Embedded CDNs in 2023
- A status

Jac Kloots
Feb 2023

A bit about the CDN market

Growth in the market based on the increased consumer demand for entertainment and OTT services and in the later years on increased demand for e-learning and work from home solutions.

The legacy players like Akamai, Lumen, Tata, Edio etc are well established as reliable and with global reach.

The newer major CDNs like Cloudflare, Fastly and Stackpath are offering a more diverse and specialised set of services like security services and DDOS protection, high performance real time delivery services etc.

And everyone is now looking at edge computing

Large content producers who use comercial CDNs often use multiple CDNs. This both to support their different types of content, to leverage different geographical scope of the CDN and finally - more than one providers means better leverage when negotiating contracts…

Who offers embedded servers?

- Akamai
- Netflix
- Google
- Amazon

- Facebook
- Cloudflare
- CDN77
- Microsoft

- Apple
- Qwilt
- And more to come

Statistics

- Most ISPs in the sample have 6 or less embedded CDNs

- 95 percentile is 6

- The **top 6** are:

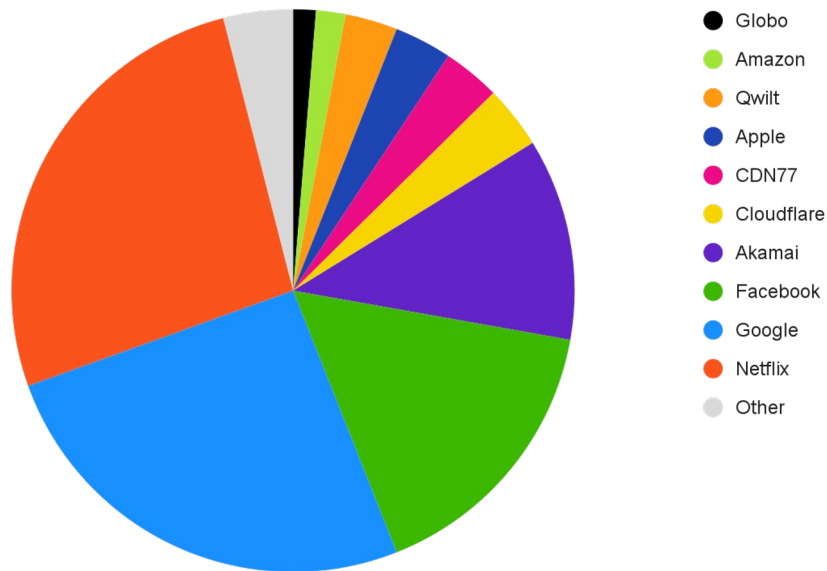|	
Google **(25.5%)**	Netflix **(26.5%)**
Facebook **(16.2%)**	Akamai **(11.6%)**
Apple **(3.3%)**	CDN77 **(3.3%)**

Not quite FAANG but almost ;)

Legend:
- Globo
- Amazon
- Qwilt
- Apple
- CDN77
- Cloudflare
- Akamai
- Facebook
- Google
- Netflix
- Other

Percentages of the ISPs in the sample that have embedded servers from the CDN

Total distribution of traffic marked as CDN traffic

- **~74%** over the external network border
- **~14%** from embedded servers
- **~7%** Embedded to embedded internally in the networks
- **~5%** Externally to embedded servers
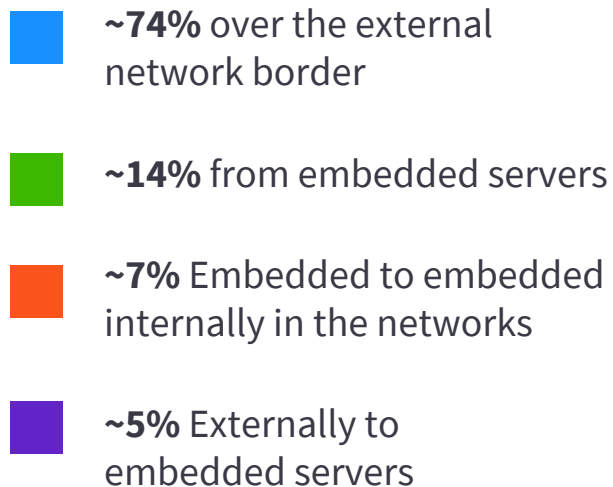
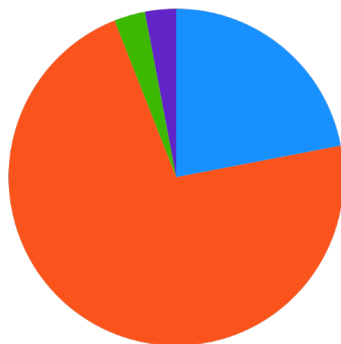*Average traffic over 24 hours that includes a sunday evening.*

Statistics

Google
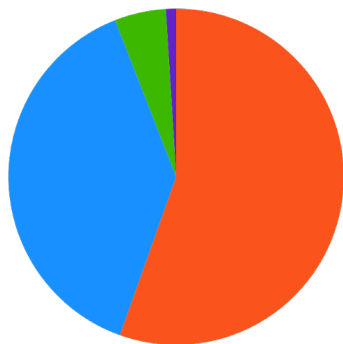
- Outside to end-users **(39%)**
- Embedded to end-users **(56%)**
- Outside to Embedded **(5%)**
- Embedded to Embedded **(1%)**
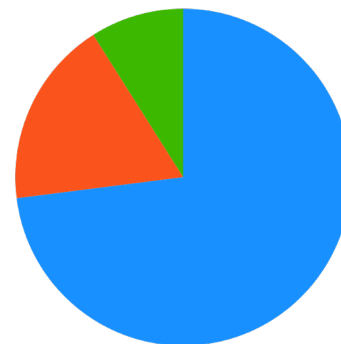
Netflix

- Outside to end-users **(22%)**
- Embedded to end-users **(72%)**
- Outside to Embedded **(3%)**
- Embedded to Embedded **(3%)**

Facebook

- Outside to end-users **(43%)**
- Embedded to end-users **(54%)**
- Outside to Embedded **(4%)**
- Embedded to Embedded **(0%)**

Akamai

- Outside to end-users **(73%)**
- Embedded to end-users **(18%)**
- Outside to Embedded **(9%)**
- Embedded to Embedded **(0%)**

Steer to the closest cache

BGP

- BGP is used to signal the subnets that a cache or cluster should serve

- The most control is given for the systems where every cache has a BGP session and the steering system respects common BGP parameters.

- A setup where all Servers have an identical session will work for most deployments.

- Prefixes are sent from the ISP to the cache

DNS Server

- End-users are mapped to a cluster based on the DNS server they are using to request the content

Anycast

- Anycast addresses are announced from the caches and the ISPs routing decides which is the closest

Geolocation

- Geolocation is used by some as the primary way if mapping an end-user to a cluster. Some use the geolocation of the end-user IP address, but some rely on the geo location of the DNS server the end-user is using to request the content.

And then some of the magic

Latency and other QoS data

- The CDN creates a mapping of ISP DNS servers to POPs or clusters based on latency measurements
- Some CDNs includes internet connectivity from the cluster to the DNS servers or IP addresses in their mapping algorithm
- The content owned CDNs also include signals from the content clients in the algorithm

Load

- The load of the individual servers or clusters is part of the decision of exactly which server the end-user will end up using.
- Load balancing can take place in the cluster or in the CDN response to (where is?)
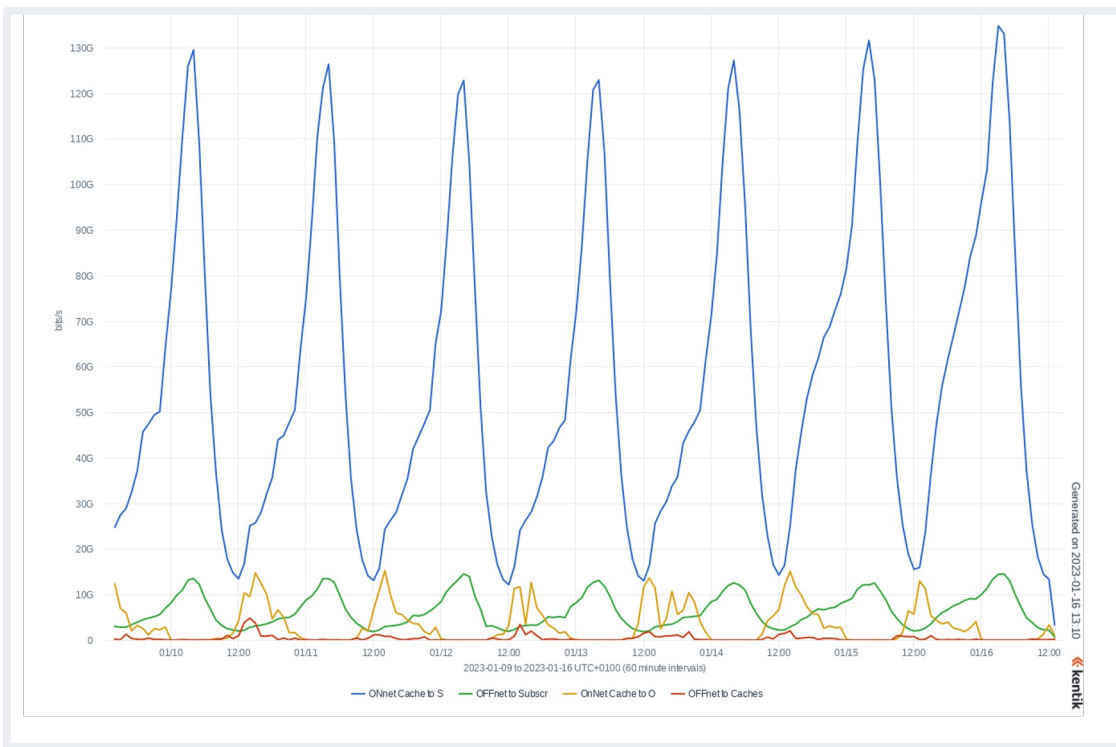
Content

- Not all content is on all servers
- Content that is pushed out is distributed according to CDN magic
- Not all content from a CDN will be cached when the content is pulled
- This means where the content is/can be is also part of the steering decision.

Traffic profiles for different ways of placing content
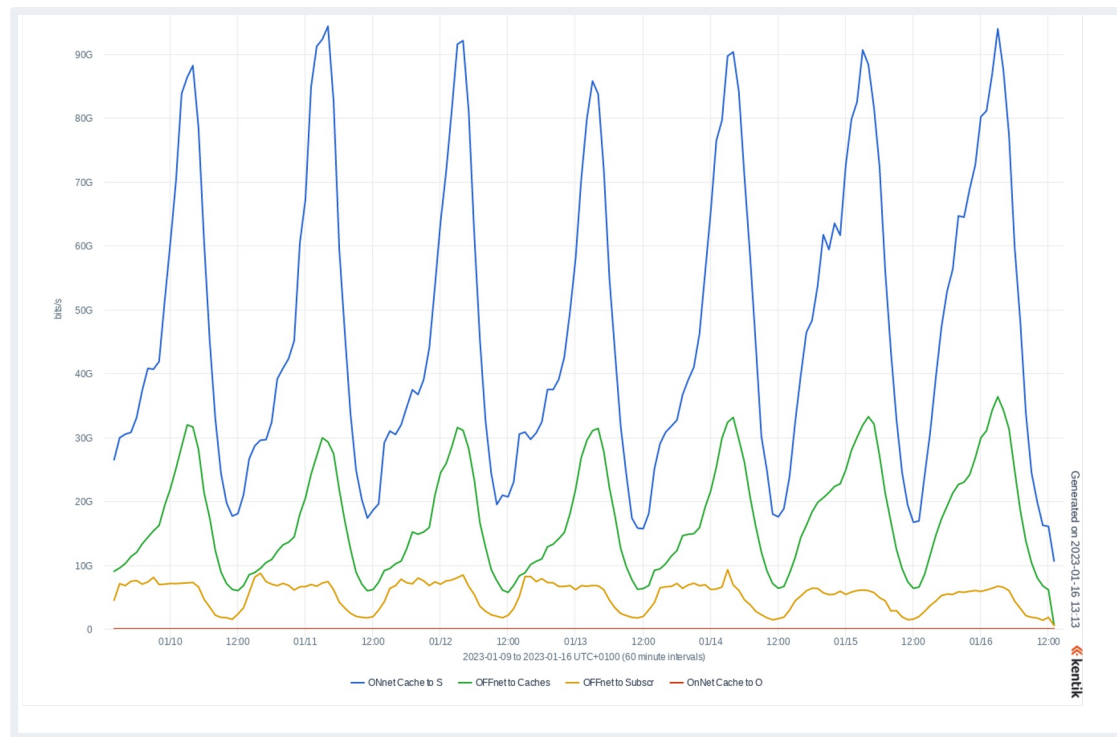
- Centrally calculated placement

- Fill window

- Optimized to reduce edge fill traffic as much as possible

Typical Netflix traffic profile in well dimensioned deployment

Traffic profiles for different ways of placing content

- Proxy and akamai magic

- Relatively dynamic content

- Not all content will be at the embedded servers

Typical Akamai traffic profile in well dimensioned deployment
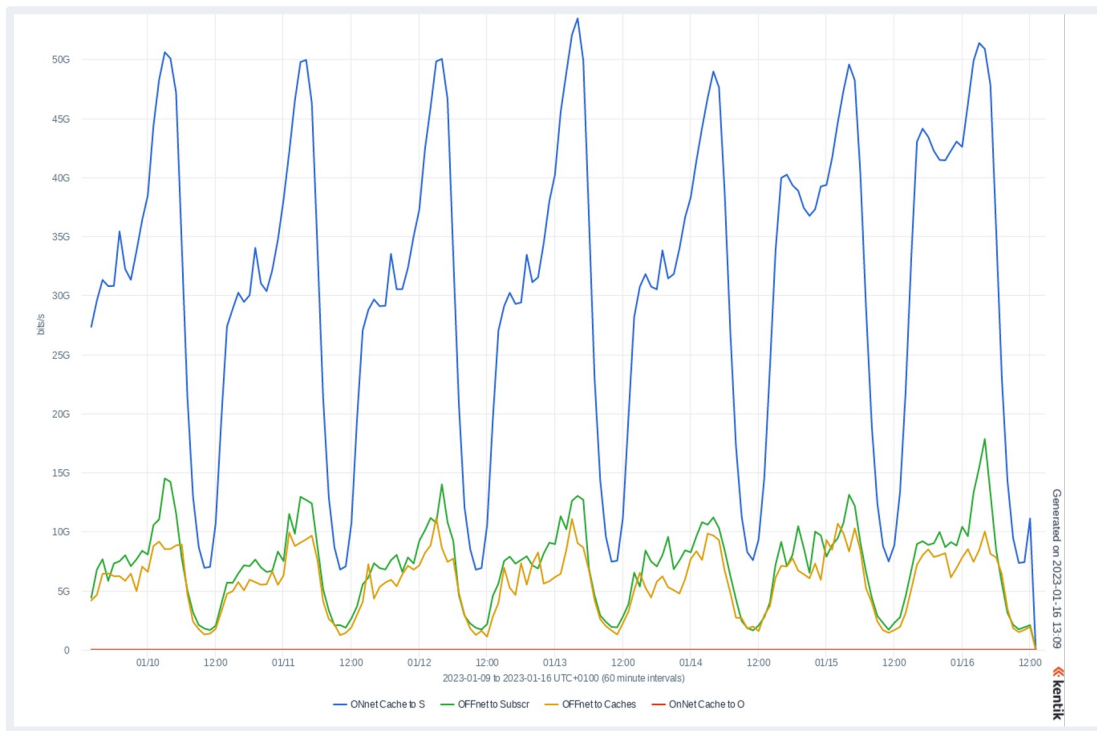
Traffic profiles for different ways of placing content

- Proxy and Facebook magic
- Relatively dynamic content
  - *Only the static is cached*
- Not all content will ever
  be at the embedded servers

Feeds

Likes

Anything that needs to
be globally consistent

Typical Facebook traffic profile in well dimensioned deployment

Deploying embedded servers deep into the network

- Understand if the end-user mapping is supported by your network implementation
  - DNS servers
  - IP address plan
  - Use of CGN
- Be aware of clustering
- Off load/hit rate for the local cluster might be lower than for a big cluster because of less disk space in the cluster

Is it easier or harder in 2023 compared to 2012?

2012

- Large boxes
- Off load unpredictable
- Few players
- No CGN support

2023

- Small boxes
- More predictable offload
- More predictable and flexible end-user mapping
- Many players
- Some CGN support

kentik®

THANK YOU

Jac Kloots
https://www.linkedin.com/in/jackloots/