IDC Link
Real-Time IDC Research® Opinion on Industry News, Trends, and Events

Flexera Announces Expansion to Its AI Cost Management Platform

October 27, 2026

By: Jevin Jensen

IDC'S QUICK TAKE

On October 20, 2026, Flexera announced an expansion of its AI Cost Management platform. The expansion is built on existing AI capabilities, enabling a more automated and proactive approach to managing an enterprise's AI costs. Flexera's push into the AI data cloud and infrastructure cost control confirms that FinOps for AI has outgrown simplistic token tracking, as the depth of complexity has also increased. These changes mean that tracking and optimization must now span every layer where AI spend occurs. IT leaders who continue to manage SaaS, cloud, and AI costs in separate silos risk repeating the asset sprawl and spiraling costs of previous eras.

PRODUCT ANNOUNCEMENT HIGHLIGHTS

Flexera's recent expansion announcement can be found here: AI Cost Management platform. The Flexera One platform now includes cost controls for AI cloud infrastructure and AI data cloud platforms. The AI ecosystem that Flexera now manages and optimizes includes AI gateways, observability, model routing, and economic control systems. Finally, an AI maturity model assessment tool was introduced to help companies improve their internal AI practices. The AI Cost Management maturity assessment produces a scorecard and a peer comparison with near-instant results for enterprises using the tool.

New autonomous capabilities in Flexera's Data Cloud Optimization and Workload Optimization offerings use agents to balance cost and performance across clouds, containers, Databricks, and Snowflake. Snowflake automation is expected by the end of 2026, with Databricks automation following in early 2027.

These new capabilities are part of the Flexera One platform, providing full-stack AI Cost Management (AICM). Existing AICM capabilities include Token Cost Control and Cloud Cost Optimization. Flexera's Token Cost Control pulls cost and usage data from AI vendors' APIs, normalizes tokens, and provides a single view of SaaS-based AI. Cloud Cost Optimization provides cost visibility, allocation, governance, and optimization for AI services and supporting cloud infrastructure purchased through hyperscalers.

Flexera also announced collaborations with Revenium and Ascerta (formerly Pay-i) to extend integrations and support for open source community and commercial AI infrastructure. These efforts will help organizations gain greater visibility, governance, and financial accountability across their AI investments.

IDC'S POINT OF VIEW

Flexera's AICM capabilities help enterprises address their biggest current struggle: understanding what AI really costs. IDC finds that by August, 90.5% of enterprises are over their 2026 token budget. Organizations lack real-time visibility and actionable

recommendations. Only 9% of companies can proactively review their AI costs weekly or more often (source: IDC's *Intelligent Cloud and Edge Operations with FinOps Survey,* September 2026; n = 254.)

Tracking AI cloud infrastructure matters. GPUs, data platforms, APIs, networks, token-based charges, and other elements are new and expensive costs that FinOps teams must optimize. This occurs amid increasing complexity that IT and FinOps teams must address. IDC recommends a platform approach that extends FinOps across the full stack and consolidates multiple tools, avoiding the training and overhead of multiple FinOps solutions. Flexera's reach across SaaS, cloud, data clouds, and compute aligns with this model.

Buyers should still look past the bold headlines. Token costs will continue to fall as companies' usage grows exponentially. Flexera offers desperately needed visibility with its expanded AICM solution, and its recommendations and optimization will continue to improve through 2027. Autonomous optimization, with guardrails, is the right direction. The agents that can apply recommendations quickly will beat the monthly review of static reports every time.

Another risk IDC has warned about is agent sprawl. IDC has seen this movie before, with virtual machines and zombie cloud resources driving up costs and attracting unwanted attention from the CFO. Enterprises deployed 28.8 million agents in 2025, and IDC expects them to manage roughly 80 times that number by the end of 2026, with daily agent actions nearing 217 billion by 2029 (see *Agent Economics: Adoption and Delivery Model for Agents, Actions, and Tokens,* IDC #US54034725, December 2025).

The winners will understand AI economics, avoid shadow IT sprawl, and proactively manage total cost of ownership. A winner will maximize business outcomes and ROI, not simply spend the most on AI.

Advice for IT leaders

- **Put FinOps inside every AI project.** This needs to occur before final project approval to ensure proper ROI and accountability. Without it, overruns surface only after the invoice arrives.
  - **Next step:** Add a FinOps review to the AI project intake and forecast costs before funding.
- **Rightsize GPUs, storage, network, AI APIs, and data platforms.** Leverage a cloud cost solution to find and optimize these critical cloud infrastructure costs.
  - **Next step:** Review GPU and data platform utilization monthly. Use the FinOps platform solution to find and optimize these costs. Put Flexera on the evaluation list if tool consolidation is an appropriate strategy for the enterprise.
- **Budget for AI's new cost elements.** AI-ready networks, workload placement, governance reviews, and ongoing model costs rarely appear in traditional business cases.
  - **Next step:** Build an AI total cost of ownership template and require it for every pilot. IDC tracks more than 20 infrastructure cost elements unique to AI.

- **Tag AI resources and measure allocation accuracy.** Cost accountability starts with proper and accurate tags.
  - **Next step:** Automate tagging for AI services and target 90% or better cost allocation accuracy, as per IDC's FinOps maturity model. Run an independent analysis to determine the organization's maturity level.
- **Treat tokens differently from subscriptions.** A usage-based token charge can dwarf a time-based compute cost. Review token costs more than once a week.
  - **Next step:** Put Flexera AICM on the short list for review of capabilities such as alerting on high token usage, coaching teams on input versus generated tokens, and normalizing vendor billing units**.**
- **Use AI for FinOps.** Forecasting and anomaly alerts speed up decision-making, and 44% of enterprises name automated AI portfolio monitoring as a top cost control strategy.
  - **Next step:** Pilot agent-assisted remediation with approval guardrails.
- **Inventory every agent now.** Unmonitored agents create outage, security, and cost risk.
  - **Next step:** Define agent standards, manage A2A and MCP protocols on a platform, and plan disaster recovery.
- **Benchmark, train, and report.** Peer comparison exposes stretch goals that internal history hides.
  - **Next step:** Complete the maturity assessment, certify FinOps staff, and show the CFO unit economics such as cost per transaction.

Subscriptions Covered:
Intelligent Cloud and Edge Operations with FinOps

Please contact the IDC Hotline at 800.343.4952, ext.7988 (or +1.508.988.7988) or sales@idc.com for information on applying the price of this document toward the purchase of an IDC or Industry Insights service or for information on additional copies or Web rights. Visit us on the Web at www.idc.com. To view a list of IDC offices worldwide, visit www.idc.com/offices. Copyright 2026 IDC. Reproduction is forbidden unless authorized. All rights reserved.