IT@Intel
Enterprise Testing

IT@Intel: Granulate Optimizes vCore Utilization & Data Throughput in Intel IT's Databricks Platform

Intel IT continues to improve our big data platform efficiency through AI and automation.

As a large IT shop, we continuously look for ways to reduce both operational expenses and infrastructure purchases, while simultaneously improving the user experience. Unfortunately, workload optimization can be disruptive and time consuming. Most optimization efforts require developers to take time away from business applications and product development to rewrite code – and sometimes, the tuning can lead to sub-optimal performance.

Learning from past experiences and challenges, our developers created a reusable framework for our applications. They configured jobs to run with the largest memory and compute configurations to handle worst-case scenarios, such as month-end processing with the biggest datasets. However, this led to an over-allocation of resources during daily operations when the datasets were smaller.

After our successful proof-of-concept (POC) testing of our on-premise platform, where we found that the Granulate software reduced 25% of memory and 44% of CPU utilization, while offering potential cost savings, we began testing in the cloud Databricks data lake. Collaborating with our technology partner Accenture, Intel IT performed a four-week POC with the active Granulate agent in our Production and Quality Assurance (QA) environment running on the Databricks platform securely in the public cloud. The results included a **23% improvement in vCore compute consumption, a 17% increase in data throughput**, and reduced CO2 emissions, aligning with the on-premises benefits we observed previously for another big data workload. This POC showed us we can run more jobs within the same compute power of the current Databricks cluster. Alternately, we can run the same number of jobs with reduced capacity to save cost and CO2 emissions.

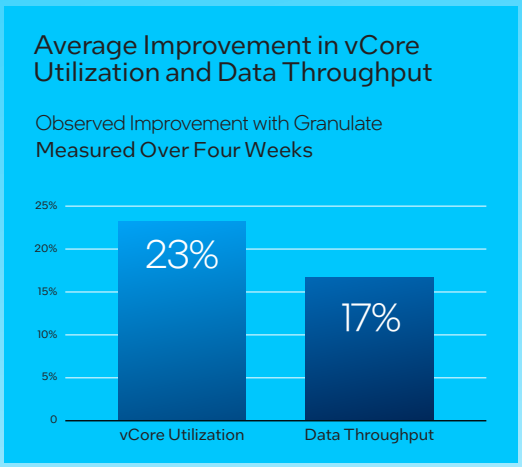


Figure 1: Granulate substantially improved vCore utilization (by 23%) and data throughput (by 17%).

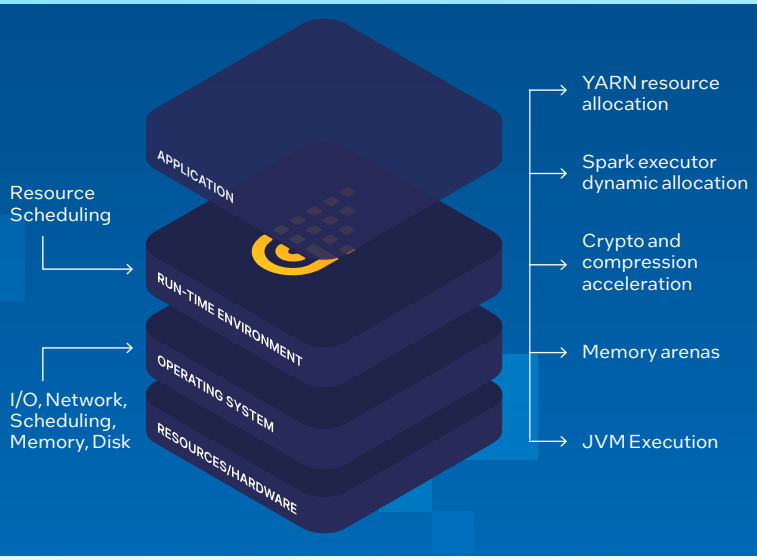


Figure 2: Five key Granulate optimizations improve platform efficiency.

How Granulate Works

Granulate is an autonomous workload optimization solution. It continuously profiles and learns the specific characteristics of each job and then optimizes resource scheduling to reduce inefficiencies and improve throughput of big data workloads – both on-premises and in the cloud. Granulate optimizes resources for many run-time environments, including Java, Python, Scala, Go, Node, Kotlin, and Clojure. The optimizations are enabled at both the runtime and operating system levels.

Assessing Granulate Capabilities in the Public Cloud

We decided to test Granulate's capability to optimize big data platforms, with no hands-on development, in our hybrid cloud environment. First, an assessor from our Information Security team performed a risk assessment of Granulate. He reviewed several assessments by independent third parties, including Granulate's SOC 2 Type II Report, the results of Granulate's penetration testing and remediation actions, and SecurityScorecard's report. In that report, Granulate earned an "A" rating of 97. The Intel Information Security team assessed that Granulate met or exceeded the pertinent standards.

Granulate was quickly and securely implemented into Intel IT's Databricks environment. Installation was followed with two main steps: first, implementation of the optimization agent via workspace initialization script for all workspaces targeted for optimization; the second step included securely sharing a Service Principal token to allow for Granulate's enhanced autoscaling capabilities to take effect. Once done, a brief learning period was conducted which resulted in the activation of the Granulate agent and meaningful cost reduction being felt while maintaining or improving SLA metrics.

We initially deployed the Granulate agent in our Quality Assurance environment to validate no disruption to business operations. The agent learned about the platform's Databricks jobs, and then we deployed it in our production environment. Again, the agent learned the environment and identified optimization opportunities at the job level behind the scenes, with no human involvement. We then activated the Granulate agent and moved to the benchmark phase, resulting in the 23% vCore utilization improvement and 17% throughput improvement mentioned in Figure 1.

Figure 4 was taken within the Granulate user interface. It shows a time shift benchmark comparison, visualizing the core count improvements. Specifically, the orange line is the most recent nine days of data where Granulate is *active* and optimizing, and the blue line is the previous nine days of data where Granulate was completely *passive*. Once Granulate is active, the core count consistently remains much lower while processing the same jobs present during the passive period. This type of benchmark analysis is delivered to all Granulate clients upon initial activation to measure total improvement.

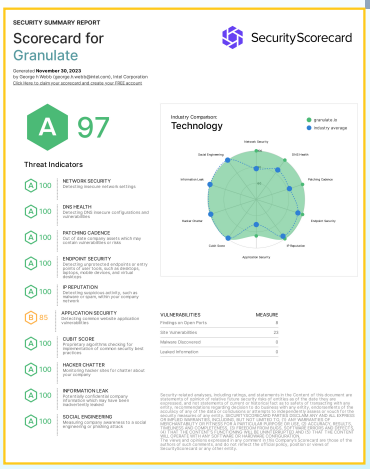
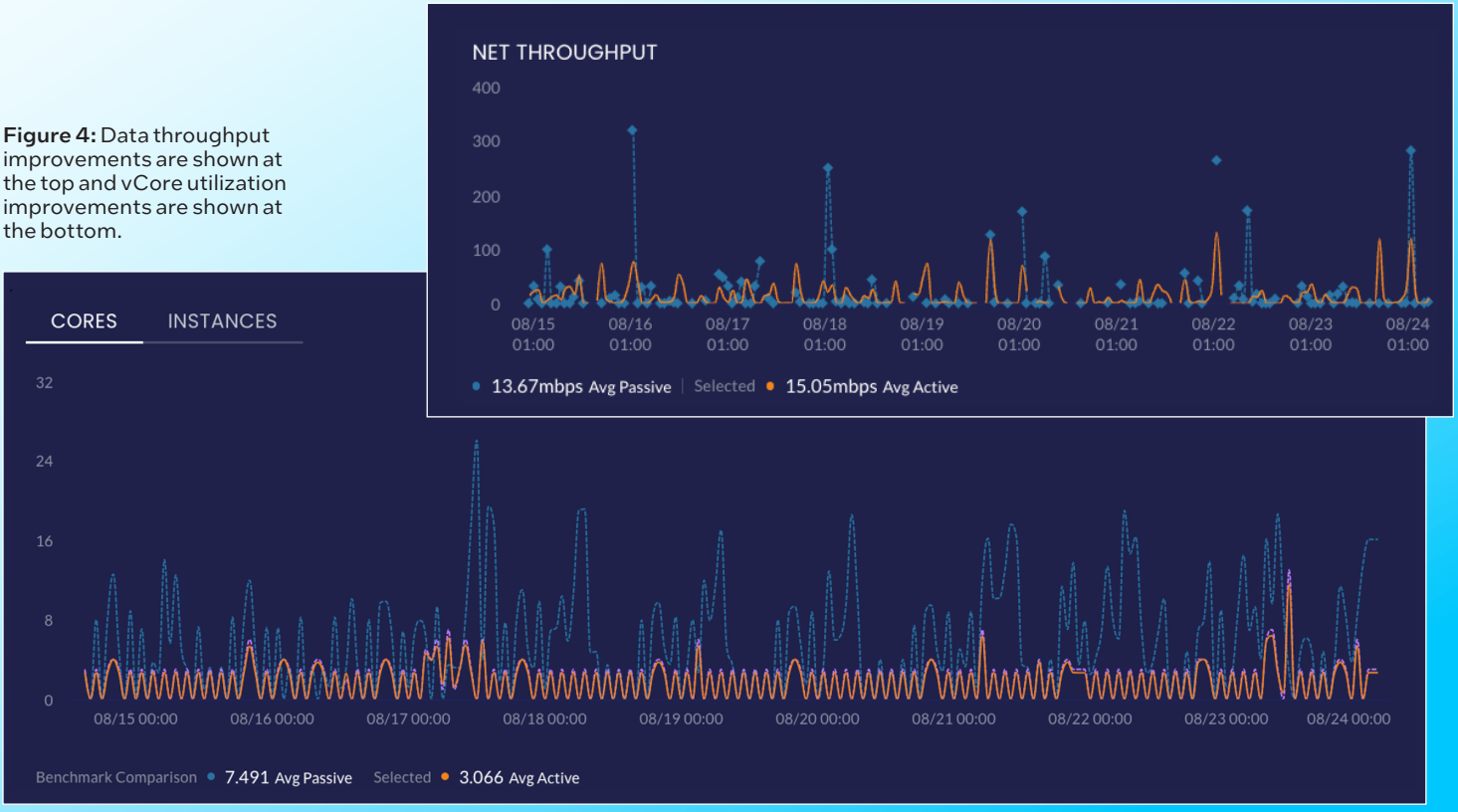


Figure 3: SecurityScorecard is a well-known third-party cybersecurity risk assessment provider.

Figure 4: Data throughput improvements are shown at the top and vCore utilization improvements are shown at the bottom.



Combining the Best of Data Lakes and Warehouses into One Platform

Databricks provides a unified analytics platform for working with massive datasets in the cloud. It is powered by Apache Spark, an open-source cluster computing framework optimized for extremely fast processing of big data workloads. Spark was developed at UC Berkeley before becoming an Apache project. The founders of Databricks were the creators of Spark while at Berkeley. Their core innovation is the "lakehouse" architecture. It combines the best of data lakes (for storing vast amounts of raw data) and data warehouses (for structured analytics). Databricks sits on cloud object stores as a unified interface for data engineering, data science, and analytics.

Enterprise Cloud Analytics Reference Architecture

Supporting business data-intensive applications used across the Intel, Intel's Enterprise Cloud Analytics Reference Architecture supports real-time and batch ingestion from various sources with different formats like structured, semi-structured and unstructured data. Data is loaded and curated in Delta Lake using Apache Spark (written in SparkSQL and PySpark), powered by Databricks. After data curation, data gets moved to Cloud Data Warehouse and ready for consumption. Business users can use BI/Client tools to query Data Warehouse for analytics and reporting/visualization. Data scientists leverage the Databricks platform to run machine learning algorithms on the same Data Lake, without copying or duplicating the data into another environment.

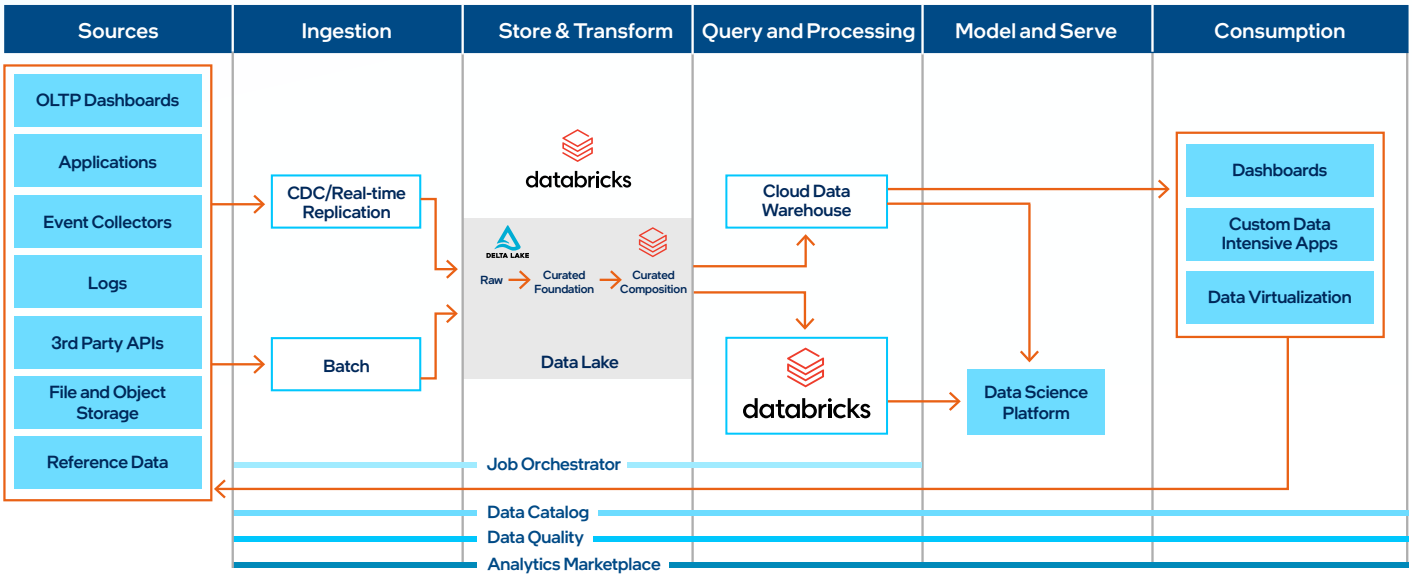


Figure 5: The Intel IT Enterprise Cloud Analytics Reference Architecture

Conclusion

Granulate autonomously learns application behavior, optimizing memory and CPU usage for each job continuously and without the need for manual intervention by application teams. This means one-time activation/change enables on-going optimization every time jobs are executing. Now, by simply activating the Granulate agent, our platform team can optimize resources and reduce excess allocation during runtime. This is achieved without any commitment of additional resources or code changes from the application teams.

When we look at these POCs together, our testing shows Granulate can drive value across hybrid cloud environments at enterprise scale. Although the team needed to reboot the cluster for the initialization script (Granulate agent) to get installed, no other downtime or code changes are required. Once the optimization agent was implemented, Granulate's enhanced autoscaling capabilities took effect. A brief learning period led to the activation of the Granulate agent.

Over the four weeks of testing, this POC clearly demonstrated Granulate's ability to enhance and optimize workload efficiency in our Databricks implementation, allowing more job execution per compute unit. We have now started a phased deployment across our remaining Databricks environment, with further rollout to follow upon successful measurement of key performance indicators.

Driving Sustainability through Technology

Intel IT's Commitment to Leading the Way Towards a Greener Future

Intel IT recognizes that sustainability must be at the core of everything we do. As the organization responsible for Intel's global technology infrastructure and services, we have an obligation to reduce our environmental impact and help Intel achieve its ambitious sustainability goals.

Intel IT is proud to be pursuing sustainability through both our operations and technology. We are investing in renewable energy, optimizing energy efficiency in our data centers, reducing waste, and utilizing sustainably sourced materials. We aim to reach net-zero greenhouse gas emissions, net-positive water use, and zero waste to landfills by Intel's stated deadlines.

But our commitment goes beyond our own footprint. We are also dedicated to helping Intel's customers and partners improve sustainability through our technology. By developing energy efficient hardware and software, we can empower others to reduce their energy use and carbon emissions. We are also collaborating across industries to discover new ways technology can address environmental challenges.

For Intel IT, sustainability is integral to our mission of building a better future through computing. We believe in the potential for technology to help create a more sustainable world. And we are determined to lead by example, driving sustainability not just through our words, but our concrete actions. We encourage others in the industry to join us in this critical effort.

Learn more

IT@Intel: Green Computing at Scale

For more information

IT@Intel

Granulate™ Optimizes Memory and CPU Utilization in Intel IT's Cloudera Platform

Intel Granulate Optimizes Databricks Data Management Operation

Accenture and Intel: Accelerating Client Transformation

Databricks Technology Partners

Granulate, an Intel Company

To learn more about Granulate, please contact your Intel account executive, and for assistance implementing Granulate solutions, Accenture is ready to partner with you.

intel. GRANULATE™ An Intel Company databricks accenture

DISCLAIMERS:

1. Granulate identified and optimized 23% reduction in vCore utilization and 17% data throughput increase when running on Intel IT's Databricks production platform in the public cloud. The POC tests ran from 14/Aug/23 to 12/Sep/23, comparing vCore efficiency as well as data throughput before and after enabling Granulate. Intel IT's Databricks platform configurations: 10 instances with 3rd Generation Intel® Xeon® Platinum 8370C (Ice Lake) processor. Total platform memory: 768 GB. Total platform storage: 152 TB. Software: Databricks version 11.3 LTS, Java 8; Ubuntu 20.04.5 LTS and Granulate version 4.5.0. The utilization results are the average of over 44000 executed Spark job/applications, optimized by Granulate. Testing performed by Intel.
2. IT@Intel Proof of Concept with Granulate on Intel IT's Cloudera Platform, May 2023

Performance varies by use, configuration, and other factors. Learn more at www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in the disclaimer above and may not reflect all publicly available updates. See disclaimer above for configuration details. No computer system can be absolutely secure.

Granulate technologies may require enabled hardware, software, or service activation.