Performance of Met Office Weather and Climate Codes on Cavium ThunderX2 Processors

Adam Voysey, Maff Glover

HPC Optimisation Team

Met Office

Contents

- Introduction
  - The Met Office and why we use HPC
- UM and NEMO
- Results
- Understanding the results
- Considering TX2 for an operational system (at the Met Office)
- Conclusions

Introduction

What is the Met Office?

UK Government
Part of UK Government - all Government employees

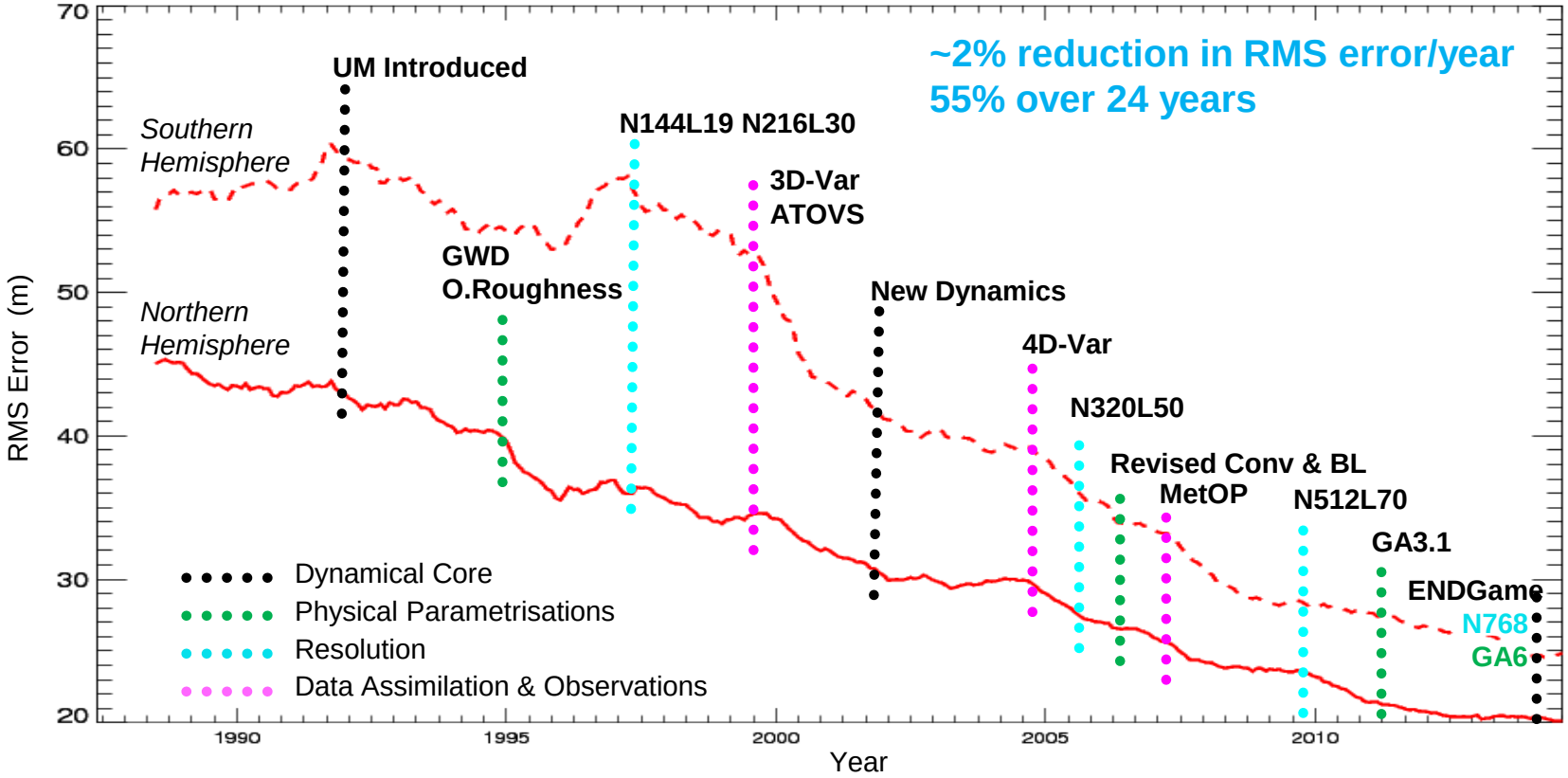
Science
A world-leading science institute in Earth science

Commercial Business
As a trading fund – able to participate in competitive markets

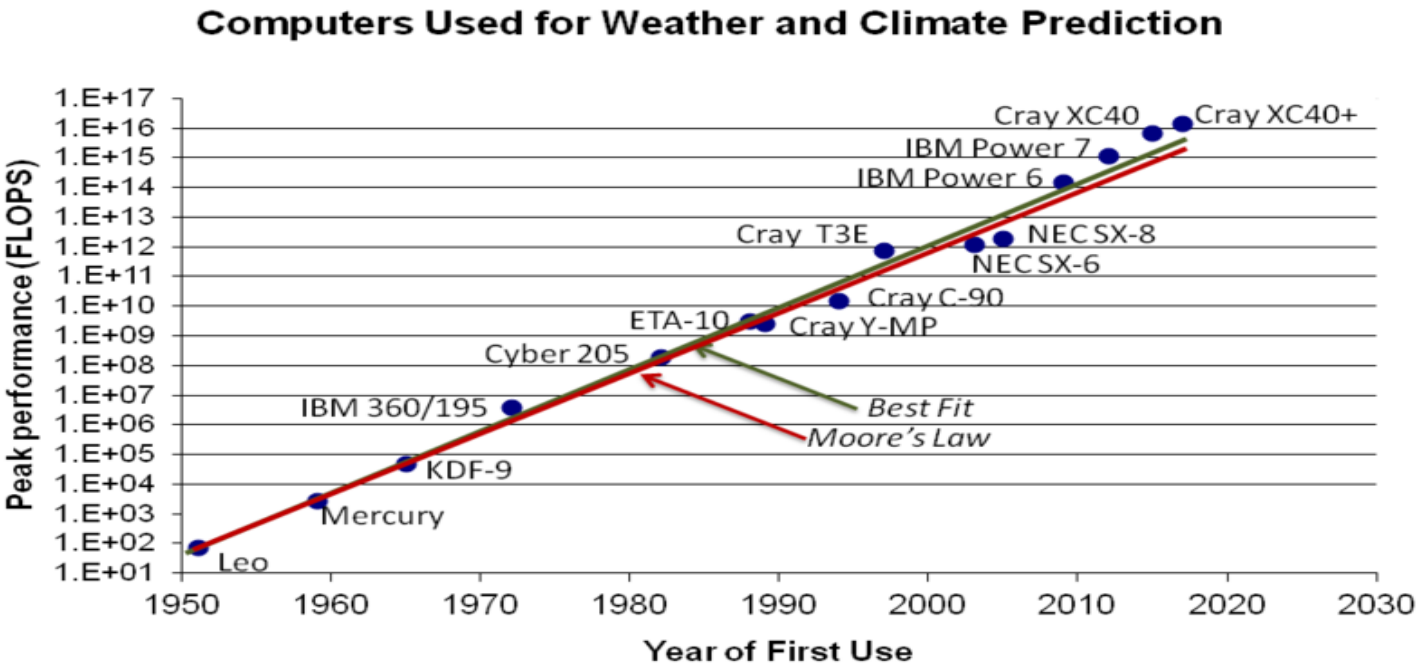
UN International Community
Part of a ~200 strong international community, UN treaties. UM partnerships

Improving Forecasts

500hPa Height Day 3 RMS Error vs. Analyses



- 3 top 40 systems at launch
- Architected for reliability
- 7PF and 2x 2.8PF (HPL)
- 6 Lustre filesystems totalling 24 PB
- Cray Aries interconnect and MPI
- PBS Pro scheduling and Cray/Intel compilers
- Challenging forecast availability targets



Isambard

GW4

University of BRISTOL

UNIVERSITY OF BATH

UNIVERSITY OF EXETER

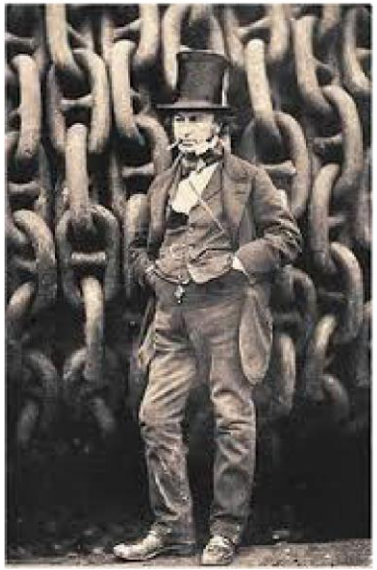
CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Met Office

EPSRC

CRAY
THE SUPERCOMPUTER COMPANY

ARM



Attended the “Raising Steam” and “Stoking The Fire” hackathons.

Worked on porting/running UM and NEMO benchmarks.

(Also in attendance: Maff Glover, HPC Optimisation Team)

Met Office

The Unified Model (UM) & NEMO

Met Office

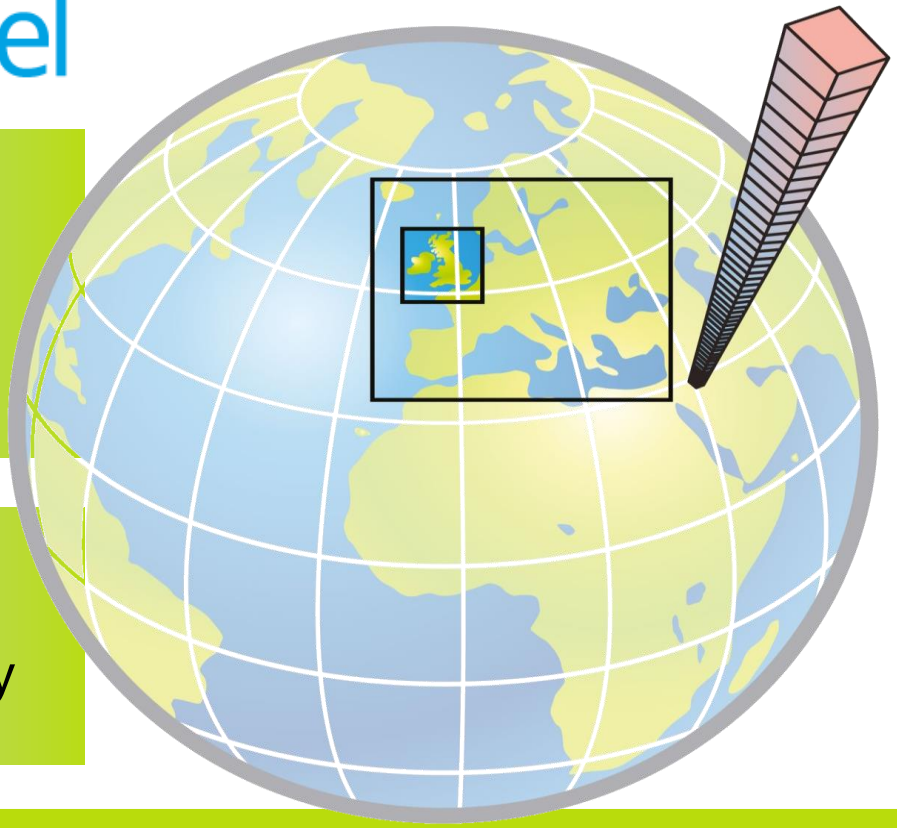
UM | Unified Model

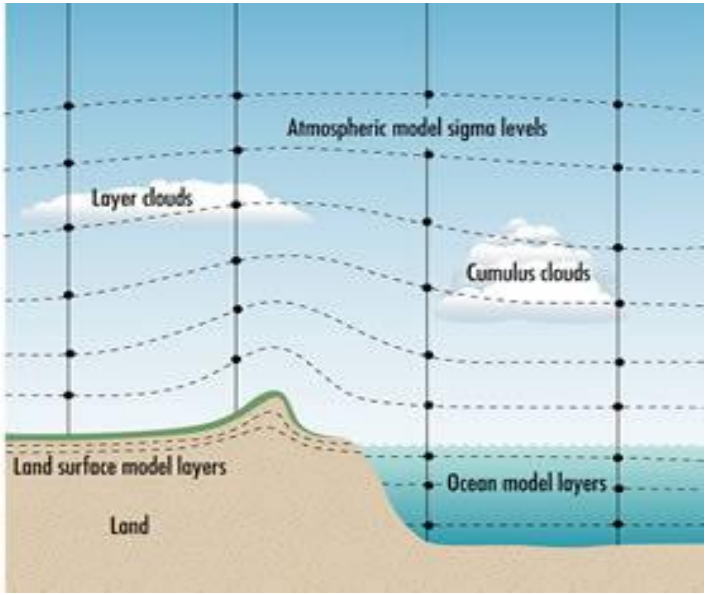
UM Atmosphere

- Well over 20 years old
- Fortran / MPI / OpenMP
- Global Collaboration
- Rapidly changing

Coupled Systems

- 4DVAR Assimilation
- UKCA Atmospheric Chemistry
- NEMO Ocean



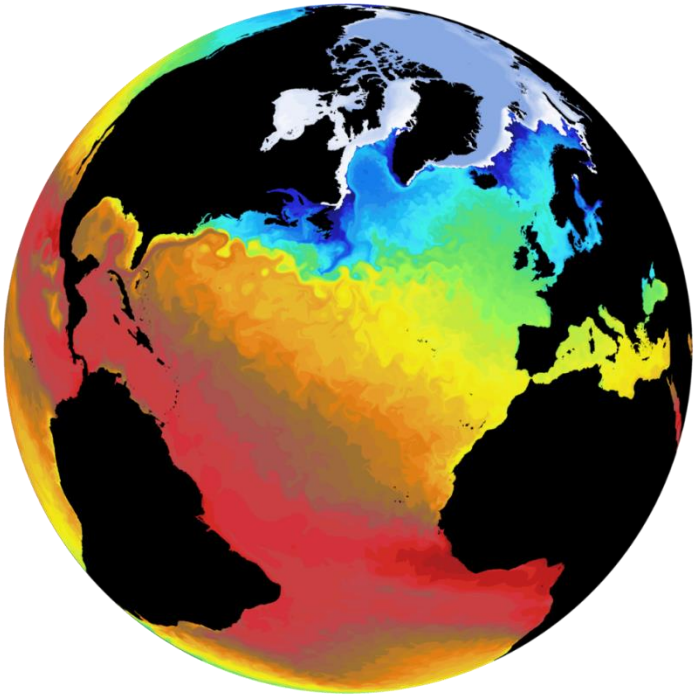


- Regular lat-long grid with vertical sigma levels
- Navier-Stokes on rotating globe (model dynamics)
- Sub-grid parametrizations of rain/sun/... (model physics)
- Initial value problem; Climate is the long-term state

NEMO

- NEMO (Nucleus for European Modelling of the Ocean) is a state-of-the-art ocean modelling framework that includes components for ocean dynamics, for sea-ice and for ocean biogeochemistry.

- NEMO also comes with a nesting package allowing to set-up regional zooms and a versatile data assimilation interface (see https://www.nemo-ocean.eu/).



Benchmark Configurations

UM

- AMIP configuration
- Benchmark based on vn10.8
- Low resolution (N48)
- Little I/O (PMSL diagnostic only)

NEMO

- GYRE_PISCES configuration (idealised)
- Benchmark based on vn3.6 + additional changes (development version towards vn4.0)
- Fairly high resolution (¹⁄₁₂°); but few levels

Met Office

Single Node Performance

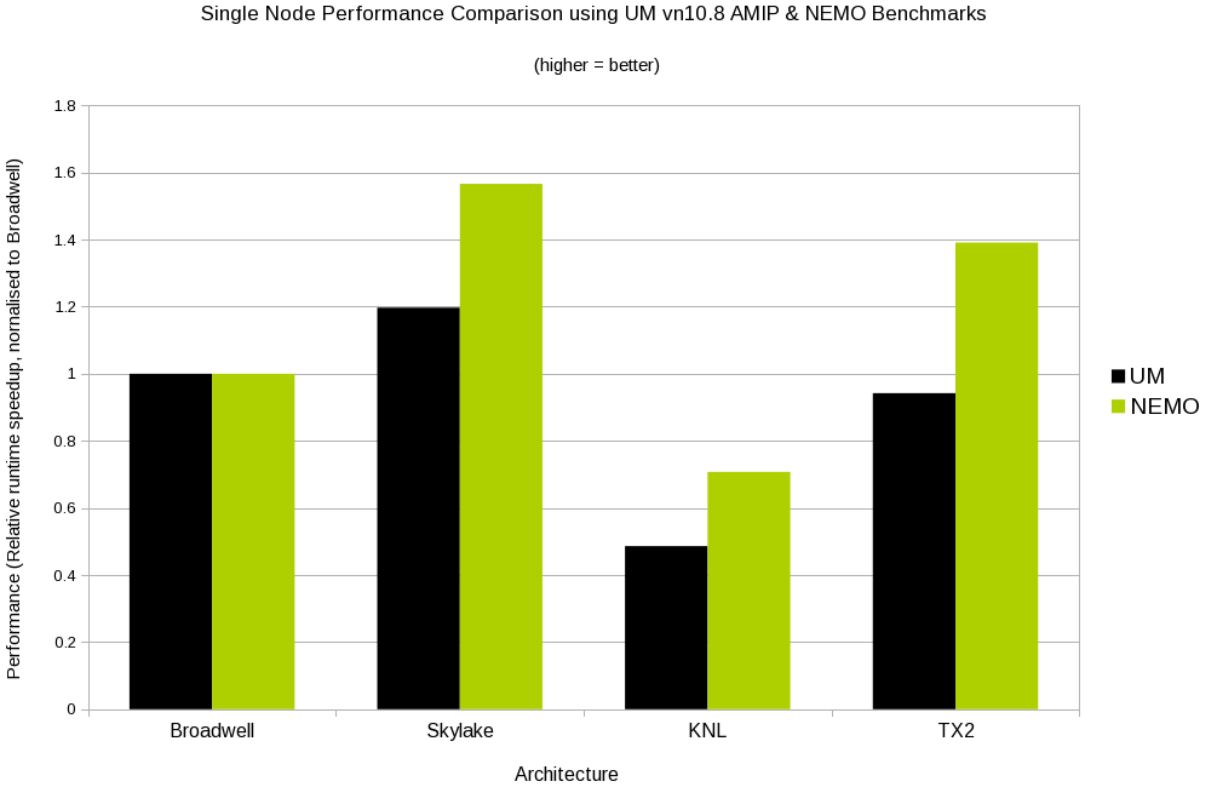
Results

Broadwell
Swan - Intel Xeon (Broadwell),
2 × 22-core @ 2.2GHz

Skylake
Swan - Intel Xeon (Skylake),
2 × 28-code @ 2.1GHz

KNL
XCK – Intel Xeon Phi (Knights landing),
64-core @ 1.3Ghz

TX2
Isambard – Cavium ThunderX2
2 × 32-core @ 2.2GHz



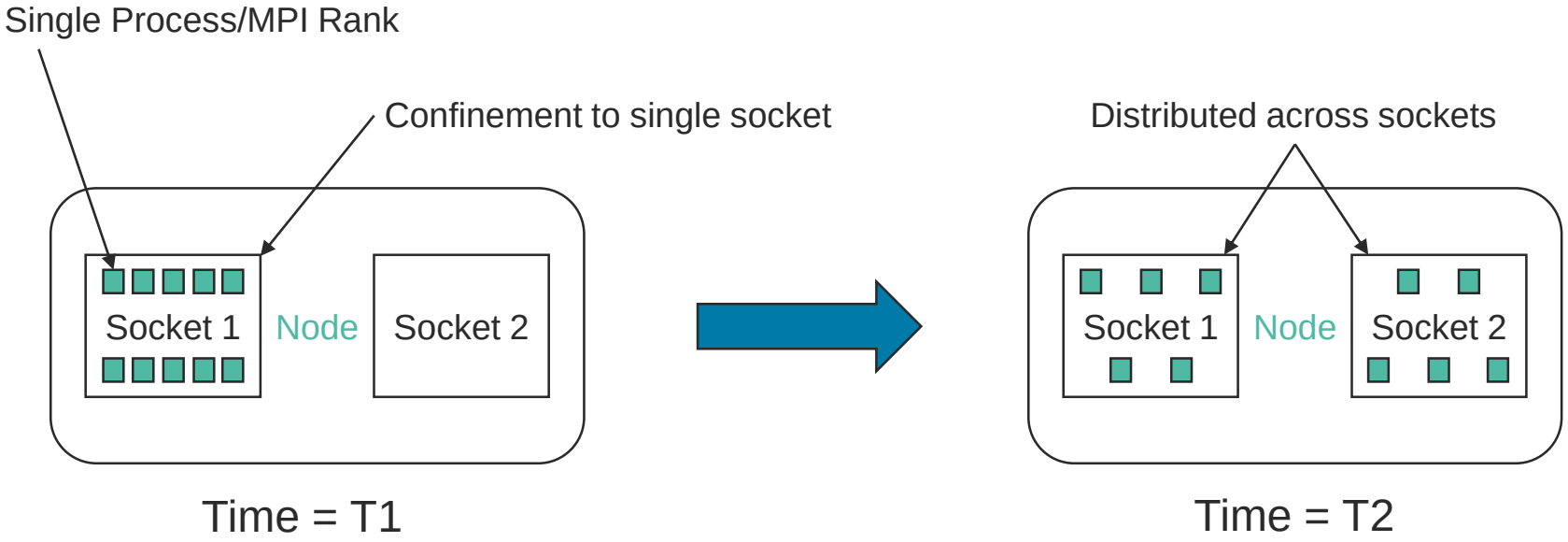
Met Office

Digging Deeper:

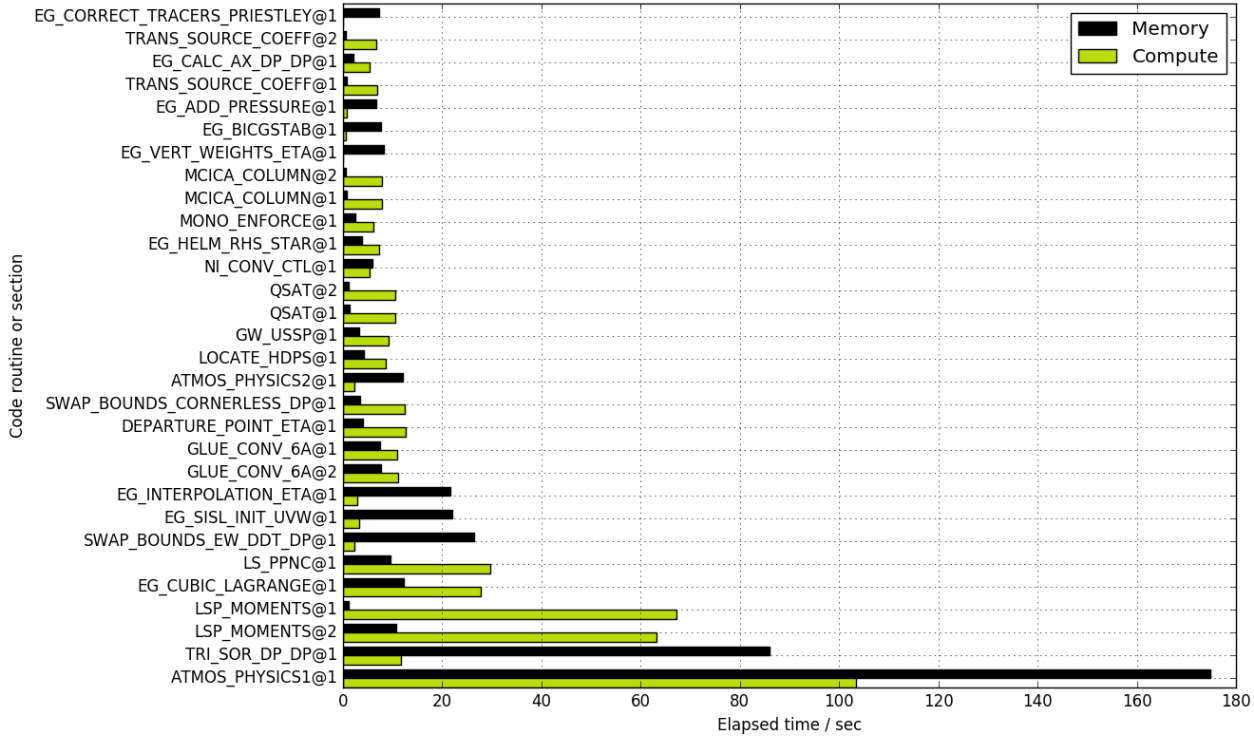
Understanding the results

Estimating Compute/Memory Bound Code

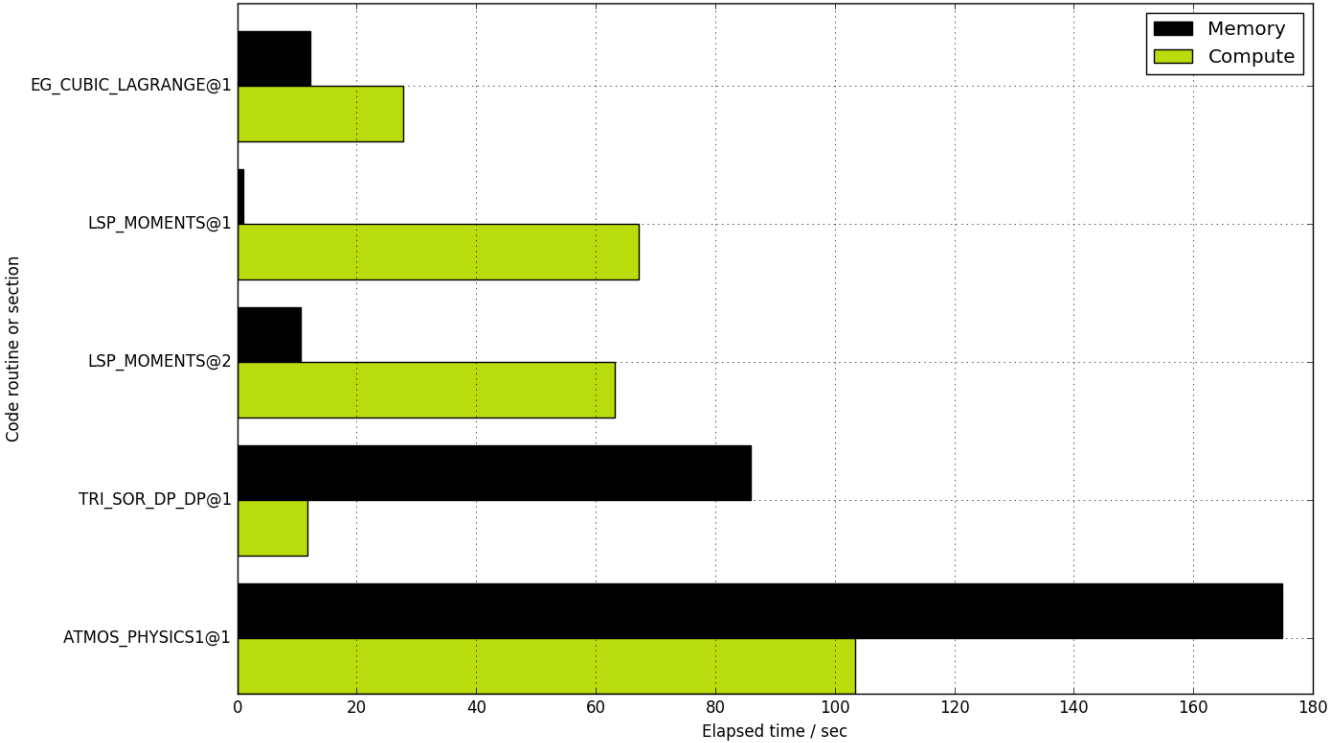
Compute = (2 × T2) − T1
Memory = 2 × (T1 − T2)



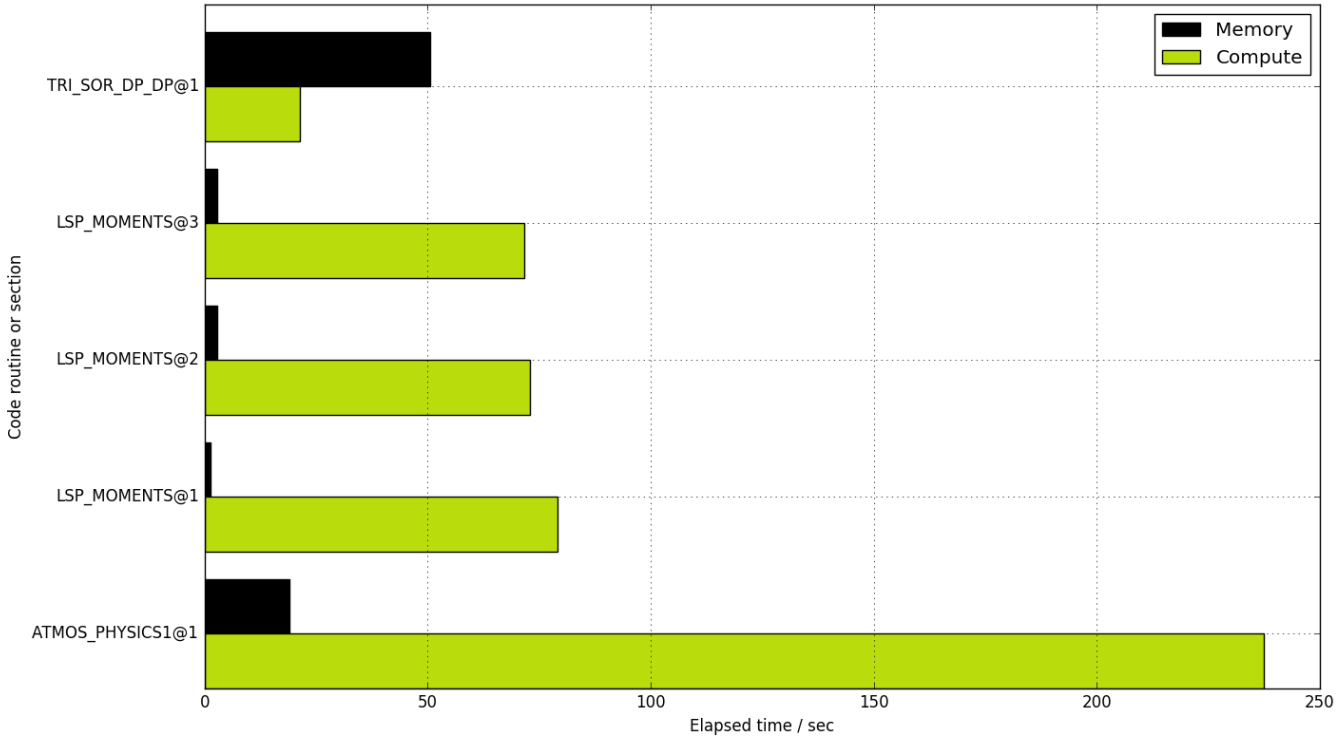
UM On Broadwell (top 30)



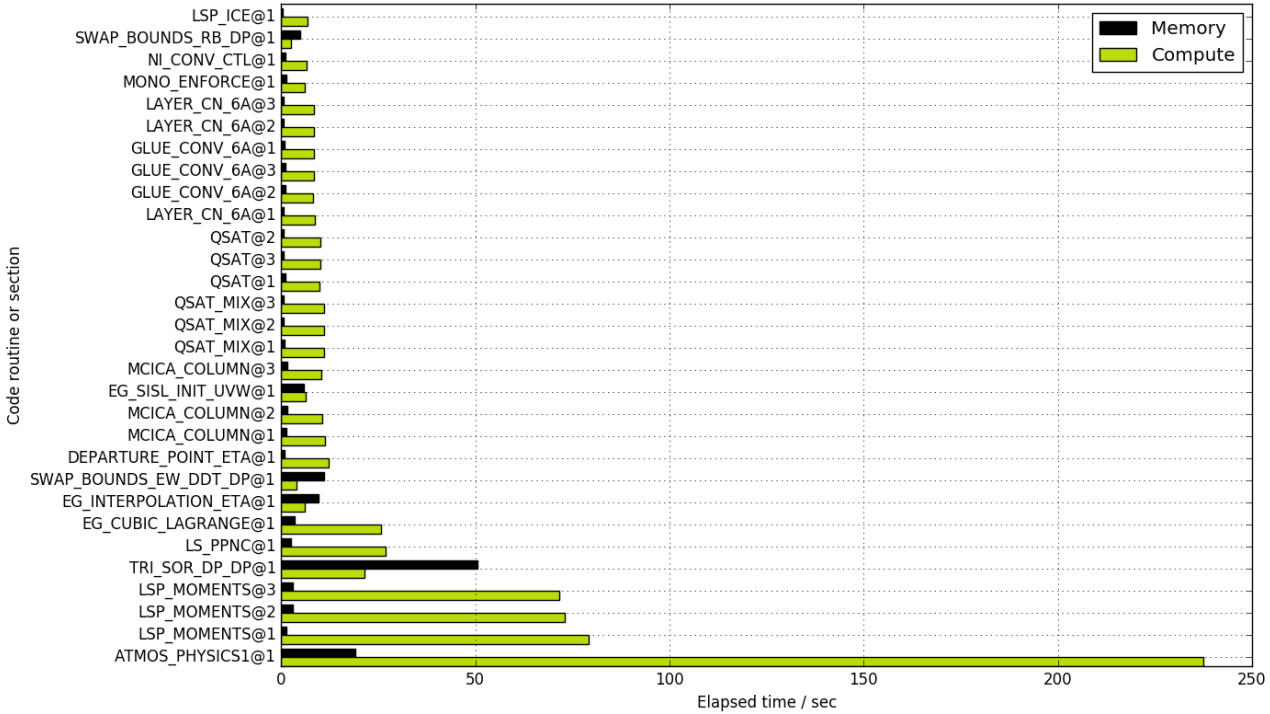
UM On Broadwell (top 5)



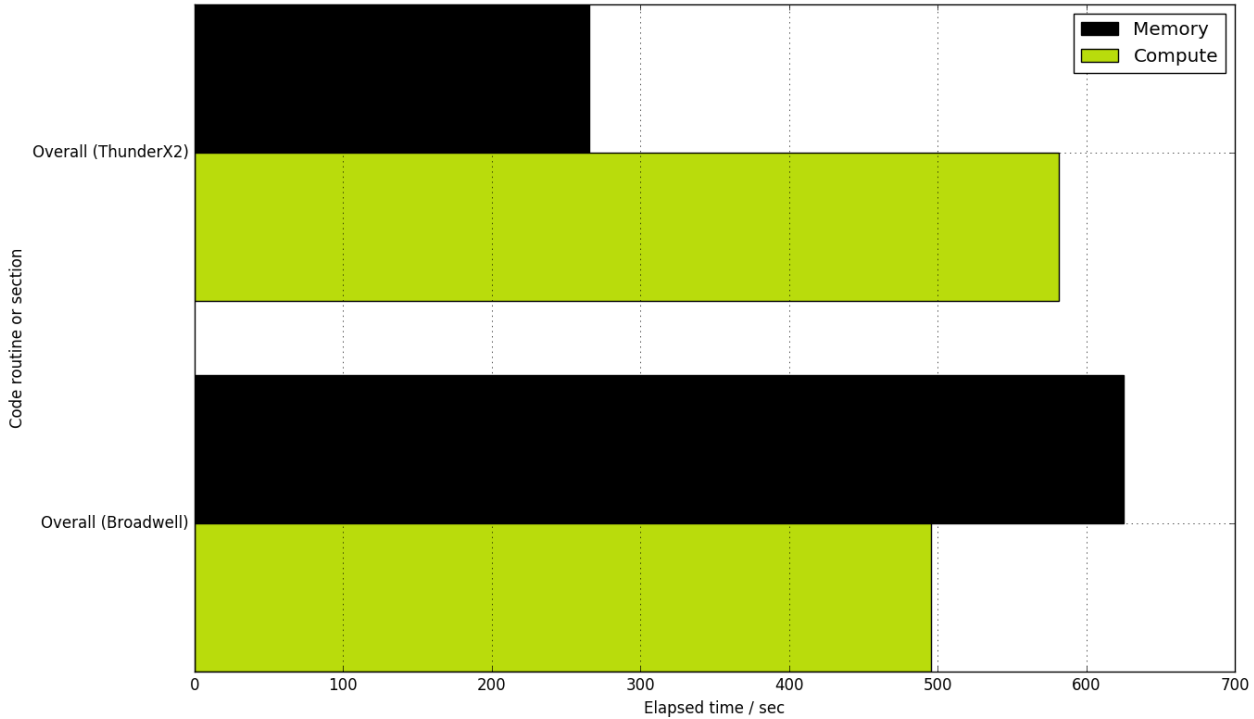
UM On TX2 (top 5)



UM On TX2 (top 30)

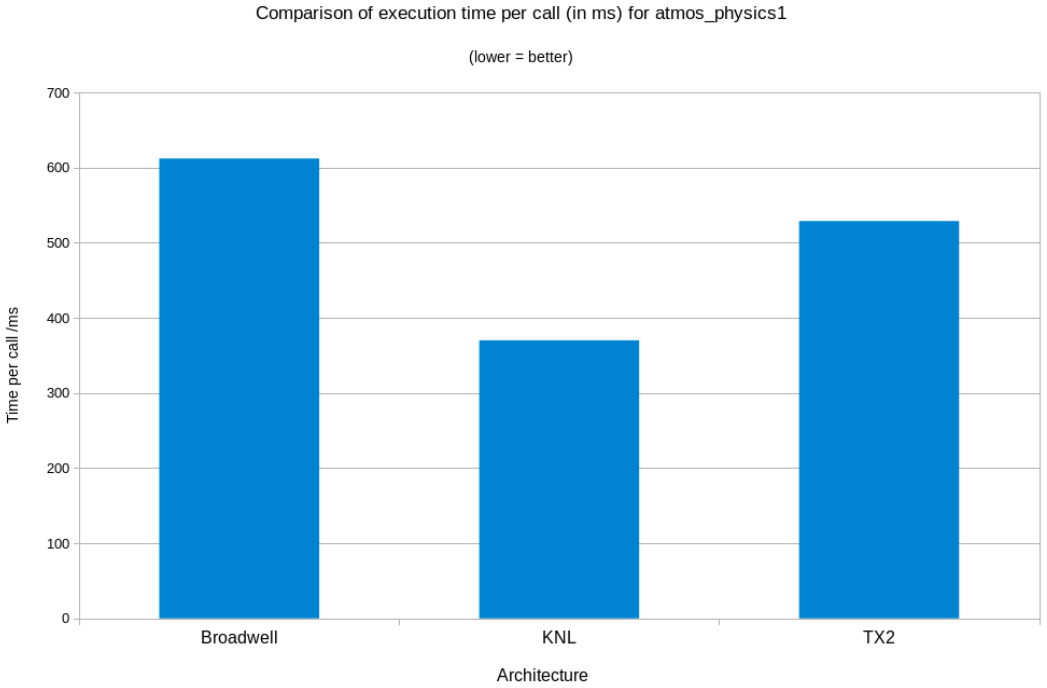


UM Overall



ATMOS_PHYSICS1() is a microcosm!

- Contains a mixture of compute and memory bandwidth bound sections
- Contains some sections with OpenMP
- Some sections vectorise well; others don't.
- Contains (calls to) MPI communications
- [But no I/O]



Why does KNL do well?

Vectorisation!

- Can use the Cray compiler to generate loop marking and optimisation reports for the code
- [-hlist=]

TX2

```
1223.   M                !$OMP DO SCHEDULE(STATIC)
1224.  + M m-----------------< DO j = 1, rows
1225.   M m Vr8------------<   DO i = 1, row_length
1226.   M m Vr8                p_layer_boundaries(i,j,0) = p_star(i,j)
1227.   M m Vr8                p_layer_centres(i,j,0) = p_star(i,j)
1228.   M m Vr8------------>   END DO
1229.   M m----------------> END DO
1230.   M                !$OMP END DO NOWAIT
```

KNL

```
1223.   M                !$OMP DO SCHEDULE(STATIC)
1224.  + M m-----------------< DO j = 1, rows
1225.   M m Vr2------------<   DO i = 1, row_length
1226.   M m Vr2                p_layer_boundaries(i,j,0) = p_star(i,j)
1227.   M m Vr2                p_layer_centres(i,j,0) = p_star(i,j)
1228.   M m Vr2------------>   END DO
1229.   M m----------------> END DO
1230.   M                !$OMP END DO NOWAIT
```

M – multithreaded
m – partitioned
r – unrolled
V – vectorised
+ - generated additional text information

Compiler Optimisation

Cray Scientific Library Calls

TX2

```
1209.       $GVN_3068 = and( -t$224, -t$185 ) >= 0
1209.       if ( int( $GVN_3068 ) == 0 ) then !  99.50%
1209.         $I_L1209_S182 = 0
1209.         $T_$I_L1209_912_H6 = t$185 * t$224
1209.         if ( $T_$I_L1209_912_H6 >= 2 ) then !  99.50%
1209.           $TC_343 = and( -2, $T_$I_L1209_912_H6 )
1209.           $LC_S180 = -$TC_343
1209.           $SI_S181 = 0
1209.           $LIS_snowdepth_surft_498 = int( 0[loc( snowdepth_surft ),0].L )
1209.           $LIS_oloc_snowdepth_1849 = loc( snowdepth_p )
1209.           if ( $LC_S180 < -15 ) then
1209.             do
1209.               $GVN_3537 = $LIS_snowdepth_surft_498 + $SI_S181
1209.               $GVN_3538 = $LIS_oloc_snowdepth_1849 + $SI_S181
1209.               0[$GVN_3537:2:1,a16].L = 0[$GVN_3538:2:1,a8].L
1209.               2[$GVN_3537:2:1,a16].L = 2[$GVN_3538:2:1,a8].L
1209.               4[$GVN_3537:2:1,a16].L = 4[$GVN_3538:2:1,a8].L
1209.               6[$GVN_3537:2:1,a16].L = 6[$GVN_3538:2:1,a8].L
1209.               8[$GVN_3537:2:1,a16].L = 8[$GVN_3538:2:1,a8].L
1209.               10[$GVN_3537:2:1,a16].L = 10[$GVN_3538:2:1,a8].L
1209.               12[$GVN_3537:2:1,a16].L = 12[$GVN_3538:2:1,a8].L
1209.               14[$GVN_3537:2:1,a16].L = 14[$GVN_3538:2:1,a8].L
1209.               $SI_S181 = 128 + $SI_S181
1209.               $LC_S180 = 16 + $LC_S180
…
```

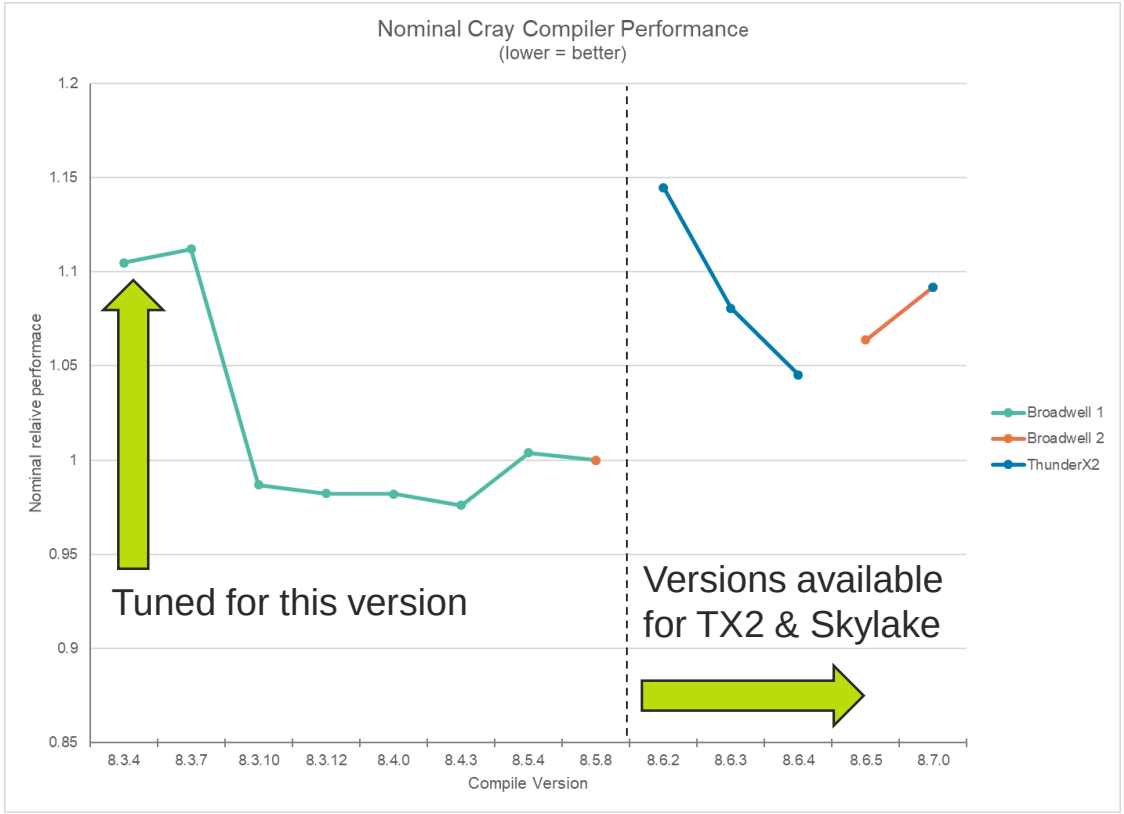
KNL

```
1209.       if ( and( -t$224, -t$185 ) < 0 ) then !  99.50%
1209.         $GCS_101 = t$185 * t$224
1209.         if ( $GCS_101 > 256 ) then
1209.           __cray_dcopy_knl( 0[int( snowdepth_surft[0].L ),0].L, 0[loc( snowdepth_p ),0].L,
                                                                                $GCS_101 )
1209.         else
1209.           $INDUC_S185 = 0
1209.           if ( 0 < $GCS_101 ) then
1209.             if ( $GCS_101 >= 8 ) then !  99.50%
1209.               $TC_372 = and( -8, $GCS_101 )
1209.               $LC_S183 = -$TC_372
1209.               $SI_S184 = 0
1209.               $LIS_b1472 = int( snowdepth_surft[0].L )
1209.               $LIS_b1473 = loc( snowdepth_p )
1209.               if ( $LC_S183 < -15 ) then
1209.                 do
1209.                   $GCS_99 = $LIS_b1472 + $SI_S184
1209.                   $GCS_100 = $LIS_b1473 + $SI_S184
1209.                   0[$GCS_99:8:1,a32].L = 0[$GCS_100:8:1,a8].L
1209.                   8[$GCS_99:8:1,a32].L = 8[$GCS_100:8:1,a8].L
1209.                   $SI_S184 = 128 + $SI_S184
1209.                   $LC_S183 = 16 + $LC_S183
1209.                   if ( $LC_S183 >= -15 ) exit
1209.                 enddo
…
```

Analysing the NEMO results

- NEMO is highly memory-bandwidth bound
- Not much vectorisation (in benchmark version)
- This should favour TX2
- … but … work being done on NEMO code to both improve vectorisation and reduce effect of memory bandwidth
- This may reduce the advantages of processors with greater memory bandwidth

Cray Compiler Version

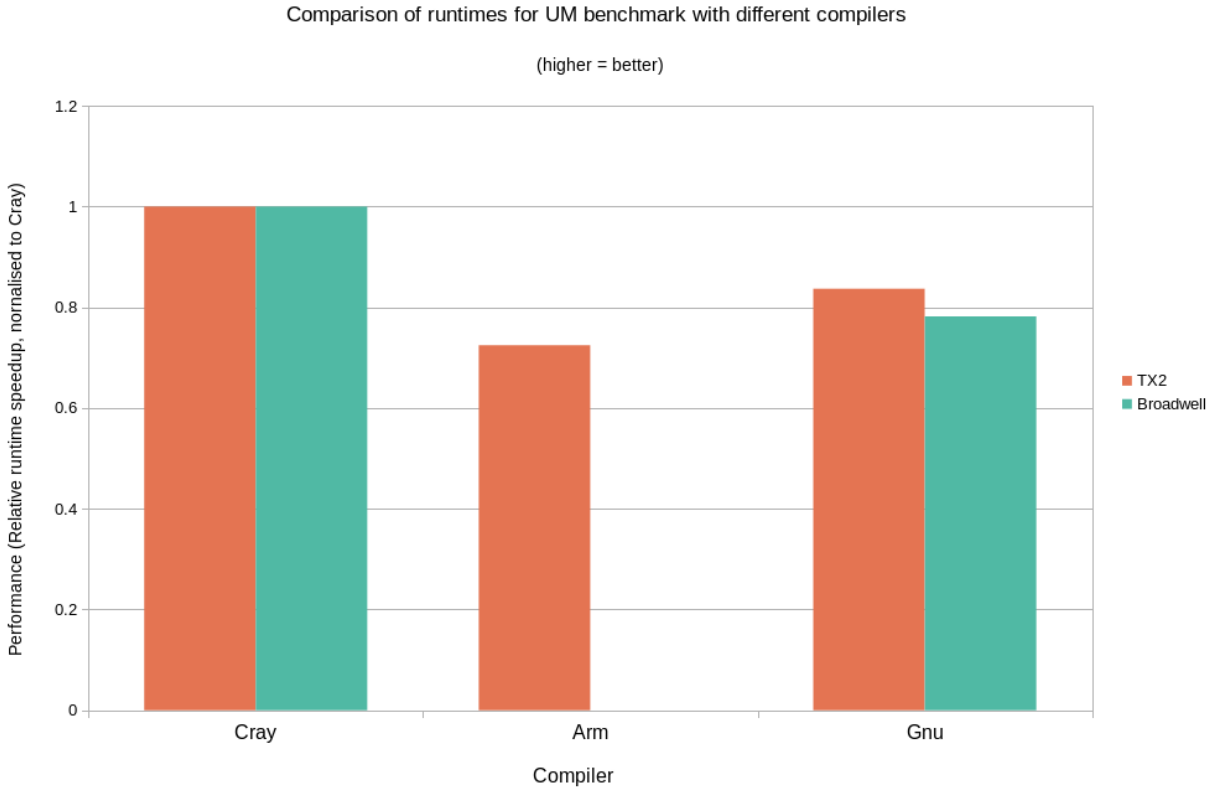


Arm & Gnu (with UM)

Compiler Versions

TX2
Cray = 8.6.4
Arm = 18.1
Gnu = 7.2

Broadwell
Cray = 8.5.8
Gnu = 6.3



Future Work

- Test on full system (higher resolution; multiple nodes)
- This will require new higher resolution UM setup
- Collect information on power consumption
- Try to re-tune compiler flags to latest Cray compiler version

TX2 as an Operational System?

Met Office

Pros

- Easy portability
- Competitive performance
- Good cost/flop

Cons

- Would probably want better vectorisation (SVE?)
- More variability in run length? (Possibly just due to test nodes?)

There is nothing to prevent serious consideration of the use of an operational machine based on the Cavium Thunder X2

Conclusions

Met Office

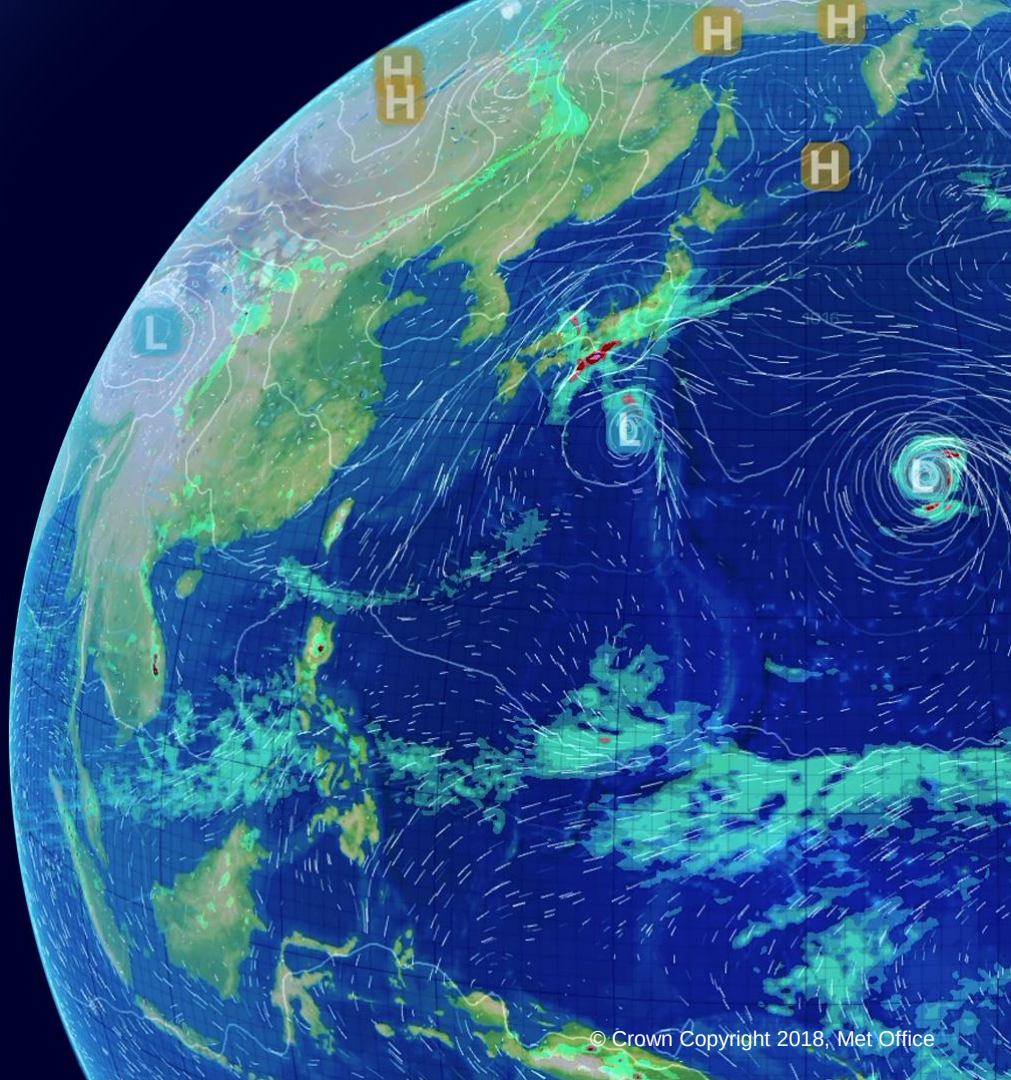
- The UM and NEMO have successfully been ported to run on Isambard on the Cavium Thunder X2 CPU
- This was relatively easy and trouble free to do
- Possible to produce runs with Cray CCE, Gnu, and Arm Compiler toolchains
- Performance is competitive with Intel CPUs
- But the performance characteristics and details differ
- Higher memory bandwidth gives the Thunder X2 a performance boost

Acknowledgements

- GW4 Partners
  - Simon McIntosh-Smith, James Price
- Cray
  - Lucian Anton
- Isambard Support
  - Joe Heaton
- HPC Optimisation Team
  - Sam Cusworth, Maff Glover, Michele Guidolin, Andy Malcom, Paul Selwood
- Other UM and NEMO developers
- Arm

Thank You

Questions?



Spare Slides

What is the Met Office purpose?

Protection

Prosperity

Well-being

- £30bn of economic value to the UK *(2005 – 2015 London Economics, 2015)*
- 14:1 return on tax-payer investment *(London Economics, 2015)*
- Additional £1.2bn to be delivered from 2015 £100m Capital Grant
- Largest weather and climate dedicated supercomputer in the world

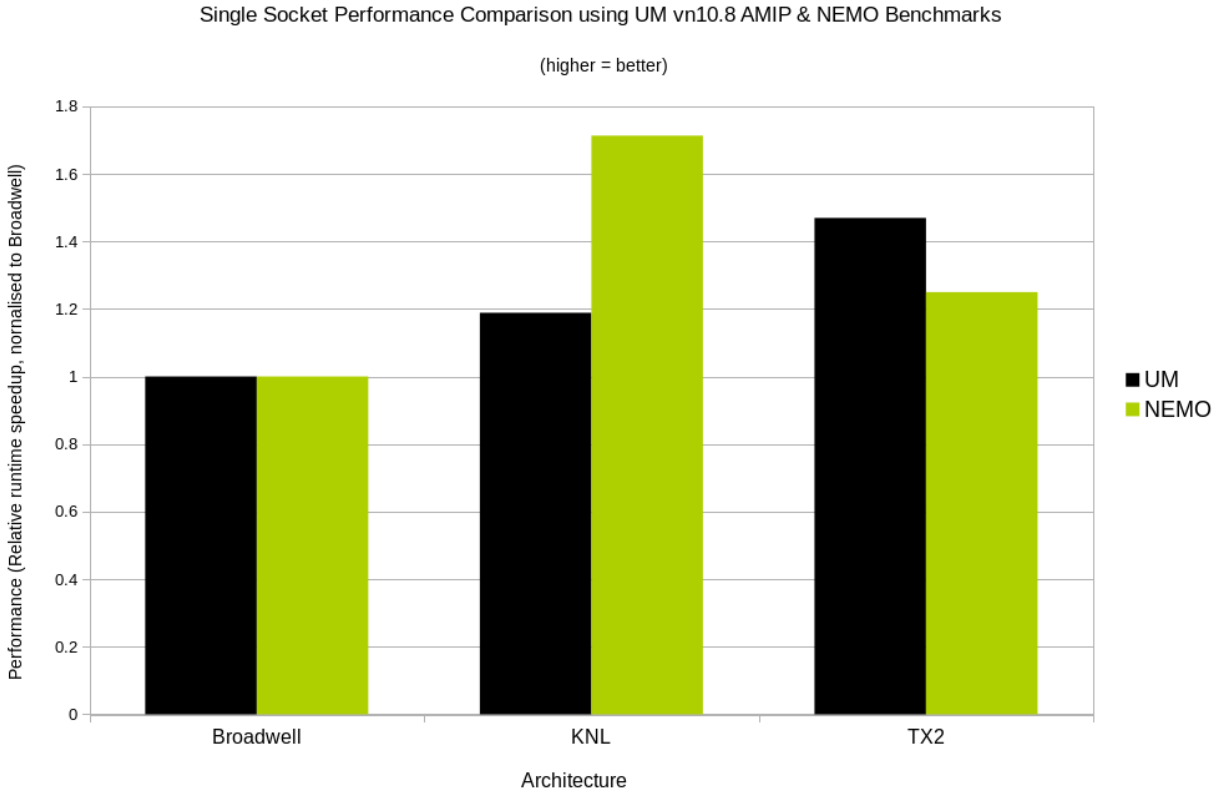
Results

Broadwell
XCE/F – Intel Xeon (Broadwell),
18-core @ 2.1GHz

KNL
XCK – Intel Xeon Phi (Knights landing),
64-core @ 1.3Ghz

TX2
Isambard – Cavium ThunderX2
UM: 32-core @ 2.5GHz
NEMO: 28-core @ 2.0GHz

NB: NEMO on KNL used Intel Compiler



UKV and MOGREPS-UK

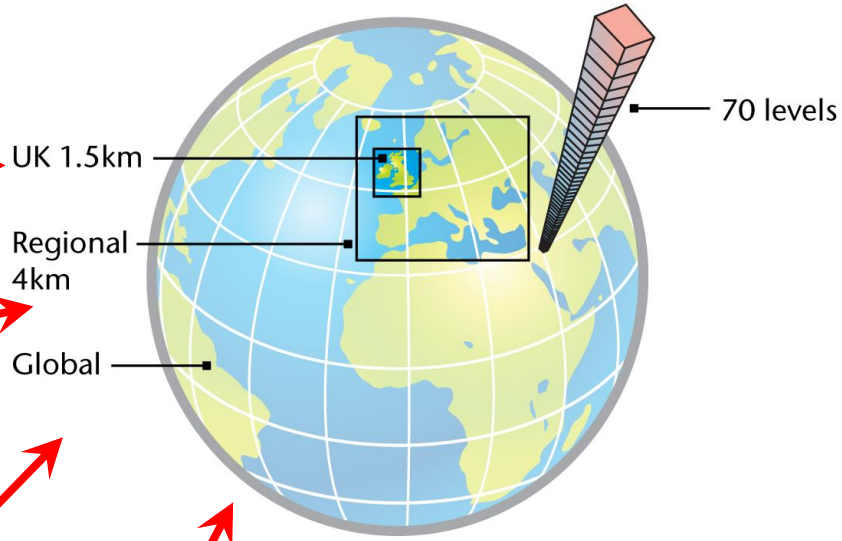
- 1.5km 70L (40km model top)
- 12hr forecast 16 times/day
- 54hr forecast 6 times/day
- 120hr forecast 2 times/day
- 12-member Ensemble - 2.2km 4x/day 54h

Euro4

- 4km 70L (40km model top)
- 66hr forecast twice/day
- 144hr forecast twice/day

Global and MOGREPS-G

- 10km 70L (80km model top)
- 66hr forecast twice/day
- 144hr forecast twice/day
- 18-member Ensemble - 20km 4x/day 7d



Seasonal: GloSea5

- 60km 85L (85km model top)
- ¼ degree Ocean
- 14-member Ensemble
- 7month forecast once/week

NEMO

- NEMO is used by a large community in Europe and world-wide for a wide range of applications : oceanographic research, operational oceanography, seasonal forecast and climate projections

- NEMO is in particular used in 6 Earth System Models within CMIP6 and in Copernicus Marine Services (CMEMS) model-based product.