

# **MPP M3D and Large Scale Calculations**

**L.E. Sugiyama**

**Massachusetts Institute of Technology, Cambridge MA 02139-4307**

**CEMM Meeting  
APS-DPP, Philadelphia PA  
28 October 2006**

# TOPICS

- **MPP M3D for ELMs, changes from standard M3D**

Plasma/vacuum/wall

Initial MHD Equilibrium and packed poloidal plane meshes

BC's and poisson solvers

FFT's

Upwind advection of density

- **Using (or not using) the MPP computers for DoE Theory Milestone 2006**

Seaborg

Bassi

Jacquard

Jaguar

Ewok

# M3D MPP for ELMs

- Transfer OMP ELM model to MPP code

- New ideal MHD equilibrium

Generate from GEQDSK (including density profile, if avail) using OMP M3D code (Strauss/GY Park)

Data from experiment or XGC code

- Poloidal meshes, packed around plasma edge and adjacent vacuum region (linear triangular FE's)

Closer packing on outboard side

Up-down asymmetry if needed

Plasma core excluded to save cpu; inner magnetic flux surface boundary

- MPP load balancing for poloidal mesh

- New boundary conditions, new optimization of poisson solvers

Increase dissipation near boundaries for numerical stability

Poisson solvers different from OMP code (iccg/jacobi)

- **Toroidal periodicity and restart changing periodicity**
- **Toroidal direction FFTW Fourier transform package gives typical 3x speedup over discrete FT for numbers of pol planes not  $2^k$ .**
- **Upwind advection of density for nonlinear evolution (using OMP algorithm).  
Edge density change is larger than temperature change in ELM!**

# RESULTS

- Linear perturbations at given toroidal mode number  $n$  work well.  
MHD or 2F  $\omega_{*i}$ , some full 2F
- Efficient with small number of processors up to medium high  $n \simeq 20$   
Toroidal periodicity  $n$ , 8 pol planes, 4 proc/domains in each pol plane  $\rightarrow$  32 processors  
Enable successive step-up to more pol planes
- Ready to combine several simultaneous linear runs to test multiple  $n$ 's.
- Pol plane mesh requirement is sensitive to  $n$ .
- DIII-D and C-Mod Type I ELM linear perturbations  
C-Mod Type I ELMs described by J. Terry, APS-DPP G11 Tuesday morning
- Work gives basis for scaling to 1000+ processors.

# C-Mod Resistive MHD and Two-Fluid Growth Rates

$n = 6$	12	20
<b>Equilibrium 2 (844)</b>		
<b>MHD</b>		
3.08e-2	4.56e-2*	6.78e-2
<b>2F</b>		
NA	4.16e-1	5.13e-02 <sup>†</sup>
<b>Equilibrium 1 (960)</b>		
<b>MHD</b>		
1.70e-2	2.47e-2	NA
<b>2F</b>		
1.1e-2	2.01e-2 <sup>3</sup>	NA

\* $\gamma = 4.566\text{e-}02$  at  $S = 4 \times 10^7$  is similar to value at  $S = 10^6$ .

<sup>†</sup> $\gamma = 5.156\text{e-}2$  at smaller  $H = 0.02$ .

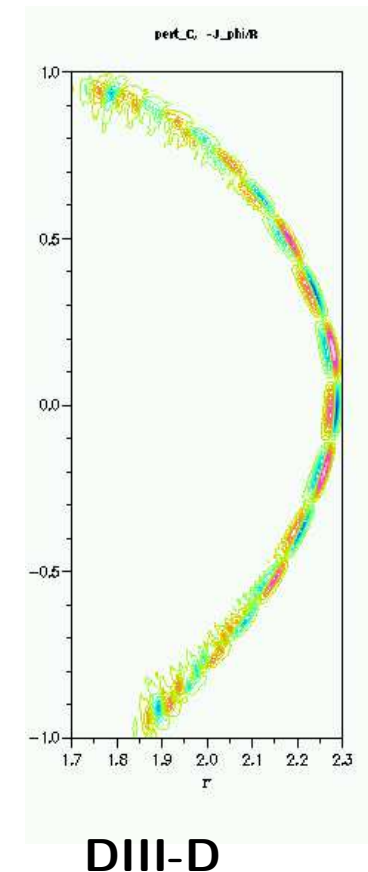
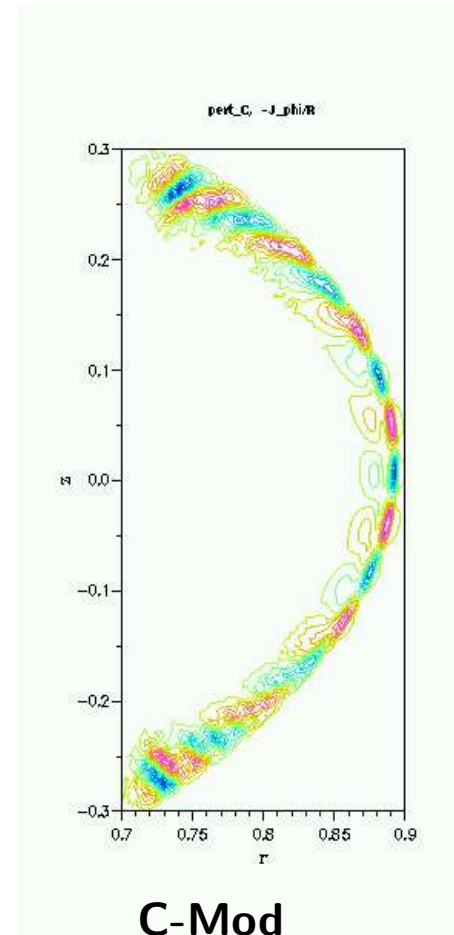
<sup>3</sup> 2F electrons and ions.

- Two-fluid keeps ion gyroviscous effect  $\omega_{*i}$ .
- Results: Growth rates are smaller for Equilibrium 1, well before ELM crash. No major differences in profile shape, but pedestal and current peak are smaller.
- Ion gyroviscosity ( $\omega_{*i}$ -effects) reduces growth rate by small amount for  $n = 6$  and 12, more for  $n \geq 20$ , for two-fluid strength  $H = c/(\omega_{pi}R) = 0.04$ . For  $n = 12$ , results are similar for smaller  $H = 0.02$  and  $0.04$ .
- Other factors help stabilize higher- $n$  modes in C-Mod?

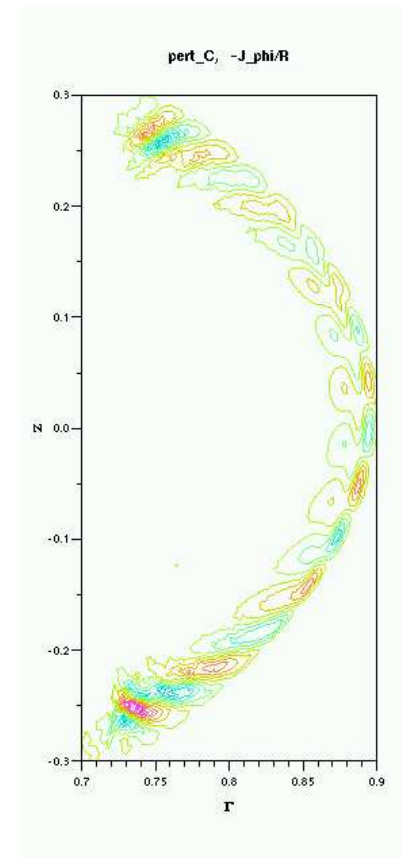
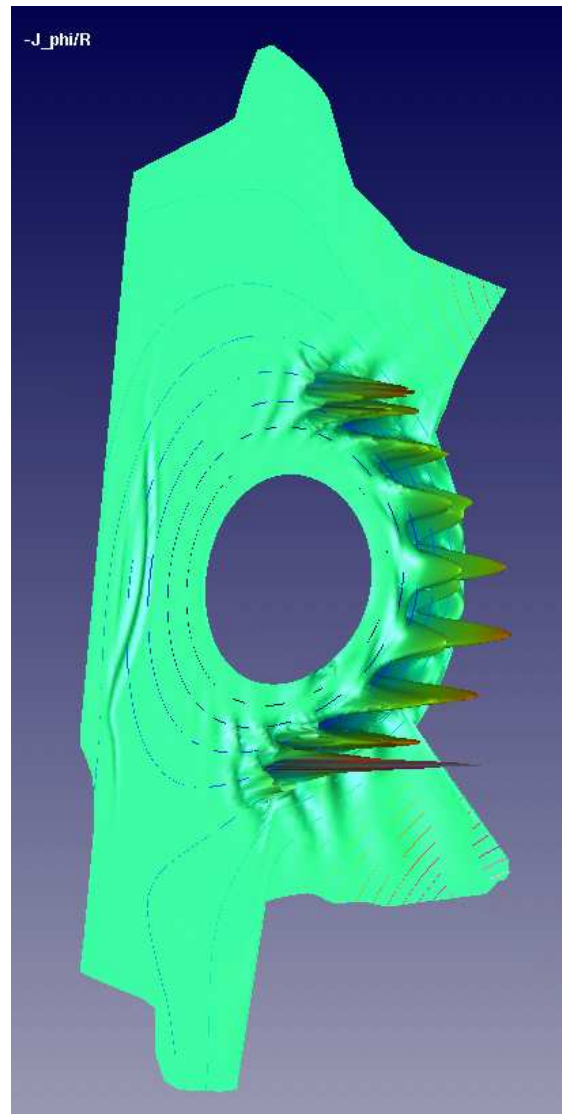
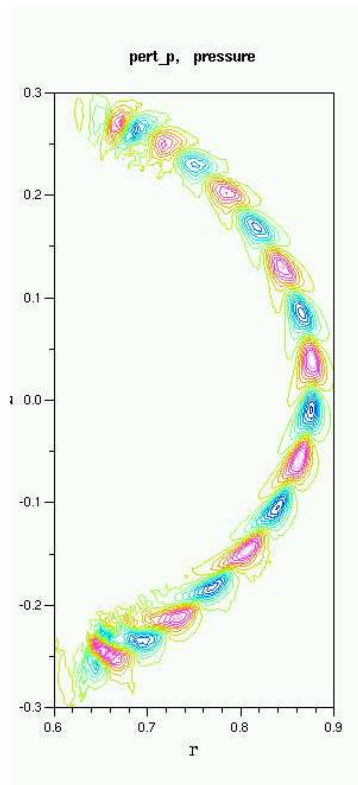
# C-Mod Instability Differs from DIII-D Type I ELM

- Different plasma shapes
- C-Mod growth rates are lower:  
 $n = 12$  MHD mode has  $\gamma \simeq 0.2$  in DIII-D, at somewhat higher resolution.
- DIII-D strongly ballooning
- $J_\phi$  and flows tilt more toward horizontal at top, bottom of disturbed region in C-Mod, rather than radially inward (more peeling than ballooning?)

Shown: perturbed  $J_\phi$  for  $n = 12$  mode in C-Mod Eq-2 (left) and DIII-D 113317.

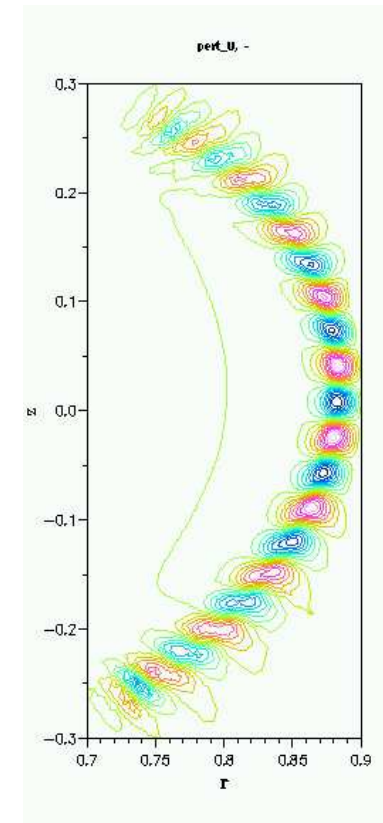
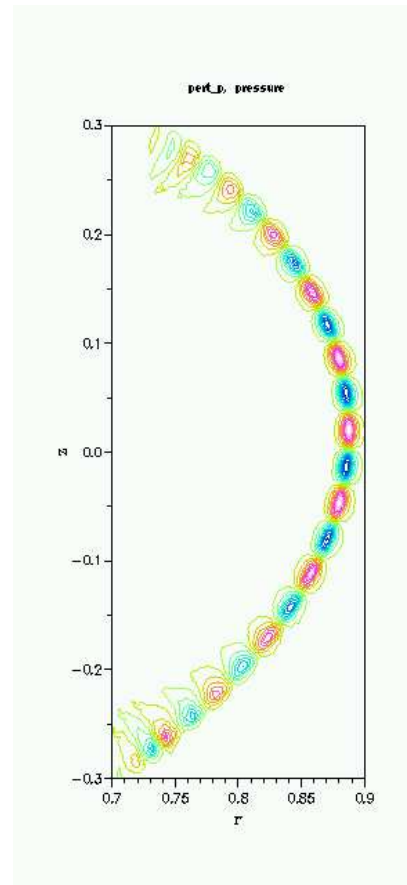
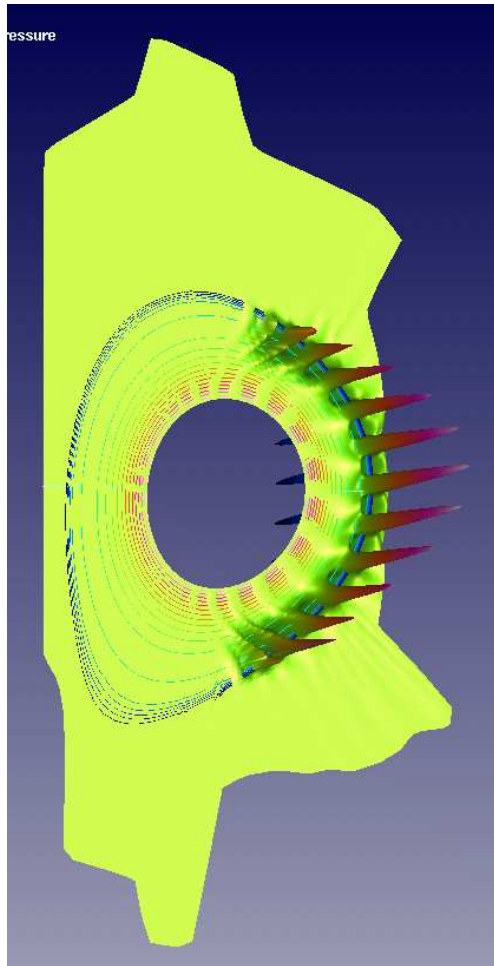


**C-Mod Two-Fluid  $n = 12$  Eigenfunction:  
Pressure and Toroidal Current Density (2 views) against equilibrium pressure)**

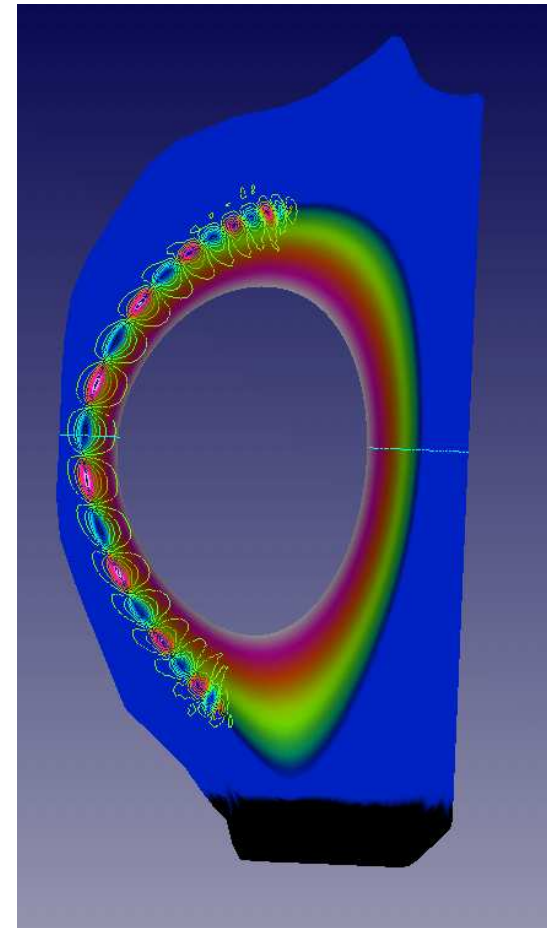
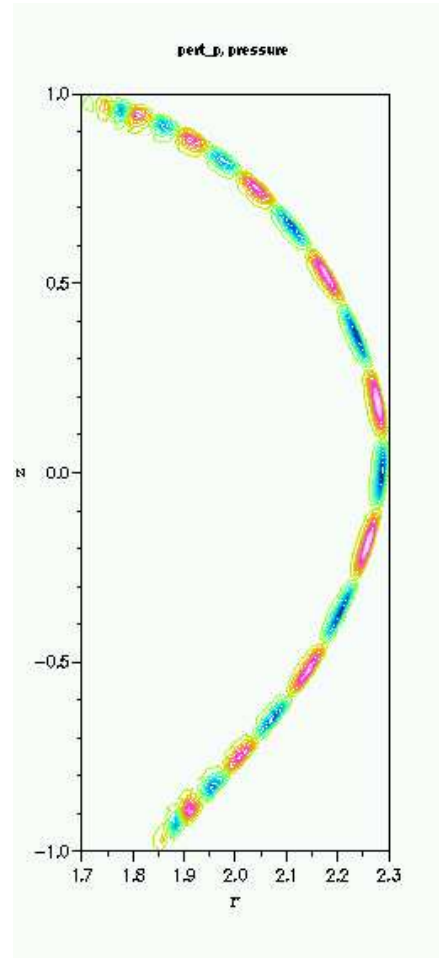
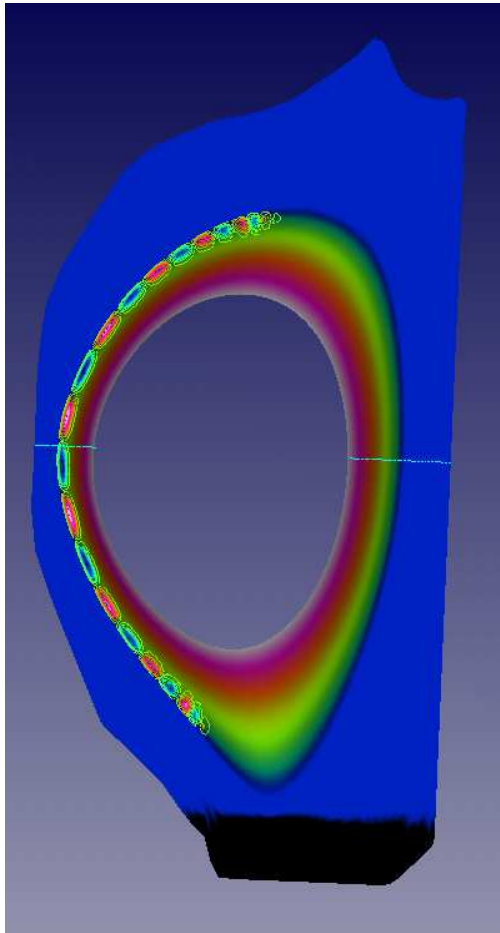




**C-Mod MHD  $n = 20$  Eigenfunction:  
Pressure (2 views) and Velocity Stream Function (Eq-1 shown, similar to Eq-2)**



DIII-D MHD  $n = 12$  Eigenfunction:  
Pressure (2 views) and Velocity Stream Function



# Large MPP Runs

- **NO full size MPP nonlinear runs were successfully performed for the DoE Milestone (NERSC, NCCS systems)**

**Very few full size runs got through the batch queues at all on any MPP system**

**Full size,  $n=0-40$  toroidal mode run for Milestone was done on NCCS Ram with OMP code (80 proc)**

- **Problems with MPP computer systems for large runs**  
 **$\sim 400$  processors with 96 pol planes, 60 radial by 400 poloidal (slightly larger than OMP run)**

**Batch queue waits and job turnaround, machine downtimes**

**Software and operating system (OS) problems**

**Poor interprocessor communication on most systems slows input/output and sometimes computation**

**Little or no useful MPP debugging facilities,**

**Newer runs show that probably need larger, higher resolution poloidal grid for this number of toroidal modes (48, smoothed at end of time step to 40).**

- **Since end of milestone, more large scale runs have gone through, but couldn't look at output from China (IAEA conference) — files too big to download onto my laptop for visualization!**

## Seaborg

- Standard code uses PETSc MPI libraries, HDF5 files for visualization (S. Klasky's AVS IDAVE), FFTW Fourier transform package for pseudospectral  $\partial/\partial\phi$  derivatives.
- Good turnaround on small/medium size jobs (16–64 proc) began in Dec 2005 until mid-July 2006, then poor.
- Large size jobs difficult: for 400 proc job, even a +7 day priority increase for the Milestone, mid-Aug to end-Sept, helped only marginally.
- Excellent interactive small jobs (up to 64 proc).
- OMP M3D runs well on up to 16 processors (1 node). Used to generate ELM input files for MPP code. Also used Ram.
- Consultants helpful.
- Down for security incident early Oct.  
Locked out until this Friday, since on travel.

## Bassi

- **Slow PETSC matrix solves (Poisson operators), speedup only 2x over Seaborg, while charge factor is 6x. Confirmed by NERSC consultants...**
- **Slow MPI parallel file reads/writes (checkpoint files).**
- **Otherwise, runs well for smaller jobs, relatively stable, similar to Seaborg.**
- **400 proc jobs have longer waits, but often less than Seaborg.**
- **Poor interactivity, usually only 16 proc jobs possible (nominal 32).**

## Jaguar

- M3D only working effectively on Jaguar in mid-July 2006.  
Small test runs looked promising, but then machine down for 3 weeks for upgrade.
- After upgrade, found FFTW package didn't work for M3D (FFTW was first used for M3D ELMs during downtime, but package existed on Jaguar long time, works for other codes. For M3D, it works on all other machines where installed.) Switched back to regular FFT and discrete FT, works well.
- Fast compared to Seaborg, Bassi!
- Many downtimes.
- No debugging (totalview, etc exists, but way too slow to use).  
No interactive runs.
- Can't run long jobs (eg 24 hour) on smaller numbers of processors – multiple restarts required.

- **Lustre file system problems.**  
Very slow MPP parallel read/writes (checkpoint files).  
Recently, random job failures during checkpoint writes.
- **Consultants less experienced than Seaborg.**
- **OK turnaround for 400 proc jobs (16 hour) recently, but:**
- **Oct 6 OS upgrade. Results now differ from Bassi! (Linear ELM growth rates not too different.)**



## Jacquard

- **Can't run useful jobs: MPI parallel checkpoint writes fail (segmentation fault).  
Also read/writes extremely slow.**
- **Single jobs run OK if don't write a checkpoint file.**
- **Poor debugging, interactivity.**
- **NERSC consultants give Jacquard low priority.**

## Ewok

- Medium size system (160 proc) at NCCS, similar to Jacquard, used for CPES.
- PGI compilers work particularly poorly with PETSc.
- Extremely slow interprocessor communications problems after Infiniband was installed (M3D ran well before).  
Not just the checkpoint file read/write, but also PETSc.
- Possible to run M3D, but very slow (should be approx 2x faster than Seaborg, but probably 10x slower).
- No debugger.

## MPP computers – Summary

- Available MPP computers are difficult to use for practical results.

Small and very small jobs (to 64 proc) run well on most systems.

Medium to largish jobs run sluggishly or are limited in duration (queue restrictions).

Difficult to do the runs and groundwork needed to scale real jobs to 1000 or more processors and petascale computing.

- Most systems (except Seaborg) have serious interprocessor communication problems that make parallel input/output slow. In some, communication also interferes with calculation speed.
- Poor debugging facilities.

All the installed debuggers are poor for MPP codes, slow over a network, work poorly with PETSc.

Most debugging must be done by write statement.

Good interactivity for multiple processor jobs is a must (at least 2 nodes on multiple proc/node systems, more on dual-proc systems)

- **Better reliability of software/OS is required!**  
Large amounts of time spent moving code/files around to test systems after changes.
- **Jaguar, Jacquard, Ewok use PGI MPP compilers — not a good choice. In particular, problems with PETSc.**

# SUMMARY

- MPP M3D for ELMs, changes from standard M3D

Plasma/vacuum/wall

Initial MHD Equilibrium and packed poloidal plane meshes

BC's and poisson solvers

FFT's

Upwind advection of density

- MPP computers' scorecard on ELM Milestone, spring thru Sept 2006:

The only successful large scale nonlinear run for the DoE nonlinear ELM Milestone was made on Ram with M3D OMP code! Zero full size runs on Jaguar.

Interprocessor communication and MPP file read/write — generally poor, except Seaborg good.

Compilers: beware PGI compilers, especially with PETSc

Poor to nonexistent debugging for MPP codes:

Debuggers poor, use write statements (!)

Interactivity poor (Seaborg very good)

- **Allowance for code development and testing process is needed to effectively set up very large MPP runs (Seaborg OK, Jaguar poor).**
- **Need to improve reliability, accessibility, interactivity, and user-friendliness of MPP computer systems to**
  - (1) do physics**
  - (2) move to petascale computing for real problems****INCITE allocations are not the answer.**