Running CMAQ on Microsoft's Azure Cloud

UNC – IE Team: Christos Efstathiou, Liz Adams, Carlie Coats, Sarav Arunachalam

UNC-ITS Team: Rob Zelt, Mark Reed, John McGee

EPA Team & Microsoft "Blackbelts"

"Ninjas"

6/16/2022

CMAQ on Azure: The Big Picture

Goal

- Demonstrate our efforts with Azure's HPC solutions, in continuation of our experience with AWS and UNC ITS HPC infrastructure

- Documented through a series of tutorials available to the community:

  https://cyclecloud-cmaq.readthedocs.io/en/latest/index.html

Structure

➤ **Background & Resources:** Serve as a portal to provide a quick introduction to Azure basics, and share useful links for newcomers

 (See Chapters: *1. Introductory Tutorial & 13. Additional Resources*)

➤ **Model Configuration & Simulation Design:** Having as a baseline a 2-day CMAQ 5.3.3 simulation for a US CONUS 12-Km resolution domain
  - Note: Testing phase with the 2-day SouthEast US domain (less demanding) on single node Virtual Machines (VMs) highlighted performance differences between the two main architectures HC44rs (Intel Xeons) and HB120v2-3 (AMD EPYC) offered
  - System Requirements and suggestions discussed in *Chapter 2*

➤ **Core of the tutorial is broken into two branches:**
  - ➤ Deployment on single VMs (*3. Intermediate Tutorial*) and
  - ➤ Setting up multi-node CycleCloud instances (*4. Advanced Tutorial*)

➤ **Post-Process workflow,** prerequisites (R & packages) and QA discussed separately (*Chapters 5-9*) – to be consolidated

➤ **Clean-up workflow,** including moving the output to an S3 bucket (*Chapter 10*), logging out and deleting the CycleCloud instance (*Chapter 11*)

➤ **Benchmarking and Performance Optimization** in *Chapter 12*

➤ **Serve as a repository** for useful information and findings: Additional resources, future work and ways for others to contribute to this tutorial (*Chapters 13-15*)
Issues · CMASCenter/cyclecloud-cmaq · GitHub

UNC - Institute for the Environment

Notes on accounts and hardware availability

- The Azure Portal controls access to VMs using quota limits.
    - For a pay-as-you-go account, the quota limits are by default set to zero for VMs.
    - The beginner tutorial provides instructions to create an account and request a quota limit increase.
    - This process may take a week or so and requires a $29.99/month service subscription account.
- The level of access to virtual machines for accounts will depend on the type of support agreement that your organization subscribes to, and limits may be unique to agreements made for each enterprise account https://azure.microsoft.com/en-us/support/plans/
- Using the UNC Subscription, we did not meet such hard quota limits
- However, certain roles must be assigned, and permissions need to be granted to users within institutional accounts
- Availability of certain HPC clusters and cost can be fluctuating and is region-specific
- Cloud services are continuously evolving, so some things can temporarily break

Current Status with single VM efforts

- Explored CentOS and Ubuntu distributions successfully – current recommendation from Microsoft is AlmaLinux which is HPC targeted but the post process is different (R install mods)
- Created VM images of the single-node setup for the two HPC hardware solutions, shared in our internal Azure Gallery
- Used successfully by the UNC team – alleviated the need for setting up everything, including the storage provisioning and linking steps
- Given access to share publicly through a Public Compute Gallery and exploring a few details before sharing (license agreement in our page)
- Minor fixes to be streamlined in the cloud_init scripts & attach the Post and QA process steps
- Target a release – looking for beta testers, tell us more about you

Current Status with the CycleCloud efforts

- Currently working on deriving a CycleCloud template to be shared with others
- Adding "safety" features to the SLURM jobs
  - i.e., Cluster shuts down completely after a certain period of inactivity to deal with potential charges from that runs that crashed but do not exit the queue properly
- Streamlining the Post and QA steps adds another level of complexity and OS-specific differences
- Exploring additional CycleCloud template options adding spot instances

Azure vs AWS

- Account administration and hardware/service provisioning is significantly different: Azure is quite restrictive with initial quota limits on a singe-user account compared to AWS (evolving)

- CycleCloud requires more compute resources than the Parallel Cluster

    Parallel Cluster requires

    - VM for the head node (always running)
    - VMs for the compute nodes

    CycleCloud requires

    - VM for hosting the application, (always running)
    - VM to run the SLURM scheduler (always running)
    - VMs for the HPC queue (compute nodes - running on demand)
    - VMs for the HTC queue (post-process -running on demand nodes)

- The VMs for hosting the CycleCloud Application and the Scheduler node are always running. We can terminate and restart the scheduler node to avoid accumulating costs, but may not be able to stop and restart the application VM. For AWS Parallel Cluster, it was easy to stop and restart the cluster to reduce costs

- Cyclecloud is a licensed application, building on it and sharing it must be done through Azure Marketplace after a review process – on AWS the ebs volume with data was made public along with the YAML files to generate the pcluster

- Performance comparison: The Intel-based Azure and AWS offers relying on similar hardware (Xeon) performed and scaled similarly using the same number of processors. AMD EPYC processors show improved efficiency (44 -> 120cpus, decreased run time to half, while costs were about 10% more). In the case of spot instances HB120v2 costs the same as HC44rs!

- Availability of compute nodes is another question that may impact different regions in a different way

Questions - Poll

- Does your group have an Azure account (institutional or individual?)
- Previous experience setting up CMAQ? Level?
- Is your goal to run on a single VM? Experience level?
- Is your goal to run on multi-node clusters ? Experience level?
- Do you have an existing cloud storage solution? Which one?
- Has your group tried running CMAQ on the Parallel Cluster since the release of the tutorial? If so, please provide feedback on your experience.
- Link to poll: https://forms.gle/Civicg3rdzkZBHz49