M3D-C1 ZOOM Meeting

5/10/2020

Agenda

1. Announcements
2. CS Issues
    1. Yang Liu May 8 email
    2. Autotune May 4 email Yang
    3. Review of preconditioner reuse (solve2) - Jin
    4. Review of GPU status-Jin
    5. Local adaption status-Seegyoung
    6. Status of new modules on portal -Jin
    7. NERSC Time
    8. other
3. Physics Studies
    1. 3D VDE Benchmark Update: New sideways force result
    2. 3D ITER VDE Force calculation status
    3. Fluid runaway electrons: stability and sources – Chen Zhao
    4. Summit proposal—Chang Liu
    5. Other

Announcements

- CTTS talk series Schedule
  - 4/29 Chang Liu M3D-C1-K (posted)
  - 5/6 NIMROD CS
  - 5/13 Sovinec/Jardin 3D VDE benchmark
  - 5/20 Eric Howell NTM with NIMROD
  - 5/27 Roman Samulyak SPI simulations
  - 6/3 Chen Zhao MHD stability with RE
  - 6/10 Cesar Clauser Mitigation in NSTX with EMPI
  - 6/17 Charlson Kim NIMROD SPI Simulations
  - 6/24 open
- Opening new web site today: m3dc1.pppl.gov
  - Will contain documentation and vgs from weekly meeting

From Yang Liu

6 cori knl node, 32	planes, 384 total MPIs, (15 nodes needed)		
NS=20	default: 32 jacobi-blocks 8 jacobi-blocks or less: out-of-memory 64 jacobi-blocks or more: hanging		
	PS=[ROWPERM, COLPERM, log2(nprows), NSUP, NREL]	OS=[petsc time]	
	['NOROWPERM', 'METIS_AT_PLUS_A', 0, 63, 17]	3.91E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 1, 151, 27]	2.87E+01	
	['LargeDiag_MC64', 'MMD_AT_PLUS_A', 2, 203, 12]	2.76E+01	
	['LargeDiag_MC64', 'METIS_AT_PLUS_A', 2, 234, 23]	2.62E+01	
	['LargeDiag_MC64', 'METIS_AT_PLUS_A', 2, 84, 38]	2.63E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 0, 299, 35]	3.42E+01	
	['LargeDiag_MC64', 'MMD_AT_PLUS_A', 1, 52, 32]	3.44E+01	
	['NOROWPERM', 'MMD_AT_PLUS_A', 1, 262, 20]	3.12E+01	
	['LargeDiag_MC64', 'MMD_AT_PLUS_A', 1, 262, 22]	3.15E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 65, 15]	2.88E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 65, 40]	2.74E+01	
	['LargeDiag_MC64', 'MMD_AT_PLUS_A', 2, 224, 38]	2.62E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 153, 13]	2.71E+01	
	['NOROWPERM', 'MMD_AT_PLUS_A', 2, 53, 13]	3.20E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 106, 27]	2.65E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 121, 30]	2.63E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 140, 36]	2.57E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 137, 40]	2.60E+01	
	['LargeDiag_MC64', 'METIS_AT_PLUS_A', 2, 141, 34]	2.60E+01	
	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 118, 34]	2.61E+01	
default:	['NOROWPERM', 'METIS_AT_PLUS_A', 2, 128, 20]	2.87E+01	

Autotune May 4 email from Yang

- Hi, Steve, Nate and Jin,
  I'm running the autotuning with openmpi+GNU compiled m3dc1 now. However, in addition to the significant slowdown compared to intel+craympi, there is something wrong from the run log and it would nice if Nate or Jin can look into this.
- 1. The calculated physical quantities seem wrong (e.g. the energy appears to be 10^20 whereas using intel+craympi it's 127). I attached a runlog using the Run_05B_XY example.
- 2. Also, I sometimes get a segmentation fault in hdf5_write_* of output.f90, I currently comment out them. I saw this error when I run larger examples than Run_05B_XY, for example Run_05A_XY
- I'm not sure whether you have tried GNU (even with craympi) for compiling m3dc1. If so, have you seen something similar before? I would guess this is a problem related to the compiler rather than mpi.
- To reproduce it, I've prepared the detailed instructions on building m3dc1 with openmpi+GNU here (see the heading Instruction on building petsc with openmpi/4.0.1 and gcc/8.3 on knl cori) :
- https://docs.google.com/document/d/15mv6cfYSsC8un8pHB7ZNjszAyEonS54Ta-957v-ojXE/edit

- Could you try to look into this?

Preconditioner reuse (from 4/27/20)

Here is the summary of SOLVE2 and instructions to call it.

pskip: C1input

the number of ntime steps to use the previous preconditioner

=1, default, the traditional run

Small Test

ntimes=18

1 2 (3) 4 5 ((6)) 7 8 (9) 10 11 ((12)) 13 14 (15) 16 17 ((18))

pskip=1 SymbolicFC=39 NumericalFC=39

pskip=3 SymbolicFC=28 NumericalFC=28

pskip=6 SymbolicFC=25 NumericalFC=25

overall solve time decrease: ~10%

(this is a small case)

on 2 nodes: OUT OF MEMORY

on 4 nodes: OK

Jin Chen

code committed.

GPU Timings (from 4/13/20)

1. comparison solve times of using 1-gpu, 8-gpu, and 16-gpu:

16-gpu run is twice as fast as a 1-gpu run.

log.mpi16_**gpu1**_O2
LOOP TIME 5 Tot 3.8548E+02 compute 3.6474E+02 solve 2.0741E+01
log.mpi16_**gpu8**_O2
LOOP TIME 5 Tot 3.7791E+02 compute 3.6452E+02 solve 1.3394E+01
log.mpi16_**gpu16**_O2
LOOP TIME 5 Tot 2.3248E+02 compute 2.2135E+02 solve 1.1138E+01

2. Comparison solve times of code compiled with INTEL & PGI compiler:

INTEL compiler is 3 times faster than PGI compiler

log.mpi16_gpu8_**O2**
LOOP TIME 5 Tot 3.7791E+02 compute 3.6452E+02 solve 1.3394E+01
log.mpi16_gpu8_**pgi**
LOOP TIME 5 Tot 2.8449E+02 compute 2.5149E+02 solve 3.3002E+01

3. Comparison solve time on CORI HASWELL & CORI GPU

Code runs faster on HASWELL than on CORI GPU

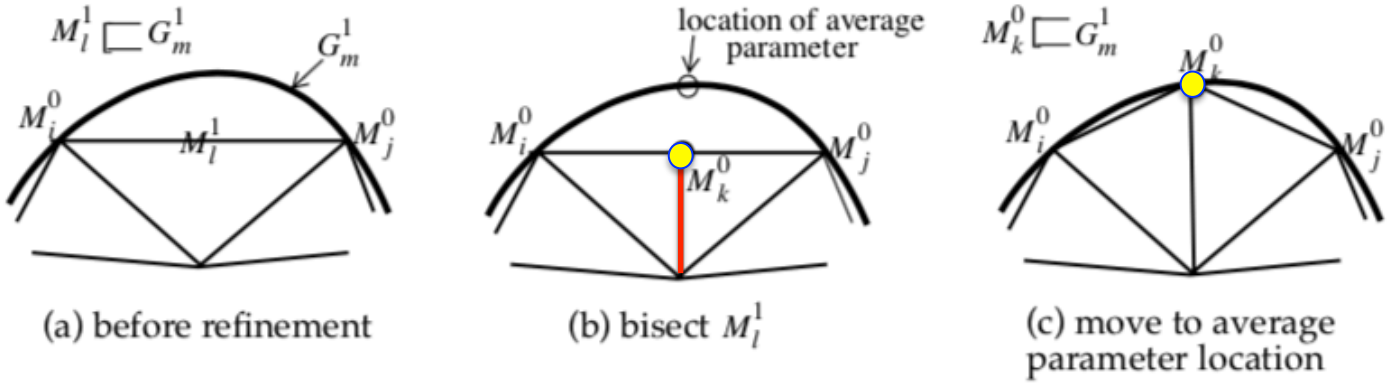
log.mpi16_2d_**hsw**:
LOOP TIME 5 Tot 9.5476E+00 compute 2.6574E+00 solve 6.8902E+00
log.mpi16_gpu16_**O2**
LOOP TIME 5 Tot 2.3248E+02 compute 2.2135E+02 solve 1.1138E+01

2D Mesh Modification Operation in MeshAdapt

2D Operations

- Edge/Face splitting
- Edge swapping
- Edge Collapsing
- Vertex Snapping (not used in M3DC1)

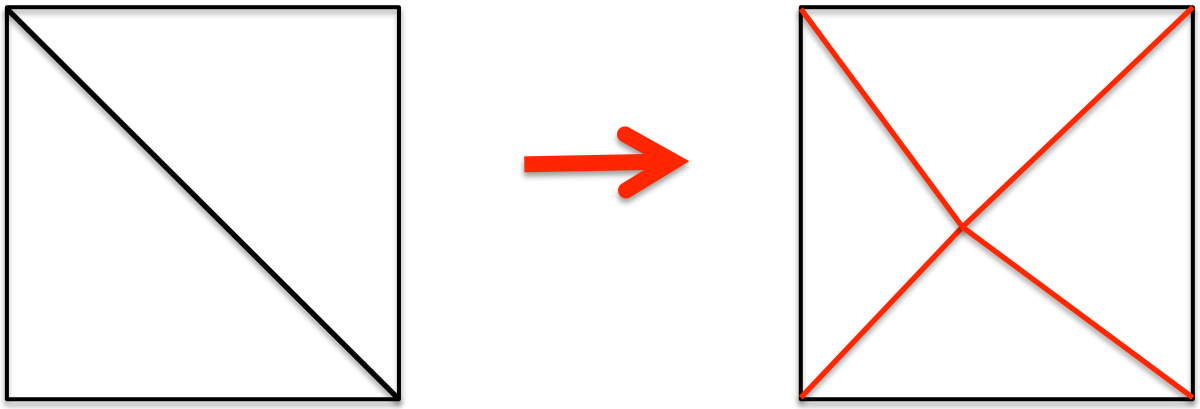
Edge Splitting followed by Vertex Snapping



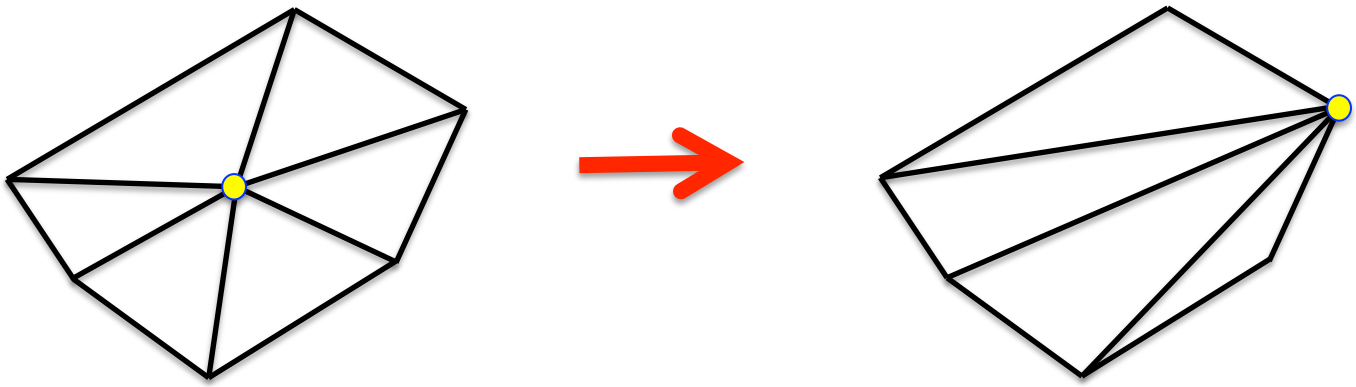
(a) before refinement (b) bisect M_l^1 (c) move to average parameter location

Figure 5.3 An example of target location calculation.

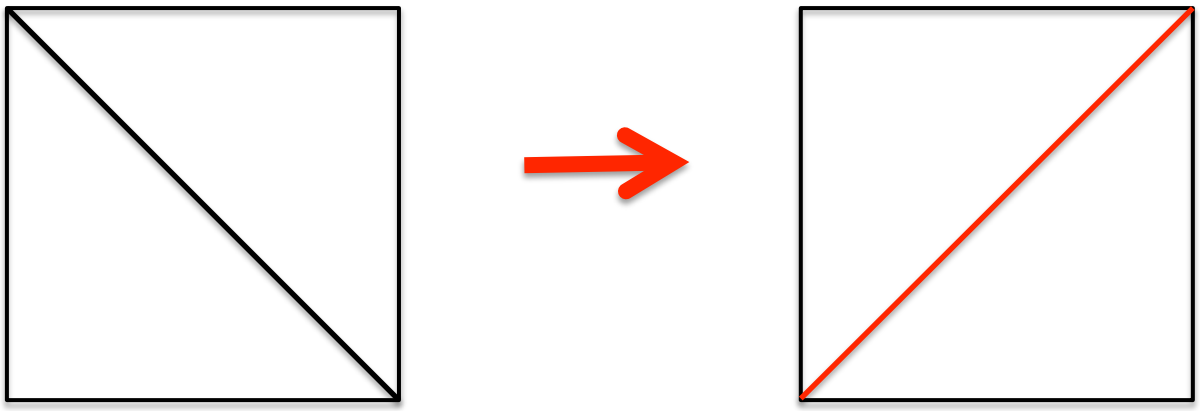
Face Splitting
(equivalent to a series of edge splitting)



Edge Collapsing



Edge Swapping



Edge Swapping followed by Vertex Snapping

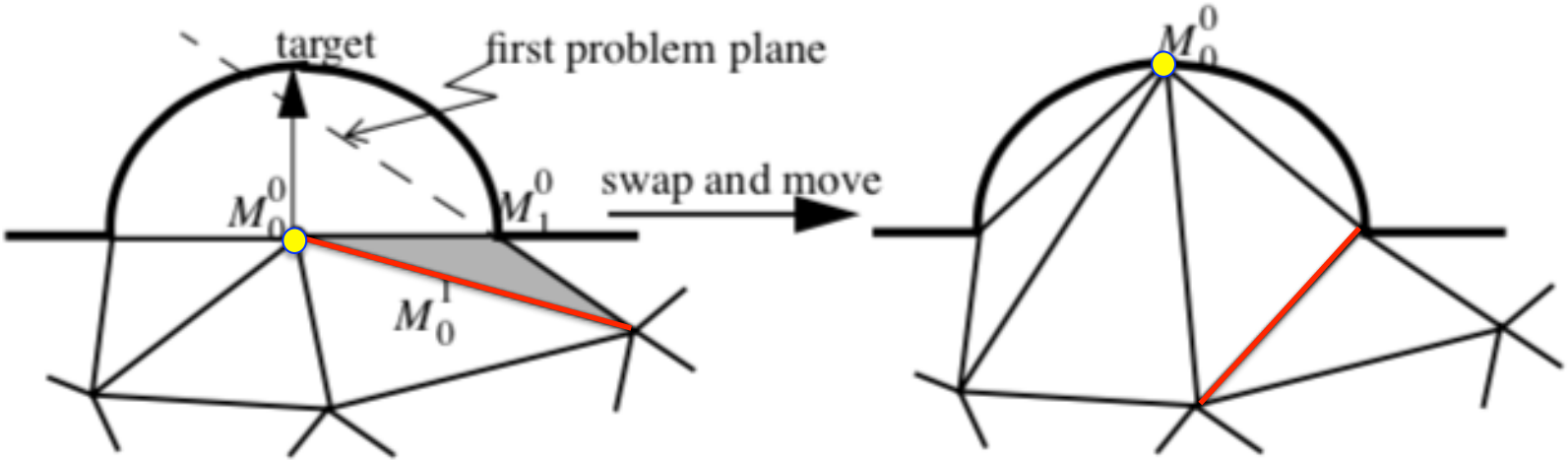


Figure 5.9 Application of a swap operation to eliminate a problem face.

New Modules on Portal

Code works for nskip and pskip with the following:

sunfire.openmpi-4.0.1.mk

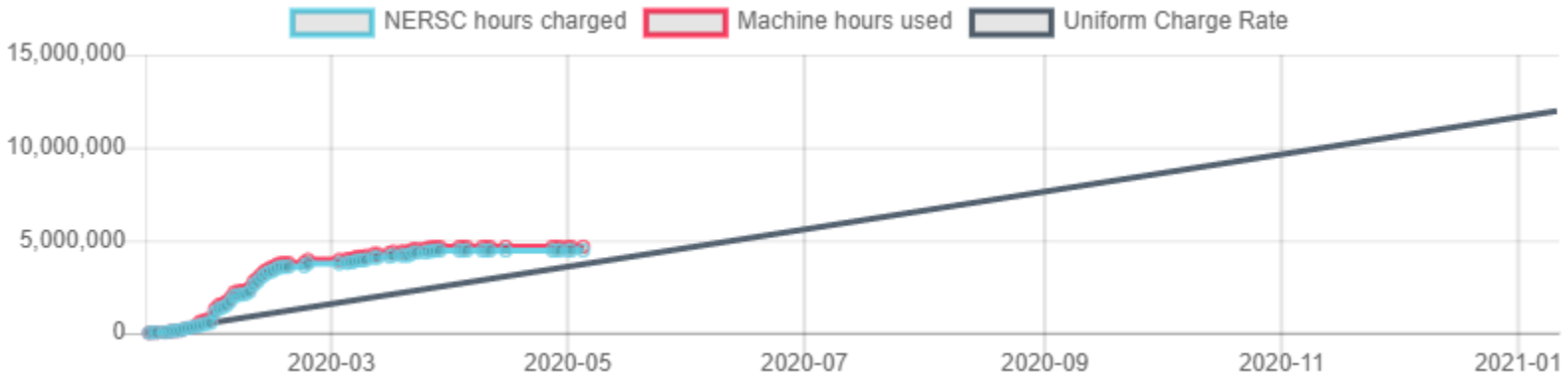
Follow instructions in:
README/readme.portal.openmpi-4.0.1

However, problem with the complex version…MPI errors

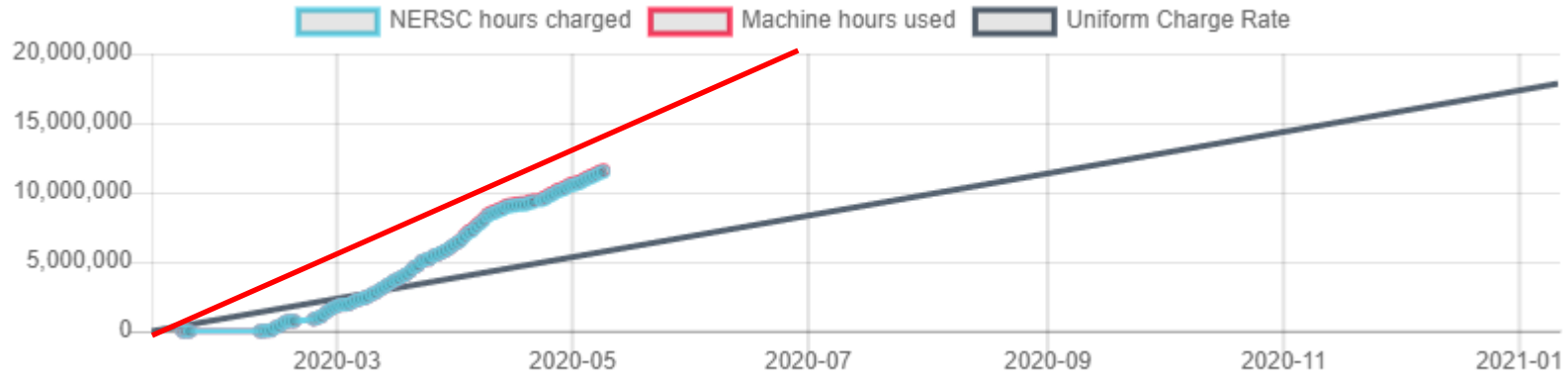
Jin Chen

NERSC

MP288



M3163



Need to use less mp288. m3163 approaching linear usage rate

Add to batch file: #SBATCH –account=3163

(clauser, kleiner, lyons, strauss)