SoundCloud Used Pythian as an Extension of Its Team, Bringing Support to Internal Teams Faster

The Production Engineering team at SoundCloud uses Pythian for on-demand support. This means SoundCloud has the resources to support its internal teams and work on business-wide customization and upgrades.

SoundCloud is a music entertainment company based in Berlin, Germany. The company is known for letting people discover and enjoy music from a diverse community, and its creators are at the core of that mission. Since it was founded in 2008, SoundCloud has become the world's largest music and audio platform—and it's on a quest to become the world's biggest social sound platform.

More than 100 million people in the United States alone listen to music on SoundCloud. The platform is also used by content creators and businesses who own content. SoundCloud's goal to become an even larger platform is driven by a desire to enable creators to make a living through the platform.

A capacity-strapped central team couldn't respond to internal customers quickly enough

SoundCloud runs on tight capacity, which meant the Production Engineering team was too capacity-strapped to respond to internal customers quickly enough.

Like many high growth companies, SoundCloud has a lean but nimble technology team. When business runs as expected, that works well. But when teams need help unexpectedly, this model makes it hard to provide the necessary support.

A small capacity constraint can be fixed by hiring a new team member. That solution works once or twice, but it's not scalable. And that's exactly the predicament that SoundCloud faced. "How do we get our feature teams unblocked?" Jake Maizel asked.

Pythian

Jake Maizel is the Vice President of Infrastructure at SoundCloud. His teams are responsible for the relational database design, which is based in MySQL.

You build it; you run it

At SoundCloud, teams are primarily responsible for their own databases. If a team needs a database, they build and run it. The teams manage all aspects of their databases, including scaling, schema changes, query performance, optimization, outage planning, capacity planning—everything. Then, if something unusual comes up—such as performance problems or complications building a new database—they consult Jake's central team.

The central team, called Production Engineering, comprises the veterans who support other teams on specific projects within the company. They also build tools to improve the efficiency of the relational databases. They're also responsible for any customization that's needed for the environment.

When the Production Engineering team couldn't respond to internal requests for help as quickly as they wanted or needed, they turned to Pythian for help.

Elasticity to handle fluctuating support needs

- Provided on-demand consulting services
- Gave SoundCloud the elasticity to handle fluctuating support needs
- Created a Slack channel that allowed SoundCloud to ask questions as they arose

SoundCloud looked at Pythian's standard support options, which include things like running a database. For international companies like SoundCloud, compliance with GDPR and EU data privacy laws can present additional challenges. Although Pythian's range of expertise includes experience working with GDPR, ensuring compliance with additional international regulations would still be a hurdle for SoundCloud. Overall, there wasn't an out-of-the-box support option that jumped out as being a perfect fit.

So, Pythian suggested something different.

"Pythian created a model where we could subscribe to hours over a period of time. The Pythian team is like an extension of our team. They're like an on-tap consulting company,"

– Jake Maizel, VP of Infrastructure

Now, SoundCloud relies on Pythian for its general knowledge. Pythian functions like an extension of the Production Engineering team – without eating into valuable resources. To be available in that on-demand manner, Pythian and SoundCloud set up a Slack channel. "We tap into Pythian's expertise and ask questions about how we should do things," said Jake Maizel.

Improving read latencies from 250ms to 40ms

One challenge where SoundCloud asked for support was decreasing the high read latencies on the 99th percentile, which were about 250 milliseconds. Although Cassandra can have high latencies for reads in some cases, 250ms is unusually high and indicated that something else was causing problems. After investigating the issue, Pythian found several opportunities for improvement. To optimize the cluster performance, Pythian:

- Disabled the incremental repairs in Cassandra

- Disabled the table compression, which wasn't efficient and was generating compression/decompression overhead

- Changed Java garbage collection to G1 GC and increased heap size

- Tuned OS kernel settings

- Changed the read_ahead_kb to 8

After tuning the cluster, the read latencies on the 99th percentile improved substantially—from 250ms to 40ms.

Optimized performance, reduced latencies and upgraded databases

- **Optimized the cluster performance.** When the SoundCloud team implemented Cassandra Reaper, they faced several issues making the tool operational. The Pythian team stepped in, and now Cassandra performs anti-entropy repairs without issues.

- **Enabled internal speed by improving tail read latencies**. Pythian improved the read latencies on the 99th percentile from 250ms to 40ms. This allowed the SoundCloud team to provide internal customers the support they needed in a timely fashion.

- **Changed a cluster from a single datacenter to a multi-datacenter topology.** Pythian increased availability for a cluster by helping provision a new datacenter. The team did this while adhering to strict topological constraints.

When SoundCloud contacted Pythian, it had two clusters that were underperforming. SoundCloud wanted to change the topology of one cluster, which was tricky because of its Cassandra settings.

SoundCloud was also curious about whether a different Snitch could be used, which would help avoid storing multiple replicas of data on the same rack. The engineers asked Pythian to investigate whether their driver logic could be set up to use a different consistency level for a multi-coast deployment.

Optimizing the cluster performance

[Many of Cassandra's default settings for repairs are not ideal. This can be overcome by using Cassandra Reaper, an alternative tool that doesn't come with Cassandra](#). The repairs are essential in keeping the data consistent and in sync when spread across multiple nodes. Ideally, the process is automated.

When the SoundCloud team implemented Cassandra Reaper, they faced several issues making the tool operational. The Pythian team stepped in; now the process is automated, and Cassandra performs anti-entropy repairs without issues.

By fixing and tuning the Cassandra properties and the environment surrounding Cassandra, Pythian helped SoundCloud improve the cluster performance. To optimize performance on several fronts, Pythian:

- Tuned the OS kernel settings

- Disabled incremental repairs

- Fixed Cassandra Reaper issues

Enabling internal speed by improving long-tail read latencies

After Pythian optimized compressed read latencies from 250ms to 40ms, the SoundCloud team was better situated to meet internal deadlines. By improving the long-tail read latencies in the drakkar-upload cluster, Pythian helped SoundCloud comply with the SLAs. SoundCloud was once again able to provide internal customers the support they needed in a timely fashion.

Changing a cluster from a single datacenter to a multi-datacenter topology

SoundCloud also wanted support switching from a single datacenter to a multi-datacenter topology. The Pythian team looked at the architectural reference inherent to Cassandra, which dictates how it distributes data and requests. The team also looked at how the Cassandra clients communicate with the datacenter to see if it could be optimized to serve global users more effectively. Specifically, how could they increase availability and benefit from using multiple datacenters?

Pythian increased availability for a cluster by helping provision a new datacenter. This was done while adhering to strict topological constraints.

Typically, adding a new datacenter to Cassandra is straightforward, but clusters using RackInferringSnitch have added complexities. This Snitch implementation infers topology from IP addresses, which meant the Pythian team had to make sure the nodes were distributed evenly among racks against the rack inferring. There should also be strict controls on IPs when new nodes are added to a cluster. Without that, the racks or data centers will be created unevenly.

To improve cluster performance and fix the Cassandra anti-patterns, the Pythian team:

- Changed read/write consistency from one/all to local_quorum
- Added a new datacenter to the cluster
- Added nodes to the new datacenter when RackInferringSnitch was being used

Helping backend teams configure consistency levels

Pythian also investigated the consistency levels in Cassandra. Pythian created a report explaining how the DataStax Java driver for Cassandra could be optimized while using the multi-cluster. This helped SoundCloud benefit from the high availability.

Providing internal teams the support they need, whenever they need it

The Production Engineering team at SoundCloud is strong, but they don't have unlimited resources to support all the teams that need help.

By working with Pythian, SoundCloud improved performance on multiple fronts. Pythian helped SoundCloud increase data consistency by helping set up Cassandra repairs and avoiding incremental repairs. Neither issue is substantial on its own, but both can cause performance issues later. By adding a new data center and refactoring consistency levels, Pythian helped make SoundCloud's data more available.

Using the Pythian team as on-tap consultants means SoundCloud can provide its current and future internal teams the support they need, whenever they need it.

Hiring additional team members isn't a scalable solution, especially as SoundCloud continues to grow. Bringing new team members on board can be time consuming and expensive. With Pythian, the Production Engineering team has the expert support they need without the burden of hiring and training new colleagues.

Even though SoundCloud has seen consistent growth, internal production needs can be more variable. Working with Pythian gives SoundCloud the flexibility to use as much consultancy support as they need – without being burdened to the cost of more support than they use.

Most importantly, because the Production Engineering team has Pythian's on-demand support, they have the resources to support their own internal teams and to work on customization for business-wide customization and upgrades.

linkedin.com/company/Pythian   twitter.com/Pythian   Contact us at +1-866-798-4426 or info@pythian.com

ABOUT PYTHIAN

Founded in 1997, Pythian is a global IT services company that helps organizations transform how they compete and win by helping them turn data into valuable insights, predictions and products. From cloud automation to machine learning, Pythian designs, implements and supports customized solutions to the toughest data challenges.

© Pythian Services Inc. 2022

OFFICES

Ottawa, Canada   New York City, USA   Minneapolis, USA   London, England   Hyderabad, India

Pythian