



Universidad
Zaragoza

Trabajo Fin de Grado

DYNAMIC LOADING OF PLUGINS IN RUST IN THE ABSENCE OF A STABLE APPLICATION BINARY INTERFACE

Autor:

MARIO ORTIZ MANERO

Director:

JAVIER FABRA CARO

Grado en Ingeniería Informática
Departamento de Ingeniería e Ingeniería de Sistemas
Escuela de Ingeniería y Arquitectura

Junio 2022

AGRADECIMIENTOS

Padre y madre o familia

Amigos, de uni y de siempre

Profesores, en especial Fabra

Mentores de Tremor, comunidad open source, fundación de linux, wayfair?

RESUMEN

Una mañana, tras un sueño intranquilo, Gregorio Samsa se despertó convertido en un monstruoso insecto. Estaba echado de espaldas sobre un duro caparazón y, al alzar la cabeza, vio su vientre convexo y oscuro, surcado por curvadas callosidades, sobre el que casi no se aguantaba la colcha, que estaba a punto de escurrirse hasta el suelo. Numerosas patas, penosamente delgadas en comparación con el grosor normal de sus piernas, se agitaban sin concierto. - ¿Qué me ha ocurrido? No estaba soñando. Su habitación, una habitación normal, aunque muy pequeña, tenía el aspecto habitual. Sobre la mesa había desparramado un muestrario de paños - Samsa era viajante de comercio-, y de la pared colgaba una estampa recientemente recortada de una revista ilustrada y puesta en un marco dorado. La estampa mostraba a una mujer tocada con un gorro de pieles, envuelta en una estola también de pieles, y que, muy erguida, esgrimía un amplio manguito, asimismo de piel, que ocultaba todo su antebrazo. Gregorio miró hacia la ventana; estaba nublado, y sobre el cinc del alféizar repiqueteaban las gotas de lluvia, lo que le hizo sentir una gran melancolía. “Bueno -pensó-; ¿y si siguiese durmiendo un rato y me olvidase de”

ABSTRACT

Una mañana, tras un sueño intranquilo, Gregorio Samsa se despertó convertido en un monstruoso insecto. Estaba echado de espaldas sobre un duro caparazón y, al alzar la cabeza, vio su vientre convexo y oscuro, surcado por curvadas callosidades, sobre el que casi no se aguantaba la colcha, que estaba a punto de escurrirse hasta el suelo. Numerosas patas, penosamente delgadas en comparación con el grosor normal de sus piernas, se agitaban sin concierto. - ¿Qué me ha ocurrido? No estaba soñando. Su habitación, una habitación normal, aunque muy pequeña, tenía el aspecto habitual. Sobre la mesa había desparramado un muestrario de paños - Samsa era viajante de comercio-, y de la pared colgaba una estampa recientemente recortada de una revista ilustrada y puesta en un marco dorado. La estampa mostraba a una mujer tocada con un gorro de pieles, envuelta en una estola también de pieles, y que, muy erguida, esgrimía un amplio manguito, asimismo de piel, que ocultaba todo su antebrazo. Gregorio miró hacia la ventana; estaba nublado, y sobre el cinc del alféizar repiqueteaban las gotas de lluvia, lo que le hizo sentir una gran melancolía. «Bueno -pensó-; ¿y si siguiese durmiendo un rato y me olvidase de

Índice

1. Introducción	1
1.1. Contexto	1
1.2. Objetivo	2
1.3. Motivación	2
1.3.1. Tiempos de compilación	2
1.3.2. Modularidad	3
1.3.3. Aprender de otros	3
1.4. Metodología	3
1.4.1. Organización	3
1.4.2. Desarrollo	4
1.4.3. Recursos públicos	4
2. Entendiendo Tremor	5
2.1. Sistemas de Procesado de Eventos	5
3. Tecnologías para implementar un sistema de plugins	7
4. Implementación	9
5. Conclusiones	11
Lista de Figuras	15
Lista de Tablas	17
Anexos	18
A. Un anexo	21

Capítulo 1

Introducción

1.1. Contexto

Este proyecto se ha realizado en colaboración con *Tremor*¹, un sistema de procesamiento de eventos de alto rendimiento, escrito en el lenguaje de programación *Rust*². Tremor es un programa de código abierto bajo la fundación *Cloud Native Computing Foundation (CNCF)*³, que es también parte de la organización *Linux Foundation (LFX)*⁴.

Formalmente, el trabajo se ha llevado a cabo gracias a la iniciativa *LFX Mentorship*, con el título “CNCF - Tremor: Add plugin support for tremor (PDK)”⁵⁶. Esta iniciativa promueve el aprendizaje de desarrolladores de código abierto, proporcionando una plataforma transparente y facilitando un sistema de pagos.

Finalmente, *Wayfair* es una empresa estadounidense de comercio digital de muebles y artículos del hogar⁸. Actualmente, ofrece 14 millones de ítems de más de 11.000 proveedores globales [1] y es el principal financiador tanto de Tremor como de este proyecto.

¹<https://tremor.rs>

²<https://www.rust-lang.org>

³<https://www.cncf.io/>

⁴<https://www.linuxfoundation.org/>

⁵Página oficial de la iniciativa: <https://mentorship.lfx.linuxfoundation.org/project/b90f7174-fc53-40bc-b9e2-9905f88c38ff>

⁶*Tracking issue* en GitHub: <https://github.com/tremor-rs/tremor-runtime/issues/791>

⁷RFC en la documentación de Tremor: <https://www.tremor.rs/rfc/accepted/plugin-development-kit/>

⁸<https://www.wayfair.com/>

1.2. Objetivo

La tarea a llevar a cabo es la implementación de un sistema de plugins, denominado *Plugin Development Kit (PDK)*, para la base de código ya existente en Tremor.

Esto es una tarea no-trivial, dado que Rust no tiene un *Application Binary Interface (ABI)* estable. Es decir, que si se compila la *runtime* (el binario principal encargado de cargar funcionalidad externa) y los *plugins* (los binarios individuales con la funcionalidad) de forma separada, no hay garantía de que la representación binaria de los datos o la convención de llamada a funciones — entre otros — sea la misma.

Esto implica que *dynamic loading* es imposible de forma segura puramente con Rust, debiéndose recurrir a otro ABI que sí sea estable, como el del lenguaje de programación C. Por tanto, se deben escribir *bindings* (la definición de la interfaz compartida entre runtime y plugins) completas en C y transformar tipos de Rust a C y viceversa cuando se interactúe con plugins.

1.3. Motivación

1.3.1. Tiempos de compilación

Actualmente, el problema más importante en Tremor es sus tiempos de compilación. En un ordenador de gama media de ~600 € como el Dell Vostro 5481, compilar el binario `tremor` desde cero requiere de más de 7 minutos en modo debug. Incluso en el caso de cambios incrementales (una vez las dependencias ya han sido compiladas), hay que esperar unos 10 segundos. Esto no es una buena experiencia de desarrollo e impide que nuevos programadores se unan a la comunidad de Tremor.

Debido a la naturaleza del programa, este problema solo empeorará con el tiempo. Tremor debe tener soporte para un gran número de protocolos (e.g., TCP o UDP), software (e.g., Kafka o PostgreSQL) y codecs (e.g., JSON o YAML). El número de dependencias continuará incrementando hasta que imposibilite la creación de nuevas prestaciones en Tremor.

Los problemas relacionados con tiempos de compilación excesivamente largos no se limitan a Tremor. Es uno de las mayores críticas que recibe Rust y un 61 % de sus usuarios declaran que aún se necesita trabajo para mejorar la situación [2].

1.3.2. Modularidad

Otra ventaja que provee un sistema de plugins es modularidad; ser capaz de tratar la runtime y los plugins de forma separada suele resultar en una arquitectura más limpia [3]. También hace posible el desacoplamiento del ejecutable y sus componentes; algunas dependencias tienen un ciclo de versionado más rápido que otras y generalmente es más conveniente actualizar únicamente un plugin, en lugar del programa por completo.

1.3.3. Aprender de otros

Otros proyectos maduros con características similares a las de Tremor, como *NGINX* [4] o *Apache HTTP Server* [5], llevan beneficiándose de un sistema de plugins desde hace mucho. Informan mejoras en flexibilidad, extensibilidad y facilidad de desarrollo [6][7]. Aunque las desventajas también mencionen un pequeño impacto en el rendimiento y la posibilidad de caer en un *dependency hell*, sigue siendo una buena idea al menos considerarlo para Tremor.

1.4. Metodología

1.4.1. Organización

El proyecto ha tenido una duración de unos 10 meses, comenzando en agosto de 2021 y terminando en mayo/junio de 2022. Su realización ha sido completamente remota y con horarios muy flexibles. Se usó el servidor de Discord de Tremor⁹ como plataforma principal para comunicarse, tanto por texto como por videollamada. Se programó una llamada por semana, en la que explicaba mi progreso y recibía ayuda de mis mentores en caso de que me hubiera quedado atascado en algún momento.

Disponía de tres mentores, que me guiaban en el proceso de desarrollo: Darach Ennis (*Principal Engineer and Director of Tremor Project*), Matthias Wahl (*Staff Engineer*) y Heinz N. Gies (*Senior Staff Engineer*), todos empleados por Wayfair.

La organización de forma más estructurada para las tareas que tenía pendientes, en las que estaba trabajando en ese momento, y las que ya había realizado, se basó principalmente en un Kanban en GitHub¹⁰.

⁹<https://discord.com/invite/Wjqu5H9rhQ>

¹⁰<https://github.com/marioortizmanero/tremor-runtime/projects/1>

1.4.2. Desarrollo

Para reducir el coste de desarrollo y asegurarse de que el proceso sea completamente seguro (en memoria y concurrencia), el sistema de plugins aprovecha librerías existentes en Rust y herramientas como macros procedurales. El sistema de compilación usado es solución oficial de Rust: Cargo, que también incluye un *formatter*, *linter*, y extensiones instalables creadas por la comunidad. Adicionalmente, existe una gran cantidad de tests y *benchmarks* que se han de tener en cuenta para mantener el *Code Coverage* (la cantidad de código cubierta por los tests) y el rendimiento.

1.4.3. Recursos públicos

Este trabajo está disponible públicamente al completo. Además, a medida que he investigado e implementado el sistema de plugins, he ido escribiendo todo en mi blog personal, *NullDeref*; gran parte de los contenidos de este documento se han obtenido de ahí. Tiene una serie con un total de 5 artículos, donde entro en gran detalle y suma unas 2 horas adicionales de lectura.

- El repositorio de GitHub para el binario de Tremor:
<https://github.com/tremor-rs/tremor-runtime>
- Mi *fork*, con ramas adicionales usadas durante el desarrollo:
<https://github.com/marioortizmanero/tremor-runtime>
- Mi repositorio con experimentos antes de implementar la versión definitiva:
<https://github.com/marioortizmanero/pdk-experiments>
- La serie de artículos en mi blog personal:
<https://nullderef.com/series/rust-plugins/>.

Capítulo 2

Entendiendo Tremor

2.1. Sistemas de Procesado de Eventos

Tremor es un *Sistema de Procesado de Eventos*, que consiste en “el monitorizado y análisis (procesado) de *streams* de información (datos) sobre cosas que pasan (eventos)” [8]. Tremor fue creado como una alternativa de alto rendimiento a herramientas como *Logstash* [9] o *Telegraf* [10], pero ha evolucionado para soportar casos de uso más complejos. Al contrario que esas piezas de software, Tremor también tiene soporte para *agregación* y *rollups*, e incluye un lenguaje *ad hoc* para *Extract, Transform, and Load (ETL)* y *queries*.

Para más información sobre Tremor se puede consultar *Tremor Getting Started* [11], que contiene una explicación de sus conceptos más básicos y sus posibles usos — o cuándo *no* usarlo, en *Tremor Constraints and Limitations* [12]. *Tremor Recipes* [13] lista un total de 32 ejemplos de cómo configurar y emplear el software.

Capítulo 3

Tecnologías para implementar un sistema de plugins

Capítulo 4

Implementación

Capítulo 5

Conclusiones

Bibliografía

- [1] Chris Sweeney. *Inside Wayfair's Identity Crisis*. 2019. URL: <https://www.bostonmagazine.com/news/2019/10/01/inside-wayfair/>.
- [2] *Rust Survey Results*. 2021. URL: <https://blog.rust-lang.org/2022/02/15/Rust-Survey-2021.html#challenges-ahead>.
- [3] Carliss Young Baldwin y Kim B Clark. *Design rules: The power of modularity*. Vol. 1. MIT press, 2000.
- [4] NGINX. 2004. URL: <https://www.influxdata.com/time-series-platform/telegraf/>.
- [5] *Apache HTTP Server*. 1995. URL: <https://httpd.apache.org/>.
- [6] *Dynamic Modules Development*. 2016. URL: <https://www.nginx.com/blog/dynamic-modules-development/#demand>.
- [7] *Dynamic Shared Object (DSO) Support*. 2021. URL: <https://httpd.apache.org/docs/2.4/dso.html#advantages>.
- [8] David C Luckham. *Event processing for business: organizing the real-time enterprise*. John Wiley & Sons, 2011.
- [9] *Logstash*. 2011. URL: <https://www.elastic.co/logstash/>.
- [10] *Telegraf*. 2015. URL: <https://www.influxdata.com/time-series-platform/telegraf/>.
- [11] *Tremor Getting Started*. 2021. URL: <https://www.tremor.rs/docs/next/getting-started/getting-started>.
- [12] *Tremor Constraints and Limitations*. 2021. URL: <https://www.tremor.rs/docs/next/ConstraintsLimitations>.
- [13] *Tremor Recipes*. 2021. URL: <https://www.tremor.rs/docs/next/recipes/README>.

Lista de Figuras

Lista de Tablas

Anexos

Anexos A

Un anexo

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur? At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere