

Предисловие

В 24-м томе Трудов Института системного программирования РАН представлены статьи исследователей, аспирантов и студентов из ИСП РАН, Московского государственного университета и нескольких других организаций. В статьях описываются результаты исследований, выполненных в конце 2012 г. и первой половине 2013 г. Статьи данного тома сгруппированы в пять тематических блоков. Первый блок посвящен различным аспектам современных высокопроизводительных вычислений.

В статье А.О. Кудрявцева и др. «Разработка и реализация облачной системы для решения высокопроизводительных задач» описаны основные проблемы, возникающие при переносе высокопроизводительных вычислений в облако. Рассматривается подход к организации высокопроизводительного облачного сервиса с использованием виртуализации. Описана архитектура разработанной системы «виртуальный суперкомпьютер», основанной на облачной платформе OpenStack и системе виртуализации KVM/QEMU.

И.А. Дудина, А.О. Кудрявцев и С.С. Гайсарян представили статью «Разработка и реализация облачного планировщика, учитывающего топологию коммуникационной среды при высокопроизводительных вычислениях», в которой отмечается, что для эффективного решения высокопроизводительных задач в облаке необходимо планирование виртуальных машин на серверах, учитывающее особенности высокопроизводительных вычислений. Рассматривается подход к планированию, основанный на оценке размещения с помощью Hop-Byte метрики.

В статье М.С. Акопяна и Н.Е. Андреева «Исследование и разработка шаблонов неэффективного поведения в параллельных MPI, UPC приложениях» обсуждаются шаблоны в параллельных программах, приводящие к потере производительности. Рассматриваются шаблоны как в параллельных MPI приложениях для вычислительных систем с распределенной памятью, так и в параллельных UPC программах для систем с разделенным глобальным адресным пространством (PGAS). Предложен метод автоматизированного обнаружения шаблонов неэффективного поведения.

В статье «Анализ эффективности итерационных методов решения систем линейных алгебраических уравнений, реализованных в пакете OpenFOAM» предложен способ определения коэффициентов усиления гармоник, основанный на использовании дискретного преобразования Фурье. В качестве примера приведен анализ эффективности метода BiCGStab с ILU и многосеточным предобуславливанием из пакета OpenFOAM при решении разностных аналогов уравнений Гельмгольца и Пуассона, возникающих при моделировании течения вязкой несжимаемой жидкости в квадратной каверне.

В статье Я.В. Загуменного и Ю.Д. Чашечкина «Расчет течений непрерывно стратифицированной жидкости с использованием открытых вычислительных пакетов на базе технологической платформы UniHUB» описывается авторский опыт использования технологической платформы UniHUB при численном моделировании и проведении расчетов течений непрерывно стратифицированной жидкости с применением свободных прикладных вычислительных пакетов OpenFOAM, Salome и ParaView. Внимание уделяется вопросам построения высокоразрешающих расчетных сеток, постановки сложных граничных с помощью встроенных и расширенных утилит пакета OpenFOAM, разработки собственных решателей, обработки и визуализации расчетных данных, а также проведения расчетов в параллельном режиме на вычислительном кластере МСЦ РАН.

Завершает первый блок статья Ю.А. Румянцев «Прямая передача данных между ПЛИС Virtex-7 по шине PCI Express», в которой рассматривается передача данных по шине PCI Express с одновременным участием нескольких ПЛИС. В статье, в частности, показано, что при одновременных передачах через общий канал скорость отдельных передач не уменьшается до тех пор, пока суммарная скорость передачи не превышает пропускную способность общего канала; затем канал используется на 100%, а его пропускная способность делится поровну между устройствами.

Второй тематический блок содержит статьи, в которых обсуждаются методы компиляции и анализа программ, а также их приложений.

Статья Ш.Ф. Курмангалеева «Методы оптимизации Си/Си++ - приложений распространяемых в биткоде LLVM с учетом специфики оборудования» посвящена методам оптимизации Си/Си++ приложений, которые применяются в системе двухэтапной компиляции, позволяющей распространять такие приложения в промежуточном представлении LLVM. Описывается метод статического инструментирования с неполным покрытием всех дуг потока управления программы и последующей его коррекцией для сбора профиля, позволяющий получать профиль, сравнимый по качеству с классическим подходом, но при этом обеспечивающий существенное снижение накладных расходов на сбор профиля. Предлагается подход к построению специализированного облачного хранилища приложений, позволяющий решать вопросы оптимизации и защиты программ, а также обеспечивающий снижение накладных расходов на компиляцию и оптимизацию приложений в облачной инфраструктуре.

В статье Ш.Ф. Курмангалеева и др. «О методах деобфускации программ» отмечается, что в связи со схожестью проблем, стоящих перед оптимизирующим компилятором и деобфускатором, представляется разумным использование компиляторной инфраструктуры LLVM в качестве ядра деобфускатора. Хотя из-за высокой зашумленности выходного кода авторам пришлось разработать свою инфраструктуру, которая позволяет добиться более чистого выходного кода, применение LLVM или аналогичной

разработки остается одним из перспективных направлений при разработке деобфусцирующего программного обеспечения.

И.Е. Бронштейн в статье «Вывод типов для языка Python» приводит обзор описанных в научной литературе алгоритмов вывода типов для языков с параметрическим полиморфизмом. Затем описывается новый алгоритм, являющийся модификацией алгоритма декартова произведения. Демонстрируется, как модуль вывода типов, использующий новый алгоритм, анализирует различные конструкции языка Python.

Наконец, в статье С.П. Вартанова и М.К. Ермакова «Автоматический поиск ошибок синхронизации в приложениях на платформе Android рассматривается задача автоматического поиска ошибок синхронизации при проведении динамического анализа приложений в рамках платформы Android. Приводится обзор существующих инструментов для анализа компонентов Android-приложений. Описывается механизм обнаружения ошибок синхронизации, используемый инструментом динамического анализа байт-кода Coffee Machine, и его основные аспекты: инструментирование, генерация трассы отношений предшествования и использование инструмента ThreadSanitizer Offline для обнаружения состояния гонок по сгенерированной трассе для платформы Android.

Третий блок статей посвящен использованию формальных методов для тестирования программного обеспечения.

В.П. Иванников, А.К. Петренко и др. в статье «Опыт использования UniTESK как зеркало развития технологий тестирования на основе моделей» описывают наиболее значимые применения технологии UniTESK в промышленных проектах, суммируют их опыт и оценивают перспективные направления развития компонентных технологий тестирования на основе моделей.

М.У. Мандрыкин, В.С. Мутилин и Хорошилов А. В. в статье «Введение в метод CEGAR – уточнение абстракции по контрпримерам» указывают, что успешный автоматизированный анализ программных систем среднего размера с использованием проверки моделей, получаемых при помощи предикатной абстракции, стал возможен благодаря развитию метода уточнения абстракции по контрпримерам – CEGAR. Этот метод так или иначе используется в таких инструментах, как SLAM, BLAST, SATABS и CPAchecker. В статье рассматривается метод CEGAR в том виде, как он реализован в инструментах статической верификации BLAST и CPAchecker.

Статья Е.М. Новикова «Построение спецификаций программных интерфейсов в открытой системе покомпонентной верификации ядра Linux» содержит описание нового подхода к построению спецификаций программных интерфейсов, который позволяет достаточно эффективно применять инструменты статической верификации для проверки выполнения правил использования программных интерфейсов в условиях неполноты модели окружения.

В статье В.Н. Федотова «Автоматизация регрессионного тестирования при помощи анализа трасс событий» представлен алгоритм process mining, предназначенный для применения в инструментированной распределенной системе. Система инструментруется таким образом, что взаимодействия между модулями системы проходят через единую шину, связывающую все компоненты системы. Записанные при обработке на шине события анализируются представленным в работе алгоритмом, результатом работы которого является модель взаимодействий между модулями системы. Полученная модель используется для разработки покрывающего регрессионного тестового набора.

Четвертый тематический блок содержит ряд достаточно разнородных статей, так или иначе связанных с управлением данными и их анализом.

В статье С.Д. Кузнецова и А.В. Посконина «Распределенные горизонтально масштабируемые решения для управления данными» отмечается, что новые потребности современных приложений привели к появлению большого числа специализированных распределенных систем, способных удовлетворительно справляться с возникающими задачами. Предлагается обзор некоторых современных решений, обеспечивающих масштабируемость при работе с большими объемами данных под высокими нагрузками.

С.В. Герасимов и др. в статье «Инструментальные средства оценки качества научно-технических документов» предлагают комбинированный подход к оценке качества научно-технических документов, учитывающий различные категории автоматически рассчитываемых характеристик качества документов – как существующие библиометрические и наукометрические характеристики, так и новые типы характеристик, основанные на семантическом анализе текстов научно-технических документов, применении эвристических правил и методов оценки наличия прямых текстовых заимствований.

Статья В.А. Золотова и В.А. Семенова «Современные методы поиска и индексации многомерных данных в приложениях моделирования больших динамических сцен» посвящена современным методам поиска и индексации многомерных данных применительно к задачам моделирования больших динамических сцен. Для решения этих задач требуется применение специальных схем индексации многомерных данных. В статье обсуждаются два фундаментальных подхода к организации подобных схем, а также проводится их сравнительный анализ в контексте комплексных требований, предъявляемых к приложениям обсуждаемого класса.

В статье И.В. Блудова «Особенности табличных выражений SQL и их соответствие с концепциями реляционной модели данных» анализируется соответствие операторов SQL реляционной модели данных. Рассмотрены особенности табличных выражений в стандарте SQL и показаны случаи, в которых эти выражения противоречат реляционной модели. Приведены

варианты использования таких табличных выражений в SQL, которые наиболее точно отражают концепции реляционной модели.

Статья Л.Е. Карпова и В.Н. Юдина «Роль предыстории при оценке сложного объекта в управлении по прецедентам» посвящена принципам построения системы управления для нестационарных объектов – развитию подхода к управлению объектами по прецедентам, разработанного в рамках более ранних работ авторов, применительно к нестационарным объектам с неполным описанием. Упор сделан на разработку методики, позволяющей снизить неоднозначность оценки объекта до и после воздействия. Источником дополнительной информации может служить предыдущее поведение объекта.

В статье тех же авторов «Гибридный подход к построению систем поддержки решений» отмечается, что в существующих на данный момент гибридных системах поддержки принятия решений основным инструментом вывода являются порождающие правила, а прецеденты используются лишь для обработки исключений. В описываемом подходе рассуждения по правилам и прецедентам дополняют друг друга в условиях неоднозначной оценки неполностью описанного случая. Разработана методика двухэтапной оценки случая, где сначала используется прецедентный подход с целью получить представление о возможной принадлежности случая тому или иному классу. На основе этой информации, на втором этапе, уже по правилам вывода проводится обратный логический вывод от возможного заключения как гипотезы – к фактам, подтверждающим эту гипотезу.

Последний блок тома посвящен математическим проблемам компьютерной науки и содержит одну статью: М.А. Трушникова «Вероятностный анализ нового алгоритма упаковки прямоугольников в полосу». В статье говорится, что в 1993 году Коффман и Шор предложили онлайн-алгоритм упаковки прямоугольников в полосу с оценкой $O(N^{2/3})$ для математического ожидания незаполненной площади полосы в стандартной вероятностной модели. С тех пор вопрос о возможности улучшения этой оценки в классе онлайн-алгоритмов оставался открытым. В данной работе проанализирован принципиально новый онлайн-алгоритм упаковки, предложенный ранее автором. Для него доказана оценка $O(N^{1/2}(\log N)^{3/2})$ для математического ожидания незаполненной площади полосы.

Академик РАН В.П. Иванников