

Dos al Cubo decreased cloud costs by 30% with **Varnish**

Background

Dos Al Cubo, a digital media technology company based in Argentina, has spent the last 20 years developing 360° solutions for all stages of the digital media evolution. This includes both the Thinkindot content management system (CMS), designed for journalists and the digital media industry as a whole, as well as consulting on process and technology organization, editorial strategy, audience-generation optimization, and customized technology offers.

The challenge

Redundancy and resilience

A key aspect of Dos al Cubo's consultancy is to ensure performance, speed, and stability for websites using the CMS in even the highest traffic peaks. At the same time, they needed more control over their caching and cache invalidation strategies and the ability to add an origin shield behind their CDN.

The Dos al Cubo team already used Varnish Cache for some of these use cases but some of the improvements they wanted to make required a more robust solution and Varnish support.

Dos al Cubo at a glance

Organization

 Dos al Cubo is a leading provider of content management systems (CMS) for mass media.

Challenge

- Launch an origin shield solution for improved service availability and faster response times for uncached content
- Control TTL at origin and ensure cache persistence
- Manage complex purging patterns

Varnish Enterprise

- Varnish Enterprise and Varnish Support
- Varnish Enterprise as origin shield
- Professional support from Varnish core developers for integration and incident support
- Varnish High Availability
- Varnish Massive Storage Engine
- Varnish Broadcaster
- Varnish Controller
- A range of VMODs, such as brotli, ykey, headerplus and purge



The solution

The need to solve several challenges prompted the choice to adopt Varnish Enterprise:

- · Improvement of cache efficiency and persistence
- Flexibility of available purge methods
- Single API to purge content across every node in the cluster
- · Cache replication across cluster nodes
- VCL replication across cluster nodes

The team decided to build on their work with Varnish

Cache and existing VCL cache policies by moving forward with Varnish Enterprise. They focused on working with Varnish core engineers to set up their Varnish cluster and enterprise features, such as cache storage configuration, replication and purge API. They also worked on delivering a failover solution and self-healing of each node across the cluster.

To be the perfect partner for publishing companies in supporting the success of their content strategies and achieving relevance and performance, we needed to take caching and origin shielding to the next level.

-Leandro Armas, CEO. Dos al Cubo

Architecture and failover solution

The Dos al Cubo implementation features a multi-region Varnish Enterprise cluster that works as an origin shield in the middle of a third-party CDN and their app servers. This infrastructure is hosted across multiple US regions, and each region has its own autoscaling cluster of Varnish Cache nodes, without affecting latency thanks to cache replication. Failover events are managed with health checks and DNS rerouting policies at two levels:

- In a CDN outage, traffic is rerouted directly to the Varnish Enterprise cluster
- In a complete outage on a specific region of the Varnish cluster, traffic is rerouted to the closest region.

We found that every component of the Varnish Enterprise solution has great documentation, which supported our smooth setup along with the help of the Varnish Team.

> -Leandro Armas, CEO, Dos al Cubo

Origin shield to circumvent outages

As a high-performance, flexible solution, Varnish Enterprise as an origin shield was placed behind Dos al Cubo's CDN to improve their service availability in case of a CDN outage. The origin shield would ensure that content could continue to be served even in outage conditions, and response times for serving uncached content would be improved.

Controlling TTL: Persisting objects in cache at origin

With a third-party CDN, Dos al Cubo could not guarantee that objects would persist in cache for the specific TTLs they specified at origin, as is typical for a multi-tenant CDN solution. The origin shield with Varnish put control back into their hands.

Better cache and purge control

Working with news and media sites, fresh content is constant. Content is being created, updated, and deleted in a nonstop flow. With that level of content generation and a caching policy living in front, Dos al Cubo needed to be able to purge the specific modified content and every piece of related content efficiently at every update. Varnish enables purging by the specific path, headers, or cache keys/tags.



Results

Varnish Enterprise: Using fewer resources for better performance

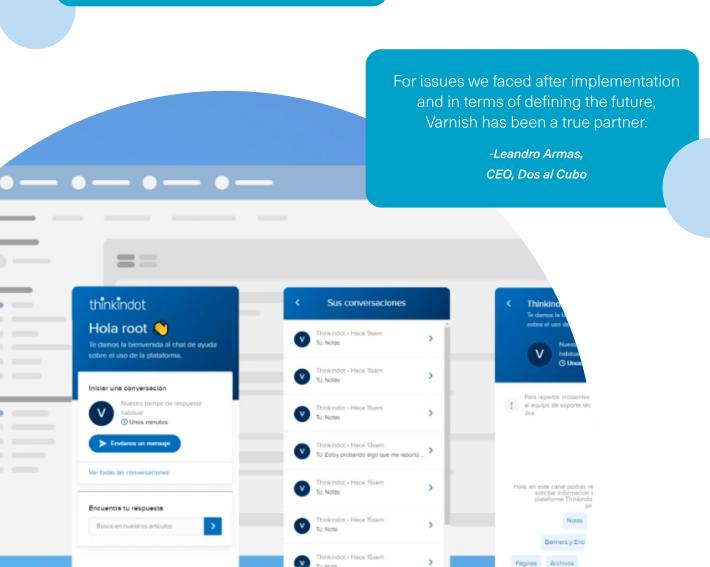
Varnish Enterprise delivered control and resilience for Dos al Cubo, which extended their ability to provide uptime and performance to their customers. Varnish has delivered on Dos al Cubo's expectations.

Dos al Cubo has been able to progressively decrease the number of resources dedicated to their frontend app servers without losing the capacity to add new customers and maintain an average resource usage level of 30% CPU capacity.

The Varnish rollout boosted efficiency with fewer resources without affecting performance.

In terms of costs, comparing our consumption in May 2021 versus August 2021, we decreased our app server cloud costs by 46.51%. If we take Varnish cluster costs into account, we decreased this by almost 30% during these three months.

-Leandro Armas, CEO, Dos al Cubo At the same time, Varnish flexibility has increased other efficiencies, as the CMS requires specific headers to pass from their CDN. One example is the user's geolocation information. Instead of managing the geolocation database in their origin, Dos al Cubo uses simple Varnish VCL logic to catch every CDN header and send a unique one to their CMS as well as generate that information in Varnish directly using available VMODs. This results in the added benefit of being able to make their CMS work the same way regardless of the CDN provider.



Tighter content control and performance with Varnish

In the future, Dos al Cubo would like to see more feature richness added to Varnish Controller to enhance their control over cache node management and more granular historical and sampled Varnish statistics for each node.

The setup of our Varnish cluster spans four AWS regions in the US. Considering that we have nodes on both US coasts, one of our main concerns was how Varnish High Availability and Broadcaster would work with replication and latencies between 70 to 100ms. These latencies didn't affect them at all, and the cache replication works as expected.



