(TWO) APPLICATION SHOW CASES ON INTEL® XEON PHI™ PROCESSORS

Dr.-Ing. Michael Klemm
Senior Application Engineer
Software and Services Group

Legal Disclaimer & Optimization Notice

INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS”. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

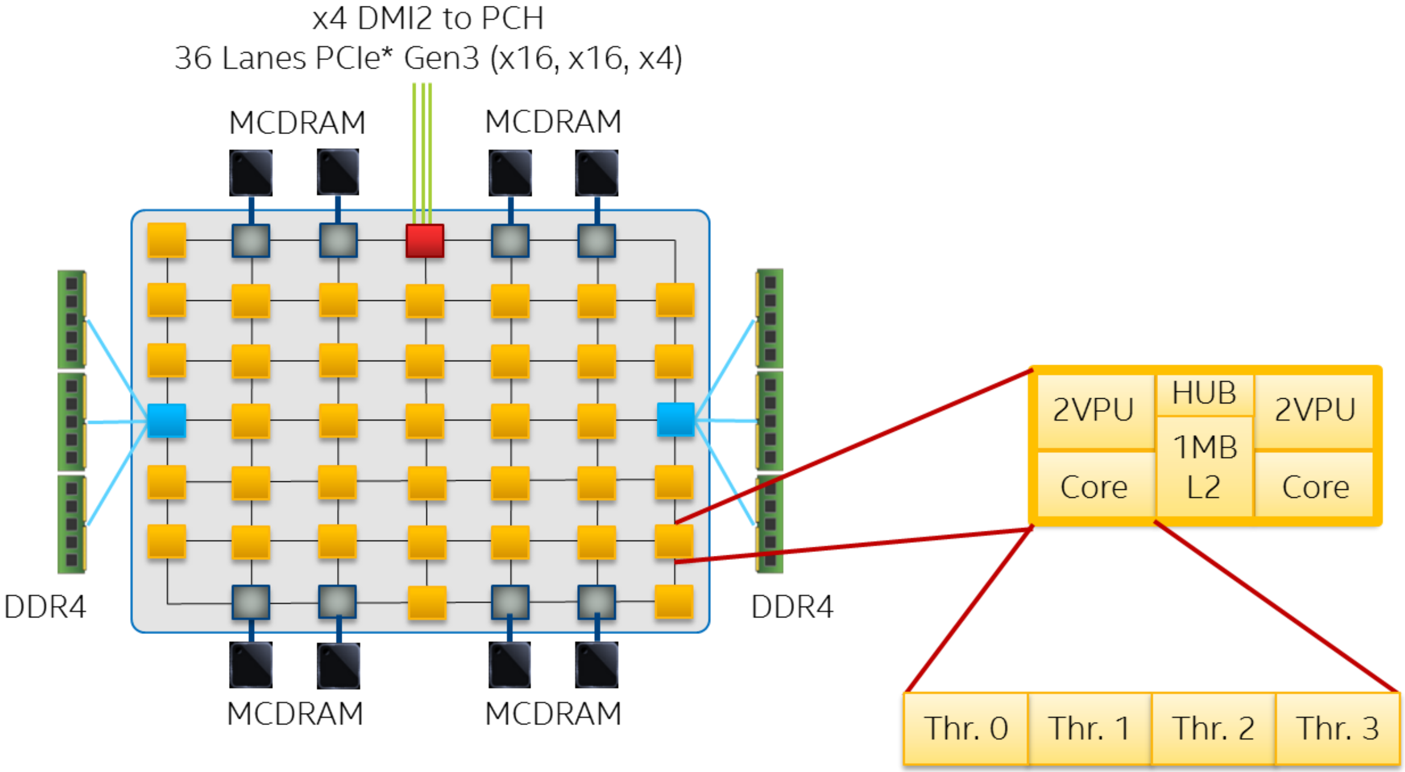
Copyright© 2017, Intel Corporation. All rights reserved. Intel, the Intel logo, Atom, Xeon, Xeon Phi, Core, VTune, and Cilk are trademarks of Intel Corporation in the U.S. and other countries.

Optimization Notice

Intel’s compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804

Intel® Xeon Phi™ Processor Architecture



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

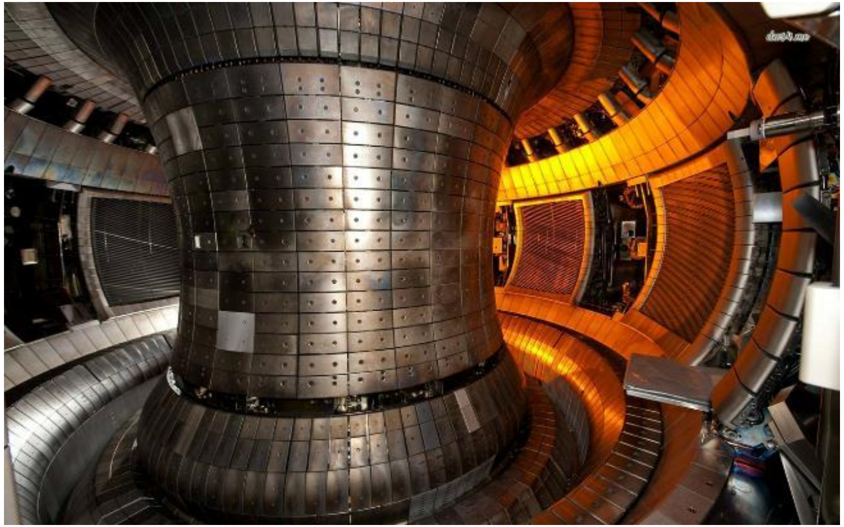
GTC-P

Tokamak plasma physics particle-in-cell (PIC) code

Work by:
Jason Sewall (Intel)

Princeton Gyrokinetic Toroidal Code

- Plasma turbulence simulation
  - Motion of ions through Tokamak
  - Vlassov-Poisson equation using particle-in-cell (PIC)
  - Well-studied in HPC
  - Many ‘leadership-class’ runs and ports



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Algorithm

Charge

- Particles deposit charge onto grid

O(Particles)

Poisson

- Solve Poisson equation over grid

O(Grid)

Field

- Reconstruct electric field over grid

O(Grid)

Smooth

- Filter grid fields

O(Grid)

Push

- Transfer field to particles
- Move particles (in phase space)

O(Particles)

Shift

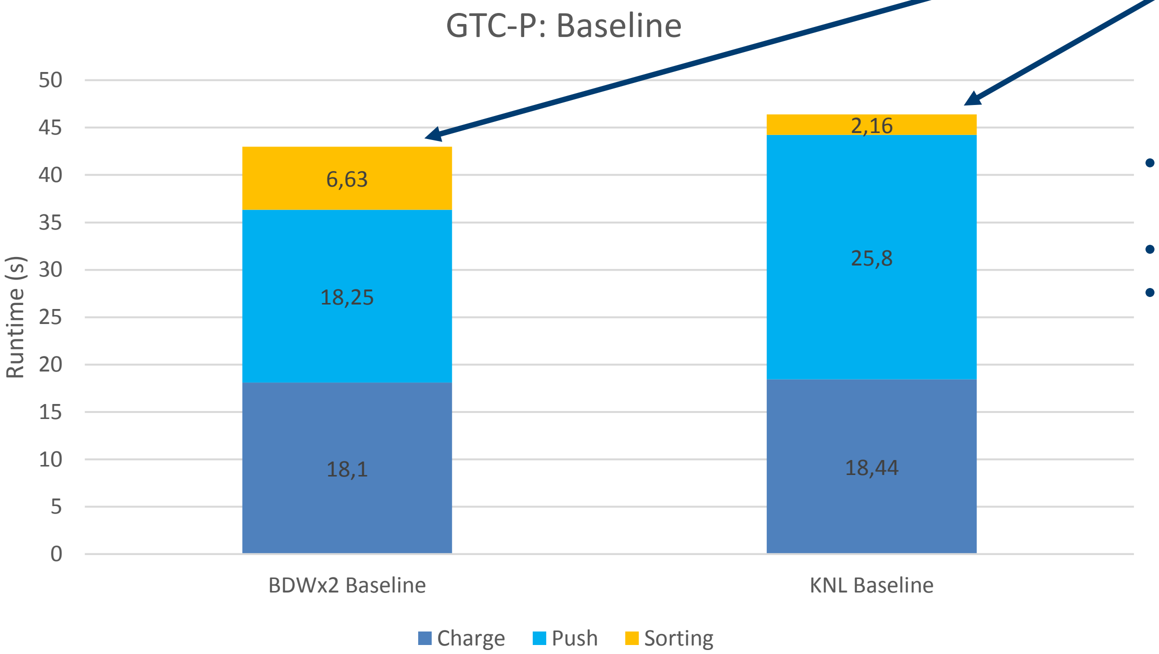
- Move particles between MPI ranks

O(Particles)

Particles >>> Grid
(for this code)

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Optimizations (BASELINE)



KNL ~8% slower than 2xBDW

- B-1rank-half problem:
  - Run with 1 rank and 400 particles
- KNL results from Xeon Phi 7250
- BDW results from 2x Xeon E5-2698 v4

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Optimizations to help vectorization

- Avoid excessive memoization
  - Gathers expensive, can be avoided sometimes

```
im  = ii;
im2 = ii + 1;

tdumtmp = pi2_inv * (tflr - zetatmp * qtinv[im]) + 10.0;
tdumtmp2 = pi2_inv * (tflr - zetatmp * qtinv[im2]) + 10.0;

tdum = (tdumtmp - ( int )tdumtmp) * delt[im];
tdum2 = (tdumtmp2 - ( int )tdumtmp2) * delt[im2];

j00 = abs_min_int(mtheta[im] - 1, ( int )tdum);
j01 = abs_min_int(mtheta[im2] - 1, ( int )tdum2);

jtion0tmp = igrid[im] + j00;
jtion1tmp = igrid[im2] + j01;
```

- Minimize type conversions

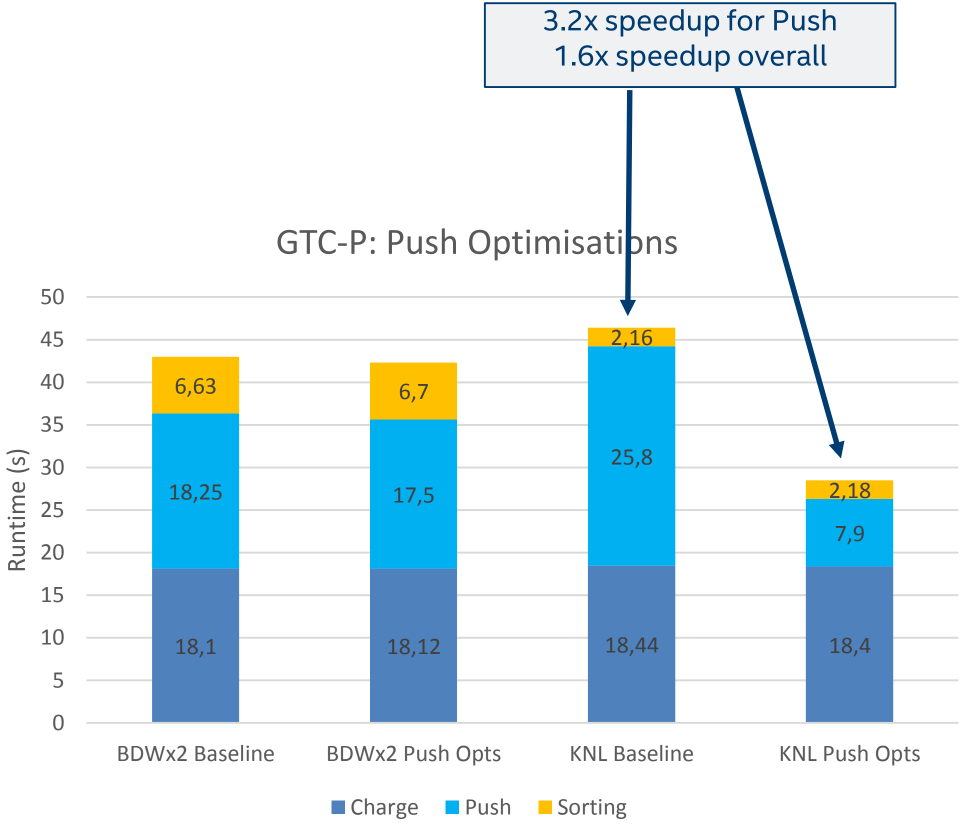
```
const real im_r  = ii_r;
const real im2_r = ii_r + 1.0;

const real mth_im_r  = poloidal_mtheta(im_r, mtheta_a, mtheta_b, mthetamax_r);
const real mth_im2_r = poloidal_mtheta(im2_r, mtheta_a, mtheta_b, mthetamax_r);

const real pgrid_base = igrid[(int) im_r];
const real pgrid_next = pgrid_base + mth_im_r + 1.0;

const real qtinv_m  = poloidal_qtinv(im_r, q0, q1, q2, ainv, a0, deltar,
mth_im_r);
const real qtinv_m2 = poloidal_qtinv(im2_r, q0, q1, q2, ainv, a0, deltar,
mth_im2_r);

const real tdumtmp  = tflr - zetatmp_pi2 * qtinv_m + 10.0;
const real tdumtmp2 = tflr - zetatmp_pi2 * qtinv_m2 + 10.0;

const real tdum  = fmod(tdumtmp, 1.0) * mth_im_r;
const real tdum2 = fmod(tdumtmp2, 1.0) * mth_im2_r;

const real j00 = abs_min_real(mth_im_r - 1.0, floor(tdum));
const real j01 = abs_min_real(mth_im2_r - 1.0, floor(tdum2));

const int jtion0tmp = (int) (pgrid_base + j00);
const int jtion1tmp = (int) (pgrid_next + j01);
```

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Optimizations (PUSH)

- Large ‘diagnostic’ branch in code
  - Only active for certain iterations
  - Multiversion code so extra code not in ‘normal’ loop
- Strip-mining loop can help alignment
- Narrowing masks from whole-loop to just write-masking
- Marking reductions essential for correctness

```
#pragma omp for nowait
for (int mo = 0; mo < mi; mo += 16) {
    real *__restrict__ z0mo  = particle_data->z0 + mo;
    real *__restrict__ z1mo  = particle_data->z1 + mo;
    real *__restrict__ z2mo  = particle_data->z2 + mo;
    ....
    #pragma omp simd aligned(z0mo, z1mo, z2mo, ... : 64) \
        simdlen(16) \
       reduction(+ : particles_energy_a, ...)
    for (int v = 0; v < 16; v++) {
         const real zion2m = z2mo[v];
         const int valid = v + mo < mi && !gtc_hole(zion2m);
         ...
    }
 }
```



3.2x speedup for Push
1.6x speedup overall

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Optimizations (Charge)

```
#pragma omp for
for (m = 0; m < mi; m++) {
       zetatmp = z2[m];
       if (zetatmp == HOLEVAL) {
           continue;
       }
       <later>
       densityi_part[ij1] += d1;
       densityi_part[ij1 + 1] = +d2;
       densityi_part[ij1 + mzeta + 1] += d3;
       densityi_part[ij1 + mzeta + 2] += d4;

       densityi_part[ij2] += d5;
       densityi_part[ij2 + 1] = +d6;
       densityi_part[ij2 + mzeta + 1] += d7;
       densityi_part[ij2 + mzeta + 2] += d8;
}
```

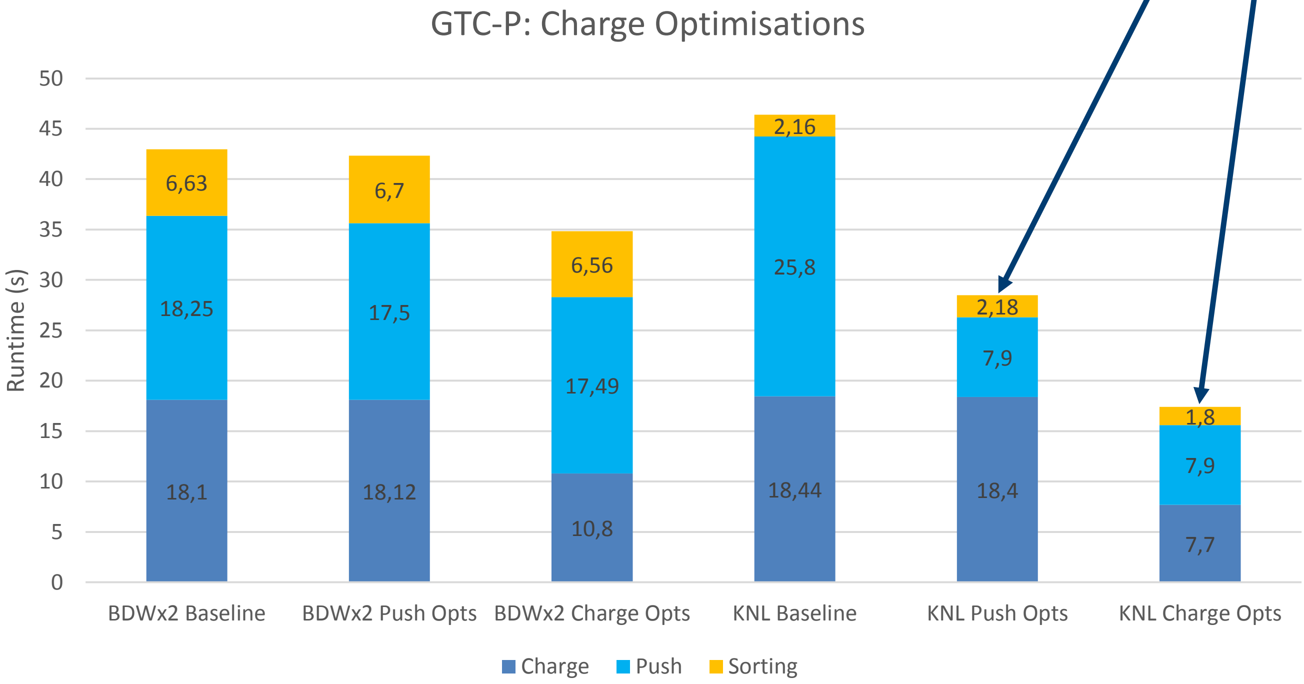
```
#pragma omp declare simd simdlen(16)
static void chargei_update(const int offs, wreal *addr, const real del) {
    #pragma omp ordered simd
    { addr[offs] += del; }
}


#pragma omp for
for (int mo = 0; mo < mi; mo += 16) {
       real *__restrict z0mo = particle_data->z0 + mo;
       real *__restrict z1mo = particle_data->z1 + mo;
        real *__restrict z2mo = particle_data->z2 + mo;
        real *__restrict z4mo = particle_data->z4 + mo;
        real *__restrict z5mo = particle_data->z5 + mo;

#pragma omp simd aligned(z0mo, z1mo, z2mo, z4mo, z5mo : 64) simdlen(16)
          for (int v = 0; v < 16; ++v) {
               const real zetatmp      = z2mo[v];
               const int  valid        = v + mo < mi && !gtc_hole(zetatmp);
               <lots of code>
               if (valid) {
                      chargei_update(ij1, densityi_part, wz0 * wt00);
                      chargei_update(ij1 + 1, densityi_part, wz1 * wt00);
                      chargei_update(ij1 + mzeta + 1, densityi_part, wz0 * wt10);
                      chargei_update(ij1 + mzeta + 2, densityi_part, wz1 * wt10);

                      chargei_update(ij2, densityi_part, wz0 * wt01);
                      chargei_update(ij2 + 1, densityi_part, wz1 * wt01);
                      chargei_update(ij2 + mzeta + 1, densityi_part, wz0 * wt11);
                      chargei_update(ij2 + mzeta + 2, densityi_part, wz1 * wt11);
               }
       }
}
```

- Strip-mining loop can help alignment
- Narrowing masks from whole-loop to just write-masking helpful
- Write-conflicts can be helped with ordered simd
    - Or vconflict + scatter

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Optimizations (Charge)

2.3x speedup for Charge
1.6x speedup overall



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Optimizations (Sorting)

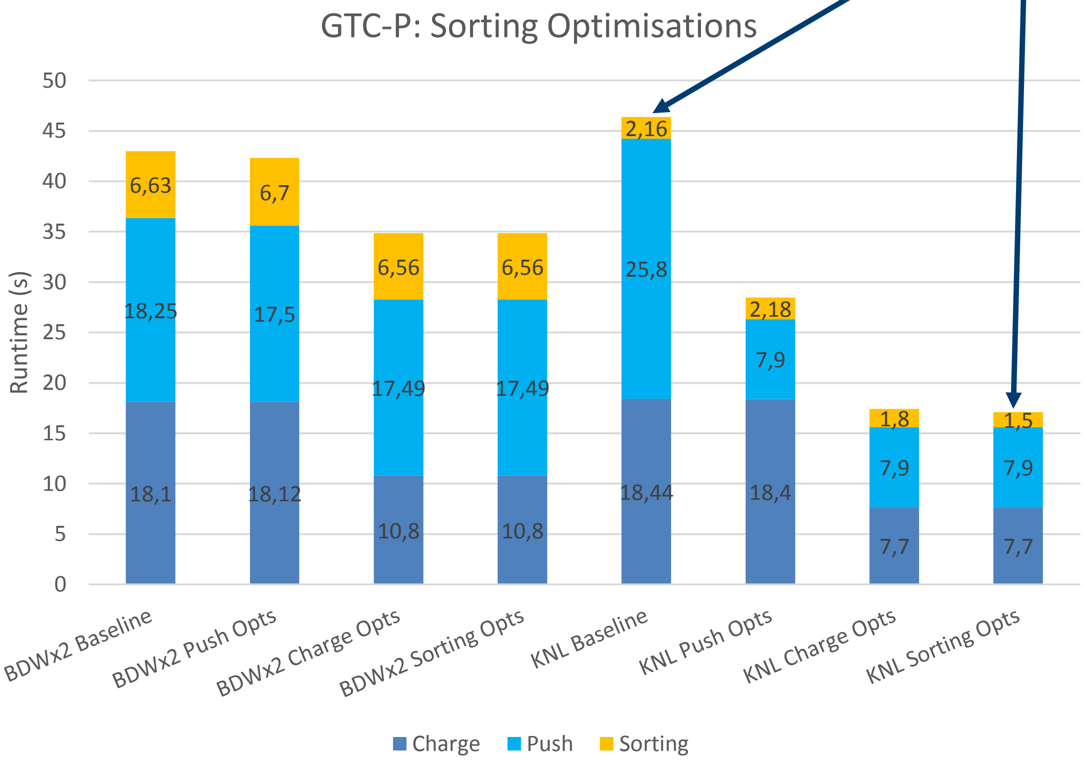
1.2x speedup for Sort
KNL now ~2x faster than 2xBDW
On KNL optimisations deliver ~2.6x cumulative speedup

Unnecessary pressure on TLB:

```
#pragma omp for
        for (m = 0; m < mi_new; m++) {
            z0[m] = z00[m];
            z1[m] = z01[m];
            z2[m] = z02[m];
            z3[m] = z03[m];
            z4[m] = z04[m];
       }
```

Use vectors, alignment, and copy 1 at a time:

```
#pragma omp for simd schedule(static:simd) aligned(z0,z00:64) nowait
        for (m = 0; m < mi_new; m++)
            z0[m] = z00[m];
#pragma omp for simd schedule(static:simd) aligned(z1,z01:64) nowait
        for (m = 0; m < mi_new; m++)
            z1[m] = z01[m];
#pragma omp for simd schedule(static:simd) aligned(z2,z02:64) nowait
        for (m = 0; m < mi_new; m++)
            z2[m] = z02[m];
#pragma omp for simd schedule(static:simd) aligned(z3,z03:64) nowait
        for (m = 0; m < mi_new; m++)
         z3[m] = z03[m];
#pragma omp for simd schedule(static:simd) aligned(z4,z04:64) nowait
        for (m = 0; m < mi_new; m++)
            z4[m] = z04[m];
```



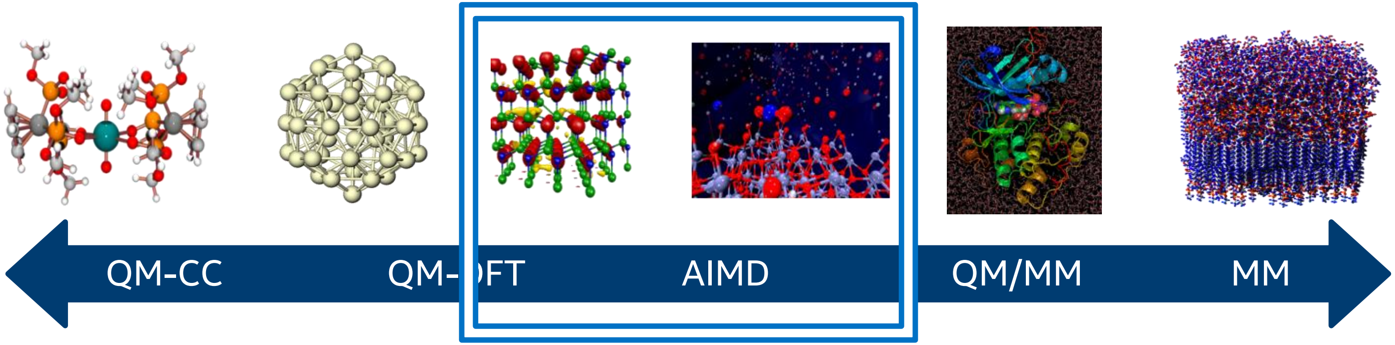
Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

NWCHEM AIMD

NWChem Ab-initio Molecular Dynamics

Work by:
E. Bylaska (PNNL), Matthias Jacquelin (LBL), Bert de Jong (LBL), Michael Klemm (Intel)

Introduction: Plane Wave Methods



$$(-1/2)\nabla^2\Psi + V_{ext}\Psi + V_H\Psi + V_{xc}\Psi + V_{x,exact}\Psi = E\Psi$$

$$\langle\Psi_i|\Psi_j\rangle = \delta_{ij}$$

- 100-1000 atoms, uses plane wave basis
- Many FFTs and DGEMM operations
- “Meaty”: Lots of FLOPs, but also bandwidth sensitive

N_eN_g

$(N_aN_g + N_gLogN_g + N_eN_g) + N_aN_eN_g$

$N_eN_gLogN_g+N_eN_g+2N_gLogN_g+N_g+N_eN_g$

$N_eN_gLogN_g + N_eN_g$

$N_e(N_e+1)N_g LogN_g$

$N_e^2N_g + N_e^3$

N_a - number of atoms
N_e - number of electrons
N_g – size of FFT grid

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiarie r names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Strong Scaling is Key

- 20 psec of simulation time ≈ 200,000 steps
  - 1 sec/step = 2-3 days simulation time
  - 10 sec/step = 23 days simulation time
  - 13 sec/step = 70 days simulation time
- Mesoscale phenomena at longer time scales
  - Assume 1 sec/step
  - 100 psec = 10-15 days simulation time
  - 1 nsec = 100 - 150 days simulation time
- Strong scaling required to reduce time per time step as much as possible
  - At least below 1sec/step

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

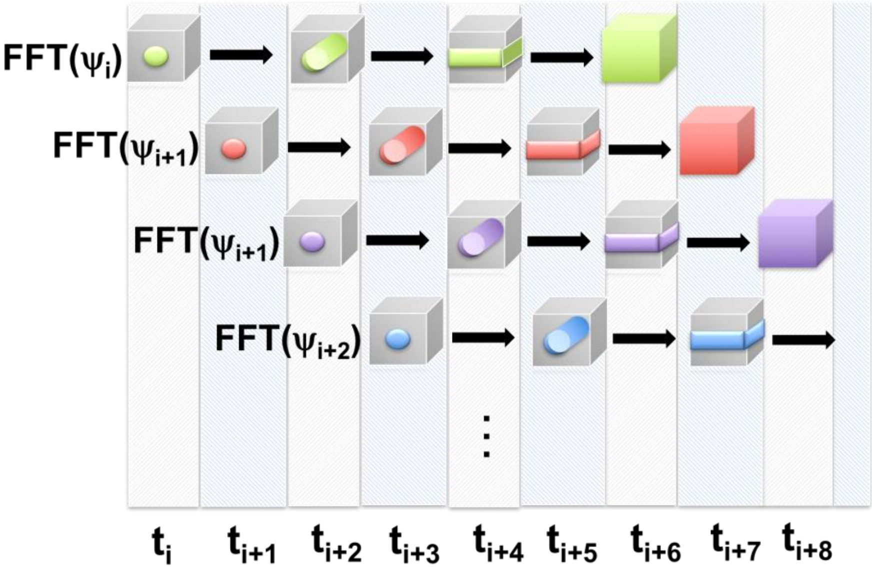
Strong Scaling is Key

- 20 psec of simulation time ≈ 200,000 steps
  - 1 sec/step = 2-3 days simulation time
  - 10 sec/step = 23 days simulation time
  - 13 sec/step = 70 days simulation time
- Mesoscale phenomena at longer time scales
  - Assume 1 sec/step
  - 100 psec = 10-15 days simulation time
  - 1 nsec = 100 - 150 days simulation time
- Strong scaling required to reduce time per time step as much as possible
  - At least below 1sec/step

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

3D FFTs – Pipelined Implementation

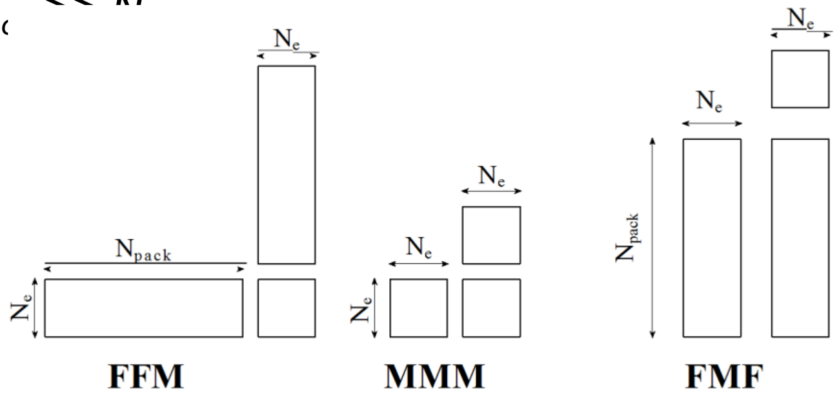
- Performed at each step
  - 2 Ne 3D FFTs for DFT
  - Plus (Ne+1)*Ne 3D FFTs for hybrid DFT
- In reciprocal space, sphere of radius Ecut is stored
- 3D FFTs are pipelined
  - Overlap communication and computation
  - Latency reduction
  - N2 1D FFTs per stage execute in parallel



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

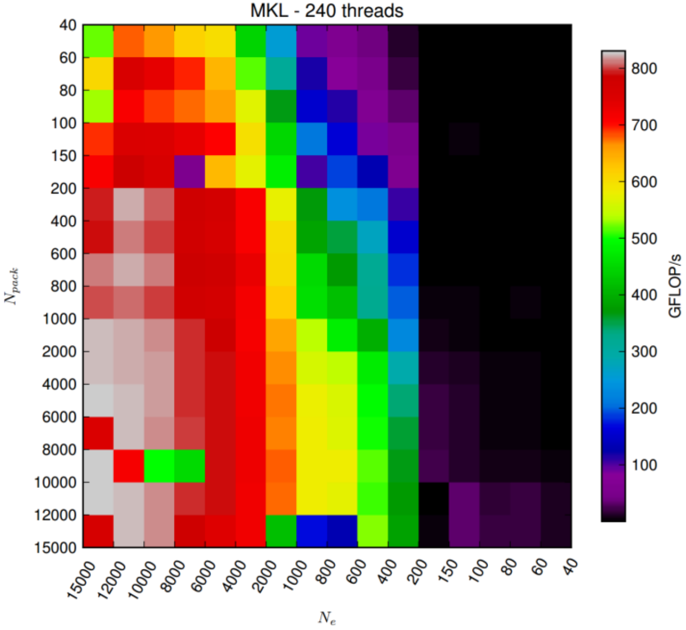
Lagrange Multiplier

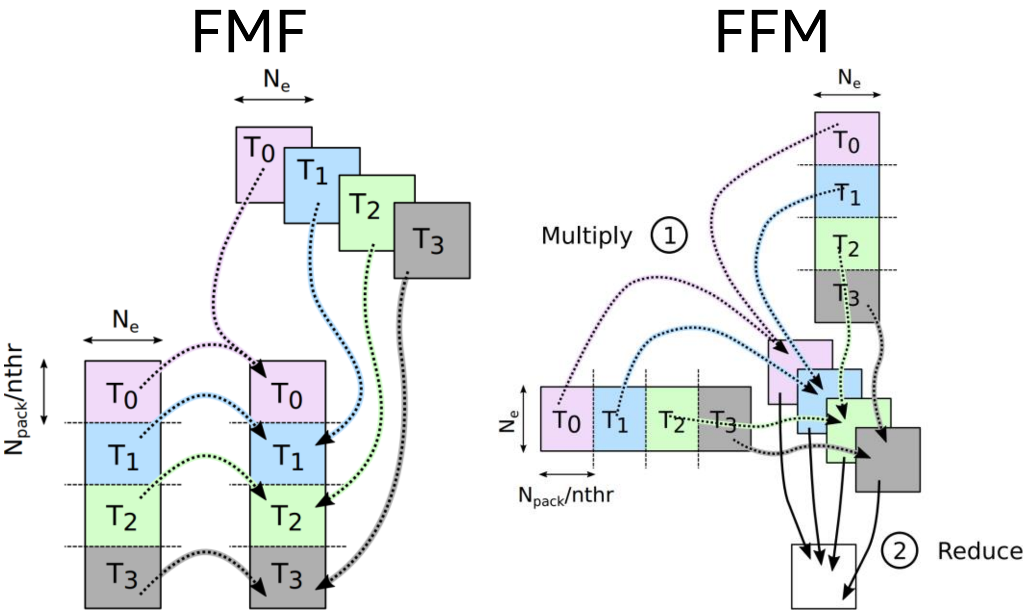
- Sequence of matrix products of shape F or M
  - F: N_{pack} x N_e or N_e x N_{pack} matrix (tall & skinny)
  - M: N_e x N_e matrix
  - In general: N_{pac} >> N



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Lagrange Multiplier – Parallelization





Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Experimental Setup – NERSC Cori

- “Haswell”, HSW
  - Cray* XC40
  - 2S Intel® Xeon® E5-2698v3 processors
  - 32 cores, no Hyper-Threading
  - 2.3 GHz clock frequency
  - 128 GB of DDR4 at 2133 MHz
  - Cray* Aries* w/ Dragonfly

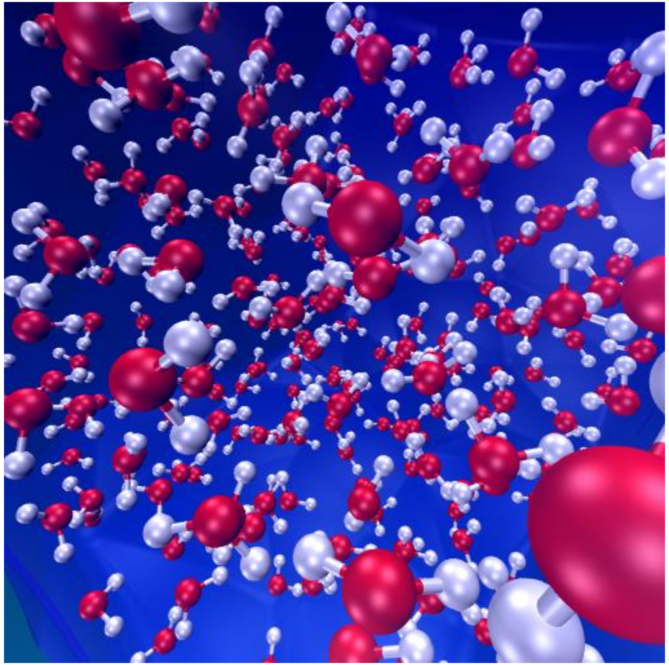
- “Knights Landing”, KNL
  - Cray* XC40
  - Intel® Xeon Phi™ 7250 processors
  - 68 cores w/ 4 hardware threads
  - 1.4 GHz clock frequency
  - 96 GB of DDR4 at 2400 MHz
  - Cache mode
  - Quadrant cluster mode
  - Cray* Aries* w/ Dragonfly

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

*Some names and brands may be claimed as the property of others.

Experimental Setup – Benchmarks

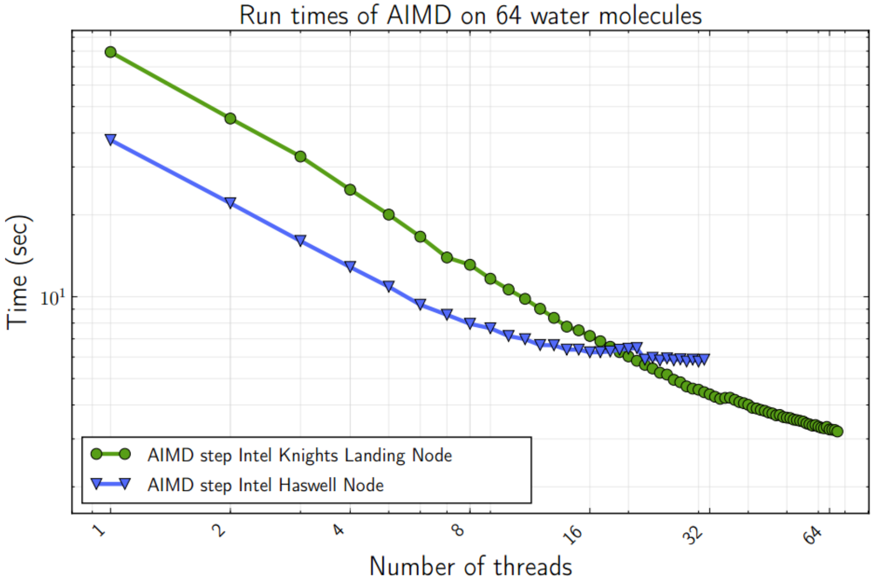
- water64:
  - 64 water molecules in a box
  - test intra-node strong scaling
- water256:
  - 256 water molecules
  - test cluster strong scaling
  - N_e=2056
  - N_g=5,832,000 (180^3)
  - N_{pack}=437,000



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Intra-node Performance

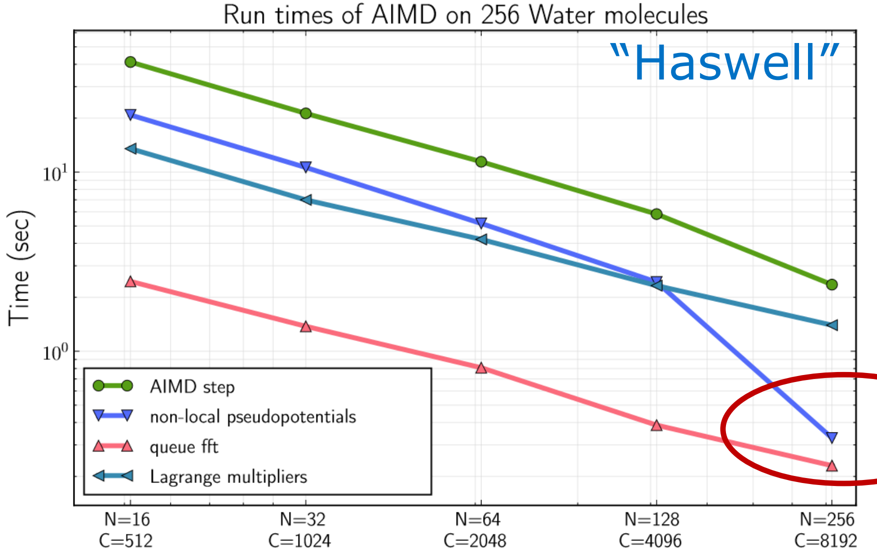
- Insight into performance without fabric effects
- Xeon node saturates at about 16 cores, reaching memory bandwidth limits
- Xeon Phi node keeps strong scaling due to the on-package cache memory
- 1.8x speed-up of KNL over HSW node



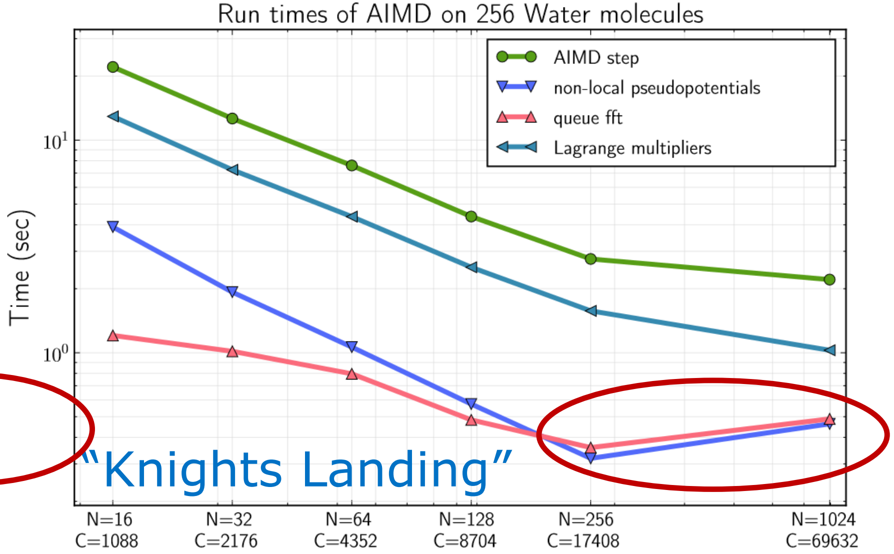
Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. System configuration: Cray* XC40 system, 2S Intel® Xeon® E5-2698v3 processor, Intel® Hyper-Threading technology disabled, 128 GB of DDR4 (8x 16 GB, 2133 MHz), Cray* Aries interconnect with Dragonfly topology; Cray* XC40 system Intel® Xeon Phi™ 7250 processors, 96 GB of DDR4 (6x 16GB, 2400 MHz), quadrant cluster mode, MCDRAM in cache mode, Cray* Aries interconnect with Dragonfly topology.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Performance



"Haswell"



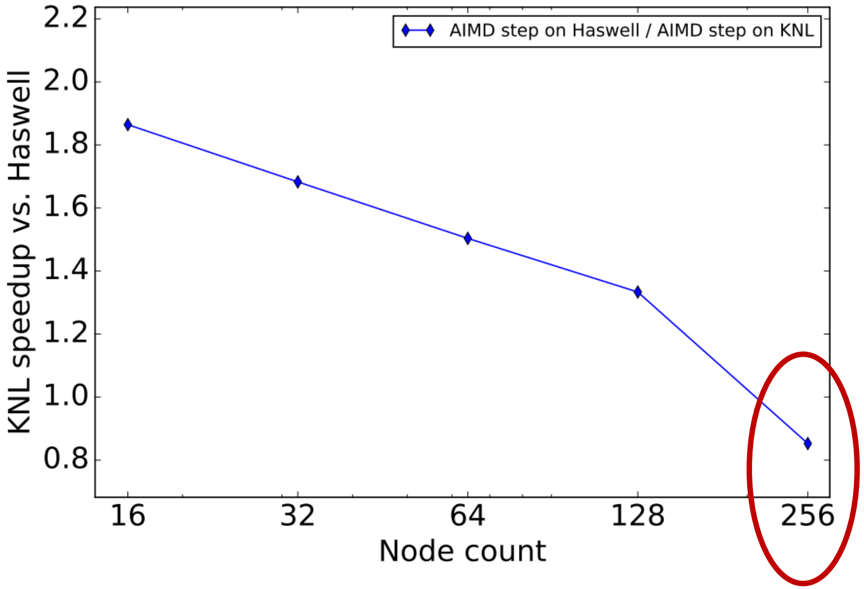
"Knights Landing"

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. System configuration: Cray* XC40 system, 2S Intel® Xeon® E5-2698v3 processor, Intel® Hyper-Threading technology disabled, 128 GB of DDR4 (8x 16 GB, 2133 MHz), Cray* Aries interconnect with Dragonfly topology; Cray* XC40 system Intel® Xeon Phi™ 7250 processors, 96 GB of DDR4 (6x 16GB, 2400 MHz), quadrant cluster mode, MCDRAM in cache mode, Cray* Aries interconnect with Dragonfly topology.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Relative Performance – HSW vs KNL

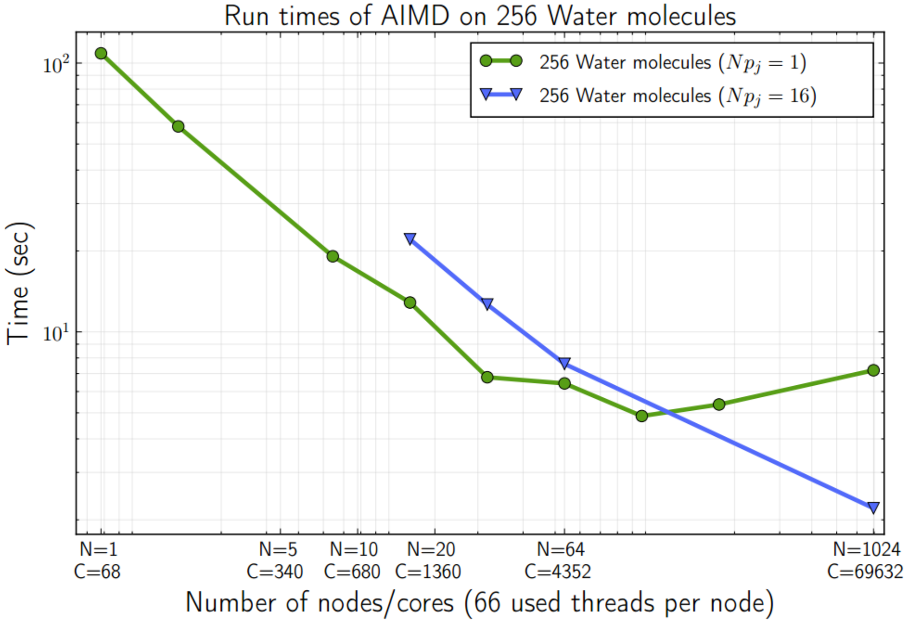
- Strong scaling regime
- Interconnect latency becomes visible
- Less occupancy of the network
- KNL seems to suffer from this more than HSW does



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

Performance – Effect of the Processor Grid

- Processor grid is a tradeoff
- 2D processor grid:
  $N_p = N_{pi} * N_{pj}$
- Large N_{pj} favors FFTs and non-local pseudopotentials
- Lagrange multiplier suffers from large N_{pi}
- Balancing N_{pi} and N_{pj} is required
  - problem size
  - number of ranks



Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

SUMMARY

Summary –

- Much of Knights Landing's throughput comes from parallelism:
  - Codes will need to be modernized to fully exploit the features of the chip
  - Usually: thread-parallel *and* SIMD-parallel execution key to performance

- Optimizations for Knights Landing usually also pay off on Xeon processors

- Plain library approaches are not good enough at times due to special requirements of application kernels

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. * Other names and brands may be claimed as the property of others. All products, dates, and figures are preliminary and are subject to change without any notice. Copyright © 2017, Intel Corporation.

experience
what’s inside™