

ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ ХАРАКТЕРИСТИК ПРИКЛАДНЫХ СЕТЕВЫХ СЕРВИСОВ

В.О. Писковский

*Московский государственный университет им М.В. Ломоносова, Факультет вычислительной
математики и кибернетики*

Россия, 119991, ГСП-1 Москва, Ленинские горы, МГУ имени М.В. Ломоносова, д.1, стр. 52, 2-й учебный
корпус, факультет ВМК

E-mail: vpiskovski@lvk.cs.msu.ru

А.А. Грушо

Федеральный исследовательский центр «Информатика и управление» Российской академии наук

Россия, 119333, Москва, ул. Вавилова, д. 44, кор. 2

E-mail: grusho@yandex.ru

Ключевые слова: прогнозирование характеристик выполнения задач, распределенные прикладные сетевые сервисы, бессерверные вычисления, туманные вычисления, соглашение об уровне обслуживания, технология виртуализации, случайный лес, матричные разложения, сверточные нейронные сети, рекомендательные системы, технологии виртуализации.

Аннотация: В настоящее время в развитии современных вычислительных систем наблюдаются тенденции к повышению их функциональности за счёт увеличения числа аппаратных и виртуальных ресурсов, привлекаемых для обеспечения распределенной обработки данных, с которыми этим системам приходится работать.

Технологии виртуализации, распределенные облачные вычисления приводят к возможности использования вычислительных мощностей транспортных сетей и сетей доступа, что, в свою очередь, делают необходимыми анализ и прогнозирование оптимального использования сетевых ресурсов.

1. Введение

С каждым годом вычислительные системы всё больше проникают во все сферы человеческой деятельности и отрасли промышленности. Компьютерная сеть, населенная множеством сервисов и услуг, стала неотъемлемой частью информационных и автоматизированных систем. Современные вычислительные системы увеличивают свою функциональность за счёт повышения числа используемых аппаратных модулей и компонентов, сложности программного обеспечения для своевременной обработки растущего объёма данных, в том числе с привлечением методов машинного обучения.

Доклад посвящен разработке методов прогноза временных характеристик прикладных сетевых сервисов в зависимости от данных об эксплуатации сетевого сервиса на используемых сетевых ресурсах, текущего и прогнозируемого состояния аппаратного и программного обеспечения.

Развитие технологии распределенных облачных вычислений, реализация отказоустойчивых сетевых сервисов в виде распределенных сетевых услуг приводит к необходимости перехода от атомарных систем, локализованных на выделенных для этого аппаратных единицах, к так называемым сетевым распределенным системам,

которые работают, пока обеспечивается внутренняя и внешняя сетевая связность их компонент. Такие системы используют преимущества технологии виртуализации и распределены по различным серверным мощностям. Сеть, обеспечивающая их функционирование, должна при этом обладать достаточной степенью избыточности для обеспечения заданного уровня качества обслуживания. Если экземпляр сервиса или его составляющей части оказался недоступен, то в зависимости от уровня обслуживания управление сервисом восстанавливает работоспособность, активизируя резервную копию или соответствующий клон сервиса или его составляющей части.

2. Прогноз выполнения задач

Сетевые сервис, функция, услуга существуют и работают не на конкретной единице оборудования, а в сети и пока существует сеть, опирающаяся на определенный уровень избыточности в ИКТ инфраструктуре. При этом пользователям сети предоставляется широкий спектр качественно новых информационных услуг:

- организация информационного взаимодействия между автоматизированными системами, как унаследованными в качестве поставщиков информации, так и вновь включаемыми в контур управления;
- доступ к данным, при необходимости, к первичным, в рамках реализации перехода от централизованной схемы «загрузка в централизованное хранилище с очисткой — анализ — распространение» к схеме «распределенные размещение и предобработка данных — загрузка — анализ — распространение»;
- интеграция информационных ресурсов звеньев управления;
- возможность создания гибкой распределенной информационной инфраструктуры, легко адаптируемой под оперативно создаваемые сегменты и изменения структуры сети;
- обеспечение заданного уровня безотказности и живучести.

Обеспечение функционирования распределенных вычислительных систем в соответствии с предъявляемыми к ним техническими требованиями определяется эффективностью размещения сетевого сервиса на оборудовании в сети. Программируемость полученной таким образом вычислительной среды — один из ключевых факторов развития информационных инфраструктур. Широкий спектр приложений, таких как службы Интернета вещей, методы машинного обучения, анализа данных и научных вычислений и прочие, используют модель вычислений «Функции как услуга» (Function as a Service, FaaS) облачных платформ Amazon Lambda, Azure, Google Functions и другие. Широкое распространение получают «бессерверные вычисления», что предполагает разбиение приложения на небольшие «функции», локализованные на разных узлах сети. Работа таких приложений управляется единой облачной платформой.

Ключевым вопросом для организации работы распределенных приложений, особенно в труднодоступных малонаселенных районах, например, Крайнего севера или горных массивов, является решение задачи локализации прикладных сервисов, сетевых функций по узлам сети. Задача такой локализации должна учитывать доступные вычислительные ресурсы, прогноз времени выполнения, характеристики энергоснабжения, вопросы охлаждения, состояние сетевых соединений, принятые требования качества сервиса (QoS, SLA), а также другие характеристики.

В докладе рассматривается построение прогнозирования времени выполнения задачи [1, 2] на узлах сети с учетом указанных выше характеристик. В качестве исходных данных для построения и тестирования методов прогнозирования

используются опубликованные данные трасс работы распределенных сетевых сервисов на типовом оборудовании облачных сетевых структур и сетей доступа к ним [3].

В докладе содержатся результаты исследования способов прогнозирования реализации сетевого сервиса, исходя из априорных данных его работы на оборудовании сетей доступа и заданных вероятностно-временных характеристик предоставления услуги, а именно методы, основанные на использовании:

- случайного леса,
- матричных разложений,
- глубоких сверточных нейронных сетей (ГНС, CNN) на платформе машинного обучения ГосНИИАС «Plat2023»/«Платформа-ГНС» [4, 5].

В качестве примеров рассмотрены задачи прогнозирования работы сетевых приложений доставки видео-содержимого, а также задачи хранения и управления данными в СУБД типа «ключ-значение».

3. Заключение

В докладе предложены методы прогнозирования результатов выполнения распределенных прикладных сервисов на узлах сети с целью соблюдения предустановленных показателей соглашения об уровне качества обслуживания, а также с учетом текущих характеристик состояния сети и оборудования.

Список литературы

1. Bishop C.M. Pattern Recognition and Machine Learning, Springer Science+Business Media, 233 Spring Street, New York, NY 10013, USA, 2006.
2. Френкель С.Л., Захаров В.Н., Модели учета влияния статистических характеристик трафика вычислительных сетей на эффективность прогнозирования средствами машинного обучения // Информатика и ее применения. 2023. Т. 17, Вып. 4. С. 71-80.
3. Data traces from a data center testbed – Kaggle. <https://www.kaggle.com/datasets/jaliltaghia/data-traces-from-a-data-center-testbed> (дата обращения: 11.09.2023).
4. Визильтер Ю.В., Вишняков Б.В., Желтов С.Ю. Современные технологии искусственного интеллекта и их применение в авиационных комплексах // Материалы XVI Всероссийская мультikonференция по проблемам управления (МКПУ-2023). Волгоград, Россия, 11-15 сентября 2023. Т. 3. Управление аэрокосмическими системами (УАКС-2023). С. 13-16.
5. Платформа-ГНС. Унифицированная программная платформа машинного обучения. Материалы к семинару-конференции «PLAT-2023» ФАУ «ГосНИИАС» 15 февраля 2023 https://www.gosniias.ru/pages/d/plat-2023-brosh_gns.pdf (дата обращения: 17.01.2024).