GeoIP HTTPS Redirector

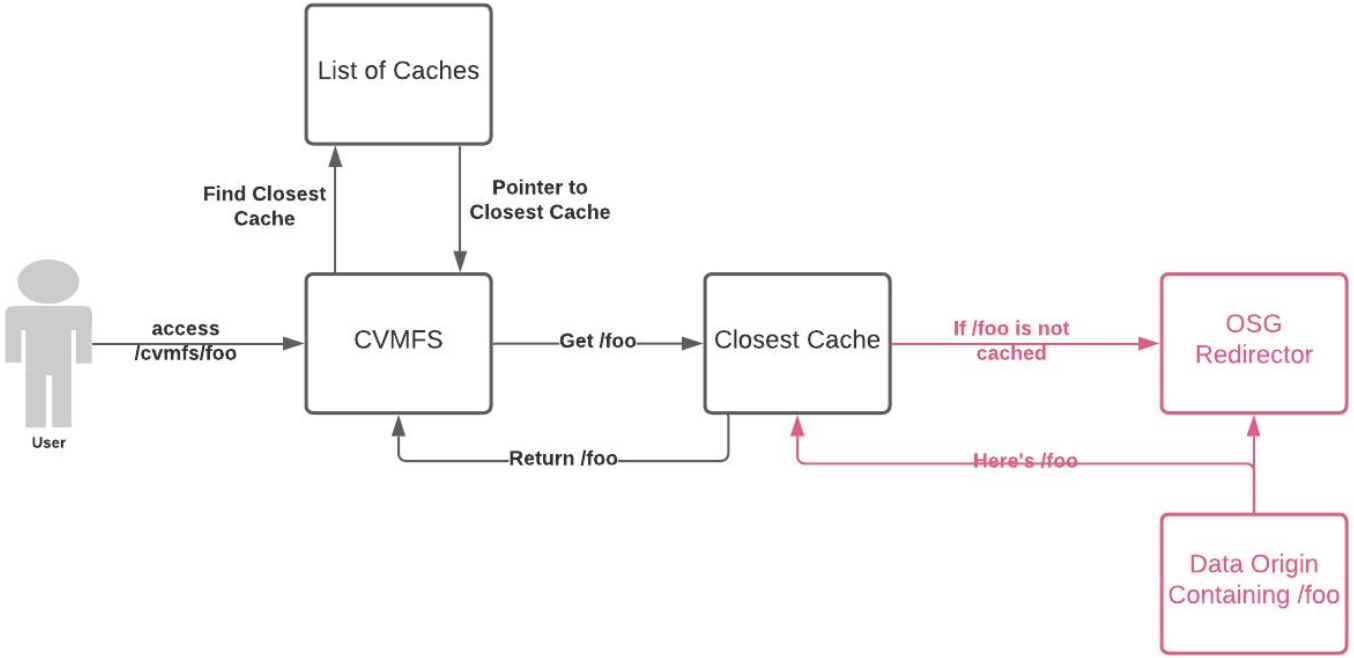
Aashay Arora, Edgar Fajardo

UC San Diego

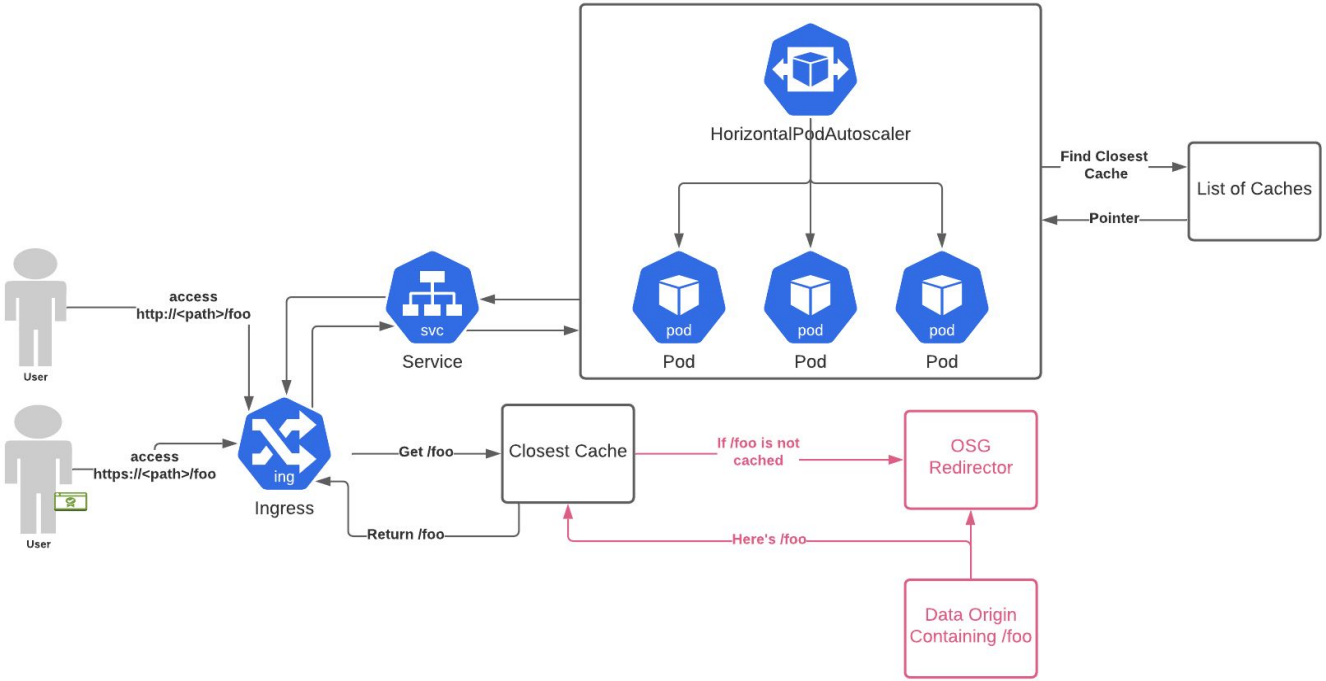
Motivation

- Data is moved from origin to the job site over the network.
- Even if the data file is reused by multiple users, it still has to be transferred over the network n times.
- Resolution: Stashcache [see more]
- Create a proxy web-app which redirects users to the XCache site closest to the job site based on GeoIP.
- Deployed using Kubernetes Pods, the redirector can scale vertically and horizontally.

How Stashcache Works



Our Proposal



Overview

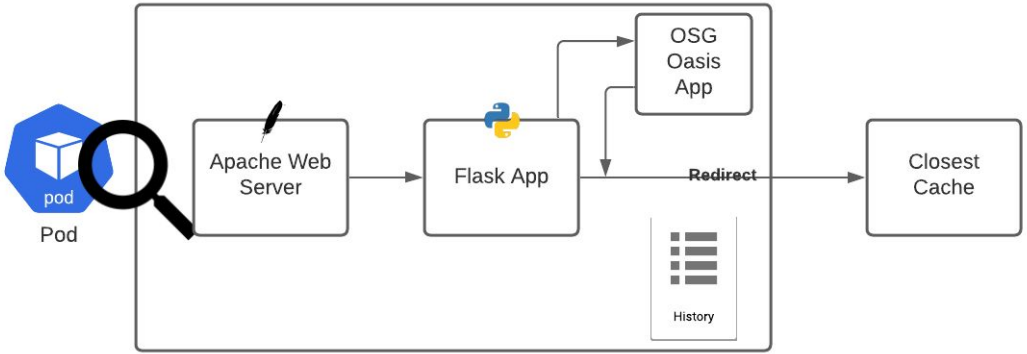
- Deploy Python (Flask based) app behind an Apache web-server on a Kubernetes Pod.
- Create Ingress to host the web-app on a public url.
- Use the Horizontal Pod Autoscaler API to make sure pod is automatically replicated when:
  - CPU Usage exceeds threshold.
  - Requests-per-second for the ingress exceed threshold.

  Important for when put into production

- Run tests to make sure things work with HTCondor.

Python Web App

- HTTPS GeoIP Redirector: [Github]
  - Flask based web app which returns a redirect request.
  - Calls OSG-Oasis GeoIP Web App to locate the nearest Cache of the client's IP
  - https://geoip-https.nautilus.optiputer.net
  - File at subpath /*foo* can be accessed like:
    - https://geoip-https.nautilus.optiputer.net/foo
  - User IP and closest cache is memorized so that re-visiting users can be looked up quickly.

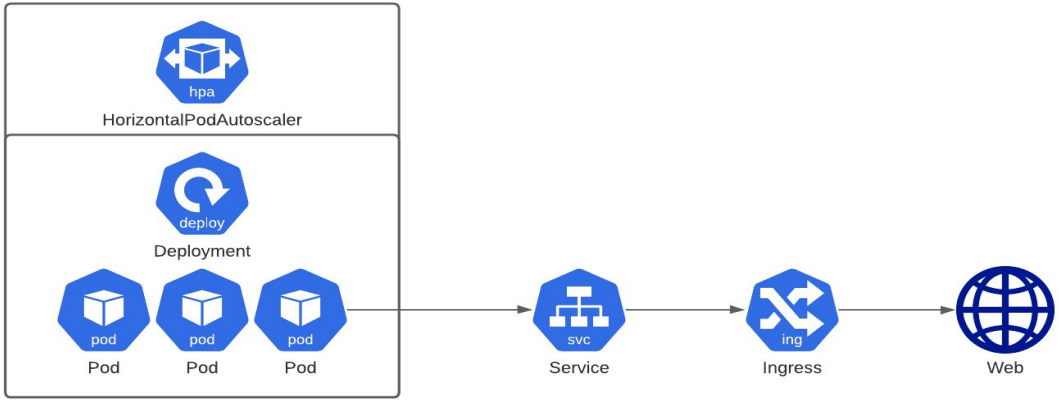


Merits

- Eliminates the need for cvmfs
  - Allows GeoIP federation to work with standard tools like curl and gfal-copy as opposed to stashcp
  - Allows us to do all this in about 60 lines of code as opposed to modifying XRootD source code.
- In HTCondor submit file, transfer_input_files = https://<url>/foo
  - Issue with HTTPS authentication options

Deployment

- Minimalistic deployment built over OSG Software base image.
- Dockerfile
- Pod Port forwarded to Kubernetes Service which is exposed to the web using a Kubernetes Ingress.



Autoscaling

- To avoid a central bottleneck, the redirector can be autoscaled using built-in Kubernetes tools
- Kubernetes Horizontal Pod Autoscaler replicates pods when the defined thresholds are exceeded.
-

NAME	REFERENCE	TARGETS	MINPODS	MAXPODS	REPLICAS	AGE
geoip-http-hpa	Deployment/geoip-http	<unknown>/5k, 0%/50%	1	2	1	23d
geoip-https-hpa	Deployment/geoip-https	<unknown>/5k, 0%/50%	2	4	2	13d

```yaml
metrics:
- type: Resource
  resource:
    name: cpu
    target:
      type: Utilization
      averageUtilization: 50
- type: Object
  object:
    metric:
      name: requests-per-second
    describedObject:
      apiVersion: networking.k8s.io/v1
      kind: Ingress
      name: geoip-https-ing
    target:
      type: Value
      value: 5k
```

Integration with HTCondor (8.9.11)

- Getting files works using gfal-copy and curl:
  - **X509_USER_PROXY=./cert.pem gfal-copy** https://geoip-https.nautilus.optiputer.net//foo file:///bar
  - **curl** https://geoip-https.nautilus.optiputer.net//foo **-o** ./bar **--cert** ./cert.pem
- Condor Integration works only with http.
  - Condor get transfer_input_files using curl but there is no option to specify cert

Use Case (HTTPS)

```
 curl -L -v https://geoip-https.nautilus.optiputer.net//user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf -o /dev/null --cert proxyForAashay
* Server certificate:
*  subject: DC=org; DC=incommon; C=US; ST=California; L=La Jolla; O=University of California, San Diego; OU=UCSD/SDSC; CN=nautilus.optiputer.net
*  start date: Jun  1 00:00:00 2020 GMT
*  expire date: Jun  1 23:59:59 2021 GMT
*  subjectAltName: host "geoip-https.nautilus.optiputer.net" matched cert's "*.nautilus.optiputer.net"
*  issuer: C=US; O=Internet2; OU=InCommon; CN=InCommon IGTF Server CA
*  SSL certificate verify ok.
{ [5 bytes data]
< HTTP/1.1 302 Found
< Content-Length: 431
< Content-Type: text/html; charset=utf-8
< Date: Thu, 25 Feb 2021 00:51:50 GMT
< Location: https://stashcache.t2.ucsd.edu:8444/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf
< Server: Apache/2.4.6 (CentOS) mod_wsgi/3.4 Python/2.7.5
<
* Ignoring the response-body
{ [431 bytes data]

100   431  100   431    0     0  11342      0 --:--:-- --:--:-- --:--:-- 11342
* Connection #0 to host geoip-https.nautilus.optiputer.net left intact
* Issue another request to this URL: 'https://stashcache.t2.ucsd.edu:8444/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf'
*   Trying 198.17.101.66...
* TCP_NODELAY set
* Connected to stashcache.t2.ucsd.edu (198.17.101.66) port 8444 (#1)
* ALPN, offering http/1.1
* Cipher selection: ALL:!EXPORT:!EXPORT40:!EXPORT56:!aNULL:!LOW:!RC4:@STRENGTH
* successfully set certificate verify locations:
  CAfile: /etc/pki/tls/certs/ca-bundle.crt
  CApath: none
* TLSv1.2 (OUT), TLS header, Certificate Status (22):
} [5 bytes data]
* TLSv1.2 (OUT), TLS handshake, Client hello (1):
} [512 bytes data]
* TLSv1.2 (IN), TLS handshake, Server hello (2):
{ [62 bytes data]
* TLSv1.2 (IN), TLS handshake, Certificate (11):
{ [4644 bytes data]
```

In HTCondor (HTTP)

Submit File

```
executable = test.sh

transfer_input_files = https://geoip-http.nautilus.optiputer.net/pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M
error = test.error
output = test.output
log = test.log

request_cpus = 1
request_memory = 2MB
request_disk = 2MB
queue 1
```

Executable

```
#!/bin/bash
echo "Starting";
md5sum test.stashdes.1M;
echo "Done";
```

Output

```
Starting
b6d81b360a5672d80c27430f39153e2c  test.stashdes.1M
Done
```

Correct Hash

HTCondor HTTPS

```
...
028 (247618.000.000) 2021-02-17 14:21:04 Job ad information event triggered.
Cluster = 247618
EventTime = "2021-02-17T14:21:04.407"
EventTypeNumber = 28
HoldReason = "Error from slot1_1@glidein_1497226_453064495@sdsc-69.t2.ucsd.edu: FILETRANSFER:1:non-zero exit (1) from /condor/condor_local/execute/dir_1497207/glide_P5VnYR/main/condor/libexec/curl_plugin.
Error: NSS: client certificate not found (nickname not specified) (https://geoip-https.nautilus.optiputer.net/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf)"
HoldReasonCode = 12
HoldReasonSubCode = 0
JOB_GLIDEIN_ClusterId = "$$(GLIDEIN_ClusterId:Unknown)"
JOB_GLIDEIN_Entry_Name = "$$(GLIDEIN_Entry_Name:Unknown)"
JOB_GLIDEIN_Factory = "$$(GLIDEIN_Factory:Unknown)"
JOB_GLIDEIN_Name = "$$(GLIDEIN_Name:Unknown)"
JOB_GLIDEIN_ProcId = "$$(GLIDEIN_ProcId:Unknown)"
JOB_GLIDEIN_Schedd = "$$(GLIDEIN_Schedd:Unknown)"
JOB_GLIDEIN_Site = "$$(GLIDEIN_Site:Unknown)"
JOB_GLIDEIN_SiteWMS = "$$(GLIDEIN_SiteWMS:Unknown)"
JOB_GLIDEIN_SiteWMS_JobId = "$$(GLIDEIN_SiteWMS_JobId:Unknown)"
JOB_GLIDEIN_SiteWMS_Queue = "$$(GLIDEIN_SiteWMS_Queue:Unknown)"
JOB_GLIDEIN_SiteWMS_Slot = "$$(GLIDEIN_SiteWMS_Slot:Unknown)"
JOB_Site = "$$(GLIDEIN_Site:Unknown)"
MyType = "JobHeldEvent"
Proc = 0
Subproc = 0
TriggerEventTypeName = "ULOG_JOB_HELD"
TriggerEventTypeNumber = 12
...
012 (247618.001.000) 2021-02-17 14:21:04 Job was held.
        Error from slot1_2@glidein_1497226_453064495@sdsc-69.t2.ucsd.edu: FILETRANSFER:1:non-zero exit (1) from /condor/condor_local/execute/dir_1497207/glide_P5VnYR/main/condor/libexec/curl_plugin. Error: NSS:
client certificate not found (nickname not specified) (https://geoip-https.nautilus.optiputer.net/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf)
        Code 12 Subcode 0
```

UC San Diego

Future Work

- cvmfs runs integrity checks on all files, adding something similar to our deployment
- Alternate Solution:
  - Since every pilot has XCACHE_CLOSEST in its environment, the same can be done directly by condor,
    - Having Closest_Cache in the environment of the Job Sites.
    - Having something like transfer_input_files = ENV(XCACHE_CLOSEST)/myinputfile in the submit job to do it directly
- Make sure HTTPS works
- Integration with sci-tokens
  - If job has a token, condor uses token to interact with the web app.

Acknowledgement

The authors would like to thank Brian Bockelman for proposing the idea and Nautilus PRP for giving us the resources to host the pods.

Backup

Use Case (HTTP)

```
curl -L -v https://geoip-http.nautilus.optiputer.net/pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M
* Server certificate:
*  subject: DC=org; DC=incommon; C=US; ST=California; L=La Jolla; O=University of California, San Diego; OU=UCSD/SDSC; CN=nautilus.optiputer.net
*  start date: Jun  1 00:00:00 2020 GMT
*  expire date: Jun  1 23:59:59 2021 GMT
*  subjectAltName: host "geoip-http.nautilus.optiputer.net" matched cert's "*.nautilus.optiputer.net"
*  issuer: C=US; O=Internet2; OU=InCommon; CN=InCommon IGTF Server CA
*  SSL certificate verify ok.
{ [5 bytes data]
< HTTP/1.1 302 Found
< Content-Length: 387
< Content-Type: text/html; charset=utf-8
< Date: Thu, 25 Feb 2021 00:16:58 GMT
< Location: http://stashcache.t2.ucsd.edu:8000/pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M
< Server: Apache/2.4.6 (CentOS) mod_wsgi/3.4 Python/2.7.5
<
* Ignoring the response-body
{ [387 bytes data]

100   387  100   387    0     0   7301      0 --:--:-- --:--:-- --:--:--  7301
* Connection #0 to host geoip-http.nautilus.optiputer.net left intact
* Issue another request to this URL: 'http://stashcache.t2.ucsd.edu:8000/pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M'
*   Trying 198.17.101.66...
* TCP_NODELAY set
* Connected to stashcache.t2.ucsd.edu (198.17.101.66) port 8000 (#1)
> GET /pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M HTTP/1.1
> Host: stashcache.t2.ucsd.edu:8000
> User-Agent: curl/7.59.0
> Accept: */*
>
< HTTP/1.1 200 OK
< Connection: Keep-Alive
< Content-Length: 1048576
* Connection #1 to host stashcache.t2.ucsd.edu left intact
```

Same Command From East Coast (HTTP)

```
curl -L -v https://geoip-http.nautilus.optiputer.net/pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M -o test.stashdes.1M
* Server certificate:
*  subject: DC=org; DC=incommon; C=US; ST=California; L=La Jolla; O=University of California, San Diego; OU=UCSD/SDSC; CN=nautilus.optiputer.net
*  start date: Jun  1 00:00:00 2020 GMT
*  expire date: Jun  1 23:59:59 2021 GMT
*  subjectAltName: host "geoip-http.nautilus.optiputer.net" matched cert's "*.nautilus.optiputer.net"
*  issuer: C=US; O=Internet2; OU=InCommon; CN=InCommon IGTF Server CA
*  SSL certificate verify ok.
> GET /pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M HTTP/2
> Host: geoip-http.nautilus.optiputer.net
> user-agent: curl/7.68.0
> accept: */*
>
< HTTP/2 302
< content-type: text/html; charset=utf-8
< date: Thu, 25 Feb 2021 00:23:24 GMT
< location: http://osg-gftp2.pace.gatech.edu:8000/pnfs/fnal.gov/usr/des/persistent/stash/test.stashdes.1M
< server: Apache/2.4.6 (CentOS) mod_wsgi/3.4 Python/2.7.5
< content-length: 393
<
100   393  100   393    0     0    913      0 --:--:-- --:--:-- --:--:--   913
* Connection #0 to host geoip-http.nautilus.optiputer.net left intact
```

Same with gfal-copy

```
X509_USER_PROXY=/home/users/aaarora/proxyForAashay gfal-copy -vv https://geoip-https.nautilus.optiputer.net/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf file:///dev/null

INFO [gfal_module_load] plugin /usr/lib64/gfal2-plugins//libgfal_plugin_http.so loaded with success
INFO [gfal_module_load] plugin /usr/lib64/gfal2-plugins//libgfal_plugin_srm.so loaded with success
INFO [gfal_module_load] plugin /usr/lib64/gfal2-plugins//libgfal_plugin_gridftp.so loaded with success
INFO [gfal_module_load] plugin /usr/lib64/gfal2-plugins//libgfal_plugin_xrootd.so loaded with success
INFO [gfal_module_load] plugin /usr/lib64/gfal2-plugins//libgfal_plugin_file.so loaded with success
INFO Event triggered: BOTH GFAL2:CORE:COPY LIST:ENTER
INFO Event triggered: BOTH GFAL2:CORE:COPY LIST:ITEM https://geoip-https.nautilus.optiputer.net/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf => file:///dev/null
INFO Event triggered: BOTH GFAL2:CORE:COPY LIST:EXIT
INFO file:///dev/null is a special file (20000), so keep going
INFO Event triggered: BOTH GFAL2:CORE:COPY:LOCAL TRANSFER:ENTER https://geoip-https.nautilus.optiputer.net/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf => file:///dev/null
INFO Event triggered: BOTH GFAL2:CORE:COPY:LOCAL TRANSFER:TYPE streamed
INFO Event triggered: BOTH GFAL2:CORE:COPY:LOCAL TRANSFER:EXIT https://geoip-https.nautilus.optiputer.net/user/ligo/frames/O3/V1Online/V-V1Online-12487/V-V1Online-1248726000-2000.gwf => file:///dev/null
```