



Our Common Journey to a Federated User Directory

Maghen

Software Engineer · Tchap/DINUM



Kai

IT Architect · BWI GmbH



BundesMessenger

A Simple Task

Given two homeservers, let each user find any other user 🔍



Homeserver A

Alice Bob



Homeserver B

Charlie

For current Matrix search only includes:

- local users
- users that share a room
- users in public rooms

Vector art by Vecteezy.com

The “Solution”

- Introduction of the *Discovery Room* strategy ~2022
- Users are discoverable when they share rooms with federated users
- Auto-join users to discovery room
- Hide discovery room from users

JOIN ALL THE USERS



The Problem

- User search over federation is not satisfactory
- DINUM and BWI independently stumbled over discoverability problems (2022–2026)
 - Difficulty to maintain: slowness of batches to synchronize users in discovery room
 - Major issues related to a huge number of members within a room: sync, profile update, state resolution...
- End up deactivating the Discovery Room strategy on the biggest homeservers
 - Removal of the discovery room was quite painful
 - Federated User Search became a degraded feature

A Story of Collaboration

- Timeline:
 - ~ 2024-2025 : Contributions to content-scanner(antivirus) from Tchap/BWI
 - 2025-Q4 BWI/Tchap: Strasbourg – Matrix Conference 2025
 - 2026-Q1 BWI/Tchap: Connection
 - 2026-Q2/Q3 BWI/Tchap/Element: discussion around MSC4258 and design
 - 2026-Q4 BWI/Element: MSC4259 – Profile Update EDU implementation design
- A journey
 - MSC4258: started the discussion and worked as a support
 - Tchap: implementation of the feature on a fork
 - BWI: suggestion of an implementation on upstream
 - Brainstorming BWI/Tchap
 - Difficulties to analyze and scope the feature along data privacy and security matters
 - Learning curve for Synapse codebase
 - Element: discussion around a long term solution
 - BWI/Tchap joined by Element still learning and collaborating

MSC4258 Approach

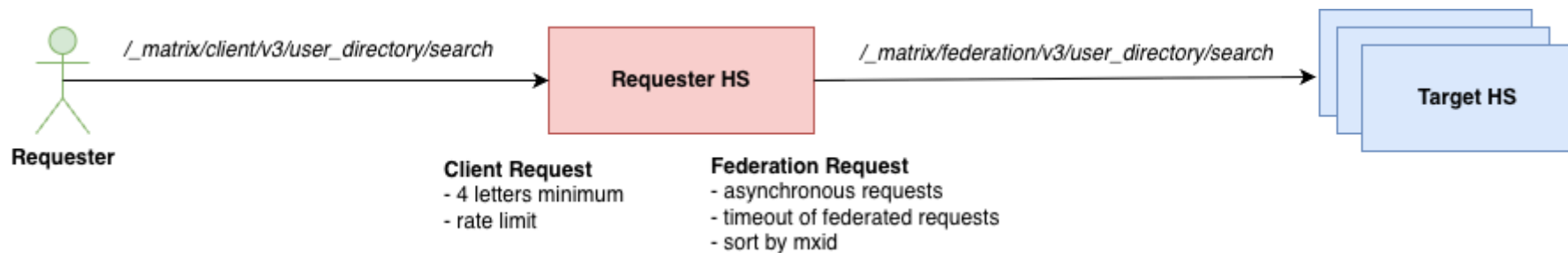
- Goal: search for all users in a federation
- Idea:
 - A user performs a search on their local HS
 - The local HS triggers a search to the remote HS
 - Remote HS replies to the search without propagating the search further
- Proposal:
 - Allow a local homeserver to perform a search request on a remote homeserver (new federation endpoint)
 - Ability to restrict the search to local users, known users or remote users
 - Give the users the ability to control their visibility in the user directory

Immediate Resolution

- Short term approach
- Build on [MSC4258](#) (2025)
- Closed federation
- Tight integration between federated homeservers
- Trusted homeservers $\Rightarrow$ Full list of participants
- Implementation:
 - Tchap: Perform a live search on each remote HS, with a short timeout
 - BWI: Perform search by storing remote profile data locally

Tchap

- Our requirements
 - Simple and loose coupling
 - Closed federation
 - Production issue: needs to be release ASAP
- Design solution (on a fork)



- On production: it does work for more than of 450k of actives users 🎉

BWI

- Requirements
 - Simple architecture
 - Closed federation
 - Eventual replacement of production solution
- Pre-fetch user data from all participating homeservers
- Create search index over huge pile of collected user profile info
- Custom federation endpoint
- Implemented as [Synapse PR #19887](#) in collaboration with upstream
- Will be improved upon by migrating to push architecture ([MSC4259](#))

Data Privacy

Two aspects.

- **Scraping:** Can public data be scraped en-masse? Is there a way to prevent effective scraping?
- **Policy:** Which user is allowed to be found by each homeserver?

Data Privacy – Scraping – User Attacker

- Can join/read rooms to gather profile data
- Can use systematic search
 - Model attacker with search oracle e.g. 10 new users per request
 - * e.g. enabled by dictionary attacks, might have diminishing returns
 - Rate-limits apply e.g. ~ 1000/days
 - 400.000 users → 40 days
- Both can be amplified by fishing/social engineering

Data Privacy – Scraping – Server Attacker

- Acts as proxy for all users on homeserver
- → Rate-limits are magnitudes less efficient
- → Keeping profile data distributed with online search yields no real benefit over proactively collecting all data directly through the federation

Data Privacy – Policy

- Handling of custom profile data (MSC4133)
 - Role, status, organization, email, phone
 - Maybe similar in character to IAM attributes
- Who can set policy?
 - Per-user/per-server
- What can be accessed by whom?
 - Conditions/granularity of visibility
- Will become relevant for actual standardization
- Challenge is to fit actual use-cases
- Add all knobs that are useful, but don't add too much unnecessary ones



Do you have a:

- *use-case for custom profile data?*
- *need for a special access policy*

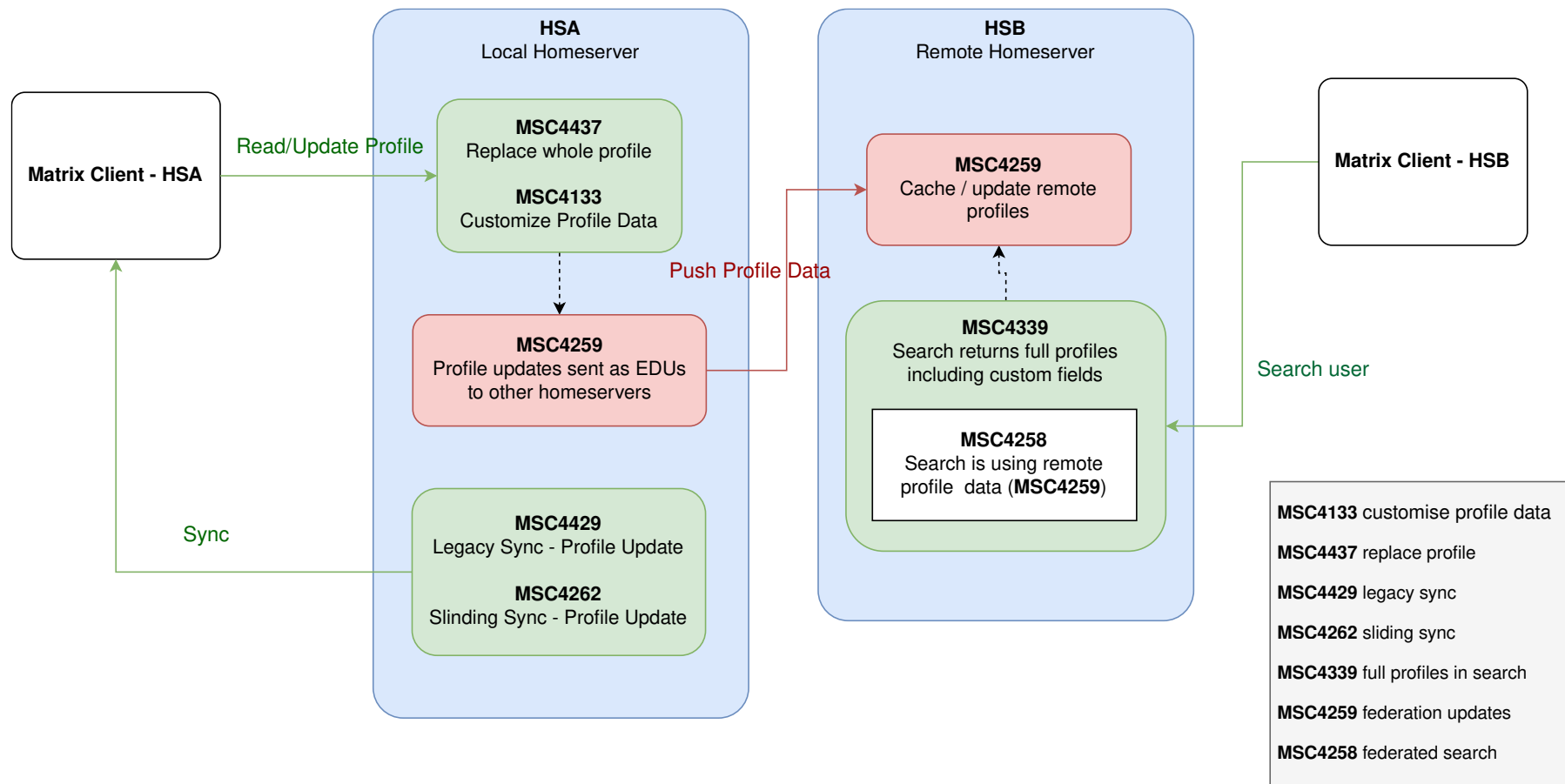
⇒ ***We want to hear from you!!***

The Next Steps

Specification

- What people want: Discoverability beyond known room members and own homeserver
- Long term approach
- Make it as a community feature (please help us!)
- Topics are still in discussion
 - Privacy
 - Performance
 - Open federation

Overview of Profile Data MSCs



MSC4259: Profile Update EDUs for Federation

- User search relies on synchronized profile data
- Push semantics as more elegant replacement for current pull semantics
- More performant synchronization of profile data
- Robust and more maintainable design
- No new data created, just remodeling existing data flows
- Enable other feature on federation:
 - Display profile data with their custom field $\Rightarrow$ User status
 - Advanced search on custom profile fields

Wrap up

- First collaborative work between French and German Public Sector
- Different perspectives helps to better understand a problem
- Quality has improved (exchange ideas, different requirements, more robust approach)
- Getting more confident by working together
- Interaction between French and German teams are becoming more frequent

Do you have questions?

Contact

Kai A. Hiller

@V02460:matrix.org

kaialexander.hiller@bwi.de

Maghen Calinghee

@mcalinghee:matrix.org

maghen.calinghee.ext@numerique.gouv.fr