Inca Control Infrastructure

Shava Smallen
ssmallen@sdsc.edu

Inca 2.0 Workshop
February 23, 2006

Control infrastructure goals

• Minimal impact on monitored resources

• Flexible reporter scheduling and configuration options

• Easy installation and maintenance

• Valid proxy credential available to reporters
**Architecture overview**

- incat
- Reporter
- Reporter Repository
- Data Consumers
- Reporter Agent
- Suite
- Depot
- Reporter Manager
- Resource 1
- Resource 2
- Resource N
- Reporter Managers
- Reporter Agents
- Reporters
- Agents
- Data Consumers

**Reporter Agent**

- Implements configuration specified by Inca administrator
- Stages and launches reporter managers (local, SSH, Globus)
- Handles reporter distribution and updates
- Distributes reporter configuration and schedules
- Sends configuration to Inca depot
Reporter Manager

- Previously known as the “controller” in v1

- Receives from reporter agent that started it:
  - Reporters and libraries
  - Reporter configuration and schedules

- Executes reporters periodically (cron) or on-demand and forwards reports to the depot

- Profiles reporter system usage and enforces timeouts

Deployment configuration

- Reporter repositories (URLs)

- Resources

- Report series and suites
Resource specification

- **Resource** - cluster, supercomputer
  - Reporter manager executes on login node

- **Resource group** - two or more resources
  - Shared characteristic (e.g., ia64 arch)
  - Site
  - VO

---

Sample Resources

- **TeraGrid**
  - **SDSC**
    - **Blue Gene**
    - **IA-64 Cluster**
  - **NCSA**
    - **IA-64 Cluster**

---

Resource macros

- Resource or resource group attributes/variables
- Specify values in one place & re-use
- Specify resources differences
- Most specific value wins
- Can contain multiple values

- **TeraGrid**
  - **myproxyHost** = mp.teragrid.org

- **DataStar**
  - gatekeeper = dslogin.sdsc.edu
  - nativeCC = xlc
  - charged = SDSC-02

- **NCSA IA-64 Cluster**
  - gatekeeper = tg-login.ncsa.edu
  - nativeCC = icc
  - charged = NCSA-24
Resource access methods

- **Local** - via Java Runtime exec
  - *Required information:* installation directory

- **SSH** - via SSHTool’s Java SSH API
  - Currently assumes passwordless key
  - Encrypted key and agent authentication to be added
  - *Required information:* installation directory, ssh server, remote username
  - *Optional information:* ssh server port, public key

- **Globus** - via Java CoG 1.2
  - Support Globus 2.4.3 servers; support for Globus 4 to be added
  - *Required information:* installation directory, GRAM server, GridFTP server
  - *Optional information:* GRAM port, GridFTP port, GRAM DN, GridFTP DN, GASS port

---

“Single” resource option

- A set of hosts are treated as a single resource (e.g., round-robin DNS)
- Inca will pick one host to execute reporters on (soon)
- If host goes down, will try to restart on another host (soon)
- Dependent on access method
**Proxy support**

- Proxy retrieved from MyProxy server to:
  1. Launch Reporter Manager (Globus access method)
  2. Maintain proxy credential on reporter managers (SSL only) - Independent of access method

- **Required Information**: MyProxy hostname, username, password

- **Optional information**: MyProxy server DN, credential lifetime

---

**Deployment configuration**

1. Reporter repositories (URLs)
2. Resources
3. Report series and suites
What is a report series?

A set of reports collected at different points in time by executing a reporter with a set of arguments in a context on a particular host.

Specifying reporter arguments

- E.g., wget reporter (from hands-on session)

% net.benchmark.wget \
   -page=http://www.cnn.com \
   -version=no \ 
   -help=no \ 
   -verbose=1 \ 
   -log=3 \
**Specifying reporter execution context**

- Optional execution string can be used to set the context the reporter runs under.

  - E.g., run reporter under fresh shell:
    
    ```bash
    /bin/sh -l -c '@EXEC@'
    
    => /bin/sh -l -c 'net.benchmark.wget -page=…'
    ``

- E.g., softenv/modules configuration:
  
  ```bash
  soft add +atlas; @EXEC@
  
  => soft add +atlas; cluster.math.atlas.version …
  ```

**Specifying resources**

- Use a single resource name or resource group.

  - E.g.,
    
    - `tg-login.sdsc.edu`
    - SDSC
    - TeraGrid
    - IA-64
Using resource macros in arguments

TeraGrid

grid.globus.unit \ -host=@gatekeeper@

Resource’s macro value substituted at execution time

Blue Gene

grid.globus.unit \ -host=bglogin.sdsc.edu

NCSA IA-64 Cluster

grid.globus.unit \ -host=tg-login.ncsa.edu

Using resource macros in arguments (cont.)

NCSA IA-64 Cluster

grid.performance.ping \ -host=@hosts@

Reporter will be executed once for each value in macro.

hosts = tg-login.sdsc.edu, tg-login.uc.edu, tg-login.psc.edu

NCSA IA-64 Cluster

grid.performance.ping \ -host=tg-login.sdsc.edu

grid.performance.ping \ -host=tg-login.uc.edu

grid.performance.ping \ -host=tg-login.psc.edu
Using resource macros in arguments (cont.)

- Multiple multi-valued macros ⇒ cross product
  - E.g.,

        data.transfer.unit -host=@gridftpServers@ -dir=@dirs@
    • @gridftpServers@ = bglogin.sdsc.edu, tg.ncsa.edu
    • @dirs@ = /gpfs/inca, /users/inca, /scr/inca

  ⇒ Will expand to:

1. data.transfer.unit -host=bglogin.sdsc.edu -dir=/gpfs/inca
2. data.transfer.unit -host=bglogin.sdsc.edu -dir=/users/inca
3. data.transfer.unit -host=bglogin.sdsc.edu -dir=/scr/inca
4. data.transfer.unit -host=tg.ncsa.edu -dir=/gpfs/inca
5. data.transfer.unit -host=tg.ncsa.edu -dir=/users/inca
6. data.transfer.unit -host=tg.ncsa.edu -dir=/scr/inca

Limit resource usage of reporter

- Wall clock time
  - E.g., no more than 10 seconds

- Cpu seconds
  - E.g., no more than 2 cpu seconds

- Memory
  - E.g., no more than 20 MB

- Reporter will be killed and an error report will be sent indicating the resource usage exceeded
**Reporter scheduling options**

- Periodic execution
  - Cron syntax (supports Vixie syntax)
    - `45 * * * *` - on the 45th minute every hour
    - `*/10 * * * *` - every 10 minutes
    - `2-59/10 * * * *` - every 10 minutes starting at 2

- Immediate execution
  - Trying out new reporters
  - Debugging reporter

- Other reporter execution methods to be added
  - Queue interaction
  - Coordinated execution - e.g., network probes

**Specifying a report series nickname**

- Convenience for querying and data display
- Can contain macros
- E.g., `atlas_version`
- E.g., `gridftp_test_to_@site@`
What is a suite?

- A set of report series configurations that share a common theme
  - E.g., GITS
  - E.g., CTSS Version 3
  - E.g., GRASP

- Suites can be shared across Inca deployments (just fill in the macro values)

Summary

- Staging, scheduling, and execution of reporters

- Receives commands from Inca administrator via incat (next talk)

- Deployment configuration options
**Future Work**

- **Pre-release**
  - Improved fault tolerance - e.g., automatic reporter manager restart
  - Apply resource changes to already running reporters

- **Post-release**
  - Other reporter execution methods
  - Suites
  - Profiling of Inca components
  - Reporter version selection