DM32.2 Ad Hoc on SQL Extensions for Property Graphs

A brief overview/update

Jan Michels – Ad Hoc Chair
Data Management Standards
Oracle Database
jan.michels@oracle.com

July 27, 2017

Safe Harbor Statement

The following is intended to outline the general direction. It is intended for information purposes only. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making decisions. The development, release, and timing of any features, functionality, or standards described remains at the sole discretion of whoever develops or releases such features, functionality, or standards.

DM32.2 Ad Hoc – The what

- (ANSI/INCITS/)DM32.2 is the US committee responsible for standardizing SQL
  - Mirrors/tracks ISO SQL committee activities (ISO/IEC JTC1/SC32/WG3)
- Charter: Develop SQL extensions to support (querying of) property graphs
- Weekly 2-hour telecons since end of April ’17
- Participants (~20 individuals per call):
  - IBM
  - JCC Consulting
  - Microsoft
  - Neo Technology
  - Oracle
  - SAP
  - Sybase, an SAP Company
  - Teradata
  - Individual experts (3)

DM32.2 Ad Hoc – The why

- 1987: 1st ISO SQL Standard

  >> FFW >>

- December 2016: 8th Edition (aka SQL:2016) is published
- Possible work items for the next edition:
  - Better support for Big Data applications.
  - **Graph queries.**
  - Approximate queries/aggregates and uncertain data.
  - Integration of statistical packages (i.e./e.g., "R").
  - MapReduce support.
  - Streaming/continuous queries.
  - Support for blockchains.
  - BASE transactions.
  - SQL macros
  - Publish/subscribe
  - Fine-grained access control (row/column-level access control)
- June 2017: ISO made work on SQL Extensions for Property Graphs "official":
  - ISO/IEC 9075-16 SQL/Property Graph Queries (SQL/PGQ)

DM32.2 Ad Hoc – The story so far

- First few meetings allowed participants to give overviews of existing products/implementations
- Initial discussions on how to represent property graphs in tables
  - Various approaches
  - No consensus yet
  - Deferred
- Focus is shifting towards:
  - Logical property graph data model
  - Query extensions
    - Precedent in integrating/querying XML/JSON/row pattern matching

Querying property graphs

- Many open questions on various levels
- Many questions not yet discussed (or even asked), e.g.:
  - Integrate how deeply?
  - Labels
    - Required? If so, one or more per vertex/edge?
    - Do they imply a type or a set of properties?
    - Or, are they just some "special" property with special query syntax ?
  - Typing
    - SQL is strongly typed
    - Can vertices (of the same type/label) still differ wrt. type of a given property?
    - Property may be present for some vertices, absent for others
  - Etc.

DM32.2 Ad Hoc – A best case timeline*

- April 2017
  - Start of weekly DM32.2 ad hoc telecons
  - Majority of work will be done in these
- Mid-~~June~~ October
  - Maybe 1st skeleton working draft (WD)
  - Maybe some concept sections filled in
- Mid-~~October~~ February ‘18
  - 1st real WD
- ~~Mid-January~~ May 2018
  - CD-ready WD
- ~~Fall~~ December 2018
  - DIS
- ~~Spring~~ Summer 2019
  - FDIS
- ~~Summer~~ Fall 2019
  - IS

*subject to change at any time