

РАЗРАБОТКА МОДЕЛЕЙ ОБЛАЧНОЙ СИСТЕМЫ, ПОСТРОЕННОЙ НА БАЗЕ ТЕХНОЛОГИИ ПРОГРАММНО-КОНФИГУРИРУЕМЫХ СЕТЕЙ

Л. С. Адрова, П. Н. Полежаев

Кафедра компьютерной безопасности и математического обеспечения информационных систем, центр информационных технологий, кафедра геометрии и компьютерных наук, Оренбургский государственный университет

Оренбург, Российская Федерация

E-mail: {newblackpit, kroshic555}@mail.ru

В работе описываются созданные модели облачной системы, включая структурную и имитационную модели самоорганизующейся гетерогенной облачной платформы, модели облачного сервиса и приложения. Они лягут в основу разрабатываемых алгоритмических решений, обеспечивающих эффективное функционирование облачных систем с программно-конфигурируемыми сетями и инфраструктурами.

ВВЕДЕНИЕ

В настоящее время существующие решения для интеграции распределенных центров обработки данных в единую вычислительную инфраструктуру на основе грид-платформ (Oracle Grid Engine, Condor-G, Globus Toolkit, HPC-NASIS и др.) обладают целым рядом недостатков, включая: недостаточно высокую производительность для задач, требующих интенсивного обмена данными; сложность конфигурирования и обеспечения безопасности; невозможность четкого выделения сетевых ресурсов для выполнения конкретных сетевых операций; сложность в адаптации грид-систем под новые вычислительные парадигмы, например, "map-reduce".

Существующие облачные платформы для научных вычислений (AWS HPC, Penguin Computing On-Demand, IBM Platform Computing, SURFsara, Sabalcore HPC Cloud) имеют ряд преимуществ по сравнению с грид-системами: упрощение конфигурирования облачных ресурсов и приложений через портал самообслуживания или API; виртуализацию вычислительных ресурсов; возможность создания облачных приложений, использующих виртуальные вычислительные сервисы для эффективного выполнения вычислительных алгоритмов и решения стандартных подзадач; возможность использования различных вычислительных парадигм; значительную масштабируемость и гибкость - возможность, в зависимости от загрузки, значительно увеличить или сократить количество виртуальных машин.

Однако, облачные платформы для научных вычислений также обладают рядом недостатков: сложность перемещения данных, в особенности между облачными платформами различных производителей; сложность обеспечения их взаимодействия; отсутствие сведений о топологии облачной системы для выполняющихся приложений и их пользователей, что делает невозмож-

ным предсказание характеристик сети с течением времени.

Отдельно следует заметить, что существующие грид и облачные платформы для распределенных вычислений ориентированы в основном на планирование вычислительных задач без учета подстройки сети под их коммуникационные схемы, под загрузку и сохранение исходных, промежуточных и результирующих данных. Таким образом можно сформулировать следующие основные задачи, которые планируется решить в рамках данного исследования:

1. Разработка эффективных алгоритмов планирования виртуальных машин, облачных приложений, размещения и организации доступа к данным гетерогенной облачной платформы, ориентированных на учет топологии вычислительной системы, коммуникационных схем заданий, расположение и характер обрабатываемых входных, выходных и промежуточных данных, а также на учет гетерогенного характера облачных платформ.
2. Разработка эффективных методов адаптивной маршрутизации сетевых коммуникаций облачных приложений с учетом коммуникационных схем заданий, наличия нескольких автономных систем, связанных через сеть Интернет, а также требований к качеству обслуживания (QoS, Quality of Service).

Для их решения необходима разработка ряда моделей, которые будут далее описаны.

I. СТРУКТУРНАЯ МОДЕЛЬ САМООРГАНИЗУЮЩЕЙСЯ ГЕТЕРОГЕННОЙ ОБЛАЧНОЙ ПЛАТФОРМЫ

Данная модель описывается в виде взвешенного ориентированного мультиграфа. Вершины - автономные системы, дуги - сетевые связи между ними. Веса дуг - характеристики соответствующих каналов передачи между автоном-

ными системами (пропускные способности, задержки).

Каждая автономная система, в свою очередь, описывается в виде взвешенного ориентированного мультиграфа, в котором вершинами являются все устройства сети, включая сервера, виртуальные машины, запущенные на серверах, граничные шлюзы, коммутаторы и контроллеры программно-конфигурируемой сети [1], облачные контроллеры, СХД, точки входа в VPN и VXLAN туннели и др. Дуги представляют сетевые соединения между ними. В реальных сетях между двумя физическими устройствами может быть несколько параллельных сетевых соединений, поэтому две вершины могут быть соединены параллельными дугами. Точкой подключения дуги к узлу является сетевой порт. Весами вершин являются статические параметры и динамические характеристики соответствующих устройств сети, а весами дуг - пропускные способности, задержки, проходящие через них маршруты и т.п.

Структурная модель позволяет описать статические и динамические параметры современных распределенных гетерогенных облачных платформ. Сведения для структурной модели берутся из данных, получаемых алгоритмом обнаружения топологии сети и ее состояния [2].

II. МОДЕЛЬ ОБЛАЧНОГО ПРИЛОЖЕНИЯ

Облачное приложение представляет собой взвешенный ориентированный ациклический граф зависимостей по данным, в котором вершинами являются задачи, сервера облачной БД, СХД и прочие ресурсы хранения и кэширования данных, дуги - зависимости по данным между соответствующими вершинами. Каждая вершина-задача характеризуется предъявляемыми ресурсными требованиями (к количеству ядер, архитектуре команд ядер, размерам оперативной и дисковой памяти, наличию специальных библиотек или оборудования на физических или виртуальных узлах, используемых для запуска процессов), количеством запускаемых процессов, оценкой времени выполнения, коммуникационными шаблонами передачи данных между процессами. Каждая дуга характеризуется типом доступа (доступ к файлу в СХД, к локальному файлу, к распределенной БД, к сервису данных и пр.), оценкой объемов передаваемых данных, требованиями к QoS.

III. МОДЕЛЬ ОБЛАЧНОГО СЕРВИСА

Облачный сервис, как и облачное приложение, описывается в виде ориентированного графа зависимостей по данным, отличие заключается в том, что с точки зрения пользователя облачный сервис представляет собой закрытую систему. Также все его приложения распределены на предустановленном наборе виртуальных машин или физических серверов, новые их

экземпляры масштабируются динамически в зависимости от количества поступающих запросов на выполнение функции конкретного сервиса со стороны облачного приложения, конечного пользователя или других облачных сервисов. Данная модель, как и предыдущая, используется алгоритмом планирования распределения облачных сервисов и приложений на выделенные виртуальные машины и физические сервера гетерогенной облачной платформы, а также алгоритмом планирования размещения и организации доступа к данным облачных сервисов и приложений.

IV. ИМИТАЦИОННАЯ МОДЕЛЬ ГЕТЕРОГЕННОЙ ОБЛАЧНОЙ ПЛАТФОРМЫ

Имитационная модель гетерогенной облачной платформы. Данная модель основана на структурной модели и включает в себя подобранные законы распределения случайных величин для параметров заявок пользователей на запуск облачных приложений и сервисов, процессов, возникающих при их работе, включая изменения загруженности вычислительных ресурсов, передачу сетевых пакетов, работу программно-конфигурируемой инфраструктуры и пр. Также модель содержит имитационную схему функционирования распределенной гетерогенной облачной платформы, отражающую принципы работы всех облачных ресурсов и сетевых связей. Данная модель ляжет в основу разрабатываемого симулятора гетерогенной облачной платформы. Отличительная черта данной модели - использование первых двух моментов распределения времени между пакетами, что позволяет аппроксимировать законы распределения времени с помощью метода двумерной диффузионной аппроксимации.

ЗАКЛЮЧЕНИЕ

В рамках данной работы описаны основные модели облачной системы, которые станут основой для разрабатываемых алгоритмических решений, обеспечивающих эффективное функционирование облачных систем с программно-конфигурируемыми сетями и инфраструктурами.

Исследования выполнены при поддержке Президента Российской Федерации, стипендии для молодых ученых и аспирантов (СП-2179.2015.5).

1. McKeown, N. Openflow: enabling innovation in campus networks / N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, J. Turner // ACM SIGCOMM Computer Communication Review. – 2008. – Vol. 38. – PP. 69–74.
2. Polezhaev, P. Network Resource Control System for HPC based on SDN / P. Polezhaev, A. Shukhman, Yu. Ushakov // Proceedings of 14th International Conference, NEW2AN 2014 and 7th Conference ruSMART 2014, St. Petersburg, Russia. Lecture Notes in Computer Science – vol. 8638. – PP. 219–230.