

[Speak with an expert](#)

Run every workload on the lowest-cost compute

> Intelligent automation for containers and VMs across EKS, AKS, GKE, and ECS

Container and VM environments are consistently overprovisioned, *and it's getting worse*. **Kubernetes**, home to **95% of new AI deployments by 2028**, is the clearest case:

8% **Average CPU utilization**
This number has *declined* from 10% last year.

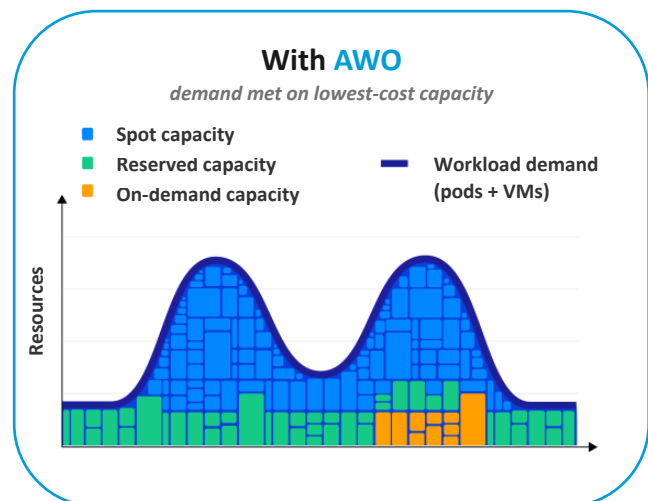
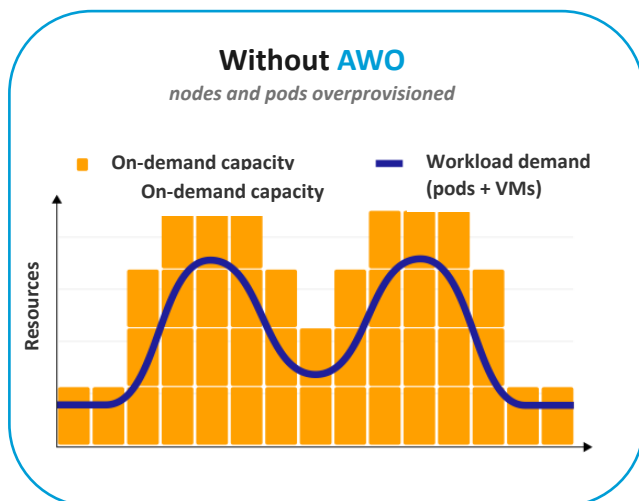
5% **Average GPU utilization**
Even worse than CPUs.

20% **Avg Memory utilization**
This number has *declined* from 23% last year.

2% **Spot usage for GPUs**
Spot Instance use for GPUs has been essentially nonexistent.

With **Autonomous Workload Optimization** you get:

- **Intelligent spot automation:** Safely maximizes spot instance adoption with predictive interruption management, automated capacity replacement, and built-in reliability guardrails.
- **Application-aware autoscaling:** Continuously analyzes workload requirements to intelligently scale nodes, clusters, and compute resources based on real-time demand.
- **Advanced bin-packing:** Identifies underutilized containers and efficiently consolidates workloads to maximize infrastructure utilization and reduce waste.
- **Autonomous pod rightsizing:** Automatically tunes resource requests and limits based on actual consumption, eliminating excess capacity without disrupting running applications.



> What teams achieve with **Autonomous Workload Optimization**



Fuel business growth while achieving the greatest cost savings and resource efficiencies.



Get accurate, actionable data to make the right decisions to power your business.



Ensure continuous optimization for performance, availability, and cost efficiency.



Free DevOps to innovate instead of spending time on operational tasks.



Power a strong FinOps culture through stakeholder alignment and collaboration.

JUMQ

“We were able to migrate most of our workloads to spot instances and save greatly on our OpEx. **Flexera brought our costs down by about 70%.**”

- Mattheiu Antoine, Sys Dev Eng

Chegg®

“Chegg’s successful adoption of microservices and containers, in large part, can be attributed to **Flexera keeping our infra cost and management to a bare minimum.**”

- Steve Evans, VP Engr

[Read the case study](#)

Customers save up to

70%

Global Ticket Vendor

“As of today, 100% of our K8s workers are running on Flexera. **We are not thinking about cost anymore**, Flexera does that for us!”

- Shane Savoie, Chief Architect

[Read the case study](#)



flaschenpost.de

“Flexera **automates workload level rightsizing – without restarting our pods** – so we can scale as we grow.”

- Denis Pesikov, SRE/DevOps Engr Mgr

> Where teams go next: **Unified Autonomous Optimization**

Rate optimization (ProsperOps) autonomously manages RIs, SPs, and CUDs, and continuously rebalancing coverage as Autonomous Workload Optimization reshapes infrastructure based on workload demand. Together they deliver precise optimization and compounded savings in one closed-loop system.

Autonomous Rate Optimization



ProsperOps
a Flexera company



Autonomous Workload Optimization

Formerly Spot Ocean  and Elastigroup 

[Learn more](#)