Condor Week 2007

Glidein Factories
(and in particular, the glideinWMS)

by Igor Sfiligoi
Anybody heard of “The Grid”? 

- “The Grid” is the current way forward in most sciences
  - Certainly in High Energy Physics (and in particular CMS)

- “The Grid” is the sum of “Grid Sites”, each offering a moderate amount of (mostly) computing resources
  - Each site has a standard “Gatekeeper”, responsible for regulating access to the site
    (How the “Gatekeeper” handles the computing resources, is anyone's guess)
Dear public, “The Grid”

And “The User”

“The Grid” is not an easy place to live in!
Compare this to **Condor**

- A single system from the user point of view
  - User submits to a local scheduler
  - Condor does all the magic
So let **Condor** manage “**The Grid**”!

Life is good again!
So let **Condor** manage “**The Grid**”!

**But how do we get here?**

Life is good again!
The answer: **Condor glide-ins**

**Legenda:**
- Central manager
- Execute daemon
- Submit node and user(s)
- Gatekeeper
- Worker node
The answer: **Condor glide-ins**

**Legenda:**
- Central manager
- Execute daemon
- Submit node and user(s)
- Gatekeeper
- Worker node

Condor Week 2007

Glidein Factories - by I. Sfiligoi
What exactly is “a glidein”?  

• “A glidein” is just a regular `condor_startd` daemon, submitted as a Grid job  
  
• The glidein-Grid job needs to:  
  – validate the worker node (for example against memory and disk problems)  
  – discover or fetch the condor binaries  
  – configure the Condor daemons  
  – start the Condor daemons  

• For simple use-cases, you can use `condor_glidein`
The glidein factory

- Needs to know how to submit to the "Grid Sites"
  - ... how to obtain the list of sites
  - For each site:
    - how to talk to the "Gatekeeper"
    - what is the configuration of the site
      (network, security, software, etc.)

- Needs to know when to submit new glideins
  - Slots are not free
  - Resources not used by my pool could be used by others
    - Submit only if users need more resources
      (modulo speculative submissions)
    - Submit only to sites who declare that can run at least a subset of user jobs
The glideinWMS
http://home.fnal.gov/~sfiligoi/glideinWMS/

• A glidein-based Workload Management System (WMS) developed for USCMS
  – Derived from the CDF GlideCAF (Presented at CondorWeek2006)
  – But meant to be generic enough to support different communities

• Uses the dividi-et-impera approach
  🏭 Glidein Factories know how to submit to the Grid Sites
  🍀 VO* Frontends monitor jobs and direct the factories

• Condor Collector used for message passing

* VO = Virtual Organization ~ Condor Pool
The **glideinWMS**

http://home.fnal.gov/~sfiligoi/glideinWMS/

**Legenda:**
- Central manager
- Execute daemon
- Submit node and user(s)
- Gatekeeper
- Worker node

**WMS Legenda:**
- Collector
- Glidein factory
- VO frontend
The glideinWMS

http://home.fnal.gov/~sfiligoi/glideinWMS/

Legenda:
- Central manager
- Execute daemon
- Submit node and user(s)
- Gatekeeper
- Worker node

WMS Legenda:
- Collector
- Glidein factory
- VO frontend
The glideinWMS
http://home.fnal.gov/~sfiligoi/glideinWMS/

Legenda:
- Central manager
- Execute daemon
- Submit node and user(s)
- Gatekeeper
- Worker node

WMS Legenda:
- Collector
- Glidein factory
- VO frontend
glideinWMS internals

Count jobs that match factory attributes

Factory Name
Requested idle glideins

Factory Name Attributes

Keep requested idle glideins in the queue

Legenda:

Condor-G scheduler
Everything else like previous slide

More details in the backup slides
• Glidein startup script simply loads other scripts
- HTTP used for network transfers (cacheable, works when no privacy issues)

signature.sha1
file.lst
condor_bin.tgz
configs.cfg
validate.sh
myscript.sh
start_condor.sh

Web Cache

All files signed

Web Server

Startup script + arguments

Worker Node

glidein_startup

load file list
load files
execute scripts

Errors?
No

Downloaded scripts do all the real work

Yes

Startd

This batch slot would not be able to run a user job

See backup slides for text description.
Network security concerns

- Traffic on WAN insecure by definition
- Using x509 (GSI) service proxies for authentication
  - **Condor tools** securing communication between
    - VO Frontend and Glidein Factory
    - Startd and Collector/Schedd
  - **Condor** supports integrity checks to prevent data tampering and encryption for privacy
- HTTP-accessed data checked via SHA1 checksums
  (no privacy possible here)
Security on the Worker Nodes

• Glide-in **Condor** not running as a privileged user
  – Cannot change UID without help from the system
  – **Condor daemons not protected from user jobs**

• **Open Science Grid (OSG)** starting to deploy gLEExec on its worker nodes
  – A x509-based Apache-suexec derivative
  – **Condor** can use the service proxy to run the user job under a different UID
  – **Same security as if Condor running as root**
Working over Firewalls

- **Condor** is based on the peer-to-peer principle
  - Needs two-way network traffic
- Most **Grid Sites** behind firewalls
  - Most have only outgoing connectivity
  - Some only proxied traffic
- **Condor GCB** can help at such sites
  - See GCB talks for more details
- VPNs could be another option, but are less trivial to use in user-space
Conclusion

- "The Grid" 🌍 has a lot of resources (even for free)
  - Why not use them?
- Glideins 🦚 allow you to use those resources without a single change in your jobs
  - You can even submit standard universe jobs!
- glideinWMS 🦚 can help you automatize the maintenance of a glidein pool
  - Let me know if you are interested
    sfiligoi@fnal.gov
    http://home.fnal.gov/~sfiligoi/glideinWMS/
Glidein Factories

Backup slides
VO Frontend ClassAd

Due to Condor limitations, define also GlideinMyType

Target a specific Entry Point

Request a steady stream of glideins starting

Costumize the submitted glideins.
GlideParamXXX must match the names published by the factory

Monitoring data like:
Idle="546", Running="222"

Published classad

MyType="glideclient"
Name="reqX@client"
ClientName="client"
ReqName="reqX"
ReqGlidein="entry@factory"
ReqIdleGlideins.nr
ReqMaxRun.nr
ReqMaxSubmitXHour.nr
GlideinParamWWW="val1"
...
GlideinParamZZZ="valY"
GlideinMonitorNNN="valN"
...
GlideinMonitorMMM="valM"
Glidein Factory ClassAd

Published classad

MyType="glidefactory"
Name="entry@factory"
FactoryName="factory"
GlideinName="entry"
Attribute1="...
...
AttributeN="...
GlideinParamXXX="val1"
...
GlideinParamYYY="valZ"
GlideinMonitorNNN="valN"
...
GlideinMonitorMMM="valM"

Due to Condor limitations, define also GlideinMyType

Attributes that describe the glidein like:
ARCH="INTEL", MaxHours=72, Site="Florida"

Parameters set glidein parameter defaults like:
CONDOR_HOST="UNDEFINED", SEC_DEFAULT_ENCRYPTION=OPTIONAL
MinDisk=16G, CheckFilesExist="/tmp/CMS,$DATA/OSG"

Monitoring data like:
TotalStatusIdle="234", TotalStatusRunning="1356"
TotalRequestedIdle="50"
Glidein Factory essentially a publish-read-submit loop

Details about ClassAd content in the backup slides
• VO Frontend acts as a matchmaker

- VO Schedd
- Query Schedd(s)
- Jobs Attributes
- Match and count
- Nr jobs x Factory
- Publish requests
- VO Collector
- WMS Collector
- Details about ClassAd content in the backup slides
Glidein details

- Dummy startup script
  - Just loads other files and execute the ones marked as executable
- File transfer implemented using HTTP
  - Easy cacheable, standard tools available (Squid)
  - Proven to scale, widely used in Industry
- All sensitive file transfers signed (SHA1)
  - Prevent tampering, as HTTP travels in clear
Glidein details

- Standard sanity checks provided
  - Disk space constraints
  - Node blacklisting
- Generic Condor configure and startup script provided, too
- Factory admins can easily add their own customization scripts (both for checks and configs)
  - Allowing Frontends to add custom scripts envisioned, but not yet implemented
Open a permanent connection

One way firewall

Reuse the permanent connection
**glideinWMS support**

- **glideinWMS** developed by and for the CMS collaboration
  - No funding to support other users

- However:
  - Having other users would bring in new ideas
    - Best-effort support will always be there for everybody
  - Collaboration with other groups welcome
    - both for development and support