Advanced scripting for HPC resource management

C. Onime

Overview

- Introduction
- O.S. high-level scheduler
- HPC shell scripts Case studies
  - Scheduler failures
  - Single Logon
  - Distributing configuration files
  - Batch Node management
  - UPS Management

High-level scheduler - cron

- O.S. Level time based scheduler
- Linux implementations
  - Time based (atd, cron)
  - Non-time based (anacron)
- Interfaces
  - Scripts & executables
    - /etc/cron.hourly/
    - /etc/cron.daily/
    - /etc/cron.weekly/
    - /etc/cron.monthly/
  - Special syntax
    - /etc/crontab
    - /etc/cron.d/
    - Per user editor (crontab –e)

System-wide crontab

- Syntax

  * * * * * user application/tool to run

- Fields
  - Minute (0-59)
  - Hour (0-23)
  - Day of month (1-31)
  - Month (1-12)
  - Day of week (0-7)

Case study 1 – Scheduler failures

- Situation:
  - Avoiding job scheduler lockups for heterogeneous (multi-queue) clusters. Maui is running, torque is running, jobs are not starting. Maui thinks nodes are over-committed.
- Discussion
  - Share your ideas and thoughts?

Case study 1 – Scheduler failures

- Solution
  - Cron based shell script
    - Test maui's idea of cluster resources every 5 minutes
    - Restart maui and/or tourque, if it is not what is expected.

Case study 2 – Single Logon

- Situation
  - Authenticating and authorizing HPC users from an external source.
  - Only important for Logon nodes (or master-node).
- Discussion
  - Share your thoughts and ideas?

Case study 2 – Single Logon

- Solution
  - HPC resources only available to users in special Group.
  - Authentication
    - Integrated via PAM (UNIX+Kerberos)
  - Periodic tasks (Crontab – 15 minutes)
    - Query external source for members of special group
    - Add new-users to local password database, remove old users
    - Distribute to cluster nodes
      - What if node was down, while updating ?

Case study 3 – Distributing configuration files

- Situation
  - Consistent configuration files across cluster: user-list, fstab, autofs. Typical problem some nodes crash on jobs for new users..
- Discussion
  - What tools can be used for this?

Case study 3 – Distributing configuration files

- Solution
  - Tools could be C3tools or rsync or even scp
  - One-way periodic push (Master –> nodes)
  - What of speed?
  - Is it really scalable?

Case study 4 – Node management

- Situation
  - Automating solutions to common problems affecting batch-queue nodes including disk full, kernel panics, memory leaks
- Discussion
  - Is it important?
  - How?

Case study 4 – Node management

- Solution
  - Detect problem
    - Log-file analysis (Generic script, central logging)
      - What of outdated info?
    - Loop output of pbsnodes (generic script parsing info)
      - Reliability & Scalability?
    - Individual Node Health check
      - Built-into batch system
  - Tasks
    - Take node offline, complete current job, reboot node and put back on-line.

Case study 5 – UPS Management

- Situation
  - Automating response to power-outages. Single UPS for all cluster devices is monitored via network or USB cable and no-backup power-system
- Discussion
  - Ideas and thoughts?
  - Is it a waste of our time?

Case study 5 – UPS management

- Solution
  - Central solution
    - Monitoring by single machine and management decisions applied to all nodes
  - Independent/individual monitoring
  - What of nodes running jobs?

That’s all