

GOJIRA

Worlds Largest Fortune 500 Cloud Cluster Run



When HGST decided to build their next generation of drives using helium they had hit on a way to dramatically increase drive capacity and simultaneously reduce power costs. In order to do this they needed to model airflow on a supercomputer. Trouble is they had just lost access to theirs due to the recent merger. So they turned to AWS.

Enter Gojira a 70,000 core Cloud supercomputing workload deployed on AWS. HGST leveraged AWS along with Cycle Compute and Trace3 to build a next-generation cloud-first high-performance compute platform to model optimal drive characteristics and deliver the leading drive on the market.

Like many organizations HGST built their initial environment utilizing the standard AWS offering - EC2 instances. This allowed for a very flexible deployment but resulted in a costly, sprawling, and fractured environment. After analyzing their workload, raw Amazon bills and performance data Trace3 rationalized it against customer usage models and implemented a new design. The improved design guarantees service-levels, reduces monthly costs by 30%, and introduces security controls to guard against IP theft. Gojira is the largest fortune-500 cloud cluster running today.

Now that's a monster anyone would think twice before challenging.



COOL FACTS

HGST is a pioneer in helium drive technology. Helium drives reduce power consumption by 23% and increase capacity by 40% per drive. This game-changer results in a 45% increase in efficiency per TB and took HGST over 6 years to develop.

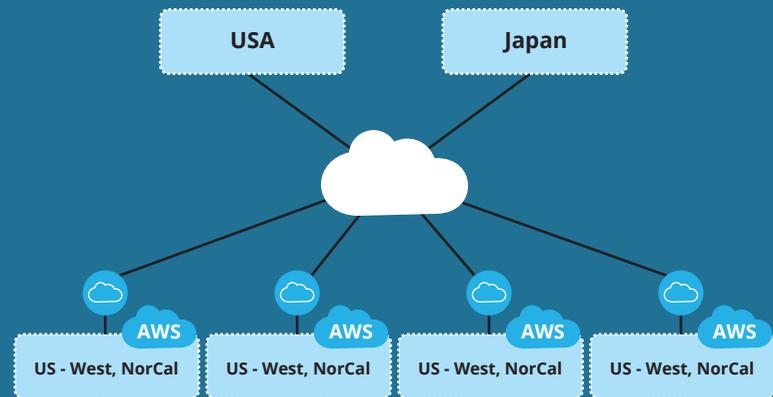
Gojira is a 70,980 core / 728.95 teraflop AWS Cloud HPC workload deployed to simulate the dynamics of air turbulence within a drive. One of the simulations HGST needed deployed would have cost millions to build in-house but the raw virtual compute on AWS cost only \$5,600 per run.

The design will be the model for all HGST builds on AWS and will guarantee production SLAs that meet requirements for production usage. The methodology co-developed with Trace3 and Cycle Compute will be adopted for future Public Cloud target environments such as Microsoft Azure and Google Cloud Platform.

SOLUTION HIGHLIGHTS

SERVICE LEVEL ASSURANCE

Trace3 engaged with HGST Network and Operations teams to develop a new VPC Network design that would allow for resilient SSL access as well as future use of AWS Direct Connect Services. The design model will be the core for all HGST AWS instances moving forward and will guarantee a level of service that meets corporate requirements for production usage.



COST CONTROL AND ANALYSIS

In addition to the use of AWS Reserved Instances Trace3 and HGST explored the use of SPOT instances for workloads that could tolerate a more dynamic compute environment. Spot Instances allow you to bid on spare Amazon EC2 instances and run them whenever your bid price is viable. Trace3 developed a financial model for HGST to determine the best mix of spot and reserved instances to maximize their return on investment for AWS compute instances.

CLOUD MOBILITY AND SECURITY

Trace3 developed a model for HGST on AWS that will guarantee production SLAs and meet other production requirements like security and access management. A centralized method to control identity access and integrate corporate controls on the public cloud was necessary. This methodology will be adopted for future Public Cloud environments such as Microsoft Azure and Google Cloud Platform for HGST moving forward.