

Flujo de Agregadores de Información y de Aprobaciones Financieras en Colombia

Julián Eduardo Arias Barrera

Noviembre 2021

1 Introducción

Desde siempre la toma de productos financieros (el estudio y toma de decisión) ha sido un proceso extremadamente largo y tedioso tanto para el cliente que solicita un producto de estos como para la entidad financiera al ejecutar el estudio (calcular el riesgo). Se ha logrado llegar al punto de digitalizar procesos de obtención de información a base de datos externas (servicios de terceros) para completar dicha información, pero siguen habiendo procesos manuales, debido a que no se tiene una infraestructura lo suficientemente flexible o transparente que permita optimizar el proceso sin incurrir en gastos adicionales y que deje de ser rentable ser implementado. O bien porque se tiene un sistema que puede realizar dicho proceso de forma digital a completo pero no tiene las prestaciones de la programación funcional y herramientas/frameworks más adecuados, optimizados y entendibles que actuales implementaciones. Se logrará dar con una solución mas adecuada y optimizada utilizando Kafka con Scala.

2 Problema

Para las aprobaciones financieras en este momento se necesitan tres pasos esenciales que son: la obtención de información, transformarlo a una única estructura (también limpiar datos a no utilizar), y su procesamiento en "motores" de riesgo. Poder garantizar la flexibilidad, rapidez (respuesta en lo mas rápido posible) y no incurrir en gastos adicionales será el objetivo.

En la siguiente figura podemos observar la solución que actualmente se tiene y sus redundancias.

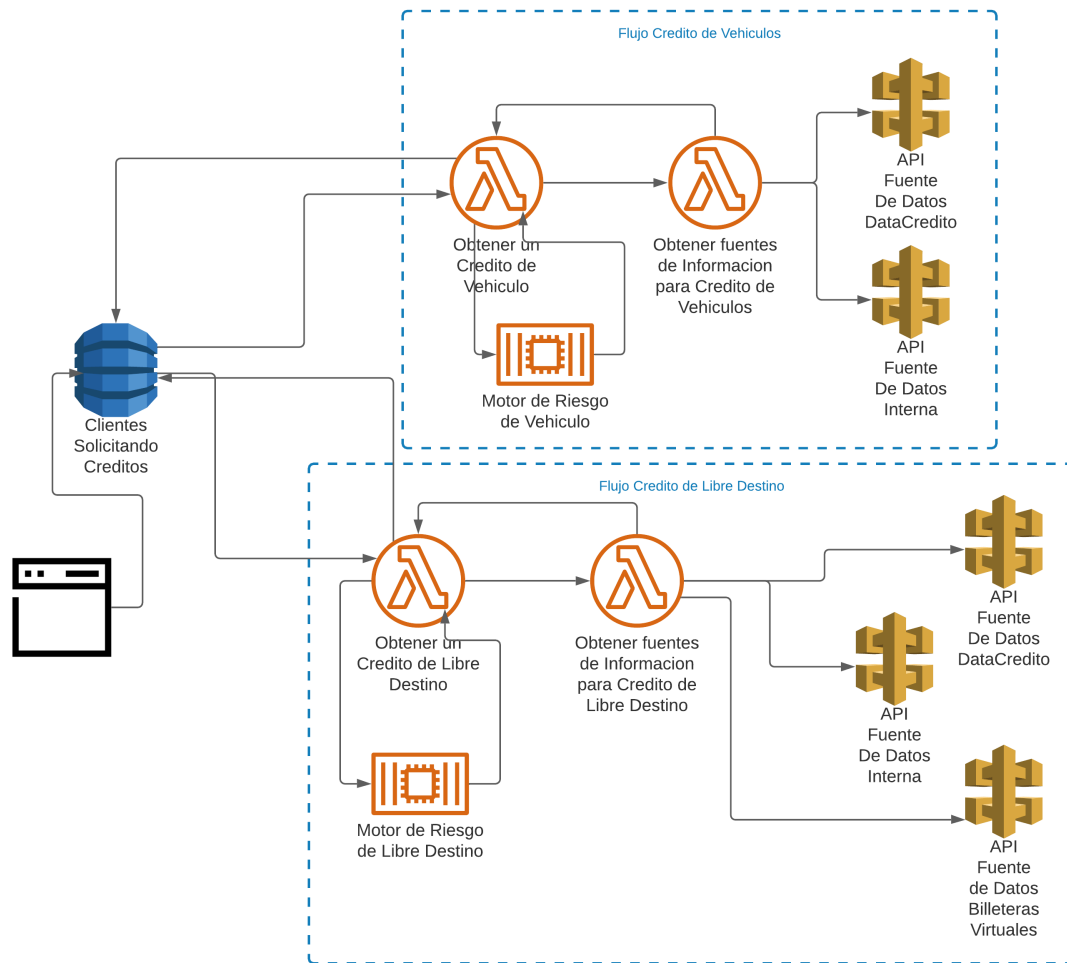


Figure 1: Vieja Implementación

Se tiene aplicaciones por cada tipo de crédito, es decir no es centralizado este proceso, su tiempos de respuesta son demorados, y es difícil que es aplicar concurrencia en algunos lenguajes y restricción de paralelismo que tiene los lambdas.

2.1 La obtención de Información

Cuando un cliente pide un crédito se obtiene información de 3 tipos de fuentes: El que cliente aporta, las externas, y la que la entidad financiera posee. Las entidades consultadas dependerá del tipo de credito como también repetir este proceso relacionado otro cliente si incluye por ejemplo un cotitular o cónyuge en el crédito.

En el caso de la información que el cliente aporta generalmente son datos que no se encuentran en otras fuentes, por ejemplo ganancias de consultorias (freelance por ejemplo) o contratos de prestación de servicios realizados.

Para fuentes externas tenemos las centrales de riesgo (por ejemplo DataCredito), fuentes de información financiera de terceros (Publicas: de recolección de impuestos, o Privadas como base de datos con información de Billeteras Virtuales), alivios o campañas realizadas para inversión o "colchón" (por ejemplo los Alivios ofrecidos a algunas empresas por la pandemia COVID).

Finalmente las internas que son las que la entidad financiera posee, si fue cliente o es cliente : como fue el manejo de sus productos y la información suministrada previamente.

Este paso es complejo ya que en algunos casos se encontraba que cada equipo de producto, por ejemplo "Credito de Vehiculos", "Credito de Vivienda" , "Credito para Pymes" , "Credito para Libre Inversion", cada uno de estos tenian su propia implemnetacion y artefactos ya que para cada uno de estos las fuentes de información requeridas no es siempre la misma y la consulta de cada una de las fuentes de información (APIs) genera costos.

2.2 Empaquetamiento de Información

En este paso el objetivo es hacer lo mas compacto y simple la puesta de parámetros a un "Motor" de aprobación. Por lo tanto las 3, 4, o 5 fuentes de información consultadas en el previo paso deben ser reunidas en un sola estructura, habrán campos que no serán de utilidad que responden cada una de las APIs requeridas a consultar y que no deben ser enviadas al motor de riesgo. Como también agrupar esta información con la de su cotitular o cónyuge relacionado.

Se puede encontrar que este proceso tampoco era centralizado como se quería por la complejidad de interpretar los diferentes fuentes de datos que consulta cada crédito.

2.3 Ejecución en Motores de Riesgo

Los motores de riesgo suelen ser una "caja negra" que es entregada por parte del área de riesgo de la entidad financiera que se supone que a partir de variables del mercado, dado un input de variables sobre un(os) cliente(s) y un tipo de crédito en particular, da una respuesta sobre si el cliente es viable a tomar el tipo de crédito (alguna veces con que monto) o no (pero con algún monto menor). Por lo tanto reduce 50-100 variables a una respuesta sintética.

3 Propuesta

La propuesta para optimizar este proceso y que pueda dar realmente respuesta en tiempo real es utilizar las bondades del paralelismo de Scala y la framework ofrecido por Kafka.

Siguiendo la idea de los 3 procesos esenciales:

3.1 Agregadores

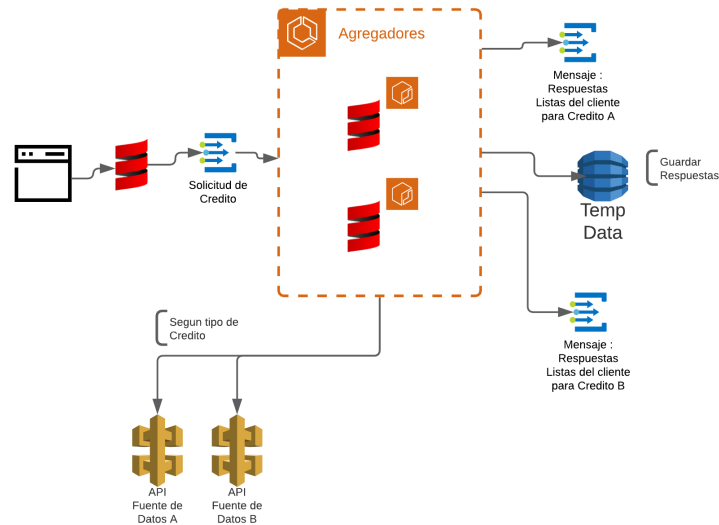


Figure 2: Agregadores

Se tendría un topic inicial donde llega la información básica del cliente y el tipo de crédito que solicita (y información adicional por parte del cliente opcional), estarían suscritos un grupo que llamaremos "Agregadores" los cuales consultarán concurrentemente las APIs necesarias para ese tipo de crédito, y enviarán un mensaje a un topic por tipo de credito cuando estas respuestas estén listas y guardadas en un base de datos No-Sql para ser consumidas por parte del siguiente grupo.

3.2 Empaquetadores

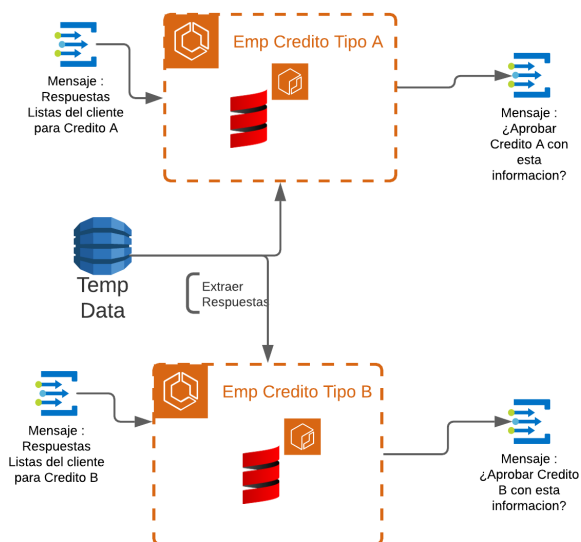


Figure 3: Empaquetadores

Los siguientes grupos los "Empaquetadores", leerán y empaquetaran las respuestas que previamente se guardaron en la base de datos (mapear a diferentes nombres variables y desechar variables que no se utilicen) l y publicaran un mensaje en el topic de aprobación de crédito correspondiente.

3.3 Aprobadores

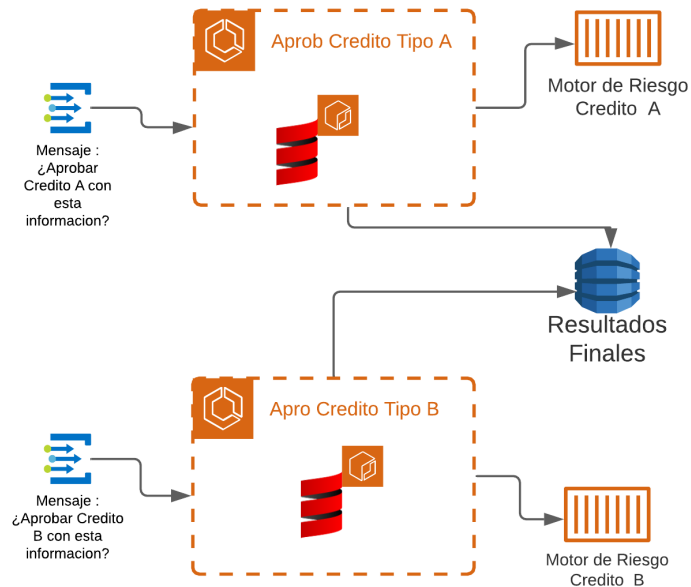


Figure 4: Aprobadores

Y finalmente los últimos grupos que ejecutaran el motor de riesgo según crédito, y alojarán la respuesta en la base de datos.

3.4 Implementación completa

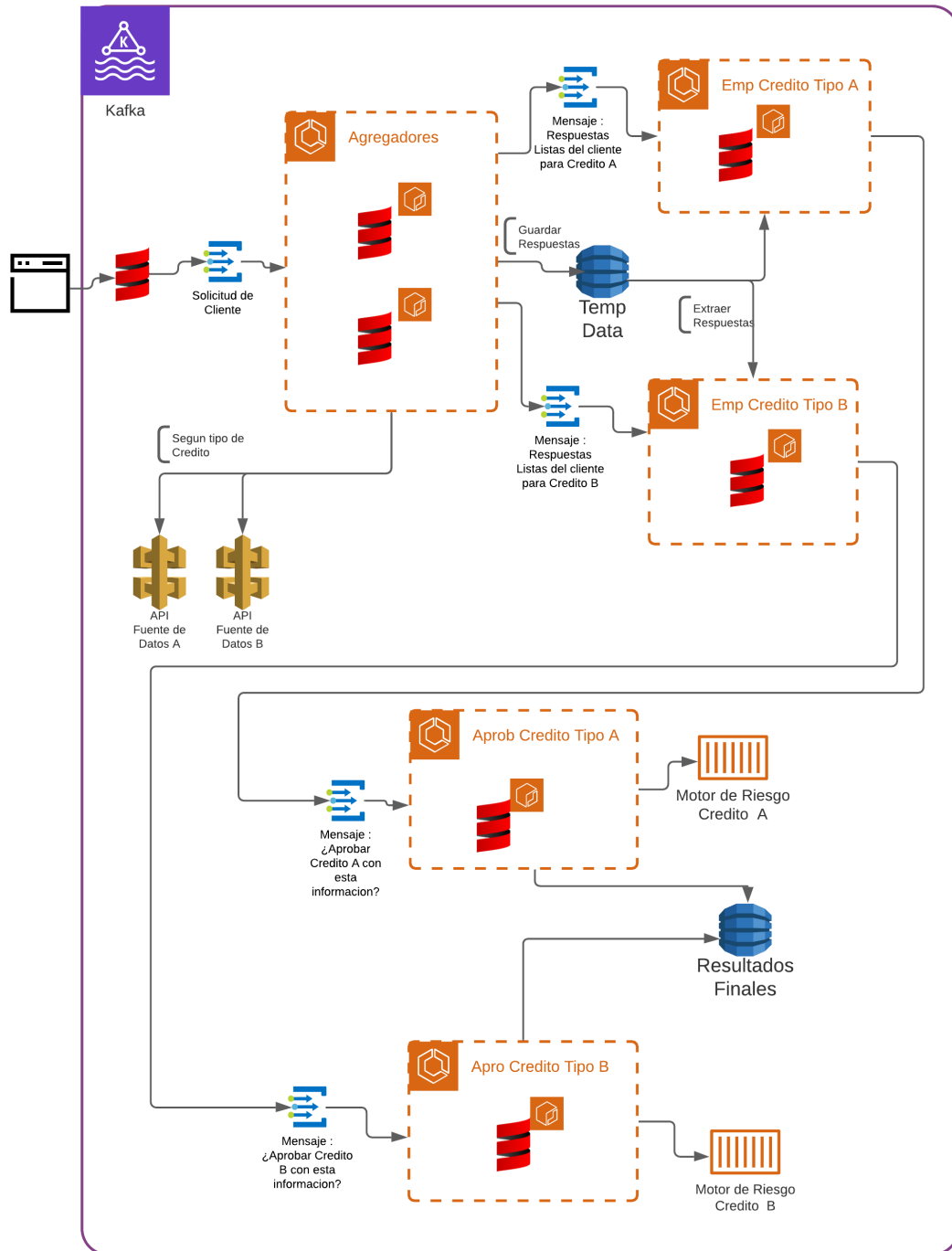


Figure 5: Nueva Implementación

Ya en esta nueva implementación se puede ver centralizado la primera parte de obtención de información, las bondades del paralelismo y la facilidad de supervisión/configuración centralizada que nos ofrece Kafka.

3.5 Flexibilidad

Adicionalmente también mejor facilidad en caso dado que se necesite agregar un nuevo motor o en el caso que se realice campañas (por un tiempo limitado se cambian los parámetros de aprobación de crédito), y tener la posibilidad de hacer pruebas estilo A-B.

Por ejemplo en este flujo final se podría comparar cuantos clientes se aprobaron con campaña y sin campaña (Comparar las tablas de guardado). Por lo tanto se podría comparar para un mismo tipo de crédito cuantos clientes se aprobaron con respecto a otro motor de riesgo.

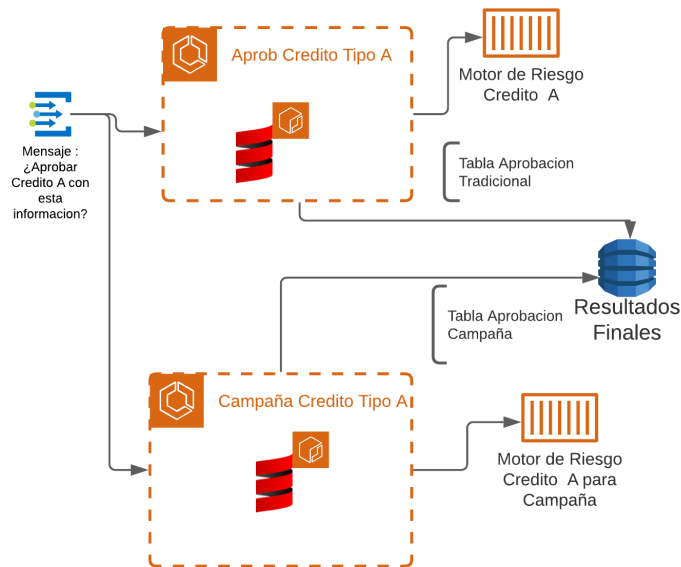


Figure 6: Configuración de dos aprobadores para un mismo crédito

4 Resultados Esperados

Se espera que las respuestas de aprobaciones sean mas rápidas, que sea mas fácilmente escalable la solución, mantenible y monitoriable para todos los productos crediticios, es decir obtener todos los beneficios de un orquestador de clusters, con el fin que los clientes tengan una mejor experiencia, y también para quienes mantienen esta solución si: se quita fuentes, agrega fuentes o cambia los parámetros de las fuentes o procesamiento para uno o varios productos, tener varios tipos de motores para un mismo crédito y tener la capacidad de reutilizar o direccionar fácilmente este flujo.