



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Escola Tècnica
Superior d'Enginyeria
Informàtica

Escola Tècnica Superior d'Enginyeria Informàtica
Universitat Politècnica de València

Implementación y despliegue de un proxy inverso para una arquitectura de microservicios

TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

Autor: Alejandro Carrión Sanmartín

Tutor: Patricio Letelier Torres

Curso 2020-2021

Resumen

????

Palabras clave: ?????, ???

Resum

????

Paraules clau: ????, ?????????

Abstract

????

Key words: ?????, ?????

Índice general

Índice general	IV
Índice de figuras	V
<hr/>	
1 Introducción	1
1.1 Motivación	2
1.2 Objetivos	3
1.3 Estructura del documento	3
2 Estado del arte	5
2.1 <i>API Gateway</i>	5
2.2 <i>Kong</i>	5
2.3 <i>Tyk</i>	6
2.4 <i>Ocelot</i>	6
2.5 <i>NGINX</i>	6
2.6 Comparativa	7
2.7 Tecnología utilizada	7
2.7.1 Ejemplo de uso básico	7
3 Desarrollo de la solución	9
3.1 Especificación de requisitos	9
3.1.1 Diagrama de casos de uso	9
3.1.2 Pruebas de aceptación	9
3.2 Diseño	9
3.2.1 Diagrama de componentes	9
3.2.2 Diagrama de clases	9
3.3 Programación	10
3.3.1 Construcción de prototipos	10
3.3.2 Consolidación del microservicio autogenerado	10
3.3.3 Primeros despliegues	10
3.4 Pruebas	11
3.4.1 Pruebas de regresión	11
3.4.2 Pruebas finales	11
4 Conclusiones	13
Bibliografía	15

Índice de figuras

- 1.1 Esquema de un *proxy* frente a un *proxy* inverso. Imágenes de <https://www.ionos.es/digitalguide/servidores/know-how/que-es-un-servidor-proxy-inverso>. 2
- 1.2 Arquitectura de microservicios básica frente a una con *proxy* inverso. Imágenes de <https://www.adictosaltrabajo.com/2020/05/27/introduccion-al-api-gateway-pattern>

CAPÍTULO 1

Introducción

El uso de arquitecturas de microservicios es una práctica ya consolidada hoy en día en el mundo del desarrollo de *software*. Esta consiste en la construcción de servicios independientes, ejecutados en procesos diferentes, que se encargan de realizar funciones concretas y que trabajan de forma conjunta para lograr el objetivo u objetivos globales de la aplicación que constituyen. Los beneficios que otorga este enfoque frente a la aproximación tradicional monolítica son muchos y muy variados. Algunos de ellos son:

- **Uso de diferentes tecnologías.** Cada microservicio puede estar construido con una tecnología diferente y puede utilizar distintos mecanismos de persistencia.
- **Maniobrabilidad en los despliegues.** Ante cualquier cambio no es necesario desplegar la aplicación entera, solamente los microservicios implicados.
- **Tolerancia a fallos.** La posibilidad de desplegar la aplicación de forma que quede repartida en diferentes máquinas, incluso duplicando microservicios, otorga cierta capacidad para tolerar fallos.
- **Escalabilidad y mantenibilidad.** Los microservicios, y la separación funcional que otorgan, facilitan el escalado de las diferentes partes de la aplicación de manera independiente. Lo mismo sucede con el mantenimiento, pudiendo crear equipos especializados.

Se puede obtener más información acerca de los microservicios en el artículo de James Lewis y Martin Fowler titulado "*Microservices*" [1]. Para una lectura con más profundidad, el libro "*Building Microservices* [2]" de Sam Newman.

Por contra, este tipo de arquitecturas aumentan la complejidad del desarrollo en algunos aspectos concretos como pueden ser el versionado de los microservicios o la coordinación de las comunicaciones entre ellos. Es aquí donde entra este trabajo, pues está enfocado a paliar otra de sus desventajas: la exposición de múltiples puntos de entrada al *backend* formado por los microservicios. Para ello, se pretende desarrollar un *proxy* inverso que sea su puerta de acceso.

Para comprender correctamente que es un *proxy* inverso es conveniente ver su relación con su patrón hermano: el *proxy* de reenvío, o *proxy* a secas. Un *proxy* es un componente intermediario que se encarga de proteger una red cliente haciendo que estos clientes no tengan comunicación directa con los servidores a los que se conectan a través de Internet. Por otro lado, un *proxy* inverso hace la misma función pero protegiendo a un grupo de servidores, ocultándolos así de sus clientes. Ambos pueden coexistir, de hecho suelen hacerlo. Para ilustrar mejor esta diferencia, la figura ?? presenta un esquema conceptual de un *proxy* y otro de un *proxy* inverso.

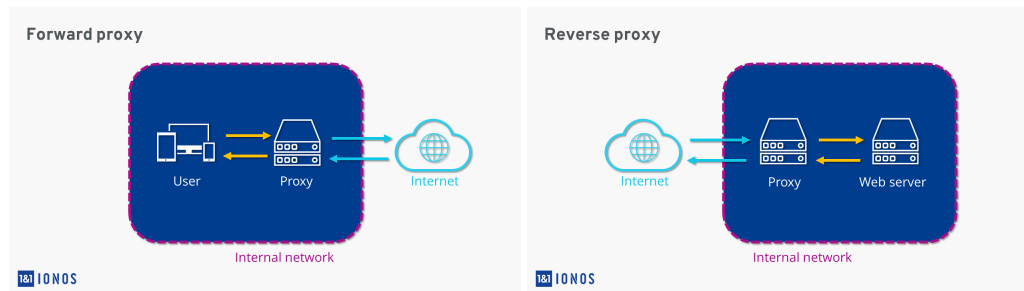


Figura 1.1: Esquema de un *proxy* frente a un *proxy* inverso. Imágenes de <https://www.ionos.es/digitalguide/servidores/know-how/que-es-un-servidor-proxy-inverso>.

Como se puede inferir de las definiciones anteriores, el uso de un *proxy* inverso no queda restringido a ciertas arquitecturas, pudiéndose utilizar para ocultar el servicio o servicios que consume cualquier aplicación. Sin embargo, este componente adquiere una gran importancia en el enfoque de microservicios, pues es importante no exponer estos al exterior. Además, se suele utilizar también para realizar tareas de balanceo de carga. La figura ?? muestra la comparativa de una arquitectura de microservicios básica y otra que utiliza el componente comentado. En la segunda se observa que con el uso de un *proxy* inverso los clientes no acceden directamente a los microservicios, ni siquiera los conocen.

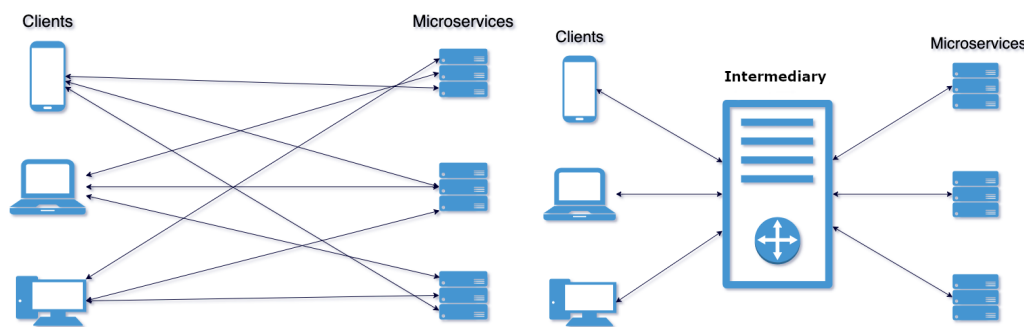


Figura 1.2: Arquitectura de microservicios básica frente a una con *proxy* inverso. Imágenes de <https://www.adictosaltrabajo.com/2020/05/27/introduccion-al-api-gateway-pattern>.

1.1 Motivación

Este proyecto surge en el contexto de una práctica en empresa. El autor ha formado parte del equipo de *I+D+i* de una compañía española del sector sociosanitario durante un periodo de tiempo de más de un año. Esta desarrolla y comercializa un *software* de gestión geriátrica y actualmente se encuentra en proceso de renovación del mismo. El departamento nombrado se dedica a reconstruirlo desde cero con tecnología de vanguardia como el uso de arquitecturas de microservicios, el desarrollo de software dirigido por modelos y la generación automática de código.

La temática del trabajo es el desarrollo de un *proxy* inverso que actúe de intermediario entre la interfaz de usuario de la nueva aplicación en desarrollo y sus microservicios. Se pretende construir este nuevo componente porque se cree necesario proteger el *backend* de la aplicación y ocultar los microservicios que lo forman. Por otro lado, también se quiere tener la posibilidad de lanzar a ejecución múltiples instancias de los microservicios con el fin de conseguir cierta tolerancia a fallos y poder también realizar balanceo de carga.

La temática comentada fue elegida debido a la estrecha relación que guarda con el desempeño del autor en las prácticas mencionadas, esto es, contribuciones a un micro-

servicio específico destinado a orquestar el despliegue de la aplicación. Por otro lado, el desarrollo a llevar a cabo le va a otorgar una visión más global de la aplicación sobre la que se trabaja así como aumentar el nivel de conocimiento acerca de la misma con la motivación de seguir contribuyendo al proyecto por mucho tiempo más. Por último, la tecnología a utilizar, *.NET*, es de su interés y aspira así a crecer como desarrollador de ese *framework*. En cuanto a la motivación de la empresa, esta quiere dar la posibilidad de aprovechar, en cierta medida, el trabajo realizado y a realizar por el autor del documento dentro del marco de trabajo de la entidad, viéndose esta beneficiada también.

1.2 Objetivos

El objetivo principal de este trabajo es construir un *proxy* inverso. La construcción de este tiene las siguientes aspiraciones sobre la aplicación en construcción:

- **Ocultar los microservicios** que forman la aplicación para que la interfaz de usuario no acceda directamente a estos por motivos de seguridad.
- Permitir la **instanciación múltiple** de los microservicios que, a su vez, tiene por finalidad:
 - Conseguir que la aplicación sea **tolerante a fallos**, gracias a la posibilidad de tener un mismo microservicio desplegado en máquinas diferentes.
 - **Aumentar la eficiencia**, al poder crear o parar instancias dinámicamente según el tráfico que reciba la aplicación.

1.3 Estructura del documento

?????

CAPÍTULO 2

Estado del arte

En la actualidad existen en el mercado muchas aplicaciones y servicios que se pueden utilizar como *proxy* inverso. Algunas de estas soluciones son de pago, otras gratuitas e incluso algunas de código abierto y se pueden dividir en dos tipos: productos *software* ya contruidos y librerías. Los primeros suelen ser fáciles de configurar y se pueden poner en marcha de una manera muy rápida. Las segundas requieren una parte de programación pero se adaptan mejor a las necesidades particulares, pues permiten tener más control al usuario.

2.1 *API Gateway*

A modo de aparte, es interesante mencionar un patrón parecido al *proxy* inverso y que también podría haber sido utilizado para resolver el problema que atañe a este trabajo: el *API Gateway*. Ambos comparten algunos casos de uso, por lo que sus diferencias causan confusión y no suelen quedar claras. Generalmente, se entiende que un *API Gateway* es una especialización de un *proxy* inverso, proporcionando así funcionalidades extra. Las más aceptadas e importantes son:

- Interpretan los mensajes que reciben y pueden hacer transformaciones sobre ellos; los *proxy* inversos solo los redirigen donde corresponda.
- Suelen ofrecer agregaciones de peticiones, esto es, aunar dos o más llamadas al *backend* y exponer esta composición a través de un único *endpoint*.
- Realizan tareas transversales a todos los *endpoints* como autenticación, autorización o monitorización.

El patrón recién comentado fue descartado desde el primer momento porque no se pretendía otorgar la funcionalidad de ninguno de los tres puntos clave que caracterizan esta aproximación. Por este motivo, y por la búsqueda de sencillez que se quería, era más coherente decantarse por un *proxy* inverso.

2.2 *Kong*

Tal y como dice su página web, *Kong* [3] es el *API Gateway* más conocido a nivel mundial. Cuenta con las siguientes características:

- Está optimizado para arquitecturas de microservicios.

- Está concebido para ser altamente escalable.
- Ha sido creado sobre el *proxy* inverso *NGINX*, lo que hace que su eficiencia destaque.
- Es gratuito, aunque ofrece la versión *Enterprise*, que facilita la gestión del API Gateway, y la *SaaS*, llamada *Kong Cloud*.
- Posee un gran número de *plugins* desarrollados por la propia compañía y por su comunidad, que es bastante grande al ser una aplicación tan popular.
- Actualmente es utilizado por grandes empresas como *Intel*, *Ferrari* o *Yahoo!*.

2.3 Tyk

Tyk [4] es un *API Gateway* de código abierto.

- Permite alojar sus servicios tanto en la nube como en una infraestructura propia. También ofrece un modelo híbrido.
- Dispone de extensiones para integrarlo con diferentes mecanismos de autenticación como *OAuth* o *LDPA*.
- Necesita una base de datos *Redis*.
- Puede escalar horizontalmente y verticalmente.

2.4 Ocelot

Ocelot [5] es una librería para *.NET Core* que permite a una aplicación de ese *framework* actuar como *API Gateway*. Posee las características siguientes:

- Está pensada para arquitecturas orientadas a servicios o a microservicios.
- Al tratarse de una librería, se utiliza de forma sencilla añadiéndola como un paquete *NuGet* más.
- Su configuración es muy básica, teniendo que especificarla en un fichero *json*.
- No permite cambiar su configuración de manera dinámica.

2.5 NGINX

Originariamente *NGINX* [6] fue construido para ser un servidor web, como por ejemplo *Apache*, pero más tarde ofreció la posibilidad de actuar como *proxy* inverso, balanceador de carga o *proxy* para protocolos de correo electrónico. Desde la vertiente que interesa a este trabajo:

- Se trata de un *proxy* inverso ligero y de alto rendimiento
- Ofrece una versión gratuita y otra de pago, *NGINX Plus*, la cual ofrece funcionalidades extra.
- Se configura a través de un fichero el cual puede ser recargado durante su ejecución, es decir, es configurable dinámicamente.

2.6 Comparativa

?????

2.7 Tecnología utilizada

La tecnología a utilizar para llevar a cabo este proyecto va a ser *YARP* [7]. Se trata de una librería hecha por el propio *Microsoft*. Su objetivo es facilitar la creación de *proxies* inversos. Sus aspectos más destacados son:

- Se encuentra todavía en desarrollo y la única versión lanzada hasta la fecha es una *preview*. Se prevé que vayan saliendo a la luz más versiones con más funcionalidades basadas en la experiencia de la propia empresa pero también en las opiniones de la gente.
- Fue diseñada para construir *proxies* inversos dentro de la propia infraestructura de *Microsoft* ya que en muchas áreas de la compañía necesitaban algo así.
- A raíz de la heterogeneidad de casos de uso que debe cubrir derivado del punto anterior, está pensada para ser muy personalizable y flexible.
- Permite cambiar su configuración de forma dinámica.

2.7.1. Ejemplo de uso básico

?????

CAPÍTULO 3

Desarrollo de la solución

?????

3.1 Especificación de requisitos

- Redirección de peticiones
- Carga dinámica de rutas
- Eliminación de rutas
- Multi-instacia de microservicios
- Versionado de microservicios

3.1.1. Diagrama de casos de uso

?????

3.1.2. Pruebas de aceptación

?????

3.2 Diseño

?????

3.2.1. Diagrama de componentes

?????

3.2.2. Diagrama de clases

?????

3.3 Programación

Para la construcción del *proxy* inverso que protagoniza este trabajo se ha seguido una serie de pasos que se van a comentar a continuación.

3.3.1. Construcción de prototipos

El desarrollo comenzó con la elaboración de dos prototipos de microservicio con la funcionalidad mínima para redirigir peticiones. Estos prototipos tenían la misión de demostrar si era posible seguir la estructura general de todos los microservicios de la aplicación, con gran parte del código generado a partir de modelos. Precisamente este código autogenerado es lo que generaba ciertas dudas con lo referente al rendimiento final, el cual se podía ver perjudicado debido a un código menos específico y más genérico, así como a un exceso de características innecesarias para el caso en cuestión pero que sí debían poseer el resto de microservicios.

El primero de los prototipos se hizo lo más simple posible, es decir, con la lógica justa y necesaria para desempeñar su trabajo y sin seguir ninguna estructura concreta. El segundo, fue construido utilizando las mismas herramientas con las que fueron construidos el resto de microservicios de la aplicación, esto es, generación de código a partir de modelos. Cabe destacar que, para no realizar esfuerzos en vano sin saber qué prototipo iba a ser la opción elegida, se dejaron algunos detalles correspondientes a la generación de código pendientes. Esto se debe a que el proceso de generación de código no se adaptaba al cien por cien a las necesidades del nuevo microservicio.

Una vez construidos los prototipos, se procedió a medir el impacto de seguir la norma para decidir qué prototipo desechar y con cuál seguir adelante.

Los resultados de las mediciones fueron claros: ambos prototipos tenían un rendimiento casi idéntico. De esta forma, se decidió continuar con la versión autogenerada.

3.3.2. Consolidación del microservicio autogenerado

Una vez elegida la opción de seguir el estándar, hubo que resolver los aspectos de la generación de código que no quedaron resueltos. Así pues, se tuvo que modificar algunas plantillas de código a partir de las cuales se generan los microservicios. Por otro lado, también fue necesario abordar otros problemas derivados de seguir el patrón de microservicio estándar.

3.3.3. Primeros despliegues

A estas alturas del desarrollo ya se tenía una primera versión del *proxy* inverso perfectamente funcional pero antes de llevar a cabo los primeros despliegues había que adaptarlo al proceso de despliegue que se seguía hasta el momento. La interfaz de usuario de la aplicación se conectaba directamente a los microservicios por lo que hizo falta interponer el nuevo componente entre ella y el *backend*.

Más tarde, se hizo que los propios microservicios utilizaran el *proxy* inverso para comunicarse entre ellos.

El único microservicio que no se pudo reemplazar fue el de autorización porque tenía una dirección específica, configurada a través del DNS.

Como paso previo a incluir el *proxy* inverso al proceso de despliegue primero se hicieron pruebas de manera local. Se desplegó el *proxy* inverso en una máquina a parte y se le

cargó con rutas que apuntaban a los microservicios ya desplegados para que la interfaz de usuario lo utilizara. Se comprobó que todo funcionaba correctamente.

Un problema que surgió fue el de los timeouts. Dos microservicios tenían un timeout muy grande debido al elevado coste de sus peticiones. El *proxy* inverso cortaba las llamadas antes de que terminaran. Al final esto resultó culpa de *Yarp*, que tiene un timeout por defecto de 100 segundos. La solución fue aumentar el timeout para las peticiones que se dirigen a los microservicios en cuestión. Para ello, fue necesario añadir un parámetro opcional para el timeout en la creación de rutas. De esta forma se permite la personalización del timeout en función de la ruta que se utilice.

3.4 Pruebas

Para comprobar el correcto funcionamiento del *proxy* inverso se han realizado dos tipos de pruebas: unas más sencillas, hechas cada vez que se producía un incremento considerable de funcionalidad, y otras más formales, al final del proyecto.

3.4.1. Pruebas de regresión

Se trata de una serie de comprobaciones definidas formalmente que han servido para verificar que las funcionalidades anteriormente implementadas no dejaban de funcionar al añadir nuevas.

3.4.2. Pruebas finales

Con el propósito de garantizar la corrección del comportamiento del producto *software* desarrollado se llevaron a cabo pruebas finales. A las pruebas de regresión anteriormente comentadas se les añade...

CAPÍTULO 4

Conclusiones

En un futuro se puede querer extender la funcionalidad y convertirlo en un *API Gateway*.

Bibliografía

- [1] Artículo de J. Lewis y M. Fowler acerca de los microservicios. Disponible en:
<https://martinfowler.com/articles/microservices.html>
- [2] Libro de referencia sobre microservicios Sam Newman Building Microservices
- [3] Documentación oficial de *Kong*:
<https://konghq.com/kong>
- [4] Documentación oficial de *Tyk*:
<https://tyk.io>
- [5] Documentación oficial de *Ocelot*:
<https://threemammals.com/ocelot>
- [6] Documentación oficial de *NGINX*:
<https://www.nginx.com/>
- [7] Documentación oficial de *YARP*:
<https://microsoft.github.io/reverse-proxy>