

LDBC Financial Benchmark Introduction

Heng LIN @ Ant Group

June 16 2022

1 / Background

2 / FinBench Design

3 / FinBench Progress

1 / Background

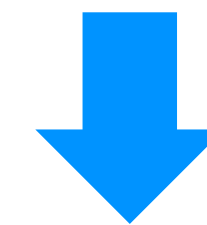
LDBC Benchmarks



SPB An RDF-based benchmark for semantic databases

Graphalytics Graph algorithms benchmark

SNB - Benchmark based on social network scenarios
- Limited when applying to financial field.



FinBench Benchmark based on financial scenarios

Differences From SNB

Application Scenarios

- SNB: friend recommendation, ...
- FinBench: risk control, AML, KYC, ...

Schema Characteristics

- SNB:
 - Single edge
- FinBench:
 - Multiple edges
 - More properties in edge

Workload Characteristics

- Tight latency bounds (e.g. P999 20ms)
- Read-write query
- Temporal window: constrained by start_time
- Different subgraph patterns (e.g. cycles, paths, chains)
- Recursive path filtering: Recursive path filtering and regular path query
- ...

2 / **FinBench Design**

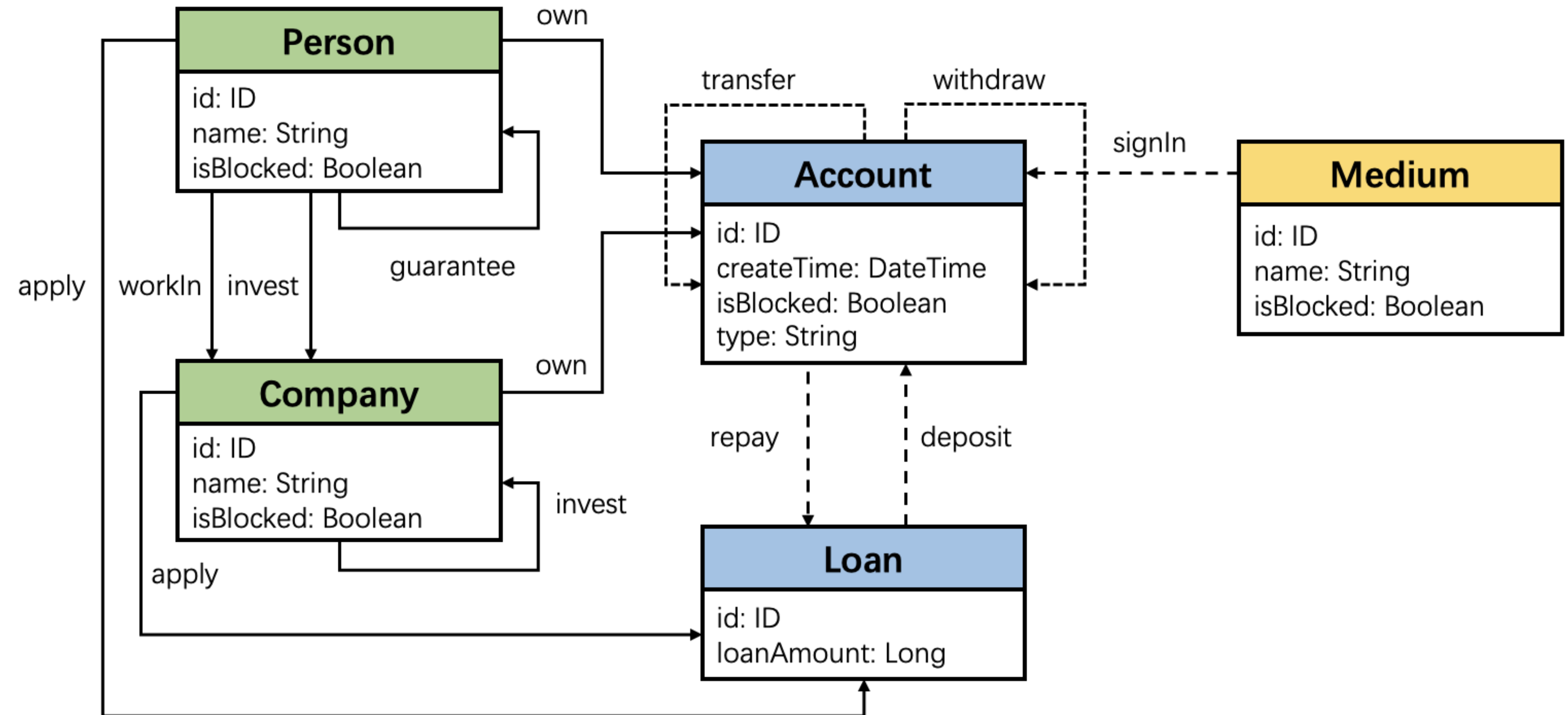
FinBench Data Schema (Proposed)

Vertices:

Person, Company,
Account, Loan, Medium

Edges:

transfer, withdraw, repay,
deposit, signIn, apply,
workIn, invest, guarantee,
own



Note: The dashed arrows represent multiple edges from source vertex to the destination

FinBench Workload (Proposed)

Workloads

TP Workload

Latency:
Tight bound

Query:
1-4 hops

AP Workload

Latency:
unlimited

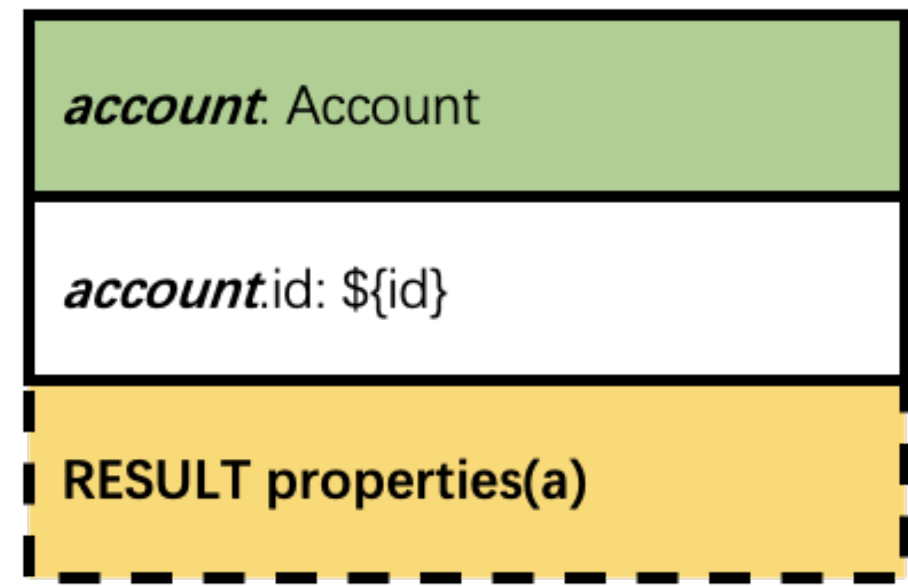
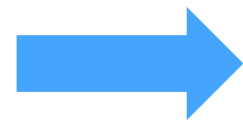
Query:
Iterative graph analytics

Load Definition

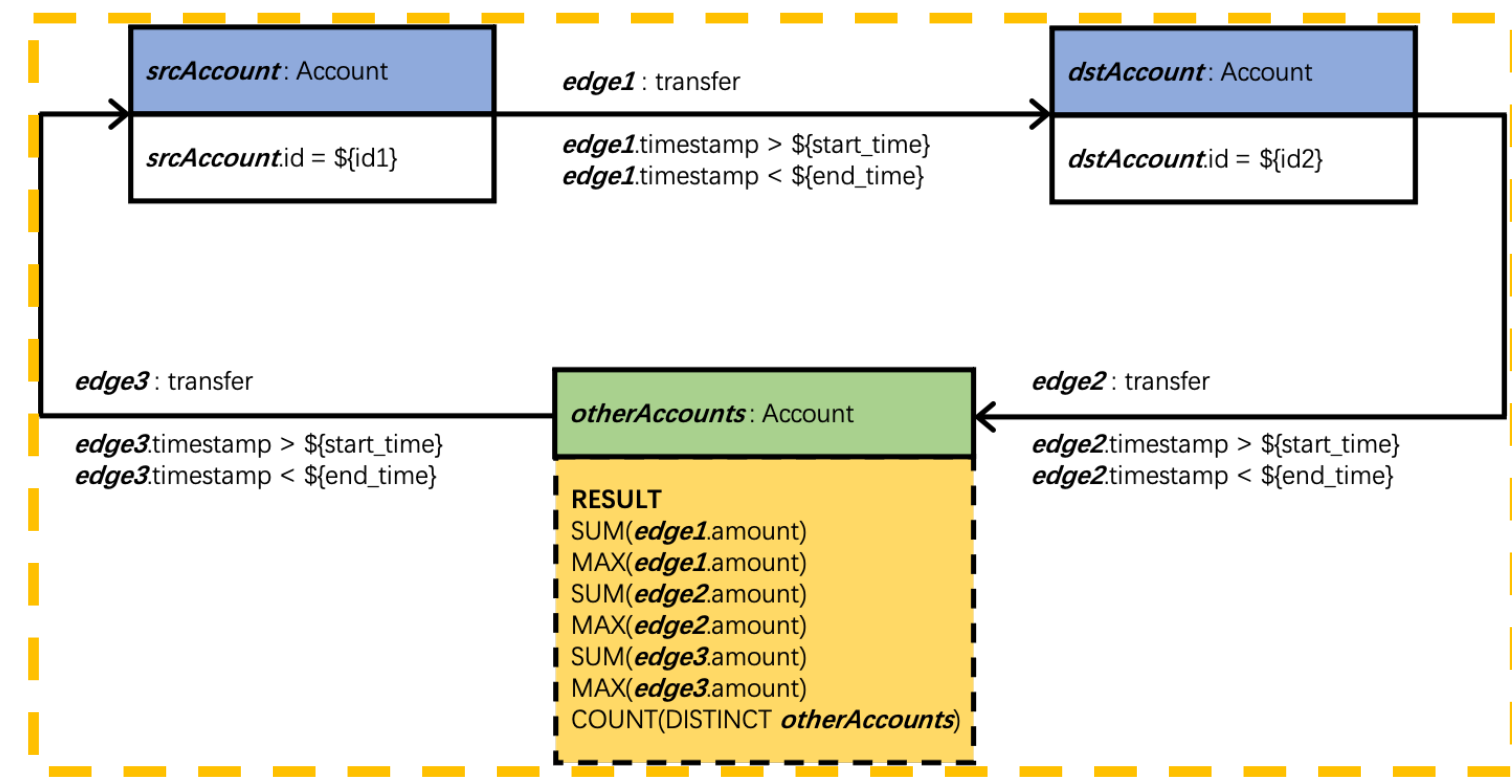
- Strong transactional requirement
- Peaks and troughs
- ...
- Time-to-live data management
- Periodical data influx

FinBench Query Example: Read-Write Query

Read Query



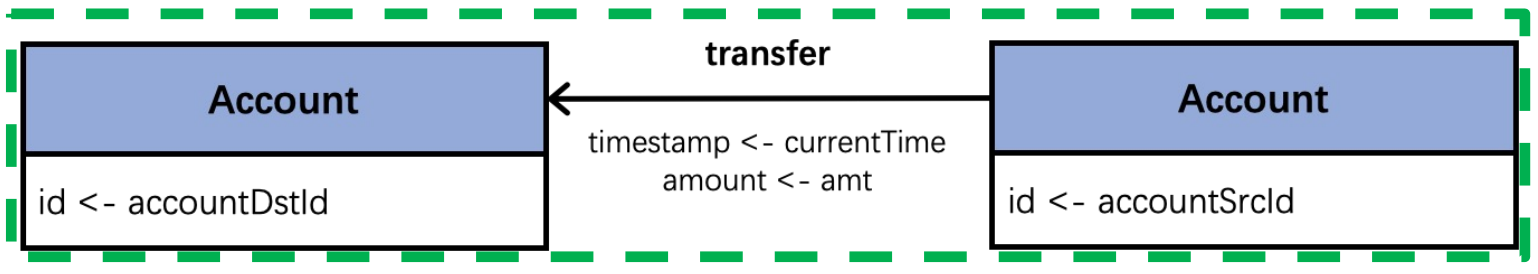
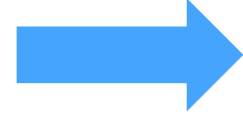
Exact Account Property Query
[Ref: read / 22]



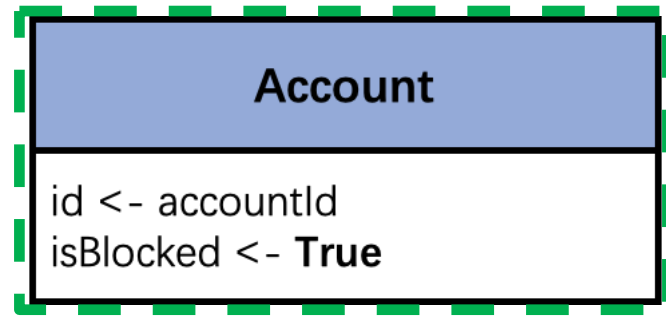
AML: Three accounts in a cycle
[Ref: read / 4]

Read-Write Query

Write Query



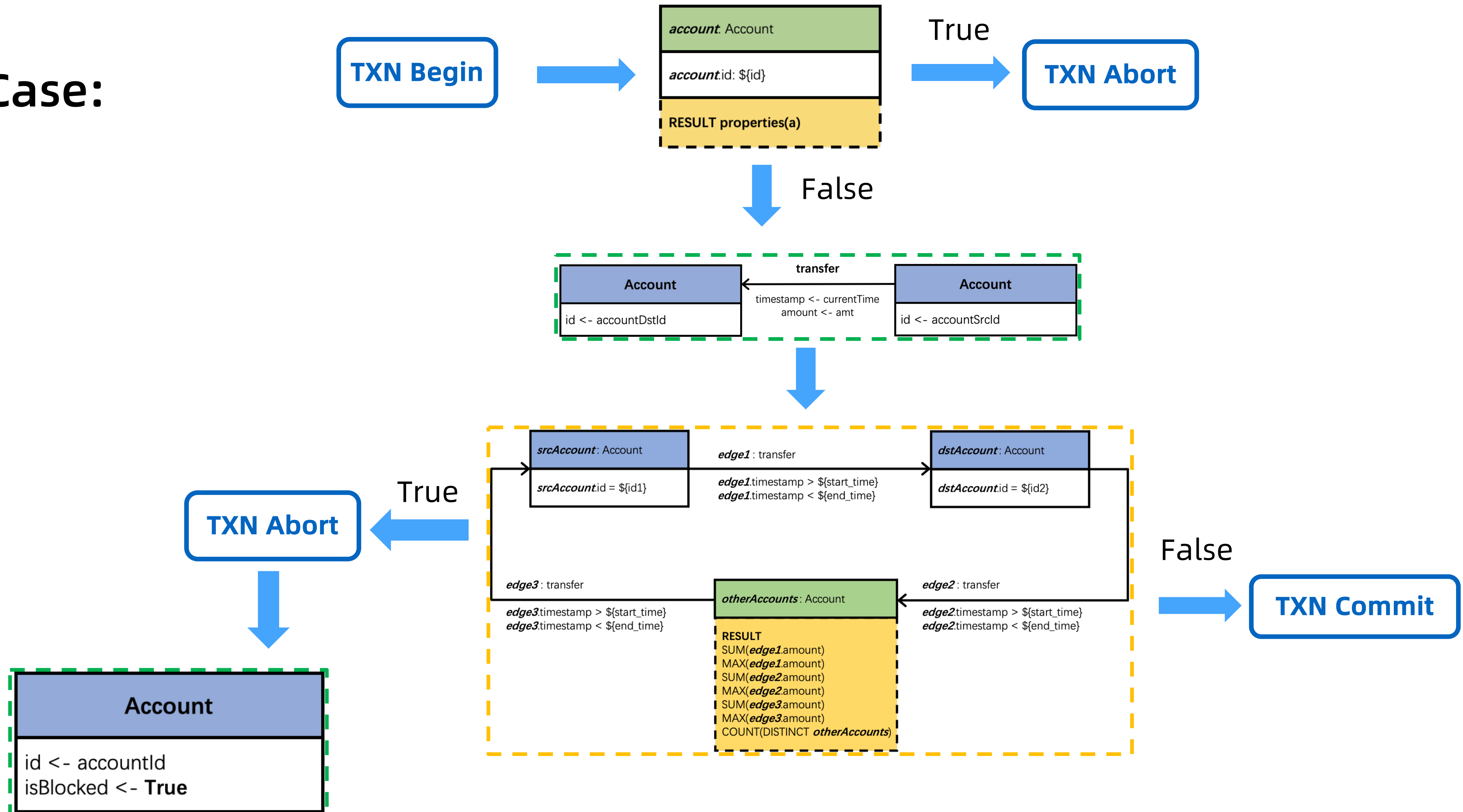
Add a transfer edge
[Ref: write / 3]



Mark an account as blocked
[Ref: write / 10]

FinBench Query Example: Read-Write Query

AML Case:



FinBench Query Example: Temporal Window

Financial Scenarios

With the passage of time:

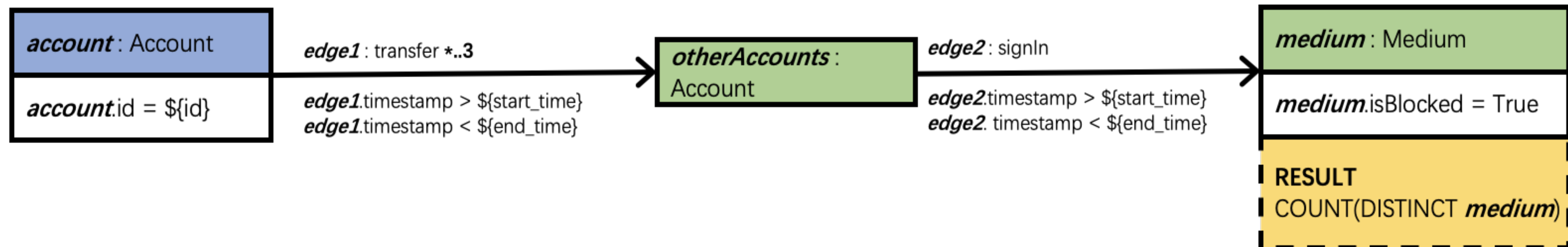
- Queries only look back in a temporal window



FinBench Design

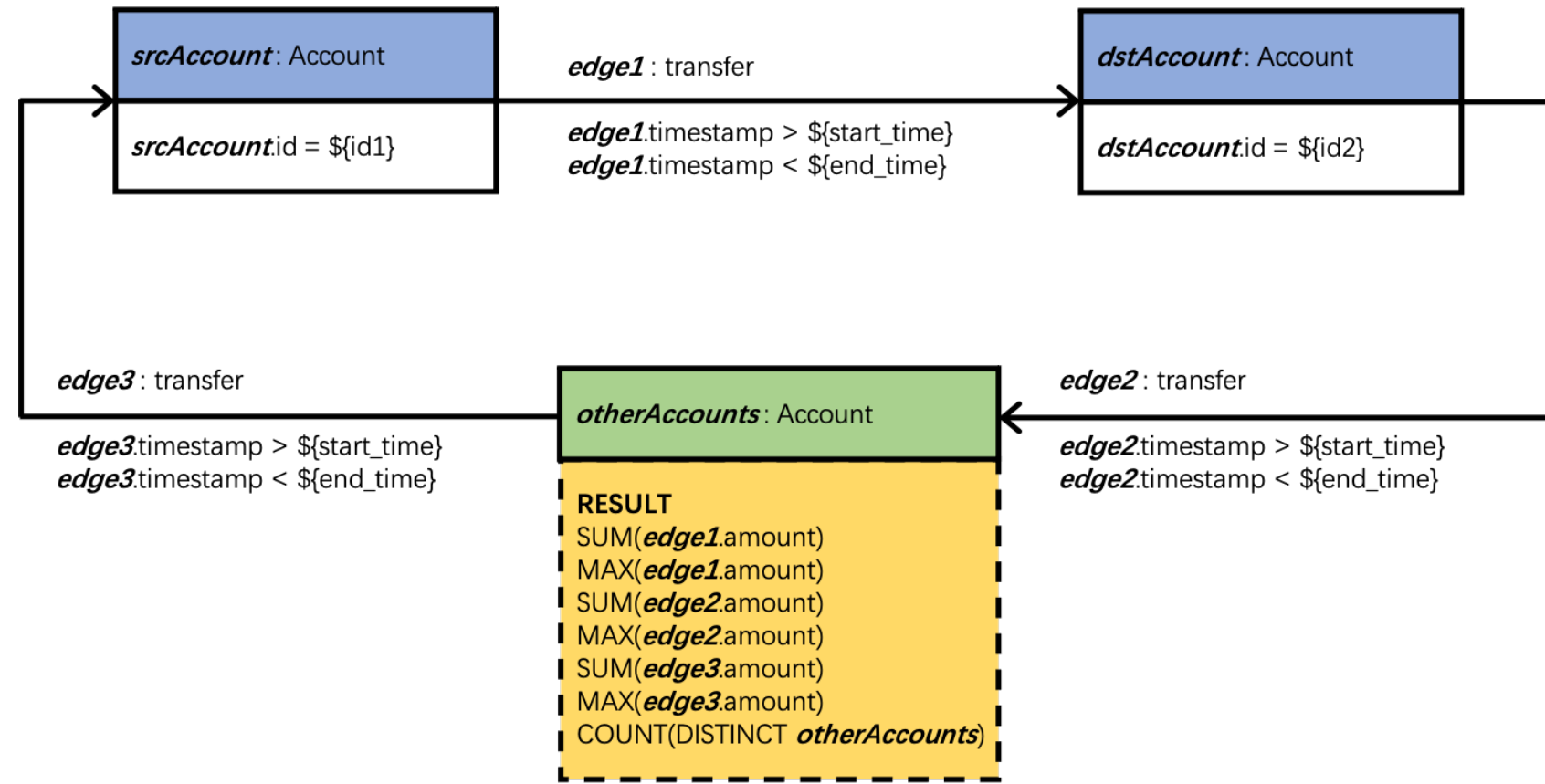
Queries are constrained by:

- After `start_time`
(In implementation, systems can choose to deprecate old data.)

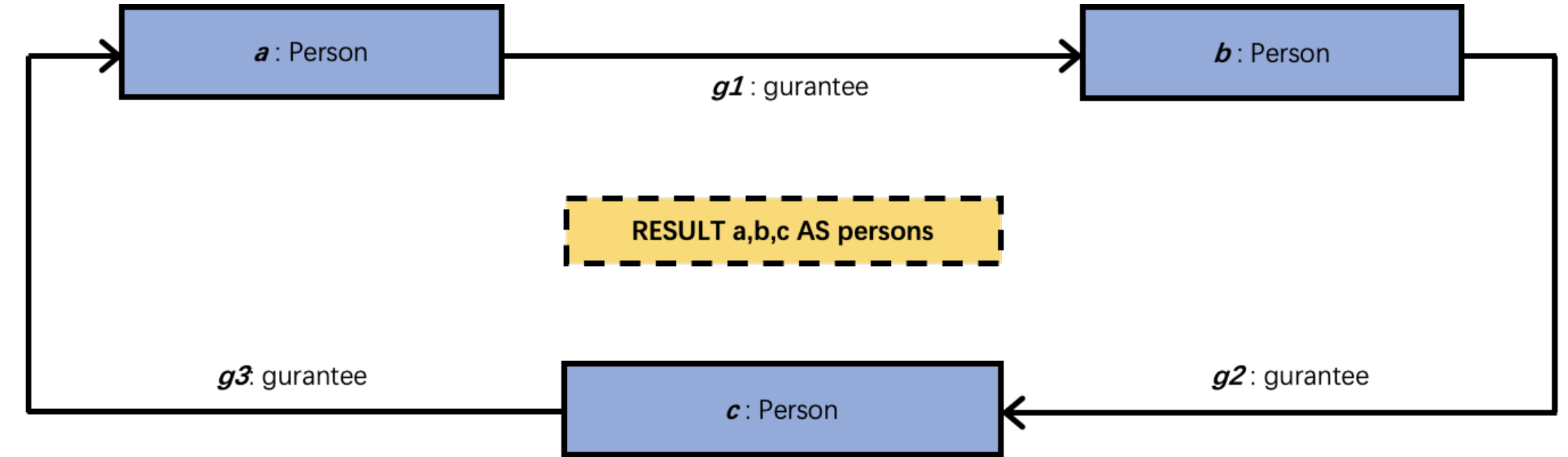


Blocked medium related accounts
(query constrained by `start_time` and `end_time`)
[Ref: read / 1]

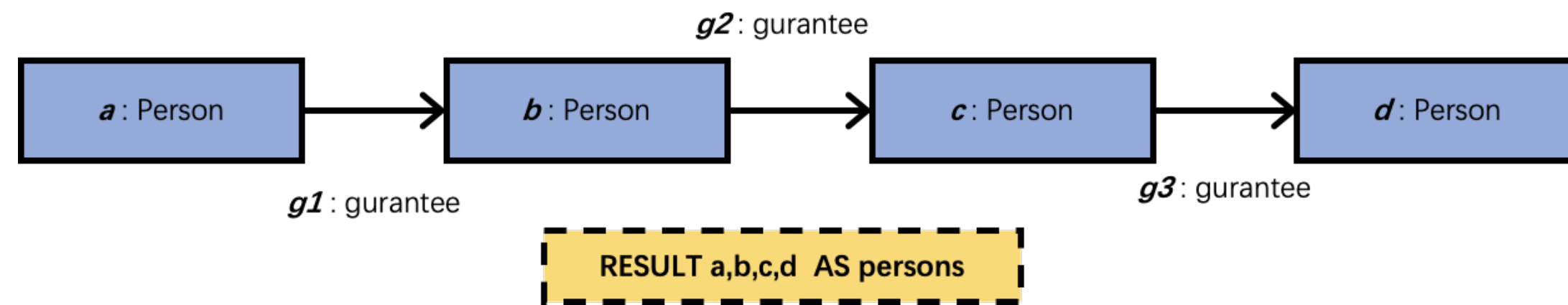
FinBench Query Example: Subgraph Pattern



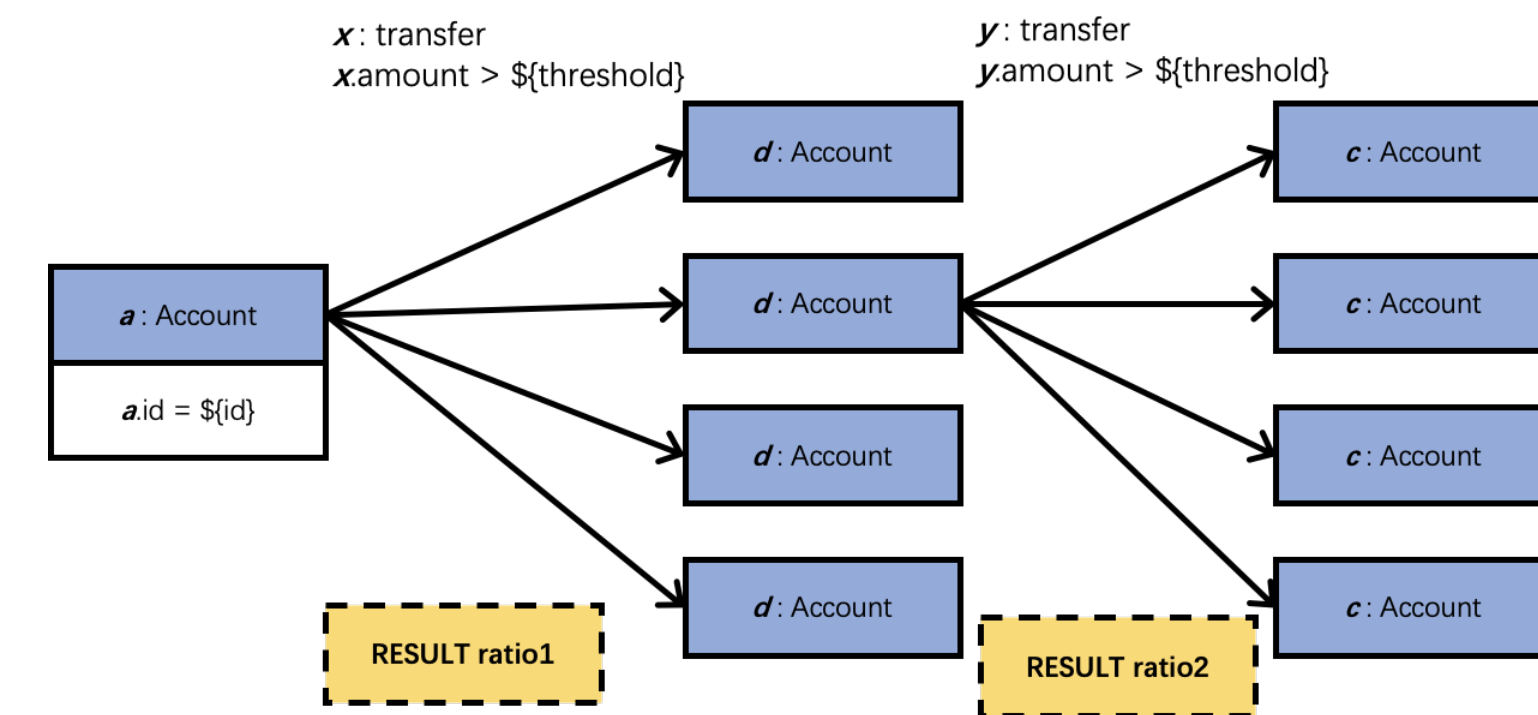
Subgraph pattern: Transfer cycle
[Ref: read / 4]



Subgraph pattern: Guarantee cycle
[Ref: read / 18]

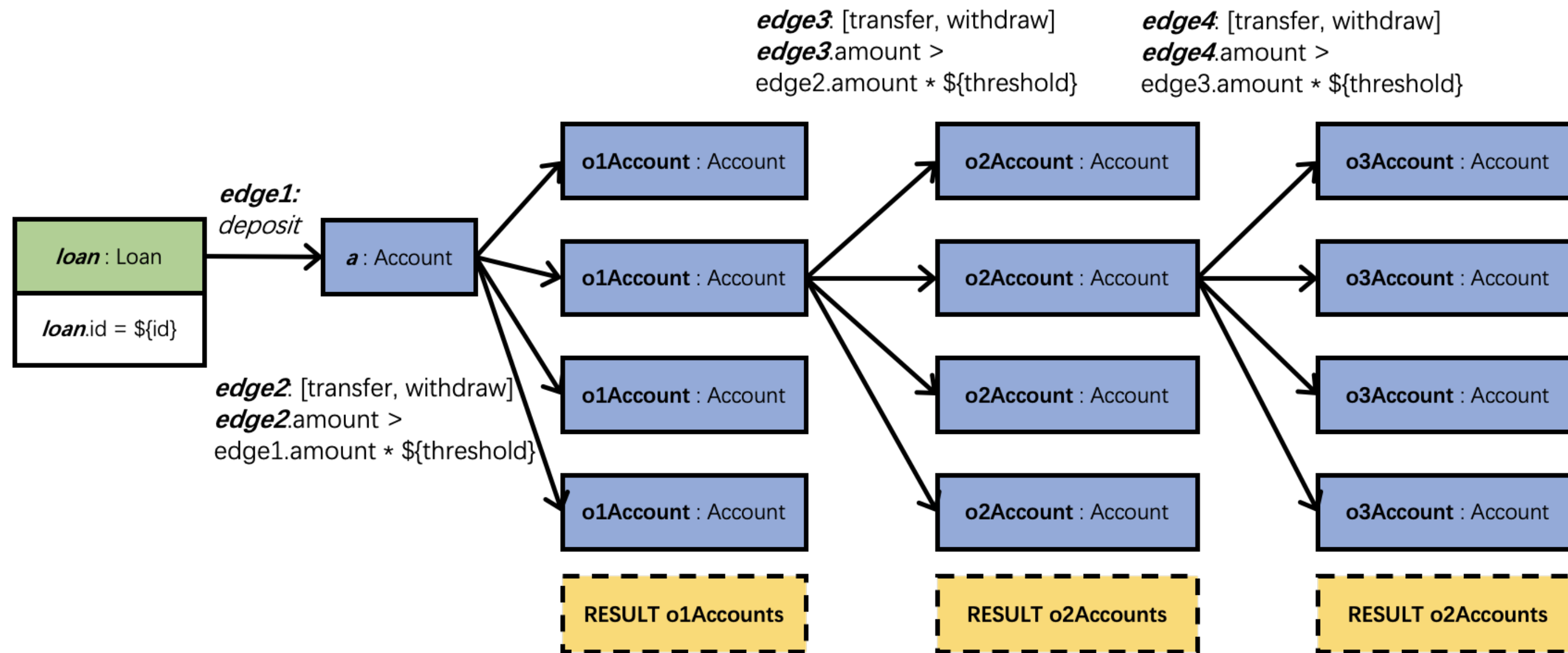


Subgraph pattern: Guarantee chains
[Ref: read / 19]



Subgraph pattern: Transfer paths
[Ref: read / 9]

FinBench Query Example: Recursive Path Filtering



Transfer trace after loan applied
(Path Filtering)
[Ref: read / 11]

Regular Path Queries

- Edges in path: deposit, transfer, withdraw -> multi-type edges
- Flexible Expression: RPQs

Recursively Path Filtering

Assuming: $A \xrightarrow{e_1} B \xrightarrow{e_2} \dots \rightarrow X$

- Timestamp order: $e_1 < e_2 < \dots < e_i$
- Amount order: $e_1 > e_2 > \dots > e_i$
- Time window: $e_{i-1} < e_i < e_{i-1} + \Delta$

3 / **FinBench Progress**

Work Charter Established

Name of task force: FinBench Task Force

Proposed or current leader: Zhihui Guo, Ant Group, guozhihui.gzh@antgroup.com

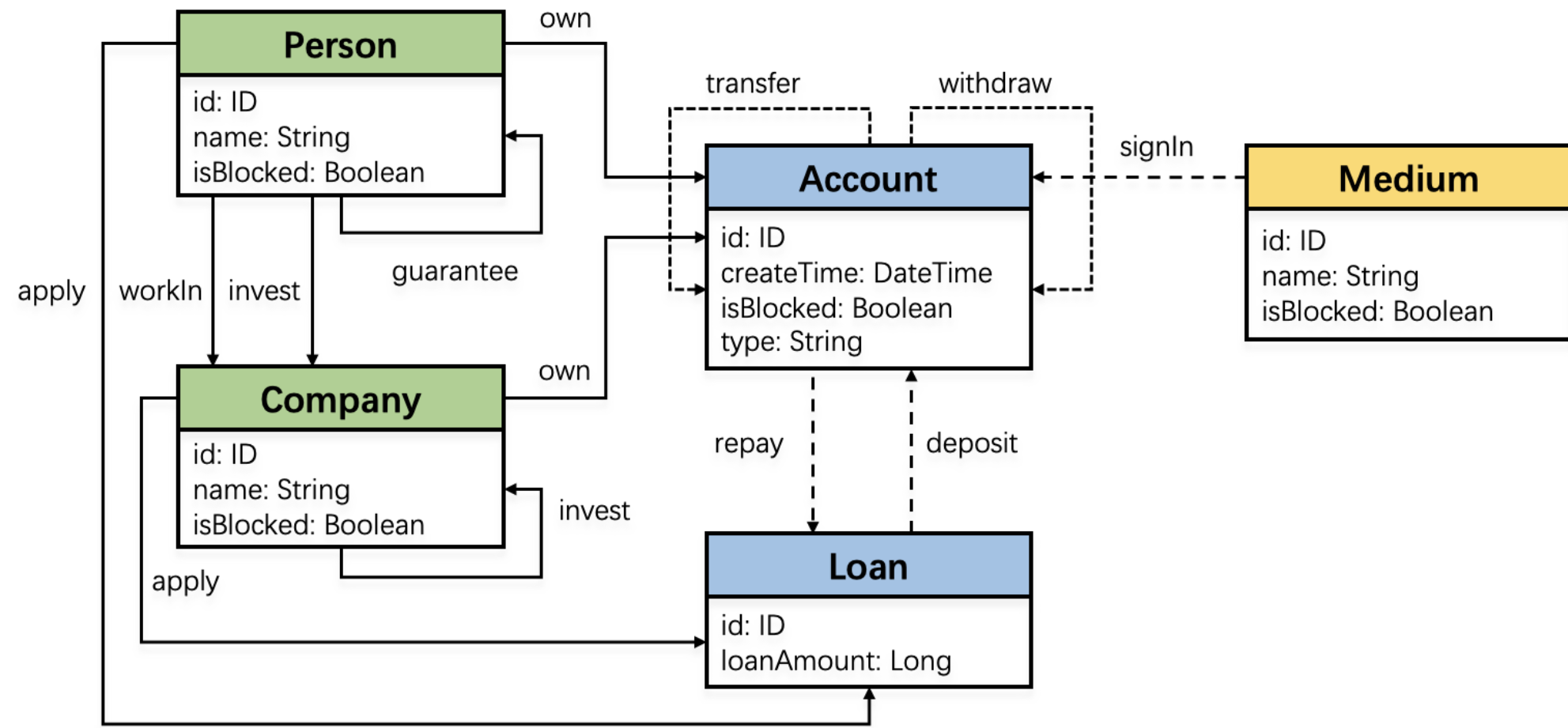
Scope of work:

- The FinBench project aims to define a graph database evaluating benchmark and develop a data generation process and a query driver to make the evaluation of the graph database representative, reliable and comparable, especially in financial scenarios.

Members:



FinBench Draft version 0.1.0-RC



Schema

Online / read / 23

query	Online / read / 23		
title	Accounts with the same transfer sources of exact account		
pattern			
desc.	Given an Account, find all the blocked accounts that connect to a third-party account which the given account has transfer-in from. Return all the accounts' id.		
params	1	id	ID id of the Account
	2	start_time	DateTime begin of the time window
	3	end_time	DateTime end of the time window
result	1	COLLECT(DISTINCT dstAccount.id)	[ID] R
sort	1	todo	↑
limit	todo		
CPs	0.0		
relevance	This query is a typical analysis for gang related accounts in risk control.		

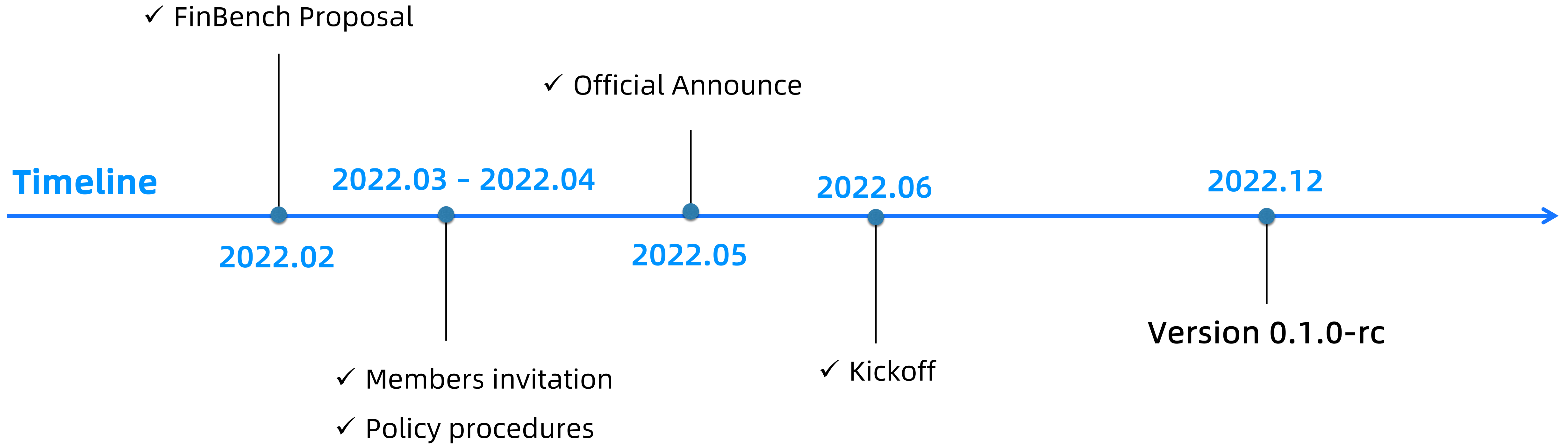
Read Query *24

Write Query *14

Read-Write Query * 5

* https://ldbcouncil.org/ldbc_finbench_docs/ldbc-finbench-specification.pdf

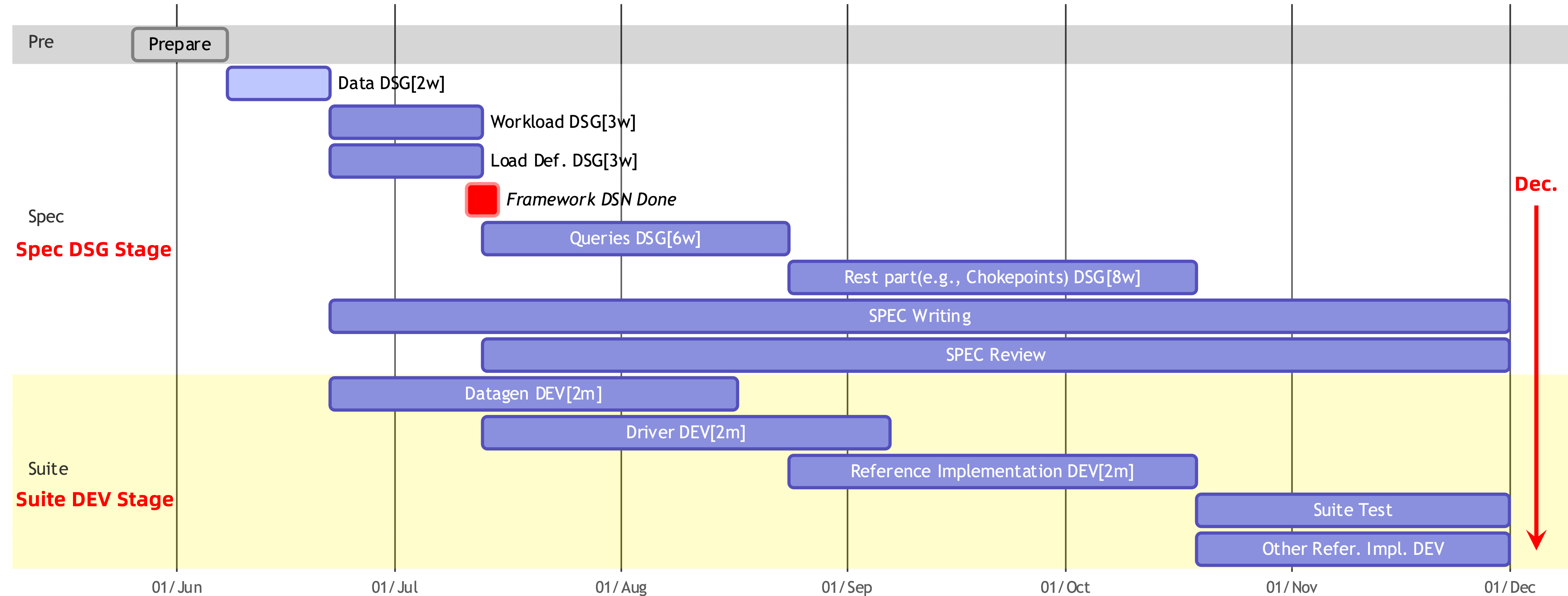
FinBench Progress



Timeline

Note:
 - DSG short for design
 - DEV short for development

FinBench v0.1.0-rc



- Version v0.1.0-rc DSN&&DEV: from Jun to Dec
- v0.1.0 approved by LDBC BoD based on v0.1.0-rc

- Spec DSN and Suite DEV works in parallel
- Online meeting for every 2-3 week at design stage

THE END

Welcome to Join us!



Heng LIN:
linheng.lin@antgroup.com



Zhihui GUO
guozhihui.gzh@antgroup.com

Useful Links:

- [1] LDBC FinBench Webpage: <https://ldbncouncil.org/benchmarks/finbench/>
- [2] LDBC Announcement: <https://ldbncouncil.org/post/announcing-the-ldbc-financial-benchmark-task-force/>
- [2] FinBench Spec on Github: https://github.com/ldbc/ldbc_finbench_docs
- [3] FinBench Initial Draft: https://ldbncouncil.org/ldbc_finbench_docs/ldbc-finbench-specification.pdf

THE END

THANK YOU!