

從人類到代理： 未知人工智慧身分的風險與防護策略

Cisco Duo for Agentic Identity Security



Lance 朱育民
思科台灣資安事業部 總監

James 林彥均
思科台灣資安事業部 資深技術顧問

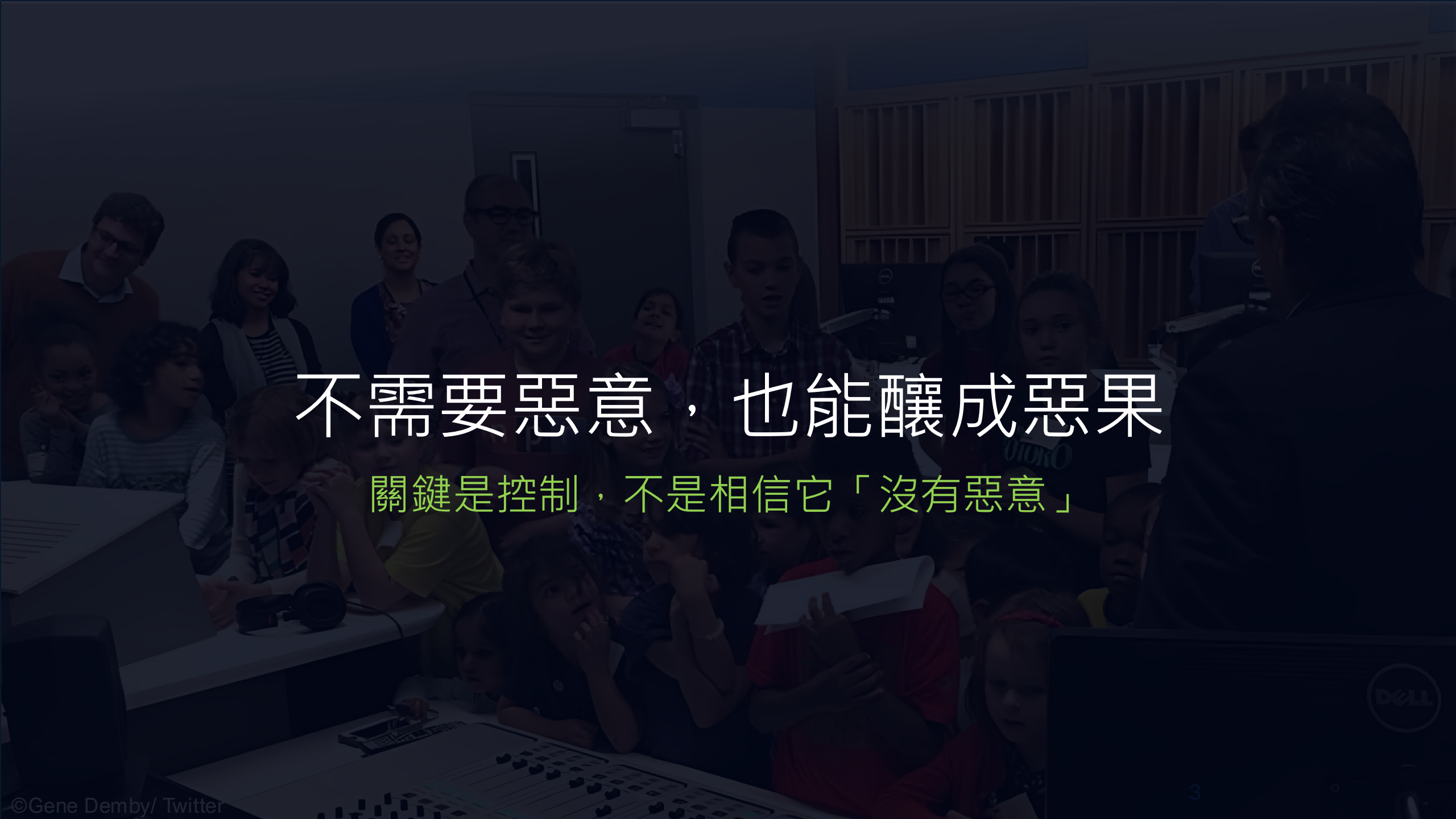
Protecting the World from the Agents



Lance 朱育民
思科台灣資安事業部
總監



James 林彥均
思科台灣資安事業部
資深技術顧問

A group of children and adults are gathered in a classroom or computer lab. They are looking towards a screen on the right side of the frame. The children are of various ages, from young kids to teenagers. Some are sitting at desks with computers, while others are standing. The adults are also looking in the same direction. The overall atmosphere is one of focused attention.

不需要惡意，也能釀成惡果

關鍵是控制，不是相信它「沒有惡意」

我們需要為 Agentic Workforce 重新設計安全機制

保護 人工智慧代理

保護 Agent, 不被外界侵害

保護外界, 不受Agent所害

用機器級的速度, 即時偵測、即時反應

AI agent，是一種全新的身分類別



人類 Humans

廣泛存取資源的權限

運作速度慢

具備判斷與倫理



機器 Machines

存取資源有限

運作速度快

僵化、照規則執行



代理 Agents

廣泛存取資源的權限

運作速度快

缺乏有效的判斷力跟外部約束

AI agents 「致命三要素 (lethal trifecta)」



能存取私密資料
(Access to private data)



會接觸不可信內容
(Exposure to untrusted content)



能對外通訊／上傳
(Ability to communicate externally)

Source: <https://simonwillison.net/2025/Jun/16/the-lethal-trifecta>

Zero Trust 必須進化



Humans

Access

證明 身分
信任 使用者



Agents

Actions

驗證 行為
控制 Agent

Agent環境複雜，要把 Zero Trust 做到每一個動作，真的很難



AI agent 無所不在
(Agents Everywhere)



任務驅動的行為
(Task-driven Behavior)



破碎、分散的管控
(Fragmented Enforcement)



CISCO

Security



Cisco Duo 為 AI Agent 打造的 Zero Trust

Cisco Duo 為 AI Agent 打造的 Zero Trust

認得每一個 Agent
(Know every agent)

授權每一個動作
(Authorize every action)

如果有風險即時應對
(Adapt to risk in real time)

← 在AI Agent存取工具與資料的接入點，統一管控對應的授權 →



認得每一個 Agent
(Know every agent)

授權每一個動作
(Authorize every action)

如果有風險即時應對
(Adapt to risk in real time)

Cisco Duo 能為每一個與你環境互動的 AI agent ,
建立 能見度(Visibility) 、身分(Identity) 與 權責歸屬(Ownership)



Agentic IAM by Duo

Agent 與工具探索
(Agent & Tool Discovery)

Agent 目錄
(Agent Directory)

Agent 存取政策
(Agent Access Policy)

人為權責歸屬
(Human Accountability)

生命週期管理
(Lifecycle Management)

認得每一個 Agent
(Know every agent)



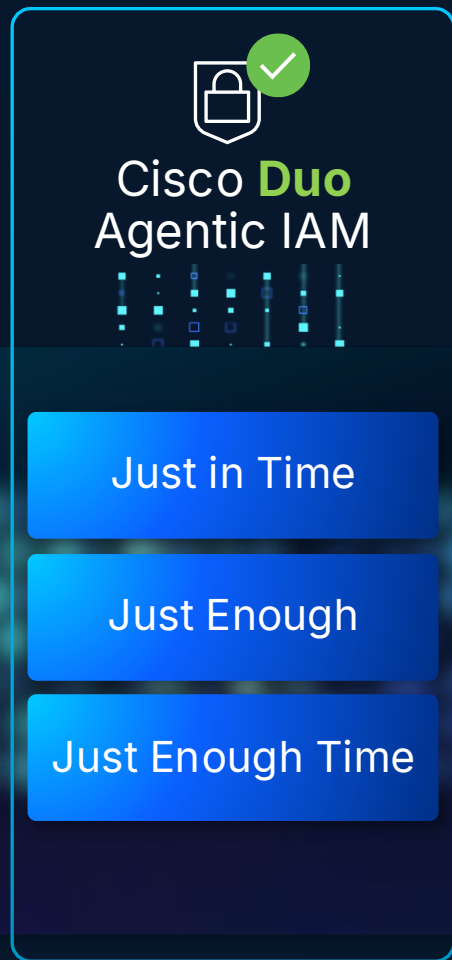
授權每一個動作
(Authorize every action)

如果有風險即時應對
(Adapt to risk in real time)

Cisco Duo 能在 AI Agent 存取 企業工具 與 資料 的接入點 統一管控對應的授權



Agentic Workforce



工具、資源與資料





Cisco Duo 能在「每次發起需求時，自動發 token」 ——每次授權都**即時、夠用、剛剛好**



Just in Time

要用的那一刻才產生，用完即逝；
平常沒有長期憑證放著。



Just Enough

只開「這一個資源、這一個動作」
靠 Resource Indicator 限定目標。



Just Enough Time

幾分鐘就過期；
就算被偷也來不及用。

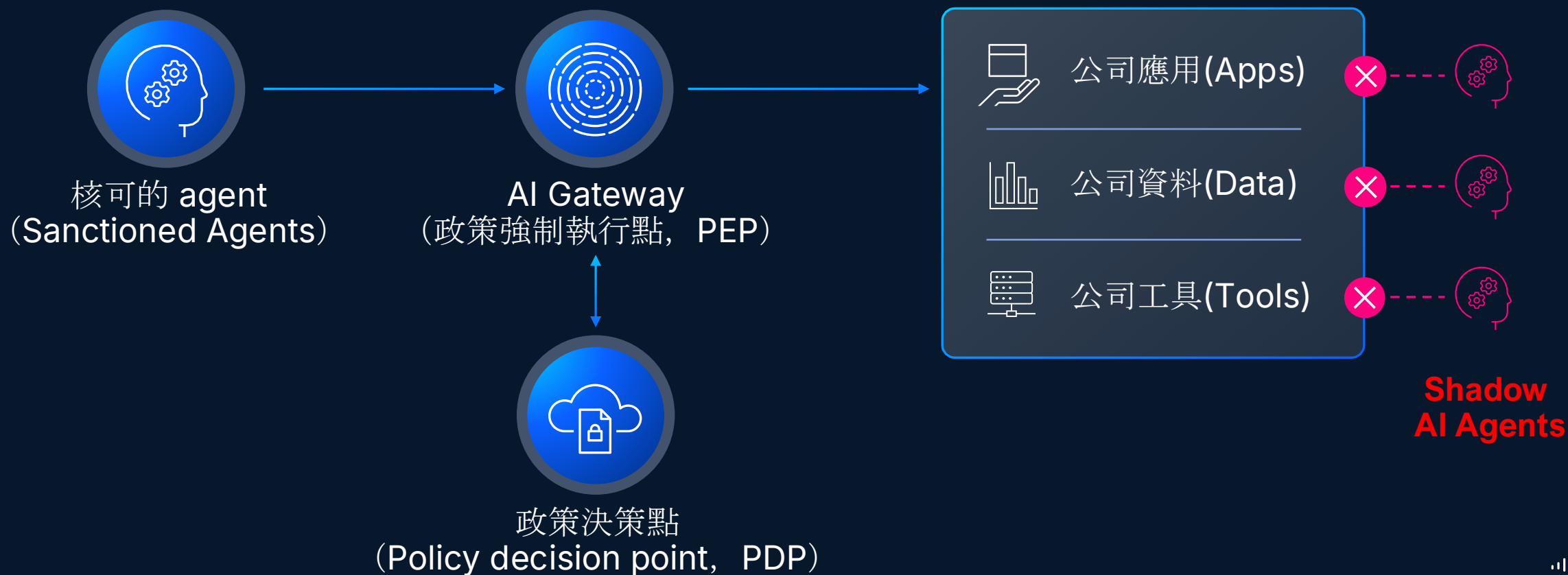
傳統做法十分危險：永久、萬能、靜態的 API 存取授權；一旦洩漏即永久、全開。

這三個 J 提供的是機制能力，而不是『預設就保證很嚴格』。

嚴不嚴，看 owner 的設定：scope 設太寬、效期設太長，等於自己架空了它。所以真要追究，那是組態過於寬鬆的問題，不是機制無效。



Cisco Duo能在「企業工具邊界」落實管控：SaaS + Private Apps



認得每一個 Agent
(Know every agent)

授權每一個動作
(Authorize every action)



如果有風險即時應對
(Adapt to risk in real time)

Cisco Duo能以**意圖**、**威脅**與**行為**情境來動態調整 AI Agent 的防護決策

針對安全與資安威脅的 **runtime** 防護

威脅情境
(Threat context)

行為情境
(Behavioral context)



通訊
(Communication)

即時監控與管控

認得每一個 Agent
(Know every agent)

授權每一個動作
(Authorize every action)

如果有風險即時應對
(Adapt to risk in real time)

無論哪一種 Agent, 存取資料與應用 Cisco Duo 都建立在 identity-aware 的 Zero Trust 之上



認得每一個 Agent
(Know every agent)

授權每一個動作
(Authorize every action)

如果有風險即時應對
(Adapt to risk in real time)

無論哪一種 Agent, 存取資料與應用 Cisco Duo 都建立在 identity-aware 的 Zero Trust 之上

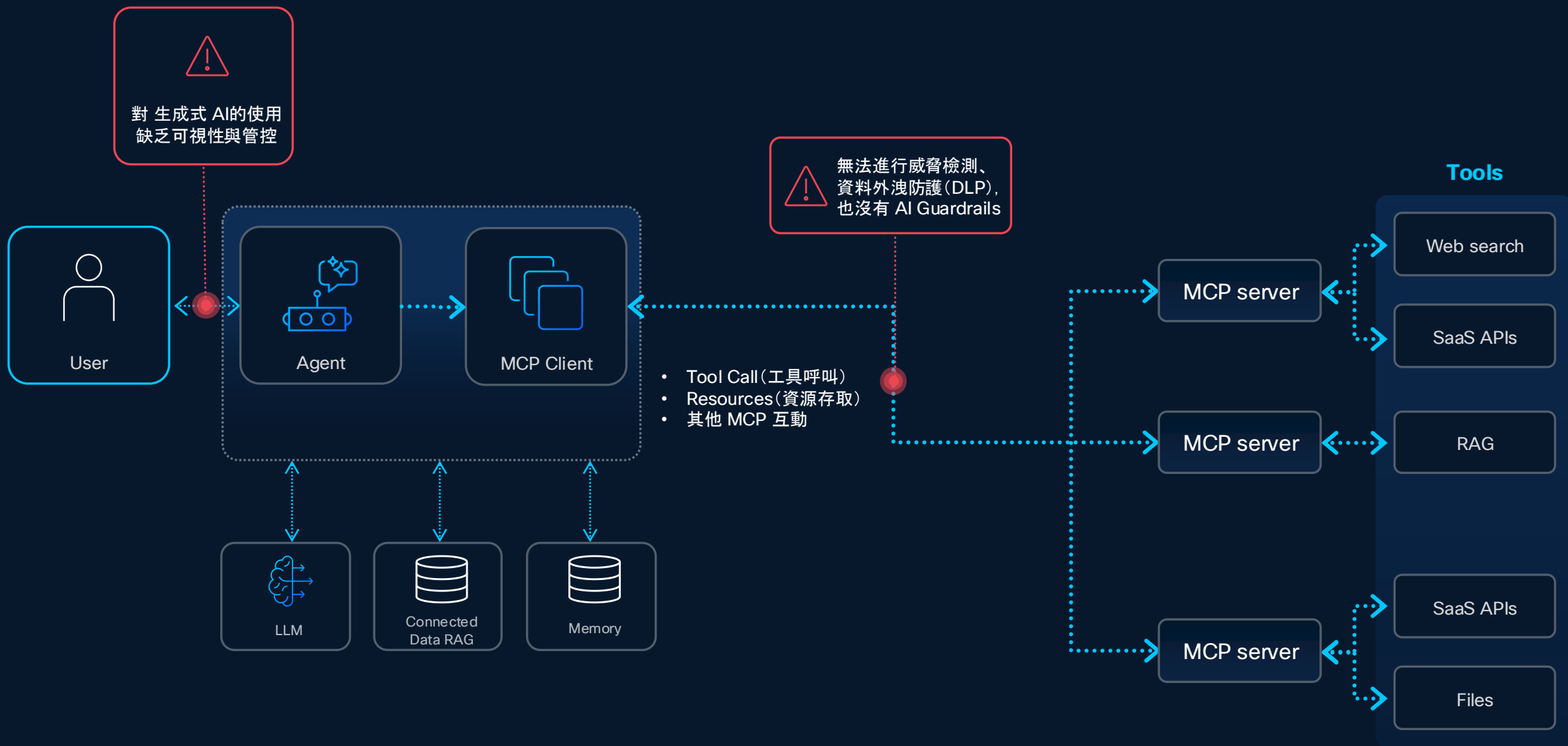




Demo

Duo IAM、MCP 查核機制 架構說明

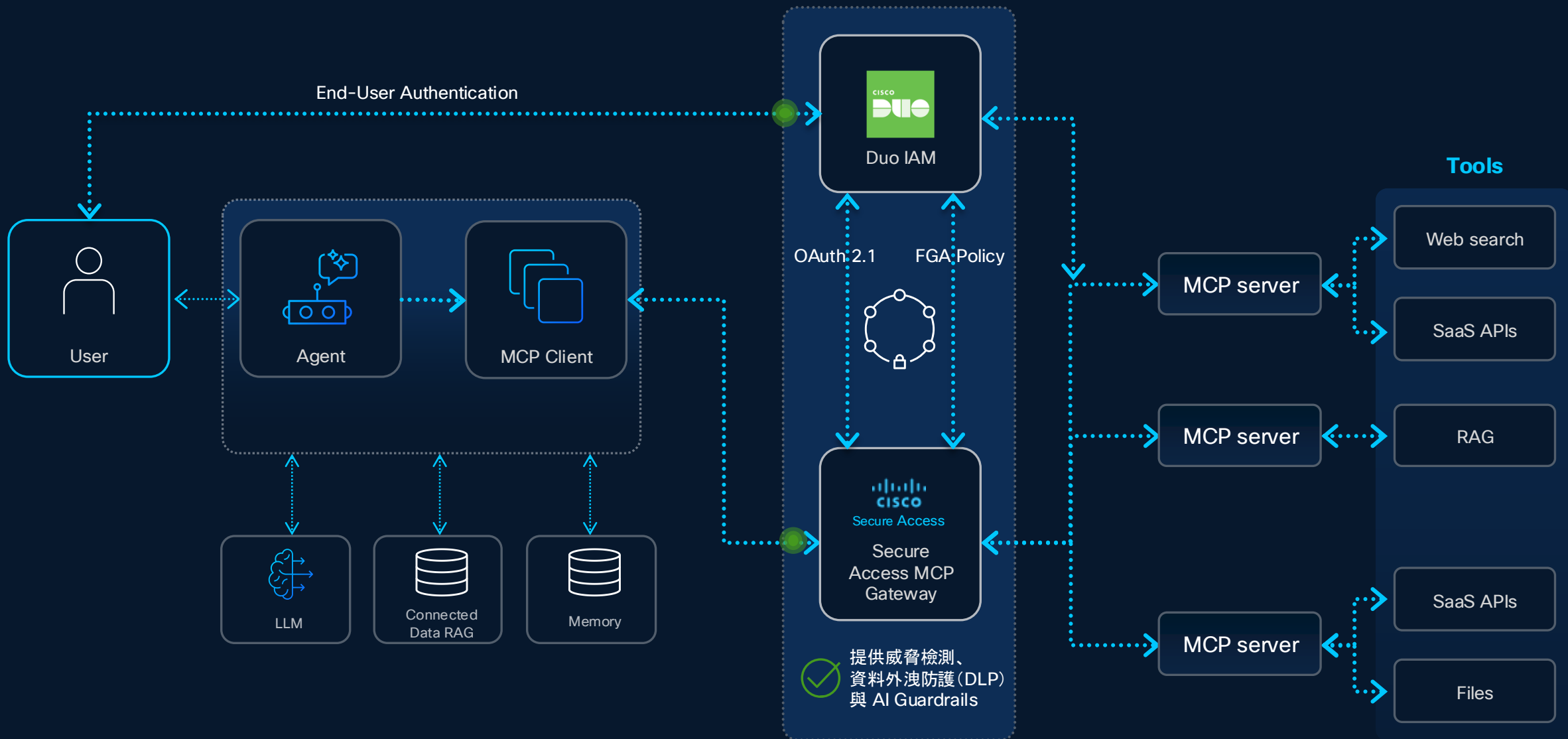
AI agent 流量檢測與威脅防護



AI agent 流量檢測與威脅防護

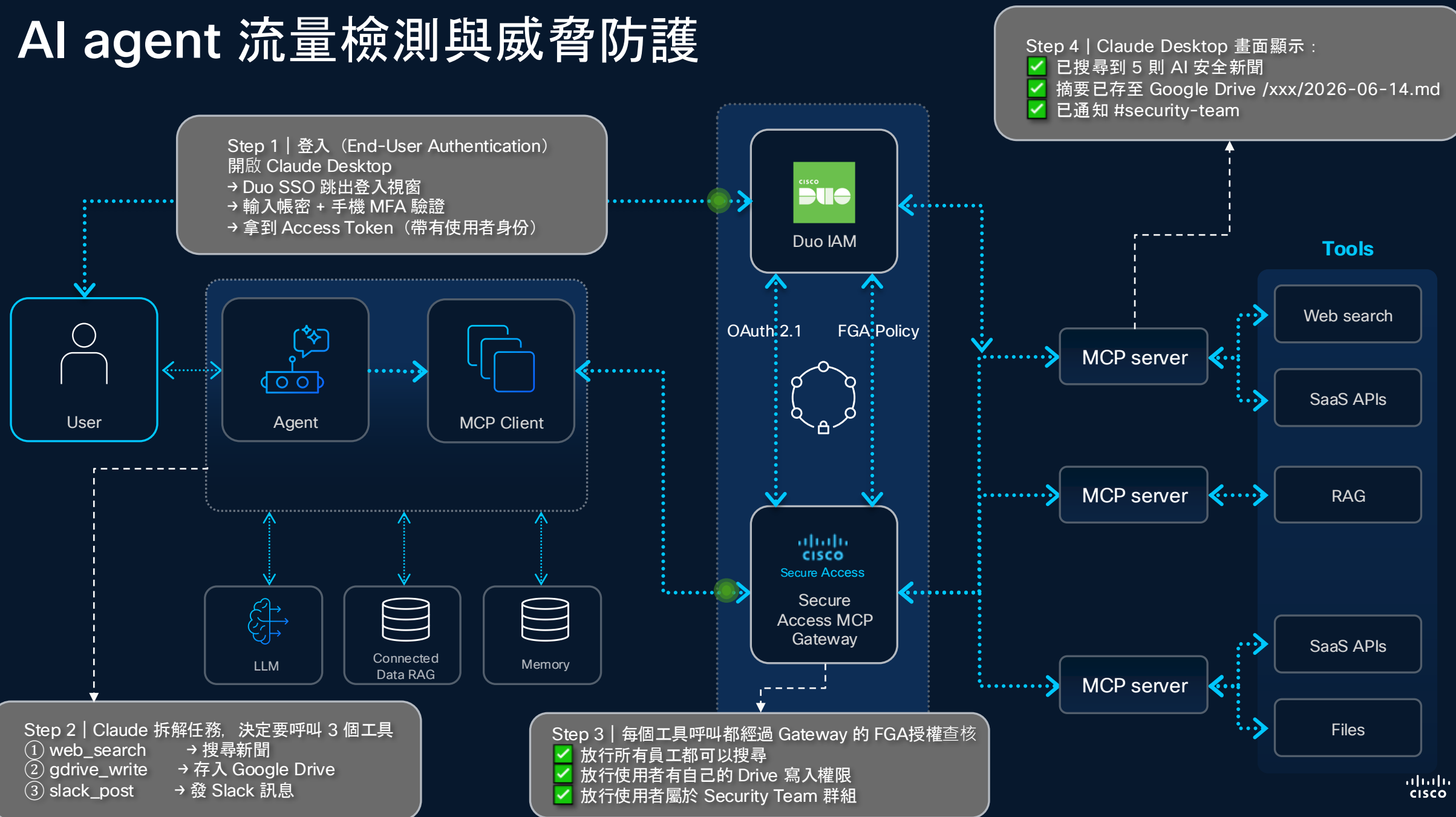


提供對 生成式 AI
使用的完整可視性與管控



提供威脅檢測、
資料外洩防護 (DLP)
與 AI Guardrails

AI agent 流量檢測與威脅防護





Demo

Duo MCP 授權控制 與 語意分析

🔍 Chat

📑 Cowork

📄 Code

+ New chat

📁 Projects

🔗 Artifacts

⚙️ Customize

Pinned

Cisco Zero Trust Agentic AI probing問題

Snort防火牆開單自動化製作

🗨️ Ask Joker AI Lab

Snort Everywhere

Recents

CiscoFMC MCP demo for ZTAi Agent

AI Agent資安三大支柱

RTP-FTD-UZTNA 設備的規則列表

RTP-FTD-UZTNA 設備的規則列表

進展到哪了

ISE VM遷移至Nutanix備份方案

查詢 IP 198.18.5.31 的存取規則

查詢 IP 198.18.5.31 的存取規則

Access Control Policy 規則統計

Secure Workload 防火牆存取規則

思科交换机系列与IOS版本差异

Secure Access internet backhaul tunnel 5

工具識別

Syslog設定與Splunk資料接收問題

CSCO股價查詢

使用 Math-MCP 計算简单加法

用 Math-MCP 計算乘法

Syslog設定與Splunk資料接收問題

使用 math-mcp 計算乘法

使用 math-mcp 計算乘法

Show 10 more

J James · Joker AI Lab

⌵

📄

🌟 Golden hour thinking

Claude Fable 5 is currently unavailable.

[Learn more](#)



How can I help you today?



Sonnet 4.6 Low



📄 Code

📝 Write

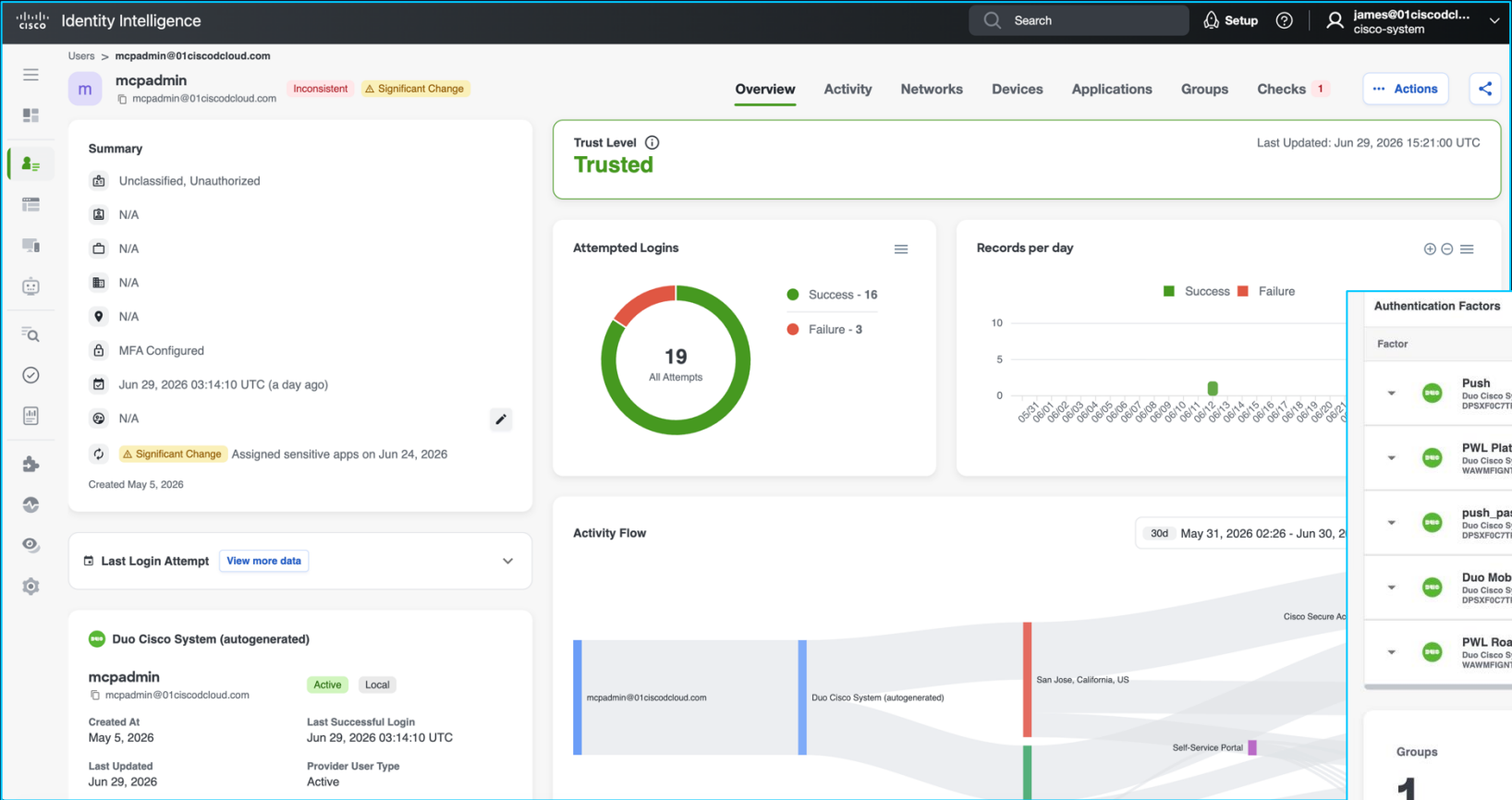
📁 From Drive

📅 From Calendar

📧 From Gmail

Session ID: BRKSEC-XXXX

透過 Cisco Duo Identity Intelligence 全面解析身份存取歷程



Authentication Factors

Factor	Assurance Level	Status	# Changes	Usage Count	Device	Phone Number	Last
Push Duo Cisco System (autogenerated) DPSXF0C7TPPDE1OF69FX	Medium	ACTIVE	0	7	Google Pixel 10 Pro	N/A	Jun
PWL Platform Authenticator Duo Cisco System (autogenerated) WAWMPFGNT90347TXYE0X	High	ACTIVE	1	6	iCloud Keychain	N/A	Jun
push_passwordless Duo Cisco System (autogenerated) DPSXF0C7TPPDE1OF69FX	Medium	N/A	0	3	Google Pixel 10 Pro	N/A	Jun
Duo Mobile Duo Cisco System (autogenerated) DPSXF0C7TPPDE1OF69FX	Medium	ACTIVE	0	0	Google Pixel 10 Pro	N/A	N/A
PWL Roaming Authenticators Duo Cisco System (autogenerated) WAWMPFGNT90347TXYE0X	High	ACTIVE	1	0	iCloud Keychain	N/A	N/A

mcpadmin@01ciscodcloud.com 的
連線時間、來源、目的、信任等級、單次事件紀錄

信任等級的分類

Cisco Identity Intelligence weighs several factors together in a proprietary algorithm to produce a User Trust Level for each user in your organization, which ranges from Untrusted to Trusted:



- **Trusted** 表示該使用者具有極高的安全程度



- **Favorable** 表示該使用者具有良好的安全程度



- **Neutral** 表示該使用者已被評估，且沒有正面或負面的行為表現



- **Questionable** 表示該使用者展現出的行為可能暗示風險，或可能屬於不理想的行為



- **Untrusted** 表示該使用者展現出的行為非常不佳、惡意，或屬於不理想的行為



- **Unknown** — 使用者也可能具有 Unknown Trust Level，這表示該使用者先前未被評估，或缺乏近期帳號活動，特別是成功登入活動，因此無法判定其信任等級。



Demo

Duo 細緻化授權 (Fine-Grained Authorization)

ChatCoworkCode

New chat

Projects

Artifacts

Customize

Pinned

Cisco Zero Trust Agentic AI probing問題

Snort防火牆開單自動化製作

Ask Joker AI Lab

Snort Everywhere

Recents

RTP-FTD-UZTNA 設備的規則列表

進展到哪了

ISE VM遷移至Nutanix備份方案

CiscoFMC MCP demo for ZT AI Agent

查詢 IP 198.18.5.31 的存取規則

查詢 IP 198.18.5.31 的存取規則

Access Control Policy 規則統計

Secure Workload 防火牆存取規則

思科交换机系列与IOS版本差异

Secure Access internet backhaul tunnel S

工具識別

Syslog設定與Splunk資料接收問題

CSCO股價查詢

使用 Math-MCP 計算简单加法

用 Math-MCP 計算乘法

Syslog設定與Splunk資料接收問題

使用 math-mcp 計算乘法

使用 math-mcp 計算乘法

查找case的bug ID和hotfix

Secure Firewall SD-WAN的QoS功能限制

Show 11 more

James · Joker AI Lab

Coffee and Claude time?

Claude Fable 5 is currently unavailable.

Learn more

How can I help you today?

+

Sonnet 4.6 Low

</> Code

Write

From Drive

From Calendar

From Gmail

MCP Gateway 要求 - 通過列表

[illegible]

Event Details

Action

✔ Allowed

動作:驗證通過

Block Reason

Time

Jun 30, 2026 10:05 AM

發生時間

Access details

Identity

 **mcpadmin (mcpadmin)**

這是誰:已註冊的服務帳號/agent

Resource/Application

Destination

CiscoFMC MCP

Destination IP

198.18.5.35

Destination Port

52081

Transaction ID

2f105b180846117c

MCP 存取要求觸發過程詳細內容

Event Type	Status	Event ID ⓘ	Source	Destination	Reason Code	Rule Name	Time	
▼ MCP Access	✔ Allowed	2f105b180846117c	mcpadmin (mcpad...	CiscoFMC MCP	-	SPA	Jun 30, 2026 10:05 AM	
誰發出請求 用什麼方式連接 安全檢查結果 存取目的地								
Source AD Users: mcpadmin (mcpadmi... Agent name: - Agent ID: 0 Source IP: 160.79.106.35 Location: - Connection Type: MCP Gateway (via Z... Security Controls MCP Session ID: - ^ MCP Gateway (1) Allow: 1 View all Action: Allowed ∨ Tool calls Destination FQDN: CiscoFMC MCP MCP Server: - Destination IP: 198.18.5.35 Destination Port: 52081								

MCP Gateway 要求 - 阻擋列表

[illegible]

Event Details

Action

Blocked

動作:驗證不過, 阻擋

Block Reason

auth_policy_blockaccess blocked by policy rule 0

Time

Jun 30, 2026 10:05 AM

發生時間

Access details

Identity

 **mcpadmin (mcpadmin)**

這是誰:已註冊的服務帳號/agent

Resource/Application

Destination

CiscoFMC MCP

Destination IP

198.18.5.35

Destination Port

52081

Transaction ID

c7d49bac1a92dd45

MCP Details

MCP 存取要求觸發過程詳細內容

Event Type	Status	Event ID ⓘ	Source	Destination	Reason Code	Rule Name	Time	
 MCP Tool Call	 Blocked	c7d49bac1a92dd45	mcpadmin (mcpad...	CiscoFMC MCP	auth_policy_blockac...	—	Jun 30, 2026 10:05 AM	

誰發出請求



✔ Source

AD Users: [mcpadmin \(mcpadmi...](#)
Agent name: agent
Agent ID: 1
Source IP: 160.79.106.35
Location: -

用什麼方式連接



✔ Connection

Type: MCP Gateway (via Z...

安全檢查結果



✖ Security Controls

MCP Session ID: e6c0c7fa7dde4643a...

▼ MCP Gateway (1)

^ Tool calls

Action: Blocked
Tool name: list_fmc_profiles
Duo rule name: —

存取目的地



➖ Destination

FQDN: [CiscoFMC MCP](#) 

MCP Server: -
Destination IP: 198.18.5.35
Destination Port: 52081



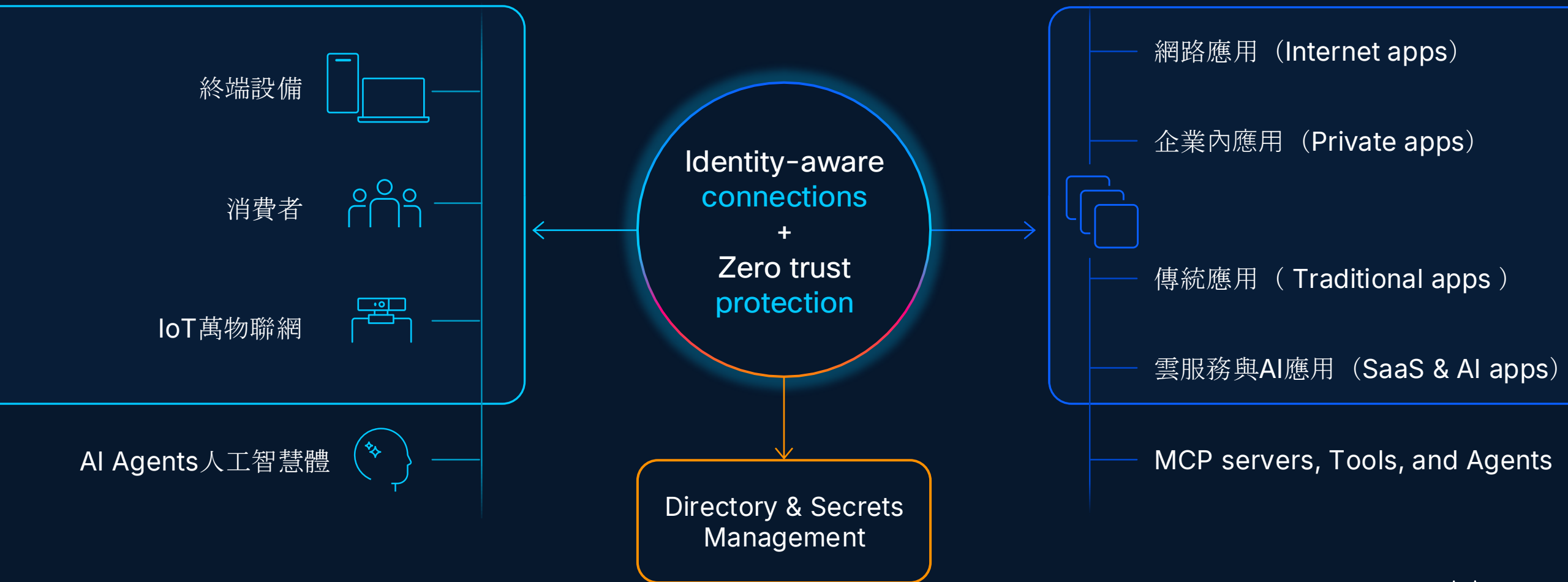


Demo

Duo Agent Directory &
Cisco Identity Intelligence for Non-Human Identities

Cisco Duo 將 Zero Trust Access 延伸至 Agentic AI

全面保護 AI Agent 的存取安全



CISCO Connect

Thank you

