



HPC Cloud Computing in High Tech Manufacturing

Steve Phillpott, CIO
steve.phillpott@hgst.com

David Hinz, Global Director, HPC/Cloud Solutions
david.hinz@hgst.com

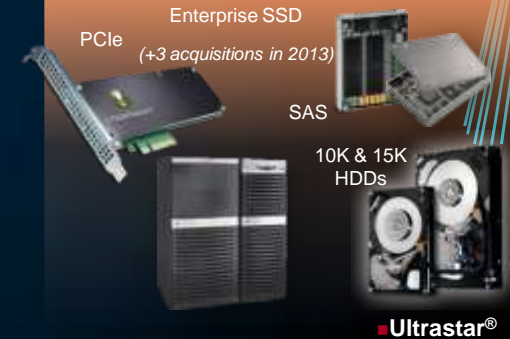


HGST History

- Founded in 2003 through the combination of the hard drive businesses of IBM, the inventor of the hard drive, and Hitachi, Ltd (“Hitachi”)
- Acquired by Western Digital in 2012
- More than 4,200 active worldwide patents
- Headquartered in San Jose, California
- Approximately 42,000 employees worldwide
- Develops innovative, advanced hard disk drives (HDD), enterprise-class solid state drives (SSD), external storage solutions and services
- Delivers intelligent storage devices that tightly integrate hardware and software to maximize solution performance

Cloud & Datacenter

Performance Enterprise



Capacity Enterprise



Ultrastar® Helium Platform
Enterprise Hard Drive

Zero to Cloud <<12 Months

April 2013

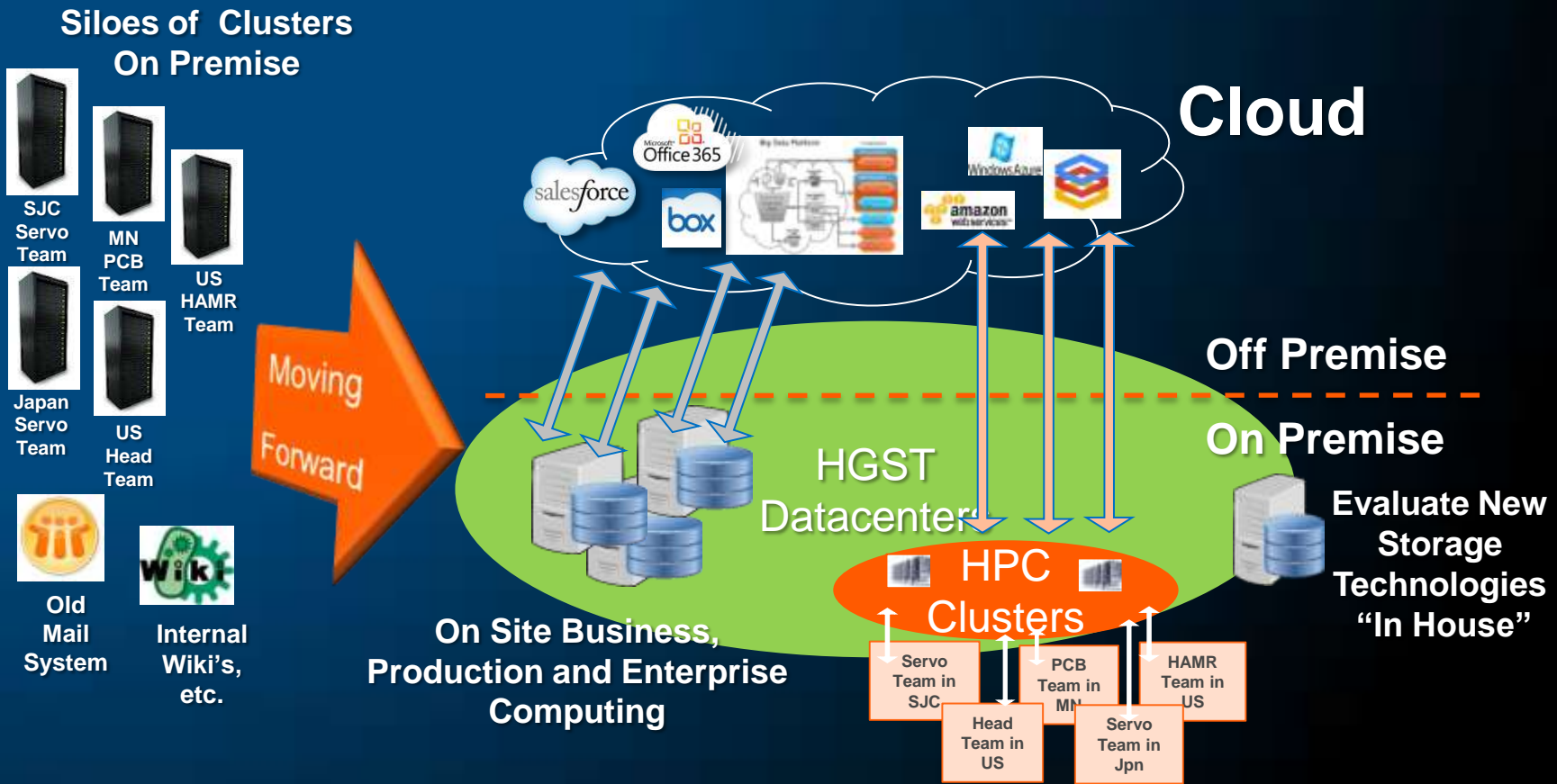


By 31 Mar 2014:

- ✓ Cloud eMail – Microsoft Office365
- ✓ Cloud eMail archiving/eDiscovery
- ✓ External Single SignOn (off VPN)
- ✓ Cloud File/Collaboration – BOX
- ✓ Salesforce.com
 - ✓ Integrated to save files in BOX
- ✓ Cloud–High Performance Computing (HPC) on Amazon’s AWS
- ✓ Cloud – Big Data Platform on AWS
 - ✓ Extract insights from manufacturing data
- ✓ Cloud - Data mart and provisioning service using AWS Red Shift
- ✓ Cloud – HR Transformation via Workday Deployment



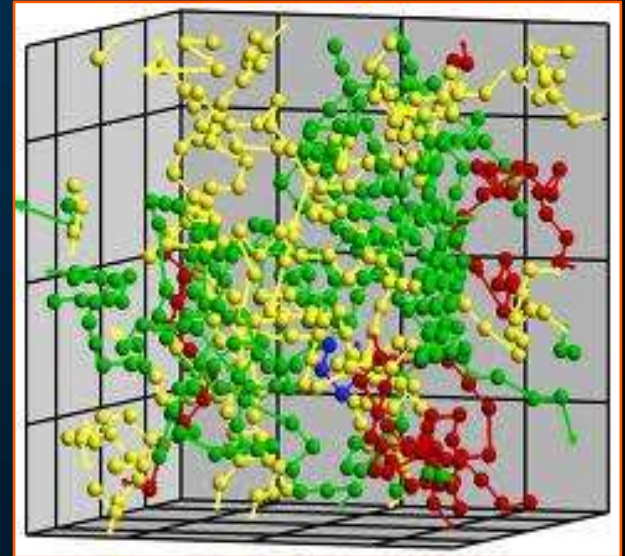
Evolution of HGST's Data Centers and HPC



An Agile Enterprise Datacenter Integrating On-Premise and Cloud HPC and Enterprise Solutions

Cloud HPC Use Case: Molecular Dynamics Simulation

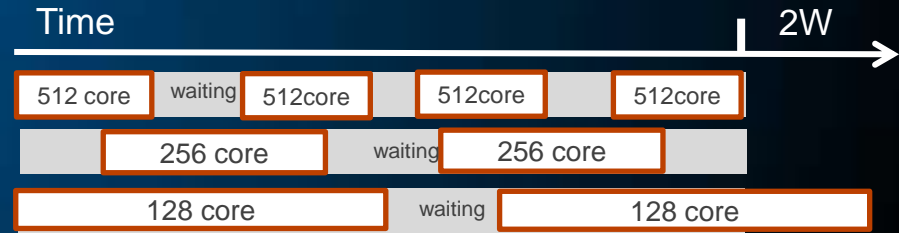
- HGST uses a custom in-house built Molecular Dynamics Simulation for RnD of materials and lubricants needed for HDD's
- Research to achieve higher memory densities, faster read/write capabilities, smaller form factors and lower power consumption



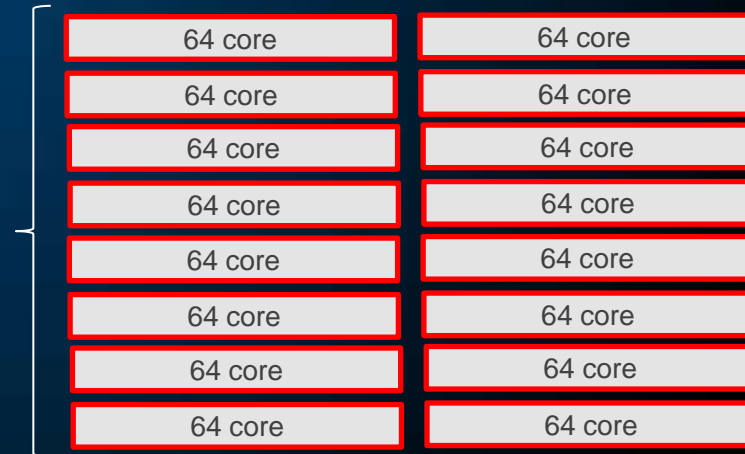
“Job Size” at HGST	Complexity [atoms]	Number of Time Steps	Job Type “Frequency”
Small	300,000	100	200 per day, 2 days per week once or twice a month
Medium	300,000	1000	20 Medium jobs during the day, 4 days per month
Large	300,000	30000	3 large jobs per day, 6 days per month
Very Large	300,000	3000000	1 large job per month

Lessons Learned: *Cloud HPC allows us to Shape Compute To Match Work To Be Done*

Before:
Shared
Super Computer



Today:
Amazon AWS EC2
CC2

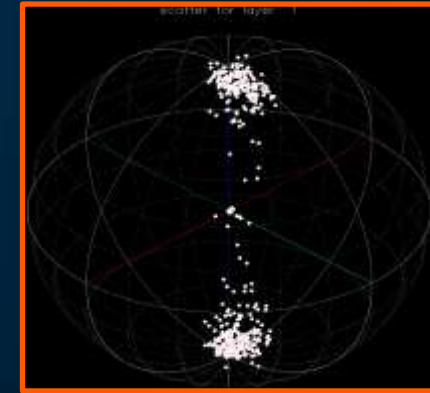


Shape compute to match work: running
parallel jobs across 64 core clusters
yielding 1.67x faster throughput

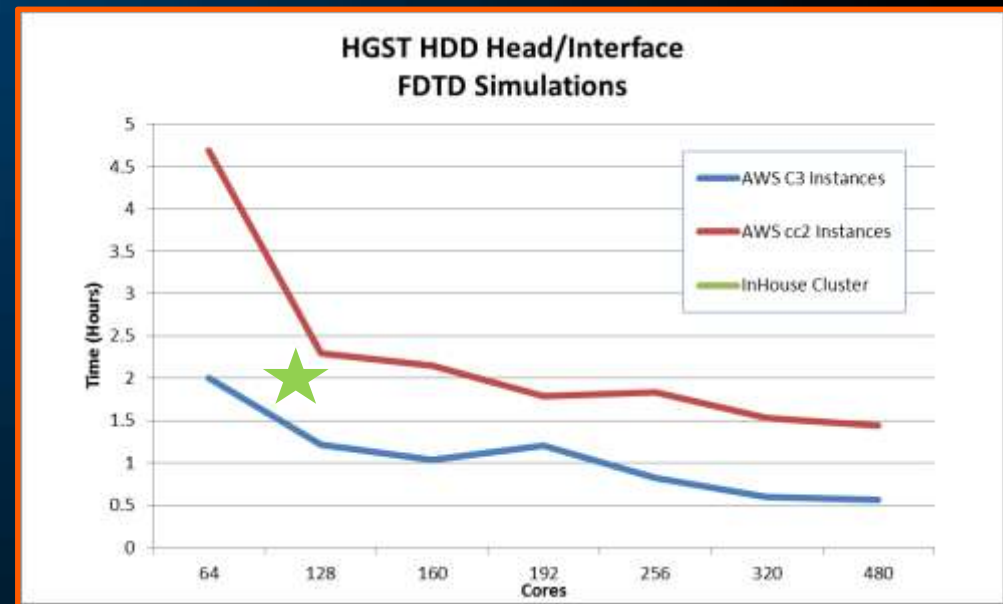


Lessons Learned: *Dynamic Compute Environment*

- RnD Teams model new technologies for future HGST HDD products
 - Lumerical Finite-difference time-domain (FDTD) solver simulates large, complex models of many variable parameters and materials while scaling across large clusters



- New AWS C3 Instances have significant improvement for single simulation performance (throughput) and scalability



On-premise solutions are “performance static” for 3-5 years
 HGST RnD groups will see AWS infrastructure “upgrades” every 12-18 months

Lessons Learned:

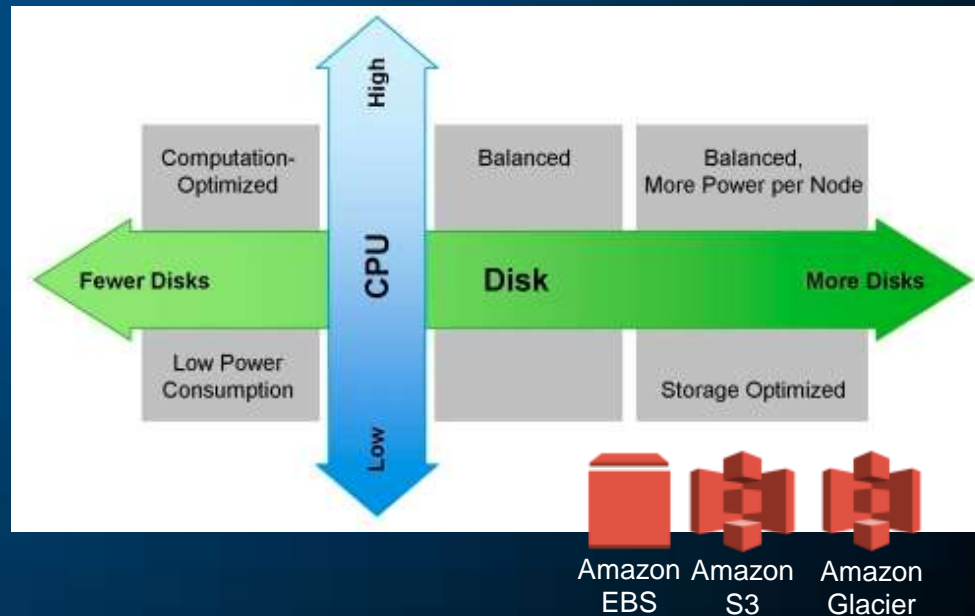
Finding the right balance of Compute Needs

Hadoop

MapReduce

Compute Bound Operations and Workloads

- Clustering/Classification
- Complex text mining
- Natural-language processing
- Feature extraction



Hadoop MapReduce I/O Bound Operations and Workloads

- Indexing and Grouping
- Data importing and exporting
- Data movement and transform

Cloud's Large Variety of Compute and I/O Operations and Storage allows for Optimization/Tuning

HPC Computing: Evolution @ HGST

- Lessons Learned
 - Ramping up new technology quickly
 - » Use cloud to “road test” and Proof of concept a variety of cluster sizes and options for optimal \$/performance
 - Burstable capability to manage “crunch times” before holidays or “catch up” over weekends
 - Cloud provides flexibility to move locations

Challenge/Opportunities: Are Commercial HPC Applications Cloud Ready?

- “Hybrid” HPC Model

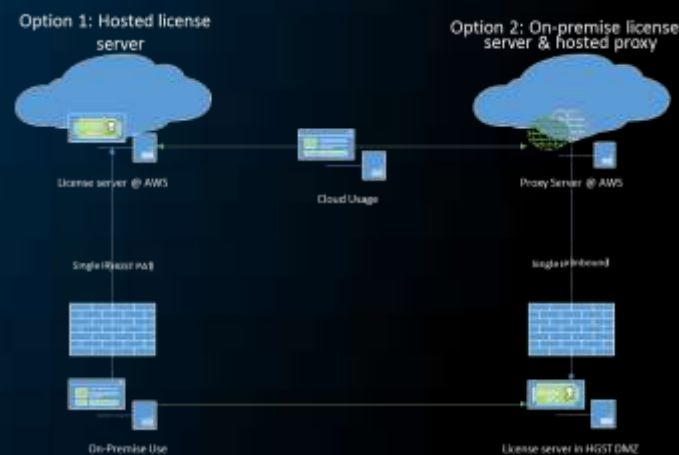
- On-premise with mix of burst/ over-flow / direct queue connection to Cloud

- EULAs

- Comprehend concurrent on-premises and Cloud
- Allow single license server for both on-premise and cloud compute

- Pricing Models

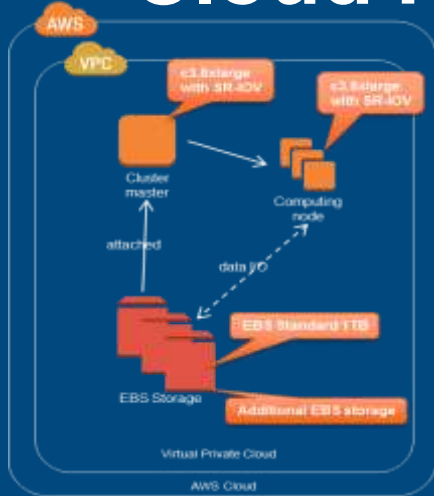
- No consistency across vendors
- Most not aligned with time based consumption pricing of cloud services



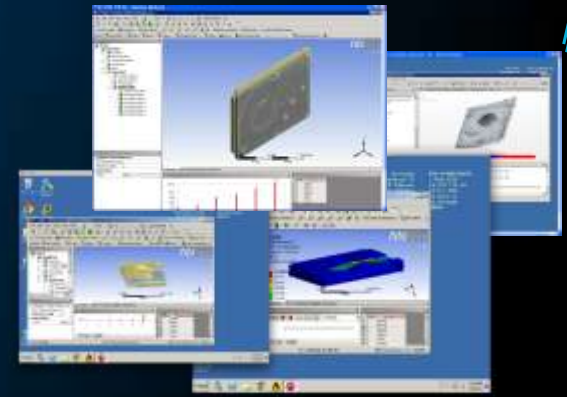
Additional Challenges

- *One Application in the Cloud may not be enough*
 - Multiple applications in user workflow
 - Data management and data access key for user efficiency
- *Don't underestimate the effort for Change Management*
 - Non technical issues can be more difficult to solve than technical blockers

Cloud HPC: What's Next For HGST



- Utilize Latest Compute Capability
 - Deploy in AWS C3 Environments to Improve overall performance and reduce monthly AWS compute bill



- Deploy GUI HPC Applications for pre and post simulation work in Cloud
 - Use Remote Visualization to minimize data migration from/to local systems
 - Collaboration across multiple sites
- Applications Being Evaluated Also Include:
 - Modeling applications to simulate HDD arm movement and settling times
 - Computation Fluid Dynamics for modeling of airflow inside HDD
 - GPU computing to significantly accelerate simulation speeds for advanced HDD design work

“We’ve Only Just Begun....”

- Current Results < 12 months
- Re-aligning Business Group Leadership, Development Teams, Research and Development Teams on New Capabilities Mode
- Demands and Uses Expected To Grow And Accelerate Market Success
- Cloud Compute Model is Maturing at a rapid pace

2013: Set the “Cloud Foundation”
2014: Acceleration And Expand Use

Q&A