DWF HMC: MSPCG, EOFA, (RM,LA)HMC, etc

Tuning/optimizing for summit

Kate Clark, Bob Mawhinney, Chulwoo Jung (presenter), David Murphy, Jiqun Tu .....

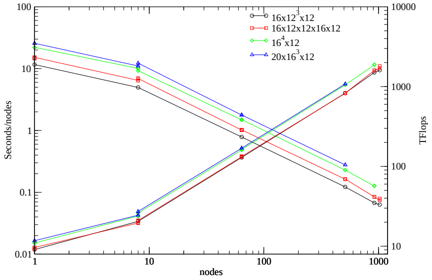
April 27, 2019

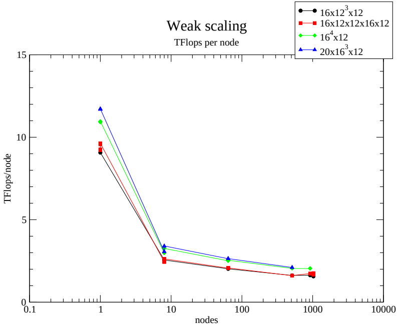
Summit preparation: INCITE 2019

850K node hours, shared with MILC
2+1 flavor physical , $96^3 \times 192 \times 12$, a^{-1} ~2.77Gev, $L \sim 6.8$fm
$16 \times 12^3 \times 12$ on $(1 \times 8 \times 8 \times 16 = 1024)$ ndoes $\times$ 6 GPUs
Strategy: CPS + QUDA inverters

- Tuning with (currently 6) Hasenbusch masses, Force Gradient...
- Started from a thermalized $32^3 \times 64$ lattice duplicated in all directions
- QUDA inverter (CG+Multimass) interface (re)checked against CPS/BFM. Exact One Flavor Algorithm(EOFA) added.
- Previously only had interface to asymmetric preconditioner. Symmetric added for Multisplitting-preconditioned CG (MSPCG: arXiv:1804.08593).

QUDA Mobius on Summit at the INCITE submission

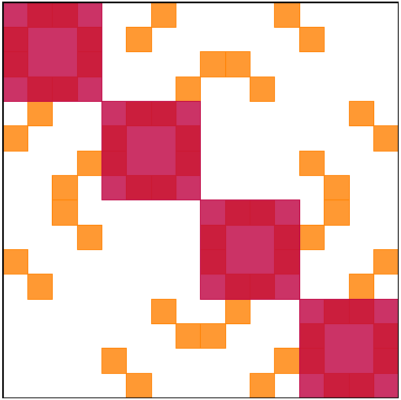




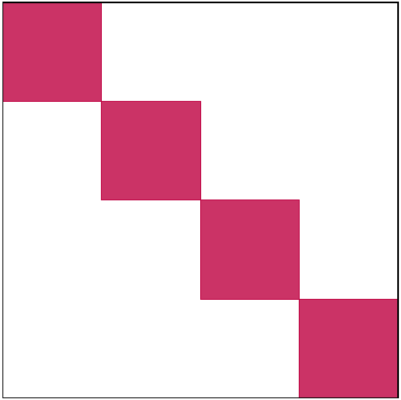
Performance limited by network bandwidth.
Not-so-great weak scaling. less than 2TF per node on 1024 nodes, compared to >10TF for single node with same local volume.

Multisplitting-preconditiond Conjugate Gradient(MSPCG)

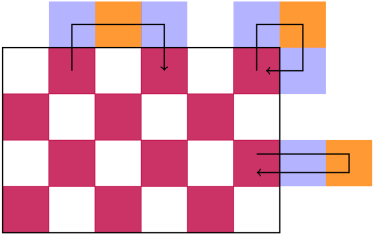
Additive Schwarz 'done right' for Mobius CGNE ($D^\dagger D\phi = D^\dagger\chi$).
Fixed iteration preconditioner per outer iteration.



$$A$$

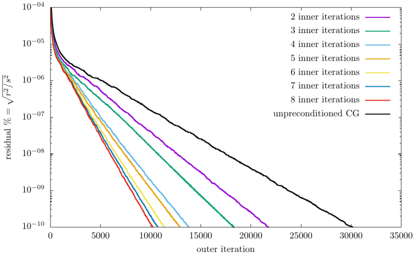


$$P = \bigoplus_s A_s$$



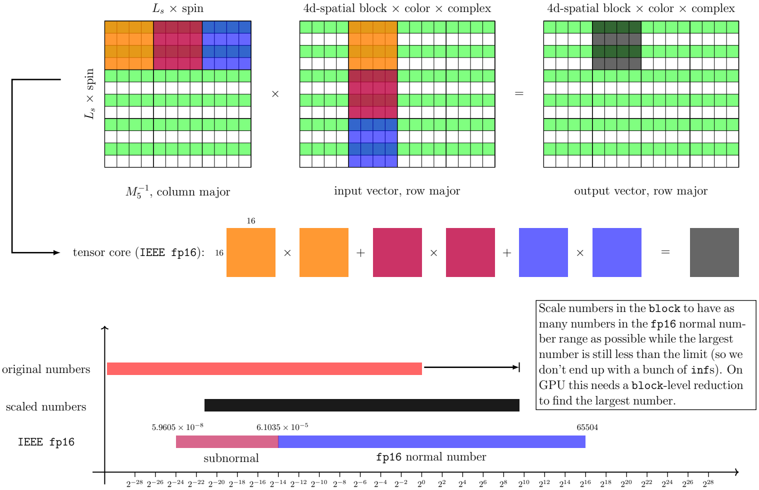
from J. Tu, arXiv:1811.08488

The preconditioner inversion P^{-1} does not need to be solved to arbitrarily high precision for the algorithm to work.



Up to factor of 3 reduction in outer iteration

Using Tensor core



From J. Tu

Scale numbers to be in the range for fp16, avoid overflow and too much underflow

- QUDA Dslash rewrite (K. Clark,..),
- 5D part fused to achieve better overlap with communication (J. Tu)
- Network itself seems to have improved, albeit still without GDR.

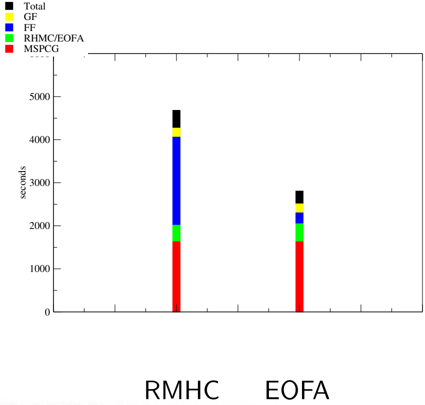
Time in CG	$< 10\%$	25-35%	$> 60\%$	
$64^3 \times 128 \times 12$	Double	Half	Precon	min/traj.
$(4\times)4^2 \times 8 = 128$	160	570	3230	86
$(4\times)4^2 \times 16 = 256$	260	861	6230	53
$(4\times)4 \times 8 \times 16 = 512$	360	1165	11630	36
$96^3 \times 192 \times 12$				
$(6\times)4^2 \times 16 = 256$	420	1340	9400	
$(6\times)4 \times 8 = 512$	770	2300	18810	79
$(6\times)8^2 \times 16 = 1024$	1140	3700	36300	47

Table: Aggregate QUDA Mobius performance on summit, in TFLOPS/s

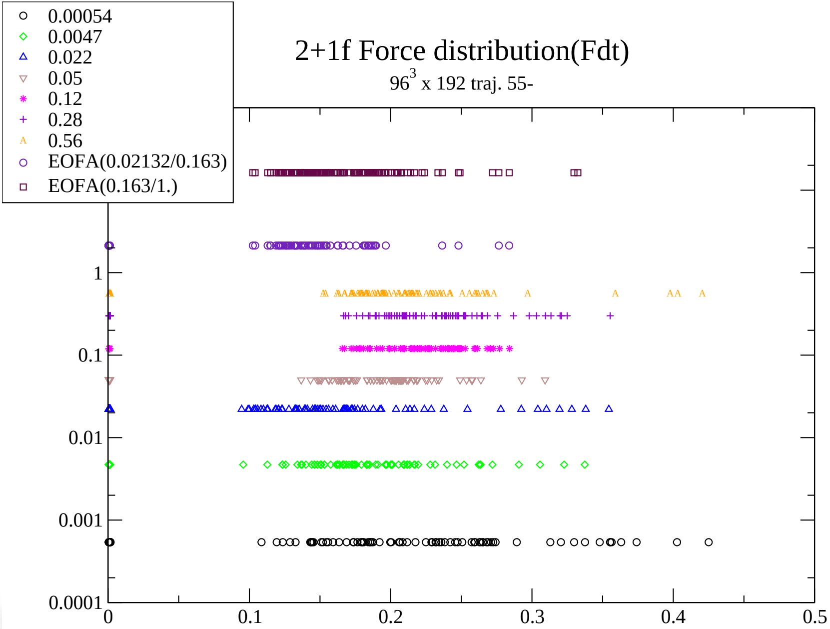
We could use faster tensor core!
Currently MSPCG is giving $20 - 30\%$ speedup

Exact One Flavor Algorithm(EOFA)

Timing breakdown of 96I evolution on OLCF Summit(1024 nodes)



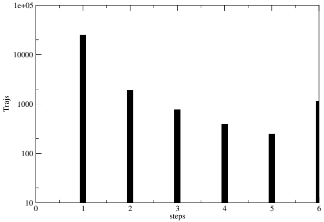
While Multimass solver does not take much time, fermion force term calculation becomes time consuming. Arithmetic intensity low (no smearing).
Exact One Flavor Algorithm(Chiu et. al.) with the improved pre-conditioner (D. Murphy et al. arxiv:1706.05843) allows effective use of mixed precision solvers, efficient mass preconditioning, and reduce the number of pseudofermions significantly. Necessary routines implemented in QUDA by J. Tu in collaboration with D. Murphy.


2+1f Force distribution(Fdt)
96^3 x 192 traj. 55-
0.00054
0.0047
0.022
0.05
0.12
0.28
0.56
EOFA(0.02132/0.163)
EOFA(0.163/1.)
1
0.1
0.01
0.001
0.0001
0
0.1
0.2
0.3
0.4
0.5

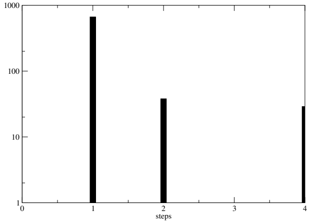
Further improvement?

- We need better hardware!! Our 1024-node run has been causing problems with PSU. Have been working with IBM. Most of the production runs confined to 512-node. Used only 140K node hours so far. 2nd stream??.
- GDR?
- Split Grid?: Current Hasenbusch tuning has highly unbalanced iteration counts. Combined with (relatively) good strong scaling, it doesn't look promising with the current tuning. A different tuning with a lot more intermediate masses, different preconditioning operators are being explored.
- (RM,LA)HMC?
  RMHMC(by Y. Jang) and LAHMC(CJ) implemented in CPS.
  Small volume (32^3) test ensembles with RMHMC($1/a \sim$3.1Gev) and Look-ahead($1/a \sim$2.8Gev) have been generated. So far it has not shown enough improvement over HMC to adopt for the current summit run. Study ongoing.

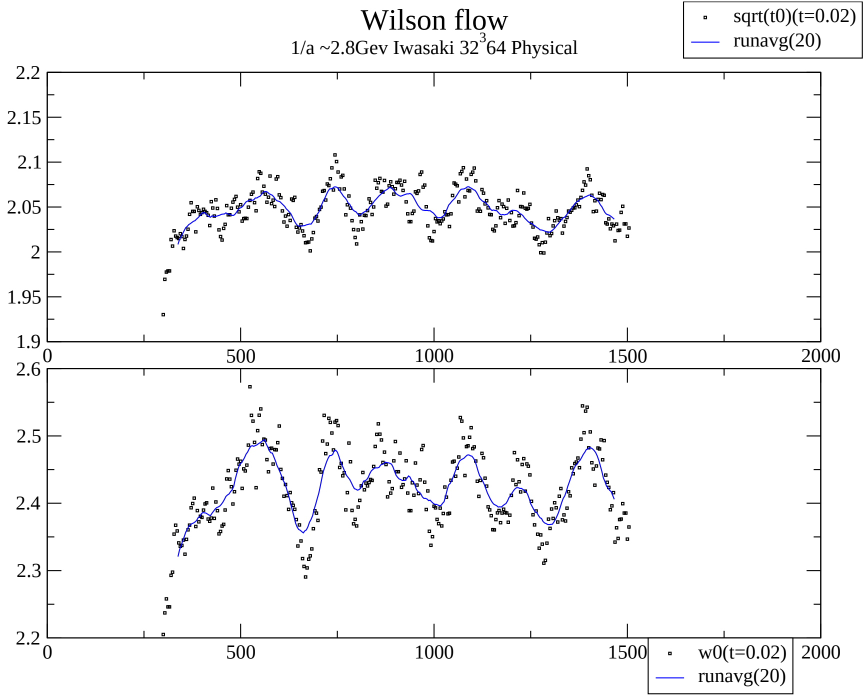
- Look-ahead HMC(LAHMC): Try going further without momentum refresh even after the initial MD was rejected, up to a predetermined number. Relaxes detailed balance, to require preserving N$\leftrightarrow$ N probability distribution instead of 2 $\leftrightarrow$ 2 in traditional HMC.
  So far, it seems harder to tune LAHMC to do multiple steps with fermions.



Wilson $\beta = 6.4,\ 32^4, 4MD \times 5$



DWF+I $\beta = 2.31,\ 32^3 \times 64, 2MD \times 3$



Wilson flow
1/a ~2.8Gev Iwasaki 32^3 64 Physical
sqrt(t0)(t=0.02)
runavg(20)
2.2
2.15
2.1
2.05
2
1.95
1.9
0
500
1000
1500
2000
2.6
2.5
2.4
2.3
2.2
0
500
1000
1500
2000
w0(t=0.02)
runavg(20)

Thank you!

Riemann Manifold Hybrid Monte Carlo (RMHMC)

Origin: Duane & Pendleton (Phys. Lett. B206, 101106 (1988))
In normal HMC, slow-varying (in spacetime) modes has a small force in MD time, resulting in a critical slowing down. Changing the hamiltonian can accelerate the slow modes. Accelerated 'mass term' should be gauge invariant $\rightarrow$ field-dependent metric.
RMHMC (Girolami & Calderhead, 2011)):
Typical integrator algorithms such as leapfrog is non-reversible for non-separable hamiltonian: implicit integrator to maintain reversibility

$$p^{n+\frac{1}{2}} = p^n - \frac{\epsilon}{2}\frac{\delta H}{\delta\theta}(\theta^n, p^{n+\frac{1}{2}})$$

$$\theta^{n+1} = \theta^n + \frac{\epsilon}{2}\left[\frac{\delta H}{\delta p}(\theta^n, p^{n+\frac{1}{2}}) + \frac{\delta H}{\delta p}(\theta^{n+1}, p^{n+\frac{1}{2}})\right]$$

$$p^{n+1} = p^{n+\frac{1}{2}} - \frac{\epsilon}{2}\frac{\delta H}{\delta\theta}(\theta^{n+1}, p^{n+\frac{1}{2}})$$