Gaining a Competitive Edge with HPC Cloud Bursting

William J Edsall
The Dow Chemical Company
R&D Information Research – High Performance Computing Team
September 5th 2018
About the presenters

William J. Edsall – The Dow Chemical Company
High Performance Computing Architect
- BA Applied Mathematics
- 11 years HPC experience, 15 years IT
- 3 kids ages 0, 2 and 4
- Self-proclaimed “ex cloud critic”
About Dow High Performance Computing

On premise HPC capabilities
- Two Intel CPU based HPC Clusters
- >85% AVG utilization
- Nvidia K80 and V100 coprocessors
- High end HPC desktops

HPC workload at Dow
- Computational Chemistry
- Computational Fluid Dynamics
- Molecular Dynamics
- Computer Aided Engineering
- Finite Element Analysis
- Optimization
- Deep Learning / Machine Learning
- X-ray Tomography

Executive Summary

Proof of Concept with Rescale and Microsoft Azure: 3/1/2018 - 7/1/2018

- ISV Licensing Model: 7/1/2018 - 3/1/2019
- Data Transfer (Dow to Azure): 7/1/2018 - 3/1/2019

Users Trained: 6/29/2019

Success: Working On Premise / Cloud Hybrid Environment: 7/7/2019

Today: 7/7/2019

Mar May Jul Sep Nov 2019 Mar May Jul
Cloud Bursting Opportunity

- **CPU Hours Demand**
  - A few CPU
  - 100s of CPU
  - 1,000s of CPU

- **Urgency**
  - No Urgency
  - Time Sensitive

- **Special Cases**
  - Extreme memory, massive GPU farms and massive parallel jobs (we don’t have resources to solve on premise)

- **Computational Chemistry Today**
- **Computational Chemistry Tomorrow**

- **Resources**
  - Dow Enterprise Workstation
  - High-End Desktop
  - HPC Cluster
  - Cloud Bursting Opportunity
Challenges with on premise HPC Model

- True demand difficult/impossible to capture
- Utilization always follows capacity trend
Challenges With HPC Cloud Bursting Model

- HPC Cloud Data Security Model
- Cost Recovery Model
- Azure Active Directory Integration
- Independent Software Vendor (ISV) Licensing
- Data Transfer (Dow to/from Azure)
Rescale Cloud Bursting Proof of Concept at Dow

~2x the total CPUs on premise at Dow

**Computational Chemistry**
- 1,400 molecules from ChEMBL database
- Each molecule calculated with 7 CPU using Azure H16r instance
- Entire ~9,800 CPU cluster running in <30 minutes
  - 0 queue time (no queue exists)
  - Cluster was shut down as jobs finished
  - Charges stopped automatically
- **Goal met in 5 hours**

Work performed by Dr. Peter Margl, The Dow Chemical Company

~2x larger than the biggest parallel job at Dow

**Computational Fluid Dynamics**
- Ansys FLUENT simulation
- Scaled up 20 -> 200 -> 2,000 CPUs
- Entire 2,000 CPU run on Azure H16r instances
  - One Infiniband fabric
- Cluster was stood up – job was verified and then stopped
- **Goal met in 15 minutes**

Work performed by Dr. Chang Kai (Lance) Wu, The Dow Chemical Company
Scale Beyond our POC

Massive Scale Example: Univa (June 27th 2018)

• “Univa leveraged AWS to deploy 1,015,022 cores in a single Univa Grid Engine cluster”
• “The cluster was built in approximately 2.5 hours”

Reference Link: