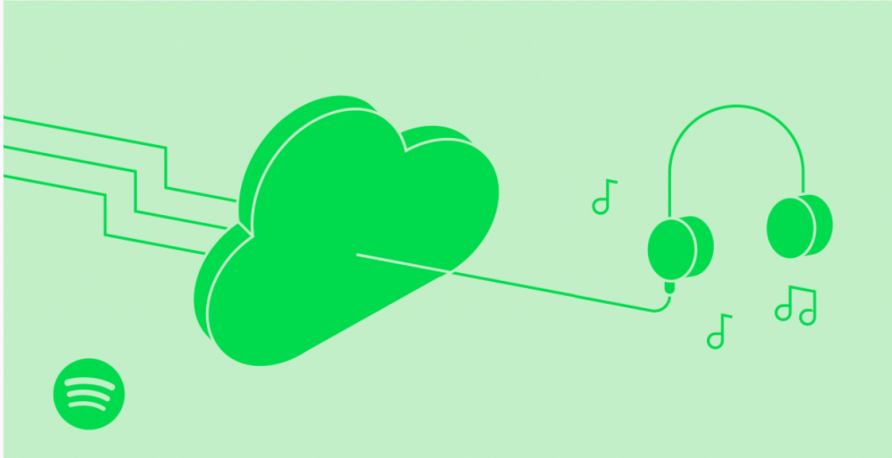




Views From The Cloud: A History of Spotify's Journey to the Cloud, Part 1



December 9, 2019
Published by Niklas Gustavsson



Spotify's Chief Architect, Niklas Gustavsson, was at the heart of the company's journey to migrate its data centers to the Google Cloud Platform. In the first of two articles, he tells his story of the migration and how it enables engineers to focus on delivering a better audio experience for customers.

Life before the cloud

It might come as a surprise that Spotify wasn't always a cloud business. However, it's important to remember we were founded in 2006 and launched two years later, at a time when cloud solutions from providers such as Amazon, Google, and Microsoft were only just getting off the ground.

Fast forward a few years and our on-premise database clusters had expanded to meet customer demand, handling some 2,000 services; 20,000 daily data pipeline runs; and more than 100 Petabytes of data.

As the technology expanded so did our number of employees. There were about 100 teams managing their systems, hosted in four geographical regions around the world. Each region comprising one or more collocated data centers. We were getting to the size where we had to make serious investments into building out our data center infrastructure.

We are fundamentally in the music business and not to build data centers. Our goal is to make all the songs in the world—and now podcasts—available to our users. This audience doesn't get any value out of us running a massive data center operation. And we want our engineering teams to be able to focus on building the next big thing at Spotify rather than spending their time building low-level infrastructure. We want to continuously move up the stack.

Thus, in early 2015 we started exploring what a cloud strategy would look like for Spotify.

Getting to the cloud

In the initial planning phases, we had to choose the cloud hosting model and migration strategy that was right for us. Let me give you an overview of our approach, and how we drew up our plan.

We were positioned to make a decision between three options: work with multiple cloud providers, move to a hybrid setup with some combination of cloud providers and on our own data centers, or go all-in with a single cloud provider.

Working with multiple providers is great for minimizing lock-in effects with any single cloud, but means you need to invest in abstractions across multiple providers. This was something that we wanted to avoid because it can prevent you from 'moving up the stack' and getting a greater return on your investment.

The hybrid option is usually advocated when you want to retain on-site ownership and control of your data (e.g. due to regulatory reasons), or when you have systems that would be infeasible to move to a cloud environment. Neither case applies to Spotify.

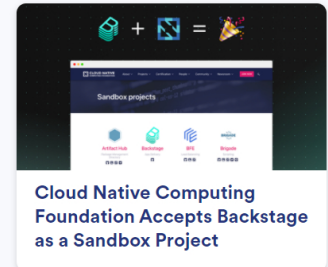
SHARE



CATEGORIES

Infrastructure

Related articles



@SpotifyEng on Twitter

Tweets by @SpotifyEng



Spotify Engineering
@SpotifyEng

The recipe for efficient, scalable audio pipelines? #Python, @ApacheBeam and a cloud. Read more about [Klio.io](#), our open framework for processing media at scale, on the [@GCPcloud](#) blog featuring @roguelynn: [cloud.google.com/blog/products/...](#) <https://twitter.com/GCPcloud/status/1334303207107989511>



Simplify creating data pipelines ...
Spotify open-sources Klio: scalabl...
[cloud.google.com](#)



16m

Having evaluated the potential outcomes, we decided on the third option, an 'all-in' move to just one provider where we could build a deeper working relationship that went beyond simply offloading infrastructure to a third party. This led us to commit to the Google Cloud Platform (GCP).

Supporting our engineers

Now, having 100 teams move thousands of components into the cloud was an ambitious plan, and you can't possibly hope to achieve these goals if you're not aligned. So it was particularly encouraging to see the approach of our infrastructure teams. They ended up being one of the loudest voices behind a project that, after all, would fundamentally reshape their role at Spotify.

At this time, GCP was still a nascent platform, and together with Google, we identified a set of missing features to support a large and complex customer like Spotify. Examples included scalable VPN and IAM/project organization. We worked closely with Google engineers to close these gaps, leading to some of the products that we still use heavily, for example, Shared VPC.

We also built a use case where we demonstrated the ability to both stream music and process data to calculate royalty payments. Doing so further increased confidence in the migration at all levels of the organization.

Lift and shift or rewrite?

As described above, operating a hybrid solution at Spotify's scale is complex and comes with some inherent risks. As we would need such a setup during the migration, we strived to balance between keeping the migration period as short as possible, while remaining confident in our ability to operate the service without disruption.

We, therefore, used a mix of lift and shift, moving components without any redesign, and rewrites where needed.

Our approach here was highly pragmatic. The first part of the migration focused on user-facing services and here we used lift and shift throughout to ensure that there was no disruption to music streaming at any time. In our capacity provisioning systems, our data centers and our new GCP regions would look and behave alike, making it straightforward for teams to deploy their services in both with minimal impact on productivity.

The second workstream involved migrating our data processing. Teams had the option to rewrite code if they had enough time. This again ensured that we hit our project timelines while delivering an uninterrupted service to customers. Of course, teams still have the option to rewrite code at a later date when workloads allow.

Focus, focus, focus

The project was managed by a small team of about half-a-dozen Spotify engineers who collaborated with a team from Google. They had three key tasks:

- Build a visualization of the migration state: This was a simple color-coded, real-time representation of the migration state of different operations. We found this was a great way of motivating people. Once you've completed a task, you could immediately see the result of your efforts in the visualization.
- Build a standardized migration sprint program: This boiled down to a two-week sprint for all engineering teams. During this time, they would focus fully on migrating their systems.
- Create teaching material for common migration cases and our cloud infrastructure.

So how did it go? Remarkably smoothly. There were times where we needed the support of the team at Google to work through some of the challenges. It's probably fair to say that in projects of this scale, the partner needs to be flexible with their approach – and their product. Working closely with Google, we were able to close the gaps and enhance their platform to meet our requirements along the way.

Maximizing success, minimizing stress

By May 2017, we had achieved each of the migration sprints and traffic was fully routed towards GCP. In December 2017 we closed the first of our four on-premise data centers and the remaining three were retired in 2018.

To sum up, moving to the cloud allowed us to move away from on-premise infrastructure and move up the stack. We can innovate faster, by building services in this environment and finding better ways to build solutions that benefit our massive user base, taking full advantage of machine learning, data processing, and other opportunities.

In my next article, I'll look in more detail at how the Google Cloud Platform empowers our engineers and some of the innovative systems that we've been able to design and build in this environment.

Spotify Engineering Rewritten



#2020Wrapped is finally here!
spotify.com/wrapped



Embed

View on Twitter

SHARE THIS:



Sign up for engineering updates

By clicking sign up you'll receive occasional emails from Spotify. You always have the choice to adjust your interest settings or unsubscribe.

Sign up!



[Spotify Jobs](#)
[Spotify.com](#)
[Newsroom](#)

[Spotify R&D Research](#)
[Spotify R&D Design](#)

[Legal](#) [Privacy](#) [Cookies](#) [About Ads](#)

© 2020 Spotify AB