Red Hat MRG

Matthew Farrellee

Grid Technical Lead, Engineering Manager

Outline

- **Background**

- **MRG Status**

- **Condor and Deltacloud**

- **Condor/Red Hat Collaboration**

- **Condor at 10K+ scale**

Background

- **Red Hat is investing in a set of defining technologies, actively being developed in conjunction with partners. These key projects include among others...**

  - Grid/Distributed computing (Partnership with UW around Condor)

  - Real-time (provides pre-emption and latency determinism)

  - Virtualization (provides scale, VM, KVM, and ability to provide environment with Job)

  - Distributed computing / Messaging (built on AMQP) www.apache.org/qpid www.amqp.org

  - Cloud APIs with Deltacloud www.deltacloud.org

  - .... just to name a few

- **The combination of these types technologies in a standard Linux distribution will provide a new generation of compute platforms.**

MRG Status

- ## Coming up on the 3rd update, MRG 1.3

  - http://www.redhat.com/mrg

  - Qpid graduated to top level Apache project

  - Enhanced Condor manageability

- ## Red Hat Global Professional Services

  - Engaging around MRG Grid

  - Strategic planning

  - Pilot and integration testing

  - Integration into standard operating environment

  - Full life-cycle management and migration

  - Custom development around Condor

Enhanced Condor manageability

- **MRG Management Console**
  - http://www.youtube.com/watch?v=kPi8ickYN84
  - Improved job visualization with rollup into submissions
  - Integrating dynamic EC2 resources into pool views
  - Addition of user focused interface
  - Scale by removing backend load from Schedd

- **Alerting infrastructure over Messaging**

- **Configuration management**
  - Model params, features, nodes, groups
  - Configuration validation, feature param resolution
  - Ensure configuration application
  - See Will and Rob's talk later

Condor and Deltacloud

- **Deltacloud is an API that abstracts the differences between clouds**

  - http://www.deltacloud.org/drivers.html

  - EC2, RHEV-M, Rackspace, RimuHosting, OpenNebula, …

- **Deltacloud from Condor-G**

  - grid_resource = dcloud

- **Deltacloud to Condor resources**

  - Build on VM Universe, Condor's resource management and scheduling capabilities

  - Decisions to make: image management, VM sizing, realms

- **Positioning Condor as a core component in Cloud building**

Condor Week 2007
-Red Hat/Condor collaboration
*Condor open source

*Red Hat Madison office

Condor Week 2009
-Collaboration update

Condor Week 2008
-EC2
-Management

Condor Week 2010

Condor Project and Red Hat Collaboration

Condor 7.0 Release
-1st release with source
-Condor into Fedora
-Beta MRG release

Condor 7.4 Release
-Open Source advancement
-UW/Red Hat features

Condor 7.2 Release
-UW/Red Hat features
-Red Hat MRG 1.1 GA

Open Source advancement

- **UW source code repository now available via a public repository**

  - git clone http://condor-git.cs.wisc.edu/repos/condor.git

- **UW ticket tracking system now available to anonymous users**

  - **http://condor-wiki.cs.wisc.edu**

  - Ticket/issue tracking (bugs, features, todo, doc)

  - Wiki with developer, user, administrator documentation

- **Transparency**

  - Ticket numbers (#XYZ)

  - http://condor-wiki.cs.wisc.edu/index.cgi/tktview?tn=XYZ

Open Source advancement

- **Tickets (from 13 Jan 2009 to 14 April 2010)**

  - 1,352 total tickets

  - 289 resolved defects

  - 143 resolved enhancements

- **Commits (as of 14 April 2010)**

  - 3,196 commits in the last 260 working days by 34 individuals

Count commits over how many days:	365	Go

Developer	Number of commits
~ TOTAL ~	3196
Jaime Frey	527
Dan Bradley	409
Matthew Farrellee	392
Peter Keller	260
Nick LeRoy	208
Kent Wenger	204
Timothy St. Clair	156
Greg Thain	146
Ziliang Guo	100
Robert Rati	98
Todd Tannenbaum	82

UW/Red Hat features

- **VM Universe enhancements**

  - The condor_vm-gahp now supports KVM and links with libvirt, rather than calling virsh command-line tools. (#688, #1079, …)

  - VM Universe support for helper scripts to generate libvirt XML (#618)

- **DAEMON_LIST to be reconfigurable (#1306)**

  - Add multiple Startds or Schedds without restarting

- **kill_sig for Vanilla Universe (#1289)**

  - Allows programs a chance to cleanup or checkpoint

- **Modernizing build chain with cmake (was imake) (#779)**

  - For developers mostly, proper dependencies, consistent across platforms (Windows), enhanced packaging

- **Configuration table, enhanced usability (#123, #988)**

- …

Condor at 10K+ scale

- **Negotiator and Collector**

  - Both want fast CPUs and modest amounts of RAM

  - Negotiator speed directly influences utilization, through speed in producing matches and interrupt (blocking) interaction with Schedds

  - Collector speed impacts utilization through Negotiator interaction and the accuracy of its pool view

- **Schedd and Shadows**

  - Submit host for 10K CPU pool needs >16GB of memory

  - Shadow private memory O(1.5MB)

  - 8GB DDR3 DIMMs at ~$1,000 each, need 3 to be >16GB

  - Double that for an HA Schedd configuration

  - Triple it to allow for backup hardware

  - That's ~$0.90 per CPU for memory

Improvements for scale

- **Negotiator –**

  - Development of asynchronous negotiation protocol on the development roadmap (#928)

- **Collector –**

  - When anything leaves the pool it sends an INVALIDATE message to the Collector

  - Invalidation was O(number of ads in the Collector)

  - From 7.4.3 on the Collector has an O(1) algorithm (#1316)

- **Shadow –**

  - Memory & code analysis needed

- **Schedd –**

  - Ongoing efforts reducing file system operations, optimizing event loop

  - #921, #749, #1061, …

Schedd scale

- **Jobs: 20,000 sleep 45 (no log), SPOOL on NFS**

- **Patched –**
  - #1287, #1288 (March)

- **Duration –**
  - 7.4.3-0.5: 5,853 sec
  - Patched: 1,261 sec

- **Concurrent (mean) –**
  - 7.4.3-0.5: 544 shadows
  - Patched: 1,255 shadows

- **More! 10x speedup**
  - #1348 (just this week)
  - ~27 to ~277 jobs/sec
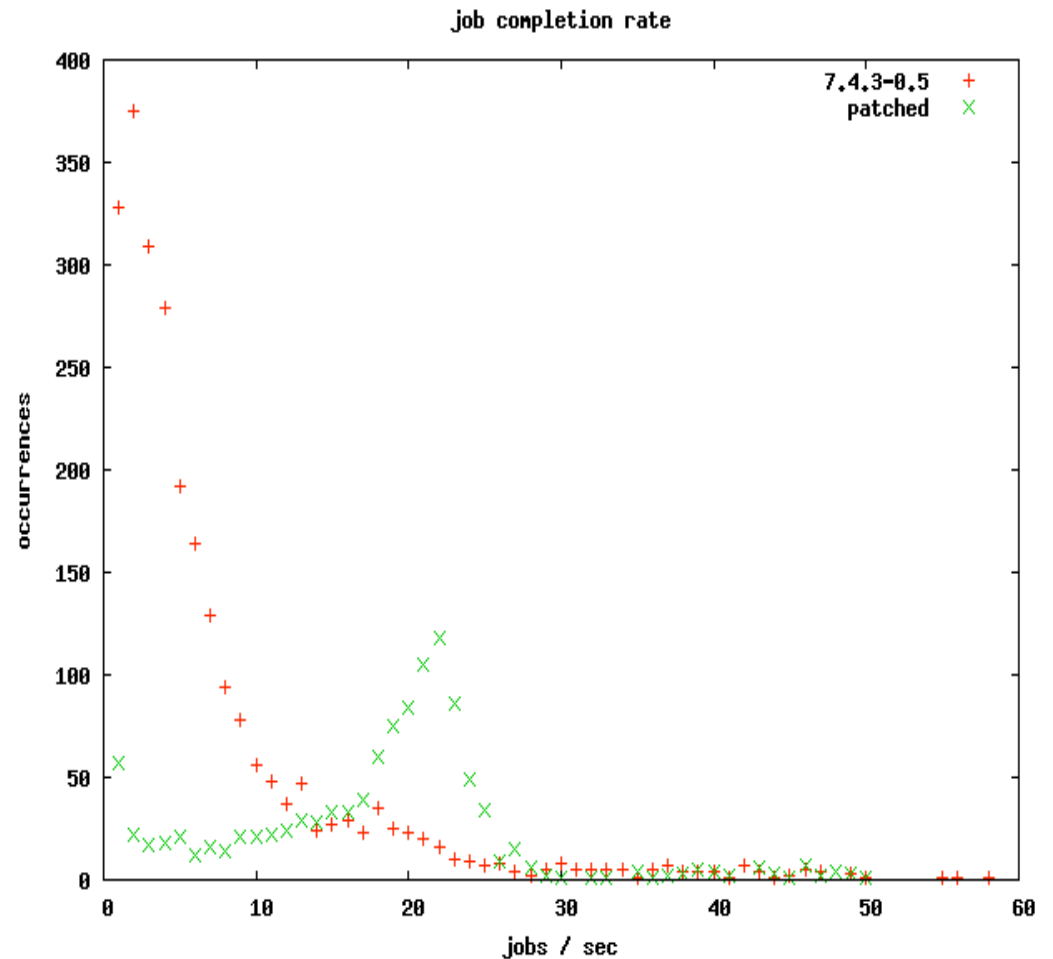
redhat.

Thank You

**We are still hiring, contact
matt@redhat.com**