# Cast AI named a Technology Innovation Leader by Frost & Sullivan

## 01 What Frost & Sullivan said




Frost & Sullivan awarded Cast AI the **2026 North American Technology Innovation Leadership Recognition** in the cloud cost optimization platform industry.

> **"A truly autonomous, agent-driven platform that unifies performance, efficiency, and scalability across the full compute stack."**
>
> FROST & SULLIVAN, 2026 REPORT CONCLUSION

The analysts credit Cast AI with applying **optimization across the entire compute stack: CPU, GPU, memory, storage, and network resources**. They also highlight its reach into **GPU and AI-token optimization**, and its fit for **large, complex enterprise estates**.

Beyond the scope of optimization, analysts recognize Cast AI's approach as **built for the AI era**: instead of alerts that wait for human attention, Cast AI **autonomously fixes waste and risk** before they turn into incidents.

## 02 Why it matters

WHAT YOU GET

### Freed budget and engineering time back, without trading away performance.

Frost & Sullivan describes over-provisioning waste being removed continuously while applications stay stable, results showing within minutes of deployment, and engineers free to focus on strategic work instead of manual tuning. The payoff grows with scale: the more clusters, teams and GPUs a company runs, the less any manual process can cover.

HOW IT WORKS

### Workloads and resources in one optimization cycle.

Tuning one layer alone tends to backfire: trimming what pods request leaves nodes oversized, and shrinking nodes puts workloads at risk. Cast AI optimizes the Kubernetes workloads and the resources beneath them (CPU, GPU, memory, storage, network) in one cycle, so every change accounts for the others and performance stays intact. The result: budget that stays freed and toil that stays gone.

## 03 What makes Cast AI different: full-loop automation, end to end

Most tools optimize one resource or one layer. Frost & Sullivan describes Cast AI as automating the whole stack and coordinating the application with its resources, so performance stays intact.

| LAYER | WHAT CAST AI AUTOMATES (PER FROST & SULLIVAN) | CAST AI PRODUCT |
|---|---|---|
| CPU and memory | Real-time autoscaling, bin-packing, Spot Instance orchestration, drift remediation | Kubernetes Optimization |
| Application performance | Out-of-memory prevention, policy-based enforcement, security patching, predictive adjustments, before users feel an issue | Kubernetes Optimization |
| GPU and AI workloads | GPU allocation matched to live demand for LLM inference, training and AI apps | OMNI Compute |
| Storage and databases | The same continuous optimization applied beyond compute | Storage Autoscaler, Database Optimizer |
| Token and AI economics | Lower token cost and higher throughput through the coding agent | Coding Platform (Kimchi) |
| Every environment | AWS, Azure, GCP, hybrid and on-prem under one consistent approach | Cast AI Anywhere |

READ THE FULL REPORT

### See how Frost & Sullivan evaluated Cast AI.

cast.ai/reports/frost-sullivan-north-american-technology-innovation-leadership-recognition

Source: Frost & Sullivan, 2026 Technology Innovation Leadership, North American Cloud Cost Optimization Platform.

© Cast AI 2026 / cast.ai