Azure HPC Integration

Final Project Report for Phase 1: Pilot Project

CREATED FOR

XYZ (anonymized)
Address

1 Executive Summary

UberCloud has been helping customer XYZ to evaluate and implement engineering simulations on Microsoft Azure. During this project, which started on October 1, 2018, UberCloud project team worked with XYZ IT and subject matter experts to design and configure XYZ's Azure environment to support running ANSYS software in an HPC configuration, test XYZ simulation benchmarks, understand engineering simulation usage of HPC resources, and related workflows.

2 Project Activities

2.1 Activity: Personnel On-Boarding and Test Environment Setup

Project activities started with the on-line kick-off meeting on October 3, 2018. Project teams were established and project was reviewed in detail.

Project Teams:

- XYZ
  - IT Lead: …
  - Technical Lead (ANSYS+HPC): …

  - Technical Expert (Rocky): …
  - Global Licensing: …
- UberCloud
  - Project Manager: Reha Senturk
  - Technical Lead:
  - Lead Developer:
  - Test Engineer:
  - Business Lead:

Additionally, on October 22, UberCloud Project Manager had an on-site face-to-face meeting with key XYZ team to map XYZ IT and in-house engineering workflows.

For active team collaboration and efficient project management, weekly project team meetings were scheduled (Tuesdays at 10:00 AM CET) and an HPC POC SharePoint Site was created to keep project files like architecture documents, benchmark files, meeting notes, and the project plan.

Deliverables

- Project schedule including implementation plan was established and updated through the project (Appendix 1).

2.2 Activity: Execution of Pilot Project

UberCloud worked with XYZ Team to define the requirements for HPC cluster. Based on the information a reference HPC network topology was established and deployed on Azure.

HPC Cluster in Azure

XYZ IT Team created the pilot cluster in XYZ's Azure subscription, and established access to UberCloud team. Pilot cluster has IP based firewall security to allow proper testing and access. HPC cluster will be moved to XYZ subnet when it is rolled out to production.

Based on XYZ project team's decision, West Europe and Central India data centers were identified as POC sites. Azure West Europe Data Center is the main site for XYZ Data Center migration. India was chosen as the second site to provide HPC resources that eliminate their latency issues.

Based on XYZ IT requirements and engineering usage needs, UberCloud created a flexible global HPC Cluster Architecture (Appendix 2). XYZ IT naming standards were used to properly identify the resources. Resource groups were created as ams-

vnet-rg, 1 TB centralized storage account (ams-fileshare) was created for storing test cases and shared files, nodes were created as head node (ams-hn), compute nodes (ams-cnXY), visualization nodes (ams-vnXY), container registry was created as ams-registry-rg resource group, license server was created as ams-license. Resources in India cluster were created with pnq2 prefixes.

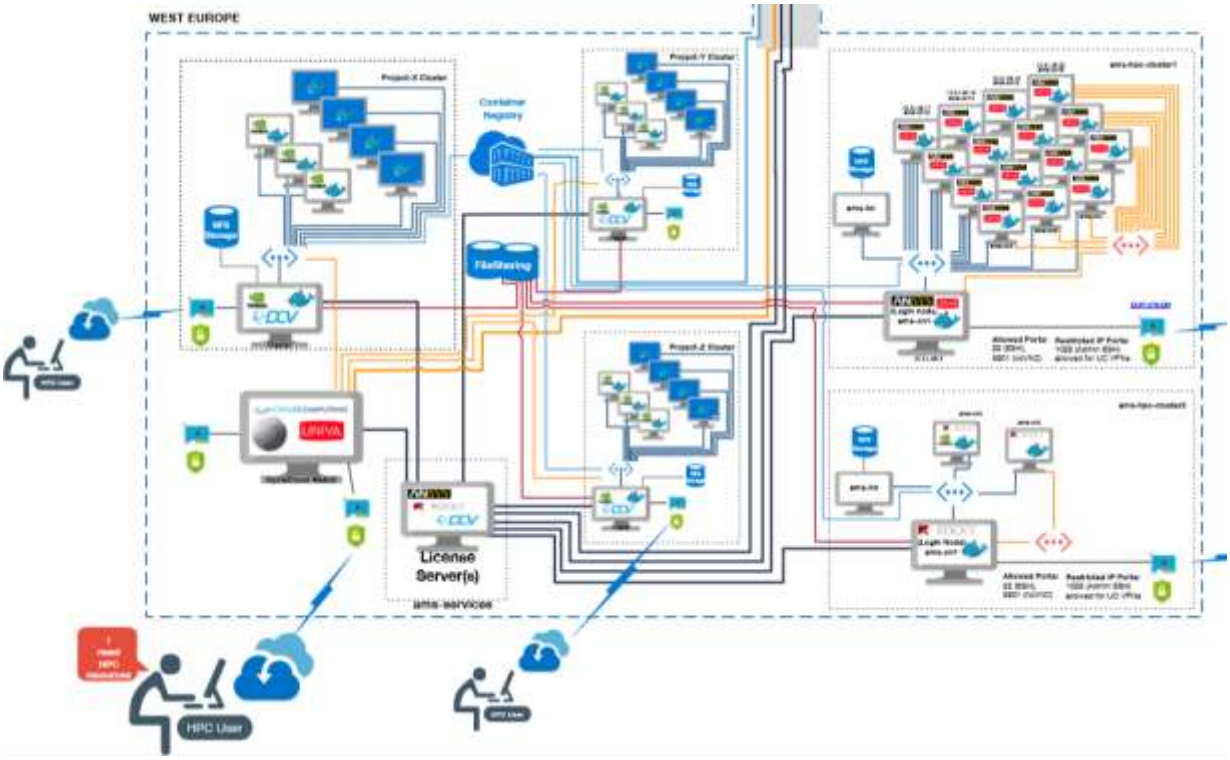

Figure: XYZ Cloud HPC Architecture (West Europe)

Azure provides globally accessible and flexible HPC resources. H Series high-performance nodes were chosen for CPU computations, N Series GPU nodes were chosen for GPU computations and visualization purposes.

Application and Technology Testing

A container registry was created in XYZ's Azure HPC environment and UberCloud ANSYS Container images (ANSYS19 and ANSYS19.2 with VNC and GPU versions) were stored. Containers were deployed on the XYZ's Pilot cluster. As a part of this pilot project, enabling tools and devices like Nice DCV, GPU and IB/RDMA were configured and successfully tested.

The test clusters were created for XYZ Engineers to perform the benchmark tests using ANSYS Mechanical, Fluent, ANSYS CFX. UberCloud assisted XYZ engineers during the tests by setting up license servers, providing CLI and UI support, setting

up RSM, troubleshooting computation errors, and resolving Azure resource issues. Benchmark tests were successfully run and results were reported (Appendix 3).

Performance

ANSYS Fluent, CFX, and Mechanical Benchmark tests were performed using H16 nodes in Azure which have the same CPUs as XYZ's on-premise HPC cluster. Single node runs brought very close results (Azure was 1-3% slower).
Multinode runs showed good scalability when using many compute nodes for large simulations, performance on Azure was 10-30% slower due to InfiniBand speed for interconnect (customer's ID has 56 Gb FDR speed vs H16r's have 20 Gb FDR). Ability to scale up to large number of resources brings ability to run multiple jobs simultaneously reducing multi job duration and significantly increasing HPC throughput for engineers. Benchmark results and details are provided in Appendix 3.

CycleCloud and Univa Grid Engine (UGE)

Microsoft CycleCloud offers an easy to use, web-based cluster provisioning and management interface for engineers to create their own resources on-demand. CycleCloud's policy-based templates allow administrators to set up the rules and usage limitations to control HPC utilization of different groups.

Univa Grid Engine provides a powerful multi-cloud workload management platform. Combined with CycleCloud, on-demand clusters can be created with built in workload management capabilities. This combination delivers self-service HPC capability to engineers to create and control their own HPC resources which should reduce IT dependency of HPC resources.

Licenses

- ***ANSYS:*** UberCloud provided test ANSYS licenses during the project. XYZ engineers were able to perform benchmark tests to compare Azure HPC performance to XYZ's on-premise cluster, experience usability of cloud CAE workflow. In production environment, XYZ will be able to use existing site licenses in the cloud. UberCloud ANSYS Containers can connect and use local license servers.
- ***NiceDCV***: UberCloud provided test Nice DCV licenses during the project. XYZ recently upgraded Nice DCV licenses to 10 concurrent users. This should cover the main usage for ANSYS engineers, however, additional 3-4 Nice DCV licenses would be needed for Rocky users.
- ***Docker CE***: Docker is the industry standard open-source containerization platform for software applications. UberCloud's unique HPC containers use

Docker platform. Docker Community Edition's open source components are licensed under the Apache 2.0 license.

- ***Microsoft CycleCloud***: CycleCloud software is provided free with Azure Enterprise subscription.
- ***Univa Grid Engine***: UberCloud provided test UGE licenses during the project. License costs for expanded roll-out is included in the cost estimate.
- ***UberCloud HPC Containers***: License for UberCloud's software container technology was also provided for duration of this project. XYZ can extend the use of UberCloud's software container technology beyond the project period by paying the corresponding fee for the number of copies and the number of users.

IT and Engineering Processes

During this pilot phase of HPC migration to Azure cloud, XYZ's existing processes were evaluated and XYZ Azure HPC Operations and Enterprise Roll-Out Map Document was created (Appendix 4).

HPC Utilization and Azure Consumption

ANSYS and most other simulation tools depend on CPU usage. Based on 40% average utilization of customer's on-premise 512 core HPC cluster, a reserved cluster of 256 CPU cores for queuing daily jobs, and on-demand clusters for projects and larger usages were designed to provide the best performance in most cost-effective way.

GPU based Azure nodes allow engineers to perform their full simulation workflow with high resolution graphics for pre and post-processing while running computations on the CPU clusters. 6 GPU visualization nodes (NV series) were reserved to provide high quality interface for this usage. Additional visualization nodes are also available for on-demand short term usages.

High-end GPU nodes are also utilized for computing in newer simulation applications like Rocky. 4 GPU compute nodes (NC series) were reserved for Rocky simulations. Additional nodes are also available for on-demand short term usages.

Based on in-house cluster utilization and Pilot phase Azure HPC utilization, UberCloud and XYZ teams created XYZ Cloud HPC Usage Projection for 2020 (Appendix 5). This projection provides a basis for enterprise roll-out cost estimation.

Deliverables

- UberCloud ANSYS Containers with FEA and CFD applications were deployed to XYZ's Azure subscription. Benchmark results provided by XYZ engineers are provided in Appendix 3.
- XYZ Azure HPC Operations and Enterprise Roll-Out Map Document (Appendix 4).
- XYZ Cloud HPC Usage Projection for 2020 (Appendix 5).
- Using XYZ Simulation Applications with UberCloud Software Containers on Azure HPC Environment (Appendix 6) was created and shared with XYZ project team. Engineers and XYZ's IT team was trained on the use of the UberCloud containers.

3 Conclusion and Next Steps

During Phase 1 project, XYZ's engineering applications (ANSYS Mechanical, Fluent, CFX, and ESSS Rocky) were implemented in Azure HPC and tested in West Europe and Central India data centers. Cloud HPC will allow XYZ engineers to use simulation applications and vast resources on demand allowing faster results and turn around time while providing complete control and management to IT. Major outcomes of this project are summarized below:

- Azure provides a variety of HPC compute resources for different application needs of XYZ engineers. New compute resources are regularly added, providing faster and cheaper resources for XYZ.
- Azure HPC resources are available in many data centers around the world, providing low-latency regional access for XYZ sites.
- UberCloud provides complete engineering workflow capability in cloud including full graphical user interface for pre and post-processing, and batch.
- UberCloud HPC containers provide portability to XYZ's engineering applications (including complete and complex multi-physics workloads), allow multi-location and multi-cloud usability to application workloads.
- UberCloud HPC containers allow engineers to collaborate and increase productivity by working on the same simulation from any location.
- UberCloud HPC containers provide a standard HPC Application Management for all HPC applications including ability to containerize other simulation applications including in-house codes.
- Cloud access allows worldwide usage of XYZ's applications from all sites
- Cloud HPC provides almost unlimited number of resources giving XYZ engineers freedom to innovate. They can use the resources to speed-up their

simulations, run multiple iterations, models, (multi-) physics, perform parametric sweeps and experiments.
- Cloud HPC allows global access to software licenses, reducing the number of local licenses.
- Cloud HPC eliminates the hardware bottlenecks during peak usage and fluctuating project loads.

Microsoft Azure with UberCloud Containers provides the modern Cloud HPC environment for XYZ. During Phase 1, XYZ's IT and Engineering workflows were evaluated and the following operational tasks were identified as critical processes to be developed before enterprise roll-out.

- Engineering Workflow and Usage Optimizations: Cloud HPC usage requires better planning of the simulations, results and data storage, job-based vs interactive usage and processing. UberCloud and XYZ teams started creating the list of Engineering Best Practices in Cloud for XYZ Engineers and documented in XYZ Azure HPC Operations and Enterprise Roll-Out Map for next Phase 2 project (Appendix 4).
- Consumption monitoring and billing: User/Group based consumption should be monitored to ensure proper use of cloud resources and minimize idle resources.
- Simulations generate large amount of data. XYZ HPC Architecture is designed with regional storage and also replication for data that needs global access. Storage, archival and cleanup policy must be developed in alignment with XYZ Data Center Migration project.

As the result of Phase 1, HPC Cloud Architecture was built and tested successfully. Enterprise Roll-Out map was also created identifying the next steps necessary to establish steady-state operation. Phase 2 project will focus on the Enterprise Roll-Out tasks and help XYZ complete cloud migration of all its HPC workloads.

4 Appendix

1. Project Plan
2. XYZ Cloud HPC Architecture
3. ANSYS Benchmark Results
4. XYZ Azure HPC Operations and Enterprise Roll-Out Map for Phase 2 project
5. XYZ Cloud HPC Usage Projection for 2020
6. Using XYZ Simulation Applications with UberCloud Software Containers on Azure HPC Environment