

Акселерационная программа «Стартапы LETI» 2023



# Архитектурно-ориентированные методы оптимизации машинного кода



Трекер проекта: Машурова Вера Александровна

Эксперт проекта: Кулибанова Валерия Вадимовна

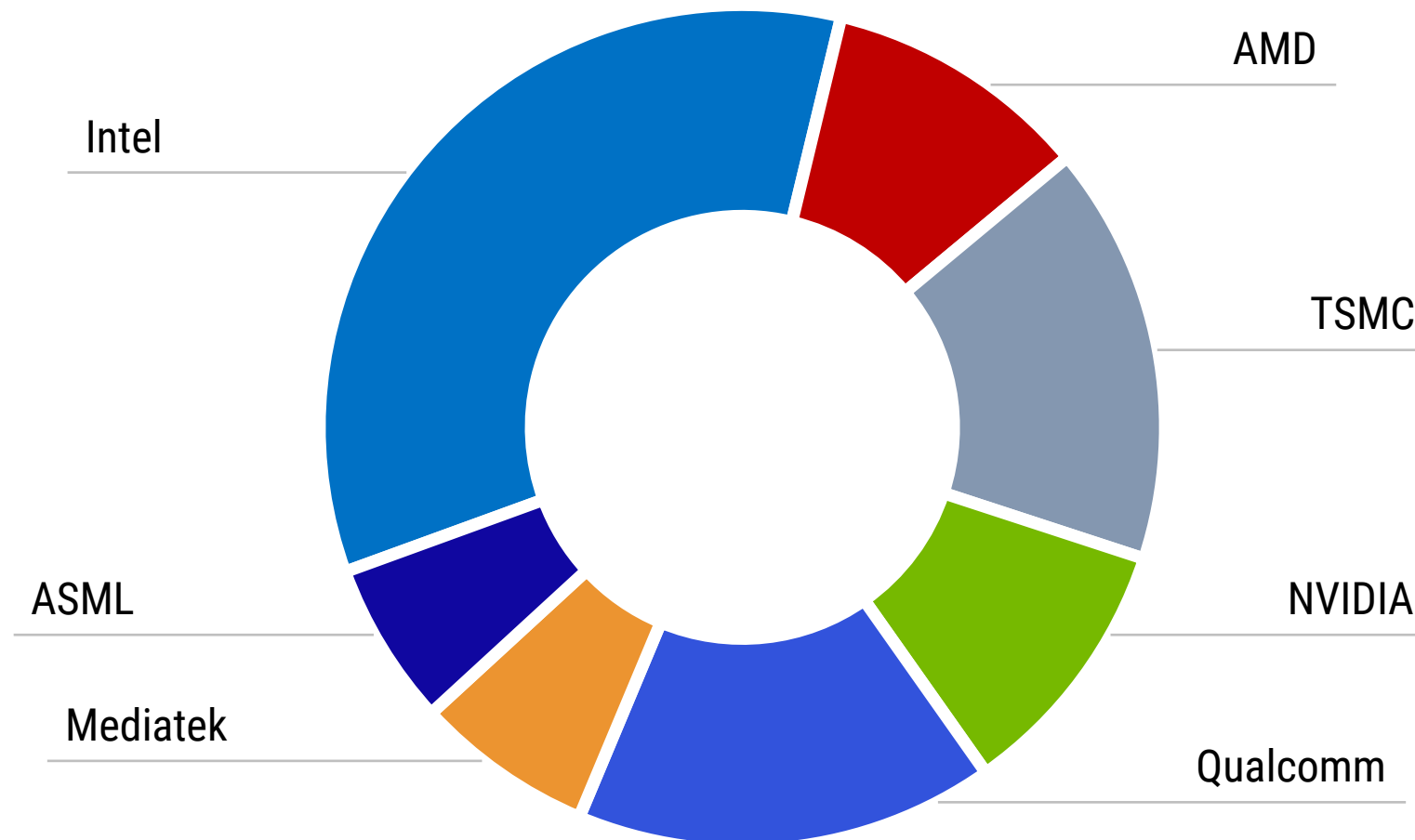
# Где есть вычисления?



# Объёмы инвестиций

Более **\$50 млрд** потратили ведущие мировые компании на R&D в 2022 году.

Почти все эти средства используются для разработки новых **процессоров, видеокарт и систем на кристалле.**





# Машина для EUV-литографии



- **Передовая** технология производства микрочипов
- **20 лет разработки** специалистами со всего мира
- **\$20 млрд** использовано на R&D
- **\$170 млн** — стоимость одного экземпляра
- Производит только **одна компания**
- Тем не менее, **большой спрос**

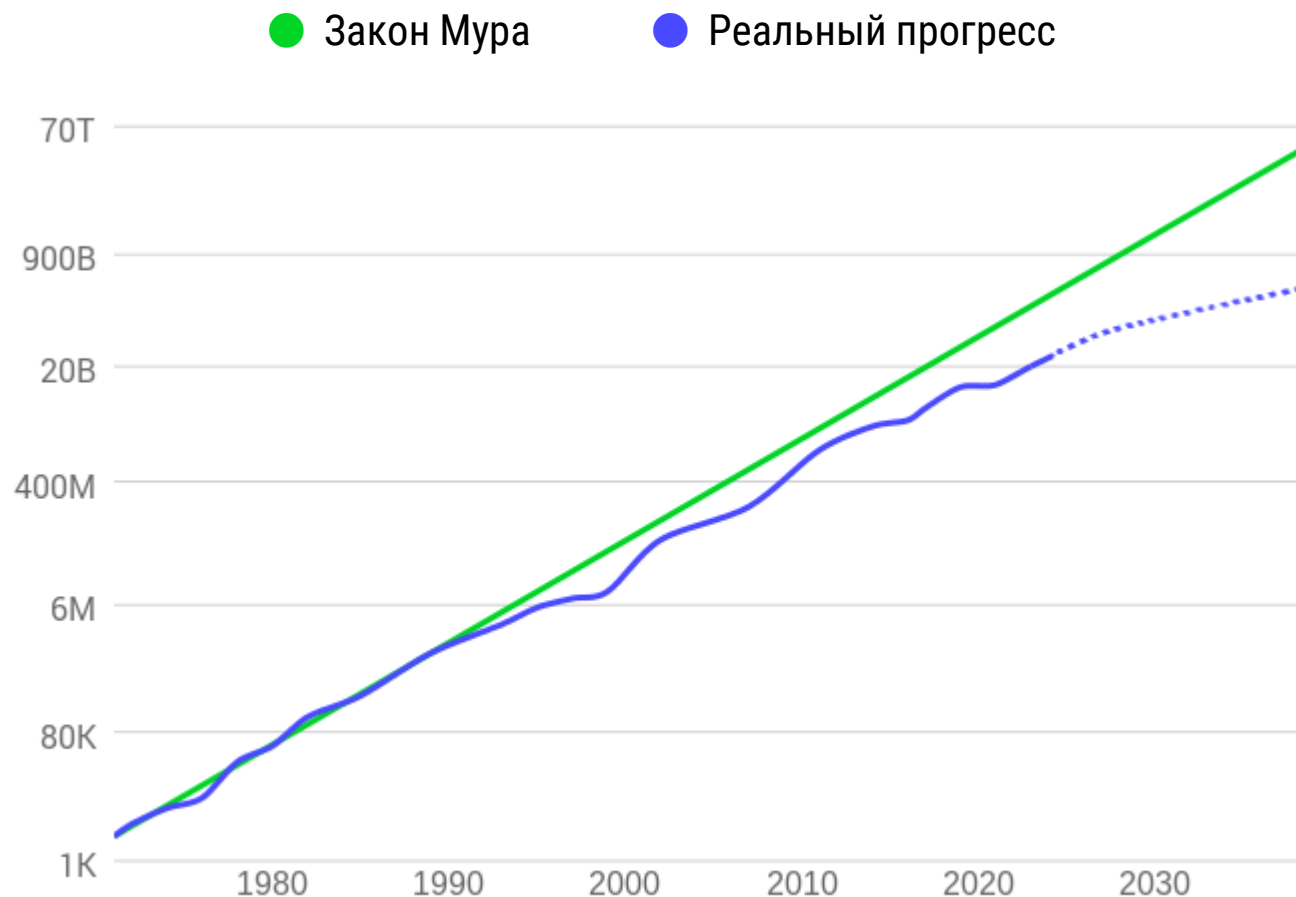


# Закон Мура

Согласно **закону Мура** число транзисторов на кристалле **удваивается каждые два года**.

Поддерживать его становится всё сложнее из-за **физических ограничений**.

**Оптимизация программ** становится **решающим фактором** в гонке за производительностью.



# Что мы предлагаем?



Мы разрабатываем продукт, который позволит людям и организациям по всему миру сделать свои программы **быстрее** и **энергоэффективнее**.

Наша разработка будет востребована практически во всех сегментах, связанных с вычислениями. В первую очередь это обработка больших данных, телекоммуникационное оборудование и встраиваемые системы.

Актуальность проекта обусловлена **растущими потребностями** во всё более и более быстрых вычислениях.

## Почему это нужно?



Для обеспечения работы высоконагруженных сервисов необходима **дорогостоящая инфраструктура**. Чем больше вычислений, тем больше нужно оборудования – то есть материальных вложений.

Оптимизации программного обеспечения позволяют **существенно снизить финансовые издержки**.

Пример из практики: шесть лет назад базу данных «ВКонтакте» переписали три чемпиона мира по программированию с использованием более эффективных структур данных. Это позволило сэкономить **\$5 млн + \$750 тысяч в год** за счёт операционных расходов.

## Как это работает?



Мы предлагаем **решение** для **автоматической** оптимизации программ. Оно не требует каких-либо вмешательств со стороны разработчика: достаточно просто запустить программу в специальной среде.

**Оптимизация** заключается в **перестройке машинного кода** программы **с учётом особенностей** аппаратного обеспечения, обрабатываемых данных и окружения, в котором запущена программа.

В ходе экспериментов вы обнаружили, что такой подход позволяет существенно ускорить вычисления — вплоть до **трёхкратного ускорения**.



Мы считаем, что ПО подобного рода должно быть открытым, доступным и иметь свободную лицензию — это позволит сделать его максимально **полезным** и **эффективным**.

Такой идеологии уже следуют многие программные пакеты — и являются **лидирующими** в своём классе.

Открытость позволяет развивать продукт силами специалистов **со всего мира**, а не локальной группы людей.

**Непосредственных** конкурентов у проекта нет – все продукты подобного класса можно использовать **параллельно** с предлагаемым решением, так как они оптимизируют другим способом.

	<b>Цена</b>	<b>Что делает</b>	<b>Как делает</b>	<b>Потенциал ускорения</b>
<b>Наш проект</b>	Бесплатно	Оптимизация	Динамически	3x
<b>BOLT</b>	Бесплатно	Оптимизация	Статически	1,5x
<b>PGO в компиляторах</b>	Бесплатно	Оптимизация	Статически	1,3x
<b>QEMU</b>	Бесплатно	Эмуляция	–	–
<b>Unicorn</b>	Бесплатно	Эмуляция	–	–
<b>Dynarmic</b>	Бесплатно	Трансляция	–	–

# Стратегии продвижения

- Форумы, митапы, конференции
- Статьи в журналах
- Интеграция в существующие инфраструктуры оптимизации
- Встречи с представителями компаний



## S

- Отсутствие прямых конкурентов
- Большой потенциал
- Открытость и доступность
- Может частично компенсировать технологическое отставание

## W

- Необходимость запуска программы для оптимизации – усложняет работу с микроконтроллерами

## O

- Интеграция в компиляторы и тулчейны
- Кросс-архитектурная трансляция
- Задействование видеокарт и FPGA
- Разработка новых процессоров

## T

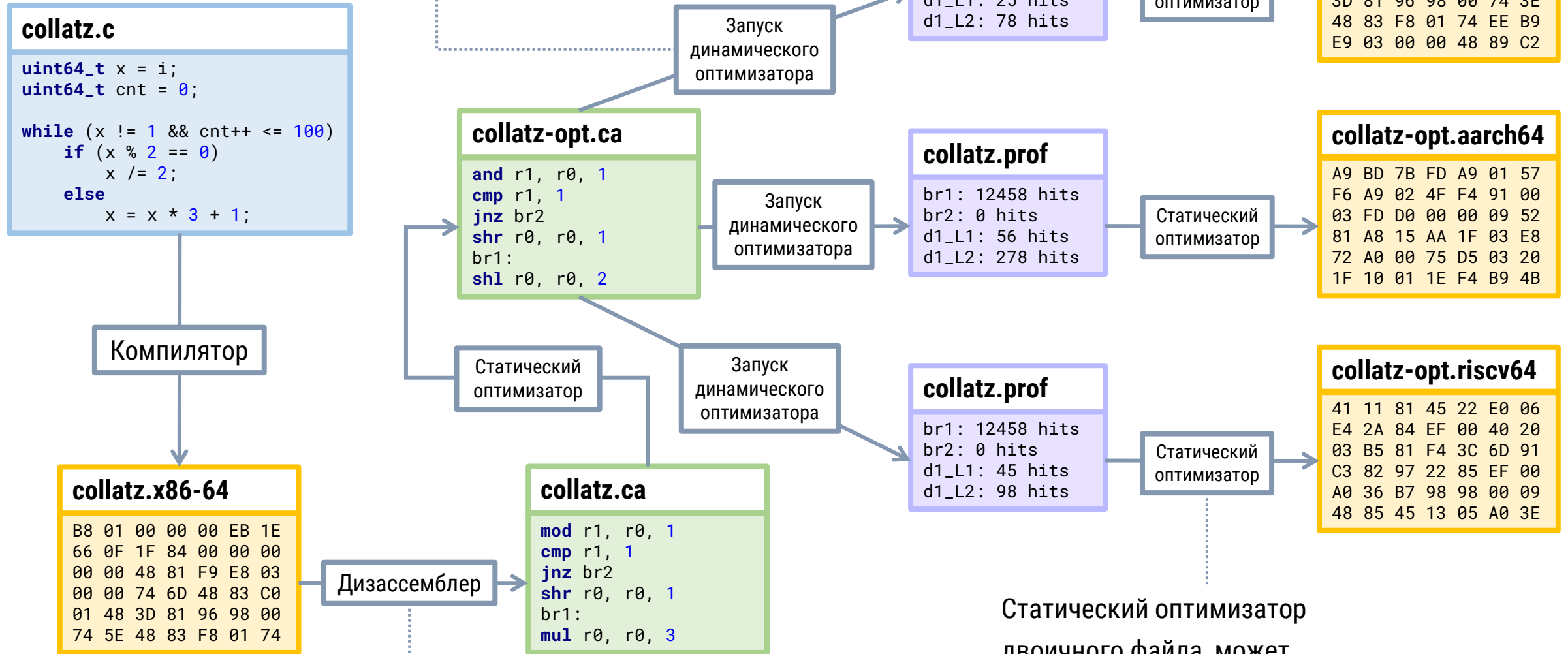
- Не факт, что некоторые из идей удастся реализовать
- Инструмент может остаться неизвестным

## Результаты исследований и разработок можно развивать в нескольких направлениях:

- Интеграция в существующие компиляторы и тулчейны
- Улучшение средств эмуляции архитектур
- Задействование более эффективных вычислителей для решения задач, таких как видеокарты, нейронные ядра и FPGA
- Разработка более эффективных архитектур процессоров, рассчитанных на предлагаемый способ оптимизации
- Создание собственной инфраструктуры для оптимизации и трансляции кода

# Инфраструктура

Динамический оптимизатор,  
собирающий информацию о  
работе программы



Конвертер в промежуточное представление

Статический оптимизатор  
двоичного файла, может  
использовать собранный профиль



# Статус проекта



В **настоящее время** – есть научно-исследовательский задел, разрабатывается MVP.

Разработка парсера машинного кода для x86-64

Создание тестовых программ и подготовка SPEC-тестов

Разработка генератора машинного кода для x86-64

Реализация простейших оптимизирующих проходов

1-й квартал

2-й квартал

3-й квартал

4-й квартал

# Финансовые потребности и обеспечение



Желательно иметь **парк машин** для тестирования.

Основные расходы – заработная плата исследователей и разработчиков.

**Необходимые финансы** для получения MVP в течение года с учётом квалификации участников и налоговых взносов:

$$(200\ 000 \text{ рублей} \times 1 \text{ человек} + 100\ 000 \text{ рублей} \times 5 \text{ человек}) \times 12 \text{ месяцев} \times (1 + 0,13 + 0,3) \approx \mathbf{12\ 000\ 000 \text{ рублей}}$$

# Команда проекта



<b>ФИО</b>	<b>Должность</b>	<b>Функция в проекте</b>	<b>Образование и опыт</b>
Когутенко Андрей Александрович	Магистрант	Тимлид	НИР с Huawei
Пазников Алексей Александрович	Доцент	Консультант	НИР с Huawei и РФФ, статьи, конференции
Осипов Ярослав Сергеевич	Магистрант	Аналитик	Низкоуровневая оптимизация
Репин Степан Александрович	Магистрант	Исследователь	НИР с Huawei
Манжиков Леонид Павлович	Аспирант	Исследователь	НИР с Huawei
Кадыров Тимур Валерьевич	Аспирант	Исследователь	НИР с Huawei



**Спасибо за внимание!**

[and.kogutenko@yandex.ru](mailto:and.kogutenko@yandex.ru)

[apaznikov@gmail.com](mailto:apaznikov@gmail.com)