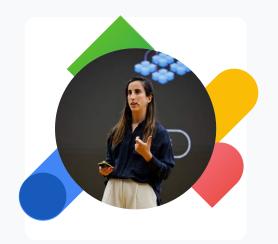
Google Cloud
Cloud Day
Lisbon

### The Connected Agent:

Mastering ADK for Advanced A2A Collaboration





Julia Hernandez

Al Specialist Google Cloud



André Vieira

Presales Sr Manager Celfocus

Google Cloud 02

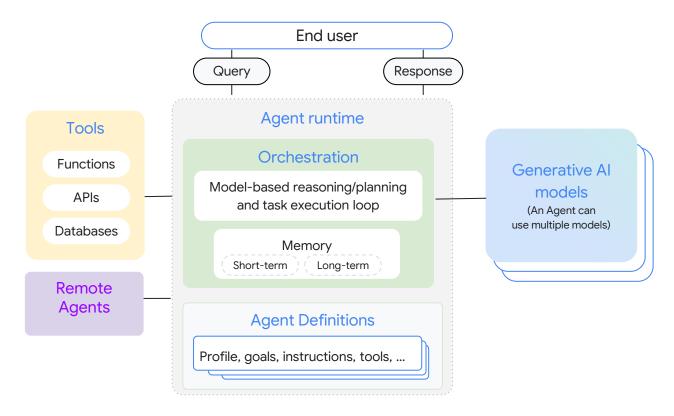


## What is an Al Agent?



Google Cloud Next

### Al Agents reason, plan, and execute tasks for users



Google Cloud Next Proprietary

# An Evolution Al Agents





LLM + Prompt



LLM + Retrieval



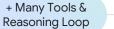




LLM + Retrieval + Actions









Multi Agent Systems



Multi Agent Systems with Remote Agents

What is next?

Google Cloud Next

# **ADK**

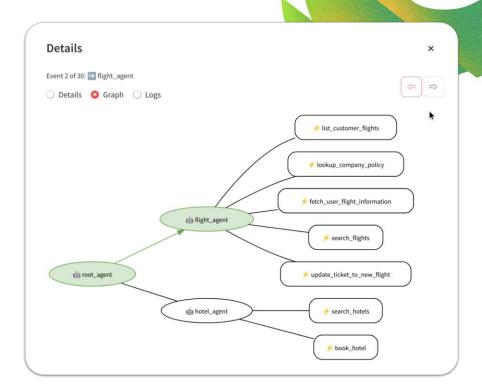
Agent Development Kit



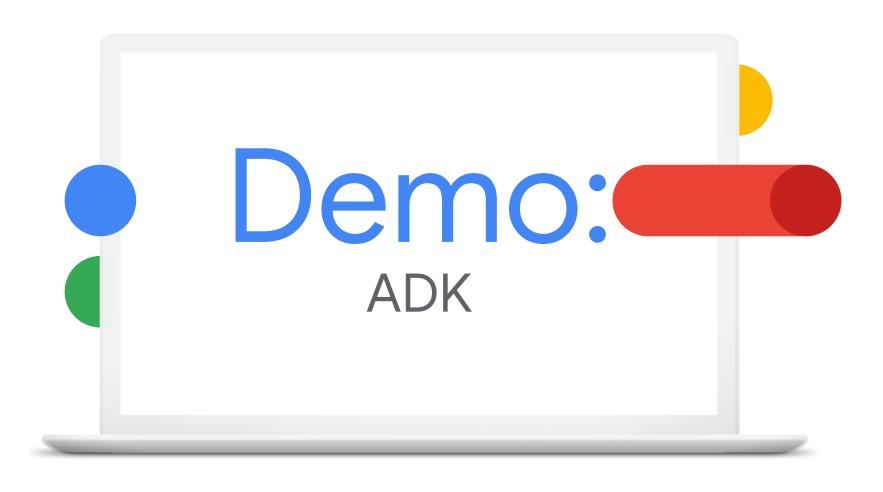
### Agent Development Kit

### Framework and SDK to build multi-agent solutions

- Define multi-agent applications orchestrating actions across many agents and tools
- Intuitive local dev UI for fast iteration; visualize agent topology and trace agent's actions
- Interleave deterministic logic with actions driven by gen Al reasoning for effortless, hybrid agents
- Built-in support for long-running sync tools / human-in-theloop
- Gemini is the default, but any generative Al model is supported, including fine-tuned models
- Support to call remote agents and thousands of existing tools via MCP
- Python and Java SDKs



Google Cloud Next Proprietary



# A2A

Agent to Agent



# We are entering a multi-agent world

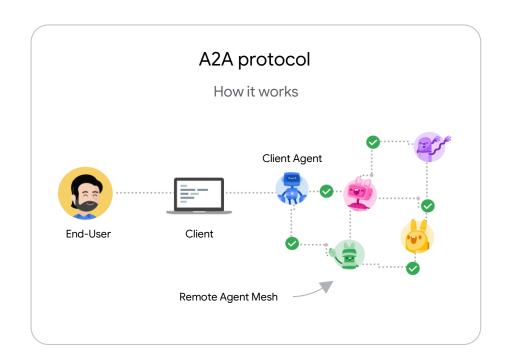
- Multiple frameworks and unique deep agents from vendors
- Blackbox Agents across boundaries cannot share tools, memory, plans, thoughts, etc.
- Difficult to represent "agent as a tool" given multimodal, unstructured, dynamic interaction with users. (OpenAPI specs or MCP not sufficient)



# A new standard for Agent-to-Agent collaboration is needed

### Introducing A2A (Agent2Agent) protocol

- Open protocol designed specifically to handle the dynamic and long-running nature of agent collaboration
- Enables agent discovery using Agent Card
- Designed with enterprise-use in mind for a secure and governed experience.
- Built on top of existing, popular standards including HTTP, SSE, JSON-RPC



### A2A compliments MCP

### Connect Agents to Tools and Other Agents using MCP and A2A

#### MCP (Model Context Protocol)

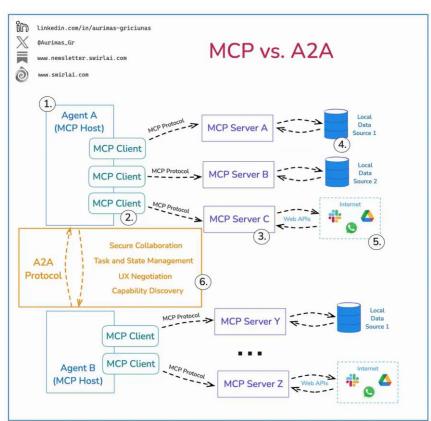
#### for tools and resources

- Connect agents to tools, APIs, and resources with structured inputs/outputs.
- ADK supports MCP tools.

### **A2A (Agent2Agent Protocol)**

#### for agent-agent collaboration

- Dynamic, multimodal communication between different agents without sharing memory, resources, and tools
- Open standard driven by community.





014

# Making Data Actionable

As a highly specialised technology company, we help clients undergo their innovation path to become Al-driven, providing professional services focused on creating business value through Analytics and Cognitive solutions.



Gartner.

REGIONAL TELECOM IT SERVICE PROVIDERS Europe Gartner.

TIER 1 CSP PARTNER for Generative AI FUTURENET

2024 AWARDS Winner - Operator Award ow ignite

FINALIST EXCELLENCE in autonoumous network



- 4 AWARDS WINNER 2024
- Innovative & Futuristic
   Application
- Application of Al & Automation
- . Beyond Telco
- · Tech for Good



TOP 20 -TELCO AI CHAMPIONS

Google Cloud 07

# The journey begins with the vision but Agentic AI is the enabler to achieve it

### **Customer Service**

From reactive tickets to intelligent resolution

Resolve faster with agents that understand intent, customer history, and real-time data.

"How can I navigate real-time data and resolve issues without switching tools?"

### Operational Efficiency

From manual workflows to smart coordination

Use agents to automate repetitive tasks, reduce tool overload, and empower teams without retraining.

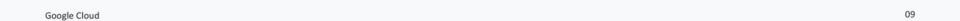
"How can I automate more without rebuilding everything?"

### **Supply Chain**

From slow decisions to fast data-driven decisions

Agents detect, diagnose, and dispatch actions using real-time insights and multimodal Al.

"How do I reduce truck rolls and improve my inventory visibility?"



### Operational Efficiency: Automating with Intent, Not Scripts

### Agentic AI replaces static workflows with goal-driven orchestration.

No predefined flows. No RPAs. All actions are triggered by user intent.

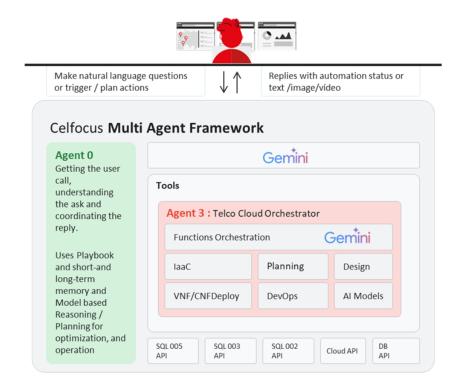
### Agent 0

Interprets the request and uses Chain of Thought reasoning to route it to the most appropriate agent.

#### Agent 3

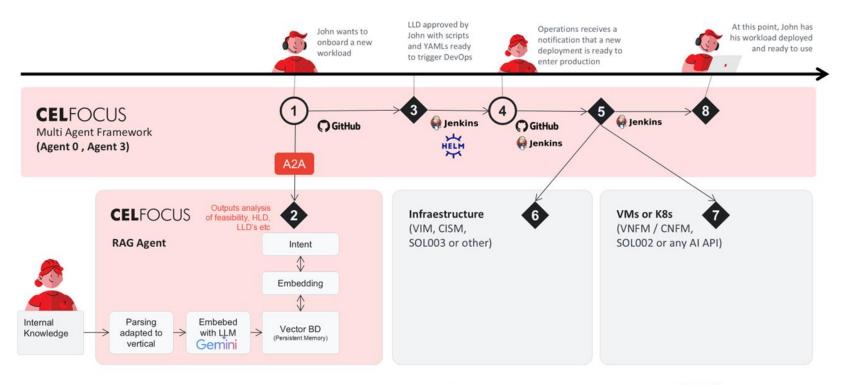
Executes actions across its domain, using declarative tools and Al-based planning.

This model elevates human operators, turning them into supervisors of intelligent systems instead of executors of repetitive logic.



Google Cloud 017

### The use case: 10 days SLA to 30 min







### The change in the operating model

Dimension	Today	Agentic Al Operating Model
Decision Execution	Engineers need to know multiple tools and take decisions	Agents act based on delegated intent and engineers became approvers.
Knowledge Sharing	Siloed, static documentation	RAG-based memory graph, shared context.  Everyone contributes to the knowledge to grow across the company and the Agentic AI is just the bridge between teams.
Automation	RPA islands, hard-coded flows	CoT agents coordinating dynamic actions. Intent based. Validated by engineers, made by Agentic AI.
Personalization and scalability	Limited by product vendors	Limited by engineers' imagination.

Google Cloud 015



03

## What is Next?

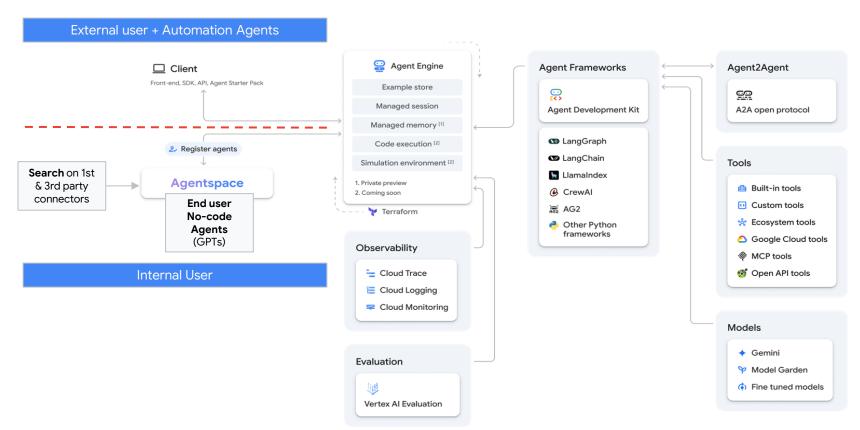
020

# AE

Agent Engine



### System Architecture of Agents in Google Cloud



### Vertex Al Agent Engine

Integrated (and independent) features to deploy and manage agents in production





#### Runtime

- Fully managed **runtime** with CRUD access to agents
- Scale up and down based on your needs
- Security and authentication
- Framework agnostic: ADK, LangGraph, CrewAl, LlamaIndex, LangChain, AG2, custom



### **Context Management**

- Sessions to manage short term memory in a conversation
- Memory Bank to store and retrieve info from sessions to personalize agent interactions



#### Quality

- Measure quality with the Vertex AI Evaluation service
- Improve agent performance with the Example Store
- Optimized with Gemini model training runs



### **Monitoring**

Understand agent behavior with open telemetry integration with

- Google Cloud Trace, Monitoring, and Logging
- 3Ps\*\* like LangSmith, Arize, and Weights & Biases



#### **Advanced Tools**

- Code execution sandbox\*\* to run agent code
- Computer use sandbox\*\* for agents
- Agent simulation\*\* to proxy user traffic to enable agent evaluation and monitoring

Google Cloud Next Proprietary

<sup>\*</sup> private preview

<sup>\*\*</sup> coming soon

### Beyond context windows, bringing memory to agents

A more strategic approach to managing past interactions by storing and retrieving relevant information as needed.

#### **Public Preview**

Agent Engine Sessions
Store, manage, and retrieve the ongoing conversation history (state) within a single session as short-term memory

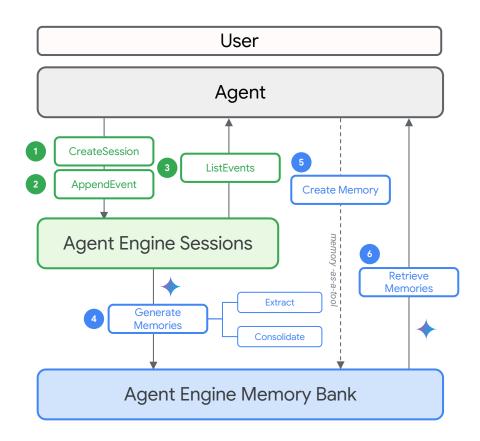
#### **Public Preview**

Agent Engine Memory Bank
Store, transform, and retrieve
memories (state), particularly across
multiple sessions as
long-term memories

Sessions help developers manage the state required to maintain conversational flow

Memory Bank enables personalized LLM agents that can remember and utilize past interactions without the overhead and performance limitations of constantly feeding entire conversation histories into the context window.

### Agent Engine Memory Bank in action



# Thank You

