Cloud High Performance Computing

Justin Freeman
Model Systems Team Leader
Bureau of Meteorology

Australian Government
Bureau of Meteorology

Research and Development Plan 2020-2030

Themes

CUSTOMISED IMPACT-BASED FORECASTS AND WARNINGS WHEN AND WHERE IT COUNTS

Develop higher-resolution, more localised, customised and accurate forecasts, updated frequently for cities and regional areas,

RELIABLE AND TRUSTED FORECASTS

Enhance the quality of, and increase, when and where appropriate, the quantity of observational data for analyses and for assimilation into our models

AN EARTH SYSTEM NUMERICAL PREDICTION CAPABILITY

Develop a full Earth system numerical prediction capability with coupled subcomponents

SEAMLESS WEATHER AND CLIMATE INSIGHTS

Historical observations and predictions, from minutes to decades

Australian Government
Bureau of Meteorology

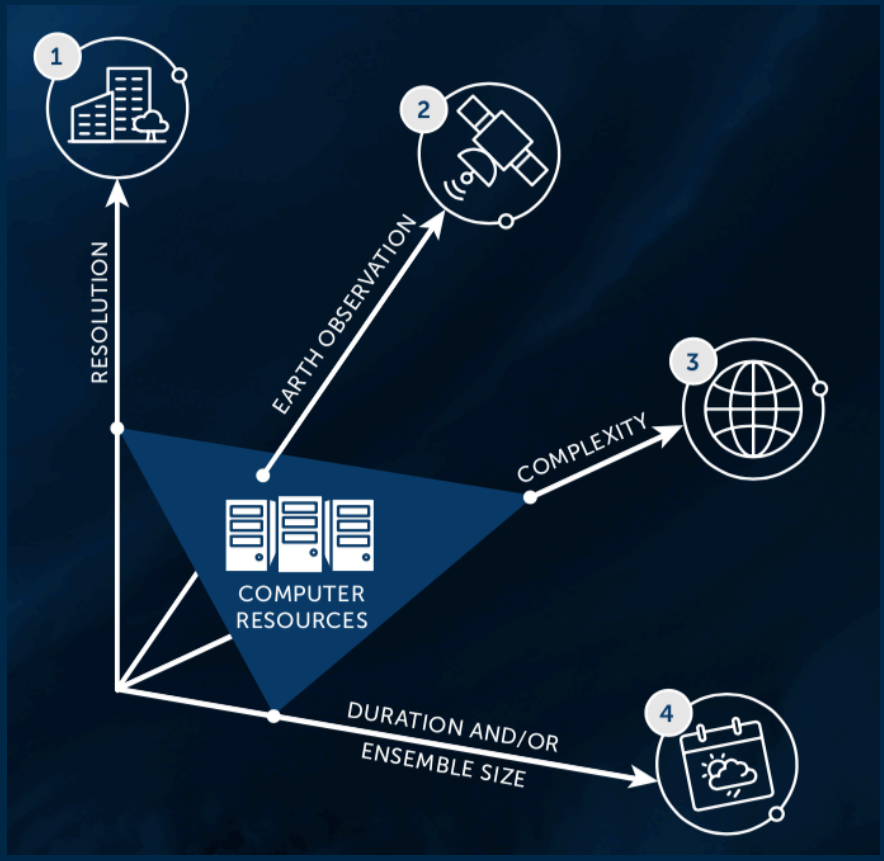
Research and Development Plan 2020-2030

Key Drivers

Computing infrastructure is fundamental to the successful delivery of our research and development plan.

Our numerical prediction systems will need to evolve to take advantage of next generation compute hardware and infrastructure.

Achieving these objectives is reliant on the adequate availability of HPC resources and its optimal utilisation.



Australian Government
Bureau of Meteorology

Research and Development Plan 2020-2030

AN EARTH SYSTEM NUMERICAL PREDICTION CAPABILITY

Develop a full Earth system numerical prediction capability with coupled subcomponents

3-YEAR TARGET	5-YEAR GOAL	10-YEAR VISION
Initial implementation of future computing environment, Earth system coupling and efficient research to operations	Introduction of Next Generation Modelling System, capable of running on future HPC architectures	Nation-wide numerical full Earth system ensemble prediction system

Australian Government
Bureau of Meteorology

Research and Development Plan 2020-2030

3-YEAR TARGET	5-YEAR GOAL	10-YEAR VISION
DELIVERED THROUGH:		
• Establishment of a research to operations program; • Upgraded seasonal, ocean and wave predictions; • Prototype coupled model implemented for experimental numerical weather and ocean prediction; • Targeted development of exascale algorithms and tools and software refactored for graphics processing unit (GPU) accelerated computing; • Prototype Next Generation Modelling System components implemented and tested in collaboration with the UK Met Office;	• Seamless atmospheric, ocean, land and hydrological modelling across time scales; • Antarctic coupled prediction systems, including sea-ice forecasting capability; • Transition of development and services to cloud computing; • Improved representation of physical processes in ACCESS from process study results; • Accurate coupled fire-atmosphere model; • New observational capabilities for Antarctica; • Significant reduction in sea surface temperature bias for Pacific Ocean, improved	• Fully coupled Earth system models, including hydrology, ensembles, with regional kilometre-scale, sub-hourly updating; • Use of information from throughout the water cycle to initialise coupled models; • Convection-allowing global model; • Significantly improved foundational Earth system components coupled for near-seamless information for all time scales; • Significant reductions in model biases in regions such as the Maritime continent and

Australian Government
Bureau of Meteorology

HPC + Cloud

R&D Plan Benefits

- Greater efficiency in model development
- Modelling Systems that are aligned to adopt future HPC architecture
- New technologies and HPC enable a flexible response to new software requirements

Challenge

Can we match the UM atmospheric model performance on AWS?

Australian Government
Bureau of Meteorology

UM Benchmark

ACCESS-G	
Spatial Resolution	0.175° longitude by 0.117° latitude ~12km mid-latitudes ~17km in tropics
Grid Resolution	1536 x 2048 70 vertical levels top level ~80km
Operational Configuration	9792 cores 25 minutes 3-day forecast (short FC)

Australian Government
Bureau of Meteorology

UM Benchmark

UM benchmark for N1024L70 (12km) based on PS39 (UM10.6+)

Benchmark Configuration	
Grid decomposition	16 x 36 MPI procs 4 OpenMP threads per proc
IO decomposition	6 x 8 MPI procs 4 OpenMP threads per proc
Total	4x(16 x 36) + 4x(6x8) = 2496 cores

Australian Government
Bureau of Meteorology

HPC System

- 3,024 nodes each containing two 24-core Intel Xeon Cascade Lake processors.
- 48 cores per node
- 192 Gigabytes of memory per node.
- Up to 200 Gb/s InfiniBand.
- Lustre parallel file system.

Reference System

c5n.18xlarge instance

- Intel Xeon Skylake-SP processors.
- 36 cores per node (hyperthreading disabled).
- 192 GB Memory per node.
- 100 Gbps of dedicated point-to-pont network bandwidth.
- AWS Parallel Cluster.
- Elastic Fabric Adapter.
- FSx Lustre 10 TB scratch filesystem.

AWS ParallelCluster

Australian Government
Bureau of Meteorology

Reference and AWS HPC

CentOS 8 operating system

Intel Parallel Studio version 2019.0.815
icc (ICC) 19.0.5.281
ifort (IFORT) 19.0.5.281

gcc 8.3.1

Reference System

Amazon Linux 2

Intel Parallel Studio 2020.0.166
icc (ICC) 19.1.2.254
ifort (IFORT) 19.1.2.254

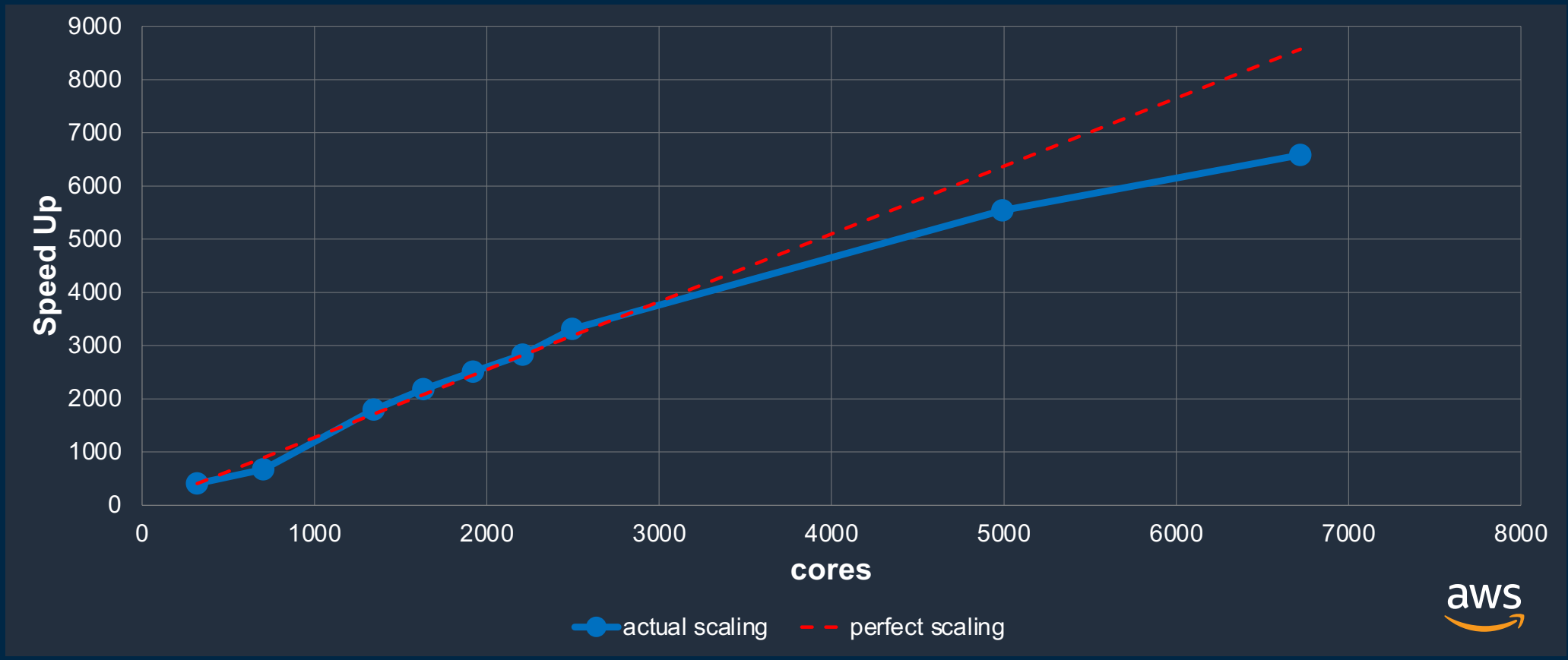
gcc 9.3.0

AWS

Australian Government
Bureau of Meteorology

Unified Model AWS

Reference System (2496 cores)	AWS (2496 cores)
51 min 16 sec	48 min 30 sec
47 min 10 sec	

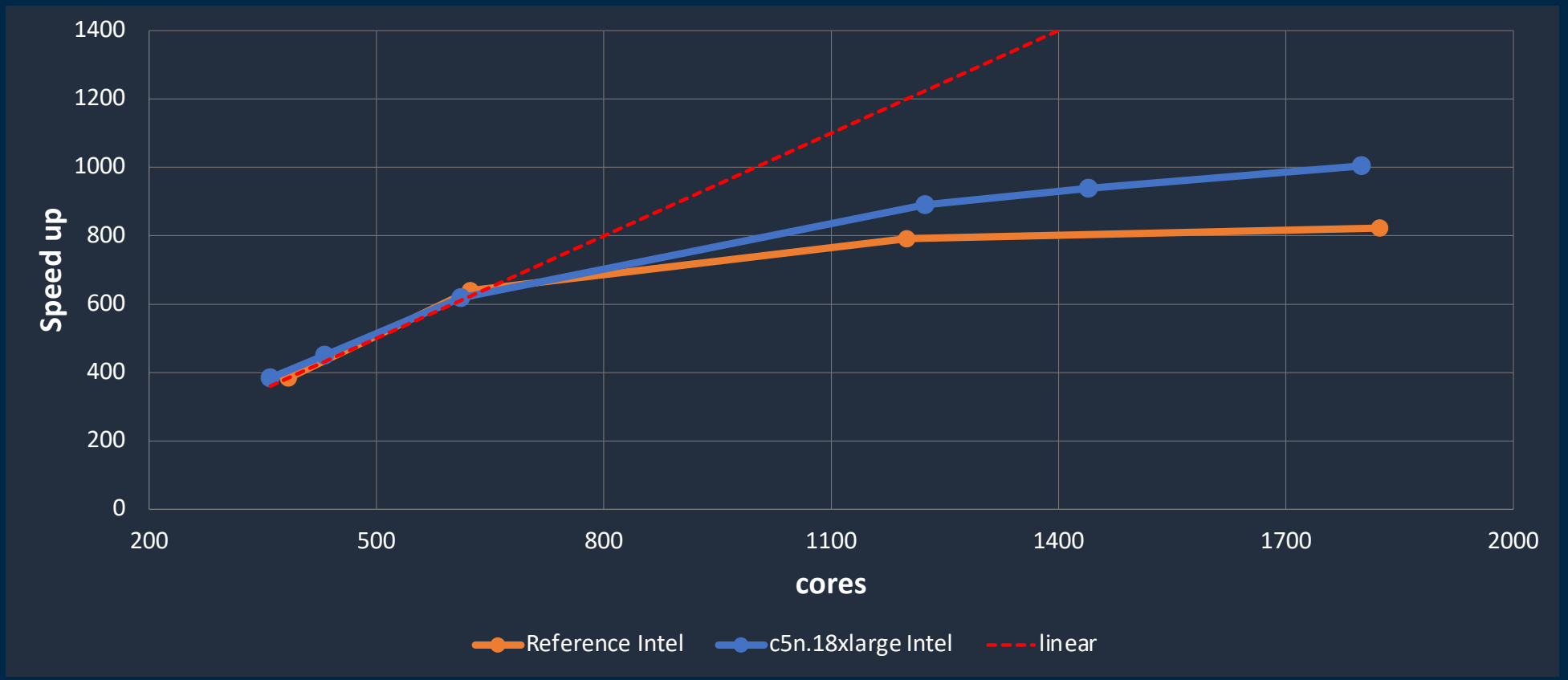


Australian Government
Bureau of Meteorology

What about other HPC codes on AWS?

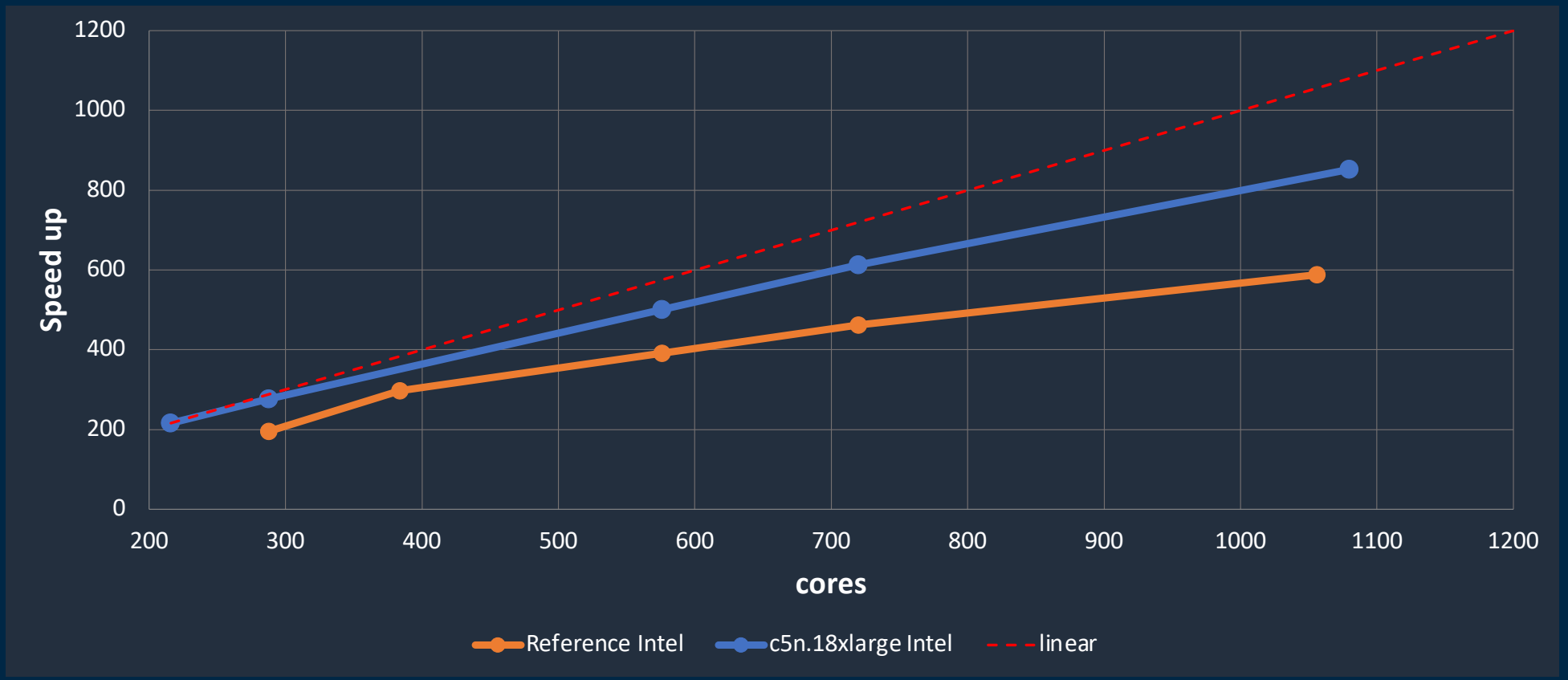
Australian Government
Bureau of Meteorology

OceanMAPS Global Ocean Forecasting MOM5 + SIS



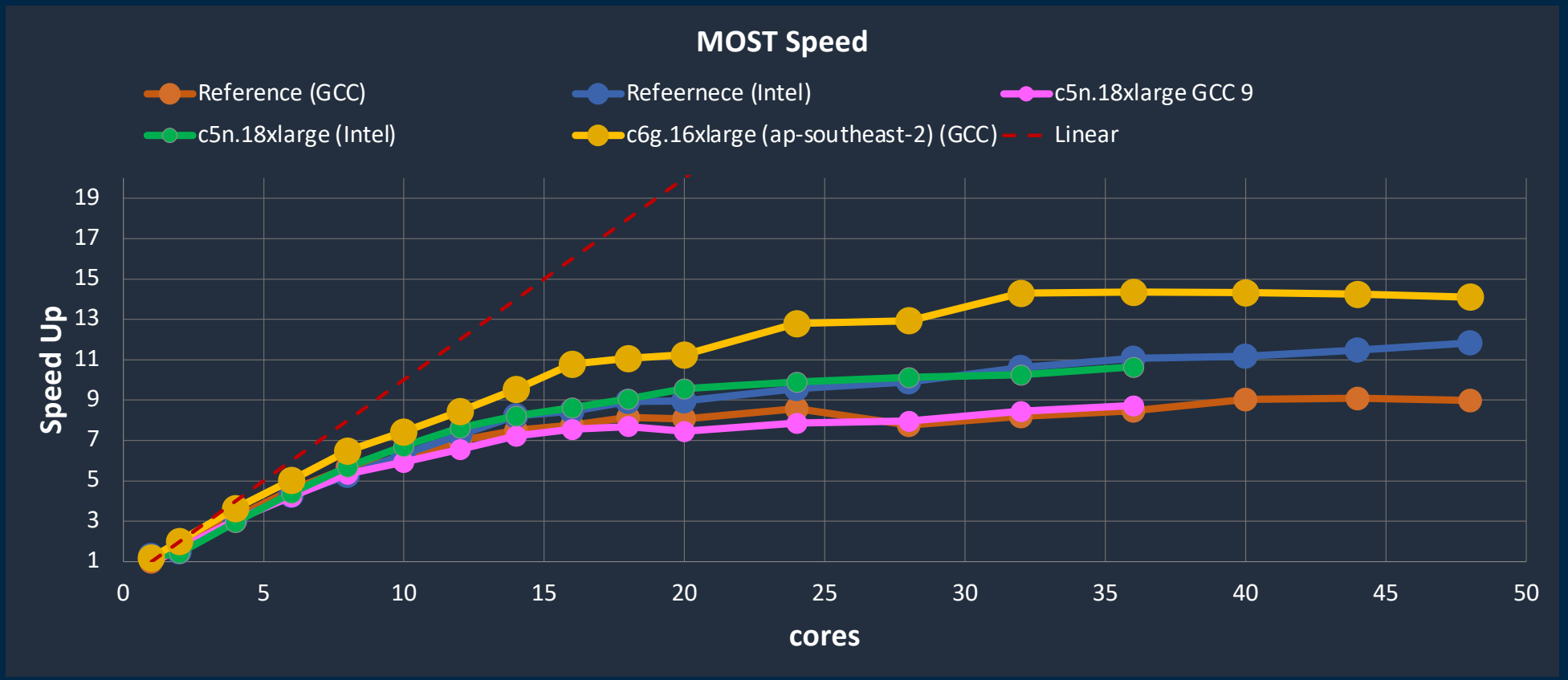
Australian Government
Bureau of Meteorology

ADEPT – WAVEWATCHIII Wave Forecasts



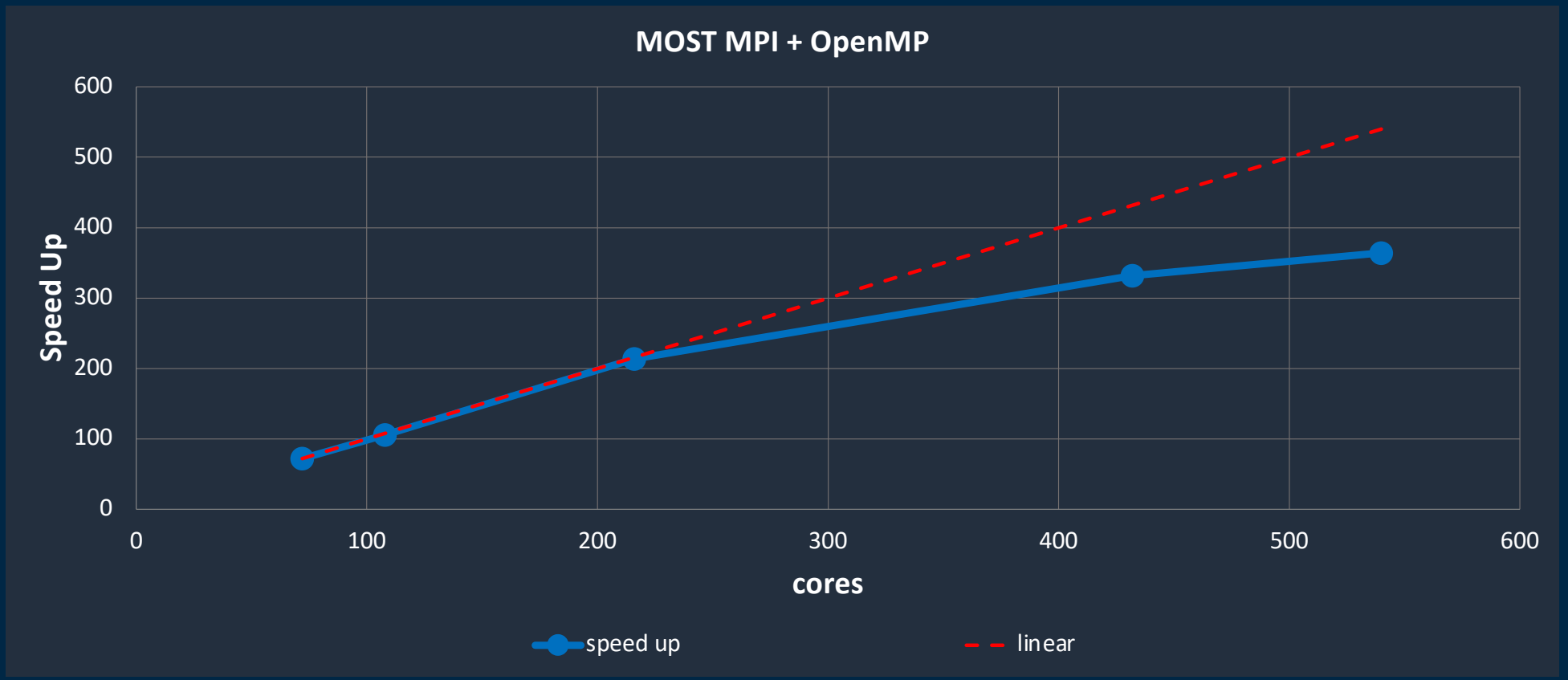
Australian Government
Bureau of Meteorology

MOST Tsunami Propagation OpenMP Performance



Australian Government
Bureau of Meteorology

MOST Tsunami Propagation Hybrid MPI + OpenMP



Australian Government
Bureau of Meteorology

HPC + AWS

- AWS matched current HPC performance
- AWS ParallelCluster is easy to configure and use
  - Familiar environment with access to the tools you currently know and use
  - Flexible to configure cluster to meet the characteristics of the workload
  - Choice of instance types – data optimized, compute optimized, GPU's, ARM, AMD,…
  - Scalable compute and scalable storage
  - Dedicated high performance networking and high performance filesystems
  - Cluster per workload
  - Destroy when you are done
- Next Steps
  - Next Generation Modelling Systems Benchmarking (LFRic)
  - Cloud Optimisation

Australian Government
Bureau of Meteorology

Thank you

Justin.Freeman@bom.gov.au

Kevin Jorissen, Linda Eitelberg and Steve Gillard