

# Scaling NWP workloads on AWS to achieve your research goals

aws

Timothy Brown, Sean Smith aws-hpc-weather@amazon.com

© 2022, Amazon Web Services, Inc. or its Affiliates.

## Agenda

- 1. AWS HPC Overview
- 2. Parallel Cluster Manager
- 3. Installing Applications
- 4. Application Performance Results



## AWS ParallelCluster





## x86 Instance Snapshot for Weather and Climate



#### Hpc6a.48xlarge

- OD: 3c/core-hr
- 1Y AURI: 2.21/core-hr
- 3Y AURI: 1.41c/core-hr

#### C6i.32xlarge

- OD: 8.5c/core-hr
- 1Y AURI: 5.43c/core-hr
- 3Y AURI: 3.26c/core-hr

#### C5n.18xlarge

- OD: 10.8c/core-hr
- 1Y AURI: 6.35c/core-hr
- 3Y AURI: 3.41c/core-hr

\* OD == On-Demand \* AURI == All Upfront Reserved Instances \* All pricing info is for Ohio (us-east-2) Region © 2022, Amazon Web Services, Inc. or its Affiliates. All rights reserved. Amazon Confidential and Trademark





# Elastic Fabric Adapter – Networks built to scale



- ✓ OS bypass
- ✓ GPUdirect and RDMA
- ✓ Libfabric core supports wide array of MPIs and NCCL



ECMP-enabled packet spraying and cloud-scale congestion control

#### 1.4B cell Siemens Simcenter STAR-CCM+ automotive CFD simulation



At ~40,000 cores (400 nodes), Hpc6a+EFA shows nearly 100% scaling efficiency





## **15-minutes to create a HPC Cluster**

အာ AWS ParallelCluste	Manager		<b>&amp; seaam@amazon.com ▼</b> us-east-2 ▼
×	Clusters (7)		+ Create Cluster
Home	Q		< 1 >
Clusters			
Custom Images	Name		✓ Version
Official Images	accounting	⊘ CREATE COMPLETE	3.1.2
	• weather-spack-cache		3.1.2
	Gromacs	UPDATE COMPLETE	3.1.2
	⊖ cfd	⊘ CREATE COMPLETE	3.1.2
avvs	O weather-2	⊘ CREATE COMPLETE	3.1.1
	hpc6a	UPDATE COMPLETE	3.0.0
	Cluster: weather-spack-cache     Image: storage       Details     Instances     Storage	Edit Stop Delete Filesystem	Shell DCV 💿 🗙
	Properties		
	cloudformationStackArn cloudformation:us-east- 2:822857487308:stack/weather-spack- cache/feac5c80-b443-11ec-8822-0a7eca5ee532	clusterStatus CREATE COMPLETE computeFleetStatus RUNNING	lastUpdatedTime 4/4/2022, 10:21:07 AM region us-east-2

Create-Cluster Wizard

# Connect via the browser (or SSH)

## Stop a cluster (preserve all data) Then resume later

...



## Slurm







# **Installing Software**

## Common ways

- Custom Amazon Machine Image (AMI) via packer
- Package managers (Spack, Easybuild, ...)
- By hand

https://geek-and-poke.com/geekandpoke/2010/5/14/how-to-become-invaluable.html



## Spack – Package Manager

### We are going to use Spack<sup>1</sup> to install all our packages.

- \$ export SPACK ROOT=/shared/spack
- \$ git clone -c feature.manyFiles=true https://github.com/spack/spack \$\$PACK ROOT
- \$ echo "export SPACK ROOT=/shared/spack" >> \$HOME/.bashrc
- \$ echo "source \\$SPACK ROOT/share/spack/setup-env.sh" >> \$HOME/.bashrc
- \$ source \$HOME/.bashrc

### Spack binary cache

https://binaries.spack.io/develop

- \$ spack mirror add binary mirror https://binaries.spack.io/develop
- \$ spack buildcache keys --install --trust

1. Todd Gamblin, Matthew P. LeGendre, Michael R. Collette, Gregory L. Lee, Adam Moody, Bronis R. de Supinski, and W. Scott Futral. The Spack Package Manager: Bringing Order to HPC Software Chaos.

In Supercomputing 2015 (SC'15), Austin, Texas, November 15-20 2015. LLNL-CONF-669890.

🔴 🔴 🌒 🏟 Spack — Spa	ck 0.18.0.dev0 doct ×	+								
- → C ŵ	C A https://spack	.readthedocs.io/en/latest/	E \$	⊻	*	969	٢	æ	<b>U</b>	-
🚸 Spa	ck	Docs » Spack			¢	<b>)</b> Edit	on G	iitHub	)	
latest		Spack								
Search docs		These are docs for the Spack package	e manager. For sphere p	acking	, see	oyspa	ck.			
BASICS Feature Overview Getting Started Basic Usage Tutorial: Spack 101 Using Spack to Replace Homebrew/Conda		Spack is a package management tool of configurations of software on a wide w designed for large supercomputing cer- share common installations of softwar libraries that do not have a standard A version does not break existing installa same system.	lesigned to support mi ariety of platforms and iters, where many use e on clusters with exoi Bl. Spack is non-destri ations, so many configu	ultiple d envir rs and tic arcl uctive: uratior	versie onme appli hitect insta	ons a ents. I cation ures, illing a n coex	nd t was n tear using a new kist or	ns ; / n the		
REFERENCE Configuration Files Basic Settings		Most importantly, Spack is simple. It of versions and configuration options cor package files are written in pure Pytho single file for many different builds of	fers a simple <i>spec</i> synt ncisely. Spack is also si n, and specs allow pac the same package.	ax so mple f kage a	that u or pa autho	isers ckage rs to	can si e auth maint	pecify Iors: ain a	/	
Build Customization Environments		See the Feature Overview for example	s and highlights.							
Container Images Monitoring		Get spack from the github repository a	ind install your first pa	ckage						
Mirrors Modules		<pre>\$ git clone -c feature.manyFiles=tr \$ cd spack/bin \$ ./spack install libelf</pre>	ue https://github.com,	spack,	/spac⊦	.git				
Package Repositories Build Caches Command Reference		If you're new to spack and want to sta manual below.	rt using it, see Getting	Starte	d, or	refer	to the	e full		
	un lataat -	Basics								

Featu Gettin

Home

Confi

Enviro Conta Monit Mirro Modu Packa Build C Comm Read t



## Spack External Packages

## We define the external packages in \$SPACK\_ROOT/etc/spack/packages.yam1

	packages:
2	libfabric:
3	<pre>variants: fabrics=efa,tcp,udp,sockets,verbs,shm,mrail,rxd,rxm</pre>
4	externals:
5	- spec: libfabric@1.16.0 fabrics=efa,tcp,udp,sockets,verbs,shm,mrail,rxd,rxm
6	<pre>prefix: /opt/amazon/efa</pre>
7	buildable: False



## **UFS Weather Model**

- Version 2.0.0
- Global C768 with 65 vertical levels
- 2 day forecast
- Model timestep is 150 seconds



## $C768 \rightarrow 768$ cells per direction per face $\rightarrow 13$ km resolution



## **Building UFS Weather Model With Spack**

spack install –j 12 ufs-weather-model%intel^intel-oneapi-mpi+external-libfabric

Spack Arguments and Flags Description				
install	Install a nackaga			
Install	install a package.			
-j 12	Parallel build with 12 cores.			
ufs-weather-model	The package to install.			
%intel	Specify the Intel compiler (icc and ifort).			
^intel-oneapi-mpi	Use the Intel OneAPI MPI library.			
+external-libfabric	Use an external fabric for MPI.			

# **Running UFS Weather Model**

### I\_MPI\_OFI\_LIBRARY\_INTERNAL must be set to 0 before loading Intel MPI \_\_\_\_\_\_ Specify EFA as the fabric adaptor \_\_\_\_\_

Confirm EFA is used, when I\_MPI\_DEBUG=4 (slurm job output text)

#### #!/bin/bash

```
#SBATCH --job-name=
```

- #SBATCH --output=slurm-ufs-%j.out
- 5 #SBATCH --nodes=150
- 6 #SBATCH --ntasks-per-node=24

```
7 #SBATCH --exclusiv
```

```
ouroant
```

```
export I_MPI_DEBUG=4
export I MPI FABRICS=ofi
```

- 11 export I MPI OFI LIBRARY\_INTERNAL=0
- 12 export I MPI OFI PROVIDER=efa
- 13 export I\_MPI\_PIN\_DOMAIN=omp
- 14 export KMP\_AFFINITY=compact
- 15 export OMP\_NUM\_THREADS=4

```
.6
```

- 17 spack load intel-oneapi-mpi
- 18 spack load ufs-weather-model
- 19 module load libfabric-aws
- 20 set -x
- 21 ulimit -s unlimited
- 22 ulimit -a
- 23
- 24 time mpiexec.hydra ufs\_weather\_model

1 [0] MPI startup(): libfabric version: 1.16.0~amzn3.0

2 [0] MPI startup(): libfabric provider: efa



## Scale Up Performance

Time to Results UFS v2.0.0, C768, 2 day forecast



## Scale Up Cost

Cost to Results UFS v2.0.0, C768, 2 day forecast





## Weather Workshops <a href="https://weather.hpcworkshops.com">https://weather.hpcworkshops.com</a>





# Thank you

If you have any additional questions, please feel free to reach out.

Timothy Brown, Sean Smith aws-hpc-weather@amazon.com

© 2022, Amazon Web Services, Inc. or its Affiliates.

## Weather Models Run On AWS

- WRF ARW
- MPAS Atmosphere
- UFS Atmosphere (FV3GFS)
- Unified Model
- Harmonie
- ICON