Azure
Hand on Lab

HPC
Machine Learning
Cognitive Services

Hands-On Lab

SLURM Linux Cluster HOL.html

Docker and Azure Container Service HOL.html

doAzureParallel: Take advantage of Azure’s flexible compute directly from R session

- https://azure.microsoft.com/en-us/blog/doazureparallel/

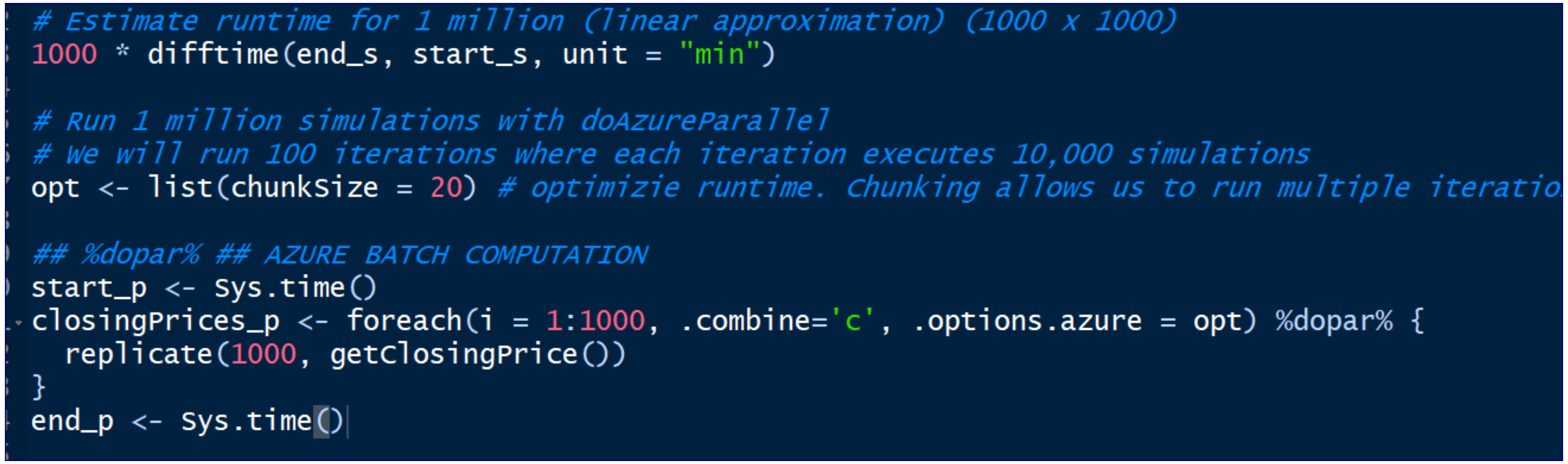
- doAzureParallel – a lightweight R package built on top of Azure Batch, that allows you to easily use Azure’s flexible compute resources right from your R session. doAzureParallel compliments Microsoft R Server by providing the infrastructure for running massive compute parallel simulations.
- At its core, the doAzureParallel package is a parallel backend, for the widely popular foreach package, that lets you execute multiple processes across a cluster of Azure virtual machines. In just a few lines of code, the package helps you create and manage a cluster in Azure, and register it as a parallel backend to be used with the foreach package

- https://github.com/Azure/doAzureParallel/tree/master/samples

Monte Carlo Pricing - HPC Simulation - %dopar%

8-node cluster (standard D2v2: 2 vCPU, 7 Gb)

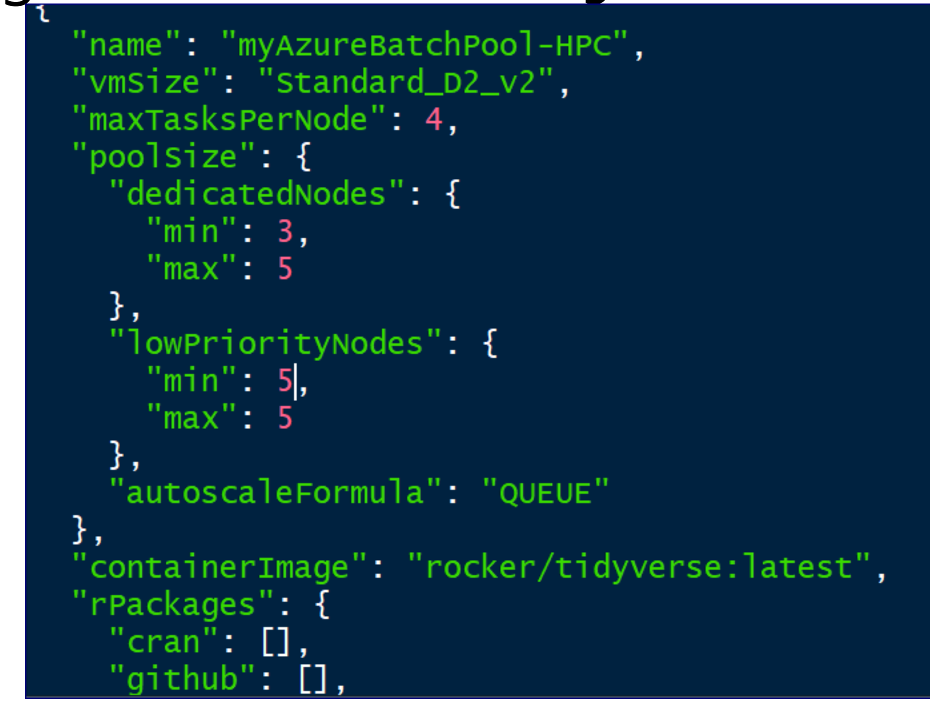
- specify VM class in `cluster.json`
- specify credentials for Azure Batch and Azure Storage in `credentials.json`



```
# Estimate runtime for 1 million (linear approximation) (1000 x 1000)
1000 * difftime(end_s, start_s, unit = "min")

# Run 1 million simulations with doAzureParallel
# We will run 100 iterations where each iteration executes 10,000 simulations
opt <- list(chunkSize = 20) # optimizie runtime. Chunking allows us to run multiple iteratio

## %dopar% ## AZURE BATCH COMPUTATION
start_p <- Sys.time()
closingPrices_p <- foreach(i = 1:1000, .combine='c', .options.azure = opt) %dopar% {
  replicate(1000, getClosingPrice())
}
end_p <- Sys.time()
```



```
{
  "name": "myAzureBatchPool-HPC",
  "vmSize": "Standard_D2_v2",
  "maxTasksPerNode": 4,
  "poolSize": {
    "dedicatedNodes": {
      "min": 3,
      "max": 5
    },
    "lowPriorityNodes": {
      "min": 5,
      "max": 5
    },
    "autoscaleFormula": "QUEUE"
  },
  "containerImage": "rocker/tidyverse:latest",
  "rPackages": {
    "cran": [],
    "github": [],
```

Azure Machine Learning Pipeline with AzureBatchStep

This Jupyter notebook is used to demonstrate the use of AzureBatchStep in Azure Machine Learning Pipeline. An AzureBatchStep will submit a job to an AzureBatch Compute to run a simple windows executable.

- https://github.com/Azure/MachineLearningNotebooks/blob/master/how-to-use-azureml/machine-learning-pipelines/intro-to-pipelines/aml-pipelines-how-to-use-azurebatch-to-run-a-windows-executable.ipynb

Guidance and framework for running HPC applications on Azure

- https://github.com/az-cat/az-hpcapps

- This repository provides automation scripts for creating Azure Batch pools and Azure CycleCloud clusters that you can use to run common high-performance computing (HPC) applications. This repo also serves as a catalog of HPC applications that you can use for testing. More than a dozen common HPC applications are currently supported, including several ANSYS solvers and Star-CCM+, and you can add more as needed as described in this guide.

Microsoft

The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.