

End-To-End Solution for Integrated Workload and Data Management using GlideinWMS and Globus Online

Parag Mhashilkar, Fermi National Accelerator Laboratory
(on behalf of all the authors)

Overview

- CEDPS Project
- Motivation for the Integration
- GlideinWMS Overview
- Globus Online Overview
- Condor File Transfer Plugin Architecture
- Globus Online Plugin
- Interfacing GlideinWMS with Globus Online
- Stress Tests
- Conclusion

Center for Enabling Distributed Petascale Science (CEDPS)

- CEDPS: Five year project 2006 - 2011, funded by the Department of Energy (DOE)
- Goals
 - ◎ Produce technical innovations for rapid and dependable data placement within a distributed high performance environment and for the construction of scalable science services for data and computing from many clients.
 - ◎ Address performance and functionality troubleshooting of these and other related distributed activities.
- Collaborative Research
 - ◎ Argonne National Laboratory
 - ◎ Fermi National Accelerator Laboratory
 - ◎ Lawrence Berkeley National Laboratory
 - ◎ University of Wisconsin Madison
 - ◎ Information Sciences Institute

CEDPS Activities at Fermilab

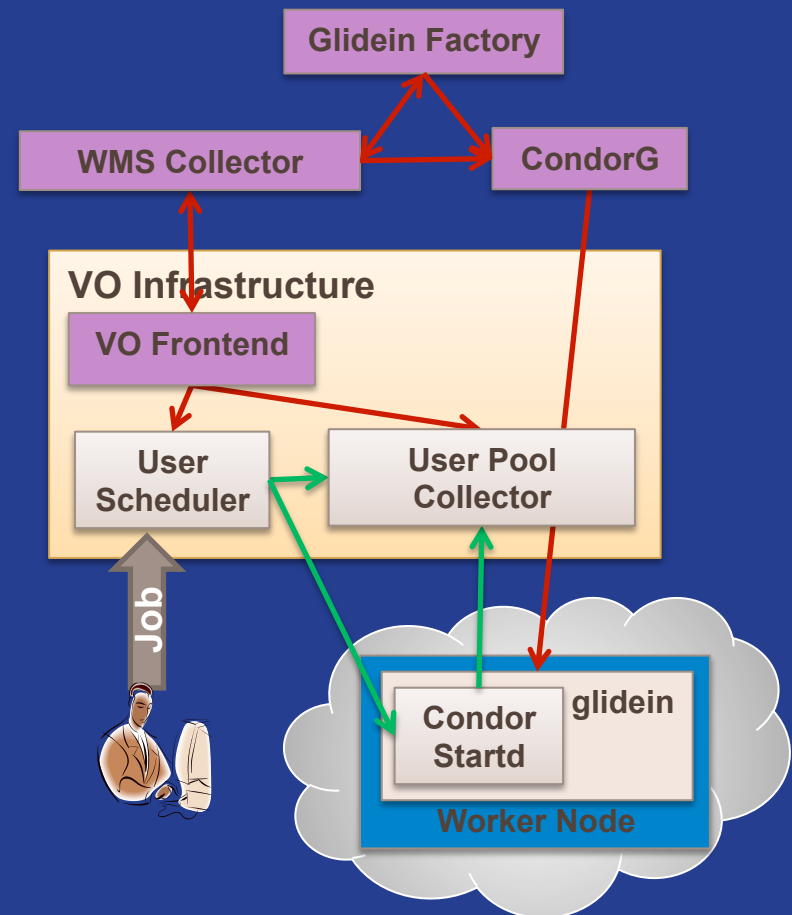
- Phase 1:
 - ⊙ Provide data corruption detection in the Gridftp service
- Phase 2:
 - ⊙ Integration of asynchronous data stage-out mechanisms in overlay workload management systems
 - ⊙ Integrate support for Globus Online in GlideinWMS (This talk)
 - ⊙ Release resources at job termination. Delegate data stage-out to external agents.
 - ⊙ Supporting the integration of data movement mechanisms with scientific DH frameworks
 - ⊙ Supporting the integration of Globus Online with Dark Energy Survey (DES) data handling system
 - ⊙ Investigating different data movement mechanisms for data stage-out on Grids

Motivation

- Provide an End-To-End Solution
 - ◎ Large VOs have custom data management solutions
 - ◎ Too expensive for small VOs
 - ◎ Data transfer done by the infrastructure and not by the job
 - ◎ Integrates with the WMS
 - ◎ Integrates with the data transfer layer
 - ◎ Independent of the transfer protocol/service
 - ◎ Can be extended to different data transfer services/protocol like SRM, IRODS, Gridftp, rcp, scp, etc.
- Solution
 - ◎ GlideinWMS as WMS for user jobs
 - ◎ Globus Online as data movement service
 - ◎ Use Condor's file transfer plugin architecture

GlideinWMS

- Pilot-based WMS that creates on demand a dynamically-sized overlay condor batch system on Grid resources to address the complex needs of VOs in running application workflows
- Components
 - ◉ WMS Collector
 - ◉ Glideinfactory
 - ◉ User Pool Collector + CondorG
 - ◉ User Scheduler
 - ◉ VO Frontend
- Factory knows about the sites and how to submit glideins to the sites
- VO frontend knows about the user job details
- WMS Collector acts as a dashboard for Factory - VO Frontend communication.



Globus Online

- Initial implementation of XUAS (XSEDE User Access Services)
- Reliable file transfer
 - ⊙ Easy “fire and forget” file transfers
 - ⊙ Automatic fault recovery
 - ⊙ High performance
 - ⊙ Across multiple security domains
- No IT required
 - ⊙ Software as a Service (SAAS)
 - ⊙ No client software installation
 - ⊙ New feature automatically available
 - ⊙ Works with existing GridFTP servers
 - ⊙ Globus Connect
 - ⊙ Consolidated support and troubleshooting




Command line interface
ls alc#dtn:~
scp alc#dtn:~/myfile \
nersc#dtn:~/myfile

HTTP REST interface
POST https://transfer.api.
globusonline.org/v1.0/
transfer <transfer-doc>

Globus Online – hosted service for high speed, reliable, secure, data movement

- Easy fire-and-forget transfers
- Automatic fault recovery
- Across multiple security domains



Globus Connect Multi-User
at campus clusters

GridFTP server at large
Compute facilities

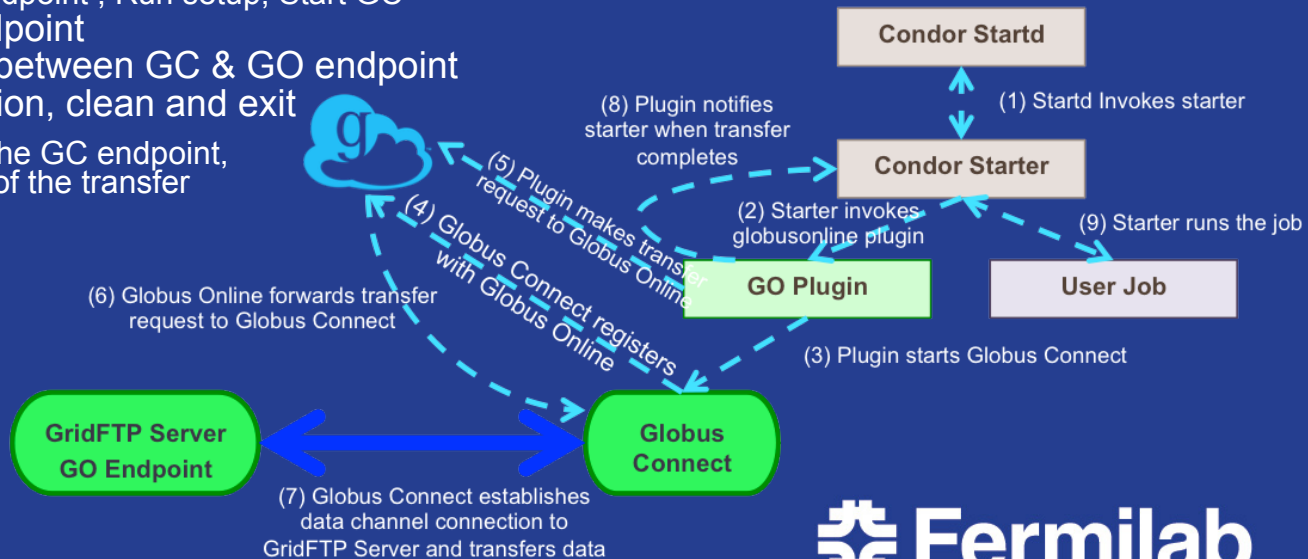
Globus Connect
on local computers

Condor File Transfer Plugin Architecture

- Condor File Transfer Plugins
 - ⊙ Supported since condor v7.6
 - ⊙ Extend transfer protocols supported
- Plugins
 - ⊙ Executable scripts written in your favorite language
 - ⊙ Easy configuration
 - FILETRANSFER_PLUGINS = /path/to/curl_plugin, /path/to/protocolx_plugin**
 - ⊙ Two modes of operation
 - ⊙ Query mode: Register the plugin and protocols supported
 - \$./curl_plugin -classad**
 - PluginVersion = "0.2"
 - PluginType = "FileTransfer"
 - SupportedMethods = "http, ftp, file"
 - ⊙ Action mode: Perform an actual transfer
 - \$./curl_plugin http://home.fnal.gov/~parag/test.file /tmp/test.file**

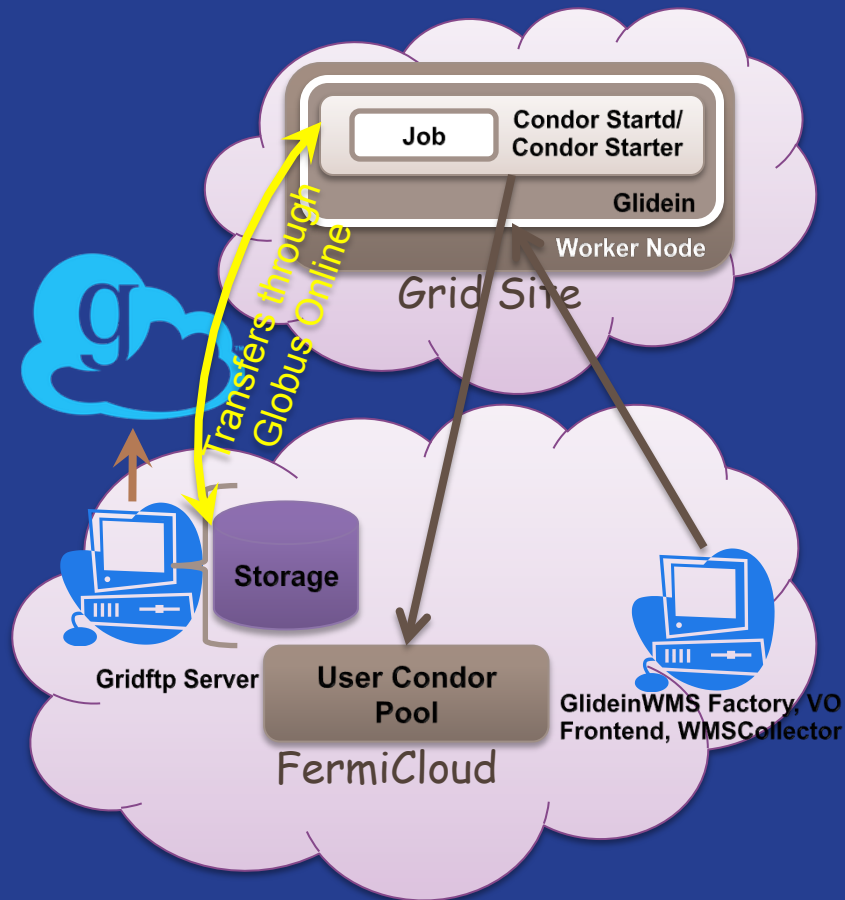
Globus Online Plugin

- Modes of operation
 - ⊙ Query mode:
`$./globusonline_plugin.py -classad`
`PluginVersion = "0.2"`
`PluginType = "FileTransfer"`
`SupportedMethods = globusonline, file"`
 - ⊙ Action mode:
`$./globusonline_plugin.py globusonline://parag:if-go-ep:/tmp/datafile /tmp/test.file`
- Figure best possible pre-requisites
 - ⊙ X509_USER_PROXY, Globus Connect (GC) to use, GO user, GO endpoint & file location from the URI
- Bootstrap Globus Connect
 - ⊙ Unpack, Add GC endpoint , Run setup, Start GC
- Activate the GO endpoint
- Initiate scp/transfer between GC & GO endpoint
- On transfer completion, clean and exit
 - ⊙ Stop GC, Remove the GC endpoint, Exit with the status of the transfer



End-To-End Solution: GlideinWMS & GO

- Test setup on FermiCloud
 - ⊙ Infrastructure-as-a-Service private cloud for Fermilab's Scientific Program
- Setup a gridftp server
- Create a GO end point for the gridftp server
- Configure VO Frontend to use Globus Online plugin
 - ⊙ Custom Scripts
 - ⊙ globusonline_plugin.py,
 - ⊙ globusconnect
 - ⊙ setup script
- Run Test job



Test job: Condor JDF

```
universe = vanilla
executable = /local/home/testuser/testjobs/info-from-go.sh
initialdir = /local/home/testuser/testjobs/joboutput
output = out.$(cluster).$(process)
error = err.$(cluster).$(process)
log = log.$(cluster).$(process)
should_transfer_files = YES
when_to_transfer_output = ON_EXIT_OR_EVICT

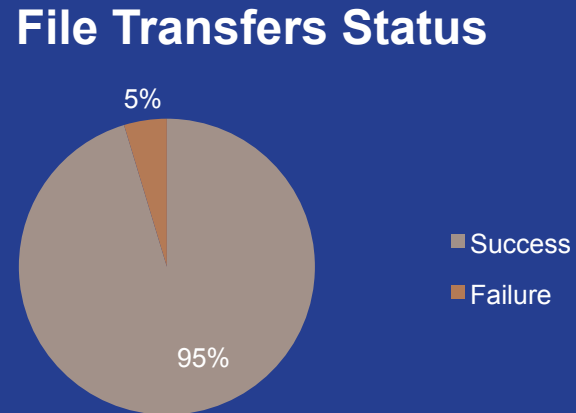
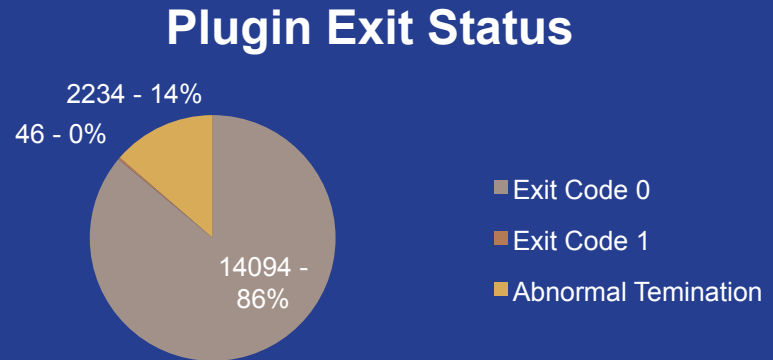
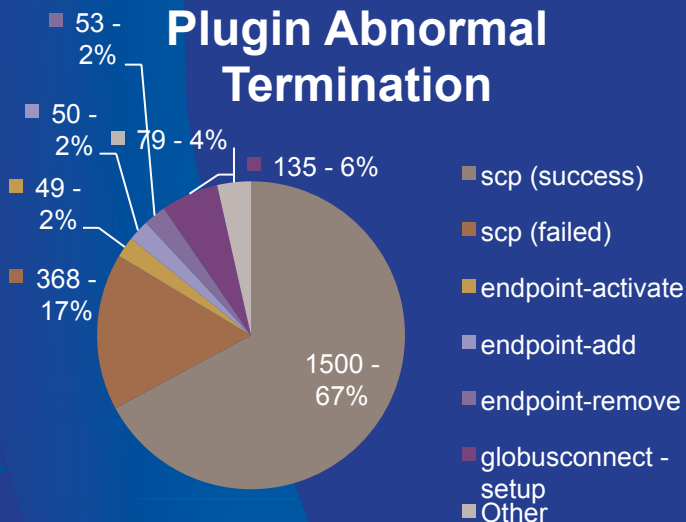
# Input files
transfer_input_files = globusonline://parag:paragtest:/grid/data/parag/data

# Output sandbox transfer
output_destination = globusonline://parag:if-go-ep:/tmp/uploaded_data-fnalgrid
transfer_output_files = data_to_upload

x509userproxy=/local/home/testuser/testjobs/x509up_u11017.fermilab
+DESIRED_Sites = "GR7x1"
Requirements = (stringListMember(GLIDEIN_Site,DESIRED_Sites)) && (Disk > 0 && (Arch == "INTEL" || Arch ==
"X86_64"))
queue
```

Stress Tests & Results

- Identify the bottleneck when WMS and data services are integrated
- Tests: Plugin's first prototype
 - ◉ 2636 Minerva Jobs submitted
 - ◉ 16359 Output files transferred back through GO



Advantages of the Approach

- You can support any protocol!
 - ◎ Needs protocol specific plugin
 - ◎ Required binaries/scripts
 - ◎ Transfer using GlideinWMS custom scripts OR
 - ◎ Pre-installed on the worker nodes
- Independent of the GlideinWMS release
- Independent of the condor release
 - ◎ Just requires condor v7.6+

Limitations

- Condor
 - ⊙ X509_USER_PROXY: Inconsistent between input & output sandbox transfers
 - ⊙ Getting Logs/Debugging Info back to the submitted
 - ⊙ All the files go to the GO endpoints and nothing gets transferred back to submit node
 - ⊙ If there is a problem accessing GO itself, logs are lost.
 - ⊙ No support for multiple transfers
 - ⊙ Syntax too rigid: plugin <source> <destination>
 - ⊙ Reusing/Reducing number of Globus Online endpoints
 - ⊙ Each file transfer creates a new GO endpoint.
- Globus Online
 - ⊙ Fail earlier than 1 day default for Globus Online
 - ⊙ Current limit of 3 active transfers per user. You can have more queued.
 - ⊙ Stress tests stressed the relay servers
- Transfer Plugin
 - ⊙ Which GLOBUSCONNECT
 - ⊙ We need to do a better job of figuring out which GLOBUSCONNECT to use.
 - ⊙ Start the transfer in background and periodically poll it. Fail accordingly rather than the default GO deadline of 1day

Conclusion

- End-To-End Solution that works!
 - ◎ Integrates WMS with the data management layer
 - ◎ Data transfer done by the infrastructure and not by the job
 - ◎ Independent of the glideinWMS release
- Solution using Condor's transfer plugin architecture,
 - ◎ currently supports GO
 - ◎ can be easily extended to different data transfer services/protocols

References

- CEDPS
<http://www.cedps-scidac.org>
- Globus Online:
<https://www.globusonline.org/>
- Condor: <http://cs.wisc.edu/condor>
- Glideinwms: <http://www.uscms.org/SoftwareComputing/Grid/WMS/glideinWMS/doc.prd/index.html>
- End-To-End Solution Project:
<https://cdcvs.fnal.gov/redmine/projects/cedps-glideinwms>
<http://cd-docdb.fnal.gov/cgi-bin/ShowDocument?docid=4406>
- ◉ Test Results:
<http://cd-docdb.fnal.gov/cgi-bin/ShowDocument?docid=4474>
- ◉ Condor File transfer plugin: <http://www.cs.wisc.edu/condor/CondorWeek2011/presentations/zmiller-cw2011-data-placement.pdf>