

The Impact of AI in the WAN network.

A data-driven analysis

-

Javier Antich | Principal Product Management Engineer
CTO Office, Provider Connectivity
Cisco Data Center & Internet Infrastructure

March 9th – NetNod 2026



AI Traffic Is Rising — The Real Wave Is Still Ahead

Agentic workflows will power the next wave of demand

LAYER	DRIVER	IMPACT
Adoption X	Commoditization 	↑ Lower cost & demand growth
	Better LLMs 	↑ Addressable tasks
Capability X	Multi-modal 	↑ Payload size
	Reasoning 	↑ Tokens & Compute
Interaction	AI Agents 	↑ Transactions per request
	Multi-agent 	↑ Requests per session

1448%

Increase in tokens processed by AI models in the last 12 months¹

75%+

Of inference-driven data creation and processing at the edge by 2030²

10x

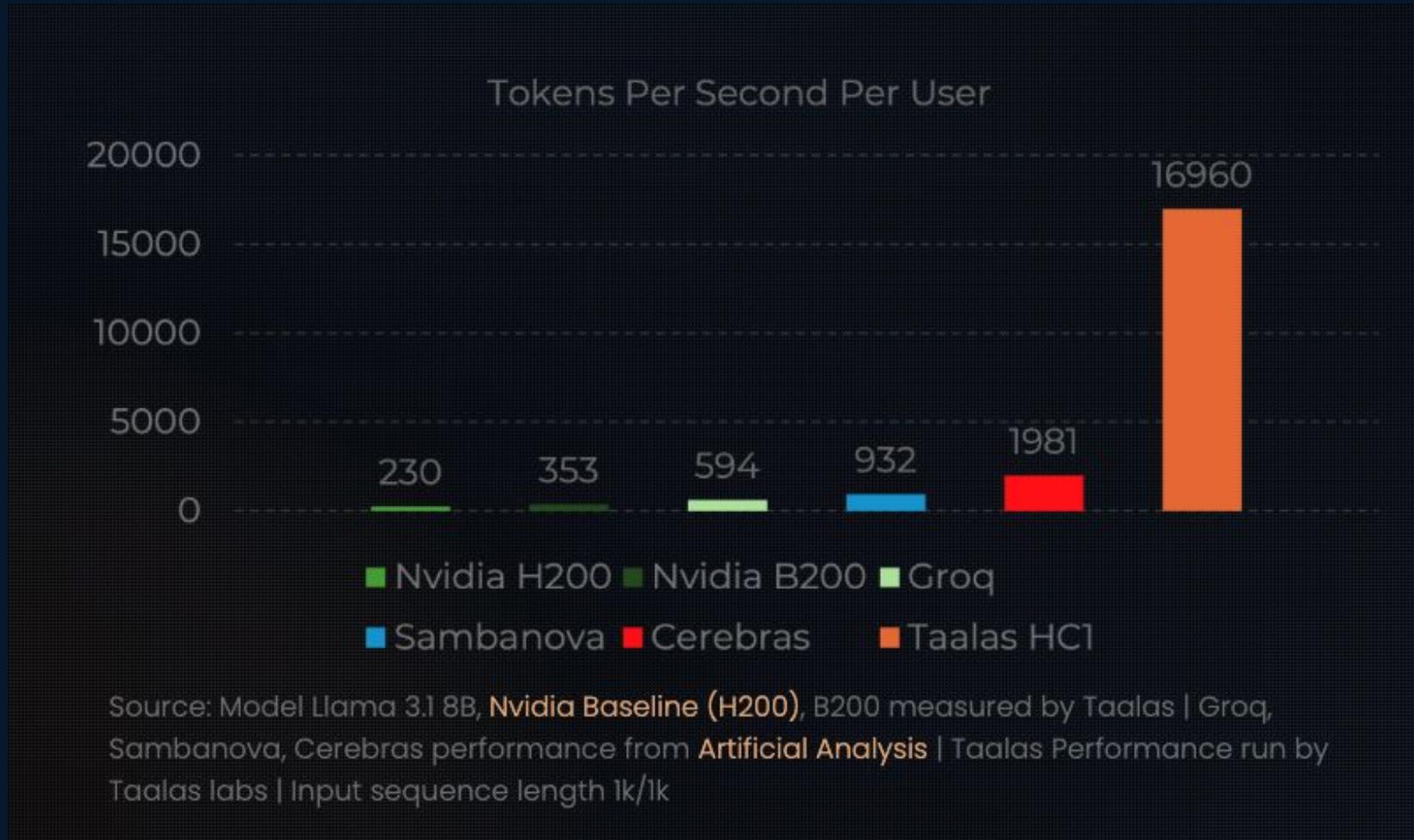
Bandwidth required upstream compared to downstream

36x

Increase in AI traffic as early as 2023-2024¹

1: Openrouter.ai/rankings
2: KPMG

Inference speed accelerating



How does AI Inference traffic look like?

Inference client

HTTPS (TCP or QUIC)

Inference server (LLM)



Figure 2. AI inference traffic volume evolution
Source: Cisco data

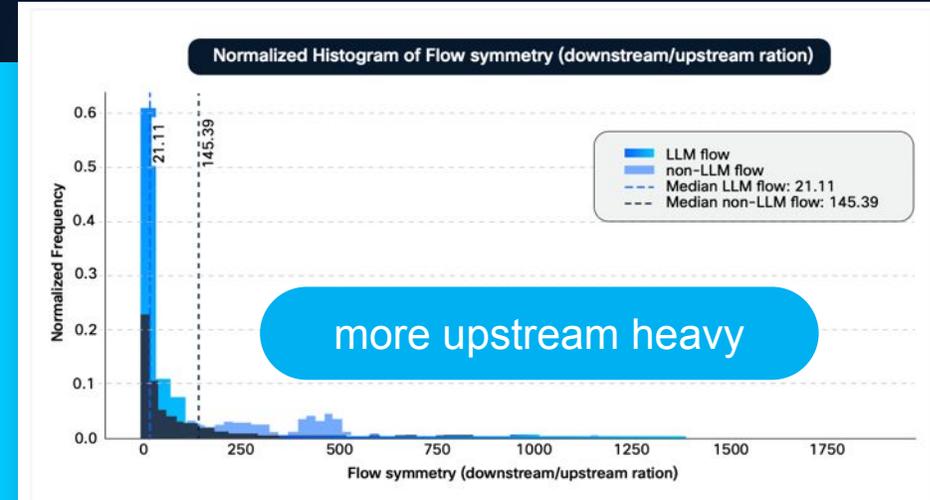


Figure 5. Traffic flow symmetry of AI inference flows vs. non-AI web transactions
Source: Cisco data

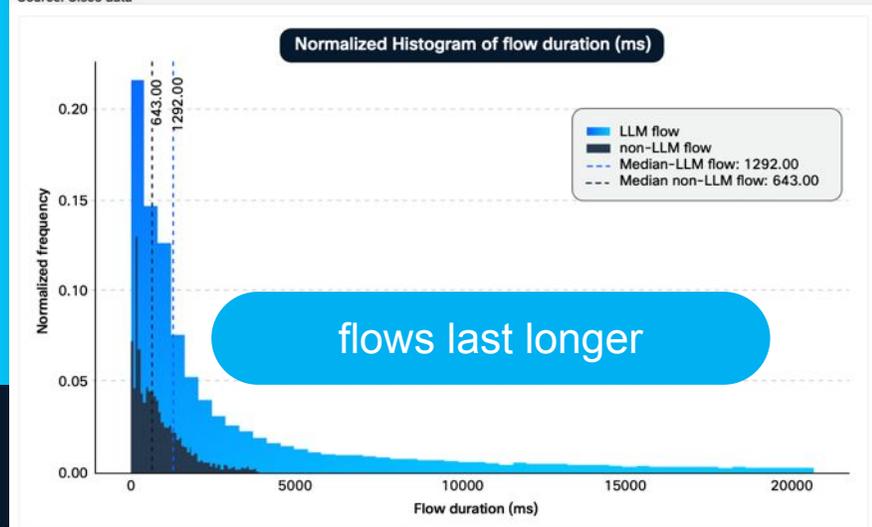


Figure 3. Flow duration distribution of AI inference flows vs. non-AI web transactions
Source: Cisco data

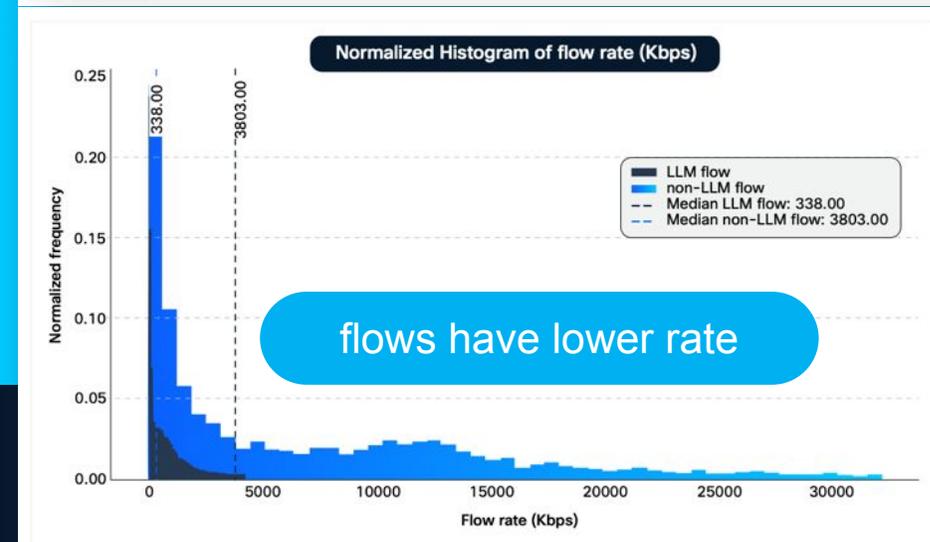
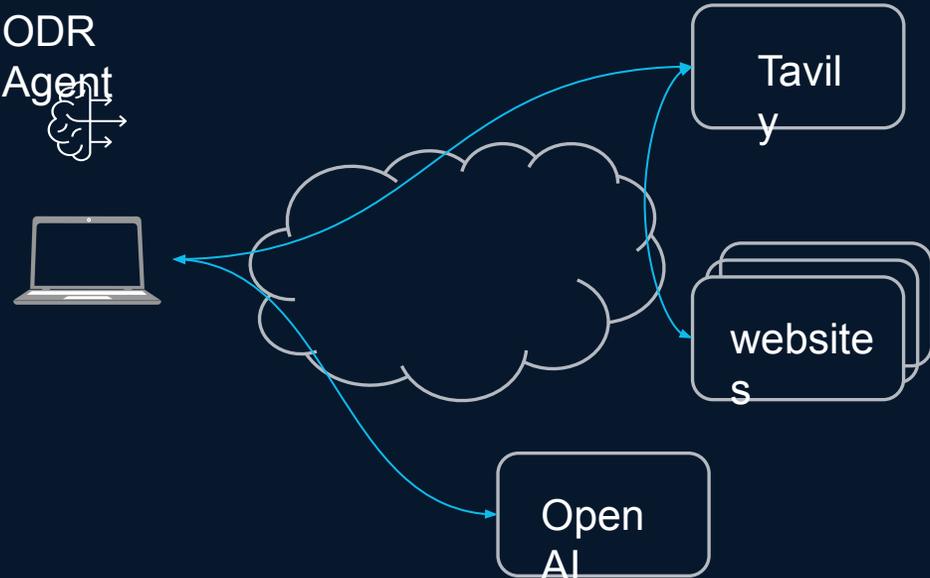


Figure 4. Flow rate distribution of AI inference flows vs. non-AI web transactions
Source: Cisco data

AI Agent traffic flows

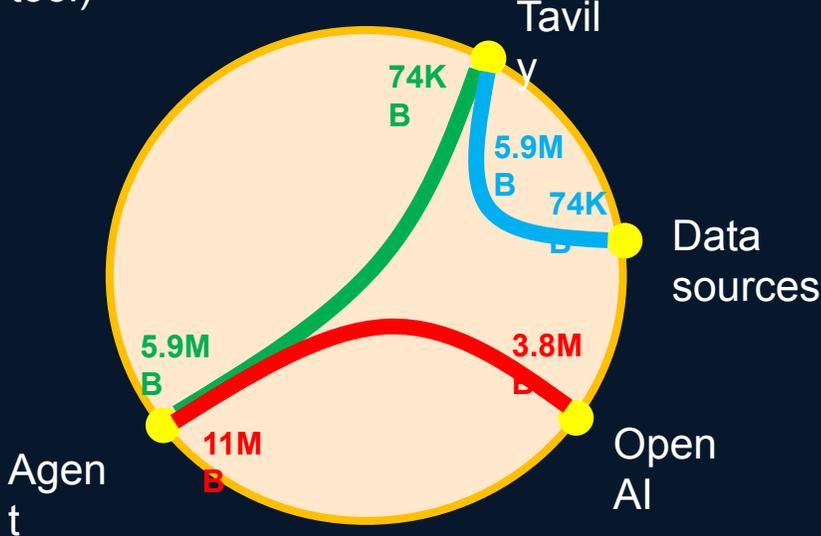
Soon in your network...

Example: Langchain Open Deep Research agent*



Task: "For Telia, gather public IR, socials, news, bets, partnerships, exec bios and social posts, organization, announcements, competitive"

44 different sources explored (via Taviily tool)



26.8M B
Agent triggered traffic

70%
LLM Inference traffic

4
web sources analyzed

450%
Agent driven traffic growth

Deep Network Troubleshooting agents



Agent Transaction details

Number of agents: 13	Number of LLM requests: 261	Total prompt tokens: 1033391	Total completion tokens: 121348
Transaction duration (s): 175.267037			

Agent(s)

261 LLM requests

1.12 million tokens

~4 bytes per token

4.5 MBytes

Human(s)

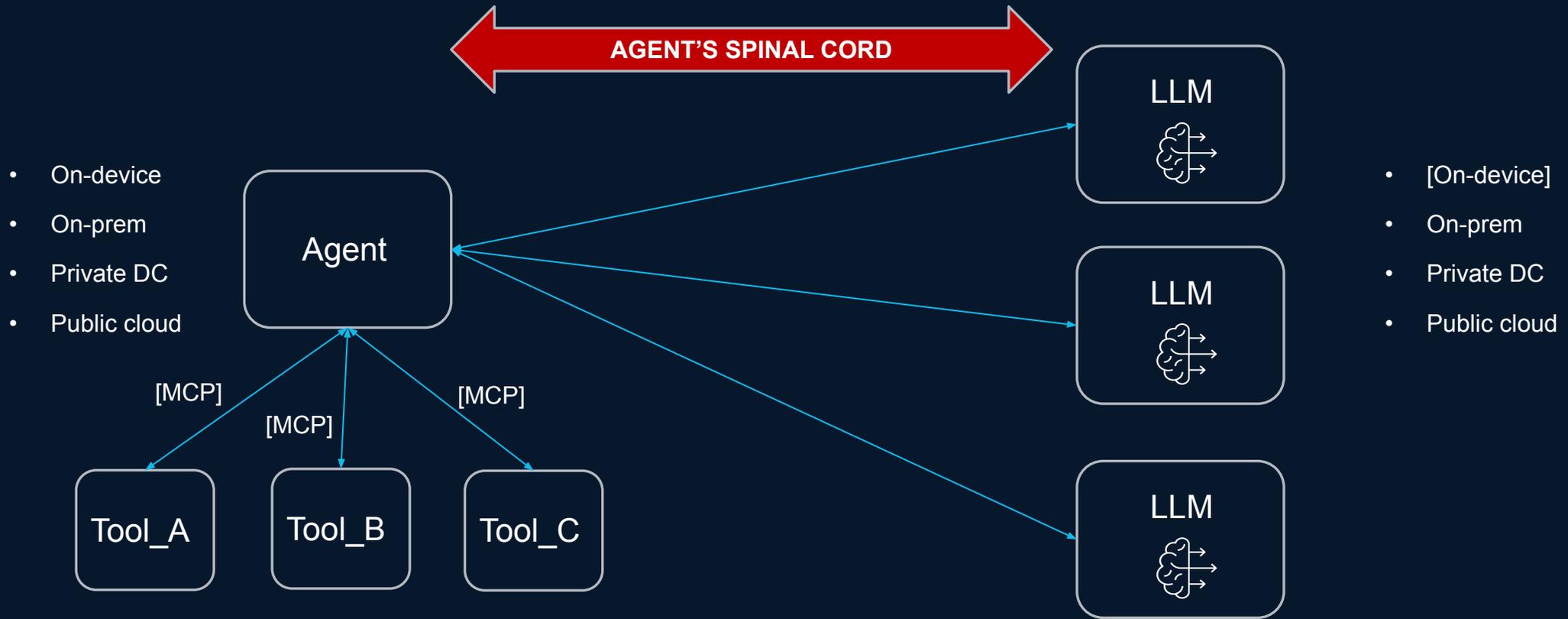
A few hundred Kbytes in CLI commands

Assume 1MByte

Agent multiplies network traffic by

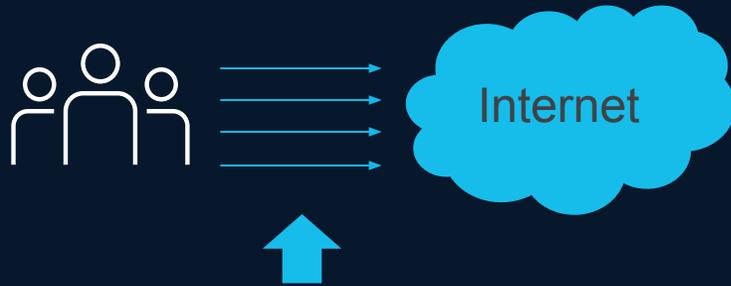
4.5x

The new critical path



Agents driving traffic multiplication

Network traffic in the Digital Era



Requested and processed at "human speed"

Network traffic in the Agentic Era



Requested and processed at "agent speed"



Gen AI

AI Agents

Digital Workers

Knowledge Graphs
Virtual Environments
LLMs
Evaluations on Research Domains

Reasoning Tools
LLMs
MCP
Observability
A2A

AI LLMs
SLMs
Deep Learning

The Digital Employee era is happening

NEWS | COMPENSATION, BENEFITS AND PAYROLL

Digital employees | BNY Mellon gives agentic AI 'staff' logins & human managers



BNY Mellon has revealed that dozens of its new AI-powered "digital employees" now operate with system logins and direct human line managers.

These tools are already performing tasks such as code writing and validating payment

Pro

Intel set to transfer marketing jobs to an AI which could ironically be running on Intel processors

News By Wayne Williams published June 26, 2025

Intel tells staff their jobs are being handed to Accenture

Comments (0)

AI EFFECT

Shopify CEO says staffers need to prove jobs can't be done by AI before asking for more headcount

PUBLISHED MON, APR 7 2025-2:16 PM EDT | UPDATED MON, APR 7 2025-2:58 PM EDT

Annie Palmer
@ANNIEPALMER/
#ANNIEPALMER

SHARE f X in

KEY POINTS

- Shopify CEO Tobin Lutke told employees in a memo that they'll have to show jobs can't be done by artificial intelligence before asking for more headcount and resources.
- Lutke said there's a "fundamental expectation" that employees are using AI in their day-to-day work.
- Tech companies have poured money into developing AI at the same time that they continue to cut jobs.



... doing up to 50% of the work at Salesforce, CEO Marc Benioff says

THU, JUN 26 2025-9:49 AM EDT | UPDATED THU, JUN 26 2025-1:46 PM EDT

Anthya Subin
ANTHYA_SUBIN

SHARE f X in

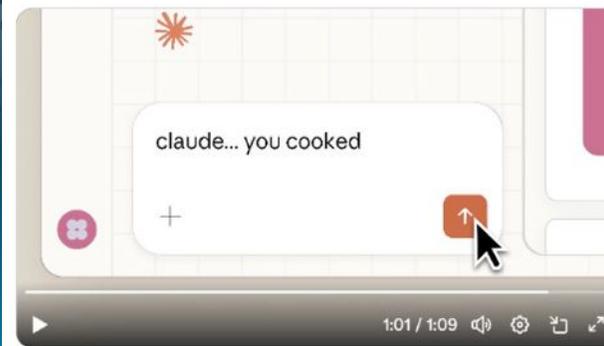
- Salesforce CEO Marc Benioff said artificial intelligence is accounting for 30% to 50% of the company's workload.

Technology

Claude
@claudeai

Introducing Cowork: Claude Code for the rest of your work.

Cowork lets you complete non-technical tasks much like how developers use Claude Code.



10:06 PM · Jan 12, 2026 · 24.7M Views

10:06 PM · 19U JS' S05B · 343M AIGAR

OpenAI launches consulting division... Making Mckinsey, Accenture, Big 4s nervous

TFS By The Finance Story Published Date: 17th July, 2025 Last Edited Date: 24th July, 2025





McKinsey CEO Bob Sternfels says the firm now has 60,000 employees: 25,000 of them are AI agents



Linas Beliūnas  · 2nd

Building a Safer Internet with AI  | Scouting for top startups to in...

[Visit my website](#)

14h · Edited · 

[+ Follow](#)

AI is eating consulting? McKinsey just confirmed it now counts AI as part of its workforce - 40,000 humans & 25,000 AI agents. That's 65,000 "employees" 🤖

Chief Executive Bob Sternfels made it official:

- AI already powers 40% of McKinsey projects
- Agents produced 2.5M charts in 6 months
- 1.5M human hours saved per year

The goal is to have 1 AI agent per human 🤖

McKinsey didn't fire people (yet). But they deleted entry-level work.

The boring parts are now gone:

- Desk research
- Slide drafting
- Data cleaning
- First-pass analysis

That work now happens at machine speed. Humans move up the stack instantly.

But the real bet the consulting giant is making isn't about productivity.

It's about optionality.

McKinsey is building a firm that can:

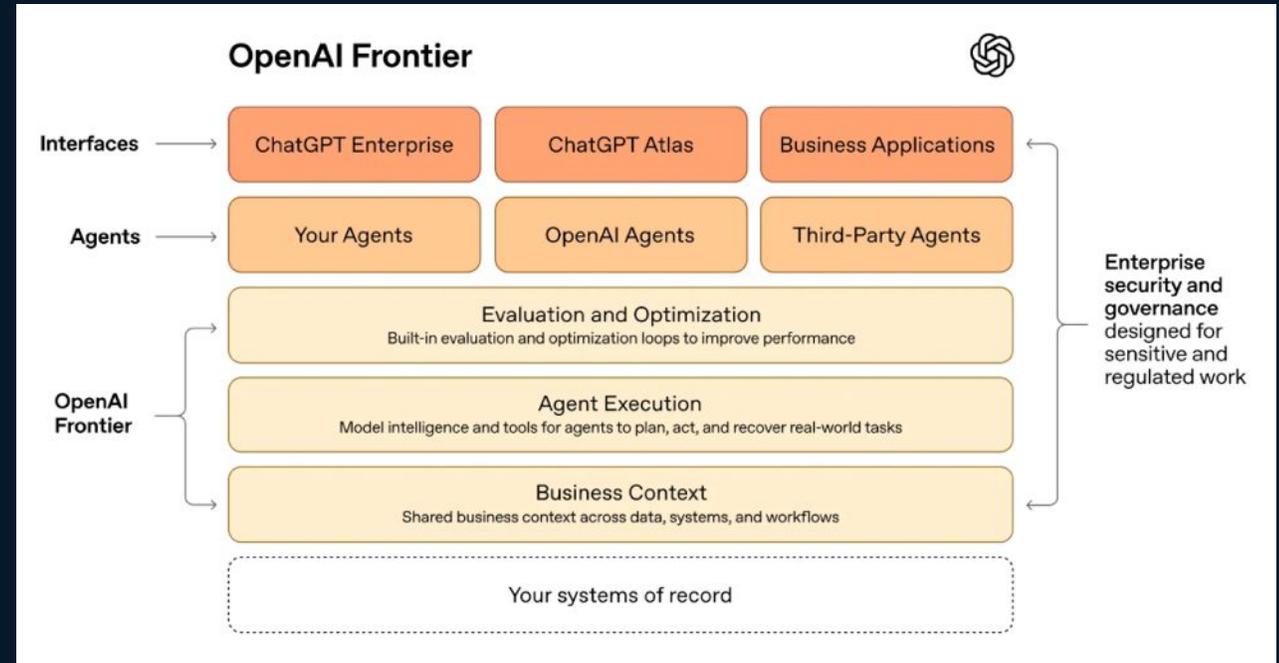
- Scale without hiring
- Serve more clients with the same headcount
- Recompose teams instantly
- Price expertise, not hours

Consulting then becomes software-like.

Today, we're introducing **Frontier**, a new platform that helps enterprises build, deploy, and manage AI agents that can do real work. Frontier gives agents the same skills people need to succeed at work: shared context, onboarding, hands-on learning with feedback, and clear permissions and boundaries. That's how teams move beyond isolated use cases to AI coworkers that work across the business.

For AI coworkers to actually work, a few things matter:

- They need to understand how work actually gets done across systems.
- They need access to a computer and tools to plan, act, and solve real-world problems.
- They need to understand what good looks like, so quality improves as the work changes.
- And they need an identity, permissions, and boundaries teams can trust.



The digital employee Era

Mechanize, a new A.I. start-up that has an **audacious goal of automating all jobs** — yours, mine, those of our doctors and lawyers, the people who write our software and design our buildings and care for our children.

“Our goal is to fully automate work,” said Tamay Besiroglu, 29, one of Mechanize’s founders. “We want to get to a fully automated economy, and make that happen as fast as possible.”

mechanize.work

Investors*

\$80/hr Average pay 44.3k Roles created \$1.5M+ Daily payouts

Shape the future of work

Find top-tier, remote, AI roles for your expertise.
Available only on Mercor.

Start working

Search roles

mercor.com

Opensource RL environments

PRIME Intellect

- Deploy GPU Instance
- Multi-Node Cluster
- Instances
- Templates
- Reserved Instances

Explore

- Environments
- Community Pools
- Compute Contributions

Account

- Profile
- Inbox
- Billing
- Keys

Support

- Chat
- Documentation

Terms of Service | Privacy Policy

Sign In

Environments Hub

A community hub for discovering and sharing environments, both for RL training and downstream evaluation

Explore | My Stars | Environments

Search by name, author, description, tags...

Coding | Data & ML | Games | Math & Reasoning | Multimodal | Science & Medicine | Search & Tool Use

Featured

Author	Stars	Environment Name	Description	Tags	Version	Updated
whyphylabs	6	vf-multi-swe-bench-openhands	Multi-SWE-Bench adapter environment for verifiers		0.1.0	Updated about 1 month ago
daniel	15	deepcoder	DeepCoder environment for coding problems with executable verification in Prime Intellect sandbox.	coding, sandbox, +1	0.2.3	Updated 3 days ago
will	6	tau2-bench	Verifiers implementation of tau2-bench	tool-use, user-sim, +3	0.1.5	Updated 25 days ago
arcee-ai	20	ifeval	IFEval single-turn chat environment using RLVF-IFEval with JSON constraint rewards. Heavily...	rivr, chat, +6	0.1.0	Updated about 1 month ago
adi-kmt	1	nyt-connections			0.1.1	Updated about 2 months ago
hud	21	hud-text-2048	Text-based 2048 game for training agents to reach target tiles through strategic moves	game, text, +2	0.1.0	Updated about 1 month ago

Community

Author	Stars	Environment Name	Description	Tags	Version	Updated
jayb	1					
plank-ai	0					
hud	15					
primeintellect	3					

How does the future look like?

Network impact of Digital
Workers

Capacity impact

Agentic AI adoption*

*source: Cisco research

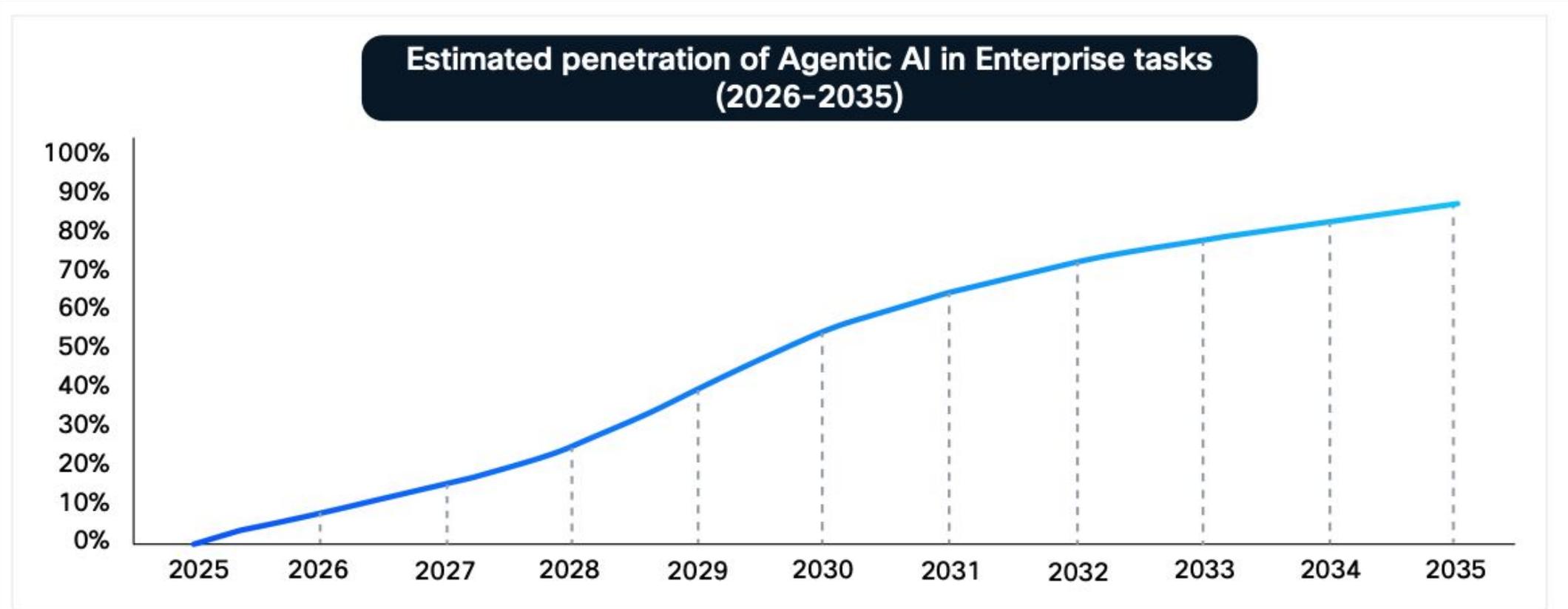


Figure 9. Projected adoption of agentic AI in enterprise tasks

Source: Cisco model

Agentic AI adoption*

*source: Cisco research

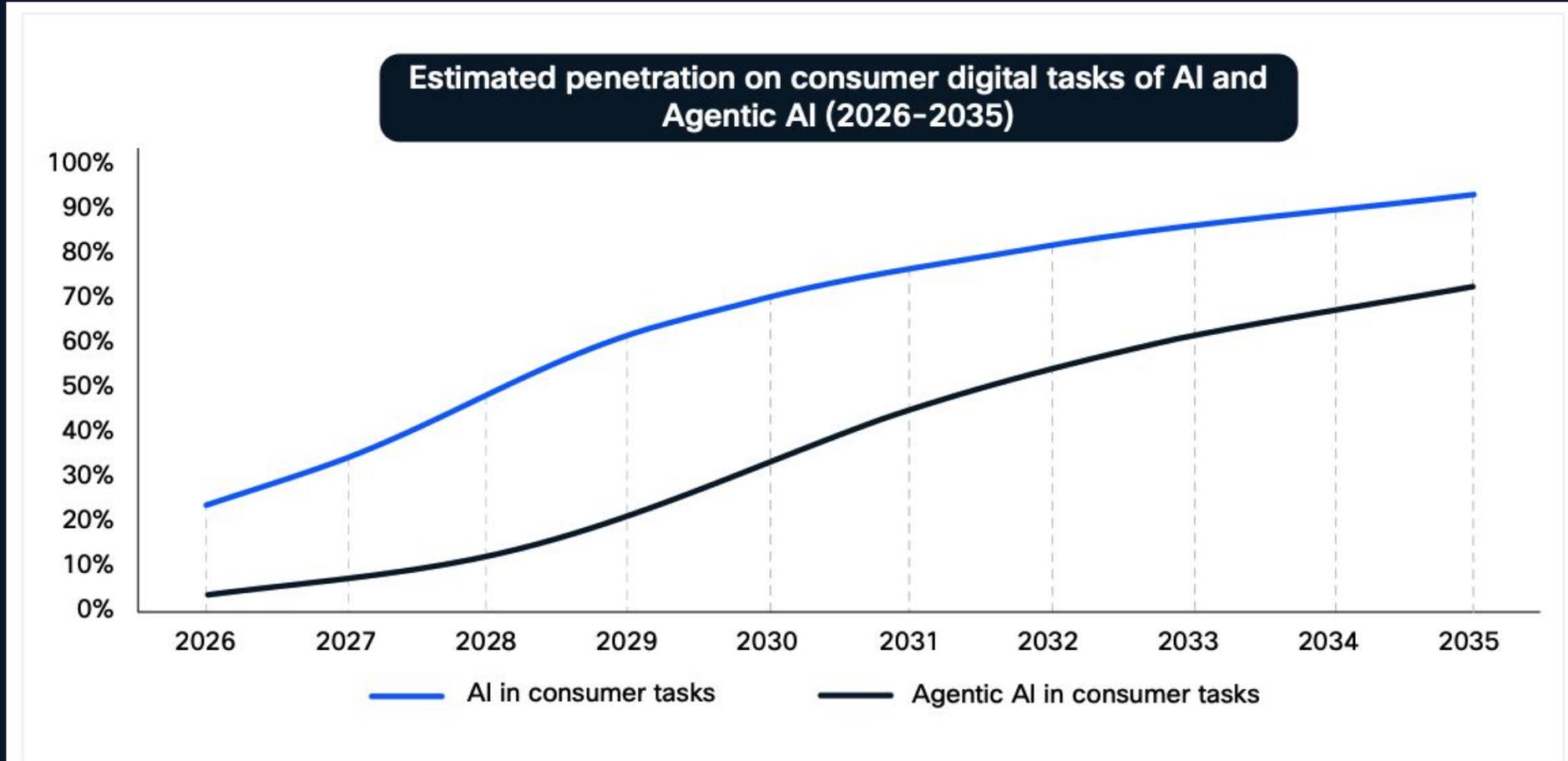


Figure 11. Projected adoption of AI and agentic AI in consumer digital tasks

Source: Cisco Model

Internet network traffic evolution with Agentic AI*

*source: Cisco research

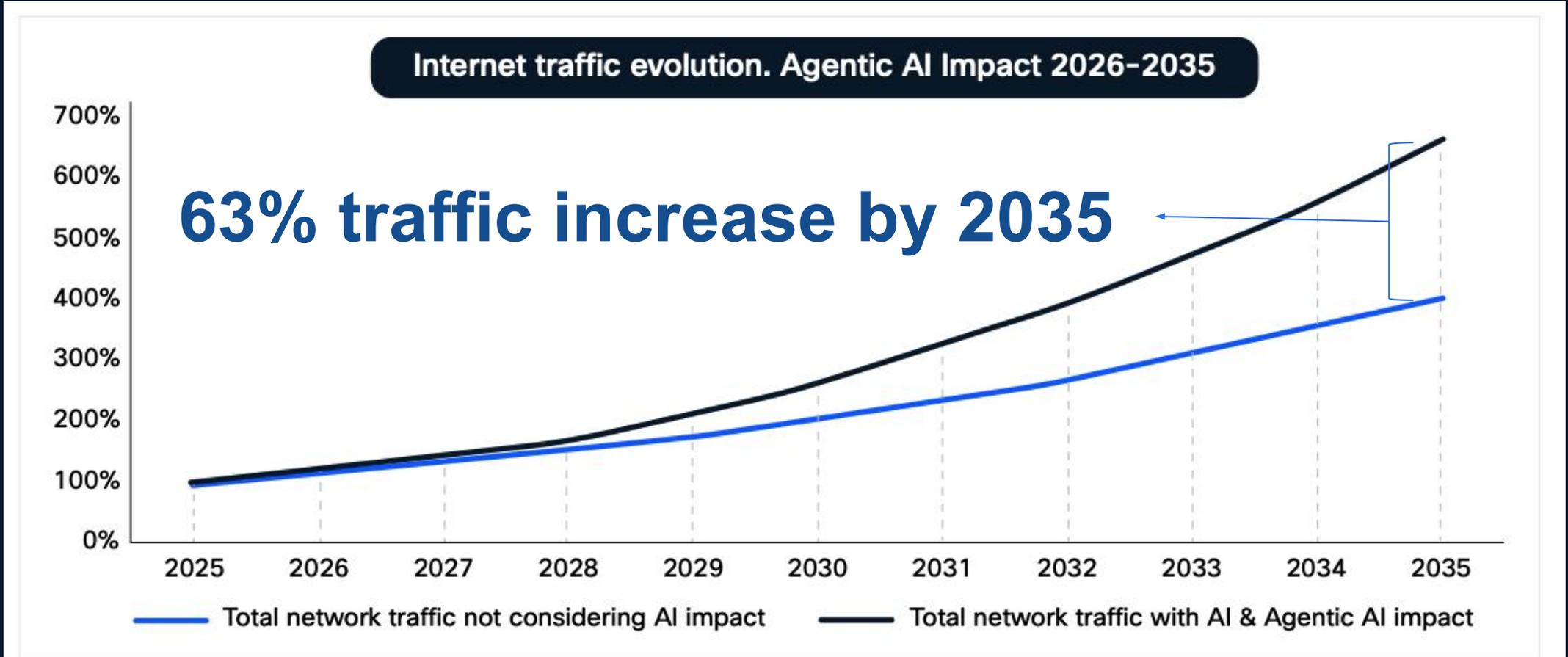


Figure 12. Projected impact of AI and agentic AI in internet network traffic

Source: Cisco model

AI Inference traffic growth*

*source: Cisco research

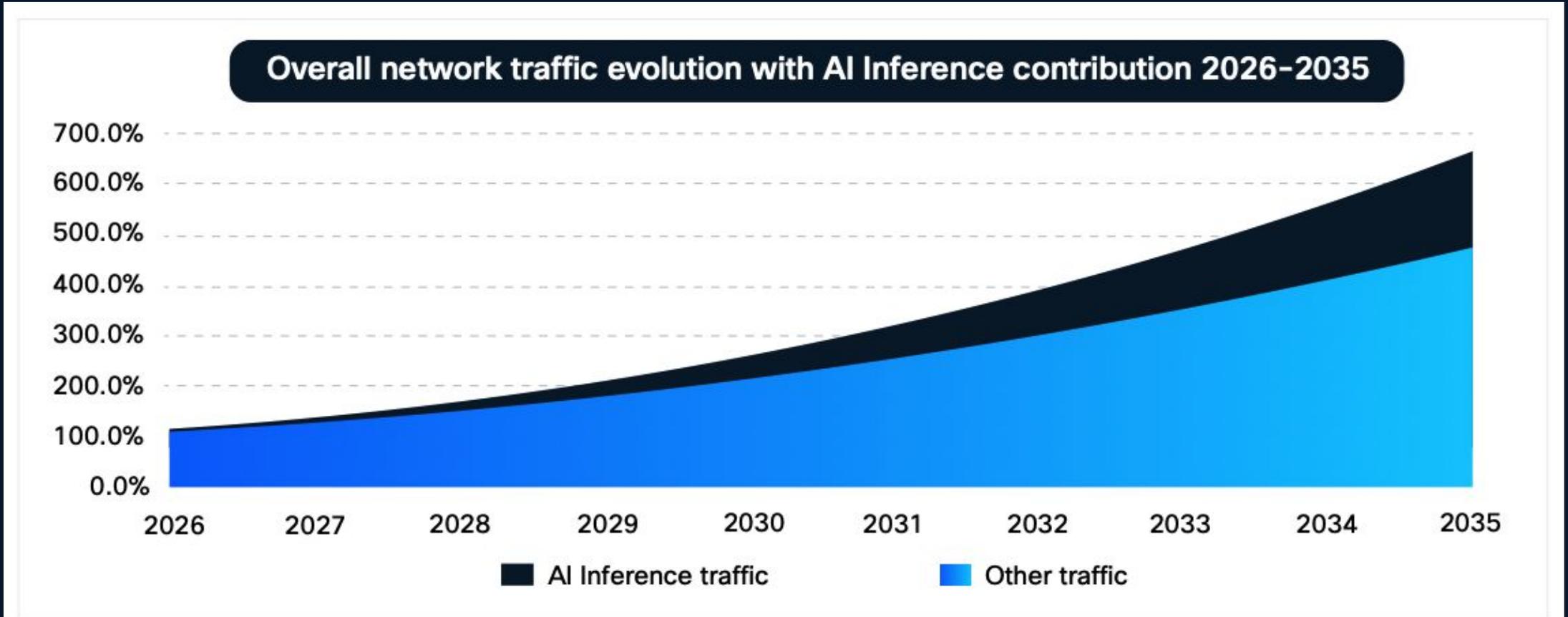


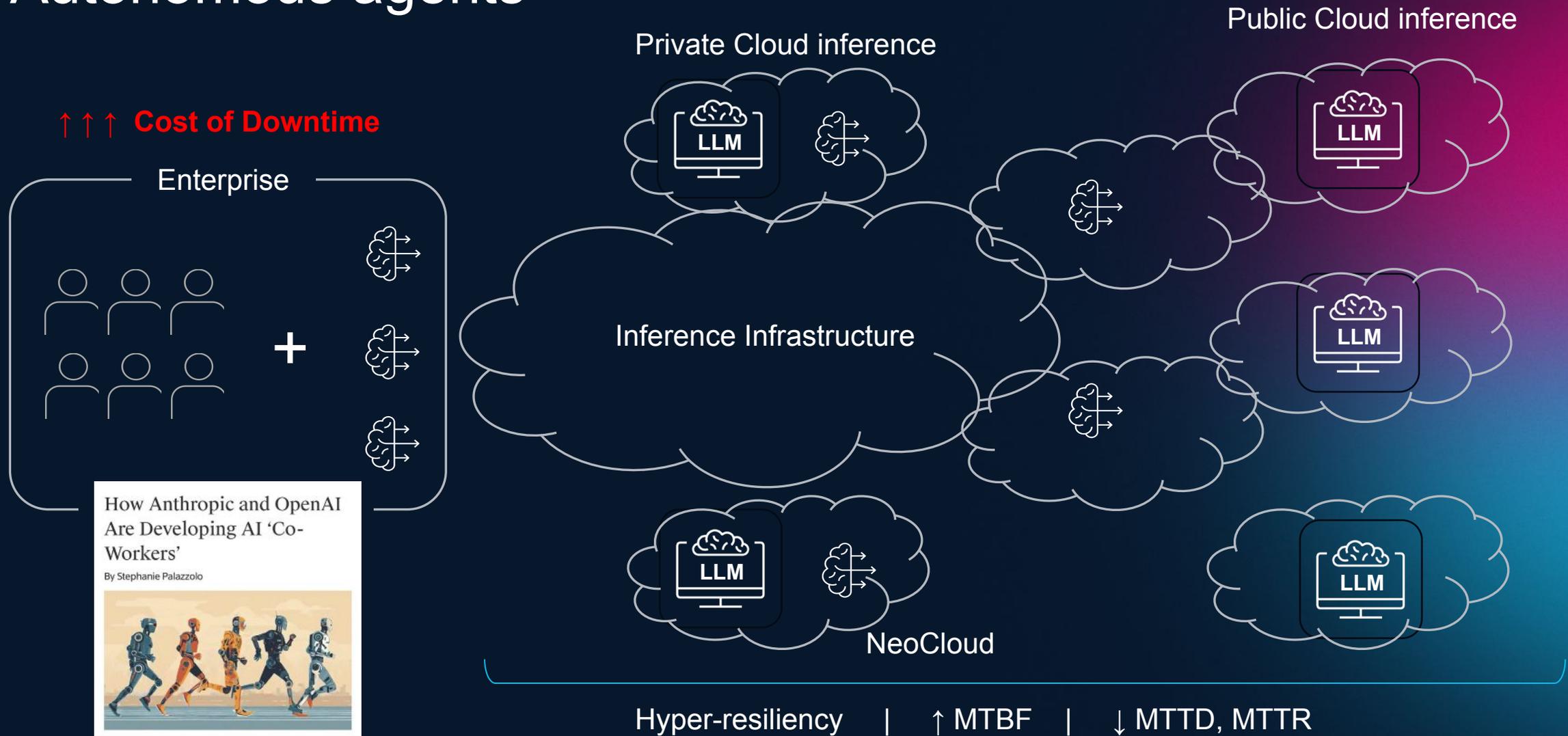
Figure 13. Projected network traffic growth, AI inference traffic versus other traffic

Source: Cisco model

Network impact of Digital Workers

The need for hyper resiliency

Digital employees enabled by AI Autonomous agents



Key takeaways

AI Inference traffic is here to stay and grow

Agentic AI is a major traffic multiplier

Inference will be critical to multiple applications

New opportunities emerging for SPs around AI Inference

