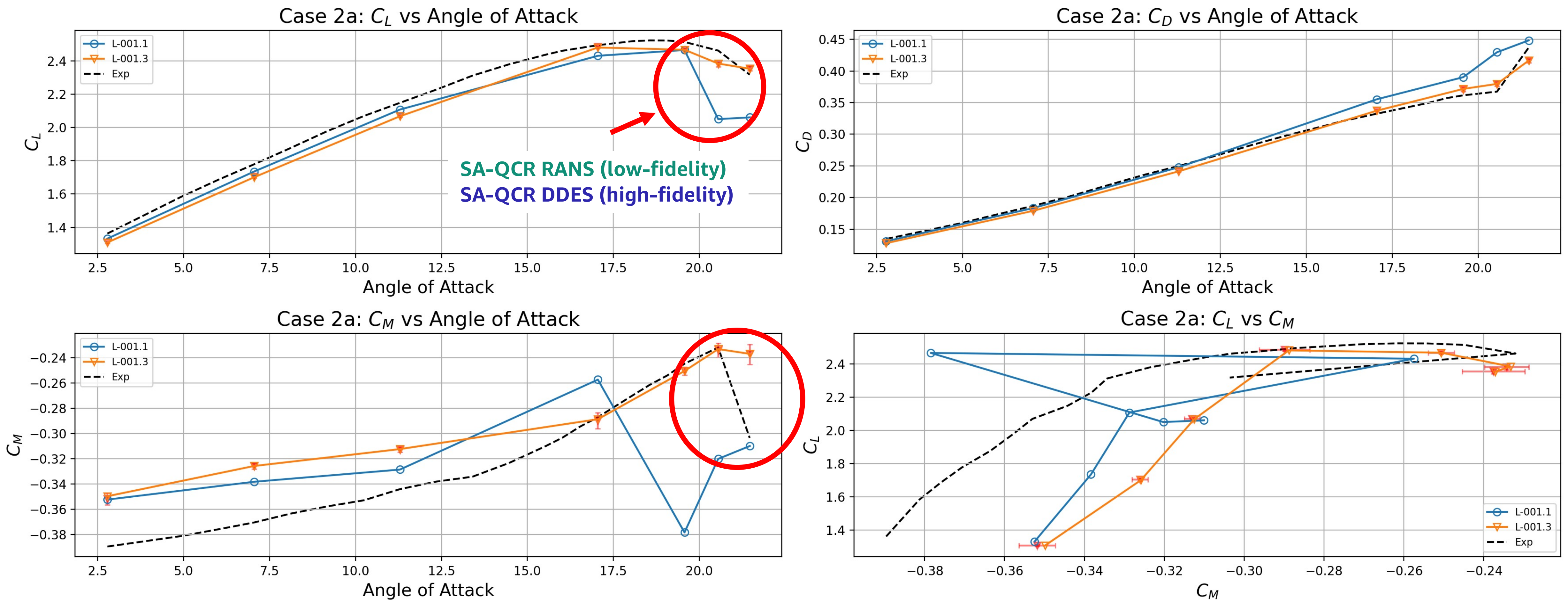
HLPW4 workshop

4th High Lift Prediction Workshop

Advance the state of the art for CFD with Boeing/NASA and many others

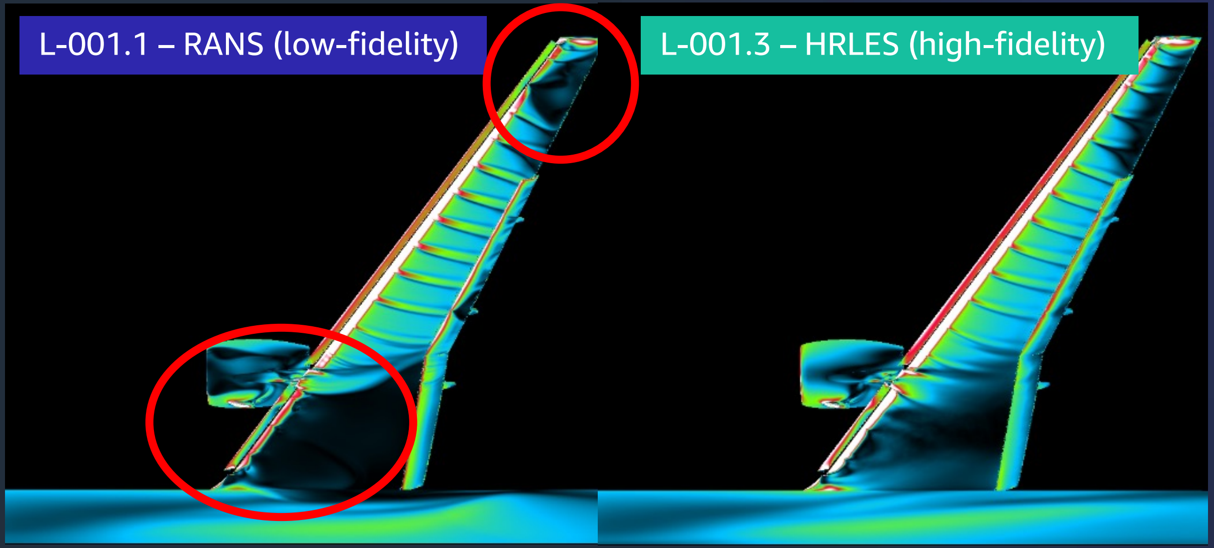


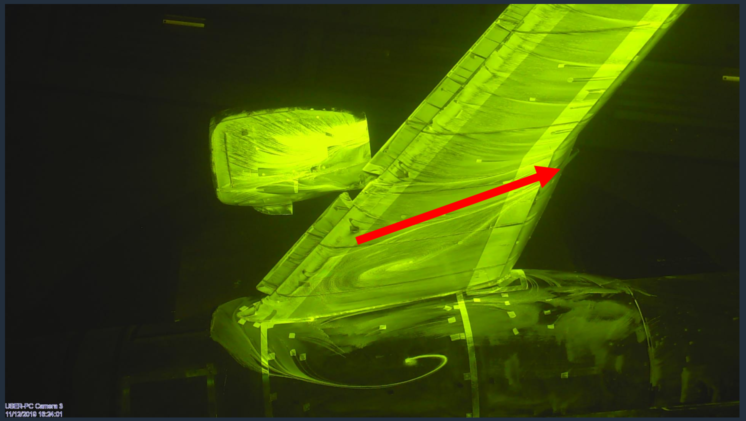
Why High-Fidelity CFD?

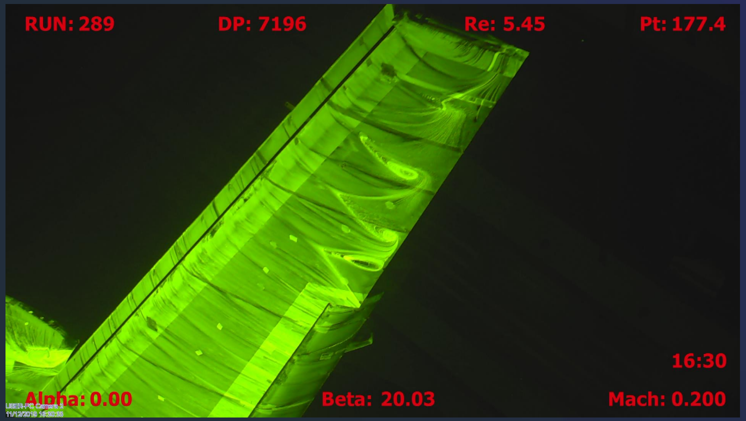


Supporting Evidence

21.47 AoA







On-prem HPC pain points

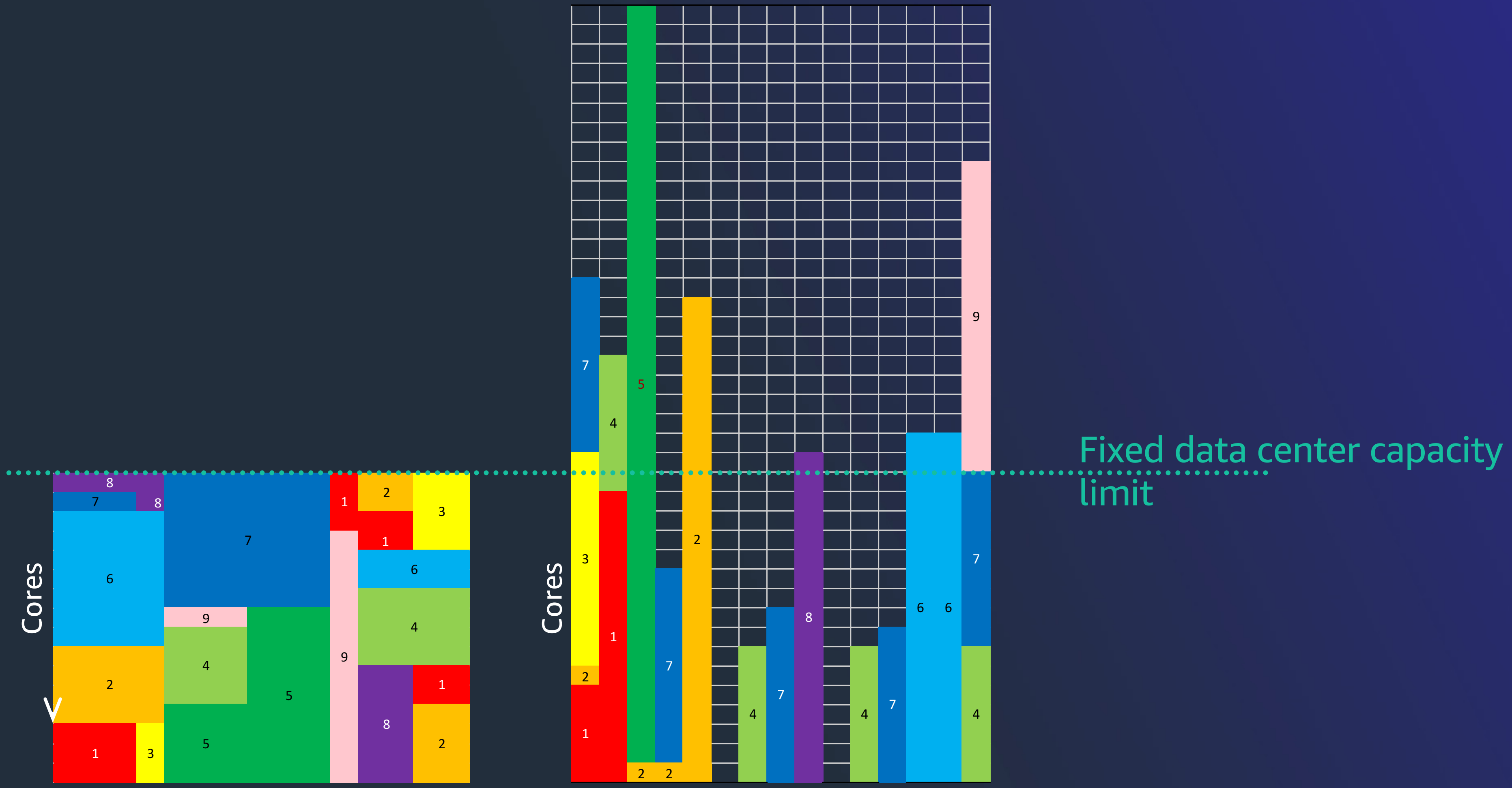
I’m stuck in the queue → **Can you bump up my priority?**

Why is UserX always hogging the queue?

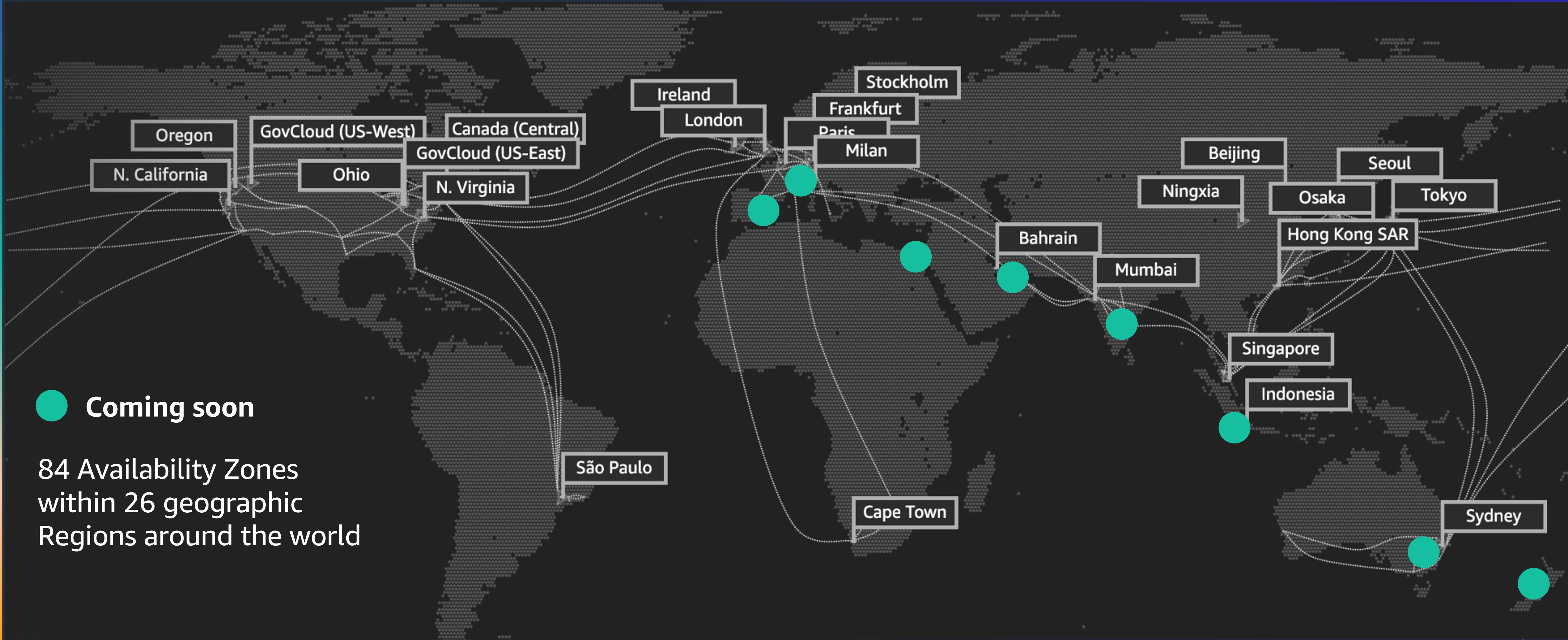
I’ve run out of disk space → **Can I have 1 TB more?**

I promise to delete some files

Cloud offers greater flexibility



Global infrastructure

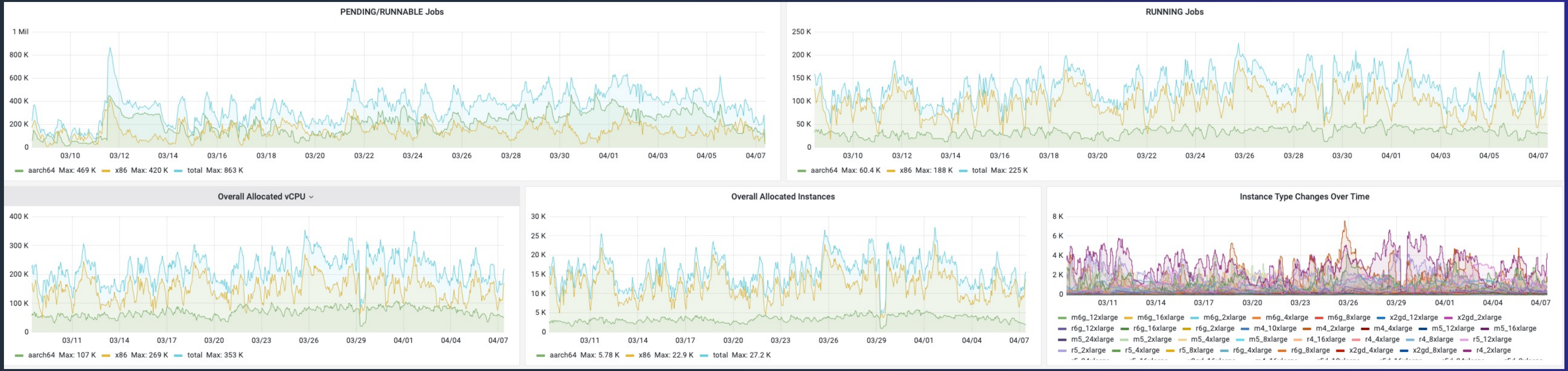


84 Availability Zones within 26 geographic Regions around the world

Arm HPC use

Max # Jobs per Week	Max # Jobs per Day	Average Job Runtime	Max # vCPUs	Max # Instances
63.6M	10.9M	33 Mins	353K*	27.2K*

* all running on spot instances



https://aws.amazon.com/solutions/case-studies/arm-limited-case-study/

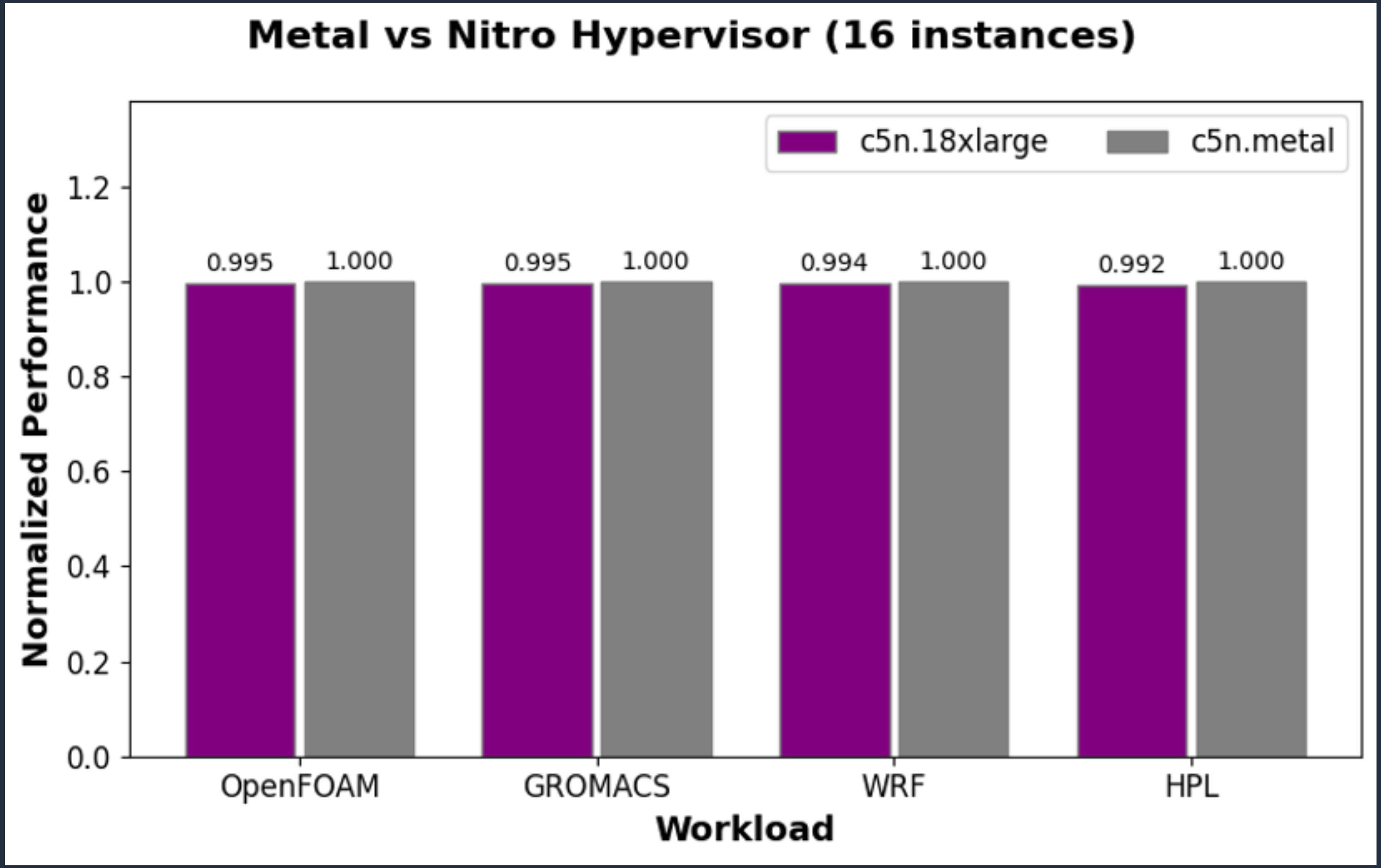
Security/ITAR : AWS GovCloud (US)

Isolated AWS infrastructure and services for customers with strict regulatory and compliance requirements and sensitive data. In addition AWS Secret Regions for Classified, Secret, Top Secret.

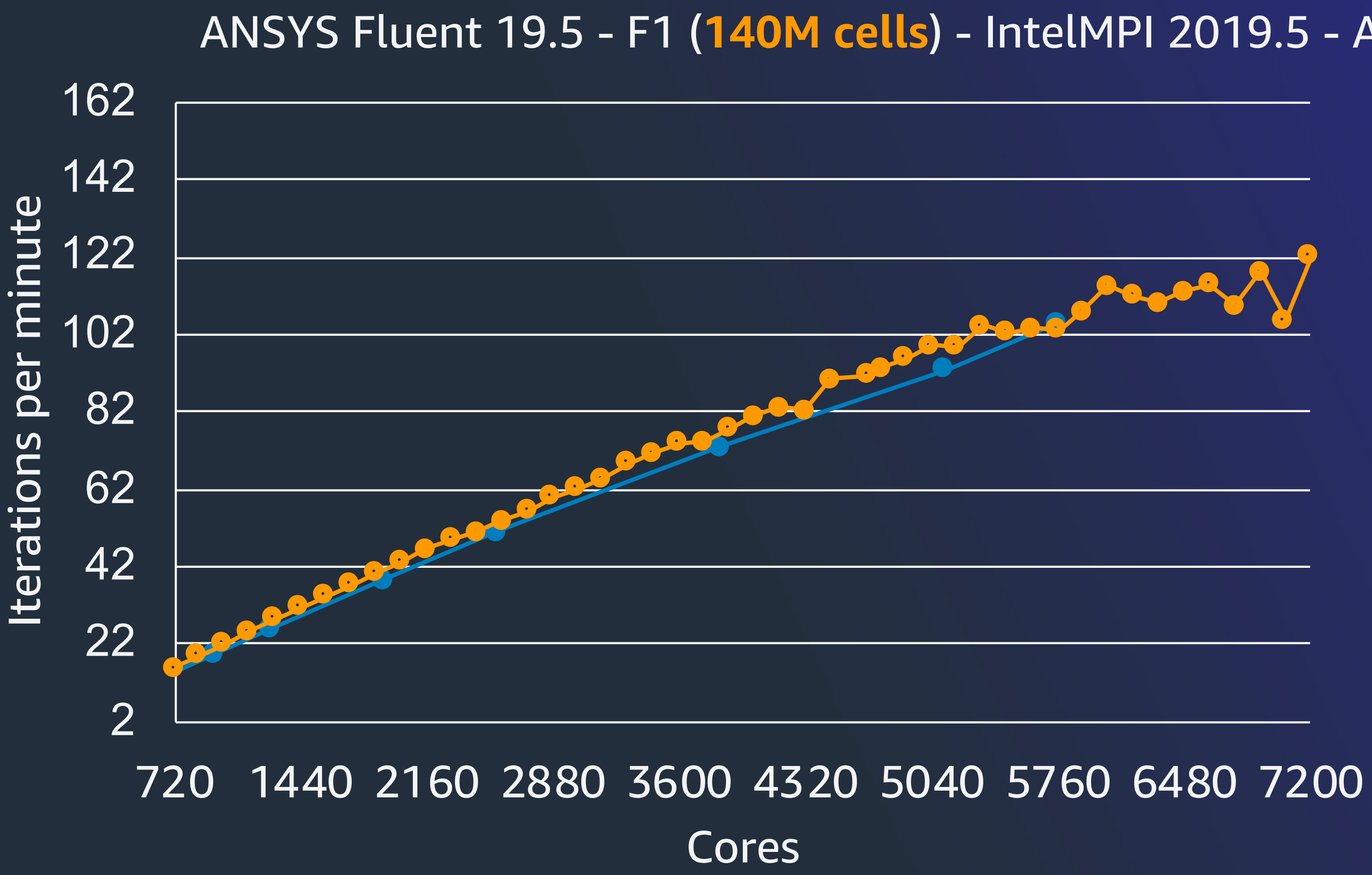
Widely used NASA/US Gov codes (i.e OVERFLOW + FUN3D) running on tens of thousands of cores every day in AWS GovCloud

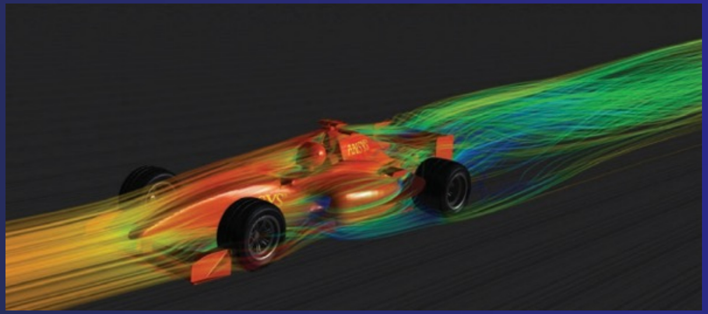
Addresses the most stringent US Government regulations, policies and security requirements

AWS Hypervisor (Nitro) gives bare-metal performance

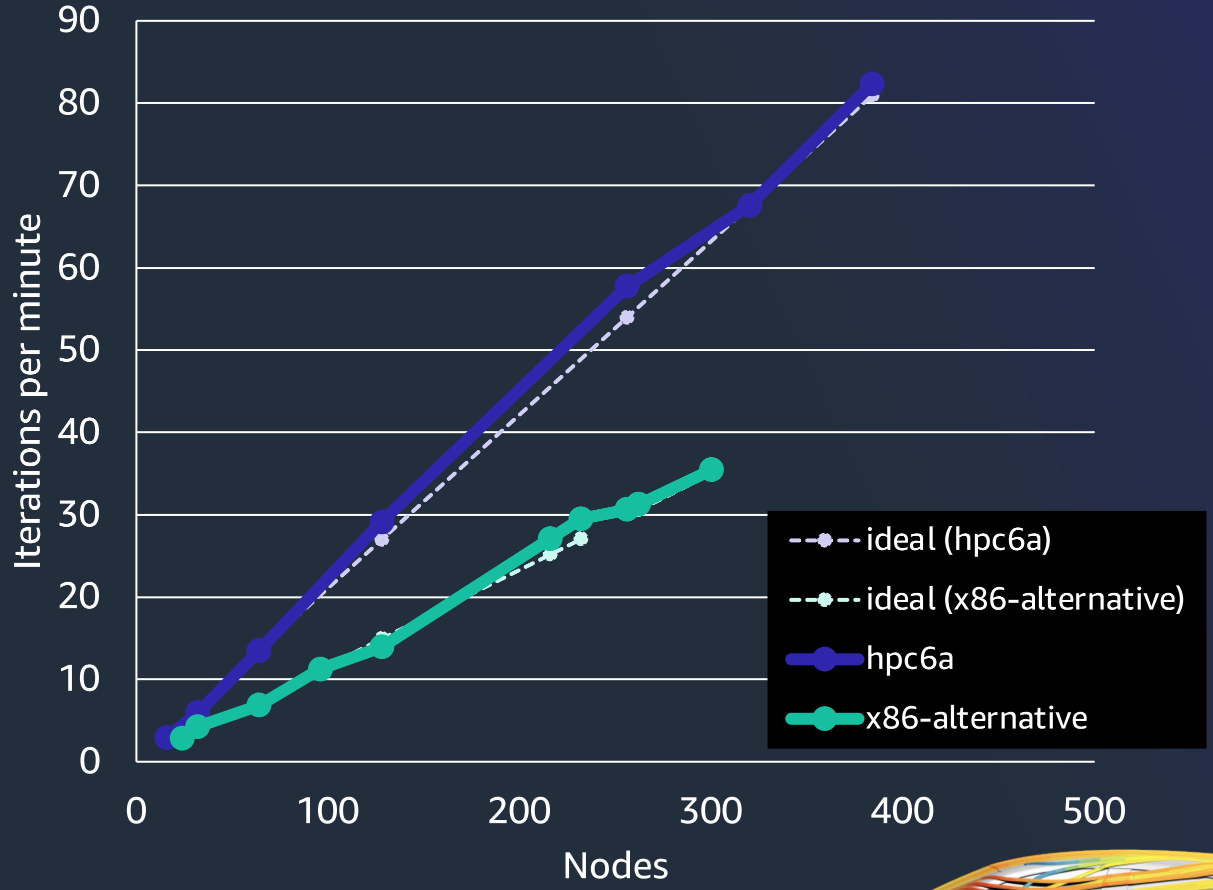


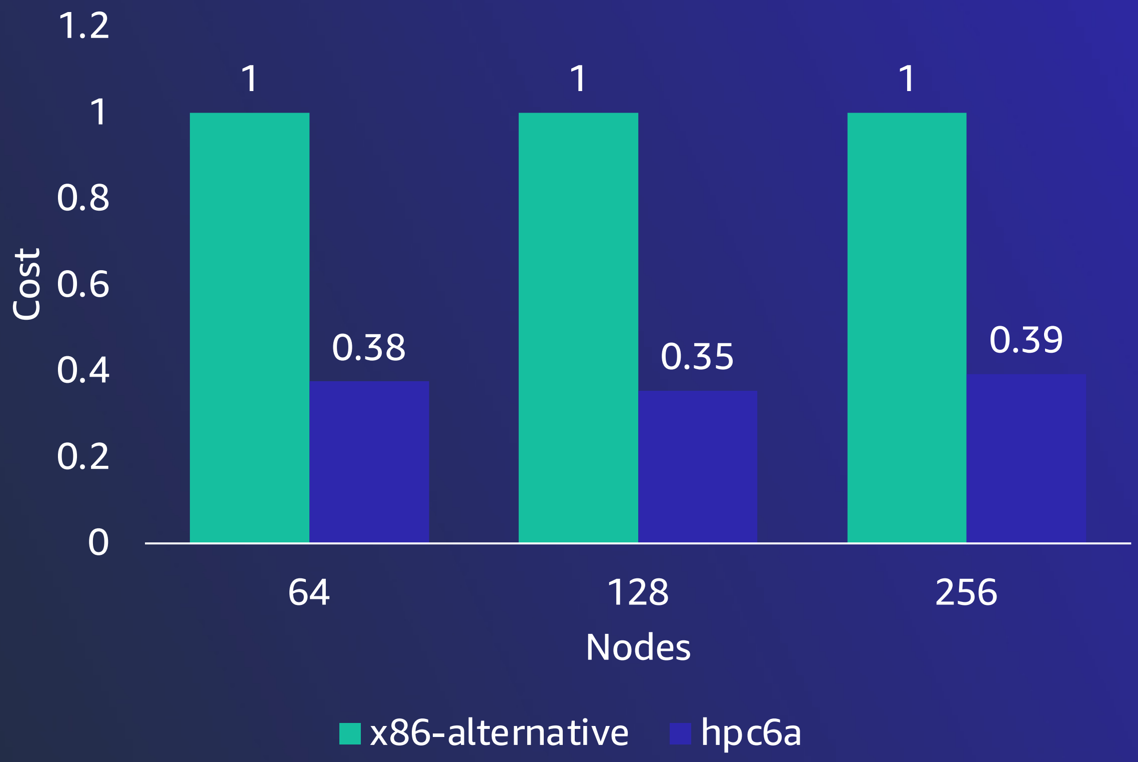
Scaling on AWS – Fluent





HPC6a – 1.4B cell Siemens Simcenter STAR-CCM+

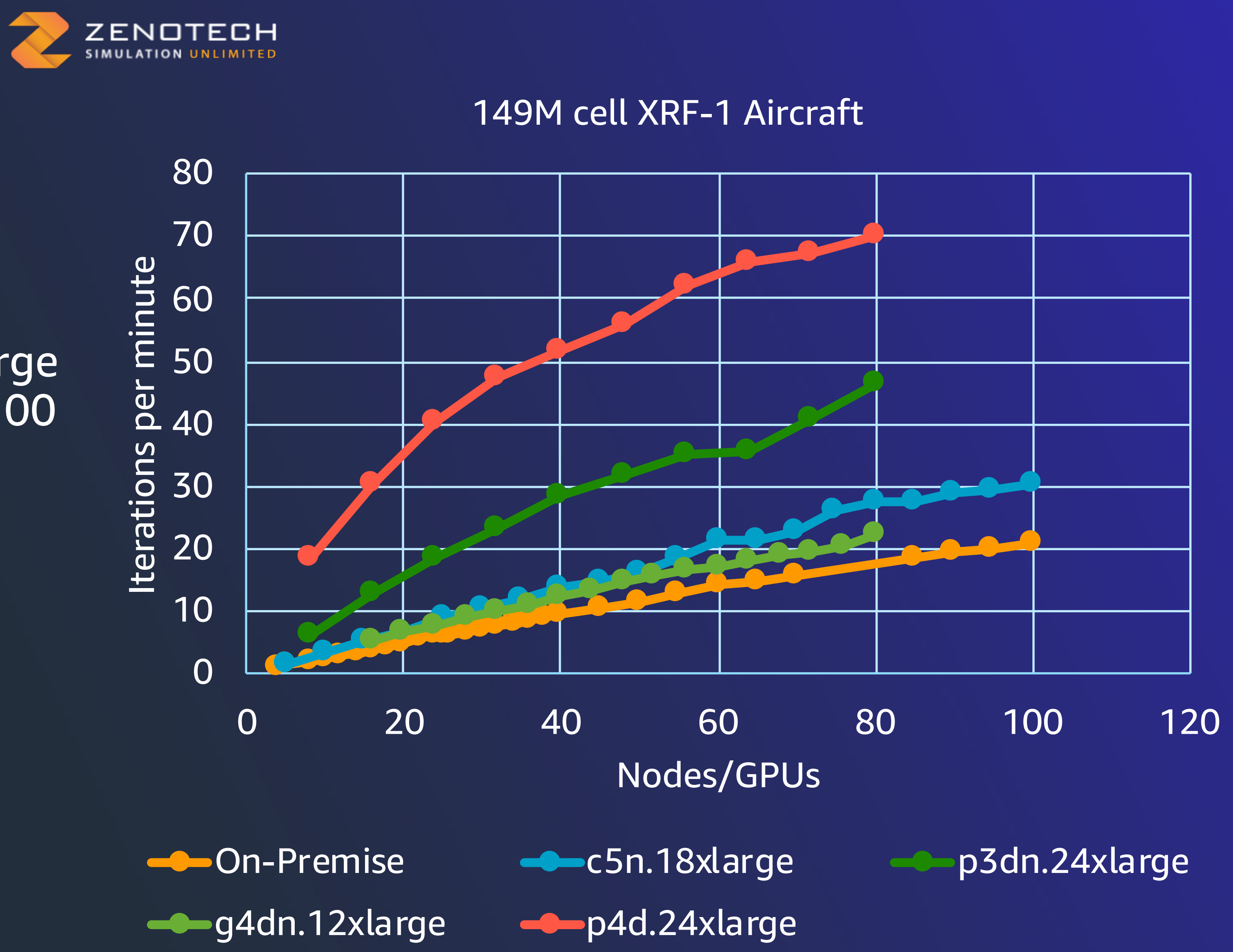




aws

GPUs?

- Recent work with Zenotech (zCFD) is very interesting
- GPUs (P3.24xlarge and G4dn.16xlarge Amazon EC2 instances – NVIDIA v100 and T4) deliver faster and cheaper results
- P4d.24xlarge (NVIDIA A100) gave even higher boost for a lower cost

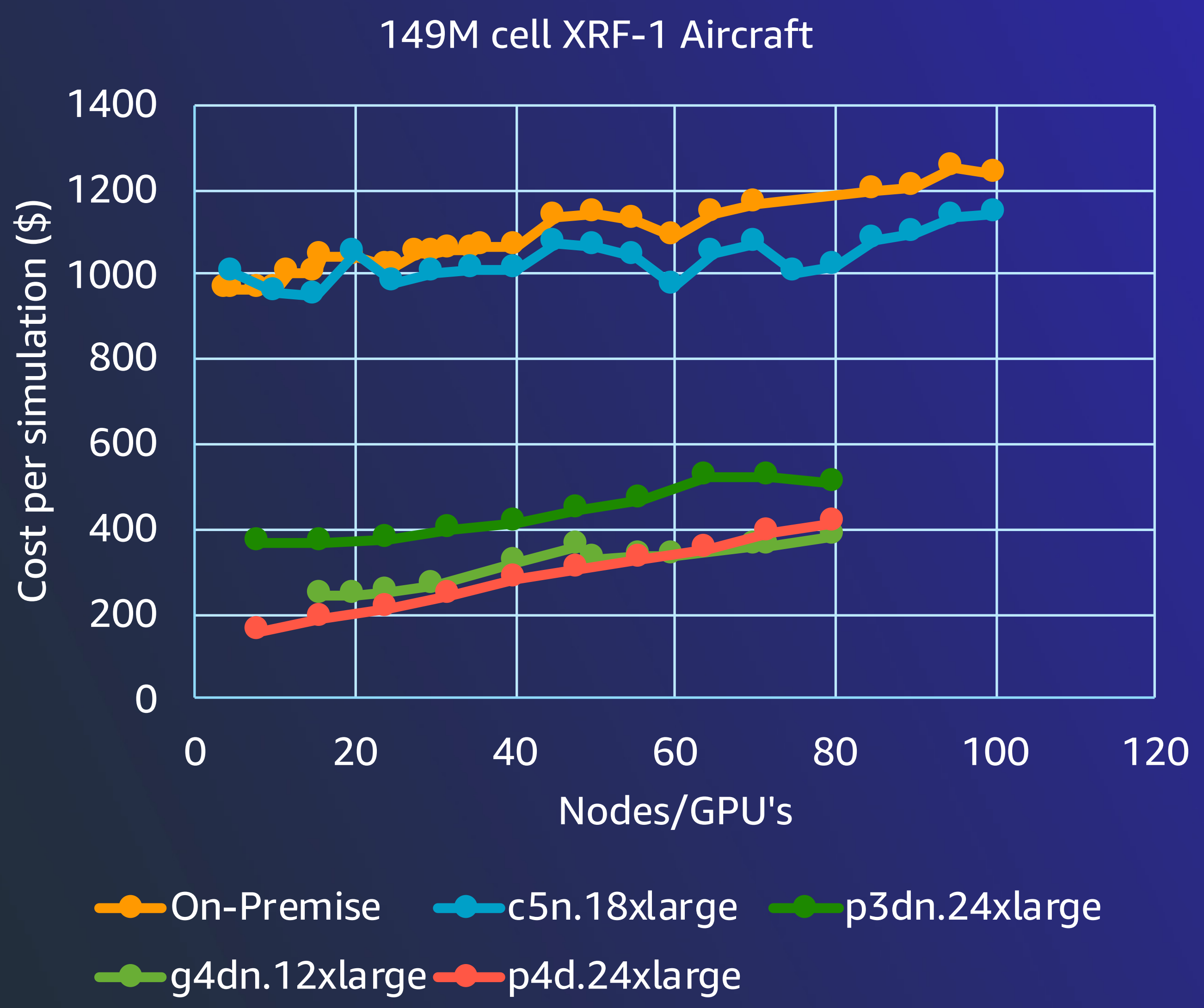


GPUs?

- Recent work with Zenotech (zCFD) is very interesting

- GPUs (P3.24xlarge and G4dn.16xlarge Amazon EC2 instances – NVIDIA v100 and T4) deliver faster and cheaper results

- P4d.24xlarge (NVIDIA A100) gave even higher boost for a lower cost



Key services that enable HPC on AWS

AWS ParallelCluster

Amazon EC2

Elastic Fabric Adapter (EFA)

Amazon FSx for Lustre

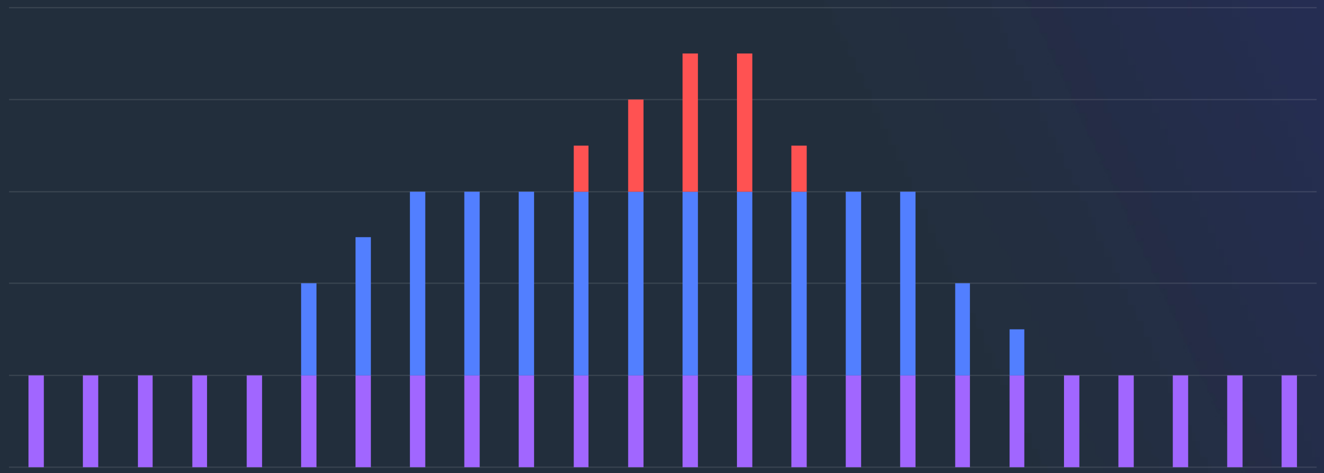
AWS Batch

Infrastructure as code i.e ability to update your entire environment via API calls

Simplifying capacity and cost optimization

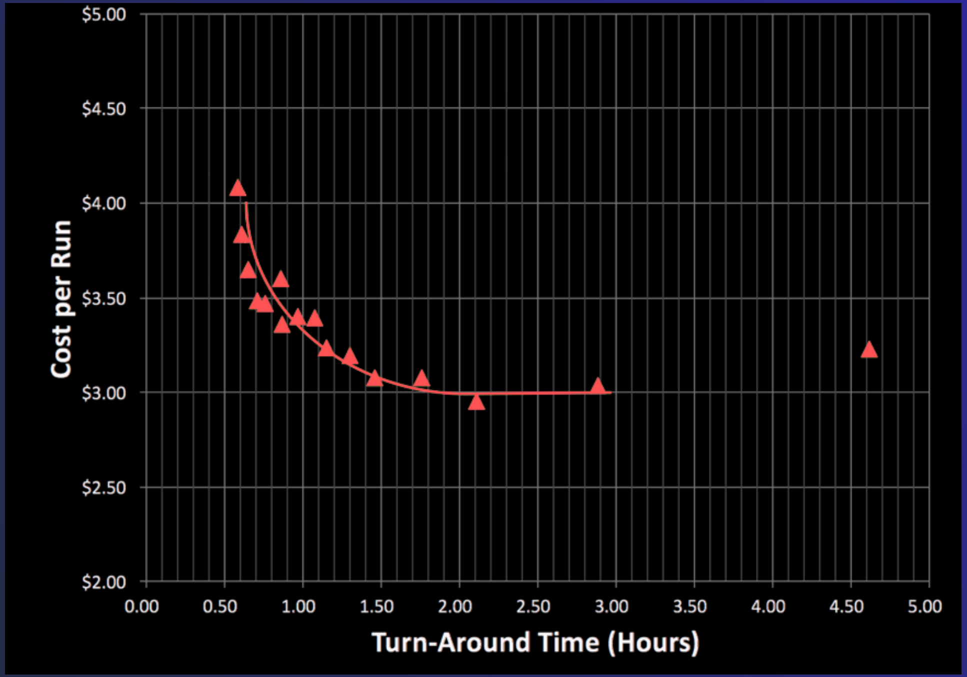
Scale using Spot Instances, On-Demand Instances, or both

Use Reserved Instances/Saving Plans for known/steady-state workloads



Range of AWS budget tools to help you control your spend

Evaluate the trade-off of time to solution vs. cost for scaling



Customer case study – Joby Aviation

- Novel aircraft meant limited prior knowledge – CFD was key to the design
- Conduct thousands of simulations on hundreds of cores using Amazon EC2 C5n.18xlarge instances
- AWS ParallelCluster to easily deploy an HPC environment
- Amazon S3 Glacier to store hundreds of terabytes of data

"Using AWS helps us get results from our CFD work faster. In some cases, we get results in 24 hours that may have previously taken a week."

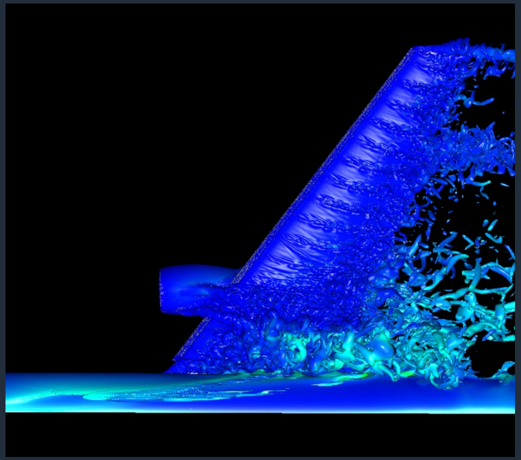
Alex Stoll,
Aeromechanics Lead, Joby Aviation

Boom Supersonic goes all in on AWS

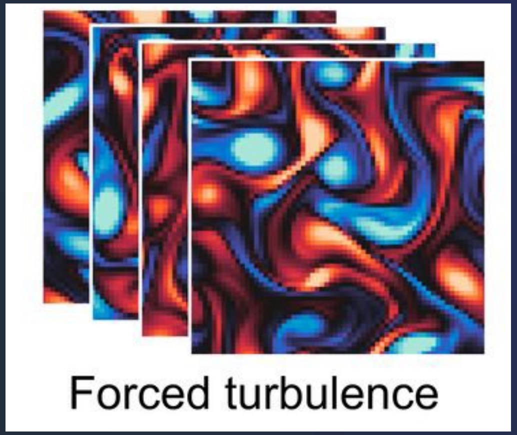
- Boom Supersonic used AWS to **accelerate the design and construction** of its supersonic aircraft
- Boom can run thousands of advanced computer simulations concurrently, resulting in an estimated **6x increase in productivity versus an on-premises environment**
- Boom used more than **53 million compute hours** on **AWS to complete design and testing** of its Overture airliner
- 525TB of data stored in Amazon S3

The future: ML and HPC possibilities

Extract relationships from simulation data

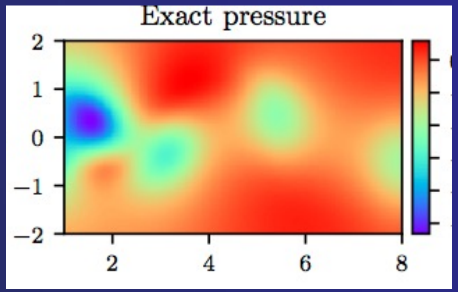


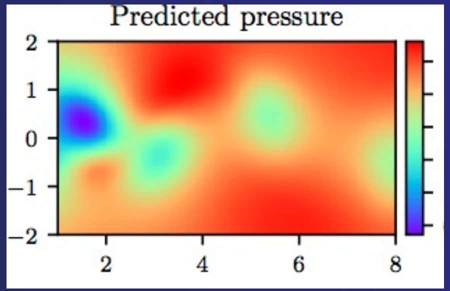
Accelerate HPC simulations with embedded ML models



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

Learn solutions to governing equations directly



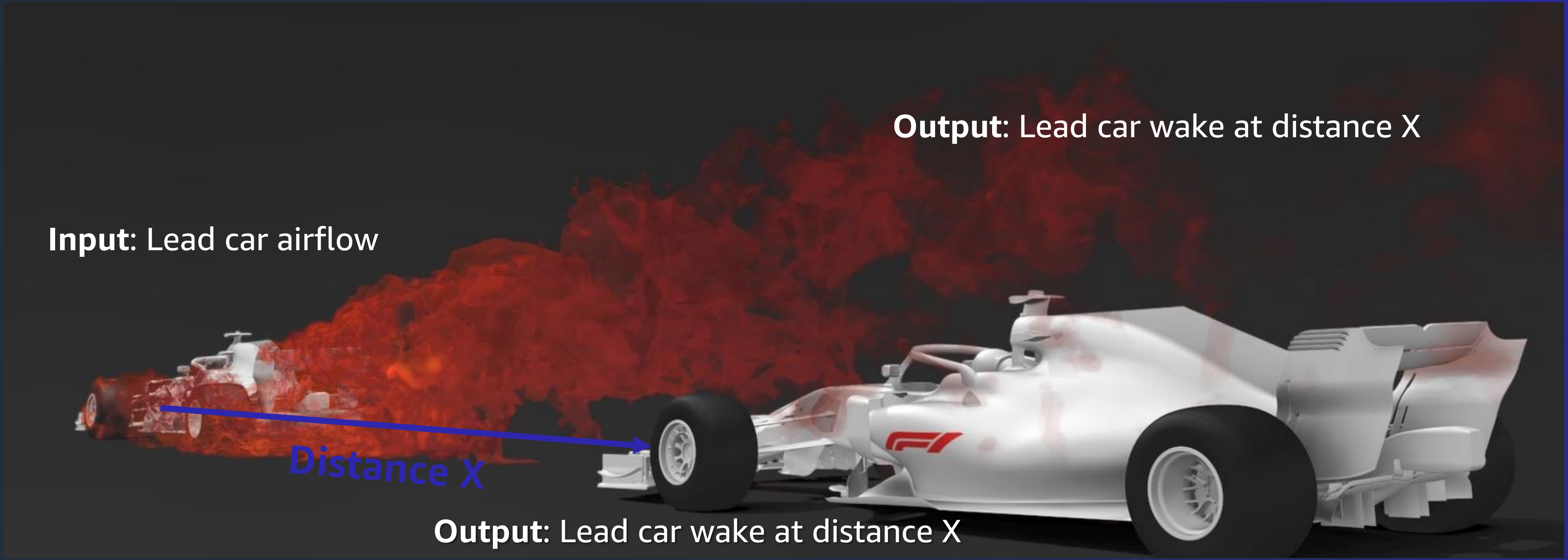


Credit: Physcs Informed Deep Learning,
maziarraissi.github.io/PINNs/

AWS hardware (Amazon EC2 P4d, AWS Trainium, AWS Inferentia) and services (Amazon SageMaker, Amazon Deep Learning AMIs) to accelerate

Demonstrating ML + HPC

Extract relationships from HPC simulation results



Problem statement:

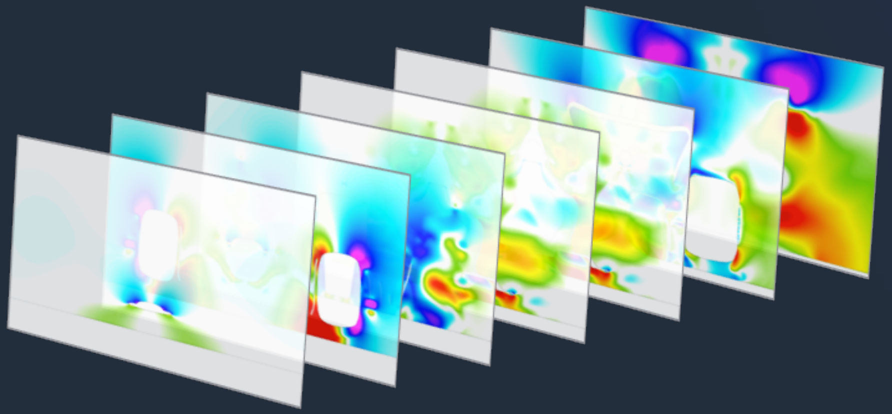
Determine optimal lead car features to obtain target flow profile at distance X

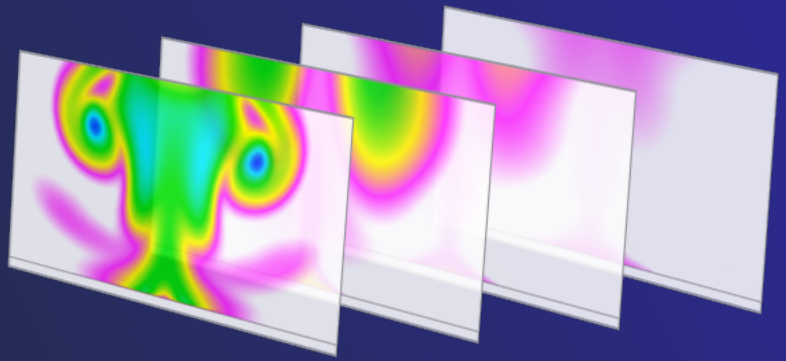
ML + HPC Solution

3D Geometry

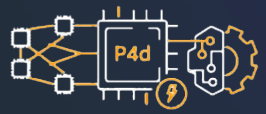
Input Flow Field

→ Deep Learning Model → Output Flow Field





EC2 UltraClusters



Summary

- **Large-scale, high-fidelity CFD** is needed to reduce need for wind tunnel and physical testing
- HPC is the main bottleneck to achieve current ‘low-fidelity’ (RANS-like)’ same business-day turnaround time
- Cloud model can help to broaden HPC access to companies/researchers and allow testing on latest hardware
- You can spin up **tens of thousands of cores and hundreds of TB HPC systems** in a manner of minutes and change their configuration with a line of code
- **Provides a great way to combine HPC+ML+Digital Twins**

Thank you!