



Traffic Switchover 2018 vs 2020

From Varnish to ATS

Emanuele Rocca
Wikimedia Foundation

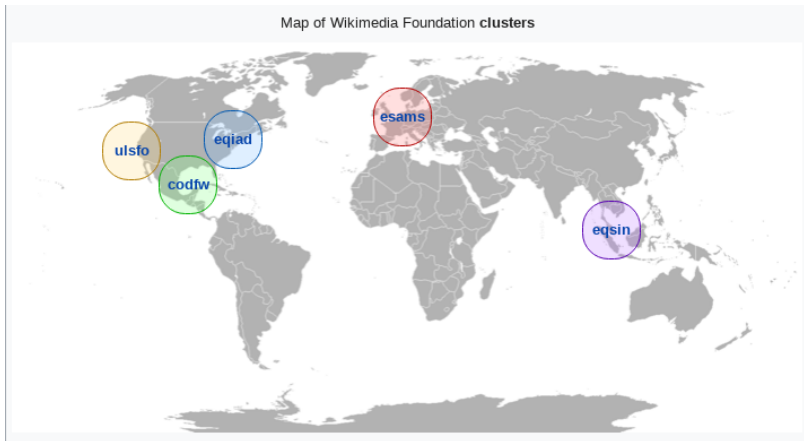
September 21st 2020

How we spent 1 year to save 2 minutes

Outline

- ▶ Old CDN Architecture with Varnish
- ▶ New CDN Architecture with ATS
- ▶ Traffic Switchover 2018
- ▶ Traffic Switchover 2020

Cluster Map



eqiad: Ashburn, Virginia - cp10xx

codfw: Dallas, Texas - cp20xx

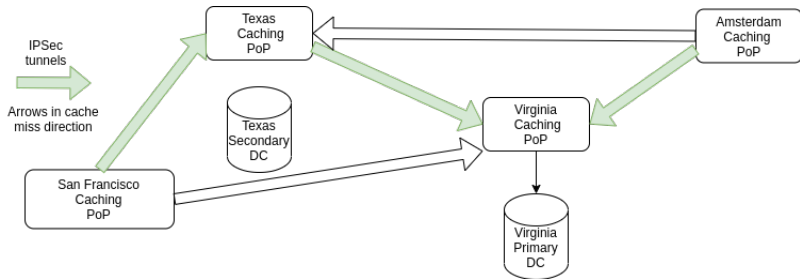
esams: Amsterdam, Netherlands - cp30xx

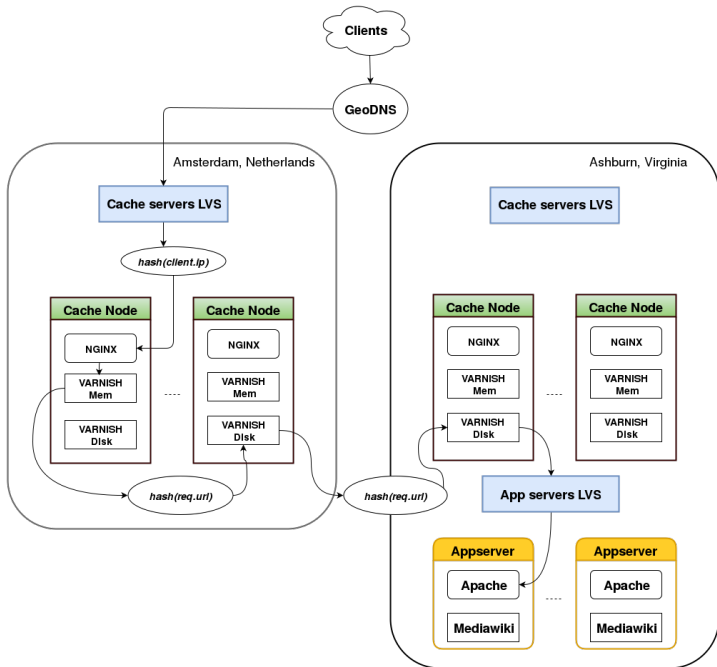
ulsfo: San Francisco, California - cp40xx

eqsin: Singapore - cp50xx

Old CDN (Varnish)

CDN Architecture with Varnish and IPSec





Inter-DC Traffic

- ▶ Encryption between data centers necessary
- ▶ IPsec between cache servers
- ▶ Minimal hitrate on "remote DCs"
- ▶ Architectural constraints due to Varnish not supporting outgoing TLS

Problems with Varnish

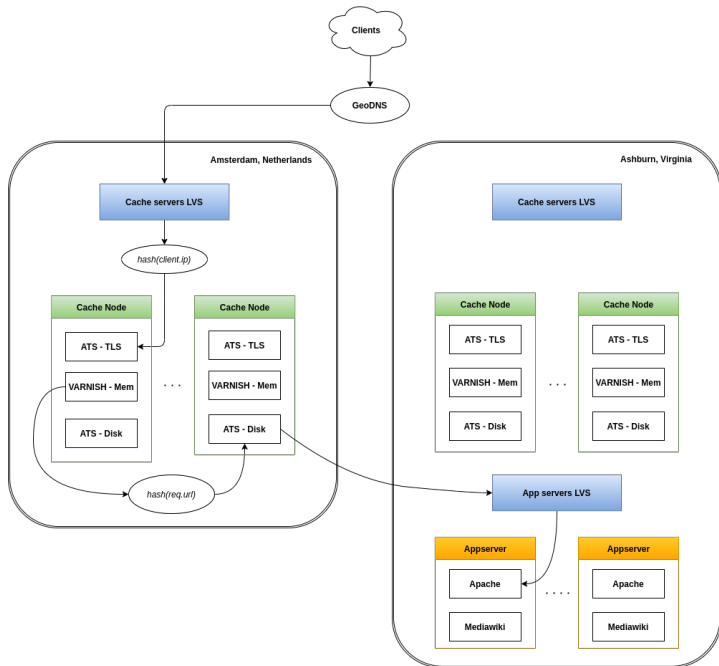
Bad issues with the "file" storage backend

- ▶ Cron job restarting Varnish backend twice a week (true story)
- ▶ Frequent outages, manually restarting services during the weekend on a recurring basis
- ▶ Ephemeral storage

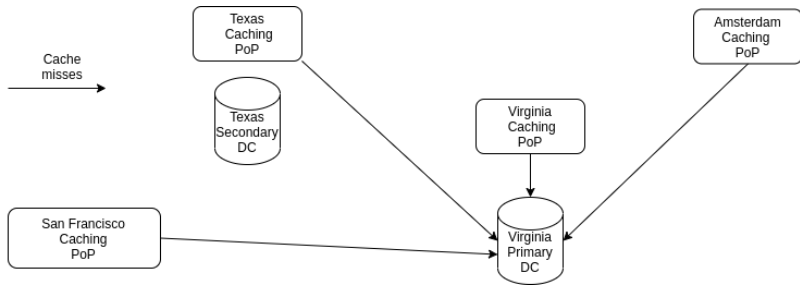
Enter Apache Traffic Server

- ▶ Around since 1996 (before dot-com bubble)
- ▶ Used by Apple, Verizon, LinkedIn, Comcast
- ▶ Reliable persistent storage backend
- ▶ Supports TLS

New CDN (ATS)



CDN Architecture with ATS, no IPsec

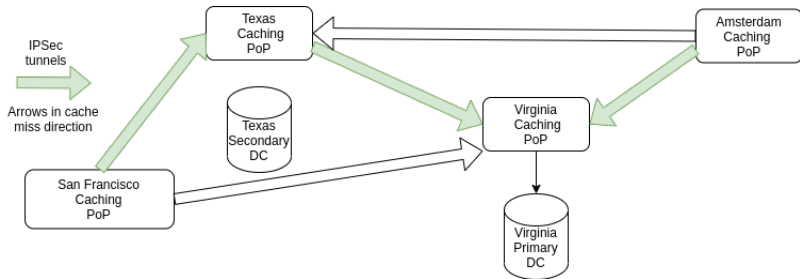


Simpler architecture

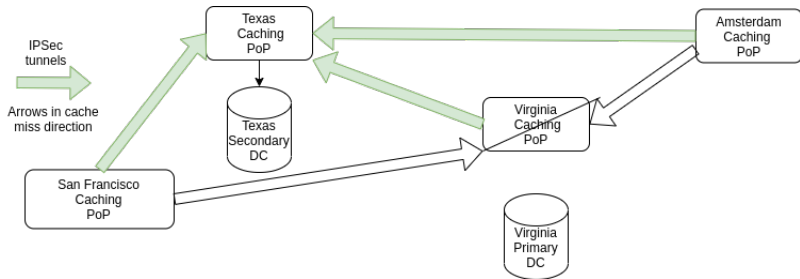
- ▶ IPsec removed entirely thanks to outbound TLS support in ATS
- ▶ No need for caches in different DCs to be aware of each other
- ▶ Saving primary DC caches lots of requests
- ▶ No need to change inter-DC routing when depooling a site
- ▶ Nothing to do during a DC switchover

Traffic Switchover 2018

Switchover 2018 (Varnish and IPsec)



Switchover 2018 (Varnish and IPsec)



Step 1: route esams caches to codfw

In role/common/cache/text.yaml and upload.yaml:

```
cache::route_table:  
  eqiad: 'codfw'  
  codfw: 'eqiad'  
  ulsfo: 'codfw'  
  esams: 'eqiad' # Replace 'eqiad' with 'codfw' here  
  eqsin: 'codfw'
```

- ▶ puppet-merge
- ▶ sudo cumin 'A:cp-esams' 'run-puppet-agent -q'

Step 4: move swift to codfw

Skipping the dance to set swift A/A to simplify, in `role/common/cache/upload.yaml`:

```
swift:
  backends:
    #eqiad: 'ms-fe.svc.eqiad.wmnet' <- commented
    codfw: 'ms-fe.svc.codfw.wmnet'
swift_thumbs:
  backends:
    #eqiad: 'ms-fe-thumbs.svc.eqiad.wmnet' <- commented
    codfw: 'ms-fe-thumbs.svc.codfw.wmnet'
```

- ▶ puppet-merge
- ▶ run puppet on `cp-upload_eqiad`

Step 5: move services to codfw

- ▶ restbase
- ▶ cxserver
- ▶ eventstreams
- ▶ wdqs
- ▶ ...

Step 6: switchover MediaWiki and API

In role/common/cache/text.yaml:

```
appservers:  
  backends:  
    # comment eqiad, uncomment codfw  
    eqiad: 'appservers.svc.eqiad.wmnet'  
    # codfw: 'appservers.svc.codfw.wmnet'  
api:  
  backends:  
    # comment eqiad, uncomment codfw  
    eqiad: 'api.svc.eqiad.wmnet'  
    # codfw: 'api.svc.codfw.wmnet'
```

-
- ▶ puppet-merge with puppet disabled
 - ▶ run puppet on cp-text_codfw FIRST
 - ▶ run puppet on cp-text_eqiad

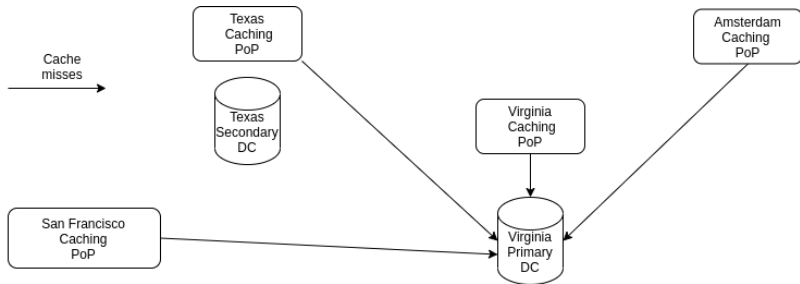
Switchover 2018: MediaWiki and API

```
14:44 MediaWiki: read-only period starts  
14:48 START: - sre.switchdc.mediawiki.04-switch-traffic  
14:50 END: - sre.switchdc.mediawiki.04-switch-traffic  
14:51 MediaWiki: read-only period ends
```

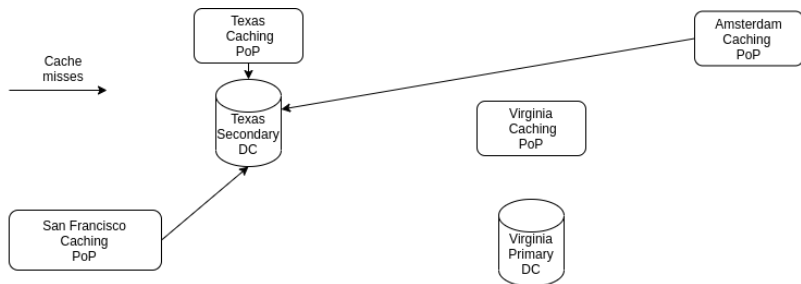
04-switch-traffic.py took 2 out of 7 read-only minutes.

Traffic Switchover 2020

Switchover 2020 (ATS)



Switchover 2020 (ATS)



No puppet change needed, ATS follows DNS and connects directly to the origins via TLS.

Conclusions

Conclusions

- ▶ CDN architecture greatly simplified
- ▶ Removed a frequent source of outages
- ▶ Removed shameful cron job
- ▶ Sunday mornings now spent with family instead of `sudo varnish-backend-restart`
- ▶ DC switchovers made boring

Step 2: set temporary A/A for text A/P

In role/common/cache/text.yaml:

```
restbase_backend:  
  be_opts:  
    port: 7231  
    max_connections: 5000  
  backends:  
    eqiad: 'restbase.svc.eqiad.wmnet'  
    #codfw: 'restbase.svc.codfw.wmnet' <- uncomment
```

- ▶ puppet-merge, run puppet on cp-text_codfw
- ▶ Make very sure the above ran successfully on all nodes or we might get loops

Step 3: set temporary A/A for upload A/P

In role/common/cache/upload.yaml:

```
swift:
  backends:
    eqiad: 'ms-fe.svc.eqiad.wmnet'
    # codfw: 'ms-fe.svc.codfw.wmnet' <- uncomment
swift_thumbs:
  backends:
    eqiad: 'ms-fe-thumbs.svc.eqiad.wmnet'
    codfw: 'ms-fe-thumbs.svc.codfw.wmnet'
```

-
- ▶ puppet-merge, run puppet on cp-upload_codfw
 - ▶ Make very sure the above ran successfully on all nodes or we might get loops

Switchover 2020 ATS config

```
- type: map
  target: http://upload.wikimedia.org
  replacement: https://swift.discovery.wmnet
```

No puppet change needed, ATS follows DNS and connects directly to the origins via TLS.