

Powered by  **GRANULATE**

Application Runtime Optimization

**Improving performance and reducing cost with
optimizing autonomous, continuous workloads**



Autonomous, seamless workload optimization

INTRODUCTION

Granulate improves application performance and cuts costs by up to 60%
 - with no code changes involved

Real-Time Continuous Optimization in a Nutshell

Granulate's real-time continuous optimization is a workload optimization solution that reduces compute spending by up to 60% and improves key performance metrics by up to 40%.

How can you reduce costs and achieve better performance?

Granulate automatically learns an application's specific resource usage patterns and data flow to identify contended resources, bottlenecks, and prioritization opportunities. Then, it tailors OS-level scheduling and prioritization decisions to lower costs without compromising on performance. This creates fast time to value regardless of application infrastructure and previous cost-saving measures.

- **60%** LOWER COSTS
- **40x** FASTER RESPONSE TIME
- **5x** INCREASED THROUGHPUT
- **0** CODE CHANGES

Performance Points

Application-Driven: Tailored resource management based on application needs

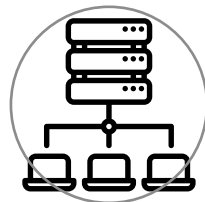
Dynamic: Real-time process level scheduling and prioritization

Automatic: Autonomous optimization no configuration or maintenance

BACKGROUND

Environment Agnostic

- All architectures
- All environments



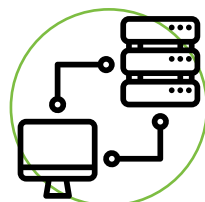
Past Initiatives

- Discount purchasing
- Cloud hygiene
- Rightsizing



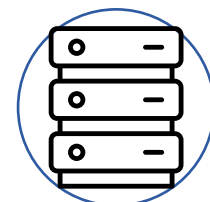
Simple Installation

- Single command line
- Start from a single VM
- Completely automated



No R&D Efforts

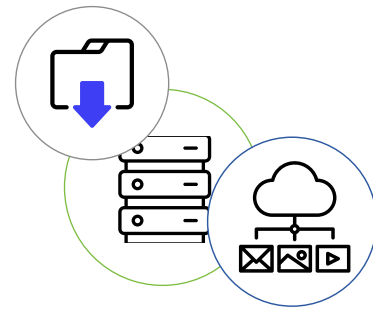
- No code changes
- No maintenance
- No configuration





How it works

- Install the Granulate agent on your workload of choice
- Agent learns application data flow and processing stages
- Agent optimizes resource management decisions



Granulate agents automatically analyze and learn the running environment processing stages and dataflow to identify application and resource bottlenecks. Once activated, Granulate agents perform AI-powered application-driven OS adaptations to improve the application performance and reduce the amount of required compute resources.

By learning the application's specific resource usage patterns and analyzing CPU scheduling order, oversubscribed locks, memory, network, and disk access patterns, the agents are able to identify contended resources, bottlenecks, and prioritization opportunities and solve them in real-time using scheduling and prioritization decisions.

Industry - Autonomous Mobility

HQ - Israel

Employees - 1300+



45%

REDUCED COSTS

45.2%

FASTER JOB COMPLETION TIME

The Challenge

Infrastructure - AWS / Workload Type - Big Data / Application Language - Python

Mobileye is running core workloads on AWS for greater speed, agility, and compute power. AWS enables Mobileye to innovate rapidly using AWS's broad and deep portfolio of services, including compute, storage, database, analytics, machine learning, and edge computing. These services help Mobileye supply automakers with the most advanced selfdriving applications. Mobileye runs big-data workloads on AWS constructing data lakes leveraging Amazon Simple Storage Service (Amazon S3) to ingest, process, and analyze hundreds of petabytes of vehicle data gathered from sensors, images, and video feeds. Insights gained from this data give Mobileye the ability to fine tune its technology in significantly shorter cycles and iterate on its autonomous vehicle capabilities.

Mobileye runs big-data workloads over AWS leveraging PySpark to support its RoadBook™ of localized drivable paths and visual landmarks using REM™; and provide mapping for autonomous driving. AWS's EMR services are enabling Mobileye to quickly innovate on top of our highly scalable, fault-tolerant infrastructure but are also becoming a significant cost driver for the organization.

Why Granulate

While searching for an optimization solution, Mobileye identified the great potential in implementing Granulate's Real-Time Continuous Optimization to achieve better performance for their PySpark workloads as well as significant cost reduction. Granulate was an attractive



solution for Mobileye as it offers a unique value proposition for big-data workloads, improving performance without migration to a 3rd party data lake/analytic platform.

Grofers Cuts Compute Costs by 57% In 10 days

Learn how Grofers leveraged Granulate's real-time continuous optimization to increase application performance and save %57 on EC2 costs without code changes or R&D effort.

Vertical: Online Retail

HQ: Gurugram, India

Employees: 3,000+



Grofers is India's largest low price online supermarket in the grocery space. The company uses its in-house technology platform to manage a network of over 5,000 partner stores that enable the company to run a fast and lean supply chain – from manufacturers straight to customers in 27+ cities across India.

Grofers utilizes its efficient supply chain to deliver over 25 million products to customers every month and its customer base is growing 4x month after month.

57%

COMPUTE COST REDUCTION

42%

RESPONSE TIME REDUCTION

29%

INCREASED THROUGHPUT

36%

REDUCED CPU UTILIZATION

The Challenge

As a growing online retail and deliveries company facing millions of visitors every month, the technical team faced many architectural challenges. Growing infrastructural systems is never easy, but when you grow as fast as Grofers has, the challenges and costs tend to pile up faster than estimated. “With our rapid growth, we quickly realized the need to address our compute costs. We were looking for a solution that would not require lots of development on our side” - Vaidik Kapoor, VP Engineering at Grofers.

For any online retail company, achieving operational efficiency, and in particular, lowering the cost of goods sold (COGS) is always essential. In Grofers' case, a company that drives its competitive advantage from being a low price retailer, this is especially important. Grofers recognized the need to optimize their AWS compute costs and to increase the application performance, but with their fast pace of growth, increasing scale, and complex architecture the need for a new optimization approach arose.

“The results were impressive and better than expected, in a few days we were able to reduce our AWS compute costs by nearly 60%”

**Vaidik Kapoor,
VP Engineering at Grofers**

Solution

With the need to meet the growing costs and performance challenges, Grofers turned to Granulate's real-time continuous optimization solution. Granulate agents were installed on an API service handling 7,200 requests per second, with an average response time of 7ms running at an average of 60% CPU utilization.



Following 3 days of autonomous learning of Grofers workload, the agents were activated to begin real-time optimization. This optimization resulted in immediate significant performance.

Mobile Attribution & Marketing Analytics

HQ: San Francisco,

USA - Employees: 1000



25%

SPARK COST REDUCTION

35%

KAFKA COST REDUCTION

30%

ANNUAL COST REDUCTION

The Challenge

As the market leader of the growing mobile attribution market, AppsFlyer heavily relies on AWS usage, and with the rapid growth also comes an increased demand for their AWS infrastructure and growing compute costs. AppsFlyer infrastructure is comprised of a wide range of services and technologies. Starting with one service out of around 40 in each cluster type, Granulate was deployed to optimize 2 services: Spark and Kafka consumers. The results of the first two services demonstrate the cost reduction potential that AppsFlyer can achieve with Granulate's application-driven resource management optimization.

Why Granulate

AppsFlyer's highly professional team was looking for an innovative complementary solution to the already implemented best-practices and cost-management solutions. The team needed

a new approach that will support the company's fast-growth and scalability requirements without requiring R&D efforts or application changes. Granulate was able to meet all of AppsFlyer's needs, offering an additional significant cost reduction without requiring any code changes and with a simple installation. Granulate was able to optimize the performance and resource utilization of both services allowing AppsFlyer to handle these services with significantly less computing resources leading to a dramatic cost reduction.

“With Granulate, we were able to save hundreds of thousands of dollars on two of our clusters without any code changes”

**Eli Zilbershtein,
Platform Group Team Leader**