```
vpcmpatd
vpand
vmovdau vmm11, ymmword ptr [rsp]
vpcmpgtd
vpand
vpand
vmovdqu ymm11, ymmword ptr [rsp - 128]
vpcmpqtd
vpand
vpand
vmovdqu ymm10, ymmword ptr [rsp + 96]
vpcmpgtd
vmovdqu ymm11, ymmword ptr [rsp - 32]
vpcmpgtd
vpcmpgtd
```

# 5-3: Arch intrinsics and inline assembly (Theory)

Artem Pavlov, TII, Abu Dhabi, 18.04.2025

#### Instruction Set Architecture (ISA)

- "Language" in which CPUs talk
- Most prominent examples: x86, Arm64, RISC-V
- ISA users require strong backward compatibility

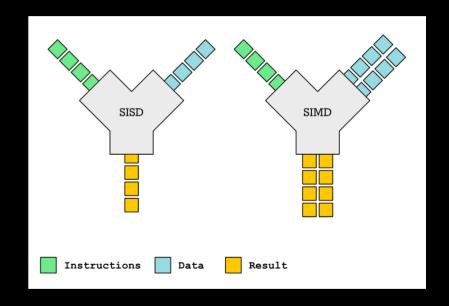
```
example::f:
                       .LBB0 1
              and
                       rsi, 8
                      .LBB0 4
                       .LBB0_6
11
     .LBB0 1:
12
              ret
     .LBB0 4:
                       rsi, -8
              and
17
```

#### ISA extensions

- Instructions sets which extend the "base" set
- x86 examples:
  - SIMD: SSE, SSE2, SSE3, SSSE3, SSE4, AVX, AVX2, AVX-512
  - Cryptography: AES-NI, SHA-NI, CLMUL, RDRAND
  - Bit manipulation: BMI, BMI2
- By default Rust uses only SSE and SSE2

#### SIMD: Single Instruction, Multiple Data

- Instructions which provide data-level parallelism
- x86 CPUs use SIMD instructions with fixed vector sizes (128, 256, and 512 bits)
- ARM and RISC-V provide "vector" extensions



# Problem: using ISA extensions

- Most user CPUs have "modern" extensions (e.g. AES-NI and AVX2)
- Extensions can make program much faster and even more secure
- But using an unavailable extension will cause CPU exception in the best case, and Undefined Behaviour in the worst

# Using CPU extensions in Rust

- Use of ISA extension has to be explicitly enabled
- It's done using target features
- Target feature can be enabled for a whole program or for a separate function
- In the latter case, we can call the function only after we have checked at runtime that CPU supports the necessary target feature

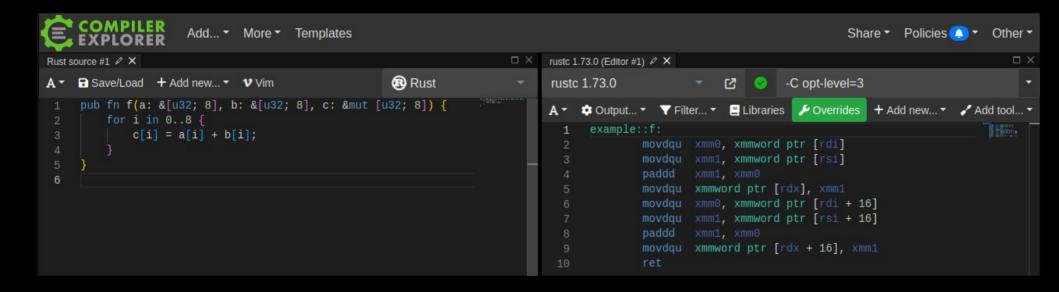
#### Enabling target feature for a whole program

- Selected target features can be enabled with
   -C target-feature=+aes,+avx2 compiler flag
- You can enable all target features supported by your CPU using -C target-cpu=native compiler flag
- Warning: the latter may result in a worse performance in some cases. Do not forget to benchmark!

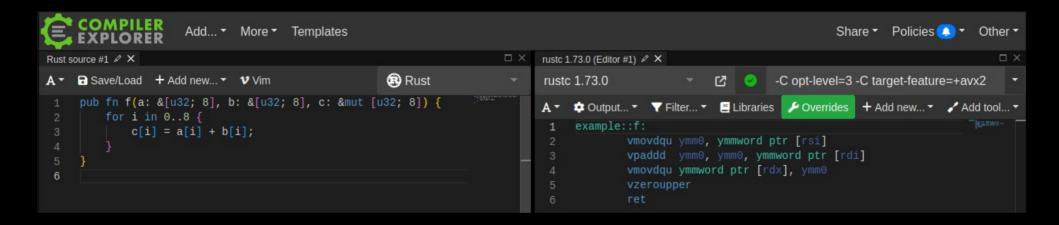
#### Example: scalar code

```
COMPILER
                     Add... ▼ More ▼ Templates
                                                                                                                             Share Policies T Other
    EXPLORER
Rust source #1 0 X
                                                                              rustc 1.73.0 (Editor #1) & X
                                                                                                                                                       \square \times
                                                      Rust
                                                                                                                    opt-level=3 -C target-feature=-sse.-sse2
A - Save/Load + Add new... - V Vim
                                                                               rustc 1.73.0
     pub fn f(a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8]) {
                                                                                   Output... T Filter...
                                                                                                          ■ Libraries  Poverrides + Add new... Add tool...
         for i in 0..8 {
             c[i] = a[i] + b[i];
                                                                                                     eax, dword ptr [rsi]
                                                                                                     eax, dword ptr [rdi]
                                                                                                     dword ptr [rdx], eax
                                                                                                     eax, dword ptr [rsi + 4]
                                                                                                     eax, dword ptr [rdi + 4]
                                                                                                     dword ptr [rdx + 4], eax
                                                                                                     eax, dword ptr [rsi + 8]
                                                                                                     eax, dword ptr [rdi + 8]
                                                                                                     dword ptr [rdx + 8], eax
                                                                                                     eax, dword ptr [rsi + 12]
                                                                                                     eax, dword ptr [rdi + 12]
                                                                                                     dword ptr [rdx + 12], eax
                                                                                                     eax, dword ptr [rsi + 16]
                                                                                                     eax, dword ptr [rdi + 16]
                                                                                                     dword ptr [rdx + 16], eax
                                                                                                     eax, dword ptr [rsi + 20]
                                                                                                     eax, dword ptr [rdi + 20]
                                                                                                     dword ptr [rdx + 20], eax
                                                                                                     eax, dword ptr [rsi + 24]
                                                                                                     eax, dword ptr [rdi + 24]
                                                                                                     dword ptr [rdx + 24], eax
                                                                                                     eax, dword ptr [rsi + 28]
                                                                                                     eax, dword ptr [rdi + 28]
                                                                                                     dword ptr [rdx + 28], eax
```

# Example: autovectorization (SSE2)



# Example: autovectorization (AVX2)



#### <u>Autovectorization fragility</u>

```
Add... ▼ More ▼ Templates
                                                                                                                            Share ▼ Policies △ ▼ Other ▼
Rust source #1 @ X
                                                                              rustc 1.73.0 (Editor #1) & X
                                                     R Rust
    ■ Save/Load + Add new... ▼ Vim
                                                                                                                   -C opt-level=3 -C target-feature=+avx2
                                                                              rustc 1.73.0
     pub fn f(a: &[u32; 8], b: &[u32; 8], c: &mut [u32]) {
                                                                                  Output... T Filter... Filter... Add tool... Add tool...
         for i in 0..8 {
             c[i] = a[i] + b[i];
                                                                                                    .LBB0 1
                                                                                                    eax, dword ptr [rsi]
                                                                                                    dword ptr [rdx], eax
                                                                                                    .LBB0 3
                                                                                                    eax, dword ptr [rdi + 4]
                                                                                                    eax, dword ptr [rsi + 4]
                                                                                                    dword ptr [rdx + 4], eax
                                                                                                    rcx, 2
                                                                                                    .LBB0 5
                                                                                                    eax, dword ptr [rdi + 8]
                                                                                                    eax, dword ptr [rsi + 8]
                                                                                                    dword ptr [rdx + 8], eax
                                                                                                    eax, dword ptr [rdi + 12]
                                                                                                    eax, dword ptr [rsi + 12]
                                                                                                    dword ptr [rdx + 12], eax
                                                                                                    .LBB0 9
                                                                                                    eax, dword ptr [rdi + 16]
                                                                                                    eax. dword ptr [rsi + 16]
                                                                                                    dword ntr [rdy + 16] eav
```

#### "Fixing" autovectorization

```
Add... ▼ More ▼ Templates
                                                                                                                   Rust source #1 @ X
                                                                        rustc 1.73.0 (Editor #1) & X
                                                 R Rust
A - Save/Load + Add new... - V Vim
                                                                                                          -C opt-level=3 -C target-feature=+avx2
                                                                        rustc 1.73.0
    pub fn f(a: &[u32; 8], b: &[u32; 8], c: &mut [u32]) {
                                                                        A • Output... • V Filter... • E Libraries / Overrides + Add new... • Add tool... •
        assert!(c.len() >= 8);
        for i in 0..8 {
                                                                                            rcx, 8
            c[i] = a[i] + b[i];
                                                                                             .LBB0 2
                                                                                     vmovdqu ymm0, ymmword ptr [rdi]
                                                                                     vpaddd ymm0, ymm0, ymmword ptr [rsi]
                                                                                     vmovdqu ymmword ptr [rdx], ymm0
                                                                              .LBB0 2:
                                                                                            rdi, [rip + .L unnamed 1]
                                                                                            rdx, [rip + .L_unnamed 2]
                                                                                             esi, 30
                                                                                            gword ptr [rip + core::panicking::panic@GOT
                                                                              .L unnamed 3:
                                                                                     .ascii "/app/example.rs"
                                                                              .L unnamed 1:
                                                                                     .ascii "assertion failed: c.len() >= 8"
                                                                              .L unnamed 2:
                                                                                            .L_unnamed 3
```

#### Target feature detection at runtime

- Target features can be detected at runtime using std::is\_x86\_feature\_detected! Macro
- The macro is available (for now) only on std and x86 targets
- An alternative: the cpufeatures crate

# Selectively enabling target features

- Target feature can be enabled for a function with #[target\_feature(enable = "...")] attribute
- Such function has to be unsafe, since to call it we have to check availability of the required ISA extensions
- Inside the function compiler can use instruction from enabled extensions
- RFC: https://rust-lang.github.io/rfcs/2045-target-feature.html

# Example: #[target\_feature(enable = "...")]

```
COMPILER
                    Add... ▼ More ▼ Templates
                                                                                                                   Share ▼ Policies △ ▼ Other ▼
Rust source #1 & X
                                                                        rustc 1.73.0 (Editor #1) 0 X
                                                                                                                                          \square \times
   ■ Save/Load + Add new... ▼ Vim
                                                 R Rust
                                                                                                         -C opt-level=3
                                                                        rustc 1.73.0
     pub fn add default(
                                                                            a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
                                                                              example::add default:
                                                                                     movdqu xmm0, xmmword ptr [rdi]
         for i in 0..8 {
                                                                                     movdqu xmm1, xmmword ptr [rsi]
             c[i] = a[i] + b[i];
                                                                                     movdqu xmmword ptr [rdx], xmm1
                                                                                     movdqu xmm0, xmmword ptr [rdi + 16]
                                                                                     movdqu xmm1, xmmword ptr [rsi + 16]
     #[target feature(enable = "avx2")]
     pub unsafe fn add avx2(
                                                                                     movdgu xmmword ptr [rdx + 16], xmm1
         a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
         for i in 0..8 {
                                                                              example::add avx2:
             c[i] = a[i] + b[i];
                                                                                     vmovdqu ymm0, ymmword ptr [rsi]
                                                                                     vpaddd ymm0, ymm0, ymmword ptr [rdi]
                                                                                     vmovdqu ymmword ptr [rdx], ymm0
                                                                                     vzeroupper
```

#### Example: autodetection 1

```
fn add inner(
         a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
         for i in 0..8 {
             c[i] = a[i] + b[i];
     #[target_feature(enable = "avx2")]
    unsafe fn add_avx2(
         a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
         add_inner(a, b, c)
     pub fn add(
         a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
         if std::is x86 feature detected!("avx2") {
             unsafe { add avx2(a, b, c) }
          else {
21
             add_inner(a, b, c)
```

#### Example: autodetection 2

```
example::add avx2:
             vmovdqu ymm0, ymmword ptr [rsi]
             vpaddd ymm0, ymm0, ymmword ptr [rdi]
             vmovdqu ymmword ptr [rdx], ymm0
             vzeroupper
             ret
     example::add:
11
12
                     rax, qword ptr [rip + std_detec
             mov
                     rax, gword ptr [rax]
             mov
                     .LBB1 1
                     .LBB1 4
21
     .LBB1 3:
```

```
.LBB1 3:
               xmm0, xmmword ptr [r15]
                xmm1, xmmword ptr [r14]
       paddd
               xmmword ptr [rbx], xmm1
               xmm0, xmmword ptr [r15 + 16]
               xmm1, xmmword ptr [r14 + 16]
               xmmword ptr [rbx + 16], xmm1
.LBB1 1:
                gword ptr [rip + std detect::det
                .LBB1 3
.LBB1 4:
                example::add avx2
```

#### Example: autodetection 3

```
Add... ▼ More ▼ Templates
                                                                                                                  Share Policies
                                                                                                                                      Other *
Rust source #1 / X
                                                            \square \times
                                                                 rustc 1.73.0 (Editor #1) 0 X
                                           R Rust
                                                                                                   -C opt-level=3 -C target-feature=+avx2
   Save/Load + Add new... Vim
                                                                  rustc 1.73.0
     fn add inner(
                                                                 a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
                                                                             vmovdqu ymm0, ymmword ptr [rsi]
         for i in 0..8 {
                                                                             vpaddd ymm0, ymm0, ymmword ptr [rdi]
            c[i] = a[i] + b[i];
                                                                             vmovdqu ymmword ptr [rdx], ymm0
     #[target feature(enable = "avx2")]
     unsafe fn add avx2(
         a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
         add_inner(a, b, c)
     pub fn add(
         a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
         if std::is x86 feature detected!("avx2") {
             unsafe { add_avx2(a, b, c) }
         } else {
             add_inner(a, b, c)
```

#### arch intrinsics

- Special low-level arch-specific unsafe functions which compile down to one instruction\*
- For example, \_mm256\_add\_epi32 usually compiles to vpaddd
- Defined in the std::arch module
- Callers must ensure that CPU has the required target feature

#### std::arch::x86 64



#### Module x86\_64

Structs

Constant

Functions

Type Aliases

```
fxsave (x86 or x86-64) and fxsr
Izent u32<sup>△</sup> (x86 or x86-64) and lzent
mm256 abs epi8<sup>♠</sup> (x86 or x86-64) and avx2
mm256 abs epi16<sup>A</sup> (x86 or x86-64) and avx2
mm256 abs epi32<sup>A</sup> (x86 or x86-64) and avx2
mm256_add_epi8^ (x86 or x86-64) and avx2
_mm256_add_epi16<sup>\text{\Delta}</sup> (x86 or x86-64) and avx2
_mm256_add_epi32<sup>Δ</sup> (x86 or x86-64) and avx2
_mm256_add_epi64<sup>A</sup> (x86 or x86-64) and avx2
_mm256_add_pd^ (x86 or x86-64) and avx
mm256_add_ps (x86 or x86-64) and avx
mm256_adds_epi8 (x86 or x86-64) and avx2
_mm256_adds_epu8<sup>♠</sup> (x86 or x86-64) and avx2
```

\_mm256\_adds\_epu16<sup>△</sup> (x86 or x86-64) and avx2

TXTSTOT (x86 or x86-64) and fxsr

```
Restores the XMM, MMX, MXCSR, and X8/ FPU registers from
the 512-byte-long 16-byte-aligned memory region mem_addr.
Restores the XMM, MMX, MXCSR, and x87 FPU registers from
the 512-byte-long 16-byte-aligned memory region mem addr.
Saves the x87 FPU, MMX technology, XMM, and MXCSR registers
to the 512-byte-long 16-byte-aligned memory region mem addr.
Saves the x87 FPU, MMX technology, XMM, and MXCSR registers
to the 512-byte-long 16-byte-aligned memory region mem addr.
Counts the leading most significant zero bits.
Counts the leading most significant zero bits.
Computes the absolute values of packed 8-bit integers in a.
Computes the absolute values of packed 16-bit integers in a.
Computes the absolute values of packed 32-bit integers in a.
Adds packed 8-bit integers in a and b.
Adds packed 16-bit integers in a and b.
Adds packed 32-bit integers in a and b.
Adds packed 64-bit integers in a and b.
Adds packed double-precision (64-bit) floating-point elements in
a and b.
Adds packed single-precision (32-bit) floating-point elements in
a and b.
Adds packed 8-bit integers in a and b using saturation.
Adds packed 16-bit integers in a and b using saturation.
Adds packed unsigned 8-bit integers in a and b using
saturation.
Adds packed unsigned 16-bit integers in a and b using
```

#### Intel Intrinsics Guide

- The official resource about x86 intrinsics: https://www.intel.com/content/www/us/en/docs/intrinsics-guide/index.html
- Contains detailed description of intrinsics and additional information (e.g. latency, throughput, etc.)



#### Intrinsics vs std::arch::asm!

- You do not manually allocate registers with an intrinsicsbased code, the compiler can handle stack spilling if necessary
- The compiler "understands" intrinsics to a certain extent
- The compiler may change order of intrinsic calls
- For example, \_mm256\_sub\_epi32(x, x) can be compiled as vxorps xmm0, xmm0, xmm0

#### Vector types

- x86 arch intrinsics work with "vector types"
- For example, \_\_m128, \_\_m128i, \_\_m256, \_\_m256d, \_\_m256i, etc.
- \_\_m256i depending on intrinsic an be interpreted as u8x32, i8x32, u16x16, i16x16, u32x8, i32x8, u64x4, i64x4
- m256 is interpreted as f32x8, and m256d as f64x4
- Variables of vector types are usually processed in associated SIMD registers: XMM, YMM, and ZMM

#### Example: intrinsics 1

```
#[cfg(target arch = "x86")]
use std::arch::x86::*:
#[cfg(target arch = "x86 64")]
use std::arch::x86 64::*;
pub fn add(
    a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8],
    if is_x86_feature_detected!("avx2") {
        unsafe { add avx2(a, b, c) }
     } else if is x86 feature detected!("sse2") {
        unsafe { add sse2(a, b, c) }
        add soft(a, b, c)
fn add_soft(a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8]) {
    for i in 0..8 {
        c[i] = a[i] + b[i];
```

```
#[target feature(enable = "sse2")]
unsafe fn add sse2(a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8]) {
   let a1: m128i = mm loadu si128(a.as_ptr().cast());
   let b1: m128i = mm loadu si128(b.as ptr().cast());
   let res1: m128i = mm add epi32(a1, b1);
   mm storeu si128(c.as mut ptr().cast(), res1);
   let a2: __m128i = _mm_loadu_si128(a.as_ptr().add(4).cast());
   let b2: m128i = mm loadu_si128(b.as_ptr().add(4).cast());
   let res2: m128i = mm add epi32(a2, b2);
   mm storeu si128(c.as mut ptr().add(4).cast(), res2);
#[target_feature(enable = "avx2")]
unsafe fn add_avx2(a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8]) {
   let a: __m256i = _mm256_loadu_si256(a.as_ptr().cast());
   let b: __m256i = _mm256_loadu_si256(b.as_ptr().cast());
   let res: __m256i = _mm256_add_epi32(a, b);
   mm256 storeu si256(c.as mut ptr().cast(), res);
```

#### Example: intrinsics 2

```
eax, dword ptr [r14]
eax, dword ptr [r15]
dword ptr [rbx], eax
eax, dword ptr [r14 + 4]
eax, dword ptr [r15 + 4]
dword ptr [rbx + 4], eax
eax, dword ptr [r14 + 8]
eax, dword ptr [r15 + 8]
dword ptr [rbx + 8], eax
eax, dword ptr [r14 + 12]
eax, dword ptr [r15 + 12]
dword ptr [rbx + 12], eax
eax, dword ptr [r14 + 16]
eax, dword ptr [r15 + 16]
dword ptr [rbx + 16], eax
eax, dword ptr [r14 + 20]
eax, dword ptr [r15 + 20]
dword ptr [rbx + 20], eax
eax, dword ptr [r14 + 24]
eax, dword ptr [r15 + 24]
dword ptr [rbx + 24], eax
eax, dword ptr [r14 + 28]
eax, dword ptr [r15 + 28]
dword ptr [rbx + 28], eax
```

```
example::add sse2:
       movdqu xmm0, xmmword ptr [rdi]
       movdau
               xmm1, xmmword ptr [rsi]
        paddd
               xmmword ptr [rdx], xmm1
       movdau
               xmm0, xmmword ptr [rdi + 16]
       movdau
               xmm1, xmmword ptr [rsi + 16]
       movdqu
       paddd
                xmmword ptr [rdx + 16], xmm1
       movdau
       ret
example::add avx2:
        vmovdqu ymm0, ymmword ptr [rsi]
        vpaddd ymm0, ymm0, ymmword ptr [rdi]
        vmovdqu ymmword ptr [rdx], ymm0
        vzeroupper
        ret
```

#### cfg-based switching

```
pub fn add(a: &[u32; 8], b: &[u32; 8], c: &mut [u32; 8]) {
    cfg_if::cfg_if!{
        if #[cfg(target_feature = "avx2")] {
            unsafe { add_avx2(a ,b, c) }
        } else if #[cfg(target_feature = "sse2")] {
            unsafe { add_sse2(a ,b, c) }
        } else {
            add_default(a ,b, c)
        }
}
```

#### Example: bounding boxes

- We have 8\*N points given as u32 and list of bounding boxes
- We want to find points which are inside of at least one bounding box

```
pub unsafe fn foo(
   x: &[ m256i; N],
   y: &[__m256i; N],
   z: &[__m256i; N],
   bboxes: &[[__m256i; 6]],
) -> [__m256i; N] {
   let mut res = [_mm256_setzero_si256(); N];
   for bbox in bboxes {
        for i in 0..N {
            let tx = _mm256_and_si256(
                mm256 cmpqt epi32(x[i], bbox[0]),
                mm256_cmpgt_epi32(bbox[1], x[i]),
            let ty = mm256 and si256(
               _mm256_cmpgt_epi32(y[i], bbox[2]),
                _mm256_cmpgt_epi32(bbox[3], y[i]),
            let t = mm256 and si256(tx, ty);
            let tz = mm256 and si256(
                mm256_cmpgt_epi32(z[i], bbox[4]),
                _mm256_cmpgt_epi32(bbox[5], z[i]),
            let t = mm256 and si256(t, tz);
            res[i] = _mm256_or_si256(res[i], t);
    res
```

#### Target feature disadvantages

- #[target\_feature(enable = "...")] have to be marked unsafe (changed in Rust 1.86)
- Runtime target feature checks are not enforced by compiler
- It's currently impossible to specify that target feature will NOT be present during execution
- Compiler sometimes generates suboptimal code
- No vector extensions support
- On x86 SIMD-based code badly interacts with soft-float targets

#### Future of target features

 Reduction of unsafe amount in target feature v1.1:

https://rust-lang.github.io/rfcs/2396-target-feature-1.1.html

- Vector extensions support: https://github.com/rust-lang/rfcs/pull/3268
- Portable simd (a.k.a std::simd): https://github.com/rust-lang/portable-simd

# Questions?