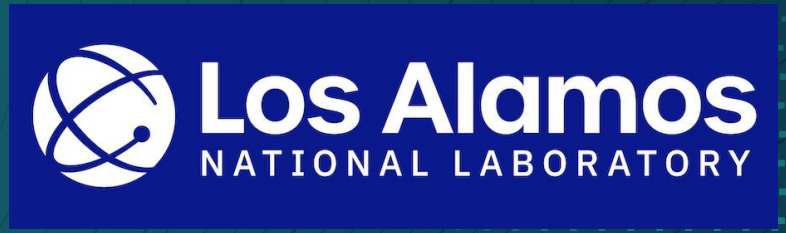


Characterizing the impact of compiler and MPI version differences in Containers with Spack

David Bernado[1] and Martha Dix[2]

Mentored by: Jordan Ogas, Paul Ferrell, Reid Priedhorsky, Nick Sly and Megan Phinney

1 

bernadda@whitman.edu
Whitman College

2 

martha.k.dix@gmail.com
St. Olaf College

LA-UR-21-28004

What Are Containers?

- Isolated Linux namespaces
- Useful for running apps across different machines in a tailored, isolated environment
- Charliecloud = unprivileged container manager, built for HPC

Granular Package Management

- Spack = unprivileged package manager
- User controls own software stack instead of system admin
- Different versions of a package can be installed and switched between at will

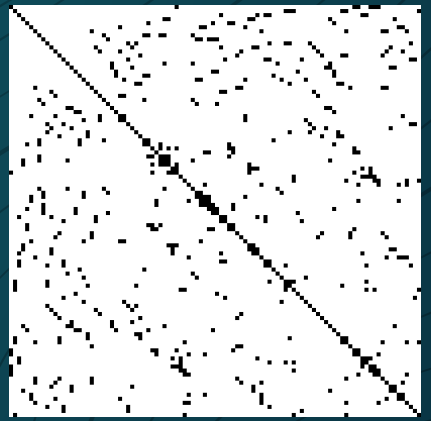
CTS-2 Benchmarks

- Serves as acceptance tests for next generation hardware
- 4 Figure of Merit benchmarks
  - HPCG, LAGHOS, Quicksilver, SNAP
- 4 microbenchmarks
- Simulates computational workload of scientific applications

HPCG

HPCG
High Performance Conjugate Gradients

- Used in supercomputer ranking
- Multigrid preconditioned conjugate gradient algorithm
- Calculates metrics around problem aspects like:
  - Symmetric Gauss-Seidel multigrid method
  - Sparse matrix vector multiplications
  - Dot product
  - Vector updates with sum of 2 scaled vectors

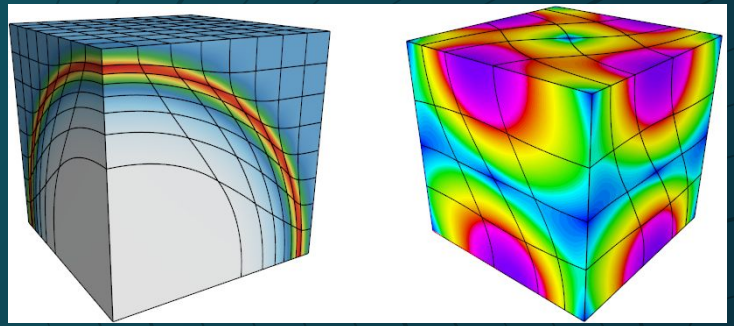


Sparse Matrix

= Non zero number

LAGHOS

- Mimics simulation of compressible gas dynamics
- Good proxy for scientific apps
- Measures performance in execution time of certain algorithms
  - Inversion of global kinematic mass matrix
  - Inv. of local thermodynamic mass matrices
  - Force evaluation
  - Physics calculation



The Question:

How do different versions of GCC and OpenMPI impact the performance of apps running inside containers?

1. Which versions perform better than others?
2. What is the optimal combo of GCC/OpenMPI versions?
3. How do the containers perform vs. bare-metal?

Related Work

- Kovács 2017, Le and Paz 2017, Younge et al. 2017, Torrez et al. 2019 all found containers compared to bare metal system have minimal or no impact
- To our knowledge, no previous studies have investigated how version combinations affect benchmark performance

Testing Setup

		GCC Version			
		9.3.0	10.3.0	11.1.0	Bare Metal (9.3.0)
	3.1.6				
OpenMPI	4.0.6				
Version	4.1.1				
	Bare Metal (3.1.6)				

(Or “bear metal”, if you prefer)

Testing Setup

Bare Metal OpenMPI:

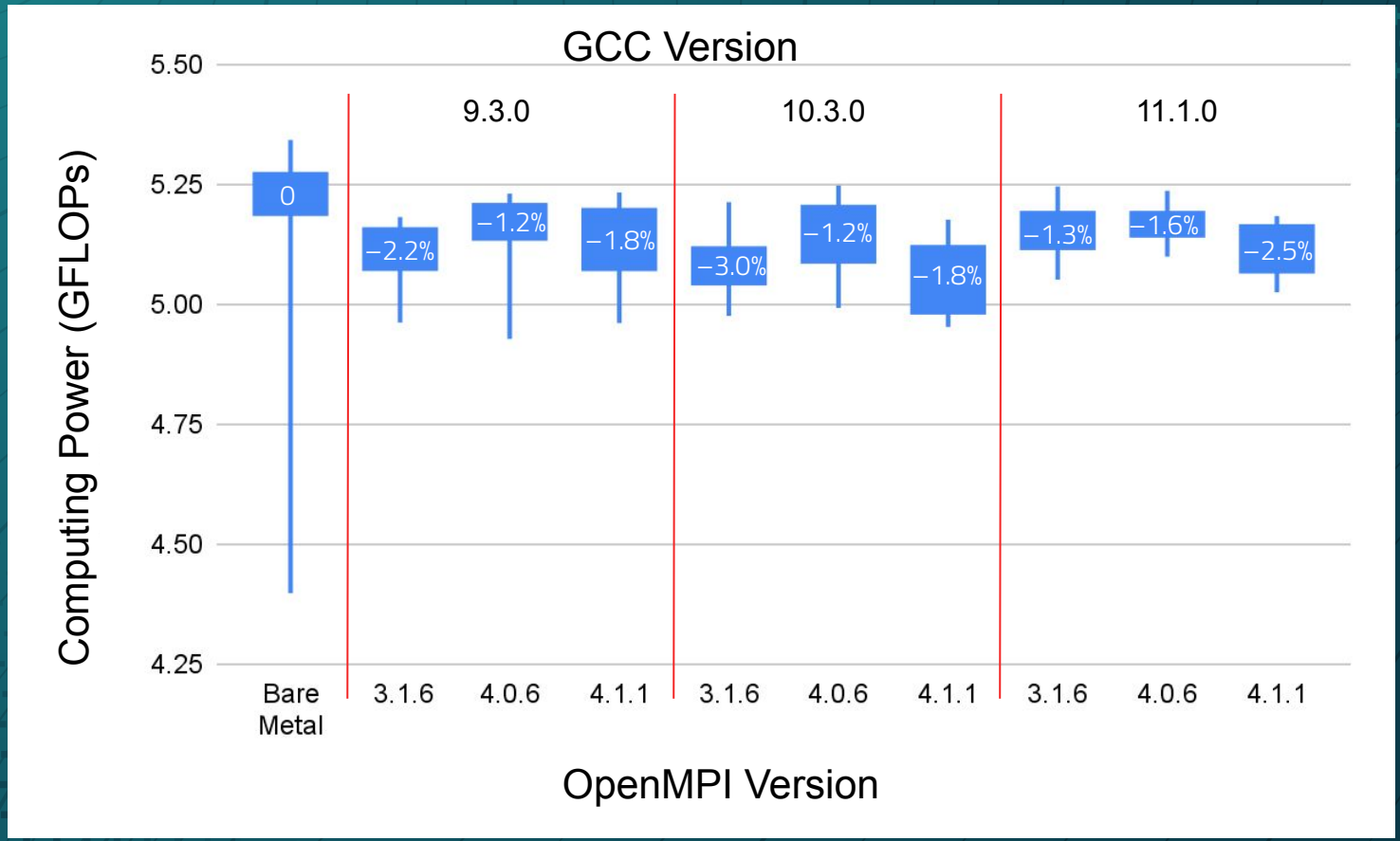
- --with-slurm
- --with-pmix
- --with-ucx
- --without-psm2
- --with-verbs

Spack OpenMPI:

- --with-slurm
- --with-pmix
- --without-ucx
- --with-psm2
- --without-verbs

Results

HPCG 1-Node: Computing Power



Higher is Better

~0.02 GFLOP difference between Bare Metal and Versions

Statistical Tests

- T-test - Determines if there is a significant difference between 2 groups
- If significant, allow us to say that the difference did not occur by chance (but could occur for other reasons)
- Level of significance 0.05
- Bonferroni correct - corrects for multiple comparisons
- Statistically significant if:
  - HPCG- p-value < 0.00005 (corrected for 1012 comparisons)
  - LAGHOS- p-value < 0.0001 (corrected for 450 comparisons)

    p-value < 0.0008 (corrected for 60 comparisons)

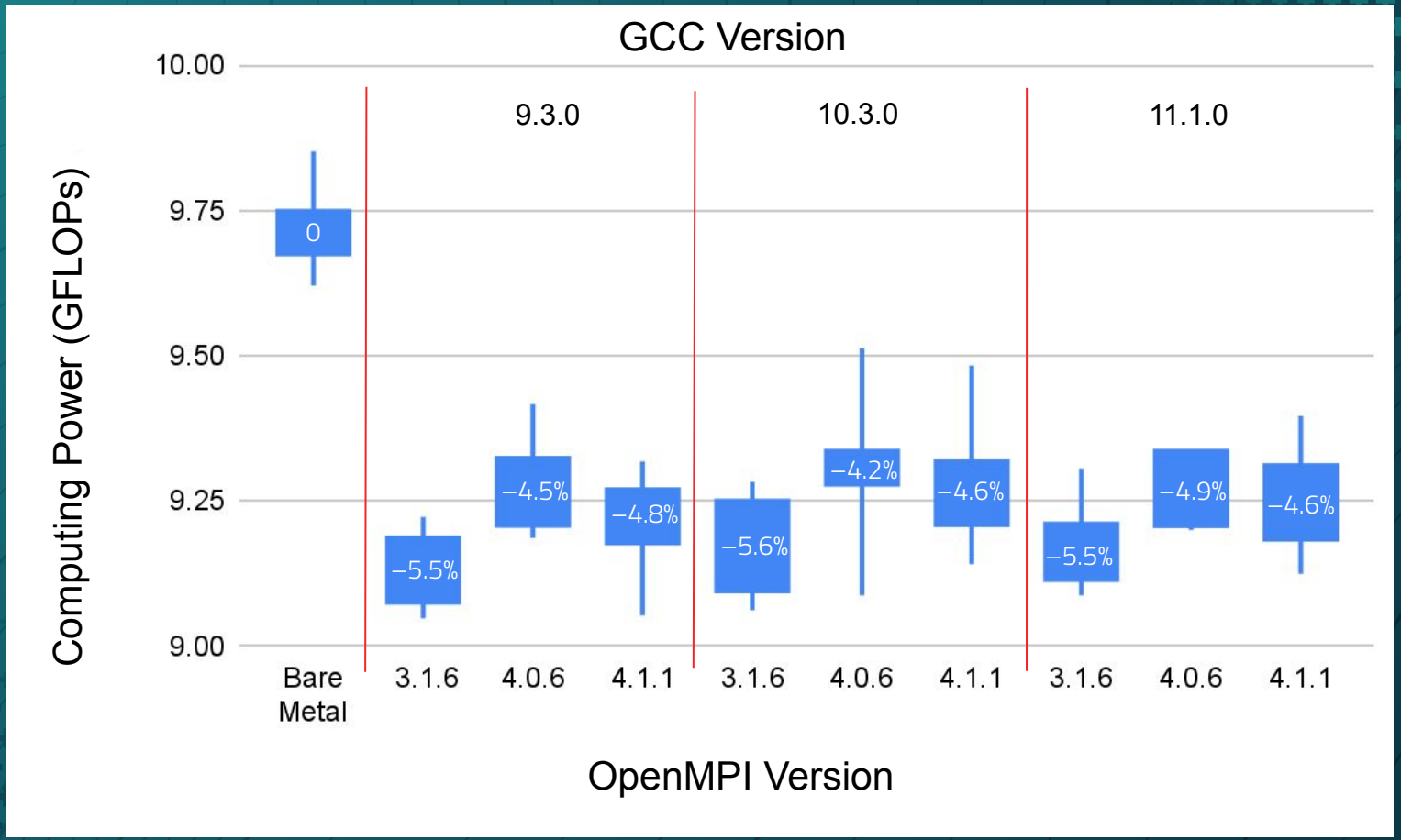
HPCG 1-Node T-Test P-values

= significant
p-value < 0.00005

= not significant

GCC-OpenMPI	Bare Metal (9.3.0-3.1.6)	9.3.0-3.1.6	9.3.0-4.0.6	9.3.0-4.1.1	10.3.0-3.1.6	10.3.0-4.0.6	10.3.0-4.1.1	11.1.0-3.1.6	11.1.0-4.0.6	11.1.0-4.1.1
Bare Metal										
9.3.0-3.1.6										
9.3.0-4.0.6										
9.3.0-4.1.1										
10.3.0-3.1.6										
10.3.0-4.0.6										
10.3.0-4.1.1										
11.1.0-3.1.6										
11.1.0-4.0.6										
11.1.0-4.1.1										

HPCG 2-Node: Computing Power



~0.5 GFLOP difference between Bare Metal and Versions

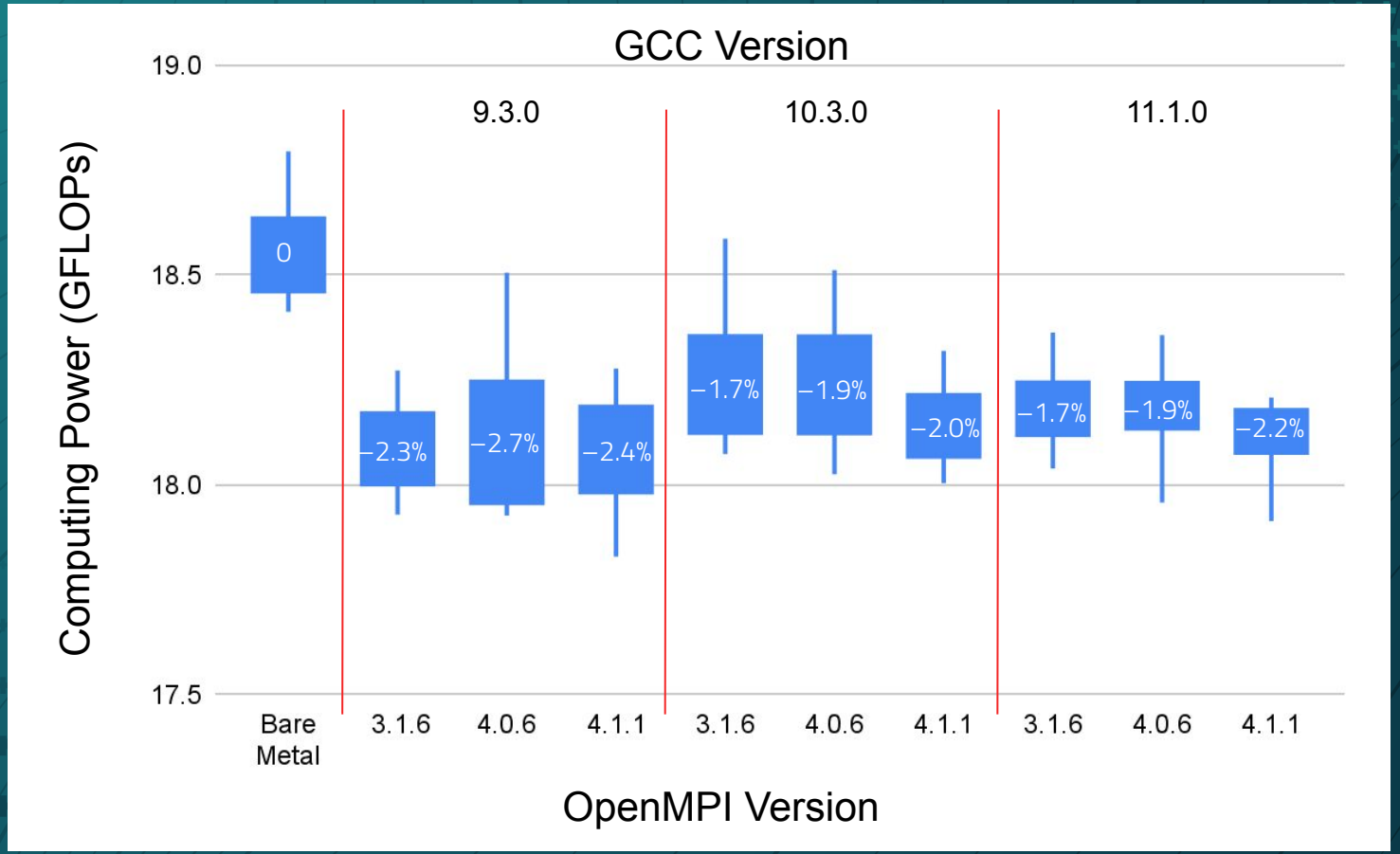
HPCG 2-Node T-Test P-values

= significant
p-value < 0.00005

= not significant

GCC-OpenMPI	Bare Metal	9.3.0-3.1.6	9.3.0-4.0.6	9.3.0-4.1.1	10.3.0-3.1.6	10.3.0-4.0.6	10.3.0-4.1.1	11.1.0-3.1.6	11.1.0-4.0.6	11.1.0-4.1.1
Bare Metal		√	√	√	√	√	√	√	√	√
9.3.0-3.1.6										
9.3.0-4.0.6										
9.3.0-4.1.1										
10.3.0-3.1.6										
10.3.0-4.0.6										
10.3.0-4.1.1										
11.1.0-3.1.6										
11.1.0-4.0.6										
11.1.0-4.1.1										

HPCG 4-Node: Computing Power



~0.5 GFLOP difference between Bare Metal and Versions

HPCG 4-Node T-Test P-values

= significant
p-value < 0.00005

= not significant

GCC-OpenMPI	Bare Metal	9.3.0-3.1.6	9.3.0-4.0.6	9.3.0-4.1.1	10.3.0-3.1.6	10.3.0-4.0.6	10.3.0-4.1.1	11.1.0-3.1.6	11.1.0-4.0.6	11.1.0-4.1.1
Bare Metal		√	√	√	0.0001	p < 0.0001	√	√	√	√
9.3.0-3.1.6										
9.3.0-4.0.6										
9.3.0-4.1.1										
10.3.0-3.1.6										
10.3.0-4.0.6										
10.3.0-4.1.1										
11.1.0-3.1.6										
11.1.0-4.0.6										
11.1.0-4.1.1										

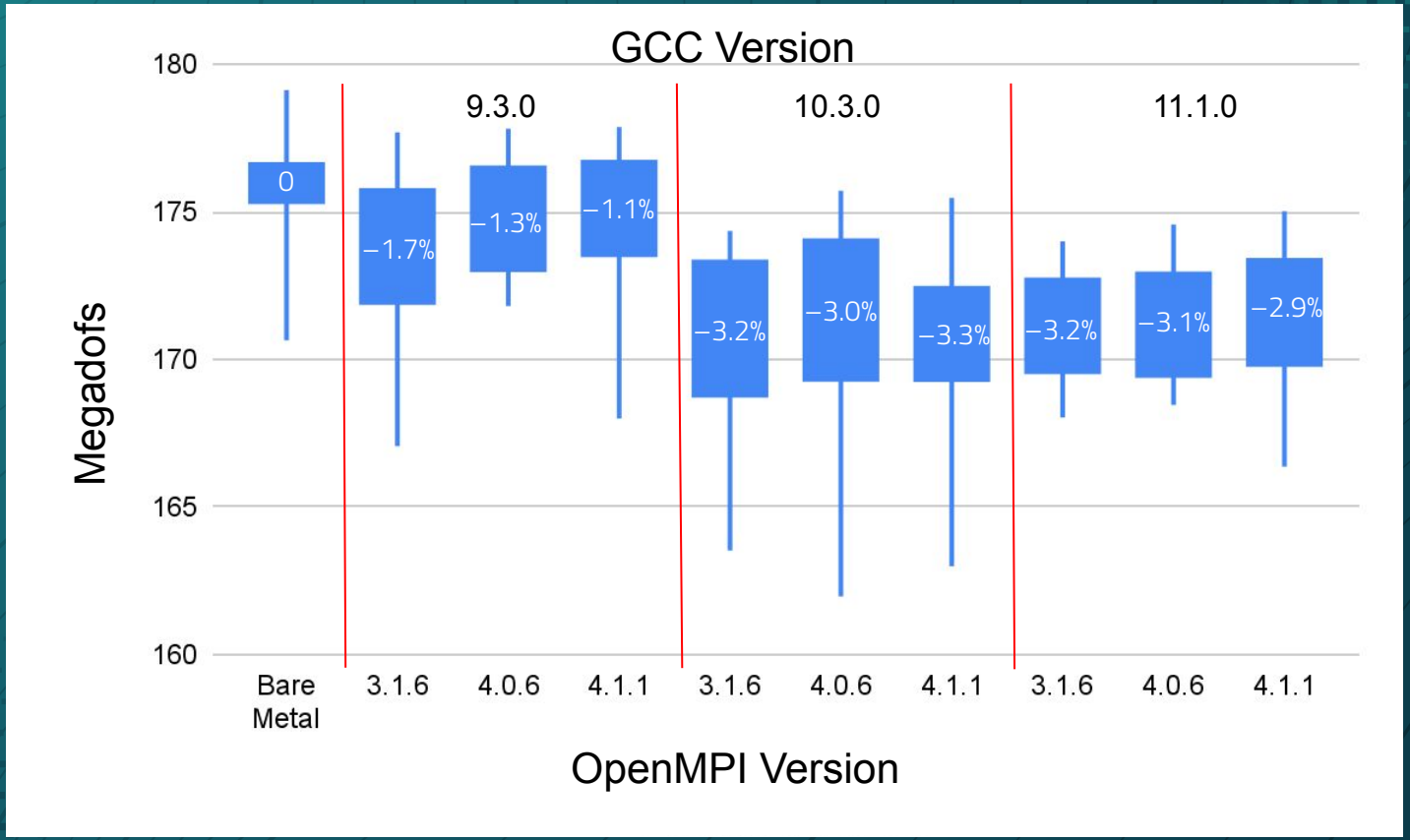
HPCG Summary

1. **Which versions perform better than others?**
   - ➔ GCC or OpenMPI version appear to have no effect.

2. **What is the optimal combo of GCC/OpenMPI versions?**
   - ➔ N/A

3. **How do the containers perform vs. bare-metal?**
   - ➔ On multi-node configurations there was an average of ~6% decrease in performance for 40% (12 of 29) metrics when containers were used.
   - ➔ The greater difference seen here than in other similar studies could be due to lack of MPI implementation optimization and use of built in OpenMPI versus the Spack OpenMPI used in the containers.

LAGHOS Results: Update Quad Data Rate



Higher is Better

LAGHOS T-Test P-values

= significant
p-value < 0.0001

= not significant

GCC-OpenMPI	Bare Metal	9.3.0-3.1.6	9.3.0-4.0.6	9.3.0-4.1.1	10.3.0-3.1.6	10.3.0-4.0.6	10.3.0-4.1.1	11.1.0-3.1.6	11.1.0-4.0.6	11.1.0-4.1.1
Bare Metal		√			√	√	√	√	√	√
9.3.0-3.1.6					√		√	√	√	
9.3.0-4.0.6					√	√	√	√	√	√
9.3.0-4.1.1					√	√	√	√	√	√
10.3.0-3.1.6										
10.3.0-4.0.6										
10.3.0-4.1.1										
11.1.0-3.1.6										
11.1.0-4.0.6										
11.1.0-4.1.1										

LAGHOS T-Test Grouped P-values

= significant
p-value < 0.0008

= not significant

GCC Version

GCC	Bare Metal	9.3.0	10.3.0	11.1.0
Bare Metal		√	√	√
9.3.0			√	√
10.3.0				
11.1.0				

OpenMPI Version

OpenMPI	Bare Metal	3.1.6	4.0.6	4.1.1
Bare Metal				
3.1.6				
4.0.6				
4.1.1				

LAGHOS Summary

1. **Which versions perform better than others?**
   - ➔ GCC version slightly affected 70% (8 of 11) of the metrics
   - ➔ Versions 10.3.0 and 11.1.0 decreased average performance by ~2%

2. **What is the optimal combo of GCC/OpenMPI versions?**
   - ➔ GCC version 9.3.0 outperformed the others

3. **How do the containers perform vs. bare-metal?**
   - ➔ Containers only slightly hindered 30% (4 of 11) of the metrics in an average performance decrease of ~2%.
   - ➔ The greater difference seen here than in other similar studies could be due to lack of MPI implementation optimization and use of build in OpenMPI versus the Spack openMPI used in the containers.

Future Work

- Optimize MPI implementation
- Test the other figure of merit benchmarks (SNAP and Quicksilver)
- Look at performance beyond 4 nodes
- Examine other compilers and MPIs

Acknowledgements

We would like to thank our mentors, Jordan Ogas, Paul Ferrell, Reid Priedhorsky, Nick Sly and Megan Phinney, for their help, advice and mentorship throughout our project.

We would also like to thank Julie Wiens, Evan Donato, Francine Lapid, Travis Cotton and all who were involved in organizing the Supercomputer Institute.

Works Cited

Ákos Kovács, “Comparison of different Linux containers,” in *Proc. TSP*, 2017.

E. Le and D. Paz, “Performance analysis of applications using Singularity container on SDSC Comet,” in *Proc. PEARC*, 2017.

A.J. Younge *et al.*, “A tale of two systems: Using containers to deploy HPC applications on supercomputer and clouds,” in *Proc. CloudCom*, 2017.

Torrez *el al.*, “HPC container runtimes have minimal or no performance impact,” in *IEE CANOPIE-HPC*, 2019.

External images used in this presentation:

https://ih0.redbubble.net/image.7772971.2289/flat,1000x1000,075,f.jpg

Thank You!

David Bernado
bernadda@whitman.edu

Martha Dix
martha.k.dix@gmail.com