Case Study:
Fusion PIC Code XGC1 on Cori KNL

Charlene Yang

Application Performance Group
NERSC
XGC1: Particle-In-Cell Simulation

PI: CS Chang (PPPL) | ECP: High-Fidelity Whole Device Modeling of Magnetically Confined Fusion Plasma
Gather Fields from Mesh to Ions

Solve Fields on Mesh

Deposit Charge From Particles to Mesh

Collision Operator

Ion Push

Electron Push Sub-Cycling
push electrons without updating fields or collisions – only field gather and push

~50x

*Computation
*Mapping
Unoptimized XGC1 Timings on 1024 Cori KNL nodes in Quad-Flat mode
XGC1: Code Profile on Roofline

- Self Elapsed Time: 5.322 s
- Total Time: 89.318 s

- Force
- Interpolate
- Search
1. **Search** for nearest 3 mesh nodes to the particle position (for multi-mesh refinement)

2. **Interpolate** fields from 3 mesh points to particle position

3. **Calculate force** on particle from fields

4. **Push** particle for time step $dt$
ToyPush: Baseline Profile - Timings

Unoptimized ToyPush Timings on Cori KNL in Quad-Cache mode
Roofline Model shows information timings alone cannot show
- Kernel Force Calculation: close to vector peak performance
- Kernel Interpolate and Search: less than scalar peak performance

- Data collected with Intel Advisor and analyzed with pyAdvisor
- Single thread rooflines on Cori KNL
- Should focus optimization efforts on Interpolate and Search kernels
Interpolation: Vectorization – L1 Blocking

- **veclength optimizations**
  - Baseline: $2^{9}$
  - Optimized: $2^{6}$

<table>
<thead>
<tr>
<th>Function / Call Stack</th>
<th>Clockticks</th>
<th>Instructions Retired</th>
<th>L1 Hit Rate</th>
<th>L2 Hit Rate</th>
<th>L2 Hit Bound</th>
<th>L2 Miss Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>e_interpol_tri</td>
<td>105,271,600,000</td>
<td>64,954,400,000</td>
<td>80.8%</td>
<td>94.4%</td>
<td>36.7%</td>
<td>29.5%</td>
</tr>
<tr>
<td>eom_eval</td>
<td>73,858,400,000</td>
<td>65,283,400,000</td>
<td>67.3%</td>
<td>99.9%</td>
<td>100.0%</td>
<td>0.8%</td>
</tr>
<tr>
<td>b_interpol_analytic</td>
<td>60,141,200,000</td>
<td>23,109,800,000</td>
<td>90.3%</td>
<td>100.0%</td>
<td>4.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>_intel_mic_avx512f_memset</td>
<td>35,288,400,000</td>
<td>3,441,200,000</td>
<td>42.1%</td>
<td>100.0%</td>
<td>0.8%</td>
<td>0.0%</td>
</tr>
<tr>
<td>rk4_push</td>
<td>20,528,200,000</td>
<td>14,898,800,000</td>
<td>31.9%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

- Low L1 Hit Rate, L2 Hit Bound
- High L1 Hit Rate

~1.5x improvement (MCDRAM Flat); ~2x improvement (DDR Flat)
Interpolation: Vectorization - Memory Access

Problems:

- Field data is stored on grid nodes, particles access nearest 3 grid nodes **indirectly** via triangle index.
  
efield(j, tri(i, itri(iv)))

- Interpolation loop is vectorized but **inefficiently** because of gather loads

Intel Compiler Vectorization Report

LOOP BEGIN at interpolate_aos.F90(67,48)
  reference itri(iv) has unaligned access
  reference y(iv,1) has unaligned access
  reference y(iv,3) has unaligned access
  reference evec(iv,icomp) has unaligned access
  reference evec(iv,icomp) has unaligned access

.....

irregularly indexed load was generated for the variable <grid_mapping_(1,3,itri(iv))>, 64-bit indexed, part of index is read from memory

.....

LOOP WAS VECTORIZED
unmasked unaligned unit stride loads: 6
unmasked unaligned unit stride stores: 3
unmasked indexed (or gather) loads: 18

.....

18 Gathers per loop iteration
(3 nodes x 3 components x 2)
Interpolation: Vectorization - Memory Access

Optimizations:

• Group particles that access the same triangle together, access grid nodes directly with a scalar index

• Single mesh: Trivial

• Multiple mesh: Feasible for number of particles >> number of grid nodes

• Align arrays during compile time.

Intel Compiler Vectorization Report

LOOP BEGIN at interpolate_aos.F90(72,51)
reference y(iv,1) has aligned access
reference y(iv,3) has aligned access
reference evec(iv, icomp) has aligned access

.....
SIMD LOOP WAS VECTORIZED
.....
unmasked aligned unit stride loads: 5
unmasked aligned unit stride stores: 3

~1.6x improvement
Problem:

- Initialization of large arrays with `avx512_memset` at every time step before entering vector loop becomes memory bandwidth bound.

Optimizations:

- Initialize array inside the vector loop (if you can)
- Use threads for initialization

Intel Compiler Vectorization Report

```
LOOP BEGIN at interpolate_aos.F90(57,5)
  memset generated
  loop was not vectorized:
  loop was transformed to memset or memcpy
LOOP END
```

~5% improvement
Higher if no. of particle increases
Interpolation: Optimization Path on Roofline

- Kernel moved to compute bound regime
- AI increased due to memory access pattern change
- Peak compute performance is nearly reached

Baseline Case (w/ Indirect access)

Replace Indirect Access with Scalar Access

Optimize Vector Length

Access Grid Data in Scalar Chunks
Search: Vectorization – ‘cycle’ + SIMD

Problems:
• Multiple exits due to ‘cycle’ statement prevents vectorization
• Assumed read after write (RAW) dependency prevent vectorization

Optimization:
• Replace exit condition with a logical mask
• Vectorize with omp simd directive, declare private arrays simd private

Intel Compiler Vectorization Report

LOOP BEGIN at search.F90(62,8)
loop was not vectorized: loop with multiple exits cannot be vectorized unless it meets search loop idiom criteria

LOOP BEGIN at search.F90(66,8)

reference y(iv,1) has aligned access
reference y(iv,3) has aligned access
reference id(iv) has aligned access
reference continue_search(iv) has aligned access
data layout of a private variable bc_coords was optimized, converted to SoA
OpenMP SIMD LOOP WAS VECTORIZED
unmasked aligned unit stride loads: 4
unmasked aligned unit stride stores: 1

1.5x improvement
Forcing SIMD vectorization doesn’t work initially due to multiple exits. Once exits are eliminated, code vectorizes.
ToyPush: Optimized Performance

- **Force Kernel**: still good performance, close to vector add peak
- **Interpolate Kernel**: 10x speedup, closer to vector FMA peak
- **Search Kernel**: 3x speedup, closer to L2 bandwidth roof

Marker size $\sim$ CPU time

- Code is available at https://github.com/tkoskela/toypush
XGC1: Optimization Speedups (WIP)

XGC1 Timings on 1024 Cori KNL nodes in Quad-Flat mode
XGC1: Optimized Performance on Roofline

Performance (GFLOPS)

Interpolate

Force

Search

Self Elapsed Time: **0.120** s  Total Time: **0.660** s

Arithmetic Intensity (FLOP/Byte)
XGC1: Code Profile on Roofline

Performance (GFLOPS)

Self Elapsed Time: 5.322 s  Total Time: 89.318 s

Arithmetic Intensity (FLOP/Byte)
Summary

• **XGC1 -> ToyPush -> XGC1**

• **Roofline Model can help**
  – Identify performance bottlenecks (compute, bandwidth, latency, *etc*).
  – Prioritize optimization efforts (routines, vectorization, memory access, *etc*).
  – Tell when to stop (realistic achievable performance, distance to roofs).

• **Intel Advisor can take care of the rest!**
  – Integrated compiler reports, static binary analysis (instruction set, data types, *etc*.) and dynamic analysis (CPU sampling).
  – FLOPS/trip counts, vectorization efficiency, dependency and memory access pattern.
  – Roofline charts of various flavors 😊
    • Original DRAM-based Roofline (DRAM <-> Core).
    • Cache-aware Roofline (L1 <-> Core) ✔.
    • Cache-simulator based Roofline (L1, L2, LLC, MCDRAM and DRAM <-> Core) ✔.