Economies of Scale for MCP and Agents

Why You Need an Identity Broker

Magnus Jungsbluth · Jan Brennenstuhl · Zalando SE

AGNTCon + MCPCon Europe 2026 · Amsterdam

Connecting a tool is only the start.

Identity work repeats across integrations

MCP integration A	MCP integration B	MCP integration C
Discover authorization	Discover authorization	Discover authorization
Validate token audience	Validate token audience	Validate token audience
Handle scopes and errors	Handle scopes and errors	Handle scopes and errors

Repeated integration work

Resource servers can use a separate authorization server.

The problem is repeated implementation, not OAuth itself.

We learned this while scaling microservices

Before

Controls in each service.

Authentication, transport security, retries, and rate limits.
Product engineering teams repeated the work.
Implementations drifted.

After

Controls in the platform.

Gateways and identity services provided common controls.
Product engineering teams focused on domain logic.

Product teams should not rebuild a platform

Agent A	Agent B	Agent C
Runtime	Runtime	Runtime
Session state	Session state	Session state
Tool routing	Tool routing	Tool routing
Registration	Registration	Registration

Shared concerns, rebuilt by each team

The same product engineering effort can repeat across agent projects.

Shift down the shared work

Product team A	Product team B	Product team C
Agent behavior Domain tools	Agent behavior Domain tools	Agent behavior Domain tools

Product ownership

Shared platform

Agent Runtime	MCP Gateways	Authn & Authz	LLM Gateway	Observability
Run agents	Reach tools	Delegated access	Consume LLMs	Trace actions

Shift down: move common implementation into the platform.

Reuse changes the cost of the next integration

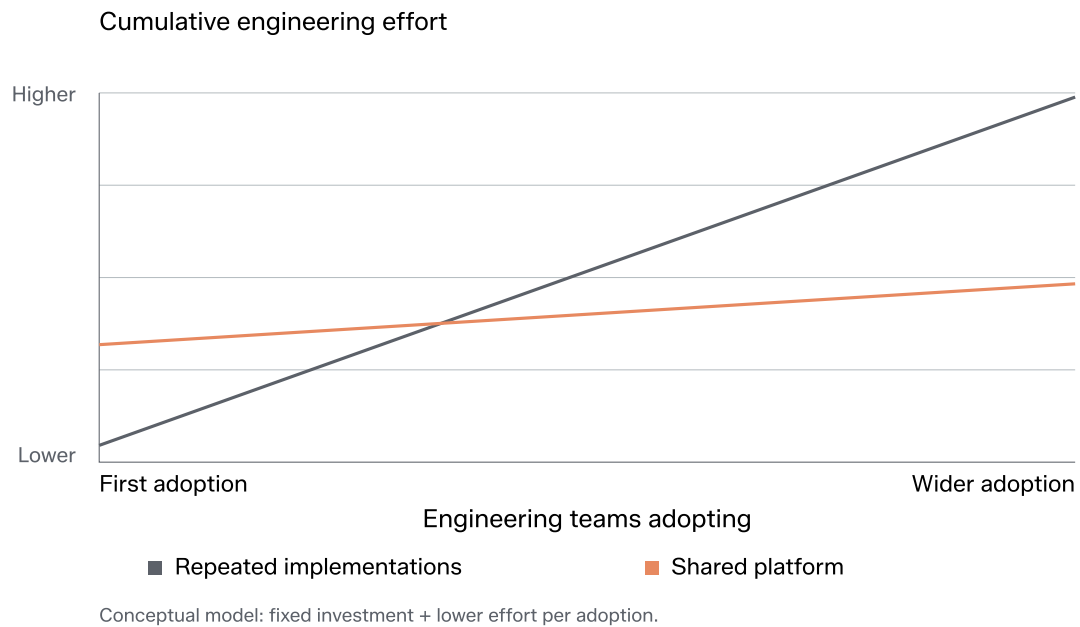
Economies of scale

The platform needs an initial investment + maintenance.

Each new agent then reuses security properties + infrastructure.

Tool providers publish via the platform.

The benefit is less repeated engineering and fleet management.

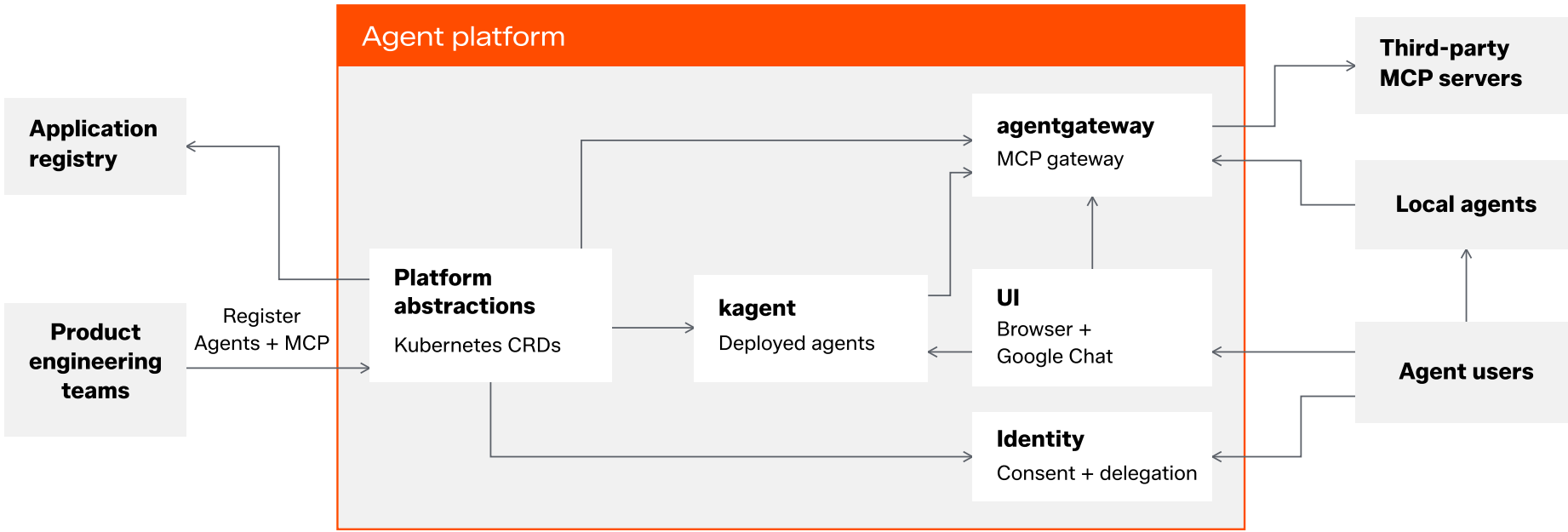


Conceptual model: fixed investment + lower effort per adoption.

Illustrative model, not measured cost data. Integration and operating work remain.

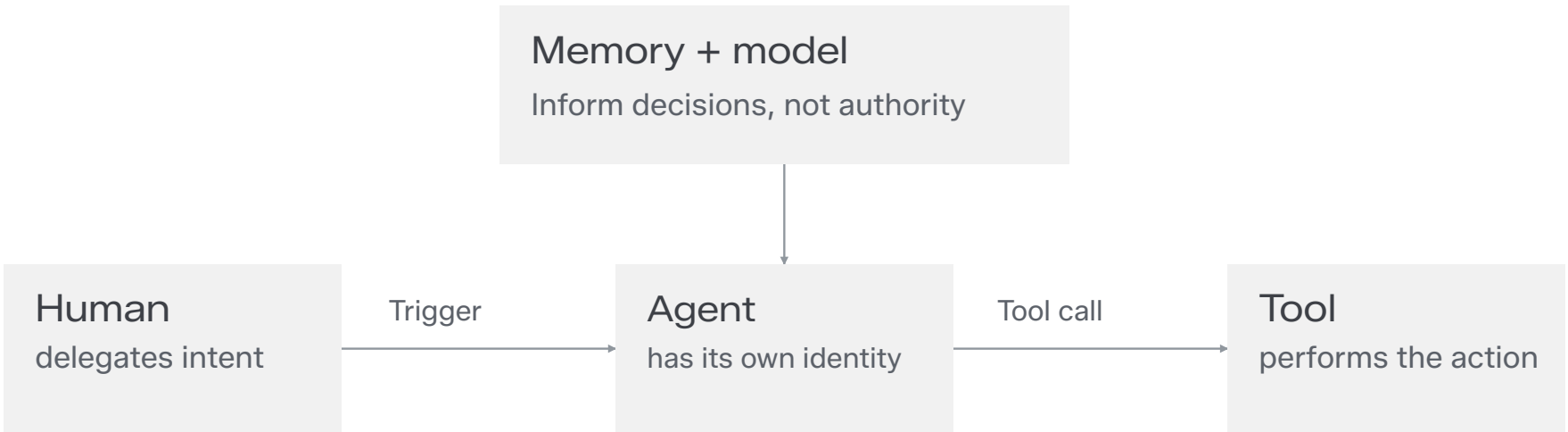
Central infrastructure. Local autonomy in business decisions.

A shared platform connects the pieces



A reusable control plane unifies registration, runtime, routing, and delegated access.

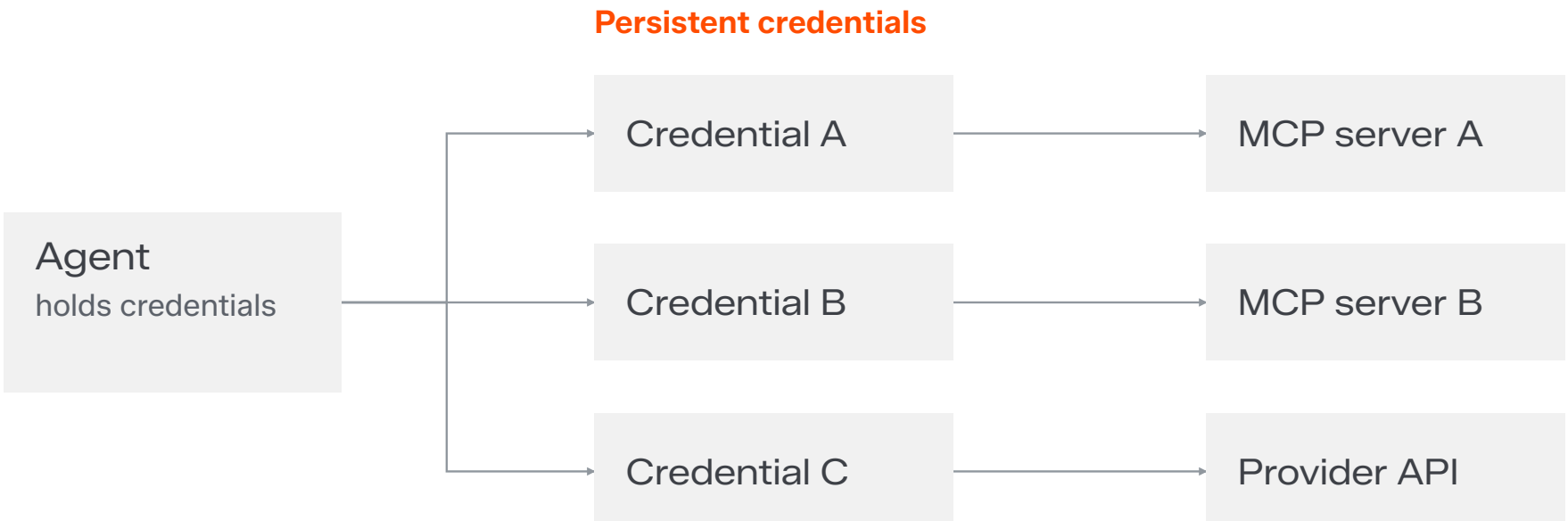
An agent is not a person or a fixed script



Attribute the action to both the user and the agent

Autonomy does not remove the need for attribution.

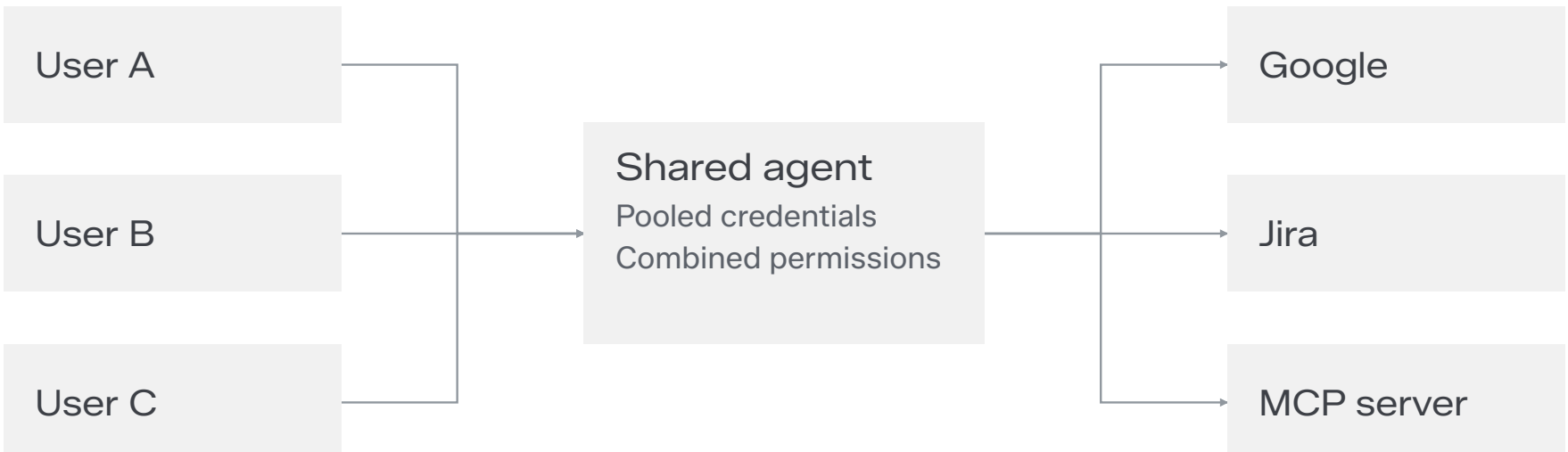
A robot account is a tempting shortcut



A separate account does not solve delegated access.

Shared credentials pool users' permissions

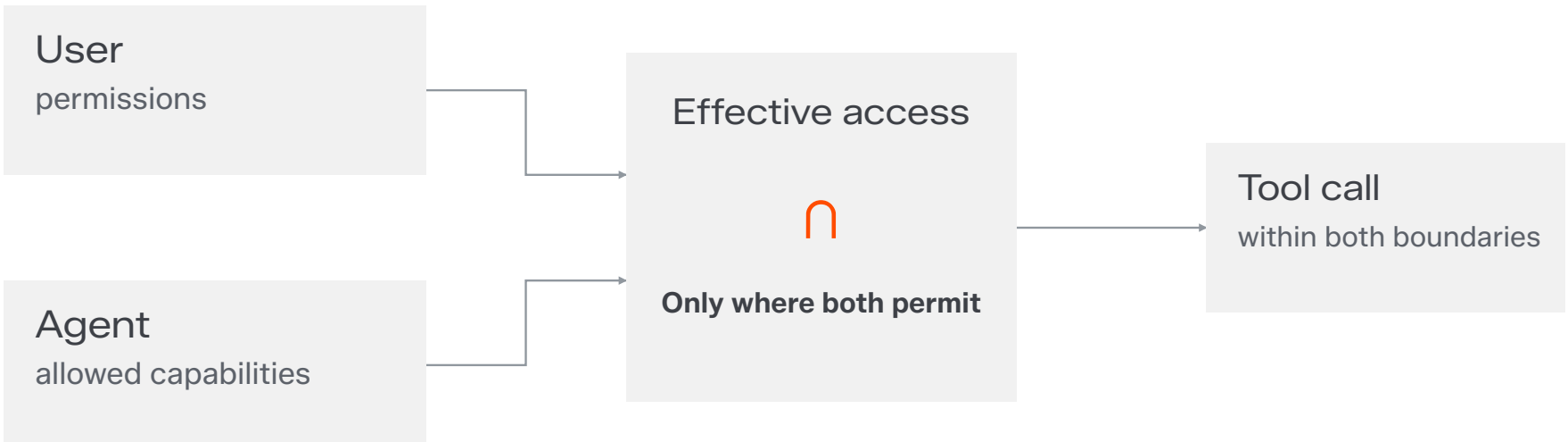
Anti-pattern: pooled authority



Not scoped to the current user

The current user should not gain the combined permissions of other users.

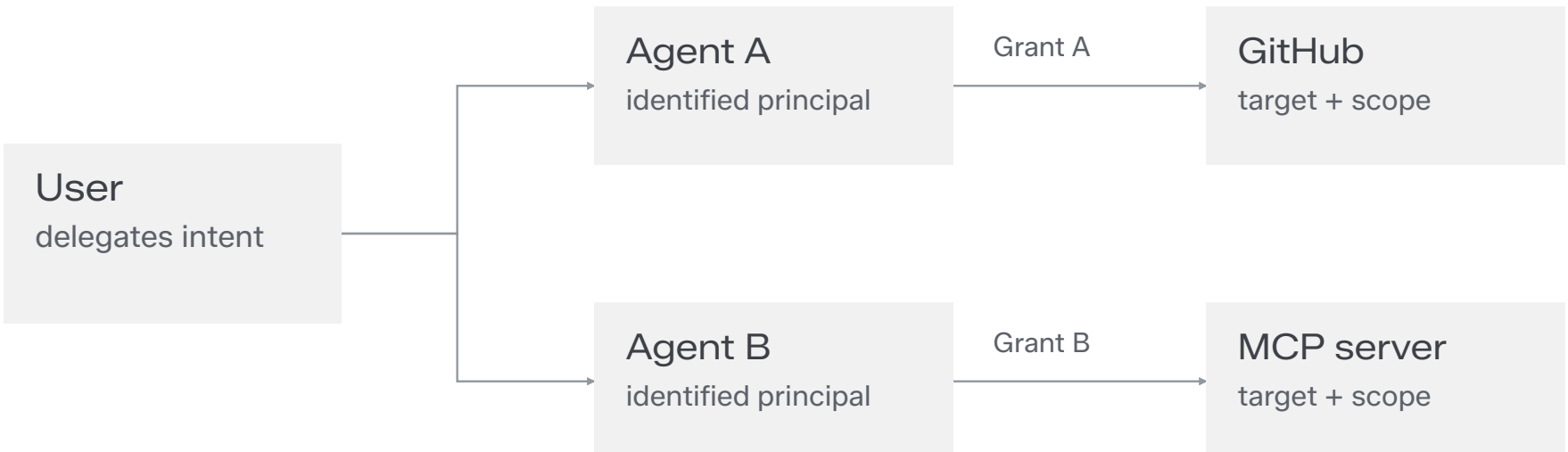
Identify the agent. Delegate the user's access.



Keep user identity and agent identity in the authorization context.

User permission AND agent permission. A live delegation is also required.

Delegation is a graph, not one credential



User + agent + target + scope

User + agent + target + scope. Keep every delegation attributable.

Agentic Identity Broker

A shared service for delegated access and to capture user consent.
Built and used in the Zalando Agent Platform.
Open source for agents and MCP integrations.

Keep the IdP. Add delegated access.

Identity Provider

Who is the user?

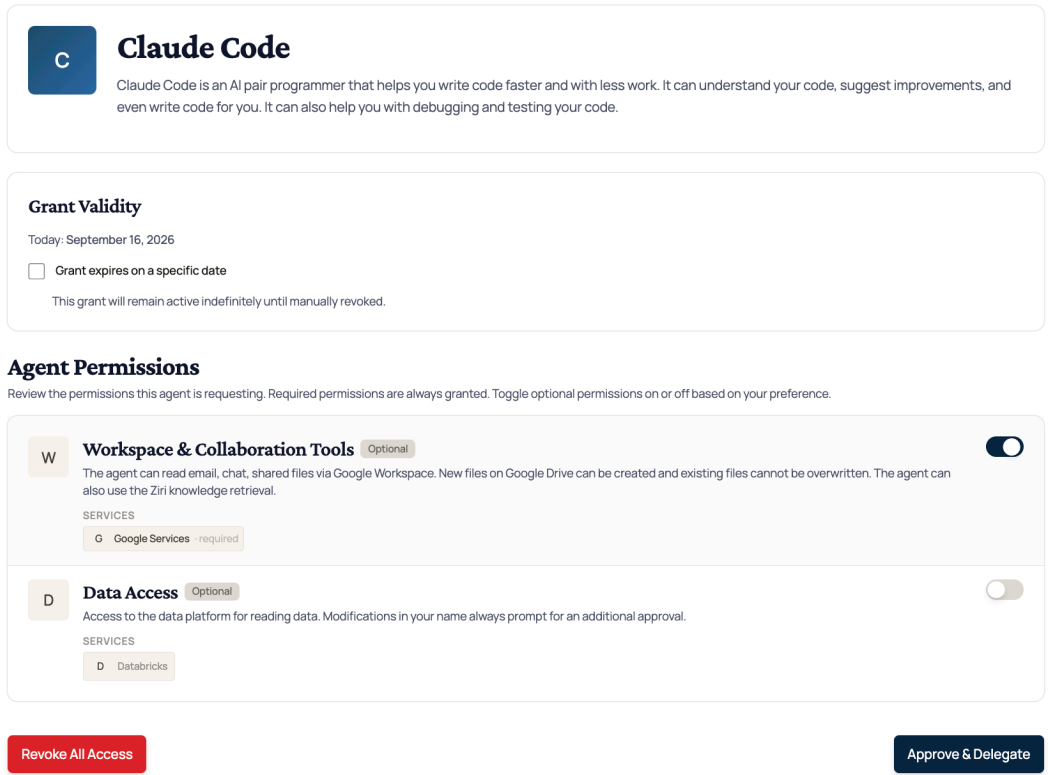
Keep the existing login system. A trusted proxy authenticates the user and forwards that identity to the broker.

Identity Broker

What is delegated?

Resolve the user, agent, service, and live grant. Manage provider sessions, consent, and scoped access centrally.

Consent Screen



Users choose the agent permissions and lifetime of a delegated grant.

Consent becomes a bounded grant

Choose the agent

Know who is asking

Show validated agent metadata. Identify the product and operator behind the request.

Choose capabilities

Understand the requested access

Use a permission set in business language, not an unexplained provider scope string.

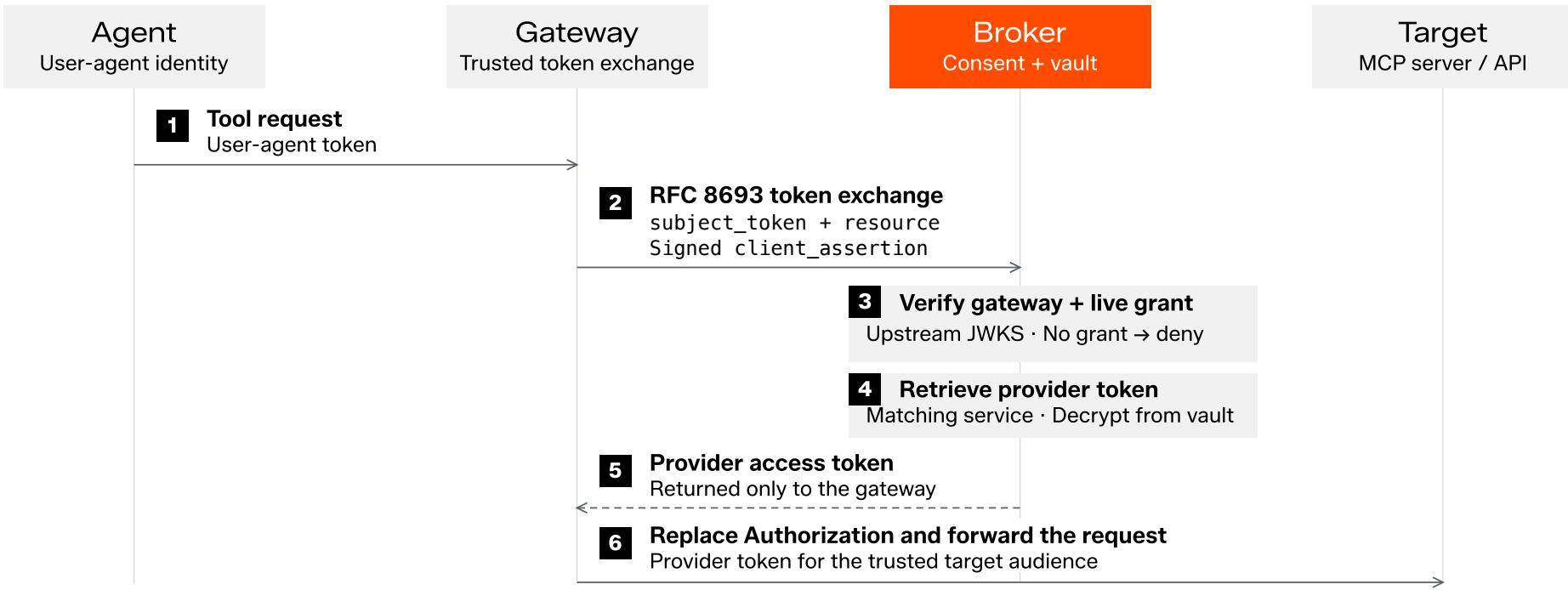
Approve the grant

Scope and duration

Record the user, agent, and permissions. Apply an expiry where appropriate. Keep revocation available.

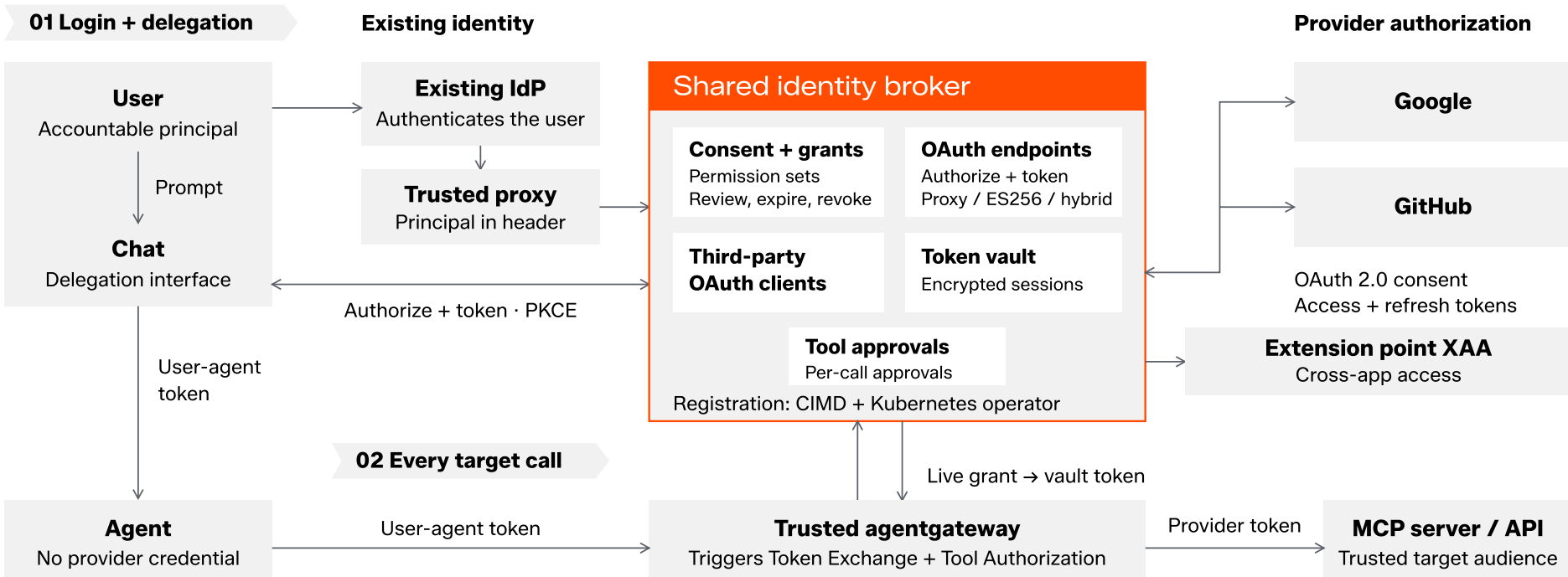
Consent management remains available outside the initial login/ consent flow. Users can review and revoke grants.

The agent never receives a provider token



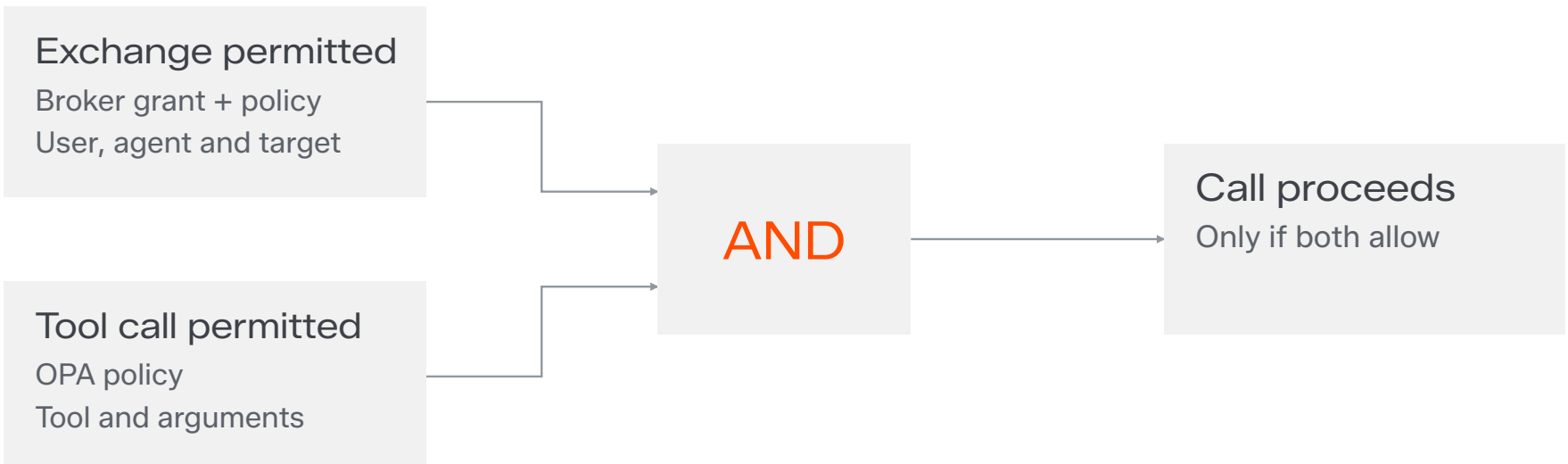
RFC 8693. A missing or revoked grant denies the exchange.

The full broker architecture



Provider credentials stay between vault, trusted gateway and target.

Beyond the token exchange: Authorization



Policy: allow, deny or challenge + risk indication

Available today: allow or deny. Both checks must pass.

Beyond consent: central tool approvals

In Progress

Tool approvals

Apply OPA policies to tool invocations. MCP elicitation with the approval URL. Resume only after approval in the broker.

Roadmap

Strong affirmation

Use CIBA backchannel authorization. Bind an authenticator approval to the exact high-risk transaction.

Outlook

Intent-based access

Carry the user's intended task as policy input. Evaluate tool calls for actions outside that intent.

Audit Trails vs User Transparency

Central

Monitor for central risk

More traditional security focused monitoring with the twist that sessions tie individual logs.

User Facing

Build Trust

What users ask for is to see what agents did on teir behalf beyond their chat sessions. Audit logs must be also available per agent-user pair and in user friendly language.

Core of the Zalando Agent Platform

Runtime

kagent

Kubernetes-native agent execution and session control.

Gateway

agentgateway

Shared MCP and model routing. ExtProc connects request enforcement to the identity broker.

Identity

Agentic Identity Broker

Delegation, provider sessions, token exchange, and consent. Reused by each integrated agent and tool.

Delegate access.
Apply Authorization.

Agentic Identity Broker

Delegated access for agents and MCP.

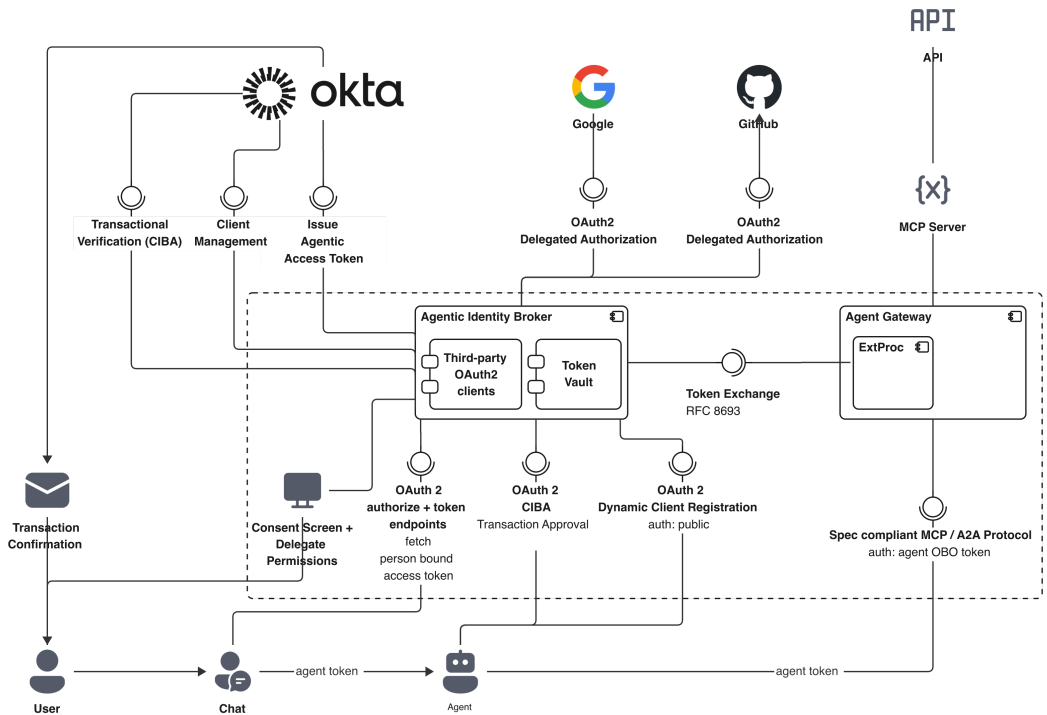
Keep your identity provider and existing IAM investment. Extract authentication & authorization concerns from agents and MCP Servers.

Magnus Jungsbluth · Jan Brennenstuhl
Zalando SE

Read the documentation, explore the implementation, and contribute:

agenticidentitybroker.dev

Appendix: the original architecture



Historical design reference. Current registration uses CIMD instead of DCR. CIBA remains outlook.