# Create RedHat 7.4 image with RDMA drivers

## Obtain RDMA drivers and mpi libraries from CentOS-HPC:7.4 on a Standard\_H16r node

#> tar -zcf rdma.tgz /opt/mircrosoft

#> tar -zcf intel.tgz /opt/intel

## Start build VM with RedHat7.4

#> az group create -n osimages -l westeurope

#> az vm create -g osimages -n compnode --image RedHat:RHEL:7.4:7.4.2018010506 --size Standard\_H16r --storage-sku Standard\_LRS

#> az vm list-ip-addresses -o table

## Modify OS disk on VM

Copy rdma.tgz and intel.tgz to the VM.

Install this exact kernel to match the rdma kernel drivers and reboot to run the correct kernel

#> sudo yum install kernel-3.10.0-693.17.1.el7.x86\_64 libibverbs dapl

#> sudo reboot

Unpackag the rdma drivers and install; disable firewall and modify waagent config file to enable rdma; update limits.conf for mpi.

#> tar zxvf rdma.tgz

#> cd opt/microsoft/rdma/rhel74

#> sudo yum install ./kmod-microsoft-hyper-v-rdma-4.2.3.1.144-20180209.x86\_64.rpm ./microsoft-hyper-v-rdma-4.2.3.1.144-20180209.x86\_64.rpm

#> sudo systemctl disable firewalld

#> sudo vi /etc/waagent.conf

# Enable RDMA management and set up, should only be used in HPC images

OS.EnableRDMA=y

#> sudo vi /etc/security/limits.conf

\* soft memlock unlimited

\* hard memlock unlimited

Untar the Intel MPI libraries

#> cd /

#> sudo tar zxvf <path to>/intel.tgz

Reboot the vm and verify the rdma interface (eth2) has a ip-address (172.16.\*.\*)

#> reboot

#> ip a

## Create a VM Image from the VM

#> az vm deallocate -g osimages -n compnode

#> az vm generalize -g osimages -n compnode

#> az image create -g osimages -n redhat-hpc2 --source compnode

## Build scaleset to test

#> az group create -n hpccluster -l westeurope

#> az vmss create -n vmsscluster -g hpccluster \

--vm-sku Standard\_H16r \

--instance-count 2 \

--disable-overprovision \

--image /subscriptions/12345-123-123-12345/resourceGroups/osimages/providers/Microsoft.Compute/images/redhat74-hpc

Log into the first node and run pingpong

#> source /opt/intel/impi/5.1.3.223/bin64/mpivars.sh

#> mpirun -hosts node1,node2 -ppn 1 -n 2 -env I\_MPI\_FABRICS=dapl -env I\_MPI\_DAPL\_PROVIDER=ofa-v2-ib0 -env I\_MPI\_DYNAMIC\_CONNECTION=0 IMB-MPI1 pingpong

## Add on: expanding with Cuda GPU drivers on a Standard\_NC24 during step 3.

Installing the kernel with the correct veriosn to stay compatible with the RDMA drivers above

#> sudo yum install -y kernel-3.10.0-693.17.1.el7.x86\_64

#> sudo reboot

Add dependencies for the cuda drivers: kernel headers and devel package, LIS and dkms

#> sudo yum install -y kernel-devel-3.10.0-693.17.1.el7.x86\_64 kernel-headers-3.10.0-693.17.1.el7.x86\_64

#> wget https://aka.ms/lis

#> tar xvzf lis

#> cd LISISO

#> sudo ./install.sh

#> sudo reboot

#> sudo rpm -Uvh https://dl.fedoraproject.org/pub/epel/epel-release-latest-7.noarch.rpm

#> sudo yum -y install dkms

Now we can download and install the Cuda packages and drivers

#> CUDA\_REPO\_PKG=cuda-repo-rhel7-10.0.130-1.x86\_64.rpm

#> wget http://developer.download.nvidia.com/compute/cuda/repos/rhel7/x86\_64/${CUDA\_REPO\_PKG} -O /tmp/${CUDA\_REPO\_PKG}

#> sudo rpm -ivh /tmp/${CUDA\_REPO\_PKG}

#> rm -f /tmp/${CUDA\_REPO\_PKG}

#> sudo yum install -y cuda-drivers

#> sudo reboot

#> sudo yum install -y cuda

#> sudo reboot

You can test the nvidia drivers with nvidia-smi:

$ nvidia-smi

+-----------------------------------------------------------------------------+

| NVIDIA-SMI 410.48 Driver Version: 410.48 |

|-------------------------------+----------------------+----------------------+

| GPU Name Persistence-M| Bus-Id Disp.A | Volatile Uncorr. ECC |

| Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util Compute M. |

|===============================+======================+======================|

| 0 Tesla K80 Off | 00006F75:00:00.0 Off | 0 |

| N/A 42C P0 71W / 149W | 0MiB / 11441MiB | 0% Default |

+-------------------------------+----------------------+----------------------+

| 1 Tesla K80 Off | 00008970:00:00.0 Off | 0 |

| N/A 50C P0 54W / 149W | 0MiB / 11441MiB | 0% Default |

+-------------------------------+----------------------+----------------------+

| 2 Tesla K80 Off | 0000A8F3:00:00.0 Off | 0 |

| N/A 49C P0 58W / 149W | 0MiB / 11441MiB | 0% Default |

+-------------------------------+----------------------+----------------------+

| 3 Tesla K80 Off | 0000C82A:00:00.0 Off | 0 |

| N/A 39C P0 78W / 149W | 0MiB / 11441MiB | 1% Default |

+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+

| Processes: GPU Memory |

| GPU PID Type Process name Usage |

|=============================================================================|

| No running processes found |

+-----------------------------------------------------------------------------+