

ОПТИМИЗАЦИЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛЕНИЙ, ПРЕДОСТАВЛЯЕМАЯ В КАЧЕСТВЕ ВЕБ-СЕРВИСА

Д.А. Пузырев, С.А. Немнюгин, А.Ю. Петров

В настоящее время производители вычислительной техники перешли на новый тип компьютерных архитектур - архитектуры, основанные на многоядерных процессорах. Это с необходимостью влечет массовый переход на многоядерные, а, по сути, параллельные архитектуры и в тех учреждениях, где активно используется вычислительная техника. Это научные учреждения, проектно-конструкторские бюро, индустрия, государственные учреждения, университеты, структуры военно-промышленного комплекса, банки, крупный, средний и малый бизнес и т. д. Общепринятой практикой стало приобретение высокопроизводительных параллельных вычислительных кластеров для решения трудоемких и критически важных задач.

Эффективное использование многоядерных и параллельных архитектур требует использования методов многопоточного и параллельного программирования. Совокупность этих методов определяет параллельную модель программирования [1, 2], которая является альтернативой традиционной, ориентированной на классическую фоннеймановскую архитектуру, последовательной модели. Параллельная модель программирования качественно отличается от последовательной, поскольку выполнение параллельной программы происходит на нескольких разных узлах вычислительной системы и требует корректного распределения вычислительной работы по процессорам/ядрам, эффективной организации обмена промежуточными и окончательными результатами вычислений или обработки информации. Существуют стандартные de facto инструменты разработки параллельных и многопоточных программ, такие как Message Passing Interface (MPI), OpenMP, POSIX Threads, Windows API и некоторые другие. Вместе с тем, подавляющая часть разработанного и используемого в настоящее время программного обеспечения разработана в рамках последовательной модели программирования.

Когда время научного или инженерного расчета может занимать несколько часов или дней, вопрос о возможности ускорении расчета не может сбрасываться со счета. В случае предоставления результатов расчета в форме веб-сервиса время ожидания становится критичным параметром, даже если речь идет о минутах. В связи с этим возникает проблема модификации (оптимизации) программного обеспечения для использования имеющегося программного обеспечения на новых (параллельных) архитектурах. Эффективная оптимизация возможна как с учётом особенностей конкретного приложения, так и с учётом архитектуры процессора (параметрическая оптимизация) [3].

Исторически сложилось так, что разработчики наукоёмкого программного обеспечения занимаются оптимизацией вычислений не в первую очередь. Это логично, поскольку основная квалификация ученых и программистов, работающих с прикладными высокопроизводительными приложениями, лежит в области их профессиональных интересов. Разработчики, как правило, также ограничены в возможности оптимизации из-за недостатка свободного времени или людских ресурсов.

Существующие технологии оптимизации программного кода представлены, во-первых, оптимизирующими компиляторами (Intel, PGI), функциональность которых достаточно ограничена, во-вторых, системами полуавтоматического распараллеливания кода (ParaWise, VAST/Parallel), которые либо устарели, либо, наоборот, недостаточно развиты на данный момент (Intel Parallel Advisor). Все эти системы являются "коробочными" решениями, привязаны к отдельным платформам и не обладают достаточной гибкостью.

В связи с этим, возникает необходимость в такой системе оптимизации программного обеспечения, которая не была бы напрямую привязана к определенной платформе, была бы достаточно гибкой для применения в различных ситуациях (различные объём и сложность кода), не требовала бы от использующего ее программиста глубоких знаний о процессе оптимизации.

Логичным решением является предоставление услуг по оптимизации в качестве веб-сервиса. В этом случае заказчик в процессе написания программы загружает в систему требуемые фрагменты исходного кода, которые обрабатываются на стороне сервера (в зависимости от объемов и сложности кода возможно использование облачных вычислений при анализе, обработке и верификации переработанного кода), затем получает обратно оптимизированный исходный код в удобном виде.

Тем самым программист получает удобный инструмент для оптимизации программного кода "на лету", без внесения кардинальных изменений в рабочий процесс.

Веб-сервисом может предоставляться как услуга автоматической и полуавтоматической оптимизации, так и ручная оптимизация кода специалистами, имеющими профессиональные навыки в нужных областях. При этом сохраняется фактор удобства использования услуги, но варьируется число доступных способов оптимизации: автоматический режим на первом этапе подразумевает анализ и распараллеливание циклов, в полуавтоматическом режиме пользователь задает системе дополнительные подсказки, позволяющие более эффективно провести распараллеливание кода, в ручном режиме специалисты могут применять все существующие способы оптимизации, вплоть до переработки алгоритмов, переноса части вычислений на

графический процессор и в перспективе, настройка связки приложение-программируемая архитектура выполнения.

ООО "ПАРСЕР" разрабатывает веб-сервис для вывода на рынок системы, включающей в себя автоматический анализатор кода и параллелизатор, основанный на технологии OpenMP, с последующим включением других способов оптимизации и совершенствованием платформы веб-сервиса. Работа поддерживается Фондом содействия развитию малых форм предприятий в научно-технической сфере по программе "СТАРТ". Авторы выражают благодарность учебно-исследовательской лаборатории СПРИНТ СПбГУ-Intel за поддержку исследований.

ЛИТЕРАТУРА

1. В.В.Воеводин, Вл.В.Воеводин "Параллельные вычисления"-СПб.: БХВ-Петербург, 2002.
2. С.А.Немнюгин, О.Л.Стефик "Параллельное программирование для многопроцессорных вычислительных систем" -СПб.: БХВ-Петербург, 2002.
3. Richard Gerber, Aart J.C. Bik, Kevin B. Smith and Xinmin Tian. The Software Optimization Cookbook, Second Edition. Intel Press, 2005