intel. tiber.
App-Level Optimization

Nylas Case Study

Nylas Improves Latency by 58% and Reduces Cloud Costs by More Than 35%

Watch now →

nylas

The Challenge

Infrastructure: Python on GCP

As a leading business automation infrastructure solution, improving API performance and providing the best quality of service is a crucial KPI for Nylas. Nylas handles billions of API requests per day on GCP and has a constantly growing environment. As Nylas' user base is growing by 2-4x year over year, they needed a way to dramatically improve performance quickly with no code changes or R&D efforts.

Their API platform provides real-time bi-directional sync with user's email , calendar and contact data, with an intelligent automation process built on top. Nylas' Sync service is the key component of the platform. When an account is connected to Nylas, which happens thousands of times per second, the sync engine starts pulling in email messages for the account, pulling new message arrivals in parallel, and prioritizing the most recent ones to achieve a better user experience.

As a core piece of Nylas' API puzzle, the Sync service is required to operate quickly and efficiently. That means that it is extremely sensitive to performance and requires an aggressive SLA of "time to sync" or latency.

The Sync service runs on GCP and includes ~200 running instances and over 4K cores.

35% Cost Reduction

58% Reduced Latency

35% Increased Throughput

15% Reduced CPU Utilization

nylas

SaaS
HQ: California, USA
Employees: 200+

Nylas is the Communication Platform-as-a-Service (CPaaS) company that empowers businesses worldwide to unlock the true power of their communications data.

Nylas simplifies developers' lives by enabling them to automate manual, repetitive, everyday tasks with little to no code.

> "It never occurred to us that we might improve performance so much that reducing cost was an option. But with (Intel Tiber App-Level Optimization), we were able to leverage these results into a cost reduction of 35%."
>
> **Caleb Geene, Staff Site Reliability Engineer**

Why Intel® Tiber™ App-Level Optimization

Nylas searched for a solution that would improve its infrastructure performance, reduce CPU utilization and response time, all without requiring code changes or R&D overhead.

As a real-time continuous optimization solution, providing improved performance with no code changes, Intel Tiber App-Level Optimization was the perfect fit to answer their Sync cluster needs. After testing App-Level Optimization on a few machines and seeing great value, Nylas was on board to expand optimization to the entire cluster.

Nylas initially experimented with Intel's open source Continuous Profiler in order to locate production code bottlenecks and identify optimization opportunities. The profiling tool allowed Nylas to predict the improved performance potential App-Level Optimization can achieve.

Results

By implementing App-Level Optimization on a few Sync machines, Nylas dramatically improved the service performance, cutting 35% off the cost of their compute spend and increasing the number of sync processes executed per machine by a large margin. This was a convincing demonstration of the potential cost reduction that can be achieved using App-Level Optimization's application-driven resource management optimization.

After learning the Sync service resource usage patterns and data flow, App-Level Optimization began making real-time decisions at the runtime level to prioritize resource allocation mechanisms and queues to achieve optimized performance. Immediately after activation, Nylas experienced a 58% reduction in latency, allowing the same cluster to handle 2X more syncing requests on the cluster level while reducing CPU utilization by more than 10%.

These results allowed Nylas' cluster to handle the sync queue much faster, handling more requests with fewer machines and reducing CPU utilization per machine. This led to substantial cost reduction, alongside improved performance on the cluster level.

Request a demo →