

Evolveum

**Project midScale
Coaching Session #2**

May 2021

Agenda



- Objectives
- Progress (START-M4, M5)
- Outcomes and achievements
- Plans (M5-FINISH)

Project Objectives

- Increase midPoint scalability by an order of magnitude

Routinely manage: ~1 million identities

Possible deployments: ~10 million identities

- Identify scalability and performance obstacles

Performance testing environment

- Overcome scaling obstacles

Data store: data model, indexing, partitioning, etc.

Clustering / multithreading

Stability: thread safety, tooling

Visibility & usability: UX, GUI performance, diagnostics

Project Progress Overview

Milestone	Goal	Planned date	Status
START	Project start	15 Oct 2020	DONE
M1	Performance environment	26 Nov 2020	DONE (report)
M2	Performance evaluation and repository analysis/design	15 Jan 2021	DONE (report)
M3	Multithreading, Schrodinger and Query language	26 Feb 2021	DONE (report)
M4 (midPoint 4.3)	Performance Repo prototype, UI basic tests, Multinode Tasks	12 Apr 2021	DONE (report)
M5	PostgreSQL, Performance environment (2), UX Analysis & Design	31 May 2021	In progress
M6	Repository optimization, performance evaluation	9 Jul 2021	
M7	Migration Procedure, GUI Improvements, Auto-scaling	31 Aug 2021	
M8 (midPoint 4.4)	Project finish	14 Oct 2021	

Overall Status

- On track, progress according to plan
- Small deviations from original plan
 - Some tasks are done ahead of time
 - Scope of some tasks was adjusted to account for work done ahead of time
- Extra unfunded resources: Kamil Jireš, Radovan Semančík
 - Internal funding from Evolveum, not funded by NGI

Major Achievements (START-M4)

M4

- Architecture and Design (M1, continuously updated)
- Performance evaluation (preliminary)
- Axiom query language (M2-M3)
- Schrodinger testing framework & tests (since M2)
- Multi-threading, multi-node tasks, diagnostics and robustness improvements
- Testing environment (since M3)
- Native PostgreSQL repository prototype (M4, key functionality)
- Survey (M3-M4)
- MidPoint 4.3 Release

Axiom Query Language

- Human-friendly query language
- Goal: administration of system with millions of identities
- Based on concepts of Axiom data modeling

Axiom is outcome of midPrivacy: Data Provenance

- Community workshop/webinar (May 2021)

<https://www.youtube.com/watch?v=UACuiXQ8w-Y>

`name = 'smith' or familyName = 'Smith'`

```
<filter>  
<or>  
  <equal>  
    <path>name</path>  
    <value>smith</value>  
  </equal>  
  <equal>  
    <path>familyName</path>  
    <value>Smith</value>  
</or>  
</filter>
```

Schrodinger Testing Framework

- Automated testing of MidPoint user interface
- Simulating user working in web browser
- Fluent, easy-to-write test code
- Maintainable, even if user interface design changes
- Goals

Primary: user interface stability

Secondary: evaluating user interface scalability

```
usersPage
    .table()
    .search()
    .byName()
    .inputValue("mike")
    .updateSearch()
    .and()
... ..
    .clickSave()
    .feedback()
    .isSuccess();
```


Survey

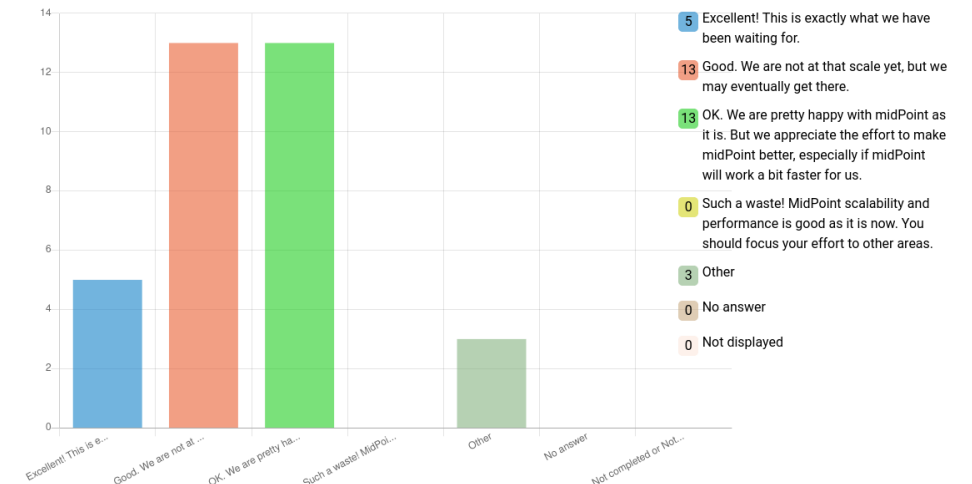
- MidPoint community survey, February-March 2021

<https://docs.evolveum.com/midpoint/projects/midscale/survey/>

- Relatively low number of responses (34)
- Validating midScale purpose and goals
- Positive results

Results of midScale will be a benefit to midPoint community and Evolveum customers

The goal of midScale project is to increase the total number of identities that midPoint instance can handle. We hope to be able to routinely support deployments with tens of millions of identities. How do you feel about this plan?



MidPoint 4.3 “Faraday”

- Released during M4 of midScale project (9th April 2021)
<https://docs.evolveum.com/midpoint/release/4.3/>
- All midScale functionality integrated, available to the community
Some functionality marked as *experimental*, gathering community feedback
PostgreSQL repository prototype code is present, but not enabled by default
- Community workshop/webinar (April 2021)
Recording available on Evolveum YouTube channel
<https://www.youtube.com/watch?v=vrrrEOusZrE>

| Outcomes (M5, in progress)

- Tests running in testing environment
- Better performance evaluation
- PostgreSQL repository closer to completion
- Public on-line workshops/webinars

M5

Testing Environment

- Dynamic environment
- Kubernetes/Docker
- Jenkins
- Grafana
- Gitops – Gitlab, Flux – verification in progress

Performance Gains (preliminary results)

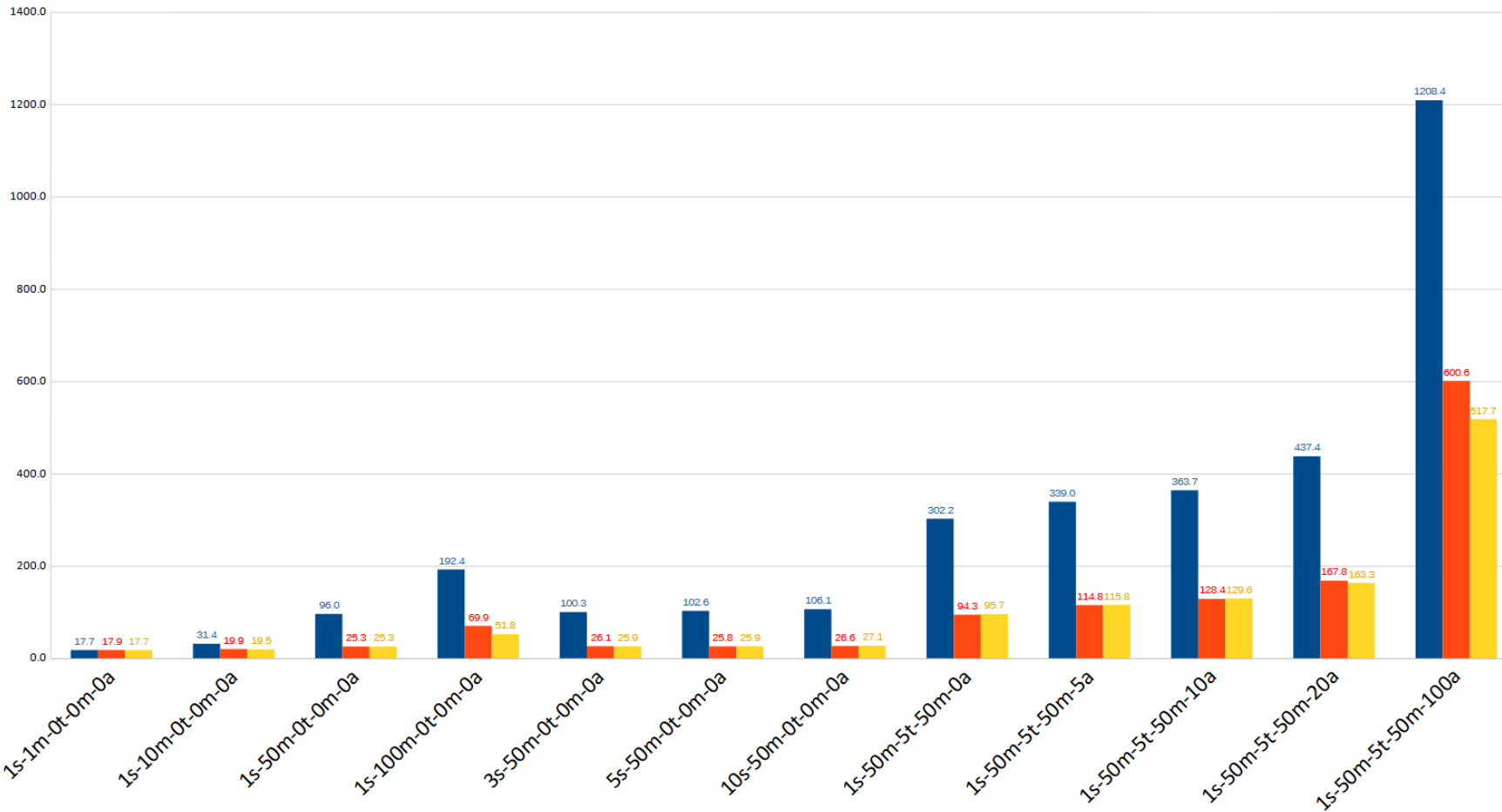
Initial import of accounts

Decrease of processing time in the most complex case:

4.2 → 4.3: **50.3%**

4.2 → master: **57.2%**

■ support-4.2
■ support-4.3
■ master



Scenario [s = sources, m = mappings, t = targets, a = assignments]

Performance Gains (preliminary results)

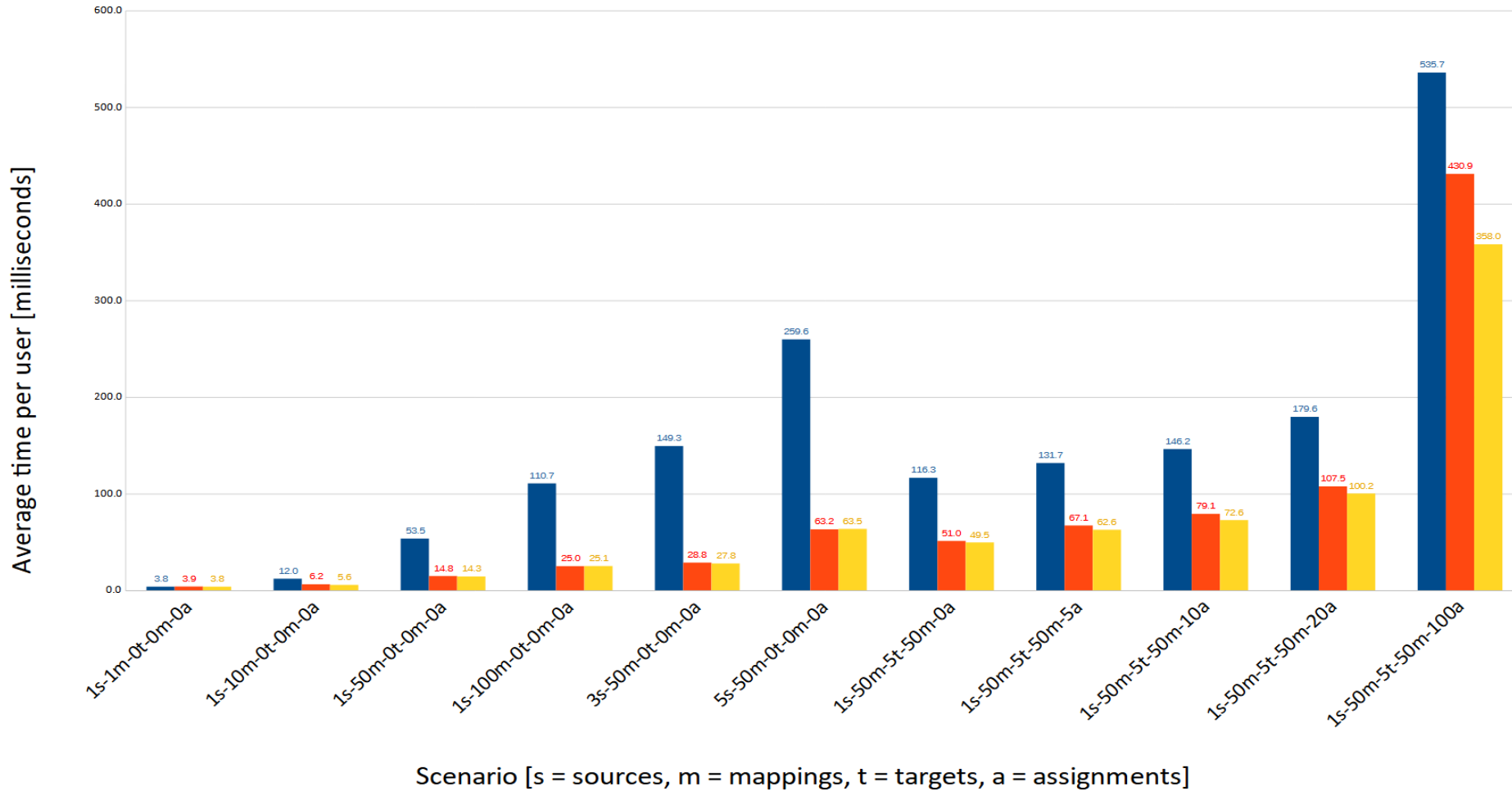
User recomputation

Decrease of processing time in the most complex case:

4.2 → 4.3: **19.6%**

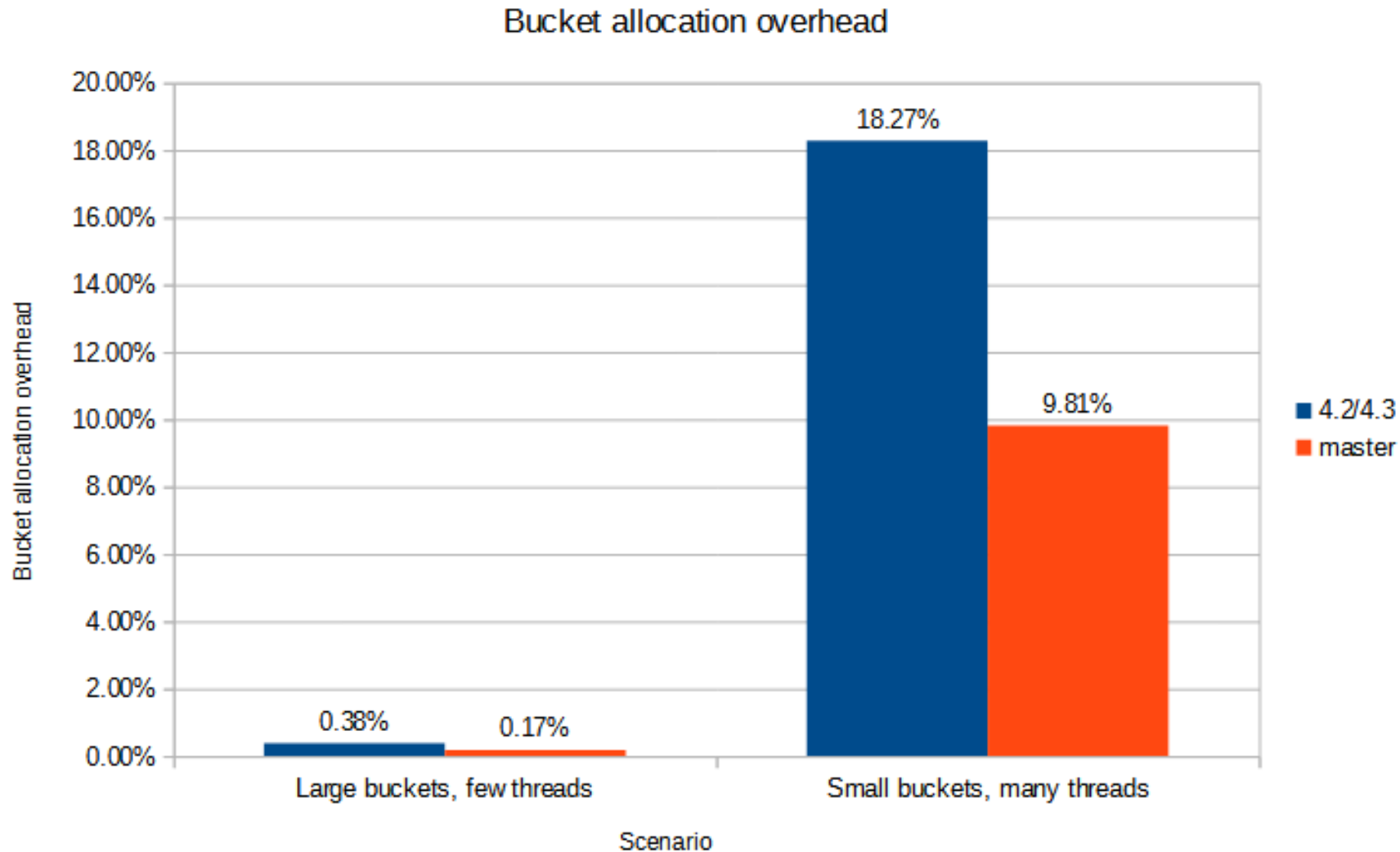
4.2 → master: **33.2%**

■ support-4.2
■ support-4.3
■ master

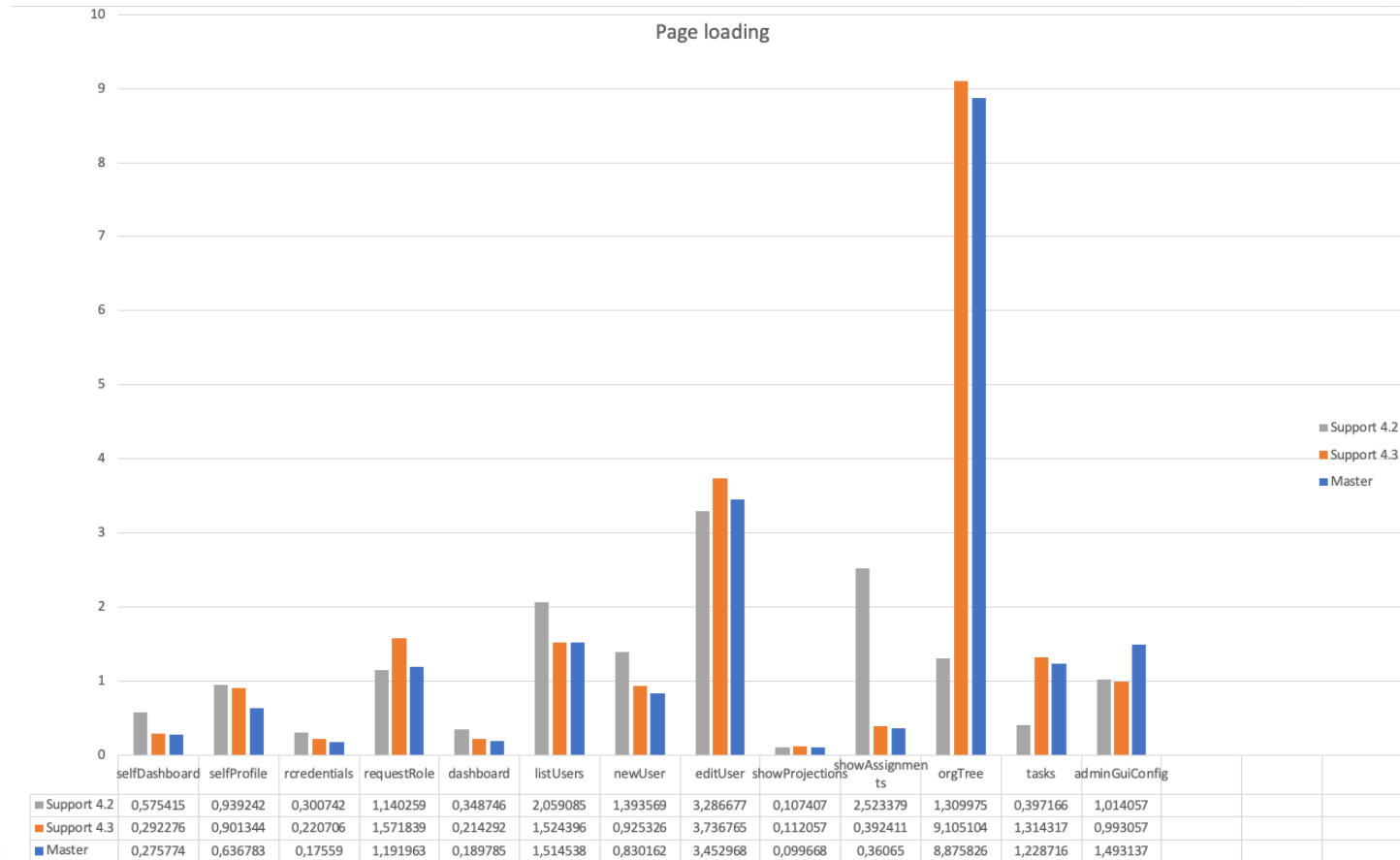


Scenario [s = sources, m = mappings, t = targets, a = assignments]

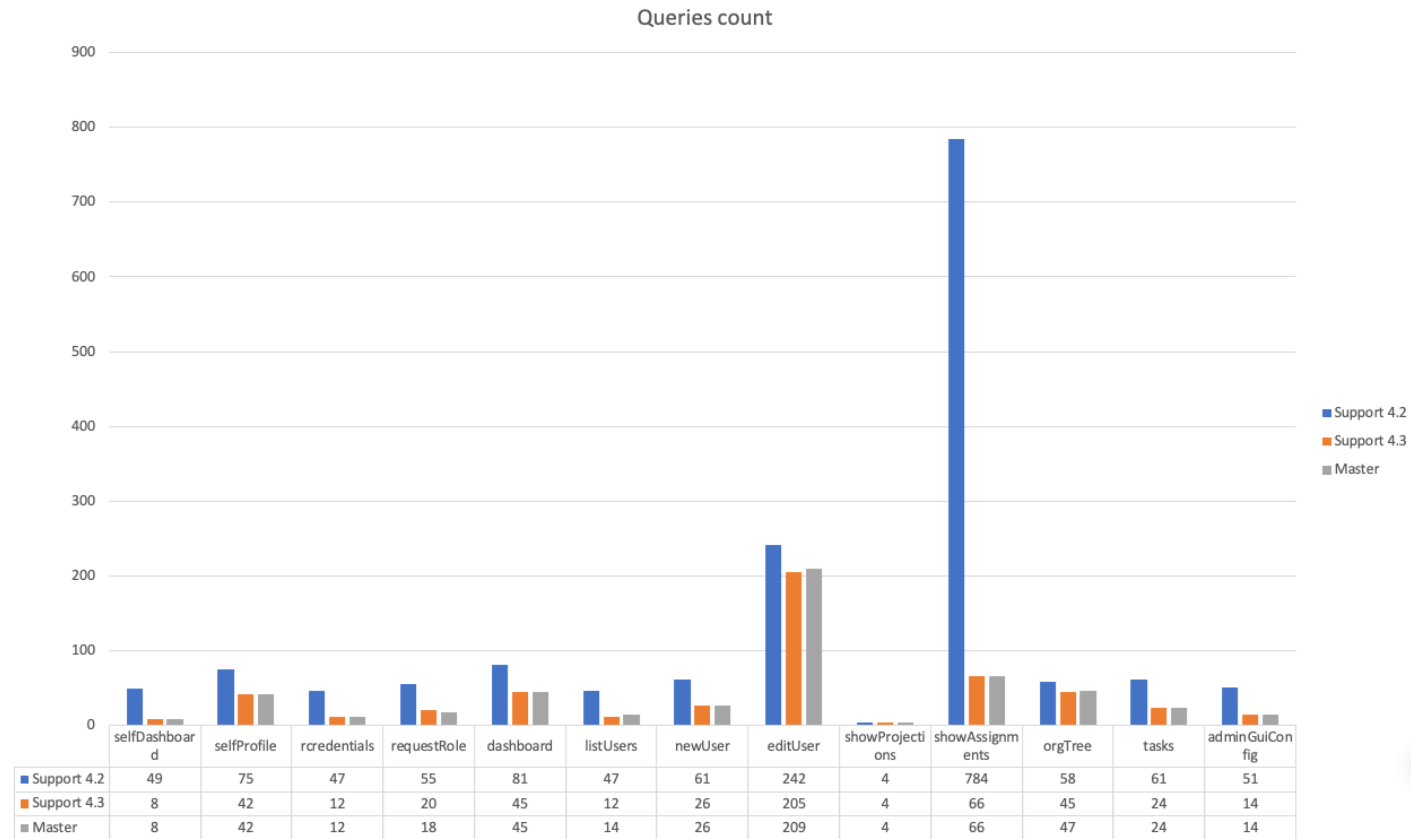
Performance Gains Tasks (preliminary results)



Performance Gains UI (preliminary results)

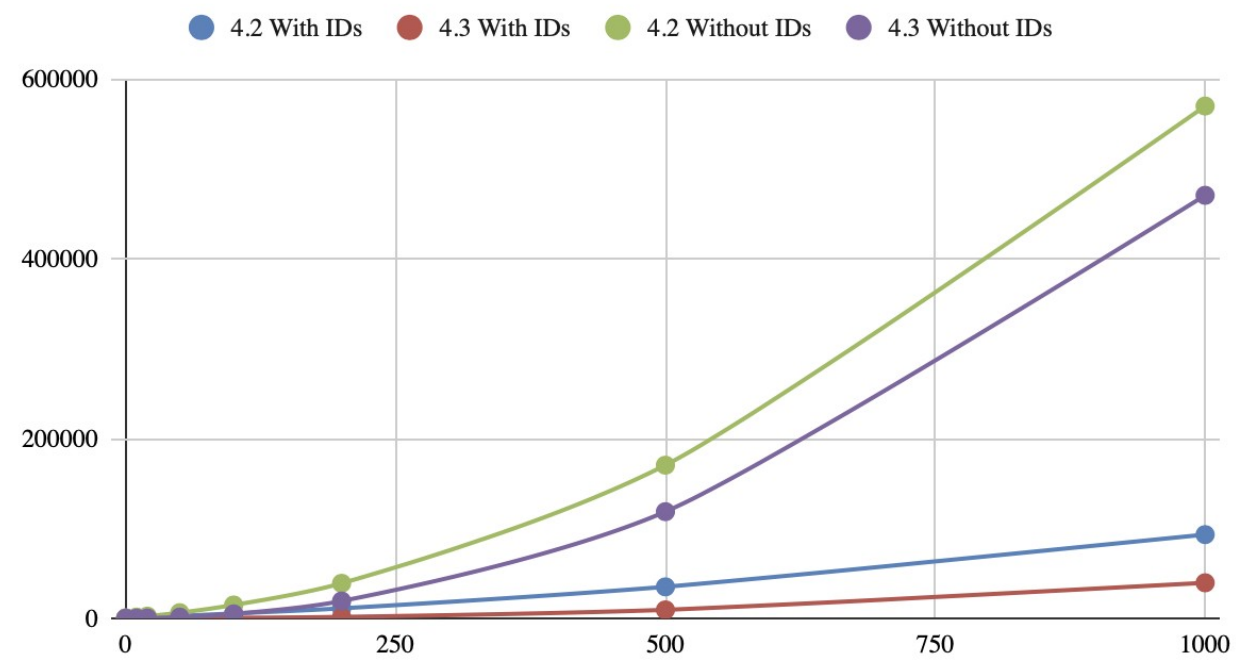


Performance Gains UI (preliminary results)

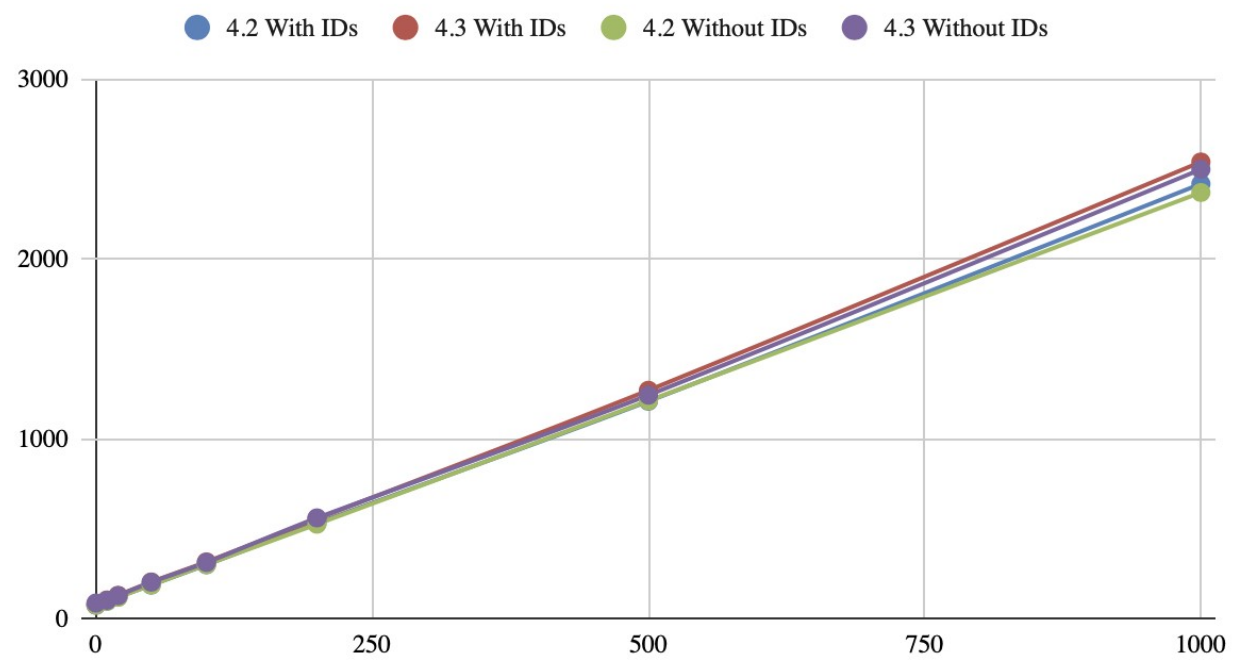


Prism performance analysis #1

Prism Unmarshaling

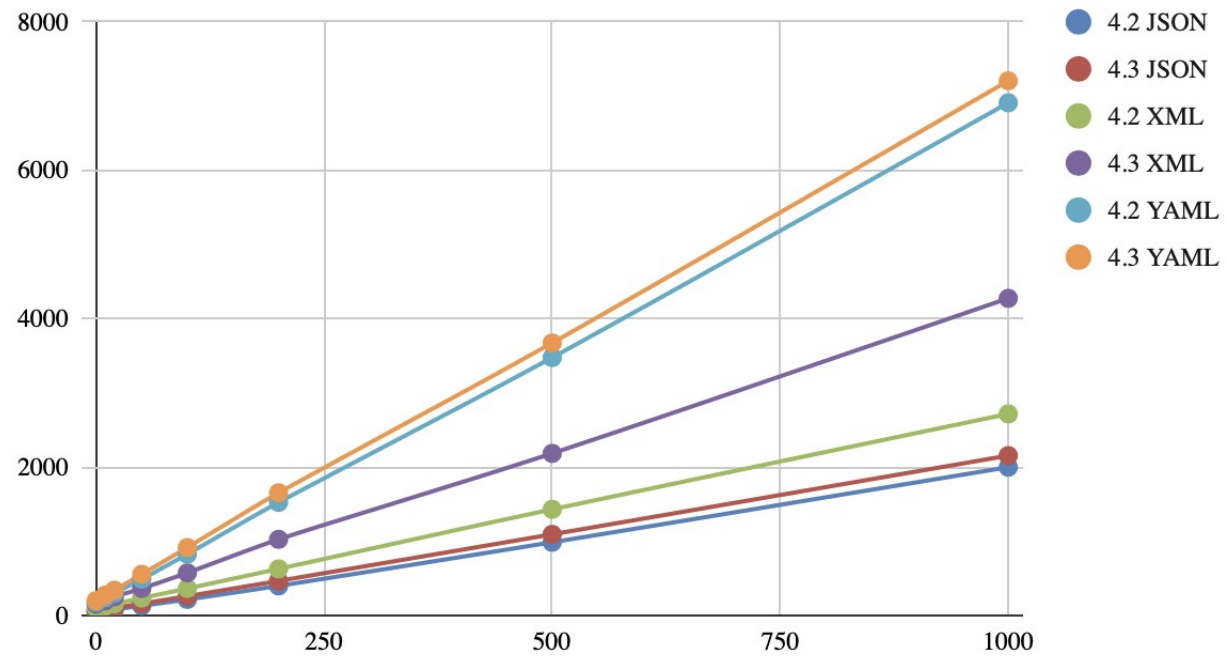


Prism Marshalling

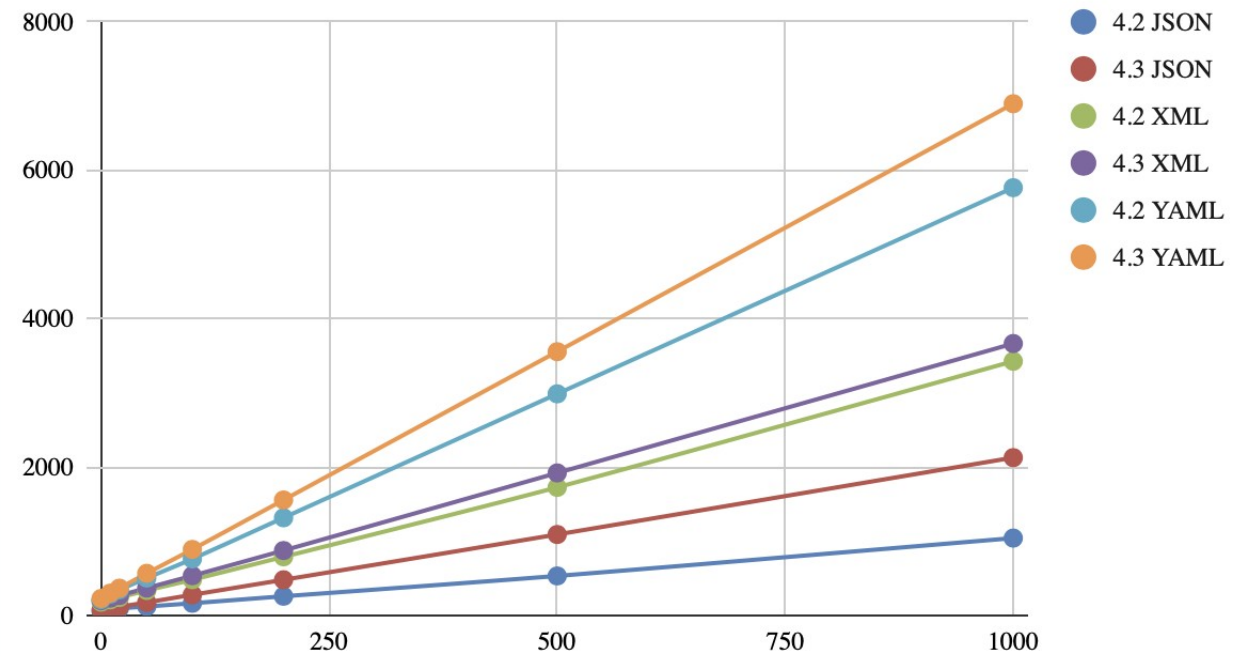


Prism performance analysis #2

XNode Deserialization

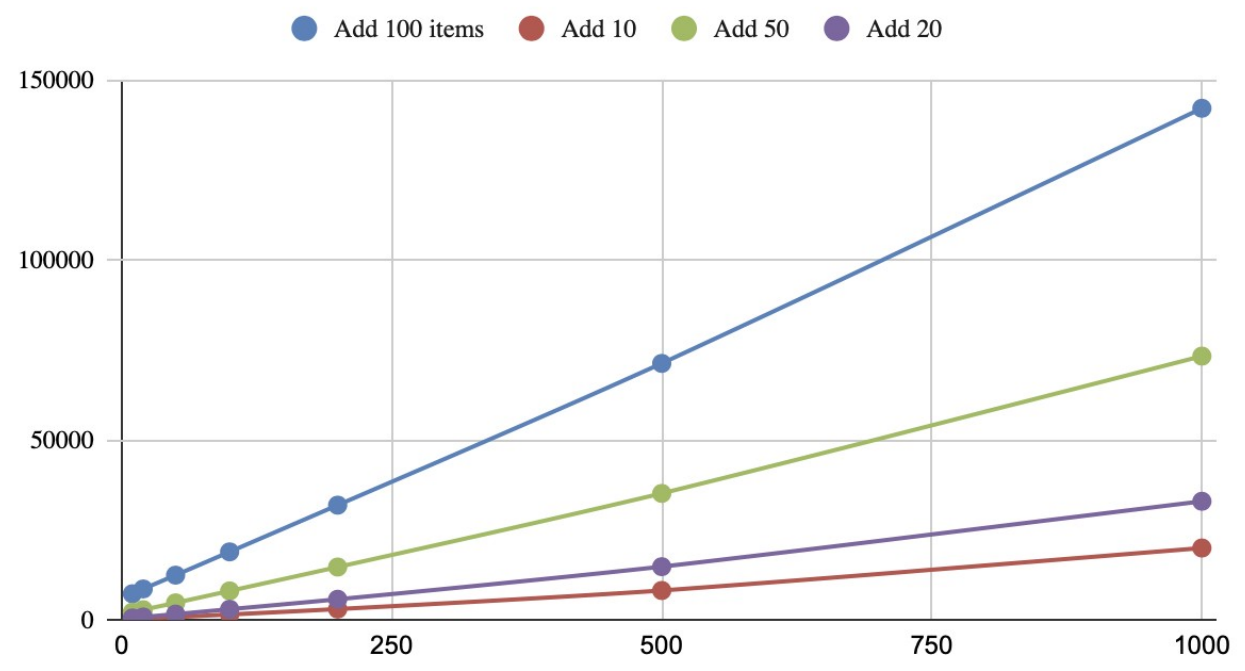


XNode Serialization

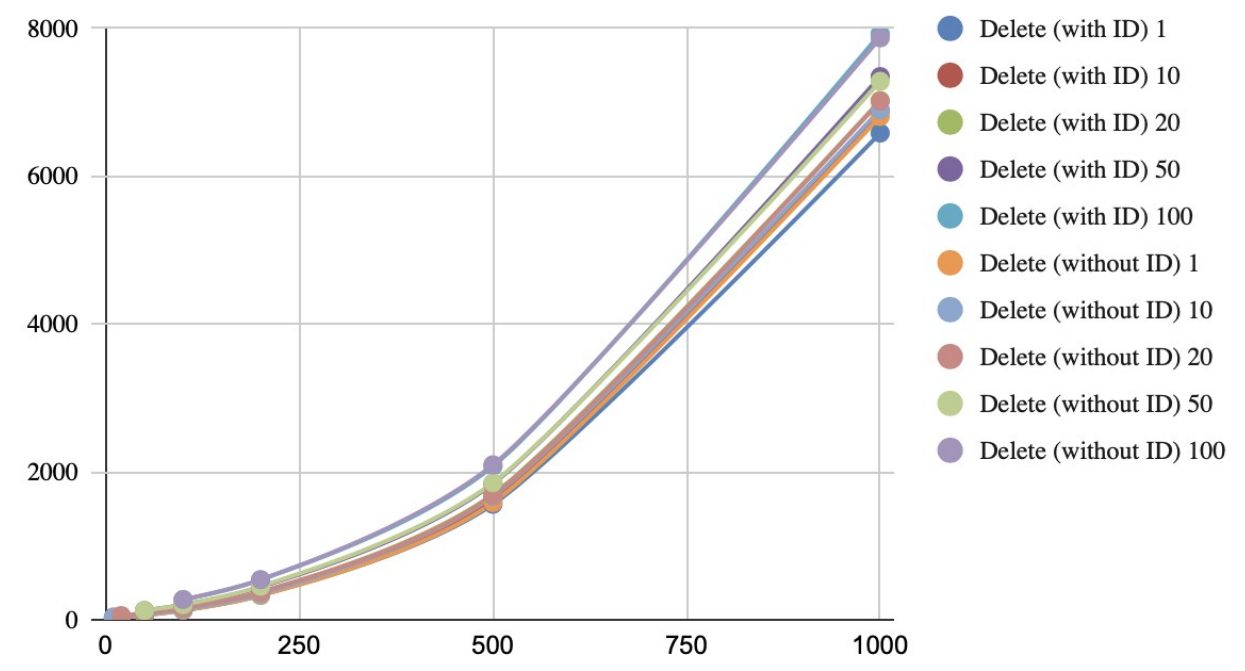


Prism performance analysis #3

Delta Add



Delta Delete



Native PostgreSQL Repository

- Prototype delivered in midPoint 4.3
- Key scalability functionality (high development priority)
- Still work in progress
- Development progress according to plan
- Added extra resource to lower the risk

Project and Outcomes Propagation

- Evolveum blog posts
- Evolveum mailing lists
- Public on-line workshops/webinars
 - Recordings will be available shortly
- Evolveum YouTube channel
- Internet2/InCommon communication channels (selective)

Budget

Organization	Total Budget (EUR)	Used Budget (EUR)	Used Budget (%)
Evolveum	249 200,-	138 934,-	55

Plans: M5-FINISH

- Finish PostgreSQL repository implementation, test & optimize
- GUI and UX improvements
- Upgrade/migration tools (old DB → PostgreSQL repository)
- Auto-scaling, clustering, task management improvements
- Test, test & test
 - Evaluate and re-evaluate performance and scalability
 - Continuous improvement of tests and test environment
- Stabilization, documentation, prepare for midPoint 4.4 release

Project Resources

- Project home page

<https://docs.evolveum.com/midpoint/projects/midscale/>

- Solution Architecture

<https://docs.evolveum.com/midpoint/projects/midscale/architecture/>

- MidPoint source code

<https://github.com/Evolveum/midpoint>

- Evolveum Blog

<https://evolveum.com/blog/>

Discussion

Questions & Answers

Thank you for your time

See other talks at <https://docs.evolveum.com/talks>

Also **follow us** on our social media for further information!



/Evolveum



/Evolveum



/Evolveum



@Evolveum



/Evolveum

Evolveum