 ****

**Propuesta de proyecto**

**Titulación:**

Máster en Software de Sistemas Distribuidos y Empotrados.

**Tipo de proyecto:**

Proyecto de Desarrollo de Ingeniería.

**Título del Proyecto Fin de Master:**

Desarrollo de un algoritmo para agrupación de eventos con solapamiento en el tiempo a través de técnicas distribuidas.

**Tutores(as) Coordinadores(as):**

Isabel Muñoz Fernández, [imunoz@eui.upm.es](mailto:imunoz@eui.upm.es).

Sergio Arévalo Viñuales, [sergio.arevalo@eui.upm.es](mailto:sergio.arevalo@eui.upm.es).

**Tutores profesionales:**

Dusan Skovajsa, [dusan@ayscom.com](mailto:dusan@ayscom.com).

José Alberto Rodríguez, jalberto\_rodriguez@ayscom.com.

**Alumnos:**

Lizandro Ramirez Difo, [lizandro.rdifo@alumnos.upm.es](mailto:lizandro.rdifo@alumnos.upm.es).

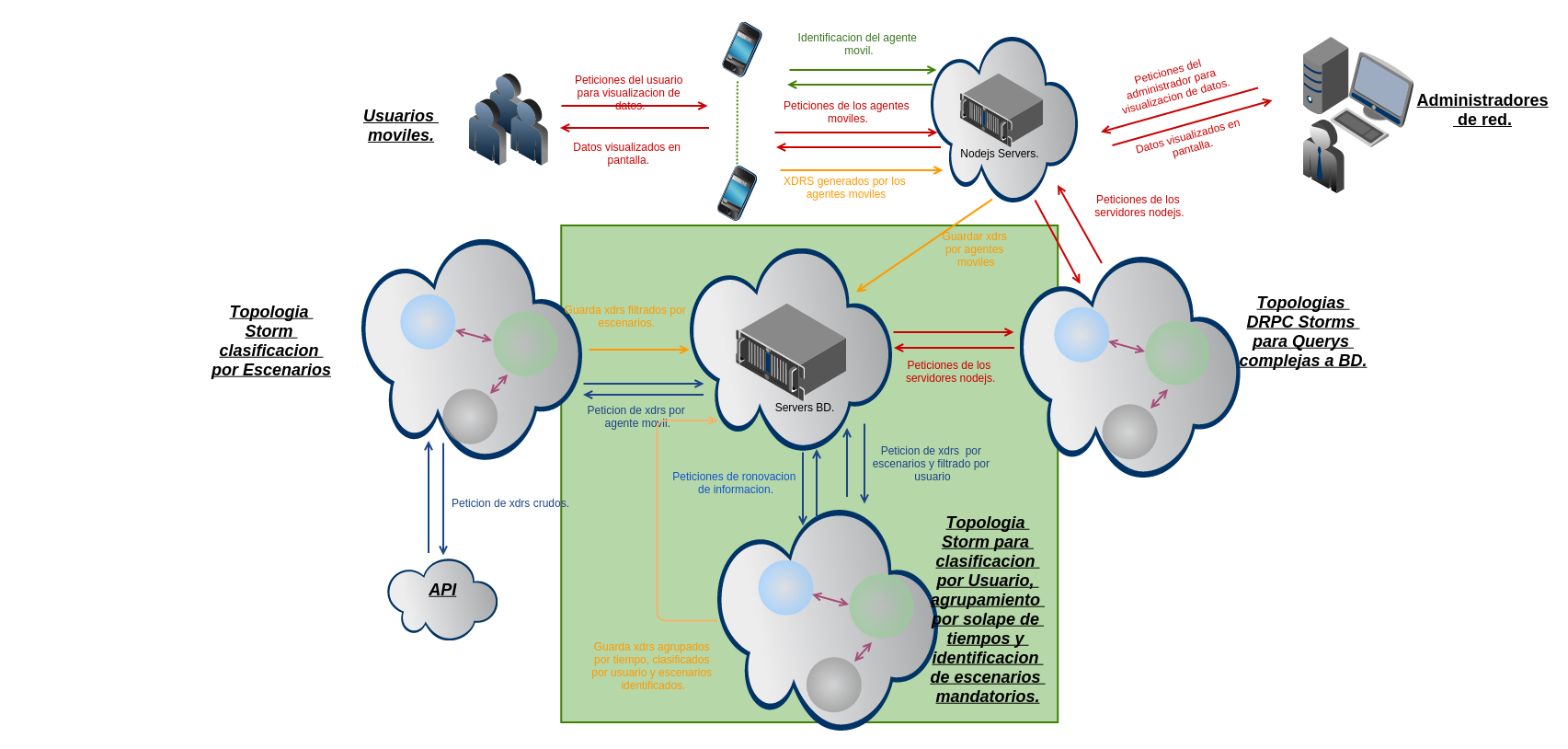
**Objetivos:**

**Objetivo General**

* Desarrollar y evaluar un algoritmo de agrupación de eventos con solape en el tiempo sobre diferentes tecnologías de sistemas distribuidos (Apache Storm, Cassandra y Zookeeper).

**Objetivos Específicos**

* Implementar un algoritmo para agrupación de eventos con solapamiento en el tiempo a través de técnicas distribuidas en un entorno de pruebas.
* Utilizar Apache Storm como sistema de computación en tiempo real distribuido.
* Incorporar Apache Cassandra como sistema de base de datos distribuida con alta escalabilidad para escritura y lectura de información.
* Aplicar Apache Zookeeper para tolerancia a fallos en los sistemas Apache Storm y Cassandra.

**Resumen**

El proyecto presentado en estos párrafos forma un parte del desarrollo de un sistema de captura, procesamiento y visualización de datos de medidas de los distintos parámetros de calidad en la red de acceso de los operadores de telefonía móvil. Esta información se encuentra en el ámbito de Big Data será necesario implementación de tecnologías distribuidas para su tratamiento como son Apache Storm para procesamiento distribuido en tiempo real, Apache Zookeeper para tolerancia a fallos y Apache Cassandra como sistema de base de datos con alta escalabilidad.

El objetivo principal del sistema es la agrupación de los datos obtenidos por los agentes móviles con los datos del núcleo de la red en lo referente a los XDR's para proporcionar una plataforma de diagnóstico de los posibles problemas de los usuarios de las redes móviles. La correlación de estos datos estará basada en el solapamiento en el tiempo de los distintos XDR's registrados, de aquí la importancia en la implementación de un algoritmo que realice este procesamiento e identifique los solapamientos generando grupos solapados en tiempo pertenecientes a un mismo usuario.

Para la realización del sistema se plantea tener los datos generados por el agente móvil y los datos del núcleo de red almacenados en una base de datos NoSQL(Apache Cassandra) denominada BD en crudo, estos datos en crudo serán procesados periódicamente por un conjunto de servidores a los que llamaremos “correladores”. Los correladores generarán diagnóstico de los diferentes problemas que puedan tener los usuarios, correlando los datos de las diferentes fuentes de información almacenados en la base de datos en crudo. También servirán para generar información agregada que se presentará en distintos tipos de informes ejecutivos. La información generada por los correladores se almacenará en la base de datos “correlada” para posteriores consultas y análisis.

Fig. Representacion grafica del sistema de captura, procesamiento y visualización.

La figura anterior representa un bostejo del funcionamiento del sistema completo en la parte de Backend, la parte sobreada con color en el centro de imagen representa el alcance del proyecto para propuesta de Trabajo Fin de Master. El API representa la interfaz con la fuente de datos en nuestro caso la empresa Telefonica que implemente el sistema.

**Palabras Claves:** Celdas telefónicas; Correlación de External Data Representation (XDR's); Apache Storm; Apache Zookeeper; Usuarios móviles; Apache Cassandra; Base de datos Distribuidas; Tolerancia a fallos; Diagnostico Escenarios; NoSQL; Correladores; Base de datos en crudo; Escenarios Mandatorios; (API).

**Alumnos / Duración:**

1 alumno / 6 meses

**Aprobación de la propuesta:**

Fecha: 20/05/2015

VB, Tutor

Fecha: 20/05/2015

VB, Presidente de la Comisión Académica de Postgrado