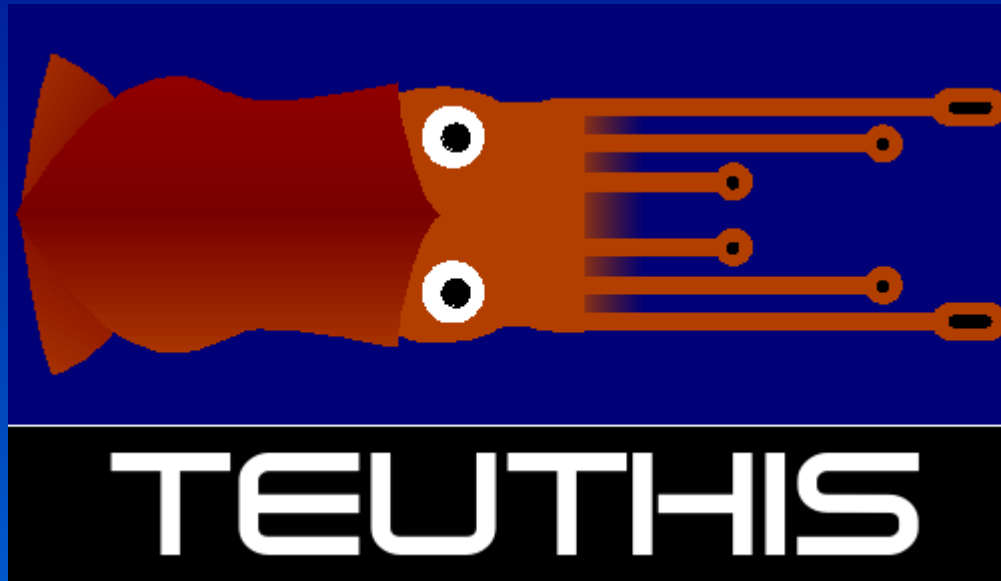


Managing Astrophysical Simulations with



Paul Ricker

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

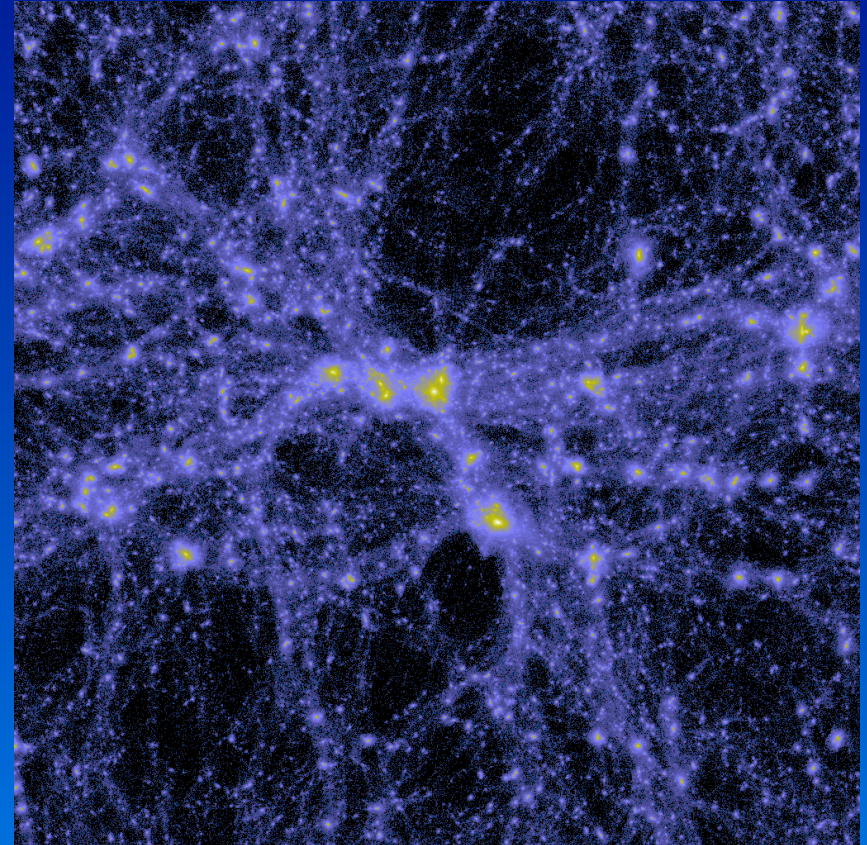
SC|05, Seattle, WA
November 2005





What do astrophysical simulations require?

- **Multiphysics**
 - Gravitational N -body problem
 - Gasdynamics
 - Reactive source terms
 - Radiation transport
- **Huge range of scales**
 - Length: typically $> 10^5$
 - Mass: typically $> 10^9$
- **HPC resources a must**
 - 100s - 1000s of CPUs
 - Datasets several TB per run
 - 10 – 1000 runs/project to realize scientific goals



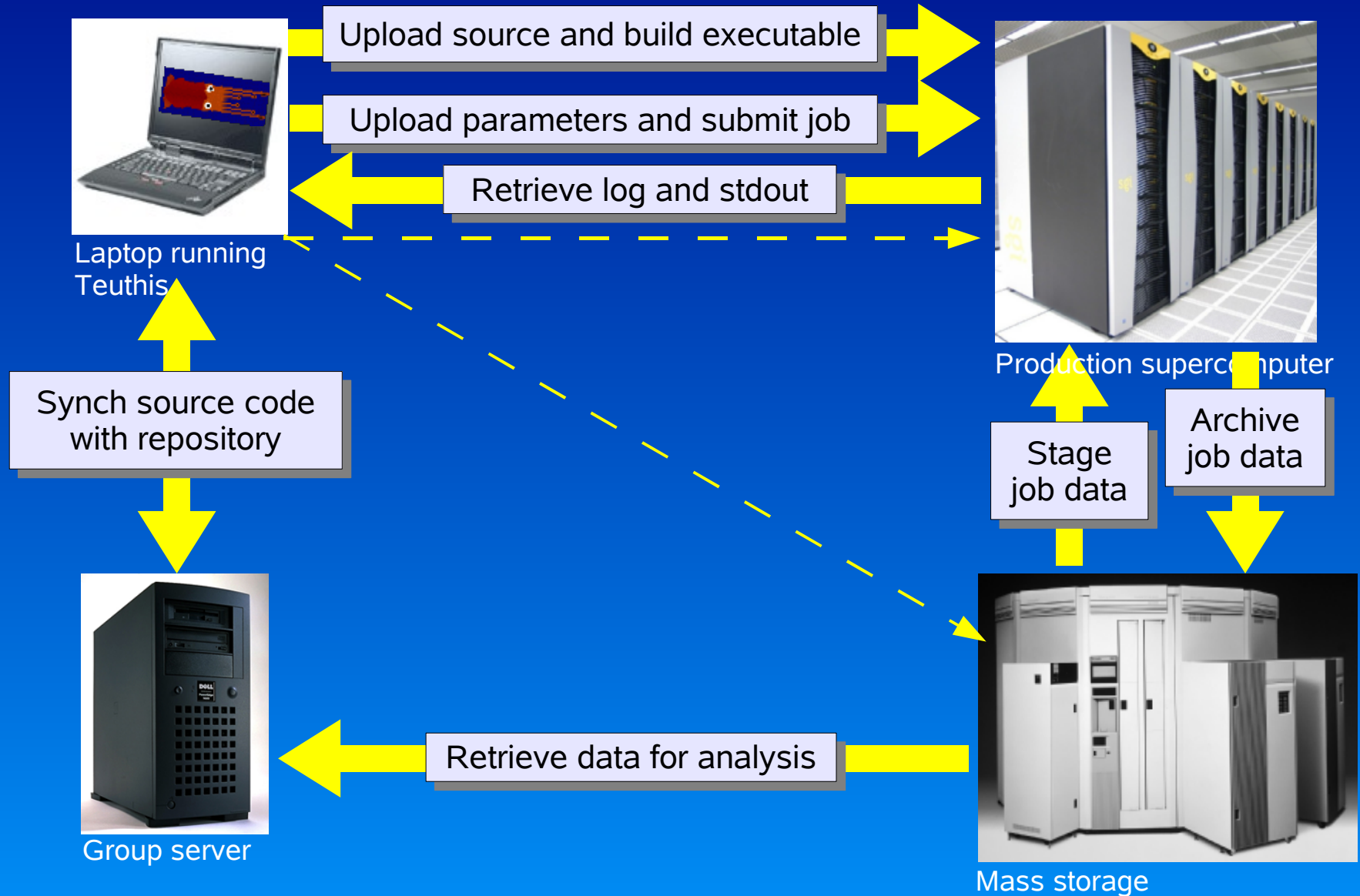


What is Teuthis?

- **A control panel**
 - Remotely configure and build applications
 - Submit and track remote jobs
 - Painlessly create parameter studies and restart jobs
- **A data manager**
 - Stage and archive data
 - Keep track of where datasets are stored
- **A notebook**
 - Organize job metadata by purpose and disposition
- **An aid to collaboration**
 - Share notebook files with collaborators



Running simulations: Teuthis approach





Resources manipulated by Teuthis

● Projects

- Collections of related scientific questions
- e.g. “Galaxy cluster scaling relations”

● Experiments

- Vary one or more parameters with a specific code to answer a particular question
- e.g. “How does the level of galaxy feedback affect the mass-temperature relation?”

● Runs

- Realizations of a single set of parameters, carried to completion
- e.g. “Run with 10x fiducial energy input”

● Jobs

- Individual attempts to complete a run
- e.g. “Job 123456 on 128 processors for 18 hours”



Resources manipulated by Teuthis

● Applications

- Scientific codes that accept text-based parameter files on input
- Execute noninteractively
- Create a log file, messages to stdout, other data files
- May or may not need to be recompiled for each experiment
- e.g., simulation code, visualization frame generator, database processing

● Machines

- Login host name
- Access method
- Queuing system
- Paths (where to build, run, etc.)



Under the hood

- **Python/PyGTK**
 - Cross-platform (including GUI)
 - Facilitates rapid prototyping
 - Modules for OS access, process control, Grid services
 - Widely used in physics/astrophysics community
- **Access methods**
 - Plain ssh – option to store password
 - ssh-agent
 - Kerberos
 - GSI authentication – local certificates or MyProxy
- **File transfer methods**
 - scp/gsiscp
 - uberFTP
 - globus-url-copy
- **Job submission methods**
 - PBS, LSF, LoadLeveler
 - Unix process control (no queue)
 - User-specified job template



Project view

Simulation Manager 1.0			
File View Settings Help			
Name	Description	Status	Date last modified
▼ FLASH testing	Testing Simulation Manager using FLASH		Tue Oct 4 17:1
▼ Basic Sedov test (local)	Test of local jobs		Tue Oct 4 19:14:5
Run A			Tue Oct 4 19:15:0
▼ Basic Sedov test (cobalt)	Test of jobs on a machine with PBS queuing and using :		Wed Oct 5 01:43:
Run A	Single run using default parameters to test job submission	Complete	Wed Oct 5 01:17:
Job A0004	Original	28909 [21:26 10/04/2005] 1 CPU/00:10 (Complete) Successful completion	Wed Oct 5 01:50:2
Job A0004	Restart of 28909	28910 [21:27 10/04/2005] 1 CPU/00:10 (Complete) Successful completion	Wed Oct 5 01:16:7
▼ Sedov scaling test (cobalt)	Test with varying number of processors		Wed Oct 5 01:43:
Run A1		Complete	Wed Oct 5 02:07:
Job A10001	Original	28924 [01:44 10/05/2005] 1 CPU/00:10 (Complete) Successful completion	Wed Oct 5 02:05:3
Run A2		Complete	Wed Oct 5 02:07:
Job A20001	Original	28925 [01:45 10/05/2005] 2 CPUs/00:10 (Complete) Successful completion	Wed Oct 5 02:05:2
Run A4		Complete	Wed Oct 5 02:07:
Job A40001	Original	28926 [01:45 10/05/2005] 4 CPUs/00:10 (Complete) Successful completion	Wed Oct 5 02:05:7
Run A8		Complete	Wed Oct 5 02:07:
Job A80001	Original	28927 [01:45 10/05/2005] 8 CPUs/00:10 (Complete) Successful completion	Wed Oct 5 02:05:0
▼ Sedov test with varying parameter (cobalt)	Test of jobs with a single varying parameter (lrefine_mi		Wed Oct 5 02:09:
Run A		Complete	Wed Oct 5 02:08:
Job A0001	Original	28928 [01:52 10/05/2005] 1 CPU/00:10 (Complete) Successful completion	Wed Oct 5 01:58:3
Job A0001 Copy	Original	29231 [15:01 10/05/2005] 1 CPU/00:10 (Complete) No data	Wed Oct 5 15:01:5
Run B		Complete	Wed Oct 5 02:08:
Job B0001	Original	28929 [01:52 10/05/2005] 1 CPU/00:10 (Complete) Successful completion	Wed Oct 5 01:58:4
Run C		Complete	Wed Oct 5 02:08:
Job C0001	Original	28930 [01:53 10/05/2005] 1 CPU/00:10 (Complete) Successful completion	Wed Oct 5 01:58:5
Run D		Complete	Wed Oct 5 02:08:
Job D0001	Original	28931 [01:53 10/05/2005] 1 CPU/00:20 (Complete) Successful completion	Wed Oct 5 01:59:0
Run E		In progress	Wed Oct 5 02:08:
Job E0001	Original	28932 [01:53 10/05/2005] 1 CPU/00:20 (Complete) Successful completion; k	Wed Oct 5 02:02:5
Run F		In progress	Wed Oct 5 02:09:
Job F0001	Original	28933 [01:53 10/05/2005] 1 CPU/00:30 (Complete) Exceeded MAXBLOCKS	Wed Oct 5 01:55:2



Configuring applications

Applications

FLASH 2.4

Application name: FLASH 2.4

Description: Astrophysical AMR hydro/N-body code

Local source path: /home/ricker/workspace/FLASH2_devel

Executable name: flash2 Configuration file: Modules

Configure command: ./setup Build command: make

Build executable Move to remote exec dir Parallel executable

Parameter file: flash.par Log file: flash.log

To restart an interrupted job:

Expect a control file named

Use a command-line argument

Copy runtime parameters from file: flash.par.restart



Configuring machines

The screenshot shows a window titled "Machines" with a list of machines on the left: "cobalt", "tungsten", and "tungsten (ssh-agent)". The "cobalt" machine is selected, and its configuration is shown in the main area. The configuration is divided into several sections:

- Basic information:** Machine name: cobalt; Description: SGI Altix at NCSA; Type: Compute, Online data, Archival data.
- Access:** Login host: cobalt.ncsa.uiuc.edu; User ID: ricker; Method: gsissh+uberftp; X.509 certificate subject: /C=US/O=National Center for Supercomp; Auth init command: grid-proxy-init -cert %K -key %k -valid %; Remote exec command: gsissh -o BatchMode=yes -o GSSAPIAut; File upload command: gsiscp -B -r -p -q "%f" %u@%h:"%r"; File download command: gsiscp -B -r -p -q %u@%h:"%r" "%f"; Third-party transfer cmd: uberftp %i "lopen %h; put %f %r".
- Job submission:** Job template: [empty]; Parallel exec: mpirun -np %n; Queuing: pbs; CPUs/node: 1; Default: 1; Mem/node (MB): 1000; Accounts: [empty]; A R Queues: standard; A R.
- Paths:** Remote build root: /u/ac/ricker/build/test; Remote exec dir: /u/ac/ricker/exec; Remote run root: /scratch/users/ricker; Remote Teuthis dir: /u/ac/ricker/teuthis; OS type: Unix.
- Actions:** New, Clone, Remove buttons.

At the bottom of the window are buttons for Cancel, Apply, and OK.



Project dialog

Project properties [X]

Project name:

Description:

Creator: Paul Ricker

Begin date: Fri Oct 7 13:26:24 2005

Last modified: Fri Oct 7 13:26:24 2005

Notes locator:

Wiki URL Edit text file

Actions:



Experiment dialog

Experiment properties

Information

Experiment name: Sedov test with file transfer

Description: Sedov test on a remote machine with two third-party file transfers.

Created: Fri Nov 4 12:50:12 2005 Modified: Sun Nov 6 16:24:29 2005

Application

Application: FLASH 2.4 Upload source

Configuration file: Browse... Upload file

Remote build dir: /u/ac/ricker/build/test/FLASH_2.4

Config command: ./setup sedov -auto Do it Output

Build command: make; mv flash2 /u/ac/ricker/exec Do it Output

Executable to use: /u/ac/ricker/exec/flash2

Exec arguments:

Data

Src machine: tungsten

Src files: /u/ac/ricker/input/la128.tar
/u/ac/ricker/input/lcdms64.tar

Dest machine: tungsten

Dest path: /u/ac/ricker/test

Execution

Exec machine: cobalt

Queue: normal Account:

of CPUs (range): 1 Tiling: 1 Mem/node (MB): 1000

Actions

Generate runs Clone experiment

Parameters

Template: /home/ricker/flash.par Browse... Edit

Parameter 1	lrefine_max	Range	1-6		Parameter 5		Range		
Parameter 2		Range			Parameter 6		Range		
Parameter 3		Range			Parameter 7		Range		
Parameter 4		Range			Parameter 8		Range		

Cancel Apply OK



Configuring and building applications

Experiment properties

Information

Experiment name: Sedov test with file transfer

Description: Sedov test on a remote machine with two third-party file transfers.

Created: Fri Nov 4 12:50:12 2005 Modified: Sun Nov 6 16:24:29 2005

Application

Application: FLASH 2.4 Upload source

Configuration file: Browse... Upload file

Remote build dir: /u/ac/ricker/build/test/FLASH_2.4

Config command: ./setup sedov -auto Do it Output

Build command: make; mv flash2 /u/ac/ricker/exec Do it Output

Executable to use: /u/ac/ricker/exec/flash2

Exec arguments:

Data

Src machine: tungsten

Src files: /u/ac/ricker/input/la128.tar
/u/ac/ricker/input/lcdms64.tar

Dest machine: tungsten

Dest path: /u/ac/ricker/test

Execution

Exec machine: cobalt

Queue: normal Account:

of CPUs (range): 1 Tiling: 1 Mem/node (MB): 1000

Actions

Generate runs Clone experiment

Parameters

Template: /home/ricker/flash.par Browse... Edit

Parameter 1: refine_max Range: 1-6	Parameter 5:
Parameter 2:	Parameter 6:
Parameter 3:	Parameter 7:
Parameter 4:	Parameter 8:

Cancel Apply OK



Generating runs

Experiment properties

Information

Experiment name: Sedov test with file transfer

Description: Sedov test on a remote machine with two third-party file transfers.

Created: Fri Nov 4 12:50:12 2005 Modified: Sun Nov 6 16:24:29 2005

Application

Application: FLASH 2.4 Upload source

Configuration file: Browse... Upload file

Remote build dir: /u/ac/ricker/build/test/FLASH_2.4

Config command: ./setup sedov -auto Do it Output

Build command: make; mv flash2 /u/ac/ricker/exec Do it Output

Executable to use: /u/ac/ricker/exec/flash2

Exec arguments:

Data

Src machine: tungsten

Src files: /u/ac/ricker/input/la128.tar
/u/ac/ricker/input/lcdms64.tar

Dest machine: tungsten

Dest path: /u/ac/ricker/test

Execution

Exec machine: cobalt

Queue: normal Account:

of CPUs (range): 1 Tiling: 1 Mem/node (MB): 1000

Actions

Generate runs Clone experiment

Parameters

Template: /home/ricker/flash.par Browse... Edit

Parameter 1	lrefine_max	Range	1-6	<input type="checkbox"/>	Parameter 5		Range		<input type="checkbox"/>
Parameter 2		Range		<input type="checkbox"/>	Parameter 6		Range		<input type="checkbox"/>
Parameter 3		Range		<input type="checkbox"/>	Parameter 7		Range		<input type="checkbox"/>
Parameter 4		Range		<input type="checkbox"/>	Parameter 8		Range		<input type="checkbox"/>

Cancel Apply OK



Submitting jobs

The screenshot displays a graphical user interface for job submission. On the left, a 'Job properties' window is partially visible, showing fields for 'Local job information' (Local job ID, Comments, Disposition, Created, Last modified), 'Application' (Application, Executable to use, Exec arguments), and 'Execution' (Exec machine, Queue, No. of CPUs). The 'Execution' section is highlighted in yellow.

The main window, titled 'Review and edit job script for submission', contains a text area with the following job script:

```
#!/bin/sh
#PBS -l walltime=00:10:00
#PBS -l mem=2000mb
#PBS -l ncpus=1
#PBS -q debug
#PBS -V
#PBS -A lhl
#PBS -N A0002_Copy
mkdir -p /scratch/users/ricker/New_Project/Sedov_test_-cobalt-/A/A0002_Copy
cd /scratch/users/ricker/New_Project/Sedov_test_-cobalt-/A/A0002_Copy
mpirun -np 1 /u/ac/ricker/exec/flash2 |
```

At the bottom of the window are 'Submit' and 'Cancel' buttons. To the right, another window is partially visible with 'Clone job', 'Archive data', and 'OK' buttons.



Transferring data

Job properties

Local job information

Local job ID: C0001
Comments: Original
Disposition: No data
Created: Thu Nov 10 03:06:53 2005
Last modified: Thu Nov 10 03:06:53 2005

Application

Application: FLASH 2.4
Executable to use: /u/ac/ricker/exec/flash2
Exec arguments:

Execution

Exec machine: cobalt Wall time: 00 h 00 m 00 s
Queue: normal Account:
No. of CPUs: 1 Tiling: 1 Mem/node (MB): 1000

Data

Src machine: tungsten
Src files: /u/ac/ricker/input/la128.tar
 /u/ac/ricker/input/lcdms64.tar
Dest machine: tungsten
Dest path: /u/ac/ricker/test

Actions

View parameters Submit Status Clone job
View log file View output Continue job Archive data

Remote job information

Remote job ID: 000000
Submitted: Not yet submitted
Run status: Unsubmitted

Cancel Apply OK



Checking status

Job properties

Local job information

Local job ID: C0001
Comments: Original
Disposition: No data
Created: Thu Nov 10 03:06:53 2005
Last modified: Thu Nov 10 03:06:53 2005

Application

Application: FLASH 2.4
Executable to use: /u/ac/ricker/exec/flash2
Exec arguments:

Execution

Exec machine: cobalt Wall time: 00 h 00 m 00 s
Queue: normal Account:
No. of CPUs: 1 Tiling: 1 Mem/node (MB): 1000

Data

Src machine: tungsten
Src files: /u/ac/ricker/input/la128.tar
 /u/ac/ricker/input/lcdms64.tar
Dest machine: tungsten
Dest path: /u/ac/ricker/test

Actions

View parameters Submit Status Clone job
View log file View output Continue job Archive data

Remote job information

Remote job ID: 000000
Submitted: Not yet submitted
Run status: Unsubmitted

Cancel Apply OK



Viewing log files and output files

The screenshot displays three overlapping windows from a job management interface:

- Job properties:** A sidebar on the left with sections for 'Local job information', 'Application', and 'Execution'. Fields include 'Local job ID', 'Comments', 'Disposition', 'Created', 'Last modified', 'Application', 'Executable to use', 'Exec arguments', 'Exec machine', 'Queue', and 'No. of CPUs'.
- Output for job A0002:** A window showing the start of a log file with the text: '==== Standard output =====', '-----', '!Begin PB', 'Job ID:', 'Username:', 'Group:', 'Creating B', 'Warning: n', 'Thus no jo', '[AUTOPE', '[AUTOPE', 'master1', 'flash: init', 'p_ambier', 'rho_ambi', 'gamma', 'exp_ener', 'r_init =', 'p_exp', 'xctr =', 'yctr =', 'zctr =', 'ndim', 'nsubzone'.
- Log file for job A0002:** A window showing the full log file content:

```
FLASH log file: 11-03-2005 21:24:08 Run number: 1
-----
Number of processors:      1
Dimensionality:           2
Max Number of Blocks/Proc: 1000
Number x zones:           8
Number y zones:           8
Number z zones:           1
Setup stamp:   Wed Nov 2 20:30:32 2005
Build stamp:   Wed Nov 2 20:37:15 CST 2005
System info:   Linux co-login1.ncsa.uiuc.edu 2.4.21-sgi306rp21 ia64
Version:       FLASH 2.4.20040921
Build directory: /u/ac/ricker/build/test/FLASH_2.4/object
Setup syntax:  ./setup.py sedov
f compiler flags:
ifort -c -r8 -i4 -ftz -fpp -O3 -ip -assume byterecl -DN_DIM=2 -DMAXBLOCKS=1000
-DNXB=8 -DNYB=8 -DNZB=1
c compiler flags:
icc -I /usr/apps/hdf5/include -c -O3 -ip -DN_DIM=2 -DMAXBLOCKS=1000 -DN
YB=8 -DNZB=1
loader flags:  -O3 -ip -o
-----
Comment: Sedov explosion
-----
FLASH Modules used:
```



Continuing a job

Job properties

Local job information

Local job ID: C0001
Comments: Original
Disposition: No data
Created: Thu Nov 10 03:06:53 2005
Last modified: Thu Nov 10 03:06:53 2005

Application

Application: FLASH 2.4
Executable to use: /u/ac/ricker/exec/flash2
Exec arguments:

Execution

Exec machine: cobalt
Queue: normal
Account:
No. of CPUs: 1
Tiling: 1
Mem/node (MB): 1000
Wall time: 00 h 00 m 00 s

Data

Src machine: tungsten
Src files: /u/ac/ricker/input/la128.tar
/u/ac/ricker/input/lcdms64.tar
Dest machine: tungsten
Dest path: /u/ac/ricker/test

Actions

View parameters | Submit | Status | Clone job
View log file | View output | **Continue job** | Archive data

Remote job information

Remote job ID: 000000
Submitted: Not yet submitted
Run status: Unsubmitted

Cancel | Apply | OK



Future plans

- **Data**
 - Background file transfers/reliable file transfer (RFT)
 - Data replica management
- **Job submission and monitoring**
 - GRAM job submission
 - Use of OGRE to handle staging/run/archiving/analysis
- **Metadata and analysis**
 - Link initial conditions – simulation – analysis jobs
 - Link runs to multiple experiments
 - Plugins for external visualization/analysis tools
- **Collaborative aspects**
 - Grid portal version



Getting Teuthis

1.0 beta release now available at

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

TEUTHIS

About

- Documentation
- Download
- Support
- Presentations
- Related projects
- CI home
- Internal pages

Welcome to Teuthis!

Teuthis is a tool intended to improve the efficiency with which computational scientists make use of computing resources, particularly high-performance computers. It is designed especially for the needs of astrophysical simulations, but any computational task that takes a set of input parameters from a file and runs noninteractively can be managed using Teuthis.

With Teuthis you can:

- Remotely configure and build applications from local source code
- Submit and track jobs on remote computing resources
- Painlessly schedule and track multiple restart jobs
- Stage and archive data on different machines
- Create large parameter studies with a few simple operations
- Organize job metadata by purpose and disposition
- Share calculation records with collaborators

Teuthis Job

Name	Descr
FLASH testing	
Sedov test with file transfer Sedo	
Run A	
Job A0001	Orig
Sedov scaling test	
Run A1	
Job A1001	Orig
Job A1001 Copy	Orig
Run A2	
Job A2001	Orig
Run A4	