

# The Open High Throughput Computing Content Delivery Network

Dave Dykstra

CernVM Workshop

31 January 2018

# WLCG Content Delivery Network

- At the last CernVM Workshop I talked about a new way to find squids for caching content at WLCG sites: WLCG Web Proxy Auto Discovery
  - At <http://wlcg-wpad.cern.ch/wpad.dat>
  - Since then the service has been updated:
    - Now based on CMS & ATLAS squid registrations, cross-checked against registrations in GOCDDB & OIM, plus a few additional ones manually added
    - Different answers can be given for different address ranges in same GeoIP Organizations - used to distinguish CERN Meyrin & Wigner
    - Adds backup proxies @ CERN & FNAL for CVMFS and CMS Frontier services
    - <http://wlcg-wpad.fnal.gov/wpad.dat> replica added
  - Used in production by U.S. CMS opportunistic computing
  - A variation of the service is used in production by LHC@Home at <http://lhchomeproxy.{cern.ch|fnal.gov}/wpad.dat> to find squids at CERN & FNAL in the right order
- WLCG standard extended to look first for <http://grid-wpad/wpad.dat> to provide local override & offload
  - implemented at CERN, including IPv6 support

# Cloudflare CDN

- Cloudflare is a commercial Content Delivery Network vendor
  - Has easy-to-use web interface
    - Works with any domain when they host the DNS
    - Content from any hosted DNS alias to any other server can be cached on their huge network of distributed servers
  - Works for CVMFS & Frontier
  - Has a free tier that allows unlimited bandwidth
  - Includes DDoS protections, IPv6 support, DNSSEC, and more

# AnyCast

- Cloudflare uses a small number of IP addresses automatically routed to their nearest data center with AnyCast
  - No need to abuse DNS caching (as I suggested at the last CernVM workshop) with different responses based on geoIP
  - Requires a lot of effort on their part to set up, to arrange with many ISPs
    - Not something that could be used to direct traffic to squids at many grid sites; needs one entity controlling the distributed network
- Cloudflare directs to the correct origin server based on 'Host' header from client

# Sample measurements - CVMFS

- Some measurements collected on starting up a sample CMS job.
  - CVMFS cache cleared before each run. Client at FNAL. 17600 CVMFS queries, 677MB compressed, 1.5GB uncompressed, usually 2 runs. No squid for cvmfs, but Frontier always cached in local squid.
    - FNAL stratum 1: 8m54s/8m33s
    - FNAL+Cloudflare: 14m54s/10m34s
    - RAL stratum 1: 44m18s (no second run)
    - RAL+Cloudflare: 77m30s/8m30s

# Sample measurements - Frontier

- Same CMS job, same client at FNAL
  - CVMFS pre-cached and no local squid for Frontier.  
2800 Frontier queries, 115MB compressed, 225MB uncompressed.
    - CERN frontier server: 11m48s/10m19s (there is a squid on the frontier server so second run avoids contacting DB)
    - CERN+Cloudflare: 18m27s/3m38s
  - CVMFS pre-cached & Frontier in local squid: 3m11s
- Conclusion: once objects cached in Cloudflare, performance almost as good as local server

# openhtc.io

- New domain hosted in Cloudflare free tier
- For use when no local squids are available
- “Page Rules” set to Cache Everything, Respect server expiration headers
- Domain is set up for long term sustainability, with multiple people having access
- Initial aliases are made for CVMFS stratum 1s and CMS frontier servers. For example:
  - s1fnal-cvmfs.openhtc.io, s1cern-cvmfs.openhtc.io
  - cms-frontier.openhtc.io
- Extendable to other HTC applications

# Use cases

- Now configured for any OSG cvmfs installation that has CVMFS\_HTTP\_PROXY=DIRECT
- Now configured for Frontier for U.S. CMS opportunistic use, when no WLCG squid
  - Hasn't actually run at such sites, but was successfully used to serve a large grid site for a weekend when the local squid was failing
- Working on setting up for CMS LHC@Home, slowly so we can measure impact
- Planned for OpenData
- Planned for CernVM cvmfs default



# Disadvantages

- Much less detailed monitoring
- Uses more WAN bandwidth than on-site squids
- See more hits on origin servers than using squid alone, but much fewer than the number of clients
- Only caches ports 80 & 8080
- Probably will need to disable If-Modified-Since on Frontier; apparently not possible to purge errors
- We don't know if Cloudflare will eventually object or if we will trigger DDoS protection as usage goes up

# Info

- <http://openhtc.io>
- [How does Cloudflare work?](#)
- [How does Anycast work?](#)