Intel Workload Optimizer by Granulate improves performance by tailoring infrastructure decisions to specific application needs

The real-time continuous optimization that Intel Workload Optimizer by Granulate delivers is a new approach to optimizing production workloads by leveraging resource usage patterns and data flow to automatically adapt kernel-level and runtime-level resource management to better fit the application needs.

Real-time continuous optimization allows organizations to handle compute workloads with 60 percent fewer servers while improving performance by 40 percent with no code changes or R&D efforts required.¹

Workload optimization can be costly, disruptive, and time-consuming. Most optimization efforts require users to rewrite their code to achieve performance improvements, resulting in time away from a business’s core product development. Once installed, Intel Workload Optimizer by Granulate agents learn the resource usage patterns and the data flow of the application and operating system. The agents use US-patented algorithmic models to identify instances of data bottlenecks and resource contention within workloads and then adjust resource management decisions to accelerate data flow through an application. This process involves adapting resource allocation at the operating system and runtime level to continuously and autonomously optimize memory allocation, process swapping, thread scheduling, storage access, and network communications. Users perpetually experience performance increases as the Intel Workload Optimizer by Granulate continuously recognizes usage patterns and optimizes each workload.

Intel Workload Optimizer by Granulate optimizes Intel® processor performance, helping to ensure performance superiority over the competition

The workload optimizer is purpose-built for Intel® processors, delivering differentiated value for Intel customers.
How does it work?

1. Deploy the optimizer’s open source gProfiler for up-front assessment of the expected performance improvement. Assessment results in less than 24 hours.

2. Deploy the workload optimizer on instances of a workload of choice.

3. Intel Workload Optimizer by Granulate learns the application data flow and processing stages and then optimizes resource management decisions.

Use cases:
Intel Workload Optimizer by Granulate optimizes compute everywhere

- **CUSTOM APPLICATIONS – JAVA, SCALA, CLOJURE & KOTLIN**
  - EKS/AKS/GKE, ECS/ACS containers, monolith

- **CUSTOM APPLICATIONS – GO, PYTHON AND RUBY**
  - EKS/AKS/GKE, ECS/ACS containers, monolith

- **BIG DATA**
  - Spark, Hadoop, EMR, PySpark, Dataproc, HDInsight, Elasticsearch

- **STREAM PROCESSING**
  - Kafka consumer/producer, ActiveMQ consumer/producer, RabbitMQ consumer/producer, Kafka broker

- **CUSTOM APPLICATIONS – NODE.js**
  - EKS/AKS/GKE, ECS/ACS containers, monolith

<table>
<thead>
<tr>
<th>IMPROVED PERFORMANCE</th>
<th>1 WEEK</th>
<th>2 WEEKS</th>
<th>3 WEEKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>40%–60%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25%–40%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%–25%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TIME TO VALUE**

- **1 WEEK**
  - 40%–60%
  - 25%–40%
  - 10%–25%

- **2 WEEKS**
  - 40%–60%
  - 25%–40%

- **3 WEEKS**
  - 40%–60%
Mobileye and AWS

Mobileye runs core workloads on AWS for greater speed, agility, and compute power. AWS enables Mobileye to innovate rapidly using its compute, storage, database, analytics, machine learning, and edge computing services. These services help Mobileye supply automakers with advanced self-driving applications. Mobileye runs big-data workloads on AWS, constructing data lakes on Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud (EC2) Intel-based instances. Mobileye can ingest, process, and analyze significant amounts of vehicle data gathered from sensors, images, and video feeds. Insights gained from the data give Mobileye the ability to fine-tune its technology in significantly shorter cycles and iterate on its autonomous vehicle capabilities.

The challenge

Infrastructure: AWS
Workload type: Big data, PySpark
Application language: Python

Mobileye also runs big-data workloads on AWS, using PySpark to support its RoadBook of localized drivable paths and visual landmarks to provide centimeter-accurate mapping for autonomous driving. Amazon EC2 R4.8/16/24XL and R5.8/16/24XL on Spot instances, specifically Intel-powered instances, enable Mobileye to quickly innovate on top of a highly scalable, fault-tolerant infrastructure, but these services have also become a significant cost driver for the organization.

Solution

Granulate agents were deployed on a Mobileye EC2 workload as container sidecars and began learning the workload, data flow, and resource usage patterns within minutes. Following a few days of learning, the agents were ready to be activated to start optimizing the application’s performance.

Results

Mobileye monitored the performance and results throughout the Granulate activation. Mobileye documented a 44.5 percent reduction in average job completion time. The immediate impact of the shorter job completion times was a 45 percent cost reduction, due to fewer instance hours required to run the same workload.

Why Intel Workload Optimizer by Granulate?

While searching for an optimization solution, Mobileye identified Intel Workload Optimizer by Granulate as having the potential to achieve better performance for PySpark workloads and significantly reduce costs. Granulate was an especially appealing solution for Mobileye, given the optimizer’s unique ability to improve big-data workload performance without migrating to a third-party data lake or analytics platform.

There were other unique advantages that stood out to the Mobileye team:

- **Intel Workload Optimizer by Granulate has native integration with AWS services:** Mobileye could enjoy the cost reduction and performance improvements to its PySpark workload, without changing the way its DevOps team operated.

- **The workload optimizer requires no implementation efforts:** Granulate agent can be deployed using Docker containers within several minutes and without any code changes.

- **The workload optimizer provides better performance and lower costs:** Mobileye could benefit not only from significant cost reduction, but improved agility and SLA performance by implementing Intel Workload Optimizer by Granulate.

---

**About Mobileye**

Industry: Mobility
HQ: Israel
Employees: 1,300+
Size: Over 60K cores for REM activity

Mobileye is the global leader in the development of computer vision and machine learning, data analysis, localization, and mapping for Advanced Driver Assistance Systems and autonomous driving. Mobileye technology keeps passengers safer on the road, reduces the risk of traffic accidents, and has the potential to revolutionize the driving experience by enabling autonomous driving.

Mobileye’s proprietary software algorithms and EyeQ chips perform detailed interpretations of the visual field to anticipate possible collisions with other vehicles, pedestrians, cyclists, animals, debris, and other obstacles. Mobileye’s products can also detect roadway markings such as lanes, road boundaries, barriers, and similar items; identify and read traffic signs, directional signs, and traffic lights; create a RoadBook of localized drivable paths and visual landmarks using Road Experience Management (REM); and provide mapping for autonomous driving.

---

Mobileye reduced PySpark costs by 45 percent

<table>
<thead>
<tr>
<th>Simple and seamless integration</th>
<th>Performance benefit in less than 14 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very little time investment</td>
<td></td>
</tr>
<tr>
<td>No code changes required</td>
<td></td>
</tr>
</tbody>
</table>

REDUCED COST

FASTER JOB COMPLETION TIME

45% 44.5%

1 Based on average job completion time.

2 Based on average job execution time.

There are other unique advantages that stood out to the Mobileye team:

- **Intel Workload Optimizer by Granulate has native integration with AWS services:** Mobileye could enjoy the cost reduction and performance improvements to its PySpark workload, without changing the way its DevOps team operated.

- **The workload optimizer requires no implementation efforts:** Granulate agent can be deployed using Docker containers within several minutes and without any code changes.

- **The workload optimizer provides better performance and lower costs:** Mobileye could benefit not only from significant cost reduction, but improved agility and SLA performance by implementing Intel Workload Optimizer by Granulate.

---

3
Intel Workload Optimizer by Granulate decision tree

Choosing a candidate workload

- Is the workload Linux based?
  - Yes
  - Java, Scala, Clojure, Go, Python, Ruby, Node.js
  - Batch, transactional, analytics, web, back end, HPC
- What is the Runtime environment of the workload?
  - C, C#, C++, .NET, PHP
- What is the workload type?
  - Database
- Is the workload GPU based?
  - No
  - Not supported
  - Significant performance improvements

Granulate deployment

Complete POC lifecycle in less than two weeks

<table>
<thead>
<tr>
<th>Up-front assessment</th>
<th>Proof of value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DAY 1</strong></td>
<td></td>
</tr>
<tr>
<td>gProfiler Assessment</td>
<td></td>
</tr>
<tr>
<td>Run Intel Workload Optimizer by Granulate open source gProfiler on several VMs</td>
<td></td>
</tr>
<tr>
<td>Get up-front analysis of expected performance improvement and cost reduction</td>
<td></td>
</tr>
<tr>
<td><strong>DAY 2</strong></td>
<td></td>
</tr>
<tr>
<td>Workload optimizer deployment</td>
<td></td>
</tr>
<tr>
<td>Intel Workload Optimizer by Granulate deployment on selected workloads</td>
<td></td>
</tr>
<tr>
<td>Less than one minute</td>
<td></td>
</tr>
<tr>
<td><strong>DAY 6</strong></td>
<td></td>
</tr>
<tr>
<td>Learning and activation</td>
<td></td>
</tr>
<tr>
<td>Four days of autonomous workload learning</td>
<td></td>
</tr>
<tr>
<td>On day 6 activation, see immediate performance improvement</td>
<td></td>
</tr>
<tr>
<td><strong>DAY 12</strong></td>
<td></td>
</tr>
<tr>
<td>Performance gains and cost reduction</td>
<td></td>
</tr>
<tr>
<td>Benchmark and validate performance gains</td>
<td></td>
</tr>
<tr>
<td>Full cluster deployment</td>
<td></td>
</tr>
<tr>
<td>Automatic cost reduction</td>
<td></td>
</tr>
</tbody>
</table>
Customer offer

• Offer expires December 31, 2021
• Intel technical sales specialist to provide a one-day diagnostic to estimate the expected customer savings and performance benefits for the chosen workload
• Fully funded POC by Intel and Granulate—takes from 7 days to 21 days
• Will deploy in less than 30 days from first meeting
• Can deploy immediately after POC—existing case studies show 45 percent reduction1 in AWS fees

Contact your Intel account team and Kevin Johnson (kevin.d.johnson@intel.com), General Manager of Intel's Cloud & Enterprise Solutions & Technology Sales Group to set up a meeting.

Our security commitment

Security
We conduct ongoing rigorous security testing and third-party assessments. Deployment model is available without internet connectivity.

Privacy
We offer world-class data protection and settings that protect your infrastructure data.

Compliance
Granulate policies and products are compliant with GDPR privacy regulations.

---

1. Internal measurements provided by Mobileye and Granulate and tested on AWS EC2 instances R4.8xlarge, R4.16xlarge, R5.8xlarge, R5.16xlarge, R5.24xlarge. For more information on configurations, see https://aws.amazon.com/ec2/instance-types/.

Notices and disclaimers
Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Intel® technologies may require enabled hardware, software, or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

0821/JW/CMD/PDF