Current status of CESM on Xeon Phi systems.

Srinath Vadlamani*(srinathy@ucar.edu)
John Dennis*, Youngsung Kim*, Jim Edwards*
*NCAR
Outline

• We will review Community Earth Systems Model (CESM).
  – It’s nice to know what FC5 really means.
• We will review Xeon Phi systems of interest.
• We will see what the CESM landscape looks like on the Xeon Phi systems.
  – Brief discussion about verifications
• We have some initial Babbage performance results
Review of CESM

- Coupled framework for components consisting of different models.
  - >1.5M lines of Fortran code
- Many collaborators with acute accuracy requirements.
- Few standalone drivers exist.
  - Kernel generator (KGEN) is addressing this.
- HOMME (dynamical core) + atmosphere physics and chemistry + active land (CLM) = FC5 ~ CAM-SE
- FIDEL is only dynamical core with all fluxes from data files [HOMME + coupler].
- Scientist want a “push-button” code.
- Tuning the application must eventually be hidden in application configure scripts.
  - This is very very hard.
Having more than one Xeon Phi cluster aids the sanity.

We have a **CESM BRANCH** that builds out of the box on the following systems (SNB & KNC):

- **Production system**
  - Stampede @TACC – CESM runs
    - MPSS 3.3, intel `{13.1.163,14.0.3,15.0.0}`, impi `{4.1.0.036,4.1.3.049,5.0.0.28}`
    - **FC5 ne16_ne16**: Verified **

- **Research test beds**
  - Pronghorn @NCAR – CESM not running yet.
    - MPSS upgrading from 2.1 to 3.3, intel `{14.0.3}`, impi `{4.0.3.008,5.0.0.28}`
  - Babbage @ NERSC- CESM runs
    - MPSS 3.3, intel `{13.1.2,14.0.3_update,15.0.3_beta}`, impi `{4.1.1,4.1.3,5.0.1_update}`
Variability of throughput due to software stack?

• We have seen KNC OS changes impact throughput rates
  • On the Stampede KNC OS (MPSS @ 2.1) we experienced FC5 KNC throughput = .3 * SNB throughput. Now the CXP repo is at .17*SNB throughput after MPSS upgrade to 3.3.

• Code base evolution -> moving target
  • Getting KN<C/L> application runs in regression suites with performance metrics.

• A bug fix in the 14.0.3 compiler not sufficient
  • Using -mP2OPT_hpo_matrix_opt_framework=0 flag still appears to be necessary CESM
Verification process

• Run to Run reproducibility
  – -no-opt-dynamic-align needs to be tested to allow for full vectorization on SNB
• CESM Statistical Verification Package (CSVP)
  – Create ensemble of all variables (118) on trusted platform by 101 initial temperature perturbations.
  – 1 year simulation
  – Particular compset and resolution
  – 101 member ensemble establishes [min,max] RMSZ score or global mean interval per variable: baseline intervals
  – Compare 3 similar perturbations on ported platform or new algorithm.
  – Success: Less than 10% variables outside of baseline intervals with no repetition among the 3 experiments.
  – *FC5 ne16.... On stampede with compilers.... Has been verified
Don’t only chase slowest component. *not optimal PE-component layout*

Both FC5 and BC5 throughput on Babbage KNC is quite less and needs to be investigated.
Hybrid MPI/OpenMP on KNC

Babbage, 2 KNC cards, FC5, ne16_ne16, intel/15.0.0.beta, impi/5.0.1_udpate1

{16x12, 24x8, …} did not run, possible OMP_STACKSIZE issue
CESM thread scaling

- CAM-SE (HOMME) on KNC appears to prefer MPI only versus Hybrid MPI + OpenMP
- Unexpected results:
  - Edison/Mira: 6-8 threads/node
- Area of future investigation
Conclusion and forward thinking remarks.

- We have a version of CESM ported to three Xeon Phi systems.
  - An increase of users can expand the system list.
  - We need to get BC5 running with latest compiler set to use new features.
  - Experiments are scripted and version controlled for reproducibility.

- Threading performance will be investigated
  - Intel’s Inspector to the rescue?

- Vectorization and memory access will be attacked
  - Profiling tools for Xeon Phi need more support.
    - *Edison* can be a great place to explore craypat.
  - KGEN like kernel generators will be integral.

- Speedup on Xeon Phi is also necessary to pursue verification efforts.
  - BC5 with “working compiler set” is candidate.

- Time to explore VTUNE on Babbage.

- Intel15 ports need to be analyzed.
  - Hopefully the hints can be implemented with success.
References


