

Universidad de La Habana  
Facultad de Matemática y Computación



# Implementación de un sistema de autenticación de usuarios basado en LDAP

Autor: Eric Nordelo Galiano  
Lian Ulloa Mc-kion

Tutor: **Lic. Gilberto Garcia**

Tutor: **Lic. Darío Álvarez Arteaga**

Trabajo de Diploma  
presentado en opción al título de  
Licenciado en Ciencias de la Computación



Junio de 2019



# Agradecimientos

## Opinión del tutor

# Resumen

# Abstract

# Índice general

<b>Introducción</b>	<b>1</b>
<b>1. Directorio Único</b>	<b>6</b>
1.1. Surgimiento del Directorio Único . . . . .	6
<b>2. Estado del arte</b>	<b>8</b>
2.1. EA1 . . . . .	8
2.2. EA2 . . . . .	8
2.3. EA3 . . . . .	8
2.4. EA4 . . . . .	8
2.5. EA5 . . . . .	8
<b>3. Implementación del sistema</b>	<b>9</b>
3.1. Consideraciones previas . . . . .	9
3.2. SI2 . . . . .	9
3.3. SI3 . . . . .	9
3.3.1. SI3.1 . . . . .	9
3.3.2. SI3.2 . . . . .	9
3.3.3. SI3.3 . . . . .	9
3.4. Desarrollo de la solución . . . . .	9
<b>4. Experimentación y Resultados</b>	<b>10</b>
4.1. Configuración . . . . .	10
4.2. Experimentación . . . . .	10
4.3. Metodología para medir rendimiento de Squid . . . . .	10
<b>Recomendaciones</b>	<b>11</b>
<b>Conclusiones</b>	<b>12</b>

# Índice de figuras



# Introducción

Desde la propuesta de Leonard Kleinrock en el año 1961 en un artículo titulado "Information Flow in Large Communication Nets" (traducido ha "Flujo de Información en Grandes Redes de Comunicación"), y el uso del término "paquete" en 1965 por Donald Davies para describir datos enviados entre computadoras en una red, que impulsarían el desarrollo de ARPANET entre 1966 y 1969, el planeta entró en una nueva etapa que revolucionó el desarrollo en todos los campos tecnológicos: la "Era Digital" o "Era de la Información". Esta era gira en torno a las nuevas tecnologías e Internet y está llevando a cabo cambios profundos y transformaciones en una sociedad donde la automatización de los procesos mejora considerablemente la calidad, rapidez y robustez de los mismos, y donde la conectividad mediante las redes de dispositivos (no puede hablarse solo de computadoras) facilitan la comunicación y la organización tanto de personas como de empresas y organismos.

Las oraciones en el párrafo anterior son largas y la idea detrás de ellas es difícil de seguir. Todo lo que este en un formato similar al del párrafo anterior debe ser separado en oraciones cortas y conexas.

Una de las ventajas que trae consigo el establecimiento de redes de dispositivos es la posibilidad de descentralizar el almacenamiento de información manteniendo un adecuado acceso a la misma. Actualmente la red de la Universidad de La Habana cuenta con un gran número de servicios consumidores o fuentes de diferentes tipos de información, pero carece de un servicio que "administre" el flujo de comunicación con los mismos. De estos servicios se benefician no solo las facultades pertenecientes a la Universidad, también se benefician otras instituciones asociadas a la misma como el Instituto Superior de Diseño (ISDI) y el Instituto Superior de Ciencias y Tecnologías Aplicadas (InSTEC). La gran mayoría de estos servicios requiere que el la verificación previa del usuario que solicita acceder o agregar información a los mismos, por lo que con el objetivo de centralizar el proceso de verificación

Lleva referencia

Lleva referencia

Fue la invención del término lo que provocó el impulso de ARPANET? Además, esto realmente viene al caso?

No me parece adecuado el uso de la palabra planeta aca. Realmente fue solo la humanidad (parte de ella, ni siquiera me queda claro que porcentaje)

se han ido acumulando sobre un dominio web ([directorio.uh.cu](http://directorio.uh.cu)) varias actualizaciones aisladas que conforman lo que se conoce como "Directorio Único de la Universidad de La Habana" para poder integrar y administrar servicios que funcionan sobre diferentes tecnologías. Esto trae consigo dificultad a la hora de integrar nuevos servicios al directorio, así como inestabilidad y falta de robustez en el servicio de verificación y solicitud de información almacenada.

Reescribir  
por favor.

El párrafo de la explicación del problema debería ser re escrito para su mejor comprensión.

Dicho esto, el objetivo de este proyecto es proveer a la Universidad de La Habana de un servicio que se encargue de administrar estos procesos de manera confiable y eficiente, que además sea fácil de integrar y fácil de modificar (implementar mejoras).

Para lograr esto surgen varias preguntas a responder como por ejemplo: ¿qué método se debería utilizar para almacenar la información?, ¿qué método se debería utilizar para acceder a la misma?, ¿qué método usar para mantener la información actualizada?...

Como respuesta a la primera pregunta, después de analizar distintos métodos para gestionar información, debido a las características del problema que enfrentamos llegamos a la conclusión que lo mejor es almacenarla en una estructura de directorios. La principal ventaja de esto sobre las bases de datos relacionales (y no relacionales) es la velocidad para realizar consultas sobre grandes volúmenes de datos (perdiendo en velocidad de modificación de los mismos, tanto en actualizaciones, como en inserciones y eliminaciones).

Si vas a  
usa esto  
como un  
objetivo  
general,  
'estos  
procesos'  
debe  
ser cla-  
ramente  
definido.  
Cuales  
son los  
procesos?

Este es el  
punto de  
conducir  
la inves-  
tigación,  
es mejor  
desarro-  
llar una  
separa-  
ción ló-  
gica de  
estas pre-  
guntas  
mas que  
presentar-  
las plana-  
mente al  
lector.

Estas son  
las res-  
puestas  
que de-  
beríamos

Decidido esto y después de analizar diversas maneras de manejar una estructura de directorios optamos por el LDAP (Lightweight Directory Access Protocol o Protocolo Ligero de Acceso a Directorios) como protocolo de acceso y OpenLDAP como implementación del mismo.

Este protocolo brinda un esquema similar al de una guía telefónica, implementando un enfoque jerárquico para el almacenamiento de la información. Además existe una interfaz de autenticación para la mayoría de los servicios que administran usuarios y permisos, que utiliza como fuente de información a servidores que implementen este protocolo. Partiendo de la problemática existente la autenticación de los servicios que se brindan en el recinto universitario, se ha formulado la siguiente hipótesis de investigación: mediante el sistema de autenticación basado en protocolo LDAP se solucionará la administración eficiente de la información y los servicios en la Universidad de La Habana.

Con esta tesis queremos sustituir el actual sistema de directorio único por un sistema de autenticación basado en el protocolo LDAP.

Para esto es necesario investigar sobre el estado del arte de los sistemas que implementan este protocolo, así como la disponibilidad que brindan. Además es necesario automatizar el proceso de recopilar la información provista por estas fuentes, para lo cual se diseñará una esquema de base de datos, guiado por el protocolo LDAP, que se adapte a la estructura de los datos almacenados actualmente. Por último, se plantea implementar una API acorde con el protocolo OpenId para facilitar la autenticación para futuros servicios.

















Evitemos usar **Debido a esto y Decido esto**, se puede reformular como **Habiendo analizado las ventajas de usar una estructura de directorio se decidió** ...

En la introducción no toca explicar el protocolo LDAP. Tiene que haber una sección del capítulo 2 para explicar el concepto de LDAP y como funciona.

Esto esta aqui para plantear algun objetivo?

Generar la hipótesis a partir

# Notes

	Lleva referencia . . . . .	1
	Lleva referencia . . . . .	1
	Fue la invención del término lo que provocó el impulso de ARPA- NET? Además, esto realmente viene al caso? . . . . .	1
	No me parece adecuado el uso de la palabra planeta aca. Realmente fue solo la humanidad (parte de ella, ni siquiera me queda claro que porciento) . . . . .	1
	Que cambios? Como calificas cuales de ellos son profundos? . . . .	1
	Este es el resultado natural de la automatización. . . . .	1
	Las oraciones en el párrafo anterior son largas y la idea detrás de ellas es difícil de seguir. Todo lo que este en un formato similar al del párrafo anterior debe ser separado en oraciones cortas y conexas. . . . .	1
	Recomiendo cambiar por: 'sin comprometer el acceso' . . . . .	1
	gran número? . . . . .	1
	Ahora mismo no lo hacen, parte del objetivo de esta Tesis es sentar las bases para que lo hagan. . . . .	1
	Agregar informacion no requiere de esta autenticación . . . . .	1
	Reescribir por favor. . . . .	2
	El párrafo del la explicación del problema debería ser re escrito para su mejor comprensión. . . . .	2
	Si vas a usa esto como un objetivo general, 'estos procesos' debe ser claramente definido. Cuales son los procesos? . . . . .	2
	Este es el punto de conducir la investigación, es mejor desarrollar una separación lógica de estas preguntas mas que presentarlas planamente al lector. . . . .	2
	Estas son las respuestas que deberiamos ir intercalando con las preguntas del parrafo anterior. . . . .	2

■	Esta afirmacion necesita una referencia que la apolle. No tiene que ser a un artículo, puede ser a la sección del documento donde se explica la razón y dicha sección tendria referencias a cualquier otro documento. . . . .	2
■	Evitemos usar <b>Debido a esto</b> y <b>Decido esto</b> , se puede reformular como <b>Habiendo analizado las ventajas de usar una estructura de directorio se decidió ...</b> . . . . .	3
■	En la introducción no toca explicar el protocolo LDAP. Tiene que haber una sección del capitulo 2 para explicar el concepto de LDAP y como funciona. . . . .	3
■	Esto esta aqui para plantear algun objetivo? . . . . .	3
■	Generar la hiposteis a partir del objetivo que puede ser: <b>la implementación de un sistema LDAP para la administración de la información dentro de la UH</b> y hacer notar que de las deficiencias citadas (que debieron haber sido mencionadas anteriormente) saldrán las mejoras inmediatas . . . . .	3
■	Esta último parrafo esta bien. . . . .	3
■	En qué fecha se hizo esto? . . . . .	6
■	Cuáles son los nombres oficiales de las fuentes de datos?, Qué significa SIGENU? si es que son siglas . . . . .	6

# Capítulo 1

## Directorio Único

### 1.1. Surgimiento del Directorio Único

Como resultado de la necesidad de acceso a la información de los usuarios de la red de la Universidad de la Habana, tanto estudiantes, como profesores y externos, varios servicios han sido implementados de manera aislada, durante los últimos años con el fin de suplir esta necesidad. A raíz del proceso de centralización de los servicios llevado a cabo en el período de <fechainicio-fechafinal>, dichos servicios, relacionados con la creación, almacenamiento y la actualización de la información referente a los usuarios, fueron integrados en lo que se conoce hoy como Directorio Único.

En qué fecha se hizo esto?

Entre las principales funcionalidades que componen el Directorio Único, se encuentra la de constituir una fuente, centralizada, de información sobre toda aquella persona que pretenda hacer uso de las facilidades disponibles en la red universitaria. Dicha centralización representa un elemento necesario para garantizar el control apropiado del uso de dichas facilidades, ya que la fragmentación de los datos, inherente a la estructura actual de la Universidad y a sus actividades, dígame la composición por facultades, departamentos de investigación, etc, así como la continua realización de eventos para fomentar la investigación y el intercambio de experiencias, aumenta considerablemente la complejidad a la hora administrar la información almacenada y la restricción o concesión de acceso a los servicios brindados.

Sin embargo, esta información no se genera de forma centralizada, sino que se encuentra esparcida entre distintas fuentes como son el SIGENU, .... Por esta razón, esta información centralizada, es actualizada a través de diversos procesos con una frecuencia no tan regular como debiera ser, lo cuál se debe principalmente a que parte de estos procesos mencionados

Cuáles son los nombres oficiales de las fuentes de datos?, Qué significa SIGENU? si es que son siglas

anteriormente, ni siquiera son realizados automáticamente, sino que existe una persona(o varias) encargada transportar, generalmente en medios físicos, la información desde donde se genera hasta donde es necesitada (en casa del herrero, cuchillo de palo).

Entre los datos más relevantes, almacenados en este sistema, se encuentra aquellos que permiten la restricción y/o concesión de acceso a los servicios brindados, es decir, las credenciales de los usuarios de la red. Esta información permite:

1. **Autenticar al usuario:** Comprobar que la persona que solicita un servicio es quien dice ser.
2. **Administrar el acceso:** En correspondencia del rol de la persona autenticada, permitir o no el acceso a ciertos servicios

## Capítulo 2

### Estado del arte

2.1. EA1

2.2. EA2

2.3. EA3

2.4. EA4

2.5. EA5



## Capítulo 3

# Implementación del sistema

3.1. Consideraciones previas

3.2. SI2

3.3. SI3

3.3.1. SI3.1

3.3.2. SI3.2

3.3.3. SI3.3

3.4. Desarrollo de la solución

## Capítulo 4

# Experimentación y Resultados

4.1. Configuración

4.2. Experimentación

4.3. Metodología para medir rendimiento de Squid

# Recomendaciones

# Conclusiones

