

**Containerization of numerical ocean model  
for computational reproducibility and  
portability in the cloud computing**

**Kwangwoog Jung**

*Seoul National University  
Marine Environmental Prediction Lab.*

- Cloud Computing
- Containerization
- Container Orchestration
- Containerization of ROMS
- Result and Summary
- Demo (Optional)

# What is Cloud Computing?

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models (The NIST Definition of Cloud Computing, NIST, 2011).

## ■ Essential Characteristics

1. On-Demand
2. Broad NW access
3. Resource Pooling
4. Rapid elasticity

## ■ Service Models

1. IaaS (Infra as a Service)
2. PaaS (Platform as a Service)
3. SaaS (Software as a Service)

## ■ Deployment Models

1. Public Cloud (Cloud Service Provider, e.g. AWS, GCP, Azure)
2. Private Cloud (Your Data Center and virtualized Infrastructure)
3. Hybrid Cloud (Public Cloud + Private Cloud)
4. Community cloud (For specific purpose)

# Why cloud computing is important in Ocean Science?

## ■ Support Need of High Resolution & Model Scale

1. High Resolution
2. Model Scale
3. Ensemble Modelling
4. Data Analytics

## ■ Solve the Management and technical Issues

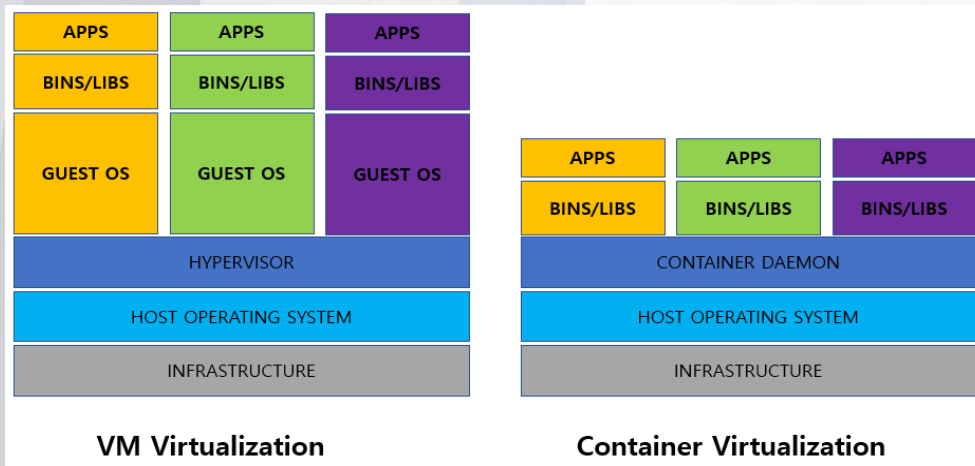
1. Space & Power
2. Time & Cost
3. Technical Complexity
4. Maintenance & Engineers

More Science,  
Less IT Issues

# What is Container and containerization?

Containers are lightweight packages of your application code together with dependencies such as specific versions of programming language runtimes and libraries required to run your software services (*What are Containers?*, Google, <https://cloud.google.com/learn/what-are-containers>).

Containerization is a software deployment process that bundles an application's code with all the files and libraries it needs to run on any infrastructure. Traditionally, to run any application on your computer, you had to install the version that matched your machine's operating system (*What is containerization?*, AWS, <https://aws.amazon.com/what-is/containerization>).



## What is Container Orchestration?

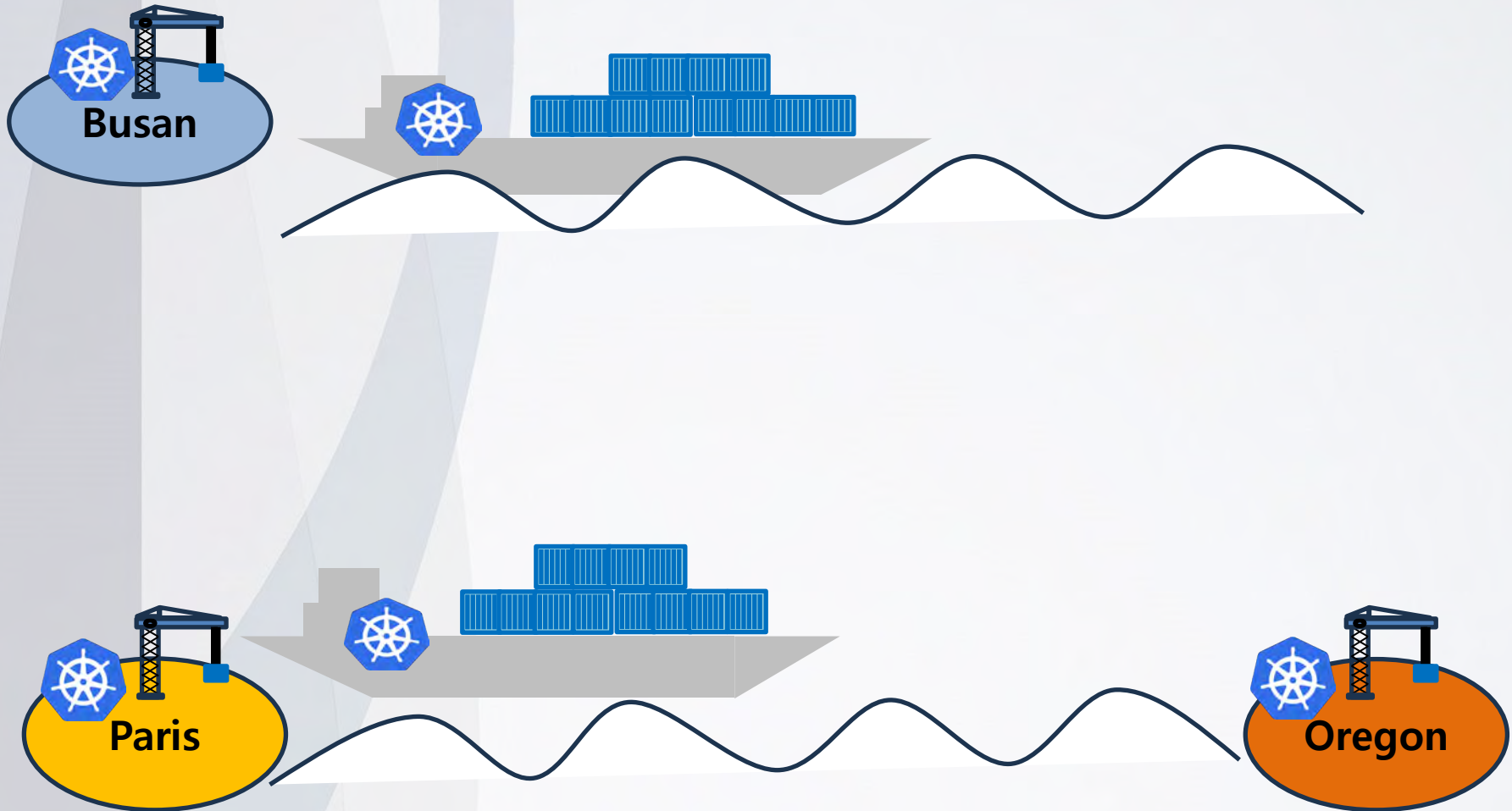
Container orchestration means automatically provisions, deploys, scales, and manages containerized applications without worrying about the underlying infrastructure. Users can implement container orchestration anywhere containers are, allowing them to automate the lifecycle management of containers (What is container orchestration?, Google, <https://cloud.google.com/discover/what-is-container-orchestration>).

Kubernetes is defacto standard orchestration S/W as an opensource(Linux Foundation). (The name Kubernetes (k8s) originates from Greek, meaning helmsman or pilot)



**kubernetes**

# What is Container Orchestration?



# Cloud Service Locations

AWS Local Zones locations

- Generally Available
- Announced



(AWS, 2023)

54 regions worldwide 140 available in 140 countries



(Microsoft, 2023)

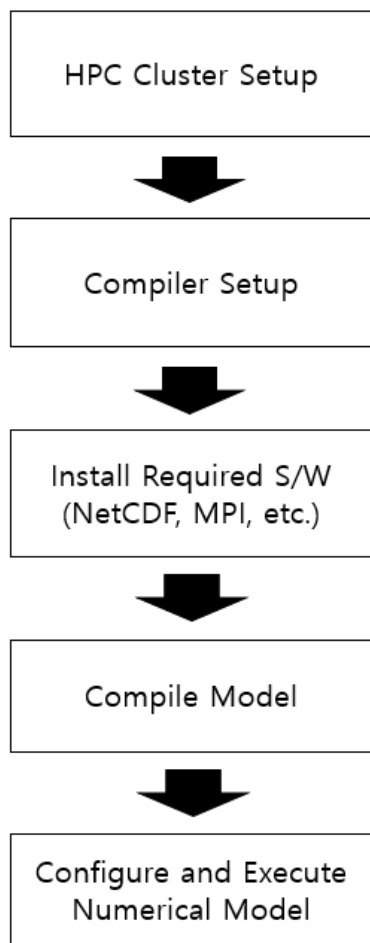


(Google, 2023)



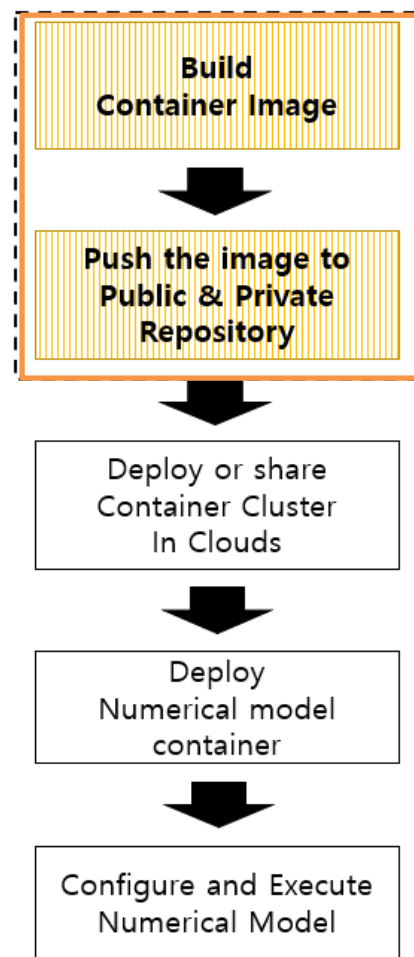
# Traditional HPC Cluster vs Ad-Hoc HPC Cluster

## Workflow of traditional HPC cluster



Workflow of technician or model developer for the sharable image

## Workflow of container-based HPC cluster



Workflow of end-user's setup and running the model in the public or private clouds

# How to create container Image and register?

```
147.46.82.230 ~ ubuntu@gluster: /XOMS/next7885/roms/next7885
File Edit Setup Control Window Help
# Install ubuntu
FROM          ubuntu:18.04
MAINTAINER   next7885@snu.ac.kr
RUN          apt-get -y update
RUN          apt-get install -y openssh-server

#replace sshd_config
RUN sed -ri 's/^#?PermitRootLogin\s+.*#PermitRootLogin yes/' /etc/ssh/sshd_config
RUN sed -ri 's/#UsePAM yes/#UsePAM yes/g' /etc/ssh/sshd_config

#make_ssh
RUN mkdir /root/.ssh
RUN mkdir -p /var/run/ssh

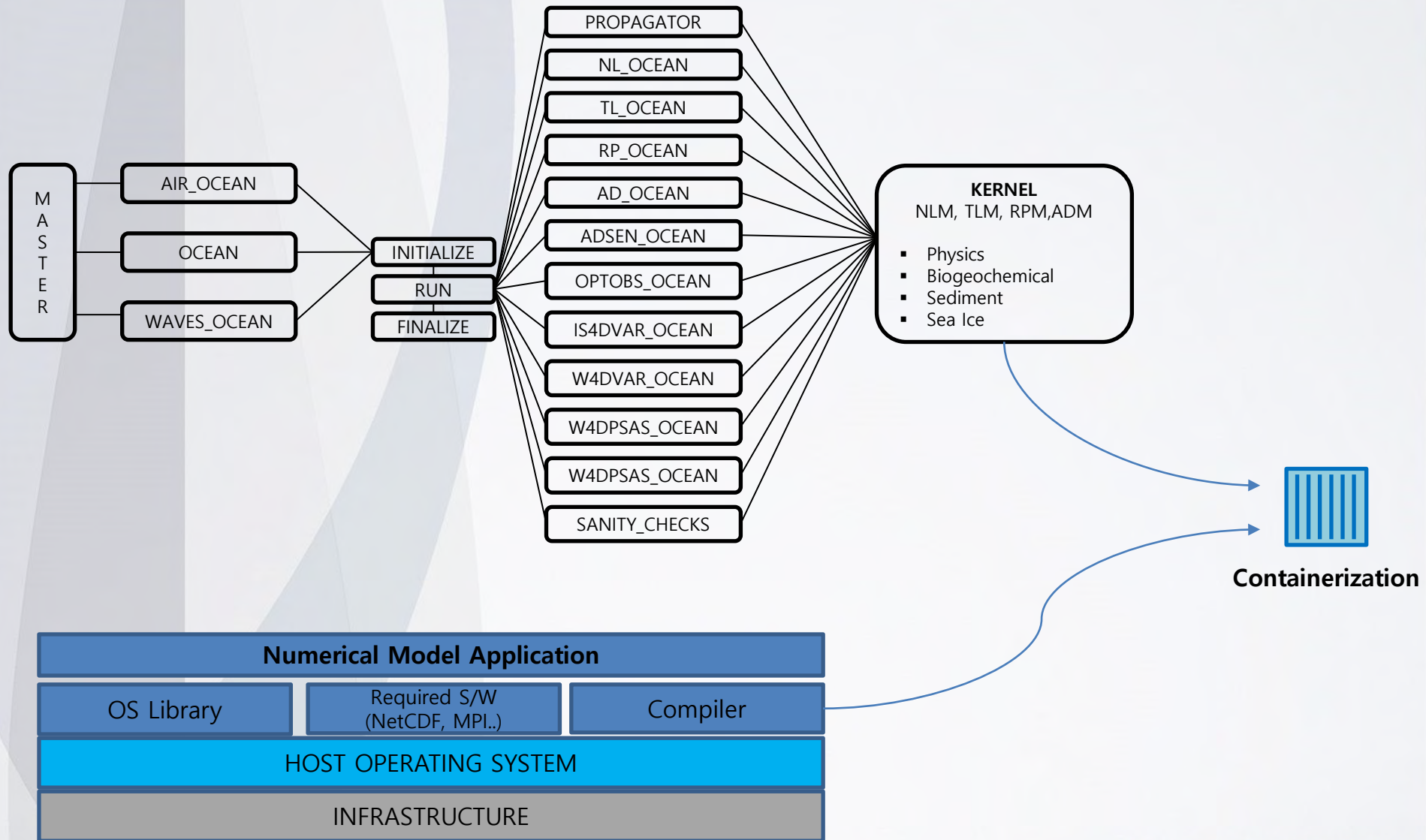
#install kubectl
RUN apt-get install -y curl
RUN curl -LO https://storage.googleapis.com/kubernetes-release/release/v1.18.0/bin/linux/amd64/kubectl
RUN chmod +x ./kubectl
RUN mv ./kubectl /usr/local/bin/kubectl
RUN kubectl version --client

# Install gcc
RUN apt-get -y install apt-utils
ENV DEBIAN_FRONTEND noninteractive
RUN apt-get -y install gcc
RUN apt-get -y install g++
RUN apt-get -y install gfortran
RUN apt-get -y install wget
RUN apt-get -y install file
RUN wget https://download.openmpi.org/release/open-mpi/v3.1/openmpi-3.1.4.tar.gz
RUN tar -xvf ./openmpi-3.1.4.tar.gz
```

The screenshot shows the Docker Hub interface for the user 'next7885'. The page lists several public repositories:

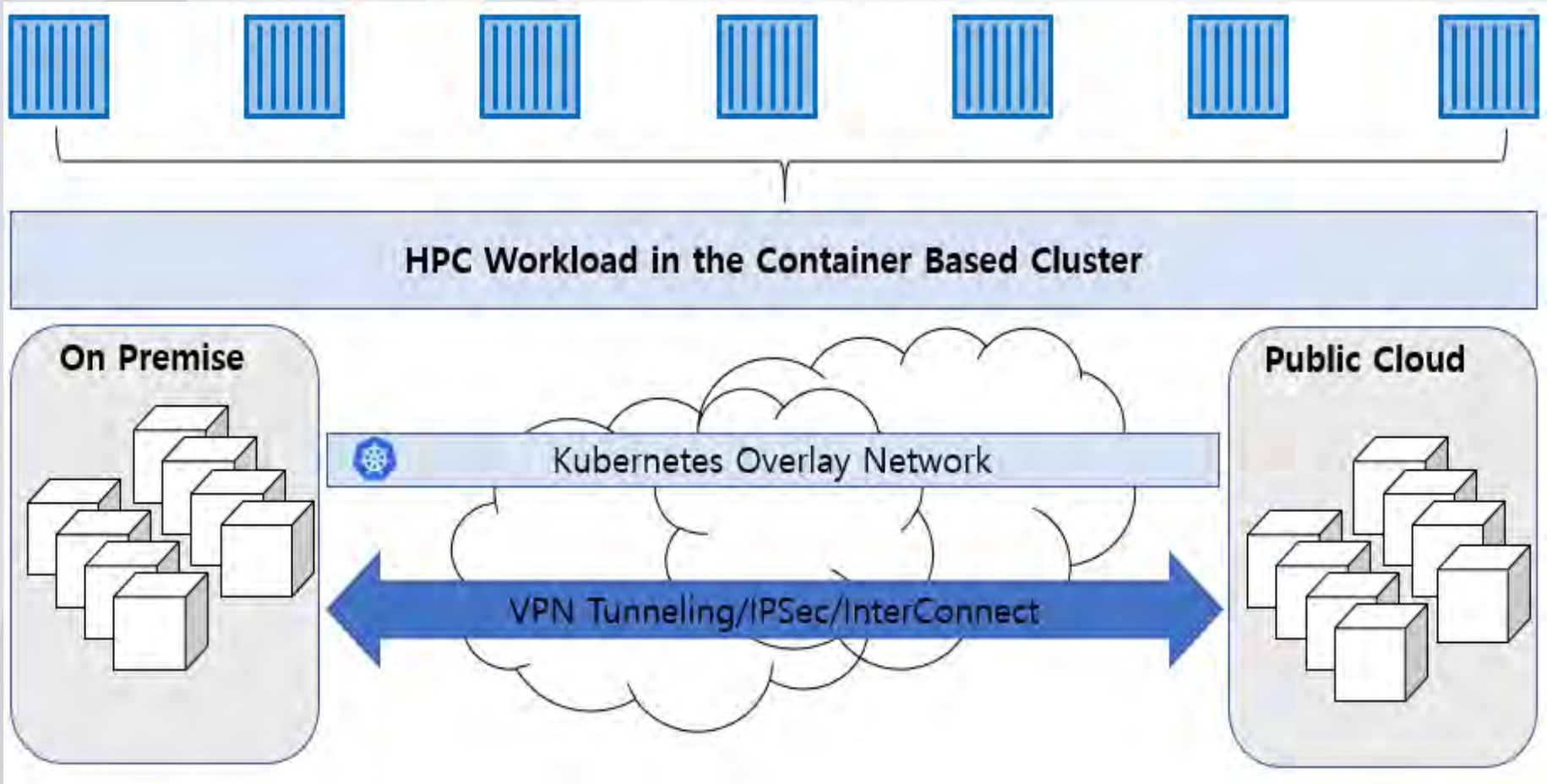
Repository Name	Scanned	Stars	Downloads	Visibility
next7885 / ubuntu_roms_k8s_hpc	Not Scanned	0	220	Public
next7885 / ubuntu_roms_k8s	Not Scanned	0	677	Public
next7885 / kube_docker	Not Scanned	0	180	Public
next7885 / ubuntu_mpi	Not Scanned	0	6.7K	Public
next7885 / ubuntu_ssh	Not Scanned	0	24	Public
next7885 / ubuntu_gfortran_mpi_netcdf_roms	Not Scanned	0	67	Public

# Containerization of Numerical Ocean Model(ROMS)

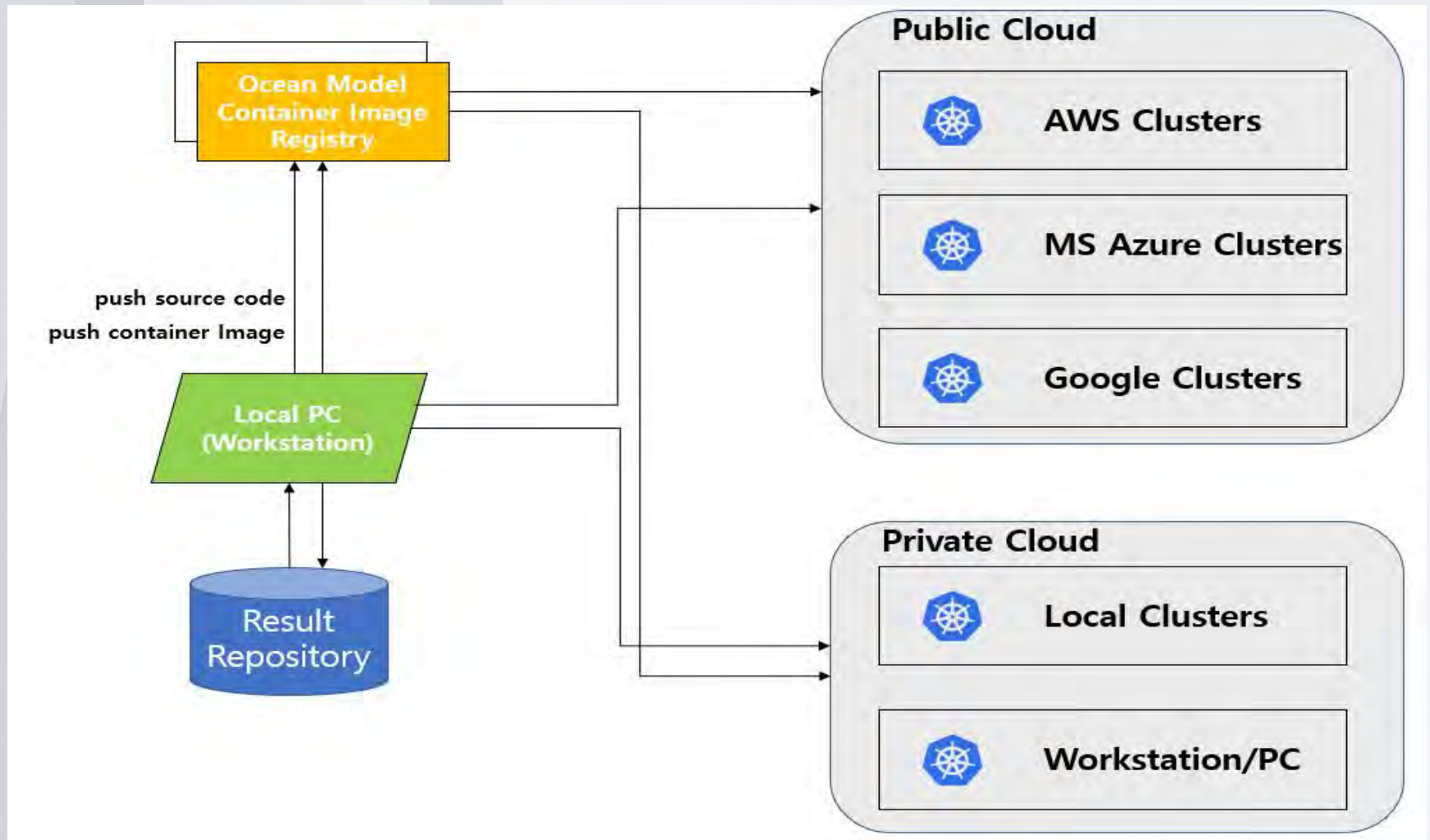


# Containerization of Numerical Ocean Model(ROMS)

- ▶ Increase the portability of numerical model



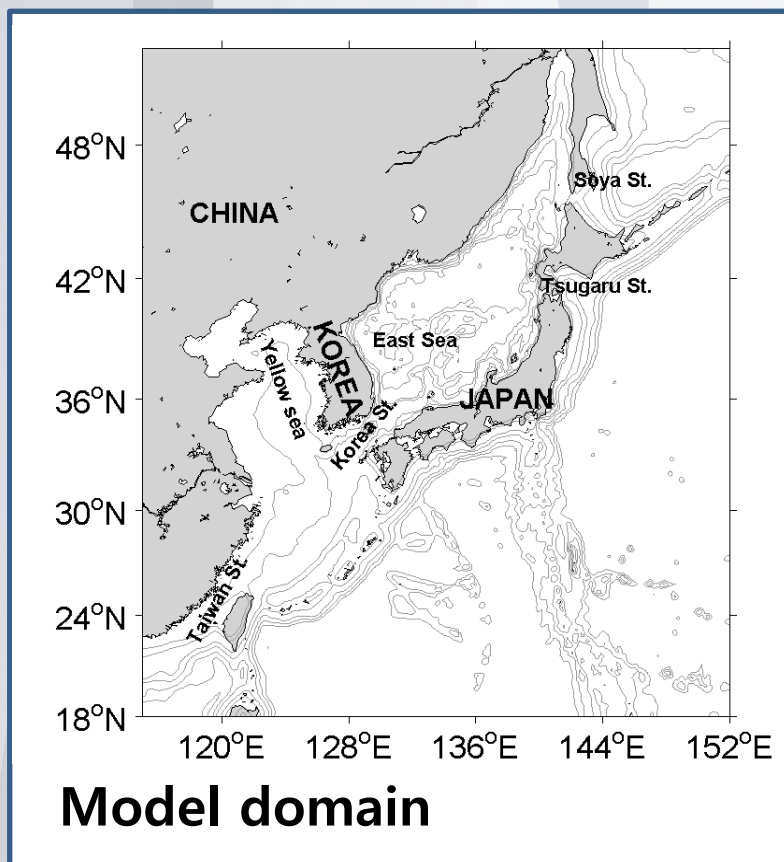
# Containerization of Numerical Ocean Model(ROMS)



## Interoperability of numerical modelling (System Configuration)

	Laptop-PC	Local Cluster#1	Local Cluster#2
<b>CPU Type</b>	Intel-i7	Intel Xeon	Intel Xeon
<b>Hypervisor</b>	VirtualBox	KVM	KVM
<b>Guest OS</b>	Windows 10	CentOS 7.5	Ubuntu 18.04
<b>Nodes</b>	1 Node	4 Nodes	3 Nodes
<b>vCores/Memory</b>		8Cores/16G	4Cores/8G
<b>Container Runtime</b>	Containerd	Docker v19.03.12	Docker v19.03.6
<b>N/W Interface</b>	LAN	LAN	LAN
<b>Orchestration Tool</b>	Microk8s (Kubernetes Compatible)	Kubernetes v1.18.3 minikube	Kubernetes v1.18.3
	Amazon-AWS	Google-GCP	MS-Azure
<b>CPU Type</b>	Intel Xeon, AMD	Intel Xeon	Intel Xeon, AMD
<b>OS</b>	Ubuntu 18.04, CentOS 7.5	Ubuntu 18.04	Ubuntu 18.04, CentOS 7.5
<b>Container Runtime</b>	Docker v19.03.6	Docker v19.03.6	Docker v19.03.6
<b>Orchestration Tool</b>	Kubernetes v1.18.3	Kubernetes v1.18.3	Kubernetes v1.18.3

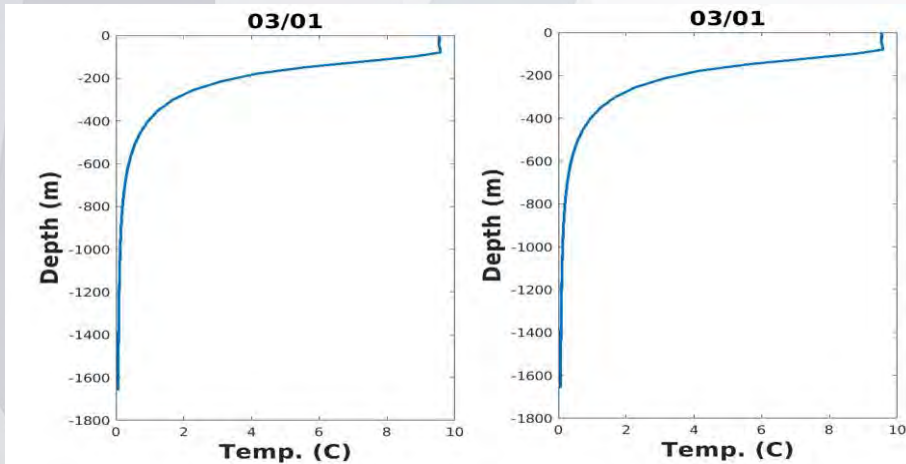
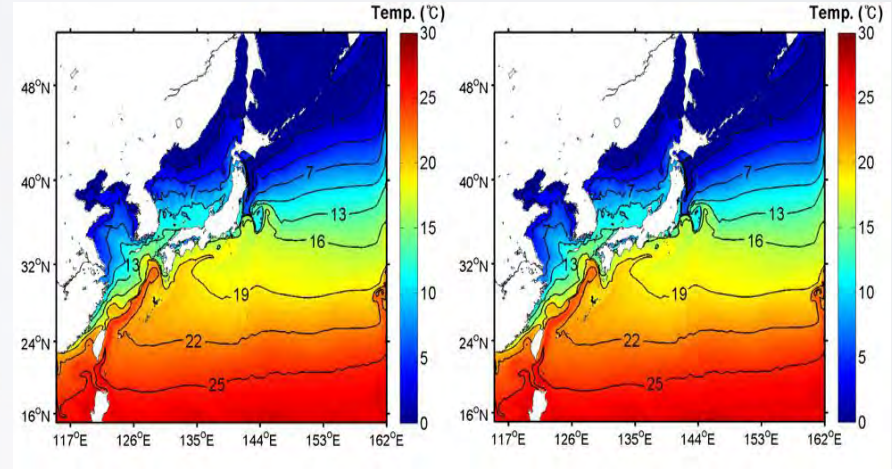
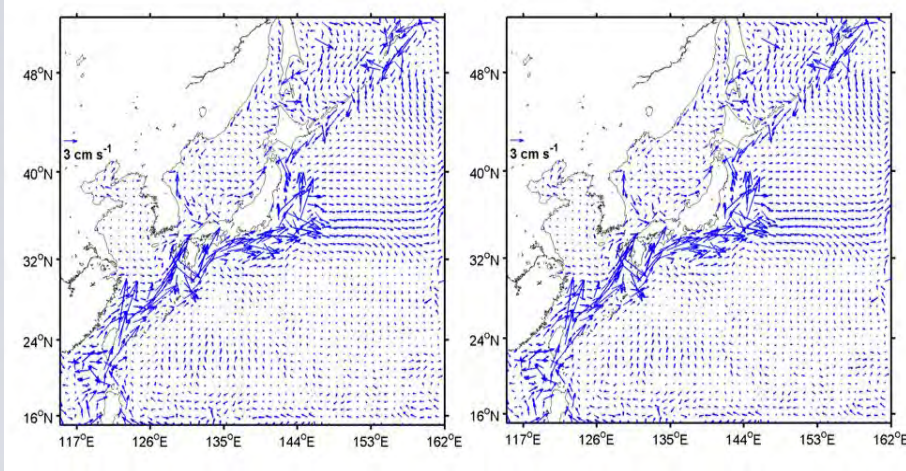
## ROMS modelling Configurations



<b>Model</b>	<b>ROMS v3.6</b>
<b>Resolution</b>	<b>5/10/20 km</b>
<b>Topography</b>	<b>ETOPO5</b>
<b>Vertical layer</b>	<b>40 layers</b>
<b>Initial condition</b>	<b>WOA 1998</b>
<b>Open boundary</b>	<b>SODA (2001-2010)</b>
<b>Surface boundary</b>	<b>ECMWF-interim (2010)</b>
<b>Tidal forcing</b>	<b>TPXO6</b>
<b>Vertical mixing scheme</b>	<b>KPP</b>
<b>Heat flux</b>	<b>Bulk flux parameterization</b>

Type	Coarse	Medium	Fine
<b>Dimension of grid</b>	210×206×40	422×412×40	846×826×40
<b>Degree of Freedom</b>	1,730,400	6,954,560	27,951,840

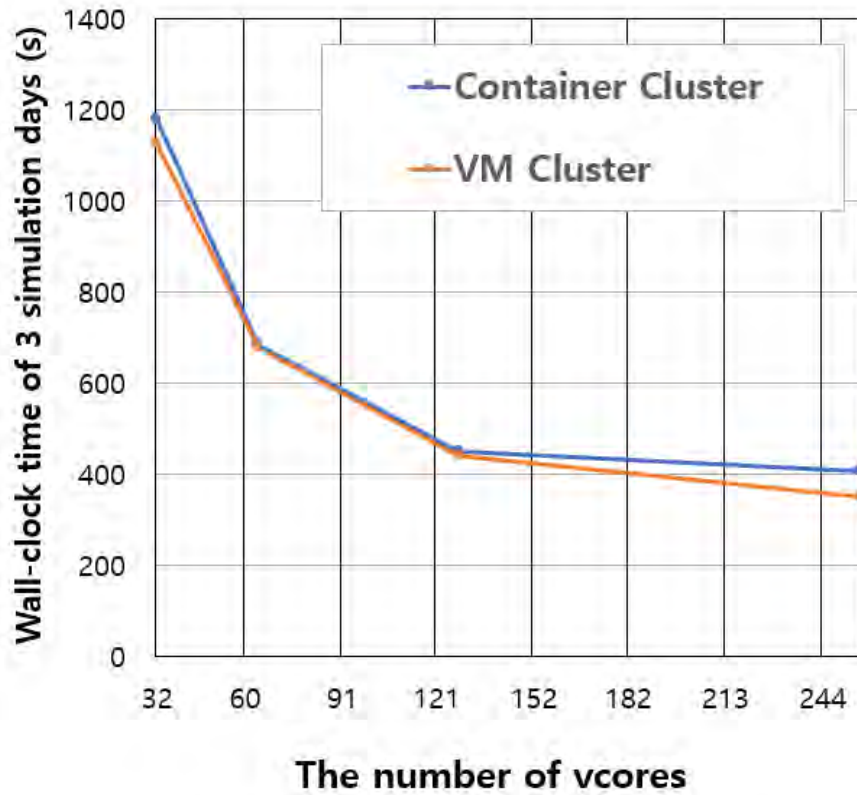
# Result-Reproducibility of numerical ocean model



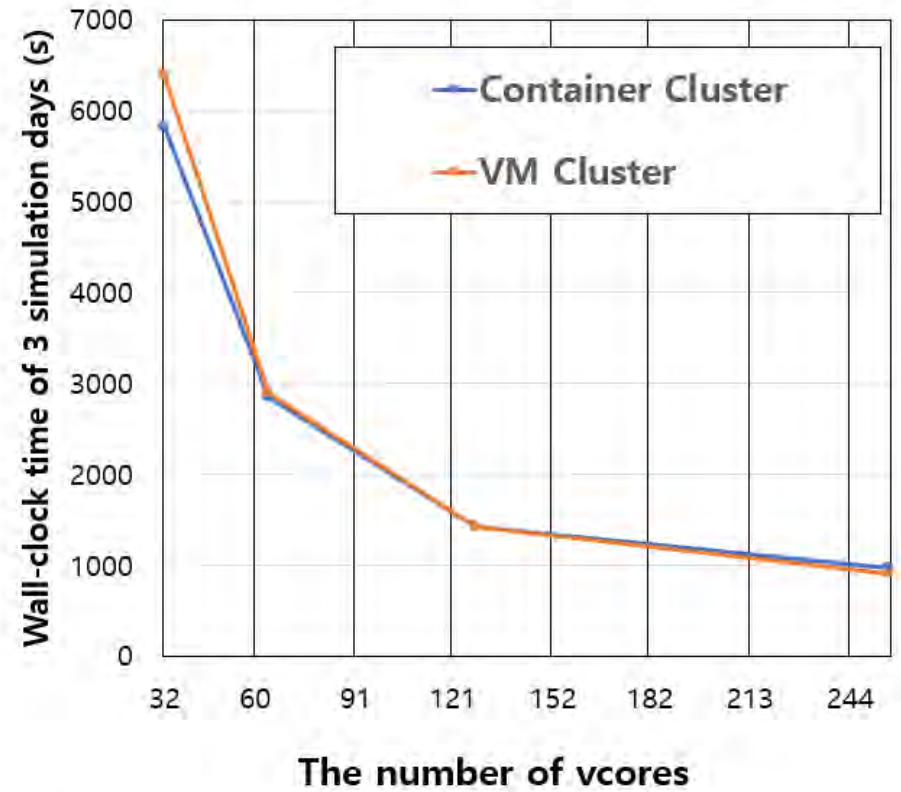
- Computational reproducibility  
All Root Mean Square Errors were commonly 0.0 °C for both SST and vertical temperature, and they were 0.0 ms<sup>-1</sup> along the surface velocity



# Result-Performance of Container Cluster

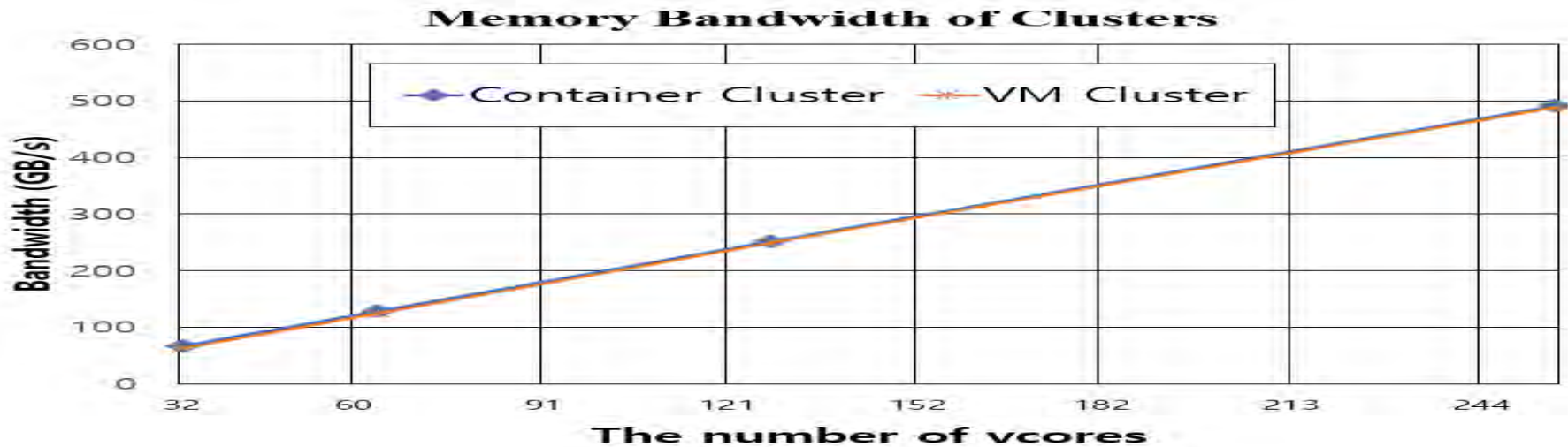
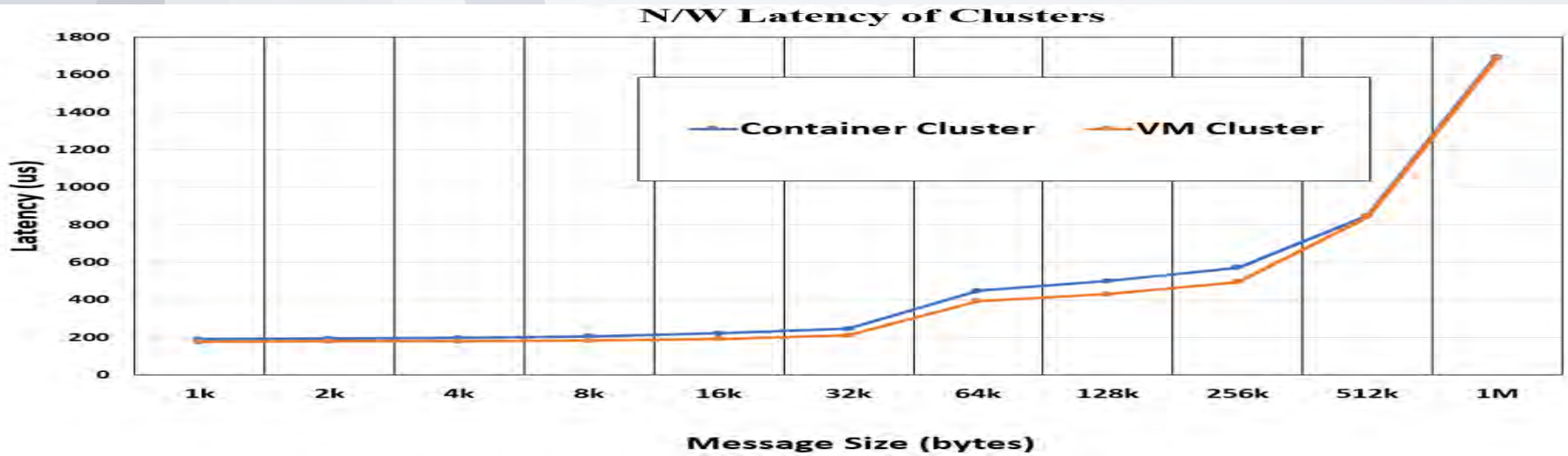


(a) Wall-clock time of medium grid simulation



(b) Wall-clock time of fine grid simulation

# Result-Performance of Container Cluster



- Design and implement containerization of numerical ocean model on the public and private clouds.
- Evaluate the performance of numerical ocean model with grids on the public clouds.
- Present a good alternative solution for the computational reproducibility of numerical model.



**Thank you  
for your attention**

**Q&A**

The screenshot displays the Google Cloud console interface for VM instances. The browser address bar shows the URL: `console.cloud.google.com/compute/instances?cloudshell=false&project=romscloud-184304`. The page title is "VM instances - Compute Engine". The left sidebar contains navigation options for Virtual machines, Storage, and Instance groups. The main content area is titled "VM instances" and includes a filter bar, a table of instances, and a "Related actions" section.

**VM instances Table:**

Status	Name	Zone	Recommendations	In use by	Internal IP	External IP	Connect
<input type="checkbox"/>	<a href="#">gcp-roms01</a>	us-west1-a			10.0.0.2 ( <a href="#">nic0</a> )		SSH ▾ ⋮
<input checked="" type="checkbox"/>	<a href="#">opst-demo-master</a>	us-west1-a			10.0.0.11 ( <a href="#">nic0</a> )	34.168.106.221 ( <a href="#">nic0</a> )	SSH ▾ ⋮
<input checked="" type="checkbox"/>	<a href="#">opst-demo-worker01</a>	us-west1-a			10.0.0.12 ( <a href="#">nic0</a> )	35.212.184.5 ( <a href="#">nic0</a> )	SSH ▾ ⋮

**Related actions:**

- Explore Backup and DR** (NEW): Back up your VMs and set up disaster recovery.
- View billing report**: View and manage your Compute Engine billing.
- Monitor VMs**: View outlier VMs across metrics like CPU and network.
- Explore VM logs**: View, search, analyze, and download VM instance logs.
- Set up firewall rules**: Control traffic to and from a VM instance.
- Patch management**: Schedule patch updates and view patch compliance on VM instances.
- Load balance between VMs**: Set up Load Balancing for your applications as your traffic and users grow.

Compute Engine - ROMSCLOUD

console.cloud.google.com/compute/instancesAdd?cloudshell=false&project=romscloud-184304

Google Cloud ROMSCLOUD Search (/) for resources, docs, products, and more

### Create an instance

To create a VM instance, select one of the options:

- New VM instance**  
Create a single VM instance from scratch
- New VM instance from template**  
Create a single VM instance from an existing template
- New VM instance from machine image**  
Create a single VM instance from an existing machine image
- Marketplace**  
Deploy a ready-to-go solution onto a VM instance

**Name**  
roms-test-k8s

**MANAGE TAGS AND LABELS**

**Region**  
us-central1 (Iowa)  
Region is permanent

**Zone**  
us-central1-a  
Zone is permanent

### Machine configuration

General purpose Compute optimized Memory optimized GPUs

Machine types for common workloads, optimized for cost and flexibility

Series	Description	vCPUs	Mem
C3	Consistently high performance	4 - 176	8 - 1.4
C3D	Consistently high performance	4 - 360	8 - 2.8
<b>E2</b>	Low cost, day-to-day computing	0.25 - 32	1 - 121
N2	Balanced price & performance	2 - 128	2 - 86
N2D	Balanced price & performance	2 - 224	2 - 891
T2A	Scale-out workloads	1 - 48	4 - 19
T2D	Scale-out workloads	1 - 60	4 - 241
N1	Balanced price & performance	0.25 - 96	0.6 - 6

**Monthly estimate**  
\$289.91  
That's about \$0.40 hourly  
Pay for what you use: no upfront costs and per second billing

Item	Monthly estimate
16 vCPU + 16 GB memory	\$288.91
10 GB balanced persistent disk	\$1.00
<b>Total</b>	<b>\$289.91</b>

[Compute Engine pricing](#)

**EQUIVALENT CODE**

```
gcloud compute instances create roms-test-k8s \
  --project=romscloud-184304 \
  --zone=us-central1-a \
  --machine-type=e2-highcpu-16 \
  --network-interface=network-tier=PREMIUM, \
  stack-type=IPV4_ONLY,subnet=default \
  --maintenance-policy=MIGRATE \
  --provisioning-model=STANDARD \
  --service-account=104048672002-compute@developer.gserviceaccount.com \
  --scopes=https://www.googleapis.com/auth/devstorage.read_only,https://www.googleapis.com/auth/logging.write,https://www.googleapis.com/auth/monitoring.write,https://www.googleapis.com/auth/servicecontrol,https://www.googleapis.com/auth/service.management.readonly,https://www.googleapis.com/auth/trace.append \
  --create-disk=auto-delete=yes,boot=yes,device-name=roms-test-k8s,image=projects/debian-cloud/global/images/debian-11-bullseye-v20231010,mode=rw,size=10,type=projects/romscloud-184304/zones/us-central1-a/diskTypes=pd-balanced \
  --no-shielded-secure-boot \
  --shielded-vm \
  --shielded-integrity-monitoring \
  --labels=goog-ec-spc=vm_add-gcloud \
  --reservation-affinity=any
```

**CREATE** **CANCEL** **EQUIVALENT CODE** **COPY** **RUN IN CLOUD SHELL** [View gcloud reference](#)

The screenshot displays the Google Cloud Console interface for a Compute Engine instance named 'opst-demo-worker01'. The 'OBSERVABILITY' tab is active, showing various performance metrics. The left sidebar contains navigation options for Virtual machines, Storage, Instance groups, Health checks, Marketplace, and Release Notes. The main content area features a 'METRICS' sidebar and six charts: CPU Utilization (at 100%), Memory Utilization (requires Ops Agent), Network Traffic (peaks at 5MB/s), Disk Space Utilization (requires Ops Agent), New Connections with VMs/External/G... (peaks at 1/s), and Disk Throughput (peaks at 200KIB/s). All charts share a common x-axis from 1:00 PM to 1:50 PM UTC+9. The top navigation bar includes options for EDIT, RESET, CREATE MACHINE IMAGE, CREATE SIMILAR, and START / RESUME.

ssh.cloud.google.com/v2/ssh/projects/romscloud-184304/zones/us-west1-a/instances/opst-demo-master?authuser=0&hl=en\_US&pro...

ssh.cloud.google.com/v2/ssh/projects/romscloud-184304/zones/us-west1-a/instances/opst-demo-master?authuser=0&hl=en\_...

SSH-in-browser      [↑ UPLOAD FILE](#)      [↓ DOWNLOAD FILE](#)      !      ⌨      ⚙

```
top - 04:48:12 up 15:51, 0 users, load average: 4.08, 4.19, 4.02
Tasks: 16 total, 5 running, 11 sleeping, 0 stopped, 0 zombie
%Cpu(s): 98.8 us, 1.1 sy, 0.0 ni, 0.2 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
KiB Mem : 15337764 total, 9541844 free, 2650084 used, 3145836 buff/cache
KiB Swap: 0 total, 0 free, 0 used. 12390868 avail Mem
```

PID	USER	PR	NI	VRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
210	roms	20	0	906892	584988	11940	R	99.3	3.8	44:43.42	gcc_nwp_no_acc
211	roms	20	0	901696	579084	11884	R	99.0	3.8	44:43.07	gcc_nwp_no_acc
209	roms	20	0	906896	585108	12060	R	98.3	3.8	44:43.20	gcc_nwp_no_acc
208	roms	20	0	929892	600620	14968	R	97.3	3.9	44:44.14	gcc_nwp_no_acc
285	roms	20	0	36624	3052	2588	R	0.3	0.0	0:00.02	top
1	root	20	0	4632	880	812	S	0.0	0.0	0:00.01	sh
15	root	20	0	72304	3196	2444	S	0.0	0.0	0:00.00	sshd
16	root	20	0	4536	812	748	S	0.0	0.0	0:00.00	sleep
17	root	20	0	18512	3416	2972	S	0.0	0.0	0:00.01	bash
33	root	20	0	60084	3368	2920	S	0.0	0.0	0:00.00	su
34	roms	20	0	20644	3340	2872	S	0.0	0.0	0:00.01	bash
71	roms	20	0	4572	900	840	S	0.0	0.0	0:03.40	tail
202	roms	20	0	358684	16108	13584	S	0.0	0.1	0:00.18	mpirun
257	root	20	0	18512	3316	2928	S	0.0	0.0	0:00.01	bash
272	root	20	0	60084	3368	2916	S	0.0	0.0	0:00.00	su
273	roms	20	0	18512	3388	2940	S	0.0	0.0	0:00.00	bash



ssh.cloud.google.com/v2/ssh/projects/romscloud-184304/zones/us-west1-a/instances/opst-demo-master?authuser=0&hl=en\_US&pro...

ssh.cloud.google.com/v2/ssh/projects/romscloud-184304/zones/us-west1-a/instances/opst-demo-master?authuser=0&hl=en\_...

SSH-in-browser      [↑ UPLOAD FILE](#)    [↓ DOWNLOAD FILE](#)

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
211	roms	20	0	901696	578292	11884	R	99.3	3.8	47:59.04	gcc_nwp_no_acc
209	roms	20	0	906896	584580	12060	R	99.0	3.8	47:59.50	gcc_nwp_no_acc
208	roms	20	0	929892	600620	14968	R	98.7	3.9	47:59.89	gcc_nwp_no_acc
210	roms	20	0	906892	584460	11940	R	98.0	3.8	47:59.51	gcc_nwp_no_acc
285	roms	20	0	36624	3052	2588	R	0.3	0.0	0:00.16	top
1	root	20	0	4632	880	812	S	0.0	0.0	0:00.01	sh
15	root	20	0	72304	3196	2444	S	0.0	0.0	0:00.00	sshd
16	root	20	0	4536	812	748	S	0.0	0.0	0:00.00	sleep
17	root	20	0	18512	3416	2972	S	0.0	0.0	0:00.01	bash
33	root	20	0	60084	3368	2920	S	0.0	0.0	0:00.00	su
34	roms	20	0	20644	3340	2872	S	0.0	0.0	0:00.01	bash
71	roms	20	0	4572	900	840	S	0.0	0.0	0:03.42	tail
202	roms	20	0	358684	16108	13584	S	0.0	0.1	0:00.18	mpirun
257	root	20	0	18512	3316	2928	S	0.0	0.0	0:00.01	bash
272	root	20	0	60084	3368	2916	S	0.0	0.0	0:00.00	su
273	roms	20	0	18512	3388	2940	S	0.0	0.0	0:00.00	bash

```

roms@roms-ssh-statefulset-0:~$
roms@roms-ssh-statefulset-0:~$
roms@roms-ssh-statefulset-0:~$ gcc
gcc: fatal error: no input files
compilation terminated.
roms@roms-ssh-statefulset-0:~$ █

```

```

ssh.cloud.google.com/v2/ssh/projects/romscloud-184304/zones/us-west1-a/instances/opst-demo-worker01?authuser=0&hl=en_US&pro...
ssh.cloud.google.com/v2/ssh/projects/romscloud-184304/zones/us-west1-a/instances/opst-demo-worker01?authuser=0&hl=en_...
SSH-in-browser
UPLOAD FILE
DOWNLOAD FILE
238518 ubuntu 20 0 906892 584460 11940 R 106.7 3.8 50:44.00 gcc_nwp_no_acc
238517 ubuntu 20 0 906896 584580 12060 R 100.0 3.8 50:43.91 gcc_nwp_no_acc
238519 ubuntu 20 0 901696 578292 11884 R 100.0 3.8 50:43.36 gcc_nwp_no_acc
251859 next7885 20 0 11248 3956 3272 R 6.7 0.0 0:00.01 top
1 root 20 0 169676 13160 8564 S 0.0 0.1 0:05.35 systemd
2 root 20 0 0 0 0 S 0.0 0.0 0:00.08 kthreadd
3 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 rcu_gp
4 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 rcu_par_gp
5 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 slub_flushwq
6 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 netns
8 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 kworker/0:0H-kblockd
10 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 mm_percpu_wq
11 root 20 0 0 0 0 S 0.0 0.0 0:00.00 rcu_tasks_rude_
12 root 20 0 0 0 0 S 0.0 0.0 0:00.00 rcu_tasks_trace
13 root 20 0 0 0 0 S 0.0 0.0 0:00.49 ksoftirqd/0
14 root 20 0 0 0 0 I 0.0 0.0 0:08.71 rcu_sched
15 root rt 0 0 0 0 S 0.0 0.0 0:00.38 migration/0
16 root -51 0 0 0 0 S 0.0 0.0 0:00.00 idle_inject/0
18 root 20 0 0 0 0 S 0.0 0.0 0:00.00 cpuhp/0
19 root 20 0 0 0 0 S 0.0 0.0 0:00.00 cpuhp/1
20 root -51 0 0 0 0 S 0.0 0.0 0:00.00 idle_inject/1
21 root rt 0 0 0 0 S 0.0 0.0 0:00.74 migration/1
22 root 20 0 0 0 0 S 0.0 0.0 0:00.63 ksoftirqd/1
24 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 kworker/1:0H-kblockd
25 root 20 0 0 0 0 S 0.0 0.0 0:00.00 cpuhp/2
26 root -51 0 0 0 0 S 0.0 0.0 0:00.00 idle_inject/2
27 root rt 0 0 0 0 S 0.0 0.0 0:00.73 migration/2
28 root 20 0 0 0 0 S 0.0 0.0 0:00.46 ksoftirqd/2
30 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 kworker/2:0H-events_highpri
31 root 20 0 0 0 0 S 0.0 0.0 0:00.00 cpuhp/3
32 root -51 0 0 0 0 S 0.0 0.0 0:00.00 idle_inject/3
33 root rt 0 0 0 0 S 0.0 0.0 0:00.75 migration/3
34 root 20 0 0 0 0 S 0.0 0.0 0:00.44 ksoftirqd/3
36 root 0 -20 0 0 0 I 0.0 0.0 0:00.00 kworker/3:0H-events_highpri
next7885@opst-demo-worker01:~$ gcc
Command 'gcc' not found, but can be installed with:
apt install gcc
Please ask your administrator.
next7885@opst-demo-worker01:~$ █

```