

Propuesta de proyecto de investigación

Maestría en Ciencias y Tecnologías de la Información

1. Nombre del proyecto

Propuesta de un mecanismo para despliegue del sistema ALEPH usando contenedores e infraestructura como código.

2. Responsable(s)

Dr. Ricardo Marcelín Jiménez

(Depto. de Ing. Eléctrica, UAM Iztapalapa)

rmarcelin@izt.uam.mx

Dr. Humberto Cervantes Maceda

(Depto. de Ing. Eléctrica, UAM Iztapalapa)

hcm@xanum.uam.mx

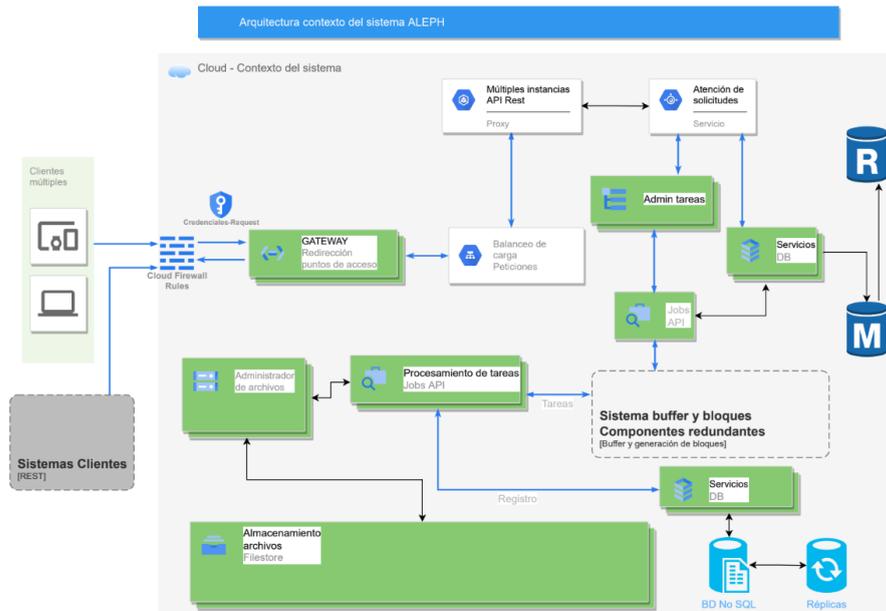
3. Perfil deseable del alumno

El alumno(a) participante en este proyecto debe tener conocimientos de redes de computadoras e ingeniería de software, mostrar facilidad para la programación de algoritmos distribuidos y la lectura de artículos en inglés.

4. Presentación del contexto e identificación de la problemática

En la UAM-I se ha desarrollado un sistema de almacenamiento masivo llamado ALEPH. Dicho sistema es un sistema de cómputo distribuido sobre el que se realizan operaciones de tratamiento de información (incluyendo operaciones tales como cifrado, compresión, generación de redundancia) que tienen como objetivo procesar un archivo que debe almacenarse (y posteriormente recuperarse), para garantizar su seguridad y alta disponibilidad [1,2] .

ALEPH tiene una arquitectura mixta donde parte de los componentes son microservicios y otra parte son extensiones de Hadoop [3], como se muestra en la figura siguiente:



Actualmente, el despliegue de ALEPH se realiza de forma manual, lo cual es una tarea complicada y propensa a errores. Esto es un problema pues existen oportunidades de llevar ALEPH a distintos entornos de operación, dentro y fuera de la universidad, pero la falta de mecanismos modernos de despliegue del sistema complican su implantación y, además, limitan la posibilidad del equipo de desarrollo de adoptar métodos modernos de desarrollo como DevOps [4]. DevOps es un enfoque moderno de desarrollo que busca cerrar la brecha que existe entre los equipos de desarrollo y la operación del sistema en producción. Para ello, se busca automatizar en la medida de lo posible el proceso de despliegue y liberación del sistema.

Adicionalmente, buscamos incorporar un sencillo mecanismo o procedimiento para soportar el requerimiento de escalabilidad que deben ofrecer este tipo de sistemas. Lo anterior significa que cuando el conjunto de dispositivos de almacenamiento llega a un límite de su capacidad se debe agregar un nuevo lote de dispositivos y, de inmediato, invocar un mecanismo para recuperar el balance de carga [2].

5. Objetivos

Objetivo general

El objetivo de este proyecto es investigar, desarrollar y aplicar mecanismos que faciliten el despliegue de ALEPH y el soporte de actividades asociadas con la metodología de desarrollo DevOps.

En particular es deseable:

- Investigar la manera de empaquetar los distintos elementos del sistema en contenedores y de soportar la ejecución del sistema una vez empaquetado.
- Investigar la manera de describir la topología de implantación de Aleph para poder desplegarla usando mecanismos de “infraestructura como código”.

Objetivos específicos:

1. Entender la arquitectura de ALEPH y la manera en que se despliega actualmente.

2. Entender los conceptos necesarios para empaquetar y ejecutar ALEPH en contenedores y para desplegarlo usando un enfoque de infraestructura como código.
3. Proponer un enfoque para soportar el despliegue y ejecución de ALEPH en contenedores y usando un enfoque de infraestructura como código.
4. Construir un prototipo que valide el enfoque.
5. Evaluar la propuesta.

6. Metodología

Act. 1. Realizar un estudio de la literatura y tecnologías necesarias.

Act. 2. Desarrollar propuesta e implementar prototipo.

Act. 3. Evaluar .

Act. 4. Comunicar los resultados.

Calendario

La siguiente tabla presenta un calendario tentativo para el proyecto. Nota: Los meses son meses “efectivos” de trabajo.

	01	02	03	04	05	06	07	08	09	10	11	12
Act1												
Act2												
Act3												
Act4												

7. Resultados

Al finalizar el proyecto el alumno habrá desarrollado y creado un prototipo que facilite la implantación de ALEPH y su desarrollo usando un enfoque de DevOps.

Al término del primer trimestre el alumno entregará al coordinador el protocolo de investigación que, al menos, contendrá una descripción del estado del conocimiento, la problemática que se aborda y la metodología. Este protocolo se presentará en el seminario de avances organizado al cierre del trimestre. Éste contendrá la caracterización de los requerimientos funcionales y no funcionales, la descripción de la arquitectura a usar, así como una propuesta de diseño e implantación, incluida una agenda de pruebas.

Al término del segundo trimestre el alumno entregará al coordinador un reporte de avance en forma de artículo . Este reporte se presentará en el seminario de avances y contendrá los resultados de las pruebas programadas anteriormente.

Al finalizar el tercer trimestre el alumno entregará al coordinador un primer borrador de su idónea comunicación de resultados. De igual forma, el alumno presentará por última vez en el seminario de avances, el resultado de su trabajo.

En el curso del cuarto trimestre el alumno entregará la versión final de su idónea comunicación de resultados.

8. Infraestructura necesaria y disponible

El participante requiere de una estación de trabajo con acceso a internet, así como el acceso al código de ALEPH. A nivel de software, todas las aplicaciones que se utilizarán son gratuitas. Toda esta infraestructura está disponible actualmente.

9. Lugar de realización

Una vez restablecidas las actividades académicas en la universidad los trabajos se podrán realizar en Laboratorio T326B de la UAM-Iztapalapa También, el proyecto puede realizarse en forma no presencial y con acceso remoto a los recursos del laboratorio. De igual forma, el seguimiento se puede realizar usando recursos de comunicación sincrónicos y asincrónicos (i.e. videoconferencia & email).

10. Bibliografía inicial

[1] Yanilos and S. Sobti. The Evolving Field of Distributed Storage. IEEE Internet Computing. September. 2001. pp 35-39.

[2] Quezada-Naquid M., Marcelín-Jiménez R., González-Compeán J. L. Babel: The Construction of a Massive Storage System, IJWSR, Volume 13, Issue 4, October-December 16 (JCR impact factor 0.469) DOI: 10.4018/IJWSR.2016100103.

[3] White, T. (2012). *Hadoop: The definitive guide*. " O'Reilly Media, Inc."

[4] Bass, L. and Klein, J. "Deployment and Operations for Software Engineers: A DevOps Engineering Text", ASIN B09XW7TKFK, April, 2022