СБОРНИК ТЕХНОЛОГИЙ

2019

ОГЛАВЛЕНИЕ

ИСП РАН: 25 ЛЕТ РАЗВИТИЯ И РОСТА

АРУТЮН АВЕТИСЯН

доктор физико-математических наук, академик РАН, директор ИСП РАН

В 2019 г. ИСП РАН отмечает юбилей: 25 января исполнилось 25 лет с того дня, как академик В.П. Иванников основал Институт системного программирования, который стал ведущим центром компетенций в этой области в России. В основе экосистемы ИСП РАН — научная школа, которая создавалась ещё в 1960-1970 гг. в ИТМиВТ под руководством академика С.А. Лебедева. Бизнес-моделью стал «треугольник знаний», объединяющий образование, исследования и инновации (известный также как «модель Физтеха»).

Несмотря на универсальность этой модели, её успешная реализация существенно зависит от ряда факторов – например, экономических. Первое пятилетие после создания Института все усилия были направлены на сохранение научной школы и способности к воспроизведению кадров высшей квалификации. Несмотря на колоссальные проблемы 1990-х и смену 80% сотрудников ИСП РАН в связи с массовой «утечкой мозгов» из страны, наша бизнес-модель показала свою жизнеспособность. Мы продолжали работать в нескольких направлениях, разрабатывая компиляторные технологии, операционные системы и базы данных; старались сохранить и увеличить поток студентов из МГУ и МФТИ. С самого начала мы широко использовали свободное программное обеспечение как базу для долгосрочного конкурентоспособного развития.

Всё это помогло нам найти зарубежных заказчиков и заключить первые контракты с крупными индустриальными партнёрами — например, с канадской компанией Nortel Networks Corporation. Вместе мы реализовали ряд исследовательских проектов — в частности, в области формальной верификации программ. Постепенно мы выработали правильную модель международного сотрудничества, позволяющую нам получать финансовую поддержку фундаментальных исследований и обратную связь от индустрии.

В начале 2000-х гг. самое сложное время закончилось, и у Института начался период стабилизации. Здесь

бизнес-модель ИСП РАН продемонстрировала не только устойчивость, но и способность к быстрому развитию. Заработанная репутация позволила нам привлечь новых партнёров: Intel, HP, Dell. Мы стали активно развивать новые научные направления: анализ программ на уязвимости, анализ бинарного кода, естественных языков, позже — анализ социальных сетей. Сформировался устойчивый коллектив. К 2009 г. число сотрудников ИСП РАН превысило 150 человек, средняя зарплата — 47 тысяч рублей (против 22 тысяч в 2003 г.), а доля договорных работ — 73%. Экосистема ИСП РАН перешла на новый уровень: в 2008-2009 гг. у нас начался период роста.

За прошедшие десять лет ИСП РАН продемонстрировал успешный трансфер знаний и технологий за рамки Института. Постепенно мы перешли от фазы совместных исследований с крупными компаниями непосредственно к внедрению наших технологий, права на которые остаются у ИСП РАН. В 2009 г. у нас появился долгосрочный партнёр — Samsung, с которым мы организовали совместную лабораторию. Примером сотрудничества стал анализатор Svace, разработанный в ИСП РАН. Сейчас это основной инструмент статического анализа для поиска ошибок в исходном коде в Samsung. В то же время стало расти и число наших российских партнёров, началось многолетнее сотрудничество с такими компаниями, как «Вымпелком», «РусБИТех» и др.

В рамках трансфера знаний мы начали создавать сети региональных лабораторий системного программирования: в 2008 г. — в Ереване, в 2009 г. — в Великом Новгороде, в 2019 г. — в Орле. Такой распределённый центр компетенций позволяет разделять ресурсы и знания, а также решать масштабные задачи в области разработки и внедрения ПО.

В 2015 г. к двум кафедрам системного программирования в МФТИ и МГУ присоединилась третья: базовая кафедра ИСП РАН в ВШЭ. Увеличился поток студентов. С 2016 г. Институт начал проводить ежегодные Открытые конференции, собирающие сотни участников. Мы значительно расширили список наших международных партнёров. В частности, с 2019 г. мы сотрудничаем с компанией Huawei, с которой созданы две совместные лаборатории.

Для иллюстрации нашего прогресса достаточно привести лишь некоторые статистические данные. По итогам 2019 г. объем привлечённого финансирования ИСП РАН составит более 800 миллионов рублей — это в 3 раза больше, чем 10 лет назад. В настоящее время доля договорных работ превышает 88% (причём половина — с российскими, половина — с зарубежными компаниями). Число сотрудников Института постоянно растёт и сейчас составляет около 300 человек.

В настоящее время в Институте активизированы исследования в рамках таких перспективных направлений, как искусственный интеллект и анализ больших данных. Одним из главных направлений работы Института остаётся кибербезопасность. В настоящее время осуществляется

переход от отдельных технологий к комплексным предметно-ориентированным платформам, которые нацелены на обеспечение технологической независимости страны. Это платформа жизненного цикла безопасного ПО (в частности, включает в себя статические анализаторы Svace и BinSide, инструменты динамического анализа ИСП Фаззер и Anxiety), а также платформа анализа текстов и социальных медиа (включает в себя фреймворк Talisman, платформу для извлечения семантики из текста Texterra, систему поиска SciNoon).

Наши технологии входят в Единый реестр российского ПО и внедряются в крупных мировых компаниях (Samsung, Huawei), а также на российском рынке («РусБИТех», ГосНИИАС и др.). При этом мы стараемся диверсифицировать риски, и никогда не опираемся только на бизнес и только на одного заказчика.

В области кибербезопасности ИСП РАН сотрудничает со ФСТЭК России в целях создания специализированных стандартов и методик. Кроме того, мы реализуем совместные проекты с крупными образовательными и научно-исследовательскими центрами (израильский Технион, тайваньский ITRI, Белградский университет и др.). В наших дальнейших планах — развитие междисциплинарных исследований (в этом году мы начали работать в сфере цифровой медицины), увеличение потока студентов, развитие предметно-ориентированных платформ.

За 25 лет существования Институт смог создать экосистему, обеспечивающую постоянную генерацию кадров и инноваций в области системного программирования. Наша бизнес-модель продемонстрировала успешную работу в самых разных условиях. Имеющийся технологический задел, кадровый потенциал, репутация и связи с индустрией позволяют нам с оптимизмом смотреть в будущее.

В этом сборнике приводится информация о главных достижениях ИСП РАН в 2019 г., подробнее рассказывается о разрабатываемых платформах, а также о бизнес-модели и научно-образовательной деятельности Института. Основная часть сборника посвящена детальному описанию инновационных технологий ИСП РАН.

ИСП РАН: ГЛАВНЫЕ СОБЫТИЯ 2019 ГОДА

В 2019 г. Институт начал переход от разработки отдельных технологий к созданию комплексных предметно-ориентированных платформ — с возможностью адаптации под требования заказчиков. Такое решение обусловлено необходимостью объединения взаимодействующих технологий в профильные стеки, способные обеспечить эффективную, продуктивную и безопасную работу ПО.

Кроме того, в текущем году был заключён ряд соглашений о сотрудничестве — как с организационными, так и с технологическими партнёрами. В частности, подписаны договоры с Научно-исследовательским институтом промышленных технологий Тайваня (ITRI) и с Белградским университетом, а также соглашение с Национальным исследовательским ядерным университетом «МИФИ», чьи студенты смогут проходить преддипломную практику и писать выпускные работы в ИСП РАН.

В рамках стратегического партнёрства с компанией Huawei созданы две совместные лаборатории: для проведения исследований и разработок в области компиляторных технологий и компонентов операционных систем, а также для проведения работ в области статического и динамического анализа программ.

Одной из важнейших активностей ИСП РАН в 2019 г. стала разработка и реализация совместно со ФСТЭК России программы дополнительного профессионального образования. Двухнедельные учебные курсы, проведённые экспертами Института для трёх учебных групп с мая по октябрь, охватили все технологии, упомянутые в новой «Методике выявления уязвимостей и недекларированных возможностей в программном обеспечении». В течение года были обучены 56 специалистов — сотрудников органов по сертификации и испытательных лабораторий, аккредитованных ФСТЭК России.

В этом году Институт был выбран оператором двух взаимосвязанных цифровых платформ, которые создаются по инициативе Минобрнауки РФ. Во-первых, это Единая цифровая платформа научного и научно-технического

взаимодействия, организации и проведения совместных исследований в удалённом доступе, в том числе с участием зарубежных учёных (ЦПСИ). Она будет обеспечивать эффективное взаимодействие исследователей и работу виртуальных лабораторий с доступом к разнообразным сервисам. Во-вторых, это Цифровая система управления сервисами научной инфраструктуры коллективного пользования (АС УСНИКП), в том числе уникальными научными установками (УНУ) и центрами коллективного пользования (ЦКП). В числе её главных функций — предоставление вычислительных мощностей по запросу, облачного хранилища и пакетов прикладных программ.

Кроме того, ИСП РАН выиграл конкурс, проведённый РФФИ совместно с Министерством науки и технологии Израиля, и теперь займётся реализацией проекта в области медицины вместе с представителями Израильского технологического института (Технион). Проект посвящен разработке новых методов автоматического распознавания электрокардиограмм, полученных с 12-канального кардиографа. Специалисты ИСП РАН будут разрабатывать и поддерживать мобильное приложение и облачный сервис.

В 2019 г. Институт организовал две конференции при поддержке IEEE (Открытая конференция ИСП РАН и «Иванниковские чтения»), провёл круглый стол «Системное программирование как ключевое направление противодействия киберугрозам» на форуме «Армия-2019», принял участие в организации Научно-практической конференции разработчиков OS DAY, а также в ряде других мероприятий. Сотрудники ИСП РАН провели конкурс «Лучший свободный диплом» в финале XII Международной олимпиады в сфере информационных технологий «IT-планета». Кроме того, ИСП РАН совместно с МИФИ и рядом зарубежных научно-исследовательских центров стал организатором международной конференции «Интеллектуальные технологии в робототехнике» (ITR-2019).

Важным этапом развития Института в этом году стало расширение ереванской и новгородской лабораторий системного программирования, открытых ИСП РАН в 2008 и в 2009 гг. соответственно, а также создание аналогичной лаборатории в Орле. Помимо работ в области анализа программ, активизировано направление анализа больших данных. Достигнута также договорённость о расширении научно-образовательной деятельности в НовГУ в связи с запуском магистерской программы «Информационные технологии больших данных». При участии сотрудников ереванской лаборатории с этого учебного года начата трансформация высшего образования в крупнейших вузах Армении. В частности, проводится синхронизация курсов Российско-Армянского университета и филиала МГУ в Ереване с курсами, которые читаются на кафедре системного программирования ВМК МГУ.

В 2019 г. получены пять свидетельств о регистрации программы для ЭВМ и один патент на изобретение (способ

верификации формальной автоматной модели поведения программной системы). Шесть программ включены в Единый реестр российского ПО:

- DigiTEF (№5377),
- ТРАЛ (№5323),
- Talisman.Биография (№5547),
- Talisman.Поток (№6045),
- Asperitas (№5921),
- Fanlight (№6066).

15 ноября 2019 г. директор ИСП РАН А.И. Аветисян стал первым академиком РАН по специальности «прикладная математика и информатика, кибербезопасность» (Отделение математических наук РАН). Двое сотрудников Института победили в конкурсе 2019-2021 гг. на получение стипендии Президента РФ молодым учёным и аспирантам по направлению «Стратегические информационные технологии, включая вопросы создания суперкомпьютеров и разработки программного обеспечения». Один из сотрудников получил Премию правительства Москвы молодым учёным.

ИСП РАН: ЭКОСИСТЕМА ИННОВАЦИЙ

Деятельность ИСП РАН нацелена на внедрение результатов фундаментальных исследований в индустрию. Бизнес-модель Института состоит из трёх тесно связанных активностей, 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 на создание новых технологий;
- внедрение новых технологий в компаниях-партнёрах и формирование инновационных продуктов после получения обратной связи от индустрии;
- обучение студентов и аспирантов на основе разработанных технологий (с обязательным участием в исследовательских и промышленных проектах Института).

Такая модель хорошо известна и применяется в исследовательских лабораториях ведущих университетов (Stanford, MIT, Berkeley, Carnegie Mellon) и индустриальных гигантов (IBM, Intel), а также в государственных исследовательских центрах (INRIA, Fraunhofer). При условии эффективной реализации данная модель позволяет решить проблему разрыва между наукой и промышленностью, а также организовать подготовку кадров высшей квалификации.

ФУНДАМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Фундаментальные исследования и проведение экспериментальных работ — необходимые элементы деятельности Института, позволяющие двигаться в русле самых новых тенденций в мире ИТ, а также генерировать собственные идеи для проектов с бизнес-партнёрами. ИСП РАН ведёт большое число научных и образовательных программ и сотрудничает с ведущими российскими и зарубежными научными и университетскими центрами (ITRI (Тайвань), Университет Пассау (Германия), Израильский технологический институт Технион, Белградский университет и др.). Это позволяет обеспечивать высокий уровень результатов исследований, а репутация в акаде-

мических и университетских кругах открывает перспективу внедрения отечественных технологий на международных рынках.

В рамках научной деятельности ИСП РАН осуществляет выпуск собственного издания «Труды Института системного программирования РАН» (индексируется в РИНЦ и Scopus). Институт отвечает также за выпуск и редактуру журнала РАН «Программирование» (индексируется в Web of Science и Scopus). Оба издания входят в перечень ВАК.

Кроме того, при Институте функционирует российский Центр верификации ОС Linux, созданный для развития и продвижения открытых стандартов Linux.

ВНЕДРЕНИЕ

ИСП РАН внедряет результаты своих исследований через крупные промышленные и исследовательские организации, которые одновременно используют новые технологии Института и продвигают их в широкую практику. Большая часть работ по контрактам ведётся с долговременными партнёрами, которые сотрудничают с ИСП РАН более пяти лет. В числе главных зарубежных партнёров – Samsung, Huawei, HP, Intel, Nvidia, Rogue Wave, Bentley Systems (ранее Synchro Software), Linux Foundation; в числе отечественных — «РусБИТех», ГосНИИАС, «Вымпелком», «Базальт СПО», «МВП Свемел».

НАУЧНОЕ СОТРУДНИЧЕСТВО

Одна из форм организации долгосрочного сотрудничества в ИСП РАН — это совместные лаборатории. При наличии постоянного финансирования они позволяют гибко планировать имеющиеся ресурсы, а также наращивать компетенции во вновь образующихся направлениях системного программирования и организовывать подготовку молодых специалистов с компетенциями в интересующих партнёров областях.

В настоящее время в Институте функционируют совместные лаборатории с такими компаниями, как Samsung (нацелена на компиляторные технологии, в том числе, на обеспечение безопасности в контексте ОС Android и Tizen) и Huawei (первая лаборатория нацелена на проведение исследований и разработок в области компиляторных технологий и компонентов операционных систем, вторая — в области статического и динамического анализа). Кроме того, на базе технологической сервисно-ориентированной облачной платформы Fanlight создана и успешно функционирует лаборатория для решения задач механики сплошных сред, реализующая исследовательские проекты в интересах промышленных предприятий.

ИНТЕЛЛЕКТУАЛЬНАЯ СОБСТВЕННОСТЬ

В бизнес-модели ИСП РАН права на интеллектуальную собственность остаются за Институтом, либо же они могут передаваться сообществу разработчиков свободного программного обеспечения (СПО) в рамках специальных соглашений (например, с Free Software Foundation). С учётом специфики данной модели была разработана оригинальная лицензия, базирующаяся не на получении роялти, а на прямом финансировании со стороны заказчика дальнейших исследований и разработок, направ-

ленных на развитие технологии. Заказчику передаются неисключительные права по использованию, при этом исключительные остаются за Институтом. В отдельных ситуациях решение по управлению правами принимается индивидуально с учётом перспектив долгосрочного развития. Пример такого исключения – контракт с Фондом перспективных исследований (ФПИ), по которому все права передаются заказчику.

СВОБОДНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ (СПО)

Один из важнейших компонентов созданной экосистемы — широкое использование СПО, без которого невозможно представить себе современное системное программирование. СПО рассматривается как:

- инструмент, предоставляющий легитимный свободный доступ ко всем современным технологиям, включая готовые к использованию программные продукты и открытые стандарты;
- возможность вести инновационное развитие без аутсорсинга благодаря взаимодействию с глобальным рынком продуктов и услуг;
- мощный образовательный ресурс — среда и инфраструктура международных СПО-проектов могут использоваться для подготовки специалистов.

Научная деятельность подразумевает открытость результата и «видимость» его автора, что часто приходит в противоречие с корпоративной политикой ИТ-компаний. Для ИСП РАН открытость результатов исследований — это одновременно и стимул к работе, и инструмент продвижения технологий Института. Открытость приводит к тому, что каждый молодой исследователь «виден» в международном сообществе ИТ-специалистов. Его вклад и репутация — это его капитал, и Институт делает все возможное, чтобы этот капитал рос максимально быстро.

ОБРАЗОВАНИЕ

Краеугольный камень экосистемы инноваций ИСП РАН — образовательная деятельность, которая осуществляется в нескольких направлениях:

- Интеграция ИСП РАН с ведущими вузами. Кафедры системного программирования, на которых работают сотрудники Института, открыты в МГУ им. М.В. Ломоносова, МФТИ и ВШЭ. В первый год обучения в ИСП РАН студенты-третьекурсники слушают лекции специалистов, посещают спецсеминары, знакомятся с исследовательской тематикой по научным направлениям Института и получают специальную стипендию. Во второй год обучения студенты участвуют в исследовательских проектах и получают зарплату, отвечающую требованиям рынка. К моменту выпуска многие учащиеся имеют научные публикации и уже являются реальными специалистами по системному программированию.
- Собственная аспирантура ИСП РАН, предусматривающая одновременно накопление практического опыта и изучение новых технологий. Кроме того, аспиранты активно вовлекаются в процессы обучения: ведут семинарские и практические занятия со студентами, руководят подготовкой курсовых и дипломных работ. Накопив такой опыт, выпускник аспирантуры, как правило, ста-

новится руководителем небольшой исследовательской группы.

- Стипендиальная программа. В рамках поддержки образовательных процессов ИСП РАН запустил стипендиальную программу, которая охватывает студентов ряда образовательных организаций, в числе которых Новгородский государственный университет им. Ярослава Мудрого, Российско-Армянский университет, филиал МГУ в Ереване и Ереванский государственный университет.
- Развитие сети лабораторий системного программирования. В настоящее время функционируют и развиваются три лаборатории, открытые ИСП РАН в Ереване, в Великом Новгороде и в Орле. Лаборатории привлекают к работе успешных студентов и аспирантов, которые занимаются разработкой перспективных технологий в тесном сотрудничестве с индустрией.

В 2017 г. ИСП РАН совместно с компанией Samsung открыл на базе МФТИ «IoT Академию Samsung», в рамках которой студенты проходят спецкурс, направленный на изучение реальных случаев использования технологий Интернета вещей в различных отраслях, а также создают собственные прототипы IoT-устройств. В 2018 г. была запущена вторая часть этого проекта на Факультете аэромеханики и летательной техники (ФАЛТ) МФТИ в Жуковском.

ИСП РАН
В ЦИФРАХ

ОБЩИЙ ОБЪЁМ ПРИВЛЕЧЁННЫХ СРЕДСТВ ЗА ПЕРИОД 2015–2019 ГГ.:	БОЛЕЕ 2,5 МЛРД. РУБ.
ДОЛЯ ДОГОВОРНЫХ РАБОТ ЗА ПЕРИОД 2015–2019 ГГ.:	88,23%
ДОЛЯ НАУЧНЫХ СОТРУДНИКОВ МОЛОЖЕ 39 ЛЕТ:	74,5%
ЧИСЛЕННОСТЬ СТУДЕНТОВ БАЗОВЫХ КАФЕДР ИСП РАН В 2019–2020 УЧЕБНОМ ГОДУ:	МФТИ — 42 ЧЕЛ.; ВМК МГУ — 60 ЧЕЛ.; ВШЭ — 38 ЧЕЛ.
ЧИСЛЕННОСТЬ АСПИРАНТОВ В 2019–2020 УЧЕБНОМ ГОДУ:	21 АСПИРАНТ В БЮДЖЕТНОЙ АСПИРАНТУРЕ 11 В ОЧНОЙ ДОГОВОРНОЙ
ОБЩИЙ ОБЪЁМ СТИПЕНДИЙ АСПИРАНТОВ ИСП РАН ЗА 10 МЕСЯЦЕВ 2019 Г.:	2,095 МЛН. РУБ.
ОБЩИЙ ОБЪЁМ СРЕДСТВ, НАПРАВЛЕННЫХ НА СТИПЕНДИАЛЬНУЮ ПОДДЕРЖКУ СТУДЕНТОВ (2018–2019 УЧ. Г.):	2,940 МЛН. РУБ.
ЧИСЛО ПУБЛИКАЦИЙ, НАПИСАННЫХ СОТРУДНИКАМИ ИСП РАН И ИНДЕКСИРОВАННЫХ В МЕЖДУНАРОДНЫХ БАЗАХ ДАННЫХ В 2018 ГОДУ:	WEB OF SCIENCE: 65. SCOPUS: 91. РИНЦ: 92. GOOGLE SCHOLAR: 77.

ГРАНТЫ

В НАСТОЯЩЕЕ ВРЕМЯ В ИСП РАН ВЕДЁТСЯ НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА ПО 24 ГРАНТАМ РФФИ И ОДНОМУ ГРАНТУ РНФ, А ТАКЖЕ ПО РЯДУ ДРУГИХ ГРАНТОВ И СУБСИДИЙ. КРОМЕ ТОГО, ПРИ ФИНАНСОВОЙ ПОДДЕРЖКЕ РАН В ИНСТИТУТЕ ДЕЙСТВУЕТ ЛАБОРАТОРИЯ ПО СИСТЕМНОМУ ПРОГРАММИРОВАНИЮ И ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ.

ЧТО ДАЛЬШЕ? ПЕРЕХОД ОТ ТЕХНОЛОГИЙ К ПЛАТФОРМАМ

В настоящее время Институт осуществляет переход от отдельных технологий к комплексным предметно-ориентированным платформам. В дальнейшем их развитие должно стать надёжной основой технологической независимости страны, а также успешного развития цифровой экономики. Создание таких платформ приобретает особую актуальность в свете реализации в России программы импортозамещения электронной компонентной базы (ЭКБ). Успешное осуществление такого масштабного проекта возможно только при создании и постоянном развитии отечественных инструментальных средств разработки и отладки ПО.

Специалисты ИСП РАН имеют все необходимые компетенции для создания платформ, которые аккумулируют в себе уже существующие инновационные технологии Института, внедренные в крупных индустриальных проектах. Сборка каждой платформы осуществляется под конкретные нужды и охватывает весь спектр перспективных технологий системного программирования. В частности, специалисты ИСП РАН разрабатывают следующие типы платформ, объединяющих профильные разработки:

- Платформа жизненного цикла безопасного ПО. Включает в себя статические анализаторы Svace и BinSide (применяются в процессе разработки ПО), ИСП Фаззер и Anxiety (используют фаззинг и динамическое символьное исполнение, применяются в процессе тестирования ПО), AstraVer Toolset (система дедуктивной верификации ключевых компонентов), ИСП Обфускатор (обеспечивает защиту от эксплуатации уязвимостей) и др.;
- Платформа создания распределённых систем. Включает в себя комплекс решений для создания сервис-ориентированных ЦОД, предоставляющих возможность хранения данных и совершения сложных ресурсоёмких вычислений, а также программный комплекс DigiTEF для создания сложных цифровых моделей промышленных устройств;
- Платформа анализа текстов и социальных медиа. Включает в себя платформу анализа социальных медиа Talisman, систему поиска SciNoon, платформу для извлечения семантики из текста Texterra и др.

Наряду с разработкой платформ в Институте продолжается и дальнейшая оптимизация существующих технологий. Все они являются инновационными, активно внедряются в индустрии и предусматривают ряд выгодных возможностей для заказчиков:

- Быстрая кастомизация базовой технологии в соответствии с конкретными требованиями;
- Внедрение с одновременным обучением разработчиков;
- Оперативная настройка преобразований и инновационная технологическая доработка продукта в соответствии с новыми задачами и вызовами. В частности, возможно расширение набора инструментов и функций, а также адаптация для использования в различных предметных областях;
- Возможность получить полностью отчуждаемый продукт (для промышленного развёртывания на собственном оборудовании).

ТЕХНОЛОГИИ

ДИНАМИЧЕСКИЙ АНАЛИЗАТОР ANXIETY

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

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

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:

- Создание инструментов анализа (чекеров), которые основаны на методе динамического символьного выполнения и предназначены для конкретных типов ошибок;
- 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;
- Интеграция со статическими анализаторами исходного или машинного кода для реализации направленного анализа, позволяющего выборочно тестировать компоненты целевой программы. Проверка ошибок, 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 (в частности, с ИСП Фаззером) для повышения эффективности Anxiety. 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 символьное 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).

– Реализация специфических задач анализа программ на базе динамического символьного выполнения (в частности, определение достижимости определенной функции или операции в программе);
– Возможность использования для реализации обеспечительных мер ГОСТ Р 56939-2016 (при необходимости сертификации ПО для использования на территории России).

ДЛЯ КОГО ПРЕДНАЗНАЧЕН ANXIETY?

– Компании, нацеленные на разработку ПО с высоким уровнем надёжности и безопасности;
– Компании, отвечающие за аудит или сертификацию ПО.

ОПЫТ ВНЕДРЕНИЯ

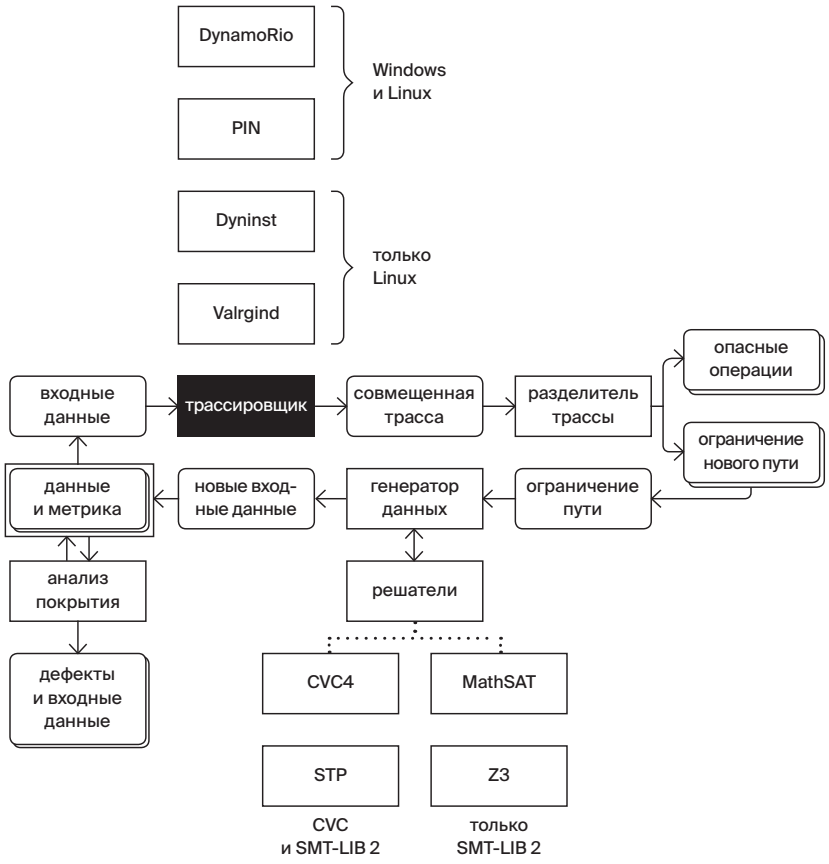
Инструмент Anxiety используется при тестировании программ, входящих в поставку ОС Astra Linux.

ПОДДЕРЖИВАЕМЫЕ СРЕДЫ И ИНСТРУМЕНТЫ

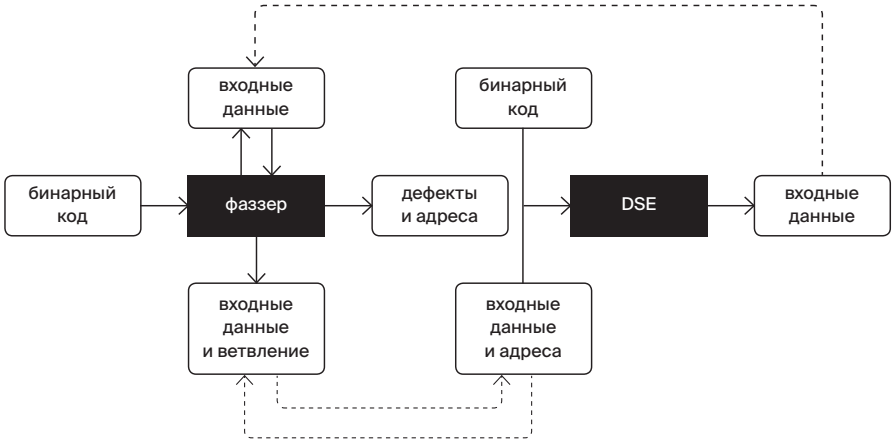
Поддерживает анализ в ОС Windows (начиная с версии XP) и ОС семейства Debian Linux, а также работу различных типов SMT-решателей (STP, Z3, MathSAT и др.). Основан на среде динамической инструментации DynamoRIO (поток инструкций обрабатывается инструментом Triton, обеспечивая поддержку ОС Windows) и среде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

СХЕМА РАБОТЫ

Динамический символьный анализ



Фаззинг


входные данные
бинарный код
бинарный код
фаззер
дефекты и адреса
DSE
входные данные
входные данные и ветвление
входные данные и адреса

СИСТЕМА ВЕРИФИКАЦИИ ASTRAVER TOOLSET

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

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

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 (Microsoft VCC, Frama-C WP), однако благодаря специфической доработке обладает технологической уникальностью: возможностью верификации ключевых компонентов системы безопасности ядра Linux. Выложен в открытый доступ (http://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, приведение целочисленных типов к указательным);
- 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.

– Возможность доработки комплекса под конкретного заказчика (в плане поддержки верификации компонентов на языке С).

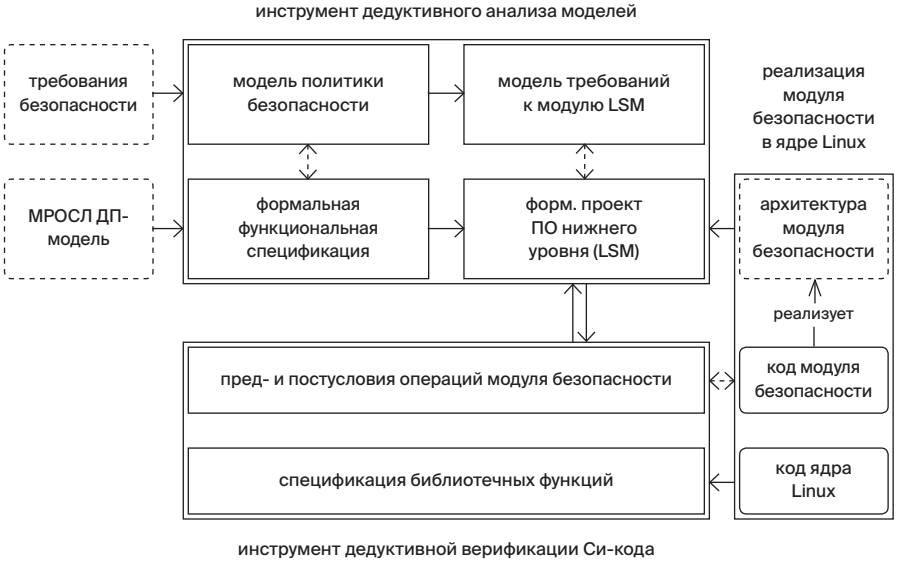
ДЛЯ КОГО ПРЕДНАЗНАЧЕН ASTRAVER TOOLSET?

– Компании, нацеленные на разработку ПО с высокой степенью надёжности и безопасности – как информационной, так и функциональной (ПО для самолётов, 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

BINSIDE: СТАТИЧЕСКИЙ АНАЛИЗАТОР БИНАРНОГО КОДА

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

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

BinSide — инструмент для анализа бинарного кода, разрабатываемый на основе фреймворка BinNavi, который переводит ассемблерный код в представление REIL. Данное представление позволяет анализировать код независимо от процессорной архитектуры и операционной системы. Интегрирован с интерактивным дизассемблером IDA PRO, который используется для обратной разработки.

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).
- Поиск ошибок в работе с памятью на куче (в том числе, использование после освобождения и двойного освобождения).
- Мощный гибкий движок с основными типами анализа:
  - анализ значений и указателей, отслеживание помеченных данных, моделей 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);

- возможность вручную размечать функции в IDA Pro как источники помеченных данных и неаккуратной работы с памятью;
- бинарные файлы импортируются из IDA Pro, что позволяет выполнять анализ над нестандартными или обфусцированными бинарными файлами.

– Плагин для статического анализа патчей программ.
– Высокая скорость (на бинарном файле из более 3000 функций работа инструмента длится около 2000 секунд).
– Возможность конвертировать результаты в формат Svace (при наличии отладочной информации) для отображения в веб-интерфейсе в целях навигации по исходному коду.

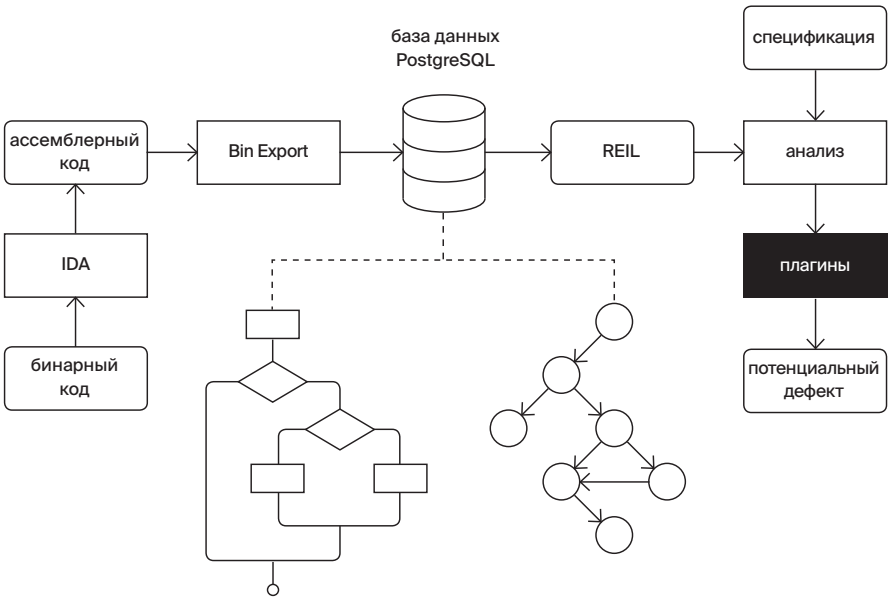
ДЛЯ КОГО ПРЕДНАЗНАЧЕН BINSIDE?

– Компании, которые нуждаются в тщательной проверке стороннего ПО при отсутствии доступа к исходному коду;
– Разработчики, которым требуется повысить качество работы инструментов динамического анализа за счёт дополнительных данных, полученных с помощью статического анализа.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

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

СХЕМА РАБОТЫ



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

Constructivity 4D — технология для создания перспективных программных систем и сервисов, оперирующих динамическими сценами и большими массивами пространственно-временных данных. 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.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

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).
- 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 объёмов, 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;
- 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);
- Оригинальный метод маршрутизации в глобальном динамическом окружении. Основан на извлечении пространственной, метрической и топологической информации, а также на её согласованном использовании при планировании путей;
- Различные возможности расширения библиотеки, которая может использоваться при разработке новых приложений, а также для оптимизации и расширения функций уже существующих.

ДЛЯ КОГО ПРЕДНАЗНАЧЕНА CONSTRUCTIVITY 4D?

Технология используется 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 Synchro (https://www.synchroltd.com), предназначенной для визуального 4D-моделирования, планирования и управления масштабными индустриальными проектами в сфере строительства зданий, инфраструктурных объектов и др. Применяется более чем 300 компаниями в 36 странах (в том числе в России).

ПЛАТФОРМА ДЛЯ СОЗДАНИЯ ЦИФРОВЫХ ДВОЙНИКОВ DIGITEF

DigiTEF — программный комплекс на базе OpenFOAM и утилит других открытых проектов, а также уникальных модулей и библиотек ИСП РАН. Платформа позволяет решать прикладные задачи газовой динамики, аэродинамики, гидродинамики и акустики. Предназначена для создания сложных цифровых моделей промышленных устройств. Включена в Единый реестр российского ПО (№5377).

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Платформа решает те же задачи, что и мировые аналоги. Сравнительные исследования производительности и точности ядра DigiTEF с Ansys Fluent и Star CCM+ показали сопоставимые (а в некоторых случаях и более низкие) вычислительные затраты при одинаковой точности. Вокруг платформы DigiTEF сформировано сообщество инженеров, исследователей и разработчиков промышленных проектов.

DigiTEF — это:

- открытый исходный код (позволяет контролировать и адаптировать реализованные алгоритмы);
- развитие параллельно с веткой OpenFOAM+;
- наличие средств автоматизации вычислений и интеграции моделей для комплексного исследования технических объектов;
- возможность разработки дополнительных компонентов в соответствии с конкретными требованиями.

КОМПЛЕКС СОСТОИТ ИЗ ДВУХ ОСНОВНЫХ БЛОКОВ:

1 OpenDTEF — ядро программного комплекса на основе OpenFOAM. Содержит основные алгоритмы, процедуры и функции, а также набор сторонних библиотек на языке C++. Находится в открытом доступе (https://github.com/unicfdlab) и состоит из следующих компонентов:

- Компонент инструментов для моделирования сжимаемых течений;
- Компонент расширенных настроек расчетного случая на основе swak4Foam;

— Компонент для параметризации на базе Python. Позволяет проводить автоматизацию расчётных случаев, а также осуществлять интеграцию в DigiTEF программных комплексов Salome, Paraview и CodeAster;

2 Компоненты, разработанные в ИСП РАН:

— компонент анализа данных для визуализации и извлечения информации. Предназначен для анализа результатов и построения моделей пониженной размерности с использованием методов обработки данных (FFT, POD, DMD, Hilbert transformations);
— для расчёта сжимаемых течений на основе квазигазодинамических (КГД) уравнений, позволяющих использовать процедуру пространственно-временного осреднения для определения основных газодинамических величин (плотности, скорости и температуры и др.);
— для расчёта несжимаемых течений на основе КГД-уравнений. Компонент применим в задачах океанологии, конвекции и дозвуковых течений;
— для расчёта несжимаемых и сжимаемых течений на основе гибридного алгоритма Pimple и Курганова-Тадмора;
— для расчёта дозвуковых турбулентных течений с использованием гибридного URANS/LES подхода и низко диссипативных численных схем;
— для проведения акустического анализа. В компоненте реализованы аналогии Керла и Фокс Уильямса — Хокинга.

ДЛЯ КОГО ПРЕДНАЗНАЧЕН DIGITEF?

DigiTEF предназначен для использования на предприятиях ресурсоёмких отраслей промышленности. Использование цифровых моделей позволяет повысить эффективность проектирования, а также снизить стоимость и сложность реализации индустриальных проектов.

ОПЫТ ВНЕДРЕНИЯ

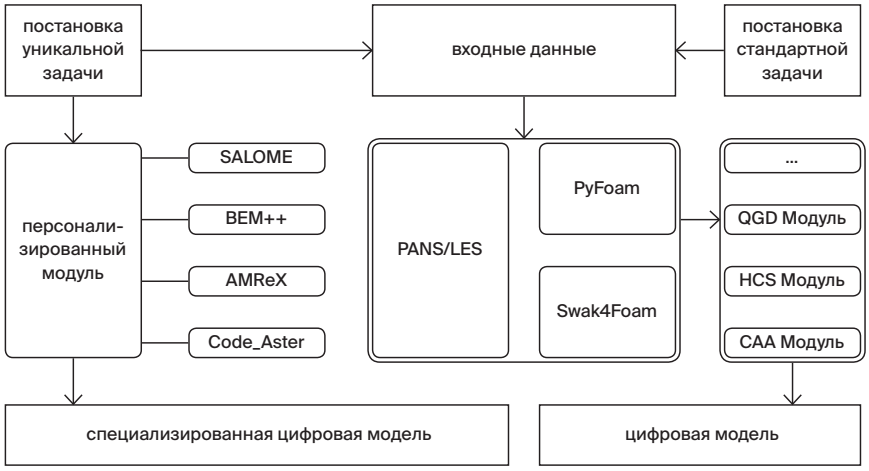
DigiTEF используется в ряде проектов в области ветроэнергетики, космонавтики, авиации, металлургии, а также в нефтегазовой отрасли.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

ОС Linux. Могут также использоваться другие ОС, которые поддерживают виртуальную машину Oracle VirtualBox (на Microsoft Windows 10 — с применением оболочки Bash). В случае их использования падение производительности не превышает 5%.

Требуемая оперативная память — не менее 16 Гб. DigiTEF поддерживает параллельные вычисления, что существенно ускоряет его работу. Кроме того, поддерживается возможность использования высокопроизводительных систем вычислений (суперкомпьютеров и кластеров) для ускорения расчётов. Проверенное количество вычислительных ядер — до 1536.

СХЕМА РАБОТЫ



Digital Test Facility

KLEVER: ТЕХНОЛОГИЯ ВЕРИФИКАЦИИ МОДЕЛЕЙ КРУПНЫХ ПРОГРАММНЫХ СИСТЕМ

Klever — система верификации моделей, полученных на основе исходного кода крупных программных систем, разработанных на языке программирования Си, для проверки требований безопасности, надёжности и производительности.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Klever — технология, основанная на научных исследованиях в области выполнения полностью автоматического доказательства корректности моделей программ, извлекаемых из исходного кода (в том числе без участия пользователя).

В основе системы лежат методы для покомпонентной верификации исходного кода программных систем размером в миллионы строк кода на языке Си. Технология выложена в открытый доступ (https://forge.ispras.ru/projects/klever).

Klever — это:

- Высокоточный консервативный анализ исходного кода любого сложного ПО (выявление всех возможных ошибок искомых видов при явно заданных предположениях);
- Проверка расширяемого набора требований к программе (проверка правил безопасного программирования на языке Си и корректности использования интерфейса, специфичного для проверяемой программы);
- Масштабируемость. Модульная верификация программ позволяет применять наиболее точные методы для анализа исходного кода — в частности, методы проверки моделей и символьного выполнения к большому объёму кода;
- Подробные сведения об ошибках. Система верификации не просто указывает на место ошибки в исходном коде, а предоставляет всю последовательность команд для воспроизведения ошибки в удобном пользовательском интерфейсе. При необходимости на основе таких трасс могут быть автоматически сгенерированы тесты;

– Возможность доработки технологии под конкретные нужды. Оперативное расширение списка обнаруживаемых ошибок. Разработка набора спецификаций для формализации специфичных для программы требований, а также для моделирования окружения и в некоторых случаях – плагинов;
– Удобный многопользовательский веб-интерфейс для выполнения статической верификации, а также хранения, анализа и сравнения результатов.

ДЛЯ КОГО ПРЕДНАЗНАЧЕН KLEVER?

– Компании, нацеленные на разработку ПО с высоким уровнем надёжности и безопасности;
– Сертификационные лаборатории.

ОПЫТ ВНЕДРЕНИЯ

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

Для демонстрации возможностей системы были выполнены работы по верификации драйверов устройств и подсистем операционной системы Linux. В результате удалось обнаружить более 300 ошибок, подтверждённых 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 структур данных ядра Linux.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

Ubuntu 18.04, 16 гигабайт оперативной памяти, от 100 Гб свободного места на диске.

СХЕМА РАБОТЫ

LINGVODOC: 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 — lingvodoc.ispras.ru.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Lingvodoc — кроссплатформенная технология с открытым исходным кодом (https://github.com/ispras/lingvodoc и https://github.com/ispras/lingvodoc-react), основанная на научных исследованиях.

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 поиска данных на карте с автоматическим построением изоглосс;
- Возможность бесконфликтной двусторонней отложенной синхронизации;
- Повышенный уровень автоматизации (по сравнению с аналогичным проектом Kielipankki);

– 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 VMEmperor), 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 с расширением набора функций, а также адаптация под другую научную отрасль.

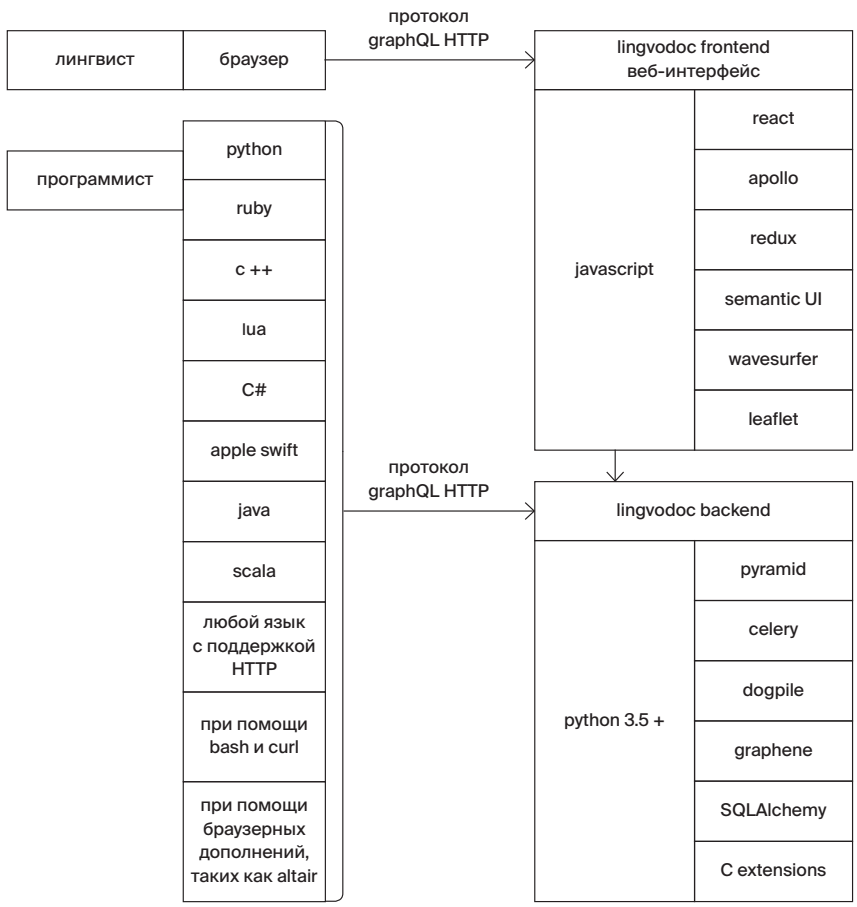
ДЛЯ КОГО ПРЕДНАЗНАЧЕН LINGVODOC?

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

ОПЫТ ВНЕДРЕНИЯ

В настоящее время Lingvodoc используется в рамках совместных проектов с Институтом языкознания РАН и Томским государственным университетом.

СХЕМА РАБОТЫ



MASIW: ПОДДЕРЖКА ПРОЕКТИРОВАНИЯ ОТВЕТСТВЕННЫХ СИСТЕМ

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.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

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. 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.

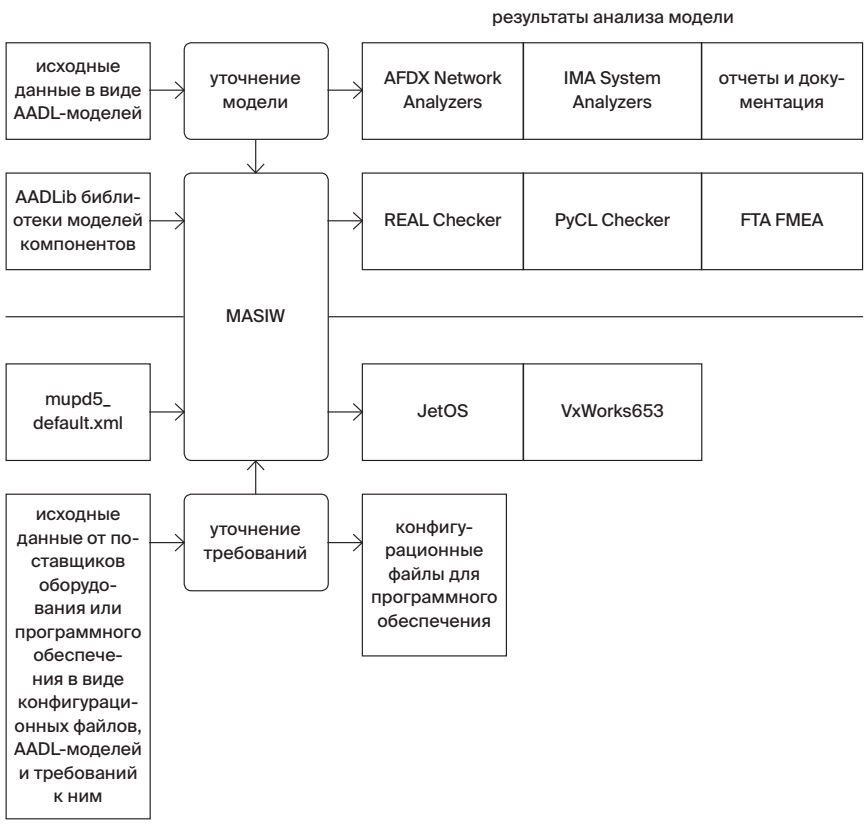
MASIW — это:

- 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 РВ 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.
– Возможность расширения набора инструментов путём создания собственных модулей (благодаря модульной архитектуре в основе технологии).

СХЕМА РАБОТЫ



ГЕНЕРАТОР ТЕСТОВЫХ ПРОГРАММ MICROTESK

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).

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

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 (GenesysPro и RAVEN), однако отличается от них повышенной производительностью и удобством использования, а также распространением по лицензии открытого исходного кода.

Выложен в открытом доступе на сайте ИСП РАН: https://forge.ispras.ru/projects/microtesk. Описание технологии доступно на сайте 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 удобны 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 RISC-V, ARM, MIPS, PowerPC;
    - поддерживается многоядерность целевой микропроцессорной архитектуры.
- 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.

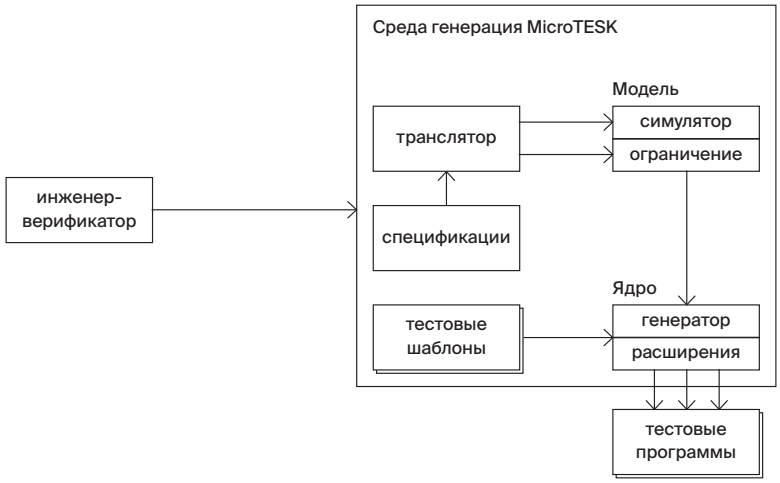
СИСТЕМНЫЕ ТРЕБОВАНИЯ

ОС Windows или ОС на базе ядра GNU/Linux, Java 8.

ОПЫТ ВНЕДРЕНИЯ

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

СХЕМА РАБОТЫ



СИСТЕМА АНАЛИЗА СЕТЕВОГО ТРАФИКА PROTOSPHERE

Protosphere — система глубокого анализа сетевого трафика (DPI). Может встраиваться как компонент в системы мониторинга, классификации, защиты от вторжений и утечек информации. Регистрирует несоответствия между реализацией протокола и фактическим трафиком. Позволяет быстро добавлять поддержку новых (в том числе закрытых) протоколов благодаря универсальности внутреннего представления.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Protosphere — инновационная система, основанная на научных исследованиях технологий анализа сетевого трафика. Объединяет ключевые особенности иностранных аналогов (Wireshark, Microsoft Message Analyzer) с универсальным внутренним представлением, позволяющим быстро расширять возможности анализа.

Protosphere — это:

- Оптимальные возможности ядра системы:
  - универсальная модель представления данных при разборе 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 связанных потоков.
- Поддержка всех этапов анализа сетевой трассы — каждый этап с отдельным компонентом визуализации, все компоненты синхронизированы:
  - локализация одного или нескольких исследуемых сетевых соединений на графе сетевых взаимодействий и в дереве сетевых потоков;
  - детализация выделенных соединений на временной диаграмме;
  - наглядное представление выделенных в сетевых пакетах полей в дереве разбора сетевого потока;
  - выявление несоответствий между реализацией протокола и фактическим трафиком в журнале диагностики;
  - извлечение и анализ данных произвольного уровня (L7+).
- Быстрое расширение списка поддерживаемых протоколов:
  - API доступ к результатам разбора;
  - локализация ошибок разбора;
  - возможность отладки разрабатываемого модуля на потоке, позволяющая существенно ускорить поддержку новых протоколов.
- Поддержка двух режимов работы: на потоке и в отложенном режиме.

- Продвинутый графический интерфейс, позволяющий выбирать наиболее удобный вариант представления результатов проводимого анализа.
- 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/точностью анализа и потребляемыми ресурсами.

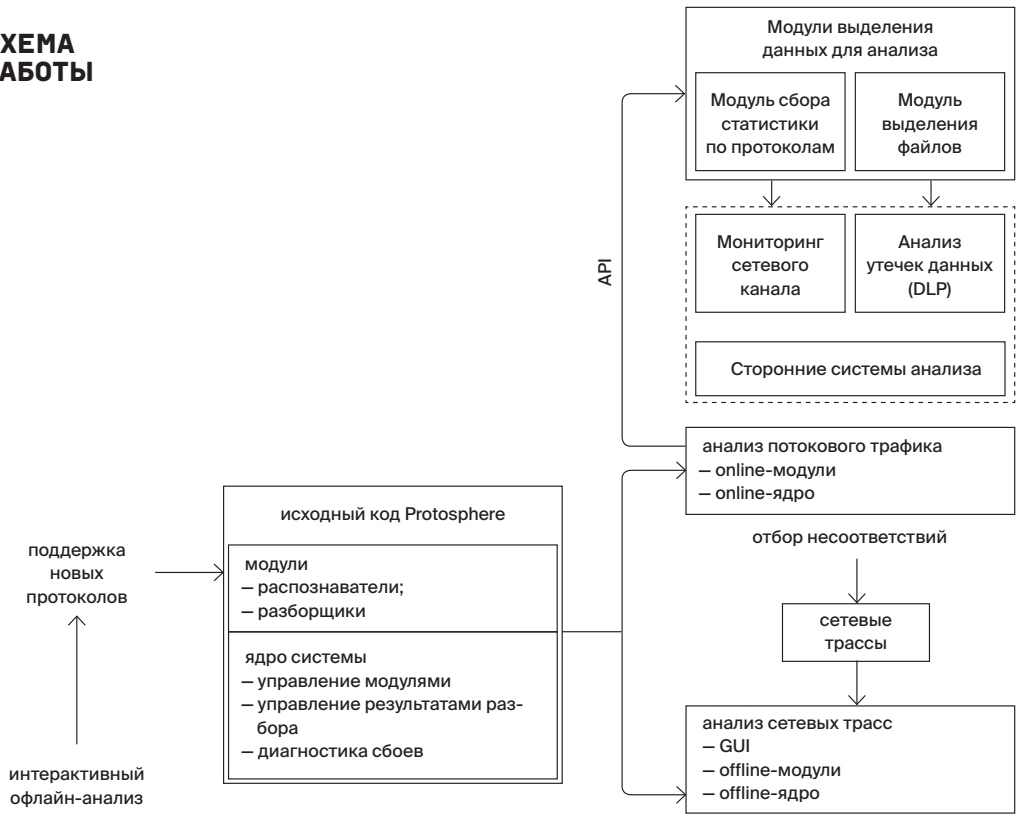
ДЛЯ КОГО ПРЕДНАЗНАЧЕНА PROTOSPHERE?

- 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;
- Компании, которым требуется контроль и мониторинг сетевых каналов в режиме реального времени.

ПОДДЕРЖИВАЕМЫЕ ПЛАТФОРМЫ И АРХИТЕКТУРЫ

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

СХЕМА РАБОТЫ



ПЛАТФОРМА ДЛЯ АНАЛИЗА ПРОГРАММ НА ОСНОВЕ ЭМУЛЯТОРА QEMU

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

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

QEMU поддерживает более 10 архитектур процессоров (i386 и Intel 64, ARM и Thumb, MIPS, PowerPC и др.). Реализует отладку по удаленному GDB-протоколу и совместим с IDA Pro, GDB и средами разработки. В режиме полносистемной эмуляции подходит для отладки низкоуровневого ПО — такого, как загрузчик и ОС. Исходный код QEMU систематически проверяется двумя статическими анализаторами (Coverity и Svace), что делает анализ потенциально вредоносного ПО в эмуляторе более безопасным.

Эмулятор с поддержкой обратной отладки и интроспекции доступен на GitHub: https://github.com/ispras/swat, как и набор инструментов автоматизации: https://github.com/ispras/qdt, https://github.com/ispras/i3s.

Платформа ИСП РАН на основе QEMU — это:

- Запись и воспроизведение работы виртуальной машины:
  - При каждом воспроизведении виртуальная машина ведёт себя одинаково и точно так же, как при записи. Все воздействия извне зафиксированы и повторяются самим эмулятором, что упрощает отладку ошибок, связанных с параллельной работой приложения (состояние гонки, взаимные блокировки);
  - На базе воспроизведения реализована GDB-совместимая обратная отладка, которая заключается в откате к предыдущим снимкам состояния виртуальной машины и поиске предпоследнего срабатывания точки останова или предыдущей инструкции;
  - Записывается минимум информации, что позволяет вести длительную запись, необходимую для отладки редко повторяющихся ошибок;
  - Низкое относительное замедление, вносимое записью, позволяет контролировать ПО, требующее взаимодействия с удалённой системой в режиме реального времени.

- Получение высокоуровневой информации о работе гостевой ОС (интроспекция ВМ) без внесения каких-либо изменений в ядро ОС или установки программ мониторинга:
  - Возможность получить последовательность совершаемых системных вызовов, обращений к именованным функциям в динамических библиотеках, список работающих процессов, список открытых файлов и загруженных в память модулей;
  - Поддержка любого образа виртуальной машины на основе Linux, в том числе — образов встраиваемого ПО различных устройств;
  - Отладка с помощью встроенного в эмулятор сервера WinDbg, что позволяет отображать информацию о гостевом ПО в терминах абстракций ядра Windows. При этом не требуется включение отладочного режима работы гостевой ОС;
- Ускорение разработки расширений для QEMU:
  - Сокращение времени на подготовку средств динамического анализа для образцов кода, требующих специализированной аппаратуры;
  - Автоматизированное добавление процессорных архитектур с использованием генератора декодеров машинных команд и C-подобного языка описания семантики инструкций;
  - Система автоматического первичного тестирования виртуальной машины. Для работы системы требуются только утилиты GNU Binutils и компилятор языка C;
  - Автоматизированная разработка моделей устройств;
  - Генерация виртуальной машины (в форме исходного кода модуля QEMU) как из существующих, так и из новых устройств по описанию на языке Python с использованием графического интерфейса пользователя со схематичным изображением машины;
  - API для автоматизации процесса отладки на языке Python по протоколу GDB RSP: отладка гостевого кода, кода эмулятора и обоих одновременно.
- Удобство практического использования:
  - Свободное расширение возможностей QEMU благодаря открытому исходному коду и собственным инструментам ускоренной разработки ИСП РАН;
  - Анализ бинарного кода без внедрения программ в гостевую систему;
  - Модульная структура механизма интроспекции с возможностью расширения за счёт новых плагинов;
  - Удобное API для самостоятельной разработки плагинов интроспекции;
  - Возможность адаптации под конкретные нужды пользователя;
  - Поддержка актуальных версий QEMU с новой периферией и процессорными ядрами.

ДЛЯ КОГО ПРЕДНАЗНАЧЕНА ПЛАТФОРМА НА БАЗЕ QEMU?

- Разработчики загрузчиков, драйверов, ОС и другого системного ПО;
- DevOps-команды (воспроизводимость ошибок, кросс-разработка, масштабирование тестирования в облачной среде);
- Аналитики потенциально вредоносного ПО;
- Специалисты по сертификации ПО.

ПОДДЕРЖИВАЕМЫЕ ГОСТЕВЫЕ СРЕДЫ

Эмулируемые платформы: i386, x86-64, ARM, MIPS, PowerPC и другие.
Гостевые системы, поддерживаемые интроспекцией: Windows XP (x86), Windows 10 (x86-64) и Linux 2.x-4.x на платформах x86, x86-64, ARM, AArch64.

ОПЫТ ВНЕДРЕНИЯ

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

СХЕМА РАБОТЫ

RETRASCOPE: ИНСТРУМЕНТ СТАТИЧЕСКОГО АНАЛИЗА HDL-ОПИСАНИЙ

Retrascope — инструмент функциональной верификации модулей цифровой аппаратуры. Retrascope предоставляет автоматизированные средства анализа кода, извлечения формальных моделей и генерации функциональных тестов. В качестве входных данных инструмент принимает описания модулей цифровой аппаратуры на синтезируемых подмножествах языков Verilog и VHDL, а также спецификации поведения.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Retrascope — открытый инструмент функциональной верификации модулей цифровой аппаратуры. Инструмент реализует ряд методов извлечения и анализа формальных моделей, а также генерации функциональных тестов. Модульная архитектура Retrascope позволяет разрабатывать гибридные техники верификации HDL-описаний за счёт комбинирования различных средств анализа формальных моделей. Retrascope доступен на сайте ИСП РАН: https://forge.ispras.ru/projects/retrascope.

Retrascope — это:

- Извлечение формальных моделей из исходного кода:
  - граф потока управления;
  - решающая диаграмма охраняемых действий;
  - высокоуровневая решающая диаграмма;
  - расширенный конечный автомат.
- Генерация функциональных тестов:
  - случайные тесты;
  - выявление недостижимого кода;
  - выявление типовых ошибок;
  - проверка пользовательских свойств.
- Проверка формальных моделей (model checking) на соответствие спецификациям:
  - PSL;
  - SystemVerilog Assertions.
- Графический интерфейс на основе Eclipse IDE (также доступен интерфейс командной строки):
  - запуск инструмента с параметрами;
  - визуализация извлеченных моделей (Zest, GraphML).

- Открытый исходный код (лицензия Apache License Version 2.0);
- Расширяемость на уровне исходного кода:
  - добавление новых моделей;
  - расширение набора средств анализа.
- Открытые интерфейсы взаимодействия позволяют использовать различные средства для достижения целей анализа и верификации без изменения кода инструмента:
  - SMT-решатели — язык SMT-LIB v2;
  - Средства проверки моделей — язык SMV.

ДЛЯ КОГО ПРЕДНАЗНАЧЕН RETRASCOPE?

- Компании, занимающиеся проектированием цифровой аппаратуры;
- Коллективы, проводящие исследования в области функциональной верификации цифровой аппаратуры.

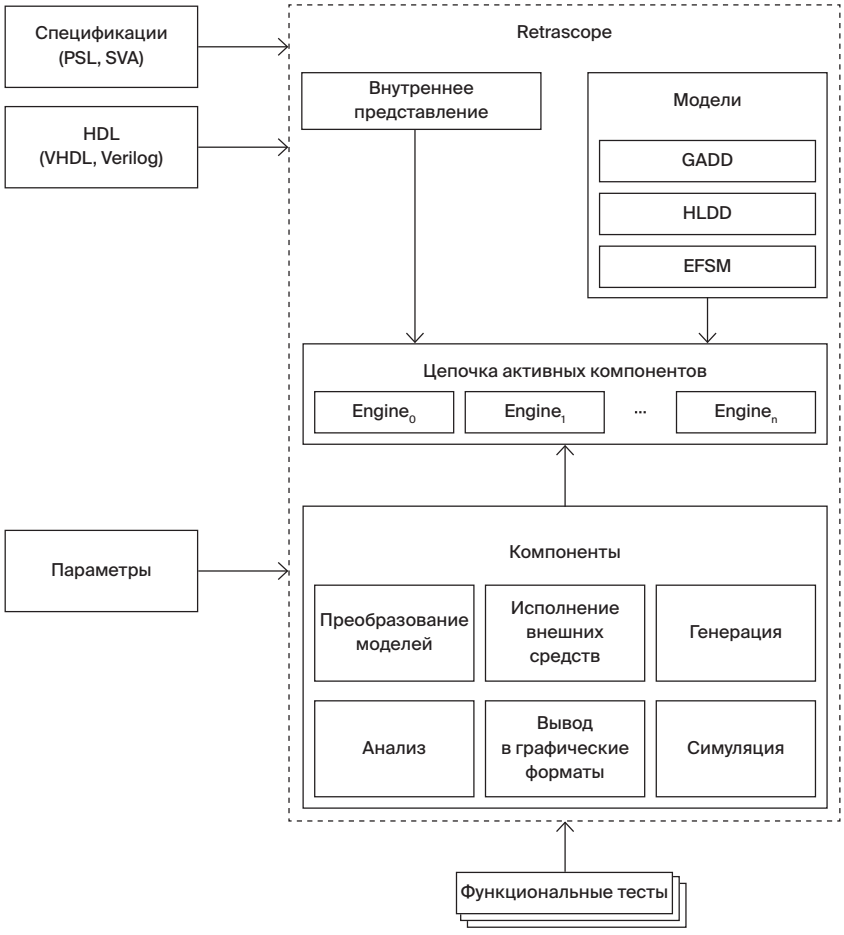
ОПЫТ ВНЕДРЕНИЯ

Инструмент находится на стадии исследовательского прототипа, ведётся разработка.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

Программное обеспечение: ОС Windows или ОС на базе ядра GNU/Linux, Java 8.

СХЕМА РАБОТЫ



СИСТЕМА ИССЛЕДОВАТЕЛЬСКОГО ПОИСКА SCINOON

SciNoon — система совместного исследовательского поиска научных статей. Позволяет группе исследователей быстро погружаться в новую предметную область и находить ответы на свои вопросы, а затем отслеживать новые публикации по изучаемой тематике.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

SciNoon — инновационная система, созданная с целью оптимизации длительной командной работы с научными публикациями. Статьи в SciNoon можно добавлять как из широко известных поисковых систем и электронных библиотек (Google Scholar, arxiv.org, Semantic Scholar, PubMed), так и с помощью загрузки PDF-файлов. Уникальная особенность — графические карты исследований, на которые все члены группы могут добавлять найденные ими статьи.

SciNoon — это:

- Общее рабочее место для совместной обработки публикаций;
- Возможность масштабирования карты исследований для управления степенью подробности отображаемой информации о статьях;
- Нормализация и дедупликация метаданных загружаемых статей благодаря внутренней базе данных. Построение связей между статьями и авторами;
- Классификация контекстов цитирований в один из пяти классов (с точки зрения цели цитирования):
  - Background (цитируемая статья содержит общую информацию, относящуюся к области исследования в рассматриваемой статье);
  - Use (рассматриваемая статья использует методы, данные и т.д. из цитируемой статьи);
  - Compare (рассматриваемая статья указывает на различия/сходства с цитируемой статьёй);
  - Extend (рассматриваемая статья продолжает развитие методов из цитируемой статьи);
  - Weak (рассматриваемая статья критикует цитируемую статью, указывает на ошибки авторов).
- Возможность находить релевантные исследованию статьи без использования поиска по ключевым словам (благодаря встроенной рекомендательной системе);
- Настройка собственного перечня вопросов, на которые надо искать ответы в статьях. В зависимости от полученных ответов можно по-разному отображать статью на карте исследований;
- Объединение близких статей в кластеры;

- Возможность получать уведомления о действиях каждого исследователя в команде, а также оперативно обмениваться мнениями и помогать друг другу;
- Анализ всех собранных ответов на вопросы с помощью встроенной табличной формы представления, а также экспорт в формате CSV (если требуется более сложная обработка);
- Отслеживание новых статей по рассмотренной тематике после завершения исследования и быстрая актуализация первоначально полученных результатов.

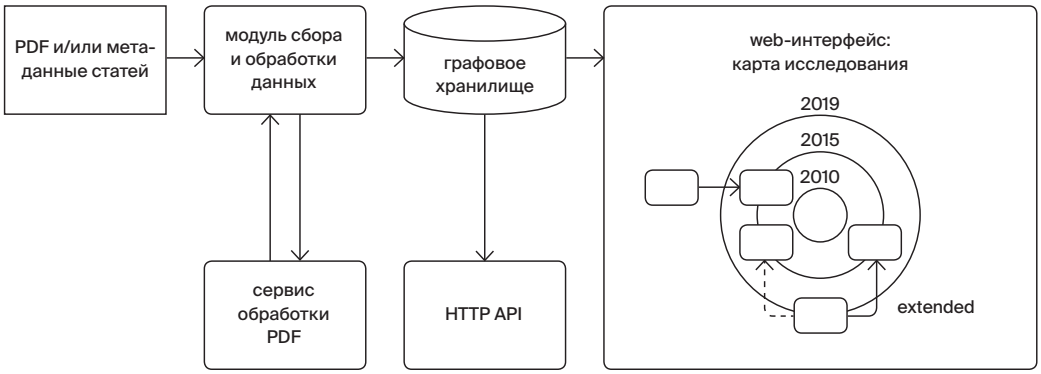
ДЛЯ КОГО ПРЕДНАЗНАЧЕН SCINOON?

- Сотрудники R&D отделов корпораций, которым нужно быстро найти решение возникшей научной задачи;
- Сотрудники научно-исследовательских институтов, нуждающиеся в инструменте для командной работы;
- Преподаватели и студенты ВУЗов, занимающиеся исследовательским поиском для подготовки научных работ.

ОПЫТ ВНЕДРЕНИЯ

SciNoon используется в ИСП РАН при проведении исследований и при руководстве студентами.

СХЕМА РАБОТЫ



СТАТИЧЕСКИЙ АНАЛИЗАТОР SVACE

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#, Java. Добавление поддержки Kotlin и Go планируется в конце 2020 г. Включён в Единый реестр российского ПО (№4047).

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Svace — постоянно развивающийся инновационный продукт, основанный на многолетних исследованиях. Объединяет ключевые качества иностранных аналогов (Coverity Scan Static Analysis, Fortify Static Code Analyzer, Klocwork Static Code Analysis) с уникальным использованием открытых промышленных компиляторов в целях максимальной поддержки новых стандартов языков программирования.

Svace — это:

- Высокое качество анализа:
  - 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 всех доступных 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.
- 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);
- Полная совместимость с нормативными документами и требованиями регуляторов (ФСТЭК РФ).
- Возможность использования для реализации обеспечительных мер ГОСТ Р 56939-2016 (при необходимости

сертификации ПО для использования на территории России), в том числе полная совместимость с новой «Методикой выявления уязвимостей и недекларированных возможностей в программном обеспечении» ФСТЭК России.

ДЛЯ КОГО ПРЕДНАЗНАЧЕН SVACE?

- Компании, нацеленные на разработку ПО с высокой степенью надёжности и безопасности;
- Компании, которые нуждаются в сертификации разрабатываемого ПО;
- Сертификационные лаборатории.

ОПЫТ ВНЕДРЕНИЯ

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. Внедрён в ОАО «РусБИТех» и «ОКБ Сухого».

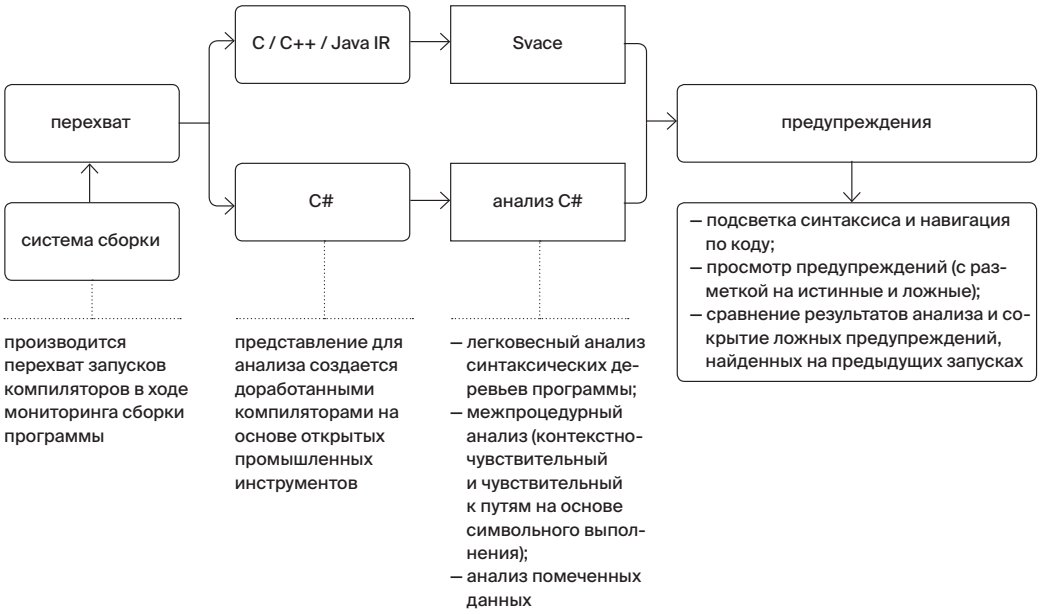
ПОДДЕРЖИВАЕМЫЕ ПЛАТФОРМЫ И АРХИТЕКТУРЫ

- Платформы, для которых предназначается анализируемый код: ОС на базе ядра Linux (начиная с версии 2.6), ОС Windows (начиная с XP).
- Архитектуры: 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



ФРЕЙМВОРК ДЛЯ АНАЛИЗА СОЦИАЛЬНЫХ МЕДИА TALISMAN

Talisman — фреймворк для анализа данных о людях, сообществах, продуктах и организациях. Основан на современных методах машинного обучения, компьютерной лингвистики, анализа сложных сетей и обработки больших данных. Выявляет закономерности во взаимосвязях с помощью анализа графов из сотен миллионов узлов.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Talisman интегрирован с платформой для извлечения семантики из текста (Texterra) и оригинальной технологией сбора данных ИСП РАН. 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 облачной среды Asperitas (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;
- Работа как в режиме облачного сервиса, так и на оборудовании заказчика;
- Оперативная адаптация и расширение функционала для использования в различных предметных областях (информационная безопасность, медицина, аудит и др.).

ФРЕЙМВОРК ПРЕДСТАВЛЕН ДВУМЯ ВЗАИМОДОПОЛНЯЮЩИМИ ПРОДУКТАМИ:

1. TALISMAN. ПОТОК.

Система для предобработки потока больших данных из социальных медиа. Масштабируемый программный фреймворк с микросервисной архитектурой, разработанный на основе свободного ПО. Повышает продуктивность разработки прикладных систем анализа за счет объединения нескольких обработчиков потоковых данных. Включён в Единый реестр российского ПО (№6045). http://talisman.ispras.ru/talisman-se-поток/

2. TALISMAN. БИОГРАФИЯ.

Система для анализа больших данных из социальных медиа. Обеспечивает автоматизированное пополнение анкетных данных сотрудников на основе информации из социальных медиа и других источников. Включена в Единый реестр российского ПО (№5547). http://talisman.ispras.ru/talisman-se-биография/

ОБЛАСТИ ПРИМЕНЕНИЯ

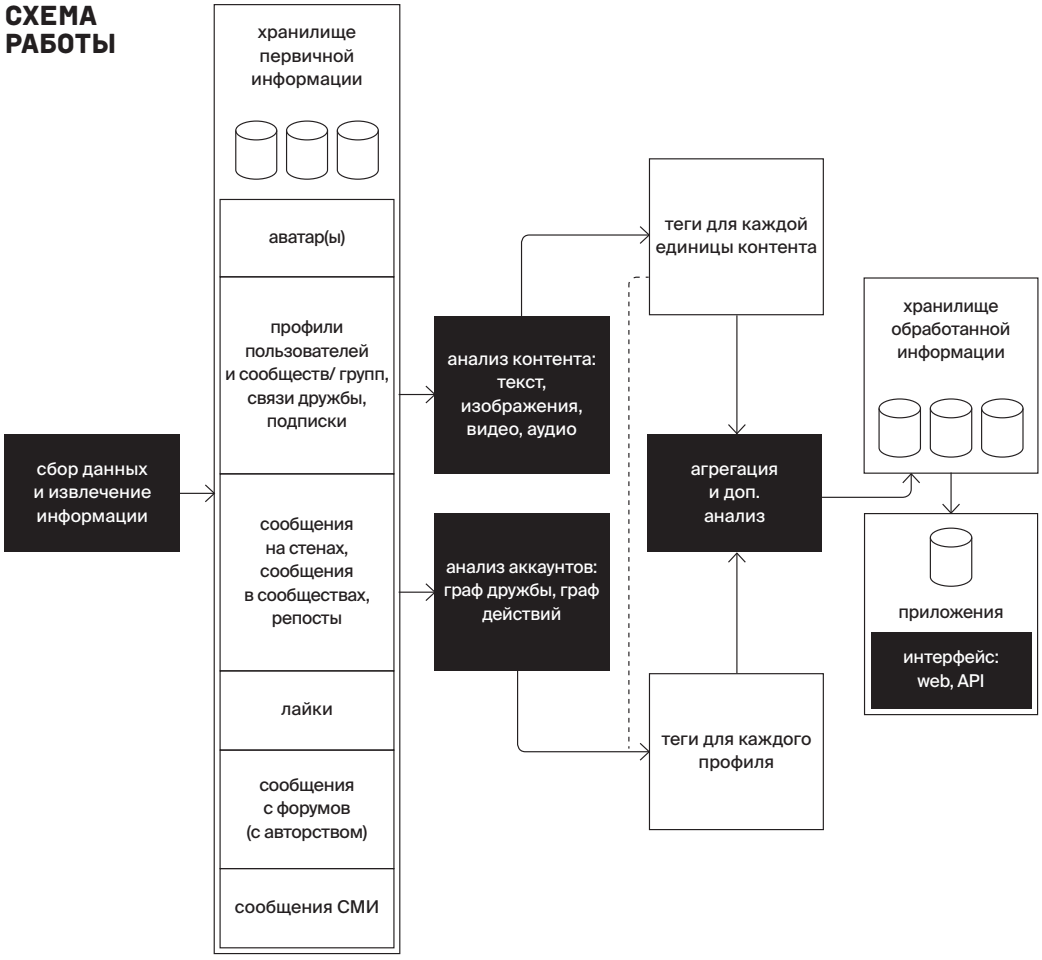
- Выявление групп по интересам на основе анализа текстов социальных медиа, в том числе: определение целевой аудитории (как в маркетинговых целях, так и при формировании политических программ), выявление точек социального напряжения и злободневных проблем с наибольшим числом недовольных;
- Выяснение общественного 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 долговременных 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



БАЗОВЫЙ СЕМАНТИЧЕСКИЙ АНАЛИЗАТОР TEXTERRA

Texterra — масштабируемая платформа для извлечения семантики из текста. Базовый комплекс технологий для создания многофункциональных прикладных приложений. Анализирует тексты с помощью выделения концептов. Включена в Единый реестр российского ПО (№4048).

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

Texterra осуществляет уникальный анализ русскоязычных текстов на основе выделения концептов, а не только слов. Отличается от иностранных аналогов преимущественным вниманием к русскому языку. Базируется 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 облачной среды Asperitas (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.
- 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 https://texterra.ispras.ru/;
- 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 прикладных систем обработки текста.

ОПЫТ ВНЕДРЕНИЯ

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.

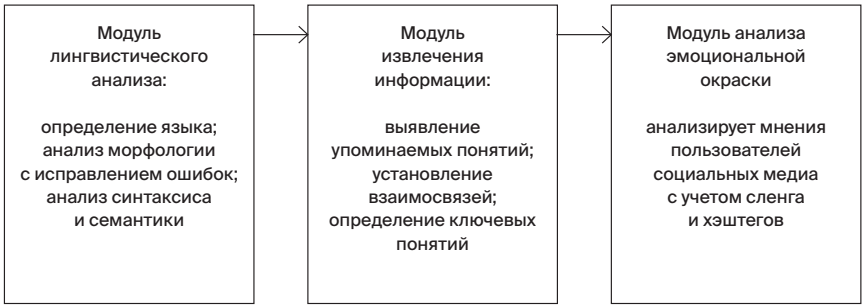
ПОДДЕРЖИВАЕМЫЕ ЯЗЫКИ

Texterra анализирует тексты на русском и английском языках.

СИСТЕМНЫЕ ТРЕБОВАНИЯ

- Любые платформы, поддерживаемые Java 8;
- Не менее 16 Гб оперативной памяти для каждого из анализируемых языков;
- Рекомендуется применение 64-битной версии ОС.

СХЕМА РАБОТЫ



ИСП 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.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

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;
- Использование оригинального метода обеспечения целостности потока управления (CFI), который успешно противодействует большинству атак с повторным использованием кода (ROP, JOP, ret-to-plt и др.). На базе компилятора GCC реализован прототип CFI, который показал среднее замедление на наборе тестов SPEC CPU2006 около 2%, что заметно ниже, чем у традиционных методов;
- 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 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.
- Метод обеспечения целостности потока управления (CFI).

— 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.

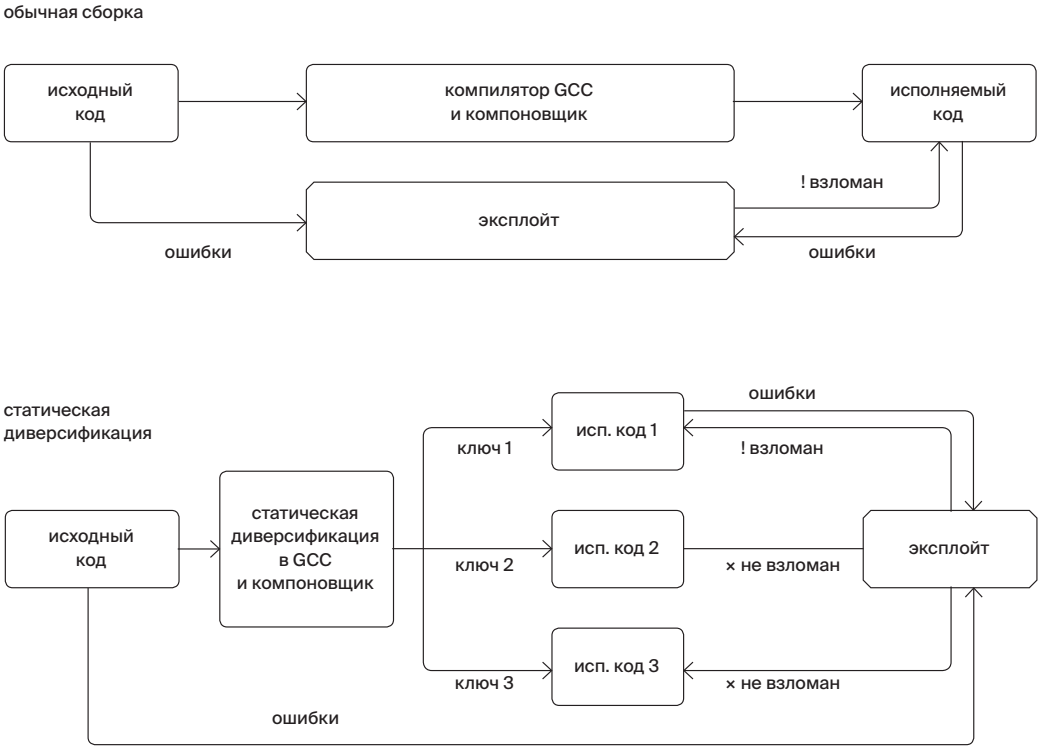
ОПЫТ ВНЕДРЕНИЯ

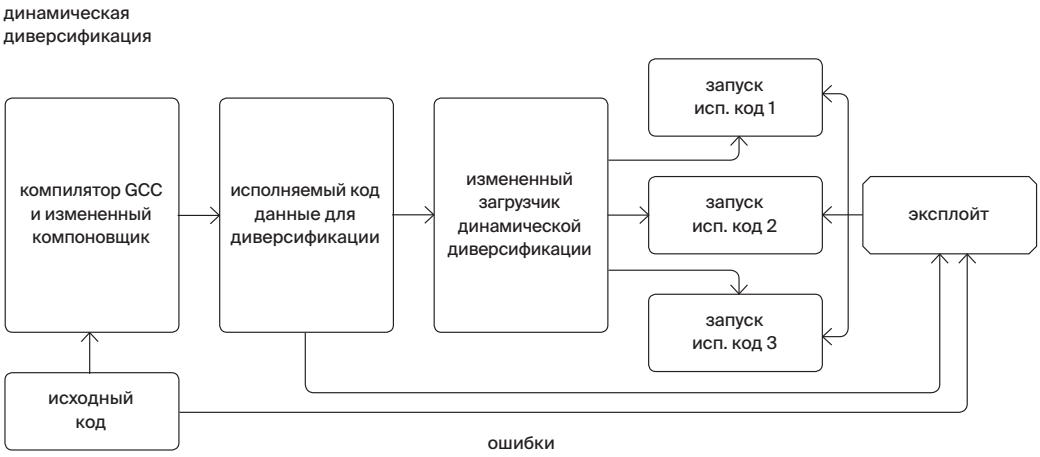
ИСП Обфускатор внедрен в ОС «Циркон», которую используют МИД и Пограничная служба ФСБ России.

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




динамическая
диверсификация
компилятор GCC
и измененный
компоновщик
исполняемый код
данные для
диверсификации
измененный
загрузчик
динамической
диверсификации
запуск
исп. код 1
запуск
исп. код 2
запуск
исп. код 3
эксплойт
исходный
код
ошибки

ТРАЛ: СРЕДА АНАЛИЗА БИНАРНОГО КОДА

ТРАЛ — уникальный промышленный инструмент для анализа свойств бинарного кода. Позволяет работать с кодом различных целевых процессорных архитектур. Не требует наличия отладочной информации и исходных кодов. Применим для анализа всего программного стека от загрузчика до прикладного ПО. Включён в Единый реестр российского ПО (№5323).

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

ТРАЛ — комплекс технологий, основанный на многолетнем опыте разработчиков компиляторов и специалистов по информационной безопасности. В отличие от аналогичных научно-исследовательских технологий в области анализа бинарного кода, доработан до промышленного использования.

Ключевые возможности:

- восстановление потоков данных и управления на уровне машинных команд без ограничений;
- локализация в коде отдельных алгоритмов, формальное представление их структуры и семантики;
- автоматизация ручного анализа, полностью автоматическое решение многих прикладных задач.

ТРАЛ — это:

- Модульная архитектура среды (позволяет расширять набор поддерживаемых целевых платформ и развивать функциональное наполнение среды);
- Поддержка автоматизации анализа с помощью сценариев и открытого API (предоставляет возможность интегрировать среду с другими инструментами: IDA Pro и Wireshark);
- Глубокий анализ:
  - для анализа достаточно наличия лишь исполняемого бинарного кода;
  - в основе подхода – динамический анализ по трассам выполнения, при необходимости дополняемый статическим анализом снимков памяти;
  - предварительное автоматическое повышение уровня представления;
  - восстановление статического представления программ, входящих в состав анализируемой системы, в том числе по нескольким запускам;
  - точный анализ потоков данных, учитывающий особенности аппаратуры (конвейер команд, прерывания, трансляция виртуальных адресов, DMA);

- интерактивное восстановление блок-схемы алгоритма, основанное на построении срезов информационных потоков;
- подход, реализованный в среде, невосприимчив к большинству известных приёмов противодействия анализу.

- Высокая производительность:
  - параллельный анализ с высокими показателями масштабируемости на многоядерных рабочих станциях;
  - возможность анализа длительных сценариев работы анализируемой системы.
- Развитый графический интерфейс:
  - просмотр трасс выполнения с обширными возможностями поиска и навигации, аналогичными классическому отладчику, но с возможностью мгновенного перемещения по потокам данных как вперёд, так и назад во времени;
  - автоматическая разметка высокоуровневой структуры трассы: процессов и потоков выполнения, обработчиков прерываний, стеков вызовов, динамически загружаемых модулей и символов в них;
  - просмотр значений параметров и возвращаемых значений вызванных функций;
  - разметка трассы с указанием внешних событий (сетевые взаимодействия и пользовательский ввод-вывод) и событий, связанных с работой аппаратуры.

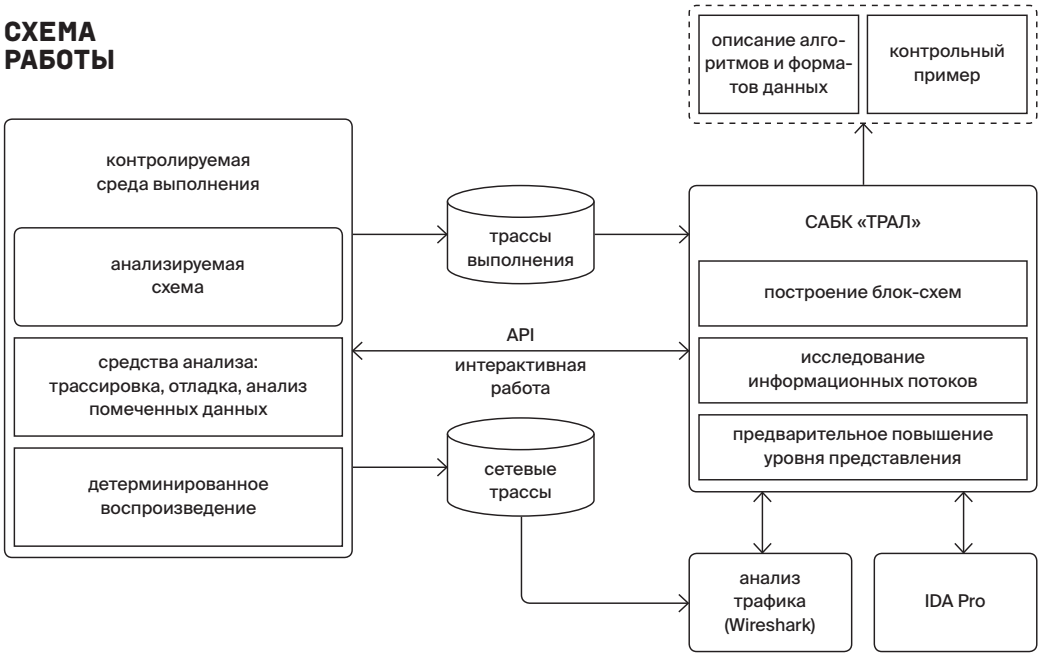
ДЛЯ КОГО ПРЕДНАЗНАЧЕН ТРАЛ?

- Лаборатории, проводящие анализ вредоносного кода;
- Компании-разработчики встраиваемого ПО и компонентов ОС;
- Сертификационные лаборатории.

ПОДДЕРЖИВАЕМЫЕ ПЛАТФОРМЫ И АРХИТЕКТУРЫ

- Системные требования среды анализа ТРАЛ: ОС Windows или ОС на базе ядра Linux, 64-разрядный процессор архитектуры x86, 16 и более Гбайт ОЗУ.
- Целевые процессорные архитектуры: x86, x86-64, ARMv6, ARMv7.
- Целевые ОС: семейство Windows, семейство Linux, поддерживается возможность работы с неопознанной ОС и с кодом, работающим вне ОС.

СХЕМА РАБОТЫ



ИНСТРУМЕНТ ТЕСТИРОВАНИЯ ИСП ФАЗЗЕР

ИСП Фаззер — инструмент проведения фаззинг-тестирования. Позволяет осуществлять динамический анализ ПО. Обнаруживает ошибки или закладки как при наличии, так и при отсутствии исходного кода. Позволяет построить процесс разработки в соответствии с ГОСТ Р 56939-2016.

ОСОБЕННОСТИ И ПРЕИМУЩЕСТВА

ИСП Фаззер — инструмент динамического анализа, необходимый на всех этапах разработки, тестирования и эксплуатации ПО. Решает те же задачи, что и мировые аналоги (Synopsys Codenomicon, beSTORM, Peach Fuzzer), однако более удобен для российских компаний в условиях процесса импортозамещения.

ИСП Фаззер — это:

- Осуществление фаззинг-тестирования через различные источники внешних данных (файл, аргументы командной строки, стандартный поток ввода, аргументы переменных окружений, сеть);
- Возможность добавления пользовательских мутационных преобразований (для генерации новых входных данных и увеличения эффективности тестирования);
- Наличие модулей пред- и постобработки входных данных для осуществления константных преобразований над данными перед их отправкой в анализируемое ПО;
- Поддержка многопоточного анализа и на одной машине, и на распределённых;
- Поддержка пользовательских плагинов отправки данных по сети (плагины позволяют осуществлять взаимодействие с клиентским или серверным ПО и отправлять мутированные данные);
- Возможность интеграции с рядом необходимых инструментов жизненного цикла разработки безопасного ПО, созданных в ИСП РАН:
  - использование динамического анализатора Anxiety (ИСП РАН) для преодоления условных переходов, которые долгое время не получается пройти с помощью фаззинг-тестирования;
  - возможность получать входные данные, на которых проявляются ошибки, размеченные инструментом статического анализа BinSide в автоматическом режиме;
  - отображение трассы последовательности функций, приводящих к аварийному завершению в статическом анализаторе Svace.

- Совместная работа с дизассемблером IDA PRO:
  - Сохранение покрытия для плагина Lighthouse, которое отображает покрытые базовые блоки в ПО;
  - Вывод процента покрытых базовых блоков.
- Возможность проведения анализа серверного и клиентского ПО, работающего по протоколам с состояниями и без состояний;
- Лёгкая расширяемость и добавление новых методов в рамках существующей инфраструктуры; оперативная адаптация под новые задачи;
- Возможность использования для реализации обеспечительных мер ГОСТ Р 56939-2016 (при необходимости сертификации ПО для использования на территории России).

СИСТЕМНЫЕ ТРЕБОВАНИЯ

Поддержка ОС семейства Linux и Windows. ИСП Фаззер способен проводить фаззинг-тестирование встроенных устройств (контроллеры, устройства интернета вещей), а также сервисов и COM-объектов ОС Windows.

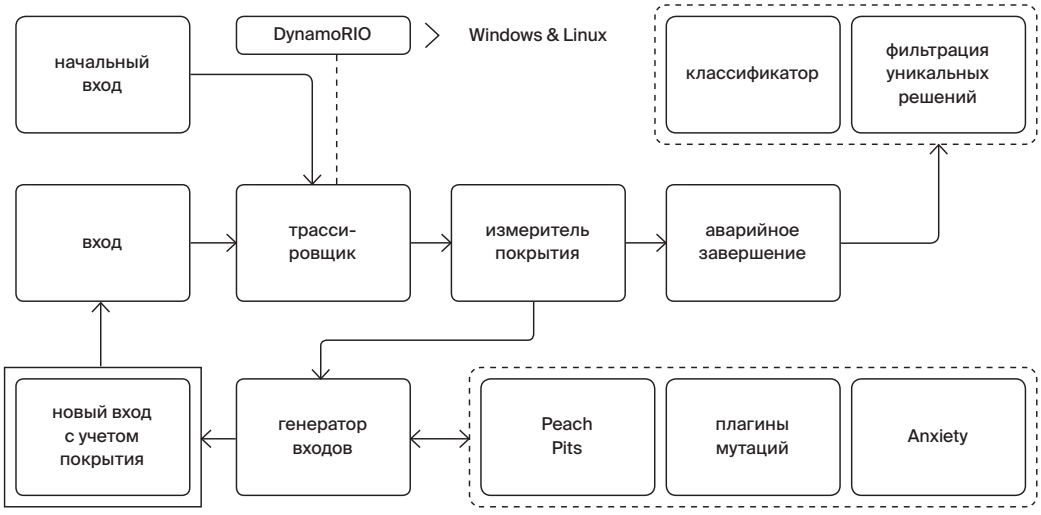
ДЛЯ КОГО ПРЕДНАЗНАЧЕН ИСП ФАЗЗЕР?

Компании, нацеленные на разработку ПО с высокой степенью надёжности и безопасности.

ОПЫТ ВНЕДРЕНИЯ

ИСП Фаззер внедрён в компаниях АО «НПО РусБИТех», «Код безопасности» и МВП «Свемел», а также задействован в осуществлении проекта в области анализа и кибербезопасности программ, который реализуется ИСП РАН совместно с Институтом индустриально-технологических исследований Тайваня (ITRI).

СХЕМА РАБОТЫ



КОМПЛЕКС РЕШЕНИЙ ДЛЯ СОЗДАНИЯ СЕРВИС-ОРИЕНТИРОВАННЫХ ЦОД

Комплекс предоставляет возможность хранения данных и совершения 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 (решение задач механики сплошных сред, анализ больших данных, анализ программ на уязвимости 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);

В НАСТОЯЩЕЕ ВРЕМЯ КОМПЛЕКС ПРЕДСТАВЛЕН ЧЕТЫРЬМЯ РЕШЕНИЯМИ:

I. ОБЛАЧНАЯ СРЕДА ASPERITAS НА БАЗЕ OPENSTACK, KUBERNETES И CEPH.

Создана на основе совместного проекта с компанией Dell. Предназначена для кратковременных вычислений с большими доступными ресурсами. Подход к развёртыванию облачной среды из локальных источников реализован в виде заранее подготовленной виртуальной машины, обладающей всеми необходимыми инструментами для запуска процесса развёртывания. Среда Asperitas включена в Реестр российского ПО (№5921).

- Развёрнута на базе открытых современных технологий, которые являются основными для построения больших приватных облачных систем;
- Предоставляет пользователям весь необходимый функционал:
  - управление виртуальными сетями и вычислительными кластерами с использованием систем Keystone, Neutron, Nova (аналог Amazon EC2);
  - блочное хранение данных, а также расширяемое объектное хранилище на основе распределённой файловой системы Ceph;
  - управление контейнерными окружениями на базе Kubernetes.

II. УНИВЕРСАЛЬНЫЙ ОРКЕСТРАТОР.

Инструмент управления жизненным циклом программных систем. Предоставляет возможность разработки и внедрения различных сервисов уровня PaaS:

- для анализа больших данных с полностью настроенными системами Apache Spark, Apache Hadoop и Apache Ignite, а также с 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.

ВОЗМОЖНОСТИ:

- Работает как сервис с системой пользователей, групп и REST API;
- Регистрирует историю действий и состояний программной системы;
- Работает с облачными системами виртуализации;
- Способен развёртывать сложные распределённые системы со всеми возможными комбинациями сервисов по запросу (Spark, Hadoop, Ignite, Cassandra, Jupyter, shared remote FS, remote FS server, Nextcloud);
- Способен учитывать совместимость сервисов и версий;
- Способен использовать локальные источники при развёртывании сервисов.

III. РЕШЕНИЕ ДЛЯ УПРАВЛЕНИЯ ВИРТУАЛЬНЫМИ МАШИНАМИ VMEMPEROR

Разработан в ИСП РАН для решения внутренних задач, находится в открытом доступе (https://github.com/ispras/vmemperor). 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.

IV. ПЛАТФОРМА ДЛЯ ОРГАНИЗАЦИИ WEB-ЛАБОРАТОРИЙ FANLIGHT

Создана в результате участия ИСП РАН в программе «Университетский кластер» и в международном проекте Open Cirrus (учреждён HP, Intel и Yahoo!). Предназначена для развёртывания SaaS-инфраструктур для вычислительных web-лабораторий средствами Docker Compose. Построена на контейнерных технологиях и предоставляет виртуальные рабочие места 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. Включена в Реестр российского ПО (№6066).

- Демонстрирует высокую эффективность работы с облачными вычислениями благодаря использованию контейнеров:
  - комфортная работа с тяжёлыми 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.

ОПЫТ ВНЕДРЕНИЯ

Вычислительный кластер на базе Asperitas используется для анализа информационных потоков в технологии анализа социальных медиа Talisman и для работы других технологий ИСП РАН (в частности, для анализа ОС Android с помощью Svace). Реализован совместный проект с компанией Huawei (анализ больших графов с помощью технологий обработки больших данных), а также инфраструктура поддержки жизненного цикла ОС Tizen, позволяющая организовать процесс совместной разработки компонентов ОС и автоматизировать регулярную сборку и тестирование образов. Кроме того, осуществляется ряд работ при участии Минобрнауки РФ.

Возможности платформы Fanlight использовались в ряде совместных проектов по развёртыванию web-лабораторий с ФГУП «РФЯЦ-ВНИИЭФ», ООО «РРС-Балтика», ИПМ им. М.В. Келдыша РАН (разработка технологий для увеличения и эффективного использования ресурсного потенциала углеводородного сырья Союзного государства), а также с Лабораторией механики сплошных сред ИСП РАН (https://unicfd.ru).

VMEmperor не применялся во внешних коммерческих проектах, однако используется во внутренних проектах ИСП РАН.