

Индекс УДК: 004.048

Наименование: “Исследование обратной задачи планирования композитных приложений для оценки производительности алгоритмов в зависимости от структуры композитных приложений”

Исполнитель: Говоруха Оксана Сергеевна, Университет ИТМО, г. Санкт-Петербург

Научный руководитель: Мельник Михаил Алексеевич, Университет ИТМО, г. Санкт-Петербург

Основная часть тезиса:

Для решения сложных научных задач активно применяются распределенные вычислительные системы и облачные сервисы. При проведении сложных расчетов в таких средах часто используются приложения, состоящие из набора различных задач и зависимостей между ними, - композитные приложения или workflow. Распределение таких задач по вычислительным ресурсам и составление расписания их выполнения для минимизирования общего времени выполнения приложения, представляет собой задачу планирования.

Задача планирования композитных приложений относится к классу NP-трудных задач. На данный момент предложено большое количество алгоритмов для решения этой задачи в распределенных гетерогенных и облачных средах. Существующие алгоритмы можно разделить на три основные группы: эвристические, мета-эвристические и гибридные. Задача планирования может включать дополнительные условия, такие как ограничение времени выполнения или ограничение бюджета. При выполнении композитных приложений, используя облачные сервисы, важно учитывать стоимость ресурсов при составлении расписания и минимизировать не только общее время, но и затраты на выполнение приложения. Со временем появились алгоритмы планирования, оптимизирующие несколько целевых функций и учитывающие заданные ограничения при составлении расписания. Множество исследований посвящено анализу и сравнению алгоритмов планирования, однако на данный момент недостаточно изучено влияние входных параметров, таких как структура workflow, и характеристик вычислительной среды на работу алгоритма и эффективность составления расписания.

Цель данной работы исследовать производительность алгоритмов планирования композитных приложений в зависимости от топологических особенностей структуры workflow, а также характеристик и состояния вычислительной среды и нагрузки, что позволит в дальнейшем вносить изменения в структуру композитного приложения и настраивать определенные характеристики среды, для улучшения эффективности составления расписания конкретным алгоритмом.

Для оценки производительности алгоритма выделяют определенные метрики, в зависимости от критериев планирования. Заданные критерии влияют на применяемый подход и техники, используемые для планирования композитных приложений. В облачной среде ключевыми критериями являются общее время выполнения расписания, стоимость, надежность работы алгоритма, энергозатратность используемых ресурсов и безопасность. Для анализа алгоритма и его производительности могут быть использованы следующие метрики: makespan, отображающее время окончания выполнения приложения, согласно составленному расписанию, также применяется нормализованная метрика - SLR. Для оценки стоимости используют метрику - MCR, отображающую затраты на использованные ресурсы и передачу данных при выполнении задач, которые рассчитываются на основе ценовой модели облачного сервиса. Также часто для оценки алгоритма, оптимизирующего несколько заданных целевых функций, вводят дополнительную метрику, включающую makespan и стоимость, с заданными весами. Для анализа надежности алгоритма и эффективности использования им ресурсов,

могут применяться такие метрики: GR, определяющее процент задач, закончивших выполнение до установленного дедлайна, среди всех распределенных задач, NAN и NAT, отображающие количество активных ресурсов в течении выполнения задачи и общее время работы ресурсов, соответственно, и RTA, представляющее собой отношение общего времени выполнения задачи ко времени активной работы ресурсов.

Для тестирования алгоритмов и получения значений метрик, используется симулятор инфраструктуры реальной системы для выполнения планирования композитных приложений. Использование симулятора позволяет варьировать конфигурации состояния среды и нагрузку. Предлагаемая модель для оценки производительности алгоритма планирования в зависимости от структуры композитного приложения и других внешних параметров выглядит следующим образом: в качестве входных параметров симулятору передается сгенерированное или реальное состояние структуры workflow, определенный алгоритм, производительность которого необходимо оценить, задаются характеристики вычислительной среды и передается список измеряемых метрик. Получаемые данные представляют собой показатели метрик, оценивающие эффективность составляемого расписания конкретным алгоритмом для разных структур композитного приложения при заданном состоянии вычислительной среды. Предполагается также интегрирование генетического алгоритма для изменения структуры workflow, согласно особенностям данного алгоритма планирования, чтобы улучшить его производительность.

Таким образом, в данной работе проанализированы существующие критерии и метрики для оценки производительности алгоритмов планирования композитных приложений и разработана методология для оценки производительности алгоритма в зависимости от входных данных, таких как структура композитного приложения, и характеристик вычислительной среды.