

Compile-time Deadlock Detection in Rust using Petri Nets

Horacio Lisdero Scaffino

Facultad de Ingeniería
Universidad de Buenos Aires

June 30, 2023

Agenda

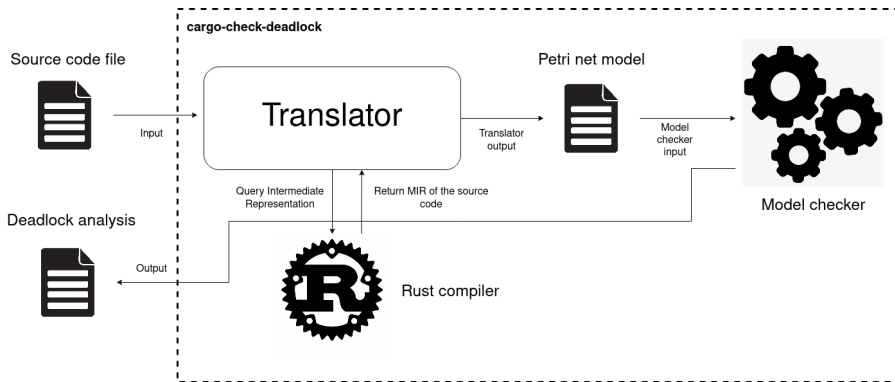
- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

A bird's-eye view of the tool

The Translator is the core component. The model checker and the Rust compiler, *rustc*, are dependencies.



Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

What is Rust?

Rust is a multi-paradigm, general-purpose programming language that aims to provide developers with a safe and efficient way to write low-level code.

What is Rust?

Rust is a multi-paradigm, general-purpose programming language that aims to provide developers with a safe and efficient way to write low-level code.

- Memory-safe
- Compiled to machine code, no runtime needed
- High-level simplicity
- Low-level performance (on the same level as C or C++)

Brief timeline of Rust

- 2007** Started as a side project by Graydon Hoare, a programmer at Mozilla
- 2009** Mozilla officially started sponsoring the project
- 2015** First stable version 1.0
- 2016** Mozilla releases Servo, a browser engine built with Rust
- 2019** `async/await` support stabilized
- 2021** The Rust Foundation is founded by AWS, Huawei, Google, Microsoft, and Mozilla
- 2021** The Android Open Source Project encourages the use of Rust for the SO components below the ART
- 2022** The Linux kernel adds support for Rust alongside C
- 2023** 8 years in a row the most loved programming language in the Stack Overflow Developer Survey

Memory safety

It achieves memory safety without using a garbage collector or reference counting. Instead, it uses the concept of **ownership** and **borrowing**.

Memory safety

It achieves memory safety without using a garbage collector or reference counting. Instead, it uses the concept of **ownership** and **borrowing**.

It prevents a wide variety of error classes at compile-time:

- Double free
- Use after free
- Dangling pointers
- Data races
- Passing non-thread-safe variables

If a violation of the compiler rules is found, the program will simply not compile.

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - **How does it look like?**
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Immutability by default

In other languages, immutability is the exception or an afterthought (e.g. `const`-ness in C/C++).

```

1  fn main() {
2      let x = 1; // Immutable by default
3      x = x + 1;
4  }
```

The Rust compiler points out exactly where the error is and provides help on how to fix it.

```

error[E0384]: cannot assign twice to immutable variable `x`
--> src/main.rs:3:5
  |
2 |     let x = 1;
  |         -
  |         |
  |         first assignment to `x`
  |         help: consider making this binding mutable: `mut x`
3 |     x = x + 1;
  |     ^^^^^^^^^ cannot assign twice to immutable variable
```

Move semantics by default

Each value has only one owner. If a variable is passed to another function or assigned to a different variable, the owner of the value changes.

```
1 fn main() {
2     let name = String::from("Alice");
3     print_name(name);
4     println!("The name is: {}", name); // Compilation error
5 }
6
7 fn print_name(name: String) {
8     println!("Name: {}", name);
9 }
```

Values have copy semantics only if they are marked as `Copy`. This is the case for numbers by default. Compare this with the default in C++ vs the best practices.

Move semantics by default: Error message

```

error[E0382]: borrow of moved value: `name`
--> src/main.rs:4:33
|
2 |     let name = String::from("Alice");
|         ---- move occurs because `name` has type `String`,          |
3 |     print_name(name);
|         ---- value moved here
4 |     println!("The name is: ", name); // Compilation error
|                                     ^^^^^ value borrowed here after move

note: consider changing this parameter type in function `print_name` to borrow
instead if owning the value isn't necessary
--> src/main.rs:7:21
|
7 | fn print_name(name: String) {
|     -----          ^^^^^^ this parameter takes ownership of the value
|     |
|     in this function
= note: this error originates in the macro `$crate::format_args_nl`
  which comes from the expansion of the macro `println` (in Nightly builds, run with -Z macro-backtrace for more info)
help: consider cloning the value if the performance cost is acceptable
|
3 |     print_name(name.clone());
|                   ++++++++

```

Algebraic Data Types, aka enums with fields

```

1  enum Shape {
2      Circle { radius: f64 },
3      Rectangle { width: f64, height: f64 },
4      Triangle { base: f64, height: f64 },
5  }
6
7  fn main() {
8      let shapes = vec![
9          Shape::Circle { radius: 5.0 },
10         Shape::Rectangle { width: 10.0, height: 8.0 },
11         Shape::Triangle { base: 7.0, height: 4.0 },
12     ];
13
14     for shape in shapes {
15         match shape {
16             Shape::Circle { radius } => {
17                 let circle = Circle { radius };
18                 // Do something with the circle...
19             },
20             Shape::Rectangle { width, height } => {
21                 let rectangle = Rectangle { width, height };
22                 // Do something with the rectangle...
23             },
24             Shape::Triangle { base, height } => {
25                 let triangle = Triangle { base, height };
26                 // Do something with the triangle...
27             },
28         }
29     }
30 }

```


A more advanced match statement

A **match** statement works *with* the type system, while a mere **if** can do anything and it is not bound to the type system. Match statements are always *exhaustive*: They must handle all possibilities.

```
1  fn main() {
2      let number = 42;
3
4      match number {
5          0 => println!("The number is zero"),
6          1 | 2 | 3 => println!("The number is a small prime"),
7          n @ 4..=9 => println!("The number is between 4 and 9: {n}"),
8          n if is_even(n) => println!("The number is even: {n}"),
9          n if is_odd(n) => println!("The number is odd: {n}"),
10         _ => panic!("The number doesn't match any specific case!"),
11     }
12 }
```

Python 3.10 introduced a similar feature ([PEP 636](#)). Java 17 has a limited version of this ([JEP 406](#)).

Modeling data in Rust

- Leverage the type system in your favor
- Make invalid states unrepresentable
- Define new types for the entities in your domain
- Use enums when variables can take different values

```
1 struct FakeCat {  
2     alive: bool,  
3     hungry: bool,  
4 }
```

```
1 enum RealCat {  
2     Alive { hungry: bool },  
3     Dead,  
4 }
```

By directly modeling the business domain with Rust's expressive type system, the compiler is able to verify the business logic and we catch more errors at compile-time.

Error handling with Result

```
1  use std::fs::File;
2  use std::io::Read;
3  // This definition is part of the standard library
4  // It does not need to be imported
5  enum Result<T, E> {
6      Ok(T),
7      Err(E),
8  }
9
10 fn read_file_contents(path: &str) -> Result<String, std::io::Error> {
11     let mut file = File::open(path)?;
12     let mut contents = String::new();
13     file.read_to_string(&mut contents)?;
14     Ok(contents)
15 }
16
17 fn main() {
18     let file_path = "example.txt";
19     let result = read_file_contents(file_path);
20
21     match result {
22         Ok(contents) => {
23             println!("File contents:\n{}", contents);
24         }
25         Err(error) => {
26             eprintln!("Error reading file: {}", error);
27         }
28     }
29 }
```

The enum Option: No need for null pointers

```
1  // This definition is part of the standard library
2  // It does not need to be imported
3  pub enum Option<T> {
4      None,
5      Some(T),
6  }
7
8  fn main() {
9      let mut list = vec![1, 2, 3, 4, 5];
10
11     while let Some(element) = list.pop() {
12         println!("Popped element: {}", element);
13     }
14     // List::pop() returned `None`
15     println!("List is empty!");
16 }
```

No OOP: Just structs with methods

```
1  struct Rectangle {
2      width: u32,
3      height: u32,
4  }
5
6  impl Rectangle {
7      fn new(width: u32, height: u32) -> Rectangle {
8          Rectangle { width, height }
9      }
10
11     fn area(&self) -> u32 {
12         self.width * self.height
13     }
14
15     fn is_square(&self) -> bool {
16         self.width == self.height
17     }
18
19     fn double_size(&mut self) {
20         self.width *= 2;
21         self.height *= 2;
22     }
23 }
```

Define traits to share an interface

```
1  trait Container {
2      fn get_value(&self);
3  }
4
5  struct Storage {
6      value: i32,
7  }
8
9  impl Container for Storage {
10     fn get_value(&self) {
11         println!("Value: {}", self.value);
12     }
13 }
14
15 impl PartialEq for Storage {
16     fn eq(&self, other: &Self) -> bool {
17         self.value == other.value
18     }
19 }
20
21 impl std::fmt::Display for Storage {
22     fn fmt(&self, f: &mut std::fmt::Formatter) -> std::fmt::Result {
23         write!(f, "Value: {}", self.value)
24     }
25 }
```

Nearly everything is an expression as in Lisp

```
1  fn main() {
2      let numbers = vec![1, 2, 3, 4, 5];
3
4      let sum_of_squares: i32 = numbers
5          .iter()
6          .fold(0, |acc, x| acc + x * x);
7
8      let result = if sum_of_squares > 50 {
9          "Sum of squares is greater than 50"
10     } else {
11         "Sum of squares is not greater than 50"
12     };
13
14     println!("Result: {}", result);
15 }
```

Generics

```
1  struct Pair<T, U> {
2      first: T,
3      second: U,
4  }
5
6  impl<T, U> Pair<T, U> {
7      fn new(first: T, second: U) -> Self {
8          Pair { first, second }
9      }
10
11     fn get_first(&self) -> &T {
12         &self.first
13     }
14
15     fn get_second(&self) -> &U {
16         &self.second
17     }
18 }
```


Lifetimes

```
1 struct StringHolder<'a> {
2     value: &'a str,
3 }
4
5 impl<'a> StringHolder<'a> {
6     fn new(value: &'a str) -> Self {
7         StringHolder { value }
8     }
9
10    fn get_value(&self) -> &'a str {
11        self.value
12    }
13 }
14
15 fn main() {
16     let input_string = String::from("Hello, lifetimes!");
17
18     let holder;
19     {
20         let local_string = String::from("Local string");
21         holder = StringHolder::new(local_string.as_str());
22         println!("Holder value: {}", holder.get_value());
23     }
24
25     println!("Input string: {}", input_string);
26     println!("Holder value: {}", holder.get_value());
27 }
```

Lifetimes: Error message

```

error[E0597]: `local_string` does not live long enough
--> src/main.rs:21:36
|
20 |         let local_string = String::from("Local string");
|         ----- binding `local_string` declared here
21 |         holder = StringHolder::new(local_string.as_str());
|                                     ^^^^^^^^^^^^^^^^^^^^^ borrowed value does not
|                                                             live long enough
22 |         println!("Holder value: ", holder.get_value());
23 |     }
|     - `local_string` dropped here while still borrowed
...
26 |     println!("Holder value: ", holder.get_value());
|                                     ----- borrow later used here

```

For more information about this error, try ``rustc --explain E0597``.

Only one active mutable reference at any given time

```

1  struct Item {
2      value: i32,
3  }
4
5  fn main() {
6      let mut item = Item { value: 42 };
7
8      let reference1 = &mut item; // First mutable reference
9      let reference2 = &mut item; // Second mutable reference - COMPILATION ERROR
10
11     reference1.value += 1;
12     reference2.value += 1;
13
14     println!("Reference 1: {}", reference1.value);
15     println!("Reference 2: {}", reference2.value);
16 }

```

error[E0499]: cannot borrow `item` as mutable more than once at a time

--> src/main.rs:9:22

```

8 |     let reference1 = &mut item; // First mutable reference
  |                      ----- first mutable borrow occurs here
9 |     let reference2 = &mut item; // Second mutable reference - COMPILATION ERROR
  |                      ^^^^^^^^^ second mutable borrow occurs here
10 |
11 |     reference1.value += 1;
  |     ----- first borrow later used here

```

For more information about this error, try `rustc --explain E0499`.

A mutable reference is allowed only if no immutable references are present

```

1  fn main() {
2      let mut item = 42;
3
4      let reference1 = &item; // First mutable reference
5      let reference2 = &mut item; // Second mutable reference - COMPILATION ERROR
6
7      if *reference1 == 1 {
8          println!("Item is set to one");
9      }
10     *reference2 += 1;
11
12     println!("Reference 1: {}", reference1);
13     println!("Reference 2: {}", reference2);
14 }

```

error[E0502]: cannot borrow `item` as mutable because it is also borrowed as immutable

--> src/main.rs:5:22

```

4 |     let reference1 = &item; // First mutable reference
  |                       ----- immutable borrow occurs here
5 |     let reference2 = &mut item; // Second mutable reference - COMPILATION ERROR
  |                       ^^^^^^^^^ mutable borrow occurs here
6 |
7 |     if *reference1 == 1 {
  |           ----- immutable borrow later used here

```

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - **Why Rust?**
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Memory safety is critical for reliability and security

Empirical investigations have concluded that around 70% of the vulnerabilities found in large C/C++ codebases are due to memory handling errors. This high figure can be observed in projects such as:

- Android Open Source Project [1],
- the Bluetooth and media components of Android [2],
- the Chromium Projects behind the Chrome web browser [3],
- the CSS component of Firefox [4],
- iOS and macOS [5],
- Microsoft products [6, 7],
- Ubuntu [8]

Rust adoption is increasing fast

- The Android Open Source Project encourages the use of Rust for the SO components below the ART [9].
- The Linux kernel introduces in version 6.1 official tooling support for programming components in Rust [10, 11].
- At Mozilla, the Oxidation project was created in 2015 to increase the usage of Rust in Firefox and related projects. As of March 2023, the lines of code in Rust represent more than 10% of the total in Firefox Nightly [12].
- At Meta, the use of Rust as a development language server-side is approved and encouraged since July 2022 [13].
- At Cloudflare, a new HTTP proxy in Rust was built from scratch to overcome the architectural limitations of NGINX, reducing CPU usage by 70% and memory usage by 67% [14].
- At Discord, reimplementing a crucial service in Rust provided great benefits in performance and solved a performance penalty due to the garbage collection in Go [15].
- At npm Inc., the company behind the npm registry, Rust allowed scaling CPU-bound services to more than 1.3 billion downloads per day [16].
- A study of Rust-based code found it runs so efficiently that it uses half as much electricity as a similar program written in Java, a language commonly used at AWS [17].

The tooling is great and works out of the box

- `cargo`, the official package manager: Format, build, test, lint, update, and publish packages.

The tooling is great and works out of the box

- **cargo**, the official package manager: Format, build, test, lint, update, and publish packages.
- **Test harness** for unit tests, integration tests, and tests in documentation comments (doctests). No third-party libraries needed.

The tooling is great and works out of the box

- **cargo**, the official package manager: Format, build, test, lint, update, and publish packages.
- **Test harness** for unit tests, integration tests, and tests in documentation comments (doctests). No third-party libraries needed.
- An official public registry for Rust packages (called “crates”): <https://crates.io/>.

The tooling is great and works out of the box

- **cargo**, the official package manager: Format, build, test, lint, update, and publish packages.
- **Test harness** for unit tests, integration tests, and tests in documentation comments (doctests). No third-party libraries needed.
- An official public registry for Rust packages (called “crates”): <https://crates.io/>.
- Automatic generation of a static website from the doc comments of the project. It is published automatically to <https://docs.rs/>.

The tooling is great and works out of the box

- **cargo**, the official package manager: Format, build, test, lint, update, and publish packages.
- **Test harness** for unit tests, integration tests, and tests in documentation comments (doctests). No third-party libraries needed.
- An official public registry for Rust packages (called “crates”): <https://crates.io/>.
- Automatic generation of a static website from the doc comments of the project. It is published automatically to <https://docs.rs/>.
- An official linter included with the default installation that catches even more errors and spots non-idiomatic code: **clippy**.

The tooling is great and works out of the box

- **cargo**, the official package manager: Format, build, test, lint, update, and publish packages.
- **Test harness** for unit tests, integration tests, and tests in documentation comments (doctests). No third-party libraries needed.
- An official public registry for Rust packages (called “crates”): <https://crates.io/>.
- Automatic generation of a static website from the doc comments of the project. It is published automatically to <https://docs.rs/>.
- An official linter included with the default installation that catches even more errors and spots non-idiomatic code: **clippy**.
- Integration with git, GitHub, VSCode, IntelliJ is great and easy to use.

The tooling is great and works out of the box

- **cargo**, the official package manager: Format, build, test, lint, update, and publish packages.
- **Test harness** for unit tests, integration tests, and tests in documentation comments (doctests). No third-party libraries needed.
- An official public registry for Rust packages (called “crates”): <https://crates.io/>.
- Automatic generation of a static website from the doc comments of the project. It is published automatically to <https://docs.rs/>.
- An official linter included with the default installation that catches even more errors and spots non-idiomatic code: **clippy**.
- Integration with git, GitHub, VSCode, IntelliJ is great and easy to use.
- A new stable compiler release every 6 weeks [18].

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets**
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets**
 - What is a Petri net?**
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Informal definition

A Petri net is a mathematical modeling tool used to describe and analyze the behavior of concurrent systems. It provides a graphical representation of the system's state and its transitions, allowing for visual and formal analysis of complex processes.



- Places: Represent states in the system (*circles*)
- Transitions: Represent usually events or actions that occur in the system (*rectangles*)
- Tokens: Marks inside of places that are created and consumed by transitions (*points inside of places*)

Mathematical definition

A Petri net is a 5-tuple, $PN = (P, T, F, W, M_0)$ where:

$P = \{p_1, p_2, \dots, p_m\}$ is a finite set of places,

$T = \{t_1, t_2, \dots, t_n\}$ is a finite set of transitions,

$F \subseteq (P \times T) \cup (T \times P)$ is a set of arcs (flow relation),

$W : F \rightarrow \{1, 2, 3, \dots\}$ is a weight function for the arcs,

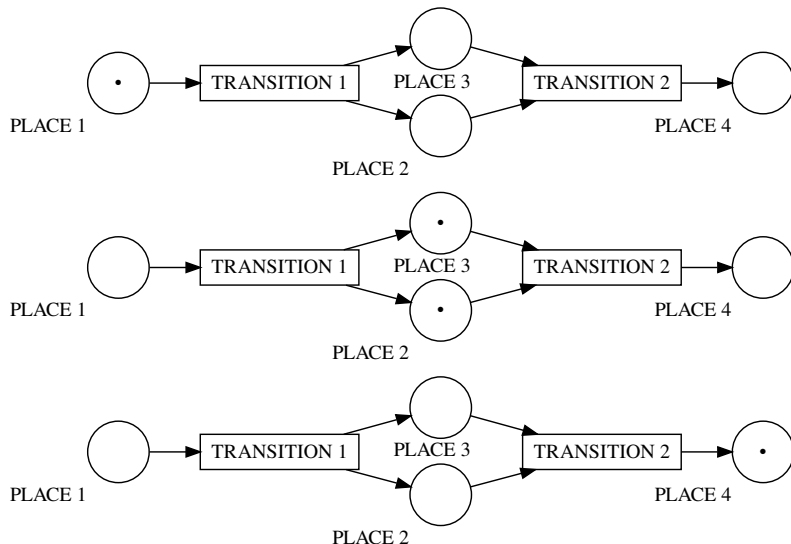
$M_0 : P \rightarrow \{0, 1, 2, 3, \dots\}$ is the initial marking,

$P \cap T = \emptyset$ and $P \cup T \neq \emptyset$

The graph is by definition *bipartite*. There can only be edges:

- from places to transitions or
- from transitions to places

Transition firing rule

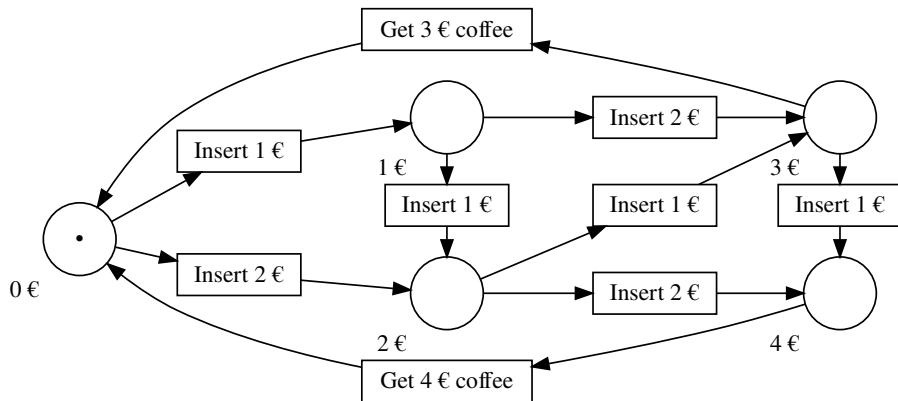


Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets**
 - What is a Petri net?
 - Examples**
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

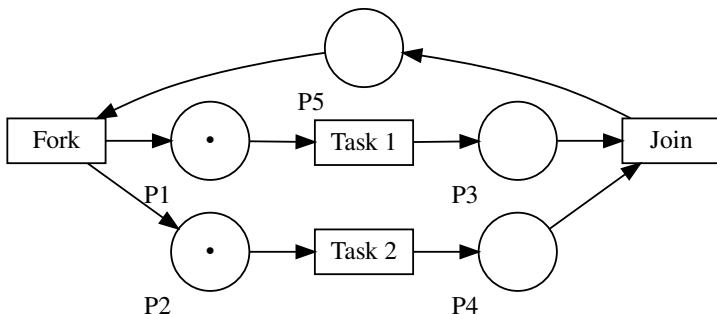
Vending machine

This is a finite-state machine (FSM), a subclass of Petri nets.



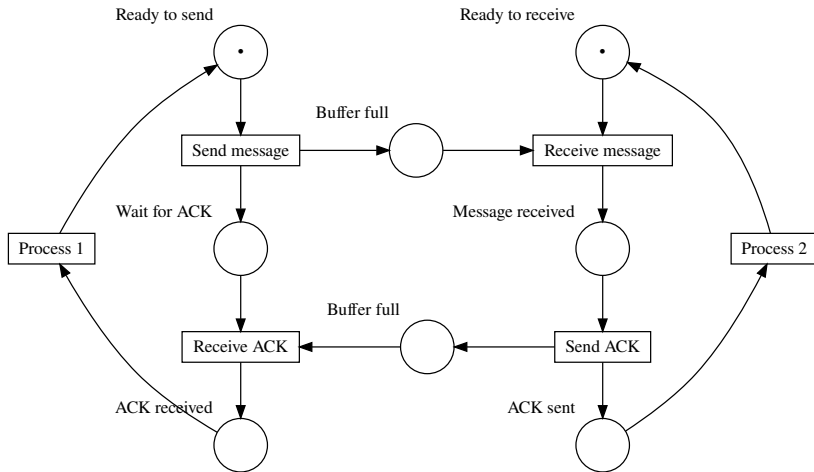
Parallel activities: Fork/Join

This is a marked graph (MG), a subclass of Petri nets. Observe the concurrency between Task 1 and Task 2. This cannot be modeled by a single finite-state machine.



Communication protocols: Send with ACK

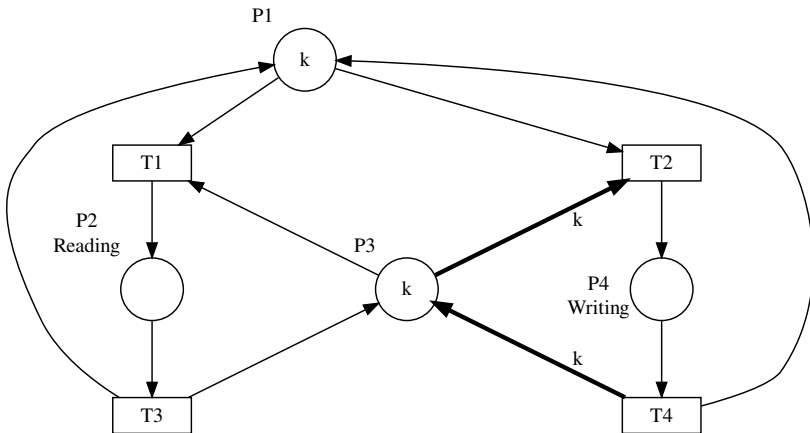
A simple protocol in which Process 1 sends messages to Process 2 and waits for an acknowledgment to be received before continuing. For simplicity, no timeout mechanism was included.



Synchronization control: Readers and writers

A Petri net system with k processes that either read or write a shared value.

- If one process writes, then no process may read.
- If a process is reading, then no process may write.
- There can only be zero or one process writing at any given time.



Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets**
 - What is a Petri net?
 - Examples
 - Why Petri nets?**
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Reachability analysis

Petri nets can be analyzed using formal methods to conclude whether the net can reach a deadlock or not. There is a notion of liveness analogous to the one found in computer systems.

Reachability analysis

Petri nets can be analyzed using formal methods to conclude whether the net can reach a deadlock or not. There is a notion of liveness analogous to the one found in computer systems.

Several model checkers are being developed and there is even a Model Checking Contest that takes place every year. State-of-the-art tools can handle Petri net models with more than **70 000 transitions** and **one million places**.

Reachability analysis

Petri nets can be analyzed using formal methods to conclude whether the net can reach a deadlock or not. There is a notion of liveness analogous to the one found in computer systems.

Several model checkers are being developed and there is even a Model Checking Contest that takes place every year. State-of-the-art tools can handle Petri net models with more than **70 000 transitions** and **one million places**.

Translating source code to a Petri net has been done before for other programming languages [19, 20] and also for Rust [21, 22]. The difficulty lies in supporting more synchronization primitives than simple mutexes and translating code from real-world applications.

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation**
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

Compilation stages in *rustc*

- Lexing: The source text is turned into a stream of atomic source code units known as tokens.

Compilation stages in *rustc*

- Lexing: The source text is turned into a stream of atomic source code units known as tokens.
- Parsing: The stream of tokens is converted into an **Abstract Syntax Tree**.

Compilation stages in *rustc*

- Lexing: The source text is turned into a stream of atomic source code units known as tokens.
- Parsing: The stream of tokens is converted into an **Abstract Syntax Tree**.
- High-level Intermediate Representation (HIR):
 - Desugar loops: **while** and **for** to simple **loop**.
 - Type inference: The automatic detection of a type of an expression.
 - Trait solving: Ensuring that each implementation block (**impl**) points to a valid trait.
 - Type checking.

Compilation stages in *rustc*

- Lexing: The source text is turned into a stream of atomic source code units known as tokens.
- Parsing: The stream of tokens is converted into an **Abstract Syntax Tree**.
- High-level Intermediate Representation (HIR):
 - Desugar loops: **while** and **for** to simple **loop**.
 - Type inference: The automatic detection of a type of an expression.
 - Trait solving: Ensuring that each implementation block (**impl**) points to a valid trait.
 - Type checking.
- Mid-level Intermediate Representation (MIR):
 - Checking of exhaustiveness of pattern matching.
 - Desugar method calls to function calls
(`x.method(y)` becomes `Type::method(&x, y)`).
 - Add implicit dereferencing operations.
 - Borrow checking.

Compilation stages in *rustc*

- Lexing: The source text is turned into a stream of atomic source code units known as tokens.
- Parsing: The stream of tokens is converted into an **Abstract Syntax Tree**.
- High-level Intermediate Representation (HIR):
 - Desugar loops: **while** and **for** to simple **loop**.
 - Type inference: The automatic detection of a type of an expression.
 - Trait solving: Ensuring that each implementation block (**impl**) points to a valid trait.
 - Type checking.
- Mid-level Intermediate Representation (MIR):
 - Checking of exhaustiveness of pattern matching.
 - Desugar method calls to function calls
(`x.method(y)` becomes `Type::method(&x, y)`).
 - Add implicit dereferencing operations.
 - Borrow checking.
- Code generation:
 - *rustc* relies on LLVM as a backend.
 - It leverages many optimizations of the LLVM intermediate representation.
 - LLVM takes over from this point on.
 - At the end, object files are linked to create an executable.

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - **MIR**
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

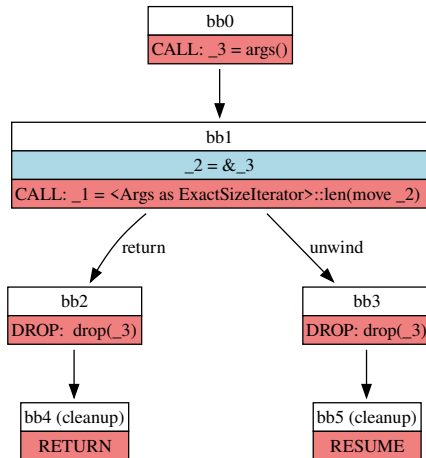
Hello World in MIR

BB means “basic block”. Each one is formed by statements and one terminator statement. The terminator statement is the only place where the control flow can jump to another basic block.

```
1  fn main() -> () {
2      let mut _0: ();
3      let _1: ();
4      let mut _2: std::fmt::Arguments<'_>;
5      let mut _3: &[&str];
6      let mut _4: &[&str; 1];
7
8      bb0: {
9          _4 = const _;
10         _3 = _4 as &[&str] (Pointer(Unsize));
11         _2 = Arguments::<'_>::new_const(move _3) -> bb1;
12     }
13
14     bb1: {
15         _1 = _print(move _2) -> bb2;
16     }
17
18     bb2: {
19         return;
20     }
21 }
```

MIR as a graph that shows the flow of execution

The MIR is a form of control flow graph (CFG) used in compilers. In this form, the translation to a Petri net becomes evident.



Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - **Modeling threads**
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

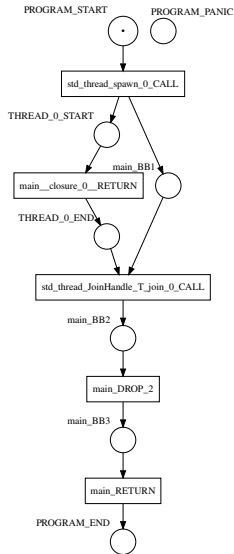
Example program

Let's consider a trivial program that spawns a thread that does nothing and immediately joins it.

```
1  fn main() {  
2      let thread_join_handle = std::thread::spawn(move || {  
3          // some work here  
4      });  
5      // some work here  
6      let _res = thread_join_handle.join();  
7  }
```

- `std::thread::spawn` should create an additional token that models the program counter of the second thread.
- The joining thread should wait until the spawned thread finishes.

Petri net model for a thread



Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation**
 - MIR
 - Modeling threads
 - Modeling mutexes**
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources

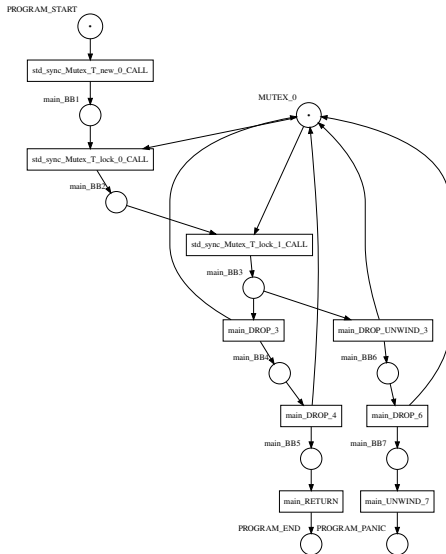
Example program

Consider a simple program that locks a mutex twice. The second lock operation will deadlock because the lock handle returned by the first call to `std::sync::Mutex::lock` is not dropped until it falls out of scope.

```
1  fn main() {  
2      let data = std::sync::Mutex::new(0);  
3      let _d1 = data.lock();  
4      let _d2 = data.lock(); // cannot lock, since d1 is still active  
5  }
```

- There should be a single place that models the mutex.
- Locking the mutex is taking the token from the mutex place.
- Unlocking the mutex is setting the token back in the mutex place.

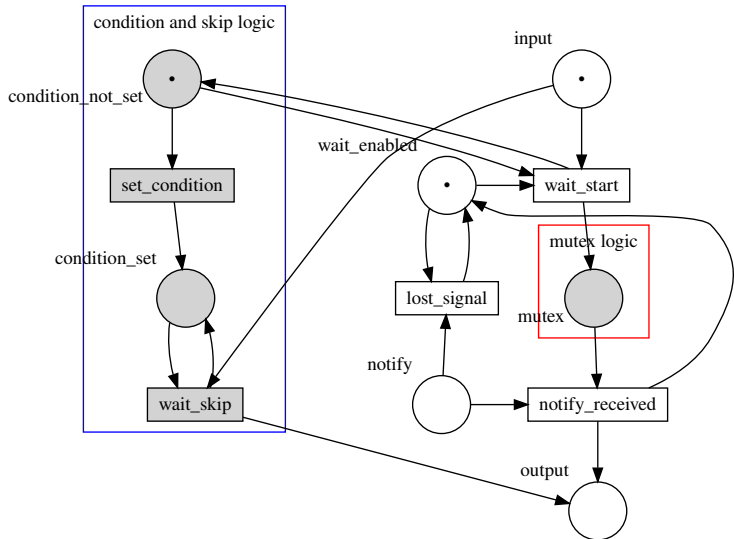
Petri net model for a mutex



Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation**
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables**
- 5 Bibliography
- 6 Additional resources

How to model a condition variable



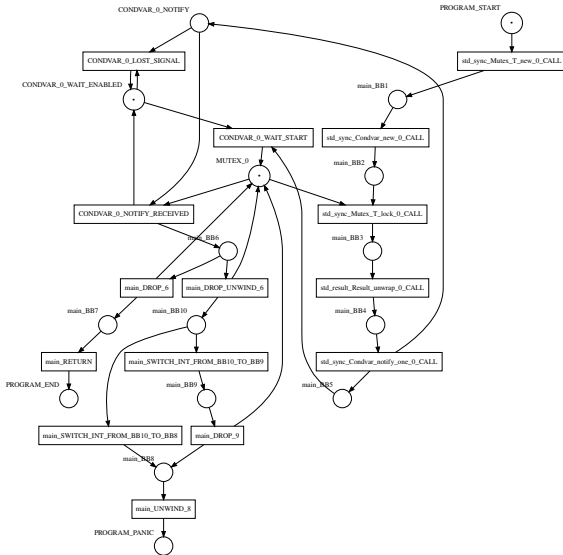
Example program

We have to use a very simple example program to keep the net small. In this case, the thread is trying to notify itself, which leads to a lost signal.

```
1  fn main() {  
2      let mutex = std::sync::Mutex::new(false);  
3      let cvar = std::sync::Condvar::new();  
4      let mutex_guard = mutex.lock().unwrap();  
5      cvar.notify_one();  
6      let _result = cvar.wait(mutex_guard);  
7  }
```

- The model for the condition variable should appear in the Petri net.
- The notify place should be set.
- But the signal gets consumed because `std::sync::Condvar::wait` was not called.

Petri net model for the example program



Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography**
- 6 Additional resources

Bibliography I



E. Stepanov, "Detecting Memory Corruption Bugs With HWSan." <https://android-developers.googleblog.com/2020/02/detecting-memory-corruption-bugs-with-hwasan.html>, 2020.
Accessed on 2023-02-24.



J. V. Stoep and C. Zhang, "Queue the Hardening Enhancements." <https://android-developers.googleblog.com/2020/02/detecting-memory-corruption-bugs-with-hwasan.html>, 2020.
Accessed on 2023-02-24.



The Chromium Projects, "Memory safety."
<https://www.chromium.org/Home/chromium-security/memory-safety/>, 2015.
Accessed on 2023-02-24.



D. Hosfelt, "Implications of Rewriting a Browser Component in Rust."
<https://hacks.mozilla.org/2019/02/rewriting-a-browser-component-in-rust/>, 2019.
Accessed on 2023-02-24.



P. Kehrer, "Memory Unsafety in Apple's Operating Systems."
<https://langui.sh/2019/07/23/apple-memory-safety/>, 2019.
Accessed on 2023-02-24.



M. Miller, "Trends, Challenges, and Strategic Shifts in the Software Vulnerability Mitigation Landscape."
<https://www.youtube.com/watch?v=PjbG0jjnBZQ>, 2019.
Accessed on 2023-02-24.



S. Fernandez, "A proactive approach to more secure code."
<https://msrc.microsoft.com/blog/2019/07/a-proactive-approach-to-more-secure-code/>, 2019.
Accessed on 2023-02-24.

Bibliography II



A. Gaynor, "What science can tell us about C and C++'s security."

<https://alexgaynor.net/2020/may/27/science-on-memory-unsafety-and-security/>, 2020.

Accessed on 2023-02-24.



J. V. Stoep and S. Hines, "Rust in the Android platform."

<https://security.googleblog.com/2021/04/rust-in-android-platform.html>, 2021.

Accessed on 2023-02-22.



J. Corbet, "The 6.1 kernel is out." <https://lwn.net/Articles/917504/>, 2022.

Accessed on 2023-02-24.



S. D. Simone, "Linux 6.1 Officially Adds Support for Rust in the Kernel."

<https://www.infoq.com/news/2022/12/linux-6-1-rust/>, 2022.

Accessed on 2023-02-24.



Mozilla Wiki, "Oxidation Project." <https://wiki.mozilla.org/Oxidation>, 2015.

Accessed on 2023-03-20.



E. Garcia, "Programming languages endorsed for server-side use at Meta." <https://engineering.fb.com/2022/07/27/developer-tools/programming-languages-endorsed-for-server-side-use-at-meta/>, 2022.

Accessed on 2023-02-24.



Y. Wu and A. Hauck, "How we built Pingora, the proxy that connects Cloudflare to the Internet." <https://blog.cloudflare.com/how-we-built-pingora-the-proxy-that-connects-cloudflare-to-the-internet/>,

2022.

Accessed on 2023-03-20.

Bibliography III



J. Howarth, "Why Discord is switching from Go to Rust."

<https://discord.com/blog/why-discord-is-switching-from-go-to-rust>, 2020.
Accessed on 2023-03-20.



The Rust Project Developers, "Rust case study: Community makes rust an easy choice for npm," 2019.

<https://www.rust-lang.org/static/pdfs/Rust-npm-Whitepaper.pdf>.



R. Pereira, M. Couto, F. Ribeiro, R. Rua, J. Cunha, J. P. Fernandes, and J. Saraiva, "Energy efficiency across programming languages: how do energy, time, and memory relate?," in *Proceedings of the 10th ACM SIGPLAN international conference on software language engineering*, pp. 256–267, 2017.



P. Albin, "RustFest Barcelona - Shipping a stable compiler every six weeks."

<https://www.youtube.com/watch?v=As1gXp5kX1M>, 2019.
Accessed on 2023-02-24.



K. M. Kavi, A. Moshtaghi, and D.-J. Chen, "Modeling multithreaded applications using petri nets," *International Journal of Parallel Programming*, vol. 30, pp. 353–371, 2002.



A. Moshtaghi, "Modeling Multithreaded Applications Using Petri Nets," Master's thesis, The University of Alabama in Huntsville, 2001.



T. Meyer, "A Petri Net semantics for Rust," Master's thesis, Universität Rostock — Fakultät für Informatik und Elektrotechnik, 2020.

<https://github.com/Skasselbard/Granite/blob/master/doc/MasterThesis/main.pdf>.



K. Zhang and G. Liua, "Automatically transform rust source to petri nets for checking deadlocks," *arXiv preprint arXiv:2212.02754*, 2022.

Agenda

- 1 Introduction
- 2 Rust
 - What is Rust?
 - How does it look like?
 - Why Rust?
- 3 Petri nets
 - What is a Petri net?
 - Examples
 - Why Petri nets?
- 4 Translation
 - MIR
 - Modeling threads
 - Modeling mutexes
 - Modeling condition variables
- 5 Bibliography
- 6 Additional resources**

Resources to learn Rust

- [The Rust Book](#): Available online and locally with the default Rust installation.
- [Rust by Example](#): Another official book with a more practical approach.
- [Rustlings](#): Small exercises to get you used to reading and writing Rust code!
- [Comprehensive Rust](#): A three-day Rust course developed by the Android team.
- [Take your first steps with Rust](#): A simple course on Microsoft Learn.
- [Rust Programming Course for Beginners](#) by freeCodeCamp.org.
- [No Boilerplate](#): A Youtube channel mainly dedicated to topics connected with Rust. Some ideas were used for this presentation.

Online Petri net simulators

- A simple simulator by Igor Kim can be found on <https://petri.hp102.ru/>. A tutorial video on Youtube and example nets are included in the tool.
- A complement to this is a series of interactive tutorials by Prof. Wil van der Aalst at the University of Hamburg. These tutorials are Adobe Flash Player files (with extension `.swf`) that modern web browsers cannot execute. Luckily, an online Flash emulator like the one found on https://flashplayer.fullstacks.net/?kind=Flash_Emulator can be used to upload the files and execute them.
- Another online Petri net editor and simulator is <http://www.biregal.com/>. The user can draw the net, add the tokens, and then manually fire transitions.

Questions?

Links

Thesis <https://github.com/hlisdero/thesis>

Tool <https://github.com/hlisdero/cargo-check-deadlock>

Prasentation https://github.com/hlisdero/thesis/tree/main/presentation_en

Published crate <https://crates.io/crates/cargo-check-deadlock>