Яндекс Такси

Незаменимый С++

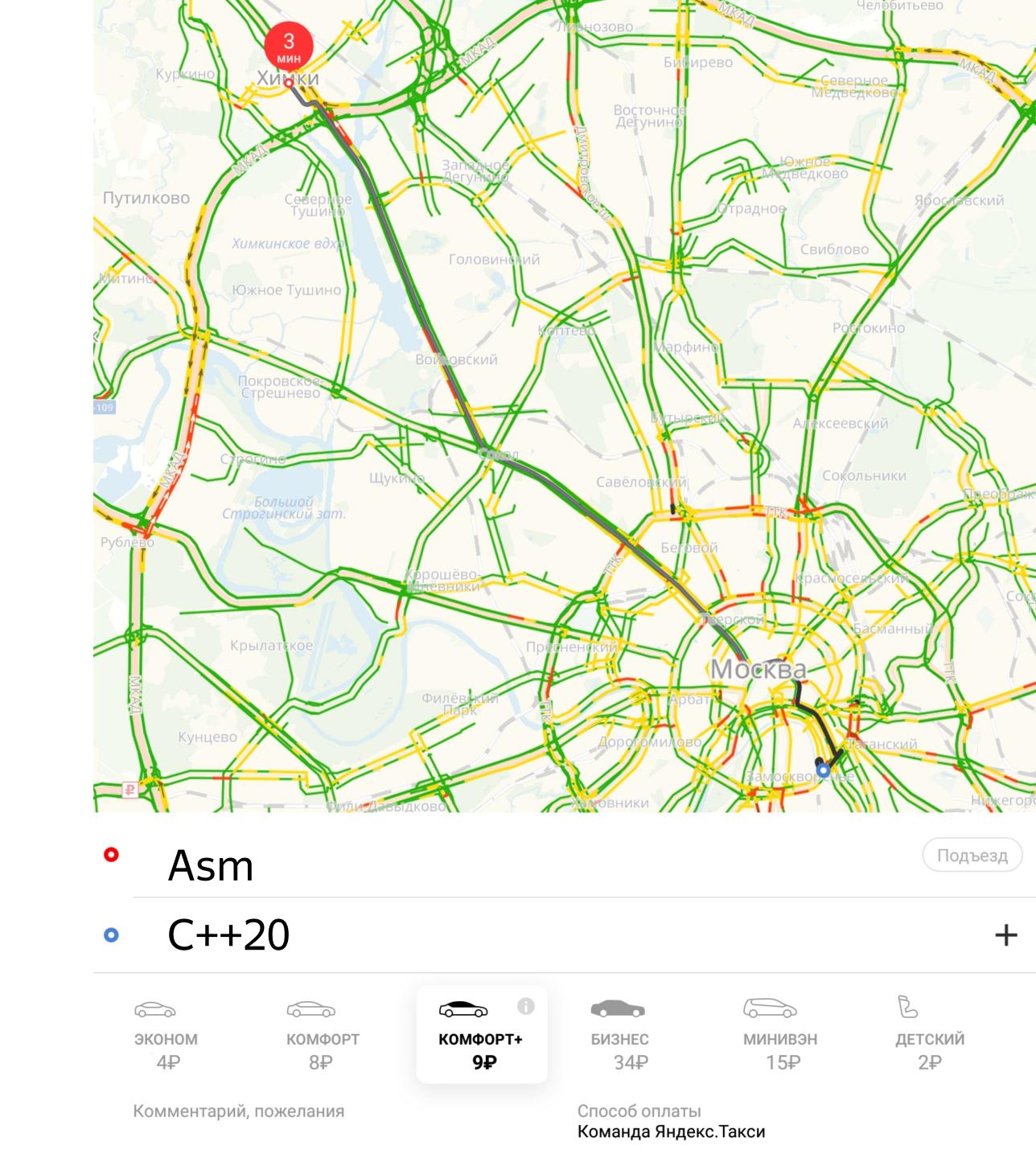
Полухин Антон

Antony Polukhin

Яндекс Такси

Содержание

- На С++ больше не пишут?
- C++ vs. Asm
- C++ vs. Rust
- C++ vs. Go
- C++ vs. (Java + C#)
- Слабые места С++
- Что с этим делать



Заблуждение N°1

На С++ больше не пишут программ

Hезаменимый C++

• Поисковые движки

- Поисковые движки
- Большинство высоконагруженных програм (Такси)

Незаменимый С++ 7 / 135

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры

Hезаменимый C++

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация

Незаменимый C++ 10 / 135

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)
- «Виртуальные машины»

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)
- «Виртуальные машины»
- Научные программы (CERN и Бозон Хиггса)

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)
- «Виртуальные машины»
- Научные программы (CERN и Бозон Хиггса)
- Части ОС (Драйверы, userspace)

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)
- «Виртуальные машины»
- Научные программы (CERN и Бозон Хиггса)
- Части ОС (Драйверы, userspace)
- Автопром

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)
- «Виртуальные машины»
- Научные программы (CERN и Бозон Хиггса)
- Части ОС (Драйверы, userspace)
- Автопром
- Заводы

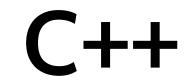
- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)
- «Виртуальные машины»
- Научные программы (CERN и Бозон Хиггса)
- Части ОС (Драйверы, userspace)
- Автопром
- Заводы
- Биржа

Незаменимый C++ 17 / 135

- Поисковые движки
- Большинство высоконагруженных програм (Такси)
- Игры
- Браузеры
- Спецэффекты и анимация
- Компиляторы (не только компиляторы для С++)
- «Виртуальные машины»
- Научные программы (CERN и Бозон Хиггса)
- Части ОС (Драйверы, userspace)
- Автопром
- Заводы
- Биржа
- Офисные приложения

Незаменимый C++ 18 / 135

ОК, чем C++ привлекает сегодня?



+ zero-overhead

- + zero-overhead
- + неограниченные возможности

- + zero-overhead
- + неограниченные возможности
- + поддержка огромного количества платформ

- + zero-overhead
- + неограниченные возможности
- + поддержка огромного количества платформ
- + безопасность

- + zero-overhead
- + неограниченные возможности
- + поддержка огромного количества платформ
- + безопасность?

- + zero-overhead
- + неограниченные возможности
- + поддержка огромного количества платформ
- + безопасность?
- + небольшой рантайм

Незаменимый C++ 26 / 135

- + zero-overhead
- + неограниченные возможности
- + поддержка огромного количества платформ
- + безопасность?
- + небольшой рантайм
- + стабильность

Незаменимый C++ 27 / 135

Сравним с другими языками!

+ Ассемблер позволяет выжать максимум из железа

+ Ассемблер позволяет выжать максимум из железа

- непереносимый код

+ Ассемблер позволяет выжать максимум из железа

- непереносимый код
- очень медленная разработка

+ Ассемблер позволяет выжать максимум из железа

- непереносимый код
- очень медленная разработка
- не всегда быстрее

+ Ассемблер позволяет выжать максимум из железа

- непереносимый код
- очень медленная разработка
- не всегда быстрее

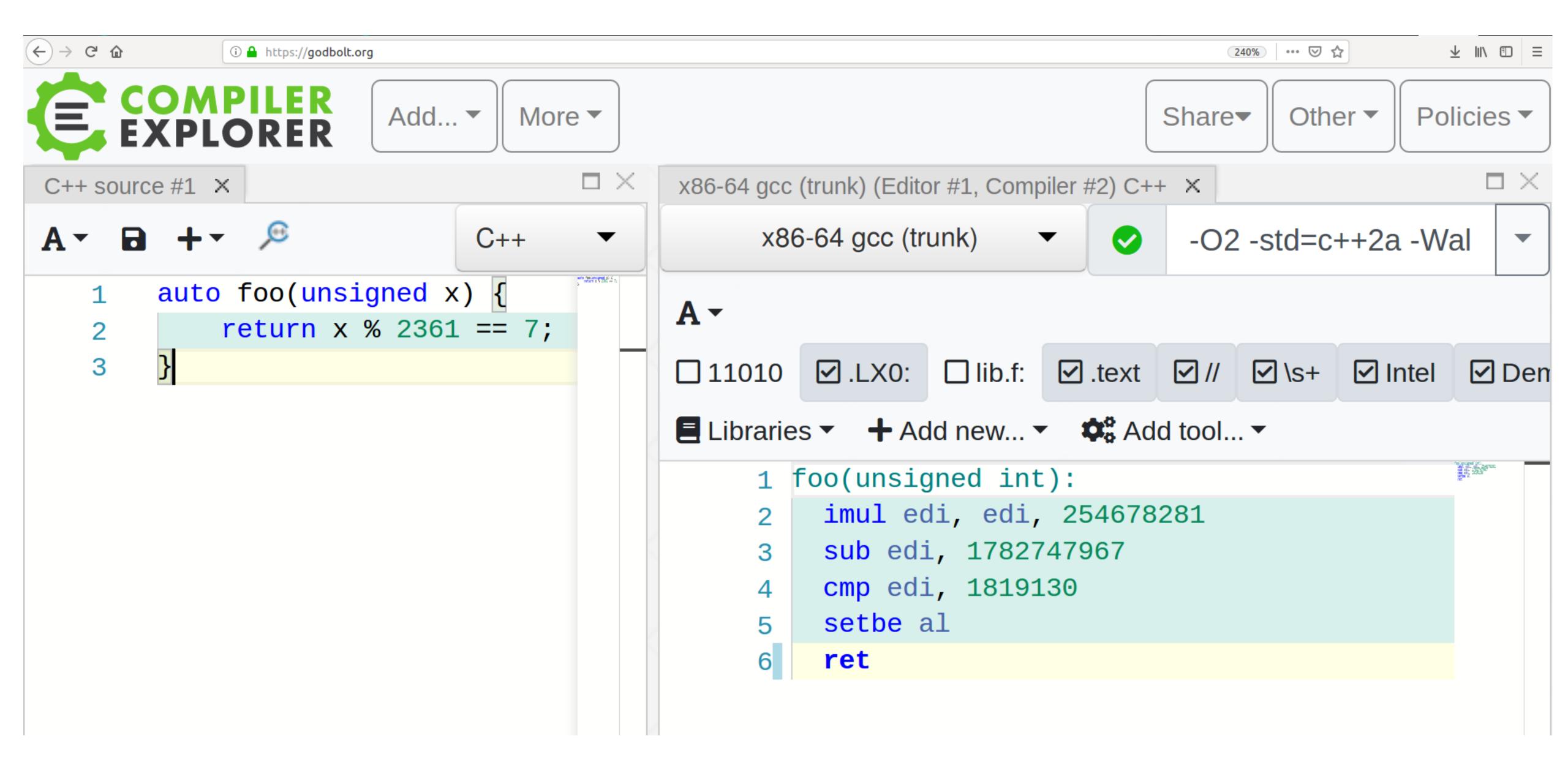
Заблуждение N°2

Большие программы на C++ работают медленнее чем на ASM

C++ vs. Asm (или как завалить Asm разработчика)

$$x \% 2361 == 7$$

C++ vs. Asm (или как завалить Asm разработчика)



C++ vs. Asm (или как завалить Asm разработчика)

```
TT1 OF
              unstanca).
              (X * C3) + C4 <= 2 * C4, where
  11751
              C3 is modular multiplicative inverse of (unsigned) C1 and 1<<pre>prec and
  11752
              C4 is ((1 << (prec - 1) - 1) / C1).
  11753
              If C1 is even, S = ctz(C1), use
  11754
              ((X * C3) + C4) r>> S <= (C4 >> (S - 1))
  11755
              where C3 is modular multiplicative inverse of (unsigned)(C1>>S) and 1<<pre>prec
  11756
              and C4 is ((1 << (prec - 1) - 1) / (C1 >> S)) & (-1 << S).
  11757
  11758
              See the Hacker's Delight book, section 10-17. */
· · · L1759
  11760
          enum tree_code
          maybe_optimize_mod_cmp (enum tree_code code, tree *arg0, tree *arg1)
  11761
  11762
```

Hезаменимый C++

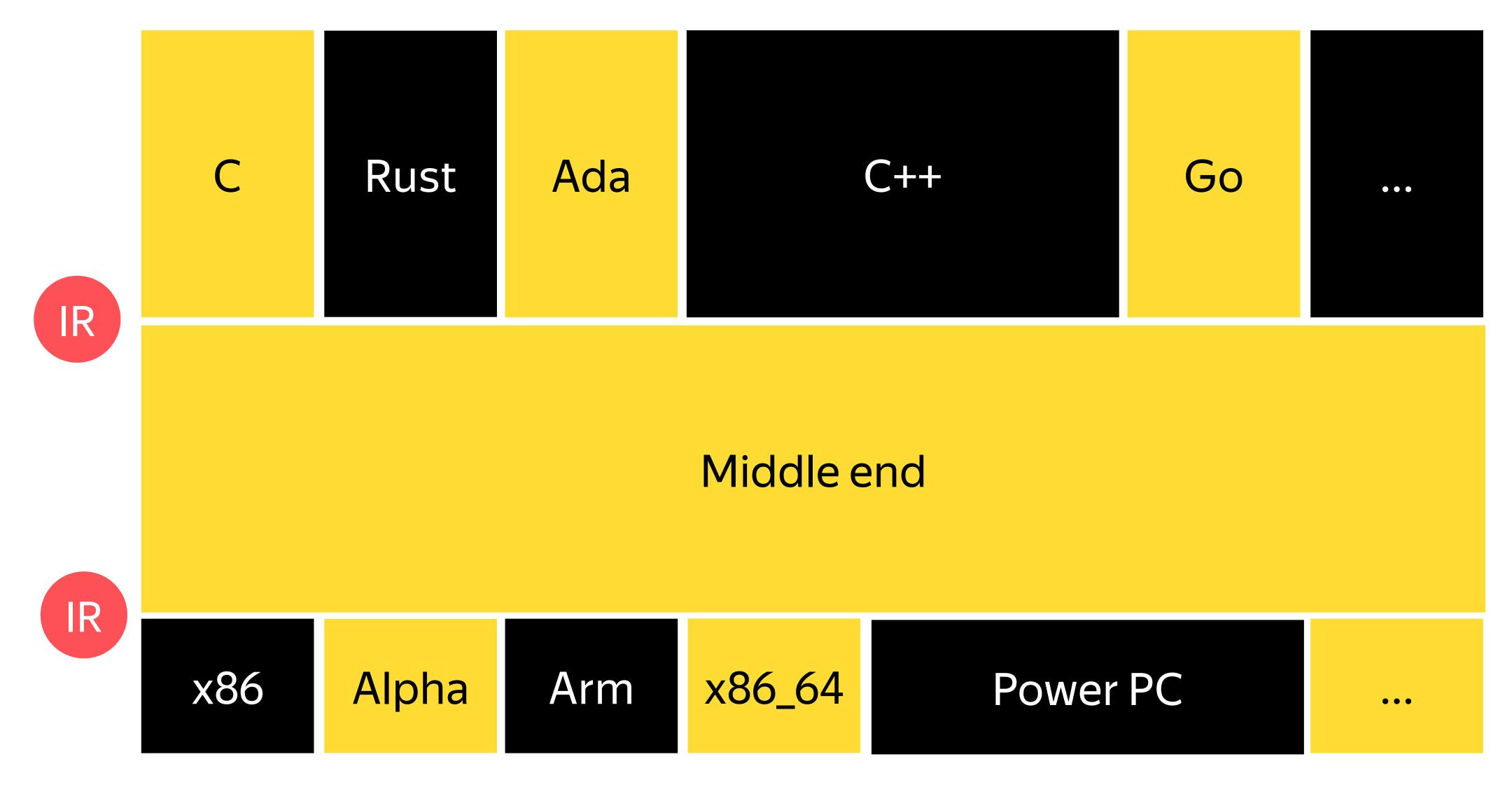
+ Отличная безопасность?

- + Отличная безопасность?
- + Потенциальная возможность оптимизировать лучше C++

Заблуждение N°3

Rust более/такой же производительный как С++

Анатомия компилятора (упрощённо)



```
\square \times
Rust source #2 × rustc nightly (Editor #2, Compiler #2) Rust ×
                                                                         C++ source #1 × x86-64 clang (trunk) (Editor #1, Compiler #1) C++ ×
                                                                              Rust
                                                                                                                                     C++
    ■ Save/Load + Add new... ▼
                                                                                                                                              \blacksquare
                                                                                  #include <cstdint>
                                                                                  auto square(std::int32_t num) {
        pub fn square(num: i32) -> i32 {
                                                                             3
             return num * num
                                                                                      return num * num;
                                                                             4
                                                                             5
    5
                                                                                                                                             \square \times
Diff rustc nightly vs x86-64 clang (trunk) ×
                                              Right:
     Left:
                                                      x86-64 clang (trunk) -O2 -s... ▼
            rustc nightly -O
A -
   1 - example::square:
                                                                           1 + square(int): # @square(int)
         mov eax, edi
                                                                                 mov eax, edi
         imul eax, edi
                                                                                 imul eax, edi
         ret
                                                                                 ret
```

Где UB?

```
\square \times
Rust source #2 × rustc nightly (Editor #2, Compiler #2) Rust ×
                                                                        C++ source #1 × x86-64 clang (trunk) (Editor #1, Compiler #1) C++ ×
                                                                        C++
                                                          Rust
    Save/Load + Add new... ▼
                                                                                                                                            \blacksquare
                                                                                 #include <cstdint>
        pub fn square(num: i32) -> i32 {
                                                                                 auto square(std::int32_t num) {
                                                                            3
             return num * num
                                                                                     return num * num;
                                                                            4
                                                                            5
    5
                                                                                                                                           \square \times
Diff rustc nightly vs x86-64 clang (trunk) ×
     Left:
                                              Right:
                                                      x86-64 clang (trunk) -O2 -s... ▼
            rustc nightly -O
A٠
                                                                          1 + square(int): # @square(int)
   1 - example::square:
         mov eax, edi
                                                                               mov eax, edi
         imul eax, edi
                                                                               imul eax, edi
         ret
                                                                               ret
```

Hезаменимый C++ 47 / 135

```
\square \times
Rust source #2 × rustc nightly (Editor #2, Compiler #2) Rust ×
                                                                     C++ source #1 × x86-64 clang (trunk) (Editor #1, Compiler #1) C++ ×
                                                                          Rust
                                                                                                                               C++
    Save/Load + Add new... ▼
                                                                                                                                       ▾
                                                                             #include <cstdint>
        pub fn foo(max: i32, num: i32) -> i32 {
                                                                             auto foo(std::int32_t max, std::int32_t num) {
            return max * num
                                                                                 return max * num;
                                                                         5
        pub fn bar(max: i32, num: i32) -> i32 {
                                                                             std::int32_t bar(std::int32_t max, std::int32_t num
            return bar(max, num) * bar(max, num)
                                                                                 return bar(max, num) * bar(max, num);
                                                                         9
   9
                                                                                                                                      \square \times
Diff rustc nightly vs x86-64 clang (trunk) ×
    Left:
                                            Right:
           rustc nightly -O
                                                    x86-64 clang (trunk) -O2 -s... ▼
   1 - example::foo:
                                                                       1 + foo(int, int): # @foo(int, int)
        mov eax, edi
                                                                            mov eax, edi
        imul eax, esi
                                                                            imul eax, esi
                                                                            ret
        ret
                                                                       5 + bar(int, int): # @bar(int, int)
   6 - example::bar:
                                                                            ret
        ret
```

```
\square \times
                                                                    C++ source #1 × x86-64 clang (trunk) (Editor #1, Compiler #2) C++ ×
Rust source #2 × rustc nightly (Editor #2, Compiler #1) Rust ×
                                                                        C++
                                                      Rust
   Save/Load + Add new... ▼
                                                              \blacksquare
                                                                                                                                 ▾
                                                                           #include <cstdint>
                                                                           auto foo(std::int32_t max, std::int32_t num) {
       pub fn foo(max: i32, num: i32) -> i32 {
           return max * num
                                                                               return max * num;
   5
       pub fn bar(max: i32, num: i32) -> i32 {
                                                                           std::int32_t bar(std::int32_t max, std::int32_t nu
           return foo(max, num) * foo(max, num)
                                                                               return foo(max, num) * foo(max, num);
   9
                                                                       9
                                                                                                                                \square \times
Diff rustc nightly vs x86-64 clang (trunk) ×
                                          Right:
    Left:
          rustc nightly -O -C overflo... ▼
                                                 x86-64 clang (trunk) -O2 -s... ▼
                                                                   8 + bar(int, int): # @bar(int, int)
 12 -
 13 - example::bar:
       push rax
 14 -
       mov eax, edi
                                                                         mov eax, edi
 15
       imul eax, esi
                                                                         imul eax, esi
                                                                  10
 16
       jo .LBB1_3
                                                                         jo .LBB1_3
                                                                  11
        imul eax, eax
                                                                         imul eax, eax
       jo .LBB1_2
                                                                  13 + jo .LBB1_3
                                                                                  20 -
       pop rcx
```

```
\square \times
                                                                                                                                      \square \times
                                                                       C++ source #1 × x86-64 clang (trunk) (Editor #1, Compiler #2) C++ ×
Rust source #2 × rustc nightly (Editor #2, Compiler #1) Rust ×
                                                                           C++
   ■ Save/Load + Add new... ▼
                                                         Rust
                                                                                                                                        ▾
                                                                               include <array>
       pub fn mutate(num: i32) -> [i32; 6] {
                                                                               uto mutate(std::int32_t num) {
                                                                                  std::array < std::int32_t, 6 > numbers = {1, num,}
            let mut numbers = [1i32, num, 3, 4, 5, 6];
            for (j, elem) in numbers.iter_mut().enumerate()
                                                                                  for (std::size_t j = 0; auto& elem : numbers) {
                                                                          5
                *elem += j as i32
                                                                                      elem += j;
                                                                          6
                                                                                      ++j;
            return numbers;
   8
                                                                                  return numbers;
   9
                                                                          9
                                                                         10
                                                                                                                                      \square \times
Diff rustc nightly vs x86-64 clang (trunk) ×
    Left:
                                            Right:
                                                    x86-64 clang (trunk) -O2 -s... ▼
           rustc nightly -O
        mov dword ptr [rdi], ecx
        mov ecx, dword ptr [rsp + 4]
 14 -
        mov dword ptr [rdi + 4], ecx
 15 -
        mov rcx, qword ptr [rsp + 8]
 16 -
        mov qword ptr [rdi + 8], rcx
 17 -
 18 - mov rcx, qword ptr [rsp + 16]
        mov qword ptr [rdi + 16], rcx
        add rsp, 24
  20 -
```

```
\square \times
Rust source #2 × rustc nightly (Editor #2, Compiler #1) Rust ×
                                                                     C++ source #1 × x86-64 clang (trunk) (Editor #1, Compiler #2) C++ ×
                                                                         Rust
                                                                                                                            C++
    ■ Save/Load + Add new... ▼
                                                                                                                                    \blacksquare
                                                                            #include <cstdint>
       pub struct Stats { x: u32, y: u32, z: u32, }
                                                                            struct Stats { std::uint32_t x, y, z; };
       pub fn sum(a: &Stats, b: &Stats) -> Stats {
                                                                            auto sum(const Stats& a, const Stats& b) {
            return Stats \{x: a.x + b.x, y: a.y + b.y, z: a
                                                                                 return Stats \{a.x + b.x, a.y + b.y, a.z + b.z\}
Diff rustc nightly vs x86-64 clang (trunk) ×
           rustc nightly -O
                                                   x86-64 clang (trunk) -O2 -s... ▼
                                           Right:
A -
     Left:
   1 - ZN7example3sum17h2331b594b0172e4aE:
                                                                     1 + sum(Stats const&, Stats const&): # @sum(Stats const&, St
        mov ecx, dword ptr [rdx]
                                                                          mov eax, dword ptr [rsi]
        mov r8d, dword ptr [rdx + 4]
                                                                          add eax, dword ptr [rdi]
                                                                          mov ecx, dword ptr [rsi + 4]
        add ecx, dword ptr [rsi]
        add r8d, dword ptr [rsi + 4]
                                                                          add ecx, dword ptr [rdi + 4]
        mov edx, dword ptr [rdx + 8]
                                                                          mov edx, dword ptr [rsi + 8]
        add edx, dword ptr [rsi + 8]
                                                                          add edx, dword ptr [rdi + 8]
        mov rax, rdi
                                                                          shl rcx, 32
        mov dword ptr [rdi], ecx
                                                                          or rax, rcx
       mov dword ptr [rdi + 4], r8d
        mov dword ptr [rdi + 8], edx
        ret
                                                                    10
                                                                          ret
 12
```

Вроде норм.

Вроде норм.

Переходим на Rust?

Вроде норм.

Переходим на Rust?

Oh, wait!..

Вроде норм.

Переходим на Rust?

Oh, wait!..

$$C \rightarrow C++ - noop$$

C → Rust — PAIN!!!!!!

• unsafe {} → нет безопасности

- unsafe {} → нет безопасности
- UB: https://doc.rust-lang.org/reference/behavior-considered-undefined.html

- unsafe {} → нет безопасности
- UB: https://doc.rust-lang.org/reference/behavior-considered-undefined.html

It is the programmer's responsibility when writing unsafe code to ensure that any safe code interacting with the unsafe code cannot trigger these behaviors. unsafe code that satisfies this property for any safe client is called *sound*; if unsafe code can be misused by safe code to exhibit undefined behavior, it is *unsound*.

⚠ *Warning:* The following list is not exhaustive. There is no formal model of Rust's semantics for what is and is not allowed in unsafe code, so there may be more behavior considered unsafe. The following list is just what we know for sure is undefined behavior. Please read the

- unsafe {} → нет безопасности
- UB: https://doc.rust-lang.org/reference/behavior-considered-undefined.html
- Нет возможности использовать C headers
 - Надо генерировать свои
 - Обновление библиотек БОЛЬ!
 - Hago headers руками допатчивать
 - Мучительные страдания с borrow checker на сложных С проектах [https://hackernoon.com/why-im-dropping-rust-fd1c32986c88]

Заблуждение N°4

Программа написанная на языке Rust X не содержит ошибок

Anything

• unsafe или аналоги → нет безопасности

Anything

- unsafe или аналоги → нет безопасности
- если ваша программа компилится, это ещё не значит что всё ОК

- <великое множество>

- <*Bеликое множество*>

+ Асинхронность и многопточность на основе корутин

- <великое множество>

- + Асинхронность и многопточность на основе корутин
 - Boost.Fibers

- <*Bеликое множество*>

- + Асинхронность и многопточность на основе корутин
 - Boost.Fibers
 - Yandex. Taxi userver

- <*Bеликое множество*>

- + Асинхронность и многопточность на основе корутин
 - Boost.Fibers
 - Yandex. Taxi userver
 - Quantum

- <великое множество>

- + Асинхронность и многопточность на основе корутин
 - Boost.Fibers
 - Yandex. Taxi userver
 - Quantum
 - Folly fibers

- <*Bеликое множество*>

- + Асинхронность и многопточность на основе корутин
 - Boost.Fibers
 - Yandex. Taxi userver
 - Quantum
 - Folly fibers
 - Coroutines TS

- <великое множество>

- + Асинхронность и многопточность на основе корутин
 - Boost.Fibers
 - Yandex. Taxi userver
 - Quantum
 - Folly fibers
 - Coroutines TS
 - C++20

Python vs. Go

Go скорее конкурент Python, чем С++

Заблуждение N°5

Бенчмарки показывают что прогарммы на X быстрее С++

Добро пожаловать в мир «честных» бенчмарков!

• Отключается сборщик мусора

- Отключается сборщик мусора
- Код написан не на С++

Hезаменимый C++

- Отключается сборщик мусора
- Код на C++ написан в стиле float* f = new float;

Незаменимый C++ 78 / 135

- Отключается сборщик мусора
- Код на C++ написан в стиле float* f = new float;
- На X написана другая программа

- Отключается сборщик мусора
- Код на C++ написан в стиле float* f = new float;
- На X написана другая программа
- Тестируется библиотека а не язык

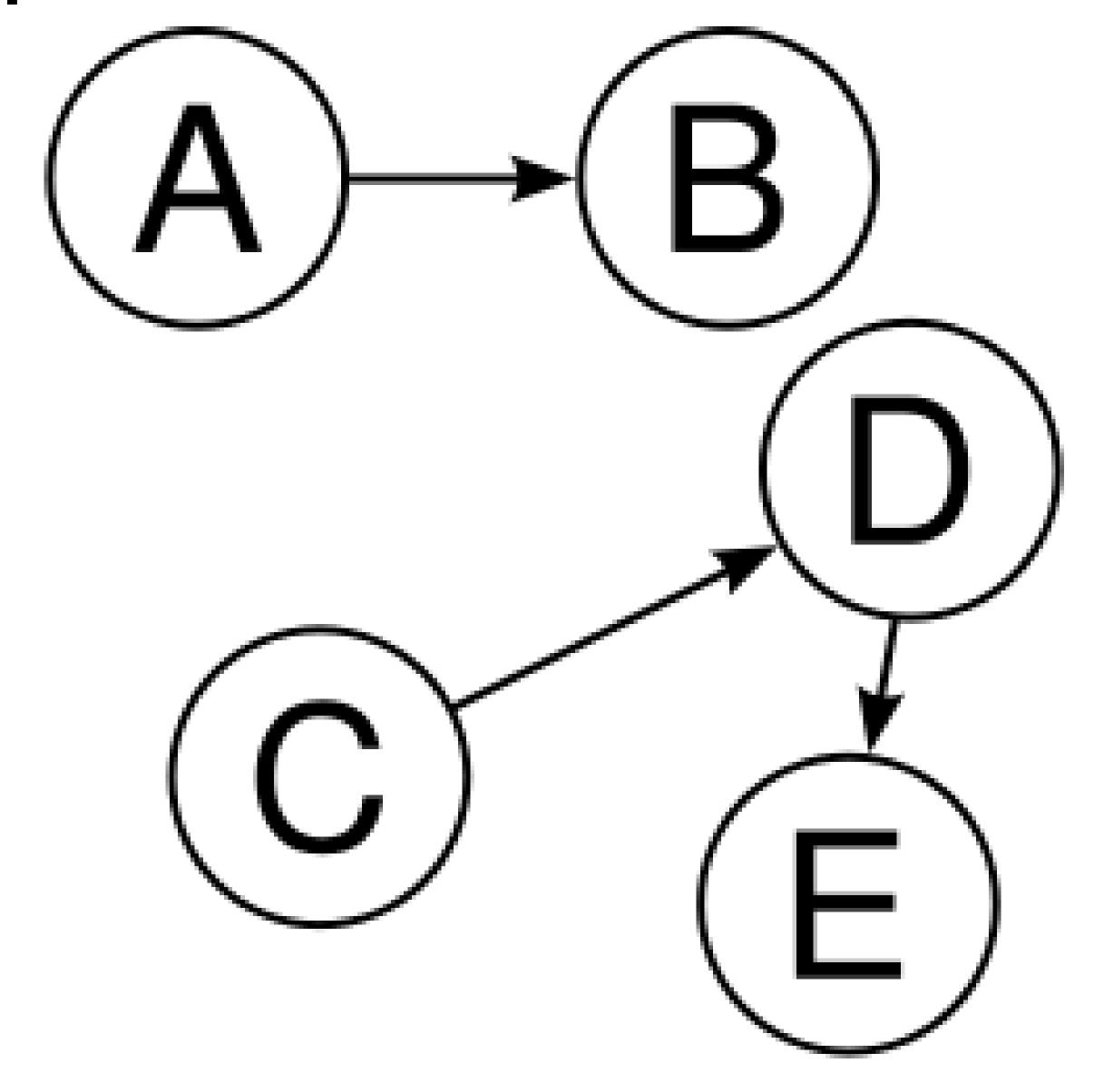
- Отключается сборщик мусора
- Код на C++ написан в стиле float* f = new float;
- На X написана другая программа
- Тестируется библиотека а не язык
- Сравниваются разные оптимизаторы, а не языки:

```
_mm_sub_pd(
    _mm_mul_pd(distance, _mm_set1_pd(1.5)),
    _mm_mul_pd(
    _mm_mul_pd(_mm_mul_pd(_mm_set1_pd(0.5), dsquared), distance),
    _mm_mul_pd(distance, distance),
    ),
)
```

Заблуждение N°6

Сборщик мусора не добавляет накладных расходов

Mark and sweep



```
struct list_node {
    list_node* next;
    list_node* prev;
};
struct slist_node {
    slist_node* next;
};
```

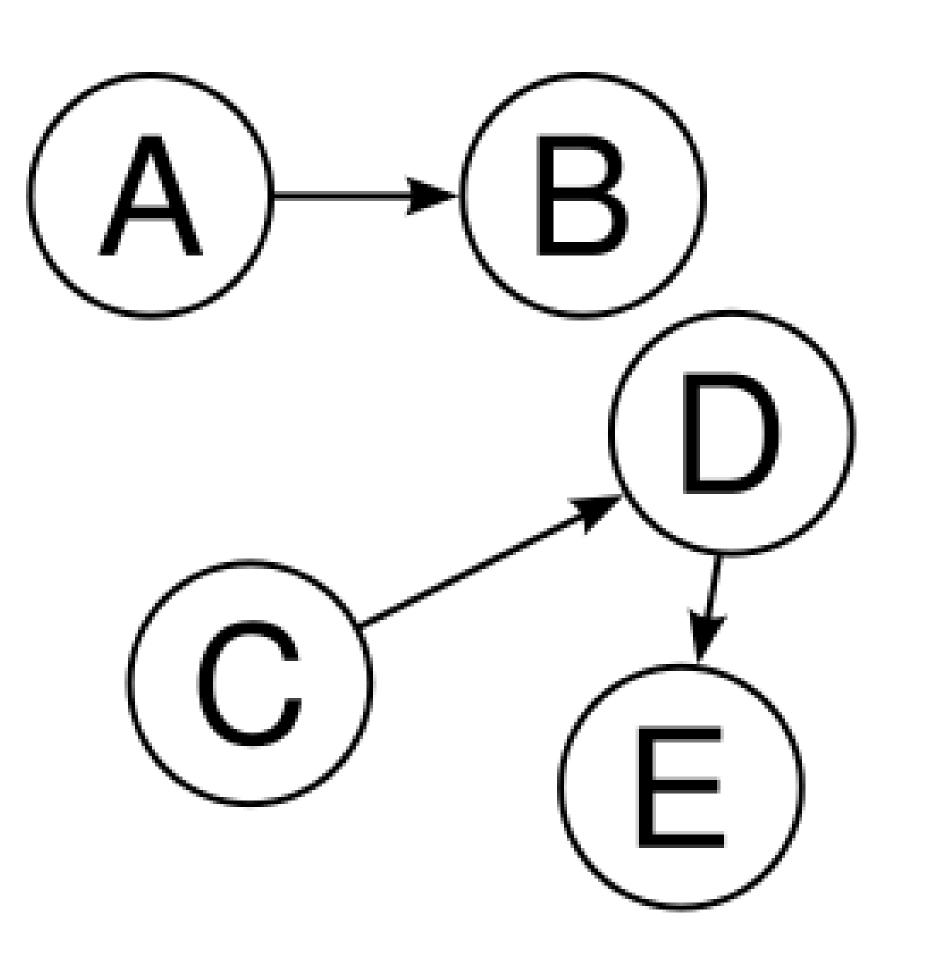
vector<void*> root;

vector<void*> root; // root[0] — это slist_node или list_node?

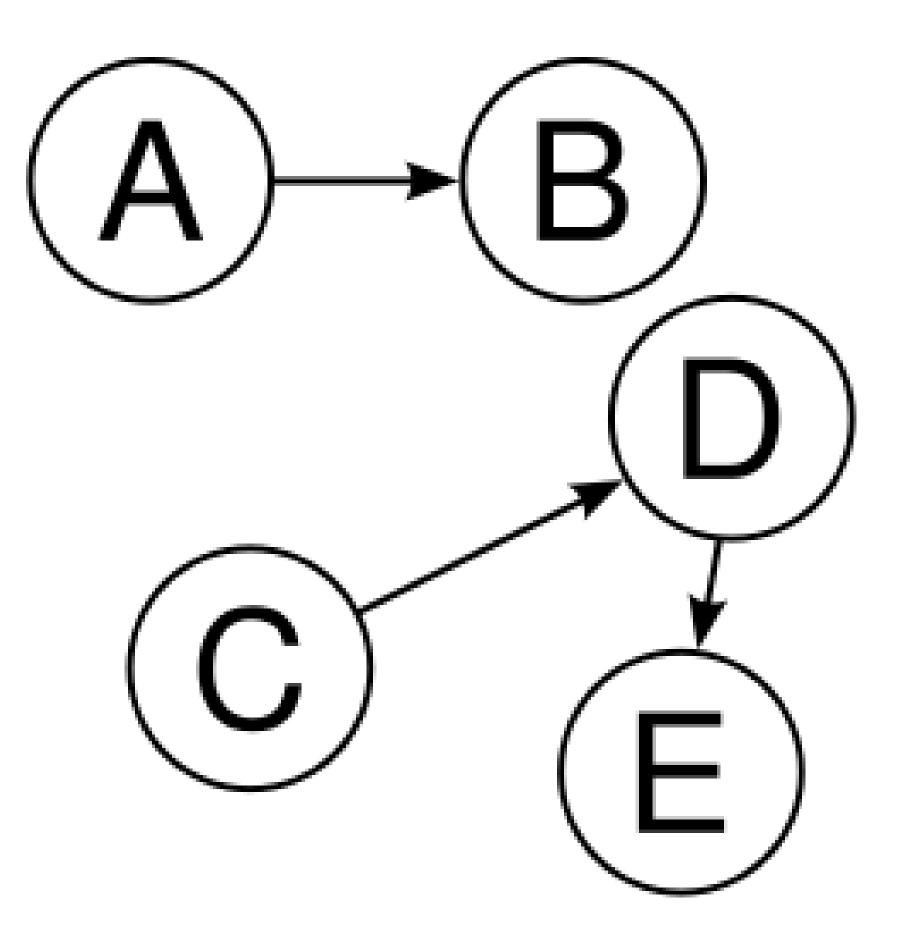
```
vector<void*> root; // Сколько указателей и где они? vector<Object*> root;
```

```
struct list_node {
    __meta vptr;
    list_node* next;
    list_node* prev;
};
struct slist_node {
    __meta vptr;
    slist_node* next;
```

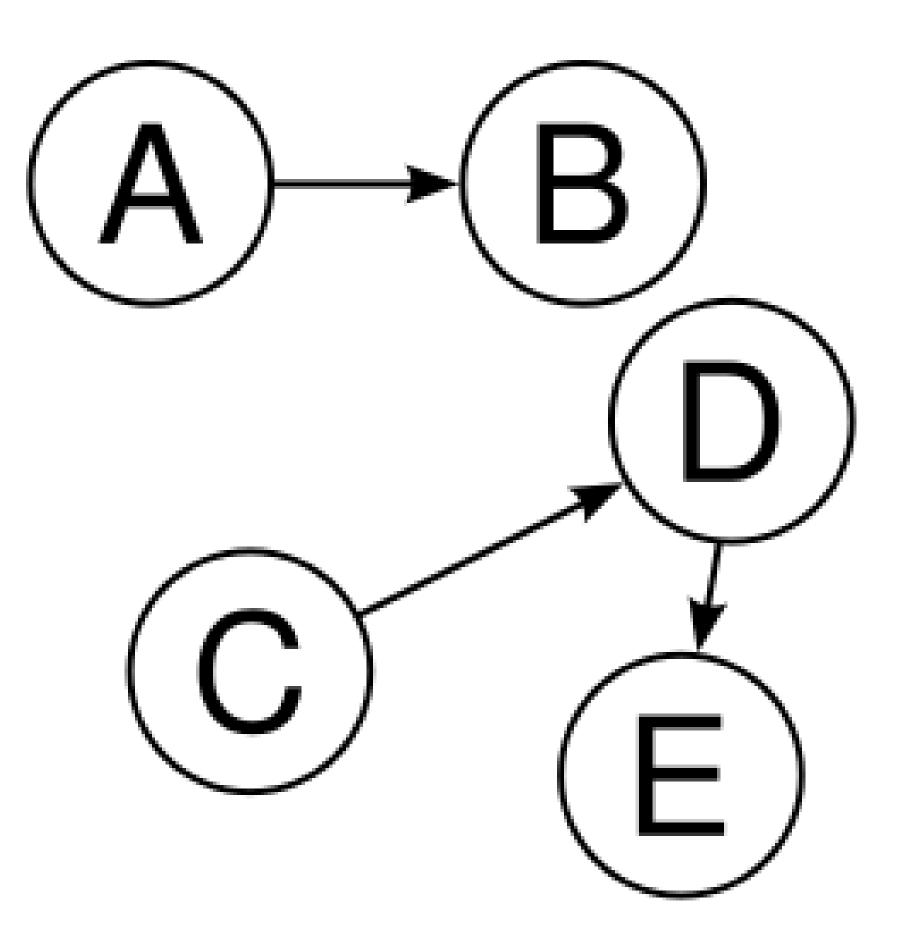
• Все потоки останавливаются



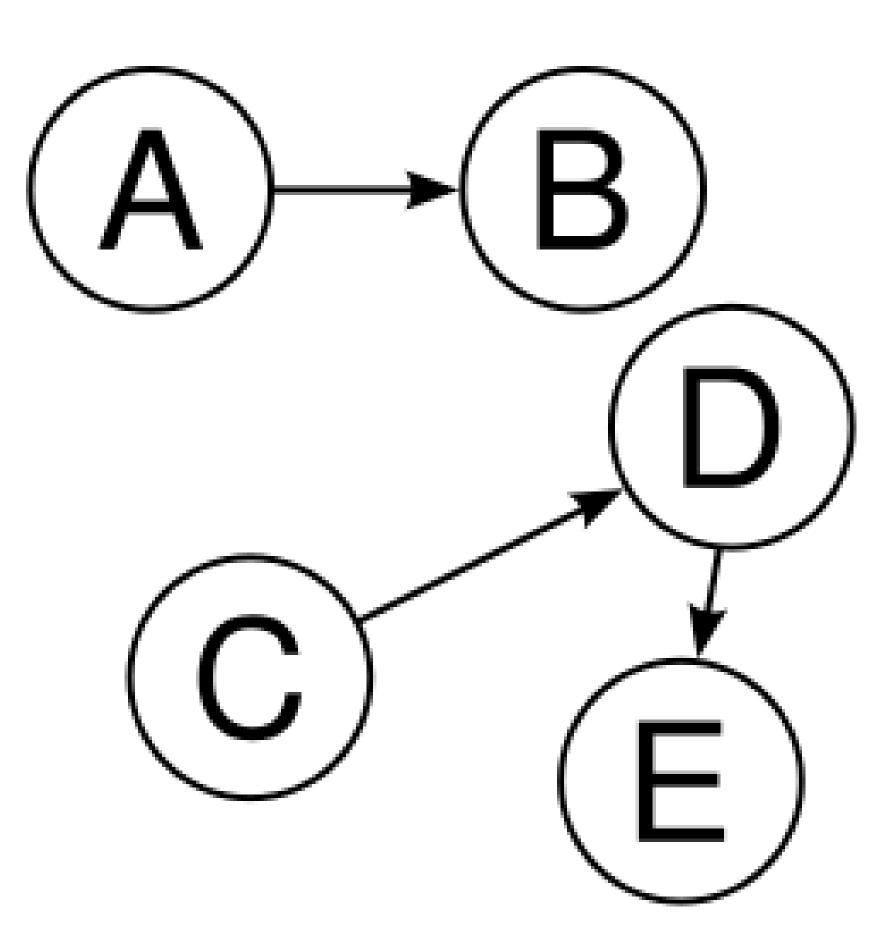
- Все потоки останавливаются
 - Проходим по всем узлам графа, помечая достижимые узлы



- Все потоки останавливаются
 - Проходим по всем узлам графа, помечая достижимые узлы
 - ????

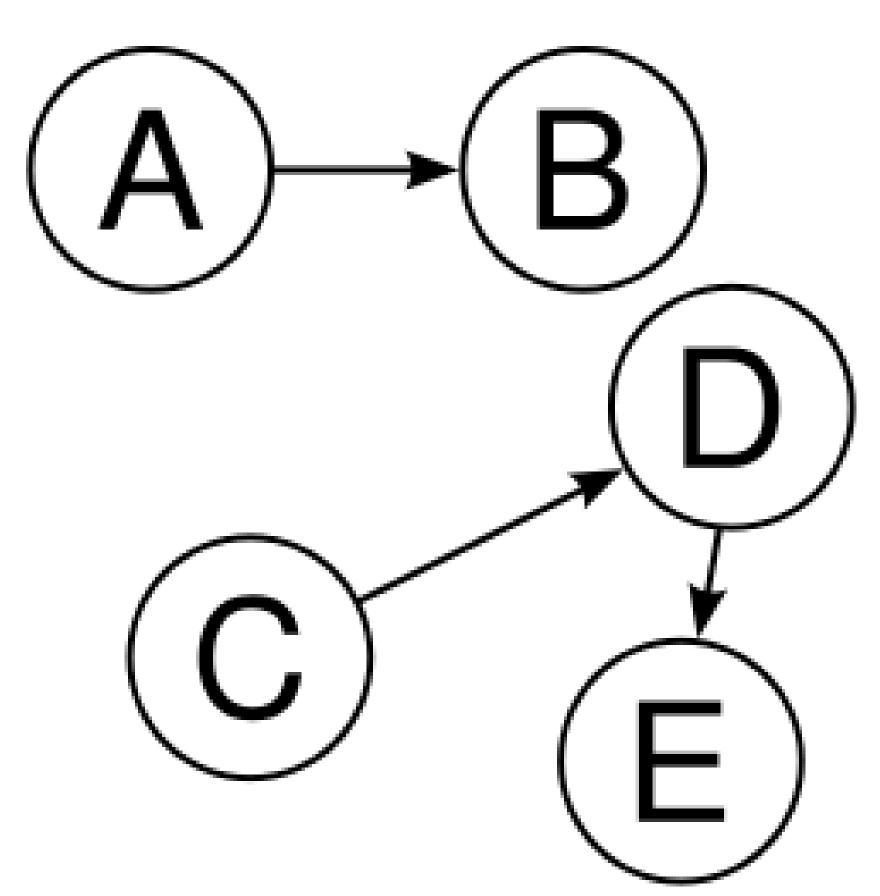


- Все потоки останавливаются
 - Проходим по всем узлам графа, помечая достижимые узлы
 - ????
 - Profit



- Все потоки останавливаются
 - Проходим по всем узлам графа, помечая достижимые узлы
 - ????
 - Profit

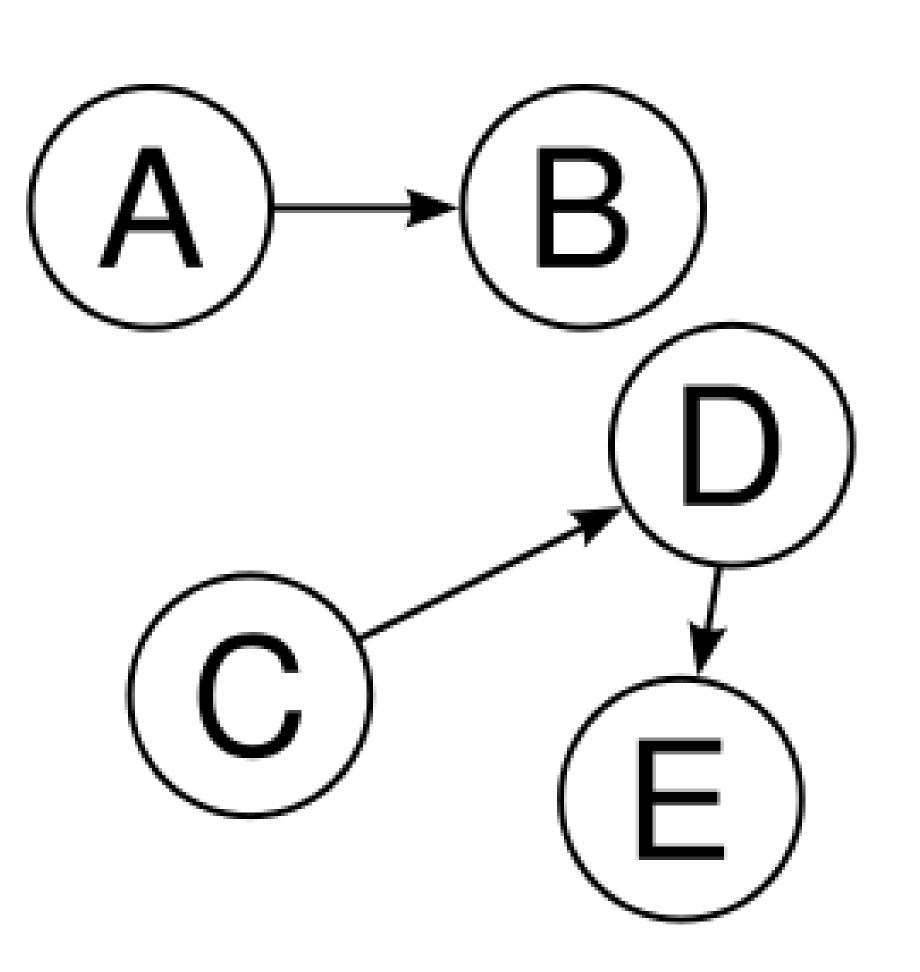
• На современном железе больше 1 ядра!



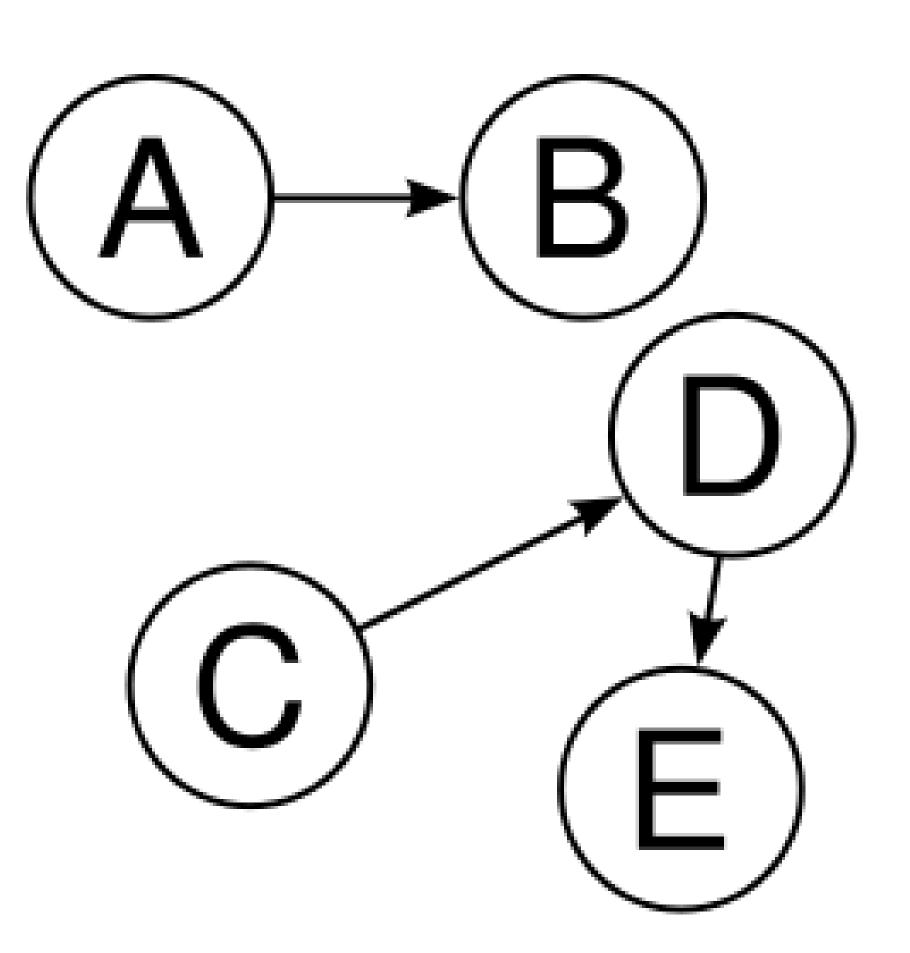
- Все потоки останавливаются
 - Проходим по всем узлам графа, помечая достижимые узлы
 - ????
 - Profit

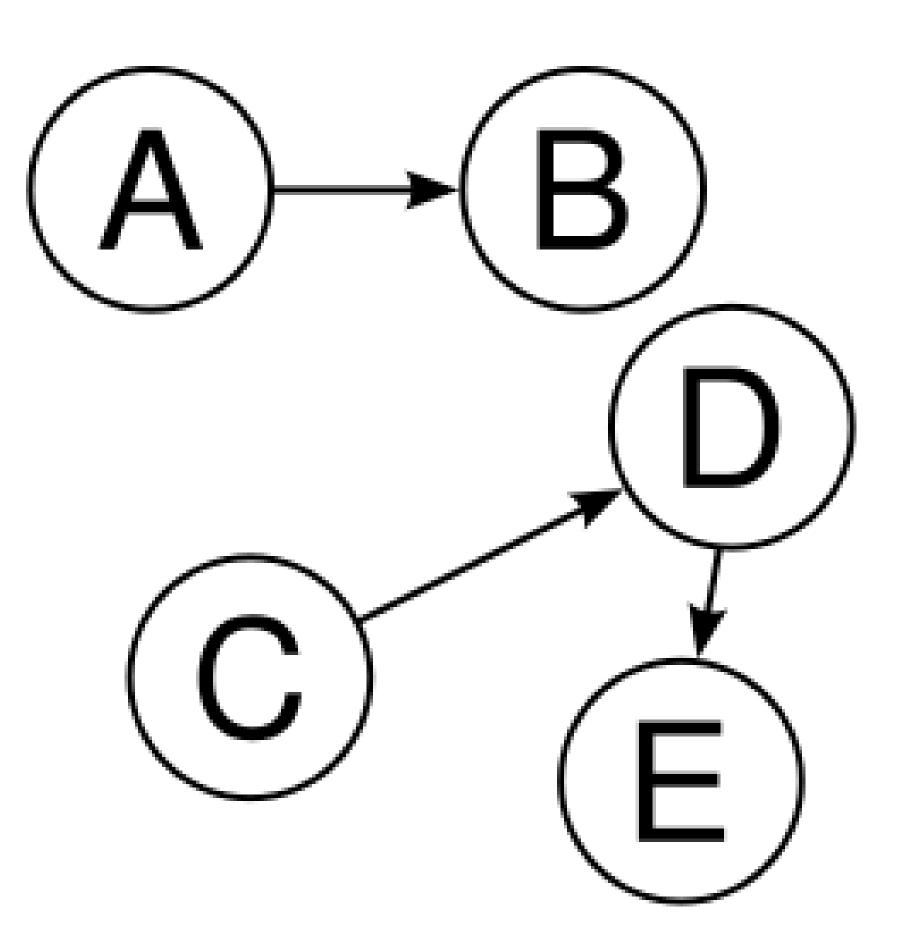
- На современном железе больше 1 ядра!
- При каждой сборке мусора мы проходимся по всем узлам → постоянно перепроверяя живые узлы

• Ничего не останавливается

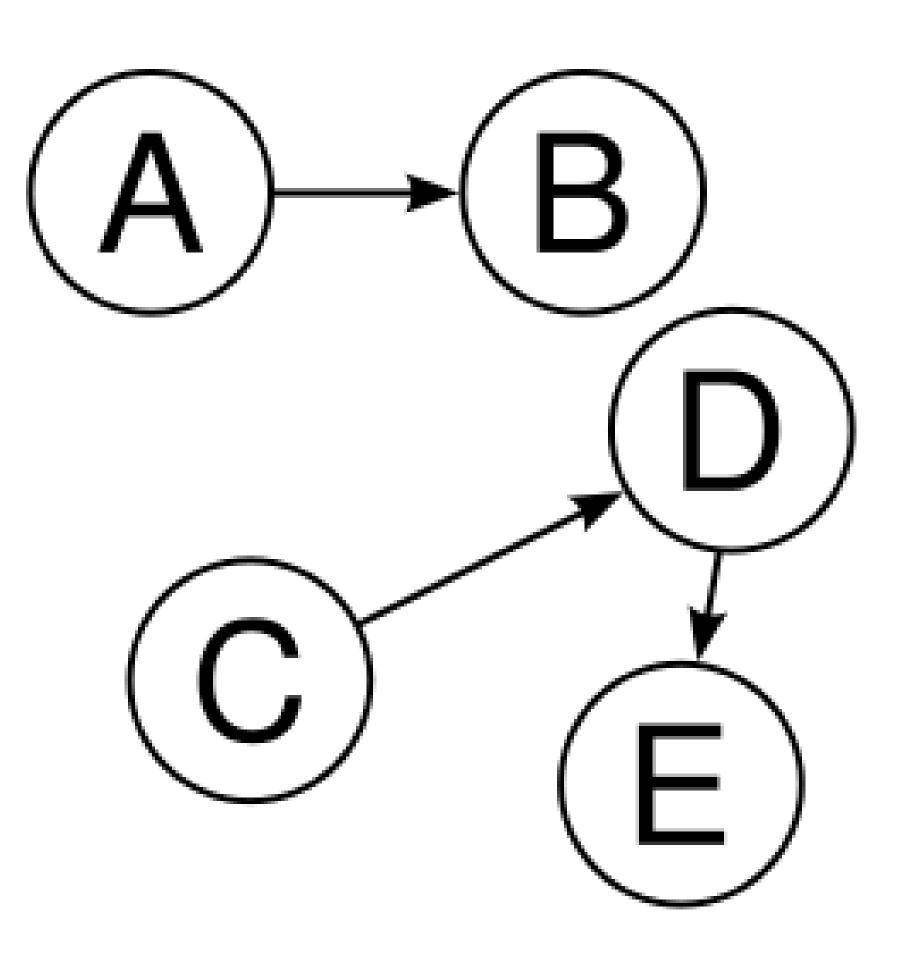


- Ничего не останавливается
 - Многопоточно разбираем мусор

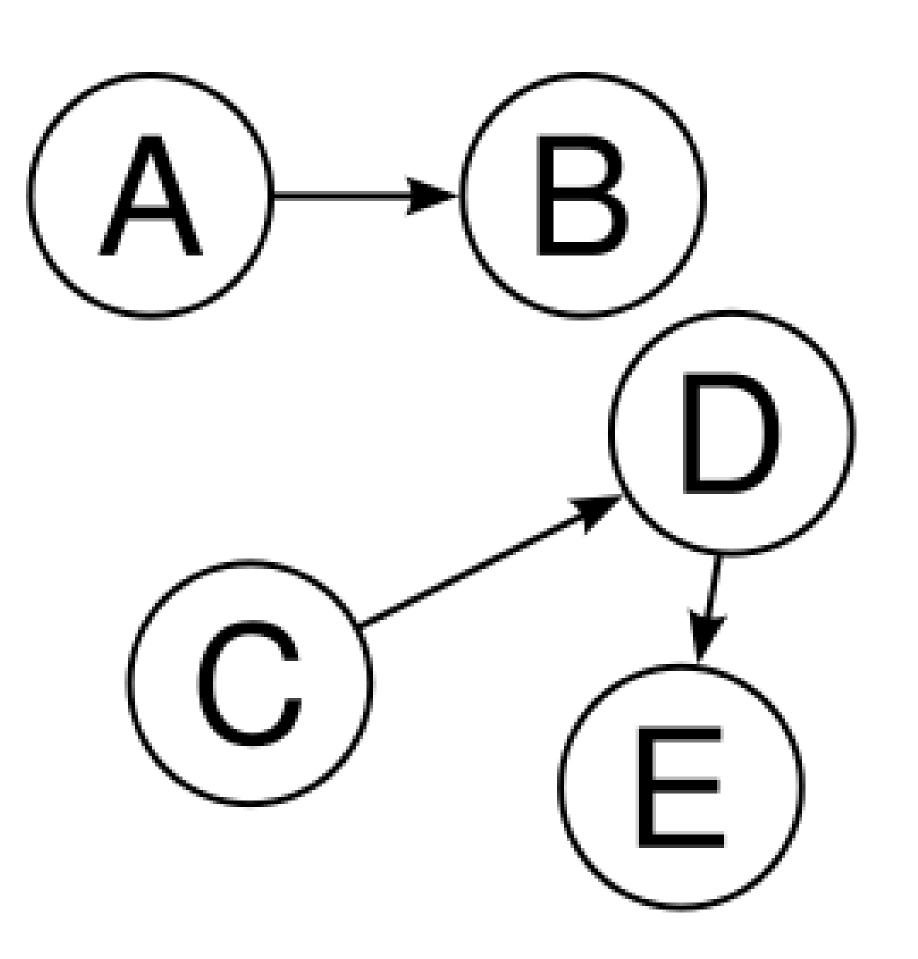




- Ничего не останавливается
 - Многопоточно разбираем мусор
 - Ура!

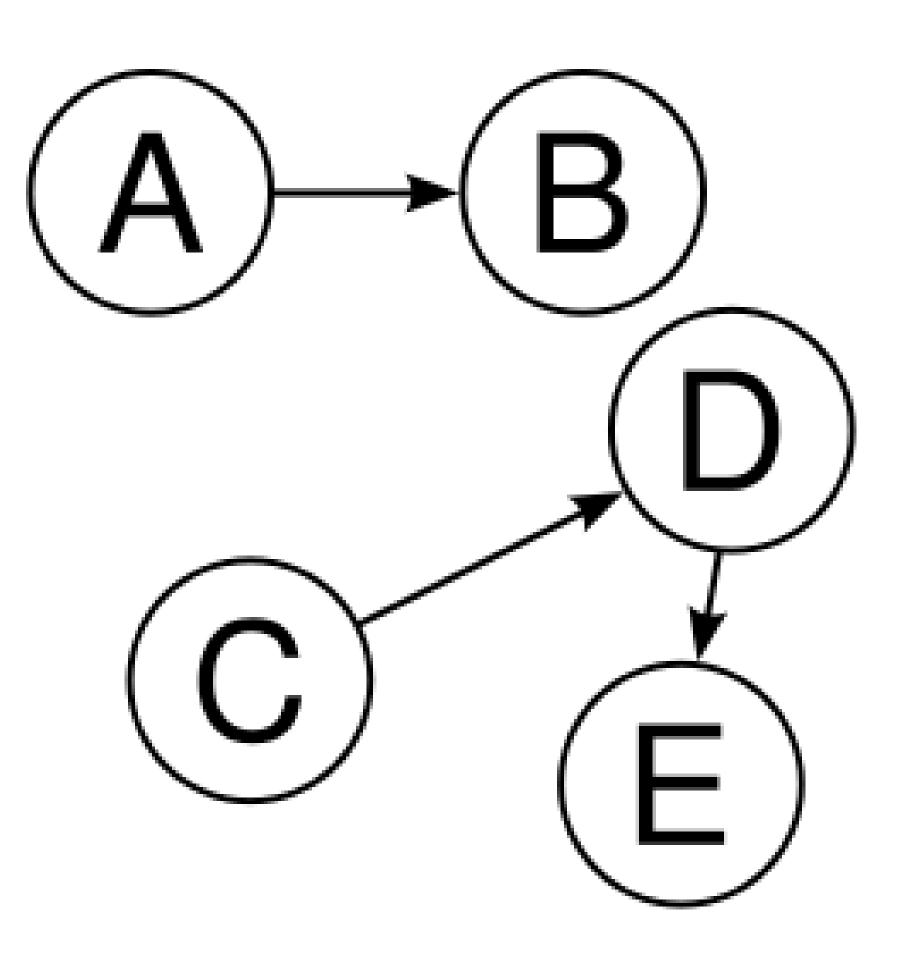


- Ничего не останавливается
 - Многопоточно разбираем мусор
 - Ура!
 - HO

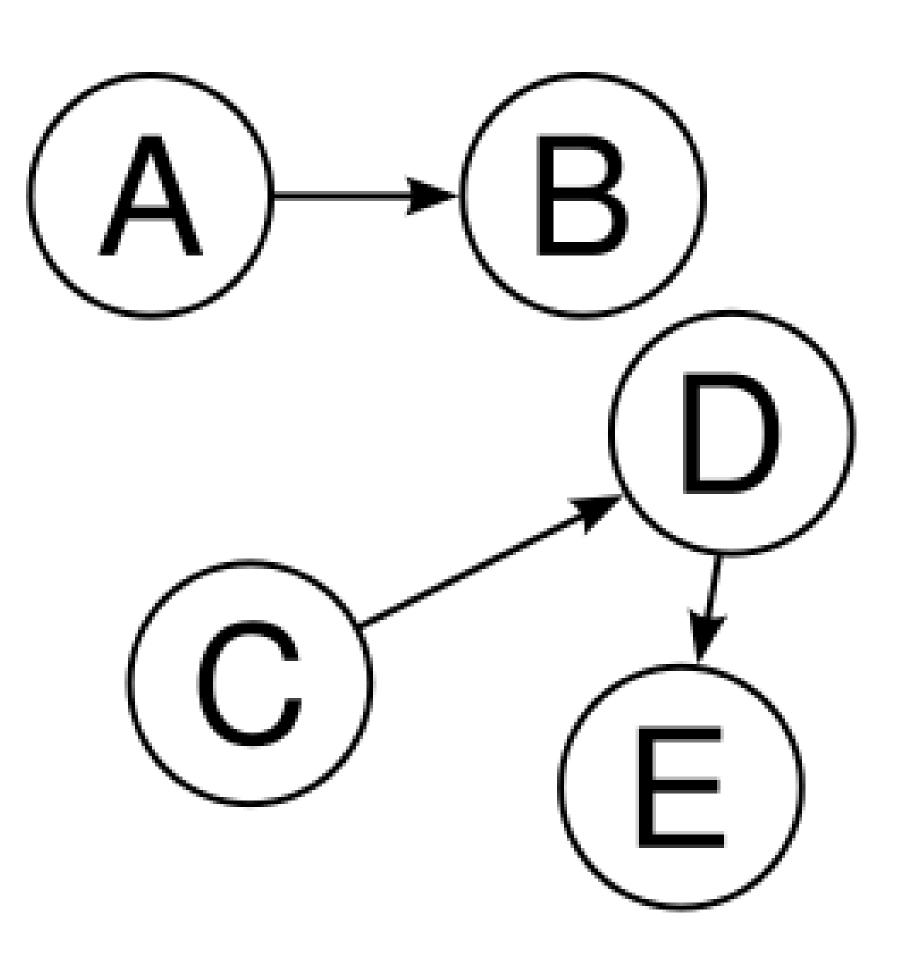


- Ничего не останавливается
 - Многопоточно разбираем мусор
 - Ура!
 - HO
 - Нам нужно синхронизировать все потоки через атомарные инструкции

Незаменимый C++ 99 / 135



- Ничего не останавливается
 - Многопоточно разбираем мусор
 - Ура!
 - HO
 - Нам нужно синхронизировать все потоки через атомарные инструкции
 - Это очень дорого, если не мы последними трогали ресурс



- Ничего не останавливается
 - Многопоточно разбираем мусор
 - Ура!
 - HO
 - Нам нужно синхронизировать все потоки через атомарные инструкции
 - Это очень дорого, если не мы последними трогали ресурс
 - Мы делаем это при каждой сборке мусора!

+ Не проходим лишние разы

shared_ptr<Object> some;

```
shared_ptr<Object> some; // Скоро будет больно!
```

```
shared_ptr<Object> some; // Скоро будет больно! shared_ptr<Object> prev = some.a; // ...
```

```
shared_ptr<Object> some; // Скоро будет больно! shared_ptr<Object> prev = some.a; // ... prev.a = some; // Циклические ссылки!
```

Trade-offs

• Плавающий мусор

- Плавающий мусор
 - ... или удар по производительности

- Плавающий мусор
 - ... или удар по производительности

• Останавливать все потоки и тормозить с непредсказуемыми задержками

- Плавающий мусор
 - ... или удар по производительности

- Останавливать все потоки и тормозить с непредсказуемыми задержками
 - ... или ничего не останавливать и тратить CPU такты впустую

- Плавающий мусор
 - ... или удар по производительности

- Останавливать все потоки и тормозить с непредсказуемыми задержками
 - ... или ничего не останавливать и тратить CPU такты впустую
- Плата за Object*

C++ vs. (Java + C#)

C#??Java

C#?? Java

• См. «Сборщики Мусора»

Logstash

 Программа для сбора, трансформации и складирования логов.

Бесплатное и очень популярное Open Source приложение на Java.

Что может пойти не так?

Logstash

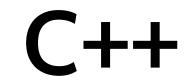
 Программа для сбора, трансформации и складирования логов.

Бесплатное и очень популярное Open Source приложение на Java.

Что может пойти не так?

%CPU %MEM COMMAND 505,5 2,9 java 66,9 1,4 daemon

Слабые места С++



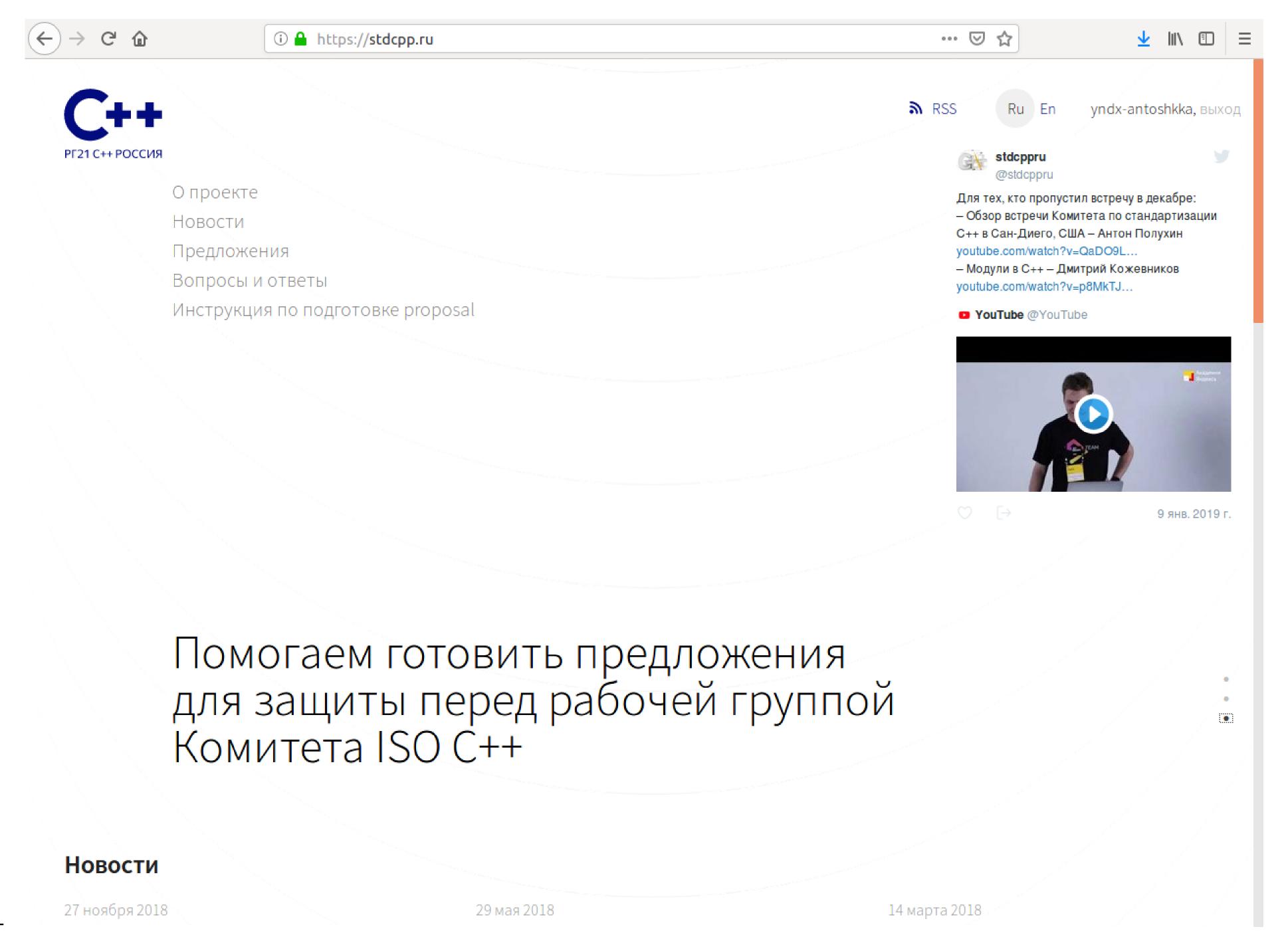
C++

- скрытый высокий порог вхождения
 - неограниченные возможности
 - проблемы с безопасностью

C++

- скрытый высокий порог вхождения
 - неограниченные возможности
 - проблемы с безопасностью
- отсутствие изкоробочности
 - крошечная стандартная библиотека
 - отсутствие готовой инфраструктуры

ОК, а что делать то?



• С++ везде

- С++ везде
- С++ крайне популярен

- С++ везде
- С++ крайне популярен
- С++ самый производительный инструмент для написания больших приложений

- С++ везде
- С++ крайне популярен
- C++ самый производительный инструмент для написания больших приложений
- Относитесь со скепсисом к бенчмаркам

- С++ везде
- С++ крайне популярен
- С++ самый производительный инструмент для написания больших приложений
- Относитесь со скепсисом к бенчмаркам
- He верьте managed языкам, которые говорят что они быстрее C++

- С++ везде
- С++ крайне популярен
- C++ самый производительный инструмент для написания больших приложений
- Относитесь со скепсисом к бенчмаркам
- He верьте managed языкам, которые говорят что они быстрее C++
- С++ не идеален!..

- С++ везде
- С++ крайне популярен
- C++ самый производительный инструмент для написания больших приложений
- Относитесь со скепсисом к бенчмаркам
- He верьте managed языкам, которые говорят что они быстрее C++
- С++ не идеален!..
 - ...но это исправимо

Спасибо

Полухин Антон

Старший разработчик Yandex. Тахі



antoshkka@gmail.com



antoshkka@yandex-team.ru



https://github.com/apolukhin



https://stdcpp.ru/

