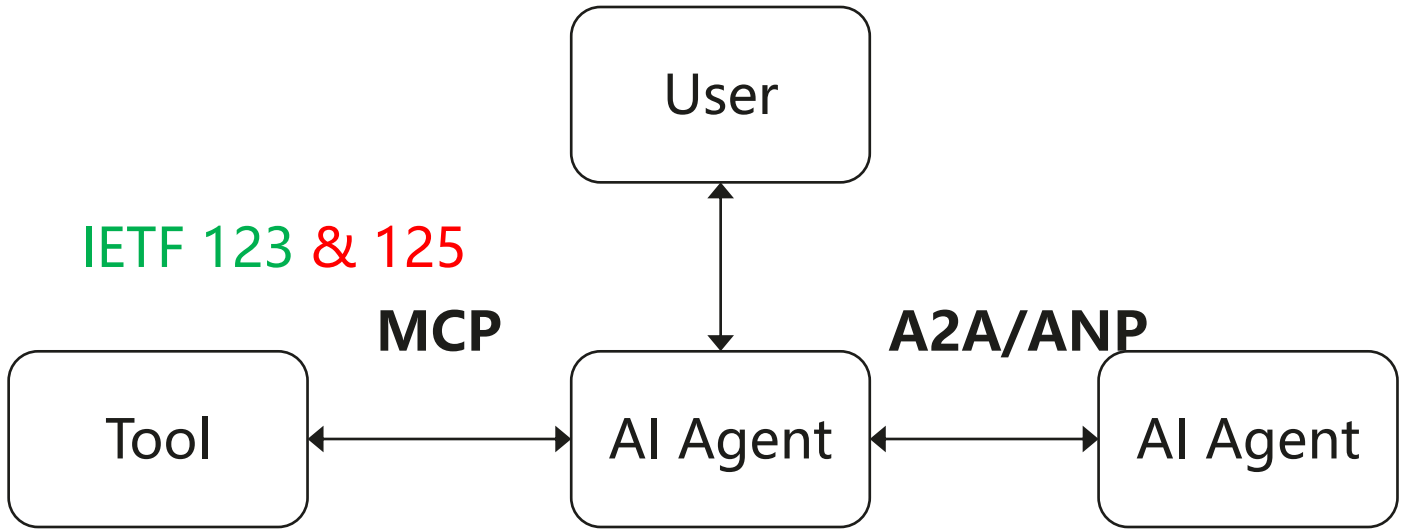
AI Agent Protocol Security

Security Aspects of Agent Protocol



Security aspects include:

- Identity
- Authentication
- Authorization
- Privacy

What We Achieved: Identity

In IETF #123, we developed IDM, which can assign DID to AI agent and tool box, and publish to VDR.
Github: https://github.com/rainie-song/agent-protocol-security/

AI agent

DID

- Identifier
- Document

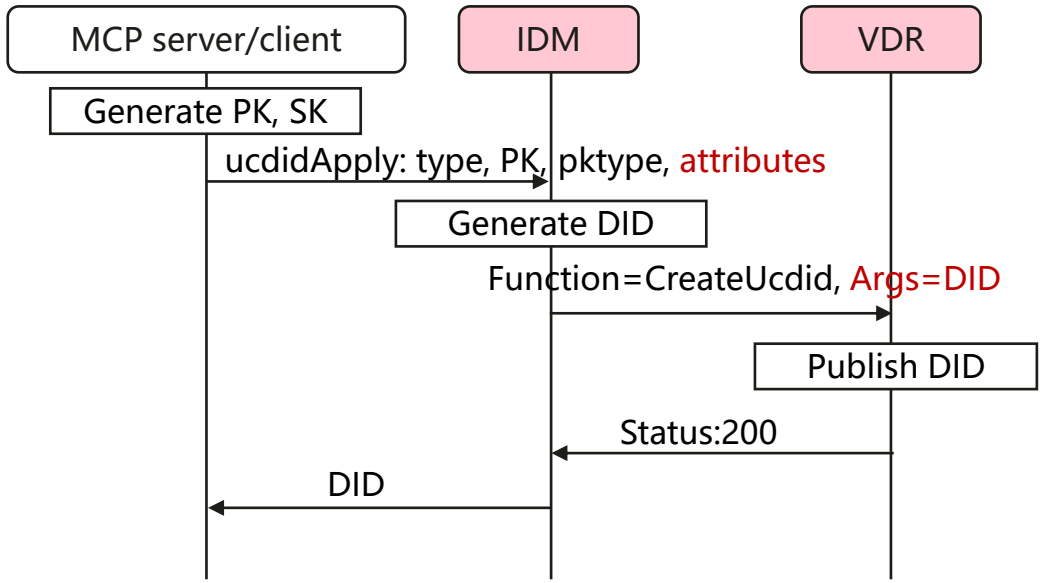
Document
- **Type:** AI agent
- **Provider**: vendor of AI agent
- **Description**: function/capabilities of AI agent (AI assistant)
- **Protocol**: (MCP/A2A)
- **Transparency**: methods to implement AI transparency (no/database/PDL)

ToolBox

DID

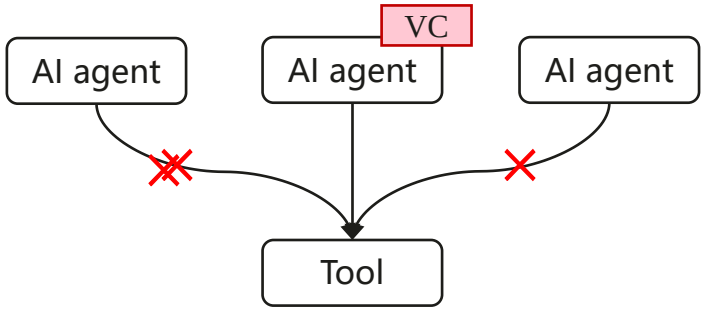
- Identifier
- Profile

Document
- **Type:** toolbox
- **Provider:** vendor of tool
- **Description:** tools names



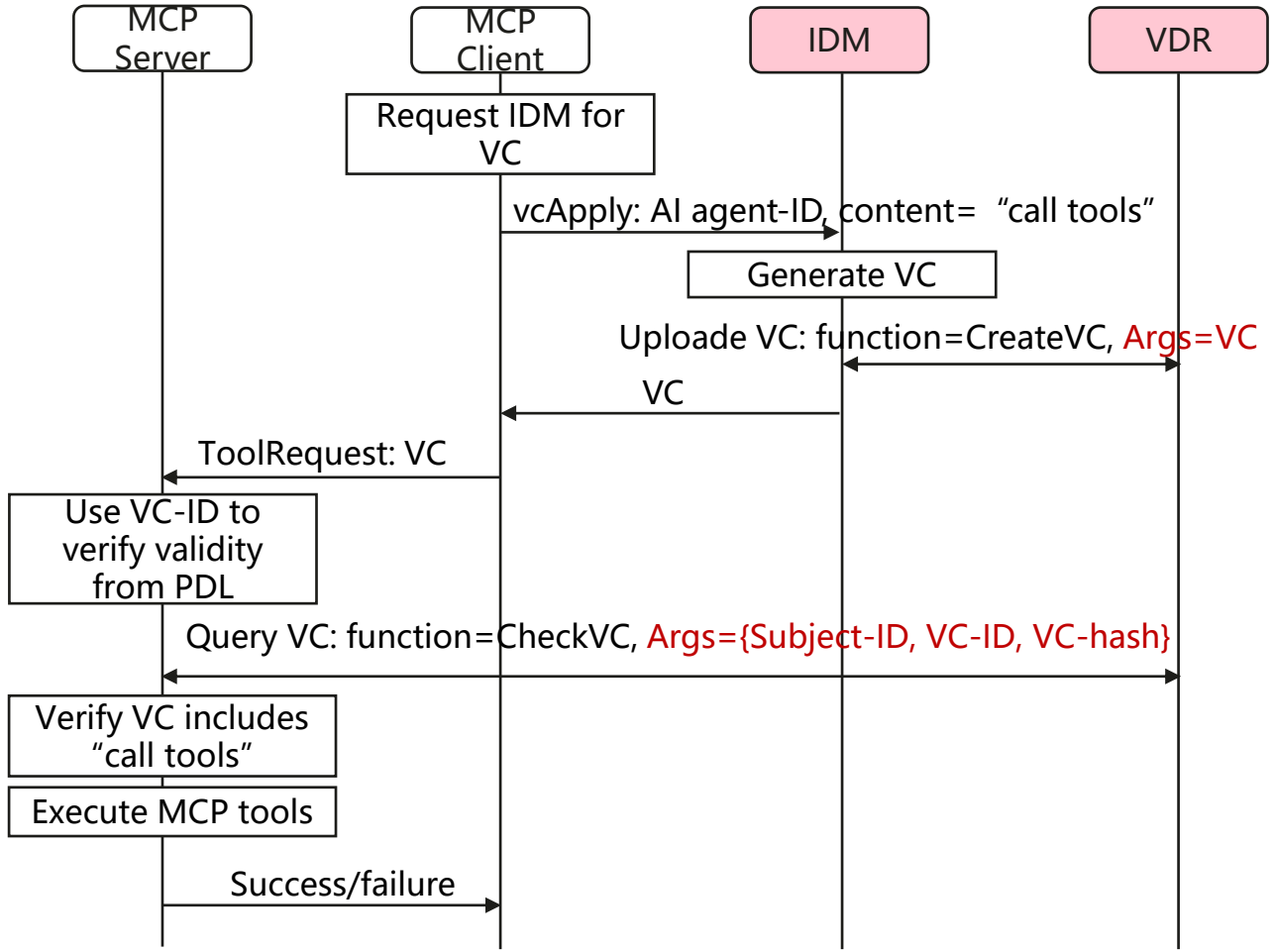
What We Achieved: Authorization

In IETF #123, we developed a simple authorization procedure based on VC.
Only AI agents with specific attributes can access tool



VC

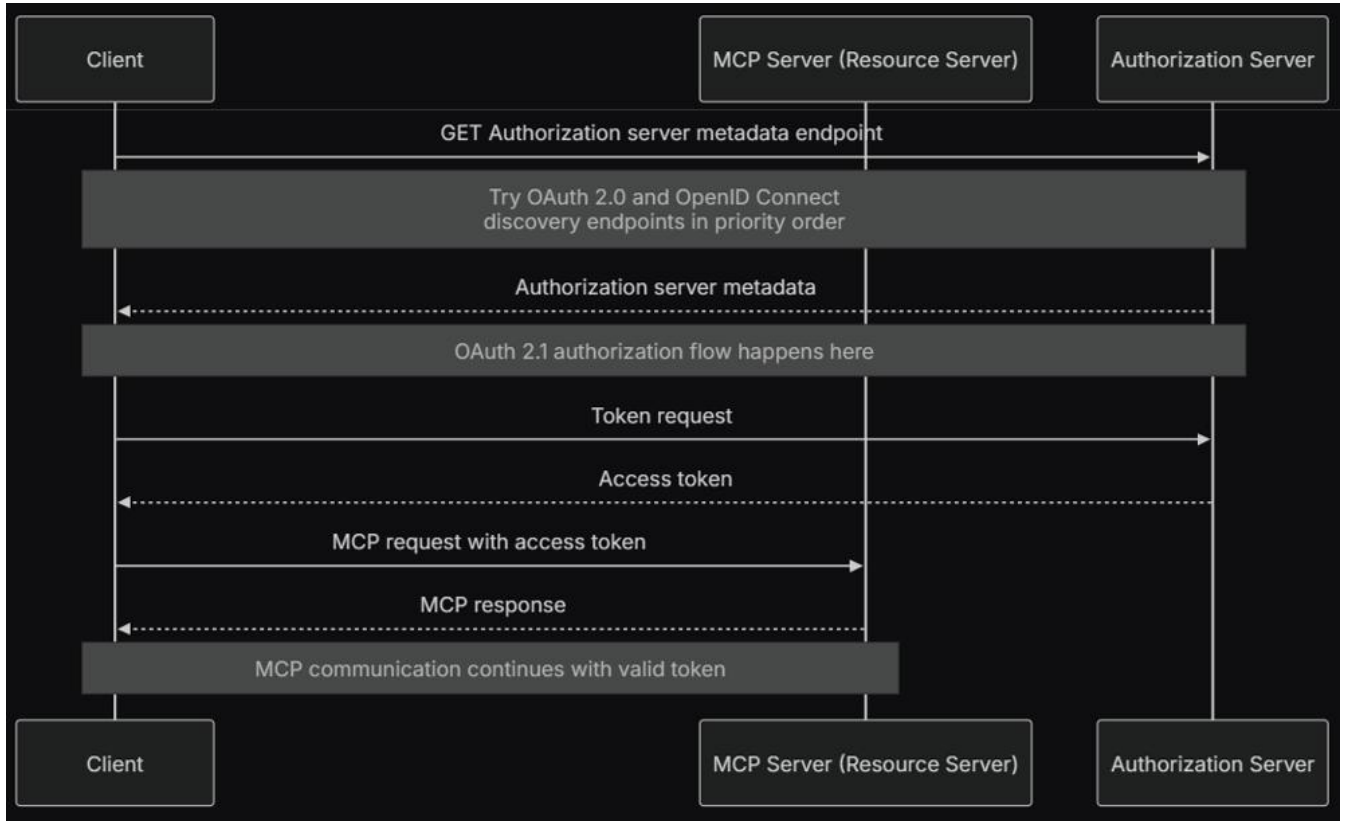
- Issuer
- Subject
  - ✓ subjectID
  - ✓ Attribute
- Proof
  - ✓ Signature



How about OAuth?

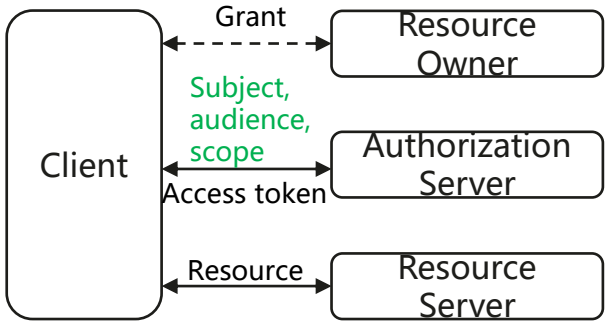
MCP Uses OAuth for Authorization

- MCP uses OAuth for secure authorization between MCP clients and MCP servers.
- OAuth acts as a secure intermediary, allowing MCP clients to access protected tools on an MCP server without directly sharing user credentials.

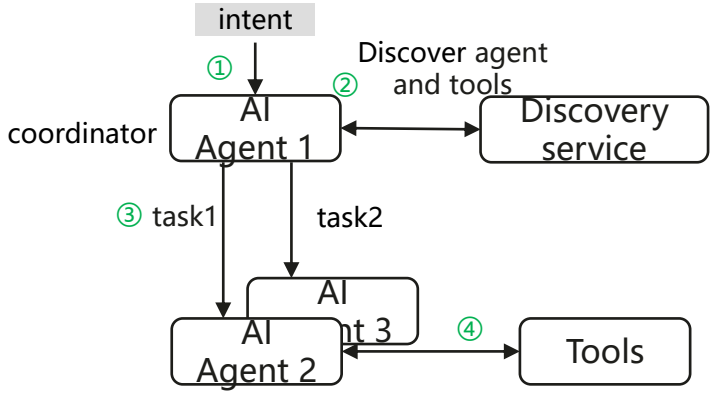


Security Requirements for Group Authorization

OAuth 2.0



Multi-agent System



- Coordinator -> **the entity that apply token**
- Task -> Selected AI agents -> **the entity that use the token**
- Task -> Tool server -> audience
- Task -> tools -> scope

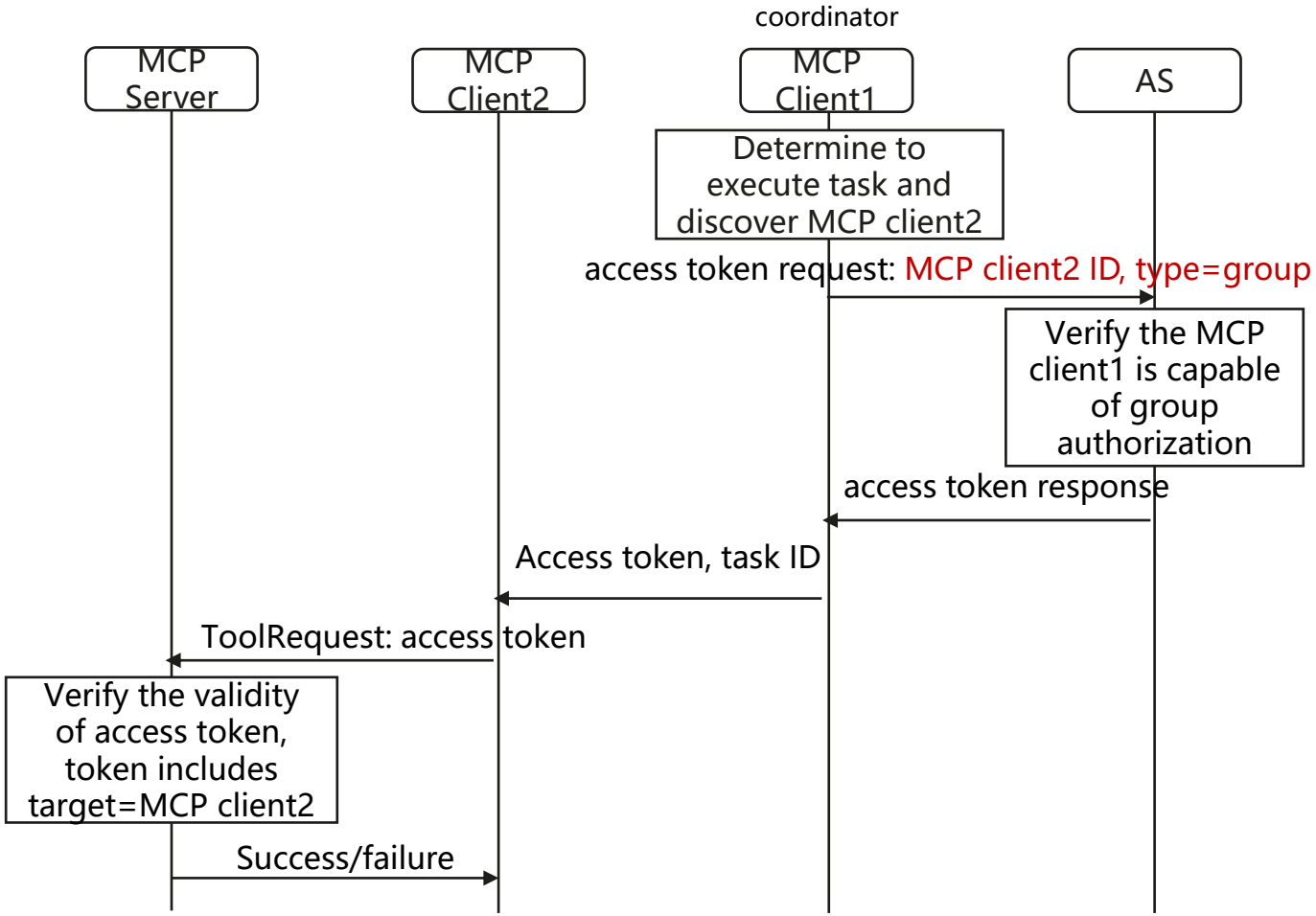
Conclusion: **The coordinator can apply for an access token for a task**

1. The coordinator receives the intent from the user, resolves the intent, and determines the tasks to be executed.
2. The coordinator uses task ID to discover and select appropriate AI agents and required tools to complete these tasks.
3. The coordinator then assigns tasks to the corresponding AI agents.
4. When executing tasks, AI agents may need to invoke various tools

What We Achieved: Authorization

This time, we developed group authorization procedure based on OAuth.

- Draft: draft-song-oauth-ai-agent-collaborate-authz-01 - OAuth2.0 Extension for Multi-AI Agent Collaboration
- Github: https://github.com/rainie-song/agent-protocol-security/



Access token

- iss: AS ID
- **sbj: MCP client1 ID**
- tar: MCP client2 ID
- aud: MCP server ID

...

option1

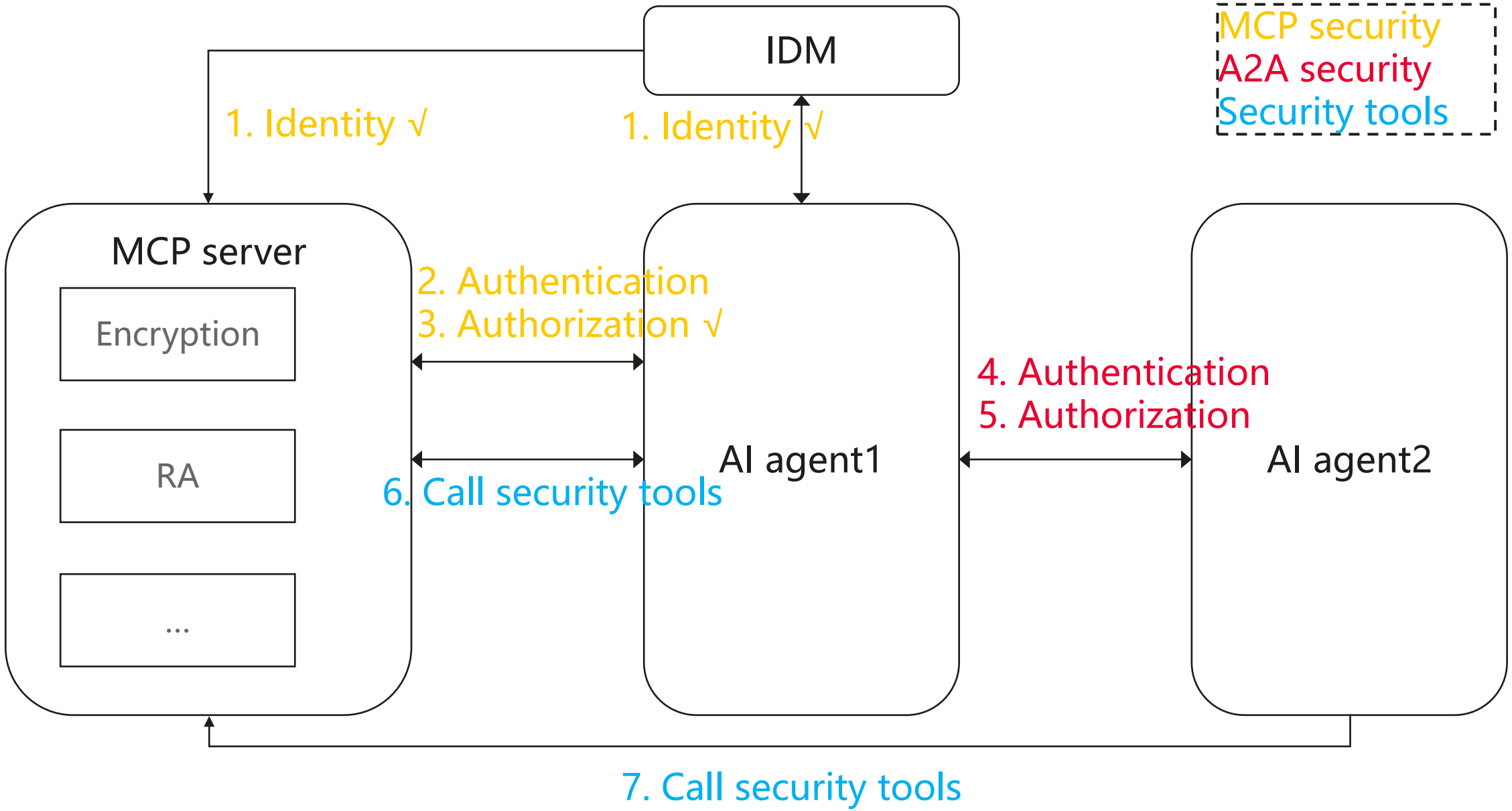
Access token

- iss: AS ID
- **sbj: MCP client2 ID**
- app: MCP client1 ID
- aud: MCP server ID

...

option2

What Next


IDM
MCP security
A2A security
Security tools
1. Identity √
1. Identity √
MCP server
Encryption
RA
...
2. Authentication
3. Authorization √
6. Call security tools
AI agent1
4. Authentication
5. Authorization
AI agent2
7. Call security tools

Thank you.

Yurong Song (songyurong1@huawei.com)
Chuan Ann Foo (foo.chuan.ann1@h-partners.com)