

AI WORKLOADS AT SCALE

Kubernetes Cluster with Supermicro's Systems with AMD EPYC[™] 7002 Series processors

TABLE OF CONTENTS

Executive Summary	1
System Configuration	2
Introduction to Kubernetes	4
Kubernetes Cluster Deployment	4
Scale-up With Kubernetes Cluster	8
Conclusion	9



Supermicro AMD based WIO System

Executive Summary

The Deep Learning (DL) benchmark results in the previous white paper¹ clearly show that a DL workload in Docker containers performs the same as on a BareMetal system. Building an on-prem Kubernetes cluster with GPU workers and AI framework-specific Docker containers can help organizations run projects or productions in a highly reliable and scalable platform. This white paper introduces Supermicro AMD based WIO systems, AS-1114S-WTRT, as Kubernetes Admin and master nodes. Along with AS-2023US-TR4, we build an NVIDIA GPU capable Kubernetes cluster that uses Cloud-native CEPH storage as persistent volumes and demonstrates how a DL workload can scale the Kubernetes cluster.

SUPERMICRO

Supermicro (Nasdaq: SMCI), the leading innovator in high-performance, highefficiency server and storage technology is a premier provider of advanced server Building Block Solutions® for Enterprise Data Center, Cloud Computing, Artificial Intelligence, and Edge Computing Systems worldwide. Supermicro is committed to protecting the environment through its "We Keep IT Green®" initiative and provides customers with the most energy-efficient, environmentally-friendly solutions available on the market.

¹ White paper: <u>supermicro® system combines and epyc™processors and nvidia gpus to achieve</u> <u>consistent deep learning performance with linear scaling</u>

System Configuration

AS-1114S-WTRT is one of the Supermicro AMD EPYC[™] 7002 series based WIO series servers, offering a wide range of I/O options to deliver truly optimized systems for specific requirements. For more detailed system information, please go HERE. Customers can optimize the storage and networking alternatives to accelerate performance, find the perfect fit for their applications. In our case, it uses the VMWare host and Kubernetes master nodes. Figure 1 and Figure 2 provide an overview of the system:

Table 1 provides a sample configuration for AS-1114S-WTRT



Figure 1



Figure 2

System	Part Number	QTY	Part Description
Infrastructure	AS-1114S-WTRT	1	H12 SSW-NT,CSV116TS-R504WBP
CPU	PSE-ROM7552	1	AMD EPYC™ 7552 DP/UP 48C/96T 2.2G 192M 200W 4094
Memory	MEM-DR416L-HL01-ER32	8	16GB DDR4-3200 2Rx8 ECC REG DIMM
HDD/SSD (Storage)	HDS-X2A-XS7680TE70004	2	[NR]Seagate Lange 7.68TB SAS 12Gb/s, 15mm, 2.5", 0.8DWPD SSD, HF
HDD/SSD (OS)	HDS-SMN1-MZ1LB3T8HMLA07	2	Samsung PM983 3.84TB NVMe PCIe3x4 V4 M.2 22x110mm (1.3 DWPD)
AOC	MCX4121A-ACAT	2	Standard Low-profile Mellanox 25GbE card with 2x SFP28 ports
Storage Controller	AOC-S3008L-L8i	1	AOC-S3008L-L8i Retail Pack
Cable	CBL-SAST-0593	1	Internal Mini-SAS HD to Mini-SAS HD 60cm,30AWG,12Gb/s
Software License	SFT-DCMS-SINGLE	1	Supermicro System Management Software Suite node license, HF, RoHS/REACH, PBF

Table 1

Table 2 provides cluster system roles and components. For more information on solution reference architectures, which consists of different level bundle configurations with various software vendors depending on the customers' workloads, refer to our team's solution page https://verticalsolutions.supermicro.com for more information.



System	Role	QTY	Software
AS-1114S-WTRT WIO server	Virtual host with Kubernetes admin node	1	VMWare 6.5
AS-1114S-WTRT WIO server	Kubernetes Master nodes	3	Kubernetes 1.19
AS-2023US-TR4 with 2xV100 32GB PCIe GPU Per Node	Kubernetes Worker nodes	3	Kubernetes 1.19 Nvidia-docker
SSE-X3648T	Network Switches TOR	2	SMCI OS

Table 2

Figure 3 shows the reference architecture.







Introduction to Kubernetes

Kubeflow is a machine learning toolkit that runs on Kubernetes by abstracting machine learning solutions and best practices utilizing the features of Kubernetes. It makes deployments of machine learning workflows portable and scalable and runs in distributed environments. Kubeflow's core and ecosystem critical user journeys (CUJs) provide software solutions for end-to-end workflows to build, train, deploy, and/or develop a model and create and run a workflow pipeline. Details regarding Kubernetes are at: https://www.kubeflow.org/docs/started/kubeflow-overview/

The AI/ML applications can be deployed in either a single node cluster or multi-node clusters, with each node loaded with multiple GPUs. Single node clusters can be used to build proof of concept solutions and smaller-scale training tasks utilizing multiple GPUs. The single-node clusters are usually used in development environments and for non-production use. Multi-node clusters are used in production deployments as they provide high availability, significant improvement in speed of execution, and efficient scalability.

Kubernetes Cluster Deployment

Making GPU work with the Kubernetes cluster on-prem can be a complicated task. Therefore, NVIDIA has developed a tool named DeepOps, to automate Kubernetes cluster provisioning, configuring, and scaling. Further information regarding DeepOps can be found <u>here</u>.

Figure 4 outlines the process to prepare and deploy the NVIDIA GPU cluster on-prem.



Figure 4



The tool is based on two open-source projects: Ansible and Kubespray, which are very flexible and scalable. It can deploy the Kubernetes cluster onto a single GPU node or hundreds of systems except for the cluster-admin node. Once DeepOps is installed on the admin node, modifying the inventory file according to the design. Figure 6 is the inventory file located under the deepops/config directory. In addition, the k8s-cluster.yml in the deepops/playbooks provides cluster options.

# Als [kube gpu2 gpu3 #gpu0 #gpu0	o add mo -node] 1 2	gmt/master	nodes	here	if	they	will	run	non-con	ıtrol	plane	jobs	
[k8s- kube- kube-	cluster: master node	children]											
##### # SLUI ##### [sluri #logi	# RM # m-master n01	-]											
[slur #gpu0 #gpu0	m-node] 1 2												
[slurn slurm slurm	m-cluste -master -node	er:childre	n]										
##### # SSH #####	# connect #	tion confi	guratio	on									
[all: # SSH #ansi #ansi	varsj User ble_usei ble_ssh_	∽=ubuntu _private_k	ey_file	e='~/	. ssł	n/id_	rsa'						
# SSH #ansi ubunt	bastior ble_ssh_ u@deepop	n/jumpbox _common_ar os:~/deepo	gs='-o ps/con	Prox fig\$	yCor	nmand:	="ssh	-W 9	%h:%p -q	l npni	ntu@10	.0.0.1	
					Г	iyur	83						

To set up the Kubernetes cluster with GPU installed worker nodes, run the following command:

ansible-playbook --limit k8s-cluster playbooks/k8s-cluster.yml -k -K³

² Link to the k8s-cluster.yml file <u>https://drive.google.com/file/d/1EZLH4lfp1IVtaJWhPOvcjvNyut7L5ady/view?usp=sharing</u>

³ More instructions can be found in <u>https://verticalsolutions.supermicro.com</u> and <u>HERE</u>



After the installation process is completed, the Kubernetes cluster should be up and running. To validate the cluster, type the following command on the admin node: kubectl get node. Figure 6 shows the output.

ubuntu	@deepops	:~\$ kubec	tl aet	node	
NAME	STATUS	ROLES	AGE	VERSION	N
gpu2	Ready	master	56d	v1.18.9	9
gpu3	Ready	<none></none>	56d	v1.18.9	9
ubuntu	@deepops	:~\$ kubec	tl get	namespac	ce
NAME			STATU	IS AGE	
cert-m	nanager		Activ	/e 56d	
cluste	er-gpu-ve	rify	Activ	7e 13m	
defaul	lt		Activ	/e 56d	
gpu-op	berator-r	esources	Activ	/e 56d	
istio-	system		Activ	re 56d	
knativ	ve-servin	g	Activ	re 56d	
kube-r	node-leas	e	Activ	re 56d	
kube-p	oublic		Activ	re 56d	



Figure 7 shows the overall architecture discussed in this paper.



Figure 7

NVIDIA DeepOps is a handy open-source tool to deploy GPU Cloud-Native clusters. The tested environment consists of one VM as a deployment host and two AS-2023US-TR4 as GPU nodes. You can find the detailed features and configurations <u>here</u>.



The tool also installs and configures a Cloud-native certified rook-ceph CEPH storage system as the cluster's persistent volumes, which is a must for all Deep Learning workloads since the datasets and intermediate training results have to be shared among the GPU nodes. Figure 8 and Figure 9 show the dynamic CEPH StorageClass information for the Kubernetes cluster.

ubuntu@deepops:~\$ kubectl get pods -n rook-ceph				
NAME	READY	STATUS	RESTARTS	AGE
csi-cephfsplugin-4xghp	3/3	Running	0	56d
csi-cephfsplugin-b7hcr	3/3	Running	Θ	56d
csi-cephfsplugin-provisioner-68b5fb6499-m2tsf	4/4	Running	0	56d
csi-cephfsplugin-provisioner-68b5fb6499-x5z5k	4/4	Running	1	56d
csi-rbdplugin-btsp4	3/3	Running	0	56d
csi-rbdplugin-dhgq9	3/3	Running	0	56d
csi-rbdplugin-provisioner-675d74c797-pn4sw	5/5	Running	Θ	56d
csi-rbdplugin-provisioner-675d74c797-szn8z	5/5	Running	3	56d
rook-ceph-agent-qhbpc	1/1	Running	0	56d
rook-ceph-agent-xmtkl	1/1	Running	0	56d
rook-ceph-mds-cephfs-a-56dcd84ddc-scqd7	1/1	Running	0	56d
rook-ceph-mds-cephfs-b-7d666c9755-l8kp2	1/1	Running	0	56d
rook-ceph-mgr-a-7ff587789b-k5mtx	1/1	Running	0	56d
rook-ceph-mon-a-754458fd4-l9czm	1/1	Running	Θ	56d
rook-ceph-mon-b-66d96d95d-wm9zl	1/1	Running	0	56d
rook-ceph-mon-c-89c7569fd-gg2mm	1/1	Running	0	56d
rook-ceph-operator-768dfb98c4-9hzvx	1/1	Running	2574	56d
rook-ceph-osd-0-75b46fcfb6-ld2w8	1/1	Running	0	56d
rook-ceph-osd-1-7747c8df4f-skqjs	1/1	Running	Θ	56d
rook-ceph-osd-prepare-gpu2-sp4qk	0/1	Completed	0	35m
rook-ceph-osd-prepare-gpu3-fbjxq	0/1	Completed	0	35m
rook-ceph-tools-8d9d7c8f4-nnmxm	1/1	Running	0	56d
rook-discover-2f6w4	1/1	Running	0	56d
rook-discover-tb2lw	1/1	Running	0	56d
ubuntu@deepops:~\$				

Figure 8

NAMECAPACITYACCESS MODESRELAIM POLICYSTATUSCLAIMSTORAGECLASSREASONAGE26iRWXRetainBoundkubeflow/benchmark-pv-claimrook-ceph-block54d16iRWORetainBoundkubeflow/mnist-pv-claimrook-ceph-block56d206iRWORetainBoundkubeflow/mnist-pv-claimrook-ceph-block56d56d16iRWORetainBounddefault/busybox-1rook-ceph-block55d16iRWORetainBounddefault/busybox-2rook-ceph-block55d16iRWORetainBounddefault/busybox-2rook-ceph-block54d16iRWORetainBounddefault/busybox-2rook-ceph-block54d54d106iRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block54d56d106iRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block55d206iRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block55d206iRWORetainBoundkubeflow/metadata-mysqlrook-ceph-block55d206iRWORetainBoundkubeflow/metadata-mysqlrook-ceph-block55d206iRWORetainBoundkubeflow/mist1-pv-claimrook-ceph-block56d206iRWO <t< th=""><th>ubuntu@deepops:</th><th>:~/kubtest\$ kubectl get pv</th><th></th><th></th><th></th><th>second associated</th><th></th><th></th></t<>	ubuntu@deepops:	:~/kubtest\$ kubectl get pv				second associated		
ASSREASONAGEbenchmark-pv2GiRWXRetainBoundkubeflow/benchmark-pv-claimrook-cephblock5dd1GiRWORetainBoundkubeflow/mnist-pv-claimrook-cephblock5dd1GiRWORetainBoundkubeflow/mnist-pv-claimrook-cephblock5dd2GiRWORetainReleasedkubeflow/mnist-pv-claimrook-cephblock5dd1GiRWORetainBounddefault/busybox-1rook-cephblock5dd1GiRWORetainBounddefault/busybox-2rook-cephblock5dd1GiRWORetainBounddefault/busybox-2rook-cephblock5dd1GiRWORetainBounddefault/busybox-2rook-cephblock5ddpvc-72d3clf-9013-d874-a9b4-6325577caac510GiRWORetainReleasedkubeflow/katib-mysqlrook-cephblock5dd5d10GiRWORetainReleasedkubeflow/benchmark-pv-claimrook-cephblock5d5d1GiRWORetainReleasedkubeflow/benchmark-pv-claimrook-cephblock5d5d1GiRWORetainBoundkubeflow/metadata-mysqlrook-cephblock5d5d1GiRWORetainBoundkubeflow/metadata-mysqlrook-cephblock5d5d1GiRWORetainBoundkubeflow/metadat	NAME		CAPACITY	ACCESS MODES	RECLAIM POLICY	STATUS	CLAIM	STORAGECL
benchmark-pv 26i RWX Retain Bound kubeflow/benchmark-pv-claim rook-ceph -block 54d "rook-ceph -block 56d 2061 RWO Retain Bound kubeflow/mnist-pv-claim rook-ceph -block 56d pvc-244a88dd-9eaa-4d1f-b6b2-483d8ac88b5c 16i RWO Retain Bound default/busybox-1 rook-ceph -block 56d pvc-244a88dd-9eaa-4d1f-b6b2-483d8ac88b5c 16i RWO Retain Bound default/busybox-1 rook-ceph -block 56d pvc-244a88dd-9eaa-4d1f-b6b2-483d8ac88b5c 16i RWO Retain Bound default/busybox-2 rook-ceph -block 56d pvc-2362919-022C-4546-b69f-b06bef2f7fd9 16i RWO Retain Bound default/busybox-2 rook-ceph -block 54d pvc-6c14fa1f-cfb6-4d1f-bf1d-24f45808d5c3 106i RWO Retain Released kubeflow/katib-mysql rook-ceph -block 56d pvc-7b2d3c1f-9013-4874-a9b4-6325577caac5 106i RWO Retain Released kubeflow/katib-mysql rook-ceph -block 56d pvc-7b2d3c1f-9013-4874-a9b4-6325577caac5 106i RWO Retain Released kubeflow/benchmark-pv-claim rook-ceph -block 56d pvc-7b9c834-9111-45a2-bdc4-9a96e19e4b49 26i RWO Retain Released kubeflow/benchmark-pv-claim rook-ceph -block 54d pvc-8062982d-e06c-40d6-9d78-aba84de69f40 206i RWO Retain Bound kubeflow/mnist1-pvc rook-ceph -block 56d pvc-438d7173-741b-49ac-b76f-f0db2abd02ce 106i RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-438d1713-741b-49ac-b76f-f0db2abd02ce 106i RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-438d1713-741b-49ac-b76f-f0db2abd02ce 106i RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-438d1713-741b-49ac-4762-4435-be27-ba0077446369 106i RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d	ASS REASON	AGE						
-block54dmnist-pv1GiRWORetainBoundkubeflow/mnist-pv-claimrook-ceph-block56dpvc-03346976-32fa-4ffd-alab-903e8a0f04a9206iRWORetainReleasedkubeflow/minio-pv-claimrook-ceph-block56drook-ceph-block56drook-ceph-blockrook-ceph-block56dS6dBounddefault/busybox-1rook-ceph-block56drook-ceph-block56drook-ceph-block56dS6dBounddefault/busybox-2rook-ceph-block56dS6dRWORetainBounddefault/busybox-2rook-ceph-block56dS6dBounddefault/busybox-2rook-cephrook-ceph-block56dS6dRWORetainReleasedkubeflow/matala-mysqlrook-ceph-block56dS6dRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block56dS6dS6dRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block56dS6dS6dRWORetainBoundkubeflow/mist1-pvcrook-ceph-block56dS6dS6dRWORetainBoundkubeflow/mist1-pvcrook-ceph-block56dS6dS6dRWORetainBoundkubeflow/mist1-pvcrook-ceph-block56dS6dS6dRWORetainBound <t< td=""><td>benchmark-pv</td><td></td><td>2Gi</td><td>RWX</td><td>Retain</td><td>Bound</td><td>kubeflow/benchmark-pv-claim</td><td>rook-ceph</td></t<>	benchmark-pv		2Gi	RWX	Retain	Bound	kubeflow/benchmark-pv-claim	rook-ceph
mist-pv1GiRWORetainBoundkubeflow/mmist-pv-claimrook-ceph-block56dpvc-03346976-32fa-4ffd-alab-903e8a0f04a920GiRWORetainReleasedkubeflow/minio-pv-claimrook-ceph-block56dpvc-244a88dd-9eaa-4dlf-b6b2-483d8ac88b5c1GiRWORetainBounddefault/busybox-1rook-ceph-block55dpvc-538c9319-c92c-4546-b69f-b06bef2f7fd91GiRWORetainBounddefault/busybox-2rook-ceph-block54d54d1GiRWORetainReleasedkubeflow/katib-mysqlrook-ceph-block54d54d10GiRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block56d56d10GiRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block56d56d10GiRWORetainReleasedkubeflow/metadata-mysqlrook-ceph-block56d56d10GiRWORetainBoundkubeflow/metadata-mysqlrook-ceph-block55d56d1GiRWORetainBoundkubeflow/ministl-pvcrook-ceph-block56d56d1GiRWORetainBoundkubeflow/ministl-pvcrook-ceph-block56d56d1GiRWORetainBoundkubeflow/mistl-pvcrook-ceph-block56d56d1GiRWORetainBoundkubeflow/mistl-pv	-block	54d						
-block 56d pvc-03346976-32fa-4ffd-alab-903e8a0f04a9 20Gi RWO Retain Released kubeflow/minio-pv-claim rook-ceph -block 56d pvc-244a88dd-9eaa-4dlf-b6b2-483d8ac88b5c 1Gi RWO Retain Bound default/busybox-1 rook-ceph -block 54d -block 54d -block 54d pvc-538c9319-c92-c4546-b69f-b06bef2f7fd9 1Gi RWO Retain Bound default/busybox-2 rook-ceph -block 54d -block 54d -block 54d -block 56d pvc-7c2d3clf-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/katib-mysql rook-ceph -block 56d pvc-7c2d3clf-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7b9c834-911-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/benchmark-pv-claim rook-ceph -block 55d pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/ministl-pvc rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/mistl-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/mistl-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/mistl-pv-claim rook-ceph -block 56d pvc-eaff52d4-2272-405-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/mistl-pv-claim rook-ceph -block 56d pvc-eaff52d2-a240-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mistl-pv-claim rook-ceph -block 56d pvc-eaff52d2-a240-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mistl-pv-claim rook-ceph -block 56d	mnist-pv		161	RWO	Retain	Bound	kubeflow/mnist-pv-claim	rook-ceph
pvc-03346976-32fa-4ffd-alab-903e8a0f04a9 2061 RWO Retain Released kubeflow/minio-pv-claim rook-ceph -block 55d pvc-244488dd-9eaa-4d1f-b6b2-483d8ac88b5c 16i RWO Retain Bound default/busybox-1 rook-ceph -block 55d pvc-538c9319-c92c-4546-b69f-b06bef2f7fd9 16i RWO Retain Bound default/busybox-2 rook-ceph -block 54d pvc-6c14fa1f-cfb6-4d1f-bf1d-24f45808d5c3 106i RWO Retain Released kubeflow/katib-mysql rook-ceph -block 56d pvc-7c2d3c1f-9013-4874-a9b4-6325577caac5 106i RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7b9c834-9111-46a2-bdc4-9a96e19e4b49 26i RWX Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7b9c834-9111-46a2-bdc4-9a96e19e4b49 26i RWX Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-8bf205c-de3b-4106-b5ef-82528fc56526 16i RWO Retain Bound kubeflow/mnist1-pvc rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 106i RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 106i RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-e3f52d2-a24b-4c1d-83e-bdc4-9a967446369 106i RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e3f52d4-2a72-4b3-be27-ba0077446369 106i RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 206i RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 206i RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d	-block	56d						
-block56dpvc-244888dd-9eaa-4d1f-b6b2-483d8ac88b5c1GiRWORetainBounddefault/busybox-1rook-ceph-blockS5dpvc-538c9319-c92c4546-b69f-b06bef2f7fd91GiRWORetainBounddefault/busybox-2rook-ceph-blockS4dS4dS4dS4drook-cephsdd-blockS6dRWORetainReleasedkubeflow/katib-mysqlrook-ceph-blockS6dS6drook-cephsddrook-ceph-blockS6dS6drook-cephsddrook-ceph-blockS6dS6drook-cephrook-ceph-blockS6dS6drook-cephrook-ceph-blockS6drook-cephrook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6drook-cephrook-ceph-blockS6d <t< td=""><td>pvc-03346976-32</td><td>2fa-4ffd-alab-903e8a0f04a9</td><td>20G1</td><td>RWO</td><td>Retain</td><td>Released</td><td>kubeflow/minio-pv-claim</td><td>rook-ceph</td></t<>	pvc-03346976-32	2fa-4ffd-alab-903e8a0f04a9	20G1	RWO	Retain	Released	kubeflow/minio-pv-claim	rook-ceph
pvc-2344a88dd-9eaa-4d1f-b6b2-483d8ac88b5c 1G1 RWO Retain Bound default/busybox-1 rook-ceph -block 54d pvc-5380319-c92c-4546-b69f-b06bef2f7fd9 1Gi RWO Retain Bound default/busybox-2 rook-ceph -block 54d pvc-6c14falf-cfb6-4d1f-bf1d-24f45808d5c3 10Gi RWO Retain Released kubeflow/katib-mysql rook-ceph -block 56d pvc-7c2d3clf-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7fbc834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/metadata-mysql rook-ceph -block 55d pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/minist1-pvc rook-ceph -block 55d pvc-60c982dd-e06c-4066-9d78-aba84de69f40 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-sade-b76f-f0db2abd02ce 20Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d	-block	56d						
-block 55d pvc-538c9319-c92c 4546-b69f-b06bef2f7fd9 1Gi RWO Retain Bound default/busybox-2 rook-ceph -block 54d pvc-6c14falf-cfb6-4d1f-bf1d-24f45808d5c3 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7c2d3c1f-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7fb9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/metadata-mysql rook-ceph -block 55d pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mnist1-pvc rook-ceph -block 54d pvc-d09082dd-e06c-40d6-978-aba84de69f40 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-e3f52d4a-2a72-4b3-be2r-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d	pvc-244a88dd-96	eaa-4d1f-b6b2-483d8ac88b5c	1G1	RWO	Retain	Bound	default/busybox-1	rook-ceph
pvc-538c9319-c92c-4546-b697-b06bef2f7fd9 1G1 RWO Retain Bound default/busybox-2 rook-ceph -block 54d pvc-6c14falf-cfb6-4d1f-bf1d-24f45808d5C3 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7c2d3c1f-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7fb9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7b9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWO Retain Bound kubeflow/mistl-pvc rook-ceph -block 54d -block 56d pvc-8b7205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-ed53d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d	-block	55d						
-block 54d pyc-6c14falf-cfb6-4d1f-bf1d-24f45808d5c3 10Gi RWO Retain Released kubeflow/katib-mysql rook-ceph -block 56d pyc-7c2d3c1f-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pyc-7fb9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/benchmark-py-claim rook-ceph -block 55d pyc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mnist1-pyc rook-ceph -block 54d pyc-d9c982dd-e06c-40d6-9d78-aba84de69f40 20Gi RWO Retain Released kubeflow/mysql-py-claim rook-ceph -block 56d pyc-d930f173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pyc-d830f173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pyc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pyc-eaff52d2-a24b-4c1d-83aed-866e70072bdf 20Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d	pvc-538c9319-c9	92c-4546-b69t-b06bet2t7td9	161	RWO	Retain	Bound	default/busybox-2	rook-ceph
pvc-7c1241a1r-cTbb-4d11r-br1de-24f45808d5c3 10G1 RWO Retain Released kubeflow/katib-mysql rook-ceph -block 56d pvc-7c2d3c1f-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pvc-7fb9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/benchmark-pv-claim rook-ceph -block 55d pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mnist1-pvc rook-ceph -block 54d pvc-d0c982dd-e06c-40d6-9d78-aba84de69f40 20Gi RWO Retain Released kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d	-block	540		-			1 1 11 11 11 11	
-Dlock 56d pyc-7c2d3clf-9013-4874-a9b4-6325577caac5 10Gi RWO Retain Released kubeflow/metadata-mysql rook-ceph -block 56d pyc-7b96c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/metadata-mysql rook-ceph -block 55d pyc-8b67205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mnist1-pyc rook-ceph -block 54d pyc-d02082dd-e06c-40d6-9d78-aba84de69f40 20Gi RWO Retain Released kubeflow/mysql-py-claim rook-ceph -block 56d pyc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pyc-e35d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pyc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pyc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-py-claim rook-ceph -block 56d	pvc-6c14falf-c1	rb6-4d1t-bt1d-24t45808d5c3	1061	RWO	Retain	Released	kubeflow/katib-mysql	rook-ceph
pvc-7fb9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWU Retain Released kubeflow/metadata-mysql rook-ceph -block 5Gd pvc-7fb9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/mnist1-pvc rook-ceph -block 5Gd pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mnist1-pvc rook-ceph -block 54d -block 56d pvc-d30d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-d30d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/msql-pv-claim rook-ceph -block 56d	-block	560					1.1.0	
-Dicck 56d pvc-7fb9c834-9111-46a2-bdc4-9a96e19e4b49 2Gi RWX Retain Released kubeflow/benchmark-pv-claim rook-ceph -block 55d pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mnist1-pvc rook-ceph -block 54d -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d -block 56d -block 56d pvc-e38d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d -block 56d -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d	pvc-/c2d3c1T-90	013-4874-a9b4-6325577caac5	1061	RWO	Retain	Released	kubeflow/metadata-mysql	rook-ceph
pvc-7302C34-9111-46a2-0624-9390e19e4049 201 NWX Retain Released Ruberlow/Denchmark-pv-Claim rook-cepn -block 55d pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mnistl-pvc rook-ceph -block 56d pvc-d038d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d -block 56d -block 56d pvc-e8f52d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d -block 56d -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d	-DLOCK		201	5 1.04	A18155	8.10.000	L.L.B. destant and the	
-DICK 550 pvc-8b0f205c-de3b-4106-b5ef-82528fc56526 1Gi RWO Retain Bound kubeflow/mmist1-pvc rook-ceph -block 54d pvc-d06C982dd-e06c-40d6-9d78-aba84de69f40 20Gi RWO Retain Released kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-eff4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph	pvc-/Tb9c834-9	111-4682-D0C4-989681984049	261	RWX	Ketain	Released	kubertow/benchmark-pv-claim	rook-ceph
pvc-800f205-0e30-4106-058F-82228F250526 101 RWU Retain Bound Rubeflow/mmisti-pvc rook-cepn -block 54d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d -block 56d pvc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e8f52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d	-DLOCK		202	000	n and the	and the second	hat discount and	and such
-Diock 540 pvc-d0c982dd-e06c-40466-9d78-aba84de69f40 20Gi RWO Retain Released kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-d33d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e83f94a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d	pvc-8001205c-de	3D-4100-D50T-82528TC50520	161	RWU	Retain	Bound	Rubertow/mnisti-pvc	rook-ceph
-block 56d pvc-d38d7173-741b-49ac-b76f-f0db2abd02ce 10Gi RWO Retain Bound kubeflow/mysql-pv-ctaim rook-ceph -block 56d pvc-e85d194a-2272-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-eff4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph	-DLOCK	540 and a particular of the	2003	DLIO	Detain	Palaasad	kuhoflou/mucal nu claim	raak conh
-DUCK 3900 pvc-d38d7173-741b-49ac-b76f-f9db2abd02ce 10Gi RWO Retain Bound kubeflow/metadata-mysql rook-ceph -block 56d pvc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d -block 56d pvc-f7a4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph	pvc-duc982dd-et	10C-4000-9078-aba840e69140	2001	RWU	Retain	Released	Rubertow/mysqt-pv-ctaim	rook-cepn
pvc-e85d194a-2a72-4b35-be27-ba0077446369 1061 RWO Retain Bound Rubeflow/metadata-mysqt rook-ceph -block 56d -block 56d pvc-e8f52d2-a24b-4c1d-8aed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d -block 56d pvc-f7a4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph		11h 40ac h76f fodhachdoaco	1005	PLIO	Potoin	Bound	kuhaflau/matadata musal	rook conh
-DUCK 360 pvc-e85d194a-2a72-4b35-be27-ba0077446369 10Gi RWO Retain Bound kubeflow/katib-mysql rook-ceph -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-f7a4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph	block	+10-49ac-0701-10002a0002ce	1001	RWU	Retain	Bound	Kubertow/metadata-mysqt	rook-ceph
block 56d -block 56d pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-f7a4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph	ovc- 095d1040-20	72 4b25 ba27 ba0077446260	1001	PHO	Patain	Round	kubaflow/katib_musal	rook-conh
pvc-eaff52d2-a24b-4c1d-83ed-866e70072bdf 20Gi RWO Retain Bound kubeflow/mysql-pv-claim rook-ceph -block 56d pvc-f7a4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph	hlock	56d	1001	NWO.	Netatii	Bound	Kubercow/kacib-mysqc	rook-ceph
pro-en 1222-2240-46240-46240-46240-46240-46240-12001 2001 1000-0001 1000-0001 1000-0001 1000-0001 1000-0001 100 	nyc_eaff52d2_a	24b-4c1d-92ed-966e70072bdf	2001	PW0	Patain	Round	kubaflow/musal-pu-claim	rook-conh
-0100, 300 prc-f7a4732a-4f62-4a55-b918-1963fbc127ed 20Gi RWO Retain Bound kubeflow/minio-pv-claim rook-ceph	hlock	56d	2001	NNU.	Necarii	Bound	Kuber cow/mysqc-pv-ctaim	100k-ceph
	nyc . f7a4732a. 4	f62-4255-b018-1063fbc127ed	20Gi	RWO	Retain	Bound	kuheflow/minio.nv.claim	rook-coph
-block Sed	hlock	56d	2001	THIC .	no cozn	ujuanu	Kuber com/minito-pw-c caim	rook-cepit

Figure 9

Scale-up with Kubernetes Cluster

Not only does a Cloud-native platform can provide reliable, resilient, and distributed microservices, but it also can scale up on demand. In the test environment, we have tested Deep Learning benchmark throughput with the number of POD replicas. Figure 10 is the snippet of the Kubernetes YAML configuration of the Deep Learning benchmark application. Figure 11 depicts the benchmark under different Deep learning frameworks.



Figure 10

Numbers of Replica and Throughput



Different Neural Networks



Conclusion

The foundation of Cloud-native technologies is Kubernetes can be complicated to deploy, maintain and upgrade. Adding GPU into the mix can introduce another layer of complexity. However, with Supermicro AS-1114S-WTRT WIO systems, NVIDIA GPU appliances, and NVIDIA DeepOps tool, it becomes more straightforward for an organization to build a private Cloud-native platform along with MLOps applications. From a designing perspective, AS-1114S-WTRT as a flexible cloud node can meet compute, storage, and virtualization requirements for the hardware infrastructure; From an operational aspect, DeepOps makes the on-prem Cloud-native platform easier to deploy, maintain, and scale. This paper also demonstrated how the Kubernetes PODs could speed up Deep Learning processes that can reduce time-to-market.

Reference

https://docs.nvidia.com/datacenter/cloud-native/index.html https://www.kubeflow.org/docs/started/kubeflow-overview/

AMD, the AMD Arrow logo, EPYC, and combinations thereof are trademarks of Advanced Micro Devices, Inc.

