Grid Computing
at
St Jude Children's Research Hospital

David Coss

BOINC Workshop

18 Aug 2011

Outline

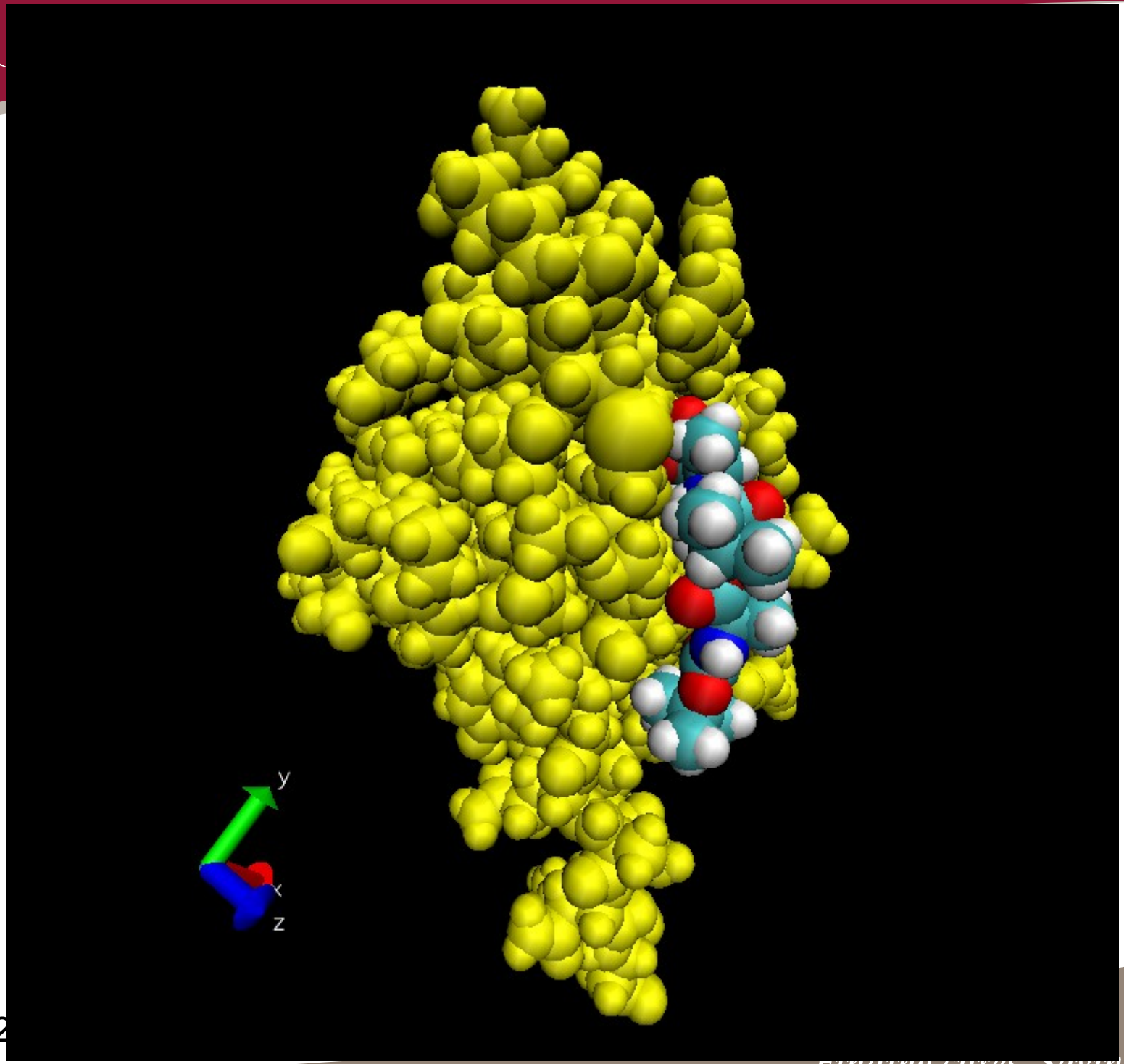
- Introduction
- Science Applications
- Queue System

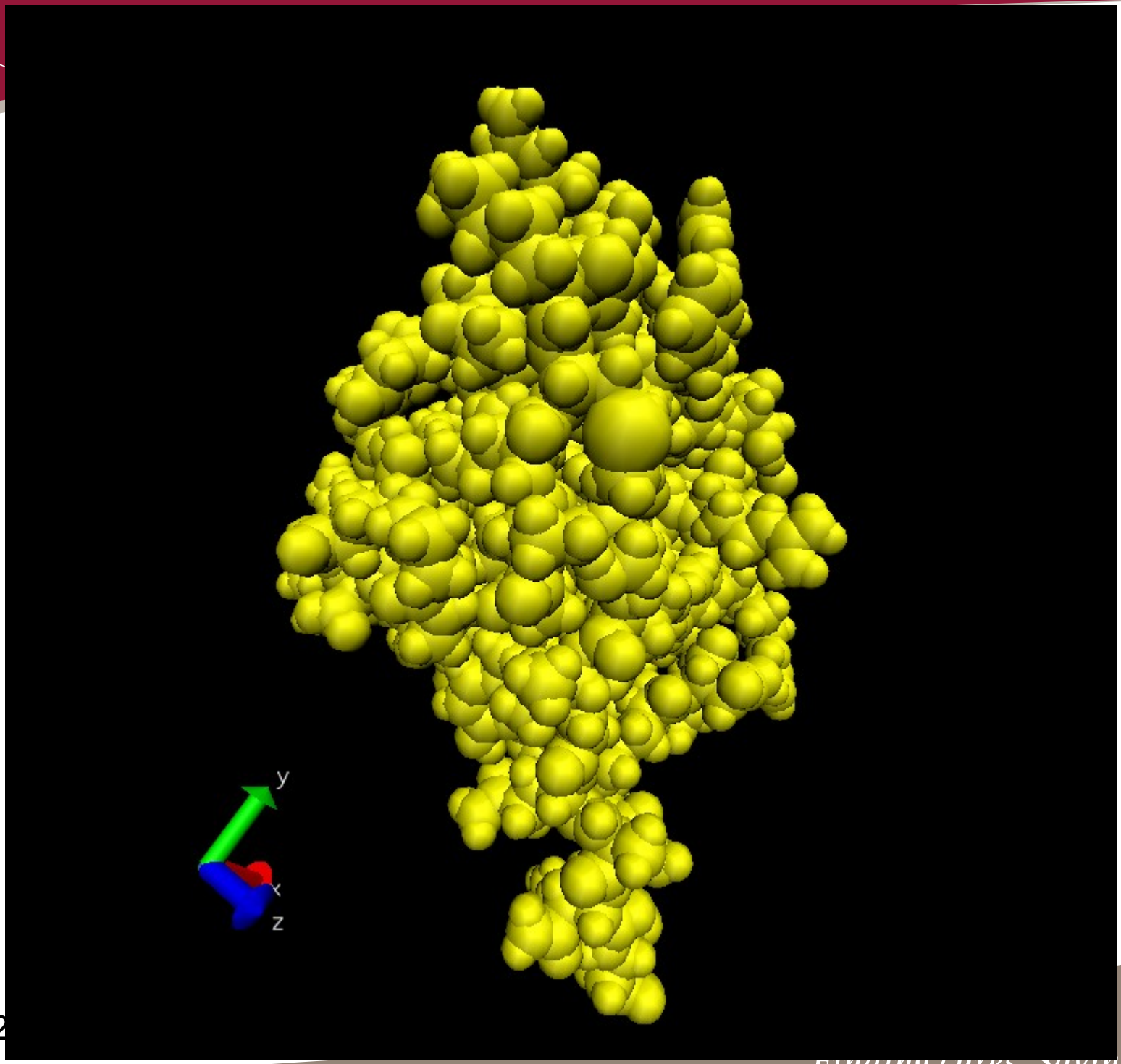
Introduction

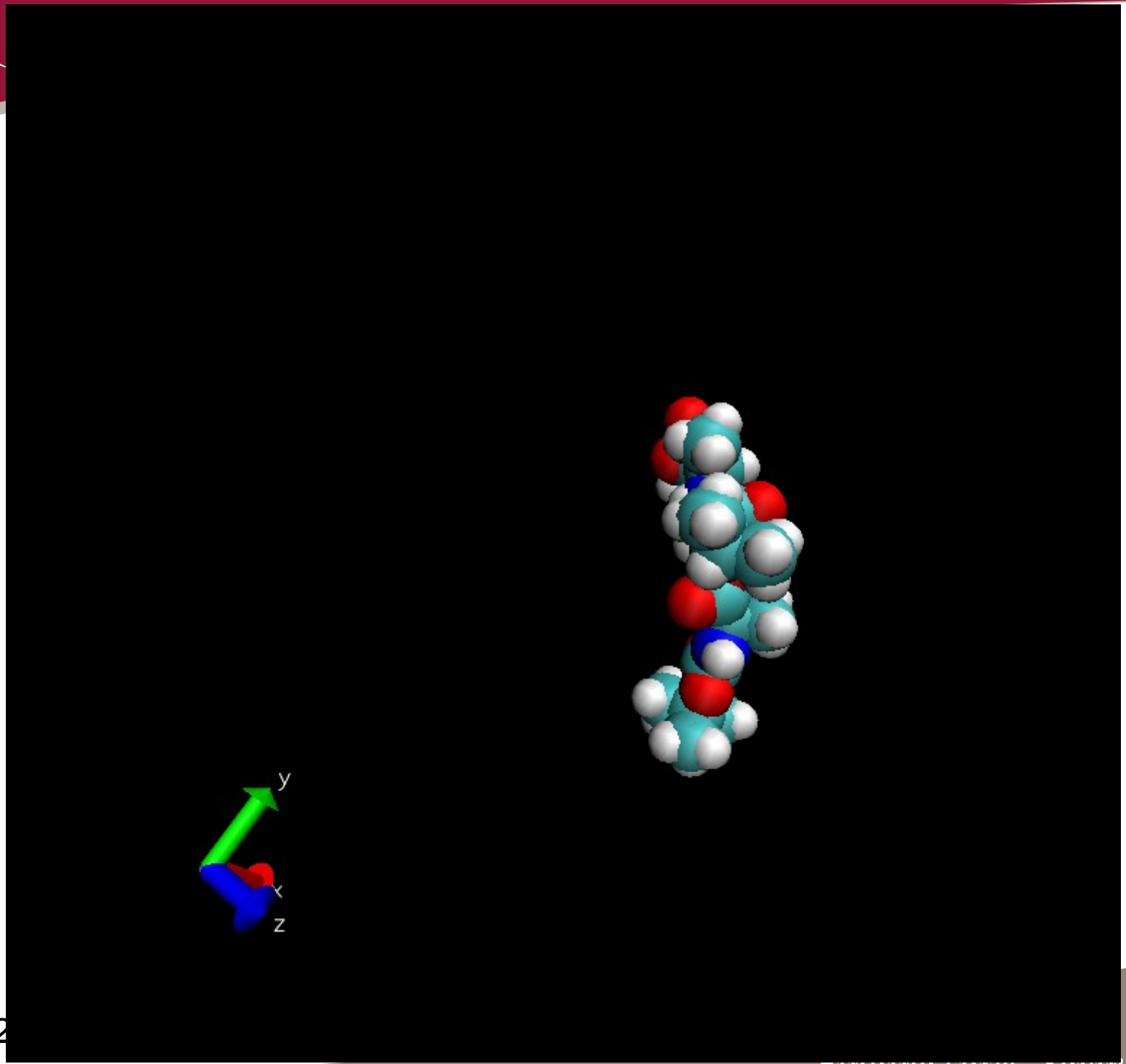
- Grid Computing by Bashford Lab at St Jude
- Started 2010
- Goal
  - Computer Aided Drug Discovery using Molecular Dynamics based Free Energy calculation
- Utilize non-clinical St Jude computers
  - ≤2000 computers

Science Applications

- Molecular Dynamics (MD)
  - *Sander* from *Amber 11*
- Free Energy Calculation
  - MMPBSA
  - In-house program, which...
    - performs Molecular Mechanics
    - Solves Poisson-Boltzmann equation using *libmead*
    - Calculates Surface Area
  - Able to use *Amber* and *Gromacs* input files


y
x
z


y
x
z


y
x
z

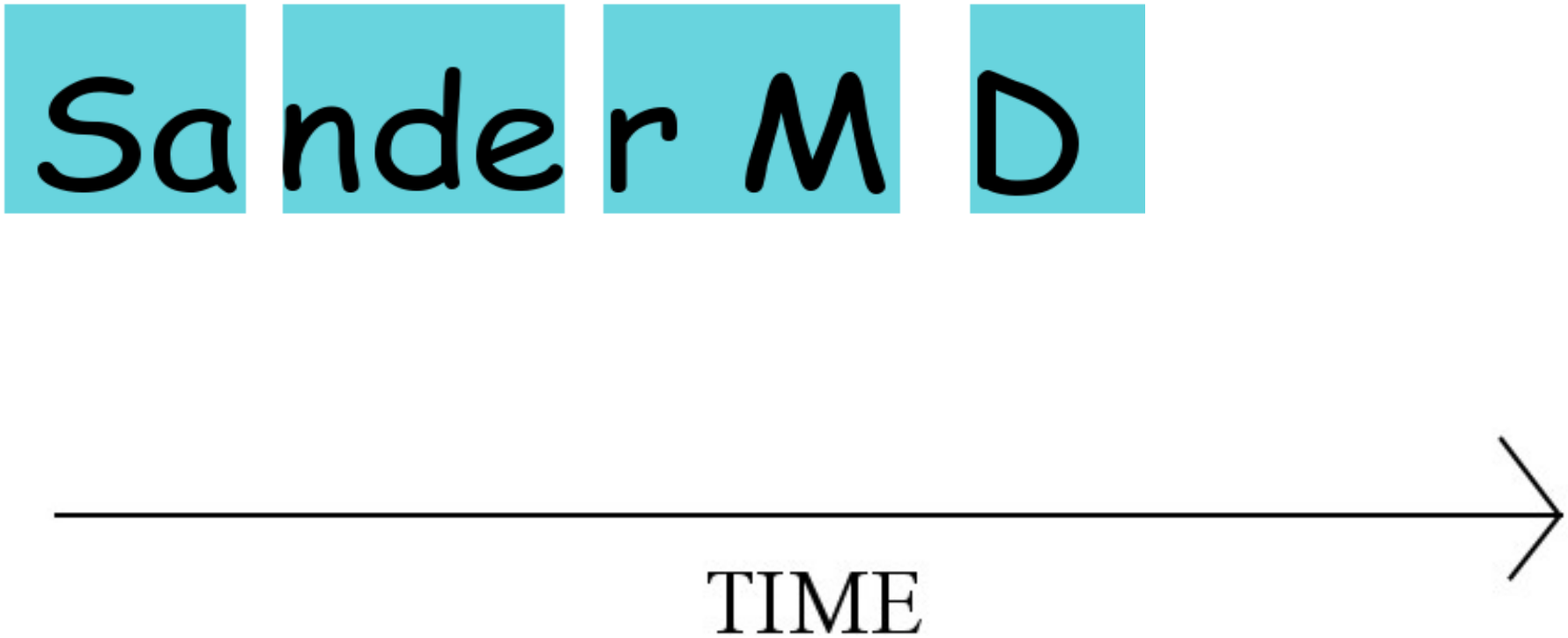
Parallelization

- MD calculations cannot be run in parallel on the grid
- MMPBSA segments can be run in parallel
- Solution:
  - Segment MD calculations
  - Follow MD segments will parallel MD calculations

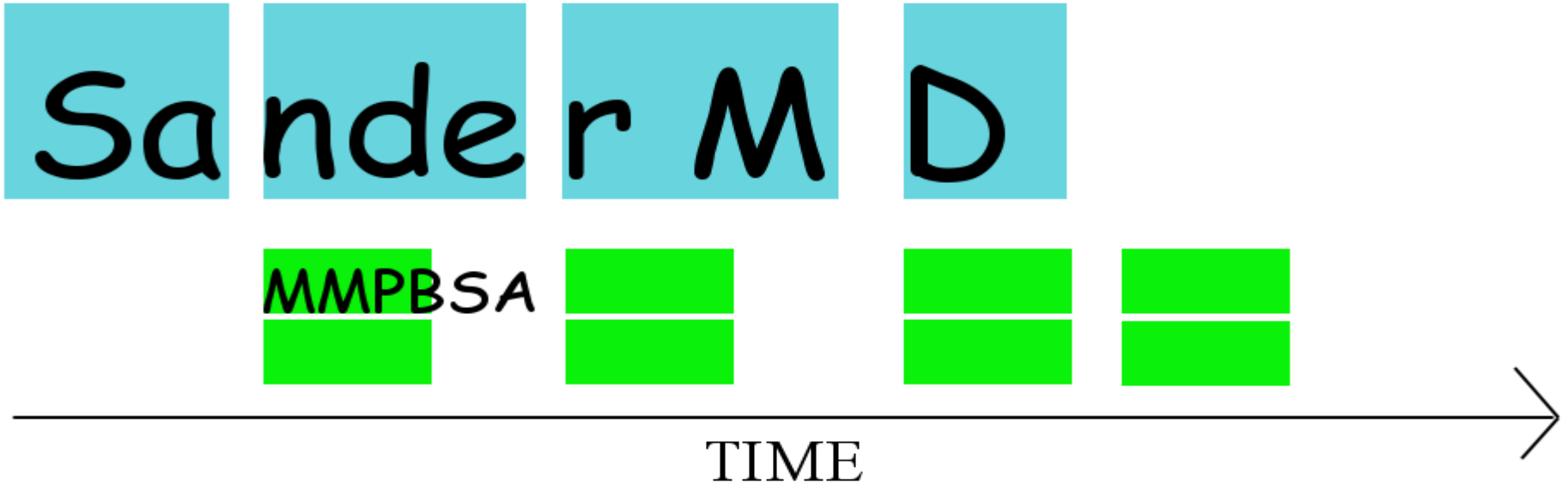
Parallelization

Sander MD

Parallelization



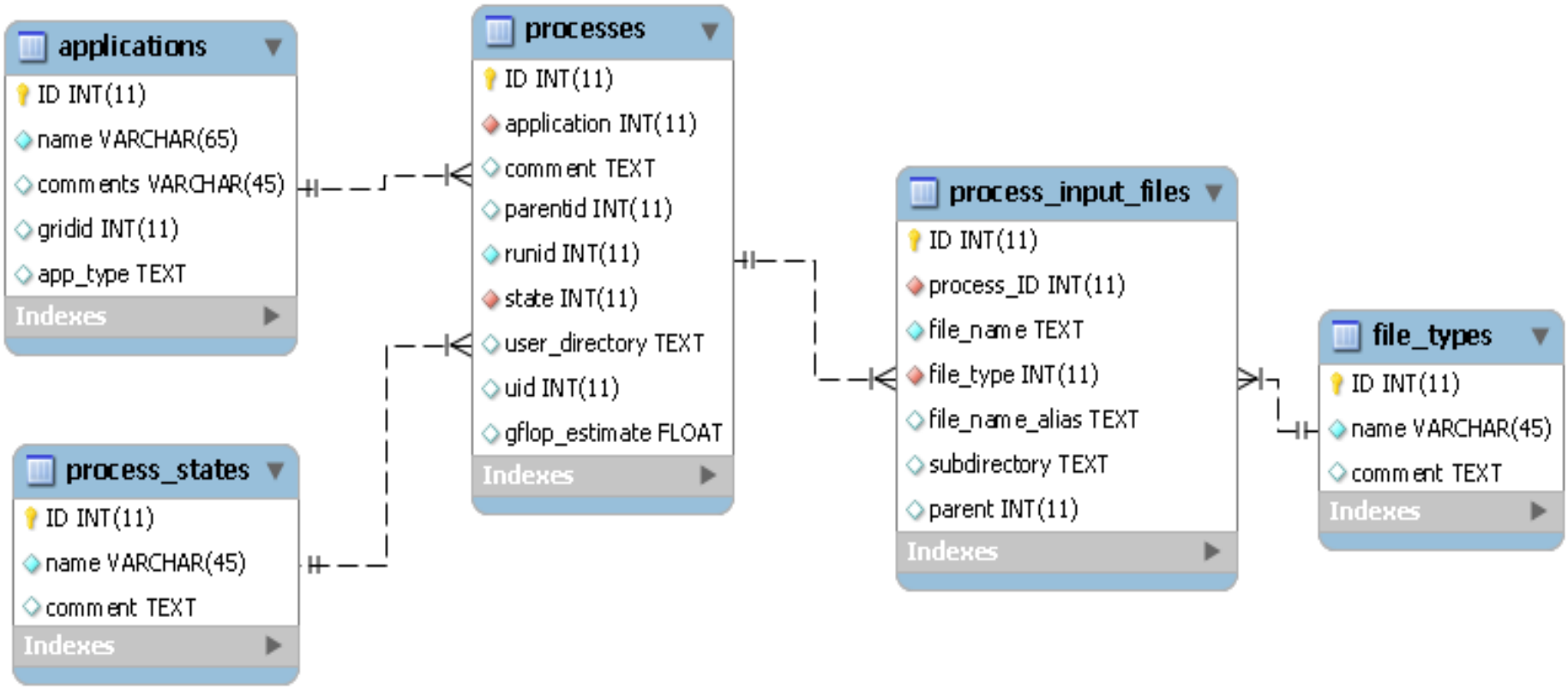
Parallelization



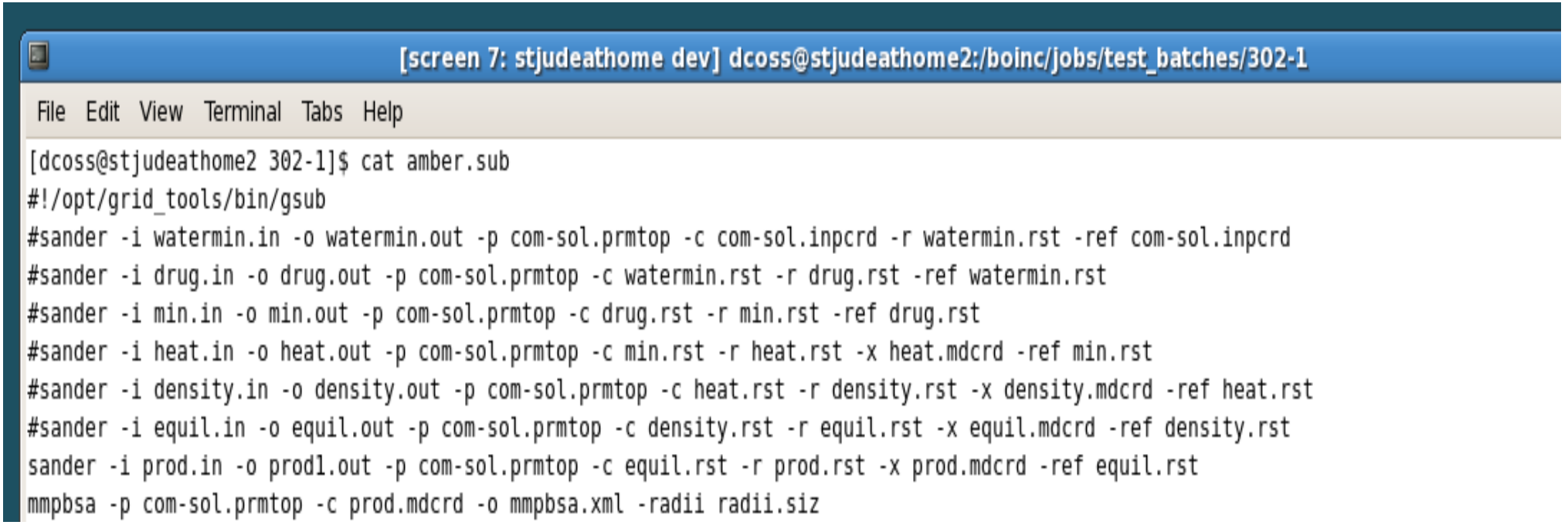
Job Submission

1. Commands written normally, similar to shell script
2. Commands are parsed and interdependency is resolved
3. Processes are segmented and stored in DB
4. Processes are assigned owner based on UID of person submitting work

Queue Database



Job Submission



Job Submission

```
[dcoss@stjudeathome2 302-1]$ grid_master
grid_master -- Program for managing jobs on the grid.
Usage: grid_master [options] COMMAND PARAMETER

Commands:
close -- cancels the job.
process  -- determines whether or not the job should run or continue to wait and acts accordingly.
recombine -- recombines files that belong specified recombination process ID.
start -- starts a job if it is waiting. If it depends on other process, those are started first, if they have not yet finished.
status  -- prints the status of the process.
list  -- prints the jobs in the queue system owned by the user. If a number is specified, all process in that job batch are listed.

Options:
-s, --simulate  Simulates the database queries, without actually sending the job to the grid.
-v, --verbose=[INTEGER] Sets the verbosity of the program. Higher the level, the more verbose.
-h, --help      This help dialog
    --usage     Same as --help
[dcoss@stjudeathome2 302-1]$
```

Acknowledgments

- Work supported financially by
  - American Lebanese Syrian Associated Charities
  - National Institutes of Health grant ARRA R01GM086243