



Санкт-Петербургский государственный университет
Кафедра системного программирования

Планирование исполнения подзапросов в PosDB

Щека Дмитрий Вадимович, группа 24.M41-мм

Научный руководитель: Г.А. Чернышев, ассистент кафедры ИАС

Санкт-Петербург
2025

Постановка задачи

- ❶ Провести теоретический обзор операторов, необходимых для поддержки подзапросов.
- ❷ Описать механизмы взаимодействия операторов для поддержки подзапросов в условиях различных стратегий материализации.
- ❸ Составить планы различных типов запросов с содержанием подзапросов, выполнение которых возможно в системе PosDB.



- PosDB [CGG⁺17] — распределенная колоночная СУБД.
- Предназначена для эффективного выполнения аналитических запросов.
- Реализует механизм материализации, позволяющий переходить от позиционного представления данных к кортежному.

Подзапросы:

- * Это запросы, вложенные в другие запросы.
- * Могут располагаться практически в любых частях внешнего запроса.
- * Позволяют улучшить простоту понимания смысла запроса.
- * Некоторые классы подзапросов сложно выполнять эффективно.

Особую сложность в эффективной реализации представляют коррелированные подзапросы. Коррелированный подзапрос — это подзапрос, который ссылается на один или несколько столбцов основного запроса. Способ эффективной реализации был предложен командой HyperDB [NLK17].

Определения необходимых операторов

$$T_1 \overline{\bowtie}_p T_2 = \{ t_1 \circ t_2 \mid t_1 \in T_1 \wedge t_2 \in T_2(t_1) \wedge p(t_1, t_2) \} \quad (1)$$

$$T_1 \overline{\bowtie}_p^1 T_2 = \begin{cases} \text{ошибка,} & \text{если } \exists t_1 \in T_1 : |\{t_1\} \bowtie_p T_2(t_1)| > 1. \\ T_1 \overline{\bowtie}_p T_2, & \text{иначе.} \end{cases} \quad (2)$$

$$T_1 \overline{\bowtie}_p^{M:m} T_2 = \{ t_1 \circ (m : (\exists t_2 \in T_2(t_1) : p(t_1, t_2)))) \mid t_1 \in T_1 \} \quad (3)$$

Зависимое соединение

select Name, Total
from Professors, Lateral (select sum(ECTS) as Total
from Courses
where Professors.PersId = Courses.Lecturer)

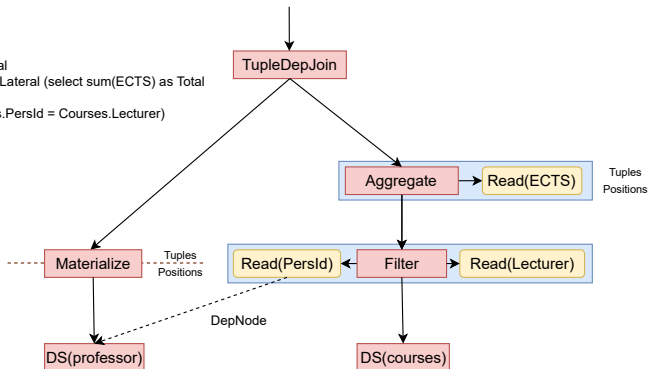


Рис.: План запроса с lateral

Одиночное соединение

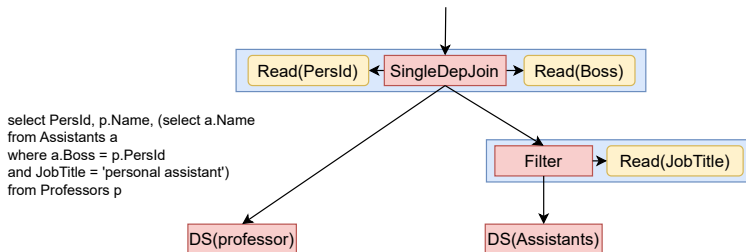


Рис.: План запроса с одиночным соединением

Соединение с маркировкой

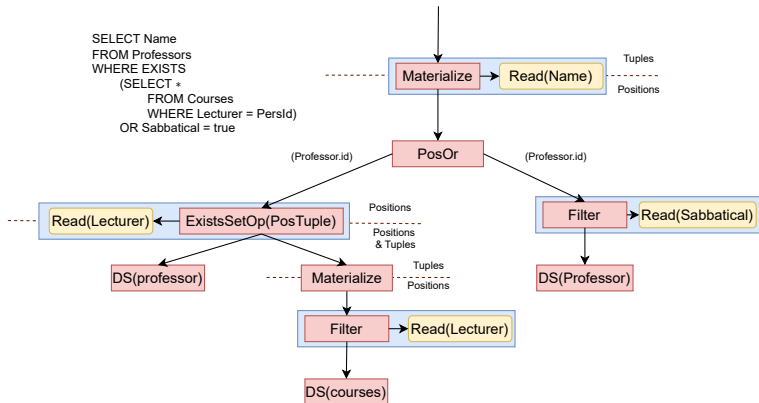


Рис.: План запроса с соединением с маркировкой

- ❶ Проведен теоретический обзор операторов, необходимых для поддержки подзапросов.
- ❷ Описаны механизмы взаимодействия операторов для поддержки подзапросов в условиях различных стратегий материализации.
- ❸ Составлены планы различных типов запросов с содержанием подзапросов, выполнение которых возможно в системе PosDB.

Список литературы



George A. Chernishev, Viacheslav Galaktionov, Valentin D. Grigorev, Evgeniy Klyuchikov, and Kirill Smirnov.

PosDB: A distributed column-store engine.

In Alexander K. Petrenko and Andrei Voronkov, editors, *Perspectives of System Informatics - 11th International Andrei P. Ershov Informatics Conference, PSI 2017, Moscow, Russia, June 27-29, 2017, Revised Selected Papers*, volume 10742 of *Lecture Notes in Computer Science*, pages 88–94. Springer, 2017.



Thomas Neumann, Viktor Leis, and Alfons Kemper.

The complete story of joins (in hyper).

In Bernhard Mitschang, Daniela Nicklas, Frank Leymann, Harald Schöning, Melanie Herschel, Jens Teubner, Theo Härder, Oliver Kopp, and Matthias Wieland, editors, *Datenbanksysteme für Business, Technologie und Web (BTW 2017), 17. Fachtagung des GI-Fachbereichs „Datenbanken und Informationssysteme“ (DBIS), 6.-10. März 2017, Stuttgart, Germany, Proceedings*, volume P-265 of *LNI*, pages 31–50. GI, 2017.