

Исследование алгоритмов планирования задач и управления потоками данных для облачных вычислительных систем

Полежаев Петр Николаевич

Оренбургский государственный университет (Оренбург), Россия

e-mail: newblackpit@mail.ru

Аннотация

В наши дни облачные вычислительные системы активно используются для развертывания высокопроизводительных систем – виртуальных вычислительных кластеров и грид-систем. Прежде всего это связано с их экономичностью – нет необходимости закупать дорогое аппаратное обеспечение, оплата производится только за фактически потребленные вычислительные и сетевые ресурсы. Также отсутствуют затраты на обновление системы, замену выходящего из строя оборудования.

Существующие облачные системы, такие как OpenStack, Eucalyptus и др. не обеспечивают должного уровня производительности для всей облачной системы в целом. Прежде всего, это связано с использованием неэффективных алгоритмов планирования задач. Другим основным недостатком является отсутствие встроенных средств планирования сетевого трафика между процессами вычислительных задач. В данном проекте предложены алгоритмы планирования задач для облачной системы с учетом многопроцессорности и виртуализации вычислительных узлов, топологии системы, сетевой конкуренции и существующих коммуникационных схем исполняемых задач. Основная идея алгоритмов – использование данной информации для компактного размещения процессов вычислительных задач с целью снижения сетевой конкуренции.

Проведено их экспериментальное исследование с помощью разработанного симулятора облачной системы, которое показало их эффективность для различных характеристик потока задач.

Также предложен подход к управлению потоками данных между процессами вычислительных задач, основанный на применении программно-конфигурируемой сети (ПКС) в качестве физической сети для облачной системы. Основная его идея – передача планировщиком коммуникационной схемы вычислительной задачи и схемы отображения процессов на узлы контроллеру ПКС, что позволяет ему использовать ее для прокладки маршрутов передачи данных между каждой задействованной парой процессов еще до запуска задачи.

Экспериментальное исследование показало, что совместное использование разработанных алгоритмов планирования задач и предложенных методов управления потоками данных между ее процессами с использованием ПКС обеспечивает снижение времени выполнения потока задач до 7%. Исследования выполнены при поддержке Министерства образования и науки Российской Федерации в рамках проекта № 07.514.11.4153 ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы» и РФФИ (проекты №12-07-31089 и № 13-07-97046).