جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

SHAHEEN
SUPERCOMPUTING LABORATORY

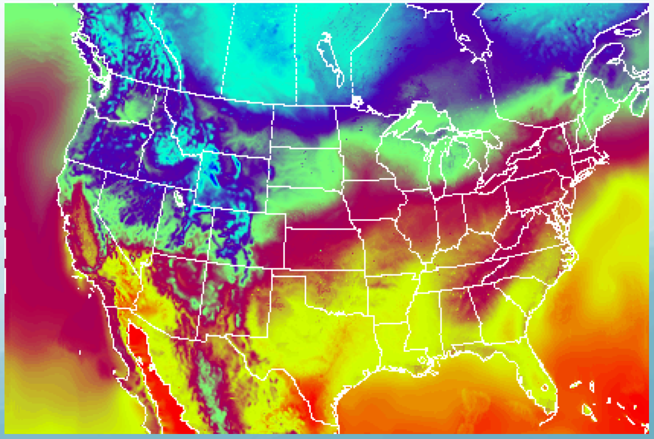
Performance evaluation of Weather Research and Forecast (WRF) on Intel Knights Landing Processor (KNL)

George S. Markomanolis, Saber Feki
KAUST Supercomputing Laboratory
King Abdullah University of Science and Technology

IXPUG Birds of a Feather (BOF),
Supercomputing 2016 (SC16)
Salt Lake City, November 16, 2016

WRF

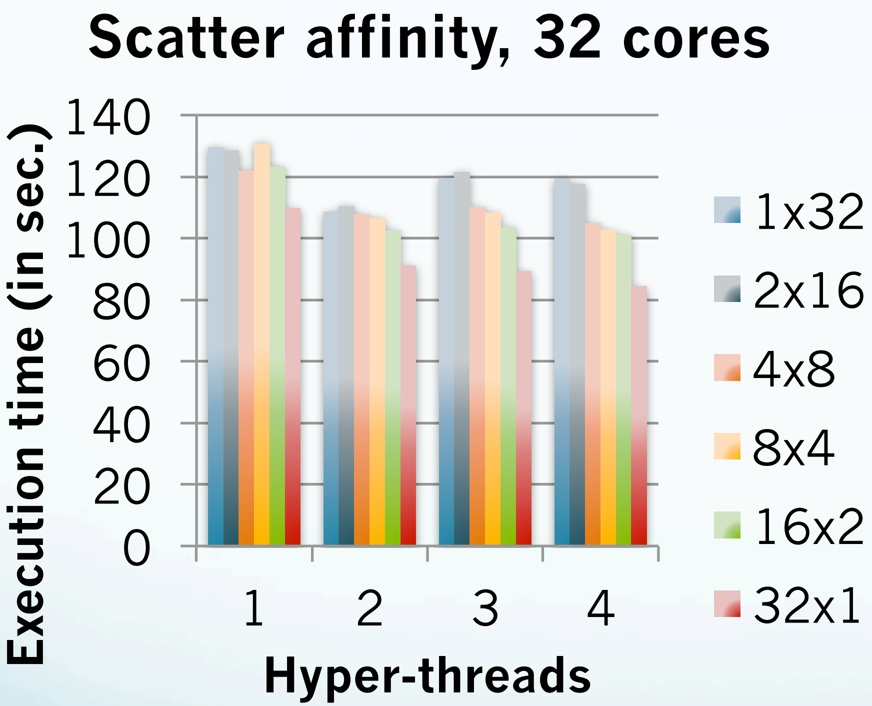
- Weather Research and Forecasting Model (WRF) is one of the most used models on our supercomputer Shaheen II
- We use the domain Conus 12km:
  - All the simulation fits in MCDRAM
  - One Intel KNL node (Phi 7210, 64 cores) was used for all the experiments
  - 3 hours of simulation, saving output files after 3 hours

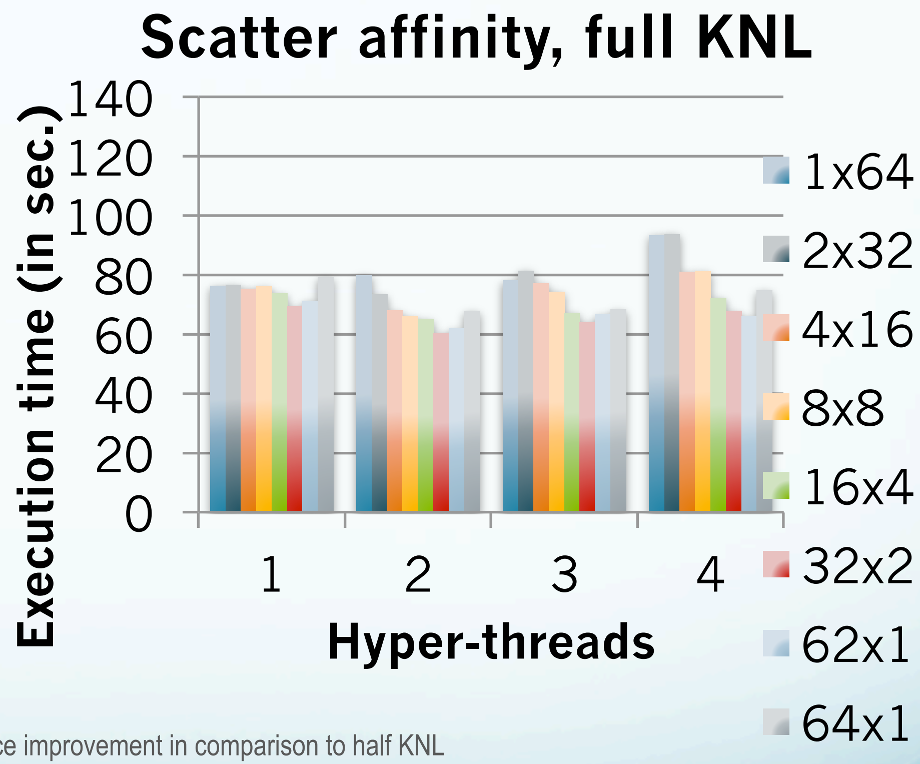


Clustering and memory modes

- We did experiments on various memory modes**, flat**/cache, and clustering modes **Quadrant**/Hemisphere/SNC-4/SNC-2
- Extensive experiments from 1 to 64 MPI processes while using 1 to 64 OpenMP threads (using at least 32 cores totally per simulation), 1-4 hyper threads and OpenMP affinity scatter/compact/balanced

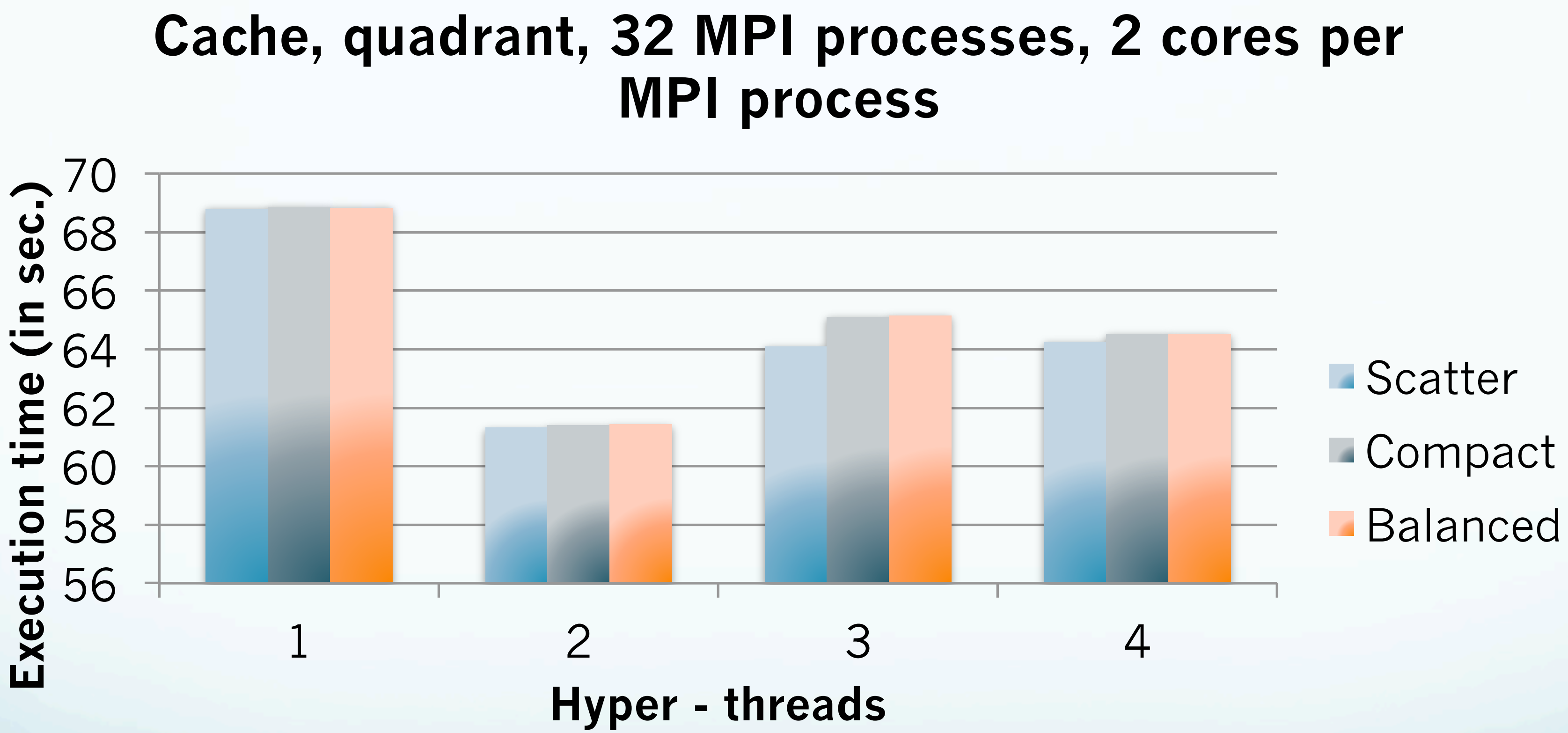
Flat - Quadrant





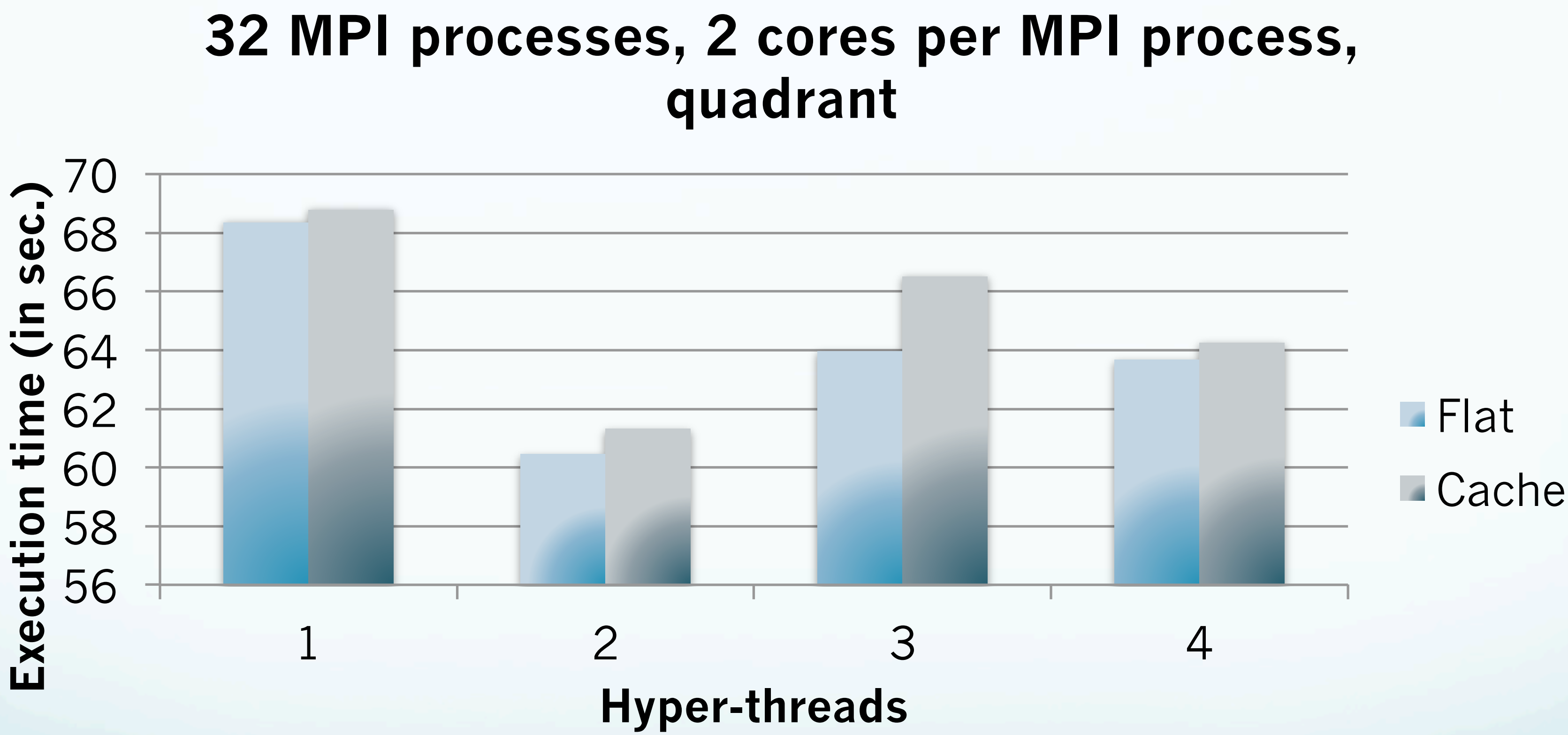
- Using all the physical cores of the KNL, provide at least 19.6% performance improvement in comparison to half KNL
- We have the best performance with 32 cores for MPI, and 4 OpenMP threads by using 2 cores and 2 hyper threads per MPI process (KMP_PLACE_THREADS=2C,2T, numactl –m 1)
- With 4 hyper-threads we usually have better performance than 3 hyper-threads
- We use only MCDRAM for these experiments

Affinity



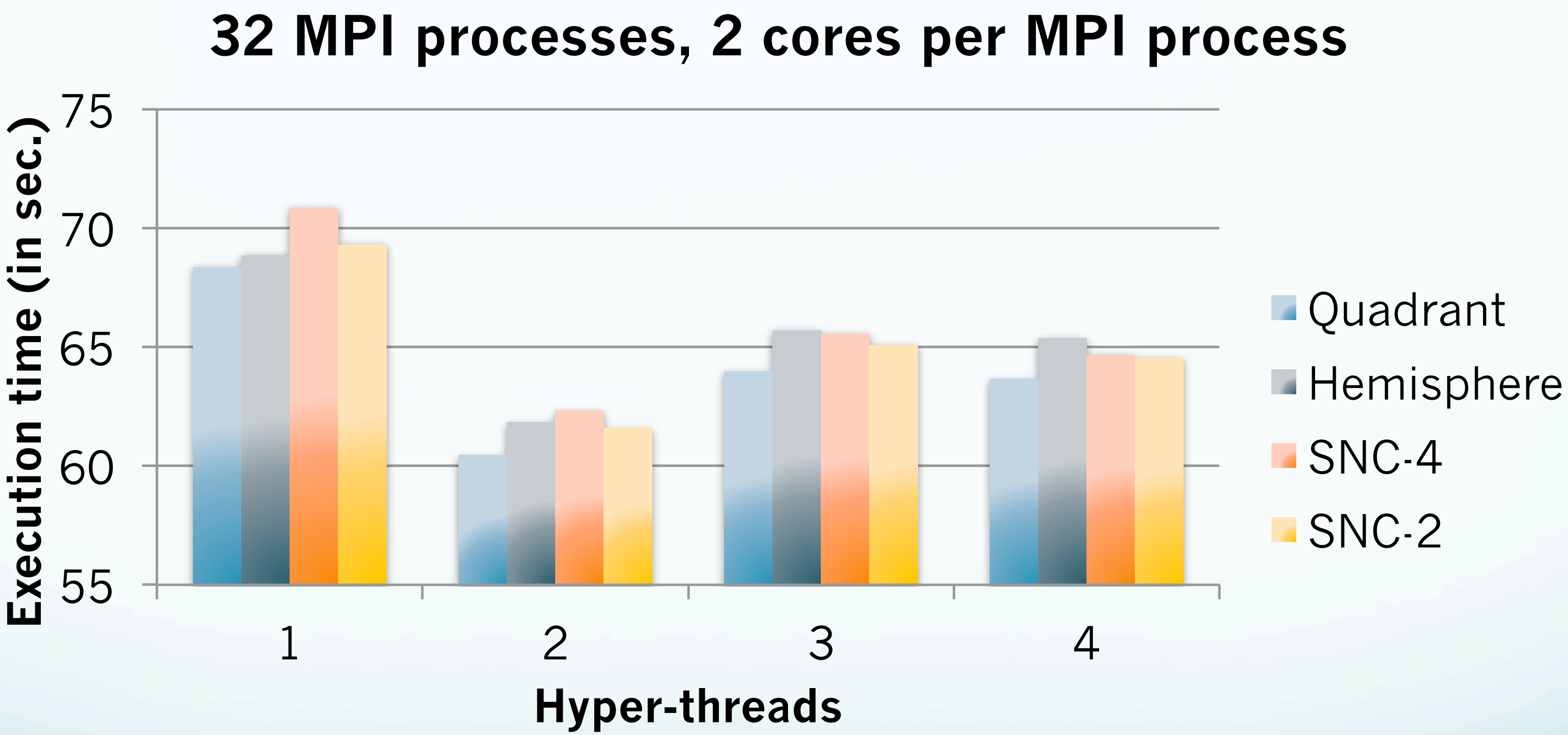
- The scatter affinity behaves better than the rest options, but overall there is variation.
- In this case, with 4 hyper-threads we have similar or better performance than 3 hyper-threads.

Flat vs Cache mode



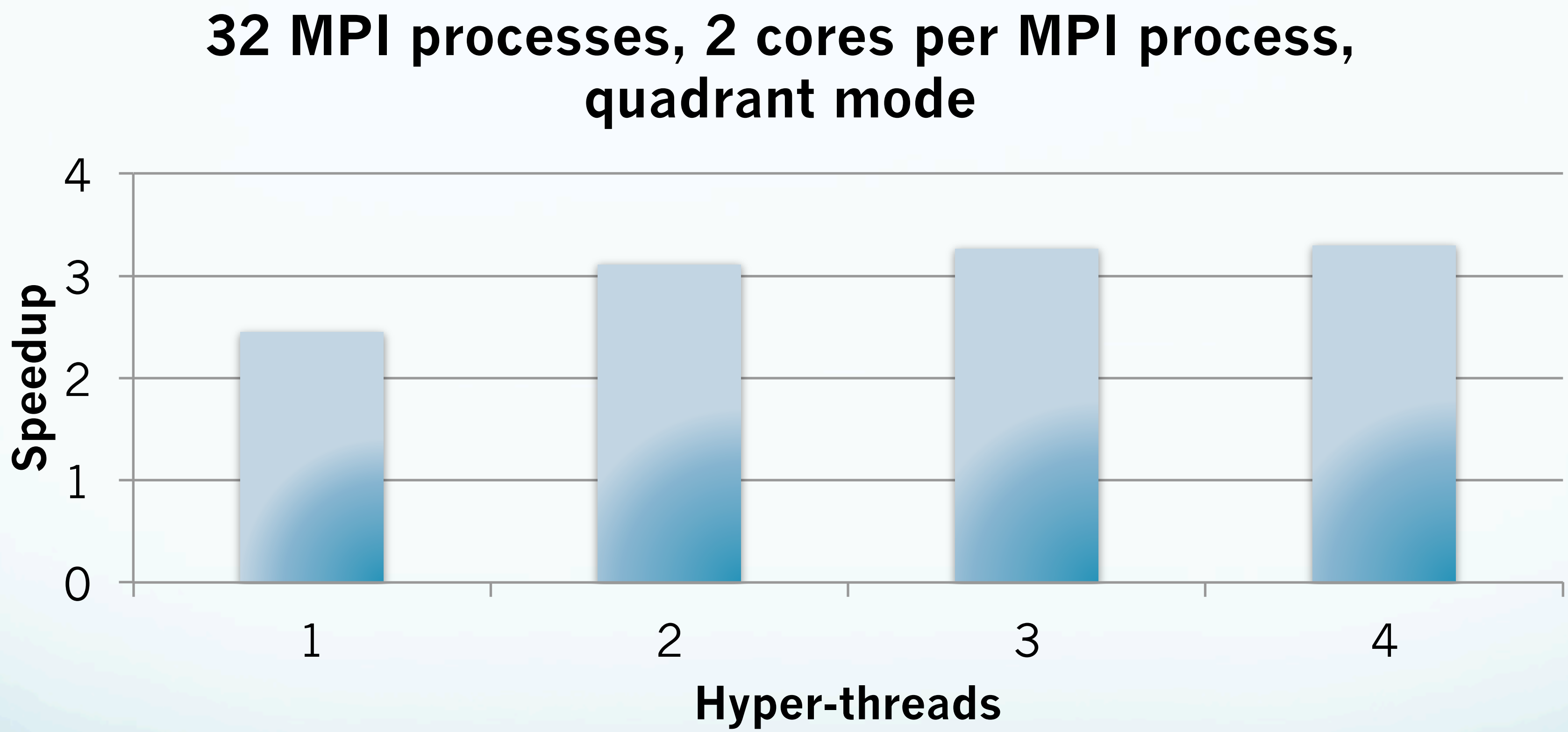
- Flat is overall better than Cache mode for this application up to **4%**.

Comparison of clustering modes



- Quadrant mode behaves more efficient across all the comparisons
- We use an instance with the best provided performance. We select 32 MPI processes, with 2 cores per MPI process

MCDRAM vs DDR4

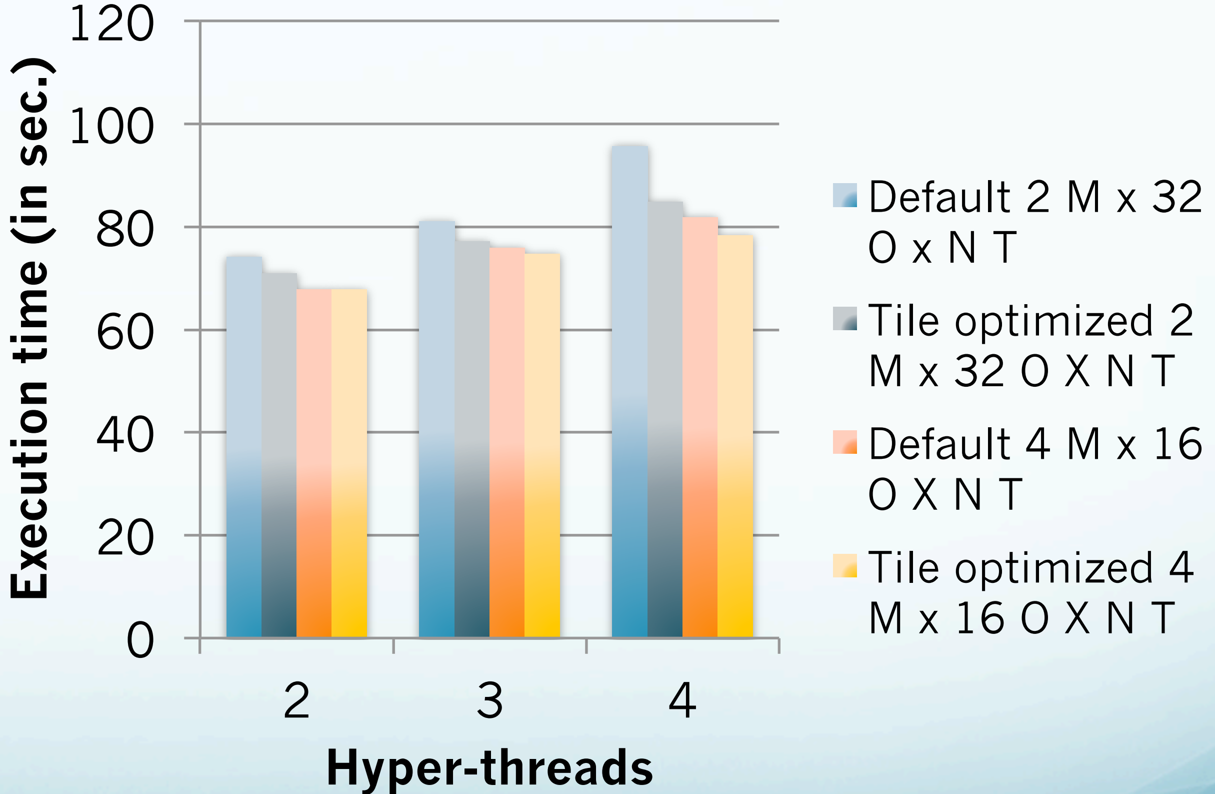


- MCDRAM is at least around to 3 times faster than DDR4 when we use WRF.
- While increasing the hyper-threads, the speedup varies from 3 to 3.3 times.

WRF – Patches and tiles

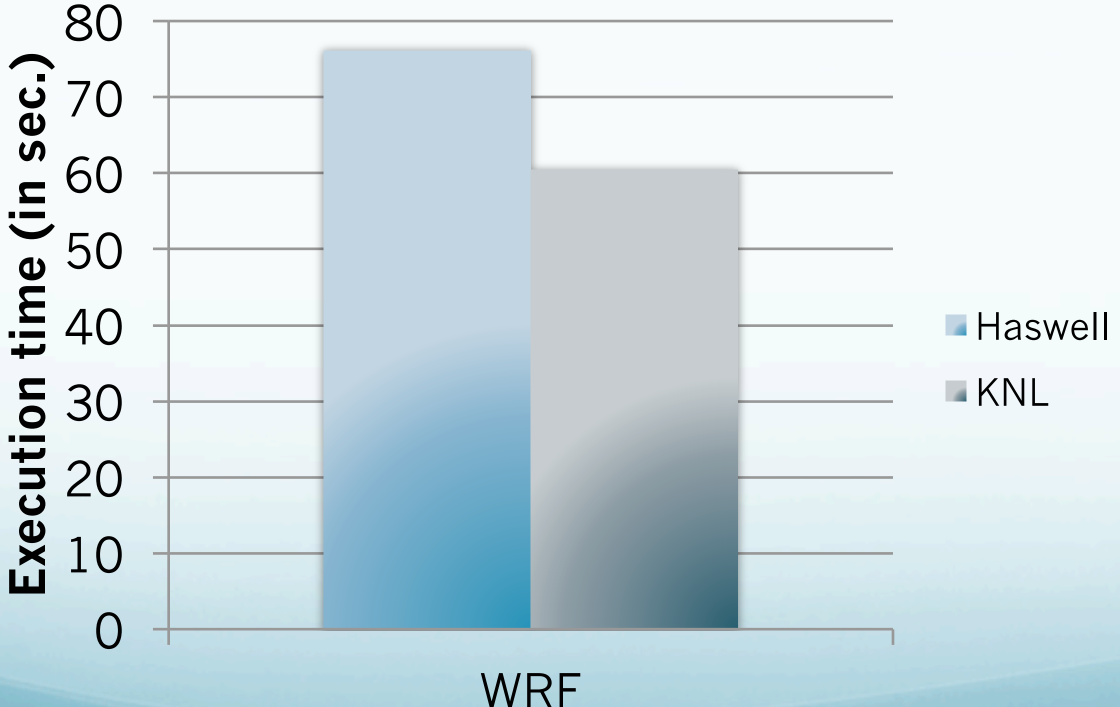
- The domain is distributed in rectangular pieces on different cores and theses pieces can be subdivided into smaller rectangular pieces that are called tiles and can be assigned to threads. We test different size of tiles.

- 2 M x 32 O x 2 T means 2 MPI, 32 OpenMP threads and 2 hyper-threads. The default tile has dimensions 1 x 64. If we change it to 2 x 32 (Tile optimized 2 M x 32 O x 2 T), then we get 4% performance improvement.



Comparison with Haswell

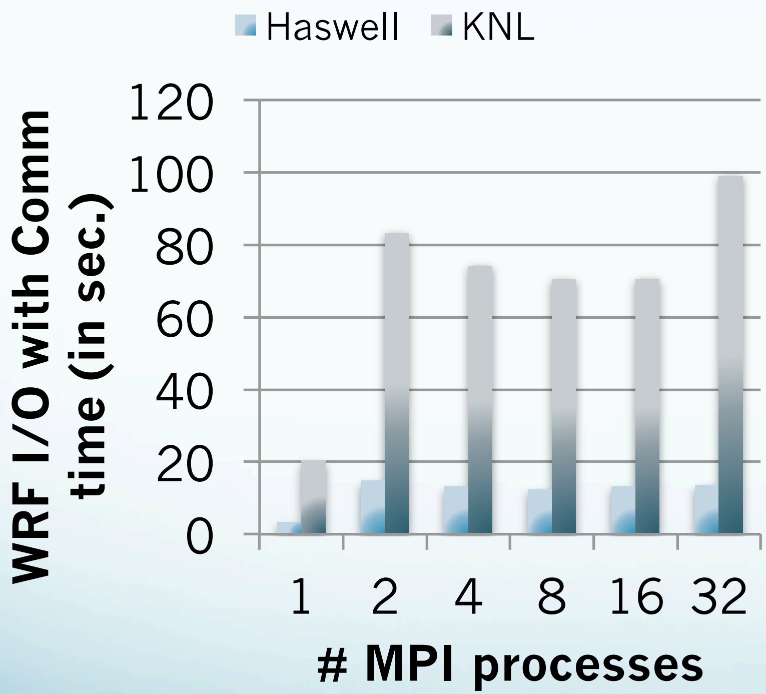
- The performance on KNL is 20,54% faster than dual socket Haswell excluding the I/O time. This is achieved with Flat/Quadrant modes, mapping the application on MCDRAM, with 32 MPI processes, 2 cores per MPI process and 2 threads for each process.

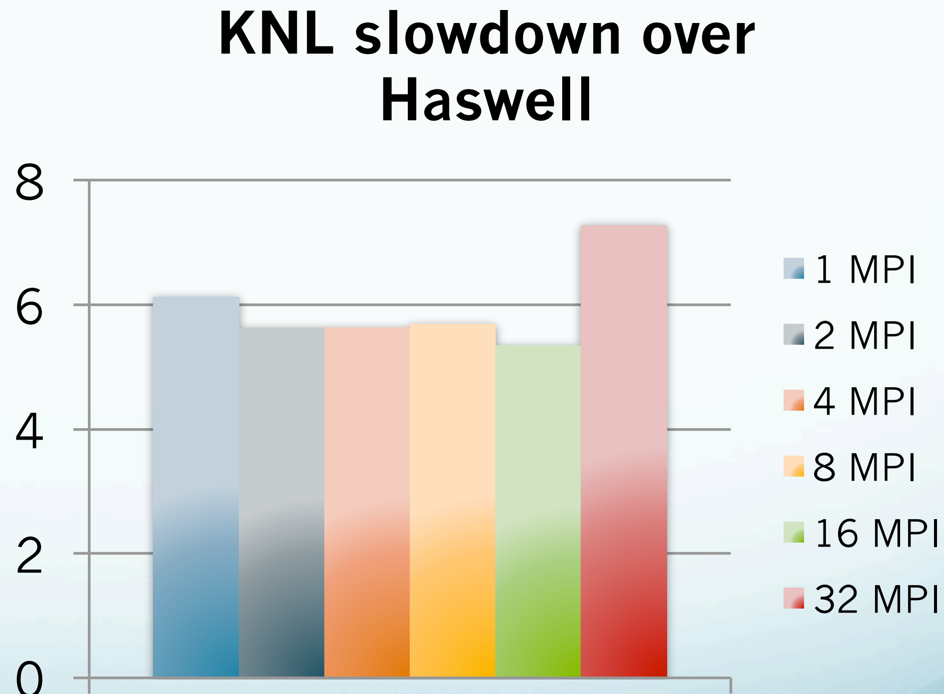


Thank You !

WRF – I/O with Comm

- Serial NetCDF
  - All the MPI processes send data to the root rank and this one writes the data serially (totally ~2GB). The results include the communication and the I/O times





The communication between the participated MPI processes to rank 0 delays a lot the total execution, that's why the time for the total procedure of I/O with 1 MPI process is faster than having more MPI processes.

Summary and Future work

- We did an extensive study of WRF on single node of KNL
- We checked various memory and clustering modes
- We did check the accuracy of the results
- WRF on KNL is 20% faster than dual socket Haswell
- It would be interesting to study the WRF I/O with Intel KNL and Lustre filesystem
- We need to test the scalability of WRF across KNL nodes