

One Platform, Many Models: Scaling Sanctioned AI Across the Global Enterprise



Andrew Buss

Senior Research Director,
SaaS, Platform Strategies,
and Ecosystems, IDC



Luis Fernandes

Senior Research Manager
SaaS, Platform Strategies,
and Ecosystems, IDC



Ewa Zborowska

Research Manager,
AI, IDC

To address the dual challenge of rapid employee AI adoption and enterprise data-governance risk, Cisco built Circuit – an internal, model-agnostic AI OS now used by more than 100,000 annual active users worldwide.

Solution snapshot

Organization:

Cisco Systems, Inc.

Organizational challenge:

Employees were increasingly turning to external AI tools, creating data governance, compliance, and security risks. Cisco needed a single internal platform capable of scaling the latest AI capabilities safely across a global, diverse workforce and without locking into any single vendor's model or technology stack.

Solution:

Circuit is Cisco's internal AI OS. It is a unified, secure, model-agnostic platform that delivers assistants, agents, retrieval-augmented knowledge, and enterprise workflow automation across all functions and geographies.

Project duration:

The project involves multiyear, ongoing platform development with continuous capability releases. Broad enterprise availability has been achieved and scaled over the past 18–24 months.

Benefits:

- Over 100,000 annual active users, 59,000 weekly active users, and 125,000 prompts submitted per workday
- About 15% broad-based productivity improvement
- Approximately 3,600 agents created in the first week of Agent Studio, with about 1,250 deployed, each generating five to six user interactions per day
- Named the number 1 internal AI application at Cisco

Introduction

The pace of enterprise AI adoption has shifted faster than most organizations anticipated. Employees are discovering, trialing, and embedding AI tools into their daily work at a rate that is outstripping the ability of IT, security, and legal functions to govern them. The productivity upside is substantial, but the unsanctioned use of public AI services carries real data governance, compliance, and security implications.

Cisco, a global technology company with a workforce of approximately 130,000 spread across every major geography and most enterprise functions, engineering, sales, HR, finance, legal, and operations, encountered this challenge at scale. Employees were already using external AI services, creating a growing risk surface as sensitive internal information flowed into tools Cisco did not control. Internal legal, security, and data classification teams raised the implications, and the company began seeking a sanctioned alternative.

Rather than adopting a single off-the-shelf AI product, Cisco chose to build an internal platform.

- Data and security controls needed to align with Cisco's own multitier classification standards,

including the ability to handle confidential, highly confidential, and restricted data, a level of control the company concluded it could not fully obtain from generic third-party AI products.

- No single vendor offered the combination of frontier model access, deep workflow integration, enterprise-grade observability, and governance Cisco required for a workforce as large and functionally diverse as its own.
- Committing to a single large language model (LLM) provider was viewed as a long-term risk, given how quickly the model landscape has been evolving.

The result is Circuit, Cisco's internal AI OS. Rather than a single chatbot or point assistant, Circuit is a layered platform spanning interface, orchestration, builder tools, intelligence, and infrastructure, through which planned AI experiences inside Cisco are delivered. Over the past 18 to 24 months, it has scaled from an internal initiative to what Cisco reported as its most widely used internal AI application, with more than 100,000 annual active users, 59,000 weekly active users, and approximately 125,000 daily prompts from engineers, sellers, HR specialists, operations staff, and employees across most functions.



Solution overview: What is Circuit?

Cisco describes Circuit internally as its AI OS. The framing is deliberate. An OS is not an application but the layer through which applications are built, deployed, secured, observed, and governed. Cisco's stated intent with Circuit was to create a single foundation spanning the AI stack so that AI-powered assistants, agents, and workflows could be built once and reused across the organization rather than reimplemented by individual teams.

Circuit is organized in five integrated layers

UI layer

The top layer is the user-facing interface. It provides a natural language experience across both text and recently added speech-to-speech interaction. Users can conversationally invoke agents, retrieve information from internal knowledge bases, generate content, and trigger multi-step enterprise workflows. An "intent engine" interprets each request and dynamically loads the most relevant information so employees do not need to know which tool to open for a given task. Circuit offers access through desktop agents for engineers, web applications for nontechnical users, and productivity plug-ins integrated into platforms, such as Microsoft 365 and Webex. This multi-channel approach delivers a consistent, AI-powered experience that supports complex technical workflows while remaining accessible and user-friendly for business users.

Orchestration layer

Behind the UI sits a custom orchestration layer that routes each user prompt to an LLM selected at runtime, balancing accuracy, latency, and inference cost. This orchestration supports multiagent systems through agent-to-agent interactions and manages control-to-data plane operations. This is the mechanism by which Circuit avoids committing to a single fixed model provider. It also incorporates efficiency features, including response caching and request batching, intended to contain inference spend as usage scales.

Builder hub

The third layer is where employees and engineering teams extend the platform through a suite of no-code Studios, such as Connector Studio and Agent Studio. These all connect to build controlled, end-to-end AI-powered workflows at scale. The platform also provides developer APIs and Command-Line Interfaces (CLIs) for teams building more complex AI-powered products, alongside an agent registry with governance controls intended to prevent duplication and manage the agent life cycle. The registry plays a central role in sustaining quality as the number of agents grows.

Intelligence layer

Cisco characterizes the intelligence layer as Circuit's "secret sauce." It combines an internal knowledge base, retrieval-augmented generation as a service for both the Circuit UI and third-party applications, a prompt hub of shared and community-created prompts, and long-term memory capabilities. A cost efficiency and intelligence engine serves as the backbone to optimize across data onboarding, infrastructure, and builder tools through model routing, total cost of ownership analysis, context optimization, caching, and compression.

This engine determines which model (proprietary or open source, cloud or on-premises) handles each request, optimizing cost, speed, and performance.

Infrastructure and security layer

At the foundation sits a cloud-agnostic infrastructure spanning public cloud environments and on-premises resources. On-premises deployments leverage Cisco AI PODs, purpose-built hardware optimized for model inference and workload execution. The platform employs a multicloud infrastructure strategy for containers, compute, model inference, and data loss prevention capabilities, ensuring flexibility in deployment while maintaining security and performance standards.

Multi-tier data classification is enforced across every layer, alongside prompt injection scanning, model governance policies, approval workflows, and detailed audit logging. Cisco's own AI Defense product is integrated to provide continuous AI-specific security controls. This unified approach enables sensitive data to remain in private on-premises environments while expanding workloads across hybrid deployments as needed.

Applying an AI Factory approach, in which a combination of hardware, software, and services is used to deploy workloads in a hybrid environment, allows projects to scale coherently from pilots and limited deployment to generalized projects across the production ecosystem.

This standardized, AI-ready infrastructure deploys seamlessly on-premises or in the cloud, which allows for greater reliability and flexibility as it scales.

Three design principles cut across every layer

1. Model abstraction as a first-class principle

Circuit is designed to prevent vendor lock-in at the model level. It can route to open source or proprietary models, on-premises or in the cloud, for any given request, abstracting across major frontier models and Cisco's own internally developed LLMs. Cisco reported that new frontier models are integrated within one to two weeks of release, supported by an internal service-level agreement. Users and developers can also choose or switch models manually. In a market where relative model performance shifts rapidly, this architecture preserves optionality that a single-vendor approach would limit. However, it also places the burden of evaluation, routing logic, and ongoing model governance on Cisco's own team.

2. "Cisco on Cisco" philosophy

Circuit also functions as a customer-zero deployment for Cisco's own products. Cisco AI Defense handles prompt injection and skill-level security scanning, Splunk provides observability, and Cisco's identity and access management tools handle authentication and authorization.

3. Innovation incubator for new projects

Circuit allows for a cutting-edge environment for prototyping new ideas or innovative projects. A sandbox within the secure ecosystem allows for productivity gains and project expansion without outside resources.

The platform therefore plays three roles: an internal productivity environment for Cisco employees, an innovation incubator, and a setting in which Cisco's own AI, security, and observability portfolio is exercised at enterprise scale. According to Cisco, several customers who have seen Circuit demonstrated have asked

whether they could deploy something similar. The platform is not currently productized for external use, but having been shown to a small number of customers, it sparked interest among those briefed on the architecture and Cisco's approach.

Challenges

There are three distinct categories of challenge, together offering a realistic view of what enterprise-scale AI platform development entails.

Technology selection in a fast-moving landscape

The biggest technical challenge is not building any individual component but deciding what to build, buy, and integrate. The AI ecosystem is evolving rapidly on multiple dimensions simultaneously. Model capabilities, agent frameworks, memory architectures, orchestration patterns, and interface standards are all changing in short cycles. As the market moves to autonomous AI, new security challenges emerge. Agent identifiers, digital worker authentication, and autonomous behavior tracking require solutions to maintain control and accountability at scale. Cisco's explicit response is a "fail fast, succeed sooner" mindset, accepting that some decisions will need to be revisited within months rather than years. The rollout of long-term memory is a concrete example of this iterative approach. Rather than waiting to build a perfect, all-encompassing memory system, the Circuit team prioritized getting a working version into users' hands. For the initial release, the team scoped the feature to support only LLM models and Live Web Data, deliberately leaving support for complex Cisco Data and granular memory editing for future updates. With this approach, the team can learn through real-world implementation and iterate based on actual user behavior.

People and change management at scale

The largest overall challenge is organizational rather than technical. With a workforce of roughly 130,000 and a market full of generic AI tools, keeping people aligned with Circuit as the platform for AI requires continuous internal communication, cross-functional working groups, and a clear narrative about what Circuit delivers. The UX team's in-depth user research and usability testing for major features ensure designs are intuitive and broadly adoptable across diverse user personas and skill levels. This mirrors the broader industry experience. AI adoption challenges are predominantly people challenges, not technology issues. IDC research consistently finds that organizational readiness, targeted enablement, and governance clarity are the factors that most often determine whether enterprise AI programs succeed or stall.

Agent proliferation and quality control

Democratizing agent creation through Agent Studio may be a clear strength of the platform, but it introduces a governance challenge. When any employee can build an agent in minutes, the result can quickly be hundreds of agents, many overlapping, some partially complete, some abandoned. Cisco has addressed this with a structured registration process and a dedicated agent management capability (referred to internally as an "agent registry") that manages the agent life cycle, identifies duplication, and removes incomplete or abandoned agents. Cisco is candid that this is an area still being developed and the balance between empowerment and oversight is one of the harder problems in large-scale AI platform operation.

Rollout and adoption

One notable choice in Circuit's rollout was to make the platform broadly available rather than restricting early access to a single pilot team. Once core capabilities were in place, Cisco made the platform available to employees across all functions and geographies. This approach produced two key effects: a faster feedback loop for the engineering team and the opportunity to surface use cases organically across the organization rather than prescribing them centrally. However, a broad rollout also carries risks, including reduced control over early use case selection and increased pressure on governance. Therefore, Cisco invested in governance tooling in parallel to support scale.

Deep workflow integration with enterprise systems

A meaningful share of Circuit's reported usage comes from integration with the enterprise systems on which employees already rely. IT teams across Cisco have built agent integrations that connect Circuit to key platforms, including:

- **Workday**, for HR self-service tasks such as logging time off, reviewing compensation, and resolving HR cases through an internal assistant
- **Salesforce**, for sales intelligence, account lookups, opportunity analysis, and lead scoring
- **Box and SharePoint**, for document access, retrieval, and summarization
- **Jira**, for developer workflows including ticket creation, status lookup, and closure
- **Outlook**, for accessing calendar and scheduling functionalities
- **Webex**, for communicating, summarizing conversations, and accessing meeting transcripts

Tasks that previously required navigating multiple enterprise systems can now be completed through a single natural language prompt. Cases that previously required opening a ticket and waiting 24 hours (and, in some instances, up to a week) for human triage can now be resolved directly in conversation with Circuit or an integrated agent, although coverage varies by system and workflow.

Community-driven adoption

The prompt hub plays a central role in Circuit's adoption curve. It is a shared library of community-created prompts, with hundreds of prompts stored and reused across the company. Employees build on each other's work rather than starting from scratch, which accelerates learning and spreads best practices across functions. The agent registry also performs a similar role for AI agents. Once an Agent Studio agent is approved, it becomes discoverable and reusable across Cisco, turning individual employee innovation into shared enterprise capability.

No-code builder tools

Low- and no-code capabilities are a recurring theme in Circuit's design. Agent Studio allows any employee to build an agent in roughly 30 minutes without writing code. Connector Studio, a Model Context Protocol builder, extends the same principle to connectors, allowing nontechnical users to wire Circuit into additional enterprise systems and data sources. This is a deliberate choice to broaden the builder community beyond engineering and place agent and workflow creation in the hands of the people who best understand each domain.

Benefits

Circuit's impact is visible in both behavioral metrics (how widely and frequently the platform is used) and productivity outcomes reported by the workforce.

Enterprisewide adoption

Cisco reported more than 100,000 annual active users on Circuit, covering the majority of its approximately 130,000 workforce, with 59,000 weekly active users and approximately 125,000 prompts submitted per workday through the UI (with additional traffic through APIs). Circuit was named Cisco's number 1 internal AI application in FY25.

Usage is described by Cisco as voluntary rather than mandated, which, if sustained, suggests that the platform is genuinely useful to employees across the workforce rather than driven by policy. Still, there exists a strong partnership with learning and development teams in HR to integrate Circuit as part of quarterly learning labs so that the training necessary for adoption becomes interactive and customized.

Measurable productivity gains

At Cisco's scale, any productivity gain reported represents a meaningful share of aggregate time. A broader Q1FY26 internal survey reported a 15% productivity improvement across all users, with gains distributed across functions rather than concentrated in one area. Additionally, 77% of users agreed that Circuit saves time, with 74% declaring it increases productivity and 75% confirming it helps better serve customers. Cisco noted that it is continuing to refine its measurement framework with additional behavioral and outcome-based metrics.

Strong builder community uptake

Within the first week of Agent Studio's launch, Cisco employees created approximately 3,600 agents, roughly 1,250 of which were fully deployed within that same week. Each deployed agent generated an average of five to six user interactions per day. The volume of builder activity in the first week points both to a bottom-up appetite for the no-code approach and quality control challenge covered earlier.

Workflow transformation

Tasks that previously required navigating multiple enterprise systems, such as logging time off in Workday, resolving HR cases, looking up Salesforce accounts, managing Jira tickets, and retrieving documents from Box or SharePoint, can now be completed through a single natural language prompt. Beyond direct time savings, this reduces the cognitive load on employees and lowers the barrier to using enterprise systems that were previously seen as friction-heavy.

Cost optimization at scale

Model routing, response caching, and request batching allow Cisco to manage inference costs as usage grows. Cisco's stated goal is for cost per inference to scale sublinearly with usage, which can be harder to achieve when relying on a single external vendor in which the cost per token is effectively fixed. As inference becomes a more significant line item for enterprises, cost-aware orchestration is likely to become a recurring requirement in enterprise AI architecture decisions.

Future development

Cisco's road map for Circuit reflects a consistent trajectory: from productivity tool to platform to broader enterprise substrate. A selection of planned and in-flight capabilities includes:

- Fully autonomous agents, for specific, bounded domains in which supervised automation is appropriate
- Agent orchestration a virtual team-style coordination layer that chains specialized agents across the full development life cycle, from idea to launch
- Storylines, an AI-powered presentation generation capability
- AI workflows, orchestrated, multi-step processes that connect multiple AI agents, connectors, schedules, and channels via natural language or through a drag-and-drop interface
- Vertical personalization, a custom AI workspace tailored to an employees' roles so the tools and AI capabilities presented are highly relevant from first login
- Expanded channel reach, including mobile and plug-in integrations to bring Circuit into the tools employees already use every day

Several of these capabilities are currently in pilot or limited release. Collectively, they suggest a platform evolving from an AI assistant for employees to an orchestration layer for enterprise work. How much of this materializes as planned will depend on factors common to any in-progress road map, including shifting market conditions and internal priorities.

IDC's opinion

Circuit is a compelling example of an enterprise responding pragmatically to the realities of AI adoption.

The commitment to independence from any single model provider is well-judged. As the model landscape continues to shift rapidly, routing across open source and proprietary models or cloud and on-premises preserves flexibility and helps to avoid potential commercial or technical lock-in. The discipline of reusability across Circuit is similarly worth noting: The prompt hub and agent registry turn individual prompts and Agent Studio creations into shared assets rather than one-off artifacts, reducing duplication and surfacing working practices across the workforce.

Circuit also represents a measured response to shadow AI. Rather than attempting to block external tool usage, which is common but often ineffective at scale, Cisco offered a sanctioned internal alternative with the data controls, governance, and integrations that public services cannot provide. For many enterprises wrestling with the same exposure, this pattern is more likely to succeed than restriction alone.

Two cautions are worth highlighting. First is the learning curve. A platform of this scope is genuinely powerful for technically confident users but can be overwhelming for employees who simply want to complete a task. Investment in onboarding, role-specific entry points, and the planned personalized spaces will be important to prevent a gap opening between power users and the rest of the workforce. Second is agent proliferation. Democratized creation is a strength, but 3,600 agents in one week points to future curation work.

A sustained discipline in life-cycle management will be essential to ensure the registry remains a source of signal rather than noise.

The customer-zero approach is a clear strength and points to a natural next step. Running Cisco's own AI, security, and observability products inside Circuit at enterprise scale subjects them to sustained real-world conditions and builds a credible foundation for future commercialization.

For Cisco and for customers who have already expressed interest, the logical next step is to consider how elements of Circuit could be packaged as a replicable offering. Not every component will translate directly, but the architecture, governance patterns, and builder tooling represent a commercial opportunity worth exploring.

Methodology

The project and company information contained in this document was obtained from multiple sources, including information supplied by Cisco and a detailed briefing and platform demonstration conducted between IDC analysts and Cisco's Circuit leadership team in March 2026.



About the IDC analysts



Andrew Buss

Senior Research Director,
Digital and Datacenter Infrastructure and Services, IDC

Andrew covers datacenter infrastructure services as part of IDC's Digital and Datacenter Infrastructure Strategies. His research focuses on datacenter facilities owned and operated by SPs, hyperscalers, and enterprises. He also tracks enterprise tech buyer sentiment and strategies to drive business outcomes. Andrew collaborates closely with other infrastructure teams to produce a comprehensive view of the dynamics, competitive positions, and market opportunities created by various IT and non-IT datacenter technologies.

[More about Andrew Buss →](#)



Ewa Zborowska

Research Director, AI, IDC

Ewa is an experienced technology professional with 25 years of expertise in the European IT industry. Since 2003, she has been a member of the IDC team, based in Warsaw, researching IT services markets. In 2018, she joined the European team with a specific emphasis on cloud and AI. Ewa is currently the lead analyst for IDC's European AI Innovations and Strategies CIS.

[More about Ewa Zborowska →](#)



Luis Fernandes

Senior Research Manager,
Digital and Datacenter Infrastructure and Services, IDC

Luis covers datacenter infrastructure services as part of IDC's Digital and Datacenter Infrastructure Strategies. He develops research that spans from edge to core to cloud, looking at trends and strategy shifts on how companies buy, build, and run their digital infrastructure. Additionally, Luis serves as a go-to resource in Europe, the Middle East, and Africa (EMEA) for IDC's enterprise infrastructure. In this role, he produces EMEA-centric insights on topics, such as hybrid cloud adoption, flexible consumption models, and ESG strategies. Luis has worked with European Union governing bodies, such as the European Commission, on projects surrounding the development of the region's cloud-to-edge strategy, edge deployment and proliferation, and digital sovereignty.

[More about Luis Fernandes →](#)

Message from the sponsor



Enterprise AI is moving from pilots to production, and organizations are working through how to scale it in ways that balance employee enthusiasm, business ambition, AI economics, and the need to protect sensitive data.

Cisco's approach is to treat AI as foundational rather than additive - reinventing how teams build, run, and measure their work. Circuit, Cisco's internal enterprise AI platform, sits at the heart of that approach, giving every employee an AI partner to augment their work and unlock measurable value. Through a single, sanctioned environment, it brings together curated models, enterprise data, and purpose-built agents - anchored by responsible AI principles and human oversight in the loop.

As customer zero, we run Cisco's own AI, security, and observability portfolio inside Circuit - turning enterprise-scale learning into better products for our customers.

The perspective in this paper will be useful for any organization looking to adopt a similar approach on their AI journey.

[Learn more](#)

IDC Custom Solutions

IDC Custom Solutions produced this publication. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis that IDC independently conducted and published, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies.

This IDC material is licensed for external use and in no way does the use or publication of IDC research indicate IDC's endorsement of the sponsor's or licensee's products or strategies.

[IDC.com](#)[IDC on LinkedIn](#)[IDC Blog](#)

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight help IT professionals, business executives, and the investment community make fact-based technology decisions and achieve their key business objectives.

©2026 IDC. Reproduction is forbidden unless authorized. All rights reserved. [CCPA](#)