

Research project structure and layout

Abstract

Introduction

Introduce RBAC and its usage in databases, explain how it works overall.

Explain the motivation behind the project, that I am collecting data and researching implementation of RBAC. The purpose for the research is to later figure out which features, and performance can be expected for existing RBAC systems, with the intention for later developing an open standard and looking at how to possibly implement it in data lake architectures.

Analysis

Explain the reason to look at existing implementations, and their features and usage. Explain my choices for choosing 4 open source and 2 proprietary database providers.

PostgreSQL

Shortly explain PostgreSQL any distinct features in its service.

Explain its implementation of RBAC.

Creation of roles,

Grant structure who can grant what and how.

Snowflake

CockroachDB

SurrealDB

MariaDB

Microsoft Synapse

Analysis Conclusion

Show my feature matrix and the overall results of my research into the different providers

Performance experiments

Explain the reason why I want to do experiments. And why I chose to do performance testing on PostgreSQL, MariaDB and Snowflake.

Experiments

Explain the different experiments that I performed. First explain the tree types and the reasoning behind that choice. Since we want to see if the role structure had an impact on performance. Then explain the experiments the create roles over 15 min. explain create 1000 roles, and then do select table and info schema.

Experimental setup

Explain the setup. How MariaDB and PostgreSQL was hosted locally on my machine and how I create a python scripts connect and run the tests and their repetition. Also including a description of my pc's specifications.

Then write about how Snowflake was hosted remotely and that I connected to their service via a python script to run the experiments, then also explain that for the 15 min experiment was also repeated directly in snowflake by running an python script in their service remotely

Results of experiment

show all my results and experiments visually write about whether it was expected or unexpected. Also, for PostgreSQL and snowflake calc the median and show the value.

Discussion

Talk about the feature matrix and which features are common, which should be kept if I were to develop an open standard.

Talk about the results of the performance tests and what kind of performance tendencies should an implementation show over time and the number of roles.

Conclusion

Summarization of all my main points from the report.

That common sql statements that are almost universal like **create role <name>, grant <privilege> on <role>, grant all on <role>, grant <role> to <role>, revoke** should be a part or supported by the open standard.

That its possible to create a data structure that can create and grant roles at an fast and almost constant time.