UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA
CARRERA DE INGENIERIA EN SISTEMAS COMPUTACIONALES

TEMA:

Servicio de logs centralizado para los servidores del Centro de Cómputo de la UCSG, visualizado por medio de una aplicación web dentro de la red interna.

AUTOR
FIALLOS ROMERO LUIS FERNANDO

Trabajo de Titulación previo a la obtención del título de:
INGENIERO EN SISTEMAS COMPUTACIONALES

TUTOR:
Ing. Gallardo Posligua Vicente Adolfo, Mgs.

Guayaquil, Ecuador
13 de marzo de 2018

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

CERTIFICACIÓN

Certificamos que el presente trabajo fue realizado en su totalidad por **Fiallos Romero Luis Fernando** como requerimiento parcial para la obtención del Título de **INGENIERO EN SISTEMAS COMPUTACIONALES**

TUTOR

f. _____

Ing. Gallardo Posligua Vicente Adolfo, Mgs.

DIRECTORA DE LA CARRERA

f. _____

Ing. Guerrero Yépez Beatriz, Mgs

Guayaquil, a los 13 días del mes de marzo del año 2018

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

DECLARACIÓN DE RESPONSABILIDAD

Yo, **Fiallos Romero Luis Fernando**

DECLARO QUE:

El Trabajo de Titulación, **Servicio de logs centralizado para los servidores del Centro de Cómputo de la UCSG, visualizado por medio de una aplicación web dentro de la red interna** previo a la obtención del título de **INGENIERO EN SISTEMAS COMPUTACIONALES**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias bibliográficas. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 13 días del mes de marzo del año 2018

EL AUTOR

f. _____

Fiallos Romero Luis Fernando

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

AUTORIZACIÓN

Yo, **Fiallos Romero Luis Fernando**

Autorizo a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Titulación, **Servicio de logs centralizado para los servidores del Centro de Cómputo de la UCSG, visualizado por medio de una aplicación web dentro de la red interna**, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

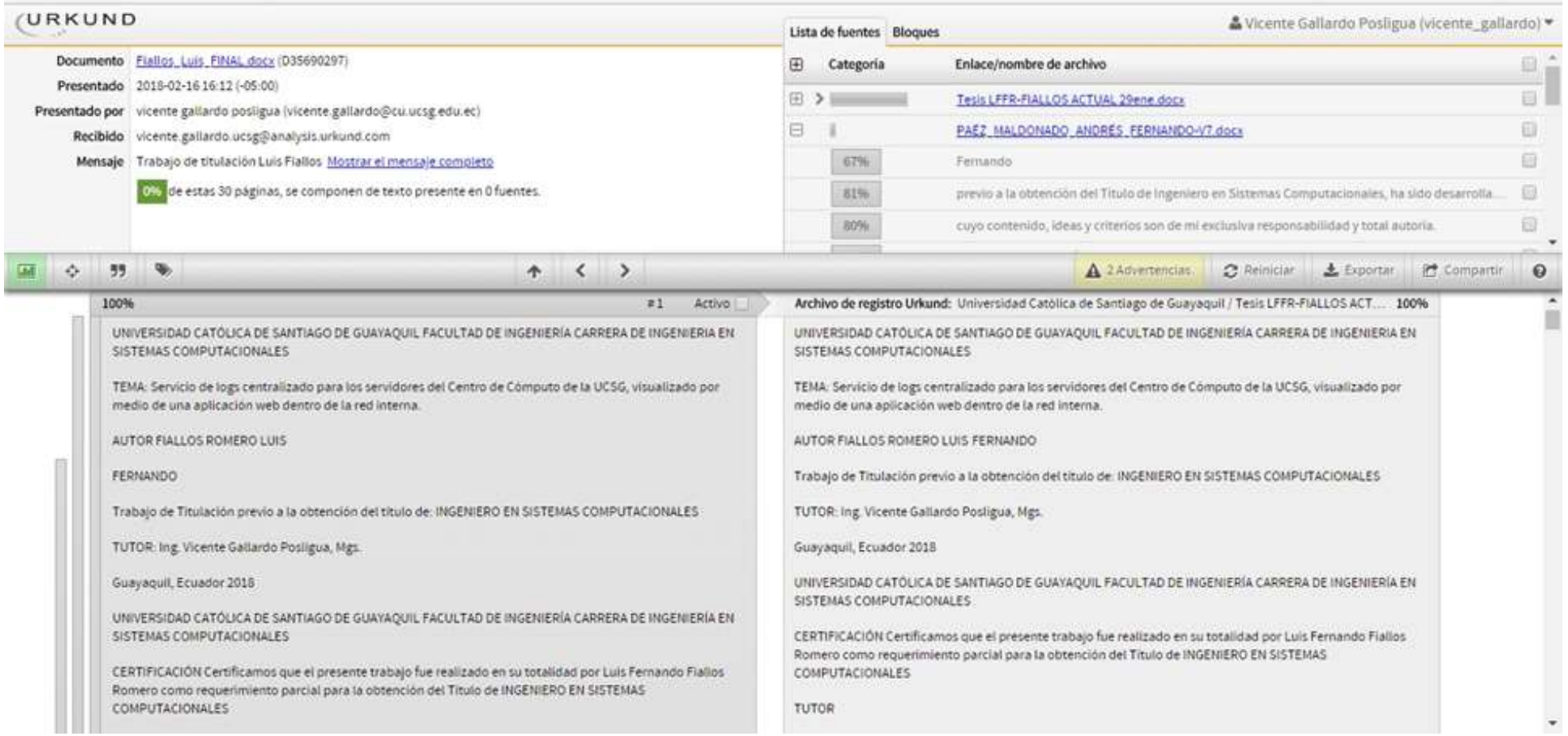
Guayaquil, a los 13 días del mes de marzo del año 2018

EL AUTOR:

f. _____

Fiallos Romero Luis Fernando

REPORTE URKUND



AGRADECIMIENTO

Los resultados de este proyecto, están dedicados a todas aquellas personas que, de alguna forma, han sido parte de mi vida a lo largo de mi carrera universitaria.

A mi mamá, por haberme inculcado desde muy pequeño valores, sobre todo el de la responsabilidad y enseñarme lo importante que es valerse por uno mismo.

A mi papá que a lo largo de toda mi vida ha apoyado y motivado mi formación académica.

Luis Fernando Fiallos Romero

DEDICATORIA

Dedico el desarrollo de este proyecto a mi madre, Laura Romero y a mi hermana Vanessa León, por el apoyo incondicional, consejos y cuidados hacia mi durante cada etapa de mi vida, este logro va dedicado a ellas, ya que depositaron su confianza en cada reto que se ponía enfrente, sin dudar de mi capacidad y es debido a ellas que soy quien soy ahora.

También dedico la culminación de este proyecto a todas las personas que creyeron en mí y me apoyaron a lo largo de mi carrera universitaria.

Luis Fernando Fiallos Romero

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

TRIBUNAL DE SUSTENTACIÓN

f. _____

Ing. Beatriz Del Pilar Guerrero Yépez, Mgs.

DIRECTORA DE CARRERA

f. _____

Ing. Alex Adrián Almeida Campoverde, Mgs.

COORDINADOR DEL ÁREA

f. _____

Ing. Byron Severo Yong Yong

OPONENTE

ÍNDICE GENERAL

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

RESUMEN

Este proyecto abarca la problemática de lo necesario que es para una empresa tener implementada una estructura de centralización de logs, para lograr una mejor visualización y administración de eventos generados por los equipos y servicios manejados por la organización. En la actualidad existen múltiples protocolos y herramientas de visualización y administración de logs centralizados, es por esta razón que se procedió a investigar antecedentes y definiciones básicas del tema en general, con el fin de desarrollar un proyecto de administración y visualización de logs enfocándose en las necesidades del Centro de Cómputo de la Universidad Católica de Santiago de Guayaquil, ya que es el departamento encargado de gestionar la plataforma tecnológica que soporta los diferentes servicios académicos y administrativos dentro de la universidad. Como parte de los objetivos de este proyecto se analizaron los requerimientos y características que fueron solicitados por el Centro de Cómputo para la implementación de un equipo centralizado de logs. Para la selección de la herramienta utilizada se compararon las más conocidas con las características requeridas, implementándose un servidor LogAnalyzer sobre un ambiente operativo CentOS que capturó información de diferentes tipos de equipos y servidores haciendo uso del protocolo SYSLOG. Con la finalidad de que este proyecto sea utilizado por el Centro de Cómputo en un ambiente de producción, se presenta un manual de instalación y configuración de LogAnalyzer para la implementación del servidor centralizado en el Centro de Cómputo considerando los diferentes servidores y servicios que posee.

Palabras clave: LOG; SYSLOG; RSYSLOG; LOGANALYZER; CENTRALIZACION DE LOGS; SOFTWARE LIBRE.

ABSTRACT

This project covers the problem of how necessary it is for a company to have a logs centralization structure implemented to achieve a better visualization and management of events generated by the equipment and services managed by the organization. Currently there are multiple protocols and tools for visualization and administration of centralized logs, that is why we proceeded to investigate background and basic definitions of the topic in general, in order to develop a project for administration and visualization of logs focusing on the needs of the Centro de Cómputo of the Universidad Católica de Santiago de Guayaquil, since it is the department in charge of managing the technological platform that supports the different academic and administrative services within the university. As part of the objectives of this project, the requirements and characteristics that were requested by the Centro de Cómputo for the implementation of a centralized logs team were analyzed. For the selection of the tool used, the most known ones were compared with the required characteristics, implementing a LogAnalyzer server on a CentOS operating environment that captured information on different types of equipment and servers using the SYSLOG protocol. In order for this project to be implemented by the Centro de Cómputo in a production environment, an installation and configuration manual of LogAnalyzer for the implementation of the centralized server in the Centro de Cómputo is presented considering the different servers and services you have.

Keywords: LOG; SYSLOG; RYSLOG; LOGANALYZER; CENTRALIZATION OF LOGS; SOFTWARE OPENSOURCE.

INTRODUCCIÓN

El Centro de Cómputo de la Universidad Católica de Santiago de Guayaquil (UCSG) se encarga de desarrollar, administrar, supervisar y dar mantenimiento tecnológicamente a los múltiples servicios que ofrece la organización para cumplir con los objetivos, metas, estándares y requerimientos que la organización se ha propuesto. Algunos de los servicios que administra el Centro de Cómputo son los del de correo electrónico institucional, internet, intranet, Proxys, VPN, Dominio institucional, entre otros. En total son alrededor de 110 servidores con distintos servicios que el Centro de Cómputo maneja, por lo que sería complicado para el personal encargado de la administración de estos servidores observar cada uno de ellos para poder detectar algún fallo en el servicio, alteración de información, inicio de sesión no autorizado o algún evento crítico que se haya presentado.

Un servidor de logs centralizado permitirá que los eventos que se muestran en los logs de los servidores del Centro de Cómputo de la UCSG se dirijan a este repositorio centralizado para una mejor y rápida administración, complementándolo con una aplicación web dentro de la red del Centro de Cómputo para visualizarlos de una forma más ordenada, esto facilitaría el trabajo y ahorra tiempo al personal encargado de estos servidores para poder identificar problemas y actuar rápidamente a eventos críticos.

Los resultados de esta investigación orientada a ofrecer una solución a la situación planteada, se presentan en este documento como sigue: en el capítulo I se hace referencia a la problemática detectada y todos los elementos necesarios que guían la investigación tales como los objetivos, justificación y alcance. En el capítulo II se menciona los conceptos y definiciones que respaldan la importancia de la implementación de este proyecto. En el capítulo III se hace referencia a los métodos y al diseño aplicado para la elaboración de este proyecto, con el fin de recolectar la información suficiente para la correcta implementación. En el capítulo IV se demuestra la solución tecnológica de este problema, para finalmente ofrecer algunas conclusiones y recomendaciones.

CAPITULO 1

EL PROBLEMA

Los resultados de esta investigación tuvieron como punto de partida la identificación de una problemática a resolver y que se relaciona directamente con el Centro de Cómputo de la Universidad Católica de Santiago de Guayaquil (UCSG), por lo que toda la información relevante con respecto al problema y la solución propuesta se describe en este capítulo.

1.1 Formulación del Problema.

Hoy en día es muy importante estar prevenido de ataques, sesiones sin autorización o manipulación mal intencionada de la información que manejan los servidores que proporcionan diversos servicios en una organización, por eso es primordial hacer un seguimiento constantemente a los logs que estos servidores emiten, para así tener en cuenta si existen vulnerabilidades en el sistema o servidor, también sirve para identificar qué usuarios han ingresado al servidor para interactuar con la información que estos manejan.

Los logs emiten información muy valiosa que puede ser de utilidad al momento de un inconveniente, pero hacer seguimiento a estos logs se hace muy complicado cuando la organización ya posee varios servidores y servicios, y por consecuencia puede generar que se pase por alto identificar algún evento critico que se haya registrado en el log de un servidor, por lo que podría existir fallas o caídas en algún servicio.

El Centro de Cómputo de la UCSG posee más de 110 servidores puestos en producción con diversos servicios, tal cantidad de equipos hace que el seguimiento o la observación del detalle de los logs sea un trabajo bastante complejo de realizar para el personal del área de Producción Informática encargado de la administración de estos servidores, por consecuencia se convierte en una tarea muy compleja el poder identificar y/o prevenir ataques, detectar vulnerabilidades, identificar sesiones, entre otros.

Por tanto, se requiere centralizar los logs de los servidores que han sido seleccionados según su nivel de importancia.

1.2 Justificación

La centralización de los logs ayudará al personal de Producción Informática del Centro de Cómputo de la UCSG a tener una mejor administración de los eventos generados por sus servidores y aplicaciones, ya que se encuentra necesario que se realice esta implementación por la cantidad de equipos y servicios que manejan. Se vuelve muy dificultoso poder revisar y analizar los logs de cada servidor en su respectivo formato, por esa razón también se implementará una aplicación web con una herramienta OpenSource, para poder visualizar de una manera más ordenada y en un solo formato los logs que se han centralizado. Esta implementación mejorará la prevención y detección de ataque, vulnerabilidades, fallos de aplicaciones, entre otros registros.

Para resolver la problemática de la visualización y seguimiento de los logs que presenta el Centro de Cómputo de la UCSG, se implementará un servidor que actuará como repositorio centralizado para los logs que generen los servidores clientes escogidos por el personal del área de Producción Informática del Centro de Cómputo, esto con el fin de obtener un mejor control sobre los eventos que se generan diariamente en los logs de los servidores, ya que si eventos críticos como el fallo de alguna aplicación, el mal funcionamiento del hardware de un servidor, el inicio de sesión no autorizado, entre otros, no se detectan a tiempo y no se toman medidas de prevención, sobre todo en servidores que son indispensables para el Centro de Cómputo, puede generar pérdida de información o mal funcionamiento de los diversos servicios en producción que ofrece el Centro de Cómputo a la comunidad universitaria.

Este trabajo de titulación se enmarca en la línea *uso de software libre,* establecida en el quehacer de la carrera Ingeniería en Sistemas Computacionales de la UCSG.

1.3 Delimitación del Tema

El proyecto pretende cubrir los siguientes puntos:

- Instalar un servidor con sistema operativo CentOS 6.2 que sirva de repositorio central para los servidores clientes.
- Centralizar por medio del protocolo Syslog los eventos generados de los logs de los 5 servidores del Centro de Cómputo escogidos por el personal del área de Producción Informática de la UCSG.
- Instalar la base de datos MySQL Server para almacenar los logs de los servidores escogidos en el repositorio central por un periodo mínimo de un mes.
- Implementar un ambiente web dentro de la red interna del Centro de Cómputo para Instalar la herramienta OpenSource LogAnalyzer que permitirá la visualización de la información generada a partir de los logs centralizados de una manera estructurada.
- Establecer un usuario y contraseña administrador para la autorización y autenticación del inicio de sesión de la aplicación web.
- Realizar búsquedas de logs filtrado por tipo de evento, fecha, servidores.
- Generar reportes básicos con la herramienta LogAnalyzer de acuerdo a lo solicitado por el personal del área de Producción Informática.

1.4 Objetivos

Los objetivos que guiaron esta investigación y permitieron establecer los parámetros para solucionar la problemática planteada, fueron:

1.4.1 General

Implementar un repositorio centralizado para los logs generados por los servidores del Centro de Cómputo de la UCSG, con funciones indispensables para los servicios que ofrece la universidad, a través de herramienta OpenSource, con visualización por medio de una aplicación web en la red interna.

1.4.2 Específicos

- Analizar con el personal del área de Producción Informática del Centro de Cómputo de la UCSG la selección de los servidores que se utilizarán para la recolección de la información de los logs.

- Realizar el diseño e instalación de un servidor con las herramientas OpenSource necesarias para elaborar el proceso de centralización.

- Implementar un ambiente web para la visualización y generación de reportes básicos de todos los logs centralizados elaborados por los servidores clientes escogidos.

Para el cumplimiento de los objetivos establecidos y diseñar la solución al problema planteado, se procedió a realizar un levantamiento de información relativa a situaciones similares y a analizar opiniones de expertos en este ámbito de aplicación.

CAPITULO 2

MARCO TEÓRICO

En este capítulo se pretende abarcar los elementos teóricos y conceptuales que sustentan el uso de los logs como alternativa de solución a posibles intervenciones exteriores en la información de una institución. Se incluye también algunos elementos legales.

2.1 Fundamentación Teórica y Conceptual

La teoría con la cual se desarrolla este proyecto se basa principalmente en las definiciones claves para la implementación de un repositorio de logs centralizados. Esto cubre los temas que se relacionan con el Big Data, Data machine, los logs, método para el transporte de logs (Protocolo Syslog), métodos para asegurar el transporte de logs (Rsyslog), herramientas de análisis y administración que ayuda a la visualización y organización de los logs que han sido centralizados.

2.1.1 Big Data & Machine Data

Para hacer referencia al **Big Data & Machine Data**, es necesario reconocer que en la actualidad las medianas y grandes empresas se enfrentan diariamente a situaciones particulares en sus sistemas y/o servidores, la información que éstas manejan van creciendo conforme la empresa vaya ampliándose o manteniéndose en el tiempo, por lo que en estos días el poder analizar, descubrir y entender los reportes de la información que los equipos y los sistemas manejan de manera estándar se ha vuelto un poco complicado por el gran volumen de registros que hay que observar en cada uno de los servidores o sistemas que las empresas manejan.

Lo primero que hay que tener en cuenta para entender más sobre el manejo de gran volumen de datos es: ¿Qué es Big Data? Y ¿Qué es Machine Data?

Big Data abarca “La gestión y análisis de grandes volúmenes de datos que no pueden ser analizados de manera convencional, ya que superan los límites y capacidades de las herramientas de software habitualmente

utilizadas para la captura, gestión y procesamiento de datos”(Barranco Ricardo & IBM, 2012).

Big data cubre la necesidad de consolidar la gran cantidad de información que se presenta en una organización para ayudar en la toma de decisiones (Begoña & Díez, 2016).

Por otro lado, Machine Data tiene relación con “La información digital creada por la actividad de computadoras, teléfonos móviles, sistemas integrados y otros dispositivos en red. Dichos datos se hicieron más frecuentes a medida que avanzaban tecnologías como la identificación por radiofrecuencia (RFID) y la telemática. Más recientemente, los datos de la máquina han ganado más atención a medida que ha ido creciendo el uso de Internet y otras tecnologías de gestión de Big Data”(Rouse, 2014).

Existen diferentes tipos de datos que se generan de distintas formas como los de audio, video, imágenes, datos generado por sensores o alarmas como los de los automóviles, temperatura, movimiento, entre otras, que nos sirven para analizar y poder tomar medidas sea de prevención o simplemente para tener un control sobre nuestros dispositivos o sistemas, es importante tener en cuenta que además del gran volumen de datos que se generan en distintos equipos, existe una gran variedad de tipos de datos por lo que el procesamiento de estos registros debe tener una gran velocidad, caso contrario se puede pasar por alto una falla, vulnerabilidad o alteración del servicio que el sistema o equipo ofrece.

Cada organización maneja diversos tipos de datos según al entorno en el que estén dedicados, por lo que es necesario establecer por medio de políticas y procedimientos qué datos son los más importantes o críticos para la empresa, de esta manera se puede enfocar en prevenir posibles problemas o pérdidas de información en los sistemas o equipos principales que utiliza el personal de la organización.

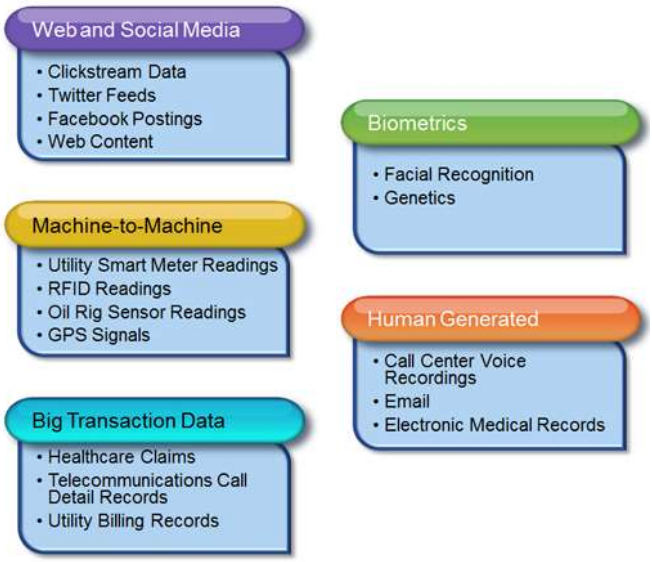


Figura 1: Tipos de datos de Big Data. Adaptado de “Not Your Type? Big Data Matchmaker On Five Data Types You Need To Explore Today”, Por (Soares, 2012).

1.- Web and Social Media: Este tipo de datos incluye todo contenido que se genera en el internet e información que se obtiene por medio de las redes sociales como lo son Twitter, Facebook, LinkedIn, entre otros. Nos sirve para crear pautas en las cookies, sobre todo en cookies de terceras personas que visitan a nuestra web o redes sociales y rastrear sus usuarios e interacciones realizadas.

2.- Machine-to-Machine (M2M): M2M Son las tecnologías que permiten conectar a los sistemas de una organización con otros dispositivos en diferentes lugares por medio de una red cableada o inalámbrica. M2M utiliza diferentes herramientas que capturan algún evento generado por el sistema y lo transporta a otras aplicaciones que traducen estos eventos en información significativa.

3.- Big Transaction Data: Incluye registros de facturación, reclamos, registros detallados de las llamadas realizadas o recibidas, etc. Estos datos son enfocados a ser transaccionales y están elaborados principalmente en formatos como lo son el semi-estructurado y el no estructurado. También se puede aplicar cuando se habla de metadatos, calidad de datos, privacidad y administración del ciclo de vida de la información.

4.- Biometrics: Es la Información biométrica generada por equipos en la que están incluidas las huellas digitales, escaneo de la retina, reconocimiento facial, etc. Estos datos son usados principalmente por el área de seguridad, para identificar usuarios.

5.- Human Generated: Diariamente las personas generamos múltiples cantidades de datos como grabaciones de voz, mensajes de texto, correos electrónicos, encuestas, entre otras, las cuales pueden ser información importante para una organización.

En la actualidad los datos de máquina son uno de los activos más importantes que posee cualquier organización, por lo cual son los más utilizados al momento de realizar algún tipo de levantamiento de información, auditoria o simplemente para determinar el correcto funcionamiento de un equipo o sistema. Cabe mencionar que gran parte de la información más importante que puede conseguir la organización se encuentra escondida en estos datos: El ¿Dónde se produjeron fallos?, cómo optimizar y garantizar la experiencia del usuario e identificar manipulaciones indebidas de la información de la empresa. Toda esta información la cual es muy importante para toda empresa, se las puede capturar en los Data Machine que son generados por las operaciones diarias que se realizan en una empresa.

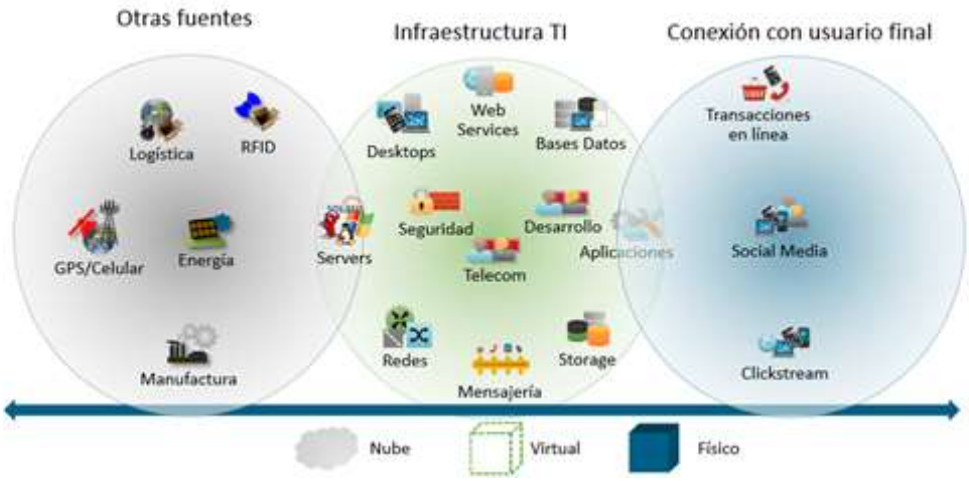


Figura 2: La información empresarial es machine data. Adaptado de “Presentación de la herramienta Splunk”, por (Splunk, 2013).

Los Data Machine son muy importantes para una organización porque contienen un registro definitivo de todos los eventos donde se detalla la actividad y el comportamiento de los clientes, los usuarios, las transacciones, las aplicaciones, los servidores, las redes entre otras. Incluyen configuraciones, datos de API, colas de mensajes, eventos de cambio, resultados de comandos de diagnóstico, registros de detalles de llamadas, datos de sensores de sistemas industriales, etc.(Splunk, s/f-a)

Un evento es una notificación creada por algún equipo o sistema informático que contiene información de una actividad o situación que haya ocurrido. Estos eventos pueden contener información crítica sobre alguna actividad en específico que no debe pasar por alto, pero la mayoría de veces pasa desapercibida por el personal de la organización.

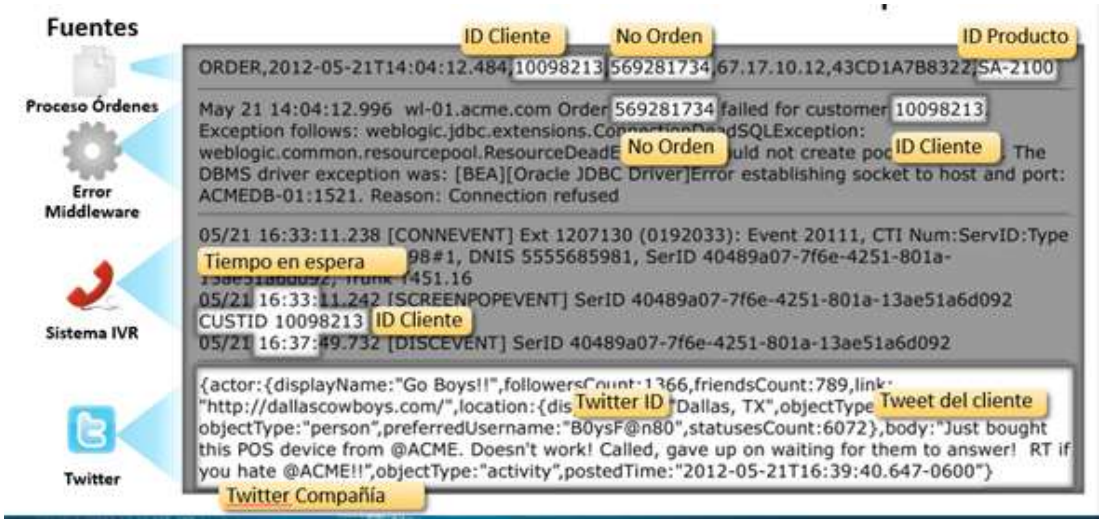


Figura 3: Información crítica en los datos de máquina. Adaptado de “Presentación de la herramienta Splunk”, por (Splunk, 2013).

Si las informaciones de estos eventos no se analizan constantemente puede ocasionar un incidente (conjunto de eventos de seguridad no planificados) en los equipos o sistemas, lo que puede ocasionar perdida de información o mal funcionamiento de aplicaciones. Por esta razón se debe establecer políticas o algún método para analizar lo eventos que son generados por los logs que emite cada servidor o sistema.

2.1.2 Logs

Los logs son registros de los eventos que se generan en los servidores, aplicaciones, redes y sistemas de una organización. Cada uno de estos archivos contiene información relacionada a un evento específico que ocurrió dentro de un equipo, sistema o red.(Vieda, 2013)

Hay que tener en claro ciertas definiciones y conceptos para entender un poco más sobre los logs:

- La palabra “Log” significa “bitácora” en español.

- El log es un registro oficial de eventos que se guardan en un equipo por un periodo de tiempo establecido.

- Cuando un evento ocurre el log generado debe responde las siguientes preguntas: ¿Quién?, ¿Qué?, ¿Cuándo?, ¿Dónde? y Por qué?

- Los logs también son considerados como ficheros de texto donde se guarda información o datos importantes como inicios de sesión, conexiones remotas o eventos generados por el servidor o sistema.

- **Administración de log:** Es el proceso en el cual se efectúa la generación, transporte, almacenamiento, el estudio, monitoreo y reportes de los logs.

- **Análisis de log:** Es el análisis que se le realizan a los logs para identificar los eventos que interesan al administrador o descartar eventos de menos importancia.

- **Retención y reducción de logs:** Los logs generados por los equipos y sistemas son almacenados por un periodo de tiempo establecidos en las políticas de la empresa, esto con el fin de archivar los eventos para el posterior análisis.

- **Rotación de log:** Cerrar un log registrado y abrir otro nuevo generado por el equipo o sistema de acuerdo al tiempo establecido en las políticas y el espacio de almacenamiento del repositorio centralizado.

  Una gran cantidad de los logs son almacenados o desplegados en un formato estándar, el cual es un conjunto de caracteres para dispositivos y aplicaciones comunes. De esta forma cada uno de los logs generados por un dispositivo en particular puede ser interpretado y almacenado en otro equipo diferente (Gómez, 2014).

  Los logs eran utilizados anteriormente para ayudar en la búsqueda y solución de errores en los sistemas, hoy en día con el aumento de la cantidad de información y procesos que maneja cada organización los logs poseen varias funcionalidades que ayudan al personal encargado de la seguridad de TI de la organización como lo son los procesos de optimización en los sistemas, registro de las actividades de los usuarios, detección de vulnerabilidades en equipos o en la red, entre otras.

2.1.2.1 Administración de logs

Toda organización debe manejar sus políticas y proceso sobre los logs, dependiendo de las necesidades y objetivos de la empresa. Se tiene que priorizar el tipo de información que le interesa revisar contantemente al personal encargado de los servidores o servicios para que esta información este integra y disponible por un periodo de tiempo establecido en las políticas.

Existen algunas formas de poder administrar los logs de los servidores, equipos, sistemas, pero son 2 los más importantes o los más significativos:

Un solo servidor central que sirva como repositorio para la administración de Logs que sean importantes o que deban ser analizados constantemente. Al tener solamente un servidor de logs puede ser blanco fácil a un atacante, por lo que se podría poner en duda la integridad de la información, para evitar esto se deben tomar medidas de protección, como la creación de roles de usuario para el análisis y la administración de los logs, bloquear puertos que

estén abiertos innecesariamente, colocar una protección de antivirus de ser necesario, entre otras cosas que aumenten la seguridad de este servidor que actuara como único repositorio de logs.

Otro de los métodos para administrar los logs es implementar diferentes servidores actuando de repositorios que almacenan los logs de acuerdo a una clasificación establecida por la organización.

Esta forma de administrar los logs, tendrá una inversión más fuerte para la organización, ya que se manejan varios equipos repositorios dependiendo de cómo se estableció la clasificación y la distribución de los logs, también se tiene en cuenta por medida de seguridad que cada equipo repositorio tenga un clon en una red diferente en caso de falla o alteración de la información en uno de los servidores de logs.

2.1.2.2 Análisis de logs

La fase del análisis de logs, quizás sea el proceso más complejo para el personal encargado de los servidores de una organización, ya que, si no se tiene establecido un método para administrar los logs, esta tarea se puede volver difícil de realizar y más aún si la organización maneja múltiples servidores con distintos servicios.

Existen 3 conceptos importantes cuando hablamos del análisis de los logs:

Correlación de eventos: Se enfoca en encontrar la relación entre dos o más entradas que posea un log. Por lo general la correlación consiste en reglas establecidas que determinan que eventos están relacionados, incluso si tienen diferentes fuentes. Un ejemplo puede ser las peticiones a un servidor web desde una misma dirección IP, generando entradas en los logs del Firewall, Balanceador de Carga, SO del servidor web y la propia aplicación web. (Vieda Manuel, 2013)

Visualización: Los logs tienen distintos formatos dependiendo del sistema, equipo o servidor que se esté analizando, en la mayoría de veces la información de estos logs no es tan amigable (Ver figura 4) para el usuario que se encargara de revisar los eventos, por lo que puede pasar por alto

información valiosa en muchos casos. Es necesario tener una visualización ordenada y clara de todos los eventos que han generados nuestros logs de los servidores que han sido centralizados, para así evitar que se pase por alto un detalle importante que pueda comprometer el correcto funcionamiento de los servicios de la organización.

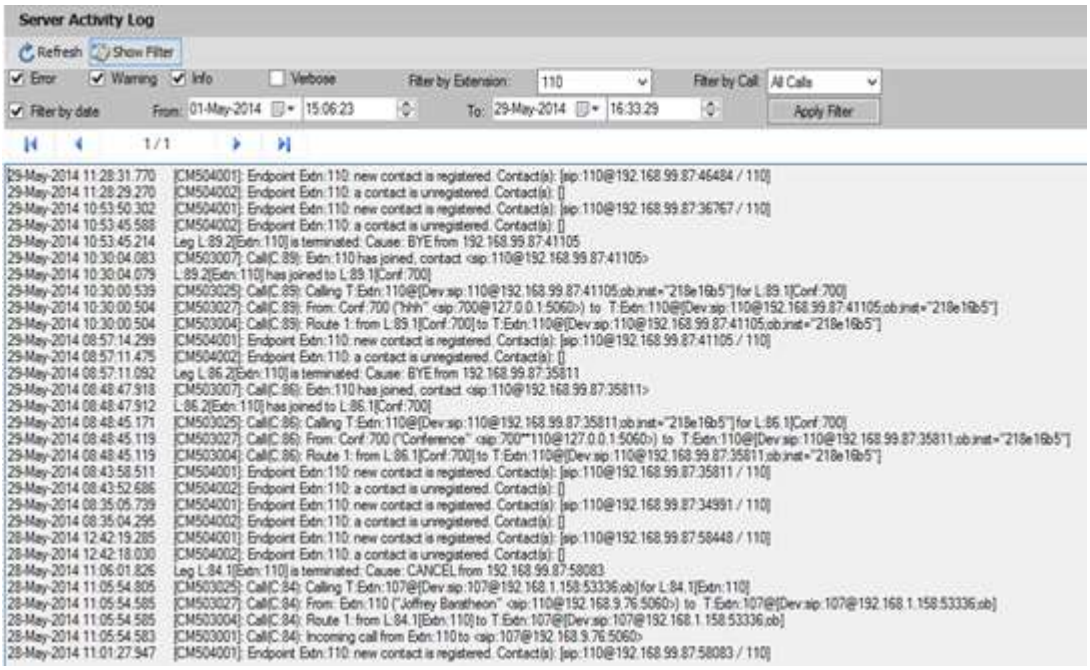


Figura 4: The 3CX Server Activity Log. Adaptado de “Using the 3CX Server Activity Log to Troubleshoot Issues”, Por (Psyllos Andeas, 2012).

Reportes: Los reportes permite visualizar al administrador la información de los logs que ha sido generado por el análisis que realizo el repositorio centralizado de logs. En la mayoría de ocasiones los reportes se realizan a partir de un rango de fechas, un evento en particular que se quieran identificar o por servidor administrado. El administrador debe elaborar o visualizar un reporte a partir de eventos (Información, Noticia, Advertencias estados críticos).

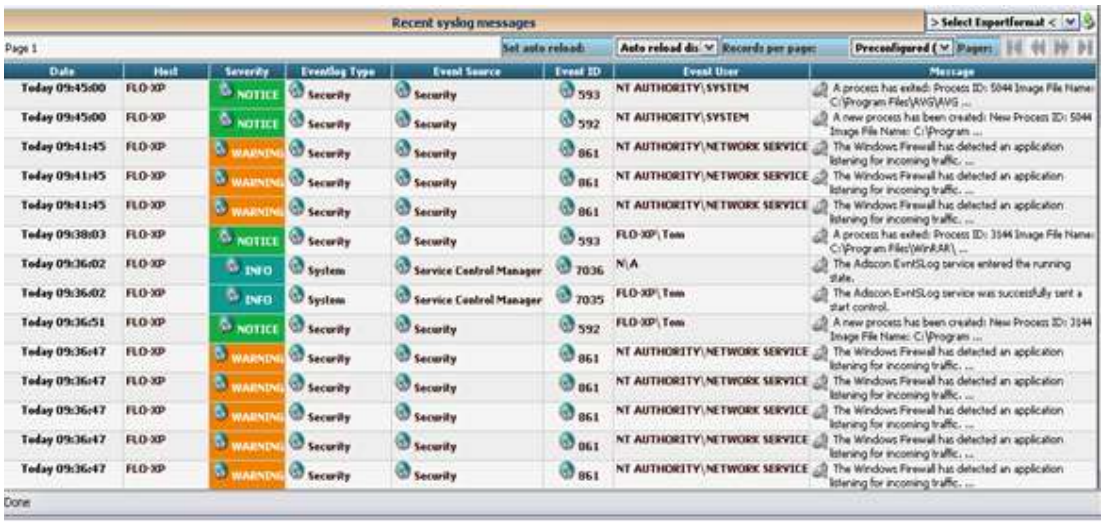


Figura 5: Recent syslog messages. Adaptado de “Instalando Rsyslog y web log analyzer en debían Wheezy - Monitoreo Centralizado de LOG”, por (MikrotikPeru, 2016).

2.1.2.3 Retención y reducción de logs

En la mayoría de casos el gobierno de cada ciudad o país obliga a las organizaciones a respaldar la información generada por los logs por un periodo de tiempo establecido, en medios de almacenamientos removibles (discos duros, cintas) o en servidores de respaldo esto con el fin de que la información manejada por la empresa esté disponible en caso de presentarse algún tipo de investigación.

Existen varios métodos para realizar el almacenamiento de logs, pero principalmente hay un procedimiento para archivar la mayor cantidad de logs de cierta forma en que se aproveche todo el espacio disponible en los medios de almacenamiento.

La retención de logs consiste en guardar los tipos de logs que han sido establecidos por algún procedimiento operativo de la empresa. Esto con el fin de preservar la información que ya debe ser descartada de la organización y que puede servir para una situación en particular como un tipo de investigación o el manejo de incidentes dentro de la empresa.

Si se retiene los logs en un medio de almacenamiento hay que tener en cuenta el espacio que estos necesitan, por lo que es necesario llevar a cabo un proceso de compresión de archivos, esto con el fin de minimizar el

espacio que requieren los logs y mejorar el volumen de información que se maneja.

Después de un periodo es necesario ir descartando logs antiguos para dar entrada a nuevos logs, por lo que es necesario el método de la “Reducción de logs” que consiste en eliminar eventos que ya no son necesarios para la empresa, eliminando las entradas de logs que son ya innecesarios al momento de que se crean nuevos logs.

2.1.2.4 Rotación de logs

La rotación es un proceso que consiste en crear nuevos logs cuando se determina que el archivo ya está completo. Particularmente este proceso se realiza en un periodo determinado (Cada hora, día, semana, mes.) o cuando el archivo ha alcanzado un tamaño determinado. Es muy importante que las organizaciones establezcan el método de rotación ya que tiene algunas ventajas, como lo es la preservación de la información en caso de que un archivo se dañe. Los archivos también se vuelven fácilmente manejables por el sistema operativo y las aplicaciones que administran los logs. Una ventaja más es poder realizar un análisis en los archivos que tienen un menor tamaño y menos información.

Para realizar una administración y un análisis eficiente de los logs que se van a centralizar en un repositorio, es necesario implementar un protocolo para el envío de los registros de cada servidor o máquina hacia el o los equipos que serán utilizados como repositorios, el protocolo más conocido para realizar el proceso de centralización es el Syslog.

2.1.3 Syslog

Syslog es un estándar utilizado para la captura, el procesamiento y el trasporte de mensajes de registro del sistema, es decir los logs del sistema. Es tanto un protocolo de red como a la aplicación o biblioteca compartida que sirve para procesar y mensajes de registro del sistema.(Gómez, 2014)

Syslog es una herramienta que se creó principalmente para sistemas operativos Linux, pero también existen versiones para diferentes sistemas

operativos como lo son Microsoft o Solaris por medio de herramientas que simulan un protocolo syslog, con el fin de poder centralizar los logs de diferentes equipos sin importar el sistema que tengan instalado.

Los servidores que soportan el protocolo Syslog para el envío de sus eventos hacia un repositorio centralizado, tienen la capacidad de transmitir cualquier tipo de mensaje, por lo general los mensajes que se envían contienen eventos relacionados a la seguridad de los equipos o aplicaciones, errores, advertencias, entre otros.

2.1.3.1 Configuración de un repositorio centralizado Syslog

El protocolo Syslog como se ha mencionado, ayuda en la centralización y consolidación de los archivos logs generados por equipos y aplicaciones.

La configuración del Syslog no es muy compleja, se necesita configurar un servidor Syslog que es identificado como Syslogd que es el Daemon de Syslog, el equipo cliente debe enviar un mensaje de texto que tenga un tamaño menor a 1024 Bytes. Una vez establecido el servidor Syslog, se provee una interfaz API para aplicaciones y mensajes, es los que estarán establecidos diferentes niveles de seguridad y se crea grupos de mensajes que estarán seccionados por diferentes tipos (Begoña & Díez, 2016).

Este protocolo puede ser configurado en servidores con sistemas operativos Windows, Linux o también en equipos de red como Switches y Routers.

En los servidores tanto clientes como el central el puerto UDP 514 debe estar abierto ya que, por este medio, los mensajes saldrán de los servidores clientes y entrarán del servidor centralizado.

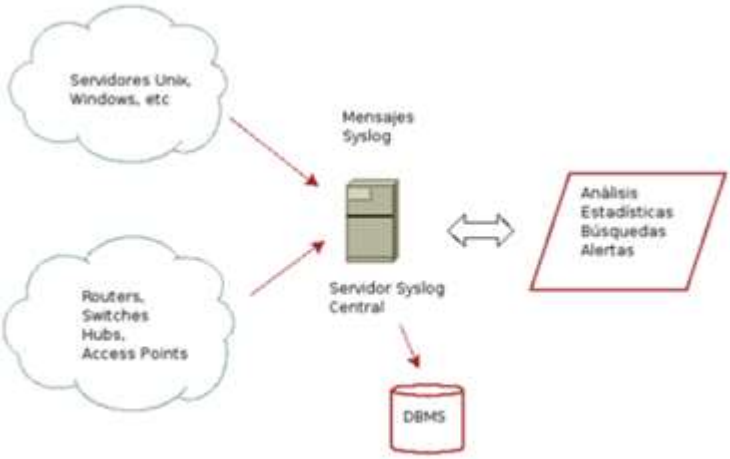


Figura 6: Servidor de log central. Adaptado de “Gestión de Logs”, por (Begoña & Díez, 2016)

2.1.3.2 Rsyslog

El protocolo Syslog es sencillo de implementar por lo que su seguridad no es tan elevada. Syslog se puede complementar con una herramienta llamada Rsyslog, que implementa el protocolo Syslog y le agrega filtros, con una configuración que es fácil administrar. Este complemento aumenta la seguridad que el protocolo Syslog no lo ofrece, por lo que se vuelve necesario de implementar al momento de activar el protocolo.

Sysklog y Syslog eran los predecesores de Rsyslog, las anteriores versiones de Rsyslog no eran tan completas como ésta, ya que no permitían el ingreso de los datos desde diferentes fuentes, no transformaban los datos y tampoco se podía enviar los resultados a varios destinos. Rsyslog es versátil, fácil de usar y completo que puede ser utilizado tanto en grandes, como pequeñas organizaciones.

El Rsyslog permite guardar los archivos de log tanto en formato de texto como un mensaje o algo más administrable como una base de datos MySQL, también se puede elegir una ruta alternativa en caso de que se presente alguna falla en la ruta normal de los logs desde los clientes hacia el servidor centralizado.

Para realizar la configuración del Rsyslog se debe trabajar con ciertos comandos en el archivo /etc/rsyslog.conf. Para realizar la conversión del

dominio Rsyslog en un repositorio de logs remoto se debe implementar un plugin imudp ya que para realizar este proyecto utilizaremos puertos UDP - 514.

Los mensajes y los niveles de los logs que han sido centralizados por medio de los protocolos Syslog y Rsyslog muestran un nombre que hace referencia al tipo de evento que se ha generado esto con el fin de identificar rápidamente qué situación o que evento se debe analizar primero en caso de que se deban revisar los logs generados por los servidores y aplicaciones de la organización. Adicionalmente al mensaje que emite un Syslog, también se envían dos atributos incluidos en dichos mensajes que son: Facilidad y nivel. En las figuras 7 y 8 se menciona algunos mensajes Syslog diferenciados por su atributo.

LOG_AUTH	Mensajes de seguridad/autenticación (descontinuado)
LOG_AUTHPRIV	Mensajes de seguridad/autenticación (privado)
LOG_CRON	Servicio CRON
LOG_DAEMON	Daemons del sistema
LOG_FTP	Daemon FTP
LOG_KERN	Mensajes del Kernel
LOG_LOCAL[0-7]	Reservados para uso local
LOG_LPR	Sub-sistema de impresión
LOG_MAIL	Sub-sistema de correo
LOG_NEWS	Sub-sistema de noticias USENET
LOG_SYSLOG	Mensajes generados internamente por Syslogd
LOG_USER (default)	Mensajes de nivel de usuario genéricos
LOG_UUCP	Sub-sistema UUCP

Figura 7: Grupos Syslogs (Facilidades). Adaptado de “Gestión de Logs”, por (Begoña & Díez, 2016)

LOG_EMERG	Sistema en estado inútil
LOG_ALERT	Se requiere acción inmediata
LOG_CRIT	Condiciones críticas
LOG_ERR	Condiciones de Error
LOG_WARNING	Condiciones de precaución
LOG_NOTICE	Condición normal, pero significativa
LOG_INFO	Mensaje informativo
LOG_DEBUG	Mensaje de depuración

Figura 8: Grupos Syslogs (Niveles). Adaptado de “Gestión de Logs”, por (Begoña & Díez, 2016)

En la mayoría de casos el protocolo Syslog debe ir acompañado de una herramienta que permita la administración y visualización de los logs que han sido centralizado, con el fin de identificar y realizar las búsquedas de los logs según su tipo, servidor, fecha, etc.

Existen varias herramientas pagadas u OpenSource para la administración de los logs, las cuales analizaremos a continuación.

2.1.4 Herramientas de análisis y visualización de logs centralizados.

Las herramientas para la visualización de los logs que han sido centralizados por los protocolos Syslog y Rsyslog son necesarias para llevar una administración de logs de manera ordenada y fácil de interpretar, existen muchas herramientas que nos ayudan con este propósito, pero se escogerá 3 herramientas para analizarlas y concluir posteriormente en cual será implementada en este proyecto.

2.1.4.1 Herramienta LogAnalyzer

LogAnalyzer es una herramienta de código libre que ayuda a analizar las entradas de *"Sterling Configurator Visual Modeler"* **debs.log**. La herramienta proporciona una vista de indicadores que son claves en el análisis de logs: Memoria utilizada, solicitudes y sesiones de los servidores, así como tiempos de respuesta clasificados por usuario y tipo de solicitud que se han realizado desde un equipo hacia el servidor.

LogAnalyzer es un interfaz web para el análisis y la administración de mensajes syslog y eventos de Linux y Windows mantenido por Adiscon (también proporciona el mantenimiento de Rsyslog). Utiliza como repositorio nativo la base de datos MySQL.(Ferrer & Ortega, 2013)

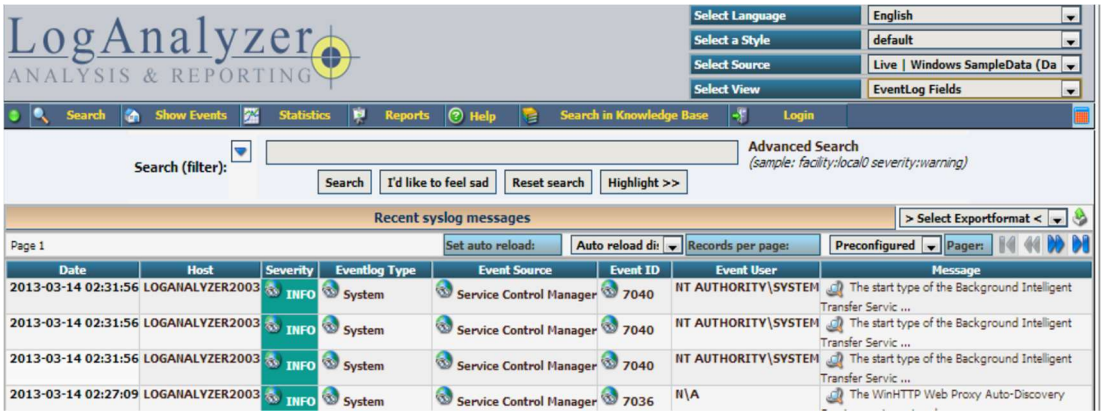


Figura 9: Interfaz web LogAnalyzer. Adaptado de “Registro, Centralización y Análisis de Eventos en un entorno Corporativo Multiplataforma”, por (Ferrer & Ortega, 2013)

Las características más indispensables de la herramienta LogAnalyzer son las siguientes:

- Selección de análisis
- Selección de vistas
- Categorización de los eventos generados
- Búsqueda avanzada por tipología de log, severidad de los eventos, servicios y fuentes.
- Búsqueda general de mensajes UOC.
- Integración nativa del protocolo syslog
- Exportación de reportes en formato CSV y XML
- Estadísticas de eventos que se han producidos, fuentes de origen, tipología y marco temporal
- Generación de informes de eventos producidos, fuentes de origen, tipología y marco temporal
- Integración con la base de datos MySQL
- Integración Apache Web Server

Report Summary

Event Summary	
Total Events	2293
INFO	1173
NOTICE	1012
WARNING	79
ERR	29

Computer Summary
XPTEST(1029), W2003R2(980), W2KTESTING(247), MACHINENAME(37),

Events Consolidated per Host

XPTEST

No.	First Event	Last Event	Process	Type	Event ID	Count
1	2008-09-16 15:14:42	2008-09-16 15:17:58	VMTools	INFO	105	29
	Description		The service was started.			
2	2008-09-16 15:14:48	2008-09-16 15:14:49	Windows Update Agent	INFO	18	23
	Description		Installation Ready: The following updates are downloaded and ready for installation. This computer is currently scheduled to install these updates on Mittwoch, 16. August 2006 at 03:00: - Security Update for Windows XP (KB890859) - Security Update for Windows XP (KB914389) - Security Update for Windows XP (KB920683) - Security Update for Windows XP (KB908519) - Update for Windows XP (KB894391) - Cumulative Security Update for Outlook Express for Windows XP (KB911567) - Security Update for Windows XP (KB896428) - Security Update for Windows XP (KB913580) - Security Update for Windows XP (KB905749) - Security Update for Windows XP (KB908531) - Security Update for Windows XP (KB904706) - Update for Windows XP (KB916595) - Security Update for Windows XP (KB912919) - Security Update for Windows XP (KB900725) - Security Update for Windows XP (KB888302) - Security Update for Windows XP (KB917422) - Security Update for Windows XP (KB901214) - Security Update for Windows XP (KB917953) - Security Update for Windows XP (KB905414) - Security Update for Windows XP (KB917344) - Security Update for Windows XP (KB914388) - Security Update for Windows XP (KB899589) - Security Update			

Figura 10: LogAnalyzer, generación de reportes. Adaptado de “Registro, Centralización y Análisis de Eventos en un entorno Corporativo Multiplataforma”, por (Ferrer & Ortega , 2013)

2.1.4.2 Herramienta Splunk

Splunk es una de las herramientas más famosas para la búsqueda, monitorización y análisis de logs que son generados por los servidores y aplicaciones a través de un ambiente web, Splunk actúa siempre en tiempo real, lo que la herramienta encuentra almacenado en su repositorio se puede visualizar por medio de reportes, gráficos, alertas, entre otros tipos de notificaciones que han sido establecidas por el usuario.

Splunk tiene diversas aplicaciones adicionales para mejorar aún más la administración y análisis de los logs con cualquier formato en el que se genere, sean de aplicaciones, sistemas, servidores, páginas web, sistemas operativos, base de datos, entre otras.

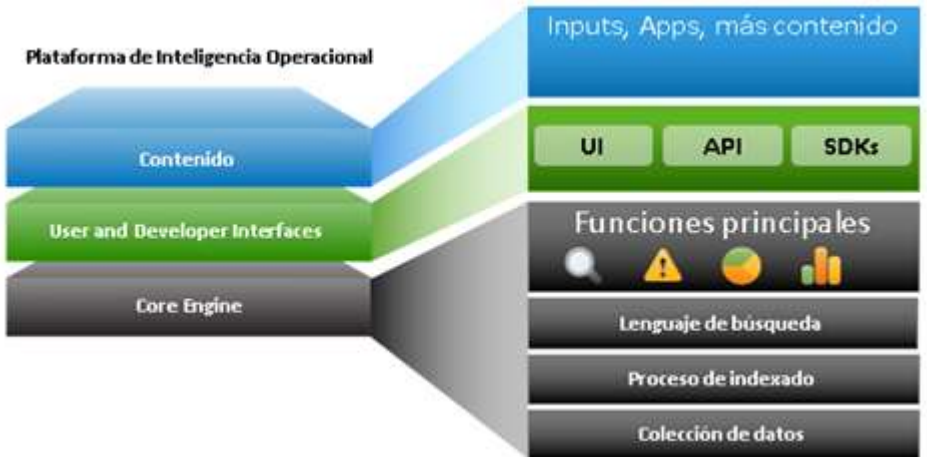


Figura 11: Cómo se compone la plataforma de Splunk. Adaptado de “Presentación de la herramienta Splunk”, por (Splunk, 2013).

La plataforma Splunk Enterprise consta de 2 capas:

- Un motor central y una capa de interfaz.
- Además de una plataforma que puede ejecutar un amplio espectro de contenido que admita casos de uso.

Los casos de uso van desde la aplicación mgmt. y operaciones de TI, cumplimiento de ES y PCI, análisis web, entre otros.

El motor central administra los servicios que son básicos para la entrada de datos en tiempo real, indexación y búsqueda, así como alertas, procesamiento distribuido a gran escala y acceso basado en roles.

La capa de interfaz consiste en la interfaz de usuario básica para la búsqueda, generación de informes y visualización, contiene interfaces de desarrollador, API REST y SDK. Los SDK brindan un acceso conveniente a los servicios básicos del motor en una variedad de entornos de lenguaje de programación. Estas interfaces programáticas permiten:

- Extender Splunk
- Integra Splunk con otras aplicaciones
- Crear aplicaciones completamente nuevas desde cero que requieren OI o servicios analíticos que Splunk proporciona

2.1.4.3 Herramienta Loggly

Loggly es otra herramienta muy completa para la administración y análisis de logs, este software tiene un enfoque a la simplicidad y la facilidad de uso orientado a un público que practica la ingeniería de software (DevOps).

La creación de Loggly fue con el objetivo de ayudar a los DevOps a encontrar y resolver los problemas operacionales. Esto hace que la herramienta sea muy amigable para todo tipo de desarrolladores. Procesos como la elaboración de un desempeño personalizado y paneles de administración DevOps resultan fáciles de realizar.

Una de las características que diferencia a Loggly de las otras herramientas es que tiene un servicio de logging en la nube, donde puede centralizar sus logs, elaborar gráficos, entre otras utilidades.

Algunas de las utilidades que tiene la herramienta de administración de logs son las siguientes:

Monitoreo proactivo: Visualizar el rendimiento de la aplicación, el comportamiento del sistema y la actividad inusual en el equipo. Controlar los recursos y las métricas claves, y eliminar los problemas antes de que afecten a los usuarios finales.

Solución de problemas en registros: Traza los problemas hasta identificar la causa raíz. Ver cómo interactúan los componentes, identificar las correlaciones y compartir los resultados con el equipo.

Integraciones DevOps: Ayuda a trabajar mejor en los equipos usando datos y análisis dentro de las herramientas de DevOps. Loggly se integra con Slack, HipChat, GitHub, Jira, New Relic, PagerDuty, entre otras.

Análisis e informes de datos: Analiza y visualiza los datos para responder preguntas claves, seguir el cumplimiento del SLA y detectar tendencias. Loggly también simplifica la investigación y los informes de KPI.

2.2 Fundamentación Legal

Ya que este proyecto se realizará utilizando herramientas libres (Software OpenSources) es importante mencionar la institución legal que respalda la utilización de estas herramientas con fines de investigación y conocimiento.

El Instituto Ecuatoriano de la Propiedad Intelectual (IEPI) por medio de la Dirección Nacional de Derechos de Autor está fomentando el uso de políticas de licenciamiento libre para de este modo garantizar el acceso al conocimiento y la investigación.

Adicionalmente el IEPI ha establecido una política institucional que permita el paso de los equipos de software licenciado a software libre, por medio de un plan de migración. Es así que actualmente un porcentaje de los equipos de los equipos con los que trabaja la institución ya están funcionando con software libre, esto con el fin de dar el primer paso a que las organizaciones utilicen software libre para implementar sus proyectos e investigaciones a su manera, y no regirse por un software de propietarios.

Es necesario apreciar al software libre como una manera o una herramienta que ayuda a implementar o mejorar una investigación que nos proporcione una aportación para la sociedad y no verlo como si fuera algo gratis que podemos distribuir sin ningún fin específico. Como indica Richard Stallman en su libro *Software libre para una sociedad libre:*

El software libre (OpenSource) es una cuestión de libertad no de precio. Para entender mejor el concepto debemos pensar en la acepción de libre como “libertad de expresión” y no como “barra libre de cerveza”. Con software libre nos referimos a la libertad de los usuarios para ejecutar, copiar, distribuir, estudiar, modificar y mejorar el software. Nos referimos a 4 clases de libertad para los usuarios de software. (Stallman & Lessig, 2004, Pg. 31).

CAPITULO 3

METODOLOGÍA DE LA INVESTIGACIÓN Y ANÁLISIS DE RESULTADOS

Este capítulo abarca todo lo referente a las características de la investigación realizada, metodología de trabajo, análisis de la información recogida y los resultados logrados.

3.1 Tipo de investigación

Este proyecto se basará en la implementación de un servidor donde se instalará un software en ambiente web que permitirá la administración de logs, por lo que se va a trabajar con los paradigmas de investigación que tiene la Ingeniería de Software según como lo mencionan en el paper académico *Research in software engineering: an analysis of the literature* (Glass, Vessey, & Ramesh, 2002), de los cuales cual se elegirá y analizara el **paradigma de investigación descriptivo** que está basado en 3 diferentes supuestos, los cuales serán descritos a continuación:

Supuesto Ontológico: El software que se está estudiando ya existe y lo que se busca en la investigación es conocer cómo funciona.

Supuesto epistemológico: La persona que elabora la investigación debe ser un observador para así conseguir el conocimiento suficiente del tema a investigar.

Supuesto metodológico: Se basa en la experimentación, observación, análisis y uso de las estadísticas o reportes, con el fin de validar los resultados que se han obtenido del proyecto desarrollado.

Con el desarrollo de estos paradigmas en la implementación de este proyecto, se podrá observar los logs generados por los servidores centralizados en un ambiente y tiempo real, de tal forma que se obtendrán datos verdaderos de los eventos generados por los servidores que han sido centralizados en el Centro de Cómputo de la UCSG en el que se observaran

hora y fecha de los eventos, nombre del servidor, nombres de usuarios, tipo de eventos, entre otros mensajes.

3.2 Diseño de la investigación

Para el desarrollo de este proyecto el diseño de la investigación será guiado por el libro *Metodología de la investigación* (Bernal, 2013) *,* se analizará y se escogerá el diseño **no experimental transeccional descriptivo**.

La investigación no experimental se basa en un diseño que se realiza sin manipular directamente las variables y está enfocado en la observación natural del comportamiento de los fenómenos que tiene la investigación a realizar, esto con el fin de analizar dichos acontecimientos.(Dzul, 2013)

Esto servirá para recolectar información y datos que se necesitarán para la implementación del servidor de logs centralizado, la información sólo será recolectada una vez, en un único momento y estará centrada en analizar el nivel o el estado de las variables en un momento dado.(Dzul, 2013)

De esta manera con la información ya recolectada mediante técnicas establecidas se puede seleccionar el software correcto que se va a utilizar en la implementación del servidor de logs centralizado, que dependerá del tipo de logs que generan los 5 servidores seleccionados por la unidad de Producción Informática del Centro de Cómputo de la UCSG y de las características que ofrecen las distintas herramientas de software, adicionalmente la información que será recolectada nos ayudará a escoger las correctas características necesitadas para el equipo que actuará como repositorio centralizado.

3.2.1 Técnicas e instrumentos para obtención de información

Existen diferentes tipos de técnicas que nos ayudan con la recolección de información, con el fin de elaborar una investigación o implementar un proyecto.

Para la elaboración de este proyecto se utilizó 2 técnicas las cuales se explicarán a continuación:

3.2.1.1 Entrevista

Los conceptos y definiciones de este método de recolección de datos están basadas en una investigación realizada sobre las entrevistas. (Garcia, Martinez, Naira, & Sanchez, 2009)

La entrevista es una de las técnicas de recolección de información más utilizada por los investigadores, igual que las encuestas y lo que diferencia a las entrevistas de las encuestas, es que una es cualitativa y la otra cuantitativa respectivamente.

La entrevista no está considerada como una conversación normal entre dos personas, es más enfocada a una conversación formal que tiene una intencionalidad que lleva implícitamente a los objetivos que están englobados a la investigación que el entrevistador está realizando.

Las entrevistas tienen diferentes tipos de estructuración, por ejemplo:

Entrevista no estructurada: Es un tipo de entrevista más libre, se realiza con preguntas abiertas y que se van elaborando mediante las respuestas que da el entrevistado en el desarrollo de la conversación.

Entrevista semi-estructurada: Es un tipo de entrevista más elaborada, donde el entrevistado tiene temas escogidos de los que se va a dialogar en la entrevista, además se utilizan preguntas estructuradas y preguntas que van saliendo de la conversación que se está realizando.

Entrevista estructurada: Es el tipo de entrevista más rígido que existe, ya que tiene que seguir una estructura previamente elaborada con el contexto de los temas en los que se basa la investigación del entrevistador.

Para la implementación de este proyecto se escogerá la **entrevista semi-estructurada** con un propósito individual, ya que no se elaborará un cuestionario formal, pero el dialogo se desarrollará bajo preguntas que se realizarán las cuales estarán enfocadas a los siguientes puntos:

- Selección de los 5 servidores para la centralización
- Tipo de servicios de los equipos escogidos para la centralización

- Tipos de logs generados por cada servidor
- Tiempo de almacenamiento de los logs
- Espacio en disco que consumen los logs de los servidores
- Sistemas operativos de los servidores escogidos
- Software de visualización y administración de logs
- Estructura de red de los servidores escogidos.
- Protocolo Syslog.

Se escogieron estos temas para realizar la entrevista, ya que se pretende conocer mejor los requerimientos que se necesitan cubrir en la implementación del repositorio de logs centralizado en el Centro de Cómputo de la UCSG, ya que conociendo el tipo de servicio y logs que generan los mismos se puede escoger la mejor herramienta para la administración y análisis de los diferentes tipos de logs, además de escoger el correcto espacio en disco del servidor central, que dependerá del tamaño total que generen los logs de los cinco servidores seleccionados.

3.2.1.2 Observación

Los conceptos y definiciones de este método de recolección de datos están basadas en una investigación realizada sobre la observación. (Díaz L., 2010)

La observación es otra forma de recolección de información de la metodología de la investigación, este método consiste en observar, almacenar y aclarar los comportamientos y acciones que realizan habitualmente personas u objetos.

Existen diferentes maneras de realizar la técnica de la observación, las cuales se definirán a continuación:

Sistematización

Observación sistemática: El investigador define previamente todos los elemento y comportamientos que quiere observar, lleva una estructura definida. Este método es utilizado cuando se tiene definido con claridad el problema de la investigación.

Observación natural: El investigador debe supervisar todos los aspectos relacionados al problema que está estudiando, sin necesidad de especificar los detalles antes de realizar la observación. Este método es utilizado cuando no se precisa cual es el problema de la investigación.

Participación del investigador

No participante: El investigador es no participante cuando recolecta la información desde un lugar distinto a los eventos, sin la necesidad de intervenir en los hechos del problema investigado.

Participante: El investigador es participante cuando obtiene los datos interviniendo presencialmente en los fenómenos o hechos del problema investigado.

Presencia del investigador

Indirecta: El investigador observa los eventos o hechos sin tener contacto directo con ello, sino a través de informes de otras personas que han observado el mismo acontecimiento anteriormente, libros, etc.

Directa: El investigador tiene contacto directo con las características o eventos de los que se quiere investigar.

Para el desarrollo de esta investigación se realizó una recopilación de información por medio de la técnica de **observación sistemática** ya que se tiene una estructura definida de los eventos y características de lo que se va a observar en los servidores escogidos y del software que se podría utilizar para la implementación del repositorio centralizado, por lo tanto, la observación sistemática que se realizará será **directa** e **indirectamente.**

Se hizo una revisión directa a los servidores que han sido seleccionados por la unidad de Producción Informática del Centro de Cómputo de la UCSG con el fin de identificar las características que tienen cada uno de ellos, para así tomar en cuenta la forma en que serán centralizados.

Para la observación de los detalles que se han especificado anteriormente, se recolecto la información con el siguiente formato:

Tabla 1: Formato de Características de servidores a centralizar

Servidor	Virtual/físico	S.O.	Versión S.O.	Syslog Activo/Inactivo

También se procedió a realizar una revisión indirecta a 3 programas de visualización y análisis de logs, con el fin de escoger el mejor software que se ajuste a los requerimientos necesitados para el análisis de los logs en el Centro de Cómputo de la UCSG.

Para la observación de los detalles que se han especificado anteriormente, se recolectó la información con el siguiente formato que fue elaborado según los requerimientos necesarios que solicita el Centro de Cómputo de la UCSG para la correcta administración y análisis de logs, adicionalmente se escogió tres herramientas de análisis de logs, basándose en la popularidad y buenas referencias que estas tienen, con el fin de escoger la mejor herramienta que contenga los requerimientos solicitados del Centro de Cómputo.

Tabla 2: Formato de Características de programas a evaluar

Características a evaluar	LogAnalyzer	Splunk	Loggly
Software libre			
Visualización Web			
Generación de Informes de estados			
Generación de Informes estadísticos			
Visualiza eventos de Windows			
Visualiza eventos de Linux			
Visualiza eventos de Oracle			
Base de Datos incluida			
Emisión de alertas			
Monitoreo en tiempo real			
Descarga de reportes generados			
Envió de mensajes vía Email en caso de un evento critico			
Resuelve problemas operacionales			
Servicio de logs en la nube			

3.3 Análisis de resultados

El resultado de la entrevista que se realizó en el Centro de Cómputo de la UCSG, ayuda al desarrollo e implementación de este proyecto, ya que se da a conocer el tipo de características y requerimientos que se necesitan para la implementación del servidor que actuará como repositorio central para los logs que son generados por los servicios que fueron escogidos en la entrevista realizada.

Se procedió a realizar una observación remota a los servidores que fueron seleccionados, con el fin de analizar sus características e identificar si tienen

activado el protocolo Syslog, que es la herramienta fundamental que se debe utilizar para proceder con la centralización de logs.

En la tabla 3 se presentan los detalles que fueron encontrados durante la observación.

Tabla 3: Características de servidores a centralizar

Servidor	Virtual/físico	S.O.	Versión S.O.	Syslog Activo/Inactivo
Intranet	Virtual	Red Hat	7.1	Inactivo
Dominio de usuarios	Virtual	Windows Server	2012 R2 Standard	Inactivo
Firewall	Físico	Cisco ASA 5515	7.1	Activo
Proxy principal 1	Físico	CentOS	7.1	Activo
Proxy principal 2	Virtual	CentOS	7.2	Activo

Se identificó que tres servidores poseen un SO basado en Linux, uno en Windows y un Equipo de firewall marca Cisco, en caso de los servidores basados en Linux y el equipo Cisco se deberá a proceder con la activación del protocolo Syslog en sus sistemas en caso de no tenerlo activado, en los servidores con sistemas operativos Windows se debe instalar herramientas adicionales que permitan una simulación de un protocolo Syslog, ya que el Syslog está basado en Linux y es necesario para realizar la centralización de logs en un servidor Linux.

En el análisis de la observación que se realizó a los programas de visualización y administración de logs, se pudo identificar las características de cada programa, esto con el fin de escoger la mejor herramienta basada en los requerimientos que se necesiten para la administración de logs en el Centro de Cómputo de la UCSG.

Tabla 4: Características de programas a evaluar

Características a evaluar	LogAnalyzer	Splunk	Loggly
Software libre	X		
Visualización Web	X	X	X
Generación de Informes de estados	X	X	X
Generación de Informes estadísticos	X	X	X
Visualiza eventos de Windows	X	X	X
Visualiza eventos de Linux	X	X	X
Visualiza eventos de Oracle	X	X	X
Base de Datos incluida	X	X	
Identificador de tipo de logs	X	X	X
Monitoreo en tiempo real	X	X	X
Descarga de reportes generados	X	X	X
Envío de mensajes vía Email en caso de un evento critico		X	X
Resuelve problemas operacionales			X
Servicio de logs en la nube		X	

Analizando los resultados que se dieron durante la observación de las características de las 3 herramientas seleccionadas y evaluando las características que son indispensables para el análisis de logs, se escogió la herramienta LogAnalyzer, ya que cumple los requerimientos indispensables y que son necesarios por el Centro de Cómputo de la UCSG para la administración y análisis de los logs que generan los servidores seleccionados, adicionalmente esta herramienta se diferencia de las otras ya que es OpenSource, lo que hace que sea libre de su uso sin costo adicional.

3.4 Software de análisis y visualización de logs LogAnalyzer

Como se ha mencionado anteriormente, LogAnalyzer es una herramienta OpenSource que es utilizada para la administración, análisis y visualización de logs, la herramienta cuenta con múltiples funciones que son utilizadas para el análisis de los logs que han sido centralizados en el servidor de logs principal.

Algunas de las funciones que posee esta herramienta son la generación de estadísticas, generación de informes, búsqueda de mensajes, integración con la base de datos de MySQL, búsquedas avanzadas por tipos de logs fechas, servidores.

Se analizó en profundidad algunas características y funciones de LogAnalyzer, con base en la documentación de su herramienta. (Adiscon, 2013)

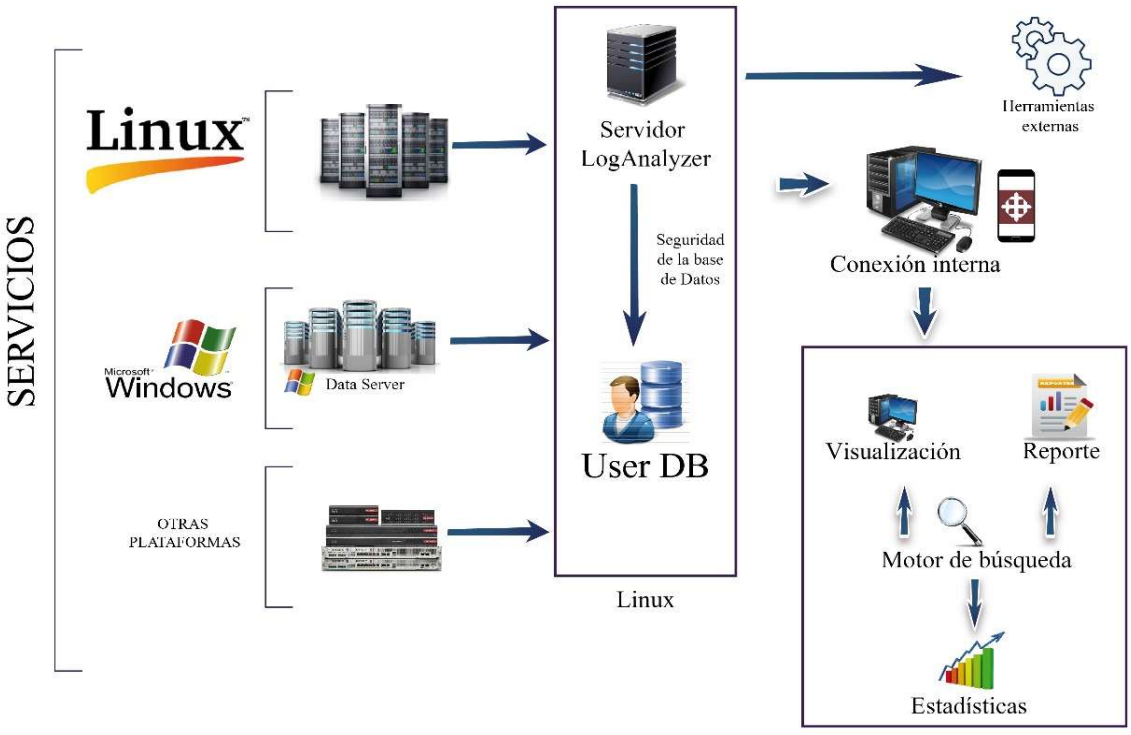


Figura 12: Estructura LogAnalyzer.

Configuración de LogAnalyzer

LogAnalyzer es una herramienta OpenSource que sirve para el análisis y la administración de diferentes tipos de logs que son manejados por servidores con sistemas operativos Linux, Windows, Oracle, Solaris, entre otros y también dispositivos de red como Routers y Switches

LogAnalyzer tiene una configuración que se la realiza por medio de un archivo maestro. Si el servidor donde la herramienta va a ser instalada tiene habilitado el sistema UserDB, entonces la mayoría de las configuraciones se las puede realizar por medio de la aplicación web, sin realizar demasiados cambios en el archivo de configuración.

El sistema UserDB

El sistema userDB permite al administrador del servidor a crear diferentes cuentas de usuario y realizar configuraciones específicas dependiendo del rol asignado a cada usuario. Esta función es una gran ayuda si varias personas interactúan con la herramienta LogAnalyzer.

El sistema userDB se encuentra deshabilitado por defecto. Esto se debe a que LogAnalyzer necesita una base de datos para que el sistema userDB funcione, ya que los perfiles o roles de usuario se almacenan dentro de dicha base de datos. La configuración de las tablas de la base de datos requiere que el administrador intervenga en su elaboración, dependiendo de los datos que se desean tener por usuario.

Básicamente LogAnalyzer tiene dos tipos de usuarios en el sistema UserDB: administradores y no administradores. Los administradores pueden modificar cualquier cosa dentro de la herramienta, los usuarios que no son administradores solo pueden cambiar sus preferencias personales.

El sistema userDB aún no es una herramienta que posea una seguridad sólida, pero si ayuda mucho con necesidades básicas de seguridad. Se pueden crear grupos de usuarios y las fuentes de datos se pueden asignar a

un grupo de usuarios específico. Entonces, solo los usuarios de este grupo pueden acceder a la fuente de datos en cuestión.

Esto es útil si se posee en la organización un grupo de personas que se encargan de manejar los registros del firewall y otro grupo que se encarga de manejar los registros del correo. También se puede definir dos grupos diferentes y asignar una misma fuente de datos en consecuencia, luego se puede asignar al usuario al grupo en que debería pertenecer. Esto se realiza con el fin de que cada usuario que interactúe con la herramienta LogAnalyzer solo pueda observar y modificar lo que tiene permitido dependiendo el rol que posea.

Base de datos de LogAnalyzer.

La herramienta puede funcionar y almacenar los logs sin la necesidad de una base de datos. LogAnalyzer solo necesita la instalación y configuración de una base de datos si se planea usar el sistema userDB o usar fuentes de datos de bases de datos. Si no se planea utilizar el sistema UserDB, no se necesita la instalación de una base de datos adicional. Un escenario típico, por ejemplo, es la revisión privada de los archivos syslog basados en el servidor, para este caso de uso, no se requiere base de datos.

Seguridad de LogAnalyzer

Los datos que poseen los logs son muy valiosos para un atacante que intenta ingresar o robar información en una empresa. Por lo tanto, es recomendable asegurar el acceso de cualquier manera al servidor central de logs donde esté instalado LogAnalyzer, especialmente si se posee datos de registro en tiempo real. Es recomendable instalar LogAnalyzer únicamente en servidores locales en la organización, que no sean accesibles a través de Internet. Si por casos especiales es necesario colocarlo en un servidor accesible a través de Internet, el acceso debe protegerse en la capa http al menos. En cualquier caso, se sugiere el uso de https para evitar la pérdida accidental de la confidencialidad.

El sistema userDB se puede utilizar como una herramienta para ajustar la capacidad del usuario para ver las fuentes de datos (los usuarios solo pueden ver las fuentes que pertenecen a uno de sus grupos). Sin embargo, esto se considera actualmente un mecanismo de control de acceso secundario. Si LogAnalyzer es accesible en Internet, se debe implementar otras fuentes de protección.

Fuentes de datos LogAnalyzer

La fuente de datos de LogAnalyzer es un conjunto de datos Syslog (y otros tipos de datos) que son recopilados. Las fuentes de datos pueden ser almacenados en archivos de texto o en una base de datos. Se debe tener en cuenta que los registros pueden ser generados en un solo archivo y no en múltiples, esto dependerá de la aplicación o equipo que genere los logs. En las bases de datos, LogAnalyzer admite tablas en formato MonitorWare o en el formato utilizado por php-syslog-ng.

Al realizar la configuración inicial de LogAnalyzer desde, se debe asegurar de utilizar el esquema de MonitorWare. Si se utiliza Rsyslog para crear la base de datos, se debe tener en cuenta que Rsyslog también trabaja con el esquema de MonitorWare por defecto. Entonces el administrador no debe realizar configuraciones especiales en caso de que se implemente Rsyslog.

Búsquedas en LogAnalyzer

La herramienta puede realizar búsquedas de cualquier fuente de datos para una variedad de propiedades. Por defecto, el texto se busca dentro del mensaje. Sin embargo, se puede realizar búsquedas mucho más complejas. Esto se lo debe realizar por medio del botón de "búsqueda avanzada" para construir estas búsquedas especializadas.

Hay que tener en cuenta que las búsquedas en LogAnalyzer se las realizan por medio de solicitudes de obtención del protocolo http. Lo que significaría que se puede copiar y pegar una URL, la cual contendrá una fuente completa. Esta podría ser una manera eficiente de enviar búsquedas

especializadas de logs a un compañero de trabajo o tener algún proceso automatizado para extraer datos específicos en un horario periódico.

La sección de búsqueda de LogAnalyzer se la utiliza de forma muy parecida a cualquier motor de búsqueda importante, ya que es bastante intuitiva. Se debe tener en cuenta que una búsqueda en la herramienta es limitada a una única fuente de datos.

Datos de registro de eventos de Windows

LogAnalyzer tiene incorporado un soporte automático para mostrar los datos del registro de eventos de Windows en un formato especifico, siempre y cuando los datos son generados por EventReporter o por los agentes de reenvío del Agente de MonitorWare . Esta herramienta incluye la detección adecuada y la capacidad de filtrar en propiedades específicas del registro de eventos (como el Id del evento).

Integración con herramientas externa

Como ya se ha mencionado anteriormente, LogAnalyzer puede acceder a herramientas externas, y lo más importante, a la base de conocimientos de MonitorWare para ayudar en el análisis de logs. El objetivo de trabajar con una herramienta externa es proporcionar información útil que ayude a realizar un mejor análisis y una identificación más rápida.

CAPITULO 4

PROPUESTA TECNOLÓGICA

Este capítulo abarca la propuesta tecnológica que se utilizó para la implementación del servidor de logs centralizado en el Centro de Cómputo de la UCSG, cumpliendo los requerimientos solicitados por el personal de Producción Informática.

4.1 Instalación y configuración de la herramienta LogAnalyzer para la centralización de logs

4.1.1 Pre Requisitos

La versión de LogAnalyzer que se utilizó en este proyecto es la 4.1.6, la cual se instaló en una plataforma Linux, y ya que se utilizó un ambiente web para la visualización y administración de los logs se necesitó del uso de PHP y apache para tener una conexión web. En este proyecto se realizó la instalación de LogAnalyzer con un gestor de base de datos (MySQL) que es donde se va a tener almacenado todos los logs que han sido centralizados.

Hardware:

- Memoria RAM 8GB
- Disco duro de 700 GB
- Procesador Intel Core I7

Software:

- Sistema operativo de ambiente Linux. (CentOS 6.2).
- Apache 2.2.3
- MySQL 5.1.73
- PHP 5.3.3
- Rsyslog 5.8.10

4.2 Instalación de LogAnalyzer

Antes de comenzar con la instalación de la herramienta LogAnalyzer se debe instalar los siguientes pre-requisitos que serán descargados vía

comandos en el terminal de la máquina virtual (CentOS 6.2) que se ha creado.

```
[root@syslog ~]# yum install httpd mysql mysql-server php php-mysql php-gd rsysl
```

Una vez finalizadas las instalaciones de los pre-requisitos se procede a poner el siguiente comando: '/ usr / bin / updatedb' que se encarga de actualizar el índice de los archivos para que los comandos 'find' y 'locate' funcionen correctamente.

```
Complete!
[root@syslog ~]# /usr/bin/updatedb
```

Después de esto se procedió a inicializar el servicio de MySQL.

```
[root@syslog ~]# /sbin/chkconfig --levels 235 mysqld on
[root@syslog ~]# /etc/init.d/mysqld start
```

Al finalizar el segundo comando se debe esperar hasta que el servicio sea inicializado.

```
                                                           [  OK  ]
Starting mysqld:                                           [  OK  ]
```

Una vez inicializada la base de datos se procedió a configurar la seguridad básica, como contraseña, quitar usuarios anónimos, impedir conexiones remotas a la base de datos y cargar las tablas de privilegios, por lo que se debe colocar “Y” a todas las preguntas que se generen.

```
[root@syslog ~]# /usr/bin/mysql_secure_installation


Set root password? [Y/n] y
New password:
Re-enter new password:
Password updated successfully!
Reloading privilege tables..
 ... Success!
```

```
Remove anonymous users? [Y/n] y
 ... Success!

Normally, root should only be allowed to connect from 'localhost'.  This
ensures that someone cannot guess at the root password from the network.

Disallow root login remotely? [Y/n] y
 ... Success!

By default, MySQL comes with a database named 'test' that anyone can
access.  This is also intended only for testing, and should be removed
before moving into a production environment.

Remove test database and access to it? [Y/n] y
 - Dropping test database...
 ... Success!
 - Removing privileges on test database...
 ... Success!

Reloading the privilege tables will ensure that all changes made so far
will take effect immediately.

Reload privilege tables now? [Y/n] y
 ... Success!
```

Establecida la seguridad básica, se procedió a configurar la base de datos y sus tablas.

Primero ingresando a la base de datos con el usuario root y colocando la contraseña definida en el servidor.

```
[root@syslog ~]# mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 12
Server version: 5.1.73 Source distribution

Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>
```

Se procedió a crear un usuario en la base de datos y designarle una contraseña.

```
mysql> CREATE USER rsyslog;
Query OK, 0 rows affected (0.00 sec)

mysql> SET PASSWORD FOR rsyslog= PASSWORD('password here');
```

Luego se procedió a crear la base de datos designada para el Rsyslog y se la selecciona.

```
mysql> CREATE DATABASE rsyslogdb;
Query OK, 1 row affected (0.00 sec)

mysql> USE rsyslogdb;
Database changed
mysql>
```

Manualmente, se tiene que crear dos tablas en la base de datos rsyslogdb: “SystemEvents” que servirán de repositorio de almacenamiento de los logs recibidos de los servidores clientes y el localhost. La otra tabla que se creó es “SystemEventsProperties” donde se guardarán el nombre e identificador de todos los logs guardados en la base de datos. El esquema de las tablas de la base de datos fue estructurado en base a un post de esta instalación.(Scobles, 2012)

```
mysql> CREATE TABLE SystemEvents
    -> (
    ->  ID int unsigned not null auto_increment primary key,
    ->  CustomerID bigint,
    ->  ReceivedAt datetime NULL,
    ->  DeviceReportedTime datetime NULL,
    ->  Facility smallint NULL,
    ->  Priority smallint NULL,
    ->  FromHost varchar(60) NULL,
    ->  Message text,
    ->  NTSeverity int NULL,
    ->  Importance int NULL,
    ->  EventSource varchar(60),
    ->  EventUser varchar(60) NULL,
    ->  EventCategory int NULL,
    ->  EventID int NULL,
    ->  EventBinaryData text NULL,
    ->  MaxAvailable int NULL,
    ->  CurrUsage int NULL,
    ->  MinUsage int NULL,
    ->  MaxUsage int NULL,
    ->  InfoUnitID int NULL ,
    ->  SysLogTag varchar(60),
    ->  EventLogType varchar(60),
    ->  GenericFileName VarChar(60),
    ->  SystemID int NULL
    -> );
Query OK, 0 rows affected (0.02 sec)

mysql> CREATE TABLE SystemEventsProperties
    -> (
    ->  ID int unsigned not null auto_increment primary key,
    ->  SystemEventID int NULL ,
    ->  ParamName varchar(255) NULL ,
    ->  ParamValue text NULL
    -> );
Query OK, 0 rows affected (0.01 sec)

mysql>
```

Una vez creadas las tablas se debe dar privilegios a la cuenta de usuario Rsyslog que se creó anteriormente.

```
mysql> GRANT ALL PRIVILEGES ON `rsyslogdb`.* TO 'rsyslog'@'%' IDENTIFIED BY 'password here';
```

```
mysql> flush privileges;
Query OK, 0 rows affected (0.00 sec)

mysql>
```

Terminado esto, se procedió a salir de la base de datos para continuar con la configuración del archivo Rsyslog.

```
mysql> exit
Bye
[root@syslog ~]#
```

Se debe editar en 3 pasos el archivo principal del Rsyslog (rsyslog.conf), con el fin de permitir el paso de los mensajes syslog que se envíen desde los servidores clientes al central, también se deberá configurar las redes o sub redes de las cuales serán permitidas enviar mensajes syslog hacia el servidor central, y por último se debe dar paso a que los mensajes syslog que lleguen de los clientes al central sean guardados en la base de datos de Rsyslog.

Paso 1: En la sección “Modules” se debe habilitar el puerto 514 en TCP y UDP para recibir mensajes Syslog.

```
[root@syslog ~]# vi /etc/rsyslog.conf
```

```
# rsyslog v5 configuration file

# For more information see /usr/share/doc/rsyslog-*/rsyslog_conf.html
# If you experience problems, see http://www.rsyslog.com/doc/troubleshoot.html

#### MODULES ####

$ModLoad ommysql  # provides support for MySQL
$ModLoad imuxsock # provides support for local system logging (e.g. via logger comma
$ModLoad imklog   # provides kernel logging support (previously done by rklogd)
#$ModLoad immark  # provides --MARK-- message capability

# Provides UDP syslog reception
$ModLoad imudp.so
$UDPServerRun 514

# Provides TCP syslog reception
$ModLoad imtcp.so
$InputTCPServerRun 514
```

Paso 2: Antes de llegar a la sección ### begin forwarding rule ### Se debe configurar las direcciones IP o las subredes a las que el servidor deberá aceptar mensajes Syslog vía UDP y TCP, en este proyecto se utilizaron 2 sub-redes (172.16.100.0/24 y 172.16.1.0/24) más la IP local 127.0.0.1

```
$AllowedSender TCP, 127.0.0.1, 172.16.1.0/24, 172.16.100.0/24
$AllowedSender UDP, 127.0.0.1, 172.16.1.0/24, 172.16.100.0/24

# ### begin forwarding rule ###
# The statement between the begin ... end define a SINGLE forwarding
# rule. They belong together, do NOT split them. If you create multiple
# forwarding rules, duplicate the whole block!
# Remote Logging (we use TCP for reliable delivery)
#
```

Paso 3: En la sección ### begin forwarding rule ### del archivo rsyslog.conf se debe agregar la siguiente línea, para que los mensajes Syslogs sean guardados en la base de datos Rsyslog:

```
*.*       :ommysql:127.0.0.1,rsyslogdb,rsyslog,<password here>
# ### end of the forwarding rule ###
-- INSERT --
```

Terminada estas configuraciones, se debe guardar y salir del archivo para reiniciar el servicio de Rsyslog.

```
[root@syslog ~]# service rsyslog restart
Shutting down system logger:                               [  OK  ]
Starting system logger:                                    [  OK  ]
[root@syslog ~]#
```

Una vez reestablecido el servicio de Rsyslog, se debe realizar una prueba para verificar que los logs se están enviando a la base de datos de Rsyslog creada anteriormente.

Primero se ejecutó el comando: “tail -f / var / log / messages” para visualizar los logs generados por el propio servidor Syslog.

```
[root@syslog Desktop]# tail -f /var/log/messages
Jan 18 15:45:07 syslog rsyslogd: db error (1045): Access denied for user 'rsyslo
using password: YES)

Jan 18 15:48:25 syslog kernel: Kernel logging (proc) stopped.
Jan 18 15:48:25 syslog rsyslogd: [origin software="rsyslogd" swVersion="5.8.10"
nfo="http://www.rsyslog.com"] exiting on signal 15.
Jan 18 15:48:25 syslog kernel: imklog 5.8.10, log source = /proc/kmsg started.
Jan 18 15:48:25 syslog rsyslogd: [origin software="rsyslogd" swVersion="5.8.10"
nfo="http://www.rsyslog.com"] start
Jan 18 15:55:49 syslog kernel: Kernel logging (proc) stopped.
Jan 18 15:55:49 syslog rsyslogd: [origin software="rsyslogd" swVersion="5.8.10"
nfo="http://www.rsyslog.com"] exiting on signal 15.
Jan 18 15:55:49 syslog kernel: imklog 5.8.10, log source = /proc/kmsg started.
Jan 18 15:55:49 syslog rsyslogd: [origin software="rsyslogd" swVersion="5.8.10"
nfo="http://www.rsyslog.com"] start
```

Luego se accedió al motor de base de datos MySQL.

```
[root@syslog Desktop]# mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 44
Server version: 5.1.73 Source distribution

Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
```

Se seleccionó la base de datos de Rsyslog.

```
mysql> use rsyslogdb;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
```

Se realizó una consulta a la tabla SystemEvents creada anteriormente, si el resultado de la búsqueda es algo diferente a “empty set” significa que la base de datos está guardando los logs generados localmente, por lo que se determina que las instalaciones y configuraciones hasta ahora están correctas.

```
mysql> select * from SystemEvents;
+----+------------+---------------------+---------------------+----------+----------+----------+
------------------------------------------------------------------------------------------------
--------------------+------------+------------+-------------+-----------+---------------+------
---+----------------+--------------+-----------+----------+----------+-----------+-----------
----------------------+--------------+-----------------+----------+
| ID | CustomerID | ReceivedAt          | DeviceReportedTime  | Facility | Priority | FromHost |
 Message
                    | NTSeverity | Importance | EventSource | EventUser | EventCategory | Event
ID | EventBinaryData | MaxAvailable | CurrUsage | MinUsage | MaxUsage | InfoUnitID | SysLogTag
                      | EventLogType | GenericFileName | SystemID |
+----+------------+---------------------+---------------------+----------+----------+----------+
------------------------------------------------------------------------------------------------
--------------------+------------+------------+-------------+-----------+---------------+------
---+----------------+--------------+-----------+----------+----------+-----------+-----------
----------------------+--------------+-----------------+----------+
|  1 |       NULL | 2018-01-18 15:48:25 | 2018-01-18 15:48:25 |        0 |        6 | syslog   |
 imklog 5.8.10, log source = /proc/kmsg started.
                    |       NULL |       NULL | NULL        | NULL      |          NULL |    NU
LL | NULL           |         NULL |      NULL |     NULL |     NULL |          1 | kernel:
                     | NULL         | NULL           |     NULL |
|  2 |       NULL | 2018-01-18 15:48:25 | 2018-01-18 15:48:25 |        5 |        6 | syslog   |
  [origin software="rsyslogd" swVersion="5.8.10" x-pid="2993" x-info="http://www.rsyslog.com"] s
tart               |       NULL |       NULL | NULL        | NULL      |          NULL |    NU
```

Ahora se debe configurar el sistema operativo para iniciar el servidor web en el arranque y también se debe dar inicio manualmente al servicio httpd:

```
[root@syslog Desktop]# chkconfig --levels 235 httpd on
[root@syslog Desktop]# service httpd start
Starting httpd: httpd: apr_sockaddr_info_get() failed for syslog
httpd: Could not reliably determine the server's fully qualified domain name, us
ing 127.0.0.1 for ServerName
                                                           [  OK  ]
[root@syslog Desktop]#
```

Una vez levantado el servicio, se debe modificar dos líneas del archivo **httpd.conf** en la ruta /etc/httpd/conf/httpd.conf

```
[root@syslog Desktop]# vi /etc/httpd/conf/httpd.conf
[root@syslog Desktop]#
```

Primero se debe cambiar la línea de “Listen:80” a “Listen: IP-servidor:80”

```
# Listen: Allows you to bind Apache to specific IP addresses and/or
# ports, in addition to the default. See also the <VirtualHost>
# directive.
#
# Change this to Listen on specific IP addresses as shown below to
# prevent Apache from glomming onto all bound IP addresses (0.0.0.0)
#
#Listen 12.34.56.78:80
Listen 172.16.100.8:80
```

Después se debe modificar la línea de “#ServerName www.example.com:80” a “ ServerName nombre-servidor: 80”

```
#
# ServerName gives the name and port that the server uses to identify itself.
# This can often be determined automatically, but we recommend you specify
# it explicitly to prevent problems during startup.
#
# If this is not set to valid DNS name for your host, server-generated
# redirections will not work.  See also the UseCanonicalName directive.
#
# If your host doesn't have a registered DNS name, enter its IP address here.
# You will have to access it by its address anyway, and this will make
# redirections work in a sensible way.
#
ServerName syslog:80
```

Una vez modificada estas dos líneas, se procede a guardar los cambios en el archivo y reiniciar el servicio.

```
[root@syslog conf]# /etc/init.d/httpd restart
Stopping httpd:                                            [  OK  ]
Starting httpd:                                            [  OK  ]
[root@syslog conf]#
```

Terminado el reinicio del servicio, se procedió a modificar el archivo IPtables

```
root@syslog conf]# vi/etc/sysconfig/iptables
```

Se debe agregar al final del archivo, antes del COMMIT lo siguiente:

```
INPUT -p tcp --dport 80 -m state --state NEW,ESTABLISHED -j ACCEPT
OUTPUT -p tcp --sport 80 -m state --state ESTABLISHED -j ACCEPT
```

Y a continuación de estas líneas se deben agregar las redes de las cuales están permitidas el paso de mensajes Syslog, para este proyecto el archivo quedaría de la siguiente manera:

```
# Firewall configuration written by system-config-firewall
# Manual customization of this file is not recommended.
*filter
:INPUT ACCEPT [0:0]
:FORWARD ACCEPT [0:0]
:OUTPUT ACCEPT [0:0]
-A INPUT -m state --state ESTABLISHED,RELATED -j ACCEPT
-A INPUT -p icmp -j ACCEPT
-A INPUT -i lo -j ACCEPT
-A INPUT -m state --state NEW -m tcp -p tcp --dport 22 -j ACCEPT
-A INPUT -j REJECT --reject-with icmp-host-prohibited
-A FORWARD -j REJECT --reject-with icmp-host-prohibited
-I INPUT -p tcp --dport 80 -m state --state NEW,ESTABLISHED -j ACCEPT
-I OUTPUT -p tcp --sport 80 -m state --state ESTABLISHED -j ACCEPT
-I INPUT -p tcp --dport 514 -s 172.16.100.0/24 -j ACCEPT
-I INPUT -p udp --dport 514 -s 172.16.100.0/24 -j ACCEPT
-I INPUT -p tcp --dport 514 -s 172.16.1.0/24 -j ACCEPT
-I INPUT -p udp --dport 514 -s 172.16.1.0/24 -j ACCEPT
COMMIT
```

Terminando de modificar el archivo, se debe guardar y reiniciar los servicios de red y las de IPtables

```
[root@syslog conf]# /etc/init.d/network restart
Shutting down interface eth0:  Device state: 3 (disconnected)
                                                           [  OK  ]
Shutting down loopback interface:                          [  OK  ]
Bringing up loopback interface:                            [  OK  ]
Bringing up interface eth0:  Active connection state: activated
Active connection path: /org/freedesktop/NetworkManager/ActiveConnection/2
                                                           [  OK  ]
[root@syslog conf]# /etc/init.d/iptables restart
iptables: Flushing firewall rules:                         [  OK  ]
iptables: Setting chains to policy ACCEPT: filter          [  OK  ]
iptables: Unloading modules:                               [  OK  ]
iptables: Applying firewall rules:                         [  OK  ]
[root@syslog conf]#
```

Si todas estas configuraciones han sido correctamente instaladas y modificadas, se debe proceder a descargar el programa LogAnalyzer, que para la elaboración de este proyecto se utilizará la versión más reciente (4.1.6)

```
[root@syslog Desktop]# cd /tmp
[root@syslog tmp]# wget http://download.adiscon.com/loganalyzer/loganalyzer-4.1.
6.tar.gz
--2018-01-22 14:52:42--  http://download.adiscon.com/loganalyzer/loganalyzer-4.1
.6.tar.gz
Resolving download.adiscon.com... 138.201.116.127, 2a01:4f8:c17:44a6::2
Connecting to download.adiscon.com|138.201.116.127|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 2852860 (2.7M) [application/x-gzip]
Saving to: “loganalyzer-4.1.6.tar.gz”

100%[======================================>] 2,852,860    136K/s   in 32s

2018-01-22 14:53:14 (87.9 KB/s) - “loganalyzer-4.1.6.tar.gz” saved [2852860/2852
860]

[root@syslog tmp]#
```

Una vez descargado el LogAnalyzer, se procedió a descomprimir la carpeta que lo contiene.

```
[root@syslog tmp]# tar -xvzf loganalyzer-4.1.6.tar.gz
```

Cuando se haya terminado de descomprimir la carpeta, se debe copiar el directorio de origen en el directorio HTML de apache y proceder a crear el archivo config.php

```
[root@syslog tmp]# cd loganalyzer-4.1.6/src
[root@syslog src]# rm -R -f /var/www/html
[root@syslog src]# mkdir /var/www/html
[root@syslog src]# cp -R * /var/www/html
[root@syslog src]# cd /tmp/loganalyzer-4.1.6/contrib/
[root@syslog contrib]# cp * /var/www/html
[root@syslog contrib]# cd /var/www/html
[root@syslog html]# chmod +x configure.sh secure.sh
[root@syslog html]# ./configure.sh
```

Finalizada esta configuración se debe realizar un “ls” para observar si el archivo config.php se ha creado.

```
[root@syslog html]# ls
admin               configure.sh  favicon.ico  lang                statistics.php
asktheoracle.php    convert.php   images       login.php           templates
BitstreamVeraFonts  cron          include      reportgenerator.php themes
chartgenerator.php  css           index.php    reports.php         userchange.php
classes             details.php   install.php  search.php
config.php          export.php    js           secure.sh
[root@syslog html]#
```

Se debe realizar un último paso para finalizar correctamente la instalación de LogAnalyzer. Se creó una base de datos LogAnalyzer MySQL y su usuario correspondiente.

```
[root@syslog Desktop]# mysql -u root -p
```

```
mysql> create database loganalyzerdb;

mysql> CREATE USER loganalyzer;


mysql> SET PASSWORD FOR loganalyzer= PASSWORD('setpasswordhere');

mysql> GRANT ALL PRIVILEGES ON `loganalyzerdb`.* TO 'loganalyzer'@'%' IDENTIFIED
 BY 'setpasswordhere';

mysql> flush privileges;
```

Terminada la configuración de la base de datos de LogAnalyzer, se puede incluir un módulo Ommail dentro del archivo Rsyslog.conf el cual enviara una notificación de algún log generado por correo electrónico, por medio de un servidor SMTP. En esta configuración se estableció que se envié un correo electrónico cuando el servicio Rsyslog sea reiniciado.

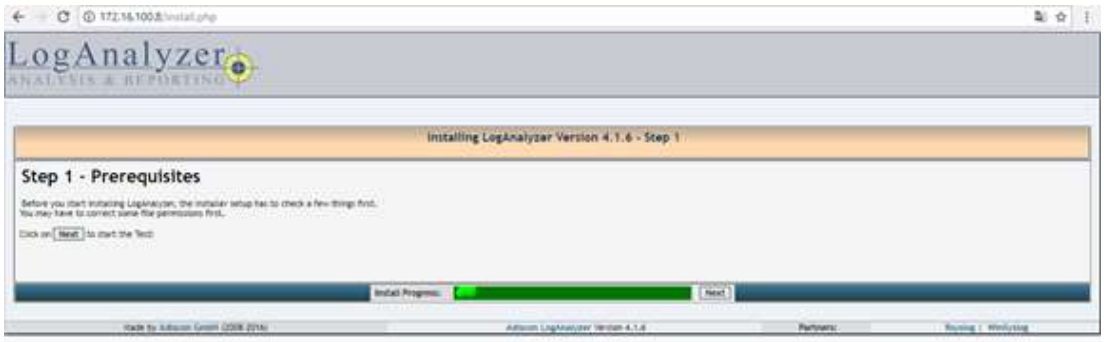
```
$ModLoad imtcp.so
$InputTCPServerRun 514

#mail
$ModLoad ommail
$ActionMailSMTPServer 172.16.1.125
$ActionMailFrom rramirep@ucsg.edu.ec
$ActionMailTo luferfiro@hotmail.com
$template mailSubject,"UCSG ALERTA DE LOG SERVIDOR: %hostname%"
$template mailBody,"RSYSLOG Alert\r\nMensaje='%msg%'"
$ActionMailSubject mailSubject
# make sure we receive a mail only once in six
# hours (21,600 seconds ;))
$ActionExecOnlyOnceEveryInterval 60
# the if ... then ... mailBody must be on one line!
if $msg contains 'imklog 5.8.10, log source = /proc/kmsg started.' then :ommail:
;mailBody
# re-set interval so that other actions are not affected
$ActionExecOnlyOnceEveryInterval 0
```

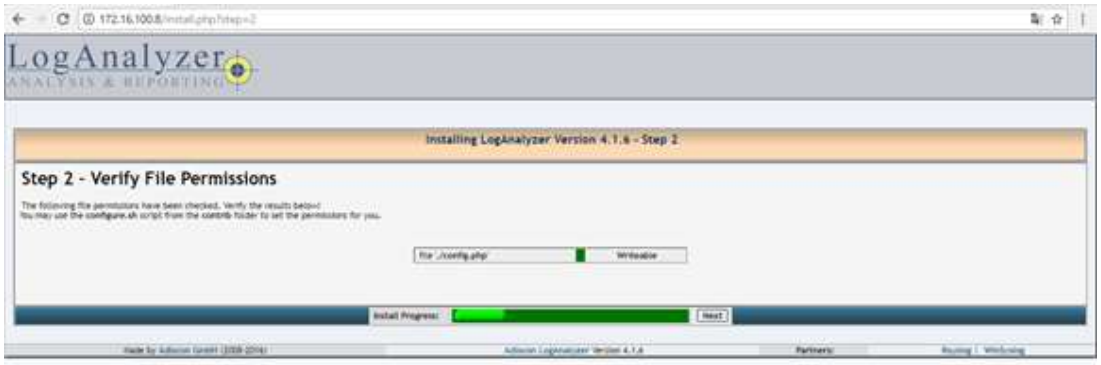
Una vez configurado el modulo Ommail, ha concluido la instalación de la herramienta, ahora se debe proceder con la configuración respectiva vía web.

4.3 Configuración de LogAnalyzer

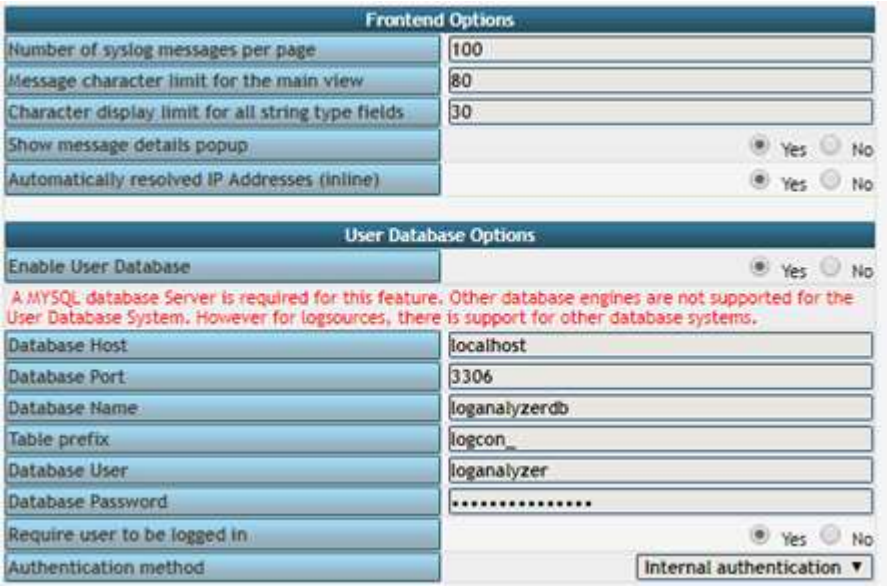
Para realizar el primer paso de la configuración de LogAnalyzer se necesita ir a un equipo cliente dentro de la red y acceder por vía web apuntando a la dirección IP establecida en el servidor. Debe aparecer un mensaje de error critico que indica que “Falta el archivo de configuración principal”, para comenzar la configuración se deberá dar click en “Here”.

Luego se debe dar dos veces click en “Next” en las próximas dos pantallas que aparecerán que serán la de prerrequisitos y la verificación de permisos establecidos al archivo config.php.





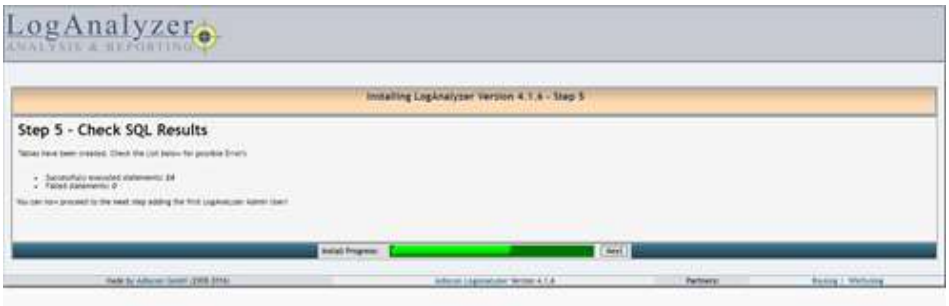
La configuración del paso 3 para este proyecto quedara de la siguiente manera:



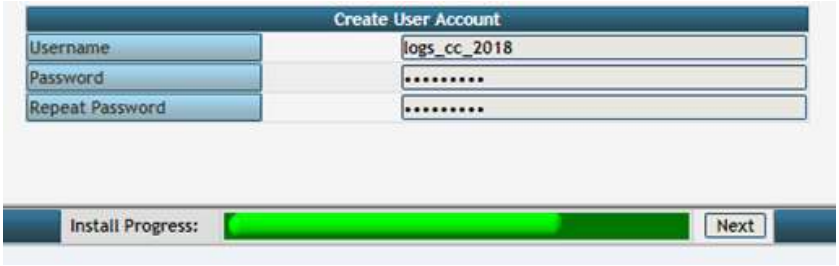
Esta configuración es muy importante para el correcto funcionamiento de la aplicación web, ya que se debe selecciona la base de datos donde se almacenan los logs del host y clientes, se debe establecer el número de logs por página en la aplicación web, entre otras configuraciones elementales.

En el cuarto paso se debe dar click en “Next”

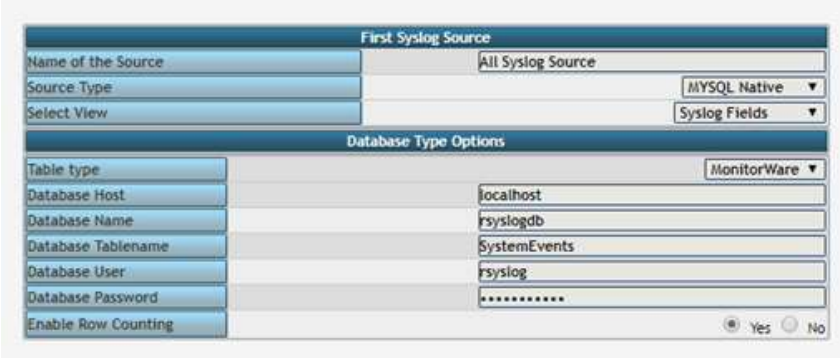
En el quinto paso también se debe dar click en “Next”



En el paso 6 se procedió a crear un usuario administrador para la aplicación web de LogAnalyzer, con el fin de visualizar, administrar y configurar los reportes de los logs generados por los servidores.

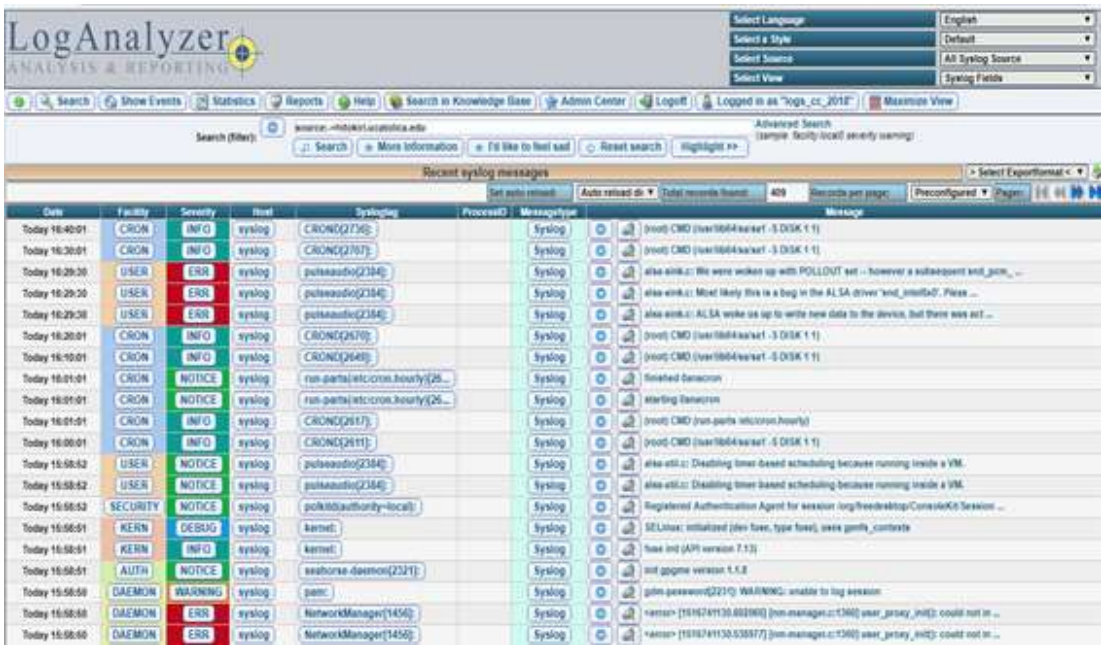


El paso 7 es el último de la configuración inicial de LogAnalyzer, en este paso se debe especificar en donde se va a buscar los mensajes Syslog que tenga el servidor central de logs, y se debe configurar la base de datos Rsyslog creado en los pasos de instalación de LogAnalyzer



El paso 8 solo indica que todo ha sido configurado correctamente, se debe dar click en “Finalizar” y se procedió a colocar el usuario y la contraseña de administrador establecido para la aplicación web.

Si toda la instalación y configuración ha sido realizada correctamente, se abrirá el visualizador de logs, en este caso se podrá observar los logs que son generados por el propio servidor centralizado.



4.4 Reportes y estadísticas de LogAnalyzer

Una de las características principales de LogAnalyzer es la elaboración libre de reportes y estadísticas personalizadas, de acuerdo a las necesidades o requerimientos de las personas encargadas de la administración y análisis de los logs.

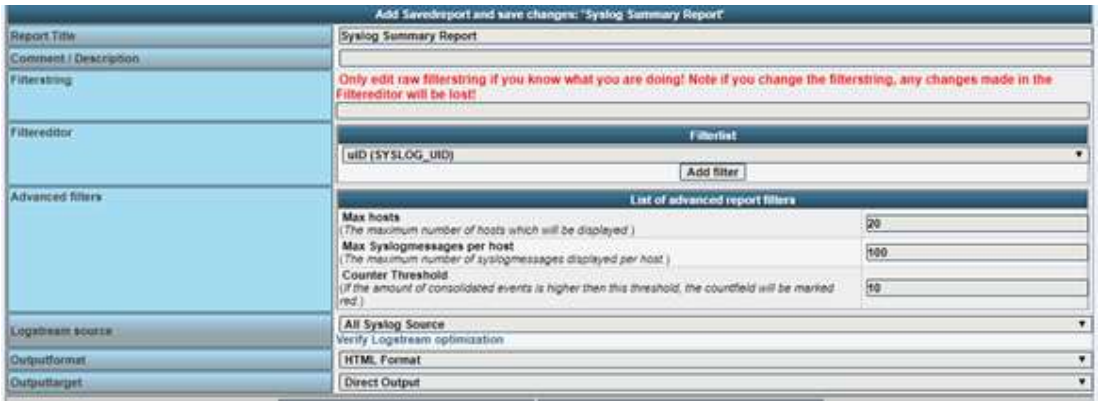
4.4.1 Reportes

Para elaborar un reporte se debe ingresar a la aplicación web como un usuario con privilegios altos y dar click en “Admin Center” lo cual habilita opciones avanzadas de administrador.

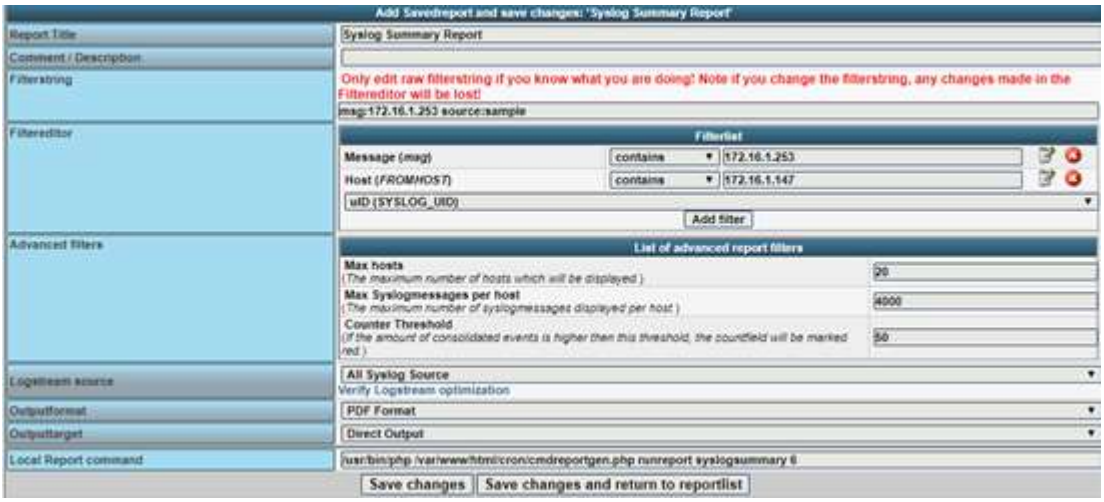
Luego se debe dar click en la opción “Report Modules”

Este proyecto está basado en el protocolo syslog, por lo que los reportes que se deben elaborar son “Syslog Summary”.

Para elaborar un reporte nuevo se debe dar click en la opción “Add Savedreport and save changes



Los datos para el siguiente ejemplo fueron llenados de acuerdo a un reporte solicitado por el personal del Centro de Cómputo de la UCSG en el cual se quieren identificar cuáles son las peticiones de nombres más solicitadas por los usuarios por medio del DNS.



Los datos que se establecieron para este ejemplo fueron:

Dos filtros, el Host que en este caso es el firewall (172.16.1.147) que envía los logs al repositorio centralizado y el contenido del mensaje que es el servidor DNS (172.16.1.253).

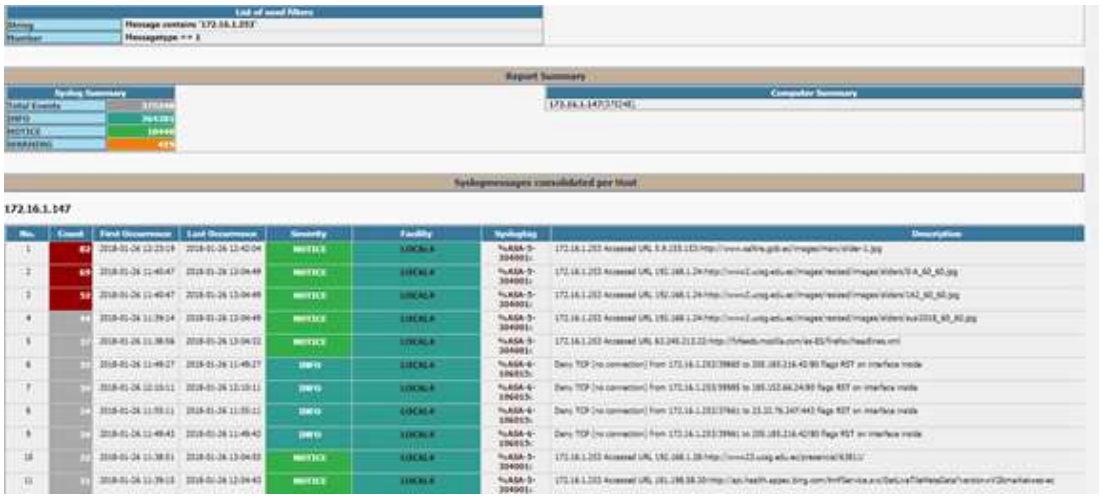
En este caso el campo “Max Host” es irrelevante ya que solo se trabaja con un host en este reporte.

En el campo “Max Syslogmessages per host” se estableció que sea 4000, ya que son varios mensajes Syslog que se ven en este reporte.

En el campo “Counter Threshold” se colocó 80, lo que identificara con un color rojo las nombres que hayan sido pedidas más de 80 veces.

En los dos últimos campos se establece si el reporte será generado en formato PDF o HTML y si será descargado o visualizado primero en web en caso de que se haya elegido PDF.

El reporte final se vera de la siguiente manera en formato HTML:



En formato PDF será de la siguiente manera:

List of used filters	
String	**Message contains '172.16.1.253'**
Number	**Messagetype == 1**

Report Summary

Syslog Summary	
Total Events	**375248**
INFO	**364381**
NOTICE	**10448**
WARNING	**419**

Computer Summary
172.16.1.147(375248).

Syslogmessages consolidated per Host

172.16.1.147

No.	Count	First Occurrence	Last Occurrence	Severity	Facility
1	82	2018-01-26 12:23:19	2018-01-26 12:42:04	NOTICE	LOCAL4
	Description		172.16.1.253 Accessed URL 5.9.155.153:http://www.salitre.gob.ec/images/mary/slider-1.jpg		
2	69	2018-01-26 11:40:47	2018-01-26 13:04:49	NOTICE	LOCAL4
	Description		172.16.1.253 Accessed URL 192.168.1.24:http://www2.ucsg.edu.ec/images/resized/images/slide rs/0-A_60_60.jpg		
3	52	2018-01-26 11:40:47	2018-01-26 13:04:49	NOTICE	LOCAL4
	Description		172.16.1.253 Accessed URL 192.168.1.24:http://www2.ucsg.edu.ec/images/resized/images/slide rs/1A2_60_60.jpg		
4	44	2018-01-26 11:39:14	2018-01-26 13:04:49	NOTICE	LOCAL4
	Description		172.16.1.253 Accessed URL 192.168.1.24:http://www2.ucsg.edu.ec/images/resized/images/slide rs/susi2018_60_60.jpg		

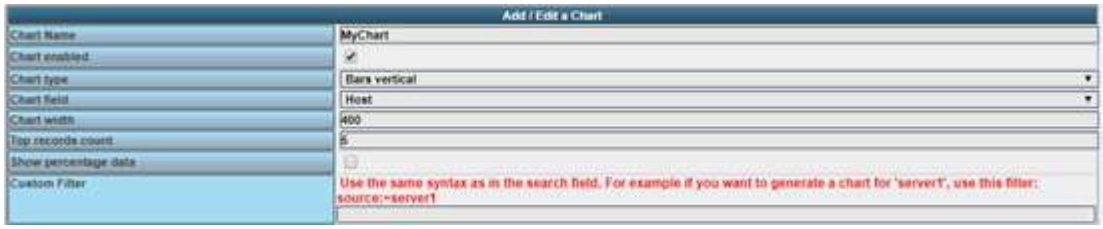
4.4.2 Estadísticas

Para elaborar una estadística, al igual que para generar un reporte se debe ingresar a la aplicación web como un usuario con privilegios altos y dar click en “Admin Center” lo cual habilita opciones avanzadas de administrador.

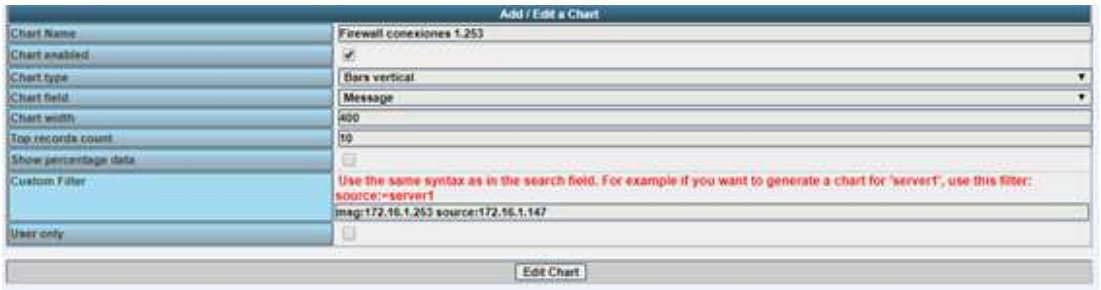
Luego se debe dar click en la opción “Charts”

En la herramienta vienen Pre-configurada algunas estadísticas básicas como la cantidad de logs de cada servidor cliente que este enviando los reportes al servidor centralizado, la cantidad de logs con severidad “Error”, entre otros. Para configurar una nueva estadística se debe dar click en “Add new Chart”.

Los datos en los campos deben ser llenado de acuerdo al tipo de estadística que se realiza.



Los datos para el siguiente ejemplo fueron llenados de acuerdo a una estadística solicitada por el personal del Centro de Cómputo de la UCSG en el cual se quieren identificar por medio de un gráfico cuáles son las diez peticiones de nombres más solicitadas por los usuarios por medio del DNS.



Los campos fueron llenados de la siguiente manera:

En el campo “Chart Type” se escogió que el grafico sea generado en barras verticales, LogAnalyzer posee 3 tipos de gráficos: Barras verticales, horizontales y tipo pastel.

En el campo “Chart Field” se escoge el tipo de búsqueda para generar el grafico, para el ejemplo se escogió tipo “Message”.

Se escogió 10 el número total de barras que se mostraran, ya que el requerimiento del Centro de Cómputo es que se muestre el top 10 de peticiones al servidor DNS.

En el campo “Custom Filter” es donde se debe filtrar la búsqueda para generar el grafico de la estadística. En este ejemplo se filtró por medio de la dirección IP del DNS que se encuentra en el servidor de Firewall.

Una vez configurada la estadística se debe dar click en “Edit Chart” para guardarlo.

Para visualizar la estadística elaborada se debe dar click en la opción “Stadistics” en la aplicación web.

Esta opción muestra las estadísticas que se fueron elaboradas y las predeterminadas.

Lo que se buscó en el desarrollo de esta propuesta tecnológica es centralizar los logs que generan los múltiples servicios que maneja el Centro de Cómputo de la UCSG y por medio de la instalación y configuración de la herramienta LogAnalyzer poder tener un mejor orden, búsqueda y visualización de los logs.

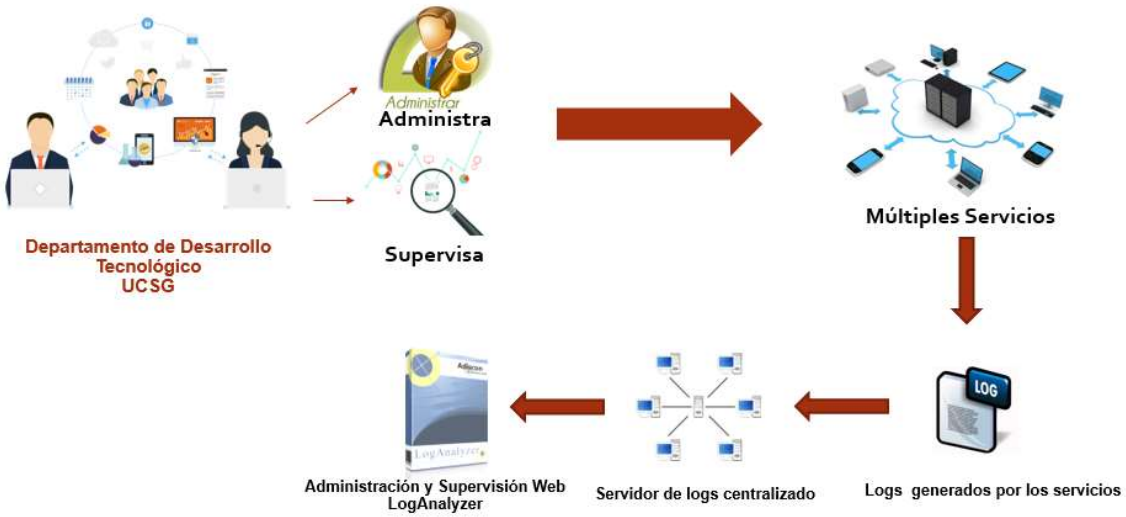


Figura 13: Diseño de implementación.

Conclusiones y recomendaciones

Conclusiones

Para el respectivo proyecto se logró realizar un análisis junto con el personal de Producción Informática del Centro de Cómputo con el fin de centralizar los logs de 5 equipos que ofrecen servicios a la UCSG.

Se procedió a realizar por medio de una observación directa, características de los equipos escogidos, como lo fueron:

- Servicio prestado
- Virtual o físico
- Sistema operativo
- Versión del sistema operativo
- Syslog activo o inactivo

Analizando estas características, se logró diseñar un ambiente óptimo para la centralización de logs con los servidores escogidos, ya que se determinó el sistema operativo y recursos de hardware que fueron necesarios para implementar un servidor centralizado.

Se realizó una observación y análisis a las herramientas de administración de logs centralizados más utilizadas y que cumplieran con los requerimientos necesitados por el Centro de Cómputo de la UCSG

Se seleccionó la herramienta LogAnalyzer ya que cumplía parámetros establecidos por el personal de Producción Informática y además es una herramienta OpenSource gratuita, donde adicionalmente se pueden elaborar reportes y estadísticas de acuerdo a las necesidades que se vayan presentando en el Centro de Cómputo, incluso se puede establecer para que la herramienta trabaje con otros protocolos diferentes al Syslog.

Por medio de la herramienta escogida se implementó una configuración para establecer un ambiente web dentro de la red de la UCSG, con el fin de visualizar los logs que han sido centralizados, de una manera ordenada, filtrando la búsqueda de logs específicos que dependerá de lo que se requiere observar en el momento.

Se elaboraron reportes y estadísticas especializados, cumpliendo con los requerimientos que necesitaba cubrir el personal de Producción Informática del Centro de Cómputo.

Se puede determinar que los objetivos de la investigación fueron cumplidos de acuerdo al requerimiento suscitado en la problemática que posee el Centro de Cómputo de la UCSG con la administración de logs.

Recomendaciones

El servicio de logs centralizado debería implementarse en un servidor de producción sea físico o virtual, ya que cuando se desarrolló este proyecto, se lo implementó en un equipo de pruebas. Se recomienda implementarlo en un servidor con el fin de una vez que el servicio vaya a ser puesto en producción, se debe buscar las características adecuadas, dependiendo de cuantos servidores más direccionarán sus logs al servidor centralizado y así también poder realizar los diferentes tipos de respaldo que se manejan dentro del departamento del Centro de Cómputo.

LogAnalyzer no tiene una configuración para emitir una alerta en caso de que se genere un log con severidad 0, lo que significaría que un servicio ha dejado de funcionar correctamente en algún equipo que este enviado los logs al repositorio centralizado. Es recomendable adicionar configuraciones al módulo “Ommail” del archivo de configuración Rsyslog para enviar un Email por medio de un servidor SMTP en caso de que se genere un log de emergencia o algún log importante para el personal de Producción Informática del Centro de Cómputo. Para configurar el módulo Ommail se debe ingresar a archivo Rsyslog.conf y modificarlo dependiendo del tipo de log que se desee recibir una notificación.

Figura 14: Alerta de mensaje Syslog

El ambiente web del servidor centralizado puede ser visualizado dentro de la red interna de la UCSG, por lo que se recomienda instalar un módulo en la aplicación móvil de herramientas administrativas para el personal de producción informática que está en proyecto en el Centro de Cómputo, con el fin de visualizar los logs, reportes o estadísticas en un teléfono Smartphone que esté conectado en alguna red interna dentro de la Universidad.

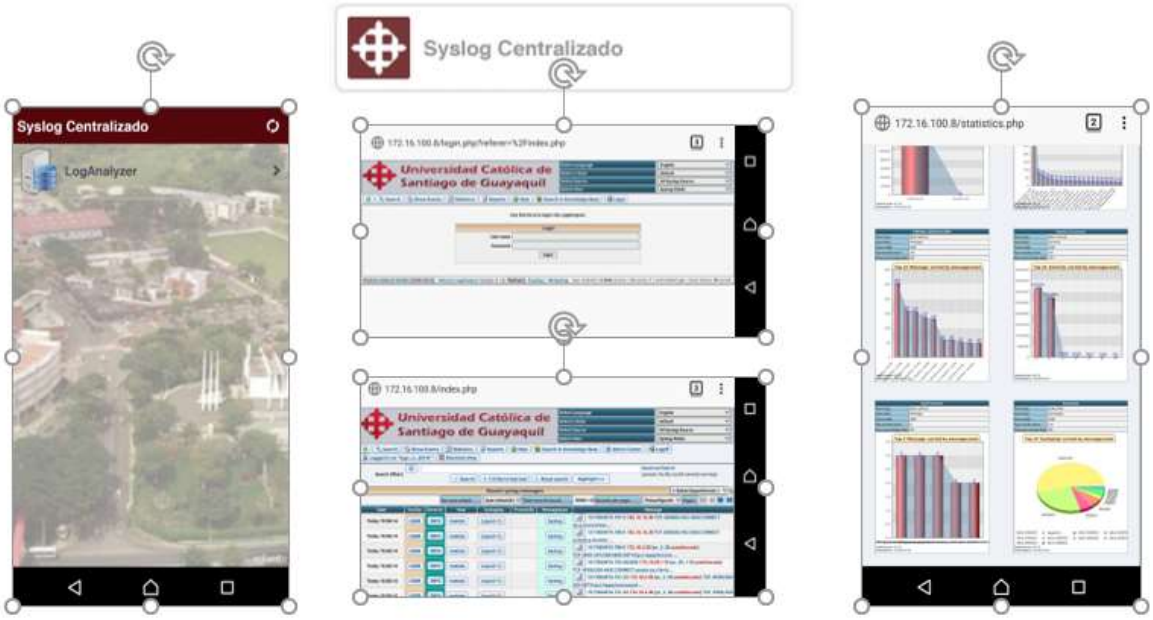


Figura 15. Aplicación móvil – Modulo Loganalyzer

En el caso de la seguridad de la base de datos, se recomienda instalar un protocolo de encriptación tipo AES de la base utilizada en la implementación de este proyecto (MySQL), con el fin de encriptar por bloques los datos de las tablas configuradas en la base de datos.

MySQL es parte de Oracle lo que permite complementarse o trasladar información de una base a otra sin complicaciones. El Centro de Cómputo por políticas de empresa trabaja con base de datos Oracle en la mayoría de sus servicios, ya que se maneja una intranet con los diferentes módulos que se desarrollan dependiendo de los requerimientos que le solicitan los diferentes departamentos de la UCSG al Centro de Cómputo. Se recomienda establecer un módulo en el sistema universitario (SIU) que permita la elaboración de reportes más institucionales en caso de requerirlo.

Figura 16. Intranet – SIU

Referencias bibliográficas

Adiscon. (2013). LogAnalyzer Basics. Recuperado a partir de http://loganalyzer.adiscon.com/doc/basics.html

Administración de Logs. (s/f). Recuperado el 8 de noviembre de 2017, a partir de https://www.clm.com.co/soluciones/administracion-registros.htm

Barranco Ricardo, & IBM. (2012, junio 18). ¿Qué es Big Data? Recuperado el 9 de noviembre de 2017, a partir de http://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/index.html

BERNAL CESAR AUGUSTO. (2013). *METODOLOGIA DE LA INVESTIGACION.* S.l.: NIELSEN BOOKDATA.

Big data | ¿Qué es big data? | Oracle España. (s/f). Recuperado el 9 de noviembre de 2017, a partir de https://www.oracle.com/es/big-data/index.html

Big Data: la información es poder. (2014, noviembre 19). Recuperado el 9 de noviembre de 2017, a partir de http://comunidad.iebschool.com/it4all/2014/11/19/big-data-la-informacion-es-poder/

De Maya, D. (2017). CONTROL DE LOGS EN SISTEMAS LINUX. Recuperado a partir de https://hardsoftsecurity.es/logsLinux.pdf

Diaz, L. (2010, noviembre). La Observación. Recuperado a partir de http://www.psicologia.unam.mx/documentos/pdf/publicaciones/La_observacion_Lidia_Diaz_Sanjuan_Texto_Apoyo_Didactico_Metodo_Clinico_3_Sem.pdf

Dzul, M. (2013, octubre). *Aplicación básica de los métodos científicos "Diseño No-Experimental"*. Recuperado a partir de https://www.uaeh.edu.mx/docencia/VI_Presentaciones/licenciatura_en_mercadotecnia/fundamentos_de_metodologia_investigacion/PRES38.pdf

Ferrer Ignasi, & Ortega Dario. (2013, marzo 8). *Registro, Centralización y Análisis de Eventos en un entorno Corporativo Multiplataforma*. Universitat Oberta de Catalunya. Recuperado a partir de http://openaccess.uoc.edu/webapps/o2/bitstream/10609/22750/6/iortegavTFC0313memoria.pdf

Garcia, M., Martinez, C., Naira, M., & Sanchez, L. (2009, junio 4). La entrevista. Recuperado a partir de https://uam.es/personal_pdi/stmaria/jmurillo/Met_Inves_Avan/Presentaciones/Entrevista_(trabajo).pdf

Glass, R., Vessey, I., & Ramesh, V. (2002, marzo 20). Research in software engineering an analysis of the literature. Recuperado a partir de www.elsevier.com/locate/infsof

Gómez, J. (2014, septiembre 25). *Servidor de Logs Centralizado*. Escola Tècnica Superior d'Enginyeria Informàtica Universitat Politècnica de València, Valencia, España. Recuperado a partir de https://riunet.upv.es/bitstream/handle/10251/43428/Memoria.pdf?sequence=1

Loggly. (s/f). *See it. Analyze it. Inspect it. Solve it.* Recuperado a partir de https://www.loggly.com/product/

López José. (2014, febrero 27). La moda del Big Data: ¿En qué consiste en realidad? - elEconomista.es. Recuperado el 9 de noviembre de 2017, a partir de http://www.eleconomista.es/tecnologia/noticias/5578707/02/14/La-moda-del-Big-Data-En-que-consiste-en-realidad.html

María Begoña Alonso, & Alegre Díez. (2016, febrero). *Gestión de Logs*. Universidad Internacional De La Rioja. Recuperado a partir de http://reunir.unir.net/bitstream/handle/123456789/3618/ALONSO-ALEGRE%20DIEZ%2C%20MARIA%20BEGO%C3%91A.pdf?sequence=1

Porven, J., & Montesino, R. (s/f). Marco de trabajo para la gestión centralizada de trazas de seguridad usando herramientas de código abierto. *Julio 2015*, *9*(3). Recuperado a partir de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2227-18992015000300002

PowerData, G. (s/f). Big Data: ¿En qué consiste? Su importancia, desafíos y gobernabilidad. Recuperado el 9 de noviembre de 2017, a partir de https://www.powerdata.es/big-data

Psyllos Andeas. (2012). *Using the 3CX Server Activity Log to Troubleshoot Issues*. Recuperado a partir de https://www.3cx.com/blog/docs/3cx-server-activity-log/

Rodriguez, D. O. (2008, enero 31). SEGURIDAD INFORMATICA: MEJORES PRACTICAS Y HERRAMIENTAS PARA EL MONITOREO DE BITACORAS EN UNIX. Recuperado el 8 de noviembre de 2017, a partir de

http://danielomarrodriguez.blogspot.com/2008/01/mejores-practicas-y-herramientas-para.html

Rodríguez Hernández, C., Gavio, B., Felipe, R., Román Bu, Y., Rivero, D., Manuel, C., & Cortés Cortés, M. (2014). GeReport: Sistema de Gestión de Reportes Dinámicos. *Revista Cubana de Ciencias Informáticas*, *8*(4), 116–129.

Rouse Margaret. (2014, diciembre). What is machine data? Recuperado el 9 de noviembre de 2017, a partir de http://internetofthingsagenda.techtarget.com/definition/machine-data

Rumbos Salomón, R. E. (2012). *El Gran libro de Debian GNU/Linux*. Barcelona: Marcombo.

Scobles, A. (2012, mayo 8). Installing LogAnalyzer and rsyslog on CentOS. Recuperado a partir de http://itmanager.blogs.com/notes/2012/05/setting-up-a-loganalyzersyslog-server.html

Splunk. (2013, noviembre). *Presentación de la herramienta Splunk*.

Splunk. (s/f-a). Splunk hace que los datos de máquina sean accesibles, útiles y valiosos para todo el mundo. Recuperado el 10 de noviembre de 2017, a partir de https://www.splunk.com/es_es/resources/machine-data.html

Splunk. (s/f-b). Splunk hace que los datos de máquina sean accesibles, útiles y valiosos para todo el mundo. Recuperado el 10 de noviembre de 2017, a partir de https://www.splunk.com/es_es/resources/machine-data.html

Stallman, R. M., & Lessig, L. (2004). Software libre para una sociedad libre. Recuperado a partir de https://www.gnu.org/philosophy/fsfs/free_software2.es.pdf

Sunil Soares. (2012, junio 3). Not Your Type? Big Data Matchmaker On Five Data Types You Need To Explore Today. Recuperado el 9 de noviembre de 2017, a partir de http://www.dataversity.net/not-your-type-big-data-matchmaker-on-five-data-types-you-need-to-explore-today/

Vieda, M. (2013, julio 20). Administración de LOGS. Recuperado a partir de https://manuelvieda.com/blog/administracion-de-logs/

ANEXO 1

UNIVERSIDAD CATÓLICA DE SANTIAGO DE GUAYAQUIL

Guayaquil, 23 de octubre de 2017

Ing.
Beatriz Guerrero, M.S.
Directora de carrera
Ingeniería en sistemas
En su despacho

Autorización:

Yo, el Mgs. Vicente Adolfo Gallardo Posligua con numero de cedula: 1202433817 autorizo al estudiante **Luis Fernando Fiallos Romero** con documento de identidad **0950904722** a que realice el tema de titulación "*Servicio de logs centralizado para los servicios del Centro De Computo de la UCSG, visualizado por medio de una aplicación web*" dentro del centro de cómputo usando los equipos disponibles del mismo para lograr un mejor desempeño del centro de cómputo.

Agradezco de antemano su amable atención a la presente.

Atentamente,

Ing. Vicente Gallardo Posligua, Mgs.
Director del Centro de Cómputo

Apartado 09-01-4671
Teléfono 2206951
Guayaquil – Ecuador

ANEXO 2

UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

Guayaquil, 20 de febrero de 2017

Ing.
Beatriz Guerrero, M.S.
Directora de carrera
Ingeniería en sistemas
En su despacho

Aceptación:

Yo, el Mgs. Vicente Adolfo Gallardo Posligua con numero de cedula: 1202433817 constato que el estudiante **Luis Fernando Fiallos Romero** con documento de identidad **0950904722** ha realizado el desarrollo e implementación propuesta en el tema de titulación *"Servicio de logs centralizado para los servicios del Centro De Computo de la UCSG, visualizado por medio de una aplicación web"* cumpliendo los requerimientos y objetivos planteados por el Centro de Cómputo.

Particular que informo para los fines pertinentes.

Atentamente,

Ing. Vicente Gallardo Posligua, Mgs.
Director del Centro de Cómputo

C.C: Archivo

Apartado 09-01-4671
Teléfono 2206951
Guayaquil – Ecuador

Recibido
Feb 20/17
21 hs.

ANEXO 3

UNIVERSIDAD CATÓLICA DE SANTIAGO DE GUAYAQUIL

Guayaquil, 22 de febrero de 2017

Ing.
Beatriz Guerrero, M.S.
Directora de carrera
Ingeniería en sistemas
En su despacho

De mis consideraciones:

Yo, el estudiante **Luis Fernando Fiallos Romero** con documento de identidad **0950904722** que ha realizado el desarrollo e implementación propuesta en el tema de titulación *"Servicio de logs centralizado para los servicios del Centro De Computo de la UCSG, visualizado por medio de una aplicación web"* cumplo en informarle las credenciales necesarias para acceder desde la red interna al servicio que fue implementado en el Centro de Cómputo de la UCSG:

Dirección IP privada: *********
Usuario administrador: logs_cc_2018
Contraseña: *********

Cabe mencionar que, al ser un servicio privado y crítico para el Centro de Cómputo, las credenciales completas fueron entregadas vía correo electrónico al director del departamento.

Particular que informo para los fines pertinentes.

Atentamente,

Luis Fernando Fiallos Romero

Apartado 09-01-4671
Teléfono 2206951
Guayaquil – Ecuador

C.C: Archivo
Ing. Vicente Gallardo

DECLARACIÓN Y AUTORIZACIÓN

Yo, **Fiallos Romero Luis Fernando** con C.C: # 0950904722 autor/a del trabajo de titulación: **Servicio de logs centralizado para los servidores del Centro de Cómputo de la UCSG, visualizado por medio de una aplicación web dentro de la red interna** previo a la obtención del título de **Ingeniero en Sistemas Computacionales** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, **15** de **febrero** de **2018**

f. _____

Nombre: **Fiallos Romero Luis Fernando**

C.C: **0950904722**

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TEMA Y SUBTEMA:	Servicio de logs centralizado para los servidores del Centro de Cómputo de la UCSG, visualizado por medio de una aplicación web dentro de la red interna		
AUTOR(ES)	Luis Fernando Fiallos Romero		
REVISOR(ES)/TUTOR(ES)	Vicente Adolfo Gallardo Posligua		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Ingeniería		
CARRERA:	Ingeniería en Sistemas Computacionales		
TITULO OBTENIDO:	Ingeniero en Sistemas Computacionales		
FECHA DE PUBLICACIÓN:	7 de marzo de 2018	**No. DE PÁGINAS:**	88
ÁREAS TEMÁTICAS:	Tecnologías y sistema productivo, tecnologías de la información y comunicación		
PALABRAS CLAVES/ KEYWORDS:	Log, Syslog, Rsyslog, LogAnalyzer, Centralización de logs, Software libre.		

RESUMEN/ABSTRACT: Este proyecto abarca la problemática de lo necesario que es para una empresa tener implementada una estructura de centralización de logs, para lograr una mejor visualización y administración de eventos generados por los equipos y servicios manejados por la organización. En la actualidad existen múltiples protocolos y herramientas de visualización y administración de logs centralizados, es por esta razón que se procedió a investigar antecedentes y definiciones básicas del tema en general, con el fin de desarrollar un proyecto de administración y visualización de logs enfocándose en las necesidades del Centro de Cómputo de la Universidad Católica de Santiago de Guayaquil, ya que es el departamento encargado de gestionar la plataforma tecnológica que soporta los diferentes servicios académicos y administrativos dentro de la universidad. Como parte de los objetivos de este proyecto se analizaron los requerimientos y características que fueron solicitados por el Centro de Cómputo para la implementación de un equipo centralizado de logs. Para la selección de la herramienta utilizada se compararon las más conocidas con las características requeridas, implementándose un servidor LogAnalyzer sobre un ambiente operativo CentOS que capturó información de diferentes tipos de equipos y servidores haciendo uso del protocolo SYSLOG. Con la finalidad de que este proyecto sea utilizado por el Centro de Cómputo en un ambiente de producción, se ha diseñado un manual de instalación y configuración de LogAnalyzer para la implementación del servidor centralizado en el Centro de Cómputo considerando los diferentes servidores y servicios que posee.

ADJUNTO PDF:	☒ SI	☐ NO
CONTACTO CON AUTOR/ES:	**Teléfono:** +593-4-2133370	E-mail: luis.fiallos@cu.ucsg.edu.ec
CONTACTO CON LA INSTITUCIÓN (C00RDINADOR DEL PROCESO UTE)::	**Nombre:** Yanza Montalvan Angela Olivia	
	Teléfono: +593-9-83035702	
	E-mail: angela.yanza@cu.ucsg.edu.ec	

SECCIÓN PARA USO DE BIBLIOTECA	
Nº. DE REGISTRO (en base a datos):	
Nº. DE CLASIFICACIÓN:	
DIRECCIÓN URL (tesis en la web):	