

Интеллектуализация разработки распределенных пакетов прикладных программ на основе непрерывной интеграции программного обеспечения и конфигурирования ресурсов

И.В. Бычков¹, А.Г. Феоктистов¹, С.А. Горский¹, Р.О. Костромин¹

¹Институт динамики систем и теории управления им. В.М. Матросова СО РАН, Лермонтова 134, Иркутск, Россия, 664033

Аннотация

В докладе предложен подход к поддержке непрерывной интеграции программного обеспечения и конфигурирования ресурсов в инструментальном комплексе Orlando Tools при подготовке и проведении крупномасштабных вычислительных экспериментов. Представлены аспекты интеллектуализации этих процессов с использованием концептуального моделирования и мультиагентных технологий.

Ключевые слова

Непрерывная интеграция, конфигурация ресурсов, интеллектуализация

1. Введение

Потребность в проведении крупномасштабных научных экспериментов в распределенных средах неуклонно растет. В настоящее время средства разработки научных пакетов, как правило, не поддерживают непрерывную интеграцию программного обеспечения и конфигурирование необходимых ресурсов в полной мере. Это негативно отражается на качестве программного обеспечения и увеличивает как сроки его разработки, так и время проведения экспериментов [1]. Проблемы, возникающие при использовании известных инструментариев непрерывной интеграции в распределенных средах с разнородными ресурсами, обсуждаются в [2, 3]. Доклад посвящен разработке новых средств создания распределенных пакетов прикладных программ (РППП) в распределенной среде на основе мультиагентных технологий, интеллектуального анализа ретроспективных и текущих данных, а также машинного обучения агентов.

2. Поддержка вычислений в распределенной среде на основе непрерывной интеграции программного обеспечения

Предложен подход к непрерывной интеграции как прикладного, так и системного программного обеспечения в процессе разработки РППП на основе модульного подхода. Он реализован в инструментальном комплексе Orlando Tools (ОТ) [4]. В нем выделяются два уровня непрерывной интеграции. На первом уровне решаются задачи автоматизации разработки модулей путем интеграции средств ОТ с системой GitLab [5]. Второй уровень, связанный с объединением модулей в рамках РППП и размещением пакета на ресурсах среды, полностью реализуется ОТ. Инструментальный комплекс обеспечивает автоматизацию как разработки программного кода, его сборки и тестирования на уровне модулей, так и развертывания пакета на ресурсах среды с его последующим тестированием в целом. Концептуальная модель среды, поддерживаемая в ОТ и представляющая знания о программно-аппаратном обеспечении, является основой поддержки всех процессов непрерывной интеграции.

3. Конфигурирование ресурсов среды и их настройка

ОТ обеспечивает формирование требуемой конфигурации ресурсов среды и их настройку. В соответствии с новыми требованиями к ресурсам среды, необходимым в процессе решения

задач, в рамках ОТ разработана специализированная подсистема управления конфигурациями ресурсов, обеспечивающая автоматизацию их виртуализации, запуска и настройки. В процессе своей работы данная подсистема использует средство управления конфигурациями Ansible [6], а также платформу OpenStack [7] для запуска виртуальных машин (ВМ) и управления виртуализированными ресурсами. В качестве таких ресурсов могут выступать ресурсы центров коллективного пользования, а также ресурсы облачных и туманных платформ. Для запуска ВМ на ресурсах коллективного пользования формируются системные задания, направляемые в общие очереди локальных менеджеров этих ресурсов. Распределение ВМ между узлами среды реализуется мультиагентной системой управления потоками заданий. Агенты используют систему мониторинга ресурсов для получения, обработки, хранения, интеллектуального анализа и применения данных (как текущих, так и ретроспективных) о процессах их функционирования.

4. Заключение

Практическая значимость результатов исследования заключается в уменьшении числа ошибок и сбоев при разработке и применении РППП, сокращении времени, необходимого для проведения экспериментов, и повышении эффективности использования ресурсов. Данные, извлекаемые в рамках непрерывной интеграции и конфигурации ресурсов используются в гибридной имитационной модели анализа функционирования компонентов информационно-вычислительных систем и их программно-аппаратных инфраструктур, а также в процессе машинного обучения агентов, представляющих данные компоненты.

Исследования проведены в рамках выполнения научной темы «Технологии разработки и анализа предметно-ориентированных интеллектуальных систем группового управления в недетерминированных распределенных средах».

5. Литература

- [1] Poola, D. A taxonomy and survey of fault-tolerant workflow management systems in cloud and distributed computing environments / D. Poola, M.A Salehi, K. Ramamohanarao, R. Buyya // Software architecture for big data and the cloud. – Morgan Kaufmann, 2017. – P. 285-320.
- [2] Möller, S. Robust cross-platform workflows: how technical and scientific communities collaborate to develop, test and share best practices for data analysis / S. Möller, S. Prescott, L. Wirzenius, P. Reinholdtsen, B. Chapman, P. Prins, S. Soiland-Reyes, F. Klötzl, A. Bagnacani, M. Kalaš, A. Tille, M.R. Crusoe // Data Science and Engineering. – 2017. – Vol. 2(3). – P. 232-244.
- [3] Shahin, M. Continuous integration, delivery and deployment: a systematic review on approaches, tools, challenges and practices / M. Shahin, M. Babar, L. Zhu // IEEE Access. – 2017. – Vol. 5. – P. 3909-3943.
- [4] Tchernykh, A. Orlando Tools: Development, Training, and Use of Scalable Applications in Heterogeneous Distributed Computing Environments / A. Tchernykh, A. Feoktistov, S. Gorsky, I. Sidorov, R. Kostromin, I. Bychkov, O. Basharina, A. Alexandrov, R. Rivera-Rodriguez // Communications in Computer and Information Science. – 2019. – Vol. 979. – P. 265-279.
- [5] DevOps Platform Delivered as a Single Application – GitLab [Электронный ресурс]. – Режим доступа: <https://about.gitlab.com> (26.01.2021).
- [6] Kostromin, R. Survey of software configuration management tools of nodes in heterogeneous distributed computing environment / R. Kostromin // Proceedings of the 2nd International Workshop on Information, Computation, and Control Systems for Distributed Environments. – 2020. – Vol. 2638. – P. 156-165.
- [7] Feoktistov, A. Virtualization of heterogeneous HPC-clusters based on OpenStack platform / A. Feoktistov, I. Sidorov, V. Sergeev, R. Kostromin, V. Bogdanova // South Ural State University Bulletin: Computational Mathematics and Software Engineering series. – 2017. – Vol. 6(2). – P. 37-48.