

## ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

на диссертацию Бучацкого Рубена Артуровича  
«Метод динамической компиляции SQL-запросов для реляционных СУБД»,  
представленную на соискание ученой степени кандидата технических наук по  
специальности 2.3.5 – математическое и программное обеспечение вычислительных  
систем, комплексов и компьютерных сетей

В последние годы по мере увеличения производительности, роста объема и снижения стоимости оперативной и внешней памяти производительность СУБД для некоторых классов запросов определяется непосредственно скоростью обработки запросов процессором. Для сложных запросов дальнейшее ускорение их исполнения возможно за счёт оптимизации вычислений, выполняемых на процессоре, в том числе с применением компиляторных оптимизаций. Эти обстоятельства определили цель диссертационной работы Р. А. Бучацкого – разработать и реализовать метод динамической компиляции SQL-запросов в реляционных СУБД для оптимизации выполнения запросов для более эффективного использования процессора за счёт трансформации операторов плана запроса из Volcano-модели в модель явных циклов.

Актуальность выбранной темы обусловлена тем, что для получения эффекта от применения динамической компиляции запросов необходимо также изменить и модель выполнения запроса, так как при использовании реализованной в большинстве СУБД модели Volcano не удастся получить выигрыш в производительности от динамической компиляции из-за недостатков, присущих данной модели. Разработка метода динамической компиляции запросов с трансформацией на лету операторов плана запросов из модели Volcano в модель явных циклов, более подходящую для генерации кода, является актуальной задачей.

Используя опыт разработки статических и динамических компиляторов Р. А. Бучацкий предложил и разработал метод динамической компиляции запросов, который применим к используемой в большинстве современных СУБД Volcano-модели выполнения за счёт трансформации операторов плана запроса в модель явных циклов. Описан принцип декомпозиции алгоритмов операторов модели Volcano на функции интерфейса в модель явных циклов для реализации генераторов промежуточного представления. Разработан метод динамической компиляции выражений, позволяющий избавиться от накладных расходов на интерпретацию выражений в запросах. Разработан метод кэширования для повторного использования сгенерированного динамическим компилятором кода у одинаковых запросов, тем самым нивелируя расходы на компиляцию и оптимизацию. Также разработаны эвристики для оценки предполагаемой сложности запроса, потенциальных накладных расходов на его компиляцию, и выбора на основе данных оценок одной из стратегий выполнения запроса, позволяющих использовать интерпретацию, либо динамическую компиляцию, а также принимать решение о необходимости кэширования скомпилированного кода запроса.

Все предложенные диссертантом методы были реализованы в виде программного расширения к объектно-реляционной СУБД с открытым исходным кодом PostgreSQL, которая является наиболее развитой открытой альтернативой коммерческому СУБД. Динамический компилятор реализован на базе компиляторной инфраструктуры LLVM – популярной системы для построения компиляторов. По результатам тестирования было получено ускорение в несколько раз на тестах из промышленных тестовых наборов TPC-H и TPC-DS по сравнению с интерпретатором, что показывает перспективность разработанных методов в контексте СУБД.

Р. А. Бучацкий много лет занимается разработкой компиляторов, успешно ведет научную и исследовательскую деятельность, руководит командой разработчиков. Результаты диссертации обсуждались как на российских, так и на международных конференциях. По теме работы опубликовано 7 статей, в том числе в журналах, входящих в международные системы цитирования Web of Science и Scopus. Исходный код динамического компилятора выражений опубликован в открытом доступе, а элементы его реализации были использованы при разработке JIT-компилятора выражений, реализованного сообществом разработчиков СУБД PostgreSQL. Результаты, полученные в рамках диссертационной работы, внедрены и используются российскими (ООО «РусБИТех-Астра») и зарубежными (Huawei) компаниями, а также в научно-исследовательских и учебных проектах ИСП РАН.

Считаю, что диссертационная работа соответствует всем требованиям, предъявляемым ВАК РФ к работам на соискание ученой степени кандидата технических наук по специальности 2.3.5 – математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей, а её автор, Бучацкий Рубен Артурович, заслуживает присуждения ему учёной степени кандидата технических наук.

Научный руководитель:  
директор ИСП РАН,  
академик РАН

А. И. Аветисян

5 сентября 2022 г.