

What It Took To Build Our Own MCP Gateway

Nick Pordash — Sr. MTS, Software Engineer

Vijay Samuel — Principal MTS, Architect

© 2026 eBay. All rights reserved. Confidential and proprietary.

#WeAreeBay!

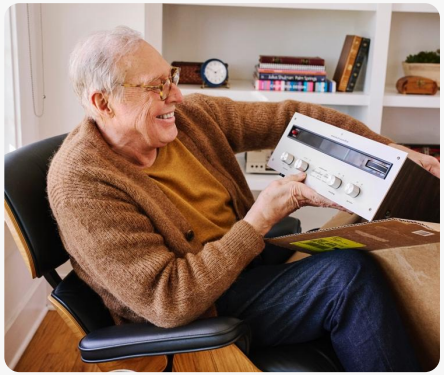
© 2026 eBay. All rights reserved. Confidential and proprietary.

eBay Today!



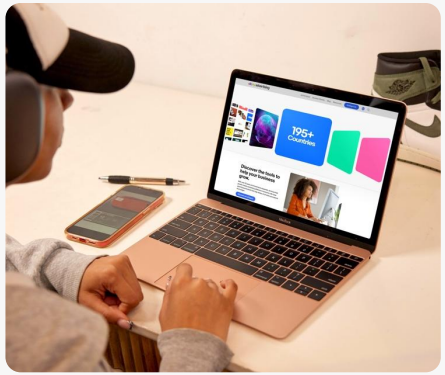
2.8B

Live Listings



147M

Total Active Buyers Worldwide



190+

Markets



$79.6B

GMV

**** All data approximate as of Q3 2026, except for GMV which shows full year 2025**

© 2026 eBay. All rights reserved. Confidential and proprietary.

01

One Protocol, A Dozen Ways To Log In...

© 2026 eBay. All rights reserved. Confidential and proprietary.

Every Server Wanted Its Own Key

- **One server, one config entry**, per person, per client, per laptop
- Long-lived **personal tokens** pasted into JSON files
- Public repos: **24,008** secrets in MCP config files, **2,117** still valid
- No central place to see, rotate or revoke any of it

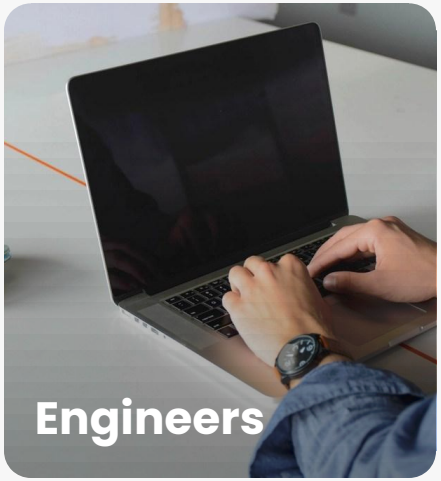
The easiest path was also the least safe one.



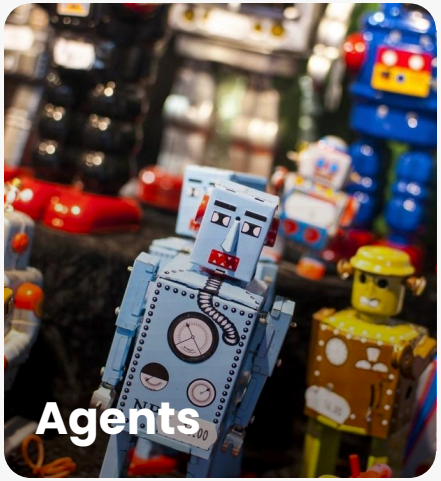


© 2026 eBay. All rights reserved. Confidential and proprietary.

Who's Calling, And Who Has To Trust Them?



Company SSO, from IDEs and chat clients



Workload identity, no human at the keyboard



Trust our identity platform, want *their own* audience and roles



GitHub, Google, Slack, Atlassian… trust only their own OAuth

Four kinds of party, and no one credential works for all of them.

© 2026 eBay. All rights reserved. Confidential and proprietary.

Where Did The User Go?

Shortcut 1 · shared identity



Shortcut 2 · forward the token



Either the backend can't tell who you are, or every backend can be you.

© 2026 eBay. All rights reserved. Confidential and proprietary.

Hundreds Of Tools, One Context Window

- **~100 servers × dozens of tools** = thousands of tool definitions
- Clients cap tool counts: one popular IDE started at **40**
- Tool names capped at **~64 characters**, so prefixed names get rejected
- More tools, worse picks: accuracy slips from **25–50** tools

Aggregating naively breaks the client before it helps anyone.



© 2026 eBay. All rights reserved. Confidential and proprietary.

We're Not The Only Ones...



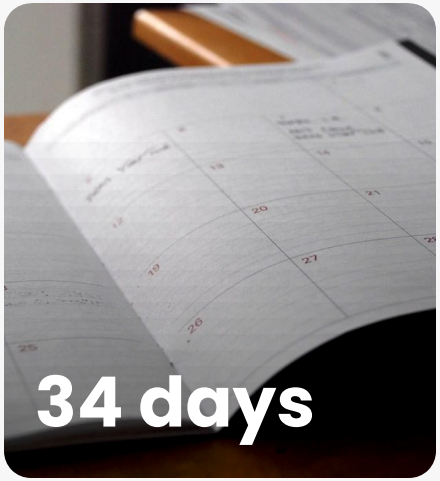
Of software leaders name security as the top barrier to MCP adoption



Of ~5,200 open-source MCP servers rely on long-lived keys or tokens



Found a shadow or misconfigured MCP server in the past year



One SaaS MCP server could expose data across tenants before the bug was found

Security, not capability, is what's slowing MCP down.

Stacklok State of MCP 2026 · Astrix Security, 5,205 servers · Salt Security 2H 2026 · public incident report, 2025.

© 2026 eBay. All rights reserved. Confidential and proprietary.

So Why Build Our Own?

© 2026 eBay. All rights reserved. Confidential and proprietary.

What We Couldn't Buy

- **Not the hard part:** routing and inbound login are well covered by existing gateways
- **Our identity:** our SSO, our workload identity, our network boundaries
- **The real user at every backend:** token exchange wired into our identity providers and backends, in a token meant only for it
- **SaaS consent stored server-side**, in our own secrets store
- **Onboarding as data**, owned by the teams who own the servers

The protocol is the easy part. Identity is yours to integrate.





© 2026 eBay. All rights reserved. Confidential and proprietary.

02

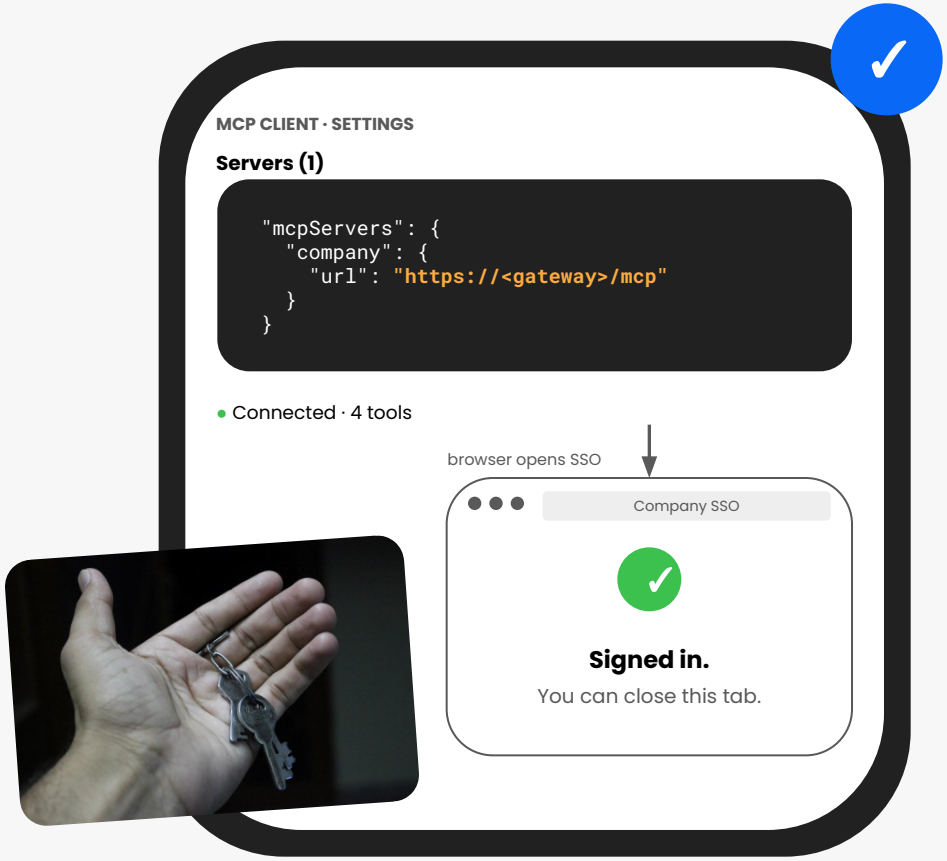
One URL. Log In Once!

© 2026 eBay. All rights reserved. Confidential and proprietary.

Add One URL, Get Everything

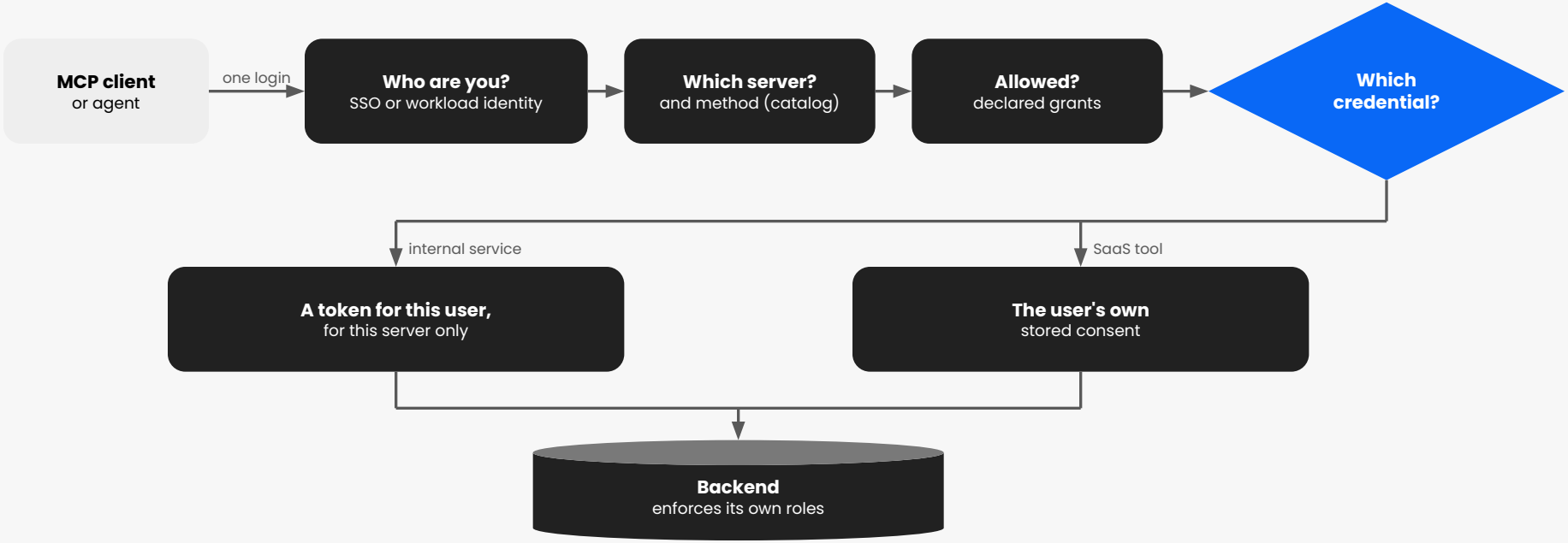
- **One config entry** for about 100 servers
- **Add the URL:** the browser opens SSO, and you're in
- **No tokens on disk.** Nothing to paste, nothing to rotate
- **Connect a SaaS tool once;** it follows you to every client and machine
- **Errors say what to do:** "Connect GitHub at …"

The best login is the one nobody has to configure.



© 2026 eBay. All rights reserved. Confidential and proprietary.

What Happens On Every Call

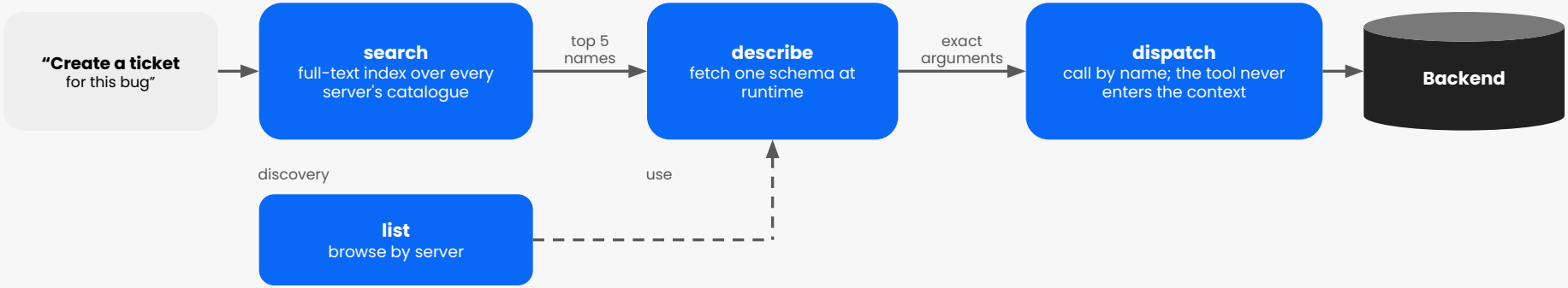


Separately from calls, background discovery lists each server's tools using the gateway's own identity, never on a user's behalf.

Log in once at the edge. Get the right credential at every hop.

© 2026 eBay. All rights reserved. Confidential and proprietary.

Four Tools Instead Of Thousands



- **Works for every kind:** tools, prompts, resources, resource templates and skills
- **Policy-aware:** search and describe only show what you're allowed to call
- **New servers are searchable** on the next refresh, with no client restart

Discovery is a tool, not a preload.

© 2026 eBay. All rights reserved. Confidential and proprietary.

03

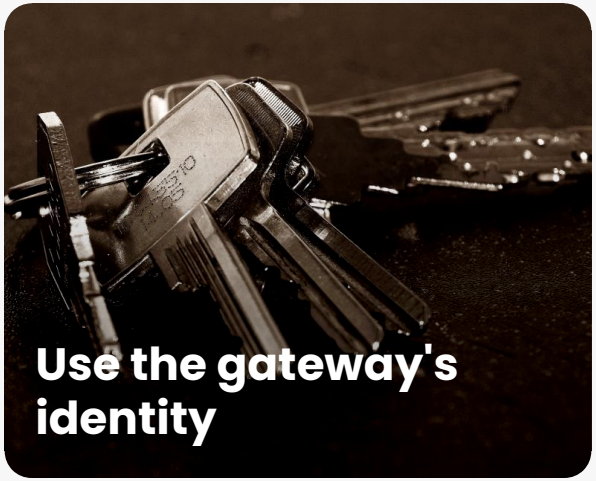
Identity Is The Product...

© 2026 eBay. All rights reserved. Confidential and proprietary.

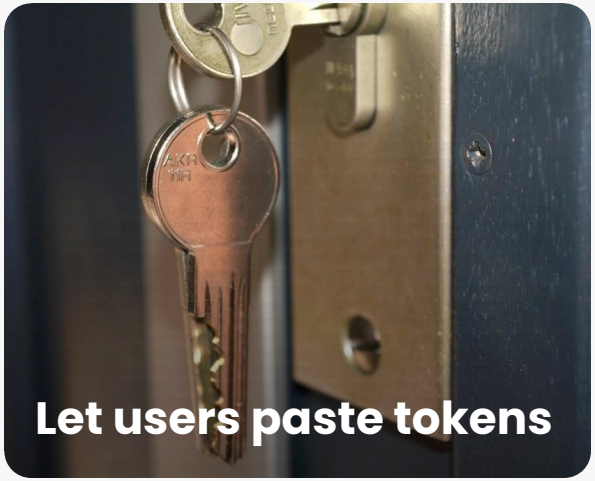
Three Shortcuts We Didn't Take



Wrong audience, replayable anywhere. The MCP spec says **MUST NOT**



Everyone gets the gateway's rights. Every audit log says "gateway"

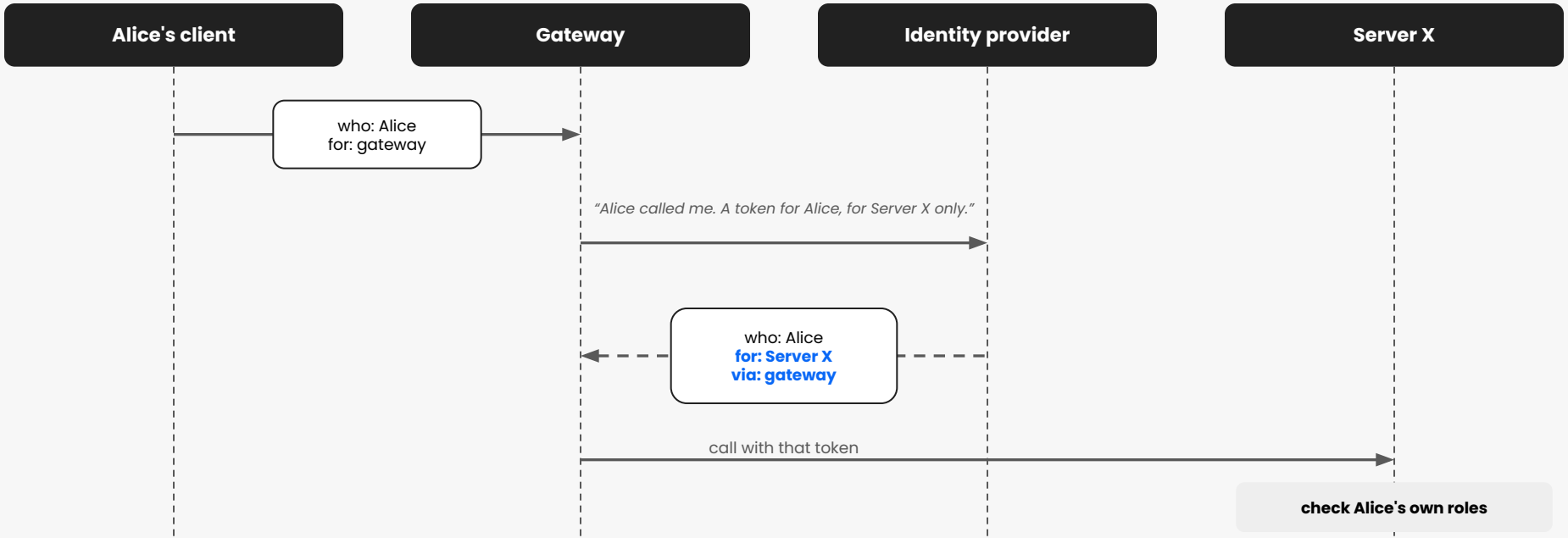


Secrets on laptops. No central revoke. No idea who has what

Each one works in a demo and fails in an audit.

© 2026 eBay. All rights reserved. Confidential and proprietary.

Same User, Right Audience



45 of 49 production servers get a per-user exchanged token today.

The backend sees Alice, with Alice's rights, in a token meant only for it.

© 2026 eBay. All rights reserved. Confidential and proprietary.

SaaS: Consent Once, Locked Twice

- SaaS providers don't trust our identity provider, so there's nothing to exchange
- **Consent once** per SaaS provider, in the browser, with the provider's own login
- **Stored server-side:** encrypted by our key service, inside our secrets store
- **Refreshed by the gateway;** disconnect in one place
- **15** SaaS providers connected today

A SaaS token never has to live on a laptop again.



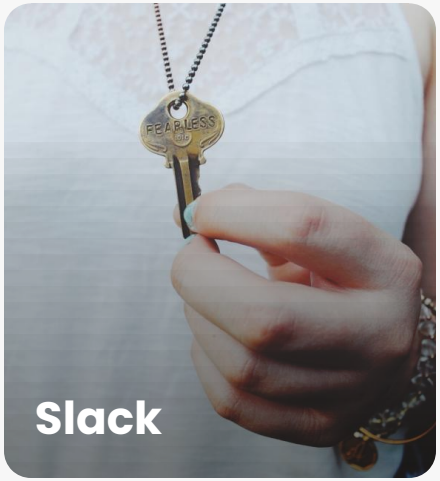


© 2026 eBay. All rights reserved. Confidential and proprietary.

Every Product Speaks Its Own OAuth



Refresh token only on *first* consent, and only if you ask for offline access



User permissions go in a separate `user_scope` parameter



Needs an `audience`; every refresh replaces the refresh token

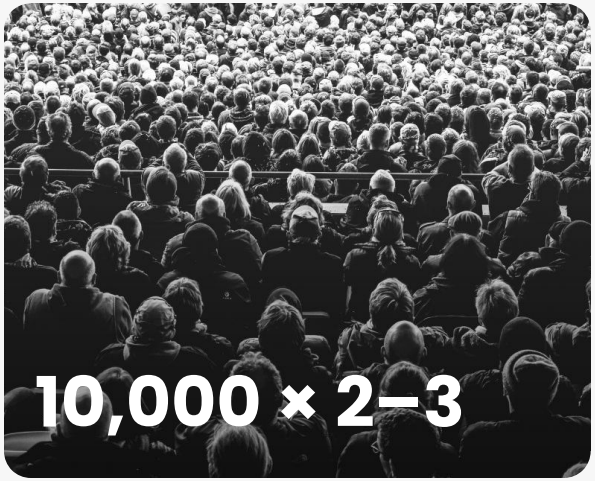


Plain-form response unless you ask for JSON; refresh tokens not guaranteed

One generic provider. The quirks live in configuration, not code.

© 2026 eBay. All rights reserved. Confidential and proprietary.

The Registration Flood



Developers, times the MCP clients each one uses



SSO client registrations if every client registers itself



Client our SSO sees, because the gateway handles registration

Absorb the churn at the gateway, so your identity provider doesn't have to.

The MCP spec now treats dynamic registration as a fallback and recommends Client ID Metadata Documents.

© 2026 eBay. All rights reserved. Confidential and proprietary.

Agents Get Their Own Badge

- **One registry record per agent:** identity, owner, purpose and tool grants, declared together
- **Grants say which tools an agent may call,** by name, pattern or hint: approved by our identity platform, enforced by the gateway
- **Not registered and ready? It doesn't run**
- **In production:** workload identity only. No passwords, no stored secrets
- **For SaaS:** the agent has its own account, and the gateway holds its credential

Humans and agents share one door, but never one identity.



© 2026 eBay. All rights reserved. Confidential and proprietary.

04

Onboarding Is A Pull Request!

© 2026 eBay. All rights reserved. Confidential and proprietary.

Configuration Is Data

- Servers, SaaS providers and access grants live in **one catalog repo**
- **Add a server = open a pull request.** Live on the next refresh
- **No gateway release**, no restart, no gateway-team code review
- **Dozens** of onboarding PRs a month, often several servers at a time
- No MCP server yet? **Point us at an OpenAPI spec**

The platform team can't be the bottleneck for every new server.





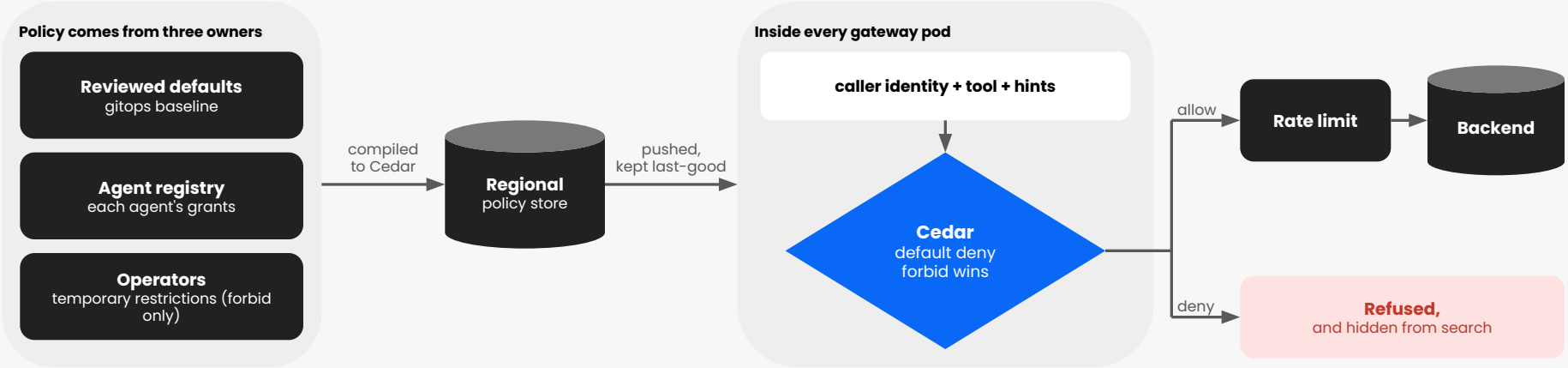
© 2026 eBay. All rights reserved. Confidential and proprietary.

05

Who May Call What, And How Often?

© 2026 eBay. All rights reserved. Confidential and proprietary.

Every Call Is Admitted



- **Rules by caller and pattern:** a tool name, a prefix like `get_*`, or a hint like *read-only* or *destructive*
- **Evaluated in process,** against a local copy, with no network call per request

Policy is data from many owners, decided locally on every call.

© 2026 eBay. All rights reserved. Confidential and proprietary.

Rate Limits Without A Global Bottleneck

- **Two budgets per server:** one per caller, one for everyone combined
- **Only real work counts:** tool calls and resource reads, never discovery
- **Counted per cluster, summed across clusters,** synced every quarter-second
- A policy denial never spends quota; a store outage never looks like one
- **Denials come back as a tool result:** "retry after N seconds", so the model backs off

The realistic abuse case isn't an attacker. It's an agent stuck in a loop.





© 2026 eBay. All rights reserved. Confidential and proprietary.

So, What Should You Take Home?

© 2026 eBay. All rights reserved. Confidential and proprietary.

Three Takeaways

- **Never forward a token you weren't meant to receive:** exchange it, one per backend
- **Make discovery a tool,** and admission a decision on every call
- **Keep onboarding declarative:** servers and provider quirks are reviewed configuration, not code

One URL for the user. The real user at every backend.



© 2026 eBay. All rights reserved. Confidential and proprietary.



Thank You!