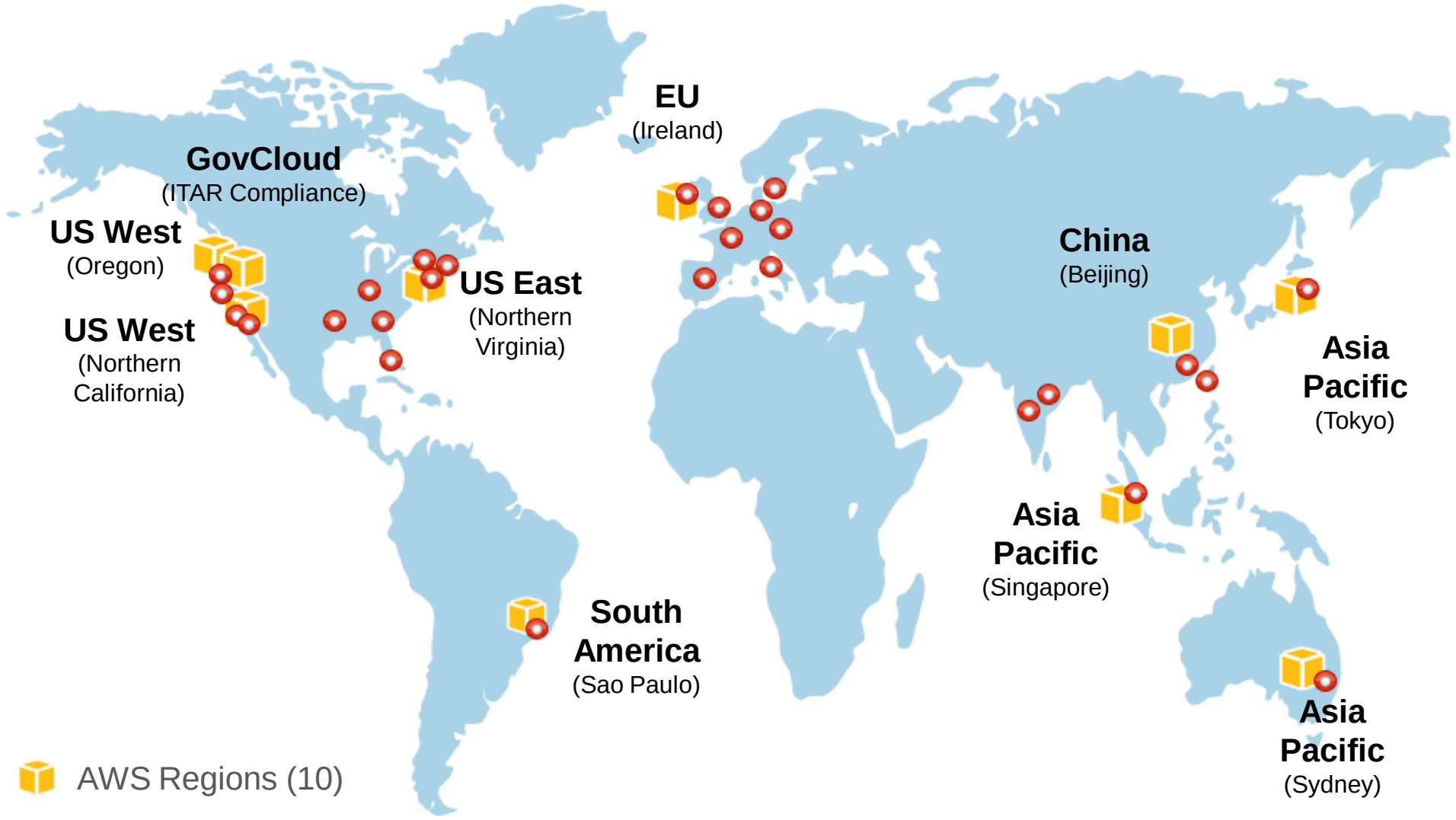
AWS Cloud for HPC and Big Data

David Pellerin, Business Development Principal
IDC HPC User Forum

September 16, 2014

AWS Regions



amazon web services

Customer Applications

Development & Deployment

Monitoring

CloudWatch

Deployment & Management

BeanStalk, OpsWork, Cloud Formation, CloudTrail

Identity & Access

IAM, Federation

Application Services

Applications

SES, SNS, SQS, Elastic Transcoder, CloudSearch, SWF, AppStream

Foundation Services

Databases

RDS, Dynamo, ElastiCache, RedShift

Analytics

EMR, DataPipeline, Kinesis

Content Delivery

CloudFront

Compute

EC2, WorkSpaces

Storage

S3, EBS, Glacier, Storage Gateway

Networking

VPC, Direct Connect, ELB, Route53

API

Human Interaction

Support

Interaction

Web Console

Command Line

Libraries, SDK's

node, ruby, Java, JS, python, Microsoft .net, php, iOS, android

Regions

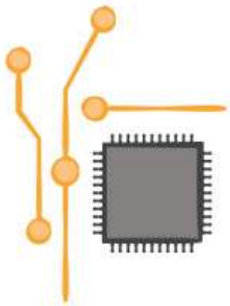
Availability Zones

Edge Locations

AWS Global Infrastructure

EC2 Instance Type History

Increasing customer choice…



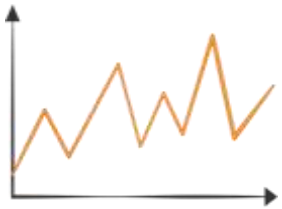
new

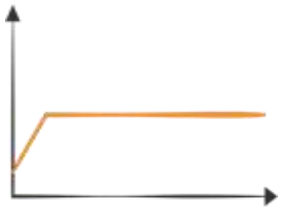
existing

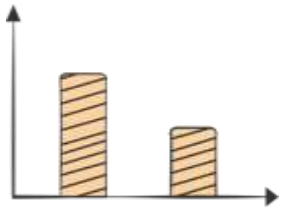
2006	2007	2008	2009	2010	2011	2012-2013	September, 2014
m1.small	m1.xlarge m1.large m1.small	c1.medium c1.xlarge m1.xlarge m1.large m1.small	m2.2xlarge m2.4xlarge c1.medium c1.xlarge m1.xlarge m1.large m1.small	cc1.4xlarge cg1.4xlarge t1.micro m2.xlarge m2.2xlarge m2.4xlarge c1.medium c1.xlarge m1.xlarge m1.large m1.small	cc2.8xlarge cc1.4xlarge cg1.4xlarge t1.micro m2.xlarge m2.2xlarge m2.4xlarge c1.medium c1.xlarge m1.xlarge m1.large m1.small	cr1.8xlarge hs1.8xlarge m3.xlarge m3.2xlarge hi1.4xlarge m1.medium cc2.8xlarge cg1.4xlarge t1.micro m2.xlarge m2.2xlarge m2.4xlarge c1.medium c1.xlarge m1.xlarge m1.large m1.small	t2.micro t2.small t2.medium t1.micro hs1.8xlarge m3.xlarge m3.2xlarge hi1.4xlarge m1.medium cc2.8xlarge cr1.8xlarge cg1.4xlarge m2.xlarge m2.2xlarge m2.4xlarge c1.medium c1.xlarge m1.xlarge m1.large m1.small g2.2xlarge hs1.xlarge hs1.2xlarge hs1.4xlarge c3.large c3.xlarge c3.2xlarge c3.4xlarge c3.8xlarge m3.medium m3.large i2.large i2.xlarge i2.4xlarge i2.8xlarge r3.large r3.xlarge r3.2xlarge r3.4xlarge r3.8xlarge

Multiple Purchase Models

On-Demand	Reserved	Spot
Pay for compute capacity by the hour with no long-term commitments	Make a low, one-time payment and receive a significant discount on the hourly charge	Bid for unused capacity, charged at a Spot Price which fluctuates based on supply and demand
For spiky workloads, or to define needs	For committed utilization	For time-insensitive or transient workloads





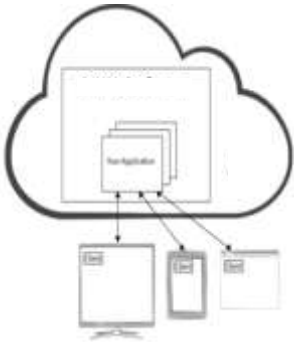


AWS Spot is a game-changer for HPC

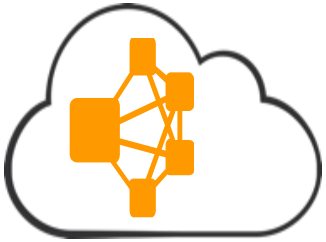
Motivators for HPC in the Cloud

Cloud for HPC Scalability

Cloud for Secure Global Collaboration



Cloud for Big Data



HGST
a Western Digital company

"HGST is using AWS for a higher performance, lower cost, faster deployed solution vs buying a huge on-site cluster."

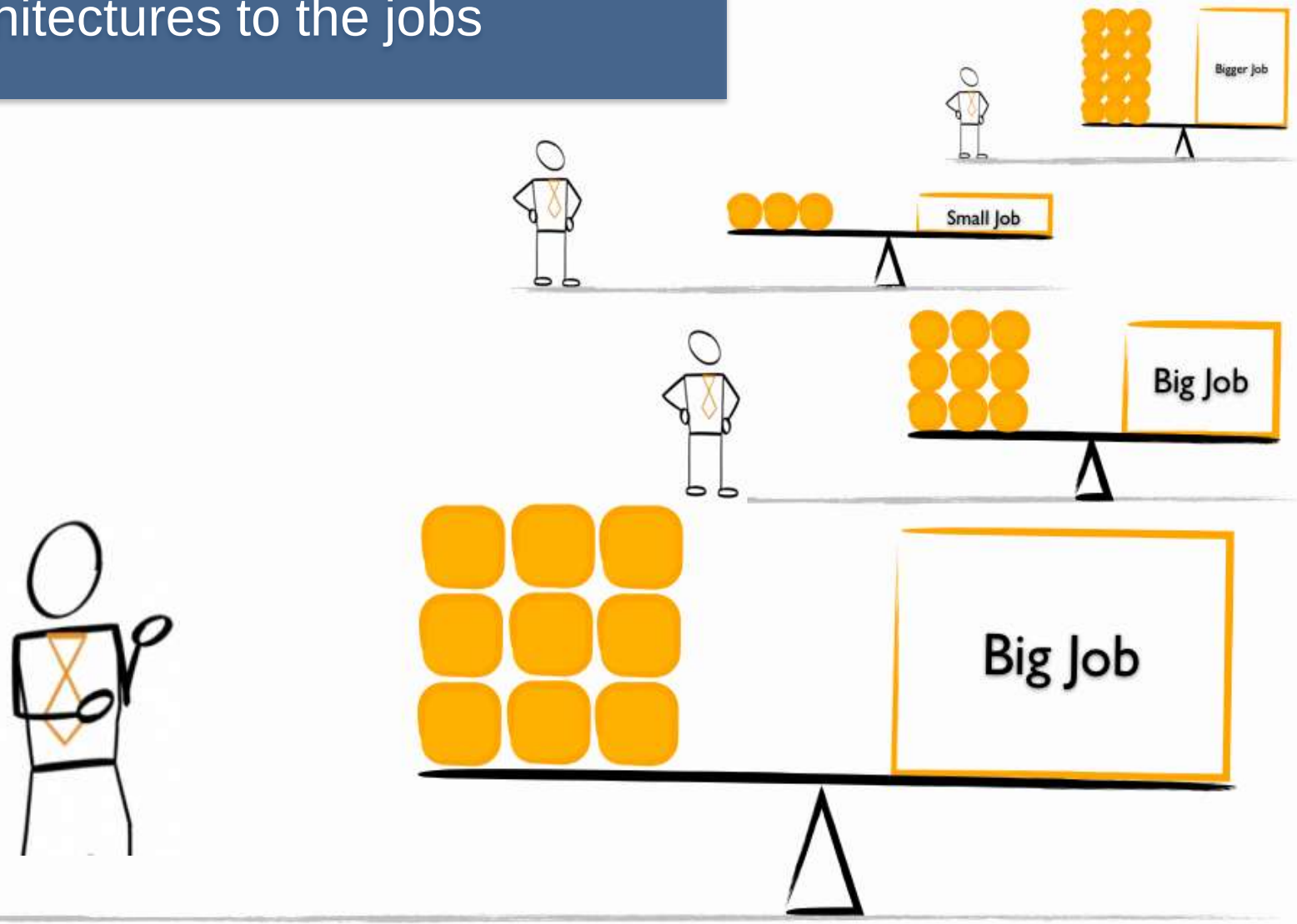
- Steve Philpott, CIO

HGST application roadmap:

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

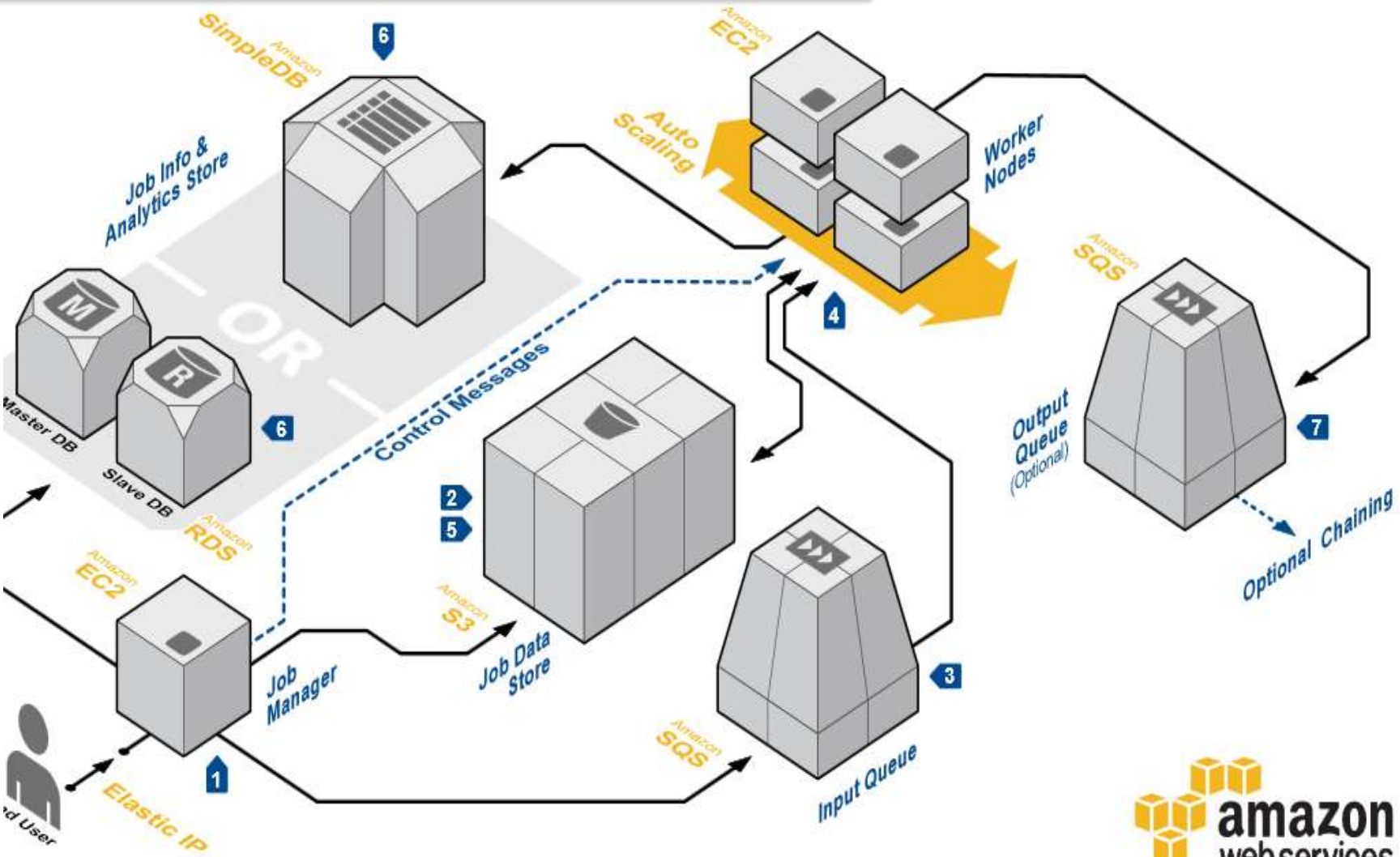
amazon
web services

On AWS, deploy multiple clusters running at the same time and match the architectures to the jobs



Use automation to manage cluster sizing and monitor jobs and costs

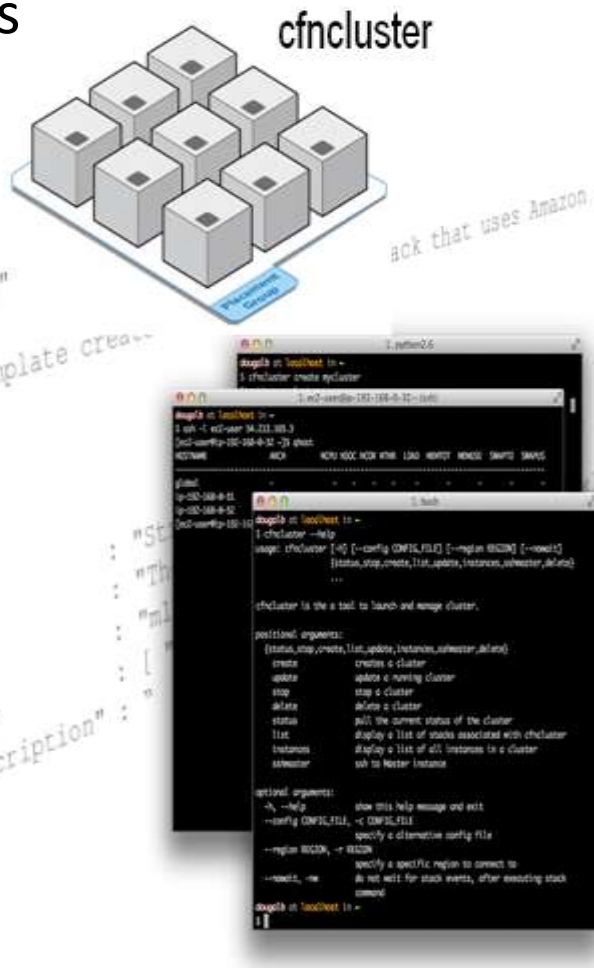
AWS Auto Scaling can work with existing HPC scheduling software



amazon web services

Many HPC Deployment Methods

- Traditional HPC schedulers and cluster managers
- “Born in the cloud” tools
  - MIT StarCluster
  - Cycle Computing CycleServer
- AWS-provided tools and APIs
  - Cloudformation, Auto Scaling
  - cfncluster (github.com/awslabs/cfncluster)
- For many use-cases…



Touch-Sensor Modeling on AWS

for TRUETOUCH® Touchscreen Controllers

A. Gourevitch

Cypress Semiconductor Corp., San Jose, CA, USA

We report an implementation of parallel computing on Amazon Web Services™ (AWS) for touch-sensor modeling. COMSOL Multiphysics® was used to simulate an electromagnetic field distribution in a capacitive sensor assembly. Multiple COMSOL jobs were deployed on separate AWS instances and were executed in parallel. The simulation results indicate that implementation of parallel computing for COMSOL simulations can significantly reduce the computational time required for optimization of capacitive touch sensor patterns.

Files Available for Download

Abstract

Paper

Courtesy of Cypress Semiconductor

amazon web services

Reservoir Simulation on AWS

Search

+61 (08) 9446 2099

Stochastic Simulation

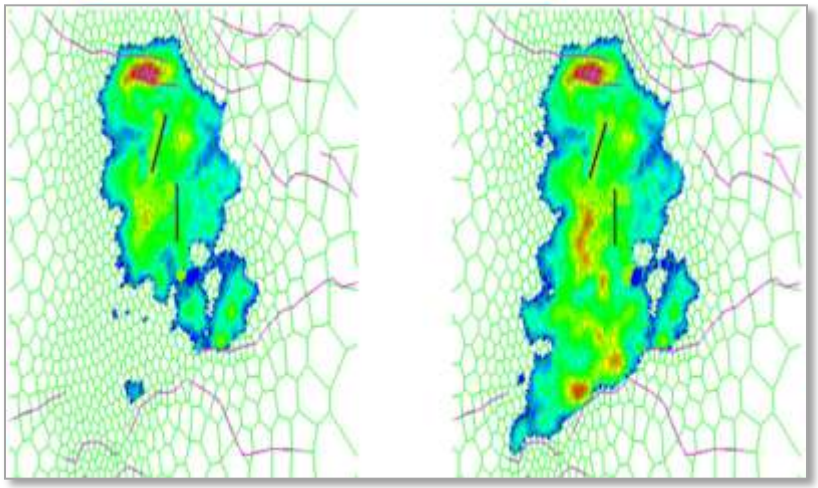
Home | About Us | ResAssure | GasAssure | News | Resources | Investors | Contact

Technical Overview | Unique Features | Million Realisations in 24 Hours | Features and Benefits | Consulting | FAQs | Register Your Interest | Benchmark Studies

Technical Overview

What is ResAssure?

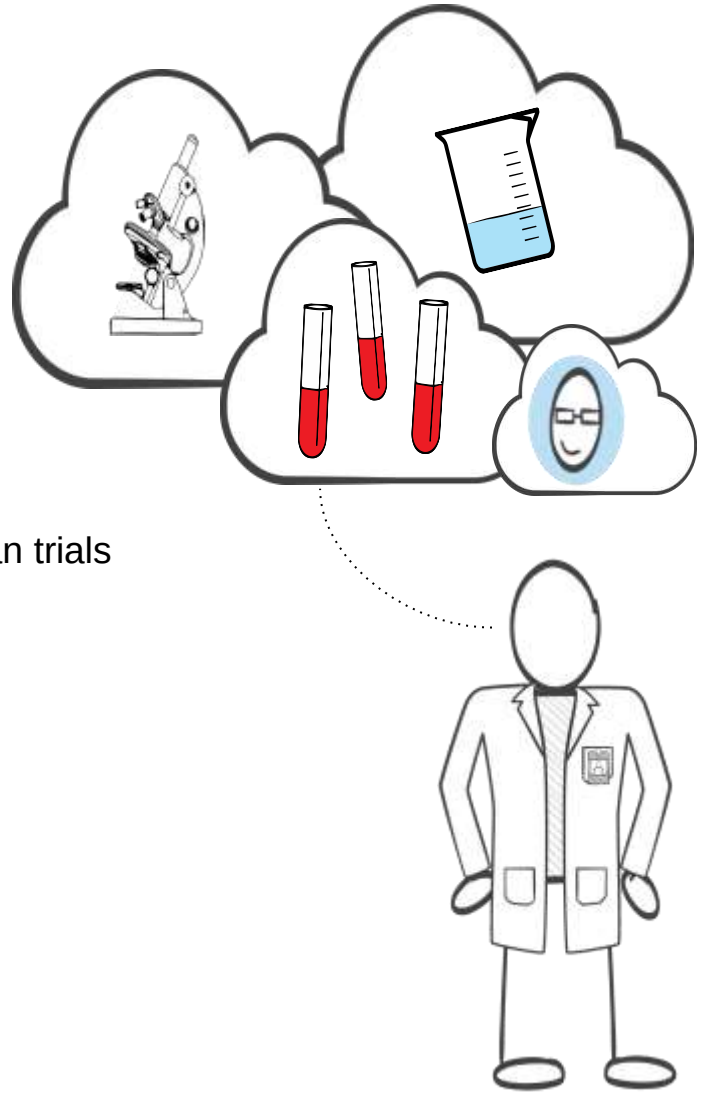
ResAssure is a Stochastic Simulation software solution and it's innovation in reservoir simulation currently solves fully-implicit, dynamic three-phase fluid flow equations for every geological realisation. ResAssure marks a significant milestone in the history of reservoir simulation, leading technological advancements in the oil and gas industry. It has been developed to work with standard reservoir simulation package datasets such as Eclipse, CMG and VIP.



amazon web services

Bristol-Myers Squibb Clinical Trials on AWS

	On-Premises	Cloud
# of Simulations	2000	2000
# of Servers	2	256
Total Run Time (hr)	**60**	**1.2**



Clinical trial simulations took 98% less time

More efficient and iterative simulations results in fewer human trials

64% savings on clinical trial costs

We're using fewer subjects in these trials, and needing fewer blood samples.

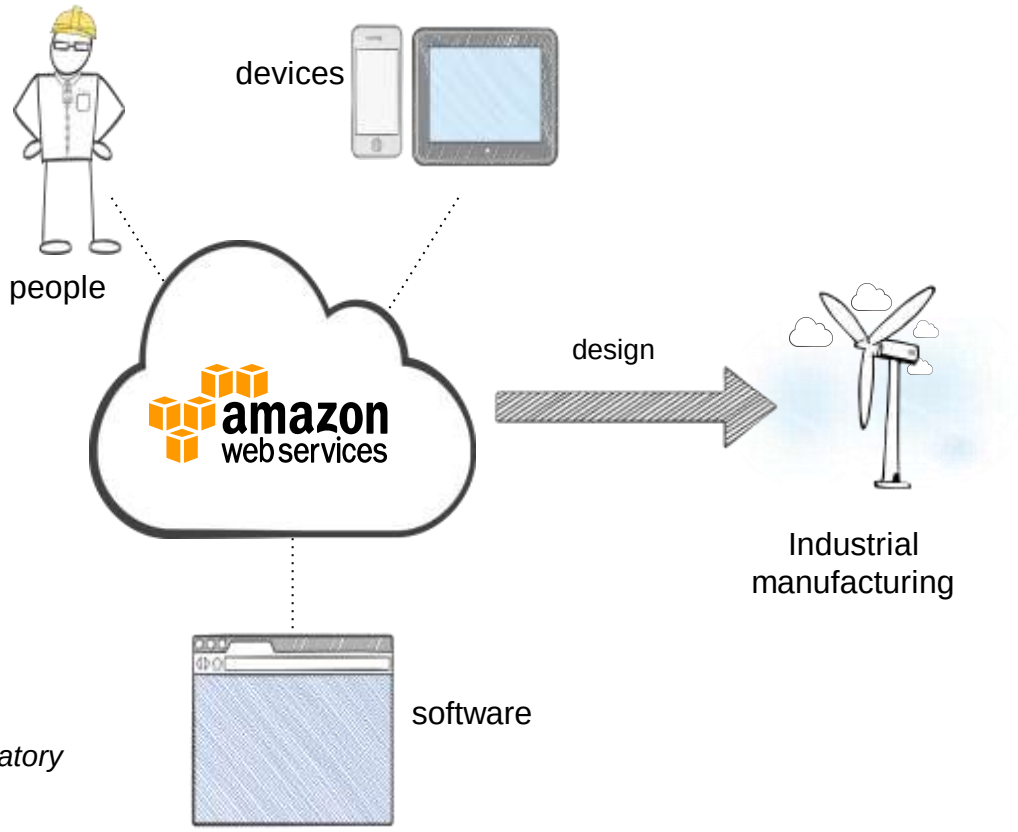
Russell Towell
Senior Solutions Specialist

Bristol-Myers Squibb

amazon web services

Collaboration and Design in the Cloud

Cross-functional collaboration app
Helps design around manufacturing
Allows users to define how they work
Users can spin-up their own environments

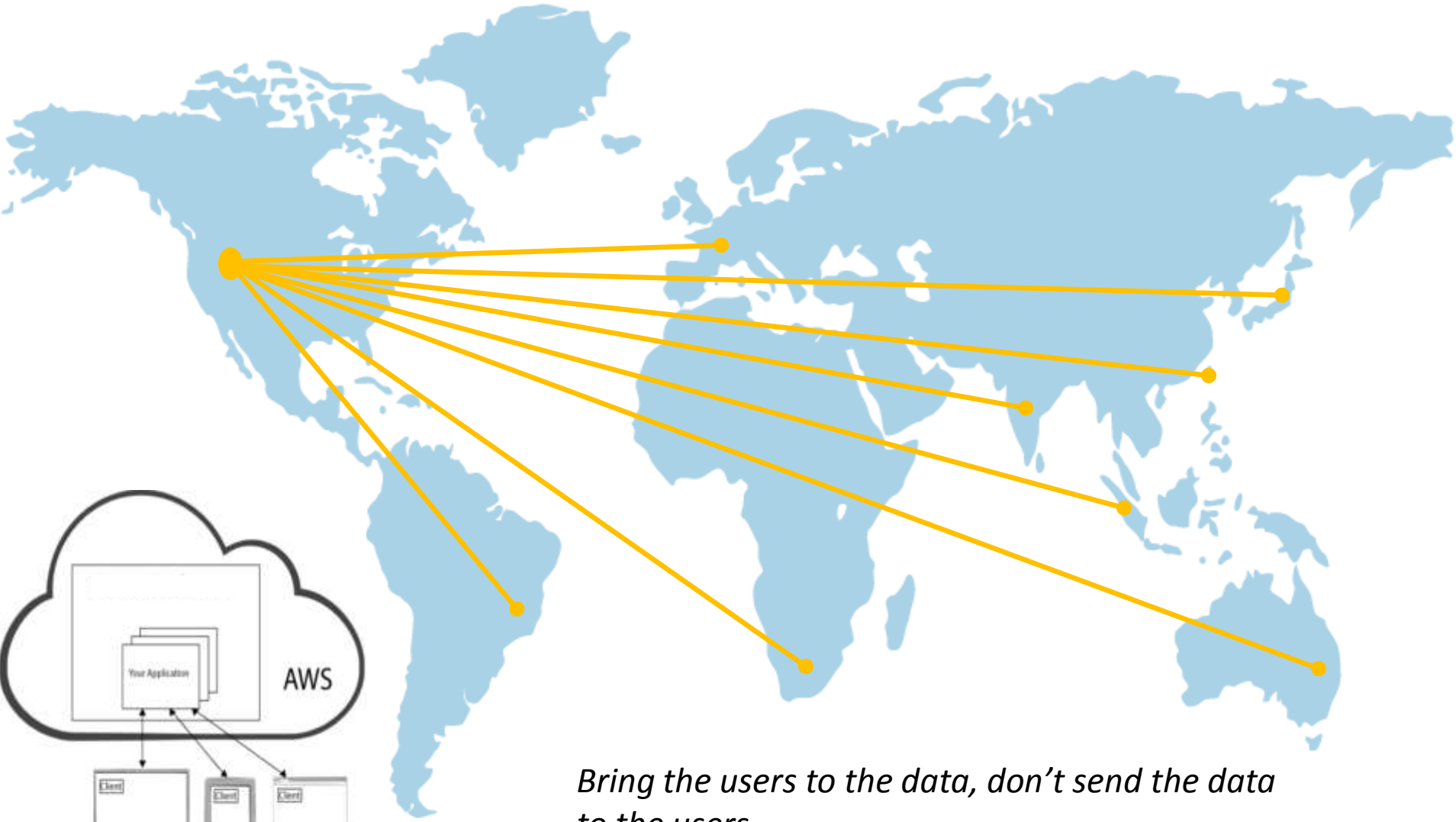


This could change the way manufacturing is architected.

Joe Salvo
Manager, Business Integration Technologies Laboratory

amazon
web services

Enabling Global Collaboration



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

amazon web services

Thin Client Remote Collaboration

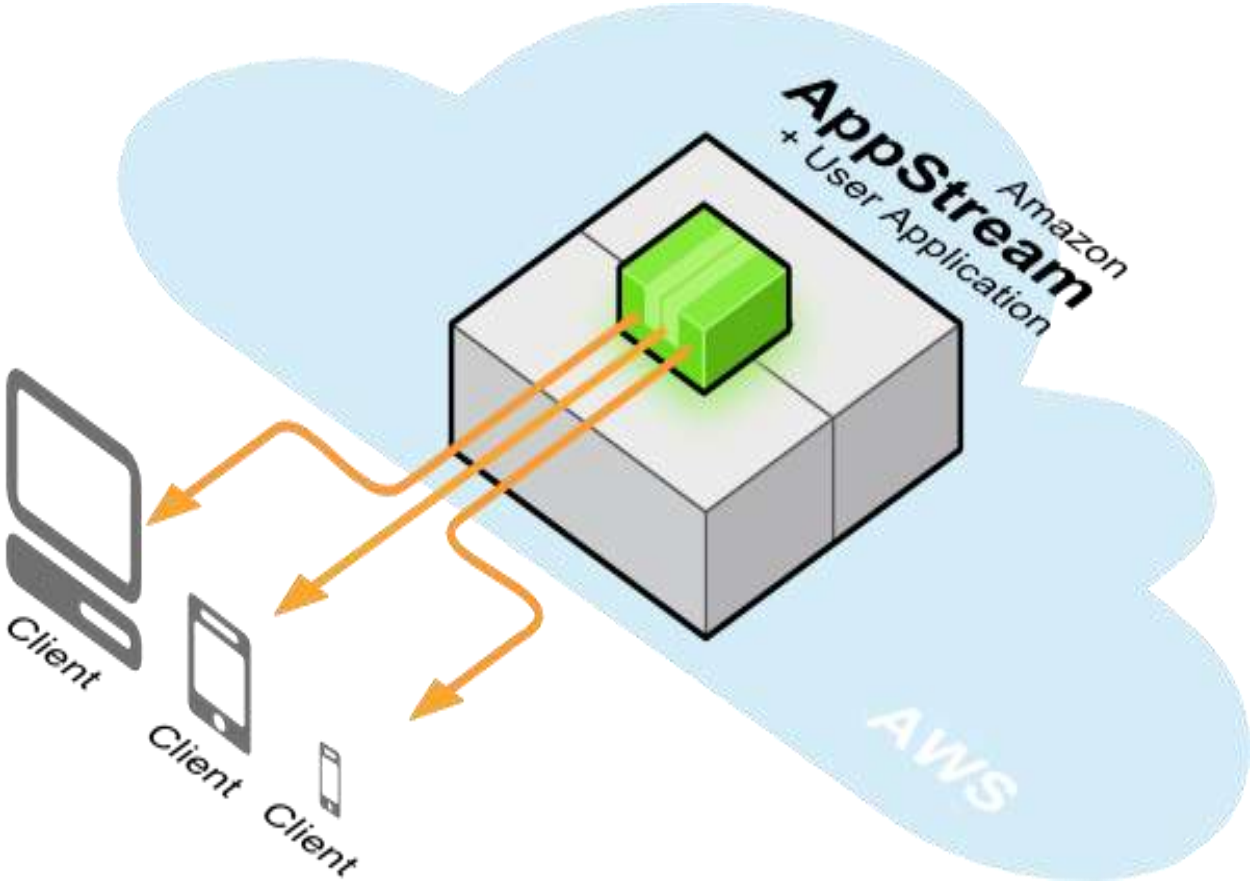
Calgary Scientific PureWeb™

www.calgaryscientific.com/resolutionmd/web/

demos.getpureweb.com/

Amazon AppStream

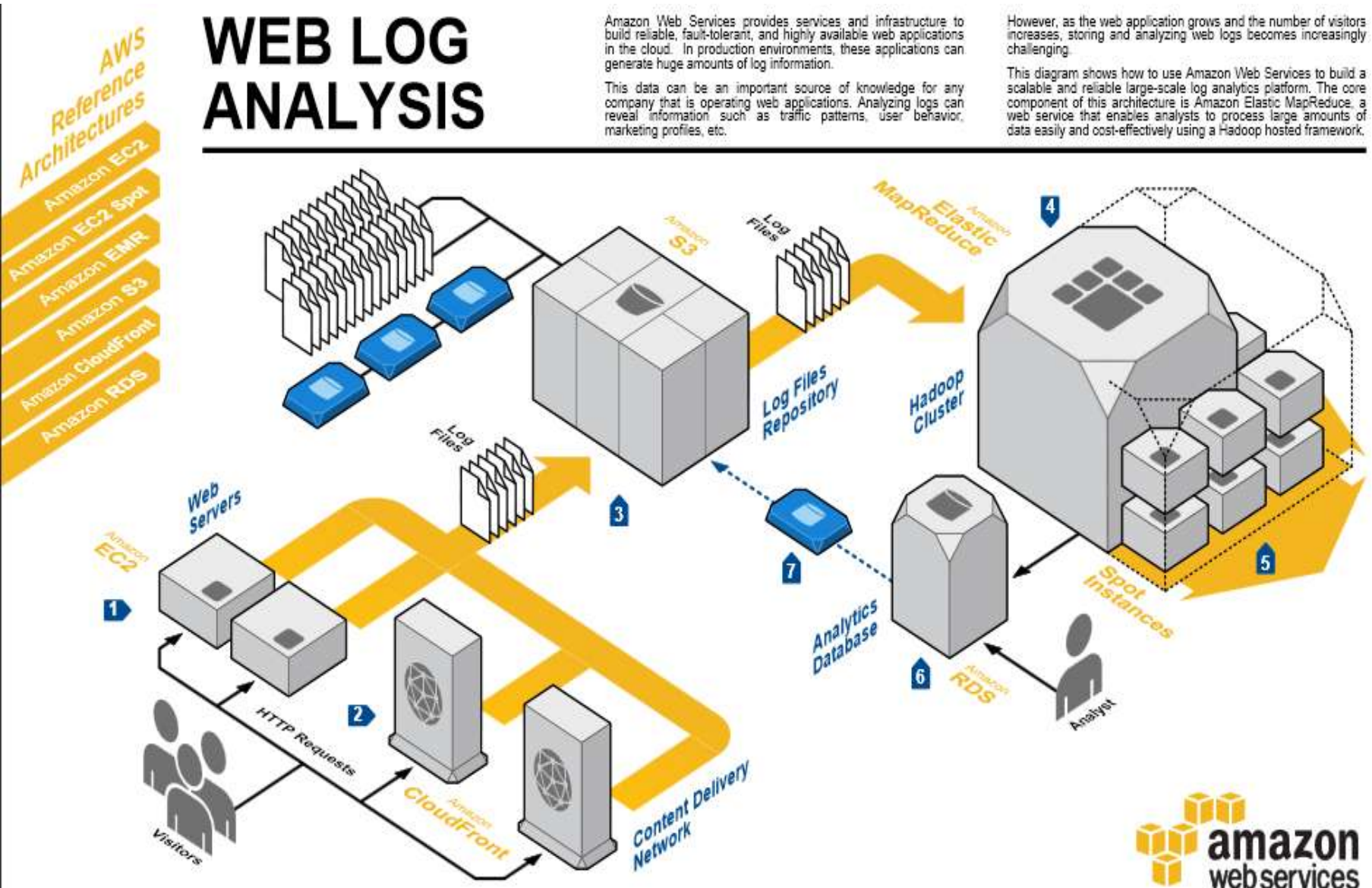
- Application Streaming
- Remote visualization
- Thin client 3D applications



amazon web services

Big Data
Plus Cloud
Equals Awesome

AWS Has Always Been About Big Data



Big Data for Financial Market Analysis

AWS Case Study: FINRA

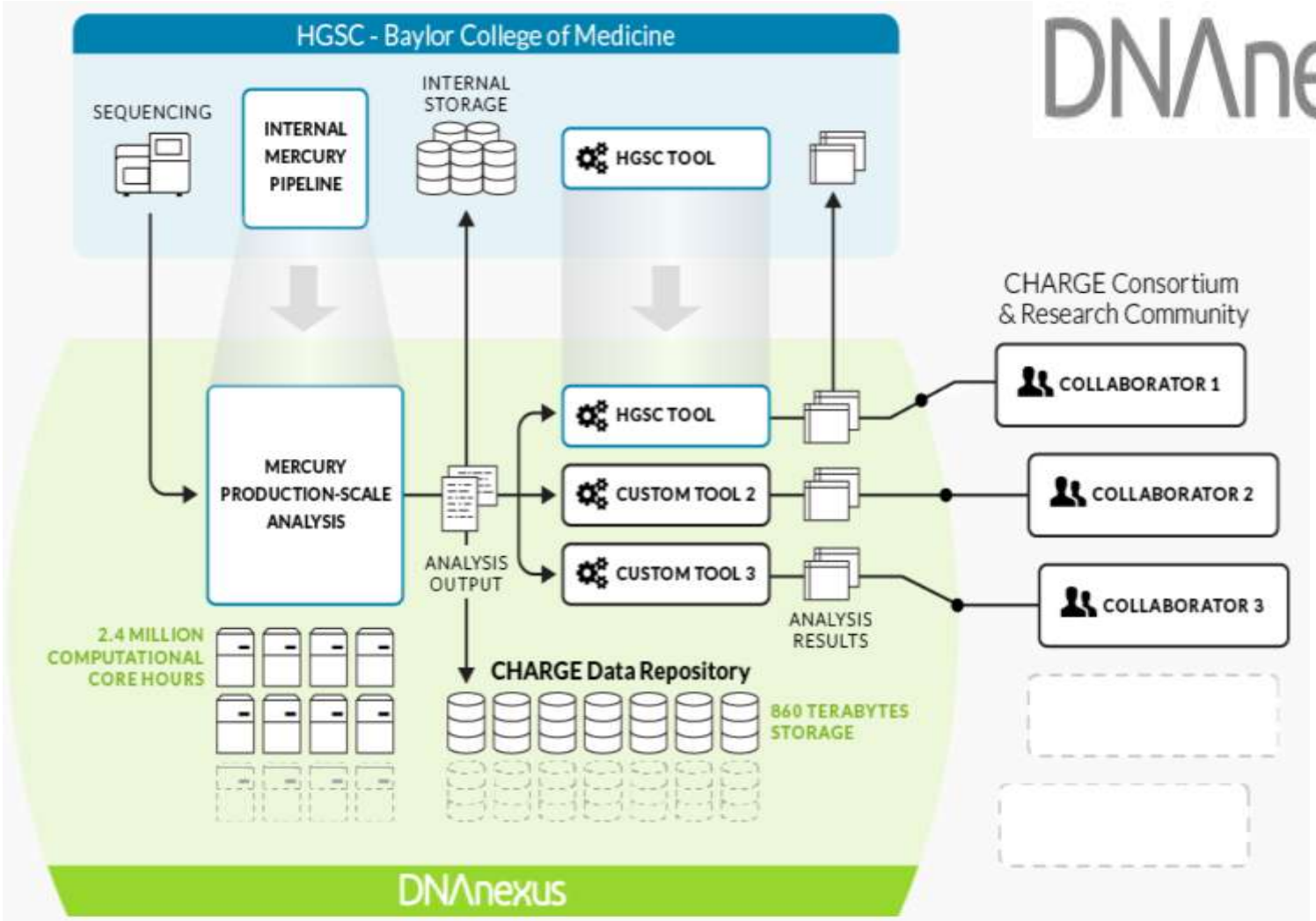
About FINRA

The Financial Industry Regulatory Authority (FINRA), one of the largest independent securities regulators in the U.S., was established to help watch and regulate financial trading practices. To respond to rapidly changing market dynamics, FINRA is moving its platform to Amazon Web Services (AWS) to analyze and store approximately 30 billion market events every day.

FINRA selected AWS because it offered the right services while fulfilling the company's security requirements. By using dynamic clusters (Hadoop, Hive, and HBase), and services such as Amazon Elastic MapReduce (Amazon EMR) and Amazon Simple Storage Service (Amazon S3), FINRA was able to create a flexible platform that can adapt to changing market dynamics. By using the AWS Cloud, FINRA has been able to increase agility, speed and cost savings while allowing them to operate at scale. The company estimates it will save $10 to $20 million annually by using AWS.

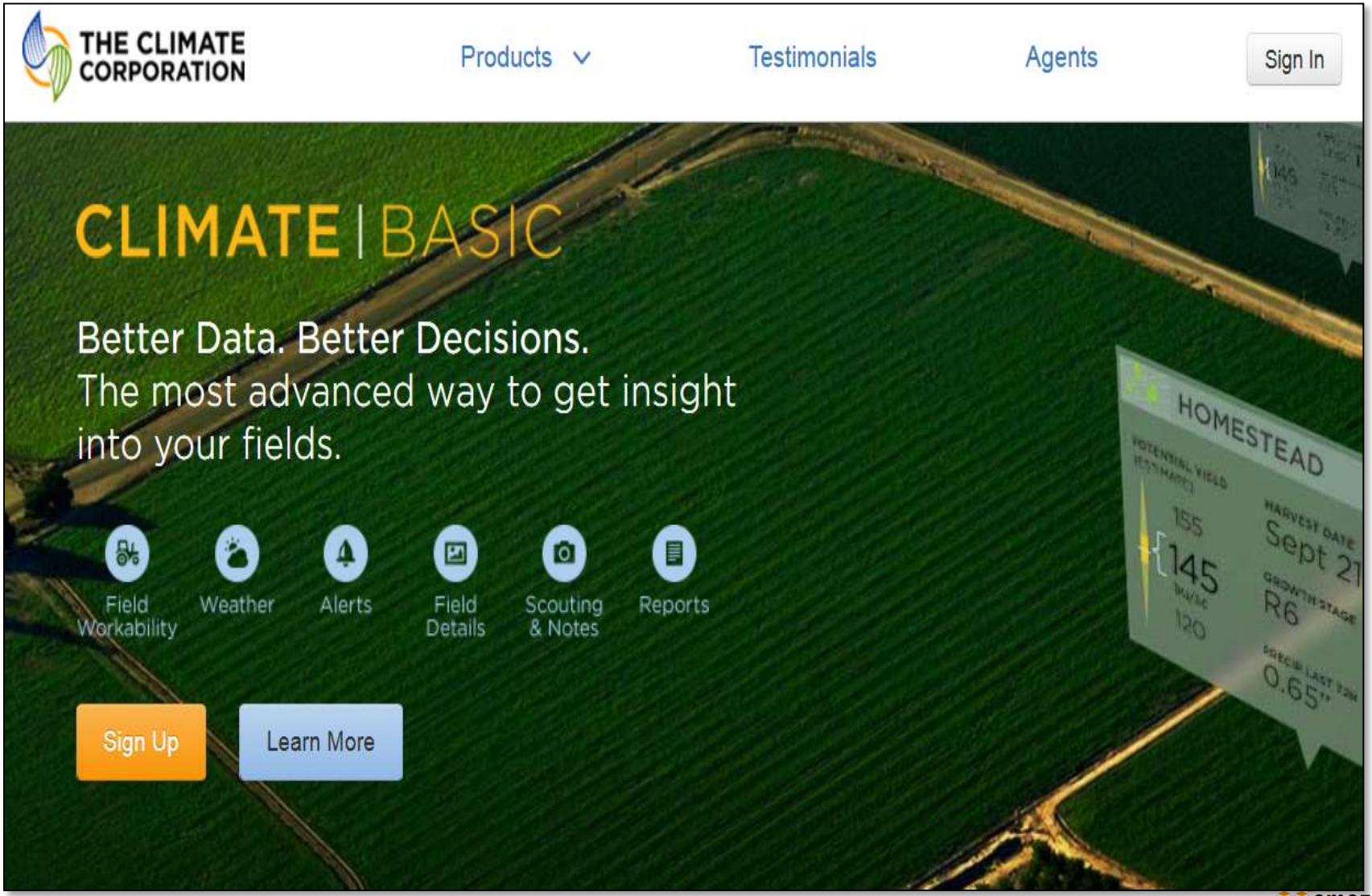
Big Data for Genome Analysis

Baylor College of Medicine, Amazon Web Services, and DNAnexus: cloud-based analysis of genomic data from over 14,000 patients



amazon web services

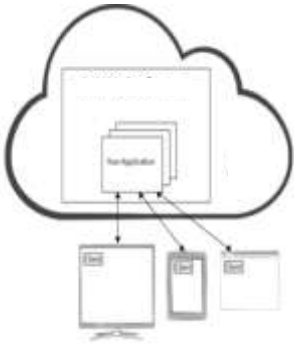
Big Data is Everywhere!



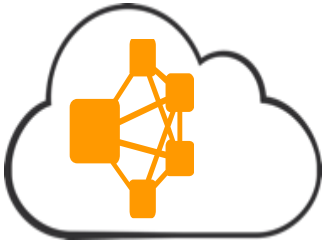
Summary

Cloud for Scalability

Cloud for Global Collaboration



Cloud for Big Data



Thank You