# How to Get the **Most** Out of Your Cloud **Without** the Expense.

Moving to the cloud is an imperative to most businesses, but optimization is a fundamental requirement for the benefit to outweigh the cost. Cloud optimization can help save you money on your cloud usage. It can also give you one thing you need more than anything else: reassurance that you're making the right choices.
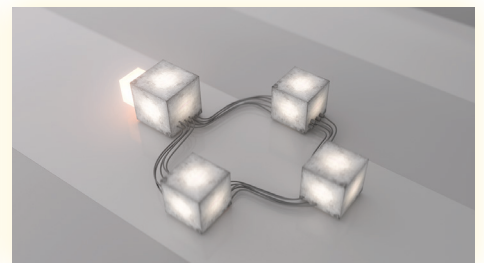
Why is the Cloud so Expensive?
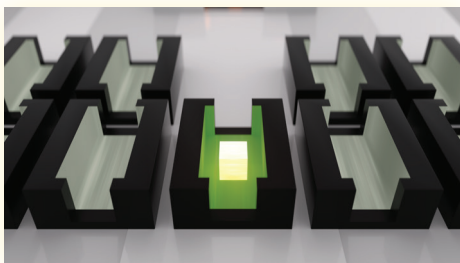
Overprovisioning
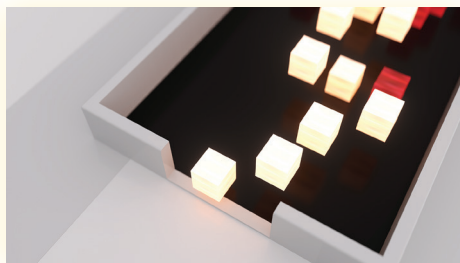
Poor Cloud Density

Workloads Running on Old Hardware

Higher than expected cloud bills can be a real problem. And that's not the only impact. Surprises in the IT bill create conflicts between groups within the business that can leave IT with few resources that it needs (this is what we call the "shock and squeeze"). It also reflects poorly on the IT pros who are on the hook to make the call about provisioning cloud resources.

- The #1 inhibitor to customer cloud programs is cost.[1]
- Focus should be on cloud optimization, not cost.[2]
- The question is not, "Is it too expensive?" It should be, "How do we run it efficiently?"[2]

MULTICLOUD STRATEGIES at 89% of organizations[3]

GLOBAL WHOLE-STRATEGIES 16.9% CAGR through 2025[3]

GLOBAL WHOLE-CLOUD SPEND $1.3T by 2025[4]

Moving to the cloud can **save** you money or **cost** you money.

Intel can help you begin your cloud engagement with the optimizations and policy-driven automation that could help you achieve the savings your business is looking for. And you don't have to disrupt your operations to do it. The flexible cloud optimization we offer can deliver results while adapting your unique business strategy and goals.

BUY SMART

- AI-guided automation learns relationships among applications to improve migration motion

- Take guesswork out of migration with a structured path that can reduce time, cost and risk

Dr Migrate

As you begin the run-up to cloud migration, we can help you find the best cloud services and instances for your workloads. Most leading cloud service providers (CSPs) offer plenty of choices and instance options — some of which can seem almost identical on the surface. CSPs also change these offerings regularly, making it hard to keep up and find an optimal path.

By working with Intel as you begin your cloud journey, you can establish transparency and governance around your workload placement. Once you do, you can employ automation and data-driven processes to manage data and cloud provisioning decisions.

In your on-premises data center, operating at **10%–40%** CPU utilization might be needed to accommodate peak traffic.

But in the cloud—where you can scale as needed—it can reduce your profits.

What you should worry more about than **costs in the cloud.**

Every business is unique, but we can help you find your path to cutting cloud costs and maximizing performance with policy-driven, automated optimization. In some cases, cloud optimization can save money with only a few tweaks.

After deploying to the cloud, many businesses like yours realize that their choices aren't delivering optimal results. When cloud expenses exceed what you expected, it can lead to organizational misalignment that drives up your costs and reduces the effectiveness of your digital transformation.

This can happen for a number of reasons, including developer overprovisioning, less-than-ideal workload performance, overlooked bottlenecks and unexpected demand.

We can help you unify your cloud strategy with metrics and reporting that clearly identify a path to optimal performance and maximized value. And we can help you map the approach that can help deliver the results you want.

MAKE IT BETTER

- Employ advanced machine learning and analytics to assess cloud, container and server resource utilization

- Get precise recommendations for cloud instance cost and performance improvement

- Optimize instance levels and address purchasing strategies simultaneously

Densify

Cuts Cloud Costs up to 40%[5]

The acceleration advantage that AWS Graviton can't match

The accelerators built into Intel® Xeon® processors are a more efficient way to drive up performance in the public cloud than increasing virtual CPU (vCPU) count, moving into a higher-priced specialized instance or re-platforming to a different architecture. Ultimately, the result is better workload performance and better price performance (the amount of work that can be accomplished per dollar at an hourly rate).

Read more

Production profiling with g:Profiler

g:Profiler gives you a view of your code-level performance to identify bottlenecks and map optimization opportunities.

- **Optimize your cloud spend with always-on analysis**. Once deployed, g:Profiler continuously analyzes code performance across your entire environment, allowing you to optimize the most resource-consuming parts of your code, improve application performance and reduce costs.

- **Deploy with ease**. Enable performance investigation with minimal overhead. It takes only minutes to deploy g:Profiler, with no code changes or deployment modifications.

- **Unify your cloud strategy**. g:Profiler combines multiple sampling profilers for a unified visualization of what your CPU is spending time on.

Optimize Automatically,
With **No Code Changes**.

We can help you implement cloud optimization and policy-driven automation that can help you find problems, identify opportunities and cut costs. And you don't have to disrupt your operations to do it.

When your business is fully deployed in the cloud, you need scalable, high-performance software to make the most of your hardware resources. While cloud computing and microservices offer you flexibility through distributed applications and deployment scalability, modern architectures have introduced more-complex performance issues that are not easily managed by traditional operating systems and runtimes.

Additionally, businesses can sometimes deploy older Linux distributions and application libraries that are not up to date with the latest advancements in today's high-performance CPUs. We can help ensure that your compute clusters, instance types and cloud deployments are rightsized for your business objectives.

MAKE IT FASTER

Granulate uses AI and machine learning to map your service's data flows and processing patterns, so it can automatically optimize runtime-level resource management.

- Easy to implement; no code changes needed
- Helps even if you're already optimizing
- Find savings automatically with automated continuous optimizations to maintain peak performance without intervention or maintenance

GRANULATE™
An Intel Company

Reduced costs by
45%[6]
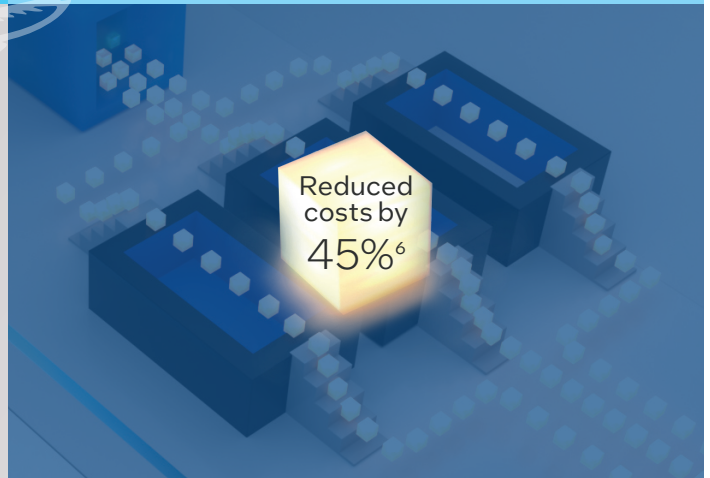
Comprehensive Reporting with Intel® System Health Inspector

Your first step to performance analysis is a baseline of your current system configuration. Intel System Health Inspector offers performance profiling and much more. With this tool, we can help you:

- Perform a large set of configuration checks, health checks and micro benchmarks on physical, virtual or public cloud instances.
- Diagnose common configuration problems and pinpoint performance bottlenecks.
- Find ways to boost application performance and cost efficiency.

| Collect, check, and baseline characters that may affect operation and performance | Suitable for single-node and multinode inspection | User chooses depth of inspection, from quick to detailed |

Performance and configuration insight with Intel® VTune™ Profiler

VTune Profiler optimizes application performance, system performance and system configuration for HPC, cloud, IoT, media, storage and more. With Intel VTune Profiler, you can locate hotspots, identify memory or I/O bottlenecks and examine how efficiently code is threaded.

Optimize your GPU offload schema and data transfers for SCYL, OpenCL code, Microsoft DirectX or OpenMP offload code. You can also characterize performance aspects of large-scale message passing interface (MPI) and OpenMP workloads.

Algorithm

Parallelism

Platform analyses

Accelerators

Microarchitecture

I/O
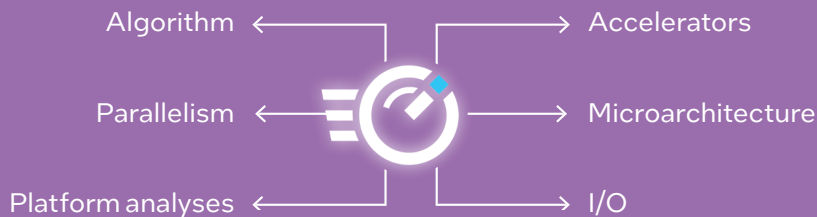
Key metrics analysis and visualization with Intel® Telemetry Collector

Which applications are using your memory? Do you have a lot of resource contention in your cache? Do you have a properly hyperthreaded application that's being well distributed, or is it only using one core? How much power are you using and where? ITC answers critical performance questions by providing comprehensive monitoring of a wide variety of performance metrics that you can ingest, analyze and visualize.

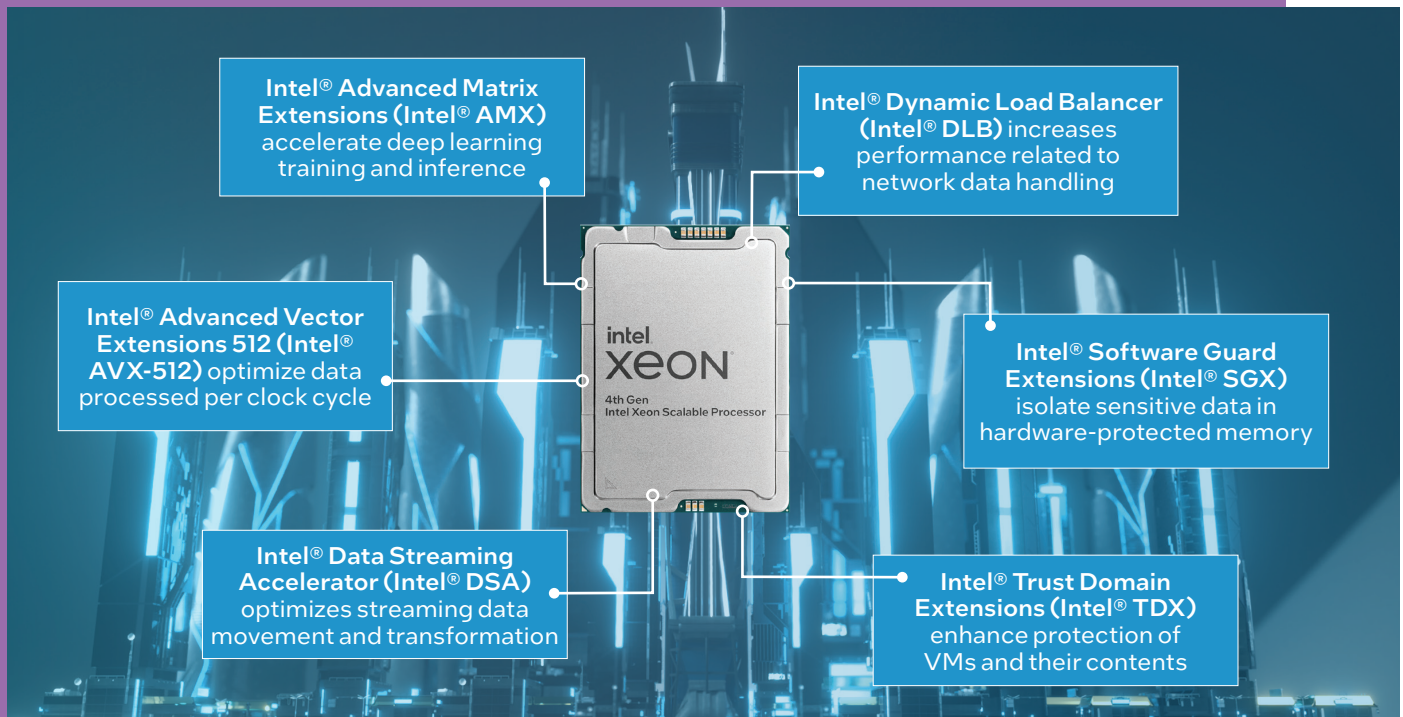ITC helps maximize cloud and data center value:

▪ Save time and money with help diagnosing and troubleshooting thorny problems.

▪ Clearly differentiate between problems such as degradations caused by a hardware failure versus workload imbalances.

▪ Detect the underlying issues of your performance problems and avoid continually putting out fires.

Accelerate your path to cloud value.

The tools described in this paper provide the means to take full advantage of the cost-efficiency opportunities that await today's businesses on their cloud journeys. They are built with open, flexible technologies to drive greater value from the infrastructure you already have. Beyond these, 5th Gen Intel® Xeon® processors feature the industry's largest set of built-in accelerators to deliver performance and power-efficiency advantages across the fastest-growing cloud workloads, including AI, analytics, networking, storage and high performance computing.

Built-In accelerators in Intel® Xeon® processors

Intel® Advanced Matrix Extensions (Intel® AMX) accelerate deep learning training and inference

Intel® Dynamic Load Balancer (Intel® DLB) increases performance related to network data handling

Intel® Advanced Vector Extensions 512 (Intel® AVX-512) optimize data processed per clock cycle

Intel® Software Guard Extensions (Intel® SGX) isolate sensitive data in hardware-protected memory

Intel® Data Streaming Accelerator (Intel® DSA) optimizes streaming data movement and transformation

Intel® Trust Domain Extensions (Intel® TDX) enhance protection of VMs and their contents

intel
XEON
4th Gen
Intel Xeon Scalable Processor

To enable new built-in accelerator features within a hyper-scaled environment, Intel supports the ecosystem with all of the most common cloud APIs, libraries and OS-level software. This results in more efficient CPU utilization, lower cloud electricity consumption and higher services ROI, while helping businesses achieve their sustainability goals.

With Intel, businesses can speed up time to deployment with the largest ecosystem of partners they know and use. Hardware and software vendors and solution integrators around the world build their products on Intel Xeon processors, offering maximum choice and interoperability with the reassurance of thousands of real-world implementations.

Contact your Intel representative to optimize your cloud — and save your wallet.

More Information

5th Gen Intel® Xeon® processors

Dr. Migrate by LAB3

Intel® Cloud Optimizer by Densify

Granulate, an Intel company

intel.

[1]HashiCorp, "HashiCorp 2022 State of Cloud Strategy Survey." https://www.hashicorp.com/state-of-the-cloud.

[2]SiliconANGLE Media, June 18 2021. "Cloud costs ignite fresh debate in tech community over where to run the enterprise." https://siliconangle.com/2021/06/18/cloud-costs-ignite-fresh-debate-tech-community-run-enterprise/.

[3]Flexera. "2022 State of the Cloud Report." https://resources.flexera.com/web/pdf/Flexera-State-of-the-Cloud-Report-2022.pdf.

[4]IDC Market Forecast, September 2021 (IDC #US47397521). "Worldwide Whole Cloud Forecast, 2021–2025: The Path Ahead for Cloud in a Digital-First World." https://www.idc.com/getdoc.jsp?containerId=US47397521.

[5]Densify. "Cloud Cost Optimization with NTT DATA and Densify." https://www.densify.com/resources/cloud-cost-optimization-nttdata/.

[6]Granulate, an Intel Company. "Mobileye Reduced 45% On Their AWS Costs Leveraging Granulate." https://granulate.io/case-studies/moblieye/.