

**СБОРНИК
ТЕХНОЛОГИЙ**

2018

ОГЛАВЛЕНИЕ

6

ИСП РАН – ЭКОСИСТЕМА
ИННОВАЦИЙ

16

СТАТИЧЕСКИЙ
АНАЛИЗАТОР SVACE

19

BINSIDE: СТАТИЧЕСКИЙ
АНАЛИЗАТОР
БИНАРНОГО КОДА

21

ДИНАМИЧЕСКИЙ
АНАЛИЗАТОР ANXIETY

24

ПЛАТФОРМА ИСП РАН
ДЛЯ АНАЛИЗА ПРОГРАММ
НА ОСНОВЕ ЭМУЛЯТОРА
QEMU

27

ИСП ОБФУСКАТОР

30

АНАЛИЗАТОР СЕТЕВОГО
ТРАФИКА PROTOSPHERE

33

KLEVER:
ТЕХНОЛОГИЯ СТАТИЧЕ-
СКОЙ ВЕРИФИКАЦИИ
ПРОГРАММ НА ЯЗЫКЕ GNU
Си

36

ГЕНЕРАТОР ТЕСТОВЫХ
ПРОГРАММ MICROTesk

38

RETRASCOPE:
ИНСТРУМЕНТ ОБРАТНОЙ
ИНЖЕНЕРИИ
HDL-ОПИСАНИЙ

40

СИСТЕМА ВЕРИФИКАЦИИ
ASTRAVER TOOLSET

43

MASIW:
ПОДДЕРЖКА
ПРОЕКТИРОВАНИЯ
ОТВЕТСТВЕННЫХ
СИСТЕМ

46

CONSTRUCTIVITY 4D:
ТЕХНОЛОГИЯ ИНДЕКСИ-
РОВАНИЯ, ПОИСКА
И АНАЛИЗА БОЛЬШИХ
ПРОСТРАНСТВЕННО-
ВРЕМЕННЫХ ДАННЫХ

48

БАЗОВЫЙ СЕМАНТИ-
ЧЕСКИЙ АНАЛИЗАТОР
TEXTERRA

51

ТЕХНОЛОГИЯ АНАЛИЗА
СОЦИАЛЬНЫХ МЕДИА
TALISMAN

54

LINGVODOC:
ВИРТУАЛЬНАЯ
ЛАБОРАТОРИЯ
ДЛЯ ДОКУМЕНТАЦИИ
ИСЧЕЗАЮЩИХ ЯЗЫКОВ

56

СИСТЕМА
ИССЛЕДОВАТЕЛЬСКОГО
ПОИСКА SCINOON

58

КОМПЛЕКС РЕШЕНИЙ ДЛЯ
СОЗДАНИЯ СЕРВИС-
ОРИЕНТИРОВАННЫХ ЦОД

**АРУТЮН АВЕТИСЯН
ДОКТОР ФИЗИКО-МАТЕМАТИЧЕСКИХ
НАУК, ЧЛЕН-КОРРЕСПОНДЕНТ РАН,
ДИРЕКТОР ИСП РАН**



Институт системного программирования Российской академии наук — ведущий центр компетенций в области системного программирования в России. Специалисты ИСП РАН создают технологии мирового уровня, которые позволяют Институту конкурировать с центрами разработки международных ИТ-корпораций и лучшими научно-исследовательскими организациями мира в различных областях системного программирования — анализ кода, верификация, анализ данных, операционные системы и др.

В основе успеха Института — экосистема, которая поддерживает полную цепочку разработки от генерации идей и фундаментальных исследований до технологий и продуктов, готовых к внедрению, и позволяет готовить кадры высшей квалификации в сфере ИТ. Благодаря основателю Института Виктору Петровичу Иванникову, в ИСП РАН сохранена и адаптирована к современным реалиям научная школа, которая начала формироваться еще в советское время. Сегодня она показывает свою жизнеспособность и эффективность в условиях высокой мобильности научных кадров, идей, открытой глобальной конкуренции.

Ключевой механизм удержания передовых позиций — нацеленность на наукоемкие инновации. Они базируются на многолетних научно-исследовательских проектах и долгосрочных партнерских отношениях. Технологии разрабатываются в тесной интеграции с индустрией, внедряются и используются крупными отечественными компаниями и транснациональными корпорациями. Среди долговременных отечественных партнеров

Института – ГосНИИАС, Вымпелком, НИИСИ РАН, РусБИТех, Базальт СПО, МВП Свемел. Среди зарубежных – Samsung, Huawei, Dell EMC, HPE, Intel, Nvidia, Rogue Wave, Bentley Systems (ранее Synchro Software), Linux Foundation, ITRI. Со многими из них созданы совместные лаборатории. ИСП РАН реализует совместные проекты с ведущими университетскими и исследовательскими центрами, такими, как Кембридж (Великобритания), Карнеги-Меллон (США), INRIA (Франция), Университет Пассау (Германия), Технион (Израильский технологический институт) и другие.

Также на базе Института функционируют собственная аспирантура и кафедры в МГУ, МФТИ и ВШЭ. Студенты и аспиранты с самого начала вовлекаются в реальные научные проекты. К моменту выпуска многие студенты имеют научные публикации и уже являются реальными специалистами по системному программированию.

Сборник открывает описание бизнес-модели ИСП РАН, где подробно рассказано, как устроена экосистема Института, раскрыты наши подходы к подготовке кадров. В основной части сборника приводится описание ряда технологий и продуктов ИСП РАН, которые уже используются в индустрии.

Современные наукоёмкие продукты и технологии требуют постоянных инноваций и, как правило, – адаптации под конкретные запросы. Невозможно развивать и внедрять передовые разработки без постоянного взаимодействия с потребителями. В связи с этим сотрудничество с Институтом имеет несомненные плюсы для российских заказчиков. Большинство наших технологий являются уникальными для России и конкурентоспособными на мировом рынке. И мы всегда готовы кастомизировать их в соответствии с запросами потребителей.

В связи с нацеленностью ИСП РАН на тесную работу с промышленными партнерами и спецификой коммерческих проектов, разработки Института долгое время были известны только узкому кругу специалистов. Впервые описанные технологии были продемонстрированы широкой общественности на Первой научно-практической конференции ИСП РАН в конце 2016 г.

ИСП РАН — ЭКОСИСТЕМА ИННОВАЦИЙ

Исследовательская деятельность ИСП РАН нацелена на передачу результатов фундаментальных исследований в промышленность или в другую сферу использования. Это означает, что все виды деятельности института фокусируются на том, чтобы разрабатываемые технологии, программные продукты, методы решения задач системного программирования, создающиеся в Институте, отвечали современным требованиям и были в максимальной степени готовы для внедрения.

Бизнес-модель Института состоит из трех тесно связанных активных элементов, которые в совокупности дают синергетический эффект:

- проектно-ориентированные фундаментальные и прикладные исследования в области системного программирования (по контрактам с российскими и зарубежными компаниями, Минобрнауки РФ, Программам РАН, грантам РФФИ и ФПИ и т.п.), нацеленные, в первую очередь, на создание новых технологий;
- инновации — проектные внедрения результатов передовых исследований в промышленных компаниях-партнерах. Инновационный продукт может появиться только после получения обратной связи от индустрии;
- образование — обучение студентов и аспирантов на основе разработанных и используемых в работах ИСП РАН современных технологий, с обязательным участием студентов в исследовательских и промышленных проектах института.

Такая модель хорошо известна и применяется как в исследовательских лабораториях ведущих университетов (Stanford, MIT, Berkeley, Carnegie Mellon), так и в лабораториях индустриальных гигантов (например, IBM и Intel) и в государственных исследовательских центрах, таких как INRIA (Франция), Fraunhofer (Германия) и других. Эффективно реализованная, модель позволяет решить проблему разрыва между наукой и промышленностью и готовить кадры высшей квалификации, способные создавать и внедрять новые технологии.

ФУНДАМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

В модели Института представлены все элементы цепи от генерации идей и фундаментальных исследований до технологий и продуктов, готовых к внедрению. При этом фундаментальные исследования, проведение поисковых и экспериментальных работ — необходимые элементы деятельности, позволяющие двигаться в русле самых новых идей и еще только зарождающихся технологий. Фундаментальные исследования — это также источник идей для новых проектов с бизнес-партнерами и заказчиками. Все инновации, созданные в Институте, базируются на фундаментальных исследованиях.

ИСП РАН ведет большое число научных и образовательных программ в тесной кооперации как с ведущими российскими, так и зарубежными научными и университетскими центрами, в частности, с университетами Кембриджа (Великобритания), Карнеги-Меллон (США), INRIA (Франция), университетом Пассау (Германия), Израильским технологическим институтом Технион и другими. Это позволяет обеспечивать высокий уровень результатов исследований. Кроме того, репутация в академических и университетских кругах открывает перспективу внедрения отечественных технологий на международных рынках.

ВНЕДРЕНИЕ

ИСП РАН не использует модель стартапов для внедрения результатов исследований. Вместо этого Институт развертывает внедрение своих результатов через крупные промышленные и исследовательские организации, которые одновременно являются пользователями новых технологий и партнерами по продвижению разработок института в широкую практику.

В настоящее время более 90% работ по контрактам ведутся с отечественными и зарубежными организациями, которые являются долговременными партнерами ИСП РАН (5 лет и более). Из зарубежных партнеров Института по инновационной активности можно назвать Samsung, HP, Intel, Nvidia, Rogue Wave, Bentley Systems (ранее Synchro Software), ITRI и Linux Foundation, из отечественных — ГосНИИАС, МВП Свемел, Базальт СПО, Вымпелком. Доля отечественных компаний в общем объеме финансирования устойчиво растет на протяжении последних десяти лет.

НАУЧНОЕ СОТРУДНИЧЕСТВО

Одна из форм, которая применяется в ИСП РАН для организации долгосрочного сотрудничества, – совместные лаборатории. Они позволяют при наличии постоянного финансирования гибко планировать имеющиеся ресурсы, а также наращивать компетенции во вновь образующихся направлениях системного программирования, организовывать подготовку молодых специалистов с компетенциями в интересующих партнеров областях.

В настоящее время в Институте функционируют совместные лаборатории с такими компаниями, как Samsung (Южная Корея) – нацелена на компиляторные технологии, в том числе, на обеспечение безопасности в контексте мобильных платформ (ARM, Android, Tizen); Rogue Wave (США) – технологии статического анализа кода на выявление уязвимостей безопасности; Dell (США) – анализ и обработка больших данных; Bentley Systems (ранее Synchro Software, Великобритания) – комплексное планирование и nD-моделирование масштабных проектов; CUDA Research Center компании NVIDIA – технологии параллельных вычислений для современных и перспективных гетерогенных систем сверхвысокой производительности. Ведутся совместные проекты с ведущими университетскими и исследовательскими центрами, такими как Кембридж (Великобритания), Карнеги-Меллон (США), INRIA (Франция), Университет Пассау (Германия) и др.

На базе технологической сервисно-ориентированной облачной платформы Fanlight создана и успешно функционирует лаборатория для решения задач механики сплошных сред. В ней совместно работают системные и прикладные программисты, что позволяет эффективно вести исследовательские проекты в интересах промышленных предприятий. Такая лаборатория обеспечивает интеграцию науки, образования и индустрии на современном технологическом уровне.

ИНТЕЛЛЕКТУАЛЬНАЯ СОБСТВЕННОСТЬ

В модели, созданной ИСП РАН, права на интеллектуальную собственность остаются за ИСП РАН при использовании собственных технологий и имеющегося задела по фундаментальным исследованиям, либо могут передаваться в рамках специальных соглашений (например, с Free Software Foundation) сообществу разработчиков свободного ПО. С учетом специфики бизнес-модели Института была разработана оригинальная лицензия, базирующаяся не на получении роялти, а на прямом финансировании со стороны заказчика дальнейших исследований и разработок, направленных на развитие технологии. Заказчику передаются неисключительные права по использованию, исключительные права остаются за Институтом.

В особых ситуациях решение по управлению правами интеллектуальной собственностью принимается индивидуально с учетом перспектив долгосрочного развития направления исследований и коллектива Института в целом. Примером такого исключения может являться контракт с Фондом перспективных исследований (ФПИ), по которому все права должны быть переданы заказчику (ФПИ), в том числе, заказчику будут переданы неисключительные права на патент, который принадлежит ИСП РАН и будет использоваться в данном проекте.

СВОБОДНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Один из важнейших компонентов созданной экосистемы – широкое использование свободного программного обеспечения (СПО), без которого невозможно представить современное системное программирование. СПО Институтом рассматривается как:

- инструмент, предоставляющий легитимный свободный доступ ко всему многообразию современных технологий, включая готовые к использованию программные продукты, технологии и открытые стандарты;
- возможность взаимодействия с глобальным рынком продуктов и услуг, что позволяет вести инновационное развитие вместо аутсорсинга;
- мощный образовательный ресурс – среда и инфраструктура международных СПО-проектов могут использоваться для подготовки специалистов высокой квалификации.

Научная деятельность подразумевает открытость научного результата и «видимость» автора этого результата, что часто приходит в противоречие с корпоративной политикой ИТ-компаний. Для ИСП РАН открытость результатов исследования, в частности, активное использование возможностей модели открытого программного кода, является одновременно и стимулом к работе, и инструментом продвижения Института и технологий, которые в нем разрабатываются. Открытость приводит к тому, что каждый молодой исследователь, даже если он работает в большом коллективе, «виден» в международном сообществе ИТ-специалистов. Его вклад, его репутация – это его капитал, а Институт делает все возможное, чтобы этот капитал рос максимально быстро.

ОБРАЗОВАНИЕ

Краеугольный камень экосистемы инноваций ИСП РАН – образовательная деятельность.

- Интеграция ИСП РАН с ведущими вузами: кафедры в МГУ, МФТИ и ВШЭ. Каждый год через кафедры в Институт приходят 50-60 студентов-третьекурсников. В первый год обучения в ИСП РАН они слушают лекции специалистов, занимаются на спецсеминарах и знакомятся с исследовательской тематикой

по научным направлениям Института. Во второй год обучения в ИСП РАН студенты уже участвуют в реальных проектах Института. К моменту выпуска многие студенты имеют научные публикации и уже являются реальными специалистами по системному программированию в своих областях исследований.

- Аспирантура ИСП РАН. Учеба в аспирантуре ИСП РАН – это одновременно накопление практического опыта и изучение новых технологий. Кроме того, аспиранты активно вовлекаются в процессы обучения. Они ведут семинарские и практические занятия со студентами, руководят подготовкой курсовых и дипломных работ. Накопив такой опыт, выпускник аспирантуры, как правило, уже становится руководителем небольшой исследовательской группы.

За десять лет количество сотрудников, вовлеченных в образовательную деятельность, увеличилось больше, чем в два раза. Было разработано более десятка новых курсов и значительно переработаны существующие. С учетом высокого приоритета образовательной деятельности с целью повышения эффективности учебных курсов наряду с сотрудниками института, официально участвующими в преподавании, Институт привлекает к этой деятельности примерно столько же сотрудников дополнительно.

Начиная с первого года обучения в ИСП РАН, студенты начинают получать стипендию, а со второго года – зарплату, которая в настоящее время сопоставима с зарплатами в высокотехнологичных ИТ-компаниях. Вовлечение студентов и аспирантов в реальные исследовательские проекты – очень серьезная мотивация, привлекающая молодых специалистов в созданную Институту экосистему. Исследовательская работа подразумевает, что научная проблема или задача, которую предстоит решить, является настоящим вызовом, она требует не только владения самыми передовыми технологиями, но требует выхода за границы известных методов решений подобных задач.

Институт совместно с компанией Samsung Electronics открыл на базе Московского физико-технического института «IoT Академию Samsung», в рамках которой по разработанным учебно-методическим материалам учащиеся проходят годичный учебный курс, направленный на изучение реальных случаев использования технологий Интернета вещей в различных отраслях, а также создают собственные прототипы IoT-устройств.

ИСП РАН получил и развивает успешный опыт распространения своей экосистемы за рамки Института. Под научным руководством ИСП РАН осуществляется работа трёх лабораторий на базе других организаций:

- Лаборатория системного программирования имени В.П. Иванникова Российско-Армянского (Славянского) государственного университета (РАУ, Армения) занимается исследованием и разработкой программ интеллектуального анализа текстов;
- Лаборатория системного программирования

- Ереванского государственного университета (ЕГУ, Армения);
- Лаборатория системного программирования Новгородского государственного университета имени Ярослава Мудрого.

Сотрудники этих лабораторий выполняют совместные с ИСП РАН промышленные и исследовательские проекты, образуя распределенную среду разработки и исследований, работают над совместными публикациями и выступлениями на конференциях, читают современные курсы в своих университетах. Молодые специалисты лабораторий проходят в ИСП РАН очную стажировку, в том числе семестровые стажировки для подготовки бакалаврских и магистерских дипломных работ.

ИСП РАН В ЦИФРАХ

В общем объеме финансирования около 2160 млн. руб., привлечённых за последние пять лет на выполнение фундаментальных и прикладных исследовательских работ, доля договорных работ составила более 87%. Отметим, что в общем объеме договорных работ доля работ с отечественными организациями составила 42,8%. При этом, в 2018 году средняя зарплата научного сотрудника составляет 220 080 руб.

На июль 2018 года среднесписочная численность работников Института составляет 209,2. Численность научных сотрудников, выполняющих исследования и разработки, составляет 216 человек (из них 14 докторов и 44 кандидатов наук, более 10 профессоров). Фактически количество научных сотрудников с учетом оформления части из них на неполную ставку значительно увеличилось. При этом лица моложе 39 лет составляют 73% от общего числа научных сотрудников. Отметим, что в связи с нехваткой штатных единиц с 2008 г. введены внебюджетные штатные единицы. В 2011 г. в рамках правительственной программы «1000 штатных единиц молодым ученым» Институт получил 5 дополнительных бюджетных единиц. Несмотря на это, в настоящее время более 50% сотрудников Института работает во внебюджетном штате и есть постоянная тенденция к его увеличению.

Постоянно растет количество сотрудников Института, вовлеченных в образовательную деятельность (на текущий момент более 30). Было разработано свыше 10 новых курсов и значительно переработаны существующие. Надо отметить, что фактическое финансирование всех этих работ ведётся за счет Института. При этом с учетом высокого приоритета образовательной деятельности с целью повышения эффективности учебных курсов наряду с сотрудниками института, официально участвующими в преподавании, Институт привлекает к этой деятельности примерно столько же сотрудников дополнительно.

Существенно выросло количество организуемых Институтом всероссийских и международных семинаров

и конференций: в 2017 г. – 8, из них 5 международных, часть из которых поддержаны международными сообществами IEEE и IEEE Computer Society.

Сотрудники Института приняли участие в работе более 300 ведущих российских и зарубежных конференций. Опубликовано свыше тысячи научных статей и десятки монографий, значительная часть которых размещаются в электронной библиотеке IEEE Xplore и индексируются в международных индексах научного цитирования Web of Science и Scopus, получено около 70 свидетельств РИД, 2 программы включены в Единый реестр российских программ для электронных вычислительных машин и баз данных Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации.

Молодые сотрудники Института систематически получают стипендии Правительства и Президента РФ. Также двумя молодыми учеными были получены медали РАН с премиями. Двоим сотрудникам присуждены Госпремии Российской Федерации. Институт получил грант Президента РФ по поддержке ведущих научных школ «Методы и средства обеспечения информационной безопасности».

Издание Института «Труды Института системного программирования РАН» было признано научным сообществом, в том числе – международным. Труды ИСП РАН включены в список ВАК с 2010 г., и в настоящее время находятся в процессе включения в международную систему Scopus. Часть статей включается в английскую версию журнала «Программирование» (издательство Springer).

КОРОТКО О ГЛАВНОМ

В совокупности созданная в ИСП РАН экосистема, базируясь на передовых подходах к организации исследований и разработок, в том числе, широком использовании СПО и управлении интеллектуальной собственностью, обладает высоким уровнем адаптивности и динамизма, позволяя адекватно реагировать на вновь возникающие технологические и организационные вызовы. Это значительно расширяет горизонты планирования и позволяет строить отношения с заказчиками и партнерами на долговременной основе, существенно снижая риски для обеих сторон. Это важно, так как прямое взаимодействие с организациями реального сектора экономики является ключевой составляющей экосистемы Института. Без такого сотрудничества невозможно достичь и постоянно удерживать высокий технологический уровень исследований и разработок.

За время существования Института накоплен опыт и получены результаты мирового уровня в ряде передовых областей системного программирования, включая разработку собственных технологий и инструментов, успешно выполнены десятки производственных проектов, подготовлены сотни студентов и аспирантов, защищено более 50 докторских и кандидатских диссер-

таций, выпущены сотни публикаций, получены патенты в области ИКТ. Работа ИСП РАН как эффективного научного и производственного центра в области системного программирования получила мировое признание, что, в частности, выражается во включении ученых ИСП РАН в программные комитеты многих международных конференций, предложениях выступить в качестве приглашенных докладчиков, участии в экспертных группах международных консорциумов. При ИСП РАН созданы два центра компетенции по специализированным направлениям – Центр верификации ОС Linux и Центр по параллельным и распределенным вычислениям, которые успешно интегрированы в российское и мировое сообщество и во все виды деятельности по модели ИСП РАН.

ТЕХНОЛОГИИ

СТАТИЧЕСКИЙ АНАЛИЗАТОР

1/17

SVACE Умный поиск
ошибок и уязвимостей

Svace – необходимый инструмент жизненного цикла разработки безопасного ПО, основной статический анализатор компании Samsung. Обнаруживает более 50 классов критических ошибок в исходном коде. Поддерживает языки C, C++, C# и Java. Зарегистрирован в Едином реестре российского ПО.

ПОЧЕМУ SVACE?

- 1** Максимальная гибкость и адаптивность под российского заказчика

Svace – необходимый инструмент жизненного цикла разработки безопасного ПО, основной статический анализатор компании Samsung. Обнаруживает более 50 классов критических ошибок в исходном коде. Поддерживает языки C, C++, C# и Java. Зарегистрирован в Едином реестре российского ПО.

 - Ускоренная кастомизация (конфигурация существующих детекторов, а также написание индивидуальных, доступных только данному заказчику; создание специфических интерфейсов);
 - Ускоренная адаптация к работе с новым окружением (добавление новых компиляторов в течение 1-2 недель, в сложных случаях до 2 месяцев);
 - Непрерывное обучение разработчиков заказчика, постоянное взаимодействие и технологическая доработка продукта в соответствии с новыми задачами во время всего жизненного цикла разработки безопасного ПО;
 - Гибкие условия лицензирования с учетом потребностей заказчика (в частности, возможность получить исходный код продукта);
 - Полная совместимость с нормативными документами и требованиями регуляторов (ФСТЭК РФ).
- 2** Качество на уровне мировых аналогов

Svace – постоянно развивающийся инновационный продукт, основанный на многолетних исследованиях. Он объединяет ключевые качества иностранных конкурентов (Coverity Static Analysis, HP Fortify, RogueWave Klocwork Static Analysis) с уникальным использованием открытых промышленных компиляторов в целях максимальной поддержки новых стандартов языков программирования.

- Svace — это:
 - Качественный глубокий анализ:
 - точное представление исходного кода (благодаря интеграции с любой системой сборки);
 - полное покрытие всех путей с учетом связей между функциями для поиска сложных ошибок;
 - высокий процент истинных срабатываний (60-90%).
 - Масштабируемость и высокая скорость:
 - параллельный анализ с использованием до 64 процессорных ядер;
 - возможность анализировать системы из десятков миллионов строк кода (анализ Android 6 из 8 миллионов строк занимает 5-6 часов);
 - поддержка не только полного, но и инкрементального анализа системы (подразумевает быструю повторную проверку недавно измененного кода).
 - Удобный интерфейс просмотра предупреждений:
 - подробное описание ошибок с навигацией по коду;
 - разделение срабатываний на истинные и ложные;
 - миграция результатов между запусками и сокрытие ложных срабатываний.

ДЛЯ КОГО ПРЕДНАЗНАЧЕН SVACE?

- Компании, нацеленные на разработку ПО с особым акцентом на высокую надежность и безопасность;
- Компании, которые нуждаются в сертификации разрабатываемого ПО;
- Сертификационные лаборатории.

SVACE И SAMSUNG

Svace — основной анализатор Samsung с 2015 года. Применяется для проверки собственного ПО компании на базе ОС Android и исходного кода ОС Tizen, которая используется в смартфонах, информационно-развлекательных системах и бытовой технике Samsung. С 2017 года Svace проверяет все изменения, присланные для рецензирования и включения в ОС Tizen.

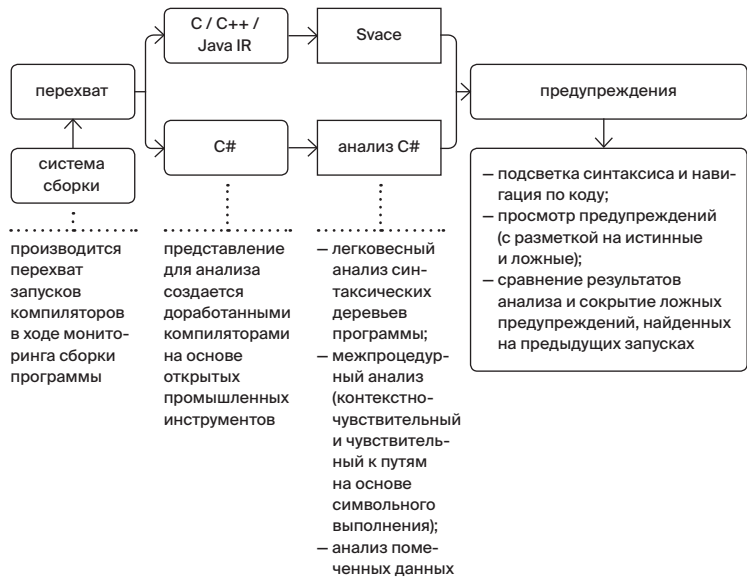
ПОДДЕРЖИВАЕМЫЕ ПЛАТФОРМЫ И АРХИТЕКТУРЫ

- Платформы, на которых работает анализатор: ОС на базе ядра Linux (начиная с версии 2.6), ОС Windows (начиная с XP).
- Целевые архитектуры для анализируемого кода: Intel x86/x86-64, ARM, ARM64, MIPS, MIPS64, Power PC, Hexagon.

ПОДДЕРЖИВАЕМЫЕ КОМПИЛЯТОРЫ

- Для C/C++: GCC (GNU Compiler Collection), Clang (LLVM compiler), Microsoft Visual C++ Compiler, RealView/ARM Compilation Tools (ARMCC), Intel C++ Compiler, Wind River Diab Compiler, NEC/Renesas CA850, CC78K0(R) C Compilers, C/C++ Compiler for the Renesas M16C Series and R8C Family, Panasonic MN10300 Series C Compiler, C compiler for Toshiba TLCS-870 Family, Samsung CalmSHINE16 Compilation Tools, Texas Instruments TMS320C6* Optimizing Compiler и др.
- Для C#: Roslyn, Mono.
- Для Java: OpenJDK Javac Compiler, Eclipse ECJ compiler, Jack Compiler for Android.

СХЕМА РАБОТЫ



BINSIDE: СТАТИЧЕСКИЙ АНАЛИЗАТОР БИНАРНОГО КОДА

Ищет ошибки
без доступа
к исходному коду

BinSide – инструмент обнаружения дефектов в программе методами статического анализа исполняемого кода. Необходим, когда нет доступа к исходному коду (например, при анализе закрытых библиотек).

ПОЧЕМУ BINSIDE?

- 1 **Высокотехнологичный подход к анализу кода**

BinSide – инструмент для анализа бинарного кода, разрабатываемый на основе фреймворка BinNavi, который переводит ассемблерный код в представление REIL, чей язык позволяет анализировать код независимо от процессорной архитектуры и операционной системы.

BinSide – это:

 - Лёгкая расширяемость:
 - детекторы отдельных ошибок пишутся как плагины, которые можно оперативно добавлять и менять;
 - используется представление REIL из 17 инструкций без побочных эффектов (каждая ассемблерная инструкция транслируется в набор из REIL-инструкций).
 - Плагины для самых критичных типов ошибок (в том числе, уязвимостей форматной строки и ошибок работы с указателями);
 - Поиск двух типов переполнения буфера (основным критерием наличия уязвимости является возможность получения контроля над входным буфером):
 - возникает при копировании информации из большего буфера в меньший (например, при использовании таких небезопасных функций, как `strcpy`, `memcpy` и т. д.);
 - возникает при копировании одного буфера в другой без проверки границ первого буфера (например, пока не встретится символ конца строки).
 - Мощный гибкий движок с основными типами анализа:
 - интервальный анализ, отслеживание помеченных данных, модели статической и динамической памяти, а также анализ графов потока данных и управления;

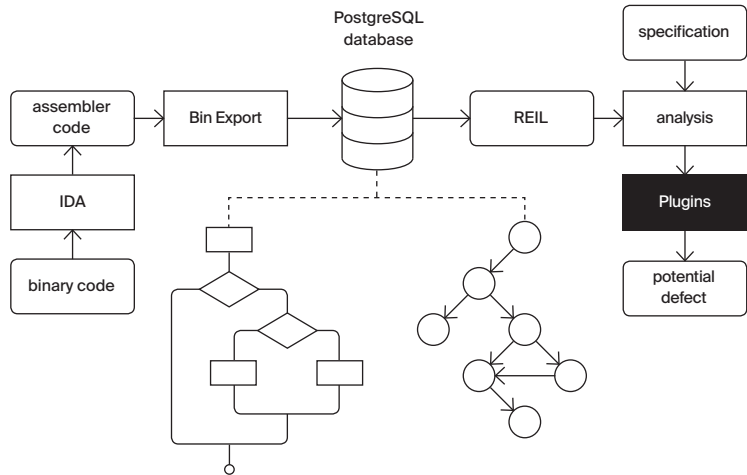
- поиск и нахождение ошибок на всех путях (в том числе, не покрытых тестированием или динамическим анализом).
- Плагин для статического анализа патчей программ;
- Высокая скорость работы (на бинарном файле из более 3000 функций длится около 2000 секунд).
- Навигация по графам потока управления, потока данных и вызовов программы в рамках среды BinNavi.

2 Особое удобство для пользователя BinSide интегрирован с интерактивным дизассемблером IDA PRO, который используется для обратной разработки.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

Поддерживает анализ бинарных файлов и библиотек архитектур x86, x64, ARM, PowerPC и MIPS.

СХЕМА РАБОТЫ



ДИНАМИЧЕСКИЙ АНАЛИЗАТОР

3/17

ANXIETY

Комбинированный
подход к поиску
ошибок

Anxiety – среда для обнаружения ошибок и потенциально опасных ситуаций в процессе разработки, приёмочного тестирования и эксплуатации ПО. Работает на основе динамического символического выполнения, позволяющего автоматически генерировать входные данные без наличия исходного кода и отладочной информации.

ПОЧЕМУ ANXIETY?

1 Сочетание
важнейших
функций

Особенность Anxiety – комбинированный подход к динамическому анализу, который заключается в интеграции со статическими анализаторами и инструментами фаззинга. Удачное сочетание технологий позволяет Anxiety решать те же задачи, что и ведущие мировые аналоги (CA Veracode Dynamic Analysis, Synopsys Dynamic Application Security Testing и Rogue Wave CodeDynamics)

Anxiety – это:

- Высокий уровень производительности анализа за счет поддержки распределённого и параллельного режимов работы, интеграции с фаззером, а также фильтров на поток входных данных и анализируемые функции;
- Создание инструментов анализа, основанных на методе динамического символического выполнения;
- Интеграция со статическими анализаторами исходного или машинного кода для реализации направленного анализа, позволяющего выборочно тестировать компоненты целевой программы. Проверка дефектов, ранее обнаруженных с помощью статического анализа (в частности, обнаружение деления на ноль, разыменования нулевого указателя, зацикливания, нарушения пользовательских утверждений и др.);
- Интеграция с инструментами рандомизированного тестирования программ (фаззинга) для повышения его эффективности. В частности, интеграция решает проблемы при прохождении условных переходов, зависящих от сравнения с константами. Использование фаззинга позволяет добиться покрытия исходного кода программы тестовыми наборами входных данных значительно быстрее, чем динамическое символическое выполнение;
- Модульная инфраструктура (трассировщик, чекер и генератор входных данных), позволяющая производить замену компонентов системы и расширять её функциональность;

- Поддержка различных источников внешних данных программы (файлы, сетевые сокеты, переменные окружения, стандартный поток ввода).
- 2** Удобство для
российского
заказчика
- Anxiety – высокотехнологичная разработка ИСП РАН, основанная на результатах многолетних исследований и предназначенная для промышленного использования. Гибкая базовая среда с возможностью всесторонней адаптации под нужды заказчика. В числе преимуществ:
- Реализация специфических задач анализа программ на базе динамического символьного исполнения (в частности, определение достижимости определенной функции или операции в программе);
 - Возможность получить отчуждаемый продукт;
 - Возможность использования для реализации обеспечительных мер ГОСТ Р 56939-2016 (при необходимости сертификации ПО для использования на территории России).

ДЛЯ КОГО ПРЕДНАЗНАЧЕНА ANXIETY?

- Компании, нацеленные на разработку ПО с особым акцентом на высокую надежность и безопасность;
- Компании, отвечающие за аудит или сертификацию ПО.

ГДЕ ИСПОЛЬЗУЕТСЯ ANXIETY?

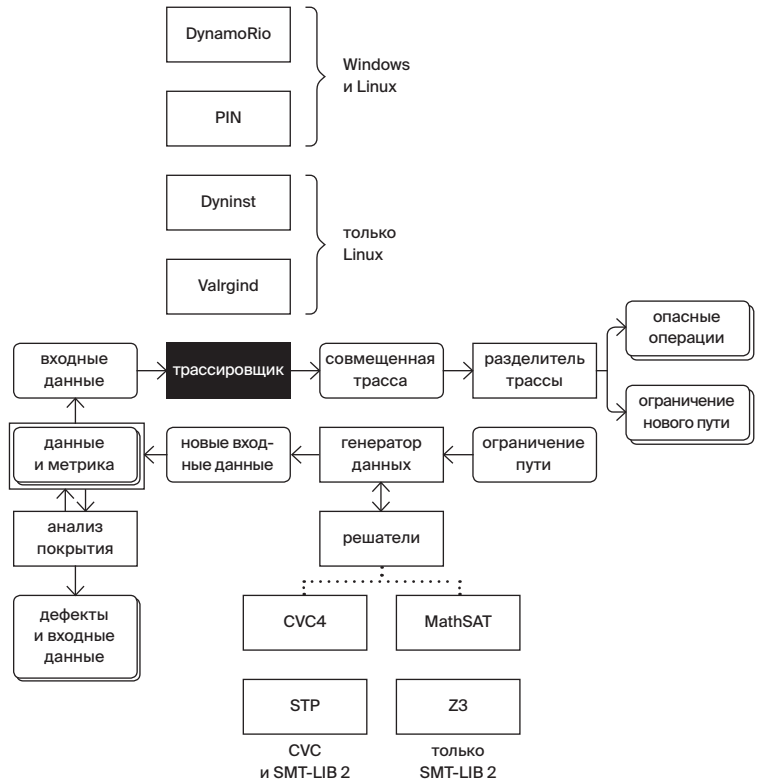
Инструмент Anxiety используется при тестировании программ, входящих в поставку ОС Astra Linux.

ПОДДЕРЖИВАЕМЫЕ СРЕДЫ И ИНСТРУМЕНТЫ

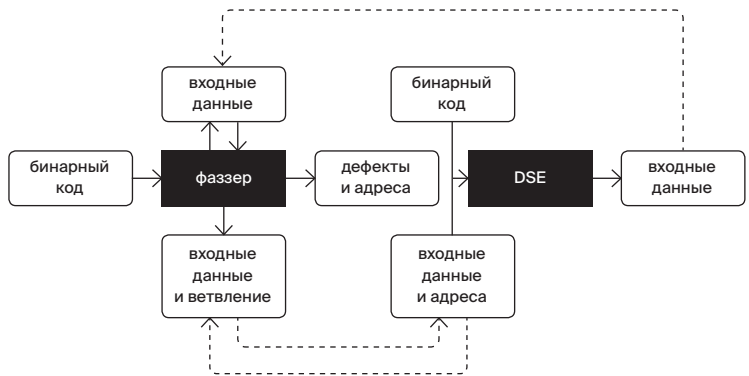
Поддерживает анализ в ОС Windows (начиная с версии XP) и ОС семейства Debian Linux, а также работу различных типов SMT-решателей (STP, Z3, MathSAT и др.). Основана на средах динамической инструментации DynatomoRIO (что обеспечивает поддержку ОС Windows, поток инструкций обрабатывается инструментом Triton) и динамической бинарной трансляции Valgrind, для которой разработаны плагины сбора трассы и вычисления покрытия базовых блоков.

СХЕМА РАБОТЫ

Динамический
символьный
анализ



Фаззинг



ПЛАТФОРМА ИСП РАН ДЛЯ АНАЛИЗА ПРОГРАММ НА ОСНОВЕ ЭМУЛЯТОРА



Тщательный
анализ работы
программ

Платформа ИСП РАН для анализа программ построена на базе открытого эмулятора QEMU, который используется при необходимости кроссплатформенной разработки.

ПОЧЕМУ ПЛАТФОРМА НА ОСНОВЕ QEMU?

1 Широкий ряд
возможностей

В полносистемном режиме эмулятор Qemu поддерживает более 10 семейств процессоров (i386, x86-64, ARM, MIPS, PowerPC и другие платформы). Может применяться для создания новых виртуальных платформ, отладки загрузчиков и ядерного кода ОС (в том числе, драйверов эмулируемого оборудования), анализа вредоносного ПО, а также записи работы виртуальной машины для последующего воспроизведения. Поддерживает удалённую отладку через IDA, GDB или Eclipse.

Особенностью Платформы ИСП РАН является уникальное сочетание трёх реализованных технологий:

- Воспроизведение работы виртуальной машины, необходимое при анализе или отладке сложно воспроизводимого дефекта или сценария работы:
 - возможность разовой записи возникновения состояния гонки в многопоточном приложении или работы сетевого приложения с сервером и последующей

- многократной отладки сценария в GDB, Eclipse или IDA (без перенастройки или создания тех же условий возникновения дефектов);
 - возможность анализировать драйверы и прошивки (благодаря записи работы BIOS и ядра);
 - поддержка жесткого диска, клавиатуры и мыши, работы с сетью, аудиоадаптером и последовательным портом.
 - Обратная отладка:
 - использование команд обратной отладки GDB на основе записанного сценария;
 - использование промежуточных снимков состояния системы (QEMU переходит от снимка с снимку и воспроизводит записанный сценарий, чтобы определить последнее место срабатывания точки останова).
 - Анализ программ:
 - использование механизма интроспекции, который извлекает высокоуровневую информацию о гостевой системе без вмешательства в её работу (инструмент с поддержкой интроспекции находится в открытом доступе: <https://github.com/ispras/qemu/tree/plugins>);
 - возможность получить последовательность совершаемых системных вызовов, обращений к именованным функциям (в .dll и .so), список работающих процессов, а также список открытых файлов и загруженных в память модулей;
 - возможность запустить любой образ виртуальной машины на основе Linux (например, прошивку маршрутизатора) и получить, в частности, последовательность вызовов функций сразу для всей системы.
- 2 Удобство для пользователя
- Свободное расширение возможностей QEMU благодаря открытому исходному коду;
 - Анализ бинарного кода без внедрения программ в гостевую систему;
 - Модульная структура механизма интроспекции с возможностью расширения за счет новых плагинов;
 - Удобное API для самостоятельной разработки плагинов интроспекции;
 - Возможность адаптации под конкретные нужды заказчика;
 - Поддержка актуальных версий QEMU с новой периферией и процессорными ядрами.

ПОДДЕРЖИВАЕМЫЕ ГОСТЕВЫЕ СРЕДЫ

- Эмулируемые платформы: i386, x86-64, ARM, MIPS, PowerPC.
- Гостевые системы, поддерживаемые интроспекцией: Windows XP/7/8/10 и Linux 2.x-4.x на платформе x86.

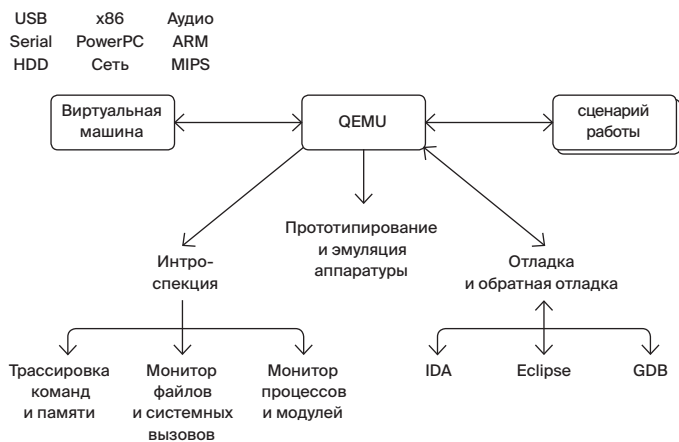
ДЛЯ КОГО ПРЕДНАЗНАЧЕНА ПЛАТФОРМА ИСПРАН НА БАЗЕ QEMU?

- Для разработчиков драйверов;
- Для разработчиков ОС и системного ПО;
- Для аналитиков вредоносного программного обеспечения;
- Для специалистов по сертификации программного обеспечения.

ОПЫТ ВНЕДРЕНИЯ

Реализованный механизм воспроизведения был принят мировым сообществом разработчиков QEMU и включен в версию 3.1.

СТРУКТУРА ПЛАТФОРМЫ



ОБФУСКАТОР

Обфускатор – комплекс технологий по противодействию массовой эксплуатации уязвимостей, возникающих в результате ошибок или закладок. Если злоумышленник смог атаковать одно из устройств с одинаковым ПО, остальные останутся под защитой благодаря изменениям, внесённым в код.

ПОЧЕМУ ОБФУСКАТОР?

- 1** Быстрая доработка в соответствии с нуждами заказчика
 - Обфускатор основан на многолетних исследованиях запутывающих преобразований кода. Благодаря наличию всех необходимых компетенций, наша команда способна создать индивидуальное промышленное решение для любого заказчика:
 - Оперативная адаптация технологии под конкретный компилятор или инструментатор бинарного кода (в среднем доработка занимает 2 месяца, в сложных случаях – до полугода);
 - Возможность получить полностью отчуждаемый продукт;
 - Бесплатное тестирование и демонстрация работы (с использованием программы заказчика или какой-либо другой);
 - Непрерывная настройка преобразований и технологическая доработка продукта в соответствии с новыми задачами и вызовами.
- 2** Оптимальное сочетание необходимых функций и особенностей
 - Обфускатор защищает систему от массовой эксплуатации уязвимостей с помощью различных методов диверсификации кода и позволяет собирать код полного дистрибутива ОС.
 - Обфускатор – это:
 - Тонкая настройка баланса между степенью запутывания и уровнем производительности (при применении с целью защиты от обратного анализа). Минимальное замедление работы – в 1,2 раза, максимальное – в 8 раз;
 - Полная автоматизация (не требуется специальная подготовка исходного кода программы и дополнительные усилия со стороны билд-инженеров заказчика);
 - Использование набора открытых компиляторов GCC, который позволяет корректно собирать код полного дистрибутива ОС;
 - Два метода диверсификации:
 - Динамическая диверсификация кода при запуске программы. Применяется, когда заказчику обязательно

нужен один и тот же код на всех устройствах (например, из-за обязательной сертификации). Этот метод позволяет перемещать до 98% кода с небольшим увеличением его объёма и ухудшением производительности примерно на 1,5%. Преимущества Обфускатора по сравнению с аналогичными продуктами:

- Перемешивание до функции (в отличие от технологий ASLR и Pagerando, которые перемещают только крупные блоки кода);
 - Перемешивание функций во всей системе, кроме ядра, а также отсутствие вероятности конфликта с антивирусами (преимущества перед аналогичной технологией Selfrando, разработанной для Tor Browser);
- Статическая диверсификация кода. Каждый раз при компиляции в зависимости от заданного ключа получается новый исполняемый файл. Преимущества данного метода:
- не увеличивается объём бинарного кода (в частности, важно для интернета вещей);
 - ухудшение производительности стремится к нулю;
 - благодаря работе внутри компилятора, а не постфактум в компоновщике, можно применять расширенный набор диверсифицирующих преобразований и более гибко его настраивать.
- Бесконфликтное совмещение с другими средствами защиты ПО (в том числе, с системным механизмом ASLR).

ДЛЯ КОГО ПРЕДНАЗНАЧЕН ОБФУСКАТОР?

- Разработчики специализированных дистрибутивов операционных систем;
- Разработчики прикладного ПО.

ГДЕ ИСПОЛЬЗУЕТСЯ ОБФУСКАТОР?

ИСП Обфускатор внедрен в ОС «Циркон», которую используют МИД и Пограничная служба ФСБ России.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

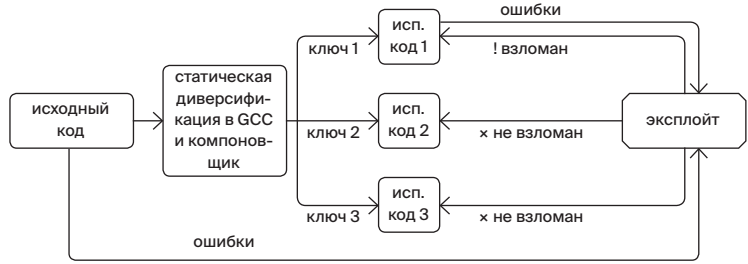
Обфускатор – универсальный продукт, который можно адаптировать под любые системные требования. В настоящее время основная версия работает в ОС на базе ядра Linux (начиная с версии 2.6) с поддержкой архитектуры Intel x86/x86-64.

СХЕМА РАБОТЫ

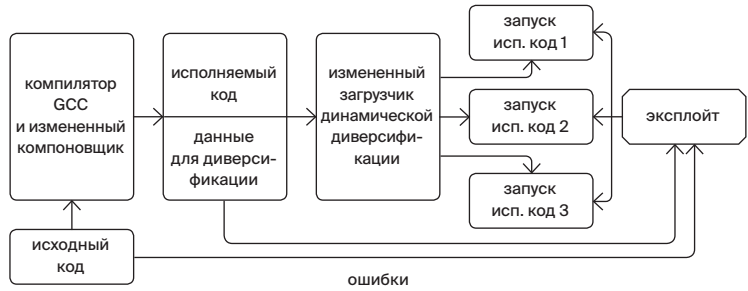
обычная сборка



статическая диверсификация



динамическая диверсификация



АНАЛИЗАТОР СЕТЕВОГО ТРАФИКА

6/17

Анализирует трафик,
выявляет аномалии

PROTOSPHERE

Protosphere – система глубокого анализа сетевого трафика (DPI). Часть защиты от вторжений и утечек информации. Выявляет несоответствия между спецификацией протокола и конкретной реализацией. Позволяет быстро добавлять поддержку новых (в том числе, и закрытых) протоколов благодаря универсальности внутреннего представления.

ПОЧЕМУ PROTOSPHERE?

1 Оптимальное сочетание необходимых функций

Protosphere – инновационная система, основанная на научных исследованиях технологий анализа сетевого трафика. Объединяет ключевые особенности иностранных конкурентов (Wireshark, Microsoft Network Monitor) с универсальным внутренним представлением, позволяющим быстро расширять возможности анализа.

Protosphere – это:

- Оптимальные возможности ядра системы:
 - универсальная модель представления данных в процессе разбора сетевого трафика;
 - обработка данных, содержащих искажения, потери, перестановки и дублирование пакетов, а также асимметричный трафик;
 - поддержка анализа сжатых и зашифрованных данных;
 - поддержка туннелей произвольной конфигурации.
- Поддержка всех этапов исследования сетевого трафика за счёт синхронизированных средств визуализации:
 - локализация одного или нескольких исследуемых сетевых соединений на графе сетевых взаимодействий и дереве сетевых потоков;
 - детализация выделенных соединений на временной диаграмме;
 - наглядное представление результатов разбора на дереве разбора;
 - диагностика несоответствий между спецификацией протокола и фактическим трафиком в журнале диагностики сбоя;
 - извлечение и анализ данных, в том числе, приклад-

- ного уровня за счёт совместного использования окон содержимого данных, списка фрагментов и списка объектов.
 - Быстрое расширение списка поддерживаемых протоколов:
 - API доступа к результатам разбора;
 - локализация ошибок разбора;
 - возможность отладки разрабатываемого модуля на потоке, позволяющая существенно ускорить внедрение новых протоколов.
 - Поддержка двух режимов работы: на потоке и в отложенном режиме. За счёт единой кодовой базы ускоряется внедрение и проверка новых протоколов;
 - Развитый графический интерфейс, который позволяет выбирать наиболее удобный вариант представления результатов проводимого анализа.
- 2** Гибкость и адаптивность под нужды заказчика
- Ускоренная кастомизация благодаря универсальности внутреннего представления (поддержка новых протоколов, извлечение новых типов данных, настройка формата выдачи результатов анализа);
 - Адаптация под сетевой канал и доступные вычислительные ресурсы (гибкая система конфигурирования позволяет находить баланс между детализацией и точностью анализа и потребляемыми ресурсами);
 - Возможность заказчика получить отчуждаемый продукт.

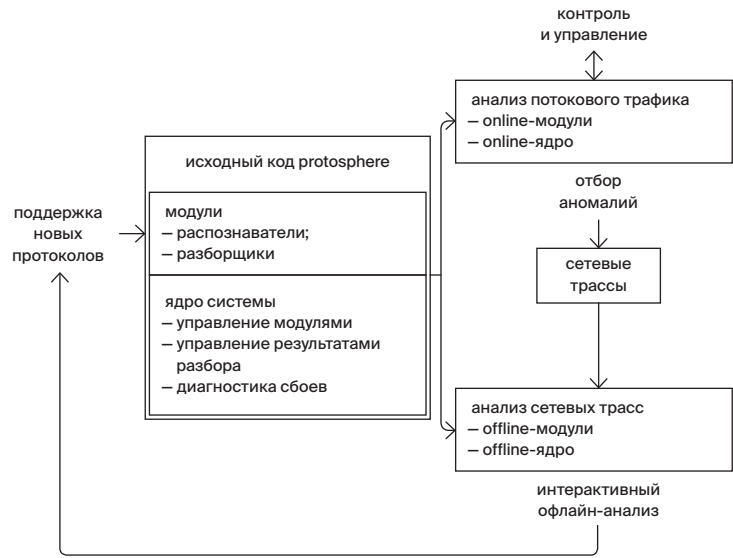
ДЛЯ КОГО ПРЕДНАЗНАЧЕНА СИСТЕМА PROTOSPHERE?

- Компании, занимающиеся тестированием реализаций сетевых протоколов (в том числе, во встраиваемых ОС и сетевой аппаратуре);
- Компании-разработчики средств сетевой безопасности (межсетевых экранов, а также систем обнаружения и предотвращения вторжений);
- Компании по производству техники, нуждающейся в повышенном уровне безопасности из-за обязательной сертификации.

ПОДДЕРЖИВАЕМЫЕ ПЛАТФОРМЫ И АРХИТЕКТУРЫ

Архитектуры: Intel x86-64.
Платформы: ОС Windows, ОС на базе ядра Linux.

СХЕМА РАБОТЫ



ТЕХНОЛОГИЯ СТАТИЧЕСКОЙ ВЕРИФИКАЦИИ ПРОГРАММ НА ЯЗЫКЕ GNU СИ

Klever — система статической верификации с использованием передовых инструментов для тщательной проверки безопасности, надёжности и производительности программных систем, разработанных на языке GNU Си. В частности, применяется для верификации ОС реально-го времени.

ПОЧЕМУ KLEVER?

1 Высокая точность и расширенные возможности анализа

Klever — технология, основанная на научных исследованиях и предназначенная для промышленного использования. Позволяет применять высокоточные формальные методы для верификации больших программных систем. Решает те же задачи, что и мировые аналоги (в частности, Microsoft SDV), отличаясь от них способностью проводить верификацию любого сложного ПО, а не только узкого класса программ (драйверов или встраиваемого ПО). Технология выложена в открытый доступ (<https://forge.ispras.ru/projects/klever>).

Klever — это:

- Высокоточный консервативный анализ любого сложного ПО (выявление всех возможных ошибок искомым видом);
- Проверка различных требований к программе (проверка правил безопасного программирования на языке Си и корректного использования интерфейса, специфичного для проверяемой программы);
- Масштабируемость формальных методов (модульная статическая верификация позволяет масштабировать применение соответствующих инструментов и анализировать проекты, которые содержат тысячи строк кода на языке GNU Си);

- Поддержка верификации разных версий программы (не требуется модифицировать исходный код проверяемой программы, что позволяет легко проводить верификацию различных версий и проверять исправление ошибок сразу после их обнаружения и устранения);
 - Использование инкрементального процесса уточнения результатов верификации (для управления декомпозицией и генерацией моделей окружения фрагментов программ применяется постоянно уточняемый набор спецификаций).
- 2** **Высокий уровень адаптивности и удобства**
- Доработка технологии под нужды заказчика. Оперативное расширение списка обнаруживаемых ошибок. Разработка набора спецификаций для формализации специфичных для программы требований, а также спецификаций для моделирования окружения и в некоторых случаях – плагинов;
 - Удобный многопользовательский веб-интерфейс для выполнения статической верификации, а также хранения, анализа и сравнения результатов;
 - Возможность заказчика получить отчуждаемый продукт (после подготовки, адаптации, разработки спецификаций и поиска ошибок). Обучение разработчиков заказчика.

ОПЫТ ВНЕДРЕНИЯ

Технология Klever разработана в рамках Центра верификации ОС Linux (inuxtesting.org), организованного на базе ИСП РАН при поддержке Linux Foundation. В настоящее время Klever используется для верификации различных операционных систем.

В процессе верификации драйверов устройств и подсистем операционной системы Linux удалось добиться следующих результатов:

- Обнаружены более 300 ошибок, подтвержденных разработчиками: ошибки выхода за границу буфера, разыменование нулевого указателя, использование неинициализированной памяти, повторное или некорректное освобождение памяти, состояния гонки и взаимные блокировки, утечки специфичных ресурсов ядра Linux, некорректные вызовы функций в зависимости от контекста, некорректная инициализация специфичных структур данных ядра Linux;
- Достигнуто 50% покрытие кода драйверов устройств и подсистем ядра. Для поиска указанных ошибок при выполнении различных сценариев взаимодействия драйверов ядра Linux с их окружением Klever построил достаточно точную модель окружения (более 20 наиболее широко используемых интерфейсов драйверов, таких как обработчики прерываний и таймеры, интерфейсы USB и PCI устройств, интерфейсы сетевых и символьных устройств).

ДЛЯ КОГО ПРЕДНАЗНАЧЕН KLEVER?

- Компании, нацеленные на разработку ПО с особым акцентом на высокую надежность и безопасность;
- Компании, которые нуждаются в сертификации разрабатываемого ПО;
- Сертификационные лаборатории.

СХЕМА РАБОТЫ



ГЕНЕРАТОР ТЕСТОВЫХ ПРОГРАММ

8/17

MICROTESK

Верифицирует
микропроцессоры

MicroTESK – реконфигурируемая и расширяемая среда генерации тестовых программ для функциональной верификации микропроцессоров. Позволяет автоматически конструировать генераторы тестовых программ для целевых архитектур микропроцессоров на основе их формальных спецификаций. MicroTESK применим для широкого спектра архитектур (RISC, CISC, VLIW, DSP).

ПОЧЕМУ MICROTESK?

1 Продуманная перспективная концепция

MicroTESK – комплекс технологий для промышленного использования, включающий в себя базовую среду моделирования (строит модели микропроцессоров на основе формальных спецификаций) и среду генерации (строит тестовые программы на основе шаблонов). По решаемым задачам близок к мировым аналогам (Genesys Pro и RAVEN), однако отличается от них повышенной производительностью и удобством использования, а также распространением по лицензии открытого исходного кода. В открытом доступе на сайте ИСП РАН: forge.ispras.ru/projects/microtesk. Кроме того, описание технологии доступно на [сайте microtesk.org](http://www.microtesk.org).

MicroTESK – это:

- Использование формальных спецификаций в качестве источников знаний о конфигурации верифицируемого микропроцессора:
 - спецификации архитектуры на nML (регистры, память и режимы адресации, логика инструкций, текстовый/бинарный формат инструкций);
 - дополнительные спецификации подсистемы памяти на mmuSL (свойства буферов памяти (TLB, L1 и L2), логика трансляции адресов и логика операций чтения и записи);
 - потенциальная возможность перехода к формальной верификации, а также генерации набора инструментов для разрабатываемого микропроцессора (дизассемблер, эмулятор и др.)
- Генерация тестовых программ на основе объектно-ориентированных тестовых шаблонов:

- тестовые шаблоны на языке Ruby (за счёт чего шаблоны наглядны и просты в поддержке);
 - возможность одновременного использования различных техник генерации наборов инструкций и тестовых данных (случайная генерация, комбинаторная генерация, генерация на основе разрешения ограничений и др.);
 - масштабируемость среды генерации (возможность разрабатывать сложные шаблоны при небольших затратах за счет повторного использования).
- Широкий набор поддерживаемых архитектур микропроцессоров:
- поддержка особенностей различных классов архитектур на уровне среды разработки генераторов (RISC, CISC, VLIW, DSP);
 - разработаны генераторы тестовых программ на основе MicroTESK для таких архитектур, как ARM, MIPS, PowerPC, RISC-V;
 - поддерживается многоядерность целевой микропроцессорной архитектуры.
- 2** Максимальное удобство для заказчика
- Оперативная настройка среды под новые архитектуры с минимальными затратами и автоматическое извлечение информации о тестовых ситуациях (благодаря формальным спецификациям);
 - Удобный язык разработки тестовых шаблонов, позволяющий быстро описывать сложные сценарии верификации;
 - Возможность интеграции широкого набора различных методов генерации (случайной, комбинаторной, на основе разрешения ограничений и др.);
 - Оперативная локальная техподдержка, а также обучение разработчиков заказчика;
 - Возможность заказчика получить отчуждаемый продукт.

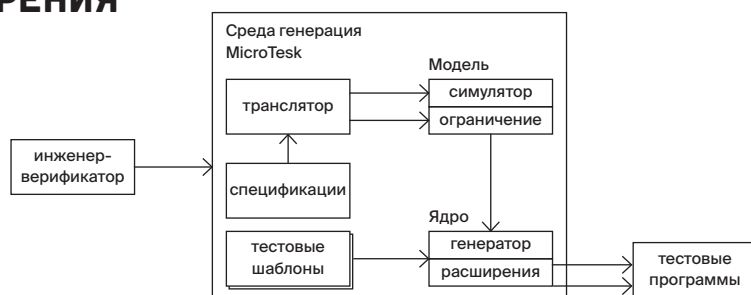
СИСТЕМНЫЕ ТРЕБОВАНИЯ

OS Windows или ОС на базе ядра GNU/Linux, Java Runtime Environment версии 8.

ОПЫТ ВНЕДРЕНИЯ

MicroTESK разрабатывается с 2007 года. Использовался в российских и международных проектах по разработке современных промышленных микропроцессоров (в частности, в промышленных проектах по верификации микропроцессоров ARMv8 и MIPS64).

ОПЫТ ВНЕДРЕНИЯ



ИНСТРУМЕНТ ОБРАТНОЙ ИНЖЕНЕРИИ HDL-ОПИСАНИЙ

Статический
анализ описаний
цифровой аппаратуры

Retrascope – инструмент обратной инженерии и функциональной верификации описаний цифровой аппаратуры. Предоставляет автоматизированные средства извлечения и анализа формальных моделей исходного кода. Поддерживает синтезируемые подмножества языков Verilog, VHDL.

ПОЧЕМУ RETRASCOPE?

- | | | |
|---|-----------------------------|---|
| 1 | Сочетание важнейших качеств | <p>Retrascope – расширяемый инструмент, позволяющий разрабатывать гибридные техники верификации HDL-описаний за счёт комбинирования различных средств анализа формальных моделей.</p> <p>Retrascope – это:</p> <ul style="list-style-type: none"> – Извлечение формальных моделей из исходного кода и их визуализация. Поддерживаются следующие виды моделей: <ul style="list-style-type: none"> – граф потока управления; – решающая диаграмма охраняемых действий; – высокоуровневая решающая диаграмма; – расширенный конечный автомат. – Генерация функциональных тестов для модулей аппаратуры (случайная генерация, обход расширенного конечного автомата, ограниченная проверка моделей); – Проверка формальных моделей (model checking) с помощью внешних средств проверки (NuSMV, nuXmv) на соответствие пользовательским PSL-спецификациям. |
| 2 | Удобство для заказчика | <ul style="list-style-type: none"> – Графический интерфейс на основе Eclipse IDE (также доступен интерфейс командной строки); – Открытый исходный код (лицензия Apache License Version 2.0); – Расширяемость на уровне исходного кода (возможность добавления новых представлений описаний аппаратуры, |

- а также средств анализа);
- Открытые интерфейсы взаимодействия (языки SMT-LIB, SMV) позволяют использовать различные средства проверки моделей и решатели для достижения целей анализа и верификации

ДЛЯ КОГО ПРЕДНАЗНАЧЕН RETRASCOPE?

- Компании, нацеленные на разработку цифровой аппаратуры;
- Коллективы, проводящие исследования в области функциональной верификации цифровой аппаратуры.

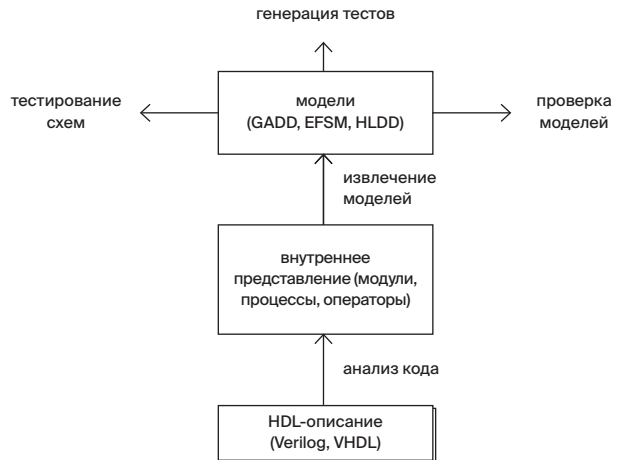
ОПЫТ ВНЕДРЕНИЯ

Инструмент находится на стадии исследовательского прототипа, ведётся разработка.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

Аппаратное обеспечение: IBM-совместимый ПК.
 Программное обеспечение: ОС Windows или ОС на базе ядра GNU\Linux, Java Runtime Environment версии 8.

СХЕМА РАБОТЫ



СИСТЕМА ВЕРИФИКАЦИИ ASTRAVER TOOLSET

10/17

Тщательная
проверка уровня
безопасности

AstraVer Toolset – система дедуктивной верификации ключевых компонентов. Позволяет разрабатывать и верифицировать модели политик безопасности, а также проводить доказательство корректности ключевых компонентов на языке С. Необходимый инструмент достижения целей семейств доверия ADV_SPM и ADV_FSP, определенных в ГОСТ Р ИСО/МЭК 15408-3-2013.

ПОЧЕМУ ASTRAVER TOOLSET?

1 Высокое качество верификации

AstraVer Toolset – комплекс инструментов, предназначенный для промышленного использования и основанный на многолетних научных исследованиях. Объединяет два подхода к верификации: на уровне моделей и на уровне кода. Решает те же задачи, что и аналогичные инструменты (например, Microsoft VCC, Frama-C WP), однако благодаря специфической доработке обладает технологической уникальностью: возможностью верификации ключевых компонентов системы безопасности ядра Linux. Выложен в открытый доступ (linuxtesting.ru/astraver).

AstraVer Toolset – это:

- Комплексный подход к верификации, начиная с формализации требований верхнего уровня и до анализа поведения кода;
- Моделирование функциональных требований (формализация функциональных требований к системе, доказательство внутренней согласованности требований и недостижимости небезопасных состояний);
- Верификация ключевых компонентов на языке С (формализация требований к ключевым компонентам, доказательство корректности работы компонента на всех возможных входных данных);
- Поддержка индустриального кода (нестандартные расширения компилятора GCC, арифметические операции с побитовой точностью, адресная арифметика (включая поддержку конструкции `container_of`), функциональные указатели, переинтерпретация типа указателя между целочисленными типами);
- Решение важнейших задач профилей защиты:

- формальное моделирование политики безопасности;
 - формальное доказательство внутренней непротиворечивости модели политики безопасности и недостижимости небезопасных состояний;
 - разработка полуформальной или формальной функциональной спецификации;
 - формальное или полуформальное доказательство соответствия между моделью политики безопасности и функциональной спецификацией;
 - формальное или полуформальное доказательство соответствия между различными представлениями целевого ПО, такими как функциональная спецификация, проект ПО и его реализация.
- 2 Удобство для российского заказчика
- Инструментарий AstraVer Toolset позволяет решить задачи, сформулированные в требованиях к средствам защиты информации (в семействах доверия ADV_SPM и ADV_FSP, определённых в ГОСТ Р ИСО/МЭК 15408-3-2013 «Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий»);
 - Возможна доработка комплекса под конкретного заказчика (в плане поддержки верификации компонентов на языке C).

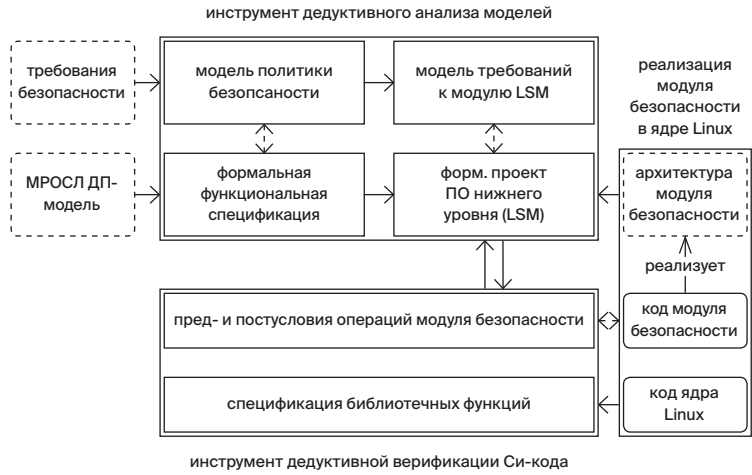
ДЛЯ КОГО ПРЕДНАЗНАЧЕН ASTRAVER TOOLSET?

- Компании, нацеленные на разработку ПО с особым акцентом на высокую надежность и безопасность – как информационную, так и функциональную (ПО для самолетов, АЭС и др.);
- Компании, которые нуждаются в сертификации разрабатываемого ПО в соответствии с ГОСТ Р ИСО/МЭК 15408;
- Испытательные лаборатории средств защиты информации в соответствии с требованиями безопасности.

ОПЫТ ВНЕДРЕНИЯ

Система AstraVer Toolset применялась при разработке средств защиты информации ОС Astra Linux Special Edition (АО «НПО РусБИТех»), которая успешно прошла сертификацию на соответствие требованиям безопасности информации ФСТЭК России к операционным системам по профилю защиты «2А». В основу отечественной разработки была положена МРОСЛ-ДП модель безопасности, а реализация ее новых возможностей в ОС Astra Linux Special Edition продолжает верифицироваться с помощью AstraVer Toolset.

СХЕМА РАБОТЫ



→ ручная разработка

--> автоматизированная верификация

ПОДДЕРЖКА ПРОЕКТИРОВАНИЯ ОТВЕТСТВЕННЫХ СИСТЕМ

Прогнозирует качество, защищает от ошибок

MASIW – набор инструментов для разработки программно-аппаратных комплексов ответственных систем в сфере авиации, медицины и др. Создан для инженеров-конструкторов комплексов бортового оборудования для авиационных судов, разрабатываемого с применением интегрированной модульной авионики (ИМА). Оперативно адаптируется под другие предметные области.

ПОЧЕМУ MASIW?

- 1 Сочетание важнейших качеств

MASIW – технология для оптимизации разработки сложных программно-аппаратных комплексов, а также их верификации. Позволяет провести предварительную оценку качества изделия до появления опытного образца, а также анализ на отказоустойчивость. Экономит время и деньги, снижает риск появления ошибок и дефектов. Разрабатывается совместно с ФГУП «ГосНИИАС».

Несмотря на наличие инструмента OSATE на момент начала разработки, на сегодняшний день MASIW превосходит его по функциональности в плане верификации, а также статического и динамического анализа.

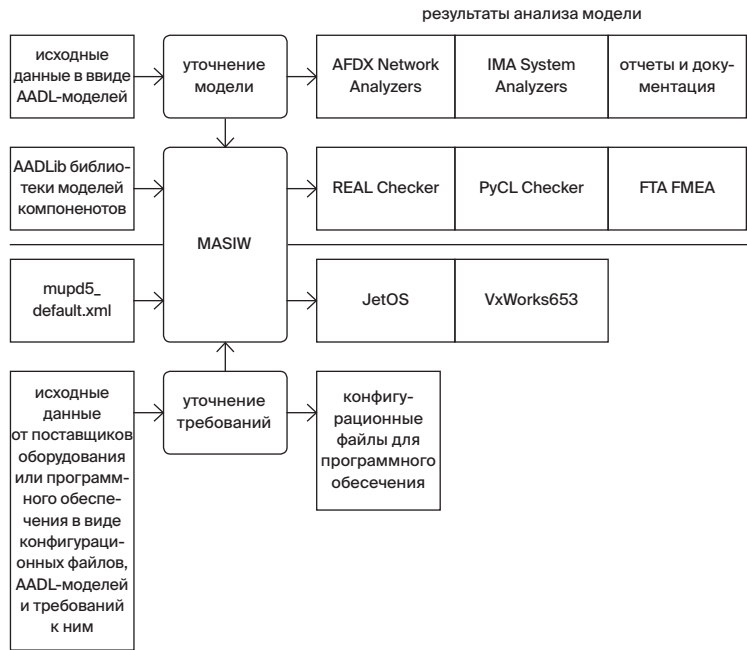
В текущей реализации MASIW – это:

- Создание, редактирование и управление моделями на языке AADL:
 - создание/редактирование моделей посредством текстового или графического редактора;
 - поддержка командной разработки с возможностью отслеживания и внесения изменений для отдельных элементов модели;

- поддержка переиспользования AADL-моделей сторонних разработчиков.
- Анализ моделей:
 - анализ структуры программно-аппаратного комплекса (достаточности аппаратных ресурсов, согласованности интерфейсов и т. п.);
 - проверка разрабатываемого программно-аппаратного комплекса на соответствие требованиям;
 - анализ характеристик передачи данных в сети AFDX (времени доставки сообщений от отправителя к получателю, глубины очередей передающих портов и т. п.);
 - построение дерева неисправностей и его численный анализ для определения вероятности отказного события верхнего уровня;
 - анализ видов и последствий отказов на основе архитектурной модели комплекса бортового оборудования, включая построение таблицы видов и последствий отказов;
 - симуляция модели программно-аппаратного комплекса с генерацией пользовательских отчётов по результатам работы симулятора, в том числе, совместная симуляция работы прикладных разделов под управлением ОС PV в эмуляторе Qemu и универсального симулятора AADL моделей.
- Синтез моделей:
 - распределение функциональных приложений по вычислительным модулям с учётом ограничений ресурсов аппаратной платформы и с учётом дополнительных ограничений, касающихся вопросов надёжности и безопасности программно-аппаратного комплекса;
 - генерация распределения вычислительного времени процессора между функциональными приложениями (циклограмма расписания запуска приложений для ARINC-653 совместимых ОС реального времени).
- Генерация конфигурационных данных:
 - разработка специализированных инструментов конфигурационных данных на основе предоставляемого программного интерфейса (API);
 - генерация конфигурационных файлов для компонентов КБО.

- | | | |
|---|------------------------|--|
| 2 | Удобство для заказчика | <ul style="list-style-type: none"> - Возможность расширения набора инструментов путём создания собственных модулей (благодаря модульной архитектуре в основе технологии); - Оперативная адаптация технологии под другие предметные области; - Возможность получить отчуждаемый продукт; - Непрерывное взаимодействие с локальной техподдержкой и обучение разработчиков заказчика. |
|---|------------------------|--|

СХЕМА РАБОТЫ



CONSTRUCTI- VITY 4D:

12/17

Основа для 4D-моделирования
индустриальных проектов

ТЕХНОЛОГИЯ ИНДЕКСИРОВАНИЯ, ПОИСКА И АНАЛИЗА БОЛЬШИХ ПРОСТРАНСТВЕННО- ВРЕМЕННЫХ ДАННЫХ

Constructivity 4D – технология для создания перспективных программных систем и сервисов, оперирующих большими массивами пространственно-временных данных и динамическими сценами. Способна проводить визуальный анализ миллионов объектов с различным геометрическим представлением и индивидуальным динамическим поведением. Внедрена в систему Synchro, предназначенную для 4D-моделирования крупных строительных объектов.

ПОЧЕМУ CONSTRUCTIVITY 4D?

1 Оптимальное сочетание функций для анализа объектов максимальной сложности

Constructivity 4D – технология для промышленного использования, объединяющая оригинальные методы пространственно-временного индексирования, поиска, а также качественного и количественного анализа данных с учётом особенностей их геометрического представления, сложной организации и предопределённого характера динамики.

Constructivity 4D – это:

- Использование развитых наборов операций для эффективного исполнения запросов:
 - Темпоральные операции (реализуют классическую интервальную алгебру Аллена применительно к временным штампам дискретных событий и их интервалам);
 - Метрические операции (позволяют определять индивидуальные свойства геометрических объектов и характеристики их взаимного расположения: диаметр, площадь, объем, центр масс, планарные проекции и др.);
 - Топологические операции (предназначены для классификации взаимного расположения объектов и установления фактов их совпадения, пересечения, покрытия, касания, перекрытия или коллизии). Допускают конструктивную имплементацию и применимы для анализа сложных объектов (в отличие от известных топологических моделей DE-9IM, RCC-8 и RCC-3D);
 - Ориентационные операции (обобщают известные системы исчисления направлений Франка, Фрексы, CDC, OPRA и применимы для анализа объектов с протяженными границами).

- Объектно-ориентированная реализация на языке C++ (расширяемый набор классов, интерфейсов и связанных с ними методов для задания пространственно-временных данных и исполнения типовых запросов к ним);
 - Эффективное исполнение запросов и решение типовых задач (реконструкция сцены на заданный момент времени, выборка объектов в заданной пространственной области, поиск ближайших соседей, определение статических и динамических столкновений, бесконфликтная маршрутизация в глобальном динамическом окружении);
 - Система пространственно-временного индексирования (бинарные деревья событий, октарные деревья пространственной декомпозиции, деревья ограничивающих объемов, объектных кластеров, занятости пространства);
 - Комбинированная вычислительная стратегия для определения столкновений в сценах. Объединяет методы точного определения столкновений, методы локализации на основе пространственной декомпозиции, иерархии ограничивающих объемов и методы темпоральной когерентности;
 - Оригинальный метод маршрутизации в глобальном динамическом окружении. Основан на извлечении пространственной, метрической и топологической информации, а также на её согласованном использовании при планировании путей.
- 2** Удобство для заказчика
- Различные возможности расширения библиотеки (может использоваться при разработке новых приложений, а также для оптимизации и расширения функций уже существующих);
 - Непрерывное взаимодействие с локальной техподдержкой и обучение разработчиков заказчика;
 - Возможность заказчика получить отчуждаемый продукт.

ДЛЯ КОГО ПРЕДНАЗНАЧЕНА CONSTRUCTIVITY 4D?

Технология предназначена для создания приложений в самых разных предметных областях, в частности: компьютерная графика и анимация, геоинформатика, научная визуализация, автоматизация проектирования и производства, робототехника, логистика, планирование и управление проектами.

ОПЫТ ВНЕДРЕНИЯ

Технология успешно используется в составе программной системы Synchrono (<https://www.synchrono.com/>), предназначенной для визуального 4D-моделирования, планирования и управления масштабными промышленными проектами в сфере строительства зданий, инфраструктурных объектов и др. Применяется более чем 300 компаниями в 36 странах (в том числе, в России).

БАЗОВЫЙ СЕ- МАНТИЧЕСКИЙ АНАЛИЗАТОР TEXTERRA

Умный анализ
текстов

Texterra – масштабируемая платформа для извлечения семантики из текста. Базовый комплекс технологий для создания многофункциональных прикладных приложений. Анализирует тексты с помощью выделения концептов. Входит в Единый реестр российского ПО.

ПОЧЕМУ TEXTERRA?

- 1 Уникальное сочетание функций
 - Texterra осуществляет уникальный анализ русскоязычных текстов на основе выделения концептов, а не просто слов. Отличается от иностранных аналогов преимущественным вниманием к русскому языку. Базирована на результатах фундаментальных исследований и предоставляет возможность интеграции с поисковой системой Elasticsearch, существенно расширяя ее возможности. Удачное сочетание технологий позволяет платформе конкурировать с проектами уровня IBM Watson Natural Language Understanding.
 - Texterra – это:
 - Высокая скорость обработки текста (морфологический анализ – 69 000 слов в секунду, синтаксический – 39 100 слов/сек, разрешение кореферентности – 10 100 слов/сек, полный разбор текста – приблизительно 13 600 слов/сек);
 - Максимальное внимание к русскому языку (в отличие от аналогичных проектов spaCy и UDPipe, а также IBM Watson Natural Language Understanding, который не поддерживает анализ эмоций и концептов в русскоязычных текстах);
 - Большой объем знаний (более 7 миллионов понятий);
 - Построение базы знаний без привлечения экспертов (автоматическое пополнение с помощью Wikipedia, MediaWiki, Linked Open Data и др.);
 - Масштабируемость как по скорости обработки текстов, так и по объему знаний (с помощью Apache Ignite и оригинальной облачной технологии ИСП РАН);
 - Высокая точность анализа текста благодаря ряду ключевых особенностей:
 - Многоуровневый поиск по смежным понятиям;

- Адаптивность к сленгу, хэштегам и ошибкам;
- Анализ эмоциональной окраски (с разделением отношения к объектам и их атрибутам);
- Определение взаимосвязей людей и компаний (на основе информации в тексте);
- Определение неявных упоминаний объектов в дискуссиях.

2 Максимальная адаптивность под российского заказчика

Texterra – высокотехнологичный продукт, сочетающий передовые научные разработки с возможностью их промышленного использования. Наша локальная техподдержка работает с максимальным вниманием к российскому заказчику. В числе главных преимуществ:

- Высокая скорость разработки индивидуального решения;
- Два варианта использования:
 - в качестве отчуждаемого продукта на локальном сервере заказчика с доступом как по протоколу HTTP (REST-архитектура), так и по протоколу RMI;
 - онлайн на сайте texterra.ispras.ru;
- Непрерывное обучение разработчиков заказчика, а также инновационная технологическая доработка продукта в соответствии с новыми задачами и вызовами;
- Простое и быстрое освоение специфичных предметных областей и возможность интеграции новых языков для анализа (благодаря современному подходу к машинному обучению).

ДЛЯ КОГО ПРЕДНАЗНАЧЕНА TEXTERRA?

- Разработчики корпоративного ПО (в частности, чат-ботов);
- Разработчики систем семантического поиска для специфических предметных областей (информационная безопасность, медицина, аудит и т.п.);
- Разработчики произвольных прикладных систем обработки текста.

С КЕМ МЫ СОТРУДНИЧАЕМ?

Texterra доработана до промышленного уровня в рамках сотрудничества с HP и Samsung (цель совместных проектов – получение технологий для анализа корпоративной отчетности и поддержки работы смарт-телевидения). В настоящее время на базе платформы работает ряд оригинальных разработок ИСП РАН (в частности, технология анализа социальных медиа Talisman). Texterra используется также рядом государственных ведомств России.

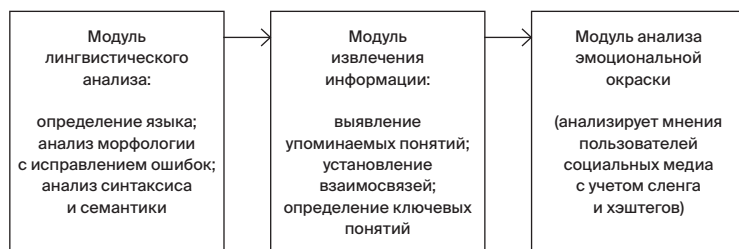
ПОДДЕРЖИВАЕМЫЕ ЯЗЫКИ

Texterra анализирует тексты на русском и английском языках.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

- Любые платформы, поддерживаемые Java 1.8;
- Не менее 16 Гб оперативной памяти для каждого из анализируемых языков;
- Рекомендуется применение 64-битной версии ОС.

СХЕМА РАБОТЫ



ТЕХНОЛОГИЯ АНАЛИЗА СОЦИАЛЬНЫХ МЕДИА

TALISMAN

Анализирует всё,
находит главное

Talisman — технология анализа больших данных о людях, сообществах, продуктах и организациях. Выявляет закономерности во взаимосвязях с помощью анализа больших графов из сотен миллионов узлов.

ПОЧЕМУ TALISMAN?

1 Уникальное сочетание функций

Talisman — промышленная технология, интегрированная с платформой для извлечения семантики из текста (Texterra) и оригинальной технологией сбора данных ИСП РАН. По технологическому уровню Talisman сопоставим с лучшими мировыми аналогами (Palantir Gotham и IBM Watson Content Analytics). Преимущество — автоматизация рутинных процессов с помощью последних научных достижений (сокращает затраты на аналитиков).

Talisman — это:

- Сочетание важнейших функций, в частности:
 - Семантический анализ с использованием возможностей платформы Texterra (определение эмоциональной окраски сообщений, уникальная для русского языка работа с концептами, возможность анализировать комментарии пользователей, выявлять неявные упоминания объектов в дискуссиях и др.);
 - Анализ больших графов из сотен миллионов узлов (в том числе, автоматическое построение графов распространения информации с определением ролей: первоисточник, распространитель, лидер мнения, читатель).
 - Автоматическая группировка сообщений в информационные сюжеты (карта всех обсуждаемых тем в информационном пространстве с учетом перетекания между различными ресурсами);
 - Выявление истинных параметров пользователей соцсетей. Уточнение пола, возраста (с точностью до года), образования, семейного положения, региона

- проживания на основе анализа профилей и активности пользователей (расширяемый список);
- Автоматическое определение параметров целевой аудитории (агрегация по демографическим атрибутам и выявление доминирующих значений);
- Инструменты проверки достоверности информации (выявление ботов, фильтрация спама, обнаружение признаков манипуляции мнением аудитории).
- Получение отчетов по объектам мониторинга в течение нескольких минут после публикации информации благодаря технологиям анализа больших данных стека Apache Hadoop и эластичной масштабируемости системы с использованием оригинальной облачной технологии ИСП РАН;
- Анализ любых больших данных: корпоративных, новостных, информации из социальных сетей (ВКонтакте, Facebook, Twitter, Instagram, Одноклассники, Youtube, LinkedIn и др.), блогов (LiveJournal), открытых каналов мессенджера Telegram и ресурсов Dark web. Для проведения анализа Talisman может интегрироваться как с оригинальной технологией сбора данных ИСП РАН, так и со внешними сборщиками.

2 Максимальное удобство для российского заказчика

- Talisman предусматривает ряд выгодных возможностей для российского заказчика:
- Промышленное развёртывание решений на оборудовании заказчика (Talisman — полностью отчуждаемый продукт);
 - Работа в режиме облачного сервиса;
 - Непрерывное оперативное взаимодействие с локальной российской техподдержкой (включая внедрение с обучением разработчиков заказчика);
 - Оперативная адаптация технологии и расширение функционала для использования в различных предметных областях (информационная безопасность, медицина, аудит и др.).

ОБЛАСТИ ПРИМЕНЕНИЯ

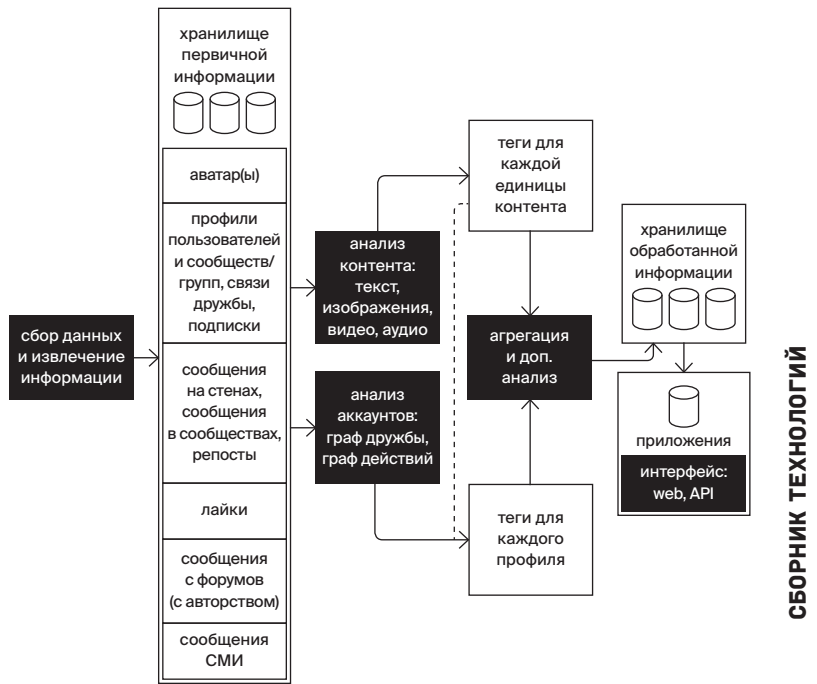
- Выявление групп по интересам на основе анализа текстов социальных медиа, в том числе: определение целевой аудитории (как в маркетинговых целях, так и при формировании политических программ), выявление точек социального напряжения и злободневных проблем с наибольшим числом недовольных;
- Выяснение общественного мнения об организациях, людях и товарах;
- Определение ключевых трендов и прогнозирование эффективности интернет-рекламы;
- Оптимизация управления персоналом (эффективный подбор сотрудников, верификация данных, помощь в разработке систем мотивации на основе текущих и долгосрочных интересов, выявление скрытой деятельности и скрытых связей, а также мониторинг утечек и разглашения внутренней информации);
- Решение задач в области репутационного менеджмента (в частности, выявление причин недовольства сотрудников и клиентов);

- Выявление информационных кампаний, манипулирующих мнением целевой аудитории, а также определение целевой аудитории, на которую направлена кампания.

ИСПОЛЬЗУЕМЫЕ ЯЗЫКИ

В настоящее время Talisman использует языки, распознаваемые анализатором Texterra (русский и английский).

СХЕМА РАБОТЫ



ВИРТУАЛЬНАЯ ЛАБОРАТОРИЯ ДЛЯ ДОКУМЕНТАЦИИ ИСЧЕЗАЮЩИХ ЯЗЫКОВ

Lingvodoc — система для совместной многопользовательской документации исчезающих языков, создания многослойных словарей и научной работы с полученными звуковыми и текстовыми данными. Совместный проект с Институтом языкознания РАН и Томским государственным университетом. Разрабатывается с 2012 года. Сайт проекта — lingvodoc.ispras.ru.

ПОЧЕМУ LINGVODOC?

- 1 Уникальное сочетание важнейших функций

Lingvodoc — кроссплатформенная технология с открытым исходным кодом (<https://github.com/ispras/lingvodoc> и <https://github.com/ispras/lingvodoc-react>), основанная на научных исследованиях и объединяющая ряд необходимых функций. Система постоянно дорабатывается. В настоящее время в рамках совместного проекта с ТГУ завершается разработка уникального решения для построения изоглосс.

Lingvodoc — это:

 - Совместная работа пользователей над пополнением словарных данных (в отличие от аналогичного проекта Starling, где такая работа не предусмотрена);
 - Сохранение полной истории действий пользователей;
 - Одновременная работа с аудиотекстовыми корпусами и словарями на основе интеграции с программой ELAN, разработанной Институтом психолингвистики Макса Планка (Нидерланды);
 - Расставление однонаправленных и двунаправленных связей между лексическими входами внутри словарей, а также между словарями;
 - Запись, проигрывание и хранение звуков с разметкой (в форматах wav, mp3 и flac), а также построение формант гласных с последующей визуализацией;
 - Продвинутый поиск, который позволяет искать данные в словарях по множеству параметров (в отличие от аналогичного проекта TypeCraft);
 - Возможность бесконфликтной двусторонней отложенной синхронизации;
 - Повышенный уровень автоматизации (по сравнению с аналогичным проектом Kielipankki).

- 2 Широкие возможности для пользователей
 - Создание словарей любой структуры, как типичных двуслойных (слой лексических входов и слой парадигм), так и многослойных. Кроме того, существует функция импорта для готовых словарных структур;
 - Работа как с привлечением облачных ресурсов ИСП РАН

(в настоящее время клиент-серверная архитектура оптимизирована под облачную инфраструктуру VMEmpreor), так и с развёртыванием локальной версии с изоляцией собственных данных;

- Наличие программы для веб-просмотра и десктопной версии;
- Открытая регистрация (с подтверждением);
- Оперативная доработка технологии под любого заказчика с расширением набора функций, а также адаптация под другую научную отрасль.

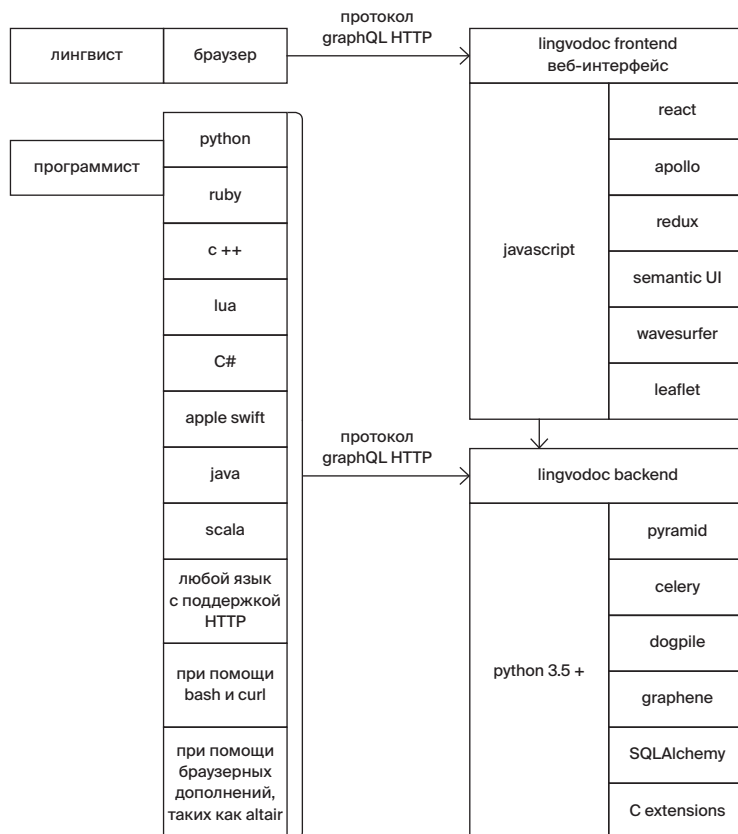
ДЛЯ КОГО ПРЕДНАЗНАЧЕН LINGVODOC?

В первую очередь, Lingvodoc предназначен для лингвистов, ведущих научную работу в сфере документации исчезающих языков. Однако возможна доработка технологии под другие цели.

ГДЕ ИСПОЛЬЗУЕТСЯ LINGVODOC?

В настоящее время Lingvodoc используется в рамках совместных проектов с Институтом языкознания РАН и Томским государственным университетом. Ведутся также переговоры с рядом научно-исследовательских институтов.

СХЕМА РАБОТЫ



СИСТЕМА ИС- СЛЕДОВАТЕЛЬ- СКОГО ПОИСКА

SCINOON

Умный поиск научных публикаций

SciNoon — система для исследовательского поиска научных статей. Объединяет ряд уникальных функций для оптимизации процесса поиска и анализа полученных результатов. В частности, позволяет работать в команде и сохранять историю действий пользователей. Поддерживает работу с большими данными.

ПОЧЕМУ SCINOON?

- 1 Оптимальное сочетание важнейших функций
- SciNoon — инновационная система, созданная с целью оптимизации длительной командной работы с научными публикациями. Помогает исследовать новые предметные области, а также поддерживать информированность в какой-либо одной. Решает те же задачи, что и крупные мировые аналоги (Microsoft Academic Search, Google Scholar, Semantic Scholar), однако вместе с тем обладает уникальными функциями.

SciNoon — это:

- Уникальная функция поддержки совместного поиска, необходимого для успешной командной работы:
 - общее рабочее место, доступное всем участникам исследовательской группы;
 - осведомлённость всех членов группы о работе каждого из них с помощью интегрированного чат-бота.
- Оптимизированная система сбора релевантных научных статей:
 - агрегация данных из разных источников (загруженных пользователем PDF-файлов с полными текстами статей, а также метаданных, собранных с помощью браузерного плагина, интегрированного с Google Scholar);
 - сервис рекомендаций на основе ранее отобранных статей;
 - возможность навигации по графу цитирования.
- Обзор накопленных результатов с помощью удобного пользовательского интерфейса:
 - визуализация всех отобранных статей в виде графа цитирования с возможностью его масштабирования;
 - возможность задания специфичных для исследо-

- вательской задачи аспектов и соответствующая разметка статей;
- визуализация отдельных статей с учетом их важности и введенных значений аспектов;
- табличное представление отобранных статей с учетом как метаданных, так и значений аспектов;
- возможность полуавтоматической кластеризации.
- Поддержка работы с большими данными:
 - собственная графовая модель для представления знаний обо всех статьях, авторах и проводимых исследованиях. Масштабирование на графы размером в десятки миллионов узлов за счет использования графовой базы данных, развёрнутой поверх Apache Cassandra;
 - масштабирование бизнес-логики за счет использования Akka;
 - интеграция с Apache Spark.

2 Максимальное удобство для пользователей

- Два варианта использования:
 - в качестве отчуждаемого продукта на локальном сервере заказчика;
 - онлайн на сайте scinoon.at.ispras.ru.
- Оперативная адаптация технологии и расширение функционала для использования в различных предметных областях.

ДЛЯ КОГО ПРЕДНАЗНАЧЕН SCINOON?

- Сотрудники R&D отделов корпораций;
- Сотрудники научно-исследовательских институтов, нуждающиеся в инструменте для командной работы;
- Преподаватели и студенты ВУЗов, занимающиеся исследовательским поиском для подготовки научных работ.

СХЕМА РАБОТЫ



КОМПЛЕКС РЕШЕНИЙ ДЛЯ СОЗДАНИЯ СЕР- ВИС-ОРИЕНТИ- РОВАННЫХ ЦОД

Максимальное упрощение
сложных задач

Предоставляет возможность хранения данных и совершенствования сложных ресурсоёмких вычислений с использованием как контейнеров, так и виртуальных машин. В частности, предназначен для развёртывания облачных сред.

ГЛАВНЫЕ ПЛЮСЫ

- Возможность адаптации под решение конкретных классов задач (CFD, аналитика больших данных, анализ программ на уязвимости и дефекты и др.);
- Технологическая безопасность и отчуждаемость решений (возможность воссоздания инфраструктуры в изолированной среде с полным контролем над ней за счет использования открытых стандартов, свободного ПО и научных разработок ИСП РАН);
- Оперативная работа локальной техподдержки, обладающей всеми необходимыми компетенциями.

В НАСТОЯЩЕЕ ВРЕМЯ КОМПЛЕКС ПРЕДСТАВЛЕН ТРЕМЯ РЕШЕНИЯМИ:

- 1 Облачная среда на базе кастомизированной технологии Openstack
 - Создана в рамках совместного проекта с компанией Dell. Предназначена для кратковременных вычислений с большими доступными ресурсами. Надёжно функционирует с 2014 года.
 - Развёрнута на базе открытой технологии Openstack, являющейся стандартом построения больших облачных систем;

- Предоставляет пользователям весь необходимый функционал для кратковременных вычислений с большими доступными ресурсами:
 - управление виртуальными сетями и вычислительными кластерами с использованием систем Keystone, Neutron, Nova (аналог Amazon EC2);
 - блочное хранение данных на основе системы Cinder (аналог Amazon Elastic Block Storage);
 - легко расширяемое объектное хранилище на основе Openstack Swift (аналог Amazon S3).
 - Предоставляет возможность разработки и внедрения различных сервисов уровня PaaS:
 - вычислительный кластер Big Data Open Lab для анализа больших данных с полностью настроенными системами Apache Spark, Apache Hadoop и Apache Ignite и произвольным количеством вычислительных узлов (запуск одного кластера занимает около 5 минут). Находится в открытом доступе (<https://github.com/ispras/spark-openstack>);
 - для исследований в области искусственного интеллекта с использованием Tensorflow, Caffe и др., а также современного аппаратного обеспечения (серверов с NVIDIA Tesla V100 на шине SXM2);
 - для работы с HPC.
- 2** Решение для управления виртуальными машинами VMEmpereor
- Разработан в ИСП РАН для решения внутренних задач, находится в открытом доступе (github.com/ispras/vtempereor). Предназначен для управления виртуальными ресурсами на уровне IaaS. С 2012 года бесперебойно работает на базе платформы XCP-ng/Citrix XenServer, предоставляя пользователям простой доступ к получению виртуальных ресурсов по запросу и их оркестрации.
- 3** Платформа для организации web-лабораторий Fanlight
- Создана в результате участия ИСП РАН в программе «Университетский кластер» и в международном проекте Open Cirrus (учреждён Hewlett-Packard, Intel и Yahoo!). Предназначена для развёртывания SaaS-инфраструктур для вычислительных web-лабораторий средствами Docker Compose. Построена на виртуальных контейнерах и работает на базе виртуальных рабочих мест в модели DaaS (Desktop as a Service). Доступна для пользователей на сайте fanlight.ispras.ru. Поддерживает только приложения, разработанные для ОС на базе ядра Linux.
- Демонстрирует высокую эффективность работы с облачными вычислениями благодаря использованию контейнеров:
 - комфортная работа с тяжелыми инженерными CAD-CAE приложениями, требующими поддержки аппаратного ускорения 3D-графики для сложной визуализации;
 - поддержка выполнения MPI, OpenMP, CUDA приложений за счет доступа к HPC-кластерам, многоядерным процессорам и графическим ускорителям NVIDIA.
 - Расширяет вычислительные возможности на уровне PaaS за счет подключения аппаратных ресурсов (HPC/BigData кластеры, системы хранения, сервера с графическими ускорителями);

- Позволяет провести кастомизацию под заданную прикладную область за счет интеграции специализированных расчётных прикладных пакетов. В частности, есть опыт внедрения:
 - в области МСС: OpenFOAM, SALOME, Paraview и др.;
 - в области Gas&Oil: tNavigator, Eclipse, Roxar, Tempest и др.
- Позволяет пользователю работать через любой тонкий клиент (включая мобильные устройства) без вспомогательного ПО;
- Может быть развёрнута на сервере, вычислительной ферме, в облаке (с уровня IaaS) или в собственном облачном ЦОД.