MPI Performance Analysis

Dr Michèle Weiland, Senior Research Fellow

m.weiland@epcc.ed.ac.uk

THE UNIVERSITY of EDINBURGH

Thanks to Nick Brown @EPCC for all the results and data that are presented here!

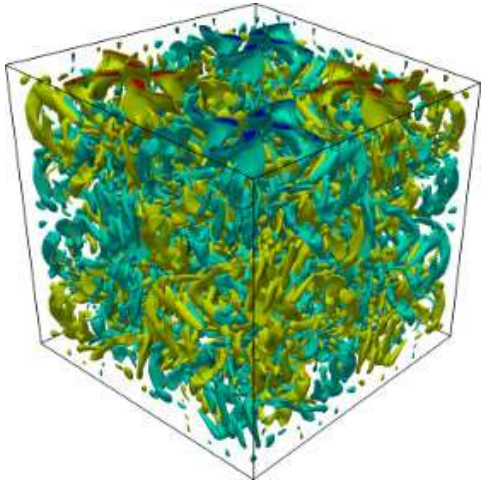
epcc

Fulhame: a research system

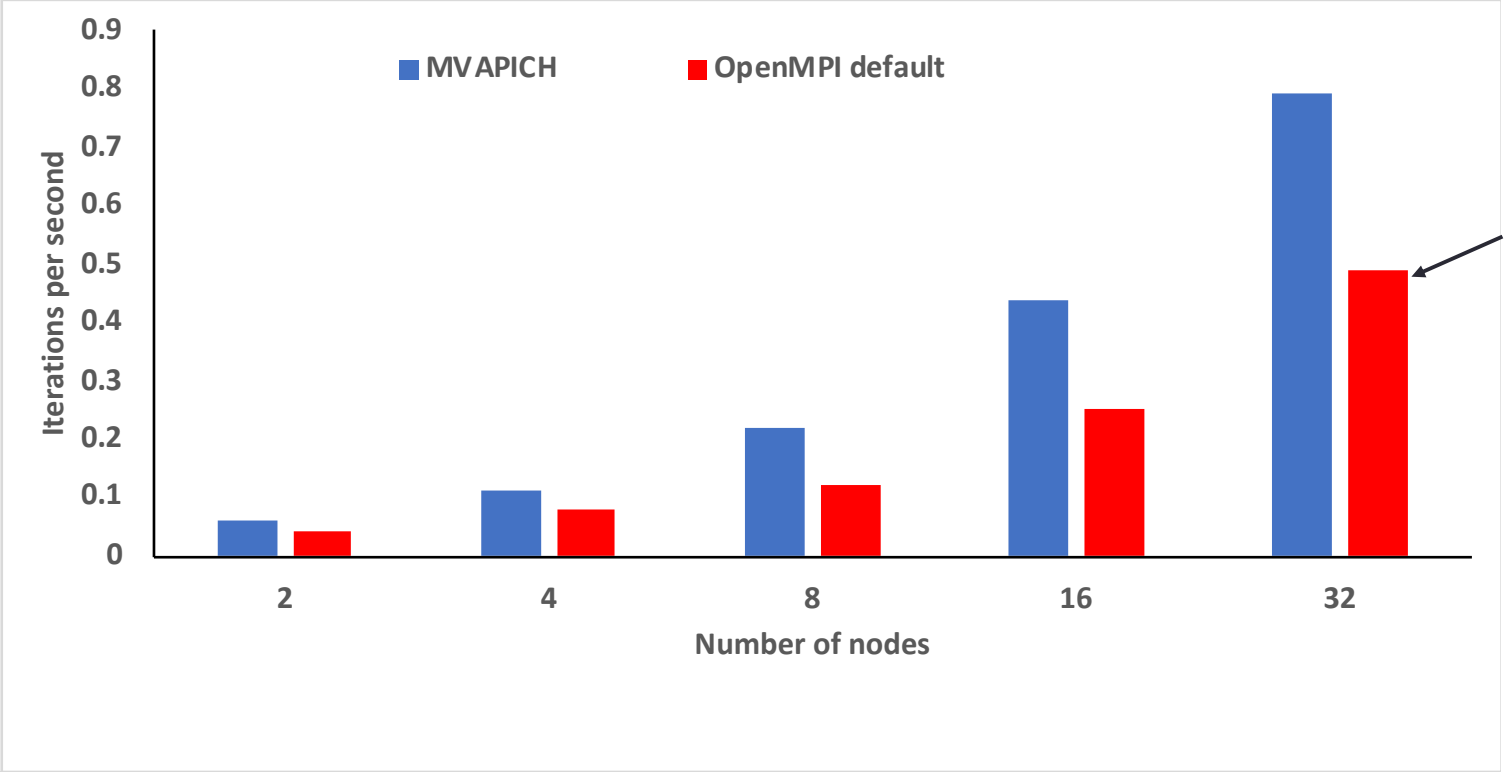
- The system is currently being used to explore and demonstrate the role of ARM in HPC
- A key question is what the performance characteristics of different MPI implementations are on the machine, and how do these differ from other systems?
  - We have MVAPICH2 (2.3.2), OpenMPI (4.0.0) and HPE's MPT(2.20) installed on Fulhame
    - Which one should I use for my application?
- In this talk comparing the communication properties of four applications/benchmarks with the different MPI implementations
  - OpenSBLI, Met Office NERC Cloud model (MONC), and GROMACS
  - Using the OSU micro-benchmarks to understand what we see with performance

OpenSBLI

- Python-based modelling framework that expands a set of differential equations written in Einstein notation, and generates C code that performs the finite difference approximation
  - Generated C code uses the OPS library for target specific parallelisation technology such as MPI or GPUs
- Using Taylor-Green vortex problem in a cubic domain of length 2π as benchmark problem
  - Strong scaling with problem size of 1024^3
  - Contains a mixture of P2P communications and collectives (reduction and gather)
  - Written in C++
  - On Fulhame with MPT there is a build error which means we only have results for MVAPICH and OpenMPI



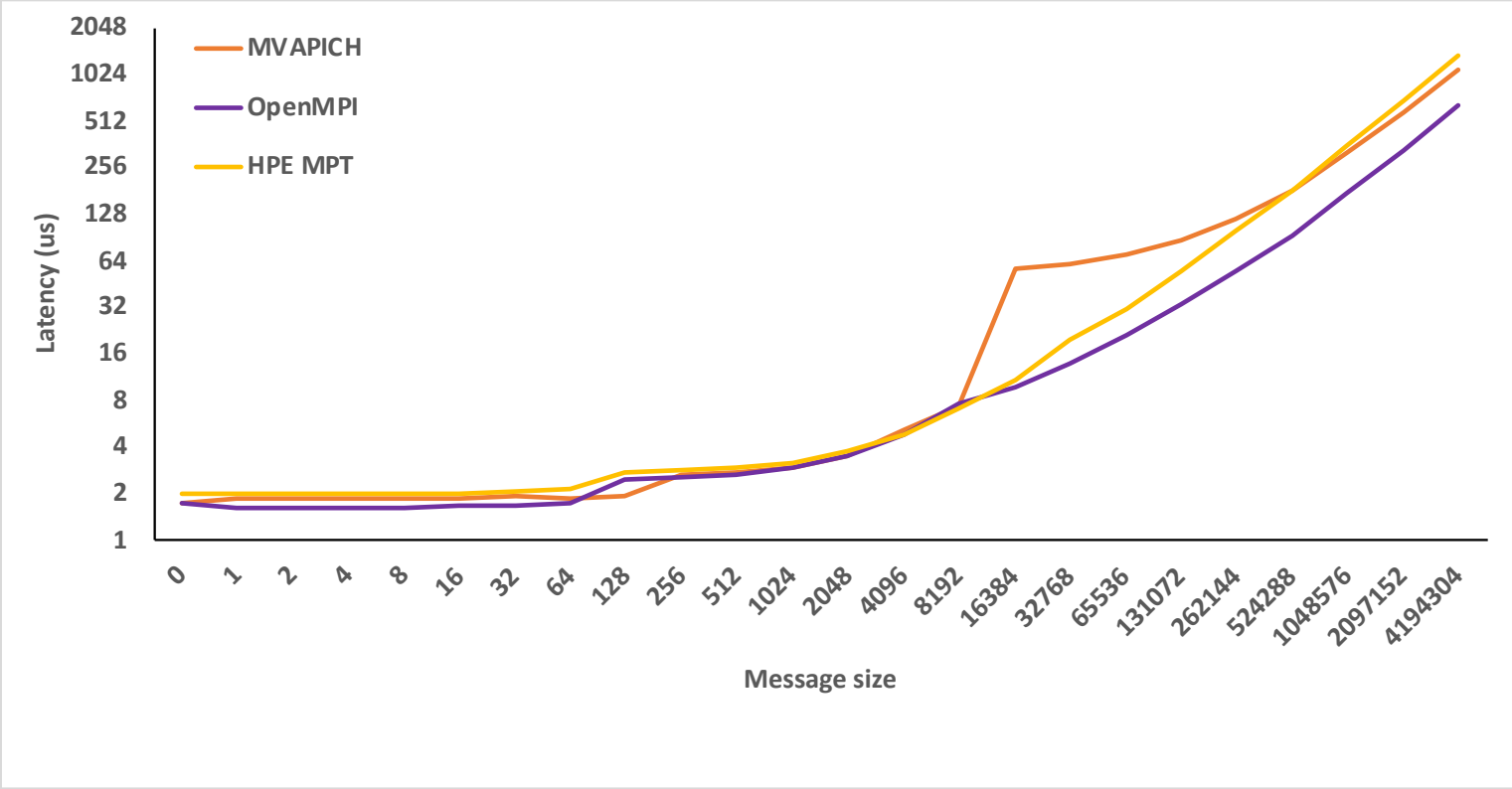
OpenSBLI



Default is for UCX to use UD transport, instead forcing it to use RC improved performance significantly *UCX_TLS=shm, rc*

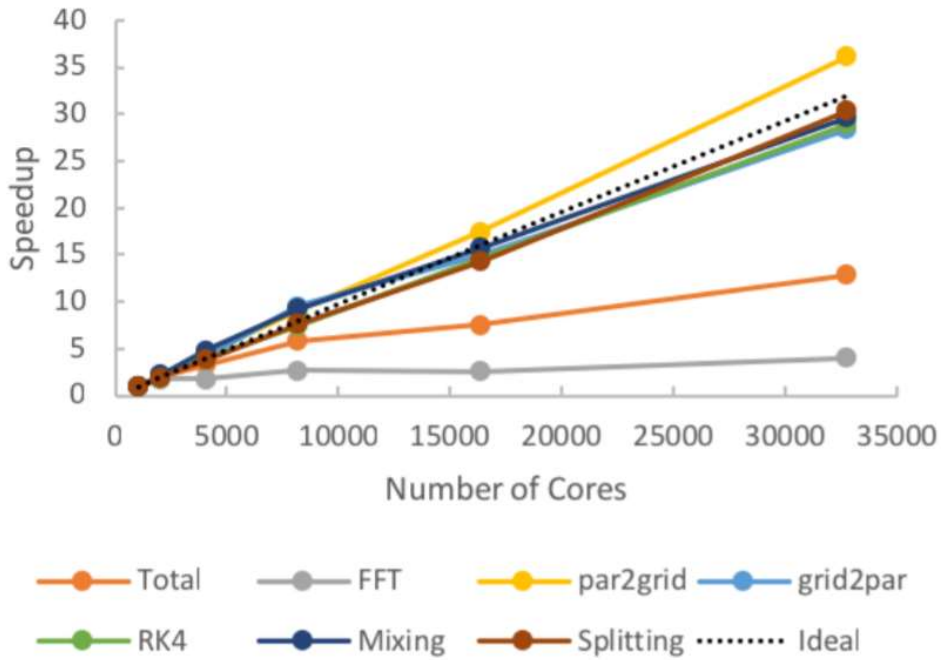
Compiled with GCC compiler 8.2

P2P Inter-node latency

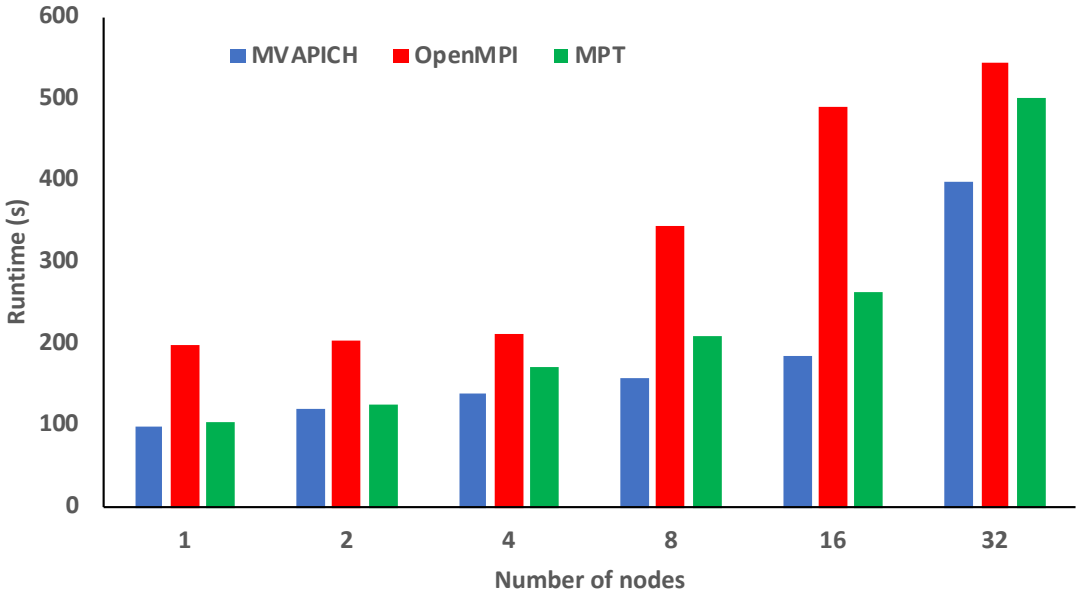


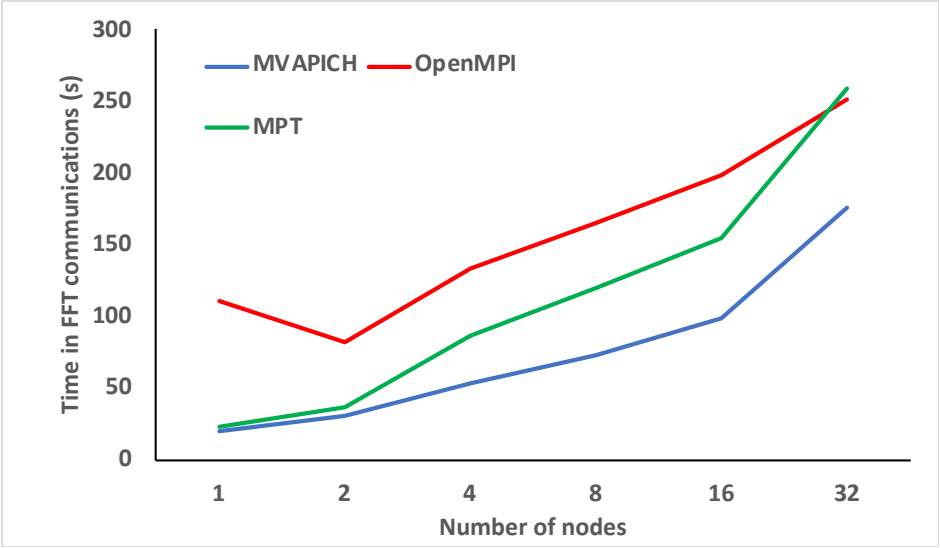
Met Office NERC Cloud (MONC) model

- MONC is a model we developed with the Met Office for simulating clouds and atmospheric flows
  - Written in Fortran 2003 and uses MPI (and OpenMP) for parallelisation
  - Stencil based computation, halo swapping between timesteps, but very heavy use of FFTs



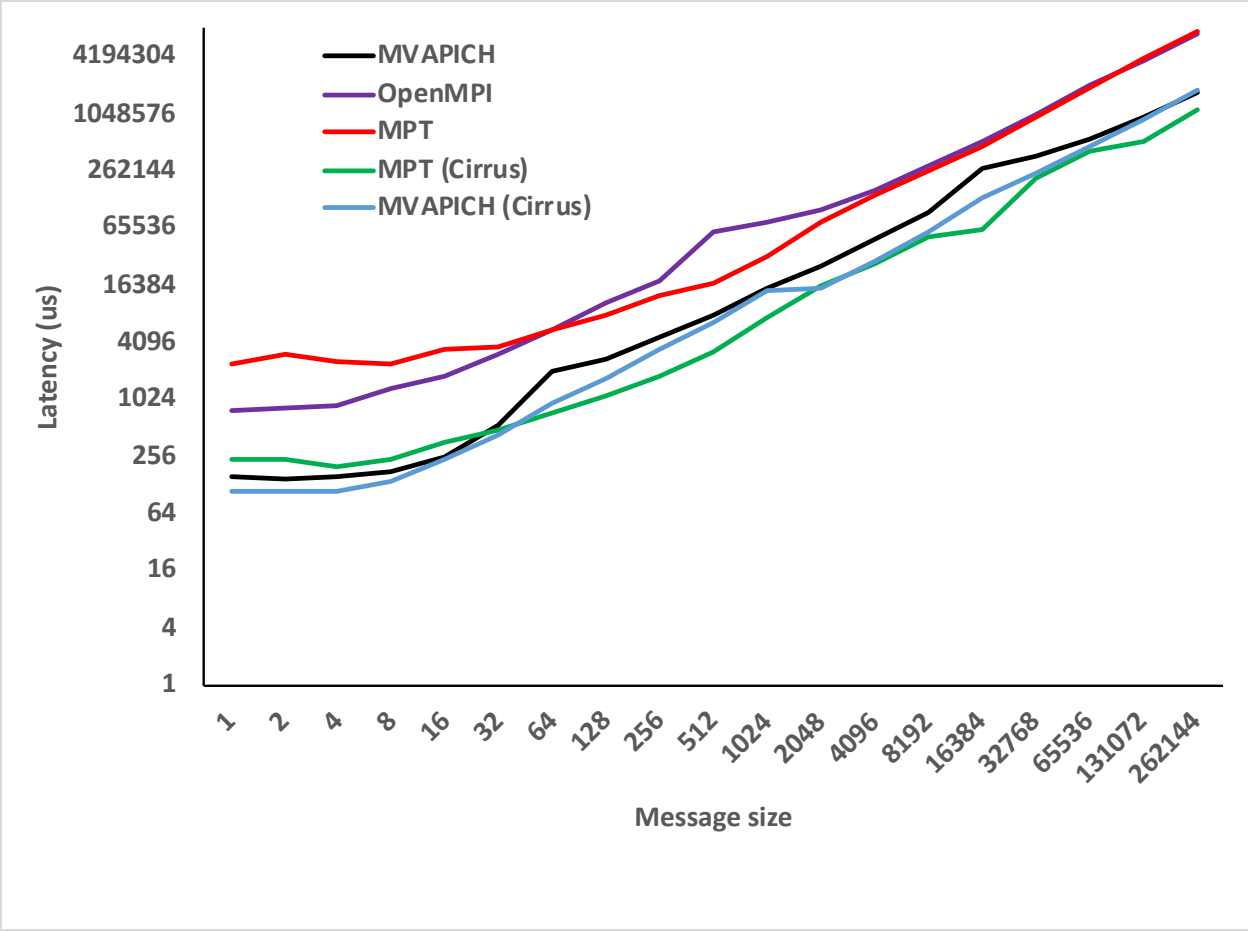
Met Office NERC Cloud model





Compiled with GCC version 8.2, weak scaling with local problem size of 131,072 grid cells

MPI_Alltoall latency (16 nodes)



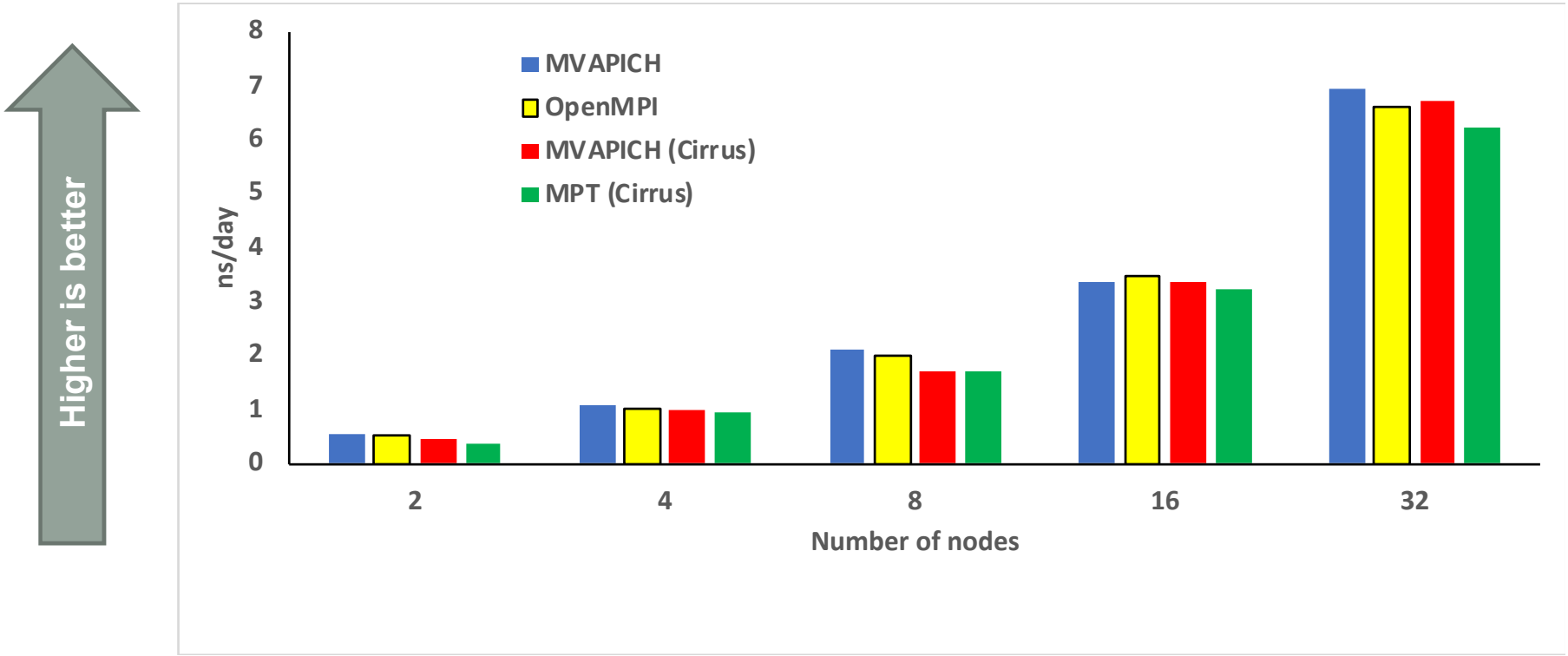
GROMACS

- Package to perform molecular dynamics with hundreds to millions of particles
  - Primarily designed for biochemical molecules like proteins that have a lot of complicated bonded interactions
  - Since GROMACS is extremely fast at calculating the nonbonded interactions many groups are also using it for research on non-biological systems
- GROMACS 1400k atom benchmark taken from the HEC BioSim website
  - Strong scaling
  - GCC 8.3 compiler on Fulhame, GCC 4.8 on Cirrus

GROMACS
FAST. FLEXIBLE. FREE.

epcc

GROMACS


Higher is better
ns/day
8
7
6
5
4
3
2
1
0
MVAPICH
OpenMPI
MVAPICH (Cirrus)
MPT (Cirrus)
2
4
8
16
32
Number of nodes

Questions?

epcc