glexec/SCAS pilot service

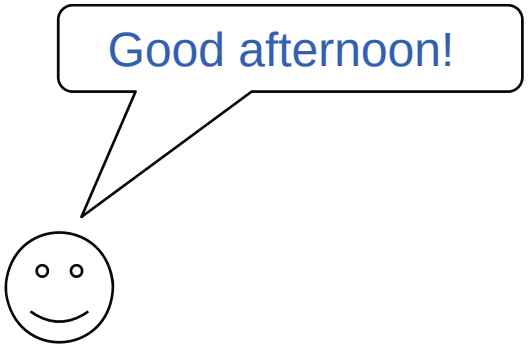
Status and short-term plans

Antonio Retico, Gianni Pucciani

GDB 11-Mar-09 - CERN

Agenda

- **Description of the glexec/SCAS pilot**
  - Subject
  - Objectives
  - Partners
  - Success conditions
  - Access info
- **Recent history**
  - Software versions
  - Deployment
  - Integration works
  - Stress testing
- **Next Steps**



Agenda

- **Description of the glexec/SCAS pilot**
  - Subject
  - Objectives
  - Partners
  - Success conditions
  - Access info

Subject

- **Set-up worker nodes enabling glexec/SCAS**
  - Production nodes
  - lcg-CE
  - PROD BDII
- **Versions**
  - Starting from version in gLite PPS PPS Update 43
  - Newer versions deployed if required (both on the pilot and in certification)

Objectives

- **Integration with job management frameworks PanDA (ATLAS), Dirac3 (LHCb), AliEn (Alice)**
  - Basic use case (identity switch) in real production conditions
  - New error codes (“internal failure” vs. “user non-authorised”)
- **Operability and scalability**
  - Set of SA3 scalability tests of SCAS
  - Resilience of glexec to short-duration failures in the backend
  - Configuration of SCAS in load-balancing

Partners

- **Coordination: Antonio Retico, Gianni Pucciani (CERN)**
- **JRA1: Oscar Koeroo (NIKHEF)**
  - Development, support
- **SA3: Gianni Pucciani (CERN)**
  - Stress testing
- **SA1: Angela Poschlad (FZK), Pierre Girard (IN2P3), Ronald Starink (NIKHEF)**
  - Site installations
- **Atlas: Jose Caballero, Maxim Potekhin (BNL)**
  - PanDA Integration
- **LHCb: Andrei Tsaregorodtsev, Stuart Paterson (CERN)**
  - Dirac3 Integration
- **Alice: Latchezar Betev**

Success conditions

- No major issues present in glexec and SCAS
- Stable activity (multi user pilot jobs) for ~2 weeks
- Achieved integration with experiments' frameworks
- Positive feedback of site managers about operability

Access info

- **Home Page**
  - https://twiki.cern.ch/twiki/bin/view/LCG/PpsPilotSCAS
- **Meetings**
  - 05-Feb-09: Kick off. Minutes at PPIslandKickOff2009x02x05
  - 19-Feb-09: Follow-up. Minutes at PPIslandFollowUp2009x02x19
  - 26-Feb-09: Follow-up. Minutes at PPIslandFollowUp2009x02x26
- **Contacts**
  - egee-pps-pilot-scas@cern.ch

Agenda

- **Recent history**
  - Software versions
  - Deployment
  - Integration works
  - Stress testing

Software versions

- **Software versions**
  - Glexec: patch #2770 -> #2829 (27th-Feb).
    - Consistent error codes plus improvements to SCAS client.
  - SCAS service: patch #2767

New releases:SCAS client (#2829)

- **lcmaps-plugins-scas-client 0.2.7-1**
  - Improved:
    - Back off tactics improved on TCP/IP and SSL layer
    - Treats SOAP level failures properly
  - New feature:
    - Fail-over and/or load balancing options (diff strategies):
      - *Round-robin:*
        - follow the configured list of endpoint (top to bottom)
      - *Round-robin random start (load balancing):*
        - **(default)** Like previous, but the the first endpoint to try is randomly selected, adequately balancing between the endpoints
      - *Random (least effective, but load balancing):*
        - Pseudo random endpoint selection (not smart enough to not try a failed endpoint yet).

New releases:SCAS client (#2829)

- **gLExec 0.6.6-2**
  - Improved stability:
  - Fixing open issues
    - file locking technique has more options (for CREAM CE)
    - removed rudimentary safety checks
  - New feature:
    - Consolidating exit codes of gLExec:
      - *201: User error*
        - Input proxy wasn't setup correctly, command not found, other items
      - *202: System error*
        - glexec.conf has wrong perms, or init problem in LCAS or LCMAPS
      - *203: AuthZ failed*
        - Calling user wasn't a pilot/prod user, proxy verification failed, SCAS didn't return an account (blacklisted user or no account available)
      - *204: Child process executed and returned overlapping exit code*
        - Child process return with exit codes 201, 202, 203 or 204
      - *126: Shell returned that the executable can't be executed*

Deployment

- **FZK (ready since 19th-Feb)**
  - Dedicated CE test-mw-2-fzk.gridka.de accounted for FZK-LCG2
  - CE-Status is set to SCASPilot
  - currently published in pre-production. SITE_NAME=FZK-PPS
  - can be used by
    - lhcb:/lhcb/Role=pilot (queue lhcbXXL)
    - atlas:/atlas/Role=production + atlas:/atlas/usatlas/Role=pilot (queue atlasXXL)
    - cms:/cms/Role=production (queue cmsXL)
    - alice:/alice/Role=pilot (queue aliceXL)
- **Nikhef EL-Prod (ready since 27th-Feb)**
  - All the WNs are gLExec-enabled and accessible by all prod CEs
  - multiple SCAS endpoints for fault tolerance. All production CEs are configured to use WNs with gLExec.
  - can be used by:
    - lhcb:/lhcb/Role=pilot
    - atlas:/atlas/Role=production
- **Both sites run (#2829)**
- **IN2P3 ready to step-in after the 15th Mar**

Integration works: Atlas

- **Testing in progress at FZK since 26th- Feb**
- **installation of gLExec/SCAS at gridka works fine**
  - If a myproxy server is used to pass the credentials, myproxy-logon has to be installed on the WN
  - Old problems with proxies found in previous versions have been fixed.
- **More info in the summary of PanDA initial testing**
- **Issues to be addressed by the exp framework**
  - When gLExec is invoked, the environment (belonging the pilot) vanishes
  - when gLExec is invoked the current directory is moved to the new user HOME directory
  - the new user has no permissions to execute programs in the pilot directories nor to write (i.e. the output and log files which then the pilot will be looking for...)

Integration works: LHCb

- **Testing started at NIKHEF on the 9th- Mar**
  - No progresses reported so far. Error messages from LCMAPS
  - Waiting to be white-listed at FZK to cross-check

Integration works: Alice

- **Taking part to pilot check-points**
- **Currently more focused on CREAM**
- **Integration tests tentatively for the beginning of summer**
  - Pilot layout may evolve in the meantime

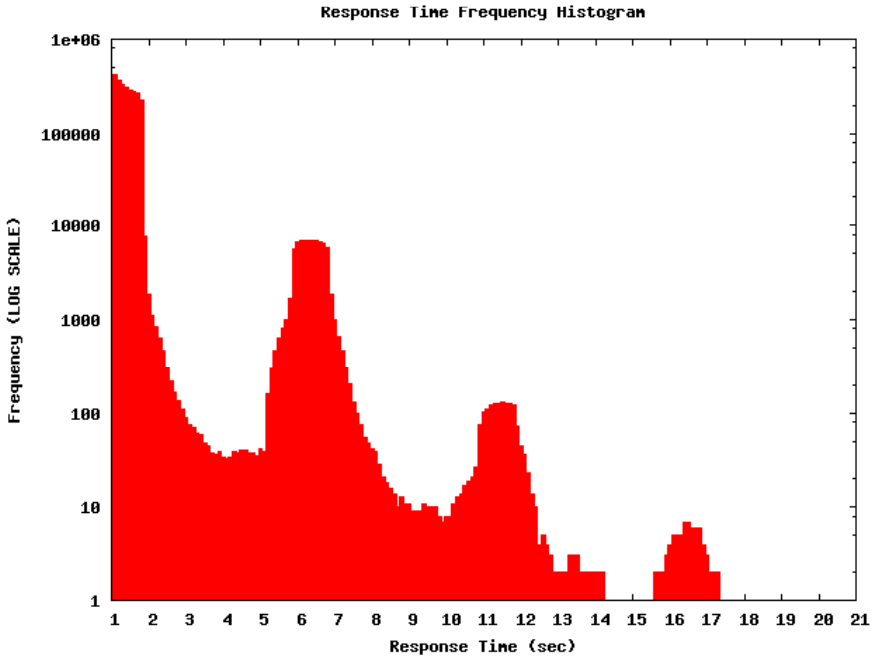
Stress testing

- **Stress testing:**
  - Details at: SA3-SCAStestsresults
  - Latest results summary: the memory leak in the server is still present but the new SCAS client rpm nearly removes the errors due to the internal SCAS server refresh.

From **(6M request, 14K errors)** *to* **(3M requests, 1 error)**

Glexec response time on a WN:

- Zone [0,2): 98.16%
- Zone 2 [2,10): 1.80%
- Zone 3 [10, +inf): 0.04%



Next steps

- **Finish Panda and Dirac integration test**
  - Start working on return codes
- **CE at FZK published in production**
- **Deployment of SCAS/glexec at IN2P3**
  - Site available starting from 15th of March
  - Focus on load balancing
- **Next check-point : 19th of March**

Questions?