The Convergence of Big Compute and Big Data in Cloud-Based HPC

David Pellerin, HPC Business Development Principal

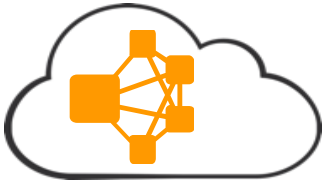
June, 2016

Motivators for the Cloud in HPC

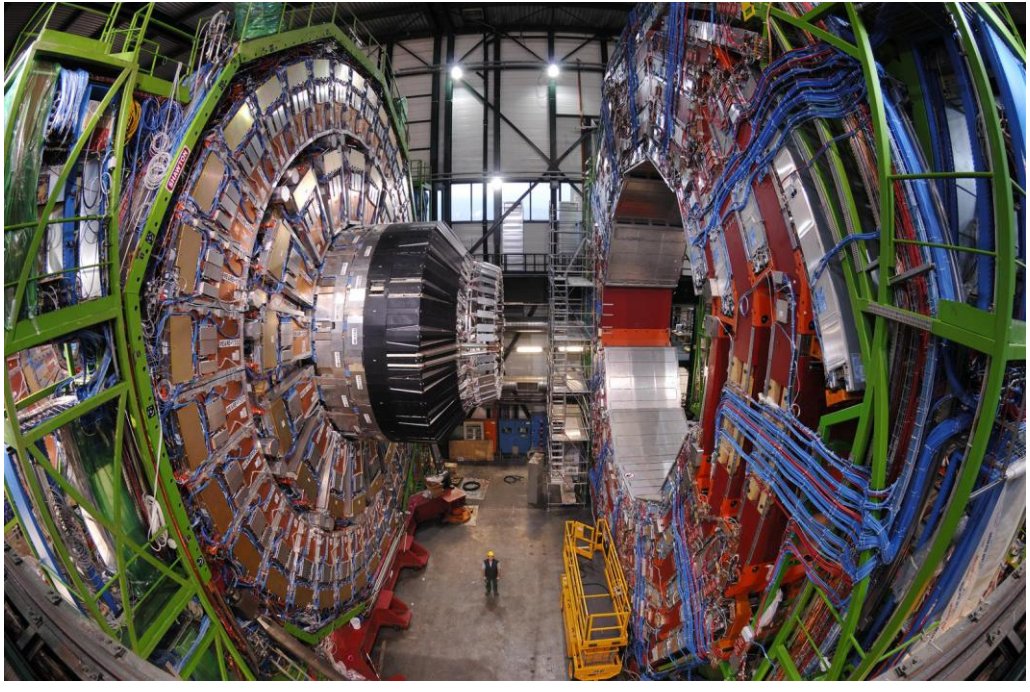
Cloud for HPC Scalability

Cloud for Secure Global Collaboration

Cloud for Big Data and IoT



Cloud Enables Scale for Big Data and Big Compute



Finding Patterns in the Data

This is

Big Data

TECH TIMES PERSONAL TECH BIZ TECH FUTURE TECH

Scientists; Evidence Of New, Unknown Particle?

By **Jim Algar**, Tech Times | January 12, 9:52 AM

Like Follow Share Tweet Reddit 1 Comment •••

SUBSCRIBE

An anomaly in data from particle collisions in the Large Hadron Collider has researchers scratching their heads. Is this evidence of new particles that could turn the Standard Model of Physics on its head?
(Photo : Getty Images)

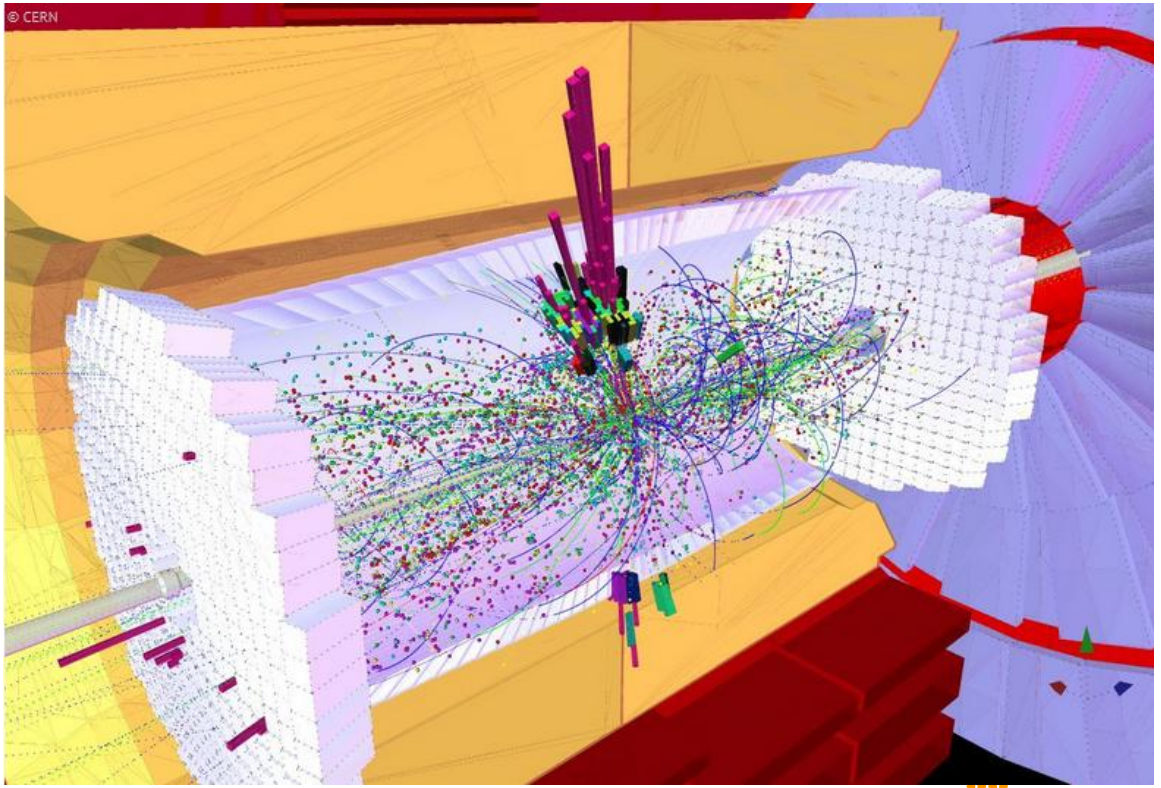
An unexpected "bump" in the data coming out of collision experiments with the Large Hadron Collider in Switzerland has scientists wondering if they've witnessed evidence of previously-unknown subatomic particles.

The collision, which cannot be explained by the Standard Model of physics, may have been evidence of a previously-undiscovered particle, or maybe even two particles, researchers say.

Building Computer Models and Running Simulations

This is

Big Compute



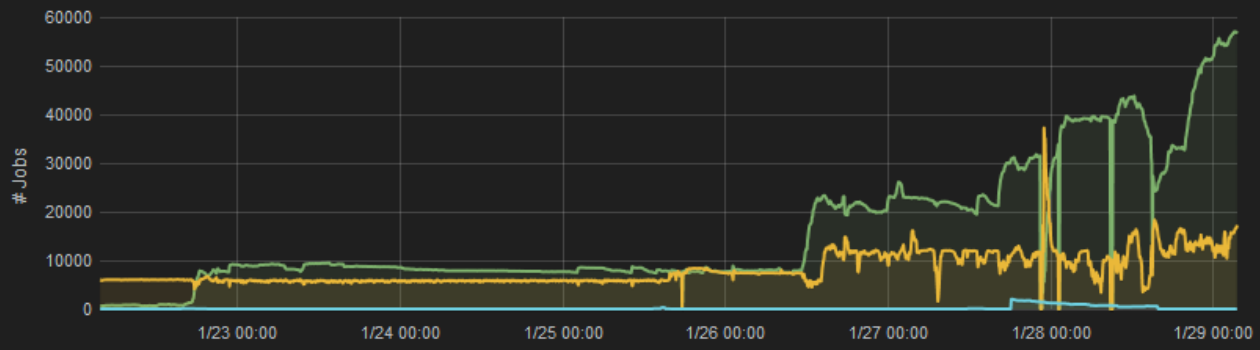
amazon web services

HEP Cloud Summary

Zoom Out | Jan 22, 2016 03:37:38 to Jan 29, 2016 03:32:38 | Refresh every 1m

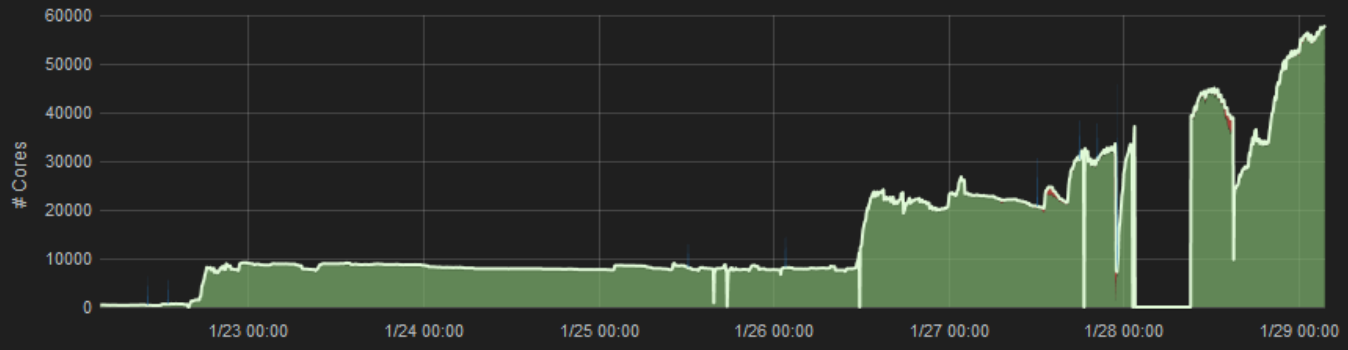
AWS VM Status | GCloud VM Status | HEP Cloud HTCondor Status | HEP Cloud Slots

Job Status



	min	max	avg	current
Running	0	56994	15861	56624
Idle (total)	0	37228	8082	17250
Idle (Fifebatch / nova)	0	2000	136	0

Slots Summary



	min	max	avg	current
01_cpu	0	57535	14214	56991
04_cpu	0	0	0	0
08_cpu	0	480	50	0
Unclaimed	0	32546	171	449
Total	0	57708	14435	57440
Idle	0	37454	46	0

Scalability and Performance for Simulations

Examples

- High-energy physics
- Weather modeling
- Fluids, structures, materials analysis
- Thermal and electromagnetic simulations
- Genomics, proteomics and molecular dynamics
- Seismic and reservoir simulations
- 3D rendering and visualizations

Cloud unlocks data-driven simulations at massive scale

Image Capture and Image Processing on AWS

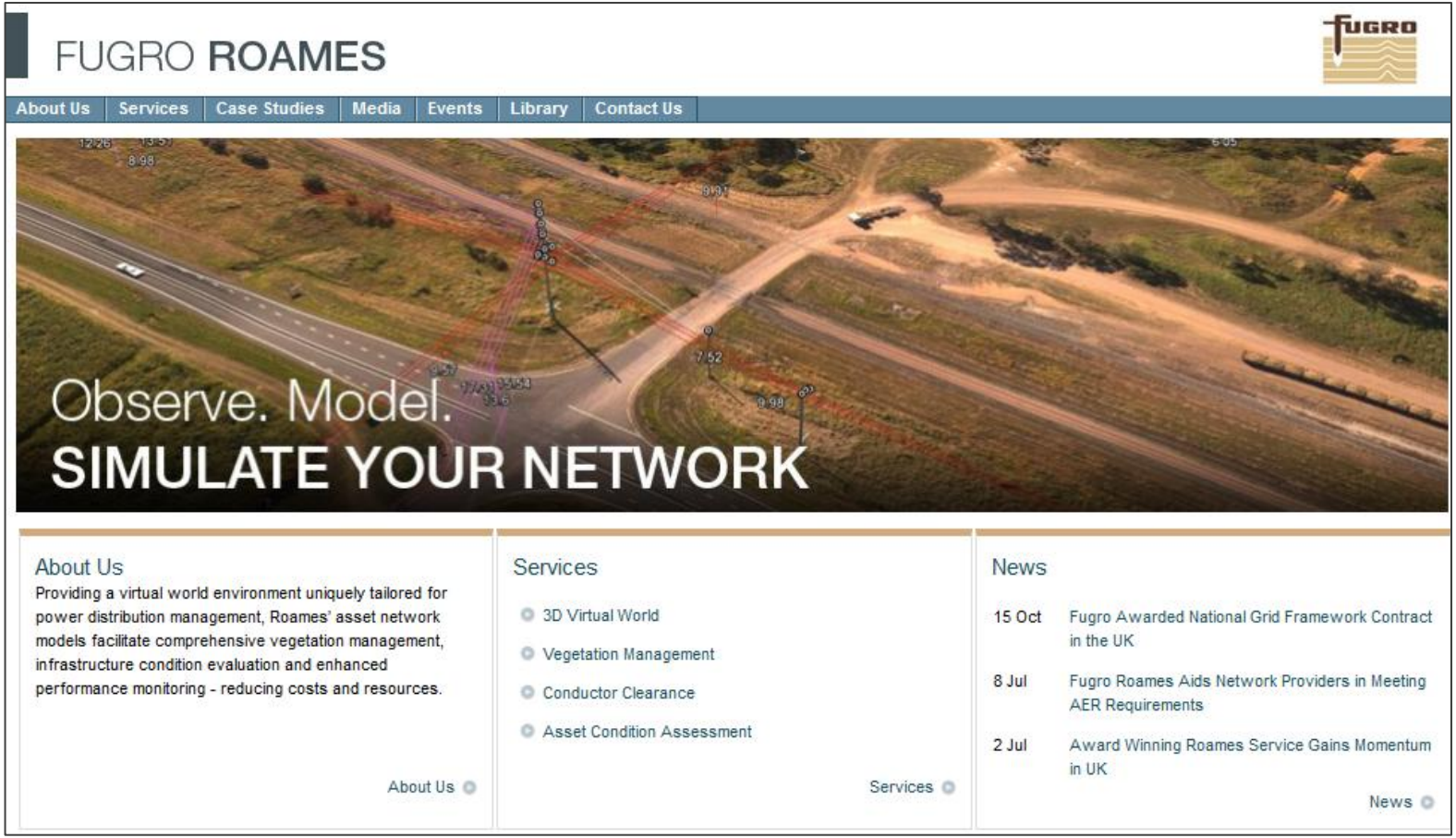
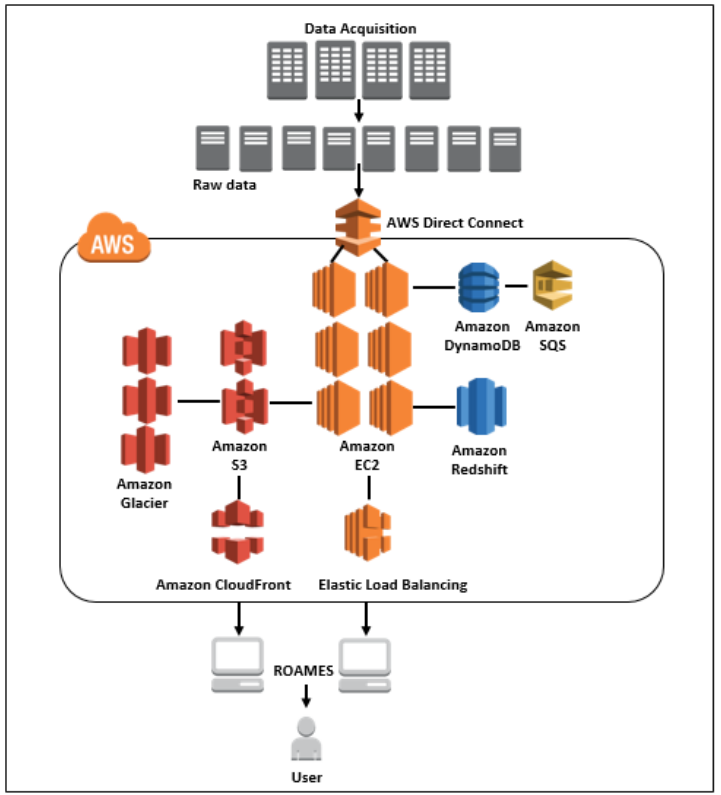


Image Capture and Processing

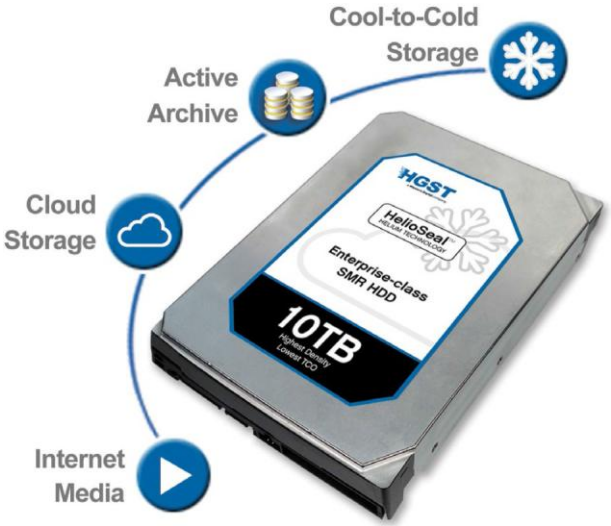
- Aircraft equipped with cameras, laser sensors
- Repeated overflights of power networks
- Captured data is used to render detailed 3D models of the power lines, and the environment
- Analytics and simulations are run to generate actionable reports
- Goal: directing post-disaster repair and prioritizing ongoing maintenance



"Fugro Roames has enabled Ergon Energy to reduce the cost of vegetation management from AU$100 million to AU$60 million per year."

- Josh Passenger, Technical Architect, Fugro Roames

Big Data and HPC in Product Engineering



HGST applications for engineering:

- ✓ Molecular dynamics, CAD, CFD, EDA
- ✓ Collaboration tools for engineering
- ✓ **Big data for manufacturing yield analysis**

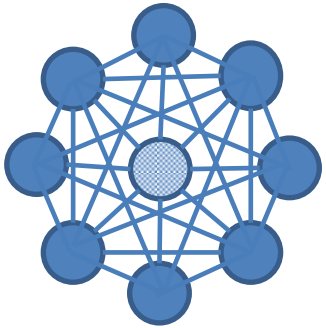
Running drive-head simulations at scale:

Millions of parallel parameter sweeps, running months of simulations in just hours.

Over 85,000 Intel cores running at peak, using EC2 Spot instances

amazon web services

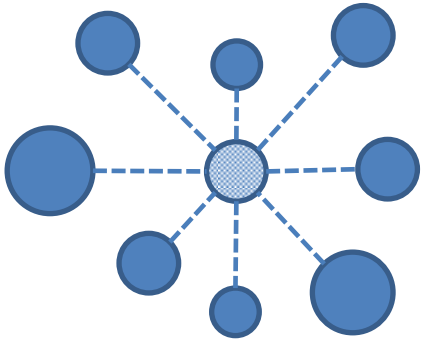
Cluster HPC and Grid HPC on the Cloud



Cluster HPC

Tightly coupled, latency sensitive applications

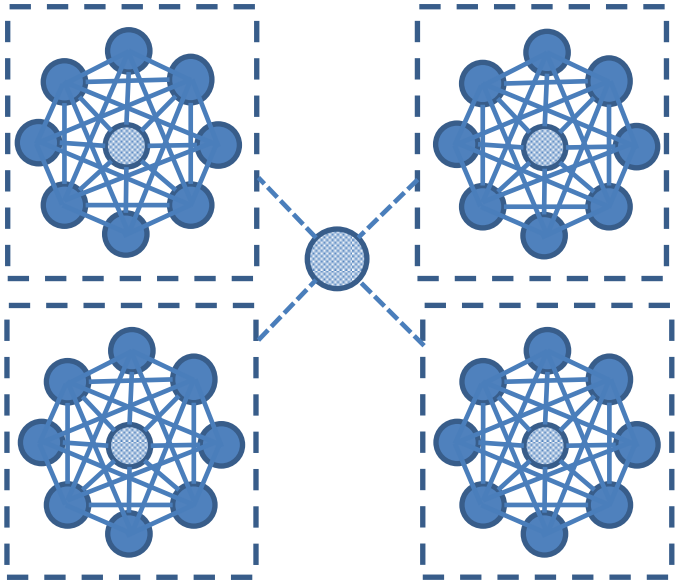
Use larger EC2 compute instances, placement groups, Enhanced Networking



Grid HPC

Loosely coupled, pleasingly parallel

Use a variety of EC2 instances, multiple AZs, Spot, Auto Scaling, SQS



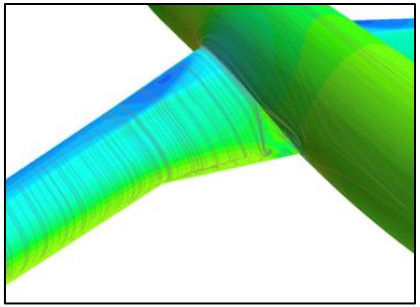
Grids of Clusters

Use a grid strategy on the cloud to run a group of parallel, individually clustered HPC jobs

amazon web services

Scaling Fluid Dynamics on AWS

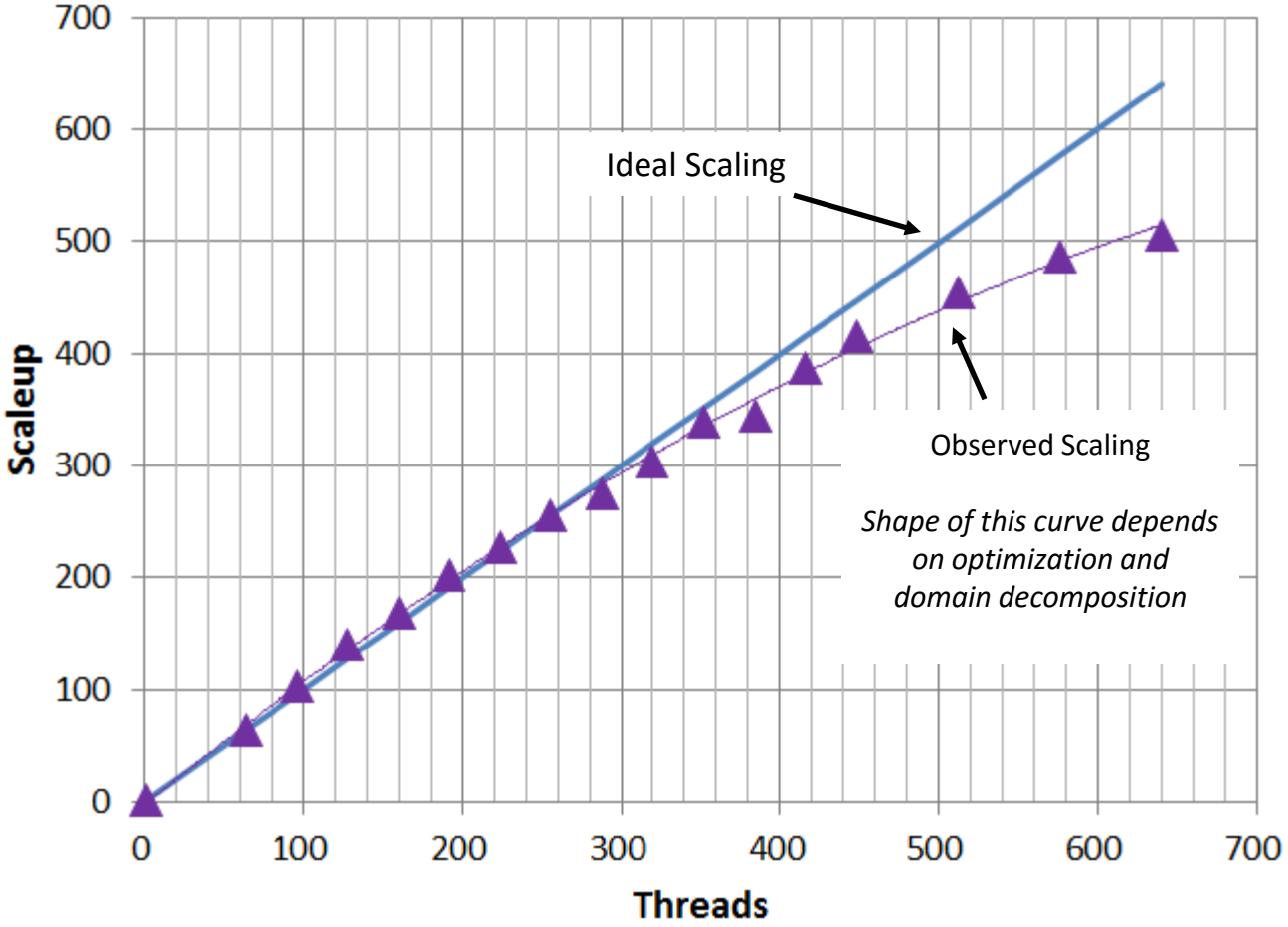
TLG AEROSPACE
THE LOADS GROUP



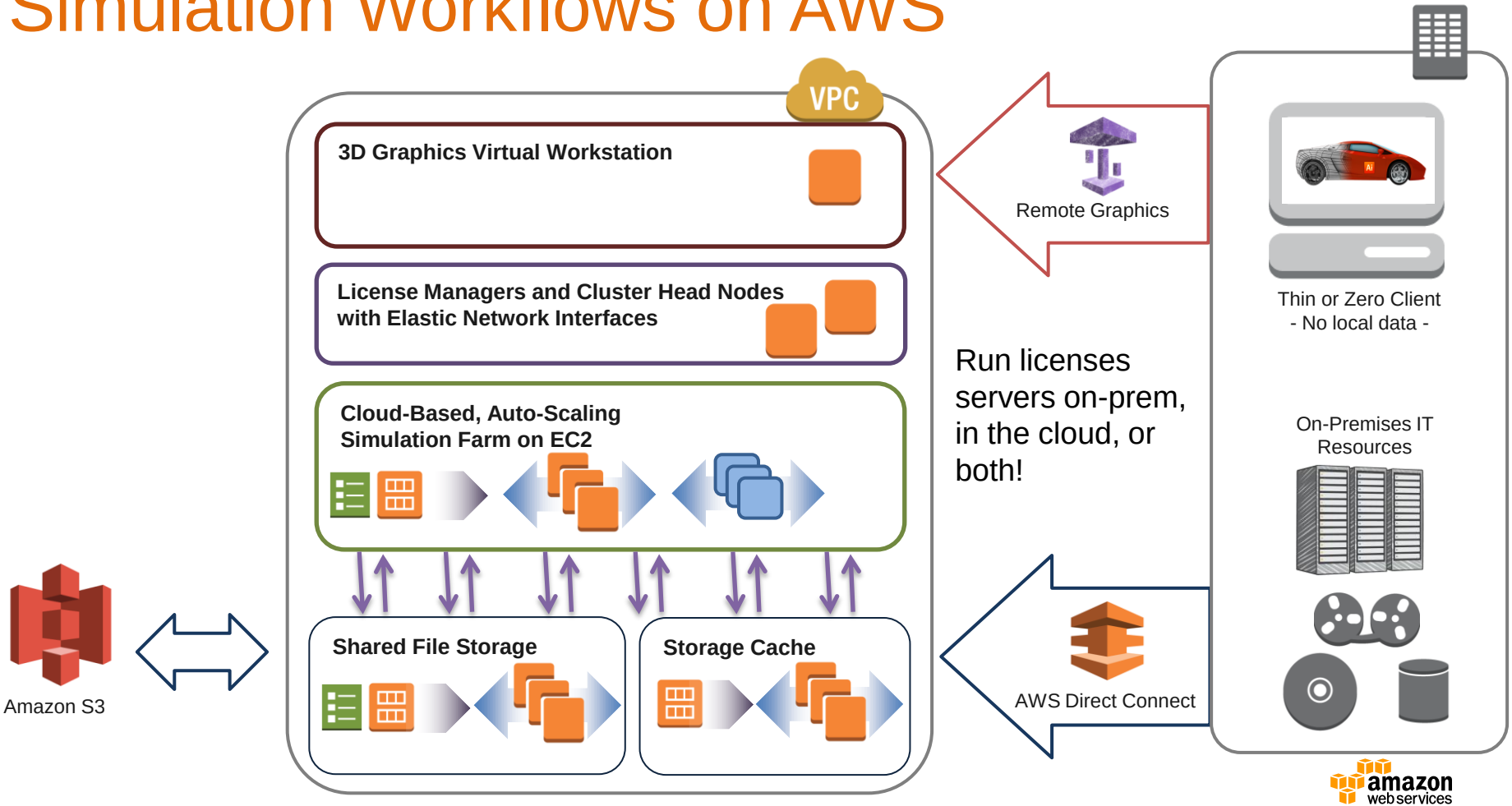
16M cell, polyhedral, external aero case, STAR-CCM+

Running on threads, c4.8xlarge instances

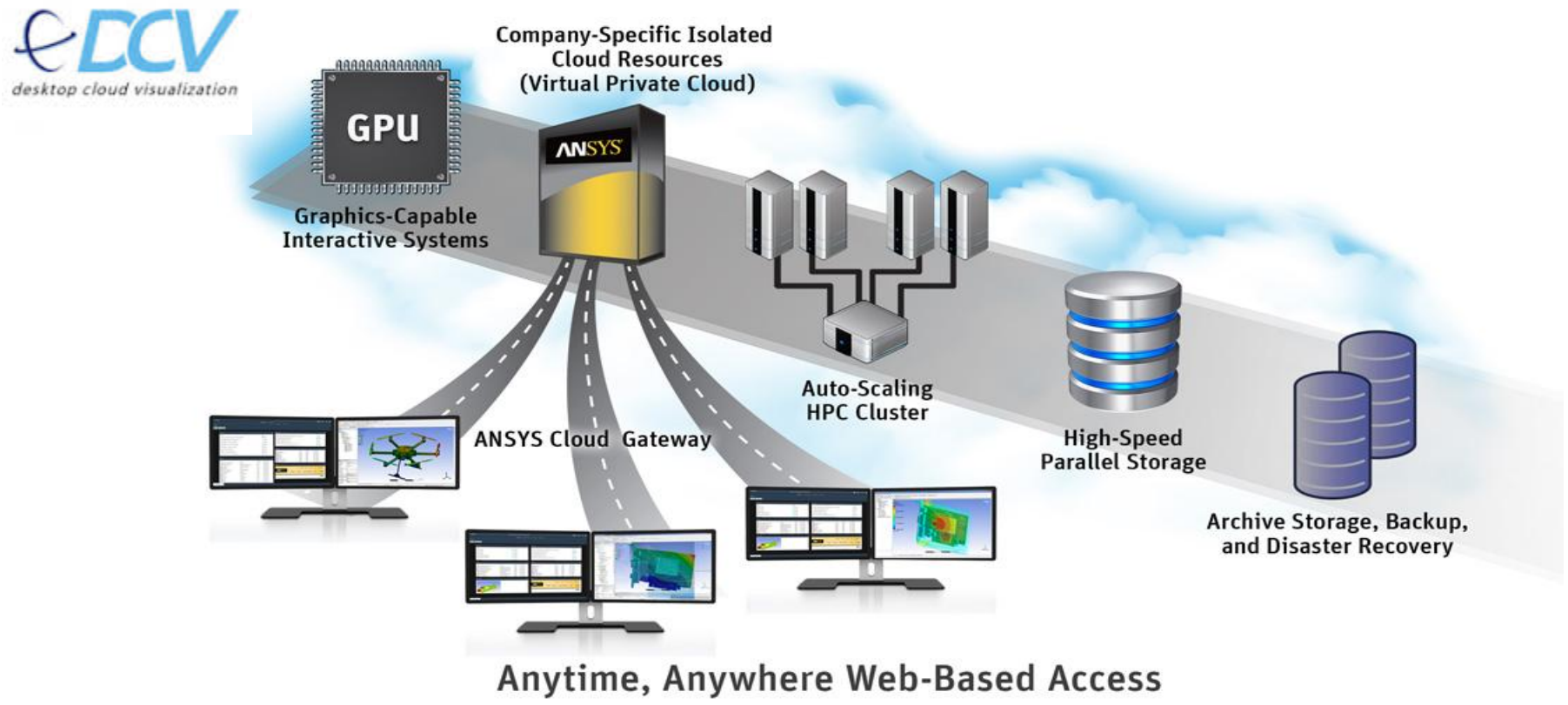
Demonstrates excellent scalability for typical CFD models



Simulation Workflows on AWS

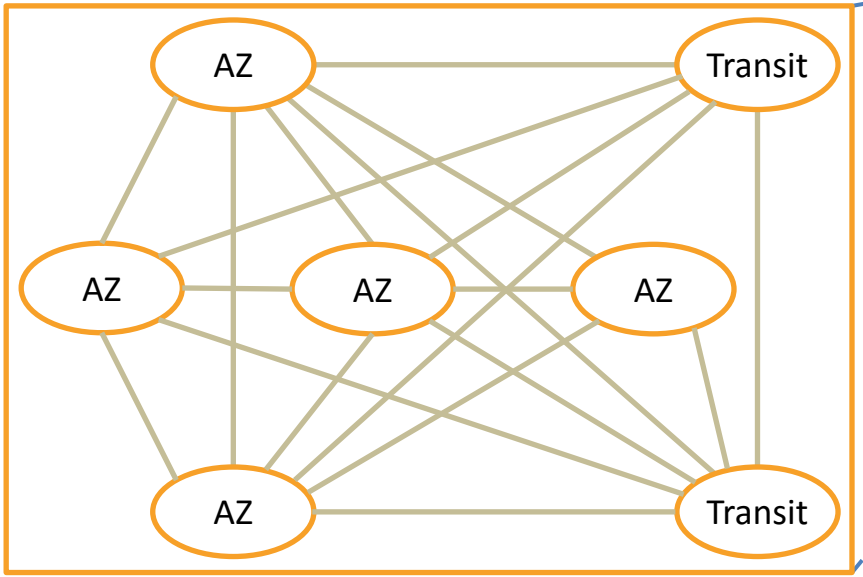


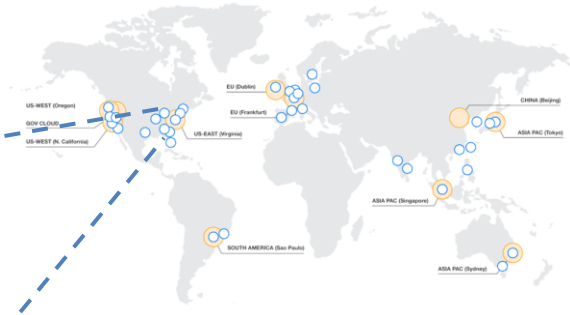
Example: ANSYS Enterprise Cloud on AWS



Global Cloud Services – Regions and AZs

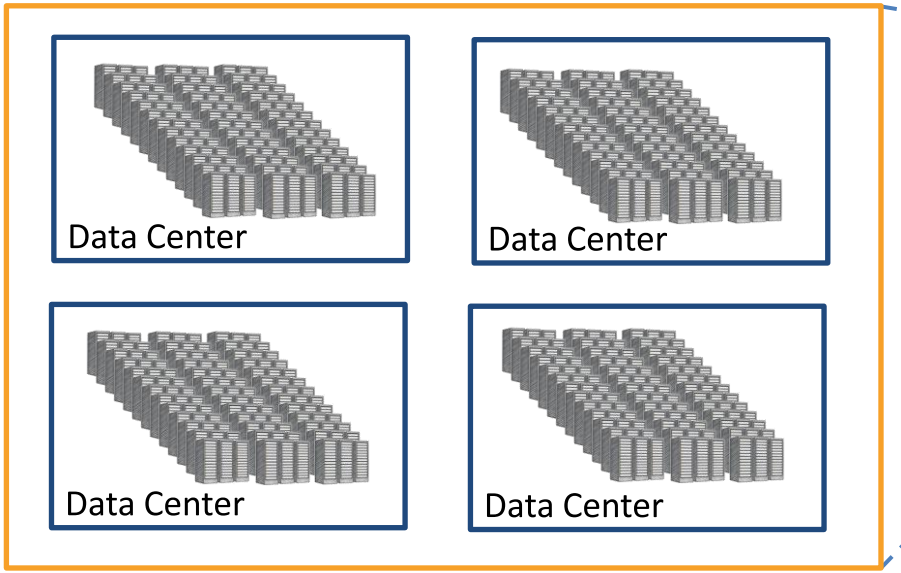
Example AWS Region

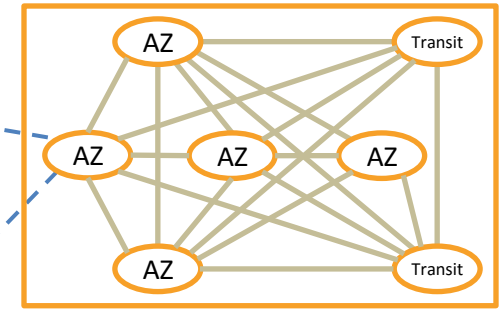




- Mesh of Availability Zones (AZ) and Transit Centers
- **Redundant** paths to transit centers
- Transit centers connect to:
  - Private links to other AWS regions
  - Private links to customers
  - Internet through peering & paid transit

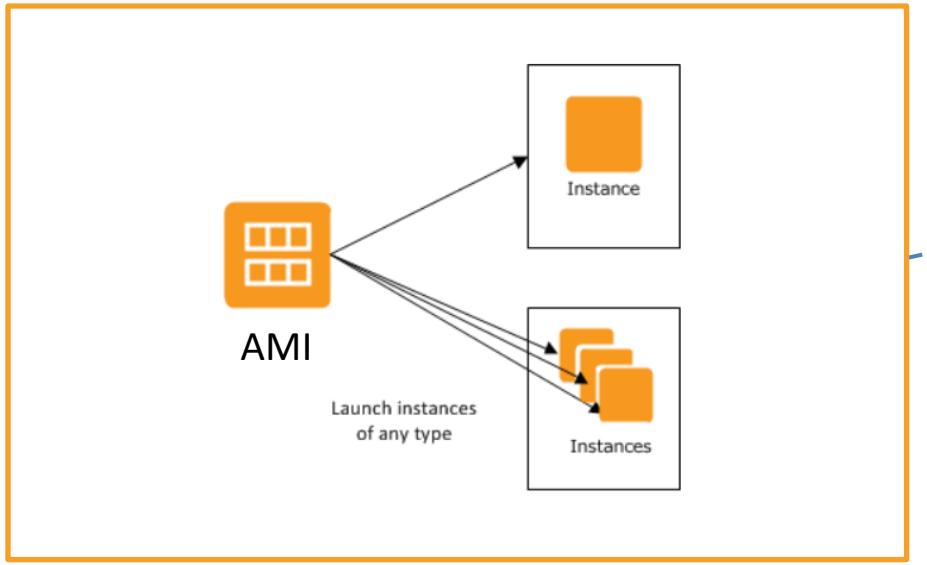
Example AWS Availability Zone

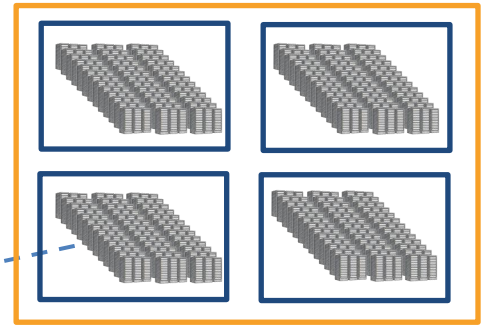




- 30+ AZs world-wide
- All regions have 2 or more AZs
- Each AZ is 1 or more DC
  - No data center is in two AZs
  - Some AZs have as many as 6 DCs

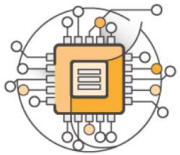
AWS Machine Images and Instances





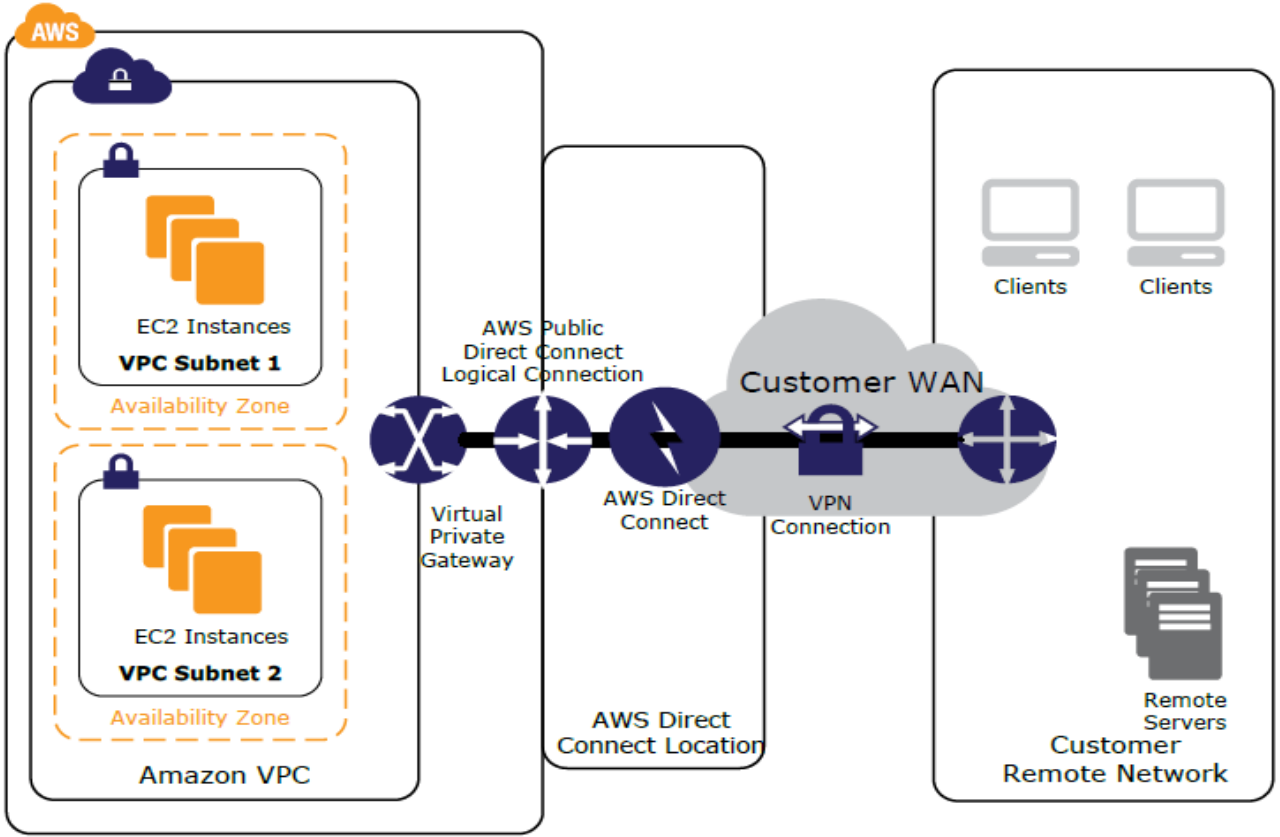
Instance types

General Purpose:	M1, M3, M4, T2
Compute Optimized:	C1, CC2, C3, C4
Memory Optimized:	M2, CR1, R3, X1
Storage Optimized:	HI1, HS1, I2
GPU:	CG1, G2
Micro:	T1, T2



amazon web services

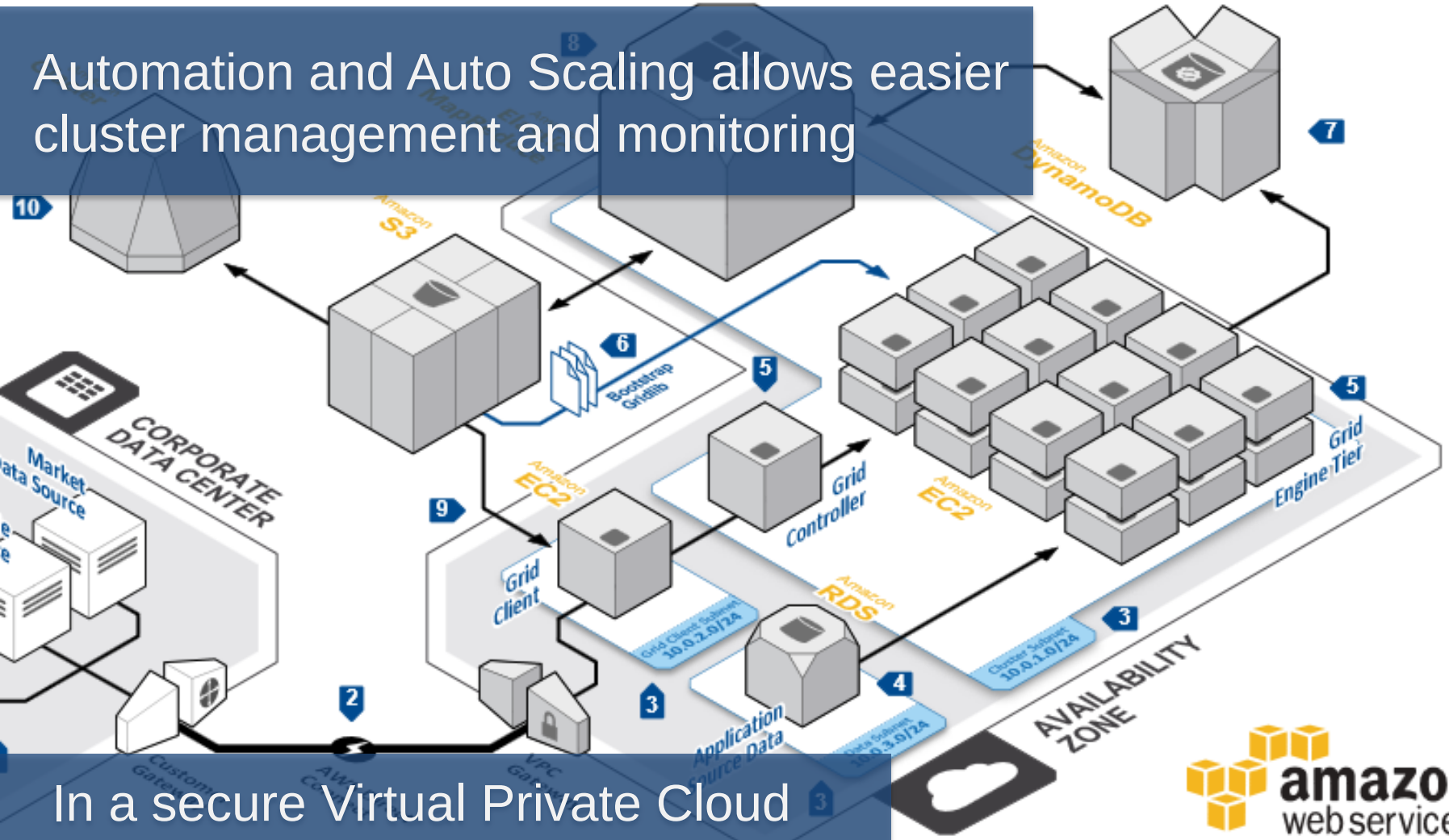
Virtual Private Cloud (VPC)



VPC Connectivity options:
http://media.amazonwebservices.com/AWS_Amazon_VPC_Connectivity_Options.pdf

amazon web services

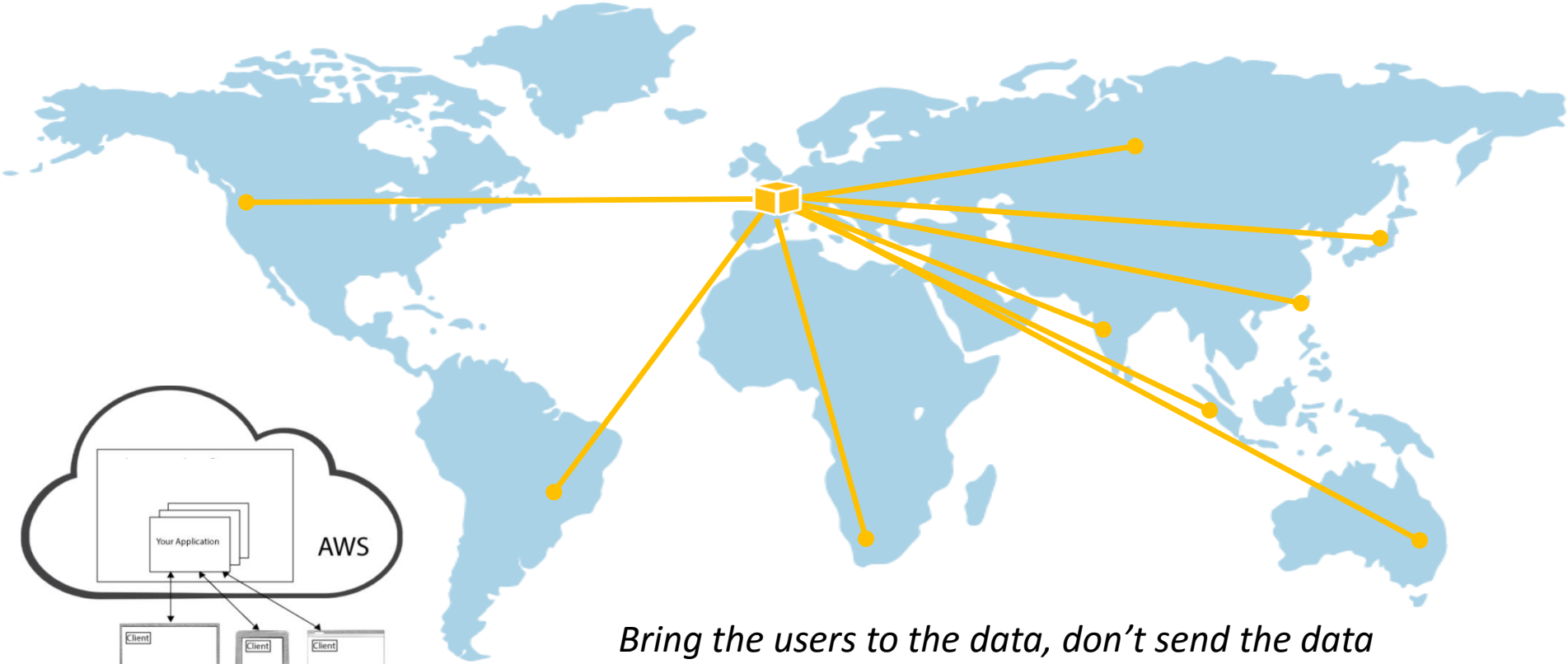
Automation and Auto Scaling allows easier cluster management and monitoring



In a secure Virtual Private Cloud

amazon web services™

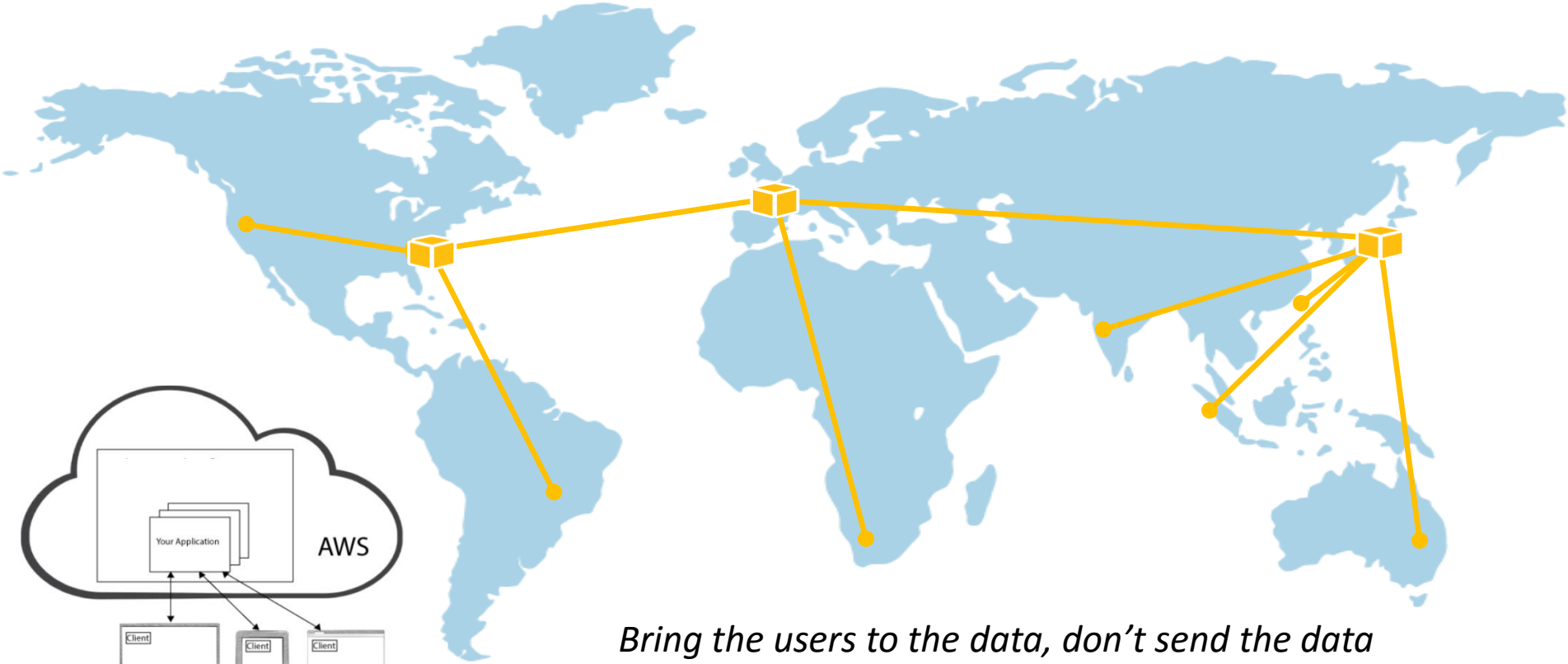
Enabling Global Collaboration



Bring the users to the data, don't send the data to the users

amazon web services

Enabling Global Collaboration



Bring the users to the data, don't send the data to the users

amazon web services

Cloud is not the first platform shift…

There was a time when…

- Technical workstations were turnkey, single-purpose, vertically integrated, and more truly “bare metal”

What happened?

- General-purpose Unix workstations and servers became available, and…
- The problem spaces outgrew single workstations, giving rise to the centrally managed, time-sliced HPC cluster

Now?

- The problem spaces are fast outgrowing the centrally managed, special-purpose cluster
- The answer is cloud, including high performance virtualization and containers

History Favors Economies of Scale

1985
Application-specific technical workstations

1995
Economics of scale via general-purpose, high performance Unix workstations

Cloud is the new, more scalable technical computing platform

2005
Application-specific, datacenters for HPC

Today
Economies of scale via general-purpose, high performance cloud

amazon web services

Resources

aws.amazon.com/hpc

aws.amazon.com/big-data/

dpelleri@amazon.com