

ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ МНОГОЯДЕРНЫХ СОПРОЦЕССОРОВ INTEL XEON PHI В СУБД ПРИ ВЫПОЛНЕНИИ АГРЕГАТНЫХ ФУНКЦИЙ

А.С. Порозов, П.С. Костенецкий

Южно-Уральский государственный университет

В Южно-Уральском государственном университете на суперкомпьютере «Торнадо ЮУрГУ» имеется 384 современных многоядерных сопроцессора Intel Xeon Phi, применение которых в различных областях науки не может не вызывать интерес. Параллельная обработка баз данных является одной из актуальных задач к которой возможно применение многоядерных сопроцессоров Intel Xeon Phi. На данный момент уже опубликован ряд работ [1, 2], описывающих эффективность использования многоядерных сопроцессоров Intel Xeon Phi для выполнения некоторых специфичных для СУБД алгоритмов с использованием сопроцессоров.

В данной работе описывается исследование, посвященное оценке эффективности использования многоядерных сопроцессоров Intel Xeon Phi, при параллельной обработке запросов, требующих вычисления агрегатных функций. На данный момент авторами был разработан эмулятор параллельной СУБД, выполняющий запросы агрегатных функций к базе данных. В данной статье работа эмулятора рассматривается применительно к функциям SUM, AVG, MAX, MIN, часто используемым в реальных СУБД. Эмулятор включает в себя версии запроса для CPU и для многоядерного сопроцессора Intel Xeon Phi. Алгоритмы для CPU и для Intel Xeon Phi реализованы с помощью технологии OpenMP. Работа с многоядерным сопроцессором производится в режиме offload. Время передачи данных на многоядерный сопроцессор и время, требуемое на инициализацию потоков на ЦПУ и сопроцессоре не учитывается. На Рис. 1. проведены результаты вычислительных экспериментов.

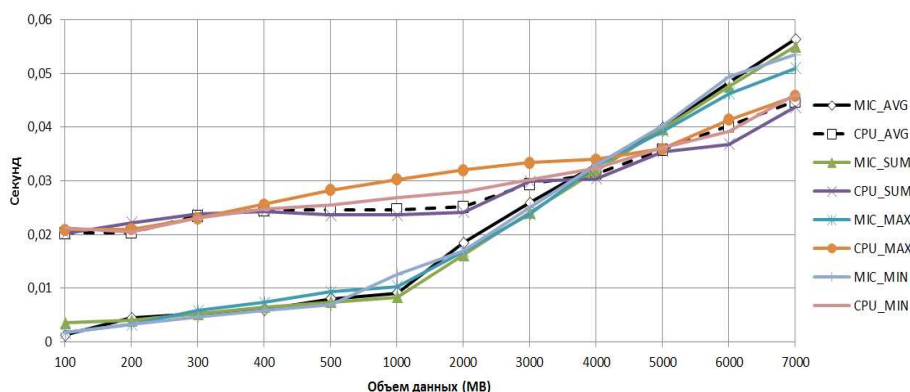


Рис. 1. Время выполнения агрегатных функций в зависимости от объема обрабатываемых данных

На основе проведенных экспериментов можно сделать вывод, что многоядерные сопроцессоры Intel Xeon Phi могут эффективно использоваться в СУБД для вычисления агрегатных функций. При размере обрабатываемого отношения базы данных менее 4 Гб сопроцессор значительно выигрывает в скорости обработки у центрального процессора. Также, во время данного исследования было выявлено, что первая передача данных на сопроцессор занимает гораздо больше времени, чем все последующие. Вероятнее всего, это связано с особенностями архитектуры памяти многоядерного сопроцессора Intel Xeon Phi.

Направлением дальнейших исследований будет реализация и тестирование разработанных [1] методов обработки данных на многоядерных сопроцессорах Intel Xeon Phi при объемах данных, превышающих размер памяти сопроцессора, путем разбиения и сжатия данных перед передачей на сопроцессор. Предполагается, что при сжатии и разбиении данных архитектура Intel MIC покажет гораздо лучшие результаты, чем ЦПУ, что позволит увеличить производительность при работе со сверхбольшими данными при выполнении запросов.

ЛИТЕРАТУРА:

1. Беседин К.Ю., Костенецкий П.С. Моделирование обработки запросов на гибридных вычислительных системах с многоядерными сопроцессорами и графическими ускорителями // Программные системы: теория и приложения: электрон. научн. журн. института программных систем им. А.К. Айламазяна РАН. 2014. Т. 5, № 1(19), с. 91–110 2014. Т. 5, № 1(19), с. 91–110
2. Kim C., Chhugani J., Satish, N. et al., Designing fast architecture-sensitive tree search on modern multicore/many-core processors, // ACM Trans. Database Syst. 2011 –Vol 36,–No 4., –P 22:1–22:34