

The (p) in (p)SLIMpy

Darren Thomson
SINBAD Sponsor Meeting
August 29, 2006



The (p) is silent

Seismic Laboratory for
Imaging and Modeling

- (p)SLIMpy is unique in the English language
 - "ps" beginnings are common
 - NO other word starts with "ps" followed by a consonant
- Enough with that psilly business...
 - time to get pstarted



Parallel Extensions

Seismic Laboratory for
Imaging and Modeling

- **Why?**
 - Seismic data can be HUGE!
- **What?**
 - Scalable
 - Full SLIMpy functionality
 - Minimal work to go from serial to parallel
- **How?**
 - MPI in Python
 - Automatic overlapping domain decomposition



Quick Start

Seismic Laboratory for
Imaging and Modeling

- **dnoise.py**
 - serial

```
dnoise.py data=data.rsfs output=res.rsfs
```

- parallel

```
mpirun -np 4 -hostfile $PBS_NODEFILE dnoise.py  
data=data.rsfs output=res.rsfs nwin=0202 eps=16  
localtmpdir=/tmp globaltmpdir=/scratch
```

- presto!



Break It Down

Seismic Laboratory for
Imaging and Modeling

- Launch (4) MPI process across your nodes

```
mpirun -np 4 -hostfile $PBS_NODEFILE ...
```

- Call dnoise.py as you would like

```
... dnoise.py data=data.rsfs output=res.rsfs ...
```

- Provide parallel parameters to SLIMpy

```
... nwin=0202 eps=16 localtmpdir=/tmp globaltmpdir=/scratch
```



Details...

Seismic Laboratory for
Imaging and Modeling

- **nwin**
 - defines number of windows along each axis
 - 2 digits per axis, no spaces
- **eps**
 - depth of overlap into neighbouring windows
 - eps=0 performs non-overlapping windowing
- **localtmpdir, globaltmpdir**
 - locations to store temporary binaries
 - ensures optimal performance in file I/O



Nuts & Bolts

Seismic Laboratory for
Imaging and Modeling

- SLIMpy detects MPI
 - splits data files
 - parallel operators & vectors inherit from serial
 - minor refinements/tweaks
 - rebuilds output files
- Hidden from user
 - NO changes to applications, algorithms



Requirements

Seismic Laboratory for
Imaging and Modeling

- MPI
 - i.e. MPICH, LAM, etc.
- PETSc
 - libraries for parallel numerical processing
- pypar
 - simple MPI wrapper for Python
- RSF
 - new seismic processing package



A Quick Demo

Seismic Laboratory for
Imaging and Modeling

Look Left!!!



But...

Seismic Laboratory for
Imaging and Modeling

- (non-) overlapping domain decomposition not always desired
 - that's ok!
- All manner of embarrassingly parallel problems are easily solved with SLIMpy
- pypar makes it straightforward