

DOI:

## МЕТОДИКИ ОЦЕНКИ ЭФФЕКТИВНОСТИ ИНТЕГРАЦИИ ПРОГРАММНО-ТЕХНОЛОГИЧЕСКИХ РЕШЕНИЙ В ЦИФРОВЫЕ ПЛАТФОРМЫ<sup>1</sup>

Ильин Д. Ю., Никульчев Е.В.

*МИРЭА – Российский технологический университет, Россия, г. Москва, пр. Вернадского, д. 78; Российская академия образования Россия, г. Москва, ул. Погодинская, д. 8*  
nikulchev@mail.ru, i@dmitryilin.com

*Аннотация: Цифровые платформы позволяют осуществлять сбор и обмен информацией между огромным количеством пользователей, объединяя результаты в большие данные. Доклад посвящен разработке методик и виртуальных имитационных информационных инфраструктур для оценки эффективности интеграции программно-технологических решений в цифровые платформы (на примере цифровой платформы массовых исследований в системе образования). Использование концепции «инфраструктура как код» сокращает затраты на проведение вычислительных экспериментов на заданном стеке технологий и позволяет адекватно оценить технологическое решение в условиях интеграции в стек технологий, при этом учитывается инфраструктура и заданные условия функционирования разрабатываемой цифровой платформы. Предложенный подход позволяет на этапе выбора технологий оценить их эффективность и снизить затраты на вычислительные эксперименты. Результаты внедрения демонстрируют эффективность предложенных методик.*

Ключевые слова: цифровые платформы, стек технологий, интегрированные архитектуры.

### Введение

Распространение веб-приложений, обусловленная их платформи- и аппаратно-независимостью, массовостью использования, широкими возможностями применяемых в интерфейсов и протоколов обмена данных, программируемыми возможностями определило развитие ИТ-направления – создание цифровых платформ (ЦП). Использование ЦП позволяет осуществлять сбор и обмен информацией между огромным количеством пользователей, объединяя результаты в большие данные.

Несмотря на обилие различных применений ЦП, сформировался комплекс используемых в них современных информационных технологий (технологических решений), который принято назвать «стек технологий». Важной особенностью элементов этого стека является их заменяемость, то есть одна из технологий может быть заменена на альтернативную или вновь созданную.

Цифровые платформы, предназначенные для сбора и обработки больших объемов данных, создали не только новые задачи для исследования при проектировании интегрированных многоуровневых архитектур (такие как нагрузочные эксперименты, выбор компонентов, обеспечение кроссплатформенности интерфейсов, учет различных протоколов передачи данных), но и критерии оценки – масштабируемость, обеспечение гарантированной доставки данных, восстанавливаемость. Актуальными являются задачи интеграции, характерные для широкого класса ЦП, возникающих на разных этапах разработки и модернизации [1–3]: выбор и оценка ресурсной эффективности способов хранения данных; выбор способов передачи данных в условиях ограничений на каналы связи и возможности вычислительной инфраструктуры, подбор эффективно взаимодействующих компонентов свободно-распространяемого программного обеспечения с учетом специфики решаемых задач, задачи повышения эффективности за счет обновления версии или миграции на новое программно-технологическое решение.

Цель исследования – разработка методик, информационных моделей и виртуальных имитационных информационных инфраструктур для оценки эффективности интеграции программно-технологических решений в цифровые платформы (на примере цифровой платформы массовых исследований в системе образования).

Цифровые платформы как правило разрабатываются не с чистого листа, а путем интеграции ряда существующих и вновь созданных технологических решений. В условиях необходимости обслуживания большого числа пользователей на передний план выходят требования к эффективности и надежности функционирования [4, 5] платформы. На эффективность и надежность могут влиять архитектура системы [6], отдельные компоненты [7, 8], параметры инфраструктуры [9].

Для построения архитектурных решений ЦП применяются методы анализа и синтеза. Первичный этап, анализ требований к архитектуре (architecture requirement analysis), проводится с целью выявления основных вариантов использования, функциональных и нефункциональных требований

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (грант 17-29-02198)

(см. ГОСТ Р 57100-2016/ISO/IEC/IEEE 42010:2011; ГОСТ Р ИСО/МЭК 25010-2015). На основании полученной информации проводится синтез архитектуры (architecture synthesis) с целью формирования наиболее эффективного архитектурного решения на основе определенного набора слабосвязанных компонентов ЦП, их связей и эффективных способов обмена данными.

При разработке веб-интерфейсов ЦП применяются инструменты, позволяющие проверить работоспособность той или иной функциональной составляющей. Устоявшимся подходом к экспериментальной оценке и верификации технологических решений является тестирование [10, 11] программного обеспечения. Также для экспериментальной оценки и верификации технологических решений используют ретроспективные данные мониторинга.

В работе предложено для анализа технологических решений формировать виртуальную имитационную среду для оценки требуемых показателей, учитывающую особенности инфраструктуры вычислительного комплекса, окружение и режим функционирования. Данный подход позволит обеспечить выбор решений на основе количественных характеристик, соответствующих требованиям ЦП.

## 2 Средства оценки эффективности отдельных технологических решений в виртуальной конфигурируемой имитационной среде

Выбор технологических решений в ходе проектирования и разработки ЦП влияет на последующие этапы, при этом чем позднее будет выявлена ошибка – тем выше будет стоимость ее исправления. Среди распространенных подходов по устранению проблемы – использование инкрементальных моделей с коротким циклом разработки новой версии ЦП, например, гибкие методологии.

Чтобы сократить затраты, предлагается проводить экспериментальные исследования в виртуальной конфигурируемой имитационной среде (рис. 1) с целью верификации интегрируемых в ЦП технологических решений.

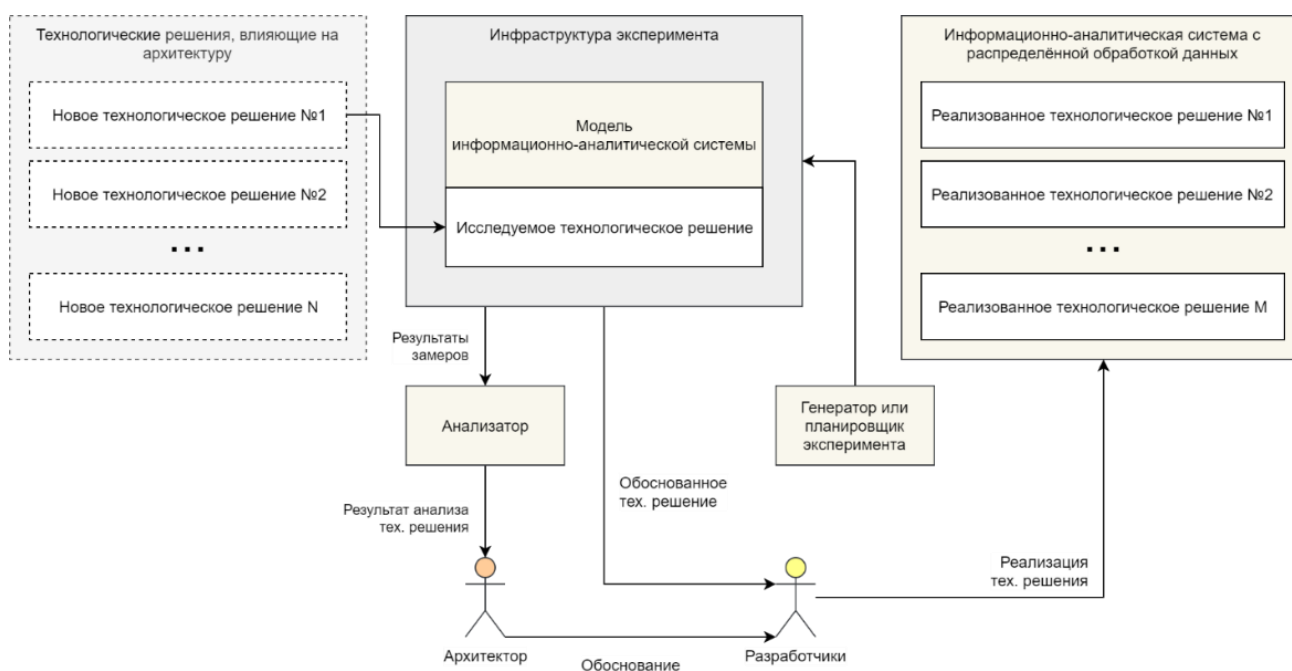


Рис. 1. Схема добавления верифицированных технологических решений в стек технологий ЦП

Задача и методика по ее решению формулируется следующим образом. Требуется найти стек технологий эффективный по заданным критериям: цифровая платформа с интегральной архитектурой; ограничения – заданная масштабируемая архитектура, условия функционирования и инфраструктура; набор технологических решений (компоненты, протоколы, СУБД, структуры данных); критерии качества функционирования ЦП.

Решение этой задачи возможно с помощью разработанной методики, предполагающей реализацию следующих шагов.

1. Определение ограничений, влияющих на выбор ТР.
2. Формирование множества альтернатив.

3. Определение существенных критериев качества для заданной ЦП.
4. Планирование эксперимента.
5. Разработка виртуальной имитационной инфраструктуры.
6. Проведение серии вычислительных экспериментов.
7. Принятие решения.
8. Внедрение выбранного ТР в стек.

При постановке эксперимента важна минимизация влияния наблюдателя на объект. Для изоляции исследуемого технологического решения от посторонних воздействий необходимо сформировать независимую инфраструктуру. Ее можно подготовить как аппаратно, используя физические вычислительные устройства (компьютеры, серверы, роутеры и др.), так и программно, с помощью виртуальных машин. Второй вариант стоит считать предпочтительным, поскольку реализация инфраструктуры программными средствами отличается более рациональным использованием ресурсов и переносимостью. Кроме того, использование виртуальных машин предоставляет возможность многократного использования инфраструктуры.

Следует принять во внимание факт, что инфраструктура, подготовленная с помощью виртуальных машин, имеет ряд недостатков: требуется большой объем дискового пространства; тяжело отслеживать текущее состояние виртуальных машин, а вносимые изменения требуется документировать отдельно.

Для нивелирования недостатков можно воспользоваться подходом «инфраструктура как код». Подход реализуется с помощью таких систем, как:

- системы создания и конфигурирования виртуальной среды разработки (например, система Vagrant);
- системы для автоматизации развертывания и управления приложениями в средах с поддержкой контейнеризации (например, Docker);
- системы управления конфигурациями (например, Ansible, Puppet).

Как показывают исследования, для постановки экспериментов системы контейнеризации подходят меньше. Они предоставляют меньший уровень изоляции вычислительных ресурсов, что может сказаться на результатах.

Для проведения сравнительных экспериментов с целью выбора эффективных технологических решений разработан инструмент экспериментальной оценки и верификации технологических решений ЦП (рис. 2).

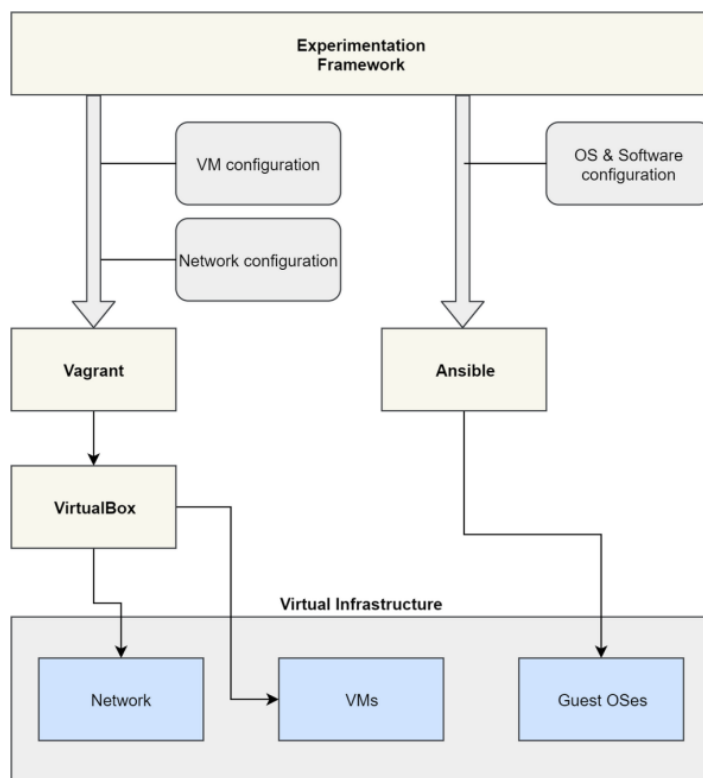


Рис. 2. Схема подготовки виртуальной инфраструктуры для экспериментальной оценки и верификации технологических решений

На рис. 2 отражена общая схема экспериментального стенда. В качестве исходных данных используются конфигурационные файлы на языках YAML, Ruby. На их основе производится запуск виртуальных машин с заданными характеристиками и настройка сетевых подключений. Далее на основе соответствующей настройки производится загрузка эталонного образа операционной системы и ее запуск. После запуска с помощью системы управления конфигурациями (Ansible) осуществляется установка необходимого программного обеспечения на гостевые операционные системы.

Важной задачей, являющейся общей для разработчиков ЦП, является выбор структуры хранения данных. В последнее десятилетие многие системы сбора данных представляют собой веб-приложения. Это открывает возможности подключения к вычислительным сервисам большого количества пользователей независимо от используемого ими гаджета и операционных систем. Изменения в дизайне и архитектуре программного обеспечения и требованиях к приложениям определило требования хранения больших данных. Требования к масштабированию определили развитие нереляционных баз данных – NoSQL. Однако при построении ЦП необходимо осуществить выбор между СУБД и формой хранения данных. Для выбора СУБД разработана методика экспериментального исследования ресурсной эффективности записи данных в условиях ограниченности вычислительных ресурсов.

Методика состоит из следующих шести шагов.

1. Инициализация виртуальных машин.
2. Загрузка массива данных в оперативную память виртуальной машины генератора клиентских запросов.
3. Последовательная запись инвариантного набора данных (ResearchSubject).
4. Последовательная запись тестового набора данных (ResearchResult, ResearchResultAnswer).
5. Получение результатов о количестве сохраненных записей.
6. Анализ результатов и выбор эффективного варианта.

Разработан алгоритм интегральной оценки качества стека технологий на основе нечеткой логики и роевого интеллекта.

1. Формализация задачи выбора стека технологий [9].
2. Формирование правил нечеткого вывода принятия решений об эффективности стека технологий ЦП.
3. Исследование системы нечеткого вывода на полноту покрытия правилами диапазона входных значений, отсутствие избыточности правил, исключение ситуации неоднозначного выбора за счет настройки весов правил.
4. Выбор способа нормировки значений частных показателей качества технологий.
5. Организация экспериментов в виртуальной имитационной информационной инфраструктуре для получения нормированных значений для конфигурации виртуальной имитационной информационной инфраструктуры.
6. Использование алгоритма роевого интеллекта для организации направленного поиска эффективного стека для конфигурации решения.

## Результаты внедрения

Методики выбора стека технологий и результаты их применения внедрены в федеральную ЦП массовых психологических веб-исследований «DigitalPsyTools» [12] Российской академии образования. «DigitalPsyTools» использовалась для информационной поддержки веб-исследований в системе образования: мониторинга психологической службы среднего образования, психологического тестирования студентов вузов РФ, популяционного психологического тестирования школьников Ленинградской области, психологического тестирования по COVID-19. Анализ результатов показал, что выбранные технологические решения обеспечивают надежное функционирование при заданной нагрузке во время эксплуатации федеральной системы.

На рис. 3 показано, как изменялось ежедневное число пользователей, отправляющих анкетные данные. В сутки максимальное значение превысило 3500 тысячи уникальных записей.

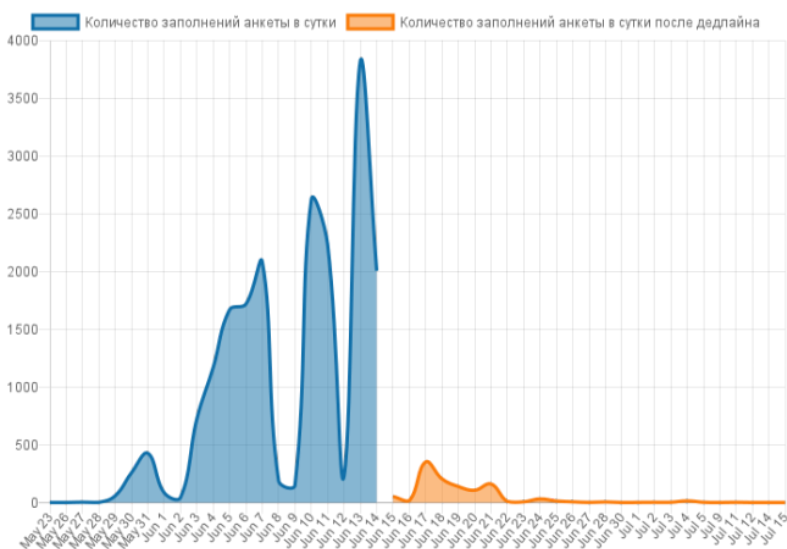


Рис. 3. График участия в опросе

Полученные результаты внедрения и развития ЦП массовых исследований в системе образования демонстрируют эффективность выбранных технологических решений, что в свою очередь, обеспечено разработанными методиками экспериментальной оценки технологических решений в виртуально имитационной инфраструктуре.

## Заключение

Проведен обзор методов повышения эффективности интеграции технологий и дизайна архитектур цифровых платформ. Объединены задачи выбора компонентов, выбора СУБД и др. в задачу выбора стека технологий. Сформулирована задача исследования, позволяющая сократить размерность и повысить скорость сходимости алгоритмов поиска эффективного решения за счет включения в качестве ограничений режима функционирования ЦП.

Разработана виртуальная программно-конфигурируемая имитационная среда как инструмент оценки эффективности стека технологических решений в ЦП. Использование концепции «инфраструктура как код» сокращает затраты на проведение вычислительных экспериментов на заданном стеке технологий и позволяет адекватно оценить технологическое решение в условиях интеграции в ЦП.

Разработана методика оценки интеграции технологических решений при создании ЦП. Методика позволяет оценить значения критериев для оценки эффективности ТР в конкретных условиях конкретной ЦП. Впервые учитывается инфраструктура и заданные условия функционирования ЦП, что снижает множество возможных решений. Предложенный подход «инфраструктура как код» позволяет снизить затраты на эксперименты.

Сформирована методика выбора эффективных технологических решений на основе использования нечеткой логики. Для осуществления направленного поиска и сокращения перебора вариантов выбора стека технологий применен поиск решений на основе роевого алгоритма. Полученное решение позволило сократить количество итераций по оценке показателей в условиях заданной вычислительной инфраструктуры. Для одного из примеров число итераций сокращено с 1080 до 312.

Методики выбора стека технологий и результаты их применения внедрены в федеральную ЦП массовых психологических веб-исследований «DigitalPsyTools». Анализ результатов показал, что выбранные технологические решения обеспечивают надежное функционирование при заданной нагрузке во время эксплуатации федеральной системы.

## Литература

1. Комлева Н., Вилявин Д.А. Цифровая платформа для создания персонализированных адаптивных онлайн курсов // Открытое образование. 2020. Т. 24, № 2. – С. 65–72.
2. Лукинова О. Вопросы проектирования цифровых платформ в парадигме открытых систем // Управление развитием крупномасштабных систем (MLSD'2018). – М. : ИПУ РАН, 2018. – С. 304–306.

3. *Barak M., Ziv S. Wandering: A Web-based platform for the creation of location-based interactive learning objects // Computers & Education. 2013. Vol. 62. – P. 159–170.*
4. *Bishung J., Koyejo O., Okezie A. et al. A critical analysis of topics in software architecture and design // Advances in Science, Technology and Engineering Systems Journal. 2019. Vol. 4, No. 2. – P. 211–220.*
5. *Almarimi N., Ouni A., Bouktif S., Mkaouer M.W., Kula R.G., Saied M.A. Web service API recommendation for automated mashup creation using multi-objective evolutionary search // Applied Soft Computing. 2019. Vol. 85. – P. 105830.*
6. *Ramírez A., Parejo J.A., Romero J.R., Segura S., Ruiz-Cortés A. Evolutionary composition of QoS-aware web services: a many-objective perspective // Expert Systems with Applications. 2017. Vol. 2. – P. 357–370.*
7. *Roy-Hubara N., Sturm A. Design methods for the new database era: a systematic literature review // Softw Syst Model. 2020. Vol. 19, № 2. – P. 297–312.*
8. *Ilin D., Gusev A., Nikulchev E. Software Design using Genetic Quality Components Search // International Journal of Advanced Computer Science and Applications. 2019. Vol. 10, No. 12. – P. 48-54*
9. *Gusev, A., Ilin D., Kolyasnikov P., Nikulchev E. Effective Selection of Software Components Based on Experimental Evaluations of Quality of Operation // Engineering Letters. 2020. Vol. 28, No. 2. – P. 420-427*
10. *Garousi V., Giray G., Tüzün E., Catal C., Felderer M. Aligning software engineering education with industrial needs: a meta-analysis // Journal of Systems and Software. 2019. Vol. 156. – P. 65–83.*
11. *Lemos O.A.L., Silveira F.F., Ferrari F.C., Garcia A. The impact of Software Testing education on code reliability: An empirical assessment // Journal of Systems and Software. 2018. Vol. 137. – P. 497–511.*
12. *Никольчев Е.В., Ильин Д.Ю., Колясников П.В., Исмагуллина В.И., Захаров И.М., Малых С.Б. Разработка открытой цифровой платформы масштабных психологических исследований Вестник РФФИ. 2019. № 4 (104). – С. 99–113.*