HPGMG on the Knights Landing Processor

Samuel Williams

Lawrence Berkeley National Laboratory (LBNL)

SWWilliams@lbl.gov

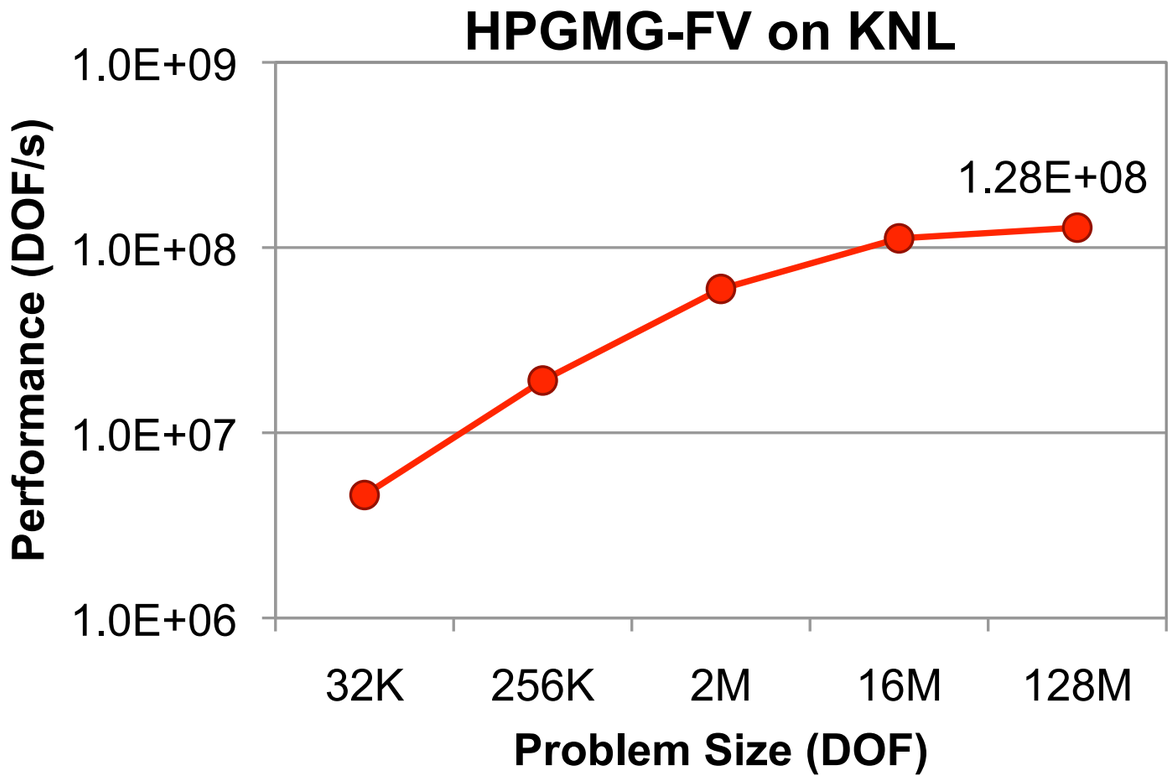
KNL Testbed

- Single node, self-hosted KNL whitebox
- 68 cores, 1.4GHz, dual 512b vector units
- hierarchical memory architecture (16GB HBM + 96GB DDR)
- “quadflat” configuration of the memory hierarchy and cache directory
- single-node experiments

Baseline KNL Performance

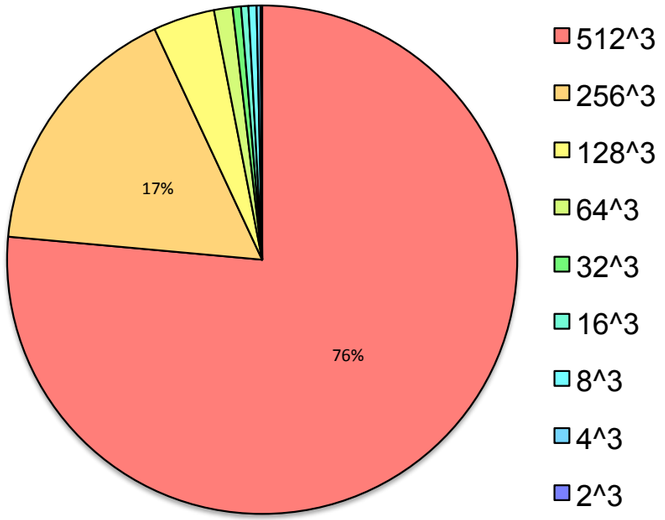
- HPGMG-FV from repo without modifications or tuning
  - single process, 64 OpenMP threads
  - **“numactl –m 1”** to bind memory allocation to the HBM numa node
- Good performance on moderately large problems (256^3 and 512^3) that fit in 16GB of HBM
- Limited dynamic range

  (hard to get solve under 10ms)

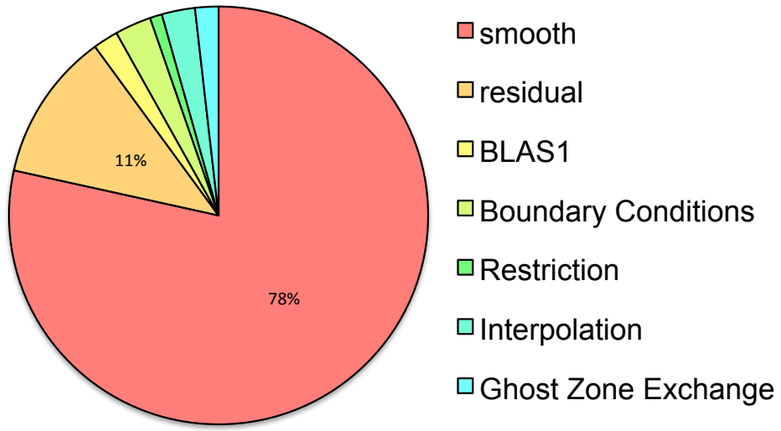


Baseline Breakdown and Analysis

- Single process runs so…
  - no MPI for single process runs
  - bulk of run time is on level 0
  - bulk of the run time is in the smoother and residual stencils

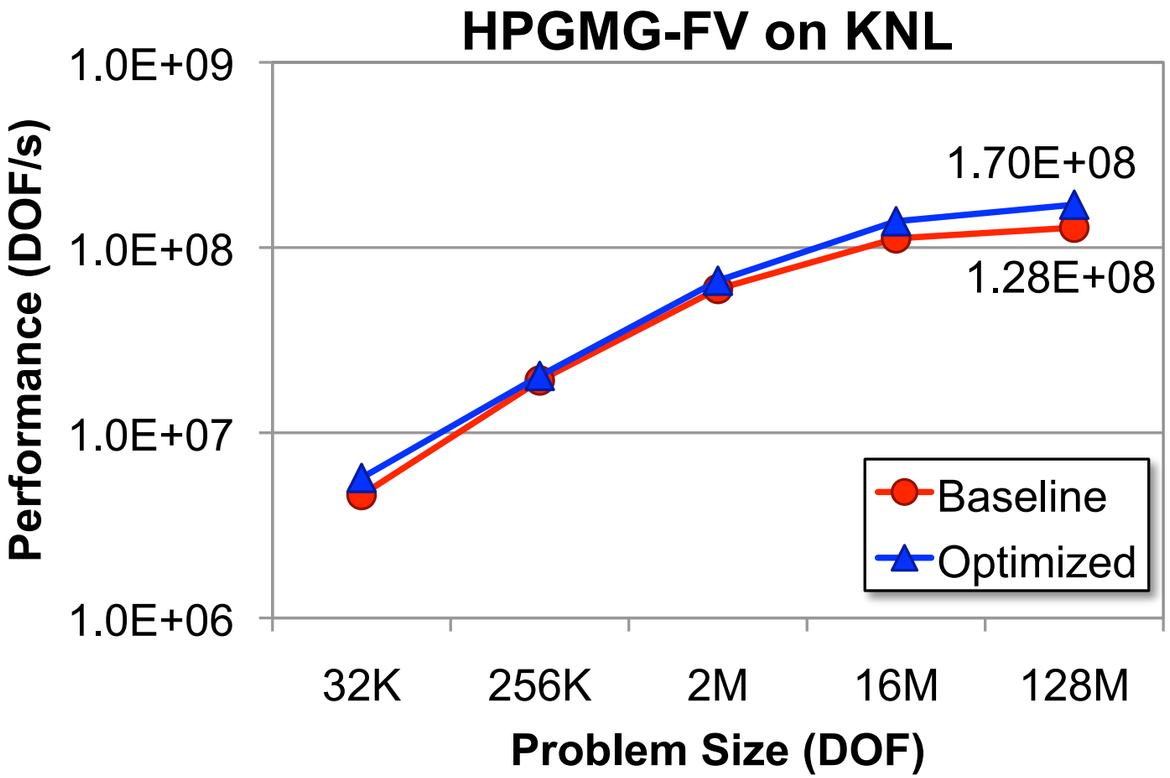


- However, smoother performance was less than ideal (Roofline)
  - **172GB/s of "effective" bandwidth**
  - STREAM bandwidth exceeds 400GB/s in quadflat



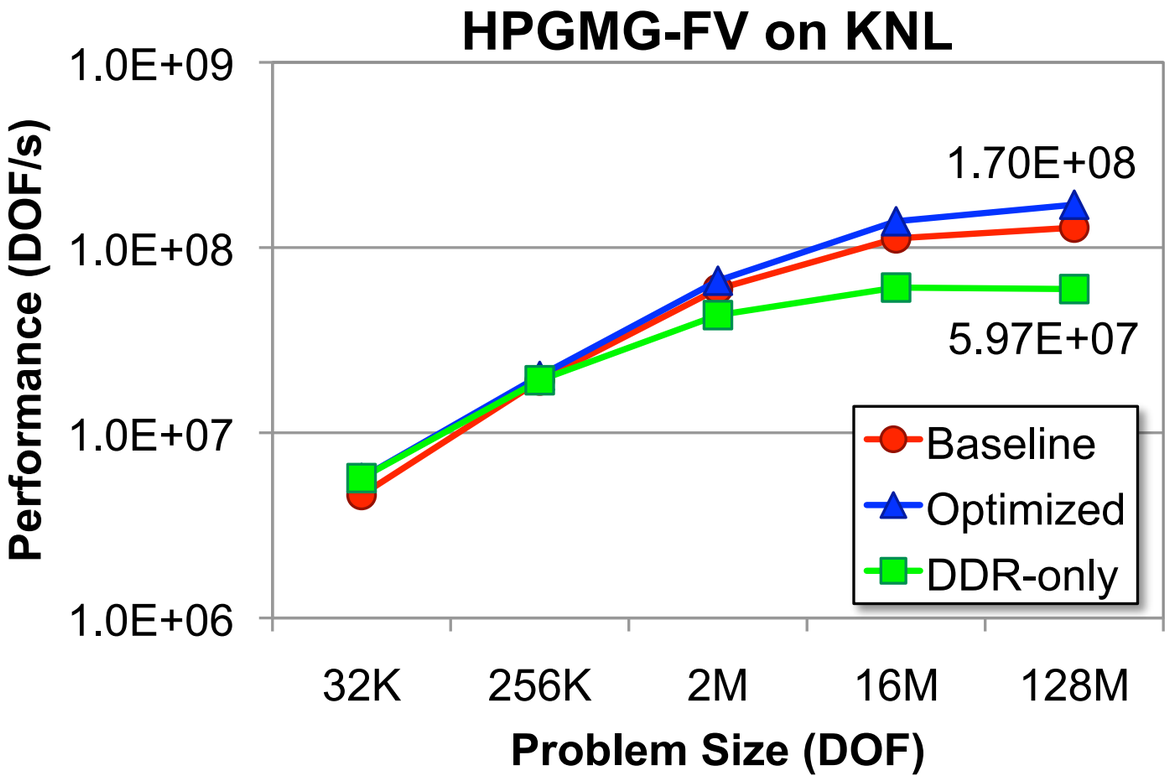
Optimized KNL Performance

- Use flux-optimized variants …
  - operators.flux.c is already in the repo
  - fissions/reorganizes loops around flux calculations
  - eliminates redundant calculations
  - leverages omp4 simd clauses
  - uses ICC-specific clauses for alignment
  - hand-crafted array padding to avoid L2 conflicts.
- **Improved smoother by ~40%**
- No change in dynamic range



Benefit of KNL's HBM over DDR

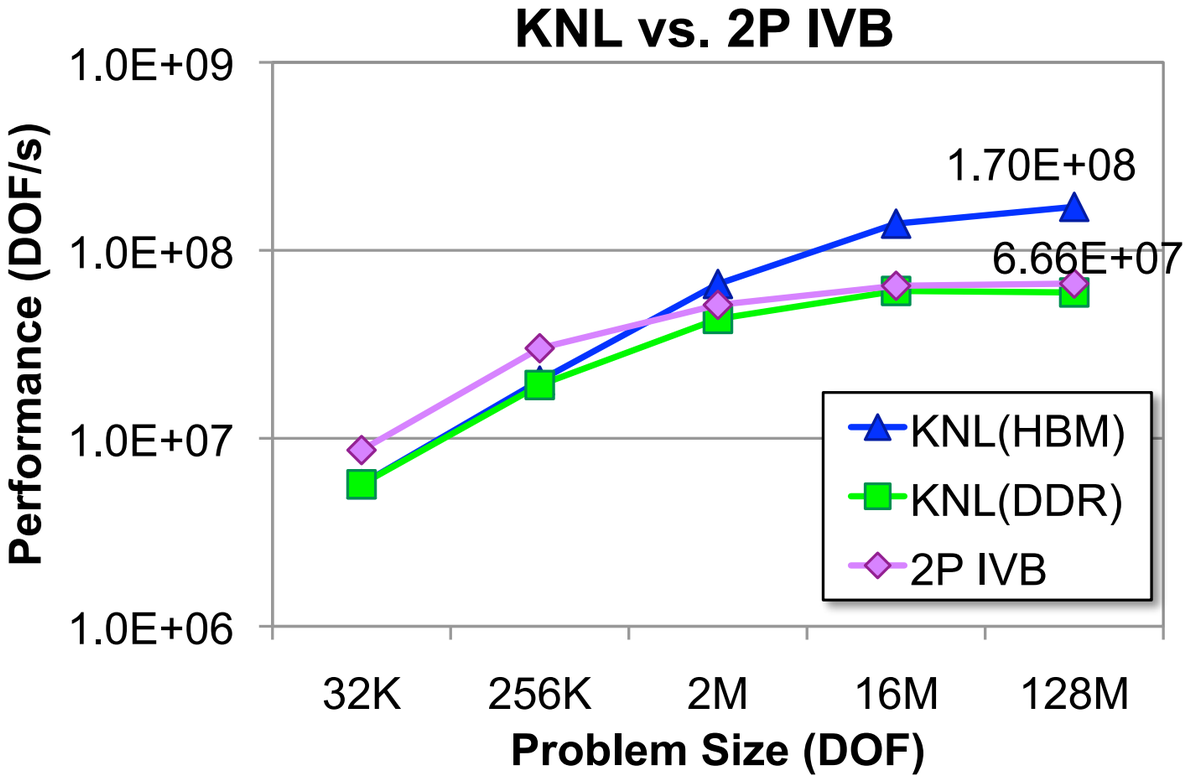
- numactl –m 0 to pin memory allocation to DDR to quantify the benefit of HBM.
- For moderately large problems (<16GB), **HBM provides a nearly 3x boost in performance**
- For small problems, HBM has no benefit over DDR
  - high memory latencies
  - high OpenMP overheads
  - lack of parallelism (compared to HW)



KNL Performance vs. Ivy Bridge

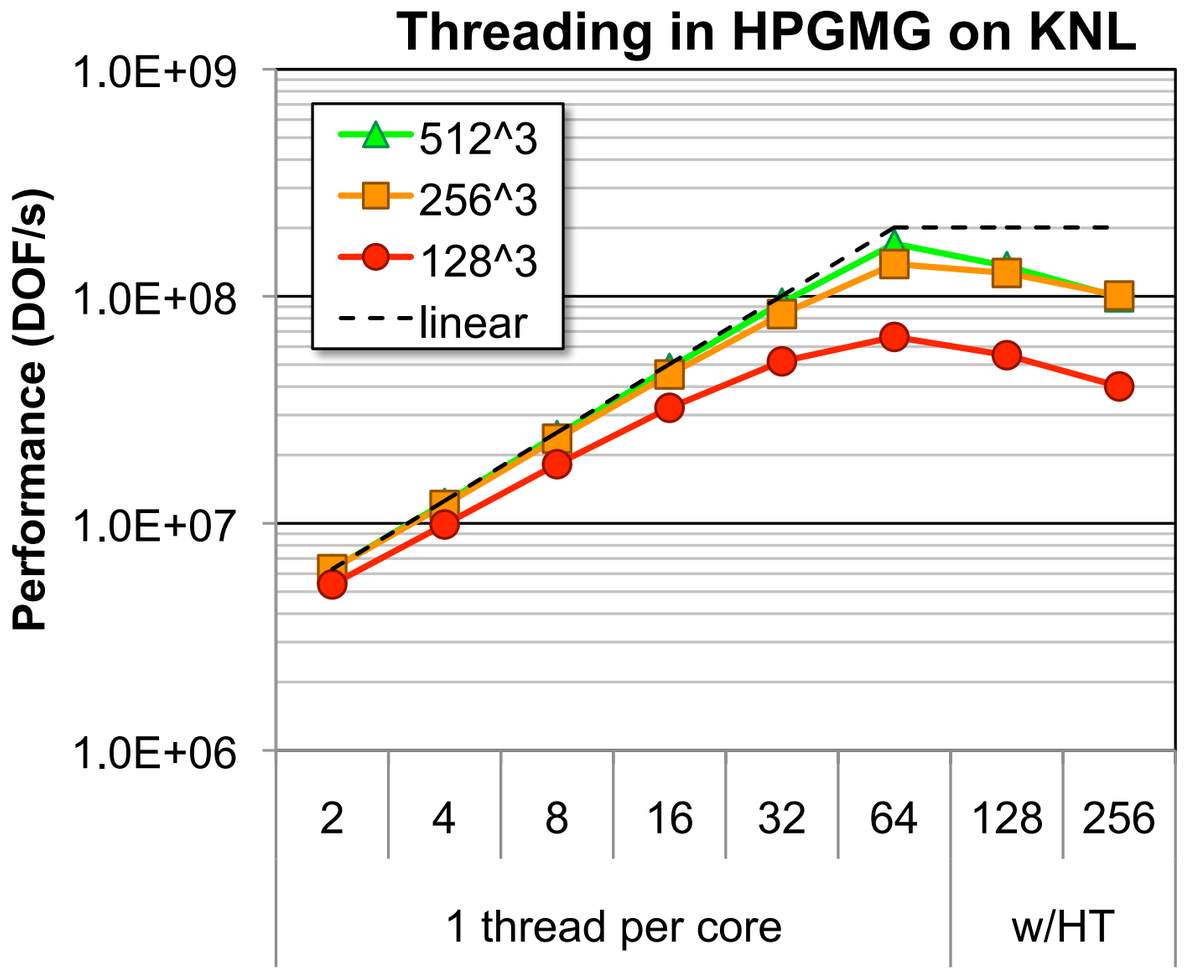
- For large problems,
  - **If using HBM, Knights Landing is 2.5x faster than a 2P Ivy Bridge node (2x12)**
  - If constrained to only use DDR, KNL is no faster than an Ivy Bridge (no surprise if bandwidth-limited)
- **Ivy Bridge is moderately faster than KNL for small problems.**

  (big concern when strong scaling)



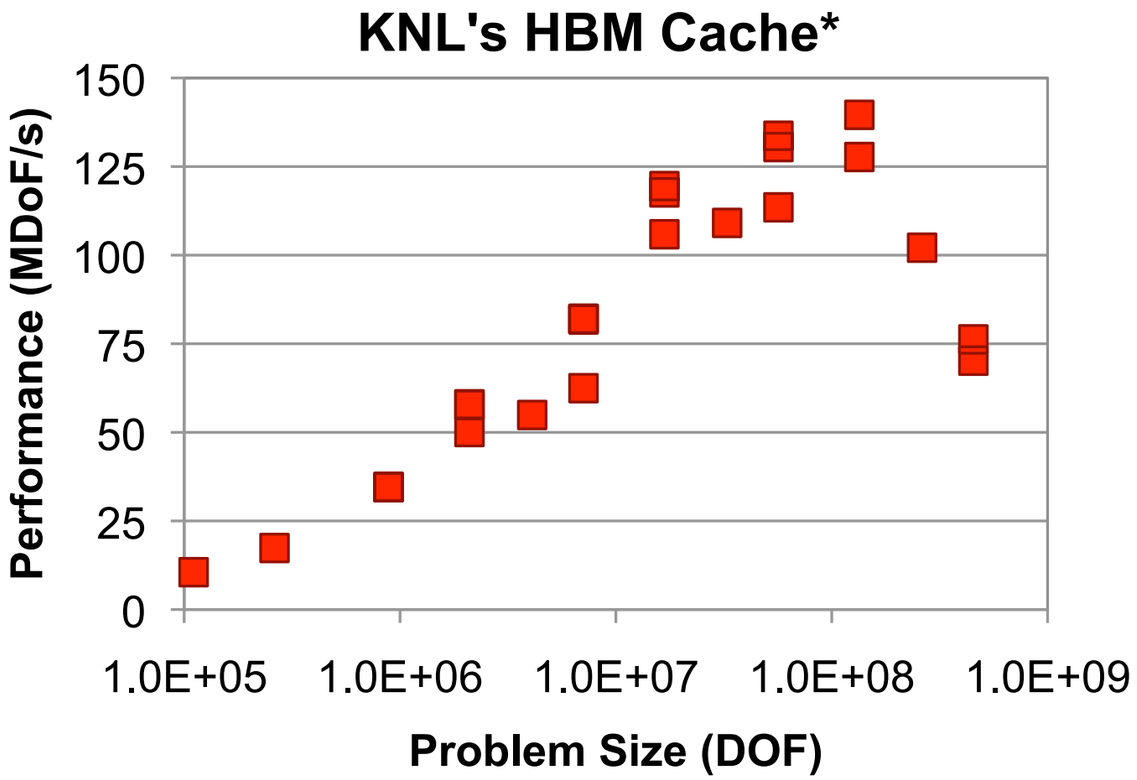
OpenMP Scalability on KNL

- Generally speaking, performance scales well with threads for moderately large problems.
- **HyperThreading generally hurts performance**
  - benefit on ghost exchanges/boundary conditions dominated by irregular/strided access
  - substantial penalty on cache-sensitive smoothers



KNL Performance using HBM as a Cache

- KNL can configure its HBM as a large, HW-managed L3 cache
- allows seamless execution of a wide range of problem sizes without the productivity of a hierarchal memory
- We observe a strong cache effect when the problem size exceeds HBM capacity and is streamed from DDR.



*Data was collected on a 64c preproduction node. quadcache STREAM bandwidth is ~90% of quadflat. 68c production nodes are 5-10% faster than this system.

November 2016 Ranking

HPGMG Rank	System Site	System Name	10^9 DOF/s	MPI	OMP	Acc	DOF per Process	Top500 Rank	Notes
1	ALCF	Mira	500	49152	64	0	36M	6	
2	HLRS	Hazel Hen	495	15408	12	0	192M	9	
3	OLCF	Titan	440	16384	4	1	32M	3	K20x GPU
4	KAUST	Shaheen II	326	12288	16	0	144M	10	
5	NERSC	Edison	296	10648	12	0	128M	49	
6	CSCS	Piz Daint	153	4096	8	1	32M	8	K20x GPU
7	Tohoku University	SX-ACE	73.8	4096	1	0	128M	-	vector
8	LRZ	SuperMUC	72.5	4096	8	0	54M	27	
9	NREL	Peregrine	10.0	1024	12	0	16M	-	
10	NREL	Peregrine	5.29	512	12	0	16M	-	
11	HLRS	SX-ACE	3.24	256	1	0	32M	-	vector
12	NERSC	Babbage	0.762	256	45	0	8M	-	KNC
13	*NERSC*	*KNL white box*	*0.170*	*1*	*64*	*0*	*128M*	*-*	*KNL*

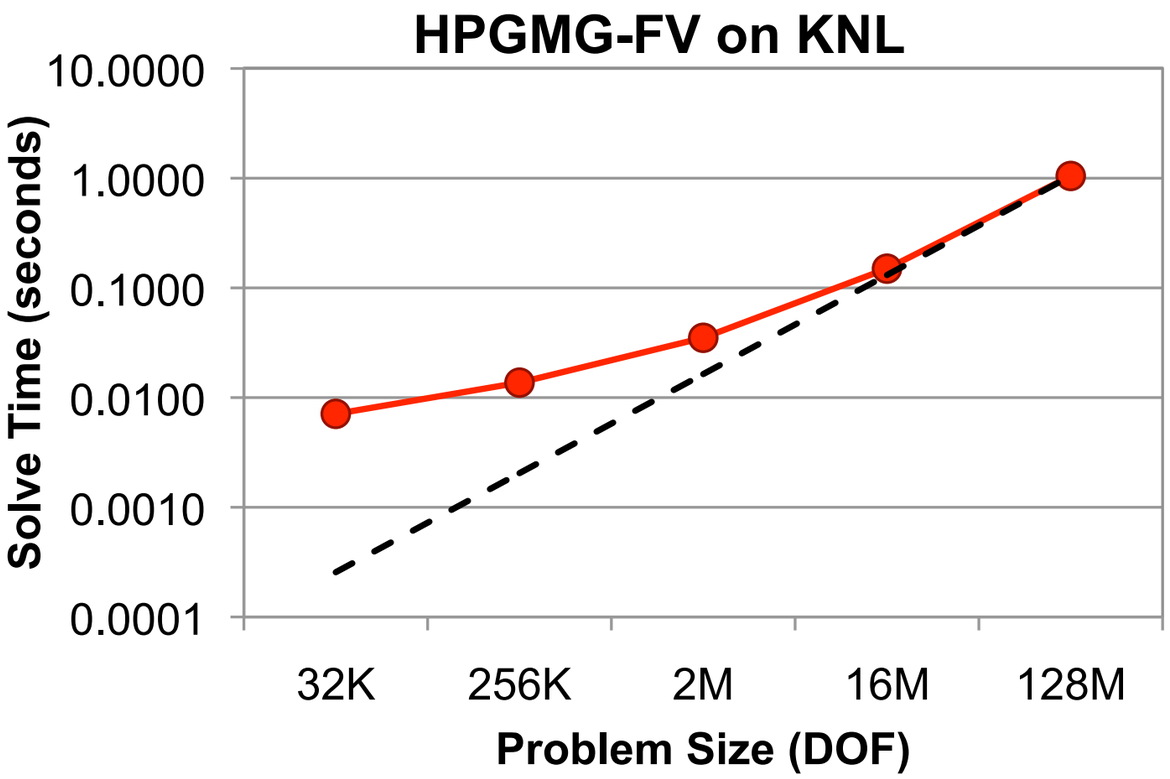
KNL Summary

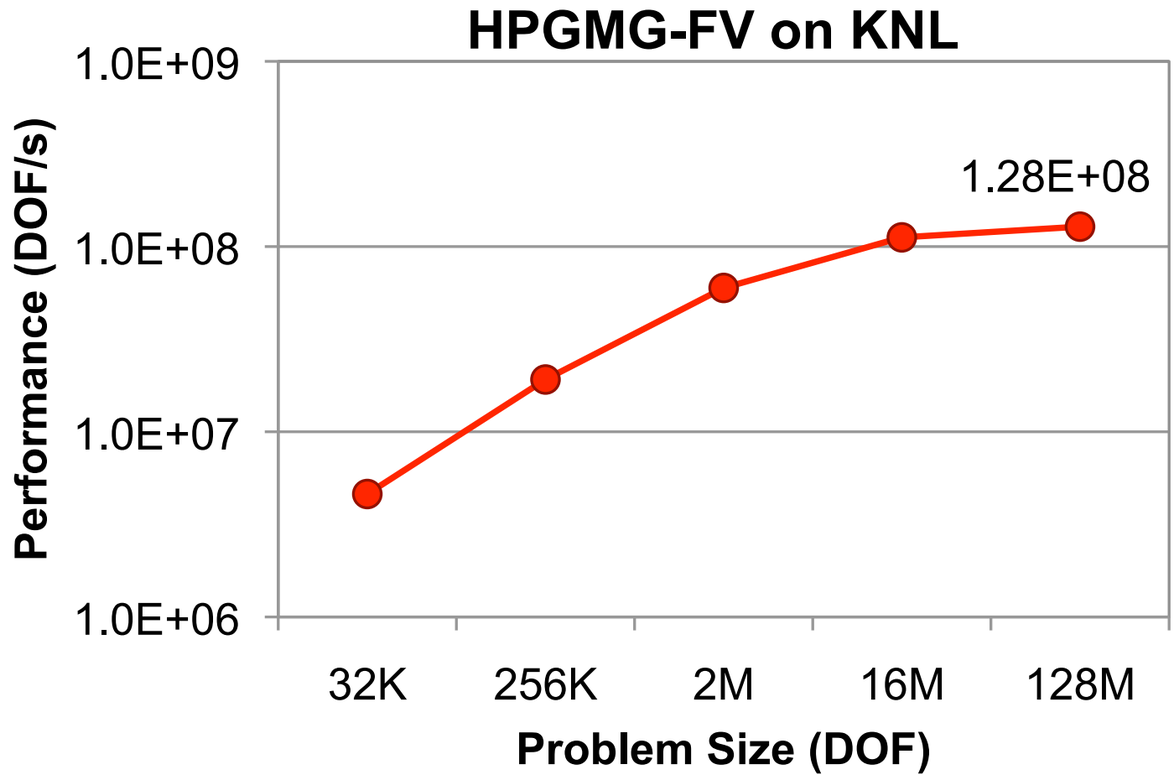
- **KNL is 2.5x faster than a 2P IVB**
  (required some loop fission and omp4 simd)
- The HBM cache worked well (and was productive)

- KNL was unable to exploit its full HBM bandwidth (in 4^{th} order)
- KNL was unable to keep capacity misses to a minimum
- KNL was even more sensitive to dynamic range and generally required huge problems for best performance.
- *looking forward to evaluating HPGMG on Cori…*

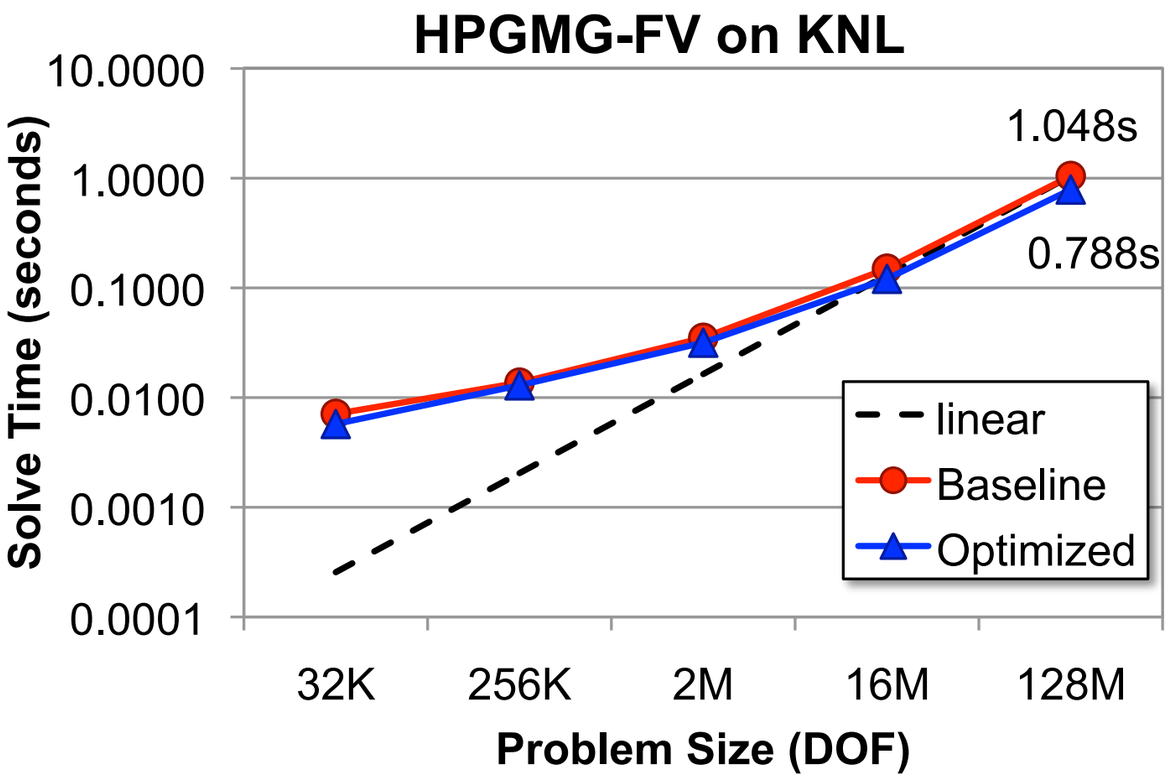
Backup Slides

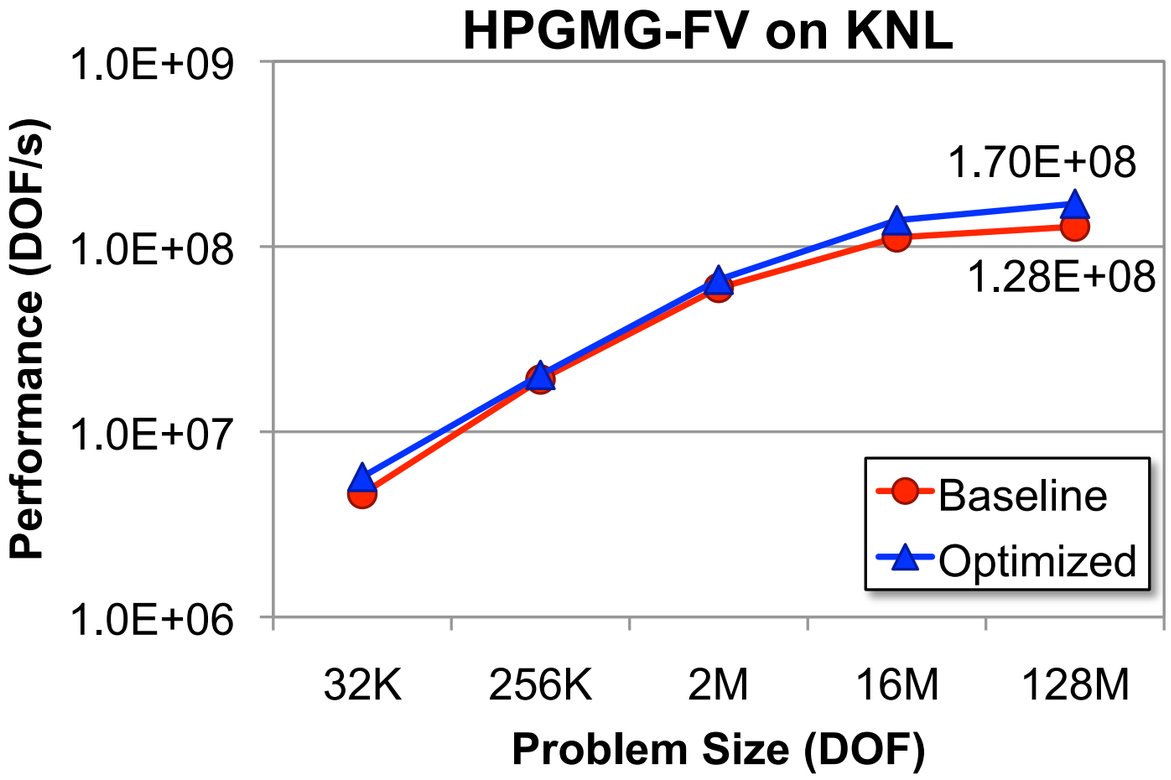
Baseline KNL Performance



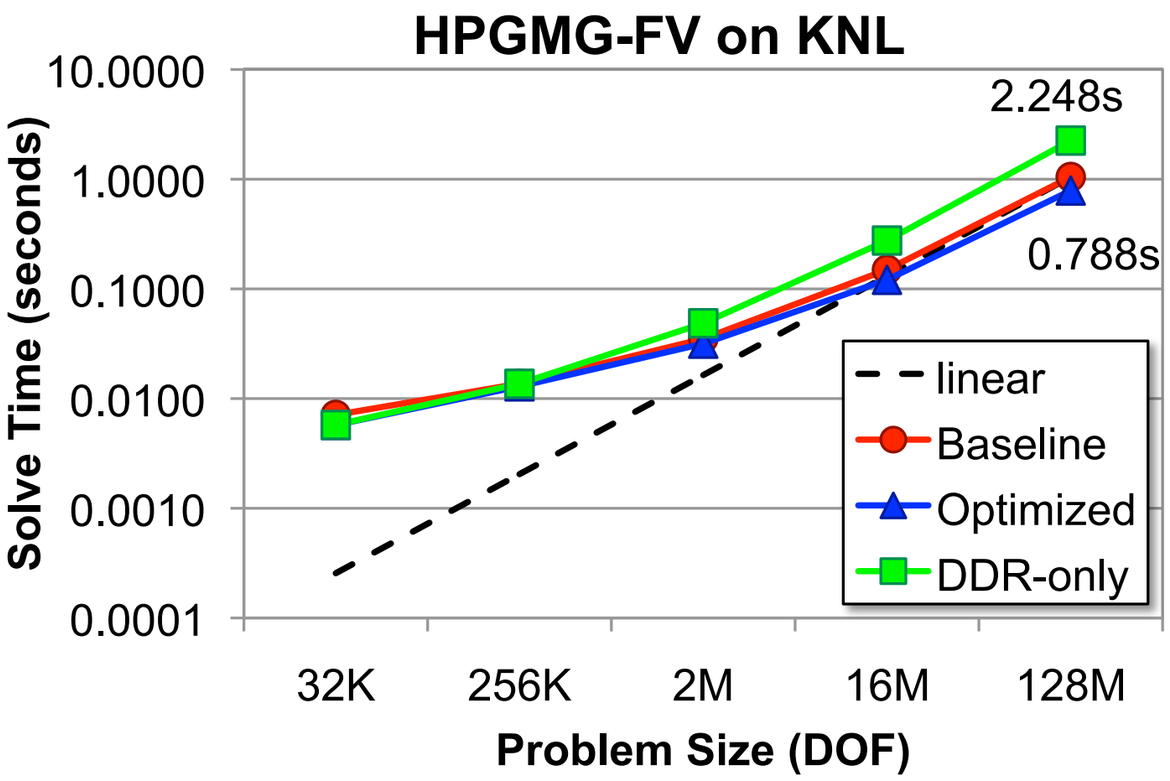


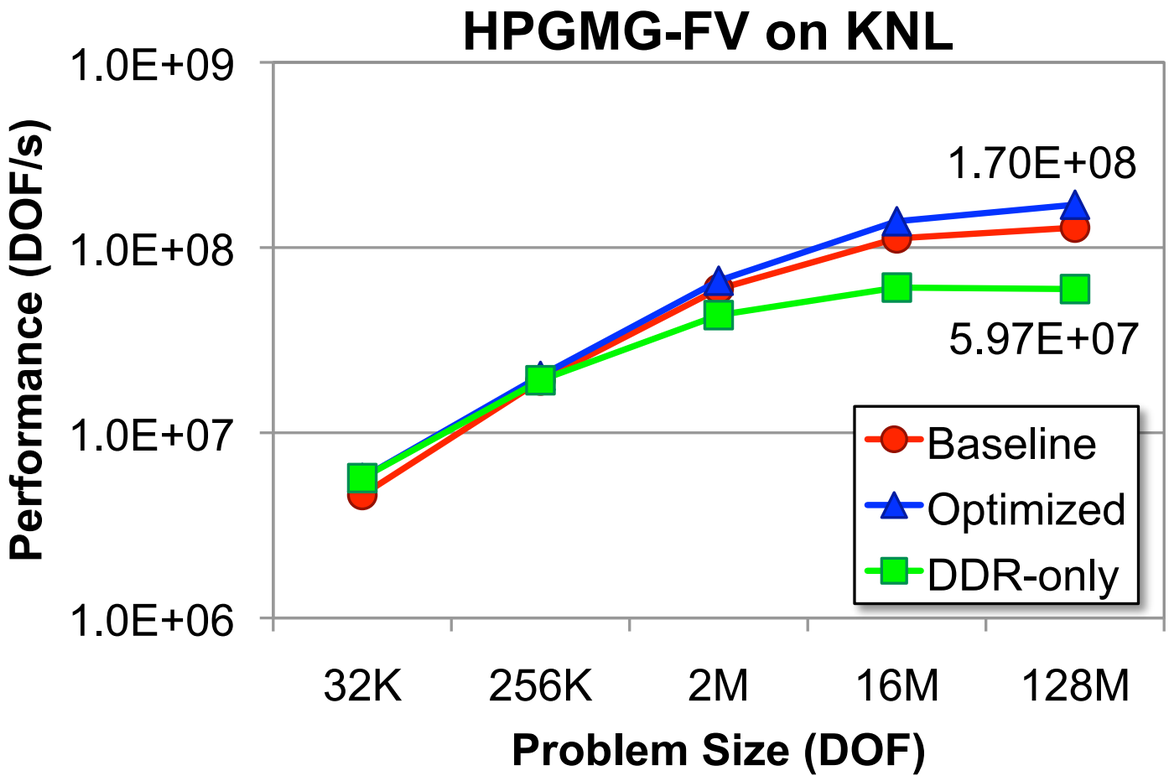
Optimized KNL Performance





Benefit of KNL's HBM over DDR





KNL Performance vs. Ivy Bridge