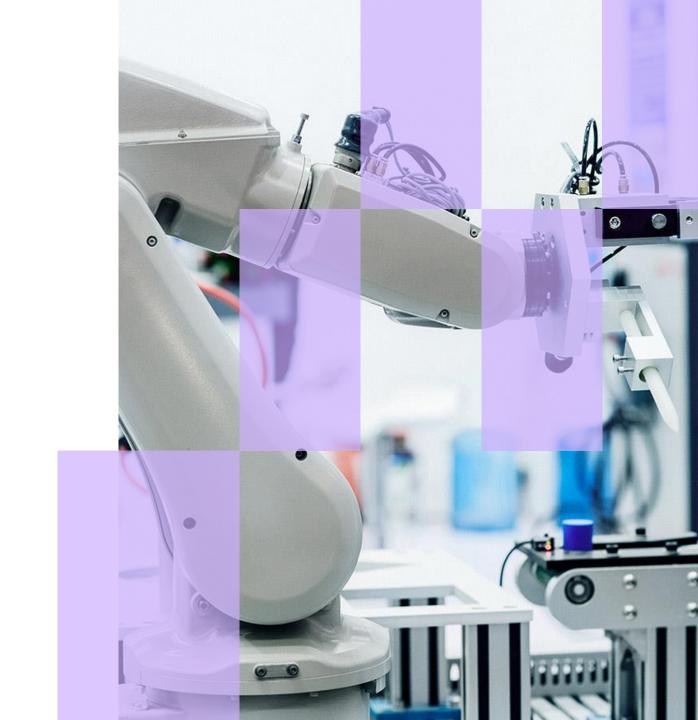
Building Bridges

C++ Interop., Foreign Function Interfaces & ABI





Gareth Williamson Software Engineer



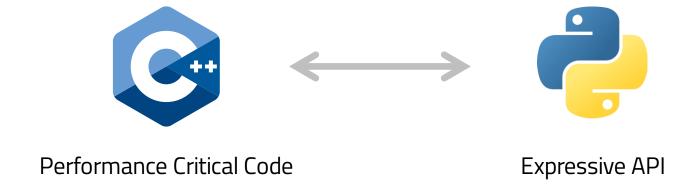
Why build a multi-language system?

Why build a multi-language system?

Use the right tool for the job

Why build a multi-language system?

Use the right tool for the job



Why build a multi-language system?

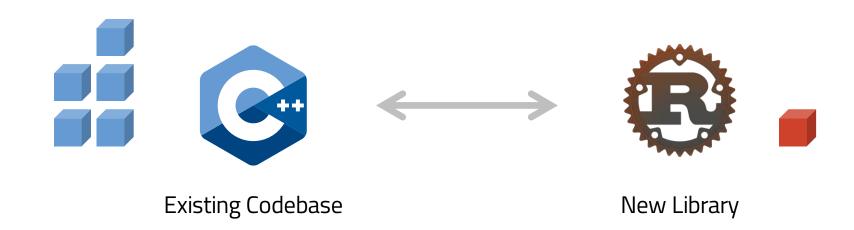
Use the right tool for the job

Re-use existing code

Why build a multi-language system?

Use the right tool for the job

Re-use existing code



How can we build a multi-language system?

How can we build a multi-language system?

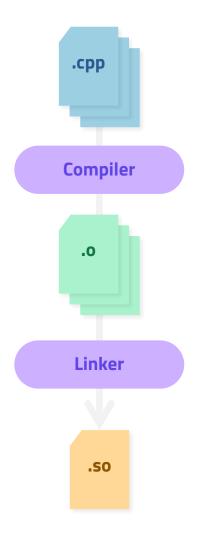
Inter-process

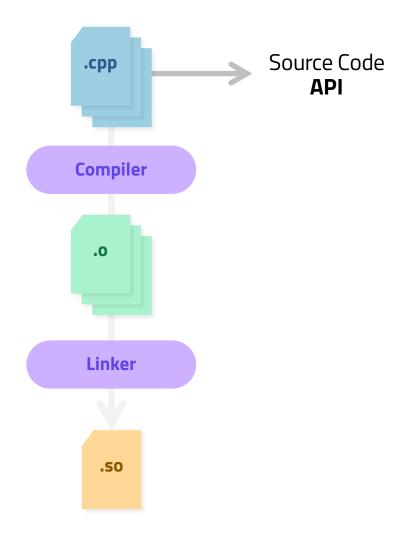
How can we build a multi-language system?

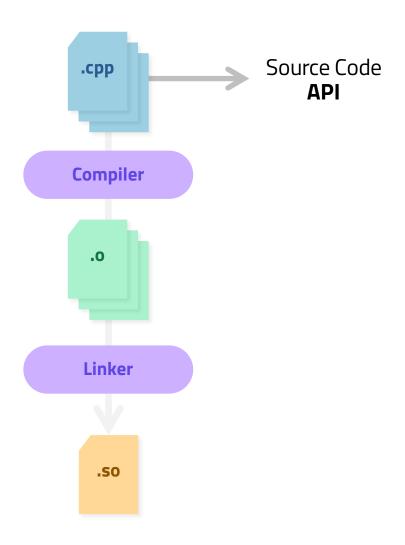
Inter-process

Link Object Code

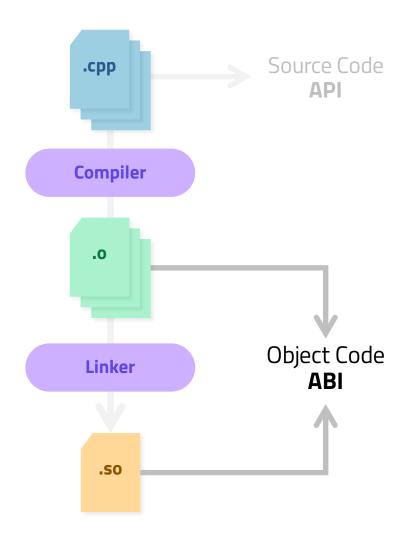
Application Binary Interface



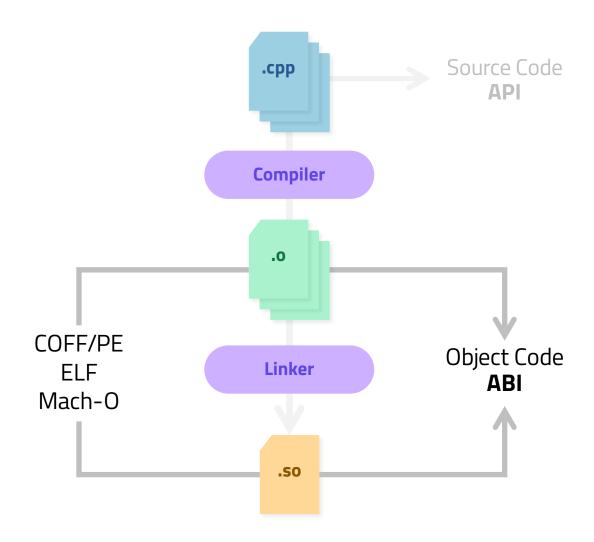




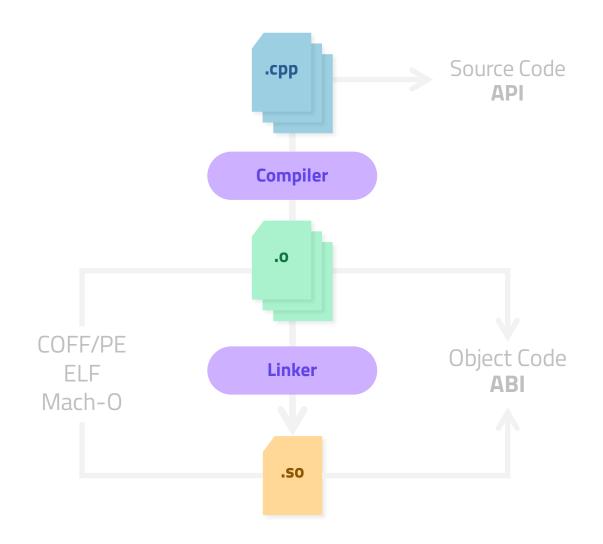
API is the **source code** level interface



API is the **source code** level interface

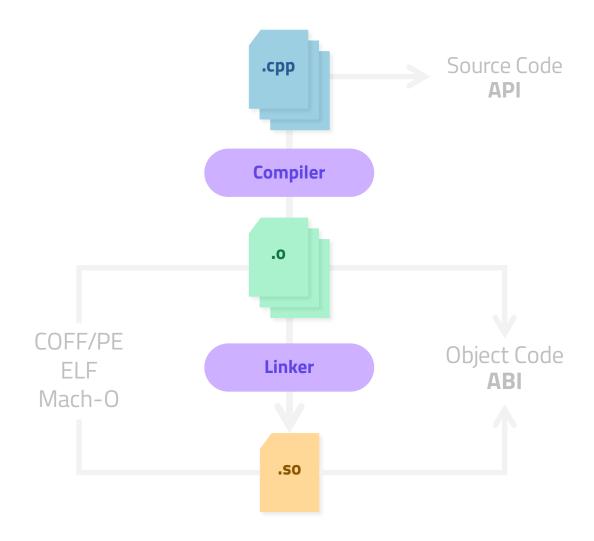


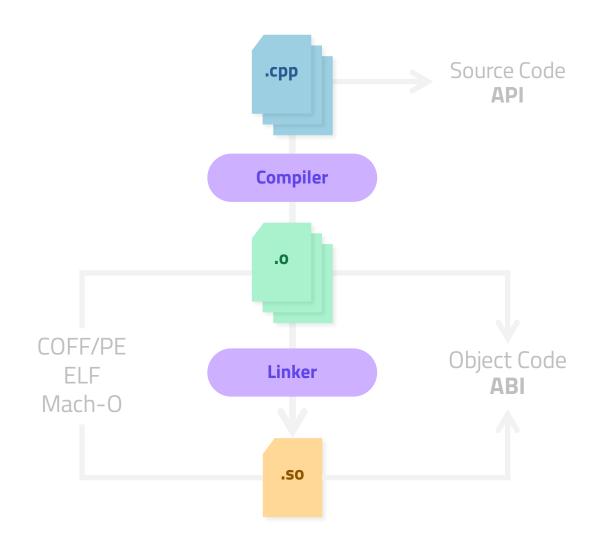
API is the **source code** level interface



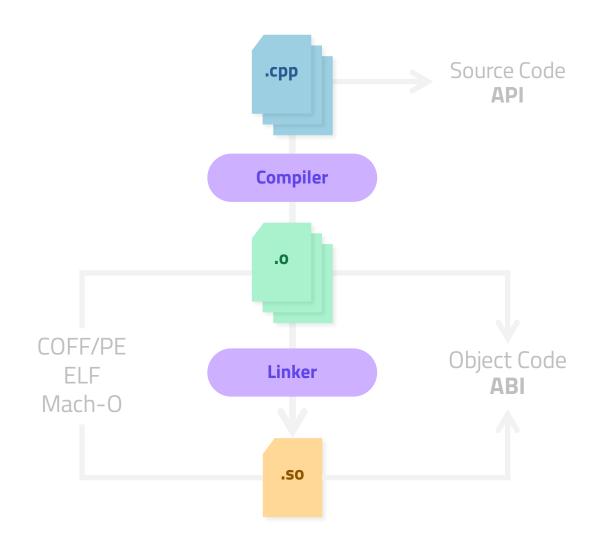
API is the **source code** level interface

ABI is the **object code** level interface



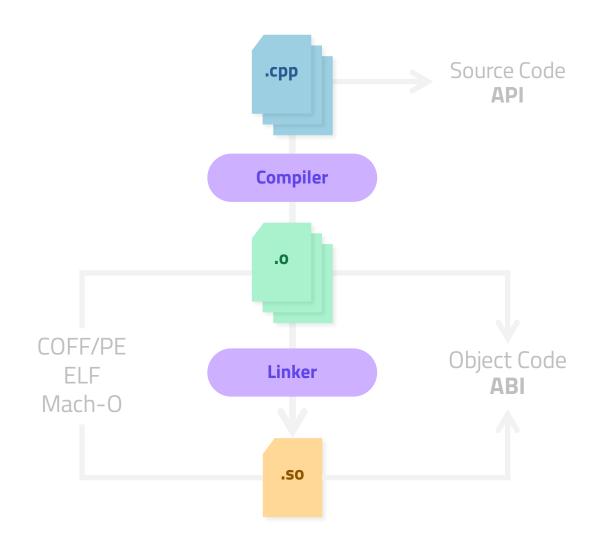


What's inside object files?



What's inside object files?

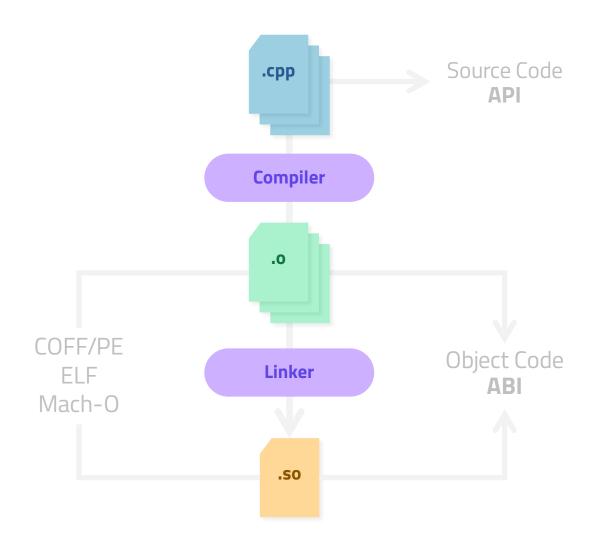
Machine Code



What's inside object files?

Machine Code

Static Data

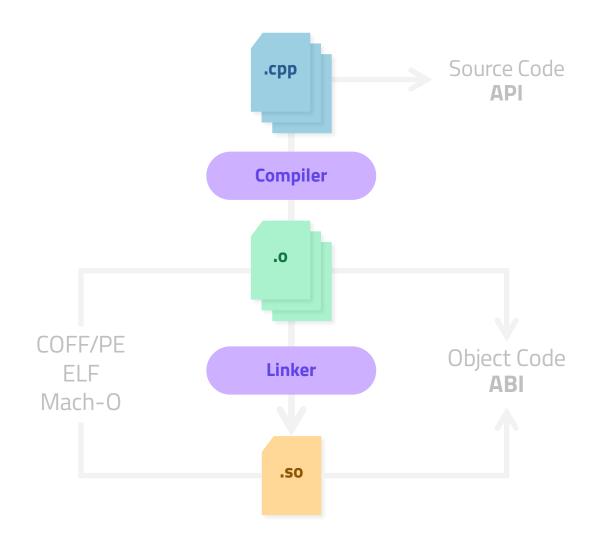


What's inside object files?

Machine Code

Static Data

Symbol Table



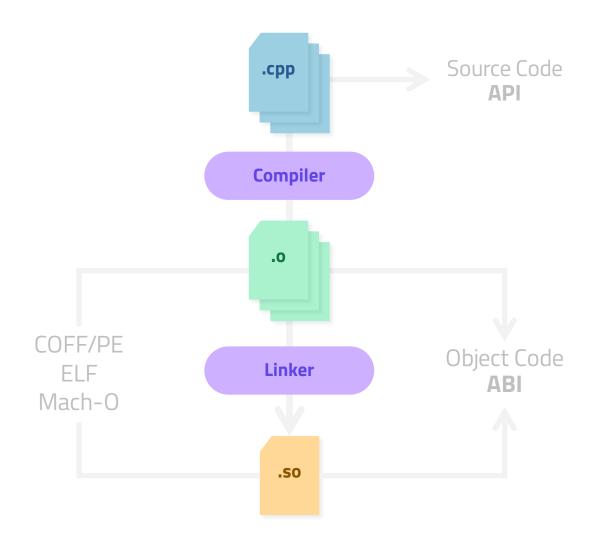
What's inside object files?

Machine Code

Static Data

Symbol Table

linux \$ nm --extern-only lib.so | c++filt



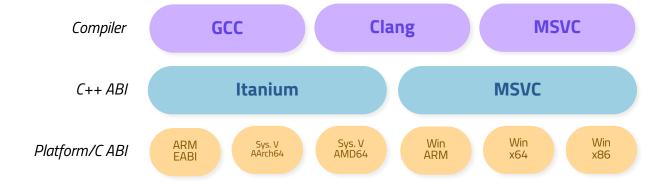
What's inside object files?

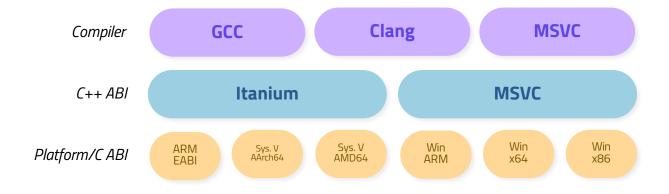
Machine Code

Static Data

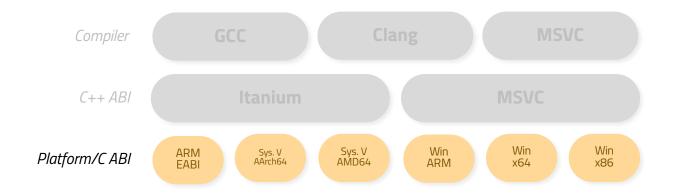
Symbol Table

```
linux $ nm --extern-only lib.so | c++filt
windows $ dumpbin /exports lib.dll
```



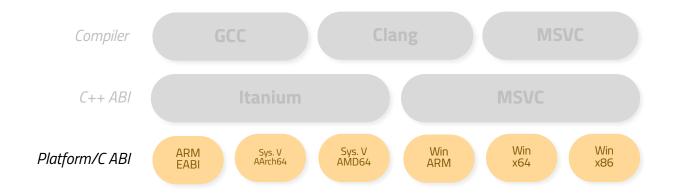


The Layers of ABI



The Layers of ABI

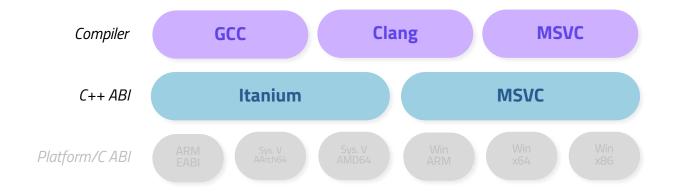
Processor Architecture



The Layers of ABI

Processor Architecture

Operating System

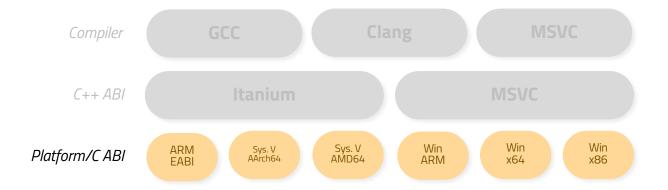


The Layers of ABI

Processor Architecture

Operating System

Compiler



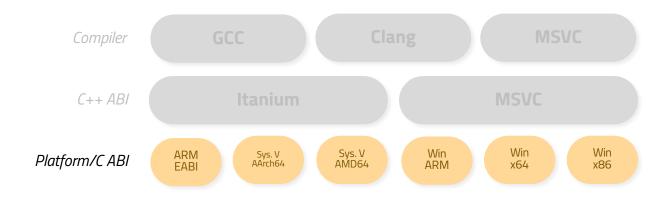


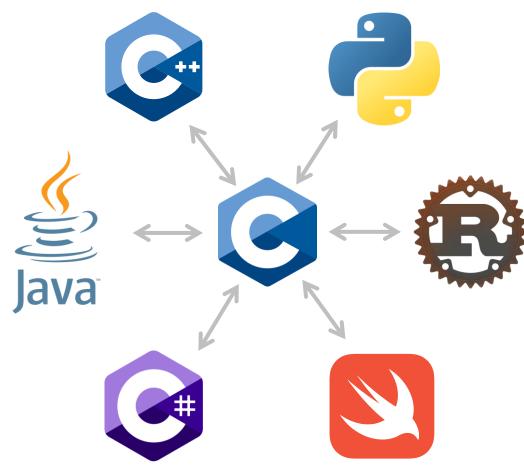
Platform/C ABI is the lowest common denominator



Platform/C ABI is the lowest common denominator







Python/C

Python/C

```
01
02 import ctypes
03
04 libc = ctypes.cdll.LoadLibrary("libc.so.6")
05
06 libc.printf.restype = ctypes.c_int
  libc.printf.argtypes = [
08
       ctypes.c_char_p,
      ctypes.c_char_p,
09
10
11
12 libc.printf(b"Hello, %S\n", "World!")
13
14
```

Python ctypes

Python/C

```
01
02 import ctypes
03
04 libc = ctypes.cdll.LoadLibrary("libc.so.6")
05
06 libc.printf.restype = ctypes.c_int
   libc.printf.argtypes = [
08
       ctypes.c_char_p,
       ctypes.c_char_p,
09
10 7
11
12 libc.printf(b"Hello, %S\n", "World!")
13
14
```

Python ctypes

Load Shared Libraries

```
01
02 import ctypes
03
04 libc = ctypes.cdll.LoadLibrary("libc.so.6")
05
06 libc.printf.restype = ctypes.c_int
   libc.printf.argtypes = [
08
       ctypes.c_char_p,
       ctypes.c_char_p,
09
10 7
11
12 libc.printf(b"Hello, %S\n", "World!")
13
14
```

Python ctypes

Basic C Types

```
01
02 import ctypes
03
04 libc = ctypes.cdll.LoadLibrary("libc.so.6")
05
06 libc.printf.restype = ctypes.c_int
  libc.printf.argtypes = [
08
       ctypes.c_char_p,
       ctypes.c_char_p,
09
10
11
12 libc.printf(b"Hello, %S\n", "World!")
13
14
```

Python ctypes

C-like Function Signatures & Calls

```
01
03 import ctypes
04
   class Point(ctypes.Structure):
       _fields_ = [
06
           ("x", c_int),
          ("y", c_int),
10
11 point = Point(10, 20)
14
```

Python ctypes

```
01
03 import ctypes
04
   class Point(ctypes.Structure):
       _fields_ = [
06
           ("x", c_int),
          ("y", c_int),
09
10
11 point = Point(10, 20)
14
```

Python ctypes

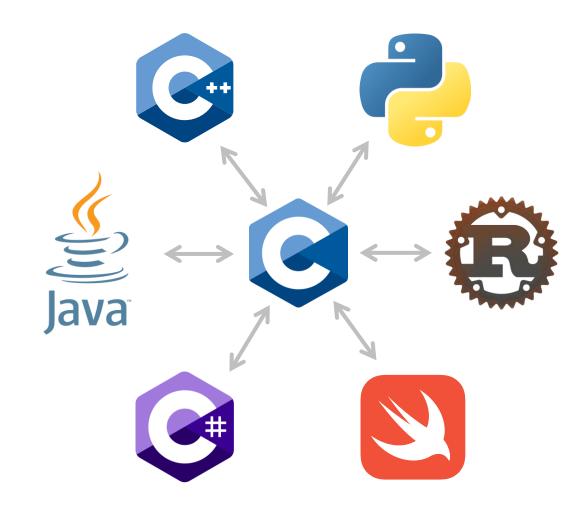
C-like Structure Types

```
01
03 import ctypes
04
  class Point(ctypes.Structure):
      _fields_ = [
    ("x", c_int),
     ("y", c_int),
10
11 point = Point(10, 20)
14
```

Python ctypes

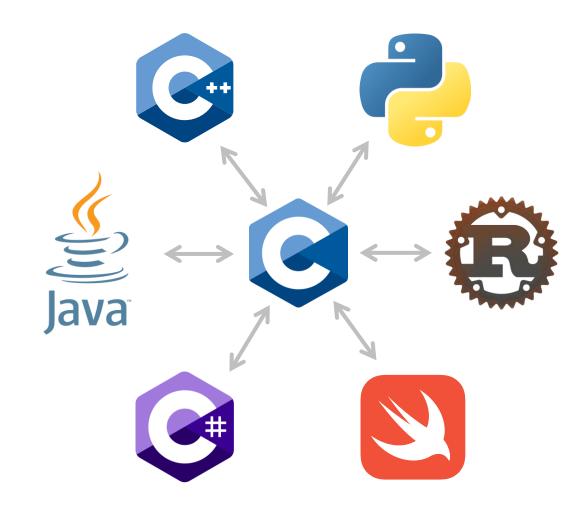
Most languages have similar APIs

Platform/C ABI



Platform/C ABI

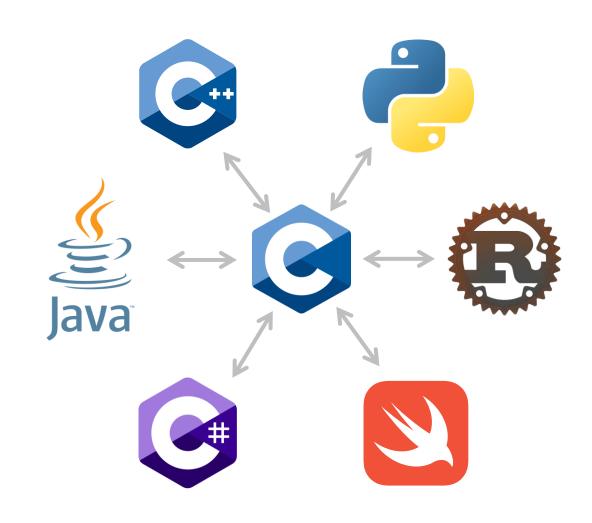
A good foundation for cross-lang. interop.



Platform/C ABI

A good foundation for cross-lang. interop.

Calling Convention & Type Layout



Platform/C ABI: Calling Convention



Calling Convention

Calling Convention

Caller & Callee Saved Registers

Calling Convention

Caller & Callee Saved Registers

Passing Parameters & Return Values

Calling Convention

Caller & Callee Saved Registers

Passing Parameters & Return Values

Stack Frame Layout

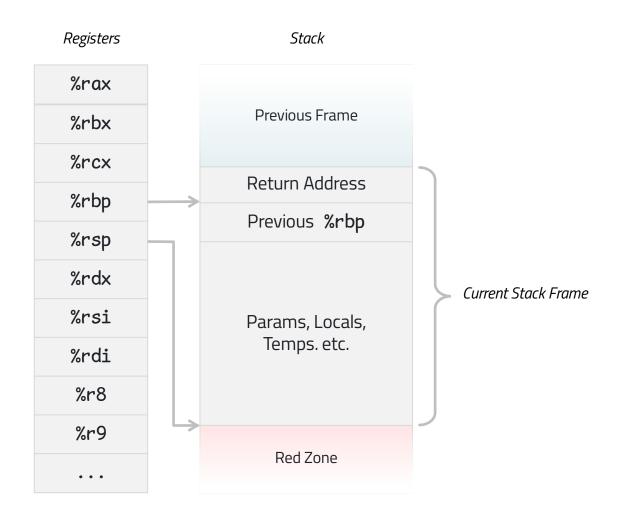
Registers

%rax %rbx %rcx %rbp %rsp %rdx %rsi %rdi %r8 %r9

Stack

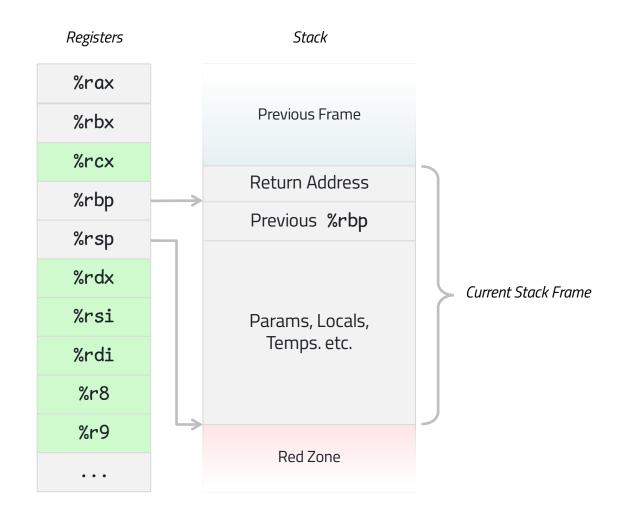
Previous Frame

System V AMD64 ABI



System V AMD64 ABI

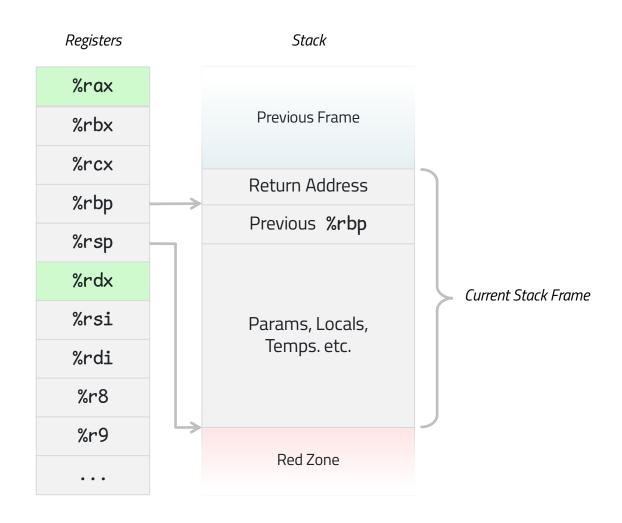
Push Return Address & Base-Pointer on Stack



System V AMD64 ABI

Push Return Address & Base-Pointer on Stack

Parameters in %rdi %rsi %rdx %rcx %r8 %r9*

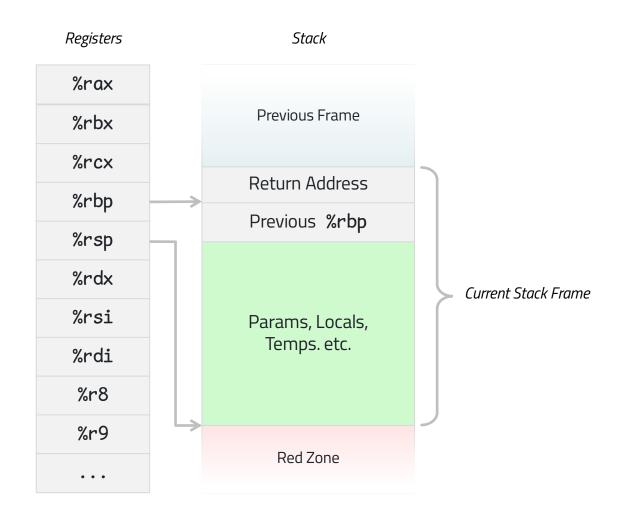


System V AMD64 ABI

Push Return Address & Base-Pointer on Stack

Parameters in %rdi %rsi %rdx %rcx %r8 %r9*

Return Values in %rax %rdx**



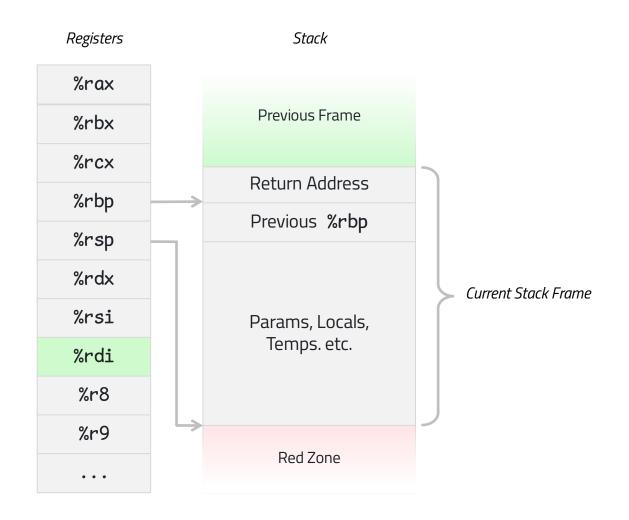
System V AMD64 ABI

Push Return Address & Base-Pointer on Stack

Parameters in %rdi %rsi %rdx %rcx %r8 %r9*

Return Values in %rax %rdx**

*Large params spill onto stack



System V AMD64 ABI

Push Return Address & Base-Pointer on Stack

Parameters in %rdi %rsi %rdx %rcx %r8 %r9*

Return Values in %rax %rdx**

*Large params spill onto stack

**Large return values & non-trivial types by ref.

C++ ABI: Non-trivial

C++ ABI: Non-trivial

Non-trivial for the purposes of calls*

C++ ABI: Non-trivial

Non-trivial for the purposes of calls*

*This is different from std::is_trivial, which is deprecated in C++26 (P3247).

C++ ABI: Non-trivial

Non-trivial *for the purposes of calls**

Non-trivial copy-constructor, move-constructor, or destructor

*This is different from **std::is_trivial**, which is deprecated in C++26 (P3247).

C++ ABI: Non-trivial

Non-trivial *for the purposes of calls**

Non-trivial copy-constructor, move-constructor, or destructor

Or all copy and move constructors are deleted

*This is different from **std::is_trivial**, which is deprecated in C++26 (P3247).

```
01
03
04
05 #ifndef _MSC_VER
     define __stdcall __attribute__((stdcall))
07 #endif
08
09 extern "C" void __stdcall baz();
10
14
```

C++ ABI: Calling Convention

```
01
03
04
05 #ifndef _MSC_VER
     define __stdcall __attribute__((stdcall))
07 #endif
08
09 extern "C" void __stdcall baz();
10
14
```

C++ ABI: Calling Convention

No Calling Convention in C++ Standard

```
01
03
04
05 #ifndef _MSC_VER
     define __stdcall __attribute__((stdcall))
07 #endif
08
09 extern "C" void __stdcall baz();
10
14
```

C++ ABI: Calling Convention

No Calling Convention in C++ Standard

Implementation defined keywords/attributes

```
01
03
04
05 #ifndef _MSC_VER
     define __stdcall __attribute__((stdcall))
07 #endif
08
09 extern "C" void __stdcall baz();
10
14
```

C++ ABI: Calling Convention

No Calling Convention in C++ Standard

Implementation defined keywords/attributes
__stdcall

```
01
03
04
05 #ifndef _MSC_VER
     define __stdcall __attribute__((stdcall))
07 #endif
08
09 extern "C" void __stdcall baz();
10
14
```

C++ ABI: Calling Convention

No Calling Convention in C++ Standard

Implementation defined keywords/attributes

```
__stdcall
```

__attribute__((stdcall))

```
01
03 extern "C" void foo();
04 void foo();
05
06 namespace bar {
       extern "C" void foo();
       void foo();
09 }
10
11 extern "C" foo(std::string str);
12 extern "C" bar(int &num);
13
14
```

C++ ABI: Language Linkage

```
01
03 extern "C" void foo();
04 void foo();
05
06 namespace bar {
       extern "C" void foo();
      void foo();
09 }
11 extern "C" foo(std::string str);
12 extern "C" bar(int &num);
13
14
```

C++ ABI: Language Linkage

extern "C" is a language linkage specification

```
01
03 extern "C" void foo();
04 void foo();
05
06 namespace bar {
       extern "C" void foo();
      void foo();
09 }
10
11 extern "C" foo(std::string str);
12 extern "C" bar(int &num);
13
14
```

C++ ABI: Language Linkage

extern "C" is a language linkage specification

Names refer to the same entity

```
01
03 extern "C" void foo();
04 void foo();
05
06 namespace bar {
     extern "C" void foo();
     void foo();
09 }
10
11 extern "C" foo(std::string str);
12 extern "C" bar(int &num);
13
14
```

C++ ABI: Language Linkage

extern "C" is a language linkage specification

Names refer to the same entity

No effect on calling convention

```
01
03 extern "C" void foo();
04 void foo();
05
06 namespace bar {
      extern "C" void foo();
       void foo();
09 }
10
11 extern "C" foo(std::string str);
12 extern "C" bar(int &num);
13
14
```

C++ ABI: Language Linkage

extern "C" is a language linkage specification

Names refer to the same entity

No effect on calling convention

C++ language linkage is the default

Calling Convention

Calling Convention

Platforms usually have a default

Calling Convention

Platforms usually have a default

Default is usually denoted *C* or **cdec1**

Calling Convention

Platforms usually have a default

Default is usually denoted *C* or cdecl

Windows x86 has stdcall, fastcall, thiscall

Platform/C ABI: Type Layout

Type Layout

Few guarantees in the C Standard

Type Layout

Few guarantees in the C Standard

A char is 1 byte (CHAR_BIT bits) with weakest (usually 1) alignment

Type Layout

Few guarantees in the C Standard

A char is 1 byte (CHAR_BIT bits) with weakest (usually 1) alignment

struct members are laid out in declaration order

Type Layout

Few guarantees in the C Standard

A char is 1 byte (CHAR_BIT bits) with weakest (usually 1) alignment

struct members are laid out in declaration order

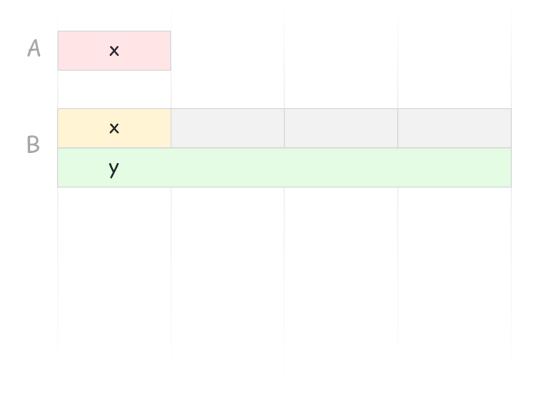
structs can have padding in the middle or at the end

```
01
03
04 struct A { char x; };
05
07 struct B { char x; int y; };
08
09
10 struct C { char x; int y; char z; };
11
14
```

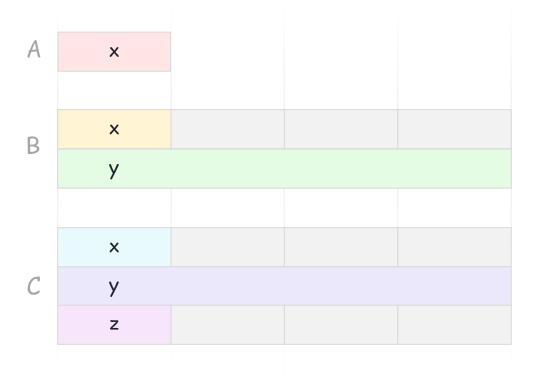
```
01
03
04 struct A { char x; };
05
07 struct B { char x; int y; };
08
09
10 struct C { char x; int y; char z; };
11
14
```



```
01
03
04 struct A { char x; };
05
07 struct B { char x; int y; };
08
09
10 struct C { char x; int y; char z; };
11
14
```



```
01
03
04 struct A { char x; };
05
07 struct B { char x; int y; };
08
09
10 struct C { char x; int y; char z; };
11
14
```



```
01
03
04 struct A { char x; };
05
07 struct B { char x; int y; };
08
09
10 struct C { char x; int y; char z; };
11
14
```

```
01
03
04 struct A { char x; };
05 static_assert(sizeof(struct A) == 1);
06
07 struct B { char x; int y; };
08 static_assert(sizeof(struct B) == 8);
09
10 struct C { char x; int y; char z; };
11 static_assert(sizeof(struct C) == 12);
12
14
```

Type Layout

Use **static_assert** to guarantee type layout

```
01
03
04
05 #include <stddef.h>
06
07 struct B { char x; int y; };
08 static_assert(offsetof(struct B, y) == 4);
09
14
```

Type Layout

Use **static_assert** to guarantee type layout

Use **offsetof** for member offsets

```
01 #include <type_traits>
02
03 struct A { int m; };
04 static_assert(std::is_standard_layout_v<A> == true);
05
07
09
14
```

C++ Standard-Layout Types

```
01 #include <type_traits>
02
03 struct A { int m; };
04 static_assert(std::is_standard_layout_v<A> == true);
05
07
09
14
```

C++ Standard-Layout Types

Standard-layout classes are useful for communication with code written in other programming languages. – [class.prop]p6

```
01 #include <type_traits>
02
03 struct A { int m; };
04 static_assert(std::is_standard_layout_v<A> == true);
05
06
07
09
10
14
```

C++ Standard-Layout Types

Standard-layout classes are useful for communication with code written in other programming languages. – [class.prop]p6

All NSDMs are also standard-layout

```
01 #include <type_traits>
02
03 struct A { int m; };
04 static_assert(std::is_standard_layout_v<A> == true);
05
06 class B: public A { int m; };
  static_assert(std::is_standard_layout_v<B> == false);
08
09
10
13
14
```

C++ Standard-Layout Types

Standard-layout classes are useful for communication with code written in other programming languages. – [class.prop]p6

All NSDMs are also standard-layout

All NSDMs are in the same base class.

```
01 #include <type_traits>
02
03 struct A { int m; };
04 static_assert(std::is_standard_layout_v<A> == true);
05
06 class B: public A { int m; };
07 static_assert(std::is_standard_layout_v<B> == false);
08
09 class C { int m; public: int n; };
10 static_assert(std::is_standard_layout_v<C> == false);
11
13
14
```

C++ Standard-Layout Types

Standard-layout classes are useful for communication with code written in other programming languages. – [class.prop]p6

All NSDMs are also standard-layout

All NSDMs are in the same base class

All NSDMs have the same access control*

```
01 #include <type_traits>
02
03 struct A { int m; };
04 static_assert(std::is_standard_layout_v<A> == true);
05
06 class B: public A { int m; };
07 static_assert(std::is_standard_layout_v<B> == false);
08
09 class C { int m; public: int n; };
10 static_assert(std::is_standard_layout_v<C> == false);
11
13
14
```

C++ Standard-Layout Types

Standard-layout classes are useful for communication with code written in other programming languages. – [class.prop]p6

All NSDMs are also standard-layout

All NSDMs are in the same base class

All NSDMs have the same access control*

*Re-ordering based on access control is removed in C++23 (P1847), but this requirement remains for standard-layout...

```
01 #include <type_traits>
02
03 struct A { int m; };
04 static_assert(std::is_standard_layout_v<A> == true);
05
06 class B: public A { int m; };
07 static_assert(std::is_standard_layout_v<B> == false);
08
09 class C { int m; public: int n; };
10 static_assert(std::is_standard_layout_v<C> == false);
11
12 struct D { virtual void foo(); };
13 static_assert(std::is_standard_layout_v<D> == false);
14
```

C++ Standard-Layout Types

Standard-layout classes are useful for communication with code written in other programming languages. – [class.prop]p6

All NSDMs are also standard-layout

All NSDMs are in the same base class

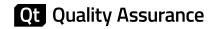
All NSDMs have the same access control*

No virtual functions or virtual base classes

*Re-ordering based on access control is removed in C++23 (P1847), but this requirement remains for standard-layout...



Can we do better than a C API?



Can we do better than a C API?

Need to define APIs in terms of C

Can we do better than a C API?

Need to define APIs in terms of C

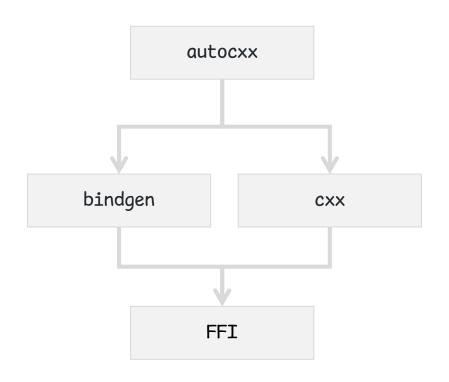
Remove higher-level abstractions

Can we do better than a C API?

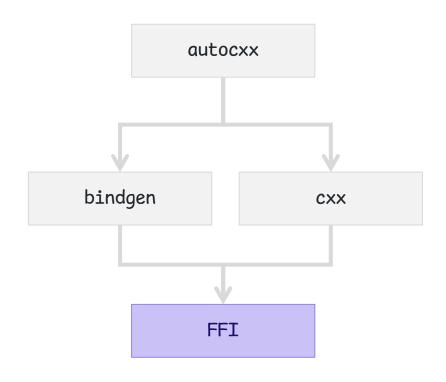
Need to define APIs in terms of C

Remove higher-level abstractions

Case Study: Rust/C++



Rust/C++ Ecosystem



Built-in C-like FFI

```
01 use std::os::raw::*;
02
03 #[repr(C)]
04 struct Point {
      x: c_int,
      y: c_int,
07 }
08
09 unsafe extern "C" {
       fn dist(first: Point, second: Point) -> c_float;
11 }
13
14
```

Rust FFI

```
01 use std::os::raw::*;
02
03 #[repr(C)]
04 struct Point {
       x: c_int,
      y: c_int,
07 }
08
09 unsafe extern "C" {
       fn dist(first: Point, second: Point) -> c_float;
11 }
13
14
```

Rust FFI

Basic C types

```
01 use std::os::raw::*;
02
03 #[repr(C)]
04 struct Point {
      x: c_int,
      y: c_int,
07 }
08
09 unsafe extern "C" {
       fn dist(first: Point, second: Point) -> c_float;
11 }
13
14
```

Rust FFI

Basic C types

C-like structure types

```
01 use std::os::raw::*;
02
03 #[repr(C)]
04 struct Point {
      x: c_int,
      y: c_int,
07 }
08
09 unsafe extern "C" {
       fn dist(first: Point, second: Point) -> c_float;
11 }
13
14
```

Rust FFI

Basic C types

C-like structure types

C-like functions decls.

```
01 int *foo() {
02    return nullptr;
03 }
```

```
04 unsafe extern "C" {
05     fn foo() -> Option<&i32>;
06 }
07
08 unsafe fn bar() {
09     if let Some(it) = foo() {
10         println!("{it}");
11     }
12 }
13
```

Rust FFI

```
01 int *foo() {
02    return nullptr;
03 }
```

```
04 unsafe extern "C" {
05     fn foo() -> Option<&i32>;
06 }
07
08 unsafe fn bar() {
09     if let Some(it) = foo() {
10         println!("{it}");
11     }
12 }
13
```

Rust FFI

Nullable Pointer Optimization

```
01 int *foo() {
02    return nullptr;
03 }
```

```
04 unsafe extern "C" {
05     fn foo() -> Option<&i32>;
06 }
07
08 unsafe fn bar() {
09     if let Some(it) = foo() {
10         println!("{it}");
11     }
12 }
13
```

Rust FFI

Nullable Pointer Optimization

Option::None has same repr. as nullptr

```
01 int *foo() {
02   return nullptr;
03 }
```

```
04 unsafe extern "C" {
05     fn foo() -> Option<&i32>;
06 }
07
08 unsafe fn bar() {
09     if let Some(it) = foo() {
10         println!("{it}");
11     }
12 }
13
```

Rust FFI

Nullable Pointer Optimization

Option::None has same repr. as nullptr

Idiomatic handling in Rust

```
01 void foo() try {
       bar();
03 } catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
      unsafe { baz(); }
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
13 }
```

Rust FFI: Unwind ABIs

```
01 void foo() try {
02    bar();
03 } catch (const std:: runtime_error& e) {
04    std::println("{}", e.what());
05 }
```

```
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
08    let n = Box::new(42);
09    unsafe { baz(); }
10 }
```

```
11 void baz() {
12     throw std::runtime_error("Hello from Rust!");
13 }
```

Rust FFI

Rust has "-unwind" ABIs

```
01 void foo() try {
       bar();
    catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
04
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
      unsafe { baz(); }
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
13 }
```

Rust FFI

Rust has "-unwind" ABIs

C++ exceptions can passthrough Rust

```
01 void foo() try {
       bar();
03 } catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
04
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
08
      unsafe { baz(); }
09
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
13 }
```

Rust FFI

Rust has "-unwind" ABIs

C++ exceptions can passthrough Rust

Correct cleanup in Rust

```
01 void foo() try {
       bar();
03 } catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
04
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
08
09
      unsafe { baz(); }
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
13 }
```

Rust FFI

Rust has "-unwind" ABIs

C++ exceptions can passthrough Rust

Correct cleanup in Rust

Cannot catch foreign exceptions

```
01 void foo() try {
       bar();
03 } catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
04
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
      unsafe { baz(); }
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

```
01 void foo() try {
02    bar();
03 } catch (const std:: runtime_error& e) {
04    std::println("{}", e.what());
05 }
06 #[no_mangle]
```

```
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
08    let n = Box::new(42);
09    unsafe { baz(); }
10 }
```

```
11 void baz() {
12    throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

Unwind Tables

```
01 void foo() try {
       bar();
03 } catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
04
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
      unsafe { baz(); }
09
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

Unwind Tables

```
01 void foo() try {
02    bar();
03 } catch (const std:: runtime_error& e) {
04    std::println("{}", e.what());
05 }
06 #[no_mangle]
```

```
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
08    let n = Box::new(42);
09    unsafe { baz(); }
10 }
```

```
11 void baz() {
12    throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

Unwind Tables

Base Exception ABI

```
01 void foo() try {
02    bar();
03 } catch (const std:: runtime_error& e) {
04    std::println("{}}", e.what());
05 }

06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
08    let n = Box::new(42);
09    unsafe { baz(); }
10 }
```

```
11 void baz() {
12    throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

Unwind Tables

Base Exception ABI

Unwinds the stack using unwind tables: libunwind

```
01 void foo() try {
       bar();
03 } catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
04
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
      unsafe { baz(); }
09
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
```

Exception ABI

Unwind Tables

Base Exception ABI

Unwinds the stack using unwind tables: libunwind

13 }

```
01 void foo() try {
02    bar();
03 } catch (const std:: runtime_error& e) {
04    std::println("{}", e.what());
05 }
```

```
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
08    let n = Box::new(42);
09    unsafe { baz(); }
10 }
```

```
11 void baz() {
12     throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

Unwind Tables

Base Exception ABI

Unwinds the stack using unwind tables: libunwind

```
01 void foo() try {
02    bar();
03 } catch (const std:: runtime_error& e) {
04    std::println("{}", e.what());
05 }
```

```
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
08    let n = Box::new(42);
09    unsafe { baz(); }
10 }
```

```
11 void baz() {
12     throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

Unwind Tables

Base Exception ABI

Unwinds the stack using unwind tables: libunwind

Language Exception ABI

```
01 void foo() try {
       bar();
03 } catch (const std:: runtime_error& e) {
       std::println("{}", e.what());
04
05 }
06 #[no_mangle]
07 pub extern "C-unwind" fn bar() {
      let n = Box::new(42);
08
      unsafe { baz(); }
09
10 }
11 void baz() {
       throw std::runtime_error("Hello from Rust!");
13 }
```

Exception ABI

Unwind Tables

Personality Routine

Personality Routine

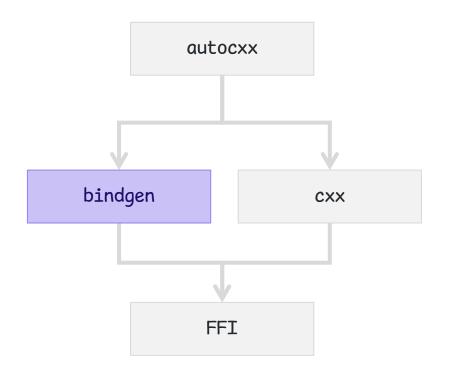
Base Exception ABI

Unwinds the stack using unwind tables: libunwind

Language Exception ABI

Personality routine

Rust/C++: bindgen



Automated binding generator

```
01 struct Point {
02    int x;
03    int y;
04 };
```

```
06
07 #[repr(C)]
08 #[derive(Debug, Copy, Clone)]
09 pub struct Point {
10    pub x: c_int,
11    pub y: c_int,
12 }
13
```

bindgen

```
01 struct Point {
02    int x;
03    int y;
04 };
```

```
06
07 #[repr(C)]
08 #[derive(Debug, Copy, Clone)]
09 pub struct Point {
10    pub x: c_int,
11    pub y: c_int,
12 }
13
```

bindgen

Built with libclang

```
01 struct Point {
02    int x;
03    int y;
04 };
```

```
06
07 #[repr(C)]
08 #[derive(Debug, Copy, Clone)]
09 pub struct Point {
10    pub x: c_int,
11    pub y: c_int,
12 }
13
```

bindgen

Built with libclang

Auto generates C/C++-Rust bindings

```
01 struct Point {
02    int x;
03    int y;
04 };
```

```
06
07 #[repr(C)]
08 #[derive(Debug, Copy, Clone)]
09 pub struct Point {
10    pub x: c_int,
11    pub y: c_int,
12 }
13
```

bindgen

Built with libclang

Auto generates C/C++-Rust bindings

Command line tool or library for build integration

```
01 struct Point {
02    int x;
03    int y;
04 };
```

```
06
07 #[repr(C)]
08 #[derive(Debug, Copy, Clone)]
09 pub struct Point {
10    pub x: c_int,
11    pub y: c_int,
12 }
13
```

bindgen

Built with libclang

Auto generates C/C++-Rust bindings

Command line tool or library for build integration

```
$ bindgen foo.hpp > out.rs
```

```
01
03 const _: () = {
      ["Size of Point"]
          [::std::mem::size_of::<Point>() - 8usize];
      ["Alignment of Point"]
          [::std::mem::align_of::<Point>() - 4usize];
      ["Offset of field: Point::x"]
          [::std::mem::offset_of!(Point, x) - Ousize];
      ["Offset of field: Point::y"]
           [::std::mem::offset_of!(Point, y) - 4usize];
08 };
09
10
```

bindgen

```
01
03 const _: () = {
      ["Size of Point"]
          [::std::mem::size_of::<Point>() - 8usize];
      ["Alignment of Point"]
          [::std::mem::align_of::<Point>() - 4usize];
      ["Offset of field: Point::x"]
          [::std::mem::offset_of!(Point, x) - Ousize];
      ["Offset of field: Point::y"]
           [::std::mem::offset_of!(Point, y) - 4usize];
08 };
09
10
```

bindgen

Static assertions for size and offset

```
01 int foo(char x);
02 int foo(char x, int y);
```

```
03
04
05 unsafe extern "C" {
       #[link_name = "_Z3fooc"]
       pub fn foo(x: c_char) -> c_int;
08
       #[link_name = "_Z3fooci"]
09
       pub fn foo1(x: c_char, y: c_int) -> c_int;
10
11 }
13
```

bindgen: Function Overloading

```
01 int foo(char x);
02 int foo(char x, int y);
03
04
05 unsafe extern "C" {
       #[link_name = "_Z3fooc"]
06
       pub fn foo(x: c_char) -> c_int;
08
       #[link_name = "_Z3fooci"]
09
       pub fn foo1(x: c_char, y: c_int) -> c_int;
11 }
12
13
```

bindgen: Function Overloading

Auto-generates mangled names

```
01 int foo(char x);
02 int foo(char x, int y);
03
04
05 unsafe extern "C" {
       #[link_name = "_Z3fooc"]
       pub fn foo(x: c_char) -> c_int;
08
       #[link_name = "_Z3fooci"]
09
       pub fn foo1(x: c_char, y: c_int) -> c_int;
11 }
12
13
```

bindgen: Function Overloading

Note **foo1**: No overloading in Rust

```
01 template<typename T>
02 struct 5 {
03    T value;
04 };
05 void foo(S<int> s);
```

```
06 #[repr(C)]
07 pub struct S<T> {
08     pub value: T,
09 }
10 unsafe extern "C" {
11     #[link_name = "_Z3foo1SIiE"]
12     pub fn foo(s: S<c_int>);
13 }
```

bindgen: Templates

```
01 template<typename T>
02 struct S {
03   T value;
04 };
05 void foo(S<int> s);
```

```
06 #[repr(C)]
07 pub struct S<T> {
08     pub value: T,
09 }
10 unsafe extern "C" {
11     #[link_name = "_Z3foo1SIiE"]
12     pub fn foo(s: S<c_int>);
13 }
```

bindgen: Templates

Simple Templates use Rust Generics

```
01 template<typename T>
02 struct S {
03   T value;
04 };
05 void foo(S<int> s);
```

```
06 #[repr(C)]
07 pub struct S<T> {
08     pub value: T,
09 }
10 unsafe extern "C" {
11     #[link_name = "_Z3foo1SIiE"]
12     pub fn foo(s: S<c_int>);
13 }
```

bindgen: Templates

Simple Templates use Rust Generics

Correct Usage as Parameter

```
01 template<typename T>
02 struct S {
03   T value;
04 };
05 void foo(S<int> s);
```

```
06 #[repr(C)]
07 pub struct S<T> {
08     pub value: T,
09 }
10 unsafe extern "C" {
11     #[link_name = "_Z3foo1SIiE"]
12     pub fn foo(s: S<c_int>);
13 }
```

bindgen: Templates

Simple Templates use Rust Generics

```
01 template<typename T>
02 struct S {
03   T value;
04 };
05 void foo(S<int> s);
```

```
06 #[repr(C)]
07 pub struct S<T> {
08     pub value: T,
09 }
10 unsafe extern "C" {
11     #[link_name = "_Z3foo1SIiE"]
12     pub fn foo(s: S<c_int>);
13 }
```

bindgen: Templates

Simple Templates use Rust Generics

No method generation

```
01 template<typename T>
02 struct S {
03   T value;
04 };
05 void foo(S<int> s);
```

```
06 #[repr(C)]
07 pub struct S<T> {
08     pub value: T,
09 }
10 unsafe extern "C" {
11     #[link_name = "_Z3foo1SIiE"]
12     pub fn foo(s: S<c_int>);
13 }
```

bindgen: Templates

Simple Templates use Rust Generics

No method generation

No static assertions generated

01 struct 5 { 02 int i; 03 virtual void foo(); 04 }; 05 06 static_assert(sizeof(S) == 16);

Polymorphic Types

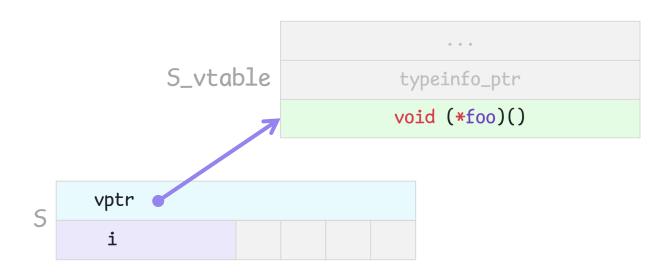
```
S vptr
```

```
01 struct 5 {
02    int i;
03    virtual void foo();
04 };
05
06 static_assert(sizeof(S) == 16);
```

Polymorphic Types

Vptr inserted into layout

C++ ABI



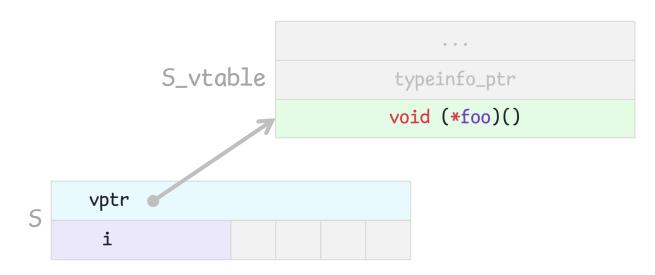
```
01 struct 5 {
02    int i;
03    virtual void foo();
04 };
05
06 static_assert(sizeof(5) == 16);
```

Polymorphic Types

vptr inserted into layout

vtable contains method overrides, RTTI, & more

C++ ABI



```
01 struct 5 {
02    int i;
03    virtual void foo();
04 };
05
06 static_assert(sizeof(S) == 16);
```

Polymorphic Types

vptr inserted into layout

vtable contains method overrides, RTTI, & more

vtable layout & behavior is complicated

```
01
03 struct Base {
       int i;
       virtual void foo();
06 };
07
08 struct Derived : Base {
       int j;
       void foo() override;
11 };
14
```

bindgen: Polymorphic Types

```
01 #[repr(C)]
02 pub struct Base__bindgen_vtable(c_void);
03
04 #[repr(C)]
05 pub struct Base {
       pub vtable_: *const Base__bindgen_vtable,
       pub i: c_int,
08 }
09
10 #[repr(C)]
11 pub struct Derived {
       pub _base: Base,
       pub j: c_int,
14 }
```

bindgen: Polymorphic Types

```
01 #[repr(C)]
02 pub struct Base__bindgen_vtable(c_void);
03
04 #[repr(C)]
05 pub struct Base {
       pub vtable_: *const Base__bindgen_vtable,
      pub i: c_int,
08 }
09
10 #[repr(C)]
11 pub struct Derived {
       pub _base: Base,
       pub j: c_int,
14 }
```

bindgen: Polymorphic Types

Base has a vptr

```
01
03
04 unsafe extern "C" {
       #[link_name = "_ZN4Base3fooEv"]
       pub fn Base_foo(this: *mut c_void);
07
       #[link_name = "_ZN7Derived3fooEv"]
       pub fn Derived_foo(this: *mut c_void);
09
10 }
14
```

bindgen: Polymorphic Types

```
01
03
  unsafe extern "C" {
       #[link_name = "_ZN4Base3fooEv"]
       pub fn Base_foo(this: *mut c_void);
07
       #[link_name = "_ZN7Derived3fooEv"]
       pub fn Derived_foo(this: *mut c_void);
09
10 }
14
```

bindgen: Polymorphic Types

Pure FFI functions; no methods

```
01
03
04 unsafe extern "C" {
       #[link_name = "_ZN4Base3fooEv"]
       pub fn Base_foo(this: *mut c_void);
07
       #[link_name = "_ZN7Derived3fooEv"]
       pub fn Derived_foo(this: *mut c_void);
09
10 }
14
```

bindgen: Polymorphic Types

Raw void pointers

```
01
03
04
05 fn main() {
      let mut d = get_derived();
      unsafe {
           Derived_foo(&mut d as *mut _ as *mut c_void);
10 }
14
```

bindgen: Polymorphic Types

```
01
03
04
05 fn main() {
      let mut d = get_derived();
      unsafe {
           Derived_foo(&mut d as *mut _ as *mut c_void);
10 }
14
```

bindgen: Polymorphic Types



Unsafe... and a bit messy 🙉

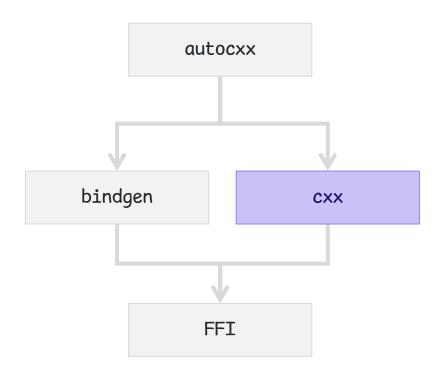


```
01
03
04
05 fn main() {
      let mut d = get_derived();
      unsafe {
           Derived_foo(&mut d as *mut _ as *mut c_void);
10 }
14
```

bindgen: Polymorphic Types

bindgen only tries to get the layout correct

Rust/C++: cxx



Automated bindings + bridged types

CXX

CXX

Bindings for types like **std::string** and **std::vector**

CXX

Bindings for types like std::string and std::vector

Special consideration given to ownership

CXX

Bindings for types like std::string and std::vector

Special consideration given to ownership

Patterns for error handling

```
01
03
04 #[cxx::bridge]
05 mod ffi {
       unsafe extern "C++" {
           include!("foo.hpp");
          type Foo;
           fn new_foo() -> UniquePtr<Foo>;
11 }
14
```

CXX

```
01
03
04 #[cxx::bridge]
05 mod ffi {
      unsafe extern "C++" {
           include!("foo.hpp");
           type Foo;
           fn new_foo() -> UniquePtr<Foo>;
11 }
14
```

CXX

Include C/C++ files

```
01
03
04 #[cxx::bridge]
05 mod ffi {
      unsafe extern "C++" {
           include!("foo.hpp");
          type Foo;
           fn new_foo() -> UniquePtr<Foo>;
11 }
14
```

CXX

Declare types & functions

```
01
03
04 fn foo(s: String) {}
05
06 fn main() {
      let s = "Hello, Meeting C++".to_string();
      foo(s); // Move
       foo(s); // Error: use of moved value: `s`
10 }
14
```

Move Semantics

```
01
03
04 fn foo(s: String) {}
05
06 fn main() {
      let s = "Hello, Meeting C++".to_string();
      foo(s); // Move
       foo(s); // Error: use of moved value: `s`
10 }
14
```

Move Semantics

Rust has destructive moves

```
01
03
04 fn foo(s: String) {}
05
06 fn main() {
      let s = "Hello, Meeting C++".to_string();
       foo(s); // Move
       foo(s); // Error: use of moved value: `s`
10 }
14
```

Move Semantics

No destructor call for s

```
01
03
04 void foo(std::string s) {}
05
06 int main() {
      auto s = std::string("Hello, Meeting C++");
     foo(std::move(s)); // Move
     foo(s); // Ok
10 } // `s.~S()`
14
```

Move Semantics

```
01
03
04 void foo(std::string s) {}
05
06 int main() {
      auto s = std::string("Hello, Meeting C++");
      foo(std::move(s)); // Move
      foo(s); // Ok
10 } // `s.~5()`
14
```

Move Semantics

C++ leaves s in a valid state

```
01
03
04 void foo(std::string s) {}
05
06 int main() {
      auto s = std::string("Hello, Meeting C++");
      foo(std::move(s)); // Move
     foo(s); // Ok
10 } // `s.~S()`
14
```

Move Semantics

Destructor is still run

cxx: Types

cxx: Types

Shared

cxx: Types

Shared Opaque

cxx: Types

Shared Opaque Extern

cxx: Types

Shared

Opaque

Extern

Visible to both languages

Passed by value

Trivial types

cxx: Types

Shared

Opaque

Extern

Visible to both languages

Visible to one language

Passed by value

Handled behind indirection

Trivial types

Non-trivial types

cxx: Types

Shared

Opaque

Extern

Visible to both languages

Visible to one language

Defined outside of cxx

Passed by value

Handled behind indirection

Opaque or Trivial

Trivial types

Non-trivial types

cxx: Types

Shared

Opaque

Extern

Visible to both languages

Visible to one language

Defined outside of cxx

Passed by value

Handled behind indirection

Opaque or Trivial

Trivial types

Non-trivial types

```
01 struct 5 {
02    ~S();
03 };
04 std::unique_ptr<S> make();
```

```
05 #[cxx::bridge]
06 mod ffi {
07     unsafe extern "C++" {
08         include!("foo.hpp");
09         type S;
10         fn make() -> UniquePtr<S>;
11     }
12 }
```

cxx: Opaque C++ Types

```
01 struct 5 {
02     ~5();
03 };
04 std::unique_ptr<S> make();
```

```
05 #[cxx::bridge]
06 mod ffi {
07     unsafe extern "C++" {
08         include!("foo.hpp");
09         type 5;
10         fn make() -> UniquePtr<S>;
11     }
12 }
13
```

cxx: Opaque C++ Types

Handled behind an indirection

```
01 struct 5 {
02     ~S();
03     int foo() const;
04 };
```

```
05 #[cxx::bridge]
06 mod ffi {
07     unsafe extern "C++" {
08         include!("foo.hpp");
09         type 5;
10         fn foo(self: &5);
11     }
12 }
13
```

cxx: Opaque C++ Types

Declare methods as expected, but...

```
01 struct 5 {
02     ~S();
03     void bar();
04 };
```

```
05 #[cxx::bridge]
06 mod ffi {
07    unsafe extern "C++" {
08         include!("foo.hpp");
09         type 5;
10         fn bar(self: Pin<&mut 5>);
11    }
12 }
```

cxx: Opaque C++ Types

```
01 struct 5 {
02    ~5();
03    void bar();
04 };
```

```
05 #[cxx::bridge]
06 mod ffi {
07    unsafe extern "C++" {
08         include!("foo.hpp");
09         type S;
10         fn bar(self: Pin<&mut S>);
11    }
12 }
```

cxx: Opaque C++ Types

Non-const methods require Pin

```
01 struct 5 {
02    ~S();
03    void bar();
04 };
```

```
05 #[cxx::bridge]
06 mod ffi {
07     unsafe extern "C++" {
08         include!("foo.hpp");
09         type 5;
10         fn bar(self: Pin<&mut S>);
11     }
12 }
13
```

cxx: Opaque C++ Types

&mut T can be used to swap memory

```
01 struct 5 {
02    ~S();
03    void bar();
04 };
```

```
05 #[cxx::bridge]
06 mod ffi {
07    unsafe extern "C++" {
08         include!("foo.hpp");
09         type S;
10         fn bar(self: Pin<&mut S>);
11    }
12 }
13
```

cxx: Opaque C++ Types

&mut T can be used to swap memory

```
01 use std::mem::{swap, take, replace};
01 pub const fn swap<T>(x: &mut T, y: &mut T);
```

cxx: Types

Shared

Opaque

Extern

Visible to both languages

Visible to one language

Defined outside of cxx

Passed by value

Handled behind indirection

Opaque or Trivial

Trivial types

Non-trivial types

```
01 struct 5 {
02     ~5();
03 };
```

```
04 struct S(i32);
05 unsafe impl ExternType for S {
       type Id = cxx::type_id!("S");
       type Kind = cxx::kind::Trivial;
08 }
09 #[cxx::bridge]
10 mod ffi {
      unsafe extern "C++" {
          include!("foo.hpp");
          type S = super::S;
14
15 }
```

cxx: Trivial Extern C++ Types

```
01 struct 5 {
02  ~5();
03 };
```

```
04 struct 5(i32);
05 unsafe impl ExternType for S {
      type Id = cxx::type_id!("S");
      type Kind = cxx::kind::Trivial;
08 }
09 #[cxx::bridge]
10 mod ffi {
     unsafe extern "C++" {
         include!("foo.hpp");
          type S = super::S;
14
15 }
```

cxx: Trivial Extern C++ Types

Existing (bindgen or hand-written) bindings

```
01 struct 5 {
02   ~S();
03 };
```

```
04 struct S(i32);
05 unsafe impl ExternType for S {
      type Id = cxx::type_id!("S");
      type Kind = cxx::kind::Trivial;
08 }
09 #[cxx::bridge]
10 mod ffi {
     unsafe extern "C++" {
          include!("foo.hpp");
          type S = super::S;
14
15 }
```

cxx: Trivial Extern C++ Types

Existing (bindgen or hand-written) bindings

Trivial = subject to Rust move semantics*

```
01 struct 5 {
02  ~5();
03 };
```

```
04 struct S(i32);
05 unsafe impl ExternType for S {
      type Id = cxx::type_id!("5");
      type Kind = cxx::kind::Trivial;
08 }
09 #[cxx::bridge]
10 mod ffi {
     unsafe extern "C++" {
          include!("foo.hpp");
    type S = super::S;
14
15 }
```

cxx: Trivial Extern C++ Types

Existing (bindgen or hand-written) bindings

Trivial = subject to Rust move semantics*

*C++26 introduces **std:**:is_trivially_relocatable (P2786)

```
01 fn foo() -> Result<(), String> {
       Err("Rust Error".to_string())
03 }
04
05 #[cxx::bridge]
06 mod ffi {
       extern "Rust" {
           fn foo() -> Result<()>;
      unsafe extern "C++" {
           fn bar() -> Result<()>;
13 }
14
```

cxx: Error Handling

```
01 fn foo() -> Result<(), String> {
      Err("Rust Error".to_string())
03 }
04
05 #[cxx::bridge]
06 mod ffi {
       extern "Rust" {
          fn foo() -> Result<()>;
      unsafe extern "C++" {
          fn bar() -> Result<()>;
13 }
14
```

cxx: Error Handling

C++ exceptions to Rust Result<T, E>

```
01 fn foo() -> Result<(), String> {
      Err("Rust Error".to_string())
03 }
04
05 #[cxx::bridge]
06 mod ffi {
       extern "Rust" {
          fn foo() -> Result<()>;
      unsafe extern "C++" {
          fn bar() -> Result<()>;
13 }
14
```

cxx: Error Handling

C++ exceptions to Rust Result<T, E>

Converted at the boundary

```
01 namespace rust {
02 namespace behavior {
03
04 template <typename Try, typename Fail>
05 static void trycatch(Try &&func, Fail &&fail) noexcept
06 try {
       func();
08 } catch (const std::exception &e) {
       fail(e.what());
10 }
11
12 } // namespace behavior
13 } // namespace rust
14
```

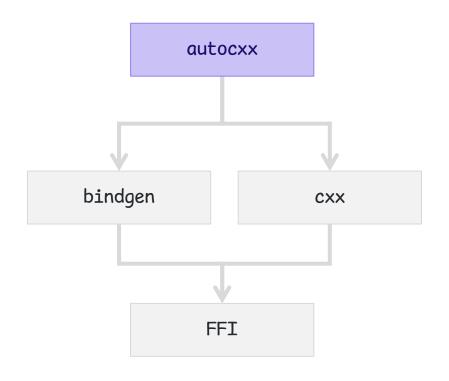
cxx: Error Handling

C++ exceptions to Rust Result<T, E>

Converted at the boundary

Custom conversion function

Rust/C++: autocxx



Glues bindgen and cxx together

```
01 include_cpp! {
       #include "foo.hpp"
       safety!(unsafe)
03
       generate!("Foo")
04
       generate!("Bar")
05
06 }
07 fn main() {
       let foo = ffi::Foo::new(10.into());
       println!("{}", foo.val);
09
10
11
       let mut bar = ffi::Bar::new(20.into())
           .within_unique_ptr();
       println!("{}", bar.pin_mut().val());
13
14 }
```

autocxx

```
01 include_cpp! {
       #include "foo.hpp"
       safety!(unsafe)
03
       generate!("Foo")
04
       generate!("Bar")
05
06 }
07 fn main() {
       let foo = ffi::Foo::new(10.into());
       println!("{{}}", foo.val);
09
10
       let mut bar = ffi::Bar::new(20.into())
           .within_unique_ptr();
       println!("{}", bar.pin_mut().val());
13
14 }
```

autocxx

Glues bindgen and cxx together

```
01 include_cpp! {
       #include "foo.hpp"
       safety!(unsafe)
03
       generate!("Foo")
04
       generate!("Bar")
05
06 }
07 fn main() {
       let foo = ffi::Foo::new(10.into());
       println!("{}", foo.val);
09
10
11
       let mut bar = ffi::Bar::new(20.into())
           .within_unique_ptr();
       println!("{}", bar.pin_mut().val());
13
14 }
```

autocxx

Glues bindgen and cxx together

Provides a fluent bridge

```
01 include_cpp! {
       #include "foo.hpp"
       safety!(unsafe)
03
       generate!("Foo")
04
       generate!("Bar")
05
06 }
07 fn main() {
       let foo = ffi::Foo::new(10.into());
       println!("{}", foo.val);
09
10
11
       let mut bar = ffi::Bar::new(20.into())
           .within_unique_ptr();
       println!("{}", bar.pin_mut().val());
13
14 }
```

autocxx

Glues bindgen and cxx together

Provides a fluent bridge

Trivial types are simple values

```
01 include_cpp! {
       #include "foo.hpp"
       safety!(unsafe)
03
       generate!("Foo")
04
       generate!("Bar")
05
06 }
07 fn main() {
       let foo = ffi::Foo::new(10.into());
       println!("{}", foo.val);
09
10
11
       let mut bar = ffi::Bar::new(20.into())
           .within_unique_ptr();
       println!("{}", bar.pin_mut().val());
13
14 }
```

autocxx

Glues bindgen and cxx together

Provides a fluent bridge

Non-trivial types behind an indirection + Pin



Summary

Use C as a foundation

Summary

Use C as a foundation

Static assert type layout

Summary

Use C as a foundation

Static assert type layout

Use code generation to overcome differences

Summary

Use C as a foundation

Static assert type layout

Use code generation to overcome differences

Handle exceptions at the boundary



Danke schön! Fragen?