test harness and reporting framework

Shava Smallen
San Diego Supercomputer Center

Grid Performance Workshop
6/22/05
Is the Grid Up?

- Can user X run application[s] Y on Grid[s] Z? Access dataset[s] N?
  - Are Grid services the application[s] use available? Compatible versions?
  - Are dataset[s] N accessible to user X? Credentials?
  - Is there sufficient space to store output data?
  - ...
- Community of users (VO)?
- Multiple communities of users?
Testing a Grid

• If you can define “Grid up” in a machine-readable format, you can test it
• User documentation, users, mgmt

Grid up example

Develop and optimize code at Caltech

Run large job at SDSC, store data using SRB.

Run large job at NCSA, move data from SRB to local scratch and store results in SRB

Run larger job using both SDSC and PSC systems together, move data from SRB to local scratch storing results in SRB

Move small output set from SRB to ANL cluster, do visualization experiments, render small sample, store results in SRB

Move large output data set from SRB to remote-access storage cache at SDSC, render using ANL hardware, store results in SRB
What type of testing?

- Deployment testing
  - Automated, continuous checking of Grid services, software, and environment
  - E.g., gatekeeper ping or scaled down application

http://inca.sdsc.edu
Who tests?

- Grid/VO Management
  - Run from default user account
  - Goal: user-level problems detected & fixed before users notice
  - Results available to users
- User-specific
  - Debug user account/environment issues
  - Advanced usage: feedback tests

http://inca.sdsc.edu
Inca

- Framework for the automated testing, benchmarking and monitoring of Grid systems
  - Schedule execution of information gathering scripts (reporters)
  - Collect, archive, publish, and display results

http://inca.sdsc.edu
Outline

✓ Introduction
  • Inca architecture
  • Case study: V&V on TeraGrid
  • Current and Future Work
  • Feedback
Inca Reporters

- Script or executable that outputs XML conforming to Inca specification
- Context of execution is required - important for repeatability
  - What commands were run?
  - What machine?
  - What inputs?
- Communicate more than pass/fail
  - Body XML can be reporter specific - flexibility
  - E.g., package version info (software stack availability)
  - E.g., SRB throughput (unusual drop in SRB performance)
- Users can run it independently of framework

http://inca.sdsc.edu
Reported Execution Framework

- How often should reporters run
  - Boot-time, every hour, every day?
- Modes of execution:
  - One shot mode:
    - boot-time, after a maintenance cycle, user checking their specific setup
  - Continuous mode: cron scheduling
- Data can be queried from a web service and displayed in a web page

http://inca.sdsc.edu
Outline

✓ Introduction
✓ Inca architecture
  • Case study: V&V on TeraGrid
  • Current and Future Work
  • Feedback

http://inca.sdsc.edu
TeraGrid

• **TeraGrid** - an “enabling cyberinfrastructure” for scientific research
  - ANL, Caltech, Indiana Univ., NCSA, ORNL, PSC, Purdue Univ., SDSC, TACC
  - 40+ TF, 1+ PB, 40Gb/s net

• **Common TeraGrid Software & Services**
  - Common user environment across heterogeneous resources
  - TeraGrid VO service agreement

[Map of TeraGrid network areas]

http://inca.sdsc.edu
Validation & Verification

- Common software stack:
  - **20 core packages**: Globus, SRB, Condor-G, MPICH-G2, OpenSSH, SoftEnv, etc.
  - **9 viz package/builds**: Chromium, ImageMagick, Mesa, VTK, NetPBM, etc.
  - **21 IA-64/Intel/Linux packages**: glibc, GPFS, PVFS, OpenPBS, intel compilers, etc.

  - **50 version reporters**: compatible versions of SW
  - **123 tests/resource**: package functionality

- **Services**: Globus GRAM, GridFTP, MDS, SRB, DB2, MyProxy, OpenSSH
- **Cross-site**: Globus GRAM, GridFTP, OpenSSH

http://inca.sdsc.edu
Validation & Verification (cont.)

• Common user environment
  - $TG_CLUSTERSCRATCH,
    $TG_APPS_PREFIX, etc.
  - SoftEnv configuration - manipulate user environment

→ Verify environment vars defined in default environment

→ Verify Softenv keys defined consistently across sites

http://inca.sdsc.edu
Inca deployment on TeraGrid

- 9 sites/16 resources
- Run under user account inca
Detailed Status Views

Click on the column headers in each table for more information.

The following table shows the environment variables that should be defined in the environment:

<table>
<thead>
<tr>
<th>Variables</th>
<th>resource1</th>
<th>resource2</th>
<th>resource3</th>
<th>resource4</th>
<th>resource5</th>
<th>resource6</th>
<th>resource7</th>
<th>resource8</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLOBUS_LOCATION</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>GLOBUS_PATH</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>HDF4_HOME</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>HDF5_HOME</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>MYPROXY_SERVER</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>SASL_PATH</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>SRB_HOME</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_APPS_PREFIX</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_CLUSTER_GPFS</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_CLUSTER_HOME</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_CLUSTER_PFS</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_CLUSTER_PVFS</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_CLUSTER_SCRATCH</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_COMMUNITY</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>TG_EXAMPLES</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>HDF5-1.6.2-r1</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>unavailable</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>HDF5-1.6.2-r2</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
<tr>
<td>Irix 500</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
<td>available</td>
</tr>
</tbody>
</table>
Drill-down capability

<table>
<thead>
<tr>
<th>Reporter details:</th>
</tr>
</thead>
<tbody>
<tr>
<td>reporter name</td>
</tr>
<tr>
<td>description</td>
</tr>
<tr>
<td>version</td>
</tr>
<tr>
<td>status</td>
</tr>
<tr>
<td>url</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Execution information:</th>
</tr>
</thead>
<tbody>
<tr>
<td>inputs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ran at (GMT)</td>
</tr>
<tr>
<td>age</td>
</tr>
<tr>
<td>runs every</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reporter system command log:</th>
</tr>
</thead>
<tbody>
<tr>
<td>The following are the &quot;system&quot; commands executed by the reporter. Note that the reporter may execute other actions in between system commands (e.g., change directories). Please click the on reporter name above for the full reporter code.</td>
</tr>
<tr>
<td>% globusrun -a -r test_hostname 2&gt;&amp;1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Host information:</th>
</tr>
</thead>
<tbody>
<tr>
<td>hostname</td>
</tr>
<tr>
<td>ipaddr</td>
</tr>
<tr>
<td>uname</td>
</tr>
</tbody>
</table>
Summary Status

Summary of Common TeraGrid Software and Services 2.0
Page generated by Inera: 07/13/04 18:39 CDT

This page offers a summary of results for critical grid, development, and cluster tests (view list of tests). Details about a resource's test results are available by clicking on the resource name in the "Site-Resource" column of the table.

<table>
<thead>
<tr>
<th>Site-Resource</th>
<th>Grid</th>
<th>Development</th>
<th>Cluster</th>
<th>Total Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>site1-resource1</td>
<td>Pass: 32 Fail: 1</td>
<td>Pass: 23 Fail: 0</td>
<td>Pass: 1 Fail: 1</td>
<td>Pass: 56 Fail: 2</td>
</tr>
<tr>
<td></td>
<td>96% passed</td>
<td>100% passed</td>
<td>50% passed</td>
<td>96% passed</td>
</tr>
<tr>
<td></td>
<td>75% passed</td>
<td>100% passed</td>
<td>50% passed</td>
<td>85% passed</td>
</tr>
<tr>
<td>site2-resource1</td>
<td>Pass: 1 Fail: 18</td>
<td>Pass: 2 Fail: 10</td>
<td>n/a</td>
<td>Pass: 3 Fail: 28</td>
</tr>
<tr>
<td></td>
<td>5% passed</td>
<td>16% passed</td>
<td>9% passed</td>
<td></td>
</tr>
</tbody>
</table>

Expanded View of Errors

**site1-resource1**

**Grid**

1. globus-2.4.3-intel-r3: failed: duroc_mpi_helloworld_to_jobmanager-pbs test

Key

- Green: All tests passed: 100%
- Red: One or more tests failed: < 100%
- Gray: Tests not applicable to machine or have not yet been ported

History of percentage of tests passed in “Grid” category for a 6 month period
Measuring TeraGrid Performance

- GRASP (Grid Assessment Probes)
  - test and measure performance of basic grid functions
- Pathload [Dovrolis et al]
  - measures dynamic available bandwidth
  - uses efficient and lightweight probes
Lessons learned

• Initially focused on system administrative view

• Moving towards user-centric view
  ▪ File transfer functionality and performance
  ▪ File system availability
  ▪ Job submission
  ▪ SRB performance
  ▪ Interconnect bandwidth
  ▪ Applications: NAMD, AWM
Integration with Knowledge Base

Are you having problem(s) with:

• Data
• Job Management
• Security
…

YES:  Are you having trouble transferring a file?
YES:  Are you seeing poor performance?
…

1. Check to see if you have valid proxy…

grid.middleware.globus.unit.proxy

Reporter passed.

Reporter details:

reporter  grid.middleware.globus.unit.proxy  (click on name, reporter name to view reporter script)
description Verifies that user has valid proxy; attempts to create if not
version 1.5
status production
url http://www.globus.org/security/proxy.html

Execution information:

inputs help no
      log 3
      verbose 1

run at Wed Jun 15 23:31:56 2005 (GMT)

Reporter system command log:

The following are the *system* commands executed by the reporter. Note that the reporter may execute other actions in between system commands (e.g., change directories). Please click the on reporter name above for the full reporter code.

% grid-proxy-info -timeleft 2>61

Host information:

hostname tg-login1.sdsc.teragrid.org
ipaddr 198.202.112.33
uname Linux tg-login1 2.4.21.SuSE_286.6ef2 #1 SMP Wed May 4 09:24:24 CDT 2005 ia64 unknown
Outline

- Introduction
- Inca architecture
- Case study: V&V on TeraGrid
  - Current and Future Work
  - Feedback

http://inca.sdsc.edu
Inca Today

• Software available at: http://inca.sdsc.edu
• Current version: 0.10.3
• Also available in NMI R7
• Users:

- TERAGRID
- GEON
- DEISA
- NGS
Inca 2.0

- Initial version of Inca focused on basic functionality
- New features:
  - Improved storage & archiving capabilities
  - Scalability - control and data storage
  - Usability - improved installation and configuration control
  - Performance - self-monitoring
  - Security - SSL, proxy delegation
  - Condor integration
- Release in 3-6 months
### Error:

call to globus-url-copy failed: error: the server sent an error response: 425 425 Can't open data connection.  
data_connect_failed() failed: an authentication operation failed.

### This error has been detected:

- 7 times in the past week
- 20 in the past month

[view graph of past error occurrences]

### Submit information or possible solution for this error:

[search for information submitted for an error or reporter]

### Reporter details:

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporter name</td>
<td>data.transfer.gridftp.unit.copy (click view reporter script)</td>
</tr>
<tr>
<td>Description</td>
<td>This test verifies the globus-url-copy and destination. If the source file does not exist, a small test file.</td>
</tr>
<tr>
<td>Version</td>
<td>1.9</td>
</tr>
<tr>
<td>Status</td>
<td>production</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://www.ncsa.uiuc.edu/People/jbasney/teragrid-setup-test.html">http://www.ncsa.uiuc.edu/People/jbasney/teragrid-setup-test.html</a></td>
</tr>
</tbody>
</table>
View Resource Usage

**Execution information:**

```plaintext
inputs
  verbose 1
  dstURL gsiftp://resource1.teragrid.org/~inca/gridftp.test
  srcURL file:/tmp/inca/gridftp.test
  help no
  log 3

ran at (GMT)  Thu Jun 16 06:34:03 2005
run time (seconds)  5.24  [graph run time history]
CPU time (seconds)  3.12  [graph CPU time history]
memory used (MB)  5  [graph memory used history]
age  16 hours 7 mins
runs every  24 hour(s)
```

**Reporter system command log:**

The following are the *system* commands executed by the reporter. Note that the reporter may execute other actions in between system commands (e.g., change directories). Please click the on reporter name above for the full reporter code.

```plaintext
% globus-url-copy file:/tmp/inca/gridftp.test
gsiftp://resource1.teragrid.org/~inca/gridftp.test 2>1
```

**Host information:**

```plaintext
hostname  ran.on.hostname
ipaddr  192.000.00.000
uname  Linux unanumenun #2 SMP Fri Jun 3 11:44:48 EST 2005 i686
       i686 i386 GNU/Linux
```
Summary

- Inca is a framework that provides automated testing, benchmarking, and monitoring
  - Grid-level execution to detect problems and report to system administrators
  - Users can view status pages and compare to problems they see
  - Users can run reporters as themselves to debug account/environment problems
- Currently in-use for TeraGrid V&V, GEON, and others

http://inca.sdsc.edu
Outline

✓ Introduction
✓ Inca architecture
✓ Case study: V&V on TeraGrid
✓ Current and Future Work
  • Feedback

http://inca.sdsc.edu
Feedback

• How are you monitoring your Grid infrastructure?

• What do you need to test?

• What diagnostic/debugging tools are available to users?

• Displaying test results to users? In what format? How much detail?

http://inca.sdsc.edu
More Information

http://inca.sdsc.edu

- Current Inca version: 0.10.3
- New version in 3-6 months
- Email: ssmallen@sdsc.edu