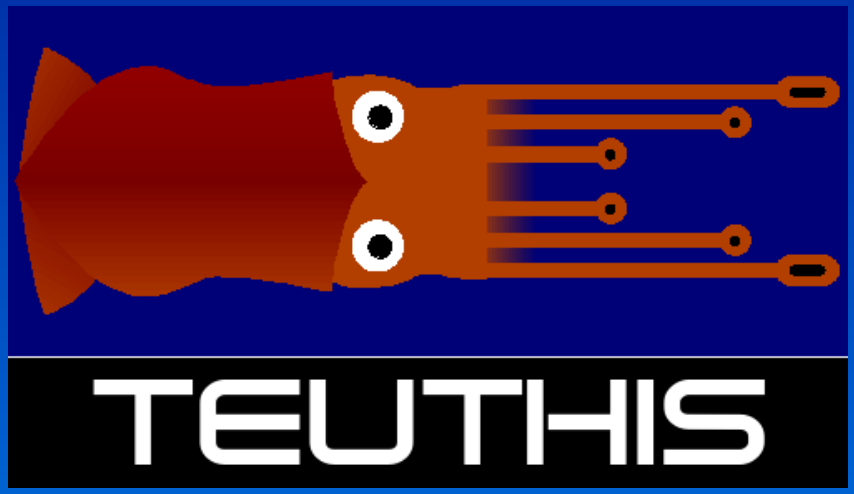
Teuthis: Simulation Management Cyberinfrastructure



Paul Ricker
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign

DES Collaboration Meeting
December 12, 2006

NCSA

What do simulators need to do?

- Control simulation jobs
  - ✓ Configure and build applications on remote systems
  - ✓ Submit and track remote jobs
  - ✓ Create parameter studies and restart jobs
- Manage simulation data and tasks
  - ✓ Stage and archive data
  - ✓ Keep track of where datasets are stored
  - ✓ Organize job metadata by purpose and disposition
- Share data with collaborators
  - Make data available
  - Convert data among different formats

✓Teuthis addresses these aspects of simulation tasks with a small, simple, cross-platform GUI (~ 7,000 lines of Python + PyGTK).

Objects manipulated by Teuthis

Workflow hierarchy

Project
Galaxy cluster scaling relations

Experiment
How does the level of galaxy feedback affect the mass-temperature relation?

Run
Run with 10x fiducial energy input

Job
Job 123456 on 128 processors for 18 hours

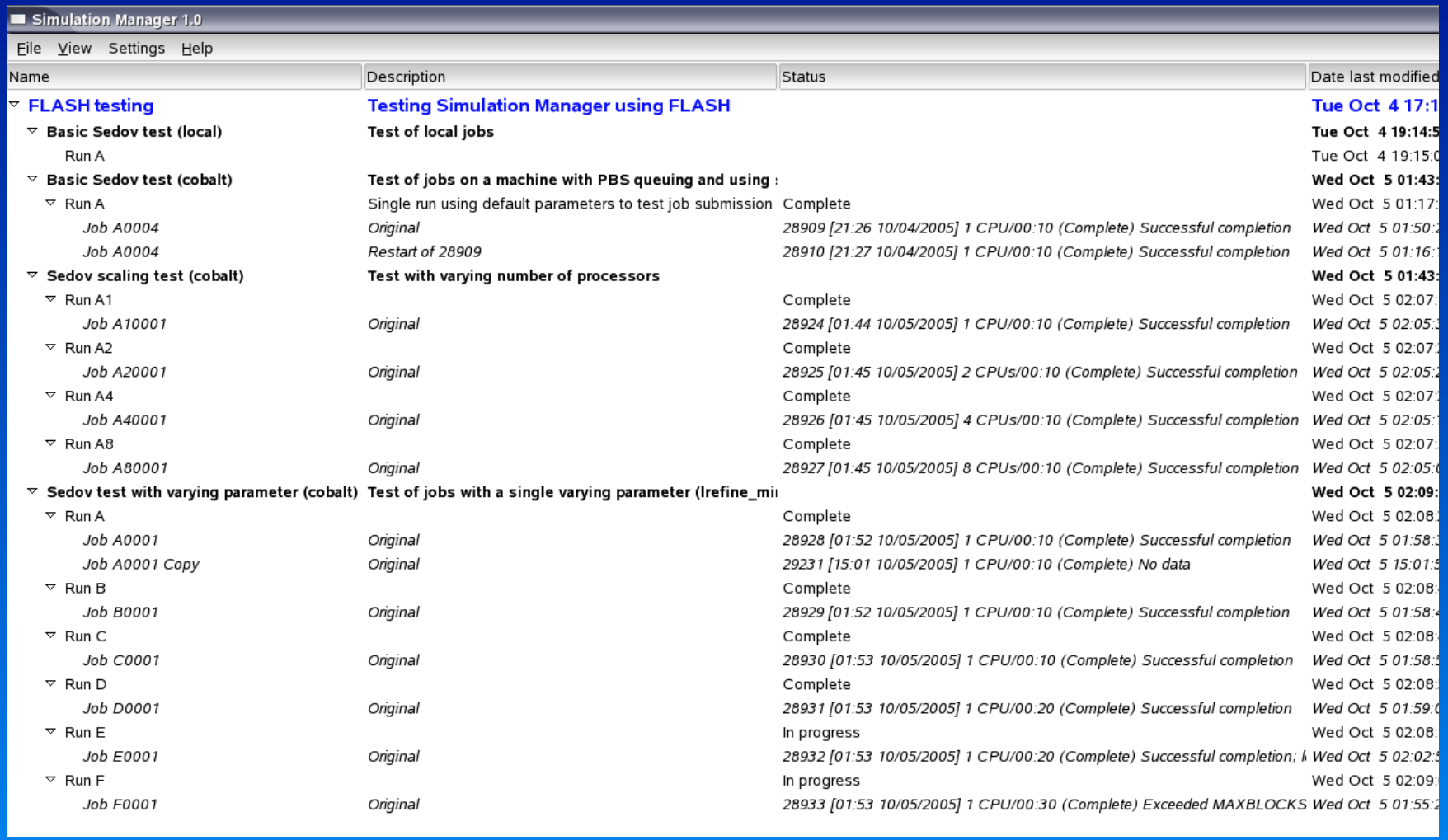
Job
Restart job 789012 on 256 processors for 18 hours

Resources

Application
- Accepts text parameter file
- Executes noninteractively
- May need to be compiled
- Produces log file, screen output, data files

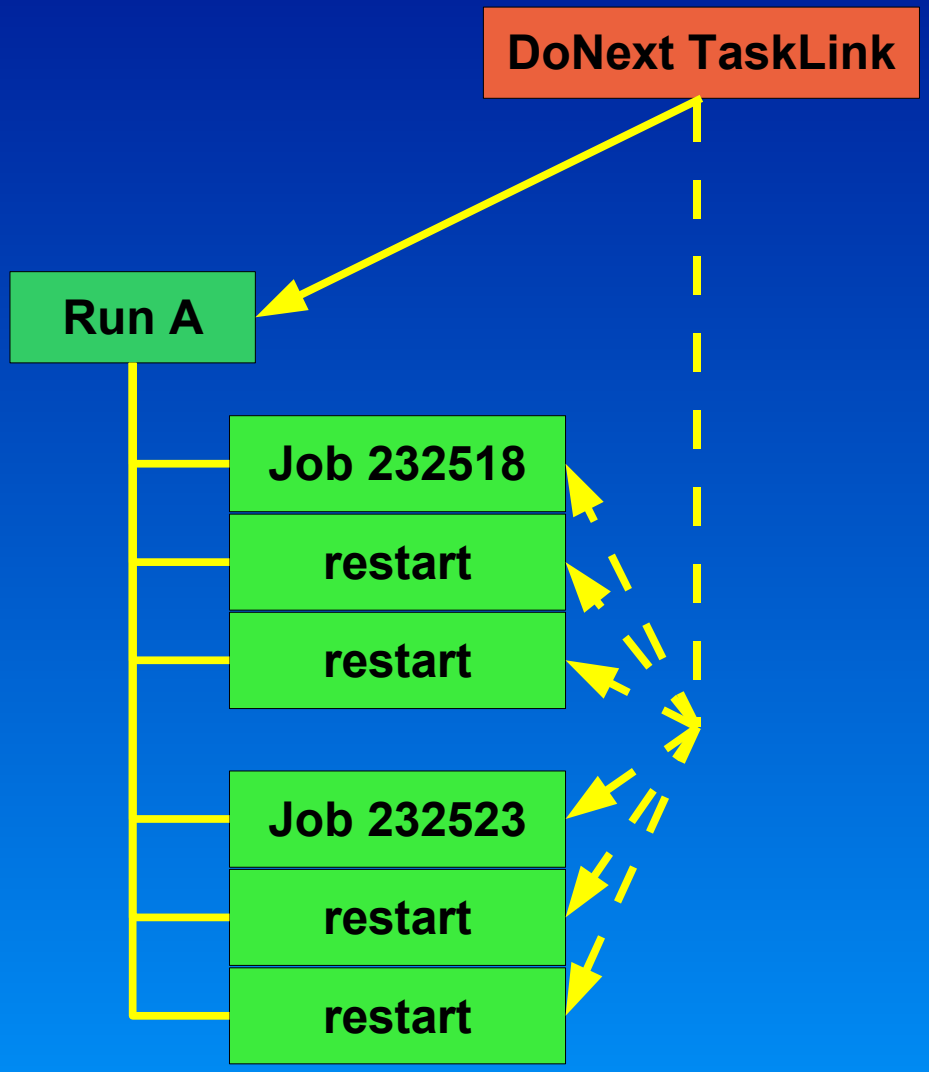
Machine
- Login host
- Access method
- Queuing system
- Paths and commands

Project view

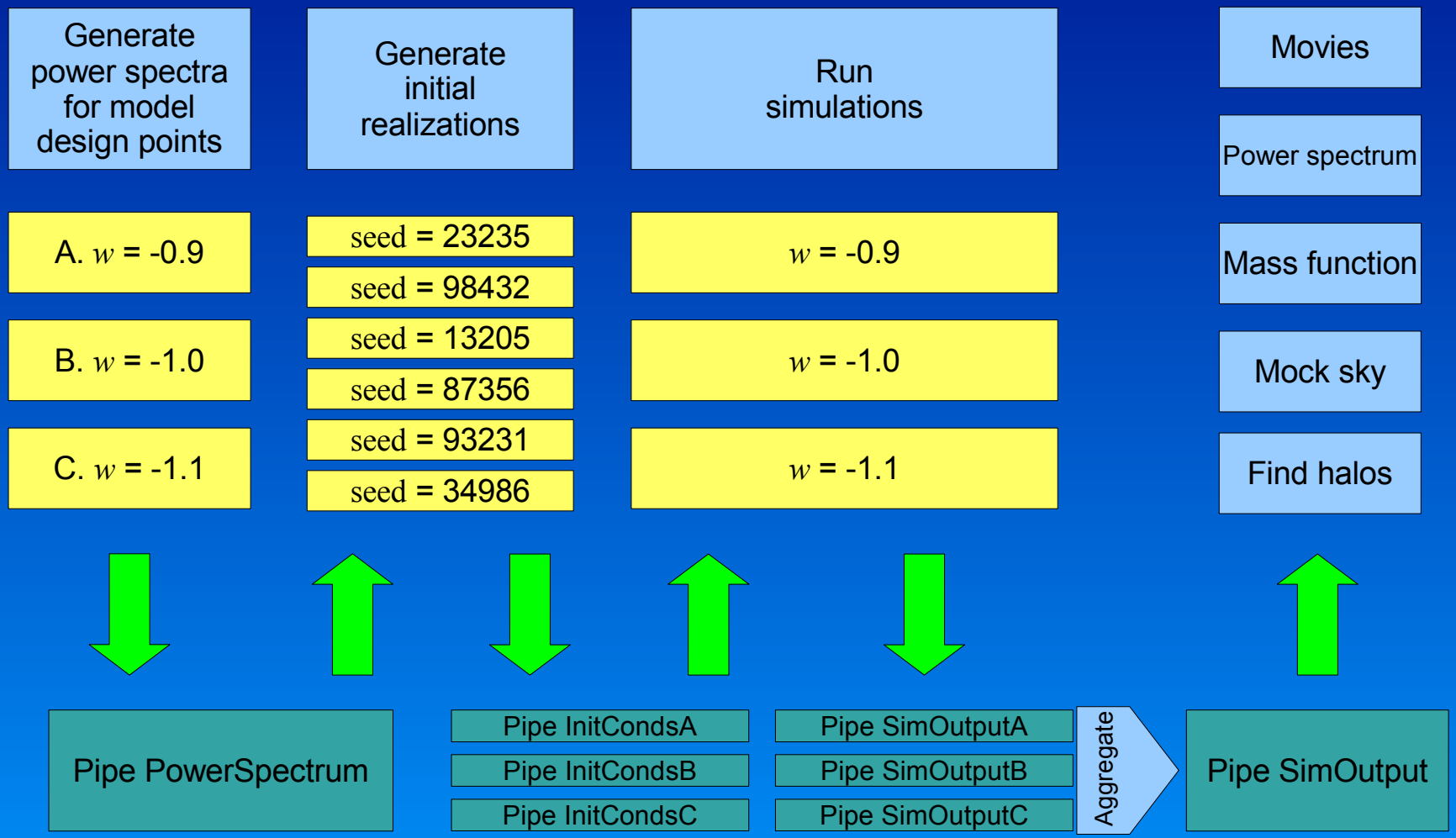


Teuthis workflow management

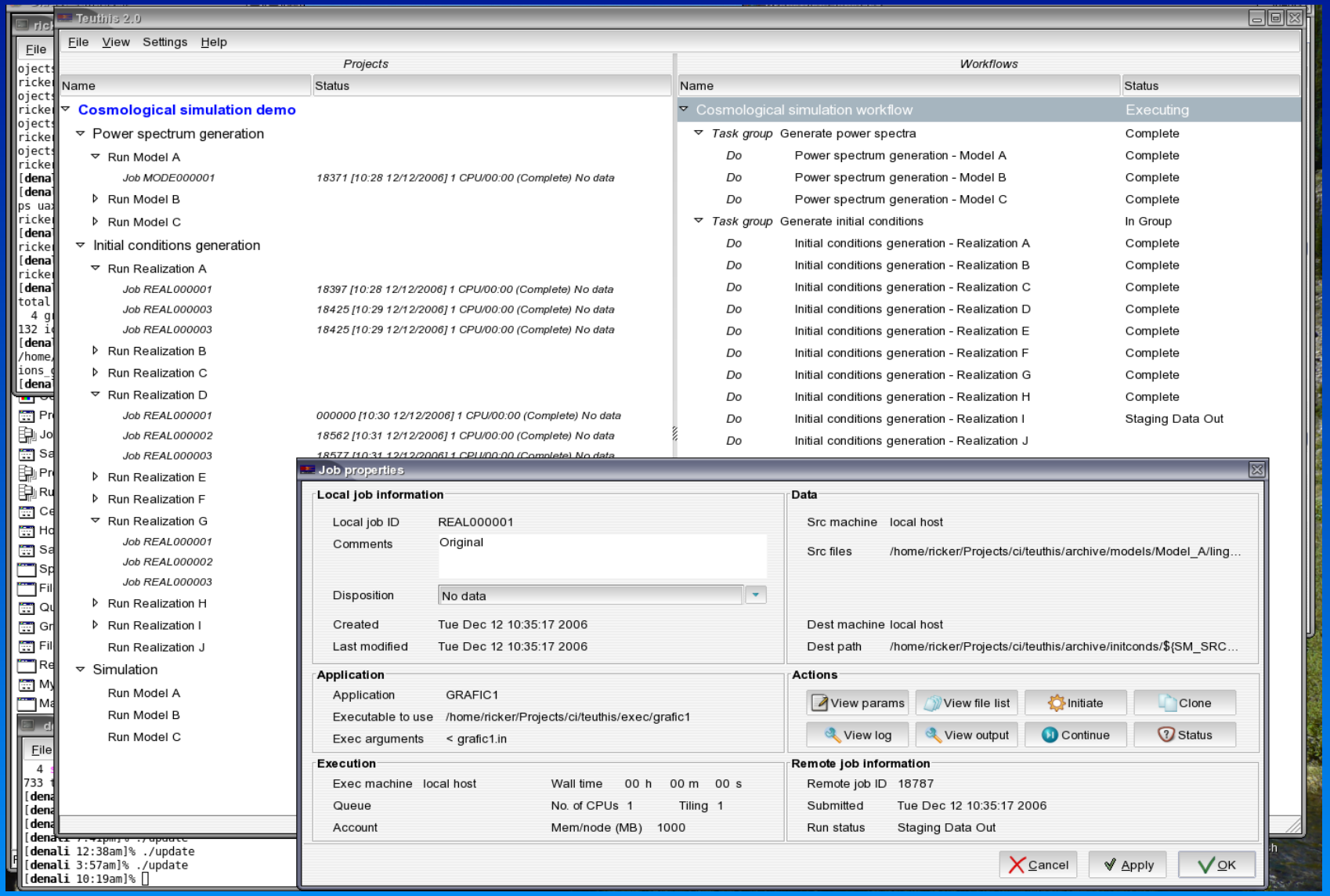
- Task links
  - Atomic workflow unit
  - Linked to a particular run
  - Invocation triggers new job: stage in – exec – stage out
  - Automatic job continuation
  - Pattern matching conditionals
  - Types – Static, Aggregate, DoNext, DoTogether, DoNTimes, DoIf, WhileDo, TaskGroup
- Pipes
  - Dynamically named sources / destinations of data



Example workflow



Example workflow – Teuthis implementation



Forthcoming work

- More sophisticated experiment designs
  - Latin hypercube, random samples
- More sophisticated workflows
  - DoTogether with multiple execution threads
  - Wait rather than fail if pipes are empty
  - GRAM job submission with Reliable File Transfer (RFT)
- Integration with other tools
  - External workflow engines: export workflow script and submit
  - Data management: archiving and tracking files
  - Initial conditions, analysis, visualization

Getting Teuthis

1.0 release available; 2.0 coming soon...

http://mazama.ncsa.uiuc.edu/projects/teuthis