

Текущие возможности и перспективы развития кластерного планировщика CSched

Р.А. Ахметшин, Е.А. Ермолаев, А.Л. Штангеев, А.В. Юлдашев

Уфимский государственный авиационный технический университет

Эффективная эксплуатация суперкомпьютерной техники невозможна без использования развитого системного программного обеспечения, предназначенного для предоставления интерфейса взаимодействия с суперкомпьютером, управления, планирования, мониторинга выполнения вычислительных задач и т.д. В Уфимском государственном авиационном техническом университете (УГАТУ) для этого используется как стороннее системное программное обеспечение, так и собственные разработки. К последним, в частности, относятся «Программный комплекс автоматизированных расчетов на кластерных системах CCS», а также «Кластерный планировщик задач CSched».

CSched предназначен для планирования выполнения задач на кластерных вычислительных системах в связке с системой очередей TORQUE. Планировщик поддерживает учет различных ресурсов вычислительных узлов (вычислительных ядер, оперативной памяти и т.д.), а также плавающих ресурсов, например, сетевых лицензий на программное обеспечение, что позволяет использовать его при проведении расчетов средствами коммерческих пакетов компьютерного моделирования, в том числе ANSYS, SIMULIA Abaqus, Schlumberger Eclipse, Roxar Tempest More и т.п. Поддерживается настройка политик планирования для различных классов задач, в качестве которых могут выступать упомянутые выше пакеты.

Одной из основных целей разработки собственного планировщика являлось повышение производительности вычислений за счет использования алгоритмов планирования, позволяющих минимизировать задержки при конкурентном доступе к общим ресурсам кластерной системы: оперативной памяти и коммуникационной среде [1]. Необходимо отметить, что в известных планировщиках подобные алгоритмы не реализованы. Единственной системой, в которой произведена попытка учесть влияние конкуренции при доступе к оперативной памяти, является Platform LSF, где имеется поддержка требований задач к пропускной способности памяти в виде ресурсного запроса.

Наш подход предполагает выполнение автоматической априорной оценки характеристик поступающих задач, мониторинг загрузки общих ресурсов многоядерных узлов в динамике работы кластерной системы и накопление статистики использования общих ресурсов, а также возможность оценки влияния конкуренции за общие ресурсы на время выполнения задач для более эффективного планирования, что в альтернативных планировщиках не поддерживается. В то же время проведенные экспериментальные исследования на суперкомпьютере УГАТУ подтверждают, что планирование с учетом конкуренции при доступе к оперативной памяти многоядерных узлов позволяет получить ускорение при решении реальных задач около 1,5 раз.

В настоящее время ведутся работы по созданию специализированного модуля для планировщика CSched, который позволит проводить планирование с учетом конкуренции за общие ресурсы многоядерных узлов. Кроме этого ведется реализация алгоритмов справедливого распределения ресурсов между пользователями кластерной системы (fairshare [2]), что позволит обеспечить высокий уровень доступности кластерной системы для всех ее пользователей.

Литература

1. Юлдашев А.В. Минимизация времени выполнения MPI-программ с учетом конкуренции за каналы передачи данных коммуникационной среды кластерной системы // Вестник УГАТУ, 2011. – Т.15. – №2 (42). – С. 99-105.
2. Jackson D., Snell Q., Clement M. Core Algorithms of the Maui Scheduler. URL: <http://www.eecs.harvard.edu/~chaki/bib/papers/jackson01maui.pdf> (дата обращения: 02.12.2012).