

аспирант Кобицкая Ю.А.,
аспирант Иванисенко И.Н.
Харьковский национальный университет радиоэлектроники
e-mail: KobitskayaJulya@gmail.com
e-mail: Ivanisenko79@yahoo.com

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ОБЛАЧНЫХ СЕРВИСОВ С УЧЕТОМ САМОПОДОБНЫХ СВОЙСТВ ВХОДЯЩИХ ПОТОКОВ

В работе предложен подход к анализу и балансировке нагрузки для облачных вычислений, который основан на имитационном моделировании и учитывает самоподобные свойства поступающих на сервер запросов клиентов.

В настоящее время технологии «облачных» вычислений приобретают все большую популярность, а концепция Cloud Computing является одной из самых модных и актуальных тенденций развития информационных технологий. Под облачными вычислениями обычно понимают возможность получения необходимых вычислительных мощностей по запросу из сети, причем пользователю не важны детали реализации этого механизма и он получает из этого «облака» все необходимое.

При проектировании и создании облачных решений важное место занимает обеспечение надежности и оптимальности использования ресурсов – мониторинг, или получение информации о доступности и загрузке аппаратных ресурсов платформ с копиями распределенного приложения, и балансировка, то есть распределение поступающих пользовательских запросов между имеющимися аппаратными платформами. Одной из проблем облачных вычислений является неравномерность запроса ресурсов со стороны клиентов. [1]

Целью представленной работы является разработка подхода к анализу и балансировке нагрузки для облачных вычислений, который учитывает характерные особенности поступающих на сервер запросов клиентов.

Информационные потоки в облачных сервисах. Динамическая природа облаков приводит к формированию новой картины трафика, отличающейся от той, что наблюдается в обычных телекоммуникационных сетях. В различных облачных сетях свои особенности передачи данных, но трафик, формирующийся между центрами обработки данных (ЦОД) всегда складывается из двух основных компонентов:

– статичный компонент – устойчивая, предсказуемая часть трафика, которая соответствует главным образом неизменяющимся взаимодействиям внутри ЦОД или соединениям, идущим из облака или в облако;

– постоянно меняющийся, динамичный компонент, который менее предсказуем в отношении отправителя/получателя потока данных, его продолжительности и объема.

В некоторых крупных облачных сетях динамичный компонент трафика может превалировать над статичным, меняясь в течение часа или на протяжении дня. Основными причинами этих изменений являются масштабная миграция баз данных, значительное разовое повышение интенсивности трафика, связанное с конкретным событием, подключение новых клиентов и т. п. В этом случае необходимы инфраструктурные решения, способные оперативно

реагировать на изменения и предоставлять «эластичную» полосу пропускания в определенной точке сети в определенное время.

Самоподобие сетевых трафиков. Многочисленные исследования процессов в информационных сетях показали, что реализации сетевого трафика обладают свойством масштабной $X(t)$ инвариантности (самоподобием). Стохастический процесс $X(t)$ является статистически самоподобным с параметром

самоподобия H (показателем $a^{-1}X$ at Херста), если процесс обладает теми же статистическими характеристиками второго порядка, что и $X(t)$.

Одним из важнейших свойств трафика, как случайного процесса, является наличие «тяжелых хвостов» его функций распределений. Самоподобный трафик имеет особую структуру, сохраняющуюся на многих масштабах – в реализации всегда присутствует некоторое количество очень больших выбросов при относительно небольшом среднем уровне трафика. Эти выбросы вызывают значительные задержки и потери пакетов, даже когда суммарная потребность всех потоков далека от максимально допустимых значений.

Самоподобные свойства обнаружены в локальных и глобальных сетях, в частности в трафиках Ethernet, ATM, приложений TCP, IP, VoIP и видеопотоках. Причина такого эффекта заключается в особенностях распределения файлов по серверам, их размерах, в типичном поведении пользователей. Оказалось, что изначально не проявляющие свойств самоподобия потоки данных, пройдя обработку на узловых серверах и активных сетевых элементах, начинают подавать ярко выраженные признаки самоподобия. [2]

Наличие у передаваемых клиентами информационных потоков свойств самоподобия оказывает большое влияние на эффективность работы облачных сервисов. Особенно важную роль это играет для работы сервисов, обеспечивающих передачу мультимедийного трафика и трафика реального времени.

Имитационное моделирование облачных вычислений. Основным инструментом исследования и прогнозирования поведения самоподобных потоков данных является имитационное моделирование, для которого необходима модель входной самоподобной нагрузки.

Существует множество моделей самоподобного трафика [2]. В работе [3] предложена модель агрегированного самоподобного потока, учитывающая степень самоподобия и «тяжесть хвоста» функции распределения. Параметрами модели являются интенсивность трафика, показатель Херста и коэффициент вариации, большие значения которого соответствуют большим выбросам в реализации. Данная модель была использована для проведения численного моделирования.

Современные информационные сети построены на основе мультиплексирования потоков данных. Согласно классической теории массового обслуживания, множество потоков данных со случайными вариациями распределений вероятностей дадут в результате некий усредненный сглаженный трафик. Однако этот подход не применим к потокам данным, обладающим свойствами самоподобия. Особое значение это приобретает при построении облачных инфраструктур, где важной задачей является оптимальное распределение нагрузки между компонентами.

В работе показано, что длина очереди в буфере и величина потерь в коммутационном узле сети определяются значениями параметров Херста и коэффициентом вариации объединенного потока данных. Численные исследования продемонстрировали, что суммарный самоподобный поток приобретает свойства потока с максимальным показателем Херста. Кроме этого мультиплексированный трафик сохраняет пиковые значения выбросов.

При распределении нагрузки на узлы сети необходимо перенаправлять объединенные потоки данных, принимая во внимание не только суммарную интенсивность трафика, но и показатели самоподобия и всплесков.

Симуляторы распределенных вычислений не могут в достаточной степени моделировать облачную структуру и в настоящее время появились приложения имитации работы облачных серверов [4].

Рассмотрим основные симуляторы облачных сервисов.

– CloudSim –новый, обобщенный и расширяемый инструментарий моделирования, который позволяет симулировать и проводить опыты над развивающимися системами облачных вычислений, инфраструктурами и прикладными средами для одиночных и межсетевых облачных серверов. CloudSim представляет собой веб-приложение с открытым исходным кодом, которое запускает предварительно настроенные распространенные проекты с открытым исходным кодом, связанных с робототехникой, например, симулятор Gazebo.

– CDOSim – имитирует время отклика, нарушения SLA (Service Level Agreement) и затраты на CDO (Collateralized debt obligations). CDO - это технологии, касающиеся симулятора, который принимает решения о выборе поставщика облака, конкретных стратегиях адаптации выполнения, развертывания компонентов виртуальной машины и конфигурации ее экземпляров. Развертывание компонентов на виртуальную машину включает в себя возможность формирования новых компонентов из уже существующих. CDOSim имеет способность представлять точку зрения пользователя, а не провайдера.

– TeachCloud – это облачный симулятор, который создан специально для образовательных целей. TeachCloud предоставляет простой графический интерфейс, через который студенты и ученые могут изменять конфигурацию в облаке и выполнять простые опыты. TeachCloud использует CloudSim как основную платформу.

– iCanCloud – облачный симулятор, который основан на Simcan. iCanCloud является основой программного обеспечения для моделирования больших сетей хранения данных. iCanCloud может предложить компромисс между затратами и производительностью конкретного приложения на конкретном оборудовании с целью информирования пользователей о связанных с этим расходах. Он имеет полный графический интерфейс пользователя, с помощью которого могут быть разработаны и запущены различные эксперименты, однако, существующие программные системы можно моделировать только вручную. Он также позволяет параллельное выполнение одного эксперимента на нескольких машинах.

– SPECI (Simulation Program for Elastic Cloud Infrastructures) является инструментом моделирования, который позволяет анализировать и исследовать свойства масштабирования поведения большого ЦОД, используя политику размеров и дизайна промежуточного программного обеспечения как входные данные. Это инструмент моделирования, который дает возможность изучить аспекты масштабирования, а также эксплуатационные свойства будущих ЦОД. SPECI состоит из двух пакетов: макета ЦОД и его топологии и пакета компонентов для выполнения экспериментов и измерений.

– GroudSim – симулятор на основе событий, который необходим при моделировании для научных приложений в сетевых и облачных средах на основе ядра дискретных событий, независимого от масштабируемого моделирования. Он в основном сосредоточен на IaaS (Infrastructure as a Service), но в то же время легко расширяется для поддержки дополнительных моделей, таких как PaaS (Platform as a Service), DaaS (Desktop as a Service) и TaaS (Testing as a Service). Пользователь может моделировать свои эксперименты в той же среде, используемой для реальных приложений за счет интеграции GroudSim в среду Askalon. GroudSim предоставляет полный набор функций для сложных сценариев моделирования, таких как выполнение работы на арендованных вычислительных ресурсах, расчета затрат и фоновой нагрузки на ресурсы.

– DCSim (DataCenter Simulator) сосредоточен на виртуализированном ЦОД, который распределяет IaaS на несколько арендаторов в целях достижения симулятором оценки и разработки методов управления ЦОД.

В дальнейшей работе предполагается с помощью имитационного моделирования на симуляторе облачных вычислений исследовать свойства мультиплексированных самоподобных потоков, поступающих на облачные сервисы и, на основе полученных результатов, разработать подход к балансировке нагрузки, который учитывает самоподобные свойства поступающих на сервер потоков данных.

1. Klement'ev I.P. Introduction to Cloud Computing / I.P. Klement'ev, V.A. Ustinov // UGU. – 2007. – 233 p.

2. Sheluchin O. I. Self-Similar Processes in Telecommunications / O. I. Sheluchin, S. M. Smolskiy, A. V. Osin // New York : John Wiley & Sons. – 2007. – 320 p.

3. Kirichenko L. Mathematical simulation of self-similar network traffic with aimed parameters Anale / L.Kirichenko, T.Radivilova, Abed Saif Alghawli // Seria Informatică. Vol. XI fasc. 1 – 2013. P.17-22.

4. Malhotra R. Study and Comparison of Various Cloud Simulators Available in the Cloud Computing / Dr. R. Malhotra, Prince Jain // International Journal of Advanced Research in Computer Science and Software Engineering. Vol. 3. Issue 9. 2013. P. 347-350.